

Ecological forecasting under climatic data uncertainty: a case study in phenological modeling

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2010 Environ. Res. Lett. 5 044014

(<http://iopscience.iop.org/1748-9326/5/4/044014>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 141.213.23.37

The article was downloaded on 21/12/2010 at 13:42

Please note that [terms and conditions apply](#).

Ecological forecasting under climatic data uncertainty: a case study in phenological modeling

Benjamin I Cook^{1,2}, Adam Terando³ and Allison Steiner⁴

¹ NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA

² Lamont-Doherty Earth Observatory, 61 Route 9W, Palisades, NY 10964, USA

³ Biodiversity and Spatial Information Center, NC State University, Raleigh, NC 27695-7617, USA

⁴ Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI 48109-2143, USA

E-mail: bc9z@ldeo.columbia.edu, ajterand@ncsu.edu and alsteiner@umich.edu

Received 9 September 2010

Accepted for publication 30 November 2010

Published 15 December 2010

Online at stacks.iop.org/ERL/5/044014

Abstract

Forecasting ecological responses to climate change represents a challenge to the ecological community because models are often site-specific and climate data are lacking at appropriate spatial and temporal resolutions. We use a case study approach to demonstrate uncertainties in ecological predictions related to the driving climatic input data. We use observational records, derived observational datasets (e.g. interpolated observations from local weather stations and gridded data products) and output from general circulation models (GCM) in conjunction with site based phenology models to estimate the first flowering date (FFD) for three woody flowering species. Using derived observations over the modern time period, we find that cold biases and temperature trends lead to biased FFD simulations for all three species.

Observational datasets resolved at the daily time step result in better FFD predictions compared to simulations using monthly resolution. Simulations using output from an ensemble of GCM and regional climate models over modern and future time periods have large intra-ensemble spreads and tend to underestimate observed FFD trends for the modern period. These results indicate that certain forcing datasets may be missing key features needed to generate accurate hindcasts at the local scale (e.g. trends, temporal resolution), and that standard modeling techniques (e.g. downscaling, ensemble mean, etc) may not necessarily improve the prediction of the ecological response. Studies attempting to simulate local ecological processes under modern and future climate forcing therefore need to quantify and propagate the climate data uncertainties in their simulations.

Keywords: ecology, climate, forecasting

1. Introduction

There is wide interest within the global change research community in using model projections of future climate to forecast ecological responses to climate change. To date, the vast majority of such studies have used large scale, globally calibrated vegetation and carbon cycle models (e.g., [1, 2]) that

have input requirements compatible with the typical spatial (hundreds of kilometers) and temporal (monthly) resolution of climate model output. These models are most commonly used to investigate changes in vegetation dynamics and carbon storage over large spatial scales (regional to global), with the implicit understanding that these results will have much lower fidelity and utility at any single location or grid cell.

Increasingly, however, researchers are using output from global climate models to force site-specific ecological models that predict species phenology, fecundity, and population dynamics (e.g. [3–10]). This is often done without quantifying the fidelity of the climate simulations at the study location, and typically using output from only a few (or even a single) climate models.

Calibrating and validating these ecological models using climate observations presents a unique challenge to the ecological research community. In many cases, local climate data are simply not available at the same location as ecological observations (e.g. [11]). As an alternative, data from nearby weather stations may be used for model calibration and validation, assuming that these stations will be representative of the climate at the location of the ecological observations. If contemporaneous nearby weather stations are not available, there are a variety of secondary, derived data products that can be substituted. These products are interpolated from sparse observations to uniform and continuous spatial and temporal grids using statistical or process based models (e.g. [12, 13]). The usefulness of these interpolated datasets for local modeling, however, can be limited by the quantity and quality of the underlying observations, their often coarse temporal (e.g. monthly) and spatial (e.g. hundreds of kilometers) resolutions, and the nature of the variable itself (e.g. continuous versus discrete). While global vegetation models can be constructed relatively easily using these coarse global gridded climate datasets, localized models typically require climate data at much finer spatial and temporal resolution than either these gridded datasets or climate models can provide.

