

The Role of Long Time Series of Ecological Data for the Calibration and Evaluation of Ecological Models

Mathematical models of ecological processes play an important role in the assessment of Global Change impacts on terrestrial ecosystems as well as for predicting the future of the Earth System as a whole. Over the coming decades to centuries, ecosystem function (e.g., biogeochemical cycling) will change directly as a consequence of drivers of Global Change, and indirectly as a consequence of changes in ecosystem structure (e.g., composition, life forms, etc.). The latter tend to be slow and lag behind environmental change, making difficult any experimental or observational approach to understanding these dynamics. Models are therefore required to project the longerterm consequences of environmental change on ecosystems and the entire Earth System.

Models may need to be calibrated, and their projections need to be evaluated against independent data. Calibration denotes the process of fitting model output to a set of observational data e.g. by manipulating those model parameters that cannot be estimated unequivocally from direct measurements. The failure to find a good fit between model predictions and observations may be indicative of a lack of understanding of the system, which often leads to changes in the model itself. Evaluation (often also termed “validation”) refers to the process of testing model projections against data sets that were not used to estimate the model structure or the model’s parameters. Failure to pass a model evaluation test usually leads to model reformulation, whereas successful model evaluations usually are followed by model applications, e.g. studies of Global Change impacts under scenarios of driving variables such as climate, atmospheric chemistry, or land use.

Models of longterm ecosystem changes are best calibrated or evaluated based on long records of biotic changes. A key source of such records are environmental archives. Below, the current state of research is highlighted based on a few examples of model evaluation activities in one particular and widespread class of models, the “forest gap models” (Shugart 1984). These models simulate the establishment, growth and mortality of individual trees on small

