Verification of Road Surface Temperature Forecasts Assimilating Data from Mobile Sensors

VIRVE KARSISTO

Finnish Meteorological Institute, and University of Helsinki, Helsinki, Finland

LAURI LOVÉN

University of Oulu, Oulu, Finland

(Manuscript received 28 September 2018, in final form 24 January 2019)

ABSTRACT

The advances in communication technologies have made it possible to gather road condition information from moving vehicles in real time. However, data quality must be assessed and its effects on the road weather forecasts analyzed before using the new data as input in forecasting systems. Road surface temperature forecasts assimilating mobile observations in the initialization were verified in this study. In addition to using measured values directly, different statistical corrections were applied to the mobile observations before using them in the road weather model. The verification results are compared to a control run without surface temperature measurements and to a control run that utilized interpolated values from surrounding road weather stations. Simulations were done for the period 12 October 2017–30 April 2018 for stationary road weather station points in southern Finland. Road surface temperature observations from the stations were used in the forecast verification. According to the results, the mobile observations improved the accuracy of road surface temperature forecasts when compared to the first control run. The statistical correction methods had a positive effect on forecast accuracy during the winter, but the effect varied during spring when the daily temperature variation was strong. In the winter season, the forecasts based on the interpolated road surface temperature values and the forecasts utilizing mobile observations with statistical correction had comparable accuracy. However, the tested area has high road weather station density and not much elevation variation, so results might have been different in more varying terrain.

1. Introduction

Vehicles providing road condition information will have an important role in the future intelligent transport systems (ITS). They enable detailed knowledge of road conditions along the road network, which is critical to ensure the safety of autonomous vehicles in winter conditions. Traditionally, road condition forecasts have been based on the observations of stationary road weather stations (RWS), but these are sparsely located especially in less populated areas. Road surface temperature is very dependent on local conditions, especially on the openness of the location. There can be even 10°C surface temperature variability at different parts of the road network (Shao et al. 1996; Bogren et al. 2000). Local temperature anomalies may remain unnoticed if the monitoring network is too sparse. However, installing traditional RWSs to also cover the smaller

roads would be very expensive. Vehicle-based observations enable much more detailed road condition monitoring in a cost-effective way. In an ideal situation, the observations could be obtained in a great volume from common vehicles, but a lot of data can be also gathered using selected vehicles with a lot of driving kilometers such as buses and taxis.

A prerequisite for accurate forecasts is information about the current conditions at the forecast area. It is important that the initial conditions in the forecasting model represent real conditions, because even small deviations can change the forecast. Mobile observations have the potential to improve the forecast accuracy for road stretches without RWSs. Accurate road condition forecasts enable well-timed road maintenance actions, which reduce the road accident risk and potentially lead to economical savings as unnecessary actions can be avoided. For example, roads can be salted beforehand if the forecast correctly estimates when the road surface temperature will drop below the freezing point. This prevents

Corresponding author: Virve Karsisto, virve.karsisto@fmi.fi

DOI: 10.1175/WAF-D-18-0167.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

the roads from freezing and becoming dangerous for drivers. However, the roads should not be salted too early, as traffic causes the salt to wear off from the road. More accurate forecasts also produce less false alarms, and there will be fewer cases when the roads are salted although the surface temperature remains above 0°C. Accurate forecasts also enable detailed warnings about dangerous road conditions beforehand so that road users can select their route and adjust their driving behavior accordingly.

Thermal mapping has been commonly used to estimate surface temperatures on road sections without RWS measurements (Thornes 1991; Shao et al. 1997; Gustavsson 1999). The method is based on the idea that the spatial behavior of surface temperature is reproduced in similar weather situations. The surface temperature along the target road network is measured several times using a measurement instrument attached to a vehicle. The measured temperature variability can be used later to determine the surface temperature along the road network. However, the increasing availability of mobile observations will make it possible to obtain real-time observations along roads in the future. Modern cars are equipped with a multitude of sensors and can provide road condition and weather information. Several earlier studies discuss the potential use of cars as sensors (Allegretti and Bertoldo 2014; Petty and Mahoney 2007). A disadvantage of thermal mapping and mobile observations is that they provide only a snapshot of the current situation and do not tell about the temporal behavior of surface temperature (Chapman and Thornes 2006). This might be changed if the majority of cars would provide observations in the future, but currently the availability of data is still a problem. Obtaining data from the vehicles' internal communication systems is difficult, because in normal cars an unauthorized accessing of the car's data can even cause the rejection of the vehicle's warranty. Data can be more easily collected from heavy vehicles, but there still might not be detailed instructions for how to interpret the data contents.

The aim of the research presented in this paper was to find out if the road surface temperature forecast accuracy can be improved by using observations from vehicles. A disadvantage of mobile observations is that they are more exposed to disturbances, such as heating from the vehicle or drifting snow. The quality of the measurements made with mobile sensors might not be as high as measurements made at RWSs. The data quality of mobile observations should be studied before using them for verification, postprocessing, or assimilation in a forecasting system. It should also be ensured that the forecast accuracy increases when the forecast utilizes the new data.

This study presents verification results of road surface temperature forecasts that assimilate observations made

with an optical Teconer RTS411 instrument (Haavasoja et al. 2012). The next section describes the Finnish Meteorological Institute's (FMI) road weather model. Section 3 gives information of the datasets used in the simulations and explains the statistical calibration methods applied to the mobile observations. Results of the forecast verification are described in section 4. Section 5 contains a discussion, conclusions, and suggestions for further research.

2. Road weather model

The Finnish Meteorological Institute's road weather model (RWM) has been in operation since 2000 (Kangas et al. 2015). The model is a one-dimensional heat balance model and aims to predict the temperature at the surface. Other model outputs are the amounts of water, ice, snow, and frost on the road, friction, road condition, and traffic index (Fig. 1). The road condition describes the status of the road, which can be for example "wet" or "icy." The traffic index depicts the overall driving conditions.

The calculation of surface temperature is based on the surface energy balance (Brutsaert 1984):

$$G = I_{\text{net}} - H - \text{LE} + \text{Tr}, \quad (1)$$

where G is heat flux into the ground, I_{net} is net radiation at the surface, H is sensible heat flux, LE is latent heat flux, and Tr describes heating caused by traffic. Figure 2 illustrates the model energy fluxes. The ground is divided into several layers, which are also shown in Fig. 2. The calculation of heat transfer between the ground layers is based on (Patankar 1980):

$$\rho_g c_g \frac{\delta T(z, t)}{\delta t} = \frac{\delta}{\delta z} K \frac{\delta T(z, t)}{\delta z}, \quad (2)$$

where T is temperature, z is vertical distance in the ground, t is time, K is heat conductivity, ρ_g is density, and c_g is specific heat capacity of the ground. To calculate the temperature at the next time step, Eq. (2) is integrated over the time step and the volume of the layer and solved using the forward difference explicit method:

$$T_i^{j+1} = T_i^j + \frac{1}{\rho_g c_g \frac{z_{i+1} - z_{i-1}}{2\Delta t}} \times \left(K_i \frac{T_{i+1}^j - T_i^j}{z_{i+1} - z_i} - K_{i-1} \frac{T_i^j - T_{i-1}^j}{z_i - z_{i-1}} \right), \quad (3)$$

where index i refers to the ground layer and index j to time, Δt is the model time step, and K_i means heat conductivity between layers $i + 1$ and i .

