

Assessing Near Surface Variability with a Wireless Sensor Network on the Small Scale

DISSERTATION
ZUR ERLANGUNG DES DOKTORGRADES
DER NATURWISSENSCHAFTEN IM FACHBEREICH
GEOWISSENSCHAFTEN
DER UNIVERSITÄT HAMBURG

VORGELEGT VON

KATHARINA LENGFELD

AUS

Bad Oldesloe

Hamburg
2012

Als Dissertation angenommen vom Fachbereich Geowissenschaften der Universität
Hamburg

Auf der Grund der Gutachten von Jun.-Prof. Dr. Felix Ament
und Prof. Dr. Bernd Leitl

Hamburg, den 10.01.2012

Prof. Dr. Jürgen Oßenbrügge
Leiter des Fachbereichs Geowissenschaften

Contents

List of Abbreviations	I
Abstract	III
Zusammenfassung	V
1 Introduction	1
2 Sensor Network	5
3 Characterisation of the Sensor Network	11
3.1 Humidity and Temperature Sensors	11
3.2 Infrared Thermometers	16
3.3 Wind Speed and Direction Sensors	17
3.4 Solar Radiation Sensors	19
3.5 Precipitation Sensors	20
3.6 Soil Moisture and Temperature Sensors	22
3.7 Profile Measurements	23
4 FLUXPAT 2009	27
4.1 Deployment	28
4.2 Variability	32
4.3 Explaining Factors for the Variability	40
4.4 Cluster Analysis	46
4.5 Comparison with FOOT3DK Model	57
4.6 Comparison with TERRA Standalone Model	68
5 Hamburg Airport	75
5.1 Deployment	75
5.2 Variability	80
5.3 Explaining Factors for the Variability	88
5.4 Cluster Analysis	93
5.5 Comparison with DWD Data	101
6 Comparison and Conclusion	109

7 Outlook	115
A Regression Coefficients	117
B FLUXPAT2009	121
C Technical Problems	123
Literature	127
Acknowledgement	133
Affirmation	135

List of Abbreviations

adp	Air Dryness Point
ALB	Albedo
c_H	Absolute Heat Capacity
COND	Conductivity
COSMO	Consortium for Small Scale Modelling
c_p	Specific Heat Capacity
D	Hydraulic Diffusivity
Δq_{90-10}	Quantile Range (Difference between 10% and 90% Quantile)
dr	Distance to the River
DWD	Deutscher Wetterdienst (German Meteorological Service)
E	Saturated Water Vapour Pressure
EPFL	Ecole Polytechnique Federale de Lausanne
ϵ	Infrared Emissivity
f	Vertical Part of the Coriolis Parameter
FOOT3DK	Flow Over Orographically Structured Terrain - 3 Dimensional (Kölner Version)
F_{SM}	Soil Moisture Flux
F_{SW}	Shortwave Radiation
FSW	Shortwave Radiation Sensor
FLUXPAT	Fluxes and Patterns in the Soil-Vegetation-Atmosphere System
g	Gravitational Constant
H	Sensible Heat
I	Infrared Radiation
k	Kurtosis
K	Hydraulic Conductivity
LAI	Leaf Area Index
LE	Latent Heat
LUC	Land Use Class
LUCE	Lausanne Urban Canopy Experiment
m	Mean Difference
MAC	Media Access Control
μ	Mean
MSR	Minimal Stomatal Resistance
ω	Angular Velocity

p	Pressure
φ	Latitude
π	Exner Pressure
pv	Pore Volume
q	Specific Humidity
Q	Saturated Specific Humidity
q10	10% Quantile
q90	90% Quantile
R	Reference Station
RAIN	Rain Sensor
RC	Radiance Coefficient
RH	Relative Humidity
R_L	Gas Constant for Dry Air
RMSE	Root Mean Square Error
RMSD	Root Mean Square Distance
RR	Rain Rate
s	Skewness
SEAL	Sealing
σ	Standard Deviation
σ_B	Stefan Boltzmann Constant
SM	Soil Moisture and Temperature Sensor
T_{Air}	Air Temperature
θ	Potential Temperature
TIR	Infrared Surface Temperature Sensor
TransRegio32	TransRegional Collaborative Research Centre 32
TRH	Temperature and Humidity Sensor
T_{Surf}	Surface Temperature
v	Wind Speed
VEG	Vegetation
WIND	Wind Sensor
z_0	Roughness Length

Abstract

In this study, atmospheric conditions near the surface, e.g. air temperature, specific humidity, surface temperature and wind speed, are investigated using a low cost wireless sensor network. A sensor intercomparison demonstrates that the measurements of the network, although not highly accurate, are consistent. Thus, it is an appropriate tool to study the small-scale spatial and temporal variabilities in the near surface atmosphere.

The influence of the heterogeneous surface on the atmospheric conditions is examined by deploying the network in a heterogeneous and a homogeneous area. The first campaign took place in western Germany as part of the project FLUXPAT2009 in August 2009. The deployment consisted of 13 stations set up as a 2.3 km double transect covering various surface types, including a small river. The latter campaign took place at Hamburg Airport in April and May 2010. Twelve stations were deployed over homogeneous grassland along both runways.

At the airport, the stations in the northern, more rural, environment are colder than the stations in the southern urban region, leading to a mean difference in air temperature of 0.5 K across a distance of 2.7 km. At the FLUXPAT site, the river is the predominant source of influence on air temperature. Stations close to the river are colder than stations far away from the river. The inter-station variation is distinctly higher with a difference of 0.8 K across a distance of only 0.8 km. The variability in air temperature is mainly influenced by the wind speed for both sites. For specific humidity, the mean difference between the driest and the wettest station at the heterogeneous site is five times the mean difference at the homogeneous site. The driving factors are wind speed and the distance to open waters (the river at the FLUXPAT site and a water reservoir at the airport). The two sites are particularly distinguishable through surface temperature. Different vegetation types at the heterogeneous site cause a mean difference of more than 3 K across a distance of 1.9 km. At the homogeneous grassland site, the mean difference is three times smaller and is primarily influenced by wind speed. Wind speed at a height of 2 m is unaffected by land-use heterogeneities, but is influenced by nearby obstacles.

Setting up a network of stations in a relatively small area is not always feasible. Cluster analysis identifies stations that are more or less redundant in both set-ups. For air temperature, the stations at the FLUXPAT site are clustered according to their distance to river; at the airport, according to their position (in the north, at the centre or in the south of the airport). The clustering of surface temperature is similar to air temperature at the airport. At the heterogeneous site, however, it depends on the land-use type. An estimation of

the error made by considering the effect of a reduced quantity of stations is given. For example, using only one station for observing air temperature results in an error of 0.9 K at the heterogeneous site and 0.7 K at the homogeneous site.

The observations carried out during the FLUXPAT campaign are compared to air temperature and specific humidity simulations of the mesoscale atmospheric model FOOT3DK. This comparison indicates that the FOOT3DK model simulates either air temperature or specific humidity satisfactorily at the lowest model level depending on the dominating land use classes at each grid cell. The influence of heterogeneous forcing and vegetation on heat flux modeling is studied using the soil-vegetation-atmosphere transfer scheme TERRA. The measurements of the FLUXPAT campaign are used as input for four different runs with homogeneous and heterogeneous forcing and vegetation. Heterogeneous vegetation reduces the bias between the stations, heterogeneous forcing reduces the variability for each station.

Zusammenfassung

In dieser Arbeit werden bodennahe atmosphärische Größen, z.B. Lufttemperatur, spezifische Feuchte, Oberflächentemperatur und Windgeschwindigkeit, mit Hilfe eines kostengünstigen, drahtlosen Sensornetzwerks untersucht. Ein Vergleich der verwendeten Sensoren zeigt, dass die Messungen des Netzwerks zwar nicht exakt, aber konsistent sind. Folglich ist das hier verwendete Netzwerk ein probates Mittel um kleinskalige räumliche und zeitliche Variabilitäten in der bodennahen Atmosphäre zu beobachten.

Der Einfluss von heterogenen Landoberflächen auf die atmosphärischen Größen wird untersucht, indem das Netzwerk in einem heterogenen und einem homogenen Gebiet aufgebaut wird. Die erste Messkampagne fand im Westen Deutschlands als Teil des FLUXPAT2009 Projekts im August 2009 statt. Das Netzwerk bestand aus 13 Stationen, die als 2.3 km langer Doppeltransekt angeordnet wurden. Dieser Transekt umfasste verschiedene Landoberflächen und einen kleinen Fluss. Die zweite Kampagne fand am Flughafen Hamburg im April und Mai 2010 statt. Zwölf Stationen wurden auf homogenen Wiesenflächen entlang der beiden Startbahnen aufgebaut.

Die Stationen im Norden des Flughafens befinden sich in eher ländlicher Umgebung und sind kälter als die Stationen im südlichen städtischen Bereich. Das führt zu einer mittleren Temperaturdifferenz von 0.5 K auf einer Distanz von 2.2 km. Im FLUXPAT Gebiet wird die Lufttemperatur hauptsächlich von dem Fluss beeinflusst. Stationen in Flussnähe sind kälter als weit entfernte Stationen. Die Variation in der Lufttemperatur ist deutlich größer mit einer mittleren Differenz von 0.8 K auf einer Distanz von nur 0.8 km. Die Variabilität ist bei beiden Messkampagnen stark von der Windgeschwindigkeit beeinflusst. Für spezifische Feuchte ist die mittlere Differenz zwischen der trockensten und der feuchtesten Station im heterogenen Gebiet fünfmal so groß wie im homogenen Gebiet. Die entscheidenden Einflussfaktoren sind die Windgeschwindigkeit und die Gewässer (der Fluss im FLUXPAT Gebiet und ein Löschwasserbecken am Flughafen). Der Unterschied zwischen beiden Messgebieten ist jedoch für die Oberflächentemperatur am deutlichsten. Unterschiedliche Vegetation im heterogenen Gebiet führt zu einer mittleren Differenz von über 3 K auf einer Distanz von 1.9 km. Auf dem homogenen Wiesenland am Flughafen ist die mittlere Differenz nur ein Drittel davon und wird hauptsächlich von der Windgeschwindigkeit beeinflusst. Die Windgeschwindigkeit wird von der Heterogenität in der Landnutzung nicht beeinflusst, aber durch nahe gelegene Hindernisse.

Ein Netzwerk von Stationen in einem relativ kleinen Gebiet aufzubauen ist oft nicht möglich. Daher werden mit Hilfe von Clusteranalyse in beiden Messaufbauten Stationen

bestimmt, die mehr oder weniger überflüssig sind. Für die Lufttemperatur werden die Stationen im FLUXPAT Gebiet gemäß ihres Abstands zum Fluss in Gruppen eingeteilt, am Flughafen gemäß ihrer geographischen Lage (im Norden, im Zentrum oder im Süden des Flughafens). Die Gruppeneinteilung für Oberflächentemperatur ist am Flughafen ähnlich der für Lufttemperatur. In dem heterogenen Gebiet basiert sie jedoch auf der unterschiedlichen Vegetation. Außerdem wird eine Schätzung des Fehlers vorgenommen, der entsteht, wenn die Anzahl der Stationen verringert wird. Wird für Lufttemperatur nur eine Station genutzt, beträgt der Fehler 0.9 K im heterogenen und 0.7 K im homogenen Gebiet.

Die Messungen der FLUXPAT Kampagne werden mit Simulationen der Lufttemperatur und spezifischen Feuchte des mesoskaligen Atmosphärenmodells FOOT3DK verglichen. Dieser Vergleich deutet darauf hin, dass das Modell im untersten Level nur entweder die Lufttemperatur oder die Feuchte zufriedenstellend simulieren kann, abhängig von der dominierenden Landnutzungsklasse in der Gitterzelle. Der Einfluss von heterogenem Antrieb und Vegetation auf die Modellierung der Wärmeflüsse wird mit Hilfe des Boden-Vegetation-Atmosphären Transferschemas TERRA untersucht. Die Messungen der FLUXPAT Kampagne werden als Antrieb für vier verschiedene Modellläufe mit heterogenem und homogenem Antrieb und Vegetation verwendet. Heterogene Vegetation verringert die Abweichungen der Mittelwerte zwischen den Stationen, heterogener Antrieb verringert die Variabilität an jeder Station.

Chapter 1

Introduction

Atmospheric conditions near the surface, e.g. air temperature, specific humidity, surface temperature and wind speed, vary on small spatial and temporal scales. The reasons for these variabilities are surface heterogeneities. The classical approach to measuring meteorological conditions close to surface is to use a single, highly accurate station. However, using only one station provides limited area-representativeness. A large number of stations need to be deployed over a small area to monitor the variability caused by different land surfaces and environments.

In earlier and more recent studies, networks of meteorological stations have been used for monitoring the meteorological conditions. For most of these studies, the inter-station distance typically exceeds 1 km. For example, Zemel and Lomas (1976) deployed a network of 70 stations in the Huleh Valley in Israel. They examined the patterns of air temperature. Kawashima and Ishida (1992) studied the air temperature close to the surface in a 250 km x 300 km wide area in Japan with a network consisting of 130 stations. They investigated the maximum distance over which the hourly air temperatures of two stations remain correlated: The distance is up to 93 km in summer and up to 98 km in winter; under wet conditions it is larger than under dry conditions. In another study, Hubbard (1994) used a network with inter-station distances of several 10 km to determine the spatial variability of daily measurements in the High Plains (U.S.A.). The maximum distance between stations needed to explain 90% of the variance is investigated: For the maximum air temperature, the stations should not be more than 60 km apart. The maximum distance between sites for minimum air temperature, humidity and solar radiation is 30 km, for wind speed 10 km and for precipitation only 5 km.

To monitor the variability and to investigate local features of meteorological conditions at the microscale up to 2 km (Orlanski, 1975), smaller networks are necessary. With such networks air temperature fluctuations and, especially nocturnal cooling, are frequently explored. Bodine et al. (2009) studied this cooling in the Lake Thunderbird Micronet (Shapiro et al., 2009) area with 26 stations in a 120 m x 320 m wide domain. They found a difference of 3 K to 4 K within a distance of 70 m to 100 m and an elevation range of 6 m in temperature observed at a height of 2 m. Hunt et al. (2007) detected a rapid nocturnal cooling and strong inversions at the El Reno Oklahoma Mesonet site using a

transect of four portable automated micrometeorological stations. Under calm wind conditions, low humidity and clear sky, particularly one site was up to 5 K colder than other stations, which were located a few 100 m away. Nocturnal cooling along a 22 km long transect in complex terrain including forest was investigated by Gustavson et al. (1998). Mahrt (2006) studied the spatial variability of surface air temperature in complex terrain within a horizontal range of 200 m up to 1.4 km using six networks consisting of four to 14 temperature sensors. A quite dense network of measurements of air-, surface- and soil temperature and soil moisture was used by Xu et al. (2002). They constructed a 10 km transect with stations every 10 m to examine the spatial variability of the meteorological conditions along this transect and found that 22% to 52% of the variations could be explained by the topography.

To be in charge of such a network is a challenging and time consuming task because the data has to be collected at each individual station. In recent years, wireless sensor networks have become a common method to investigate small scale variabilities. The advantage of these wireless networks is their autonomy, which permits their use in almost any environment. For example, sensors were placed on zebras for studying wildlife-tracking systems in Kenya (Zhang et al., 2005). A network of 16 sensors was set up on an active volcano in Ecuador to collect seismic and acoustic data (Werner-Allen et al., 2006a,b). Polastre et al. (2004) placed 43 stations in bird's nests on Great Duck Island (U.S.A.) for habitat monitoring.

In this study, the influence of the environment on meteorological conditions near the surface, e.g. air temperature, specific humidity, surface temperature and wind speed, will be examined. Two different kinds of environments are chosen. The first site is a heterogeneous terrain located in western Germany covering different kinds of land use, hedge banks, settlement and a small river. The second campaign takes place at Hamburg Airport. This area is homogeneous with grasslands only. The site's extension is less than 5.5 km. The following questions will be addressed:

- How large are the variabilities in air temperature, specific humidity, surface temperature and wind speed for both campaigns?
- What causes these variabilities? Are there different explaining factors for the two campaigns?
- Which sites are effectively redundant and need to be observed only once?
- How large is the resulting error compared to the full network?
- The German Meteorological Service operates a climate reference station at the centre of the Hamburg Airport. Are the observations at the centre of the airport representative for the whole area?

Another field of application for such a sensor network is the validation of model simulations and parametrisation approaches. Two examples will be discussed in this thesis: The observations of the sensor network will be compared to the output of the atmospheric

model *FOOT3DK* (Flow Over Orographically Structured Terrain - 3 Dimensional Kölner Version (Brücher et al., 2003)). This mesoscale model offers a spatial resolution of 100 m. The following questions arise:

- How accurate are the model simulations?
- Is it possible to simulate small scale variabilities with the *FOOT3DK* model?

For simulating heat fluxes, recent studies use the mosaic or tile approach (e.g. Seth et al., 1994) by considering surface heterogeneities, but neglecting atmospheric variability. The soil-vegetation-atmosphere transfer scheme *TERRA* operated by the German Meteorological Service (DWD) is used to simulate sensible and latent heat flux driven by the network measurements of the *FLUXPAT2009* campaign. A model study is carried out that deals with the question:

- How important are heterogeneous forcing and vegetation for the simulation of sensible and latent heat fluxes?

In Chapter 2, the concept of the sensor network is described. The sensors used in this study are characterised in Chapter 3 and a description of the calibration of these sensors is given. Chapter 4 deals with the campaign over the heterogeneous terrain (see also Lengfeld and Ament (2011)). The site and the set up are described. The variability of air temperature, specific humidity and surface temperature and their explaining factors are analysed. Based on these analyses, redundant stations are determined using cluster analysis and the error made by using less dense networks is estimated. The observations of the sensor network are also compared to simulations of the *FOOT3DK* model and the *TERRA* model. The *TERRA* model is used to examine the influence of heterogeneous forcing and vegetation on heat flux simulations. Chapter 5 provides the same analyses of the variabilities in atmospheric conditions, their explaining factors and redundant stations at the second site at the Hamburg Airport. A comparison between both sites and concluding remarks are made in Chapter 6. In Chapter 7, an outlook for further research is given.

Chapter 2

Sensor Network

The sensor network used in this study is called *SensorScope*. It has been developed at the *École Polytechnique Fédérale de Lausanne* (EPFL) in Switzerland. It is an autonomous network with low-cost stations. Fig. 2.1a shows a picture of such a station. Each station is composed of an aluminium pole with a base plate and guy-ropes, a solar panel for power supply, external sensors and a hermetic box, in which the electronics are enclosed. A Shockfish TinyNode (tinynode.com, 2011) is used consisting of a Texas Instrument MSP430 16-bit microcontroller (ti.com, 2011) and a Semtech XE1205 radio transceiver (semtech.com, 2011). The latter operates at 868 MHz. A *SensorScope* network consists of two different kinds of stations: Slave stations and master stations. Slave stations send their data via radio communication to a master station. Master stations are equipped with a SIM-card to transmit the collected data via GSM/GPRS to a central server. The stations build up the communication network automatically. A scheme of the communication within the network is displayed in Fig. 2.1b. The *SensorScope* stations can be deployed in almost every environment, but the distance between two stations should not exceed the maximum communication range of 1200 m for clear line of sight (Ingelrest et al., 2010). The communication became a problem during one of the measurement campaigns, despite an inter-station distance which was distinctly smaller than the maximum communication range. Another problem was the power supply. All stations were equipped with a solar panel and rechargeable batteries. For longer periods of overcast sky, the batteries could not be recharged and data was lost. For detailed information about the network, see Barrenetxea et al. (2008) and Ingelrest et al. (2010).

The communication stack within a *SensorScope* datalogger consists of four layers (Fig. 2.1c):

1. The application layer collects data, which has to be sent to the master station, e.g. measurements from attached sensors or battery levels.
2. The transport layer creates and receives packets. Data packets have to be sent to the master stations. Control packets, e.g. synchronisation packets, have to be sent to neighbouring nodes. If it is necessary, the transport layer queues the packets whereby priority is given to control packets.

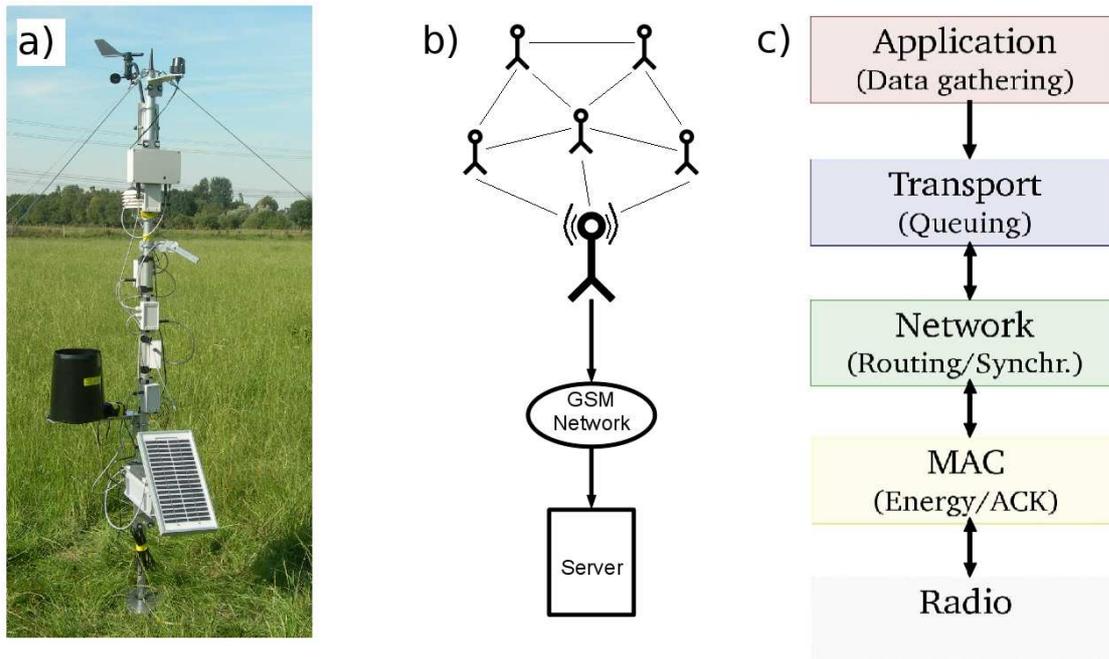


Figure 2.1: a) *SensorScope* station, b) communication inside the network, c) communication stack

3. At the network layer, all routing decisions are made. After choosing a next hop for slave stations, this layer passes the data packets to the Media Access Control (MAC) layer. At the master stations, the data packets are forwarded to the serial port. Control packets are always passed to the MAC layer.
4. The MAC layer turns the radio on and off. It sends and receives messages. Due to power consumption, the radio is switched off for as long as possible. Therefore, packets, which must be sent while the radio is off, are stored and sent later. In case of data packets, an acknowledgement is sent back to the previous hop.

Every station stores the stations whose signals it can receive in a *neighbourhood* table. Neighbours, that can only be heard, are detected with the help of acknowledgement messages. If a data packet is successfully sent to a neighbour, but no acknowledgement is sent back, the neighbour is blacklisted and will not be considered as a next hop again.

A sequence number is assigned to every data packet. In case of unsuccessful delivery, the packet is resent. By counting missing sequence numbers, the quality of links is estimated. This quality of links indicates whether a station is a good neighbour and whether a station forwards packets to the master station effectively. Slave stations choose their next hop randomly with two restrictions: The next hop must be closer to the master station and it should be a high quality neighbour in terms of the quality of links. In case there is no high quality neighbour, the slave stations choose a low quality neighbour. Therefore, no routing backbone to the master station exists that might be overstrained and the protocol

is consequently more robust.

The stations are supplied with energy by solar panels. Hence, the available power is limited. The biggest energy consumer is the radio. The energy consumption is eight times higher when the radio is turned on (Barrenetxea et al., 2008). Therefore, the radio should be turned off for as long as possible. All stations switch their radios on synchronously to send and receive packets. To make sure all stations are awake at the same time, their clocks need to be synchronised. The time at the master station serves as reference time. A slave station sends a request to a neighbour closer to the master station when it wants to update its clock. This neighbour passes the time to all its neighbours further from the master station, which regularly send messages with its time to the central server. The server calculates the offset between network and actual time and corrects the timestamp. Stations can be equipped with meteorological sensors according to the researchers requirements. For this study six different kinds of sensors are used:

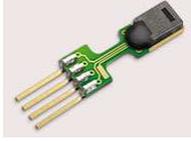
- **SHT75 Humidity and Temperature Sensor** made by *Sensirion*: Relative humidity is measured by a capacity sensor element, a band-gap sensor is used for measuring air temperature. For more information see sensirion.com (2011).
- **TN901 Infrared Thermometer Module** made by *ZyTemp*: The sensor measures infrared radiation I of a target by an infrared mirror through an IR filter. The infrared temperature of the surface T_{surf} is calculated according to the *Stefan-Boltzmann-Law*:

$$I = \varepsilon \cdot \sigma_B \cdot T_{surf}^4, \quad (2.1)$$

where $\sigma_B = 5.67 \cdot 10^{-8} \text{ W m}^{-2}\text{K}^{-4}$ is the *Stefan-Boltzmann constant*. The infrared emissivity ε is set to 0.95. The sensor has a field of view of 53.2° . For more details see zytemp.com (2011).

- **Anemometer 6410** made by *Davis*: Wind speed is measured by wind cups and a magnetic switch, wind direction is measured by a wind vane and a potentiometer (davisnet.com, 2011a).
- **Solar Radiation 6450** made by *Davis*: The sensor is equipped with a silicon photodiode with spectral response in the range of 400 to 1100 nm (davisnet.com, 2011c). It measures global radiation, the sum of direct and diffuse solar radiation, at a certain point.
- **Rain Collector II 7852** made by *Davis*: The sensor is a tipping bucket (214 cm² collection area and 16.5 cm diameter) with magnetic switch for measuring precipitation. A detailed description can be found on davisnet.com (2011b).
- **EC-TM Probe** made by *Decagon*: The sensor measures temperature of soil and dielectric permittivity to determine the water content and, therefore, the soil moisture (decagon.com, 2011).

Table 2.1: Sensor specifications

Sensor Name	Photo	Measuring Range	Accuracy
SHT75 Humidity and Temperature Sensor		RH: 0 - 100%, T: -40 - 125°C	RH: $\pm 1.8\%$, T: $\pm 0.3^\circ\text{C}$
TN901 Infrared Thermometer		-10 - 50°C	$\pm 0.6^\circ\text{C}$ between 15°C and 35°C, $\pm 2^\circ\text{C}$ or $\pm 2\%$ at full range
Anemometer 6410		Speed: 1.5-79 m s ⁻¹ , Dir.: 0 - 360°	Speed: $\pm 0.1 \text{ m s}^{-1}$ or $\pm 5\%$, Dir.: $\pm 4^\circ$
Solar Radiation 6450		0 - 1800 W m ⁻² (operating temperature: -40 - 65°C)	$\pm 5\%$ of full scale
Rain Collector II 7852		0 - 999.8 mm per day 0 - 9999 mm total	$\pm 4\%$ (resolution of 0.252 mm)
EC-TM Probe		Moist.: 0 - 100 %, T: -40 - 50°C	Moist.: $\pm 3\%$ VWC, T: $\pm 1^\circ\text{C}$

A list of photos of the sensors, their measuring range and their accuracy is given in Table 2.1.

The *SensorScope* network was first tested within the *Lausanne Urban Canopy Experiment*

(LUCÉ) from July 2006 to May 2007 (Nadeau et al., 2009). 100 stations were deployed at the campus of EPFL to study the interactions between an urban environment and the lower atmosphere. From August to September 2007 a network of 16 stations was set up at a rock glacier located on Le Généri in Switzerland, which is a source of dangerous mud streams during intense rain. With the network observations, a correlation of precipitation with wind and temperature based on the shape of the landscape is studied. More information about these campaigns can be found in Ingelrest et al. (2010).

Chapter 3

Characterisation of the Sensor Network

The *SensorScope* stations are equipped with low-cost sensors. These sensors might not be as accurate as more expensive ones. Therefore, a comparison against a more accurate reference is necessary to estimate the quality of the network measurements. For each type of sensor, a calibration has been done. A description of the set-ups for the different sensor types is given in Chapters 3.1 to 3.6. The accuracy compared to the reference and the consistency of the network sensors has been investigated. Based on these comparisons calibration coefficients are calculated for each sensor.

A second experiment is described in Chapter 3.7. Air temperature and humidity sensors and wind sensors are mounted at three different heights to carry out profile measurements. This set-up is used to examine the viability of resolving vertical profiles of meteorological quantities.

3.1 Humidity and Temperature Sensors

The SHT75 Humidity and Temperature Sensors were calibrated on the rooftop of the Meteorological Institute of the University of Hamburg. They were all attached to a pole with distances of around 20 cm in between (see Fig. 3.1). The network sensors were compared to a HMP45 Humidity and Temperature Sensor made by Campbell (campbellsci.com, 2011c), which was mounted close to the network sensors (front sensor in Fig. 3.1). This sensor has a certified accuracy of 1.5% for relative humidity and 0.13 K for air temperature. The HMP45 measurements served as references for the calibration. The calibration period took place between the FLUXAPT (see Chapter 4) and the airport campaign (see Chapter 5). It started on September 8th and ended on October 8th, 2009.

The network sensors are surrounded by a small radiation shield and passively ventilated. However, solar radiation influences the temperature measurements. As an example, the comparison between sensor number 11 and the reference sensor is shown in Fig. 3.2a as a scatter plot. The comparison between the reference and the other network sensors gives similar results. At night, shown as blue dots, the measurements of sensor 11 and the



Figure 3.1: SHT75 Humidity and Temperature Sensors on the rooftop of the Meteorological Institute during the calibration period. The reference sensor is attached at front of the pole.

reference sensor match well. But during the day, represented by red dots, sensor 11 overestimates especially the high temperatures. The distribution of the differences between sensor 11 and the reference is presented in Fig. 3.2c. To determine the performance of the network sensor, the mean difference m between the network and the reference sensor, the *Root Mean Square Error* (RMSE) and 10% and 90% quantiles, q_{10} and q_{90} , are calculated. The RMSE between two time series of length N from stations X_a and X_b is defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (X_{a,i} - X_{b,i})^2}{N}}. \quad (3.1)$$

The mean bias is 0.16 K and the RMSE 0.34 K. The 10% and 90% quantiles are -0.09 K and 0.6 K leading to a quantile range Δq_{90-10} of 0.69 K. This overestimation is caused by solar radiation that heats the network sensor. Therefore, a correction formula for air temperature dependent on solar radiation is applied:

$$T_{corrected} = T_{measured} + a \cdot F_{SW} + b, \quad (3.2)$$

where $T_{corrected}$ is the corrected air temperature, $T_{measured}$ is the air temperature measured by the network sensor, F_{SW} is solar radiation and a and b are regression coefficients. A list of the coefficients a and b can be found in Table A.1 in Appendix A. The mean bias is eliminated by choosing the coefficients accordingly. The comparison between the corrected temperature measurements of sensor 11 and the reference is shown in Fig. 3.2b and the distribution of their differences in Fig. 3.2d. The influence of solar radiation is strongly reduced and the differences are arranged symmetrically around their mean of 0 K. The RMSE is reduced by half to 0.19 K. The 10% and 90% quantiles are also almost symmetrical to 0 K: q_{10} amounts to -0.18 K, q_{90} amounts to 0.17 K.

To correct the measurements of relative humidity, first the difference between the network sensors and the reference sensor is calculated. Averages are taken for every one-

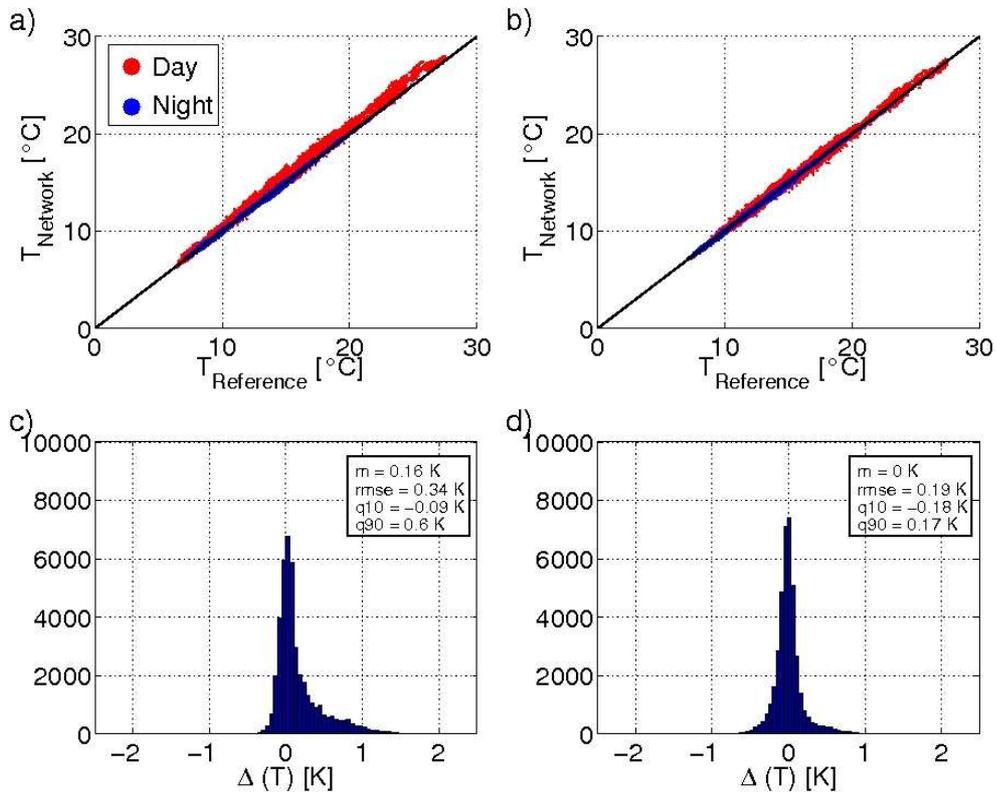


Figure 3.2: Comparison of air temperature measurements from network sensor 11 and the reference sensor for measured data (a) and corrected data (b). Daytime temperatures are represented by red dots, nighttime temperatures by blue dots. Distribution of the difference between sensor 11 and the reference sensor for measured data (c) and corrected data (d).

percent interval between 20% and 100% relative humidity. Smaller values did not occur during the calibration period. These mean values are given in Fig. 3.3a as black dots for sensor 11 as an example. All other sensors show similar results. Coloured graphs represent a linear (blue), a quadratic (red) and a cubic fit (green). The linear fit underestimates relative humidity less than 45% and between 70% and 90% and overestimates relative humidity between 45% and 70% and more than 90%. The same applies to the quadratic fit, although it is more suitable for relative humidity between 50% and 70% than the linear fit. The cubic fit matches the differences between sensor 11 and the reference sensor best, especially for very high and very low values of relative humidity. Therefore, a correction formula of the third order, ΔRH_{fit} , is applied to the differences between humidity measurements of the network sensors and the reference:

$$\Delta RH_{fit} = a \cdot (\Delta RH)^3 + b \cdot (\Delta RH)^2 + c \cdot \Delta RH + d, \quad (3.3)$$

where ΔRH is the difference between the measured relative humidity of the network

sensors and the reference sensor and a , b , c and d are regression coefficients. A list of the coefficients a , b , c and d can be found in Table A.2 in the Appendix A. Based on ΔRH_{fit} , a bias δRH is calculated for every sensor that has to be added to the measurements:

$$RH_{corrected} = RH_{measured} + \delta RH, \quad (3.4)$$

with $RH_{corrected}$ being the corrected relative humidity and $RH_{measured}$ being the measured relative humidity by the network sensor.

