Royal Netherlands
Meteorological Institute
Ministry of Infrastructure and the Environment

Interpolation of Makkink evaporation in the Netherlands

Paul Hiemstra and Raymond Sluiter

De Bilt, 2011 | Technical report; TR-327

Interpolation of Makkink evaporation in the Netherlands

Versie 1.0

Datum 25 oktober 2011
Status Definitief

Interpolation of Makkink evaporation in the Netherlands

Dr. Paul Hiemstra
Dr. Raymond Sluiter

December 15, 2011

Contents

Preface

The aim of this study was to find an optimal interpolation method for Makkink reference evaporation. This study is part of a larger project within the KNMI to improve the way we interpolate meteorological data. Recording which interpolation method we choose to interpolate Makkink evaporation and providing the scientific foundation for this choice is its main goal.

Our colleagues at KS-KA provided feedback on our interpolated maps and provided essential climatological background information. In particular we thank Gerard van der Schrier, Rudmer Jilderda, Rob Sluijter and Jules Beersma.

All interpolations, data processing and visualization where done using the statistical computing environment R (R Development Core Team, 2010). Without this open source project this study could not have been performed.

De Bilt, February 2011

Paul Hiemstra and Raymond Sluiter

Summary

Computer models, e.g. hydrological models, are one of the key components in determining effect policy and estimating impact of water management. Many of these models require gridded information on meteorological variables such as evaporation and precipitation. Creating such gridded maps from observations is done using interpolation methods such as splines or linear regression. The aim of this study was to find the optimal interpolation method for Makkink reference evaporation.

The evaporation dataset consisted of three temporal support levels: yearly normalen, monthly normalen and daily observations. For each level we interpolated using a number of interpolation methods: inverse distance weighted interpolation, kriging, linear regression, interpolation with the mean, nearest neighbor and thin plate splines. To choose which method was optimal we looked at two aspects: summary statistics of cross-validation and the visual appearance of the map.

From our results we conclude that for this dataset thin plate splines is the optimal interpolation method.

Chapter 1

Introduction

In an ever changing environment, effective policy development to counter any adverse changes is of great importance. For example, should the height of the North Sea dikes be increased in response to a rise in sea level? Or how does climate change influence the availability of water for drinking and industrial applications? Computer models, e.g. hydrological models, are an important tools in assessing the impact of changes. Forecasts of meteorological parameters, e.g. rainfall or evaporation, are fed into the model. The behavior of the hydrological model in response to these changes in driving forces provides insight in the impact of these changes.

One of the driving forces for hydrological models is the reference evaporation. In the Netherlands, roughly 70% of the precipitation that falls evaporates. This fact makes evaporation an important driving force for any hydrological model. There are a number of methods to calculate the reference evaporation (Winter et al., 1995) from a number of meteorological parameters. At the KNMI, Makkink is used to calculate the reference evaporation. Makkink requires the incoming shortwave radiation and the mean daily temperature to calculate the reference evaporation for a particular day. The incoming radiation and temperature are measured at monitoring stations. However, hydrological models require gridded maps which include estimates of reference evaporation at unmeasured locations. Generating gridded maps from point data, i.e. the monitoring stations, is commonly referred to as interpolation. Many interpolation methods are available and where previously used for meteorological data, e.g. multiple linear regression (Gurtz et al., 1999), inverse distance weighted interpolation (Ni et al., 2006; Menzel, 1999), splines (McVicar et al., 2007; Jeffrey et al., 2001) or kriging (Jeffrey et al., 2001). Each of these interpolation methods has strong and weak points.

The purpose of this study was to find the optimal method for interpolating Makkink reference evaporation. Based on our own experience and previous research at the KNMI (Sluiter, 2009; Buishand et al., 2008) we defined a number of research questions. First, the number of stations for which Makkink can be calculated varies between 5 and 36 stations. Especially when the number of points is low, interpolation of these observations presents challenges. How many stations do we need to get an acceptable interpolation? Second, the evaporation shows a clear *temporal* trend throughout the year. Evaporation is much higher in summer than in winter, which is mainly caused by the small amount of incoming shortwave radiation. What is the influence of this temporal trend on the interpolation? Does the performance of the different interpolation methods change? Third, the data is available on three temporal supports: daily evaporation, 30 year yearly averages, or yearly *normalen*, and 30 year monthly averages, or monthly *normalen*. How does the performance of the interpolation methods change between these temporal supports. Fourth, literature suggest a number of *spatial* trends that might influence

evaporation (Schuurmans, 2009; McVicar et al., 2007; Jiang et al., 2004; Menzel, 1999; Gutman and Rukhovetz, 1996). Which of these trends improve the interpolation significantly? Does the improvement depend on the temporal support of the data, time of the year or the number of available observations?

To answer these questions we interpolated all observations for the different temporal supports and judged the quality of the resulting map. We determined the quality of a map by looking at two aspects. First, we looked at cross-validation statistics to quantify the performance of each interpolation method. Second, we used expert judgment on how the patterns in the interpolated map fitted expected patterns as a qualitative performance measure. Such an expert judgment is to some degree subjective. However, by defining the patterns we expect to find in an interpolated map we tried to make the expert judgment as objective as possible.

This report is divided into a number of chapters. Chapter 2 discusses the individual interpolation methods assessed in this study. In addition, we describe how we judge the quality of a particular map, which is paramount for accurately choosing the optimal interpolation method. The interpolation results are shown in chapter 3. This chapter is subdivided into the three temporal supports of the observations. Some of the results are shown inside this chapter, but the majority is shown in a number of appendixes at the end of this report (appendixes A to B). The final chapter will provide a discussion of the research questions and place them in the context of previous research.

Chapter 2

Methods

2.1 Makkink reference evaporation

The description of Makkink evaporation in this section is mainly based on Schuurmans (2009) and Buishand and Velds (1980). For additional detail, the reader is referred to these publications.

The primary driving force in case of evaporation is the incoming solar radiation. The following energy balance can be constructed at the surface:

$$Q_n = S^{\downarrow} - S^{\uparrow} + L^{\downarrow} - L^{\uparrow} \tag{2.1}$$

where Q_n is the net radiation flux, $S^{\downarrow}$ the incoming shortwave radiation (0.1 - 3 μm), $S^{\uparrow}$ the outgoing shortwave radiation, $L^{\downarrow}$ the incoming longwave radiation (3-50 μm) and, $L^{\uparrow}$ the outgoing longwave radiation. Part of the incoming shortwave radiation is reflected by the Earth's surface. The relation between $S^{\uparrow}$ and $S^{\downarrow}$ is given by:

$$S^{\uparrow} = \alpha \cdot S^{\downarrow} \tag{2.2}$$

where α is the fraction of radiation that is reflected, also known as the albedo. The albedo depends on the type of surface. Snow has a very high albedo (0.95), whereas grass has an albedo of about 0.25. The longwave radiation is mainly emitted by the earth's surface, but clouds can reflect the longwave radiation back to the surface. An alternative energy balance for Q_n subdivides the net radiation into three processes in which it is used:

$$Q_n = \lambda E + H + G \tag{2.3}$$

where λE is the latent heat flux (Wm^{-2}, energy used for evaporation of water), λ the heat of vaporization (Jkg^{-1}), E the evaporation flux ($\mathrm{kgm}^{-2}\mathrm{s}^{-1}$), H the sensible heat flux (Wm^{-2}, energy used to heat up the air), and G the soil heat flux (Wm^{-2}, energy used to heat up the soil).

Makkink relates daily reference evaporation to the incoming shortwave radiation (de Bruin, 1987):

$$ET_{ref} = C \cdot \frac{s}{s+\gamma} \cdot \frac{S^{\downarrow}_{day}}{\lambda * \rho} \tag{2.4}$$

where ET_{ref} is the Makkink reference evaporation (md^{-1}), C a constant equal to 0.65 (de Bruin, 1987), s the slope of the curve of saturation water vapor pressure ($\mathrm{kPa}^{\circ}\mathrm{C}^{-1}$), γ the psychometric

constant (kPa°C^{-1}), $S^{\downarrow}_{day}$ the daily incoming shortwave radiation (Jm^{-2}d^{-1}), and ρ the bulk density of water, i.e. 1000 kgm^{-3}.

The slope of the curve of saturation water vapor pressure is related to mean daily temperature (T_{day}, °C):

$$s = \frac{7.5 \cdot 237.3}{(237.3 + T_{day})^2} \cdot \ln 10 \cdot e_s \tag{2.5}$$

where e_s equals:

$$e_s = 0.6107 \cdot 10^{\frac{7.5 \cdot T_{day}}{237.3 + T_{day}}} \tag{2.6}$$

the psychometric constant (γ) is also related to T_{day}:

$$\gamma = 0.0646 + 0.00006 \cdot T_{day} \tag{2.7}$$

as is the heat of vaporization (λ):

$$\lambda = (2501 - 2.375 \cdot T_{day}) \cdot 1000 \tag{2.8}$$

This parametrization of Makkink only requires $S^{\downarrow}_{day}$ and T_{day}, which are both measured at monitoring stations throughout the country.

2.2 Interpolation methods

In a spatial context, interpolation is the process of estimating a variable at unmeasured locations using surrounding locations that do have a measurement of that variable. To obtain a gridded map, an interpolation estimate is made at each grid node until the map filled. A very basic interpolation method is to take the mean of all observations and assign it to all unmeasured grid nodes. This interpolation method is called global mean interpolation. An improvement on global mean interpolation is to take into account local deviations from the global mean by limiting the range up to which points are used to calculate the mean value for an unmeasured grid node. This method is known as local mean interpolation. Another option is to assign the nearest measured value to each grid node, also known as nearest neighbor interpolation. These three interpolation methods present the most basic methods in this study. Therefore, they act as a baseline for the other interpolation methods. In the next few section we will provide details on four additional interpolation methods.

2.2.1 Inverse distance weighted interpolation

Inverse distance weighted interpolation (IDW) extends on local mean interpolation. Local mean interpolation assigns the same weight to each observation within the search radius. In IDW the weight of an observation is inversely related to its distance to the prediction location. The estimate $\hat{Z}(x_0)$ at an unmeasured location x_0 is given by (Bivand et al., 2008):

$$\hat{Z}(x_0) = \frac{\sum_{i=1}^{n} w(x_i) Z(x_i)}{\sum_{i=1}^{n} w(x_i)} \tag{2.9}$$

where $w(x_i)$ is the weight that the observation at location x_i receives, and $Z(x_i)$ the observation at location x_i. The weight is related to the distance by:

$$w(x_i) = ||x_i - x_0||^{-p} \tag{2.10}$$

where $||\cdot||$ is the Euclidian distance and p is a power that determines how fast the weight drops with distance. Note that when the prediction location is equal to one of the observation locations, the weight of that observation becomes infinite. Consequently, the prediction at that location equals the observation. This means that IDW is an exact interpolator. Furthermore, the weights are always in the range [0,1], so predicted values never fall outside the range of the observations.

In this study we tried three different IDW interpolation with different values for p: 0.5, 2 and 4. When p becomes larger the weight drops faster with distance and the pattern becomes more locally oriented.

2.2.2 Linear regression

Linear regression relies on external variables which show a relation to that which we want to interpolate. For example, if pollution is caused by a leaking underground tank, the concentration of pollutant could be related with distance to the tank. To interpolate, a regression model is fitted between the observations and the external variable. When predicting at an unmeasured location, we invert this model. A simple regression model, or linear model, using one predictor is expressed as:

$$Z(x) = \beta_0 + \beta_1 f(x) + e(x) \tag{2.11}$$

where β_0 is the intercept of the regression model, β_1 the slope, $f(x)$ the value of the external variable at location x, and e the residual to the regression model at location x. An assumption in linear regression is that the residuals are uncorrelated. This means that there is no structure left in the residuals and we can use Ordinary Least Squares fitting (OLS) for the coefficients of the regression model. Note that regression models are not limited to using just one predictor. Multiple predictors can be used jointly for interpolation. Furthermore, predictors can be both continuous or discrete variables.

Literature suggests a number of possible external variables, or predictors, which could be of interest:

- Altitude (Menzel, 1999; McVicar et al., 2007)
- Slope and aspect (McVicar et al., 2007)
- Distance to the coast (Sluiter, 2009)
- Landuse (Menzel, 1999)
- Vegetation type (Jiang et al., 2004)
- Radiation input from remote sensing (Schuurmans, 2009; Gutman and Rukhovetz, 1996)
- Wind (Schuurmans, 2009)

Within the scope of this research we chose to focus on distance to the coast and altitude. Figure 2.1 shows the two maps we used for distance to shore and altitude. The altitude map was based on the Actueel Hoogtemodel van Nederland (AHN) and the distance to shore map was created by calculating the distance to a digitised shoreline in a GIS.

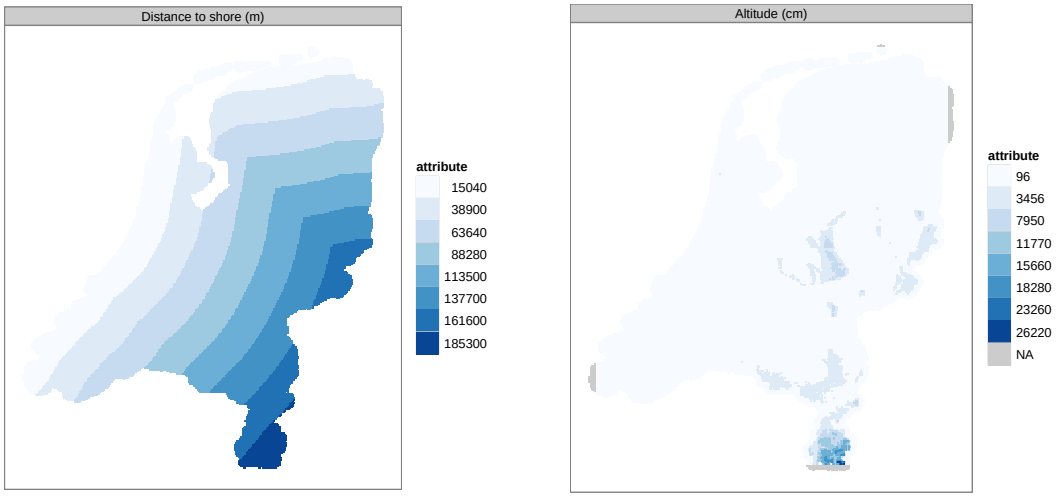


Figure 2.1: Maps of the predictors used in this study: a) distance to shore in meters and b) altitude based on the AHN in cm.

2.2.3 Kriging

Kriging (Chilès and Delfiner, 1999; Christensen, 1996) is a form of linear regression. The difference between kriging and 'normal' linear regression lies in how it deals with the residuals. In linear regression these are assumed to hold no further information and are not used for prediction. In kriging the residuals are spatially correlated and are taken into account during prediction. The underlying linear model is very similar to the model in section 2.2.2:

$$Z(x) = \beta_0 + \beta_1 f(x) + e(x) \tag{2.12}$$

where the difference is that the residuals, $e(x)$, are no longer uncorrelated but have a spatial correlation: $\text{Cov}(e(x), e(x+h)) = C(h)$. The function $C(h)$ is also called the covariance function . An important assumption in kriging is that the covariance between points only depends on the distance between points, not on the location within the study area. This assumption, together with the assumption that the mean of the residual is zero, is called *stationarity*. This form of kriging is known as kriging with external drift (KED) or universal kriging (UK). In this study we will use the term KED. In case the trend component in equation 2.12 is limited to a an *unknown* mean value, the equation reduces to:

$$Z(x) = m(x) + e(x) \tag{2.13}$$

this form of kriging is is known as ordinary kriging (OK). Simple kriging (SK) simplifies even further by assuming that the trend is limited to a *known* mean value. In this study we focused on using KED and OK, and disregarded SK as the mean value was unknown.

An important aspect of kriging is estimating the covariance function of the residuals. In general the covariance function is defined as $C(h) = C(0) - \gamma(h)$, where $\gamma(h)$ is the variogram function, or variogram model, and $C(0)$ equals the sill, or maximum variance, or the variogram function. In case of KED, fitting both the coefficients of the trend, β_0 and β_1, and the variogram

model involves the following procedure. First, we estimate the coefficients using least squares fitting. Second, we determine the residuals to the trend using the fitted coefficients. Third, we construct a so called sample variogram based on the estimated residuals. The sample variogram shows a plot of distance versus semivariance for the observation, i.e. how much do points differ at a certain distance. An option is to show all the point pairs in the plot, but a more common approach is to bin the distance into classes and calculate the average semivariance for those bins. The sample variogram is calculated using the following expression:

$$\hat{\gamma}(\bar{h}_k) = \frac{1}{2N_k} \sum_{i=1}^{N_k} (\hat{e}(x_i) - \hat{e}(x_i + h))^2 \tag{2.14}$$

where $\hat{\gamma}(\bar{h}_k)$ is the estimated average semivariance for bin k, $\bar{h}_k$ is the average distance of bin k, N_k is the number of point pairs in bin k and $\hat{e}(\cdot)$ is the estimated residual at a certain spatial location.

The sample variogram provides the estimated spatial structure of the data. In order to estimate values at unknown locations we need this spatial structure defined not only at the locations of the bins, but throughout the entire range of distance values. This is accomplished by fitting a mathematical function to the sample variogram, the variogram model. The variogram model cannot be any mathematical function, but needs to be conditionally positive definite (Chilès and Delfiner, 1999). This ensures that no negative kriging error variances are produced during prediction of the map.

Fitting the variogram model to the sample variogram can be either by a visual fit or by using a fitting algorithm to minimize some cost function. In this study we used the automatic fitting routine from the R package `automap` (Hiemstra et al., 2009), which builds heavily on the `gstat` package (Pebesma, 2004). The fitting routine uses least squares fitting to find an optimal fit between the variogram model and the sample variogram. In addition, we used a Stein type mode (Ste) because this model can cover a wide range of possible function shapes by varying the kappa parameter.

The coefficients of a variogram model are commonly termed nugget, sill and range. The nugget is the short scale variation and is defined as the point where the variogram model crosses the y-axis. The sill is the maximum semivariance of the variogram model and represents the total variation in the residuals. Finally, the range is the value where the variogram model reaches the sill. Beyond this distance the spatial correlation between points is no longer present.

Once the trend coefficients and the variogram model are estimated from the data, all prediction locations can be visited sequentially to obtain a interpolated map of Makkink evaporation.

2.2.4 Thin plate spline

Splines (Wahba, 1990; Hutchinson and Gessler, 1994) interpolate by fitting a series of polynomials through the observations. *Cubic* splines interpolate using third order polynomials. An important aspect of splines is whether to fit the global pattern or the local pattern. Figure 2.2 illustrates this concept by fitting a 2d cubic spline to a fake dataset. The splines in figure 2.2 show an increased local fit from left to right, better reproducing the local variations. Whether local or global interpolation is preferable depends on the objective of the study and on the data.

To interpolate the evaporation data we used the two-dimensional equivalent of the cubic spline, thine plate splines (Wahba, 1990). Thin plate splines can be compared to bending a thin sheet of metal to match the evaporation data. As with cubic splines, thin plate splines balances local accuracy versus global accuracy. This optimization is done using a cost function which involves a minimum error and the observation locations combined with a minimum in bending

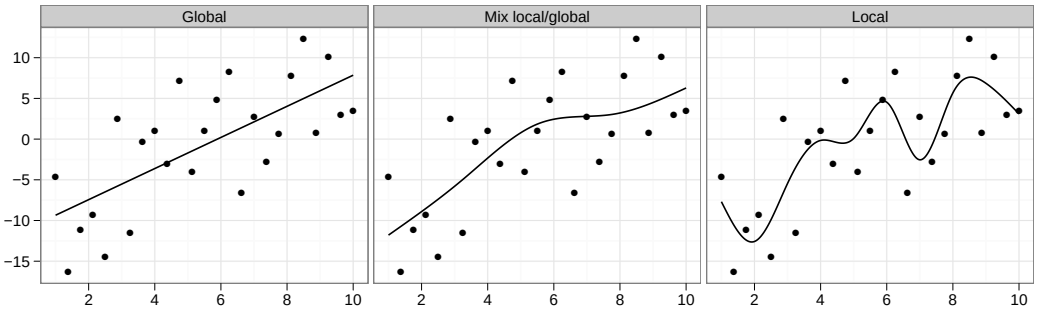


Figure 2.2: Cubic splines fitted to random data focusing on either global or local accuracy.

of the sheet. The minimum error at the observation locations maximizes local accuracy, whilst minimizing the bending of the sheet penalizes over-maximizing the local accuracy, stimulating global accuracy. How strong this effect is is controlled by the λ parameter. In this study we used the function `Tps` from the fields package to calculate the thin plate splines interpolation.

In general λ is estimated form the observations using generalized cross-validation (GCV). For the interpolation of normalen we wanted to keep the ratio of local versus global interpolation consistent to make the interpolated comparable. We chose a λ equal to 0.004, based on expert judgment of the maps.

2.3 Interpolation quality

The purpose of this study was choose an optimal interpolation method for Makkink evaporation. Which interpolation method is optimal partly depends on how we define optimality. In this study we defined optimal as a mix of cross-validation summary statistics and an expert judgment on the maps.

2.3.1 Quality of the fitted trend

A common measure for the quality of a fitted trend is the R^2 which indicates the fraction of variance in the data which is explained by the model. The R^2 for a given regression model is defined as:

$$R^2 = 1 - \frac{SS_e}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}(Z_i - \hat{Z}_i)^2}{\sum_{i=1}^{n}(Z_i - \bar{Z})^2} \tag{2.15}$$

where Z_i is the measured evaporation, $\hat{Z}$ the evaporation estimated by the regression model, and $\bar{Z}$ the mean evaporation. There is no way to link a certain R^2 to an objective measure of quality. Whether or not an R^2 is *good* depends on, among others, the type of data which is modeled and the personal preference of the modeler. The R^2 is primarily used to compare alternative regression models.

2.3.2 Cross-validation

Cross-validation uses all data to train the model. After that, observations are taken out of the dataset one by one, i.e. leave-one-out cross-validation. By predicting that point using the other observations we can asses how effective the interpolation is in predicting itself. By calculating statistics on the cross-validation residuals we can quantify the accuracy of a interpolation method. A common summary statistic is the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{Z}_{cv,i} - Z_i)^2} \tag{2.16}$$

where n is the number of observations, ($\hat{Z}_{cv,i}$ the cross-validation estimate and Z_i the measured observation. A second statistic is the mean error (ME) or bias. The ME indicates whether there is a tendency towards over- or underestimation in the cross-validation residuals ($\hat{Z}_{cv} - Z$):

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\hat{Z}_{cv,i} - Z_i) \tag{2.17}$$

Calculating RMSE and ME for the different interpolation methods allows us to compare their performance. However, these statistics only say something on *relative* performance. It could be that all methods perform poorly, but some perform less poorly than others. To quantify the *actual* performance we defined:

$$RMSE_{SD} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{Z}_{cv,i} - Z_i)^2}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2}} \tag{2.18}$$

where $\bar{Z}$ is the mean of the observations. $RMSE_{SD}$ divides the cross-validation RMSE by the standard deviation of the observations. This indicates how the variance of the cross-validation residuals compares to the variance of the observations. Consequently, the $RMSE_{sd}$ quantifies the performance of an interpolation method relative to interpolating using the mean of the observations.

We used a similar approach for ME by diving it by the mean of the observations:

$$ME_{mean} = \frac{\frac{1}{n}\sum_{i=1}^{n}(\hat{Z}_{cv,i} - Z_i)}{\bar{Z}} \tag{2.19}$$

ME_{mean} provides a sense of scale of the ME relative to the scale of the observations (the mean value).

