Temperature Reconstruction using Observations and Reanalysis Data

Marieke Dirksen [1,2], Raymond Sluiter[1], Henk van den Brink[1], Albert Klein Tank[2]

[1] Royal Netherlands Meteorological Institute. PO Box 201, 3730 AE De Bilt, The Netherlands. [2] Wageningen University, Droevendaalsesteeg 4, 6708 PB Wageningen

Abstract

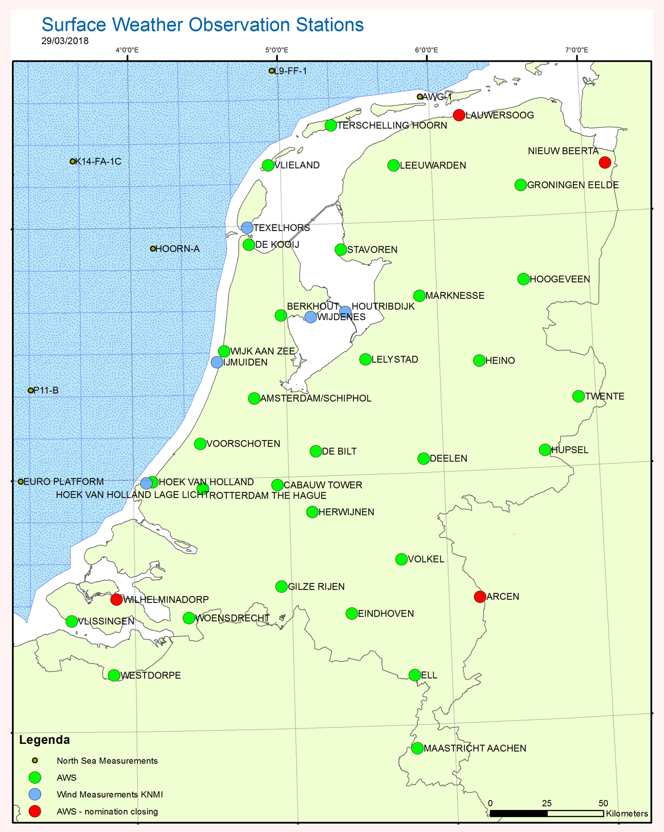
High-resolution, regularly-gridded temperature maps are frequently used in various scientific fields. Within the Netherlands the spatial density of the official automatic weather stations (AWS) is insufficient for this purpose. The AWSs are combined with re-analysis data. In order to explain spatial variations and reduce the size and noise of the re-analysis data set, a Principle Component Analysis (PCA) was performed. The Principle Components (PCs) are related to parameters which determine local temperature patterns, namely: distance to the sea, population density, height, albedo, insolation, roughness, precipitation and vegetation index. This study shows that combining observations and re-analysis data improves the interpolation results.

Temperature, Observations, PCA, Climatology, Spatial Prediction

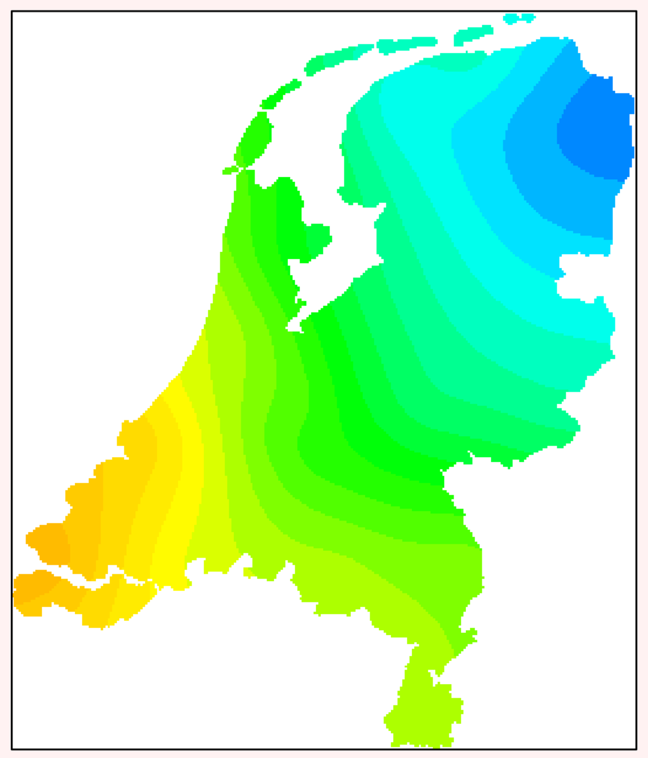
Observations and the Conventional Interpolation Method

