Advances in Water Resources 33 (2010) 215–230

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

An ecohydrological approach to predicting regional woody species distribution patterns in dryland ecosystems

Trenton E. Franz [a,*], Kelly K. Caylor [a], Jan M. Nordbotten [a,b], Ignacio Rodríguez-Iturbe [a], Michael A. Celia [a]

[a] Department of Civil and Environmental Engineering, Engineering Quadrangle, Princeton University, Princeton, NJ 08544, USA
[b] Department of Mathematics, University of Bergen, N-5008 Bergen, Norway

ARTICLE INFO

Article history:
Received 4 May 2009
Received in revised form 11 December 2009
Accepted 12 December 2009
Available online 21 December 2009

Keywords:
Ecohydrology
Kenya
Dryland
Optimization
Climate change
Acacia, rainfall patterns

ABSTRACT

This paper presents a quantitative ecohydrological framework for predicting regional distribution patterns of woody species in dryland ecosystems. The framework is based on an existing stochastic model for the daily mass balance of water that represents the interactions between soils, climate, and vegetation. Individual species selection is based on an optimality trade-off hypothesis, which states that dryland vegetation patterns are constrained by maximization of water use and simultaneous minimization of water stress. The relative importance of water use and stress avoidance to the overall fitness of three *Acacia* species is determined from the heterogeneous basin, the Upper Ewaso Ng'iro river basin, of the central Kenya highlands. The model results indicate that overall fitness is more strongly influenced by water use than stress avoidance but that consideration of both stress avoidance and water use is critical to predicting basin-scale patterns of species distribution. We identify a linear trend in the frequency and intensity of storms with the same annual total using a basin-wide gauge precipitation dataset. After calibration, we apply the basin average linear trends in time for average rain per storm and storm arrival rates. The model results indicate the upslope migration of two species, *Acacia tortilis* and *Acacia xanthophloea* to areas with higher total rainfall. Lastly, we explore the modeled changes of species cover in the basin influenced by changes in rainfall total holding growing season rainfall variability constant and changes in growing season rainfall variability holding total rainfall constant. We find that changes in dryland species distribution patterns and relative abundance may be as sensitive to growing season rainfall variability as they are to changes in total rainfall amounts.

1. Introduction

Drylands are extensive, covering 30% of the Earth's land surface and 50% of Africa [62,66]. Furthermore, dryland ecosystems support a large fraction of the human population and most pastoralist societies [62]. Two challenges facing the sustainability of pastoralist societies are land use and climate change [32]. Although changes in land use are an important and a major concern, we focus our efforts on predicting natural ecosystem responses to climate change. In this study, we present a modeling framework for addressing the impacts of rainfall on the distribution of woody vegetation species. We apply our framework to an African dryland watershed in central Kenya.

Prior studies demonstrate that dryland ecosystems are sensitive to shifts in rainfall climatology [12,25,59,66,69,70]. Regional rainfall patterns are primarily governed by larger scale phenomena, such as the Intertropical Convergence Zone (ITCZ), Hadley cells [21], and global teleconnections [46]. Recent changes and predicted future changes in rainfall patterns vary greatly around the globe [32]. For example, in the southwest United States, rainstorms have become more frequent but less intense, with decreasing storm depths caused by changes in the Hadley cell over the Pacific Ocean [21,28]. Using the observed changes in rainfall patterns, climate model predictions indicate a decrease of 15% in water availability (annual rainfall minus evapotranspiration) over the next two to three decades [69]. Comparatively, rainstorms in sub-Saharan Africa are predicted to become more intense and less frequent [29]. Explanations for these changes include increases in surface albedo and increases of dust particle concentrations and drop nucleation [77]. Decreased infiltration and expansion of bare soils, amplified by changes in the distribution of rainfall for sub-Saharan Africa, will lead to greater runoff and decreased soil moisture.

Dryland vegetation responds strongly and rapidly to changes in soil moisture through shifts in stomatal conductance [33], and soil moisture dynamics are governed by the daily arrival of storms and the dynamics of plant water use between storm arrivals. Despite the importance of daily rainfall processes in determining dryland soil moisture and plant water use, a review of 41 recent models

* Corresponding author. Tel.: +1 609 651 3346.
E-mail address: tfranz@princeton.edu (T.E. Franz).

0309-1708/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2009.12.003

on arid and semi-arid grazing found only five models that included daily rainfall as a forcing variable [72]. The vegetation response to highly temporally variable rainfall pulses has been found to be nonlinear [2,33] and temporal averaging can lead to erroneous solutions of the vegetation response. Previous studies have found daily rainfall to be well represented by stochastic processes [33,59]. In addition to temporal variations, rainfall in many dryland ecosystems is highly heterogeneous in space, which is caused by the occurrence of localized convective storms [12].

Besides rainfall, the representation of vegetation needs to be further refined. Although fractional woody cover is the most common description of dryland vegetation structure, the ubiquity of water limitation indicates that differential water use across functional groups or individual species may also be critical. Plant water use is species-dependent [73], which suggests that differences in dryland species composition can result in substantial differences in basin water use [39,53]. Given the complexity of ecological and environmental factors that together determine the spatial patterns of species occurrence, a general framework for describing factors that govern geographical patterns of species extent has not yet been developed. However, the field of biogeography has addressed species distributions and local inter-species competition using the niche concept [18,24,27,30]. Here, we adopt Hutchinson's definition, which defines a niche as the subset of environmental conditions that affect a particular organism and determine its absolute fitness [30,34]. We refer to fitness as the ability of an individual to grow, reproduce, and survive and will provide a formal definition later in the paper. Modeling of species distributions using niche-based approaches has been primarily studied in two ways. The first method uses correlation between abiotic drivers (e.g. soil, temperature, rainfall) and species occurrence to statistically represent the niches in multidimensional space [3,50]. While this method obtains satisfactory correlation with observed vegetation patterns, it fails to address specific causal mechanisms governing species distribution patterns. In addition, this framework is difficult to extrapolate to novel groups of species [3,34]. The second method adopts a mechanistic approach that describes plant fitness, explicitly modeling growth, reproduction, and survival. However, at regional and continental scales, these models require many input parameters making them difficult to parameterize [43,44,54].

This paper is motivated by a need to develop more detailed frameworks of coupled water and vegetation systems in drylands while avoiding overparameterization and model complexity in order to address issues such as climate change. The two main challenges we address with our proposed framework are the representation of rainfall as a daily stochastic process and the refinement of fractional woody cover into predictions of the spatial distribution of species.

In this work, we use a previously-developed stochastic soil water balance model [35,36,55,58] to represent the interactions between climate, soils, and plants. Such models have been used on a precipitation gradient in the Kalahari desert [63–65,74–76], the Upper Rio Salado basin of New Mexico [7], and a California oak savanna [9]. The model is nonspatial and determines average, steady-state growing season values of runoff, leakage, interception, and evapotranspiration of a single species. We expand the model into a spatial context by applying an optimality trade-off hypothesis which states that dryland vegetation patterns are constrained by maximization of water use and simultaneous minimization of water stress [8]. The model is forced by daily rainfall, which is represented as a marked Poisson process described by the mean depth of daily rainfall and the mean arrival rate of storms. For the basin presented in this study, we generate spatial estimates of each parameter and analyze temporal trends in the variables using a long-term daily precipitation dataset. The stochastic soil water balance model is used to determine each component of a fitness vector for all species, where growth and reproduction are estimated as the ratio of evapotranspiration to growing season rainfall, and plant survival is estimated with the dynamic water stress over the growing season. The contribution of each of these components to overall plant fitness is unknown. The skill and accuracy of our model results are tested against two separate model cases that represent the end members on a continuum of possible model skill and accuracy. The lower bound case is described by selecting a species at random. The upper bound case, a neural network, is a powerful predictive model for linear and nonlinear data [26,37].

This research has three objectives. The first is to extend an existing mechanistic approach in order to predict changes in the distribution of woody plant species in a central Kenyan watershed caused by specific climate change scenarios based on historical rainfall observations. Our second objective is to predict changes in species patterns and relative abundance in response to changes in total growing season rainfall or changes in the variability of growing season rainfall. Our third objective is to determine the relative importance of plant water use and plant water stress in determining the overall distribution of three *Acacia* tree species, *Acacia drepanolobium, Acacia tortilis*, and *Acacia xanthophloea*.

The remainder of the paper is organized as follows: first, in the methods section, we describe our study site and key meteorological data necessary to apply the stochastic water balance model. We also introduce our modeling framework and calibration procedure we use to apply the fitness vector framework we have adopted for predicting individual species distributions. The methods section ends with a brief examination of model sensitivity to key parameters. Our results provide model predictions of species distribution patterns based on observed trends in rainfall variability, as well as the differing effects of changing either the mean or variability of growing season rainfall. Our results conclude with an examination of our fitness vector framework's performance as well as the inferred relative importance of water use and water stress in governing species distribution within our study basin. We conclude with a discussion section that explores the implications of our fitness vector results, the potential for our approach to explain observed shifts in dryland species composition, and a discussion of model limitations and plans for future refinement.

2. Methods

2.1. Basin description

Our study basin is the Upper Ewaso Ng'iro river basin (15,200 km^2), which is located in the central Kenya highlands. The basin spans gradients of elevation, temperature, precipitation, and contains eight different soil texture classes (Figs. 1 and 2). Because the basin is located on the equator, the annual climate consists of two rainy seasons caused by the Intertropical Convergence Zone (ITCZ). Temperature and precipitation patterns are heavily influenced by elevation (Figs. 2a and 4a). Soil texture ranges from sandy clay to clay soils (Fig. 2b, 1980 UNESCO Soil Map). The gradients of temperature, precipitation and different soil texture classes lead to a complex mosaic of vegetation species composition.

The Upper Ewaso Ng'iro river basin is a typical savanna ecosystem with woody vegetation (mostly *Acacia*) and grasses (Fig. 2c). In this analysis, we focus on the 60% of the basin dominated by three *Acacia* species, *A. drepanolobium*, *A. tortilis*, and *A. xanthophloea*. We have chosen to study these species because they represent a diversity of water use strategies, and each occurs across a wide range of the basin. Moreover, these species are described by prior studies, so their water use characteristics are well-known. *A. drepanolobium* is the dominant species on the black cotton clay soils and has been

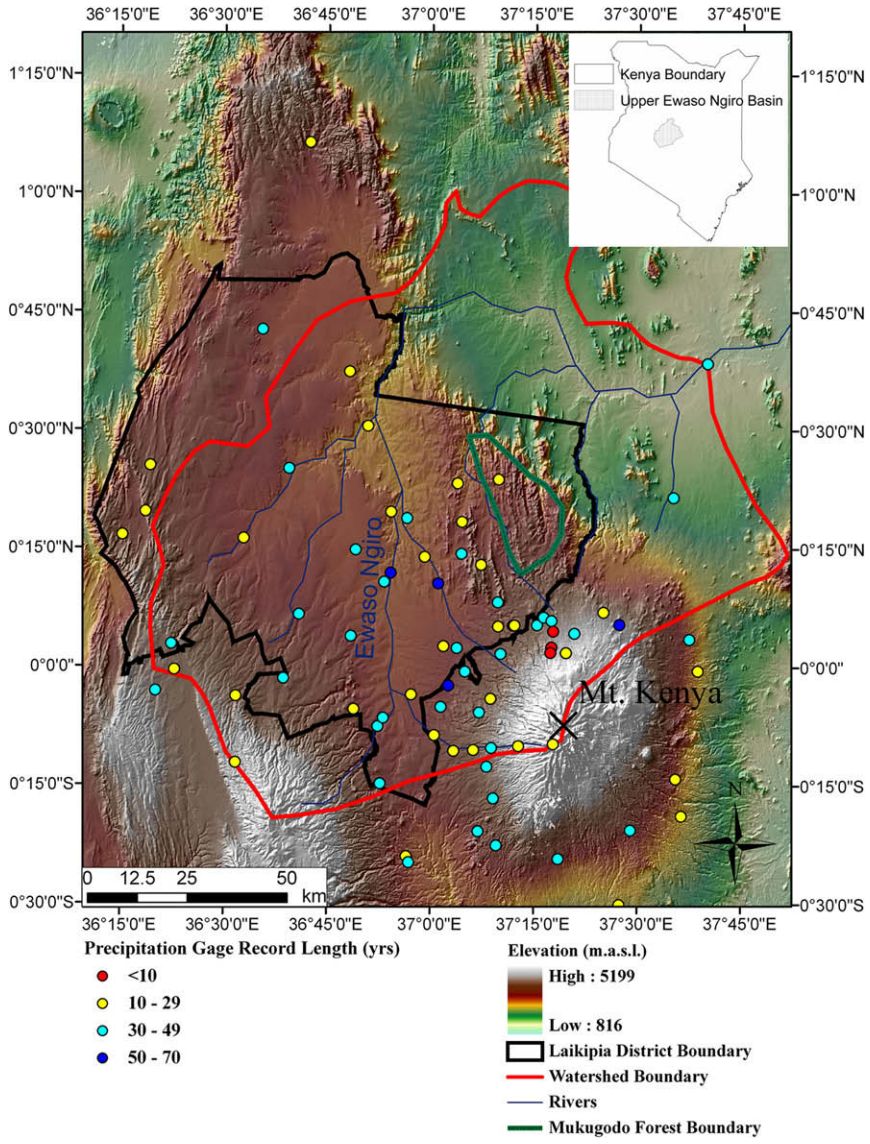

Fig. 1. Location of the study area in the central Kenyan highlands. Mount Kenya (5199 m.a.s.l.) is located in lower right hand corner. The figure illustrates the steep elevation gradient, Upper Ewaso Ng'iro river watershed boundary (15,200 km^2) (red line), Laikipia District boundary (black line), Mukogodo Forest boundary (green line), and daily precipitation gauges organized by record length. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