For interpolating daily evaporation the amount of cross-validations is very large. To visualize the large large amount of cross-validation results we defined the following approach. For each day 10 points are divided amongst the interpolation methods. The division is done based on the $RMSE_{SD}$ score. Interpolation method i on day t will receive a score defined by:

$$\text{score}_{it} = \frac{RMSE_{SD,i}}{\sum_{j=1}^{k} RMSE_{SD,j}} * 10 \tag{2.20}$$

where k is an index for all interpolation methods. Summing the score for a certain time interval (e.g. 1 year or 30 years) will provide a cumulative score for each interpolation method. Dividing the score of each interpolation method by the score of interpolation using the mean provides the skill of the interpolation method to just using interpolation with mean.

2.3.3 Expert judgment

Choosing between interpolation methods using expert judgment on its visual appearance can be somewhat subjective. To remedy some of the subjectiveness we tried to formalize the expert judgment by defining a number of patterns there should or should not be present in the map:

- The detail in the spatial pattern should be reasonable given the amount of observations. If the pattern is too detailed the map suggests an amount of accuracy that is not valid. On the other hand, if a pattern is present it should be represented in the map. Furthermore, too much detail puts emphasis on the local patterns, which might be a problem when we are interested in a globally accurate pattern.

- For the yearly and monthly normals, evaporation is higher at the sea. This is caused by, on average, less cloud cover close to the shore. This effect is strongest at the western shore and less pronounced at the northern wadden shore. Within the year this pattern shifts from west-east to south-west north-east orientation.

- On a daily timescale the patterns should be much more erratic than on a time scale of normals.

Chapter 3

Results

In this chapter we present the results of this study per temporal support level: yearly normals, monthly normals and daily observations of Makkink reference evaporation. The unit of *daily* evaporation used in this study is 10^{-1} mm, the normals have unit mm. We used this odd unit for daily evaporation to be able to store the evaporation data in the database as integers. Note that this choice also influences the units of the derived quantifies such as the sample variogram and the fitted regression parameters.

3.1 Interpolation of yearly normals

Some basic statistics on the yearly normalen are given by table 3.1. Note that for calculation of the yearly normals we only included stations which had a timeseries of Makkink evaporation of more than 20 years.

mean	581.31	mm
sd	21.29	mm
min	544.70	mm
max	614.60	mm

Table 3.1: Basic statistics for yearly normals.

An important matter is how the explanatory variables explain the variation in the evaporation data (`evap`). We fitted linear regression models for all combinations of both candidate explanatory variables: distance to the shore (`dist_shore`) and altitude (`alt`). How well the regression models perform is vital for the trend-based interpolation methods: linear regression and kriging with external drift (KED). Table 3.2 shows the results for each regression model. It is obvious that `log(dist_shore)` performs better than `dist_shore` as a predictor for evaporation, the R^2 becomes almost twice as large. Adding `alt` increases regression model performance slightly: the R^2 increase somewhat. However, the coefficients fitted for altitude are all non-significant (p-value < 0.025). In addition, the regression model with only `alt` performs poorly with a very low R^2 and very high p-value.

Fitting a variogram model is an important step for the four kriging based interpolations we tested. These where ordinary kriging (OK) and kriging with external drift using three regression

model	R^2	dist_shore	dist_shore	dist_shore
evap ∼ dist_shore	0.2153	0.0946		
evap ∼ dist_shore + alt	0.3511	0.0433		0.1575
evap ∼ log(dist_shore)	0.3924		0.0165	
evap ∼ log(dist_shore) + alt	0.5070		0.0082	0.1380
evap ∼ alt	0.0434			0.4745

Table 3.2: R^2 and p-value for linear regression models for each combination of the candidate explanatory variables.

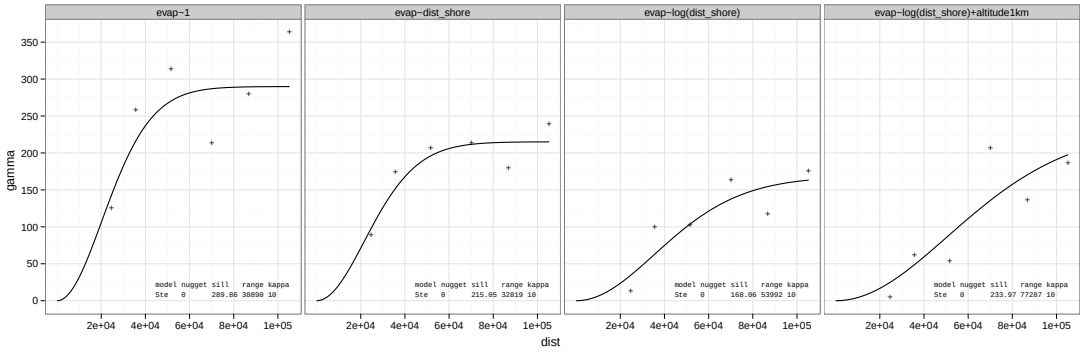


Figure 3.1: Sample variogram and fitted variogram models for the four kriging interpolations.

models: evap ∼ dist_shore, evap ∼ log(dist_shore) and evap ∼ log(dist_shore) + alt. Figure 3.1 shows the variogram models for each kriging option.

A number of important trends are visible in this figure. For the first three variograms there is a clear drop in sill. This is logical because some of the variability in the observations is explained by the trend. The sill is lowest for evap ∼ log(dist_shore), for evap ∼ log(dist_shore) + alt the sill increases again. The trend in the sill seems to following the trend in the R^2, although the lowest sill is not associated to the highest R^2 (evap ∼ log(dist_shore) + alt), but to evap ∼ log(dist_shore). The range stays roughly the same for evap ∼ 1 and evap ∼ dist_shore. However, the range increases for evap ∼ log(dist_shore) and even more for evap ∼ log(dist_shore) + alt, indicating that the spatial trend in the residuals becomes more and more globally oriented.

To choose an optimal interpolation method we looked at two aspect: how does the map look and how does the map perform in cross-validation. First we will look at the visual appearance of the map. The twelve interpolation options we selected are shown in figure 3.2. Judging the visual appearance is done using the aspects mentioned in section 2.3.3. Ordinary kriging (OK), kriging with external drift (KED) with dist_shore, all inverse distance weighted interpolations, nearest neighbor interpolation and interpolation with the mean show a non continuous pattern parallel to the shore. This pattern does not correspond with the expected pattern with high evaporation at the coast. KED with log(dist_shore), KED with log(dist_shore) + alt, both regressions and the thin plate splines (TPS) do show the expected pattern at the coast. TPS distinguishes itself further in the north of the Netherlands. KED and regression show isolines parallel to the coast, TPS shows lines perpendicular to the coast.

In addition to the pattern at the coast, an important aspect was how globally or locally

	Global pattern
1	`mean:evap` $\sim$ `1`
2	`idw05:evap` $\sim$ `1`
3	`regress:evap` $\sim$ `log(dist_shore)`
4	`krige:evap` $\sim$ `log(dist_shore) + alt` (KED)
5	`tps:evap` $\sim$ `1`
6	`krige:evap` $\sim$ `log(dist_shore)` (KED)
7	`krige:evap` $\sim$ `1` (OK)
7	`krige:evap` $\sim$ `dist_shore` (KED)
7	`idw:evap` $\sim$ `1`
10	`krige:evap` $\sim$ `log(dist_shore) + alt` (KED)
11	`idw4:evap` $\sim$ `1`
12	`nn:evap` $\sim$ `1`
	Local pattern

Table 3.3: Ordinal ranking of the interpolated maps in figure 3.2 on a scale from a locally to a globally oriented pattern.

oriented the pattern was. Table 3.3 shows a subjective ordinal ranking of the interpolation methods according to how globally or locally oriented they are. Interpolation with the mean and IDW with inverse distance power 0.5 show a very global pattern. Nearest neighbor and IDW with a power of 4 show a much locally oriented map. TPS and KED with `log(dist_shore)` present a compromise between the global pattern and the local variations. The KED and the regression which use altitude as a predictor show some very local trends. These trends are most likely artifacts of the underlying altitude map used to perform the interpolation. The amount of detail in this map suggests too much accuracy in the map.

In addition to the visual appearance of the maps, we also performed cross-validation to quantify the performance of the interpolation. Figure 3.3 shows the bias and absolute error of each interpolation method, see section 2.3.2 for details on their definition. Judging from the RMSE_{SD} The best performing interpolations are TPS and KED `evap` $\sim$ `log(dist_shore) + alt`, closely followed by KED `evap` $\sim$ `log(dist_shore)`. The regressions perform very poorly. Of the IDW options IDW with a power equal to 4 performs best. Note that interpolations that are a compromise between the global and local pattern perform well in cross-validation. The bias for all interpolation methods is very low compared to the mean value of the observations.

Cross-validation statistics only give the average performance of method. Figure 3.4 shows a spatial plot of the cross-validation residuals, enabling us to visualize any patterns. We want to highlight the following patterns:

- The residuals of interpolation with the mean (`mean:evap` $\sim$ `1`) show underestimation at the coast and overestimation more inland. This confirms that evaporation should be higher at the coast.

- Increased performance of KED by adding altitude (`krige:evap` $\sim$ `log(dist_shore)` vs `krige: evap` $\sim$ `log(dist_shore) + alt`) is mainly caused by an improved performance at the southern most stations (in Limburg).

- The performance of TPS versus the two KED's which perform well (`krige: evap` $\sim$ `log(dist_shore)` and `krige:evap` $\sim$ `log(dist_shore)+ alt` is mainly better in the

north. This indicates that the trend with distance to the coast is not as prominent in the north. TPS performs better there because it does not assume any trend.

- The regression cross-validation results still show a lot of pattern, much more than the KED with the same regression model. This indicates that interpolation of the residuals from the regression model boosts interpolation performance.

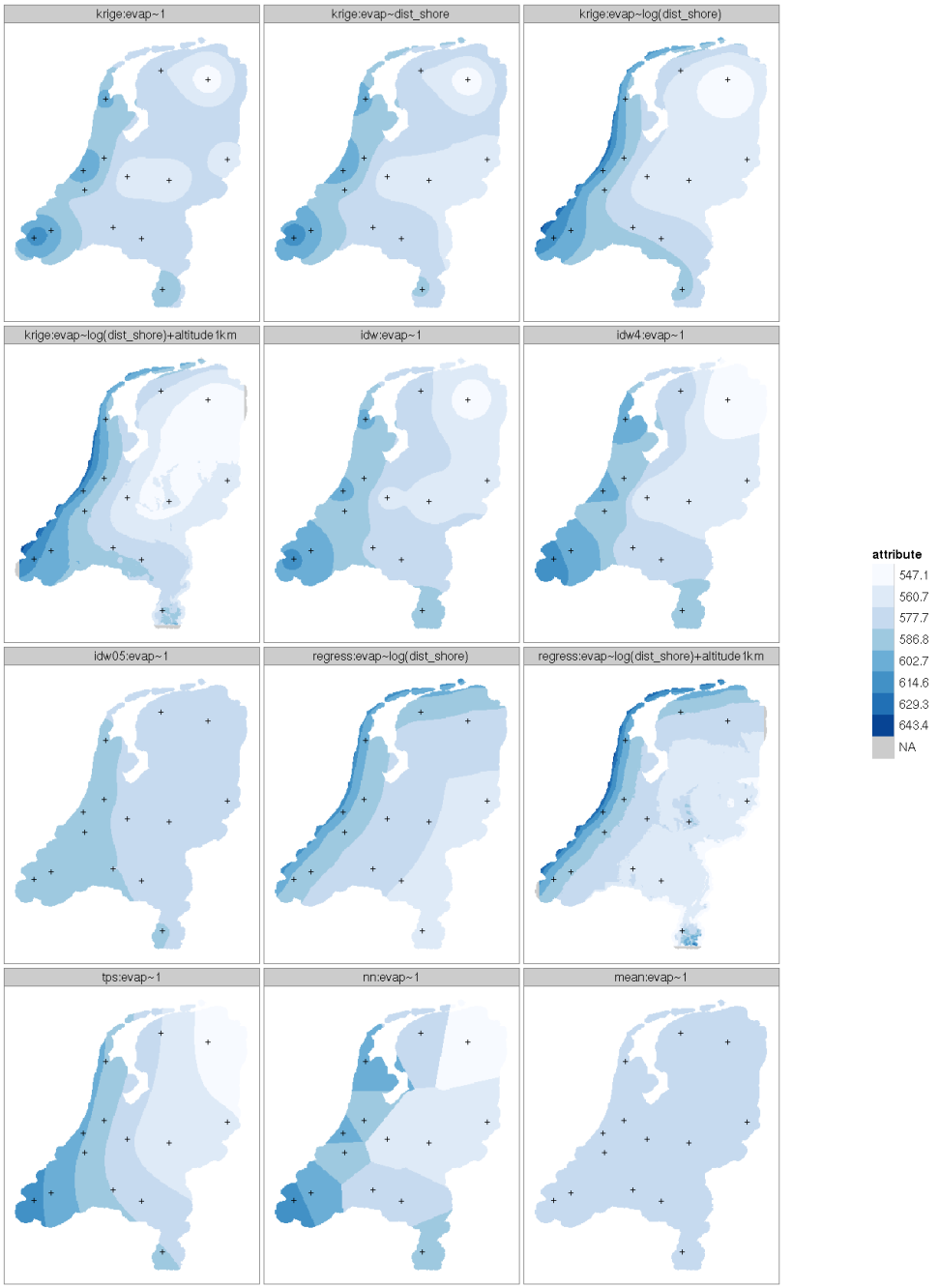


Figure 3.2: Interpolated maps of yearly normal Makkink reference evaporation. The names used in this figure are: krige for kriging, idw for inverse distance weighted interpolation, regress for linear regression interpolation, tps for thin plate splines, nn for nearest neighbor and mean for mean interpolation. The text behind the $\sim$ shows the explanatory variables used.

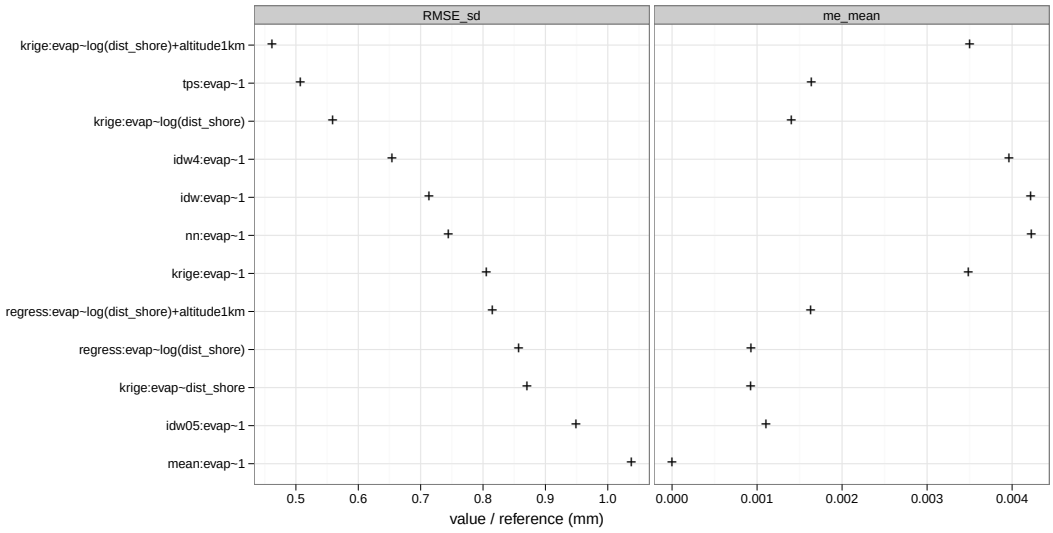


Figure 3.3: Cross-validation statistics for the different interpolation options: (root mean squared error / sd observed) and (bias / mean observed).

Figure 3.4: Map of the cross-validation residuals for all interpolation methods considered. Red is overestimation, green is underestimation.

3.2 Interpolation of monthly normals

Interpolation of monthly normals shows a lot of similarity to the interpolation of yearly normals. This section will focus on the differences and similarities between the monthly and yearly interpolation results. As with yearly normals, we excluded stations with only 20 years of available data. In addition, we only included stations for which we had normals for each month. Finally, we excluded stations `CABAUW` and `WILHELMINADORP`. Both these stations showed large local variations with surrounding stations, making variogram modeling difficult. Taking into account these considerations, we calculated the sample variograms based on 12 observations for each month.

month	mean_evap	sd_evap
January	8.41	0.74
February	15.15	1.01
March	33.99	1.60
April	60.16	1.74
May	86.98	2.24
June	93.66	2.96
July	98.25	2.97
August	82.53	2.85
September	50.88	1.82
October	28.07	1.31
November	11.17	0.95
December	6.32	0.45

Table 3.4: Mean and standard deviation of Makkink evaporation monthly normals (mm).

Table 3.4 shows the mean and standard deviation of the observations for each month. The general trend is obvious: high evaporation in summer and low in winter.The results in appendix A.1 show the regression results for each month. In the 'warm' months, March to September, the trend with distance to the coast shows a relatively high R^2 ($>\sim$0.25) and small p-values. In addition, the results are much better for `log(dist_shore)` than for `dist_shore`. This R^2 and p-values indicate that distance to the shore is a significant trend in these warm months. However, in the remaining 'cold' months the R^2 is low and the p-values are large. So, in these cold months the trend with distance to the shore is not significant.

In the warm period the variogram models, shown in appendix A.2, show roughly the same pattern as the yearly normals. The sill drops when more predictors are used and the range becomes larger. In the cold period the sill and range stay roughly the same, indicating that the predictors do not explain part of the variance in the observations. This confirms the low R^2 and high p-value in appendix A.1: in the cold period the predictors are not successful.

The maps for each month in appendix A.3 show a comparable relation to the yearly normals. In the warm period the maps show a pattern similar to the yearly normals: high evaporation near the shore, low evaporation more inland. In the cold period the spatial pattern becomes more north-south oriented, which again confirms that the trend with distance to shore is not present in the cold period. This is confirmed by the fact that the maps of the four kriging options look roughly same for the cold period. The pure regression based interpolations still show a shore oriented trend in the cold period, whilst it is clear that there is no trend. Ordinary kriging and TPS both show a good compromise between local accuracy and global accuracy. The IDW interpolations with power 2 and 4 show a tendency towards a local pattern.

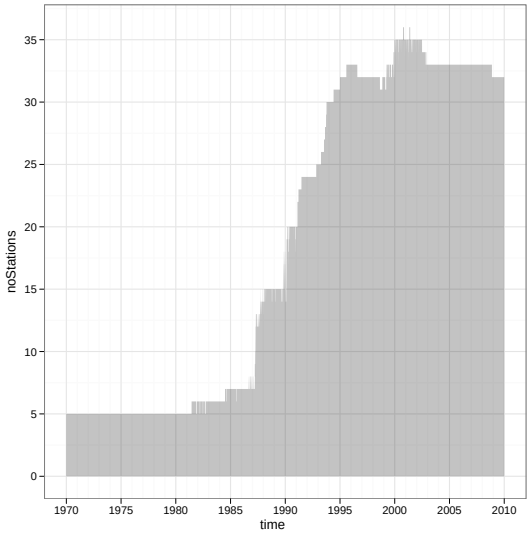


Figure 3.5: Number of stations which can be used for interpolation versus time.

Finally, we quantified the performance of each interpolation method using cross-validation. The general pattern is that TPS and ordinary kriging show the best performance. This indicates that they are most successful in adapting to the changing pattern during the year. In the warm period, methods that take into account the trend show a sharp increase in performance, e.g. the results for May. Overall, the observation oriented methods such as TPS and OK perform better than the trend oriented and mixed trend/observations methods such as KED and linear regression.

3.3 Interpolation of daily observations

The sheer number of interpolated maps makes interpreting the results of daily observations more complex than that of the normalen. So, the results for daily observations are not focused on variogram models and maps, but more on aggregated cross-validation results. However, some differences and similarities between the daily observations and the normalen are important to note:

- The number of available monitoring stations is not constant, figure 3.5 show how the number of stations changes in time. This is caused by a limited number of global radiation observations for those years.

- The yearly temporal trend is similar to that in the monthly normals, high evaporation in summer, low in winter. Figure 3.6 shows a timeseries of the mean daily evaporation from 1970 to 2010.

- The trend of high evaporation at the coast and low inland is not consistently present in the daily maps. Of number of regression models from 2008 illustrates this point nicely, see figure 3.7. Some days there is no trend at all, and on the days that the trend is present it is not always a higher evaporation at the coast.

To provide a sense of how interpolated maps look like on a daily basis, appendix B provides a few weeks of maps. The figure shows the first week of January and July for 1970, 1989 and 2008. Although the maps we selected to show are arbitrary, they provide a nice illustration. Most prominent is that the patterns are much more erratic than for the maps of normals. Going from 1970 to 2008 shows an increased amount of detail as the number of available observations increases.

To quantify interpolation performance across the daily evaporation dataset, we calculated the skill each method had in contrast to interpolation using the mean. Figure 3.8 shows this skill. The results have split out into several categories: low (< 12), medium (> 12 and ≤ 18) and high (> 18) amounts of observations and period in the year, warm (March-September) or cold. For all categories TPS performs best. A striking outlier is the performance of the interpolations that include altitude as a trend for low amounts of observations. The skill relative to the mean interpolation increases when the number of observations increases. This is logical given that more observations allow an interpolation method to insert much more detail into a map, which is nog present in the mean interpolation. In addition, performance in the warm period is a little better than in the cold period.

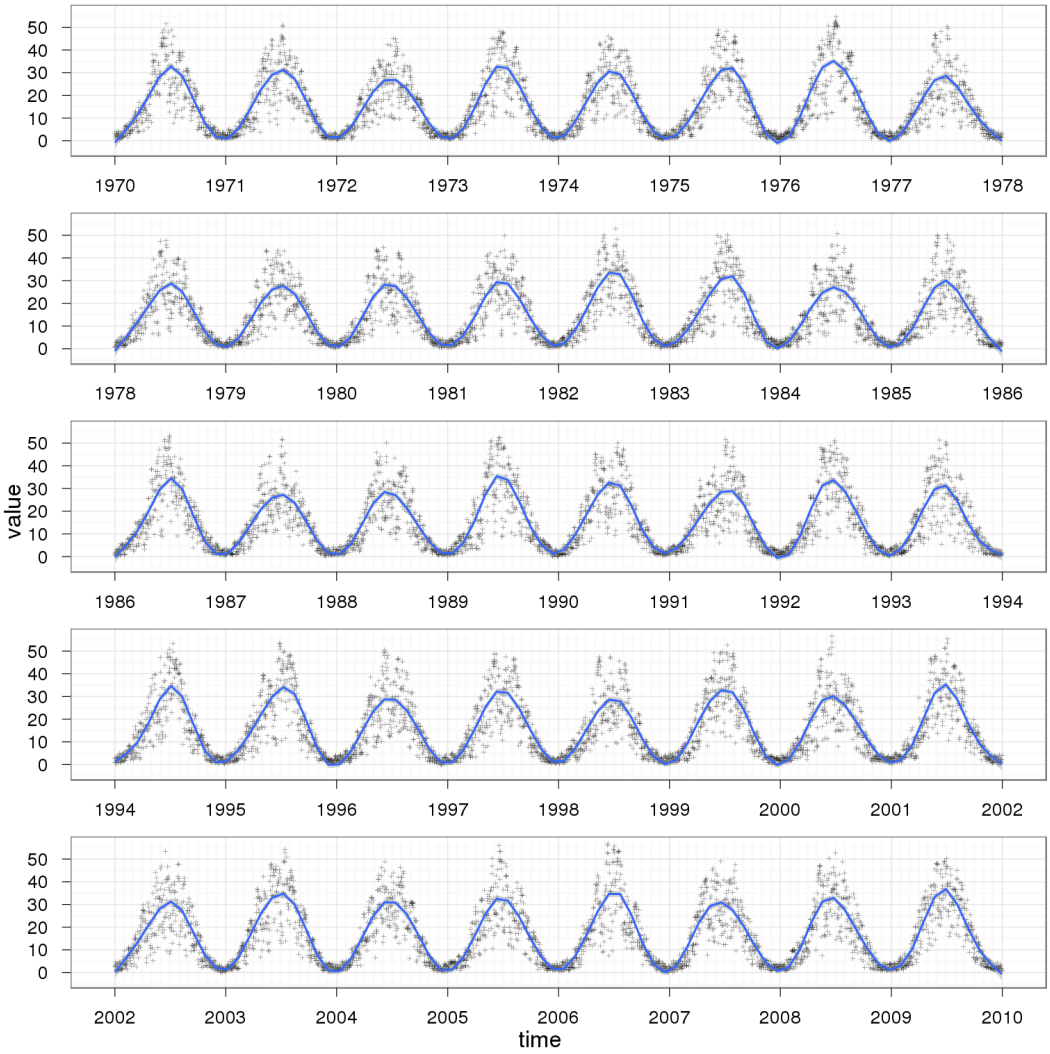


Figure 3.6: Timeseries of daily mean Makkink evaporation (10^{-1}mm) for the period 1970 to 2010. The blue line is a fitted spline.

