Ecosystems (2011) 14: 183–197
DOI: 10.1007/s10021-010-9403-9

ECOSYSTEMS

A Benchmark Test for Ecohydrological Models of Interannual Variability of NDVI in Semi-arid Tropical Grasslands

Philippe Choler,[1,2]* William Sea,[1] and Ray Leuning[1]

[1]CSIRO Marine and Atmospheric Research, PO Box 3023, Canberra, Australian Capital Territory 2601, Australia; [2]Laboratoire d'Ecologie Alpine, UMR CNRS-UJF 5553, Université J. Fourier—Grenoble I, BP53, 38041 Grenoble, France

Abstract

Pulses of aboveground net primary productivity (ANPP) in response to discrete precipitation events are an integral feature of ecosystem functioning in arid and semi-arid lands. Yet, the usefulness of nonlinear, ecohydrological pulse response functions to predict regional-scale patterns of annual ANPP at decadal scales remains unclear. Here, we assessed how different pulse response (PR) models compete with simple linear statistical models to capture variability in yearly integrated values of Normalized Difference Vegetation Index ($NDVI_{int}$), a remotely sensed proxy of annual ANPP. We examined 24-year-long time series of $NDVI_{int}$ calculated from Advanced Very High Resolution Radiometer (AVHRR) NDVI for 350,000 km^2 of tropical grasslands in northern Australia. Based on goodness-of-fit statistics, PR models clearly outperformed statistical models when parameters were optimized for each site but all models showed the same error magnitude when all sites were combined in ensemble simulations or when the models were evaluated outside the calibration period. PR models were less biased and their performance did not deteriorate in the driest areas compared to linear models. Increasing the complexity of PR models to provide a better representation of soil water balance and its feedback with plant growth did not improve model performance in ensemble simulations. When error magnitude, bias, and sensitivity to parameter uncertainty were all considered, we concluded that a low-dimensional PR model was the most robust to capture $NDVI_{int}$ variability. This study shows the potential of long time series of AVHRR NDVI to benchmark process-oriented models of interannual variability of $NDVI_{int}$ in water-controlled ecosystems. This opens new avenues to examine at the global scale and over several decades the causal relationships between climate and leaf dynamics in the grassland biome.

Key words: C_4 grasslands; ecohydrology; phenology; precipitation pulses; semi-arid ecosystems.

Introduction

A cornerstone of arid land ecology is that annual primary productivity results from a sequence of pulsed responses to discrete rainfall events (Noy-Meir 1973). Following each pulse, part of the production is diverted into vegetative or reproductive

Received 14 May 2010; accepted 1 November 2010;
published online 7 December 2010

Author Contributions: P.C. analyzed data and wrote the paper; P.C, B.S., and R. L. designed the study.

*Corresponding author; e-mail: philippe.choler@ujf-grenoble.fr

storage compartments to ensure survival of organisms during unfavorable, dry periods. This is the well-known Westoby–Bridges 'Pulse-reserve' hypothesis published by Noy-Meir (1973). Recently, Ogle and Reynolds 2004) built on this hypothesis and proposed a low-dimensional threshold-delay model to capture plant response to discrete rainfall events. Essentially, they proposed a pulse response (PR) model to describe vegetation dynamics. Growth is triggered, when after some time-delay, the main climate driver (precipitation or soil water content) exceeds a lower threshold. The growth rate depends on the current vegetation state and water availability and the rate coefficient for plant decay is either a constant or a function of soil moisture content. These PR models were primarily intended to describe fine-scale temporal dynamics of plant growth (Reynolds and others 2004). To what extent this kind of model structure is also relevant to simulate annual aboveground net primary productivity (ANPP) and its variability at decadal scales and across large areas, remains unclear.

Time series of NDVI, the Normalized Difference Vegetation Index, have been used extensively as an indirect estimate of ANPP (Tucker and others 1983; Tucker and Sellers 1986; Prince and others 1995; Goetz and others 1999). These remote sensing-based estimates of ANPP have their origins in the work of Monteith (1977) who found strong relationships between ANPP of crops and annual integrals of absorbed photosynthetically active radiation (PAR). Significant positive relationships between the annual integrated values of NDVI, hereafter $NDVI_{int}$, and ANPP have also been observed in grasslands (Paruelo and others 1997). Furthermore, many regional and global studies have demonstrated significant statistical relationships between NDVI-derived ANPP and precipitation in water-controlled ecosystems (Farrar and others 1994; Potter and Brooks 1998; Richard and Poccard 1998; Lotsch and others 2003; Ji and Peters 2004). Thus, long time series of remotely sensed vegetation indices provide opportunities to track the response of seasonal and interannual variability of ANPP to climate (Pettorelli and others 2005).

Regression models using annually integrated values of precipitation and evapotranspiration as linear predictors of $NDVI_{int}$ have been used extensively (Paruelo and Lauenroth 1998; Ji and Peters 2004; Pineiro and others 2006; Propastin and Kappas 2008). These studies did not make explicit the linkages between rainfall pulses, soil water balance, and plant growth and as such are rather disconnected from more process-oriented eco-hydrogical modeling frameworks (Porporato and others 2002; Daly and Porporato 2005; Caylor and others 2009). As a result, we currently have strikingly different modeling approaches to address the same biological question.

Compared with the many regional and global studies that examined NDVI-climate relationships in water-controlled ecosystems, it is striking not to have a benchmark for competing models. The lack of thorough comparisons between different modeling approaches is a main caveat of the current literature in ecosystem phenology as it is the case more generally with land-surface models (Pitman and others 1999; Abramowitz 2005). The exercise is not trivial because (i) there are many different metrics to evaluate the performance and the robustness of competing models (Alewell and Manderscheid 1998; Willmott and Matsuura 2006; Liu and Gupta 2007; Gulden and others 2008), (ii) the acceptance criteria are rarely stated explicitly (Rykiel 1996), and (iii) the usefulness of a model strongly depends upon the temporal and spatial scales at which it is evaluated. For example, a high-dimensional process-oriented model may be a reasonable choice for site-specific application where information is readily available, but of limited use for larger scale applications if it is highly sensitive to parameter uncertainties or if unknown site variability is high. In this regard, a significant improvement would be to have a low-dimensional model that provides a good representation of the fundamental biological and physical processes and the possibility of using the same model across a wide range of conditions while maintaining acceptable predictive accuracy. To our knowledge, there has been no comparative study examining how linear versus PR models of $NDVI_{int}$ meet these rather contradictory requirements.

This paper addresses three questions: (i) How can we assimilate remotely sensed NDVI to calibrate and test models of ANPP using pulse response functions? (ii) What level of model complexity in the representation of water balance and plant growth dynamics do we need to include to get accurate and robust estimates of $NDVI_{int}$ at regional scales? and (iii) What are the criteria needed to benchmark these models of $NDVI_{int}$? We used NDVI measurements from AVHRR sensors aboard NOAA satellites because it is the only global dataset with temporal coverage of more than two decades. First, this is invaluable for tracking regional to continental long-term changes in vegetation cover and phenology (for example, Myneni and others 1997; Anyamba and Tucker 2005; Heumann and others 2007; Donohue and others 2009). Second, long time series of NDVI allow us to investigate the

response of vegetation to multiannual cycles of climate, such as those induced by the El Nino-Southern Oscillation (Hashimoto and others 2004). Finally, it enables robust evaluation of competing models by splitting observed data into multi-year calibration and validation periods.

Materials and Methods

Study Area

Our case study corresponds to 350,000 km^2 of tropical tussock grasslands located in the Northern Territory and Queensland (Australia) (Figure 1). The grasslands are dominated by perennial C_4 grasses, among which Mitchell grass (*Astrebla* spp.), feather-top wire grass (*Aristida* spp.) and Blue grass (*Dichanthium* spp.) are the commonest. These grasslands exhibit strong responses to precipitation events (Christie 1981; Hunter 1989; Choler and others 2010) making them ideal candidates to calibrate and test PR models. The dominant soil type is cracking clay. Eastern and northern sites exhibit deep cracking gray soils whereas the south-western sites in Queensland exhibit brown-red clay soils (McKenzie and Hook 1992; McKenzie and others 2000). A mean value of 450 mm for soil available water capacity (W_{cap}) was calculated from available data of saturated soil water content (W_{sat}) and soil depth (Z) (McKenzie and others 2000).

