Process based modelling of ecosystem structure and dynamics with aDGVM2
–
A case study of whole-plant trait trade-offs & shrubs in African savannas

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 11
der Johann Wolfgang Goethe -Universität
in Frankfurt am Main

von
Camille Gaillard
aus Monclar d'Agenais

Frankfurt (2020)
(D 30)

vom Fachbereich 11 Geowissenschaften/Geographie der

Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan:

Prof. Dr. Georg Rümpker

Gutachter:

Dr. Simon Scheiter, Prof. Dr. Thomas Hickler

Datum der Disputation: 16 Dezember 2020

TABLE OF CONTENTS

ACRONYMS & GLOSSARY

aDGVM2: “adaptive Dynamic Global Vegetation Model 2” (DGVM used for this thesis)

AGBM (or AGB): “Above Ground Biomass”

CRU: “Climatic Research Unit”

DBH: “Diameter at Breast Height”

DGVM: “Dynamic Global Vegetation Model”. A type of ecological model.

DVM: “Dynamic Vegetation Model” (i.e., a non-“Global” DGVM)

E.S.: “Ecosystem Service”

ESM: “Earth System Model”

IPCC: “Intergovernmental Panel on Climate Change”

IPBES: “Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services”

LAI: “Leaf Area Index”

LIDAR: “Light Radar” or “Laser Imaging, Detection & Ranging”

LLL: “Limpopo Living Landscape”

LSM: “Land Surface Model”

MAP: “Mean Annual Precipitation”

NDVI: “Normalized Difference Vegetation Index”

NPP: “Net Primary Productivity”

P50: “Matric potential at 50% loss of conductance”

PFT: “Plant Functional Type”

SLA: “Specific Leaf Area”

SUMMARY

Shrubs are a characteristic component of savannas, where they coexist with trees and grasses. They are often part of woody encroachment phenomena, which have been observed globally, and the determinant of shrub encroachment cases, which are particularly of concern in African savannas. In response to climate change and land use change, African savannas are vulnerable to biome shifts and shrub encroachment is a process driving and explaining this risk.

We contribute to furthering the understanding of shrubs biogeography and ecology by considering the number of stems of woody plants to characterise shrubs phenotype and strategy. We postulate that shrubs are multi-stemmed, compared to single-stemmed trees and integrate this assumption in aDGVM2 (adaptive Dynamic Global Vegetation Model 2). Modelling a trait representing the number of stems of a woody plant implies a trade-off between single-stemmed plants having higher height growth potential and multi-stemmed plants having higher hydraulic capacity but limited height growth. Multi-stemmed individuals, being shorter, are more likely to suffer severe damage from fires than tall single-stemmed trees managing to grow their crown out of the flame zone.

We simulate potential vegetation over sub-Saharan Africa at 1° spatial resolution, with aDGVM2 and compare it to simulations without our shrub model turned on. We also test the impact of fire by including or excluding it from our simulations. To assess the accuracy and relevance of our approach, we benchmark our overall model's performance against multiple satellite derived products of above ground biomass (AGBM), and against specific field measurements of AGBM. We further benchmark our results against vegetation cover type derived from satellite data.

We demonstrate that shrubs can be modelled as multi-stemmed woody plants in African savannas based on whole-plant trait trade-off without being predefined as static functional types. Indeed, the addition of our shrub model to aDGVM2 allows for shrubs to emerge dynamically through community assembly processes without *a priori* categorisation. Our shrub model also improves the simulated vegetation patterns simulated by aDGVM2 in sub-Saharan Africa, particularly in savannas. The simulated pattern of stem number per woody individual broadly follows our assumptions about biogeographic patterns as it is lowest in equatorial African forests and increases in savannas and grasslands

as precipitation decreases. Shrubs are more abundant in more water-stressed regions where they have a competitive advantage over trees due to their increased relative water transport potential. However, in arid and hyper-arid regions, further investigations are required. Simulated shrub prevalence is higher in more open and fire prone landscapes, where woody cover and biomass are reduced.

Adding shrubs to aDGVM2, while increasing complexity allows for greater simulated diversity. As resilience and resistance of ecosystems have been shown to be influenced by diversity, such model development is necessary to improve our ability to forecast ecosystems responses to changes. However, there are challenges to fully tap this benefit. Assessing the accuracy and relevance of our approach is challenging. Data and simulations are conceptually different which limit the possibility to conclude based on comparison. Benchmarking challenge is exacerbated by the variability existing among satellite derived products and site studies observations. In areas of extremely low biomass and vegetation cover, such as deserts and semi-deserts, the accuracy of our model is more concerning as small differences in absolute values are relatively more important.

Categorisation of life-forms shapes our understanding of their ecology and biogeography, thus, consensus about their definition is direly needed. To contribute to this debate, we investigate how vegetation distribution patterns arising from our shrub model inform our understanding of shrub biogeography. First, shrub distribution in trait space (considering stem number), relatively to environmental drivers, concurs with our assumptions. Second, shrub spatial distribution is consistent with our characterisation assumptions. Third, the role of simulated shrubs in an ecosystem supports realistic ecological dynamics. Our model allows for, shrubs to exhibit a specific phenotype, but also a specific life-strategy, which we characterise in terms of persistence strategy (shrubs are mainly resprouters, in contrast to trees, which can be either resprouters or reseeders) and in terms of resource acquisition (rooting strategy) and allocation (carbon investment). Adding stem count as a trait to aDGVM2 increase the range of simulated functional diversity.

Our shrub model allows for aDGVM2 to simulate realistic ratio of grass to woody vegetation across sub-saharan Africa. Similarly, it simulates ratio of shrubs to trees consistent with our hypotheses. Shrubs mediate the competitive balance between grasses and trees and have a complex interplay with fire. Agreement with benchmarking data is shaped by model stochasticity and by the perspective taken on simulations results, as shrub prevalence depends on the variable considered (plant number, canopy cover, AGBM share), and on the thresholds used to classify shrubs. Thus, our non-categorical approach, in contrast to PFT (Plant Functional Type) based approaches, is more challenging to benchmark; yet, it offers greater opportunities to study ecological dynamics. While investigating shrubs as a category provides meaningful information, questioning the traits and parameters underlying it provides further information about ecosystems states and dynamics. We argue that considering the continuum of life-forms can help circumvent the challenge of unclear categories.

Increasing realism typically implies to increase model complexity, which challenges model development and model operation. Initialisation can be solved by increasing data availability. However, at present, model complexity exceeds data availability for complete model initialisation. We argue that, the more variables are initialised based on real data and not randomly, the better the performance of a model. Therefore, when stochastically initialising a model and conducting replicate simulations, the iteration producing the best fit with benchmarking data should be the one having randomly selected the most realistic value for initialisation, and thus could be used to reverse engineer parameters values.

Ecosystems are dynamic systems, yet models are typically benchmarked to constant values or fixed states not fitting with ecological timescales. This gap can be bridged through reconstruction of past ecological states and dynamics. The lack of long-term data for benchmarking share similarities with the lack of holistic model benchmarking. In both cases it is possible to misinterpret a model relevance and accuracy by focusing on a limited subset of parameters. A cautious approach should consider that a model can be deemed satisfactory only regarding parameters for which it has been benchmarked against observation data. However, optimizing the fit for one parameter might decrease the fit for another; thus, highlighting points which require research effort. We argue that more holistic benchmarkings are more promising as they expand the relevance of models and can foster model development.

Ecosystem history, ecological timescale and overall context dependencies, question determinism in ecology. Determining to which extent rules can be derived from observations and to which extent observed ecological states and parameters are due to chaotic processes is a critical question for model development. We argue that the finer the details of a model (e.g.: modelling photosynthetic rate at the cell level instead of NPP (Net Primary Productivity) at landscape scale) the more likely modelled processes are to be deterministic. However, it comes at the expense of model complexity. We propose to investigate evolutionary processes, and specifically co-evolution. This would allow to determine the level of complexity necessary to ensure sufficient model performance and to assess the importance of determinism in ecosystems history. Additionally, we argue that iteratively sampling the relative weight of stochasticity in a model is necessary to assess its simulations confidence. Concomitantly, averaging replicate simulations does not necessarily provide more accurate evaluation. In contrast, we argue that stochasticity can be used to represent ecosystem heterogeneity at the sub-grid cell scale. Such approach implies to question the consequences of scaling up models' simulations from the basic unit at which they are performed to the grid-cell scale they represent. We demonstrate that the basic unit at which a model operates (for aDGVM2: the number of simulated individuals) modifies the relative importance of stochasticity.

Model assumptions are implicitly considered valid when simulations results can be satisfactorily benchmarked, however, we argue that they should be explicitly discussed as well, and not just their outcome. Similarly, we argue that data constraints should be more extensively integrated into the evaluation of models, be it for initialisation or benchmarking, in order to avoid overlooking unaccounted for limitations. We also argue that benchmarking methodologies should be adapted to new model developments to remain relevant. More holistic benchmarking should account for all dimensions of a model (space, time, traits, variables) and for the relative strength of a process depending on the mixture of other processes included in a model with which it can interact. Benchmarking could also go beyond assessment of models' accuracy and be used to fit model post-process and inform initialisation.

Shrub encroachment has been increasingly reported to occur over vast expanses of multiple ecosystems globally. Despite the increasing number of studies about this topic over the last ten years, no consensus emerged so far to explain this phenomenon. We argue that, while general trends can be highlighted, shrub encroachment is a polymorphic phenomenon and disentangling the relative importance of proposed drivers of shrub encroachment can only be achieved locally. Even if explanations of the phenomenon relevant for its management can only be achieved locally, the development of models able to represent them has to move away from context dependent studies and has to take into account multiple drivers and their interactions. We investigate shrub encroachment understood as the increase in shrub prevalence. Precisely, we consider mean stem count as a proxy for shrub prevalence. An array of environmental and anthropogenic factors and ecological mechanisms have been proposed to explain shrub encroachment. They range from shrub suppression mechanisms to alternative stable state theory to the direct effect of biotic and abiotic drivers. The aDGVM2 can represent a subset of those drivers. Namely we assess the relative importance and interactions of fire, grazing, MAP and atmospheric [CO_2] for shrub prevalence. This provides insights to forecast future shrub encroachment scenario and thus to inform potential management strategies.

We focus on 35 sites across southern African savannas, where cases of woody cover expansion have been documented. We test 11 scenarios considering 3 levels for fire, grazing and atmospheric [CO_2], while MAP is implicitly considered across the diversity of sites for which we run simulations. Additionally, we account for stochasticity by conducting 10 replicate simulations for each site and each scenario, for a total of 3850 unique simulation runs. The aDGVM2 simulates higher shrub prevalence in the presence of natural fire and in the absence of grazing under pre-industrial atmospheric [CO_2]. Shrub prevalence tends to increase with decreasing MAP depending on scenario and site considered, albeit non-linearly, with a distinct threshold around MAP of 700mm/year for multiple scenarios. In parallel, literature shows that MAP impact on shrub encroachment can range from positive to negative. Shrub prevalence decreases with increasing atmospheric [CO_2]. As aDGVM2 has been demonstrated to simulate increased vegetation height with increasing atmospheric [CO_2], this result indicates that CO_2 fertilisation effect leads to woody cover expansion where trees dominate at the expense of shrubs. Fire can suppress shrubs both at high frequency and when absent. The interaction between fire and grazing is complex. More precise investigations, considering higher number of pressure level, are required to set the upper and lower limits of fire frequency driving shrubs prevalence and to design precise management plans, according to local specificities. This is particularly critical when threshold effect might have a major impact, as simulated for MAP, depending on scenario and site.

Our simulation experiment highlights that shrub encroachment is multi-factorial. While general trends can be identified, further investigations are required to consider potential explanatory drivers which were excluded from our analysis, such as elephants or wood fuel harvesting, and to consider additional factors level, as well as transitive scenarios.

However, complexity and computation time would increase beyond our available computation resource. Additionally, while mean stem count is a meaningful proxy for shrub prevalence, considering other variables, such as canopy cover or biomass, can alter our perception of shrub prevalence. Finally, focusing solely on shrubs does not inform about overall vegetation community response to tested drivers. This could be achieved by considering trait space response to stress factors. In conclusion, management plans should avoid relying on general trends of shrub encroachment and instead draw upon precise and site-specific investigations to disentangle the relative weight of each potential drivers according to their context dependencies.

ZUSAMMENFASSUNG

Sträucher sind ein charakteristischer Bestandteil von Savannen, in denen sie mit Bäumen und Gräsern koexistieren. Sie sind häufig Teil des weltweit zu beobachtenden Vordringens von Gehölzen und von Verbuschung, die in afrikanischen Savannen besonders besorgniserregend ist. Aufgrund des Klimawandels und der Landnutzungsänderungen sind afrikanische Savannen anfällig für Biomverschiebungen. Die Verbuschung ist ein Prozess, der dieses Risiko antreibt und erklärt.

Die vorliegende Arbeit trägt zum weiteren Verständnis der Biogeographie und Ökologie von Sträuchern bei, indem die Anzahl der Stämme von Holzpflanzen berücksichtigt wird, um den Phänotyp und die Strategie von Sträuchern zu charakterisieren. Im Vergleich zu Bäumen mit nur einem Stamm sind Sträucher immer mehrstämmig. Diese Annahme wird in das aDGVM2 (adaptive Dynamic Global Vegetation Model 2) integriert. Die Modellierung eines Merkmals, das die Anzahl der Stämme einer Holzpflanze darstellt, impliziert einen Trade-off zwischen einstämmigen Pflanzen mit größerem Höhenwachstumspotential und mehrstämmigen Pflanzen mit größerer hydraulischer Kapazität, aber begrenztem Höhenwachstum. Mehrstämmige Individuen, die generell kleiner sind, erleiden mit größerer Wahrscheinlichkeit schwerere Schäden durch Brände als hohe einstämmige Bäume, die es schaffen, ihre Krone aus der Flammenzone herauszuwachsen.

Mit dem aDGVM2 wurde eine potentielle Vegetation in Subsahara-Afrika mit einer räumlichen Auflösung von 1° simuliert und diese mit Simulationen verglichen, bei denen das Strauchmodell nicht berücksichtigt war. Getestet wurden auch die Auswirkungen von Feuer, indem dieses in den Simulationen entweder einbezogen oder ausgeschlossen wurde. Um die Genauigkeit und Aussagekräftigkeit des Ansatzes zu beurteilen, wurde die Leistung des Gesamtmodells mit mehreren satellitengestützten Daten der oberirdischen Biomasse (AGBM) und mit standortspezifischen Feldmessungen der AGBM verglichen. Des Weiteren wurden die Ergebnisse der Vegetationsbedeckung, die aus Satellitendaten abgeleitet wurde, gegenübergestellt.

Die vorliegende Arbeit zeigt, dass es möglich ist, auf Basis des Trade-offs zwischen Merkmalen ganzer Pflanzen Sträucher als mehrstämmige Gehölze in afrikanischen Savannen zu modellieren, ohne dass die Sträucher als statische Funktionstypen vordefiniert

werden müssen. In der Tat können durch das Hinzufügen des Strauchmodells zum aDGVM2 Sträucher ohne vorherige Kategorisierung dynamisch durch Prozesse zur Bildung von Artengemeinschaften entstehen. Dieses Strauchmodell verbessert auch die Vegetationsmuster, die vom aDGVM2 in Subsahara-Afrika simuliert wurden, insbesondere in Savannen. Das simulierte Muster der Stammanzahl pro holzigem Individuum folgt weitgehend den Annahmen über biogeografische Muster, da es in äquatorialafrikanischen Wäldern am niedrigsten ist und in Savannen und Graslandschaften mit abnehmendem Niederschlag zunimmt. Sträucher sind in Regionen mit Wasserknappheit häufiger anzutreffen, wo sie aufgrund ihres erhöhten relativen Wassertransportpotenzials einen Wettbewerbsvorteil gegenüber Bäumen haben. Für aride und hyperaride Regionen sind jedoch weitere Untersuchungen erforderlich. Die simulierte Prävalenz von Sträuchern ist in offeneren und feueranfälligeren Landschaften höher, wo die Holzbedeckung und Biomasse reduziert sind.

Das Hinzufügen von Sträuchern zum aDGVM2 und die zunehmende Komplexität des Modells ermöglichen eine größere simulierte Diversität. Da die Resilienz und Resistenz von Ökosystemen durch Diversität beeinflusst werden, ist eine solche Modellentwicklung notwendig, um besser vorherzusagen, wie Ökosysteme auf Veränderungen reagieren. Diesen Vorteil voll auszuschöpfen, stellt jedoch eine Herausforderung dar. Es ist schwierig, die Genauigkeit und Aussagekräftigkeit des Modellansatzes zu beurteilen. Daten und Simulationen sind konzeptionell unterschiedlich, was die Möglichkeit einschränkt, auf der Basis eines Vergleichs zwischen Daten und Simulationen zu schlussfolgern. Die Herausforderungen des Benchmarking werden zusätzlich durch die Schwankungen zwischen satellitengestützten Daten und Beobachtungen aus Feldstudien verschärft. In Gebieten mit extrem geringer Biomasse- und Vegetationsbedeckung, wie Wüsten und Halbwüsten, wird die Genauigkeit des Modells zunehmend bedeutsamer, da bereits kleine Unterschiede in den absoluten Werten verhältnismäßig wichtig sind.

Da die Kategorisierung von Lebensformen unser Verständnis ihrer Ökologie und Biogeographie beeinflusst, ist ein Konsens über ihre Definition dringend erforderlich. Um zu dieser Debatte beizutragen, wurde untersucht, wie Vegetationsverteilungsmuster, die sich aus dem Strauchmodell ergeben, das Verständnis der Biogeographie der Sträucher beeinflussen. Erstens entspricht die Strauchverteilung im Merkmalsraum unter Berücksichtigung der Stammanzahl und in Abhängigkeit von Umweltfaktoren den oben genannten Annahmen. Zweitens stimmt die räumliche Verteilung von Sträuchern mit den

Charakterisierungsannahmen überein. Drittens unterstützt die Rolle, welche die simulierten Sträucher in einem Ökosystem spielen, eine realistische ökologische Dynamik. Unser Modell erlaubt es, dass Sträucher einen bestimmten Phänotyp, aber auch eine bestimmte Überlebensstrategie aufweisen, die wir in Bezug auf die Persistenz (Sträucher treiben vor allem nach Waldbränden wieder aus (Resprouter), im Gegensatz zu Bäumen, die entweder Resprouters sind oder sich durch Keimung von Samen neu etablieren (Reseeder)) sowie in Bezug auf Ressourcenbeschaffung (Wurzelstrategie) und Allokation (Kohlenstoffinvestition) charakterisieren. Das Hinzufügen der Stammzahl als Merkmal im aDGVM2 erhöht die Bandbreite der simulierten funktionellen Vielfalt.

Das Strauchmodell ermöglicht es dem aDGVM2, ein realistisches Verhältnis von Gras zu Gehölzvegetation in Subsahara-Afrika zu simulieren. In ähnlicher Weise simuliert es das Verhältnis von Sträuchern zu Bäumen gemäß der aufgestellten Hypothesen. Sträucher beeinflussen das Wettbewerbsgleichgewicht zwischen Gräsern und Bäumen und weisen ein komplexes Zusammenspiel mit Feuer auf. Die Übereinstimmung mit Benchmarking-Daten variiert je nach berücksichtigten Daten. Die Ergebnisse des Benchmarking sind zudem durch die Zufälligkeit des Modells und durch die Perspektive, die in Bezug auf die Simulationsergebnisse eingenommen wird, geprägt, da die Strauchprävalenz von den berücksichtigten Variablen (Pflanzenanzahl, Baumdeckung, AGBM-Anteil) und den zur Klassifizierung der Sträucher verwendeten Schwellenwerten abhängt. Daher ist der hier verfolgte nicht-kategorische Ansatz im Gegensatz zu PFT (Plant Functional Type)-basierten Ansätzen zwar beim Benchmarking anspruchsvoller, bietet allerdings größere Möglichkeiten zur Untersuchung der ökologischen Dynamik. Während die Untersuchung von Sträuchern als Kategorie aussagekräftige Informationen liefert, bietet die Infragestellung der zugrunde liegenden Merkmale und Parameter weitere Informationen über den Zustand und die Dynamik von Ökosystemen. In der vorliegenden Arbeit wird die Auffassung vertreten, dass die Berücksichtigung des Kontinuums der Lebensformen dazu beitragen kann, die Herausforderung unklarer Kategorien zu umgehen.

Zunehmender Realismus bedeutet typischerweise eine Erhöhung der Modellkomplexität, was sowohl für die Modellentwicklung als auch für den Modellbetrieb eine Herausforderung ist. Die Initialisierung kann durch das Erhöhen der Datenverfügbarkeit gelöst werden. Derzeit übersteigt die Modellkomplexität jedoch die Datenverfügbarkeit für eine vollständige Modellinitialisierung. Die These ist, dass die Leistung eines Modells umso besser ist, je mehr Variablen auf der Basis von realen Daten und nicht zufällig initialisiert werden. Daher sollte bei der stochastischen Initialisierung eines Modells und der wiederholten Durchführung von Simulationen die Iteration, die am besten zu den Benchmarking-Daten passt, diejenige sein, die zufällig den realistischsten Wert für die Initialisierung ausgewählt hat und somit für das Reverse-Engineering der Parameterwerte verwendet werden kann.

Ökosysteme sind dynamische Systeme, aber Modelle werden normalerweise mit konstanten Werten oder festen Zuständen verglichen, die nicht mit ökologischen Zeitskalen übereinstimmen. Diese Lücke kann durch die Rekonstruktion vergangener ökologischer Zustände und Dynamiken geschlossen werden. Das Fehlen von Langzeitdaten für das Benchmarking weist Ähnlichkeiten mit dem Fehlen eines ganzheitlichen Modell-Benchmarking auf. In beiden Fällen ist es möglich, die Aussagekraft und Genauigkeit eines Modells falsch zu interpretieren, indem man sich auf eine begrenzte Teilmenge von Parametern konzentriert. Bei einem vorsichtigen Ansatz sollte berücksichtigt werden, dass ein Modell nur in Bezug auf Parameter als zufriedenstellend angesehen werden kann, für die es mit Beobachtungsdaten verglichen wurde. Durch Optimierung der Anpassung für einen Parameter kann jedoch die Anpassung für einen anderen Parameter verringert werden, wodurch indes jene Punkte hervorgehoben werden, die erforscht werden müssen. Es wird hier argumentiert, dass ganzheitlicheres Benchmarking vielversprechender ist, da es die Aussagekraft von Modellen erweitert und die Modellentwicklung fördern kann.

Die Geschichte von Ökosystemen, die ökologische Zeitskala und die allgemeinen Kontextabhängigkeiten stellen den Determinismus in der Ökologie in Frage. Inwieweit Regeln aus Beobachtungen abgeleitet werden können und inwieweit beobachtete ökologische Zustände und Parameter auf chaotische Prozesse zurückzuführen sind, sind kritische Fragen für die Modellentwicklung. Es wird hier argumentiert, dass je feiner die Details eines Modells sind (z. B. die Modellierung der Photosyntheserate auf Zellebene anstelle von NPP (Net Primary Productivity) im Landschaftsmaßstab), desto wahrscheinlicher ist es, dass modellierte Prozesse deterministisch sind. Dies geht jedoch zu Lasten der Modellkomplexität. Daher wird vorgeschlagen, evolutionäre Prozesse und insbesondere die Koevolution zu untersuchen. Dies würde es ermöglichen, den Komplexitätsgrad zu bestimmen, der erforderlich ist, um eine ausreichende Modellleistung sicherzustellen, und die Bedeutung des Determinismus in der Geschichte der Ökosysteme zu bewerten. Darüber hinaus wird die Auffassung vertreten, dass eine iterative Stichprobenprüfung der relativen Bedeutung zufälliger Prozesse in einem Modell erforderlich ist, um die Verlässlichkeit der Simulationen zu bewerten. Gleichzeitig liefert der Mittelwert von wiederholten Simulationen nicht unbedingt eine genauere Bewertung. Im Gegensatz dazu wird hier argumentiert, dass zufällige Prozesse verwendet werden können, um die Heterogenität von Ökosystemen in einem Maßstab feiner als Gitterzellen darzustellen. Ein solcher Ansatz impliziert, dass die Konsequenzen der Hochskalierung von Modellsimulationen aus der Grundeinheit, in der sie durchgeführt werden, auf die von ihnen dargestellte Gitterzellenskala in Frage gestellt wird. Es konnte gezeigt werden, dass die Grundeinheit, mit der ein Modell arbeitet (für aDGVM2: die Anzahl der simulierten Individuen), die relative Bedeutung der Zufallsprozesse verändert.

Modellannahmen werden implizit als gültig angesehen, wenn Simulationsergebnisse zufriedenstellend bewertet werden können. Es wird hier jedoch argumentiert, dass die Annahmen ebenfalls explizit diskutiert werden sollten und nicht nur ihre Ergebnisse. In ähnlicher Weise sollten Datenbeschränkungen stärker in die Bewertung von Modellen einbezogen werden, sei es zur Initialisierung oder zum Benchmarking, um zu vermeiden, dass nicht berücksichtigte Einschränkungen übersehen werden. Es wird hier die Auffassung vertreten, dass Benchmarking-Methoden an neue Modellentwicklungen angepasst werden sollten, um relevant zu bleiben. Ein ganzheitlicheres Benchmarking sollte alle Dimensionen eines Modells (Raum, Zeit, Merkmale, Variablen) und die relative Stärke eines Prozesses berücksichtigen, die abhängig ist von den in im Modell integrierten, anderen Prozessen, mit denen der untersuchte Prozess interagieren kann. Das Benchmarking könnte auch über die Bewertung der Modellgenauigkeit hinausgehen und zur Anpassung des Modellnachbearbeitungsprozesses und zur Information für die Initialisierung verwendet werden.

Es wird zunehmend berichtet, dass die Verbuschung global und in weiten Gebieten mehrerer Ökosysteme auftritt. Trotz der Fülle verfügbarer Forschungen zu diesem Thema in den letzten zehn Jahren gab es bisher keinen Konsens, wie dieses Phänomen erklärt werden kann. Es wird hier argumentiert, dass, obwohl allgemeine Trends hervorgehoben werden können, die Verbuschung ein polymorphes Phänomen ist und die Entflechtung der relativen Bedeutung der vorgeschlagenen Treiber für Verbuschungen nur lokal erreicht werden kann. Auch wenn für das Management des Phänomens relevante Erklärungen nur lokal erreicht werden können, muss sich die Entwicklung von Modellen, die in der Lage sind, diese darzustellen, von kontextabhängigen Studien entfernen und mehrere Treiber und ihre Interaktionen berücksichtigen. Untersucht wurde die Verbuschung, die als Zunahme der Strauchprävalenz verstanden wird. Genau genommen wurde die mittlere Stammanzahl als Proxy für die Strauchprävalenz betrachtet. Bislang ist eine Reihe von umweltbedingten und anthropogenen Faktoren sowie ökologischen Mechanismen vorgeschlagen worden, um die Verbuschung zu erklären. Sie reichen von Mechanismen, die das Strauchwachstum unterdrücken, über Theorien alternativer Systemzustände bis hin zur direkten Wirkung biotischer und abiotischer Treiber. Das aDGVM2 kann eine Teilmenge dieser Treiber darstellen. Bewertet werden die relative Bedeutung und die Wechselwirkungen von Feuer, Beweidung, MAP und atmosphärischen [CO_2] für die Strauchprävalenz. Dies bietet Einblicke in die Vorhersage zukünftiger Szenarien der Verbuschung und damit Information über mögliche Managementstrategien.

Die vorliegende Arbeit konzentriert sich auf 35 Standorte in südlichen afrikanischen Savannen, an denen Fälle von Verbuschung dokumentiert wurden. Getestet wurden 11 Szenarien unter Berücksichtigung von 3 Ebenen für Feuer, Beweidung und atmosphärischen [CO_2], während MAP implizit für die Vielfalt der Standorte, für die Simulationen durchgeführt wurden, berücksichtigt wurde. Zusätzlich berücksichtigt wurden die Zufallsprozesse, indem 10 wiederholte Simulationen für jeden Standort und jedes Szenario durchgeführt wurden, was insgesamt 3850 Simulationsläufen entspricht. Das aDGVM2 simuliert eine höhere Strauchprävalenz bei natürlichem Feuer und ohne Beweidung unter vorindustriellem atmosphärischen [CO_2]. Die Strauchprävalenz steigt tendenziell mit abnehmendem MAP in Abhängigkeit von Szenario und Standort, wenn auch nicht linear, mit einem eindeutigen Schwellenwert um den MAP von 700 mm/Jahr für mehrere Szenarien. Parallel dazu zeigt die Literatur, dass der MAP-Einfluss auf die Verbuschung von positiv bis negativ reichen kann. Die Strauchprävalenz nimmt mit zunehmenden atmosphärischem [CO_2] ab. Da das aDGVM2 mit zunehmenden atmosphärischem [CO_2] eine erhöhte Vegetationshöhe simuliert, deutet dieses Ergebnis darauf hin, dass der CO_2-Düngungseffekt zu einer Ausdehnung der Gehölzbedeckung führt, bei der Bäume auf Kosten von Sträuchern dominieren. Feuer kann Sträucher sowohl bei hoher Frequenz als auch bei Abwesenheit unterdrücken. Die Wechselwirkung zwischen Feuer und Beweidung ist komplex. Unter Berücksichtigung einer höheren Anzahl von Druckniveaus sind genauere Untersuchungen erforderlich, um die oberen und unteren Grenzen der Häufigkeit von Feuer, das die Strauchprävalenz determiniert, festzulegen und präzise Bewirtschaftungspläne gemäß den lokalen Besonderheiten zu erstellen. Dies ist besonders kritisch, wenn der Schwellenwerteffekt je nach Szenario und Standort erhebliche Auswirkungen haben kann, wie es für MAP simuliert wurde.

Das Simulationsexperiment zeigt, dass die Verbuschung multifaktoriell ist. Während allgemeine Trends identifiziert werden können, sind weitere Untersuchungen erforderlich, um potenziell erklärende Treiber zu berücksichtigen, die von der durchgeführten Analyse ausgeschlossen wurden, wie z. B. Elefanten oder die Entnahme von Feuerholz, sowie um zusätzliche Faktoren und transitive Szenarien zu berücksichtigen. Komplexität und Rechenzeit würden jedoch über unsere verfügbaren Rechenressourcen hinausgehen. Während die mittlere Stammanzahl ein aussagekräftiger Indikator für die Strauchprävalenz ist, können andere Variablen wie die Baumbedeckung oder die Biomasse unsere Wahrnehmung der Strauchprävalenz verändern. Schließlich sagt die Konzentration auf Sträucher nichts über die allgemeine Reaktion der Vegetationsgemeinschaft auf getestete Treiber aus. Eine solche Aussage könnte erreicht werden, indem die Reaktion des Merkmalsraums auf Stressfaktoren berücksichtigt wird. Zusammenfassend sollten Managementpläne vermeiden, sich auf allgemeine Trends der Verbuschung zu stützen, und stattdessen auf präzise und standortspezifische Untersuchungen zurückgreifen, um die relative Bedeutung jedes potenziellen Treibers entsprechend seiner Kontextabhängigkeit zu entwirren.

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

<u>Chapter 4</u>

Chapter 5

Chapter 6

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

CHAPTER 1 - INTRODUCTION

This chapter is the work of the sole author of this thesis.

1.1 - NATURAL WORLD - STUDY FOCUS

1.1.1 - AFRICAN VEGETATION

Figure 1.1: Africa seen from space (Image: Landsat/Copernicus, Data: SIO, NOAA, U.S. Navy, NGA, GEBCO, Image: IBCAo; obtained with Google Earth Pro, 2019).

Africa is the second largest continent (behind Asia), home to the second largest human population globally, forecasted to grow the fastest in the coming decades (United Nations, 2019) and with the longest history of human settlement (McDougall, Brown and Fleagle, 2005; Smith *et al.*, 2007; Scerri *et al.*, 2018). Despite its size and population, it hosts many of the world poorest countries, but also some of the fastest growing economies (World Bank, 2019). Along its economic diversity the continent hosts an even greater cultural diversity, with over 2000 languages (Lewis, Simons and Fennig, 2016), 1000 ethnicities (Childs, 2003), 54 states, and a rich, yet under documented history (Davidson, 1966; J. Desmond Clark, 2008). Africa's chances, notably a wealth of resources (Jensen and Wantchekon, 2004), and challenges, among which violence and conflicts are prevalent (Elbadawi and Sambanis, 2000; Salehyan *et al.*, 2012), will be further stressed as climate change is likely to affect Africa more severely than other continents (Hendrix and Salehyan, 2012; Niang *et al.*, 2015).

Among Africa’s natural resources are nine biodiversity hotspots, such as the Cape floristic region, the succulent Karoo or the Guinean forests of West Africa (Myers *et al.*, 2000; M. Hoffman *et al.*, 2016). These hotspots are spread among a wide range of climates (Kottek *et al.*, 2006; Beck *et al.*, 2018) and landscapes. Africa’s large biodiversity can be classified into eight biomes (Olson *et al.*, 2001): Deserts and xeric shrublands; Tropical and subtropical grasslands, savannas and shrublands; Mediterranean forests, woodlands and scrubs; Montane grasslands and shrublands; Flooded grasslands and savannas; Mangroves; Tropical and subtropical dry broadleaf forests; Tropical and subtropical moist broadleaf forests. Other classifications exist (Monserud and Leemans, 1992; Prentice *et al.*, 1992; Friedl *et al.*, 2002), which subdivide vegetation into finer categories (Mayaux *et al.*, 2004; Mucina and Rutherford, 2006), highlighting the African landscape diversity. Yet, one powerful and emblematic image comes to mind when thinking about Africa: a sunset over the savanna.

1.1.2 - SAVANNA

Figure 1.2: Artistic depiction of a sunset over an African savanna (art by Gee).

The iconic African savanna landscape, as depicted on figure 1.2 is part of the Tropical and subtropical grassland, savanna and shrubland biome as defined by Olson *et al.* (2001). Savannas are defined as regions where grasses and woody plants co-exist to create a landscape that is neither a grassland or a shrubland nor a forest or a woodland. According to this definition, it is estimated that savannas cover approximately one sixth of the global land surface (Scholes and Archer, 1997; Grace *et al.*, 2006). However, characterisation of savannas is multifaceted. The Cambridge Advanced Learner's Dictionary describes savannas as "a large, flat area of land covered with grass, usually with few trees, that is found in hot countries, especially in Africa" (Cambridge University Press, 2008). Alternatively, Sankaran *et al.* (2005), Lehmann *et al.* (2011, 2014), Staver, Archibald and Levin (2011), Staver, Botha and Hedin (2017), explained savannas distribution by its drivers, particularly rainfall, soil and fire. Rainfall drives savannas both in terms of Mean Annual Precipitation (MAP) and in terms of seasonality, as savannas typically have a dry season and a rainy season. Pennington, Lehmann and Rowland (2018) provided a precise definition of savanna which classifies a fifth of the global land surface as savanna, based on trees and grasses cover share and including fire and herbivory as drivers. Other key characterisation schemes of savanna focus on its distinction from other biomes; for example, Ratnam *et al.* (2011) focused on its differences from forests based on plant types. Discounting the variety of definitions, savannas are rich and diverse ecosystems (Figure 1.3).

Savannas have been pivotal in human evolution, being associated with the development of the upright posture of the *Homo* genus (Clark, 1987; Lieberman *et al.*, 2007; Stewart and Stringer, 2012). They support iconic species, such as rhinoceroses, elephants, giraffes and lions. They provide traditional resources such as wild food, construction wood, fuel wood, and medicines (Twine *et al.*, 2003; Boafo, Saito and Takeuchi, 2014). They typically have strong and complex relations to large grazers and browsers herds (Archibald *et al.*, 2005; Codron *et al.*, 2007; Holdo, Holt and Fryxell, 2009; Kimuyu *et al.*, 2014; Baumgartner *et al.*, 2015). Fire strongly drives vegetation dynamics in savannas and is often used for ecosystem management (Freckleton, 2004; Bond, Woodward and Midgley, 2005). However despite their long term historical relevance for both Humanity and biodiversity, they are largely threatened (Cardoso Da Silva and Bates, 2002; Eriksen and Watson, 2009; Boon et al., 2016). Indeed, the savanna biome contains the majority of the world's remaining potential arable land (Pennington, Lehmann and

Rowland, 2018), and as land use changes rapidly, particularly in Africa, vast swathes of land are at risk of conversion (Aleman, Blarquez and Staver, 2016; Ordway, Asner and Lambin, 2017). This threat is aggravated by the impending consequences of climate change (Scholze *et al.*, 2006), notably changing rainfall regimes. Therefore, there is an urgent need to improve our understanding of savannas as it harbors vital resources for both Humankind and biodiversity.

To assess the risks and dangers faced by African savannas, we need to understand savannas ecosystems. This implies to identify and characterise their key processes and their key life-forms. Savannas are characterised by the coexistence of grassy and woody vegetation. Grasses mainly exhibit a C4 photosynthetic pathway in savannas (Sage and Monson, 1999), but C3 species can also be abundant. A few examples of the extensive grass diversity in savannas are the common finger grass (*Digitaria eriantha*), tussock grasses such as *Hyparrhenia diplandra* and *Heteropogon contortus* or lovegrasses (*Eragrostis*). The woody component of savannas is equally diverse, with life-forms ranging from bushes to trees through shrubs. The most common type of trees are of the *Acacia* genus, but Baobabs, Jackal berry tree (*Diospyros mespiliformus*), and even succulent trees such as *Euphorbia ingens* are also a trademark of African savannas. Interestingly, each has a unique strategy to deal with prevailing environmental constraints (e.g.: Acacias shed their leaves in drought conditions, Baobabs and Euphorbias store water).

Figure 1.3: Glimpses of savanna landscape diversity (photography by the author). From left to right and from top to bottom: Rocky hillside dominated by Euphorbia trees (Euphorbia ingens). Open savanna at sunset. Sunset over a white rhinoceros (Ceratotherium simum) herd on a sandy soil savanna. South African giraffes (Giraffa camelopardalis giraffe) in a semi-open savanna. African elephants (Loxodonta africana) herd in an opening in a closed savanna. Savanna recently opened by African elephants. Thicket in a savanna, hosting Impalas (Aepyceros melampus). Recently burned savanna. Scrubland savanna. Pictures taken in 2016.

1.1.3 - SHRUBS

In addition to grasses and trees, shrubs are ubiquitous through African savannas (Figure 1.3) (Cole, 1982; Lloyd *et al.*, 2008) and a distinct ecological strategy (Condit, Hubbell and Foster, 1995; Hoffmann and Solbrig, 2003). Shrubs play a key role in many biomes and ecological processes (Parmenter and MacMahon, 1983; Vieira, Uhl and Nepstad, 1994; Sirami *et al.*, 2009). They are a keystone in the phenomenon of woody encroachment (Maestre *et al.*, 2009; Stevens, Lehmann, *et al.*, 2016), which can lead to either forest expansion (Wiegand, Ward and Saltz, 2005; Mitchard *et al.*, 2009) or desertification (Grover and Musick, 1990; Archer, Boutton and Hibbard, 2001; Van Auken, 2009). Despite their relevance we do not know enough about their physiology and their ecology to fully assess shrub encroachment phenomenon. This shortcoming is related to the lack of clear consensual definition of what shrubs are. In various studies shrubs are characterised as short woody plants, however, the threshold to discriminate them against trees typically ranges from 3m to 4m, going more rarely up to 6m. This characterisation does not allow to separate them from bushes (also short woody plants) and from small trees. Recent studies proposed new criteria to characterise shrubs. Zizka, Govender and Higgins, (2014) defined shrubs based on height and stem number as multi-stemmed short woody individuals. This definition allows to distinguish clearly between trees and shrubs, but creates an intermediate category of "scrubs" (tall shrubs or multi-stemmed trees, "tree sometime shrub" in Zizka, Govender and Higgins (2014)) and does not distinguish from bushes (Figure 1.5). Nonetheless, this definition is particularly relevant and holds new potential to study shrubs as a specific life-form. It is indeed a common sight through African savannas (Figure 1.3 & 1.4).

Figure 1.4: Shrub thicket in a South African savanna during a drought (2016).

Figure 1.5: Schematic representation of life-forms of the grassland-savanna-forest ecosystems. From left to right: bush, shrub, "scrub" (tall shrub, multi-stemmed tree), tree (art by Gee).

This definition of shrubs supports the classification into distinct vegetation cover types relevant for the present thesis (Figure 1.6). Forests (closed forest) are dominated by trees with a closed canopy where light competition is among the main drivers. Savannas are landscapes where woody vegetation, made up of a mixture of trees and shrubs, coexists with a grass layer and where fire is among the main drivers. Grasslands are dominated by grasses with sparse woody plants. Shrublands are dominated by shrubs, with sparse grasses and other woody plants. Xeric shrublands are mainly barren, with vegetation layer dominated by shrubs and bushes, where aridity is the main driver. Woodlands are dominated by woody vegetation but without a closed canopy, where grasses are suppressed but not completely outcompeted (can be seen as open forest or closed savanna).

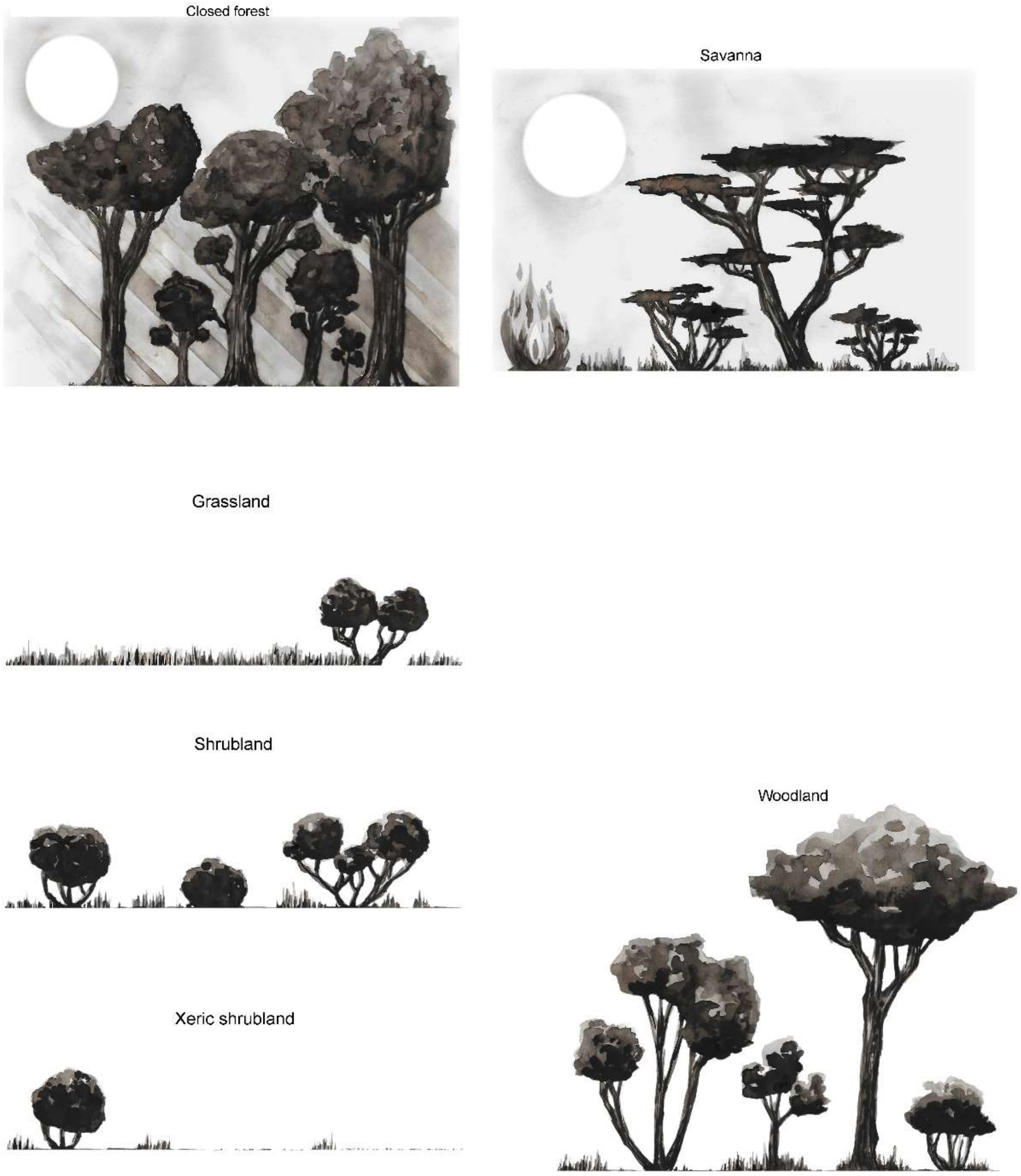


Figure 1.6: Vegetation cover types relevant to the present thesis (art by Gee).

1.2 - NUMERIC REPRESENTATION OF NATURE

1.2.1 - MODELS & ECOLOGY

The modern “model” concept can be read on multiple levels. Directly related to the word Latin origin, a model is a “reference”/”example”. Its second interpretation is further derived from the Latin origin. It says that a model is a simplified representation of an object or a system. It is a schematic visualization or a conceptualization of an item, and it is this second interpretation which is of interest to us here. Additionally, “model” is used to describe either a look-alike or function-alike representation of an object.

Ecological models are rooted historically in mathematical models representing relations between variables. Their foundations are made of observations, from which equations representing ecosystems behaviour are obtained through interpolations & linear regressions, and through theoretical models, such as the Lotka-Volterra equations (Volterra, 1926; Lotka, 1927). Their development and increase in complexity were made possible thanks to their digitalization, as it allowed for computation intensive processes to be modelled. Increase in complexity allowed for model to represent more than the sum of their parts; i.e., ecosystem dynamics are represented by multiple sub-processes instead of being directly modelled by a unique equation. This allows for greater flexibility and reduces the approximation due to data regression, thanks to the possibility to test the interplay between parameters and mechanisms. Process based ecological models are made up of many equations derived from multiple observations datasets, each limited to a specific “frame” (i.e., context or study) and not universal. This can lead to different model performing better in some cases than others even if they include the same processes, due to stemming their equations from different contexts. These assumptions, being fundamental to the functioning of these models, should be explicitly discussed and investigated more extensively (Zaehle *et al.*, 2005; Quillet, Peng and Garneau, 2010; Bachelet, Rogers and Conklin, 2015). We aim to contribute to this discussion and argue that reflecting on this shift from mathematical relation to a system representation can foster new progress in the field.

1.2.2 - DGVM - ESM - LSM

Ecological models can be divided into two broad categories based on their methodological and conceptual approach. First are statistical models, such as niche models (Monahan, 2009; Williams, Anandanadesan and Purves, 2010). Second are models simulating ecosystem states by simulating ecosystem dynamics, such as gap models (Shugart *et al.*, 2018) or DGVMs (Dynamic Global vegetation Models) (Prentice *et al.*, 2007; Fisher *et al.*, 2017), which are our specific focus. They aim to reproduce ecosystem state variables and processes. This makes them particularly relevant to perform experimental studies, which would be too costly or even impossible to perform in empirical studies, and to study past or potential future ecosystem. However, to have confidence in models results, they need to be thoroughly assessed on two points. First, the extent of models' relevance need to be accurately and explicitly framed (hereafter extent of validity). Second, the models performance and accuracy need to be benchmarked against observation data (Kelley *et al.*, 2013). However, models evaluation and relevance is intrinsically relative (Oreskes, Shrader-Frechette and Belitz, 1994).

DGVMs have a long history, stemming from a mixture of different research fields, making them trans-disciplinary (Prentice *et al.*, 2007). They first appeared in the 1990's with models such as IBIS (Foley *et al.*, 1996) or VECODE (Brovkin, Ganopolski and Svirezhev, 1997). Since then, DGVMs have been linked with ESMs (Earth System models) (Quillet, Peng and Garneau, 2010) and LSMs (Land Surface Models) (Li *et al.*, 2011) enabling the simulation of closed biogeochemical cycles (e.g.: carbon, water, energy).

Numerous DGVMs have been developed during recent decades, such as: JSBACH, JEDI, LPJ, LPJ-GUESS, aDGVM2, SEIB-DGVM, LM3-PPA, CLM(ED), ED, ED2, aDGVM, aDGVM2, ORCHIDEE (Fisher *et al.*, 2017). This leads to conflicting classification of these models. For example, van Oijen *et al.* (2018) identifies three types of models: ecological, process-based and integrated. Fisher *et al.* (2018) refers to ESM (Earth System Models) which include first and second generation DGVMs, further sub-divided into forest gap models, and vegetation demographic models, again sub-divided in individual-based models and cohort-based models. However, they rely on a common architecture framework (Prentice *et al.*, 2007). They all are process based model simulating terrestrial vegetation ecosystems relatively to environmental (such as climate variability or soil characteristics) and internal (such as competition for light, space, water or nutrients) parameters. Most of them use a set of PFTs (Plant Functional Types) representing vegetation archetypes based on unique sets of parameters controlling their behaviour and architecture (Smith, Prentice and Sykes, 2001; Prentice *et al.*, 2007), for example "temperate broadleaf deciduous tree" or "tropical humid evergreen tree". Typically, DGVMs simulate the relative share of each PFT they consider in grid-cell they are run for, based on the input data they rely on (such as climatic variables) and which define the resolution at which they can be used. Due to their similarities, they face common challenges and share common questions (Quillet, Peng and Garneau, 2010). Fisher *et al.* (2018) reviewed the most recent progress and priorities concerning ESMs with a focus on DGVMs. For example, among DGVMs various methodological and conceptual approaches exist to model light competition, water competition, vegetation demographics (Fisher *et al.*, 2017). Additionally, limitations due to data availability have to be overcome, and addressing computation informatics challenges would improve model's reliability and potentials.

The purpose of a DGVM is to be a mathematical representation of ecological processes and thus simulate ecosystem dynamics while recreating ecosystems states as realistic variables. This holistic approach is the strength and challenge of modern DGVMs. This methodological approach clearly distinguishes DGVMs from purely statistical models, such as ecological niche models.

1.2.3 - ADGVM2

Defining life-forms based on traits and not based on categories is a modern idea. Lavorel and Garnier (2002) hypothesized that plant traits can be used to predict ecosystem behaviour. This hypothesis is the key concept for designing trait-based vegetation models. While DGVMs, embedding this assumption, developed and diversified rapidly over the last two decades, attempts for a fully trait-based model emerged only recently with the JEDI model (Pavlick *et al.*, 2013) and with the aDGVM2 (Scheiter, Langan and Higgins, 2013). As first generation DGVMs and the PFT approach are not adapted to study the role of biodiversity in great details and as biodiversity has been identified as crucial for ecosystems stability, notably in the face of climate change (Tilman, 1995; Tilman *et al.*, 1997; Loreau *et al.*, 2001; Isbell *et al.*, 2015), research intensifies to go beyond PFTs and first generation DGVMs. The number of studies advocating for such paradigm shift is increasing (Yang *et al.*, 2015; Fisher *et al.*, 2017; van Oijen *et al.*, 2018).

The aDGVM2 is among the first DGVMs fully based on traits trade-off which allow, by stochastic iterative sampling of the potential trait space, to select the most adapted life-forms (each potentially with a unique set of traits values) to prevailing environmental conditions and biotic interactions, *via* trait selection through generations (considering crossovers and mutations). This makes it of the first models able to study biodiversity effects on vegetation communities and ecological processes by relying on individual functional traits and not on PFTs. This strength is also relevant to study socio-economic interplays with biodiversity. For example, it can be used to study the effect of preferential harvesting of wild plants based on specific traits on the overall vegetation community assembly through time. It can also be used to study speciation processes and their consequences on functional diversity and trait space. For example it could be used to study coevolution or evolution directionality (Gould, 1988; Hawkins, Grover and Wendel, 2008; Svensson and Berger, 2019). The aDGVM2's key strength relies on its ability to simulate potentially unique trait combination for each individual and to select adapted individuals through traits filtering (Figure 1.7). This feature allows to study the impact of traits on vegetation community assembly and to study traits based biodiversity (Langan, 2019). The aDGVM2 is described extensively in Scheiter, Langan and Higgins (2013) and Langan, Higgins and Scheiter (2017), and on specificities in Pfeiffer *et al.* (2019) (annual and perennial grasses and grazing), Gaillard *et al.* (2018) (shrubs), second chapter of this thesis, and Kumar *et al.* (2020) (C3 grasses, plant level leaf photosynthesis).

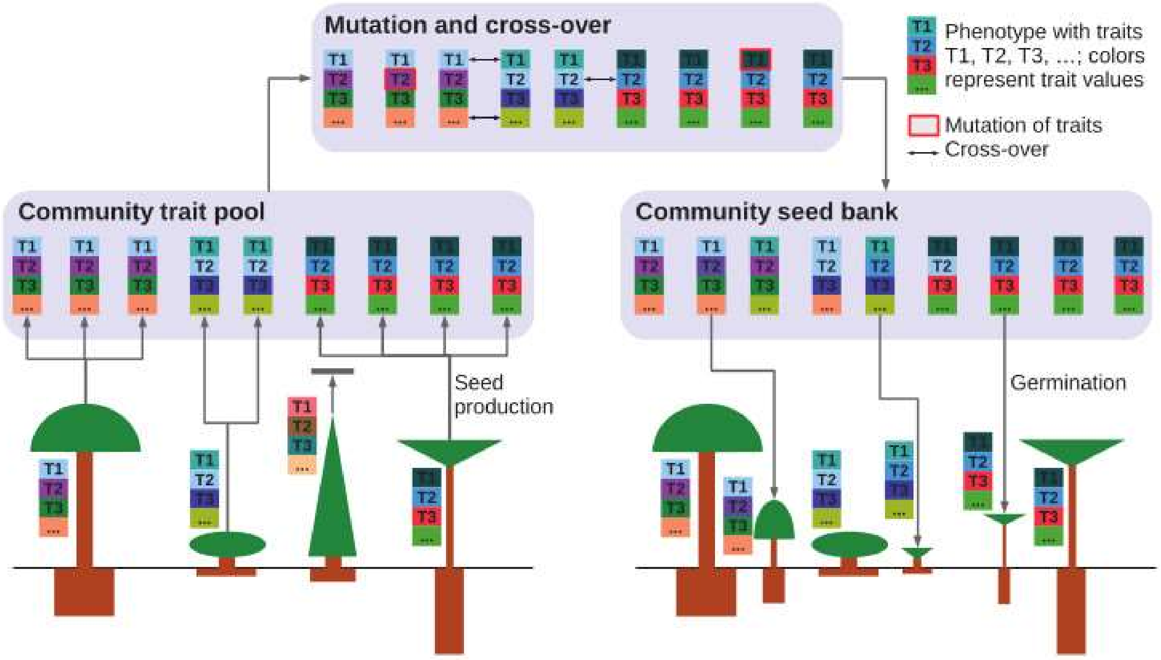


Figure 1.7: In aDGVM2 each individual is represented by a potentially unique combination of traits which is dynamically selected through traits filtering processes depending on environmental conditions, mutation and cross-over (Scheiter, Langan and Higgins, 2013).

1.3 - ECOSYSTEMS & LIFE-FORMS - STATES & DYNAMICS

1.3.1 - SHRUB ENCROACHMENT

Shrub encroachment is a globally observed phenomenon (Grover and Musick, 1990; Rivest *et al.*, 2011; Stevens, Lehmann, *et al.*, 2016). It can be associated with desertification processes (Grover and Musick, 1990; Van Auken, 2009) or with forest expansion processes (Mitchard *et al.*, 2009). It is one aspect of woody encroachment, which covers forest expansion and bush encroachment. The common feature is the observation of an increase in woody cover, over recent years (or decades) in a given region. These regions are typically grasslands (Van Auken, 2000; D'Odorico, Okin and Bestelmeyer, 2011) or savannas (Roques, O'Connor and Watkinson, 2001), but can also be more arid biomes (Jeltsch *et al.*, 1997). Shrub encroachment has been increasingly identified as a concern, as shown by the increasing number of relevant publications (Figure 1.8). Shrub encroachment impacts many stakeholders. Park managers are concerned as it alters habitat suitability for wildlife (Sirami *et al.*, 2009; Stanton *et al.*, 2018). Rangeland farmers are concerned as it alters forage availability and quality (Rivest *et al.*, 2011) as well as predation risk (Blaum *et al.*, 2007). Shrub encroachment can also have a wider impact on ecosystem functioning, local climatic conditions and hydrological regimes by modifying water cycling, albedo and surface roughness (Eldridge *et al.*, 2011, 2012; He *et al.*, 2011). It could impact the global Earth system by modifying carbon sequestration potential (Goodale and Davidson, 2002).

Despite being extensively studied, no consensus emerges at present regarding shrub encroachment, its drivers and its outcomes (Cao *et al.*, 2018). The lack of consensus could be explained partly due to the differences between case studies (from desertification to forest expansion) and partly due to the ambiguity of the definition of shrub. These differences are underlined by the fact that multiple factors can cause such phenomenon. Shrub encroachment is a multifaceted phenomenon covering widely different cases. As such, it is challenging to find a consensus about its causes. Finally, shrub encroachment is intrinsically a phenomenon for which historic legacy plays a role as it happens over decades and is directional, and thus hysteresis effects need to be considered. This also implies that analysing shrub encroachment accurately is a data intensive process, and thus, data availability is a limiting factor (Cao *et al.*, 2018).

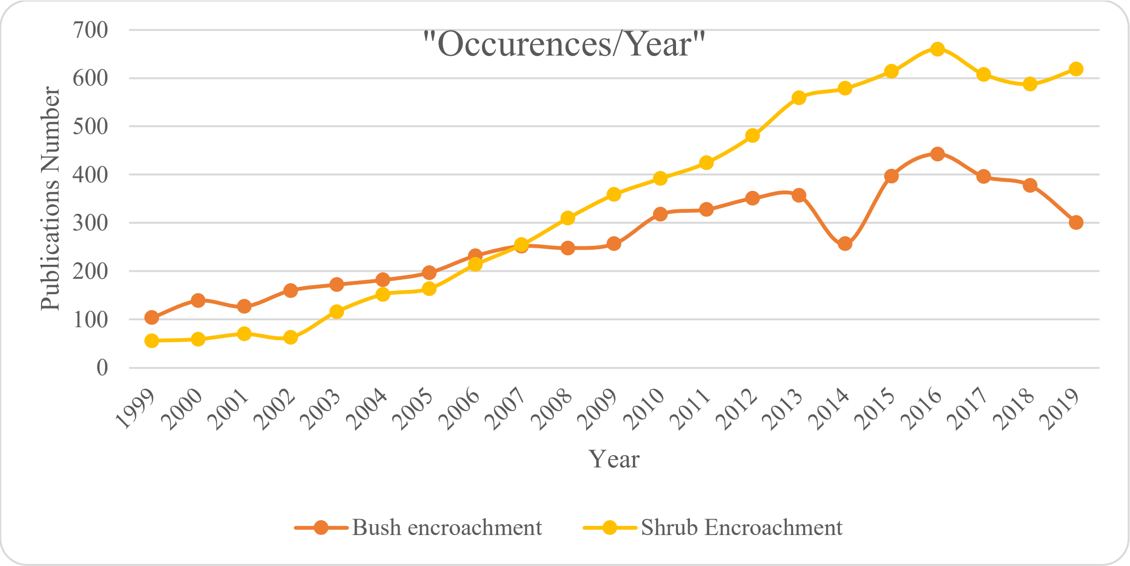


Figure 1.8: Number of peer review publications returned by a Google scholar search for the terms "shrub encroachment" (yellow) and "bush encroachment" (orange) over the last 20 years.

1.3.2 - ECOSYSTEMS VULNERABILITY

Disclaimer: This section is based on Scheiter et al. (2018), which the author of this thesis co-authored.

Ecosystems and biodiversity around the world are facing multiple threats (Cole and Landres, 1996; N Joppa *et al.*, 2016), and the level of these threats is increasing rapidly in many ecosystems (McKee *et al.*, 2004; Lenzen *et al.*, 2012). It is of paramount importance to assess threats effects on ecosystems and to forecast their trajectories, as it is likely to impact E.S. (Ecosystem Service) provision and livelihood of societies around the world in the near to mid-term future (Egoh *et al.*, 2012; Maron *et al.*, 2017). Factors such as climate change, atmospheric [CO_2] increase and land use change have been identified as large scale drivers of ecosystem change (Biggs *et al.*, 2008; Chown *et al.*, 2010; IPCC, 2013, 2014). Interactions between these factors can be complex. For example, extreme events such as droughts can reduce biomass (Ogutu and Owen-Smith, 2003; Mooney *et al.*, 2009), while atmospheric [CO_2] increase might foster vegetation growth (Wigley, Bond and Hoffman, 2010; Buitenwerf *et al.*, 2012). In Scheiter *et al.* (2018), we showed that grasslands are particularly at risk of conversion to savanna while savannas might convert to woodlands, and we showed that woody encroachment phenomenon is likely to lead to large scale change in vegetation cover. This risk is higher in more extreme climate change scenarios, i.e., when [CO_2] rises unimpaired at current rate until the end of the century. Kumar and Scheiter (2019) showed similar trends for South and South-East Asia. Globally, vast expenses of diverse ecosystems are at risk of conversion due to climate change and atmospheric [CO_2] increase (Cramer *et al.*, 2001; Gonzalez *et al.*, 2010)

This shift in vegetation cover will have cascading consequences as it will modify habitat suitability. Thus, by changing vegetation cover, climate change (and particularly CO_2 increase), will impact biodiversity; components of which might exhibit different responses to similar changes. This might lead to unprecedented ecosystem states.

In Scheiter *et al.* (2018) we showed that grassland is the biome most likely to shift to a different state. However, changes inside biomes might be larger than shifts between biomes, given that a change of 2 or 3 percent can be enough to cross the threshold between biome categories, and given that a biome might encompass a wide range of values for a given parameter. The savanna biome encompasses a wide range of woody cover, thus, a larger increase in woody cover can happen without being recognized by looking at biome shifts only. Thus, we need to further our analysis by taking a non-categorical perspective on vegetation. Additionally, simulations in Scheiter *et al.* (2018) were conducted with aDGVM, which could not distinguish between shrubs and trees according to a trait-based trade-off. We turned to aDGVM2 to develop such capability as it allows for community assembly processes to dynamically select the most adapted traits combinations. This allows us to investigate woody encroachment in greater details and enables us to distinguish between shrub encroachment and forest expansion, adding multiple layers of knowledge to our analysis. For example, woody cover expansion means higher potential for carbon sequestration. Carbon sequestration potential is different between shrubs and trees. If we know to which extent woody cover expansion is due to shrubs or trees we can more accurately estimate change in carbon sequestration potential. Similarly, habitat suitability is different between shrub thicket and forests, thus impacting biodiversity. Finally, drivers of shrub encroachment might be found to be different from drivers of bush encroachment or forest expansion (Ward, 2005; Wiegand, Ward and Saltz, 2005; Mitchard *et al.*, 2013; O'Connor, Puttick and Hoffman, 2014; Stevens, Lehmann, *et al.*, 2016).

1.3.3 - MODELLING SHRUBS

At present, there are few DGVMs explicitly simulating shrubs, and they typically rely on a PFT approach (Hickler *et al.*, 2006; Zeng, Zeng and Barlage, 2008; Clark *et al.*, 2011; Lawrence *et al.*, 2011), which differs from a mechanistic representation of shrubs. Shrubs are understudied (relatively to trees), poorly categorised in terms of life-form and in terms of life strategy, and, consequently, their ecological role is unclear (Gaillard *et al.*, 2018). The PFT based representation of shrubs in DGVMs does not allow to define a trait-based trade-off specific to shrubs. Shrub encroachment studies, despite their growing number, do not usually provide better information about the general physiological or ecological specificities of shrubs, partly because of a lack of common definition in the literature (Wilson, 1995; Zizka, Govender and Higgins, 2014; Götmark, Götmark and Jensen, 2016) and partly because they focus on shrub encroachment phenomenon more than on shrub life-forms (Jeltsch *et al.*, 1997; Roques, O'Connor and Watkinson, 2001; Knapp *et al.*, 2008; Eldridge *et al.*, 2011; Eldridge and Soliveres, 2014). However, we can use field observations to establish defining criteria for shrubs. In this regard, Zizka, Govender and Higgins (2014) characterised shrubs, in African savannas, as multi-stemmed individuals. This definition is particularly relevant as it is drawn from our focus region. It informs our view on shrubs as a life-form but does not directly show a mechanistic trait-based trade-off related to a specific life strategy. However, the consequences of this phenotypical difference (Figure 1.9) motivate a translation into the aDGVM2 framework. Due to allometric relations, for a given biomass, an increasing number of stems implies a decreasing height, simultaneously, an increasing number of stems means an increasing sapwood area (assuming a constant ratio). We question to which extent this allows to simulate shrub distribution and which consequences it has on plants processes and on how they react to environmental constraints. This is explained in great details in chapter 2.

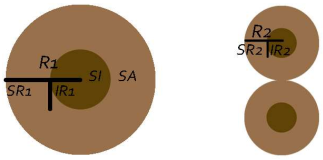


Figure 1.9: Relation of stem number to sapwood area relative to total diameter, the sapwood area of two stems is greater than the sapwood area of one stem for the same total stem radius. SI: Heartwood (dark); SA: Sapwood Area (light); R: Radius; SR: Sapwood "Radius"; IR: Internal (heartwood) Radius.

1.3.4 - ASSESSING MODELS

Models are benchmarked by comparing their output to observation data to ensure that models represent what they aim to and to assess their accuracy. An increasing number of parameters and variables are available to benchmark models; which can improve models reliability. In parallel, new model development require new benchmarking (Prentice *et al.*, 2007; Fisher *et al.*, 2010; Kelley *et al.*, 2013). We identify two complementary approaches to model benchmarking. First, model benchmarking can be done to support model development; i.e., benchmarking for testing model accuracy and relevance, and to improve a model (Fisher *et al.*, 2010). Second, model benchmarking can be done to assess the accuracy of the representation of given specific variables or processes; i.e., benchmarking for testing an ecological hypothesis (typically among multiple different models). The second approach can imply model intercomparison projects to test different approaches to model a given ecological question (Cadule *et al.*, 2010; Loew *et al.*, 2013; Peng *et al.*, 2015; Zhang *et al.*, 2018) or focused on the benchmarking process itself (Kelley *et al.*, 2013; Best *et al.*, 2015). Intercomparison projects foster the homogenization of benchmarking approaches, which is widely advocated for (Warszawski *et al.*, 2014; Eyring *et al.*, 2016; Rosenzweig *et al.*, 2017). The call for benchmarking standardization goes beyond the DGVM community and encompasses the whole Earth system modelling community (Abramowitz, 2012). Benchmarking trait-based DGVMs has many similarities with benchmarking PFTs based DGVMs, but it also has unique specificities. In trait-based models, vegetation is not categorised *a priori* and it needs to be characterised *a posteriori* into life-forms and life strategies. This is a unique challenge, but it also implies unique opportunities which are discussed through this thesis, such as: more detailed representation of ecosystem functions, higher flexibility in vegetation communities, or evolutionary trajectories of plant communities.

1.4 - STUDY’S RELEVANCE

1.4.1 - MODELS

The target of this thesis is to unravel ecological processes and dynamics explaining biogeographic patterns. It implies the intertwining of two items: model and ecology. A model is both a tool and a research question in itself, as investigating our ecological questions raises questions about modelling methodology and requires questioning them explicitly. At present, second generation DGVMs are the most common state of the art process based ecological models (Fisher *et al.*, 2017). They can be individual-based, which ensures a level of realism in the representation of plant level processes (Sato, Itoh and Kohyama, 2007; Scheiter, Langan and Higgins, 2013). Others are cohort based, typically using PFTs, which limits their ability to represent the diversity of life-forms and life strategies diversity as well as community assembly processes (i.e., natural selection), compared to individual based models. This impairs the development of a shrub model that relates forms and function (plant strategy) in a mechanistic way and not *a priori* determined. Thus, despite several attempts to model shrubs in DGVMs (Hickler *et al.*, 2006; Zeng, Zeng and Barlage, 2008; Clark *et al.*, 2011; Lawrence *et al.*, 2011), none fitted our needs so far. Therefore, we needed to rely on the first third generation DGVM, aDGVM2 and develop a new compatible shrub model.

Investigating ecological questions leads to questioning modelling concepts and benchmarking methodologies which can prove relevant for the development of DGVMs and supports and informs the research effort towards fully trait-based models. We argue that more process-based and non-categorical models, such as aDGVM2, offer unprecedented opportunities to investigate ecological questions. We argue for more holistic benchmarking approaches to explore the challenges and opportunities opened by such innovative model architecture. Our approach also highlights conceptual challenges to be solved to improve models reliability and capabilities. Most importantly, we demonstrate that it is possible to model a specific life-form and its associated life strategy without directly defining it *a priori*.

1.4.2 - ECOLOGY

Woody encroachment has been identified as a major threat to many African biomes (Midgley and Bond, 2015; Stevens, Lehmann, *et al.*, 2016; Scheiter *et al.*, 2018). Going beyond the PFT approach, our shrub representation based on mechanistic trait trade-off allows to model dynamically ecosystem processes and thus enable a deeper investigation of shrub encroachment processes. Many studies contributed to disentangle the mesh of causes and consequences of this global phenomenon (Roques, O'Connor and Watkinson, 2001; A Carla Staver, Archibald and Levin, 2011; Stevens, Lehmann, *et al.*, 2016) and to test potential management scenarios (Tocco *et al.*, 2013; Boon *et al.*, 2016; Smit *et al.*, 2016; Case and Staver, 2017b; Daryanto, Fu and Zhao, 2019). With our present model, we further these investigations. Besides shrub encroachment as an ecologically relevant phenomenon, shrubs themselves have direct effects on ecosystems functions. For example, shrubs define habitat suitability for different wildlife species (Popp *et al.*, 2007; Sirami *et al.*, 2009; Smit and Prins, 2015; Soto-Shoender *et al.*, 2018) and can modify animals behaviours, for example by impacting hunting behaviour (Blaum *et al.*, 2007).