FIG. 1. Road weather model input and output values. Precipitation phase and road surface temperature are optional. Input road surface temperature is used only in model initialization.

More detailed descriptions of the model physics are given by [Kangas et al. \(2015\)](#) and [Karsisto et al. \(2017\)](#). The model has been recently updated, and the model version used has some differences when compared to the models described in the publications. One of the most influential changes is the model time-step reduction to 30 s for the whole simulation. The original model had a time step of 5 min. Another major change was done to the ground temperature profile calculation. In the original model, the temperature profile for the next time step was determined iteratively by solving the temperature for each layer simultaneously ([Kangas et al. 2015](#)). In the new model version, new temperature for each layer is calculated directly implementing Eq. (3) similarly as in the METRo model ([Crevier and Delage 2001](#)).

The RWM requires atmospheric values (air temperature, humidity, wind speed, precipitation, and short-wave and longwave radiation) as input. The model run consists of the initialization phase and the forecast phase. In the initialization phase, the atmospheric values are obtained from the observations, and the model is run for circa 2-day period to get a good initial state for the ground temperature profile. However, the radiation parameters are obtained from the 3D numerical weather prediction model (NWP) due to the lack of observations. In the forecast phase, all the aforementioned atmospheric values are obtained from the NWP forecast. The FMI RWM is not a stand-alone model, but it relies on the NWP model in regard to the large-scale weather prediction. Using the values from the NWP forecast, the model aims to predict the road conditions at the forecast point.

The RWM uses mobile road surface temperature observations in the initialization by the coupling method ([Crevier and Delage 2001](#); [Karsisto et al. 2016](#)). The method iteratively determines a correction coefficient for the radiation so that the surface temperature in the

simulation fits to the observed road surface temperature. This iterative radiation adjustment period was set to 3 h in the model. The coefficient is given for either short-wave or longwave radiation depending on which has the higher value at the start of the coupling period. The effect of the coefficient decreases exponentially as the forecast advances. The coupling method makes it possible to use even a single road surface temperature observation efficiently.

3. Data and methods

The road surface temperature forecasts were produced for the RWS points. The forecasts were not run in

FIG. 2. The depths of the ground layers in the road weather model (m) and a schematic representation of energy fluxes.

real time but as hindcasts afterward. Surface temperature observations from the stations were used for forecast verification. The results are compared to two types of control runs that did not use the mobile observations to find out the effect of the new data on the forecast. The first type of control run did not use any kind of road surface temperature observations. The second type utilized interpolated road surface temperature values, which were obtained by using observations from independent RWSs, as in a leave-one-out cross-validation experiment. The first control run simulates a forecast situation where the road weather station network is very sparse and the second run a situation with multiple stations nearby the forecast target point.

a. Mobile measurements

Teconer RCM411 and RTS411 are optical measurement instruments that can be attached to vehicles (Haavasoja et al. 2012). RCM411 measures optically water layer thickness, road condition, and friction. The functionality of the device is based on the absorption of radiation at near-infrared wavelengths. Dry, wet, and icy roads reflect the radiation differently at these wavelengths, and thus the road condition can be determined. RTS411 measures road surface temperature by detecting infrared radiation. The measurements are done every second. The mobile measurements used in the simulations were from Teconer devices attached to vehicles that passed selected RWS points in southern Finland. An average of road surface temperature measurements within 50 m from the RWS during a pass was used as the RWM model initialization value.

Before this study, statistical analysis was performed for RWS and RTS411 road surface temperature measurements to find out possible differences. Mobile sensors are more exposed to disturbances than RWS, like heating from the car and drifting snow. For example, the infrared radiation emitted by the vehicle can reflect from the road to the optical surface temperature sensor and thus cause a warm bias. The statistical analysis was based on linear mixed effect models and used data from three winter periods between years 2014 and 2017. The data were gathered with a rendezvous sensor fusion method similarly as in this study, identifying rendezvous cases where a vehicle passed an RWS. The averages of road surface temperature measurements within a 50-m radius of the RWSs during each rendezvous were calculated. The analysis is covered in detail by Lovén et al. (2019). According to the results, the Teconer observations are on average 0.62°C warmer than the RWS observations in dry conditions and at 0°C temperature. The temperature difference was found to be dependent also on the road status (dry, wet, or icy) and the individual

device. Water, ice, and snow on the road surface can cause the optical sensor to measure the temperature of the substance rather than the actual road surface temperature. This causes the difference in comparison with the RWS measurements, because at most RWSs in Finland the surface temperature observations are done with asphalt embedded sensors. The device dependency can be partially caused by the device's location in the vehicle. For example, nearby warm engine bodies can emit infrared radiation that disturbs the surface temperature measurements when this radiation is reflected from the road surface to the sensor.

b. Statistical calibration methods

Three sensor calibration methods for the mobile surface temperature data were tested, which were all based on linear statistical models. The first one used the calibration provided by a linear mixed effect model (Robinson 1991; West et al. 2007) constructed by Lovén et al. (2019). The model was fitted with Teconer and RWS observations from winter periods between the years 2014 and 2017. Winter period 2017–18 used in this study was not included in the model development. The model assumed that RWS and mobile road surface temperature observations are linearly related, with the road status (i.e., dry, wet, icy, etc.) affecting the coefficients of the linear relationship. Further, the model assumed that the current calibration levels of both RWS and mobile sensors (i.e., their reading when actual road surface temperature was 0°C), each follow individual Gaussian distributions:

$$\mathbf{T}_{s,\text{RWS}} = \mathbf{X}_{\text{Mob}}\boldsymbol{\beta} + \mathbf{Z}_{\text{id}}\mathbf{b} + \boldsymbol{\varepsilon}. \quad (4)$$

Here, $\mathbf{T}_{s,\text{RWS}}$ corresponds to the RWS road surface temperatures, $\mathbf{X}_{\text{Mob}}$ is a design matrix of the mobile road status and surface temperature observations as well as the overall intercept, corresponding to dry road status, while $\boldsymbol{\beta}$ is the fixed effect coefficient. The term $\mathbf{Z}_{\text{id}}$ includes dummy sensor identifiers for both the mobile and the RWS sensors, with the random coefficient $\mathbf{b} = (b_1, b_2)$, corresponding to mobile (b_1) and RWS (b_2) calibration levels, distributed normally as $b_1 \sim N(0, I\sigma_1^2)$ and $b_2 \sim N(0, I\sigma_2^2)$. The noise term $\boldsymbol{\varepsilon}$ is also assumed Gaussian with zero mean.

The calibration in Eq. (5), corresponding to the expectation of $\mathbf{T}_{s,\text{RWS}}$, adjusts the mobile observations with a linear function of the fixed coefficients:

$$T_{s,\text{Adj}} = \beta_0 + \beta_1 T_{s,\text{Mob}}. \quad (5)$$

Here, $T_{s,\text{Adj}}$ is the calibrated mobile road surface temperature, while $T_{s,\text{Mob}}$ is the original mobile observation,

TABLE 1. Status-based coefficients used in Eqs. (2)–(4).

Road status	β_0	β_1	β_2	β_4	β_5
Dry	0.03	1.13	1.08	0.00	1.13
Moist	−0.79	1.02	1.08	−0.72	1.02
Wet	−0.96	1.04	1.08	−0.90	1.04
Snow/frost	−1.00	0.91	1.08	−0.86	0.86
Ice	−0.97	0.97	1.08	−1.07	0.95
Slush	−1.01	0.96	1.08	−1.02	0.98

TABLE 2. Device-based coefficients used in Eqs. (3) and (4).

