Sampling Design Optimization for Space–Time Kriging

ARTICLE · OCTOBER 2012
DOI: 10.1002/9781118441862.ch9

4 AUTHORS:

Gerard Heuvelink
Wageningen University
252 PUBLICATIONS 3,528 CITATIONS

Daniel A. Griffith
University of Texas at Dallas
207 PUBLICATIONS 2,950 CITATIONS

Tomislav Hengl
ISRIC - World Soil Information
99 PUBLICATIONS 1,671 CITATIONS

S.J. Melles
Ryerson University
25 PUBLICATIONS 296 CITATIONS

Available from: S.J. Melles
Retrieved on: 28 August 2015
1

Sampling design optimization for space-time kriging

G.B.M. Heuvelink, D.A. Griffith, T. Hengl and S.J. Melles

Wageningen University, The University of Texas at Dallas, ISRIC World Soil Information, and Trent University

1.1 Introduction

Space-time geostatistics has received increasing attention over the past few decades, both on the theoretical side with publications that extend the set of valid space-time covariance structures (e.g., Fonseca and Steel (2011); Fuentes et al. (2008); Ma (2003)), and on a practical side with publications of real-world applications (e.g., Gething et al. (2007); Heuvelink and Griffith (2010); see citations in Cressie and Wikle (2011)). Public domain software solutions to space-time kriging are becoming more user-friendly (e.g., Cameletti (2012); Pebesma (2012); Rue (2012); Spadavecchia (2008); Yu et al. (2007)), although not as much as is the case with spatial kriging. These developments, combined with almost all environmental variables varying in both space and time, indicate that the use of space-time kriging will likely increase dramatically in the near future.

One topic that to our knowledge has not yet been addressed in space-time geostatistics concerns the choice of the number and configuration of observation points in space and in time. Sampling design optimization in space has been well covered in the geostatistical literature (e.g., Angulo et al. (2005); Brus and Heuvelink (2007); Müller (2007); Müller and Stehlik (2010); Xia et al. (2006)). From a methodological point of view, no principal difficulty should prevent an extension of these approaches to the space-time domain, and existing techniques may be used. However, there are important differences, too. In space-time kriging, often strong anisotropies must be taken into account because variation in time can be quite different from variation in space. With space-time applications, one may choose from a much richer set of sampling designs (de Gruijter et al. (2006), Chapter 14). Data collection costs may be different in space-time applications because it may be relatively cheap to collect a time series of data at fixed spatial locations (e.g., using permanently
installed devices and data-loggers), or it may be relatively cheap to collect observations at multiple spatial locations at (nearly) the same time instant (e.g., during the same field campaign). Thus, there are interesting aspects to be explored when extending sampling design optimization from the spatial to the spatio-temporal domain.

Sampling design optimization requires that an optimality criterion be defined to evaluate and compare different designs. Common criteria are minimizing (1) the errors associated with estimation of the trend and covariance structure of the statistical model that is adopted, and (2) the average kriging prediction error variance. Of these two, minimization of the estimation error of the covariance structure is the most difficult to address. Solutions for spatial optimization have been derived (e.g., Diggle and Ribeiro (2007); Müller (2007); Zhu and Stein (2006)), but because of the (numerical) difficulties involved, this issue is not our first priority in an extension to the space-time case. In this chapter, we restrict ourselves to a criterion that simultaneously minimizes the variance of the estimation error of a linear trend and that of the interpolation error of the kriging residual. This criterion is essentially equivalent to minimizing the average universal kriging variance. See Brus and Heuvelink (2007) for a spatial analogue.

A solution algorithm is used to minimize the defined criterion and obtain an optimal sampling design. Both analytical and numerical techniques have been developed and used (e.g., Melles et al. (2011); Müller (2007); Pilz and Spock (2006)), although analytical methods typically involve simplifications to make a solution tractable. Here we work with one particular numerical optimization technique, known as spatial simulated annealing (e.g., Guedes et al. (2011); van Groenigen et al. (1999)), which is generic and easily implemented, although computationally demanding, and yields the global optimum only in theory. More attractive and efficient algorithms exist, such as genetic algorithms and greedy algorithms (e.g., Baume et al. (2011); ter Braak (2006); Vrugt et al. (2008)). However, because the emphasis of this chapter is on statistical aspects, we restrict ourselves to the one technique for which we have much experience.