We conceptualised and tested representation of shrubs as a specific life-form with a unique strategy, specifically for African savannas. We identified their key characteristics based on their architecture and dynamics and provide a comprehensive characterisation of shrubs, more adapted for DGVMs, namely, a trait-based process-based definition of life-strategies without *a priori* constraints.

As systems closer to critical thresholds respond more sensitively to external perturbations (Scheffer *et al.*, 2009), it is necessary to increase models accuracy and precision. This implies to improve models ability to grasp finer details. The development of our shrub model allows for greater details based on a dynamic, flexible and innovative modelling approach; as such, it represents a major contribution towards a better understanding and forecasting of scenarios with more detailed ecosystem responses to large scale climatic and anthropogenic changes.

1.4.3 - SOCIETIES

The ecological questions investigated and ecosystem dynamics modelled for this thesis are embedded in the social, political and economic context of our modern society with manifold actual or potential impacts. Particularly of concern is the fact that ecosystems resistance, robustness and resilience are challenged by environmental drivers which can lead ecosystems to cross thresholds, or tipping points, implying a point of no return between distinct alternative dynamic equilibriums (Gillson, 2004, 2015; Lenton *et al.*, 2008; Lenton, 2011; Veraart *et al.*, 2012; M. Scheffer *et al.*, 2015). These ecological switches can be driven by shifts in prevailing environmental conditions. For example, a one year drought is an extreme event, however, a long term increase in drought frequency and/or drought length is a shift in bioclimatic conditions, which can induce longer lasting effect of greater magnitude on societies. An example is the impact of drought on the Syrian society, which contributed to its destabilization leading to the current war (De Châtel, 2014; Gleick and Gleick, 2014; Kelley *et al.*, 2015), even if it is not the only cause. Many historical cases when environmental changes impacted social structures have been documented, notably in "Collapse: How Societies Choose to Fail or Succeed" (Diamond, 2005). This question is prone to debate due to the specificities of studying historical events (Page, 2005; Peiser, 2005; Hunt, 2006; Schwartz and Nichols, 2010; Butzer, 2012). Studying potential future scenarios has the additional difficulty that we are at present facing unprecedented environmental challenges (IPCC, 2014; Díaz *et al.*, 2019). This study is a step towards addressing the concerns stemming from current and future environmental and climatic changes. To do so, we provide higher level of detail to represent shrubs, specifically in African savannas, and thus enable higher capability to study shrub encroachment phenomenon, while supporting the development of new advancements in ecological modelling.

Shrub encroachment is a relatively slow process, at human life time scale; yet it is relatively fast at ecological time scale. It is a typically gradual process, which makes it difficult to distinguish between a pre-encroachment and a post-encroachment vegetation state. Therefore, defining a tipping point between these two states requires more detailed insights, which our approach contributes to. Our shrub model also enables the study of ecosystems resilience and resistance against shrub encroachment phenomenon, and the directionality of the phenomenon. This is particularly relevant to societies as shrub encroachment has direct impact on what a land can provide. For example by decreasing fodder availability for grazers, it reduces rangeland farming (MacLeod, Brown and Noble, 1993; Mugasi, Sabiiti and Tayebwa, 2000) and/or hunting potential. This would consequently affect economical assets of commercial farmers or, more dramatically, reduce the robustness of the natural safety net for local population, thus increasing the danger of following extreme climatic events or economic crisis (Sallu, Twyman and Thomas, 2009; Richardson, 2010). Besides potentially devastating or game changing thresholds effects, having the ability to finely manage ecosystems is a key to sustainable development.

1.5 - QUESTIONS & OBJECTIVES

1.5.1 - AIMS & GOALS

(1) The cornerstone of this thesis was the investigation of plants communities' structures and dynamics in African savannas. (2) To investigate this topic the representation of African savannas in the aDGVM2 was improved. (3) Specifically a mechanistic shrub model based on whole plant trait trade-offs was developed. This technical goal had a combined conceptual goal, which was to (4) demonstrate that unique life-forms with a specific strategy can emerge dynamically from modelled ecological processes without *a priori* explicit characterisation. This development allowed to expand the reach of this thesis in two directions. (5) Models structure and development are supported by the comprehensive investigation of their underlying hypotheses and concepts. (6) Drivers of shrub encroachment are investigated, to inform society about management of African savannas vegetation, and to further ecological understanding of this phenomenon.

1.5.2 - CHAPTERS & QUESTIONS

"How vulnerable are ecosystems in the Limpopo to climate change?" (Scheiter *et al.*, 2018), introduced the questions investigated in this thesis. It showed that vegetation in the Limpopo province, South Africa, is likely to face an increase in woody cover, leading to biome shifts. The biomes most at risk are grasslands converting to savannas, and aDGVM also modelled extensive woody increase in savannas. Part of the uncertainty about these results can be explained by the dynamics of this vegetation cover change, as aDGVM could not distinguish between trees and shrubs. According to these findings, we turned to aDGVM2 to develop a shrub model. In chapter 2, "African shrub distribution emerges via a trade-off between height and sapwood conductivity" (Gaillard *et al.*, 2018) we introduced a novel shrub model. We demonstrate that shrubs can be modelled as a distinct life-form with a specific strategy dynamically emerging from a trait-based trade-off based on stem number. The development of this model underlies our subsequent chapters. Using the novel implementation of shrubs we answered the following questions:

1. Does the extended model allow us to reproduce observed contemporary shrub distribution patterns in Africa?

2. In which precipitation and fire regimes are shrubs dominant, co-dominant or outcompeted?

3. Do shrub and tree life-forms emerge from community assembly based on the implemented multi-stemmed architecture?

The answers to these questions show that key traits and whole-plant trade-off between traits are necessary to simulate the biogeographical distribution of shrubs and trees. Therefore, in chapter 3, “Shrub form & function in ecosystems structure & dynamics, simulated with aDGVM2 in African savannas”, we refined the analysis of the model behaviour and representation of key ecosystem dynamics. Specifically, we explored how the shrub strategy emerging from the implemented trade-off behaves relatively to the categorisation scheme considered. We found that modelled shrub behaviour supports our key assumptions and allows identifying shrubs as a distinct strategy. We reached this goal by answering the following questions:

1. How do we ensure the relevance and accuracy of our shrub model and in which cases and under which assumptions can we use it?

2. How do shrubs, as we model them, shape ecosystems simulated by aDGVM2?

3. How do distribution pattern of shrubs across Africa responds to environmental variables?

4. How do categorisation schemes of life-form shape our understanding of these life-forms and how can we ensure the relevance of a classification scheme?

5. Can we characterise shrubs as a specific life-strategy similarly to how we characterise them as a specific life-form? Specifically, we ask: how do shrubs behave in terms of resprouters *vs.* reseeders, in terms of rooting patterns and in terms of carbon allocation?

We explored DGVMs current challenges and opportunities in chapter 4, “Challenges and opportunities of models integrating traits and processes to simulate ecosystems states and dynamics”, based on our findings and general modelling concepts. We discuss approaches to improve DGVMs while highlighting the need for a better consideration of model foundations. We also focused on the complexity of validating models and on how to better connect models and observations. To this end, we answered the following questions:

1. How are assumptions underlying model structure and operation limiting models relevance?

2. How are the limits to data and models interacting and impairing model benchmarking?

3. Are there limits to modelling which cannot be overcome and why?

4. Which limits to modelling can be overcome and how?

We used aDGVM2 new ability to model shrubs as a life-form with a specific strategy to perform an assessment of shrub encroachment drivers in chapter 5, “Modelling impacts of atmospheric [CO_2], fire, grazing and MAP on shrub encroachment with aDGVM2”. We tested a matrix of scenarios to find out how shrubs prevalence responds to different levels of drivers pressure. In these scenarios we considered atmospheric [CO_2], fire, grazing and MAP. This allowed us to contribute to the assessment of shrub encroachment future trajectories by asking:

1. How do atmospheric [CO_2], grazing, fire and rainfall shape shrubs prevalence across southern African savannas?

2. Are the interactions between these factors synergistic or antagonistic?

3. How understanding shrub encroachment drivers shapes our understanding of the future of shrub encroachment cases?

4. Can fire management and grazing contribute to suppress undesired shrub encroachment?

Finally, we discuss the outcomes and reach of our thesis. We review and highlight the major results of our research and their relevance for the advancement of science. We present branching studies for which we conducted preliminary investigations highlighting their potential. The most promising project: “Modelling vegetation structure for Ecosystem Services - Benchmarking aDGVM2 with LIDAR data” aims to use modelled vegetation structure to estimate Ecosystem Services provision potential. It targets southern African ecosystems, specifically in parks where data is available about the interaction between wildlife tourism and vegetation structure. Benchmarking and calibrating vegetation structure simulated by aDGVM2 can be achieved with LIDAR (Light Radar) data locally available. This study questions: 1. Can we use LIDAR data to assess and constrain aDGVM2’s simulation of vegetation structure? 2. How can we translate aDGVM2 representation of vegetation structure into E.S. (Ecosystem Service) supply potential for wildlife tourism? 3. How our methodology can inform decision making for natural parks management? We identify three additional topics holding great potential which can be investigated with aDGVM2 based on the developments presented in this thesis. First, the relation between crown base height, light competition and fire avoidance can be investigated to improve our understanding of the relation between phenotypes as well as fire and shade adaptation strategies. Second, modelling the effect of grass self-shading can improve our representation of their life cycle and their competitive interaction with woody plants across savannas. Third, focusing on resprouting as a fire-adaptation mechanism would improve our representation of shrubs *per se* and in contrast to reseeders. We propose further openings, drawing on the specific strengths of aDGVM2 as well as going beyond its scope, to study key ecological questions from a modelling perspective. First, questioning the relation between shrubs and fire, relatively to crown fires would be particularly relevant to model shrublands such as the Fynbos and bring new capabilities to aDGVM2. Second, to investigate historic dynamics and biodiversity trajectories we propose to focus on coevolution constraint on trait space and ecosystems dynamics. We propose to study such dynamics by focusing on competitive interactions in mixed forest through vertical foliage distribution strategies. We conclude on questions underlying our thesis.

CHAPTER 2 - AFRICAN SHRUB DISTRIBUTION EMERGES VIA A TRADE-OFF BETWEEN HEIGHT AND SAPWOOD CONDUCTIVITY

This chapter, and its annex, were published as
Camille Gaillard[1], Liam Langan[1], Mirjam Pfeiffer[1], Dushyant Kumar[1], Carola Martens[1,2], Steven I. Higgins[3], Simon Scheiter[1]; "*African shrub distribution emerges via a trade-off between height and sapwood conductivity*", Journal of Biogeography, 2018; 1-12, DOI:10.1111/jbi.13447

1: Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt am Main, Germany ; 2: Institut für Physische Geographie, Johann Wolfgang Goethe-Universität Frankfurt am Main, Frankfurt am Main, Germany ; 3 Lehrstuhl für Pflanzenökologie, Universität Bayreuth, Bayreuth, Germany

C.G. lead the manuscript and the study; C.G., L.L., and S.S. conceived and implemented the idea, performed the simulations, and performed the analysis; S.S. lead the analysis; all authors contributed to the manuscript and overall model development.

ABSTRACT

Aim: Shrubs are multi-stemmed woody plants and are a successful growth form in many ecosystems globally. Ecosystems such as the Fynbos in South Africa or mediterranean shrublands are dominated by a dense shrub cover. In savannas, shrubs are a major vegetation component. Despite the importance of shrubs at the global scale, they are, in contrast to trees, often understudied both in empirical and modelling studies. We define shrub and tree strategies by a trade-off between water uptake capacity and height growth, and aim to explore if this trade-off allows us to explain shrub distribution.

Location: Sub-Saharan Africa

Methods: We improve a dynamic vegetation model, the adaptive Dynamic Global Vegetation Model version 2 (aDGVM2), to simulate shrubs as multi-stemmed woody plants, based on a trade-off between rapid height growth in single-stemmed trees and efficient water uptake in multi-stemmed shrubs.

Results: We show that, in aDGVM2, (a) the implemented trade-off allows a multi-stemmed shrub strategy to emerge and is sufficient to simulate the broad distribution of shrubs in African savannas; (b) fire and aridity promote shrubs at the expense of trees and grasses; and (c) the presence of shrubs influences competitive interactions between grasses and woody vegetation.

Main conclusion: We provide a novel approach to simulate shrubs in a dynamic vegetation model. This approach enhances our understanding of the distribution of shrubs, but further work is required for arid and mediterranean shrublands. We conclude that introducing fundamental trade-offs between growth forms into vegetation models can improve vegetation representation.

Keywords: aDGVM, multi-stem architecture, plant life-form, savanna, shrub encroachment, shrubs, shrub-tree dynamics, trait trade-off

2.1 - INTRODUCTION

Shrubs are the dominant or codominant plant growth form in many ecosystems around the world (Olson *et al.*, 2001). Ecosystems with a substantial shrub component cover c. 45% of the land surface (Götmark, Götmark and Jensen, 2016). Some shrublands are biodiversity hotspots, such as the Mediterranean Maquis, dominated by aromatic shrubs (e.g., *Lavendula (L.)* and *Rosmarinus (L.)* species), or the South African Fynbos, dominated by *Proteaceae (Juss.)* species. Shrubs can be major component in other biomes such as tropical savannas (Eldridge *et al.*, 2011). In these biomes, shrubs are important for the functional diversity and ecosystem functioning (Olson *et al.*, 2001). Alterations in shrub abundance may signal transitions between alternative vegetation states, either towards woody vegetation expansion (Maestre *et al.*, 2009), or towards less woody biomass, forest degradation, and desertification (Archer, Boutton and Hibbard, 2001).

In Africa, vegetation cover has recently changed rapidly (Skowno *et al.*, 2017). These changes are expected to continue in the future (IPCC, 2014) and savannas are especially vulnerable to biome shifts (Higgins and Scheiter, 2012; Scheiter *et al.*, 2018). Shrubs are naturally occurring in savannas and grasslands, where vast regions are subject to shrub encroachment (Mograbi *et al.*, 2015). Studies have identified changes in grazing and fire regimes, climate change, elevated atmospheric CO_2, and historical legacies as the main drivers of shrub encroachment (Stevens, Erasmus, *et al.*, 2016). How these factors interact and drive shrub encroachment remains elusive. An improved understanding of the functional differences between shrubs and trees is necessary to gain insights into the mechanisms driving the distribution of these growth forms.

The interest for understanding the ecology and biogeography of shrubs is growing (Marten Scheffer *et al.*, 2015), yet the distinction between shrub and tree growth forms is not trivial. One difference between shrubs and trees is that trees are typically single-stemmed while shrubs are typically multi-stemmed (Zizka, Govender and Higgins, 2014). This difference implies a fundamental trade-off (Midgley, 1996; Götmark, Götmark and Jensen, 2016). The multi-stemmed architecture means, for a given biomass, shorter individuals. Thus, shrubs are typically smaller than trees and seldom taller than 3-6m (Zizka, Govender and Higgins, 2014; Götmark, Götmark and Jensen, 2016). However, for a given biomass, shrubs have relatively larger total stem cross-section area and therefore higher sapwood area and higher hydraulic conductivity (Götmark, Götmark and Jensen, 2016). Hence, shrubs have a competitive advantage under more water-stressed conditions. Many shrub species can cope with fire due to their ability to resprout efficiently after being damaged by fire (Hoffmann and Solbrig, 2003) or by regrowth from a persistent seed bank in systems where crown fires are frequent (Pausas *et al.*, 2004). In contrast, trees grow taller than shrubs and typically cope with fire by escaping the flame zone and thereby becoming fire resistant (Hoffmann *et al.*, 2012; Zizka, Govender and Higgins, 2014). Tree height growth also implies both a competitive advantage for trees in dense ecosystems where light is limiting and defence against herbivores by keeping leaves away from browsers.

Empirical studies provide valuable insights into the biogeography of shrubs and trees and the impacts of climate and land use. Yet, models are required to predict future vegetation dynamics in response to climate change and anthropogenic impacts. Dynamic Global Vegetation Models (DGVMs)(Prentice *et al.*, 2007) are useful tools to investigate vegetation patterns at large spatio-temporal scales. However, DGVMs often focus on tree Plant Functional Types (PFTs) while shrubs are often not explicitly considered. Several DGVMs, such as LPJ-GUESS, CLM-DGVM, or SDGVM, do include shrubs (Hickler *et al.*, 2006; Zeng, Zeng and Barlage, 2008; Clark *et al.*, 2011; Lawrence *et al.*, 2011), but their representation is often simplified and lacks explicit mechanisms and trade-offs related to the multi-stemmed architecture of shrubs (Götmark, Götmark and Jensen, 2016).

Two reasons may explain why shrubs have received little attention in DGVMs. First, shrubs seem to be a relatively understudied ecosystem component. A Google Scholar search resulted in roughly seven times fewer results for “shrub” than for “tree”. Second, defining shrubs on a conceptual level and discriminating them from trees is challenging (Zizka, Govender and Higgins, 2014; Götmark, Götmark and Jensen, 2016). We argue that the ability to model shrubs and trees as distinct growth forms could improve our understanding of the current biogeography of many ecosystems and enhance future projections of impacts of climate change and land use.

Here, we present the first DGVM (the adaptive Dynamic Global Vegetation Model version 2, aDGVM2), that explicitly simulates shrubs and trees as multi-stemmed and single-stemmed woody plants, respectively. This difference in stem numbers creates a trade-off between faster height growth of single-stemmed trees versus increased relative water uptake capacity of multi-stemmed shrubs (Zizka, Govender and Higgins, 2014; Götmark, Götmark and Jensen, 2016). Selection and trait-filtering processes in aDGVM2 dynamically assemble plant communities that are adapted to site-specific environmental conditions (Scheiter, Langan and Higgins, 2013; Langan, Higgins and Scheiter, 2017), such that the model allows us to project the distribution of shrub- and tree- dominated communities. We use the updated aDGVM2 to study grass-shrub-tree interactions in Sub-Saharan Africa. In particular, we address the following questions:

1. Does the extended model allow us to reproduce observed contemporary shrub distribution patterns in Africa?

2. In which precipitation and fire regimes are shrubs dominant, codominant or out-competed?

3. Do shrub and tree life-forms emerge from community assembly based on the implemented multi-stemmed architecture?

By addressing these questions, we show that fundamental traits and whole-plant trade-off between traits are necessary to understand the biogeographical patterns of shrubs and trees.

2.2 - MATERIALS AND METHODS

We used the aDGVM2 described in Scheiter *et al.* (2013) and updated by Langan *et al.* (2017). We provide details about the general features of aDGVM2 and updates related to ecophysiology in the Supporting Information Appendix S1. Here, we only describe our novel implementation of multi-stemmed woody plants.

2.2.1 - MULTI-STEMMED WOODY PLANTS

The number of stems of a woody individual (Ns) is modelled as a trait. It is fixed for each simulated individual and does not change during its lifetime. However, it can be modified via mutation and exchange of traits while being passed on to the next generation. By these processes, aDGVM2 simulates communities with a variety of plants with different stem numbers. We constrained Ns between 1 and 10. In aDGVM2, each woody individual has carbon pools for roots, storage, reproduction, stem, leaf, and bark. To simulate multi-stemmed individuals, we split stem, leaf, and bark pools into Ns equal parts. Each of these parts is considered an individual stem of the plant. We use the existing aDGVM2 equations (Langan, Higgins and Scheiter, 2017) to calculate architecture (height, stem diameter, crown area) and sapwood area of each stem from its biomass. Total sapwood area of a plant is calculated as a sum of the sapwood areas of all stems. Following Kumagai *et al.* (2005), we assume a linear relation between stem number and plant sapwood area, although other studies suggest that this relation can be nonlinear and saturate as stem number increases(De Micco, Aronne and Baas, 2008).

A plant's crown area is calculated as the sum of the crown areas of all stems. To constrain the crown area of multi-stemmed plants, we assume that it is always less or equal to the crown area of a single-stemmed plant with the same stem biomass. This simplification considers that crowns of individual stems may overlap, which reduces the total crown area and increases self shading within the crown. Leaf area index of a plant is derived from its leaf biomass and total crown area. In contrast to stem, bark, and leaf biomass, we do not split root, storage, and reproduction biomass, and all stems are connected to a single root biomass pool (unit pipe model) (Tyree and Ewers, 1991). The structure of our model implies that multiple stems of an individual start underground, at the ground, or close to the ground.

Introducing stem number as a trait implies a trade-off between single-stemmed and multi-stemmed woody plants (Figure 2.1). For a given woody biomass and similar carbon allocation traits, single-stemmed plants are taller than multi-stemmed plants. Being tall is an advantage in dense vegetation stands with intense light competition. In fire-prone environments, rapid height growth may help tree crowns to escape the flame zone and thereby increase fire resistance (Hoffmann, Orthen and Do Nascimento, 2003). Additionally, single-stemmed trees have a higher bark thickness than multi-stemmed plants for a given amount of carbon allocated to bark. Bark thickness tends to correlate with stem diameter, and thick stems overall have thicker bark than thin stems (Williams, Witkowski and Balkwill, 2007). Higher bark thickness, in aDGVM2, reduces topkill probability (Langan, Higgins and Scheiter, 2017). In our model, shrubs need to invest more carbon in their bark than trees to achieve the same bark thickness.

In aDGVM2, fire damage of multi-stemmed plants is more pronounced because they can typically not escape the flame zone due to their lower height. However, for any given height, a shrub has higher biomass than a tree (Figure 2.1), and therefore potentially higher photosynthetic capacity. Thus, in aDGVM2, when fire frequency prevents plants from escaping the flame zone, multi-stemmed individuals are better competitors due to their ability to accumulate more carbon at low height, which entails a higher potential for regrowth due to larger root systems. Additionally, multi-stemmed individuals have a relatively enhanced photosynthetic capacity resulting from their higher sapwood area per unit stem biomass (Figure 2.1). Relatively higher sapwood area entails, in aDGVM2, an increased ability to extract soil water, such that, for a given height, multi-stemmed individuals gain a competitive advantage over single-stemmed plants in water-limited environments.

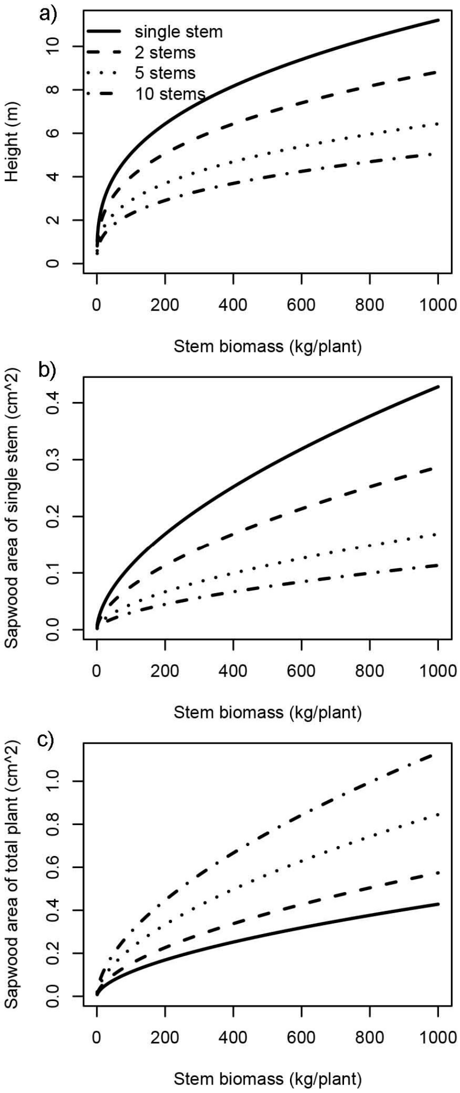


Figure 2.1: Trade-off between single- and multi-stemmed woody plants. Height (a) and sapwood area (b) of each stem of a plant are calculated from each stem's biomass using equations provided by Langan et al. (2017). Total sapwood area of a plant (c) is the sum of individual stem sapwood areas. For a given stem biomass, single-stemmed plants are taller than multi-stemmed plants, whereas plant sapwood area is higher for multi-stemmed plants.

2.2.2 - SIMULATION EXPERIMENTS

Using the extended version of aDGVM2, we simulated potential natural vegetation in Africa at 1° spatial resolution. To ensure that simulated vegetation has sufficient time to adapt to prevailing environmental conditions, we conducted simulations for 400 years using iterated monthly data from the Climatic Research Unit reference climatology for the period between 1961 and 1990 (New *et al.*, 2002). We generated daily rainfall sequences from the climatology by using the algorithm provided by New *et al.* (2002). Trial simulations showed that a 400 year period is sufficient to ensure that dynamic equilibrium conditions are reached.

We conducted four different simulation scenarios. First, we simulated vegetation in the presence of fire and multi-stemmed woodyplants. This baseline scenario is closest to natural vegetation dynamics. Second, we repeated the simulations but with fire suppressed in order to test how fire influences the distribution of multi-stemmed woody plants. Third, fire was allowed but shrubs were excluded from simulated vegetation by constraining the stem number to one for all woody plants. This scenario allows to analyse how the introduction of multi-stemmed woody plants modifies vegetation patterns and comparisons with the original version of aDGVM2 that only simulated single-stemmed trees. Fourth, we simulated vegetation in the absence of shrubs and fire.

We conducted site-specific simulations to compare model simulations, observations, and remote sensing products at site scale. For these sites, we ran 10 replicates to account for effects of stochasticity in the model. Stochasticity influences for instance the subroutines for fire, demography, cross over, and mutation of traits. We ran the simulations for eight savanna sites for which ground observation data were available (Table 2.1) (Rushworth, 1975; Rutherford, 1975, 1979; Kelly and Walker, 1976; Higgins *et al.*, 2007). For these sites, we used the 10min resolution climatology by New *et al.* (2002) as climate forcing.

2.2.3 - VEGETATION CLASSIFICATION

To analyse structure, diversity, and function in the woody layer, we classified individual simulated woody plants into different growth forms. We distinguished between four architectural types: small and tall trees, shrubs, and mixed types. Small trees are less than 4 m tall with a maximum of three stems. Tall trees exceed a height of 4 m with a maximum of three stems. Shrubs are less than 4 m tall with four or more stems. Woody plants taller than 4 m with four or more stems are classified as mixed types (shrubs, sometimes small trees). This classification allows us to be consistent with Zizka *et al.* (2014), who found that trees have on average 2.2 ± 1.4 stems, and it allows us to compare our results to vegetation maps from remote sensing (Tuanmu and Jetz, 2014). Note that patterns simulated by aDGVM2 are sensitive to the selection of the thresholds for classifying plants; however, slight modifications of these thresholds do not change our main findings and conclusions.

2.2.4 - MODEL BENCHMARKING

We compared simulated woody aboveground biomass to different remote sensing products reporting aboveground biomass (Saatchi *et al.*, 2011; Baccini *et al.*, 2012; Avitabile *et al.*, 2016). All biomass datasets were aggregated to the 1° resolution of our simulations by calculating the mean of all values within the 1° grid cell. We used the simulated mean woody biomass of the last 10 years for each grid cell. We used the “DGVMTools” R package (M. Forrest & J. Steinkamp, unpublished) to calculate statistical measures for agreement between datasets and model results (Normalized Mean Squared Error, NMSE, and coefficient of determination, R^2).

The simulated distribution of shrubs was compared to shrub prevalence from Tuanmu and Jetz (2014). The Tuanmu and Jetz (2014) dataset was aggregated to the 1° resolution of the simulations by calculating the mean of all values within the 1° grid cell. To quantify simulated shrub prevalence in each grid cell, we determined the fraction of woody individuals within the grid cell's woody plant population that are shrubs (as defined in Section 2.2.3). Furthermore, we used the average number of stems of all woody plants in a grid cell as proxy for the relative abundance of shrubs. Intermediate to high mean stem number indicates a high abundance of shrubs or mixture of shrubs and trees, whereas low mean stem number indicates a plant community dominated by trees.

The biomass and shrub cover data we use for benchmarking are similar to our model output, yet, they are structurally different from aDGVM2 output on two points. First, remote sensing data are reconstruction of data from satellite sensors based on complex models, expert knowledge, and ground truthing. Remote sensing products can deviate from reality (Friedl *et al.*, 2010) and from other products for the same variable. A perfect matching between observed variables is often not possible. Second, DGVMs simulate potential natural vegetation while remote sensing products include impacts of land use. This implies that precisely benchmarking DGVMs with remote sensing requires additional information on anthropogenic impacts.

2.3 - RESULTS

The inclusion of shrubs in aDGVM2, based on the trade-off between rapid height growth in trees and efficient water uptake in multi-stemmed shrubs, improves the simulated vegetation patterns in Africa. For the most realistic baseline scenario with shrubs, trees, and fire, simulated woody biomass patterns show better agreement with biomass estimates from remote sensing (Saatchi *et al.*, 2011; Baccini *et al.*, 2012; Avitabile *et al.*, 2016), than in scenario without shrubs (Figure 2.2, Table 2.2). Agreement between satellite products, ground studies, and aDGVM2 results varies strongly (Table 2.1). aDGVM2 is capable of broadly reproducing the distribution pattern of shrubs in African savannas and woodlands (Figure 2.2). Yet, the model appears to systematically underestimate shrub prevalence in the deserts and drylands of southern Africa (Kalahari, Karoo, Namib) and at the Horn of Africa. We can answer our first question: our approach allows us to reproduce shrub prevalence over large expanses of Africa with good agreement, in savannas and woodlands. However, model representation of shrubs in Mediterranean regions and in more arid regions is limited.

The average stem number of all woody plants simulated in a grid cell is lowest in equatorial African rain forests and increases in savannas and grasslands as precipitation decreases (Figure 2.3). The average woody plant height follows the inverse trend (Figure 2.4). The simulated distribution of shrubs reflects average water availability and the enhanced relative water uptake capacity of multi-stemmed plants in more arid areas. However, in water-stressed regions, such as the Karoo, aDGVM2 still simulates a shrub-tree mixture, while data indicate shrub prevalence only (Figure 2.2). The modelled shrub-tree communities in hyperarid regions consist of smaller individuals with a higher mean stem count (Figures 2.3 and 2.4).

Comparison of simulations with fire and without fire shows that fire generally reduces woody biomass (Figure 2.5) and woody cover (not shown) and modifies the distribution of shrubs and trees in Africa. Fire augments shrub dominance in more arid regions and increases mean stem number of woody plants (Figure 2.3) and shrub abundance (Figure 2.5). Increases in shrub abundance due to fire are most pronounced in the Sahel and the Miombo region. In contrast, compared to the scenario without fire, the model simulates a slight decrease in shrub prevalence in response to fire at the Horn of Africa and in the southern African arid regions. In summary, the results allow us to answer our second question: shrubs are favoured over trees under arid conditions, and fire promotes shrubs in savanna regions, while fire impacts are weak in high precipitation areas.

Tropical forest areas and savanna areas differ in the relative abundances of different phenotypes (Figure 2.6). In tropical forest areas, vegetation is dominated by a homogeneous layer of tall trees with a low number of stems and an understorey layer of smaller trees. In savanna areas, shrubs and trees of various sizes co-dominate, resulting in a functionally diverse woody layer. Fire plays a significant role in savannas by shifting the balance in favour of shrubs (Figure 2.6). In aDGVM2, fire does not occur, in forests, due to low grass biomass and dense canopy which prevent ignitions. Thus, fire does not influence vegetation structure in forests. Simulated patterns of architecture and community composition allow us to answer the third question: shrubs and trees do emerge as phenotypical strategies based on the fundamental trade-off between height growth and water uptake.

Table 2.1: Aboveground biomass for different savanna sites, from field observations, remote sensing products (Avitabile et al., 2016; Baccini et al., 2012; Saatchi et al., 2011), and aDGVM2 simulations. Sites 1-4: Higgins et al. (2007), site 5: Rutherford (1979), site 6: Rutherford (1975), site 7: Kelly and Walker (1976), site 8: Rushworth (1975).

Site ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Site longitude (°E)	**31.4**	**31.8**	**31.6**	**31.3**	**28.7**	**17.1**	**31.6**	**26.3**
Site latitude (°S)	**23.5**	**24.4**	**25**	**25.2**	**24.7**	**20.4**	**27.2**	**18.9**
Field observation								
Minimum (t/ha)	3	3	7	13	—	—	8.7	4.7
Mean (t/ha)	—	—	—	—	16.3	22.3	16.7	—
Maximum (t/ha)	5	12	11	23	—	—	30.8	5.8
aDGVM2 results								
Minimum (t/ha)	23	29.8	36.3	31.5	2.2	11.1	4.7	15
Mean (t/ha)	44.7	49.5	58.5	79.2	50	28.8	44.2	40.5
Maximum (t/ha)	79.4	75.9	82.3	120.3	91.8	58	100.3	71
Satellite data								
Saatchi et al. (2011) (t/ha)	17.7	21.3	26.6	49.9	16.9	12	32.8	22.8
Avitabile et al. (2016) (t/ha)	3	4.1	7.4	9.1	2.4	1.5	9.5	6.5
Baccini et al. (2012) (t/ha)	23.1	23.1	23.1	—	10.1	4.6	—	23

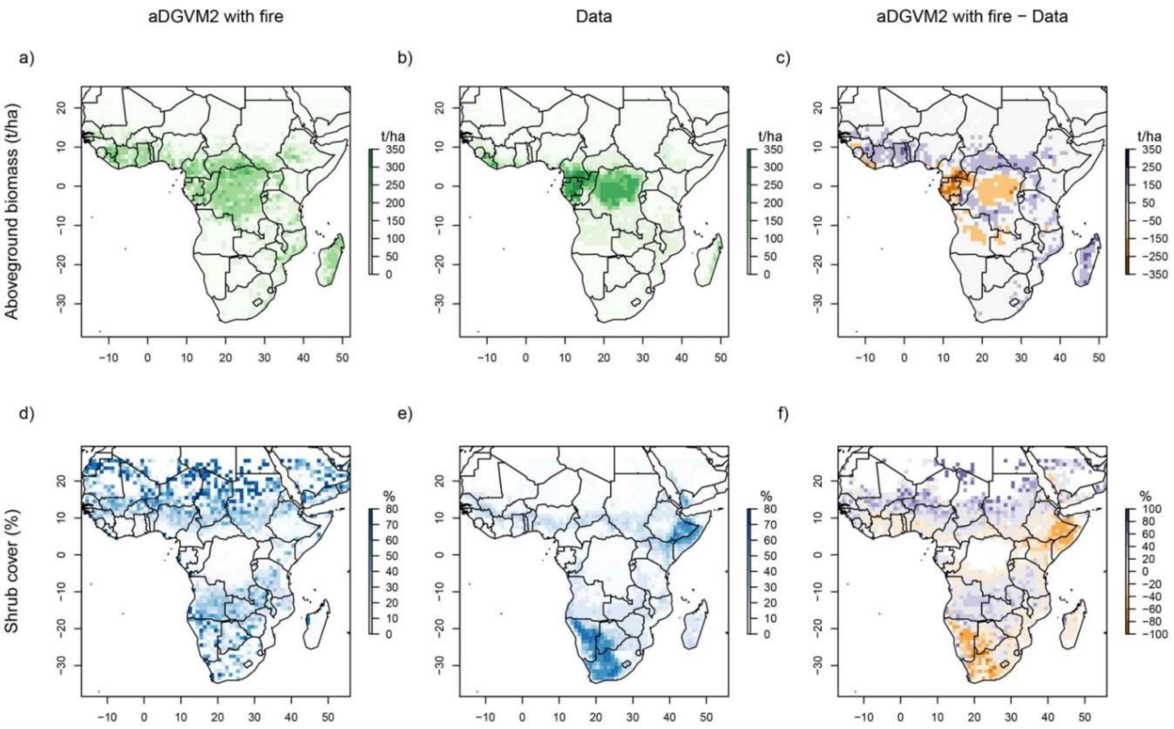


Figure 2.2: Comparison between aDGVM2 results and satellite products for biomass and shrub prevalence. Aboveground biomass data were derived from Saatchi et al. (2011) (b) shrub prevalence data were derived from Tuanmu and Jetz (2014) (e). Aboveground biomass is in tons per hectare. Shrub prevalence is in percentage of total woody plant number for aDGVM2 and in percentage of land cover for Tuanmu and Jetz (2014). Simulations were conducted in the presence of fire and shrubs, which represents the most realistic scenario. Biomass comparisons with additional satellite products are available in Supporting Information Appendix S1 and Supporting Information Figure 2.S1.

Table 2.2: Comparison of aboveground biomass between aDGVM2 results for different simulation scenarios and three remote sensing products.

Shrubs included **Fire included**	**Yes** **Yes**	**Yes** **No**	**No** **Yes**	**No** **No**
Normalized mean squared error (NMSE)				
Saatchi et al. (2011)	0.768	0.877	1.047	0.912
Baccini et al. (2012)	0.630	0.670	0.811	0.687
Avitabile et al. (2016)	0.907	1.039	1.143	1.066
Coefficient of determination (R^2)				
Saatchi et al. (2011)	0.394	0.366	0.340	0.370
Baccini et al. (2012)	0.426	0.423	0.379	0.428
Avitabile et al. (2016)	0.239	0.185	0.180	0.186

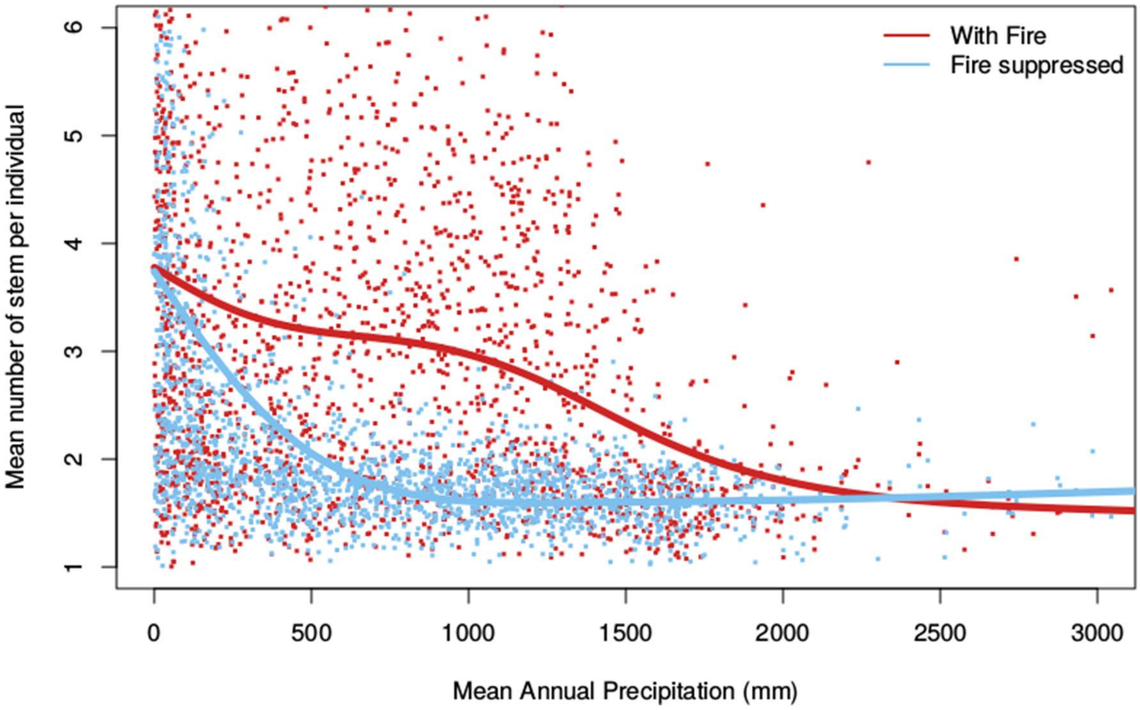


Figure 2.3: Response of mean stem number to mean annual precipitation. Mean stem number indicates the relative abundance of shrubs and trees. High stem numbers indicate dominance of shrubs while low stem numbers indicate dominance of trees. Intermediate stem numbers indicate dominance of shrubs with intermediate stem numbers or a mixture of trees and shrubs. Maximum mean stem number displayed has been constrained to six but simulations were performed with a maximum of 10 stems per plant. Each point in the graph represents the mean value of one grid cell simulated by the aDGVM2. The lines represent fitted splines.

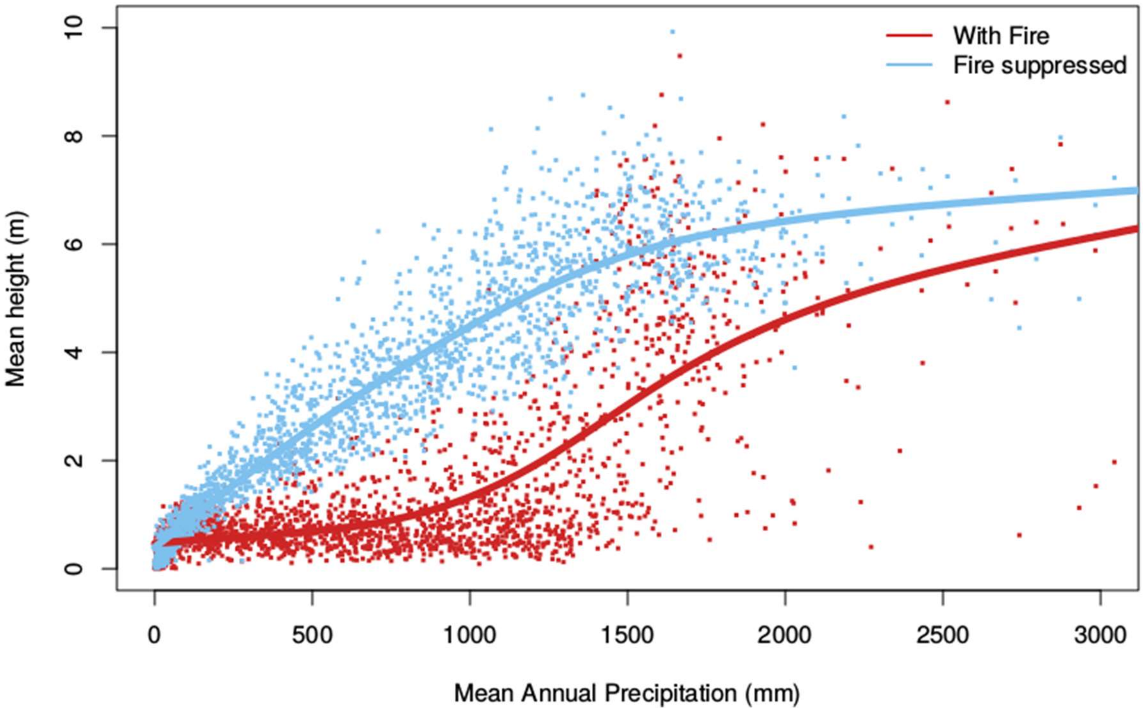


Figure 2.4: Response of mean height of woody individuals to mean annual precipitation. Higher mean height indicates dominance of trees while lower mean height indicates dominance of shrubs. Each point in the graph represents the mean value of one grid cell simulated by the aDGVM2. The lines represent fitted splines.

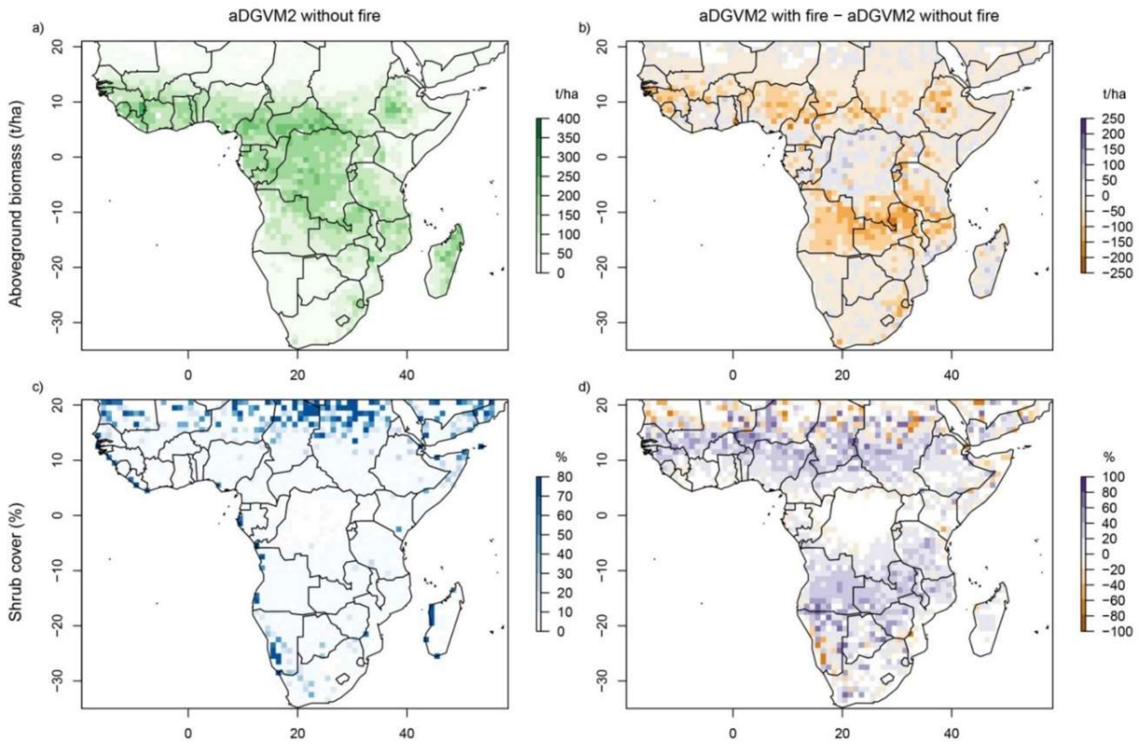


Figure 2.5: Impacts of fire suppression on aboveground tree biomass (a and b) and shrub prevalence (c and d). (a) and (c) Biomass and shrub prevalence in the absence of fire. (b) and (d) The difference between simulations with fire (Figure 2.2) and simulations without fire.

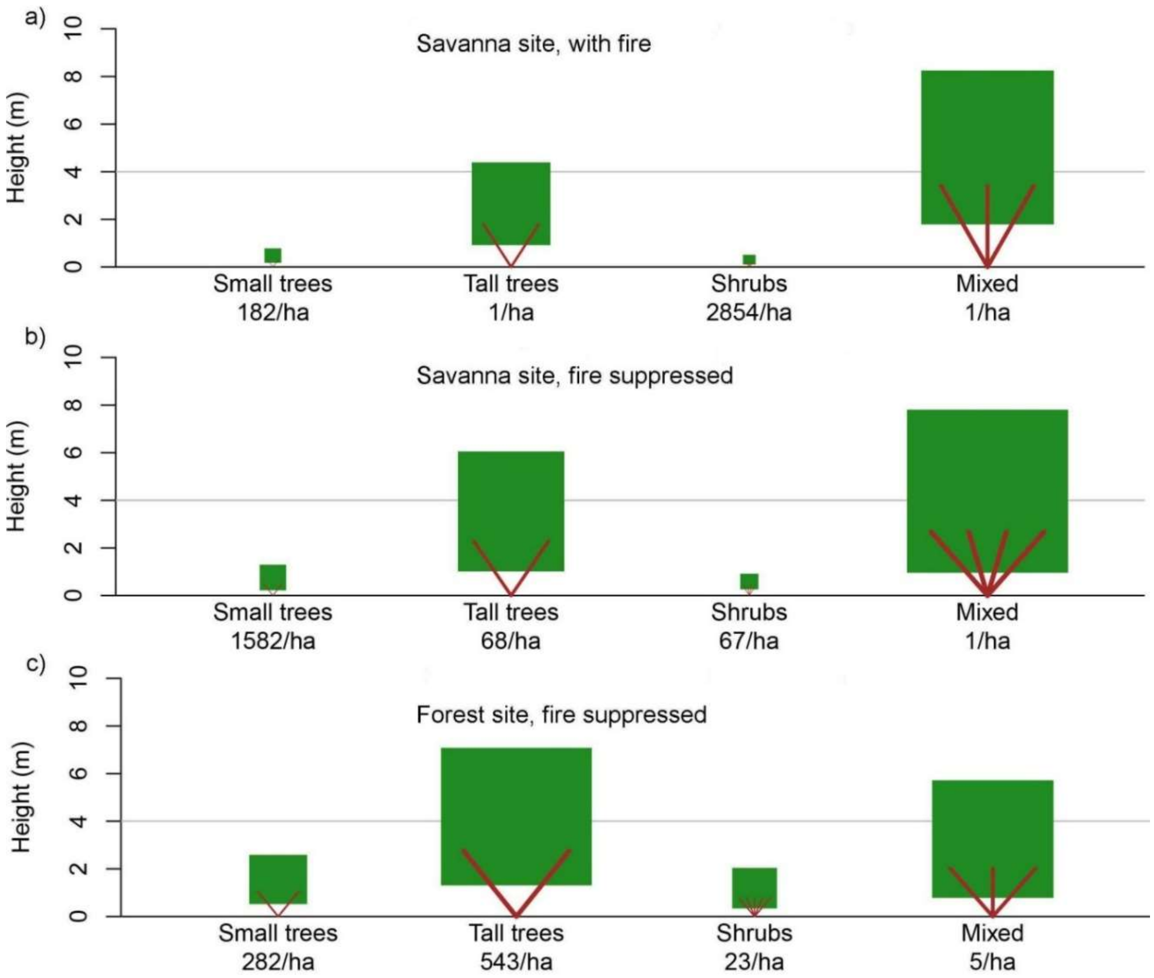


Figure 2.6: Average architecture of woody plants, at a savanna site in the presence of fire (a) and in the absence of fire (b), and at a forest site in the absence of fire (c). At the forest site, aDGVM2 does not simulate fire due to fuel and ignition limitations. For each scenario, we classified woody plants into small and tall trees, shrubs, and mixed types and we calculated mean height, mean canopy radius, mean stem number, and mean stem radius for each class. Small trees have up to three stems and are less than 4 m tall, tall trees have up to three stems and are more than 4 m tall (grey line), shrubs have four or more stems and are less than 4 m tall, remaining plants are classified as mixed types.

2.4 - DISCUSSION

Shrubs are an important vegetation growth form in many terrestrial ecosystems. Shrubs can either be the major vegetation type of an ecosystem or coexist with other vegetation types and contribute to the functional diversity of an ecosystem (Olson *et al.*, 2001). In the ecological literature, shrubs recently received more attention (Eldridge *et al.*, 2011; Zizka, Govender and Higgins, 2014; Götmark, Götmark and Jensen, 2016; Stevens, Lehmann, *et al.*, 2016). Yet, in contrast to trees, shrubs have received little attention in DGVMs. In many DGVMs, shrubs are not explicitly considered, or are included as PFTs, without explicitly accounting for their multi-stemmed architecture and the implications for growth and water uptake (Zeng, Zeng and Barlage, 2008; Clark *et al.*, 2011).

In this study, we present a mechanistic representation of shrubs and trees in the aDGVM2 by introducing stem number as a plant trait. Variable stem number allows us to simulate single- or multi-stemmed architecture of these growth forms and the associated trade-off between height growth in trees versus enhanced water availability in multi-stemmed shrubs. We show that this trade-off allows us to simulate large expanses of the current distribution of shrubs in Africa, particularly in savannas. The finding emphasizes the need to consider shrubs to understand biogeographical patterns in Africa.

2.4.1 - ADAPTATION OF SHRUBS TO WATER AND LIGHT COMPETITION

Plant growth is strongly constrained by the availability of water and light. Individual plants compete for those resources using different strategies (Craine and Dybzinski, 2013). As competition is usually driven by the most limiting resource, successful strategies vary depending on the nature of the limiting resource. Adaptation to efficient acquisition of a specific resource requires specialization. Simultaneous optimization of acquisition of all resources while developing avoidance or resistance strategies to stress factors is impossible(Stearns, 1989).

Our findings indicate that shrubs are more abundant in more water-stressed regions and that under such conditions they have a competitive advantage over trees. For a given stem biomass, a multi-stemmed plant has a higher sapwood area than a single-stemmed plant in aDGVM2 and a higher capacity to extract and transport water. In our results, this relation is reflected by the increase in average stem number and the shrub-tree ratio along gradients of decreasing mean annual precipitation, both in the presence and absence of fire. However, in aDGVM2, in regions of high water stress, multi-stemmed architecture alone is not sufficient to explain observed shrub prevalence. A multilayered soil water model and rooting depth plasticity are included in aDGVM2 (Langan, Higgins and Scheiter, 2017), allowing plants to adjust their rooting depth in response to soil water availability. Deeper rooting systems, allowing plants to reach aquifers can be an important factor in these regions (Schenk and Jackson, 2005) and are missing in aDGVM2. A stem hydraulic conductivity versus hydraulic safety trade-off is mediated by wood density in aDGVM2; however, the model does not account for the full complexity of this trade-off (Jacobsen *et al.*, 2008). Another mechanism missing in aDGVM2 to account for competition between shrubs and trees is the resistance of shrubs to aridity via redundancy of their architecture against mechanical failure (Schenk *et al.*, 2008).

Where water supply is not limiting, aDGVM2 simulates tree dominance at the expense of shrubs. Vegetation biomass and tree cover are particularly high in equatorial African forests, and are mainly driven by light competition and selection for tall individuals (Scheffer *et al.*, 2015).

2.4.2 - ADAPTATION OF SHRUBS TO FIRE

Fire is an important component in most ecosystems with substantial shrub prevalence, such as Fynbos or savanna. In savannas, fire generally reduces woody biomass and alters vegetation structure (Smit *et al.*, 2010). However, woody savanna plants are rarely killed by fire because they are adapted, by either resistance or a high ability to resprout after fire damage (Higgins, Bond and Trollope, 2000). We observed a similar behaviour in our results.

In our simulations, turning fire on reduces total woody cover and biomass in semiarid regions while increasing shrub dominance both with respect to abundance and mean stem number, compared to simulations without fire. We identified two potential mechanisms to explain how fire shifts the competitive balance between shrubs and trees. First, fire reduces recruitment success of trees and then tree prevalence. Thus, fire indirectly increases shrub growth rates by reducing tree cover and light competition by trees. Second, fire reduces mean vegetation height. This implies a competitive advantage of shrubs, which are smaller than trees and benefit more from open vegetation. It could also indicate an increased competitiveness of resprouters against reseeders, assuming that resprouters tend to be smaller than reseeders (Midgley, 1996). However, we did not assess trade-offs between resprouters and reseeders in aDGVM2. Thus, the positive response of shrub prevalence to fire at the expense of tree prevalence can be explained by mechanisms inherent to both functional types to cope with fire.

Woody plants can deal with fire either by regeneration (resprouting or reseeding) or by resisting fire (by growing tall to escape the flame zone or building enough bark for protection). Due to the built-in constraints of their respective architecture, in aDGVM2, mature trees and shrubs behave differently. Trees can survive fire by growing tall to escape the flame zone, by investing into a thick bark, or by investing in belowground reserves to increase their resprouting ability (Higgins, Bond and Trollope, 2000; Hoffmann and Solbrig, 2003; Hoffmann *et al.*, 2012). Shrubs, being shorter due to their multi-stemmed architecture, usually do not escape the flame zone. However, shrubs have higher growth rates at low size in aDGVM2, thus, they excel in post-fire resprouting (Zizka, Govender and Higgins, 2014), irrespective of their storage capacity (Shibata *et al.*, 2016).

2.4.3 - SHRUB AND TREE PHENOTYPES AND IMPLICATIONS FOR DIVERSITY

A key feature of aDGVM2 is that, by the processes of trait inheritance, mutation, and cross over, it systematically samples the trait space spanned by the simulated traits. Thereby, the model assembles communities composed of growth strategies that are suitable for the given biotic and abiotic conditions (Scheiter, Langan and Higgins, 2013). The model allows different strategies to emerge and coexist without an explicit and a priori parameterisation. By adding stem number as an additional plant trait into the model, we can simulate communities consisting of a mixture of distinct shrub and tree phenotypes. The relative abundance of these growth forms emerges as a function of light and water competition and in response to fire. Coexistence of shrub and tree phenotypes is possible in the model and it increases functional diversity both in terms of a more diverse vegetation structure and in terms of emergent phonological strategies.

Reduction in complexity is necessary in all models. Yet, resilience and resistance of ecosystems have been shown to be influenced by diversity (Isbell *et al.*, 2015). Oversimplification in vegetation models, for instance by a coarse and static representation of functional types, may cause misleading conclusions when assessing ecosystem resilience and future trajectories. Being able to simulate the coexistence of different growth forms, diversity, and adaptation of plant communities to the environment is therefore prerequisite when simulating climate change impacts on vegetation (Langan, 2019). Introducing a mechanistic representation of shrubs in aDGVM2 is therefore an important step towards a representation of functional diversity of ecosystems.

2.4.4 - FRAMING OUR RESULT'S ACCURACY

To interpret the results of our model, we need to understand its limitations and explain the discrepancies with benchmarking data. First, as explained in the methods, our model results are conceptually different from the benchmarking data. Due to this difference, discrepancies do not necessarily represent a shortcoming of our modelling approach but of the benchmarking process. Despite a disagreement in shrub prevalence between aDGVM2 results and benchmarking data in the south-west and north-east of Africa, we still observe an increase in the mean stem count per individual when aridity increases (Figure 2.3), along with a decrease in woody vegetation height (Figure 2.4). These responses indicate shrub dominance in these areas.

Second, remote sensing data potentially underestimate shrub prevalence since shrubs are typically smaller than trees and can be hidden under the canopy of taller individuals (Linderman *et al.*, 2004) or be too small to be detected by satellites. In our analysis, we consider every shrub individual, resulting in an overestimation of shrub prevalence across African savannas and woodlands by aDGVM2. Yet, this difference does not invalidate our results. Quantification of the magnitude of this effect requires in-depth studies at site level.

A third aspect is the large variability between the satellite products, and the data from site studies, used to benchmark biomass. Disagreement varies strongly between sites and can be larger than disagreement between data and our model results (Table 2.1). Variability can be explained by different spatial resolutions of data and models, different methodologies to measure, identify and classify vegetation cover, and by land use, which was not considered in this study.

The fourth aspect is the limited accuracy of our model in areas of extremely low biomass and vegetation cover, such as deserts and semideserts. In those areas, the absolute value of the discrepancies between model and data is similar as in areas with higher vegetation cover. Yet, the relative value of the deviation is much higher given the relatively low vegetation cover. This also means that the lower the vegetation cover, the more stochasticity of the model appears to be important. Stochasticity may imply larger differences between multiple simulation runs than between model simulations and data. The discrepancies between our model results and observations in south-western Africa or in the Horn of Africa suggest high uncertainty in extremely arid regions and in mediterranean biomes.

2.4.5 - FUTURE DEVELOPMENTS

Modelling plant strategies and ecosystem behaviour in a mechanistic way based on traits is an open research field (Fisher *et al.*, 2017). We propose a pathway towards a better representation of shrubs in DGVMs, which, however, leaves several question unresolved. First, there are shortcomings in our ability to discriminate trees from shrubs in our model. Single-stemmed and multi-stemmed woody individuals are more similar at very low height, due to inbuilt mathematical relations (Figure 2.1). This reduces our model's ability to simulate shrubs and trees as clearly distinct life-forms in areas where plant biomass is very low, such as in deserts or semi-deserts.

Second, our approach assumes that being multi-stemmed is a drought tolerance strategy rather than a drought avoidance strategy. Following Carvajal *et al.* (2017), in extremely arid conditions, drought tolerance (the ability to function in water-stressed condition) is more important than drought avoidance (the ability to avoid being water stressed). It appears that our representation of shrubs in extremely arid conditions is limited, hinting towards the need to consider additional mechanisms, absent from our representation of drought tolerance. Being multi-stemmed can be a drought tolerance strategy through redundancy when cavitation damages or kills single stems instead of the entire plant (Schenk *et al.*, 2008). This mechanism is related to our multi-stem model but not implemented. To explain plant communities in extremely arid areas, there are specific drought and aridity tolerance mechanisms to implement in aDGVM2, such as CAM photosynthesis, succulence, reduced stomata number, leaf wax, and leaf hairs as well as overall cavitation and embolism resistance or avoidance (Mucina and Rutherford, 2006). It seems that no single mechanism can explain plant biogeography in all arid regions (Esler and Rundel, 1999). Other region-specific mechanisms can be added into aDGVM2, such as taking into account areas of rocky surface, for example, the Karoo-Namib quartz fields and the associated adaptation mechanisms (Schmiedel and Jürgens, 1999), or including crown fire and representation of landscape heterogeneity in the Fynbos.

2.5 - CONCLUSIONS

We demonstrate that, in aDGVM2, shrub and tree growth forms dynamically emerge based on functional differences in the architectural design, without a need to predefine static functional types. To our knowledge, this is the first time that a multi-stemmed growth strategy has been simulated as an emergent property in a DGVM. The distribution of shrubs in African savannas can be modelled by relying on mechanistic principles generating trade-offs between single-stemmed and multi-stemmed plants. However, further developments are required to fully grasp the complex interactions between trees, shrubs, and grasses, specifically outside of African savannas. Shrub encroachment is assumed to be linked to increasing atmospheric CO_2 concentrations, climate change, fire, or land use changes (Midgley and Bond, 2015). How these factors interact with natural fire regimes, light, and water competition remains open. Our results highlight that further developments in DGVMs should not only focus on well-known trade-offs, such as the leaf economic spectrum (Wright *et al.*, 2004), but that they should also consider whole-plant trade-offs integrating resource availability, life history, carbon allocation, or plant architecture. Our approach is an important step towards understanding the biogeography of ecosystems where shrubs play a major role, specifically savannas and grasslands and thus towards our ability to investigate their origin and future.

ACKNOWLEDGEMENTS

C.G., M.P., and C.M. thank the BMBF SPACES initiative for funding (“Limpopo Living Landscapes” (LLL) project, grant 01LL1304B and “ARS AfricaE” project, grant 01LL1303C). S.S., L.L., D.K., and M.P. thank the DFG Emmy Noether programme for funding (grant SCHE 1719/2-1). The authors thank the SBiK-F modelling centre research group and the LLL research consortium for valuable insights.

DATA ACCESSIBILITY

Selected data are available at http://data.sasscal.org/metadata/view.php?view=li_rbis_process_step&id=6385&ident=903937203343957632. To obtain the model code please contact any of the authors.

APPENDIX S1

METHODS

THE ADGVM2

General features of aDGVM2.

aDGVM2 is an individual-based dynamic vegetation model. It simulates growth, reproduction and mortality of individual plants while keeping track of state variables, such as biomass, height and leaf area. Each plant in aDGVM2 is characterised by a potentially unique set of trait values. These traits describe plant type, leaf characteristics, leaf phenology, hydraulic characteristics, plant architecture, response to fire, reproduction and mortality. Most of these plant traits are linked by direct or indirect trade-offs. For example, specific leaf area and leaf longevity are linked to the risk of cavitation (see Scheiter, Langan and Higgins, 2013; Langan, Higgins and Scheiter, 2017; Kumar *et al.*, 2020, for full model description).

Selection and trait inheritance, in the model, assemble plant communities that are well adapted to given biotic and abiotic conditions. Plants with trait combinations that allow sufficient growth and reproduction rates can produce seeds and contribute their trait values to the community trait pool. Seeds can randomly mutate trait values or recombine within the community trait pool. Seeds are randomly drawn from the community trait pool and added to the plant population as seedlings. Plants with insufficient performance fail to contribute seeds to the seed bank and disappear from the population. Successful life history strategies emerge dynamically from this community assembly process. Plant growth is constrained by competition. Light competition is simulated by considering the impacts of neighboring plants on the light available to a target plant. Water competition is simulated via water uptake of plants from a layered common soil water pool.

Improved ecophysiology.

The model version used in this study incorporates improvements to the ecophysiology sub-models (Kumar et al., unpublished). Following Sakschewski *et al.* (2015), aDGVM2 now includes an empirical relation between specific leaf area (A_{SL}), leaf nitrogen content per unit area (N_a) and the maximum carboxylation rate per leaf area at 25°(V_{cmax25}), derived from the TRY data base (Kattge *et al.*, 2011):

$$N_a = 6.89 \cdot A_{SL}{}^{-0.571} \tag{1}$$

and

$$V_{cmax25} = 31.62 \cdot N_a{}^{0.501} \tag{2}$$

In aDGVM2, A_{SL} is calculated from P_{50}, the matric potential at 50% loss of hydraulic conductivity (see Langan, Higgins and Scheiter, 2017). In previous model versions (Scheiter, Langan and Higgins, 2013; Langan, Higgins and Scheiter, 2017), V_{cmax} and associated leaf level photosynthetic rates were equal for all plants, now, these quantities are calculated separately for each plant. We simulate the effect of temperature on Vcmax by using leaf temperature T_{leaf} while previous model versions used air temperature:

$$V_{cmax} = \frac{V_{cmax25} \cdot 2^{0.1\,(T_{leaf} - 2\ \ )}}{\left[1 + e0.3\left(T_{low} - T_{leaf}\right)\right] . \left[1 + e^{0.3\,(T_{leaf} - T_{upp})}\right]} \tag{3}$$

Here, T_{upp} and T_{low} describe the optimum temperature range of photosynthetic activity as provided in the photosynthetic model of Collatz *et al.* (1991) and Collatz, Ribas-Carbo and Berry (1992) for C_3 and C_4 plants, respectively. Leaf temperature is calculated from air temperature, wind speed, stomatal conductance, boundary layer conductance and evapotranspiration following Gates (1968).

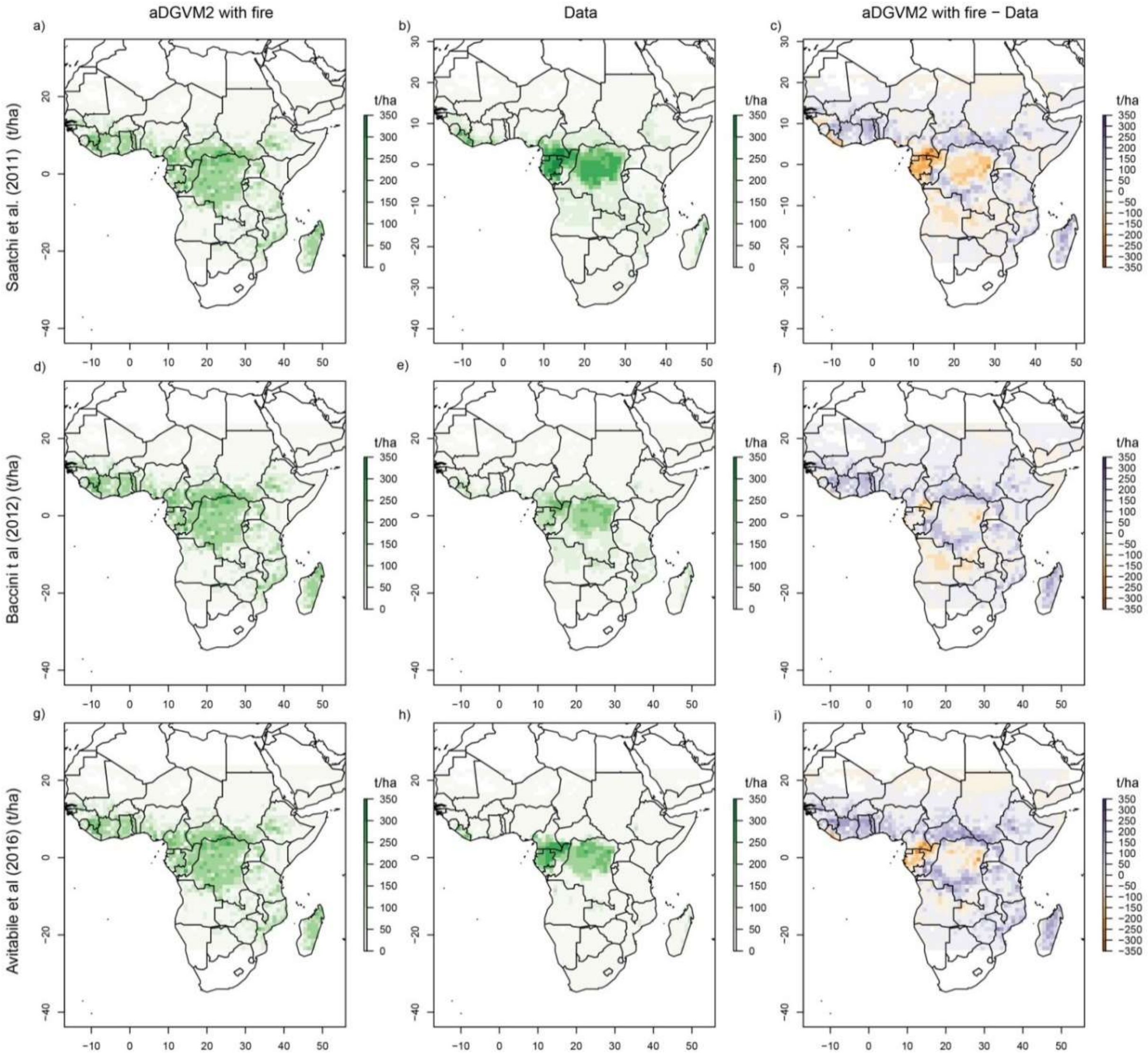


Figure 2.S1: Comparison between simulated and observed biomass. Aboveground biomass data were derived from Saatchi et al. (2011), from Baccini et al. (2012) and from Avitabile et al. (2016). Aboveground biomass is in tons per hectare. Simulations were conducted in the presence of fire and shrubs, which represents the most realistic scenario.

CHAPTER 3 - SHRUB FORM & FUNCTION IN ECOSYSTEMS STRUCTURE & DYNAMICS, SIMULATED WITH ADGVM2 IN AFRICAN SAVANNAS

This chapter is the work of the sole author of this thesis.

