1
Introduction

This book is about understanding vegetation systems in a scientific context, one topic of vegetation ecology. It is written for researchers motivated by the curiosity and ambition to assess and understand vegetation dynamics. Vegetation, according to van der Maarel (2005) 'can be loosely defined as a system of largely spontaneously growing plants.' What humans grow in gardens and fields is hence excluded. The fascination of investigating vegetation resides in the mystery of what plants 'have in mind' when populating the world. The goal of all efforts in plant ecology, as in other fields of science, is to learn more about the rules governing the world. These rules are causing patterns, and the assessment of patterns is the recurrent theme of this book.

Unfortunately, our access to the *real world* is rather restricted and – as we know from experience – differs among individuals. To assure progress in research an image of the real world is needed: the *data world*. In this we get a description of the real world in the form of numbers. (An image can

Data Analysis in Vegetation Ecology Otto Wildi

be a spreadsheet filled with numbers, a digital photograph or a digital terrain model.) Upon analysis we then develop our *model world*, which represents our understanding of the real world. Typical elements are orders, patterns or processes governing systems. It is the aim of most analytical methods to identify patterns as elements of our model view.

Finding models reflecting the real world is a difficult task due to the complexity of systems. Complexity has its origin in a number of fairly well known phenomena, one being the scale effect. Any regularity in ecosystems will emerge at a specific spatial and temporal scale only: at short spacial distance competition and facilitation among plants can be detected (Connell & Slatyer 1977); these would remain undetected over a range of kilometres. In order to study the effect of global climate change (Orlóci 2001, Walther *et al.* 2002) the scale revealed by satellite photographs is probably more promising. Choosing the best scale for an investigation is a matter of decision, experience and often trial and error. For this a multi-step approach is needed, in which intermediate results are used to evaluate the next decision in the analysis. Poore (1955, 1962) called this *successive approximation* and Wildi & Orlóci (1991) *flexible analysis*. Hence, the variety and flexibility of methods is nothing but an answer to the complex nature of the systems. Once the proper scale is found there is still a need to consider an 'upper' and a 'lower' level of scale, because these usually also play a role. Parker & Pickett (1998) discuss this in the context of temporal scales and interpret the interaction as follows: 'The middle level represents the scale of investigation, and processes of slower rate act as the context and processes of faster rates reflect the mechanisms, initial conditions or variance.'

A second source of complexity is uncertainty in data measured. Data are restricted by trade-offs and practical limitations. A detailed vegetation survey is time-consuming, and while sampling, vegetation might already be changing (Wildi *et al.* 2004). Such data will therefore exhibit an undesired temporal trend. A specific bias causes variable selection. It is easier to measure components above ground than below ground (van der Maarel 2005, p. 6), a distinction vital in vegetation ecology. Once the measurements are complete they may reflect random fluctuation or chaotic behaviour (Kienast *et al.* 2007) while failing to capture deterministic components. It is a main objective in data analysis to distinguish random from deterministic components. Even if randomness is controlled there is *nonlinearity* in ecological relationships, a term used when linearity is no longer valid. This would not be a problem if we knew the kind of relationship that was hidden in the data (e.g. Gaussian, exponential, logarithmic, etc.), but finding a proper function is usually a challenging task.

Further, spatial and temporal interactions add to the complexity of vegetation systems. In space, the problem of order arises, as the order of objects depends on the direction considered. In most ecosystems, the environmental conditions, for example elevation or humidity, change across the area. Biological variables responding to this will also be altered and become *space-dependent* (Legendre & Legendre 1998). If there is no general dependency in space, a local phenomenon may exist: *spatial autocorrelation*. This means that sampling units in close neighbourhood are more similar than one could expect from ecological conditions. One cause for this comes from biological population processes: the chance that an individual of a population will occur in unfavourable conditions is increased if another member of the same population resides nearby. It will be shown later in this book how such a situation can be detected (Section 7.3.3). Similarly, correlation over time also occurs. In analogy to space, there is *temporal dependence* and *temporal autocorrelation*. This comes from the fact that many processes are temporally continuous. The systems will usually only change gradually, causing two subsequent states to be similar. Finally, time and space are not independent, but linked. Spatial patterns tend to change continuously over time. In terms of autocorrelation, spatial patterns observed within a short time period are expected to be similar. Similarly, a time series observed at one point in space will be similar to another series observed nearby.

In summary, all knowledge we generate by analysing the data world contributes to our model world. However, this is aimed at serving society. When translating this into practice we experience yet another world, the man-made *world of values*. This is people's perception and valuation of the world, which we know from experience is continuously changing. The results we derive in numerical analysis carry the potential to deliver input into value systems, but we should keep in mind what Diamond (1999) mentioned when talking about accepting innovations: 'Society accepts the solution if it is compatible with the society's values and other technologies.' Proving the existence of global warming, as an example, can be a matter of modelling. Convincing people of the practical relevance of the problem is a question of evaluation and communication, for which different skills may be required.