Minimization of the space-time average universal kriging variance is achieved in this chapter for monthly temperatures of the Upper-Austria region (see Chapter 1). The geostatistical model that we use treats monthly temperature as the sum of a trend and a space-time correlated Gaussian stochastic residual. The trend is taken as a linear function of elevation and a remotely sensed dynamic estimate of temperature (MODIS data). The period that we consider spans three years (2001, 2002 and 2003). Time-critical sampling decisions impact the notion of optimality in terms of the number and configuration of observation points in space and time. Accordingly, we consider three optimization scenarios. The first considers optimization of a static design, in which ten locations are to be selected from 35 current stations such that increases in the average space-time universal kriging variance caused by the thinning are as small as possible. The second scenario also considers a static design, but this time the ten stations can be located anywhere within the Upper-Austria region. Establishing a fixed sampling network suffers from duplication of information over time, and information redundancy can be reduced by allowing stations to be relocated. The third optimization scenario considers a dynamic design, in which the ten station locations may be changed at the start of each year.

Before we present the case study, discuss results and draw conclusions, we briefly review space-time universal kriging and spatial simulated annealing.
1.2 Methodology

1.2.1 Space-time universal kriging

Consider a variable \( z = f(z(s, t)) \) that varies within a spatial domain \( S \) and a time interval \( T \). Let \( z \) be observed at \( n \) space-time points \((s_i, t_i), i = 1 \ldots n\). Although the number of observations \( n \) may be very large, a complete space-time coverage of \( z \) requires some form of interpolation. The objective thus is to obtain a prediction of \( z(s_0, t_0) \) at a point \((s_0, t_0)\) at which \( z \) was not observed, where \((s_0, t_0)\) is frequently associated with the nodes of a fine space-time grid. To do this, \( z \) is assumed to be a realization of a random function \( Z \) characterized by a full statistical model, including its space-time dependence structure. Next, \( Z(s_0, t_0) \) is predicted from the observations and possibly from auxiliary information.

The space-time variation of \( Z \) can be characterized by decomposing it into a deterministic trend \( m \) and a zero-mean stochastic residual \( V \) – with this decomposition being a subjective choice made by a modeler (Diggle and Ribeiro 2007) – as follows:

\[
Z(s, t) = m(s, t) + V(s, t). \tag{1.1}
\]

The trend \( m \) is a deterministic, structural component representing large-scale variation (e.g., that part of \( Z \) that can be explained physically or empirically, using auxiliary information). The residual is a stochastic component representing small-scale, 'noisy' variation (i.e., the leftover part of \( Z \), which still holds information when it is correlated in space and/or time). We assume that it is multivariate normal.

The trend component may be written as:

\[
m(s, t) = \sum_{i=0}^{p} \beta_i f_i(s, t), \tag{1.2}
\]

where the \( \beta_i \) are unknown regression coefficients, the \( f_i \) are covariates that must be exhaustively known over the space-time domain, and \( p \) is the number of covariates. Covariate \( f_0 \) is taken as unity, resulting in \( \beta_0 \) representing the intercept.

The residual \( V \) may be thought of as comprising three components: spatial, temporal, and space-time interaction. Assuming all three of these components stationary and mutually independent yields the ‘sum-metric’ space-time covariance structure (Snepvangers et al. 2003):

\[
C(h, u) = C_S(h) + C_T(u) + C_{ST}(\sqrt{h^2 + (\alpha \cdot u)^2}), \tag{1.3}
\]

where \( C(h, u) \) denotes the covariance function of \( V \) with \( h \) units of distance in space and \( u \) units of distance in time, \( C_S(h) \) denotes the purely spatial covariance function component, \( C_T(u) \) the purely temporal covariance function component, and \( C_{ST}(\sqrt{h^2 + (\alpha \cdot u)^2}) \) the space-time interaction covariance function component, with \( \alpha \) denoting a geometric anisotropy ratio that converts units of time distance, \( u \), to space distance, \( h \). It is needed because a unit of distance in space is not the same as a unit of distance in time. Note that we have taken \( h \) to be a scalar representing Euclidean distance. The first two terms on the right-hand side of Eq. 1.3 allow for the presence of zonal anisotropies (i.e., variogram sills that are not the same in all directions), which occurs when the amount of variation in time is smaller or greater than that in space, and/or that in joint space-time.
Once the trend and covariance function of the residual term have been specified, (space-time) interpolation can be calculated in the usual way. Universal kriging yields the best linear unbiased predictor (with minimum expected mean squared error among all possible predictors under multivariate normality) of $Z(s_0, t_0)$ as per ([Bivand et al. 2008], Section 8.5):

$$\hat{z}(s_0, t_0) = m_0^T \hat{\beta} + c_0^T C_n^{-1} (z - M \hat{\beta}),$$  \hspace{1cm} (1.4)

where $M$ is an $n \times p$ design matrix of predictor variables at the observation points, $m_0$ is a vector of predictors at the prediction point, $C_n$ is an $n \times n$ variance-covariance matrix for the $n$ residuals at the observation points, $c_0$ is a vector of covariances between the residuals at the observation and prediction points, $T$ denotes matrix transpose, and $z$ is a vector of observations $z(s_i, t_i)$. The regression coefficients are estimated in the usual way, using generalized least squares.