ABSTRACT

Background

In chapter 2 we described a new shrub model we developed and implemented in aDGVM2. We build upon it to test its impact on vegetation community assembly processes simulated by aDGVM2 and we test to which extent it can be used to analyse shrub life-form and life-strategy.

Objectives

We address questions arising from the shrub model we presented in chapter2. On one hand we delve into more details about the operation and understanding of the model behaviour. On the other hand we address more contextual questions, giving a broader perspective to our shrub model.

More specifically, we propose to answer the following questions: (1) Can we expand or limit the validity of our shrub model? How are these limits informative? How does it impact the representation of ecosystems by aDGVM2? (2) How introducing shrubs in aDGVM2 modifies its representation of natural vegetation? (3) Is there a broad meta-distribution pattern of vegetation across Africa? (4) How our definitions of life-form categories shape our understanding of these life-forms? How can we ensure that our categorisation actually allows to produce meaningful analyses? (5) How is shrub life-form expressed into a strategy? How do shrubs behave in terms of resprouters *vs.* reseeders, in terms of rooting patterns and in terms of carbon allocation?

These questions allow us to answer the following: How can we ensure that our analyses and model are informative and to which extent? What does our representation of shrubs implies by itself?

Methods

We perform further analyses on the simulations obtained in chapter 2. We add comparison to a new satellite data product previously unavailable. We performed 10 replicate simulations with and without fire for the greater Namibian regions.

Results

We show that shrubs distribution pattern arising from our shrub model follows the overall precipitation patterns, and that vegetation community assemblages are accordingly simulated along MAP gradient. We show that how we quantify and how we categorise our study subject impacts our ability to understand them and the conclusion we can make. Therefore we emphasize the need for more comprehensive and holistic studies and that narrow studies should be considered more carefully. We show that our shrub model allows for complex plant life-form strategies to emerge, without being pre-defined, and interact with other life-forms in a unique way, giving us the ability to simulate plants and ecosystems behaviours in unprecedented manners. For example, shrubs emerging from our model are more associated to reseeders than resprouters and exhibit rooting patterns matching *in situ* observations.

Conclusions

Categorisation of our results shapes our ability to assess them. Vegetation life-forms diversity is a continuum and categories artificially constructed and traditionally utilized are only partially relevant for assessing trait-based processed models such as aDGVM2. By going beyond the classical botanical categories and by looking at raw strategies, processes and traits, we can model shrub distribution through African savanna with aDGVM2 and gain direct insights into plant community assembly and ecosystem structure and processes.

3.1 - INTRODUCTION

3.1.1 - CHALLENGING THE MULTI-STEM SHRUB MODEL

Shrubs, as characterised in chapter 2 are a major feature of African savannas, shaping ecosystems states and dynamics (Vetaas, 1992; Le Roux, Bariac and Mariotti, 1995; Zou *et al.*, 2005). A precise understanding of their behaviour is necessary to understand how they react to biotic and abiotic drivers and thus how they can shape the future of ecosystems (Gaze *et al.*, 1998; Hagenah *et al.*, 2009; Kambatuku, Cramer and Ward, 2011; Schleicher, Wiegand and Ward, 2011). Why, when and how are they successful and how their specific life-form and strategy explain their success or failure are key questions discussed in our previous work which we propose to test. We demonstrated that shrubs emerge in aDGVM2 based on a trait-based trade-off. Here, we ask how this individual based trade-off cascades at the ecosystem level to shape vegetation community and how it impacts aDGVM2 results. We question the role of shrubs in ecosystems and how shrub life-form and strategy are embedded in vegetation communities.

In chapter 2 we presented the implementation of our shrub model in aDGVM2 and showed that it allows for shrubs to emerge dynamically from the implemented trade-off between height growth and sapwood conductivity. We benchmarked AGBM and vegetation type cover with satellite derived data and with site specific data. We found that our approach reproduces the predicted pattern of increased mean stem number along a decreasing precipitation gradient. However, our results accuracy was limited outside of savannas; e.g.: shrub cover in the Fynbos and Karoo is likely driven by mechanisms non-explicitly represented in aDGVM2 at present (such as succulence or crown fire). High levels of aridity are also challenging our vegetation simulations. This limitation drives the need to expand our analysis about our approach's validity. To do so, we consider our shrub model from three different perspectives: 1) We investigate how shrubs influence vegetation distribution and the role they play in vegetation community assembly; 2) We reconsider, refine and expand our benchmarking methodology; 3) We focus on the regions were discrepancies between our results and satellite-based products are the largest. This allows us to answer our first question:

1. How do we ensure the relevance and accuracy of our shrub model and in which cases and under which assumptions can we use it?

After determining how shrubs can be characterised specifically in the vegetation community pattern according to our modelling approach, we can investigate the role they play in an ecosystem. We test the impact of introducing shrubs in our model. In chapter 2 we tested how introducing shrubs modifies the overall performances of aDGVM2 at representing large scale ecological parameters. We now analyse how it impacts vegetation community assembly at a more detailed level. Accordingly, our second question is:

2. How do shrubs, as we model them, shape ecosystems simulated by aDGVM2?

3.1.2 - SHRUB DISTRIBUTION DEPENDS ON SHRUB CLASSIFICATION

Our understanding and perception of plant life-form impacts our ability to model them and to produce meaningful analysis. This conundrum is particularly concerning for shrubs as they are a plant type category whose definition varies and can lack precision. This fuzziness makes studies complicated due to the lack of consensus on the terminology. Our ability to draw theories and hypothesis from facts is particularly challenged when these facts are divided between various characterisation schemes. We question: Can we devise generalized rules from equivocal items?

As a botanical category, shrub plant species are classified as such according to their archetypal individual. Considered at the ecosystem, landscape or biome scale, vegetation can be classified following two different perspectives, with potentially diverging meanings. On one hand individuals can be characterised depending on their species, without considering their actual phenotype (Estabrook, Johnson and Mc Morris, 1975; Bachmann, 1995). On the other hand, individuals can be classified depending on their actual phenotype, notwithstanding of their species (Lavorel and Garnier, 2002; Kattge *et al.*, 2011). The separation between these two approaches entails a potential disconnection between actual phenotype and individual species (as a plant does not necessarily display a phenotype representative of its specie due to environmental drivers). This issue can easily be solved at site level in studies combining both approaches. However at larger scale, considering individual phenotype and species can be more challenging. Satellite derived products usually do not provide direct information about plant species. The aDGVM2 simulates life-forms adapted to ambient conditions. However, as it does not include phenotypic plasticity, the distinction between a species adapted to ambient conditions from a phenotype adapted to ambient conditions is limited. Thus, as in the aDGVM2 framework it is not possible to compare individual plant traits to plant species genotype, we need to compare actual plant phenotypes. Therefore, the way we categorise and aggregate these phenotypes is a cornerstone of our benchmarking process.

To resolve this classification conundrum, we propose to investigate how we can use vegetation distribution patterns to assess the validity of our methodology. Then we estimate the extent to which categorisation decision impacts our benchmarking results and our potential conclusions and thus, in a sense, the validity of our approach to model shrubs. Finally, we explore potential options to circumvent our present classification issue by taking an indirect approach, analysing how focusing on non-shrubs life-forms (grasses and trees) can inform us about shrubs.

Shrub distribution in non-geographic space

We demonstrate in chapter 2 that our shrub model allows aDGVM2 to simulate the spatial distribution of shrubs in African savannas. However, we observed disagreement with satellite derived products outside the savanna, woodland and forest regions. Consequently, we need to determine if this disagreement is linked to our shrub model itself or to other factors not represented in aDGVM2, such as regional specific environmental conditions and associated plant mechanisms. Is our shrub model overridden locally by other key factors or are our assumptions made to simulate shrubs lacking? To answer this question, we look at vegetation distribution in virtual space, which means non-geographically, for example in trait space or in climate-space. Doing so, we answer our third question:

3. How do distribution pattern of shrubs across Africa responds to environmental variables?

Classification

Shrub distribution has been studied regionally (Cole, 1982; Archer *et al.*, 1988; Cipriotti *et al.*, 2012) and globally (Scepan, 1999; Friedl *et al.*, 2002; Mayaux *et al.*, 2006; Bicheron P. *et al.*, 2008; Tuanmu and Jetz, 2014). However, divergences can be found among these studies. We argue that it might be related to divergences in shrub characterisation. We investigate how alternative classification schemes can impact aDGVM2 simulations and our conclusions, which might challenge our ability to produce meaningful analysis. We provide elements to answer our fourth question:

4. How do categorisation schemes of life-form shape our understanding of these life-forms and how can we ensure the relevance of a classification scheme?

Footprint of shrubs

Life-form classification schemes can differ among authors and this might directly shape the understanding of these life-forms. This raises two questions. First, it is challenging to transform uncertain categories into precise operational concepts which can be implemented in a model. Second, when it is possible to create a model of a life-form or life-strategy, our ability to assess it is impaired by the fact that data, which would be necessary to benchmark model results, can be based on different classification schemes or even non-existent. We argue that to be able to assess a model relevance and accuracy, when benchmarking data is not available or based on assumptions diverging from the model, it might be necessary to take an indirect benchmarking approach by looking, not at the studied feature, but at the impact of the feature on its environment. This approach is similar to common particle science methodology (Blum, Walter; Riefler, Werner; Rolandi, 2008). We make a first step in this direction by assessing how introducing shrubs impacts aDGVM2 simulation of vegetation communities. We make a second step by studying the abiotic environment of shrubs. Precisely, we investigate how the share of bare-ground cover reacts to the presence or absence of shrubs in aDGVM2.

3.1.3 - SHRUB LIFE-FORM *MODUS VIVENDI*

We classify shrubs as a specific life-form by characterising their phenotype. We further classify shrubs as a specific life-strategy by characterising their response to environmental parameters. We analyse shrubs interaction with other life-forms and the mechanisms on which this strategy relies. This allows us to answer our last question:

5. Can we characterise shrubs as a specific life-strategy similarly to how we characterise them as a specific life-form? Specifically, we ask: how do shrubs behave in terms of resprouters *vs.* reseeders, in terms of rooting patterns and in terms of carbon allocation?

Resprouting *vs.* reseeding

We hypothesize that, according to our assumptions on which our model relies, shrubs, as simulated by aDGVM2, should be resprouters more often than reseeders. As, in aDGVM2, shrubs are less likely than trees to escape the flame zone (due to their typically limited height growth), they are less likely to have a chance to seed. Thus, we should observe a division between trees and shrubs on this point. This relation between life-strategy and life-form has been proposed in the literature. For example, Midgley (1996) states that resprouters tend to be multi-stemmed whereas tall trees are single-stemmed reseeders and also that resprouters occur mainly in savannas and grassland (with some exceptions) and are usually smaller than reseeders. The possibility for individuals to be resprouter or reseeder is a dynamic trade-off that is natively included in aDGVM2 via different combinations of carbon allocation traits. However, it has not been evaluated yet. Therefore, this is the question we investigate in this section in order to gain insights on the behaviour of our shrub model and test if it represents only their distribution or also their strategy.

Rooting depth strategies

The reseeding or resprouting strategies are partially explained by their root systems. However, types of root systems can also be seen as strategies *per se* (e.g.: deep or shallow rooting) and concern both woody and non-woody plants. Therefore, we need to investigate if different types of root systems emerge from aDGVM2 as stand-alone or only in correlation to reseeding or resprouting strategies. Schenk and Jackson (2002) observed that herbaceaous plants have larger root systems in dry climate. This is supported and extended to many species by Phillips (1963), Pallardy (1981), Wilson (1988), Chapin, Autumn and Pugnaire (1993). The increase of root to shoot ratios with increasing aridity is documented for an array of species and regions. We analyse reseeders and reprouters strategies in terms of shoot to root ratios and rooting depths.

Carbon strategies: shrubs *vs.* trees

Resprouting and reseeding strategies diverge on how they allocate carbon and how they relate to plant phenotype. Similarly, rooting depth strategies are related to shoot to root ratio and thus to carbon allocation strategies. Before being allocated, carbon need to be accumulated, thus, carbon allocation strategies are intertwined with carbon accumulation strategies, as carbon allocation determines how carbon can be accumulated. The feedback between these two mechanisms forms the complete strategy of an individual.

Therefore, we need to investigate plant carbon accumulation strategies to achieve a comprehensive understanding of carbon strategies and shrubs and trees performances. Indeed, according to the trait-based trade-off we used as a foundation to our shrub model, when comparing two individuals, a shrub should accumulate carbon faster at an early stage whereas a tree should outcompete the shrub when reaching a later stage (Wilson, 1995). In this section, we conclude on our fifth question by answering the following: How do shrubs accumulate biomass compared to trees? Are there traits specific to old plants and traits specific to young plant?

3.2 - METHODS

3.2.1 - CHALLENGING THE MULTI-STEM SHRUB MODEL

Model dynamics

We further assessed our shrub model presented in chapter 2. We tested how the tree/grass ratio responds to MAP gradient across sub-Saharan Africa. We analysed three additional factors: the impact of the presence or absence of shrubs in our simulations; the response to different variables considered: canopy cover, number of individuals and leaf biomass; the ratio shrub/grass responds to MAP. Similarly, along MAP, we tested the relative prevalence of shrubs against trees, against all woody individuals, against all plant individuals and against the total number of potentially simulated individuals (i.e., accounting for dead cells).

We provide additional benchmarking of our model results. We use a newly available 1km resolution satellite derived AGBM map from Bouvet *et al.* (2018). We compare simulated AGBM at the site studies listed in chapter 2 to the satellite derived AGBM. We further analysed the response of satellite derived AGBM (Saatchi *et al.,* 2011; Baccini *et al.,* 2012; Avitabile *et al.,* 2016; Bouvet *et al.,* 2018) and tree height (Simard *et al.,* 2011) against MAP across sub-Saharan Africa and how our results compare to them. For tree height, we considered the mean, 50th, 80th, 90th and 95th percentiles of simulated tree height.

Data-model discrepancies were identified in chapter 2 in dry lands of Africa. The aDGVM2 simulations underestimate shrub prevalence in two regions: the Horn of Africa and the South-West of Africa (Namib to Kalahari). We conducted 10 replicate simulations focused on the South-West of Africa, parts of the Karoo and the Kalahari as well as small fractions of woodlands and savannas north of the Namib, following the same simulation protocol as in chapter 2. We investigated the following points for this region:

1 - We tested the impact of stochasticity on our ability to simulate vegetation in the study region. To this end we compared the mean, maximum, minimum, standard deviation and variance of the simulated shrub cover distribution across the region.

2 - We analysed the relative share of different plant life-forms according to the classification used in chapter 2. We then tested how different classification schemes impact our results (Table 3.1).

Table 3.1: Shrub and woody vegetation classification schemes tested for replicate simulations in south west Africa. Woody individuals are classified according to given thresholds.

	Scheme 1	**Scheme 2**	**Scheme 3**	**Scheme 4**	**Scheme 5**
Shrubs	>=4stems & <=4m		>=4stems & <=4m	<=4m	>=4stems
Tall shrubs		>=4stems & >4m			
Tree		<3stems & >4m			
Small tree		<3stems & <=4m			

The role of shrubs

We tested how introducing shrubs into aDGVM2 modifies simulated vegetation across sub-Saharan Africa by comparing simulations with & without shrubs. First, we compared canopy cover, mean height, AGBM and grass proportion (inverse indicator of woody cover) in the presence or absence of shrubs to identify the regions where the introduction of shrubs has the largest impact on indicators of ecosystem structure. We performed additional simulations for the combination of presence or absence of shrubs and/or fire which were not conducted previously. We then analysed how plant types (grasses, tall and short woody plants) respond to MAP across sub-Saharan Africa. Thus, we considered geographical patterns as well as ecological dynamics. We then analysed separately the responses of grasses and trees to the introduction of shrubs in aDGVM2 as well as the response of simulated fire frequency. Separating grasses and trees shows how the competitive balance is modified by the introduction of shrubs while considering fire frequency allows to test if there is a feedback between fire frequency and shrub cover.

3.2.2 - SHRUB DISTRIBUTION DEPENDS ON SHRUB CLASSIFICATION

Shrub distribution in non-geographic space

In order to identify a non-geographical distribution (i.e: in trait or climate space) pattern of shrubs we conduct further analyses using the simulations performed and presented in chapter 2. First we considered the relation between mean height and mean stem count, *per se*, and their relation relatively to MAP. This allows us to analyse their relation in trait space and their relation relatively to environmental driver. Then we focused on the relation between height and stem count traits & how it shapes our understanding of shrub distribution relatively to MAP by testing different classification schemes. We considered shrub distribution in terms of individual numbers and in terms of stem biomass, leaf biomass and crown area.

Classification

We tested how shrub classification cascades at biome level. To this end, we defined biome categories based on parameters relative to our shrub model (stem count and vegetation height), alongside more usual parameters (grass AGBM and woody cover). We tested different parameters combination and different thresholds combinations (Table 3.2).

We investigated how life-stages (age) and life-strategies (traits) are connected in our simulations. To do so, we tested how two traits relative to our shrub model (mean stem count and mean height) behave along MAP across sub-Saharan Africa depending on plants age. We focused on seedlings younger than 2 years old or excluded them and focused on old growth above 100 years old. We then questioned the relation between plant life-forms (woody, trees, shrubs, grasses) and life-strategies in terms of life-expectancy and turn-over dynamics. First, we analysed woody plants life duration strategy in contrast to all plants along MAP across sub-Saharan Africa. Second, we compared shrubs, trees (separated based on stem count threshold) and grasses life-expectancy strategies.

Finally, we tested how our simulated shrub distribution compares to satellite derived product (Tuanmu *et al.,* 2014) according to different shrub categorisation schemes (combinations of an array of height and stem count thresholds).

Table 3.2: Criteria used to classify simulated vegetation cover into biomes according to an array of schemes. When no values are assigned to a criterion, it means that this criterion was not used for this scheme. When no values are assigned to a biome, it means that it was not considered for this scheme.

Scheme 1								
Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*				
Grass AGBM	<0.5t/ha	>=0.5t/ha	>=0.5t/ha	>=0.5t/ha				
Woody canopy	<=15%	<=15%	>15% & <70%	>=70%				
Scheme 2								
Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*				
Grass AGBM	<1.5t/ha	>=1.5t/ha	>=1.5t/ha	>=1.5t/ha				
Woody canopy	<=20%	<=20%	>20% & <80%	>=80%				
Scheme 3								
Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*			
Grass AGBM	<0.5t/ha	>=0.5t/ha	>=0.5t/ha	>=0.5t/ha	<0.5t/ha			
Woody canopy	<=5%	<=15%	>15% & <70%	>=70%	>5% & <=15%			
Scheme 4								
Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*			
Grass AGBM	<0.5t/ha	>=0.5t/ha						
Woody canopy	<5%	<15%						
Mean height			>1m & <4m	>=4m	<=1m			
Other			Non desert & non grassland					
Scheme 5								
Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*	*Shrub savanna*	*Tree savanna*	*Arid*
Grass AGBM	<0.5t/ha	>=0.5t/ha						
Woody canopy	<5%			>=15%		>=15%	>=15%	<=15%
Mean height		<=1m	>1m & <4m	>=4m		<=1m	>1m & <4m	
Other								Non-desert & non grassland
Scheme 6								
Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*	*Shrub savanna*	*Tree savanna*	*Arid*
Grass AGBM	<1.5t/ha	>=1.5t/ha						
Woody canopy	<20%	<20%						
Mean height			>1m & <4m	>=4m	<=1m			
Other			Non desert & non grassland					

Scheme 7

Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*	*Shrub savanna*	*Tree savanna*	*Arid*
Grass AGBM	<0.5t/ha	>=0.5t/ha						
Woody canopy	<15%	<15%						
Other			Non desert & non grassland					
Mean stem			>2 & <4	<=2	>=4			

Scheme 8

Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*	*Shrub savanna*	*Tree savanna*	*Arid*
Grass AGBM	<1.5t/ha	>=1.5t/ha						
Woody canopy	<20%	>20%						
Other			Non desert & non grassland					
Mean stem			>2 & <4	<=2	>=4			

Scheme 9

Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*	*Shrub savanna*	*Tree savanna*	*Arid*
Grass AGBM	<0.5t/ha	>0.5t/ha						
Woody canopy	<5%	<15%		>=15%		>=15%	>=15%	<=15%
Other								Non desert & non grassland
Mean stem				<3		>=5	>3 & <5	

Scheme 10

Biomes *Criteria*	*Desert*	*Grassland*	*Savanna Woodland*	*Forest*	*Mixed Shrubland*	*Shrub savanna*	*Tree savanna*	*Arid*
Grass AGBM	<0.5t/ha	>=0.5t/ha						
Woody canopy	<5%	<15%		>=15%		>=15%	>=15%	<15%
Other								Non desert & non grassland
Mean stem				<=2		>=3	>2 & <3	

Footprint of shrubs

We analyse how bare ground proportion (number of dead individual cells simulated by aDGVM2) responds to the presence or absence of fire, and to the presence or absence of shrubs, relatively to MAP. Testing the effect of shrubs on bare ground allows us to assess the footprint of shrubs. Testing the effect of fire on bare ground allows us to analyse further how fire shapes vegetation (or bare ground) cover. We tested bare ground prevalence across sub-Saharan Africa as a share of simulated living plant individuals, as a share of all simulated individual cells and as a share of the simulated canopy cover (each simulated dead cell accounting for an area equal to the 1ha simulated plot divided by the number of simulated individual cells, here 3600). In aDGVM2, simulations are divided in grid cells.

These grid cells are simulated based on a representative hectare stand. We simulate a potential of 3600 individuals per hectare stand. However, not all of them might be alive, thus, leaving unoccupied spots, which we consider as bare-ground.

3.2.3 - SHRUB LIFE-FORM *MODUS VIVENDI*

Resprouting *vs.* reseeding

According to Knox and Clarke (2005), resprouters have higher carbon allocation to roots than reseeders. According to Midgley (1996) resprouters tend to be multi-stemmed and smaller than reseeders which tend to be single-stemmed. We compare our results to those provided by Knox and Clarke (2005) and by Midgley (1996). To do so, we analysed the relation between carbon allocation to storage and reproduction of woody individuals for three cases of absence/presence of shrubs and fire, and then for shrubs and trees separately in a scenario with shrubs and fire simulated. We distinguish resprouters from reseeders based on their carbon allocation strategies. Specifically, we classify resprouters as allocating relatively more to their roots and storage carbon pools and less to their reproduction carbon pool, while reseeders allocate relatively more to their reproduction carbon pool and less to their roots and storage carbon pools.

Rooting strategies

We analysed plant strategies in terms of carbon allocation to roots and to shoot. We first focused on grasses by analysing how allocations to roots and to shoot react to MAP and how allocation to roots and rooting depth are related. We then compared ratio of shoot to roots biomass for grasses, trees and shrubs (categorised following chapter 2) separately.

Carbon strategies: shrubs *vs.* trees

We investigated the growth strategies of shrubs and trees (categorised following chapter 2) relatively to their carbon allocation strategies. To this end, we compared the ratio of biomass to age (indicating carbon accumulation through time) *vs.* mean height. This shows trees and shrubs trends of carbon allocation and growth strategy.

3.3 - RESULTS

3.3.1 - CHALLENGING THE MULTI-STEM SHRUB MODEL

Model dynamics

We observe that the ratio between grass and woody individuals relative to MAP, simulated by aDGVM2, is consistent with the distribution of savannas across sub-Saharan African. However, this relation depends on the model variables used to calculate this ratio. The ratio between the number of woody individuals and grass individuals (Figure 3.1, panel a) deviates from the ratios considering leaf biomass (Figure 3.1, panel b) or canopy area (Figure 3.1, panel c). Woody plants increasingly dominate over grasses with decreasing MAP, in terms of number of individuals, whereas this relation is reversed when considering canopy cover. The ratio is less dependent on MAP when considering leaf biomass, excepted at intermediate MAP in presence of fire, where we observe an increased competitiveness of grasses. This pattern is also visible when considering canopy area. Considering the ratio between grasses and woody individuals based on leaf biomass reveals an increased competitiveness of grasses at lower MAP, relatively to woody plants (Figure 3.1, panel d). Moreover, the introduction of shrubs shifts this relation (Figure 3.1, panel a) by increasing the prevalence of woody individuals at lower MAP. This is confirmed by considering shrubs (Figure 3.1, panel e) and trees (Figure 3.1, panel f) separately. At the lowest MAP values, competitiveness of trees *versus* grasses drops whereas competitiveness of shrubs continues to increase.

We clipped ratio values in order to avoid values tending towards infinity. We limited these values to thresholds of 10, 50 and 100 to optimise readability while retaining as much information as possible. The ratio between woody plants and grasses typically tends towards infinite values everywhere but in grassland, when considering canopy area or leaf biomass. In the reverse relation, grass to woody, there are only a few grid cells exhibiting values above set limits, which shows that in most grasslands there are woody plants, whereas woodlands and forest vastly outcompete grasses.

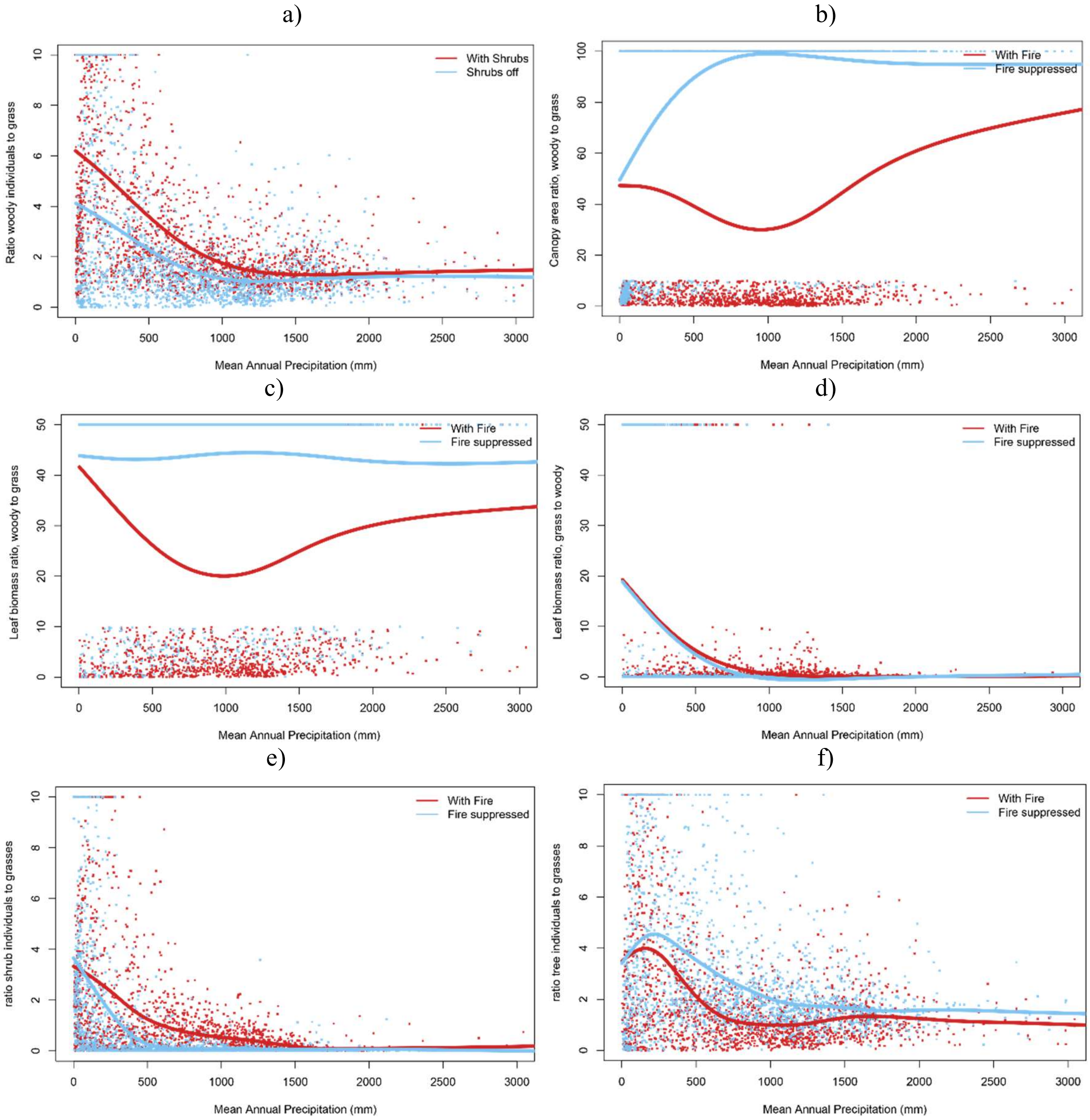


Figure 3.1: Ratios of plant types (grasses, shrubs, woody, trees) relative to number of individuals, leaf biomass and canopy area. Ratios superior to 10, 50 and 100 have been set to 10, 50 and 100, respectively. 10 for panels a, e and f, 50 for panels c and d, and 100 for panel b. Shrubs are classified as woody individuals with 4 or more stems while trees have 3 or less stems.

The aDGVM2 simulates a relation between shrubs and trees in relation to MAP that is consistent with our expectations across sub-Saharan Africa (Figure 3.1). With decreasing MAP we observe an increase in relative shrub competitiveness, first, at intermediate MAP, because of reduced tree prevalence, then, at lower MAP, because of increased shrub prevalence. At very low MAP (under 200 mm/year) trees and shrubs could be seen as co-dominants, however, it seems that actually either tree or shrubs dominate when looking at individual grid cells values (Figure 3.2, panel a). We also find (Figure 3.2, panel b) that below 500-700mm MAP, vegetation cover decreases, leaving more ground bare. Bare ground becomes a major feature of simulated land cover. Shrublands with 100% of the woody population being shrubs start to appear below MAP of *ca.* 750mm and become more frequent below *ca.* 205mm MAP (Figure 3.2, panel c).

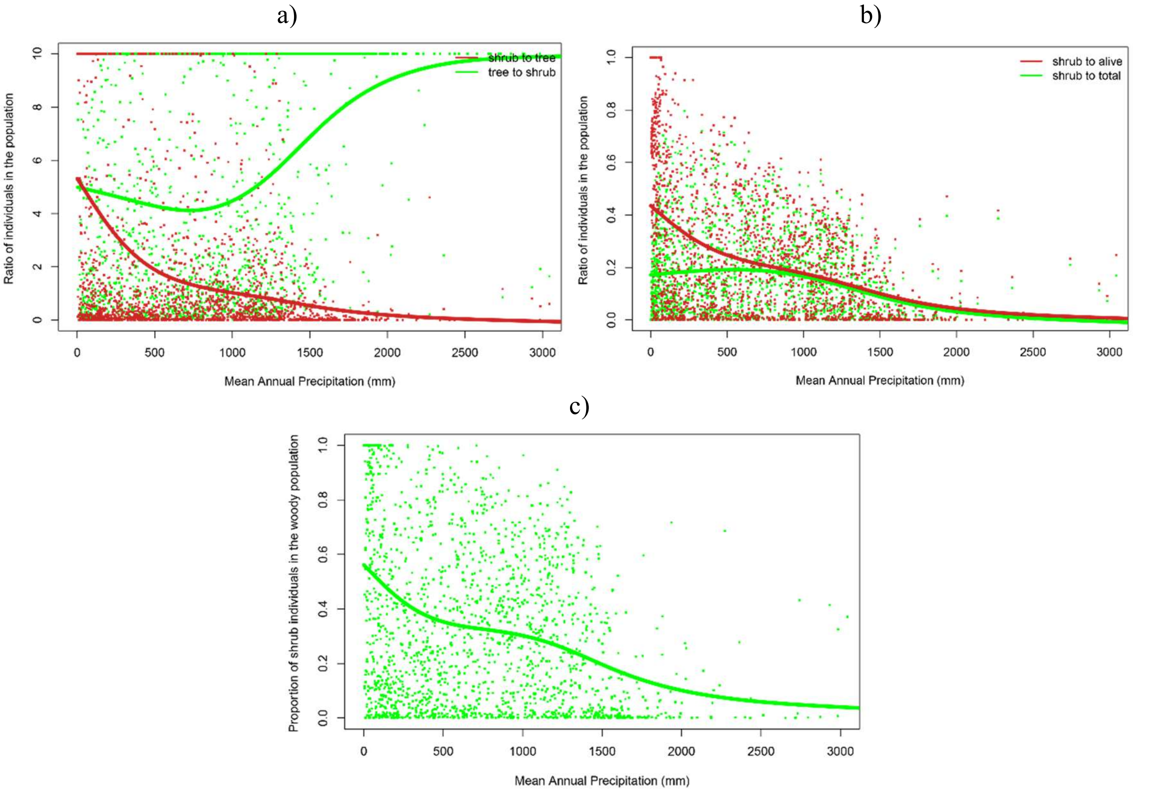

a)
b)
c)

Figure 3.2: Ratios of individuals in the population, relatively to plant types (trees, shrubs) (panel a) and relatively to total plant individuals and to total simulated individual cells (panel b). Ratio of individuals in the population (panel a) have been limited to 10 (i.e., ratio superior to 10 are set to 10) to increase readability.

Adding more satellite derived products to benchmark our model increases the range of values given for African vegetation AGBM. Specifically, Bouvet *et al.* (2018) found higher biomass for most of the sites considered compared to alternative satellite derived products (Table 3.3). Biomass provided by Bouvet *et al.* (2018) agrees better with aDGVM2's simulations results. The higher resolution of the datasets used to benchmark our results (Table 3.3) accounts for more diversity at the landscape scale than aDGVM2 captures in our simulations. However, this information is lost when simply aggregated at the 1° scale of our simulations. To mitigate this loss we tested different aggregation methods (not shown), which do not change the general pattern. However, considering the range of values (minimum and maximum) aggregated at the 1° resolution for each data set, we increase the overlap with our results.

Table 3.3: Aboveground biomass for savanna sites, from field observations, remote sensing products (Bouvet et al. 2018, Avitabile et al., 2016; Baccini et al., 2012; Saatchi et al., 2011), and aDGVM2 simulations. Field values sites 1–4: Higgins et al. (2007), site 5: Rutherford (1979), site 6: Rutherford (1975), site 7: Kelly and Walker (1976), site 8: Rushworth (1975).

Site ID	1	2	3	4	5	6	7	8
Site longitude (°E)	31.4	31.8	31.6	31.3	28.7	17.1	31.6	26.3
Site latitude (°S)	23.5	24.4	25	25.2	24.7	20.4	27.2	18.9
Field observation								
Minimum (t/ha)	3	3	7	13	__	__	8.7	4.7
Mean (t/ha)	__	__	__	__	16.3	22.3	16.7	__
Maximum (t/ha)	5	12	11	23	__	__	30.8	5.8
aDGVM2 results								
Minimum (t/ha)	23	29.8	36.3	31.5	2.2	11.1	4.7	15
Mean (t/ha)	44.7	49.5	58.5	79.2	50	28.8	44.2	40.5
Maximum (t/ha)	79.4	75.9	82.3	120.3	91.8	58	110.3	71
Satellite data								
Saatchi et al. (2011) (t/ha)	17.7	21.3	26.6	49.9	16.9	12	32.8	22.8
Baccini et al. (2012) (t/ha)	23.1	23.1	23.1	__	10.1	4.6	__	23
Avitabile et al. (2016) (t/ha)	3	4.1	7.4	9.1	2.4	1.5	9.5	6.5
Bouvet et al. (2018) (t/ha)	41	29	39	32	39	11	61	16

We observe large discrepancies between satellite products, ground studies and our results. Nonetheless, the overall pattern of biomass appears to be similar among satellite derived products and our results, i.e. biomass increases with MAP.

The aDGVM2 produces low AGBM at low MAP, and mainly high AGBM at higher MAP (Figure 3.3). In comparison, satellite derived products show a distribution of low and high AGBM throughout the MAP gradient, despite also displaying an increase with increasing MAP (Figure 3.3). Tree height simulated by aDGVM2 follows similar trend as simulated AGBM while satellite derived tree height follows similar trend to satellite derived AGBM (Figure 3.3).

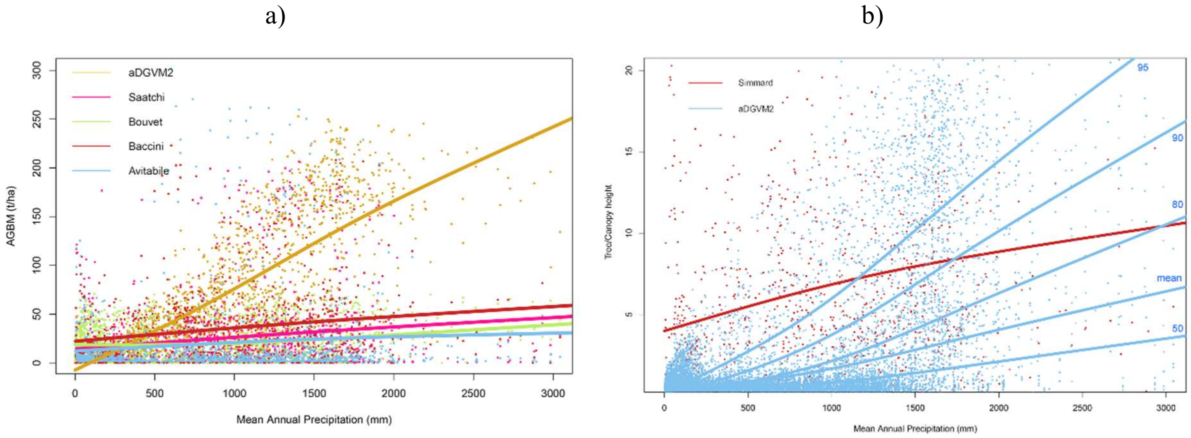


Figure 3.3: a) Aboveground biomass (t/ha) from remote sensing products (Bouvet et al. 2018, Avitabile et al., 2016; Baccini et al., 2012; Saatchi et al., 2011), and aDGVM2 simulations across sub-Saharan Africa along MAP (mm). b) Tree canopy height from remote sensing product (Simard et al. 2011) and the 95th, 90th, 80th and 50th percentiles of the tallest tree and mean tree height simulated by aDGVM2.

We conducted 10 replicate randomized simulations for the African South-West region, which confirm the discrepancies between our results and satellite derived products (Figure 3.4). However, we also observe a potential for lower discrepancies. Indeed, in some replicates, our model simulates a shrub cover more in range with satellite derived products for some grid cells (Figure 3.4, panel c) whereas on average aDGVM2 underestimates shrub cover, as shown by the standard deviation, representing up to ±40% shrub cover (Figure 3.4, panel e).

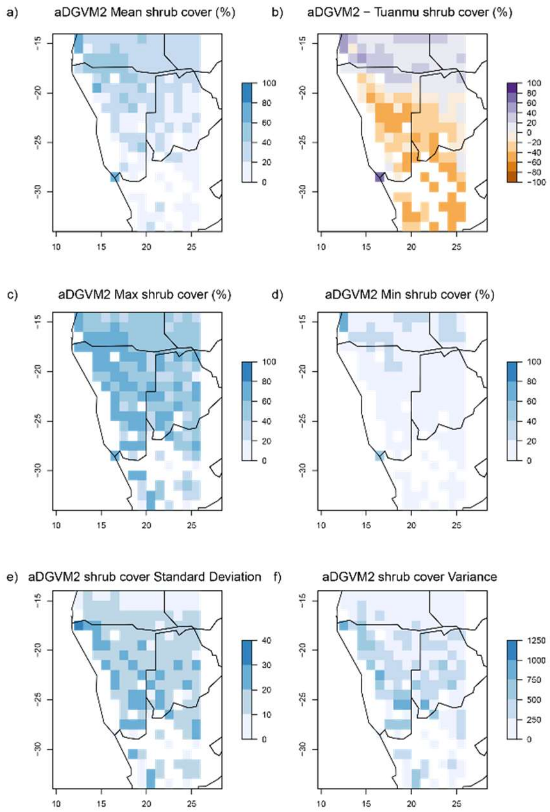
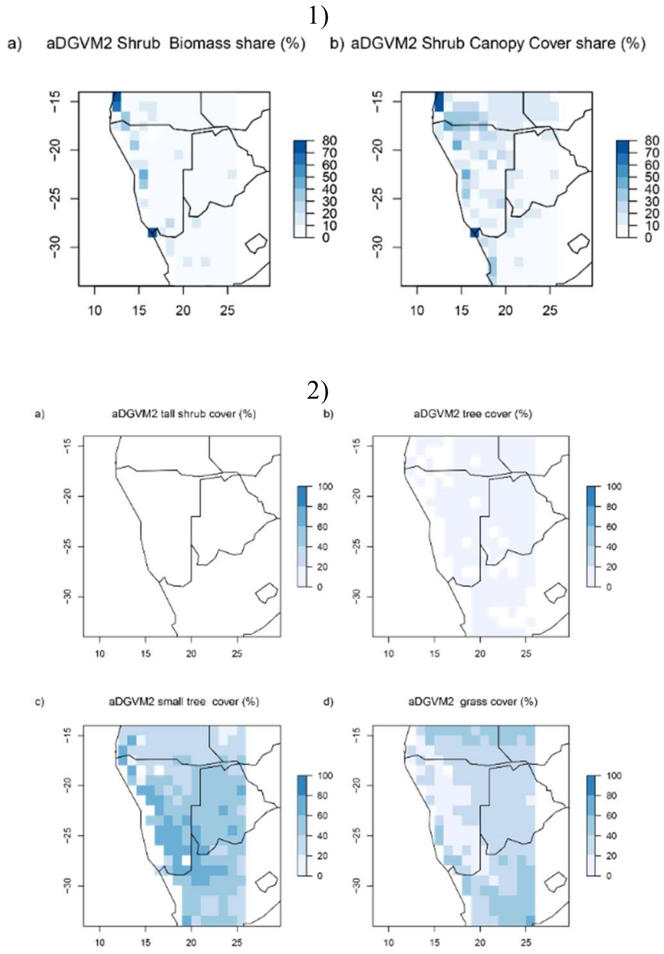


Figure 3.4: Mean (a), maximum (c), minimum (d), standard deviation (e) and variance (f) of shrub cover among 10 replicates simulations. Shrubs are classified as woody plant with 4 or more stems and up to 4m tall. b) aDGVM2 shrub cover as proportion of shrub individuals minus satellite derived vegetation cover classified as shrubs (Tuanmu et al. 2014).

Shrubs' classification schemes shape agreement between aDGVM2 and satellite derived products. The shrub cover of the satellite derived product is not directly a measurement of the canopy cover of shrubs defined as multi-stemmed plants with a height of up to 4m. It is in fact the categorisation of data from satellite sensors as area predominantly covered by shrubs or shrubland based on expert knowledge. Therefore, we used different classification methods to investigate the divergence between expert knowledge and aDGVM2 results.

We observe that the proportion of shrub of total vegetation is lower in terms of biomass and canopy cover (Figure 3.5, panel 1.a) than it is in terms of individual numbers (Figure 3.4). Most of the simulated vegetation population is made up of small trees in the Namib region, whereas tall shrubs are absent (Figure 3.5, panel 2). Classifying shrubs based on stem number leads to an underestimation of shrub cover compared to satellite derived product (Tuanmu and Jetz, 2014) (Figure 3.5, panel 5), while classifying them only based on height leads to a better fit with the benchmarking product in the most arid area and to an overestimation in area with higher vegetation cover (Figure 3.5, panel 4). It appears that when classifying shrubs, stem number overrides the effect of height, given that classification results based on height and stem together (Figure 3.5, panel 3) are more similar to classification results based on stem alone (Figure 3.5, panel 5) than based on height alone (Figure 3.5, panel 4). Finally, we observe that boundaries between vegetation cover classes are sharper in satellite derived products (Figure 3.5, panel 6) compared with aDGVM2 results.



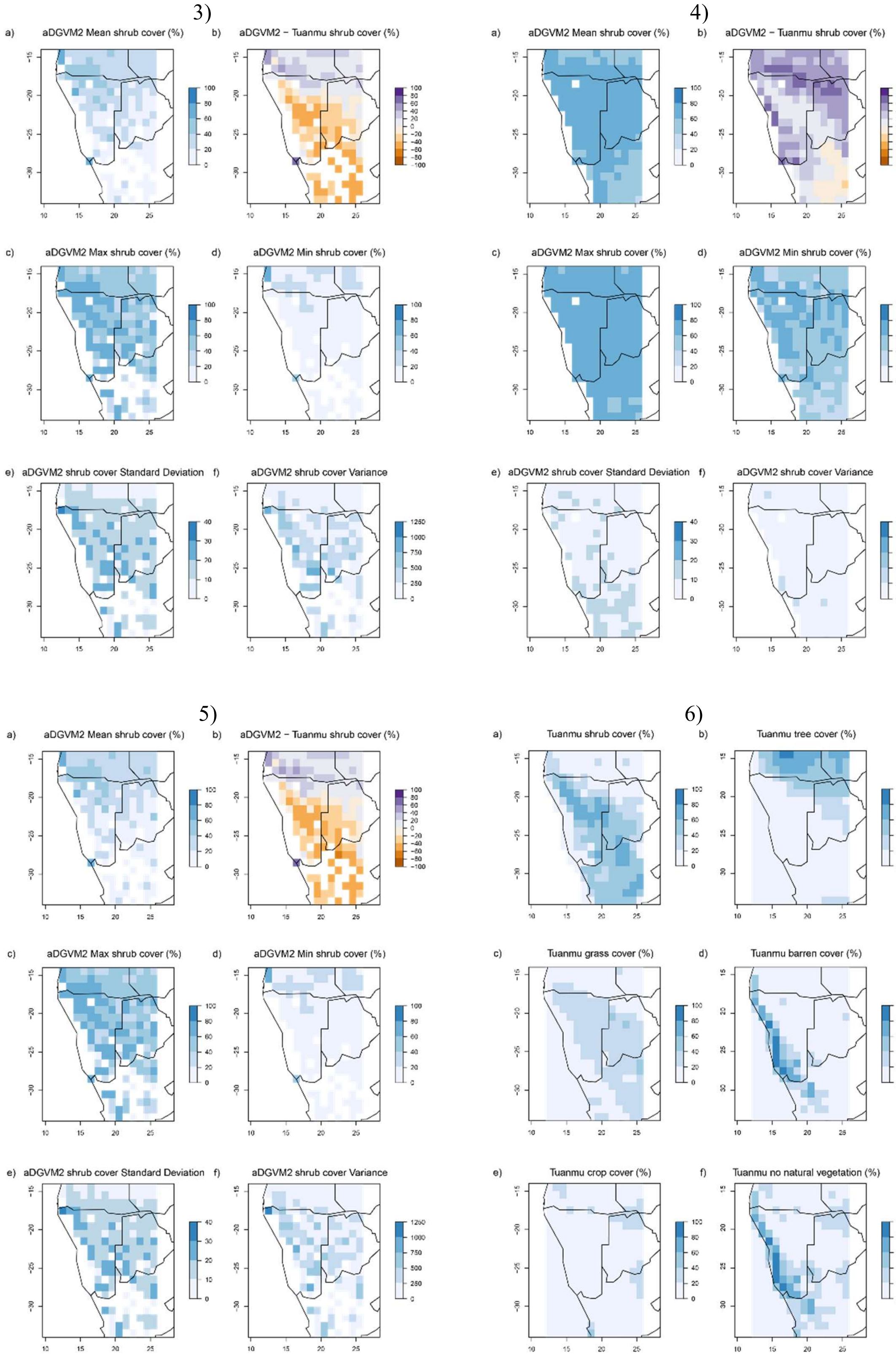


Figure 3.5: Following schemes presented in Table 1. 1) Share of a) biomass and b) canopy cover made of shrubs (Scheme 1). 2) Share of woody individuals classified (Scheme 2) as a) tall shrub, b) tree, c) small tree, d) grass. 3) Share of vegetation cover made of shrubs (Scheme 3). 4) Share of vegetation cover made of shrubs (Scheme 4) 5) Share of vegetation cover made of shrubs (Scheme 5) 6) Land cover classes as defined by Tuanmu et al. 2014; tree cover aggregates its 4 tree types; "no natural vegetation" aggregates its classes 7 to 12 (cultivated and managed, regularly flooded, urban, snow/ice, barren, open water).

The role of shrubs

By comparing simulations with and without shrubs, we can assess the effect of shrubs on simulated vegetation. Introducing shrubs modifies vegetation distribution mainly across savannas and woodland regions bordering the central African rainforest by decreasing AGBM, canopy area and to a smaller extent mean height (Figure 3.6).

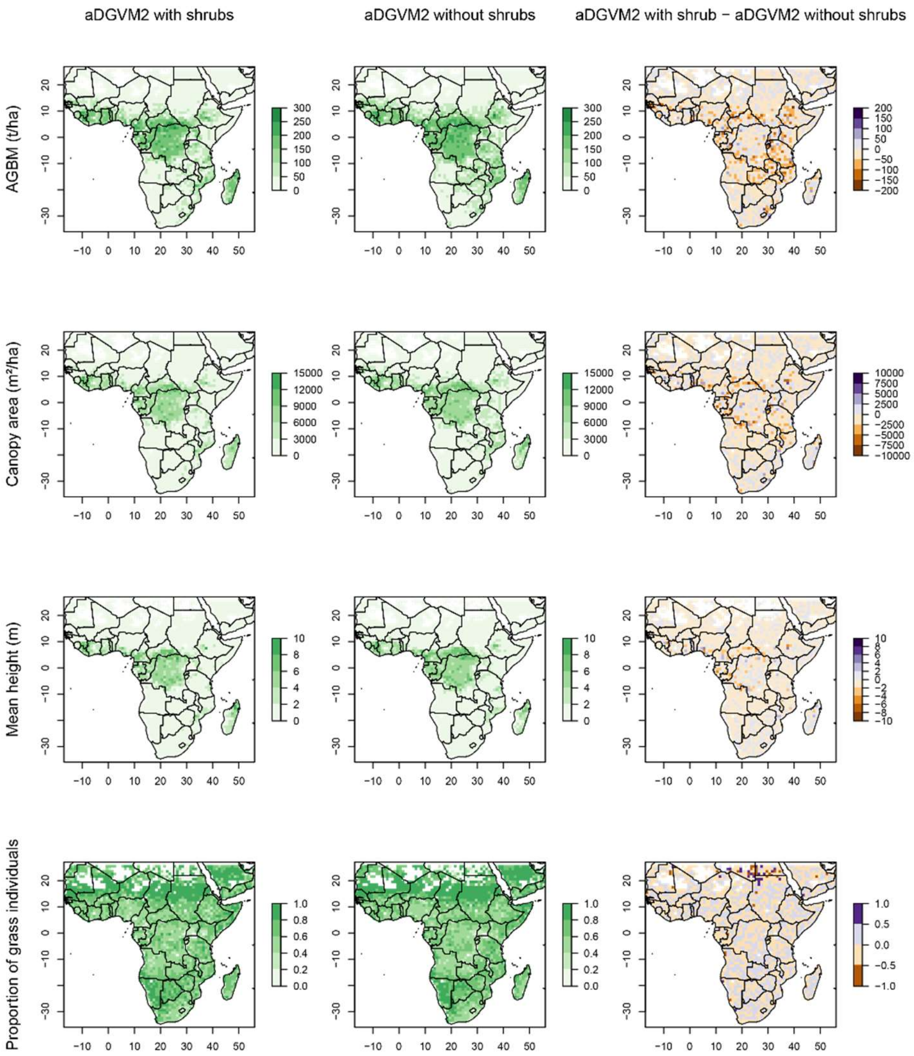


Figure 3.6: Effects of shrubs on vegetation features: AGBM (t/ha), canopy area (m²/ha), mean height (m), and proportion grass, as simulated by aDGVM2 with or without shrubs included and the difference.

Competitive interactions between grasses and woody vegetation are modified by the introduction of shrubs. Introducing shrubs in aDGVM2 induces more plant individuals per simulated hectare stand in non-hyper arid regions (Figure 3.7). This is due to an increase in grasses and small woody individuals, while the number of tall woody individuals decreases.

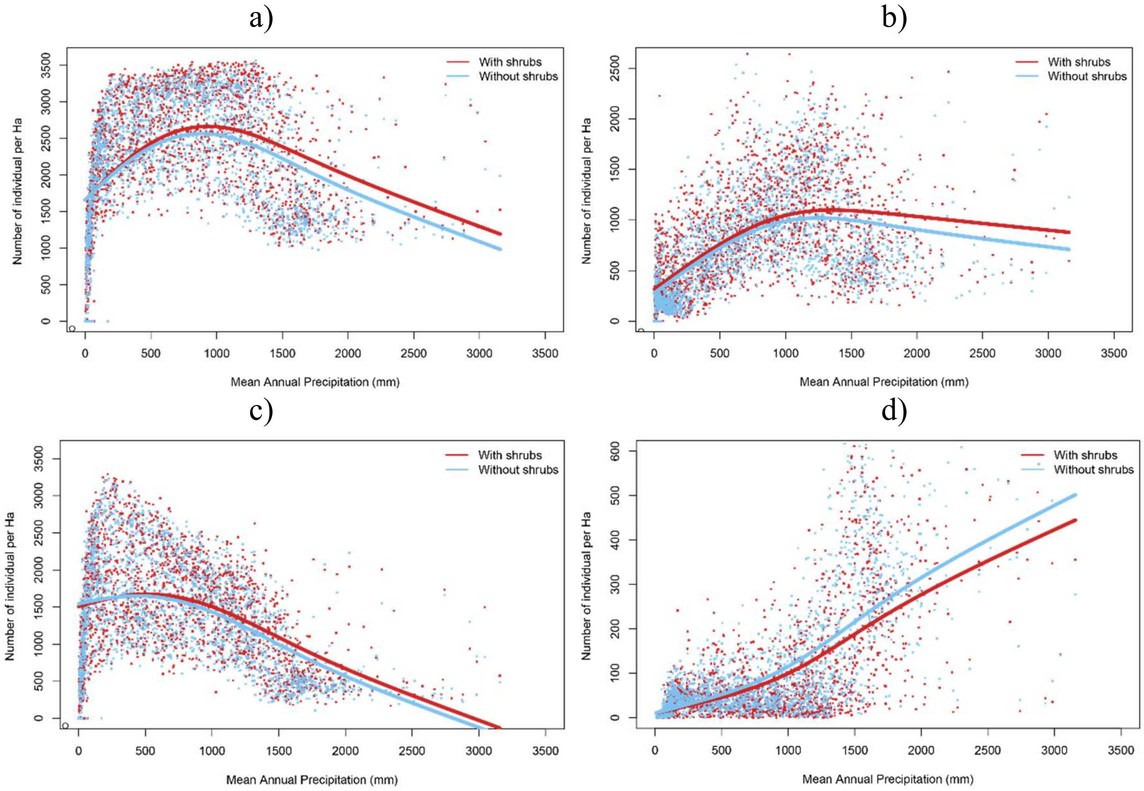


Figure 3.7: Number of individuals per hectare. a) all alive plants, b) grass, c) small woody plants (shorter than 4m), d) tall woody plants (taller than 4m).

Our results suggest a complex relation between fire and shrub cover (Figure 3.8). Adding shrubs to aDGVM2 simulations increases grass biomass all over sub-Saharan Africa while decreasing tree biomass mainly in savanna regions. Moreover, the effect of introducing shrubs on fire frequency does not follow a specific geographic pattern.

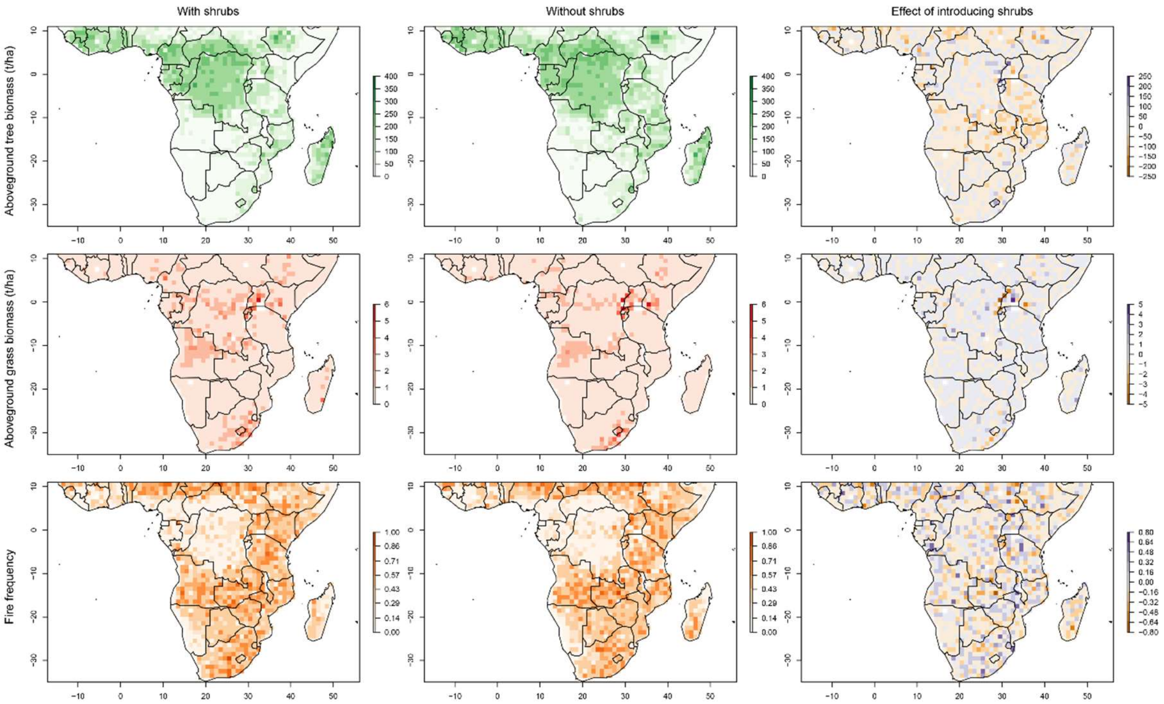


Figure 3.8: Aboveground tree and grass biomass (t/ha) and fire frequency, as simulated by aDGVM2 with or without shrubs included and the difference.

3.3.2 - SHRUB DISTRIBUTION DEPENDS ON SHRUB CLASSIFICATION

Shrub distribution in non-geographic space

The shrub strategy, as modelled by aDGVM2 according to our multi-stem scheme, allows for shrubs, as distinct life-form, to emerge. Indeed, we observe a relation between height and stem count per individual indicating that trees and shrubs are distinct life-form strategies; yet, overlapping at low height and low stem count (Figure 3.9, panel a). This overlap is found mainly at low MAP (Figure 3.9, panel b).

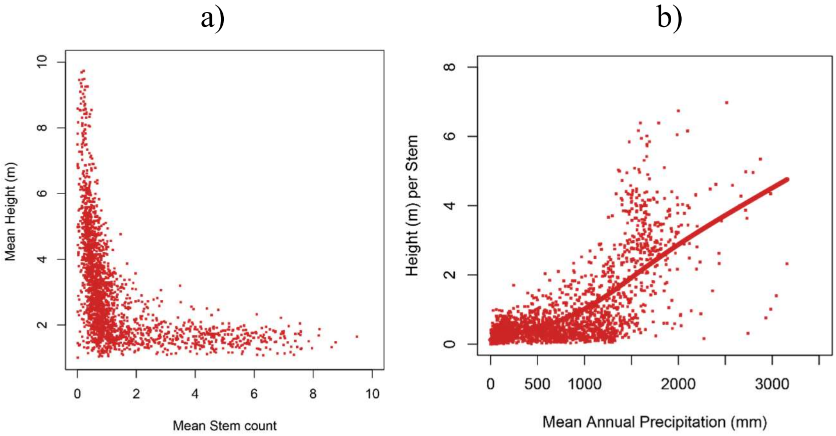

a)
b)

Figure 3.9: a) Mean vegetation height (m) versus mean stem count, b) height (m) per stem versus MAP (mm) across sub-Saharan Africa; as simulated by aDGVM2 per grid-cell.

The simulated plant life-forms ratios are unequivocal in Figure 3.2. However, it is a matter of perspective. When we use different classification schemes for shrubs and trees, the conclusions we can draw from our results are slightly different. If we consider only height to classify shrubs (Figure 3.10, panel a), their number is much higher than when considering only stem number (Figure 3.10, panel b). Simultaneously, stem number determines shrub characterisation, and thus their abundance, more strongly than height (Figure 3.10, panel c).

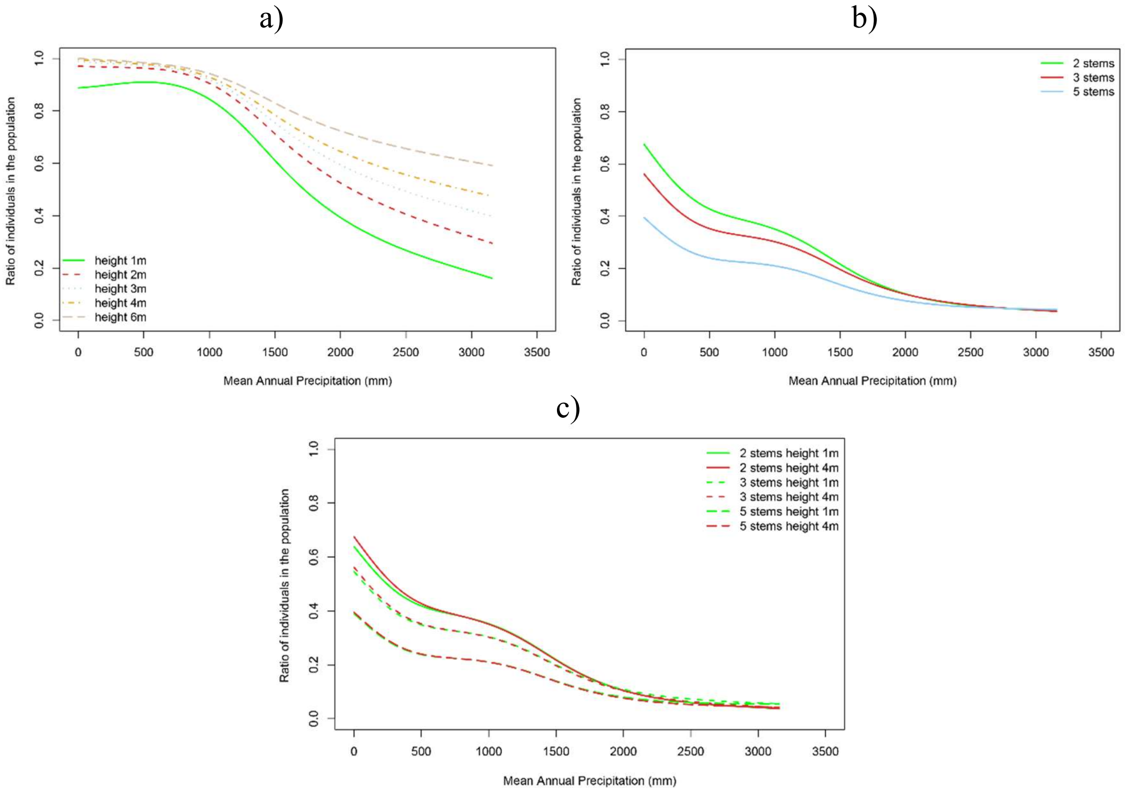


Figure 3.10: Ratio of shrub individuals in the vegetation population vs. MAP with shrubs classified as woody plants, a) with height smaller or equal to different thresholds, b) with x stems or more, c) with height smaller or equal than different thresholds and with x stems or more.

Thresholds used to classify results impact their analysis. Considering different parameters has similar effects. The simulated prevalence of shrubs, characterised based on height alone, does not follow the same pattern when considering crown area (Figure 3.11, panel a), instead of individuals number (Figure 3.2, panel a). Considering stem biomass, leaf biomass or crown area change shrub prevalence. Yet, all the characterisation schemes we tested show similar trends: no matter the parameter and thresholds considered, we observe an overall increase of shrub prevalence with decreasing MAP.

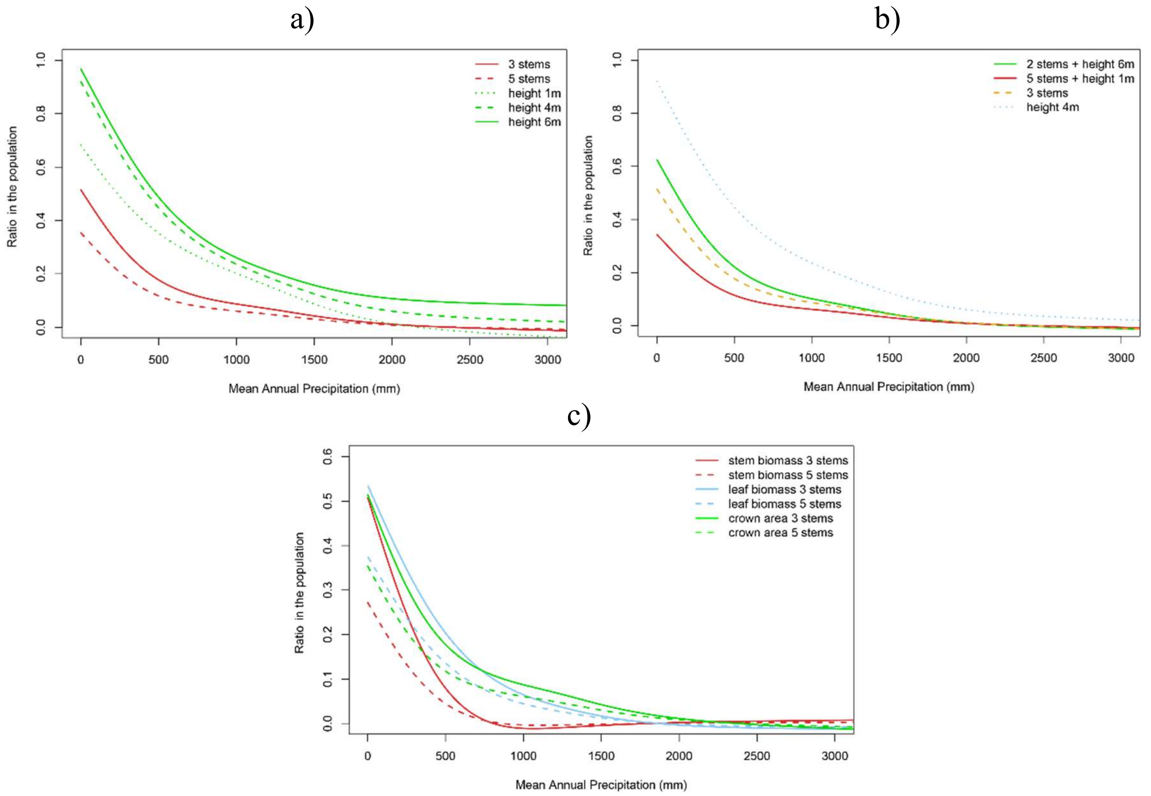


Figure 3.11: Shrub ratios in the plant population vs. MAP based on: a) crown area. Shrubs characterised as woody plants, either based on height thresholds or on stem count thresholds (shorter than height threshold; with more stem than stem threshold), b) crown area. Shrubs characterised as woody plants, either based on height thresholds or on stem count thresholds, or on both thresholds simultaneously (shorter than height threshold; with more stem than stem threshold), c) either crown area, leaf biomass or stem biomass. Shrubs characterised as woody individuals with x or more stems.

Classification

The results of aDGVM2 simulations provide directly relevant ecological information, yet, classifying them is necessary to further analyses. However, using various definitions of shrubs and trees induces uncertainties. We illustrated this effect at the individual level (Figure 3.5 & 3.11), and it cascades to the biome level (Figure 3.12). The threshold set to categorise biomes has tremendous effects at the continental scale (Figure 3.12, panels a and b). Adding more biome classes increases the precision of our interpretation of aDGVM2's results (Figure 3.12, panel c *vs.* panel a). Changing the parameters used to categorise biomes (Figure 3.12, panel a: vegetation cover; panel d: height; panel h; stem count) impacts our understanding of vegetation cover, as does changing the thresholds categorising biomes. Overall, more complex classification systems imply more complex spatial patterns of biomes at continental scale.