characterized as having a shallow wide spreading root system with a tap root [14]. *A.* tortilis is a drought tolerant species and has a deep tap root, lifting water hydraulically from deep sources [38]. The most geographically diverse species is *A. xanthophloea,* which occurs in riparian, upland clays, and a few select dry sandy areas. The remaining areas (~40%) of grass, shrubland, and woody vegetation are not included.

2.2. Basin rainfall dynamics

We examined the rainfall patterns in the Upper Ewaso Ng'iro river basin using daily precipitation gauges with record lengths of 5–75 yr (Fig. 1, data provided by NRM3 of Nanyuki, Kenya). We analyzed the data for temporal trends in two parameters: the average rain depth per storm, α (mm), and the average storm arrival rate, λ (day^{-1}), during each of the two rainy seasons (March–May and October–December). Significant temporal trends of increasing α and decreasing λ ($p < 0.05$, Table 1) were found using the Mann–Kendall statistical test [40].

While spatial variability in rainfall precludes any certain statements regarding whole-basin trends, we use the Jacobson Farm gauge, which had the longest record length at 75 yr (1934–2008), to demonstrate the potential for shifts in rainfall processes to occur in this region. Although total precipitation at the Jacobson Farm gauge did not change significantly for either rainy season, increases in α and decreases in λ were both statistically significant ($p < 0.05$, Fig. 3). A majority of the increasing trend in α and decreasing trend in λ can be attributed to the inter-annual variability in precipitation between 1975 and 1990. The inter-annual variability observed in the Jacobson Farm gauge is well within the recent historical observations over East Africa [45], and greater variability in rainfall has been documented in East Africa within the last two centuries [47].

The ensemble of global climate change models (SRES A1B scenario) predicts an increase of 0.5–0.6 mm day^{-1} by 2080–2099 compared to 1980–1999 for all of East Africa (Figs. 10–12 in [41]). In addition, the ensemble of global climate change models predicts an increase of 1.25–1.5 mm in the standard deviation of inter-annual rainfall variability by 2080–2099 compared to

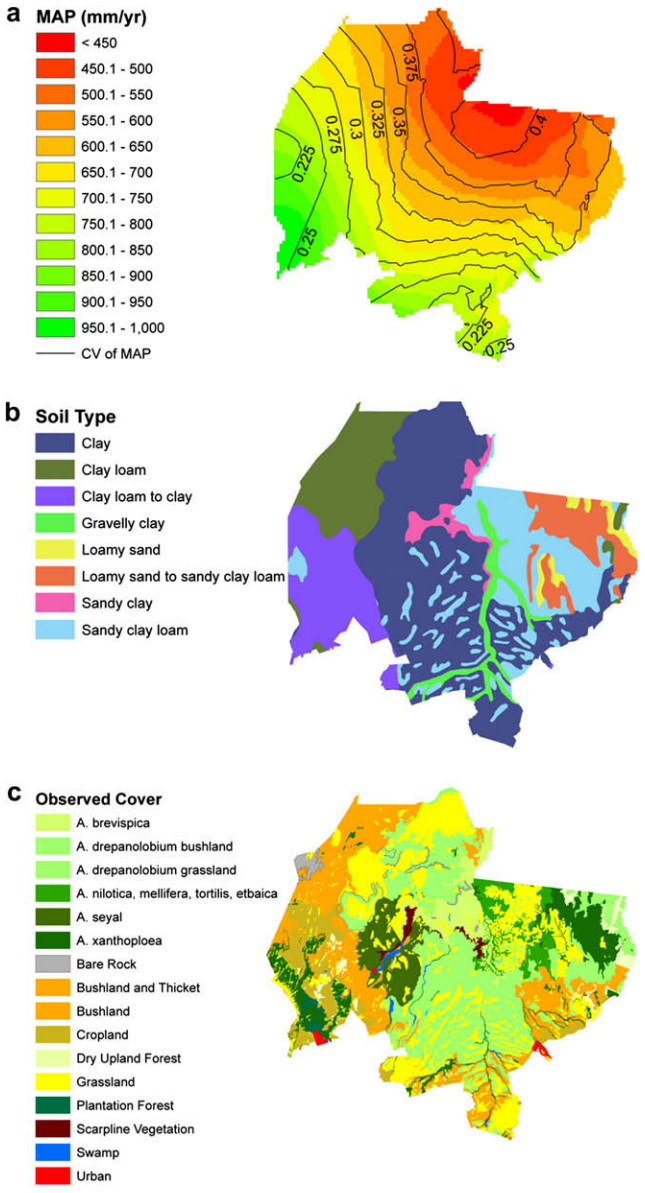

Fig. 2. The Upper Ewaso Ng'iro river basin is a heterogeneous landscape. (a) Driven by the changes in elevation, there are steep gradients in mean annual precipitation (MAP) and contours of the coefficient of variation of annual rainfall (mean divided by standard deviation). (b) The study area contains a wide range of soil texture classes from clay to sandy soils (data source 1980 UNESCO Soil Map). (c) The ecosystem is a classical tree–grass savanna with a diversity of around 10 acacia tree species, some dominant and others coexisting in areas (data source Mpala Research Centre).

1980–1999 for East Africa (Figs. 10–18 in [41]). As a simple experiment and conservative estimate of future changes in rainfall patterns, we approximated the impact of future changes in rainfall patterns on woody species vegetation distributions by linearly extrapolating observed trends of increasing α and decreasing λ.

2.3. Ecohydrological model

The mechanistic framework chosen is based on the conservation of water within the rooting zone of each species, forced by daily precipitation represented as a stochastic process. The following is a brief summary of the model, which has been presented elsewhere in full [35,36,55,58,59], see Appendix A. The rainfall $R(t)$ (mm day^{-1}), is represented as a marked Poisson process of storm arrivals in time with rate λ (day^{-1}), storm depth h (mm), where h is treated as an exponentially distributed random variable with mean α (mm). The interactions of climate, soil, and vegetation were used to determine an analytical solution to the steady-state probability distribution of soil moisture over the growing season at a point for a single species. The various components of the water balance are described by the steady-state probability distribution of soil moisture over the growing season (Eq. (A.10)). The model neglects lateral redistribution of water in the subsurface and interactions with the water table (if present). The assumption is justified for arid and semi-arid conditions where the water table is sufficiently deep (deeper than the rooting depth of each species) and subsurface flow is negligible [57]. Well data records observed at one location in the basin (Mpala Research Center, MAP = 500 mm) indicate a water table depth of at least 100 m [60].

The maximum daily potential evaporation rates, E_{max} (mm day^{-1}), are estimated for each species using the Penman–Monteith formulation, scaled by maximum canopy conductance [7]. The Penman–Monteith equation determines potential evapotranspiration (PET) according to consideration of both available energy, R_{net} (W m^{-2} day^{-1}), and vapor pressure deficit. We determine vapor pressure deficit using relative humidity, RH (%), air temperature, T_a (°C), and the assumption that leaf temperature is the same as air temperature. Our approach further assumes that PET is limited by both canopy and atmospheric conductance. We define canopy conductance as the product of species leaf area, LAI (m^2 m^{-2}), and maximum stomatal conductance, gs_{max} (mmol m^{-2} s^{-1}), and determine atmospheric conductance based on average wind speed, u (m s^{-1}), and canopy height, h_{tree} (m). Details of this approach are given elsewhere [7]. We find that growing season air temperature, T_a ($T_a = -0.0069 * z + 32.045$, $n = 14$, $R^2 = 0.95$), and pan evaporation, E_{pan} (mm day^{-1}) ($E_{pan} = -0.0024 * z + 9.8255$, $n = 14$, $R^2 = 0.70$) are linearly related to basin elevation, z (m), while growing season daily relative humidity, RH (mean = 69.5, standard deviation = 4.53, n = 10, R^2 = 0.36), and daily wind speed, u (mean = 1.95, standard deviation = 0.618, n = 14, R^2 = 0.056), are not linearly dependent on elevation (daily

Table 1
Summary of basin rainfall statistics for gauges with daily record lengths greater than 40 yr (n = 11). The data suggests that total precipitation remained constant but the intensity and frequency of storms increased through time over the period of record in the basin (data source NRM3).

Parameter description	% of stations with statistically significant trends ($p < 0.05$, Mann–Kendall test)	Mean slope of all stations through time (Sen's method)	Standard error of slopes
Total rainfall (mm), short rains (October–December)	9.1	0.298	0.333
Total rainfall (mm), long rains (March–May)	0.0	−0.609	0.286
Annual rainfall (mm)	9.1	−0.652	0.930
Average rain per storm (mm) α, short rains	27.3	0.043	0.016
Average rain per storm (mm) α, long rains	36.4	0.040	0.017
Average arrival rate of rain events (day^{-1}) λ, short rains	18.2	−0.0010	0.00043
Average arrival rate of rain events (day^{-1}) λ, long rains	45.5	−0.0018	0.00057

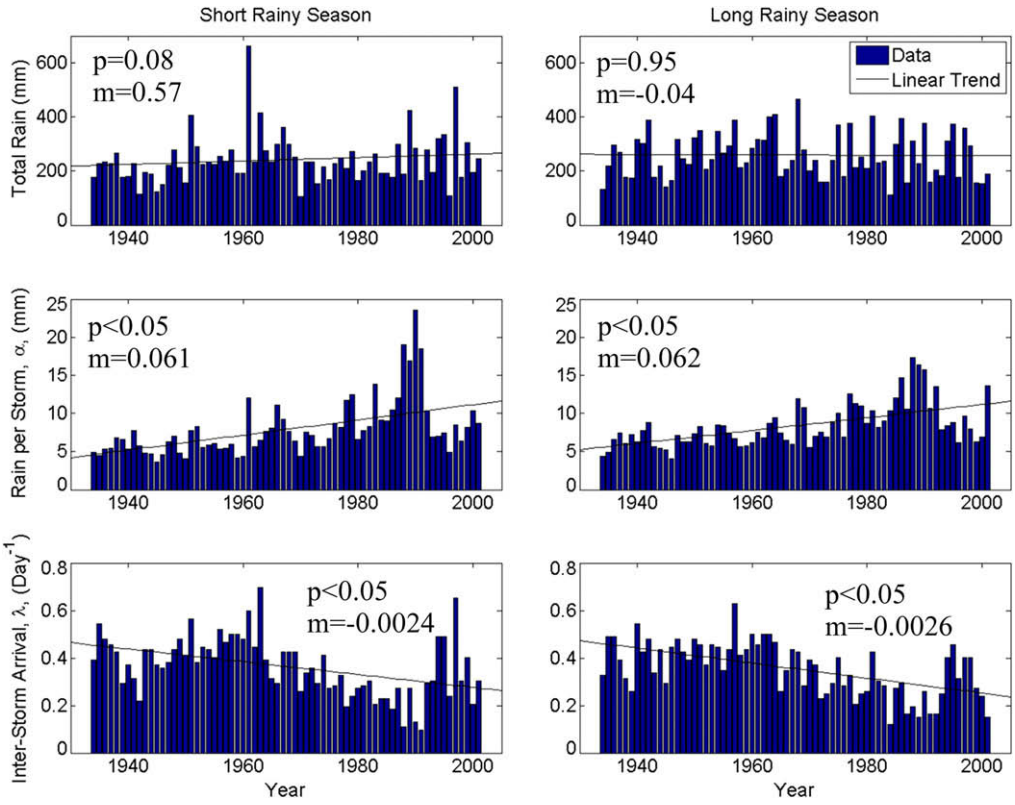

Fig. 3. Time series of total seasonal precipitation (mm), average rain per storm (mm) α, average arrival rate of rain events (day^{-1}) λ, for short (October–December) and long (March–May) rainy seasons for Jacobson Farm gauging station (record length = 75 yr). Trend analysis indicates total precipitation is not changing significantly with time but storm depth and arrival rate of storm events are changing significantly (data source NRM3). Note: Trends in Jacobson Farm data are not necessarily indicative of entire basin.

gauge data provided by NRM3 of Nanyuki, Kenya used in Fig. 4). Therefore, we assume basin average values for *RH* and *u*.

