Quantifying characteristic growth dynamics in a semi-arid grassland ecosystem by predicting short-term NDVI phenology from daily rainfall: a simple four parameter coupled-reservoir model

John F. Hermance, David J. Augustine & Justin D. Derner


To link to this article: http://dx.doi.org/10.1080/01431161.2015.1103916

Published online: 26 Oct 2015.

Article views: 65

View related articles

View Crossmark data
Quantifying characteristic growth dynamics in a semi-arid grassland ecosystem by predicting short-term NDVI phenology from daily rainfall: a simple four parameter coupled-reservoir model

John F. Hermance\textsuperscript{a*}, David J. Augustine\textsuperscript{b}, and Justin D. Derner\textsuperscript{c}

\textsuperscript{a}Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI 02912, USA; \textsuperscript{b}Rangeland Resources Research Unit, USDA—Agricultural Research Service, Fort Collins, CO 80526, USA; \textsuperscript{c}Rangeland Resources Research Unit, USDA—Agricultural Research Service, Cheyenne, WY 82009, USA

(Received 23 March 2015; accepted 26 September 2015)

Predicting impacts on phenology of the magnitude and seasonal timing of rainfall pulses in water-limited grassland ecosystems concerns ecologists, climate scientists, hydrologists, and a variety of stakeholders. This report describes a simple, effective procedure to emulate the seasonal response of grassland biomass, represented by the satellite-based normalized difference vegetation index (NDVI), to daily rainfall. The application is a straightforward adaptation of a staged linear reservoir that simulates the pulse-like entry of rainwater into the soil and its redistribution as soil moisture, the uptake of water by plant roots, short-term biomass development, followed by the subsequent transpiration of water through foliage. The algorithm precludes the need for detailed, site specific information on soil moisture dynamics, plant species, and the local hydroclimate, while providing a direct link between discrete rainfall events and consequential biomass responses throughout the growing season. We applied the algorithm using rainfall data from the Central Plains Experimental Range to predict vegetation growth dynamics in the semi-arid shortgrass steppe of North America. The mean annual rainfall is 342 mm, which is strongly bifurcated into a dominantly ‘wet’ season, where during the three wettest months (May, June and July) the mean monthly rainfall is approximately 55 mm month\textsuperscript{-1}; and a ‘dry’ season, where during the three driest months (December, January and February), the mean monthly rainfall is approximately 7 mm month\textsuperscript{-1}. NDVI data from the Moderate Resolution Imaging Spectroradiometer (MODIS) MOD13Q1 16 day, 250 m × 250 m product were used as a proxy for grassland phenology for the period-of-record 2000–2013. Allowing for temporal changes in basic parameters of the response function over the growing season, the predicted response of the model tracks the observed NDVI metric with correlation coefficients exceeding 0.92. A two-stage series reservoir is preferred, whereby the characteristic time for transfer of a rainfall event to the peak response of NDVI decreases from 24 days (early growing season) to 12 days (late growing season), while the efficiency of a given volume of rainfall to produce a correspondingly similar amount of aboveground biomass decreases by a factor of 40% from April to October. Behaviours of the characteristic time of greenup and loss of rainfall efficiency with progression of the growing season are consistent with physiological traits of cool-season C\textsubscript{3} grasses versus warm-season C\textsubscript{4} grasses, and with prior research suggesting that early season production by C\textsubscript{3} grasses is more responsive to a given amount of precipitation than mid-summer growth of C\textsubscript{4} shortgrasses. Our model explains >90% of seasonal biomass dynamics. We ascribe a systematic underprediction of observed early season greenup following drought years to a lagged or ‘legacy’ effect, as soil inorganic nitrogen, accumulated during drought, becomes available for future plant uptake.

*Corresponding author. Email: John_Hermance@Brown.Edu

© 2015 Taylor & Francis
1. Introduction and background

Many of the world’s water-limited grassland ecosystems experience dramatic, short-term (daily, weekly, and monthly) fluctuations in seasonal rainfall as a result of their hydroclimatic settings (Magliano et al. 2015). The major fraction of the total rainfall in these systems is delivered by a series of significant rainfall ‘pulses’ (Asbjornsen et al. 2011; Austin et al. 2004; Li, Zhao, and Liu 2013; Reynolds et al. 2004). Predicting impacts of short-term rainfall events on grassland ecosystems is a direct concern, not only to climate scientists and hydrologists, but to a variety of stakeholders, including range managers, water managers, ranchers, and pastoral herders.

Broad, first-order effects of interannual and intraseasonal fluctuations in rainfall on productivity of semi-arid grasslands are routinely documented from in situ field studies comparing direct harvesting of aboveground net primary production (ANPP) to local gauge measurements (Derner et al. 2008; Le Houerou 1984; Li, Zhao, and Liu 2013; Sala et al. 2012). With the advent of satellites, field measurements of ANPP have been invaluable in calibrating optical and near-infrared spectral data from low earth orbiting instruments to systematically assess the response of grassland biomass to seasonal variability of rainfall in a variety of grassland and rangeland settings, including the Sahel and East Africa (Justice, Holben, and Gwynne 1986; Nicholson, Davenport, and Malo 1990; Tucker, Justice, and Prince 1986; Tucker et al. 1985), South Africa (Wessels et al. 2006), and semi-arid and plains grasslands in the western USA (Moran et al. 2014), among others. Of increasing concern is the relationship of shorter-term, intraseasonal rainfall fluctuations to attendant behaviours of grassland phenology for modelling these interactions (Brunsell and Young 2008; Hess, Stephens, and Thomas 1996; Ogle and Reynolds 2004; Quiroz et al. 2010) and applying models to understanding dynamics of plant growth (Andales et al. 2006; Derner and Hart 2007; Proud and Rasmussen 2011). Upscaling plant–water interactions from local to catchment and regional scales (Asbjornsen et al. 2011; Choler et al. 2010), as well as the role of phenology in providing feedback to the weather–climate system (Richardson et al. 2013) are contemporary concerns for a range of applications. Understanding the intraannual/intraseasonal connection between episodic pulses of rainfall and the response of seasonal phenology of vegetation – particularly at the local level – is important for ecosystem processes and functions in semi-arid grasslands, as well as to separating effects of modulated rainfall from other factors (e.g. land management, land-use change) affecting land-cover and primary productivity. Moreover, increased knowledge of the direct intraannual rainfall–phenology relationship facilitates separation of these responses from those associated with multiyear and longer-term climate variability (Lauenroth and Sala 1992; Richardson et al. 2013).

Our primary objective here is to describe how an elementary, conceptually-based, phenomenological model can be used to predict seasonal phenology and productivity of a water-limited, semi-arid grassland from daily rainfall. Our model examines the dynamics of the North American shortgrass steppe (Figure 1), a region that is characterized by substantial daily, seasonal, and annual variability in precipitation. The computational algorithm is based on the concept of cascaded linear reservoirs, which can be viewed as a series of leaky buckets (Figure 2), well-known in the hydrology literature to represent characteristic time delays in the flow of water from one reservoir to another (Chow, Maidment, and Mays 1988; Nash 1958).

Correspondent grassland phenology is provided by estimates of vegetative biomass productivity represented by bi-weekly (16 day) values of the satellite-derived ‘greenness’ index: the normalized difference vegetation index (NDVI). This parameter has evolved from the early work of Pearson and Miller (1972), Rouse et al. (1973) and Tucker (1979) to
Figure 1. Location of study area (square symbol) for this project in the North American Short Grass Steppe (yellow-coloured hatched area). Contour intervals are 500 m, with labels at 1000 m intervals.