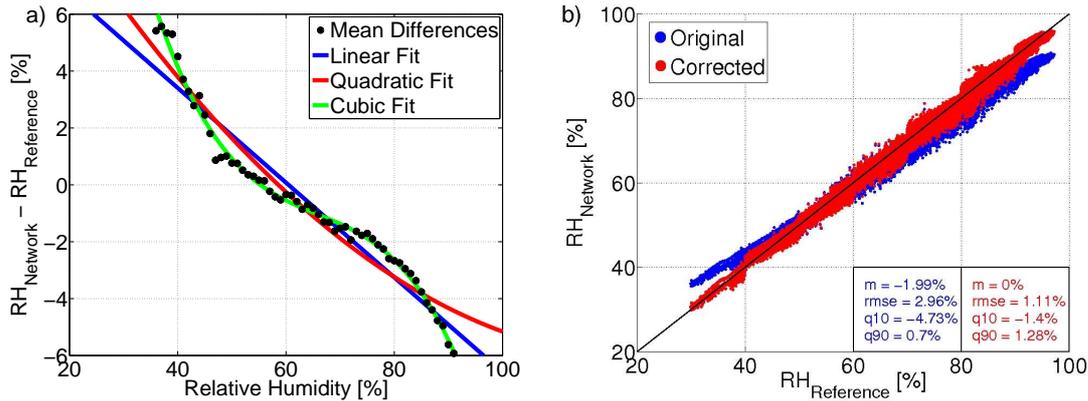


Figure 3.3: Calibration of relative humidity: a) Mean differences for 1% intervals between network sensor 11 and the reference sensor are shown as black dots, lines represent a linear fit (blue), a quadratic fit (red) and a cubic fit (green). b) Comparison of relative humidity measurements of network sensor 11 and the reference. Original data is shown in blue, corrected data in red.

A comparison of the relative humidity observed by sensor 11 and the reference sensor is shown in Fig. 3.3b. Original values measured by sensor 11 are represented by blue dots, corrected values are represented by red dots. Before the correction, sensor 11 underestimated high relative humidities and overestimated low relative humidities. The mean difference between sensor 11 and the reference is almost -2% and the RMSE is 2.96%. The corrected values of relative humidity fit better, especially for low and high relative humidity. The mean difference is eliminated and the RMSE is reduced by almost a third to 1.11%.

In general, relative humidity is highly influenced by air temperature. For the analysis of humidity data, it is reasonable to use the specific humidity because then the influence of air temperature is eliminated. Therefore, also the uncertainties in specific humidity are quantified. First, saturated vapour pressure E is computed using the air temperature T_{Air} [$^{\circ}C$] mentioned above:

$$E = 6.122 \cdot \exp\left(\frac{17.62 \cdot T_{Air}}{243.12 + T_{Air}}\right). \quad (3.5)$$

Based on air pressure measurements p observed at the *Hamburg Weather Mast* 10 km south-east of the Meteorological Institute (<http://wettermast-hamburg.zmaw.de>), the saturated specific humidity Q is calculated:

$$Q = \frac{0.622 \cdot E}{p - 0.378 \cdot E}. \quad (3.6)$$

Then, the specific humidity q [kg kg^{-1}] is generated from the relative humidity RH [%]:

$$q = \frac{RH \cdot Q}{100}. \quad (3.7)$$

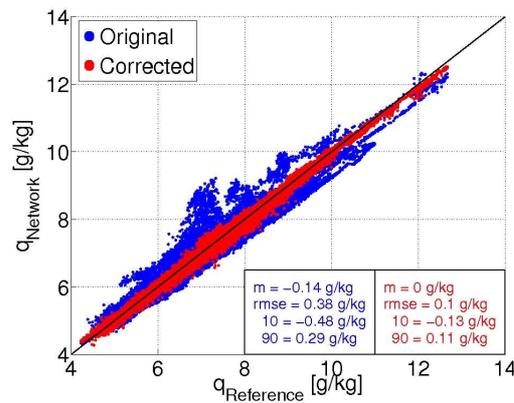


Figure 3.4: Comparison of specific humidity calculated from air temperature and relative humidity measurements of network sensor 11 and of the reference sensor. Original data is given by blue dots, corrected data by red dots.

In Fig. 3.4, the comparison of specific humidity between the reference and sensor 11 is shown. Uncorrected temperature and relative humidity measurements are used for original specific humidity (blue dots), unbiased measurements are used for corrected specific humidity (red dots). The corrected values of specific humidity clearly match better with the reference than the original ones. The mean difference between original and reference specific humidity is -0.14 g kg^{-1} . For corrected values, the bias is eliminated. The RMSE is almost four times higher (0.38 g kg^{-1}) for original specific humidity than for corrected specific humidity (0.1 g kg^{-1}). Also, the quantile range Δq_{90-10} is reduced to a third.

Although the measurements of the network sensors are not highly accurate compared to the reference sensors, their measurements are consistent and, therefore, comparable. In Table 3.1, the mean difference, the RMSEs and the quantile range between all network sensors and the reference sensor and the mean difference between the network sensors is listed for air temperature and specific humidity. For both, the mean difference between the network sensors is one order of magnitude smaller than reference-network difference. The RMSE between the network sensors is two thirds of the RMSE between network and reference and Δq_{90-10} is also smaller. A second dataset collected during one week in October 2009 confirms these improvements.

Table 3.1: Mean difference, RMSE and Δq_{90-10} for air temperature and specific humidity

		Mean Difference	RMSE	Δq_{90-10}
T_{air}	Network/Reference	0.17 K	0.4 K	0.87 K
	Network/Network	0.04 K	0.25 K	0.59 K
q	Network/Reference	0.05 g kg ⁻¹	0.26 g kg ⁻¹	0.52 g kg ⁻¹
	Network/Network	0.004 g kg ⁻¹	0.18 g kg ⁻¹	0.41 g kg ⁻¹

3.2 Infrared Thermometers

The quality of surface temperature measurements taken by TNX901 Infrared Thermometers was tested by means of a water bath. Packets of four and six sensors were put together and faced the water surface. Measurements of a *KT19* with an accuracy of 0.05 K were used (wintron.com, 2011) as references. The set-up is presented in Fig. 3.5. In the first part of the quality test, the water was frozen. The water surface temperature was measured while the ice was melting at an ambient temperature of 8°C. In the second part the water was heated up to 45°C and the surface temperature was measured during cooling down at an ambient temperature of 13°C. The sensors were aerated by a ventilator to prevent them from misting.



Figure 3.5: Calibration set-up for TNX901 Sensors. The network sensors are facing a water bath. The *KT19* used as a reference stands on the right side.

The measurements are averaged over one degree steps. Testing reveals that sensors of one packet show similar results, but different packets show different behaviour. Two sensors of the first packet and the second packet are illustrated as examples in Fig. 3.6. The mean differences in 1° steps between reference and network sensors follow a quadratic function for the two sensors of packet I, but are almost linear for the sensors of packet II. One possible reason is the different field of view of the network sensors and the *KT19*. The *KT19* has an aperture angle of 1.145°, the network sensors have an aperture angle of more

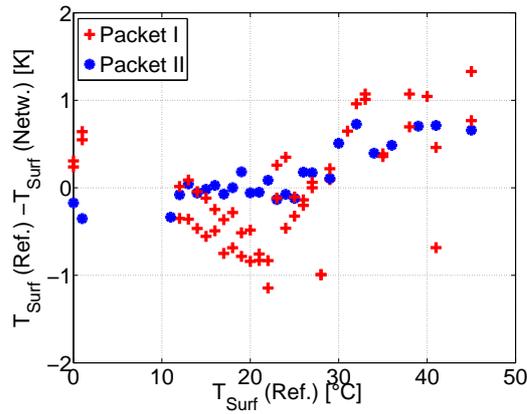


Figure 3.6: Surface temperature differences between the reference and the network sensors in 1°C steps for two sensors of packet I (red crosses) and two sensors of packet II (blue stars).

than 26° . Therefore, the KT19 measures a smaller area of the water surface. A calibration of the Infrared Thermometer is not feasible because no compensating curve can be found that fits for all sensors.

In Table 3.2, the mean difference, the RMSE and Δq_{90-10} between the network sensors and the reference is listed in the upper row. The mean difference and the RMSE, 0.11 K and 0.48 K, respectively, are both smaller than the given uncertainty of ± 0.6 K of the network sensors. The quantile range is 1.14 K. Although the mean difference between the network sensors is slightly higher than between the network sensors and the reference, their RMSE and their Δq_{90-10} is distinctly smaller. Therefore, measurements of the network sensors are comparable, although the value of surface temperature might not be highly accurate.

Table 3.2: Mean difference, RMSE and Δq_{90-10} for surface temperature

	Mean Difference	RMSE	Δq_{90-10}
Network/Reference	0.11 K	0.48 K	1.14 K
Network/Network	0.13 K	0.38 K	0.83 K

3.3 Wind Speed and Direction Sensors

The Davis Wind Sensors were calibrated in the wind tunnel *Göttinger Type* of the University of Hamburg. The tunnel generates a homogeneous, mostly turbulence-free wind field. Two sensors were placed simultaneously in the wind tunnel symmetrical to the middle of the channel, so that both were exposed to the same wind speeds. The speed of the wind tunnel was determined with a Prandl tube. The wind sensors were exposed to

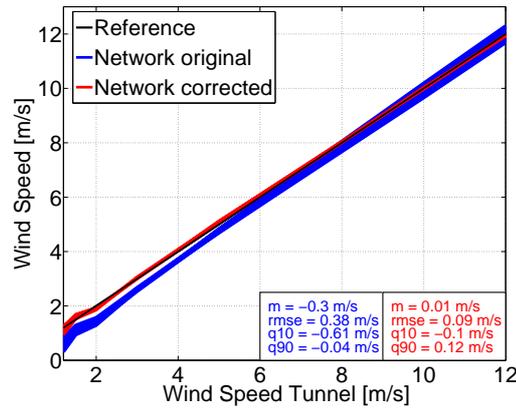


Figure 3.7: Wind speed calibration: The black line represents the reference measurements, blue lines represent wind speeds measured by the network sensors and red lines represent corrected wind speeds of the network sensors.

wind speeds of 1.2 m s^{-1} , 1.5 m s^{-1} , 2 m s^{-1} , 3 m s^{-1} , 5 m s^{-1} , 7.5 m s^{-1} , 10 m s^{-1} and 12 m s^{-1} for five minutes each wind speed.

In Fig. 3.7, the comparison of wind speeds of the wind tunnel and measured by the network sensors is shown. The uncorrected data, represented in blue, shows a clear underestimation of low wind speeds ($< 8 \text{ m s}^{-1}$). The mean difference between network sensors and reference is -0.3 m s^{-1} . Their RMSE is 0.38 m s^{-1} . The underestimation is also obvious, when taking a look at the 10% and 90% quantile: Both are negative ($q_{10} = -0.61 \text{ m s}^{-1}$ and $q_{90} = -0.04 \text{ m s}^{-1}$). To correct that bias, a linear regression of the form

$$v_{corrected} = a \cdot v_{measured} + b \quad (3.8)$$

is applied, where $v_{corrected}$ is the corrected and $v_{measured}$ the measured wind speed of the network sensors and a and b are regression coefficients. A list of a and b for each sensor is given in Table A.3 in Appendix A.

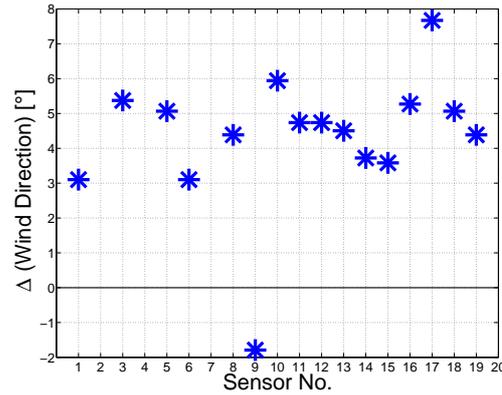
The corrected wind speeds are shown in red in Fig. 3.7. The mean bias is corrected to almost 0 m s^{-1} and the 10% and 90% quantiles are symmetrical around zero with -0.1 m s^{-1} and 0.12 m s^{-1} . The RMSE is reduced to about a quarter.

The statistics of the network sensors in Table 3.3 emphasise that all wind sensors give similar results. The mean difference between the network sensors is only 0.002 m s^{-1} , their RMSE is half the RMSE between the network sensors and the reference. Δq_{90-10} between the network sensors is also 0.13 m s^{-1} smaller than between the network and the reference.

To correct the wind direction measurements the sensors are oriented in a way that they face a northerly wind at constant wind speed for three minutes. The bias of the wind direction of each sensor is shown in Fig. 3.8. This bias is added to the wind direction measurements.

Table 3.3: Mean difference, RMSE and Δq_{90-10} for wind speed

	Mean Difference	RMSE	Δq_{90-10}
Network/Reference	-0.3 m s ⁻¹	0.38 m s ⁻¹	0.57 m s ⁻¹
Network/Network	0.002 m s ⁻¹	0.17 m s ⁻¹	0.42 m s ⁻¹

**Figure 3.8:** Bias in wind direction for each wind sensor.

3.4 Solar Radiation Sensors

The Davis Solar Radiation Sensors were attached to the same pole as the Humidity and Temperature Sensors on the rooftop of the Meteorological Institute. The sensors were 20 cm apart and measured the solar radiation from September 8th to October 8th, 2009. The measurements of the network sensors are compared to a CM11 pyranometer by *Kipp&Zonen* (kippzonen.com, 2011). The CM11 was placed approximately 1 m away at the same height.

In Fig. 3.9, the comparison between 10 minute means of solar radiation measured by network sensor 04 and by the reference sensor is shown as an example. Blue dots represent the original measurements of sensor 04. There is a distinct overestimation, especially for solar radiation greater than 400 W m⁻². The mean difference between the network sensor and the reference is 12.22 W m⁻². The RMSE is 17.81 W m⁻². The overestimation is also evident in the 10% and 90% quantiles: Both are positive with values of 6.36 W m⁻² and 28.11 W m⁻², respectively. To correct this overestimation a linear regression of the form

$$F_{SW_{corrected}} = a \cdot F_{SW_{measured}} + b \quad (3.9)$$

is applied, where $F_{SW_{corrected}}$ and $F_{SW_{measured}}$ are corrected and measured solar radiation of the network sensors, respectively, and a and b are regression coefficients. A list of these coefficients for each sensor can be found in Table A.4 in Appendix A. The corrected measurements of sensor 04 are represented as red dots in Fig. 3.9. The bias is almost re-

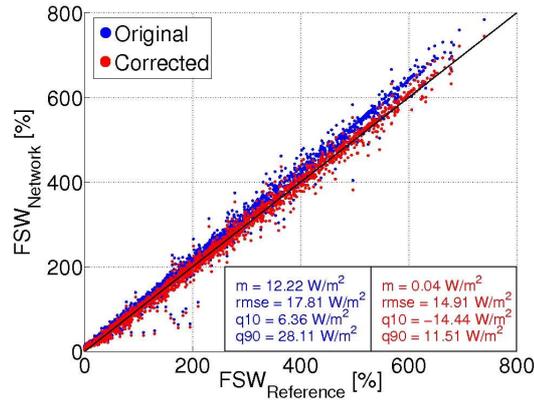


Figure 3.9: Comparison of solar radiation measurements of network sensor 04 and the reference sensor. Original data is shown in blue, corrected data in red.

moved; it is 0.04 W m^{-2} . The RMSE is slightly reduced to 14.91 W m^{-2} . The 10% and 90% quantiles are almost symmetrical to 0 with $q10 = -14.44 \text{ W m}^{-2}$ and $q90 = 11.51 \text{ W m}^{-2}$. Table 3.4 contains mean differences, RMSEs and quantile ranges between the network and the reference and between all network sensors. The mean difference between the network and the reference is more than 8 W m^{-2} , however, between all network sensors it is only 1 W m^{-2} . The RMSE of all network sensors is almost one-fourth smaller than between the network and the reference. Δq_{90-10} between the network and the reference is more than twice the quantile range between all network sensors. The RMSE is higher than Δq_{90-10} , especially for the network-reference-comparison. That is caused by a few very large outliers. Table 3.4 indicates that the network sensors are not highly accurate but comparable.

Table 3.4: Mean difference, RMSE and Δq_{90-10} for solar radiation

	Mean Difference	RMSE	Δq_{90-10}
Network/Reference	8.24 W m^{-2}	52.34 W m^{-2}	26.35 W m^{-2}
Network/Network	0.96 W m^{-2}	14.25 W m^{-2}	11.84 W m^{-2}

3.5 Precipitation Sensors

The calibration of the Davis Rain Collectors took place in the laboratory. The tipping buckets were placed underneath a bottle with a tube attached to its thin end. The droplet velocity could be regulated. A picture of the instalment is shown in Fig. 3.10. Ten millimetre water were poured into the bottle and dropped down into the rain collector. After every attempt, the adjusting screws were turned until the results were satisfactory.



Figure 3.10: *Rain Collector instalment: The rain collector is placed underneath a bottle filled with 10 mm water that functions as a drip.*

By turning the adjusting screws, the tipping point of the rocker inside the bucket can be changed. Because the resolution of the tipping bucket is 0.254 mm, it cannot measure exactly 10 mm. So every tipping bucket should measure 9.906 mm.

As an example, the first and the final result for sensor 08 is shown in Fig. 3.11. At the first attempt, sensor 08 underestimates the 10 mm of fluid by more than 1 mm. Precipitation measurement corrections were required prior to calibration during the first campaign. Therefore, a calibration factor c is calculated as follows:

$$c = \frac{10\text{mm}}{RR_1}, \quad (3.10)$$

where RR_1 is the result of the first calibration attempt. The correction for the measurements of the first campaign is of the form:

$$RR_{corrected} = c \cdot RR_{measured}, \quad (3.11)$$

where $RR_{corrected}$ and $RR_{measured}$ are corrected and measured rain rate, respectively. A list of the coefficients c can be found in Table A.5 in Appendix A.

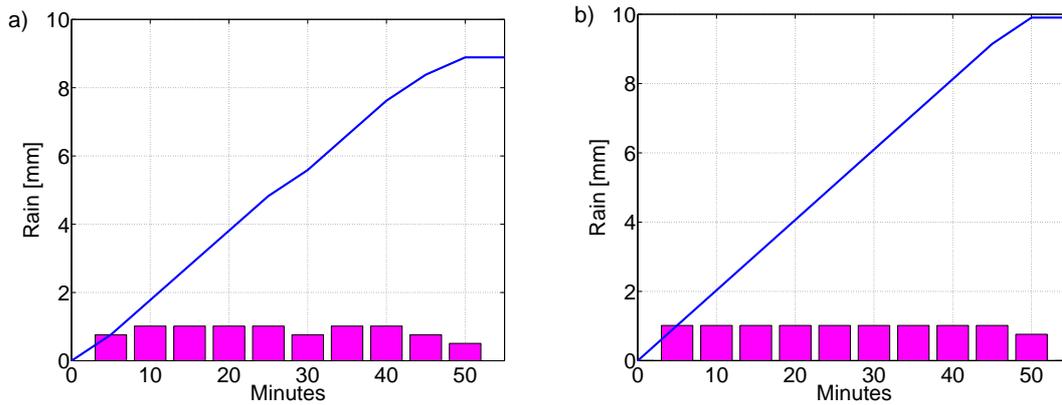


Figure 3.11: First (a) and final (b) attempt of the calibration for sensor 08. Pink bars represent the amount of rain that has reached the tipping bucket in the current five minute intervals. The blue line stands for the rain sum over the whole time period.

3.6 Soil Moisture and Temperature Sensors

To calibrate the Soil Moisture and Temperature Sensors, they were put into a box filled with sand on the rooftop of the Meteorological Institute. The sensors were placed in a holder so that they were all in the same depth and equidistant. Sand is a homogeneous medium and every sensor is expected to measure the same moisture and temperature. The calibration period began on September 23rd and ended on October 30th, 2009. Since there was no reference instrument, an absolute calibration is not possible. Only the bias between the sensors can be corrected.

In Fig. 3.12a, soil temperature measurements for the calibration period are shown for each sensor and in Fig. 3.12b, their difference to the mean is illustrated. The same for soil moisture is given in Fig. 3.12c and d. All sensors give very similar temperatures for the calibration period. The differences are mostly less than ± 0.5 K for all sensors. For soil moisture, the differences are in a $\pm 5\%$ interval around the mean of all sensors. All sensors show nearly constant biases for soil moisture and temperature. To correct these biases, the mean of all sensors for soil moisture and temperature $\bar{X}_{all\ sensors}$ is calculated. Then, the mean difference $\Delta(\bar{X})$ between $\bar{X}_{all\ sensors}$ and the mean of each sensor \bar{X}_{sensor} is computed as follows:

$$\Delta(\bar{X}) = \bar{X}_{all\ sensors} - \bar{X}_{sensor}. \quad (3.12)$$

$\Delta(\bar{X})$ has to be added to the measurements of soil moisture and temperature:

$$X_{corrected} = X_{measured} + \Delta(\bar{X}), \quad (3.13)$$

where $X_{corrected}$ and $X_{measured}$ are corrected and measured quantities. A list of $\Delta(\bar{X})$ can be found in Table A.6 in Appendix A. Because only the bias of soil moisture and soil

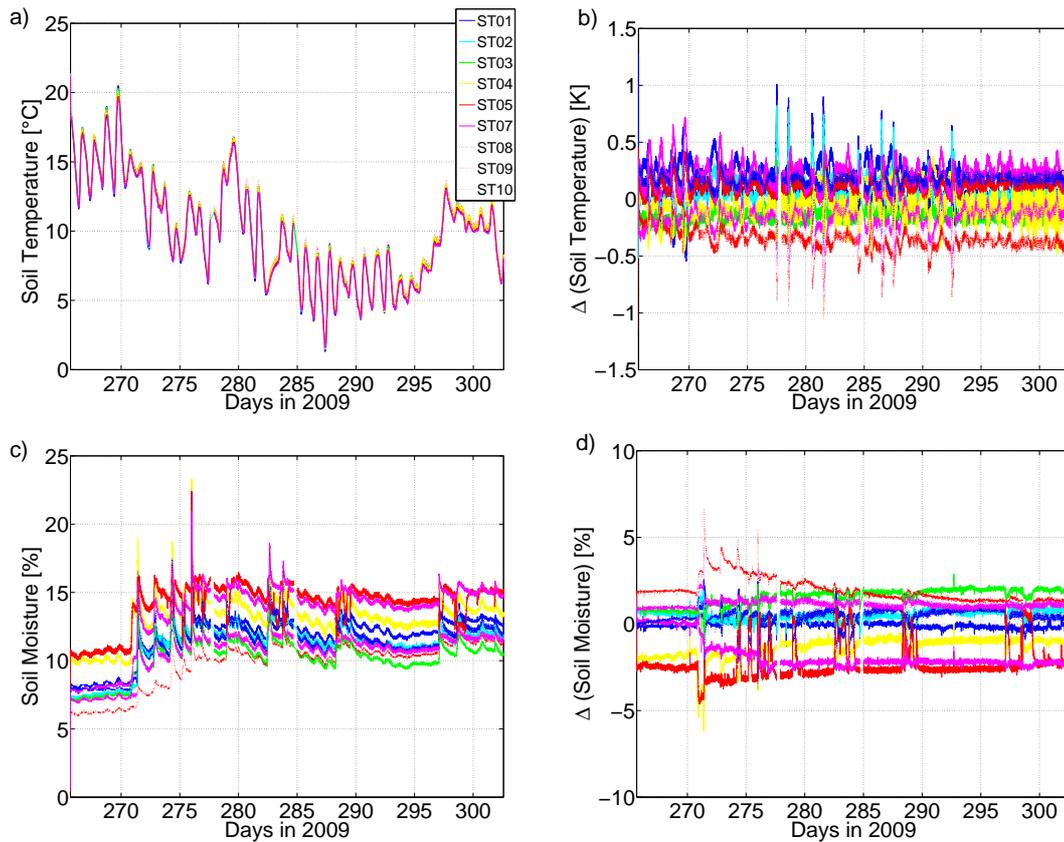


Figure 3.12: Comparison of soil moisture and temperature: a) Soil temperature measurements of all network sensors during the calibrations period. Different sensors are shown in different colours. b) Differences in soil temperature between each network sensor and the mean of all network sensors. c) Soil moisture measurements of all network sensors during the calibration period. d) Differences in soil moisture between each network sensor and the mean of all network sensors.

temperature is corrected, the absolute values are not necessarily correct, but the difference between different sites can be examined.

3.7 Profile Measurements

Another way to test the quality of the network sensors is to perform profile measurements. Three temperature and humidity sensors and three wind sensors were mounted at a mast at heights of 2 m, 2.6 m and 3.7 m. The mast was located near Merken in Western Germany. The observation period started on August 9th and ended on August 27th, 2009. Measurements were taken every minute.

In Fig. 3.13, the mean diurnal cycle for the whole period is shown for air temperature (a), specific humidity (b) and wind speed (c). Measurements of the lower level at a height of

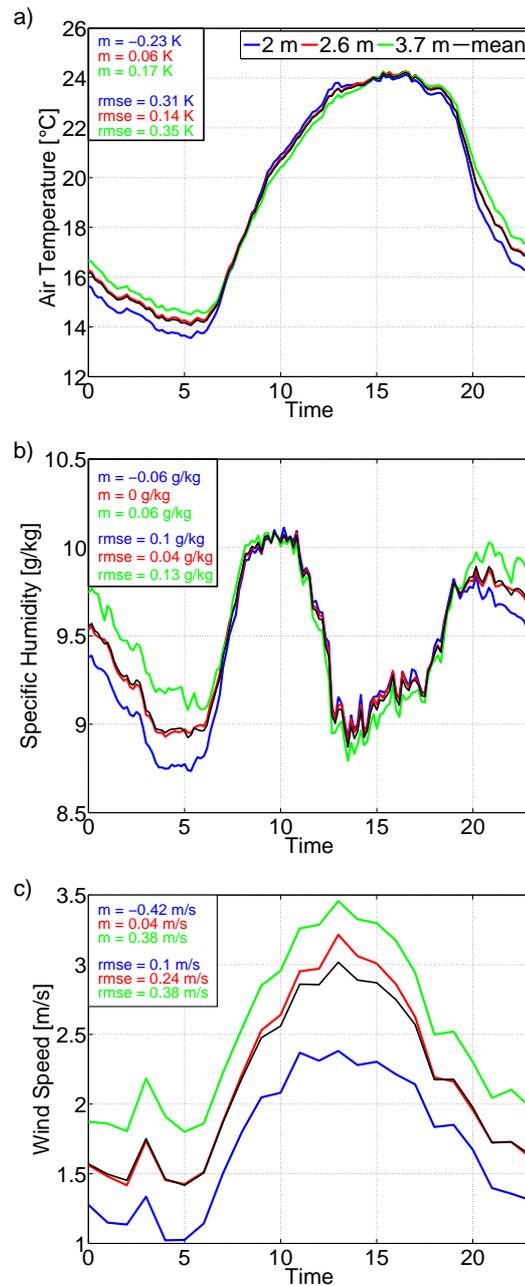


Figure 3.13: Mean diurnal cycle of a) air temperature, b) specific humidity and c) wind speed measured at heights of 2 m (blue line), 2.6 m (red line) and 3.7 m (green line) at the Merken site. The average of all three heights is presented by the black line.

2 m are presented in blue, the middle level at a height of 2.6 m in red and the upper level at a height of 3.7 m in green. The mean of all three sensors is shown as the black line.

As expected, the air temperature is spatially almost homogeneous during the day. In the afternoon, the air temperature is almost the same for all three heights. At night, the

lowest sensor measures the lowest temperatures, the highest sensor measures the highest temperatures. Therefore, the mean difference between the sensor and the mean of all sensors for the whole period is highest for the sensor at a height of 3.7 m with 0.17 K. For the middle sensor, it is only 0.06 K and the lowest sensor has a negative mean difference of -0.23 K leading to a difference between upper and lower sensors of 0.4 K. This is ten times the mean difference between all network sensors during the calibration period. The RMSEs of the differences between sensor and mean of all sensors vary between 0.35 K for the upper sensor where the wind is stronger and the air is mixed thoroughly and 0.14 K for the middle sensor that is closest to the mean.

For specific humidity in Fig. 3.13b, the picture is similar. During the day, when the atmosphere close to the surface is well-mixed, the specific humidity is almost the same for all three heights. For nighttime, the specific humidity increases with height. The middle sensor is almost identical to the mean of all sensors and has the same mean as all sensors put together. The upper sensor has a mean difference of 0.06 g kg^{-1} and the lower sensor of -0.06 g kg^{-1} . This is an order of magnitude larger than the mean difference between the network sensors during the calibration period. The RMSE lies between 0.04 g kg^{-1} for the middle and 0.13 g kg^{-1} for the upper sensor.

The wind speed in Fig. 3.13c increases with height. At the upper sensor, the mean wind speed is 0.8 m s^{-1} more than at the lower sensor. The mean difference between all sensors during the calibration period was only 0.002 m s^{-1} . The RMSE increases with height as well from 0.1 m s^{-1} at a height of 2 m to 0.38 m s^{-1} at a height of 3.7 m.

This experiment emphasises that the low cost sensors give measurements that are accurate enough to resolve differences in air temperature, humidity and wind speed in different heights.

Chapter 4

FLUXPAT 2009

The campaign *Fluxes and Patterns* in the Soil-Vegetation-Atmosphere System 2009 (FLUXPAT2009) was organised by the *Transregional Collaborative Research Centre 32* (TransRegio32). The project TransRegio32 is a collaboration of the Universities of Aachen, Bonn and Cologne and the Research Centre Jülich. The aim of this project is a better understanding of soil-vegetation-atmosphere interactions by combining geophysics, soil and plant science, hydrology, meteorology and mathematics. Processes, structures and patterns in soil, vegetation and the lower atmosphere are very complex and heterogeneous. They vary on different spatial and temporal scales. Measurements are taken on catchment to local scales to compute heterogeneous fluxes between soil, vegetation and atmosphere. These detailed measurements can be used for running high resolution models and to increase the predictive capabilities of the soil-vegetation-atmosphere-system. The catchment site for this first campaign is located in western Germany near the villages Selhausen and Merken (Fig. 4.1). It is an inhomogeneous terrain, consisting of different kinds of agriculture, settlement, a small grove and the narrow river *Rur*. On the basis of the measurements on this catchment site, the following questions are addressed: How variable are near-surface meteorological conditions? What causes these variabilities? Are there stations that are effectively redundant and need to be observed only once? How large is the error, if fewer stations are used, compared to the full network?

Another field of application of such a small scale network is the comparison with high resolution models. Air temperature and specific humidity simulations of the mesoscale atmospheric model FOOT3DK (see Chapter 4.5) are examined by comparing them to network time series. Is this model capable of reproducing the local features in these meteorological conditions? Heat flux simulations of the soil-vegetation-atmosphere transfer scheme TERRA (see Chapter 4.6) are used to address the question: How important is the consideration of heterogeneous forcing and vegetation for modeling heat fluxes?

Chapter 4.1 describes the measurement site and the deployment. Chapter 4.2 quantifies the variability of air and surface temperature, specific humidity and wind speed within the site. Explaining factors for these variabilities are investigated in Chapter 4.3. In Chapter 4.4, the stations are divided objectively into clusters and representative stations are identified. The error made by using only some of the stations is estimated. Finally a com-

parison to the FOOT3DK and a model study with the TERRA model are presented in Chapter 4.5 and 4.6.