The gauge density of precipitation (~60 stations across the basin, Fig. 1) is high enough to use an ordinary kriging algorithm to make spatial estimates of α and λ (Fig. 4). Soil parameters are estimated from the 1980 UNESCO Kenyan Soils Map, its description key, and commonly-used relationships between soil texture and soil hydraulic characteristics (Table 2 and Fig. 2). The plant parameters, gs_{max}, ψ_* (MPa), and ψ_{wilt} (MPa), are estimated for the three species from prior leaf-scale physiological measurements. We specify plant height, h_{tree}, based on field observations. Interception storage, Δ (mm), is taken to be a simple function of LAI ($\Delta = 0.043 * LAI$). Specific references for parameter estimates, relative sensitivity, and certainty can be found in the descriptions of Table 3.

The climate, soil, and vegetation parameters explained above are used in the solution of the steady-state probability distribution of soil moisture [36], which was subsequently used to describe plant water stress during the growing season (Eqs. (A.6)–(A.9)). Specifically, water stress begins to occur whenever soil moisture values drop below the value of s^*, the value corresponding to the moment when plants begin to close their stomata. Water stress values then increase to a maximum value of one as soil moisture approaches plant wilting point, s_w. At soil moisture values greater than s^*, plants are assumed to be stress-free with respect to water availability. Here we assume that all *Acacia* species respond similarly to water stress, such that soil moisture values approaching the wilting point will be weighted more heavily than values near incipient stomatal closure ($q > 1$). We address this assumption in more detail in Section 2.5. While useful as a general measure of plant water deficit, the daily static water stress (Eq. (A.6)) does not account for the seasonal distribution of the frequency and duration of stress periods below the critical value of s^*. The dynamic water stress, $\bar{\theta}$ (Eq. (A.9)), uses the crossing properties of soil moisture over the growing season and is a normalized value between 0 and 1 with 1 representing the most stressed conditions.

With a description of seasonal water use (Eq. (A.10)) and water stress (Eq. (A.9)) for a single point and species we partition the landscape into individual grid cells. The landscape is resolved as 616 unique combinations of the spatially continuous soil and climate input parameters. The size of each grid cell (mean = 10.3 km^2, standard deviation = 28.1 km^2, max = 284.5 km^2, min = 0.5 km^2), varies according to local gradients of rainfall parameters (parameter resolution of $\lambda = 0.01$, $\alpha = 0.1$ mm) and changes in soil texture across the basin. Each grid cell is independent and represented as a location containing a single set of species, soil, and climatic parameters. In each grid cell, growing season water balance, plant water use, and plant water stress are evaluated for each of the three species, *A. drepanolobium*, *A. tortilis*, and *A. xanthophloea*.

2.4. Fitness vector and calibration

The ecohydrologic water balance framework presented in the previous section has been shown to be a satisfactory model for representing the steady-state probability distribution function of soil moisture over the growing season and subsequent representations of evapotranspiration and plant water stress [49,61]. The question that remains unanswered is what are the tradeoffs between resource use (evapotranspiration of available rainfall) and resource scarcity (water stress)? The resolution of this question critically determines the predictions of species distribution. To address this problem, as well as propose a framework for species selection based on competing resources, let the vector $\vec{F}$, represent the different components of fitness for a particular species or functional

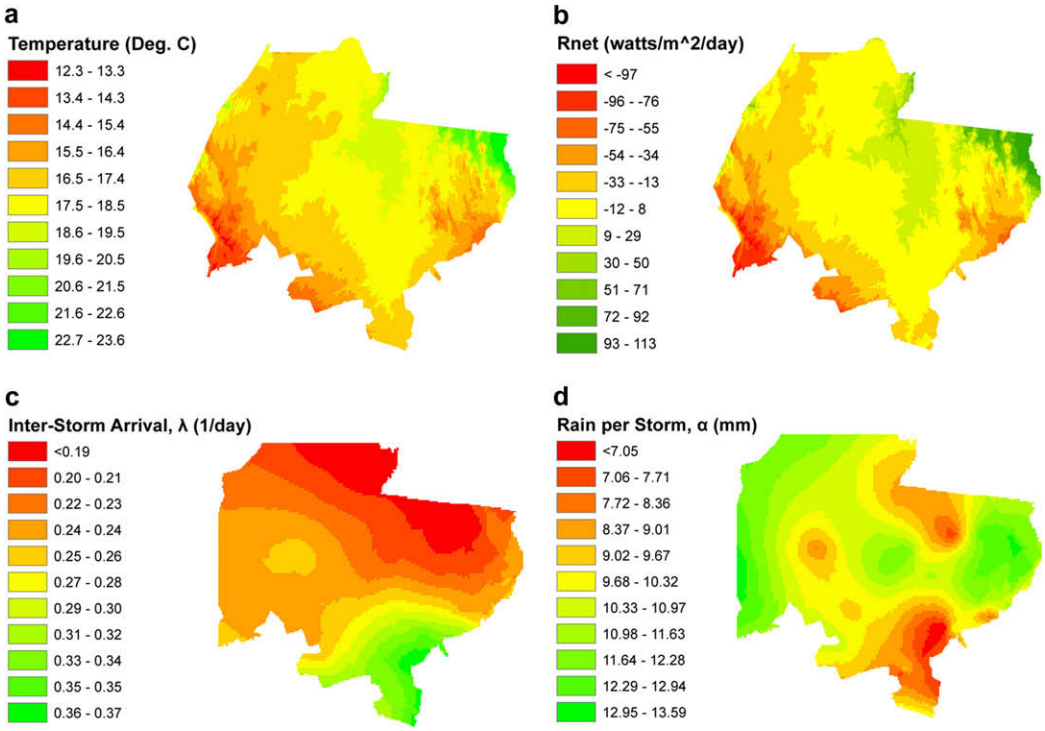

Fig. 4. Average growing season climatic forcings used to drive the model. Air temperature, T_a (°C), and pan evaporation, E_{pan} (mm day^{-1}), data are linearly correlated with elevation, z (m) (Section 2.3) (a); net radiation, R_{net} (W m^{-2} day^{-1}) is estimated from pan evaporation using the inverse of the Penman-Combination equation [15] (b); average arrival rate of storm events, λ (day^{-1}) (c); and average rain per storm, α (mm) (d); are estimated with ordinary kriging ($n = 60$) (data source NRM3).

Table 2
Summary of soil parameters used in model, where $\psi = \psi_s s^{-b}$ [10,15,59].

Soil type	S_h ($\psi = -10$ MPa)	S_{fc} ($\psi = -0.03$ MPa)	ψ_s (MPa)	b	K_s (cm day^{-1})	n
Clay	0.47	0.78	−0.00182	11.40	1	0.500
Clay loam	0.36	0.70	−0.00160	8.50	10	0.475
Gravely clay[a]	0.08	0.32	−0.00034	4.05	200	0.350
Loam	0.19	0.56	−0.00143	5.39	20	0.450
Loamy sand	0.08	0.30	−0.00017	4.38	100	0.420
Sandy clay[a]	0.43	0.74	−0.00150	10.50	5	0.480
Sandy clay loam[a]	0.32	0.66	−0.00115	8.00	45	0.465
Sandy loam	0.17	0.49	−0.00070	5.39	80	0.430

[a] Estimated value.

group. The components of the vector are dependent on the ecosystem and vegetation community in question. The following assumptions are used for this study of woody vegetation in dryland ecosystems: (1) growth and reproduction are represented by total evapotranspiration, which we determine using the modeled evapotranspiration over the growing season (Eq. (A.10)); (2) survival is represented by plant water stress, which we determine using the modeled dynamic water stress over the growing season (Eq. (A.9)); (3) nutrient use and efficiency are neglected under the assumption that changes in soil water dynamics will be the primary determinant of shifts in dryland vegetation patterns; (4) the relative importance of growth and reproduction versus survival are not weighted equally to overall plant fitness and require rescaling coefficients, which we determine using the observed vegetation cover in the study basin. The selection of the most "fit" species in a given grid cell is defined mathematically as follows:

$$Opt = \text{find species } n \text{ which is the maximum of } \|\vec{F}_n\|, \tag{1}$$

where *Opt* is the optimal species selected by the maximum value of the magnitude of the fitness vector for each species n:

$$\vec{F} = s_E \left(\frac{\langle E \rangle}{\langle R \rangle}\right)^{n_E} i + s_{\bar{\theta}}(1-\bar{\theta})^{n_{\bar{\theta}}} j, \tag{2}$$

where i, j denote the components of the fitness vector and s_E, n_E, $s_{\bar{\theta}}$, $n_{\bar{\theta}}$ are calibrated weighting coefficients all greater than zero. Both components of the fitness vector are values between 0 and 1 and defined in terms of the primary modeled state variable of soil moisture. We determined the three weighting coefficients (s_E is fixed as equal to 1 because there are only three independent coefficients) by maximizing the percent agreement between modeled and observed species cover. The optimal coefficients were found with a non-gradient search algorithm, following initial exploration of the three-dimensional coefficient space to select appropriate starting values that avoided local minima. Visualization of the coefficient space revealed smooth contours lending confidence to the coefficient selection of the global minimum solution. The fitness vector provides a framework for quantifying the tradeoff of soil moisture between evapotranspiration and stress by modeling the daily mass balance of water and selecting the optimal species in each grid cell in the basin.

Table 3
Vegetation parameter estimates associated by species with descriptions of relative sensitivity and certainty in model parameters.

Species	*Acacia drepanolobium*	*Acacia tortilis*	*Acacia xanthophloea*	Relative sensitivity	Relative certainty	Source
E_{max} (cm day^{-1})	Basin	Basin	Basin	High	Medium	1, 2, 3
E_w (cm day^{-1})	5% of E_{max}	5% of E_{max}	5% of E_{max}	Low	Low	5, 9
gs_{max} (mmol m^{-2} s^{-1})	325[a]	325	375	Low	Medium	4
h_{tree} (m)	3.0	3.0	10.0	Low	High	6, 10
K	0.5	0.5	0.5	Medium	Low	9
LAI (m^2 m^{-2})	1.57 (0.40)[b]	1.04 (0.28)[b]	2.46 (1.15)[b]	Low	High	11
q	2	2	2	Medium	Medium	9
T_{seas} (day)	185	185	185	Medium	High	10
Z_r (cm)	75	75	90	High	Medium	6, 8, 10
Δ (cm)	0.075[c]	0.050[c]	0.100[c]	Low	Medium	10
ψ_* (MPa)	−0.45[a]	−0.45	−0.08	High	Medium	4, 7
ψ_{wilt} (MPa)	−3[d]	−4	−1.9	Low	Medium	4, 7

1: [42].
2: [15].
3: [48].
4: [51].
5: [7].
6: [14].
7: [22].
8: [13].
9: [59].
10: Unpublished data (Franz) from field visits (2006, 2007, 2008, 2009).
11: MODIS 8-day, 1-km LAI/fPar product imagery from growing season June 1, 2004 DOY 153.
[a] Value adopted from *A. tortilis*.
[b] Mean and standard deviation of LAI. Pixel by pixel values were averaged using the species land cover maps and the MODIS imagery. A minimum of 1000 pixels for each species were used to estimate mean and standard deviation.
[c] Value estimated for *Acacia* trees, $\Delta = 0.043 * LAI$.
[d] Value adopted from *A. etbaica*.