The variance of the universal kriging prediction error is given by:

$$\sigma^2(s_0, t_0) = \text{Var}(Z(s_0, t_0) - \hat{Z}(s_0, t_0)) =$$

$$= C(0, 0) - c_0^T C_n^{-1} c_0 +$$

$$+ (m_0 - M^T C_n^{-1} c_0)^T (M^T C_n^{-1} M)^{-1} (m_0 - M^T C_n^{-1} c_0)$$ \hspace{1cm} (1.5)

These are familiar equations from a spatial universal kriging perspective, showing that spatio-temporal kriging and spatial kriging do not differ fundamentally either mathematically or statistically.

1.2.2 Sampling design optimization with spatial simulated annealing

Spatial simulated annealing (SSA, van Groenigen et al. (1999)) is the spatial counterpart to simulated annealing ([Kirkpatrick et al. 1983]), and as such the algorithm starts from an initial random design, makes slight perturbations of previous designs to generate candidate new designs, which are either accepted or rejected depending on whether these improve or deteriorate a selected criterion. Candidate designs that improve the criterion are always accepted, while designs that deteriorate the criterion are accepted with some probability. The latter is done in order to escape from local optima. A ‘cooling’ schedule imposes a decreasing probability of accepting worsening designs as the number of iterations increases. The iterative procedure is stopped when a fixed number of iterations has been tried, or when no improvement occurs in the criterion over a set number of iterations.

Candidate designs can be derived from the current design in various ways. In this chapter, we randomly choose a station from the selected ten and swap it with one of the remaining unused 25 stations (scenario 1, optimal selection of 10 stations from the existing 35), or displace it geographically in a random direction (scenario 2, optimal allocation of 10 stations in the entire Upper-Austria region). The size of displacement is drawn from a uniform distribution between zero and a maximum displacement, while the maximum displacement is reduced as the number of iterations increases. Scenario 3 is a dynamic design that relocates the ten stations at the start of each year first by randomly selecting a year and station number, and then by moving it, as in scenario 2.
Each iteration step of SSA requires computing the space-time average universal kriging variance. This calculation was done by defining a space-time prediction grid with 1 km spatial resolution and a time resolution of one month. This scheme yields $11,979 \times 36 = 431,244$ prediction points. The universal kriging is based on $10 \times 36 = 360$ observations, which is sufficiently moderate to allow for global kriging (i.e., kriging is not restricted to observations in the local neighborhood of a prediction point). Computational load is substantial because SSA typically requires between 1,000 and 5,000 iterations, although this number is case dependent.

Recent examples of the application of SSA for optimization of spatial sampling designs include Brus and Heuvelink (2007) and Melles et al. (2011).

### 1.3 Upper-Austria case study

![Figure 1.1](image.png)

**Figure 1.1** Locations of 35 weather stations in the Upper-Austria region. Station ‘Aspach’ is circled.

#### 1.3.1 Descriptive statistics

Figure 1.1 portrays the locations of the 35 stations in the Upper-Austria region. The average monthly temperature at these stations for 2001, 2002 and 2003 yields 35 time series, with
missing values for some stations. Figure 1.2 depicts the temperature over time at the ‘Aspach’ station, which is located in the western part of Upper-Austria at an altitude of 436 m asl (i.e., the circled site in Figure 1.1). This graph shows clear seasonal behaviour, with the lowest temperature in December-February and the highest in June-August. Figure 1.2 also shows a time series of temperature for Aspach as derived from Moderate Resolution Imaging Spectroradiometer (MODIS) remote sensing imagery. MODIS land surface temperature (LST) images were obtained directly from NASA’s FTP data service\(^1\). Next these images were mosaicked together and resampled to the local projection system. The MODIS LST images are derived from MODIS thermal bands. Wan et al. (2004) show that the accuracy of MODIS LST images is about ± 1°C. However, parts of a MODIS image can be missing pixel values, which may be attributable to cloud cover and/or other unfavourable atmospheric conditions (Neteler 2010). For this case study, we imputed values for missing pixels by averaging the values of their neighboring dates.