Datasets

AVHRR NDVI measurements were processed by the GIMMS group at NASA's Goddard Space Flight Center (Tucker and others 2005). Reflectance values in the red (ρ_r) and near infrared (ρ_{nir}) were used to calculate an 'observed' NDVI according to:

$$\mathrm{NDVI}_{obs} = (\rho_{nir} - \rho_r)/(\rho_{nir} + \rho_r) \qquad (1)$$

Data were 15-day composite values at 8 km spatial resolution. Pixels corresponding to tropical tussock grasslands were extracted using a mask from the digital map of Australia's Native Major Vegetation Groups (Australian Government, Department of the Environment, Water, Heritage and the Arts, http://www.environment.gov.au/erin/nvis/mvg/). Pixels that were directly adjacent to another type of vegetation, usually eucalypt open woodlands, were removed to avoid potential edge effects. The final dataset comprised 2,907 pixels. The analyzed period covers 24 years from July 1981 to June 2005 giving 576 observations for each pixel, that is, 24 years × 24 period values per year. Outliers, that

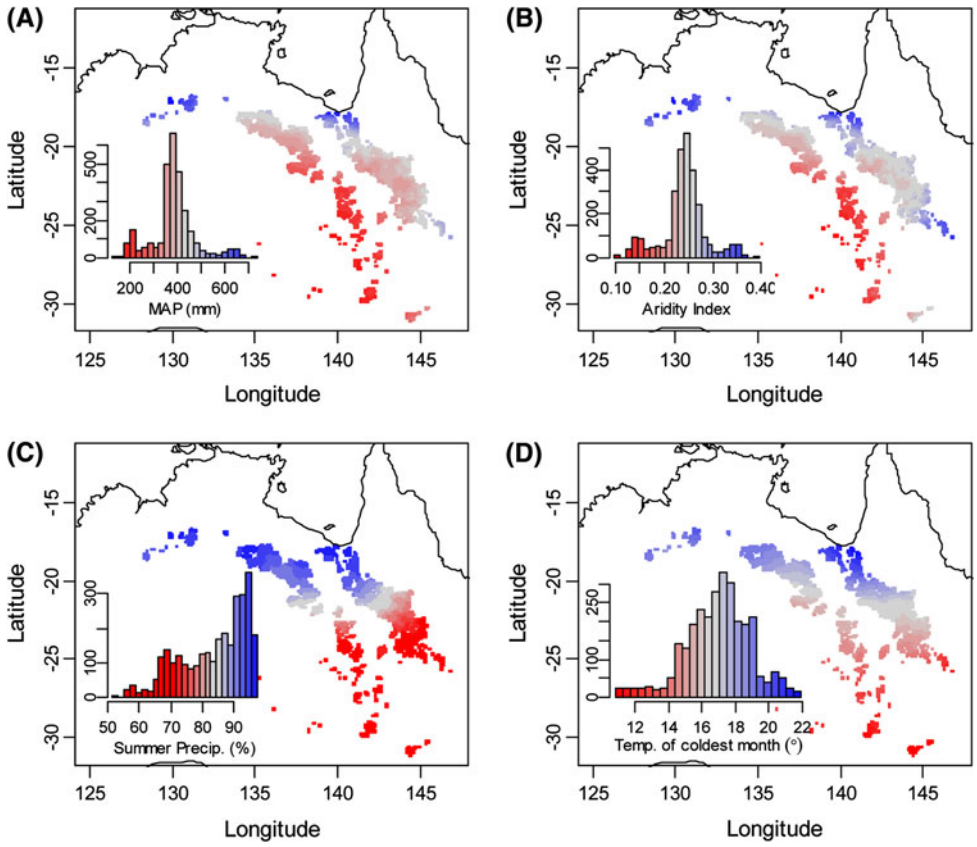

Figure 1. Geographic distribution of studied grasslands and maps of mean annual precipitation (MAP) (**A**), aridity index, the ratio between MAP and mean potential evapotranspiration (**B**), percentage of total precipitation during ''summer'' months (NDJFMA) (**C**), and mean temperature of the coldest month (**D**). *Inserts*: Frequency distribution of climate values are given for the 2,907 sites investigated. All data are averages over the period 1981–2004.

is, unusually high or low $NDVI_{obs}$, were discarded and missing values were interpolated using a second-order polynomial fitting method. An observation was considered as an outlier if its value was two times larger or smaller than the averaged value of the two preceding and the two following observations. To reduce noise in the edited signal $NDVI_{obs}$ at time t_i was replaced by the value of a quadratic polynomial fitted to $2n + 1$ points (Savitzky and Golay 1964), with n being the number of time steps preceding or following t_i. Here, we used a moving window of length $n = 2$. Small values of n allow keeping track of rapid changes in NDVI but at the expense of noise reduction efficiency (Hird and McDermid 2009). Using different values for n (data not shown), we found that a moving window of two periods was the best compromise for our data.

Following other studies (Carlson and others 1995; Seaquist and others 2003), a re-normalization of $NDVI_{obs}$ was done to attenuate various sources of remote sensing measurement error, among which are contrasting solar zenith angle, sensor degradation, and stratospheric aerosol. $NDVI_{nor}$ was calculated as follows:

$$NDVI_{nor} = (NDVI_{obs} - NDVI_0)/(NDVI_\infty - NDVI_0) \quad (2)$$

where $NDVI_0$ was the minimum $NDVI_{obs}$ corresponding to bare soil and dry vegetation and $NDVI_\infty$ was that for full vegetation cover. Averaged $NDVI_{obs}$ values during the dry season (July–August) were used to estimate site-specific $NDVI_0$ as soil properties may affect reflectances (Montandon and Small 2008). A value of 0.95 was assigned for $NDVI_\infty$, corresponding to the maximum recorded $NDVI_{obs}$ in the most productive grasslands in the northernmost area.

The integrated value of $NDVI_{nor}$, for year j was calculated according to:

$$NDVI_{int,j} = \sum_{t=1+T(j-1)}^{Tj} NDVI_{nor,t}, \quad j = 1, \ldots, N \quad (3)$$

where t is the discrete time in periods of 15 days, T is the number of 15-day period per year, and N is the number of years. The resulting $NDVI_{int}$ matrix comprised 24 years and 2,907 sites.

Matching AVHRR NDVI time periods, 15-day cumulative sums of precipitation (P), and potential evapotranspiration (E) were prepared from the 0.05° resolution SILO database (Australian Government, Bureau of Meteorology, http://www.bom.gov.au/silo/). E was calculated following Priestley and Taylor (1972), that is:

$$E = 1.26\, s\, R_n/(\lambda(s + \gamma)) \quad (4)$$

where R_n is the net radiation absorbed by vegetation and soil, λ is the latent heat of evaporation, s is the slope of water vapor saturation versus temperature curve, and γ is the psychrometric coefficient. Gridded climate data were produced by spatial interpolation of ground-based observations as described in Jeffrey and others (2001). For the sampled area, the nearest weather station or rain gauge was less than 30 km for 80% of the sites.

Modeling Approach

We distinguished two classes of models (Table 1): (i) regression models that simply relate $NDVI_{int}$ to independent variables defined by yearly integrated values of precipitation, P_{int}, and potential evapotranspiration, E_{int}, hereafter INT models, and (ii) PR models that capture the fine-scale temporal dynamics of NDVI.

The linear INT models take the following generic form

$$NDVI_{int} = \alpha_1 + \sum_{i=2}^{n} \alpha_i X_{int,i} + \varepsilon \quad (5)$$

where α_1 is the intercept, α_i's are the optimized linear coefficients corresponding to the independent variables, $X_{int,i}$, and ϵ is an error term normally distributed and with a mean of 0. Tested predictors are yearly integrated values of precipitation, P_{int}, in INT1 and P_{int} and yearly integrated values of potential evapotranspiration, E_{int} in INT2 (Table 1).

PR models were derived from the threshold-delay model proposed by Ogle and Reynolds (2004) for arid and semi-arid ecosystems. To evaluate this model in a remote-sensing perspective, we used the fractional vegetation cover, C, as the vegetation response variable following previous studies (Choudhury and others 1994; Myneni and others 1995; Carlson and Ripley 1997; Seaquist and others 2003), $NDVI_{int}$ for year j was calculated from C according to

$$NDVI_{int,j} = \sum_{t=1+T(j-1)}^{Tj} C_t^{0.5}, \quad j = 1, \ldots, N \quad (6)$$

where t is the discrete time in periods of 15 days, T is the number of 15-day periods per year, and N is the number of years (Table 1).