We use a case study approach to highlight uncertainties in site-specific ecological modeling, focusing on the climate inputs used to drive ecological models for both the modern period and future climate change scenarios. For our ecological models, we develop three species-specific models of first flowering date (FFD). Phenology models (especially FFD) are one of the most commonly used and easily understood ecological models, due in part to their relatively modest data requirements (e.g. daily temperatures), availability of long-term datasets, and reasonably high understanding of the underlying biological processes. This case study will address three questions:

- How well do the FFD models reproduce the observations when forced with various observational climate datasets?
- How well do the FFD models reproduce the general trends and variability when forced by climate model simulations of the twentieth century?
- What is the range of FFD predictions for a future climate scenario that tracks current greenhouse gas emissions?

Observations of FFD are obtained from the Mohonk Preserve at Mohonk Lake, NY (41.77°N, 74.15°W, 380 m ASL) [14], a site with an accompanying high quality local record of daily temperature and precipitation [15]. The preserve itself is located roughly 135 km north of New York City and covers an area of approximately 2800 hectares. The combination of both local weather and ecological data at this

Table 1. Observational datasets used for Test 1, simulating phenology over the latter part of the twentieth century. Common period to all datasets is 1979–2002.

Dataset	Reference	Spatial Res	Time Res	Range
Mohonk Lake (USHCN)	[15]	Local	Daily	1896–2006
CRU 2.1	[13]	0.5° × 0.5°	Monthly	1901–2002
Climate grids				
Krigged	[16]	Dscaled	3 h	1928–2006
Krigged (Elev corrected)	[16]	Dscaled	3 h	1928–2006
NARR	[17]	36 km × 36 km	3 h	1979–2006

site represents a significant and rare opportunity for evaluation, allowing the use of local phenology and climate data as a benchmark to evaluate the sensitivity of ecological forecast models to the use of alternative climate forcing data.

2. Data and methods

2.1. Climate and phenology at Mohonk Lake, NY

To parameterize the phenology models, we use records of FFD and daily temperature from the Mohonk Preserve. For the last century, the Mohonk Preserve has recorded and maintained records of plant species phenology (FFD) [14] and local climate (daily temperature and precipitation) [15]. The climate record extends from 1896 to the present and is remarkably homogeneous, largely free from biases related to instrumentation changes, observer methodology, or alterations to the surrounding environment. The phenology records begin in the 1920s using consistent observation routes [14]. If, for a given year, a species was not observed to flower at these locations, flowering for that species was not recorded. This results in the occasional missing data point, but helps to ensure high data quality.

We selected FFD observations for three woody perennial plant species: *Amelanchier arborea* (shadbush), *Kalmia latifolia* (mountain laurel), and *Sambucus racemosa* (red berried elder). Each species has at least 30 years of observations, and some of the highest sensitivity to climate of all species at Mohonk Lake, as established in a previous study [14].

2.2. Observational climate records

To evaluate the impact of using various datasets to drive the phenological models, we implement temperatures from five observationally derived datasets to assess the ability of the FFD models to reproduce the observations (table 1). In this study, we define these data sets as observationally derived to indicate that some manipulation is required to convert the original, raw observations to site-specific climate observations. This could include spatial interpolation, temporal interpolation, or other statistical smoothing such as modifying the temperature records to account for elevation.

The FFD models are initially calibrated and validated using the local Mohonk Lake climate record (OM), which serves as our baseline best case scenario. The second dataset

is derived from the CRU version 2.1 climate grids (CM), which are statistically interpolated from individual station records to uniform $0.5^\circ \times 0.5^\circ$ grids at monthly mean temporal resolution [13]. From this dataset, we select the grid cell that contains the Mohonk Lake site and perform a simple temporal repeating of the monthly mean values for each day of the month. While there are many other techniques for temporal downscaling (e.g. weather generators), we implement this simple method here to highlight the differences between monthly and daily data. The third dataset is daily data from nearby stations in the Global Historical Climatology Network (GHCN) [16], interpolated to the Mohonk Lake location from nearby stations using Kriging (KM); a fourth, alternative dataset is identically calculated, but also includes elevation as a covariate in the model (KME). Finally, the fifth dataset is the North American Regional Reanalysis (NM) [17], dynamically downscaled to 36 km spatial resolution from the NCEP–NCAR reanalysis data [12]. The NARR data are also available at a daily resolution, and for our study we use the nearest grid cell to the Mohonk Lake site. For our observational comparison, we use the period 1981–2002 as a common validation interval across all datasets.