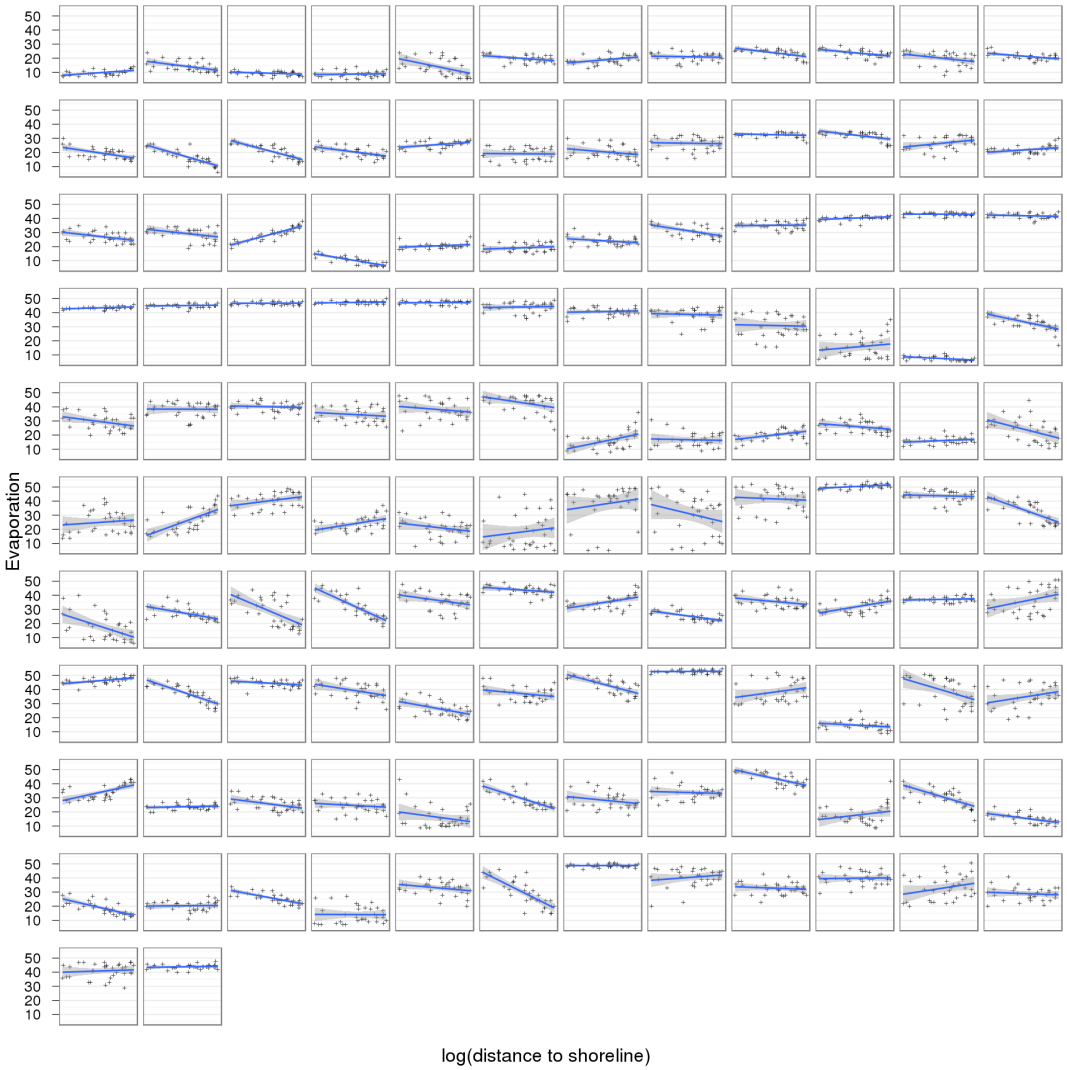


Figure 3.7: Regressions of `log(dist_shore)` versus `evap` (10^{-1}mm) per day in the period April - July 2008. Time progresses from top left to bottom right.

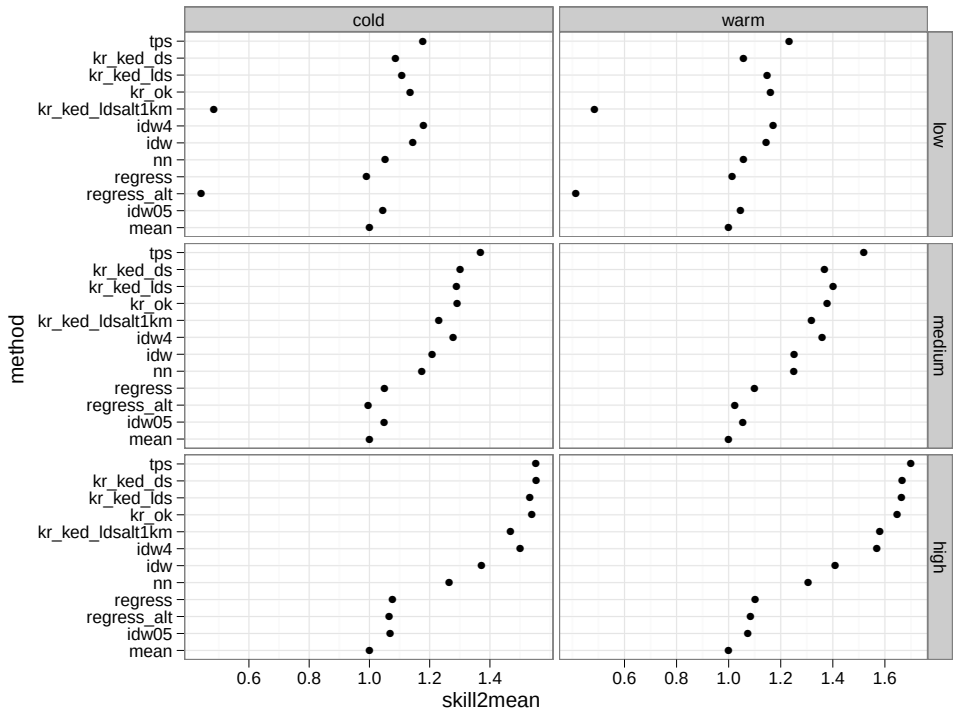


Figure 3.8: Skill score relative to the mean for each interpolation method for the warm and cold periods (columns) and low to high amounts of observations (rows).

Chapter 4

Discussion and Conclusion

4.1 Optimal interpolation

The optimal interpolation method for Makkink reference evaporation is thin plate splines (TPS). This conclusion holds under all different circumstances we took into consideration: for all amounts of available stations, for yearly and monthly normals, for daily observations and in the cold and warm period of the year. For interpolation of normals we base this conclusion on the cross-validation results combined with expert judgment on the pattern in the maps. TPS showed a low value for $RMSE_{SD}$ and showed all the patterns expected of the map. Furthermore, the map showed the amount of detail we expect based on the amount of stations used during interpolation. For daily observations we relied primarily on the cross-validation results. These show that, on average, TPS performs best.

Examples of other studies using thin plate splines for interpolation of meteorological variables include Jeffrey et al. (2001) and Price et al. (2000). These two studies used ANUSPLIN (Hutchinson, 1995). A difference between those studies and this current study is that they used altitude as an additional independent variable. We did not deem this necessary as the Netherlands is quite flat.

A number of research question where addressed in the introduction. We will discuss these issues in the next section. Finally, we will discuss some issues which are not directly related to any of the research questions.

4.2 Research questions

Number of stations
How many observations are necessary for an accurate map? This question is not straightforward to answer as this depend heavily on the type of data to be interpolated and on the use of the maps afterwards. For kriging a popular rule of thumb is to use a minimum of 30 observations to estimate the variogram (Journel and Huijbregts, 1978). According to this criterion almost none of our interpolated kriging maps are good enough. But how many observations one needs also depends on the spatial variability in the data. If the data exhibits a lot of local details, many observations are needed to interpolate a map accurately.

The normalen maps are much smoother than the daily maps. Consequently, we believe less observations are needed to interpolate those accurately. Expert judgment on the TPS maps suggested the normalen maps showed the patterns and detail we expected. Therefore, the 13

observations of normalen where enough to interpolate these normalen. Of course, such an expert judgment is to a degree subjective.

Daily evaporation maps are less smooth. The interpolated maps show that increasing the amount of stations increase the detail of the patterns in the map, especially going from low to medium amount of points. This suggests that the actual pattern is more detailed than the pattern emerging from the interpolation based on 5 points. This effect is also present when going from medium to high amount of points, but the effect is smaller. This is confirmed by the cross-validation results. The performance of the interpolation methods relative to using the mean for interpolation increases when increasing the number of points. We suggest to at least a medium amount of points ($\sim$15) should be used to interpolate daily evaporation.

Time of the year
The monthly normals show a change in spatial pattern between the warm and cold period. In the cold period the pattern is more north-sout oriented, while in the warm the pattern is west-east oriented parallel to the coast line. This is obvious from the interpolated maps and is confirmed by the cross-validation results. These results show that interpolation with regression based on distance to shore performs better in the warm period than in the cold period. The kriging interpolation that use this trend suffer less in the cold period. Interpolation of the residuals to the trend can correct some of the error introduced by using the faulty trend. The cross-validation performance of TPS is good for both the cold and warm period, almost always ranking first or second.

The cross-validation results for interpolation of daily observations show that the performance of the interpolation methods is better in the warm period. On average this is by a factor of 1.4. However, the ranking of the interpolation methods does not change, with TPS ranking first.

Temporal support
For all temporal supports TPS performs best. Important is that the pure regression based interpolations perform less for daily observations than for normalen. In addition, data oriented interpolation methods (e.g. ordinary kriging) show an increased performance when interpolating daily observations.

Spatial trends
From literature a number of predictors where deemed important for interpolating evaporation (Schuurmans, 2009; McVicar et al., 2007; Jiang et al., 2004; Menzel, 1999; Gutman and Rukhovetz, 1996). In this study we focused on distance to the shore and altitude. Although our results show that these trends influence the maps and cross-validation results significantly, we did not choose an interpolation method which incorporated them. The rationale behind is as follows. Incorporating a trend focuses the map into that trend. For example, incorporating the distance to the sea makes that trend very dominant in the resulting map. The lack of detailed validation data makes it hard to decide if this dominant pattern is realistic or not. TPS does not rely on the trend and lets the observations dictate the pattern.

In addition to dominating the pattern, incorporating a trend can also introduce a false sense of accuracy. Many studies use regression based interpolation, e.g. Agnew and Palutikof (2000). These maps tend to show a lot of detail, especially when based on high resolution DEM's. An example of such a map in our study was the mapping of yearly and monthly normals with KED using distance to shore and altitude. We find, based on expert judgment, that the detail in this map is too high. This suggests a level of map accuracy which is not realistic given the amount of stations we have available. An argument against this could be that when (cross-)validation results favor the detailed map, that it should be favored over other methods. However, the small

support of the observations combined with the large size of the study area makes validation of a detailed pattern questionable. When a DEM is used in regression, we suggest smoothing the DEM to obtain a level of detail in the DEM which matches the level of detail appropriate for the amount stations available. In our opinion, validation and expert judgment should go hand in hand when assessing the quality of an interpolated map.

The impact of distance to the shore is significant in the Netherlands. In other areas which are less influenced by the sea, this predictor is likely to be unimportant for interpolating Makkink reference evaporation. Many studies include slope and aspect as predictors (Agnew and Palutikof, 2000; Vicente-Serrano et al., 2007). We expect this to have little influence in a flat country such as the Netherlands.

4.3 Recommendations for future research

A number of research topics emerged during this study, which could prove to be interesting:

Definition of optimality
To determine which interpolation method is optimal, it is vital to define what exactly we mean by optimal. In this study we defined optimal by looking at the predictive power of the interpolation method: how well could the method reproduce the observations. The maps of daily evaporation are used for modeling studies. An alternative definition of optimality would be to see how the hydrological model performs using different interpolated datasets. This is not necessarily the same method we found optimal. The result could be that interpolation using a mean value is more than enough for the model to perform well. This in turn depends on the sensitivity of the hydrological model to evaporation.

Interpolation of anomalies
For interpolation of daily observations we interpolated the observations as they where. Another option is to first calculate long term averages, i.e. normals, for each station and correct each station for its normal. These corrected values are called anomalies. In this context we can use either yearly or monthly normals, most likely monthly normals are advisable. Interpolation would than become a stage process of calculating the normals, subtracting them from the observations, interpolating these anomalies and than add the interpolated normals. An advantage of this approach is that within the daily dataset we guarantee that the pattern prevalent in the normals is also present in the daily dataset.

Bibliography

Agnew, M. D., Palutikof, J. P., 2000. GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. Climate Research 14, 115–127.
URL PDF

Bivand, R., Pebesma, E., Gomez-Rubio, V., 2008. Applied Spatial Data Analysis with R. Springer, New York.

Buishand, T., Beersma, J., Sluiter, R., Kroon, T., 2008. Definitiestudie rasterdata meteorologie (in dutch). Tech. rep., Royal Netherlands Meteorological Institute (KNMI).

Buishand, T., Velds, C., 1980. Klimaat van Nederland 1: Neerslag en verdamping. Royal Netherlands Meteorological Institute (KNMI).

Chilès, J. P., Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. John Wiley & Sons, New York, 720p.

Christensen, R., 1996. Plane Answers to Complex Questions: The Theory of Linear Models, 2nd Edition. Springer, New York, 496p.

de Bruin, H., 1987. From penman to makkink. In: Hooghart, C. (Ed.), Evaporation and Weather: Proceedings and Information. Vol. 28. TNO committee on Hydrological Research: The Hague, pp. 5–30.

Gurtz, J., Baltensweiler, A., Lang, H., 1999. Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountainous basins. Hydrological Processes 13 (17), 2751–2768.

Gutman, G., Rukhovetz, L., 1996. Towards satellite-derived global estimation of monthly evapotranspiration over land surfaces. Advances in Space Research 18 (7), 67–71.

Hiemstra, P. H., Pebesma, E. J., Twenhöfel, C. J. W., Heuvelink, G. B. M., 2009. Real-time automatic interpolation of ambient gamma dose rates from the dutch radioactivity monitoring network. Computers & Geosciences 35 (8), 1711–1721.

Hutchinson, M., 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Science 9 (4), 385–403.

Hutchinson, M., Gessler, P., 1994. Splines - more than just a smooth interpolator. Geoderma 62 (1-3), 45–67.

Jeffrey, S., Carter, J., Moodie, K., Beswick, A., 2001. Using spatial interpolation to construct a comprehensive archive of australian climate data. Environmental Modelling and Software 16 (4), 309–330.

Jiang, H., Liu, S., Sun, P., An, S., Zhou, G., Li, C., Wang, J., Yu, H., Tian, X., 2004. The influence of vegetation type on the hydrological process at the landscape scale. Canadian Journal of Remote Sensing 30 (5), 743–763.

Journel, A., Huijbregts, C., 1978. Mining Geostatistics. Academic Press., London, 900p.

McVicar, T., Van Niel, T., Li, L., Hutchinson, M., Mu, X., Liu, Z., 2007. Spatially distributing monthly reference evapotranspiration and pan evaporation considering topographic influences. Journal of Hydrology 338 (3-4), 196–220.

Menzel, L., 1999. Modelling of evapotransporation with tran.

Ni, G., Li, X., Cong, Z., Sun, F., Liu, Y., 2006. Temporal and spatial characteristics of reference evapotranspiration in china. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering 22 (5), 1–4.

Pebesma, E. J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30 (7), 683–691.

Price, D., McKenney, D., Nalder, I., Hutchinson, M., Kesteven, J., 2000. A comparison of two statistical methods for spatial interpolation of canadian monthly mean climate data. Agricultural and Forest Meteorology 101 (2-3), 81–94.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
URL `http://www.R-project.org/`

Schuurmans, H., 2009. Penman-monteith referentieverdamping: Inventarisatie beschikbaarheid en mogelijkheden tot regionalisatie. (in dutch). Tech. rep., FutureWater report 86.

Sluiter, R., 2009. Interpolation methods for climate data: literature review. Tech. rep., Royal Netherlands Meteorological Institute (KNMI), iR 2009-04.

Vicente-Serrano, S. M., Lanjeri, S., Lopez-Moreno, J. I., 2007. Comparison of different procedures to map reference evapotranspiration using geographical information systems and regression-based techniques. Int. J. Climatol. 27 (8), 1103–1118.

Wahba, G., 1990. Spline models for observational data. Vol. 59. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Winter, T. C., Rosenberry, D. O., Sturrock, A. M., 1995. Evaluation of 11 equations for determining evaporation for a small lake in the north central united states. Water Resources Research 31 (4), 983–993.

Appendix A

Interpolation of monthly normalen

A.1 Regression results

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.0304	0.5879		
VARIABLE ~ dist_shore + altitude1km	0.2280	0.1468		0.1634
VARIABLE ~ log(dist_shore)	0.0866		0.3532	
VARIABLE ~ log(dist_shore) + altitude1km	0.2376		0.1369	0.2146
VARIABLE ~ altitude1km	0.0117			0.7383

Table A.1: Regression results for January

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.1345	0.2410		
VARIABLE ~ dist_shore + altitude1km	0.3456	0.0577		0.1226
VARIABLE ~ log(dist_shore)	0.2730		0.0814	
VARIABLE ~ log(dist_shore) + altitude1km	0.4208		0.0311	0.1640
VARIABLE ~ altitude1km	0.0018			0.8966

Table A.2: Regression results for February

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.2215	0.1226		
VARIABLE ~ dist_shore + altitude1km	0.3619	0.0590		0.1928
VARIABLE ~ log(dist_shore)	0.3530		0.0416	
VARIABLE ~ log(dist_shore) + altitude1km	0.4260		0.0345	0.3125
VARIABLE ~ altitude1km	0.0309			0.5850

Table A.3: Regression results for March

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.3960	0.0284		
VARIABLE ~ dist_shore + altitude1km	0.5053	0.0251		0.1921
VARIABLE ~ log(dist_shore)	0.6003		0.0031	
VARIABLE ~ log(dist_shore) + altitude1km	0.6481		0.0048	0.2977
VARIABLE ~ altitude1km	0.1099			0.2926

Table A.4: Regression results for April

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.3198	0.0553		
VARIABLE ~ dist_shore + altitude1km	0.3730	0.0852		0.4048
VARIABLE ~ log(dist_shore)	0.5463		0.0060	
VARIABLE ~ log(dist_shore) + altitude1km	0.5799		0.0115	0.4181
VARIABLE ~ altitude1km	0.1125			0.2864

Table A.5: Regression results for May

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.3291	0.0511		
VARIABLE ~ dist_shore + altitude1km	0.4487	0.0359		0.1959
VARIABLE ~ log(dist_shore)	0.5520		0.0056	
VARIABLE ~ log(dist_shore) + altitude1km	0.6211		0.0058	0.2321
VARIABLE ~ altitude1km	0.0768			0.3833

Table A.6: Regression results for June

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.2865	0.0729		
VARIABLE ~ dist_shore + altitude1km	0.4582	0.0275		0.1256
VARIABLE ~ log(dist_shore)	0.5277		0.0075	
VARIABLE ~ log(dist_shore) + altitude1km	0.6440		0.0036	0.1205
VARIABLE ~ altitude1km	0.0427			0.5194

Table A.7: Regression results for July

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.1970	0.1484		
VARIABLE ~ dist_shore + altitude1km	0.3973	0.0401		0.1178
VARIABLE ~ log(dist_shore)	0.3824		0.0321	
VARIABLE ~ log(dist_shore) + altitude1km	0.5216		0.0128	0.1400
VARIABLE ~ altitude1km	0.0126			0.7279

Table A.8: Regression results for August

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.1167	0.2771		
VARIABLE ~ dist_shore + altitude1km	0.3698	0.0472		0.0897
VARIABLE ~ log(dist_shore)	0.2650		0.0868	
VARIABLE ~ log(dist_shore) + altitude1km	0.4562		0.0226	0.1090
VARIABLE ~ altitude1km	0.0000			0.9920

Table A.9: Regression results for September

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.0022	0.8847		
VARIABLE ~ dist_shore + altitude1km	0.2031	0.2789		0.1662
VARIABLE ~ log(dist_shore)	0.0120		0.7348	
VARIABLE ~ log(dist_shore) + altitude1km	0.2371		0.2141	0.1377
VARIABLE ~ altitude1km	0.0856			0.3561

Table A.10: Regression results for October

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.0019	0.8940		
VARIABLE ~ dist_shore + altitude1km	0.2292	0.1874		0.1377
VARIABLE ~ log(dist_shore)	0.0351		0.5597	
VARIABLE ~ log(dist_shore) + altitude1km	0.2504		0.1599	0.1424
VARIABLE ~ altitude1km	0.0549			0.4634

Table A.11: Regression results for November

model	rsq	ds_pval	lds_pval	a1km_pval
VARIABLE ~ dist_shore	0.0119	0.7359		
VARIABLE ~ dist_shore + altitude1km	0.2384	0.1530		0.1363
VARIABLE ~ log(dist_shore)	0.0693		0.4083	
VARIABLE ~ log(dist_shore) + altitude1km	0.2774		0.1145	0.1419
VARIABLE ~ altitude1km	0.0322			0.5765

Table A.12: Regression results for December

A.2 Variogram models

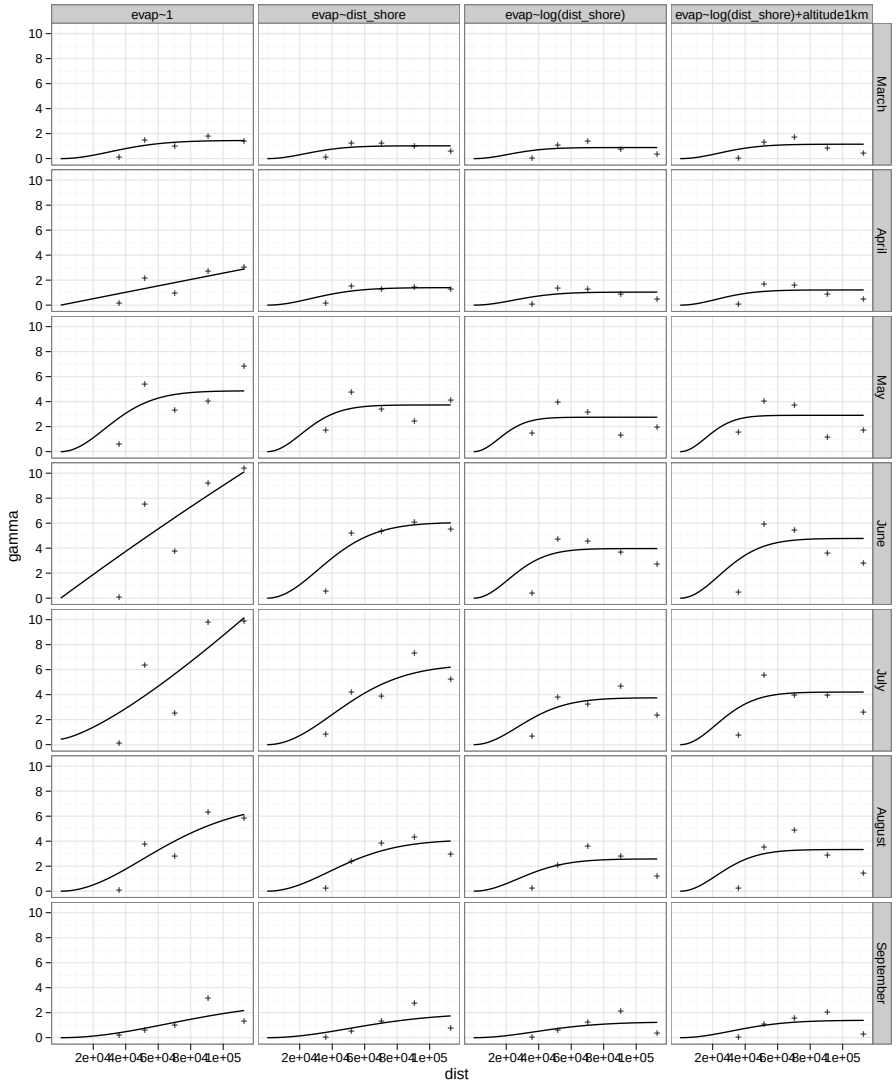


Figure A.1: Variograms for the 'warm' months, March through September.