Table 4
Summary of basin wide determinants of modeled species spatial patterns derived from principal component analysis (PCA). The PCA coefficients (EV_1, EV_2, EV_3) indicate the different loadings of each variable to the ranked eigenvalue. The first eigenvalue is weighted most by the four soil parameters, b, K_s, n, ψ_s, the second eigenvalue is weighted most by temperature, T_a, and net radiation, R_{net} (both functions of elevation, z), and the third eigenvalue is weighted most by the rainfall parameters, α and λ.

	Basin mean	Basin standard deviation	PCA coefficients		
			EV_1	EV_2	EV_3
α (cm)	0.990	0.157	−0.2218	−0.3059	0.6096
λ (day^{-1})	0.258	0.053	0.2488	0.3526	−0.5257
T (°C)	17.93	1.52	−0.2389	−0.5236	−0.3970
R_{net} (W m^{-2} day^{-1})	3.29	29.06	−0.2258	−0.5247	−0.4191
b	7.59	2.58	0.4590	−0.1685	−0.0206
K_s (cm day^{-1})	60.70	59.85	−0.4188	0.3125	−0.1103
n	0.451	0.044	0.4297	−0.2946	0.0781
ψ_s (MPa)	−0.001062	0.000582	−0.4584	0.1404	−0.0137

Calibration between the observed species (Fig. 2c) versus the modeled species required rescaling of each component of the fitness vector (Eq. (2)) because the relative importance of evapotranspiration versus stress avoidance was unknown. We assumed the relative importance of evapotranspiration versus stress avoidance for plant fitness in dryland ecosystems was independent of species, life history, space, and time, therefore allowing us to assign only one set of values for all species throughout the basin. Due to the time lag for trees to reach maturity the calibration of the fitness vector weighting coefficients was performed between the modeled species with 1950 rainfall parameters and the observed vegetation. After calibration, the performance of the model versus observed species cover was: 78.7% by total area covered by the three species, 89.4% by *A. drepanolobium*, 46.0% by *A. tortilis*, and 57.7% by *A. xanthophloea*. Assuming the rainfall patterns of the entire basin are changing similarly (Table 1), we obtained the 1950 rainfall parameters by using a linear trend with slopes for $\alpha = 0.04$ mm yr^{-1} and $\lambda = -0.0018$ day^{-1} yr^{-1} as defined by the average slope of stations with a record length greater than 40 yr. (Note: Jacobson Farm gauge slope estimates of α and λ are nearly double, Fig. 3.) The midpoint of the precipitation dataset, neglecting stations with record lengths less than 10 yr, is approximately 1975, which we assume as the long-term average values of α and λ (Fig. 4c and d).

2.5. Parameter sensitivity

Given the lack of information in the literature and difficulty of making direct parameter estimates across the basin and within species, we used multivariate statistical analyses and sensitivity analyses to rank the importance of model input parameters. We performed a principle component analysis (PCA) of the eight input parameters $(\alpha, \lambda, T_a, R_{net}, b, K_s, n, \psi_s)$ representing the interactions of climate and soil on the spatial distribution of woody species in the 616 unique grid cells defined in the basin (Table 4). Comparing the relative values of the ranked eigenvalues, we found that the first three eigenvalues account for nearly 95% of the explained variance of the three species in the basin. Evaluating the estimates of the first three PCA coefficients, EV_1, EV_2, and EV_3 (Table 4), we found that a majority of the first axis, EV_1, is weighted more heavily by the four soil texture variables (b, K_s, n, ψ_s). The first axis accounts for 53% of the species variance, which indicates that soil texture has a dominant control on current vegetation patterns

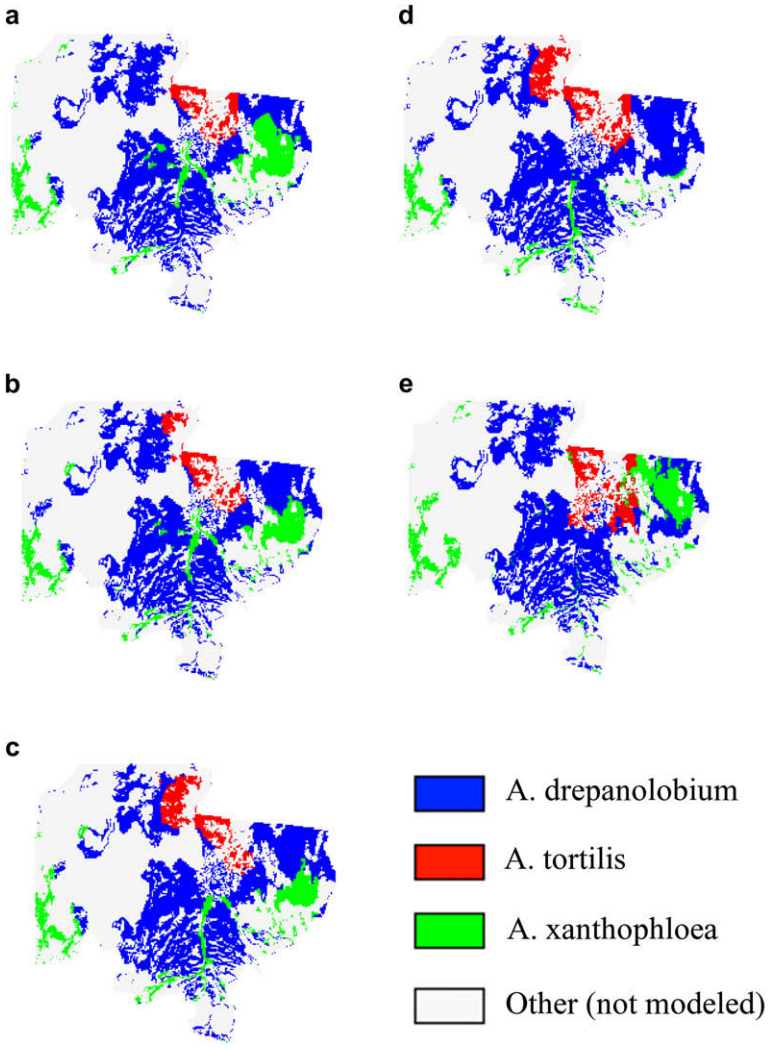

Fig. 5. Model results of woody species distribution patterns using linear trends in time for basin average rain per storm α (0.04 mm yr^{-1}) and average arrival rate of storm events λ (−0.0018 day^{-1} yr^{-1}), where (a) is rainfall from 1950, (b) 1975, (c) 2000, (d) 2025, and (e) is observed patterns. The calibration results between modeled (a) and observed land cover (e) are: 78.7% by total area, 89.4% by *A. drepanolobium*, 46.0% by *A. tortilis*, and 57.27% by *A. xanthophloea*. The changes in species cover with time suggest an upslope migration of *A. xanthophloea* and *A. tortilis* (a–d) to areas with greater total precipitation in response to decreased infiltration from more intense infrequent storm events.

within the basin. The second orthogonal axis accounts for 25% of the variance and is weighted more heavily by T_a and R_{net}, which controlled available energy for evapotranspiration across the basin. The growing season parameters T_a and R_{net} were found to be linear functions of elevation, z, in the basin (Section 2.3 and Fig. 4), which indicates the variance explained by the second axis is due to the elevation gradient. The third orthogonal axis accounts for 16% of the variance and is weighted more heavily by the rainfall parameters α and λ.

Although soil and climate play a large role in determining species distribution patterns, the model is also sensitive to plant parameters. We characterized the model sensitivity to plant parameters by changing a single parameter value in select grid cells around the basin and categorizing the relative sensitivity (Table 3). The model parameters were organized into three sensitivity categories, low, medium, and high, based on comparing the ratios of the perturbed parameter value to the mean parameter value that caused a change in species type. We found that the parameters E_{max}, Z_r, and ψ_* had the greatest impact on the modeled species with relatively small changes in magnitude. The maximum evapotranspiration rate, E_{max}, was estimated with the Penman–Montheith equation and was a function of both climate $(T_a(z), R_{net}(z), u, RH)$ and plant parameters $(gs_{max}, h_{tree}, LAI)$. The interplay between climate and vegetation made it difficult to distinguish any one parameter $(T(z), R_{net}(z), u, RH, gs_{max}, h_{tree}, LAI)$ as dominant, we therefore categorized the individual input parameters as low relative sensitivity [5]. The effective rooting zone, Z_r, was found to have high relative sensitivity as it had a large impact on modeled soil moisture. The matric potential at incipient stomatal closure, ψ_*, was very sensitive because it changed the soil moisture value at which water stress began. Because this is a water-controlled ecosystem, changing the value at which stomatal closure occurred had a large effect on evapotranspiration, which was the other determining factor for species distribution patterns (Eq. (2)). The parameters, k and T_{seas}, scaled the dynamic stress over the growing season and exhibited a medium relative sensitivity (Eq. (A.9)). The parameter controlling the concavity of the relationship between soil moisture and stress, q, was found to have a medium relative sensitivity. It is very likely that $q > 1$, so that dryland vegetation will become more stressed as soil moisture values approach the wilting point. However, see [59] for a detailed discussion of the rationale for selection of k and q values. The parameter

estimates of evaporation, E_w, near the wilting point, and ψ_{wilt}, had low relative sensitivity, which was consistent with our assumption that evapotranspiration of vegetated areas is primarily controlled by plant transpiration and not bare soil evaporation. The canopy interception, Δ, was found to have low relative sensitivity because average storm depths in the growing season were much greater than canopy interception values.

3. Results

3.1. Modeled changes in species patterns

We investigated changes in the species distribution patterns using two sets of rainfall scenarios. With the first set of scenarios we explored changes in the species patterns driven by the mean slope of linear trends in time for the basin rainfall parameters, $\alpha = 0.04$ mm yr^{-1} and $\lambda = -0.0018$ day^{-1} yr^{-1}. The changes in rainfall were applied uniformly across the basin and the new species patterns were estimated with the model and fitness vector (Fig. 5). The modeled changes in species cover for the Upper Ewaso Ng'iro River basin indicated the upslope migration of *A. xanthophloea* and *A. tortilis* to areas with higher mean annual precipitation. The modeling results also suggested the expansion of *A. drepanolobium* to areas inhabited by *A. xanthophloea* with the rainfall scenarios from 2025 (Fig. 5d). The modeling results found the territory of *A. tortilis* expanding into the drier parts of the basin occupied by *A. drepanolobium* (Fig. 5c and d).

The second set of rainfall scenarios investigated the impact of changes on the species spatial distribution patterns in response to either changes in the basin mean or variability growing season precipitation. We define basin mean growing season precipitation as $\mu_p = \bar{\alpha}\bar{\lambda}T_{seas}$ (mm), where $\bar{\alpha}$ and $\bar{\lambda}$ are the long-term average basin values given in Table 4 and T_{seas} (day) is the length of the growing season, which we assume to be constant (Table 3). We define σ_p^2 (mm^2) as the variance of growing season rainfall derived from the model representation of rainfall as a marked Poisson process, which is determined by $\sigma_p^2 = 2\bar{\alpha}^2\bar{\lambda}T_{seas}$. These two definitions allow us to test rainfall scenarios where μ_p is changing while σ_p^2 is held constant and σ_p^2 is changing while μ_p is held constant. The results in Fig. 6 illustrate basin-scale changes in spatial distribution patterns of the three modeled species due to changes in only μ_p and σ_p^2. At values of μ_p and σ_p^2 lower than the long-term mean (Fig. 6a, b, e, and f) we found that *A. tortilis* expanded its territory and the territory of *A. xanthophloea* was reduced. In contrast, at values of μ_p and σ_p^2 higher than the long-term mean (Fig. 6c, d, g, and h) we found that *A. xanthophloea* expanded its territory and *A. tortilis*'s territory was reduced in the basin. The model results indicated that increases (decreases) in μ_p (Fig. 6a–d) result in similar spatial patterns of the three species as increases (decreases) in σ_p^2 (Fig. 6e–h), suggesting changes in growing season rainfall variability are comparable to changes in growing season precipitation. We also observed a similar relationship between μ_p and σ_p^2 when looking at the relative fractional cover of the three species in the basin (Fig. 7). The implications of these patterns on how we interpret changes in mean rainfall and rainfall variability in drylands are discussed below.