Device	β_3	β_6
A	−0.60	−0.25
B	−1.08	−0.43
C	0.21	0.69
D	−0.42	0.12
E	−0.87	−0.20
F	−1.67	−1.11
G	−1.43	−1.02

β_0 is the estimated road status intercepts, and β_1 is the estimated interaction coefficient between the road status and the mobile surface temperature. The second and third models modified Eq. (3) by including individual intercepts for the mobile sensors instead of assuming them to be normally distributed. Further, the second model left out the coefficients for the road status observations. Corresponding calibration equations are provided below as Eqs. (6) and (7), while the estimated fixed-effect β coefficients for all calibration equations, obtained by fitting the models, can be found in Tables 1 and 2:

$$T_{s,Adj} = \beta_2 T_{s,Mob} + \beta_3, \quad (6)$$

$$T_{s,Adj} = \beta_4 + \beta_5 T_{s,Mob} + \beta_6. \quad (7)$$

In the following sections, the first model [Eq. (5)] is referred as “status-based correction,” the second model [Eq. (6)] as “device-based correction,” and the third model [Eq. (7)] as “status- and device-based correction.”

c. Kriging and forecast data

The RWM used interpolated air temperature, humidity, and wind speed values as atmospheric forcing in the initialization phase. The values were interpolated from surface synoptic observations (SYNOP) weather stations to a 1-km grid with a universal kriging method (Cressie 1993; Aalto et al. 2013). Elevation data, lake percentage, and sea percentage were used as explanatory variables in the interpolation. Lake percentage means the percentage of lake surface in the grid cell, and correspondingly, sea percentage means the percentage of sea surface in the grid cell. Altitude and water bodies cause spatial trends in the air temperature, humidity, and wind speed values. For example, air temperature decreases with altitude and the large heat capacity of water may cause it to change slower near lakes. These trends are removed from the data in the kriging analysis, and the spatial autocorrelation is calculated from the residuals. The gridded values were interpolated from the grid to the simulation points with bilinear interpolation. Hourly precipitation values were obtained from weather radar observations. Interpolated road surface temperature

values for the second control run were also obtained with the kriging method by using observations from road weather stations. The interpolation was done to a 10-km grid and was performed separately for each simulation point so that the road surface temperature observations from that point were not included in the calculation. After the kriging analysis, the value for the simulation point was obtained by bilinear interpolation from the nearest grid points.

The atmospheric values used in the forecast phase of the model were obtained from the forecasts made with the HARMONIE-AROME (HIRLAM–ALADIN Research on Mesoscale Operational NWP in Euromed–Applications of Research to Operations at Mesoscale) model configuration. The model is a convection-permitting nonhydrostatic NWP (Bengtsson et al. 2017). The forecasts were run four times a day for the Scandinavia region with 2.5-km resolution by FMI. The analysis times of the runs were 0000, 0600, 1200, and 1800 UTC. Local time in Finland is UTC +2 h during wintertime and UTC +3 h during summertime. The simulations were done for road weather station points located on the Helsinki–Turku motorway or on the nearby roads (Fig. 3, Table 3). Road surface temperature measurements used in the forecast verification were done with asphalt-embedded Vaisala DR511 (Vaisala 2001). As the measurement interval of RWSs is 5–10 min, the nearest measurement of the verification time was used. The used NWP data, road weather station measurements, and hourly radar values are available in a public repository (Karsisto 2018).

d. Model run construction

The model runs consist of a 52-h initialization phase and a 21-h forecast phase (Fig. 4). In the initialization phase, the air temperature, wind speed, and humidity values were obtained from the kriging data. Radar measurements were used to obtain hourly precipitation values, but the precipitation phase was taken from HARMONIE forecasts so that the values from the first six forecast hours from each run formed a continuous time series. Shortwave and longwave radiation for the initialization were obtained similarly from forecasts. The forecast phase starts three hours after the

FIG. 3. Road weather station points in Southern Finland. Blue circles represent points included in the forecast verification in this study. The size of the circle shows how many simulations in that point were able to utilize mobile observations. Brown dots show other road weather stations. All of the stations along the Helsinki-Turku motorway did not have observations during the studied period.

HARMONIE analysis time and utilizes the latest HARMONIE forecast. If there is a mobile surface temperature observation available within this 3-h period, it is utilized in the simulation with the coupling method. The simulation used the latest observation if there were multiple observations available at the forecast point during this time window. The second control run used interpolated road surface temperature values for the corresponding times so that it would not have an advantage by using more recently made observations. The coupling method's radiation adjustment period was set to 3 h. Simulations were done four times a day for the period 12 October 2017–30 April 2018. As forecasts start three hours after the HARMONIE analysis time, the corresponding forecast starting times were 0300, 0900, 1500, and 2100 UTC. The simulations were not done as real-time forecasts but as hindcasts afterward. The output surface temperature values for the forecast phase are available also in the public repository (Karsisto 2018).

e. Verification methods

Verification results were calculated separately for the periods 12 October 2017–17 February 2018 and 18 February–30 April 2018 because the daily surface temperature variation was much stronger during the springtime than winter. This caused great differences in the verification results for the two periods. The date of 18 February was chosen for the date to separate the two periods because after that the average daily surface

TABLE 3. List of road weather stations included in the study.

Latitude (°N)	Longitude (°E)	Name
60.42402	22.40233	Kaarina
60.44583	22.53482	Piikkiö
60.4237	22.68302	Paimio
60.42836	22.84023	Pitkärporras
60.43515	22.96914	Hajala
60.41729	23.10632	Halikko
60.4074	23.21123	Haukkala
60.38308	23.23238	Tupuri
60.38867	23.29192	Muurla
60.37346	23.35686	Lakiamäki
60.38049	23.4779	Kruusila
60.39117	23.5713	Suomusjärvi
60.37624	23.72575	Lahnajärvi
60.32855	23.93642	Pitkämäki
60.31271	24.00195	Karnainen
60.28811	24.15869	Lehmijärvi
60.29328	24.31916	Palojärvi
60.22805	24.59641	Nupuri

temperature variation over the simulation points used was almost consistently above 10°C. Only the cases where there was a mobile observation available within three hours before the start of the forecast phase were included in the verification. The verification concentrates on the first 10 forecast hours because the effect of the coupling method decreases as the forecast advances.

The bootstrap method was used to calculate whether the differences in the verification scores between runs were statistically significant (Efron and Tibshirani 1993; Hogan and Mason 2011). Multiple samples from the dataset were generated, and the differences between the scores were calculated from each sample. The 95% confidence interval was calculated from the distribution of differences. If 0°C was not in the obtained range, the differences between models were considered statistically significant with 95% confidence (Hogan and Mason 2011). The upper and lower limits of the range were rounded to two decimals before the check except in the case of the correlation coefficient where four decimals were used due to small differences. In this study the sample size used was the same

FIG. 4. The constructions of a model run.

FIG. 5. Scatterplots of road surface temperature measured at road weather stations (x axis) compared to mobile observations used in the initialization of model runs during the midwinter period (y axis). Bias, RMSE, and correlation coefficient values are shown in the corners of each panel. (a) Mobile observations without statistical correction, (b) device-based correction, (c) status-based correction, and (d) status- and device-dependent correction methods.

as the size of the dataset used, and the number of generated samples was 10 000. Samples were generated with replacement.

