Crowdsourcing the Urban Temperature of Amsterdam using Smartphones

M.J. Zander

Abstract

There is an increasing demand for high-resolution meteorological observations for urban areas, which are generally quite scarce. A potentially fruitful source of such data could be obtained through crowdsourcing. As smartphones have become ubiquitous and are equipped with many sensors, readings from these sensors can be a valuable 'opportunistic' source of meteorological data. Battery temperature readings from smartphones can be used to estimate the urban temperature.

In this study two large datasets from October 2016 and June 2017 are analyzed; containing more than hundred thousand smartphone readings per day. Readings from battery temperature sensors, separate ambient temperature sensors and light sensors are considered. Battery temperature readings are used in a straightforward heat transfer model to record the city-wide air temperature of Amsterdam. These are compared to temperatures measured by a network of more than 20 meteorological stations installed throughout the city.

Good results were obtained for June 2017 for the daily averaged temperature (coefficient of determination $\rho^2 = 0.84$), and reasonable results were obtained for hourly temperature estimations ($\rho^2 = 0.53$). For October, the results were less good ($\rho^2 = 0.50$), as readings from smartphones being charged could not be excluded. Readings from separate ambient temperature sensors installed in some smartphone models proved to be fruitless for estimating the urban air temperature in June ($\rho^2 = 0.33$). Light sensor readings show a distinct diurnal pattern. However, attempts to use these readings for improving the temperature estimations were ineffective.

Contents

List of Figures

List of Tables

1 Introduction

1.1 The Urban Weather & Climate

The world is rapidly urbanizing; in 2014, 54% of the world's population resided in urban areas, with an expected increase in all regions to 66% in 2050 (United Nations Department of Economic and Social Affairs, 2014). This poses a myriad of problems related to the (environmental) living conditions of urban dwellers, including resource use, sanitation, pollution and the urban climate. Meteorology and air quality form important aspects of life in the city and the urban metabolism. Inhabitants can suffer from heat in the city; during hot periods such as heat waves, the human thermal comfort can be even more drastically reduced in cities compared to rural areas, which can result in decreased labor productivity and increased mortality in cities. This is due to the phenomenon known as the Urban Heat Island (UHI). During the evening the city cools down at a slower rate than the surrounding countryside due to the high heat capacity of the materials of buildings and roads, the relative scarcity of vegetation, which all cause heat to be trapped. The nights therefore give no respite from the daytime heat (Steeneveld et al., 2016; Oke, 1982; Molenaar et al., 2016). The severe heat wave of August 2003 was for example associated with a five fold increase in the daily mortality risk in Paris (Steeneveld et al., 2016; Le Tertre et al., 2006) As the IPCC scenarios on climate change indicate an increase in heat wave frequency and severity in the future, there is urgency to adapt cities in order to cope with these changing conditions (Stocker et al., 2013; KNMI, 2014; Steeneveld et al., 2016).

To combat these environmental issues and to plan urban interventions effectively, meteorological measurements are needed on site. For urban planning, such in-situ measurements are needed at a high resolution as the urban landscape is highly heterogeneous. However, official measurements are scarce; for most cities there is often one official weather station, frequently at the airport outside the city center (Chapman et al., 2016).

Dedicated urban meteorological networks (sets of small weather stations spread out over an urban area) have been set up to in order to capture the capricious urban climate, like the Birmingham Urban Climate Laboratory (Chapman et al., 2016). In Muller et al. (2013) a review of the current state of such networks is provided. Apart from such point measurements, trajectory measurement campaigns have been held to quantify the (nocturnal) UHI in different zones of a city, like in Heusinkveld et al. (2014). Such intensive networks and campaigns provide more detail of the urban weather and climate, but are too resource-intensive to be applied broadly. Additionally, for large parts of the world, in-situ meteorological measurements are lacking altogether, while having such in-situ measurements is relevant for all areas. So, other sources of meteorological data should be explored, like data obtained by crowdsourcing.

1.2 Crowdsourcing in Atmospheric Sciences

Crowdsourcing is defined by Muller et al. (2015) as *"obtaining data or information by enlisting the services of a potentially large number of people"*. In recent years, this concept has gathered momentum within different disciplines of the scientific community. This is due to the increased willingness of the general public to participate in scientific practices and, notably, due to technological advances

(Snik et al., 2014).

Generally a distinction is made between, on the one hand, *participatory crowdsourcing* (or *citizen science, active sensors, participatory sensing*), in which citizens actively collect and submit data themselves, and on the other hand *opportunistic crowdsourcing (passive sensors)*, in which data generated for a different purpose are put to use for scientific research (Niforatos et al., 2017).

Muller et al. (2015) provide a comprehensive overview of current efforts in crowdsourcing for climate and atmospheric sciences. Compared to other scientific fields, the application of crowdsourced data in atmospheric sciences has been relatively limited, due to the discipline's strong focus on obtaining precise and representative observations. This measurement paradigm leads to a limited measurement coverage of formal weather stations because of the high costs of precise weather-monitoring equipment. This paradigm is also at odds with the concept of crowdsourcing, as with the latter, accuracy is traded-off for greater coverage; quantity over quality (Muller et al., 2015; Chapman et al., 2016).

Using data from weather stations maintained by weather enthusiasts is one way of applying crowdsourcing and thereby increasing the coverage of meteorological measurements in cities. Hobbyists install small weather stations at their homes and communicate their measurements through dedicated websites like Weather Underground (wunderground.com) and the WOW-project (wow.metoffice.gov. This data source has proven its worth in urban research (Steeneveld et al., 2011; Bell et al., 2013; Muller et al., 2015; Vos et al., 2017; Niforatos et al., 2017; Meier et al., 2017). These personal weather stations can be costly and the number of avid weather enthusiasts is relatively limited. Fortunately, there are also more low-cost options available, for example using devices which have become ubiquitous in daily life: smartphones. The abundance and popularity of smartphone weather applications ('apps') indicate a large potential for crowdsourcing in the weather domain (Niforatos et al., 2017).

1.3   Smartphone Measurements

There is now a wealth of possibilities to participate in crowdsourcing projects with one's smartphone. In the *mPING* project for example, citizens file reports on the precipitation type they see in an app, this data is used to validate radar derived precipitation estimation (Elmore et al., 2014). Another example is described by Schweizer et al. (2011) who developed an app which automatically collects noise levels from smartphones for mapping urban noise pollution. Thompson (2016) provides a review of the recent efforts of monitoring air quality by means of crowdsourcing. Many of the mentioned projects require separate sensors, smartphone sensor add-ons or mobile applications which the user installs on his/her phone. For instance, Snik et al. (2014) describe the *iSPEX* Smartphone add-on to attach to a smartphone camera to measure aerosol optical thickness. Thompson (2016) conclude that as of yet, there is no perfect sensing platform meeting the requirements of being portable, sensitive, selective and low-cost (Thompson, 2016).

Instead of connecting separate sensors to a smartphone, one could make use of the large number of sensors which are installed inside smartphones. For instance, the Samsung Galaxy S4 was advertised to have amongst others; a barometer, a separate ambient temperature & humidity sensor, geomagnetic sensor, accelerometer, gyro sensor, proximity sensor and a RGB Light Sensor. These sensors are installed for health and lifestyle applications and convenience (Samsung,2013). In

research, these sensors have been applied to determine what type of activity the smartphone user is employing (Vanini et al., 2016; Niforatos et al., 2017). Or for instance in geosciences for collecting ground level magnetic data from the geomagnetic sensor (Saltus and Nair, 2017). Explored opportunities for using these sensors for atmospheric applications include using smartphones on a hot air balloon flight to track the wind, and sensing the atmospheric pressure (Madaus and Mass, 2017; Niforatos et al., 2017; De Bruijn et al., tted). One interesting sensoring application is to use the thermometers installed in smartphone batteries to record the ambient air temperature.

Using battery temperature readings

Overeem et al. (2013) have used a dataset of smartphone battery temperature readings to model the urban ambient temperature on a daily basis for eight cities in different climate zones. These readings were collected by the smartphone app *OpenSignal*. The dataset contained 1.3 million records for the eight cities, with at least 400 readings available on 80% of the days. On a daily basis, good results were obtained when compared to official WMO station data from nearby airports, with fractions of explained variance up to $\rho^2 = 0.86$ (for the city of Rome) (Overeem et al., 2013). The analysis was repeated for two Dutch cities; Rotterdam and Amsterdam for the warm summer of 2013, with again good results: $\rho^2 = 0.77$ and 0.67 respectively. To do so, apart from measurements at the WMO stations at the nearby airports, additional in-situ temperature measurements from the local health authority were used as a reference for Amsterdam (Overeem et al., 2014).

The temperature of a smartphone battery is of course not only influenced by the ambient temperature and battery usage, as it is generally carried around in a clothing pocket. So it is assumed that the phone also receives body heat of the owner; the rate of which is dependent on the isolation between the phone and the outside air and the body. To take this into account, and convert the battery temperature readings a straightforward steady-state heat transfer model was used.

Droste et al. (2017) built upon the work of Overeem et al. (2013) by zooming in to one city: São Paulo. The authors do not only estimate daily mean temperature, but also estimate the ambient air temperature on an hourly basis. The modeled temperatures were calibrated and validated on the measurements of seven amateur weather stations in the city, and two authoritative weather stations; a micro-meteorological tower station, and the official measurements at the São Paulo airport.

So for ten cities it has been shown that smartphone battery temperatures can be used to estimate the city-wide daily averaged temperature. Additionally, for one city estimations of hourly temperatures proved fruitful (Overeem et al., 2013, 2014; Droste et al., 2017). A similar study for hourly temperature estimations for a different city can assess the robustness of these findings.

People generally spend most of their time inside buildings, whether working or resting. This is clearly a challenge when applying smartphone crowdsourcing for sampling environmental conditions. Researchers can thus choose to focus on sensors which remain mostly unaffected by the precise context of the user, like the atmospheric pressure sensor (as done in Madaus and Mass (2017)). Or alternatively, efforts can be taken to use information from additional sources to infer the context

of the user, and thereby filter for outdoors readings (Niforatos et al., 2017). One such method to infer whether a smartphone is indoors or outdoors is applied by Broyde et al. (2013); the authors used statistical analysis of path attenuation of the wireless transmission between mobile phones and telecommunication masts to distinguish between inside and outside phone usage. For this the signal strength at both the sending and receiving part is needed, so not only information from the smartphones themselves, but also information from the cellular network is necessary. Therefore, this approach cannot be applied when only collecting smartphone data via an app.

Additional smartphone sensors might therefore be used to disambiguate if the smartphone is outdoors. Many smartphones have a light intensity sensor, installed in smartphones to adjust the screen brightness. The readings from this sensor are influenced by the users context (e.g. under office lighting) and the weather conditions (e.g. sunny or overcast).

The work of Niforatos et al. (2017) is the first in which smartphone light sensor and ambient temperature sensor readings are collected for meteorological applications. Niforatos et al. (2017) launched an app, *Atmos*, to passively collect data from a limited number of smartphone sensors for weather data collection, combined with asking the volunteers to manually fill in their estimations on the weather. The authors applied supervised machine learning to predict the current and future ambient temperature from manually submitted weather data and mobile sensor data both collected through a mobile app. For the sensor data, the global ambient temperature estimations had an average error of about 10 °C compared to the reference measurements at Weather Underground stations. When combining data from both data sources, the minimal average error they found was 7.82 °C using four input variables in a regression tree analysis. However, in their research the collected data are quite spread out; just over 2×10^5 useful measurements stemming from 26 countries over a period of 32 months (Niforatos et al., 2017).

Compared to battery temperature sensors, the separate ambient temperature sensors, which are installed in some smartphone types, would be less affected by the smartphone usage. This is due to their placement within the phone (Niforatos et al., 2017). These sensors might therefore be a fruitful data source to estimate the ambient temperature.