2.3. Global and regional climate model output

Output from nine general circulation models (GCM) is taken from the World Climate Research Programme’s Coupled Model Intercomparison Project (CMIP3) database [18]. We select models for which daily data are available for both the modern 20C3M (coupled ocean, twentieth century climate forcings) and future SRES A2 (coupled ocean, high greenhouse gas emissions) scenarios. We use mean surface temperatures from the land-based grid cell closest to the Mohonk Lake site. Output for the four regional climate models (RCM) is obtained from the North American Regional Climate Change Assessment Program (NARCCAP) [19]. These are limited area regional simulations over North America, forced by boundary conditions taken from GCM simulations in the CMIP3 database. The potential advantages of using RCMs include a higher spatial resolution and specific, regional parameterizations that may improve the fidelity of the climate simulation at Mohonk Lake. Our RCM simulations use the same modern and future climate forcing scenarios as the GCMs (20C3M and SRES A2). To account for biases inherent in most climate simulations, all temperature output from the GCM and RCM simulations (20C3M and SRES A2) is converted to anomalies relative to the model mean for 1971–95. These anomalies are then added to the mean from the Mohonk Lake record (OM) for 1971–95. In this way, the mean climate of Mohonk Lake is preserved in the GCM/RCM temperature output, while allowing the models to superimpose their own trends and variability.

2.4. Phenology model calibration and validation

For our phenology model, we iterate the following series of equations through each day of the year, starting from day of

Table 2. Estimated model parameters, based on common calibration interval of 1928–1980.

Species	DayStart	F_{crit}	GDD_{thresh}	n
<i>Amelanchier arborea</i>	56	200	2	38
<i>Kalmia latifolia</i>	71	800	0	38
<i>Sambucus racemosa</i>	61	250	3	27

year (doy) DayStart:

$$GDD(\text{doy}) = \max(T_{\text{mean}}(\text{doy}) - GDD_{\text{thresh}}, 0)$$

$$GDD_{\text{sum}}(\text{doy}) = \sum_{\text{DayStart}}^{\text{doy}} GDD$$

$$\text{if } (GDD_{\text{sum}}(\text{doy}) \geq F_{\text{crit}}), \quad \text{FFD} = \text{doy}$$

where doym is the current day of the year (always greater than DayStart), $T_{\text{mean}}(\text{doy})$ is the mean temperature on doym, $GDD(\text{doy})$ is the growing degree day value for the current day of year, GDD_{thresh} is the baseline temperature for a day to be considered a GDD, $GDD_{\text{sum}}(\text{doy})$ is the cumulative sum of GDDs from DayStart to doym, and F_{crit} is the critical GDD_{sum} threshold required for FFD to be triggered. On the day GDD_{sum} matches or exceeds F_{crit} , the FFD for that species is recorded.

Values for three parameters (DayStart, GDD_{thresh} , F_{crit}) are estimated separately for each of the three species using a root mean square error minimization procedure (table 2). We use 1928–80 as the calibration interval and reserve years 1981–2002 as the common validation interval across all observationally derived climate records. We also performed a parallel model calibration incorporating a fourth parameter (chilling day requirements), but found that any improvement in model fit did not sufficiently compensate for the addition of a fourth parameter, as determined using Akaike’s Information Criterion (AIC). For the observational climate data comparison, these models are forced with varying observation temperature inputs over the common validation interval (1981–2002). For the GCM/RCM comparison, these models are forced over modern (1971–95) and future (2046–65) intervals for which the climate model output is available.

3. Results

3.1. Observationally derived climate data

Temperature, via the GDD and GDD_{sum} variables, is the primary driver for our FFD models. All datasets show significant positive trends in annual (not shown) and half year (1 January–30 June) averaged mean temperatures (figure 1, left panel). Mean temperatures calculated from the local climate data (OM) have a stronger positive trend and are slightly warmer than some of the other datasets, especially towards the end of the validation period. Abrupt and transient cooling following the 1991 eruption of Mount Pinatubo is apparent in all datasets, and the interannual variability is generally coherent between datasets. For GDDs accumulated over the first half of the year (figure 1, right panel), the observationally derived datasets tend to be much colder. Additionally, while the trends are similar if calculated for the first part of the validation period (1981–91), only OM shows a strong

