Ecological Modelling 286 (2014) 66–77

Contents lists available at ScienceDirect

Ecological Modelling

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

Modeling regeneration responses of big sagebrush (*Artemisia tridentata*) to abiotic conditions

Daniel R. Schlaepfer[a,*], William K. Lauenroth[a], John B. Bradford[b]

[a] University of Wyoming, Department of Botany, 1000 E. University Avenue, Laramie, WY 82071, USA
[b] US Geological Survey, Southwest Biological Science Center, 2255 N. Gemini Drive, Flagstaff, AZ 86001, USA

ARTICLE INFO

Article history:
Received 21 June 2013
Received in revised form
11 December 2013
Accepted 13 April 2014

Keywords:
Differential evolution optimization
Germination
Rangeland restoration
Sagebrush reclamation
Seedling survival
Soil water simulation

ABSTRACT

Ecosystems dominated by big sagebrush, *Artemisia tridentata* Nuttall (Asteraceae), which are the most widespread ecosystems in semiarid western North America, have been affected by land use practices and invasive species. Loss of big sagebrush and the decline of associated species, such as greater sage-grouse, are a concern to land managers and conservationists. However, big sagebrush regeneration remains difficult to achieve by restoration and reclamation efforts and there is no regeneration simulation model available. We present here the first process-based, daily time-step, simulation model to predict yearly big sagebrush regeneration including relevant germination and seedling responses to abiotic factors. We estimated values, uncertainty, and importance of 27 model parameters using a total of 1435 site-years of observation. Our model explained 74% of variability of number of years with successful regeneration at 46 sites. It also achieved 60% overall accuracy predicting yearly regeneration success/failure. Our results identify specific future research needed to improve our understanding of big sagebrush regeneration, including data at the subspecies level and improved parameter estimates for start of seed dispersal, modified wet thermal-time model of germination, and soil water potential influences. We found that relationships between big sagebrush regeneration and climate conditions were site specific, varying across the distribution of big sagebrush. This indicates that statistical models based on climate are unsuitable for understanding range-wide regeneration patterns or for assessing the potential consequences of changing climate on sagebrush regeneration and underscores the value of this process-based model. We used our model to predict potential regeneration across the range of sagebrush ecosystems in the western United States, which confirmed that seedling survival is a limiting factor, whereas germination is not. Our results also suggested that modeled regeneration suitability is necessary but not sufficient to explain sagebrush presence. We conclude that future assessment of big sagebrush responses to climate change will need to account for responses of regenerative stages using a process-based understanding, such as provided by our model.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Major human land use practices affect ecosystems dominated by big sagebrush (*Artemisia tridentata* Nuttall (Asteraceae)), which are the most widespread ecosystem type in semiarid western North America (McArthur and Plummer, 1978; West, 1983). Agriculture, livestock grazing (Alley, 1956; Davies et al., 2012; Sturges, 1993), renewable and non-renewable energy extraction (Walston et al., 2009), recreation, and urban development have removed or altered much of big sagebrush vegetation (Manier et al., 2013; Connelly et al., 2004; Welch, 2005). Invasive plants, particularly *Bromus tectorum*, and climate change are altering vegetation, hydrological conditions and wildfire regimes of these ecosystems (Bradley, 2010; Hanser et al., 2011). An important ecological consequence of these changes is that many sagebrush-obligate and sagebrush-associated species, such as the greater sage-grouse (*Centrocerus urophasianus*), have experienced large population declines over the last half century (Manier et al., 2013; Rowland et al., 2006).

Abbreviations: DE, differential evolution; GISSM, Germination and Individual Seedling Survival Model; MAP, mean annual precipitation; MAT, mean annual temperature; *mJSAT*, mean January air temperature; RMSE, root mean square error; VIF, variance inflation factor.

* Corresponding author. Tel.: +1 307 766 2484; fax: +1 307 766 2851.
E-mail addresses: dschlaep@uwyo.edu (D.R. Schlaepfer), wlauenro@uwyo.edu (W.K. Lauenroth), jbradford@usgs.gov (J.B. Bradford).

http://dx.doi.org/10.1016/j.ecolmodel.2014.04.021
0304-3800/© 2014 Elsevier B.V. All rights reserved.

Consequently, land managers and conservationists in the western United States are concerned about loss of big sagebrush and the decline of associated species (Manier et al., 2013; Schuman et al., 2005).

Big sagebrush is a long-lived shrub which does not re-sprout or regenerate vegetatively (Shultz, 2006). In most years, big sagebrush produces a large seed crop, which is dependent on plant status, site quality (Beetle, 1960; Goodwin, 1956; Young et al., 1989), grazing pressure (Wagstaff and Welch, 1991), insect herbivory (Takahashi and Huntly, 2010), and pathogen levels (Welch and Nelson, 1995), and weather conditions. However, successful recruitment occurs infrequently (Maier et al., 2001; Perryman et al., 2001) and achieving big sagebrush densities that satisfy legal restoration and reclamation requirements remains difficult (Schuman et al., 2005; Steward, 2006; Williams et al., 2002). The population dynamic of big sagebrush follows a Type III survivorship curve with greatest mortality early in life either at the seed or seedling stage (reviewed by Schlaepfer et al., 2014) as is also common for several grass species in sagebrush steppe ecosystems (Boyd and Lemos, 2013; James et al., 2012). Therefore, understanding controls of big sagebrush regeneration is crucial for predicting big sagebrush responses to environmental changes and economic activities and for identifying land management and restoration strategies that promote sagebrush persistence.

Regeneration processes including germination and seedling survival are fundamental to our understanding of population dynamics (Fenner, 2002; Harper, 1977) and responses to global change (Jackson et al., 2009), particularly since environmental factors typically exert stronger regulation during regeneration than adulthood (Grubb, 1977). Due to the importance of understanding regeneration processes, a diversity of models have been developed. However, building process-based models that predict germination and seedling recruitment under field conditions remains difficult due to incomplete physiological knowledge of species, complexity of biological processes, and multiple spatial and temporal scales involved (Forcella et al., 2000; Mok et al., 2012). Successful thermal-time and wet-thermal-time models of germination for species in semiarid ecosystems have been developed, but cover only a small portion of regeneration processes (Rawlins et al., 2012a,b; Roundy et al., 2007). Process-based models have been developed for crops and annual weedy plants (Forcella et al., 2000; Vleeshouwers and Kropff, 2000). Recently, a process-based germination and establishment submodel for species distribution models was presented for southeastern Australia, based on chill and heat accumulation, frost and a moisture index (Mok et al., 2012); however, this model is unsuitable for big sagebrush due to the lack of representations of snow cover and soil water distribution throughout the soil profile—factors which are necessary to understand big sagebrush regeneration (Schlaepfer et al., 2012a,b). Our literature synthesis of big sagebrush regeneration highlighted the lack of a comprehensive suitable model, which is required to better understand consequences of current management practices, global change, and economic activities in ecosystems dominated by big sagebrush (Schlaepfer et al., 2014).

Our overall objective is to understand the abiotic controls over big sagebrush regeneration and quantify how they influence regeneration potential across the distribution of this wide-ranging species. This objective required that we: (i) formulate a process-based simulation model of responses to abiotic conditions of yearly big sagebrush germination and first-year seedling survival potential, (ii) estimate uncertainty in and importance of model parameters, (iii) test whether statistical models explain relationships between climatic conditions and regeneration in big sagebrush better than our process-based model, and (iv) predict patterns of germination and seedling survival of big sagebrush across the western United States.

2. Materials and methods

2.1. Development of a process-based germination and individual seedling survival model (GISSM) for big sagebrush

We constructed GISSM around ecological processes that contribute to yearly germination success and first-season seedling survival potential of big sagebrush. We identified relevant processes based on a literature synthesis of the ecology of germination and seedling establishment of big sagebrush (Schlaepfer et al., 2014) and focused on processes that involve abiotic factors. Our analyses cover three of the four subspecies of big sagebrush recognized by Flora of North America: *tridentata* (Great Basin big sagebrush), *vaseyana* (Rydberg) Beetle (Mountain big sagebrush), and *wyomingensis* Beetle & A.M. Young (Wyoming big sagebrush, Shultz, 2006). The fourth subspecies, *parishii* (A. Gray) H.M. Hall & Clements (Mojave big sagebrush), is restricted to the southwestern part of the range and very limited specific data are available. Our implementation of GISSM is written in R version 2.15.0 (code available in Appendix B.1, R Development Core Team, 2012) and all subsequent analyses were carried out in R.

2.1.1. Overview of the GISSM algorithm and assumptions

GISSM is forced by daily values of top soil temperature, soil water potential for rooted layers, snow cover, and mean January air temperature experienced by the mother plant (*mJSAT*; here, at the location of simulation because most seeds are dispersed locally).

First, GISSM calculates the day of the year when seed dispersal starts. For each day, GISSM determines germination success based on time required to germinate, favorable and unfavorable conditions, and on how seed behave upon quiescence (Fig. 1). Then, for each day with successful emergence, a seedling (representing the cohort of all potential seedlings that emerged on the same day) is simulated and GISSM determines if it grows, ceases growing, or dies, for every day of the remainder of the current growing season (Fig. 1). GISSM results summarize days and years with potential germination success and first-season seedling survival as well as simulated causes of mortality (e.g., sample data and output, Appendices B.2 and B.3).