Hence, it is interesting to further explore not only readings from battery temperature sensors, but also from the separate temperature and light intensity sensors from smartphones. Having a dense set of smartphone readings from an urban area combined with a robust reference of air temperature observations provides a useful extension to previous endeavors as it overcomes issues on representativeness and data quality for the reference data.

2 Research objectives

This project aims to build upon the work of Overeem et al. (2013); Droste et al. (2017) and Niforatos et al. (2017) by analyzing an extensive dataset with smartphone readings from *OpenSignal* to estimate the air temperature in the city of Amsterdam. The urban temperature will be estimated per hour and per day by using battery temperature readings. It aims to assess how many readings are needed for accurate hourly temperature estimations. Additionally, it aims to assess if the urban temperature can be estimated from ambient temperature sensors in smartphones. Furthermore,

this project explores the suitability of using readings from the light sensor to reduce the number of indoor readings incorporated in the analysis in an attempt to improve the temperature estimations.

3 Materials and Methods

3.1 Case study description

The metropolitan area of Amsterdam is chosen as it is the most populated city of the Netherlands and because there are multiple urban weather stations installed throughout the city, providing a robust reference (Section 3.2.2, Figure 2). Two datasets of crowdsourced smartphone readings are available for this study; one for autumn 2016, for which the data of October are used, and for the month of June 2017. The June dataset will be analyzed more extensively (Section 3.2.1).

Weather during the case study periods

The month of June 2017 was exceptionally warm in the Netherlands; with an average temperature of 18.0 °C.

High pressure systems provided stable summer weather from June 10th, with especially warm days from June 18th to 22nd. This ended the 22nd with thunder. The following days were overcast with rain from low pressure systems brought in from the sea (see Figure 1) (Homan, 2017).

In October 2016 it was relatively sunny, cold and dry in the Netherlands with an average temperature of 9.9 °C. The month started with some rain on October 1st and 2nd. It was relatively warm on October 3 & 4, 7 and especially on October 16. This was followed by rainy days from the 17th to the 21st, after which it was dry for the rest of the month (Homan, 2016).

3.2 Data Collection

3.2.1 Smartphone data

The company OpenSignal provides a smartphone app under the same name aimed to find the best network signal strength. App users have granted OpenSignal the permission to collect information on the specifics of their phone as well as on the readings from the sensors in the devices. The app logs information from a large number of network details, processes and sensors in the smartphones. Apart from those used to provide the actual service of the app, readings from sensors like the battery thermometer are also retrieved as collateral. OpenSignal has provided two datasets with smartphone readings; from September to November 2016 and a dataset for the Amsterdam area for the month of June 2017. The readings are logged every 15 seconds when the app is activated and the smartphone screen is active.

3.2.2 Weather station data

Currently, an urban meteorological network is being set up in Amsterdam, the *Amsterdam Atmospheric Monitoring Supersite* (from now on referred to as "AMS stations"). This is part of a research project of the *Amsterdam Institute for Advanced Metropolitan Solutions*, the *Meteorology and Air Quality* chairgroup of the Wageningen University and the *Summer in the City* research project (Ronda et al., 2014; Wageningen University, 2016). When fully installed, the network will consist of 30 meteorological stations measuring wind velocity, humidity and temperature across the city. For the case study period, data from a maximum of 25 stations are available (Figure 2).
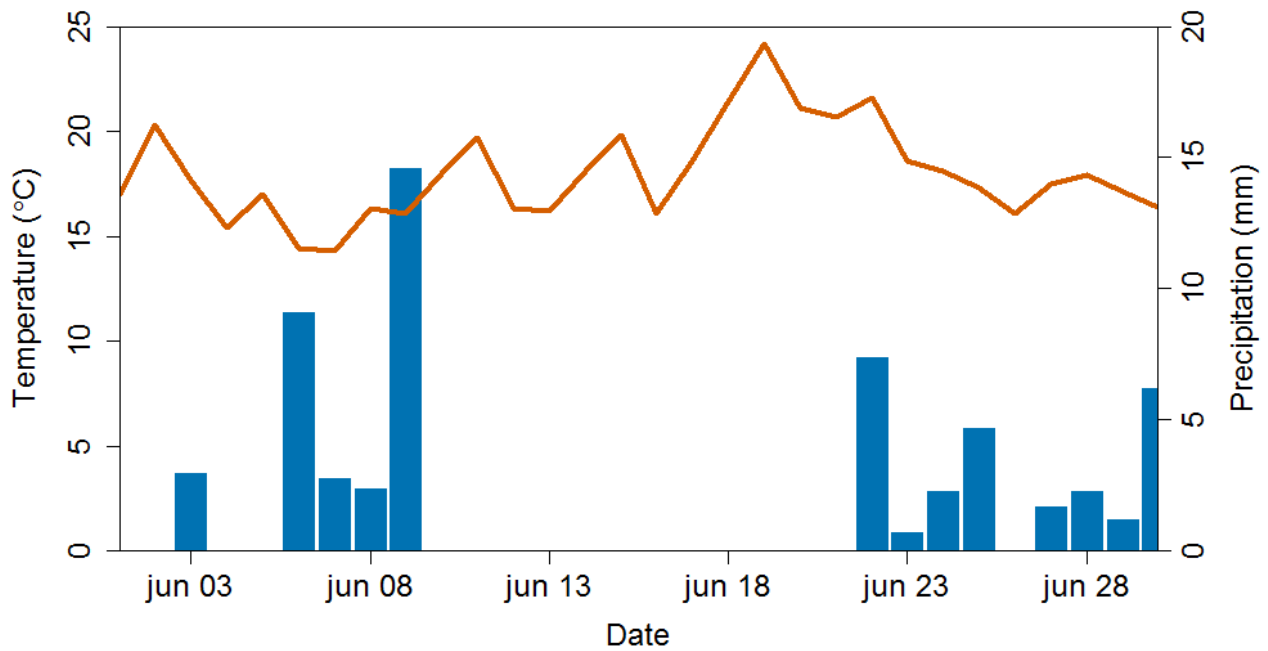
Figure 1: *Daily mean temperature (orange line) and daily precipitation sums (blue bars) at WMO station Schiphol, June 2017.*

The measurement data from these stations will be used to assess how much of the temperature signal is captured in the signal of the smartphone readings and to assess the performance of the heat transfer model. The average of all available meteorological stations is used, excluding one station located on an island in the IJ lake, as the footprint of this station is far from representative for the typical urban conditions of Amsterdam. Additionally, the nearby official WMO weather station at Schiphol airport (KNMI station 240, WMO code 6240, Figure 2, pink circle) will be used as a reference for the smartphone light sensor readings as this station is equipped with a pyranometer measuring global radiation.
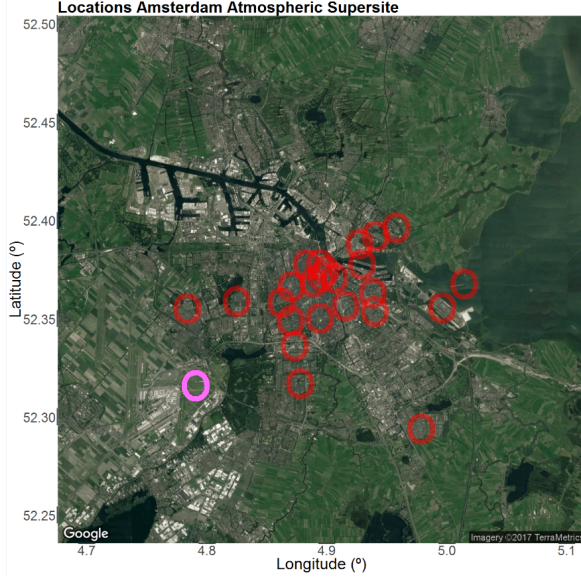
Figure 2: *Locations of the AMS stations (red circles) and the WMO station Schiphol (pink circle).*
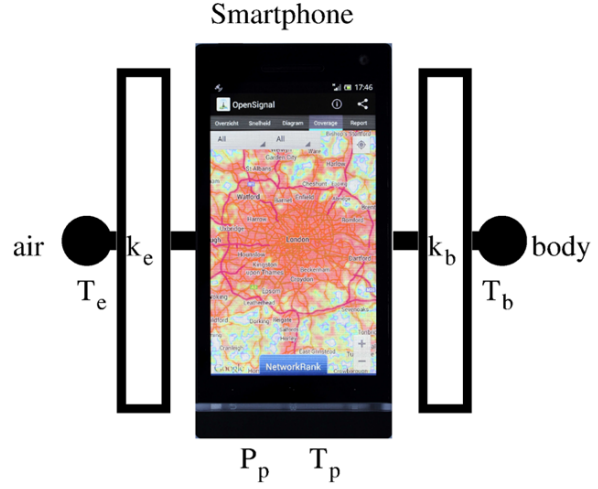

Figure 3: *Conceptual representation of the heat transfer model (Overeem et al., 2013)*

3.3 Modeling Method

3.3.1 Data filtering

As the OpenSignal app is also used by tablets, there are also tablet readings in the datasets. For the sake of clarity, these are left out for the analyses. For the available datasets, Open Signal applied a sampling strategy of collecting readings every 15 seconds while the screen is on (and the app is running). This contrasts with the sampling strategy of the previous work of Droste et al. (2017) and Overeem et al. (2013).

In these studies smartphone readings were only logged only when the smartphone was switched on or off and when it was plugged or unplugged from a charger. The readings for which the smartphone was just unplugged and discharging were filtered out as the battery temperature would be higher due to the charging. So only readings where the phone was switched on or off and was discharging and for readings when the smartphone was just plugged into the charger were selected (Overeem et al., 2013), (Droste et al., 2017).

Battery Temperature

The filtering applied in previous studies is thus not possible. So other filtering steps need to be applied to remove readings which are unlikely to be of value for the analyses. Or alternatively, to capture the readings which are assumed to be most representative for the environmental conditions: measurements taken outside. As charging the battery typically warms the phone, readings from phones being charged are filtered out (variable *BT_PLUGGED*). Furthermore, smartphones are typically charged whilst inside. Analogous to the studies of Overeem et al. (2013) and Droste et al. (2017), battery temperature readings assumed erroneous are also filtered out; the range assumed to be plausible is 10 to 47 °C.

Readings from the smartphone light intensity sensors are explored to examine the viability of using this sensor to filter out smartphone readings taken inside on an hourly basis (section 3.4).

Separate Temperature Sensor

For the same reasons as stated above and to keep the results comparable between the two types of temperature sensors, batteries being charged are also filtered out for the analysis of the separate ambient temperature sensor. A smartphone is most likely charged inside and the additional heat from charging will also influence this sensor. The range of values assumed plausible is taken to be slightly different than that of the battery temperature sensors; 5 to 40 °C, as it is assumed that the placing of the thermometer would make the measurement more direct, rendering lower readings possible and higher readings as outliers. Only readings with the highest reported temperature sensor accuracy are considered.

3.3.2 Heat transfer model

The ambient temperature of Amsterdam will be estimated on a daily and on a hourly basis. To take into account that a smartphone is far from a straightforward standalone sensor, the heat transfer model taken from Overeem et al. (2013) is used to estimate the ambient temperature via the battery temperatures. Smartphone owners typically carry their device around in a pocket of their clothing. The phone will thus often be close to the user's body and receive body heat. The processes in the phone also generate heat. Therefore the temperature of the smartphone battery T_b (°C) is assumed to be influenced by the environmental ambient temperature T_e (°C), the thermal energy generated by the phone P_p (W), and the human body temperature T_b (°C). For the statistical model, other potential heat sources are not taken into account. The heat flow between two neighboring systems is proportional to the temperature difference between the systems and the resistance between them. The thermal insulation between the phone and the environment is captured in the model as coefficient k_e ($W°C^{-1}$) and between the phone and the body, k_b ($W°C^{-1}$) (Figure 3).