Figure 2. (a) Example of a two-stage cascaded linear reservoir. (b) Theoretical time dependent fluxes of the input pulse (approximating an ‘impulse’) referred to the vertical axis at the far left; the output of Stage 1 referred to the second axis from the left; and the response output from Stage 2 referred to the axis on the right. Units are in terms of length \((L)\) and inverse time \((T^{-1})\), in our case: rainfall depth per day. Both stages of the reservoir have a characteristic response time of 5 days. As a phenomenological model, Stage 1 might be thought of as emulating the entry stage of water into the soil and its redistribution as soil moisture; whereas Stage 2 emulates the uptake of water by plant roots, short-term biomass development, and subsequent transpiration through foliage.
become a standard metric for assessing the state and vegetative productivity of grasslands (Li and Guo 2012; Rojas et al. 2011) and semi-arid rangelands of mixed vegetation (Richard and Poccard 1998; Stoms and Hargrove 2000; Svoray, Perevolotsky, and Atkinson 2013; Tucker and Nicholson 1999). First, we describe the conceptual and theoretical basis of our method; second, we demonstrate the utility of this method over a 14 year period (2000–2013) in the western portion of the North American Great Plains (Figure 1).

2. Study area and data

2.1. Study area

Our study site (40.8441° N, 104.7097° W; see Figure 1) is located in the North American shortgrass steppe at the US Department of Agriculture (USDA) Agricultural Research Service (ARS) Central Plains Experimental Range (CPER). The CPER and the surrounding Pawnee National Grassland has a mean annual precipitation of 342 mm (1939–2014), and has served as a research sites for rangeland ecosystem dynamics and livestock grazing since 1937 (Lauenroth and Burke 2008; Lauenroth and Sala 1992; Milchunas, Forwood, and Lauenroth 1994). Some of the earliest ground-based, high altitude aircraft and satellite remote-sensing studies of grassland ecosystems were calibrated against field samples at this location (Pearson and Miller 1972; Tucker and Miller 1977). Here, we examine vegetation dynamics measured using remotely-sensed imagery (see Section 2.3 below) for a 250 m × 250 m area of managed grassland located adjacent to, and approximately 400 m from, a precipitation gauge at the CPER. Shortgrass steppe-type vegetation is underlain by sandy-loam soils. The plant community is dominated by native C\textsubscript{4} (warm-season), grazing-resistant short grasses such as blue grama (Bouteloua gracilis) and buffalo grass (B. dactyloides), with additional contributions from C\textsubscript{3} (cool-season) grasses, primarily needle and thread (Hesperostipa comata) and plains pricklypear cactus (Opuntia polyacantha). The site has been managed for cattle production with grazing at moderate stocking rates (approximately 0.6 animal unit months ha\textsuperscript{-1}) during May–October each year since the 1930s (Hart and Ashby 1998).

2.2. Rainfall data

The CPER meteorological station (40.8415° N, 104.7138° W) provided near-daily rainfall data (see Figure 3 for seasonal rainfall statistics), with precipitation recorded daily Monday–Friday at 8:00 am since 1939. Monday am recordings, therefore, represent three-day cumulative rainfall amounts (Friday, Saturday, and Sunday). These data were corroborated with an on-site automated precipitation gauge (2004–present), and manually read precipitation gauges distributed across CPER checked three times weekly since 1939 (Augustine 2010). Example daily data are also used from the Global Historical Climate Network (GHCN) station USW094074, NUNN 7 NNE (40.8066° N, 104.7552° W). Figure 3 summarizes monthly rainfall statistics for the period of record 1939–2014 from the ARS-Headquarters (HQ) station. Seasonal rainfall is strongly characterized by a summer rainy season of high variability, and a winter season of low to negligible rainfall and snow. The mean monthly rainfall in the three driest months (December, January and February) is approximately 7 mm month\textsuperscript{-1}, and in the three wettest months (May, June and July) is approximately 55 mm month\textsuperscript{-1}.
2.3. NDVI data

2.3.1. Background on the NDVI

The NDVI is defined using handheld or satellite-based spectrometer data by the relation
\[
\text{NDVI} = \frac{R_{\text{NIR}} - R_{\text{Red}}}{R_{\text{NIR}} + R_{\text{Red}}}
\]
where \(R_{\text{NIR}}\) is the reflectance from a near-infrared channel and \(R_{\text{Red}}\) is the reflectance from a visible red channel (Rouse et al. 1973; Tucker 1979). Prior to the release of a data set by the respective agency, the effects of transient cloudcover and other factors are minimized for each specific ground site (or pixel) by selecting optimal data values from a number of values within a specified time interval: 10 days, 16 days, monthly, etc., depending on the data product (Holben 1986). The sampling interval is known as a ‘compositing period’.

Being based on optical and near infrared spectral information directly related to the photosynthetic capacity of vegetation, the NDVI has been used for many years to derive estimates of biomass accumulation with respect to time and has proven to be a useful remotely-sensed proxy for aboveground biomass production in grasslands (Tucker, Justice, and Prince 1986; Wessels et al. 2006). While the majority of studies using this metric to study semi-arid grasslands and related ecosystems have used satellite observations at the regional scale (An, Price, and Blair 2013; Prince, De Colstoun, and Kravitz 1998; Proud and Rasmussen 2011; Rojas et al. 2011), the actual ground-based locations sampled by satellite images are sufficiently well georeferenced that applications of NDVI data to local scale – even pasture-scale – investigations have been proven to be effective (Grigera, Oesterheld, and Pacin 2007; Li, Zhao, and Liu 2013; Sha et al. 2014). It is this local scale to which the present study is directed.

2.3.2. Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI product:

\textit{MOD13Q1}

NDVI data used in this study are the MOD13Q1 data product derived from the MODIS instrument on board the Terra spacecraft (LP DAAC 2015). Starting in early 2000, the
MODIS Land Team (MODLND) has prepared the MOD13Q1 product with a 16 day compositing period and 250 m × 250 m ground level spatial resolution. Data used here were downloaded from the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC 2014). Figure 4 shows interannual fluctuations in surface greenness at one of our study sites using the MOD13Q1 NDVI data product. The data as provided by the Land Processes Distributed Active Archive Centers (LP DAACs) are shown by the stepped time series plot, along with nine-day Gaussian smoothed accumulations of daily rainfall from the nearby GHCN gauge Nunn 7 referred to in Section 2.2.

A close inspection of the NDVI time series in Figure 4 shows data dropouts due to snow cover, and sometimes cloud cover, as well as the occasional hike-up or spike. The MODIS LP DAACs preprocess their data attempting to flag as many of these non-NDVI anomalies as possible, and distribute several levels of quality assurance (QA) files along with their NDVI base product. These QA files are essential when determining whether to further preprocess observed values. For the period of record of our analysis (2000–2013), there were 319 16 day compositing periods. Of these, five values were not produced by the MODIS Land Group, and another 16 were rejected during our analysis due to snow cover or other anomalously low or negative NDVI values. For the data used in this report, data gaps were sufficiently short and non-strategic that a linear interpolation of missing values was sufficient and preferable over such alternative ‘noise-filtering’ methods as described by Hermance et al. (2007).

2.3.3. NDVI versus alternative vegetation indices

Other vegetation indices (VIs) have been developed to deal with a variety of practical issues such as soil reflectance, atmospheric aerosols, and specific canopy characteristics. One of these is the enhanced vegetation index, or EVI (Huete et al. 2002). The EVI has lower dynamic range than the NDVI, but it is reported that the EVI has better

Figure 4. Typical time series of NDVI data from our study area: the MODIS MOD13Q1 16 day 250 m² resolution data product (light grey stepped time series plot), and nine-day smoothed accumulations of daily rainfall from a nearby gauge (dark grey shaded line plot). Note the data anomalies (snow cover and data quality) routinely flagged by the LP DAAC.
response for high biomass regions such as tropical rain forests, and is more responsive than NDVI to canopy structural variations, including leaf area index (LAI), canopy type, plant physiognomy, and canopy architecture (Huete et al. 2002). The same work suggests that EVI may be more reliable for bare soil and snow cover conditions, although NDVI typically spans a higher range in values over semi-arid sites, which is a particular advantage for our study area. In order to compare the EVI signature of vegetation for our site to the NDVI signature, we performed a first-order linear regression of EVI on NDVI for quality-filtered time series from our test site for the period of record of 2000–2014. The results provided a correlation coefficient of 0.976 and indicated that EVI had a significantly lower range than NDVI as generically predicted by Huete et al. (2002). For our time series data from the CPER, the range of EVI is lower than NDVI by a factor of 0.5752, with the best-fitting line having a relatively small, negative intercept of $-0.000842$.