GISSM output represents potential (and not realized) yearly germination and seedling success and does not predict abundance. GISSM focuses on abiotic factors and thus ignores all biotic factors (e.g. competition, herbivory, pathogens), which have been reported to decrease chances of seed and seedling survival (DiCristina et al., 2006; Jones, 1991; Owens and Norton, 1989; Reichenberger and Pyke, 1990; Schuman et al., 1998) at least after the first year (Owens and Norton, 1989). Similarly, GISSM assumes that there are enough possibilities for mycorrhization, which are reported to allow seedling to withstand drier soils than without mycorrhization (Stahl et al., 1998). Furthermore, GISSM assumes that enough seeds are available on every day after start of seed dispersal for a given year, i.e., seed production, distance of seed dispersal, seedbank dynamics, and safe sites (reviewed in Schlaepfer et al., 2014) were not limiting for our simulation purposes. GISSM assumes also that light is not a factor determining favorable germination conditions because of the assumption of sufficient viable seeds in safe microsites. Similarly, GISSM also assumes that for any given day with germination success, there is emergence success. For seedlings, GISSM does not incorporate interactions between factors affecting growth of seedlings such as the one observed between low soil moisture and freezing conditions (Lambrecht et al., 2007). GISSM assumes that above-ground biomass growth

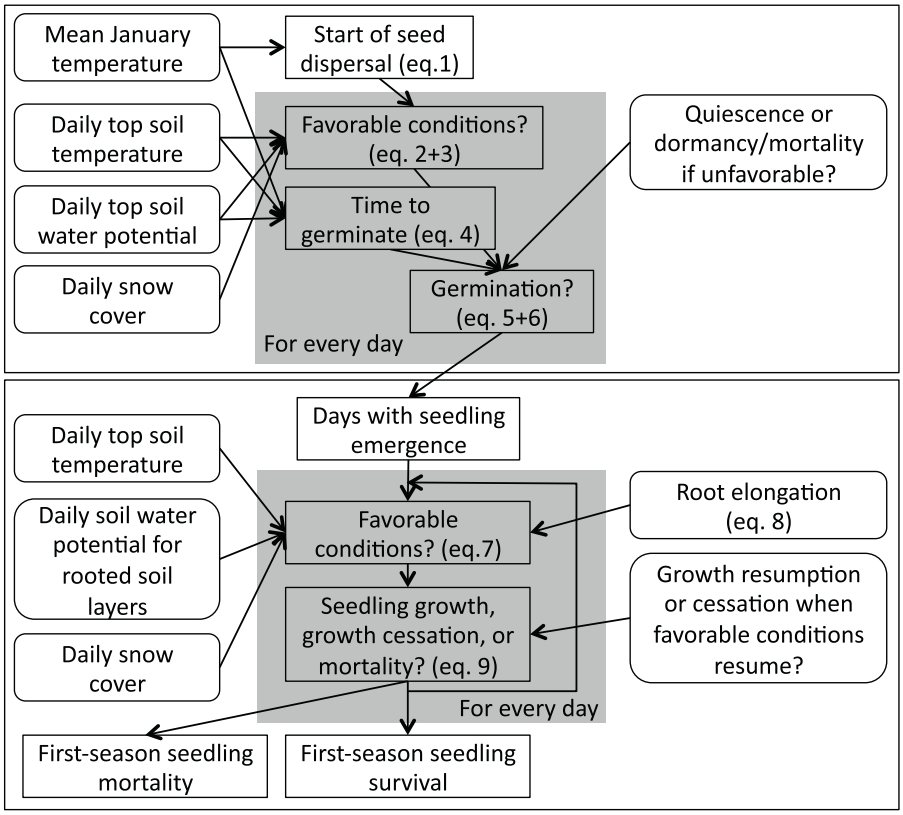

Fig. 1. Conceptual diagram of the Germination and Individual Seedling Survival Model (GISSM) for big sagebrush. Germination conditions were evaluation for every day of the simulation period; survival, growth, and mortality were simulated for each emerged seedling until the end of the first season (i.e., year of emergence).

rate is not limiting and affecting survival rates per se. Finally, GISSM assumes that seedlings have access to sufficient nutrients and that seedlings do not response to fertilization including elevated CO_2-concentrations. The consequence of these assumptions is that GISSM estimates potential, but not realized, germination and seedling survival. For instance, a location without reproductive big sagebrush in vicinity may still have positive potential for regeneration success, but no realized regeneration due to absence of seeds.

2.1.2. Modeling big sagebrush germination

GISSM incorporates four processes relevant to big sagebrush germination (Schlaepfer et al., 2014): start of seed dispersal, periods of favorable/unfavorable conditions, germination rate (which implicitly includes dormancy and stratification), and seed quiescence under unfavorable conditions.

Start of seed dispersal is influenced by the maternal environment, which is summarized by *mJSAT* (Meyer and Monsen, 1991, 1992; Meyer et al., 1990). A linear model estimates start of seed dispersal with the intercept parameter (*k1*, Table 1) describing start of seed dispersal for a site with 0°C *mJSAT* and with the slope parameter (*k2*) correcting for a relationship with *mJSAT*:

$$\text{Day of seed dispersal start (day of year)} = k1 + k2 * mJSAT \quad (1)$$

GISSM estimates germination percentage for each day *j* with a rectangular function of 100% within and 0% germination outside favorable conditions (Eq. (3)); it does not differentiate intermediate levels because it assumes sufficient viable seeds. Favorable conditions include, as considered by GISSM, daily topsoil temperatures (Eq. (2)) above minimal (*k11*) and below maximal soil temperature (*k12*), adjusted for temperatures experienced underneath snow cover (*k13*), and daily soil moisture levels above minimal soil water potential (*k14*) for the top 3 cm of soils (*k15*).

$$\begin{aligned}&\text{Adjusted top soil temperature } T_{j,\text{adj}} \text{ on day } j\,(\text{C})\\ &= T_{j,\text{soil at } k15},\ \text{if snow cover}_j = 0\\ &= k13,\ \text{if snow cover}_j > 0 \end{aligned} \quad (2)$$

$$\begin{aligned}&\text{Germination percentage } GP_j \text{ on day } j\,(\%)\\ &= 0,\ \text{if } T_{j,\text{adj}} < k11 \mid T_{j,\text{adj}} > k12 \mid SWP_{j,(\text{at } k15)} < k14\\ &= 100,\ \text{if } T_{j,\text{adj}} \geq k11 \,\&\, T_{j,\text{adj}} \leq k12 \,\&\, SWP_{j,(\text{at } k15)} \geq k14 \end{aligned} \quad (3)$$

The calculation of time to germinate (i.e., the inverse of germination rate) by GISSM accounts for seedbed conditions underneath snow cover, thereby, GISSM implicitly incorporates dormancy and stratification during the cold season (Fig. A.2, Meyer and Monsen, 1991). We included four parameters (*k3–k6*, Table 1) to describe a thermal-time model (Eq. (4a) here, Eq. (5) in Hardegree, 2006). Such models were developed for conditions of constant temperature, which do not occur under field conditions; instead, GISSM uses the means $T_{k,\text{adj}}$ and $SWP_{k,(\text{at } k15)}$ during the estimated period (*k*) of germination as estimation of incubation temperature ($T_{j,inc}$) and incubation soil water potential ($SWP_{j,inc}$). Thermal-time models were successfully applied to several grass species of sagebrush steppe under field conditions (Rawlins et al., 2012b). To better simulate field conditions other than temperature, we added linear dependencies to the basic thermal-time model with *mJSAT* (*k7*), with an interaction term between T_{inc} and *mJSAT* (k8 is representing maternal effects, Meyer and Monsen, 1991, 1992;

Table 1
Parameter definition and parameter bounds and estimates for the Germination and Individual Seedling Survival Model (GISSM) of big sagebrush. Parameters *k1–k15* concern the germination submodel, *k16–k30* apply to the seedling submodel. Parameters *k15*, *k17*, and *k24* were not optimized. Estimation of lower and upper bounds for optimization were based on a literature review (Schlaepfer et al., in review). Sets of optimized GISSM parameter values are the overall best repeat (best performance during testing), median, and mean values across the 9 out 12 optimization repeats, which tested better than chance (values for each of the 12 repeats are listed in Table A.2). Parameters were also estimated directly from observations reported in the literature (Schlaepfer et al., 2014), independent from those used for optimization. NA indicates values not available.

