

Research Article

Methodology to Forecast Road Surface Temperature with Principal Components Analysis and Partial Least-Square Regression: Application to an Urban Configuration

M. Marchetti,¹ A. Khalifa,^{1,2,3} and M. Bues³

¹Cerema, DTer Est, ERA 31, 54510 Tomblaine, France

²IFSTTAR, Centre de Nantes, route de Bouaye, CS4, 44344 Bouguenais Cedex, France

³UMR GeoRessources, Université de Lorraine, 54000 Nancy, France

Correspondence should be addressed to M. Marchetti; mario.marchetti@cerema.fr

Received 13 March 2015; Revised 9 May 2015; Accepted 21 May 2015

Academic Editor: Stefano Dietrich

Copyright © 2015 M. Marchetti et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A forecast road surface temperature (RST) helps winter services to optimize costs and to reduce the deicers environmental impacts. Data from road weather information systems (RWIS) and thermal mapping are considered inputs for forecasting physical numerical models. Statistical models include many meteorological parameters along routes and provide a spatial approach. It is based on typical combinations resulting from treatment and analysis of a database from measurements of road weather stations or thermal mapping, easy, reliable, and cost effective to monitor RST, and many meteorological parameters. A forecast dedicated to road networks should combine both spatial and time forecasts needs. This study contributed to building a reliable RST forecast based on principal component analysis (PCA) and partial least-square (PLS) regression. An urban stretch with various weather conditions and seasons was monitored over several months to generate an appropriate number of samples. The study first consisted of the identification of its optimum number to establish a reliable forecast. A second aspect is aimed at comparing RST forecasts from PLS model to measurements. Comparison indicated a forecast over an urban stretch with up to 94% of values within $\pm 1^\circ\text{C}$ and over 80% within $\pm 3^\circ\text{C}$.

1. Context and Introduction

The objective of winter maintenance is to ensure the access and the use of infrastructures in adverse weather conditions. It mainly consists in avoiding the occurrence of ice for temperatures close to freezing and the persistence of snow. The decision to treat or not a road network is often difficult and is eased by road weather outstations and forecasts, commonly in place since the 80s. These stations do provide measurements of atmospheric and road parameters such as road surface temperature (RST) to inform road managers about network status. Forecasts are obtained with numerical models either sites specific or for a whole network. A range of forecast models exists, usually site specific, and since the 1980s a significant body of literature has accrued (see [1] for a thorough review). Furthermore, because of RST variations up to 10°C , forecast interpolation is also required. This has

traditionally been achieved via thermal mapping, but of all the components contained within road weather information systems (RWIS), this interpolation has frequently been identified as the least satisfactory [2].

The thermal mapping consists of a high-resolution survey of RST by means of infrared thermometry. Measurements are conducted in various winter atmospheric conditions, and providing thermal fingerprints, used as inputs for forecast models [3–9]. These fingerprints are also used as a tool for the optimisation of routes for anti-icing [10] or to identify locations for the installation of road weather outstations.

The extent of RST variations along a route is controlled by atmospheric stability and local variations of geographic parameters and of traffic (profile and density) [3, 8]. The greatest variations along a route are observed during these stable anticyclonic conditions [3]. With decreasing atmospheric stability, the amplitude of the thermal fingerprint



FIGURE 1: Thermal mapping route (black line) (a) and illustration of its urban configuration (b).

subsequently reduces. Some authors [5] have shown that under certain weather conditions some consistency appears in the spatial variation of RST along a route, which enables thermal mapping surveys under only a few selected weather conditions (usually extreme, intermediate, and damped) [1]. A standard of five to six surveys was selected typically commissioned to provide coverage of the conditions encountered in a winter season. This approach is not optimum, and results in daily forecasts are too often “pigeon-holed” into one of the categories when used operationally [8].

However, measurement campaigns are very time-consuming. The measurement has to be partitioned into itinerary stretches done in the open time window. Some gradual change appeared in practice, moving toward a new spatial modelling based approach. Route-based forecasts take into account both meteorological and geographical data to provide a high-resolution forecast of RST and condition around the road network [2]. Besides potential significant improvement in the forecasts, it has also brought about a new set of challenges. Whereas traditional site-specific forecasts could be easily validated against sensor data from outstations located at the forecast sites, this is clearly impossible for a route-based forecast [1] where thermal mapping is still required for the verification of the spatial forecast. However, this approach is not cost effective to provide data at a high temporal resolution. Route-based forecasts can still only rely on isolated thermal fingerprints.

A forecast dedicated to road networks is the conjunction of both spatial and time forecasts needs of a vision of a full network at once. Thermal mapping is an easy, reliable, and cost effective way to monitor RST and many other meteorological parameters along routes. A numerical model will provide a temporal forecast on specific locations, and many computations will be necessary to build a forecast covering the full route or network. A new approach based on statistics was developed by Chapman et al. [11, 12] and provided RST over a full given route only through principal components analysis (PCA) but did not yet provide a RST forecast as it is usually understood, and it only relies on RST measurements which are not always available. To do so, partial least-square (PLS) regression could indeed contribute to such a goal. A similar study based on these multivariate

data analysis tools was performed by Kršmanc et al. [13] but only on specific locations.

This study will be divided in two parts. The first one will consist of a description of a RST forecast on a specific urban location based on PLS. The variables consisted in air temperature obtained from measurements with the thermal mapping vehicle, along with other meteorological and anthropogenic data. This will contribute and therefore provide a more comprehensive picture of RST variation for a broader range of atmospheric stability than that conventionally considered by such surveys. A second part will be dedicated to a comparison between forecast results and field data. Then a new approach will be described about the way a full RST profile over a route is elaborated from a T_{air} measurement or forecast of one specific location, which can have a greater reliability than RST ones. Results will be provided and discussed.

2. Methodology

2.1. Equipment, Route Investigated. DTer Est in Nancy (France) has developed and used a thermal mapping vehicle from the beginning of the existence of the technique. Along the years, a substantial quantity of thermal data for analysis has been accumulated. For this investigation, a whole route going around and within this town was chosen (Figure 1(a)). Then a stretch of a French urban street in Nancy (France) was selected for detailed analysis on RST. It is roughly several hundred meters long with a typical urban canyon configuration (Figure 1(b)). The thermal dataset is available for this route contains thirty-four of thermal fingerprints, each considered a sample, obtained under extreme and intermediate weather conditions over years 2012 and 2013, at least once a month.