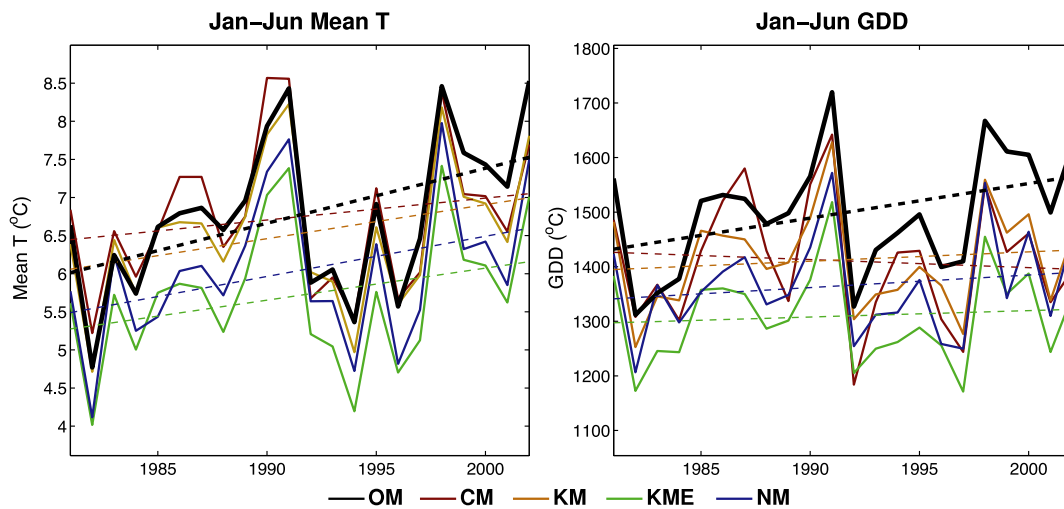


Figure 1. Mean half year (1 January–30 June) temperatures (°C) and cumulative growing degree days (°C) over the common validation interval (1981–2002) from the observational climate datasets. Trend lines are calculated using standard linear least squares methods.

positive (warming) trend in GDDs over the entire interval, indicating large differences in the ability of the different observationally derived datasets to reproduce trends in the important driving variable for the FFD models. The reason for the biases in the mean state and trends in the alternative datasets is unclear. Possibilities include micro-site effects (e.g. land cover changes) or methodological issues related to data recording (e.g. instrumentation, station moves) and data processing (e.g. quality control, interpolation techniques). Regardless, these differences do showcase the potential pitfalls in using non-local climate data for highly localized modeling applications, and strongly suggest that care must be taken in making any assumptions about the suitability of alternative data products.

3.2. Phenology simulations: observational forcing

Because of the cold bias in GDDs in the observationally derived datasets, our FFD simulations are biased towards later than observed predictions of FFD (figure 2, left panels). The fairly uniform bias across datasets translates to an ensemble median that is also biased, counter to prior climate model studies suggesting that ensemble means or medians may provide better predictive capability than any single ensemble member [20, 21]. We quantify the model matches with observations over the validation periods using Taylor diagrams [22]. All simulations capture interannual variability in the FFD, with varying degrees of success, quantified using a Pearson correlation (figure 2, right panels). Notable events captured by all simulations are delayed FFD during 1992 (associated with cooling from the Mt Pinatubo eruption) and earlier FFD during 1998 (one of the warmest years on record). For all three species, the lowest fidelity (based on Pearson) comes from the monthly dataset, CM, that required downscaling to daily resolution. While results may improve when using a more detailed temporal downscaling technique, these results highlight the need for daily temperature resolution. All other datasets show similar

correlations; however it is important to note that all are biased cold relative to OM, suggesting that the seasonal dependence of many ecological forecasting models are subject to the biases of the driving datasets. If the proper seasonal forcing is not accurately represented in the forcing datasets, then the subsequent model response will not accurately reflect the local observations.

3.3. Phenology simulations: climate model forcing

GCM simulations from the CMIP3 archive include a dynamic, prognostic ocean, resulting in internally modeled sea surface temperatures (SST) and interannual climate variability that will not necessarily match with observations. These simulations are still useful, however, for understanding how GCMs reproduce the average dates and general trends in FFD over both modern and future intervals.

