

Chapter 5

Space–Time Geostatistics for Precision Agriculture: A Case Study of NDVI Mapping for a Dutch Potato Field

G.B.M. Heuvelink and F.M. van Egmond

Abstract Many environmental variables that are relevant to precision agriculture, such as crop and soil properties and climate, vary both in time and space. Farmers can often benefit greatly from accurate information about the status of these variables at any particular point in time and space to aid their management decisions on irrigation, fertilizer and pesticide applications, and so on. Practically, however, it is not feasible to measure a variable exhaustively in space and time. Space–time geostatistics can be useful to fill in the gaps. This chapter explains the basic elements of space–time geostatistics and uses a case study on space–time interpolation of the normalized difference vegetation index (NDVI) as an indicator of biomass in a Dutch potato field. Space–time geostatistics proves to be a useful extension to spatial geostatistics for precision agriculture, although theoretical as well as practical advances are required to mature this subject area and make it ready to be used for within-season, within-field decision making by farmers.

Keywords Space–time geostatistics · Variogram · Kriging · NDVI · Interpolation · Spatial variability · Temporal variability · Prediction · Uncertainty · Mapping · GPS · Crop growth · Sampling

5.1 Introduction

Many of the soil properties that are relevant to precision agriculture vary both in space and time. Examples are soil moisture, soil nutrient concentrations and pH. Clearly, variation in space–time is not restricted to the soil, but extends to other domains such as crops and climate. For instance, the biomass and protein concentration

G.B.M. Heuvelink (✉)

Environmental Sciences Group, Wageningen University and Research Centre,
PO Box 47, 6700 AA Wageningen, The Netherlands
e-mail: Gerard.Heuvelink@wur.nl

F.M. van Egmond

The Soil Company, Leonard Springerlaan 9, 9727 KB Groningen, The Netherlands
e-mail: fenny@medusa-online.com

of the root, stem and leaves of a crop depend on location and time, and so do precipitation and temperature when the study area is large. Farmers can benefit greatly from accurate information about the status of these variables at any particular point in time and space. It can aid their management decisions on irrigation, fertilizer and pesticide applications and so on. Practically, however, it is not feasible to measure a variable exhaustively in space and time. For example, [Snepvangers et al. \(2003\)](#) used time domain reflectometry to record about 100,000 observations of the topsoil water content in a 0.36 ha grassland plot during a 30 day period, but the resulting data were still sparse and interpolation was needed to describe the space–time variation adequately. In many practical cases, even in precision agriculture where measurements may be abundant, the size of the dataset will be much smaller than in the example above and the space–time domain will be much larger. Thus, a common problem is to construct a high-resolution space–time representation of a variable that varies in space and time from a limited number of observations. Space–time geostatistics can be useful for this.

Space–time geostatistics is a natural extension of ‘spatial’ geostatistics as described in Chapter 1. It begins by characterizing the variation in space and time with variograms. Next, these variograms are used to predict the target variable at unmeasured points by kriging. The prediction error can be quantified and trends incorporated to reduce this error. Trends in space and or time are incorporated when part of the variation in the target variable can be explained by explanatory variables, such as when soil type is used to explain spatial variation in soil texture or when precipitation is used to explain temporal variation in soil moisture. The difference from ‘spatial’ geostatistics is that variation occurs in space and time, and both these sources of variation must be modelled and their effects on prediction taken into account. Variation in space might be much less than that in time. For instance, rainfall events are likely to pass over the entire field, but they vary considerably over time. The opposite is also common; soil bulk density can differ markedly between locations with different types of soil or landuse, but it changes little over a timespan of a few years. In fact, the bulk density example hints at a situation where temporal variation may be negligibly small compared to spatial variation. In such a case one might decide to disregard the temporal variation completely and return to conventional ‘spatial’ geostatistics. When spatial variation is negligible a time-series analysis will suffice. Thus, space–time geostatistics is required when neither spatial nor temporal variation can be ignored, but this does not mean that both must be equally large. Also, the lengths of spatial and temporal dependence will be different, if only because space and time have different measurement units.

This chapter explains the basic elements of space–time geostatistics and uses a case study on space–time interpolation of the normalized difference vegetation index (NDVI) as an indicator of biomass in a Dutch potato field. We begin with a description of the study site and discuss the positional correction of GPS data from the field, followed by an exploratory data analysis of the NDVI data. Next we present the theory of space–time geostatistics. We apply it to the case study, discuss results and draw conclusions.

5.2 Description of the Lauwersmeer Study Site and Positional Correction of NDVI Data

The study site is near the village of Vierhuizen in the Lauwersmeer area to the North of the Netherlands at 6.27° longitude and 53.35° latitude. The climate is temperate maritime; the annual rainfall in 2006 was 661 mm with an annual average temperature of 11.7°C . In 2006, the growing season for potatoes was relatively warm and dry except for August, which was colder and wetter than usual. The soil is characterized as a fluvisol formed on young calcareous marine sediments. Soil texture ranges from loamy sand to sandy clay loam with 1–4% organic matter. Elevation ranges from 0.5 to 1.5 m above sea level with lower areas on the W and SE parts of the field (Fig. 5.1). The lower parts of the field are more sandy. Typical crop rotations consist of seed and consumption potatoes, sugar beet, wheat and onions.

A 10 ha field with two potato varieties, Innovator and Sofista, was studied during the growing season of 2006. Innovator is a mid-early variety with fast and good biomass development, whereas Sofista is classified as mid-late with good biomass development.

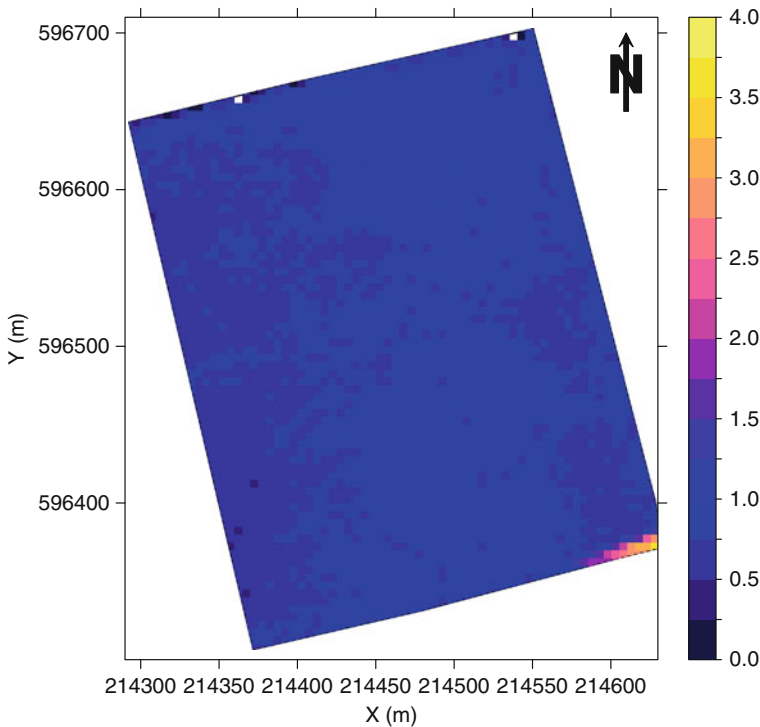


Fig. 5.1 Digital elevation map (metres above sea level) of the study site, obtained from the Dutch AHN (Actueel Hoogtebestand Nederland 2009)

A Crop Circle sensor (Holland Scientific 2008) measured spectral reflectance each time the crop was sprayed against phytophthora, a potato disease. The active spectral reflectance sensor has a spatial support of about 1 m². It emits at 650 and 880 nm and measures reflectance in the visible (between 400 and 680 nm) and near-infrared part of the spectrum (800–1100 nm). The sensor calculates the NDVI based on those reflectances. The sensor- and GPS data are logged every second. Data from headlands are not taken into account. The number of tramlines recorded and the driving speed differs with measurement date (Table 5.1).