RST was obtained by using an infrared radiometer operating in the $7\text{--}14\ \mu\text{m}$ spectral range, for a range from -30°C to $+70^\circ\text{C}$, an uncertainty of 2°C (at 23°C), fitted to the bottom of a survey vehicle, and located in a compartment at controlled temperature. The instrument measures the energy flux density emitted by the surface, and RST is calculated through the Stefan-Boltzmann equation [14]. The road surface is considered to be a grey body and as such emissivity is held constant at 0.95 [15, 16]. Air temperature (-40°C to

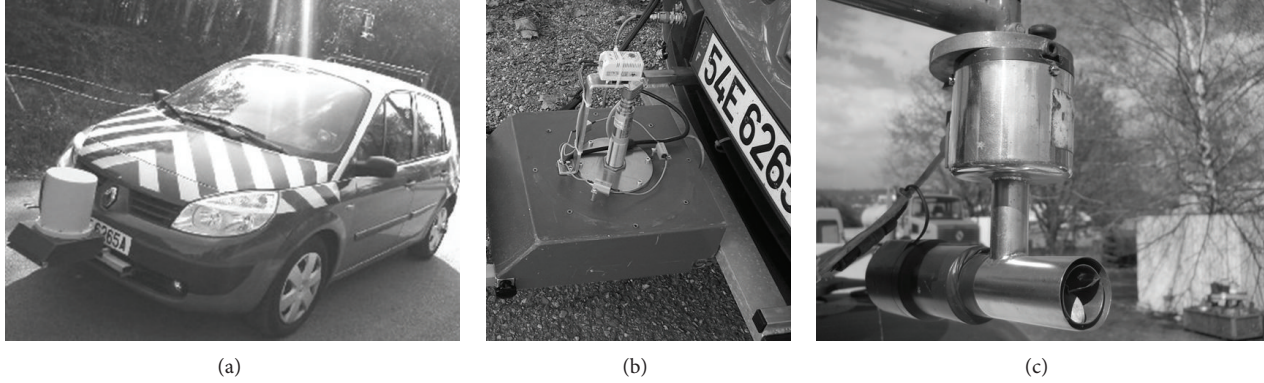


FIGURE 2: French thermal mapping vehicles and associated instruments (infrared radiometer (b) and atmospheric probe (c)).

+125°C, $\pm 0.3^\circ\text{C}$ uncertainty) and relative humidity (0–100% range, $\pm 2\%$ uncertainty) data are also obtained through an atmospheric probe located on the roof of the vehicle. An electrical turbine generated a laminar flow circulating over the sensors to improve measurements accuracy and avoid turbulences artefacts. Figure 2 illustrates the French thermal mapping vehicle with on-board measurements devices.

2.2. Partial Least-Square Regression. PLS is a statistical method that enables data reduction when dealing with large datasets [17]. It consisted in computing a calibration using a training set between air temperature matrix T_{air} and another one of variables. As pointed out by Almkvist et al. [18, 19], the ground heat flux is one variable affecting RST. This work detailed in this paper focused on a selection of variables easy to obtain, and that therefore already included this ground heat flux. It includes RST at the chosen urban location, relative humidity, wind force global radiation, nebulosity, the existence of precipitations within the previous 24 hours, the moment of the day (before midday a.m., or after midday p.m.), and the traffic intensity (weak, moderate, and high) as an anthropogenic variable. Traffic data was obtained from the appropriate service and each qualitative appreciation corresponds to a range of hours within the day. A weak traffic occurs during the night, a moderate one roughly between 10:00 a.m. and 4:00 p.m., and the highest appears during peak hours such as the ones of 8:00 a.m. and 6:00 p.m. Although RST for winter months is around 0°C , this situation contributes to make difficult the decision to apply deicers with respect to very negative RST values, and this is the reason why they were studied. This training set contributes to the generation of a statistical forecast model of RST with T_{air} . Since noise is contained in both variables and T_{air} data, this computation will conduct to the generation of a calibration matrix P and a minimized error matrix E , to obtain the relationship:

$$\text{Variables} = P \cdot T_{\text{air}} + E. \quad (1)$$

In PLS, a set of basis vectors is searched for variables and another one for T_{air} , and the understanding of how these sets are related is necessary. The eigenvectors for each of these two spaces are calculated to restore an optimum congruence

between each variable factor and its corresponding T_{air} in the least-squares sense. Since some noise is both present on variables and T_{air} measurements, eigenvectors calculated for both spaces are shifted by different amounts in different directions, due to the independence of noises in these spaces, destroying the perfect congruence between variables and T_{air} data points. PLS will try to restore the optimum congruence, defined as a perfectly linear relationship between the projections, or scores, of the variables and T_{air} data onto their corresponding factors.

In the case of thermal mapping, each thermal fingerprint is considered to be a sample, and each variable as a data point in a multidimensional space. The variables are some of the physical parameters potentially included in a numerical model and affecting RST and as explained by Hammond et al. [1, 12]. By using the data from several thermal surveys, a data matrix is generated which can then be assimilated into clusters of points in this multidimensional space. With respect to thermal mapping, PLS will determine how many factors are needed to properly forecast RST with T_{air} . Variables and T_{air} matrices, respectively, are 34 lines (one line per thermal fingerprint) per 8 columns (one column per variable) (Table 1) and 34 per 28 columns (one column per data point over the distance of the urban stretch). Once these factors are identified, they can then be used to build a forecast model for other weather conditions provided they are not significantly different from the ones used for the original PLS calculations. The first objective in this PLS approach is to identify and then to prescribe the appropriate number of thermal mapping runs required to forecast an accurate daily temperature pattern along a given route. It is then to establish the forecast thermal fingerprints, thus providing a simple spatial forecasting model, based on a cost effective and realistic data.