There are many different ways to model the water balance and the gain and loss of vegetation from an ecohydrological perspective. In this study, we distinguished three PR models of increasing

Table 1. Main Features and Equations of the Models Investigated

Model	Parameters	Equations
INT1	2	$\mathrm{NDVI}_{\mathrm{int}} = \alpha_1 + \alpha_2 P_{\mathrm{int}}$
INT2	3	$\mathrm{NDVI}_{\mathrm{int}} = \alpha_1 + \alpha_2 P_{\mathrm{int}} + \alpha_3 E_{\mathrm{int}}$
PR		$C_{t+1} = C_t + \underbrace{\beta_2 D_{t-\mathrm{lag}}(1 - C_t/C_{\max})}_{\text{gain}} - \underbrace{\delta\beta_3 C_t}_{\text{loss}}$ with (i) $C_{t+1} = \max[0, \min(C_{\max}, C_t)]$ (ii) $\delta \begin{cases} = 0 \text{ if } D_{t-\mathrm{lag}} > D_{t-\mathrm{lag}-1} \\ = 1 \text{ if } D_{t-\mathrm{lag}} \le D_{t-\mathrm{lag}-1} \end{cases}$ $\mathrm{NDVI}_{\mathrm{int,j}} = \sum_{t=1+T(j-1)}^{Tj} C_t^{0.5} \quad j = 1, \ldots, N$
PR1	3	$D_t = \max(0, P_t - \beta_1)$
PR2	4	$D_t = \max(0, W_t - \beta_1)$ with $W_{t+1} = W_t + P_t - \beta_4 \left[D_t/\left(W_{\mathrm{cap}} - \beta_1\right)\right]^2 E_t$ $W_{t+1} = \max\left[0, \min\left(W_{\mathrm{cap}}, W_t\right)\right]$
PR3	4	$D_t = \max(0, W_t - \beta_1)$ with $W_{t+1} = W_t + P_t - (1 - C_t)\left[D_t/\left(W_{\mathrm{cap}} - \beta_1\right)\right]^2 E_t - C_t \beta_4 D_t$ $W_{t+1} = \max\left[0, \min\left(W_{\mathrm{cap}}, W_t\right)\right]$

P precipitation (mm), E potential evapotranspiration (mm), W soil water content (mm), D driver of vegetation response in PR models, C fractional vegetation cover [0–1], W_{cap} soil water capacity (mm), C_{max} carrying capacity, t discrete time (in periods of 15 days), x_{int} yearly integrated value of x, N number of years (N = 24), T number of 15-day periods per year (T = 24).

complexity (Figure 2; Table 1). In general, the dynamics of C are governed by a non-linear gain function dependent on a driver D (precipitation or soil water availability) and a loss term, each multiplying the current value of C. The value of C varies in the interval [0, $C_{\max}$], where $C_{\max}$ is the carrying capacity whose value was estimated for each site as the maximum value of C for each time series. All PR models included a lag, which is the time, in number of 15-day periods, it takes for D to trigger a change in C.

In the PR1 model, the gain in C depends on a parameter β_2 (in $\mathrm{mm}^{-1}\ \mathrm{t}^{-1}$) corresponding to a precipitation use efficiency (PUE) and on effective precipitation, $P - \beta_1$ where β_1 (in mm) is a critical threshold. Models PR2 and PR3 include a soil water balance term (Figure 2; Table 1). In these models the gain in C depends on the effective soil water content, $W - \beta_1$, where β_1 (in mm) is a critical threshold and the sensitivity of plant growth to W is represented by the parameter β_2 (in $\mathrm{mm}^{-1}\ \mathrm{t}^{-1}$), corresponding to a soil water use efficiency (SWUE). The dynamics of W are modeled with a one-layer bucket model with precipitation as input, and actual evapotranspiration and run-off as outputs. W varies in the interval [0, W_{cap}]. All precipitation runs off when W reaches W_{cap}. In model PR2, we did not distinguish between soil evaporation and plant transpiration. Actual evapotranspiration is a function of E_t, relative soil water content $(W - \beta_1)/(W_{\mathrm{cap}} - \beta_1)$ and a single parameter β_4 (Table 1). In the PR3 model, we distinguished between evaporation from bare soil and plant transpiration. Bare soil evaporation is a function of E_t and relative soil water content $(W - \beta_1)/(W_{\mathrm{cap}} - \beta_1)$. Plant transpiration is modeled as a water-limited process dependent upon the effective amount of W and a parameter β_4 corresponding to a plant water extraction ability (Table 1). The partitioning between bare soil evaporation and plant transpiration is controlled by the fractional vegetation cover C. Therefore, the PR3 model includes a negative feedback of C on W (Figure 2).

The loss function of all PR models is represented by an exponential decay controlled by the parameter β_3 (in t^{-1}). We assumed that $\beta_3 = 0$ when

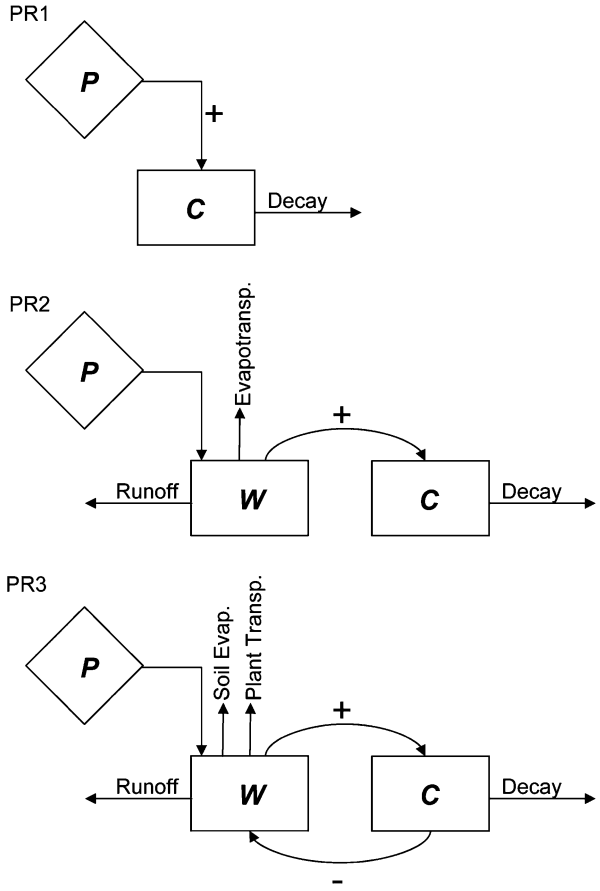

Figure 2. Relationships between precipitation (P), soil water content W, and fractional vegetation cover (C) in the three investigated pulse response models (see Table 1).

water availability increases and that β_3 is constant otherwise (Table 1, condition (ii)). This was chosen as a simple way to model an increase of leaf mortality under water shortage.

Parameter Optimization and Model Evaluation

Optimized parameter values were obtained by minimizing a cost function F. For site simulations, F was the coefficient of variation of the mean absolute error of site, hereafter CVMAE_i. For ensemble simulations, F was the average of CVMAE_i across all calibrating sites. Accordingly, we calculated F as follows:

$$F = \frac{1}{n}\sum_{i=1}^{n} \mathrm{CVMAE}_i = \frac{1}{Nn}\sum_{i=1}^{n}\sum_{j=1}^{N}\left|\mathrm{NDVI}_{\mathrm{int,p(i,j)}} - \mathrm{NDVI}_{\mathrm{int,o(i,j)}}\right| / \times \overline{\mathrm{NDVI}}_{\mathrm{int,o(i)}} \quad (7)$$

where N is the number of years, n is the number of calibrating sites, $\mathrm{NDVI}_{\mathrm{int,p}(i,j)}$ and $\mathrm{NDVI}_{\mathrm{int,o}(i,j)}$ are the predicted and observed values of $\mathrm{NDVI}_{\mathrm{int}}$ for site i and year j, respectively, and $\overline{\mathrm{NDVI}}_{\mathrm{int,o(i)}}$ is the mean $\mathrm{NDVI}_{\mathrm{int}}$ for site i. We chose to minimize CVMAE because this measure of error magnitude (i) was less sensitive to the distribution of errors compared to mean square error (MSE) or root mean square error (RMSE) (Willmott and Matsuura 2005), and (ii) ensured that each site had a similar influence on the final value of F as it compensated for differences in mean NDVI.

We used a split sample testing methodology in which parameters were calibrated for the period 1981–1990 and model validation was performed for the period 1991–2004. We switched calibration and validation periods and found no differences in model performance (data not shown). For site simulations, parameters were calibrated separately for 100 randomly sampled sites. In what is referred to as ensemble simulations, all sites were included as we searched for a common set of parameters. Here, parameter values were either the median of the distribution of site-optimized parameters or optimized values obtained in calibration tests that used subsets of 500 randomly selected sites. We found no changes in the rankings of model performance with one or the other method and present results obtained with the first one. An advantage of the first method is to reduce considerably the computation time. We used a spin-up phase of 6 months (from January to June 2001) to adjust all PR models to their forcing climate.