The NDVI data were obtained from reflectances recorded on 16 different days during the growing season of 2006. The Crop Circle sensor was mounted 8.5 m behind the tractor on a spraying boom and 6.5 m left of the centre. The GPS was located on the tractor roof. The difference in position between the sensor and GPS caused an error in the positions of recorded values that depends on the driving direction. The GPS coordinates logged were corrected for this difference (Fig. 5.2a). The accuracy of a GPS is fairly constant during any one day, but locations may shift slightly when measuring the same coordinates on different days. This effect is clearly visible from the wide transects in Fig. 5.2a. The farmer measures NDVI when he sprays the field and uses exactly the same tramlines each time to minimize crop damage. Hence the corrected coordinates measured on different days should be positioned along the same lines. The corrected coordinates were therefore corrected a second time for the ‘temporal’ GPS error. When the coordinates of a tramline differed by more than 0.5 m from the centre of the tramline, they were corrected in a perpendicular direction to the driving direction to match the centre of the tramline and to minimize unwanted displacement of the coordinates in the driving direction. The results are shown in Fig. 5.2b.

5.3 Exploratory Data Analysis of Lauwersmeer Data

The main interest of the farmer is in the distribution of the potato growth over time and space. Accurate information on growth and variation in growth enables the farmer to fine-tune fertilizer application and the timing and amount of chemicals to remove above ground biomass at the end of the growing season. The space–time distribution of below ground biomass, and hence tuber yield, is correlated to the above ground green biomass, of which NDVI is an indicator (Baret and Guyot 1991; Carlson and Ripley 1997). The values of NDVI range between 0 and 1. An NDVI of 0.2 indicates bare soil, a value of 0.65 indicates that the crop canopy is closed and a further increase in NDVI indicates an increase in the number of green leaf layers. The index ‘saturates’ at 0.9 which represents a leaf area index (LAI) of 3–4. Throughout the growing season NDVI increases until tuber formation starts and then it decreases as the above ground biomass dies and becomes yellow (Wu et al. 2007).

Figure 5.3 shows a scatter plot of NDVI against day of year (DOY). The NDVI increases rapidly in spring and gradually decreases towards the end of the summer. The spatial variation is large for all dates, as indicated by the wide spread in values. There appear to be outliers with small NDVI values when overall NDVI values

Table 5.1 Descriptive statistics of the Lauwersmeer farm data

DOY	Crop type	Number of tramlines	Number of observations	Mean	Median	Minimum	Maximum	Standard deviation	Skewness	Correlation with NDVI
NDVI										
149	Sofista	1	353	0.08	0.08	0.06	0.12	0.01	0.21	
149	Innovator	5	1689	0.16	0.16	0.05	0.32	0.04	-0.07	
154	Sofista	3	1237	0.19	0.19	0.07	0.29	0.03	0.10	
154	Innovator	5	1408	0.26	0.27	0.06	0.37	0.05	-0.75	
160	Sofista	1	370	0.28	0.28	0.08	0.40	0.05	-0.27	
160	Innovator	5	1834	0.44	0.44	0.13	0.59	0.05	-0.67	
167	Innovator	5	1605	0.69	0.70	0.08	0.75	0.05	-7.44	
168	Sofista	4	1304	0.60	0.61	0.28	0.73	0.06	-1.29	
174	Sofista	4	1231	0.72	0.72	0.45	0.77	0.03	-3.20	
174	Innovator	5	1532	0.75	0.76	0.09	0.79	0.04	-11.00	
181	Sofista	4	1312	0.76	0.76	0.35	0.79	0.02	-12.71	
181	Innovator	5	1471	0.78	0.78	0.28	0.80	0.03	-13.62	
187	Sofista	4	971	0.77	0.78	0.50	0.81	0.02	-6.10	
187	Innovator	5	1274	0.79	0.79	0.12	0.81	0.03	-17.21	
193	Innovator	5	1430	0.78	0.78	0.74	0.80	0.01	-0.49	
196	Sofista	4	1263	0.79	0.79	0.58	0.82	0.02	-3.10	
202	Sofista	4	850	0.78	0.78	0.64	0.82	0.02	-1.49	
202	Innovator	5	1210	0.80	0.80	0.70	0.82	0.01	-2.42	
208	Sofista	4	966	0.75	0.76	0.58	0.83	0.04	-0.95	
208	Innovator	5	1299	0.77	0.77	0.15	0.80	0.04	-13.14	
217	Sofista	3	747	0.73	0.73	0.50	0.81	0.04	-0.76	
217	Innovator	5	1248	0.78	0.78	0.67	0.81	0.01	-1.07	
224	Sofista	4	755	0.70	0.71	0.53	0.79	0.05	-0.83	
224	Innovator	5	1074	0.77	0.77	0.41	0.80	0.02	-5.02	

(continued)

Table 5.1 (continued)

	DOY	Crop type	Number of tramlines	Number of observations	Mean	Median	Minimum	Maximum	Standard deviation	Skewness	Correlation with NDVI
NDVI	231	Sofista	4	940	0.65	0.68	0.34	0.79	0.09	-1.05	
	231	Innovator	6	1275	0.77	0.77	0.57	0.81	0.02	-1.77	
	243	Sofista	4	1295	0.33	0.30	0.17	0.78	0.12	1.24	
	243	Innovator	5	1806	0.74	0.75	0.11	0.80	0.03	-10.37	
Clay (%)				33749	15.80	16.26	7.50	22.50	3.13	-0.44	0.035
Sand (%)				33749	62.27	60.70	55.00	84.41	6.59	0.87	-0.029
M0 (%)				33749	68.93	68.60	60.00	84.96	5.27	0.37	-0.030
Organic matter (%)				33749	2.61	2.56	1.20	4.20	0.73	0.16	0.009
Bulk density				33749	1.43	1.43	1.33	1.55	0.04	0.34	-0.026
Elevation (m ASL)				92152	0.812	0.810	-0.1	3.83	0.165	6.88	0.005