2.3. Combination of PCA, PLS, and RWIS Data/Weather Services Forecast for a Spatial and Temporal Forecast. Marchetti et al. [12] gave extensive details about the use of PCA to thermal mapping data and RST in particular. In this study, the same methodology was applied to T_{air} matrix. Each thermal fingerprint is assimilated to a sample. The matrix contains as many lines as there are samples and as many columns as there

TABLE 1: List of variables used for PLS calculations.

| Sample date | RST (°C) | Measurement time (a.m./p.m.) | Traffic (weak, moderate, and high) | Wind speed (km/h) | Relative humidity (%) | Global radiation (J/cm ²) | Nebulosity (octas) |
|-------------|-------------|------------------------------------|--|----------------------|-----------------------------|--|-----------------------|
| 2012-03-07 | 9.8 | p.m. | Moderate | 50.8 | 55 | 66 | 8 |
| 2012-03-15 | 14.4 | p.m. | Moderate | 10.8 | 29 | 170 | 1 |
| 2012-03-21 | 14.2 | p.m. | Moderate | 35.3 | 36 | 176 | 0 |
| 2012-03-29 | 22.4 | p.m. | Moderate | 40 | 33 | 137 | 7 |
| 2012-04-03 | 22.1 | p.m. | Moderate | 43.9 | 55 | 183 | 5 |
| 2012-04-30 | 24.3 | p.m. | Moderate | 26.3 | 51 | 83 | 8 |
| 2012-05-10 | 33.5 | p.m. | Moderate | 45.4 | 39 | 233 | 3 |
| 2012-05-15 | 11.3 | a.m. | Weak | 19.4 | 64 | 25 | 7 |
| 2012-06-14 | 28.0 | a.m. | Moderate | 20.2 | 50 | 288 | 5 |
| 2012-06-28 | 40.5 | a.m. | Moderate | 41.8 | 50 | 265 | 7 |
| 2012-07-04 | 24.0 | a.m. | Weak | 11.5 | 83 | 0 | 7 |
| 2012-07-16 | 29.8 | p.m. | Moderate | 34.2 | 43 | 200 | 6 |
| 2012-08-23 | 34.8 | p.m. | Moderate | 22.3 | 32 | 206 | 4 |
| 2012-09-03 | 27.9 | p.m. | Moderate | 45 | 31 | 194 | 1 |
| 2012-09-04 | 14.8 | a.m. | Weak | 10.1 | 85 | 15 | 2 |
| 2012-09-12 | 19.3 | p.m. | Moderate | 26.3 | 57 | 72 | 7 |
| 2012-10-01 | 18.9 | p.m. | Moderate | 23.4 | 44 | 120 | 4 |
| 2012-10-15 | 11.6 | p.m. | Moderate | 21.6 | 60 | 62 | 5 |
| 2012-11-23 | 6.3 | a.m. | Moderate | 32 | 83 | 24 | 8 |
| 2012-12-11 | 0.4 | p.m. | Moderate | 28.1 | 68 | 15 | 7 |
| 2013-02-14 | -6.0 | a.m. | Moderate | 10.4 | 90 | 62 | 7 |
| 2013-02-20 | -0.6 | a.m. | Weak | 19.1 | 93 | 0 | 7 |
| 2013-02-22 | -5.4 | a.m. | High | 48.2 | 66 | 85 | 1 |
| 2013-03-14 | -4.6 | p.m. | Weak | 13.3 | 83 | 0 | 0 |
| 2013-03-22 | 0.8 | a.m. | High | 29.9 | 74 | 88 | 4 |
| 2013-04-18 | 20.2 | a.m. | Moderate | 37.4 | 69 | 156 | 7 |
| 2013-08-05 | 29.9 | p.m. | Weak | 19.8 | 57 | 0 | 8 |
| 2013-08-20 | 37.0 | a.m. | High | 6.8 | 78 | 26 | 1 |
| 2013-08-30 | 29.8 | p.m. | Moderate | 27 | 37 | 223 | 2 |
| 2013-08-30 | 22.8 | a.m. | Weak | 6.8 | 91 | 0 | 8 |
| 2013-09-24 | 15.0 | p.m. | Weak | 5.8 | 98 | 0 | 8 |
| 2013-10-24 | 14.6 | a.m. | Moderate | 11.5 | 85 | 37 | 7 |
| 2013-10-24 | 12.1 | p.m. | Weak | 8.6 | 100 | 0 | 8 |
| 2013-12-09 | 3.3 | p.m. | Moderate | 9 | 93 | 38 | 5 |

are measurement locations (in this case every 3 m). Again, by multiplying the scores and the loadings matrices and adding the average profile of all samples, T_{air} profiles of the urban stretch from PCA, $T_{\text{air_PCA}}$, were built, at least correcting an eventual offset error. The work consisted in checking the number of samples mandatory to perform PCA, considering again a target of four samples.

$T_{\text{air_PCA}}$ profile was then assimilated to a one-column matrix, where each element of the column is T_{air} at a point of the itinerary ($T_{\text{air_PCA},1}, T_{\text{air_PCA},2}, \dots, T_{\text{air_PCA},L}$), L being the final point of this given route. For two $T_{\text{air_PCA}}$ profiles, $T_{\text{air_PCA},1}$ and $T_{\text{air_PCA},2}$, an interpolated T_{air} profile, $T_{\text{air_PCA_interpolated}}$, using a coefficient k ranging between 0 and

1, will be then obtained by using (2), to denote continuation of form. The values of k are selected to obtain RST profiles over the full RST range investigated and thus to obtain a RST profile even if measurements were not made. Consider

$$T_{\text{air_PCA_interpolated}} = \begin{pmatrix} k \cdot T_{\text{air_PCA},1,1} + (1-k) \cdot T_{\text{air_PCA},2,1} \\ k \cdot T_{\text{air_PCA},1,2} + (1-k) \cdot T_{\text{air_PCA},2,2} \\ \vdots \\ k \cdot T_{\text{air_PCA},1,L} + (1-k) \cdot T_{\text{air_PCA},2,L} \end{pmatrix}. \quad (2)$$