Thus the following equation is defined (Overeem et al., 2013);

$$T_e = \left(1 + \frac{k_b}{k_e}\right) \times T_p - \left(\frac{k_b}{k_e} \times T_b + \frac{P_p}{k_e}\right). \tag{1}$$

For steady-state heat flow between P_p and the heat flow to the environment and the body, and if the values of T_p, T_b, P_p, k_e, and k_b can be assumed independent over a set of readings. Averaging over a city and over a day or an hour, this leads to the following simplification;

$$T_e(t) = m \times [T_p(t) - T_0] + T_0 + \epsilon, \tag{2}$$

in which m (-) is the average of $(1 + k_b/k_e)$ from all individual observations within a given period for a certain location. The previously found values of m for different cities range from 2 to 3 (-), which implies that the battery temperatures fluctuate less than the ambient temperature. The ϵ (°C) is a random error. T_0 can be interpreted as the normal human body temperature plus a constant factor (Overeem et al., 2013; Droste et al., 2017). See the supplementary information of Overeem et al. (2013) for a more elaborate derivation of the heat model.

3.3.3   Calibration and validation

For the base run, initial values of $m = 2.1$ and $T_0 = 39.0°C$ will be taken from Overeem et al. (2014). This is the value found for when calibrating the urban air temperature in Amsterdam for the summer months of June, July and August 2013 with five meteorological stations from the public health service as a reference.

Subsequently, the month of June is divided into a calibration period (June 1 to 14) and a validation period (June 15 to 30). The model will only be calibrated on m, T_0 is assumed constant at 39.0 °C. The adjacency of these periods does infer unavoidable temporal auto-correlation. The non-linear least squares method is applied to calibrate for the value of m. The modeled temperature is compared to the averaged temperature of the AMS stations: the coefficient of determination or fraction of explained variance (ρ^2 (-)); root mean square error (RMSE (°C)); the bias or mean error (ME (°C)); and mean absolute error (MAE (°C)) will be calculated. As m is the ratio of the k coefficients of thermal insulation, it can be expected that this coefficient varies throughout the year and throughout the day. Therefore, for hourly averaged air temperatures values for m will be calibrated for each hour of the day.

To assess the dependency of the model performance on the amount of available readings for the hourly temperature estimation, the effect of setting a minimum threshold is researched; any hours for which the number of available readings is less than the threshold, are excluded from the analysis. The threshold is increased from 500 to 6000 in steps of 250 readings per hour, the resulting coefficient of determination and RMSE between the modeled temperature and the reference are calculated.

3.4   Light Intensity

Readings from the smartphone light sensors are explored, these are reported in lux ($lumen\,m^{-1}$). Only readings with the highest accuracy (*light_acc=3*) are taken into consideration. Readings from smartphones being charged are also left out of this analysis. The mean hourly light intensity from smartphones is calculated and compared to the official measurements of global radiation ($J\,cm^{-2}\,hour^{-1}$) at the WMO station Schiphol. A conversion from the one unit for light in the other is not straightforward as every wavelength of light has a different energy level, and the wavelengths which are measured by the smartphone sensors may differ amongst models. An approximate conversion, which would be valid for the Sun, is applied; 0.0079 Wm^{-2} per lux (Cumulus forum contributors, 2010).

The possibility of using these light sensor readings to make a distinction between inside and outside conditions is explored. This is done by defining thresholds; readings where the light sensor indicates a lower value than either (a) the median, (b) mean light intensity for that hour are excluded. Alternatively, a fixed threshold is applied, representative for inside office conditions (first a conservative 100, then 500 Lux, (Warwick University)). For this analysis, only daytime hours are taken into account. After filtering, the remaining readings are averaged to daily values for the daytime hours.

4 Results

The results from the October 2016 dataset are presented first, followed by the results from the June 2017 dataset. Hourly averaged ambient temperatures from battery temperature readings are discussed, after which the hourly averaged estimations are shown. The effect of the number of readings is subsequently discussed, followed by an analysis of readings from ambient temperature sensors. Finally the analysis of light sensor readings is presented.

4.1 October 2016

For October 2016 there are 3.9 million battery temperature readings within the Amsterdam area; an average of nearly 126 thousand a day. At nighttime there are substantially fewer readings than during the day; two versus six thousand per hour. The peak at 15 UTC coincides with the end of a regular workday (Figure 5a). Filtering out smartphones being charged is not possible, as readings about this variable are not included in the dataset. 4% of readings were made with the screen off and locked. There is a diurnal pattern in the battery temperature readings; the mean hourly temperature reaches a minimum of 29.7 at 4 UTC, rising to 30.9 at 22 UTC, with a distinct drop around 16 UTC (Figure 5b). These lower temperatures are most probably due to the evening commute; smartphone users being outside after work. The average battery temperature is 30.1 °C(Figure 4).

Daily temperature estimation

With the base run $m = 2.1$, the modeled temperature ranges from 18.8 °C on October 10th to 22.5 °C on October 28th, with an average of 20.3 °C; more than 9 °C higher than observed (ME = 9.2 °C). The pattern in the modeled temperature only partially follows the pattern of the observed daily temperatures at the AMS stations; there are three distinct spikes in the temperature, by which the model misses the relatively warm days the 3rd and 4th of October. The coefficient of determination ρ^2 is 0.5 (Figure 6).

As the heat transfer model is a linear model, the correspondence in signal between the modelled and measured temperature does not improve with calibration of m.

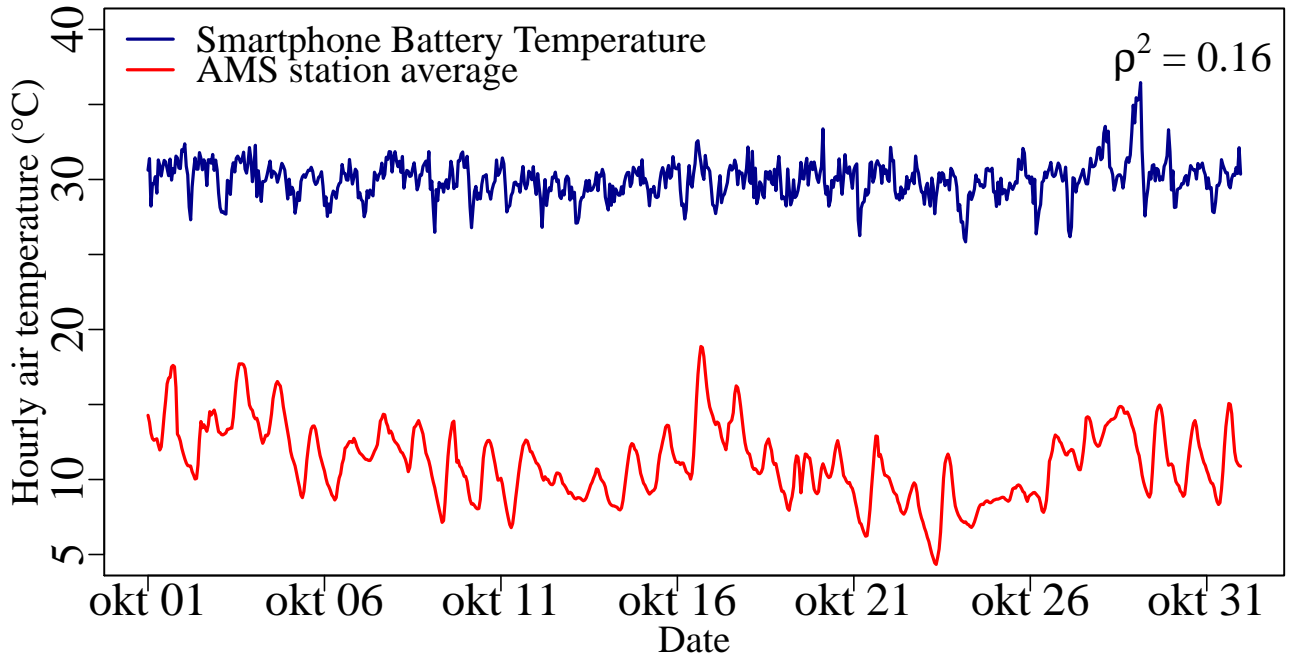Figure 4: *Pre-processed hourly battery temperature from smartphones, and the average hourly temperature from the AMS stations, October 2016.*
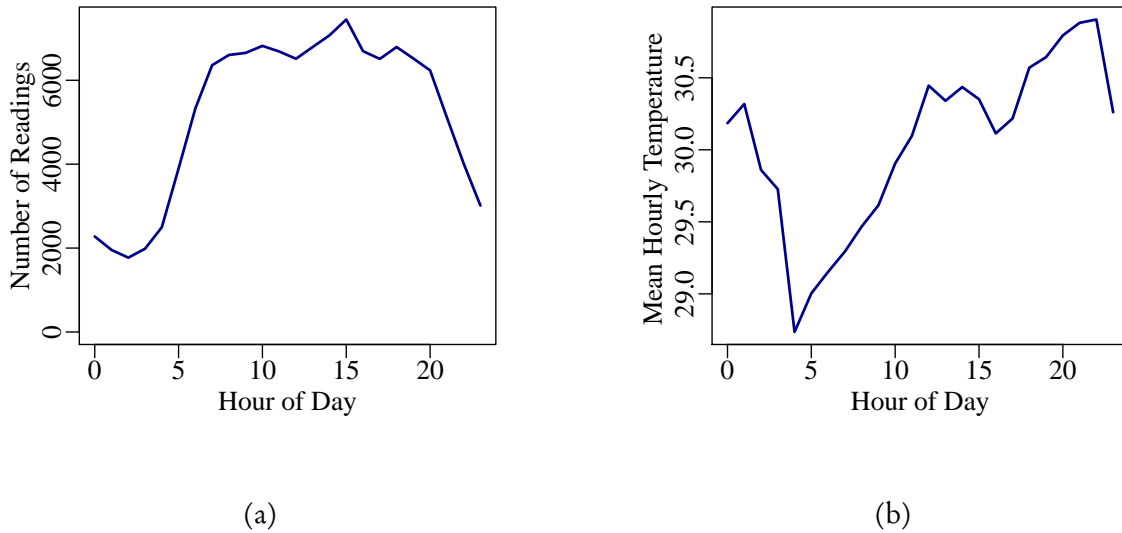

(a)

(b)