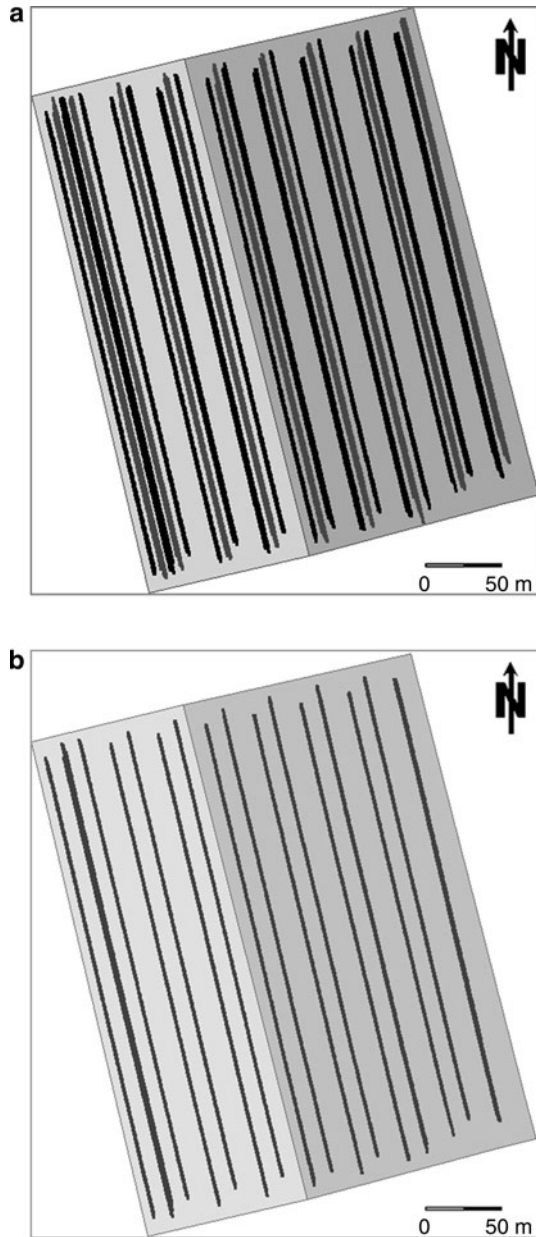


Fig. 5.2 (a) The original GPS coordinates (dark grey) and the coordinates after the first correction (*black*) and (b) GPS coordinates after second correction. Background colour represents the two potato species (*Sofista light grey*; *Innovator dark grey*)

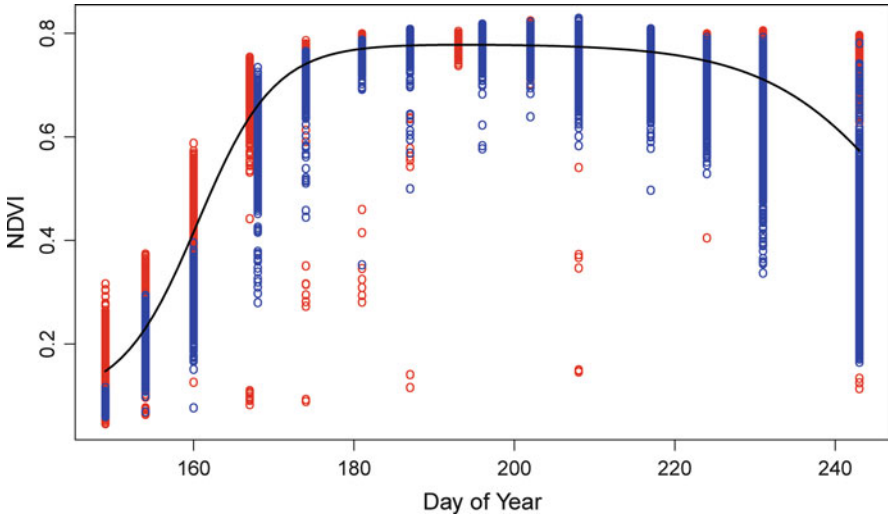


Fig. 5.3 The NDVI observations (Innovator in red, Sofista in blue) against day of year in 2006. The solid line is the fitted temporal trend

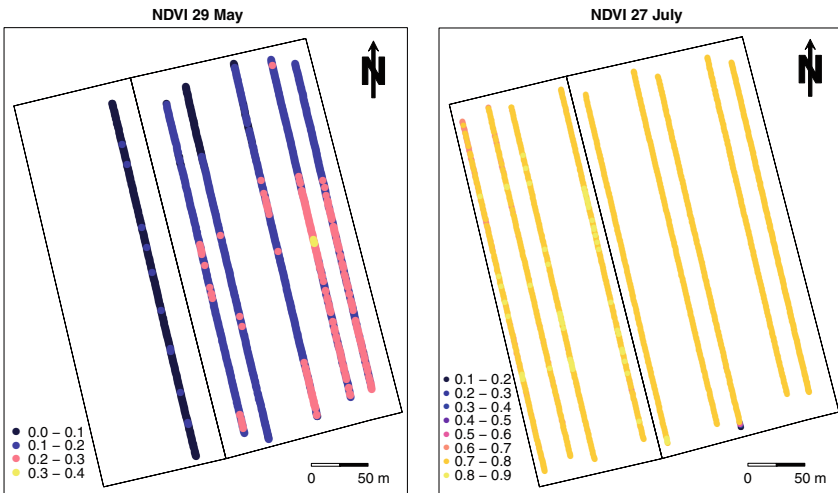


Fig. 5.4 Spatial distribution of NDVI observations for May 29 (DOY 149) and July 27 (DOY 208)

are large, showing that the crop was not fully developed in part of the field. This is confirmed in Fig. 5.4, where the spatial distribution of NDVI is shown for two dates. There are also a few small values in one area near the field boundary on July 27. Note also that the Sofista crop has developed little biomass on May 29. Figure 5.3 shows that the Sofista crop develops later than Innovator and starts to show a decline in above ground biomass earlier. Innovator has a larger average NDVI in late summer when Sofista has already started to decline.

Factors that might influence crop growth, and hence NDVI, can be divided into temporally constant but spatially varying factors and spatially constant but temporally varying factors. The first category includes soil properties, elevation and the spatial allocation of the two different crop types. Soil properties that might influence potato growth are the amount of clay, sand and organic matter, median grain size (M0), potential bulk density and potential water retention. Calculation of the latter two is based on pedotransfer functions (Wösten et al. 2001). Summary statistics of the NDVI, soil and elevation data are given in Table 5.1. Digital elevation data are available at 1-m resolution from the AHN (Actueel Hoogtebestand Nederland 2009) the national altitude dataset of the Netherlands, Fig. 5.1. The second category of spatially constant explanatory variables includes precipitation, temperature and fertilizer application. The crop received a spatially homogeneous nitrogen application on 14 June 2006 (DOY 155).