For the modern interval, two of the three species (*Amelanchier arborea* and *Kalmia latifolia*) show decreasing trends over time towards earlier FFD dates (black diamonds, figure 3). For these two species, the GCM and RCM simulations generally capture the mean timing in FFD, as well as the trends observed at the site, although the model ensembles generally have a large spread biased towards later than observed FFD and underestimated FFD trends. In the GCM ensemble, some ensemble members even indicate an opposite, positive trend in FFD. There is no notable improvement in the higher resolution RCM simulations in reproducing the ecological response, although this may be attributed to the smaller ensemble size in the RCM ensemble. Of the three species, only *Sambucus racemosa* has positive trend during the modern period, which is not reproduced by the GCM and RCM ensemble simulations. Some of the GCM and RCM ensemble members do reproduce the positive trend during the modern period, and this again highlights that the multi-model ensemble mean or median may not be the most appropriate implementation for ecological studies.

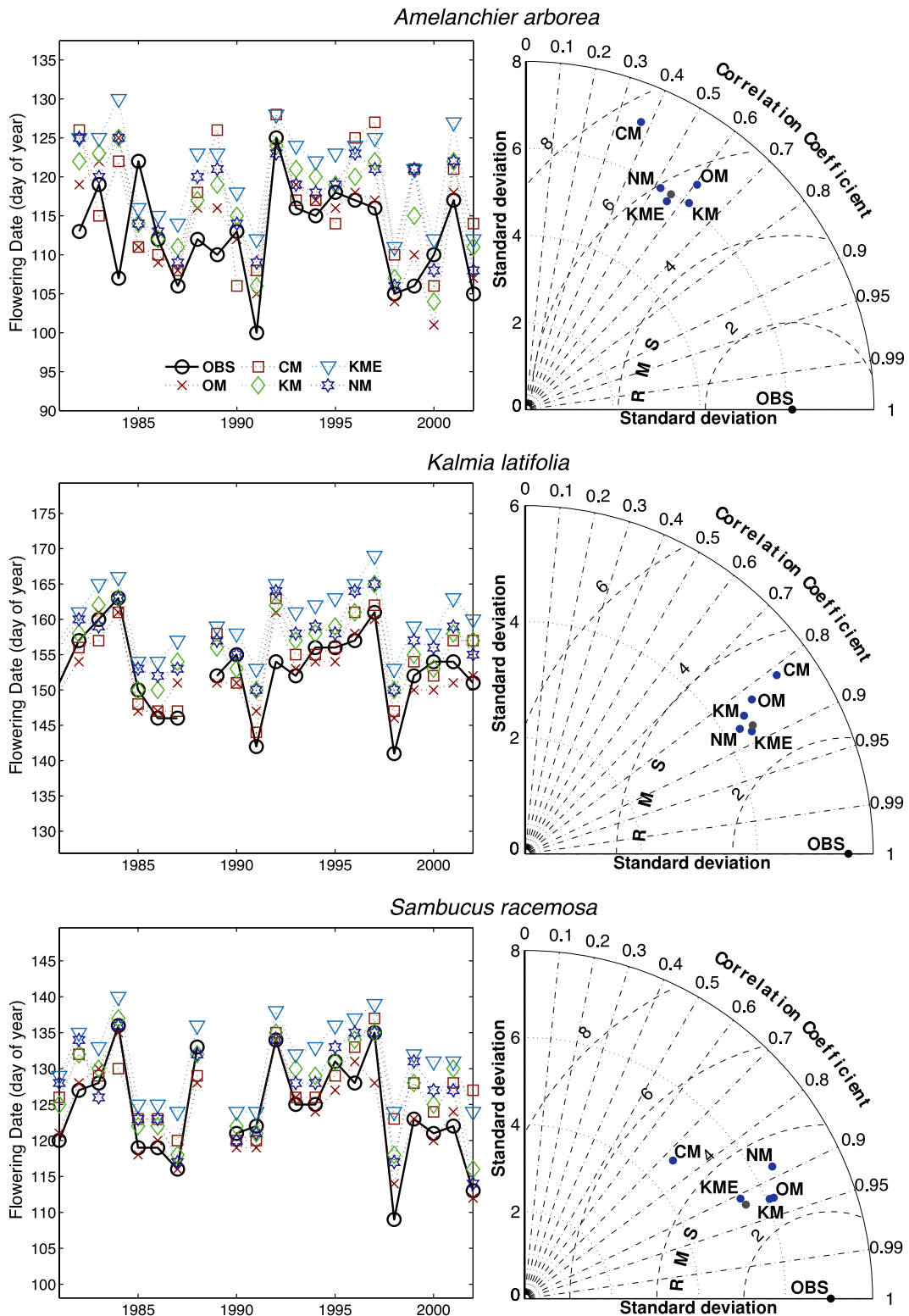


Figure 2. Time series (left column) and Taylor diagrams (right column) for simulated and observed FFD for all three species over the common validation interval (1981–2002). In the Taylor diagrams, blue dots are from simulations of the FFD models forced by the indicated observational dataset, the gray dots are based on an ensemble mean of all dataset simulations, and the true observations are indicated by the black dots.