Parameter optimization was performed with the program 'rgenoud' written by Mebane and Sekhon (2009). The method combined a genetic algorithm for searching the parameter space and a derivative-based quasi-Newton algorithm developed by Byrd and others (1995) that allows setting of lower and upper bounds for a given parameter. It is particularly efficient for difficult convergence problems where multiple local minima exist. In its present version, it does not allow optimization of real and integer parameters in the same runs. Therefore, all calibrations were repeated with step-by-step increasing values of the lag. A lag value of 30 days, that is, 2 periods of the GIMMS data gave lowest values of F in more than 80% of runs. Here, we present all results with this common lag value.

To evaluate model performance further, we partitioned the mean squared error (MSE) into

systematic and unsystematic components as described by Willmott (1982).

$$\mathrm{MSE}_i = \mathrm{MSE}_{s,i} + \mathrm{MSE}_{u,i}$$

with

$$\mathrm{MSE}_{s,i} = \frac{1}{N}\sum_{j=1}^{N}\left(\overline{\overline{\mathrm{NDVI}}}_{\mathrm{int,p(i,j)}} - \mathrm{NDVI}_{\mathrm{int,o(i,j)}}\right)^2 / \overline{\mathrm{NDVI}}_{\mathrm{int,o(i)}} \quad (8)$$

$$\mathrm{MSE}_{u,i} = \frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{NDVI}_{\mathrm{int,p(i,j)}} - \overline{\overline{\mathrm{NDVI}}}_{\mathrm{int,p(i,j)}}\right)^2 / \overline{\mathrm{NDVI}}_{\mathrm{int,o(i)}}$$

$$\overline{\overline{\mathrm{NDVI}}}_{\mathrm{int,p(i,j)}} = a + b\ \mathrm{NDVI}_{\mathrm{int,o(i,j)}}$$

where a and b are the intercept and slope of the least square regression between the predicted and observed values of $\mathrm{NDVI}_{\mathrm{int}}$ for site i, respectively. From equation (8), model bias, that is, the percentage of systematic error, was calculated as $100^*(\mathrm{MSE}_{s,i}/\mathrm{MSE}_i)$.

We examined uncertainty in model outputs given uncertainty in parameters and uncertainty in NDVI data using a sampling-based method (Helton and others 2006). For each model, we produced a mapping $[P_i,\ F(P_i),\ i = 1,\ 2,\ldots,M]$ where P_i is a vector of parameter values and M is the number of simulations. For parametric uncertainty, we sampled $M = 5{,}000$ parameter sets from an independent uniform distribution using a Latin Hypercube Sampling approach. Minimum and maximum values of each uniform distribution were determined by an arbitrary range of variation of ±20% around the optimized parameter value obtained by the calibration tests. Assessing the impact of NDVI uncertainties on model performance was not straightforward because measurement errors have multiple sources and are hard to quantify (Seaquist and others 2003). Here, we chose to modify the values of the minimum NDVI, NDVI_0, and carrying capacity, $C_{\max}$, using the same sampling procedure, that is, $M = 5{,}000$ samplings from an independent uniform distribution in a ±20% range around the observed value.

Numerical simulations, statistical analyses, and all graphics were done within the R software environment (R Development Core Team 2007). The source code is available upon request. Computationally demanding calculations were performed on the cluster HealthPhy (CIMENT, Université J. Fourier—Grenoble I).

Results

Climate Envelope of Investigated Grasslands

The climate for most of the studied grasslands was semi-arid, as evidenced by a mean annual precipitation (MAP) around 400 mm (Figure 1A) and an aridity index, the ratio between MAP and mean annual potential evapotranspiration, of around 0.25 (Figure 1C). There was a pronounced latitudinal gradient from almost exclusive precipitation during the wet summer in the north to around 70% of total precipitation during the rainy season in the south (Figure 1B). Marked contrast between dry and wet seasons was a feature of all grasslands except at the very southern tip (lat. 30°S). Temperature of the coldest month (July) ranged between 15 and 19°C (Figure 1D), corresponding to subtropical conditions. A striking feature of the area was the high interannual variability in precipitation. The averaged coefficient of variation (CV) of annual precipitation across all sites was above 40% and there was no clear increase of the CV with decreasing MAP (Figure 3A). By comparison, the averaged CV for $\mathrm{NDVI}_{\mathrm{int}}$ was around 30% and there was a significant trend ($r = -0.34$, $P < 10^{-3}$), towards higher variability in sites with low $\mathrm{NDVI}_{\mathrm{int}}$ (Figure 3B).

Model Calibration and Validation

For the INT models, the linear coefficients for P_{int} and E_{int} had a CV around 50% (Table 2). The same variations were observed for most of the parameters of PR1 and PR2 models whereas CVs were around 30% for PR3. The leaf decay rate, β_3, was remarkably homogeneous across all three PR models. The threshold value for soil water content,

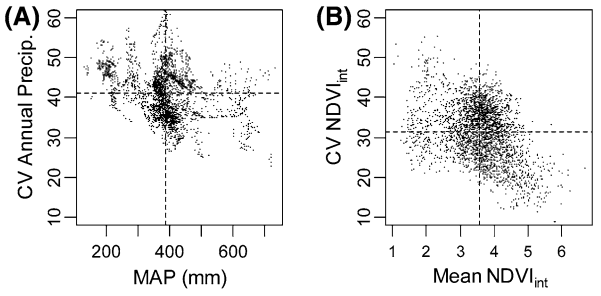

Figure 3. Relationship between mean and coefficient of variation (CV) for annual precipitation (**A**) and $\mathrm{NDVI}_{\mathrm{int}}$ (**B**). Each point represents one site. *Dotted lines* correspond to averages across sites.

Table 2. Site-calibrated Parameters for the Calibration Period 1981–1990

INT models	INT1	INT2	
α_1 Intercept	1.4 (1.1)	3.4 (1.5)	
$\alpha_2 \times 10^4$ Precipitation coeff.	53 (24)	50 (25)	
$\alpha_3 \times 10^4$ Potential evapotranspiration coeff.	na	−6.4 (±3.3)	
PR models	**PR1**	**PR2**	**PR3**
β_1 Water availability threshold	26 (17)	81 (42)	97 (45)
$\beta_2 \times 10^4$ PUE (PR1)/SWUE (PR2 − PR3)	46 (29)	28 (15)	36 (12)
β_3 Leaf decay rate	0.35 (0.13)	0.35 (0.12)	0.3 (0.11)
β_4 Parameter for evapotranspiration (PR2) or transpiration (PR3)	na	27.7 (14.9)	30.8 (11.2)

Mean (and SD) for n = 100 randomly sampled sites.

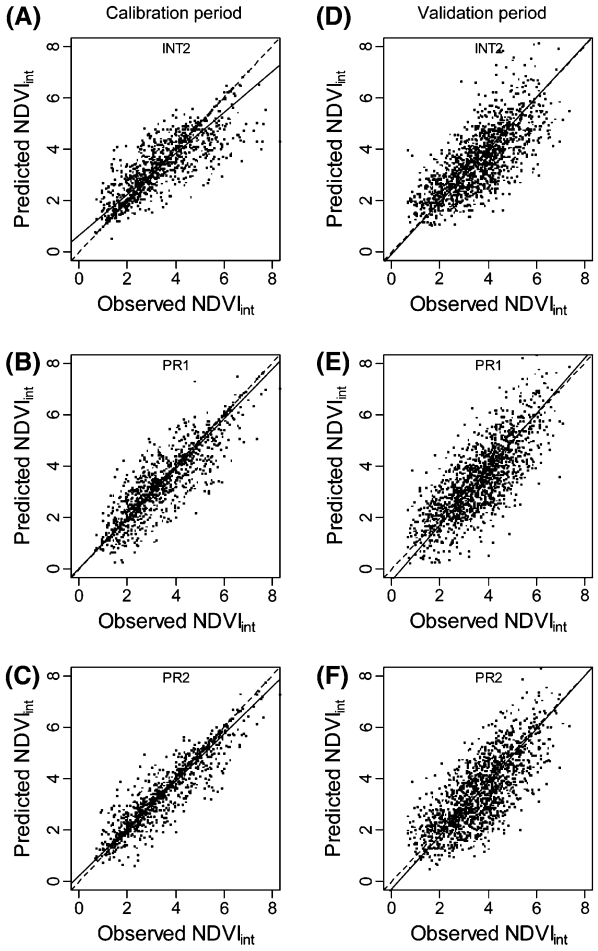

Figure 4. Predicted versus observed values of $NDVI_{int}$ in site simulations. Bivariate plots are shown for the calibration period 1981–1990 (**A–C**) and the validation period 1991–2004 (**D–F**) for 100 randomly sampled sites and for three different models. Each point represents 1 year and one site. *Dotted lines* show the 1:1 relationship and *solid lines* show the standardized major axes (SMA). See Table 3 for corresponding statistics of model performance.