5.4 Space–Time Geostatistics

Consider a variable $z = \{z(\mathbf{s}, t) | \mathbf{s} \in S, t \in T\}$ that varies within a spatial domain S and a time interval T . Let z be observed at n space–time points (\mathbf{s}_i, t_i) , $i = 1, \dots, n$. In the case study described in Sections 5.2 and 5.3, these space–time observations are transects of observations along the tramlines recorded at several instants in time. Although the total number of observations, n , may be very large, to observe z at each and every combination of time and space is not feasible. To obtain a complete space–time surface of z requires some form of prediction. Therefore, the objective is to obtain a prediction of $z(\mathbf{s}_0, t_0)$ at a point (\mathbf{s}_0, t_0) at which z was not observed, where (\mathbf{s}_0, t_0) typically is associated with the nodes of a fine space–time grid. To do this, z is assumed to be a realization of a random function Z (Webster 2000). The random function Z is characterized by a statistical model that must describe the structure of dependence in space–time. Once the model is fully defined, $Z(\mathbf{s}_0, t_0)$ may be predicted from the observations by kriging as described for the spatial variables in Chapter 1.

The space–time variation of Z can be characterized by first decomposing it into a deterministic trend m and a zero-mean stochastic residual ε as follows:

$$Z(\mathbf{s}, t) = m(\mathbf{s}, t) + \varepsilon(\mathbf{s}, t). \quad (5.1)$$

The trend m is a deterministic, structural component that represents large scale variation. The residual is a stochastic component representing small scale, ‘noisy’ variation. Alternatively, the trend may be thought of as that part of Z that can be explained physically or empirically by auxiliary information. The residual component still holds important information on the variation when it is correlated in space and or time. Note that the decomposition of Z into a trend and a residual is a scale-dependent, subjective choice made by a modeller (Diggle and Ribeiro 2007).

5.4.1 *Characterization of the Trend*

As noted in Section 5.1, often the behaviour of a variable over time is entirely different from its behaviour in space. This difference can be represented by the trend in the specification of a space–time model. For instance, crop growth in an agricultural field is causally dependent on factors such as soil, climate and management. This information should be incorporated into the model if possible when predicting crop growth or biomass accumulation in space and time. Ideally, the trend would be a process-oriented, physical-deterministic model. However, when deterministic modelling is not feasible due to a lack of understanding of the underlying governing processes or because external forces and boundary conditions are unknown or unreliably known, one may rely on a regression-type model relating the dependent to the explanatory variables in an empirical way. The simplest approach is to assume that the trend is a linear function of the (possibly transformed) explanatory variables, as in linear multiple regression:

$$m(\mathbf{s}, t) = \sum_{i=1}^p \beta_i \cdot f_i(\mathbf{s}, t), \quad (5.2)$$

where the β_i are regression coefficients, the f_i are explanatory variables that must be known exhaustively over the space–time domain and p is the number of explanatory variables.

Estimation of the regression coefficients can be done using common least squares algorithms that minimize the sum of squared differences between the observations and predictions (Montgomery et al. 2006), such as implemented in statistical software packages. Alternatively, maximum likelihood methods may be used (see Chapter 3, also).

After a trend has been specified and estimated, it may be subtracted from Z so that attention can be directed to the space–time stochastic residual ε . With this approach, uncertainties in the detrending procedure are not taken into account in the subsequent analysis. As a consequence, in kriging this causes the uncertainty in predictions to appear smaller than it is. This problem can be avoided by accounting for uncertainties in the trend coefficients, such as is done in universal kriging (Diggle and Ribeiro 2007); this method is also referred to as regression kriging (Hengl et al. 2004).

5.4.2 *Characterization of the Stochastic Residual*

Throughout this chapter, we assume that the zero-mean stochastic residual ε is multivariate normally distributed. Although a distributional assumption (i.e. normality) is not strictly needed for kriging, the assumption ascertains a completely-specified

statistical model and allows us to compute prediction intervals from the kriging results. Given this assumption, the only information lacking is the autocovariance function, C ,

$$C(\mathbf{s}_i, t_i, \mathbf{s}_j, t_j) = E[\varepsilon(\mathbf{s}_i, t_i) \cdot \varepsilon(\mathbf{s}_j, t_j)], \quad (5.3)$$

where E is the mathematical expectation. Alternatively, we may characterize the second-order properties of ε with the variogram, γ , as follows (see also Chapter 1):

$$\gamma(\mathbf{s}_i, t_i, \mathbf{s}_j, t_j) = \frac{1}{2}E[(\varepsilon(\mathbf{s}_i, t_i) - \varepsilon(\mathbf{s}_j, t_j))^2]. \quad (5.4)$$

In practice, to estimate γ from observations, some additional simplifying assumptions are necessary. A common approach is to introduce the assumption of second-order stationarity, which posits that the semivariance of $\varepsilon(\mathbf{s}, t)$ and $\varepsilon(\mathbf{s} + \mathbf{h}, t + u)$ depends only on the separating distance in space, \mathbf{h} , and that in time, u , between the points: $\gamma(\mathbf{s}, t, \mathbf{s} + \mathbf{h}, t + u) = \gamma(\mathbf{h}, u)$. Although \mathbf{h} is usually a vector in two or more dimensions, it can be replaced by Euclidean distance if isotropy is assumed in space, allowing both h and u to be treated as scalars. In practice a further simplification is needed to be able to estimate a space–time variogram from a set of observations. Here we impose a sum-metric variogram model on the space–time residual (Bilonick 1988; Dimitrakopoulos and Luo 1994; Snepvangers et al. 2003):

$$\gamma(h, u) = \gamma_S(h) + \gamma_T(u) + \gamma_{ST}(\sqrt{h^2 + (\alpha \cdot u)^2}). \quad (5.5)$$

The first two terms on the right-hand side of Eq. 5.5 allow for zonal space–time anisotropies (i.e. variogram sills that are not the same in all directions). Zonal anisotropy means that the amount of variation in time is smaller or larger than that in space, and or that in joint space–time. If variation in space dominates variation in time, then γ_S will have larger values than γ_T , which may be the case in the bulk density example given in Section 5.1. The opposite would hold for the rainfall example, also mentioned above. The third term on the right-hand side of Eq. 5.5 represents a joint space–time structure. It contains a geometric anisotropy ratio α to match distance in time with distance in space. Geometric anisotropy is needed here because a unit of distance in space is not the same as a unit of distance in time. For instance, if $\alpha = 20$ m per day, then two points that are separated by 100 m in space and zero days in time have the same correlation as two points that are 5 days apart in time and zero metres apart in space, or as two points that are separated by 60 m in space and 4 days in time (Fig. 5.5).

The sum-metric model simplifies the space–time variogram to a form such that its parameters can be estimated from a space–time dataset, but there are other approaches too. Among alternative models the product-sum model is often used (De Cesare et al. 2001). Research into which models are most appropriate for which situations is ongoing (see the discussion in Section 5.6).