Meteorological Stations

It is important to capture local temperature differences: higher temperatures within cities can cause human health issues. Although the Netherlands has a maritime climate with mild summers, during summer city temperatures can be 7 °C higher.
In the Netherlands the temporal resolution of temperature measurements is high, but the spatial density is not optimal. The 34 automatic weather stations (AWS) are approximately 30km apart, which enables a reasonable representation of the spatial variation on country scale, although the local temperature variations are not captured. For this study the daily temperature is reconstructed on a 1km grid for the period 1990-2017.



Kriging Interpolation



The conventional interpolation method only uses the observations from the meteorological stations. The figure on the left shows the climatology based on ordinary kriging. Kriging calculates spatial correlations between the observed and surrounding values. Simple kriging assumes that the covariance between the locations only depends on the distance between the locations and the mean residual is zero. Additionally, Ordinary Kriging (OK) also assumes that the trend is a known mean value ($m(x)$). The estimated value at an unmeasured location ($Z(x)$) is calculated as:

$$Z(x) = m(x) + e(x) \tag{1}$$

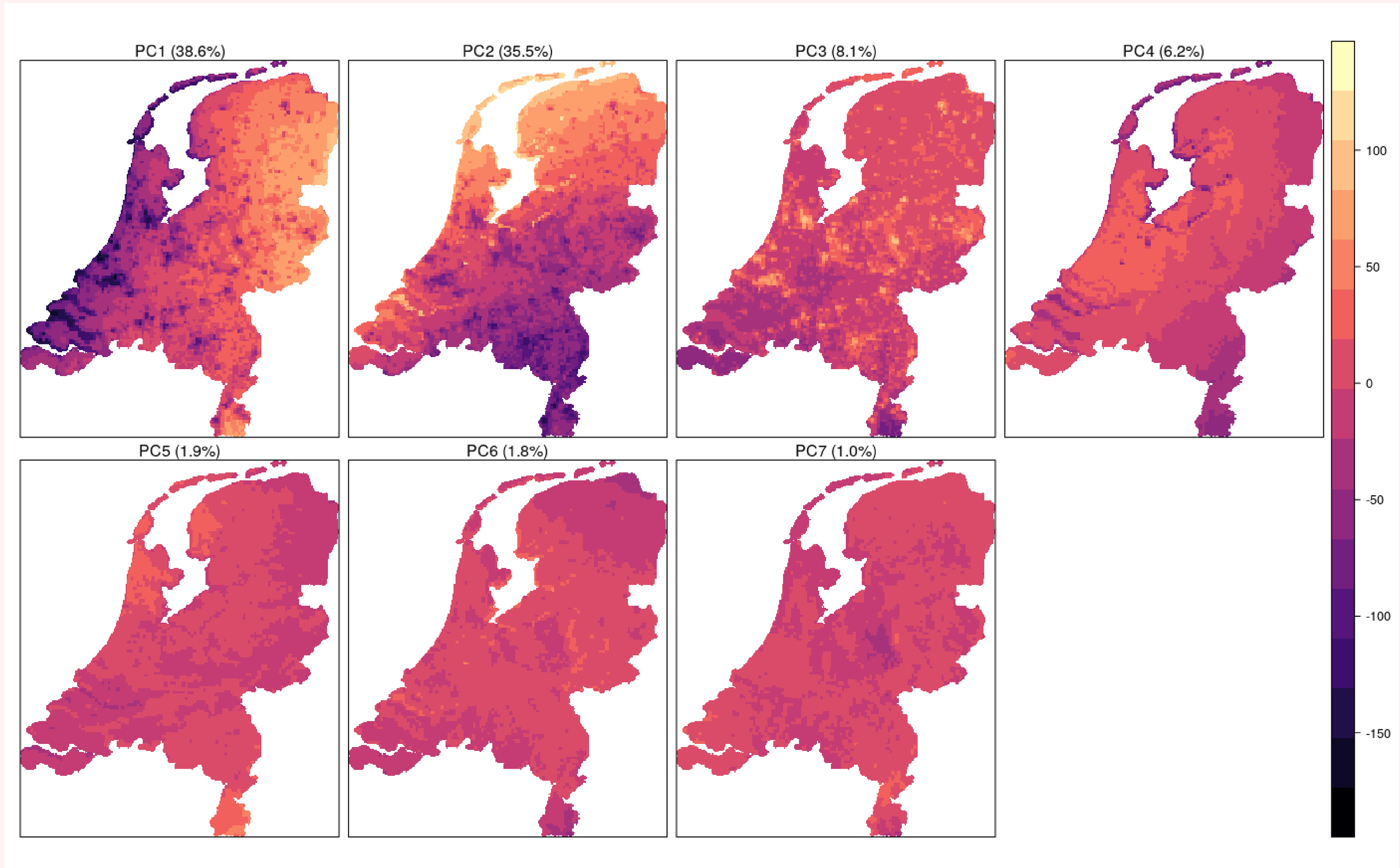
Where $m(x)$ is describing the trend and $e(x)$ is the spatially dependent residual [1].