4.1 Deployment

The domain of the FLUXPAT2009 campaign is located near Jülich in western Germany (see map in Fig. 4.1) between $50^{\circ} 51' 20''\text{N}$, $6^{\circ} 25' 35''\text{E}$ and $50^{\circ} 52' 10''\text{N}$, $6^{\circ} 27' 2''\text{E}$. It is a relatively flat terrain including a river valley with a mean elevation of ~ 100 m above sea level. The height difference between the river valley and the highest situated station, station 02, was approximately 10 m. For more information about the research area, see also Koyama et al. (2010).

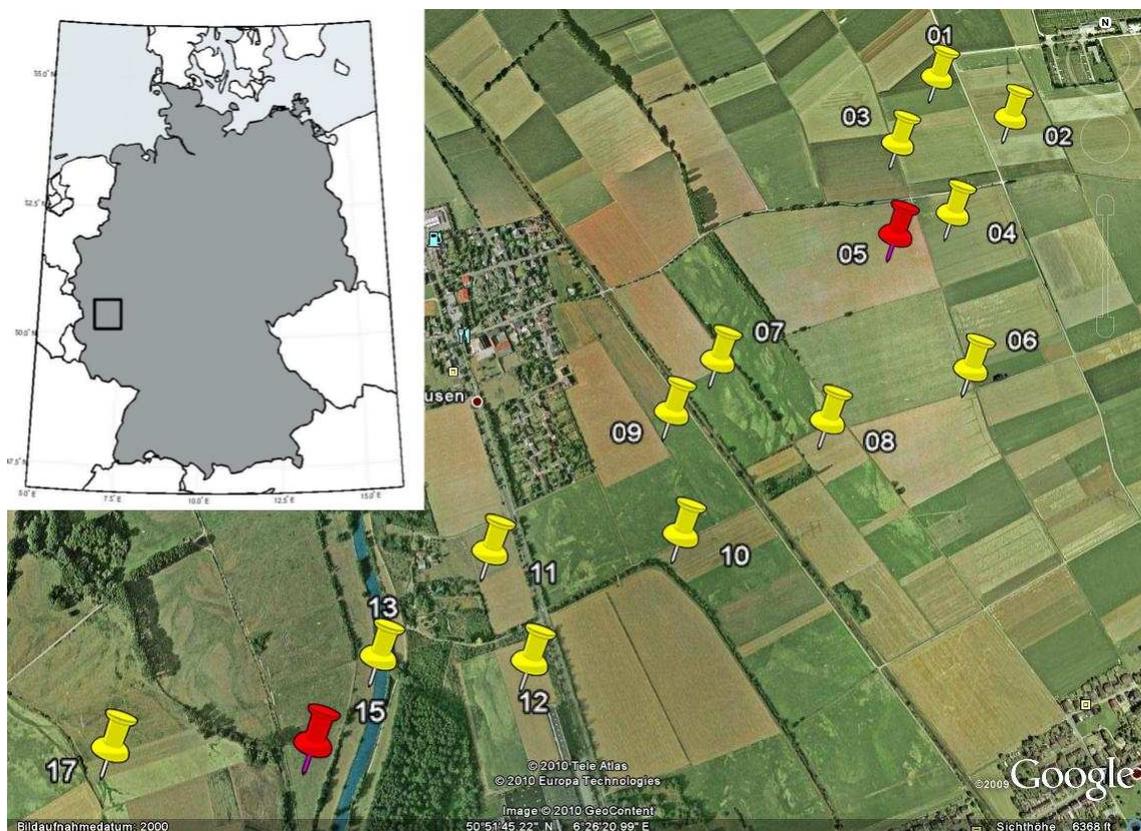


Figure 4.1: Deployment within the FLUXPAT2009 project in western Germany: Yellow markers represent slave stations, red markers represent master stations (source: Google Earth).

The measuring campaign with the sensor network operated from August 7th to August 26th, 2009. As illustrated in Fig. 4.1, two master stations (stations 05 and 15) and 13 slave stations were deployed as a double transect. Because of that concept, the network covered a variety of vegetation, e.g. grassland or arable land, and pairs of stations are comparable to identify erroneous measurements. Stations 04 and 06 were not operational during the whole measuring period and are, hence, excluded from the following analysis. Table 4.1

gives a list of all stations. The transect was 2.3 km long and the smallest distance between neighbouring stations was 140 m (between stations 07 and 09), the biggest distance was 480 m (between stations 05 and 07). Station 01 was the only site that was not used agriculturally; for research purposes, it was covered only by bare soil without any vegetation. Station 03 was sited on a potato field, station 13 was sited next to the river and station 15 on grassland close to the river. All other stations were located on acres with different kinds of crops, e.g. wheat at stations 05 and 10 or canola at stations 12 and 17. All crops were already harvested at the time of measurement. At stations 08, 09, 10, 12 and 17, the harvested crops covered the surface and the soil was not visible. At stations 02, 03, 05, 07 and 11, the harvested crops were already removed and the soil was visible.

The mean fetch for each station was calculated weighted with the frequency of the wind directions north, east, south and west. Most of the stations had a fetch between 50 m and 100 m. Station 11 had the smallest fetch with only 23 m, followed by stations 10 and 09 with 42 m and 49 m, respectively. These stations were located on the edges of fields. The longest fetches occurred at station 12 with 113 m and at station 05 with 189 m (see Table 4.1).

All stations were equipped with a wind sensor (*WIND*) facing north at a height of 2 m, an air temperature and humidity sensor (*TRH*) at a height of 1.5 m and an infrared sensor for measuring surface temperature (*TIR*). Stations 03, 09, 10, 12, 13 and 15 were additionally equipped with a sensor for measuring shortwave incoming radiation (*FSW*), a soil temperature and moisture sensor (*SM*) at a depth of 3 cm to 6 cm and a rain gauge (*RAIN*). At station 01, rain and shortwave radiation was also observed, at station 05, two soil temperature and moisture sensors at depths of 4 and 10 cm were used and station 17 was additionally equipped with a rain gauge and a soil temperature and moisture sensor. Table 4.1 gives an overview of the sensors at all 13 stations. *TRH*, *TIR*, *WIND* and *FSW* sensors displayed one minute means, *SM* sensors displayed five minute means. Accumulated precipitation was recorded in five minute intervals.



Figure 4.2: Station 15 in the front and a reference station in the back.

Table 4.1: Geographical position, land use, distance to the river (*dr*), vegetation, a mean fetch and the attached sensors for each station of the transect. TRH stands for temperature and humidity sensor, TIR stands for infrared surface temperature sensor, WIND stands for wind sensor, FSW stands for global radiation sensor, RAIN stands for rain gauge and SM stands for soil moisture and temperature sensor.

<i>No.</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Land Use</i>	<i>dr</i>	<i>Vegetation</i>	<i>Fetch</i>	<i>TRH</i>	<i>TIR</i>	<i>WIND</i>	<i>FSW</i>	<i>RAIN</i>	<i>SM</i>
01	50°52'9.3"N	6°27'1.1"E	no veg.	1.40 km	no vegetation	78 m	✓	✓	✓	✓	✓	
02	50°52'6.0"N	6°27'8.9"E	soil vis.	1.50 km	crop (harv.)	71 m	✓	✓	✓			
03	50°52'3.9"N	6°26'55.9"E	soil vis.	1.24 km	potatoes	90 m	✓	✓	✓	✓	✓	✓
05	50°51'56.8"N	6°26'54.7"E	soil vis.	1.15 km	wheat (harv.)	189 m	✓	✓	✓			✓
07	50°51'47.6"N	6°26'34.8"E	soil vis.	0.71 km	crop (harv.)	63 m	✓	✓	✓			
08	50°51'43.4"N	6°26'45.9"E	soil cov.	0.92 km	maize (harv.)	56 m	✓	✓	✓			
09	50°51'44.0"N	6°26'29.8"E	soil cov.	0.61 km	barley (harv.)	49 m	✓	✓	✓	✓	✓	✓
10	50°51'35.9"N	6°26'30.4"E	soil cov.	0.57 km	wheat (harv.)	42 m	✓	✓	✓	✓	✓	✓
11	50°51'34.8"N	6°26'11.3"E	soil vis.	0.20 km	crop (harv.)	23 m	✓	✓	✓			
12	50°51'28.0"N	6°26'15.4"E	soil cov.	0.26 km	canola (harv.)	113 m	✓	✓	✓	✓	✓	✓
13	50°51'28.3"N	6°26'0.7"E	grassland	0.02 km	grassland	57 m	✓	✓	✓	✓	✓	✓
15	50°51'23.2"N	6°25'54.9"E	grassland	0.11 km	grassland	60 m	✓	✓	✓	✓	✓	✓
17	50°51'22.9"N	6°25'35.7"E	soil cov.	0.45 km	canola (harv.)	93 m	✓	✓	✓		✓	✓

To validate the quality of the low cost network sensors, a reference station was deployed adjacent to both stations 01 and 15. The reference stations were equipped with sensors for measuring air temperature and relative humidity, surface temperature, wind speed and direction, precipitation and pressure (only at station 15) made by *Campbell Scientific*. Table B.1 in the Appendix B gives a list of the sensors. Station 15 and its reference station are shown in Fig. 4.2.

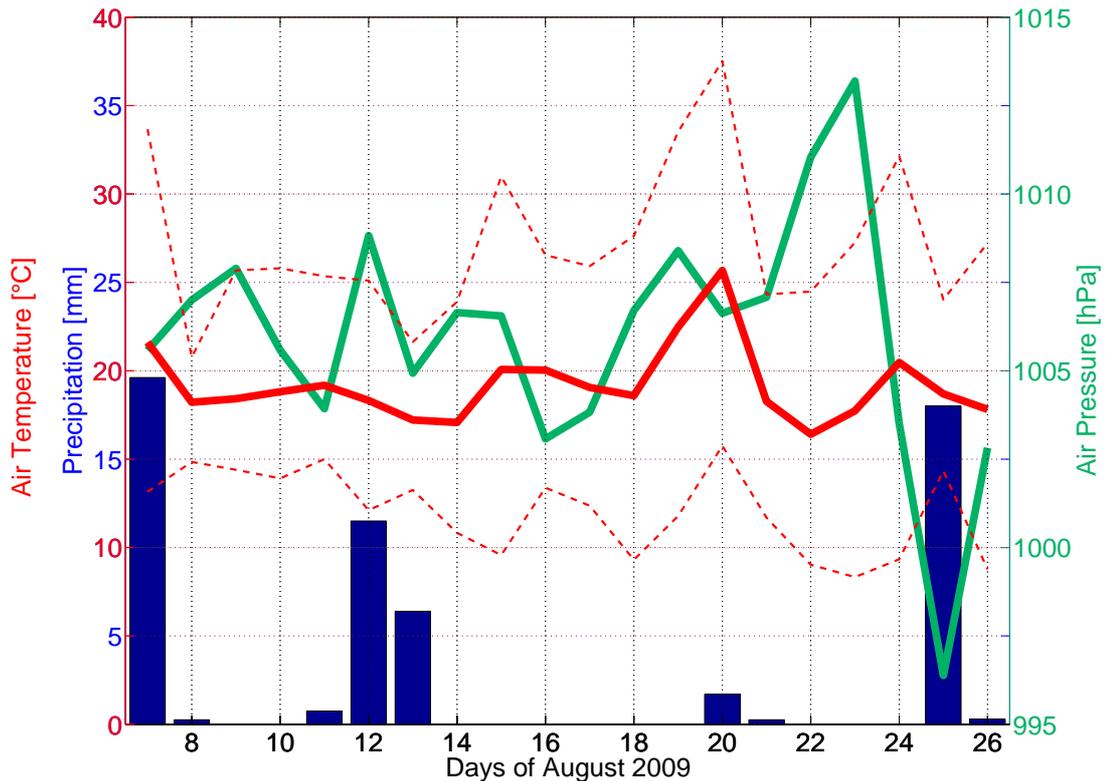


Figure 4.3: Weather observations during the FLUXPAT2009 campaign. The solid red line indicates daily mean air temperature, the dashed red lines indicate the maximum and minimum temperature, the green line indicates daily mean air pressure. Blue bars represent daily sums of precipitation.

To give an overview of the weather situation during the campaign, Fig. 4.3 shows daily mean air temperature, pressure and daily precipitation averaged over all network stations. August 2009 was a warmer than average month with a mean air temperature of 19.2°C at 2 m height. The daily mean air temperature ranged from 16.4°C on August 22nd to 25.7°C on August 20th. The highest one-minute mean temperature was also reached on August 20th with 37.5°C at station 08 at 3 p.m., the lowest being 8.3°C at station 11 on August 23rd at 4:45 a.m.. The daily mean air pressure exceeded 1005 hPa for most of the days. The wind came from the north or northwest for the first half of the campaign and

turned to the southeast direction in the second half.

The following summary of the synoptic situations during the campaign is based on six-hourly analyses of the German Meteorological Service (wetter3.de (2011)). Overall, six cold fronts, five warm fronts and three convergence lines passed the measurement site between August 7th and 26th, 2009. On the first four days, a frontal system dominated the weather situation with quick changes between warm and cold fronts and winds from the north. During August 7th, a rain event occurred with around 20 mm precipitation. On August 11th and 12th, a warm front closely followed by a cold front passed western Germany. The second of these frontal systems brought rain in the night from the 12th to 13th. The wind shifted towards the northwest. At noon on August 16th, the site was under the influence of a cold front. The wind turned to south-easterly direction on August 18th bringing warmer air masses. A convergence line crossed the area during the evening of August 20th, closely followed by a cold front bringing light rainfall. A similar situation occurred on August 24th and 25th. The air pressure dropped below 997 hPa. This time, both the convergence line and the cold front brought rain showers with up to 8 mm h^{-1} . At the end of the campaign on August 27th, the site was again under the influence of a cold front and southerly winds. On August 14th and 15th as well as on 22nd and 23rd, a high pressure system influenced the weather at the site, leading to high temperatures and low dew points. From August 18th to 20th, the site was in a warm sector between a warm and a cold front. Daily mean temperatures reached values of up to 25.7°C and the air was distinctly drier. During the whole campaign, there were two completely overcast days (August 8th and 13th) and five cloudless days (August 15th, 16th, 19th, 23rd and 24th). Overall, the weather is diversified during the FLUXPAT2009 campaign. That gives the opportunity to investigate the influence of different weather conditions, e.g. low or high wind speeds, on small scale variabilities in near surface atmospheric conditions.

4.2 Variability

One of the main goals of this experiment is to describe spatial variability of meteorological conditions close to the surface, e.g. air temperature, humidity, surface temperature and wind speed, and identify their explaining factors. To get a first idea of the magnitude of the spatial variability, mean values and mean diurnal cycles of these parameters for every station will be discussed.

The air temperature averaged over the whole measurement period is presented as circles in Fig. 4.4a. The mean temperatures of the reference systems at stations 01 and 15, labelled R 1 and R 15, are displayed as squares. Most of the sites have mean air temperatures between 19.2°C and 19.4°C . The highest temperatures occur at stations 01 and 07 with 19.7°C . These stations are located on fields with hardly any vegetation and, therefore, no shadowing. Station 13 registers the lowest temperature with 18.9°C . It is sited next to the river on grassland and the incoming solar radiation is shaded by surrounding vegetation. This results in a difference of 0.8 K in mean air temperature across a distance of 0.8 km. This difference is significantly higher than the network-to-network difference

of 0.04 K and their RMSE of 0.25 K during the calibration period (see Chapter 3.1). It is even larger than the quantile range (difference between the 10% and 90% quantile). Therefore, the variability in the network air temperature is not only caused by measurement uncertainty. External influences also contribute to the variability. This assumption is also supported by the air temperature measurements of the reference systems: Station 01 matches the reference system perfectly; both give mean temperatures of 19.7°C. Station 15, located on grassland close to the river, has a low mean temperature of 19.2°C. This is in very good agreement with the reference station, which gives a 0.1 K lower mean temperature of 19.1°C. Both systems show distinct differences between the mean air temperature at site 01 and 15: The difference amounts to 0.5 K for the network and to 0.6 K for the reference system.

Mean diurnal cycles of 10 minute mean air temperature for each station are shown in Fig. 4.4b. The temperature ranges between 13.7°C around 6:30 a.m. and 26°C around 4 p.m.. It is evident from the mean diurnal cycles that the lower mean air temperatures at stations 13 and 15 are mainly caused by strong cold air formation during the night. Both stations are sited on grassland, which is lower than the surrounding fields. Therefore, this small valley promotes nocturnal cold air formation and the temperatures at stations 13 and 15 are up to 1 K lower than the mean of all stations. In contrast, station 15 is one of the warmest sites during the day, most likely because it is sheltered from the wind by surrounding trees. Stations 01, 02 and 03 are warmest at night because they are not sheltered from wind and, therefore, they are subject to the well mixed layer immediately above the ground.

The network stations provide relative humidity data. To eliminate the influence of air temperature on the humidity measurements, specific humidity is calculated using relative humidity and air temperature measurements of the network stations and air pressure measurements of the reference system near station 15. The averaged specific humidity over the whole campaign for each site is displayed in Fig. 4.5a. Stations located in the northern part of the area are drier than stations in the southern part closer to the river. The driest site is station 05 with 9.23 g kg⁻¹, the wettest site is station 17 with 9.68 g kg⁻¹. That leads to a difference of 0.45 g kg⁻¹ across a distance of 1.8 km. The reference systems at stations 01 and 15 indicate an even larger difference of 0.63 g kg⁻¹ across a distance of 2 km. While station 15 ($q = 9.67 \text{ g kg}^{-1}$) underestimates the reference station ($q = 9.85 \text{ g kg}^{-1}$) by 0.18 g kg⁻¹, station 01 ($q = 9.4 \text{ g kg}^{-1}$) overestimates the reference station ($q = 9.22 \text{ g kg}^{-1}$) by the same amount. Therefore, it may be concluded that the variability in specific humidity might be even larger than the network indicates. The network-to-network mean difference during the calibration period (Chapter 3.1) is two orders of magnitude smaller (0.004 g kg⁻¹). Even the RMSE (0.18 g kg⁻¹) is half as large during the calibration as the difference between the mean specific humidity of stations 01 and 15 during the FLUXPAT campaign. This indicates that the differences in mean specific humidity between the network stations during the FLUXPAT campaign are site specific and not random effects of measurement uncertainties.

The mean diurnal cycles in Fig. 4.5b illustrate the typical two maxima in specific humidity

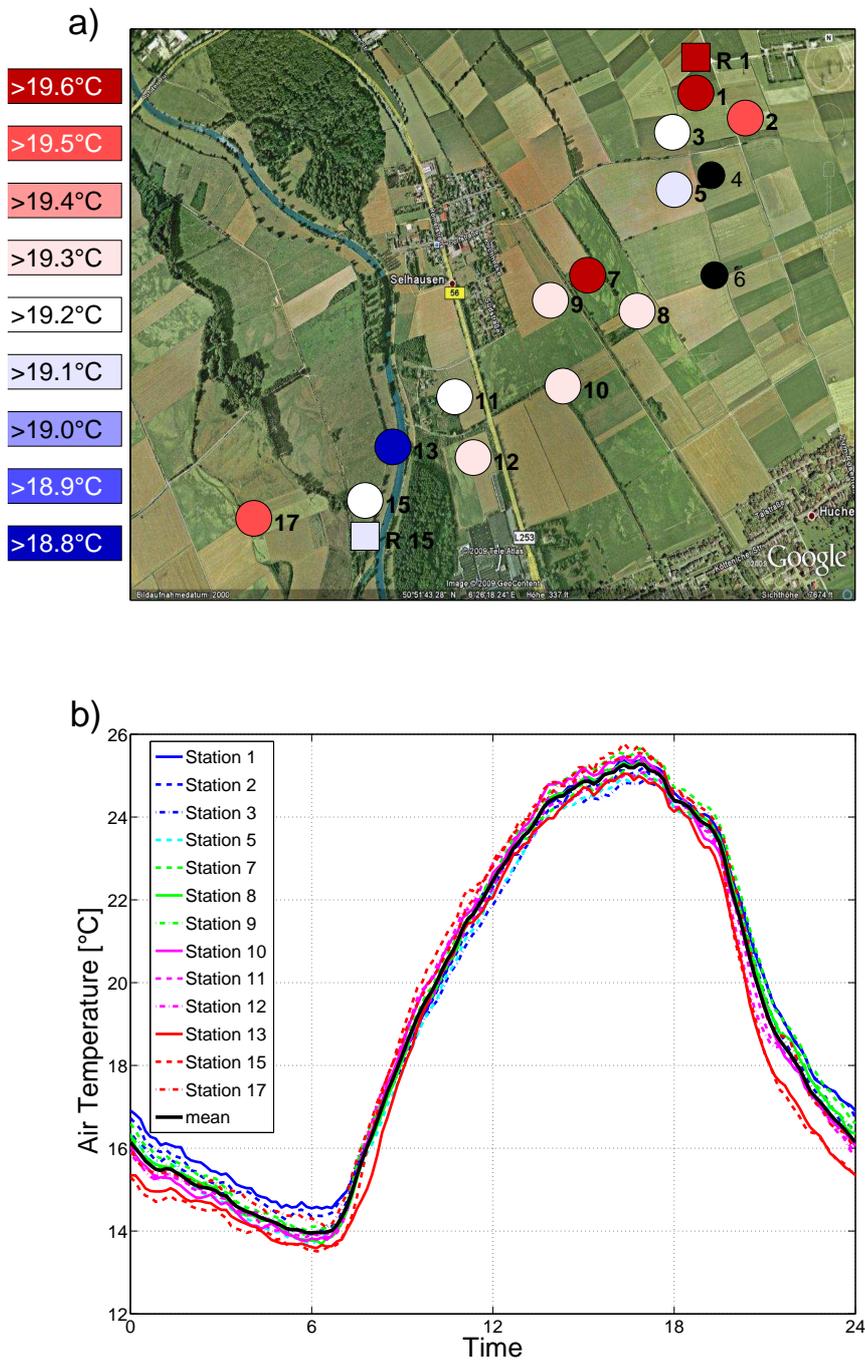


Figure 4.4: a) Map of mean air temperatures during the FLUXPAT2009 campaign: Circles stand for network stations, squares stand for reference stations. b) Mean diurnal cycle of mean air temperature for every station (coloured lines) and the mean of all stations (black line) with ten minute temporal resolution.

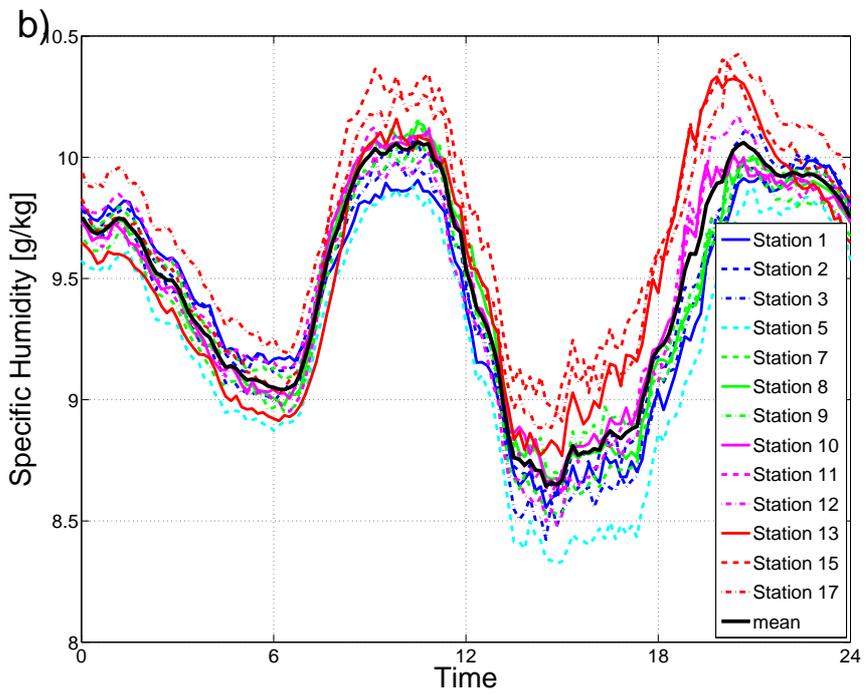
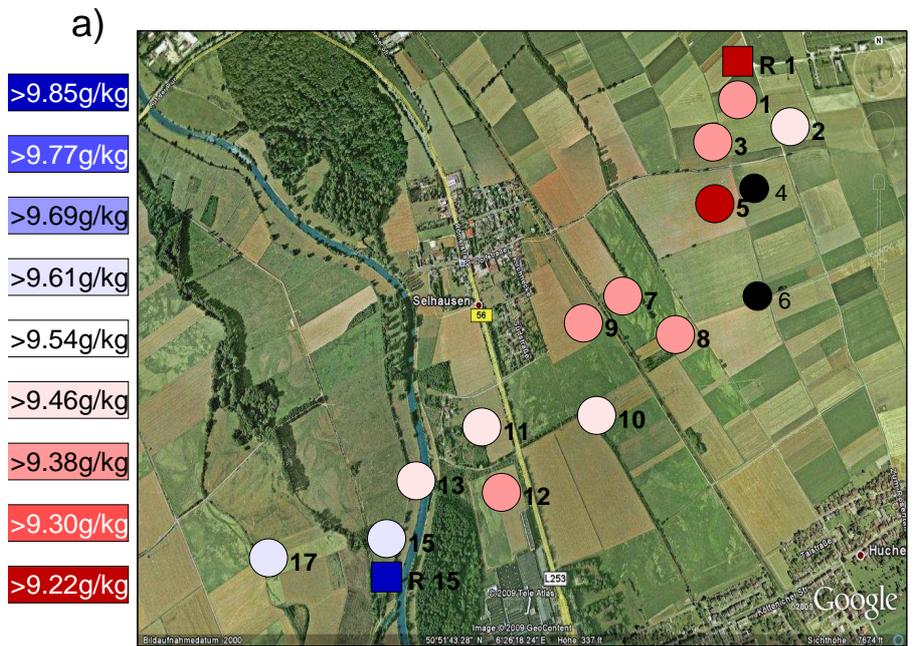


Figure 4.5: Same as Fig. 4.4, but for specific humidity.

during the day: One around 10 a.m. and one around 8 p.m.. After sunrise, evaporation increases and consequently the specific humidity also increases. In the late morning, the mixing process in the atmosphere close to the surface sets in. Therefore, the moist air is taken away to higher atmosphere layers and the specific humidity decreases. In the afternoon, the mixing process diminishes and evaporation causes the second maximum in the evening. The values of humidity range between 8.4 g kg^{-1} and 10.4 g kg^{-1} . Stations 15 and 17 measure higher humidities than all other stations throughout the day. The humidity ranges between 8.9 g kg^{-1} and 10.4 g kg^{-1} for these two stations and is thus up to 0.5 g kg^{-1} higher than the average of all stations. The diurnal cycle of station 13 has the largest amplitude. During the night, it is one of the driest stations. After sunrise, it becomes more moist and around 7 p.m. it is the wettest station. This could be caused by evaporation of the river. Evaporation increases during the day because of increasing temperatures. That leads to higher specific humidity. At night, temperature decreases and evaporation stops due to the saturated air. The site closest to the river is surrounded by hedge banks. Low wind speeds (discussed later in this chapter) and, therefore, limited mixing of the near surface atmosphere could be another reason for the high values of specific humidity during day. Station 05 is the driest site the whole day. During the night, the specific humidity is 0.2 g kg^{-1} lower than the average of all stations; during day, it is up to 0.48 g kg^{-1} lower. This station is located on a large field, where the corn has already been harvested, so there is no vegetation and the bare soil is visible. Therefore, the evaporation might be less than for other sites. It also has the largest fetch of all stations, meaning it is not sheltered from the wind and, therefore, is subject to the well mixed atmosphere close to the surface.

For surface temperature, the mean for each station is depicted in Fig. 4.6a. It varies between 18.24°C at station 15 on grassland and 21.43°C at station 03 on the potato field. That results in a difference between these two stations of 3.19 K across a distance of 1.9 km . The mean network-to-network difference during the calibration period is only 0.13 K (see Chapter 3.2). Also, the difference between the 10% and 90% quantile during the calibration period is almost four times smaller than the difference of the mean surface temperature between stations 03 and 15. Hence, the differences during the FLUXPAT campaign are induced by external influences. The surface temperature at station 13 next to the river is almost as low as at station 15. All other stations measure mean temperatures of more than 20°C . Sites where the soil is covered with harvested corn, e.g. stations 08, 09, 10, 12 and 17, show similar surface temperatures in the range of 20.1°C to 20.5°C . Highest mean surface temperatures of more than 21°C occur at sites where the bare soil is visible, i.e. stations 02, 03, 05, 07 and 11. The reference system at station 15 matches the network station perfectly; at station 01 a difference of 0.8 K is evident. The reason for this discrepancy could be different fields of view for both sensors. While the reference sensor has an aperture angle of 20° (campbellsci.com, 2011d), the network sensor has an aperture angle of more than 26° . Therefore, the network sensor averages over a larger area. The spatial variability in surface temperature on grassland at station 15 is less than the variability at station 01. Although there is no vegetation at station 01, the surface temperature might

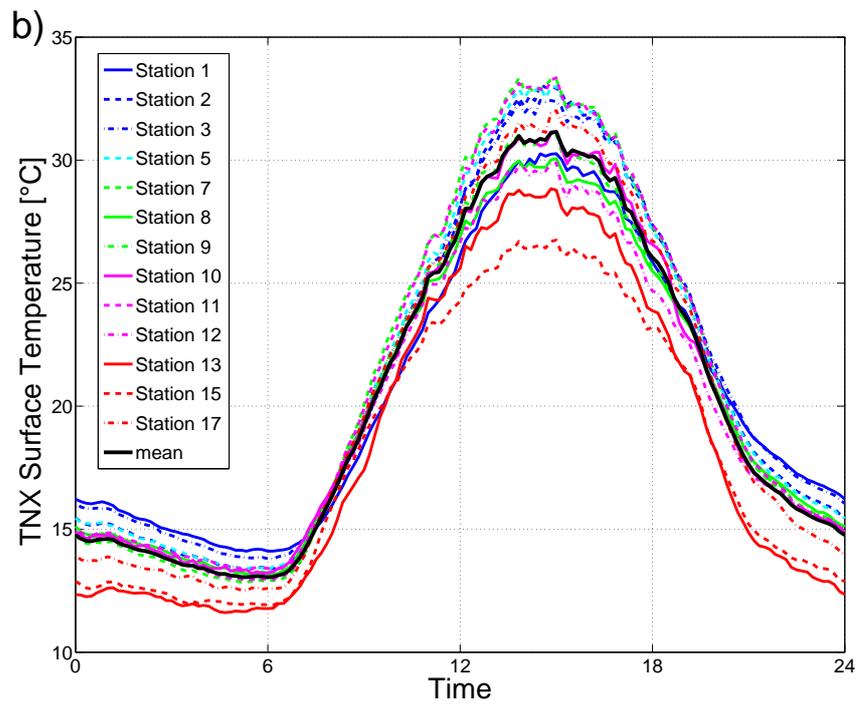
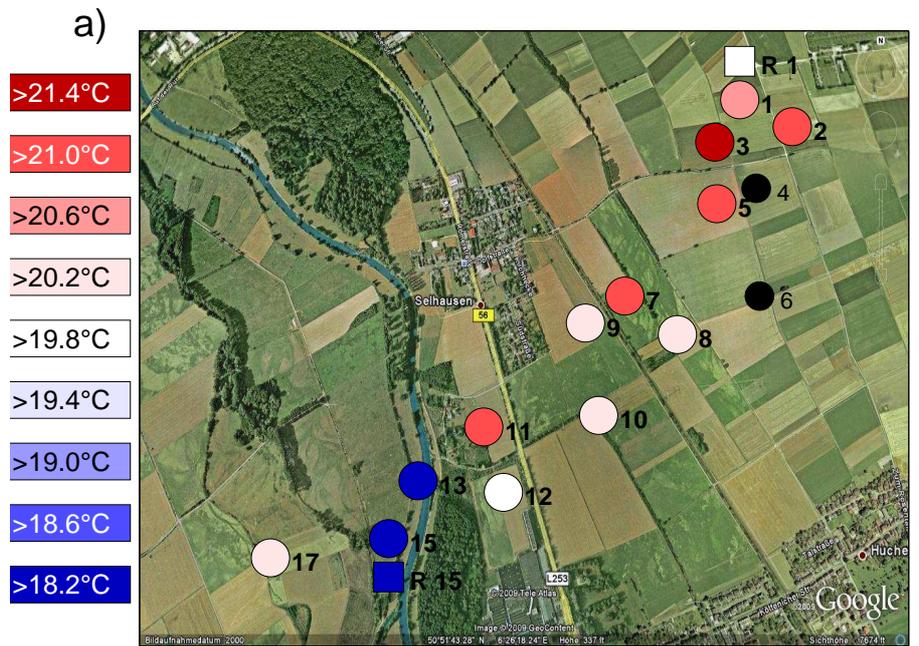


Figure 4.6: Same as Fig. 4.4, but for surface temperature.