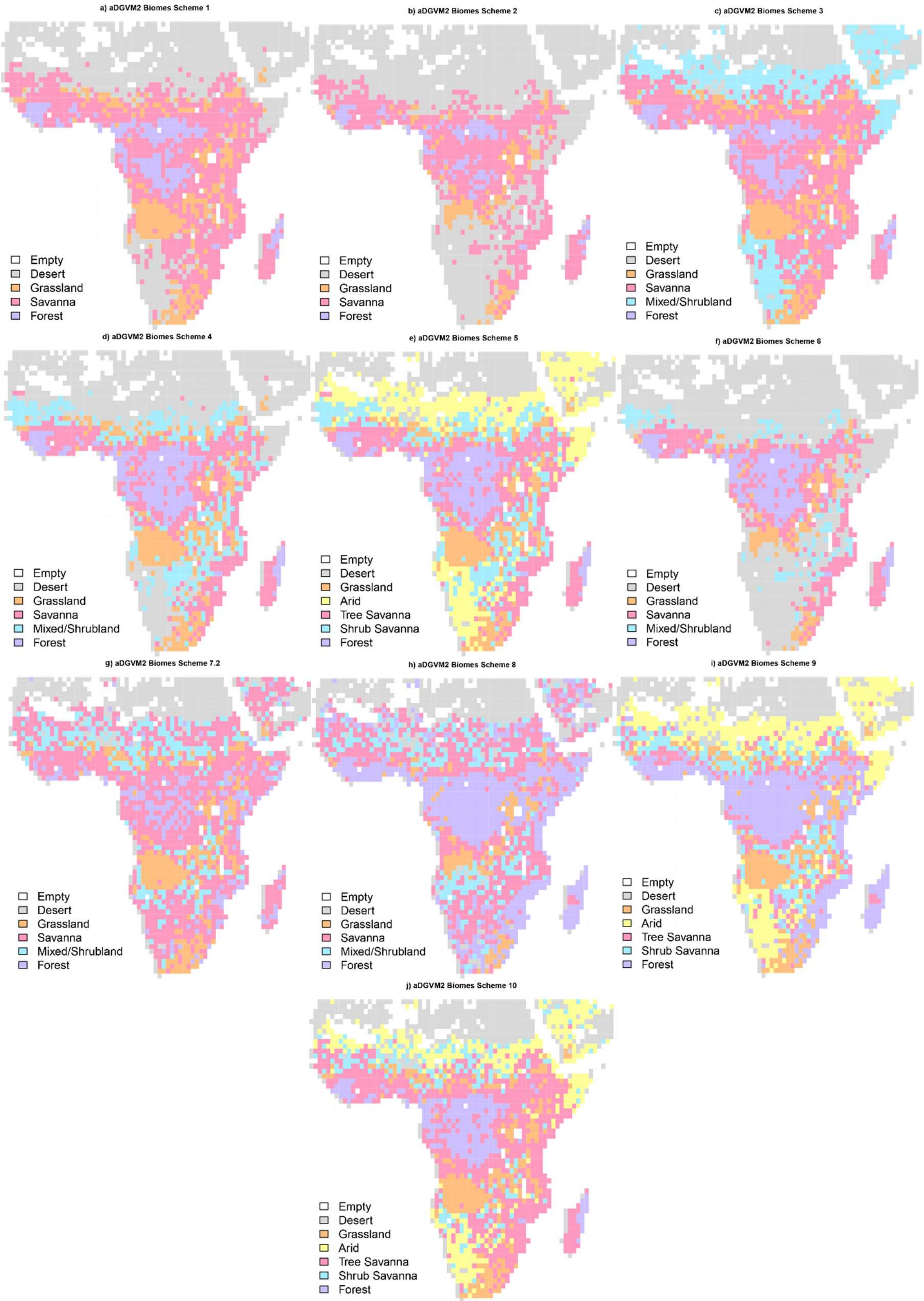


Figure 3.12: Classifications of simulated vegetation cover across sub-Saharan Africa, following schemes presented in Table 3.2.

The thresholds we use to define categories (biomes, life-forms) also depend on the categories we consider. For example, as we observe in Figure 3.13, considering plant age (e.g.: to focus on established individuals by excluding seedlings) has varying consequences depending on the parameter considered and depending on prevailing environmental conditions. Mean stem count is impacted by age more strongly at lower MAP while mean height is more strongly impacted by age at high MAP.

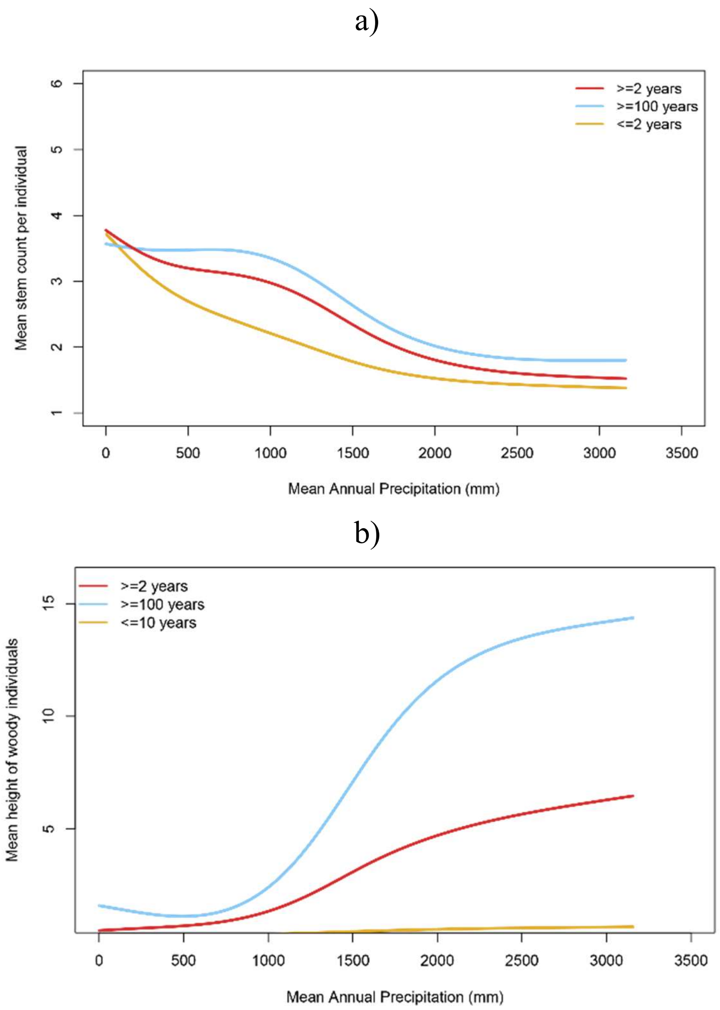


Figure 3.13: Mean stem count (panel a) and mean height (panel b) of woody individuals vs. MAP according to plant age.

Considering seedlings as trees or not, and how to define seedlings has different impacts in regions with high longevity of individuals and in regions with a higher turn-over. We observe that plants simulated by aDGVM2 reach, on average, older age at intermediate MAP, while we observe relatively higher turn-over (shorter lived individuals) at low and high MAP (Figure 3.14, panel a). The number or individuals per age value sets apart short- & long-lived strategies across sub-Saharan Africa and among plant types (Figure 3.14, panel b).

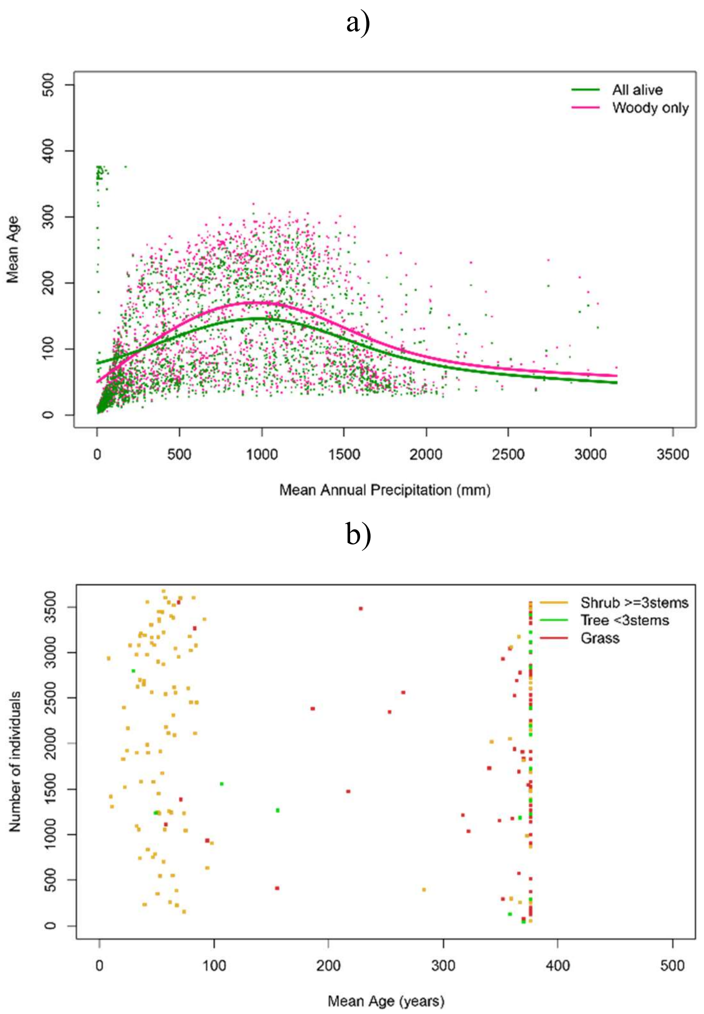


Figure 3.14: a) Mean age of all plants or of woody plants only vs. MAP, b) number of individuals vs. mean age for vegetation classes (shrubs and trees being discriminated based on stem count).

The categorisation scheme we use impacts our assessment of model performance. Depending on the thresholds we used, our results' accuracy change. Shrub cover, in terms of number of individuals, can vary vastly depending on the threshold and parameters used to characterise shrubs, which obscures the benchmarking process (Figure 3.15). This holds true if shrub cover is defined based on canopy cover or on biomass (figure not shown). The shrub category of Tuanmu *et al.* 2014, compared to aDGVM2 simulations, is as close to shrubs as we define them as it is to small woody individuals with a low stem count (i.e., shrubs with less stems), which could be classified as bushes in contrast to shrubs (Figure 3.15).

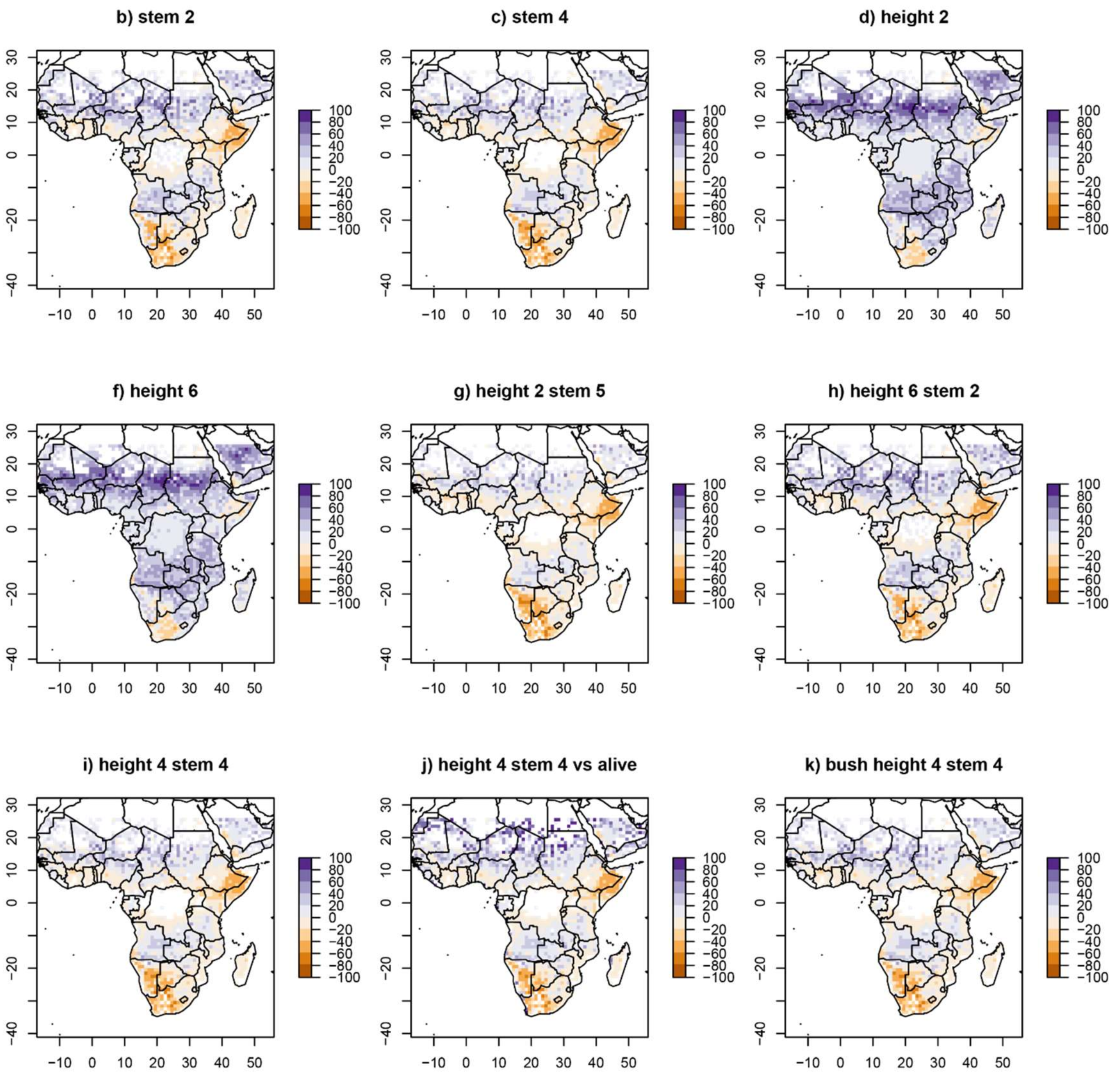


Figure 3.15: Difference between shrub cover simulated by aDGVM2 and derived from satellite data (Tuanmu et al. 2014). Panels a) to h): shrubs defined either as smaller than height threshold (meters) or as having more stems than a threshold or considering both stem number and height at the same time. Panel h): shrub cover calculated as a share of simulated alive plant individual and not as a share of the total individual space. Panel i) bush cover considered instead of shrub cover, with bush defined as shorter than 4m and with less than 4 stems

Footprint of shrubs

Investigating the imprint of our studied object can yield additional information about it. We analyse vegetation communities by looking at the bare-ground simulated by aDGVM2, either in terms of individual cells number or in terms of area cover (i.e., different from vegetation cover alone as it adds bare-ground to vegetation cover). We observe that the share of dead cells (i.e., individual cells with no living plant individual), or bare-ground, increases with increasing aridity, irrespective of the simulations settings, except for a slightly higher prevalence of bare-ground for intermediate values in the scenario without fires and with shrubs (Figure 3.16). More precisely, simulated bare-ground prevalence, in terms of share of individual cells, is higher in the tropical forest and more particularly in hyper-arid and desert regions. When considering ground cover, we see that this bare-ground is shaded by canopy cover in the tropical forest, while it represents most of the landscape in deserts and hyper-arid regions (Figure 3.17).

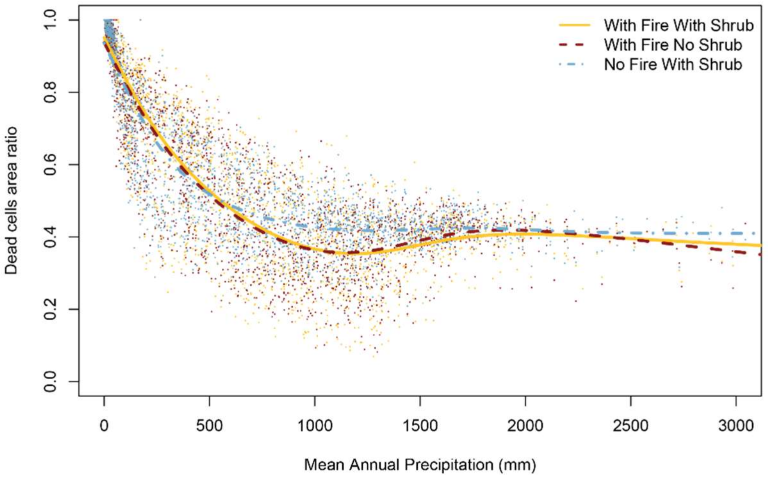


Figure 3.16: Ratio of dead individual cells (i.e., where no plant is simulated) per hectare stand vs. MAP across sub-Saharan Africa, simulated by aDGVM2 either with fires and shrubs, or with fires but without shrubs, or without fires but with shrubs.

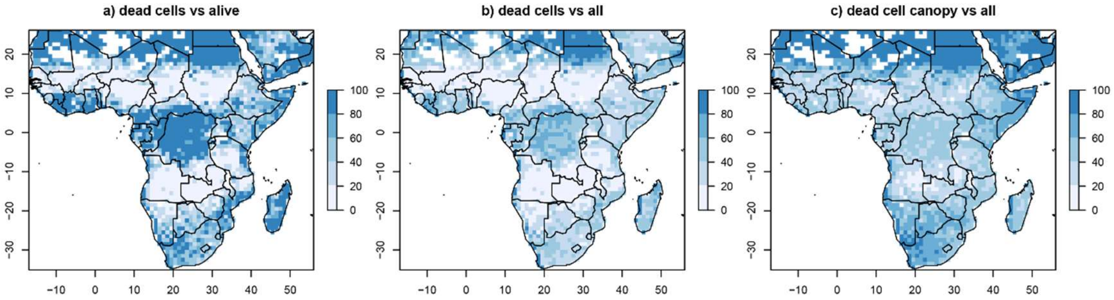


Figure 3.17: a) Percentage of dead cells (i.e., bare-ground, or empty cells) in comparison to cells with living plant individuals, per simulated hectare stand, b) percentage of dead cells in comparison to the total number of simulated cells, per simulated hectare stand, c) percentage, in terms of area, of dead cells in comparison to actual canopy cover, per simulated hectare stand (the total area can be superior to 1Ha as overlap can happen).

3.3.3 - SHRUB LIFE-FORM *MODUS VIVENDI*

Resprouting *vs.* reseeding

The aDGVM2 produces a weak distinction between resprouters and reseeders strategies as allocation to storage and allocation to reproduction are inversely related (Figure 3.18, panel a); yet, vegetation population does not show a strong division in separate clusters. Indeed, mean allocation to storage and to reproduction do not follow any clearly marked distribution pattern through Africa (figure not shown), indicating that both strategies tend to coexist. However, it appears that shrubs, characterised as multi-stemmed woody plants, have a strongly specific strategy, typically exhibiting high allocation to storage, whereas trees can allocate either more to storage or more to reproduction (Figure 3.18, panel b). It could indicate that trees can be either resprouters or reseeders while shrubs are obligatory resprouters.

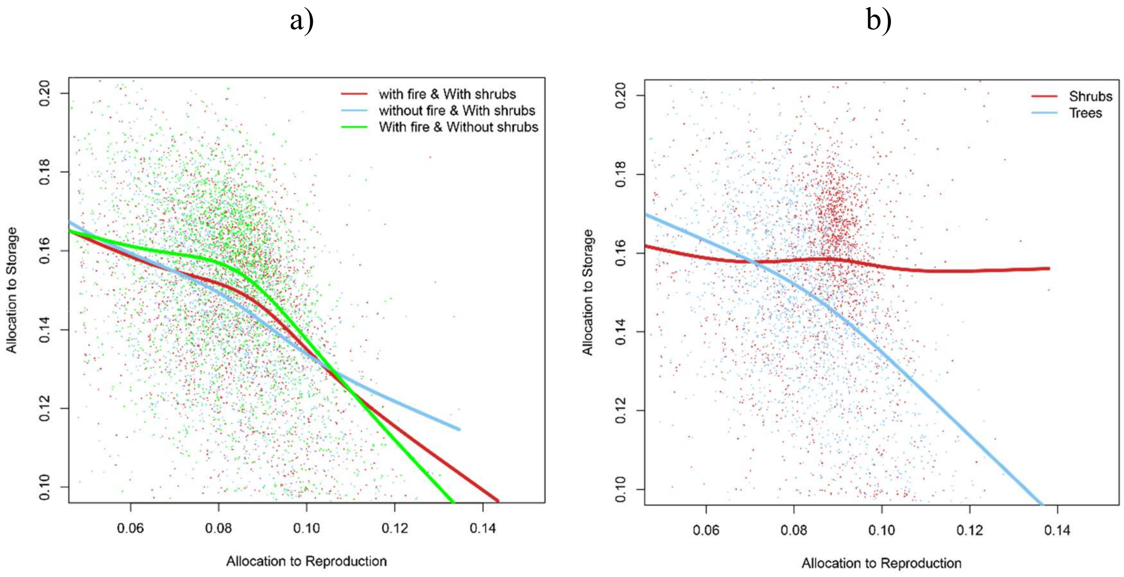
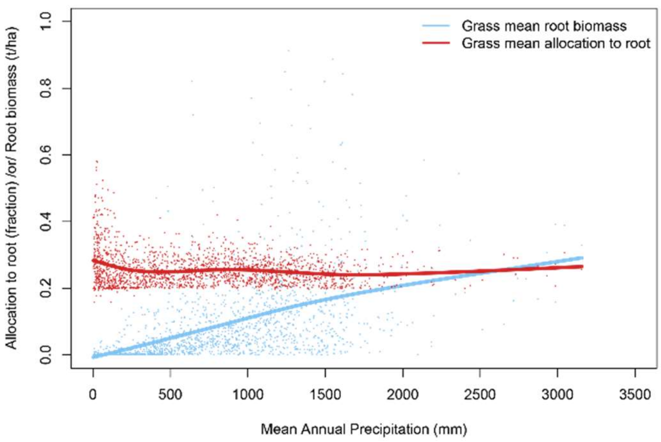


Figure 3.18: Relation between allocation to storage and allocation to reproduction, a) for all woody individuals under three different scenarios, b) for trees and shrubs separately, with shrubs defined as woody plant with 4 or more stems, for a scenario with shrubs and fire simulated.

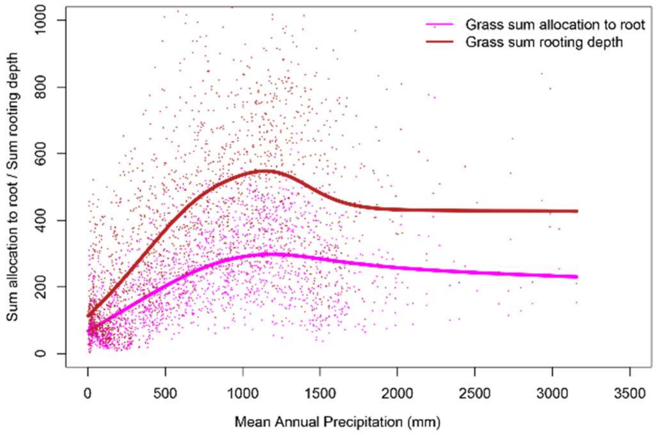
Rooting depth strategies

Schenk and Jackson (2002a) found that herbaceous plants have larger root system in dry climate. Our results partially agree with these findings. aDGVM2 simulates a slight increase in grass allocation to root at low MAP agreeing with observations, while simulating lower mean roots biomass per grass individual, disagreeing with observations (Figure 3.19, panel a). However, root biomass can translate into different rooting systems (i.e., different volume and depth). Additionally, aDGVM2 simulates higher allocation to roots and higher rooting depth of the total grass population at intermediate MAP (Figure 3.20, panel b). More generally Pallardy (1981), Wilson (1988), Chapin, Autumn and Pugnaire (1993), all observed an increase in root to shoot ratio with increasing aridity. Our results are less straightforward as shoot to root ratios are relatively constant for grasses and highest for shrubs and trees at intermediate MAP; yet, underground biomass increases relatively to aboveground biomass with decreasing MAP.

a)



b)



c)

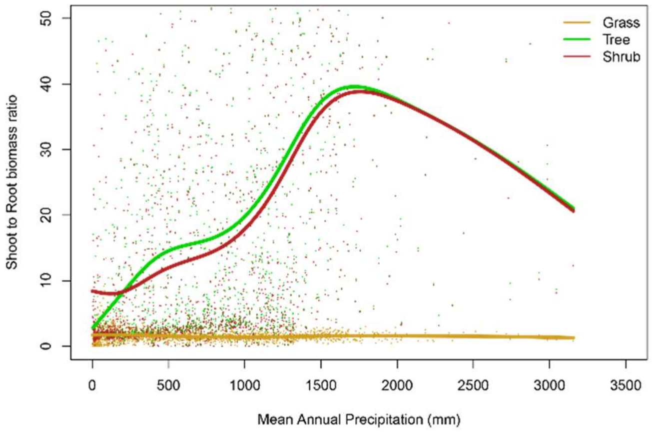


Figure 3.19: a) Grass mean allocation to root and root biomass (t/ha) in relation to MAP (mm), b) sum of grass rooting depth and of allocation to roots in relation to MAP (mm), c) ratio of shoot to root biomass for grass, tree and shrub (defined as woody plants with 4 or more stems) in relation to MAP (mm).

Carbon strategies: shrubs *vs.* trees

Plant growth rate can be understood as the biomass accumulated per year. By comparing the ratio between plant biomass and age, against its height, for shrubs and trees separately, we observe that shrubs have an increasingly faster growth rate with increasing height, whereas trees tend to level off (Figure 3.20). However, drawing relevant conclusion on these behaviours is challenging as shrubs do not reach heights similar to trees. This might indicate that big shrubs have the ability to store more carbon than small shrubs. Concomitantly, it would mean that more mature trees have a decreasing ability to store carbon, relatively to their size, probably due to their higher maintenance cost.

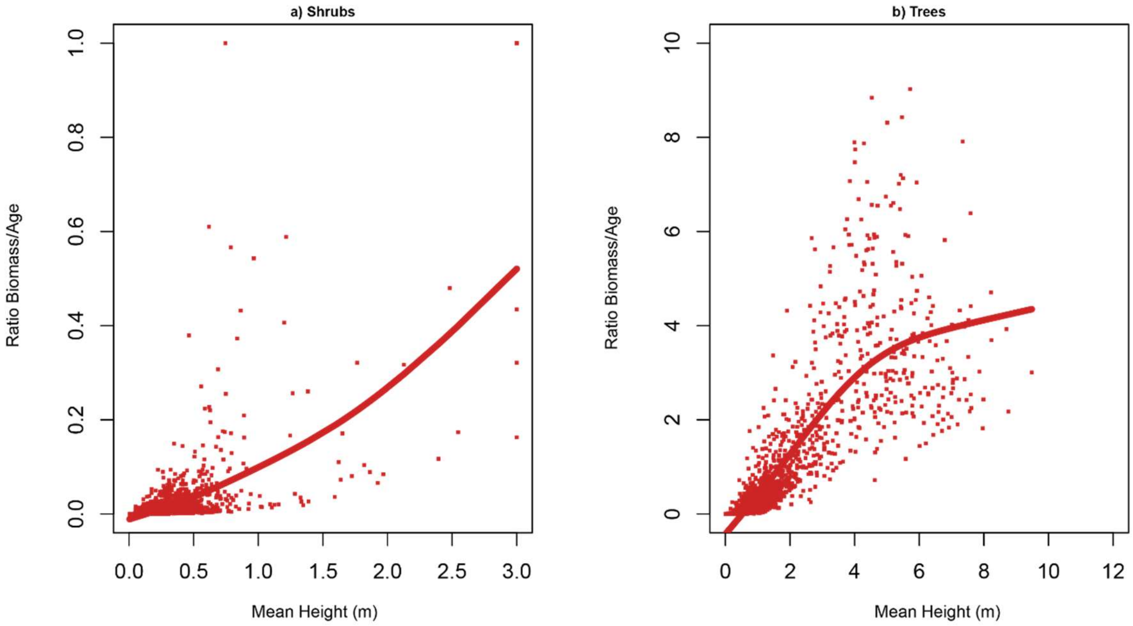


Figure 3.20: Relation between the ratio of biomass accumulated by plants at a given age to the plant mean height (m); for a) shrubs (defined as woody individual with 4 or more stems) and b) trees separately.

3.4 - DISCUSSION

3.4.1 - CHALLENGING THE MULTI-STEM SHRUB MODEL

Model dynamics

MAP has been identified as a major determinant of vegetation community assembly processes (Ma and Frank, 2006; Reed *et al.*, 2009; Báez *et al.*, 2013; He, 2014). However, impact of MAP on competitive relation between woody plants and grasses, and on each of these life-forms, varies with the parameter considered (i.e., canopy area, leaf biomass, number of individuals). For example, relative number of grasses decreases below 500mm MAP, while relative grasses canopy increases. The same is true for fire as a driver of vegetation community assembly. We can expect this to hold true for other drivers. However, while, in terms of number of individuals, woody plants increasingly dominate over grasses with decreasing MAP, this relation is reversed when considering canopy cover. As we cannot conclude on the competitive relation between woody plants and grasses based on one parameter alone, when combining multiple perspectives, we can identify plants key characteristics under different conditions and unravel how competition between life-forms, depending on drivers, shapes these life-forms.

Some parameters react similarly to a particular driver. For example, relative woody canopy area and relative woody leaf biomass decrease at intermediate MAP when fire is introduced in the system. Investigating multiple parameters responses to different drivers can be done by simulating scenarios combining different levels of these drivers and provide useful insights in ecological processes. Simultaneously, analysing parameters responses to drivers can be used to foster model development, creating a positive feedback loop. For example, we observe that, for decreasing MAP, number of woody individual increases while woody canopy cover decreases, which can be explained when considering change in phenotype and not just in plant biomass and change in vegetation community share of each of these varying phenotypes. Therefore, increasing a model's ability to represent life-forms in greater details allows to better understand the relation between ecosystems and environmental factors as well as ecosystem dynamics.

In this regard, we need to be cautious about the interpretation of models results. For example, we observe that introducing shrubs as multi-stemmed woody individuals in aDGVM2 increases competitiveness of woody plant versus grasses with decreasing MAP. However, this shift in competitiveness is relative to number of individuals and we see that canopy cover and leaf biomass exhibit different response patterns.

By focusing only on the prevalence of shrubs relatively to trees in the population, our results show that shrubs become dominant only at low MAP. However, by considering the prevalence of trees relatively to shrubs, our results show that tree prevalence already starts to decrease at intermediate MAP. This indicates that we cannot fully understand the role and strategy of a life-form by considering only one feature of this life-form. We need more holistic approaches to study vegetation community behaviour in response to particular life-form behaviour to fully grasp the strategy of this life-form. This also allows to understand land-cover dynamics. For example, by investigating shrubs, we observe that bare ground cover makes up an increasing share of the landscape at low MAP, while shrub makes up an increasing share of the vegetation population (Figure 3.2, panel b), which implies that shrubs are less negatively impacted by decreasing water availability than other life-forms. Discarding bare-ground, we could have concluded that shrubs abundance increases with increasing water scarcity, however, considering bare-ground shows that shrubs are only relatively more abundant with increasing water scarcity, as other life-forms fitness is more reduced by increasing water scarcity. Thus, analyses should not focus solely on a single object, disregarding its environment, as it can be misleading.

The aDGVM2 was first created with a focus on African savannas. The AGBM map provided by Bouvet *et al.* (2018) focuses specifically on African savannas, which can be expected to be more accurate for this region than satellite derived products having a global scope. In consequence, the better agreement of aDGVM2 results with this product than with other products for selected savannas site corresponds to our expectations. We argue that global scope satellite derived products of vegetation parameters should be used to benchmark simulated patterns while more focused products should be used to benchmark actual values. For example, benchmarking our results to Bouvet *et al.* (2018) increases our model's confidence for savannas as it is its specific focus.

The overall trend of increasing AGBM and vegetation height with increasing MAP holds true for all the satellite derived products used in this study and for aDGVM2 results. However, this effect appears more consistent among satellite derived products than with aDGVM2, irrespective of the method used to analyse aDGVM2 results. Uncertainties are present in satellite derived products. For example, Saatchi *et al.* (2011) display uncertainties going up to a ±50% in some regions. However, it seems that the discrepancy between aDGVM2 and these datasets is mainly due to aDGVM2 not being able to simulate AGBM and vegetation height as high as observations at the lowest MAP values and an overall tendency to overestimate the share of short plants.

Stochastic processes included in aDGVM2 give both flexibility and robustness. The higher the number of replicate simulations performed, the more likely aDGVM2 is to capture a realistic representation of vegetation communities. On one hand, this increases the robustness of the simulated average vegetation; on the other hand, the range of model results allows to explore vegetation diversity. In this regard, looking at average value of 10 replicates for the greater Namib region for which discrepancies were of concern confirms the gap between aDGVM2 results and benchmarking data sets. Simultaneously, it appears that aDGVM2 has the potential to simulate vegetation cover in high agreement with benchmarking data. Thus, for operational use, selecting simulations producing more "realistic" results appears as a viable shortcut, specifically if more precise and site specific climate data can be used to feed our model. Moreover, by investigating the variables responsible for variance in model results could provide useful insights for model development.

The discrepancy between simulation results and observations or satellite derived products depends directly on the parameters considered and indirectly on the methodology to analyse our results. We obtained conflicting results depending on the classification scheme we used. Ratio of shrub in plants population is different when calculated based on number of individuals or biomass or canopy cover. Satellite derived datasets are reclassification (through statistical models) of remote sensing signal based on expert knowledge and selected *in situ* observations, thus they are not direct observations of number of individuals or biomass or canopy cover. Such datasets merge both qualitative and quantitative values in a single product. Analysing aDGVM2 results solely based on individual parameters is still a purely quantitative approach. This conceptual difference and the challenge of holistically characterising aDGVM2 results (i.e., results can appear conflicting depending on characterisation schemes) call for adding qualitative approaches to model development and interpretation. This is what we aim to do when reclassifying our results. However, it does not equate to expert knowledge used to define biomes and to reclassify satellite data as our approach is purely numbers based and does not include qualitative observations, such as species, or landscape features, such as rock fields.

Additionally, it appears that the distribution of bushes, (defined as short woody individuals with low stem count), as modelled by aDGVM2 is closer to the shrub cover displayed by Tuanmu and Jetz (2014), than the shrub cover modelled by aDGVM2, irrespective of the classification scheme used to define shrubs. This indicates a conceptual difference in defining life-forms or trivially a semantic disagreement.

In conclusion, to answer our first question, it appears that aDGVM2's extent of validity, or that of any mechanistic trait-based model, depends on the parameters investigated and the definitions and classifications schemes used to characterise simulations results (or even semantics). This problem challenges our understanding of simulations results and can limit our ability to accurately assess and benchmark them. However, we can use the multifaceted nature of our simulations results to increase the ecological relevance of our results as it opens perspectives breaking common categories while informing model development (Fischer, Rödig and Huth, 2018). For example, it seems that our approach to model shrubs has a weak effect at low biomass, which synergises with the lack of drought avoidance and resistance mechanisms in aDGVM2 to reduce model accuracy in hyper arid regions. This could be counter-balanced by including these drought adaptation mechanisms or by improving representation of the relation between branching, stems and overall canopy architecture (Leaf Area Index (LAI), leaf angle) in regards to hydraulics and light interception (Tyree and Ewers, 1991; Fisher *et al.*, 2017; Trugman *et al.*, 2019).

The role of shrubs

Modelling large scale ecological dynamics and variables requires simplifications of reality, for example by reducing biodiversity to PFTs. However, the assumptions made to simplify reality entail shortcomings (Fischer, Rödig and Huth, 2018). To overcome these limitations, adding complexity to models is necessary. For models based on PFTs, it implies to increase their number. For aDGVM2 and similar modelling approach (i.e., dynamic traits variation) it implies to increase the number of traits, trade-offs and processes which allow for the emergence of life-forms that are currently not represented by the model. Our approach to model shrubs based on mechanistic trade-offs is a step in this direction.

Introducing shrubs in aDGVM2 increases the relative prevalence of woody plants against grasses with decreasing MAP (Figure 3.1, panel a). It means that aDGVM2 simulates relatively more bushes and shrubs at low MAP, where it is still underestimating shrub-land cover. This underestimation is partially explained by the fact that we benchmarked shrub cover to shrubland cover in chapter 2.

In the savannas and woodlands bordering the African tropical forest, introducing shrubs in aDGVM2 decreases simulated AGBM, which increases the agreement with data, as aDGVM2 typically simulated AGBM values above satellite derived observations in these regions. Simultaneously, woody canopy cover decreases in the same area, as expected; however, mean vegetation height does not follow the exact same pattern.

Introducing shrubs in aDGVM2 increases the number of plant individuals with increasing MAP. More precisely it increases the number of grasses and small woody individuals (shrubs and bushes) while decreasing the number of tall woody individuals (trees). Hence, when considering shrubs, simulated canopy in these regions is more open, and aDGVM2 simulates more grassy (and shrubby) savannas and woodlands. The same effect seems to apply to forests, which is unexpected and could require further investigations, albeit shrubs numbers are still low in forests (Figure 2.6).

Many studies argued that fire can contribute to suppress shrubs and reduce their prevalence (Trollope, 1980; Bond, Woodward and Midgley, 2004; Stevens, Lehmann, *et al.*, 2016), while in some ecosystems, shrubs can be adapted to fire (Keeley, 1986). Fire regime shift related to land use change has been identified as a possible explanation of shrub encroachment phenomenon across Africa (Archer, Schimel and Holland, 1995; Roques, O'Connor and Watkinson, 2001). We observe a more complex relation between fire and shrubs. Introducing shrubs in aDGVM2 reduces AGBM mainly in savannas and woodland regions. In this regard, fire has a similar effect compared to simulations without shrubs. As shrubs open up space for grasses, we expect that simulated fire frequency and/or intensity would increase as grasses are the main contributor to fire fuel in aDGVM2. Yet, as the presence of shrubs in aDGVM2 allows for greater grass biomass (compared to simulations without shrubs), fire regime does not respond strongly; which is visible in the fact that the difference in fire frequency with or without shrub does not follow a clear pattern. Therefore, we argue that investigating the inverse relation of shrub effect on fire regimes is required for a holistic conclusion on the effect of fire on shrubs. We argue that this relation needs to be studied not only for sites where shrub encroachment is happening but that more global reviews of this relation are required (Stevens, Lehmann, *et al.*, 2016).

Our step towards more accurate modelling of vegetation communities allows for a previously absent life-form strategy to emerge in aDGVM2 via the inclusion of the stem number trait and its associated trade-off. Simultaneously it shapes our overall model results, making them more realistic, in agreement with hypothesis about model complexity (Fischer, Rödig and Huth, 2018). This partially answers our second question.

3.4.2 - SHRUB DISTRIBUTION DEPENDS ON SHRUBS CLASSIFICATION

Shrub distribution in non-geographic space

Our implementation of a trade-off between height and stem count leads to the emergence of distinct strategies in aDGVM2. Tall individuals are associated to low stem count while multi-stemmed individuals are associated with low height. These strategies overlap at low height and low stem count. From another perspective, we observe that at the lowest MAP values, height per stem is constrained to low value, which implies that the relation between height and stem number is less relevant to discriminate between life-forms at this level, according to the trade-off we use for our model (Figure 2.1). Therefore, the partitioning between shrubs and trees is stronger at higher MAP. The overlap between multi-stem and single stem strategies occurs in the most arid environment, where other criteria need to be investigated to discriminate between these strategies (Turner, 1978; Cushman, 2001; Dong and Zhang, 2001; Schwinning and Ehleringer, 2001; Gorai *et al.*, 2015; Soussi *et al.*, 2016).

Height and stem count can be used to characterise shrubs phenotype (Zizka, Govender and Higgins, 2014; Götmark, Götmark and Jensen, 2016). However, these criteria do not react equally to environmental conditions. Height thresholds have a stronger effect on shrub characterisation at high MAP whereas stem count thresholds have higher effect at low MAP. Moreover, more individuals can be classified as shrubs based on height solely than on stem count solely. We could expect that using multiple criteria to characterise a life-form improves the accuracy of our classification scheme, yet, we observe that the stem count criteria overrides the height criteria. In consequence, we can simplify our shrub classification scheme and rely mainly on stem count (less at higher MAP). This is counter intuitive as shrubs are typically seen as small woody plants, yet it shows that the correlation between height and stem count is high for shrubs and confirms the hypothesis in which our model approach is grounded. Additionally, we argue that, following Figure 3.12, defining a precise and appropriate vegetation classification scheme can provide better insight into our simulation results and improve the relevance of benchmarking aDGVM2 with data.

The relation between height and stem count for shrub classification observed in aDGVM2 holds true in terms of number of individuals. However, considering, for example, leaf or stem biomass or crown area, produces different perspectives on shrub prevalence. In the case of crown area, we observe that, in contradiction with the individual numbers approach, height is a categorisation parameter that has more impact at low MAP. For most discriminating parameters, thresholds have stronger impact at low MAP. We hypothesize that the overlap between shrub and tree strategies at low height and low stem count is the reason for this pattern. The fact that the discriminating strength of various thresholds responds differently to MAP depending on the parameter considered could be used to investigate life-forms and associated strategies. We systematically observe a response of life-forms distribution to classification parameters and thresholds. Therefore, caution is necessary when comparing studies using different classification schemes and taking different perspectives on the same object, specifically for shrubs.

With these elements we conclude on our third question, strengthening our conclusion that shrub prevalence increases with aridity. We also highlight the need for caution when discussing this pattern as classification schemes and parameters considered shape the frame of this observation.

Classification

To interpret simulations results, multiple approaches are available. Individuals can be classified based on different parameters and variables (Estabrook, Johnson and Mc Morris, 1975; Bachmann, 1995), which is relevant for interpreting model's results, as we have shown. To look at larger scale, vegetation cover can be classified into biomes or bioregions based on vegetation distribution map (Olson *et al.*, 2001), and can include satellite derived observations (Tuanmu and Jetz, 2014). It has been argued that increase in complexity (Fischer, Rödig and Huth, 2018) leads to more accurate models. We observe that results are more informative when we increase the complexity of their analysis by refining categories; however, their accuracy relies on conceptual assumptions. Therefore, increase in complexity alone is not sufficient, and need to be associated with relevant conceptual development and classification scheme. In this regard, it appears that, here, a classification scheme based on both macro (population scale) and micro (individual scale) parameters holds the best potential to accurately analyse aDGVM2 results. However, it also appears that classification thresholds are spatially relative, which means that they are not equally accurate everywhere at the continental scale. We question to which extent it is due to either model's limitations, or to regionally different classification of ground observations (for example, if different expert point of views are used to classify vegetation in different regions), or different regional/local vegetation specificities. For example, Zizka, Govender and Higgins (2014) found an average number of 2,2 +/- 1,4 stems for trees. This study was site specific. The site being in an African savanna, we expect the validity of results classified according to Zizka, Govender and Higgins (2014) to be savanna centric. This means that a model (or part of a model) based on a site-specific study is more likely produce more accurate results in regions similar to the site of the reference study. However, the characterisation of regions similarities requires caution as it can also be parameter/trait relative. Alternatively, it appears possible to directly benchmark the population level parameters instead of categorised vegetation with products such as provided by Song *et al.* (2018).

Additionally, we observe that the more complex the classification system, the more continental vegetation cover appears as a mosaic. This could be seen as contradicting the large scale biomes/bioregions maps (Olson *et al.*, 2001). While obscuring this mosaic large scale maps can also integrate finer mosaics by relying on more precise classification schemes (such as Mucina and Rutherford, 2006).

We model shrubs based on height and stem count. Based on this assumption, we investigated how classifying shrubs based on height and stem count traits, shapes our understanding of this life-form. Precisely, we investigated how classification schemes shape shrub distribution patterns and how it shapes model benchmarking with existing datasets. It appears that other traits, implicitly (Tuanmu and Jetz, 2014) or explicitly (Carlson, Adams and Holsinger, 2016) considered, can also shape shrub distribution pattern. For example, plant age strongly impacts simulated mean stem count at low to intermediate MAP while it impacts mean height at high MAP; which means that excluding seedlings from our results leads to an increase in relative shrub cover at low to intermediate MAP and a decrease at high MAP. When such trait effects are explicitly disclosed we can consider them in our benchmarking process; yet, studies typically consider only a few key traits depending on their questions (Killingbeck and Whitford, 1996; Martínez-Cabrera *et al.*, 2011; Carlson, Adams and Holsinger, 2016).

Plant age and maturity are particularly of concern to our studies as age variation mainly impacts woody individuals at low to intermediate MAP, i.e., impacting the savanna region. Additionally, across sub-Saharan Africa, aDGVM2 simulates a disconnection between young and old populations. It simulates only few populations of intermediate age, showing that cohort dynamics (Oliver and Larson, 1996; Scholes, Bond and Eckhardt, 2003; Boehmer *et al.*, 2013) could have an impact on aDGVM2. This cohort effect could be linked to the model initialisation conundrum (Moncrieff *et al.*, 2015). It could also be linked to the model stochasticity, since at similar MAP some simulated grid cells have a low mean age while others have a high mean age.

The relevance of the classification scheme used to define shrubs depends on the regions considered. Thus, it could be possible to use aDGVM2 for site- or region-specific studies with higher accuracy by choosing a relevant classification of its results adapted to the context. However, benchmarking overall model behaviour will remain challenging until global datasets are available which can be directly related to aDGVM2 results. This means that they need to be conceptually equal and not just similar. For example, shrub cover as modelled can be compared to satellite derived shrub cover, yet they are conceptually different, which limits how informative such comparison can be; limitation reinforced by the differences between similar products (Tuanmu and Jetz, 2014; Gaillard *et al.*, 2018). Due to this limitation, models need to be benchmarked to multiple products and observation simultaneously; yet, as discussed above, results extent of validity appear different depending on the parameters and thresholds considered.

In conclusion, definitions of life-form categories shape our understanding of these life-forms to various extents depending on classification scheme (i.e., parameters and thresholds). Thus, there is no final answer to our fourth question; precise analyses can be made only for specific cases, implying that assessment of models results, and extent of validity need to be as holistic as possible, yet generalisation should be avoided and results need to be explicitly framed.

Shrubs footprint

In order to better frame the extent of validity of models results it might be useful, or even necessary, to look at the imprint of the studied object. Looking at an object is useful to analyse it, however, it is not sufficient as it is always part of a context, specifically in the case of ecology where both biotic and abiotic processes are at play. For example, by looking at the ratio of bare ground (i.e., “dead cells”) simulated by aDGVM2, we observe that, at low MAP, most of the simulated space is actually bare ground. This limits the relevance of any result analysis for this region as statistical validity decreases with the number of plant individuals concerned. In such case, we can conclude that simulations below a set MAP threshold are less relevant and could be discarded, and/or that further model development is required to represent ecosystems dynamics in this context. We can add to our fourth question that characterising the environment of a life-form, both biotic and abiotic, refines the definition of this life-form and thus improves the accuracy of its study.

3.4.3 - SHRUB LIFE-FORM *MODUS VIVENDI*

Resprouting *vs.* reseeding

The aDGVM2 simulates a non-linear relation between carbon allocations to storage and to reproduction. This indicates that many simulated woody individuals are allocating relatively more to either storage or to reproduction. Thus, a division emerges in aDGVM2 between resprouters (allocating more to storage) and reseeders (allocating more to reproduction). The presence or absence of shrubs and/or fire in our simulations does impact this relation. Fire appears as necessary for discrimination between resprouters and reseeders to happen. Shrubs appear as strongly determined resprouters. Our finding concurs with the literature (Canadell, Lloret and López-Soria, 1991; Midgley, 1996; Enright *et al.*, 1998; Cruz, Perez and Moreno, 2003; Knox and Clarke, 2005; Reyes, Casal and Rego, 2009). However, literature also points at the role of nutrient (Knox and Clarke, 2005; Dybzinski *et al.*, 2011) in explaining plant carbon allocation strategies; which are absent from aDGVM2. It could explain why we do not observe stronger clustering of considered traits to discriminate between resprouters and reseeders. Moreover, while strong resprouting may be common in savanna shrubs, many shrubs do not typically resprout after fire (Paciorek *et al.*, 2000; Pausas *et al.*, 2004), potentially due to environmental drivers, such as fire frequency, intensity and seasonality (Olson and Platt, 1995; Wright and Clarke, 2007) and nutrients and water availability (Cruz *et al.*, 2002; Knox and Clarke, 2005). Additionally, fire impact on seed bank could contribute to discriminate fire adapted strategies among non-resprouters (Kelly and Parker, 1990; Enright *et al.*, 1998). Hence, our approach to modelling shrubs is relevant to simulate and test how fire management strategies can interact with shrub encroachment scenarios in African savannas, but further development are required to consider other regions globally (Driscoll *et al.*, 2010). Our shrub model can also be used to investigate ecosystems response to drought as resprouters might be more adapted to cope with increased drought stress in the future (Zeppel *et al.*, 2015).

Rooting depth strategies

Similarly to shrubs, it has been observed that plants have deeper root systems in tropical dry savanna compared to tropical humid savanna (Schenk and Jackson, 2002b), yet globally, rooting depth increases with MAP for grasses (Schenk and Jackson, 2002a). The trend of increasing root to shoot ratio with increasing aridity has been documented in many studies (Wilson, 1988; Chapin, Autumn and Pugnaire, 1993). The aDGVM2 agrees partially with these findings. It simulates an inverse relation for grasses. For trees and shrubs it simulates a similar relation below *ca.* 2000 mm MAP. The aDGVM2 simulates shrubs as mainly resprouters and trees as being able to be either resprouters or reseeders, but it simulates no direct correlation between shoot to root ratio and resprouting *vs.* reseeding. However, many resprouters store carbon preferentially in their root (Pate *et al.*, 1990), leading to higher root to shoot ratios for resprouters (Verdaguer and Ojeda, 2002). However observations relate this ratio to aridity while we measure it relatively to MAP (Figure 3.19), which are related parameters but not equal (Wallén, 1967; Sahin, 2012). Therefore, comparing aDGVM2 results based on MAP to observations based on aridity is challenging. Wind, solar radiation, ground cover and soil texture all play a role in defining aridity (Spinoni *et al.*, 2015). They can impact rooting strategies through their impact on aridity, but also directly. For example, in sandy soils a larger proportion of roots is present at greater depths than in loamy and clayey soils (Jackson *et al.*, 1996; Schenk and Jackson, 2005) and rather than shifting maximum rooting depths downwards as texture changes from fine to coarse, plants in water-limited environments may shift the zone of maximum root activity downwards in the profile. These factors are considered to different extents by aDGVM2, but further model developments and analyses are required to comprehensively test their effects. Additionally, correlated factors (e.g.: nutrients, soil texture, solar radiation) are intertwined with aridity in regard to shoot to root ratio. For example root to shoot ratios decrease in soils richer in nutrients, where plants can attain higher productivity and higher biomass (e.g.: Schenk and Jackson, 2002b).

Carbon strategies: shrubs *vs.* trees

Given the relation we used to implement our trade-off between height and stem number we expect its effect to be greater at higher biomass, and height. Additionally, our hypothesis implies that shrubs should grow faster than trees, relatively to their height (Wilson, 1995) and that mature trees accumulate more biomass in absolute value. We indeed simulate this last point, but it does not allow us to conclude on trees and shrubs carbon allocation strategies as shrubs do not get as tall as trees due to their definition. As our trade-off prevents shrubs from becoming tall, it also prevents them to accumulate as much biomass as trees. To properly measure the effect we expect from this trade-off, we would need to investigate plant growth rate per biomass (e.g.: Stephenson *et al.*, 2014), which implies to study further mechanisms, such as carbon maintenance costs, respiration rates, structural carbon architecture constraints (McCree and Troughton, 1966; Merino, Field and Mooney, 1982; Ryan, 1991; Nobel, Alm and Cavelier, 1992), which is out of our study scope, but have the potential to be studied with aDGVM2.

With these observations, we can conclude on our fifth question. We observe that shrubs, contrarily to trees, are strongly determined resprouters, while following similar phenotypical relation to MAP (decreasing shoot to root ratio with decreasing MAP below a threshold). However, fully assessing their global carbon strategies will require further analysis, benchmarking and model development.

3.5 - CONCLUSION

The non-geographical distribution of shrubs and their role in vegetation community follows our assumptions, specifically in African savannas and woodlands. We specified the frame of our model's validity (e.g.: we further differentiated our modelled shrubs from bushes) and highlighted opportunities for further refinements. We argue for better consensus about vegetation categorisation. We further argue for an increasing reliance on traits and not on categories, as disagreements or agreements can be made apparent depending on the parameters and thresholds considered for classifications.

We show that complexity does not need to be high to characterise life-forms when analysis criteria are well defined. Additionally, we observed that it might be relevant to investigate parameters *a priori* not in scope with the study focus (e.g.: age). Studying an object (here, shrubs) can be achieved or enhanced by looking at this object imprint and not the object itself (e.g.: focusing on bare-ground cover at low MAP improves our understanding of vegetation dynamics in these regions). Simultaneously, high complexity is necessary at continental scale to consider the diversity of land-covers.

Our shrub model, grounded on a mechanistic trade-off allows for a greater complexity of plant strategy to emerge in aDGVM2. Shrubs are simulated as a mainly resprouters strategy without being explicitly defined as such *a priori*. One to one comparison of carbon strategies between trees and shrubs are not always feasible as they occupy distinct trait space, yet we show that one trait (stem count) has an impact on the whole plant life-form and associated strategy.

In conclusion, from a mechanistic trait-based trade-off, distinct life-forms and associated plant strategies emerge in aDGVM2 from trait filtering processes. This drives the general vegetation communities' dynamics towards higher realism and accuracy. The multi-factorial aspects of our results call for caution when generalising observations and argue for more holistic model benchmarking.

CHAPTER 4 - CHALLENGES AND OPPORTUNITIES OF MODELS INTEGRATING TRAITS AND PROCESSES TO SIMULATE ECOSYSTEMS STATES AND DYNAMICS

This chapter is the work of the sole author of this thesis.

ABSTRACT

Background

Assessing DGVMs is an extensively discussed topic among its research community. Data used to create, parameterise and benchmark models has uncertainties of its own. Thus, benchmarking a model is a process intermingling two different set of questions, and two specific perspectives: modelling and data/observation. In regards of these considerations, we ask: How can we relate data and simulations? Can we re-think this relation to ensure that simulations are effectively informative and to which extent? How can we make models more informative?

Objectives

We discuss and assess how to better connect models and observations. In this order, we propose to answer the following questions:

1. How are assumptions underlying model structure and operation limiting models' relevance?

2. How are the limits to data and models interacting and impairing model benchmarking?

3. Are there limits to modelling approach which cannot be overcome and why?

4. Which limits to modelling can be overcome and how?

Methods

As we extensively analysed and benchmarked aDGVM2 and specifically its shrub model, we have been able to gain knowledge about shrubs life-form and strategy and more generally about ecosystems processes and biogeography across African savannas. We will use these insights as baseline to go further and consider the broader picture to answer our questions. We first discus considerations specific to models, then specific to data and finally we investigate their interaction. We then propose a perspective on how to overcome challenges faced by model development and take modelling of vegetation communities to the next level.

Results

It appears that most current limitations to models can be solved through extensive research effort in terms of model development and benchmarking but also by the availability of further observation data. Given the scale of this effort and given the potential of integrated research projects between data providers and model developers, we argue for greater cooperation between research groups. This implies closer interactions and more coordinated research efforts. However, ultimately some limitations are inherent to modern modelling approaches. We need to fully consider these limitations to accurately and properly interpret models' results. Finally, we discuss opportunities to step beyond current DGVMs' limits.

Conclusions

There is room for improvement to model processes and biodiversity. Computing power and big data can be a part of it as brute force solutions. However, they cannot answer every question and the development of process and trait-based models to investigate mechanisms shaping ecosystems offer greater and smarter opportunities to step up models' performances.

4.1 - INTRODUCTION

Model and simulation are concepts with extensive records of discussions and debates, crossing many disciplines over centuries (Baudrillard, 1981, 1988; Gleiniger *et al.*, 2008). The use of mathematical models in ecology, biology, earth science and biogeography pre-dates their digital versions (Volterra, 1926; Lotka, 1927). However, the development of present day models came with the emergence of digital computers, such as the ENIAC, allowing to solve complex numerical problems (Shrukin, 1996). The increase in computing resources availability allowed for increasing models' complexity and explanatory power. Over the last decades, we saw an exponential development of such models, the list of which being too long to do them justice in a few lines (Prentice *et al.*, 2007; Fisher *et al.*, 2017). Among them and of concern to us are the DGVMs (Prentice *et* *al.*, 2007). However, despite a rapid increase in the number of models (Fisher *et al.*, 2017; Trugman *et al.*, 2019), models versions (Smith, Prentice and Sykes, 2001; Hickler *et al.*, 2006; Scheiter and Higgins, 2008; Scheiter, Langan and Higgins, 2013; Longo *et al.*, 2019) and modelling studies (Hickler *et al.*, 2005; Scheiter and Higgins, 2010; Bachelet *et al.*, 2017; Langan, Higgins and Scheiter, 2017; Pfeiffer *et al.*, 2019; Kumar *et al.*, 2020), their fundamental assumptions could benefit from more intense discussions (Song, Zeng and Zhu, 2013; Song, Zeng and Li, 2016; Fisher *et al.*, 2017; Trugman *et al.*, 2019).

We assess key questions contributing to the foundation and development of process based and trait-based models in ecology. We investigate how models are built and what their defining concept and fundamental structure entails in terms of relevance and accuracy. First, we discuss the very definition of "model", as a word definition frames the item which it refers to (i.e., the validity of a concept is limited by its definition). It means that the definition of what is a model has technical consequences for the development of DGVMs. We discuss this on a conceptual level and then relatively to our modelling approach. Second, we discuss how the way we build models has direct consequences on models' validity. Both the assumptions made to design a model and the observations and data it relies on will have consequences on its validity. Third, we discuss the entwinement of data and models concepts and structure. Finally, we discuss model development challenges and opportunities.

4.2 - MODEL FOUNDATION

4.2.1 - LIMITATIONS INHERENT TO MODEL CONCEPT

Models are a representation of reality, which means that they are a simplified version of reality, yet, they are intended to look like and function like reality. Thus, they represent (look like) and imitate (function like) reality. DGVMs are specifically useful as they combine these two traits, which allow us to handle, in a practical and controlled way, the subject of our study. It means that to create a model requires to define what is represented and, thus, to make assumptions. These assumptions define what is not in the model and what is included. Including an item in a model can be done in different ways and for different reasons. What is considered, how and why, frames the extent of validity of a model. We discuss the extent of validity of a model and, inversely, its limits.

Because a model is a simplification of an object, it is typically limited by what it does not include but which is part of the item it represents. For example, Quillet, Peng and Garneau (2010) reviewed limitations to modern DGVMs. Carbon and nitrogen cycles, competition, land-use, land-use changes and disturbances, among others, are typically missing. In a more recent review, Fisher *et al.* (2018) identified the lack of underground competition representation, for water and nutrients among the major limitation to models validity at present. This limitation goes along the fact that below-ground studies of plants are more difficult than above-ground studies of plants. Research towards addressing these limitations and others is intense and multi-faceted. For example, nitrogen cycle is implemented in an increasing number of DGVMs, following different approaches (Quillet, Peng and Garneau, 2010). There are attempts at modelling plants migration (Sato and Ise, 2012). Land use is also considered in some modelling studies (Bachelet *et al.*, 2017). Additionally, processes can be represented through different conceptual approaches, potentially conflicting (Wang, 2003; Fisher *et al.*, 2017). Besides these limitations, there are steps into a new generation of DGVMs. For example, Sato, Itoh and Kohyama, (2007), Hickler *et al.* (2008), Scheiter and Higgins, (2009), introduced representation of the processes controlling vegetation demographics. However, most modern DGVMs are based on PFTs (Hickler *et al.*, 2006; Clark *et al.*, 2011; Yang *et al.*, 2015) which do not account for intermediate life-forms and are thus not as flexible as fully trait-based models. Considering the full trait range is a promising approach to further models development (Verheijen *et al.*, 2013). Some models already explore territories beyond the PFT concept, such as aDGVM2 (Scheiter, Langan and Higgins, 2013), and others (Pavlick *et al.*, 2013; Sakschewski *et al.*, 2015) which propose to overcome the limitations related to the PFT concept by considering trade-offs as the basis of trait filtering.

Despite the continuous refinements of ecological models, DGVMs still lack explicit representation of many trait-based processes. DGVMs, considered altogether, represent a wide range of processes, yet, each represent only a fraction of these processes. Typically, models are created with a solid basis to demonstrate a modelling approach and/or to study a specific question or a specific region. They are then further refined to answer more specific questions and will implement processes specific to the given questions. However, the complexity of a model increases with the number of parameters it considers, which might lead to problematic developments or technical challenges (Fisher *et al.*, 2017), thus, it might prove more practical to fork a model into different versions. Given that models have different foundations and different development history they are more or less suited to represent specific regional vegetation processes and communities. Therefore, despite all DGVMs being benchmarked against data and simulating large scale trends and patterns reasonably well, they produce diverging results at the regional scale, as can be seen in studies using multiple models and in inter-comparison projects (Li *et al.*, 2016; Restrepo-Coupe *et al.*, 2017). Each model considers differently parameters or processes which could locally be major drivers of vegetation community. Moreover, the potential interaction between missing parameters, or between a missing parameter and a modelled parameter, might be an additional indirect limitation.

Stochasticity

Many DGVMs use randomization and stochastic computation (e.g.: Hickler *et al.*, 2008; Scheiter and Higgins, 2009; Pavlick *et al.*, 2013; Scheiter, Langan and Higgins, 2013). Stochasticity is used to approximate and account for parameters or mechanisms not explicitly considered in a model, for which no direct causal relations are drawn; either due to lack of knowledge, data or precision about a process or because explicitly simulating this process would be out of model scope or irrelevant at model scale. Typically, stochasticity represents processes which need to fluctuate in a model, but which cannot be fully mechanistically represented or for which no sufficient observation data is available. For example, in aDGVM2, during recruitment process, new individuals are randomly selected from the seed pool. By doing so, it does not need to represent the processes driving seed selection. It can avoid doing so without losing explanatory power as they can be stochastically represented at its scale even if some or all of its components might be deterministic. However, in aDGVM2, the number of seeds produced by an individual impacts the probability that one of its seed will be selected at random from the seed pool. Fisher *et al.* (2010) demonstrated that variability in demographic processes lead to major uncertainties, even when results are constrained. This is also what we observe with aDGVM2's results (Figure 3.4 and Table 2.1). Stochasticity and randomization used in DGVMs is typically a source of variability in demographic processes, and, hence, source of uncertainties which limit the validity of a model. It is not possible to build DGVMs without any stochasticity being involved. Reducing its share by incorporating more processes in models means reducing the share of natural processes that are not explicitly represented and which need to be approximated through randomization. Increasing the number of processes and mechanisms represented by a model would broaden the range of questions it could answer and the context for which it would be relevant, yet, increasing the computation power it requires.

4.2.2 - MODEL BUILDING

Models built after observations of reality embed them on two levels. One level is relative to the assumptions made as baselines for the model. It represents the scope of the model, what is included in the model, and it represents the definition of the items that are considered and how they are represented. Another level is relative to the data used to represent these assumptions. Equations used to calculate processes are derived from observations and parameters values can be constrained by observed value range.

Assumptions

The assumptions made to design a model are the conscious and desired simplification of reality. Each model is a unique concept as the sum of a unique set of assumptions, which are also the first source of its uncertainties and limitations. First, the model builders must decide which theories it relies on, or it wants to challenge, and it shapes how reality will be represented in the model. It can be theories about how to organize and structure an understanding of nature (DeAngelis and Waterhouse, 1987; Lavorel and Garnier, 2002). It can also be theories about what the drivers of ecosystems are. For example, assuming root competition to be a major driver of the grass-tree competition in savanna (Walter, 1971) or assuming that it does not play a major role (Scholes and Archer, 1997; Scheiter and Higgins, 2007) can imply two different model architectures. In the first case, it is necessary to include mechanisms through which roots can compete for space and access to water and nutrients; not in the second case. Similarly, the focus of the study drives the processes and parameters to be included and how. For example, there is a need to include cold adaptation mechanisms when studying tundra ecosystems, which is irrelevant to study tropical rainforests. *In fine*, it is the builder perception and understanding of the model which sums up these necessary simplifications, introducing a bias. Indeed, model builders have knowledge and skills which cannot encompass every single detail and it shapes how they prioritize different items for their models and studies. In resume, assumptions limit models validity *via* the interconnection of three levels: the theoretical framework, the scope of what is considered and the model builders' perception and prioritization. These three points shapes the type of model created. For example, van Oijen *et al.* (2018) identifies three types of models: ecological, process-based and integrated, to study climate change impact on biodiversity and productivity of grasslands; whereas Fisher *et al.* (2018) refers to ESM (Earth System Models) which include first and second generations of DGVMs, further sub-divided into forest gap models, and vegetation demographic models, again sub-divided in individual-based models and cohort-based models. Each type of model has inherent strengths and weaknesses shaping its relevance.

Many models, built differently and using different variables and parameters can produce a similarly fair representation of reality (Guisan and Thuiller, 2005; Todd-Brown *et al.*, 2013; Li X. *et al.*, 2016; Li F. *et. al.,* 2019), yet they can diverge when simulating future scenarios. This fact questions how meaningful the differences and similarities between these various approaches are. The possibility to create models so different yet all showing good agreement with reality could be due to the interactions between assumptions inside each model. Indeed, models built to answer specific questions and/or to study specific contexts can be precisely tailored. There are two consequences to that. First, the broader the question and/or context, the more complex a model needs to be to exhibit similarly accurate representation of reality compared to a model with a narrower focus. Second, using models outside of their designed scope should be considered very carefully.

Finally, the development of models and their quality rely on data availability to inform processes simulation. Thus, lack of data is a severe limitation, for example on the following key processes: sub-annual tree growth, dynamic biomass allocation, drought effects (Zuidema, Poulter and Frank, 2018). As such, a model can focus on interesting, relevant and meaningful questions, processes and topics, only as long as data can feed the model development as basic bricks.

Data context dependency and constraints.

As data availability is critical to support models development, data quality is equally critical to models (Li, Brimicombe and Ralphs, 2000). Even assuming data availability of the highest technical quality, their nature and their use can introduce additional biases.

While assumptions frame the outlines of the model, the data used to flesh out these frames are themselves constrained by the context from which they come from. Indeed, observations are made for a given point in time, a given region in space and with a focus on specific items (study subject and parameters measured). These three points directly limit the extent of observations validity. Specifically, ecological and biological processes have different weight in different contexts (Zaehle *et al.*, 2005). Using larger scale data, where context is less specific and/or using multiple data sources are ways to reduce this effect (Johnson and Cox, 2019). However, and despite increasing data availability, the most complex models have to rely on patchy material. The incompleteness of data implies that models will fill up the holes in these observations. For example, a model of tree distribution might use observations of several representative species of different bioregions to model global tree distribution.

Scaling up context dependent observations to broader contexts, up to global scale, is problematic in itself. Li *et al.* (2018) showed that models results change depending on which study is used as reference for parameterisation. However, none of the tested studies are incorrect, or lacking *per se*. It simply means that the effect of scaling up local observations warps the meaningfulness of these observations; i.e., models are more accurate for the context they have been designed for. Therefore, scaling up local studies to continental or global scale produces more or less valid results depending on how close local studies are to the larger scale average. This implies that parameterising a model according to the site-specific study which is the closest to the large-scale average should produce the best fit between the model output and observations. However, it does not entail a necessary good representation of local spatial heterogeneity and it does not mean that the values and patterns obtained are accurate or even correctly grounded. Because of these two reasons, a model ability to extrapolate (out of its structure and baseline data range), test scenarios and produce meaningful forecast is still limited. In conclusion, being globally valid based on local studies without having an understanding of mechanisms and an accurate model of processes interactions impairs our ability to test scenarios and forecast potential futures.

Data regression

Data is selected as reference to build models according to its relevance regarding modelled processes. To be able to use this data in a model it needs to be processed into mathematical objects. Clouds of observation points are not typically directly embedded in a model. Observation data can be transformed into either equations (e.g.: through regression or interpolation of data) or into range of parameter values, or into constants. This transformation entails a loss of information. While regression equations used by a model "represent", or capture, the whole range of data, actual model "values" are constrained to those given by the regression equation; thus, discarding all other values from the dataset, while potentially creating artificial values not corresponding to actual observation points. The strength of this effect depends on the dataset and the regression method (Figure 4.1). To dampen this effect, it is possible to either select more homogeneous data or to consider additional data to have a more globally relevant dataset. The trade-off between loss of explanatory power and context constraints should be carefully evaluated.

The degree of freedom characterising the regression of data, implies a trade-off between accurate representation of data and the relative weight of stochasticity in the model. As the degree of freedom increases, to capture an accurate picture of trait space requires an increasing number of replicate simulations (impacting computing resource requirement). This is further emphasized by the interplay between equations with varying degrees of freedom, which affects the relative weight of respectively represented processes.

Each DGVM is made up of a unique set of equations, parameters range and constants, coming from different context dependent studies. Despite common points between various models, the unique combination of sources they represent can contribute to explain why different models might perform better on some cases than others even if they include the same processes, due to stemming their equations from different contexts. For example, a model including a representation of fires based on dataset A might perform better at representing fires in region A than a model simulating fires based on dataset B.

Using observation data to build a model is necessary to represent reality. However, it introduces two sources of bias in a model. First, observations are context dependent. Second, transforming data to include it in a model reduces the realism of this data. By considering the trade-off between these two sources of bias, they can be mitigated; yet, they cannot be completely negated and should be explicitly considered when assessing the accuracy and relevance of a model.

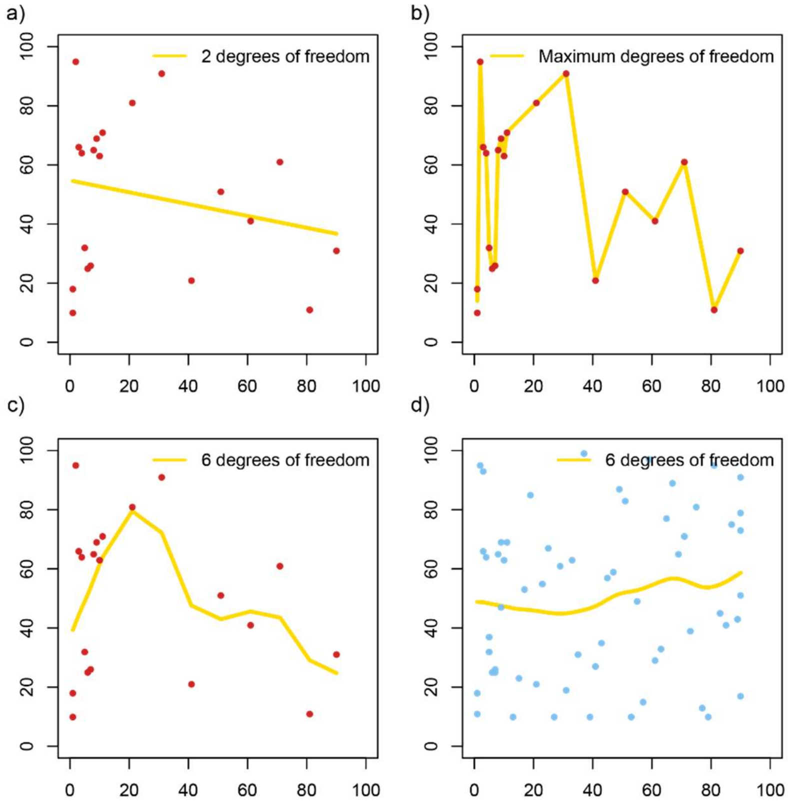


Figure 4.1: Representation of limitations inherent to data regression/interpolation. Red dots: small trended data set (random values). Blue dots: large untrended data set (random values). Yellow lines: regression lines. a) 2 degrees of freedom for regression, b) maximum degrees of freedom for regression (i.e., as many degrees of freedom as there are data points), c) 6 degrees of freedom for regression, d) 6 degrees of freedom for regression.