Figure 5: *Mean number of battery temperature readings (a) and mean battery temperature (b) per hour of day, October 2016*
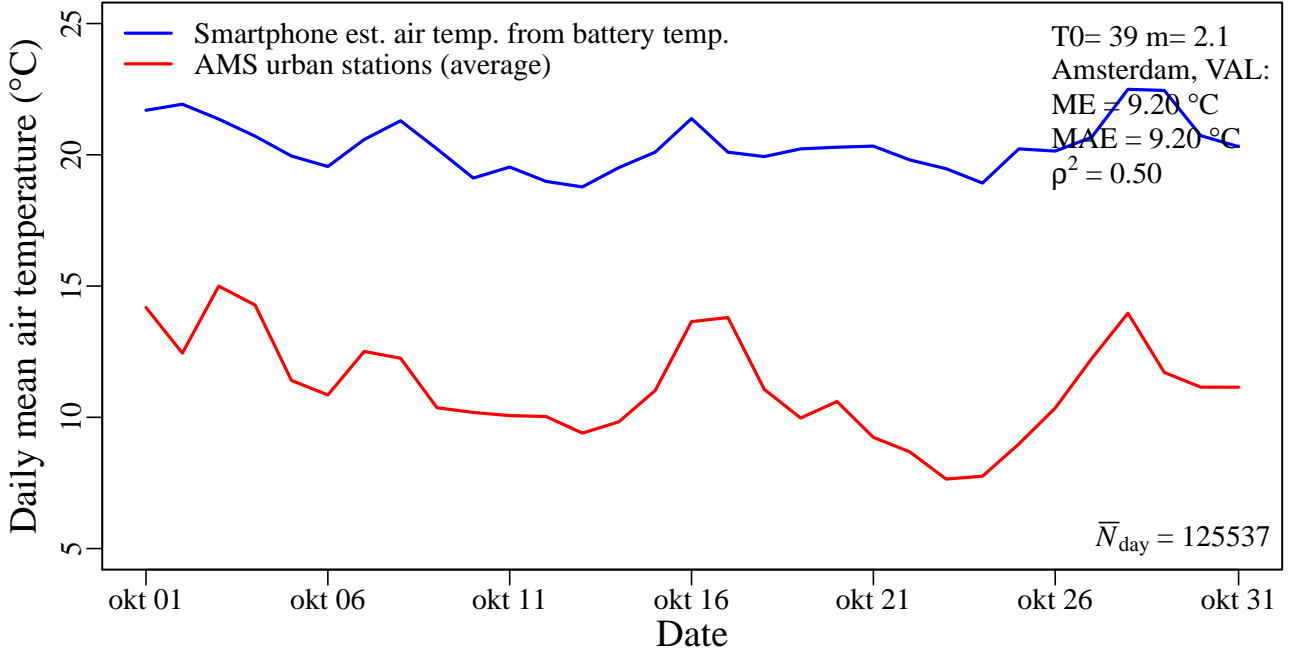
Figure 6: *Base run estimated daily temperature compared to the average temperature of the AMS stations, October 2016.*

4.2   June 2017

For the month of June 28% of readings came from smartphones being charged, and were thus filtered out. $4.2 \cdot 10^3$ readings remain after filtering; around $1.4 \cdot 10^5$ readings a day. The readings come from a total of 3199 unique smartphones. The average battery temperature reading is 31.3 °C. The battery temperatures vary less than the measured ambient air temperatures (Figure 7). There is, however, a daily cycle in both the battery temperature and the number of readings; with lower temperatures and less readings during the night (Figure 8). The minimum hourly battery temperatures is 26.7 °C on June 14th 02:00 UTC, the maximum hourly battery temperatures are however also found during nighttime; 36.3 °C on June 27th 01:00 UTC, and 35.0 °C on June 6th 3:00 UTC. These outliers are based on a lower number of readings than average for that time of day. The minimum number of readings in one hour is 146, the mean 5771, and the maximum 12041 (Figures 7 & 8). During weekends, there are slightly less readings and the number of readings starts to increase an hour later compared to weekdays (not shown).

Both the temperature measurements at WMO station Schiphol and the averaged temperature from the AMS stations are shown in Figure 7. As can be seen, during most night the temperature in the city is higher than that at Schiphol; an urban heat island. The AAMS stations give a better representation of the environmental context of the smartphone users, thus from here onwards, only the temperature from the AMS stations will be depicted.

4.2.1   Daily temperature estimation

For the base run ($m = 2.1$; $T_0 = 39.0$ °C, whole month), the pattern of warm and cold days is captured quite well ($\rho^2 = 0.84$), although the modeled temperature peaks one day earlier than measured

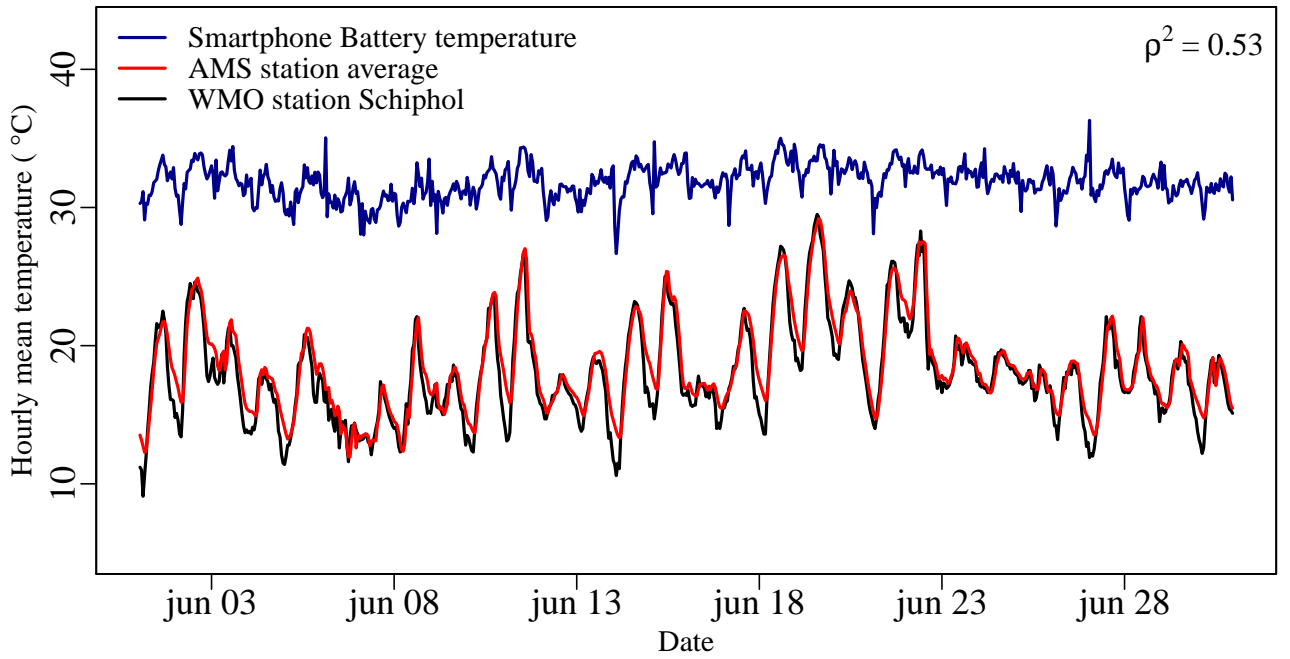Figure 7: *Pre-processed hourly battery temperature from smartphones, hourly temperature at WMO station Schiphol and the average hourly temperature from the AMS stations, June 2017.*
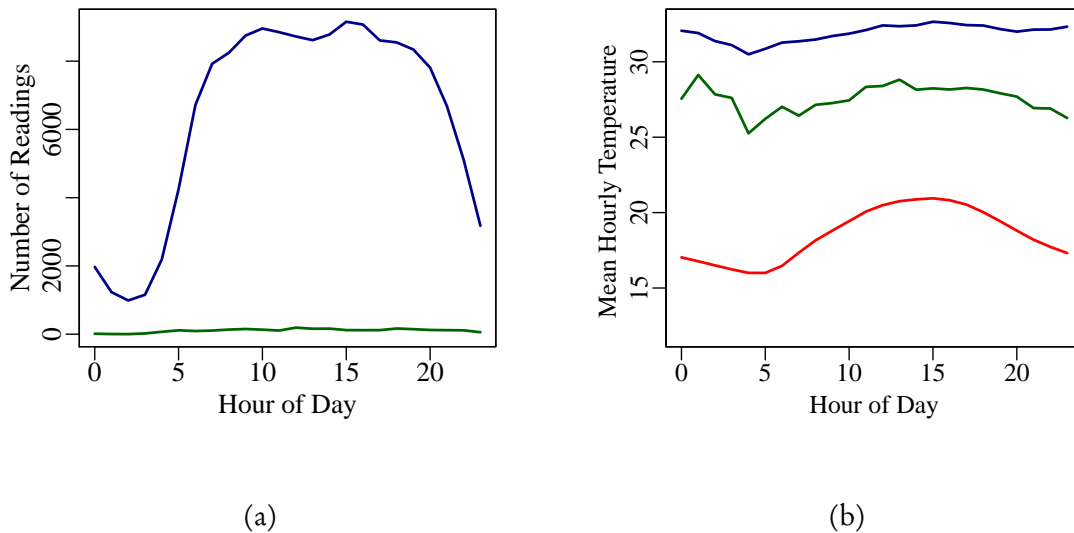

(a)

(b)

Figure 8: *Average number of readings per hour of day (a) and average temperature per hour of day (b) from battery temperature sensors (blue), ambient temperature sensors (green), and the AMS station average (red), June 2017.*
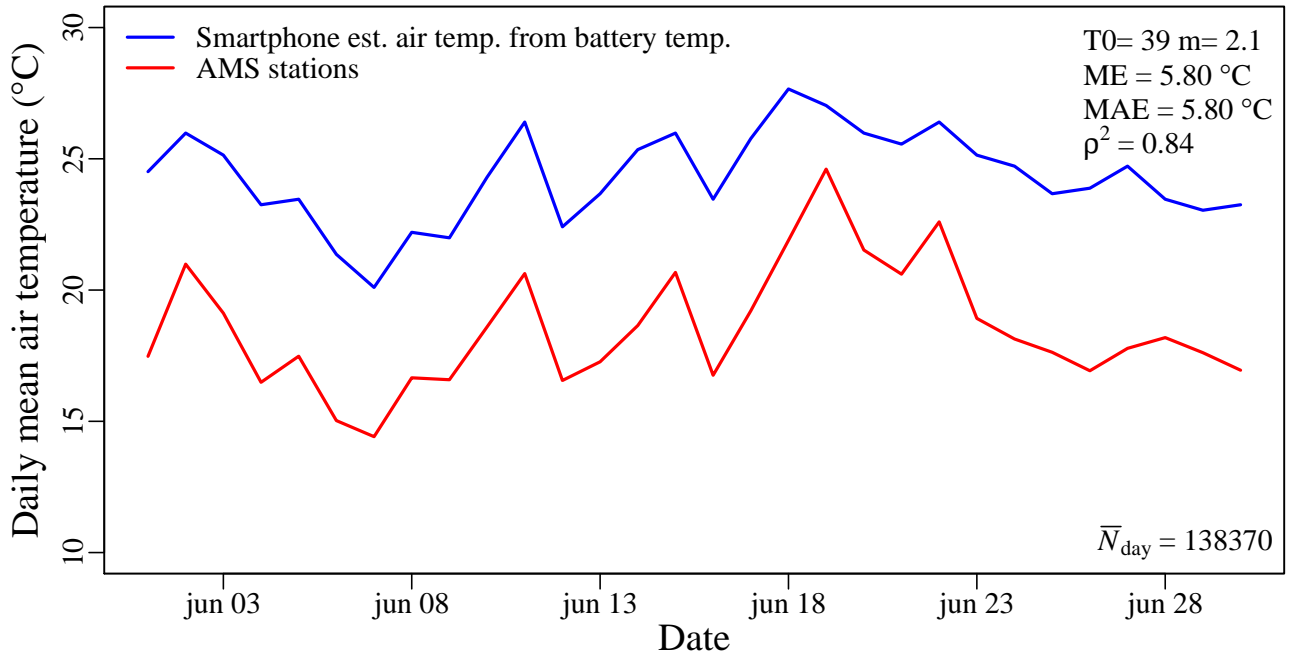
Figure 9: *Base run estimated daily temperature compared to the average temperature of the AMS stations, June 2017.*

(June 17 versus 18), resulting in a high bias for that day. There is a positive bias (ME) of 4.7 °C (Figure 9). Taking the median temperature rather than the mean for calculating hourly battery temperatures does not improve the results (Table 1). Inclusion of battery readings where the phone is being charged leads to poorer estimations, as expected (Table 1). The daily temperature is on average 0.3°C higher when the readings are not filtered for batteries being plugged in.

As can be seen in Table 1, the model performance improves if the AMS stations are used as a reference, compared to using the WMO station Schiphol, Overeem et al. (2014) found similar results.