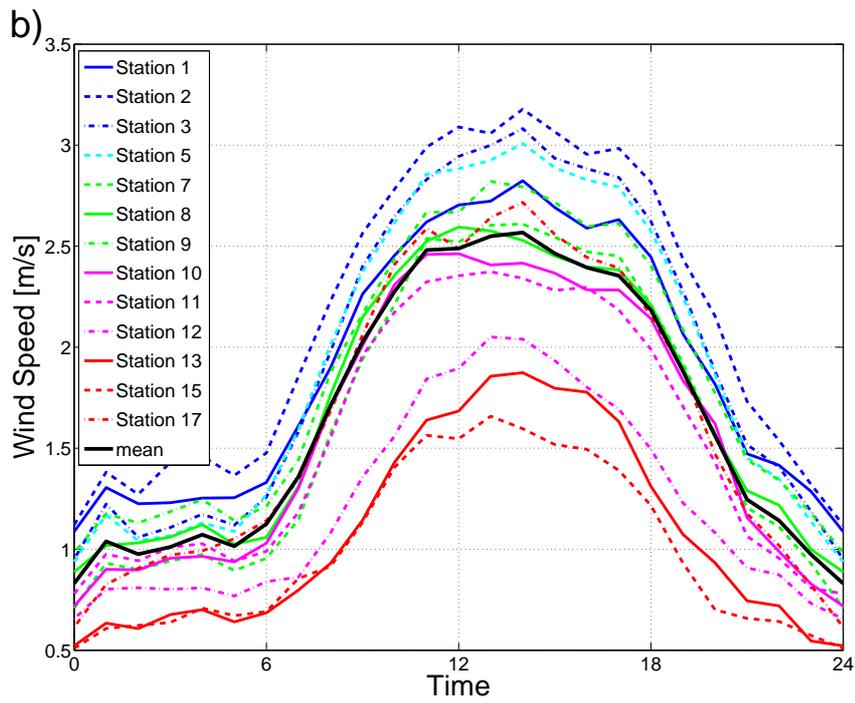
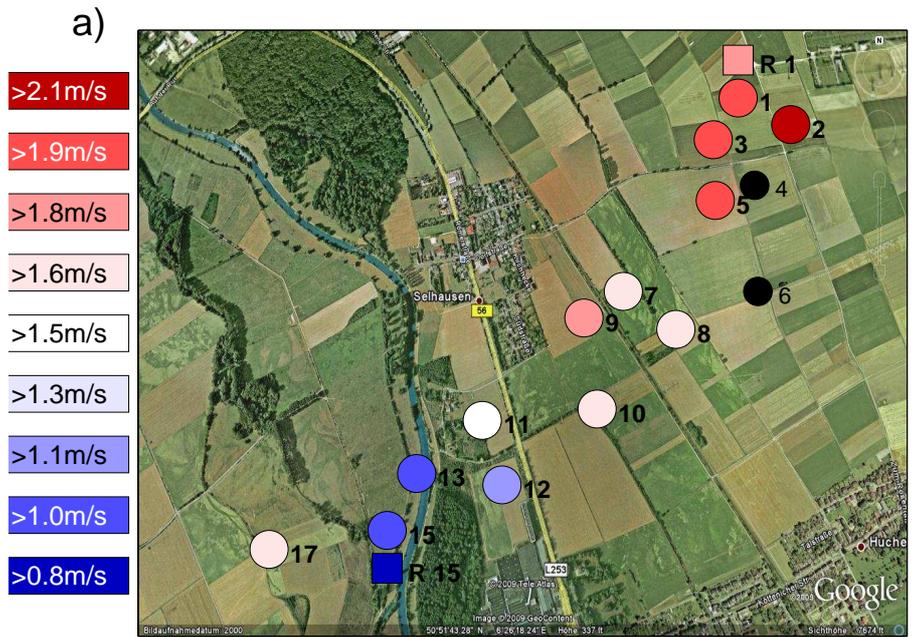


Figure 4.7: Same as Fig. 4.4, but for wind speed.

be influenced by exposed rock. Therefore, the difference between the reference and the network sensor is higher at station 01.

In Fig. 4.6b, mean diurnal cycles are presented for each network station. The surface temperatures range between 12°C at 6 a.m. and 33°C at 3 p.m.. At night, the sites with bare soil are the warmest (stations 01, 02, 03 and 05). This might be caused by more efficient heat storage. During the day, stations 09 and 11 show the highest surface temperatures, while station 01, which was warmest during the night, is colder than the average of all stations. This could be due to the surface colour. The surface at stations 02 to 05 consists of dark, dug-up soil. At station 01, the surface colour is lighter and the albedo higher. Stations 13 and 15, located above grasslands, have the lowest surface temperatures throughout the day.

An overview of mean wind speed at each site is shown in Fig. 4.7a. For the network sensors, it ranges between 1 m s⁻¹ for well-sheltered stations and 2.1 m s⁻¹ for stations in open areas. During the calibration, the mean difference between the network sensors was only 0.002 m s⁻¹. At the grassland sites at station 13 and 15, which are surrounded by trees and hedge banks, mean wind speeds are lowest (1.1 m s⁻¹ and 1.0 m s⁻¹, respectively). Also, at station 12, the wind speed is quite low with an average of 1.25 m s⁻¹. This station is located on a field surrounded by hedge banks and a building, which shelter the site from wind. Stations 07, 08, 10, 11 and 17 are all sheltered from the wind on at least one side. Their mean wind speed is in the range of 1.5 m s⁻¹ to 1.8 m s⁻¹. On the contrary, stations 01, 02, 03, 05 and 09 stand on open fields and, hence, the mean wind speed is more than 1.8 m s⁻¹. The highest mean wind speed occurs at the highest situated station of the network (station 02). The reference systems at stations 01 and 15 both give lower mean wind speeds than the network sensors: Reference 01 measured 1.81 m s⁻¹ while station 01 measured 1.92 m s⁻¹. At site 15, the difference between both systems is almost twice as large. The mean wind speed of reference 15 is 0.83 m s⁻¹, the mean wind speed of station 15 is 1 m s⁻¹. This overestimation by the network sensor might be due to the different kinds of sensors used at both systems. The reference system uses a 2D sonic anemometer. The network sensor is a cup anemometer, which keeps turning even if the wind has stopped already (*overspeeding effect*). Although the cup anemometer also has a higher start-up velocity than the 2D sonic, the overspeeding effect leads to higher averaged wind speeds.

The mean diurnal cycle of the 10 minute averaged wind speed for each station and the average of all stations is presented in Fig. 4.7b. As expected, wind speeds are lower during the night (between 0.5 m s⁻¹ and 1.5 m s⁻¹) than during the day (between 1.5 m s⁻¹ and 3.2 m s⁻¹). At sites 12, 13 and 15, which are surrounded by trees, hedge banks and buildings, the maximum wind speed (2 m s⁻¹) is distinctly lower than the average of all stations. The amplitude of the diurnal cycle is less pronounced here than for other stations. Stations 07, 08, 10, 11 and 17 are less sheltered and observe higher wind speeds throughout the day. Station 08 matches the average of all stations. At the site where station 01 is located, there are no hedge banks or trees, but the field is used for scientific purposes. Therefore, the scientific instruments and other equipment block the flow slightly,

so that the maximum wind speed reached is 2.8 m s^{-1} at 2 p.m.. The highest wind speeds occur consistently at stations 02, 03 and 05. Even at night, the mean wind speed is not less than 1 m s^{-1} .

4.3 Explaining Factors for the Variability

In the previous chapter, it is stated that the variability in the meteorological conditions close to the surface between the network stations is caused by external influences. Four of the possible external influences, namely wind speed, time of day (i.e. daytime or nighttime), surface conditions and distance to the river, will be investigated in this chapter.

The data is divided into three different wind speed classes: Low wind speeds less than 1.2 m s^{-1} , moderate wind speeds between 1.2 m s^{-1} and 2.5 m s^{-1} and high wind speeds more than 2.5 m s^{-1} . An additional differentiation between day and night leads finally to six different groups. During the day, i.e. between sunrise and sunset, mostly moderate and high wind speeds occur. These groups include 9499 and 5518 data points, respectively. Low wind speeds are only found in 3210 cases. During the night (between sunset and sunrise), the differences in sample size are even larger. While the number of low and moderate wind speed cases (9004 and 2686, respectively) are of the same order of magnitude, high wind speeds occur only in 323 cases. Since there is no reference in the deployment which can be considered as the truth for all stations, the inter-station differences ΔT_{Air} , Δq and ΔT_{Surf} between all stations are calculated for each of the six groups. The statistics of these samples are summarised in terms of mean values and 10% and 90% quantiles illustrated in Fig. 4.8 to 4.10. Mean values represent systematic deviations of a station and quantile ranges reflect the range of random deviations for an arbitrary one minute interval.

Fig. 4.8 to 4.10 address all four aspects (wind speed, time of day, surface conditions and distance to the river) and their influence on air temperature, specific humidity and surface temperature. Different surface conditions are characterised by different colours. According to Table 4.1, station 01 is located on a field with no vegetation which is used for research rather than agricultural purposes. It is marked in black. Station 13 and 15, at grassland sites close to the river, are marked in blue. Green markers represent sites where the soil is covered by harvested corn (stations 08, 09, 10, 12 and 17). Sites that are already harvested and where the soil is visible are marked with red squares (stations 02, 03, 05, 07 and 11). The stations are ordered on the x-axis according to their distance to the river, starting with station 13 closest to the river on the very left side and ending with station 02.

Mean differences in air temperature are shown in Fig. 4.8. In general, the quantile range is two to three times larger than the mean differences. This means random deviations in air temperature variability outweigh systematic effects. Station 08 is closest to the mean under all conditions; stations 13 and 15 differ the most from the other stations. With mean differences between -0.91 K and 0.48 K , the latter two stations show the largest mean differences in almost all wind speed classes during the day and during the night.

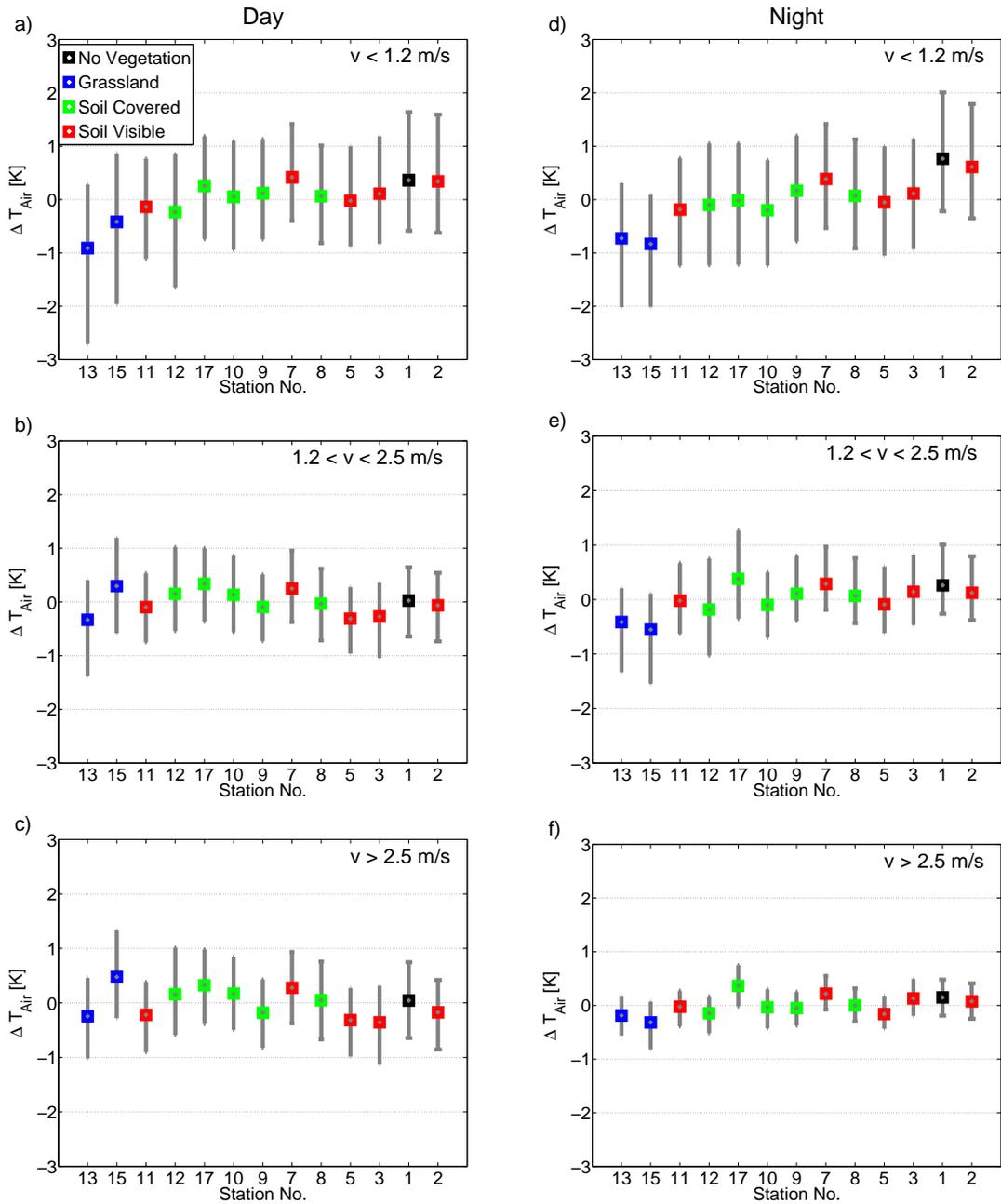


Figure 4.8: Mean difference and 10% and 90% quantile of air temperature between sunrise and sunset (Day in the left column) and between sunset and sunrise (Night in right column) for wind speeds lower than 1.2 m s^{-1} (top row), between 1.2 and 2.5 m s^{-1} (middle row) and higher than 2.5 m s^{-1} (bottom row). The colours indicate different surface types: Black for no vegetation, blue for grassland, green for covered soil and red for visible soil. On the x-axis, stations are ordered according to increasing distance to the river, starting with station 13.

The quantile range is largest at these two stations, as well.

For all groups, the mean differences range between 1 K and -1 K. They decrease with increasing wind speeds during the day as well as at night. For low wind speeds in Fig. 4.8a and d, the mean differences add up to ± 1 K. For high wind speeds in Fig. 4.8c and f, it is not more than ± 0.5 K. The quantile range Δq_{90-10} also decreases with increasing wind speeds, independent of the time of day. For low wind speeds, Δq_{90-10} ranges from 2 K to 3 K, for high wind speeds only from 0.5 K to 1.5 K. At higher wind speeds, the atmosphere close to the surface is well-mixed and the air temperature is distributed more homogeneously than at low wind speeds. Therefore, there are smaller differences between the stations at high wind speeds than at low wind speeds, leading to smaller values of Δq_{90-10} .

The variability and mean values of air temperature are almost identical for day and night for wind speeds smaller than 2.5 m s^{-1} . For high wind speeds, however, the variability is clearly lower at night. The quantile range for nighttime in Fig. 4.8f is half that for daytime in Fig. 4.8c. Obviously, wind speed is more important for mixing during stable nighttime situations. Another explaining factor for higher variability during the day, is the incoming solar radiation, which may lead to stronger heterogeneities in air temperature.

The influence of the river on air temperature is only evident for nearby stations. Stations 11 to 15 are colder than the network average during the day for low wind speeds and at night for all wind speeds. Sites that are located farther from the river are not affected. During the day, the influence of the river vanishes for wind speeds higher than 1.2 m s^{-1} . Station 15 is coldest during the night for all wind speeds. This station is located in a small valley and, therefore, the magnitude of a cold air pool is stronger than at other sites. This cold air pool is most evident at low wind speeds. Station 15 is almost 1 K colder than the network average. By contrast, however, it is one of the warmest sites for moderate and high wind speeds during the day.

Surface types seem to have no impact on air temperature regardless of wind speed. Indeed, station 13 and 15 on the grassland sites are colder than all other stations during the day for low wind speeds (Fig. 4.8a) and at night for all wind speeds, but this effect is rather due to the influence of the river than the surface type. For all other surface types, no trend is evident.

Fig. 4.9 presents mean differences and quantile ranges for specific humidity. In general, the quantile ranges are two to three times the mean differences. Therefore, random deviations dominate the variability in specific humidity.

The mean differences are mainly influenced by the time of day. During the day, they range between -0.35 g kg^{-1} at station 05 for high wind speeds in Fig. 4.9c and 0.4 g kg^{-1} at station 15 for low wind speeds in Fig. 4.9a. During the night, the mean differences are lower with values between -0.23 g kg^{-1} at station 05 and 0.33 g kg^{-1} at station 17 both for high wind speeds in Fig. 4.9f. The most obvious difference between day and night is the variability. During nighttime, Δq_{90-10} ranges from 0.43 g kg^{-1} to 1.04 g kg^{-1} . In the daytime, it is twice as large (1.06 g kg^{-1} to 1.96 g kg^{-1}). This is due to evaporation. During the day, all sites evaporate at different rates, causing diverse amounts of specific

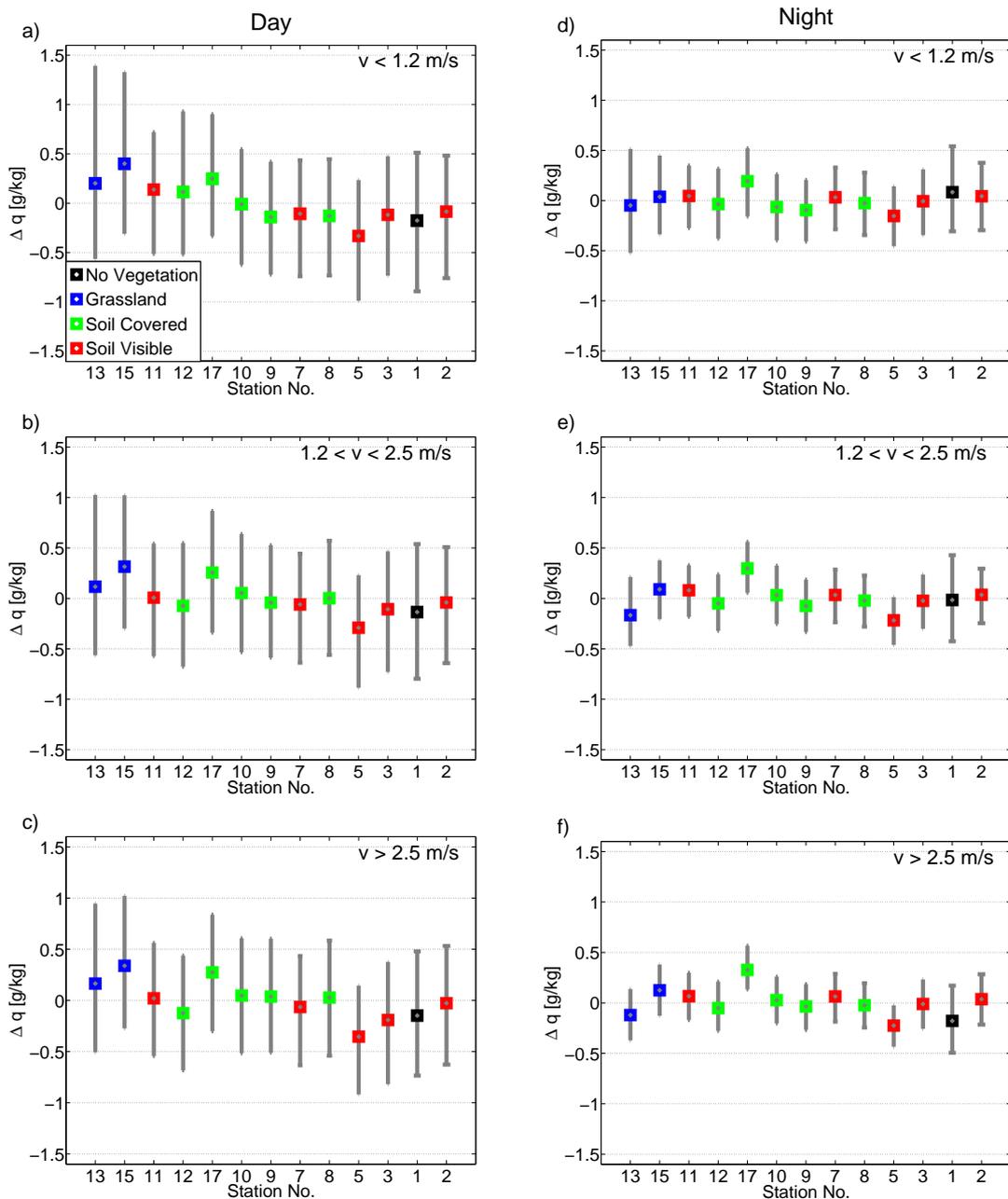


Figure 4.9: Same as Fig. 4.8, but for specific humidity.

humidity. Evaporation decreases at night. That leads to a more homogeneous pattern of specific humidity.

Wind speed, on the other hand, does not have an impact on the mean difference, neither during the day nor at night. This is a surprising result because a more homogeneous distribution of specific humidity with increasing wind speed is expected. Only station 01 seems to be influenced by wind speed during the night. For low wind speeds in Fig. 4.9d,

station 01 is 0.08 g kg^{-1} wetter than the network average. For moderate wind speeds in Fig. 4.9e, it corresponds to the average and for high wind speeds, it is 0.18 g kg^{-1} drier than average. This station also shows the largest quantile range at night for all wind speeds with up to 0.85 g kg^{-1} . In general, the variability in specific humidity is higher for wind speeds less than 1.2 m s^{-1} . The atmosphere close to the surface is mixed more thoroughly at higher wind speeds and, therefore, the specific humidity is distributed more homogeneously and the variability between the network stations decreases.

The river also seems to influence the specific humidity during the daytime, especially for low wind speeds (Fig. 4.9a) when the evaporation is strongest. Stations 11 to 17, closest to the river, are consistently wetter than the network average with mean differences between 0.11 g kg^{-1} and 0.4 g kg^{-1} . Stations 01 to 09 are 0.11 g kg^{-1} to 0.33 g kg^{-1} drier than average. At night the impact of the river on specific humidity vanishes.

Similar to air temperature, mean values and variability of specific humidity are not influenced by surface type. The fact that the grassland stations are distinctly wetter than all other stations during the day is most likely caused by the river rather than by surface type.

The mean differences in surface temperature in Fig. 4.10 range from -3.5 K at station 15 to 2.06 K at station 08, both during the day for high wind speeds. That leads to a mean difference of more than 5.5 K between these two stations. The maximum quantile range occurs also at station 15 with 7.25 K . In all six groups, Δq_{90-10} is less than twice the mean difference. For surface temperature, systematic errors play a more important role than for air temperature and specific humidity.

The time of day influences both mean differences and variability of surface temperature. During the day, the mean differences vary between -3.5 K and 2.06 K and the quantile ranges between 3.31 K and 7.25 K . Daytime surface temperature is clearly influenced by surface type. Stations 13 and 15, in the grassland area, provide the lowest surface temperatures for all wind speeds. Their mean difference to the other stations is more than -2 K . Surface temperatures at stations 08, 09, 10, 12 and 17, where the soil is covered with the stalks of the corn, is higher than for grassland stations. Their mean differences range from -1.09 K to 0.61 K . Station 01, located at the research site devoid of vegetation, has a mean difference in the same range. At this site, the soil is visible, but it has a light colour and, therefore, a high albedo leading to lower surface temperatures. At stations 02, 03, 05, 07 and 11, the bare, dark soil is visible. They show the highest surface temperatures for all wind speeds with mean differences between 0.9 K and 2.06 K . The variability is lowest for wind speeds less than 1.2 m s^{-1} for all surface types (Fig. 4.10a). This is most likely due to the fact that the wind speed is low in the morning and in the evening (see Chapter 4.2), when the solar radiation is less intense. Therefore, the variability in surface temperature is also smaller.

At night, mean differences and variability in surface temperature are distinctly lower than during the day. For low wind speeds in Fig. 4.10d, the variability is similar for day and night, but for moderate and high wind speeds (Fig. 4.10e and f), it is two to three times smaller at night. The higher the wind speed, the more homogeneous is the air

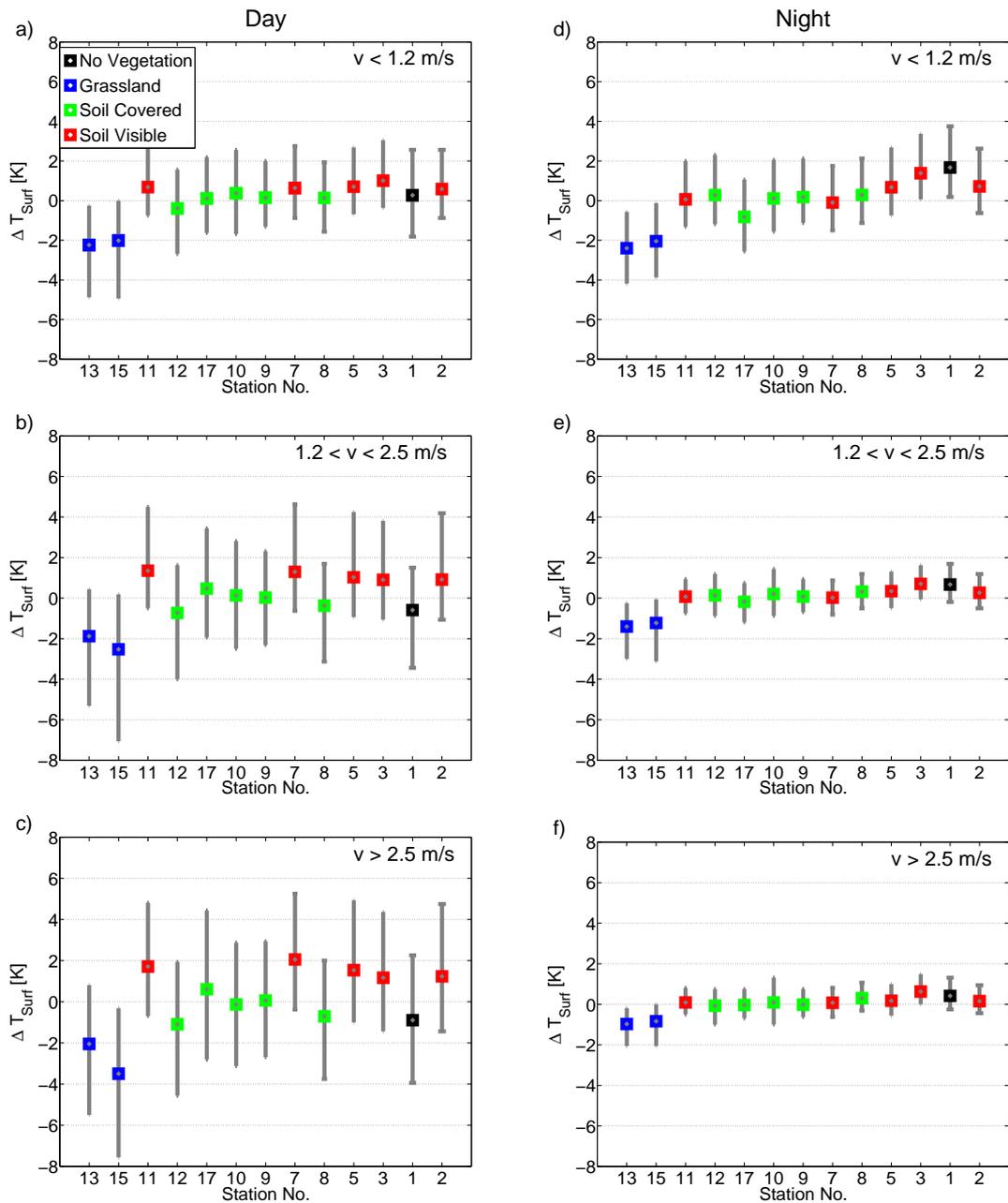


Figure 4.10: Same as Fig. 4.8, but for surface temperature.

temperature close to the surface (see Fig. 4.8). This causes also more homogeneous surface temperatures. The largest anomalies of nocturnal surface temperature are seen for wind speeds less than 1.2 m s^{-1} : Station 13 is 2.4 K colder than the network average and station 01 is 1.67 K warmer. The surface temperatures at grassland stations are again distinctly lower than the network average. For the other stations, there is no noticeable influence of the surface type. Therefore, nocturnal wind speed has a greater influence on surface

temperature than surface type, most likely because wind plays a more important role in the exchange of air masses under stable stratification.

The distance to the river does not affect the surface temperature. Indeed, the two stations closest to the river show lowest surface temperatures, but this effect is more due to the surface type than to the distance to the river.

The results in this chapter indicate that the variability in air temperature is mainly influenced by wind speed and the river, as expected. Stations close to the river are coldest, stations far away from the river are warmest. The mean differences and quantile ranges decrease with increasing wind speed. On the other hand, the mean differences in specific humidity are surprisingly unaffected by wind speed. They are mainly influenced by the distance to the river. Random deviations outweigh systematic effects for air temperature and specific humidity. For surface temperature, systematic effects dominate. The vegetation type is the dominant explaining factor for differences in surface temperature between the sites during the day. During the night, wind speed becomes more important.

4.4 Cluster Analysis

The factors explaining the variability in air temperature, specific humidity and surface temperature in the atmosphere close to the surface are examined in the previous chapter. From these results two central questions arise: Can stations that are exposed to similar external influences (e.g. surface type and wind) be represented by one single station? How much error would be introduced by representing the network with some fraction of the total number of stations? To identify groups of stations, cluster analyses are applied to the time series of meteorological conditions measured by the sensor network.

The purpose of cluster analysis is to collect elements into groups. Elements should be as similar as possible, while groups should be as diverse as possible. Different procedures exist for performing a cluster analysis. There are two principal methods: *Hierarchical* and *non-hierarchical* cluster analysis. The hierarchical method forms clusters either by stepwise refinement of the coarsest distribution or by merging the elements into clusters starting with the finest distribution. For non-hierarchical cluster analysis, the number of clusters has to be given a priori. Both methods share a common feature: Two clusters are either disjoint (i.e. no element is present in both clusters) or one cluster is containing the other. In the following, only a brief description of the used methods is given. See textbooks like Wilks (2006) for more information.

Hierarchical Cluster Analysis

There are two different kinds of hierarchical cluster analysis: Divisive (top-down) and agglomerative (bottom-up). In case of divisive cluster analysis, a stepwise refinement is carried out. The coarsest distribution is used as starting point, i.e. all elements are in one cluster, and this cluster is split into groups that are least alike until every element forms its own group. In case of the agglomerative method, it works the other way around. It starts with the finest resolution, meaning each element forms its own cluster, and then

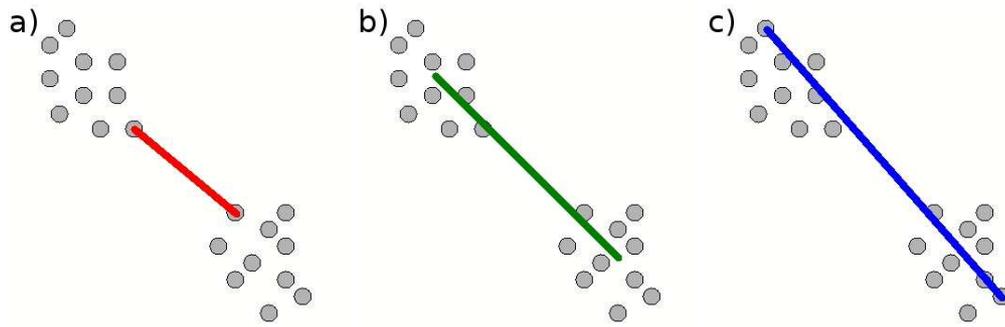


Figure 4.11: Distance measures between two clusters for a) Single Linkage Method, b) Average Linkage Method and c) Complete Linkage Method.

progressively merges elements or clusters, whose centres are closest to each other. Thus, hierarchical cluster analysis always gives the same results.

Different methods exist to determine the distance between clusters. The three most frequently used methods are *Single Linkage*, *Complete Linkage* and *Average Linkage Method*. In case of the Single Linkage or Nearest Neighbour Method, the shortest distances between elements or clusters are considered as shown in Fig. 4.11a. The element of a cluster closest to the other cluster is used for calculating the distance between two clusters. All elements within a certain distance to at least one element of the cluster are assigned to this cluster. In case of the Complete Linkage or Furthest Neighbour Method in Fig. 4.11c, the longest distance between elements or clusters is used as a distance measure. All elements are assigned to a cluster that are not more than a certain distance apart. For the Average Linkage Method, the mean of all elements (centroid) in a cluster is computed and the distance between these centroids is calculated (Fig. 4.11b). The root mean square distance (RMSD) serves as a measure of distance between two clusters X and Y with N_X and N_Y elements, respectively:

$$RMSD = \sqrt{\frac{\sum_{a=1}^{N_X} \sum_{b=1}^{N_Y} dist(X_a, Y_b)}{N_X \cdot N_Y}}, \quad (4.1)$$

where $dist(X_a, Y_b)$ is the Euclidean distance between X_a and Y_b .

The results of hierarchical cluster analyses are shown in so-called dendrograms. A sketch of such a dendrogram is presented in Fig. 4.12. It is a tree diagram, where a new cluster is formed at every node. The elements are listed on the x-axis and the distance in terms of RMSD is shown on the y-axis. Therefore, the RMSD for a certain number of clusters is evident from the dendrogram and it is up to the user to decide, how many clusters are appropriate in a certain case. However, a disadvantage of hierarchical methods is that once an element is assigned to a certain cluster it cannot be re-located, even if that would lead to better results. To avoid this effect, non-hierarchical cluster analyses have been developed.

$$x' = \frac{x - \bar{x}}{\sigma(x)}, \quad (4.2)$$

where \bar{x} is the average and $\sigma(x)$ is the standard deviation of x . Thus, the systematic differences between the stations, i.e. offset, and the diurnal cycle are eliminated. From a scientific point of view, it is interesting to investigate what role other influences play but in practise the offset and the diurnal cycle are important and need to be considered. The cluster analyses are also applied to hourly-averaged datasets. The resulting clusters are the same as for one minute measurements. Therefore, the clustering of the non-normalised one minute means is used for further investigations.

The k-means-Method is used for analyses for air temperature, specific humidity, surface temperature and wind speed to validate the results of the Average Linkage Method. The results are displayed in terms of the RMSD between the time series of the network stations and the centroids of the clusters. Each station is assigned to the cluster for which the RMSD to the centroid is smallest. To find out more about the characteristics of the formed groups and to identify certain differences between clusters, mean diurnal cycles of 10 minute means for the whole campaign are computed for each cluster.