Once the trend and variogram of the residuals have been specified, space–time prediction can be done in the usual way by ordinary or simple kriging (see Chapter 1). Kriging not only provides the best linear unbiased predictor of $Z(\mathbf{s}_0, t_0)$

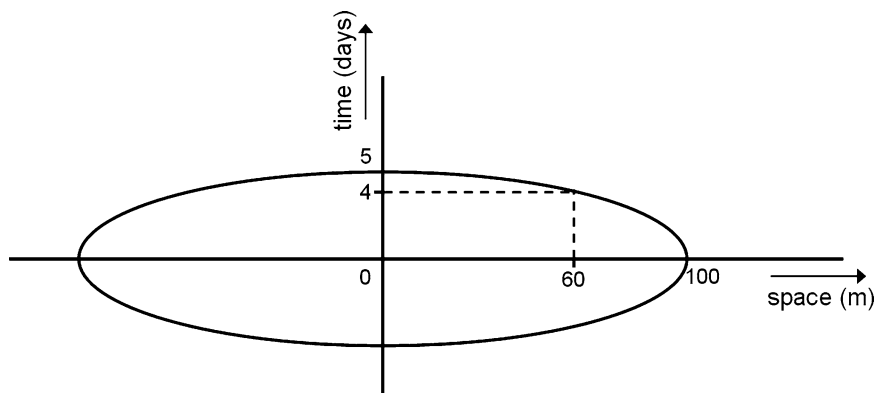


Fig. 5.5 Graphical illustration of geometric space–time anisotropy with an anisotropy ratio of 20 m per day. All points on the ellipse have the same semivariance

at any space–time point (s_0, t_0) , but it also quantifies the kriging prediction error with the kriging variance from which maps of the kriging standard deviation or lower and upper limits of prediction intervals can be derived. In space–time kriging, this gives results in three dimensions. These can be presented as a series of two-dimensional maps by taking time slices (which may be presented as a moving image when the distance between subsequent times is sufficiently small), or as time-series of predictions and prediction error standard deviations at different locations. The advantage of space–time kriging over spatial kriging for separate time points is that all observations are used rather than just the observations at the particular point in time, and that predictions in between measurement times can be made.

5.5 Application of Space–Time Geostatistics to the Lauwersmeer Farm Data

5.5.1 Characterization of the Trend

The absolute value of the correlation coefficients between soil properties and NDVI, and between elevation and NDVI are small (Table 5.1). Although these values were larger when the correlation was computed for separate days, they remained small (data not shown). However, the correlation between NDVI and crop type is greater (0.085) and there is also a clear relationship between NDVI and time (Fig. 5.3). We decided, therefore, to define the trend as a linear combination of crop type and a time effect:

$$m(\mathbf{s}, t) = \beta_S \cdot \delta(\text{crop type} = \text{Sofista}) + \beta_I \cdot \delta(\text{crop type} = \text{Innovator}) + f(\text{DOY}). \quad (5.6)$$

Here, the function δ is an indicator transform that is one if the condition is satisfied and zero otherwise; and f is a double logistic function that has been applied successfully to model temporal variation of NDVI (Fischer 1994):

$$f(DOY) = v_{\min} + v_{\max} \left(\frac{1}{1 + e^{m_1 - n_1 \cdot DOY}} - \frac{1}{1 + e^{m_2 - n_2 \cdot DOY}} \right). \quad (5.7)$$

The six parameters of the double logistic function were fitted on the NDVI data using a least squares minimization approach; the parameters are $v_{\min} = 0.091$, $v_{\max} = 0.689$, $m_1 = 33.7$, $n_1 = 0.2099$, $m_2 = 28.1$ and $n_2 = 0.1123$. The graph of the fitted temporal trend is shown in Fig. 5.3. Next, the remaining two regression coefficients were estimated on the residuals after removal of the temporal trend by unweighted least squares; the coefficients are $\beta_S = -0.0497$ and $\beta_I = 0.0335$. Estimation of the trend could be improved by the simultaneous fitting of all parameters or by using residual maximum likelihood (REML) to estimate the parameters (Chapter 3), however, this may be computationally difficult because Eq. 5.7 is nonlinear in its parameters. A more complex trend including interactions might also have been used to take into account that the effect of crop type varies with time. Absolute values of the correlation coefficients between the NDVI residuals and soil properties and elevation were all smaller than 0.10, indicating that there was no substantial predictive power left in the auxiliary information.

The histogram of the NDVI residuals is shown in Fig. 5.6. There are outliers to the left of the distribution in Fig. 5.6b. These correspond to locations at the boundary of the parcels where the crop did not develop fully (see Figs. 5.2 and 5.3, also). We decided to remove all values < -0.3 (214 observations, 0.6% of the total) because the outliers would impair the geostatistical analysis too much. The choice of the -0.3 threshold was subjective, however, if a value had been chosen closer to zero many more observations would have to be removed (Fig. 5.6b). As a consequence of discarding these observations, the resulting maps will be biased to larger

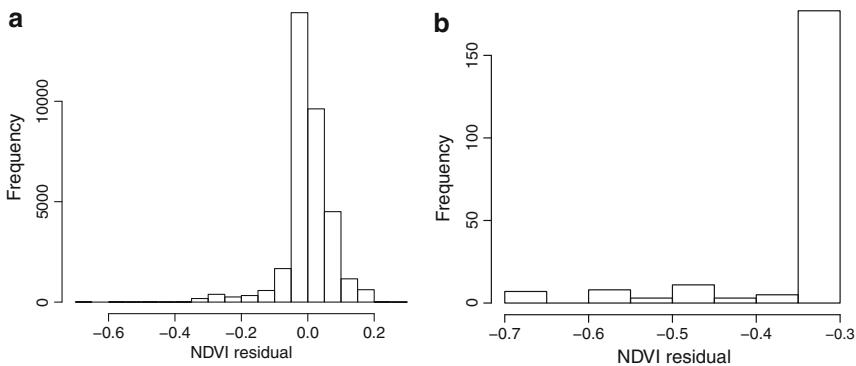


Fig. 5.6 (a) Histogram of NDVI residuals after removal of a double logistic temporal trend and crop type means and (b) histogram of a subset of the same data with only values smaller than -0.3

NDVI values at the boundary of the field. Alternatively, all the observations could be retained and transformed prior to variogram estimation and kriging, for instance using a disjunctive kriging approach with Hermite polynomial transformation as described in Chapter 1.

5.5.2 Characterization of the Stochastic Residual

The experimental variogram of the NDVI residuals was calculated for various temporal and spatial lags. The temporal lag size was chosen to be fairly large because there were only 16 measurement dates. Six time lags were used, each having a width of 10 days. The spatial lag size used was 15 m. Marginal spatial and temporal variograms were also calculated; these variograms are specific to only the space or time dimension. For instance, the marginal spatial variogram is computed on NDVI residuals for which the distance in time is zero (i.e. observations collected on the same day). The marginal and space–time experimental variograms are shown in Fig. 5.7.