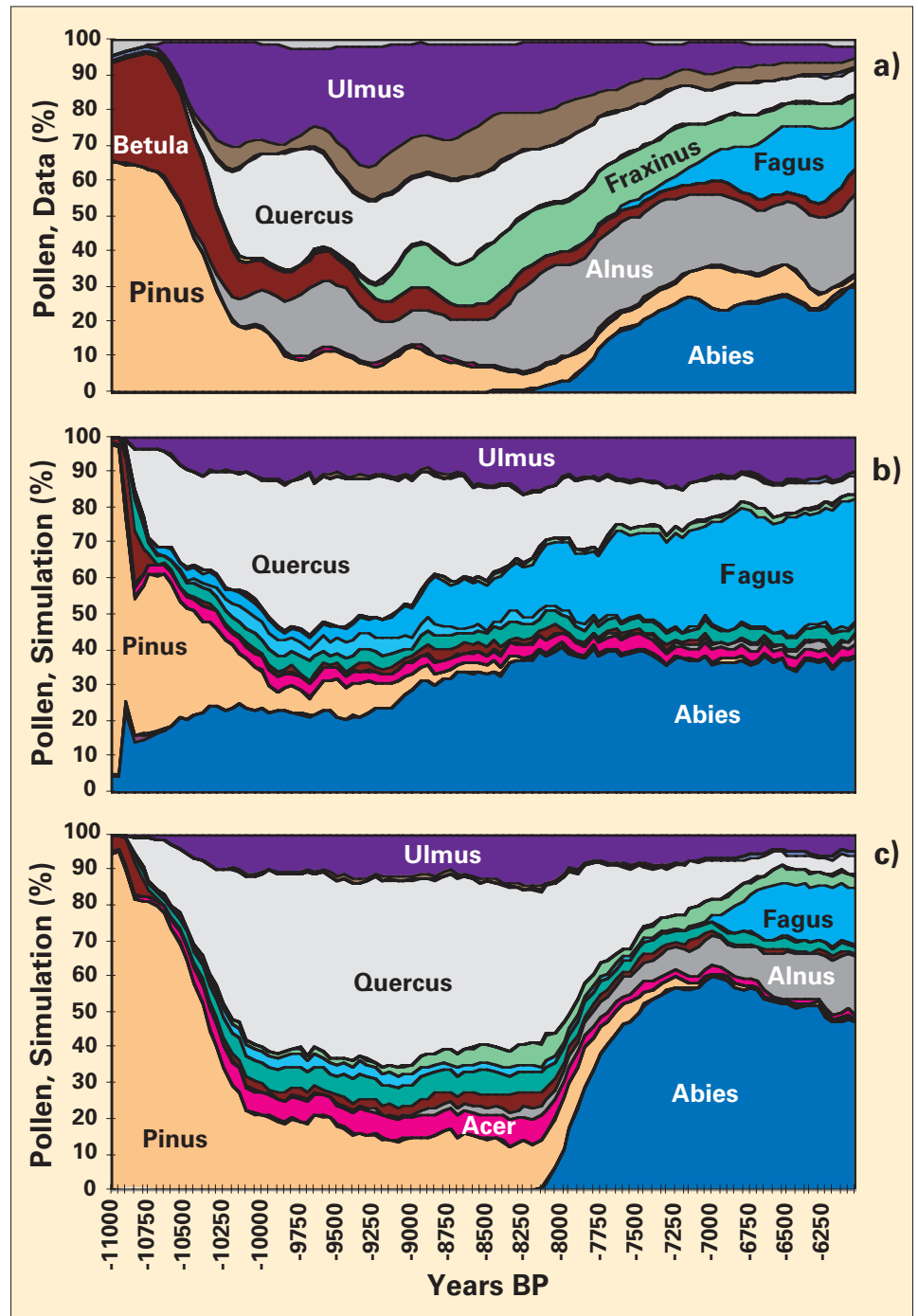


Figure 1: Pollen frequencies (ordinate, %) in the early Holocene (abscissa shows years BP) for Lake Soppensee, Switzerland (after Lotter 1989). a) Measurements from an annually laminated pollen profile. b) Simulation results of the FORCLIM forest gap model based on independently derived climatic input data (temperature and precipitation). c) as in b), but with postulated immigration dates for *Abies* and *Fagus* that were derived from the pollen record in a) (from Lischke 1998).

patches of land as a function of biotic (competition) and abiotic (climatic) factors. The mortality of a large, dominant tree produces a gap in the forest, which leads to the release of suppressed trees and enhanced recruitment, both of which drive succession; thus the term “gap” models (for a review, cf. Bug-

mann 2001). Gap models produce estimates of species-specific biomass, basal area, or leaf area index at the scale of a forest stand (i.e., 110 ha).

Although the first evaluations of gap models with paleo data were performed 20 years ago, we have not yet fully explored the potential of using paleo

data in a modeling framework, as will be shown below.

Pollen Records

Pollen data from mires and lakes have been used to evaluate the behavior of gap models on the time scale of millennia. This approach was pioneered by Solomon *et al.* (1980) and Solomon and Webb (1985). A distinct advantage is that it allows one to assess forest dynamics on very long time scales. However, pollen data usually have a low temporal resolution, making it impossible to determine rates of change that occur on time scales relevant for understanding Global Change in the 21st century. In addition, there are a number of methodological problems that have only partially been resolved in these modeling studies:

- The simulated data refer to a forest stand that is homogeneous with respect to climate and soil conditions, whereas pollen data most often represent a larger, typically heterogeneous area (catchment scale or larger). Detailed knowledge of soils and climate across the pollen source area as well as pathways of pollen transport would be required to set up simulation experiments that adequately reflect the pollen record; while such knowledge is available from paleo research, it has not been used by the modeling community yet.
- Measured records provide pollen frequencies, whereas the models simulate absolute values of species-specific biomass, basal area, leaf area index, etc., but not pollen production. To make a comparison of measured with simulated variables possible, Iversen factors (i.e., a simple linear model calibrated under current climatic and landuse conditions; Faegri and Iversen 1975) have often been used to estimate pollen production from simulated biomass. This procedure may introduce an error that is quite difficult to quantify, but likely to be significant.
- In most modeling studies, climate was either kept constant, which is not realistic across time spans of several millennia, or scenarios of climatic change were derived from the same data that were used for evaluating the behavior of the model, introducing a certain extent of circu-

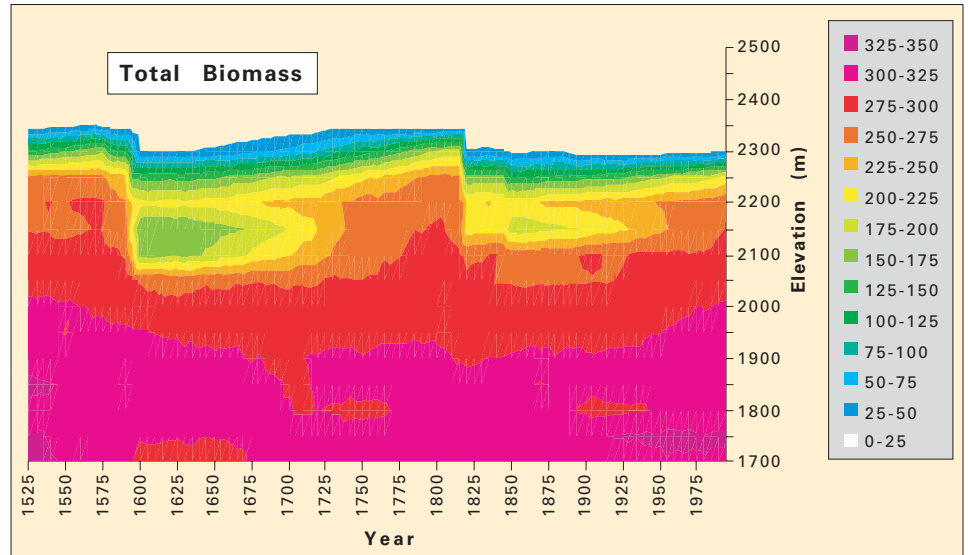


Figure 2: Contour plot of simulated total aboveground biomass (t/ha) along an elevational gradient near Davos (Switzerland) for the period 1525–1995 AD. From Bugmann and Pfister (2000).

lar reasoning. There are only a few examples of studies that were based on a reconstruction of climate that is entirely independent of the biotic record that was used for model data comparison.