4.2.3 - DATA GATHERING AND RE-ANALYSIS LIMITS - DATA QUALITY & DATA CONTEXT

Models rely on data, which introduce biases and limits relatively to their context dependency and through the data translation into models’ language. These limitations directly concern model developers, yet there are other potential limitations related to data about which modellers have little direct control. Concerns about data quality and about data creation and gathering is an intense debate (Borgman, Wallis and Enyedy, 2007; Costello *et al.*, 2013; Kosmala *et al.*, 2016), and the reliability of observations is of critical importance for modellers, impacting both the model building and model benchmarking phases of model development.

Data quality might not be assessed by modellers due to two main reasons. First, they might not be expert in the tools and methodologies used to create these data and, thus, cannot assess their quality. Second, it might be that there is no full disclosure of the data gathering process, which can be for regulatory reasons, but also for sociological reasons (Leonelli, 2017). In these cases, data quality might introduce bias in the model, but it is out of scope for the modellers. However, when data is provided with a methodology which can be understood by the modellers and with a full disclosure of their strength and weaknesses (e.g.: Saatchi et al., (2011)), modellers can and should consider such limitations.

Data availability increases (Borgman, Wallis and Enyedy, 2007), which comes with its own challenges and opportunities (Michener and Jones, 2012). For example, increasing the number of samples and sources can mitigate gaps in data due to statistical effects. However, this can be invalidated if the overall data quality decreases among the selected sources. Evaluating such conundrum is a challenge for modellers.

Even if modellers are not experts in the specific methodologies used in a given research field to gather data, debates in this field (Hanan *et al.*, 2014; Staver and Hansen, 2015; Gerard *et al.*, 2017) and expertise from collaborators in this field can be enlightening to discuss the relevance and accuracy of data sets for ESMs and particularly DGVMs. It is also possible to assess the quality of datasets without expertise in the field, either if full disclosure of data uncertainties is available (Saatchi et al., 2011), or if multiple datasets are available and can be compared (Table 2.1) (Avitabile *et al.*, 2011; Saatchi *et al.*, 2011). In such example (Saatchi et al., 2011), if global dataset quality is deemed insufficient to be used for specific modelling questions, it might be necessary to focus on site specific studies. However, it requires more precise climate data, which might be unavailable. Focusing on site specific data brings us back to the context dependency limitation which we discussed in the first section of this chapter.

4.3 - DATA: BENCHMARKING AND INITIALISING MODELS

When building models, such as ESMs and DGVMs, the relevance and quality of the model is insured via two components. One is reliance on evidences and theories; deemed correct and valid until proved otherwise. This component represents the basic bricks of models: equation and constants derived from observations. The other one is the model benchmarking, which compares its results to observed variables, preferably different from the ones fed to the model as bricks. However, adding valid bricks to a model and controlling its output does not necessarily imply that the model structure is correct and that the interplay of its basic bricks functions as intended. This is a critical part, often overlooked, for which new control methodologies need to be developed.

Models are reaching a level of complexity and refinement that is challenging and controlling all the details of a model is increasingly unwieldy. Assessing the extent of validity of a model can be done *via* benchmarking methodologies and simulations analyses, which are thus at the core of model development. Consequently, how data and model output can be related needs to be properly understood in order to prevent introducing further biases and limitations. We investigate this conundrum to identify key connections between data and models. We propose arguments and ideas about how to bridge the conceptual and practical gaps between data and models, in order to better understand and further develop more realistic models.

4.3.1 - MODEL INITIALISATION

Theoretically, if a model is initialised with observed data for all its parameters (i.e., at the beginning of a simulation, it has all its variables set to observed values), it should produce the best fit it can with benchmarking data, if benchmarking data is the same as initialisation data. In this case, divergence between model output and data would directly reflect the shortcomings of the model. In parallel, the more refined and correct a model, the closer its results to data when initialised with this data. Additionally, if benchmarked to extra data (B) not included in the initialisation (different parameters or different context), the divergence with this data (B), minus the divergence to data (A) used for initialisation, would reflect the processes lacking to represent parameters specific to extra data (B). Finally, the more complex a model, the more likely its results are to be closer to data not used for its initialisation (because it is more likely to include approximations of the missing parameters or context specific processes).

The more a model is supposed to fully represent reality, the closer it is to be reality and thus the more complex it is (given that models are a simplified version of reality). As trivial as it sounds, ultimately, only reality could be reality. However, the point of a model is to represent as much of reality as possible with as little complexity as possible. This implies to consider only a very specific and constrained subset of reality in order to produce good results without being too complex; at the cost of having limited capacity to extrapolate. Thus, we can pinpoint a trade-off in model development. For a given level of complexity, aiming at explaining as much of the variance as possible (i.e., having the best fit possible with benchmarking data) implies to be more context specific. Conversely, trying to be as global as possible, and thus less context constrained, implies a reduction of the fit with benchmarking data locally. To maximize the fit to observations and the extent of validity implies an increase in model complexity.

To initialise the parameters for which no data is available, models typically assign to them, random values taken from a realistic range. The reliance on stochasticity depends on model complexity and scale simulated. In aDGVM2, the random initialisation of the model is a major driver of the model output. Indeed, replicate simulations with identical initialisation produce nearly identical results (Results not shown) compared to replicate simulations with different random initialisations (Table 2.1).

If we assume a model to be correct, when comparing multiple simulations with different random initialisations, the one producing results with the best fit to benchmarking data is likely to be the one which has assigned values to non-data-driven initialised parameters as close to reality as possible. This relation could also be used to identify the actual values for these randomly initialised parameters through reverse engineering.

4.3.2 - ECOSYSTEMS: TRANSIENT OR EQUILIBRIUM

When we benchmark a model, we compare its output to observed values. Assuming that this comparison allows us to actually assess the relevance and accuracy of a model implies that we observe an ecosystem which is in equilibrium. This implication comes from the fact that models such as DGVMs typically produce results akin to an ecosystem in a dynamic equilibrium state. However, while the dynamic equilibrium concept has been widely used to describe ecosystem dynamics (Tuljapurkar and Semura, 1977; Winterhaider, 1984; DeAngelis and Waterhouse, 1987; Skarpe, 1992; Friend *et al.*, 1997; Svirezhev, 2000; Quillet, Peng and Garneau, 2010; Laliberté *et al.*, 2013), there are evidences about ecosystems being in a transient state (Van Minnen, Leemans and Ihle, 2000; Spagnolo and La Barbera, 2002; Hastings, 2004; Frank *et al.*, 2011). Historic dynamics can be a key factor to explain present ecosystem state, either through long term impacts (Marshall *et al.*, 2018) or through natural selection processes (such as coevolution, e.g.: Brooks (1990)). Thus, trying to model present ecosystem states based on context dependent observations might prevent from improving models' ability to extrapolate as it could lead to simulating vegetation cover agreeing with observations while misrepresenting ecological processes; particularly when dynamics timescale, amplitude and direction are unknown.

Observation data, being context dependent, cover a limited time-period. Consequently, data might not cover a time period commensurate with the dynamic equilibrium time scale of an ecosystem. Indeed, many studies are only a snapshot of an ecosystem (e.g.: Kelly and Parker, 1990; Schleicher, Wiegand and Ward, 2011; Báez *et al.*, 2013). At ecological scales, even a decade of data might be only a snapshot (Thompson, 1998; Hastings, 2004), and many datasets are made up of observations spanning only a few days to months. Only a limited number of studies provide observations for over a decade (e.g.: Laliberté *et al.*, 2013; Lawley *et al.*, 2013; Stevens, Erasmus, *et al.*, 2016). Therefore, we typically assess models based on "points" in time, relatively to typical ecological time scale. To benchmark models, long term observations are required to grasp ecological time scale trends and dynamic equilibrium range of state parameters. As observation timescale is typically a limiting factor, it is challenging to assess if a model simulates dynamic ecosystem equilibrium, since it might be close to the real ecosystem equilibrium while being quite far from available observed values constrained for a specific point in time (Figure 4.2). Additionally, identifying at which point of a dynamic equilibrium observations are made is challenging. Figure 4.2 shows that while model results might be close to benchmarking data for a time point it does not allow to conclude on the agreement with the ecosystem equilibrium through time. Therefore, when data representativeness of dynamics through time is limited, we should not conclude on a model's ability to represent the pattern and amplitude of the actual ecosystem equilibrium, irrespective of model agreement with data. This conundrum is particularly concerning at present, as land cover change is rapid and covers vast regions (Brink and Eva, 2009; Hansen *et al.*, 2013).

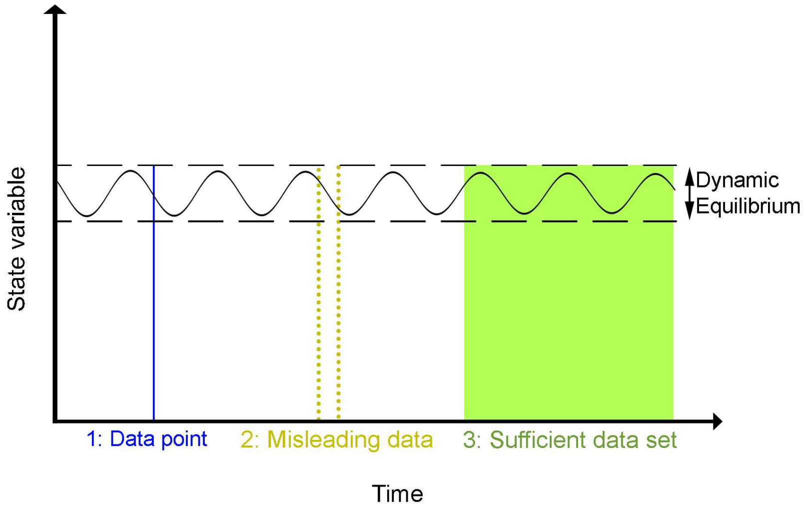


Figure 4.2: Dynamic equilibrium through time for a given ecosystem state variable and observation data. 1: One observation value through time for the given state variable. 2: Observation data for a period showing a trend not representative of the actual ecosystem dynamics. 3: Observation data over a period long enough to grasp the amplitude and frequency of state variable oscillation in a dynamic equilibrium.

4.3.3 - HOLISTIC *VS*. PATCHY

When benchmarking a model for only a fraction of the parameters it simulates, it is possible to misinterpret results and draw the wrong conclusions (Figure 4.3). Thus, a model correctly simulating a given parameter does not necessarily simulate correctly other parameters, and a model could correctly represent one parameter but for the wrong reasons.

If we assume that, in a model, only benchmarked parameters can be deemed valid and that we cannot conclude on the non-benchmarked parameters, it implies that only the ecosystem parameters for which observation data is available can be modelled. Such assumptions reduce a model extent of validity and its usefulness. There are different possibilities to avoid this limitation. First, we can consider all the non-benchmarked parameters as non-relevant for the specific question for which a model is benchmarked (for example, in chapter 2 we did not consider phenology, as it was not relevant). Second, we can simulate only parameters and processes for which there are datasets which can be used to benchmark and/or parameterise the model, implying the absence of arbitrarily or stochastically parameterised variables. For high complexity models, an intermediate approach is the most sensible option, given the difficulty to exclude the use of random initialisations. In this case, it is more meaningful to have results reasonably fitting with data for as many parameters as possible instead of trying to fit one parameter perfectly. It has been proposed to standardize model benchmarking methodologies and to systematically use all relevant data sets to have a comprehensive and holistic model assessment (Kelley *et al.*, 2013). Thus, there is an increasing need for relevant datasets. Data availability and quality can increase even without increasing observations efforts. For example, new methodologies can increase datasets accuracy and quality (Tuanmu and Jetz, 2014; Bouvet *et al.*, 2018). In general, satellite derived datasets come from an interaction between satellite imagery and ground observations. Following such approach, data inference methodologies can be developed. In this sense, a range of values for parameter A could be derived from the observed values of parameters B and C. For example, if we know vegetation height, AGBM and LAI, it might be possible to determine a realistic range of value for leaf biomass. Such methodology could benefit from Bayesian approach (Ellison, 2004).

Benchmarking multiple parameters can help to lower models' uncertainties. However, it could be that, when benchmarking two different parameters, trying to optimise the fit for one parameter decreases the fit for the other. Thus, it could be that a model cannot be parameterised to optimise benchmarking fit for multiple parameters at the same time and that the model needs to undergo additional development. Moreover, to maximize a model's fit to benchmarking data in a given context requires a specific parameterisation of the model. This parameterisation is likely to vary depending on the focus of the study, given that parameterisation accounts for adjustment of parameters and processes not implemented in the model but still contextually important. This is a major risk of misinterpretation. Indeed, if a model is parameterised to produce results fitting with benchmarking data, and if a major driver of a phenomenon or of a state parameter is absent from the model, it is unknowingly misleading. For example, if in reality a phenomenon is explained for 40% by parameter A, 30% by parameter B and 30% by parameter C, and if a model including only parameters B and C manages to explain 90% of the variance when parameterised, it is likely to overestimate the impact of parameters B and C on the phenomenon, despite appearing as more accurate than a similar model explaining only 60% of the variance. Zaehle *et al.*, (2005), showed that parameters importance varies geographically, supporting the idea that models are unlikely to perform equally at global scale. Such effect could be at play in our model and impact our ability to study shrub encroachment as a global phenomenon. In cases where shrub encroachment is due to non-modelled factors, our model of shrub encroachment should show low fit with data, but, if we parameterise the model to increase its fit with data, we might reduce confidence in its future projections.

In conclusion, the patchier the data, the more the extent of a model's validity is reduced, because it prevents the model from properly representing reality and it prevents from assessing the extent of this validity. The problem of data patchiness is particularly concerning for models representing ecosystems at large scale (continental or global) or in great details (considering many parameters); in the first case because many data are not available or not precise at such scales, in the second case because most ecosystems are studied for specific questions and not holistically. Models which need to extrapolate on patchy data sources should be considered carefully, yet, they are our best tools to model global or large-scale complex ecosystem processes.

Figure 4.3: Looking at a subset of variables does not directly allow to grasp the big picture correctly. It might not be feasible to consider enough variable to draw a holistic representation of a system, thus model's accuracy and relevance need to be properly framed (art by Gee).

4.4 - MODEL DEVELOPMENT - OPPORTUNITIES & CHALLENGES

Ecological models improved due to technical developments, and thanks to the evolution of the underlying theoretical corpus (Lavorel and Garnier, 2002; Fisher *et al.*, 2017). Limitations to present modelling approaches could be overcome through new developments (expending and refining modelled processes, improved theoretical corpus), technical improvements (code structure, computing efficiency and capacity), increase in data availability and quality (for benchmarking and parameterisation) or better model assessment. There are many opportunities to explore to take ecological modelling to the next level. We question key challenges and opportunities for the development of process-based models in ecology. We advocate for an open and active debate on these questions among modellers to foster science. We focus on the following questions: 1 - How to consider ecosystems dynamic equilibrium in model development? 2 - How to model non-deterministic natural behaviours? 3 - How can we better integrate stochasticity in model structure? 4 - Can we re-envision benchmarking methodologies for process-based models?

4.4.1 - BENCHMARKING DYNAMIC EQUILIBRIUMS & ECOLOGICAL TIMESCALE

As ecological timescale might be conflicting with human timescale (Fresco and Kroonenberg, 1992), obtaining ecological information on a meaningful timescale through reconstruction of the past is a key research field and thus bridging the gap between plant ecology and paleoecology is required (Marignani *et al.*, 2017). Research effort in this direction is intense and quickly progressing (Adams and Faure, 1998; Koch, 1998; Kröpelin *et al.*, 2008; Feurdean *et al.*, 2017, 2018; Gałka *et al.*, 2018; Zanon *et al.*, 2018; Ardenghi *et al.*, 2019). However, reconstructing past ecosystem states and dynamics implies its own questions. To reconstruct past ecosystems, we assume that “rules” derived from present observations are valid for past periods, which can be debated.

As ecological timescale challenges our ability to benchmark model results against observation data, framing a model focus is a key methodological question (e.g.: ecosystem, life-form), as it shapes how we analyse a model’s results. A process based ecological model simulates a functioning ecosystem with its processes, states and dynamics, however, they are constrained, and, typically, represent only a snapshot of reality, relatively to its input data, even if it can exhibit inter-annual variability. However, ecosystems are also an historical phenomenon (Fritts and Rodda, 1998; Fukami *et al.*, 2010; Tomscha and Gergel, 2016). In this sense, ecological timeline is equivalent to historical timeline for humans. It means that, even dynamic equilibriums of ecosystems are relative and transient. At present, models such as DGVMs, do not explicitly investigate evolutionary processes despite being able to simulate ecological timescale. This can be explained by the difficulty for the PFT approach to model evolution processes. However, trait-based model (e.g.: aDGVM2, JeDi DGVM) are well positioned to consider evolution from a DGVM perspective and to explicitly account for ecosystem history. Such approach could yield major outcomes for ecology and step up the development of DGVMs as it is theoretically and technically more correct to consider ecosystems as determined by their history (Wiens and Donoghue, 2004; Crisp, Trewick and Cook, 2011), and to model them as such.

Modelling ecosystems as an historical phenomenon is a challenging prospect. For example, ecosystems being bound by their history (Crisp, Trewick and Cook, 2011), species (and ecosystems or ecosystem states), or even functional equivalents, which disappear might not automatically reappear when the conditions under which they were existing come back. We propose two ways to take such effect into account in an ecological model. First, an ecological model can be initialised with observed values fitting with the timeframe which the model aims at simulating; which implies a limited ability to extrapolate through time. Second, an ecological model can integrate evolutionary mechanisms or be coupled to an eco-evolutionary model and take into account historical events relevant for ecosystems evolution, and thus be able to model ecosystems at ecological timescale. This second option could lead to mechanistic modelling of ecosystems evolution, based on modelling species evolution and modelling ecosystem processes. The aDGVM2 can be a first step in this direction thanks to its ability to represent mutation and crossover processes. When running simulations with aDGVM2 for ecologically relevant timescale (over 5000 years), we observe the potential for ecosystems to switch from one dynamic equilibrium to another due to intrinsic processes leading to evolutionary discontinuities (results not shown). Such perspective could allow studying questions such as: Are ecological equilibrium "true"? Are ecosystems intrinsically transitive due to history and/or specific mechanisms? However, one question would remain: To which extent can we derive rules and laws from ever-changing ecosystems?

4.4.2 - MODELLING NON-DETERMINISTIC DYNAMICS?

Models are built on observations with the assumption that rules, in the form of deterministic equations, can be derived from said observations. However, studies can propose alternative representations of a given phenomenon (Mäkelä and Sievanen, 1987; Wang, 2003), or propose conflicting observations and conclusions (Stevens, Lehmann, *et al.*, 2016). Thus, the theoretical corpus framing our understanding of the natural world is partly relative. This is due to scientific improvement but also partly due to a reorganization of scientific views on the natural world. For example, Tuanmu and Jetz (2014) produced a land cover categorisation more comprehensive than the previous land cover products it is based on, but it is also a different way of categorising land cover as the categories themselves are different between all these products. Similarly, biome, ecoregions and ecotones are different concepts which characterise vegetation cover, among other factors, in a given area. They do not change the information we have about these regions, but it changes how we understand them, and they can be characterised following different schemes. Therefore, we need to question our point of view on nature and not just the theories we use to give it meaning. Can we assess if the rules we derive from observations are natural rules or if they are a simplification of nature to make it comprehensible? In other words: Are natural processes deterministic or chaotic? Which ones? To which extent?

Chaos and average

The question of deriving general rules from observations is decreasingly relevant with narrowing focus. Not considering evolutionary processes in DGVMs narrows down focus and allows to rely on rules based on physical properties, such as carbon fixation rate through photosynthesis depending on ambient conditions (Collatz, Ribas-Carbo and Berry, 1992; Ehleringer, Cerling and Helliker, 1997). However, DGVMs typically need to simulate random processes to face uncertainties and gaps in data. Simulations replicates are then used to mitigate the weight of stochasticity by considering average values. However, average values might be a chimera, not being simulated as such by a single replicate run; thus, the replicate with the most parameters as close to the average for each parameter should be selected.

Table 4.1 shows that average, variance, minimum and maximum values vary with the number of replicates considered and with the subset of replicates considered. It shows that increasing the number of replicates considered, despite increasing the range of possibilities, increases the statistical significance of the average. However, considering the possibility of chaotic natural processes implies to consider the possibility for ecosystems to exhibit a non-deterministic dynamic; thus, deviating from observations through time. In this case, considering the average result of a model to perform predictive simulations and to test scenarios might be wrong, as the ecosystem considered might diverge from the average results under different circumstances. In the same sense, it could also be possible that an ecosystem observed at present could be in a non-average state, for example being in a transient state (Figure 4.2), or in an alternate stable state (Staal *et al.*, 2016). Theoretically, relating two parallel versions of a system (real world observations and simulations) does not imply that they would evolve in the same way due to chaos and stochasticity. Therefore, we need to ask: can we model non-deterministic processes based on deterministic equations? Concomitantly, if we are to model chaotic processes, considering average values of replicate simulations is misleading. Averaging stochastically different simulation replicates induces determinism, which goes against the simulation of non-deterministic processes. Thus, considering average values of replicate simulations implicitly assumes a deterministic behaviour, albeit non-explicitly modelled.

Modelling non-deterministic phenomenon is intrinsically challenging (Worster, 1989; Wood, 2010), and might be an unbridgeable gap between a model and reality. Explaining a share of ecosystems parameters variance can be achieved. However, to our knowledge, no model explains 100% of the variance. The non-explained fraction depends on what is considered as well as the model's quality. If we assume that natural processes are not fully deterministic, then, stochasticity can also explain part of this variance, and this, no matter how refined and complex a model is. In this sense, chaos would be the gap between reality (the full variance) and the maximum percentage of variance explained in Table 4.1.

In conclusion, solving the conundrum of deterministic *vs.* non-deterministic ecological processes might be far-off at present, but thinking about it might lead us to consider more advanced modelling experiments. For example, bridging the gap between ecological models (such as DGVMs) and eco-evolutionary models offers a possibility to integrate some of the ecological stochasticity in DGVMs. In this endeavour, aDGVM2 might be the most suitable candidate due to being the only DGVM being able to explicitly simulate mutation and cross-over to dynamically adapt vegetation communities to their environment, and thus represent non-deterministic ecosystem behaviour at large temporal scale. For example, it is able to simulate evolutionary processes leading to switch from a dynamic equilibrium to another, due to vegetation community internal processes without further external constraints over ecologically relevant time scale (results not shown).

Historical perspective

Trying to understand present ecosystem states by looking only at the mechanisms explaining the survival of any individual plant misses the point relative to how this specific plant came to exist in its specific place. Considering ecosystems history would improve our understanding of their state and dynamics (Wiens and Donoghue, 2004; Crisp, Trewick and Cook, 2011; Polly *et al.*, 2011; Staver, Bond and February, 2011). However, historical legacy introduces uncertainties related to data availability. Long term records of ecosystem states are uncommon (e.g.: Laliberté *et al.*, 2013; Lawley *et al.*, 2013; Stevens, Erasmus, *et al.*, 2016) and reconstruction of past ecosystem states is a challenging process (Adams and Faure, 1998; Koch, 1998; Kröpelin *et al.*, 2008; Feurdean *et al.*, 2017, 2018; Gałka *et al.*, 2018; Zanon *et al.*, 2018; Ardenghi *et al.*, 2019). To circumvent this limitation "rules" are derived from observations. However, different case studies focusing on the same question might lead to diverging observations. Considering more cases might smooth out these divergences. However, this might rule out underrepresented cases. This is particularly of concern for past ecosystem states, given that data availability is increasingly limited when considering increasingly older time period. Additionally, we cannot know if some cases about which no data is available ever existed. In conclusion, the fact that we typically have, at best, patchy historical data on ecosystems, implies that we need to be careful when we infer rules from ecological observations.

Coevolution

It has been argued that some species might have appeared only due to the presence of other specific species, through coevolution and mutualistic or antagonistic relations (Janzen, 1966; Stuart-Hill and Mentis, 1982; Lunau, 2004). This process was referred to as coevolution for the 1st time by Ehrlich and Raven (1964). Typically, such processes are observed at ecological time scale, capturing mutation and speciation rates. While DGVMs can simulate such timescale they typically do not explicitly investigate coevolution processes. Yet, coevolution could have a considerable impact on models' agreement with observations. Not considering coevolution might also constrain models' ability to represent specific ecosystem components (species) and ability to extrapolate model results in the past or the future. These limitations come from three points. First, model initialisation is used to take into account present or past ecosystem states; yet, it takes into account only a "state" and not a "dynamic", thus, missing coevolution processes. In consequence, it limits the model's ability to extrapolate, as it does not consider long term ecological trends. Second, during the model operation itself, as timescales of speciation and model simulation are different, the weight of processes related to speciation is typically not taken into account explicitly. Third, when benchmarking models, the weight of historic dynamics in an ecosystem is often unknown and identifying coevolution is challenging (Dybdahl, Jenkins and Nuismer, 2014), which might lead to disagreement between simulations and data unrelated to the model structure correctness. Therefore, assessing the importance of coevolution effects would contribute to improve models benchmarking.

As it has been demonstrated, model initialisation, contribute to explaining models results (Moncrieff *et al.*, 2016). In a similar fashion, weighting the impact of history on ecosystems could help to better frame models extent of validity. In this regard, testing coevolution might be a promising approach as it could build upon the strength of both evolutionary models and DGVMs. Specifically of interest for DGVMs, evolutionary models can represent the mechanisms driving speciation (Fussmann, Loreau and Abrams, 2007). Due to its unique design, aDGVM2 offers promising unique opportunities to investigate long term eco-evolutionary processes in a mechanistic ecological framework (results not shown), yet further efforts are required to better understand such processes and ensure their proper representation in aDGVM2.

In conclusion, we need to envision ecosystems as dynamics and not just as states to be able to capture coevolution and ecosystems history while considering scale effect (geographical and temporal) and its link with stochasticity. These questions offer vast possibilities for modelling research and particularly for DGVMs development, potentially leading to breakthrough in model architecture and ability to extrapolate. In this regard, aDGVM2's unique design is well suited to investigate coevolution processes and long-term speciation trends, as preliminary results demonstrate (results not shown).

4.4.3 - STOCHASTICITY: TURN A WEAKNESS INTO STRENGTH

From a conceptual perspective, models' extent of validity can be improved by using more adapted theories. The design of a model and the concepts it relies on directly limit what a model can represent and how well; limits which are further constrained by the actual structure and content of the model. Therefore, theories should be selected to provide a more accurate understanding of models design and operation, in turn allowing for more refined models. As such, developments of better model concepts provide frameworks more likely to lead to modelling breakthrough.

Stochasticity as strength

DGVMs, while being process based models, include stochasticity to provide variability and diversity to their deterministic processes and to represent heterogeneity and mechanisms not explicitly captured by their structure and content. Variability introduced by stochasticity introduces uncertainty and, therefore, can reduce models extent of validity. Fisher *et al.* (2010) argue that variability between replicate simulations is linked to uncertainty in parameterisation, and thus to stochasticity (as it is used to initialise variables for which deterministic values cannot be used), which reduces models projections reliability. Fisher *et al.* (2010) argue that models output need to be constrained by observations to allow for more reliable results. By following up on this argument, we can also see simulations variability as a chance. Indeed, greater variability between simulations could allow models to perform in wider range of context, given that results are *a posteriori* constrained by observations. As a model has wider range of results it is more likely to have results close to observations no matter how diverse and different those observations are. It can then be constrained by these observations to improve its performances.

Alternatively, selecting among many replicates the one giving the best fit with data instead of the average, means selecting the one for which stochastic processes, considered as a proxy for non-explicitly modelled processes and parameters, are the closest to these processes they are supposed to account for. This could be used for a reverse engineering approach to identify and select realistic values for variables that are not deterministically initialised.

A similar approach, yet less prone to misinterpretation, is to consider stochasticity as representing small scale heterogeneity. In chapter 2, we showed that mean value is not necessarily the most fitting with benchmarking data. Sampling trait space through stochastic simulation iteration (Table 4.1) can be used to find the most accurate output and to parameterise the model by re-constraining allowed trait space. Alternatively, aDGVM2 variability can be used to simulate the wide range of local potential heterogeneity. Following Figure 4.2 it is also possible that simulated and observed vegetation state might be an extreme in a range of potentials and not average, which questions our ability to benchmark model results with relatively limited observations.

Models typically operate on a small-scale basic unit, which is then linearly scaled up to the grid-cell resolution at which environmental data is fed to the model. However, between the grid-cell scale and the actual variability throughout landscapes there is a non-linear relation, varying geographically, which is neglected by the linear scaling from models' basic unit to grid-cell scale. Additionally, assessing models based on their average results (out of randomized replicate simulations) partially misses the interest of using stochasticity, as it prevents from capturing landscape scale variability and potentially explaining the mismatch with simulations (Fisher *et al.*, 2017). Rammig *et al.* (2018) solved the scale mismatch, between observed vegetation at the grid cell scale and models basic unit of operation, by statistically considering the small-scale variability and observation errors and using this information to correct model results. This approach supports the idea that stochasticity can be used to model sub-grid cell heterogeneity. Further, to consider sub-grid cell heterogeneity, model benchmarking would need to consider the range of potential results and how it relates with the range of vegetation observations encompassed in a simulated grid cell. However, depending on the level at which stochasticity is embedded in a model (i.e., at the cell, leaf, individual or landscape scale), it has varying consequences. For example, Fisher *et al.*, (2010), argue that stochasticity at the individual level is averaged out at the cohort scale. Therefore, stochasticity cannot be used to model vegetation heterogeneity *per se*, but only if properly parameterised, specifically when considering that it might be used to represent different types of heterogeneity depending on the region considered. Additionally, this requires a proper assessment of the weight of stochasticity in a model. For example, Table 4.1 shows that the number of simulations considered, and the sub-set of simulations considered can lead to different conclusions about model performance. Such an assessment needs to be standardized in model development to assess the effect of stochasticity and design proper experimental set-ups for simulations studies.

Stochasticity can be seen as a strength on two accounts. First, it can be used to account for landscape heterogeneity and scaling up from model basic unit to grid-cell scale. Second, it can be used to represent non-explicitly modelled processes & variables. In this second understanding, a post-processing approach could be envisioned to improve model reliability, either by considering the simulation closest to reality as the most informative or accurate among replicates instead of considering the average, or by constraining the range of results to observations. Concomitantly, a reverse engineering approach could be used to infer realistic values for randomly initialised variables. In this regard, aDGVM2's ability to produce different vegetation communities among different randomized simulation replicates (all controlled parameters being equal) could be used as a step forward in models design and performance. We propose to further investigate this question by testing if all the replicate simulations for a given region can be related to real vegetation communities, and if all the vegetation diversity is represented in models results. Such study requires precise and extensive observations to be available for the considered region, and strong collaboration between modelling team and observation team. This would allow to test the feasibility of representing fine scale variation in vegetation cover without the need for high resolution data to feed the model.

4.4.4 - FORWARD BENCHMARKING

From a technical perspective, the extent of validity of models can be directly increased by performing more numerous and more refined model assessment and benchmarking. Models are developed to represent specific ecological variables. These variables are calculated from equation representing processes. These processes are assumed to be as accurate as the observations from which they are derived (minus the loss of precision due to data regression). These are the primary elements of a model. The secondary elements of a model are derived from them and are the targets of a model. For example, xylem conductivity is a primary element directly derived from observations and explicitly coded in a model, whereas vegetation cover is the model's target, and it is calculated from xylem conductivity, and other processes, and not explicitly coded in the model. As we know the extent of validity of the primary variables included in a model, given that they are based on observations, they are not the priority target for benchmarking. Even if they can be modified by the model operation, they are assumed to remain in range with realistic values. On the other hand, we need to assess the accuracy and realism of the secondary variables simulated by the model by benchmarking them; therefore, they are a target for benchmarking.

It is possible to benchmark an increasing array of parameters, thanks to increased data availability. Many modelling researches have explored various possibilities to benchmark models depending on what the model can represent and on research questions (Luo *et al.*, 2012; Kelley *et al.*, 2013; Best *et al.*, 2015; Peng *et al.*, 2015). Most studies consider only a set of parameters to benchmark as they are designed to answer specific ecological questions. Some studies focus on modelling questions *per se* by working towards unifying and standardizing models benchmarking processes (Abramowitz, 2012; Best *et al.*, 2015; Rabin *et al.*, 2017) and/or making them more holistic (Luo *et al.*, 2012; Langan, 2019). An extensive literature is available about models benchmarking, showcasing many approaches considering a range of different parameters. This allows for the development of integrated and standardized benchmarking projects, frameworks and tools. For example, PEcAn (LeBauer *et al.*, 2013) is an "ecoinformatic workflow" aimed at facilitating the assessment of models uncertainties via statistical methods. iLAMB (F. M. Hoffman *et al.*, 2016; Collier *et al.*, 2018) is a project to streamline models benchmarking and integrate data and models to improve their validity and facilitate their development. The PALS-PLUMBER (Best *et al.*, 2015) proposed a novel benchmarking framework by estimating the target results accuracy *a priori* and benchmarking results against this estimate. Under development, the "DGVMTools" R package (M. Forrest, S. Scheiter & J. Steinkamp, unpublished) proposes a unified technical methodology of processing, visualizing, analysing and benchmarking models output. There are also research efforts towards benchmarking results of multiple models for different parameters to assess the overall relevance of these models, their strength and weaknesses. Among such models inter-comparison studies, the FireMIP project focuses on systematically comparing models performance to represent fire processes globally in ESM depending on a set of variable parameters (Rabin *et al.*, 2017). The Inter-Sectoral Impact Model Inter-comparison Project (ISI-MIP) (Warszawski *et al.*, 2014; Rosenzweig *et al.*, 2017) is a framework for models inter-comparison focusing on questions related to climate impact. It already produced results (McSweeney and Jones, 2016) and allowed for improved robustness of related research projects (Nishina *et al.*, 2015). Similarly, the Coupled Model Inter-comparison Project (CMIP) (Eyring *et al.*, 2016) coordinates and harmonize the development of models with a focus on the Amazon region, while Restrepo-Coupe *et al.* (2017) compared the ability of a set of DGVMs to represent carbon fluxes in the Amazon basin. By switching the focus, models inter-comparison studies can also be used directly to answer specific questions. For example, Sitch *et al.* (2008) used five different DGVMs to evaluate carbon cycle. Alternatively, they can provide insights on specific points of models' structure. For example, Medlyn *et al.* (2015) proposed an "assumption-centred" model inter-comparison, which evaluated models against data depending on how they represent target ecological processes.

Such research approaches are necessary to provide an overview of the state of model development and ensure a consistent quality level. There is a lot to be gained from them. However, none of them is fully holistic and there is room to expand model extent of validity to new dimensions (new parameters, new contexts, and new processes). The absence of systematic benchmarking of all the parameters for each model development step is a problem as biases and unchecked constraints could be inadvertently introduced. However, it is practically a relative problem, as models are typically benchmarked for the parameters relevant to the specific studies for which they are used. Thus, by constraining models use to specific questions, it is possible to ensure their accuracy and relevance. For example, when investigating the impact of fire on vegetation and how it might change in the future, DGVMs must represent fire processes, and be benchmarked for their simulated parameters that are related to fire. Model development efforts on these questions are intense (Spessa *et al.*, 2013; Burton *et al.*, 2019; Forkel *et al.*, 2019). They are informed by and help inform the progress of our understandings of fires globally (Hantson *et al.*, 2016; Nogueira *et al.*, 2016; Laurent *et al.*, 2019). Likewise, Luo *et al.* (2012) present a comprehensive framework for benchmarking and assessing strengths and weaknesses associated to these benchmarks for modelling biophysical processes, biogeochemical cycles and vegetation dynamics with DGVMs. Representations of processes, such as fire, are increasingly refined in DGVMs. Additionally, more and more traits and parameters are considered and benchmarked. For example, NDVI (Normalized Difference Vegetation Index) and LAI were modelled in Hickler *et al.* (2005), NPP in Hickler *et al.* (2006), SLA (Specific Leaf Area), maximum carboxylation rate at 25 ◦ C ($Vcmax_{25}$) and maximum electron transport rate at 25 ◦ C ($Jmax_{25}$) in Verheijen *et al.* (2013). While model confidence and extent of validity are increasing, many parameters require the development of specific benchmarking techniques. For example, Fisher *et al.* (2018) highlights five development priorities for VDMs (Vegetation Dynamic Model): plant competition for light and partitioning through canopies, plant hydrodynamics, below-ground resources, demographic processes, land use and fire disturbances. In conclusion, the availability of more datasets to benchmark models against and the development of model inter-comparison projects are the most promising ways to expand model extent of validity. Integrating and streamlining these two approaches should be the goal of model developers to ensure consistent quality among all models and to question their assumptions, allowing the integration of more complex dynamics into ecological models.

Scale

To improve model benchmarking, the relevance of scaling the basic unit of a simulation needs to be considered. Models are built to perform simulations on grid-cells represented by basic units. Typically, these units are of smaller size than the grid-cells they represent. For example, the basic unit for aDGVM2 is a 1ha square (Scheiter, Langan and Higgins, 2013), whereas for SEIB-DGVM it is a 30mx30m square that is used as a basic unit (Sato, Itoh and Kohyama, 2007). However, they typically simulate grid cell resolution of 0.5° to 2°. To account for this gap, the basic unit simulated by the model is typically scaled linearly to the simulated grid-cell resolution. Stochastic processes and randomized simulation replicates are used to account for the environmental heterogeneity at the grid-cell scale. This is problematic when the same level of stochasticity and replicates is supposed to scale the model from its basic unit to the grid-cell resolution, no matter what the grid-cell resolution is, and thus, no matter how much diversity it is supposed to account for. Additionally, models typically represent a fixed number of potential individuals (or individual cells, as simulated individuals can be alive or dead, or present or absent) at their basic unit scale.

Simulated proportions need to be considered given the scale of their simulation and the scale of their benchmarking, because representing plants traits proportions (the relative occurrence of each trait) for a small-scale basic unit does not imply that the model would represent the same proportions if it was to be used with a larger scale basic unit. Concomitantly, scaling up linearly the results produced by the model at its basic unit scale to the grid-cell resolution is arguable. For example, one could ask how SEIB-DGVM, with a basic unit of 30mx30m, can mechanistically simulate competition between multiple large trees and how modifying its basic unit size would impact simulations in this regard. Generally, a model built to produce realistic results when run at its basic unit scale might perform differently if its basic unit scale is modified. This is what Table 4.1 shows. Modifying the number of individual cells simulated at the 1Ha square of the aDGVM2's basic unit modifies the simulations, both in terms of variables values and in terms of model behaviour. Simulations with greater individual cells number result in greater minimum and mean simulated AGBM, associated with greater number of woody individuals. The variance between replicate simulations decreases more rapidly with increasing number of replicates when more individual cells are simulated. Changing the number of simulated individual cells does not change the spatial dimension of the basic unit but it changes its density and can therefore be assimilated to running the model for a larger (or smaller) area but considering the same number of individual cells. These results imply that models might simulate the correct trait distribution at its basic unit scale and that it can be deemed correct at larger grid-cell resolution only if linear scaling is correct. At the same time, these results question the validity of the linear scaling assumption.

Table 4.1 also shows that, to efficiently sample trait space, there is a trade-off between simulating more individuals per basic unit, which requires less replicates to constrain variance, and simulating less individuals per basic unit which requires more replicates to constrain variance. In this regard, it appears that increasing trait space (by adding more traits or allowing for greater potential trait range, including more sub-division of trait range) calls for increased computation effort (either through increased number of replicates or "larger" basic unit) to constrain variance.

In conclusion, the linear scaling used to transform model results from the basic unit scale to the grid-cell resolution needs to be questioned on two terms. First, the scaling of stochasticity, representing grid-cell diversity unaccounted for by the model at its basic unit should be questioned. The stochastic processes can be called for accounting for diversity of different magnitudes depending on the resolution at which the results are scaled up, and this, without modifying these stochastic processes. Yet, they cannot be expected to represent equally heterogeneous diversities, or scaling results at varying resolutions. Second, the structure of a model, the parameters and processes it is built on, are used to perform simulations at the scale of the model's basic unit and are expected to reproduce the correct traits distributions & proportions of a larger region. However, the linear scaling that is assumed to justify that the basic unit of a model are representative of larger grid-cell scale vegetation is typically unchecked. Our preliminary results show that the scale of the basic unit can have consequences on model performances. In consequence, we should ask: can the model perform equally, or at least similarly, if the scale of its basic unit is modified? Or, how to ensure that scaling up model's simulations from their basic unit to grid-cell level does not distort model's performance?

*Table 4.1: 100 replicate simulations were performed for a savanna site in South Africa, on the border of the Kruger National Park (coordinates: 30°E 24°S), following the procedure described in chapter 2. It was performed for 2 scenarios: a) a grid of 40*40 individual cells per hectare was computed, b) A grid of 80*80 individual cells per hectare was computed. It implies that in a) up to 1600 plant individuals could be simulated and in b) up to 6400 individuals could be simulated. A set of x simulations (x=3, 5, 10, 20, 50) was drawn at random five times for each set size. The mean AGBM, over the last 10 simulated years, was calculated for each simulation. We showed mean, minimum and maximum values among the x replicates of each set and Standard Deviation (S.D.) and Variance calculated among the 5 replicates of each set. The number of simulated individuals (# indiv.) is given for a) the full set of simulations and for b) each replicate.*

a) 40*40	Replicates	Set of 3	Set of 5	Set of 10	Set of 20	Set of 50	Set of 100	
		AGBM	AGBM	AGBM	AGBM	AGBM	AGBM	# Alive indiv.
Mean	1	62	42	59	57	56	54	568
	2	43	47	57	49	56		
	3	27	83	35	59	52		
	4	55	35	66	53	57		
	5	60	55	33	57	55		
	S. D.	14	18	15	4	2		
	Variance	168	272	180	13	2		
Min	1	41	3	15	11	1	1	277
	2	2	2	9	1	2		
	3	5	60	4	3	2		
	4	40	2	39	0	2		
	5	42	5	1	3	3		
	S. D.	20	26	15	4	1		
	Variance	333	522	177	13	0		
Max	1	82	82	101	101	119	119	1011
	2	82	75	83	118	109		
	3	64	101	83	107	107		
	4	79	72	94	101	117		
	5	89	107	107	109	109		
	S. D.	9	16	11	7	5		
	Variance	70	197	88	40	23		
Variance							2341	90195
S. D.							57	360

b) 80*80	Replicates	Set of 3		Set of 5		Set of 10		Set of 20		Set of 50		Set of 100	
		AGBM	# indiv.	AGBM	# indiv.	AGBM	# indiv.	AGBM	# indiv.	AGBM	# indiv.	AGBM	# indiv.
Mean	1	45	2404	60	3044	64	3058	72	2784	64	2964	67	2940
	2	57	3730	76	2575	64	3007	59	2765	65	2977		
	3	78	2926	76	2942	68	2946	73	2987	67	2882		
	4	59	2898	67	2904	66	2973	67	2844	68	2922		
	5	48	2489	79	3099	58	2685	64	2981	69	3015		
	S. D.	13	526	8	204	4	145	6	106	2	51		
	Variance	134	221048	50	33371	11	16879	27	8997	4	2098		
Min	1	29	1911	39	1911	29	2195	25	1789	8	1819	8	1789
	2	40	3200	52	1958	27	2223	16	1819	8	1819		
	3	43	2579	49	2057	38	1958	25	1881	20	1789		
	4	40	1819	55	2056	8	2455	8	1911	26	1819		
	5	20	1859	64	2303	26	1911	26	1819	8	1819		
	S. D.	10	604	9	151	11	220	8	50	9	13		
	Variance	73	292127	65	18329	96	38845	49	2026	60	145		
Max	1	67	2687	91	4281	82	3987	110	4235	110	4235	112	4281
	2	65	4281	95	3687	99	4281	93	3873	101	4281		
	3	98	3518	101	3711	109	4235	112	3787	109	4281		
	4	82	3675	94	3543	99	3758	108	3688	112	3987		
	5	64	3304	98	3518	97	3304	101	3873	112	4281		
	S. D.	15	579	4	310	10	400	8	207	5	128		
	Variance	172	268043	12	76795	78	127852	47	34121	17	13136		
Variance												1915	1037042
S. D.												51	1247

Demographic processes

Following up on arguments from Fisher *et al.* (2018), the representation of demographic processes is a priority for model development. We propose an approach to benchmark vegetation demographics through DBH (Diameter at Breast Height), considered as an indicator variable, following works of Muller-Landau *et al.* (2006), Lima *et al.* (2016), and Moore *et al.*, (2018). These studies show trees density as a function of their DBH for North America, Panama and tropical forests respectively. Additionally, Lima *et al.* (2016) identified a relation between DBH density, growth rate and mortality following a Weibull distribution. As shown in Moore *et al.* (2018), Muller-Landau *et al.* (2006), Crowther *et al.* (2015) and Lima *et al.* (2016), no matter the species or the site, DBH density distribution follows similar trends, albeit with some differences, therefore, we can use such material to benchmark models results. Figure 4.4 shows that aDGVM2 also simulates a decreasing number of individuals with increasing DBH, but only above a certain DBH, depending on the pool of plant individuals considered. Lima *et al.* (2016) shows that the relation between

DBH, growth rate and mortality rate varies with species, making precise benchmarking of demographics unpractical. However, it could be used to constrain model processes by linking DBH to growth rate and mortality rate based on α and β parameters of the Weibull relation identified by Lima *et al.* (2016) used as traits. DBH size class distribution observations are available for some sites in the TRY database (Kattge *et al.*, 2011) but not systematically available globally. Crowther *et al.* (2015) tree density map is available at the global scale (with the limitation that only trees with DBH>10cm are taken into account). By correlating such map with DBH distribution studies, it should be possible to derive an estimate of DBH distribution globally, which could then be used to benchmark demographic processes more precisely.

Tree density per DBH at the global scale cannot be benchmarked at present due to data availability issue. However, we argue that further research on this question could yield major results in the near future and step-up models' representation of vegetation demographics. We propose two approaches to investigate this question. First, we can benchmark the number of trees to Crowther *et al.* (2015) world map of tree density. Second, we can check if models' representation of DBH distribution is in range with observations from Moore *et al.* (2018), Muller-Landau *et al.* (2006) and Lima *et al.* (2016) or similar studies. However, this makes sense only if the considered model can properly represent the correct number of individuals, as we have seen that it can affect overall models' behaviour (Table 4.1) and that linearly scaling up results should be considered carefully. Finally, benchmarking state variables is only a proxy to benchmark processes. Directly benchmarking processes is conceptually more challenging as it revolves around a circular logic since models' representation of processes are built on equations which are regression of observations data. In this regard, it is possible to benchmark processes themselves only if there is variability in a process representation (e.g.: being modelled as the sum of various observations) or if it can be checked by studying its footprint (i.e., looking at how introducing a given process in a model modifies the behaviour of other processes in the model). For example, growth and mortality rates can be benchmarked by looking at their relation with DBH, following observations from Lima *et al.* (2016). Figure 4.4 shows that aDGVM2 can simulate similar trend of decreasing woody plant density with increasing DBH and also exhibits sites of low plant density and low DBH, where woody plants growth is constrained by environmental factors. More generally, representation of vegetation structure by aDGVM2 can be more precisely benchmarked, for example, using LIDAR data.

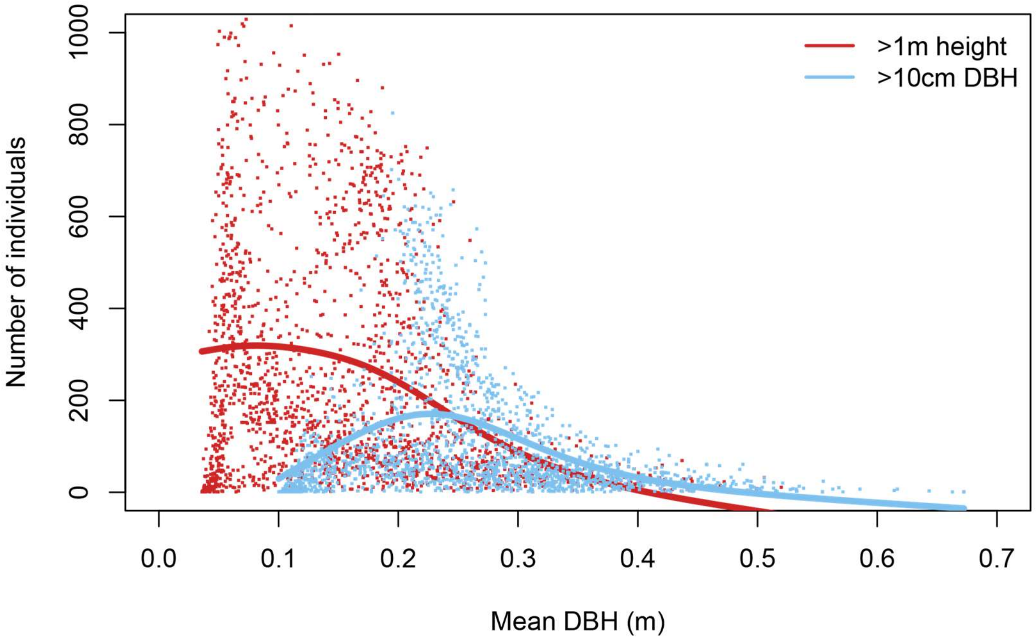


Figure 4.4: Number of simulated woody individuals per grid-cell, either taller than 1m (red) or with a DBH larger than 10cm (blue) against their mean DBH, across sub-Saharan Africa; according to simulations from chapter 2, in the presence of fire, with shrubs included in the model.

4.5 - DISCUSSION

4.5.1 - FOUNDATIONS: MODELS & DATA

Each approximation and assumption made to design a model and any additional biases, which can be introduced through different processes, have different consequences. Each of these consequences need to be considered and addressed to parameterise and accurately frame models' extent of validity, in order to compensate for their specific biases. It has been shown, albeit for hydrologic models, that models performances can be impacted more by models structure and the assumptions and observations they are based on than by their complexity, or range of considered elements (Yew Gan, Dlamini and Biftu, 1997). Similarly, Warren and Seifert (2011) show that, for ecological niche modelling, models complexity should aim for a middle ground, as models too complex are as likely as models too simple to improperly represent parameters and processes and thus to have lower accuracy and ability to extrapolate. Therefore, modellers should aim to improve models' accuracy and relevance without increasing their complexity. For example, Hayat *et al.*

(2017) propose a more refined model of trees to be used in DGVMs. Such representation could improve the validity of DGVMs by considering more plant level trade-offs and processes without necessarily making it more convoluted. However, this approach has been designed and tested for constrained geographical extents and it needs to be tested globally before being implemented in a DGVM aiming at representing global vegetation.

Models could benefit from being assessed on their structure and content and not just on their results. Being correct on a benchmarked parameter does not imply to be correct on another non benchmarked parameter. Therefore, a model extent of validity is relative to benchmarked parameters and not a general assumption on a model. Thus, model shortcomings can be hidden by misleadingly good benchmarks on at least some parameters. Alternatively, a model can be grounded on sound concepts and data yet exhibit poor performances for a given set of benchmarked parameters. Such conundrum calls for in depth technical assessment of models' structure and content.

Observations or derived products have their own sources of bias (Moilanen, 2002; Hortal *et al.*, 2008; Martin, Blossey and Ellis, 2012; Kosmala *et al.*, 2016). Modellers need to be aware of these sources of bias and carefully assess data they rely on in order to improve models' accuracy and reliability. Variability among observations can be large (Table 2.1) (Saatchi *et al.*, 2011). Figure 4.2 shows that observations can also be misleading, as they are typically shorter than ecological time scale (i.e., too short to capture the full range of variability of a dynamic equilibrium) (Hastings, 2004). The trade-off between large-scale observations which lack precision and site-specific studies which are more contextually constrained is a key determinant of data selection for modellers. Including data encompassing different scales to build and benchmark models can impact simulated ecological processes (Lawes *et al.*, 2007; Fisher, 2010; Fisher *et al.*, 2010).

4.5.2 - BENCHMARKING & INITIALISATION

The potential benefit of closer collaboration between expertise in data analysis and modelling has been demonstrated for life science at large (Carius and Findeisen, 2016), environmental questions in particular (Aspinall and Pearson, 2000; Li, Brimicombe and Ralphs, 2000) and can be expected to hold true for ecological models. The interactions between data and model, either pre or post process face challenges specific to their entanglement.

Data regression is necessary to translate observations into models, however, it introduces biases (Whittingham *et al.*, 2006). To limit this effect, an array of alternatives are available. For example, adding a varying distribution probability along the regression, in order to represent density of data points along the regression line. Alternatively, a "regression area" could be used to cover the whole range of observations points by having a distribution probability covering the entire virtual distribution space. Further opportunities are highlighted in the literature, for example by Whittingham *et al.* (2006). Such solutions are mathematically more complex than a linear regression and might be a limit to models' operation due to their higher computation intensity. Alternatively, the use of neural network architecture (Lek *et al.*, 1996; Lek and Guégan, 1999; Liu *et al.*, 2018; Brodrick, Davies and Asner, 2019) and deep learning (Christin, Hervet and Lecomte, 2019; Rammer and Seidl, 2019; Guo *et al.*, 2020) could be envisioned to provide a more suited methodology to feed data into models as well as an opportunity to increase data quality and availability, thus, counter-acting the loss of data quality through regression. For example, such methodologies could be used to feed models with "real" data distribution yet stochastically.

Models with very good statistical fit to benchmarking datasets are not necessarily the most informative (Pitt and Myung, 2002). Concomitantly, difficulty to interpret models results and to draw conclusions from them increases with models' complexity. Such phenomenon implies to re-think the assessment of models results and how to relate them to observations. Where direct statistical fit with benchmarking data might not be sufficient to assess models' validity, it might be necessary to rely more on expert knowledge (data experts and models experts), emphasizing the need for interdisciplinary model development. Additionally, considering average values of multiple replicate simulations implies to model a mean which might not exist. Similarly, aggregated data used to benchmark models' results can be an average value which does not exist; the greater the biodiversity it encompasses the greater the weight of this effect.

Models are typically fitted pre-process and benchmarked post-process, but they can also be constrained post-process. We propose three ways to envision such methodology. First, selecting among replicate simulations the one producing the results closest to observation data, instead of considering the average result of multiple replicates, might prove suitable. It would imply that stochastic parameterisation and processes produced the most realistic model behaviour for a given simulation, despite not being the model average behaviour. Second, results heterogeneity among replicate simulations could be seen as a representation of the local scale heterogeneity, which has been identified as having a strong influence on ecosystems despite being typically not explicitly modelled (Pappas *et al.*, 2015). Third, constraining results post-process with observations could be used to operationalise models and improve their reliability to inform society. Alternatively, a model's ability to extrapolate could be improved by weighting results of scenario simulations with present day difference between model and simulations.

In any case (using the replicate simulation with the best fit to observation data, or using the range of replicate to simulate local scale heterogeneity, or using the average of simulation replicates), there is an implicit assumption. That is: observed vegetation states used to parameterise and benchmark models are a valid foundation and/or target for models. However, this assumption needs to be questioned. Indeed, models are typically built to represent natural ecological and biological processes. Some models are parameterised to account for anthropological impacts, but typically not in a mechanistic way (Forrest *et al.*, 2020). Considering that the current observed vegetation states are not necessarily the "natural" or "optimal" states, given anthropological impacts, it appears that directly correlating a model built to represent potential natural vegetation and observed vegetation state, be it for model development or benchmarking, is problematic. It is possible to exclude anthropised areas from the considered datasets, but this has two issues. First, it might severely limit data availability. Second, human impact is difficult to quantify and qualify (Galli *et al.*, 2016; Kaplan and Valle, 2018; Corbau *et al.*, 2019). In this regard, the limitation through time of many datasets (e.g.: typically, global and satellite derived datasets) challenges our ability to assess human impact history. This is particularly of concern given that historic dynamics can drive ecosystems behaviour more than climate change (Thom *et al.*, 2018).

Benchmarking *vs.* Understanding

Smith *et al.* (2016) showed that ecosystem responses to drivers can lead to similar ecological responses through different plant community responses. Zaehle *et al.* (2005)

showed that models can accurately represent ecological processes yet produce highly variable results. Consequently, it is important to look at mechanisms and fine detail ecosystem structure variables and not just at upper-level indicators and to take different perspectives on models' results. Comparing data and models results gives an estimation of models' extent of validity, however, it is different from assessing models' structure. The critical part to weight models' confidence is to assess the operation of its framework by drawing connections between benchmarks and model content. We argue that this process calls for more holistic benchmarking in order to draw multiple connections between benchmarks and model content in order to weave a mesh covering as much of the models' structure and to provide information about their functioning. We propose a contribution to this approach in chapter 3. This requires in depth knowledge and understanding of a model, from its components to its benchmarks, which might be unpractical for model intercomparison projects. Yet, Medlyn *et al.* (2015), showed that, for specific questions, model intercomparison studies can inform about models' structure and functioning.

4.5.3 - CHALLENGES & OPPORTUNITIES

Depending on the model assumptions, and specifically on the data it relies on, it can be more adapted to represent certain regions or natural features than others. For example, a model considering every natural process but not those related to cold adaptation could have excellent fit with benchmarking data everywhere but in the cold regions. The importance of limits originating from a model underlying assumptions is increasing with increasing models complexity. As stated by Pitt and Myung (2002), increasing models complexity increases the difficulty to interpret models results and to draw conclusions from them. Additionally, models with better fit to benchmarking data are not necessarily the most informative. The more complex a model, the less likely it is to correctly parameterise it (Moore *et al.*, 2018). Thus, there is a paradoxical trade-off between model complexity and the ability to infer knowledge from model.

Limits to model development depend on the theoretical and observational references framing a model development. They can be mitigated by revising models design, but they can also be mitigated by novel benchmarking approaches. For example, Hartig *et al.* (2012) revisit the parameterisation and benchmarking paradigm of DVMs (Dynamic

Vegetation Model) and proposes to overcome some of its limits by switching to a new paradigm, relying on bayesian statistics and inverse inference to estimate parameters. Prentice *et al.* (2015) propose a different approach to improve parameterisation and models development at large, relying on a systematization of parameterisation and benchmarking techniques and on the use of stochastic processes.

One of the main challenges driving models performances is its parameterisation (Lawrence *et al.*, 2011; Cipriotti *et al.*, 2012; Li, Zeng and Levis, 2012; Pappas *et al.*, 2013; Verhoef and Egea, 2014; Moncrieff *et al.*, 2016; Kim *et al.*, 2019). The progress in models' realism, and added complexity, makes parameterisation more and more challenging. To solve this conundrum, it has been argued that stochastic parameterisation can be used to improve models "validity" (Prentice *et al.*, 2015). Such approach implies an *a posteriori* assessment and confirmation of a model extent of validity. Simulations made with a model randomly initialised need to be replicated until results are deemed robust enough or until the correct initialisation is found. Machine learning techniques could be used as an alternative to random initialisation. Machine learning can be used to improve parameterisation accuracy either by providing new data (Wäldchen and Mäder, 2018) or by automating parameterisation (Thessen, 2016). In this sense, machine learning can be used to find the parameterisation producing the best fit between simulations and benchmarking data. Machine learning can also be used to facilitate models development (Recknagel, 2001; Peters *et al.*, 2014; Thessen, 2016; Gobeyn *et al.*, 2019; Ma *et al.*, 2019; Mehta and Pankaj, 2019). For example, Gobeyn *et al.* (2019), proposed evolutionary algorithms, derived from machine learning, to help calibrate models, but also to reduce models complexity. Ma *et al.* (2019) envisioned to use machine learning to identify major drivers and key mechanisms explaining vegetation state. Such stirring perspectives come with a cost. First, they might require more computing resources, possibly in amount usually beyond what is available to the ecological modelling research field. Second, letting algorithms extract relevant information from data, while facilitating models' development might be detrimental to the theoretical and conceptual development of models. Third, improving models' performance does not directly equal to an improved understanding of ecosystems functions, processes and parameters. Improving the fit between a model and benchmarking data does not necessarily inform about its ability to extrapolate. Therefore, machine learning techniques hold great promises for ecological models' development, yet they need to be carefully supervised to avoid mistakes and pitfalls.

To further model development typically comes at the cost of increase computing demand. To meet this demand, it is possible to either invest more in computing resources, or by investing in improving models' efficiency. This second point means investing in programming, either to optimise code structure or to use more advanced computing techniques, such as GPU acceleration, particularly well adapted to support individual based models. To compensate for the increasing demand on computation time, optimization can also be made on the models' development workflow. This encompasses technical tools made to facilitate benchmarking of models, such as the "DGVMTools" R package (Forrest, Scheiter and Steinkamp, unpublished) as well as standardized and streamlined methodologies (LeBauer *et al.*, 2013; Warszawski *et al.*, 2014; Best *et al.*, 2015; Eyring *et al.*, 2016; F. M. Hoffman *et al.*, 2016; Rosenzweig *et al.*, 2017; Collier *et al.*, 2018).

Model development can also be enlightened and informed by ecological perspectives. For example, biodiversity metrics (Santini *et al.*, 2017) can be used to assess simulations results and answer critical ecological questions (Langan, 2019). In general, the ecological questions aimed at being answered by a model shape this model development. Therefore, we argue for model development to focus on previously unanswered questions in order to expand models' capabilities and to improve our representation and understanding of ecosystems states and behaviours. Focusing on trait trade-off and traits responses, effects and drivers holds great potential (Langan, Higgins and Scheiter, 2017; Gaillard *et al.*, 2018; Pfeiffer *et al.*, 2019; Berzaghi *et al.*, 2020), as modelling mechanisms and processes is crucial to improve models ability to represent and "forecast" scenarios.

4.5.4 - OUTLINE

To conclude on a theoretical perspective on models, we argue that, in the same way as an object exists only in a given context, a model should be equally framed by its context, which includes the constraints of the data it relies on for its design and context of the object it simulates (local biotic and abiotic factors). It means that we have to ensure that a model has the same context as the object it is supposed to represent. For example, a model should incorporate empirically derived equations obtained for the same geographical extent as the one for which the model is designed for. However, references from diverse sources, corresponding to different contexts, might be used to build a model; which can conflict with the model target context. For example, a model might be made to represent vegetation in general but all its parameters' values could come from equatorial Africa only, or a model could aim at representing specifically African savanna but its fire model could come from global observations. Consequently, a model produces a virtual ecosystem which is a chimera made of all of its references and not necessarily an existing ecosystem. From this perspective, it appears that a model, based on context-dependent observations, when tasked to represent ecosystems and scenarios outside of the range of its references is biased. Assessing the magnitude and "direction" of this bias is a challenge; which could prove highly valuable for a model's relevance. Additionally, benchmarking processes should first be used to support model development by providing information about the discrepancies between the context of the references constitutive of the model and the context of the benchmarked regions. Therefore, we need a model based on ecological processes which are valid regardless of the ecosystem state and context considered. In this regard, a DGVM based on universally valid ecological processes, compared to a model empirically derived, has the potential to have greater extent of validity and be better adapted to study scenarios (*i.e.*, cases for which we do not have observation data but which could realistically happen). DGVMs in general are going in this direction, with more process rich models. The aDGVM2 makes an extra-step in this direction (Scheiter, Langan and Higgins, 2013).

4.6 - CONCLUSION

Models simplify the object they represent by selecting a set of key characteristics, for DGVMs, these are major ecological processes and fundamental vegetation traits. Their specific structures represent a given fraction of real ecosystems and this varies geographically depending on the relative weight of its represented processes and traits locally. Models' extents of validity are inherently constrained by the data they are based on. Consequently, at large scale, DGVMs, and ESMs in general, should not be expected to have a consistent match with observation data, and should not strictly aim for one for one benchmarking fit.

Interplays between processes can have different relative importance depending on the network of processes they are related to, specifically given that vegetation communities create their own biotic environment shaping potential co-existence (Bruelheide *et al.*, 2018). Such processes are better studied with models similar to aDGVM2 (Scheiter, Langan and

Higgins, 2013; Langan, Higgins and Scheiter, 2017), which can represent trait-based community assembly processes.

As “a rising tide lifts all boats”, the emergence of trait-based models drives DGVMs development towards more realism (Pavlick *et al.*, 2013). This trend holds the potential to address concerns faced by all ecological models and potentially lead to conceptual breakthroughs. Simulating future ecosystems, specifically under changing climatic conditions, is challenging when they might be out of the range of what is known at present (Purves and Pacala, 2008; Allen *et al.*, 2010). To address such problem and to simulate accurately such ecological dynamics, we need to study processes and mechanisms and not only states. This implies to benchmark processes instead of benchmarking states.

CHAPTER 5 - MODELLING IMPACTS OF ATMOSPHERIC [CO_2], FIRE, GRAZING AND MAP ON SHRUB ENCROACHMENT WITH ADGVM2

This chapter is the work of the sole author of this thesis.

ABSTRACT

Background

Shrub encroachment is a widespread phenomenon worldwide, more generally part of the larger trend of woody cover expansion. It is of particularly high concern in Africa where vast swathes of grasslands and savannas are at risk. Despite numerous studies on the subject, there is no consensus about the mechanisms explaining it.

Objectives

We assess the relative importance of several parameters identified as major drivers of shrub encroachment events across multiple sites in southern Africa. This allows us to estimate the relative importance of each factor depending on context dependencies.

Methods

We use the aDGVM2, an individual-based and trait-based DGVM able to mechanistically simulate the emergent properties of shrub life-forms in African savannas. We test the effect of MAP, grazing, fire and atmospheric [CO_2] on vegetation community assemblages across 35 sites in southern Africa where woody cover expansion has been documented.

Results

We simulate higher shrub prevalence at lower MAP depending on considered combination of factors. Elevated atmospheric [CO_2] decreases shrub prevalence, indicating that cases identified as shrub encroachment might be a step towards woody cover expansion. Fire and grazing have a complex interaction and absence of fire as well as elevated fire frequency can limit shrub prevalence in some cases.

Conclusions

As shrub encroachment is driven by a complex interaction of drivers, management strategies should be case specific and should not rely on generalisations. A consensus should emerge about classification of “shrub encroachment”, “bush encroachment” and “woody cover expansion” to avoid overlap. This would ease further investigations and communication about each phenomenon. Our simulations indicates that in some cases shrub encroachment can be a steppingstone towards overall woody cover expansion. We recommend further simulations, including site specific ecosystem history in order to precisely test this dynamic.

5.1 - INTRODUCTION

5.1.1 - SHRUB ENCROACHMENT

Shrub encroachment became a major scientific topic in ecology and related fields around ten years ago, with around 600 publications per year between 2014 and 2019 (Figure 1.8). Vast swathes of land have been identified as experiencing shrub encroachment globally (Stevens, Lehmann, *et al.*, 2016). Shrub encroachment can impact ecosystems in multiple ways, for example by reducing rangeland farming carrying capacity, or by modifying biodiversity and potentially leading to biome change (Mugasi, Sabiiti and Tayebwa, 2000; Lett *et al.*, 2004; Knapp *et al.*, 2008; Sirami *et al.*, 2009; Eldridge *et al.*, 2012). The role of drivers of ecosystem states is also modified by shrub encroachment as it feeds back on these drivers. For example, woody plant encroachment can modify ecohydrological systems (Huxman *et al.*, 2005), which then feeds back on vegetation cover as water availability is a strong driver. Similarly, many studies show that controlled fire, along with browsing (Staver *et al.*, 2009) can suppress tree and shrub density in African savannas (Sankey *et al.*, 2012), while shrub encroachment reduces fire intensity and frequency. Further, grazing interacts with aridity to explain vegetation distribution patterns (Pfeiffer *et al.*, 2019). Several studies identified over-grazing as a cause for shrub encroachment (Archer, Schimel and Holland, 1995; Roques, O’Connor and Watkinson, 2001; Rutherford, Powrie and Husted, 2012; Stevens, Erasmus, *et al.*, 2016), whereas Tocco *et al.* (2013) contradicts these results.

The global extent of shrub encroachment follows its diversity (Eldridge *et al.*, 2012; Stevens, Lehmann, *et al.*, 2016). The first explanation for this diversity is the lack of a consensus on the definition of "shrub" in the literature. Often, shrubs are categorised as short or small woody individuals, sometimes without definitions of state parameter thresholds (e.g.: height). However, even when there is a clear numerical categorisation, comparison between studies is difficult because the threshold to classify woody plants as small, and, thus, as shrubs, varies typically between 3m and 6m. This classification fuzziness can lead to overlaps between categories of bushes (typically defined as woody plants shorter than shrubs) (Cao *et al.*, 2018), shrubs and trees, and intermediate categories. This could be related to global vegetation cover classifications, such as biomes maps (Olson *et al.*, 2001; Ellis and Ramankutty, 2008; Conradi *et al.*, 2020), typically having one or more shrubland category but no bushland category. Additionally, shrub encroachment might contribute to woody cover expansion globally, encompassing encroachment of trees, shrubs and bushes (Wigley, Bond and Hoffman, 2009; O'Connor, Puttick and Hoffman, 2014) and forest expansion (Donohue *et al.*, 2013). The intertwining of these phenomena can be explained both by authors considering them as part of the same process and by the lack of consensus among the research community on the definition of these phenomena.