3.2. Utility of the fitness vector framework

As explained above, rescaling the components of the fitness vector was necessary to estimate the relative importance of water use (fractional evapotranspiration from available rainfall) and avoidance of water stress (dynamic water stress over the growing season). The rescaled weighting coefficients suggest that water use is more important than avoidance of water stress to the overall fitness (Fig. 8). We tested the predictions of the fitness vector framework against a set of other models (Table 5). The models represented the two end members of predictive skill in the set of all possible models and served as a way to bound the results of the fitness vector framework. The measure of model skill was presented with two metrics: percent agreement of each species and Cohen's κ [19]. Cohen's κ is a measure between -1 and 1, where 0 is equivalent to a random model and 1 is perfect agreement between all species in all locations and is evaluated with $\kappa = \frac{\Pr(a)-\Pr(e)}{1-\Pr(e)}$, where $\Pr(a)$ is the probability of agreement of all species and $\Pr(e)$ is the probability of random agreement. The lower bound was represented by selection of a random species for every location or by selection of the same species in all locations. We chose a neural

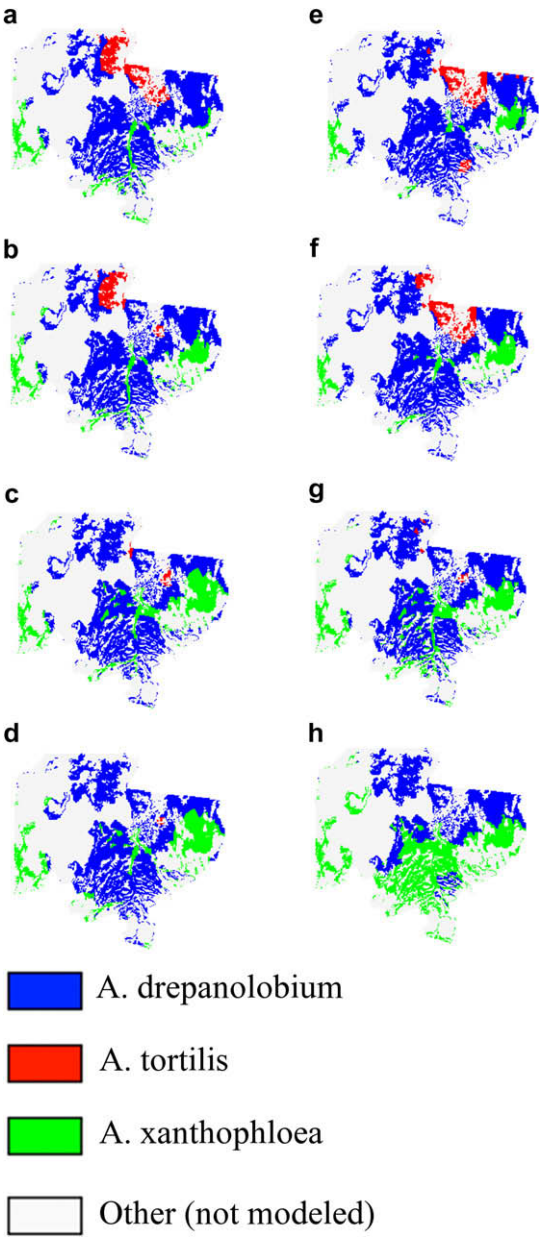

Fig. 6. Model results of woody species distribution patterns using rainfall parameters, α and λ, which vary μ_p while holding σ_p^2 constant (a–d), and vary σ_p^2 while holding μ_p constant (e–h). Increases in either σ_p^2 or μ_p lead to greater extent of *A. xanthophloea* (c, d, g, and h), while decreases in either σ_p^2 or μ_p lead to greater extent of *A. tortilis* (a, b, e, and f). Specific rainfall parameter estimates, α, λ, σ_p^2, and μ_p, for each case can be found in Fig. 7.

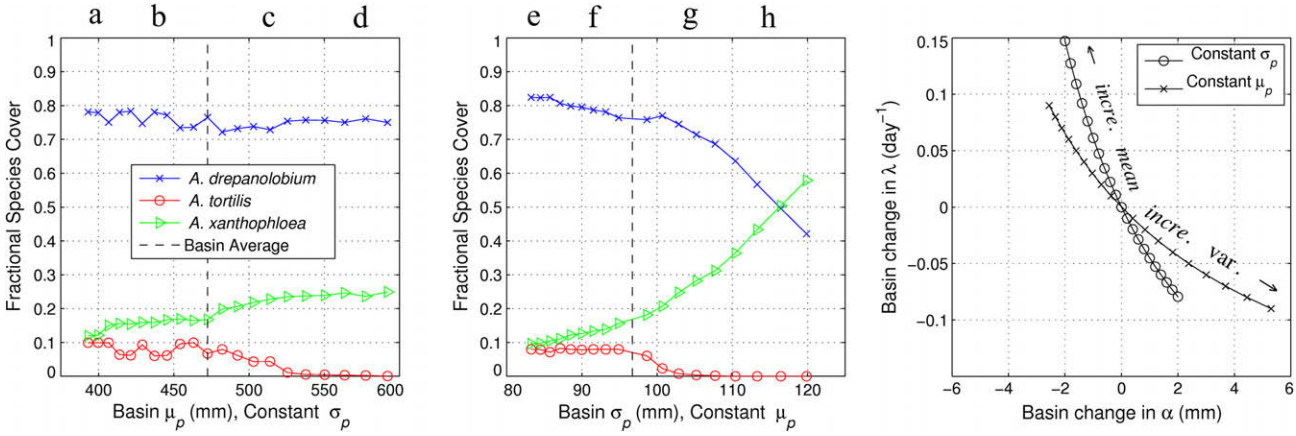

Fig. 7. Model sensitivity study of species fractional cover in the basin under changing σ_p^2 or μ_p while holding the other parameter constant (a and b). The change in species fractional cover is comparable, but inversely related. Rainstorms become more intense and less frequent for the two cases of changing μ_p or σ_p^2 while holding the other constant but the direction of the changes is opposite (c). Specific species distribution patterns in the basin are illustrated in Figs. 6 and 5 (where rainfall = 1975, is defined as the long-term average rainfall).

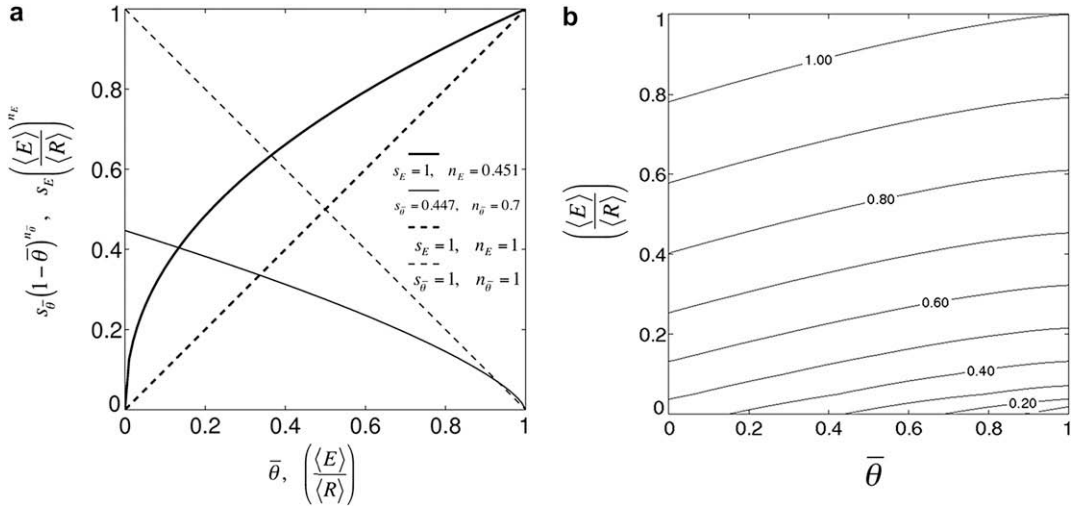

Fig. 8. Rescaled fitness vector components using the calibrated weighting coefficients versus the null hypothesis that all coefficients are equal to one (a) and the resulting contours of the magnitude of the fitness vector (b). The relative values of s_E and $s_{\bar{\theta}}$, indicate that near conditions of maximum water use and minimum stress, transpiration is twice as important as water stress. The relative values of n_E and $n_{\bar{\theta}}$, suggest that stress only begins to decrease fitness at high values, while transpiration substantially increases fitness even at low values.

Table 5
Summary of agreement between different models and observed species cover. Neural network models perform better than the fitness vector framework but at a cost of losing the physical interpretation of the system and not allowing innovative species to compete without prior calibration. A tradeoff between approaches exists between model accuracy and flexibility to innovative systems and interpretation of physical mechanisms.

Methodology	*A. drepanolobium* (%)	*A. tortilis* (%)	*A. xanthophloea* (%)	Total (%)	Cohen's κ
Fitness vector	89.4	46.0	57.7	78.7	0.506
Neural net	94.0	84.2	81.9	90.1	0.802
All *A. drepanolobium*	100.0	0.0	0.0	66.0	0.000
All *A. tortilis*	0.0	100.0	0.0	10.9	0.000
All *A. xanthophloea*	0.0	0.0	100.0	23.1	0.000
Random model	33.3	33.3	33.3	33.3	0.000

network as an upper bound because of their properties as a powerful predictive model for linear and nonlinear data [26,37]. In our implementation, the neural network is represented by a series of connected logistic equations, where the parameter estimates were optimized with nonlinear regression. The inputs to our neural network model were elevation, z (which determines both $T_a(z)$ and $R_{net}(z)$ according to Fig. 4), α, λ, and soil texture. The optimal neural network outperformed the fitness vector, with an accuracy of 90%, compared to 79% for the fitness vector approach and a Cohen's κ of 0.80 compared to a value of 0.51 for the fitness vector approach (Table 5). These results indicate the fitness vector approach fails to capture 11% of the possible information regarding species occurrence. These reasons for this discrepancy are discussed below.