For example, in a study that addressed the last of the above constraints, Lischke (1998) used an independent reconstruction of climate to drive a forest succession model at a site on the Swiss Plateau for which an annually laminated sediment record was available (Fig. 1). The simulation revealed broad agreement between simulated and measured pollen frequencies when immigration of two genera, *Abies* and *Fagus*, was delayed to the point where they appeared in the pollen record. While the simulation suggested that nonclimatic factors (such as the maximum migration rate) may be responsible for the late appearance of the two genera, the other differences between simulated and measured data are more difficult to interpret. They might derive from inadequacies in the pollen transfer functions, in the representation of the landscape by a single climatically and edaphically homogeneous site, or from inadequacies in the forest model itself. Clearly, a closer linkage between ecosystem modelers and palynologists is required to resolve these issues.

Tree Rings

Tree rings represent another data source that is characterized by long records, but they have a higher temporal resolution (annual, seasonal, or even monthly).

A distinct difference between pollen and tree-ring data is that tree rings provide an individual-based measure from which standscale growth patterns can be reconstructed across time, whereas it is more difficult to assess long-term changes of plant abundances (i.e., population dynamics).

Ring width corresponds directly to diameter increment, which is the central state variable in many gap models. Tree-ring data thus seem to be well suited for evaluating simulated growth patterns. Standardized tree-ring chronologies could be derived from simulated tree growth using dendrochronological procedures (e.g., growth trend removal). However, there are only a few studies that attempted to use tree-ring data for the evaluation of gap models (e.g., Keane *et al.* 1997, Bugmann and Pfister 2000), and there is a large potential for further exploration of this method. Because precise input data are required for driving a gap model in this mode of application, the comparison will not usually cover more than 100–150 years, but even such relatively short time series can be useful for evaluating simulated patterns with respect to the interannual or decadal variability of growth patterns.

In a larger effort to evaluate the sensitivity of a forest model to changes in climate variability, Bugmann and Pfister (2000) used the CLIMINDEX database of monthly thermic and hygric indices (Pfister 1999) to derive a climate reconstruction along a transect that crosses upper treeline in the European Alps (Fig. 2). The occurrence of the two tree-

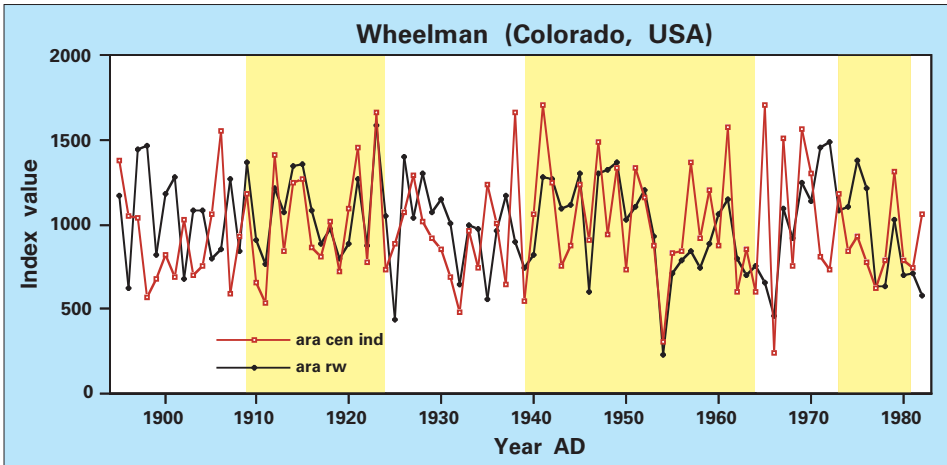


Figure 3: Comparison of tree-ring indices (rw) measured for *Pinus ponderosa* at Wheelman (Colorado, USA; cf. Kienast and Schweingruber 1986) with NPP indices (cen ind) simulated by the CENTURY model based on time series of weather data from the VEMAP database (Kittel et al. 1995) interpolated to the Wheelman site. Yellow areas denote periods with a good match between observations and simulation (H. Bugmann et al., unpublished data).

line dieback events and the growth patterns of *Picea abies* in the subalpine zone were evaluated against a number of qualitative and quantitative datasets, including several dozen treering chronologies from the Alps available from the International Tree Ring Data Bank. These tests indicated reasonable agreement between historical evidence and simulated behavior, but they were far from being conclusive, pointing to the need to conduct further tests to evaluate the usefulness of the model for studying the impacts of climate variability on tree population dynamics.