One of the ideal inputs used to take decisions in winter maintenance is the knowledge of a RST profile all over the road network which a service is in charge of. Considering as impossible to obtain either RST measurements or forecast over these profiles, these winter services rely on site-specific RWIS outstations giving atmospheric and road parameters such as air temperature, relative humidity, and RST. They can also use numerical weather models outputs to provide a forecast for this specific spot, or more recently over the whole route using route-based forecasting techniques [8]. Marchetti et al. [12] have indicated a way through PCA to obtain the whole RST profile using the RST at a single outstation, but which is not a forecast as it is usually understood. Furthermore, there is some controversy about the proper way to make RST measurement. It is done either by a probe embedded into pavement, raising the question of its too local aspect, or by a radiometer as indicated above, this temperature being radiative and not thermodynamic. The T_{air} measurement is more accepted and does not suffer the same controversy from weather services. The idea is then to use either RWIS air temperature or forecast data point from a weather service at one single outstation and then to use coefficient k to match this point with the corresponding one in $T_{\text{air-PCA,interpolated}}$. Then, the whole T_{air} profile along the route is built. Once this one is available, a PLS model is applied to forecast the RST at the location for which the PLS model was developed. Then, using again the PCA analysis and the interpolation, a RST profile along this very same route could be extracted. The principle of this work is described in Figure 3. A comparison is then conducted between measured and forecast RST profiles.

3. Results of RST Forecast Based on PCA and PLS: Discussion

3.1. Search of the Optimum Number of Measurements Sets for the PLS for RST Forecast. PCA was conducted with Unscrambler X 10.1 package software on T_{air} datasets to identify potential specific thermal fingerprints. Results indicated that over 99% of the variance could be explained with the first principal component, indicating the data homogeneity. These results are consistent with a previous published study [12]. In the case of T_{air} , an average offset of -2.3°C was identified and this correction was then applied to T_{air} profiles calculated from PCA. Generic T_{air} profiles over this route using PCA outputs (scores and loadings) were then calculated and illustrated in Figure 4.

PLS was then performed on the urban stretch, considering, respectively, first the whole dataset and then only on some samples corresponding to temperatures below 10°C considered as representative of winter in the Nancy (France) area. The same package software as for PCA was then used. The idea is to determine the minimum number of samples to build a reliable statistical model to forecast RST with T_{air} . The reliability of such a model is established mainly through their R^2 and their RMSE. The conducted evaluations and results are summarized in Table 2. The number of outliers, which corresponds to data poorly described by the statistical

models, is very small with respect to the number of points used to elaborate the model, that is, the sample number multiplied by the 28 columns of T_{air} matrix. Results indicated that at least 6 thermal fingerprints are needed to develop a reliable model with R^2 greater than 0.90, with an optimum to 8 fingerprints. This figure is consistent with the actual practice in thermal mapping. Furthermore, this PLS approach is more adapted to short stretches, with R^2 of 0.97 and 34 samples, than to a whole and longer itinerary where R^2 is below this 0.90 threshold. The consideration of both R^2 and RMSE indicated that PLS approach is specifically adapted for road environment with vegetation, where RMSE is near or below 2°C .

Once these models were developed, their performances were evaluated. To do so, they were applied to two T_{air} profiles obtained from measurements done along the same urban stretch at two different dates, along with other variables presented in Table 1, to deduce and somehow to forecast corresponding RST. A comparison with available RST field data was then conducted. A specific focus was conducted on situations corresponding to winter weather conditions in the Nancy area ($\text{RST} < 10^{\circ}\text{C}$ or near this value). Results are detailed in Table 3. PLS is therefore able to provide a RST forecast based on T_{air} measurement with a bias below 1°C for situations close to water freezing and two factors. Too many factors generated a poor bias but the standard deviation was unchanged. The worst result was obtained with the warmest situation, although it is only due to a bias.

3.2. Combination of PLS and Weather Services Forecast for a Spatial and Temporal Forecast. Some improvements were obtained in the RST forecast with PLS use. So far, there is still the need of either a T_{air} profile or other variables to obtain RST. As mentioned above, winter services are usually taking decisions on the basis of RWIS data or on the forecast of atmospheric parameters among which is air temperature. In both cases, this piece of information is very local. In some situations, there are no available RWIS or their density over the road network is not appropriate, and road managers can only rely on a forecast from a weather service, with the inconvenient of a forecast available at given hours with several hours between each. Sometimes, some locations of the network have a specific thermal behaviour because of its configuration, and the provided forecast could greatly differ from observations. The objective here is then to evaluate the performance of a RST forecast PLS-based using T_{air} values from a RWIS or T_{air} forecasts from a weather service. With the methodology described in Figure 3, a full T_{air} profile is then built using results presented in Figure 4(b). This built profile is an input in the selected PLS model elaborated in Section 3.1 to obtain RST on a specific location. Once this RST value is available, a full thermal fingerprint is obtained on the basis of data presented in Figure 4(d). The challenge is then to determine to what extent a full RST profile could be elaborated with site-specific T_{air} data.

A forecast was then undertaken at two different dates, 2013-12-11 and 2013-03-22, for which RST and T_{air} data were available along the studied urban stretch. It was obtained from Météo France weather station located in the immediate