β_1, was around 80–100 mm (PR2 and PR3), compared with 26 mm for effective precipitation in PR1.

There was a marked reduction in scatter from INT2 to PR1 or PR2 (Figure 4A, B, C) for the calibration period but these differences tended to diminish for the validation period (Figure 4D–F). For the calibration period, all goodness-of-fit statistics (MSE, MAE, CVMAE) were much lower in the PR models compared to the INT models and the bias was reduced by more than two-thirds, that is, 37–39% for INT1, INT2 against 9–11% for PR1 and PR3 (Table 3A). Based on log likelihood ratio tests, changes in model structure and complexity along the sequence INT1 to PR2 led to significant ($P < 10^{-4}$) improvement in model performance at each step. Further complexity of PR3 did not translate into higher performance for the calibration period (Table 3A). For the validation period, goodness-of-fit statistics differed little across models (Table 3B). Based on log likelihood ratio tests, we found slight improvements of PR1 and PR3 models compared to the linear models (Table 3B).

For ensemble simulations, differences in goodness-of-fit statistics were very small for both the calibration and validation periods (Table 4). We found even a slightly better fit for INT models based on MAE, CVMAE, and MSE and the lowest log likelihood was obtained for the 3-parameter model INT2 (Table 4). The main contrast between the INT and the PR models was the much higher bias in the former (more than 50% in the calibration period and around 40% for the validation period). The two INT models strongly overestimated $NDVI_{int}$ in the low range of observations and strongly underestimated it in the high range (Figure 5A, D). By contrast, the slopes of the standardized major axis (SMA) of the PR models were much closer to one and therefore the predictions did not deteriorate at

Table 3. Model Performance for Site Simulations

Model Number of parameters	INT1 2	INT2 3	PR1 3	PR2 4	PR3 4
A. Calibration period (1981–1990), n = 1,000					
r^2	0.80	0.81	0.87	**0.92**	0.90
MAE	0.59	0.57	0.50	**0.37**	0.42
CVMAE	0.17	0.17	0.15	**0.11**	0.12
MSE	0.77	0.71	0.53	**0.33**	0.38
Bias = 100 MSE_s/MSE	39	37	11	17	10
Log likelihood	−1,285	−1,250	−1,097	**−865**	−939
B. Validation period (1991–2004), n = 1,400					
r^2	0.73	0.73	**0.76**	0.74	0.74
MAE	0.73	0.74	0.74	0.77	0.73
CVMAE	0.20	0.20	0.20	0.21	0.20
MSE	0.92	0.94	0.90	0.95	0.89
Bias = 100 MSE_s/MSE	16	12	**9**	13	17
Log likelihood ($\times 10^4$)	−1,930	−1,941	−1,914	−1,953	**−1,901**

Results are shown for the calibration period (A) and for the validation period (B). Best values are in bold.

Table 4. Model Performance for Ensemble Simulations

Model Number of parameters	INT1 2	INT2 3	PR1 3	PR2 4	PR3 4
A. Calibration period (1981–1990), n = 29,070					
r^2	0.66	0.67	**0.73**	0.67	0.71
MAE	**0.80**	**0.80**	0.88	0.93	0.88
CVMAE	**0.23**	**0.23**	0.26	0.27	0.26
MSE	1.10	**1.07**	1.18	1.37	1.22
Bias = 100 MSE_s/MSE	58	54	13	18	17
Log likelihood	−42,678	**42,267**	−43,618	−45,772	−44,118
B. Validation period (1991–2004), n = 40,698					
r^2	0.69	0.70	**0.76**	0.71	0.72
MAE	**0.76**	**0.76**	0.81	0.88	0.80
(1) CVMAE	**0.21**	**0.21**	0.22	0.24	0.22
MSE	0.96	**0.94**	1.06	1.24	1.04
(2) Bias	42	36	**9**	20	23
Log likelihood	−56,952	**56,493**	−58,935	−62,083	−58,522
(3) Parametric sensitivity	**0.51**	1.69	0.67	0.83	0.74
(1) × (2) × (3)	4.51	12.77	**1.36**	4.07	3.77

Results are shown for the calibration period (A) and for the validation period (B). Best values are in bold.

the extremes of the precipitation gradient compared to the INT models (Figure 5B, C, E, F).

A quantitative estimate of model sensitivity is given by the length of the 25th to 75th percentile segment of the box plots that summarize the mappings $[P_i,\ F(P_i),\ i = 1,\ 2,\ \ldots,\ M]$ (Figure 6; Table 4B). The sensitivity of the models to parameter uncertainty was lowest for INT1 and PR1. Noticeably, the model showing the best performance in ensemble simulation (INT2) was the worst for its robustness to parametric uncertainty (Figure 6A). Sensitivity to $NDVI_0$ and C_{max} was small and did not differ among models and they all deteriorated to a similar extent when the original $NDVI_0$ and C_{max} values were replaced by randomly sampled values (Figure 6B).

Based on the results obtained for ensemble simulations and for the validation period, we benchmarked these five models by calculating the product of goodness-of-fit (CVMAE), bias, and parametric sensitivity. With similar weight given to these three criteria, the simplest PR model, that is,

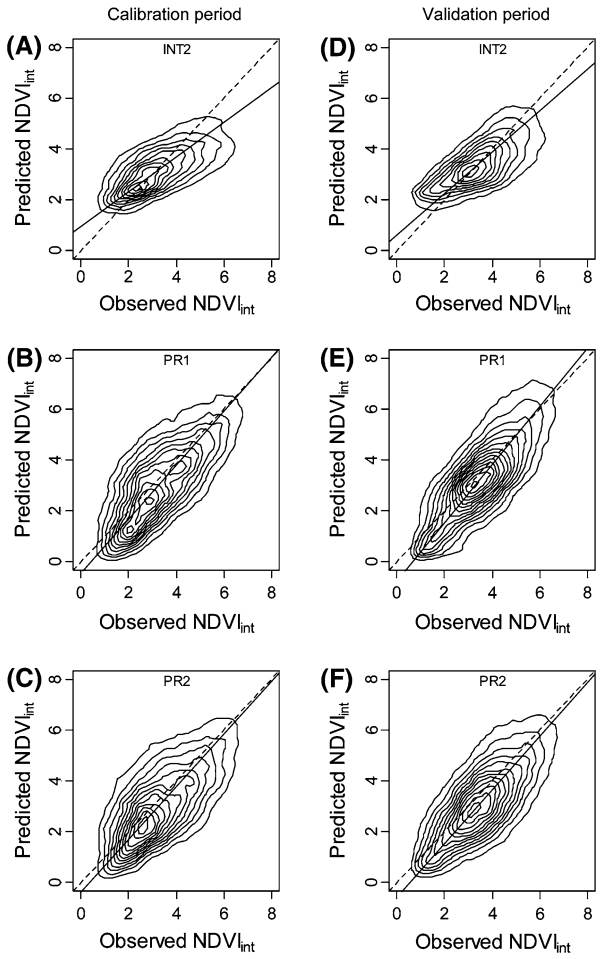

Figure 5. Predicted versus observed values of $NDVI_{int}$ in ensemble simulations. Density plots are shown for the calibration period 1981–1990 (**A**–**C**) and the validation period 1991–2004 (**D**–**F**) for the 2,907 sites and for three different models. *Dotted lines* show the 1:1 relationship and *solid lines* show the standardized major axes (SMA). See Table 4 for corresponding statistics of model performance.

PR1, was superior to the others (Table 4). Below, we further analyzed the residual of this model.

Analysis of PR1 Model Residual

The PR1 model performed worst in the driest regions (Figure 7). This is because the largest proportion of randomly chosen calibration sites are from the dominant 350–450 mm precipitation classes (Figure 1A) and thus they have an overwhelming influence on the calibrated parameter values. The performance of the PR1 model when calibrated for each precipitation class, hereafter PR1p, was consistent throughout the MAP gradient (Figure 7). Controlling for MAP was worthwhile in removing this source of systematic bias but resulted

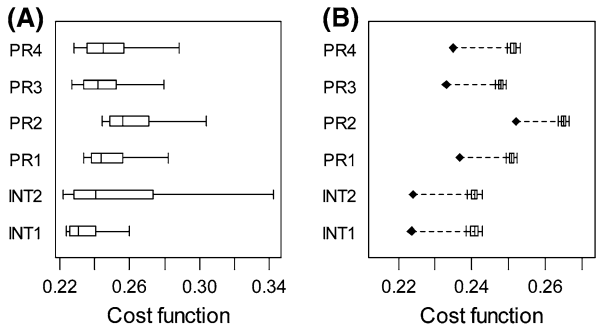

Figure 6. Sensitivity of model NDVI predictions to uncertainty in parameters (**A**) and NDVI data (**B**). *Box plots* summarize the distribution of the cost function values using 5,000 Latin Hypercube Samples of parameters. Results are shown for ensemble simulations and for the validation period 1991–2004. In **B**, *diamonds* indicate the value of the cost function with original NDVI data. The length of the *dashed lines* measures model deterioration when modifying $NDVI_0$ and C_{max}. Note the expanded scale in **B**.