These regression coefficients were used to linearly rescale the EVI time series to approximately the same scale as NDVI, finding a root mean square difference between the original NDVI and rescaled EVI time series of 0.0227 VI units, and that 90% of the NDVI values fell in the range $0.155 \leq \text{NDVI} \leq 0.464$, and 90% of the EVI values fell in the range $0.151 \leq \text{EVI} \leq 0.469$. The latter comparison, along with the Pearson correlation coefficient of $r = 0.976$ between EVI and NDVI, shows a close similarity between the two types of vegetation indices, leading us to conclude that, for purposes of our analysis of the short grass steppe at CPER, the EVI and NDVI time series (while scaled differently) provide similar information, and that factors which the EVI is designed to minimize – such as the effects of soil reflectance, atmospheric aerosols, and specific canopy characteristics – do not seem to be anomalously affecting the NDVI data from this site. However, without explicit testing, this cannot be inferred to be the case for grasslands in general, particularly, for sparsely vegetated to semi-barren areas when soil reflectance becomes a major distorting influence (Huete, Jackson, and Post 1985). For our study area, we feel that the intrinsic simplicity and wider dynamic range (NDVI $\geq 1.74$ EVI) of the NDVI are favourable attributes.

2.3.4. Differential NDVI ($\delta$NDVI)

As others (Grist, Nicholson, and Mpolokang 1997; Hess, Stephens, and Thomas 1996; Nicholson, Tucker, and Ba 1998), we view the NDVI time series ‘signal’ in Figure 4 as the superposition of two components: a ‘background’ fallow season component attributed to reflectance from fallow ground and dormant, dry season vegetation litter, upon which is added an annual transient response of biomass growth during the growing season, following the onset of rainfall in the spring. By subtracting the constant fallow season background value from the total observed NDVI values, we determine what we will refer to as the differential NDVI ($\delta$NDVI), an example of which is shown, with an application of our algorithm, in Figure 5. It is this latter component – the annually transient response of biomass growth, or its proxy: the $\delta$NDVI – that our algorithm has been designed to predict based on daily precipitation alone.

2.3.5. $\delta$NDVI as a measure of ANPP

The connection between vegetation indices measured from satellite platforms and in situ measurements of biomass production in the field is illustrated in Figure 6, where annual estimates of ANPP are plotted against seasonally integrated values of $\delta$NDVI following a procedure similar to that of Moran et al. (2014) who used EVI and ANPP values from an
area near our study site. Our target dates for integrating δNDVI values were from the typical start of the growing season, April 1 (day 91), to the time when biomass is usually harvested in the CPER, typically by the end of the first week in August (day 219). Due to the discrete compositing periods for which MODIS data were available, we used

Figure 5. (a) ‘Fit’ of observed δNDVI data (green line) from a short grass steppe site by predicted data (blue line) from a rainfall-to-NDVI second-order linear reservoir model for calendar year 2007. Black spikes are daily rainfall (P) values referred to the right-hand axis from which the predicted δNDVI shown in the figure is generated. (b) Ordinary linear least-squares regression of predicted δNDVI from daily rainfall versus observed δNDVI for 23 composite samples during calendar year 2007. Here, the predicted δNDVI intercept \( b = 0.0096 \) is virtually negligible. (A unit slope line (dashed) is shown for comparison.)

Figure 6. Seasonally integrated δNDVI as a proxy for annual estimates of ANPP from a field site in our study area of the CPER. The x-axis shows values of δNDVI seasonally summed over eight 16 day compositing periods from 1 April to early August each year. The y-axis shows ANPP from spatial averages of harvested vegetation clippings from an approximately 1 ha area near (approximately 0.33 km) our study site.
compositing periods from the start-of-period (SOP) day 97 to SOP day 209. This resulted in eight 16 day compositing periods starting on day 97 (April 7), with the last day of the eighth period on day 224 (August 12), for a total of 128 days. The seasonally-integrated metric is the simple sum of \( \delta \text{NDVI} \) computed from \( \sum_{N_{\text{SOS}}}^{N_{\text{Peak}}} \delta \text{NDVI} \), where \( N_{\text{SOS}} = 97 \) and \( N_{\text{Peak}} = 209 \). The latter is abbreviated as \( \sum_{S_{\text{P}}} \delta \text{NDVI} \) along the abscissa of Figure 6, where the subscript ‘S_{\text{P}}’ stands for a seasonal summation of eight values for the compositing periods from the start of season (SOS) to the season’s peak growth, and the subscript index ‘j’ denotes the respective members of the seasonal compositing periods.

The ordinate of Figure 6 shows spatially averaged annual estimates of \textit{in situ} biomass aboveground net primary production (ANPP) from an approximately 1 ha representative field site (40.8471°N, 104.7095°W) near the centre of Section 7 at the CPER and close (approximately 0.33 km) to the centre point (40.8441°N, 104.7097°W) of the referenced 250 × 250 m NDVI pixel (ARS-HQ NDVI S01). Clipping (with subsequent lab drying and weighing) was done inside 1 m\(^2\) moveable grazing cages from 2000 to 2011 (with 12–20 plots clipped annually, and cages moved to new random locations each year). The annual ground-based estimate of peak standing crop (ANPP), measured at the end of July or early August, represents the sum of aboveground biomass of all grasses, forbs, and subshrubs (expressed in Figure 6 in units of \( \text{g m}^{-2} \text{ year}^{-1} \)), and has been shown to provide a reasonable estimate of ANPP (Milchunas and Lauenroth 1992).

3. Model algorithm

3.1. The linear reservoir model

We define the grassland ‘system’ as the combined atmosphere, earth surface, soil, plant root, stem, and foliage composite. Our linear reservoir representation is a phenomenological model in which daily rainfall events – as input ‘impulses’ to the system – are transformed to emulate the phenology of water limited, semi-arid grasslands. We use \( \delta \text{NDVI} \) as the proxy for characterizing time-dependent aboveground vegetative biomass. At any instant of time, a member – or ‘stage’ – of our linear reservoir system is assumed to have an instantaneous volume \( V(t) \), which will be different in the different stages of a cascaded series of reservoirs such as Figure 2(a). A stage’s behaviour accords with the conservation condition that the temporal outflow of water \textit{from} a vessel \( Q_{\text{out}}(t) \) is proportional to the change in volume of water stored \textit{in} the vessel \( -dV/dt \), so that

\[
Q_{\text{out}}(t) = -dV/dt. \tag{1}
\]

The negative sign accords with the condition that a decrease of \( V \) in the vessel is related to a positive outflow. A further condition on our model is that, for each of the vessels (or leaky buckets) of the type illustrated in Figure 2(a), the outflow \( Q_{\text{out}}(t) \) of each respective stage is, in turn, proportional to the volume of water present in the stage, as expressed by:

\[
Q_{\text{out}}(t) \propto V(t). \tag{2}
\]

Such a condition would apply to each of the right circular cylinders in Figure 2(a), where for a lossy orifice, Darcy-like flow would be governed by the height of the water column in the vessel, and the height of the water column in turn governs the total volume of the
contained water. Combining the right sides of (1) and (2), using \( C \) as the constant of proportionality (to be determined), we obtain upon rearranging signs

\[
dV/V = -Cd_t. \tag{3}
\]

The left hand side integrates to the natural logarithm as in the following, and the right hand side is a time term:

\[
\ln(V(t)) = -Ct + D, \tag{4}
\]

where \( D \) is a constant of integration. In order for both sides of the latter equation to be dimensionless, the constant of proportionality must have the dimensions of inverse time, so that we define the constant \( C = 1/T \), where \( T \) is in time units. Substituting \( 1/T \) for \( C \) in Equation (4), and raising both sides of Equation (4) to a power of \( e \), we have

\[
V(t) = e^{-t/T}e^D. \tag{5}
\]

Relation (5) can be used to represent the behaviour of the first stage of Figure 2(a), if we set the second exponential term to \( V_0 = \exp(D) \), which can be used to represent the reference volume of water in the reservoir at time \( t = 0 \). The result is

\[
V(t) = V_0e^{-t/T}. \tag{6}
\]

Relation (6) shows that an exponential decay with a system’s characteristic time \( T \) is a fundamental feature of such models. Hence, \( T \) is referred to as a ‘characteristic time’ or ‘time constant’ of the respective process.