4. Results

a. Input mobile measurements

The mobile road surface temperature measurements used in the RWM initialization were compared to the road weather station measurements to find out the initial error in the model runs. In Fig. 5, there are four scatterplots with surface temperature measured at RWSs on the x axis and mobile measurements on the y axis. Different adjustment equations are applied to the mobile measurements in Figs. 5b–d. According to the figure, the Teconer observations have a warm bias compared to the RWS observations during the winter period. The warm bias is reduced when a statistical correction is applied to the Teconer measurements. The bias even turns negative when the correction based on road status is applied. In addition, all of the correction methods seem to cause too cold values in cases where RWS surface temperature was below -5°C . The correction methods are more in line with RWS observations in temperatures over 0°C , although there is

some warm bias when the device-based correction is applied. The figure shows also that all correction methods reduce the root-mean-square error (RMSE) of Teconer observations. The best results are obtained when the status- and device-based correction method is used (Fig. 5d). The RMSE and bias differences between correction methods were determined to be statistically significant with 95% confidence by using the bootstrap method as explained in section 3e. Differences between correlation coefficients were statistically significant in all cases except between the status-based correction method and in the case where no correction was used.

Figure 6 contains similar scatterplots as Fig. 5 but for the spring period. It shows that during spring time the Teconer observations are also usually warmer than RWS observations in temperatures below -5°C . The correction methods again reduce the warm bias of Teconer observations in these temperatures. Conversely, RWSs usually give warmer surface temperature measurements than Teconer instruments when the RWS observation is over 5°C . This is probably caused by the asphalt-embedded sensor heating too much during sunny days. The correction methods taking into account the road status somewhat reduce the warm bias, although the Teconer instrument might actually be more correct

FIG. 6. As in Fig. 5, but for the spring period.

in the sunny cases by showing colder temperatures. The asphalt sensor overheating is probably the reason for the higher RMSE values during the spring period than during the winter period. To test this hypothesis, the Teconer observations were compared to the road surface temperature measurements done at stations equipped with the Vaisala DST111. DST111 measures surface temperature optically similar to the Teconer RTS411, so it should not have the overheating issue. The measurements were done during time period from 1 October 2017 to 30 April 2018. For reference, the Teconer observations done with the same individual devices were compared to the measurements done with asphalt-embedded sensors at the RWSs. However, these measurements were from different stations than where the DST111 measurements were done. The Teconer observations clearly had more negative bias values when compared to measurements done with the asphalt-embedded sensors than when compared to the measurements done with DST111. Only cases where the RWS observation was above 5°C were included in the comparison. This result indicates that the reason for Teconer observations being colder than RWS measurements at high temperature values (Fig. 6) is the asphalt sensor overheating and is not an issue related to the Teconer devices.

Overall, the status- and device-based correction method (Fig. 6d) gave the best results also during the

spring period. The statistical significance of the score differences between correction methods was calculated similarly as for the winter period. The RMSE and bias differences were determined to be statistically significant at 95% confidence in all cases except between the RMSE values of the two methods utilizing status-based correction. Differences between correlation coefficients were statistically significant in all cases except when the device-based correction method was compared to the cases with no correction and between the correction methods utilizing the status-based correction.

b. Interpolated values

To better explain the forecast verification results, scatterplots in Fig. 7 compare the interpolated road surface temperature values used in the second control run to road weather station measurements. According to the figure, there is some cold bias in both the winter and spring periods. During the spring period, the RMSE value is considerably smaller than the values for statistically corrected mobile observations. Road surface temperature is measured with asphalt-embedded sensors at most road weather stations in Finland, so overheating in sunny weather affects the interpolated values. Consequently, the interpolated values are closer to RWS measurements than mobile measurements that are not affected by the overheating effect.

FIG. 7. Scatterplots of road surface temperature measured at road weather stations (x axis) compared to interpolated values from other road weather stations used in the initialization of the second control run (y axis) during the (a) midwinter period and (b) spring period. Bias, RMSE, and correlation coefficient values are shown in the corners of each panel.

c. Case study

One of the forecast cases is described in detail to demonstrate how the mobile observations contribute to the forecast. Figure 8 shows the modeled surface temperatures from all the forecast runs started at 0300 UTC 24 December for the Paimio station in addition to the RWS observations. The stars show the value used by the coupling method in the corresponding model run. All the simulations have the same modeled surface temperatures until the start of the coupling phase. In the coupling phase, the radiation correction coefficient is determined so that the modeled value fits to the observed value. The first control run that did not use the surface temperature observations and thus did not use the coupling method was too warm before the start of the forecast. The mobile observation is also too warm but brings the modeled surface temperature a little closer to the RWS value. Modifying the mobile observation with one of the three statistical corrections further decreases the modeled temperature. However, the kriged value in this case is too cold and causes also the first two forecast hours of the second control run to have too low temperature values. Nevertheless, as the observed temperature decreases, the run assimilating the kriged value is closest to the RWS observations at the third forecast hour. At the first two forecast hours, the runs using the mobile observation with a statistical

correction are closest to the RWS observations. The RWS surface temperature starts increasing at 0600 UTC, and after two hours all the runs give too low surface temperatures because the temperature increases too slowly in the model. The differences between runs decrease gradually as the lead time increases.

d. Forecast verification for all stations

Only the cases where the RWS surface temperature was below 10°C were included in the forecast verification as we are mainly interested in near-zero temperatures. This also reduces the error caused by asphalt sensor overheating, although it probably still has an effect on the verification results. Figures 9–11 show RMSE and bias values of the forecasts for both winter and spring periods. The results for the spring period are separated by the forecast start time due to the strong daily surface temperature variations.

Using the mobile observations in the RWM initialization reduced the RMSE values at the beginning of the forecast for both verification periods, when compared to the first control run (Figs. 7a and 9). Using the mobile observations without a statistical correction causes a warm bias in the model during the winter period (Fig. 7b). This is in line with the fact that the initial surface temperature values measured with the Teconer instrument tend to be warmer than the RWS

FIG. 8. Example of road surface temperature forecasts done for Paimio station for 24 December 2017. Observations are shown by the black dashed line. The black continuous line shows the control run without assimilated road surface temperatures, the red line is the control run with kriged and interpolated RWS temperatures, the blue line is the run using mobile observations, the brown line is the run using mobile observations with status-based correction, the pink line is the run using mobile observations with device-based correction, and the cyan line is the run using mobile observations with both status- and device-based correction. The vertical line shows where the initialization phase ends and the forecast phase starts. Stars show the surface temperature values assimilated in model runs by using the coupling method. The stars' colors correspond to the line colors of the model runs.

values during the winter (Fig. 5). Adding statistical correction to the Teconer observations reduced the warm bias. The negative bias of the correction method applying road status seen in the initial temperature values also causes negative bias in the beginning of the forecast during the winter period. Looking for both the bias and RMSE values for the winter period, the correction method applying both the status and device correction gives the best results of the runs utilizing mobile observations. However, the differences in the RMSE values were statistically significant with 95% confidence between these three runs only at the first forecast hour. Starting from the second forecast hour, the difference between the two runs utilizing the device-based correction was not found to be statistically significant with the bootstrap method. Starting from the fourth forecast hour, the differences between each of the three runs utilizing the statistical correction methods were not statistically significant with 95% confidence. (The results of the statistical significance tests can be found in Tables A2 and A3 in the appendix.)