Setting up 13 stations in an area of 1.7 km x 1.4 km is not possible for most deployments because it is quite expensive. The optimal number of clusters and stations depends on the researcher's requirements. One additional station increases the accuracy but also the costs. Therefore, the error emerging if not the full number of stations is used, is investigated. The following experiment is analysed: It is expected that the time series at all 13 sites are of interest, but direct measurements at certain sites are replaced by time series from another station. Two questions will be addressed: (1) Which sites are effectively redundant and need to be observed only once? (2) How large is the resulting error compared to the full network? The root mean square error (RMSE) serves as a distance measure between the time series from two stations. To identify redundant stations, the results of the Average Linkage Cluster Analysis are used. The error is expressed in two different ways: Mean RMSE and optimal RMSE. The first is calculated by averaging the RMSE of each station to all other stations within a cluster and describes the error that is made by choosing an arbitrary station from each cluster. For calculating the latter, the station with the lowest RMSE to all other stations is determined for each cluster. It describes the error that is made by choosing the most representative station in each cluster.

Results

The dendrogram of the Average Linkage Cluster Analysis for one-minute intervals for air temperature is presented in Fig. 4.13a. Stations 03 and 05 are most alike and form a cluster at a RMSD of almost 0.4 K. These two stations are merged with another cluster consisting of stations 07, 08 and 09 at a RMSD of 0.62 K. Stations 01 and 02 first form their own cluster (RMSD = 0.42 K), but later they are assigned also to the larger group. Another cluster is built by stations 10, 11, 12 and 17. They have a maximum RMSD of 0.68 K. Stations 13 and 15 form their own cluster, but with a relatively large RMSD of 0.79 K compared to the other clusters.

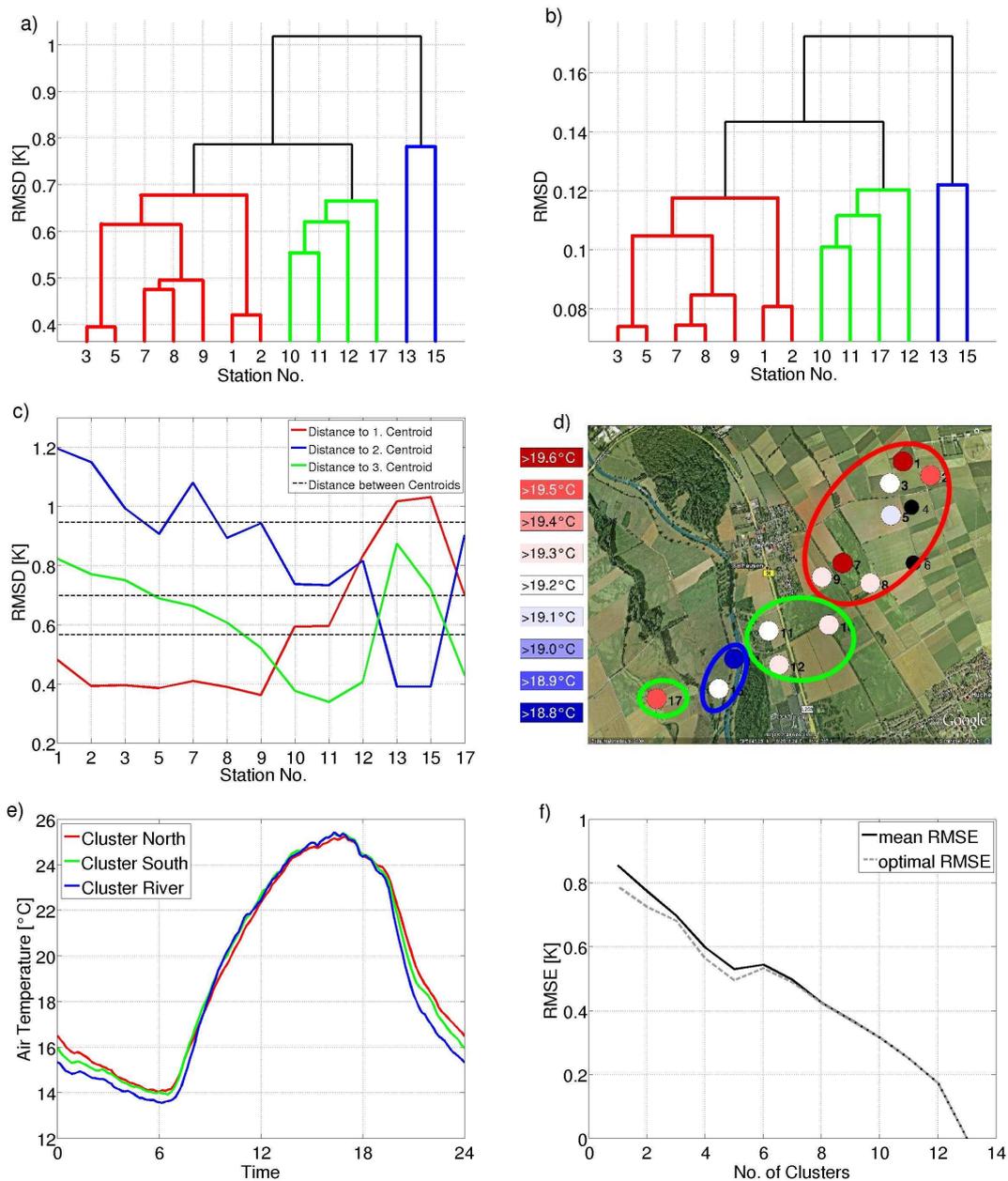


Figure 4.13: Cluster analysis for air temperature: a) Dendrogram for one-minute data, b) dendrogram for normalised data, c) k -means analysis for one-minute data, d) clusters on the map with mean air temperature, e) mean diurnal cycle of clusters with ten minute temporal resolution, f) mean and optimal RMSE for using one to 13 stations

The results of the Average Linkage Cluster Analysis of normalised one-minute intervals in Fig. 4.13b are very similar. The clusters are the same as for one-minute measurements, but the order in which the groups are formed is slightly different. For example, stations 07 and 08 are merged at a lower RMSD value (RMSD = 0.077 K) than stations 01 and

02 (RMSD = 0.08 K). In Fig. 4.13a, it is the other way around. In Fig. 4.13b, station 17 is added first to the group of stations 10 and 11, followed by station 12, instead of reverse in Fig. 4.13a.

According to these two analyses, a division into three groups, consisting of stations 01 to 09, stations 10 to 12 and 17 and stations 13 and 15, respectively, seems appropriate for air temperature. They have a RMSD of at least 0.79 K. That is more than three times the RMSE between the network stations during the calibration period in Chapter 3.1 and, hence, the clusters are significantly different. To prove the results of the Average Linkage Method, k-means Cluster Analysis is performed for three clusters. The result is displayed in Fig. 4.13c. It matches the output of the Average Linkage Analysis perfectly. Stations 01 to 09 form one cluster with a RMSD to their centroid between 0.38 K and 0.5 K. The second cluster consists of stations 10 to 12 and 17. Their RMSD to the second centroid ranges between 0.35 K and 0.42 K. Stations 13 and 15 form the third cluster with a distance to the third centroid of 0.4 K. The distances between the three centroids is more than 0.58 K and is, therefore, larger than the stations-to-cluster-centroid distances. To investigate whether the clustering is stable and not only a product of the short measurement period (20 days), bootstrapping is applied to the dataset. The data is divided into daily datasets. A thousand realisations of 20 randomly chosen days are analysed, whereby days can be used repeatedly. The result (not shown) confirms the cluster analyses of the Average Linkage and the k-means Method.

On the map in Fig. 4.13d, the three clusters are marked: Cluster *North* containing stations 01 to 09 in red, cluster *South* with stations 10, 11, 12 and 17 in green and cluster *River* consisting of stations 13 and 15 in blue. The division seems to be dependent on the geographical position of the stations and their distance to the river. All stations in cluster *North* are more than 0.6 km away from the river, stations in cluster *River* are closer than 0.11 km to the river, stations of cluster *South* are in between.

The mean diurnal cycle in Fig. 4.13e indicates that cluster *River* is distinctly colder during the night than the other two groups. This is because station 15 is located in a small valley and sheltered from the wind and, therefore, cold pool formation is strongest at this site. The area of cluster *North* is most exposed to the wind. Hence, this cluster is coldest during the day and warmest during the night. Cluster *South* is almost identical to cluster *River* during the day, but the nocturnal cooling is not as strong and, therefore, cluster *South* is warmer during the night. Overall, the differences between the clusters are more pronounced during the night.

In Fig. 4.13f, the mean and the optimal RMSE for using one to 13 clusters is shown. The use of only one station leads to a mean RMSE of 0.86 K. It decreases to 0.79 K for the most representative site: Station 09. For the three clusters *North*, *River* and *South*, the mean RMSE is 0.7 K. The station with the lowest RMSE in cluster *North* is station 02; in cluster *South*, it is station 11. Having one station at sites 02, 11 and 13 or 15 reduces the RMSE slightly to 0.69 K. Both error estimators decrease until a number of five clusters. There, the mean RMSE is 0.53 K, the optimal RMSE (using stations 07, 11, 13, 15 and 17) is 0.49 K. For six clusters, the RMSEs increase slightly. Adding more stations leads to a decrease of 0.06 K per station. For 10 clusters, the optimal equals the mean RMSE, because all clusters

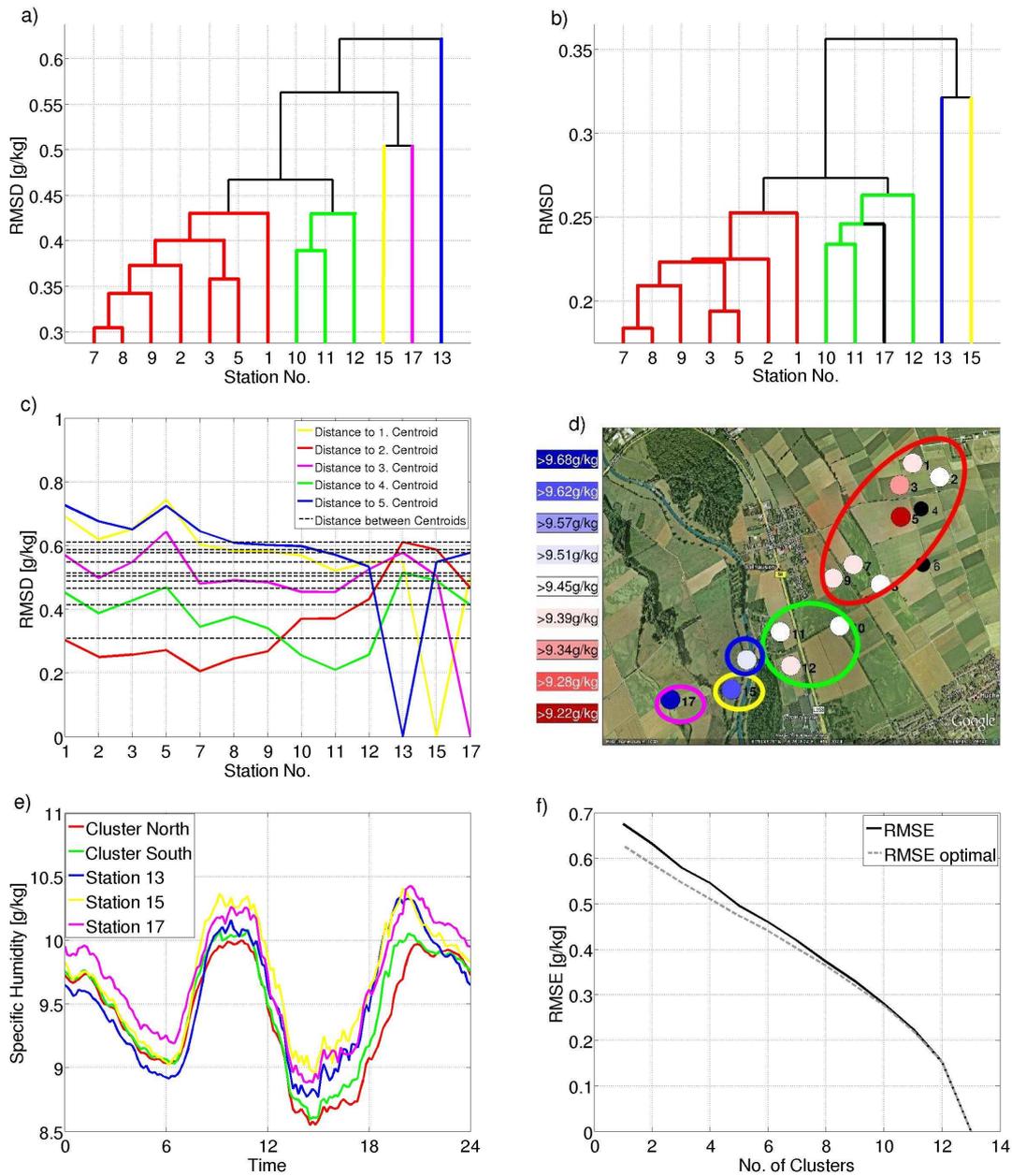


Figure 4.14: Same as Fig. 4.13, but for specific humidity.

consist of maximum two stations.

The dendrogram for one-minute intervals for the Average Linkage Analysis of specific humidity in Fig. 4.14a looks slightly different from the results for air temperature. This time, stations 07 and 08 are most alike with a RMSD of 0.3 g kg^{-1} . Stations 09 and 02 are added to this cluster before it merges with another group containing stations 03 and 05. Station 01 is added to this large cluster at a RMSD value of 0.43 g kg^{-1} . Station 10, 11 and

12 form one group with a maximum RMSD of also 0.43 g kg^{-1} . These two clusters are merged with a RMSD of 0.47 g kg^{-1} . Station 13 stays its own cluster until all stations are merged. Stations 15 and 17 also seem to be independent from the other stations.

The picture is a little different for normalised one-minute intervals (Fig. 4.14b). Indeed, stations 01 to 09 form again one large cluster, but in this case station 17 is added to the group of stations 10 to 12. Stations 13 and 15 constitute again their own clusters. Stations 03 and 05 are merged at an earlier step than for the measured one-minute intervals and station 02 is added to the largest cluster at a later step.

For a division into five groups, both clusterings differ only in the allocation of station 17. The clustering according to Fig. 4.13a for non-normalised values is used in the following, because systematic effects are important and should not be neglected. The five clusters consist of stations 01 to 09, stations 10 to 12, station 13, station 15 and station 17, respectively. The RMSD between these clusters for the Average Linkage Method is at least 0.47 g kg^{-1} . The RMSE between the network sensors during the calibration period, 0.18 g kg^{-1} , is more than two and a half times smaller. Therefore, the clusters are significantly different. The k-means Method for five clusters in Fig. 4.14c gives the same results.

It is obvious from the map in Fig. 4.14d that this is again a distribution dependent on the geographical position of the stations and, therefore, the distance to the river. Cluster *North* includes the most northern stations more than 0.6 km away from the river. The distance between the stations and the river in cluster *South* is between 0.26 km and 0.57 km and the stations are all located east of the river but southwest of cluster *North*. The two stations closest to the river constitute their own clusters. Station 13 is located less than 20 m away from the river on the western side. Station 15 stands about 100 m west of the river. Station 17, which is located more than 0.45 km west of the river, is forming its own cluster, too.

In Fig. 4.14e, mean diurnal cycles for each cluster are presented. The five clusters show clearly different behaviour. Cluster *North*, for example, is up to 0.25 g kg^{-1} drier than all other clusters between 9 a.m. and midnight. During the night, station 13 is driest, although it is located right next to the river. Station 15 on grassland close to the river is wettest during the day with a maximum of 10.4 g kg^{-1} . At night, however, station 17 is wetter than station 15. The pattern is also evident in the reference stations at site 01 and 15. Cluster *South* is close to the average of all clusters throughout the day. In contrast to air temperature, the difference between the clusters is almost constant the whole day.

Using only one instead of all 13 stations leads to a mean RMSE of 0.67 g kg^{-1} and an optimal RMSE for station 07 of 0.63 g kg^{-1} (Fig. 4.14f). Adding more stations reduces the mean RMSE by 0.05 g kg^{-1} and the optimal RMSE by 0.04 g kg^{-1} per station. The mean RMSE for the five clusters *North*, *South*, 13, 15 and 17 is 0.5 g kg^{-1} . For the optimal RMSE, stations 07 and 11 are selected as most representative for clusters *North* and *South*, respectively. The optimal RMSE for a deployment with only stations 07, 11, 13, 15 and 17 is 0.47 g kg^{-1} .

The results of the Average Linkage Analysis for surface temperature in Fig. 4.15a give a completely different clustering than for air temperature and specific humidity. The only

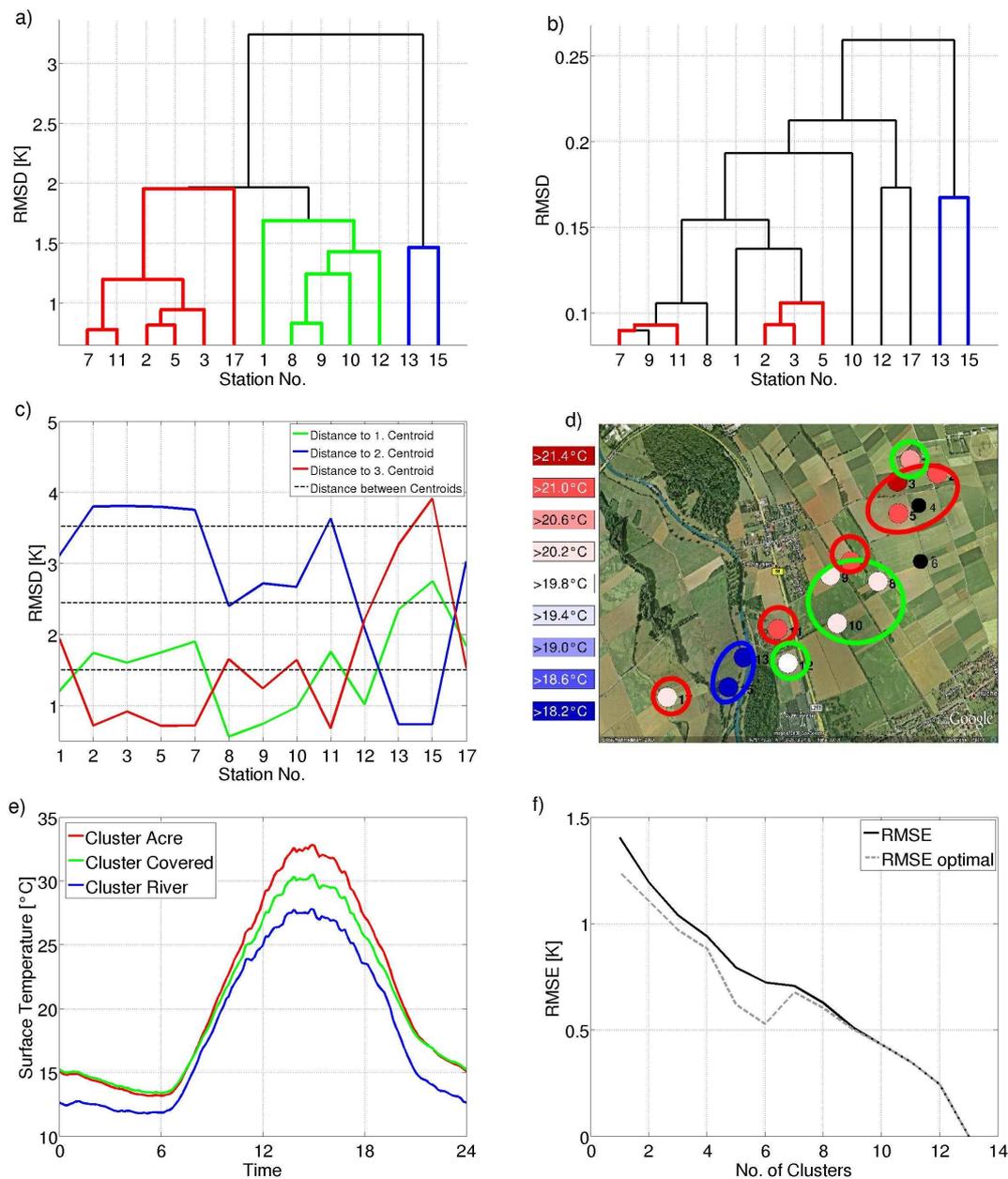


Figure 4.15: Same as Fig. 4.13, but for surface temperature.

thing all three have in common, is the separation of stations 13 and 15 from all other stations. They merge at a RMSD of 1.5 K. In general, the RMSDs for surface temperature are two to three times the RMSDs for air temperature in Fig. 4.13. For surface temperature, stations 07 and 11 are most alike. They form a cluster at a RMSD of 0.7 K. Stations 02 and 05 are also similar and build a cluster at a RMSD of 0.75 K. Station 03 is added and this cluster is merged with the former mentioned one at RMSD = 1.2 K. At last, station 17 is assign to this cluster at a RMSD level of almost 2 K. A third group consists of stations 01,

08, 09, 10 and 12. Their maximum RMSD is 1.7 K. Stations 08 and 09 are most alike in this cluster. This clustering is based on surface type. Stations 02, 03, 05, 07 and 11 are located on harvested fields, where the bare soil is visible. Only at station 17, which is assigned last to this cluster, some vegetation covers the surface partly. At the sites of stations 08, 09, 10 and 12 of the second group, the field is also harvested but the surface is covered with corn or other vegetation. Therefore, the soil is not visible. Station 01 is last assigned to this group. On this site, there is no vegetation, but the area is not agriculturally used and the soil is of lighter colour than the sites of the first cluster. Stations 13 and 15 are located on grassland and also build their own group.

The distribution for normalised values of surface temperature in Fig. 4.15b gives completely other clusters. Most alike are stations 07, 09 and 11 with a maximum RMSD of 0.09 K. Station 08 is also assigned to this cluster. The neighbouring stations 02, 03 and 05 constitute one cluster together with station 01. These groups are merged into one cluster at an RMSD of 0.15 K. Station 10 is also assigned to this large group but not until a RMSD of almost 0.2 K. The only commonality with the clustering of the measured one minute intervals is that stations 13 and 15 again form a separate cluster (RMSD = 0.165 K). Station 12 and 17 build a third group at a similar RMSD level. This distribution seems to be dependent on the distance to the river and, therefore, the geographical position. However, for surface temperature, the systematic differences induced by surface types at the sites are important. Hence, the clustering of the non-normalised measurements are used hereafter.

The three groups chosen based on the analysis in Fig. 4.15a contain stations 01, 08, 09, 10 and 12, stations 02, 03, 05, 07, 11 and 17 and stations 13 and 15. The minimum RMSD between the groups is almost 2 K. That is five times the difference between the network sensors during the calibration period. In Fig. 4.15c, the clustering for one-minute intervals of measured surface temperature according to the k-means Method is displayed. It matches the Average Linkage Analysis perfectly.

The map in Fig 4.15d indicates that the clustering is dependent on the mean value of surface temperature. All stations from cluster *Acre* (stations 02, 03, 05, 07, 11 and 17) have mean surface temperatures larger than 20.2°C, because the dark soils at these sites are efficient heat absorbers. Stations 13 and 15 in cluster *River* are located on grasslands. The soil is completely covered with grass leading to mean surface temperatures lower than 18.6°C. Cluster *Covered* contains stations 01, 08, 09, 10 and 12 with mean surface temperatures between 19.8 and 20.2 °C.

Fig. 4.15e shows that the 10 minute mean surface temperature in cluster *River* is up to 3 K lower than for the other clusters the whole day. Cluster *Acre* and cluster *Covered* have nearly the same surface temperatures at night. During the day, the surface at the stations in cluster *Acre* heats up more because of its darker colour.

The mean RMSE for using only one station is 1.4 K, the RMSE for using the optimal station 09 is reduced to 1.24 K (see Fig. 4.15f). The mean RMSE decreases to 1.04 K, if one arbitrary station from each of the three clusters mentioned above is chosen. The most representative stations are station 05 for cluster *Acre* and station 08 for cluster *Covered*. For a setup with these two stations and station 13 or 15 the optimal RMSE is 0.97 K. Adding

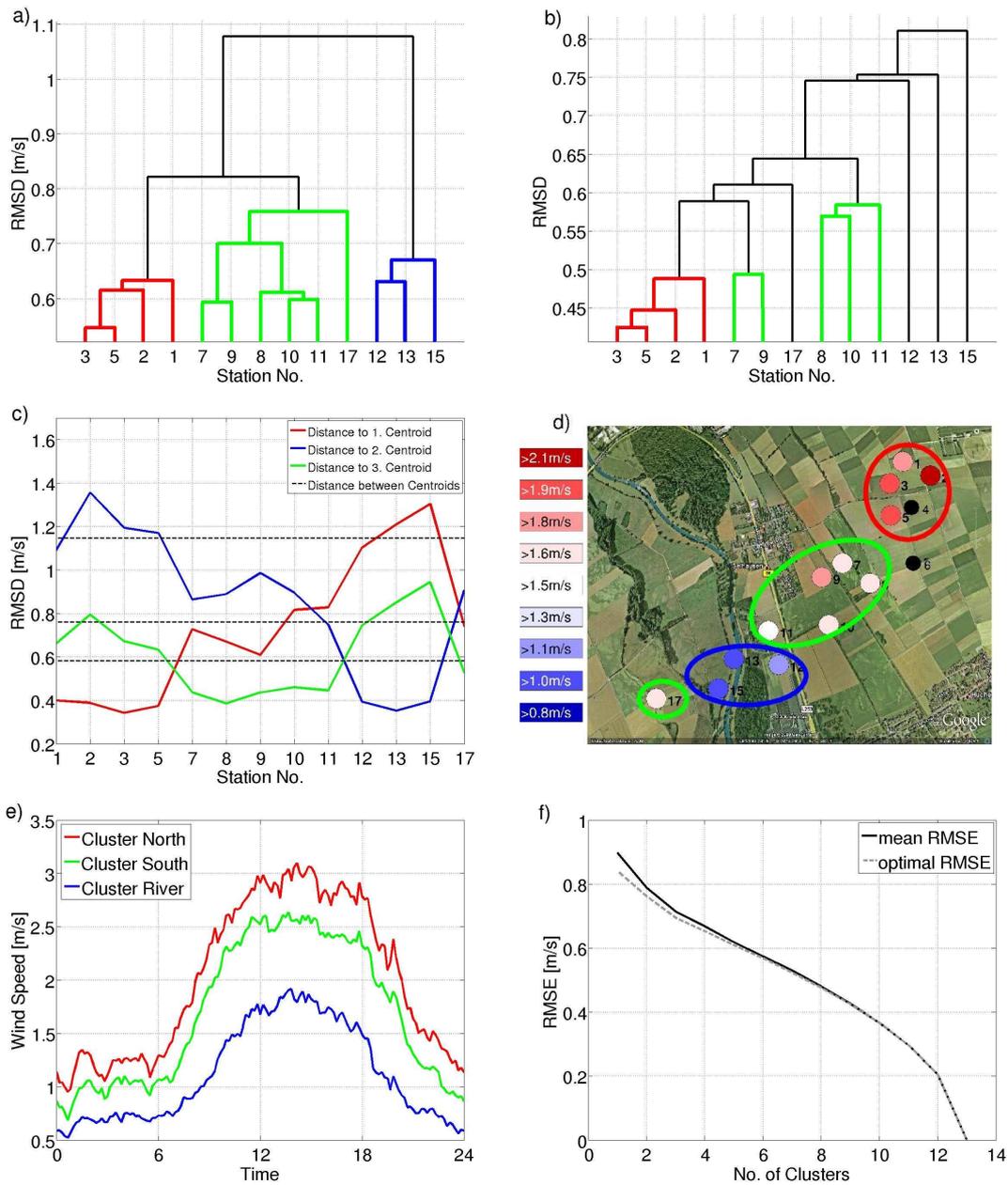


Figure 4.16: Same as Fig. 4.13, but for wind speed

more stations reduces the RMSE by 0.08 K per station.

In Fig. 4.16a, the hierarchical cluster analysis for one-minute intervals of wind speed is shown. Stations 03 and 05 are most alike with a RMSD of 0.54 m s^{-1} . They form a cluster together with stations 01 and 02 at a maximum RMSD of 0.62 m s^{-1} . A second cluster contains stations 07 to 11. These stations are merged at a maximum RMSD of 0.7 m s^{-1} . At RMSD = 0.76 m s^{-1} , station 17 is added to this group. Later, these groups

are merged into one large cluster. Stations 12, 13 and 15 build a third cluster at a RMSD value of 0.68 m s^{-1} . This clustering seems to be dependent on the location of the stations. Neighbouring sites are merged into clusters.

For normalised one-minute intervals in Fig. 4.16b, stations 01 to 05, stations 07 and 09 and stations 08, 10 and 11 form a cluster, respectively. But in contrast to measured one-minute means, the cluster consisting of stations 07 and 09 is added to stations 01 to 05 first, followed by station 17. Then this cluster is merged with stations 08, 10 and 11. Stations 12, 13 and 15 form their own cluster each and are added to the large cluster one after another.

The clustering in Fig. 4.16a seems to be more appropriate. The minimum RMSD value for three clusters is 0.81 m s^{-1} . The RMSE between the network sensors during the calibration period is four times smaller with 0.17 m s^{-1} . Therefore, the k-means Method is also applied to measured one-minute intervals. In Fig. 4.16c the results are displayed. This method gives exactly the same results.

On the map in Fig. 4.16d, the three clusters are marked in different colours. It is obvious that the groups are dependent on the strength of the wind speed. Cluster *North* consisting of stations 01 to 05 observes mean wind speeds more than 1.8 m s^{-1} because all stations are located on sites where no hedge banks or other obstacles block the wind. Stations 07 to 11 and 17 of cluster *South* stand on sites that are sheltered from the wind at least at one site. Therefore, the mean wind speeds are lower. They range between 1.5 m s^{-1} and 1.8 m s^{-1} . Cluster *River* (stations 12, 13 and 15) is well sheltered from the wind. All sites are surrounded by hedge banks or buildings and the mean wind speed is lower than 1.3 m s^{-1} . The dependency of the clustering on the location is coincidental because all stations in the north of the measurement site are less sheltered from wind than the stations in the south.

This is also demonstrated by the mean diurnal cycle of the clusters in Fig. 4.16e. Throughout the day, wind speeds are highest at cluster *North* with a maximum of 3 m s^{-1} at noon. Wind speeds are lowest at cluster *River* with a minimum of 0.5 m s^{-1} at midnight.

Using only one instead of 13 stations leads to a mean RMSE of 0.9 m s^{-1} and an optimal RMSE of 0.84 m s^{-1} for station 08 (see Fig. 4.16f). For choosing an arbitrary station out of the three clusters *North*, *River* and *South*, the RMSE decreases to 0.71 m s^{-1} . Station 03, 08 and 13 represent their cluster best. Using these three stations results in an optimal RMSE of 0.69 m s^{-1} . Any further station reduces the RMSEs by 0.05 m s^{-1} .

4.5 Comparison with FOOT3DK Model

The FOOT3DK model (Flow Over Orographically Structured Terrain - 3 Dimensional (Kölner Version)) is a mesoscale, prognostic model to simulate temporal development fields of meteorological parameters over complex terrain. It has been constructed at the University of Bonn and later developed simultaneously at the University of Cologne. It has a vertical resolution of 10 m near the ground and 500 m at the top of the model and a horizontal resolution of up to 100 m. See Brücher et al. (2003) for a more detailed

description of the FOOT3DK model.

This model is based on the equations of motion. The horizontal part of the Coriolis parameter is neglected, the vertical part $f = 2\omega\sin\varphi$ with angular velocity ω and latitude φ is assumed to be constant:

$$d_t u = -c_p \cdot \theta_v \cdot \partial_x \pi + f \cdot v, \quad (4.3)$$

$$d_t v = -c_p \cdot \theta_v \cdot \partial_y \pi - f \cdot u, \quad (4.4)$$

$$d_t w = -c_p \cdot \theta_v \cdot \partial_z \pi - g, \quad (4.5)$$

where θ_v is potential virtual temperature, $c_p = 1004 \text{ J kg}^{-1} \text{ K}^{-1}$ is specific heat capacity and g is gravitational constant.

The Exner pressure π is defined as

$$\pi = \left(\frac{p}{p_r} \right)^{\frac{R_L}{c_p}} \quad (4.6)$$

with reference pressure $p_r = 1000 \text{ hPa}$ and $R_L = 287 \text{ J kg}^{-1} \text{ K}^{-1}$ being the gas constant for dry air.

Other fundamental equations are the first law of thermodynamics (4.7), the equation of continuity (4.8) and budget equations for specific humidity q (4.9) and a passive tracer χ (4.10):

$$d_t \theta = Q_\theta, \quad (4.7)$$

$$d_t \rho = \rho \nabla \cdot \mathbf{v}, \quad (4.8)$$

$$d_t q_j = Q_{q_j}, \quad (4.9)$$

$$D_t \chi = -\chi \nabla \cdot \mathbf{v} + Q_\chi. \quad (4.10)$$

All terms named Q describe source-, sink- and transformation-processes.

The variables in equations (4.3) to (4.10) are divided into a scale and a sub-scale part by Reynolds decomposition. The variations of the scale part can be resolved temporally and spatially by the model. The turbulent part of the pressure gradient is neglected and the pressure in the vertical equation of motion (4.5) is subdivided into a hydrostatic and a non-hydrostatic part.