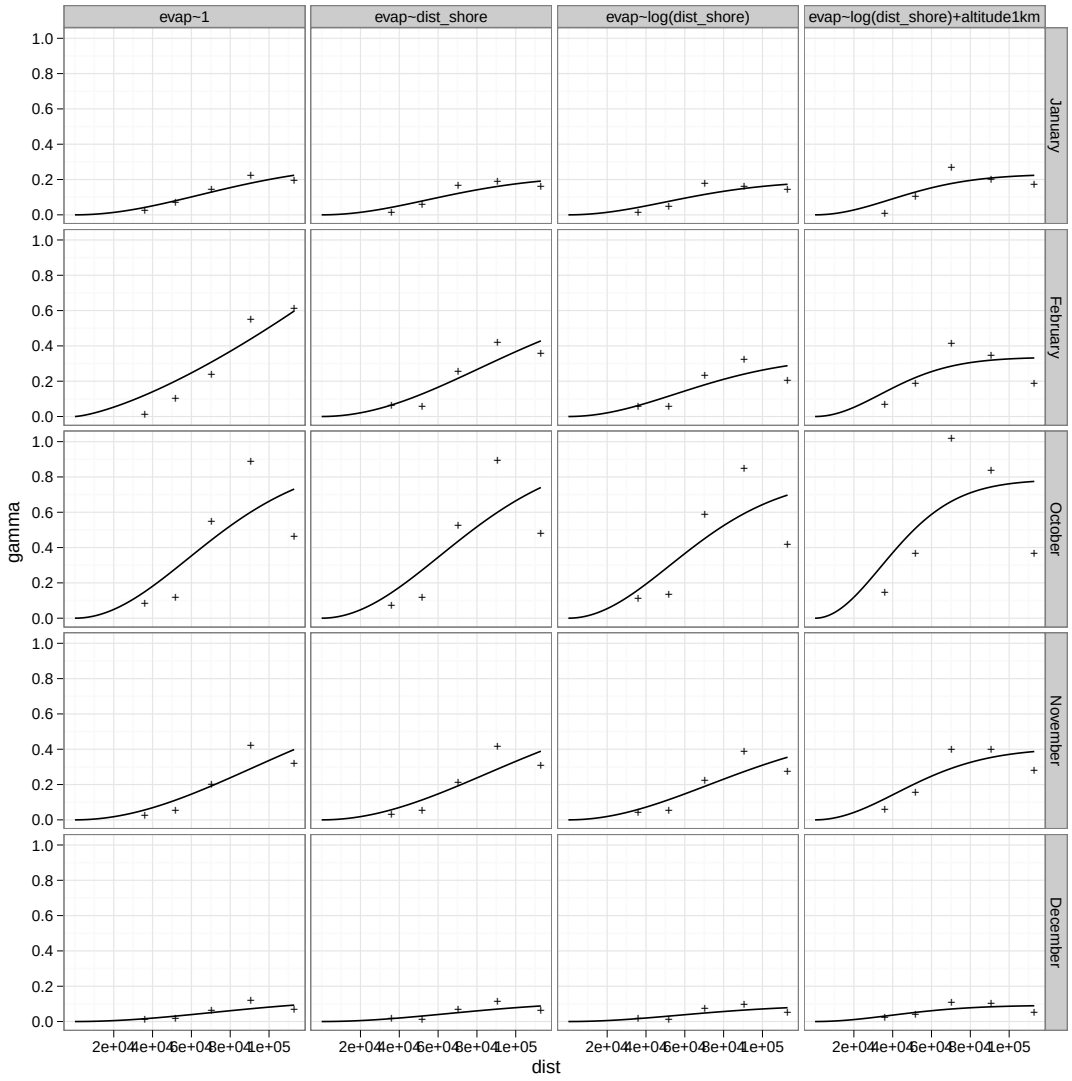


Figure A.2: Variogram for the cold months, January, February, October, November and December.