Because MODIS data are exhaustively available for the region and time periods considered, and are correlated with observed monthly temperature, they furnish a sensible covariate in the geostatistical model specification to describe the trend. In addition, because temperature is known to be correlated with elevation, elevation also is included as a covariate, even though the MODIS data already contain part of the information furnished by elevation. Elevation was obtained from a 1 km resolution SRTM DEM of the area. All covariate layers

\(^1\)https://lpdaac.usgs.gov/get_data/data_pool

\[\text{Figure 1.2} \quad \text{Time series of monthly temperature data during 2001-2003 at station ‘Aspach’. Solid circles are observations; open circles are MODIS derived temperatures.}\]
Figure 1.3  Digital elevation model of the Upper-Austria region. Elevation in meters, grid cells are 1 by 1 km.

(MODIS LST images, SRTM DEM) were prepared in the MGI Austria GK Central projection system\(^2\). All data sets and processing steps used in this chapter are available for download, together with detailed annotations, via the public repository\(^3\). The Appendix provides R code used for the space-time variography, space-time kriging and optimal removal of locations.

Figure 1.3 portrays the digital elevation model of Upper-Austria. Figure 1.4 depicts a time series of MODIS maps for 2001. Figure 1.5 presents scatter plots of monthly temperature data versus elevation and MODIS data. Note that the covariation pattern of elevation and monthly temperature shows that two stations have high altitudes while the other 33 stations are in the range between 200 to 800 m asl. Figure 1.3 shows that MODIS data have a strong correlation with monthly temperature, but are more extreme. The relationship between monthly temperature and elevation is negligible, which is due to the stronger influence of seasonality. The relationship between elevation and the residual from a linear regression of temperature on MODIS data is stronger than that between monthly temperature and elevation, but still weak (data not shown).

A multiple linear regression (ignoring spatial and temporal dependencies\(^4\)) of monthly temperature on elevation and MODIS data reveals that the MODIS data covariate is highly

\(^2\)http://spatialreference.org/ref/epsg/31255/
\(^3\)http://code.google.com/p/uppera/
\(^4\)Accordingly, the significance tests are unreliable.
Figure 1.4  Time series of MODIS monthly temperature layers for year 2001. Vertical time bar indicates month of year (row-wise display with January in upper left and December in bottom right corner). Temperature in °C.

significant, whereas elevation is just significant (at the 5% level). Together these two covariates explain 92% of the variance in monthly temperature. The estimated regression coefficients are -0.38 °C/km for elevation, and 0.68 for MODIS. Figure 1.6 portrays the histogram of regression residuals. These residuals are fairly symmetrically distributed, with a few negative outliers and a moderate negative skewness (-0.64). Based on these results, both the MODIS data covariate and elevation were included in the trend component of the geostatistical model. No transformation of residuals was considered.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Nugget</th>
<th>Sill</th>
<th>Range parameter</th>
<th>Anisotropy ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>spatial</td>
<td>0.075</td>
<td>0.381</td>
<td>25 km</td>
<td></td>
</tr>
<tr>
<td>temporal</td>
<td>0</td>
<td>0.062</td>
<td>150 days</td>
<td></td>
</tr>
<tr>
<td>space-time</td>
<td>0.035</td>
<td>0.035</td>
<td>25 km</td>
<td>75 m/day</td>
</tr>
</tbody>
</table>
Figure 1.5 Scatter plot of monthly temperature against MODIS data (left) and elevation (right). Dashed line represents the 1:1 line. All observations from 35 stations during 2001-2003 were used ($n=1,055$).

Figure 1.6 Histogram of residuals of a multiple linear regression of monthly temperature on MODIS and elevation.
1.3.2 Estimation of the space-time model and universal kriging

The universal kriging model given in Eq. (1) requires calculation of a sample variogram using the residuals rather than the monthly temperature data themselves. This task was done using R gstat software (Bivand et al. 2008). The default approach in gstat is to assume independence when calculating the residuals, although a work-around exists that calculates the residuals for an imposed correlation structure (Bivand et al. 2008), section 8.4.6), allowing an iterative approach to be used to estimate the trend and residual variogram. We used the default approach and calculated the sample residual variogram for spatial lags of 8 km and temporal lags of 1 month. The right-hand panel of Figure 1.7 summarizes these results. Next, we fitted the sum-metric model with the optim function in R, using a quasi-Newton method that allows box constraints (Byrd et al. 1995). The left-hand panel of Figure 1.7 portrays this fitted model.

Table 1.1 summarizes the parameter estimates of this fitted model.