As a coordinate system, an η system is chosen that follows the orography. The transformation into the η system can be found in Pielke (1984). An Arakawa C grid is used for discretisation of the variables in the model. Vertical, turbulent fluxes at the surface are calculated following the Monin-Obukhov-Similarity-Theory (Louis, 1979), horizontal

diffusion is formulated according to Pielke (1974), fluxes in higher levels are computed with a 1.5 order closure (Mellor and Yamada, 1982).

Incoming shortwave radiation is computed based on the solar constant and the transmissivity, i.e. dependent on humidity in the atmosphere and the clouds. For calculating upward longwave radiation, the Stefan-Boltzmann-Law (Equation 2.1) is used. For computing downward longwave radiation, the influence of clouds are taken into account according to Swinbank (1963).

In the modelling area, eight different kinds of land use occur: water, forest, grassland, crop, canola, beet, settlement and open-cast mining. Each land use is determined by nine physical parameters that will be discussed later in this chapter. A look-up table can be found in Table 4.2. To account for the different land uses a tile approach is used (Avisar and Pielke, 1989). That means each land use is considered according to its percentage in the grid box.

Table 4.2: Physical parameters for the eight land uses (water, forest, grassland, corn, canola, beet, settlement and mining) in FOOT3DK (summer): Roughness length (z_0), absolute heat capacity (c_H), conductivity (COND), albedo (ALB), sealing (SEAL), vegetation coverage (VEG), leaf area index (LAI), minimum stomatal resistance (MSR) and radiance coefficient (RC).

Param.	Water	Forest	Grassl.	Corn	Canola	Beet	Settl.	Mining
z_0 [m]	0.001	0.8	0.05	0.2	0.1	0.06	0.6	0.02
c_H [$\text{MJ m}^{-3}\text{K}^{-1}$]	4.2	2.5	2.2	2.2	2.2	2.2	2	2.2
COND [W mK^{-1}]	100	1.8	1.3	1.3	1.3	1.3	2	1.3
ALB	0.1	0.125	0.2	0.2	0.18	0.18	0.15	0.2
SEAL	0	0	0	0	0	0	0.5	0.1
VEG	0	0.98	0.9	0.1	0.1	0.92	0.35	0
LAI	0	4.5	2	0.1	0.1	3.6	2.3	0
MSR [s m^{-1}]	0	100	40	40	40	40	60	40
RC [W m^{-2}]	0	100	40	30	30	30	50	30

In the FOOT3DK model, soil heat flux is calculated as a residual of the energy balance equation with an extended force restore model (Jacobsen and Heise, 1982). Soil moisture can be computed prognostically with a force restore model or linked to the turbulent flux of sensible heat with the help of a Bowen ratio. A simple scheme is used to describe humidity, which divides liquid water content into cloud water and rainwater (Colton, 1976). As initial and boundary conditions, analyses of the model COSMO-DE (CONsortium for Small Scale MOdelling) at a 2.8 km grid resolution are used. A three-step “one-way” nesting technique increases the resolution up to 100 m.

In this thesis, simulations of the FOOT3DK model for August 18th, 2009 between 5 a.m. and midnight are used in a spatial resolution of 100 m and a temporal resolution of 6 min. The simulations are made at the University of Cologne for the FLUXPAT2009 campaign. To investigate whether the network is appropriate for model validations, a

comparison with temperature and humidity measurements of the sensor network at a height of 1.5 m, the FOOT3DK-simulations of potential temperature and specific humidity at level 01 close to the surface (at a height of approximately 1 m) and at level 02 at a height of 10 m is chosen. Potential temperature θ is converted into air temperature T_{Air} by using the following equation:

$$\theta = T_{Air} \cdot \left(\frac{p}{p_r}\right)^{\frac{R_L}{c_p}}. \quad (4.11)$$

Station 06 was not operational on August 18th, 2009. Therefore, a comparison to the observation at this station is not possible. Each network station is located in a different grid box. The percentage of each land use is determined for each grid box and listed in Table 4.3. The three dominating land uses for each box are highlighted in yellow.

Table 4.3: Percentages of land use and the resulting land use class for each station. The three dominating land uses for each box are highlighted in yellow.

Station	Water	Forest	Grassl.	Crop	Canola	Beet	Settl.	Mining	LUC
01	0	0	0	40	0	1	0	59	I
02	0	0	10	17	0	53	20	0	II
03	0	0	0	15	0	56	29	0	II
04	0	0	0	26	0	61	13	0	II
05	0	0	0	96	0	0	4	0	III
06	0	0	0	12	0	56	32	0	II
07	0	0	0	90	0	0	10	0	III
08	0	6	0	2	70	1	21	0	IV
09	0	0	0	80	0	20	0	0	III
10	0	5	0	61	0	1	33	0	III
11	0	29	34	19	0	0	18	0	V
12	0	0	0	0	81	0	19	0	IV
13	17	5	70	0	0	0	8	0	VI
15	0	25	71	0	0	0	4	0	VI
17	0	18	0	4	0	67	11	0	II

Grid boxes, which include network stations, can be divided into six different land use classes (LUC) dependent on their dominating land use types:

LUC I: Open-cast mining, crop and beet at station 01

LUC II: Beet, settlement and crop at stations 02, 03, 04, 06 and 17

LUC III: Crop, settlement and forest at stations 05, 07, 09 and 10

LUC IV: Canola, settlement and forest at stations 08 and 12

LUC V: Grassland, forest and crop at station 11

LUC VI: Grassland, forest and water at stations 13 and 15

An Average Linkage cluster analysis of the land use for each station is presented in Fig. 4.17. The cluster analysis is applied to the percentages of the eight land uses in each grid box (see Table 4.3). It results in the same six land use classes: Station 01 (yellow) and 11 (pink) form their own clusters. The biggest cluster (red) includes stations 02, 03, 04, 06 and 17. These stations are located in grid boxes, that are dominated by beet, settlement and crop. Station 05, 07, 09 and 10 build another group (green). These grid boxes mostly consist of crop and settlement. The boxes, where stations 08 and 12 are located in, are dominated by canola. Therefore, these two stations also form one cluster (blue). The sixth cluster (cyan) comprises station 13 and 15 located in grid boxes dominated by grassland.

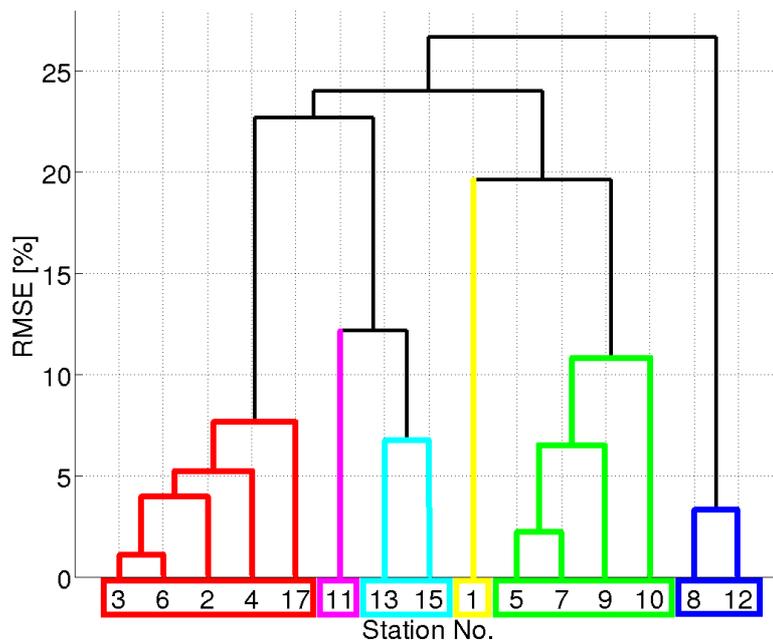


Figure 4.17: Cluster analysis of the land use in the FOOT3DK according to the Average Linkage Method.

In Fig. 4.18, a comparison between air temperature simulated by the FOOT3DK and measured by the sensor network for August 18th, 2009, between 5 a.m. and midnight is presented. Plots for each station are framed in different colours, according to their land use classes. Station 06 was not operational on August 18th. The diurnal cycle looks similar for all stations for the measured air temperature (red line) and for the calculations at level 02 (green line). Although air temperatures at level 02 are computed for a height of 10 m and measurements are taken at a height of 1.5 m, they are in good agreement. The diurnal cycle in the data at level 02 is less pronounced than in the measurements, because the

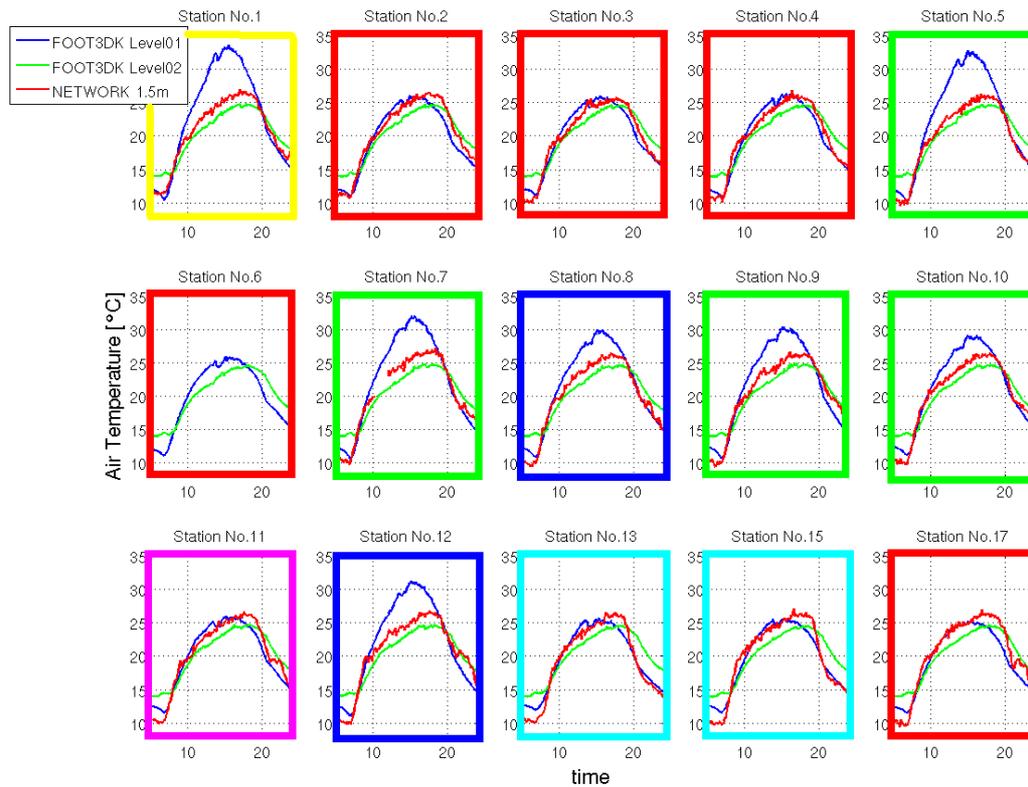


Figure 4.18: Comparison of air temperature simulated by the FOOT3DK model and measured by the sensor network for August 18th 2009. Station 06 was not operational on this date. The blue line indicates simulations at model level 01, the green line simulation at model level 02 and the red line network measurements. The plots are framed according to their land use classes: LUC I in yellow, LUC II in red, LUC III in green, LUC IV in blue, LUC V in pink and LUC VI in cyan.

atmosphere at a height of 10 m is more mixed. That leads to higher temperatures during the night and lower temperatures during the day. The mean difference between simulations and observations ranges from -0.37 K at station 01 to 0.32 K at station 05. The RMSE of the differences lies between 1.66 K for station 01 and 2.72 K for station 15. Air temperatures in model level 01, however, show large variability between the stations. During the night, the model calculations fit the measurements quite well. Before 8 a.m. the model overestimates the temperature for all stations in the order of 1 K to 2 K. After 8 p.m., the model underestimates the measurements by up to 3.5 K at station 12. These features are common for all grid boxes. In contrast, during the day, calculations at level 01 do not show the same behaviour in every grid box. For stations 02, 03, 04, 11, 13, 15 and 17, the FOOT3DK data at level 01 matches the network measurements well. Air temperatures at stations 01, 05, 07, 08, 09, 10 and 12 are clearly overestimated by the FOOT3DK model.

At station 01, the calculated temperature at level 01 is up to 8.27 K higher than the measured temperature. The mean difference between model and network is between -0.74 K at station 17 and 2.71 K at station 01. The RMSE is lowest for station 04 with 1.13 K and highest for station 01 with 3.94 K. Apparently, two groups have been formed based on the land use in the grid boxes: Stations at which measured and simulated air temperatures fit well are all located in grid boxes with land use classes II, V and VI, dominated by crop and grassland. In contrast, stations at which the model overestimates the measurements during the day are all located in grid boxes with land use classes I, III and IV, dominated by canola, beet, settlement and mining. This distinct influence of the land use classes on modeled air temperature at level 01 vanishes at level 02.

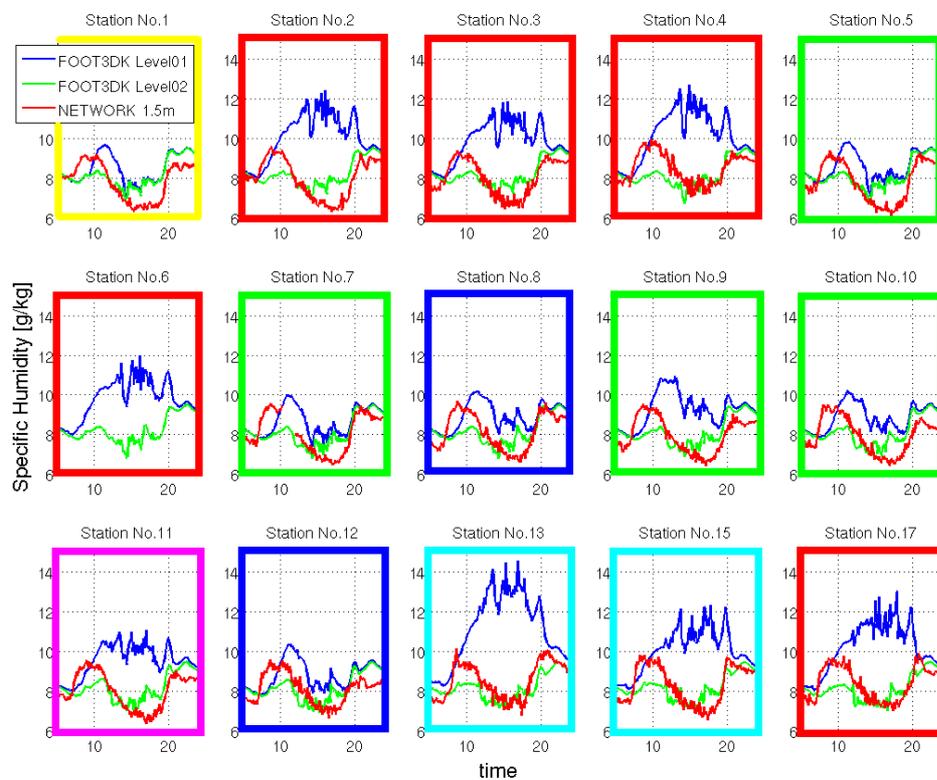


Figure 4.19: Same as Fig. 4.18, but for specific humidity.

The comparison of specific humidity between measurements and model for all stations is shown in Fig. 4.19. The measurements of all network stations (red line) are similar with specific humidities between 6 g kg^{-1} and 10 g kg^{-1} and two maxima around 10 a.m. and around 8 p.m.. Again, for half of the stations (station 01, 05, 07, 08, 09, 10 and 12), the specific humidity at level 01 (blue line) is similar to the measured humidity. In contrast to air temperature in Fig. 4.18, this time the stations are all in grid boxes with land use classes I, III and IV. For these grid boxes, specific humidity at level 01 is of the same magnitude and

has a similar diurnal cycle, although the increase in humidity in the morning starts two hours later in the model simulations than it does in the measurements. In the evening, the diurnal cycle is in phase again. In grid boxes with land use classes II, V and VI containing stations 02, 03, 04, 11, 13, 15 and 17, the simulated specific humidity of level 01 shows a completely different behaviour. The diurnal cycle does not have two maxima in the morning and in the evening, but increases from midnight to midday, fluctuates in the afternoon and decreases again in the evening. This leads to differences between measured and modeled specific humidity of up to 7.23 g kg^{-1} at station 13. That is twice as large as the measured humidity. The model overestimates the measurements for all stations. The mean difference between simulations and observations ranges from 0.49 g kg^{-1} at station 07 to 2.61 g kg^{-1} at station 13. The RMSE amounts at least to 0.98 g kg^{-1} for station 07 with a maximum of 3.37 g kg^{-1} at station 13. At level 02 (green line), the humidity is similar in every grid box. As expected, the diurnal cycle is less pronounced, because the atmosphere is better mixed at a height of 10 m. This leads to distinctly lower mean differences compared to level 01 between -0.33 g kg^{-1} (station 13) and 0.33 g kg^{-1} (station 05). The RMSE is also lower with values between 0.69 g kg^{-1} for station 13 and 0.85 g kg^{-1} for station 10.

The comparison of modeled and measured air temperature and specific humidity indicates that the FOOT3DK model is only able to simulate one of these variables, either air temperature or specific humidity, satisfactorily at level 01 close to the surface. The other variable is overestimated by the model during the day. At the second level at a height of 10 m, the simulations fit the measurements better although the height difference between model and measurements is 8.5 m. That implies an overestimation of available energy in the lowest level of the model. One possible explanation is that incoming solar radiation at the top of the model is too high. Another possibility is an erroneous calculation of the heat storage in the soil. If the soil gives off too much heat, the atmosphere especially close to the surface gets too warm. Another possible source of error is the soil moisture, which is almost constant for the model run used. Also, temperature and humidity might accumulate at the top of the model area leading to a too warm and moist atmosphere.

The quality of temperature and humidity simulations seems to be dependent on the land use classes in the grid boxes. In grid boxes with land use classes I, III and IV, the model simulates specific humidity satisfactorily but overestimates air temperature. In grid boxes with land use classes II, V and VI, it is the other way around. To find reasons for this behaviour, a closer look into the land use classes and their characteristics is necessary.

The eight land uses differ in nine physical parameters. For each grid box, these parameters are defined depending on the percentage of each land use:

- *Roughness length* (z_0): The roughness length for each grid box is computed using a weighing function based on the land-use-roughness-length z_{0_i} with proportion l_i of the respective land use:

$$z_0 = \exp\left(\sum_{i=1}^{N_{class}} l_i \cdot \ln(z_{0_i})\right) \quad (4.12)$$

- *Absolute heat capacity* (c_H): The heat capacity indicates how much thermal energy can be stored.
- *Conductivity* (COND): For water, effects of turbulent mixing and transparency of water must be considered.
- *Albedo* (ALB): The Albedo is the reflectivity potential of a surface type for shortwave radiation.
- *Sealing* (SEAL): The average fraction of each land use surface that is sealed.
- *Vegetation Coverage* (VEG): The average fraction of each land use surface that is covered by plants.
- *Leaf area index* (LAI): Ratio of leaf area to vegetation covered area (different for summer and winter).
- *Minimum stomatal resistance* (MSR): The minimum stomatal resistance is a measure of the difficulty for the plants to transpire.
- *Radiance coefficient* (RC): The radiance coefficient defines the plant specific dependency between stomatal resistance and solar radiation.

In Fig. 4.20, the nine physical parameters for each station are presented. Blue circles indicate grid boxes in which modeled specific humidity is compatible with the measurements, but air temperature is overestimated; red circles indicate grid boxes in which modeled air temperature is compatible with the measurements, but specific humidity is overestimated.

The roughness length in Fig. 4.20a shows that grid boxes in which simulations and measurements of specific humidity are in good agreement have roughness lengths larger than 0.15 m. An exception is the grid box containing station 01 with a roughness length of only 0.06 m. In contrast, the red marked stations are all located in grid boxes with roughness lengths smaller than 0.15 m except for station 11 ($z_0 = 0.23$ m). This is due to the vegetation in the respective grid boxes. Land use classes III and IV, containing stations 05, 07, 08, 09, 10 and 12, are dominated by crop and canola. The vegetation in land use classes II and VI, at stations 02, 03, 04, 06, 13, 15 and 17, is dominated by beet and grassland. Therefore, plants are lower and the roughness length is smaller in these two land use classes.

Heat capacity (Fig. 4.20b) and albedo (Fig. 4.20d) seem to have no clear affect on the quality of simulations of air temperature and specific heat. Some tendencies are evident, e.g. all stations with heat capacity greater than $2.2 \text{ MJ m}^{-3}\text{K}^{-1}$ are marked in red and all stations with albedo greater than 0.19 are marked in blue, but that does not apply to all stations. For sealing and conductivity no trend is evident.

Vegetation coverage in Fig. 4.20f, however, shows a distinct separation between grid boxes with good humidity simulations and grid boxes with good temperature simulation. In all grid boxes with vegetation cover more than 0.63 (63%), modeled and measured specific humidities match well and temperature is overestimated. In contrast, in all

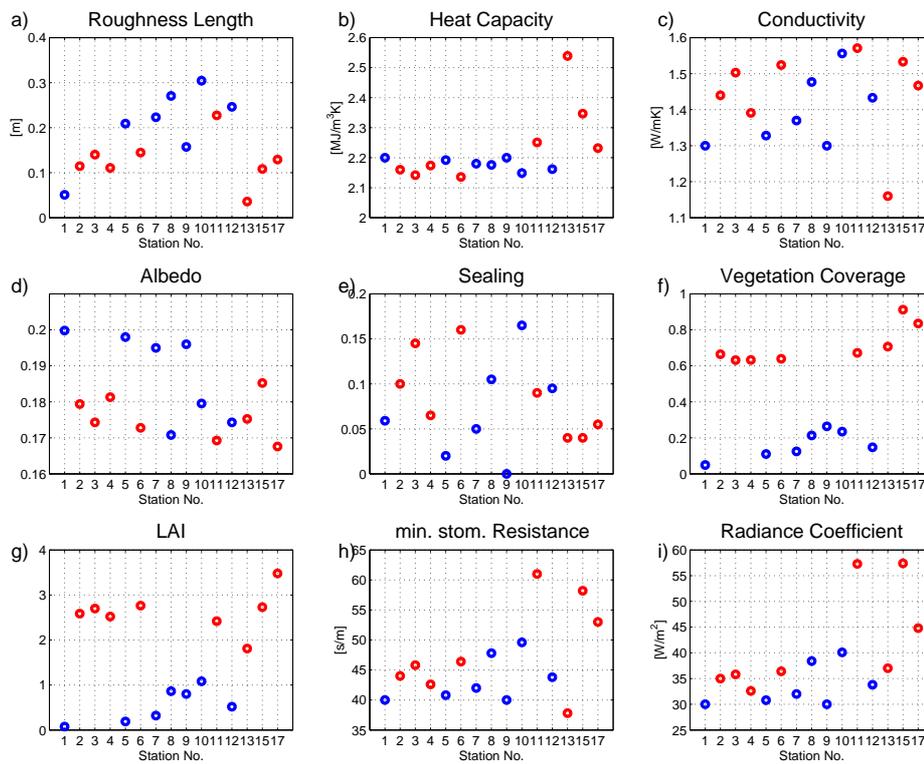


Figure 4.20: Physical parameters for every grid box, that contains a station: a) roughness length, b) heat capacity, c) conductivity, d) albedo, e) sealing, f) vegetation coverage, g) leaf area index, h) minimum stomatal resistance and i) radiance coefficient. Blue circles mark grid boxes, where humidity simulations match the measurements, red circles mark grid boxes, where temperature simulations match the measurements

grid boxes with vegetation cover less than 27%, modeled and measured air temperatures are in good agreement and humidity is overestimated by the FOOT3DK model.

The leaf area index also points out the separation of blue and red grid boxes (Fig. 4.20g). This is linked to the vegetation cover because areas with a high vegetation cover also have a high LAI. Therefore, all stations in grid boxes with land use classes II, V and VI in red have LAI values larger than 1.8 and all stations in grid boxes with land use classes I, III and IV in blue have LAI values smaller than 1.1.

For minimum stomatal resistance and radiance coefficient in Fig. 4.20h and i, no influence of the model performance regarding specific humidity and air temperature at level 01 is evident. The red marked stations 11, 15 and 17 have the highest values of minimum stomatal resistance and radiance coefficient. The remaining red stations do not differ from the blue stations.

In this study the simulations for only one day are considered. Therefore, general con-

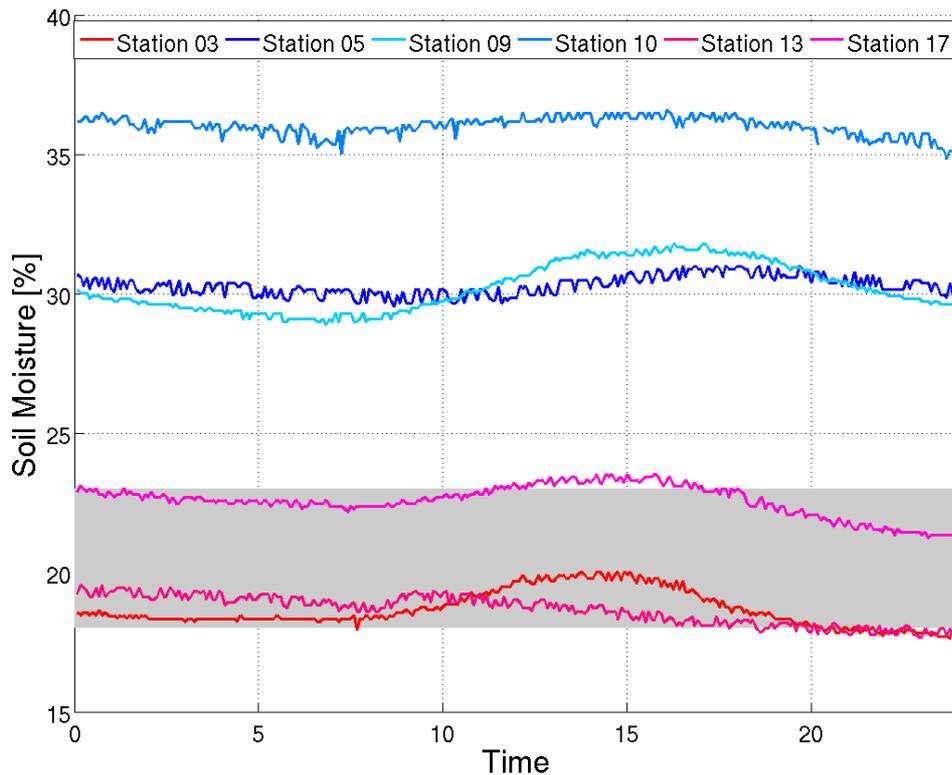


Figure 4.21: Soil moisture for August 18th 2009 at network stations 03, 05, 09, 10, 13 and 17. The grey area indicates possible soil moisture values in the FOOT3DK model.

clusions for the FOOT3DK model can only be drawn to a limited extent. However, the sensor network is an appropriate tool to investigate whether the model is able to simulate small scale variabilities realistically. It is shown that the FOOT3DK model tends to overestimate the variabilities in air temperature and specific humidity at the lowest level. Vegetation cover seems to have the largest influence on the performance of the FOOT3DK model. Vegetation cover also influences LAI, roughness length and albedo. Therefore, in these three parameters a more or less distinct separation is evident between grid boxes, in which specific humidity is modeled well, and grid boxes, in which air temperature is modeled well.

Another possible factor is soil moisture. In the FOOT3DK simulation, it is set to 23% for overgrown areas and 18% for unvegetated areas. That implies the soil moisture for each grid cell is at least 18% and not more than 23%. At the network stations 03, 05, 09, 10, 13 and 17, soil moisture is measured at a depth of 3 to 6 cm. The measurements for August 18th, 2009, are presented in Fig. 4.21. Stations 03, 13 and 17 (red lines) are located in grid boxes, where air temperature is modeled well, stations 05, 09 and 10 (blue lines) in grid boxes, where specific humidity is modeled well. The grey area in Fig. 4.21 indicates pos-

sible soil moisture values in the FOOT3DK model. All red stations have soil moistures within the grey area. At blue stations, the soil moisture is up to two times higher. Therefore, a correct assumption of soil moisture leads to good temperature simulations, but overestimation of specific humidity. In grid boxes, in which the soil moisture is underestimated, specific humidity is simulated better, but air temperature is overestimated.

4.6 Comparison with TERRA Standalone Model

The classical approach of having only one or a few stations to observe a relatively large area is often used in heat flux modeling. Homogeneous atmospheric forcing and homogeneous surface conditions are assumed as model input. Heterogeneities in atmospheric conditions and vegetation are neglected. Recent models use a mosaic approach (Seth et al., 1994) to simulate heat fluxes. In this approach, the assumption is made that only atmospheric variability can be neglected; variability in the vegetation is considered. Therefore, the models have a subgrid for the land surface in each atmospheric grid box. That brings up the question: How important is heterogeneous forcing compared to heterogeneous vegetation for heat flux modeling?

To study the influence of heterogeneous forcing and vegetation, simulations of the soil-vegetation-atmosphere transfer scheme TERRA are used. The sensor network does not observe sensible or latent heat flux. Therefore, surface temperature is used to validate the model simulation because it is available from both observations and the model. A model study examines the influence of homogeneous and heterogeneous forcing and vegetation on sensible and latent heat flux.

TERRA is the soil-vegetation-atmosphere transfer scheme used in the COSMO (Consortium for Small Scale Modelling) model operated by the German Meteorological Service (DWD). The standalone version consists of the soil module TERRA, a simplified transfer scheme, annual cycles of vegetation parameters and parameterisations of radiation interaction at the surface. A scheme of the model is shown in Fig. 4.22. The model is driven by air temperature T , relative humidity RH , wind speed v , incoming solar radiation F_{SW} , downward longwave radiation LW , air pressure p and precipitation RR measured at a reference level z . The prognostic variables are soil moisture SM , soil temperature ST , surface temperature T_{Surf} and turbulent energy fluxes.

The TERRA model has eight layers with layer depths of 0.01 m, 0.03 m, 0.09 m, 0.27 m, 0.81 m, 2.43 m, 7.29 m and 21.87 m. The lowest layer is the so-called "climate layer" with constant soil temperatures. For soil moisture, only the upper six layers are active layers with a free drainage boundary at 2.43 m. A free drainage boundary allows soil water to drain from the lowest layer, but the soil cannot be moistened from ground water below. The model is strictly one-dimensional. Horizontal transport of water and energy in the soil is neglected. Therefore, soil moisture flux F_{SM} is the sum of vertical diffusion and drainage (Dingman, 2002):

$$F_{SM} = K(SM) + D(SM) \cdot \frac{\partial SM}{\partial z}, \quad (4.13)$$

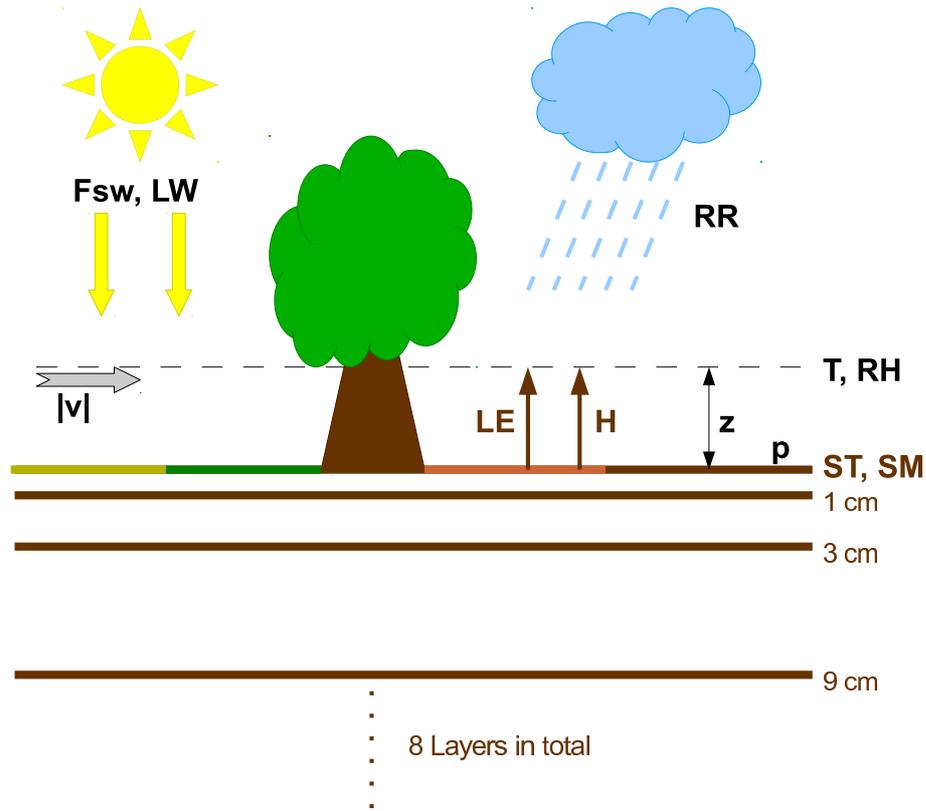


Figure 4.22: Scheme of the TERRA model indicating the forcing variables (air temperature T , relative humidity RH , wind speed v , air pressure p , precipitation RR , incoming shortwave radiation F_{SW} and downward longwave radiation LW) in black and the prognostic variables (soil temperature ST , soil moisture SM , sensible heat flux H and latent heat flux LE) in brown.

where K and D are hydraulic conductivity and diffusivity, respectively. According to Rijtema (1969), K and D are calculated as:

$$K(SM) = K_0 \cdot \exp\left(K_1 \cdot \frac{SM_{pv} - SM}{SM_{pv} - SM_{adp}}\right) \quad (4.14)$$

and

$$D(SM) = D_0 \cdot \exp\left(D_1 \cdot \frac{SM_{pv} - SM}{SM_{pv} - SM_{adp}}\right). \quad (4.15)$$

Both conductivity and diffusivity are dependent on soil type (represented by the coefficients K_0 , K_1 , D_0 , and D_1), pore volume SM_{pv} , soil moisture at the air dryness point SM_{adp} and actual soil moisture SM .