A.3 Interpolated maps

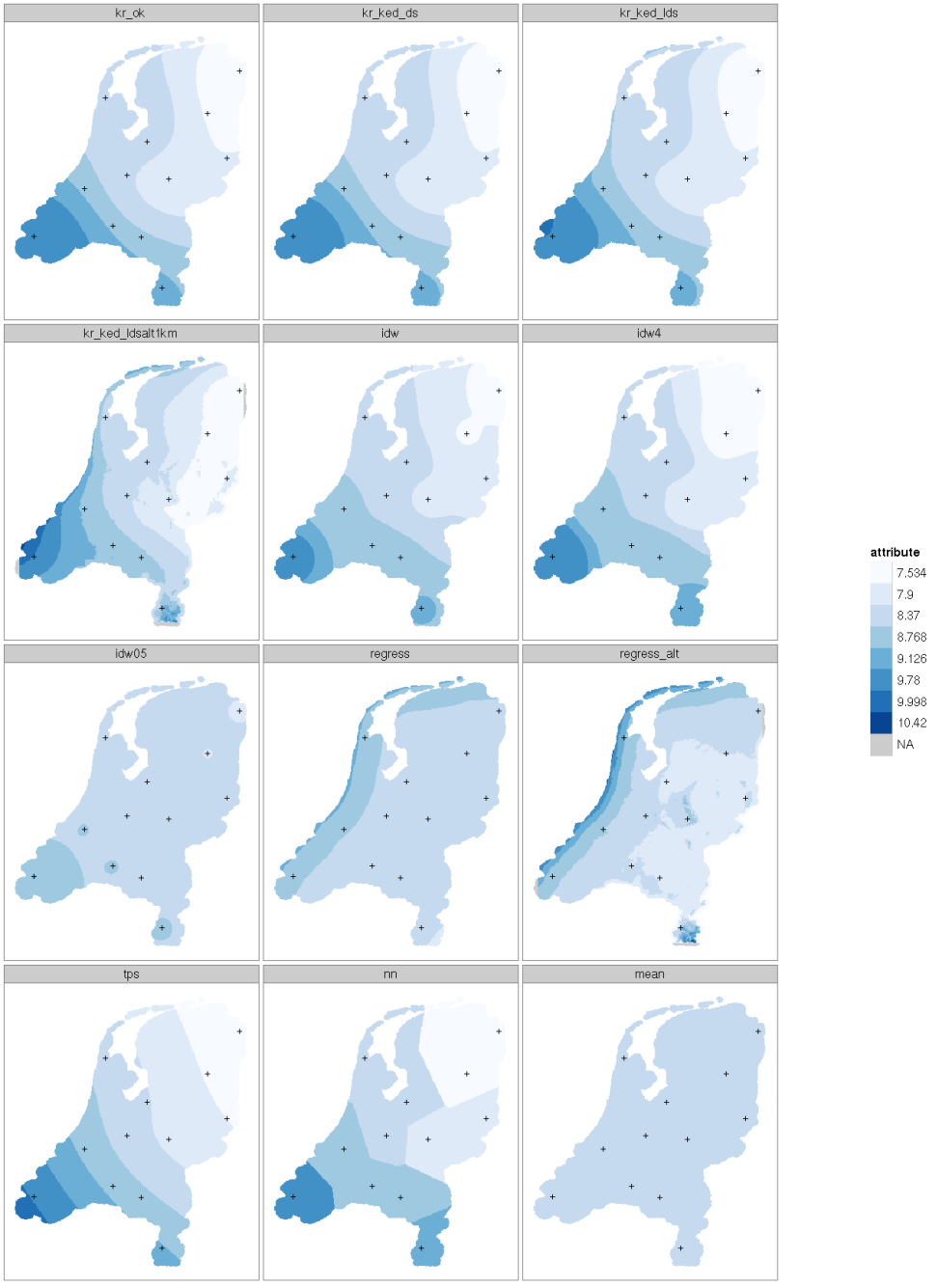


Figure A.3: Interpolated maps of 30 year monthly average for January

Figure A.4: Interpolated maps of 30 year monthly average for February

Figure A.5: Interpolated maps of 30 year monthly average for March

Figure A.6: Interpolated maps of 30 year monthly average for April

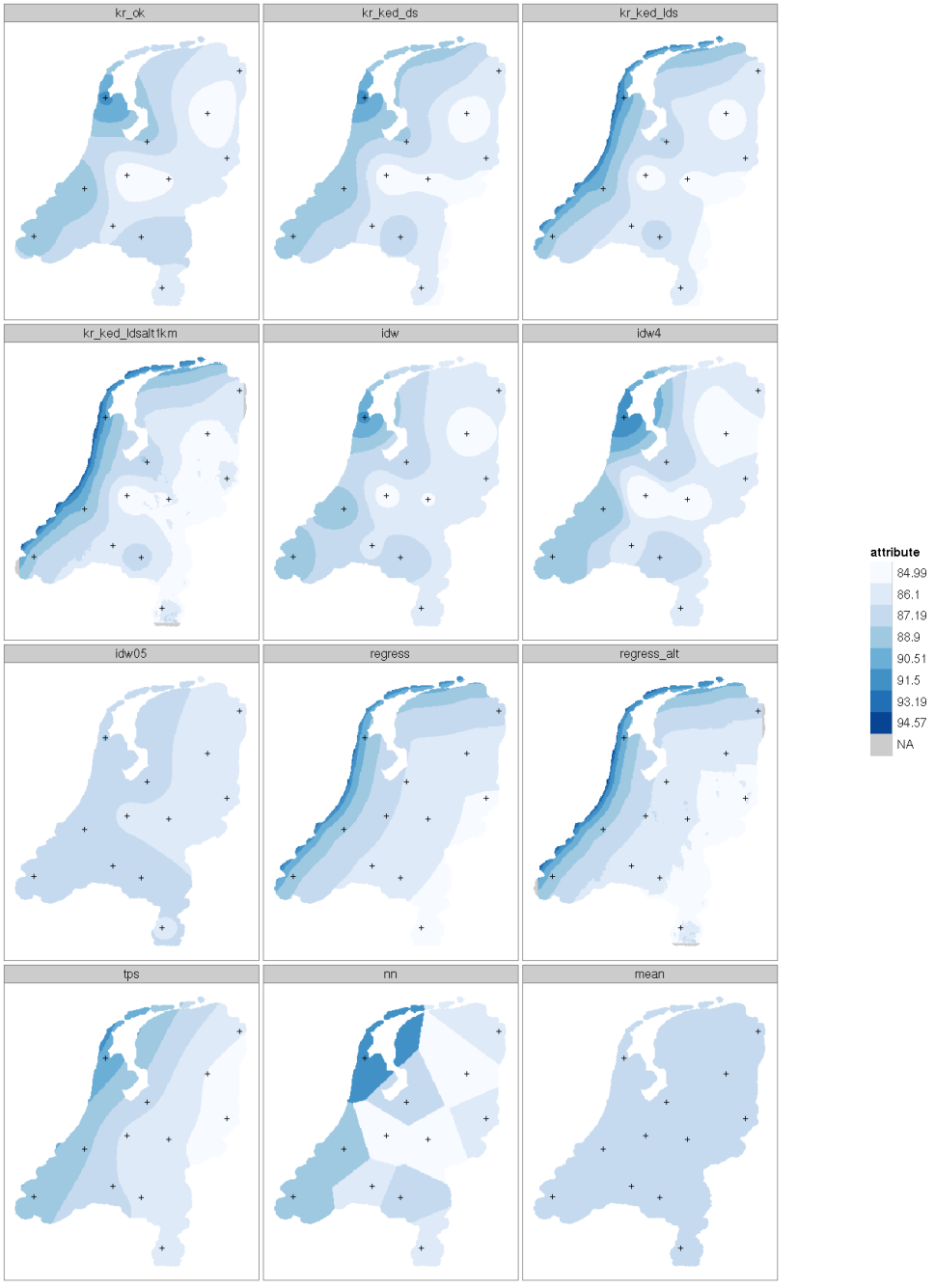


Figure A.7: Interpolated maps of 30 year monthly average for May

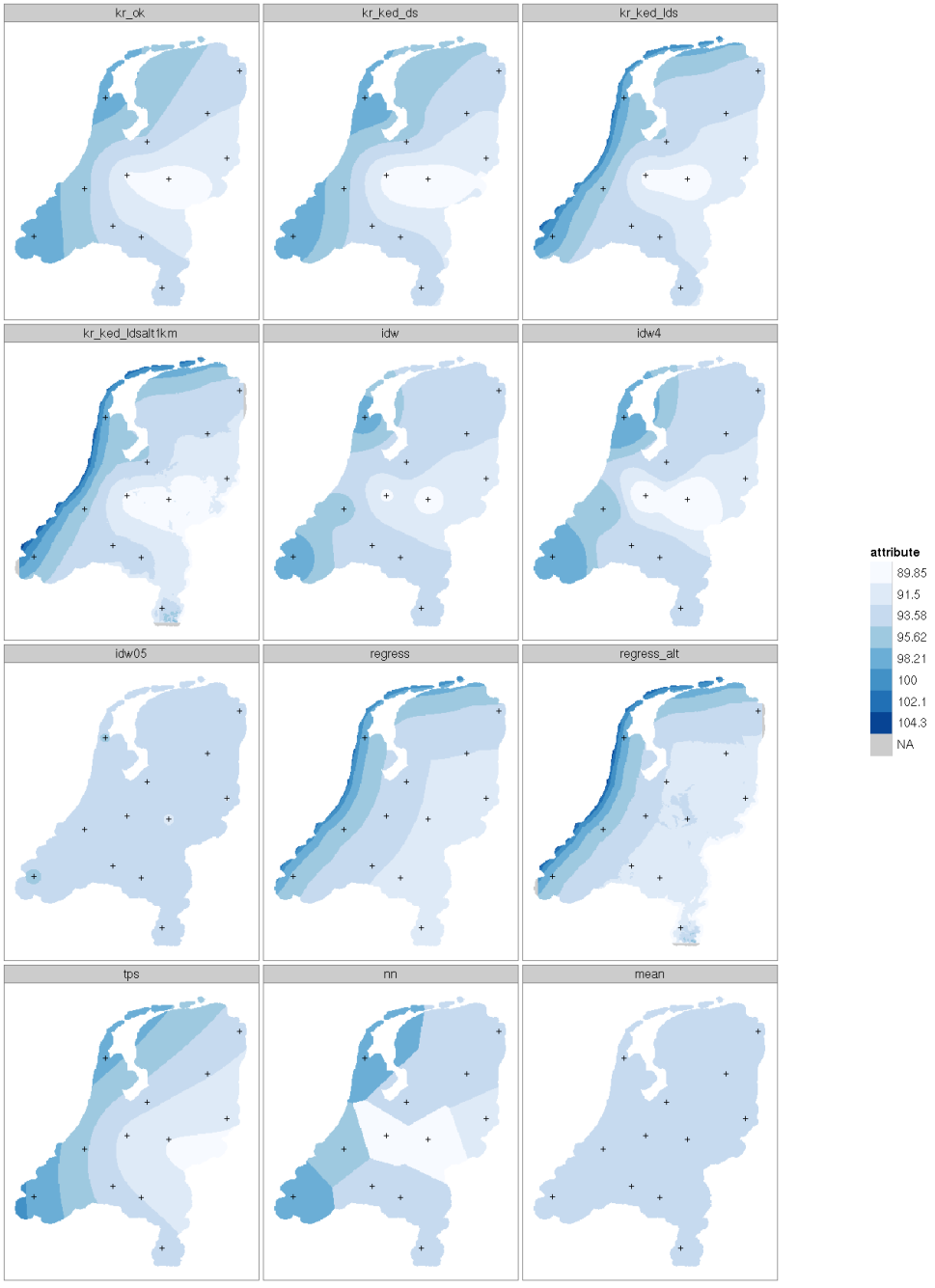


Figure A.8: Interpolated maps of 30 year monthly average for June

Figure A.9: Interpolated maps of 30 year monthly average for July

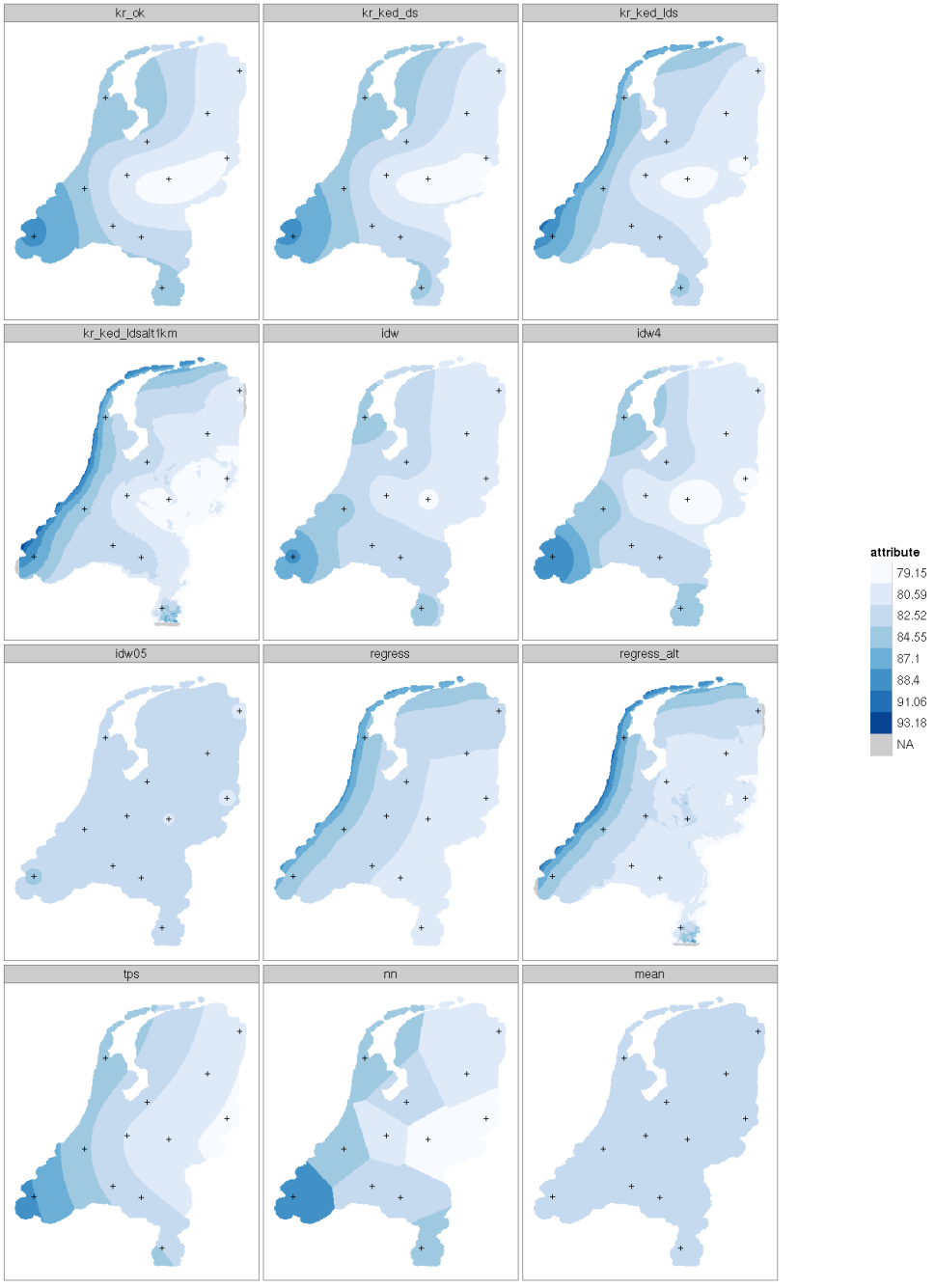


Figure A.10: Interpolated maps of 30 year monthly average for August

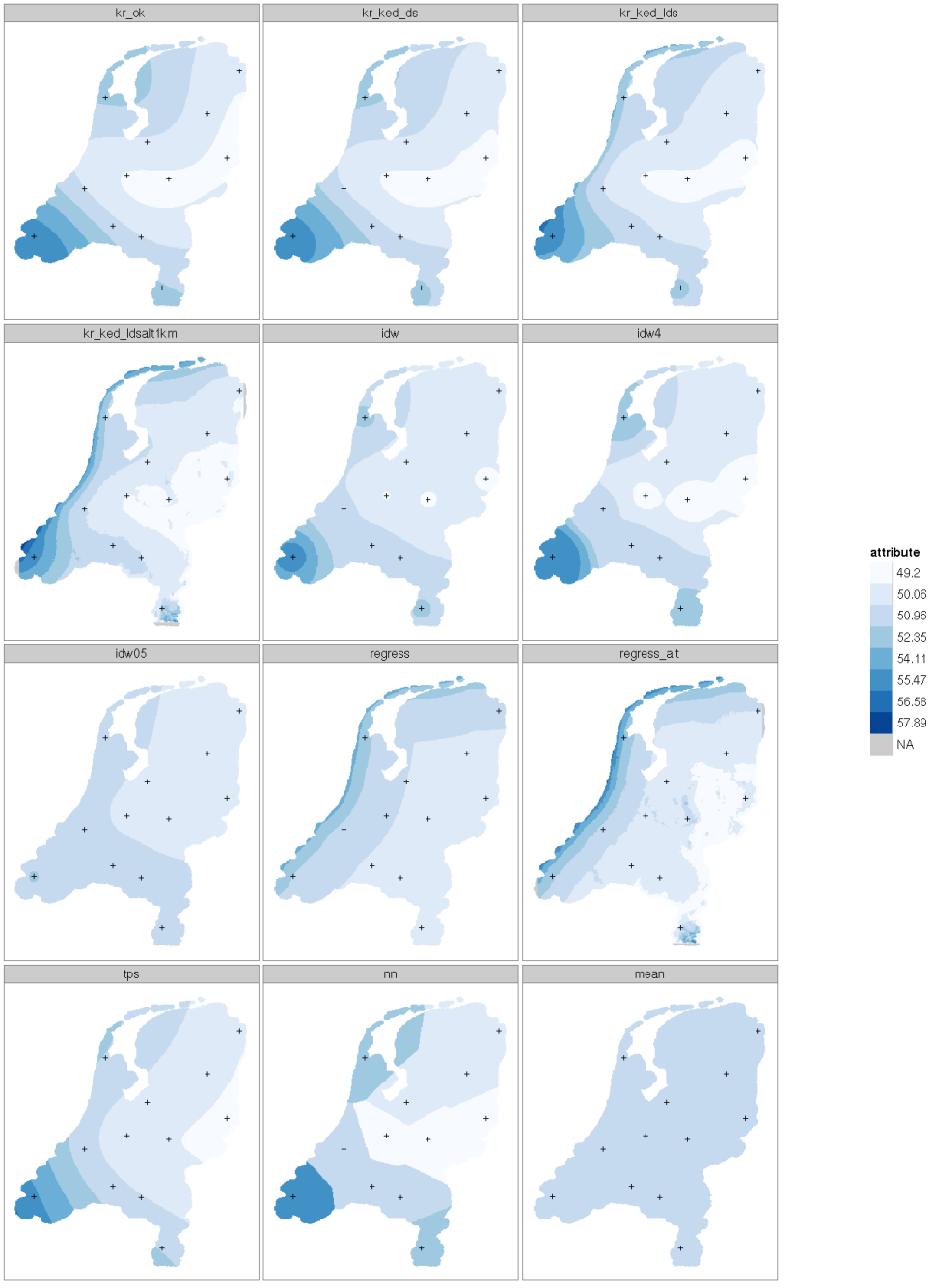


Figure A.11: Interpolated maps of 30 year monthly average for September

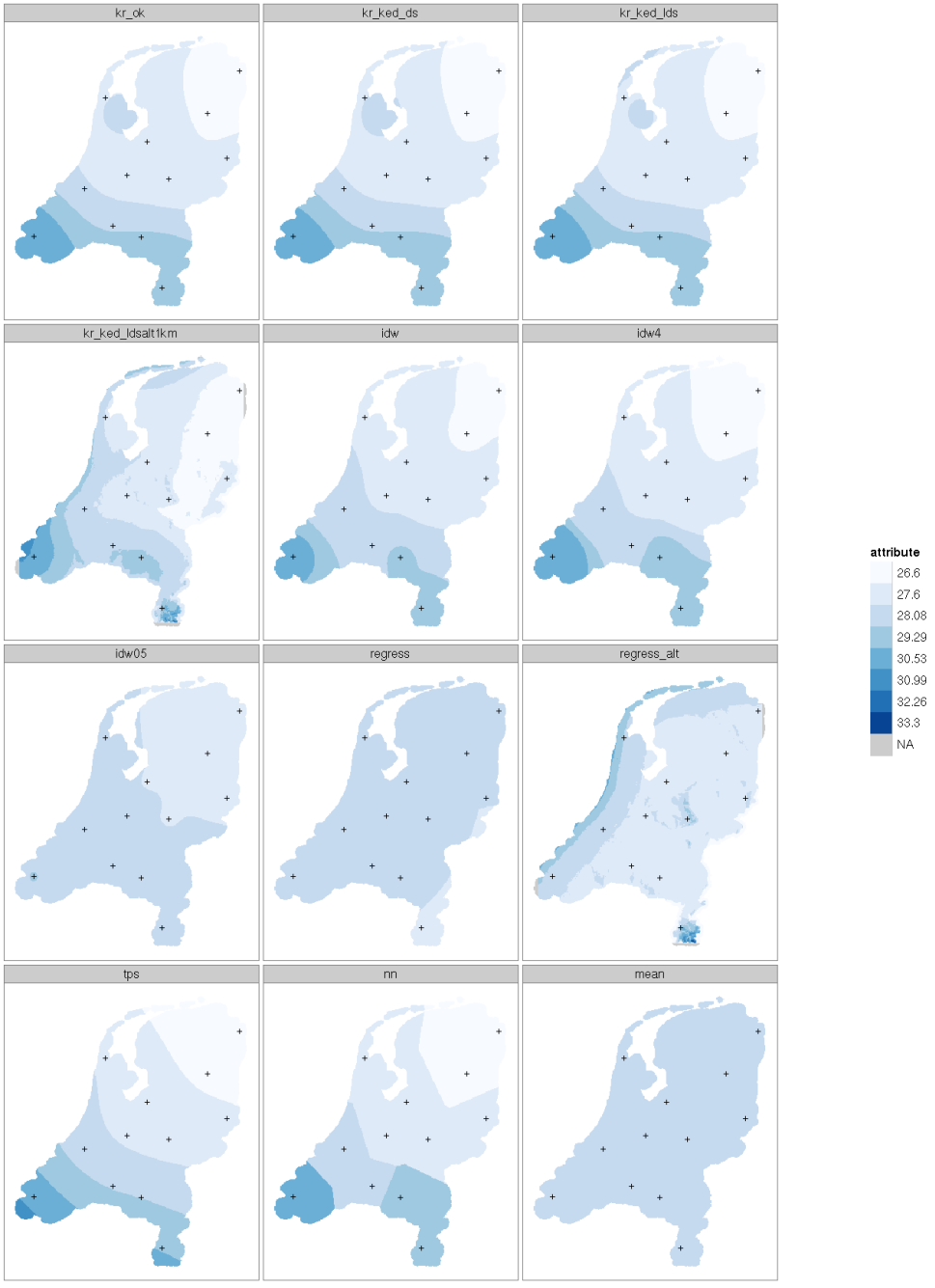


Figure A.12: Interpolated maps of 30 year monthly average for October

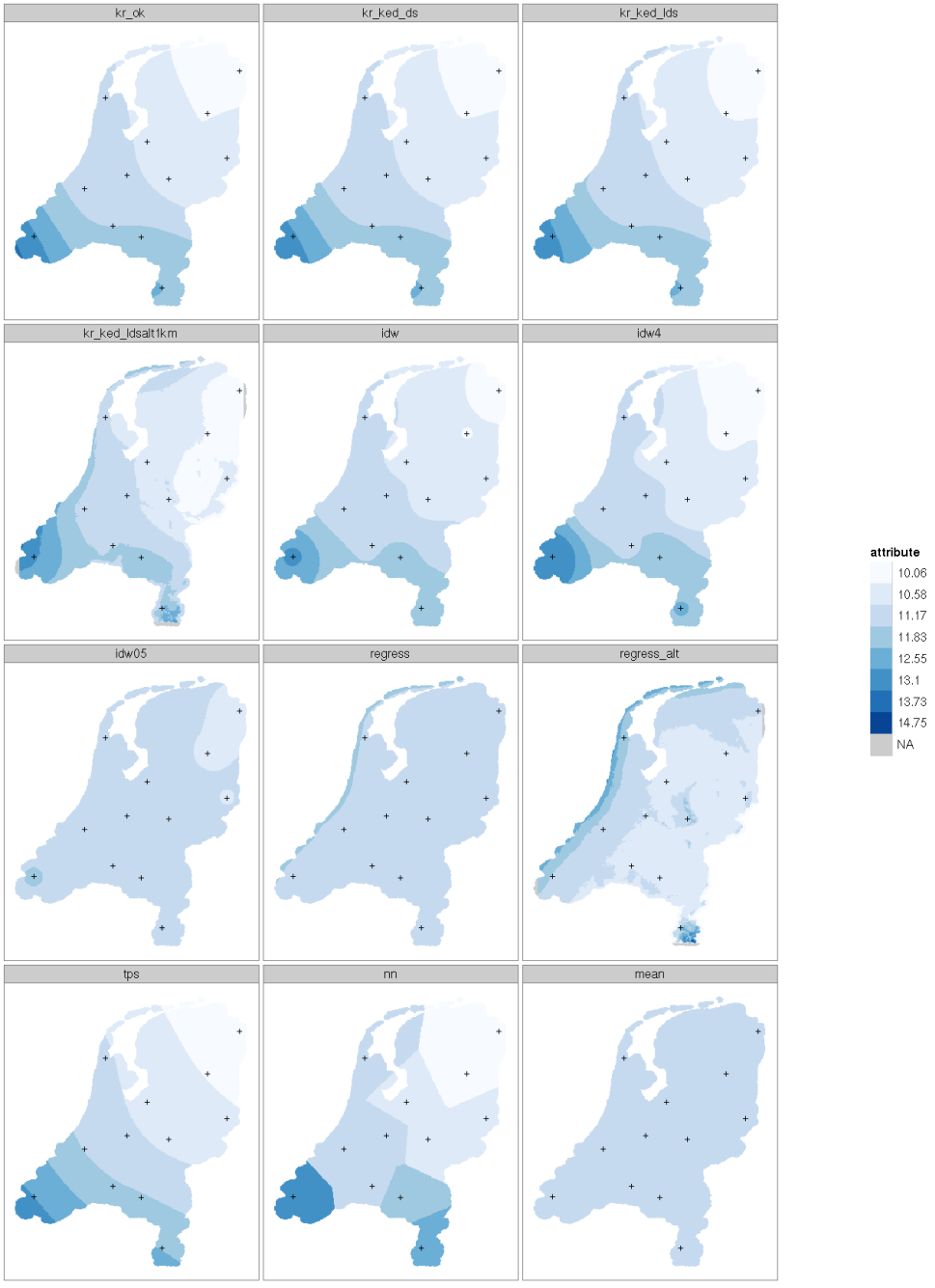


Figure A.13: Interpolated maps of 30 year monthly average for November

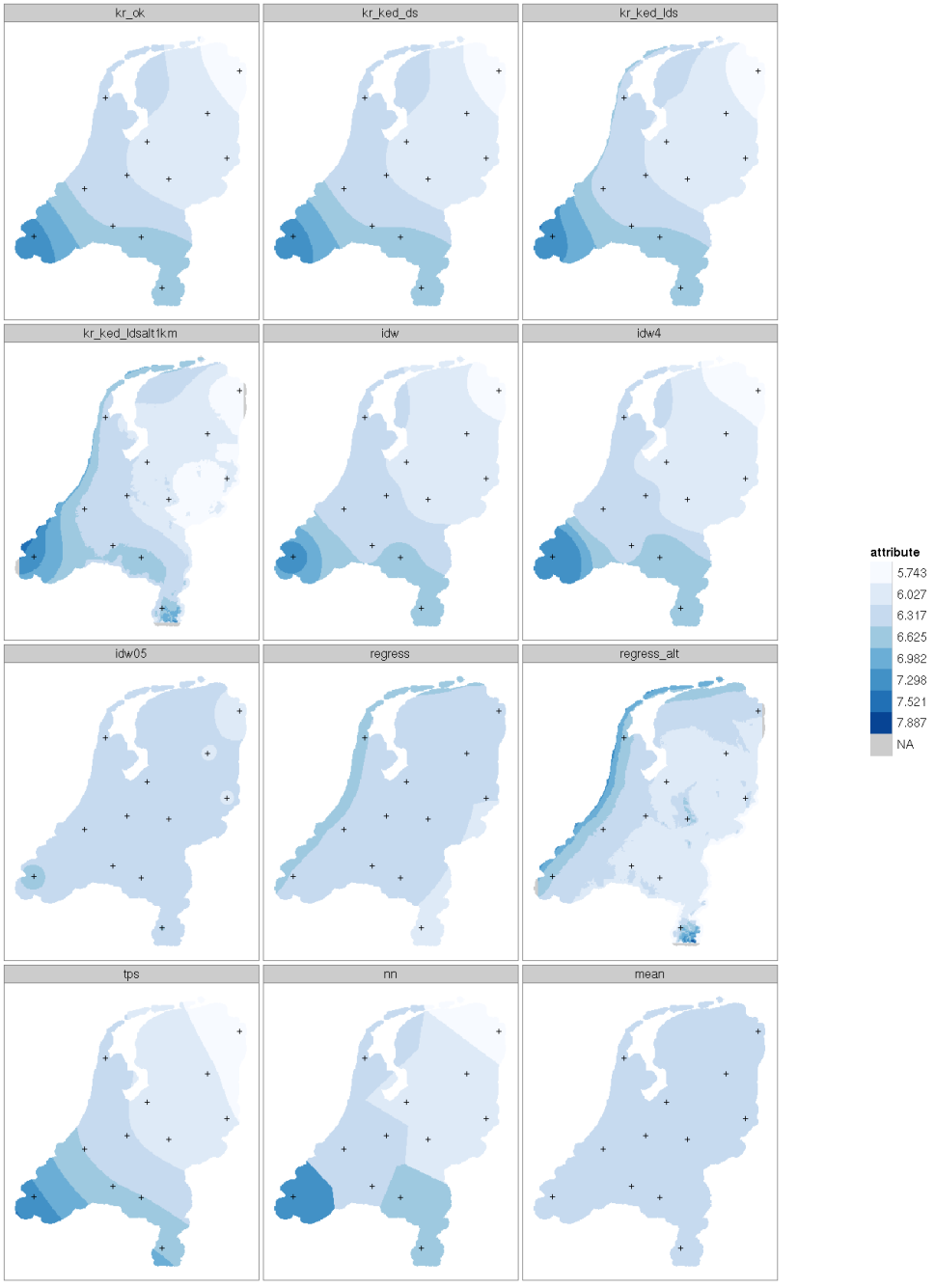


Figure A.14: Interpolated maps of 30 year monthly average for December

A.4 Cross-validation

A.4.1 Summary statistics

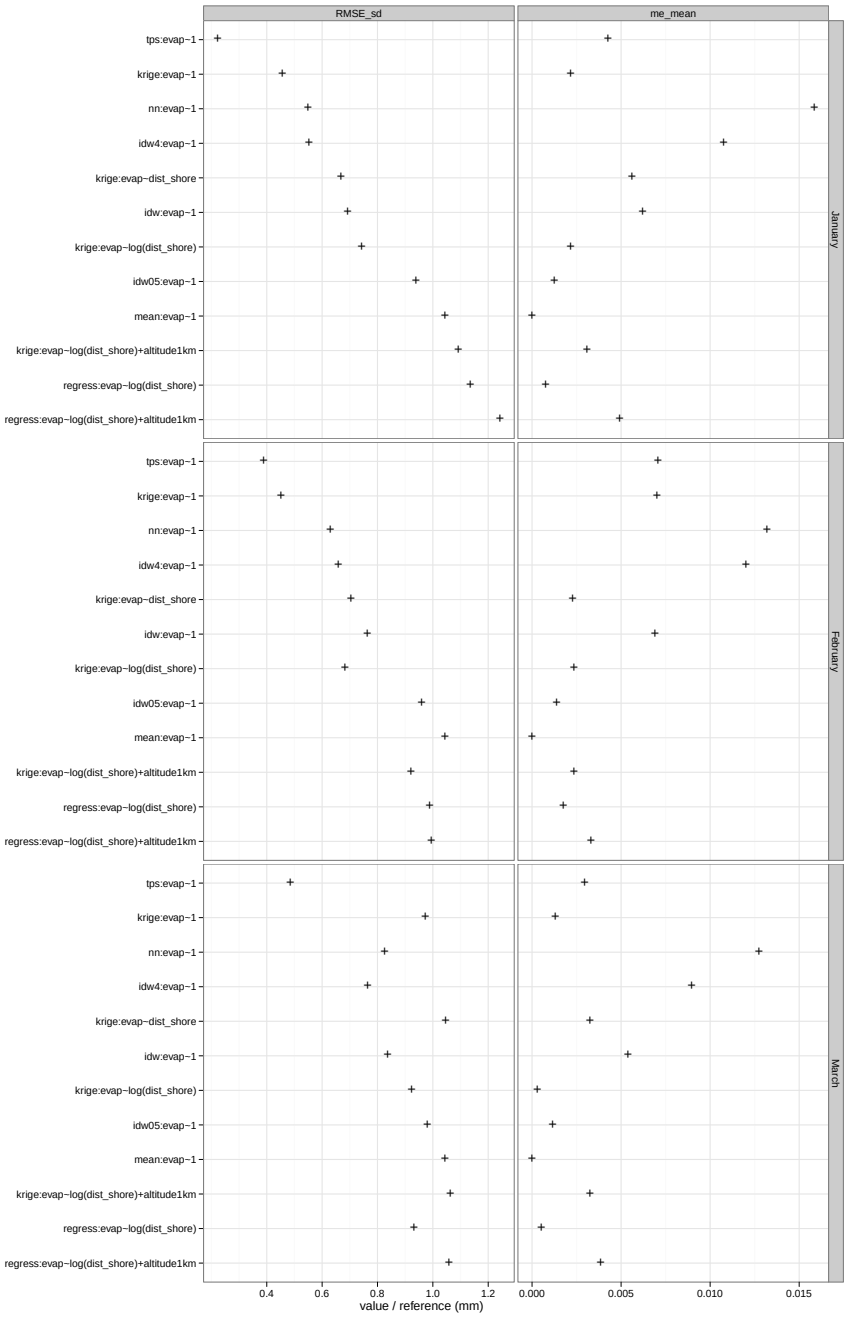


Figure A.15: Cross-validation summary statistics (RMSE divided by standard deviation and ME divided by mean) for monthly interpolated maps of January to March.

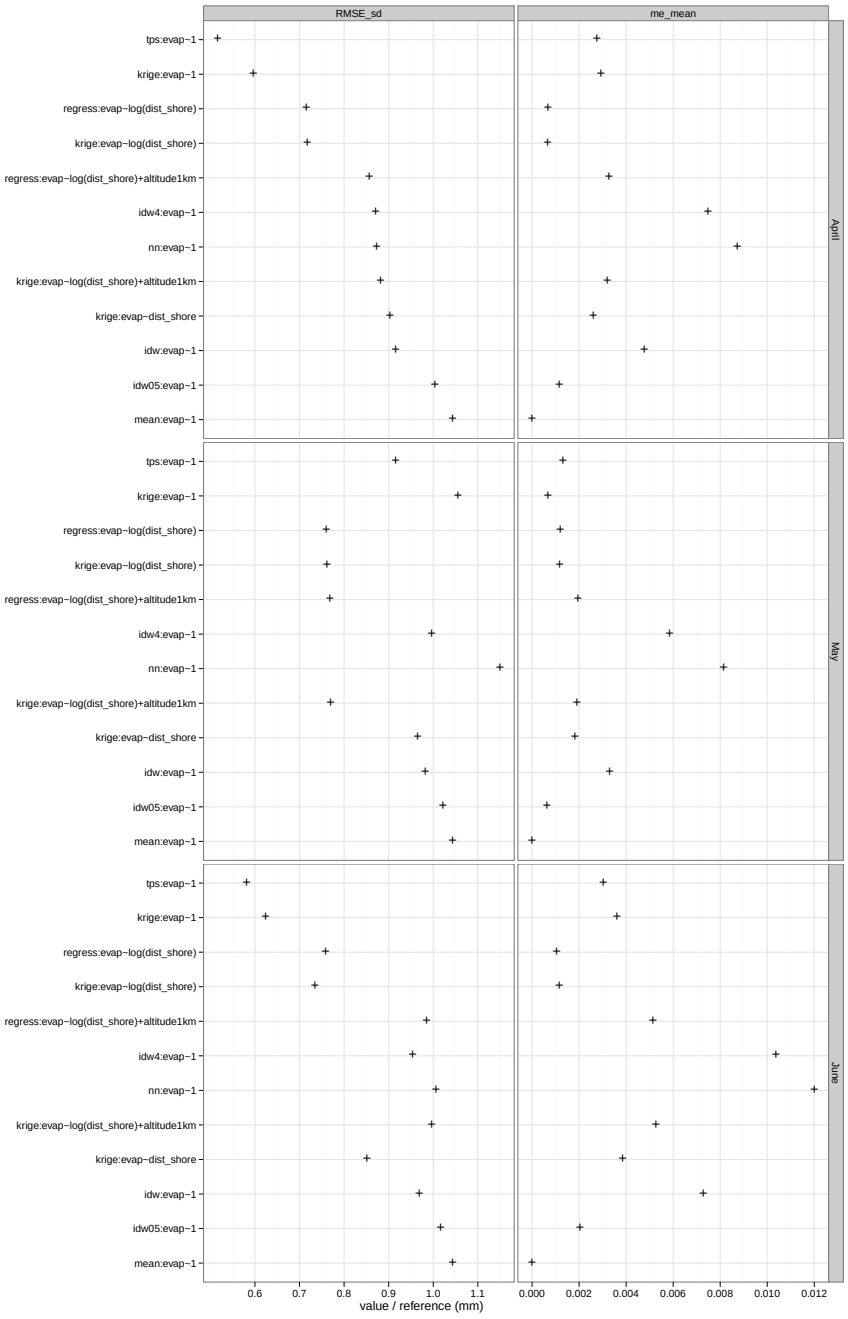


Figure A.16: Cross-validation summary statistics (RMSE divided by standard deviation and ME divided by mean) for monthly interpolated maps of April to June.

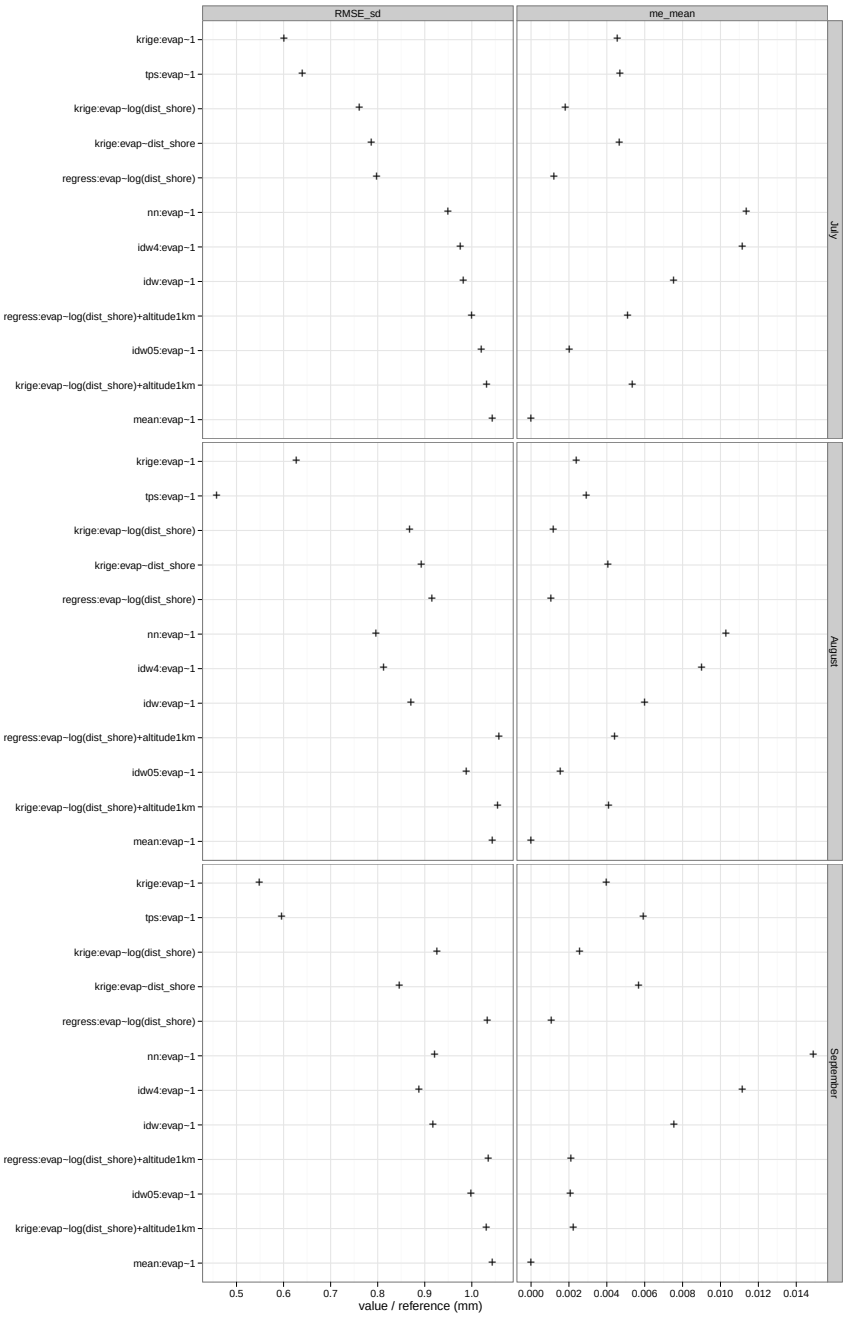


Figure A.17: Cross-validation summary statistics (RMSE divided by standard deviation and ME divided by mean) for monthly interpolated maps of July to September.