	Parameter	Unit and constraints	DE optimization: parameter bounds		DE optimization: parameter value sets			Independent: parameter estimation			
			Low	Up	Best repeat	Median[a]	Mean	Value	SE	Number of observations	Mode of estimation
k1	'Start of seed dispersal' at mean January site air temperatures (mJSAT) of 0°C	[Day of year (DOY)]	244	335	325	315	310	289	NA	5	Qualitative data
k2	Dependency of 'start of seed dispersal' on mJSAT	[DOY °C^{-1}]	0	5	2.04	0.89	1.02	2	NA	5	Qualitative data
k3	Coefficient a in Eq. (5) of Hardegree (2006)[b]	[day^{-1}], >0	0	1	0.65	0.32	0.41	0.24	0.20	284	DE optimization of RMSE with Jackknife
k4	Coefficient b in Eq. (5) of Hardegree (2006)[b]	[°C]	1	38	13.7	9.6	14	13.3	1.3	284	DE optimization of RMSE with Jackknife
k5	Coefficient c in Eq. (5) of Hardegree (2006)[b]	[°C^{-1}][c]	−200	200	−116	−115	−47	29	2.4	284	DE optimization of RMSE with Jackknife
k6	Coefficient d in Eq. (5) of Hardegree (2006)[b]	[−], $\neq 1$ and >0	0	10	0.37	2.31	3.73	2.2	1.1	284	DE optimization of RMSE with Jackknife
k7	Dependency of 'time to germinate' on mJSAT	[day °C^{-1}]	−20	0	−0.4	−2.55	−2.49	−7.53	0.382	209	Linear regression
k8	Dependency of 'time to germinate' on mJSAT × incubation temperature	[day °C^{-2}]	−1	1	0.27	0.67	0.7	0.47	0.033	209	Linear regression
k9	Dependency of 'time to germinate' on incubation soil water potential	[day MPa^{-1}]	−10	0	−3.54	−4.87	−4.02	−6.51	0.868	18	Linear regression
k10	Germination resumption/mortality after quiescence	[−], $\in \{0, 1\}$	0	1	1	1	0.67	NA	NA		
k11	Minimal germination temperature	[°C]	−5	5	3.6	−0.4	−0.1	1	NA	374	Data bounds
k12	Maximal germination temperature	[°C]	30	50	43.8	36.9	38.4	38	NA	374	Data bounds
k13	Temperature underneath snow cover	[°C]	−5	5	3.4	0.7	1	NA	NA		
k14	Minimal soil water potential allowing germination	[MPa]	−5	−0.0333	−0.4	−1.8	−2.1	−1.5	NA	26	Data bounds
k15	Depth of top soil determining germination conditions	[m]	0.03	0.03	NA	NA	NA	NA	NA		
k16	Growth resumption/cessation after favorable condition resume	[−], $\in \{0, 1\}$	0	1	1	1	0.78	NA	NA		
k17	Maximal snow water equivalents allowing growth	[m]	0	0	NA	NA	NA	NA	NA		
k18	Minimal growth temperature	[°C]	−30	0	−3	−8	−8	−5	NA	1	Data
k19	Maximal growth temperature	[°C]	30	50	34	34	37	NA	NA		
k20	Duration of snow cover allowing survival	[day]	1	60	31	23	27	NA	NA		
k21	Minimal survival temperature	[°C]	−30	0	−9	−15	−12	−15.2	NA	6	Data bounds
k22	Maximal survival temperature	[°C]	30	50	34	37	40	NA	NA		
k23	Duration of chronic wet soils	[day]	1	365	56	113	161	21	NA	1	Data
k24	Maximal chronic soil water potential allowing survival	[MPa]	−0.0333	−0.0333	NA	NA	NA	NA	NA		
k25	Duration of chronic dry soils	[day]	1	365	49	167	178	4	NA	1	Data
k26	Minimal chronic soil water potential allowing survival	[MPa]	−5	−0.0333	−2.3	−2.3	−1.6	−1.5	NA	1	Data
k27	Minimal acute soil water potential allowing survival	[MPa]	−5	−0.0333	−3.3	−2.9	−2.7	−3.88	NA	45	Data bounds
k28	Initial rooting depth	[mm]	1	100	74	36	52	36	13	18	Logistic regression
k29	Maximal rooting depth	[mm]	36	2000	1765	1298	1251	828	33	18	Logistic regression
k30	Rate of root growth	[day^{-1}]	0	0.2	0.189	0.105	0.122	0.090	0.013	18	Logistic regression

[a] 50th quantile type 3, i.e., nearest even order statistics (Hyndman and Fan, 1996), estimates which value of the 12 repeats is closest to the middle

[b] See Eq. (2a,b).

[c] Constraints for $k5$

$$\begin{cases} \text{for } k5 > 0, \text{ then } k5 \in \left]-(T-k4)*(k6^2-1)/k6, \infty\right[\\ \text{for } k5 = 0, \text{ then } k5 = 0 \\ \text{for } k5 < 0, \text{ then } k5 \in \left]-\infty, -(T-k4)*(k6^2-1)/k6\right[, \text{ with } T = \text{incubation temperature} \end{cases}.$$

Meyer et al., 1990), and with SWP_{inc} ($k9$, Eq. (4b)):

$$\text{Germination rate } R_j \text{ on day } j\,(\text{day}^{-1}) = k3 * \exp[-\ln(2)/(\ln(k6))^2 * (\ln(1 + (T_{j,inc} - k4) * (k6^2 - 1)/(k5 * k6)))^2] \quad (4a)$$

$$\text{Time to germinate } TG_j \text{ on day } j\,(\text{day}) = R_j^{-1} + k7 * mJSAT + k8 * mJSAT * T_{j,\text{inc}} + \text{k9} * \text{SWP}_{j,\text{inc}} \quad (4b)$$

Seeds in GISSM become quiescent upon conditions becoming unfavorable. Once conditions resume to be favorable, seeds may show one of three possible behaviors (Eq. (5), $k10$ parameter, Table 1): resumption of germination (continuation of the accumulation of thermal-time), permanent cessation (here, equivalent to mortality of the embryo), or restart of germination requirements (here, equivalent to moving to a seed-cohort of a different day, and thus not explicitly simulated).

$$\begin{aligned}\text{Duration of favorable conditions } F_j \text{ on day } j\,(\text{day}) &= \text{first run-length}\,(\text{T}_{k,\text{adj}} \geq k11\,\&\,\text{T}_{k,\text{adj}} \leq k12\,\&\,\text{SWP}_{k,(\text{at } k15)} \geq k14)\text{ for } k = j\ldots365,\ \textit{if } k10 = 0\,(\text{permanent cessation}) \\ &= \text{sum of}(T_{k,\text{adj}} \geq k11\,\&\,T_{k,\text{adj}} \leq k12\,\&\,\text{SWP}_{k,(\text{at } k15)} \geq k14)\text{ for } k = j\ldots365,\ \text{if } k10 = 1\,(\text{resumption of germination})\end{aligned} \quad (5)$$

Finally, GISSM determines germination success for each day (Eq. (6)).

$$\text{Germination success for day } j\,(\text{logical}) = (\text{GP}_j > 0\%\,\&\,F_j \geq TG_j) \quad (6)$$

2.1.3. Modeling big sagebrush seedlings

GISSM includes three processes relevant to big sagebrush seedling growth and survival during the first season, i.e., year of emergence, (reviewed in Schlaepfer et al., 2014): availability of conditions suitable for growth, absence of conditions leading to mortality, and above- and belowground growth rates.

Seedlings as simulated by GISSM grow during favorable conditions (Eq. (7), Fig. 1). Upon conditions becoming unfavorable, the seedlings cease growing. Once conditions resume to be favorable, GISSM can either simulate growth resumption or permanent cessation (Eq. (7), $k16$ parameter, Table 1). Favorable growth conditions are defined by the amount of snowpack ($k17$) and by minimum ($k18$) and maximum air temperatures ($k19$).

$$\begin{aligned}\text{Seedling growth SG}_\text{j} \text{ on day } j\,(\text{logical}) &= \text{FALSE, if (snow cover}_j \geq k17 \mid T_{j,air} \leq k18 \mid T_{j,air} \geq k19) \mid (k16 = 0\,(\text{permanent cessation})\,\&\,\text{any(snow cover}_k \geq k17 \mid T_{k,air} \leq k18 \mid T_{k,air} \geq k19)\text{ for k} = 1\ldots(j-1)) \\ &= \text{TRUE, if ((snow cover}_j < k17\,\&\,\text{T}_{j,air} > k18\,\&\,T_{j,air} < k19)\,\&\,k16 = 1\,(\text{growth resumption})) \mid ((\text{snow cover}_j < k17\,\&\,\text{T}_{j,air} > k18\,\&\,T_{j,air} < k19)\,\&\,k16 = 0\,\&\,\text{all(snow cover}_k < k17\,\&\,T_{k,air} > k18\,\&\,T_{k,air} < k19)\text{ for k} = 1\ldots(j-1))\end{aligned} \quad (7)$$

GISSM reflects that elongation rates of the primary root (seedlings tend to have a primary root with few, small lateral roots, Wijayratne, 2011) is critical to survival as longer roots allow access to water stored in deep soil layers during dry summer periods (Schlaepfer et al., 2012b). GISSM determines seedling growth and survival for each day starting on the day of emergence (DoE) based on the soil layer with the highest soil water potential, among those accessed by the elongating roots. The rooting depth is estimated for a given day (Eq. (8)), with favorable growing conditions, by the logistic growth function (Verhulst, 1838) with three parameters, initial rooting depth ($k28$), maximal rooting depth at the end of the first season ($k29$), and rate of root growth ($k30$), as:

$$\text{Rooting depth RD}_j \text{ on day } j\,(\text{mm}) = k28 * k29 * \exp(k30 * k)/(k29 + k28 * (\exp(k30 * \text{k}) - 1)),\ \text{with } k = \text{sum of(SG}_l)\text{ for } l = \text{DoE}\ldots j \quad (8)$$

Finally, simulated seedlings die when conditions become more extreme than survival thresholds (Eq. (9)), which are described by duration of continuous snow cover ($k20$), by minimum ($k21$) and maximum topsoil temperatures ($k22$), by duration ($k23$) of chronic high water conditions ($k24$), by duration ($k25$) of chronic dry soils ($k26$) and by acutely dry soils ($k27$).

$$\begin{aligned}\text{Seedling death SD}_j \text{ on day } j\,(\text{logical}) &= \text{FALSE, if (length of (continuous snow cover}_{\text{DoE}\ldots j} > 0) < k21\,\&\,\text{T}_{j,topsoil} < k22\,\&\,\text{length of (SWP}_{\text{DoE}\ldots j} \geq k24) < k23\,\&\,\text{length of (SWP}_{\text{DoE}\ldots j} \leq k26) < k25\,\&\,\text{SWP}_j > k27 = \text{TRUE, if (length of (continuous snow cover}_{\text{DoE}\ldots j} > 0) \geq k21 \mid T_{j,topsoil} \geq k22 \mid \text{length of (SWP}_{\text{DoE}\ldots j} \geq k24) \geq k23 \mid \text{length of (SWP}_{\text{DoE}\ldots j} \leq k26) \geq k25 \mid \text{SWP}_j \leq k27\end{aligned} \quad (9)$$

2.2. Forcing GISSM with daily data

As input forcing to simulate yearly success/failure at a given site, GISSM uses daily soil water potential in pre-defined soil layers, snow cover, soil temperature, and mean air temperature in January. We generated these inputs with SOILWAT, a daily time step, multiple soil layer, process-based, simulation model to predict ecosystem water balance for big sagebrush. SOILWAT was developed and tested in the semiarid western US shortgrass steppe (Parton, 1978; Sala et al., 1992) and has been adapted for sagebrush ecosystems (Schlaepfer et al., 2012a,b,c). Soil temperature is calculated as described in Parton (1978). As inputs for SOILWAT, we used daily precipitation and minimum and maximum air temperature from a 1/8th-degree gridded dataset for 1949–2010 (Maurer et al., 2002), mean monthly relative humidity, wind speed and cloud cover data (National Climatic Data Center, 2005), monthly vegetation (live and dead biomass, litter, and active root profile) and site-specific properties of each soil layer defined by lower depths limits at 5 cm, 10 cm, 20 cm, 30 cm, 40 cm, 60 cm, 80 cm, 100 cm, and 150 cm (1-km gridded STATSGO version CONUS-SOIL, Miller and White, 1998). Simulated water flows include interception by vegetation and litter, evaporation of intercepted water, snow melt and loss (sublimation and wind redistribution), infiltration into the soil profile, percolation and hydraulic redistribution for each soil layer, bare-soil evaporation, transpiration from each soil layer, and deep drainage (Lauenroth and Bradford, 2006; Schlaepfer et al., 2012b).