The second control run utilizing the interpolated observations gives considerably smaller RMSE values than the first control run during the winter period (Fig. 7a). However, the negative bias seen in the initial values is present also during the first forecast hours (Fig. 7b). The RMSE values of the second control run are very close to the values of the runs using mobile observations with the device-based correction methods during winter. At the first two forecast hours the results are not different enough to reach the 95% significance level. The bias values of the second control run are very similar to the run using mobile observations with the status-based correction method.

The observation error (Fig. 6) and the forecast error for the model runs assimilating mobile observations with a statistical correction (Fig. 9) are about the same order of magnitude for wintertime. The same is true for the RMSE of the kriged surface temperature values (Fig. 7) and the RMSE for the runs assimilating them (Fig. 9). The forecast error is comparable to the observation uncertainty, and thus it is possible that the observation error has a dominant effect on the forecast accuracy. This was

FIG. 9. (a) RMSE and (b) bias values of road surface temperature forecasts as function of forecast lead time. The results are calculated for the midwinter period and for all forecast starting times. Only the cases where the RWS surface temperature was below 10°C were included in the forecast verification. The black line shows results for the control run without assimilated road surface temperatures, the red line for the control run with kriged and interpolated RWS temperatures, the blue line for the run using mobile observations, the brown line for the run using mobile observations with status-based correction, the pink line for the run using mobile observations with device-based correction, and the cyan line for the run using mobile observations with both status- and device-based correction.

investigated by studying the results of a third control run that is not otherwise analyzed in this study. The model settings were similar than in the other runs, but the assimilated surface temperatures were taken from the RWS. Thus, the third control run did not involve error caused by surface temperature observation uncertainty, as the RWS observations are considered “truth” in this study. The third control run’s RMSE values for the first two forecasts hours for the winter period were only slightly lower than for the model run assimilating mobile observations with the status- and device-based correction. Starting from the third forecast hour, the RMSE values of these runs were rather similar. It can be concluded that the forecast error is the same order of magnitude regardless of the observation uncertainty involved in statistically corrected mobile observations. This is the case also with kriged values. Uncertainty related to the forecasting model and other input values seems to also

have a great effect on the forecast error. The effects of errors in the initial values on the forecast accuracy are complicated and do not directly accumulate with forecast error. The erroneous values can even improve the forecast in some cases if they compensate for some other error in the forecast.

During the springtime, the forecast accuracy was not as good as during the winter period according to the verification results (Figs. 10 and 11). The daily temperature variation was much stronger, and the RWM had difficulties in the forecasts of the daytime maximum and the nighttime minimum temperatures. The RMSE has a peak value at 1000 UTC in the forecasts started at 0300 UTC because the model greatly underestimated the daytime maximum surface temperatures (Fig. 10a). A part of the error might be caused by the asphalt sensor overheating too much during sunny days. In the forecasts started at 0900 UTC the RMSE is largest at the first

FIG. 10. RMSE values of road surface temperature forecasts as function of UTC time. The values are calculated for the spring period. Results are shown for forecasts started at (a) 0300, (b) 0900, (c) 1500, and (d) 2100 UTC. Only the cases where the RWS surface temperature was below 10°C were included in the forecast verification. The black line shows results for the control run without assimilated road surface temperatures, the red line for the control run with kriged and interpolated RWS temperatures, the blue line for the run using mobile observations, the brown line for the run using mobile observations with status-based correction, the pink line for the run using mobile observations with device-based correction, and the cyan line for the run using mobile observations with both status- and device-based correction.

forecast hours because the model had difficulties in the forecast of the daytime maximum values (Fig. 10b). The second control run has the smallest RMSE and the least negative bias values because the interpolated surface temperature values were warm enough to compensate for the otherwise too cold simulated temperatures. However, for the forecasts started at 1500 UTC the

second control run gives the worst verification results (Fig. 10c). The reason for this was the too slowly cooling surface temperature in the model during evening, as the interpolated values made the already too warm model run even warmer.

Utilizing mobile observations clearly reduces the RMSE values at the beginning of the forecasts when

FIG. 11. As in Fig. 10, but for forecast bias.

compared to the first control run. However, the effects of the initial surface temperature corrections vary between forecast starting times. Using mobile observations caused a warm bias at the start of the model run for forecasts started at 0300 and 2100 UTC (Figs. 11a,d). The correction methods reduce this bias and also improve the RMSE values. However, in a few hours the model run without a statistical correction has the lowest RMSE values for the forecasts started at 0300 UTC (Fig. 10a). The initial warm bias compensates for the otherwise too cool modeled temperatures during morning and daytime. The second control run gives slightly greater RMSE values than the runs using mobile observations with statistical correction methods at the beginning of the forecasts started at 0300 UTC. In addition, the bias values of the second

control run are slightly more negative. However, at the first forecast hour of the forecasts started at 2100 UTC the second control run gives clearly the smallest RMSE value (Fig. 10d). The bias value of the second control run is also nearest to zero of all the runs at the first forecast hour (Fig. 11d). Nonetheless, at the second forecast hour the RMSE value of the second control run is already at the same level as with the run using mobile observations with the status-based correction method.

The correction methods utilizing road status seem to somewhat increase the bias in the forecasts started at 1500 UTC (Fig. 11c), which was not seen in the verification results for the other forecast start times. These correction methods decrease the cold bias in warm temperatures according to Fig. 11. Indeed, the

20th October 2017 - 17th February 2018

FIG. 12. RMSE values of road surface temperature forecasts calculated separately for each station. The forecasts were for the time period from 20 October 2017 to 17 February 2018. Only those stations that had 100 or more forecast cases with mobile observations available within three hours before the start of the forecast are included. The total amount of forecast cases for the station are shown in the bottom-right corner of each panel. Coordinates of the stations are shown in Table 3. Only the cases where the RWS surface temperature was below 10°C were included in the forecast verification. The black line shows results for the control run without assimilated road surface temperatures, the red line for the control run with kriged and interpolated RWS temperatures, the blue line for the run using mobile observations, the brown line for the run using mobile observations with status-based correction, the pink line for the run using mobile observations with device-based correction, and the cyan line for the run using mobile observations with both status- and device-based correction.

correction methods bring the surface temperature closer to the observed values in the initialization phase of the 1500 UTC model run. However, the warmer temperatures in the initialization also cause the forecast part of the model run to be warmer. This increases the warm bias in the model as the simulated temperature cools down too slowly during the evening as mentioned above. However, the bias difference between the control run and the runs using road status in the correction method

was not statistically significant with 95% confidence (Table A10). (The statistical significance between verification scores was calculated with the bootstrap method, and the results can be found in the appendix in Tables A4–A11.)

e. Forecast verification for separate stations

Verification results presented in Figs. 9–11 are calculated over all stations. The results for individual

20th October 2017 - 17th February 2018

FIG. 13. As in Fig. 12, but for forecast bias.

stations vary due to geographical differences between station locations. As an example, Figs. 12 and 13 show RMSE and bias values calculated for separate stations during the winter period. Only those stations that had 100 or more forecast cases assimilating a mobile observation are included in the figures. The control run that did not assimilate any surface temperature observations gives the largest RMSE values at all stations, but the absolute values and the relative behavior compared to the other model runs vary between stations. The differences are biggest at Nupuri station, where the control run not using the surface temperature observations gives about 0.5°C higher RMSE values than the other runs at the first forecast hour. The control run without assimilated road

surface temperatures had also strong negative bias at the start of the forecast for the Nupuri station. The Nupuri station is located on southeast–northwest-oriented motorway with two-lane roads in both directions. There is some forested area on both sides of the motorway, which has a warming effect during wintertime as the trees prevent longwave radiation emittance from the road to the atmosphere. The surface temperature at the station was often among the warmest of the stations used during the winter period. The FMI RWM assumes open-sky conditions, which is one reason for the cold bias in the forecasts. Assimilating mobile observations in the model reduced the cold bias and RMSE values for the Nupuri station.