### 3.2. The impulse response function

There are two ways to think about \( V_0 \) in Equation (6); one is to think of the term as the volume of water stored in a reservoir waiting to be released at \( t = 0 \). Another way is to consider \( V_0 \) in Relation (6) as a sudden ‘impulse’ of water – an instantaneous volume of water \( \Delta V \) – added to the reservoir at \( t = 0 \). For a single impulse event \( \Delta V \) occurring at \( t = 0 \)

\[
V(t) = \Delta V e^{-t/T}, \tag{7}
\]

where \( V(t) \) is the instantaneous volume of water remaining in the reservoir following the singular ‘event’ \( \Delta V \). In the approach we adopt here, we want a sequence of these impulses \( \Delta V_k \), at different times \( t_k \), \( 1 \leq k \leq M \), to represent a sequence of \( M \) rainfall events over a multiyear period of record. In this sense, \( k \) is an ordinal number corresponding to the rain day in the overall period of record; \( t_k \) is the actual time date. As an example of how these parameters are used, in our observed daily rainfall time series, 2.03 mm of rain was recorded on 8 May 2000; this corresponds by happenstance to the \( k = 493 \) member of our time series, and a sampling time of \( t_k = 2000.34932 \). For our period of record from 1 January 1999 to the last rain day of 2014 (26 November 2014, or \( t_k = 2014.90274 \)), we have \( M = 5805 \) daily recorded values, many of which are zero.
If instead of occurring at \( t = 0 \), an impulse \( \Delta V_k \) occurs at \( t = t_k \), then at some time \( t \geq t_k \), the volume of water in the reservoir due to this event would be

\[
V(t) = \Delta V_k e^{-(t-t_k)/T}, \text{ for } t \geq t_k,
\]

(8)

where the subscript \( k \) on \( \Delta V_k \) simply denotes the reservoir’s ‘response’ to a singular impulse occurring at \( t_k \). The outflow from the reservoir, according to Equation (1), is given by \( Q_{\text{out}}(t) = -dV/dt \), when applied to Equation (8), leads to

\[
Q_{\text{out}}(t) = \frac{\Delta V_k}{T} e^{-(t-t_k)/T}, \text{ for } t \geq t_k.
\]

(9)

Relation (9) is known in the literature as the ‘impulse response function’ of a linear reservoir. For a sequence of impulse events (such as rain days) at arbitrary intervals of time, the response to each individual event is superposed with all the others according to

\[
Q_{\text{out}}(t) = \sum_{k=M}^{k=1} \left( \frac{\Delta V_k}{T} e^{-(t-t_k)/T} \right), \text{ for } t \geq t_k.
\]

(10)

If we think of \( \Delta V_k \) as an increment of flow at \( t_k \) over a time step \( \Delta t \), we can substitute \( \Delta V_k = Q_{\text{in}}(t_k) \Delta t \) into Equation (10), which in integral form becomes

\[
Q_{\text{out}}(t) = \int_{-\infty}^{t} \left( \frac{Q_{\text{in}}(t')}{T} e^{-(t-t')/T} \right) dt', \text{ for } t' \leq t.
\]

(11)

In applied math and the physical sciences, Equation (11) is known as a convolution integral; basically a transmutation or filtering operation that linearly transforms the input function \( Q_{\text{in}}(t') \) at time \( t' \) into a modified output function \( Q_{\text{out}}(t) \) at time \( t \). Usually, applying physically-based filter-like operations to real systems results in a time delay or lag in the output. It is the intrinsic delay of a multistage linear reservoir that is an attractive model for representing the connection between a rain event and the associated production of vegetation.

Two leaky buckets, with lossy orifices, are shown in Figure 2(a) representing the two stages of a second order \((n = 2)\) cascaded linear reservoir. We introduce a ‘slug’ or ‘impulse’ of water \( \Delta V \) into the first stage – simulating a daily ‘pulse’ of rainfall – instantaneously filling the container to an initial volume \( V_0 = \Delta V \). We should expect the volume of contained water to exponentially decrease according to Equation (6), which with \( \Delta V = V_0 \) becomes \( V(t) = \Delta V \exp(-t/T) \).

For purposes of illustration in Figure 2, we use a factor of \( T = 5 \) days for the characteristic time of the system. The composite graph in Figure 2(b) illustrates the flux \( Q(t) = \pm dV/dt \) of water at various points and times of the two-stage system in Figure 2(a). In the graph of Figure 2(b), the initial quasi-impulsive flux from an elemental volume of water \( \Delta V \) into the system is represented by a narrow, grey-shaded rectangular pulse of flow at the origin with its amplitude referred to the vertical scale at the far left of the graph. The units used in Figure 2(b) are in generalized form in terms of length \((L)\) and inverse time \((T^{-1})\). In our case, time is scaled by day. For purposes of this example, the input pulse is shown to have unit amplitude, so the axis on the far left of Figure 2(b) shows 1 unit of flow \((L T^{-1})\), crossing a unit area, lasting 1 day. For daily rainfall, we
might assume an impulsive flux of 1 mm day$^{-1}$, which is assumed to vertically infiltrate a unit area of the earth’s surface, e.g. 1 m$^2$, with a one-day duration, for an effective volume of 1 mm of rainfall-depth per unit surface area. If such an impulse of water, having volume $\Delta V$ is dropped into the first bucket in Figure 2(a) at $t = 0$ day, then the output flow from Stage 1 [according to (9)] can be represented by

$$Q_{\text{out}}^{(1)}(t) = \frac{\Delta V}{T} \frac{\text{unit}}{\text{day}} e^{-t/T}, \text{for } t \geq 0.$$  

(12)

This is shown in Figure 2(b), with the response (12) referred to the second axis from the left in the figure. Figure 2(b) shows that the flux from the first bucket begins immediately (at $t = 0$) at its peak value of $Q_{\text{out}}^{(1)}(0) = 0.2$ flow units. This accords with relation (12), since we have 1 unit volume divided by $T = 5$ days for this example. Following its initial peak flow at $t = 0$, there is an exponential decrease of flow from Stage 1 according to Equation (12) at a rate defined by the characteristic time $T = 5$ days in the exponent.

Regarding Stage 2 in Figure 2(a), its input flow will be the negative of the outflow from Stage 1, or $Q_{\text{in}}^{(2)}(t) = -Q_{\text{out}}^{(1)}(t)$. Integrating from Equation (11), from the initial introduction of the impulse into Stage 1 at $t = 0$,

$$Q_{\text{out}}^{(2)}(t) = \int_{0}^{t} \left( \frac{\Delta V}{T^2} e^{-t'/T} e^{-(t-t')/T} \right) \text{d}t',$$

(13)

and noting that $t'$ is the variable of integration and $t$ is a constant, we have

$$Q_{\text{out}}^{(2)}(t) = \left( \frac{t}{T} \right) \frac{\Delta V}{T} e^{-t/T},$$

(14)

which is the appropriate expression for the time dependent flow from Stage 2. As constrained by the leading term $(t/T)$, one would not expect any outflow at $t = 0$ since the volume of water in the second vessel is yet to accumulate, so that as expected intuitively and predicted analytically by (14), $Q_{\text{out}}^{(2)}$ starts at zero at $t = 0$, then builds to a maximum as more water is added to Stage 2 from Stage 1, whereupon $Q_{\text{out}}^{(2)}$ diminishes to zero as the level of water in Stage 2 is exhausted.