4. Discussion

4.1. Water use versus stress avoidance

While early studies on optimality regarding vegetation patterns focused solely on maximization of resource use [16,17], more recent approaches have begun to consider the additional costs of chronic resource scarcity that accompany over-consumption of limited resources [8,68]. However, we are not aware of any prior uses of ecohydrological optimality approaches to predict basin-scale patterns of individual species distribution. We address the issue of species distribution through the development of a fitness vector framework (Eq. (2)) based on a hypothesis that dryland vegetation patterns are constrained by maximization of water use and simultaneous minimization of water stress [8]. In order to apply this framework to predict the impact of rainfall variability on future vegetation patterns, we define a single set of fitness vector weighting coefficients that maximize agreement with observed species patterns in the Upper Ewaso Ng'iro River basin. The magnitude of these coefficients provides insight into the relative importance and degree of sensitivity that each fitness vector component has on species distribution. Specifically, the relative importance of water use and stress occurrence are given by the relative magnitude of s_E and $s_{\bar{\theta}}$, while the degree of nonlinearity between each of these two terms and their effect on total fitness is described by n_E and $n_{\bar{\theta}}$ (Fig. 8a). In the absence of any prior information, we might assume that water use and water stress are equally important, such that $s_E = s_{\bar{\theta}} = 1$, and that each of these components contribute to total fitness in a linear manner, $n_E = n_{\bar{\theta}} = 1$. These assumptions provide a null hypothesis against which our rescaled fitness vector components can be compared (Fig. 8a). However, when we calibrate the fitness vector weighting coefficients ($s_E = 1,\ s_{\bar{\theta}} = 0.447,\ n_E = 0.451$, and $n_{\bar{\theta}} = 0.700$) we find that instead of being the same, the relative values of s_E and $s_{\bar{\theta}}$ vary such that fractional evapotranspiration from available rainfall (Eq. (2) component i) is more important (i.e. is weighted more heavily) than dynamic water stress (Eq. (2) component j) when determining species fitness. In addition, both n_E and $n_{\bar{\theta}}$ are less than one. In the case of n_E, a value less than one indicates that increasing amounts of water use are less and less important for determining total fitness. In contrast, the fact that $n_{\bar{\theta}}$ is less than one indicates that stress only becomes an important factor in determining overall fitness when stress values are close to one because of its representation in Eq. (2) as one minus dynamic water stress. The magnitude of the calibrated fitness vector (Fig. 8b) illustrates the greater weighting due to fractional evapotranspiration from available rainfall than dynamic water stress. We acknowledge that assigning global weighting coefficients downplays the possibility that individual species have their own specific fitness vector that allows them to occupy specific niches. However, the fact that our framework includes consideration of both water use and water stress allows us to use a single representation of fitness to determine regional patterns of where species exist in the basin, even if the individual species have unique strategies of water use and stress avoidance. For example, the best possible set of weighting coefficients that only considers water use (coefficients $s_E,\ n_E$) results in a basin almost completely occupied by *A. xanthophloea*, and a reduction in total accuracy of 56% . In contrast, the best possible set of weighting coefficients that only considers stress (coefficients $s_{\bar{\theta}},\ n_{\bar{\theta}}$) results in a basin completely occupied by *A. tortilis*, and a reduction in total accuracy of 86%. Therefore, it is only possible to obtain coexistence of all three species using a set of weighting coefficients that consider both resource use and limitation. Subsequently, we must assign global rules to determine basin-scale responses and species-specific distribution patterns.

4.2. The role of changes in the mean and variance of rainfall on species spatial patterns

When observed mean trends in rainfall are incorporated into our ecohydrological model, *A. xanthophloea* and *A. tortilis* are predicted to migrate upslope to areas with higher mean annual precipitation, with *A. drepanolobium* predicted to occupy areas vacated by *A. xanthophloea*. We interpret these shifts in terms of the physiological and structural parameters of the individual species included in this analysis. *A. xanthophloea* is the tallest species (h_{tree}), has the deepest effective root zone (Z_r), and highest stomatal conductance (gs_{max}) resulting in the largest use of water (E_{max}) (Table 3).

The modeled effects to *A. xanthophloea* distribution range are most significant in the Mukogodo forest area (Fig. 1), which is completely eliminated in the model using the interpolated rainfall parameters from 2025 (Fig. 5d). The temporal changes in rainfall parameters (α and λ) lead to greater water use from available rainfall for *A. drepanolobium* and *A. tortilis*, while *A. xanthophloea*'s water use remains relatively static (Fig. 9). We also find that dynamic water stress values increase for all three species. However, because s_E is greater than $s_{\bar{\theta}}$, the fitness vector is most heavily weighted towards water use. Therefore, the predicted result of changing rainfall parameters in the Mukogodo forest is an increase of fitness for *A. drepanolobium* and *A. tortilis* with *A. drepanolobium* outcompeting *A. xanthophloea* when the 2025 rainfall parameters are used in the model. The fact that these changes are predicted even though local inter-species competition is not explicitly taken into account suggests that the Mukogodo forest is an area that may be extremely sensitive to shifts in climate over the coming decades.

Temporal changes in rainfall also suggest that *A. tortilis* will expand into the driest areas previously dominated by *A. drepanolobium* (Fig. 5c and d). The expansion of *A. tortilis* is due to the fact that this species has the lowest wilting point (ψ_{wilt}, Table 3) and stress values, which leads to higher relative fitness values when soil moisture is more frequently near the wilting point. As rainfall becomes more intense and less frequent, *A. tortilis*'s fractional water use actually increases, which leads to higher fitness values. Despite the expansion of *A. tortilis* on sandy soils, the model predicts the continued dominance of *A. drepanolobium* on the black cotton clay soils, which suggests that soil texture will continue to play a large role in explaining the distribution of species patterns in the basin.

In addition to temporal changes in rainfall patterns, we also investigated how separately varying basin mean growing season precipitation, μ_p, and growing season rainfall variability, σ_p^2, impacts species distribution patterns and fractional cover (Figs. 6 and 7). We find that the response of species patterns and fractional cover are similar when changing the mean and variance of rainfall, but that these responses are opposite in direction with respect to the frequency and intensity of rainfall. The model results suggest that the magnitude of shifts in species cover due to changing μ_p and σ_p^2 are comparable. With respect to the current climate in the basin, the model results indicate the distribution and relative abundance of the three species are stable with changes of less than 5% to either μ_p or σ_p^2 (Fig. 7a and b). The model predicts larger changes in the distribution of species and fractional cover as the basin rainfall regime moves towards the extremes of changes in μ_p or σ_p^2. Moving away from the current rainfall towards either decreasing μ_p or decreasing σ_p^2, we found *A. tortilis*'s fractional cover slightly increases while *A. xanthophloea*'s fractional cover decreases (Fig. 6a, b, e, and f). We find the opposite patterns when moving away from the current rainfall towards either increasing μ_p or increasing σ_p^2 (Fig. 6c, d, g, and h). Over most parts of the simulation range, *A. drepanolobium* remains the dominant species in the basin. The changing relative distribution of *A. xanthophloea*

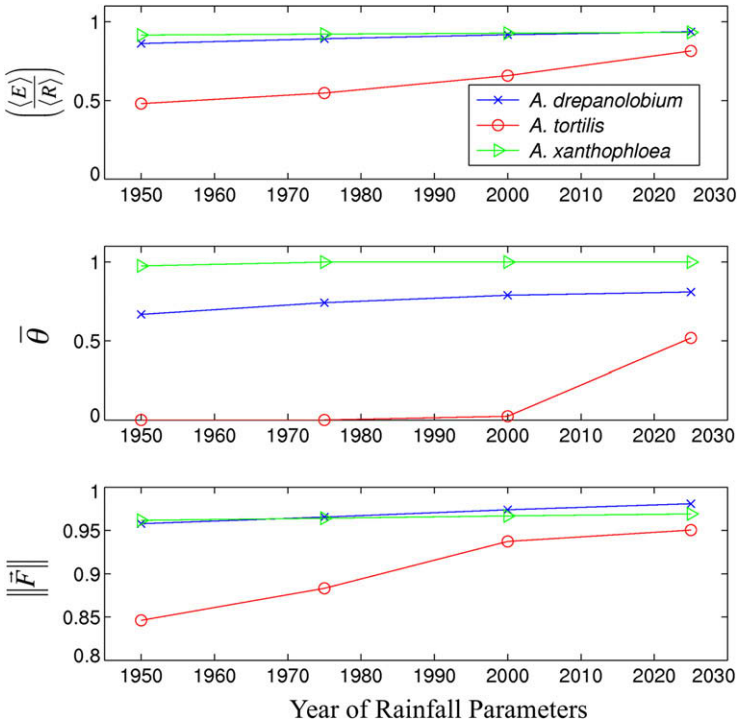

Fig. 9. Model results for fractional evapotranspiration from available rainfall, dynamic water stress, and the magnitude of the fitness vector for the Mukogodo forest using observed linear trends in time for rainfall parameters. The results illustrate the increase in fractional evapotranspiration of available rainfall for *A. drepanolobium* and *A. tortilis* while *A. xanthophloea* remains relatively static. The increases in water use result in *A. drepanolobium* having a higher fitness than *A. xanthophloea* using the rainfall parameters from 2025, which indicates that this area may sensitive to observed shifts in rainfall.

and *A. tortilis* is due to the separate effects of changing rainfall on evapotranspiration and stress. As either μ_p decreases (with σ_p^2 held constant) or σ_p^2 increases (with μ_p held constant), storms become more intense and less frequent (Fig. 7c). Initially, these larger storm depths lead to deeper infiltration fronts, but the reduction in storm frequency still causes average soil moisture values over the growing season to be lower. Overall, these changes lead to higher fractional water use relative to total rainfall, but the reduction in soil moisture causes greater stress values. Because water use is weighted more heavily to total fitness, when either the average rainfall declines or the variance in growing season rainfall increases, species that are able to maximize water use will have higher fitness. This is demonstrated by the example of *A. tortilis*, using a location from the northern portion of the basin, where rainfall is low and *A. tortilis* is currently dominant because stress values of *A. xanthophloea* and *A. drepanolobium* are high (Fig. 10). Although decreases in mean growing season rainfall lead to a dramatic increases in *A. tortilis*'s water stress, overall it still is able to maintain a lower stress value than the other two species, which exhibit chronic water stress across all values of average rainfall (Fig. 10b). Therefore, the modest increase in relative water use by *A. tortilis* (Fig. 10a) when rainfall is lowered is sufficient to allow it to remain the dominant species under the lower mean total rainfall scenarios (Fig. 10c). However, when μ_p is increased (with σ_p^2 held constant), despite further reductions in water stress (Fig. 10b), lower relative water use (Fig. 10a) causes *A. tortilis* to be replaced by species capable of using greater amounts of soil moisture (*A. drepanilobium*).

Although decreasing μ_p at constant σ_p^2 is similar to increasing σ_p^2 at constant μ_p (they both involve larger, less frequent storms, Fig. 7c), higher variances in rainfall at this northern, more arid location do not lead to higher fitness values of *A. tortilis* (Fig. 10f). This is because *A. tortilis* stress increases (Fig. 10e) without any substantial increase in water use (Fig. 10d). The reason for this increase in stress despite the lack of changes in total rainfall is the greater duration between storms, which leads to longer, more frequent periods of low soil moisture. *A. drepanolobium* is already near fully stressed conditions (Fig. 10e), so shifts in rainfall have very little impact on its total stress. The increases in stress for *A. tortilis* and stable stress values for *A. drepanolobium* lead to higher fitness values for *A. drepanolobium* in this part of the basin (Fig. 10f).

The impact of changes in rainfall variability and changes in total rainfall on dryland woody vegetation patterns and relative abundance are comparable, but inversely related. Along the gradient of increasing variance with a constant mean growing season rainfall, the model predicts the expansion of the riparian species *A. xanthophloea*, which takes advantage of its higher transpiration rates. However, as the rainfall regime approaches the lower extremes of decreasing rainfall variability, the range of the most drought tolerant species *A. tortilis* expands. The expansion and contraction of riparian and drought tolerant species is a major concern of many drylands around the world [1,4,11,31]. For example, in much of the southwest United States woody riverine species have been replaced by more drought tolerant species such as *Tamarix ramosissima* [4,11]. These changes have occurred over a period during which observations of rainfall have shown a shift towards more frequent, less intense storms [21,28]. Our results therefore provide a theoretical context for prior studies that have linked expansion of drought-tolerant invasive in the US desert southwest to altered rainfall climatology [20,56,67,78]. As an example, our model predicts both the expansion and the contraction of the drought tolerant species, *A. tortilis*, in the study basin depending on the nature of changing rainfall scenarios (Fig. 10).