Space-time universal kriging of monthly temperature for the three-year period yields prediction and prediction error standard deviation maps for each of the 36 months. As an example, Figure 1.8 depicts the result for August 2001. The spatial pattern of the prediction map closely resembles that of the MODIS data (Figure 1.4), with low values in the southern part and higher values in the central and eastern parts of the region. The prediction error standard deviation map has the lowest values near station locations, and high values along the boundaries and in the southern part of the region, where the covariates are more extreme and cause larger errors in the trend estimates. Overall, the kriging standard deviation is small, indicating that prediction errors in many instances are smaller than 1°C and rarely greater
than 2°C.

1.3.3 Optimal design scenario 1

Scenario 1 selects ten stations from the existing 35 such that the space-time average universal kriging variance is minimized. Six SSA runs were made, each producing somewhat different results, but all obtaining a very similar mean universal kriging variance (MUKV). A maximum of 2,000 iterations was imposed, but the algorithm also was stopped if 200 successive iterations did not improve the design. Figure 1.9 shows the evolution of the SSA criterion for the run that achieved the smallest MUKV. As expected, initially worsening designs are accepted because the probability of accepting worsening designs is still relatively large. After 225 iterations, the criterion only improves and some station swaps still yield a noticeable decrease in the MUKV, although the MUKV reduction from the initial to the final design is less than 7%.

Figure 1.10 portrays the corresponding optimized design. The selected stations are fairly uniformly spread over the region, with a tendency to over-represent stations near the border of the study area. These results are as expected. All six runs selected the most southern station, which can be explained by this station being at a high altitude, and hence having extreme values for elevation as well as MODIS temperature. Measurements from these stations substantially reduce the trend parameter estimation error that is part of the MUKV.
1.3.4 Optimal design scenario 2

Scenario 2 places ten stations optimally within the region without restriction. This scenario also was executed six times. The design of the run with the smallest MUKV is shown in Figure 1.11. The MUKV value is 0.32134, which is smaller than that obtained under scenario 1. The reduction is achieved by an even more uniform distribution of stations across the region, which is possible because station locations were not confined to those of the existing 35 stations. Again, a station is located in the southern ‘appendage’ where elevation is large and MODIS temperature small. This occurs in only two of the six runs, which were indeed the two runs with the smallest MUKV.

1.3.5 Optimal design scenario 3

Scenario 3 allows the repositioning of stations at the start of each calendar year. Results show that the added flexibility of this dynamic design pays off because the MUKV can be dramatically reduced to a value of 0.25301. Figure 1.12 shows the optimal design. Again, the locations are uniformly spread over the region, but this time the density can be increased because there are effectively 30 locations. However, note that for each single year the geographic spread is far from uniform. Apparently this outcome does not lead to a substantial
increase of the MUKV, which can be explained by the relatively small temporal and spatio-
temporal variation of the stochastic residual of the random process (see Table 1.1). This outcome also explains why restricting measurements to only one year does not impair the ‘information content’ of each single station. Results could have been quite different if no trend was assumed and ordinary kriging had been used, because in that case, temporal variability would likely not have been small compared to spatial variability.

Comparison of the MUKV of scenario 3 with that obtained with the original data set (35 stations ‘continuously’ measured, MUKV = 0.25548) shows that the sampling design of scenario 3 is (slightly) more accurate, while it uses far fewer observations. Even though several stations have missing values in the original dataset, which reduces the total number of observations from a potential of $3 \times 12 \times 35 = 1,266$ to $1,055$, scenario 3 has only 360 observations, which is a reduction of one-third.

**Figure 1.10** An optimal static design for selecting 10 from 35 stations within the Upper-Austria region. Filled circles represent selected stations, crosses those that were not selected.
1.4 Discussion and Conclusions

In this chapter we extend the use of spatial simulated annealing for optimization of sampling designs in space to optimization of sampling designs in space-time. This requires use of a space-time geostatistical model for which we use a common model that treats the space-time variable of interest as the sum of a trend, which is linear in known covariates, and a stationary Gaussian stochastic residual. Recent advances in space-time statistical modeling enhance flexibility and allow comparisons of optimal designs for different models. One interesting exercise would be to analyze the sensitivity of optimal designs to the structure and parameters of the space-time geostatistical model. Another interesting exercise would be to analyze how the optimal design is influenced by including parameter estimation errors in the space-time covariance structure of the stochastic residual. From a numerical perspective, evaluation of alternative optimization techniques also would be worthwhile, particularly because SSA is known to be computer-intensive, and because space-time kriging typically deals with much larger datasets than spatial kriging.