Broadly speaking, we might conceptually relate such lagged processes to the rainfall-biomass production process: Stage 1 emulates the entry stage of water into the soil and its redistribution as soil moisture; Stage 2 emulates the uptake of water by plant roots, short-term biomass development, followed by the subsequent transpiration of water through foliage (although the detailed infiltration process, accumulation and interaction with plant physiology, is much more complicated). A general expression for $n$ stages of linear reservoirs cascaded in series is

$$Q_{\text{out}}^{(n)}(t) = \left( \frac{t}{T} \right)^{n-1} \frac{\Delta V}{(n-1)! T} e^{-t/T},$$

(15)
3.3. **Superposition of the effects of rainfall pulse ‘events’: convolution**

3.3.1. **Synthesizing the response to a series of rain events**

Now, instead of a single impulse having volume $\Delta V$ at $t = 0$, we allow for an arbitrary series of impulses – daily rain events – each characterized by a specific volume $\Delta V_k$ (in units of mm) occurring at a sequence of times $t_k$. A rain event on any day $t_k$ (on or before day $t$) will add to the cumulative effect of all such events recorded on day $t$, so that summing over all rain events, we have the numerical convolution

$$Q^{(n)}_{\text{out}}(t) = \sum_k \left( \frac{t - t_k}{T} \right)^{n-1} \frac{\Delta V_k e^{-(t-t_k)/T}}{(n-1)!T} \text{where } t_k \leq t,$$  

(16)

where $n$ is the number of stages in the cascaded series of reservoirs. Relation (16) is a conservation condition that we will use to represent the migration of water through the soil–plant system from a sequence of discrete daily rainfall events $\Delta V_k$, $1 \leq k \leq M$.

3.3.2. **The biomass response**

A common postulate of ecohydrology is that water-limited vegetation in semi-arid grasslands tends to use all available water. We view the overall reservoir response $Q^{(n)}_{\text{out}}(t)$ in Equation (16) as the responsible agent for biomass development represented in our model by its proxy $\delta \text{NDVI}(t)$, and assume that the two are simply related by a linear scaling factor: $C_{\text{SF}}$. Thus, using $Q^{(n)}_{\text{out}}(t)$ from Equation (16), along with the scaling factor $C_{\text{SF}}$, one might compute a predicted value of $\delta \text{NDVI}$ according to

$$\delta \text{NDVI}_{\text{pred}}(t) = C_{\text{SF}} Q^{(n)}_{\text{out}}(t).$$  

(17)

At discrete times of observation $t = t_i$, $1 \leq i \leq N$, the latter linearly transforms Equation (16) to provide our basic $n$th-order linear reservoir model for the biomass response at time $t_i$:

$$\delta \text{NDVI}_{\text{pred}}^{(i)} = C_{\text{SF}} \sum_k \left( \frac{t_i - t_k}{T} \right)^{n-1} \frac{\Delta V_k e^{-(t_i-t_k)/T}}{(n-1)!T},$$  

(18)

where $\Delta V_k$ is the daily rainfall at time $t_k$, and $t_k \leq t_i$.

3.3.3. **Estimating the model parameters**

While the model is linear, the inverse problem of finding the optimal model parameters is not, so that at this stage of our analysis, the optimization is basically a structured search, stepping through small increments in model values over all plausible combinations of the parameters. As we do so, a fundamental constraint on our predicted $\delta \text{NDVI}$ is that the aggregated or integrated value of biomass productivity for the respective time series is constrained to equal its counterpart for the observed $\delta \text{NDVI}$. In other words, we normalize the predicted (pred) to the observed (obs) time series such that the respective area under the predicted curve (the aggregated seasonal biomass productivity) is the same as the.
observed curve. This condition determines the estimated (estim) value of $C_{SF}$ in Equation (18), which is computed from

$$C_{estim}^{SF} = \frac{\sum_{i=1}^{N_d} (\delta \text{NDVI}^{(i)}_{obs})}{\sum_{i=1}^{N_d} (\text{Unscaled} \delta \text{NDVI}^{(i)}_{pred})},$$

(19)

where $N_d$ is the number of data that are used for the computation. The number is the same for the observed and predicted samples, and depends on the application. For some contexts, $N_d$ might be the number of compositing periods in a year, for others, it might be the number of samples over a multiyear period of record.

Thus, beginning with a specific set of pre-assigned trial model parameters $n$ and $T$, we compute the model (18) under the condition imposed on $C_{estim}^{SF}$ by Equation (19). We then compute an appropriate misfit error, $e(i) = \delta \text{NDVI}^{(i)}_{pred} - \delta \text{NDVI}^{(i)}_{obs}$, dependent on the specific set of pre-assigned trial model parameters: $n$ and $T$. Our principal measures of fit will be the root-mean-square error (RMSE) and the ‘normalized’ RMSE (NRMSE), the latter defined as the ratio of the RMSE to the difference between the maximum and minimum values of the observed $\delta \text{NDVI}$. Other considerations are the Pearson correlation coefficient, $r$, between the predicted and observed time series, the coefficient of determination $R^2$, and the transfer coefficient – the slope of the regression line – between predicted and observed $\delta \text{NDVI}$ data.

### 3.3.4. Refining the linear reservoir model with time-dependent parameters

Preliminary testing of the model suggested that our response function (18) needed to be modulated during the course of the growing season. For a reservoir system of specific order $n$, it was apparent for a set of parameters ($C_{SF}$ and $T$) that a model fitting observations early in the season did not fit the observations late in the season, or vice versa. We, therefore, introduced a scheme whereby an amplitude scaling factor (an efficiency coefficient $C_{eff}$) and the characteristic time ($T$) could vary over the growing season according to the day-of-year (DOY). The volume element term in Equation (18) is accordingly defined by $\Delta V_k = C_{eff}|_{DOY_k} P_k$, where $P_k$ is the observed daily rainfall on day $t_k$ and $C_{eff}|_{DOY_k}$ is the seasonal efficiency coefficient associated with the DOY of the particular day $t_k$. To be clear, the DOY variable, as used here, can vary from 1 to 365, whereas the variable $t_k$ tracks the respective rainfall day, and can vary in principle over the complete period of record of precipitation. Similarly, we assume that the characteristic time constant $T$ varies with the DOY according to $T = T|_{DOY_k}$. It appears to be sufficient for our purposes to allow both the characteristic time and the efficiency coefficient to vary linearly over the same seasonal interval. In the shortgrass steppe, the growing season runs from 1 April (DOY = 91) to 31 October (DOY = 304). Operationally, the algorithm runs 365 days a year (for all years of record), so starting 1 January and continuing to the SOS (day 91), the efficiency coefficient and characteristic delay time are constant at specific values: $C_{eff}|_{SOS}$ and $T|_{SOS}$. Then, beginning at SOS (day 91), the two parameters are allowed to linearly transition – increasing or decreasing – to another set of values until the end of season (EOS, day 304): $C_{eff}|_{EOS}$ and $T|_{EOS}$, respectively. From the EOS to 31 December, the parameters are again held constant and then reset on 1 January. To avoid bias, we allow either parameter – $T$ or $C_{eff}$ – to increase or decrease over the season. No loss of generality is incurred if we set $C_{eff}|_{SOS} = 1$ and solve for a variable $C_{eff}|_{EOS}$. Thus,
only $C_{\text{eff}}|_{\text{EOS}}$, $T|_{\text{SOS}}$, and $T|_{\text{EOS}}$, along with $n$, the number of stages, will be used and reported as model variables.

4. Application to example data

4.1. Modelling annual data

Figure 5 shows the application of the linear reservoir model to MOD13Q1 16 day 250 m NDVI data from our shortgrass study site. The figure compares the 2007 annual phenology from $\delta$NDVI$_{\text{obs}}$ data with the inferred phenology from $\delta$NDVI$_{\text{pred}}$ data estimated from the refined version of Equation (18). The optimal model and quality of fit parameters are provided in Table 1.