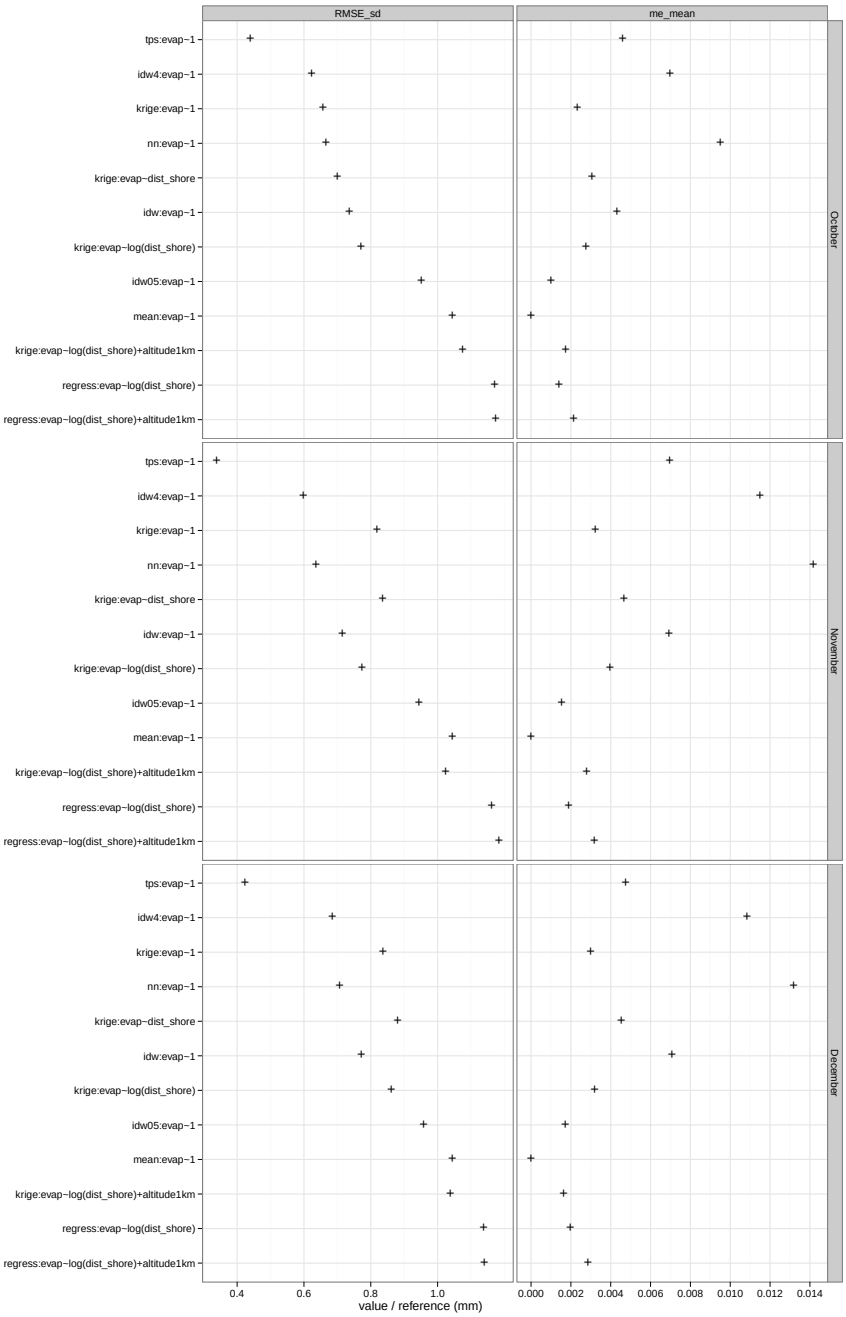


Figure A.18: Cross-validation summary statistics (RMSE divided by standard deviation and ME divided by mean) for monthly interpolated maps of October to December.