Calibration and Validation

As it is not possible to apply the same filtering steps to the datasets of October 2016 and June 2017, m is only calibrated for the month of June 2017. The month is divided into two parts to use for calibration and validation. The calibrated m is higher than the base run; 2.91 (-). As $m = (1 + k_b/k_e) = 2.91$, k_b is nearly twice as large as k_e, implying that the conductivity between the body and the phone is larger than between the phone and the environment.

The bias is reduced by calibrating m; for the validation period, the model underestimates the high temperatures from June 18 to 22, while it overestimates the temperatures of consecutive colder days, resulting in a low mean error (ME=0.19 °C). The RMSE is reduced to 1.07 °C and the MAE is 0.93 °C (Table 1, Figure 10).

As can be seen in Table 1, the choice of assigning calibration and validation periods influences the results; in the first weeks of June the battery temperatures resemble the measured temperatures more closely.
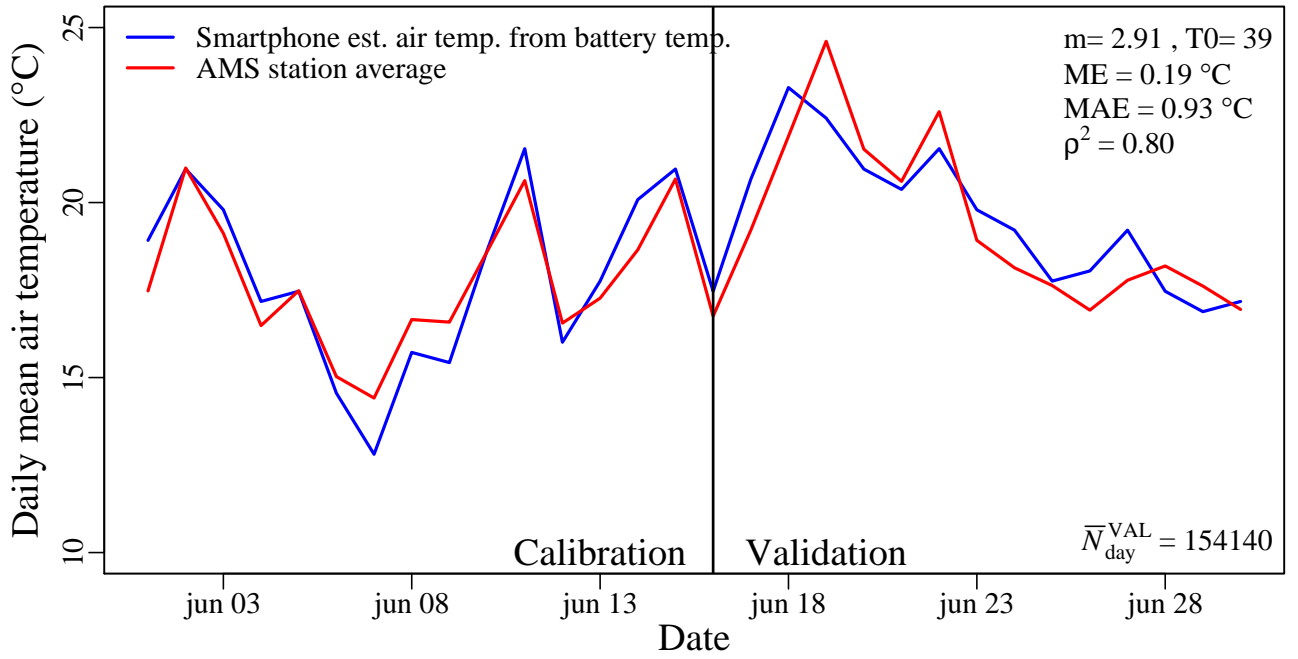
Figure 10: Calibration period and validation period and results for daily temperatures estimations.

Table 1: *Statistics for the modeled daily temperature compared to the AMS station average, for the whole month of June 2017 unless indicated otherwise.*

type	m [-]	ρ^2 [-]	ME [°C]	RMSE [°C]
base run	2.1	0.84	4.70	4.79
base run reference Schiphol	2.1	0.78	6.35	6.44
median battery temperature	2.1	0.81	5.7	5.78
base run incl. charging batteries	2.1	0.80	6.47	6.57
calibration June 1-15	2.910	0.80	0.19	1.07
calibration June 16-30	2.897	0.91	0.18	0.88

4.2.2 Hourly temperature estimation

For the base run ($m = 2.1$; $T_0 = 39.0$ °C, whole month), the model performance is worse than that for daily averaged temperatures, with $\rho^2 = 0.53$ (Table 2). Especially for the overcast days of June 23 to 26, the estimated temperature is a lot higher than the reference (Figure 11). There seems to be a delay in the evening cooling, not in the morning heating; the battery temperatures cool later than measured by the reference stations. There appears to be no systematic delay between the measured temperatures at the AMS stations and the battery temperatures (Figure 11).

Calibration and validation

Several methods of calibration are applied; first with a fixed value for m; the coefficient is kept the same for all hours of the day. Subsequently, a separate value for m is calibrated per hour of day. In both cases this is done for the same calibration period as applied for the daily averaged analysis. As
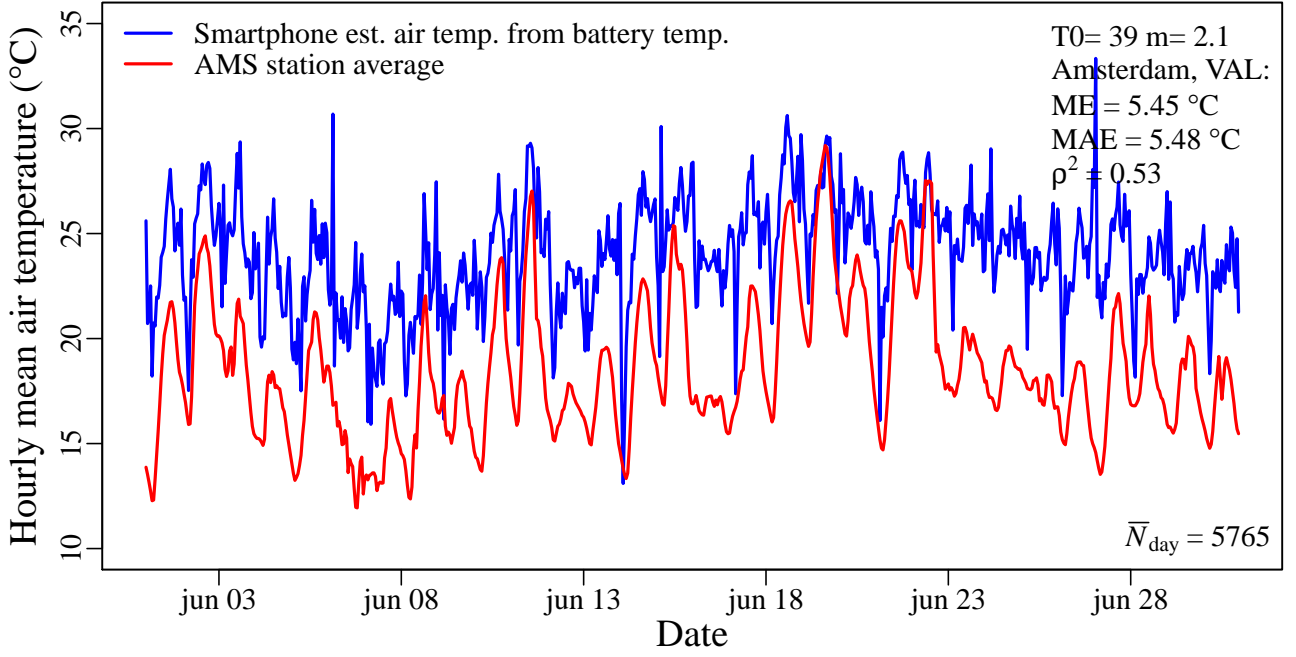
Figure 11: *Base run hourly temperature estimation compared to the average of the AMS stations, June 2017.*

Table 2: *Statistics for the modeled hourly temperature compared to reference (AMS stations), June 2017. Unless stated otherwise, the validation period is June 15 to 30.*

type	m [-]	ρ^2 [-]	MAE [°C]	RMSE [°C]
base run	2.1	0.53	5.45	5.92
only daytime hours 5-19 UTC	2.1	0.75	4.9	5.23
with value m from daily analysis	2.910	0.53	-0.34	2.64
calibrated one m value	2.804	0.42	0.59	2.73
calibrated hourly m values	varying	0.45	0.38	2.44
calibrated hourly m cal period 15-30 June	varying	0.62	-0.29	2.45

m is the ratio of heat conductivity coefficients, one can expect that these may vary throughout the day. In the work of Droste et al. (2017), this improved the model estimations.

Coefficient m varies between 2.5 and 3.2 throughout the day. Between 22 and 2 UTC (0 to 4 AM local time) the values are higher than the rest of the day (Figure A.17 in the appendix). The calibration reduces the bias between the reference and the model, the MAE decreases to less than 2 °C(Table 2, Figure 12). As the calibration period is quite short however, the assignment of the calibration and validation period has a large influence on the resulting model performance; changing the validation period to June 1 to 14, improves the model performance (Table 2). In the latter part of June there is sequence of hot days and subsequent period of overcast, rainy days. The modeled temperature estimation during this period is worse than during the first half of the month.
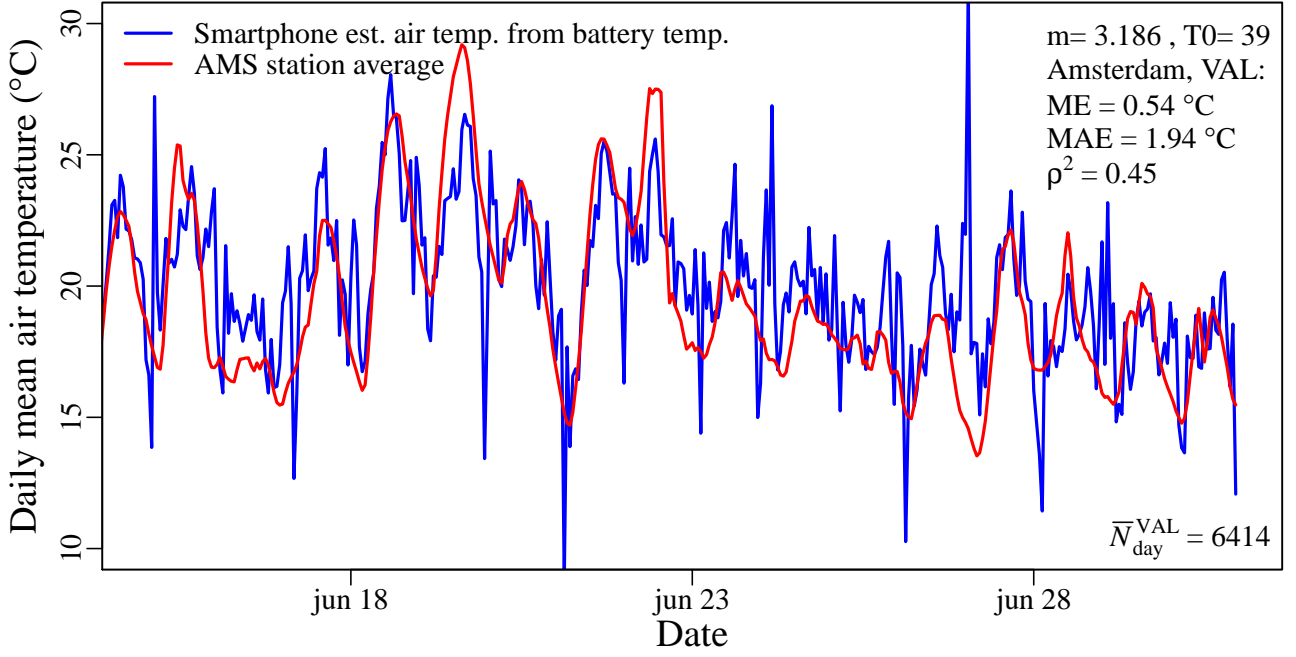
Figure 12: *Modeled hourly temperature and reference for hourly calibrated m values, validation period (June 15 to 30 2017)*

4.2.3 Number of measurements and model performance

As can be seen in Figure 15, the model performance starts to improve considerably when hours with more than 1000 readings are included. The RMSE drops from 2.64 to 2.43 °C. Setting the threshold to 1000 readings per hour excludes the two largest outliers in the dataset. After the threshold of 3250 readings, the model performance increases in smaller steps, however, increasing the threshold does not lead to convergence to a constant performance quality, as was found in Droste et al. (2017). It rather keeps improving the correlation, as less and less hours are included in the model (figure 15). The thresholds initially remove (mostly) the nighttime hours, at the highest thresholds also parts of the days are removed, with more than 40% of hours being left out (Figure 13c). Note that during the night, it is likely that more smartphones are being charged and more phones will be inside than during the day. There is no distinct optimum, but a threshold of 2000 readings per hour would seem a reasonable trade-off between model performance and temporal coverage.