The time series in Figure 5(a) demonstrates that the predicted responses of the linear reservoir model track observed $\delta$NDVI data well, both in amplitude and phase. Particularly clear in Figure 5(a) is that intraseasonal modulations of rainfall over the growing season relate directly to modulations of NDVI values. This result has important implications for inferring rainfall-to-NDVI relations at the local, pixel scale, and underscores the advantage of using compact, quick-responding functionals such as relation (15) to represent the response of a plant–soil system to a rainfall pulse. Figure 5(b) shows a scatter plot of $\delta$NDVI$_{\text{pred}}$ versus $\delta$NDVI$_{\text{obs}}$ for the 23 MOD13Q1 compositing periods in 2007. The resulting model parameters in Table 1 demonstrate decreases in the efficiency coefficient $C_{\text{eff}}(\text{DOY})$ during the growing season to 40% of its value at SOS, implying that more precipitation is needed later in the growing season to produce the same biomass as that produced with less precipitation earlier in the growing season. This finding is consistent with prior research demonstrating that early season production by $C_3$ grasses is more responsive to a given amount of precipitation than mid-summer growth of $C_4$ shortgrasses (Derner et al. 2008; Milchunas, Forwood, and Lauenroth 1994; Milchunas et al. 2008). Furthermore, there is reduced evaporative loss of water during the early, cool-season growth period (April and May) compared with increasing evaporative losses as temperatures increase over the growing season (Taylor et al. 2014). Table 1 also shows that the characteristic response time decreases over the growing season (from 15 days at SOS to 5 days at EOS), suggesting that late season vegetation greenup is quicker to respond to rainfall events than earlier in the season. This is attributable to warmer temperatures later in the growing season tending to induce more rapid growth, and to physiological traits associated with the $C_3$ versus $C_4$ species affecting temporal responses of vegetation to rainfall events. For example, on a daily basis, $C_3$ grasses respond more slowly to precipitation inputs than $C_4$ plants due to their greater height (requiring water movement through stems) and their growth occurring at cooler temperatures. The $C_4$ shortgrasses are particularly well-adapted to rapidly utilize precipitation events at warm to hot temperatures, with drought-stressed $B. gracilis$ able to respond to small (e.g. 5 mm)
events in terms of leaf water potential and conductance within 24 hours (Sala and Lauenroth 1982). Thus, the characteristic response time, and the way it is modulated over the growing season, provides a valuable means to quantify basic growth properties of this grassland ecosystem.

The results obtained from first-order ordinary least-squares (OLS) regressions of predicted versus observed data are summarized in Table 1 as a measure of the success of our algorithm, implying a better than adequate fit of the modelled results to observations. The NRMSE is 9.82%, and the variance of the residuals (or misfit error) is approximately 6% that of the variance of the phenology itself.

Time constants reported in Table 1 accord with observations of Li, Zhao, and Liu (2013) for a low rainfall grassland area in northwestern China. Using 17 carefully culled gauge-recorded rainfall events for the June to August growing seasons from 2002 to 2010, and MODIS NDVI data at 250 m resolution, an 8- to 16-day lag time between discrete rainfall events and NDVI responses was observed (Li, Zhao, and Liu 2013). Our algorithm provides an expeditious and powerful means to integrate a series of daily rain event impulses to provide a composite time-dependent, continuously evolving dynamical representation of biomass production over a growing season.

### 4.2. Modelling multiyear data

There are several ways in which our algorithm might be applied to multiyear time series, depending on the information one wants to extract from the data. Two applications are described here. The first uses a set of model parameters that vary interannually, which is to say that a different set of model parameters is computed for each year using the daily recorded rainfall for the current and prior years. The second application we describe determines a single global set of four model parameters simultaneously optimized to best fit the observed δNDVI for all years.

#### 4.2.1. Interannually variable model parameters

The fit to 14 years (2000–2013) of observed data by predicted values using interannually variable model parameters is shown in Figure 7. The 319 data values are partitioned by year, then each year’s data are fit with the linear reservoir model specific to that respective year, where each annual model has four model parameters, for an overall total of 56 model parameters. Four different coefficients are determined each year by fitting the annual model to the current year’s observed δNDVI values, using the daily recorded rain events.

![Figure 7](image.png)

Figure 7. Fit to 14 years of observed data by predicted values using interannually variable model coefficients; three different coefficients are determined each year by fitting the model to the current observed δNDVI values, using the daily recorded rain events from 2000 to 2013.
from the current and prior years. The scatter plot in Figure 8 shows the regression of $\Delta $NDVI$^\text{pred}$ on $\Delta $NDVI$^\text{obs}$ based on 319 MOD13Q1 compositing periods for the period of record from day 49 in February 2000 to day 365 in December 2013. (A unit slope line (dashed) is shown for comparison.)

Table 2. Summary of 2000–2013 interannually variable model parameters with quality of fit.

<table>
<thead>
<tr>
<th>Median of annual model parameter</th>
<th>Model fit (RMSE = 0.0352; NRMSE = 7.02%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$ (see Note)</td>
<td>2</td>
</tr>
<tr>
<td>$C_{\text{eff}}$(EOS)</td>
<td>0.3</td>
</tr>
<tr>
<td>$T$(SOS)</td>
<td>22 days</td>
</tr>
<tr>
<td>$T$(EOS)</td>
<td>9 days</td>
</tr>
</tbody>
</table>

Note: $n$ is the number of stages.

Figure 8. Scatter plot and ordinary linear least-squares regression of $\Delta $NDVI$^\text{pred}$ on $\Delta $NDVI$^\text{obs}$ based on 319 MOD13Q1 compositing periods for the period of record from day 49 in February 2000 to day 365 in December 2013. (A unit slope line (dashed) is shown for comparison.)

from the current and prior years. The scatter plot in Figure 8 shows the regression of $\Delta $NDVI$^\text{pred}$ on $\Delta $NDVI$^\text{obs}$ for 319 compositing periods from day 49 in February 2000 to day 365 in December 2013. Key elements of the analysis are summarized in Table 2, which shows that the RMSE of the time series in Figure 7 is 0.0352, and the NRMSE is 7.02%. Table 2 also contains medians of the 14 optimized values from each of the annualized four model parameters, along with the results of least squares regressions using the scatter plot of data shown in Figure 8.

Metrics comparing the results of the modelling algorithm to observations using ordinary least squares regression show a high degree of correlation between predicted and observed data, reflecting the observations made previously in Section 4.1. According to the left hand column in Table 2:

(a) Reservoirs of order $n = 2$ are preferred, which is consistent with a finite time delay between the occurrence of a rain pulse and its associated effect on peak greenup;
(b) The efficiency coefficient for the years analysed here decreases over the growing season;
(c) Typical characteristic response times decrease with progression of the growing season.

Although actual values of these model parameters differ markedly from year to year, these latter patterns of rainfall-biomass growth dynamics derived from the linear reservoir model persist. We find that for 50% of the years, early season time constants are in the range of 17–28 days (nominally 2–4 weeks), and by late season we find that for 50% of the years, the time constants decrease by a factor of 0.30–0.60 (or to 1–2 weeks). A visual inspection of the time series in Figure 7 and the accordant high correlation factors from regressions using the data in Figure 8 demonstrate that the linear reservoir algorithm with interannually variable model parameters is able to directly translate, with a high degree of certainty, intraseasonal modulations of rainfall over the course of the growing season to corresponding modulations of biomass productivity.