The extension to space-time also allows detailed analysis of the efficiency of various types of designs in space and time. Here we focus on static designs and consider only one basic
dynamic design. However, many more design options exist, such as synchronous, static-synchronous, and rotational (de Gruijter et al. 2006). The efficiency of these designs can be analyzed from a purely model-based perspective, which adds to existing studies that take a design-based or hybrid approach (e.g., Brus and de Gruijter (2011); Brus and De Gruijter (2012).

For the case study considered in this chapter, the dynamic design turns out to be much more attractive than the static designs in terms of the total number of observations that are required to achieve a given level of accuracy. However, the cost of a design is not necessarily proportional to the total number of observations. Relocating stations may be very costly, whereas maintaining a station in operation mode once it is installed may be relatively cheap. This situation may well be the case for the temperature monitoring considered here, but cost considerations may turn out differently when measurement instruments are expensive and relocating stations is cheap. Examples are the use of mobile stations that measure greenhouse gas emission, air pollution or nuclear radiation (e.g., Angelbratt et al. (2011); Heuvelink et al. (2010); Kukkonena et al. (2005)). Accessibility of locations also becomes an issue, such as when mobile stations are mounted on cars or when the study area of interest is in remote
Sampling design optimization for space-time kriging

or hostile terrain. These factors can be taken into account and the application of sampling design algorithms, such as those used here, may yield insightful solutions.

1.5 Appendix: R Code

```r
# Packages
library(rgdal)
library(gstat)
library(colorspace)
library(spacetime)
library(lattice)
library(raster)
library(sp)

# Space-time variography and kriging
# Convert space-time data frame to an STFDF
locs <- uppera.xyt.s[,c("ST_ID","X","Y")]
coordinates(locs) <- ~X+Y
proj4string(locs) <- uppera@proj4string
uppera.st <- STIDF(locs,time=uppera.xyt.s$Date,
data=uppera.xyt.s[,c("TEMP","MODIS.LST","dem","res")])
uppera.std <- as(uppera.st,"STFDF")

# Compute and plot sample variogram
vv <- variogram(TEMP~MODIS.LST+dem,data=uppera.std,width=8000,
cutoff=75000,tlags=0:20)
plot(vv, ylab="Time lag (days)"

# Initialize parameters of sum-metric variogram model
pars.init <- c(nugget.s=0.05,sill.s=0.2,range.s=25000,nugget.t=0.01,
sill.t=0.01,range.t=150,nugget.st=0.01,sill.st=0.1,range.st=25000,anis=75)
pars.l <- c(nugget.s=0,sill.s=1E-2,range.s=1000,nugget.t=0,sill.t=1E-4,
range.t=1000,anis=1)
pars.u <- c(nugget.s=0.5,sill.s=1E6,range.s=1E6,nugget.t=0.5,sill.t=1,
range.t=10000,anugget.st=0.5,snill.st=1,range.st=1E6,anis=1000)

# Function to compute variogram model
ExpVgmSumMetric=function(pars,s,t) {
  vs=ifelse(s==0,0,pars[1]+pars[2]*(1-exp(-s/pars[3])))
  vt=ifelse(t==0,y0,pars[4]+pars[5]*(1-exp(-t/pars[6])))
  h=sqrt(sˆ2+(pars[10]*as.numeric(t))ˆ2)
  vst=ifelse(h==0,0,pars[7]+pars[8]*(1-exp(-h/pars[9])))
  vs+vt+vst }

# Computes mean squared difference between sample variogram and variogram model
ExpFitFn=function(pars,gfn,v,trace=FALSE) {
  mod=gfn(pars,v$spacelag,v$timelag)
  resid=v$gamma-mod
  if (trace)
    resid$v$gamma-mod
  if (trace)
    print(c(pars,ME=mean(resid^2)))
  mean(resid^2) }

# Variogram fitting
pars.fit=optim(pars.init,ExpFitFn,gfn=ExpVgmSumMetric,v=vv,gr=NULL,
  method="L-BFGS-B",lower=pars.l,upper=pars.u,control=list(maxit=1000))

# Define space-time variogram of sum-metric model
vgm.st=vgm(pars.fit$par[1],"Exp",1E12,anis=c(0,90,0,1E-6,1E-6),add.to=vgm.st) # spatial nugget
vgm.st=vgm(pars.fit$par[2],"Exp",pars.fit$par[3]*1E6,
anis=c(0,90,0,1E-6,1E-6),add.to=vgm.st) # spatial partial sill
vgm.st=vgm(pars.fit$par[4],"Exp",1E12,anis=c(0,0,0,1,1E-5),add.to=vgm.st) # temporal nugget
vgm.st=vgm(pars.fit$par[5],"Exp",y pars.fit$par[6]*1E6,
anis=c(0,0,0,1,1E-5),add.to=vgm.st) # temporal partial sill
```
vgm.st=vgm(pars.fit$par[7],"Exp",1E-12,add.to=vgm.st) # space-time nugget
vgm.st=vgm(pars.fit$par[8],"Exp",pars.fit$par[9],
anis=c(0,0,0,1)/pars.fit$par[10],add.to=vgm.st) # space-time partial sill