An alternative, similar approach is to not use a threshold but simply use only daytime hours; if only hours between 5 and 19 UTC are included, this also increases the model performance to $\rho^2 = 0.75$ (-) (Table 2). The same holds for generating daily averaged temperatures from daytime hours; only using the hours from 5 to 19 UTC increases the coefficient of determination from $\rho^2 = 0.84$ to $\rho^2 = 0.85$ (-) (Table 3).
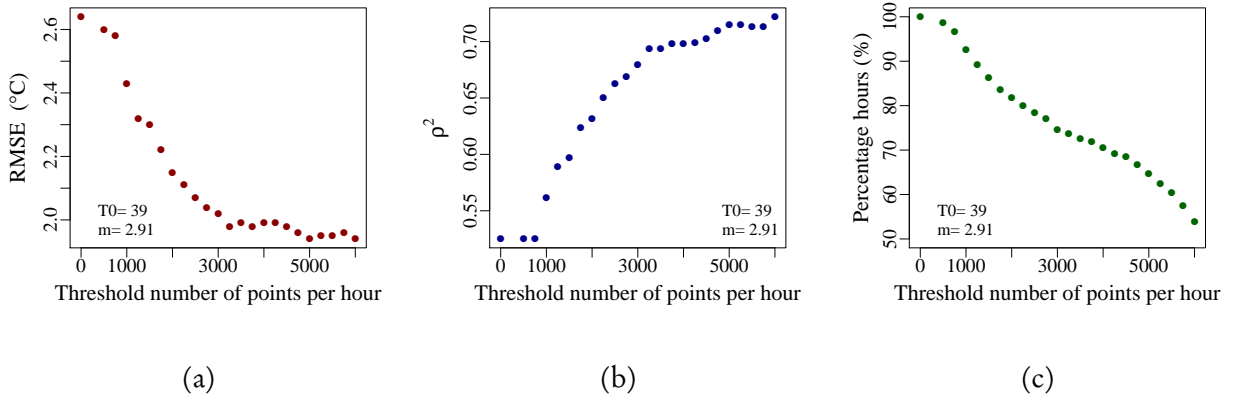
Figure 13: *RMSE (a), ρ^2 (b), and percentage of total hours used in the analysis (c) for thresholds on the number of readings per hour*

4.2.4 Temperature Sensor

A total of 76 thousand readings reported temperatures from a separate thermometer; about 2.5 thousand a day. For 148 hours (21 % of the time) there are no readings and the median number of readings is only 74 per hour. In the nighttime, there are relatively few readings available (Figure ??, Figure 8a). This limited number of readings per hour discourages analysis on an hourly timescale. The minimal hourly temperature is found at 4 UTC (6 AM local time). The readings from the temperature sensor are lower than the battery temperature (Figure 8b).

Averaged to daily mean temperatures, the correspondence with the AMS stations reference is low; ρ^2=0.33 (Figure 14). With an mean temperature of 27.6 °C, the daily averaged temperatures from temperature sensors are substantially higher than those measured at the AMS reference stations: the mean bias (ME) is 9.1 °C. The values are however 4.4 °C lower than the daily averaged battery temperature readings. As applying a linear model to the data will not improve the correspondence with the AMS station reference, no further attempts are made to model the ambient temperature from these readings. It can be concluded that there is some air temperature signal in the ambient temperature readings, but that this signal is much weaker compared to the battery temperature readings.
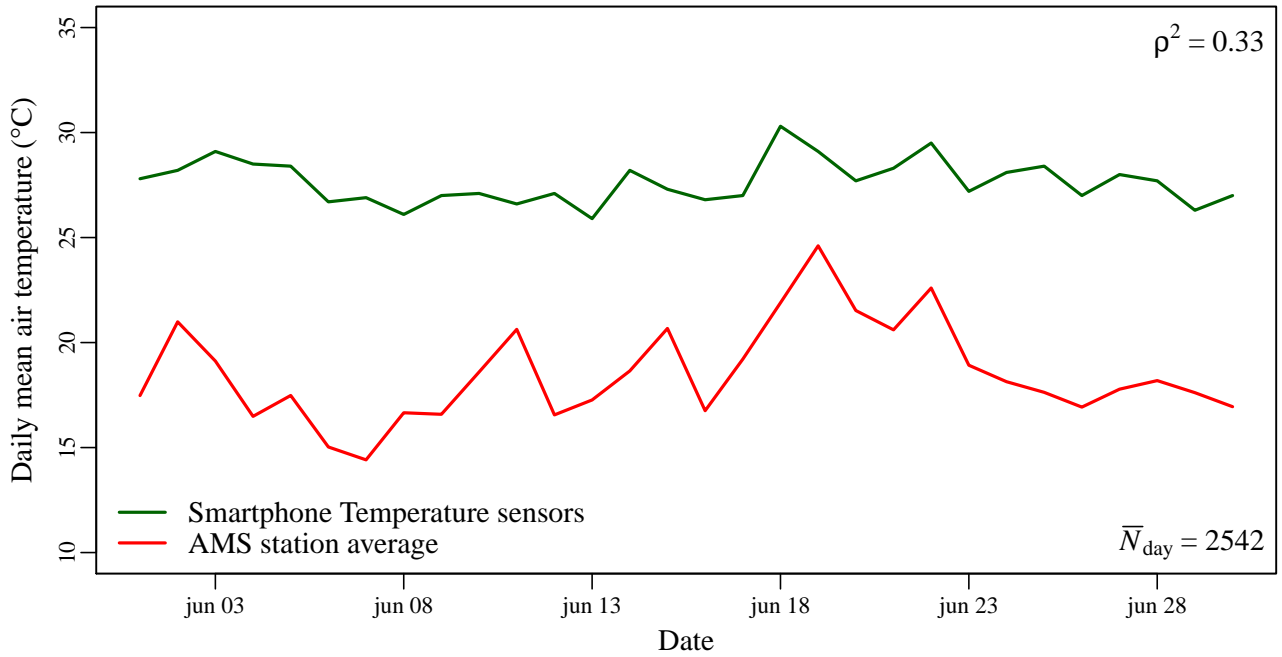
Figure 14: *Daily mean temperature from ambient temperature sensors and measured temperature at the AMS stations in June 2017*

4.3   Light Intensity

For June 2017, there are 3.0 million readings with the highest reported sensor accuracy. There is a distinct diurnal cycle in the readings from the light sensor: For the month of October, the average intensity peaks at 705 Lux at 11 UTC, while it is only 18 Lux at 0 UTC, this peak is 1860 Lux at 11 UTC for June (Figure 16). In October, the mean light intensity increases later, decreases earlier and is lower than in June. This is what one expects of an autumn month compared to summer.

The Global Radiation measured at WMO station Schiphol is two orders of magnitude larger than the mean light intensity captured by the smartphones. The correlation between the smartphone light intensity and the official measurements at WMO station Schiphol is considerable; the coefficient of determination is $\rho^2 = 0.68$ & $\rho^2 = 0.75$ for October and June respectively. To illustrate this, in Figure 16 the mean light intensity from smartphones and global radiation measured at WMO station Schiphol are shown for a couple of days in June. When converting the measured global radiation to lux, the mean light intensity from the smartphones is two orders of magnitude lower than measured at WMO station Schiphol.

The light intensity readings are not normally distributed; it is a strongly right skewed distribution (skewness 17.9 (-)), so the mean is very much shifted by a small number of very high readings.

This can be explained by the behavior of the smartphone users, as many people will be indoors most of the time, the majority of light intensity readings will not represent outside conditions. Therefore it is interesting to look at the extreme values found in the readings; in Figure 15b the 99th percentile of light intensity readings is shown per hour of day. For these readings, the smartphone user could have the phone in hands while being in broad daylight. At midday on some of the sunny clear days in June 2017, the measured global radiation surpassed 900 Wm^{-2}, which
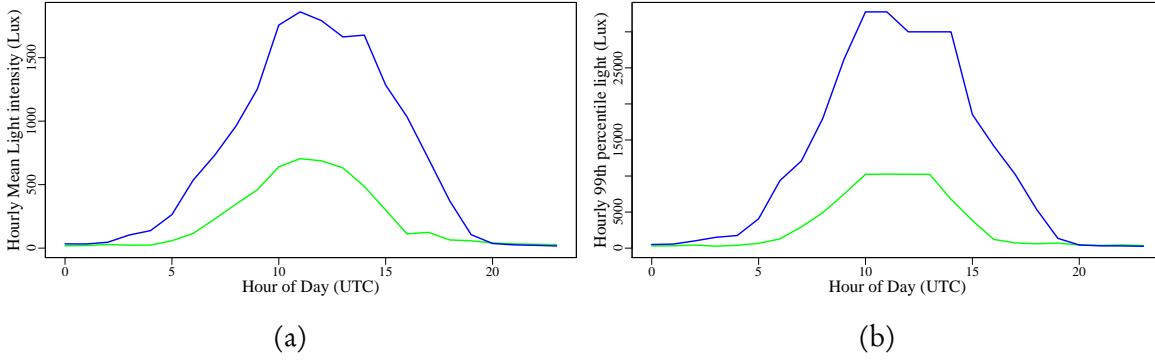
Figure 15: *Mean (a) and 99th percentile (b) of light intensity readings per hour of day in June 2017 (blue) and October 2016 (green).*

Table 3: *Statistics for different light sensor thresholds applied to calculate the daily temperature, June 2017 compared to AMS station average*

type	ρ^2 [-]	MAE [°C]	RMSE [°C]
base run	0.84	4.70	4.79
all below median lux removed	0.84	4.88	4.96
all below mean lux removed	0.79	4.82	4.96
base run 5-19 UTC	0.85	4.33	4.19
all below 100 lux removed 5-19 UTC	0.84	4.10	4.24
all below 500 lux removed 5-19 UTC	0.82	4.44	4.61

would correspond to about 1.14×10^5 Lux. The readings are still a factor four lower than found from the calculation from global radiation. This can partially be explained by the wavelengths of light which are captured by the different sensoring techniques; smartphone sensors are sensitive to a smaller range of light wavelengths. Again, the user context will also have an influence; the angle at which the smartphone is held, the orientation of the user to the Sun will also contribute to the differences between the maximal light intensity from WMO station Schiphol and the values found in the smartphone readings.