4.2.2. Global estimates of model parameters

As an alternative to the procedure in the previous section (4.2.1), where four model parameters are adjusted year-by-year to individually optimize the fit to 14 years of data (a total of 56 model parameters), there may be applications where one wants to globally fit a generic four parameter model to a number of contiguous years simultaneously. In this section, we apply such a version of our algorithm to the same 14 years of observed data (2000–2013). Our objective is not to showcase the ability of the model to readjust its four parameters to track both interannual and intraseasonal variations (as we did in Section 4.2.1), but to illustrate a particular application of the model. In this example, we ask whether a specific globally-determined, four parameter response function, derived from all 14 years of data, can adequately fit the phenologies of individual years using the respective year’s daily rainfall; and if not, then why not?

The results of our experiment are shown with the time series in Figure 9. The first step is to globally fit an nth-order cascaded linear model to the 329 observed 16 day composites shown in Figure 9, using the associated 5805 daily forcing events (rainfall values from 1999 to 2013). The optimal model, based on a minimum RMSE for all 329 predictions and observations, is a two-stage reservoir, where the characteristic storage time decreases from 22 days at SOS to 14 days at EOS, and the efficiency coefficient $C_{\text{eff}}$

Figure 9. Fit to 14 years of observed data by predicted values using a single globally determined two-stage response function to the daily rainfall during individual years 2000–2013. Interannual differences between predicted and observed data (e.g. 2007) are inferred to reflect interannual differences in site conditions (see the text).
decreases from unity at SOS to 0.35 (35%) at EOS. The global fit of this model to the 14 years of data has an optimized RMSE = 0.045, somewhat higher than the RMSE = 0.036 for 2007 in Figure 5, and the RMSE = 0.035 for the 2000–2013 time series in Figure 7. While a larger misfit is not unexpected when using fewer model parameters, there are systematic patterns in the point-by-point misfits in Figure 9, which are either absent or subdued in the year-by-year fit shown in Figure 7. In Figure 9, for example, data predicted by the global model for individual years do not fit the 2007 data as well as the individual year analysis in Figure 5 does. The years 2001 and 2003 show similar patterns of misfits. Nevertheless, the general patterns of seasonal behaviour of the model parameters are similar to those for the interannually variable model parameters, and intraseasonal tracking of the observed $\delta$NDVI by the globally predicted $\delta$NDVI is quite good according to the time series in Figure 9 and the summary statistics in Table 3.

Whether the tracking of the multiyear fluctuations of observed phenology in Figure 9 by the global model is adequate depends on its purpose. One application is to use such a model to interpolate the pattern of biomass production at sites during years when no NDVI have been recorded, or when NDVI during a critical season is missing or contaminated. Alternatively, a model based on globally determined parameters from a number of years of data might be used to detect intraannual departures of observed biomass from the expected (predicted) biomass response. We apply the latter procedure in the next section.

4.2.3. Identifying intraannual departures of observed biomass from the expected (globally predicted) biomass response

In Section 4.2.2, we pointed out that there are systematic patterns in the point-by-point misfits in Figure 9. Here, we apply our global modelling procedure in an attempt to identify and analyse the cause of one particular type of misfit: notable underpredictions of peak $\delta$NDVI for the first major growth pulse in certain years (e.g. 2001, 2003, 2004, 2007, and 2013). We begin by defining the deviation between observed and predicted peaks as

$$\Delta(\text{peak } \delta \text{NDVI})_j = (\text{peak observed } \delta \text{NDVI})_j - (\text{peak predicted } \delta \text{NDVI})_j$$  \hspace{1cm} (20)

during the earliest part (if there are two or more vegetation peaks) of each growing season, $j = 1$ to $N$, where $N = 14$ years. Accordingly, $\Delta(\text{peak } \delta \text{NDVI})_j$ for the year 2007 in the time series of Figure 9 has a large positive value. We next examine whether the magnitude of $\Delta(\text{peak } \delta \text{NDVI})_j$ for each year is related to an expected ecohydrological parameter. Our choice for the latter is the mean value of predicted $\delta$NDVI during the growing season (gs), a parameter denoted here by $\bar{\delta \text{NDVI}_{\text{pred}}}(j)$ for each of the 14 years. To be consistent with

| Table 3. Summary of the global fit of 2000–2013 NDVI data by a single four parameter linear reservoir model. |
|-------------------------------------------------|-------------------------------------------------|
| Global model parameter | Model fit (RMSE = 0.0445; NRMSE = 8.86%) |
| $n$ (see Note) | 2 | $r$ | 0.8950 |
| $C_{\text{eff}}$(EOS) | 0.3 | $R^2$ | 0.8010 |
| $T$(SOS) | 25 days | Slope | 0.8406 |
| $T$(EOS) | 13 days | Intercept | 0.0166 |

Note: $n$ is the number of stages.
the compositing periods of the MOD13Q1 database, $\delta \text{NDVI}_{\text{pred}}(j)$ is computed for each year using the predicted $\delta \text{NDVI}$ from day 97 to day 304 (the end of the 16 day compositing period beginning on day 289), a mean of 13 $\delta \text{NDVI}$ values per year. We use predicted values of $\delta \text{NDVI}$ rather than observed values, because the former is basically a filtered form of intraseasonal rainfall, and through the global modelling analysis in Section 4.2.2, the derived filter function (the impulse response function) best represents the overall (i.e. ‘global’) relation between intraseasonal patterns of rainfall (and associated climate variables) and intraseasonal patterns of biomass production. Our plan is to explore how these anomalous peak values related to the biomass-rainfall conditions predicted by the model. Dry years versus wet years will correspond to differences between the respective annual value of $\delta \text{NDVI}_{\text{pred}}(j)$ and its long-term mean value $[\text{Mean}(\delta \text{NDVI})_{\text{pred}}]$, with the difference denoted by:

$$\Delta(\delta \text{NDVI}_{\text{pred}}(j)) = (\delta \text{NDVI}_{\text{pred}}(j)) - [\text{Mean}(\delta \text{NDVI})_{\text{pred}}].$$  \hspace{1cm} (21)

Our approach is straightforward. We regress the annual values of $\Delta(\text{peak} \delta \text{NDVI})_j$ versus $\Delta(\delta \text{NDVI}_{\text{pred}}(j))$ for each of the 14 years of record: 2000–2013; results are shown in Figure 10. The coefficient of determination $R^2 = 0.0018$ ($r = 0.0430$) is poor, but a visual inspection of the pattern of misfit between the actual time series in Figure 9 suggests a possible connection between the anomalous peak values in some years and the general level of biomass productivity in the prior year. This ‘legacy’ effect (Sala et al. 2012) has been observed from analyses of ANPP in drylands regarding the influence of prior-year growing season conditions on current-year productivity (Oesterheld et al. 2001; Richardson et al. 2013; Sala et al. 2012).

Figure 10. Regressions of underpredicted peak $\delta \text{NDVI}$ for the first major growth phase in respective years versus the annual departure of the mean value of predicted $\delta \text{NDVI}$ during the growing season. Blue symbols denote data pairs from the same year ($R^2 = 0.0018$; $r = 0.043$); red symbols denote data pairs where the peak $\delta \text{NDVI}$ is lagged by one year ($R^2 = 0.341$; $r = 0.584$).
Thus, in terms of our immediate concern to explain how anomalous $\Delta(\text{peak}\, \delta\text{NDVI})_j$ values relate to the biomass-rainfall predicted conditions, it seems to be reasonable to lag $\Delta(\text{peak}\, \delta\text{NDVI})_j$ *versus* $\Delta(\bar{\delta}\text{NDVI}^{\text{gs}}_{\text{pred}}(j))$, so that $\Delta(\text{peak}\, \delta\text{NDVI})_j+1$ is plotted and correlated against the earlier $\Delta(\bar{\delta}\text{NDVI}^{\text{gs}}_{\text{pred}}(j))$ for *j* indices corresponding to years 2000–2012. The results are shown in the second scatter plot in Figure 10, where the $N_{\text{year}} = 13$ lagged pairs begin to take on a definite trend, and the coefficient of determination increases to $R^2 = 0.3410$ (or a negative correlation coefficient of $r = -0.5840$). A standard statistical analysis (e.g. Hald 1952) of the regression results implies that there is better than a 98% probability that the data are correlated, and better than a 95% probability that the slope is negative and lies between $-1.65$ and $-0.07$. It is 50% probable that the true slope lies between $-1.12$ and $-0.61$. We interpret this as a significant pattern such that $\Delta(\text{peak}\, \delta\text{NDVI})_j$ is larger for those years when the prior year mean seasonal biomass productivity (as represented by $\bar{\delta}\text{NDVI}^{\text{gs}}_{\text{pred}}(j-1)$) was lower (that is to say, $\Delta(\bar{\delta}\text{NDVI}^{\text{gs}}_{\text{pred}}(j))$ is more negative), implying drier conditions.