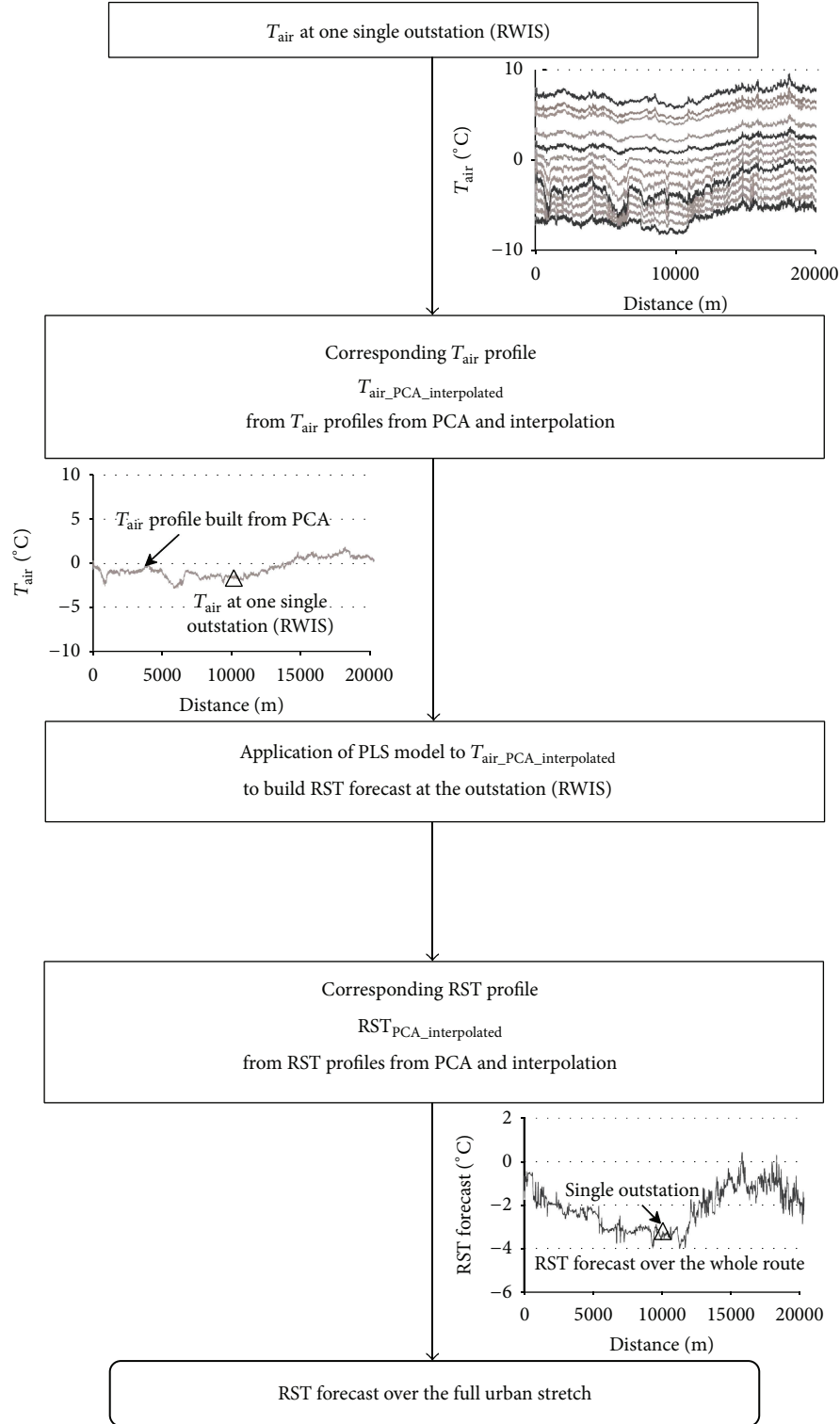


FIGURE 3: Description of the principle to forecast RST from T_{air} with PCA and PLS.

vicinity of the departure point of the urban location (Figure 1). Météo France data was provided every three hours, from 0:00 a.m. to 9:00 p.m. GMT. Since thermal fingerprints did not exactly start at the same time, Météo France data was interpolated to calculate a corresponding T_{air} . One has

to consider that this piece of information is sometimes the only reliable one obtained by road managers. A summary of these elements is provided in Table 4.

Once these elements gathered, a forecast based on the air temperature at the beginning of the stretch was used

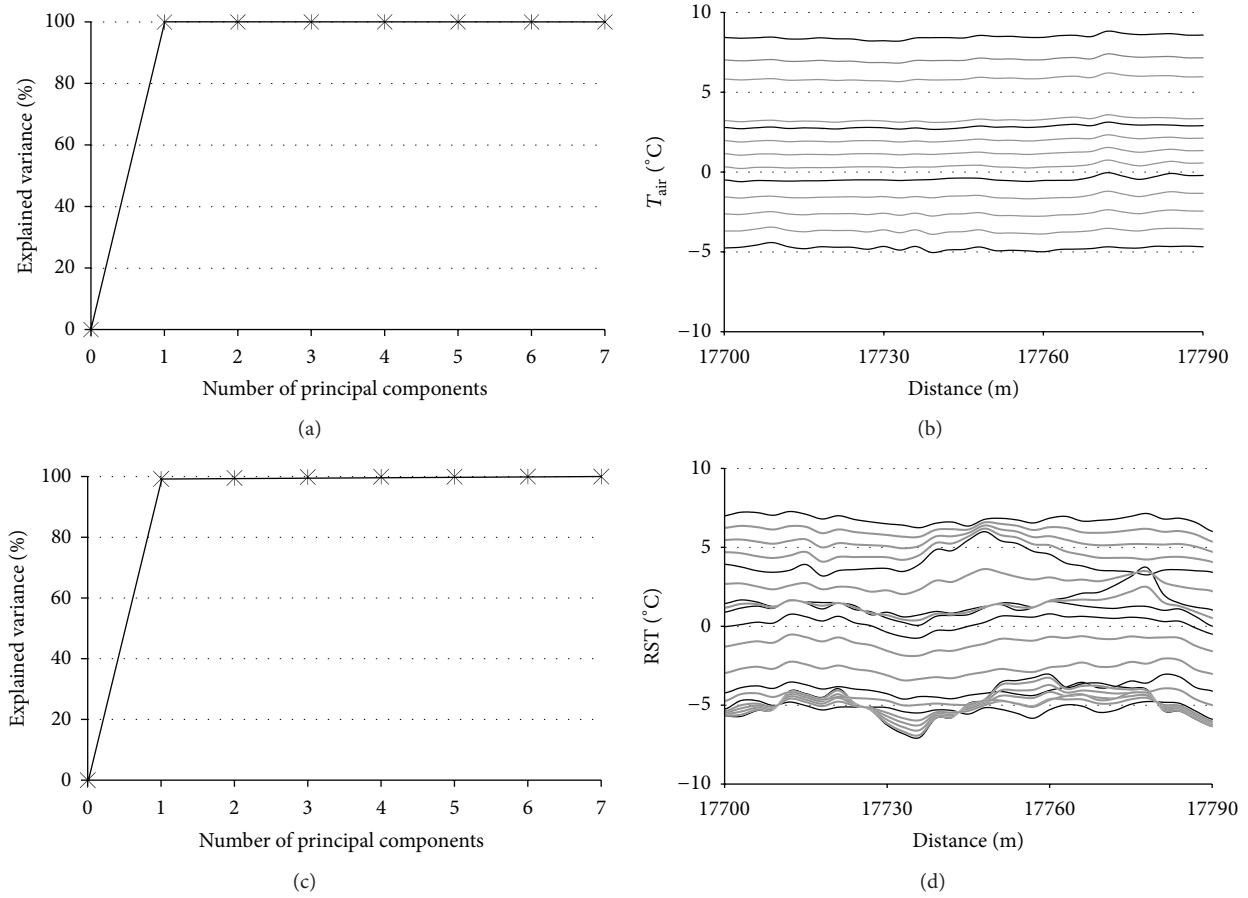


FIGURE 4: Results (explained variance (a, c) and interpolation (b, d)) of PCA applied to T_{air} and RST profiles of the urban stretch.