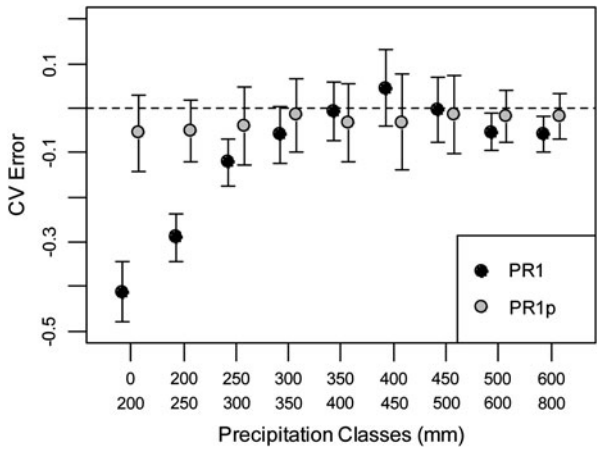

Figure 7. Effect of MAP on the error magnitude of the PR1 model. Parameters of the PR1p model were optimized for each precipitation class. Results show the averaged error across sites (±s.e.) for the calibration period 1981–1990 and for each precipitation class. Errors were normalized to the mean $NDVI_{int}$ of each site. Note that results for PR1 correspond to Figure 5B.

in only a small improvement in MAE of 0.80 compared to 0.88 for PR1 and MSE of 1.02 compared to 1.18 for PR1. Indeed, sources of uncertainty were observed within the most frequent precipitation classes (between 350 and 450 mm) and there was no reduction of this source of error from PR1 to PR1p (Figure 7). Calibrating the other PR models for each precipitation class gave similar results (data not shown).

The spatial distribution of error (normalized by the mean of observations) of PR1 showed some

clustering at the sub regional level (Figure 8A, B). Predicted values tended to be overestimated in the North-Eastern area (known locally as the Northern Downs) and underestimated in North-Western (the Barkly Tablelands) and South-Eastern areas (the Southern Downs) (Figure 8A). The same patterns were obtained for the calibration and validation periods though contrasts among sub regions were slightly attenuated in the validation period (Figure 8A, B). We found the same spatial patterns of error in all PR models (data not shown). It is noteworthy that there is little correlation between the distribution of climate variables and the distribution of PR model errors (compare Figures 1 and 8). Moreover, we found no relationship between error magnitude and the distance to the nearest weather station/rain gauge and no significant association with the two dominant soil types, that is, gray versus brown clay (data not shown).

Time series of yearly averaged errors did not show any clear multi-annual trends for either the calibration or the validation period (Figure 8B, D). Noticeable underestimations were observed for the first 3 years of simulations, raising possible issues with initial conditions. However, increasing the spin-up (and hence shortening the calibration period) did not affect the error magnitude in the first years (data not shown). Finally, yearly averages of error did not significantly correlate to any yearly averages of variables such as MAP, percentage of summer precipitation and temperature (data not shown).

Discussion

In this paper, we showed that PR models can be successfully calibrated to remotely sensed data of NDVI and, compared to linear models, can better predict the variability of $NDVI_{int}$ at a decadal scale for the studied perennial grasslands. This is supported by a benchmark analysis in which multiple criteria of a 'good' model have been examined. We also provided evidence that increasing complexity of PR models was defensible for application at a particular site but not for regional-scale simulations.

Our study has been conducted in semi-arid ecosystems where a single life form, that is, perennial grasses, significantly contributes to NDVI. Therefore, the difficulty of constructing a phenology model was reduced as we did not have to consider distinct parameters and model structure for several plant functional groups. A modeling framework for a tree–grass ecosystem would have to consider a different set of parameters for the pulse response function, a state variable for each dominant life

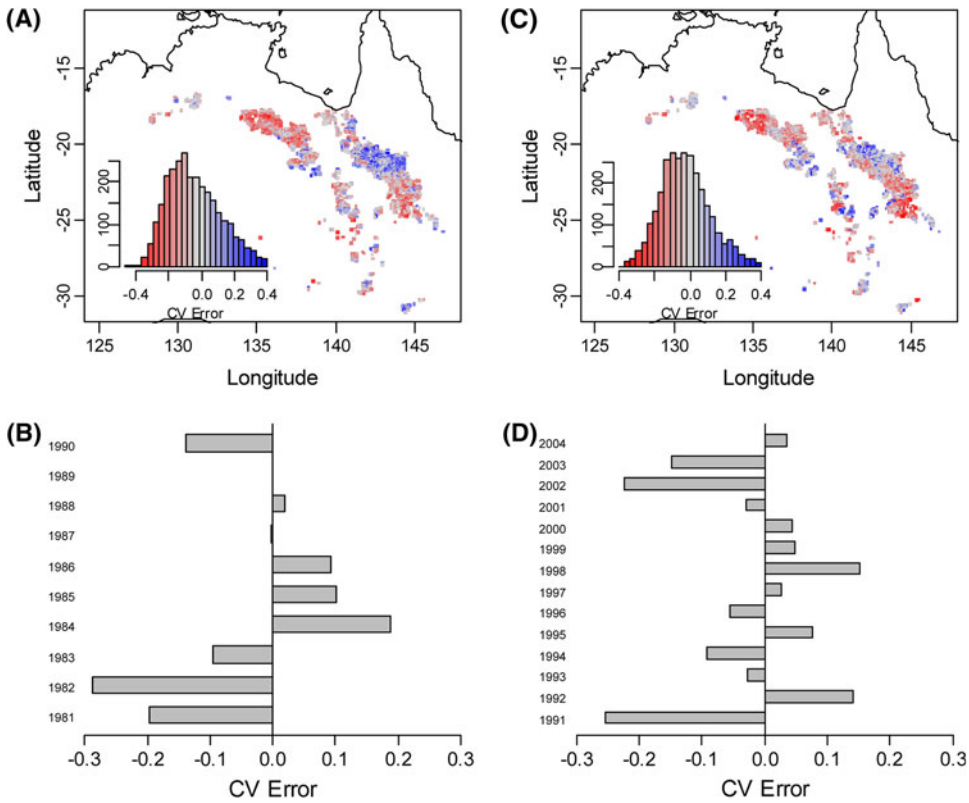

Figure 8. Spatial and temporal distribution of the error in ensemble simulation with the PR1p model. Error magnitude was normalized to the mean $NDVI_{int}$ of each site before spatial (**A**, **C**) and temporal (**B**, **D**) averaging. Results are shown for the calibration period (**A**, **B**) and the validation period (**C**, **D**).

form and additional couplings among state variables (Reynolds and others 2004; Caylor and others 2009). Another simplification was the almost lack of fires in the investigated area. The Australian fire history database indicated that no more than 1% of the area burned annually from 1997 onwards, except in 2001 where the estimate was 5% (Bastin and ACRIS Management Committee 2008). The same likely holds true before 1997 because strict control of fires is an essential part of the management of these tussock grasslands of which 95% are controlled by the cattle and sheep industry (Bastin and ACRIS Management Committee 2008). This situation drastically differed from nearby woodlands where the effect of fires on grass and tree leaf dynamics may cause losses in net annual primary productivity (Beringer and others 2007). Moreover, semi-arid tropical grasslands on clay-rich soils are particularly responsive to precipitation, with regression models likely to explain 70–80% of the observed variance in NDVI time series (Farrar and others 1994; Nicholson and Farrar 1994; Choler and others 2010), and confirmed by the present study. As reported in other grassland ecosystems (Yang and others 1998; Ji and Peters 2004), including both precipitation and evaporative demand in linear models can increase their performance. However, statistical NDVI-climate relationships remain highly empirical with strong dependencies on regional particulars of climate and vegetation (Mougin and others 1995). Other problems identified in this study included a large bias and high parametric sensitivity for INT2. These limitations clearly call for the development of more process-based PR models that account for essential features of the system dynamics.