4.3. Model limitations and future directions

Our model captured 79% of the available 90% of information derived from the neural network results (Table 5) [26,37]. While

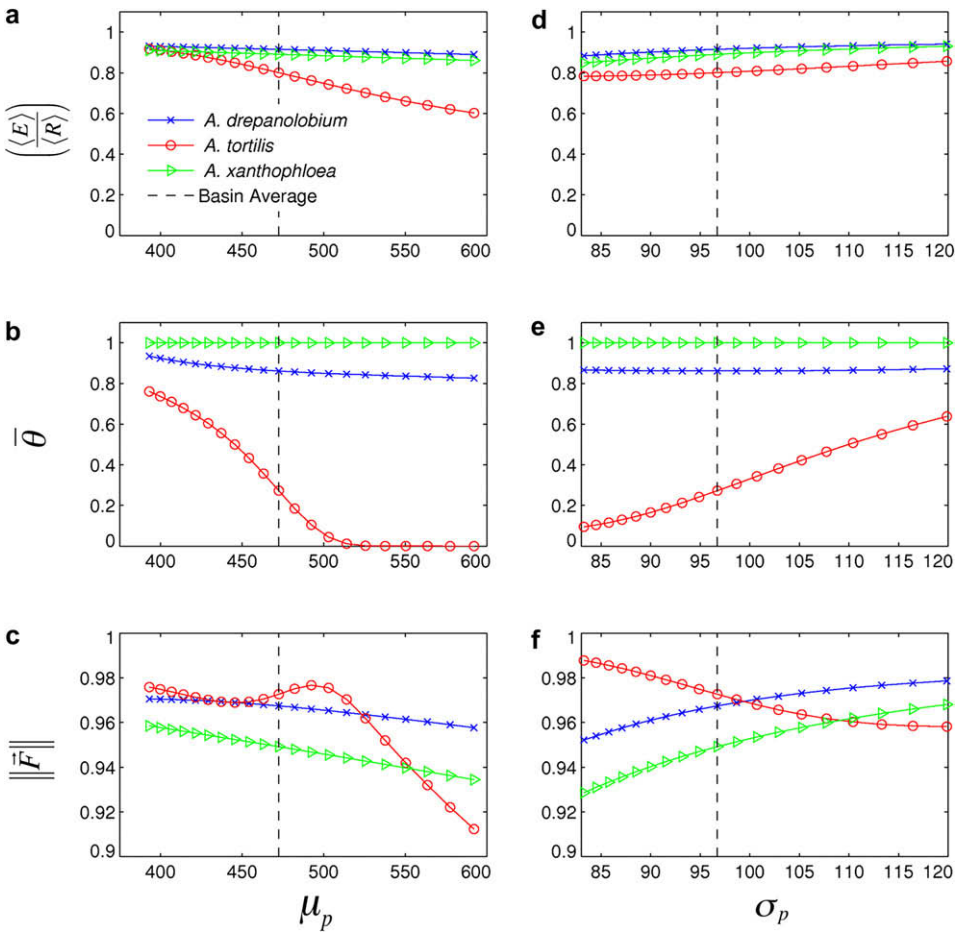

Fig. 10. Model results for fractional evapotranspiration from available rainfall (a and d), dynamic water stress (b and e), and magnitude of the fitness vector (c and f) for one point in the northern portion of the basin, which is currently dominated by the drought tolerant species *A. tortilis*. The sensitivity of the system is tested by varying μ_p, while holding σ_p^2 constant, and by varying σ_p^2 while holding constant μ_p. The results illustrate that the relative changes in water use and stress of the three species lead to different species having the highest fitness value using different rainfall scenarios.

it is not possible to determine the exact reasons for the difference in accuracy between our fitness vector approach and the neural network model, we believe the most likely factor reducing the fitness vector accuracy is the fact that, once-established, species may be able to remain in a location regardless of their fitness relative to competitors. Other factors that could possibly be necessary to improve our model's performance include the impact of local competition and facilitation for water and light, disturbance (including herbivory and fire), nutrient use and intra-species variation. Furthermore, despite the fact that the neural network performed better, it has no physical interpretation and would be extremely difficult to use with a novel dataset [6]. The basis of the fitness vector framework is the mass balance of water, which allows for physical interpretations and characterization of the mechanisms underlying model predictions.

Given the ability of the framework to capture the woody species distribution, we believe it is an useful tool for addressing climate change in dryland ecosystems on a regional scale for the following reasons: (1) the model performs well in the heterogeneous Upper Ewaso Ng'iro river basin, where vegetation patterns are influenced by complex interactions of soil, climate, and topography, (2) the model utilizes a parsimonious, analytical model of stochastic water balance, (3) the model is sensitive to subtle shifts in both rainfall frequency and rainfall depth, and (4) the model contains a relatively low number of parameters compared to other regional species models [44].

In our view, the largest challenge that remains for the modeling framework is the availability of species-specific parameters and improved understanding of how those parameters vary in space and time within species. For the current work, we were able to identify and parameterize (with some level of certainty, see Table 3) three woody species in the basin, which are dominant in approximately 60% of the area. The plant ecohydrological characteristics of the remaining woody species and functional groups in the basin are largely unknown. In particular, the influence of climate change on plant transpiration, growth, reproduction, and survival needs to be addressed with empirical studies. For example, methods capable of characterizing plant water use and plant water stress at landscape scales would help constrain parameters. Another plant parameter in need of further investigation is rooting depth, especially along natural gradients where a particular species is dominant.

Recent studies have revealed important feedbacks can occur between plant species and higher trophic levels that impact trajectories of community response to shifts in rainfall [71]. In addition, long-term experiments have shown that the removal of large herbivores has significant impacts on the growth, reproduction, and survival of woody vegetation [23]. Empirical studies have also demonstrated that herbivore and plant interactions can cascade into other communities of organisms [52]. Based on these studies, we expect that future range shifts in woody species will be greatly affected by complex ecological interactions influenced by

herbivores that are characteristic of many African savannas. However, because plant–herbivore interactions are almost always species-specific, this effort represents a necessary step towards achieving more robust estimates of whole-ecosystem response to climate change. Furthermore, although our model predicts plant species response to only changes in rainfall, we have demonstrated the high degree of sensitivity that species distributions can exhibit to subtle shifts in climate, even across a group of highly-related, drought-adapted trees. Therefore, we believe our approach demonstrates the utility of species-level predictions as a means for developing a more integrated assessment of coupled community-level interactions when attempting to describe transient vegetation responses to climate change.

Acknowledgements

We thank the Princeton Institute of International and Regional Studies (PIIRS), Princeton Environmental Institute (PEI), Princeton Grand Challenges, and the Mpala Research Conservation Center for their support. K.K.C. also acknowledges the support of the NSF through Grants DEB-742933 and EAR-0847368. The authors would also like to thank Natural Resource Monitoring, Modeling and Management Project (NRM3) of Nanyuki, Kenya for providing the extensive climatic dataset for the study. A special thank you to Dr. Elizabeth King and three anonymous reviewers for their helpful suggestions.

Appendix A. Water balance model

A full description of the water balance model is provided elsewhere [35,36,55,58,59]. Here we present a general overview of the governing equations and solutions. The governing equation for the root-zone daily soil moisture is a stochastic ordinary differential equation

$$nZ_r\frac{ds(t)}{dt}=\varphi[s(t),t]-\chi[s(t)], \tag{A.1}$$

where n is the soil porosity (dimensionless), Z_r is the effective root depth (mm), $s(t)$ is the root-zone relative soil moisture, $\varphi[s(t),t]$ is the marked Poisson process of inputs into the soil from rainfall events (mm day^{-1}), and $\chi[s(t)]$ is the soil moisture losses from the root zone (mm day^{-1}). The input function $\varphi[s(t),t]$, is given by

$$\varphi[s(t),t]=R(t)-I(t)-Q[s(t),t], \tag{A.2}$$

where $R(t)$ is rainfall, $I(t)$ is canopy interception, and $Q[s(t),t]$ is runoff (all in mm day^{-1}). The loss function $\chi[s(t)]$, is defined

$$\chi[s(t)]=E[s(t)]+L[s(t)], \tag{A.3}$$

where $E[s(t)]$ is evapotranspiration and $L[s(t)]$ is leakage through the root zone (both in mm day^{-1}). The rainfall $R(t)$ (mm day^{-1}), is represented as a marked Poisson process, as described in manuscript text (Section 2.3). The model incorporates interception from the canopy by assuming a threshold storm depth Δ (mm), below which no water penetrates the canopy. The frequency of infiltration events is given by

$$\lambda'=\lambda e^{\frac{-\Delta}{\alpha}}. \tag{A.4}$$

where λ' is the censored mean arrival rate of storms (day^{-1}). When the depth exceeds the storage capacity of the soil, saturation runoff, $Q[s(t),t]$, is produced. Leakage, $L[s(t)]$, represents the vertical percolation with unit gradient and is represented by

$$L[s(t)]=\frac{K_s}{e^{\beta(1-s_{fc})}-1}[e^{\beta(s-s_{fc})}-1], \tag{A.5}$$

where K_s is the vertical saturated hydraulic conductivity (mm day^{-1}), s_{fc} is the soil field capacity (dimensionless), and $\beta=2b+4$, where b is the pore size distribution index (dimensionless) [10]. The daily evapotranspiration, $E[s(t)]$, is modeled using a piecewise function of soil moisture [36]. The loss function is zero at and below the residual saturation of soils, s_h (dimensionless). Between the soil hygroscopic point, s_h, and the plant wilting point, s_w (saturation where permanent damage is incurred by the plant, dimensionless), the loss function varies linearly rising to a maximum soil evaporation rate, E_w (mm day^{-1}). Between the plant wilting point, s_w, and the saturation where the stomata begin to close, s_* (dimensionless), the loss function linearly increases from E_w to the maximum daily evaporation rate, E_{max} (mm day^{-1}). Between s_*, and s_{fc}, the loss function is equal to E_{max}. Finally, above field capacity, s_{fc}, the loss function is equal to the sum of E_{max} and leakage, $L[s(t)]$. The values of $s_{fc},\ s_h,\ s_w,\ s_*$ are related to the matric potential of the soil, ψ (MPa), which are defined by soil type, using the soil-retention curves [10]. Values of E_{max} are estimated for each species using the Penman–Monteith formulation of potential evapotranspiration, scaled by maximum canopy conductance [7] (Section 2.3).

The climate, soil, and vegetation parameters are used to determine the steady-state probability distribution of soil moisture over the growing season (see [36] for the full model solution). The normalized degree of stress experienced is given by

$$\zeta(t)=\begin{cases}1 & \text{if } s(t)\leqslant s_w,\\ \left[\frac{s_*-s(t)}{s_*-s_w}\right]^q & \text{if } s_w\leqslant(t)\leqslant s_*,\\ 0 & \text{if } s(t)>s_*,\end{cases} \tag{A.6}$$

where q accounts for the nonlinear relationship between plant stress and the soil moisture and assumed equal to two [59]. The daily static water stress does not account for the seasonal distribution of the frequency and duration of stress periods below the critical value of s^*. The mean duration of an excursion, $\overline{T}_\xi(\zeta)$ (day), below any given soil moisture threshold, ξ, is given by [55]

$$\overline{T}_\xi(\zeta)=\frac{P(\xi)}{\chi(\xi)p(\xi)}, \tag{A.7}$$

where $P(\xi)$ is the value of the cumulative soil moisture distribution function at the threshold value ξ, $p(\xi)$ is the value of the probability soil moisture distribution function at the threshold value ξ, and $\chi(\xi)$ is the value of the loss function at the threshold value, ξ. The mean number of excursions, $\bar{n}_\xi$ (dimensionless), below a given threshold value ξ, is given by [55]

$$\bar{n}_\xi=Ce^{-\gamma s_\xi}T_{seas}, \tag{A.8}$$

where C is an integration constant, $\gamma=\frac{nZ_r}{\alpha}$ (dimensionless), and T_{seas} (day) is the number of days in the growing season. The values of $\overline{T}_{s*}$ and $\bar{n}_{s*}$ provide a way of characterizing stress values during the growing season. The average dynamic water stress over the growing season is defined as follows [55]

$$\bar{\theta}=\begin{cases}\left(\frac{\bar{\zeta}'\overline{T}_{s*}}{kT_{seas}}\right)^{\bar{n}_{s*}} & \bar{\zeta}'*\overline{T}_{s*}<kT_{seas},\\ 1 & \text{otherwise},\end{cases} \tag{A.9}$$

where $\bar{\theta}$ is a normalized value between 0 and 1 with 1 representing the most stressed conditions, $\bar{\zeta}'$ is the mean "static" stress during an excursion below s_*, and k is the resistance of the plant to water stress and assumed equal to 0.5 for all species [7,55,59]. Each component of the seasonal water balance is defined as follows [59]

$$\begin{cases} \langle \phi \rangle = \langle R \rangle - \langle I \rangle - \langle Q \rangle, \\ \langle R \rangle = \alpha\lambda, \\ \langle I \rangle = \alpha(\lambda - \lambda'), \\ \langle Q \rangle = \alpha\left(\eta + \frac{K_s}{nZ_r}\right)p(1), \\ \langle E \rangle = \alpha\lambda' P(s_*) - \alpha\eta p(s_*) + E_{max}[1 - P(s_*)], \\ \langle L \rangle = \alpha\left[\lambda' - \lambda' P(s_{fc}) - \left(\eta + \frac{K_s}{nZ_r}\right)p(1) + \eta p(s_{fc})\right] - E_{max}[1 - P(s_{fc})], \end{cases} \tag{A.10}$$

where the brackets $\langle\ \rangle$, represent steady-state averages taken over the growing season.