With a warmer climate over the future interval (2046–65), both the GCM and RCM ensembles shift towards earlier FFD dates. Trends in FFD calculated over 2046–65 in the GCM ensemble also shift towards more uniformly negative

trends, although the ensemble spread is still quite large. The ensemble spread of the mean and trend is lower for the RCM simulations compared to the GCM simulations; again, this may be due simply to sample bias because of the lower RCM

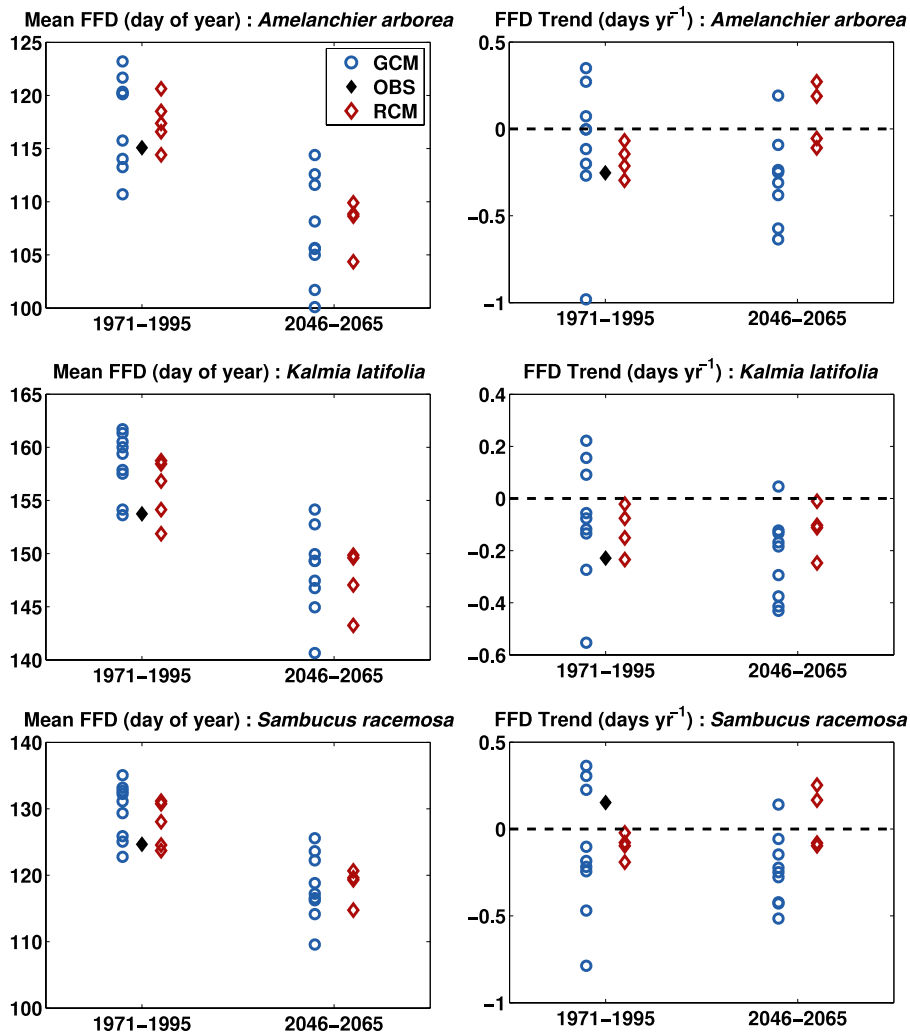


Figure 3. Mean simulated FFD (day of year) and trends (days yr⁻¹) for each species for the modern (1971–95) and future (2046–65) climate model runs. Black diamonds are the from the original FFD observations, blue circles are based on FFD predictions forced by the GCM output, and red diamonds are based on FFD predictions forced by the RCM output.

ensemble size. In addition, the robustness of these trend calculations may be in question because of the limited intervals over which we calculated the trends. We note that, for the future interval especially, we were limited in our choice of models and intervals by the availability of daily temperature output from the IPCC model archive and NARCCAP.