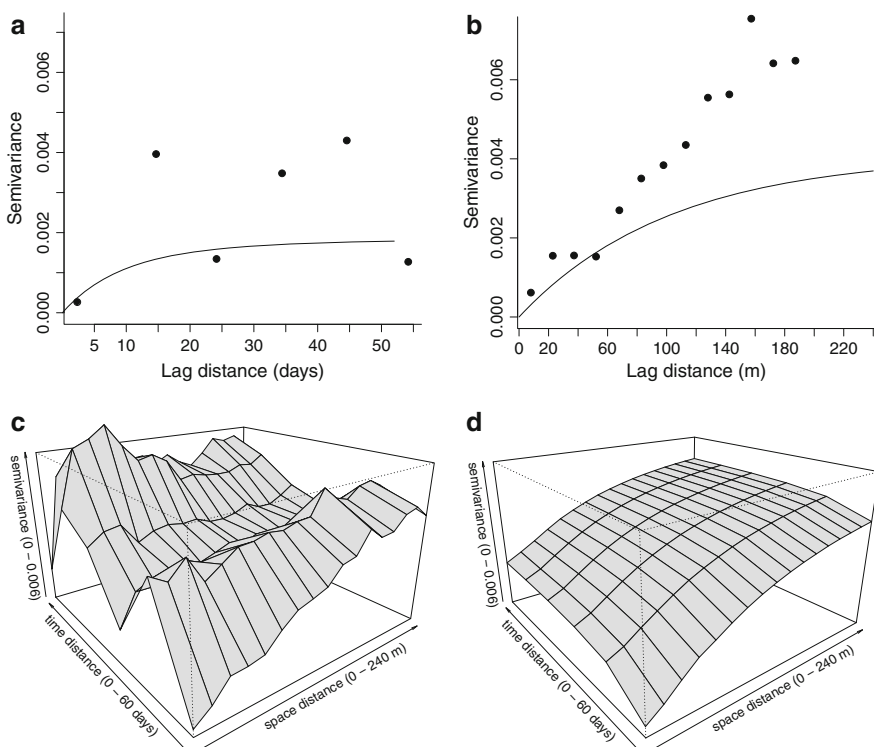


Fig. 5.7 Marginal experimental variogram (dots) and fitted model (solid lines): (a) in time direction and (b) space direction. Perspective plots of: (c) 3D experimental variogram and (d) fitted space–time variogram model

Table 5.2 Parameters of the fitted space–time variogram model

Variogram component	Nugget	Sill	Range	Anisotropy ratio
Spatial	0	0.00283	100 m	NA
Temporal	0	0.00114	8 days	NA
Spatio-temporal	0	0.00047	120 m	6 m/day

The experimental variogram in the time direction, in particular, is noisy. Nevertheless, the variograms show convincingly that the NDVI residuals are correlated in space and time. Near the origin, where distances in time and space are small, the semivariances are smaller than at larger distances. Spatial and temporal variation are of the same order of magnitude; there is somewhat more variation in space than in time. Thus, after removal of the trend, the NDVI residuals tend to vary more between locations in space than between instants in time.

The metric space–time variogram model was fitted using an exponential function for all three variogram components. Fitting was done using a quasi-Newton method with box constraints (Byrd et al. 1995). Quasi-Newton methods seek the minimum of a function by setting its gradient to zero and box constraints impose minima and maxima for the parameters in the search space (for instance, the nugget, sill and range parameter must all be equal to or greater than zero). The results are shown in Fig. 5.7b and also as the solid lines in Fig. 5.7a, b. The fitting was based on all space–time lags, which explains why the fitted curves do not reproduce the marginal variograms as well as might have been achieved if the only consideration had been to reproduce the marginal experimental variograms. The parameters of the variogram models are given in Table 5.2. The nugget variance is zero, which implies that the NDVI residual is perfectly correlated at very short distances in time and space. The spatial sill is the largest, which confirms that spatial variation in the NDVI residual dominates temporal variation. The space–time component has a moderate sill, although its contribution is not negligible.

5.5.3 Space–Time Kriging

Space–time kriging was done using the Gstat library (www.gstat.org) within the R statistical software (Bivand et al. 2008). A small subset only of the entire space–time cube of predictions and prediction error standard deviations are shown here. Maps of the predicted NDVI are shown for three time instants in Fig. 5.8. These show that NDVI is small on DOY 165, especially for Sofista, large on DOY 200 for both crops and small for Sofista and still large for Innovator on DOY 235. This agrees with the results presented in Fig. 5.3. The boundary between the two parcels is clearly visible in all the maps of predictions shown here, and so are the patches with small NDVI values near the parcel boundaries. A stripe effect is also apparent, particularly for DOY 165. This is probably an artefact caused by systematic measurement errors between the NDVI observations of neighbouring tramlines, although true systematic differences between neighbouring crop rows cannot be ruled out. One kriging

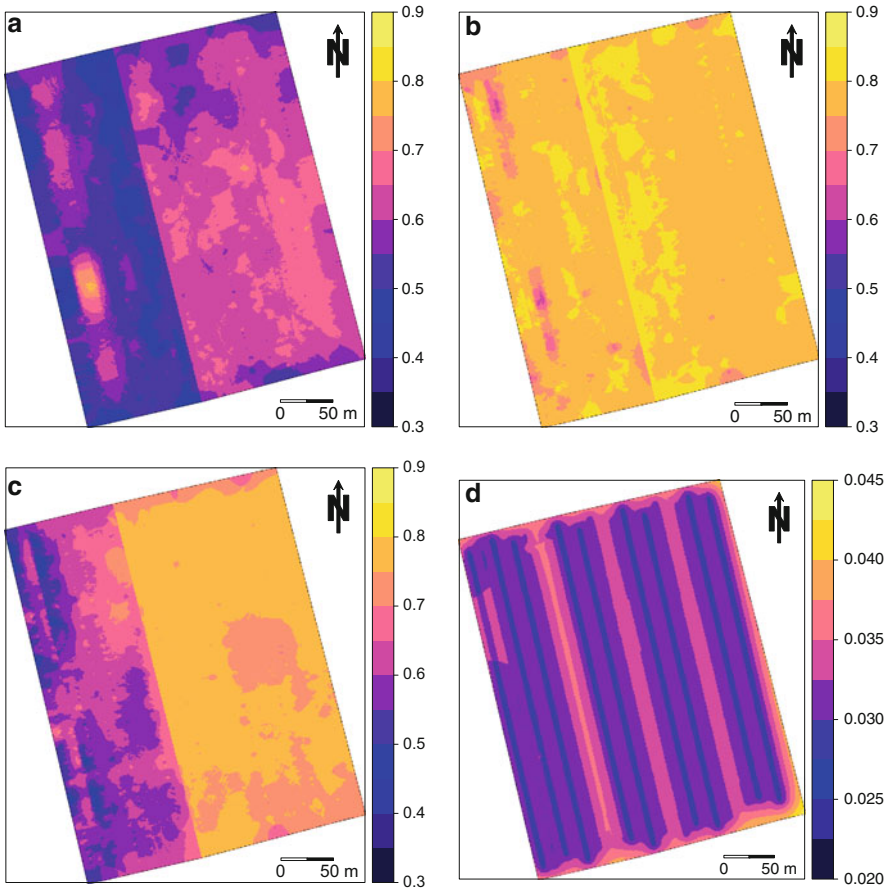


Fig. 5.8 Space–time kriged predictions for three arbitrary days: (a) DOY 165, (b) DOY 200 and (c) DOY 235; (d) kriging standard deviations for DOY 235

prediction error standard deviation map is shown for DOY 235 in Fig. 5.8d. The standard deviations are small compared to the predicted values (note the different legend entries) and are smaller near observations than further away from them. Note also that the standard deviation map indicates that only 9 of the 16 tramlines (see Fig. 5.8) were used near DOY 235 (see also Table 5.1). The standard errors are small because of the large number of observations and strong space–time correlations.