Sensible heat flux H and latent heat flux LE are calculated using a flux-gradient approach depending on the temperature and moisture gradient between the surface and the atmosphere at the reference level z . The atmospheric transfer coefficient is parametrised by a modified Louis scheme (Louis, 1979). Net evapotranspiration depends on bare soil

evaporation parametrised according to Noilhan and Planton (1989), plant transpiration parametrised according to Dickinson (1984), and sublimation of snow and evaporation from the interception store. A detailed description of the model can be found in Ament (2006), Schrodin and Heise (2001), Doms et al. (2005), and Heise et al. (2006).

The model uses eight different kinds of soil: Sand, sandy loam, loam, clay loam, clay, peat, ice and stone. They are distinguished by soil moisture saturation, soil moisture at air dryness point, soil matric potential at saturation, pore size distribution, saturated hydraulic conductivity and diffusivity. The model only considers the dominant soil type in each grid cell.

In this chapter, surface temperature simulated by the TERRA model at each station is compared to the measurements of the sensor network. The model uses one minute measurements from the network stations of air temperature, humidity, wind speed, incoming solar radiation and precipitation as input. Additionally pressure measurements are taken from the reference station close to station 15. Downward longwave radiation measurements are recorded as ten minute means at a station operated by the University of Bonn in the town of Merken, approximately 1 km south east of the network site. One minute means are linearly interpolated from these measurements. The model uses network sensor measurements at stations 03, 05, 09, 10, 12, 13, 15 and 17 as starting values for soil temperature and moisture. The measurements at site 12 are used for stations 01 and 08, and the measurements at site 17 are used for stations 02, 07 and 11, because of the similar environment at these sites. Sandy loam is chosen as soil type for all stations, because this soil type dominates the area.

Four different model runs are made to determine the influence of observations and vegetation on the surface temperature simulation:

- **RUN 1:** Homogeneous forcing and homogeneous vegetation
- **RUN 2:** Homogeneous forcing and heterogeneous vegetation
- **RUN 3:** Heterogeneous forcing and homogeneous vegetation
- **RUN 4:** Heterogeneous forcing and heterogeneous vegetation

For the runs with homogeneous forcing, the mean of the observations at all sites is used for every station. This represents the classical approach of using only one station. For heterogeneous vegetation, the four categories *no vegetation*, *grassland*, *soil covered* and *soil visible* of Chapter 4.3 are employed. A look-up table with leaf area index, plant cover, roughness length, root depth, stomatal resistance and albedo can be found in Table B.2 in Appendix B. In the runs with homogeneous vegetation, the vegetation category *soil covered* is used for all stations.

The mean difference in surface temperature between the four runs and the network stations is shown in Fig. 4.23. For this analysis, only measurements between 10 a.m. and 4 p.m. local time are used, because the land surface exchange and, thus, the surface temperatures are most heterogeneous during this time of the day (see Chapter 4.2). At night, the

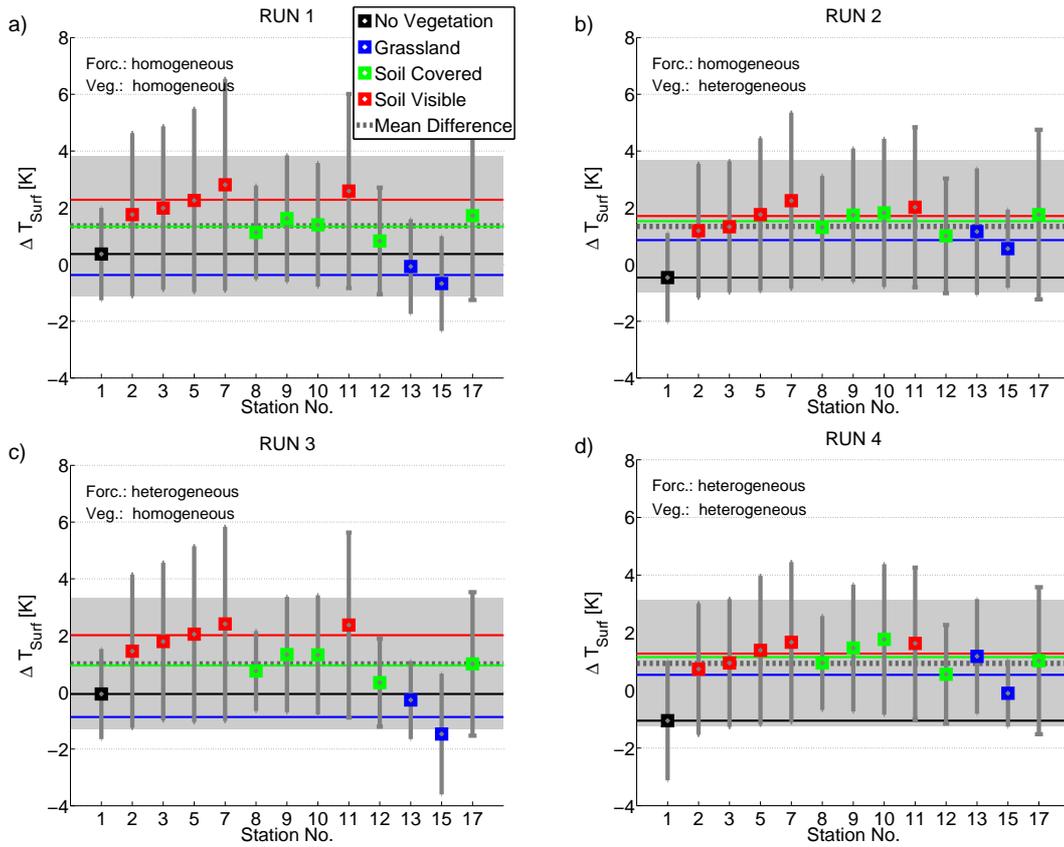


Figure 4.23: Mean differences in surface temperature and their RMSE between the network stations and a) RUN 1, b) RUN 2, c) RUN 3 and d) RUN 4. Only surface temperatures between 10 a.m. and 4 p.m. are taken into account. The errorbars represent the RMSE. The dashed grey line indicates the mean difference between measurements and simulations, the grey area indicates their mean RMSE. Different colours denote different vegetation classes, as shown in the legend.

intra-station differences, and, therefore, the influence of the heterogeneity of the site, almost vanishes. Only stations 13 and 15, close to the river where nocturnal cooling occurs, differ slightly from the other sites.

For RUN 1 in Fig. 4.23a, homogeneous forcing and vegetation are used as input for the TERRA model. The simulation overestimates the measured surface temperature at almost all stations. Only the two stations 13 and 15, located on grassland, have negative mean differences of -0.07 K and -0.67 K, respectively, because their leaf area index, plant cover and roughness length is underestimated by using the parameters of vegetation type *soil covered*. The mean difference at a single station between RUN 1 and the observations amounts to 1.36 K and the mean RMSE amounts to 2.46 K. The four different vegetation classes differ from the observations in a range of -0.37 K for grassland and 2.28 K for visible soil.

For heterogeneous vegetation in Fig. 4.23b, the mean difference between the four vegeta-

tion classes is reduced. It lies between -0.46 K for no vegetation and 1.71 K for visible soil. Also the mean RMSE is reduced to 2.32 K. The results are more uniform for all types of vegetation. However, RUN 2 still overestimates the observations by 1.34 K on average.

In RUN 3 in Fig. 4.23c, the use of heterogeneous forcing with homogeneous vegetation leads to a reduction of the mean bias to 1.00 K. However, the mean differences of the individual vegetation classes are larger than for RUN 2. They range between -0.87 K for grassland and 2.01 K for visible soil. The RMSE is still 2.31 K.

The best results are achieved by RUN 4 with heterogeneous forcing and vegetation in Fig. 4.23d. The mean bias between simulation and observation is reduced to 0.94 K and the RMSE is only 2.19 K. The RMSEs of the individual stations are also reduced compared to RUN 1 except for station 01. Therefore, RUN 4 is considered as nearest to truth in the following.

The simulations of sensible and latent heat flux of RUN 1 to 3 are compared to RUN 4 in Fig. 4.24 in terms of mean difference and RMSE. The influence of heterogeneity in forcing and in vegetation on heat flux simulations is examined. For homogeneous forcing and vegetation, the sensible heat flux in Fig. 4.24a is clearly underestimated compared to heterogeneous forcing and vegetation. This is especially true for stations located on grassland and on visible soil, which have a bias of -12.73 W m^{-2} and -17.62 W m^{-2} , respectively. The mean bias amounts to -9.22 W m^{-2} , the mean RMSE for all stations is 32.77 W m^{-2} . Again for stations on grassland and on visible soil the RMSE is larger than for stations on covered soil, because for homogeneous vegetation *soil covered* is chosen as the vegetation type for all stations.

RUN 2 with heterogeneous vegetation in Fig. 4.24b gives similar results compared to RUN 4 in terms of mean bias and mean RMSE: The sensible heat flux is underestimated by -12.68 W m^{-2} and the variability is quite high with a RMSE of 31.41 W m^{-2} . An improvement is seen for grassland stations. The mean bias of this vegetation type is reduced to -3.96 W m^{-2} and the RMSE is decreased by approximately 10 W m^{-2} .

The comparison of RUN 3 and RUN 4 in Fig. 4.24c indicates that heterogeneous forcing is important to reducing the variability. The mean RMSE decreases to 20.86 W m^{-2} , the mean bias decreases to 3.20 W m^{-2} . For the vegetation types *soil covered* and *soil visible* this improvement is particularly evident. The RMSEs for stations 02 to 12 and 17 are less than half their corresponding RMSEs for RUN 1 and their biases are all in a range between 1.12 W m^{-2} and 6.59 W m^{-2} . Only station 10 shows a larger bias with 13.85 W m^{-2} . It is obvious that for the grassland stations 13 and 15 the consideration of the right vegetation type is more important than the forcing at these stations. Their RMSEs are slightly reduced but they have a mean bias of -12.32 W m^{-2} compared to RUN 4. Station 01 is the only one that provides worse results for RUN 2 and 3 than for the homogeneous RUN 1. The differences between RUN 1 to 3 and RUN 4 in terms of latent heat flux in Fig. 4.24d-f are clearly larger than for sensible heat flux. The mean RMSE between RUN 1 and 4 amounts to 66.47 W m^{-2} (Fig. 4.24d). The mean bias of 15.92 W m^{-2} in latent heat flux is in the same range as for sensible heat flux, but the latent heat flux is overestimated when using homogeneous conditions. The modeled latent heat flux is noticeably too large at

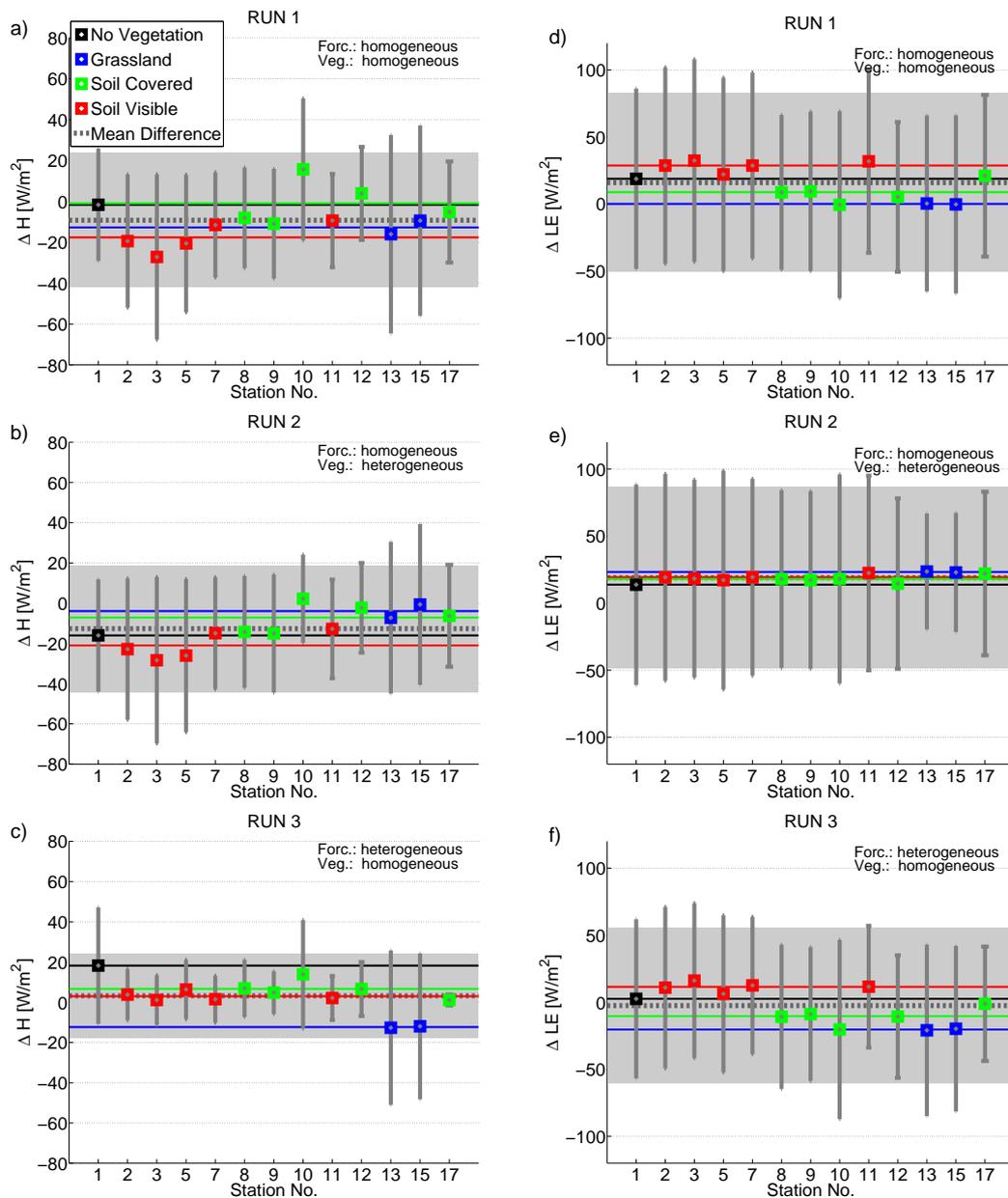


Figure 4.24: Mean differences in sensible heat flux (left column) and latent heat flux (right column) between the RUN 1 and 4 (a and d), between RUN 2 and 4 (b and e) and between RUN 3 and 4 (c and f). Only heat fluxes between 10 a.m. and 4 p.m. are taken into account. The errorbars represent the RMSE. The dashed grey line indicates the mean difference between the simulations, the grey area indicates their mean RMSE. Different colours denote different vegetation classes, as shown in the legend.

sites with *soil visible* surface cover.

This overestimation is also evident in RUN 2 with heterogeneous vegetation. The mean bias amounts to 18.99 W m^{-2} , but all vegetation types provide similar results. The biases

for the four vegetation types are all in a range of less than 10 W m^{-2} . The variability is still high with a mean RMSE of 67.55 W m^{-2} . Only for the two grassland stations is the RMSE distinctly reduced, by about 20 W m^{-2} .

Using heterogeneous forcing in Fig. 4.24f provides lower variabilities for each station. The mean RMSE is reduced to 57.81 W m^{-2} . The mean bias only amounts to -2.50 W m^{-2} , but the difference between the individual vegetation classes is larger than for RUN 2.

It has been shown that using heterogeneous vegetation with homogeneous forcing, as it is suggested by the mosaic approach, leads to more uniform simulations. The mean differences between stations and model are all in a range of 30 W m^{-2} for sensible heat flux and in a range of 10 W m^{-2} for latent heat flux. Using heterogeneous forcing with homogeneous vegetation reduces the RMSE for almost every station in terms of sensible and latent heat flux.

Chapter 5

Hamburg Airport

The Hamburg Airport is chosen as the second measuring site. In contrast to the site for the FLUXPAT2009 campaign, the airport is a homogeneous terrain. It is a wide open area. Influences of different surface types on the meteorological conditions are eliminated at this site, because all stations are located on grassland. The sensor network is used to address the following questions: (1) Is the atmosphere at the Hamburg Airport as homogeneous as the environment indicates? (2) If not, what causes variabilities between different parts of the airport? (3) In the middle of the airport the German Meteorological Service (DWD) operates a climate reference station, but is this location representative of the whole airport?

The deployment and the site are described in Chapter 5.1. In Chapter 5.2, variabilities of the meteorological conditions near the surface are characterised. Explaining factors for these variabilities are examined in Chapter 5.3. According to their environmental influences, stations are divided into clusters and the characteristics of each cluster are studied in Chapter 5.4. The representativeness of the climate reference station operated by the DWD is analysed in Chapter 5.5.

5.1 Deployment

The airport is located in the north of Hamburg in northern Germany. It extends from $53^{\circ} 36' 57''\text{N}$ to $53^{\circ} 39' 28''\text{N}$ and from $9^{\circ} 57' 23''\text{E}$ to $9^{\circ} 59' 23''\text{E}$. Fourteen stations were deployed along both runways as illustrated in Fig. 5.1. The network consisted of three master and eleven slave stations. Stations 01 to 07 were placed along the western side of the north-south runway, stations 09 to 14 along the northern side of the east-west runway and station 08 was placed in the middle of the airport, where both runways cross. Stations 04 and 12 were not operational during the whole campaign and are, therefore, excluded from the analysis. Neighbouring stations were at least 260 m but not more than 710 m apart. All stations were located on grassland. Soil type was either sand or moor. The measurement campaign started on April 24th, 2010. Due to communication problems within the network, the observation period was cancelled on May 20th, 2010. Afterwards, the deployment had to be rearranged and unfortunately it was not possible to keep the

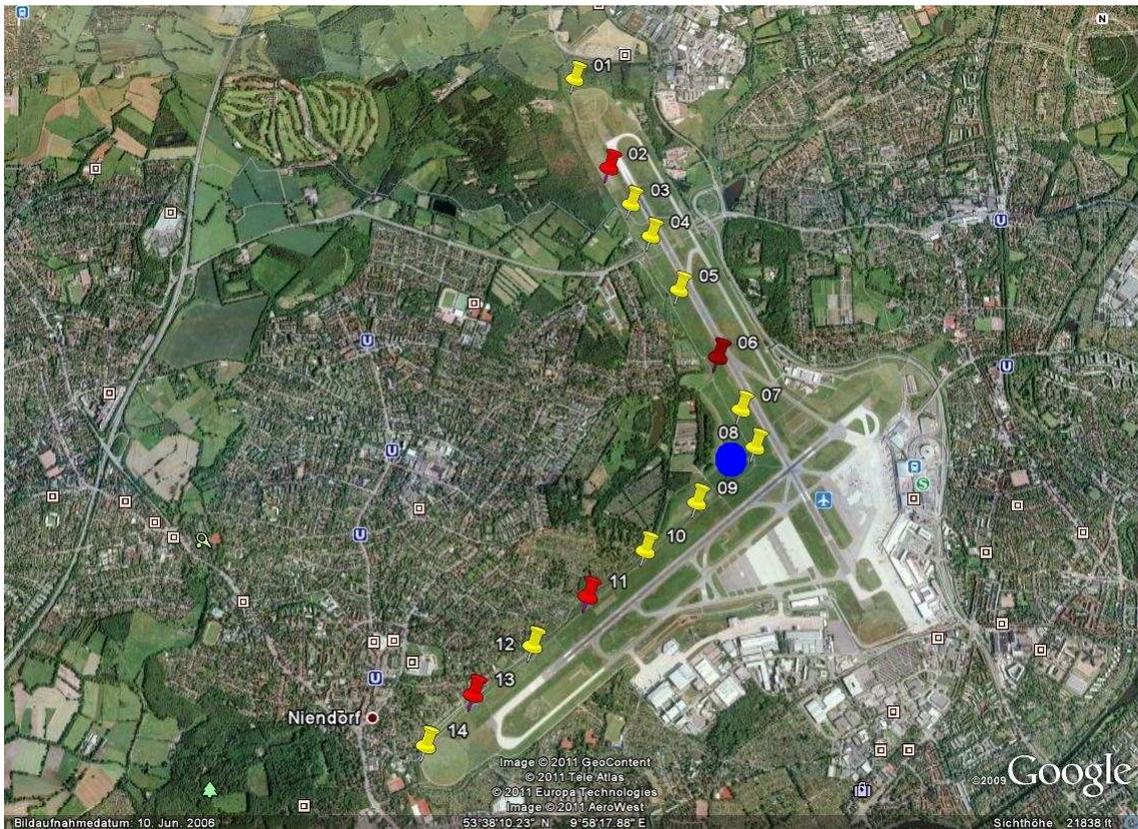


Figure 5.1: Deployment at the Airport Hamburg: Yellow markers represent slave stations, red markers represent master stations, the blue circle represents the reference station of the German Meteorological Service (source: google earth).

full network running over a longer period due to communication and battery failure. More details about technical problems during the campaign can be found in Appendix C. Mean fetches (see Table 5.1) were calculated as in Chapter 4.1 using the frequency of the four wind direction north, south, east and west. The airport is a wide open area compared to the site in the FLUXPAT2009 campaign. Therefore, the mean fetches were larger. They ranged between 84 m at station 01 and 833 m at station 08. Overall, stations at the ends of both runways had smaller fetches than stations closer to the centre of the airport.

Every station was equipped with sensors for measuring wind speed and direction (*WIND*), air temperature and humidity (*TRH*) and surface temperature (*TIR*). The first two were mounted at a height of 2 m, whereby the wind sensor faced north and the temperature and humidity sensor faced west. The latter was mounted at a height of 1.2 m and pointed to the surface in eastward direction. Due to corrosion, humidity measurements were only available at stations 01, 02, 05, 08 and 13. The temperature and humidity sensors at the other stations gave erroneous humidity measurements. Also surface temperature measurements of station 01 were incorrect and were excluded from the analysis. Stations 01, 03, 07, 08, 11 and 14 were additionally equipped with a solar radiation sensor

Table 5.1: Geographical position, land use, distance to the river (*dr*), vegetation, a mean fetch and the attached sensors for each station of the airport deployment. The checkmarks in parentheses indicate that temperature measurements are available, but humidity measurements are erroneous.

<i>No.</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Land Use</i>	<i>Fetch</i>	<i>TRH</i>	<i>TIR</i>	<i>WIND</i>	<i>FSW</i>	<i>RAIN</i>	<i>SM</i>
01	53°39'27.7"N	9°58'15.9"E	grassland	84 m	✓	✓	✓	✓	✓	✓
02	53°39'10.0"N	9°58'25.0"E	grassland	285 m	✓	✓	✓			
03	53°38'56.7"N	9°58'37.6"E	grassland	459 m	(✓)	✓	✓	✓	✓	✓
05	53°38'36.2"N	9°58'55.6"E	grassland	469 m	✓	✓	✓			
06	53°38'20.6"N	9°59'9.6"E	grassland	446 m	(✓)	✓	✓			
07	53°38'8.7"N	9°59'17.8"E	grassland.	609 m	(✓)	✓	✓	✓	✓	✓
08	53°38'0.2"N	9°59'22.6"E	grassland	833 m	✓	✓	✓	✓	✓	✓
09	53°37'48.0"N	9°59'0.6"E	grassland	401 m	(✓)	✓	✓			
10	53°37'37.6"N	9°58'41.5"E	grassland	412 m	(✓)	✓	✓			
11	53°37'27.8"N	9°58'20.6"E	grassland	355 m	(✓)	✓	✓	✓	✓	✓
13	53°37'7.3"N	9°57'40.0"E	grassland	210 m	✓	✓	✓			
14	53°36'57.0"N	9°57'23.7"E	grassland	102 m	(✓)	✓	✓	✓	✓	✓



Figure 5.2: Station 08 in the front and the climate reference station of the German Meteorological Service in the back

(FSW), a rain gauge (RAIN) and a soil moisture and temperature sensor (SM). The solar radiation sensor was mounted at a height of 2 m, the rain gauge at a height of 1 m and the soil moisture and temperature sensor at 0.1 m below the surface. Table 5.1 gives an overview of the sensors at all 14 stations. WIND, TRH, TIR and FSW provided measurements every minute, SM every five minutes. RAIN gave five minute sums of precipitation.

In the middle of the airport close to station 08, is a climate reference station of the German Meteorological Service (DWD) (see Fig. 5.1 and Fig. 5.2). To find out whether this site is most representative for the airport, datasets for air temperature, humidity and wind speed of the DWD station are compared to network measurements.

In Fig. 5.3, daily mean temperatures and pressure and daily sums of precipitation are shown. The measurements were taken by the DWD station. The lowest daily mean temperature was reached on May 10th with 6.1°C . The highest daily mean temperature was 18°C on April 29th. On days at the end of April the air temperature was more than 10°C on average, most days in May were colder than 10°C . On two nights (May 5th and 10th), the temperatures dropped below the freezing point. The mean air pressure was below 1010 hPa for the majority of days. Only at the beginning and at the end of the campaign did mean values exceed 1020 hPa. Another high pressure system passed the site on May 4th and 5th with daily mean pressure more than 1015 hPa.

For more information about the weather situation, six hour analyses of the DWD are considered (wetter3.de, 2011), which are summarised briefly in the following: During the first two days, the site was dominated by a high pressure system and south-easterly winds. On April 26th, an occlusion crossed the airport and brought slight rainfall of 2 mm

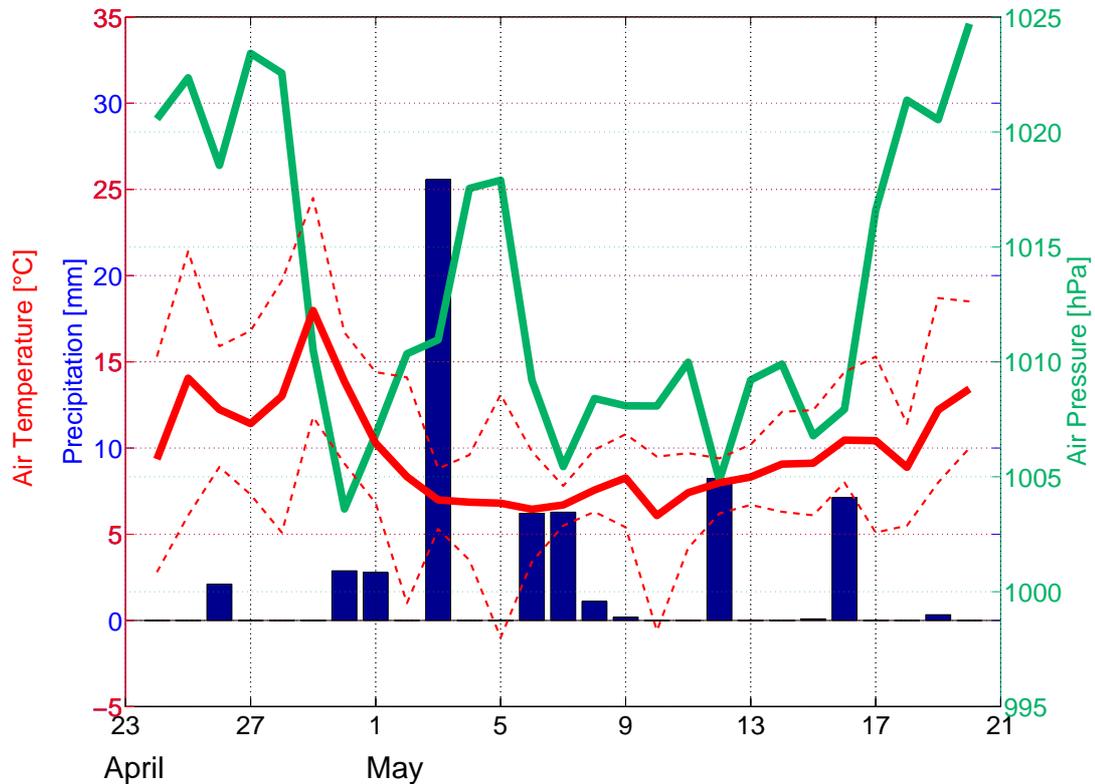


Figure 5.3: Weather Situation during the measuring campaign at Hamburg Airport. The red line indicates daily mean air temperature, the dashed red lines indicate the maximum and minimum temperature, the green line indicates daily mean air pressure. The blue bars are daily sums of precipitation.

in total. Over the next three days, the area was under high pressure influence in a warm sector with temperatures up to 25°C on the 29th. The wind shifts towards the northwest. On the following two days, two cold fronts passed the measurement site with approximately 3 mm rainfall on each day and easterly winds. For the rest of the campaign, the dominant wind direction is northwesterly. An occlusion influenced the weather on May 3rd and brought more than 15 mm precipitation. The rainfall from May 6th to 9th was also caused by an occlusion and a cold front, but it was less intense with 6 mm on the first two days and only 1 mm on the third day. On May 12th, the site was situated under the centre of a low pressure system, which brought a total rainfall of 8 mm. Another rain event, also caused by a low pressure system, occurred on May 16th. The rest of the time the site was influenced by a high pressure system with almost no precipitation. Overall, five days were completely overcast (April 26th and 30th, May 3rd, 7th and 12th) and one day, April 25th, was cloud-free. As for the first measuring campaign, the weather was quite diversified. This allows an investigation of different influences, e.g. wind speed, on the meteorological conditions close to the surface.

5.2 Variability

To investigate, whether the area of the Hamburg Airport from the meteorological point of view is as homogeneous as the environmental circumstances presume, mean values of air temperature, specific humidity, surface temperature and wind speed over the whole campaign are calculated for each station. Additionally, mean diurnal cycles of 10-minute means are computed.

The mean air temperature for each station is illustrated as circles in Fig. 5.4a. For all stations, the averaged air temperature ranges between 9.3°C and 9.8°C , leading to a mean difference of 0.5 K across a distance of 2.7 km. That is more than ten times the mean difference between the sensors during the calibration period (Chapter 3.1). Therefore, the variability in air temperature at the airport is due to external influences and not only due to measurement uncertainties. Overall, stations 01 to 07 along the north-south runway are colder on average than stations 09 to 14 along the east-west runway. The coldest site is at station 01. Stations 02 and 03 are the next warmest sites with 9.45°C . All three stations are located in the northern part of the airport, which lies in a rural environment. Stations 05 and 06 are colder than 9.6°C , only station 07 is an exception. It is the warmest of all stations. The centre of the airport, where station 08 is located, is the least sheltered from the wind in all directions. Hence, the atmosphere at this station is well mixed, leading to a moderate mean air temperature of 9.59°C . Stations 10 to 14 are warmer than 9.5°C on average. They are located in a more urban environment.

In Fig. 5.4b, the mean diurnal cycle for 10-minute means of air temperature is shown for each station and for the mean of all stations. It ranges from 5.8°C at 6 a.m. at station 01 to 13°C at 4 p.m. at station 10. The variability in air temperature is higher during the night than during the day. At night, the differences between the sites are up to 1 K; during the day, not more than 0.5 K. To clarify these differences, Fig. 5.5 displays the minimum and maximum air temperatures of the mean diurnal cycle for each station. Stations 01 to 03 are colder during the night than all other stations. The rural environment around these stations supports nocturnal formation of cold air, that leads to the lowest mean temperatures in Fig. 5.4a at these stations. In contrast, stations in the more urban environment cool down less because of urban heating effects. Stations, that are located in the middle of the airport, are warmer than the network average during the night and colder than the network average during the day, because they are most exposed to the wind. Therefore, the near surface atmosphere is well mixed. That causes higher temperatures at night and lower temperatures during the day.