A.4.2 Plots of CV residuals

Figure A.19: Spatial pattern in cross-validation residuals for January

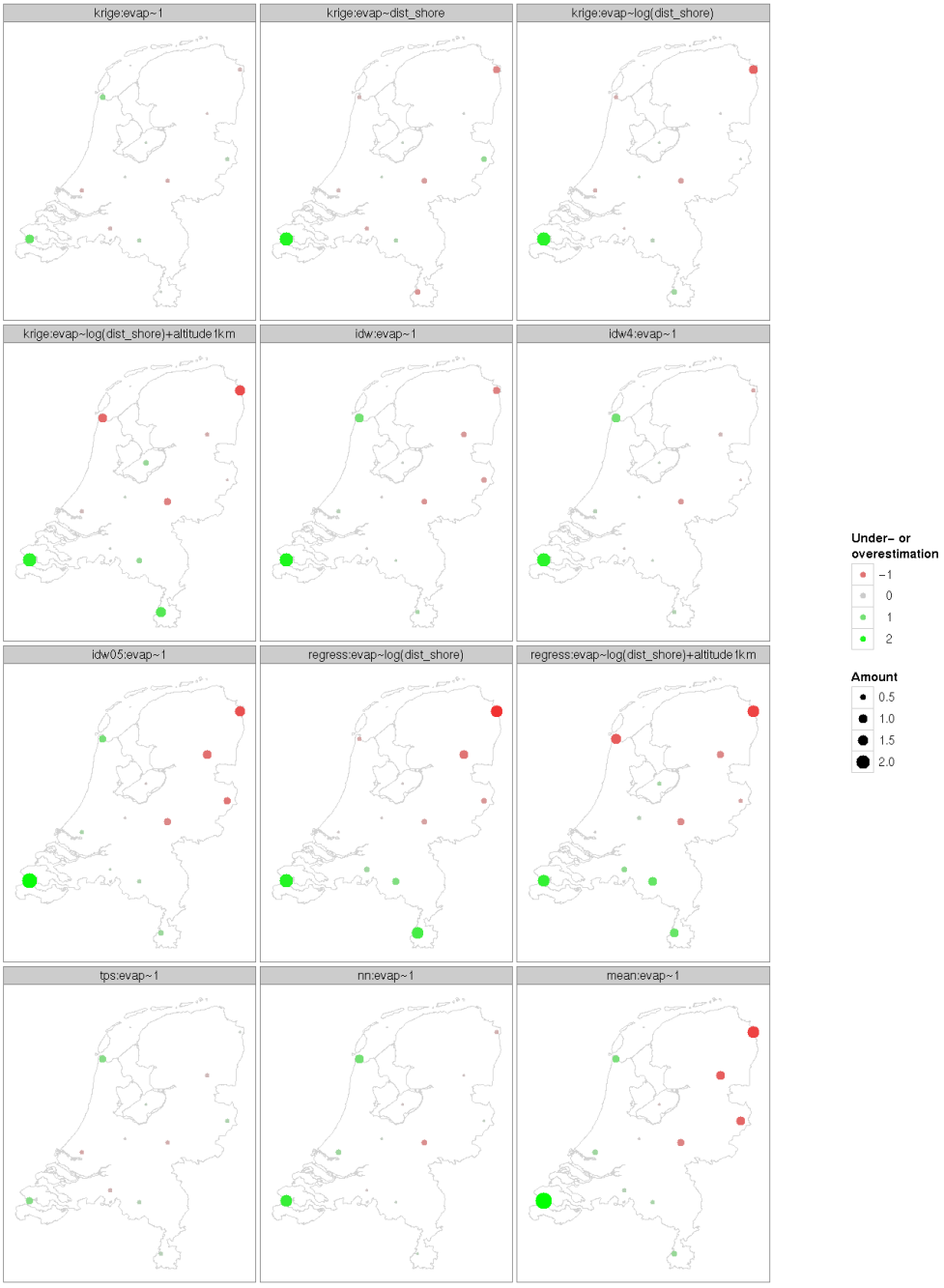


Figure A.20: Spatial pattern in cross-validation residuals for February

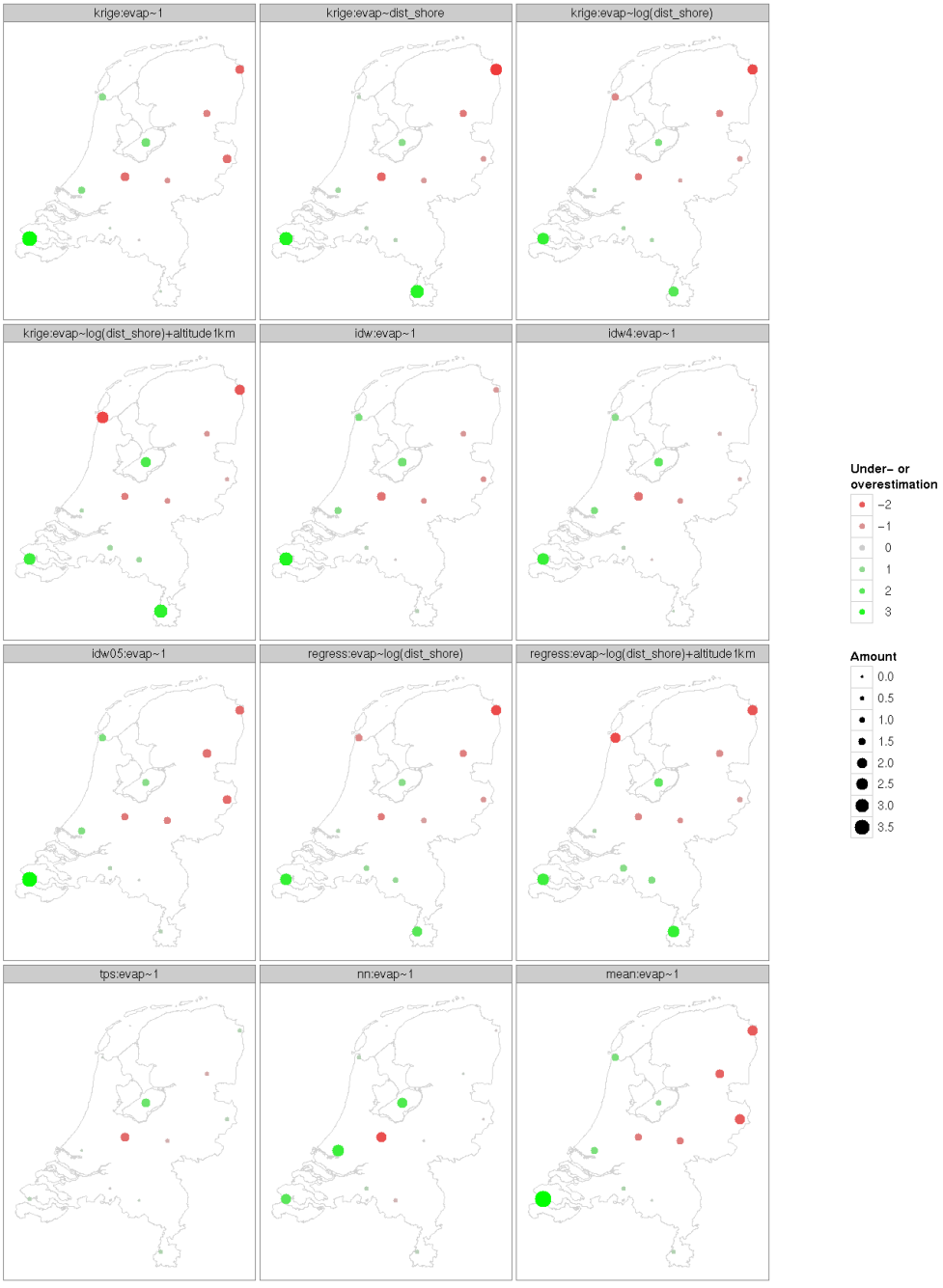


Figure A.21: Spatial pattern in cross-validation residuals for March

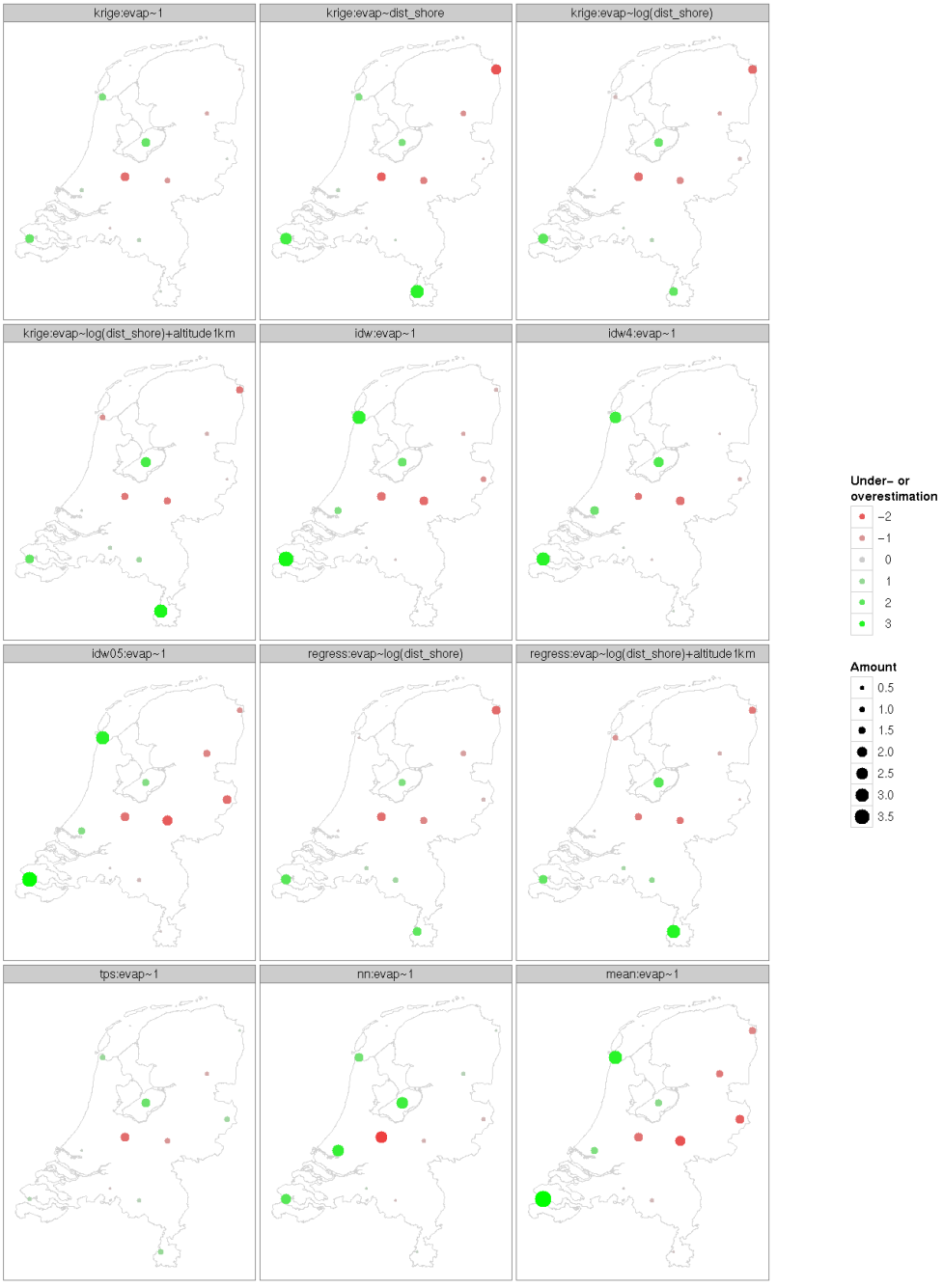


Figure A.22: Spatial pattern in cross-validation residuals for April

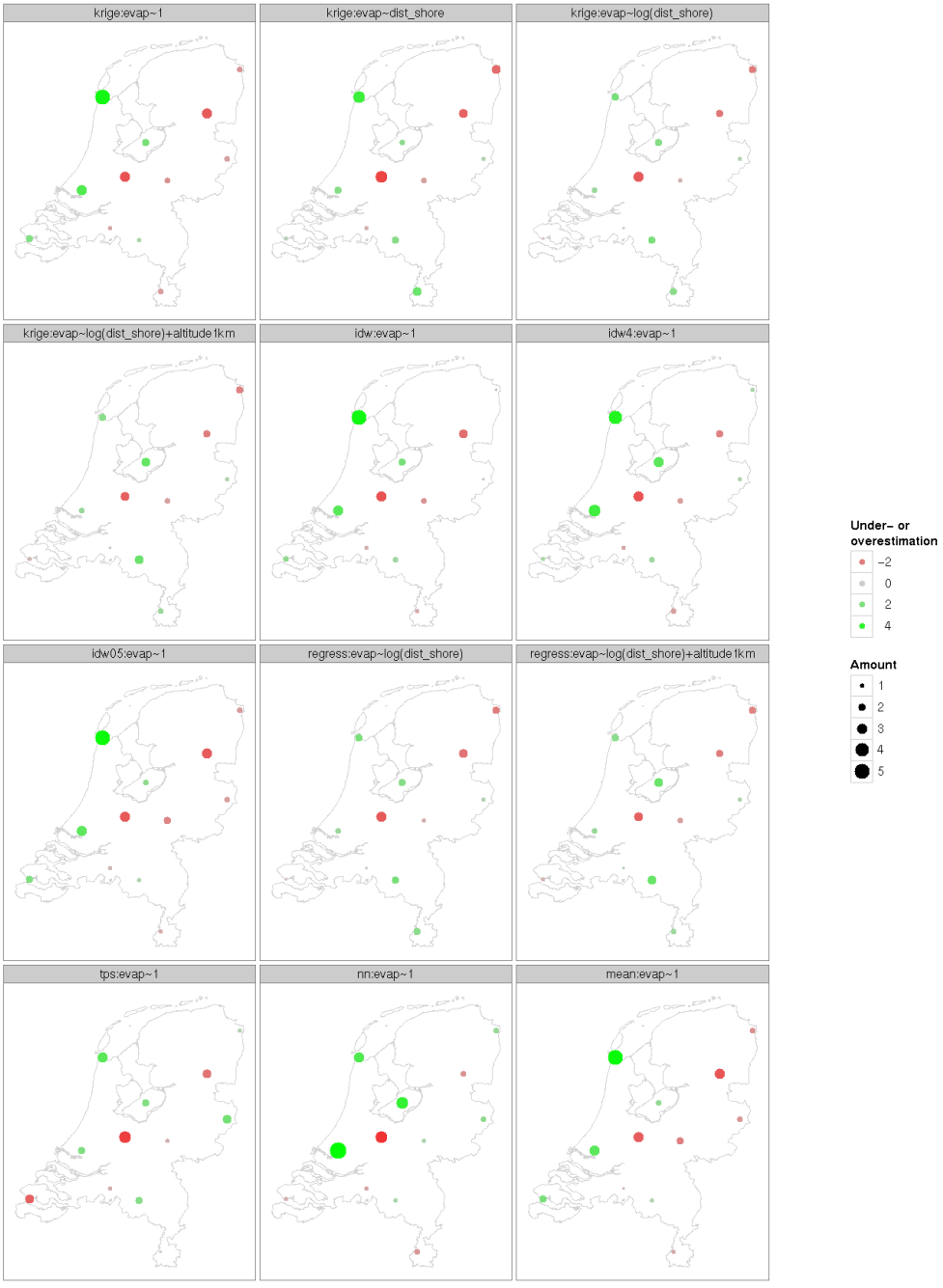


Figure A.23: Spatial pattern in cross-validation residuals for May

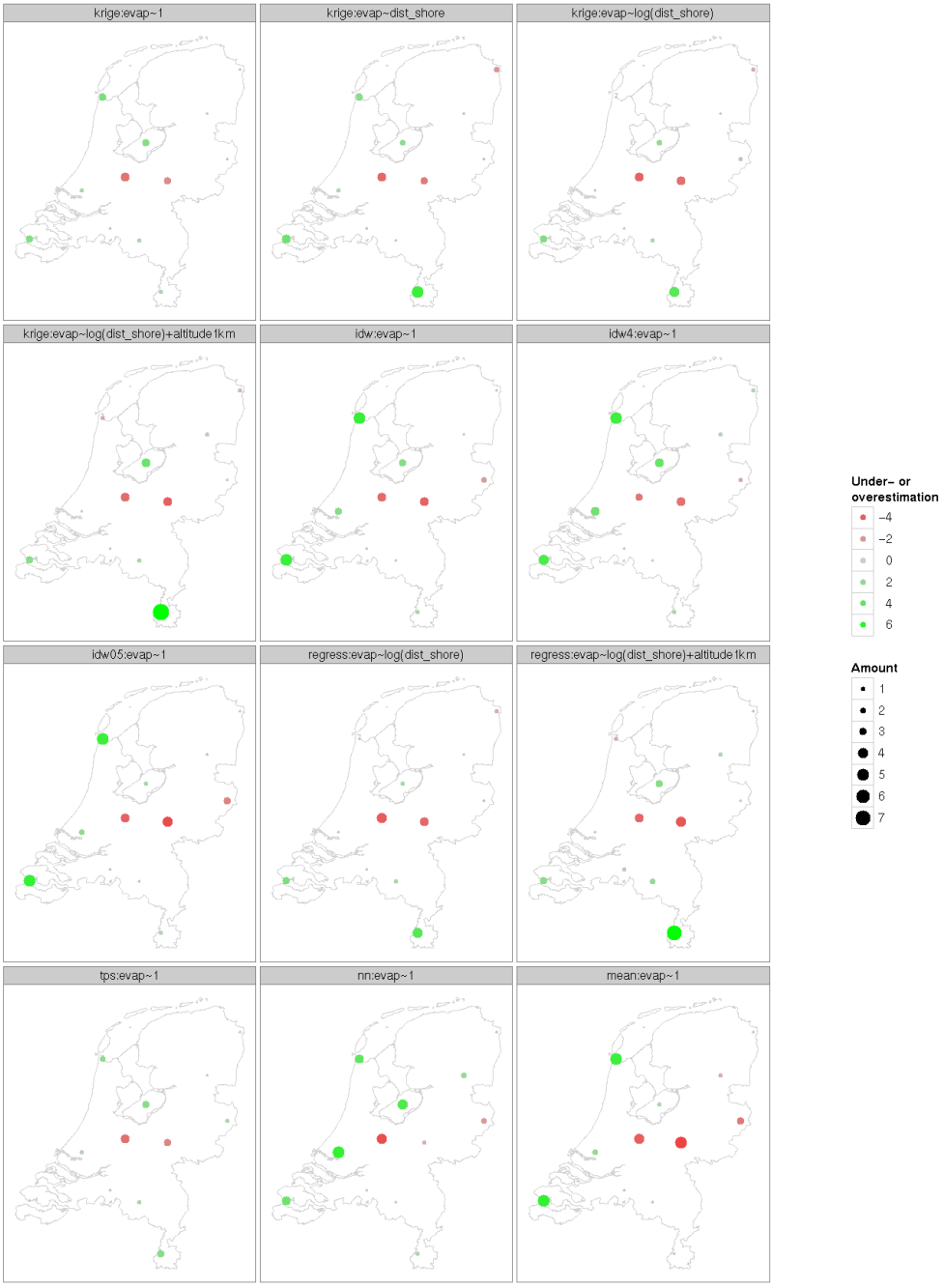


Figure A.24: Spatial pattern in cross-validation residuals for June

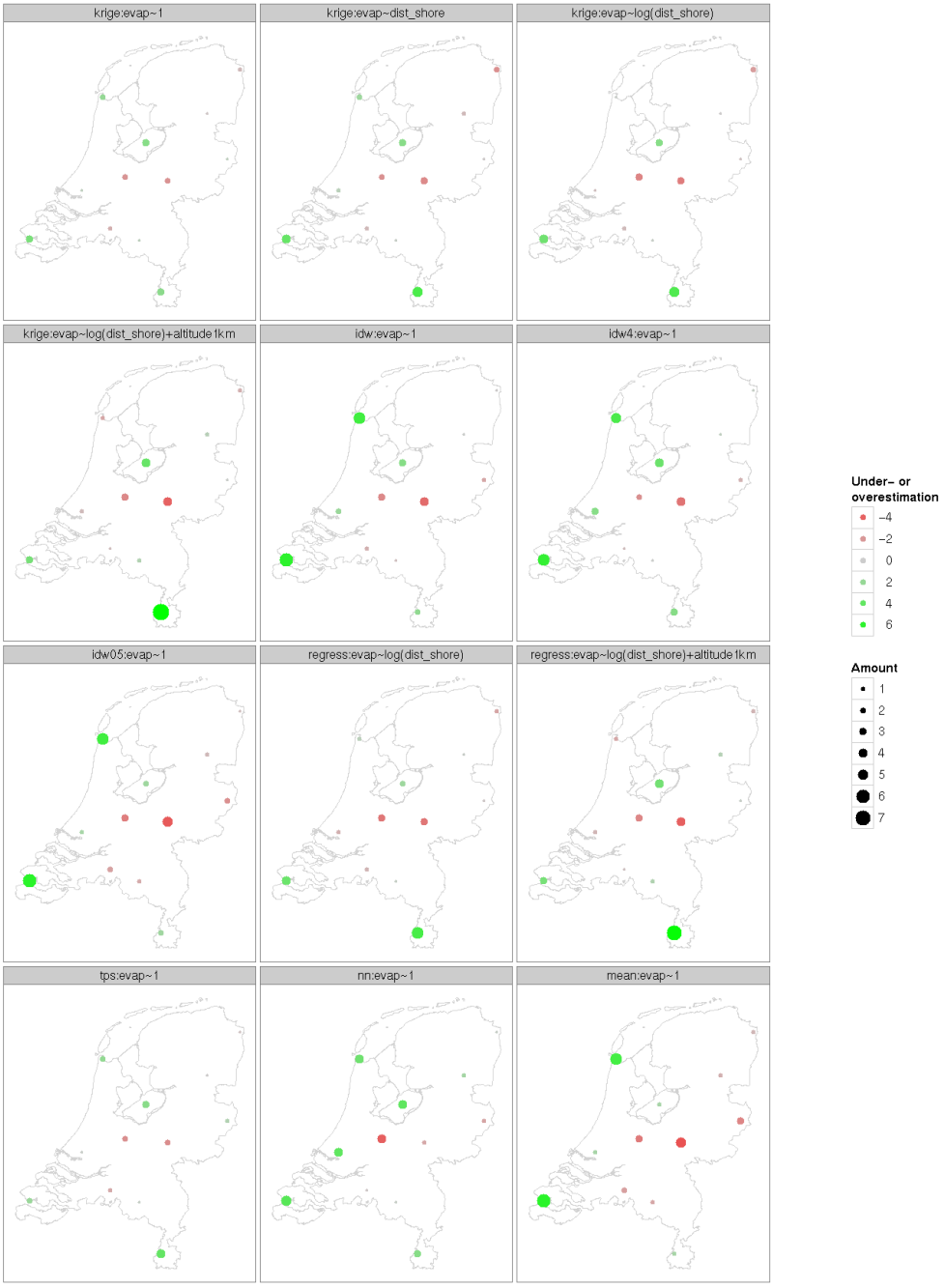


Figure A.25: Spatial pattern in cross-validation residuals for July

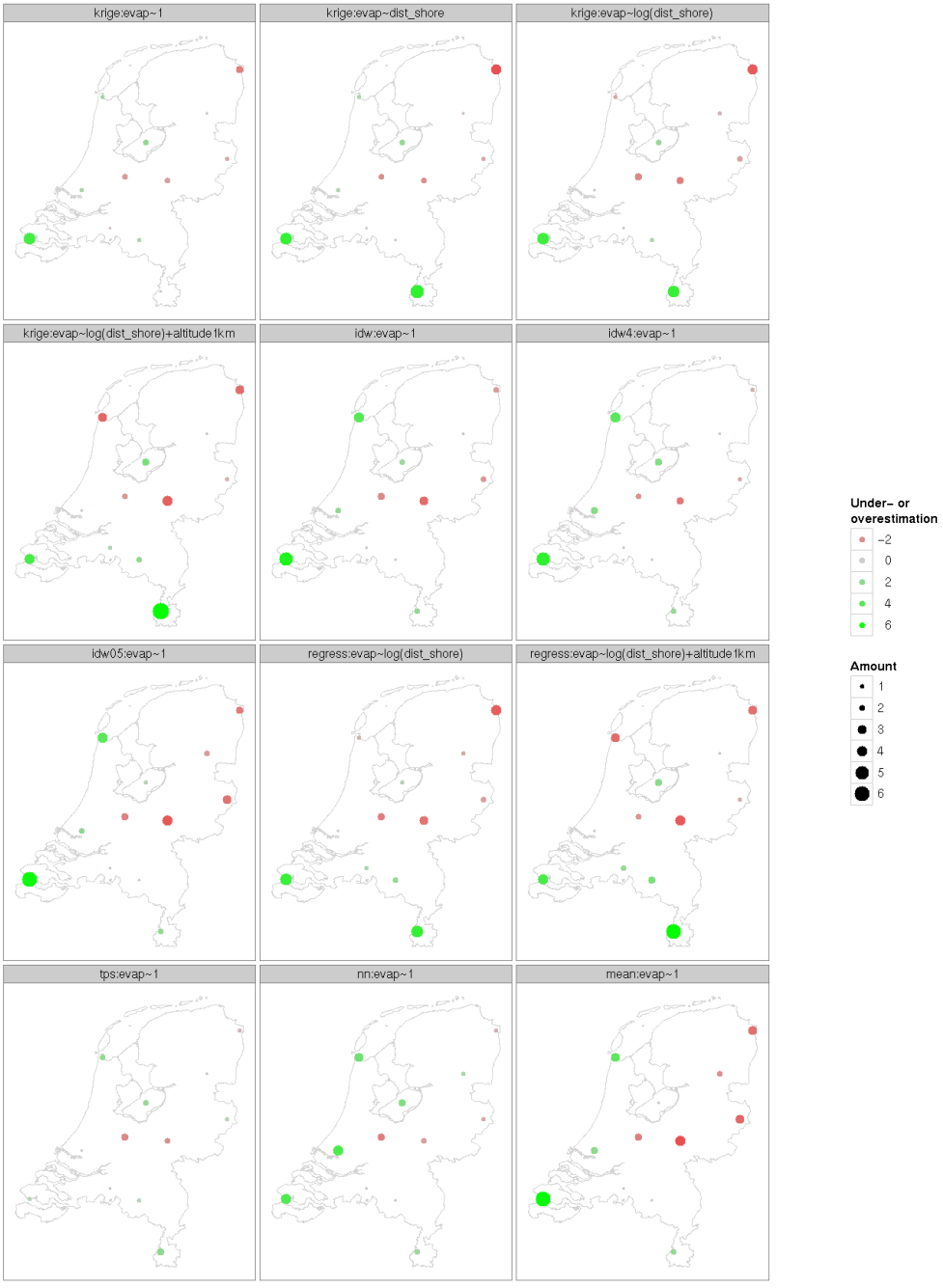


Figure A.26: Spatial pattern in cross-validation residuals for August

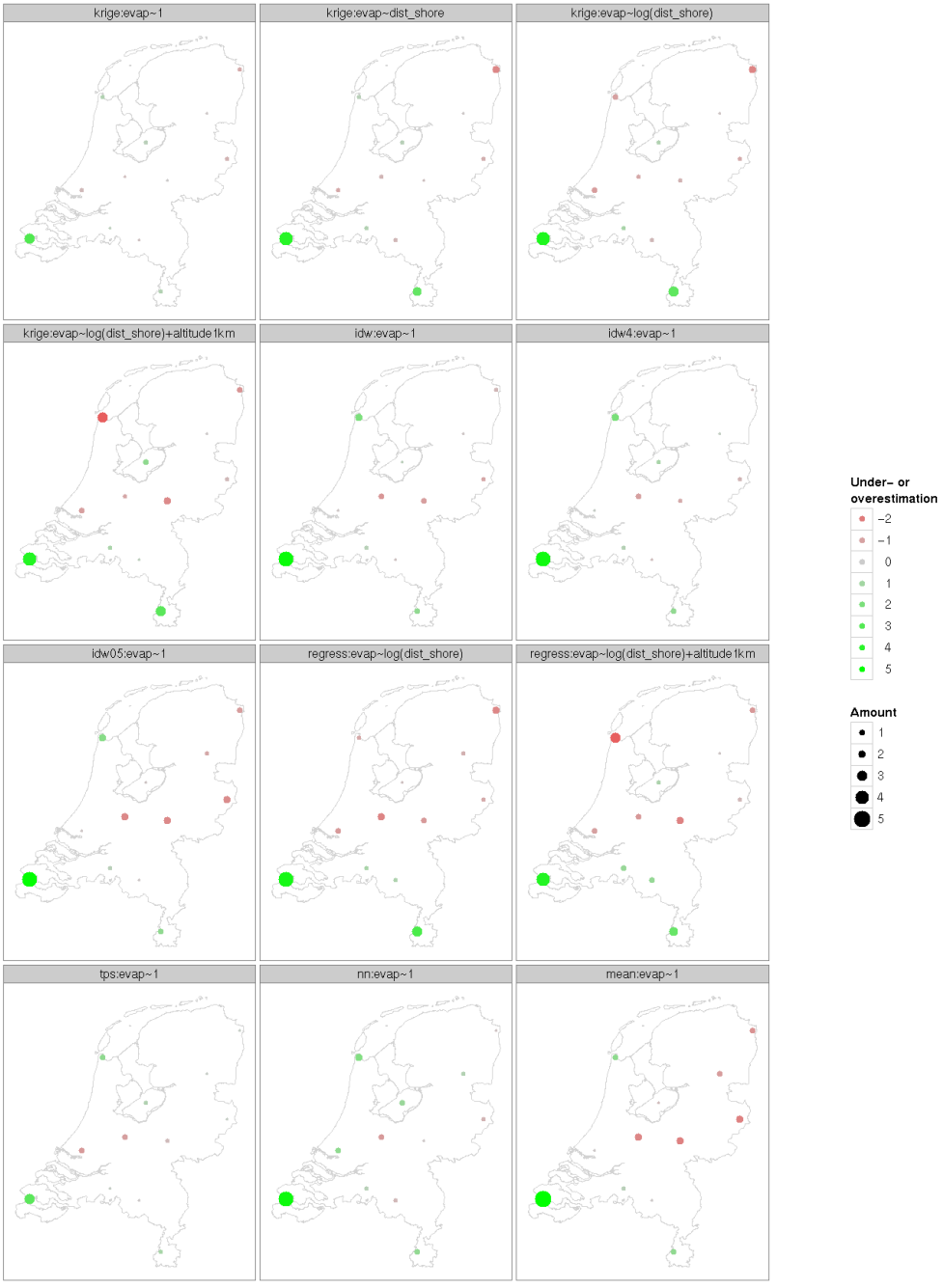


Figure A.27: Spatial pattern in cross-validation residuals for September

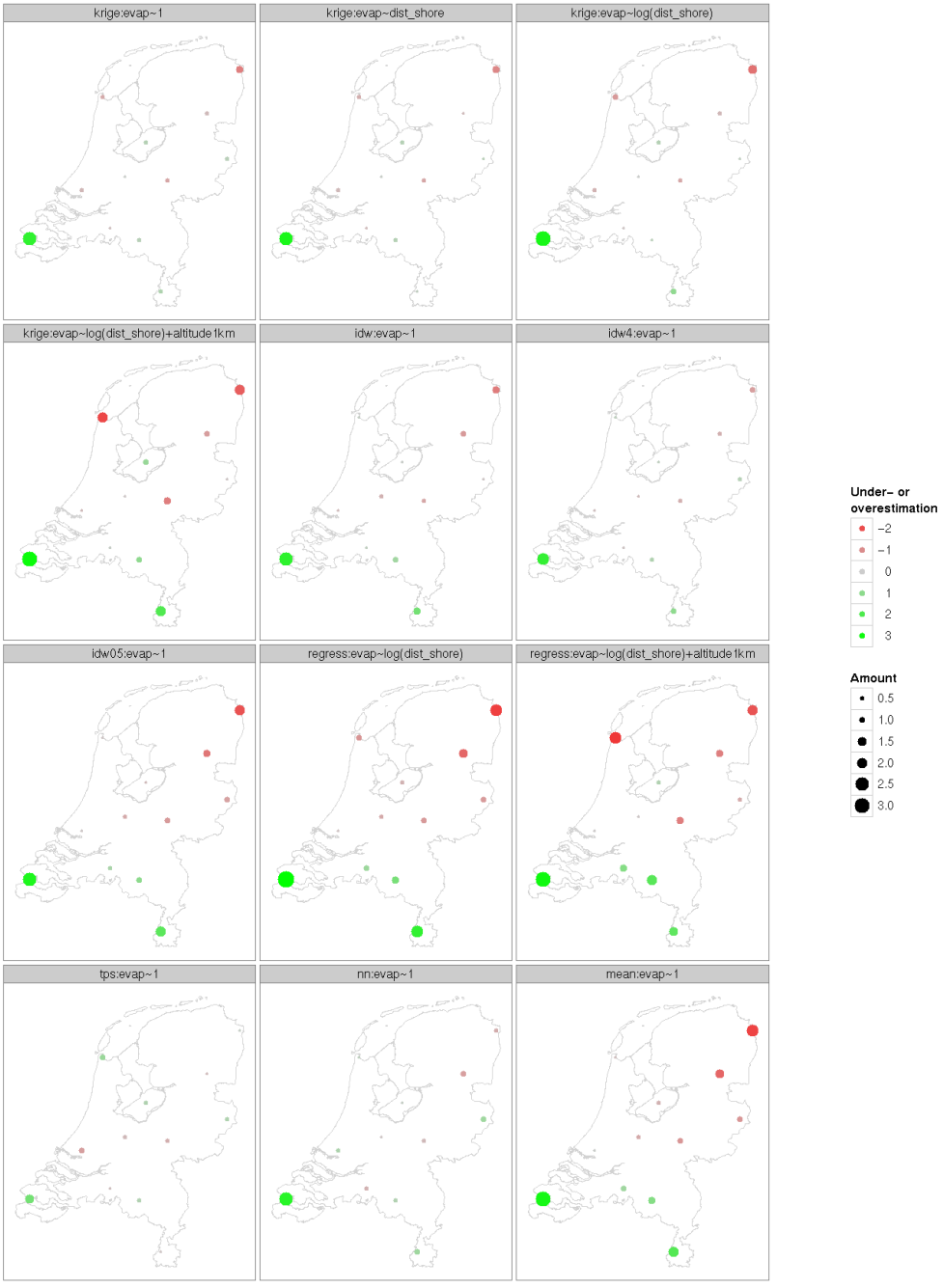


Figure A.28: Spatial pattern in cross-validation residuals for October

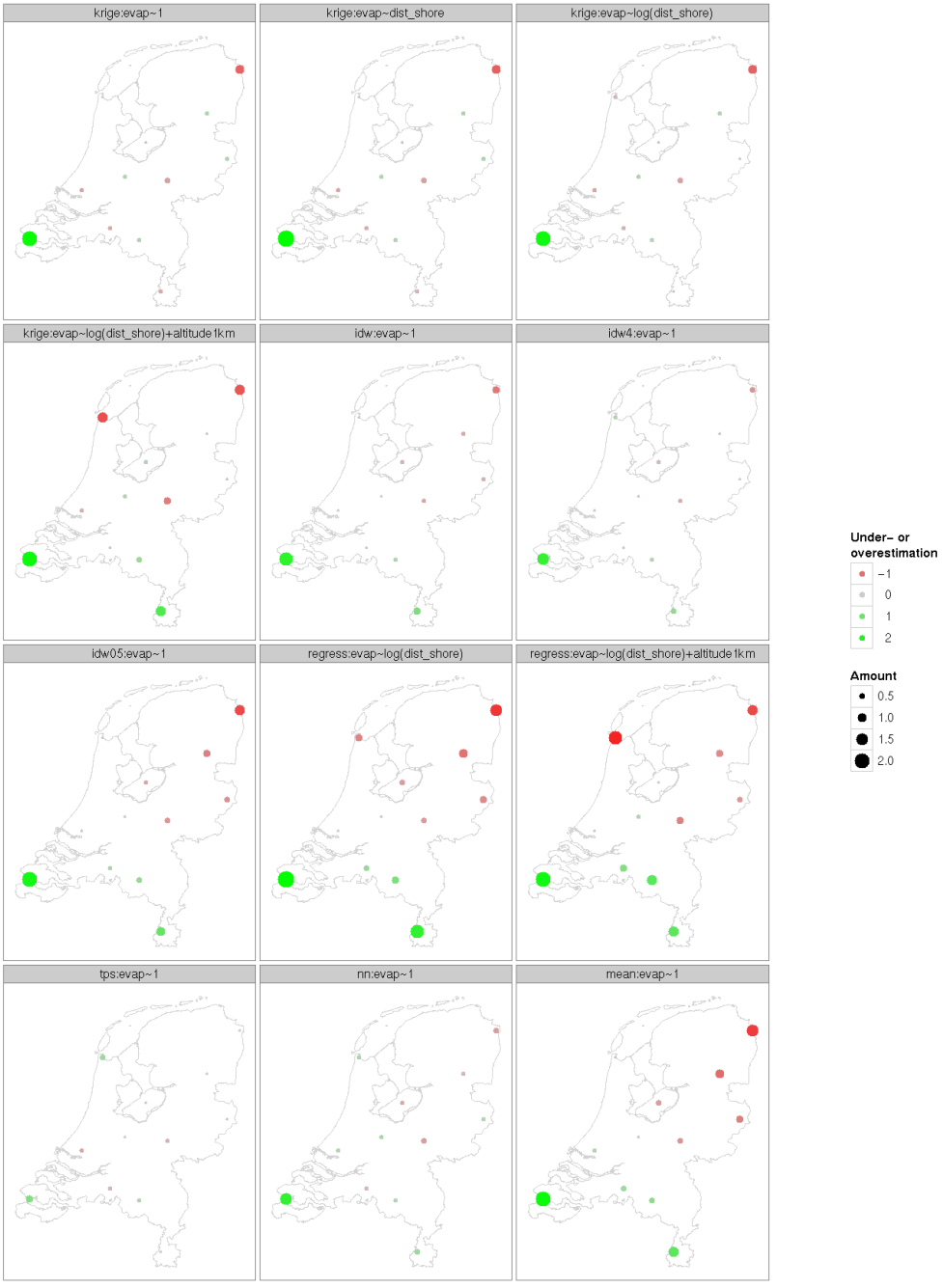


Figure A.29: Spatial pattern in cross-validation residuals for November

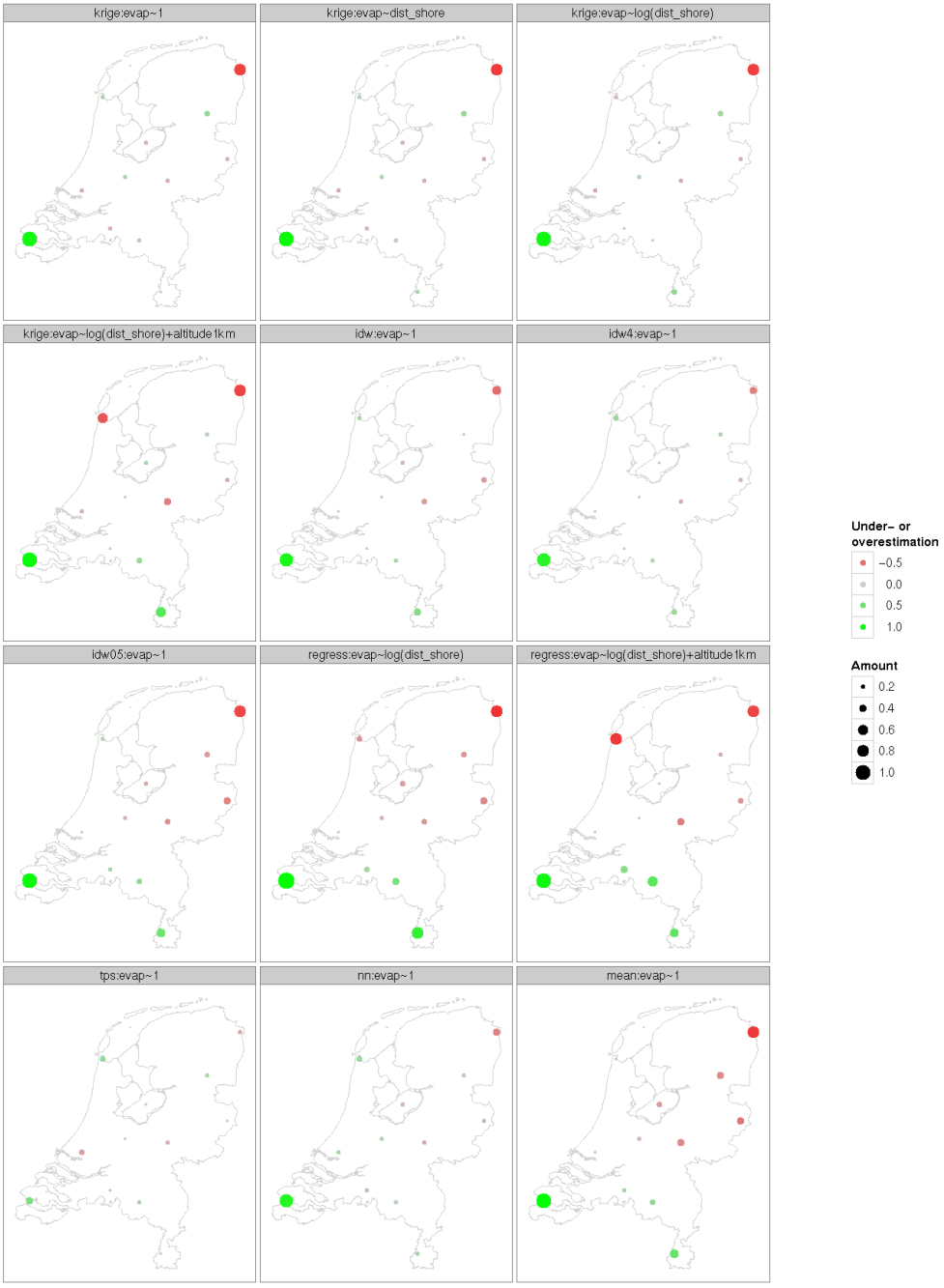


Figure A.30: Spatial pattern in cross-validation residuals for December

Appendix B

Interpolation of daily observations

B.1 Daily interpolated maps

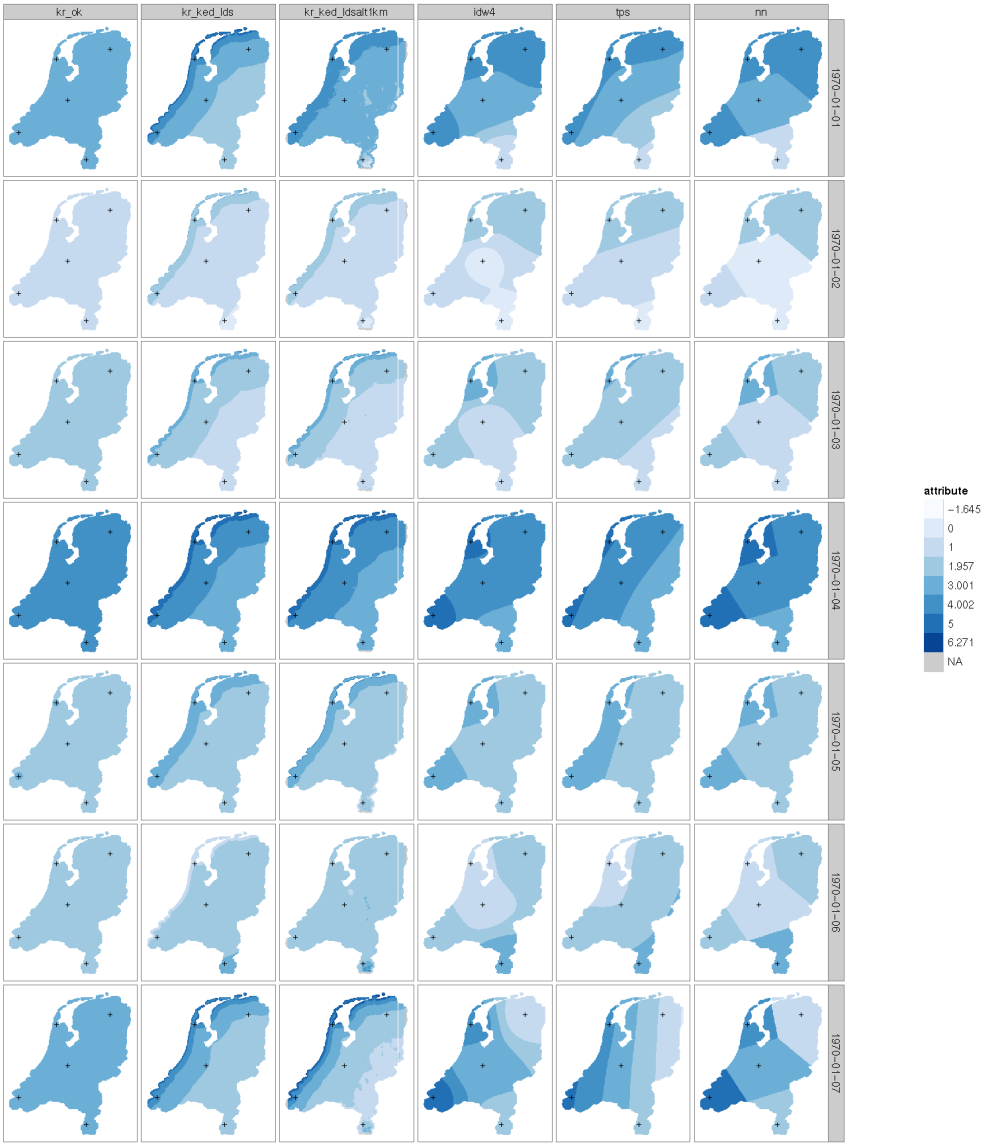


Figure B.1: Interpolated maps of daily Makkink evaporation (10^{-1}mm) for the first week of January 1970. Columns represent different interpolation methods, rows represent time.