# Define space-time prediction grid (grids1km object in workspace contains MODIS data)
n_begin <- 11 # January 2001
n_end <- 46 # December 2003
srtmdem$mask <- readGDAL("mask.sdat")$band1
range.X <- c(grids1km@coords[1]+grids1km@grid@cellsize[1]*0:grids1km@grid@cells.dim[1]-1)
range.Y <- rev(grids1km@coords[3]+grids1km@grid@cellsize[2]*0:grids1km@grid@cells.dim[2]-1))
range.timedata <- n_begin:n_end
st.grid <- expand.grid(X=range.X,Y=range.Y,
timedata=range.timedata,KEEP.OUT.ATTRS=FALSE)
st.grid[,"MODIS.LST"] <- unlist(unclass(grids1kmf@data[,n_begin:n_end]))
st.grid[,"dem"] <- as.data.frame(srtmdem)@data[,1]
st.grid[,"mask"] <- as.data.frame(srtmdem)@data[,1]
gridded(st.grid)="X"+"Y"+"timedata"

# Define set of observations
uppera.obs <- data.frame(X=uppera.std@sp@coords[,1],Y=uppera.std@sp@coords[,2],
timedata=uppera.std@time,TEMP=uppera.std$TEMP,MODIS.LST=uppera.std$MODIS.LST,
dem=uppera.std$dem,row.names=NULL)
coordinates(uppera.obs)=˜X+Y+"timedata"

# Kriging
uppera.uk <- krige(formula=TEMP˜MODIS.LST+dem,uppera.obs,st.grid,s,
model=vgm.st,debug.level=-1)

# Optimal removal of points by spatial simulated annealing

candidates # a SpatialPolygonsDataFrame - study area outline map
uppera.obs # a dataframe with columns, X, Y, timedata, TEMP, MODIS.LST and dem
grids1kmf # a SpatialGridDataFrame with monthly LST MODIS data in columns
st.grid.s # a SpatialPixelsDataFrame. This is the prediction grid in space-time
projxy <- "+proj=tmerc +lat_0=0 +lon_0=13.33333333333333 +k=1 +x_0=0
+y_0=-5000000 +ellps=bessel +units=m +no_defs"
proj4string(borders) <- projxy
nspaceDiff <- 25 # number of points to delete
ntimeSteps <- 12
range.timedata <- n_begin:n_end
num_periods <- length(range.timedata)/ntimeSteps
xy <- c("X","Y")
nn_space <- unique(uppera.obs@xy)
nn_time <- unique(uppera.obs@"timedata")
nn <- dim(nn_space)[1]-nspaceDiff # number of points to sample in space
delPts_vec <- as.vector(as.matrix(nn_space[sample(nrow(nn_space),
size=nn.spaceDiff,]),[1]))
uppera.obsNet <- subset(uppera.obs,!(uppera.obs@xy %in% delPts_vec))
coordinates(uppera.obsNet)=˜X"+"Y"+"timedata"
uppera.uk.init <- krige(formula=TEMP˜MODIS.LST+dem,uppera.obsNet,
model=vgm.st,debug.level=-1)
avkriqvar.init <- mean(uppera.uk.init$var1.var,na.rm=TRUE)
oldpoints <- unique(as.data.frame(uppera.obsNet)@xy) # save initial design

# Optimal removal of points by spatial simulated annealing
options()
nr_iterations <- 1000 # number of ssa iterations to run
max_points_shift <- 1 # number of points to move at any given time
start_p <- 0.2 # initial probability of accepting a worsening design
maxShiftFactorX <- 0.2 # multiplier of max possible shift for point in X range
minShiftFactorX <- 0 # multiplier of min possible shift for point in X range
maxShiftFactorY <- 0.2 # multiplier of max possible shift for point in Y range
minShiftFactorY <- 0 # multiplier of min possible shift for point in Y range
coolingFactor <- nr_iterations/10 # helps controls rate of cooling
Sampling design optimization for space-time kriging

*countMax <- 200*  # max nr of ssa iterations to process without change in criterion
*plotOptim <- TRUE*  # logical plot or not