2.3. Estimating GISSM parameters

We estimated values for 27 of 30 GISSM parameters with global optimization. Three parameters were not optimized because their values were fixed by the setup of the SOILWAT simulations ($k15$, $k17$, and $k24$; Table 1). We trained the model using the differential evolution (DE) algorithm (using recommended parameter

values for the 'JADE' option, Ardia et al., 2012; Price et al., 2006) against a dataset of observations of big sagebrush establishment for given years at given sites (Table A.1). We compiled this dataset of sites with either known years of success or failure for first-season seedling survival, or establishment of an individual plant in relation to absence or presence of disturbance events such as fire, sagebrush control, or grazing. The dataset includes 46 sites with 2164 unique site-years with no treatments (Fig. A.1) and 20 sites with 255 unique site-years with treatments. We excluded site-years with either a treatment (e.g., grazing, chemical/physical big sagebrush removal) or without a geo-referenced location and retained those with no treatment for a total of 46 sites and 1435 site-years. We repeated the optimization process 12-times with data partitioning whereby we used 84% of site-years for training and reserved the remaining site-years for testing (Fielding and Bell, 1997). We terminated an optimization run if the objective value was not improved for 15 iterations. We minimized as objective value one minus the mean of three evaluation criteria, 0.5-weighted kappa (Cohen's kappa, Kraemer, 2006), true skill statistic (Allouche et al., 2006), and gamma coefficient (Rousson, 2007), if overall accuracy (Franklin and Miller, 2009) was larger than expected by chance (Cohen, 1960). An objective value of 0 indicates a perfect fit; a value of 1 indicates chance performance. Performance criteria were based on the confusion matrix between observed and predicted years with or without first-season seedling survival for each site. We applied a continuity correction, i.e., addition of the square-root of machine epsilon (here $\sqrt{2^{-52}}$ for 64-bit double), if a denominator was zero. Overall performance was calculated as the mean across sites weighted by number of the observation years per site. We chose to use the mean over three evaluation criteria because individually each has been critiqued, particularly for dependence on prevalence (here, fraction of years with seedling survival) and tendencies to promote over- or under-prediction (Allouche et al., 2006; Liu et al., 2011; Mouton et al., 2010). We retained the best parameter set out of each of the 12 optimization repeats for analysis, which predicted regeneration outcome for the reserved site-years for testing better than chance, i.e., objective value smaller than 1.

We tested the stability of the optimization result, using the best repeat, by perturbing each parameter for 101 steps within the parameter bounds, while keeping the other parameters fixed at their optimized values (similar to estimating parameter importance, see below). We then adjusted the parameter that reduced the objective value the most based on the training dataset from the best repeat. We iterated this procedure until no improvement greater than 0.001 of the objective value was possible or after 15 iterations. We then compared the objective value based on the testing dataset with the initial test objective value.

Additionally, datasets collected from our literature synthesis (supplementary materials in Schlaepfer et al., 2014) allowed us to estimate a subset of GISSM parameters independently. Parameters describing upper or lower bounds, e.g., maximal temperature at which seeds germinate, were estimated as minimal or maximal value, which occurred in the literature datasets. We estimated parameters describing the relationship between seedling age and rooting depth with nonlinear regression (based on data from Table S3 in Schlaepfer et al., 2014) and the relationship between time to germinate and incubation temperature (i.e., a thermal-time model as in Eq. (5), Hardegree, 2006) was estimated with global optimization of root mean square error (RMSE) as above with leave-one-out Jackknife to estimate standard error (based on data from Table S2 in Schlaepfer et al., 2014). Because fit with observed values was insufficient (Fig. A.2d, Spearman's *rho* = 0.78), we then added linear dependencies between time to germinate, maternal mean January temperature and an interaction term with incubation temperature (Fig. A.2b, all $p < 0.001$) and incubation water potential (Fig. A.2c; $k9 = 4.338 \pm 0.588$ days/MPa, $p < 0.001$). This modified model of germination rate fitted observed data well (Fig. A.2e; Spearman's *rho* = 0.92).

2.4. Estimating uncertainty in GISSM parameters

We estimated uncertainty in the GISSM parameters in two ways. First, standard deviation, coefficients of variation as well as minimum and maximum values for the 27 optimized GISSM parameters between the 12 optimization repeats were used to estimate the degree to which the accumulated observations of big sagebrush establishment constrained parameter values. Second, we checked whether the parameter values estimated separately based on datasets collected from a literature synthesis (Schlaepfer et al., 2014) fell within the minimum-maximum range of the 12 optimization repeats.

2.5. Estimating importance of GISSM parameters

We estimated the importance of each parameter on the performance of GISSM as its relative influence on the objective value. For each parameter set, we perturbed each parameter individually in 11 steps between minimum/maximum bounds (Table 1), while keeping the others fixed at their optimized values. We determined the objective value (as above for DE optimization) based on all observations for each level of perturbation, calculated importance I_k of each parameter as a rate of maximal improvement, $I_k = [\max(\text{objective}) - \min(\text{objective})]/[\%\text{perturbation at } \max(\text{objective}) - \%\text{perturbation at } \min(\text{objective})]$ with units $[\%^{-1}]$, and ranked the parameters.

2.6. Comparing GISSM with statistical models: relationships between regeneration and climatic conditions

We built a Poisson-GLM between 24 climatic variables and regeneration events at the 46 sites available for both observation and predictions. The climate variables were the seasonal means (December–February, March–May, June–August, and September–November) of precipitation and temperature for the years previous, during, and following each establishment event. Because of high collinearity, we excluded variables with the highest variance inflation factor (VIF) until the largest remaining VIF <2.5 and all pairwise correlation coefficients <0.6 (Zuur et al., 2010). We retained 13 climate variables: all precipitation and mean temperature of December-February of the year of establishment. We tested for a significant pattern among sites in the direction of the linear relationships between climatic variables and establishment by comparing the positive and the negative slopes with a paired *t*-test.

2.7. Predicting range-wide patterns of big sagebrush germination and seedling survival potentials

We applied GISSM with the best parameter set to all ca. 30,000 10-km cells spanning the western Unites States as previously used for species distribution models of big sagebrush (Schlaepfer et al., 2010). Inputs to GISSM were SOILWAT simulation runs as described above using aggregated data for the 10-km grid cells (Schlaepfer et al., 2010) for 31 years of simulation (1980–2010). We produced range-wide maps and tested strength of overall nonlinear patterns between germination and seedling success and climatic variables with Brownian distance correlation (Szekely and Rizzo, 2009) using 999 bootstrap replicates as implemented in the *energy* R package (Rizzo and Szekely, 2013) on a random subsample of $n = 5000$.

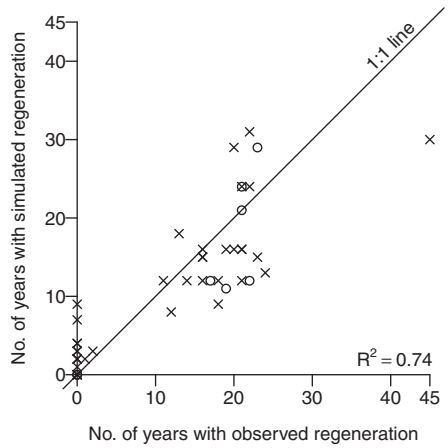

Fig. 2. Comparison of observed vs. predicted number of years with successful regeneration events at 46 sites (training sites, crosses; testing sites, circles; Fig. A.1) using GISSM with the best parameter set (i.e., best repeat).

3. Results

3.1. GISSM parameter estimation and performance

We retained 9 out of the 12 optimization repeats for analysis, the other three failed model testing, i.e., performed worse than chance. We defined four sets of parameter values, the repeat from the 9 optimization repeats with the best objective value during model testing ('best repeat'), the median and the mean values from the 9 repeats, and the values obtained using independent literature data (Tables 1 and A.2).

The 9 optimization repeats each resulted in a different set of parameters, although every optimization repeat terminated due to failure to improve the objective value after 15 iterations. Performance of all 12 repeats was better than chance during model training, i.e., overall accuracy was on average $5.8 \pm 0.7\%$ (SD) larger than chance ($T = 29.9$, $p < 0.001$; Table 2 and A.3). However, the performances of three of the repeats were not transferable to the retained testing dataset. The best repeat retained all performance measures during model testing (Table 2), and performed well predicting the number of years with regeneration successes for the 46 sites ($R^2 = 0.74$, Fig. 2).