4. Discussion and conclusions

In any modeling study, there are multiple sources of error and uncertainty that need to be characterized. Here, we illustrate some of the uncertainties contained within climate data used to drive highly localized ecological models. For the ecological community, these uncertainties are key because: (1) climate data are often not available directly at the sites of ecological observations; and (2) there is often a large disconnect in scales (spatial and temporal) between highly localized ecological data and the climate data used to develop and drive these local ecological models. This becomes increasingly important in areas with complex local

climatological conditions (e.g. microclimates) and when the climate variable of interest has low spatial covariance. In addition to these scaling issues, there are also uncertainties in the ability of globally (or even regionally) parameterized climate models to adequately simulate features of highly localized climates. Understanding and quantifying these uncertainties is a crucial task for fully understanding the robustness of various ecological and climatological forecasts.

Our observational data comparison showcases some of the inherent risks in using non-local climate data, including biases and the potential for strong disconnects between climate and ecological variables. For example, even though all the observational data sets show a similar trend in seasonally resolved mean temperatures, none reproduce the warming trend in GDDs observed in the local climate record (OM). Additionally, temporally downscaling from monthly to daily observations for the FFD model (the CM dataset) generally resulted in the lowest fidelity simulations. For models that require high temporal resolution input data, as many localized ecological models do, available climate data may

be a significant limitation because of the limited availability of quality daily data and the potential problems associated with using coarser temporal scales to calculate or represent higher resolution inputs (e.g. GDDs). The FFD simulations using GCM and RCM output also showcase the importance of showing the ensemble spread, in addition to the ensemble mean or median. Showing only the mean or median may give false confidence in the robustness of the modeling results, masking what may be a large spread in ecological variables. Recent work using Bayesian methods to quantify and propagate uncertainty in climatology and ecology (e.g. [23–25]) holds promise to provide even more robust methods to account for the structural and parametric uncertainty between and within empirical and dynamic models.

Based on our case study, we recommend that local ecological modeling studies: (1) use a suite of observational climate datasets if local climate data are not available; (2) use caution when selecting datasets that cannot explicitly resolve the temporal resolution of needed input variables; and (3) present not only the ensemble median or mean but also the full ensemble spread. Considering these issues will increase awareness of the uncertainties surrounding projections of ecological responses to climate change, and allow the community to more readily assess the robustness of various predictions.

Acknowledgments

The authors thank the editor and reviewers whose comments greatly improved the quality of this manuscript. This project resulted from work initiated at a USA National Phenology Network Research Coordination Network meeting held in October 2009 and supported by National Science Foundation Grant IOS-0639794. Special thanks to the Mohonk Preserve and the Smiley Family for providing climate and phenology data. We also wish to thank the North American Regional Climate Change Assessment Program (NARCCAP) for providing the data used in this paper. NARCCAP is funded by the National Science Foundation (NSF), the US Department of Energy (DoE), the National Oceanic and Atmospheric Administration (NOAA), and the US Environmental Protection Agency Office of Research and Development (EPA). Lamont contribution 7422.

