

27 Univariate and multivariate analysis applied on a Dutch sandy beach community

Janssen, G.M., Mulder, S., Zuur, A.F., Ieno, E.N. and Smith, G.M.

27.1 Introduction

Climate change is, beyond doubt, the most important threat facing the world's coastline and has been accompanied by intensive debate. Marine coastal ecosystems are extremely vulnerable, as they constitute the most productive and diverse communities on Earth. Coastal areas are, however, not only subject to climate change but also to many other forms of human activities. Among them, land-claim, pollution, recreation purposes and dredging activities have been threatening most of the European coasts resulting in many cases in inter-tidal habitat fragmentation and/or degradation (Raffaelli and Hawkins 1996). The consequences of these changes have been well documented in a considerable number of studies that addressed the impact and reported decreased ecosystem performance.

On a more local scale, the Dutch have fought great battles with the North Sea in order to extend their landmass as can be witnessed by the presence of dykes and sophisticated coastal defence systems. The effect of sea level rise on the ecology of the Dutch coastal system constitutes a serious issue that should not be ignored in the short term.

The Dutch governmental institute RIKZ therefore started a research project on the relationship between some abiotic aspects (e.g., sediment composition, slope of the beach) as these might affect benthic fauna. Mulder (2000) described the results of a pilot study that looked at the effects of differences in slope and grain size on fauna in the coastal zone. Using the data from this pilot study and statistical experimental design techniques, a sampling design was developed in which nine beaches were chosen by stratifying levels of exposure: three beaches with high exposure, three beaches with medium exposure and three beaches with low exposure. Sampling was carried out in June 2002, and at each beach, five stations were selected. Effort at each station was low (Van der Meer 1997).

The aim of the project was to find relationships between macrofauna of the intertidal area and abiotic variables. In this case study chapter, univariate and multivariate tools are applied in order to obtain as much information as possible. The results of the combined analyses are then used to answer the underlying question. Instead of presenting the results of the final models, we show how we got to them,

and which steps we applied (especially for the multivariate analysis). The outcome of this research will have immediate relevance for assessing and managing disturbance in marine benthic systems, with respect to degradation and biodiversity lost.

27.2 The variables

Table 27.1 gives a list of the available explanatory variables. NAP is the height of the sampling station relative to the mean tidal level. Exposure is an index that is composed of the following elements: wave action, length of the surf zone, slope, grain size and the depth of the anaerobic layer. Humus constitutes the amount of organic material. Sampling took place in June 2002. A nominal variable 'week' was introduced for each sample, which has the values 1, 2, 3 and 4, indicating in which week of June a beach was monitored. The following rules were used. Sampling between 1 and 7 June: $Week_i = 1$. Sampling between 8 and 14 June: $Week_i = 2$. Sampling between 15 and 22 June: $Week_i = 3$ and sampling between 23 and 29 June: $Week_i = 4$. The index i is the station index and runs from 1 to 45. There were nine beaches, and on each beach five stations were sampled (hence, 45 observations). Ten sub-samples were taken per station, but in this chapter we will use totals per station. $Angle_1$ represents the angle of each station, whereas $angle_2$ is the angle of the entire sampling area on the beach. Both variables were used. The variables $angle_2$, exposure, salinity and temperature were available at beach level. This means that it is assumed that each station on a beach has the same value. For $angle_2$ this assumption does not hold, and its inclusion in some of the statistical models should be done with care. A few explanatory variables contained four missing values. Most of the statistical techniques used in the analysis cannot cope with missing values, and therefore, missing values were replaced by averages.

Table 27.2 shows the Pearson correlation among the 12 explanatory variables. Only correlations significant at the 5% level are given. Except for a few variables (chalk and sorting and exposure and temperature), the correlations among the explanatory variables are relatively low, indicating that there is no serious collinearity.

As to the species, in total 75 species were measured. To simplify interpretation of graphical plots, species were grouped in the following five taxa: Chaetognatha, Polychaeta, Crustacea, Mollusca, and Insecta. Within each taxa, between 1 and 28 species were available. Species names were replaced by names taking the following form: P_1, P_2, P_3 , etc. for Polychaeta species, CR_1, CR_2, CR_3 , etc. for Crustacea species, M_1, M_2, M_3 , etc. for Mollusca species, I_1, I_2, I_3 for Insecta, and C_1 for the Chaetognatha species *Sagitta* spec. (only one Chaetognatha was measured). A list of species and notation used in this chapter are available as an online supplement to this book.

Table 27.1. List of available explanatory variables. The columns labelled “Level” indicate whether different values are available for each station on a beach (Beach) or one value for all five stations on a beach (Station).

Number	Variable	Level	Units
1	Week	Beach	-
2	Angle ₁	Station	-
3	Angle ₂	Beach	-
4	Exposure	Beach	-
5	Salinity	Beach	‰
6	Temperature	Beach	°C
7	NAP	Station	m
8	Penetrability	Station	N/cm ²
9	Grain size	Station	mm
10	Humus	Station	%
11	Chalk	Station	%
12	Sorting	Station	Mm

Table 27.2. Significant correlations ($\alpha = 0.05$) among explanatory variables.

	Week	Angle ₁	Angle ₂	Exp.	Sal.	Temp.	NAP	Pen.	Grains	Hum	Chalk	Sort.
Week	1											
Angle ₁		1										
Angle ₂			1									
Exp.				1								
Sal.					1							
Temp.						1						
NAP							1					
Pen.								1				
Grains									1			
Hum										1		
Chalk											1	

27.3 Analysing the data using univariate methods

In this section, the species data are analysed by converting them into a diversity index (Magurran 2004). We used the Shannon–Weaver index (with base \log_{10}). The shape of the Shannon–Weaver index was similar to that of the species richness. Because of this similarity the Shannon–Weaver index can also be seen as an indicator for the number of different species (at least for these data). Figure 27.1 shows a dotplot for the Shannon–Weaver index. Stations are grouped by beach. There are no stations with considerably larger or smaller values.

We now compare the Shannon–Weaver index with the explanatory variables. Quinn and Keough (2002) followed a similar approach and applied regression techniques on the diversity index. Table 27.3 shows the correlation between the Shannon–Weaver index and each of the 12 explanatory variables.