Finally, in an ongoing study treering data are being used to evaluate the inter-annual variability of ecosystem model responses across the 20th century in the context of VEMAP (Vegetation/Ecosystem Modeling and Analysis Project; Schimel et al. 1997). Bugmann et al. (unpublished) used the spatially and temporally explicit VEMAP climate database to provide sitespecific climatic input data for the CENTURY ecosystem model (Parton et al. 1994), and compared a simulated index of net primary productivity (NPP) against measured, sitespecific treering widths (Fig. 3). The results suggested that there is reasonable agreement between simulated and observed patterns across parts of the simulated time span, but the model had difficulty capturing some aspects of the measured time series. It is not clear at present whether the differences (Fig. 3) are due to methodological problems (such as comparing NPP to treering widths, the standardization procedures

that were used, the derivation of climatic input data, etc.), or whether they point at model inadequacies. Again, further research is required to develop methodologies for these modeldata comparisons.

Conclusion

It is evident from the above considerations that there is a certain mismatch between the resolution and quantity of output variables provided by ecological impact models and the availability of long measured time series that could serve as a source of climatic input data for the models and to evaluate their predictions against truly independent data. Even with an improved collaboration between paleoecologists and modelers, it is unlikely that there will ever be a single data source that can be used to fully evaluate models of longterm ecological processes. Rather, we should strive to combine several data sources for evaluating both process formulations and the overall projections that result from any given model. Long time series of paleoecological data from palynological and dendroecological studies could play a very important role in this effort.

HARALD BUGMANN

Mountain Forest Ecology, Department of Forest Sciences, Swiss Federal Institute of Technology Zurich, Switzerland
bugmann@fowi.ethz.ch

For full references please consult www.pages-igbp.org/products/newsletters/ref2003.html

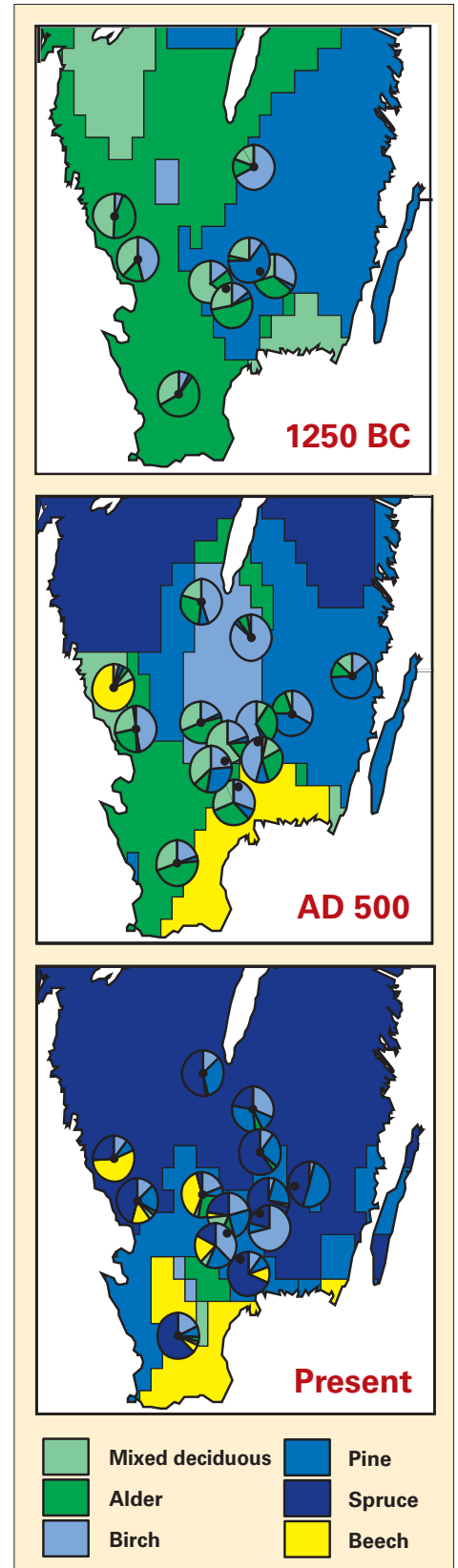


Figure 1: Forest maps of southern Sweden from 1250 BC, AD 500 and present based on regional pollen sites. The pie diagrams show data from stand-scale sites.