TABLE 2: Configurations for PLS calculations on the urban stretch.

| Samples number | Maximum explained variance percentage and associated factors number | R^2 | RMSE | Outliers number |
|----------------|---|-------|------|-----------------|
| 34 | 84%-14 factors | 0.97 | 2.6 | 50 |
| 8 | 94%-6 factors | 0.98 | 0.7 | 4 |
| 6 | 88%-4 factors | 0.99 | 0.7 | 4 |

TABLE 3: Results of PLS models with optimum number of thermal fingerprints.

| PLS configuration | Date | Measured RST ($^{\circ}\text{C}$) | RST forecast ($^{\circ}\text{C}$) | Bias β | Standard deviation σ_{β} |
|--|------------|-------------------------------------|-------------------------------------|--------------|-------------------------------------|
| RST $<10^{\circ}\text{C}$ (2 factors used out of 6 of the PLS model using 8 samples) | 2012-12-11 | 0.4 | 0.4 | 0.0 | 1.9 |
| | 2013-10-24 | 12.1 | 10.7 | -1.5 | 1.7 |
| RST $<10^{\circ}\text{C}$ (6 factors used out of 6 of the PLS model using 8 samples) | 2012-12-11 | 0.4 | 0.1 | -0.3 | 1.9 |
| | 2013-10-24 | 12.1 | 11.3 | -0.8 | 1.6 |
| RST $<10^{\circ}\text{C}$ (2 factors used out of 4 of the PLS model using 6 samples) | 2012-12-11 | 0.4 | 0.4 | 0.0 | 2.2 |
| | 2013-03-22 | 0.8 | 0.1 | -0.7 | 2.4 |
| RST $<10^{\circ}\text{C}$ (4 factors used out of 4 of the PLS model using 6 samples) | 2012-12-11 | 0.4 | -0.1 | -0.5 | 2.4 |
| | 2013-03-22 | 0.8 | -0.2 | -1.0 | 2.7 |

Where bias $\beta = \text{RST}_m - \text{RST}_{\text{measured}}$, with RST_m , $\text{RST}_{\text{measured}}$, respectively, modelled and measured road surface temperature.

TABLE 4: T_{air} data for RST forecast based on the coupling of PCA and PLS for RST forecast.

| Date | Thermal fingerprint data | | Météo France data | |
|------------|--------------------------------|---|---|--|
| | Start time thermal fingerprint | T_{air} at start time and start point ($^{\circ}\text{C}$) | Provided T_{air} ($^{\circ}\text{C}$) and time | Calculated T_{air} at start time ($^{\circ}\text{C}$) |
| 2012-12-11 | 8:30 a.m. | 2.1 | 0.6, 6:00 a.m./0.8, 9:00 a.m. | 0.7 |
| 2013-03-22 | 1:30 p.m. | 2.7 | -1.3, 12:00 a.m./4.9, 3:00 p.m. | 3.1 |

TABLE 5: Results of PLS models based on interpolated T_{air} profiles for two dates.

| PLS configuration | Date | RST measured ($^{\circ}\text{C}$) | RST forecast ($^{\circ}\text{C}$) | Bias β | Standard deviation σ_{β} |
|---|------------|-------------------------------------|-------------------------------------|--------------|-------------------------------------|
| 2 factors used out of 6 of the PLS model using 8 samples | 2012-12-11 | 0.4 | -1.9 | -2.3 | 1.3 |
| | 2013-03-22 | 0.8 | 0.5 | 0.3 | 1.2 |
| 2 factors used out of 4 of the PLS model based on 6 samples | 2012-12-11 | 0.4 | -1.6 | -2.0 | 1.0 |
| | 2013-03-22 | 0.8 | 0.5 | 0.3 | 1.2 |

TABLE 6: Evaluation of error made on RST forecast with a PLS model (2 factors used out of 4 of the model based on 6 samples) with an error of $\pm 1^{\circ}\text{C}$ on T_{air} profiles for two dates.

| Date | Error on T_{air} forecast ($^{\circ}\text{C}$) | -1.0 | -0.5 | 0.0 | 0.5 | 1.0 |
|------------|---|------|------|------|------|------|
| 2013-03-22 | RST measured ($^{\circ}\text{C}$) | | | 0.8 | | |
| | RST forecast ($^{\circ}\text{C}$) | -0.2 | 0.3 | 0.5 | 1.1 | 2.6 |
| | Standard deviation σ_{β} | 0.9 | 1.0 | 1.2 | 1.3 | 1.4 |
| 2012-12-11 | RST measured ($^{\circ}\text{C}$) | | | 0.4 | | |
| | RST forecast ($^{\circ}\text{C}$) | -2.6 | -2.0 | -1.6 | -1.1 | -0.6 |
| | Standard deviation σ_{β} | 1.1 | 1.0 | 1.0 | 1.0 | 1.0 |

to therefore establish RST forecast on the specific urban location. To do so and as detailed in Figure 3, once the air temperature in one point is identified, the corresponding full stretch T_{air} profile was established using PCA (Figure 4(b)). This profile was then used as input for the dedicated PLS model, whose performances were detailed in Table 3. Results for the two dates are given in Table 5.