As shrub encroachment is part of larger ecological dynamics (woody cover expansion), it needs to be understood in the light of these dynamics. Forest expansion has been associated with changes in human societies (Mather and Needle, 1998; Parés-Ramos, Gould and Aide, 2008). This concurs with the general observation that land use is a main driver of vegetation cover (Aleman, Blarquez and Staver, 2016), but it also questions the perception of woody cover change as it depends on stakeholders perspectives (Wigley, Bond and Hoffman, 2009). We know that biome distributions can also be explained by objective biotic and abiotic factors. For example, grassland distribution depends on availability of resource and disturbances (fire and herbivory) (Bond, 2008). Other studies point at atmospheric [CO_2] increase as a main driver of global woody cover expansion (Donohue *et al.*, 2013) and particularly of shrub encroachment in South Africa (Buitenwerf *et al.*, 2012). Atmospheric [CO_2] increase generally promotes woody vegetation over C4 grasses due to [CO_2] fertilization effects (Ehleringer, Cerling and Helliker, 1997; Kicklighter *et al.*, 1999; Bradley and Pregitzer, 2007; Donohue *et al.*, 2013; Kumar, Pfeiffer, Gaillard, Langan and Scheiter, 2020), while land use change related to overgrazing reduces grass biomass and thereby fire activity which can favour woody vegetation (Grover and Musick, 1990; Caracciolo *et al.*, 2016; Pfeiffer *et al.*, 2019). As Ringrose *et al.* (2002) show, shrub encroachment can happen due to natural or anthropogenic factors, and O'Connor, Puttick and Hoffman (2014) suggest that an interplay of factors (rainfall, history and management) drives shrub encroachment, while Aleman, Jarzyna and Staver (2018) emphasise the role of history to explain current biomes boundaries. Explaining shrub encroachment is further challenged as it feeds back on vegetation dynamics and ecosystem functions. For example, increasing woody biomass typically contributes to suppress grass growth and thereby fire, which is important to open the landscape, to suppress growth of woody vegetation and thereby to maintain vegetation in an open savanna state (Bond and Midgley, 2012). Additionally, reduced grass biomass implies reduced food availability for grazing animals (Pfeiffer *et al.*, 2019). Solving such feedback cycle is challenging. Shrub encroachment can be explained by its drivers but also by the dynamics it relies on. For example, Silva and Anand (2011) and Duarte *et al.* (2006) showed that the expansion of forest into grasslands relies on various mechanisms but that it exhibits a consistent "phenotypic" or "strategic" pattern. Pioneer trees or shrubs establish in grassland and create clusters of forest expansion. Various explanatory mechanisms have been proposed: varying nutrient concentration, absorption and use (Silva and Batalha, 2011), sheltering effect *vs.* grass competition mediated by water competition, light competition and fire resistance competition (pioneer tree excluding grass via water and light competition and sheltering seedlings from fire via low canopy branching), seed aggregation by providing preferential habitat for seed dispersers (Puyravaud, Dufour and Aravajy, 2003).

5.1.2 - OBJECTIVES

Many studies aimed at understanding shrub encroachment and many studies aimed at modelling woody vegetation cover distribution (Staal *et al.*, 2018). Due to intense research efforts, several factors driving woody cover distribution and shrub encroachment have been identified. However, if we are to comprehensively understand interactions among drivers of woody encroachment, we require process-based models and multi-factorial experiments considering all the potential drivers. Such an approach has long been limited by the lack of an explicit mechanistic representation of shrub life-forms as a unique strategy in models, distinct from tree life-forms. To be able to explore potential factors influencing shrub encroachment and to project future vegetation, we require a model that

(1) uses ecophysiological principles to simulate vegetation in response to environmental conditions, (2) includes shrubs as distinct life-form based on explicitly simulated mechanisms and (3) simulates how environmental factors, hypothesized as potential explanatory factors of shrub encroachment (such as grazing and fire), influence abundances and community structure of grasses, trees and shrubs. Despite recent research effort (Hickler *et al.*, 2006; Zeng, Zeng and Barlage, 2008; Clark *et al.*, 2011; Lawrence *et al.*, 2011), the present study is, to our knowledge, the first study that includes all these aspects. Our study was facilitated by the unique design of aDGVM2, an individual-based dynamic vegetation model (Prentice *et al.*, 2007) that allows traits values of individual plants to dynamically adjust to the biotic and abiotic environment (Scheiter, Langan and Higgins, 2013; Langan, Higgins and Scheiter, 2017). An updated version of aDGVM2 includes grazing and a detailed representation of grasses dynamics (Pfeiffer *et al.*, 2019) and includes a mechanistic representation of shrubs as multi-stemmed woody individuals (chapter 2). With these new developments, the aDGVM2 now fulfils the requirements necessary for a multi-factorial mechanistic modelling study of shrub encroachment.

For our study, following chapter 2, we classify as shrubs simulated woody plants which are multi-stemmed (>2) and up to 4m. We refer to shrub encroachment as an increase of the prevalence of shrubs in a plant community. Shrub encroachment, in this sense, is more common in ecosystems such as savannas and grasslands (Stevens, Lehmann, *et al.*, 2016). We focus on shrub encroachment in southern Africa, and particularly on savanna cases. This focus is particularly relevant due to Africa's vegetation meta-stability, originating from being disturbance-driven. Wilcox *et al.* (2018) points at alternative stable state theory and pyric herbivory to explain grasslands and open savannas distribution. Shrub encroachment impacts large expanses of these ecosystems and can lead to rapid vegetation change between alternative stable states (Midgley and Bond, 2015; Stevens, Erasmus, *et al.*, 2016; Stevens, Lehmann, *et al.*, 2016). Focusing on southern Africa also aligns with the fact that aDGVM2 has been tailored for African savannas. Stevens, Lehmann, *et al.* (2016) assembled studies quantifying shrub encroachment in southern Africa which we rely on to focus our simulations.

We aim to assess the relative importance and the mechanisms associated to each identified driver of shrub encroachment in order to derive a comprehensive explanation of the phenomenon. We use aDGVM2 to study how atmospheric [CO_2], rainfall, fire and grazing interact to drive shrub encroachment. These are the main drivers identified by Stevens *et al.* (2016), which aDGVM2 is able to simulate. Our main questions are:

1. How do atmospheric [CO_2], grazing, fire and rainfall shape shrubs prevalence across southern African savannas?

2. Are the interactions between these factors synergistic or antagonistic?

3. How understanding shrub encroachment drivers shapes our understanding of the future of shrub encroachment cases?

4. Can fire management and grazing contribute to suppress undesired shrub encroachment?

5.2 - METHODS

5.2.1 - THE ADGVM2

We use the aDGVM2 (Scheiter, Langan and Higgins, 2013; Langan, Higgins and Scheiter, 2017) for our simulations. The aDGVM2 is an individual-based model which simulates growth, reproduction and mortality of individual plants while keeping track of state variables such as biomass, height and leaf area index of individual plants. Each plant is characterised by a potentially unique set of traits describing plant type (grass or woody), leaf characteristics, leaf phenology, carbon allocation to different plant compartments, plant architecture, response to fire, reproduction and mortality. Growth, reproduction and mortality of plants are influenced by both the plant specific combination of trait values and the environmental conditions. Plant traits are linked by trade-offs to constrain overall plant performance. Mass conservation trade-offs regulate allocation to roots, stems, leaves, bark, storage and reproduction. Engineering trade-offs regulate plant architecture (Niklas and Spatz, 2010), while empirical functions define, for example, trade-offs between specific leaf area (SLA) and leaf longevity (Reich, Walters and Ellsworth, 1997) or between SLA and the capacity of a plant to extract water from the soil. The aDGVM2 simulates soil water competition and light competition via impacts of each individual plant on the total resource. Water uptake of single plants is defined by the fraction of root biomass in different soil layers, the moisture content of these soil layers and by the plant capacity to extract water from the soil (i.e., P50, matric potential corresponding to 50% loss of xylem conductivity (Langan, Higgins and Scheiter, 2017)). The light available to an individual plant is influenced by the height and LAI of neighbouring plants. Light and water availability influence the photosynthetic rate and thereby, via carbon status, the reproduction and mortality rates of each individual plant. Reproduction includes trait inheritance, cross-over and mutation of trait values. Plants that allocate enough carbon to reproduction can produce seeds. Seeds can exchange trait values thereby allowing recombination of the community trait pool (cross-over). Mutation adds new trait values to the community trait pool. Randomly selected seeds are drawn from the resulting community trait pool and are added to the plant population as seedlings. By simulating inheritance, mutation and crossover, the model generates a large variety of different trait combinations and iteratively, via mortality and reproduction, assembles a plant community that is adapted to the environmental conditions.

The aDGVM2 is particularly well suited for our study as it stems from aDGVM which was developed for African savannas (Scheiter and Higgins, 2007, 2008, 2012; Moncrieff *et al.*, 2015, 2016; Scheiter, Moncrieff, *et al.*, 2019). The aDGVM2 has been benchmarked (Langan, Higgins and Scheiter, 2017), its shrub component has been benchmarked (chapter 2) and its grazing module has been benchmarked in Pfeiffer *et al.*, (2019).

The aDGVM2 is the only DGVM that simulates shrubs based on a trait trade-off, and not as a PFT. In aDGVM2, shrubs are represented as multi-stemmed woody individuals (chapter 2), in contrast to single-stemmed trees. The advantage of multi-stemmed shrubs is the higher sapwood area for a given biomass and therefore higher competitiveness in water limited environments, compared to single-stemmed trees. The disadvantage of multi-stemmed individuals is that they are typically smaller than single-stemmed trees, which reduces their ability to compete for light. The number of stems is a trait and it is passed from one generation into the next generation. Stem number is constant during a plant's life, while in reality, the number of stems is dynamic and can change for example due to re-sprouting after fire. We use mean stem count of all woody plant individuals in a grid-cell as a proxy for shrub prevalence. The higher this value, the higher the share of multi-stemmed plants in the vegetation community. However, this does not consider the number of stem per shrub. The relative balance between grasses and woody plants is also discounted.

A recent addition to aDGVM2 allows to simulate the impact of grazing on grasses (Pfeiffer *et al.*, 2019). It represents grass biomass removal by grazers while distinguishing between annual and perennial grasses. In aDGVM2, each plant individual is simulated by a potentially unique trait combination. This allows to link grass removal to both biomass and trait values of individual plants and hence to simulate that grazers are selective (they preferentially eat fresh, nutrient rich biomass). When animals visit a grass stand, an amount of biomass is removed. This amount is defined by the daily requirement of animals and by the number of animals. Days when animals visit a vegetation stand are randomly selected and influenced by the visitation frequency of animals (see Pfeiffer *et al.*, 2019).

In aDGVM2, fire is represented mechanistically. Days with fire ignitions are randomly distributed during the years and the ignition probability is influenced by tree canopy area and precipitation. More open landscapes with low tree cover favour grasses (which are the main type of fuel in the study system, and rapid desiccation of fuel biomass). Low precipitation is used as proxy to represent low fuel moisture which promotes fire. Fuel biomass, fuel moisture and wind speed are used to calculate fire intensity (Higgins *et al.*, 2008). When potential fire intensity exceeds a given threshold, an ignition event can randomly occur and fire spreads and damages vegetation. Fire removes grass leaf biomass while tree biomass removal is influenced by height and bark characteristics.

Finally, aDGVM2 takes into account atmospheric [CO_2] and precipitation effects on vegetation due to the representation of ecophysiological processes, such as photosynthesis, respiration and plant hydrology. Therefore, we can use it for our multi-factorial experiment to test effects of MAP, [CO_2], grazing and fire on shrub encroachment and shrub distribution and competitiveness in general.

5.2.2 - SIMULATION EXPERIMENTS

We selected 35 sites across southern Africa (Figure 5.1) where woody cover expansion has been observed within savannas (Stevens, Lehmann, *et al.*, 2016). Details about site coordinates, observed shrub cover change and proposed drivers, measured MAP and simulated MAP are given in Table 5.1. We tested a matrix of scenarios to identify conditions that maximize shrub prevalence. We explicitly prescribed values for atmospheric [CO_2], fire and grazing (Table 5.2), while MAP and other climate parameters were considered through sites specific differences. Among the selected study sites, the magnitude of woody cover change ranges from a -5% to +44% (Table 5.1) and covers a range of pre-existing vegetation covers, from open grasslands to dense savanna.

We tested three levels of fire frequency: a high frequency with prescribed yearly fires, an intermediate "natural" fire frequency (i.e., simulated by the aDGVM2 fire routines), and an exclusion of fire. While the natural fire regime can occur any day of the year depending on climatic conditions, the prescribed fire has a fixed ignition date on day 90, representing late wet season fires. We tested three levels of grazing: absence of grazing, a grazing demand of 20kg/day/ha and a grazing demand of 30kg/day/ha; these values correspond to grazing intensity with low and high impact, respectively, on grasses communities, as identified by Pfeiffer *et al.* (2019). The low intensity can induce shifts in grass community assembly, while the high intensity systematically induces shifts in grass community assembly. We tested three levels of atmospheric [CO_2]: pre-industrial, with a concentration of 283ppm (Meinshausen *et al.*, 2011), present, with a concentration of 387ppm (Giorgetta *et al.*, 2013), future, with a concentration of 670ppm, following the RCP 6.0 scenario (Meinshausen *et al.*, 2011). Various levels of MAP were considered based on the MAP of each selected site, which ranged from 87mm/year to 1415mm/year according to the CRU (Climate Research Unit) dataset used to perform our simulations (New *et al.*, 2002), and from 223mm/year to 1067mm/year for the values given by the site studies (Stevens, Lehmann, *et al.*, 2016) (Table 5.1).

We conducted factorial simulations considering interactions between all the fire and grazing levels we selected. We considered atmospheric [CO_2] only for the scenario with natural fire and absence of grazing in order to limit the computation time, which implies that we used C1 (Table 5.2) for all scenarios unless stated otherwise. In aDGVM2, stochastic processes related to fire, plant demography and community assembly imply that repeated simulations for similar environmental conditions can differ. We therefore performed 10 replicate simulations for each of our scenarios. Each simulation was run for 500 years as previous tests demonstrated that this time length is necessary and sufficient for vegetation community assembly to reach a dynamic equilibrium.

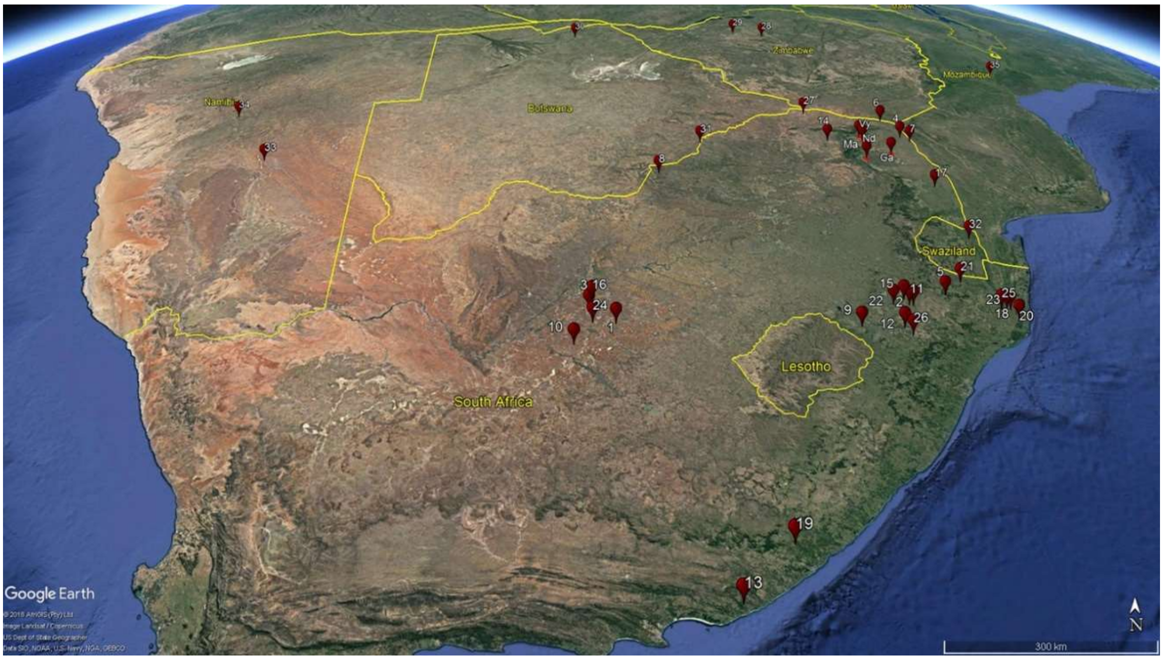


Figure 5.1: Selected sites across southern Africa where woody encroachment has been documented (Stevens, Lehmann, et al., 2016), and for which we performed simulations. (Image: Landsat/Copernicus, Data: SIO, NOAA, U.S. Navy, NGA, GEBCO, Image: IBCAo; obtained with Google Earth Pro, 2019).

Table 5.1: Coordinates of simulated sites; observed cover changes and proposed driver; MAP measured on site (Stevens, Lehmann, et al., 2016), MAP used to perform simulations (New et al., 2002) and difference between them. Cover change indicates change in the absolute value of woody cover, not relative changes (for example, 4% indicates shift from 10% to 14% and not from 10% to 10.4%).

Longitude	Latitude	Site ID	Change	Driver	MAP Stevens	MAP aDGVM2	MAP difference
25.09	-28.98	1	-5%	Rainfall & CO_2	435	442	-7
30.39	-28.21	2	4%	CO_2	775	730	45
24.63	-28.69	3	2%	-	441	417	24
31.28	-22.99	4	5%	Fire	473	497	-24
31.05	-27.88	5	6%	-	854	850	4
30.98	-22.38	6	7%	-	478	422	56
31.46	-23.12	7	7%	Grazing	438	524	-86
26.03	-25.00	8	6%	Grazing, Fire	476	595	-119
29.36	-28.81	9	9%	Grazing	869	727	142
24.37	-29.42	10	15%	Rainfall, CO_2	386	338	48
30.32	-28.21	11	9%	CO_2	774	730	44
30.13	-28.73	12	10%	-	691	730	-39
26.82	-33.50	13	13%	Grazing	605	519	86
29.64	-23.45	14	15%	-	686	466	220
30.24	-28.10	15	28%	CO_2	776	730	46
24.68	-28.50	16	9%	-	438	417	21
31.61	-24.73	17	18%	-	631	570	61
32.14	-28.03	18	19%	CO_2	891	1007	-116
27.67	-32.67	19	18%	Rainfall, Grazing	715	1093	-378
32.31	-28.23	20	36%	-	978	1007	-29
31.42	-27.50	21	34%	Elephants	816	850	-34
30.03	-28.23	22	33%	CO_2	775	730	45
32.21	-28.02	23	36%	CO_2	916	1007	-91
24.69	-28.96	24	42%	Rainfall, CO_2	450	417	33
32.04	-28.04	25	44%	CO_2	891	1007	-116
30.25	-28.87	26	38%	Grazing	771	730	41
29.30	-22.38	27	29%	Grazing	351	344	7
28.93	-18.22	28	-37%	Cultivation	761	703	58
18.12	-28.23	29	22.2-27%	Elephants, Fire, Harvesting	642	87	555
24.36	-18.80	30	-30%, +28%	Elephants	486	586	-100
26.94	-23.87	31	13%	Grazing	464	407	57
31.90	-26.27	32	1-23-29-40%	Fire, Grazing	672	1211	-539
18.00	-24.00	33	6%	CO_2	223	205	18
17.00	-22.00	34	14%	Grazing	360	366	-6
34.50	-18.79	35	10%	CO2, Elephants	1067	1415	-348

Table 5.2: Parameters used for the simulations. Grazing levels were derived from Pfeiffer et al. (2019). Fire levels represent: fire exclusion, frequent & regular artificial fire ignition, and natural fires. Atmospheric [CO$_2$] values are taken from Meinshausen et al. (2011) for pre-industrial and future values and from Giorgetta et al. (2013) for contemporary values.

Grazing intensity	G0	No grazing
	G1	20kg/day/ha
	G2	30kg/day/ha
Fire frequency	F0	No fire
	F1	Natural fire regime
	F2	Yearly late wet season fire
Atmospheric [CO$_2$]	C0	283ppm (Pre-industrial)
	C1	387ppm (Present)
	C2	670ppm (2100, RCP 6.0)

5.3 - RESULTS

The aDGVM2 simulated higher shrub prevalence (i.e., mean stem count) and standard deviation in scenario F1_G0_C0 than for any other scenario. Differences among the other scenarios were smaller (Figure 5.2). The second highest shrub prevalence was simulated for scenario F1_G0, albeit only slightly higher than the other scenarios. The lowest shrub prevalence was simulated for scenario F1_G2. The higher the simulated mean stem count, the higher the standard deviation.

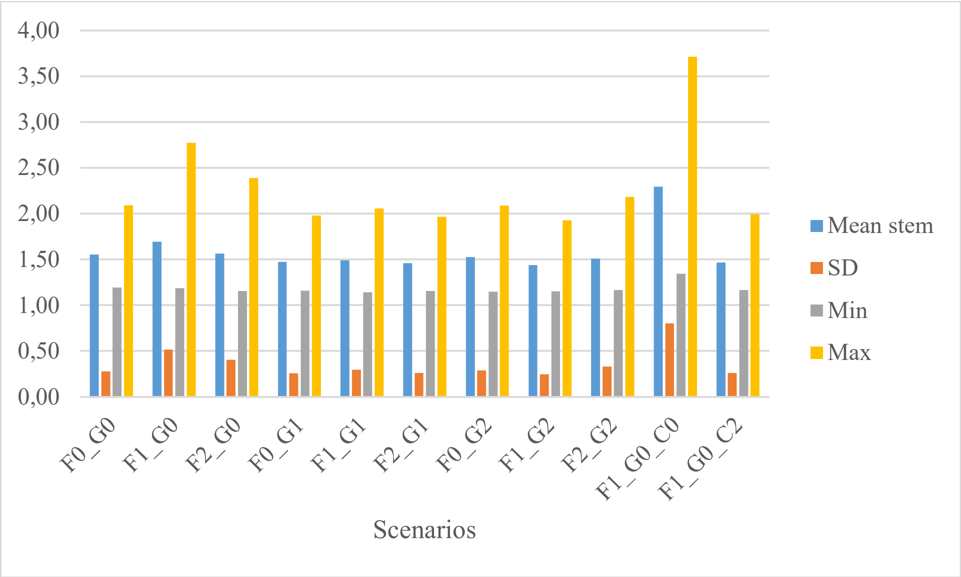


Figure 5.2: Simulated mean stem count of all alive plant individuals. Mean among all replicates and all sites (Mean), standard deviation among all mean site value (SD), minimum (Min) and maximum value among all averaged replicates per site (Max). C1 is considered for all scenarios unless stated otherwise.

By considering each site individually, we assessed if MAP explains shrub prevalence in aDGVM2. The relation between MAP and shrub prevalence was non-linear. The aDGVM2 simulated a sharp threshold in shrub prevalence between MAP of 703mm/year and 727mm/year, for most scenarios (Table 5.3). Scenarios F0_G0, F0_G1, F0_G2 and F1_G0_C2 did not follow this pattern, highlighting the importance of fire and elevated atmospheric [CO_2]. Even when discounting the threshold effect observed in most scenarios, the increase in shrub prevalence with decreasing MAP, below 703mm/year, varies for most scenarios. Some sites strongly deviate from the general trend. For example, site 10, despite having the third lowest simulated MAP, has a relatively low shrub prevalence for most scenarios. The aDGVM2 simulates higher shrub prevalence with pre-industrial atmospheric [CO_2]. Fire scenario F1 (with natural fires) has higher shrub prevalence compared to scenarios without fire F0 and with artificially increased fire frequency F2 (Figure 5.2), for grazing scenarios G0 (no grazing) and G1 (low grazing) but not for grazing scenario G2 (high grazing) (Figure 5.2). Thus, the higher the grazing pressure, the less natural fires favour shrubs. Comparing grazing scenarios, shrub prevalence is lowest for grazing scenario G1 (low grazing) and fire scenarios F0 and F2, while it decreases with increasing grazing pressure for fire scenario F1 (Figure 5.2).

Table 5.3: Mean stem count averaged among all replicates for each site and each scenario. MAP is given for each site based on the values used to perform our simulations. Values are colour coded: from light for low values to dark for high values.

Scenario	F0_G0	F1_G0	F2_G0	F0_G1	F1_G1	F2_G1	F0_G2	F1_G2	F2_G2	F1_G0_C0	F1_G0_C2	MAP
29	2.56	2.31	2.73	2.30	2.10	2.58	3.11	1.82	2.16	2.65	2.83	87
33	1.92	2.56	2.57	1.83	3.11	1.72	1.87	2.26	3.18	3.36	1.93	205
10	1.61	2.20	1.85	1.31	1.32	1.35	1.45	1.39	1.44	1.97	1.32	338
27	1.56	1.66	2.14	1.77	1.60	1.80	1.67	1.45	1.73	3.50	1.58	344
34	1.72	2.84	2.26	1.62	2.53	2.22	1.79	2.52	3.04	4.21	1.98	366
31	1.84	1.70	1.70	1.50	1.35	1.37	1.52	1.50	1.88	3.04	1.49	407
3	1.83	2.26	1.40	1.48	1.23	1.27	1.47	1.30	1.59	2.25	1.42	417
16	1.41	1.59	1.94	1.34	1.50	1.29	1.50	1.40	1.39	2.22	1.44	417
24	1.43	1.98	1.41	1.39	1.27	1.42	1.42	1.32	1.32	2.31	1.38	417
6	1.46	1.48	1.49	1.49	1.77	1.44	1.41	1.37	1.34	1.77	1.46	422
1	1.48	1.81	1.91	1.51	1.36	1.40	1.74	1.36	1.38	3.23	1.48	442
14	1.45	2.01	1.59	1.51	1.43	1.57	1.36	1.46	1.33	2.53	1.38	466
4	1.58	1.52	1.43	1.52	1.38	1.83	1.65	1.45	1.49	3.11	1.41	497
13	1.45	1.25	1.17	1.46	1.30	1.19	1.39	1.29	1.27	2.01	1.28	519
7	1.50	1.33	1.45	1.64	1.47	1.34	1.66	1.31	1.40	2.09	1.33	524
17	1.63	1.32	1.29	1.40	1.36	1.32	1.58	1.45	1.38	1.55	1.37	570
30	1.47	2.21	2.23	1.70	1.67	1.98	1.61	1.70	1.72	3.59	1.52	586
8	1.64	2.15	1.61	1.40	1.46	1.51	1.35	1.51	1.41	3.46	1.40	595
28	1.43	3.10	2.32	1.38	2.08	2.02	1.30	2.04	1.98	4.15	1.69	703
9	1.54	1.39	1.54	1.37	1.32	1.30	1.40	1.19	1.23	1.80	1.34	727
2	1.52	1.55	1.39	1.32	1.18	1.32	1.44	1.22	1.20	2.03	1.35	730
11	1.46	1.33	1.32	1.29	1.36	1.27	1.51	1.22	1.39	2.05	1.36	730
12	1.42	1.46	1.43	1.36	1.33	1.44	1.38	1.34	1.32	2.42	1.49	730
15	1.47	1.47	1.21	1.41	1.32	1.21	1.32	1.30	1.36	1.87	1.38	730
22	1.43	1.40	1.40	1.41	1.26	1.32	1.45	1.30	1.24	2.23	1.41	730
26	1.42	1.36	1.24	1.33	1.32	1.19	1.36	1.35	1.21	1.53	1.28	730
5	1.42	1.29	1.17	1.41	1.30	1.17	1.28	1.17	1.15	1.35	1.41	850
21	1.46	1.45	1.45	1.33	1.28	1.22	1.40	1.26	1.29	1.91	1.37	850
18	1.48	1.27	1.11	1.33	1.29	1.26	1.33	1.35	1.19	1.32	1.30	1007
20	1.37	1.36	1.12	1.35	1.23	1.25	1.43	1.32	1.25	1.22	1.33	1007
23	1.44	1.24	1.14	1.47	1.28	1.29	1.34	1.24	1.27	1.32	1.37	1007
25	1.44	1.33	1.11	1.28	1.36	1.23	1.33	1.26	1.21	1.43	1.27	1007
19	1.38	1.28	1.16	1.30	1.28	1.21	1.41	1.24	1.25	1.81	1.27	1093
32	1.57	1.43	1.11	1.35	1.44	1.34	1.45	1.35	1.34	1.75	1.34	1211
35	1.51	1.34	1.36	1.71	1.35	1.43	1.60	1.32	1.42	1.30	1.33	1415

5.4 - DISCUSSION

5.4.1 - EXPERIMENT RESULTS

Among the combinations of factors we tested, shrub prevalence was highest for scenario F1_G0_C0 ("natural" fires, no grazing and pre-industrial atmospheric [CO_2]) except for a few sites. We also observed that absence of fire as well as heightened fire frequency are detrimental to shrub prevalence as simulated by aDGVM2. Shrub prevalence was the lowest for scenario F1_G2, but scenarios F0_G1, F2_G1 and F1_G0_C2 had very similar values.

We argue that the decrease of shrub prevalence with increasing atmospheric [CO_2] is due to CO_2 fertilisation effect. As CO_2 fertilisation effect (Kgope, Bond and Midgley, 2010; Hickler, Rammig and C, 2015) has been simulated by aDGVM2 (Langan, 2019; Kumar, Pfeiffer, Gaillard, Langan and Scheiter, 2020) and aDGVM, its predecessor (Scheiter and Higgins, 2009; Higgins and Scheiter, 2012; Scheiter *et al.*, 2015, 2018), we understand this result as indicating a transition towards higher prevalence of trees due to their enhanced growth capacity (taller woody vegetation would suppress shorter shrubs). Therefore, the apparent contradiction between our results and studies highlighting elevated [CO_2] as a driver of shrub encroachment is deceptive, as in such cases, shrub encroachment is simulated as vegetation greening by aDGVM2, and more specifically, as woody cover encroachment, yet not shrub encroachment *per se*.

Study sites showing shrub encroachment cases (Stevens, Lehmann, *et al.,* 2016) do not show a strong correlation between rates of shrub encroachment and MAP (Table 5.1), yet change in precipitation regime was identified as a driver in a number of these studies. This implies that change in precipitation regime should drive shrub encroachment more than MAP *per se*, which means that shrub encroachment is related to ecosystems history, which was not considered in our experiment. Other factors can explain shrub distribution (chapter 2), such as elephants or land use change, which were excluded from this study. Additionally, the selection of sites where shrub encroachment was observed can introduce a bias in our study which is unaccounted for, and which could be related to parameters not considered in our simulations.

The aDGVM2 simulated a sharp threshold for shrub prevalence between MAP of 703mm/year and 727mm/year, for most scenarios (Table 5.3), but F0_G0, F0_G1, F0_G2 and F1_G0_C2 did not follow this pattern. This result highlights the importance of elevated atmospheric [CO_2]. It indicates that increasing atmospheric [CO_2] is likely to favour woody cover expansion more than shrub encroachment alone. Thus, shrub encroachment might be only a step towards woody cover expansion in a range of cases. In this regard, we can say that increasing atmospheric [CO_2] decrease the MAP threshold at which woody cover expansion can occur.

The interaction between fire and grazing is complex. However, we can conclude that shrub prevalence can be lowered both by increased fire frequency and suppressed fires, in cases with no to low grazing, as this relation is blurred at higher grazing intensity. Additionally, shrub prevalence is highest in the absence of grazing, but it is non-uniformly reduced by increasing grazing pressure as it interacts with fire regimes.

While shrub prevalence might respond linearly to a single driver, our results show that it is also possible for shrub prevalence to exhibit threshold effect relatively to certain parameters. Thus, as shrub encroachment can be explained by a combination of diverse drivers, conclusions on shrub encroachment should be drawn cautiously as it is challenging to disentangle the relative weight of each driver in diverse contexts and as shrub encroachment can be multi-faceted and confused with woody cover expansion. Conducting our experiments for sites where shrub cover decrease was observed could yield different results. However, we hypothesise that our results would remain consistent, unless site specific ecosystem history could be accounted for. This could be achieved by considering all possible transitive cases between our tested scenarios (for example, switching from scenario F1_G1 to scenario F2_G2 during a simulation), at the cost of computing time and resource. Another possibility would be to investigate only case specific scenarios for which history is documented, which would require to focus on a subset of site-specific studies.

Shrub encroachment is a multi-factorial phenomenon. Across southern Africa shrub encroachment cases associated to elevated atmospheric [CO_2] are ultimately cases of woody encroachment, associated with vegetation greening and increase in vegetation height. Shrubs thrive best in the presence of natural fires and can be stressed both by heightened and suppressed fires frequencies while grazing impact is more complex and mediated by other drivers. We hypothesise that, while it is possible to identify a primary driver of shrub encroachment, locally there is no unique explanation to such a widespread phenomenon. Thus, it is necessary to investigated local drivers, without relying on generalisation, and it is necessary to tailor precise management plans to successfully shape vegetation communities when facing shrub encroachment.

5.4.2 - EXPERIMENT LIMITATIONS

Our study is circumscribed to the combinations of factors we tested. We excluded some factors, which have been identified as relevant for shrub encroachment in other studies, because of their absence from aDGVM2, such as elephants, trampling, management strategies, wood fuel harvesting (Stevens, Lehmann, *et al.*, 2016), or browsers (Staver and Bond, 2014). We excluded some combinations of factors to keep the study scale manageable. We did not test transitivity between states, which could be important as hysteresis has been demonstrated to be relevant for many ecological processes in a wide range of ecosystems (Sternberg, 2001; Staal *et al.*, 2018). Heterogeneity among our results might be explained by local parameters impacting shrub prevalence, but this heterogeneity could also be explained by stochasticity, as 10 replicates might not be sufficient. We also did not include processes for which aDGVM2 is not precise enough, such as invasion through clustering and sheltering (Duarte *et al.*, 2006; Silva and Anand, 2011) or such as ecosystem history (Grover and Musick, 1990).

Some limitations about the number of scenarios (factors and levels) considered could be readily taken into account for atmospheric [CO_2], as historical data and future projections scenarios are available, at the cost of increased computing resources demand. Fire, grazing and precipitation would require records and projections which are patchy or absent and thus require to focus on a subset of sites. Additionally, the factors we tested have additional dimensions which we did not consider, such as seasonality and fire intensity, yet they have been considered in several field studies (e.g.: Perkins and Owens, 2003; Wright and Clarke, 2007). Finally, stochastic processes in aDGVM2 might be impacting our results, despite our 10 replicate simulations for each scenario. This could contribute to explain why the increase in shrub prevalence with decreasing MAP, below 703mm/year, has ups and downs simulated for most scenarios. However, this pattern might also be explained by local factors non-explicitly considered (e.g.: soil depth and texture, precipitation seasonality). For example, site 10, despite having the third lowest simulated MAP has a relatively low shrub prevalence for most scenarios.

While mean stem count among all simulated plant individuals represent a satisfying proxy for relative shrub prevalence, it does not further inform on ecosystem states. Therefore, a lower mean stem count, while indicating a lower shrub prevalence, can be associated with either an opening of the landscape, with an increased relative abundance of grasses (having no stem, as defined in aDGVM2), or it can be associated with canopy closure and an increased relative abundance of trees (being typically single-stemmed). Thus, further investigations are required to disentangle this conundrum. However, we expect to find decreasing mean vegetation height with decreasing MAP and increasing mean vegetation height (and canopy closure) with increasing MAP (cf. chapter 3). This would concur with present findings indicating an increase in shrub prevalence with decreasing MAP, depending on scenarios considered. Therefore, depending on scenarios, vegetation tends towards grassland or towards shrublands with decreasing MAP, while it tends towards savanna and increasing tree prevalence with increasing MAP. Given our present findings and aDGVM2 design, we propose to investigate soil texture, soil depth and seasonality (precipitation and incoming radiations) to identify the parameters driving ecosystems into grasslands or shrublands at lower MAP.

Additionally, considering shrub prevalence relatively to mean stem count indicates only a relative abundance of multi-stemmed plants. As shown in chapter 3, our understanding of vegetation cover depends on the variable considered. We argue that a more holistic assessment, considering different parameters would provide a better understanding of ecosystem responses to the relative impact of different drivers. As mean stem count response to our scenarios and sites is heterogeneous, we expect responses of other variables (e.g.: canopy cover, number of individuals, AGBM) to diverge, challenging our ability to conclude on shrub encroachment drivers.

5.4.3 - ECOLOGICAL PERSPECTIVES

Stress factors are deemed to constrain trait space (Mason and Pavoine, 2013; Lamanna *et al.*, 2014; de la Riva *et al.*, 2018). However, assessing the relative weight of different factors, of their levels, modality, interactions and direct or indirect feedbacks is challenging. Thus, we constrained our study scale in order to limit the complexity. For example, grazing can have impact on grassland community composition similar to that of aridity (i.e., overgrazing and drought often have similar effects on grasslands) (Fynn and O'Connor, 2000; Pfeiffer *et al.*, 2019). Yet, no direct equivalence has been drawn between aridity and overgrazing. Venter, Cramer and Hawkins (2018) found out that both high and low extremes of herbivory pressure, specifically browsers, can mitigate woody encroachment; which agrees with our results in the case of natural fire. They also argued that fire and grazing management can be used locally to control woody cover. However, Pellegrini *et al.* (2017) showed that fire has an increasingly limited impact on plant community assembly when biomass productivity increases, whereas elephants have a relatively stable impact on vegetation regardless of the primary productivity. Additionally, while the effects of shrub encroachment drivers can be difficult to disentangle due to their interactions, shrub cover itself feeds back on environmental factors and their drivers (Smit and Prins, 2015). The mechanism explaining such feedbacks remains to be explored.

Disentangling the precise effect of different drivers in different contexts to explain shrub encroachment is challenging. Considering the bigger picture instead of focusing on the details might allow to draw generally meaningful conclusions. For example, Axelsson and Hanan (2018) argued that the main explanation to shrub encroachment is the difference between potential woody cover and actual woody cover. This could imply that ecosystems history, including land use change, animal density, or MAP, but also shifts *per se* (e.g.: pressure release or increase), play a major role in shrub encroachment. However, Stevens, Lehmann, *et al.* (2016) did not list any study pointing explicitly at ecosystems history to explain woody cover expansion in southern African savannas.

5.4.4 - SHRUB ENCROACHMENT MANAGEMENT

Deliberate burning has been commonly proposed as a valid strategy to open savanna landscapes and to manage shrub encroachment (Trollope, 1980; Joubert, Smit and Hoffman, 2012; Sankey *et al.*, 2012; Lohmann *et al.*, 2014). However, whether controlled burnings are suitable to prevent or revert shrub encroachment in savannas is questioned on terms of fire regime parameters (frequency, intensity, seasonality) and interactions with co-drivers, such as grazing regimes (Midgley and Bond, 2015; Smit *et al.*, 2016; Stevens, Lehmann, *et al.*, 2016; Case and Staver, 2017b), which our findings support. Our simulations show that increased fire frequency decreases shrub prevalence, thus maintaining vegetation in an open state. Yet, compared to the scenario without fire, our results suggest that natural fire does not reduce shrub cover. Therefore, the presence of shrubs might be "natural" in large savanna areas (Eldridge *et al.*, 2011). Our findings indicate that proposing fire as a management tool to control shrub encroachment should be assessed with caution. However, our study is limited by the number of parameters and scenarios we tested. Further studies are necessary to fully test how fire management can influence shrub encroachment by controlling fire return intervals, fire intensity and fire seasonality. These factors have been shown to modify vegetation structure in savannas both in empirical studies (Higgins *et al.*, 2007; Savadogo *et al.*, 2009) and in previous aDGVM modelling studies (Scheiter *et al.*, 2015; Scheiter and Savadogo, 2016).

5.5 - CONCLUSION

From an ecological perspective, two further questions need to be answered to assess shrub encroachment. First, what are the ecological dynamics underlying it? Is shrub encroachment indicating that vegetation cover was below its "optimum" (Axelsson and Hanan, 2018) or related to alternative stable states (A C Staver, Archibald and Levin, 2011; D'Odorico, Okin and Bestelmeyer, 2011)? Second, what can be the future of shrub encroachment phenomenon? Does it lead to desertification (Grover and Musick, 1990; Van Auken, 2000; Archer, Boutton and Hibbard, 2001; Eldridge *et al.*, 2012) or to forest expansion (Maestre *et al.*, 2009; Donohue *et al.*, 2013; Stevens, Lehmann, *et al.*, 2016)? As shrub encroachment is a protean phenomenon, the question is: under which conditions shrub encroachment can be a step towards forest expansion and under which conditions towards desertification? To answer this question, further investigations are required. For example, by initialising aDGVM2 with different states, desert or forest, and simulating how vegetation expands or regresses when shrubs are introduced.

From a policy and management perspective, the protean nature of the shrub encroachment and the interaction between its drivers is challenging as it prevents from drawing general conclusions. Therefore, management strategies need to be specifically tailored according to the unique combination of parameters of each case and the weight and direction of each action need to be precisely assessed as it is not possible to conclude that "less grazing" or "more fire" are always the best solution.

Drawing comprehensive conclusions about shrub encroachment is further challenged by the lack of consensus in the literature. For example, “bush encroachment”, “shrub encroachment” and “woody cover expansion” are often overlapping. We call for a unification and clarification of the terminology and propose to define shrub encroachment as an increase in the prevalence of short multi-stemmed woody plants. This matters as the perception of the phenomenon determines its analysis and proposed management responses. Shrub encroachment has been characterised as a threat in many studies (e.g.: MacLeod, Brown and Noble, 1993; Blaum *et al.*, 2007; Sirami *et al.*, 2009; Eldridge *et al.*, 2012), yet, it is not clear if it is the shrub encroachment *per se* that is a threat or the causes of this encroachment that are a threat (Eldridge *et al.*, 2013). Moreover, the characterisation of this potential threat implies to define what is at risk, and thus, what is desirable. For example, while fire can be used to manage shrub cover in some cases, it can also have negative impact on biodiversity (Parr and Chown, 2003).

We argue that studying shrub encroachment as a global phenomenon part of the woody cover expansion trend can improve our understanding of ecological processes, specifically about the interaction between stress factors, and is necessary to forecast future shrub encroachment trajectories. We conclude that shrub encroachment management policies should be designed at site to regional scale, considering context specific parameters, avoiding general solutions.

CHAPTER 6 - DISCUSSION

This chapter is the work of the sole author of this thesis.

6.1 - OVERVIEW

6.1.1 - CHAPTER 2 - AFRICAN SHRUB DISTRIBUTION EMERGES VIA A TRADE-OFF BETWEEN HEIGHT AND SAPWOOD CONDUCTIVITY

We implemented a new trait in aDGVM2 representing the number of stems an individual woody plant has. This trait is controlled by a trade-off between higher potential height growth at lower stem count and higher hydraulic capacity at higher stem count. This methodological approach allows for shrubs, defined as multi-stemmed short woody plants, to emerge dynamically in aDGVM2 as a successful growth form in African savannas. Thus, we can simulate the biogeographical distribution of shrubs according to this definition across African savannas. Concomitantly, our model simulates the biogeographical distribution of the ecological relevance of the trait-based trade-off between height growth and hydraulic conductivity. By not defining plant *a priori*, as in PFT based approaches, we can highlight *a posteriori* which life-forms rely on implemented physiological and ecological mechanisms and where they matter. We envision that such an approach can be extended worldwide to all life-forms, implying a continuum between them and allowing to go beyond the traditional characterisation approach. Specifically, further developments are required to capture processes eluding our approach in the most arid biomes, where shrubs are observed as dominant. This implies that other ecological and physiological processes are more relevant for shrubs in some regions. The specificities of aDGVM2 and our methodological and conceptual approach allow to test the effect of environmental factors on plants more dynamically and flexibly than categorical approaches, and thus is well suited to test shrub encroachment drivers. Finally, our study contributes to understanding and studying life-forms and their associated strategies based on trait trade-offs and enhances our understanding of shrubs *per se*.

6.1.2 - CHAPTER 3 - SHRUB FORM & FUNCTION IN ECOSYSTEMS STRUCTURE & DYNAMICS, SIMULATED WITH ADGVM2 IN AFRICAN SAVANNAS

The aDGVM2 is a trait-based model and thus does not define *a priori* life-forms categories. Therefore, interpreting and benchmarking its results raises specific challenges which require adapted methodology. We characterised shrubs, neither as PFT nor as plant species but as life-form with a distinct biogeographical distribution due to its unique trait combination and life-strategy. Shrubs, according to our characterisation, are well adapted to African savannas, but shrubs characteristics can differ regionally (e.g.: Esler and Rundel, 1999). For example, we found out that our approach does not adequately distinguish shrubs at lower plants height, contributing to explain the discrepancies between our simulations and observations in the most arid regions of Africa.

The use of fire as a management tool against woody cover encroachment, is intensely debated (Trollope, 1980; Smit *et al.*, 2016; Case and Staver, 2017a). Contributing to this discussion, the shrubs emerging from our hypothesis are typically adapted to fire as a resprouter strategy. Shrub encroachment, and woody cover expansion at large, should not be regarded as homogeneous phenomena if they are meant to be accurately studied and efficiently managed. We argue for a consensus on terminology, considering shrub encroachment for plants that are multi-stemmed, short and adapted to fire as resprouters. Finally, our methodological approach concurs to the increasingly documented argument for more holistic and process focused models benchmarking (Luo *et al.*, 2012; Kelley *et al.*, 2013; Collier *et al.*, 2018) while supporting the trend towards trait-based non-categorical study of ecosystems (Lavorel and Garnier, 2002; Laughlin *et al.*, 2012; Pavlick *et al.*, 2013; Scheiter, Langan and Higgins, 2013; Verheijen *et al.*, 2013; Koven *et al.*, 2015; Langan, 2019). Such move could yield long-term benefit in terms of model confidence and capabilities to test scenarios.

6.1.3 - CHAPTER 4 - CHALLENGES AND OPPORTUNITIES OF MODELS INTEGRATING TRAITS AND PROCESSES TO SIMULATE ECOSYSTEMS STATES AND DYNAMICS

Our study produced insights about challenges and opportunities of process-based models. Model uncertainty can come from the very foundation of a model. Due to the concepts and data on which a model relies, there might be a mismatch between what a model is supposed to represent and what it actually represents. Trait-based, non-categorical, models are prone to such mismatch as they diverge from historical definitions of ecological objects (e.g.: species). Such models hold great potential for a more realistic and accurate simulation of ecosystems, yet, they require more careful consideration about their development and benchmarking.

Uncertainty can be introduced in modelling studies from the assessment of models results. Ensuring the meaningfulness of a model benchmarking methodology implies multiple levels of discussion; from the data gathering or creation step to the interpretation of models results. Benchmarking models can be seen as a language translation process. Observation data and models results do not correspond directly and the context and specificities of both need to be accounted for. Therefore, a qualitative assessment of models and their benchmark is as relevant as a quantitative approach.

We argue that more comprehensive and more explicit model development can contribute to addressing uncertainties. Models' foundations in terms of concept and structure should be assessed regarding their aims and purposes. Categorisation of ecosystems should shift towards trait-based approaches in order to be more correctly matched with trait-based models. This is challenging as it implies to re-think typically human-centred understandings of nature (i.e., categorical), for example, by using trait-based approaches (i.e., non-categorical). Concomitantly, adopting holistic benchmarking approaches would better frame models accuracy and extent of relevance. This could be achieved by integrating data experts in model development loops.

6.1.4 - CHAPTER 5 - MODELLING IMPACTS OF ATMOSPHERIC [CO_2], FIRE, GRAZING AND MAP ON SHRUB ENCROACHMENT WITH ADGVM2

The aDGVM2, being the first DGVM able to mechanistically simulate shrubs across African savannas, offers unprecedented opportunities to investigate shrub encroachment. Researching shrub encroachment is highly relevant to inform management policies as it affects vast swaths of lands across the Earth and impacts an array of E.S., from biodiversity and nature conservation to local livelihood and economic activities. Researching shrub encroachment is also an opportunity to further our understanding of ecological processes. Unravelling the ecological dynamics underlying this phenomenon can question ecosystems optimum, dynamic equilibriums and alternative stable states. Investigating such dynamics also furthers our understanding of afforestation and desertification processes and enhance our ability to forecast future scenarios. Finally, we argue that studying shrub encroachment goes along understanding the impact of stress factors on trait space, which is a key ecological question, specifically to further model development.

We tested a matrix of scenarios, covering a range of hypotheses to explain this phenomenon. We considered atmospheric [CO_2], fire, grazing and MAP. We found that elevated atmospheric [CO_2] might suppress shrub encroachment and replace it by woody cover expansion. We simulated that shrub encroachment is more likely at lower MAP. Fire reduces shrub prevalence only at extreme frequency (yearly) or when excluded, while natural fires are not detrimental. Interactions between drivers are complex and non-linear. Therefore, to fully disentangle the relative weight of potential shrub encroachment drivers, further scenarios need to be tested. For those, ecosystems history, transitivity, additional drivers and additional drivers' levels need to be considered. Many more studies will be necessary to fully understand the causes and consequences of shrub encroachment. We argue that these investigations should go along research about shrub, bush and tree life-forms, characterising key traits and trade-offs of various plant types and strategies, driving and supported by model development.

We conclude that shrub encroachment is a protean phenomenon and that consensus among the scientific community is required to ease its investigation and avoid confusion between closely related but distinct phenomenon (i.e., bush and shrub encroachment and woody cover expansion or afforestation). The diversity and complexity of this phenomenon, highlighted by our results, advise to avoid general solutions to manage shrub encroachment. Potentially conflicting stakeholders' interests (e.g.: rangeland farming *vs.* nature conservation?) need to be evaluated against locally specific combination of drivers of shrub encroachment.

We argue that studying shrub encroachment as a global phenomenon and part of the woody cover expansion trend can improve our understanding of ecological processes, specifically about the interaction between stress factors, and is necessary to forecast future shrub encroachment trajectories. We conclude that shrub encroachment management policies should be designed at site to regional scale, considering context specific parameters, avoiding general solutions.

6.2 - BRANCHING STUDIES

6.2.1 - ONGOING STUDY - MODELLING VEGETATION STRUCTURE FOR ECOSYSTEM SERVICES - BENCHMARKING ADGVM2 WITH LIDAR DATA

Building upon aDGVM2′s new capabilities given by our shrub model, we test its ability to represent vegetation structure and use it to translate aDGVM2 simulations into E.S. supply potential. We simulate vegetation structure with aDGVM2 for four parks across southern Africa for which we obtained data about vegetation structure and wildlife tourism (Arbieu *et al.*, 2017) and for four savanna sites in South Africa for which we obtained LIDAR data (Figure 6.1). We benchmark vegetation structure simulated by aDGVM2 against the LIDAR data we obtained. Following arguments from Fisher *et al.* (2010), we constrain our results according to this benchmarking in order to improve their reliability. Similarly, we adjust our results to consider the local specificities of each park site. Then, we translate simulated vegetation structure into E.S. supply potential for wildlife tourism according to the data gathered for each park site. Our methodology is summarized in Figure 6.2. Our conceptual framework and terminology to translate vegetation structure into ES supply potential are given in Figure 6.3. This study answers a gap in ecological modelling studies. E.S., despite being extensively investigated (Boyd and Banzhaf, 2007; Naidoo *et al.*, 2008; Harrington *et al.*, 2010; Egoh *et al.*, 2012; Martín-López *et al.*, 2014; Díaz *et al.*, 2019), are rarely modelled, specifically by DGVMs (Bachelet *et al.*, 2017; Boit *et al.*, 2019; Scheiter, Schulte, *et al.*, 2019). We question:

1. Can we use LIDAR data to assess and adjust aDGVM2's simulation of vegetation structure?
2. Can we translate aDGVM2 simulation of vegetation structure into E.S. supply potential for wildlife tourism?
3. How can our methodology inform decision making for E.S. and park management?

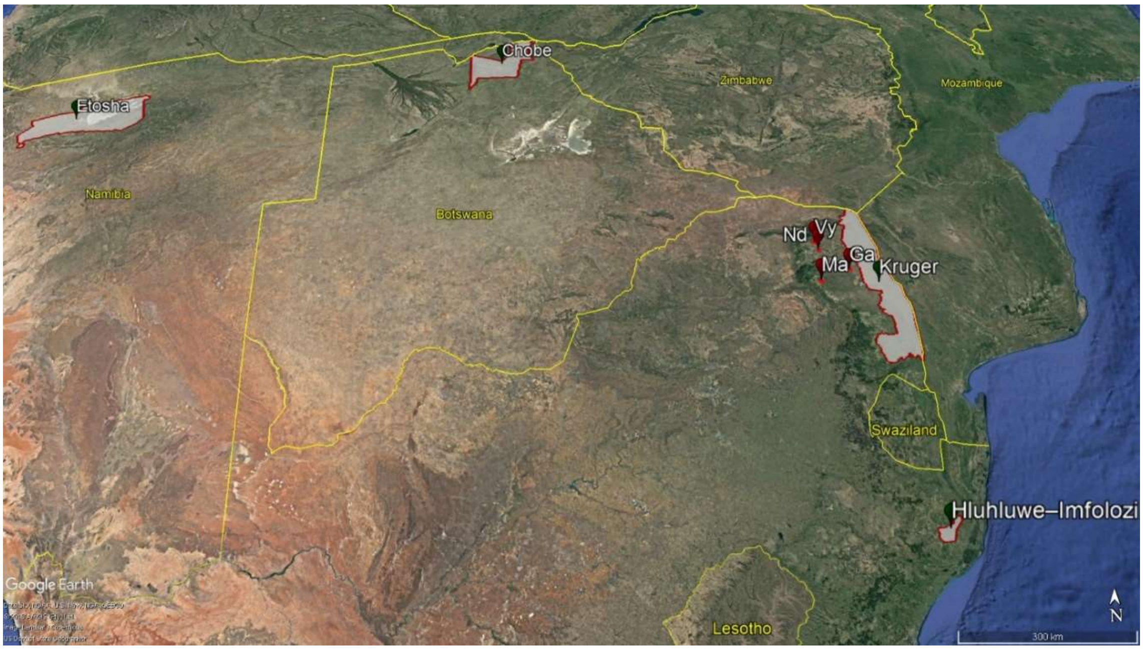


Figure 6.1: Study sites location. Red drops: sites where LIDAR data was obtained. Ga: Gaselwana. Ma: Mafarana. Nd: Ndengeza. Vy: Vyeboom. Green drops: parks where field studies were conducted. Chobe National Park. Etosha National Park. Hluhluwe-Imfolozi Park. Kruger National Park. (Image: Landsat/Copernicus, Data: SIO, NOAA, U.S. Navy, NGA, GEBCO, Image: IBCAo; obtained with Google Earth Pro).

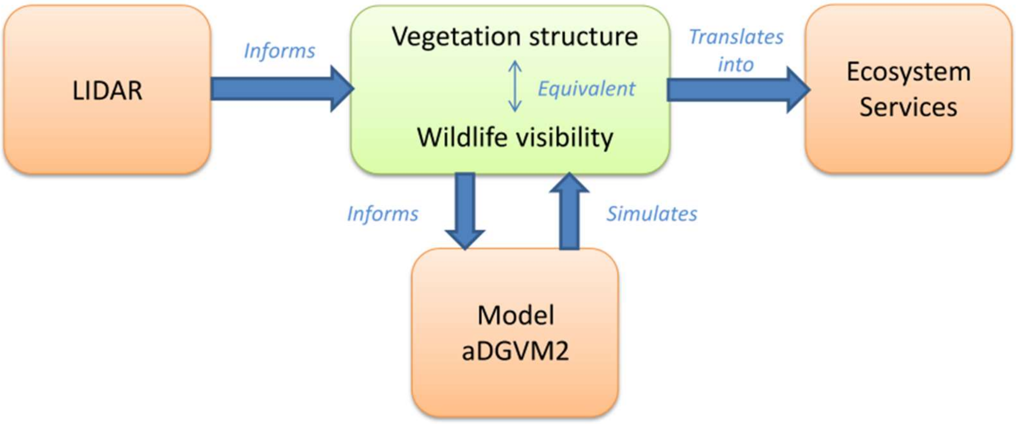


Figure 6.2: Schematic representation of our methodology. We simulate vegetation structure with aDGVM2, which we benchmark and constrain according to LIDAR data. Similarly, we adjust our results to take into account local specificities of each park site. According to observation data from the four park sites investigated, we translate simulated vegetation structure into our target E.S. (wildlife visibility).

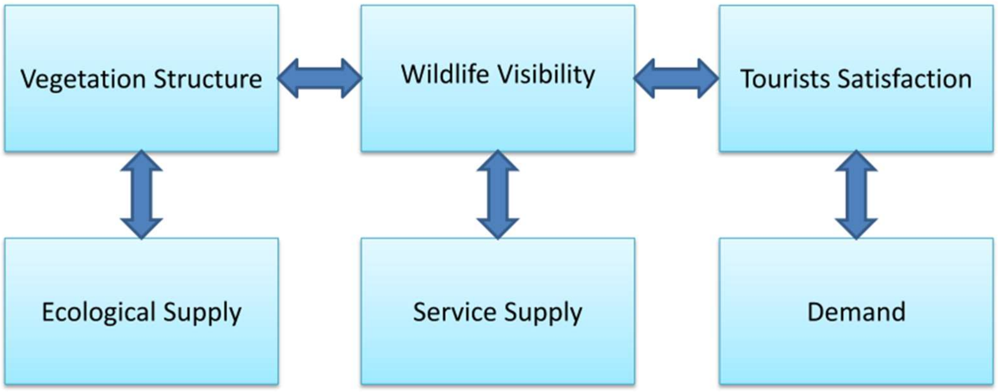


Figure 6.3: Framework of subdivisions of the E.S. concept and terminology. Vegetation structure translates into wildlife visibility which translates into tourists' satisfaction. From a wildlife tourism perspective, vegetation structure is an ecological supply, wildlife visibility is a service supply, tourists' satisfaction is a demand. Therefore, when modelling vegetation structure, we model potential E.S. supply.

6.2.2 - TESTED POTENTIAL ADGVM2'S DEVELOPMENT 1 - CROWN BASE HEIGHT, LIGHT COMPETITION & FIRE AVOIDANCE

The aDGVM2 represents tree crown form based on two shape functions, following the methodology of the first aDGVM (Scheiter and Higgins, 2008) allowing cylindrical to hemispheric geometries from top to bottom of trees. Additionally, leaves are evenly distributed through tree crowns based on LAI. Ignoring the fact that most leaves' photosynthetic activity occurs towards the top of a tree is a shortcoming particularly relevant in two cases. In savannas, fire drives crown base upward as it burns low hanging leaves. In forests, light competition drives crown base upward as lower leaves are more shaded. To account for these processes, we implemented in aDGVM2 a crown base trait and a crown base state parameter defined as a fraction of plant height. Implementing these two elements allowed for variability in crown base height to be simulated by aDGVM2 (results not shown). We further tested possibilities to constrain this trait. We considered a fractional fire effect on leaves which drives tree crown upwards based on fire intensity and frequency (results not shown). Flame height is calculated depending on fire intensity and a fraction of leaf biomass is removed calculated as the leaf biomass fraction included between crown base height and flame height, assuming a homogeneous leaf distribution through the crown. Through aDGVM2 community assembly processes, tall trees with crown base height above typical flame height are selected, while no specific selection pressure affects woody plants smaller than typical flame height as the whole plant burns regularly. Similarly, we devised a scheme for light competition to drive crown base upward (results not shown). When a woody plant crown base height is smaller than another neighbouring plant, the fraction lower than the neighbouring plant base height is shaded using the already existing light competition scheme in aDGVM2. Calibration of the strength of these two effects and benchmarking of their results are limited due to a lack of data. Indeed, extensive and precise LIDAR data would be required for several savanna and forest sites where fire intensity and frequency has been measured. With an increase in LIDAR data availability, we expect to be able to benchmark and parameterise these processes in the near future.

6.2.3 - TESTED POTENTIAL INVESTIGATION 2 - GRASS SELF-SHADING

In African savannas, non-evergreen woody plants shed their leaves when they are dormant, while grasses tend to form a layer of standing dead biomass as they typically do not shed their leaves (Mingo and Oesterheld, 2009). Grasses are key contributors to fires in savannas, notably through their standing biomass (Trollope and Potgieter, 1986; Shea *et al.*, 1996; Ward *et al.*, 1996) whereas fire frequency contributes to grasses coexisting with trees in savannas (San José and Montes, 1997; Higgins, Bond and Trollope, 2000; D’Odorico, Laio and Ridolfi, 2006; Lehsten *et al.*, 2016). The fact that grasses do not shed their dead leaves has the side effect of shading the smaller active leaves (Zimmermann *et al.*, 2010). We simulated this effect with aDGVM2 by adding a pool of standing dead grass biomass, which burns during a fire along with the dead lying grass biomass. This pool contributes preferentially to fuel load as lying biomass decays, lowering fuel load, while standing biomass does not decay, and as standing dead grass biomass dries faster than lying biomass. Standing dead grass leaves become lying biomass after a set time period. We assume that dead standing grass biomass shades grasses homogeneously over each simulated hectare stand, thus lowering indiscriminately photosynthetic activity of each grass individual. We lacked references to parameterise this effect precisely. Benchmarking this effect requires specific data which are to be obtained.

6.2.4 - TESTED POTENTIAL INVESTIGATION 3 - FIRE, SHRUBS & RESPROUTING

Many woody plants have been identified as exhibiting a resprouting strategy, in contrast to reseeding, as an adaptation to fire (Lawes and Clarke, 2011; Pausas and Keeley, 2014), many of them are classified as shrubs (Gratani and Amadori, 1991; Knox and Clarke, 2005; Reyes, Casal and Rego, 2009; Schafer and Mack, 2014). We attempted to represent this strategy by a trade-off between carbon investment in height (to escape flame zone) or fire protection *vs.* carbon investment in storage and roots. Resprouters invest more carbon in storage and roots, thus storing resources, protected from fire, which allow them to rebuild their aboveground biomass quickly after it has been destroyed by fire, by tapping in their stored resources, either right after the fire or during their next normal growth period. Reseeders by investing in height growth and fire protection avoid having their aboveground biomass removed due to fire, either by outgrowing the flame zone or by being protected from critical damages. Benchmarking data showing resprouters rate of regrowth after fire compared to reseeders and data quantifying the abundance of resprouters *vs.* reseeders across African savannas are required to assess the relevance and accuracy of our proposed approach.

6.3 - ENVISIONING FURTHER EXPANSIONS

6.3.1 - SHRUBS & CROWN FIRE

Crown fires affect many woody ecosystems around the world (Pausas *et al.*, 2004) and are even a defining characteristic for some, such as the Fynbos, where they are frequent (Archibald *et al.*, 2013). Crown fires behave differently from ground fire, typically being less frequent and more intense than grass fires (Archibald *et al.*, 2013). Shrubs are particularly prone to and contribute to crown fires as they typically do not escape the flame zone; thus, shrublands can be associated with crown fires (Martins Fernandes, 2001; Zhou, Mahalingam and Weise, 2005; Morvan, 2007; Saglam *et al.*, 2008; Tachajapong *et al.*, 2009; Lozano, 2011). The representation of shrubs we developed for aDGVM2, combined with an improved representation of tree crown architecture as we propose it could be used as a basis to a mechanistic model of crown fires in shrublands. The first step would be to parameterise our shrub model for such shrublands. The second step would be to improve fire representation by including crown fires. To do so would require taking into account shrubs as fuel for crown fire and to implement an additional fire scheme, or to adapt the existing one, taking into account the specificities of crown fires, such as higher flames and fire intensity, based on experimental data. Additionally, subsidiary factors might need to be taken into account, such as topography, influencing fire spread (Lopes, Cruz and Viegas, 2002), or flammability of living leaves, as fire adapted species might also favour fire ignition, spread and intensity.

6.3.2 - HISTORIC BIODIVERSITY TRAJECTORIES: COEVOLUTION CONSTRAINT ON TRAIT SPACE AND ECOSYSTEMS DYNAMICS

Shrubs and fire can be intertwined components characterising an ecosystem, such as in the Fynbos where shrub crown fires are a defining biome trait (Pausas *et al.*, 2004; Archibald *et al.*, 2013). This could be seen as a case of coevolution, where shrubs are adapted to a fire regime which they support and maintain and which is detrimental to their competitors (Possingham, Comins and Noble, 1995; Scheiter *et al.*, 2012). To investigate such eco-evolutionary dynamic, we can build upon our assessment of the potential drivers of shrub encroachments, by studying how stress factors constrain trait space and how a constrained trait space leads to different ecosystem dynamics. At present, aDGVM2 is the most suited DGVM to conduct such investigation as it models each individual as a potentially unique set of traits, which is a necessary ability to study trait space comprehensively. The aDGVM2 can select successful strategies by stochastically sampling trait space depending on prevailing environmental conditions, and depending on vegetation community internal dynamics (e.g.: competition for resources, cross-overs and mutations). Our study on shrub encroachment drivers can serve as a basis to study stress factors constraints on trait space, by investigating how trait space varies depending on simulated scenarios. To explicitly consider coevolution *per se* would require adding alternative scenarios. For example, comparing trait space when trait A is manually constrained from range α to range β and then do the same for trait B, with a focus on trait A and B in the overall trait space. This can be done for any number of traits and traits values through any time length. In the case of shrubs and fire in the Fynbos, instead of testing the effect of artificially forced directional evolution for a given number of traits, we simply would compare the effect of forced fire variables (e.g.: intensity, frequency) and forced shrub variables (e.g.: growth rate, leaf flammability). By assessing the importance of coevolution, ecosystems stability and trait space constraints under constant or changing environmental conditions, such study could yield major results for theoretical ecology and for nature conservation.

6.4 - COMMENTARIES

6.4.1 - LONG LASTING HISTORIC DYNAMICS & NUTRIENT LIMITATION

Human impacts on ecosystems can be traced back to prehistoric periods (Carcaillet, 1998; Yi *et al.*, 2003; Drescher-Schneider *et al.*, 2007; Nosova, Severova and Volkova, 2017; Pini *et al.*, 2017) and have been accelerating over the recent decades globally (Ordway, Asner and Lambin, 2017; Aleman, Jarzyna and Staver, 2018; Díaz *et al.*, 2019). This dynamic is a challenge for ecological models as it impacts their study object while being outside of their scope. An increasing number of research aims at resolving this issue, either by coupling ecology with social and economic sciences (Davies *et al.*, 2016; Scheiter, Schulte, *et al.*, 2019) or by incorporating it directly into ecological models (Forrest *et al.*, 2020) . Particularly challenging are the historical legacies of human impact due to their long-lasting effect shaping ecosystems dynamics. For example, while the Amazonian biodiversity has been managed by local human populations over long time period, even modifying soil structure and nutrients content (Roosevelt, 2013; Fausto and Neves, 2018), it is often mistaken as pristine. Marshall *et al.* (2018) found out that long lasting effects of human land use explain savannas mosaic landscape; fertilization by herds allows productive grass patches to exist and this effect can last centuries or even millennia. Termites mounds have been identified as vegetation drivers in savannas through differential nutrients enrichment and soil structure modification (Nutting, Haverty and Lafage, 1987; Moe, Mobæk and Narmo, 2009; Gosling *et al.*, 2012; Schaefer *et al.*, 2016). While nutrient limitation is increasingly considered by DGVMs (Quillet, Peng and Garneau, 2010), its interaction with long lasting biotic and anthropogenic effects are still missing.

Nutrients distribution is related to current and historic biotic and anthropogenic factors of long-lasting consequences, but also to geological processes of soil formation and rock weathering. Nutrients limitation is a major driver of plant distribution through savannas (Pellegrini, 2016), notwithstanding historical legacy effects. Considering that in savannas tree-grass competition is mediated by fire, while fire frequency depends on grass productivity (Shea *et al.*, 1996), itself shaped by nutrient availability (Craine, Morrow and Stock, 2008), we can argue that soil nutrient distribution is also responsible for tree-grass mosaics in savannas. Both geological soil formation and biotic dissemination of nutrients, such as grazers and browsers behaviours and distribution patterns need to be considered to fully account for nutrients distribution. Adding historic legacy and human activity to the system makes a holistic explicit representation of nutrients cycle and distribution a challenging prospect, yet, it would greatly enhance aDGVM2's capabilities and particularly its representation of savannas mosaics while allowing to study the feedbacks between plants and nutrients dispersers.

6.4.2 - MODEL PARAMETERISATION

The weight of historic legacy, human impact and nutrient limitation varies depending on their interaction, interactions with environmental drivers and plant types. Thus, depending on what a model includes and the target it simulates, a model's fit to observations data should vary. In this regard, studies, such as Zaehle *et al.* (2005) are necessary to provide a global map of the relative importance of multiple factors driving ecosystems states and dynamics. Global traits database, such as the TRY database (Kattge *et al.*, 2011), could also be used as an indicator of the relative weight of different environmental drivers if a link can be established between traits and drivers. The more model operation is geographically constrained the more useful such information can be to improve a model accuracy by adjusting model content *a priori*. Thus, model parameterisation allows a model to produce results closer to observations in a given context (i.e., it improves model results fit with specific benchmarking data).

Model parameterisation also constrains model ability to extrapolate scenarios that are outside of the context for which it has been parameterised. Thus, parameterising models implies a trade-off between context specific accuracy and ability to extrapolate. Through our study, we showed that implementing new ecological processes in a model is a valid alternative to parameterisation with the added benefit of expanding model capabilities and particularly the ability to extrapolate. We showed that our implementation of shrubs can be used across African savannas and woodlands and allows tackling previously out of reach questions, such as shrub encroachment, while improving aDGVM2's representation of vegetation structure. Parameterisation alone could not have achieved similar outcome. Simultaneously, our model development allowed to identify other key processes and parameters for African savannas and closely related biomes, which could be implemented in aDGVM2 to expand its capabilities. For example, plants succulence is relevant for plant adaptation to aridity in many drylands and savannas of Africa (e.g.: high prevalence of Euphorbia and Agave trees) (Conradi *et al.*, 2020). Additionally, specific features can be worth considering, such as rock outcroppings (e.g.: inselberg) or rock fields (e.g.: Karoo/Namib quartz field) (Porembski *et al.*, 1994; Desmet and Cowling, 1999; Schmiedel and Jürgens, 1999; Schmiedel, 2002; Siebert *et al.*, 2003), which lower the vegetation growth potential while mediating hydrology and temperature differently from other ground covers (Nobel, Miller and Graham, 1992; Valentin, 1994). Parameterisation can be seen as a data driven methodology, as it relies on testing parameters value against a fit with benchmarking data, while process implementation could be seen as a an analytical approach, as it requires to derive and test the theoretical understanding of ecosystem dynamics based on observations.