TABLE A1. Explanations of model run names used in Tables A2–A11.

Control 1	Control run without mobile observations
Control 2	Control run with interpolated road surface temperature values
Mobile obs	Run using mobile observations
Status	Run using mobile observations with the status-based correction method
Device	Run using mobile observations with the device-based correction method
Stat+Dev	Run using mobile observations with the status- and device-based correction method

The model runs assimilating mobile observations without statistical correction gave the most positive bias values at most of the stations. However, at Tupuri station the bias of the control run that did not assimilate mobile observations is a little higher. The surface temperature measurements at the Tupuri station were often among the coldest when compared to the other stations included in the study during the winter period. The station is not located on the main motorway but on an older motorway leading from Helsinki to Turku. The smaller traffic amount might be one reason for the relative coldness of the station. Although the Teconer observations had a warm bias, the model run assimilating them without statistical correction still gave slightly smaller bias values than the control run without observations.

The relative RMSE and bias differences between the three model runs assimilating mobile observations with statistical correction vary between stations. At many stations the model run assimilating mobile observation with the status-based correction gives higher RSME values than the other two. It also has more negative bias at most of the stations. The relative accuracy of these model runs and model run assimilating the kriged surface temperature values also varies between stations.

The results for the spring period should be analyzed separately for each forecast start time as the daily temperature variation has a strong effect on the verification results. However, when verification results for individual stations were calculated for each forecast hour, the total amount of forecast cases at each station was rather low. Not including the cases where the RWS surface temperature was above 10°C also affected the total number of cases. Because of the small sample size, the results are not considered reliable enough to be analyzed further in this paper.

5. Discussion and conclusions

Vehicle-based observations provide a cost-effective way for road condition monitoring on a dense spatial

TABLE A2. Results of the statistical significance test for model RMSE differences during the midwinter period computed with the bootstrap method. Each row represents a different model run combination, and the columns represent forecast lead times. If the difference between the RMSE values of two model runs was found to be statistically significant with 95% confidence, the corresponding square is marked with “x”. The model run names are explained in Table A1.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x	x	x	x	x	x
Control 1–Mobile obs	x	x	x	x	x	x	x	x	x	x
Control 1–Status	x	x	x	x	x	x	x	x	x	x
Control 1–Device	x	x	x	x	x	x	x	x	x	x
Control 1–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Control 2–Mobile obs	x	x	x	x	x	x				
Control 2–Status	x				x				x	x
Control 2–Device					x	x				
Control 2–Stat+Dev			x	x	x	x	x		x	x
Mobile obs–Status	x	x	x	x	x	x	x	x	x	x
Mobile obs–Device	x	x	x	x	x	x	x	x	x	x
Mobile obs–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Status–Device	x	x								
Status–Stat+Dev	x	x	x							
Device–Stat+Dev	x									

scale. Observations of initial road conditions are important, because even small variations can change the road condition forecast especially when the road surface temperature is near the freezing point. However, the quality of mobile observations and their effect on the forecast accuracy must be studied before operational implementation.

The effect of mobile observations on road weather forecasts was studied by including mobile surface temperature observations in the road weather model initialization. Mobile observations improved the accuracy of the road surface temperature forecasts when compared to the forecast scenario in which there would not be road weather stations in the area. Adding statistical correction to the measurements increased the accuracy further during the winter period. However, there was no clear improvement for the winter period when the results were compared to the forecast scenario with a dense RWS network. The relative differences between runs had more variation during the spring period when the daily surface temperature variation was stronger. The best performing statistical correction method of the three tested methods seemed to be the one that was dependent on the road status and the individual Teconer device when the corrected values were compared directly to the RWS observations. However, a clear winner among the correction methods could not be found when the forecast verification results were compared with each other, as the effects of the correction on the verification results are complex. For the winter period, the

TABLE A3. As in Table A2, but for bias values.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x	x	x	x	x	x
Control 1–Mobile obs	x	x	x	x	x	x	x	x	x	x
Control 1–Status	x	x	x	x	x	x	x	x	x	x
Control 1–Device	x	x	x	x	x	x	x	x	x	x
Control 1–Stat+Dev										
Control 2–Mobile obs	x	x	x	x	x	x	x	x	x	x
Control 2–Status										
Control 2–Device	x	x	x	x	x	x	x	x	x	x
Control 2–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Mobile obs–Status	x	x	x	x	x	x	x	x	x	x
Mobile obs–Device	x	x	x	x	x	x	x	x	x	x
Mobile obs–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Status–Device	x	x	x	x	x	x	x	x	x	x
Status–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Device–Stat+Dev	x	x	x	x	x	x	x	x	x	x

error in statistically corrected mobile observations and kriged values was about the same order of magnitude as the forecast error. However, the surface temperature observation error in these cases was not found to be the dominant factor in the forecast error but the other modeling errors seemed to have a strong effect on the forecast accuracy.

The verification results were affected by the asphalt sensor overheating, which may have caused too pessimistic results during the spring period. In addition, using the average of mobile measurements done within 50 m of the station does not always correspond well to the RWS measurements done at a certain spot. The mobile measurements might have been done for example in a different lane, and the road surface temperature can vary within this distance especially if the road is partially in shadow. This might have a negative effect on the accuracy of the model runs assimilating mobile observations, as the

TABLE A5. As in Table A2, but for RMSE values of forecasts started at 0900 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x	x	x		x	x
Control 1–Mobile obs	x		x	x				x		
Control 1–Status			x						x	x
Control 1–Device			x					x		x
Control 1–Stat+Dev			x						x	x
Control 2–Mobile obs		x	x			x	x	x	x	x
Control 2–Status	x	x	x	x	x	x	x		x	x
Control 2–Device	x	x	x		x	x	x	x	x	x
Control 2–Stat+Dev	x	x	x	x	x	x	x		x	x
Mobile obs–Status	x	x	x	x					x	x
Mobile obs–Device	x	x			x	x			x	x
Mobile obs–Stat+Dev	x	x							x	x
Status–Device				x			x		x	x
Status–Stat+Dev			x							
Device–Stat+Dev							x		x	x

RWS observations are used in the verification. The results of this study still provide useful information about the effects of the mobile observations and correction methods used on the road surface temperature forecasts.

An interesting topic for further studies would be to test how many stations are needed to get the same forecast accuracy as with mobile observations. This could be done by performing the road surface temperature kriging analysis with different amounts of observation points and using the values in RWM initialization. The studied area had only small elevation variations, and the effect of the mobile observations would probably have been greater in the areas with a more undulating landscape. In addition, the verification was performed only on the road weather station points. Along the roads there can be points, such as bridges, which are more prone to freezing and would be more easily monitored with mobile measurements.