Filtering on light sensor readings

Now the ability of light sensor readings to remove indoor battery temperature readings is investigated. As only using the top percentiles of light intensity readings, would obviously vastly reduce the number of available readings, the thresholds are set lower. Per hour the mean and median light intensity are calculated. These threshold therefore vary throughout the day and throughout the month. Readings with a smaller light intensity are filtered out and with the remaining readings the daily averaged battery temperature is calculated. As seen in Table 3, none of the tested light intensity filters improved the model performance. The light sensor readings may be more susceptible to differences in cloudiness and the time of day rather than if the phone is inside or outside.
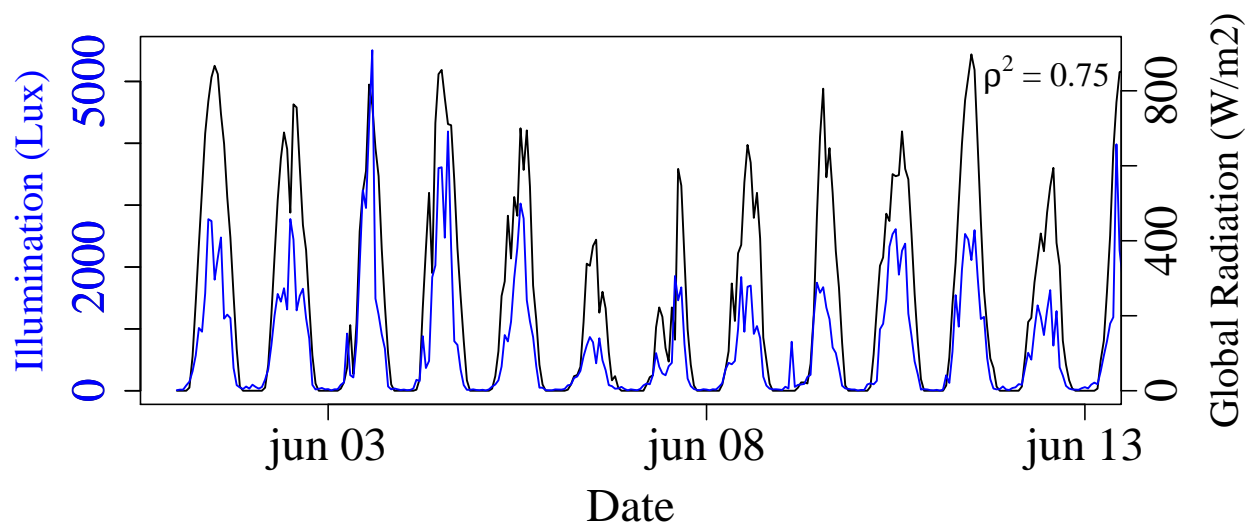
Figure 16: *Mean light intensity from smartphone readings (blue) and global radiation (black) at Schiphol 1 to 13 June 2017*

5   Discussion

5.1   Relation with other studies

The results for Amsterdam in June are comparable to those found by Droste et al. (2017) for São Paulo; ρ^2 =0.84 versus 0.86 (for the whole year of 2014). The bias is slightly smaller, ME = 0.19 versus -0.53 °C, as this study optimized the value for m for one summer month rather than for a whole year. The number of available readings per day is about ten times as high in this study compared to Droste et al. (2017). However, in this study only the city-wide temperature is considered, while for São Paulo distinctions are made for different neighborhood types ("local climate zones" (Stewart and Oke, 2012)) (Droste et al., 2017).

Compared to the previous study for Amsterdam, the temperature estimations presented here perform better; the fraction of explained variance is higher ρ^2 =0.84 versus 0.67 (-). This can mostly be attributed to the vast increase in the density of battery temperature readings; in Overeem et al. (2014) about 1×10^2 measurements were available per day, compared to over 1×10^5 for this study. The bias however diminishes only slightly; ME= 0.19 versus 0.39 °C. These results are in agreement with the findings of Overeem et al. (2014): the quality of temperature estimations (expressed in terms of fraction of explained variance) increases for increasing reading density, but with diminishing returns. For daily temperature estimations a 'plateau' seems to be reached in terms of number of readings, as the fraction of explained variance is not larger than in Droste et al. (2017). Although, the differences in sampling techniques probably also play role in this (Section 5.5).

Droste et al. (2017) found a delay of up to 4 hours for the response of the battery temperature readings to the station-measured temperature. This observed effect was explained as a combination of factors. Firstly, with a simplified differential equation it was reasoned that a typical smartphone would arrive at a steady heating or cooling rate after about 1 to 2 hours due to its heat capacity. Secondly, as a part of the readings are taken indoors this would create additional time delays; the inside temperature reacts to the temperature outdoors with a delay and a dampening of the temperature amplitude. However, such delays are observed at some points in the June dataset at the cooling, but not consistently. This is a remarkable result, as the physical explanation why one would expect a time delay would seem sound. Droste et al. (2017) found similar improvements in modeled hourly temperature results from installing a four hour time delay and from calibrating separate values for m per hour of day. Only the latter analysis was repeated here.It should furthermore be noted that the applied heat transfer model assumes stationarity, which is violated when applied on hourly timescales.

This study uses a very robust reference; a network of authoritative urban meteorological stations. By using 23 stations throughout the city, the employed reference is more representative of the environmental context where the smartphone readings were collected, compared to using a meteorological station at the city's airport. It is also more robust method than comparing two sources of crowdsourced data as done in Niforatos et al. (2017) which used smartphones and hobby weather stations data.

Like in Niforatos et al. (2017), this study could use more than just readings from the battery temperature sensors. Here, only one city is considered, the study period is shorter, less smartphone

sensors are explored, but the density and sheer number of readings is much higher. Due to the popularity of the OpenSignal app, the datasets used in this study contained the roughly the same amount of useful readings for just two days as the study of Niforatos et al. (2017) had available for a period of 32 months worldwide. Niforatos et al. (2017) also found a diurnal cycle in the light sensor readings. However, in their research two distinct peaks were observed; at 10:00 and 15:00 local time. The reason behind this is not known. The researchers also found a high standard deviation in the light intensity values. In their case this was attributed to the large study domain encompassing different latitudes and timezones. A significant effect between the collected average light intensity readings and the actual weather conditions was found for two distinct weather conditions; "Clear" and "Overcast" (Niforatos et al., 2017). This statistical analysis was not repeated in this research, a limitation of this work and a recommended potential extension. These findings pinpoint another use for the light intensity readings; instead of using them as filter for the battery temperature readings, one could use these readings for detecting cloudiness.

Niforatos et al. (2017) applied supervised machine learning to predict the current and future ambient temperature from manually submitted weather data and mobile sensor data collected through their mobile app. Multiple input variables were used to estimate the air temperature, rather than just one variable in a heat transfer model, as done in this research. The variables under consideration in Niforatos et al. (2017) are not all present in the datasets available for this research; readings from the proximity sensor, accelerometer and magnetometer are not included. This is a limitation of the research presented in this report; no repetition of the analysis of Niforatos et al. (2017) could be performed. Applying machine learning techniques on the all variables which *are* included in the dataset would not be a fruitful strategy, especially for the October 2016 dataset, which included readings from 135 different variables. These variables were nearly all related to specific network properties of the smartphones. These variables do not have a purpose which could make them interesting for meteorological applications. Nonetheless, performing additional statistical analyses between the available readings might elucidate relationships between for example the battery temperature and light sensor readings which are not uncovered in this research. This is recommended for future work.

5.2   Differences in results between October 2016 and June 2017

As 28% of smartphone readings in the June 2017 dataset were from plugged phones being charged, it is reasonable to assume a similar percentage of readings stem from smartphones being charged for October 2016. This is presumed a factor for the poorer model results in October, as demonstrated by including batteries being charged in the June dataset (Table 1). One likely supplementary explanation for the relatively poor model performance in October would be the behavior of smartphone users; people tend to stay indoors more when it is colder outside, and the indoor air temperature is more stable than the ambient temperature (as also argued in Overeem et al. (2013)) and subject to indoor heating systems. The finding that average hourly battery temperatures drop around the time of the evening commute, when more people are outside, adds weight to this interpretation (Figure 5b), such a temperature drop was also found in Niforatos et al. (2017). Also in the work of Overeem et al. (2013), for five out of eight cities model results in spring – summer outperformed

those in autumn – winter months (although for the cities of Moscow and Paris this is likely to be due the used calibration method for the value of m in a period with a large temperature gradient).

5.3   Temperature sensor

It is highly likely that readings from the separate ambient temperature sensor are not 'raw' measured temperatures from the sensors, but rather internally processed readings to compensate for heat sources within the smartphone. This is certainly the case for the smartphone model Samsung Galaxy S4, but similar compensation algorithms may also be in place in other smartphone models (Robinson, 2013). Such a compensation algorithm may influence the results. Readings from the temperature sensor and battery temperature sensor are not independent from each other. The applied algorithm used for this is not known, but applying statistical methods like principle component analysis might shed light on this. So this may be useful to apply in future research on this sensor. However, smartphones with a separate ambient temperature sensor have not 'caught on'; not many smartphone models contain this sensor (as seen in this research; only 51 of 3199 smartphones had this sensor).

5.4   Heat transfer model

The applied heat transfer model assumes that smartphones are typically kept in pockets of clothing of the user. However, as the sampling strategy of OpenSignal collects readings when the screen is on, the phone is most likely held in hand when a reading is collected. In this situation, the smartphone is still influenced by both the ambient temperature and the body temperature through the hand. It can be assumed that the smartphone is kept in a pocket, taken out, used and stored in a pocket again. Therefore the assumptions underlying the heat transfer model are still valid under this sampling strategy.

5.5   Sampling Strategy

The sampling strategy applied by OpenSignal is to collect readings every 15 seconds when the screen is on, which is undoubtedly useful for the service the app provides. On the other hand, for weather applications, this is not necessarily the optimal sampling strategy, as this generates a dataset with sets of many readings from one location and time. The readings are not independent of each other, this bloats the dataset without adding to the representativeness of the readings for the city as a whole. Resampling the dataset to only contain, for instance, one (averaged) reading per smartphone per hour, might compensate for this dataset bloating.

In this study, attention is given to the number of readings, rather than to the number of unique smartphones per unit of time, another limitation of this study. The unique identification codes of the smartphones are included in the datasets available for this research. Analyzing the number of unique smartphones providing readings per unit of time would be a step forward. From this, the number of active smartphone app users needed to attain good temperature estimations could be inferred. Now only the relation between the absolute number of readings and the model performance is researched. Additionally, time-series analysis of separate unique smartphones could be performed

to show the possible patterns of heating and cooling of the battery, or to, for instance, see patterns in the light sensor readings. Information from which unique smartphone a reading stemmed, was not available in the work of Overeem et al. (2013) and Droste et al. (2017). This type of analyses does raise privacy issues as following the 'behavior' of individual smartphone users without their explicit consent can be a privacy infringement.

An app like OpenSignal does not necessarily collect information from all smartphone sensors and variables which could be of interest for meteorological applications or for determining the user's context. With the dedicated app, Niforatos et al. (2017) could for example collect readings from the proximity sensor (measured in cm);installed to detect if the user is holding the smartphone to the ear and thereby prevent false touchscreen activity. The proximity is influenced by the smartphone usage and could help detect when a smartphone is exposed to outside conditions (Niforatos et al., 2017).

To have more control over the sampling strategy, one could thus use dedicated apps such as the *Atmos* app of Niforatos et al. (2017). Here there is a trade-off between quantity and quality; the OpenSignal apps have up to 10 million downloads, whilst the dedicated *Atmos* app has about one thousand downloads (Google Play Store, 2017). The number of ardent weather enthusiasts is apparently much more limited. This pinpoints an omnipresent challenge in crowdsourcing; 'opportunistic sensing' has a large potential as the potential number of users and therefore the number of measurements is largest, but the applicability is less than for more dedicated forms of crowdsourcing and citizen science. An additional limitation when crowdsourcing data, is that the use of apps is subject to the whims of the public; people install an app, use it shortly and, over time, lose interest, ignore it or even uninstall it, as found in Niforatos et al. (2017). This capriciousness can limit the application of this datasource to serving as a supplementary datasource. Or to use in campaigns with a limited duration for specific goals (like the mPING app). For continuous operational monitoring, using smartphone apps may thus not be the most realistic option; certain apps may only be popular for a limited period of time.