6.4.3 - CATEGORISATIONS IN ECOLOGY

The aDGVM2 is a DGVM which does not define *a priori* categories as a PFT based DGVM would. This unique feature implies specific challenges. While parameterisation can limit model ability to extrapolate, categorisation might limit its accuracy or relevance. Yet, analysing simulations results can require categorisation.

Characterising and classifying objects is consubstantial with human understanding and communication, and different approaches exist to conceptualize this relation. The scientific method followed the historic Western theoretical corpus on this question, which defines objects based on their intrinsic characteristics (Berque, 1987). In contrast, East Asian historic theoretical corpus emphasized the relation between objects and their context to define them. Merging these two perspectives would mean that an object can be characterised due to its intrinsic traits but only in a context framing it. We can envision a new approach to modelling which would rely on studying relations between objects and how they shape ecosystems without focusing specifically on *a priori* defined species.

Using increasingly refined categorisation schemes to better study objects is meaningful and allows to produce relevant analyses and to gain new insights. However, limits between categories can be uncertain, with variability potentially larger within categories than between categories. This uncertainty can be related to the fact that there is a continuum of life-forms. From a theoretical perspective this could indicate that both the Western and East Asian approaches to define objects (Berque, 1987) are actually complementary; i.e., objects are defined *per se* as well as in relation to other objects. This blur can be visualized in trait space, as hypervolumes can be related to successful strategies or species, yet trait space is not necessarily empty outside of these hypervolumes (Cornwell, Schwilk and Ackerly, 2006).

To overcome conceptual challenges related to categorisation and take ecological models to the next level, going beyond *a priori* defined categories offers great opportunities, as it implies to consider what connects objects (here, individual plants), between them and to their surroundings. Defining objects by their interaction means to investigate factors explaining their existence and thus to study processes directly. We demonstrated that aDGVM2 offers great opportunities to contribute to this endeavour. Going further could be achieved by refining model development methodology, model operation assessment and results analysis. It implies to focus on characteristics of objects instead of objects as categories. For example, instead of considering plants as either grasses or trees, we can consider plants lignin content, which implies to consider processes driving lignin content in a plant.

A non-categorical methodology might challenge the historic Western perspective. In the sense of Plato and his ontology, “beings” are or exist by themselves, and only their “image” (which is a model) is embedded in a place (Berque, 1987). This means that an object can be characterised based on intrinsic parameters regardless of its context. Such essentialist perspective could be related to the Western academic tradition to explain the world, which searches for immanent characteristics and rules. From an East Asian perspective, following, for example, Nishida Kitaro (Berque, 1987), objects exist only in relation to their surroundings and are explained and framed by these relations; thus, objects are defined by what they are not, contrarily to the Western perspective which defines objects by what they are. Developing non-categorical modelling approaches might have greater reach than simply being a more holistic benchmarking methodology.

Most DGVMs embed categories to be able to represent them; this is the conceptual basis of the PFT approach. While requiring an increasing number of categories to represent ecosystems, narrower categories can be more accurate and thus improve models' performances. At the same time, more numerous and narrower categories might impair models' ability to extrapolate into simulating states and scenarios outside of their defined ranges. The more specific a model is, the more constrained is its ability to extrapolate. Developing non-categorical models would require considering the continuum and the relations between categories and thus might imply an opposite development path to categorical approaches; by focusing on traits, it reduces the number of required categories. The aDGVM2, is at the forefront of research on this question and paves the way for stimulating science.

6.5 - CONCLUSION

Our investigations concur with the statement by White and Marshall's (2019): "Phenomenological models can have substantial predictive power, unless a prediction outside of the current known set of parameters is required, in which case mechanistic models should be superior". They also stated that "mechanistic models are particularly powerful if sufficient information exists to make *a priori* estimates of parameters and thereby make predictions that are robust to the state of a system." To which we add with our present thesis that *a posteriori* characterisation of model results can also ensure the reliability of a process-based model. The more a model is process focused and non-categorical, the more this statement is relevant. We are also in agreement with the argument by White and Marshall (2019) : "the goodness of fit of model predictions to data alone is not a sufficient test of a model. Free parameters derived from fits to data provide too much flexibility, and additional tests are therefore necessary". We demonstrate that holistic assessment of model results and behaviour is a relevant approach to solve this conundrum, considering multiple parameters and multiple perspective on each of these parameters.

The aDGVM2, being process based and less categorical than DGVMs based on PFTs, has an increased realism and ability to extrapolate which improves the reliability of simulations of future scenarios and which brings greater ecological insights and understandings. Thus, development efforts towards such modelling approach are promising and offer unprecedented research opportunities, having the potential to be more holistic than more categorical models. However, such approach comes with its own challenges, requiring adequate investigation to address them. Our investigations demonstrate the relevance to develop models which allow for life-forms to emerge dynamically, and we argue that such approach should be prioritised over modelling framework relying on increasingly refined *a priori* categories (e.g.: increasing the number of PFTs in a DGVM). The aDGVM2 approach requires more precise investigations of ecological and biological dynamics to support model development, but it allows for more in-depth studies of these processes, while requiring less parameterisation than a categorical approach.

We demonstrate that it is possible to model life-forms and life-strategies based on underlying physiological processes and that it allows to simulate and investigate ecological dynamics. This approach is valid in the framework of a trait-based DGVM (considering plant individuals non-categorically) and allowing for ecosystems to adjust to prevailing environmental conditions through selection processes, as the aDGVM2 exemplifies. We argue that further development focusing on different life-form, plant strategies and drivers of vegetation community assembly would improve model realism and ability to extrapolate, allowing to test more scenarios, while providing ecological insights on life-forms and plant strategies. Such capabilities are needed to answer critical ecological questions related to climate change, biodiversity loss and anthropogenic impacts in general.

REFERENCES

Abramowitz, G. (2012) ‘Towards a public, standardized, diagnostic benchmarking system for land surface models’, *Geoscientific Model Development*, 5(3), pp. 819–827. doi: 10.5194/gmd-5-819-2012.

Adams, J. M. and Faure, H. (1998) ‘A new estimate of changing carbon storage on land since the last glacial maximum, based on global land ecosystem reconstruction’, *Global and Planetary Change*. Elsevier B.V., 16–17, pp. 3–24. doi: 10.1016/S0921-8181(98)00003-4.

Aleman, J. C., Blarquez, O. and Staver, C. A. (2016) ‘Land-use change outweighs projected effects of changing rainfall on tree cover in sub-Saharan Africa’, *Global Change Biology*. Blackwell Publishing Ltd, 22(9), pp. 3013–3025. doi: 10.1111/gcb.13299.

Aleman, J. C., Jarzyna, M. A. and Staver, A. C. (2018) ‘Forest extent and deforestation in tropical Africa since 1900’, *Nature Ecology & Evolution*, 2(1), pp. 26–33. doi: 10.1038/s41559-017-0406-1.

Allen, C. D. *et al.* (2010) ‘A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests’, *Forest Ecology and Management*. Elsevier, 259(4), pp. 660–684. doi: 10.1016/j.foreco.2009.09.001.

Arbieu, U. *et al.* (2017) ‘The importance of vegetation density for tourists’ wildlife viewing experience and satisfaction in African savannah ecosystems’, *PLOS ONE*. Edited by H. Lutermann. Public Library of Science, 12(9), p. e0185793. doi: 10.1371/journal.pone.0185793.

Archer, S. *et al.* (1988) ‘Autogenic Succession in a Subtropical Savanna: Conversion of Grassland to Thorn Woodland’, *Ecological Monographs*, 58(2), pp. 111–127. Available at: http://www.jstor.org/stable/1942463 (Accessed: 26 April 2017).

Archer, S., Boutton, T. W. and Hibbard, K. A. (2001) ‘Trees in Grasslands: Biogeochemical Consequences of Woody Plant Expansion’, in Schulze, E.-D. et al. (eds) *Global Biogeochemical Cycles in the Climate System*. Academic Press, pp. 115–137. doi: 10.1016/B978-012631260-7/50011-X.

Archer, S., Schimel, D. S. and Holland, E. A. (1995) ‘Mechanisms of shrubland expansion: land use, climate or CO2?’, *Climatic Change*, 29(1), pp. 91–99. doi: 10.1007/BF01091640.

Archibald, S. *et al.* (2005) ‘Shaping the landscape: fire-grazer interactions in an African savanna’, *Ecological Applications*. Ecological Society of America, 15(1), pp. 96–109. doi: 10.1890/03-5210.

Archibald, S. *et al.* (2013) ‘Defining pyromes and global syndromes of fire regimes’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 110(16), pp. 6442–6447. doi: 10.1073/pnas.1211466110.

Ardenghi, N. *et al.* (2019) ‘Temperature and moisture variability in the eastern Mediterranean region during Marine Isotope Stages 11–10 based on biomarker analysis of the Tenaghi Philippon peat deposit’, *Quaternary Science Reviews*. Elsevier Ltd, 225, p. 105977. doi: 10.1016/j.quascirev.2019.105977.

Aspinall, R. and Pearson, D. (2000) ‘Integrated geographical assessment of environmental condition in water catchments: Linking landscape ecology, environmental modelling and GIS’, *Journal of Environmental Management*. Academic Press, 59(4), pp. 299–319. doi: 10.1006/jema.2000.0372.

Van Auken, O. W. (2000) ‘Shrub Invasions of North American Semiarid Grasslands’, *Annual Review of Ecology and Systematics*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 31(1), pp. 197–215. doi: 10.1146/annurev.ecolsys.31.1.197.

Van Auken, O. W. (2009) ‘Causes and consequences of woody plant encroachment into western North American grasslands’, *Journal of Environmental Management*. Academic Press, 90(10), pp. 2931–2942. doi: 10.1016/j.jenvman.2009.04.023.

Avitabile, V. *et al.* (2011) ‘Mapping biomass with remote sensing: a comparison of methods for the case study of Uganda’, *Carbon Balance and Management*. BioMed Central, 6(1), p. 7. doi: 10.1186/1750-0680-6-7.

Avitabile, V. *et al.* (2016) ‘An integrated pan-tropical biomass map using multiple reference datasets’, *Global Change Biology*, 22(4), pp. 1406–1420. doi: 10.1111/gcb.13139.

Axelsson, C. R. and Hanan, N. P. (2018) ‘Rates of woody encroachment in African savannas reflect water constraints and fire disturbance’, *Journal of Biogeography*. Wiley/Blackwell (10.1111). doi:

10.1111/jbi.13221.

Baccini, A. *et al.* (2012) ‘Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps’, *Nature Clim. Change*. Nature Publishing Group, 2(3), pp. 182–185. doi: 10.1038/nclimate1354.

Bachelet, D. *et al.* (2017) ‘Translating MC2 DGVM Results into Ecosystem Services for Climate Change Mitigation and Adaptation’, *Climate*. Multidisciplinary Digital Publishing Institute, 6(1), p. 1. doi: 10.3390/cli6010001.

Bachelet, D., Rogers, B. M. and Conklin, D. R. (2015) ‘Challenges and Limitations of Using a DGVM for Local to Regional Applications’, in *Global Vegetation Dynamics: Concepts and Applications in the MC1 Model*. Wiley Blackwell, pp. 31–40. doi: 10.1002/9781119011705.ch3.

Bachmann, K. (1995) ‘Progress and pitfalls in systematics: cladistics, DNA and morphology*’, *Acta Botanica Neerlandica*. John Wiley & Sons, Ltd, 44(4), pp. 403–419. doi: 10.1111/j.1438-8677.1995.tb00795.x.

Báez, S. *et al.* (2013) ‘Effects of experimental rainfall manipulations on Chihuahuan Desert grassland and shrubland plant communities’, *Oecologia*. Springer, 172(4), pp. 1117–1127. doi: 10.1007/s00442-012-2552-0.

Baudrillard, J. (1981) *Simulacres et Simulation*. Editions Galilée.

Baudrillard, J. (1988) ‘Simulacra and Simulations’, in Poster, M. (ed.) *Selected writings*. Stanford: Stanford University Press, pp. 166–184. Available at: https://web.stanford.edu/class/history34q/readings/Baudrillard/Baudrillard_Simulacra.html (Accessed: 6 May 2020).

Baumgartner, S. A. *et al.* (2015) ‘Can diverse herbivore communities increase landscape heterogeneity? Comparing wild and domestic herbivore assemblages in a South African savanna’, *Perspectives in Plant Ecology, Evolution and Systematics*. Elsevier GmbH, 17(1), pp. 34–43. doi: 10.1016/j.ppees.2014.11.002.

Beck, H. E. *et al.* (2018) ‘Present and future köppen-geiger climate classification maps at 1-km resolution’, *Scientific Data*. The Author(s), 5, pp. 1–12. doi: 10.1038/sdata.2018.214.

Berque, A. (1987) *Ecoumène, introduction à l’étude des milieux humains*. Paris: Belin.

Berzaghi, F. *et al.* (2020) ‘Towards a New Generation of Trait-Flexible Vegetation Models’, *Trends in Ecology and Evolution*. Elsevier Ltd, 35(3), pp. 191–205. doi: 10.1016/j.tree.2019.11.006.

Best, M. J. *et al.* (2015) ‘The Plumbing of Land Surface Models: Benchmarking Model Performance’, *Journal of Hydrometeorology*, 16(3), pp. 1425–1442. doi: 10.1175/JHM-D-14-0158.1.

Bicheron P. *et al.* (2008) ‘Globcover Products description and validation report’. Available at: http://www.theia-land.fr/sites/default/files/globcover_products_description_validation_report_i2.1.pdf (Accessed: 11 April 2018).

Biggs, R. *et al.* (2008) ‘Scenarios of biodiversity loss in southern Africa in the 21st century’, *Global Environmental Change*, 18(2), pp. 296–309. doi: 10.1016/j.gloenvcha.2008.02.001.

Blaum, N. *et al.* (2007) ‘Shrub encroachment affects mammalian carnivore abundance and species richness in semiarid rangelands’, *Acta Oecologica*, 31(1), pp. 86–92. doi: 10.1016/j.actao.2006.10.004.

Blum, Walter; Riefler, Werner; Rolandi, L. (2008) *Particle Detection with Drift Chambers*. Springer.

Boafo, Y. A., Saito, O. and Takeuchi, K. (2014) ‘Provisioning Ecosystem Services in Rural Savanna Landscapes of Northern Ghana : An Assessment of Supply , UBoafo, Y. A., Saito, O., & Takeuchi, K. (2014). Provisioning Ecosystem Services in Rural Savanna Landscapes of Northern Ghana : An Assessment of Sup’, *Journal of Disaster Research*. Fuji Technology Press, 9(4)(2014), pp. 501–515. doi: 10.20965/jdr.2014.p0501.

Boehmer, H. J. *et al.* (2013) ‘Rebuilding after collapse: evidence for long-term cohort dynamics in the native Hawaiian rain forest’, *Journal of Vegetation Science*. Edited by D. Ward. John Wiley & Sons, Ltd, 24(4), pp. 639–650. doi: 10.1111/jvs.12000.

Boit, A. *et al.* (2019) ‘Using Dynamic Global Vegetation Models (DGVMs) for Projecting Ecosystem Services at Regional Scales’, in *Atlas of Ecosystem Services*, pp. 57–61. doi: 10.1007/978-3-319-96229-0.

Bond, W. J. (2008) ‘What Limits Trees in C4 Grasslands and Savannas?’, *Annual Review Of Ecology*

Evolution And Systematics, 39, pp. 641–659.

Bond, W. J. and Midgley, G. F. (2012) ‘Carbon dioxide and the uneasy interactions of trees and savannah grasses’, *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367(1588), pp. 601–612. doi: 10.1098/rstb.2011.0182.

Bond, W. J., Woodward, F. I. and Midgley, G. F. (2004) ‘The global distribution of ecosystems in a world without fire’, *New Phytologist*. Blackwell Science Ltd, 165(2), pp. 525–538. doi: 10.1111/j.1469-8137.2004.01252.x.

Bond, W. J., Woodward, F. I. and Midgley, G. F. (2005) ‘The global distribution of ecosystems in a world without fire’, *New Phytologist*, 165(2), pp. 525–537.

Boon, R. *et al.* (2016) ‘Managing a threatened savanna ecosystem (KwaZuluNatal Sandstone Sourveld) in an urban biodiversity hotspot: Durban, South Africa’, *Bothalia*. National Botanical Institute, 46(2), pp. 1–12. doi: 10.4102/abc.v46i2.2112.

Borgman, C. L., Wallis, J. C. and Enyedy, N. (2007) ‘Little science confronts the data deluge: Habitat ecology, embedded sensor networks, and digital libraries’, *International Journal on Digital Libraries*. Springer, 7(1–2), pp. 17–30. doi: 10.1007/s00799-007-0022-9.

Bouvet, A. *et al.* (2018) ‘An above-ground biomass map of African savannahs and woodlands at 25 m resolution derived from ALOS PALSAR’, *Remote Sensing of Environment*. Elsevier, 206, pp. 156–173. doi: 10.1016/J.RSE.2017.12.030.

Boyd, J. and Banzhaf, S. (2007) ‘What are ecosystem services? The need for standardized environmental accounting units’, *Ecological Economics*, 63(2–3), pp. 616–626. doi: 10.1016/j.ecolecon.2007.01.002.

Bradley, K. L. and Pregitzer, K. S. (2007) ‘Ecosystem assembly and terrestrial carbon balance under elevated CO2’, *Trends in Ecology & Evolution*, 22(10), pp. 538–547.

Brink, A. B. and Eva, H. D. (2009) ‘Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach’, *Applied Geography*. Pergamon, 29(4), pp. 501–512. doi: 10.1016/J.APGEOG.2008.10.004.

Brodrick, P. G., Davies, A. B. and Asner, G. P. (2019) ‘Uncovering Ecological Patterns with Convolutional Neural Networks’, *Trends in Ecology and Evolution*. Elsevier Ltd, 34(8), pp. 734–745. doi: 10.1016/j.tree.2019.03.006.

Brooks, D. R. (1990) ‘Parsimony Analysis in Historical Biogeography and Coevolution: Methodological and Theoretical Update’, *Systematic Zoology*. Oxford Academic, 39(1), p. 14. doi: 10.2307/2992205.

Brovkin, V., Ganopolski, A. and Svirezhev, Y. (1997) ‘A continuous climate-vegetation classification for use in climate-biosphere studies’, *Ecological Modelling*, 101(2–3), pp. 251–261.

Bruelheide, H. *et al.* (2018) ‘Global trait–environment relationships of plant communities’, *Nature Ecology and Evolution*. Springer US, 2(12), pp. 1906–1917. doi: 10.1038/s41559-018-0699-8.

Buitenwerf, R. *et al.* (2012) ‘Increased tree densities in South African savannas: 50 years of data suggests CO2 as a driver’, *Global Change Biology*, 18(2), pp. 675–684. doi: 10.1111/j.1365-2486.2011.02561.x.

Burton, C. *et al.* (2019) ‘Representation of fire, land-use change and vegetation dynamics in the Joint UK Land Environment Simulator vn4.9 (JULES)’, *Geoscientific Model Development*, 12(1), pp. 179–193. doi: 10.5194/gmd-12-179-2019.

Butzer, K. W. (2012) ‘Collapse, environment, and society’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 109(10), pp. 3632–3639. doi: 10.1073/pnas.1114845109.

Cadule, P. *et al.* (2010) ‘Benchmarking coupled climate-carbon models against long-term atmospheric CO 2 measurements’, *Global Biogeochemical Cycles*. John Wiley & Sons, Ltd, 24(2), p. n/a-n/a. doi: 10.1029/2009GB003556.

Cambridge University Press (2008) *Cambridge advanced learner’s dictionary*. Cambridge: Cambridge University Press.

Canadell, J., Lloret, F. and López-Soria, L. (1991) ‘Resprouting vigour of two mediterranean shrub species after experimental fire treatments’, *Vegetatio*. Kluwer Academic Publishers, 95(2), pp. 119–126. doi: 10.1007/BF00045210.

Cao, X. *et al.* (2018) 'Mechanisms, monitoring and modeling of shrub encroachment into grassland: a review', *International Journal of Digital Earth*. Taylor & Francis, pp. 1–17. doi: 10.1080/17538947.2018.1478004.

Caracciolo, D. *et al.* (2016) 'Mechanisms of shrub encroachment into Northern Chihuahuan Desert grasslands and impacts of climate change investigated using a cellular automata model', *Advances in Water Resources*. Elsevier Ltd, 91, pp. 46–62. doi: 10.1016/j.advwatres.2016.03.002.

Carcaillet, C. (1998) 'A spatially precise study of Holocene fire history, climate and human impact within the Maurienne valley, North French Alps', *Journal of Ecology*. Blackwell Publishing Ltd, 86(3), pp. 384–396. doi: 10.1046/j.1365-2745.1998.00267.x.

Cardoso Da Silva, J. M. and Bates, J. M. (2002) 'Biogeographic Patterns and Conservation in the South American Cerrado: A Tropical Savanna Hotspot', *BioScience*. Oxford Academic, 52(3), pp. 225–234. doi: 10.1641/0006-3568(2002)052[0225:bpacit]2.0.co;2.

Carius, L. and Findeisen, R. (2016) 'The impact of experimental data quality on computational systems biology and engineering', *IFAC-PapersOnLine*. Elsevier B.V., 49(26), pp. 140–146. doi: 10.1016/j.ifacol.2016.12.116.

Carlson, J. E., Adams, C. A. and Holsinger, K. E. (2016) 'Intraspecific variation in stomatal traits, leaf traits and physiology reflects adaptation along aridity gradients in a South African shrub', *Annals of Botany*, 117(1), pp. 195–207. doi: 10.1093/aob/mcv146.

Carvajal, D. E. *et al.* (2017) 'Population variation in drought-resistance strategies in a desert shrub along an aridity gradient: Interplay between phenotypic plasticity and ecotypic differentiation', *Perspectives in Plant Ecology, Evolution and Systematics*. Urban & Fischer, 29, pp. 12–19. doi: 10.1016/J.PPEES.2017.10.001.

Case, M. F. and Staver, A. C. (2017a) 'Fire prevents woody encroachment only at higher-than-historical frequencies in a South African savanna', *Journal of Applied Ecology*. Edited by J. James, 54(3), pp. 955–962. doi: 10.1111/1365-2664.12805.

Case, M. F. and Staver, A. C. (2017b) 'Fire prevents woody encroachment only at higher than historical frequencies in a South African savanna', *Journal of Applied Ecology*. Edited by J. James, 54(3), pp. 955–962. doi: 10.1111/1365-2664.12805.

Chapin, F. S., Autumn, K. and Pugnaire, F. (1993) 'Evolution of suites of traits in response to environmental stress', *American Naturalist*. University of Chicago Press, 142(Suppl.). doi: 10.1086/285524.

De Châtel, F. (2014) 'The Role of Drought and Climate Change in the Syrian Uprising: Untangling the Triggers of the Revolution', *Middle Eastern Studies*. Routledge, 50(4), pp. 521–535. doi: 10.1080/00263206.2013.850076.

Childs, G. T. (2003) *An introduction to African languages*. John Benjamins Pub. Co.

Chown, S. *et al.* (2010) 'Adapting to climate change: a perspective from evolutionary physiology', *Climate Research*, 43(1), pp. 3–15. doi: 10.3354/cr00879.

Christin, S., Hervet, É. and Lecomte, N. (2019) 'Applications for deep learning in ecology', *Methods in Ecology and Evolution*. Edited by H. Ye. British Ecological Society, 10(10), pp. 1632–1644. doi: 10.1111/2041-210X.13256.

Cipriotti, P. A. *et al.* (2012) 'Understanding the long-term spatial dynamics of a semiarid grass-shrub steppe through inverse parameterization for simulation models', *Oikos*, 121(6), pp. 848–861. doi: 10.1111/j.1600-0706.2012.20317.x.

Clark, D. B. *et al.* (2011) 'The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics', *Geoscientific Model Development*, 4(3), pp. 701–722. doi: 10.5194/gmd-4-701-2011.

Clark, J. D. (1987) 'Transitions: Homo erectus and the Acheulian: the Ethiopian sites of Gadeb and the Middle Awash', *Journal of Human Evolution*. Academic Press, 16(7–8), pp. 809–826. doi: 10.1016/0047-2484(87)90025-X.

Codron, D. *et al.* (2007) 'Nutritional content of savanna plant foods: Implications for browser/grazer models of ungulate diversification', *European Journal of Wildlife Research*. Springer, 53(2), pp. 100–111. doi: 10.1007/s10344-006-0071-1.

Cole, D. N. and Landres, P. B. (1996) 'Threats to Wilderness Ecosystems: Impacts and Research Needs', *Ecological Applications*. Ecological Society of America, 6(1), pp. 168–184. doi: 10.2307/2269562.

Cole, M. M. (1982) 'The Influence of Soils, Geomorphology and Geology on the Distribution of Plant Communities in Savanna Ecosystems', in. Springer, Berlin, Heidelberg, pp. 145–174. doi: 10.1007/978-3-642-68786-0_8.

Collatz, G. J. *et al.* (1991) 'Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer', *Agriculture and Forest Meteorology*, 54, pp. 107–136.

Collatz, G. J., Ribas-Carbo, M. and Berry, J. A. (1992) 'Coupled photosynthesis-stomatal conductance model for leaves of C4 plants', *Australian Journal of Plant Physiology*, 19, pp. 519–538.

Collier, N. *et al.* (2018) 'The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation', *Journal of Advances in Modeling Earth Systems*. Blackwell Publishing Ltd, 10(11), pp. 2731–2754. doi: 10.1029/2018MS001354.

Condit, R., Hubbell, S. P. and Foster, R. B. (1995) 'Mortality Rates of 205 Neotropical Tree and Shrub Species and the Impact of a Severe Drought', *Ecological Monographs*. Ecological Society of America, 65(4), pp. 419–439. doi: 10.2307/2963497.

Conradi, T. *et al.* (2020) 'An operational definition of the biome for global change research', *New Phytologist*. doi: 10.1111/nph.16580.

Corbau, C. *et al.* (2019) 'Quantifying the impacts of the human activities on the evolution of Po delta territory during the last 120 years', *Journal of Environmental Management*. Academic Press, 232, pp. 702–712. doi: 10.1016/j.jenvman.2018.11.096.

Cornwell, W. K., Schwilk, D. W. and Ackerly, D. D. (2006) 'A trait-based test for habitat filtering: convex hull volume', *Ecology*. John Wiley & Sons, Ltd, 87(6), pp. 1465–1471. doi: 10.1890/0012-9658(2006)87[1465:ATTFHF]2.0.CO;2.

Costello, M. J. *et al.* (2013) 'Biodiversity data should be published, cited, and peer reviewed', *Trends in Ecology and Evolution*. Elsevier Current Trends, 28(8), pp. 454–461. doi: 10.1016/j.tree.2013.05.002.

Craine, J. M. and Dybzinski, R. (2013) 'Mechanisms of plant competition for nutrients, water and light', *Functional Ecology*. Edited by D. Robinson, 27(4), pp. 833–840. doi: 10.1111/1365-2435.12081.

Craine, J. M., Morrow, C. and Stock, W. D. (2008) 'Nutrient concentration ratios and co-limitation in South African grasslands', *New Phytologist*, 179(3), pp. 829–836.

Cramer, W. *et al.* (2001) 'Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models', *Global Change Biology*, 7(4), pp. 357–373.

Crisp, M. D., Trewick, S. A. and Cook, L. G. (2011) 'Hypothesis testing in biogeography', *Trends in Ecology and Evolution*. Elsevier Current Trends, 26(2), pp. 66–72. doi: 10.1016/j.tree.2010.11.005.

Crowther, T. W. *et al.* (2015) 'Mapping tree density at a global scale', *Nature*. Nature Publishing Group, 525(7568), pp. 201–205. doi: 10.1038/nature14967.

Cruz, A. *et al.* (2002) 'Resprouting in the Mediterranean-type shrub Erica australis afffected by soil resource availability', *Journal of Vegetation Science*. Opulus Press AB, 13(5), pp. 641–650. doi: 10.1111/j.1654-1103.2002.tb02092.x.

Cruz, A., Perez, B. and Moreno, J. M. (2003) 'Resprouting of the Mediterranean-type shrub Erica australis with modified lignotuber carbohydrate content', *Journal of Ecology*. John Wiley & Sons, Ltd, 91(3), pp. 348–356. doi: 10.1046/j.1365-2745.2003.00770.x.

Cushman, J. C. (2001) 'Crassulacean acid metabolism. A plastic photosynthetic adaptation to arid environments', *Plant Physiology*. American Society of Plant Biologists, 127(4), pp. 1439–1448. doi: 10.1104/pp.010818.

D'Odorico, P., Laio, F. and Ridolfi, L. (2006) 'A probabilistic analysis of fire-induced tree-grass coexistence in savannas', *The American Naturalist*, 167(3), pp. E79--E87.

D'Odorico, P., Okin, G. S. and Bestelmeyer, B. T. (2011) 'A synthetic review of feedbacks and drivers of shrub encroachment in arid grasslands', *Ecohydrology*, 5(5), pp. 520–530. doi: 10.1002/eco.259.

Daryanto, S., Fu, B. and Zhao, W. (2019) ‘Evaluating the use of fire to control shrub encroachment in global drylands: A synthesis based on ecosystem service perspective’, *Science of The Total Environment*. Elsevier, 648, pp. 285–292. doi: 10.1016/J.SCITOTENV.2018.08.140.

Davidson, B. (1966) *Africa in History*. Weidenfeld & Nicolson.

Davies, N. *et al.* (2016) ‘Simulating social-ecological systems: the Island Digital Ecosystem Avatars (IDEA) consortium’, *GigaScience*. Oxford University Press, 5(1), p. 14. doi: 10.1186/s13742-016-0118-5.

DeAngelis, D. L. and Waterhouse, J. C. (1987) ‘Equilibrium and Nonequilibrium Concepts in Ecological Models’, *Ecological Monographs*. John Wiley & Sons, Ltd, 57(1), pp. 1–21. doi: 10.2307/1942636.

Desmet, P. G. and Cowling, R. M. (1999) ‘Biodiversity, habitat and range-size aspects of a flora from a winter-rainfall desert in north-western Namaqualand, South Africa’, *Plant Ecology*. Springer, 142(1–2), pp. 23–33. doi: 10.1023/A:1009814109891.

Diamond, J. M. (2005) *Collapse : how societies choose to fail or succeed*. Viking. Available at: https://web.archive.org/web/20060721195941/http://www.metacritic.com/books/authors/diamondjared/collapse (Accessed: 27 November 2019).

Díaz, S. *et al.* (2019) *IPBES (2019): Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Bonn. Available at: https://ipbes.net/system/tdf/ipbes_global_assessment_report_summary_for_policymakers.pdf?file=1&type=node&id=35329.

Dong, X. and Zhang, X. (2001) ‘Some observations of the adaptations of sandy shrubs to the arid environment in the Mu US Sandland: Leaf water relations and anatomic features’, *Journal of Arid Environments*. Academic Press, 48(1), pp. 41–48. doi: 10.1006/jare.2000.0700.

Donohue, R. J. *et al.* (2013) ‘Impact of CO2 fertilization on maximum foliage cover across the globe’s warm, arid environments’, *Geophysical Research Letters*, 40, pp. 1–5. doi: 10.1002/grl.50563.

Drescher-Schneider, R. *et al.* (2007) ‘Vegetation history, climate and human impact over the last 15,000 years at Lago dell’Accesa (Tuscany, Central Italy)’, *Vegetation History and Archaeobotany*. Springer, 16(4), pp. 279–299. doi: 10.1007/s00334-006-0089-z.

Driscoll, D. A. *et al.* (2010) ‘Fire management for biodiversity conservation: Key research questions and our capacity to answer them’, *Biological Conservation*. Elsevier, 143(9), pp. 1928–1939. doi: 10.1016/J.BIOCON.2010.05.026.

Duarte, L. D. S. *et al.* (2006) ‘Role of nurse plants in Araucaria Forest expansion over grassland in south Brazil’, *Austral Ecology*. Blackwell Publishing Asia, 31(4), pp. 520–528. doi: 10.1111/j.1442-9993.2006.01602.x.

Dybdahl, M. F., Jenkins, C. E. and Nuismer, S. L. (2014) ‘Identifying the molecular basis of host-parasite coevolution: Merging models and mechanisms’, *American Naturalist*. University of Chicago Press, 184(1), pp. 1–13. doi: 10.1086/676591.

Dybzinski, R. *et al.* (2011) ‘Evolutionarily Stable Strategy Carbon Allocation to Foliage, Wood, and Fine Roots in Trees Competing for Light and Nitrogen: An Analytically Tractable, Individual-Based Model and Quantitative Comparisons to Data’, *American Naturalist*, 177(2), pp. 153–166. doi: 10.1086/657992.

Egoh, B. N. *et al.* (2012) ‘An African account of ecosystem service provision: Use, threats and policy options for sustainable livelihoods’, *Ecosystem Services*. Elsevier, 2, pp. 71–81. doi: 10.1016/j.ecoser.2012.09.004.

Ehleringer, J. R., Cerling, T. E. and Helliker, B. R. (1997) ‘C4 photosynthesis, atmospheric CO2 and climate’, *Oecologia*, 112(3), pp. 285–299.

Ehrlich, P. R. and Raven, P. H. (1964) ‘Butterflies and plants: a study in coevolution’, *Evolution*. Wiley-Blackwell, 18(4), pp. 586–608. doi: 10.1111/j.1558-5646.1964.tb01674.x.

Elbadawi, E. and Sambanis, N. (2000) ‘Why are there so many civil wars in Africa? Understanding and preventing violent conflict’, *Journal of African Economies*. Oxford Academic, 9(3), pp. 244–269. doi: 10.1093/jae/9.3.244.

Eldridge, D. J. *et al.* (2011) ‘Impacts of shrub encroachment on ecosystem structure and functioning: towards a global synthesis’, *Ecology Letters*. Blackwell Publishing Ltd, 14(7), pp. 709–722. doi:

10.1111/j.1461-0248.2011.01630.x.

Eldridge, D. J. *et al.* (2012) ‘A global database of shrub encroachment effects on ecosystem structure and functioning’, *Ecology*. Ecological Society of America, 93(11), pp. 2499–2499. doi: 10.1890/12-0749.1.

Eldridge, D. J. *et al.* (2013) ‘Grazing dampens the positive effects of shrub encroachment on ecosystem functions in a semi arid woodland’, *Journal of Applied Ecology*. Edited by J. Firn, 50(4), pp. 1028–1038. doi: 10.1111/1365-2664.12105.

Eldridge, D. J. and Soliveres, S. (2014) ‘Are shrubs really a sign of declining ecosystem function? Disentangling the myths and truths of woody encroachment in Australia’, *Australian Journal of Botany*. CSIRO, 62(7), pp. 594–608. doi: 10.1071/BT14137.

Ellis, E. C. and Ramankutty, N. (2008) ‘Putting people in the map: anthropogenic biomes of the world’, *Frontiers in Ecology and the Environment*. Ecological Society of America, 6(8), pp. 439–447. doi: 10.1890/070062.

Ellison, A. M. (2004) ‘Bayesian inference in ecology’, *Ecology Letters*. John Wiley & Sons, Ltd, 7(6), pp. 509–520. doi: 10.1111/j.1461-0248.2004.00603.x.

Enright, N. J. *et al.* (1998) ‘The ecological significance of canopy seed storage in fire-prone environments: a model for resprouting shrubs’, *Journal of Ecology*. John Wiley & Sons, Ltd, 86(6), pp. 960–973. doi: 10.1046/j.1365-2745.1998.00311.x.

Eriksen, S. E. H. and Watson, H. K. (2009) ‘The dynamic context of southern African savannas: investigating emerging threats and opportunities to sustainability’, *Environmental Science and Policy*. Elsevier, 12(1), pp. 5–22. doi: 10.1016/j.envsci.2008.10.009.

Esler, K. J. and Rundel, P. W. (1999) ‘Comparative patterns of phenology and growth form diversity in two winter rainfall deserts: the Succulent Karoo and Mojave Desert ecosystems’, *Plant Ecology*. Kluwer Academic Publishers, 142(1/2), pp. 97–104. doi: 10.1023/A:1009830513525.

Estabrook, G. F., Johnson, C. S. and Mc Morris, F. R. (1975) ‘An idealized concept of the true cladistic character’, *Mathematical Biosciences*. Elsevier, 23(3–4), pp. 263–272. doi: 10.1016/0025-5564(75)90040-1.

Eyring, V. *et al.* (2016) ‘Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization’, *Geoscientific Model Development*, 9(5), pp. 1937–1958. doi: 10.5194/gmd-9-1937-2016.

Fausto, C. and Neves, E. G. (2018) ‘Was there ever a Neolithic in the Neotropics? Plant familiarisation and biodiversity in the Amazon’, *Antiquity*. Antiquity Publications, 92(366), pp. 1604–1618. doi: 10.15184/aqy.2018.157.

Feurdean, A. *et al.* (2017) ‘Fire has been an important driver of forest dynamics in the Carpathian Mountains during the Holocene’, *Forest Ecology and Management*. Elsevier B.V., 389, pp. 15–26. doi: 10.1016/j.foreco.2016.11.046.

Feurdean, A. *et al.* (2018) ‘Biodiversity-rich European grasslands: Ancient, forgotten ecosystems’, *Biological Conservation*. Elsevier Ltd, 228, pp. 224–232. doi: 10.1016/j.biocon.2018.09.022.

Fischer, R., Rödig, E. and Huth, A. (2018) ‘Consequences of a Reduced Number of Plant Functional Types for the Simulation of Forest Productivity’, *Forests*, 9(8), p. 460. doi: 10.3390/f9080460.

Fisher, R. *et al.* (2010) ‘Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations’, *New Phytologist*. John Wiley & Sons, Ltd (10.1111), 187(3), pp. 666–681. doi: 10.1111/j.1469-8137.2010.03340.x.

Fisher, R. (2010) ‘Modelling vegetation responses to climate change using a second generation dynamic vegetation model : uncertainties due to ecological scale limitations’, (February), p. 2010.

Fisher, R. A. *et al.* (2017) ‘Vegetation demographics in Earth System Models: A review of progress and priorities’, *Global Change Biology*. Wiley/Blackwell (10.1111), 24(1), pp. 35–54. doi: 10.1111/gcb.13910.

Foley, J. A. *et al.* (1996) ‘An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics’, *GLOBAL BIOGEOCHEMICAL CYCLES*, 10(4), pp. 603–628.

Forkel, M. *et al.* (2019) ‘Emergent relationships with respect to burned area in global satellite observations and fire-enabled vegetation models’, *Biogeosciences*, 16(1), pp. 57–76. doi: 10.5194/bg-16-57-2019.

Forrest, M. *et al.* (2020) ‘Including vegetation dynamics in an atmospheric chemistry-enabled general circulation model: linking LPJ-GUESS (v4.0) with the EMAC modelling system (v2.53)’, *Geoscientific Model Development*. Copernicus GmbH, 13(3), pp. 1285–1309. doi: 10.5194/gmd-13-1285-2020.

Forrest, M., Scheiter, S. and Steinkamp, J. (no date) ‘“DGVMTools” R package’.

Frank, K. T. *et al.* (2011) ‘Transient dynamics of an altered large marine ecosystem’, *Nature*. Nature Publishing Group, 477(7362), pp. 86–91. doi: 10.1038/nature10285.

Freckleton, R. P. (2004) ‘The problems of prediction and scale in applied ecology: the example of fire as a management tool’, *Journal of Applied Ecology*. John Wiley & Sons, Ltd, 41(4), pp. 599–603. doi: 10.1111/j.0021-8901.2004.00941.x.

Fresco, L. O. and Kroonenberg, S. B. (1992) ‘Time and spatial scales in ecological sustainability’, *Land Use Policy*. Pergamon, 9(3), pp. 155–168. doi: 10.1016/0264-8377(92)90040-4.

Friedl, M. A. *et al.* (2002) ‘Global land cover mapping from MODIS: Algorithms and early results’, *Remote Sensing of Environment*, 83(1–2), pp. 287–302. doi: 10.1016/S0034-4257(02)00078-0.

Friedl, M. A. *et al.* (2010) ‘MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets’, *Remote Sensing of Environment*. Elsevier, 114(1), pp. 168–182. doi: 10.1016/J.RSE.2009.08.016.

Friend, A. D. *et al.* (1997) ‘A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0)’, *Ecological Modelling*. Elsevier, 95(2–3), pp. 249–287. doi: 10.1016/S0304-3800(96)00034-8.

Fritts, T. H. and Rodda, G. H. (1998) ‘The role of introduced species in the degradation of island ecosystems: A case history of Guam’, *Annual Review of Ecology and Systematics*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 29(1), pp. 113–140. doi: 10.1146/annurev.ecolsys.29.1.113.

Fukami, T. *et al.* (2010) ‘Assembly history dictates ecosystem functioning: evidence from wood decomposer communities’, *Ecology Letters*. Blackwell Publishing Ltd, 13(6), pp. 675–684. doi: 10.1111/j.1461-0248.2010.01465.x.

Fussmann, G. F., Loreau, M. and Abrams, P. A. (2007) ‘Eco-evolutionary dynamics of communities and ecosystems’, *Functional Ecology*. John Wiley & Sons, Ltd, 21(3), pp. 465–477. doi: 10.1111/j.1365-2435.2007.01275.x.

Fynn, R. W. S. and O’Connor, T. G. (2000) ‘Effect of stocking rate and rainfall on rangeland dynamics and cattle performance in a semi-arid savanna, South Africa’, *Journal of Applied Ecology*, 37(3), pp. 491–507. doi: 10.1046/j.1365-2664.2000.00513.x.

Gaillard, C. *et al.* (2018) ‘African shrub distribution emerges via a trade- off between height and sapwood conductivity’, *Journal of Biogeography*, (August), p. Accepted-available online. doi: 10.1111/jbi.13447.

Gałka, M. *et al.* (2018) ‘Response of a spring-fed fen ecosystem in Central Eastern Europe (NW Romania) to climate changes during the last 4000 years: A high resolution multi-proxy reconstruction’, *Palaeogeography, Palaeoclimatology, Palaeoecology*. Elsevier B.V., 504, pp. 170–185. doi: 10.1016/j.palaeo.2018.05.027.

Galli, A. *et al.* (2016) ‘Questioning the Ecological Footprint’, *Ecological Indicators*. Elsevier B.V., 69, pp. 224–232. doi: 10.1016/j.ecolind.2016.04.014.

Gates, D. M. (1968) ‘Transpiration and leaf temperature’, *Annual Review of Plant Physiology*, 19, pp. 211--.

Gaze, S. R. *et al.* (1998) ‘Dry season water use patterns under Guiera senegalensis L. shrubs in a tropical savanna’, *Journal of Arid Environments*. Academic Press, 40(1), pp. 53–67. doi: 10.1006/jare.1998.0435.

Gerard, F. *et al.* (2017) ‘MODIS VCF should not be used to detect discontinuities in tree cover due to binning bias. A comment on Hanan et al. (2014) and Staver and Hansen (2015)’, *Global Ecology and Biogeography*. Blackwell Publishing Ltd, 26(7), pp. 854–859. doi: 10.1111/geb.12592.

Gillson, L. (2004) ‘Testing non-equilibrium theories in savannas: 1400 years of vegetation change in {T}savo {N}ational {P}ark, {K}enya’, *Ecological Complexity*, 1, pp. 281–298.

Gillson, L. (2015) ‘Evidence of a tipping point in a southern African savanna?’, *Ecological Complexity*. Elsevier, 21, pp. 78–86. doi: 10.1016/j.ecocom.2014.12.005.

Giorgetta, M. A. *et al.* (2013) 'Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5', *Journal of Advances in Modeling Earth Systems*. John Wiley & Sons, Ltd, 5(3), pp. 572–597. doi: 10.1002/jame.20038.

Gleick, P. H. and Gleick, P. H. (2014) 'Water, Drought, Climate Change, and Conflict in Syria', *Weather, Climate, and Society*, 6(3), pp. 331–340. doi: 10.1175/WCAS-D-13-00059.1.

Gleiniger, A. *et al.* (2008) *Simulation. Presentation technique and cognitive method*. Edited by A. Gleiniger and G. Vrachliotis. Birkhäuser. doi: https://doi.org/10.1515/9783034609951.

Gobeyn, S. *et al.* (2019) 'Evolutionary algorithms for species distribution modelling: A review in the context of machine learning', *Ecological Modelling*. Elsevier, 392, pp. 179–195. doi: 10.1016/J.ECOLMODEL.2018.11.013.

Gonzalez, P. *et al.* (2010) 'Global patterns in the vulnerability of ecosystems to vegetation shifts due to climate change', *Global Ecology and Biogeography*. John Wiley & Sons, Ltd, 19(6), pp. 755–768. doi: 10.1111/j.1466-8238.2010.00558.x.

Goodale, C. L. and Davidson, E. A. (2002) 'Uncertain sinks in the shrubs', *Nature*. Nature Publishing Group, 418(6898), pp. 593–594. doi: 10.1038/418593a.

Gorai, M. *et al.* (2015) 'Rapid recovery of photosynthesis and water relations following soil drying and re-watering is related to the adaptation of desert shrub Ephedra alata subsp. alenda (Ephedraceae) to arid environments', *Environmental and Experimental Botany*. Elsevier, 109, pp. 113–121. doi: 10.1016/j.envexpbot.2014.08.011.

Gosling, C. M. *et al.* (2012) 'Effects of Erosion from Mounds of Different Termite Genera on Distinct Functional Grassland Types in an African Savannah', *Ecosystems*. Springer, 15(1), pp. 128–139. doi: 10.1007/s10021-011-9497-8.

Götmark, F., Götmark, E. and Jensen, A. M. (2016) 'Why Be a Shrub? A Basic Model and Hypotheses for the Adaptive Values of a Common Growth Form.', *Frontiers in plant science*, 7(July), p. 1095. doi: 10.3389/fpls.2016.01095.

Gould, S. J. (1988) 'Trends as changes in variance: a new slant on progress and directionality in evolution.', *Journal of Paleontology*, 62(3), pp. 319–329. doi: 10.1017/S0022336000059126.

Grace, J. *et al.* (2006) 'Productivity and carbon fluxes of tropical savannas', *Journal of Biogeography*, 33(3), pp. 387–400.

Gratani, L. and Amadori, M. (1991) 'Post-fire resprouting of shrubby species in Mediterranean maquis', *Vegetatio*. Kluwer Academic Publishers, 96(2), pp. 137–143. doi: 10.1007/BF00044975.

Grover, H. D. and Musick, H. B. (1990) 'Shrubland encroachment in southern New Mexico, U.S.A.: An analysis of desertification processes in the American southwest', *Climatic Change*, 17(2–3), pp. 305–330. doi: 10.1007/BF00138373.

Guisan, A. and Thuiller, W. (2005) 'Predicting species distribution: offering more than simple habitat models', *Ecology Letters*, 8(9), pp. 993–1009.

Guo, Q. *et al.* (2020) 'Application of deep learning in ecological resource research: Theories, methods, and challenges', *Science China Earth Sciences*. Science in China Press, pp. 1–18. doi: 10.1007/s11430-019-9584-9.

Hagenah, N. *et al.* (2009) 'Interacting effects of grass height and herbivores on the establishment of an encroaching savanna shrub', in *Herbaceous Plant Ecology: Recent Advances in Plant Ecology*. Springer Netherlands, pp. 189–202. doi: 10.1007/978-90-481-2798-6_15.

Hanan, N. P. *et al.* (2014) 'Analysis of stable states in global savannas: is the cart pulling the horse?', *Global Ecology and Biogeography*, 23(3), pp. 259–263. doi: 10.1111/geb.12122.

Hansen, M. C. *et al.* (2013) 'High-Resolution Global Maps of 21st-Century Forest Cover Change', *Science*, 342(November), pp. 850–854. doi: 10.1126/science.1244693.

Hantson, S. *et al.* (2016) 'The status and challenge of global fire modelling', *Biogeosciences*, 13(11), pp. 3359–3375. doi: 10.5194/bg-13-3359-2016.

Harrington, R. *et al.* (2010) 'Ecosystem services and biodiversity conservation: concepts and a glossary', *Biodiversity and Conservation*. Springer Netherlands, 19(10), pp. 2773–2790. doi: 10.1007/s10531-010-

9834-9.

Hartig, F. *et al.* (2012) ‘Connecting dynamic vegetation models to data - an inverse perspective’, *Journal of Biogeography*, 39, pp. 2240–2252. doi: 10.1111/j.1365-2699.2012.02745.x.

Hastings, A. (2004) ‘Transients: the key to long-term ecological understanding?’, *Trends in Ecology & Evolution*. Elsevier Current Trends, 19(1), pp. 39–45. doi: 10.1016/j.tree.2003.09.007.

Hawkins, J. S., Grover, C. E. and Wendel, J. F. (2008) ‘Repeated big bangs and the expanding universe: Directionality in plant genome size evolution’, *Plant Science*, 174(6), pp. 557–562. doi: 10.1016/j.plantsci.2008.03.015.

Hayat, A. *et al.* (2017) ‘Modeling Tree Growth Taking into Account Carbon Source and Sink Limitations’, *Frontiers in Plant Science*. Frontiers, 8, p. 182. doi: 10.3389/fpls.2017.00182.

He, Y. *et al.* (2011) ‘Coupled land-atmosphere modeling of the effects of shrub encroachment on nighttime temperatures’, *Agricultural and Forest Meteorology*. Elsevier, 151(12), pp. 1690–1697. doi: 10.1016/j.agrformet.2011.07.005.

He, Y. (2014) ‘The effect of precipitation on vegetation cover over three landscape units in a protected semi-arid grassland: Temporal dynamics and suitable climatic index’, *Journal of Arid Environments*. Academic Press, 109, pp. 74–82. doi: 10.1016/j.jaridenv.2014.05.022.

Hendrix, C. S. and Salehyan, I. (2012) ‘Climate change, rainfall, and social conflict in Africa’, *Journal of Peace Research*. SAGE Publications Ltd, 49(1), pp. 35–50. doi: 10.1177/0022343311426165.

Hickler, T. *et al.* (2005) ‘Precipitation controls Sahel greening trend’, *Geophysical Research Letters*, 32(21), pp. 1–4. doi: 10.1029/2005GL024370.

Hickler, T. *et al.* (2006) ‘Implementing plant hydraulic architecture within the LPJ Dynamic Global Vegetation Model’, *Global Ecology and Biogeography*, 15(6), pp. 567–577.

Hickler, T. *et al.* (2008) ‘CO2 fertilization in temperate FACE experiments not representative of boreal and tropical forests’, *Global Change Biology*, 14(7), pp. 1531–1542.

Hickler, T., Rammig, A. and C, W. (2015) ‘Modelling CO2 Impacts on Forest Productivity’, *Current Forestry Reports*, pp. 69–80.

Higgins, S. I. *et al.* (2007) ‘Effects of four decades of fire manipulation on woody vegetation structure in savanna’, *Ecology*, 88(5), pp. 1119–1125. doi: https://doi.org/10.1890/06-1664.

Higgins, S. I. *et al.* (2008) ‘Physically motivated empirical models for the spread and intensity of grass fires’, *International Journal of Wildland Fire*, 17, pp. 595–601.

Higgins, S. I., Bond, W. J. and Trollope, W. S. (2000) ‘Fire, resprouting and variability: a recipe for grass-tree coexistence in savanna’, *Journal of Ecology*, 88, pp. 213–229.

Higgins, S. I. and Scheiter, S. (2012) ‘Atmospheric CO2 forces abrupt vegetation shifts locally, but not globally.’, *Nature*, 488(7410), pp. 209–212.

Hoffman, F. M. *et al.* (2016) *International Land Model Benchmarking (ILAMB) Workshop Report, Technical Report DOE/SC-0186*. Richland, WA (United States). doi: 10.2172/1413397.

Hoffman, M. *et al.* (2016) ‘Biodiversity Hotspots (version 2016.1)’. doi: 10.5281/ZENODO.3261807.

Hoffmann, W. A. *et al.* (2012) ‘Ecological thresholds at the savanna-forest boundary: how plant traits, resources and fire govern the distribution of tropical biomes’, *Ecology Letters*, 15(7), pp. 759–768. doi: 10.1111/j.1461-0248.2012.01789.x.

Hoffmann, W. A., Orthen, B. and Do Nascimento, P. K. V (2003) ‘Comparative fire ecology of tropical savanna and forest trees’, *Functional Ecology*, 17(6), pp. 720–726.

Hoffmann, W. A. and Solbrig, O. T. (2003) ‘The role of topkill in the differential response of savanna woody species to fire’, *Forest Ecology and Management*. Elsevier, 180(1–3), pp. 273–286. doi: 10.1016/S0378-1127(02)00566-2.

Holdo, R. M., Holt, R. D. and Fryxell, J. M. (2009) ‘Grazers, browsers, and fire influence the extent and spatial pattern of tree cover in the Serengeti’, *Ecological Applications*. John Wiley & Sons, Ltd, 19(1), pp. 95–109. doi: 10.1890/07-1954.1.

Hortal, J. *et al.* (2008) ‘Historical bias in biodiversity inventories affects the observed environmental niche of the species', *Oikos*. John Wiley & Sons, Ltd, 117(6), pp. 847–858. doi: 10.1111/j.0030-1299.2008.16434.x.

Hunt, T. L. (2006) 'Rethinking the fall of Easter Island', *American Scientist*, 94(5), pp. 412–419. doi: 10.1511/2006.61.1002.

Huxman, T. E. *et al.* (2005) 'Ecohydrological implications of woody plant encroachment', *Ecology*. Ecological Society of America, 86(2), pp. 308–319. doi: 10.1890/03-0583.

IPCC (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by T. F. Stocker et al. Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2014) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds.), Cambridge University Press, Cambridge, United Kingdom and New York,.

Isbell, F. *et al.* (2015) 'Biodiversity increases the resistance of ecosystem productivity to climate extremes', *Nature*, 526(7574), pp. 574–577. doi: 10.1038/nature15374.

J. Desmond Clark (2008) *The Cambridge History of Africa (the Earliest Times-500 BC)*. Cambridge University press. Available at: http://libgen.is/book/index.php?md5=BDEE1D52DC2BDE4A8EFAC8BFEE34BC2D (Accessed: 23 April 2020).

Jackson, R. B. *et al.* (1996) 'A global analysis of root distributions for terrestrial biomes', *OECOLOGIA*, 108(3), pp. 389–411.

Jacobsen, A. L. *et al.* (2008) 'Comparative community physiology: nonconvergence in water relations among three semi-arid shrub communities', *New Phytologist*. Blackwell Publishing Ltd, 180(1), pp. 100–113. doi: 10.1111/j.1469-8137.2008.02554.x.

Janzen, D. H. (1966) 'Coevolution of mutualism between ants and acacias in Central America', *Evolution*. Wiley, 20(3), pp. 249–275. doi: 10.1111/j.1558-5646.1966.tb03364.x.

Jeltsch, F. *et al.* (1997) 'Analysing Shrub Encroachment in the Southern Kalahari: A Grid-Based Modelling Approach', *The Journal of Applied Ecology*, 34(6), p. 1497. doi: 10.2307/2405265.

Jensen, N. and Wantchekon, L. (2004) 'Resource Wealth and Political Regimes in Africa', *Comparative Political Studies*. Sage PublicationsSage CA: Thousand Oaks, CA, 37(7), pp. 816–841. doi: 10.1177/0010414004266867.

Johnson, S. D. N. and Cox, S. P. (2019) 'Evaluating the role of data quality when sharing information in hierarchical multi-stock assessments, with an application to British Columbia dover sole', *Canadian Journal of Fisheries and Aquatic Sciences*, p. cjfas-2018-0048. doi: 10.1139/cjfas-2018-0048.

Joubert, D. F., Smit, G. N. and Hoffman, M. T. (2012) 'The role of fire in preventing transitions from a grass dominated state to a bush thickened state in arid savannas', *Journal of Arid Environments*, 87, pp. 1–7. doi: 10.1016/j.jaridenv.2012.06.012.

Kambatuku, J. R., Cramer, M. D. and Ward, D. (2011) 'Intraspecific competition between shrubs in a semi-arid savanna', *Plant Ecology*. Springer, 212(4), pp. 701–713. doi: 10.1007/s11258-010-9856-0.

Kaplan, D. A. and Valle, D. (2018) 'A Novel Method for Quantifying Human-Induced Changes in Riverine Hydrology under Nonstationary Climate Conditions using Incomplete Data', *AGUFM*, 2018, pp. H31F-02.

Kattge, J. *et al.* (2011) 'TRY - a global database of plant traits', *Global Change Biology*, 17(9), pp. 2905–2935. doi: 10.1111/j.1365-2486.2011.02451.x.

Keeley, J. E. (1986) 'Resilience of mediterranean shrub communities to fire', in Dell, B., Hopkins, A. J. M., and Lamont, B. B. (eds) *Resilience in mediterranean-type ecosystems*. Springer Netherlands, pp. 95–112.

Kelley, C. P. *et al.* (2015) 'Climate change in the Fertile Crescent and implications of the recent Syrian drought.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 112(11), pp. 3241–6. doi: 10.1073/pnas.1421533112.

Kelley, D. I. *et al.* (2013) ‘A comprehensive benchmarking system for evaluating global vegetation models’, *Biogeosciences*, 10(5), pp. 3313–3340. doi: 10.5194/bg-10-3313-2013.

Kelly, R. D. and Walker, B. H. (1976) ‘The Effects of Different Forms of Land Use on the Ecology of a Semi-Arid Region in South-Eastern Rhodesia’, *Journal of Ecology*, 64(2), pp. 553–576.

Kelly, V. R. and Parker, V. T. (1990) ‘Seed Bank Survival and Dynamics in Sprouting and Nonsprouting Arctostaphylos Species’, *American Midland Naturalist*. JSTOR, 124(1), p. 114. doi: 10.2307/2426084.

Kgope, B. S., Bond, W. J. and Midgley, G. F. (2010) ‘Growth responses of African savanna trees implicate atmospheric [CO2] as a driver of past and current changes in savanna tree cover’, *Austral Ecology*, 35(4), pp. 451–463.

Kicklighter, D. W. *et al.* (1999) ‘A first-order analysis of the potential role of CO2 fertilization to affect the global carbon budget: a comparison of four terrestrial biosphere models’, *Tellus Series B-Chemical and Physical Meteorology*, 51(2), pp. 343–366.

Killingbeck, K. T. and Whitford, W. G. (1996) ‘High Foliar Nitrogen in Desert Shrubs: An Important Ecosystem Trait or Defective Desert Doctrine?’, *Ecology*. Ecological Society of America, 77(6), pp. 1728–1737. doi: 10.2307/2265778.

Kim, D. *et al.* (2019) ‘Improvement of Soil Respiration Parameterization in a Dynamic Global Vegetation Model and Its Impact on the Simulation of Terrestrial Carbon Fluxes’, *Journal of Climate*, 32(1), pp. 127–143. doi: 10.1175/JCLI-D-18-0018.1.

Kimuyu, D. M. *et al.* (2014) ‘Native and domestic browsers and grazers reduce fuels, fire temperatures, and acacia ant mortality in an African savanna’, *Ecological Applications*. Ecological Society of America, 24(4), pp. 741–749. doi: 10.1890/13-1135.1.

Knapp, A. K. *et al.* (2008) ‘Shrub encroachment in North American grasslands: shifts in growth form dominance rapidly alters control of ecosystem carbon inputs’, *Global Change Biology*. Blackwell Publishing Ltd, 14(3), pp. 615–623. doi: 10.1111/j.1365-2486.2007.01512.x.

Knox, K. J. E. and Clarke, P. J. (2005) ‘Nutrient availability induces contrasting allocation and starch formation in resprouting and obligate seeding shrubs’, *Functional Ecology*, 19(4), pp. 690–698. doi: 10.1111/j.1365-2435.2005.01006.x.

Koch, P. L. (1998) ‘Isotopic reconstruction of past continental environments’, *Annual Review of Earth and Planetary Sciences*. Annual Reviews, 26(1), pp. 573–613. doi: 10.1146/annurev.earth.26.1.573.

Kosmala, M. *et al.* (2016) ‘Assessing data quality in citizen science’, *Frontiers in Ecology and the Environment*. Wiley Blackwell, 14(10), pp. 551–560. doi: 10.1002/fee.1436.

Kottek, M. *et al.* (2006) ‘World Map of the Köppen-Geiger climate classification updated MARKUS’, *Meteorologische Zeitschrift,* 15(3), pp. 259–263. doi: 10.1127/0941-2948/2006/0130.

Koven, C. *et al.* (2015) ‘Expanding the use of plant trait observations and ecological theory in Earth system models’, in *DOE Workshop Report*. Rockville.

Kröpelin, S. *et al.* (2008) ‘Climate-driven ecosystem succession in the Sahara: The past 6000 years’, *Science*. American Association for the Advancement of Science, 320(5877), pp. 765–768. doi: 10.1126/science.1154913.

Kumagai, T. *et al.* (2005) ‘Sources of error in estimating stand transpiration using allometric relationships between stem diameter and sapwood area for Cryptomeria japonica and Chamaecyparis obtusa’, *Forest Ecology and Management*, 206(1–3), pp. 191–195. doi: 10.1016/j.foreco.2004.10.066.

Kumar, D., Pfeiffer, M., Gaillard, C., Langan, L. and Scheiter, S. (2020) ‘Climate change and elevated CO2 favor forest over savanna under different future scenarios in South Asia’, *Biogeosciences*, in review. Available at: https://www.biogeosciences-discuss.net/bg-2020-169/ (Accessed: 3 June 2020).

Kumar, D., Pfeiffer, M., Gaillard, C., Langan, L., Martens, C., *et al.* (2020) ‘Misinterpretation of Asian savannas as degraded forest can mislead management and conservation policy under climate change’, *Biological Conservation*. Elsevier Ltd, 241, p. 108293. doi: 10.1016/j.biocon.2019.108293.

Kumar, D. and Scheiter, S. (2019) ‘Biome diversity in South Asia - How can we improve vegetation models to understand global change impact at regional level?’, *Science of the Total Environment*. Elsevier B.V., 671, pp. 1001–1016. doi: 10.1016/j.scitotenv.2019.03.251.

de la Riva, E. G. *et al.* (2018) ‘A Multidimensional Functional Trait Approach Reveals the Imprint of Environmental Stress in Mediterranean Woody Communities’, *Ecosystems*. Springer New York LLC, 21(2), pp. 248–262. doi: 10.1007/s10021-017-0147-7.

Laliberté, E. *et al.* (2013) ‘A long-term experimental test of the dynamic equilibrium model of species diversity’, *Oecologia*. Springer, 171(2), pp. 439–448. doi: 10.1007/s00442-012-2417-6.

Lamanna, C. *et al.* (2014) ‘Functional trait space and the latitudinal diversity gradient’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 111(38), pp. 13745–13750. doi: 10.1073/pnas.1317722111.

Langan, L., Higgins, S. I. and Scheiter, S. (2017) ‘Climate-biomes, pedo-biomes or pyro-biomes: which world view explains the tropical forest-savanna boundary in South America?’, *Journal of Biogeography*, 44(10), pp. 2319–2330. doi: 10.1111/jbi.13018.

Langan, L. J. (2019) *Holism in plant biogeography - improving the representation of , and interactions between , the biosphere , hydrosphere , atmosphere and pedosphere*. Goethe-universität, Johann Wolfgang.

Laughlin, D. C. *et al.* (2012) ‘A predictive model of community assembly that incorporates intraspecific trait variation.’, *Ecology Letters*, 15(11), pp. 1291–1299. doi: 10.1111/j.1461-0248.2012.01852.x.

Laurent, P. *et al.* (2019) ‘Varying relationships between fire radiative power and fire size at a global scale’, *Biogeosciences*, 16(2), pp. 275–288. doi: 10.5194/bg-16-275-2019.

Lavorel, S. and Garnier, E. (2002) ‘Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail’, *Funcitional Ecology*, 16, pp. 545–556. doi: 10.1046/j.1365-2435.2002.00664.x.

Lawes, M. J. *et al.* (2007) ‘The effect of the spatial scale of recruitment on tree diversity in Afromontane forest fragments’, *Biological Conservation*, 139(3–4), pp. 447–456. doi: 10.1016/j.biocon.2007.07.016.

Lawes, M. J. and Clarke, P. J. (2011) ‘Ecology of plant resprouting: Populations to community responses in fire-prone ecosystems’, *Plant Ecology*, 212(12), pp. 1937–1943. doi: 10.1007/s11258-011-9994-z.

Lawley, V. *et al.* (2013) ‘Self-organization and complex dynamics of regenerating vegetation in an arid ecosystem: 82 years of recovery after grazing’, *Journal of Arid Environments*. Elsevier Ltd, 88, pp. 156–164. doi: 10.1016/j.jaridenv.2012.08.014.

Lawrence, D. M. *et al.* (2011) ‘Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model’, *Journal of Advances in Modeling Earth Systems*, 3(1). doi: 10.1029/2011MS00045.

LeBauer, D. S. *et al.* (2013) ‘Facilitating feedbacks between field measurements and ecosystem models’, *Ecological Monographs*. John Wiley & Sons, Ltd, 83(2), pp. 133–154. doi: 10.1890/12-0137.1.

Lehmann, C. E. R. *et al.* (2011) ‘Deciphering the distribution of the savanna biome’, *New Phytologist*, 191(1), pp. 197–209. doi: 10.1111/j.1469-8137.2011.03689.x.

Lehmann, C. E. R. *et al.* (2014) ‘Savanna Vegetation-Fire-Climate Relationships Differ Among Continents’, *Science*, 343(6170), pp. 548–552. doi: 10.1126/science.1247355.

Lehsten, V. *et al.* (2016) ‘The effect of fire on tree-grass coexistence in savannas: A simulation study’, *International Journal of Wildland Fire*, 25(2). doi: 10.1071/WF14205.

Lek, S. *et al.* (1996) ‘Application of neural networks to modelling nonlinear relationships in ecology’, *Ecological Modelling*. Elsevier B.V., 90(1), pp. 39–52. doi: 10.1016/0304-3800(95)00142-5.

Lek, S. and Guégan, J. F. (1999) ‘Artificial neural networks as a tool in ecological modelling, an introduction’, *Ecological Modelling*. Elsevier, 120(2–3), pp. 65–73. doi: 10.1016/S0304-3800(99)00092-7.

Lenton, T. M. *et al.* (2008) ‘Tipping elements in the Earth’s climate system’, *Proceedings of the National Academy of Sciences of the United States of America*, 105(6), pp. 1786–1793.

Lenton, T. M. (2011) ‘Early warning of climate tipping points’, *Nature Climate Change*, 1, pp. 201–209. doi: 10.1038/NCLIMATE1143.

Lenzen, M. *et al.* (2012) ‘International trade drives biodiversity threats in developing nations’, *Nature*. Nature Publishing Group, 486(7401), pp. 109–112. doi: 10.1038/nature11145.

Leonelli, S. (2017) ‘Global Data Quality Assessment and the Situated Nature of “Best” Research Practices in Biology', *Data Science Journal*, 16. doi: 10.5334/dsj-2017-032.

Lett, M. S. *et al.* (2004) 'Influence of shrub encroachment on aboveground net primary productivity and carbon and nitrogen pools in a mesic grassland', *Canadian journal of botany*. NRC Research Press, 82(9), pp. 1363–1370. Available at: http://www.ingentaconnect.com/content/nrc/cjb/2004/00000082/00000009/art00013 (Accessed: 12 May 2017).

Lewis, M. P., Simons, G. F. and Fennig, C. D. (2016) *Ethnologue : languages of africa and europe*. SIL International Publications.

Li, F. *et al.* (2011) 'Impact of spin-up forcing on vegetation states simulated by a dynamic global vegetation model coupled with a land surface model', *Advances in Atmospheric Sciences*. Springer, 28(4), pp. 775–788. doi: 10.1007/s00376-010-0009-0.

Li, F. *et al.* (2019) 'Historical (1700-2012) global multi-model estimates of the fire emissions from the Fire Modeling Intercomparison Project (FireMIP)', *Atmospheric Chemistry and Physics*. Copernicus GmbH, 19(19), pp. 12545–12567. doi: 10.5194/acp-19-12545-2019.

Li, F., Zeng, X. D. and Levis, S. (2012) 'A process-based fire parameterization of intermediate complexity in a Dynamic Global Vegetation Model', *Biogeosciences*, 9(7), pp. 2761–2780. doi: 10.5194/bg-9-2761-2012.

Li, W. *et al.* (2018) 'Gross changes in forest area shape the future carbon balance of tropical forests', *Biogeosciences*, 15(1), pp. 91–103. doi: 10.5194/bg-15-91-2018.

Li, X. *et al.* (2016) 'Estimation of gross primary production in China (1982–2010) with multiple ecosystem models', *Ecological Modelling*. Elsevier, 324, pp. 33–44. doi: 10.1016/J.ECOLMODEL.2015.12.019.

Li, Y., Brimicombe, A. J. and Ralphs, M. P. (2000) 'Spatial data quality and sensitivity analysis in GIS and environmental modelling: The case of coastal oil spills', *Computers, Environment and Urban Systems*. Pergamon, 24(2), pp. 95–108. doi: 10.1016/S0198-9715(99)00048-4.

Lieberman, D. E. *et al.* (2007) 'The evolution of endurance running and the tyranny of ethnography: A reply to Pickering and Bunn (2007)', *Journal of Human Evolution*, 53(4), pp. 434–437. doi: 10.1016/j.jhevol.2007.01.012.

Lima, R. A. F. *et al.* (2016) 'How do size distributions relate to concurrently measured demographic rates? Evidence from over 150 tree species in Panama', *Journal of Tropical Ecology*. Cambridge University Press, 32(03), pp. 179–192. doi: 10.1017/S0266467416000146.