There is a long-standing recognition that soil moisture dynamics are the primary drivers of plant response in semi-arid ecosystems (Noy-Meir 1973; Walker and Langridge 1996; Laio and others 2002). Nonetheless, the two PR models with a soil water balance did not exhibit higher performance than a purely precipitation-based PR model. Similarly, coupling soil water content and plant growth (PR3) did not reduce the error in ensemble simulations despite the intellectual appeal of including these feedbacks to capture the nonlinearities observed in plant responses and soil moisture dynamics (Hess and others 1996; Seaquist and others 2003). The reasons why increased complexity did not lead to better performance and robustness in ensemble simulations may be explained by (i) inadequate model formulation of the processes and/or their couplings, (ii) insufficient quality of the input data (climate forcing and observations) and, (iii) too coarse parameter evaluation. It is unlikely that a bad formulation or coupling of the processes was solely at stake because model performance did benefit from increased model complexity in site simulations at least for the calibration period. The effect of uncertainty in forcing variables was not formally tested. However, we believe this was not the main problem as sites closer to weather stations or rain gauges did not exhibit higher performance than those more distant. AVHRR NDVI have been used in numerous diagnostic studies examining vegetation response to long-term trends in climate (for example, Lotsch and others 2003; Nemani and others 2003; de Beurs and Henebry 2004). Our study shows that these data were appropriate to calibrate and test prognostic ecohydrological models of $NDVI_{int}$. Quality of input data was probably not the key issue in our study as site-specific calibration of PR models gave much more satisfactory results than their statistical counterparts (see Figure 4A–C). Moreover, we restricted our modeling study to time-averaged vegetation response that might be less sensitive to data uncertainty than studies focusing on particular events.

A more likely explanation for the deterioration of PR model performance outside the validation period and in ensemble simulations is that a number of unknown site-specific factors affect the system dynamics. Contrasting soil water holding capacity, soil fertility, plant functional diversity (Fensham and others 2000; O'Connor and others 2001; Tiver and others 2001), and grazing intensity (Foran and Bastin 1984; Bastin and others 1995) are among factors that may differ among sites or even change from year to year for a given site. The relative cover of perennial grasses is a dynamic variable that responds to grazing pressure changes and multiannual climatic trends. Because we made the assumption of a constant carrying capacity C_{max}, these possible long-term changes in perennial grass cover may have been overlooked.

Most, if not all, regional and global studies of phenology in water-limited ecosystems, have not properly examined the effects of different model structures on model performance and model robustness. In this study, we showed that benchmarking alternative model structure was crucial to determine (i) what physical and biological processes should be included in the modeling framework, (ii) how we should write the equations governing these processes, and (iii) how should these processes be coupled. There are many approaches to embody soil water balance and plant growth dynamics into a common ecohydrological modeling framework (for example, Walker and

Langridge 1996; Seaquist and others 2003; De Michele and others 2008). Here we advocate that, given the large unknowns in site variability listed above, adding too much complexity in PR models is, at this stage, questionable, or at least premature for regional applications. A key benefit from benchmarking models was to identify weaknesses in current model structure and consequently research directions. In this respect, we acknowledge that it would be unreasonable to refine soil water balance (for example, considering multiple layers, infiltration) without parallel effort to refine plant growth models.

Finally, we emphasized that multiple criteria are needed to evaluate competing models. There is a general tendency in ecological modeling to overlook model bias and model robustness and to primarily focus on goodness-of-fit statistics (such as r^2). Our case study showed that linear models, such as INT2, can exhibit the lowest error magnitude though being strongly biased or very sensitive to parameter uncertainty.

Conclusions

Our two main conclusions are (i) the low-dimensional pulse response model (PR1) best captures regional-scale and long-term interannual variability of $NDVI_{int}$ of tropical grasslands and (ii) increasing the dimensionality of the PR models is not justified in regional applications where site-specific information is missing. Though the investigated grasslands were a rather simple case study with its known high responsiveness to precipitation, there remained large uncertainties in the predictions. A further understanding of multi-annual changes in the carrying capacity of these grasslands is a promising way for reducing the unexplained variance in times series of $NDVI_{int}$. Our findings provide evidence that simple PR models can be tested at the global scale and over multiple decades using remotely sensed data. Building on this modeling framework, further studies focusing on the grassland biome should examine causal relationships between climate and plant growth in distinct tropical regions or across larger bioclimatic gradients.

Acknowledgments

This work was funded by a Marie-Curie fellowship to Philippe Choler and is part of the project CASOAR (contract MOIF-CT-2006-039688). W. Sea was partially funded by an office of the chief executive post-doctoral CSIRO fellowship. We thank Stephen Roxburgh and Michael Raupach for helpful comments on this work.

References

Abramowitz G. 2005. Towards a benchmark for land surface models. Geophys Res Lett 32. doi:10.1029/2005GL024419.

Alewell C, Manderscheid B. 1998. Use of objective criteria for the assessment of biogeochemical ecosystem models. Ecol Model 107:213–24.

Anyamba A, Tucker CJ. 2005. Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. J Arid Environ 63:596–614.

Bastin G, ACRIS Management Committee. 2008. Mitchell Grass Downs Bioregion (Queensland and the NT). In: Australian Collaborative Rangeland Information System (ACRIS), Ed. Rangelands 2008—taking the pulse. Canberra: National Land & Water Resources Audit, Australian Government. p 151–9.

Bastin GN, Pickup G, Pearce G. 1995. Utility of AVHRR data for land degradation assessment—a case-study. Int J Remote Sens 16:651–72.

Beringer J, Hutley LB, Tapper NJ, Cernusak LA. 2007. Savanna fires and their impact on net ecosystem productivity in North Australia. Glob Change Biol 13:990–1004.

Byrd RH, Lu PH, Nocedal J, Zhu CY. 1995. A limited memory algorithm for bound constrained optimization. Siam J Sci Comput 16:1190–208.

Carlson TN, Ripley DA. 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens Environ 62:241–52.

Carlson TN, Capehart WJ, Gillies RR. 1995. A new look at the simplified method for remote-sensing of daily evapotranspiration. Remote Sens Environ 54:161–7.

Caylor KK, Scanlon TM, Rodriguez-Iturbe I. 2009. Ecohydrological optimization of pattern and processes in water-limited ecosystems: a trade-off-based hypothesis. Water Resources Research 45. doi:10.1029/2008wr007230.

Choler P, Sea W, Briggs P, Raupach M, Leuning R. 2010. A simple ecohydrological model captures essentials of seasonal leaf dynamics in semi-arid tropical grasslands. Biogeosciences 7:907–20.

Choudhury BJ, Ahmed NU, Idso SB, Reginato RJ, Daughtry CST. 1994. Relations between evaporation coefficients and vegetation indexes studied by model simulations. Remote Sens Environ 50:1–17.

Christie EK. 1981. Biomass and nutrient dynamics in a C-4 semi-arid Australian grassland community. J Appl Ecol 18:907–18.

Daly E, Porporato A. 2005. A review of soil moisture dynamics: from rainfall infiltration to ecosystem response. Environ Eng Sci 22:9–24.

de Beurs KM, Henebry GM. 2004. Land surface phenology, climatic variation, and institutional change: analyzing agricultural land cover change in Kazakhstan. Remote Sens Environ 89:497–509.

De Michele C, Vezzoli R, Pavlopoulos H, Scholes RJ. 2008. A minimal model of soil water-vegetation interactions forced by stochastic rainfall in water-limited ecosystems. Ecol Model 212:397–407.

Donohue RJ, McVicar TR, Roderick ML. 2009. Climate-related trends in Australian vegetation cover as inferred from satellite observations, 1981–2006. Glob Change Biol 15:1025–39.

Farrar TJ, Nicholson SE, Lare AR. 1994. The influence of soil type on the relationships between NDVI, rainfall, and soil-moisture in semiarid Botswana. 2. NDVI response to soil-moisture. Remote Sens Environ 50:121–33.

Fensham RJ, Minchin PR, Fairfax RJ, Kemp JE, Purdie RW, McDonald WJF, Neldner VJ. 2000. Broad-scale environmental relations of floristic gradients in the Mitchell grasslands of Queensland. Aust J Bot 48:U27–34.

Foran BD, Bastin G. 1984. The dynamics of a Mitchell grass (*Astrebla* spp.) rangeland on the Barkly Tableland, Northern Territory. Aust Rangeland J 6:92–7.

Goetz SJ, Prince SD, Goward SN, Thawley MM, Small J. 1999. Satellite remote sensing of primary production: an improved production efficiency modeling approach. Ecol Model 122:239–55.

Gulden LE, Rosero E, Yang ZL, Wagener T, Niu GY. 2008. Model performance, model robustness, and model fitness scores: a new method for identifying good land-surface. Geophys Res Lett 35. doi:10.1029/2008GL033721.