The independent parameter set overestimated the proportion of years with regeneration by 7% and was not better than chance. The mean parameter set resulted in a trivial model with zero regeneration and was thus not better than chance. The median parameter set overestimated regeneration by 14% and was only slightly better than chance. The best repeat performed well and resulted in a small bias, i.e., 3% underestimate of the proportion of years with regeneration (Table 2). The best repeat represented the overall best parameter estimates for GISSM. However, the stability of the best repeat was low as any improvement by a single parameter resulted in a large drop of the test objective value (Table A.4), which could not be transferred to the testing sites. Thus, the best repeat remained the best performing parameter estimates for GISSM.

3.2. Uncertainty in model parameters

All estimates of the GISSM parameters retained a large degree of uncertainty after optimization (Table A.5 and Fig. A.3). Four parameters (*k1*, *k12*, *k19*, and *k22*) achieved a coefficient of variation smaller than 20%. However, these parameters, three of which were about maximum temperature, enjoyed relatively tight initial bounds and thus showed only small improvements (12– 27%) after optimization, measured as how much narrower the maximum-minimum range of values became compared to the initial bounds for optimization (Table A.5). Optimization reduced uncertainty in the ranges by more than 50% for five parameters (*k2*, *k7*, *k8*, *k18*, and *k27*), whereas it reduced uncertainty by less than 10% for three parameters (*k10, k16*, and *k20*; Table A.5). The parameter values estimated from independent literature data fell within the range of values estimated by optimization except for three parameters (*k7*, *k25*, and *k27*; Table A.5).

3.3. Relative importance of GISSM parameters

The parameter *k1*, start of seed dispersal, was the most important for every parameter set (Table A.6). The importance of the other parameters depended on the set. For the best repeat, parameters related to the thermal-time model of germination ranked high. Of medium to high importance were parameters related to minimum soil water potential for both seedlings and seeds. Parameters related to temperature, e.g., minimum temperature for seedlings, temperature influence on seed dispersal start, and duration of snow cover, achieved a medium rank. Whether germination resumed or stopped after quiescence and whether seedling growth resumed or ceased after favorable conditions resumed proved to be among the parameters with the lowest importance together with temperature and snow limits of seedling growth.

3.4. Comparing GISSM with statistical models: relationships between regeneration and climatic conditions

Relationships between regeneration events and climatic conditions were so site-specific that no general pattern emerged (Fig. A.4). Focusing on the linear component on the logit-scale of the relationships between each climate variable and regeneration indicated that there is no pattern in the direction of slopes among sites (paired *t*-test, $p > 0.05$ for both observations and predictions). Statistical models (here, Poisson-GAMM) failed (not shown), due to heterogeneity, non-linearity, residual spatial autocorrelation and strong site-specific relationships.

3.5. Range-wide prediction of potential of big sagebrush regeneration

Germination of big sagebrush (Fig. 3b) did not limit years of regeneration except for small areas in southern Nevada, whereas seedling survival (Fig. 3c) limited years of regeneration over most of the distributional range. Regeneration prediction was zero for only a small fraction of cells where the GAP land cover (Fig. 3a, US Geological Survey, 2011) indicated the presence of big sagebrush ecosystems (<0.05% by germination, <2% by seedling survival). Positive regeneration potential was predicted also for areas outside of the distributional range, e.g., Pacific coastal areas east of the Sierra Nevada and Cascades.

Fraction of years with germination was not limited at low mean annual temperature (MAT), but dropped at high MAT, whereas fraction of years with seedling survival showed a uni-modal response with a peak at intermediate MAT (Fig. 4). Germination was limited by low mean annual precipitation (MAP), whereas the response of seedling survival was uni-modal with a peak at intermediate MAP. The fraction of precipitation occurring as snowfall limited both germination and seedling survival at low and high values, though more strongly for seedling survival. All relationships were significant,

Table 2
Performance of GISSM against observations of big sagebrush regeneration at 46 sites with a total of 1435 years for the best repeat (best performance during testing), the parameter sets of median and mean values, and the independent estimates (not estimated parameters filled in with values from the best repeat, Table 1 as well as the training and testing performance of the best repeat and the mean and SD of the 9 out of 12 repeats, which tested better than chance (values for each of the 12 repeats are listed in Table A.3).

All sites	Best repeat	Median parameters	Mean parameters	Ind. parameters	Training	Best repeat	Mean	SD
Number of sites	46	46	46	46	Number of sites	38	38	0
Number of total years	1435	1435	1435	1435	Number of total years	1145	1166	44
Observed proportion of years with regeneration	0.39	0.39	0.39	0.39	Observed proportion of years with regeneration	0.38	0.38	0.01
Predicted proportion of years with regeneration	0.36	0.63	0.00	0.46	Predicted proportion of years with regeneration	0.36	0.33	0.10
Overall accuracy[a]	0.60	0.50	0.61	0.50	Overall accuracy[a]	0.60	0.60	0.03
Chance[b]	0.54	0.47	0.61	0.50	Chance[b]	0.54	0.54	0.03
Objective value[c]	0.87	0.96	0.97	1.01	Objective value[c]	0.88	0.88	0.01
TSS[d]	0.13	0.04	0.03	−0.01	TSS[d]	0.12	0.12	0.01
Gamma[e]	0.27	0.10	0.23	−0.02	Gamma[e]	0.26	0.29	0.06
Kappa[f]	0.14	0.05	0.03	0.01	Kappa[f]	0.13	0.12	0.01

Testing	Best repeat	Mean	SD
Number of sites	8	8	0
Number of total years	290	269	44
Observed proportion of years with regeneration	0.42	0.39	0.05
Predicted proportion of years with regeneration	0.38	0.29	0.12
Overall accuracy[a]	0.60	0.58	0.04
Chance[b]	0.53	0.54	0.05
Objective value[c]	0.85	0.92	0.04
TSS[d]	0.15	0.08	0.04
Gamma[e]	0.32	0.23	0.18
Kappa[f]	0.15	0.09	0.04

[a] Overall accuracy = (Correct positive predictions + correct negative predictions)/number of observations (Franklin and Miller, 2009).
[b] Chance measures expected agreement between observations and model (Cohen, 1960).
[c] 1 – mean(TSS, Gamma, Kappa), see Methods section for more details.
[d] True skill statistic (TSS) accounts for both sensitivity and specificity (Allouche et al., 2006).
[e] Gamma coefficient, a measure of ordinal association (Rousson, 2007).
[f] 0.5-weighted kappa, a calibrated and non-directional measure of association (Cohen's kappa, Kraemer, 2006).

but included also a large amount of variability among simulated cells.

4. Discussion

4.1. Process-based, deterministic simulation model of big sagebrush regeneration

We formulated a process-based simulation model, GISSM, which represents important responses of big sagebrush germination and seedling success to abiotic conditions. This represents, to the best of our knowledge, the first process-based model of big sagebrush potential germination and seedling establishment. We estimated all parameters, with associated uncertainties, and identified those parameters, which are most influential for model performance. GISSM performed better than chance and overall accuracy of yearly prediction of regeneration success/failure was 60%. However, GISSM was able to explain 74% of the variability in numbers of years with successful big sagebrush regeneration. We did not have observational data to assess performance of GISSM predictions at a daily time scale. This level of performance and the differences in performance at different levels of temporal aggregation, in the context of the amount of data for model training and testing and of the detailed process-based modeling, is interesting in itself as it highlights how to reasonably apply GISSM and also how to further improve the model, constrain

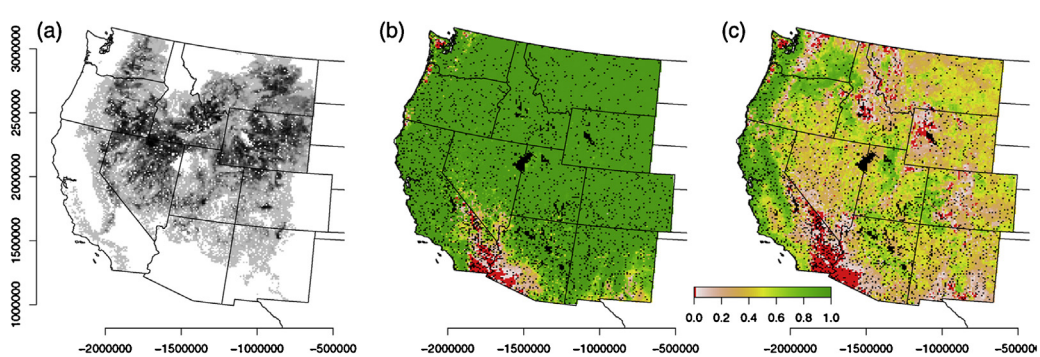

Fig. 3. Range-wide prediction of big sagebrush regeneration with GISSM on a 10-km grid. Distribution of big sagebrush ecosystems according to GAP v2 land cover data (a, darker colors represent more GAP 30-m cells within the 10-km cells, US Geological Survey, 2011). Germination (b) and seedling survival (c) as fraction of years with successes (color gradient from red (0, no years with success), tan (>0) to green (1, every year with success)). Black indicates cells with insufficient data to run SOILWAT. Albers equal-area projection.