Linderman, M. *et al.* (2004) 'Using artificial neural networks to map the spatial distribution of understorey bamboo from remote sensing data', *International Journal of Remote Sensing*. Taylor & Francis Group, 25(9), pp. 1685–1700. doi: 10.1080/01431160310001598971.

Liu, Z. *et al.* (2018) 'Application of machine-learning methods in forest ecology: Recent progress and future challenges', *Environmental Reviews*. liu2018: Canadian Science Publishing, pp. 339–350. doi: 10.1139/er-2018-0034.

Lloyd, J. *et al.* (2008) 'Contributions of woody and herbaceous vegetation to tropical savanna ecosystem productivity: a quasi-global estimate', *Tree Physiology*. Oxford Academic, 28(3), pp. 451–468. doi: 10.1093/treephys/28.3.451.

Loew, A. *et al.* (2013) 'Do we (need to) care about canopy radiation schemes in DGVMs? An evaluation and assessment study', *Biogeosciences Discussions*, 10(10), pp. 16551–16613. doi: 10.5194/bgd-10-16551-2013.

Lohmann, D. *et al.* (2014) 'Prescribed fire as a tool for managing shrub encroachment in semi-arid savanna rangelands', *Journal of Arid Environments*. Academic Press, 107, pp. 49–56. doi: 10.1016/J.JARIDENV.2014.04.003.

Longo, M. *et al.* (2019) 'The biophysics, ecology, and biogeochemistry of functionally diverse, vertically- and horizontally-heterogeneous ecosystems: the Ecosystem Demography Model, version 2.2 — Part 1: Model description', *Geoscientific Model Development Discussions*. Copernicus GmbH, pp. 1–53. doi: 10.5194/gmd-2019-45.

Lopes, A. M. G., Cruz, M. G. and Viegas, D. X. (2002) 'Firestation - An integrated software system for the numerical simulation of fire spread on complex topography', *Environmental Modelling and Software*.

Elsevier BV, 17(3), pp. 269–285. doi: 10.1016/S1364-8152(01)00072-X.

Loreau, M. *et al.* (2001) ‘Biodiversity and ecosystem functioning: Current knowledge and future challenges’, *Science*, 294(5543), pp. 804–808. doi: 10.1126/science.1064088.

Lotka, A. J. (1927) ‘Fluctuations in the abundance of a species considered mathematically’, *Nature*. Nature Publishing Group, 119(2983), p. 12. doi: 10.1038/119012a0.

Lozano, J. S. (2011) *An Investigation of Surface and Crown Fire Dynamics in Shrub Fuels*. Univesity of California Riverside.

Lunau, K. (2004) ‘Adaptive radiation and coevolution - Pollination biology case studies’, *Organisms Diversity and Evolution*. No longer published by Elsevier, 4(3), pp. 207–224. doi: 10.1016/j.ode.2004.02.002.

Luo, Y. Q. *et al.* (2012) ‘A framework for benchmarking land models’, *Biogeosciences*. doi: 10.5194/bg-9-3857-2012.

Ma, M. and Frank, V. (2006) ‘Interannual variability of vegetation cover in the Chinese Heihe River Basin and its relation to meteorological parameters’, *International Journal of Remote Sensing*. Taylor and Francis Ltd., 27(16), pp. 3473–3486. doi: 10.1080/01431160600593031.

Ma, P. *et al.* (2019) ‘Exploring the relative importance of biotic and abiotic factors that alter the self-thinning rule: Insights from individual-based modelling and machine-learning’, *Ecological Modelling*. Elsevier, 397(January), pp. 16–24. doi: 10.1016/j.ecolmodel.2019.01.019.

MacLeod, N. D., Brown, J. R. and Noble, J. C. (1993) ‘Ecological and economic considerations for the management of shrub encroachment in Australian rangelands.’, *Proceedings II of the 10th Australian Weeds Conference and 14th Asian Pacific Weed Science Society Conference, Brisbane, Australia, 6-10 September, 1993*. Queensland Weed Society, pp. 118–121.

Maestre, F. T. *et al.* (2009) ‘Shrub encroachment can reverse desertification in semi-arid Mediterranean grasslands’, *Ecology Letters*. John Wiley & Sons, Ltd, 12(9), pp. 930–941. doi: 10.1111/j.1461-0248.2009.01352.x.

Mäkelä, A. A. and Sievanen, R. P. (1987) ‘Comparison of two shoot root partitioning models with respect to substrate utilization and functional balance’, *Annals of Botany*, 59(2), pp. 129–140.

Marignani, M. *et al.* (2017) ‘Natural and human impact in Mediterranean landscapes: An intriguing puzzle or only a question of time?’, *Plant Biosystems*. Taylor and Francis Ltd., 151(5), pp. 900–905. doi: 10.1080/11263504.2016.1244121.

Maron, M. *et al.* (2017) ‘Towards a Threat Assessment Framework for Ecosystem Services’, *Trends in Ecology and Evolution*. Elsevier Ltd, 32(4), pp. 240–248. doi: 10.1016/j.tree.2016.12.011.

Marshall, F. *et al.* (2018) ‘Ancient herders enriched and restructured African grasslands’, *Nature*. Nature Publishing Group, 561(7723), pp. 387–390. doi: 10.1038/s41586-018-0456-9.

Martín-López, B. *et al.* (2014) ‘Trade-offs across value-domains in ecosystem services assessment’, *Ecological Indicators*. Elsevier, 37, pp. 220–228. doi: 10.1016/J.ECOLIND.2013.03.003.

Martin, L. J., Blossey, B. and Ellis, E. (2012) ‘Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations’, *Frontiers in Ecology and the Environment*. John Wiley & Sons, Ltd, 10(4), pp. 195–201. doi: 10.1890/110154.

Martínez-Cabrera, H. I. *et al.* (2011) ‘Integration of vessel traits, wood density, and height in angiosperm shrubs and trees’, *American Journal of Botany*. John Wiley & Sons, Ltd, 98(5), pp. 915–922. doi: 10.3732/ajb.1000335.

Martins Fernandes, P. A. (2001) ‘Fire spread prediction in shrub fuels in Portugal’, *Forest Ecology and Management*. Elsevier, 144(1–3), pp. 67–74. doi: 10.1016/S0378-1127(00)00363-7.

Mason, N. W. H. and Pavoine, S. (2013) ‘Does trait conservatism guarantee that indicators of phylogenetic community structure will reveal niche-based assembly processes along stress gradients?’, *Journal of Vegetation Science*. Edited by V. Pillar. John Wiley & Sons, Ltd, 24(5), pp. 820–833. doi: 10.1111/jvs.12033.

Mather, A. S. and Needle, C. L. (1998) ‘The forest transition: a theoretical basis’, *Area*. Blackwell Publishing Ltd, 30(2), pp. 117–124. doi: 10.1111/j.1475-4762.1998.tb00055.x.

Mayaux, P. *et al.* (2004) ‘A new land-cover map of Africa for the year 2000’, *Journal of Biogeography*, 31(6), pp. 861–877.

Mayaux, P. *et al.* (2006) ‘Validation of the global land cover 2000 map’, *IEEE Transactions on Geoscience and Remote Sensing*, 44(7), pp. 1728–1739. doi: 10.1109/TGRS.2006.864370.

McCree, K. J. and Troughton, J. H. (1966) ‘Prediction of Growth Rate at Different Light Levels from Measured Photosynthesis and Respiration Rates’, *Plant Physiology*. American Society of Plant Biologists (ASPB), 41(4), pp. 559–566. doi: 10.1104/pp.41.4.559.

McDougall, I., Brown, F. H. and Fleagle, J. G. (2005) ‘Stratigraphic placement and age of modern humans from Kibish, Ethiopia’, *Nature*. Nature Publishing Group, 433(7027), pp. 733–736. doi: 10.1038/nature03258.

McKee, J. K. *et al.* (2004) ‘Forecasting global biodiversity threats associated with human population growth’, *Biological Conservation*. Elsevier BV, 115(1), pp. 161–164. doi: 10.1016/S0006-3207(03)00099-5.

McSweeney, C. F. and Jones, R. G. (2016) ‘How representative is the spread of climate projections from the 5 CMIP5 GCMs used in ISI-MIP?’, *Climate Services*. Elsevier, 1, pp. 24–29. doi: 10.1016/J.CLISER.2016.02.001.

Medlyn, B. E. *et al.* (2015) ‘Using ecosystem experiments to improve vegetation models’, *Nature Climate Change*. Nature Publishing Group, 5(6), pp. 528–534. doi: 10.1038/nclimate2621.

Mehta and Pankaj (2019) ‘Machine Learning, Statistical Physics, and Ecological Dynamics’, *Bulletin of the American Physical Society*. American Physical Society. Available at: http://meetings.aps.org/Meeting/MAR19/Session/R51.1 (Accessed: 21 February 2019).

Meinshausen, M. *et al.* (2011) ‘The RCP greenhouse gas concentrations and their extensions from 1765 to 2300’, *Climatic Change*. Springer Netherlands, 109(1–2), pp. 213–241. doi: 10.1007/s10584-011-0156-z.

Merino, J., Field, C. and Mooney, H. A. (1982) ‘Construction and maintenance costs of mediterranean-climate evergreen and deciduous leaves I. Growth and CO2 exchange analysis’, *Oecologia*. Springer-Verlag, 53(2), pp. 208–213. doi: 10.1007/BF00545665.

De Micco, V., Aronne, G. and Baas, P. (2008) ‘Wood anatomy and hydraulic architecture of stems and twigs of some Mediterranean trees and shrubs along a mesic xeric gradient’, *Trees*. Springer-Verlag, 22(5), pp. 643–655. doi: 10.1007/s00468-008-0222-y.

Michener, W. K. and Jones, M. B. (2012) ‘Ecoinformatics: Supporting ecology as a data-intensive science’, *Trends in Ecology and Evolution*. Elsevier Ltd, 27(2), pp. 85–93. doi: 10.1016/j.tree.2011.11.016.

Midgley, G. F. and Bond, W. J. (2015) ‘Future of African terrestrial biodiversity and ecosystems under anthropogenic climate change’, *Nature Clim. Change*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 5(9), pp. 823–829. Available at: http://dx.doi.org/10.1038/nclimate2753.

Midgley, J. J. (1996) ‘Why the world’s vegetation is not totally dominated by resprouting plants; because resprouters are shorter than reseeders’, *Ecography*, 19(1), pp. 92–94. doi: 10.1111/j.1600-0587.1996.tb00159.x.

Mingo, A. and Oesterheld, M. (2009) ‘Retention of dead leaves by grasses as a defense against herbivores. A test on the palatable grass Paspalum dilatatum’, *Oikos*. John Wiley & Sons, Ltd, 118(5), pp. 753–757. doi: 10.1111/j.1600-0706.2008.17293.x.

Van Minnen, J. G., Leemans, R. and Ihle, F. (2000) ‘Defining the importance of including transient ecosystem responses to simulate C-cycle dynamics in a global change model’, *Global Change Biology*. John Wiley & Sons, Ltd, 6(6), pp. 595–611. doi: 10.1046/j.1365-2486.2000.00323.x.

Mitchard, E. T. A. *et al.* (2009) ‘Measuring Woody Encroachment along a Forest–Savanna Boundary in Central Africa’, *Earth Interactions*, 13(8), pp. 1–29. doi: 10.1175/2009EI278.1.

Mitchard, E. T. A. *et al.* (2013) ‘Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps.’, *Carbon balance and management*. Carbon Balance and Management, 8(1), p. 10. doi: 10.1186/1750-0680-8-10.

Moe, S. R., Mobæk, R. and Narmo, A. K. (2009) ‘Mound building termites contribute to savanna vegetation heterogeneity’, *Plant Ecology*. Springer, 202(1), pp. 31–40. doi: 10.1007/s11258-009-9575-6.

Mograbi, P. J. *et al.* (2015) ‘Biomass Increases Go under Cover: Woody Vegetation Dynamics in South African Rangelands’, *PloS ONE*, 10(5). doi: 10.1371/journal.pone.0127093.

Moilanen, A. (2002) ‘Implications of empirical data quality to metapopulation model parameter estimation and application’, *Oikos*. John Wiley & Sons, Ltd, 96(3), pp. 516–530. doi: 10.1034/j.1600-0706.2002.960313.x.

Monahan, W. B. (2009) ‘A mechanistic niche model for measuring species’ distributional responses to seasonal temperature gradients’, *PLoS ONE*. Public Library of Science, 4(11). doi: 10.1371/journal.pone.0007921.

Moncrieff, G. R. *et al.* (2015) ‘Understanding global change impacts on South African biomes using Dynamic Vegetation Models’, *South African Journal of Botany*. Elsevier B.V., 101, pp. 16–23. doi: 10.1016/j.sajb.2015.02.004.

Moncrieff, G. R. *et al.* (2016) ‘The future distribution of the savannah biome: model-based and biogeographic contingency’, *Philosophical Transactions of the Royal Society B-Biological Sciences*, 371(1703). doi: 10.1098/rstb.2015.0311.

Monserud, R. A. and Leemans, R. (1992) ‘Comparing global vegetation maps with the kappa-statistic’, *Ecological Modelling*, 62(4), pp. 275–293.

Mooney, H. *et al.* (2009) ‘Biodiversity, climate change, and ecosystem services’, *Current Opinion in Environmental Sustainability*. Elsevier, 1(1), pp. 46–54. doi: 10.1016/j.cosust.2009.07.006.

Moore, J. R. *et al.* (2018) ‘Equilibrium forest demography explains the distribution of tree sizes across North America’, *Environmental Research Letters*. IOP Publishing, 13(8), p. 084019. doi: 10.1088/1748-9326/aad6d1.

Morvan, D. (2007) ‘A numerical study of flame geometry and potential for crown fire initiation for a wildfire propagating through shrub fuel’, *International Journal of Wildland Fire*. CSIRO PUBLISHING, 16(5), p. 511. doi: 10.1071/WF06010.

Mucina, L. and Rutherford, M. C. (2006) *The Vegetation of South Africa, Lesotho and Swaziland*. Edited by L. Mucina and Michael C. Rutherford. Pretoria: SANBI. Available at: https://www.researchgate.net/publication/46888917 (Accessed: 12 May 2018).

Mugasi, S. K., Sabiiti, E. N. and Tayebwa, B. M. (2000) ‘The economic implications of bush encroachment on livestock farming in rangelands of uganda’, *African Journal of Range and Forage Science*. Taylor & Francis Group, 17(1–3), pp. 64–69. doi: 10.2989/10220110009485741.

Muller-Landau, H. C. *et al.* (2006) ‘Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models’, *Ecology Letters*. John Wiley & Sons, Ltd (10.1111), 9(5), pp. 589–602. doi: 10.1111/j.1461-0248.2006.00915.x.

Myers, N. *et al.* (2000) ‘Biodiversity hotspots for conservation priorities’, *Nature*, 403(February), pp. 853–858. Available at: www.nature.com.

N Joppa, B. L. *et al.* (2016) ‘Filling in biodiversity threat gaps’, *Science*, 352(6284), p. 199. doi: 10.1126/science.aaf4521.

Naidoo, R. *et al.* (2008) ‘Global mapping of ecosystem services and conservation priorities’, *Proceedings of the National Academy of Sciences of the United States of America*, 105(28), pp. 9495–9500. doi: 10.1073/pnas.0707823105.

New, M. *et al.* (2002) ‘A high-resolution data set of surface climate over global land areas’, *Climate Research*, 21(1), pp. 1–25.

Niang, I. *et al.* (2015) ‘Climate Change 2014: Impacts, Adaptation and Vulnerability: Part B: Regional Aspects: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change’, pp. 1199–1266. doi: 10.1017/CBO9781107415386.002.

Niklas, K. J. and Spatz, H. C. (2010) ‘Worldwide correlations of mechanical properties and green wood density’, *American Journal of Botany*, 97(10), pp. 1587–1594.

Nishina, K. *et al.* (2015) ‘Decomposing uncertainties in the future terrestrial carbon budget associated with emission scenarios, climate projections, and ecosystem simulations using the ISI-MIP results’, *Earth System Dynamics*, 6(2), pp. 435–445. doi: 10.5194/esd-6-435-2015.

Nobel, P. S., Alm, D. M. and Cavelier, J. (1992) ‘Growth Respiration, Maintenance Respiration and Structural-Carbon Costs for Roots of Three Desert Succulents’, *Functional Ecology*. JSTOR, 6(1), p. 79. doi: 10.2307/2389774.

Nobel, P. S., Miller, P. M. and Graham, E. A. (1992) ‘Influence of rocks on soil temperature, soil water potential, and rooting patterns for desert succulents’, *Oecologia*. Springer-Verlag, 92(1), pp. 90–96. doi: 10.1007/BF00317267.

Nogueira, J. *et al.* (2016) ‘Can We Go Beyond Burned Area in the Assessment of Global Remote Sensing Products with Fire Patch Metrics?’, *Remote Sensing*. Multidisciplinary Digital Publishing Institute, 9(1), p. 7. doi: 10.3390/rs9010007.

Nosova, M., Severova, E. and Volkova, O. (2017) ‘A 6500-year pollen record from the Polistovo-Lovatskaya Mire System (northwest European Russia). Vegetation dynamics and signs of human impact’, *Grana*. Taylor and Francis A.S., 56(6), pp. 410–423. doi: 10.1080/00173134.2016.1276210.

Nutting, W. L., Haverty, M. I. and Lafage, J. P. (1987) ‘Physical and chemical alteration of soil by two subterranean termite species in Sonoran Desert grassland.’, *Journal of Arid Environments*. Academic Press, 12(3), pp. 233–239. doi: 10.1016/s0140-1963(18)31167-4.

O’Connor, T. G., Puttick, J. R. and Hoffman, M. T. (2014) ‘Bush encroachment in southern Africa: changes and causes’, *African Journal of Range & Forage Science*. Taylor & Francis, 31(2), pp. 67–88. doi: 10.2989/10220119.2014.939996.

Ogutu, J. O. and Owen-Smith, N. (2003) ‘ENSO, rainfall and temperature influences on extreme population declines among African savanna ungulates’, *Ecology Letters*. John Wiley & Sons, Ltd, 6(5), pp. 412–419. doi: 10.1046/j.1461-0248.2003.00447.x.

van Oijen, M. *et al.* (2018) ‘Effects of Climate Change on Grassland Biodiversity and Productivity: The Need for a Diversity of Models’, *Agronomy*. Multidisciplinary Digital Publishing Institute, 8(2), p. 14. doi: 10.3390/agronomy8020014.

Oliver, C. D. and Larson, B. C. (1996) *Forest stand dynamics*. Wiley.

Olson, D. M. *et al.* (2001) ‘Terrestrial Ecoregions of the World: A New Map of Life on Earth’, *BioScience*, 51(11), p. 933. doi: https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Olson, M. S. and Platt, W. J. (1995) ‘Effects of habitat and growing season fires on resprouting of shrubs in longleaf pine savannas’, *Vegetatio*. Kluwer Academic Publishers, 119(2), pp. 101–118. doi: 10.1007/BF00045593.

Ordway, E. M., Asner, G. P. and Lambin, E. F. (2017) ‘Deforestation risk due to commodity crop expansion in sub-Saharan Africa’, *Environmental Research Letters*, 12(4). doi: 10.1088/1748-9326/aa6509.

Oreskes, N., Shrader-Frechette, K. and Belitz, K. (1994) ‘Verification, validation, and confirmation of numerical models in Earth sciences’, *Science*, 263(5147), pp. 641–646.

Paciorek, C. J. *et al.* (2000) ‘The demographics of resprouting in tree and shrub species of a moist tropical forest’, *Journal of Ecology*. John Wiley & Sons, Ltd, 88(5), pp. 765–777. doi: 10.1046/j.1365-2745.2000.00494.x.

Page, S. E. (2005) ‘Are we collapsing? A review of Jared Diamond’s Collapse: How societies choose to fail or succeed’, *Journal of Economic Literature*, 43(4), pp. 1049–1062. doi: 10.1257/002205105775362032.

Pallardy, S. G. (1981) ‘Closely related woody plants. Water deficits and plant growth.’, in.

Pappas, C. *et al.* (2013) ‘Sensitivity analysis of a process-based ecosystem model: Pinpointing parameterization and structural issues’, *Journal of Geophysical Research: Biogeosciences*, 118, pp. 1–24. doi: 10.1002/jgrg.20035.

Pappas, C. *et al.* (2015) ‘The role of local-scale heterogeneities in terrestrial ecosystem modeling’, *Journal of Geophysical Research: Biogeosciences*. John Wiley & Sons, Ltd, 120(2), pp. 341–360. doi: 10.1002/2014JG002735.

Parés-Ramos, I. K., Gould, W. A. and Aide, T. M. (2008) ‘Agricultural Abandonment, Suburban Growth, and Forest Expansion in Puerto Rico between 1991 and 2000’, *Ecology and Society*. The Resilience Alliance, 13(2), p. art1. doi: 10.5751/ES-02479-130201.

Parmenter, R. R. and MacMahon, J. A. (1983) ‘Factors determining the abundance and distribution of rodents in a shrub-steppe ecosystem: the role of shrubs', *Oecologia*. Springer-Verlag, 59(2–3), pp. 145–156. doi: 10.1007/BF00378831.

Parr, C. L. and Chown, S. L. (2003) 'Burning issues for conservation: A critique of faunal fire research in Southern Africa', *Austral Ecology*. Wiley/Blackwell (10.1111), 28(4), pp. 384–395. doi: 10.1046/j.1442-9993.2003.01296.x.

Pate, J. S. *et al.* (1990) 'Seedling Growth and Storage Characteristics of Seeder and Resprouter Species of Mediterranean-type Ecosystems of S. W. Australia', *Annals of Botany*, 65(6), pp. 585–601. doi: 10.1093/oxfordjournals.aob.a087976.

Pausas, J. G. *et al.* (2004) 'Plant functional traits in relation to fire in crown-fire ecosystems', *Ecology*, 85(4), pp. 1085–1100. doi: 10.1890/02-4094.

Pausas, J. G. and Keeley, J. E. (2014) 'Evolutionary ecology of resprouting and seeding in fire-prone ecosystems', *New Phytologist*, 204(1), pp. 55–65. doi: 10.1111/nph.12921.

Pavlick, R. *et al.* (2013) 'The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): a diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs', *Biogeosciences*. Copernicus GmbH on behalf of the European Geosciences Union, 10(6), pp. 4137–4177. doi: 10.5194/bg-10-4137-2013.

Peiser, B. (2005) 'From Genocide to Ecocide: The Rape of Rapa Nui', *Energy & Environment*. SAGE Publications Inc., 16(3–4), pp. 513–539. doi: 10.1260/0958305054672385.

Pellegrini, A. F. A. (2016) 'Nutrient limitation in tropical savannas across multiple scales and mechanisms', *Ecology*, 89(10), pp. 2712–2724.

Pellegrini, A. F. A. *et al.* (2017) 'Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity', *Nature*. Nature Publishing Group, 553(7687), pp. 194–198. doi: 10.1038/nature24668.

Peng, S. *et al.* (2015) 'Benchmarking the seasonal cycle of CO 2 fluxes simulated by terrestrial ecosystem models', *Global Biogeochemical Cycles*. John Wiley & Sons, Ltd, 29(1), pp. 46–64. doi: 10.1002/2014GB004931.

Pennington, R. T., Lehmann, C. E. R. and Rowland, L. M. (2018) 'Tropical savannas and dry forests', *Current Biology*. Elsevier, 28(9), pp. R541–R545. doi: 10.1016/j.cub.2018.03.014.

Perkins, S. R. and Owens, M. K. (2003) 'Growth and biomass allocation of shrub and grass seedlings in response to predicted changes in precipitation seasonality', *PLANT ECOLOGY*, 168(1), pp. 107–120.

Peters, D. P. C. *et al.* (2014) 'Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology', *Ecosphere*. John Wiley & Sons, Ltd, 5(6), p. art67. doi: 10.1890/ES13-00359.1.

Pfeiffer, M. *et al.* (2019) 'Grazing and aridity reduce perennial grass abundance in semi-arid rangelands – Insights from a trait-based dynamic vegetation model', *Ecological Modelling*. Elsevier, 395, pp. 11–22. doi: 10.1016/J.ECOLMODEL.2018.12.013.

Phillips, W. (1963) 'Depth of Roots in Soil', *Ecology*, 44(2), p. 424.

Pini, R. *et al.* (2017) 'From pristine forests to high-altitude pastures: an ecological approach to prehistoric human impact on vegetation and landscapes in the western Italian Alps', *Journal of Ecology*. Edited by J. Gill. Blackwell Publishing Ltd, 105(6), pp. 1580–1597. doi: 10.1111/1365-2745.12767.

Pitt, M. A. and Myung, I. J. (2002) 'When a good fit can be bad', *Trends in Cognitive Sciences*, 6(10), pp. 421–425.

Polly, P. D. *et al.* (2011) 'History matters: ecometrics and integrative climate change biology', *Proceedings of the Royal Society B-Biological Sciences*, 278(1709), pp. 1131–1140. doi: 10.1098/rspb.2010.2233.

Popp, A. *et al.* (2007) 'Simulating the impacts of vegetation structure on the occurrence of a small mammalian carnivore in semi-arid savanna rangelands', *Ecological Modelling*, 209(2–4), pp. 136–148.

Porembski, S. *et al.* (1994) 'Vegetation of rock outcrops in Guinea: Granite inselbergs, Sandstone table mountains and ferricretes - Remarks on species numbers and endemism', *Flora*. Elsevier GmbH, 189(4), pp. 315–326. doi: 10.1016/S0367-2530(17)30612-6.

Possingham, H. P., Comins, H. N. and Noble, I. R. (1995) 'The fire and flammability niches in plant-communities', *Journal of Theoretical Biology*, 174(1), pp. 97–108.

Prentice, I. C. *et al.* (1992) 'A global biome model based on plant physiology and dominance, soil properties and climate', *Journal of Biogeography*, 19(2), pp. 117–134.

Prentice, I. C. *et al.* (2007) 'Dynamic Global Vegetation Modeling: Quantifying Terrestrial Ecosystem Responses to Large-Scale Environmental Change', in Canadell, J. G., Pataki, D. E., and Pitelka, L. F. (eds) *Terrestrial Ecosystems in a Changing World*. Springer, pp. 175–192.

Prentice, I. C. *et al.* (2015) 'Reliable, robust and realistic: the three R's of next-generation land-surface modelling', *Atmospheric Chemistry and Physics*, 15(10), pp. 5987–6005. doi: 10.5194/acp-15-5987-2015.

Purves, D. and Pacala, S. (2008) 'Predictive models of forest dynamics', *Science*, 320(5882), pp. 1452–1453. doi: 10.1126/science.1155359.

Puyravaud, J.-P., Dufour, C. and Aravajy, S. (2003) 'Rain forest expansion mediated by successional processes in vegetation thickets in the Western Ghats of India', *Journal of Biogeography*. Blackwell Science Ltd, 30(7), pp. 1067–1080. doi: 10.1046/j.1365-2699.2003.00882.x.

Quillet, A., Peng, C. and Garneau, M. (2010) 'Toward dynamic global vegetation models for simulating vegetation-climate interactions and feedbacks: recent developments, limitations, and future challenges', *Environmental Reviews*, 18, pp. 333–353. doi: 10.1139/A10-016.

Rabin, S. S. *et al.* (2017) 'The Fire Modeling Intercomparison Project (FireMIP), phase 1: Experimental and analytical protocols with detailed model descriptions', *Geoscientific Model Development*, 10(3), pp. 1175–1197. doi: 10.5194/gmd-10-1175-2017.

Rammer, W. and Seidl, R. (2019) 'Harnessing Deep Learning in Ecology: An Example Predicting Bark Beetle Outbreaks', *Frontiers in Plant Science*. Frontiers Media SA, 10, p. 1327. doi: 10.3389/fpls.2019.01327.

Rammig, A. *et al.* (2018) 'A generic pixel-to-point comparison for simulated large-scale ecosystem properties and ground-based observations: an example from the Amazon region', *Geoscientific Model Development*, 11(12), pp. 5203–5215. doi: 10.5194/gmd-11-5203-2018.

Ratnam, J. *et al.* (2011) 'When is a `forest' a savanna, and why does it matter?', *Global Ecology and Biogeography*, 20, pp. 653–660. doi: 10.1111/j.1466-8238.2010.00634.x.

Recknagel, F. (2001) 'Applications of machine learning to ecological modelling', *Ecological Modelling*. Elsevier, 146(1–3), pp. 303–310. doi: 10.1016/S0304-3800(01)00316-7.

Reed, D. N. *et al.* (2009) 'The spatial distribution of vegetation types in the Serengeti ecosystem: the influence of rainfall and topographic relief on vegetation patch characteristics', *Journal of Biogeography*. John Wiley & Sons, Ltd, 36(4), pp. 770–782. doi: 10.1111/j.1365-2699.2008.02017.x.

Reich, P. B., Walters, M. B. and Ellsworth, D. S. (1997) 'From tropics to tundra: Global convergence in plant functioning', *Proceedings of the National Academy of Sciences of the United States of America*, 94(25), pp. 13730–13734.

Restrepo-Coupe, N. *et al.* (2017) 'Do dynamic global vegetation models capture the seasonality of carbon fluxes in the Amazon basin? A data-model intercomparison', *Global Change Biology*. John Wiley & Sons, Ltd (10.1111), 23(1), pp. 191–208. doi: 10.1111/gcb.13442.

Reyes, O., Casal, M. and Rego, F. C. (2009) 'Resprouting ability of six Atlantic shrub species', *Folia Geobotanica*. Springer, 44(1), pp. 19–29. doi: 10.1007/s12224-009-9029-x.

Richardson, R. B. (2010) 'Ecosystem services and food security: Economic perspectives on environmental sustainability', *Sustainability*. Molecular Diversity Preservation International, 2(11), pp. 3520–3548. doi: 10.3390/su2113520.

Ringrose, S. *et al.* (2002) 'Climate- and human-induced woody vegetation changes in Botswana and their implications for human adaptation', *Environmental Management*. Springer-Verlag, 30(1), pp. 98–109. doi: 10.1007/s00267-002-2486-0.

Rivest, D. *et al.* (2011) 'Shrub encroachment in Mediterranean silvopastoral systems: Retama sphaerocarpa and Cistus ladanifer induce contrasting effects on pasture and Quercus ilex production', *Agriculture, Ecosystems & Environment*, 141(3), pp. 447–454. doi: 10.1016/j.agee.2011.04.018.

Roosevelt, A. C. (2013) 'The Amazon and the Anthropocene: 13,000 years of human influence in a tropical rainforest', *Anthropocene*. Elsevier Ltd, 4, pp. 69–87. doi: 10.1016/j.ancene.2014.05.001.

Roques, K. G., O'Connor, T. G. and Watkinson, A. R. (2001) 'Dynamics of shrub encroachment in an African savanna: relative influences of fire, herbivory, rainfall and density dependence', *Journal of Applied Ecology*. Blackwell Science Ltd, 38(2), pp. 268–280. doi: 10.1046/j.1365-2664.2001.00567.x.

Rosenzweig, C. *et al.* (2017) 'Assessing inter-sectoral climate change risks: the role of ISIMIP', *Environmental Research Letters*. IOP Publishing, 12(1), p. 010301. doi: 10.1088/1748-9326/12/1/010301.

Le Roux, X., Bariac, T. and Mariotti, A. (1995) 'Spatial partitioning of the soil water resource between grass and shrub components in a West African humid savanna', *Oecologia*, 104(2), pp. 147–155. doi: 10.1007/BF00328579.

Rushworth, J. E. (1975) *The floristic, physiognomic, and biomass structure of Kalahari Sand shrub vegetation in relation to fire and frost in Wankie National Park, Rhodesia*. University of Zimbabwe.

Rutherford, M. C. (1975) *Aspects of ecosystem function in a woodland savanna in South West Africa*. University of Stellenbosch.

Rutherford, M. C. (1979) 'Above ground biomass subdivisions in woody species of the savanna ecosystem project study area Nylsvley', *South African National Scientific Programmes Report no. 36*, pp. 1–39.

Rutherford, M. C., Powrie, L. W. and Husted, L. B. (2012) 'Plant diversity consequences of a herbivore-driven biome switch from Grassland to Nama-Karoo shrub steppe in South Africa', *Applied Vegetation Science*, 15(1), pp. 14–25. doi: 10.1111/j.1654-109X.2011.01160.x.

Ryan, M. G. (1991) 'Effects of Climate Change on Plant Respiration', *Ecological Applications*. John Wiley & Sons, Ltd, 1(2), pp. 157–167. doi: 10.2307/1941808.

Saatchi, S. S. *et al.* (2011) 'Benchmark map of forest carbon stocks in tropical regions across three continents', *Proceedings of the National Academy of Sciences of the United States of America*. Natl Acad Sciences, 108(24), pp. 9899–9904. doi: 10.1073/pnas.1019576108.

Sage, R. F. and Monson, R. K. (1999) 'C4 Plant Biology', in Sage, R F and Monson, R K (eds). Academic Press, San Diego, USA.

Saglam, B. *et al.* (2008) 'Fire behavior in Mediterranean shrub species (Maquis)', *African Journal of Biotechnology*, 7(22). doi: 10.4314/AJB.V7I22.59536.

Sahin, S. (2012) 'An aridity index defined by precipitation and specific humidity', *Journal of Hydrology*. Elsevier, 444–445, pp. 199–208. doi: 10.1016/j.jhydrol.2012.04.019.

Sakschewski, B. *et al.* (2015) 'Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model', *Global Change Biology*, 21(7), pp. 2711–2725. doi: 10.1111/gcb.12870.

Salehyan, I. *et al.* (2012) 'Social Conflict in Africa: A New Database', *International Interactions*. Taylor & Francis Group, 38(4), pp. 503–511. doi: 10.1080/03050629.2012.697426.

Sallu, S. M., Twyman, C. and Thomas, D. S. G. (2009) 'The multidimensional nature of biodiversity and social dynamics and implications for contemporary rural livelihoods in remote Kalahari settlements, Botswana', *African Journal of Ecology*, 47(SUPPL. 1), pp. 110–118. doi: 10.1111/j.1365-2028.2008.01057.x.

San José, J. J. and Montes, R. A. (1997) 'Fire effect on the coexistence of trees and grasses in savannas and the resulting outcome on organic matter budget', *Interciencia*, 22(6), pp. 289--.

Sankaran, M. *et al.* (2005) 'Determinants of woody cover in African savannas.', *Nature*, 438(7069), pp. 846–849. doi: 10.1038/nature04070.

Sankey, J. B. *et al.* (2012) 'Quantifying soil surface change in degraded drylands: Shrub encroachment and effects of fire and vegetation removal in a desert grassland', *Journal of Geophysical Research: Biogeosciences*, 117(G2), p. n/a-n/a. doi: 10.1029/2012JG002002.

Santini, L. *et al.* (2017) 'Assessing the suitability of diversity metrics to detect biodiversity change', *Biological Conservation*. Elsevier Ltd, 213, pp. 341–350. doi: 10.1016/j.biocon.2016.08.024.

Sato, H. and Ise, T. (2012) 'Effect of plant dynamic processes on African vegetation responses to climate change: Analysis using the spatially explicit individual-based dynamic global vegetation model "SEIB-DGVM"', *Journal of Geophysical Research-Biogeosciences*, 117, p. G03017. doi: 10.1029/2012JG002056.

Sato, H., Itoh, A. and Kohyama, T. (2007) ‘SEIB-DGVM: A new dynamic global vegetation model using a spatially explicit individual-based approach’, *Ecological Modelling*, 200(3–4), pp. 279–307. doi: 10.1016/j.ecolmodel.2006.09.006.

Savadogo, P. *et al.* (2009) ‘Examination of multiple disturbances effects on herbaceous vegetation communities in the Sudanian savanna-woodland of {W}est {A}frica’, *Flora*, 204(6), pp. 409–422. doi: 10.1016/j.flora.2008.04.004.

Scepan, J. (1999) ‘Thematic validation of high-resolution Global Land-Cover Data Sets’, *Photogrammetric Engineering and Remote Sensing*, 65(9).

Scerri, E. M. L. *et al.* (2018) ‘Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter?’, *Trends in ecology & evolution*. NLM (Medline), 33(8), pp. 582–594. doi: 10.1016/j.tree.2018.05.005.

Schaefer, C. E. G. R. *et al.* (2016) ‘Termite role in soil nutrient cycling in ironstone rupestrian grasslands (Canga) in Carajás, Brazilian Amazonia’, in *Ecology and Conservation of Mountaintop Grasslands in Brazil*. Springer International Publishing, pp. 379–391. doi: 10.1007/978-3-319-29808-5_16.

Schafer, J. L. and Mack, M. C. (2014) ‘Growth, Biomass, and Allometry of Resprouting Shrubs after Fire in Scrubby Flatwoods’, *The American Midland Naturalist*, 172(2), pp. 266–284. doi: 10.1674/0003-0031-172.2.266.

Scheffer, M. *et al.* (2009) ‘Early-warning signals for critical transitions’, *Nature*, 461(7260), pp. 53–59.

Scheffer, M. *et al.* (2015) ‘Creating a safe operating space for iconic ecosystems: Manage local stressors to promote resilience to global change’, *Science*, 347(6228), pp. 1317–1319. doi: 10.1126/science.aaa3769.

Scheffer, Marten *et al.* (2015) ‘The mystery of missing trubs revisited: A response to McGlone et al. and Qian and Ricklefs’, *Trends in Ecology and Evolution*, pp. 7–8. doi: 10.1016/j.tree.2014.11.001.

Scheiter, S. *et al.* (2012) ‘Fire and fire-adapted vegetation promoted C4 expansion in the late Miocene’, *New Phytologist*, 195(3), pp. 653–666. doi: 10.1111/j.1469-8137.2012.04202.x.

Scheiter, S. *et al.* (2015) ‘Climate change and long-term fire management impacts on Australian savannas’, *New Phytologist*, 205(3), pp. 1211–1226. doi: 10.1111/nph.13130.

Scheiter, S. *et al.* (2018) ‘How vulnerable are ecosystems in the Limpopo province to climate change?’, *South African Journal of Botany*, 116, pp. 86–95. doi: 10.1016/j.sajb.2018.02.394.

Scheiter, S., Moncrieff, G. R., *et al.* (2019) ‘African biomes are most sensitive to changes in CO 2 under recent and near-future CO 2 conditions’, (October), pp. 1–37.

Scheiter, S., Schulte, J., *et al.* (2019) ‘How Does Climate Change Influence the Economic Value of Ecosystem Services in Savanna Rangelands?’, *Ecological Economics*. Elsevier, 157, pp. 342–356. doi: 10.1016/J.ECOLECON.2018.11.015.

Scheiter, S. and Higgins, S. I. (2007) ‘Partitioning of root and shoot competition and the stability of savannas’, *American Naturalist*, 170(4), pp. 587–601.

Scheiter, S. and Higgins, S. I. (2008) ‘aDGVM : An adaptive dynamic global vegetation model for tropical ecosystems . Version 1 . 0’, pp. 1–42.

Scheiter, S. and Higgins, S. I. (2009) ‘Impacts of climate change on the vegetation of Africa: an adaptive dynamic vegetation modelling approach (aDGVM)’, *Global Change Biology*, 15, pp. 2224–2246.

Scheiter, S. and Higgins, S. I. (2010) ‘How many elephants can the KNP support? A modeling study’.

Scheiter, S. and Higgins, S. I. (2012) ‘How many elephants can you fit into a conservation area’, *Conservation Letters*, 5(3), pp. 176–185. doi: 10.1111/j.1755-263X.2012.00225.x.

Scheiter, S., Langan, L. and Higgins, S. I. (2013) ‘Next generation dynamic global vegetation models: learning from community ecology’, *New Phytologist*, 198, pp. 957–969. doi: 10.1111/nph.12210.

Scheiter, S. and Savadogo, P. (2016) ‘Ecosystem management can mitigate vegetation shifts induced by climate change in West Africa’, *Ecological Modelling*, 332, pp. 19–27. doi: 10.1016/j.ecolmodel.2016.03.022.

Schenk, H. J. *et al.* (2008) ‘Hydraulic integration and shrub growth form linked across continental aridity gradients’, *Proceedings of the National Academy of Sciences*, 105(32), pp. 11248–11253. doi:

10.1073/pnas.0804294105.

Schenk, H. J. and Jackson, R. B. (2002a) ‘Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems’, *Journal of Ecology*, 90(3), pp. 480–494.

Schenk, H. J. and Jackson, R. B. (2002b) ‘The global biogeography of roots’, *Ecological Monographs*, 72(3), pp. 311–328.

Schenk, H. J. and Jackson, R. B. (2005) ‘Mapping the global distribution of deep roots in relation to climate and soil characteristics’, *Geoderma*, 126(1–2), pp. 129–140.

Schleicher, J., Wiegand, K. and Ward, D. (2011) ‘Changes of woody plant interaction and spatial distribution between rocky and sandy soil areas in a semi-arid savanna, South Africa’, *Journal of Arid Environments*. Academic Press, 75(3), pp. 270–278. doi: 10.1016/j.jaridenv.2010.10.003.

Schmiedel, U. (2002) ‘The quartz fields of southern Africa: flora, phytogeography, vegetation, and habitat ecology’, pp. 1–363.

Schmiedel, U. and Jürgens, N. (1999) ‘Community structure on unusual habitat islands: Quartz-fields in the Succulent Karoo, South Africa’, *Plant Ecology*. Springer, 142(1–2), pp. 57–69. doi: 10.1023/A:1009818210799.

Scholes, R. J. and Archer, S. R. (1997) ‘Tree-grass interactions in Savannas’, *Annual Review of Ecology, Evolution, and Systematics*, 28, pp. 517–544.

Scholes, R. J., Bond, W. J. and Eckhardt, H. C. (2003) ‘Vegetation dynamics in the Kruger system’, in du Toit, J. T., Rogers, K. H., and Biggs, H. C. (eds) *The Kruger Experience: Ecology And Management Of Savanna Heterogeneity*. Island Press, Washington.

Scholze, M. *et al.* (2006) ‘A climate-change risk analysis for world ecosystems’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 103(35), pp. 13116–13120. doi: 10.1073/pnas.0601816103.

Schwartz, G. M. and Nichols, J. J. (2010) *After collapse. The regeneration of complex societies*, *Journal of Chemical Information and Modeling*. Edited by G. M. Schwartz and J. J. Nichols. Tucson: The Uniuverisyt of Arizona Press. doi: 10.1017/CBO9781107415324.004.

Schwinning, S. and Ehleringer, J. R. (2001) ‘Water use trade-offs and optimal adaptations to pulse-driven arid ecosystems’, *Journal of Ecology*. John Wiley & Sons, Ltd, 89(3), pp. 464–480. doi: 10.1046/j.1365-2745.2001.00576.x.

Shea, R. W. *et al.* (1996) ‘Fuel biomass and combustion factors associated with fires in savanna ecosystems of South Africa and Zambia’, *Journal of Geophysical Research Atmospheres*. Blackwell Publishing Ltd, 101(19), pp. 23551–23568. doi: 10.1029/95jd02047.

Shibata, R. *et al.* (2016) ‘Relationships between resprouting ability, species traits and resource allocation patterns in woody species in a temperate forest’, *Functional Ecology*, 30(7), pp. 1205–1215. doi: 10.1111/1365-2435.12616.

Shrukin, J. (1996) *Engines of the mind: the evolution of the computer from mainframes to microprocessor*. New York: Norton.

Shugart, H. H. *et al.* (2018) ‘Gap models and their individual-based relatives in the assessment of the consequences of global change’, *Environmental Research Letters*. IOP Publishing, 13(3), p. 033001. doi: 10.1088/1748-9326/aaaacc.

Siebert, S. J. *et al.* (2003) ‘Vegetation of the rock habitats of the Sekhukhuneland Centre of Plant Endemism, South Africa’, *Bothalia*, 33(2), pp. 207–228. doi: 10.4102/abc.v33i2.454.

Silva, I. A. and Batalha, M. A. (2011) ‘Plant functional types in Brazilian savannas: The niche partitioning between herbaceous and woody species’, *Perspectives in Plant Ecology Evolution and Systematics*, 13(3), pp. 201–206. doi: 10.1016/j.ppees.2011.05.006.

Silva, L. C. R. and Anand, M. (2011) ‘Mechanisms of Araucaria (Atlantic) Forest Expansion into Southern Brazilian Grasslands’, *Ecosystems*. Springer-Verlag, 14(8), pp. 1354–1371. doi: 10.1007/s10021-011-9486-y.

Sirami, C. *et al.* (2009) ‘The impact of shrub encroachment on savanna bird diversity from local to regional scale’, *Diversity and Distributions*. John Wiley & Sons, Ltd (10.1111), 15(6), pp. 948–957. doi:

10.1111/j.1472-4642.2009.00612.x.

Sitch, S. *et al.* (2008) 'Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs)', *Global Change Biology*, 14(9), pp. 2015–2039.

Skarpe, C. (1992) 'Dynamics of savanna ecosystems', *Journal of Vegetation Science*. Wiley, 3(3), pp. 293–300. doi: 10.2307/3235754.

Skowno, A. L. *et al.* (2017) 'Woodland expansion in South African grassy biomes based on satellite observations (1990-2013): general patterns and potential drivers', *Global Change Biology*, 23(6), pp. 2358–2369. doi: 10.1111/gcb.13529.

Smit, I. P. J. *et al.* (2010) 'Effects of fire on woody vegetation structure in African savanna', *Ecological Applications*. Ecological Society of America, 20(7), pp. 1865–1875. doi: 10.1890/09-0929.1.

Smit, I. P. J. *et al.* (2016) 'An examination of the potential efficacy of high intensity fires for reversing woody encroachment in savannas', *Journal of Applied Ecology*. Edited by P. Kardol, 53(5), pp. 1623–1633. doi: 10.1111/1365-2664.12738.

Smit, I. P. J. and Prins, H. H. T. (2015) 'Predicting the Effects of Woody Encroachment on Mammal Communities, Grazing Biomass and Fire Frequency in African Savannas', *PLOS ONE*. Edited by M. S. Crowther. Public Library of Science, 10(9), p. e0137857. doi: 10.1371/journal.pone.0137857.

Smith, B., Prentice, I. C. and Sykes, M. T. (2001) 'Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space', *Global Ecology and Biogeography*, 10(6), pp. 621–637.

Smith, M. D. *et al.* (2016) 'Shared Drivers but Divergent Ecological Responses: Insights from Long-Term Experiments in Mesic Savanna Grasslands', *BioScience*. Narnia, 66(8), pp. 666–682. doi: 10.1093/biosci/biw077.

Smith, T. M. *et al.* (2007) 'Earliest evidence of modern human life history in North African early Homo sapiens', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 104(15), pp. 6128–6133. doi: 10.1073/pnas.0700747104.

Song, X. P. *et al.* (2018) 'Global land change from 1982 to 2016', *Nature*. Springer US, 560(7720), pp. 639–643. doi: 10.1038/s41586-018-0411-9.

Song, X., Zeng, X.-D. and Li, F. (2016) 'Evaluation of the individual allocation scheme and its impacts in a dynamic global vegetation model', *Atmospheric and Oceanic Science Letters*, 9(1), pp. 38–44. doi: 10.1080/16742834.2015.1128231.

Song, X., Zeng, X. and Zhu, J. (2013) 'Evaluating the tree population density and its impacts in CLM-DGVM', *Advances in Atmospheric Sciences*. Springer, 30(1), pp. 116–124. doi: 10.1007/s00376-012-1271-0.

Soto-Shoender, J. R. *et al.* (2018) 'The importance of grass cover for mammalian diversity and habitat associations in a bush encroached savanna', *Biological Conservation*. Elsevier Ltd, 221, pp. 127–136. doi: 10.1016/j.biocon.2018.02.028.

Soussi, A. *et al.* (2016) 'Plant-associated microbiomes in arid lands: diversity, ecology and biotechnological potential', *Plant and Soil*. Springer International Publishing, 405(1–2), pp. 357–370. doi: 10.1007/s11104-015-2650-y.

Spagnolo, B. and La Barbera, A. (2002) 'Role of the noise on the transient dynamics of an ecosystem of interacting species', *Physica A: Statistical Mechanics and its Applications*. North-Holland, 315(1–2), pp. 114–124. doi: 10.1016/S0378-4371(02)01245-1.

Spessa, A. *et al.* (2013) 'Evaluating the coupled vegetation-fire model, LPJ-GUESS-SPITFIRE, against observed tropical forest biomass', *EGU General Assembly 2013, held 7-12 April, 2013 in Vienna, Austria, id. EGU2013-6070*, 15. Available at: http://adsabs.harvard.edu/abs/2013EGUGA..15.6070S (Accessed: 20 February 2019).

Spinoni, J. *et al.* (2015) 'Towards identifying areas at climatological risk of desertification using the Köppen-Geiger classification and FAO aridity index', *International Journal of Climatology*. John Wiley and Sons Ltd, 35(9), pp. 2210–2222. doi: 10.1002/joc.4124.

Staal, A. *et al.* (2016) 'Bistability, Spatial Interaction, and the Distribution of Tropical Forests and

Savannas', *Ecosystems*. Springer New York LLC, 19(6), pp. 1080–1091. doi: 10.1007/s10021-016-0011-1.

Staal, A. *et al.* (2018) 'Resilience of tropical tree cover: The roles of climate, fire, and herbivory', *Global Change Biology*. Blackwell Publishing Ltd, 24(11), pp. 5096–5109. doi: 10.1111/gcb.14408.

Stanton, R. A. *et al.* (2018) 'Shrub encroachment and vertebrate diversity: A global meta-analysis', *Global Ecology and Biogeography*, 27(3), pp. 368–379. doi: 10.1111/geb.12675.

Staver, a. C., Bond, W. J. and February, E. C. (2011) 'History matters: tree establishment variability and species turnover in an African savanna', *Ecosphere*, 2(4), p. art49. doi: 10.1890/ES11-00029.1.

Staver, A C, Archibald, S. and Levin, S. (2011) 'Tree cover in sub-Saharan Africa: Rainfall and fire constrain forest and savanna as alternative stable states', *Ecology*, 92(5), pp. 1063–1072.

Staver, A Carla, Archibald, S. and Levin, S. A. (2011) 'The Global Extent and Determinants of Savanna and Forest as Alternative Biome States', *Science*, 334(6053), pp. 230–232. doi: 10.1126/science.1210465.

Staver, A. C. and Bond, W. J. (2014) 'Is there a "browse trap"? Dynamics of herbivore impacts on trees and grasses in an African savanna', *Journal of Ecology*. doi: 10.1111/1365-2745.12230.

Staver, A. C., Botha, J. and Hedin, L. (2017) 'Soils and fire jointly determine vegetation structure in an African savanna', *New Phytologist*. Blackwell Publishing Ltd, 216(4), pp. 1151–1160. doi: 10.1111/nph.14738.

Staver, A. C. and Hansen, M. C. (2015) 'Analysis of stable states in global savannas: is the CART pulling the horse? - a comment', *Global Ecology and Biogeography*. Blackwell Publishing Ltd, 24(8), pp. 985–987. doi: 10.1111/geb.12285.

Staver, C. A. *et al.* (2009) 'Browsing and fire interact to suppress tree density in an African savanna', *Ecological Applications*. doi: 10.1890/08-1907.1.

Stearns, S. C. (1989) 'Trade-Offs in Life-History Evolution', *Functional Ecology*, 3(3), p. 259. doi: 10.2307/2389364.

Stephenson, N. L. *et al.* (2014) 'Rate of tree carbon accumulation increases continuously with tree size', *Nature*. Nature Publishing Group, 507(7490), pp. 90–93. doi: 10.1038/nature12914.

Sternberg, L. D. L. (2001) 'Savanna-forest hysteresis in the tropics', *Global Ecology and Biogeography*, 10(4), pp. 369–378.

Stevens, N., Lehmann, C. E. R., *et al.* (2016) 'Savanna woody encroachment is widespread across three continents', *Global Change Biology*, 23(1), pp. 1–10. doi: 10.1111/gcb.13409.

Stevens, N., Erasmus, B. F. N., *et al.* (2016) 'Woody encroachment over 70 years in South African savannahs: overgrazing, global change or extinction aftershock?', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1703), p. 20150437. doi: 10.1098/rstb.2015.0437.

Stewart, J. R. and Stringer, C. B. (2012) 'Human Evolution Out of Africa: The Role of Refugia and Climate Change', *Science*, 335(6074), pp. 1317–1321. doi: 10.1126/science.1215627.

Stuart-Hill, G. C. and Mentis, M. T. (1982) 'Coevolution of African grasses and large herbivores', *Proceedings of the Annual Congresses of the Grassland Society of Southern Africa*. Taylor & Francis Group, 17(1), pp. 122–128. doi: 10.1080/00725560.1982.9648969.

Svensson, E. I. and Berger, D. (2019) 'The Role of Mutation Bias in Adaptive Evolution', *Trends in Ecology and Evolution*. Elsevier Ltd, 34(5), pp. 422–434. doi: 10.1016/j.tree.2019.01.015.

Svirezhev, Y. M. (2000) 'Thermodynamics and ecology', *Ecological Modelling*. Elsevier, 132(1–2), pp. 11–22. doi: 10.1016/S0304-3800(00)00301-X.

Tachajapong, W. *et al.* (2009) 'Experimental and numerical modeling of shrub crown fire initiation', *Combustion Science and Technology*. Taylor & Francis Group, 181(4), pp. 618–640. doi: 10.1080/00102200802693617.

Thessen, A. E. (2016) 'Adoption of machine learning techniques in ecology and earth science', *One Ecosystem*, 1, pp. 1–38. doi: 10.3897/oneeco.1.e8621.

Thom, D. *et al.* (2018) 'Legacies of past land use have a stronger effect on forest carbon exchange than future climate change in a temperate forest landscape', *Biogeosciences*, 15(18), pp. 5699–5713. doi: 10.5194/bg-15-5699-2018.

Thompson, J. N. (1998) ‘Rapid evolution as an ecological process’, *Trends in Ecology & Evolution*. Elsevier Current Trends, 13(8), pp. 329–332. doi: 10.1016/S0169-5347(98)01378-0.

Tilman, D. (1995) ‘Biodiversity: Population Versus Ecosystem Stability’, *Ecology*. Ecological Society of America, 77(2), pp. 350–363. doi: 10.2307/2265614.

Tilman, D. *et al.* (1997) ‘The influence of functional diversity and composition on ecosystem processes’, *Science*, 277(5330), pp. 1300–1302. doi: 10.1126/science.277.5330.1300.

Tocco, C. *et al.* (2013) ‘Pastoral Practices to Reverse Shrub Encroachment of Sub-Alpine Grasslands: Dung Beetles (Coleoptera, Scarabaeoidea) Respond More Quickly Than Vegetation’, *PLoS ONE*. Edited by B. Boldgiv. R Foundation for Statistical Computing, 8(12), p. e83344. doi: 10.1371/journal.pone.0083344.

Todd-Brown, K. E. O. *et al.* (2013) ‘Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations’, *Biogeosciences*, 10(3), pp. 1717–1736. doi: 10.5194/bg-10-1717-2013.

Tomscha, S. A. and Gergel, S. E. (2016) ‘Ecosystem service trade-offs and synergies misunderstood without landscape history’, *Ecology and Society*, 21(1). doi: 10.5751/ES-08345-210143.

Trollope, W. S. W. (1980) ‘Controlling bush encroachment with fire in the savanna areas of South Africa’, *Proceedings of the Annual Congresses of the Grassland Society of Southern Africa*. Taylor & Francis Group, 15(1), pp. 173–177. doi: 10.1080/00725560.1980.9648907.

Trollope, W. S. W. and Potgieter, A. L. F. (1986) ‘Estimating grass fuel loads with a disc pasture meter in the kruger national park’, *Journal of the Grassland Society of Southern Africa*. Taylor & Francis Group, 3(4), pp. 148–152. doi: 10.1080/02566702.1986.9648053.

Trugman, A. T. *et al.* (2019) ‘Leveraging plant hydraulics to yield predictive and dynamic plant leaf allocation in vegetation models with climate change’, *Global Change Biology*, 25(12), pp. 4008–4021. doi: 10.1111/gcb.14814.

Tuanmu, M. N. and Jetz, W. (2014) ‘A global 1-km consensus land-cover product for biodiversity and ecosystem modelling’, *Global Ecology and Biogeography*, 23(9), pp. 1031–1045. doi: 10.1111/geb.12182.

Tuljapurkar, S. D. and Semura, J. S. (1977) ‘Dynamic equilibrium under periodic perturbations in simple ecosystem models’, *Journal of Theoretical Biology*. Academic Press, 66(2), pp. 327–343. doi: 10.1016/0022-5193(77)90175-8.

Turner, N. C. (1978) ‘Arid and Semiarid Zones’, *Annual Review of Plant Physiology*, 29, pp. 277–317.

Twine, W. *et al.* (2003) ‘Consumption and direct-use values of savanna bio-resources used by rural households in Mametja, a semi-arid area of Limpopo province, South Africa : research letter’, *South African Journal of Science*. Sabinet, 99(9), pp. 467–473.

Tyree, M. T. and Ewers, F. W. (1991) ‘The hydraulic architecture of trees and other woody plants’, *New Phytologist*. Wiley/Blackwell (10.1111), 119(3), pp. 345–360. doi: 10.1111/j.1469-8137.1991.tb00035.x.

United Nations (2019) *World population prospects 2019*. New York. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12283219.

Valentin, C. (1994) ‘Surface sealing as affected by various rock fragment covers in West Africa’, *Catena*. Elsevier, 23(1–2), pp. 87–97. doi: 10.1016/0341-8162(94)90055-8.

Venter, Z., Cramer, M. D. and Hawkins, H. (2018) ‘Drivers of woody plant encroachment over Africa’, *Nature Communications*. Springer US, (2018), pp. 1–7. doi: 10.1038/s41467-018-04616-8.

Veraart, A. J. *et al.* (2012) ‘Recovery rates reflect distance to a tipping point in a living system’, *Nature*, 481(7381), pp. 357--U137. doi: 10.1038/nature10723.

Verdaguer, D. and Ojeda, F. (2002) ‘Root starch storage and allocation patterns in seeder and resprouter seedlings of two Cape Erica (Ericaceae) species’, *American Journal of Botany*. Botanical Society of America Inc., 89(8), pp. 1189–1196. doi: 10.3732/ajb.89.8.1189.

Verheijen, L. M. *et al.* (2013) ‘Impacts of trait variation through observed trait-climate relationships on performance of an Earth system model: a conceptual analysis’, *Biogeosciences*, 10(8), pp. 5497–5515. doi: 10.5194/bg-10-5497-2013.

Verhoef, A. and Egea, G. (2014) ‘Modeling plant transpiration under limited soil water: Comparison of different plant and soil hydraulic parameterizations and preliminary implications for their use in land surface models', *Agricultural and Forest Meteorology*, 191, pp. 22–32. doi: 10.1016/j.agrformet.2014.02.009.

Vetaas, O. R. (1992) 'Micro-site effects of trees and shrubs in dry savannas', *Journal of Vegetation Science*. Wiley/Blackwell (10.1111), 3(3), pp. 337–344. doi: 10.2307/3235758.

Vieira, I. C. G., Uhl, C. and Nepstad, D. (1994) 'The role of the shrub Cordia multispicata Cham. as a "succession facilitator" in an abandoned pasture, Paragominas, Amazônia', *Vegetatio*. Kluwer Academic Publishers, 115(2), pp. 91–99. doi: 10.1007/BF00044863.

Volterra, V. (1926) 'Fluctuations in the abundance of a species considered mathematically', *Nature*, 118(2972), pp. 558–560. doi: 10.1038/118558a0.

Wäldchen, J. and Mäder, P. (2018) 'Machine learning for image based species identification', *Methods in Ecology and Evolution*. Edited by N. Cooper. British Ecological Society, 9(11), pp. 2216–2225. doi: 10.1111/2041-210X.13075.

Wallén, C. C. (1967) 'Aridity Definitions and their Applicability', *Geografiska Annaler: Series A, Physical Geography*. Informa UK Limited, 49(2–4), pp. 367–384. doi: 10.1080/04353676.1967.11879765.

Walter, H. (1971) *Ecology of tropical and subtropical vegetation*. English ed. Edited by T. by M.-D. D. Brnett J. H. Edinburgh: Oliver & Boyd. doi: 10.2307/4117095.

Wang, Y. P. (2003) 'A comparison of three different canopy radiation models commonly used in plant modelling', *Functional Plant Biology*, 30(2), pp. 143–152.

Ward, D. (2005) 'Do we understand the causes of bush encroachment in African savannas?', *African Journal of Range & Forage Science*. Taylor & Francis Group, 22(2), pp. 101–105. doi: 10.2989/10220110509485867.

Ward, D. E. *et al.* (1996) 'Effect of fuel composition on combustion efficiency and emission factors for African savanna ecosystems', *Journal of Geophysical Research: Atmospheres*. John Wiley & Sons, Ltd, 101(D19), pp. 23569–23576. doi: 10.1029/95JD02595@10.1002/(ISSN)2169-8996.TRACEA1.

Warren, D. L. and Seifert, S. N. (2011) 'Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria', *Ecological Applications*. John Wiley & Sons, Ltd, 21(2), pp. 335–342. doi: 10.1890/10-1171.1.

Warszawski, L. *et al.* (2014) 'The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 111(9), pp. 3228–3232. doi: 10.1073/PNAS.1312330110.

White, C. R. and Marshall, D. J. (2019) 'Should We Care If Models Are Phenomenological or Mechanistic?', *Trends in Ecology and Evolution*. Elsevier Ltd, 34(4), pp. 276–278. doi: 10.1016/j.tree.2019.01.006.

Whittingham, M. J. *et al.* (2006) 'Why do we still use stepwise modelling in ecology and behaviour?', *Journal of Animal Ecology*. John Wiley & Sons, Ltd, 75(5), pp. 1182–1189. doi: 10.1111/j.1365-2656.2006.01141.x.

Wiegand, K., Ward, D. and Saltz, D. (2005) 'Multi-scale patterns and bush encroachment in an arid savanna with a shallow soil layer', *Journal of Vegetation Science*. Opulus Press AB, 16(3), pp. 311–320. doi: 10.1111/j.1654-1103.2005.tb02369.x.

Wiens, J. J. and Donoghue, M. J. (2004) 'Historical biogeography, ecology and species richness', *Trends in Ecology and Evolution*. Elsevier Current Trends, 19(12), pp. 639–644. doi: 10.1016/j.tree.2004.09.011.

Wigley, B. J., Bond, W. J. and Hoffman, M. T. (2009) 'Bush encroachment under three contrasting land-use practices in a mesic South African savanna', *African Journal of Ecology*, 47, pp. 62–70.

Wigley, B. J., Bond, W. J. and Hoffman, M. T. (2010) 'Thicket expansion in a South African savanna under divergent land use: Local vs. global drivers?', *Global Change Biology*, 16(3), pp. 964–976. doi: 10.1111/j.1365-2486.2009.02030.x.

Wilcox, B. P. *et al.* (2018) 'Emerging frameworks for understanding and mitigating woody plant encroachment in grassy biomes', *Current Opinion in Environmental Sustainability*. Elsevier B.V., pp. 46–52. doi: 10.1016/j.cosust.2018.04.005.

Williams, R. J., Anandanadesan, A. and Purves, D. (2010) 'The probabilistic niche model reveals the niche structure and role of body size in a complex food web', *PLoS ONE*, 5(8). doi: 10.1371/journal.pone.0012092.

Williams, V. L., Witkowski, E. T. F. and Balkwill, K. (2007) 'Relationship between bark thickness and diameter at breast height for six tree species used medicinally in South Africa', *South African Journal of Botany*. Elsevier, 73(3), pp. 449–465.

Wilson, B. F. (1995) 'Shrub Stems: Form and Functions', in *Plant Stems: Physiology and Functional Morphology*, p. 440.

Wilson, J. B. (1988) 'A Review of Evidence on the Control of Shoot: Root Ratio, in Relation to Models', *Annals of Botany*. Oxford University Press (OUP), 61(4), pp. 433–449. doi: 10.1093/oxfordjournals.aob.a087575.

Winterhaider, B. (1984) 'Reconsidering the ecosystem concept', *Reviews in Anthropology*. Taylor & Francis Group, 11(4), pp. 301–313. doi: 10.1080/00988157.1984.9977700.

Wood, S. N. (2010) 'Statistical inference for noisy nonlinear ecological dynamic systems', *Nature*. Nature Publishing Group, 466(7310), pp. 1102–1104. doi: 10.1038/nature09319.

World Bank (2019) *World Bank Annual Report 2019 - Investing in Opportunity Ending Poverty*. Available at: http://www.worldbank.org/en/about/annual-report.

Worster, D. (1989) 'The Ecology of Order and Chaos', in. Oxford University Press, pp. 1–18.

Wright, B. R. and Clarke, P. J. (2007) 'Resprouting responses of Acacia shrubs in the Western Desert of Australia - fire severity, interval and season influence survival', *International Journal of Wildland Fire*. CSIRO PUBLISHING, 16(3), p. 317. doi: 10.1071/WF06094.

Wright, I. J. *et al.* (2004) 'The worldwide leaf economics spectrum', *Nature*, 428(6985), pp. 821–827.

Yang, Y. *et al.* (2015) 'From plant functional types to plant functional traits', *Progress in Physical Geography*. SAGE PublicationsSage UK: London, England, 39(4), pp. 514–535. doi: 10.1177/0309133315582018.

Yew Gan, T., Dlamini, E. M. and Biftu, G. F. (1997) 'Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling', *Journal of Hydrology*. Elsevier, 192(1–4), pp. 81–103. doi: 10.1016/S0022-1694(96)03114-9.

Yi, S. *et al.* (2003) 'Holocene environmental history inferred from pollen assemblages in the Huanghe (Yellow River) delta, China: Climatic change and human impact', *Quaternary Science Reviews*. Pergamon, 22(5–7), pp. 609–628. doi: 10.1016/S0277-3791(02)00086-0.

Zaehle, S. *et al.* (2005) 'Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics', *Global Biogeochemical Cycles*, 19(3), pp. 1–16. doi: 10.1029/2004GB002395.

Zanon, M. *et al.* (2018) 'European forest cover during the past 12,000 years: A palynological reconstruction based on modern analogs and remote sensing', *Frontiers in Plant Science*. Frontiers Media S.A., 9, p. 253. doi: 10.3389/fpls.2018.00253.

Zeng, Xiaodong, Zeng, Xubin and Barlage, M. (2008) 'Growing temperate shrubs over arid and semiarid regions in the Community Land Model-Dynamic Global Vegetation Model', *Global Biogeochemical Cycles*, 22(3), pp. 1–14. doi: 10.1029/2007GB003014.

Zeppel, M. J. B. *et al.* (2015) 'Drought and resprouting plants', *New Phytologist*, 206(2), pp. 583–589. doi: 10.1111/nph.13205.

Zhang, H. *et al.* (2018) 'New insights on plant phenological response to temperature revealed from long-term widespread observations in China', *Global Change Biology*. Wiley/Blackwell (10.1111), 24(5), pp. 2066–2078. doi: 10.1111/gcb.14002.

Zhou, X., Mahalingam, S. and Weise, D. (2005) 'Modeling of marginal burning state of fire spread in live chaparral shrub fuel bed', *Combustion and Flame*. Elsevier, 143(3), pp. 183–198. doi: 10.1016/j.combustflame.2005.05.013.

Zimmermann, J. *et al.* (2010) 'Grass mortality in semi-arid savanna: The role of fire, competition and self-shading', *Perspectives in Plant Ecology, Evolution and Systematics*, 12(1), pp. 1–8. doi: 10.1016/j.ppees.2009.09.003.

Zizka, A., Govender, N. and Higgins, S. I. (2014) 'How to tell a shrub from a tree: A life-history perspective from a South African savanna', *Austral Ecology*, 39(7), pp. 767–778. doi: 10.1111/aec.12142.

Zou, C. B. *et al.* (2005) 'Soil moisture redistribution as a mechanism of facilitation in savanna tree-shrub clusters', *Oecologia*. Springer, 145(1), pp. 32–40. doi: 10.1007/s00442-005-0110-8.

Zuidema, P. A., Poulter, B. and Frank, D. C. (2018) 'A Wood Biology Agenda to Support Global Vegetation Modelling', *Trends in Plant Science*. Elsevier Current Trends, 23(11), pp. 1006–1015. doi: 10.1016/J.TPLANTS.2018.08.003.

ACKNOWLEDGMENTS

I thank everyone who made me who I am and supported me to achieve my doctoral thesis. I truly appreciate the experience I enjoyed along these persons and thank everyone in due respect. However, I do not name anyone personally to respect privacy and to avoid missing names. I directly thank each and every person with whom I had stimulating scientific discussions and from which I benefitted in terms of technical support. They recognise themselves and know how much I owe them. I particularly want to acknowledge the Goethe University Frankfurt am Main and the Senckenberg Biodiversity & Climate Research Center (SBiK-F) in Frankfurt am Main, where I conducted my research, and their members, first and foremost my supervisors and my teammates, as well as organisations which funded my research. I also recognise the steering drive of the Limpopo Living Landscape (LLL) research project, within which framework my research started. It involved 7 research institutions across Germany and South Africa (University of Limpopo, University of the Witwatersrand, University of Venda, University of Bonn, University of Cologne, university of Göttingen and Goethe University Frankfurt in cooperation with SBiK-F) and I appreciated the scientific collaboration with all its members.

I want to support the sci-hub project by explicitly highlighting its contribution to my research: without it, I would not have been able to access wealth of science out of reach even for organised institutions. Being illegal in some countries does not make it wrong.

I am grateful for my family and my friends for everything and for the life I had and have thanks to them. I am indebted to the many artists of all kinds which created inspiring art, to the workers who make our society function, and to all those who fought for the good of humankind.

Disclaimer:

- The form “we” has been chosen, instead of “I”, to acknowledge the inputs from fellow researchers and past scientists..