Spatial Prediction of Road Temperatures

Marieke Dirksen
KNMI & WUR
Utrechtseweg 297
De Bilt, the Netherlands
dirksen@knmi.nl

Raymond Sluiter
KNMI
Utrechtseweg 297
De Bilt, the Netherlands
sluiter@knmi.nl

Eva Kleingeld
WUR
6708 PB
Wageningen, the Netherlands
eva.kleingeld@wur.nl

Abstract

Ice formation on roads leads to a substantial increase in accidents and road damage. The Dutch Gladheidsmeldsysteem (slippery detection system) measures meteorological conditions on roads at 300+ locations. Our goal is to predict road temperatures in between the measurement locations based on environmental variables and meteorological conditions. Using a machine learning approach part of the variation in road temperatures can be explained by air- and dew point temperature, zenith angle, sky view factor and distance to the sea.

Keywords: Road Temperature, Meteorology, Prediction, Machine Learning.

1 Introduction

Ice on Dutch highways leads to an increase of 77-245% in the number of accidents (SWOV, 2012). To prevent ice formation salt is sprayed on Dutch roads. The salt causes environmental damage, vehicle corrosion and road damage. Road maintenance is expensive. For the regions of Arnhem/Nijmegen, Breda and Friesland the average total cost for winter road maintenance was over one million euros per year (De Vries, 2014).

If we know more precise where and when slipperiness occurs, we can minimize salt spraying. Roads become slippery when e.g. road temperatures drop below zero and surfaces are wet. This research focusses on the road temperature. There is currently no spatial model available which can make detailed predictions for the entire road surface. Our goal is to predict road temperatures based on environmental variables and meteorological conditions using a machine learning approach.

2 Data-sets

2.1 Road Temperature

There are currently over 300 gladheidsmeldsysteem (GMS) stations located on both highways and provincial roads (Fig. 1). The GMS stations are a collection of sensors in the road and along the road which measures among others road-, air- and dew point temperature (Fig. 2). The measurement frequency is 5 minutes. A subset of the data is used for model building (Fig. 3).

2.2 Ancillary Data

Meteorological data. The air- and dew point temperature are measured at each location. The measurements are next to the road. For the spatial predictions these variables are interpolated.

Sky View Factor. The AHN2 point measurements of elevation were regridded on a regular grid. With the R's horizon package the sky view factor was calculated.

Distance to the sea. The distance to the sea is a virtual shoreline with values around zero near the sea and high values inland.

The solar angle. The zenith angle was calculated using the R's insol package.

Figure 1: The gladheidsmeldsysteem stations for each province

Figure 2: Road Temperature measurements on a highway

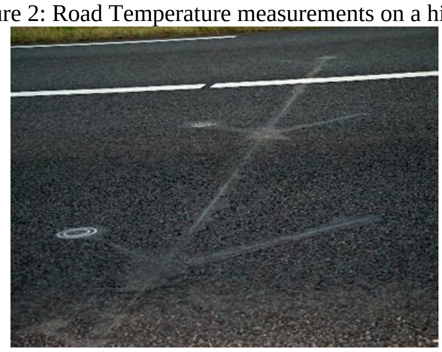
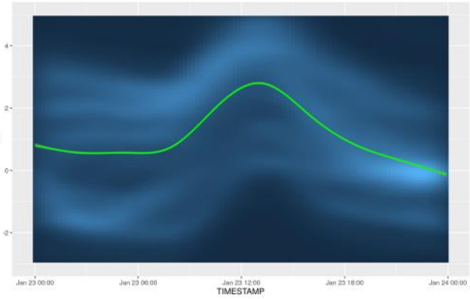


Figure 3: Measurements of all stations on 23th of January 2010, used for model training



3 Methods

Considering the large amount of available data and variables machine learning techniques are used to build a statistical model. One of the main advantages of machine learning is that machine learning algorithms learn from the data that they are fed and that researchers do not need to explicitly model all the interactions between the input variables and the output variable. Furthermore, machine learning algorithms can build models with non-linear data relationships: previous research has shown that the effect of environmental and meteorological variables on the road energy balance are non-linearly mixed (Shao et al., 1997). In this project we used the scripting language *R* (R Core Team, 2016) in combination with *RStudio* (R studio Team, 2016).

3.1 Models

The R's caret package (Kuhn et al., 2013; Kuhn et al., 2016) is used to build a linear model and a decision tree. All the data is scaled, centered and transformed using a BoxCox transformation. Missing values and large outliers are removed.