References

- [1] Cramer W *et al* 2001 Global response of terrestrial ecosystem structure and function to CO₂ and climate change: results from six dynamic global vegetation models *Glob. Change Biol.* **7** 357–73
- [2] Galbraith D, Levy P E, Sitch S, Huntingford C, Cox P, Williams M and Meir P 2010 Multiple mechanisms of Amazonian forest biomass losses in three dynamic global vegetation models under climate change *New Phytol.* **187** 647–65
- [3] Eccel E, Rea R, Caffarra A and Crisci A 2009 Risk of spring frost to apple production under future climate scenarios: the role of phenological acclimation *Int. J. Biometeorol.* **53** 273–86
- [4] Estay S A, Lima M and Labra F A 2009 Predicting insect pest status under climate change scenarios: combining experimental data and population dynamics modelling *J. Appl. Entomol.* **133** 491–9
- [5] Ghosh D, Vogiatzakis I N, Kazakis G, Dimitriou E, Moussoulis E, Maliaka V and Zacharias I 2010 Ecological changes in the highest temporary pond of western Crete (Greece): past, present and future *Hydrobiologia* **648** 3–18
- [6] Kaukoranta T, Tahvonon R and Ylamaki A 2010 Climatic potential and risks for apple growing by 2040 *Agric. Food Sci.* **19** 144–59
- [7] Monserud R A, Yang Y, Huang S and Tchebakova N 2008 Potential change in lodgepole pine site index and distribution under climatic change in Alberta *Can. J. Forest Res.* **38** 343–52
- [8] Morin X, Lechowicz M J, Augspurger C, O’Keefe J, Viner D and Chuine I 2009 Leaf phenology in 22 North American tree species during the 21st century *Glob. Change Biol.* **15** 961–75
- [9] Primack R B, Ibáñez I, Higuchi H, Lee S D, Miller-Rushing A J, Wilson A M and Silander J A 2009 Spatial and interspecific variability in phenological responses to warming temperatures *Biol. Cons.* **142** 2569–77
- [10] Yao F, Xu Y, Lin E, Yokozawa M and Zhang J 2007 Assessing the impacts of climate change on rice yields in the main rice areas of China *Clim. Change* **80** 395–409
- [11] Bradley N L, Leopold A C, Ross J and Huffaker W 1999 Phenological changes reflect climate change in Wisconsin *Proc. Natl Acad. Sci.* **96** 9701–4
- [12] Kalnay E *et al* 1996 The NCEP/NCAR 40-year reanalysis project *Bull. Am. Meteorol. Soc.* **77** 437–71
- [13] Mitchell T D and Jones P D 2005 An improved method of constructing a database of monthly climate observations and associated high-resolution grids *Int. J. Climatol.* **25** 693–712
- [14] Cook B I, Cook E R, Huth P C, Thompson J E, Forster A and Smiley D 2008 A cross-taxon phenological dataset from Mohonk lake, ny and its relationship to climate *Int. J. Climatol.* **28** 1369–83
- [15] Cook B I, Cook E R, Anchukaitis K J, Huth P C, Thompson J E and Smiley S F 2010 A homogeneous record (1896–2006) of daily weather and climate at Mohonk Lake, New York* *J. Appl. Meteorol. Climatol.* **49** 544–55
- [16] Peterson T C and Vose R S 1997 An overview of the Global Historical Climatology Network temperature database *Bull. Am. Meteorol. Soc.* **78** 2837–49
- [17] Mesinger F *et al* 2006 North American Regional reanalysis *Bull. Am. Meteorol. Soc.* **87** 343–60
- [18] Meehl G A, Covey C, Delworth T, Latif M, McAvaney B, Mitchell J F B, Stouffer R J and Taylor K E 2007 The WCRP CMIP3 multimodel dataset *Bull. Am. Meteorol. Soc.* **88** 1383–94
- [19] Mearns L and Team N 2009 The North American regional climate change assessment program (NARCCAP): overview of phase II results *IOP Conf. Ser.: Earth Environ. Sci.* **6** 022007
- [20] Hagedorn R, Doblas-Reyes F J and Palmer T N 2005 The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept *Tellus A* **57** 219–33
- [21] Pierce D W, Barnett T P, Santer B D and Gleckler P J 2009 Selecting global climate models for regional climate change studies *Proc. Natl Acad. Sci.* **106** 8441–6
- [22] Taylor K E 2001 Summarizing multiple aspects of model performance in a single diagram *J. Geophys. Res.* **106** 7183–92
- [23] Clark J S 2005 Why environmental scientists are becoming Bayesians *Ecol. Lett.* **8** 2–14
- [24] Cressie N, Calder C A, Clark J S, Hoef J M V and Wikle C K 2009 Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling *Ecol. Appl.* **19** 553–70
- [25] Smith R L, Tebaldi C, Nychka D and Mearns L O 2009 Bayesian modeling of uncertainty in ensembles of climate models *J. Am. Stat. Assoc.* **104** 97–116