In both cases, the standard deviation did not exceed 1.3°C , and the bias was larger in only one case (over 2°C) while below 1°C in the second case. The differences of performances could be easily explained. First of all, Météo France data every 3 hours has induced a poor accuracy in calculated T_{air} of Table 4. As indicated in Table 4, air temperatures measured at Météo France weather station and by the thermal mapping vehicle were different. The atmospheric probe of the thermal mapping vehicle did not measure air temperature as it is done by weather services using a standardized shelter. Furthermore, although the thermal mapping vehicle was not exactly in the immediate vicinity of the weather station at the start time, there were nearly 3 km several hundred meters between the two locations, and the thermal mapping vehicle one is constituted of buildings and pavements with respect to the absence of constructions and the presence of grass for the weather station. The further the T_{air} profile elaborated with PCA from the field measurement, the greater the risk to get an inappropriate RST.

The error in the RST forecast from PLS depends on the errors in the different variables used in its elaboration.

In a first approach, it depends on the error made on T_{air} used as input, especially in the situation when the profile of air temperature over the stretch is elaborated from one single point. To analyze to what extent the error made in the RST forecast, a set of PLS calculations was undertaken considering inputs of T_{air} profiles within $\pm 1^{\circ}\text{C}$ with respect to the measured T_{air} profile at two given dates (2013-03-22 and 2012-12-11) at start time and start point. Corresponding RST forecasts and a comparison with field measurement are given in Table 6. When the T_{air} inputs were between -0.2°C and 1°C with respect to T_{air} at start time and start point, 79% to 93% of the error on forecast RST was within $\pm 1^{\circ}\text{C}$.

The RST forecast value corresponds to the start point of the urban stretch. To obtain the whole RST profile over the very same stretch, the profiles obtained from the PCA and presented in Figure 4(d) were then used. The comparison between the forecast and the measured profiles is presented in Figure 5 for the two dates, 2013-03-22 and 2012-12-11, considering the RST forecast of Table 5 for the PLS model with 2 factors out of the 4 available ones and considering no error was made on RST.

In the 2013-03-22 case, 100% of the forecast is within $\pm 2^{\circ}\text{C}$, 94% within $\pm 1^{\circ}\text{C}$, and nearly 85% of the forecast is within $\pm 0.5^{\circ}\text{C}$, with a standard deviation of $\pm 1^{\circ}\text{C}$. The greatest difference of nearly 1.8°C between measurement and forecast can also be attributed to the presence of a specific urban configuration, with an incidence on the sky view factor that accounts for 60% of the RST value [20]. In the second

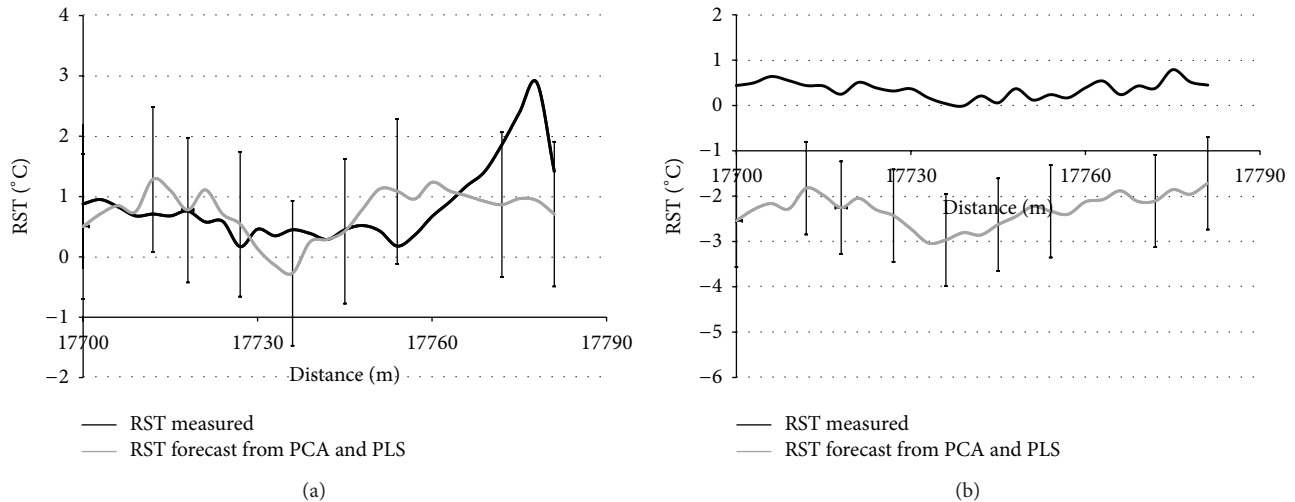


FIGURE 5: Comparison between measured RST and a RST forecast over a full urban stretch based on the combined use of PCA and PLS for two dates (2013-03-22 (a) and 2012-12-11 (b)).

case (2012-12-11), only 85% of the forecast is within $\pm 3^{\circ}\text{C}$, the difference being comprised between 3.2°C and 2.1°C . As indicated by Gustavsson et al. [21], RST is clearly affected by the urban configuration and by anthropogenic heat flux, in particular traffic, which is one of the differences between the two dates. The global radiation and the nebulosity were also among the main differences. Although good results were obtained with this PLS configuration in one situation, some variables, such as traffic, might be changed to numerical ones. Indeed, the traffic density can be related to heat flux due to tire friction, to sensible and latent flux due to vehicles engines, and to changes in convective heat exchange because of passing cars. Nevertheless, the objective is to generate a forecast based on a methodology which is different enough from the one of physical numerical models, and providing relevant and robust forecast. The extension of the forecast to a longer stretch could easily be obtained by elaborating new PLS models on the basis of an adjusted set of variables as the one illustrated in Table 1, and by iterating the PLS calculations to as many stretches as necessary.