Hashimoto H, Nemani RR, White MA, Jolly WM, Piper SC, Keeling CD, Myneni RB, Running SW. 2004. El Nino-southern oscillation-induced variability in terrestrial carbon cycling. J Geophys Res-Atmos 109. doi:10.1029/2004JD004959.

Helton JC, Johnson JD, Sallaberry CJ, Storlie CB. 2006. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliabil Eng Syst Saf 91:1175–209.

Hess T, Stephens W, Thomas G. 1996. Modelling NDVI from decadal rainfall data in the north east arid zone of Nigeria. J Environ Manag 48:249–61.

Heumann BW, Seaquist JW, Eklundh L, Jonsson P. 2007. AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982–2005. Remote Sens Environ 108:385–92.

Hird JN, McDermid GJ. 2009. Noise reduction of NDVI time series: an empirical comparison of selected techniques. Remote Sens Environ 113:248–58.

Hunter DM. 1989. The response of Mitchell grasses (*Astrebla* spp.) and Button grass (*Dactyloctenium radulans* (R Br)) to rainfall and their importance to the survival of the Australian plague locust, *Chortoicetes terminifera* (Walker), in the arid zone. Aust J Ecol 14:467–71.

Jeffrey SJ, Carter JO, Moodie KB, Beswick AR. 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ Model Softw 16:309–30.

Ji L, Peters AJ. 2004. A spatial regression procedure for evaluating the relationship between AVHRR-NDVI and climate in the northern Great Plains. Int J Remote Sens 25:297–311.

Laio F, Porporato A, Ridolfi L, Rodriguez-Iturbe I. 2002. On the seasonal dynamics of mean soil moisture. J Geophys Res-Atmos 107. doi:10.1029/2001JD001252.

Liu YQ, Gupta HV. 2007. Uncertainty in hydrologic modeling: toward an integrated data assimilation framework. Water Resour Res 43. doi:10.1029/2006WR005756.

Lotsch A, Friedl MA, Anderson BT, Tucker CJ. 2003. Coupled vegetation-precipitation variability observed from satellite and climate records. Geophys Res Lett 30:1774. doi:1710.1029/2003GL017506.

McKenzie NJ, Hook J. 1992. Interpretation of the Atlas of Australian Soils. CSIRO Division of Soils Technical Report 94/1992. CSIRO, Canberra, ACT, Australia.

McKenzie NJ, Jacquier DW, Ashton LJ, Cresswell HP. 2000. Estimation of soil properties using the Atlas of Australian Soils. CSIRO Land and Water Technical Report 11/00. CSIRO, Canberra, ACT, Australia, p. 24.

Mebane WR, Sekhon JS. 2009. R-GENetic Optimization Using Derivatives (RGENOUD). http://sekhon.berkeley.edu/rgenoud/.

Montandon LM, Small EE. 2008. The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. Remote Sens Environ 112:1835–45.

Monteith JL. 1977. Climate and efficiency of crop production in Britain. Philos Trans Roy Soc Lond Ser B-Biol Sci 281:277–94.

Mougin E, Loseen D, Rambal S, Gaston A, Hiernaux P. 1995. A regional Sahelian grassland model to be coupled with multispectral satellite data. 1. Model description and validation. Remote Sens Environ 52:181–93.

Myneni RB, Hall FG, Sellers PJ, Marshak AL. 1995. The interpretation of spectral vegetation indexes. IEEE Trans Geosci Remote Sens 33:481–6.

Myneni RB, Keeling CD, Tucker CJ, Asrar G, Nemani RR. 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. Nature 386:698–702.

Nemani RR, Keeling CD, Hashimoto H, Jolly WM, Piper SC, Tucker CJ, Myneni RB, Running SW. 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. Science 300:1560–3.

Nicholson SE, Farrar TJ. 1994. The influence of soil type on the relationships between NDVI, rainfall, and soil-moisture in semiarid Botswana. 1. NDVI response to rainfall. Remote Sens Environ 50:107–20.

Noy-Meir I. 1973. Desert ecosystems: environment and producers. Annu Rev Ecol Syst 4:25–51.

O'Connor TG, Haines LM, Snyman HA. 2001. Influence of precipitation and species composition on phytomass of a semi-arid African grassland. J Ecol 89:850–60.

Ogle K, Reynolds JF. 2004. Plant responses to precipitation in desert ecosystems: integrating functional types, pulses, thresholds, and delays. Oecologia 141:282–94.

Paruelo JM, Lauenroth WK. 1998. Interannual variability of NDVI and its relationship to climate for North American shrublands and grasslands. J Biogeogr 25:721–33.

Paruelo JM, Epstein HE, Lauenroth WK, Burke IC. 1997. ANPP estimates from NDVI for the Central Grassland Region of the United States. Ecology 78:953–8.

Pettorelli N, Vik JO, Mysterud A, Gaillard JM, Tucker CJ, Stenseth NC. 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends Ecol Evol 20:503–10.

Pineiro G, Oesterheld M, Paruelo JM. 2006. Seasonal variation in aboveground production and radiation-use efficiency of temperate rangelands estimated through remote sensing. Ecosystems 9:357–73.

Pitman AJ, Henderson-Sellers A, Desborough CE, Yang ZL, Abramopoulos F, Boone A, Dickinson RE, Gedney N, Koster R, Kowalczyk E, Lettenmaier D, Liang X, Mahfouf JF, Noilhan J, Polcher J, Qu W, Robock A, Rosenzweig C, Schlosser CA, Shmakin AB, Smith J, Suarez M, Verseghy D, Wetzel P, Wood E, Xue Y. 1999. Key results and implications from phase 1(c) of the project for intercomparison of land-surface parametrization schemes. Climate Dyn 15:673–84.

Porporato A, D'Odorico P, Laio F, Ridolfi L, Rodriguez-Iturbe I. 2002. Ecohydrology of water-controlled ecosystems. Adv Water Resour 25:1335–48.

Potter CS, Brooks V. 1998. Global analysis of empirical relations between annual climate and seasonality of NDVI. Int J Remote Sens 19:2921–48.

Priestley CHB, Taylor RJ. 1972. Assessment of surface heat-flux and evaporation using large-scale parameters. Mon Weather Rev 100:81–92.

Prince SD, Goetz SJ, Goward SN. 1995. Monitoring primary production from Earth observing satellites. Water Air Soil Pollut 82:509–22.

Propastin PA, Kappas M. 2008. Reducing uncertainty in modeling the NDVI-precipitation relationship: a comparative study using global and local regression techniques. GISci Remote Sens 45:47–67.

R Development Core Team. 2007. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Reynolds JF, Kemp PR, Ogle K, Fernandez RJ. 2004. Modifying the 'pulse-reserve' paradigm for deserts of North America: precipitation pulses, soil water, and plant responses. Oecologia 141:194–210.

Richard Y, Poccard I. 1998. A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in Southern Africa. Int J Remote Sens 19:2907–20.

Rykiel EJ. 1996. Testing ecological models: the meaning of validation. Ecol Model 90:229–44.

Savitzky A, Golay MJE. 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 36:1627–39.

Seaquist JW, Olsson L, Ardo J. 2003. A remote sensing-based primary production model for grassland biomes. Ecol Model 169:131–55.

Tiver F, Nicholas M, Kriticos D, Brown JR. 2001. Low density of prickly acacia under sheep grazing in Queensland. J Range Manag 54:382–9.

Tucker CJ, Sellers PJ. 1986. Satellite remote-sensing of primary production. Int J Remote Sens 7:1395–416.

Tucker CJ, Vanpraet C, Boerwinkel E, Gaston A. 1983. Satellite remote-sensing of total dry-matter production in the Senegalese Sahel. Remote Sens Environ 13:461–74.

Tucker CJ, Pinzon JE, Brown ME, Slayback DA, Pak EW, Mahoney R, Vermote EF, El Saleous N. 2005. An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. Int J Remote Sens 26:4485–98.

Walker BH, Langridge JL. 1996. Modelling plant and soil water dynamics in semi-arid ecosystems with limited site data. Ecol Model 87:153–67.

Willmott CJ. 1982. Some comments on the evaluation of model performance. Bull Am Meteorol Soc 63:1309–13.

Willmott CJ, Matsuura K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Res 30:79–82.

Willmott CJ, Matsuura K. 2006. On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. Int J Geograph Inf Sci 20:89–102.

Yang LM, Wylie BK, Tieszen LL, Reed BC. 1998. An analysis of relationships among climate forcing and time-integrated NDVI of grasslands over the US northern and central Great Plains. Remote Sens Environ 65:25–37.