Space–time kriging results are also shown as time-series for three arbitrary locations in Fig. 5.9. Two of these are from the Sofista parcel where NDVI predictions become smaller towards the end of the observation period. Note also that the intervals between predictions are smaller near points in time where observations were made (i.e. one of the 16 days where NDVI was measured). For the corner location in the Sofista parcel, some observations fall outside the 95% prediction interval, and

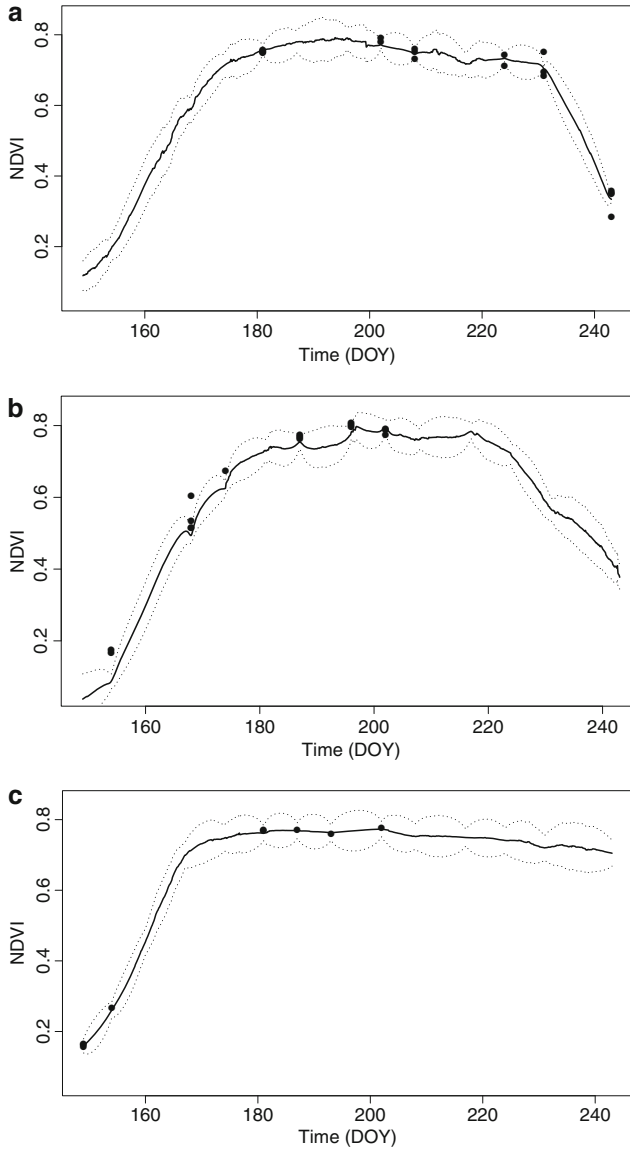


Fig. 5.9 Space–time kriging results for three arbitrary locations: (a) centre and (b) north-west corner of the Sofista parcel, and (c) centre of the Innovator parcel. The *solid line* is the regression kriging prediction, *dotted lines* represent the 95% prediction intervals derived from the kriging standard deviation. The dots are NDVI observations within a circular neighbourhood with radius 2 m

this happens more often than the expected 1 out of 20 cases. This might be caused by the fact that there may be a separation distance of up to 2 m between the observation and prediction location (a tolerance of 2 m was used to ensure that each graph

has observations), but may also be caused by the fact that outliers were removed before computing and modelling the space–time variogram (see Section 5.2), which led to a systematic underestimation of the variation.

5.6 Discussion and Conclusions

Space–time geostatistics is a useful extension to spatial geostatistics for precision agriculture because many of the variables addressed in precision agriculture vary in time as well as in space. Space–time kriging can provide predictions at a high space–time resolution and can be used, for example, to produce a time-series of spatial maps. When such maps are shown in animation mode these images may be useful to farmers to gain insight into crop growth. The data and results of the case study were only available after the growing season, but if available during the growing season the results can aid the farmer’s understanding of how crop variety and management practices affect crop growth. It will show where and when within a parcel anomalies caused by drought, wetness or diseases occur. For instance, this will enable the farmer to adjust the chemical spraying against the potato disease *Phytophthora* in space and time according to the actual occurrence of the disease. Compared to the current practice of weather dependent precautionary spraying, this can reduce substantially the amount of chemicals needed for optimal crop growth. The trend analysis, which is part of the development of a space–time model, might also provide important insight into which factors influence the variation in space and time in target variables. This can contribute to a better timing of the final fertilizer N application, or to determine the best DOY to remove the aboveground green biomass at harvest.

Space–time kriging is not intended to make forecasts in time. The primary purpose of kriging is interpolation; it is not intended for extrapolation, be that in space or time. Extrapolation would result in large uncertainties, as exemplified by the kriging variance. However, when implemented in real-time mode, space–time kriging can provide farmers with almost instant imagery of crop- and soil-related properties for the past and present. Process-based modelling, possibly augmented with data assimilation functionality (e.g. Hoeben and Troch 2000; Heuvelink et al. 2006), is needed to make reliable forecasts. Data assimilation techniques merge information from observations with information from models, taking the relative uncertainty associated with each of the sources of information into account.

The Lauwersmeer farm case study addressed in this chapter was limited to NDVI, but the methodology is generic and applies to many more properties that are relevant to precision agriculture, provided sufficient observations are available. In addition, the spatial extent and time period are not restricted. The case study addressed spatial variation within a 10 ha field during the growing season, but much larger space–time domains can be handled too. In the case study we assumed a trend that was a combination of a constant-in-time crop variety effect and a constant-in-space seasonal effect. A more elaborate model would let these effects vary in space and time, would

consider interactions between effects and would make a more careful assessment of the effects of other factors, such as soil type, previous crop growth variation and terrain form. Also, effects from parcel boundaries and anisotropy resulting from tillage, planting and fertilizer application may be included.