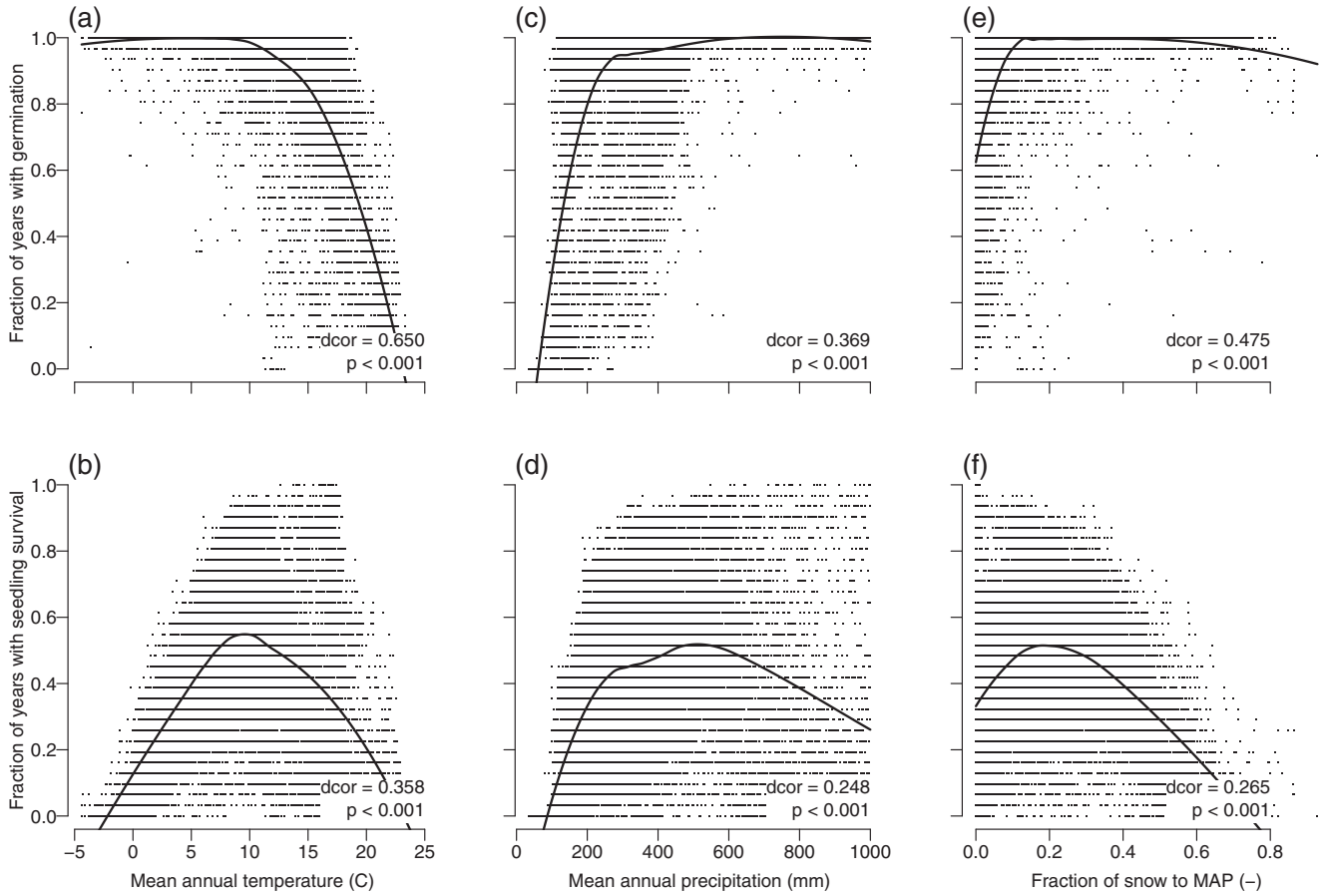

Fig. 4. Scatter between GISSM-predicted germination and seedling survival as fraction of years with successes and climate variables including mean annual temperature (a, b), mean annual precipitation (c, d), and fraction of snowfall to total annual precipitation (e, f). Bold lines summarize shape of relationships using loess function. Brownian distance correlation (dcor) tests strength of relationships (Szekely and Rizzo, 2009). Note, values on *y*-axis occur in discreet steps representing possible outcomes of the 31-year simulation period.

uncertainties, and remove deficiencies in our understanding of regeneration.

GISSM explained patterns of regeneration of big sagebrush at yearly and multi-year time scales. Our results, however, indicated that the performance of GISSM could be further improved were more or higher quality data to become available. Potential for increased performance is indicated by the high variation of parameter estimates among the 12 repeats, low stability of performance when individual parameters are perturbed, the failure of 3 of the 12 repeats during model testing, and the fact that model training constrained parameter uncertainty to a limited degree. The available data may be inadequate because all sites used for model training and testing are clustered in Wyoming, Nevada, and California, missing large parts of the geographic distribution of big sagebrush. We collected 1435 site-years of observations from 46 different sites to train and test GISSM, which provided about 50 observations per parameter. This may be not enough because these observations are not necessarily independent as they likely are correlated in time and space. Also, the available data did not allow GISSM to distinguish among subspecies, which are known to differ in some ecological responses (Schlaepfer et al., 2014). The quality of data may be inadequate for several reasons including: (1) that some of the compiled sources are based on aging big sagebrush individuals by rings, which may be inaccurate by a few years due to missing the first years of growth or by false absences due inability to age old individuals with rotten piths (Cawker, 1980), but see Ferguson (1964); (2) that first-season seedlings were missed for a given year because the seedlings died later; and (3) that the site locations may not be precise enough to extract appropriate weather and soil information to force GISSM. Furthermore, there may be a mismatch of resolution in time between the yearly data used for model building and the model formulation operating at a daily time-step; only daily observations (not available) would allow appropriate assessment of GISSM performance at a daily time-step. There may also be a mismatch between the spatial resolution of inputs used for SOILWAT and outputs used to force GISSM, e.g., daily weather input is at a 1/8th degree resolution and soil information is at a 1-km resolution, which may be too coarse to capture soil water dynamics site-specific enough for the spatial resolution at which observations were conducted. Overall, we argue that data were likely an important issue leading to the performance of GISSM. Nevertheless, these would be theoretically simple issues to overcome as they could be solved by more and better data or more intensive parameter searches.

GISSM estimates a potential maximum of big sagebrush regeneration, because it focuses on regeneration responses to abiotic factors, assumes sufficient seed availability and safe micro-sites, and ignores biotic relationships. Field observations, on the other hand, represent realized regeneration, which may be lower. Model performance, therefore, could be low if potential regeneration over or under estimated realized regeneration for particular years, e.g., due to lack of seeds or safe sites (Beetle, 1960; Owens and Norton, 1992), competition (DiCristina and Germino, 2006; Reichenberger and Pyke, 1990), diseases (Sturges and Nelson, 1986), and herbivory (Takahashi and Huntly, 2010). However, the fact that GISSM produced only small over- and under-predictions within established

big sagebrush stands, particularly when aggregated over multiple years, indicates that this may not be an important issue. Future research could establish the relative importance of abiotic versus of biotic conditions and of safe sites for limiting big sagebrush regeneration. Despite these limitations and potential future improvements, our current model represents the best tool for predicting big sagebrush regeneration.

4.2. Model parameters: uncertainty, sensitivity, and research needs

Our optimization algorithm reduced uncertainty of parameter values compared to previous knowledge by using the 1435 site-years of observations to constrain parameter values. The procedure reduced uncertainty by varying degrees depending on the parameter. Parameters that are important in GISSM (e.g., *k27*) or for which observations contain much extractable information, are likely the ones with strongly reduced uncertainty. On the other hand, unimportant parameters (e.g., *k10*, *k16*) or parameters describing rarely occurring conditions (e.g., snow cover during the growing season, *k20*), for which observations contain little information, are less well constrained. The fact that not all independent literature estimates fall within optimized ranges may be explained by the fact that much of the information in the literature is based on laboratory experiments (e.g., *k27*, which is based on results from one study), which may not be directly applicable to in situ conditions. Despite the achieved reductions in uncertainty, most parameters values remain relatively uncertain and could be further constrained with additional data.

Investments in better understanding the parameter values and relationships is only effective if the parameters and relationships in question are relevant to the model, i.e., that they show a high sensitivity for model performance. Such parameters also identify important bottlenecks for successful regeneration. Start of seed dispersal (*k1*) was the most sensitive parameter; however, this is likely due to a model artifact of having a 'regeneration' year start with seed dispersal and less likely due to a biological result. Whereas most seeds of big sagebrush are not long lived, the majority persist through the winter and germinate the following spring (Meyer and Monsen, 1991; Schlaepfer et al., 2014). Other important parameters describe the influences of soil water potential and temperature (including duration of snow cover) on both germination and seedlings. These results confirm the relevance of soil water, temperature, and snow for big sagebrush regeneration.

These results suggest specific future research that would improve our understanding of big sagebrush regeneration. In particular, start of seed dispersal, the modified wet thermal-time model of germination, and soil water potential influences are important, yet related parameters are not well constrained. These processes warrant further investigation. By contrast, temperature for seedling growth and fate of seeds and seedlings after unsuitable periods appear not to be relevant, and thus may not represent future research priorities. GISSM may be a useful framework for exploring the influence of disturbance effects such as fire or energy extraction. These efforts would benefit from training GISSM with robust datasets for manipulated sites and will likely need to account for availability of seeds, safe sites, and potentially biotic interactions. Similarly, GISSM may facilitate exploration of sagebrush regeneration responses to global change, especially if seed and seedling responses to elevated CO_2-concentrations and nitrogen deposition can be included in addition to disturbance effects.