6 Conclusions and perspective

In this study a much larger and denser dataset of smartphone readings was available than in previous work. Also, for this study a good reference was available in the form of an urban meteorological network. Readings from the smartphone battery temperature sensors contain a signal reflecting the ambient temperature. For Amsterdam the city-wide daily temperature can thus be estimated from smartphone battery temperature readings using a straightforward heat-transfer model. This model needs to be calibrated; using the value for model coefficient m calibrated for the same city during a similar, but longer time of year, resulted in a large warm bias. The necessity of calibration prevents this method to be immediately applicable to locations without a good reference network.

The signal in the battery readings shows larger correspondence with the measured temperature for June 2017 than for October 2016, largely because for the latter the dataset contains readings from smartphones being charged and due to the season causing more readings to come from indoors. Nonetheless the fraction of explained variance is $\rho^2 = 0.5$.

This is the second study in which the city-wide temperature is estimated per hour from battery temperature readings, and the first in which the relation between the number of smartphone readings and the model performance of temperature estimations has been analyzed. This showed that it is mostly during the nighttime hours, when less readings are available, that the model performance is poorer. However, it remains a question how many *unique smartphones* would be needed to obtain good temperature estimations. Such an analysis is recommended for future research.

In the readings from the separate ambient temperature sensors the signal correspondence is limited. These readings are therefore not a good information source for estimating the urban temperature. This is in part due to the limited number of smartphones models which have this sensor, which leads to substantially less available readings; only 51 unique phones occurred in the June 2017 dataset.

Readings from the light sensor show a distinct diurnal pattern. The signal in the mean hourly light intensity corresponds well with the measured global radiation at WMO station Schiphol. However, the attempt to use these readings to filter the battery temperature readings to include mostly outside readings and to exclude more inside readings did not improve the temperature estimations. In future research, these readings could be used for detecting clouds or cloudiness.

It is likely that many smartphone readings are collected inside of buildings, this remains an important drawback of the applied method. Building on the work of Niforatos et al. (2017), it may be useful to develop a system to automatically log readings from sensors when sensed that smartphone is outside. Or to monitor the change in readings over time and only log when for instance the temperature is stable or when there is a large change in the temperature.

References

Bell, S., Cornford, D., and Bastin, L. (2013). The state of automated amateur weather observations. *Weather*, 68(2):36–41.

Broyde, Y., Livschitz, M., and Messer, H. (2013). Adaptive statistical learning of cellular users behavior. *Signal Processing*, 93:3151–3158.

Chapman, L., Bell, C., and Bell, S. (2016). Can the crowdsourcing data paradigm take atmospheric science to a new level. a case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology*.

Cumulus forum contributors (2010). Cumulus forum; converting lux to watts per square meter. `http://sandaysoft.com/forum/viewtopic.php?t=3979`. last accessed 13-11-2017.

De Bruijn, E. I. F., De Haan, S., Bosveld, F., and Heusinkveld, B. (to be submitted). Measuring low altitude winds with a hot-air balloon and its validation with Cabauw tower observations.

Droste, A. M., Pape, J. J., Overeem, A., Leijnse, H., Steeneveld, G. J., Van Delden, A. J., and Uijlenhoet, R. (2017). Crowdsourcing urban air temperatures through smartphone battery temperatures in São paulo, brazil. *Journal of Atmospheric and Oceanic Technology*.

Elmore, K. L., Flamig, Z. L., Lakshmanan, V., Kaney, B. T., Farmer, V., Reeves, H. D., and Rothfusz, L. P. (2014). mping: Crowd-sourcing weather reports for research. *Bulletin of the American Meteorological Society*, 95(9):1335–1342.

Google Play Store (2017). Google Play Store. `https://play.google.com/store`. Accessed: 2017-11-12.

Heusinkveld, B. G., Steeneveld, G. J., Hove, L. W. A. v., Jacobs, C. M. J., and Holtslag, A. A. M. (2014). Spatial variability of the Rotterdam urban heat island as influenced by urban land use. *Journal of Geophysical Research: Atmospheres*, 119(2):677–692.

Homan, C. (2016). Oktober 2016; vrij koud, vrij droog en zonnig. *KNMI*, 1.

Homan, C. (2017). Juni 2017, zeer warm, zonnig en gemiddeld over het land normale hoeveelheid neerslag. *KNMI*, 2.

KNMI (2014). Knmi 14 klimaatscenarios:scenarios samengevat. `http://www.klimaatscenarios.nl/scenarios_samengevat/index.html`.

Le Tertre, A., Lefranc, A., Eilstein, D., Declercq, C., Medina, S., Blanchard, M., Chardon, B., Fabre, P., Filleul, L., Jusot, J. F., et al. (2006). Impact of the 2003 heatwave on all-cause mortality in 9 French cities. *Epidemiology*, 17(1):75–79.

Madaus, L. E. and Mass, C. F. (2017). Evaluating smartphone pressure observations for mesoscale analyses and forecasts. *Weather and Forecasting*, 32(2):511–531.

Meier, F., Fenner, D., Grassmann, T., Otto, M., and Scherer, D. (2017). Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate*, 19:170–191.

Molenaar, R. E., Heusinkveld, B. G., and Steeneveld, G. J. (2016). Projection of rural and urban human thermal comfort in the netherlands for 2050. *International Journal of Climatology*, 36(4):1708–1723.

Muller, C. L., Chapman, L., Grimmond, C. S. B., Young, D. T., and Cai, X. (2013). Sensors and the city: a review of urban meteorological networks. *International Journal of Climatology*, 33(7):1585–1600.

Muller, C. L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., and Leigh, R. R. (2015). Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology*, 35(11):3185–3203.

Niforatos, E., Vourvopoulos, A., and Langheinrich, A. (2017). Understanding the potential of human-machine crowdsourcing for weather data. *International Journal of Human-Computer Studies*, 102(Supplement C):54 – 68. Special Issue on Mobile and Situated Crowdsourcing.

Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 108(455):1–24.

Overeem, A., R., R. J. C., H., L., J., S. G., P., H. B. K., and R., U. (2014). Stadstemperatuur in beeld met smartphones. *Meteorologica*, 23:16–18.

Overeem, A., Robinson, R. J. C., Leijnse, H., Steeneveld, G. J., Horn, P. B. K., and Uijlenhoet, R. (2013). Crowdsourcing urban air temperatures from smartphone battery temperatures. *Geophysical Research Letters*, 40(15):4081–4085.

Robinson, J. (2013). The SHTC1: Inside the chip that powers WeatherSignal. `https://opensignal.com/blog/2013/06/19/the-shtc1-inside-the-chip-that-powers-weathersignal/`.

Ronda, R., Steeneveld, G. J., Attema, J., Heusinkveld, B., and Holtslag, A. A. M. (2014). Summer in the city - high resolution modelling and validation of urban weather and human thermal comfort. In *EMS Annual Meeting Abstracts*, volume 11.

Saltus, R. and Nair, M. (2017). The crowdmag app-turning your smartphone into a travelling magnetic observatory. In *EGU General Assembly Conference Abstracts*, volume 19, page 9916.

Samsung,2013. Samsung what you may not know about Galaxy S4 innovative technology.

Schweizer, I., Bärtl, R., Schulz, A., Probst, F., and Mühläuser, M. (2011). Noisemap-real-time participatory noise maps. In *Second International Workshop on Sensing Applications on Mobile Phones*, pages 1–5.

Snik, F., Rietjens, J. H. H., Apituley, A., Volten, H., Mijling, B., Di Noia, A., Heikamp, S., Heinsbroek, R. C., Hasekamp, O. P., Smit, J. M., et al. (2014). Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters. *Geophysical Research Letters*, 41(20):7351–7358.

Steeneveld, G. J., Koopmans, S., Heusinkveld, B. G., Van Hove, L. W. A., and Holtslag, A. A. M. (2011). Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the Netherlands. *Journal of Geophysical Research: Atmospheres*, 116(D20).

Steeneveld, G. J., O., K. J., A., G. R. J., and Holtslag, A. A. M. (2016). An urban climate assessment and management tool for combined heat and air quality judgements at neighbourhood scales. *Resources, Conservation and Recycling*.

Stewart, I. D. and Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12):1879–1900.

Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M., et al. (2013). IPCC, 2013: summary for policymakers. *Climate change*, pages 3–29.

Thompson, J. E. (2016). Crowd-sourced air quality studies: A review of the literature and portable sensors. *Trends in Environmental Analytical Chemistry*, 11:23 – 34.

United Nations Department of Economic and Social Affairs (2014). World urbanization prospects. `https://esa.un.org/unpd/wup/Publications/Files/WUP2014-Highlights.pdf`.

Vanini, S., Faraci, F., Ferrari, A., and Giordano, S. (2016). Using barometric pressure data to recognize vertical displacement activities on smartphones. *Computer Communications*, 87:37–48.

Vos, L. d., Leijnse, H., Overeem, A., and Uijlenhoet, R. (2017). The potential of urban rainfall monitoring with crowdsourced automatic weather stations in amsterdam. *Hydrology and Earth System Sciences*, 21(2):765–777.

Wageningen University (2016). Persbericht: De amsterdam atmospheric monitoring supersite gaat van start. `http://www.wur.nl/nl/nieuws/De-Amsterdam-Atmospheric-Monitoring-Supersite-gaat-van-start.htm`.

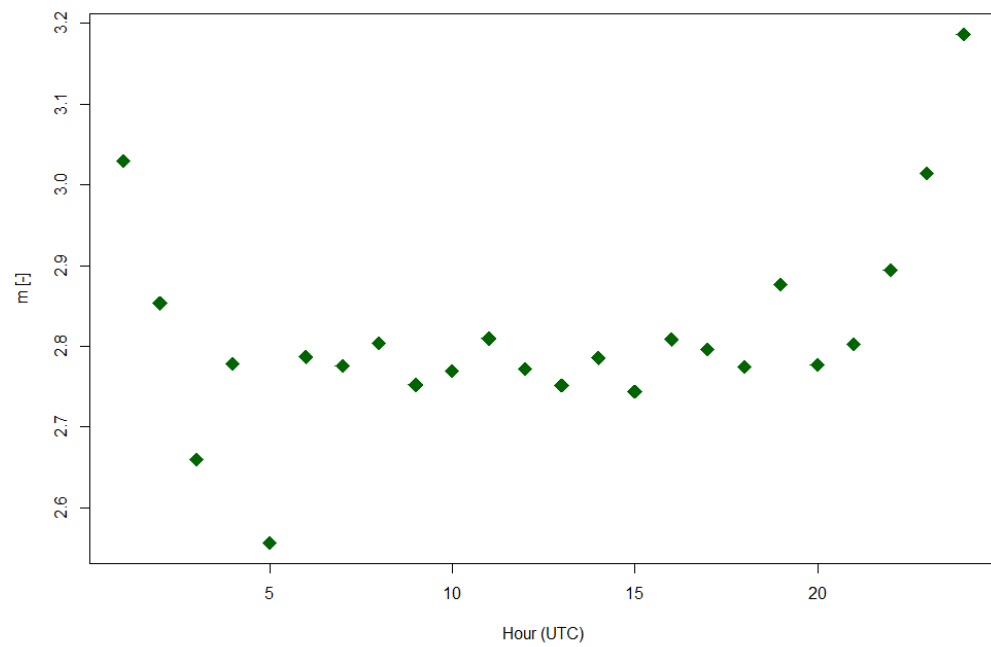Warwick University. Warwick University office lighting guidance.

Appendix

Figure A.17: *Calibrated values of m per hour of day, calibrated for 1 to 14 June 2017*