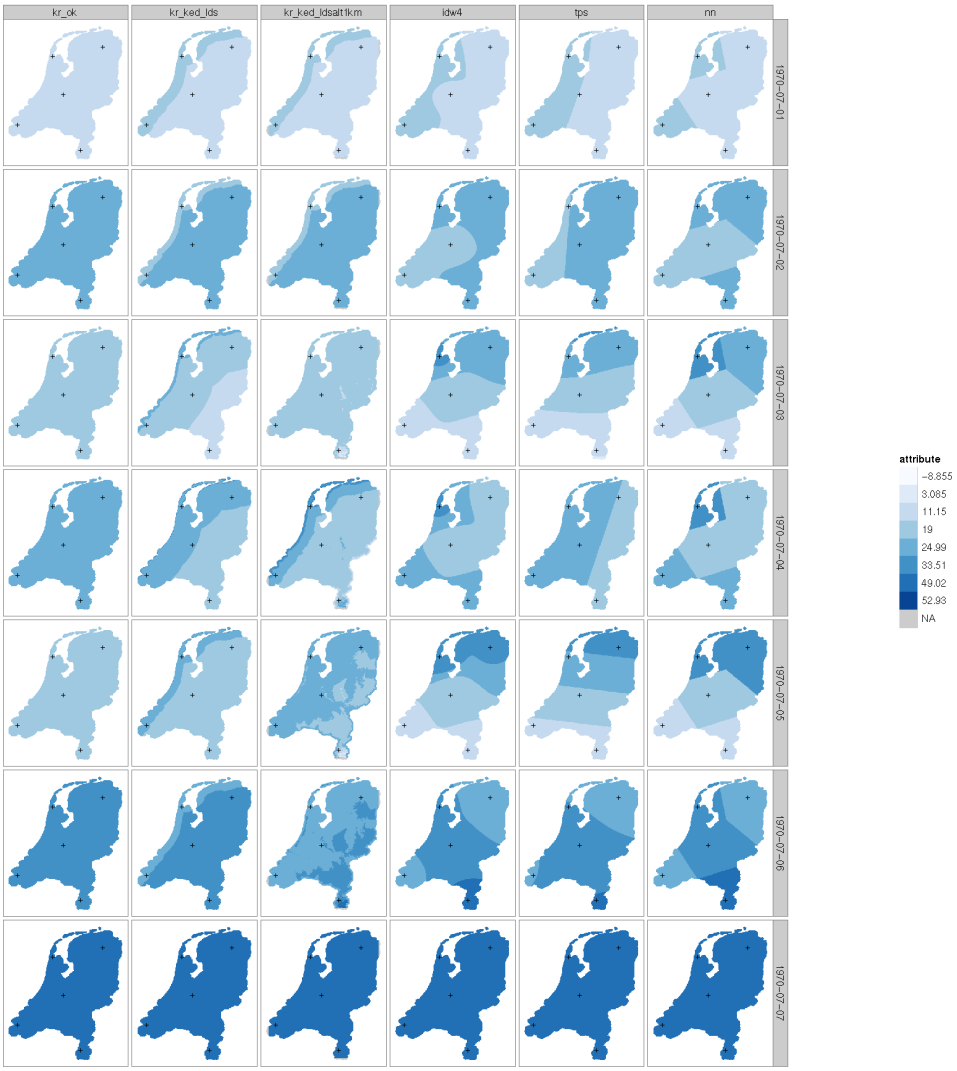


Figure B.2: Interpolated maps of daily Makkink evaporation (10^{-1}mm) for the first week of July 1970. Columns represent different interpolation methods, rows represent time.

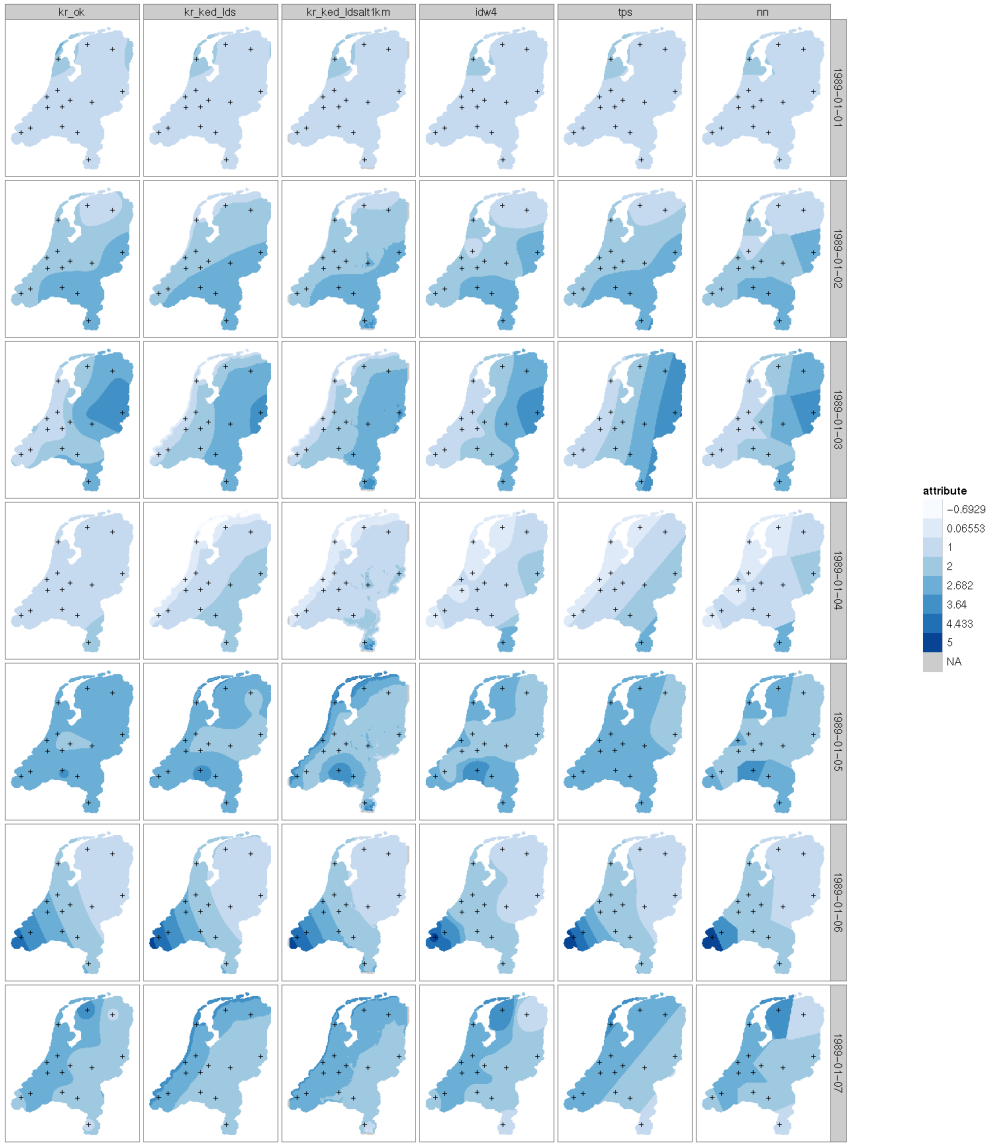


Figure B.3: Interpolated maps of daily Makkink evaporation (10^{-1}mm) for the first week of January 1989. Columns represent different interpolation methods, rows represent time.

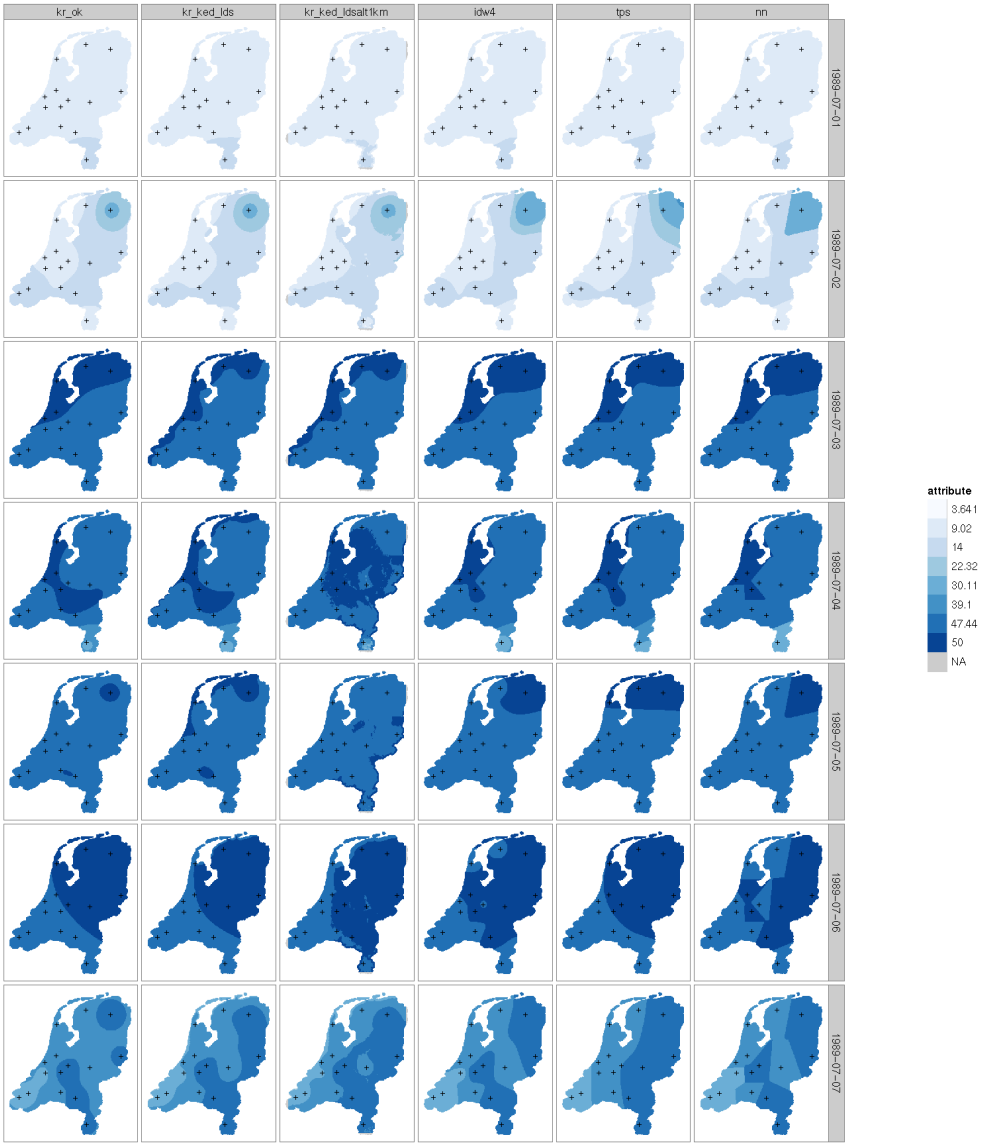


Figure B.4: Interpolated maps of daily Makkink evaporation (10^{-1}mm) for the first week of July 1989. Columns represent different interpolation methods, rows represent time.

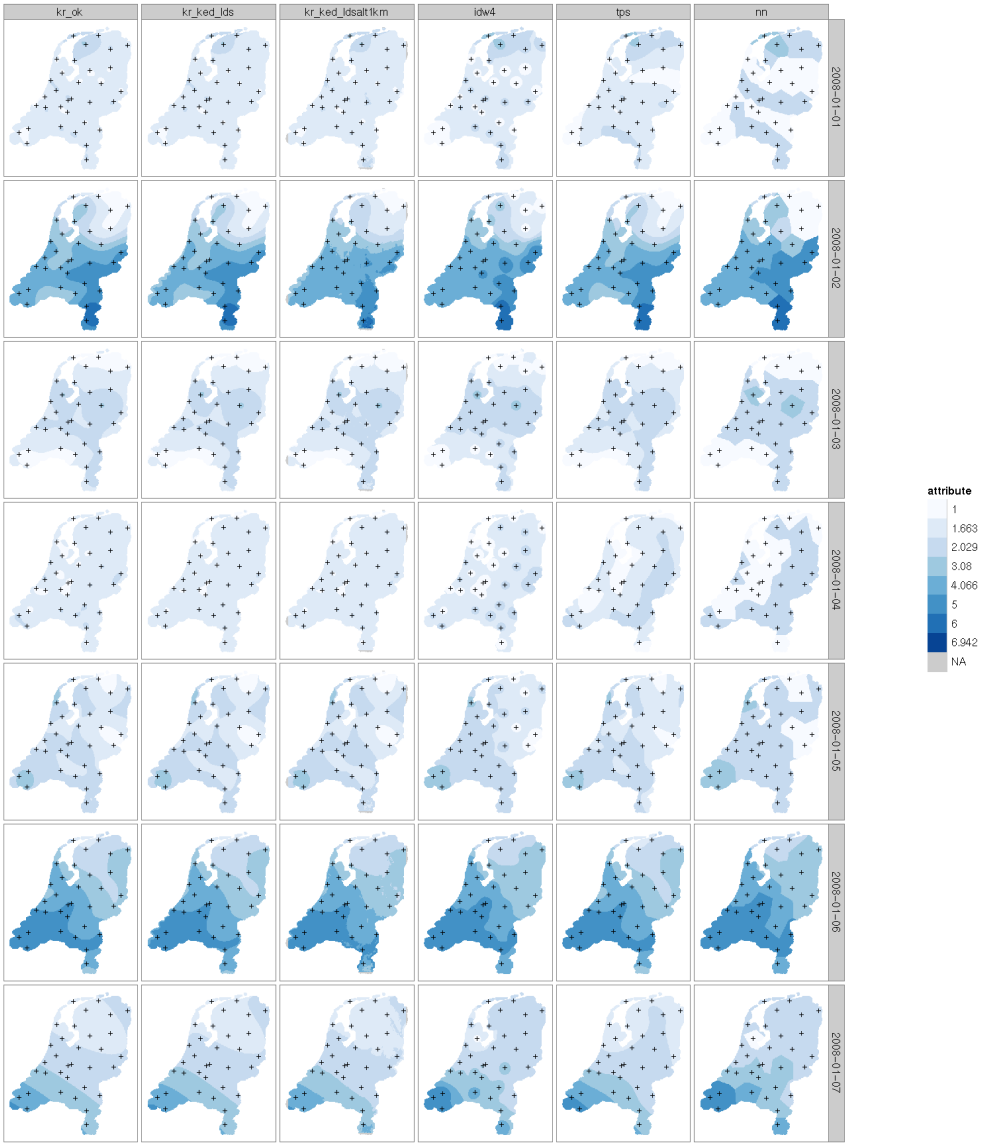


Figure B.5: Interpolated maps of daily Makkink evaporation (10^{-1}mm) for the first week of January 2008. Columns represent different interpolation methods, rows represent time.

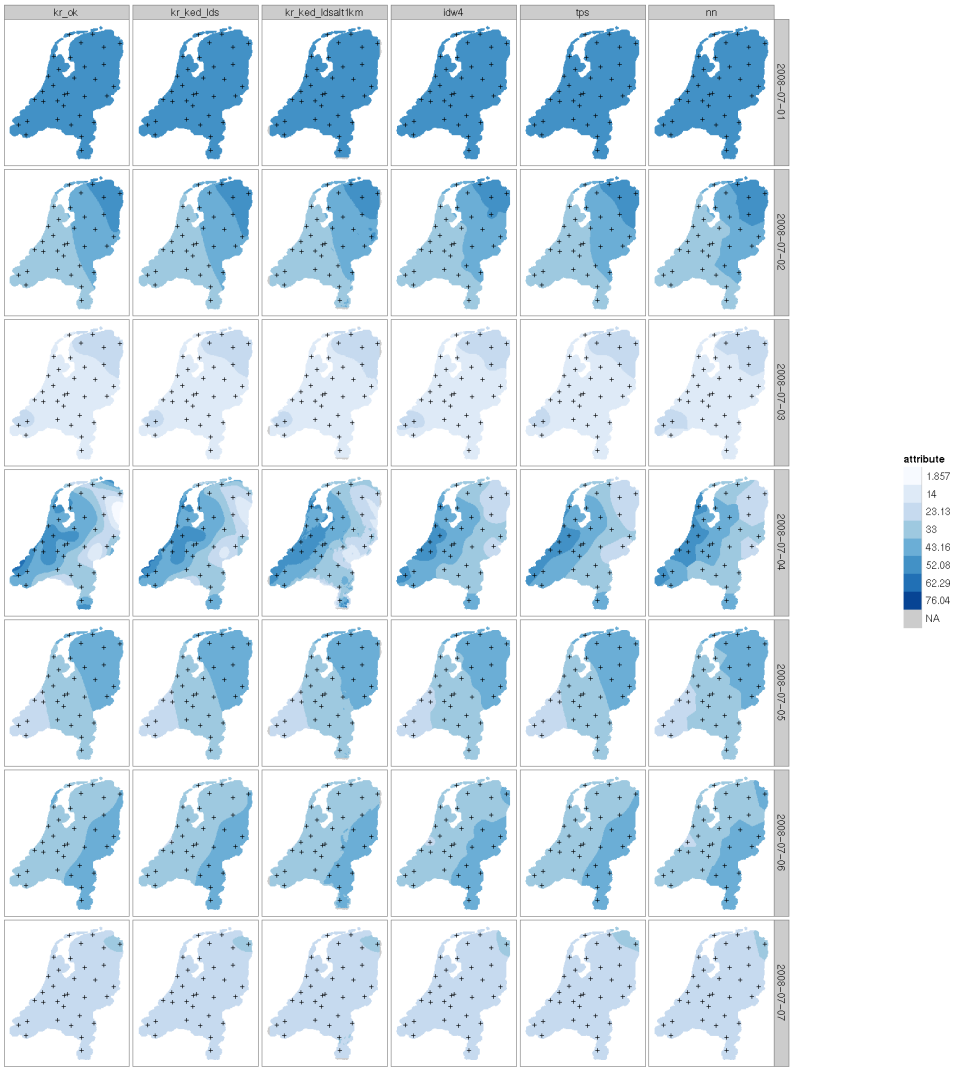


Figure B.6: Interpolated maps of daily Makkink evaporation (10^{-1}mm) for the first week of July 2008. Columns represent different interpolation methods, rows represent time.

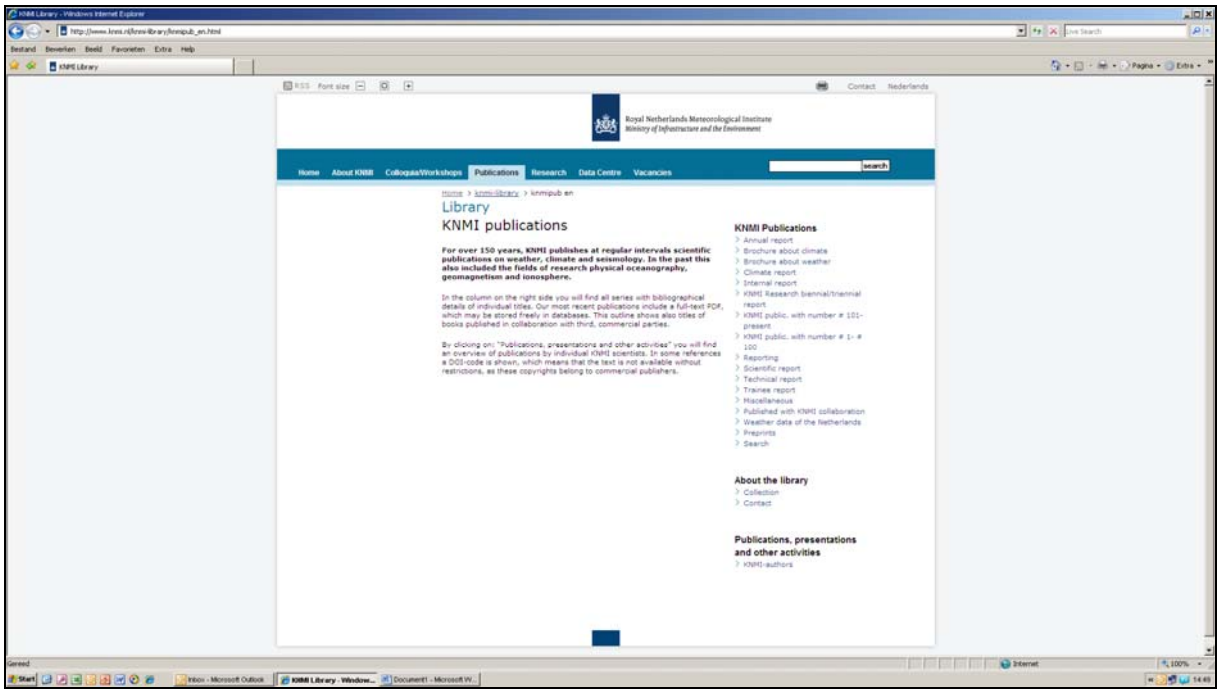
Appendix C

R settings and packages

```
> sessionInfo()
R version 2.12.0 (2010-10-15)
Platform: i686-pc-linux-gnu (32-bit)

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=C              LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] splines   grid      stats     graphics  grDevices utils     datasets
[8] methods   base

other attached packages:
 [1] fields_6.3     spam_0.23-0    xtable_1.5-6   ggplot2_0.8.8  proto_0.3-8
 [6] reshape_0.8.3  plyr_1.4       rgdal_0.6-28   automap_1.0-10 gstat_0.9-74
[11] sp_0.9-72

loaded via a namespace (and not attached):
[1] digest_0.4.2       lattice_0.19-13    RColorBrewer_1.0-2 tools_2.12.0
```

A complete list of all KNMI-publications (1854 – present) can be found on our website

www.knmi.nl/knmi-library/knmipub_en.html



The most recent reports are available as a PDF on this site.