3.2 Model validation

The models are trained using cross validation. After fine-tuning the models they are validated with an independent test set. For the linear model R^2 and RMSE values of respectively 0.31 and 1.7 $°C$ were found. The tree model has R^2 and RMSE values of 0.77 and 0.98 $°C$.

The tree model is found to be most suitable for spatial predictions of road temperature. Testing the model on the test-set shows similar values for predicted and observed values. The tree model is found to be less sensitive for measurement errors than the linear model.

4 Results

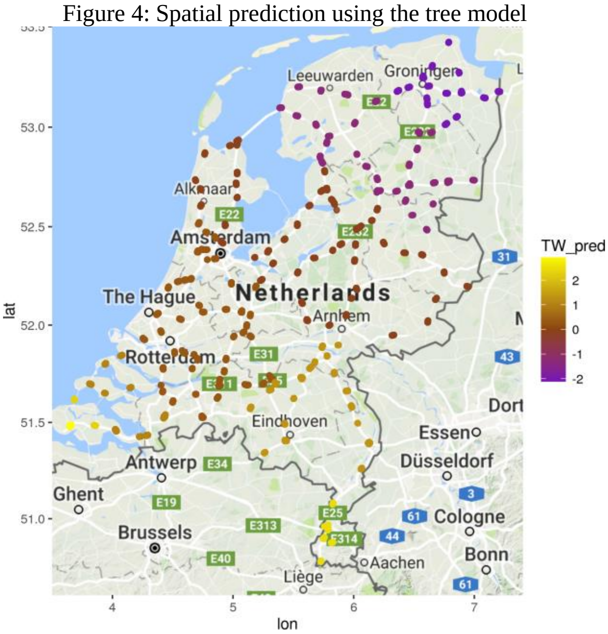
Fist analysis shows the most important variables for predicting road temperature are from most to least important: air- and dew point temperature, distance to the sea, zenith angle and sky view factor.

The tree model is used for spatial predictions (Fig. 4). Higher road temperatures are found in the southern part of the country. In the middle and northern part temperatures are around or just below zero.

5 Discussion and Conclusion

(1) The model is built for a single day, ideally several representative meteorological situation are included in the model.
(2) The model now uses five different variables. Depending on the situation other variable like traffic intensity or road type, should be included.
(3) Compared with the linear model the tree model is found to be less insensitive for outliers.
(4) Statistically the tree model shows good results and seems unbiased. First steps are made to use the model for spatial predictions.

Figure 4: Spatial prediction using the tree model



References

AHN (2015) Handleiding "AHN downloaden van PDOK". www.ahn.nl [*Accessed on January 2017*]

De Vries, D. (2012) Winter road maintenance in the Netherlands; a quantitative analysis of recent and future activities. Technical report, KNMI, De Bilt. Available from: http://bibliotheek.knmi.nl/knmipubIR/IR2014-08.pdf [*Accessed on January 2017*]

Kuhn, M., and Johnson, K. (2013) Applied Predictive Modeling [Hard-cover]. Springer Science + Business Media, New York, 5th edition. ISBN 1461468485. doi: 10.1007/978-1-4614-6849-3. [Online] Available from: http://www.amazon.com/Applied-Predictive-Modeling-Max-Kuhn/dp/1461468485/ref=pd{_}bxgy{_}b{_}img{_}z [*Accessed on January 2017*]

Kuhn, M., Contributions from, Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., the R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., and Candan, C. (2016) caret: Classification and Regression Training. [Online] Available from: https://CRAN.R-project.org/package=caret [*Accessed on January 2017*]

R Core Team (2016) R: A language and environment for statistical computing. [Online] Available from: https://www.r-project.org/ [*Accessed on January 2017*]

Rstudio Team (2016). RStudio: Integrated Development for R. [Online] Available from: www.rstudio.com [*Accessed on January 2017*]

Shao, J., Swanson, J.C., Patterson, R., Lister, P.J., and McDonald, A.N. (1997) Variation of winter road surface temperature due to topography and application of Thermal Mapping. Meteorological Applications, 4(2). Available from: http://onlinelibrary.wiley.com/doi/10.1017/S135048279700042X/pdf [*Accessed on January 2017*]

SWOV (2012) SWOV-Factsheet, De invloed van het weer op de verkeersveiligheid. Technical report, Leidsendam. [Online] Available from:https://www.swov.nl/rapport/Factsheets/NL/Factsheet{_}Invloed{_}van{_}het{_}weer.pdf [*Accessed on January 2017*]