*settings for simulated annealing*

*nr_designs <- 1*  # counter for number of accepted designs
*old.uppera.obsNet <- uppera.obsNet*  # save initial design
*criterionInitial <- avkri.var.init*  # initial criterion
*oldcriterion <- criterionInitial*  # also save current criterion
*oldDelPoints <- unique(uppera.obsDel[xy])*  # delPts in space
*old.uppera.obsDel <- uppera.obsDel*  # delPts with all times

criterionDef <- function() {  # keep track of improvement in objective function
*bestCriterion <- Inf*  # store best crit in case final design is not best design
*count <- 0*

*cnames <- dimnames(coordinates(uppera.obsNet))[[2]]*  # coordinate & time of data.frame
*vnames <- dimnames(uppera.obsNet@data)[[2]]*  # predictor variables of data.frame
*x_bounds <- bbox(candidates)[1,]*
*y_bounds <- bbox(candidates)[2,]
*x_extent <- x_bounds[2]-x_bounds[1]*
*y_extent <- y_bounds[2]-y_bounds[1]*

*max_shift_x <- maxShiftFactorX*x_extent*  # maximum shift in x-direction
*min_shift_x <- minShiftFactorX*x_extent*  # minimum shift in x-direction
*max_shift_y <- maxShiftFactorY*y_extent*  # maximum shift in y-direction
*min_shift_y <- minShiftFactorY*y_extent*  # minimum shift in y-direction

*for (k in 1:nr_iterations) {
  selected_shifts_del <- sample(nspaceDiff)*
  selected_shifts_net <- sample(nn)*
  oldDelPt <- oldDelPoints[which(selected_shifts_del<=max_points_shift),]*
  newDelPt <- oldpoints[which(selected_shifts_net<=max_points_shift),]*
  netPts <- rbind(oldDelPoints[which(selected_shifts_net>max_points_shift),], newDelPt)*
  overdelPts <- over(SpatialPoints(delPts,proj4string=CRS(as.character(projxy))),
                   ygrids1kmf)
  overdelPts.t <- subset(overdelPts, select=range.timedata)
  overDelPt <- over(SpatialPoints(delPts),st.grid.s)
  uppera.obsDel <- cbind(X=rep(delPts[“X”][1:nspaceDiff,],each=ntimeSteps),
                         Y=rep(delPts[“Y”][1:nspaceDiff,],each=ntimeSteps),
                         timedata=rep(range.timedata,nspaceDiff),
                         TEMP=rep(1,ntimeSteps*nspaceDiff),
                         MODIS.LST=overdelPts.t$MODIS.LST)
  uppera.obsDel <- as.data.frame(uppera.obsDel)
  coordinates(uppera.obsNet) <- “X”*“Y”*“timedata”
  uppera.uk <- krig(formula=TEMP~MODIS.LST+dem, uppera.obsNet, st.grid.s,
                           model=vgm.st, debug.level=-1)
  criterion <- mean(uppera.uk$var1.var, na.rm=TRUE)
  p <- runif(1)  # to allow accepting an inferior design
  if (criterion < oldcriterion) {
    oldPts <- netPts; old.uppera.obsNet <- uppera.obsNet; oldDelPoints <- delPts
    old.uppera.obsDel <- uppera.obsDel
    oldcriterion <- criterion
    nr_designs <- nr_designs+1
    count <- 0
  } else {
    if (criterion > oldcriterion & p <= (start_p*exp(-k/coolingFactor))) {
      oldPts <- netPts; old.uppera.obsNet <- uppera.obsNet
      oldDelPoints <- delPts
      old.uppera.obsDel <- uppera.obsDel
    }
    else cat(“No improvement for”,count,”iterations “)
    p <- runif(1)  # to allow accepting an inferior design
  }
  uppera.obsNet <- as.data.frame(uppera.obsNet)
  coordinates(uppera.obsNet) <- “X”*“Y”*“timedata”
}

1.6 Acknowledgements

We would like to acknowledge support provided by the Aquatic Research and Development Section of the Ontario Ministry of Natural Resources, Peterborough, Ontario, Canada.
References


Bivand RS, Pebesma EJ and Gómez-Rubio V 2008 Applied Spatial Data Analysis with R. Springer.


Cameletti M 2012 Spatio-temporal models in R Package vignettes. CRAN, Università di Bergamo.


Diggle P and Ribeiro PJ 2007 Model-Based Geostatistics. Springer.


Müller W 2007 Collecting Spatial Data 3rd edn. Springer.


Pebesma E 2012 Spatio-temporal geostatistics using gstat Package vignettes. CRAN, University of Münster.