References

[1] Archer S, Schimel DS, Holland EA. Mechanisms of shrubland expansion – land-use, climate or Co-2. Climatic Change 1995;29(1):91–9.
[2] Baudena M et al. Vegetation response to rainfall intermittency in drylands: results from a simple ecohydrological box model. Adv Water Resour 2007;30(5):1320–8.
[3] Box EO. Predicting physiognomic vegetation types with climate variables. Vegetatio 1981;45(2):127–39.
[4] Busch DE, Smith SD. Mechanisms associated with decline of woody species in riparian ecosystems of the southwestern US. Ecol Monogr 1995;65(3):347–70.
[5] Campbell GS, Norman JM. A introduction to environmental biophysics. 2nd ed. New York: Springer; 1998.
[6] Caudill M. Neural network training tips and techniques. AI Expert 1996;6(1):56–61.
[7] Caylor KK, Manfreda S, Rodriguez-Iturbe I. On the coupled geomorphological and ecohydrological organization of river basins. Adv Water Resour 2005;28(1):69–86.
[8] Caylor KK, Scanlon TM, Rodriguez-Iturbe I. Ecohydrological optimization of pattern and processes in water-limited ecosystems: a trade-off-based hypothesis. Water Resour Res 2009;45:15.
[9] Chen XY et al. Observations and stochastic modeling of soil moisture control on evapotranspiration in a Californian oak savanna. Water Resour Res 2008;44(8).
[10] Clapp RB, Hornberger GM. Empirical equations for some soil hydraulic-properties. Water Resour Res 1978;14(4):601–4.
[11] Cleverly JR et al. Invasive capacity of *Tamarix ramosissima* in a Mojave Desert floodplain: the role of drought. Oecologia 1997;111(1):12–8.
[12] D'Odorico P, Porporato A. Dryland ecohydrology. 1st ed. Dordrecht, The Netherlands: Springer; 2006.
[13] Darlington J. Termite nest structure and impact on the soil at the radar site, Embakasi, Kenya (Isoptera: Termitidae). Sociobiology 2005;45(3):521–42.
[14] Dharani N. Field guide to Acacias of East Africa. Cape Town: Struik Publishers; 2006.
[15] Dingman LS. Physical hydrology. 2nd ed. Upper Saddle River: Prentice-Hall Inc.; 2002.
[16] Eagleson PS. Ecological optimality in water-limited natural soil–vegetation systems. 1. Theory and hypothesis. Water Resour Res 1982;18(2):325–40.
[17] Eagleson PS, Tellers TE. Ecological optimality in water-limited natural soil–vegetation systems. 2. Tests and applications. Water Resour Res 1982;18(2):341–54.
[18] Elton C. Animal ecology. With an introduction by Julian S. Huxley, Animal ecology. The Macmillan Co.; 1927. p. xx + 207 p. 8 pl., 18 fig.
[19] Fleiss JL, Cohen J. Equivalence of weighted kappa and intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973;33(3):613–9.
[20] Friedman JM et al. Dominance of non-native riparian trees in western USA. Biol Invasions 2005;7(4):747–51.
[21] Fu Q et al. Enhanced mid-latitude tropospheric warming in satellite measurements. Science 2006;312(5777):1179.
[22] Gebrehiwot K et al. The use of plant water relations to characterize tree species and sites in the drylands of northern Ethiopia. J Arid Environ 2005;60(4):581–92.
[23] Goheen JR et al. Consequences of herbivory by native ungulates for the reproduction of a savanna tree. J Ecol 2007;95(1):129–38.
[24] Grinnell J. The niche-relationships of the California Thrasher. Auk Cambridge Mass 1917;34:427–33.
[25] Guenther A et al. Biogenic hydrocarbon emissions and landcover/climate change in a subtropical savanna. Phys Chem Earth Pt B: Hydrol Oceans Atmos 1999;24(6):659–67.
[26] Guisan A, Zimmermann NE. Predictive habitat distribution models in ecology. Ecol Model 2000;135(2–3):147–86.
[27] Hardin G. Competitive exclusion principle. Science 1960;131(3409):1292–7.
[28] Hu Q, Feng S. Decadal variation of the southwest US summer monsoon circulation and rainfall in a regional model. J Climate 2007;20(18):4702–16.
[29] Hulme M et al. African climate change: 1900–2100. Climate Res 2001;17(2):145–68.
[30] Hutchinson GE. Population studies – animal ecology and demography – concluding remarks. Cold Spring Harbor Symp Quant Biol 1957;22:415–27.
[31] Huxman TE et al. Ecohydrological implications of woody plant encroachment. Ecology 2005;86(2):308–19.
[32] IPCC. Working group I report, The physical science basis. I.P.o.C. Change, Geneva; 2007.
[33] Katul G, Porporato A, Oren R. Stochastic dynamics of plant–water interactions. Annu Rev Ecol Evol Syst 2007;38:767–91.
[34] Kearney M. Habitat environment and niche: what are we modelling? Oikos 2006;115(1):186–91.
[35] Laio F et al. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress – IV. Discussion of real cases. Adv Water Resour 2001;24(7):745–62.
[36] Laio F et al. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress – II. Probabilistic soil moisture dynamics. Adv Water Resour 2001;24(7):707–23.
[37] Lek S et al. Application of neural networks to modelling nonlinear relationships in ecology. Ecol Model 1996;90(1):39–52.
[38] Ludwig F et al. Hydraulic lift in *Acacia tortilis* trees on an East African savanna. Oecologia 2003;134(3):293–300.
[39] Mackay DS et al. Physiological tradeoffs in the parameterization of a model of canopy transpiration. Adv Water Resour 2003;26(2):179–94.
[40] Maidment D. Handbook of hydrology. 1st ed. New York: McGraw-Hill Professional; 1993.
[41] Meehl GA, et al. In: Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, Cambridge, United Kingdom and New York, NY, USA; 2007.
[42] Monteith JL. Photosynthesis and transpiration of crops. Exp Agric 1966;2(1):1.
[43] Morin X, Augspurger C, Chuine I. Process-based modeling of species' distributions: what limits temperate tree species' range boundaries? Ecology 2007;88(9):2280–91.
[44] Morin X, Lechowicz MJ. Contemporary perspectives on the niche that can improve models of species range shifts under climate change. Biol Lett 2008;4(5):573–6.
[45] Nicholson SE. The nature of rainfall variability over Africa on time scales of decades to millenia. Global Planet Change 2000;26(1–3):137–58.
[46] Nicholson SE, Selato JC. The influence of La Nina on African rainfall. Int J Climatol 2000;20(14):1761–76.
[47] Nicholson SE, Yin XG. Rainfall conditions in equatorial East Africa during the nineteenth century as inferred from the record of Lake Victoria. Climatic Change 2001;48(2–3):387–98.
[48] Nizinski J, Morand D, Fournier C. Actual evapotranspiration of a thorn scrub with *Acacia tortilis* and *Balanites aegyptiaca* (North Senegal). Agric For Meteorol 1994;72(1–2):93–111.
[49] Nordbotten JM. Stability analysis of probabilistic soil moisture dynamics. Adv Water Resour 2008;31(2):418–23.
[50] Norris JR, Jackson ST, Betancourt JL. Classification tree and minimum-volume ellipsoid analyses of the distribution of ponderosa pine in the western USA. J Biogeogr 2006;33(2):342–60.
[51] Otieno DO et al. Responses of *Acacia tortilis* and *Acacia xanthophloea* to seasonal changes in soil water availability in the savanna region of Kenya. J Arid Environ 2005;62(3):377–400.
[52] Palmer TM et al. Breakdown of an ant–plant mutualism follows the loss of large herbivores from an African Savanna. Science 2008;319(5860):192–5.
[53] Pataki DE, Oren R. Species differences in stomatal control of water loss at the canopy scale in a mature bottomland deciduous forest. Adv Water Resour 2003;26(12):1267–78.
[54] Pearson RG, Dawson TP. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? Glob Ecol Biogeogr 2003;12(5):361–71.
[55] Porporato A et al. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress – III. Vegetation water stress. Adv Water Resour 2001;24(7):725–44.
[56] Reynolds JF et al. Modifying the 'pulse-reserve' paradigm for deserts of North America: precipitation pulses, soil water, and plant responses. Oecologia 2004;141(2):194–210.
[57] Ridolfi L et al. Stochastic soil moisture dynamics along a hillslope. J Hydrol 2003;272(1–4):264–75.
[58] Rodriguez-Iturbe I et al. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress – I. Scope and general outline. Adv Water Resour 2001;24(7):695–705.
[59] Rodriguez-Iturbe I, Porporato A. Ecohydrology of water-controlled ecosystems. New York: Cambridge University Press; 2004.
[60] Rubenstein D. Personal communication; 2006.
[61] Salvucci GD. Estimating the moisture dependence of root zone water toss using conditionally averaged precipitation. Water Resour Res 2001;37(5):1357–65.
[62] Sankaran M et al. Determinants of woody cover in African savannas. Nature 2005;438(7069):846–9.
[63] Scanlon TM et al. Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem. Remote Sens Environ 2002;82(2–3):376–88.
[64] Scanlon TM, Albertson JD. Inferred controls on tree/grass composition in a savanna ecosystem: combining 16-year normalized difference vegetation index data with a dynamic soil moisture model. Water Resour Res 2003;39(8).
[65] Scanlon TM, Albertson JD. Canopy scale measurements of CO_2 and water vapor exchange along a precipitation gradient in southern Africa. Global Change Biol 2004;10(3):329–41.

[66] Scholes RJ, Archer SR. Tree–grass interactions in savannas. Annu Rev Ecol Syst 1997;28:517–44.
[67] Schwinning S et al. Thresholds, memory, and seasonality: understanding pulse dynamics in arid/semi-arid ecosystems. Oecologia 2004;141(2):191–3.
[68] Schymanski SJ et al. An optimality-based model of the coupled soil moisture and root dynamics. Hydrol Earth Syst Sci 2008;12(3):913–32.
[69] Seager R et al. Model projections of an imminent transition to a more arid climate in southwestern North America. Science 2007;316(5828):1181–4.
[70] Smit GN, Rethman NFG. The influence of tree thinning on the soil water in a semi-arid savanna of southern Africa. J Arid Environ 2000;44(1):41–59.
[71] Suttle KB, Thomsen MA, Power ME. Species interactions reverse grassland responses to changing climate. Science 2007;315(5812):640–2.
[72] Tietjen B, Jeltsch F. Semi-arid grazing systems and climate change: a survey of present modelling potential and future needs. J Appl Ecol 2007;44(2):425–34.
[73] Tromp-van Meerveld HJ, McDonnell JJ. On the interrelations between topography soil depth, soil moisture, transpiration rates and species distribution at the hillslope scale. Adv Water Resour 2006;29(2):293–310.
[74] Williams CA, Albertson JD. Soil moisture controls on canopy-scale water and carbon fluxes in an African savanna. Water Resour Res 2004;40(9).
[75] Williams CA, Albertson JD, Porporato A. Can water-limitation alone explain woody-herbaceous vegetation patterns along a large-scale rainfall gradient in southern Africa. Ecol Soc Am Annu Meet Abstr 2004;89:545.
[76] Williams CA, Albertson JD. Contrasting short- and long-timescale effects of vegetation dynamics on water and carbon fluxes in water-limited ecosystems. Water Resour Res 2005;41(6).
[77] Xue Y. Interactions and feedbacks between climate and dryland vegetations. In: D'Odorico P, Porporato A, editors. Dryland ecohydrology. Dordrecht, The Netherlands: Springer; 2006. p. 85–105.
[78] Zavaleta ES, Royval JL. Climate change and the susceptibility of US ecosystems to biological invasions: two cases of expected range expansion. In: Schneider SH, Root TL, editors. Wildlife responses to climate change: North American case studies. Washington, DC: Island Press; 2002. p. 277–341.