References

1. Hiemstra, P. & Sluiter, R. *Interpolation of Makkink evaporation in the Netherlands* tech. rep. (2011), 78.
2. Van den Brink, H., Baas, P. & Burgers, G. *Towards an approved model set-up for HARMONIE Contribution to WP 1 of the SBW-HB Wind modelling project* tech. rep. February (2013).
3. Kuhn, M. & Johnson, K. *Applied Predictive Modeling* 5th ed. ISBN: 978-1-4614-6848-6 (Springer Science+Business Media New York 2013, New York, 2016).

Preprocessing of Reanalysis Data

Principle Components from the HARMONIE model



The Numerical Weather Prediction (NWP) model HARMONIE re-analysis from [2], with a grid size of 2.5km, was used. HARMONIE stands for HIRLAM ALADIN research on mesoscale operational NWP in Europe. It is a non-hydrostatic model: the vertical momentum equation is solved explicitly. The boundary conditions and initial conditions come from ERA-Interim, which assimilates data from satellites, in situ measurements, radiosondes, pilot balloons, aircrafts, wind profilers and surface observations.

To reduce the dimensions of the dataset the principle components are calculated, see figure above. The proportion of variance (POV) was calculated as:

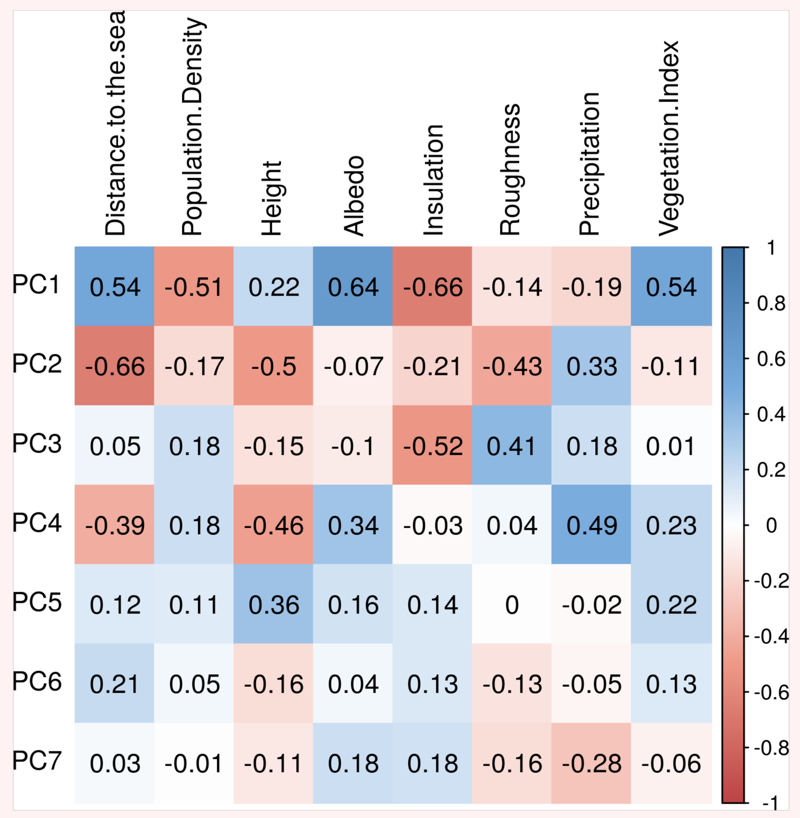
$$POV = \sigma^2 / sum(\sigma^2) \tag{2}$$

Where σ is the standard deviation of the PC.

Interpretation of the Principle Components

The PCs are related, using the Pearson correlation coefficient, to auxiliary datasets which are relevant for local temperature variations:

- **PC1** land/sea temperature gradient, albedo, insolation and vegetation index.
- **PC2** distance to the sea gradient, height, surface roughness and precipitation.
- **PC3** insolation, roughness and population density.
- **PC4** precipitation, height, distance to the sea and albedo.
- **PC5** height.
- **PC6** distance to the sea, height, roughness and insolation.
- **PC7** mainly precipitation, albedo, roughness and insolation.



Combining Observations and Reanalysis Data

Linear Model

In addition to kriging a **linear model** (lm) was used:

$$T_{1.5m} = \beta_0 + \beta_1 PC_1 + \beta_2 PC_2 + \beta_3 PC_3 + \beta_n PC_n + \beta_{...} PC_{...} \tag{3}$$

Where β_0 is the fitted constant, $\beta_{...}$ is the fitted value to the variable ($PC_{...}$). An extensive model description can be found in [3]. The climatology results are compared with **kriging with external drift** (ked), $f1$ uses $PC1$&2, $f2$ uses $PC1...4$ and $f3$ uses $PC1...7$.