4. Conclusion

The objective of this paper was to investigate a statistical approach to forecast road surface temperature on the basis of thermal mapping data. The tools used were principal components analysis and partial least-squares regression, to build a road surface temperature forecast based on air temperature and other meteorological variables along with an anthropogenic one. The forecast was first based on air temperature profile over the full urban stretch and then on one single data point. Results of the investigations indicated the possibility to elaborate a road surface temperature forecast with PLS and air temperature data. Nearly 94% of the forecasts were within $\pm 1^{\circ}\text{C}$ with respect to field measurements for a selected urban stretch in one situation and over 80% within $\pm 3^{\circ}\text{C}$. The PCA used to T_{air} profiles allowed the elaboration of generic

ones. Then, the knowledge of T_{air} from a single weather outstation was used to identify the corresponding T_{air} profile for further site-specific RST forecast with PLS. Once this value is obtained, the results of PCA to RST profiles were used to retrieve the RST profile over the full urban stretch. Results clearly depend on the proximity of this air profile from field measurements. An estimation of the error indicated that T_{air} inputs between $\pm 0.5^{\circ}\text{C}$ provided the best RST forecast results. Investigations indicated that the six to eight thermal surveys, mandatory to obtain a reliable road surface temperature forecast, and covering a large spectrum of weather situations, remain low enough to still have a cost effective approach of RST forecast within the global frame of winter maintenance. The idea will then be to iterate PLS calculations over several stretches to obtain a forecast over a full itinerary or over a full network. To obtain an analysis less statistical-mathematical and more physical, some investigations could be undertaken to apply independent component analysis (ICA) to thermal mapping data and to decompose a multivariate signal into independent non-Gaussian signals.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Authors would like to thank Météo France for its technical support, Dr. Nadège Brun, who inspired this work, for her technical advice on multivariate data analysis.

References

- [1] D. S. Hammond, L. Chapman, and J. E. Thornes, "Verification of route-based road weather forecasts," *Theoretical and Applied Climatology*, vol. 100, no. 3, pp. 371–384, 2010.

- [2] L. Chapman and J. E. Thornes, "A geomatics-based road surface temperature prediction model," *Science of the Total Environment*, vol. 360, no. 1–3, pp. 68–80, 2006.
- [3] J. E. Thornes, "Thermal mapping and road-weather information systems for highway engineers," in *Highway Meteorology*, A. H. Perry and L. J. Symons, Eds., pp. 39–67, E and FN Spon, 1991.
- [4] D. G. Belk, *Thermal mapping for a highway gritting network [Ph.D. thesis]*, University of Sheffield, Sheffield, UK, 1993.
- [5] J. Shao, P. J. Lister, G. D. Hart, and H. B. Pearson, "Thermal mapping: reliability and repeatability," *Meteorological Applications*, vol. 3, no. 4, pp. 325–330, 1996.
- [6] J. Shao, J. C. Swanson, R. Patterson, P. J. Lister, and A. N. McDonald, "Variation of winter road surface temperature due to topography and application of Thermal Mapping," *Meteorological Applications*, vol. 4, no. 2, pp. 131–137, 1997.
- [7] T. Gustavsson, "Thermal mapping—a technique for road climatological studies," *Meteorological Applications*, vol. 6, no. 4, pp. 385–394, 1999.
- [8] L. Chapman and J. E. Thornes, "The influence of traffic on road surface temperatures: implications for thermal mapping studies," *Meteorological Applications*, vol. 12, no. 4, pp. 371–380, 2005.
- [9] M. Marchetti, M. Moutton, S. Ludwig, L. Ibos, V. Feuillet, and J. Dumoulin, "Implementation of an infrared thermography camera for road thermal mapping," in *Proceedings of the 10th International Conference on Quantitative InfraRed Thermography (QIRT '10)*, pp. 245–251, Quebec, Canada, July 2010, <http://qirt.gel.ulaval.ca/archives/qirt2010/papers/QIRT%202010-083.pdf>.
- [10] H. Handa, L. Chapman, and X. Yao, "Robust salting route optimization using evolutionary algorithms," *Studies in Computational Intelligence*, vol. 51, pp. 497–517, 2007.
- [11] L. Chapman, "Probabilistic road weather forecasting," in *Proceedings of the Standing International Road Weather Conference (SIRWEC '12)*, Helsinki, Finland, May 2012, <http://www.sirwec.org/Papers/helsinki/4.pdf>.
- [12] M. Marchetti, L. Chapman, A. Khalifa, and M. Buès, "New role of thermal mapping in winter maintenance with principal components analysis," *Advances in Meteorology*, vol. 2014, Article ID 254795, 11 pages, 2014.
- [13] R. Kršmanc, A. Š. Slak, and J. Demšar, "Statistical approach for forecasting road surface temperature," *Meteorological Applications*, vol. 20, no. 4, pp. 439–446, 2013.
- [14] K. Liou, *An Introduction to Atmospheric Radiation*, Academic Press, London, UK, 2nd edition, 2002.
- [15] G. Sugrue, "Infrared thermometry—advantages and disadvantages," Atmosphere-Road Surface Interaction Group Working Paper 7, 1983.
- [16] L. Ibos, M. Marchetti, A. Boudenne, S. Datcu, Y. Candau, and J. Livet, "Infrared emissivity measurement device: principle and applications," *Measurement Science and Technology*, vol. 17, no. 11, pp. 2950–2956, 2006.
- [17] K. Kramer, *Chemometric Techniques for Quantitative Analysis*, CRC Press, New York, NY, USA, 1988.
- [18] E. S. Almkvist and C. Jansson, "How to measure the energy balance of a road surface," in *Proceedings of the Standing International Road Weather Conference (SIRWEC '04)*, Bingen, Germany, June 2004, <http://www.sirwec.org/Papers/bingen/24.pdf>.
- [19] S. E. Esben Almkvist and M. Walter, "Which variables should be measured at a road weather station—artificial intelligence gives the answer," in *Proceedings of the Standing International Road Weather Conference*, Turin, Italy, March 2006, <http://www.sirwec.org/Papers/torino/15.pdf>.
- [20] U. Postgård and M. Nunez, "Continuous measurements of sky-view factors along roads and their relationship to air and road surface temperature," in *Road Climate Variations Related to Weather and Topography*, U. Postgård, Ed., PhD Thesis, Earth Sciences Centre, University of Gothenburg, Gothenburg, Sweden, 2000.
- [21] T. Gustavsson, J. Bogren, and C. Green, "Road climate in cities: a study of the Stockholm area, South-East Sweden," *Meteorological Applications*, vol. 8, no. 4, pp. 481–489, 2001.