TABLE A4. As in Table A2, but for RMSE values of forecasts started at 0300 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x					
Control 1–Mobile obs	x	x	x	x	x	x	x	x	x	x
Control 1–Status	x	x	x	x	x	x	x		x	
Control 1–Device	x	x	x	x	x	x		x		x
Control 1–Stat+Dev	x	x	x	x	x	x			x	
Control 2–Mobile obs	x		x	x	x	x	x	x	x	x
Control 2–Status		x	x	x	x	x	x			
Control 2–Device		x	x	x	x	x				
Control 2–Stat+Dev		x	x	x	x	x			x	
Mobile obs–Status	x		x	x	x	x	x	x	x	x
Mobile obs–Device	x	x	x	x	x	x	x	x	x	x
Mobile obs–Stat+Dev	x		x	x	x	x	x	x	x	x
Status–Device										
Status–Stat+Dev										
Device–Stat+Dev										

TABLE A6. As in Table A2, but for RMSE values of forecasts started at 1500 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x					x	x	x
Control 1–Mobile obs		x								
Control 1–Status										
Control 1–Device		x	x	x	x					
Control 1–Stat+Dev										
Control 2–Mobile obs	x	x	x	x	x	x	x	x	x	x
Control 2–Status				x	x	x	x	x	x	x
Control 2–Device	x	x	x	x	x	x	x	x	x	x
Control 2–Stat+Dev		x	x	x	x	x	x	x	x	x
Mobile obs–Status	x	x	x							
Mobile obs–Device			x	x	x	x	x			
Mobile obs–Stat+Dev	x	x								
Status–Device	x	x	x	x	x		x	x	x	x
Status–Stat+Dev		x	x		x	x				
Device–Stat+Dev	x	x	x	x						

TABLE A7. As in Table A2, but for RMSE values of forecasts started at 2100 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x				x	x
Control 1–Mobile obs		x	x					x	x	x
Control 1–Status	x	x	x	x				x	x	x
Control 1–Device		x	x	x				x	x	x
Control 1–Stat+Dev		x	x	x				x	x	x
Control 2–Mobile obs	x	x				x	x		x	x
Control 2–Status	x								x	x
Control 2–Device						x			x	x
Control 2–Stat+Dev	x				x	x	x		x	x
Mobile obs–Status	x	x	x	x						x
Mobile obs–Device	x								x	x
Mobile obs–Stat+Dev										x
Status–Device			x	x						
Status–Stat+Dev	x	x	x	x	x	x	x			
Device–Stat+Dev									x	

An interesting topic for future research would also be to repeat the study in an area with only a few road weather stations such as the northern part of Finland. This study could utilize measurements from a truck fleet that has been equipped with measurement devices within the EU-funded Intelligent Arctic Trucks project (Sukuvaara et al. 2018).

The mobile surface temperature observations included in this study were done with just one sensor type. However, as the amount of available mobile observations increases, further studies should be made to assess their effect on the road weather forecasts. Different kinds of statistical correction methods will be needed as other sensors might behave differently, and road status measurements are not always available. In addition, many vehicles do not measure road surface temperature.

TABLE A8. As in Table A2, but for bias values of forecasts started at 0300 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2										
Control 1–Car obs	x	x	x	x	x	x	x	x	x	x
Control 1–Status	x	x	x	x	x	x				
Control 1–Device	x	x	x	x	x	x	x	x	x	x
Control 1–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Control 2–Car obs	x	x	x	x	x	x	x	x	x	x
Control 2–Status	x	x	x	x	x	x	x	x	x	
Control 2–Device	x	x	x	x	x	x	x	x	x	x
Control 2–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Car obs–Status	x	x	x	x	x	x	x	x	x	x
Car obs–Device	x	x	x	x	x	x	x	x	x	x
Car obs–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Status–Device	x	x	x	x	x	x	x	x		
Status–Stat+Dev	x	x	x	x	x	x	x			
Device–Stat+Dev	x	x	x	x	x					

TABLE A9. As in Table A2, but for bias values of forecasts started at 0900 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x	x	x	x	x	x
Control 1–Car obs	x	x	x	x	x	x	x			
Control 1–Status		x	x	x	x	x	x	x	x	x
Control 1–Device	x	x	x	x	x	x			x	x
Control 1–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Control 2–Car obs	x	x	x	x	x	x	x	x	x	x
Control 2–Status	x	x	x	x	x	x	x	x	x	x
Control 2–Device	x	x	x	x	x	x	x	x	x	x
Control 2–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Car obs–Status	x	x	x	x	x			x	x	x
Car obs–Device	x					x			x	x
Car obs–Stat+Dev	x						x	x	x	x
Status–Device	x	x	x	x	x		x	x	x	x
Status–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Device–Stat+Dev						x	x	x	x	x

It needs to be studied further if more commonly measured variables, like air temperature, can be used to improve road weather forecast accuracy. Air temperature measurements cannot be used in the model in the same way as road surface temperature observations were used in this study to adjust the radiation fluxes. However, they might be more usable when implemented in 3D numerical weather prediction models. There is already an ongoing project that aims to implement vehicle observations in a regional operational weather prediction model in Germany (Riede et al. 2018).

Acknowledgments. We express our gratitude to Prof. Heikki Järvinen (University of Helsinki), Dr. Timo Sukuvaara (FMI), and M.Sc. Petteri Karsisto (FMI) for their valuable comments and suggestions. The Finnish

TABLE A10. As in Table A2, but for bias values of forecasts started at 1500 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x	x	x	x	x	x
Control 1–Car obs				x	x	x	x	x	x	x
Control 1–Status		x	x	x	x	x	x	x	x	x
Control 1–Device				x	x	x	x	x	x	x
Control 1–Stat+Dev		x	x	x	x	x	x	x	x	x
Control 2–Car obs	x	x	x	x	x	x	x	x	x	x
Control 2–Status	x	x	x	x	x	x	x	x	x	x
Control 2–Device	x	x	x	x	x	x	x	x	x	x
Control 2–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Car obs–Status	x	x	x	x	x	x	x	x	x	x
Car obs–Device			x	x	x	x	x	x	x	x
Car obs–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Status–Device	x	x	x	x	x	x	x	x	x	x
Status–Stat+Dev	x									
Device–Stat+Dev	x	x	x	x	x	x	x	x	x	x

TABLE A11. As in Table A2, but for bias values of forecasts started at 2100 UTC during the spring period.

	1	2	3	4	5	6	7	8	9	10
Control 1–Control 2	x	x	x	x	x		x	x	x	
Control 1–Car obs	x	x	x	x	x	x	x	x	x	x
Control 1–Status	x	x	x	x	x	x	x	x	x	x
Control 1–Device	x	x	x	x	x	x	x	x	x	x
Control 1–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Control 2–Car obs	x	x	x	x	x	x	x	x	x	x
Control 2–Status	x	x	x	x	x	x	x	x	x	x
Control 2–Device	x	x	x	x	x	x	x	x	x	x
Control 2–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Car obs–Status	x	x	x	x	x	x	x	x	x	x
Car obs–Device	x	x	x	x	x	x	x	x	x	x
Car obs–Stat+Dev	x	x	x	x	x	x	x	x	x	x
Status–Device										x
Status–Stat+Dev	x	x	x	x	x	x	x	x		
Device–Stat+Dev	x	x	x	x	x	x	x	x		

Meteorological institute and the University of Oulu have participated in several research projects that supported this work and the authors gratefully acknowledge the support provided. The projects were Intelligent Arctic trucks and Sod5G funded by European Regional Development Fund of the European Union and Regional Council of Lapland, WiRma funded by Interreg Nord fund of European Union and Regional Council of Lapland, 5G-Safe funded by Business Finland, 6Genesis Flagship funded by Academy of Finland, and the AI Enhanced Mobile Edge Computing project funded by the Future Makers program of Jane and Aatos Erkko Foundation and Technology Industries of Finland Centennial Foundation. We want to thank also all financiers and partners who participated in these projects.