4.3. Relationships between regeneration and climatic conditions

Our model confirmed that big sagebrush regeneration is strongly dependent on climatic conditions. However, our statistical models between climatic conditions and regeneration success suggested that site-specific factors dominate relationships of regeneration. This may also indicate the benefits that would accrue if sufficient data were available to distinguish among the subspecies of big sagebrush. However, regeneration can be conceived as a function of mother plant health, seed production, germination, emergence, seedling survival (Schlaepfer et al., 2014) and thus likely represents a complex interaction among variables describing climate components at different time scales modified by site-specific edaphic, topographic, and biotic conditions. Therefore, it is not surprising that efforts to build statistical models have failed or produced contradictory results among different sites and areas. For instance, seasonal or monthly mean temperatures showed negative (Cawker, 1980; Schuman et al., 1998), zero (DiCristina et al., 2006), and positive correlations (Cawker, 1980) depending on study area. This may also be reflected by our results indicating overall uni-modal responses of regeneration with peaks at intermediate climate values (Fig. 4). This is a strong indication that big sagebrush regeneration models relying on statistical relationships will not be transferable among sites and that a process-based model such as GISSM is required.

GISSM appears to be well suited to explore responses of big sagebrush regeneration to climate change because of its process-based nature, small overall bias, and incorporation of important climatic variables. Climate change will likely affect relationships that are of high importance for big sagebrush regeneration such as start of seed dispersal, wet thermal-time germination, and dynamics of soil water potential.

4.4. Range-wide potential regeneration of big sagebrush

We successfully applied GISSM across the western United States. The fraction of years with successful regeneration, which we showed to correlate well with field observations, was greater than zero over the vast majority of the area occupied by big sagebrush ecosystems. Long-lived, iteroparous species, such as big sagebrush, do not depend on annual regeneration; thus, infrequent regeneration can be sufficient to maintain populations and the fraction of years with successful regeneration need not correlate with abundance. The geographic overestimation of our model may have arisen because we did not include absences in model development, i.e. sites with observations on big sagebrush regeneration success where big sagebrush did not occur. However, overestimation also suggests that the geographic distribution of big sagebrush is limited by processes other than abiotic influences over regeneration, such as biotic influences on regeneration (e.g., competition) or non-regeneration controls (e.g., adult mortality). Most of the current big sagebrush distribution occurs in dry and mild to cold areas, where germination does not limit regeneration. Germination may only be limiting in the warmest and driest areas, such as the Sonoran, Chihuahuan, and parts of the Mojave Deserts. However, as these driest and warmest areas will expand in the future with climate change (Karl et al., 2009), germination will likely become limiting in additional areas, particularly in those that have been identified as the trailing edge by species distribution models (Schlaepfer et al., 2012c). Seedling success showed uni-modal responses to climate variables, leading to a more complex spatial pattern of limiting big sagebrush regeneration. For instance, 20–30% of MAP occurring as snowfall elicited the highest fitted amount of seedling survival with more or less snowfall leading to reduced seedling survival. Sufficient snowmelt is required in arid and semiarid regions to recharge soil water at deeper soil layers, which are important for big sagebrush – a species with both shallow and deep roots (Sturges, 1977) – during the dry summer periods (Schlaepfer et al., 2012b). More snow, on the other hand, can lead to growing periods that are too short and could reduce seedling survival. However, Schlaepfer

et al. (2012a), focusing on mature vegetation, found that only big sagebrush ecosystems in areas where more than 30% of MAP falls as snow responded positively to future predicted climate change. Thus, our results may be an indication that big sagebrush seedlings may respond differently to climate change than mature stands. This would not be surprising as juvenile stages are often found to differ in ecological requirements from adults (Bell et al., 2014; Grubb, 1977); however, such differences can be crucial for assessments under climate change and are too often neglected by statistical approaches (Jackson et al., 2009). Evaluations of big sagebrush under future environments should account explicitly for responses of regeneration to climate conditions using process-based models such as GISSM.

Acknowledgements

We thank the US Geological Survey, Southern Rockies Landscape Conservation Cooperative, Grant G11AC20366, and the University of Wyoming for funding, and Donovan Miller for implementing the soil temperature module in SOILWAT. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2014.04.021.

References

Alley, H.P., 1956. Chemical control of big sagebrush and its effect upon production and utilization of native grass species. Weeds 4, 164–173.
Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232.
Ardia, D., Mullen, K.M., Peterson, B.G., Ulrich, J., 2012. 'DEoptim': Differential Evolution in 'R'. R Package Version 2.2-0, http://CRAN.R-project.org/package=DEoptim
Beetle, A.A., 1960. A study of Sagebrush: The Section *Tridentatae* of *Artemisia*. Laramie, WY, USA.
Bell, D.M., Bradford, J.B., Lauenroth, W.K., 2014. Early indicators of change: divergent climate envelopes between tree life stages imply range shifts in the western United States. Glob. Ecol. Biogeogr. 23, 168–180.
Boyd, C.S., Lemos, J.A., 2013. Freezing stress influences emergence of germinated perennial grass seeds. Rangeland Ecol. Manage. 66, 136–142.
Bradley, B.A., 2010. Assessing ecosystem threats from global and regional change: hierarchical modeling of risk to sagebrush ecosystems from climate change, land use and invasive species in Nevada, USA. Ecography 33, 198–208.
Cawker, K.B., 1980. Evidence of climatic control from population age structure of *Artemisia tridentata* Nutt. in southern British Columbia. J. Biogeogr. 7, 237–248.
Cohen, J., 1960. A coefficient of agreement for nominal scales. Educat. Psychol. Measure. 20, 37–46.
Connelly, J.W., Knick, S.T., Schroeder, M.A., Stiver, S.J., 2004. Conservation assessment of Greater Sage-Grouse and sagebrush habitats. Western Association of Fish and Wildlife Agencies, Cheyenne, WY, USA.
Davies, K.W., Bates, J.D., Nafus, A.M., 2012. Vegetation response to mowing dense mountain big sagebrush stands. Rangeland Ecol. Manage. 65, 268–276.
DiCristina, K., Germino, M., 2006. Correlation of neighborhood relationships, carbon assimilation, and water status of sagebrush seedlings establishing after fire. West. N. Am. Natural. 66, 441–449.
DiCristina, K., Germino, M., Seefeldt, S., 2006. Spatial and temporal patterns of sagebrush (*Artemisia tridentata* ssp. *vaseyana*) establishment following fire. In: Weber, K.T. (Ed.), Final Report: Detection, Prediction, Impact, and Management of Invasive Plants Using GIS. NASA, Greenbelt, MD, USA, pp. 99–110.
Fenner, M. (Ed.), 2002. Seeds: The Ecology of Regeneration in Plant Communities. , 2nd ed. CABI Publishing, Oxon, UK.
Ferguson, C.W., 1964. Annual rings in big sagebrush *Artemisia tridentata*. Tuscon, AZ.
Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conser. 24, 38–49.
Forcella, F., Benech Arnold, R.L., Sanchez, R., Ghersa, C.M., 2000. Modeling seedling emergence. Field Crops Res. 67, 123–139.
Franklin, J., Miller, J.A., 2009. Mapping Species Distributions. Cambridge University Press, Cambridge, UK, 338 p.
Goodwin, D.L., (Ph.D. Thesis) 1956. Autecological Studies of Artemisia Tridentata Nutt. Washington State University, Pullman, WA, USA.
Grubb, P.J., 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. Biol. Rev. 52, 107–145.
Hanser, S.E., Leu, M., Knick, S.T., Aldridge, C. (Eds.), 2011. Sagebrush Ecosystem Conservation and mManagement: Ecoregional Assessment Tools and Models for the Wyoming Basin. Allen Press, Lawrence, KS, USA, 409 pp.
Hardegree, S.P., 2006. Predicting germination response to temperature. I. Cardinal-temperature models and subpopulation-specific regression. Ann. Bot. 97, 1115–1125.
Harper, J.L., 1977. Population Biology of Plants. Academic Press, London/New York, 892 pp.
Hyndman, R.J., Fan, Y., 1996. Sample quantiles in statistical packages. Am. Stat. 50, 361–365.
Jackson, S.T., Betancourt, J.L., Booth, R.K., Gray, S.T., 2009. Ecology and the ratchet of events: climate variability, niche dimensions, and species distributions. Proc. Natl. Acad. Sci. U. S. A. 106, 19685–19692.
James, J.J., Rinella, M.J., Svejcar, T., 2012. Grass seedling demography and sagebrush steppe restoration. Rangeland Ecol. Manage. 65, 409–417.
Jones, G.P., (Ph.D. Thesis) 1991. Seedling Survival and Adult Plant–Water Relations of Black Sagebrush and Big Sagebrush in the Laramie Basin. University of Wyoming, Laramie, WY, USA, 162 p.
Karl, T.R., Melillo, J.M., Peterson, T.C., 2009. Global Climate Change Impacts in the United States. Cambridge University Press, New York, NY, USA.
Kraemer, H.C., 2006. Correlation coefficients in medical research: from product moment correlation to the odds ratio. Stat. Methods Med. Res. 15, 525–545.
Lambrecht, S.C., Shattuck, A.K., Loik, M.E., 2007. Combined drought and episodic freezing effects on seedlings of low- and high-elevation subspecies of sagebrush (*Artemisia tridentata*). Physiol. Plant. 130, 207–217.
Lauenroth, W.K., Bradford, J.B., 2006. Ecohydrology and the partitioning AET between transpiration and evaporation in a semiarid steppe. Ecosystems 9, 756–767.
Liu, C., White, M., Newell, G., 2011. Measuring and comparing the accuracy of species distribution models with presence–absence data. Ecography 34, 232–243.
Maier, A.M., Perryman, B.L., Olson, R.A., Hild, A.L., 2001. Climatic influences on recruitment of 3 subspecies of *Artemisia tridentata*. J. Range Manage. 54, 699–703.
Manier, D.J., Wood, D.J.A., Bowen, Z.H., Donovan, R.M., Holloran, M.J., Juliusson, L.M., Mayne, K.S., Oyler-McCance, S.J., Quamen, F.R., Saher, D.J., Titolo, A.J., 2013. Summary of science, activities, programs, and policies that influence the rangewide conservation of Greater Sage-Grouse (*Centrocercus urophasianus*): U.S. Geological Survey Open-File Report 2013–1098, http://pubs.usgs.gov/of/2013/1098/
Maurer, E.P., Wood, A.W., Adam, J.C., Lettenmaier, D.P., Nijssen, B., 2002. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. J. Climate 15, 3237–3251.
McArthur, E.D., Plummer, A.P., 1978. Biogeography and management of native western shrubs: a case study, section tridentatae of *Artemisia*. Great Basin Natural. Mem. 2, 229–243.
Meyer, S.E., Monsen, S.B., 1991. Habitat-correlated variation in Mountain big sagebrush (*Artemisiatridentata* ssp. *vaseyana*) seed germination patterns. Ecology 72, 739–742.
Meyer, S.E., Monsen, S.B., 1992. Big sagebrush germination patterns: subspecies and population differences. J. Range Manage. 45, 87–93.
Meyer, S.E., Monsen, S.B., McArthur, E.D., 1990. Germination response of *Artemisia tridentata* (Asteraceae) to light and chill: patterns of between-population variation. Bot. Gazette 151, 176–183.
Miller, D.A., White, R.A., 1998. A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. Earth Interact. 2, 1–26.
Mok, H.-F., Arndt, S.K., Nitschke, C.R., 2012. Modelling the potential impact of climate variability and change on species regeneration potential in the temperate forests of south-eastern Australia. Global Change Biol. 18, 1053–1072.
Mouton, A.M., De Baets, B., Goethals, P.L.M., 2010. Ecological relevance of performance criteria for species distribution models. Ecol. Model. 221, 1995–2002.
National Climatic Data Center, 2005. Climate maps of the United States, available online http://cdo.ncdc.noaa.gov/cgi-bin/climaps/climaps.pl (last accessed May 2010).
Owens, M.K., Norton, B.E., 1989. The impact of 'available area' on *Artemisia tridentata* seedling dynamics. Vegetatio 82, 155–162.
Owens, M.K., Norton, B.E., 1992. Interactions of grazing and plant-protection on Basin big sagebrush (*Artemisiatridentata* ssp. *tridentata*) seedling survival. J. Range Manage. 45, 257–262.
Parton, W.J., 1978. Abiotic section of ELM. In: Innis, G.S. (Ed.), Grassland Simulation Model. Springer-Verlag Inc., New York, pp. 31–53.
Perryman, B.L., Aaron, M.M., Hild, A.L., Olson, R.A., 2001. Demographic characteristics of 3 *Artemisiatridentata* Nutt. subspecies. J. Range Manage. 54, 166–170.
Price, K.V., Storn, R.M., Lampinen, J.A., 2006. Differential Evolution – A Practical Approach to Global Optimization. Springer Verlag, Berlin, Heidelberg, Germany.
R Development Core Team, 2012. R: A Language and Environment for Statistical Computing., ISBN 3-900051-07-0, http://www.R-project.org
Rawlins, J.K., Roundy, B.A., Davis, S.M., Egget, D., 2012a. Predicting germination in semi-arid wildland seedbeds. I. Thermal germination models. Environ. Exp. Bot. 76, 60–67.
Rawlins, J.K., Roundy, B.A., Egget, D., Cline, N., 2012b. Predicting germination in semi-arid wildland seedbeds II. Field validation of wet thermal-time models. Environ. Exp. Bot. 76, 68–73.