Practical application of space–time kriging is not as straightforward as it is for spatial kriging. This is partly because modelling variation in space–time is more difficult than modelling that in space, and because ‘off-the-shelf’ software is not yet available. The software Gstat was used in the case study, but it is not tailored to space–time geostatistics and cannot handle non-metric variograms or 3D space–time prediction. For instance, some additional programming was needed to compute the experimental variogram. More flexible alternatives are GSLIB (Deutsch and Journel 1998) extensions (De Cesare et al. 2002) and SEKS-GUI (Yu et al. 2007). The modelling of space–time variation also needs further development, both in terms of the choice of structure of the space–time variogram and in terms of the improved and user-friendly estimation of its parameters. Recently, many advanced methods have been published in the statistical literature to define classes of valid space–time covariance structures (e.g. Cressie and Huang 1999; Gneiting 2002; Huang et al. 2007; Fuentes et al. 2008; Ma 2008). A comprehensive review of these methods is beyond the scope of this chapter, but it is important to note that the variogram modelling approach used in this chapter is only one of many possibilities. One particular problem that confronts the modelling is that observations are generally spread unevenly over space and time. Data sets will be generally sparse in space and abundant in time (e.g. Snepvangers et al. 2003). In precision agriculture applications, however, data may be sparse in time and abundant in space: in the case study there were only 16 measurements in time and hundreds of spatial measurements for each instant in time. The uneven spread of observations over the space–time domain complicates the variogram modelling process and subsequent kriging. In the future the lack of temporal data could well change in agriculture and other applications. For instance, high resolution satellite or aerial imagery such as Ikonos and radar maps of rainfall may become readily available and instantly downloadable at affordable prices. Also the application and further development of ‘smart dust’ soil temperature and moisture sensors could provide a more even distribution of abundant data in space and time in precision agriculture. Space–time kriging may become a valuable tool in precision agriculture and other applications, combined with the current developing practice of using continuous soil moisture sensors and high resolution soil maps to fine-tune irrigation or the use of precise soil temperature data to determine sowing and planting dates.

In parallel with the theoretical advances, practical application of space–time geostatistics also needs further development. More real-world applications of space–time kriging will not only develop the maturity of this subject area and encourage the development of user-friendly software, but it might also reveal stable patterns in the type of model to be used for specific applications. If this were the case, then a model structure developed in one year for a given area, soil type or crop variety may be used in future years for similar areas, soil types or crop varieties, thus saving time on the cumbersome modelling stage. Thus, the farmer may use the model

developed in the case study for future use on his farm to create space–time surfaces of NDVI in near real-time. This will allow him to delineate those parts of the parcel where additional management is required (e.g. parts where the NDVI has not increased sufficiently during the past fortnight) and take action. The incentive for the further development of space–time techniques must come primarily from the field of applications. Data availability in precision agriculture is ever increasing and reliable methods are needed to distil useful information from the data and present the information in a tangible and efficient way so that farmers and their advisers can comprehend and use it. In this respect, precision agriculture may well be an important application field that can stimulate and steer the further development of space–time geostatistics.

Acknowledgements We thank Mr. Claassen for use of the Lauwersmeer farm data. We thank Allard de Wit (Alterra) for his suggestion to use the double logistic function and assistance with fitting it to the NDVI data.

References

- Actueel Hoogtebestand Nederland. (9 July 2009). <http://www.ahn.nl/index.php>.
- Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, 35, 161–173.
- Bilonick, R. A. (1988). Monthly hydrogen in deposition maps for the northeastern U.S. from July 1982 to September 1984. *Atmospheric Environment*, 22, 1909–1924.
- Bivand, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2008). *Applied spatial data analysis with R*. New York, USA: Springer.
- Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal of Scientific Computing*, 16, 1190–1208.
- Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, 62, 241–252.
- Cressie, N., & Huang, H. C. (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94, 1–53.
- De Cesare, L., Myers, D. E., & Posa, D. (2001). Product-sum covariance for space–time modeling: An environmental application. *Environmetrics*, 12, 11–23.
- De Cesare, L., Myers, D. E., & Posa, D. (2002). FORTRAN programs for space–time modeling. *Computers and Geosciences*, 28, 205–212.
- Deutsch, C. V., & Journel, A. G. (1998). *GSLIB: Geostatistical software library and user's guide* (2nd ed.). New York: Oxford University Press.
- Diggle, P. J., & Ribeiro Jr., P. J. (2007). *Model-based geostatistics*. New York: Springer.
- Dimitrakopoulos, R. & Luo, X. (1994) Spatiotemporal modeling: Covariances and ordinary kriging systems In R. Dimitrakopoulos (Ed.), *Geostatistics for the next century* (pp. 88–93). Dordrecht: Kluwer.
- Fischer, A. (1994). A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote Sensing of Environment*, 48, 220–230.
- Fuentes, M., Chen, L. & Davis, J. M. (2008). A class of nonseparable and nonstationary spatial temporal covariance functions. *Environmetrics*, 19, 487–507.
- Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97, 590–600.
- Hengl, T., Heuvelink, G. B. M., & Stein, A. (2004). A generic framework for spatial prediction of soil properties based on regression-kriging. *Geoderma*, 120, 75–93.

- Heuvelink, G. B. M., Schoorl, J. M., Veldkamp, A., & Pennock, D. J. (2006). Space–time Kalman filtering of soil redistribution. *Geoderma*, *133*, 124–137.
- Hoeben, R. & Troch, P. A. (2000). Assimilation of active microwave observation data for soil moisture profile estimation. *Water Resources Research*, *36*, 2805–2819.
- Holland Scientific (27 November 2008). www.hollandscientific.com.
- Huang, H.-C., Martinez, F., Mateu, J., & Nontes, F. (2007). Model comparison and selection for stationary space–time models. *Computational Statistics & Data Analysis*, *51*, 4577–4596.
- Ma, C. (2008). Recent developments on the construction of spatio-temporal covariance models. *Stochastic Environment Research and Risk Assessment*, *22*, S39–S47.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2006). *Introduction to linear regression analysis* (4th ed). New York: Wiley.
- Snepvangers, J. J. J. C., Heuvelink, G. B. M., & Huisman, J. A. (2003). Soil water content interpolation using spatio-temporal kriging with external drift. *Geoderma*, *112*, 253–271.
- Webster, R. (2000). Is soil variation random? *Geoderma*, *97*, 149–163.
- Wösten, J. H. M., Veerman, G. J., De Groot, W. J. M., & Stolte, J. (2001). *Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks*. Vernieuwde uitgave 2001. Wageningen, The Netherlands: Alterra, Report 153.
- Wu, J., Wang, D., & Bauer, M. E. (2007). Assessing broadband vegetation indices and QuickBird data in estimating leaf area index of corn and potato canopies. *Field Crops Research*, *102*, 33–42.
- Yu, H.-W., Kolovos, A., Christakos, G., Chen, J.-C., Warmerdam, S., & Dev, B. (2007). Interactive spatiotemporal modelling of health systems: the SEKS–GUI framework. *Stochastic Environment Research and Risk Assessment*, *21*, 555–572.