APPENDIX

Statistical Significance of the Model Bias and RMSE Differences

The results of statistical significance tests performed with the bootstrap method are presented in this appendix. Table A1 provides explanations for the model run names used in Tables A2–A11. Tables A2 and A3 show the results of RSME and bias differences for each model run pair during the winter period. Tables A4–A11 show the corresponding results for the spring period, except the results are separated by the forecast start time.

REFERENCES

Aalto, J., P. Pirinen, J. Heikkinen, and A. Venäläinen, 2013: Spatial interpolation of monthly climate data for Finland: Comparing

- the performance of kriging and generalized additive models. *Theor. Appl. Climatol.*, **112**, 99–111, <https://doi.org/10.1007/s00704-012-0716-9>.
- Allegretti, M., and S. Bertoldo, 2014: Cars as a diffuse network of road-environment monitoring nodes. *Wireless Sensor Network*, **6**, 184–191, <https://doi.org/10.4236/wsn.2014.69018>.
- Bengtsson, L., and Coauthors, 2017: The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system. *Mon. Wea. Rev.*, **145**, 1919–1935, <https://doi.org/10.1175/MWR-D-16-0417.1>.
- Bogren, J., T. Gustavsson, M. Karlsson, and U. Postgård, 2000: The impact of screening on road surface temperature. *Meteor. Appl.*, **7**, 97–104, <https://doi.org/10.1017/S135048270000150X>.
- Brutsaert, W., 1984: *Evaporation into the Atmosphere: Theory, History, and Applications*. D. Reidel Publishing Company, 299 pp.
- Chapman, L., and J. E. Thornes, 2006: A geomatics-based road surface temperature prediction model. *Sci. Total Environ.*, **360**, 68–80, <https://doi.org/10.1016/j.scitotenv.2005.08.025>.
- Cressie, N., 1993: *Statistics for Spatial Data*. John Wiley & Sons, 900 pp.
- Crevier, L. P., and Y. Delage, 2001: A new model for road-condition forecasting in Canada. *J. Appl. Meteor.*, **40**, 2026–2037, [https://doi.org/10.1175/1520-0450\(2001\)040<2026:MANMFR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<2026:MANMFR>2.0.CO;2).
- Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman and Hall, 436 pp.
- Gustavsson, T., 1999: Thermal mapping—A technique for road climatological studies. *Meteor. Appl.*, **6**, 385–394, <https://doi.org/10.1017/S1350482799001334>.
- Haavasoja, T., J. Nylander, and P. Nylander, 2012: Experiences of Mobile Road Condition Monitoring. *16th Int. Road Weather Conf. (SIRWEC)*, Helsinki, Finland, SIRWEC, http://www.teconer.fi/downloads/SIRWEC_2012_Experiences%20of%20Mobile%20RCM_ID14.pdf.
- Hogan, R. J., and I. B. Mason, 2011: Deterministic forecasts of binary events. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd ed. I. T. Jolliffe and D. B. Stephenson, Eds., John Wiley and Sons, 31–59, <https://doi.org/10.1002/9781119960003.ch3>.
- Kangas, M., M. Heikinheimo, and M. Hippä, 2015: RoadSurf—A modelling system for predicting road weather and road surface conditions. *Meteor. Appl.*, **22**, 544–533, <https://doi.org/10.1002/met.1486>.
- Karsisto V., 2018: Road surface temperature forecast study HKI-TKU 0708. Zenodo.org, accessed 25 September 2018, <https://doi.org/10.5281/zenodo.1434636>.
- , P. Nurmi, M. Kangas, M. Hippä, C. Fortelius, S. Niemelä, and H. Järvinen, 2016: Improving road weather model forecasts by adjusting the radiation input. *Meteor. Appl.*, **23**, 503–513, <https://doi.org/10.1002/met.1574>.
- , S. Tijm, and P. Nurmi, 2017: Comparing the performance of two road weather models in the Netherlands. *Wea. Forecasting*, **32**, 991–1006, <https://doi.org/10.1175/WAF-D-16-0158.1>.
- Lövén, L., V. Karsisto, H. Järvinen, M. J. Sillanpää, T. Leppänen, E. Peltonen, S. Piirtikangas, and J. Riekkä, 2019: Mobile road weather sensor calibration by sensor fusion and linear mixed models. *PLoS ONE*, **14**, e0211702, <https://doi.org/10.1371/journal.pone.0211702>.
- Patankar, S. V., 1980: *Numerical Heat Transfer and Fluid Flow*. McGraw-Hill, 197 pp.
- Petty, K. R., and W. P. Mahoney, 2007: Weather applications and products enabled through vehicle infrastructure integration (VII). U.S. Department of Transportation, 124 pp.

- Riede, H., A. Bouras, Z. Paschalidi, J. W. Acevedo-Valenciaa, T. Kratzsch, and J. Nachtigall, 2018: Flowkar: Using high-resolution data from vehicle sensors to improve operational weather products. *19th Conf. of Standing International Road Weather Commission (SIRWEC)*, Smolenice, Slovakia, SIRWEC, 95–96.
- Robinson, G. K., 1991: That BLUP is a good think: The estimation of random effects. *Stat. Sci.*, **6**, 15–51, <https://doi.org/10.1214/ss/1177011926>.
- Shao, J., P. J. Lister, G. D. Hart, and H. B. Pearson, 1996: Thermal mapping: Reliability and repeatability. *Meteor. Appl.*, **3**, 325–330, <https://doi.org/10.1002/met.5060030405>.
- , J. C. Swanson, R. Patterson, P. J. Lister, and A. N. McDonald, 1997: Variation of winter road surface temperature due to topography and application of thermal mapping. *Meteor. Appl.*, **4**, 131–137, <https://doi.org/10.1017/S135048279700042X>.
- Sukuvaara, T., K. Mäenpää, S. Kantomaa, D. Stepanova, M. Hippi, and V. Karsisto, 2018: Intelligent traffic enabled advanced road weather infrastructures in Arctic conditions. *European Geosciences Union General Assembly 2018*, Vol. 20, EGU2018-95, Vienna, Austria, European Geosciences Union, <https://meetingorganizer.copernicus.org/EGU2018/EGU2018-95.pdf>.
- Thornes, J. E., 1991: Thermal mapping and road-weather information systems for highway engineers. *Highway Meteorology*, A. H. Perry and L. J. Symons, Eds., E & FN Spon, 39–67.
- Vaisala, 2001: Vaisala Road/runway surface and depth sensor DRS511. Tech. Note, 2 pp., <https://www.vaisala.com/sites/default/files/documents/WEA-RDS-G-DRS511%20Datasheet.pdf>.
- West, B. T., K. B. Welch, and A. T. Galecki, 2007: *Linear Mixed Models: A Practical Guide Using Statistical Software*. Chapman & Hall/CRC, 374 pp., <https://doi.org/10.1002/sim.3167>.