Reichenberger, G., Pyke, D.A., 1990. Impact of early root competition on fitness components of four semiarid species. Oecologia 85, 159–166.
Rizzo, M.L., Szekely, G.J., 2013. Energy: E-Statistics (Energy Statistics). R Package Version 1.5.0, http://CRAN.R-project.org/package=energy
Roundy, B.A., Hardegree, S.P., Chambers, J.C., Whittaker, A., 2007. Prediction of cheatgrass field germination potential using wet thermal accumulation. Rangeland Ecol. Manage. 60, 613–623.
Rousson, V., 2007. The gamma coefficient revisited. Stat. Probab. Lett. 77, 1696–1704.
Rowland, M.M., Wisdom, M.J., Suring, L.H., Meinke, C.W., 2006. Greater sage-grouse as an umbrella species for sagebrush-associated vertebrates. Biol. Conser. 129, 323–335.
Sala, O.E., Lauenroth, W.K., Parton, W.J., 1992. Long-term soil–water dynamics in the shortgrass steppe. Ecology 73, 1175–1181.
Schlaepfer, D.R., Glättli, M., Fischer, M., van Kleunen, M., 2010. A multi-species experiment in their native range indicates pre-adaptation of invasive alien plant species. New Phytol. 185, 1087–1099.
Schlaepfer, D.R., Lauenroth, W.K., Bradford, J.B., 2012a. Consequences of declining snow accumulation for water balance of mid-latitude dry regions. Global Change Biol. 18, 1988–1997.
Schlaepfer, D.R., Lauenroth, W.K., Bradford, J.B., 2012b. Ecohydrological niche of sagebrush ecosystems. Ecohydrology 5, 453–466.
Schlaepfer, D.R., Lauenroth, W.K., Bradford, J.B., 2012c. Effects of ecohydrological variables on current and future ranges, local suitability patterns, and model accuracy in big sagebrush. Ecography 35, 374–384.
Schlaepfer, D.R., Lauenroth, W.K., Bradford, J.B., 2014. Regeneration processes in big sagebrush (*Artemisia tridentata*). Rangeland Ecol. Manage. (in press).
Schuman, G.E., Booth, D.T., Cockrell, J.R., 1998. Cultural methods for establishing Wyoming big sagebrush on mined lands. J. Range Manage. 51, 223–230.
Schuman, G.E., Vicklund, L.E., Belden, S.E., 2005. Establishing *Artemisia tridentata* ssp. *wyomingensis* on mined lands: science and economics. Arid Land Res. Manage. 19, 353–362.
Shultz, L.M., 2006. The genus *Artemisia* (Asteraceae: Anthemideae). In: Flora of North America Editorial Committee (Ed.), Flora of North America North of Mexico, vols. 19, 20, and 21. Oxford University Press, New York, NY, pp. 503–534.
Stahl, P.D., Frost, S.M., Williams, S.E., Schuman, G.E., 1998. Arbuscular mycorrhizae and water stress tolerance of Wyoming big sagebrush seedlings. Soil Sci. Soc. Am. J. 62, 1309–1313.
Steward, M., 2006. Revegetation of big sagebrush. In: Burget, W., Carlson, R., Dinsmoor, P., Steward, D.G.M., Stowe, R.R., Vicklund, L.E., Ferris, F.K., Kleinman, L.H., Postovit, B.C. (Eds.), Handbook of Western Reclamation Techniques, 2nd ed. Energy, Minerals Counties Coalition, Gillette, WY, USA, pp. 1–3, 5A.
Sturges, D.L., 1977. Soil water withdrawal and root characteristics of big sagebrush. Am. Midland Natural. 98, 257–274.
Sturges, D.L., 1993. Soil-water and vegetation dynamics through 20 years after big sagebrush control. J. Range Manage. 46, 161–169.
Sturges, D.L., Nelson, D.M., 1986. Snow depth and incidence of snowmold disease on mountain big sagebrush. In: McArthur, E.D., Welch, B.L. (Eds.), Proceedings—symposium on the biology of Artemisia and Chrysothamnus. Brigham Young University, Provo, UT. Intermountain Research Station USDA, Ogden, UT, pp. 215–221.
Szekely, G.J., Rizzo, M.L., 2009. Brownian distance covariance. Ann. Appl. Stat. 3, 1236–1265.
Takahashi, M., Huntly, N., 2010. Herbivorous insects reduce growth and reproduction of big sagebrush (*Artemisia tridentata*). Arthropod Plant Interact. 4, 257–266.
US Geological Survey, 2011. Gap Analysis Program (GAP). May 2011. National Land Cover, Version 2, available online URL: http://gapanalysis.usgs.gov/
Verhulst, P.-F., 1838. Notice sur la loi que la population poursuit dans son accroissement. Correspondance Math. Phys. 10, 113–121.
Vleeshouwers, L.M., Kropff, M.J., 2000. Modelling field emergence patterns in arable weeds. New Phytol. 148, 445–457.
Wagstaff, F.J., Welch, B.L., 1991. Seedstalk production of Mountain big sagebrush enhanced through short-term protection from heavy browsing. J. Range Manage. 44, 72–74.
Walston, L.J., Cantwell, B.L., Krummel, J.R., 2009. Quantifying spatiotemporal changes in a sagebrush ecosystem in relation to energy development. Ecography 32, 943–952.
Welch, B.L., 2005. Big Sagebrush: A Sea Fragmented into Lakes, Ponds, and Puddles. Fort Collins, CO, USA.
Welch, B.L., Nelson, D.L., 1995. Black stem rust reduces big sagebursh seed production. J. Range Manage. 48, 398–401.
West, N.E., 1983. Temperate Deserts and Semi-deserts. Volume 5. Ecosystems of the World. Elsevier, Amsterdam, Netherlands.
Wijayratne, U.C., (Ph.D. Thesis) 2011. Seed and Seedling Ecology of *Artemisia tridentata* in a Restoration Context. Oregon State University, Corvallis, OR, USA, 145 p.
Williams, M.I., Schuman, G.E., Hild, A.L., Vicklund, L.E., 2002. Wyoming big sagebrush density: effects of seeding rates and grass competition. Restor. Ecol. 10, 385–391.
Young, J.A., Evans, R.A., Palmquist, D.E., 1989. Big sagebrush (*Artemisia tridentata*) seed production. Weed Sci. 37, 47–53.
Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14.