5. Discussion

Understanding the relationship between precipitation measured across multiple temporal scales (intraannual, annual, and multiannual) and resulting patterns of plant productivity has been a major research focus in dryland ecosystems (Knapp and Smith 2001; Moran et al. 2014; Oesterheld et al. 2001; Sala et al. 2012; Wilcox et al. 2015). The amount as well as the timing of individual precipitation events can strongly influence ecosystem dynamics, including aboveground productivity (Cherwin and Knapp 2012; Fay et al. 2003; Heisler-White et al. 2009; Wilcox et al. 2015). Furthermore, projections of future climate conditions indicate changes in the timing and size of precipitation events (cf. National Climate Assessment 2014). Our model provides a simple, yet powerful means of integrating precipitation event size and timing using as few as four parameters to describe multiple-year vegetation dynamics in shortgrass steppe.

While the strong congruence between the observed and the precipitation-predicted biomass phenology (represented by the observed and the predicted $\delta$NDVI) in the shortgrass steppe is noteworthy, we do not imply a direct cause and effect. Instead, we infer that seasonal rainfall is, by itself, a proxy for a host of climate-related phenomena (e.g. temperature, solar insolation, humidity, soil moisture, etc.,). However, using rainfall as an independent variate (as we have done here) has allowed workers for years to draw upon parameters like the rain use efficiency (RUE) factor (Le Houerou 1984), or other linear metrics, to assess ecological function of water-limited grasslands. We recognize that our algorithm conjoins many of the details of the hydrology, soil properties, and plant physiology that are essential to understanding the fundamentals of plant ecology (Farrar, Nicholson, and Lare 1994; Jenerette, Scott, and Huete 2010; Nicholson and Farrar 1994; Richardson et al. 2013). Our model merges the behaviours of many of the specific interactive elements important to ecohydrology into a collective response of grasslands to intraseasonal modulations of rainfall. The advantage of such an elementary rainfall-based predictive model for some applications is that it obviates the need for detailed, site specific information on soil moisture dynamics, plant species, and the local hydroclimate, which, in many cases, is not available or can be difficult and expensive to obtain (Hess, Stephens, and Thomas 1996). Due to the basic and readily available nature of the data required by our algorithm, its modelling approach could be applied elsewhere, which provides a valuable means to
compare the growth dynamics of semi-arid grassland ecosystems associated with a broader range of climates and precipitation event-size distributions.

Our application of the model to the shortgrass steppe has yielded two important contributions to the dynamics of this water-limited grassland. First, there is a strong congruence of observed NDVI values and those predicted by the model from current-year daily precipitation events. Second, there are notable differences between peak predicted and peak observed NDVI, which are apparent in the early growing season following a prior drought year (e.g. 2001, 2003, 2004, 2007, 2009, and 2013, see Figure 9). Underprediction of peak δNDVI by our daily precipitation model was two-to three-fold larger in years following a drought compared to non-drought years. Over the 14 years of record, this pattern was recurrent for four of the five post-drought years. A larger-than-expected observed δNDVI (a model underprediction) following a drought year is opposite to the effect expected based on prior literature. Sala et al. (2012), for example, in their review of the phenomena, found that legacies – the association of current versus prior year conditions at a site – occur across all ecosystem types from deserts to mesic grasslands. However, they assert that the most common class of legacy is where the prior year precipitation positively controls a significant fraction of the current-year production; hence, it might be expected that a dry year would typically be followed by a year of low biomass production. In our case, we see the opposite effect; we find higher than predicted production in years following a drought year. We suggest that this underprediction of post-drought growth based on model predictions from precipitation (due apparently to the actual overproduction of vegetation based on the observed NDVI and available precipitation for that year) may occur due to accumulation of soil inorganic N (nitrogen) during droughts, which enhances plant N uptake, protein content, and photosynthetic rates in the first post-drought growing season (Augustine and McNaughton 2004; De Vries et al. 2012; Reichmann, Sala, and Peters 2013; White, Moore, and Craig 2004). Moreover, recent field experiments have specifically demonstrated substantial accumulation of soil inorganic N in the shortgrass steppe during drought (Evans and Burke 2013). Our model predictions are clearly consistent with the predictions of these field studies and yield the surprising conclusion that drought legacy can be positive through its effect on soil inorganic N. Consequences of this for explaining the anomalous increase in early-season observed δNDVI above precipitation-predicted values in post-drought years is that the enhanced ‘greenness’ of a 250 × 250 m MODIS pixel could reflect either increased integrated biomass production (i.e. an increase in plant density) or increased foliage greenness associated with enhanced leaf nitrogen content (i.e. an increase in plant photosynthetic productivity), or a combination of both effects. We suggest that broader application of our modelling approach to other grasslands worldwide will provide enhanced understanding of positive and negative drought legacies on grassland function.

6. Conclusions
The model described here represents precipitation events as impulses and simulates the response of a grassland ecosystem as a convolution – a linear superposition – of impulse response functions, each forced by a daily or short-term pulse of precipitation. We demonstrated that the linear reservoir model is an effective approach in a water-limited, semi-arid grassland where precipitation is sharply partitioned into a dominantly dry fallow season and a dominantly wet growing season.

While the dominant fraction of annual biomass production (more than 90% of the observed δNDVI variance) is related to daily precipitation in the current year, we suggest
that a significant fraction of the misfit is not random, but that underpredictions of observed NDVI by the model in 2001, 2003, 2004, 2007, 2009, and 2013 (see Figure 9) may – with the exception of 2004 – be due to a ‘legacy’ effect from prior drought years. The explanation for this is likely the accumulation of inorganic N in soils during droughts which becomes readily available for plant uptake in the first post-drought growing season. The result is an enhancement of plant growth (as represented by early-season observed δNDVI that is not predicted by the model based on the current seasonal precipitation alone) such that drought legacy in this grassland ecosystem can be positive due to higher soil inorganic N.

This model can assist with adaptive range management to promote sustainable rangeland management. The algorithm explains a major fraction of the production response of shortgrass steppe is terms of daily inputs of the current-year precipitation, and there is only a limited contribution from the effects of prior growing season precipitation through legacy effects. Adaptive management strategies can be enhanced by the knowledge that grassland biomass development early in the growing season is largely controlled by within-season precipitation inputs with a characteristic impulse response of approximately 21 days (nominally 3 weeks, or so), and that this response time declines significantly later in the growing season. These temporal intervals are more congruent with practical time-frames required by land managers needing to make effective adaptive rangeland management decisions impacting ecosystem goods (e.g. livestock production) and services.

Acknowledgements
We thank Dr Daniel Milchunas (Colorado State University) for providing the ground-based measurements of peak vegetation standing crop at the study site. The lead author’s (JFH) early interest in grassland research was sparked through several academic outreach programmes for university and college faculty organized by Director Teresa Mourad, Education and Diversity Programs Office for the Ecological Society of America, in collaboration with the US National Ecological Observatory Network, US Geological Survey, the US Department of Agriculture, and the National Center for Ecological Analysis and Synthesis. It is imperative that the international community maintains and continues to develop long-term ecological research sites and that the data from these areas be readily available to the widest possible number of scientists, educators and students.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
Funding for the ground-based measurements was provided by the Shortgrass Steppe Long Term Ecological Research (LTER) project (National Science Foundation Grant No. [0217631]).

References