Interpolation Results

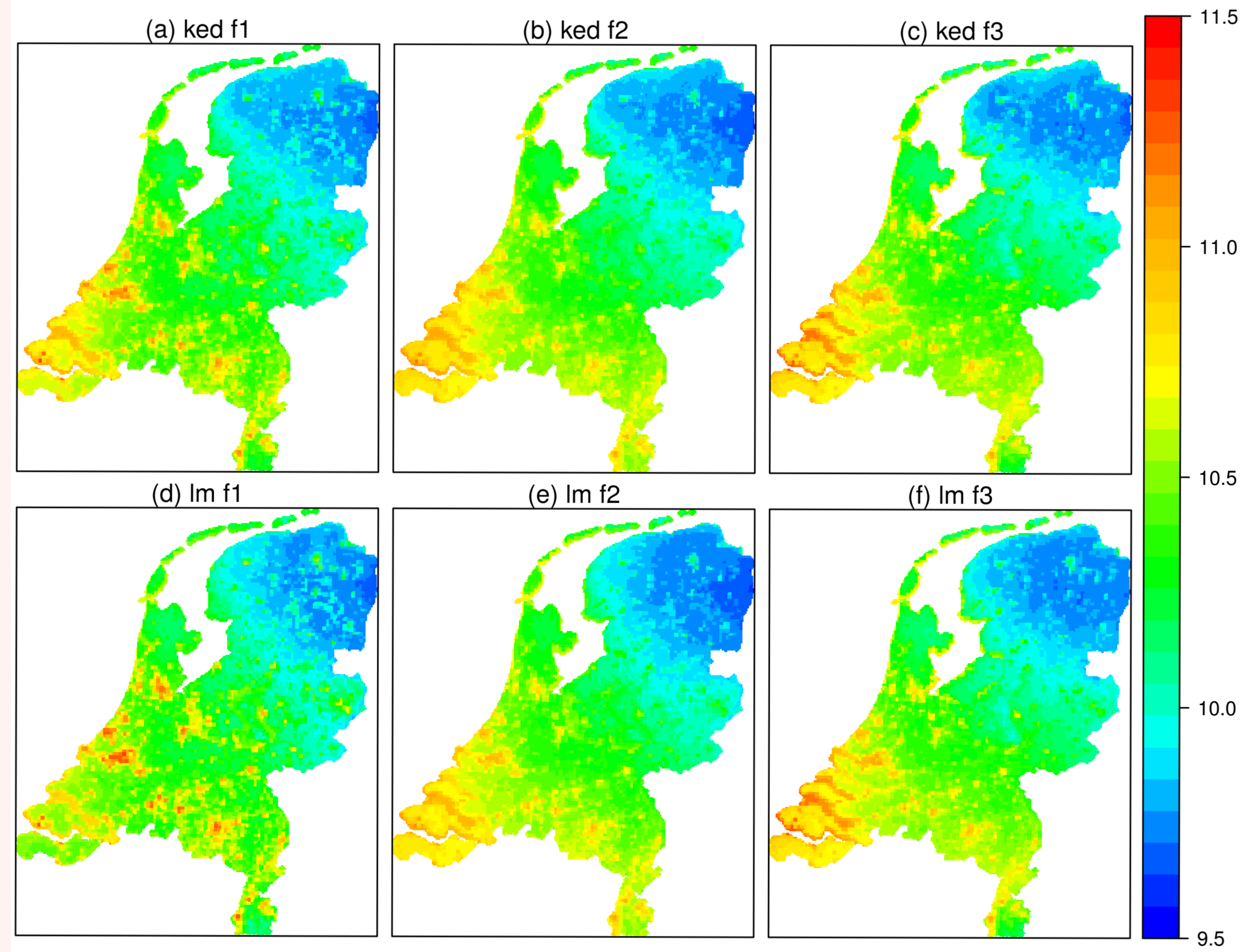


Fig.	**Model**	R^2_{val}	$RMSE_{val}$	R^2_{pred}	$RMSE_{pred}$
	ok	0.67	1.18 °C	0.93	0.11 °C
(a)	ked_{f1}	0.67	1.14 °C	0.88	1.84 °C
(b)	ked_{f2}	0.70	0.98 °C	0.93	1.73 °C
(c)	ked_{f3}	0.67	1.70 °C	0.94	1.90 °C
(d)	lm_{f1}	0.60	0.51 °C	0.63	0.45 °C
(e)	lm_{f2}	0.66	0.47 °C	0.75	0.37 °C
(f)	lm_{f3}	0.63	0.50 °C	0.81	0.32 °C

Table 1: Median R^2 and RMSE values, $_{val}$ and $_{pred}$ stand for cross-validation, in rural areas, and prediction.

Discussion and Conclusion

The temperature differs between the cities and their surroundings for lm and ked. But, the true city temperature can not be verified by the AWSs since they are located in rural areas. Future research is going to focus on crowed-sourced weather data which does have observations within cities. Although city temperatures still need to be verified, the reanalysis data has added local spatially information, and improved the interpolations statistically within the rural areas.

Contact: marieke.dirksen@knmi.nl