Specific humidity is computed based on air temperature and relative humidity measurements at the network stations and air pressure measurements at the DWD station. Unfortunately, only the humidity measurements at stations 01, 02, 05, 08 and 13 are reliable. All other stations have to be excluded from the analysis of specific humidity. The mean specific humidity in Fig. 5.6a is between 5.47 g kg^{-1} at station 01 and 5.55 g kg^{-1} at station 05. That results in a mean difference of 0.08 g kg^{-1} across a distance of 1.75 km. The mean difference between the network sensors during the calibration period is 0.004 g kg^{-1} and,

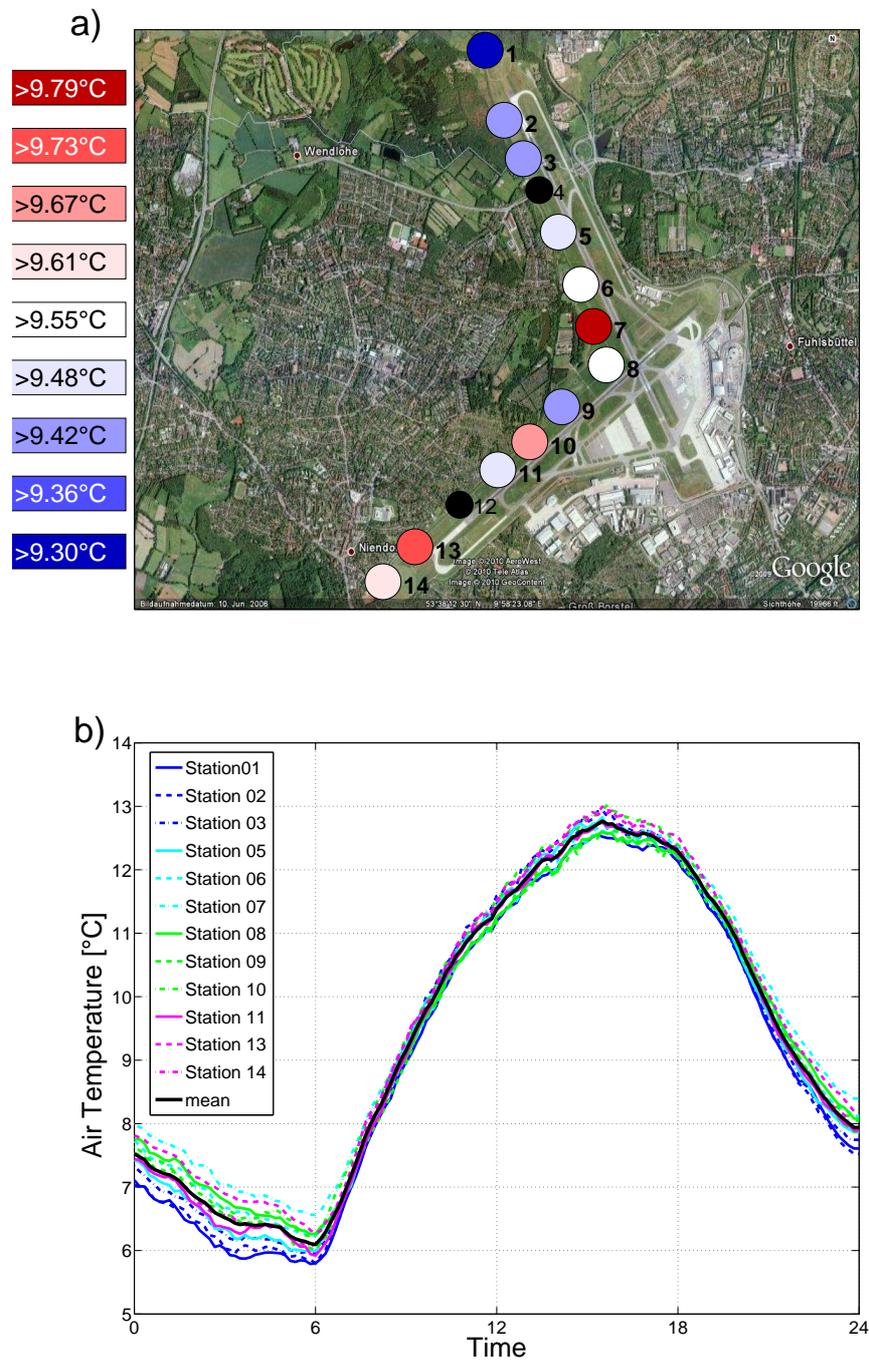


Figure 5.4: a) Map of mean air temperatures during the campaign at Hamburg Airport: Circles stand for network stations, squares stand for reference stations. b) Mean diurnal cycle of air temperature for every station (coloured lines) and the mean of all stations (black line) with ten minute temporal resolution.

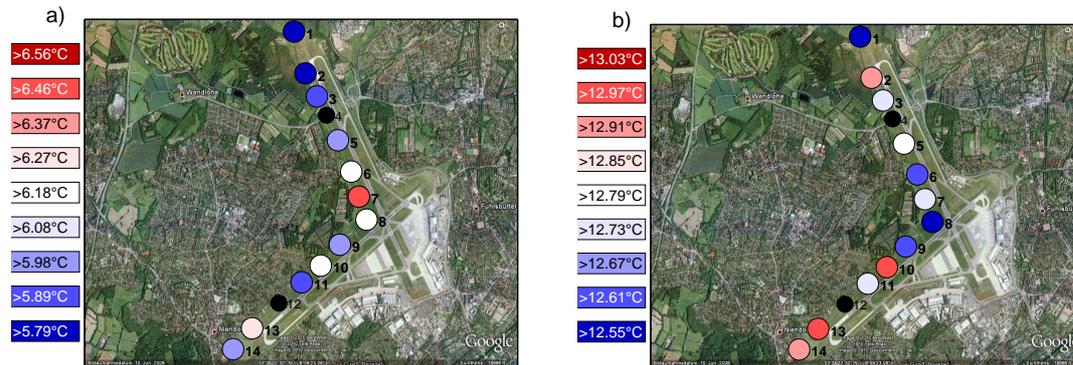


Figure 5.5: Map of (a) minimum and (b) maximum of the mean diurnal cycle of air temperatures during the campaign at Hamburg Airport for each network station.

therefore, one order of magnitude smaller. The variability in specific humidity is hence caused by external influences. Station 01 is the driest station over the whole measurement period. It is also the station with the lowest mean temperature. It is followed by stations 02 and 08. Their specific humidity is 5.49 g kg^{-1} . Station 02 is also one of the coldest stations. Station 08 is the one most exposed to the wind and, hence, the atmosphere near the surface is drier. The two most humid sites are at station 05 and 13. Station 13 is the one with the highest mean air temperature out of these five stations. It has a mean specific humidity of 5.53 g kg^{-1} . The mean specific humidity at station 05 amounts to 5.55 g kg^{-1} . This station is located close to a water reservoir, that may cause relatively high humidity at this site.

Mean diurnal cycles of 10-minute mean specific humidity are presented in Fig. 5.6b. They show two maxima for all stations: One in the morning after sunrise between 9 a.m. and 10 a.m. and one in the evening between 7 p.m. and 8 p.m., both with humidities of more than 5.6 g kg^{-1} . In between, there is a minimum in the early morning at 4 a.m. with specific humidity less than 5.1 g kg^{-1} at station 01 and a local minimum in the afternoon. The maximum difference between the stations is almost twice as large during the night (0.25 g kg^{-1}) than during the day (0.15 g kg^{-1}). Striking is the diurnal cycle of station 13. It is least pronounced because station 13 is most humid during the night due to relatively low wind speeds (see Fig. 5.8) and driest during the day. Station 05 is also wetter than the average during the night and, in contrast to station 13, also during the day. The second maximum in the evening is especially pronounced. That could be due to the water reservoir, that evaporates during the day. Stations 01 and 02 in the rural environment surprisingly are driest at night, although the soil moisture is highest at this part of the airport (50%). This could be due to a stable nocturnal stratification because of the cold air formation (see Fig. 5.4b). This might lead to formation of ground fog in a thin layer right above the surface but underneath the measurement height. Site 08 is closest to the average of all stations throughout the day. During the daytime the evaporation seems to

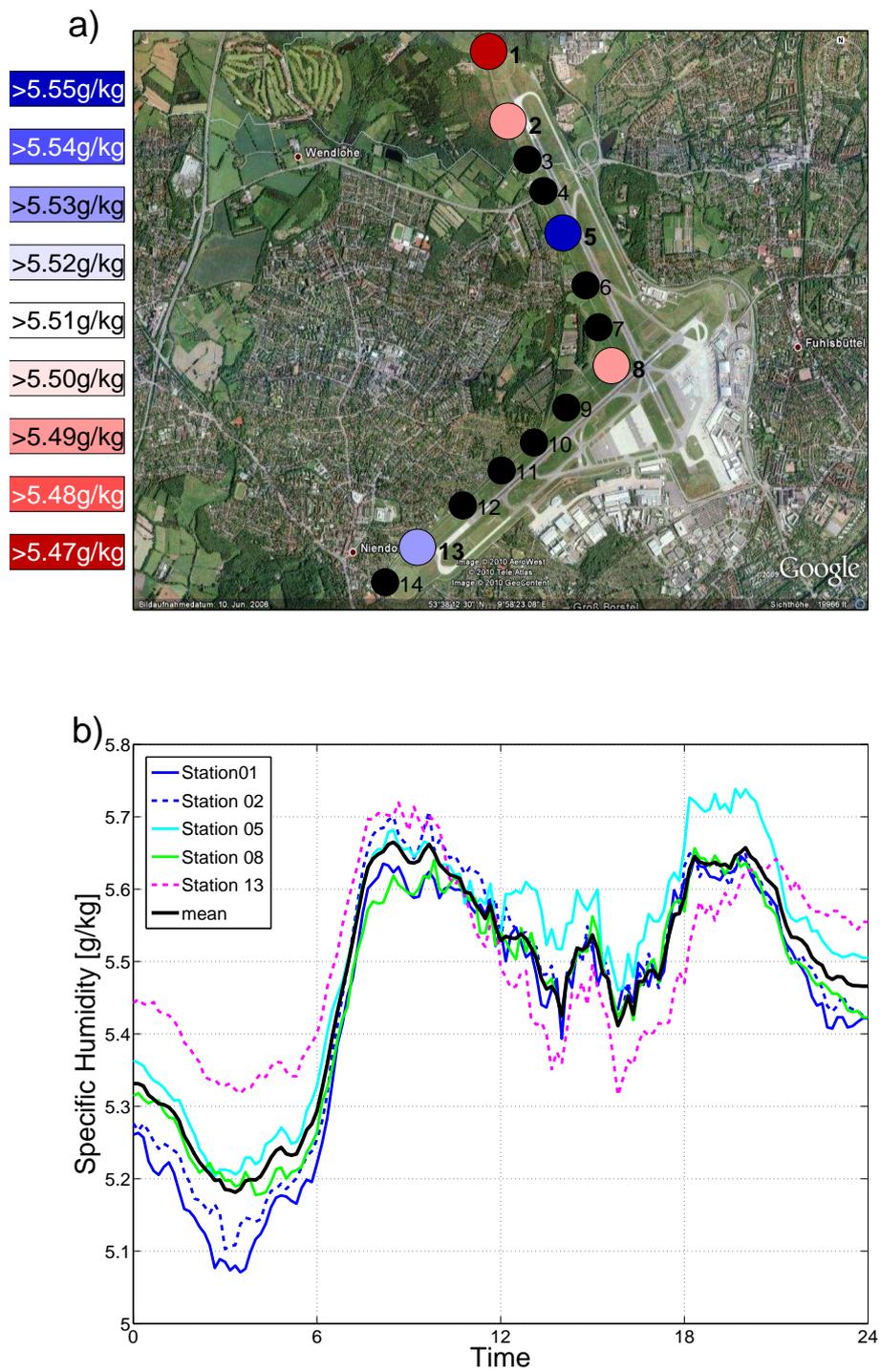


Figure 5.6: Same as Fig. 5.4, but for specific humidity.

be quite high because of relatively high soil moisture (40%) at the centre of the airport. At night, the high wind speed leads to a more mixed atmosphere than at the ends of the runway and the station is drier than station 05 and 13 in more calm areas.

For measuring surface temperature, TNX sensors are used. Station 01 did not provide reliable data and is hence excluded from the analysis. All stations are placed on grasslands, which only differ in the height of the grown grass. Mean surface temperatures over the whole campaign are shown in Fig. 5.7a for each station. Noticeable is station 08 at the centre of the airport, which is distinctly colder than the rest of the sites. It has a mean surface temperature of 9.47°C and is almost 0.43 K colder than the neighbouring station 09 and 0.65 K colder than the neighbouring station 07. To ensure, this is not due to a measurement error, a second surface temperature sensor was installed in October 2010 and provides the same results. All other sites have mean surface temperatures of more than 10°C , except for station 02 and are all in a range of 0.5 K. The warmest site is at station 03 (10.48°C), closely followed by station 06 (10.47°C). The differences in surface temperature most likely are not caused by the vegetation or surface type, because it is homogeneous at the airport. The variation between the sites might be induced by the soil moisture. It is highest in the northern part (50% at station 01) and at the centre of the airport (40% at station 08). These areas are also colder than others. At stations along the runways and in the south of the airport (measurements are available at station 03, 07, 11 and 14), the soil moisture is less than 32% and the mean surface temperatures are higher.

The mean difference between the coldest station 08 and the warmest station 03 is 1 K across a distance of 1.95 km. The mean difference of the network sensors during the calibration period is only 0.13 K. Therefore, not only measurement uncertainties are responsible for this difference. Environmental influences also play a role.

In Fig. 5.7b, the mean diurnal cycle is shown for each station. It ranges from 5°C in the early morning to almost 18°C in the afternoon. Fig. 5.7b emphasises that the surface at station 08 is up to 1 K colder than at all other sites, especially during the day. The surface at station 05, however, is up to 0.8 K warmer than the rest of the sites during the day. During the night, it is one of the coldest stations. The surface temperatures at stations 07 to 09 at the centre of the airport and at station 02 in the northern part are coldest throughout the day. The differences between stations 05 and 08 are three times larger during the day, when the sun warms the surface, than at night, where it is less than 1 K. The surface temperatures of all other stations fall within an interval of 0.5 K around the average of all stations the whole day.

The mean wind speed in Fig. 5.8a ranges between 1.78 m s^{-1} and 2.84 m s^{-1} for all stations. That leads to a mean difference of 1.06 m s^{-1} across a distance of 2.9 km, which is three orders of magnitude higher than during the calibration period. The lowest mean wind speeds occur at the very ends of both runways. These sites are sheltered by trees and hedge banks. Therefore, the wind is shielded in certain directions. At the centre of the airport, the field is more open and the wind can blow unhindered. From the centre at station 08 to the northern end at station 02, the mean wind speed decreases constantly from 2.84 m s^{-1} to 1.99 m s^{-1} . At station 01, it increases again by 0.08 m s^{-1} , leading to a

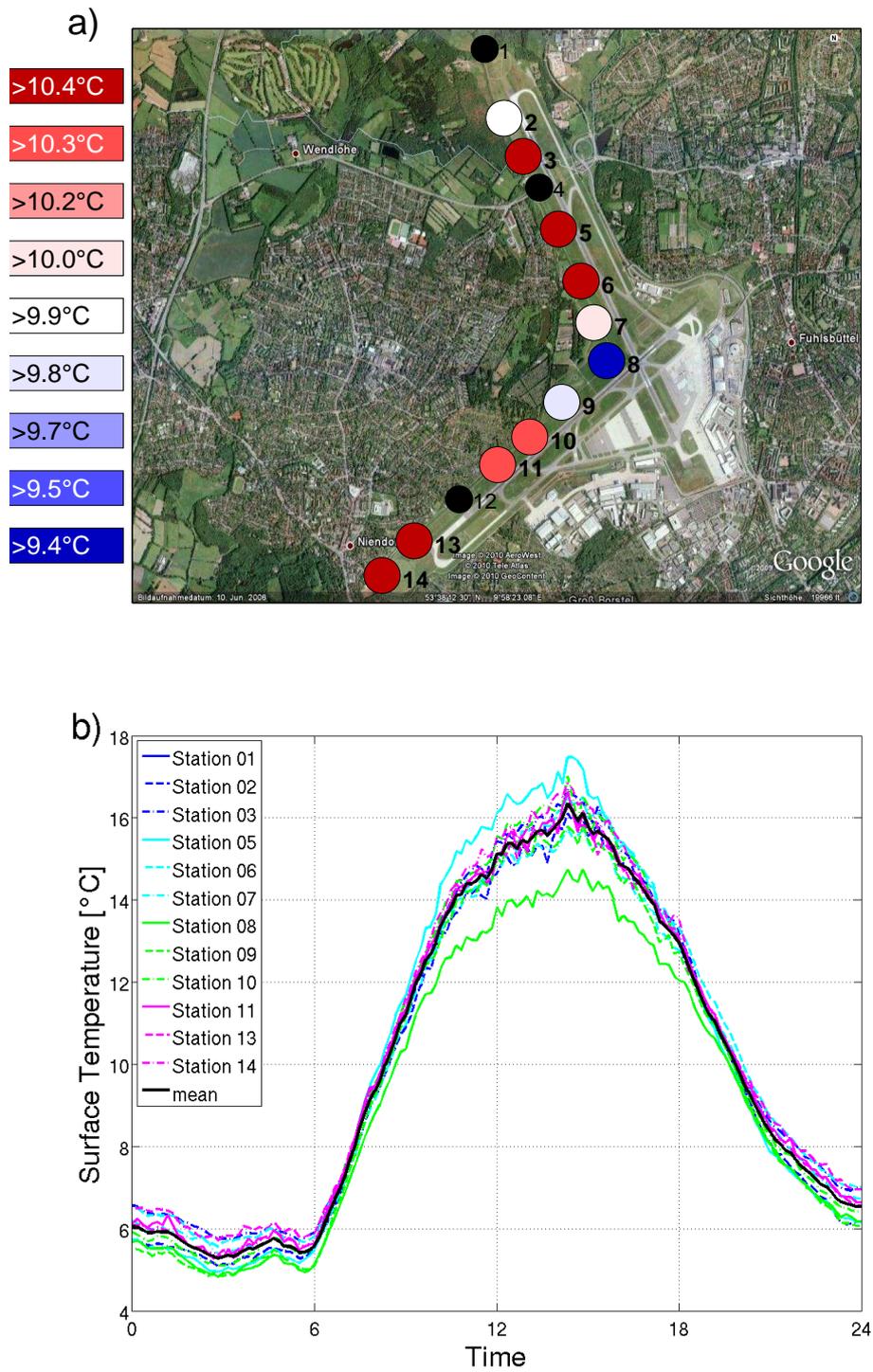


Figure 5.7: Same as Fig. 5.4, but for surface temperature.

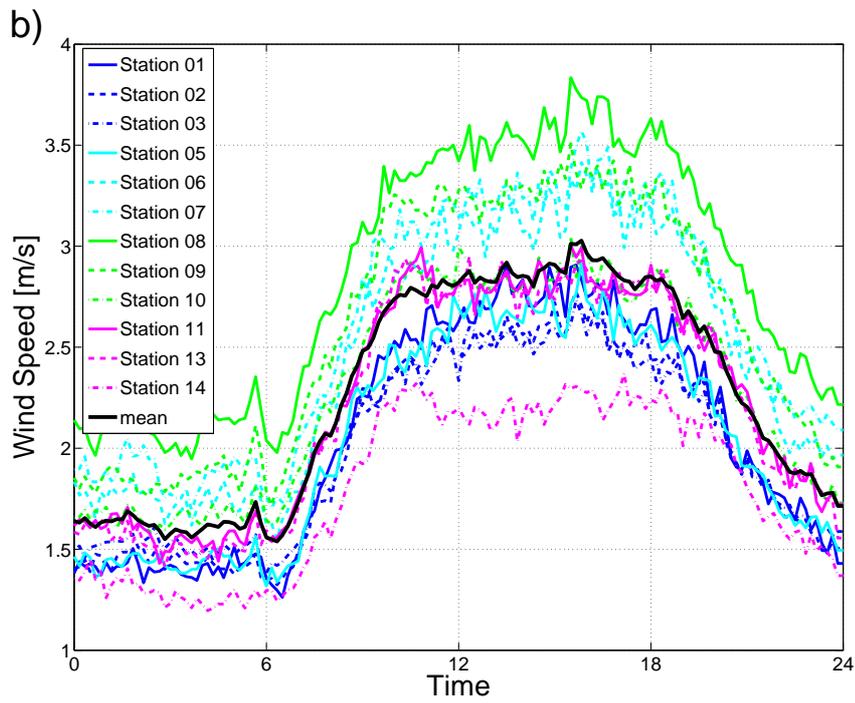
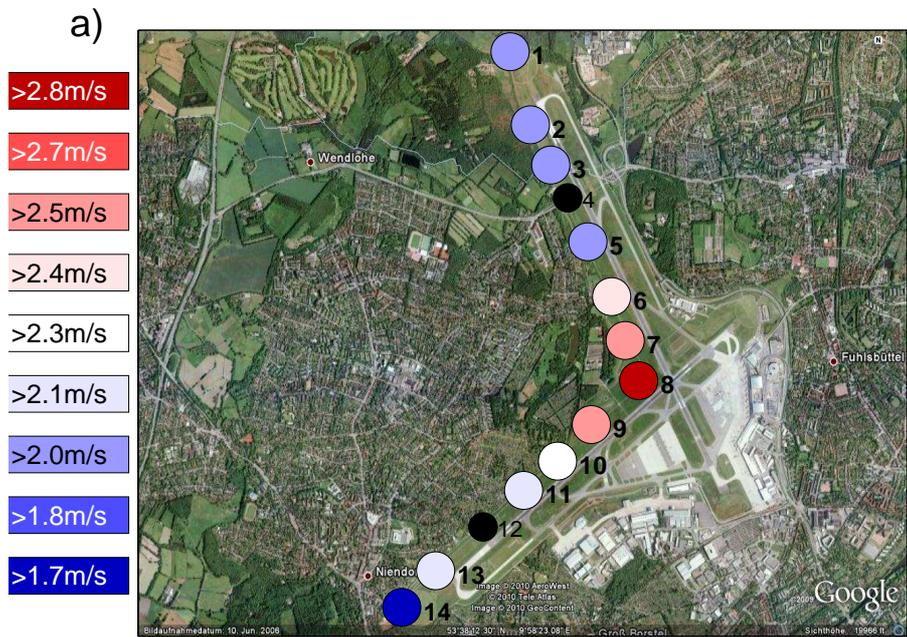


Figure 5.8: Same as Fig. 5.4, but for wind speed.

mean wind speed of 2.07 m s^{-1} . The site at station 01 is obstructed by trees only on the western side. The rest of the area is relatively open so that the wind speeds are higher than at stations 02, 03 and 05. In contrast, the very western end of the airport, where station 14 is located, is shielded in three directions because of trees and hedge banks and the wind speed is lowest of all stations. Station 08 at the centre of the airport is least shielded and, therefore, it is exposed to the highest wind speeds.

The mean diurnal cycles in Fig. 5.8b illustrate that wind speeds are lowest at the very ends of both runways and highest at the centre of the airport during the whole day. At night, wind speeds are between 1.25 m s^{-1} at station 14 and 2.4 m s^{-1} at station 01. During the day, they range between 2 m s^{-1} and 3.8 m s^{-1} for the same two stations. At station 14 (08), the diurnal cycle is least (most) pronounced. The values of wind speed during the day are even smaller at station 14 than the wind speeds during the night at station 08. Overall, wind speeds along the north-south runway are lower than along the east-west runway.

Fig. 5.9 shows the distribution of wind direction and speed for station 03 located at the north-south runway and for station 10 located at the east-west runway. During the campaign the wind came mostly from west to northwest direction. The distribution at both sites is clearly different: Station 03 is sheltered from the wind at the western site. The dominating wind direction is northwest. The wind follows the direction of the runway. For station 10, the wind came mostly from the west because in the northwest of the site trees block the wind. At both sites winds from the northeast occur. They are distinctly stronger at station 10 because the runway is oriented to this direction and the wind can blow unhindered.

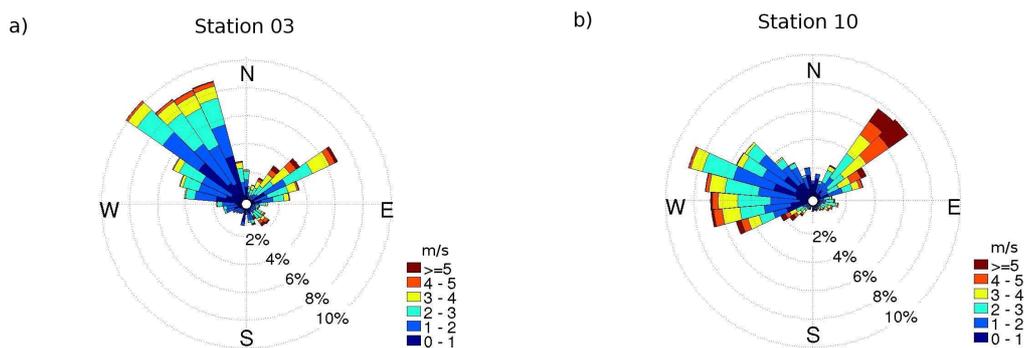


Figure 5.9: Distribution of wind direction and velocity at (a) stations 03 located at the north-south runway and (b) station 10 located at the east-west runway.

5.3 Explaining Factors for the Variability

In the previous chapter, the variability of air temperature, specific humidity and surface temperature at the Hamburg Airport is described. These differences between sites are not only caused by measurement inaccuracies but by external influences. To find out, which influences drive the variability in air temperature, specific humidity and surface temperature, the data points are categorised into the same six groups as for FLUXPAT2009 (see Chapter 4.3): It is distinguished between day (between sunrise and sunset) and night (between sunset and sunrise) and between three different wind classes (low wind speeds less than 1.2 m s^{-1} , moderate wind speeds between 1.2 m s^{-1} and 2.5 m s^{-1} and high wind speeds above 2.5 m s^{-1}). During the day, there are almost twice as many data points as during the night. During the day, the group of low wind speeds contains 2098 data points, the group of moderate wind speeds is almost five times larger with 10417 data points. The group of high wind speeds contains most data points with 12119 measurements. At night, the group of low wind speeds is largest. It includes 5522 data points. Moderate wind speeds occur 5450 times. The group of high wind speeds is only half as large with 2794 data points. Mean differences ΔT_{Air} , Δq and ΔT_{Surf} between each station and all other stations are calculated as a measure of systematic differences between the sites and their 10% and 90% quantiles as a measure of random deviations for an arbitrary one-minute interval. The results are shown in Fig. 5.10 to 5.12. Stations are sorted by the length of their mean fetches, starting with the shortest. Different colours indicate the location of the stations: All stations along the north-south runway are marked in red, stations along the east-west runway are marked in green and the station in the middle of the airport is marked in blue.

For air temperature in Fig. 5.10, the mean difference is in general approximately half the quantile range. Therefore, random deviations dominate the air temperature variability. The mean difference ranges between -0.83 K at station 01 and 0.6 K at station 07 both at night and for low wind speeds in Fig. 5.10d. Station 01 also has the largest quantile ranges with values up to 2.28 K . Stations 05 and 11 are close to the mean of all stations under all conditions. Their mean differences are between -0.17 K and 0.03 K and between -0.14 K and -0.01 K for station 05 and 11, respectively.

Wind speed clearly has an effect on the variability in air temperature. The largest mean differences occur at low wind speeds for both day and night. They decrease with increasing wind speeds. While the mean difference for wind speeds less than 1.2 m s^{-1} ranges from -0.83 K to 0.6 K , for wind speeds larger than 2.5 m s^{-1} , the amplitude of the variability of air temperature is distinctly smaller at the airport site; all stations lie within an interval of $\pm 0.55 \text{ K}$. The quantile range Δq_{90-10} is also reduced with rising wind speeds. This is evident especially at night (Fig. 5.10d-f). For low wind speeds, the air temperature is quite variable with Δq_{90-10} ranging between 1.68 K and 2.28 K . High wind speeds almost erase this variability, because the atmosphere close to the surface is well mixed; Δq_{90-10} is smaller than 0.65 K .

During the day, there is no distinct difference in air temperature between stations along the north-south runway and the east-west runway. But at night, stations 01 to 05 along the

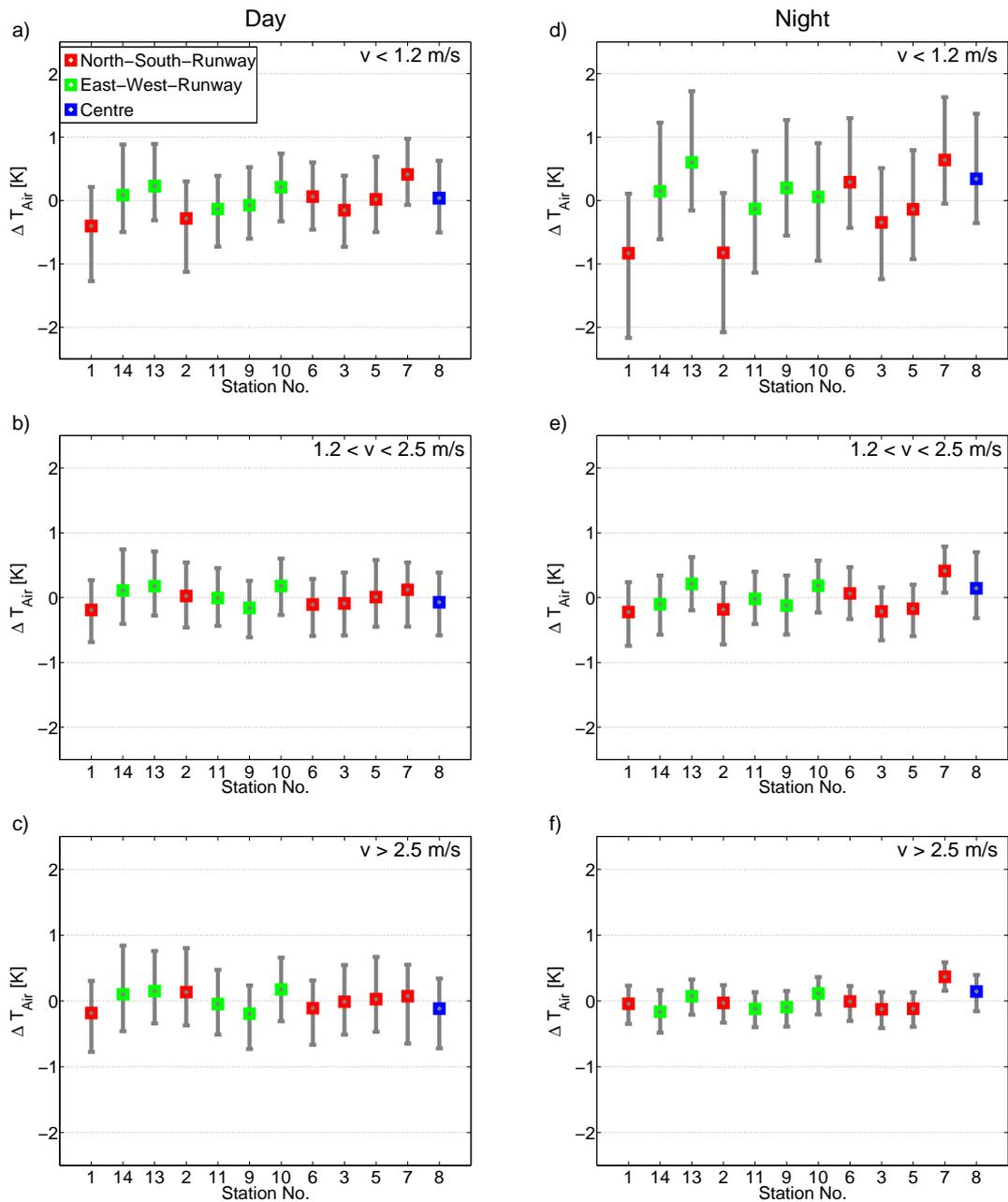


Figure 5.10: Mean difference and 10% and 90% quantile of air temperature between sunrise and sunset (Day in the left column) and between sunset and sunrise (Night in right column) for wind speeds lower than 1.2 m s^{-1} (top row), between 1.2 and 2.5 m s^{-1} (middle row) and higher than 2.5 m s^{-1} (bottom row). The colours show whether the stations are located along the north-south runway (red), along the east-west runway (green) or at the centre of the airport (blue). On the x-axis stations are ordered according to their fetch, starting with the shortest.

north-south-runway depicted in red are clearly colder for low and moderate wind speeds than all other stations. This strong nocturnal cooling is caused by the rural environment. Stations in the middle of the airport and along the east-west runway are more influenced by an urban environment and, therefore, warmer during the night. In particular station 07 is warmer than all other stations at night.

An influence of the mean fetch on the variability of air temperature is not evident. Only stations 07 and 08, having the largest fetches, are among the warmest stations during the night because the atmosphere is well mixed at these open sites. Station 01, with the shortest fetch, is the coldest station in almost all conditions, but this is more due to nocturnal cooling caused by the rural environment.

In Fig. 5.11, the results for specific humidity are shown. Due to the limited number of stations, all conclusions are very tenuous. The quantile ranges are two to three times larger than the mean differences. This means that random deviations are the explaining factors for the variability in specific humidity. Systematic influences are less important.

During the day in Fig. 5.11a-c, wind speeds have nearly no influence on the mean difference in specific humidity at stations 01 to 08. Station 05 located close to a water reservoir is wettest in all conditions, stations 01 and 08 are driest. The mean differences remain in a range of $\pm 0.09 \text{ g kg}^{-1}$. Δq_{90-10} slightly increases with increasing wind speeds from around 0.4 g kg^{-1} for low wind speeds up to approximately 0.5 g kg^{-1} for high wind speeds. Only station 13 seems to be influenced by wind speed during the day. Whereas it is the wettest station for wind speeds less than 1.2 m s^{-1} with a mean difference of 0.1 g kg^{-1} , it is the driest site for wind speeds more than 2.5 m s^{-1} with a mean difference of -0.07 g kg^{-1} . The quantile range surprisingly rises with increasing wind speed for all stations (also seen for the FLUXPAT campaign in Chapter 4.3). Normally it would be expected to decrease with increasing wind speed due to the more thoroughly mixed atmosphere close to the surface. At night in Fig. 5.11d-f, the wind plays an important role. Although station 13 is the wettest site and station 01 is drier than most of the other stations for all wind conditions, the variability decreases with increasing wind speeds. The mean difference in specific humidity ranges from -0.16 g kg^{-1} at station 01 to 0.21 g kg^{-1} at station 13 for low wind speeds. The quantile range is more than 0.47 g kg^{-1} . For moderate wind speeds, the mean difference decreases to a range of -0.06 g kg^{-1} at station 01 and to 0.08 g kg^{-1} at station 13. The quantile range is less than 0.24 g kg^{-1} because the atmosphere close to surface is well mixed and the specific humidity is distributed more homogeneously. For high wind speeds, it is about the same.

Fetch and location have no clear impact on the variability of specific humidity. Although station 13 is the driest site for almost all conditions, it is difficult to say whether this is caused by the location at the east-west runway. It is the only station with humidity measurement along this runway and, therefore, it is not clear whether other factors influence the specific humidity at this site.

The analysis for surface temperature is presented in Fig. 5.12. In contrast to the variability of surface temperature at the FLUXPAT2009 site in Chapter 4.3, the variability at the airport is dominated by random deviations rather than by systematic effects. The quantile

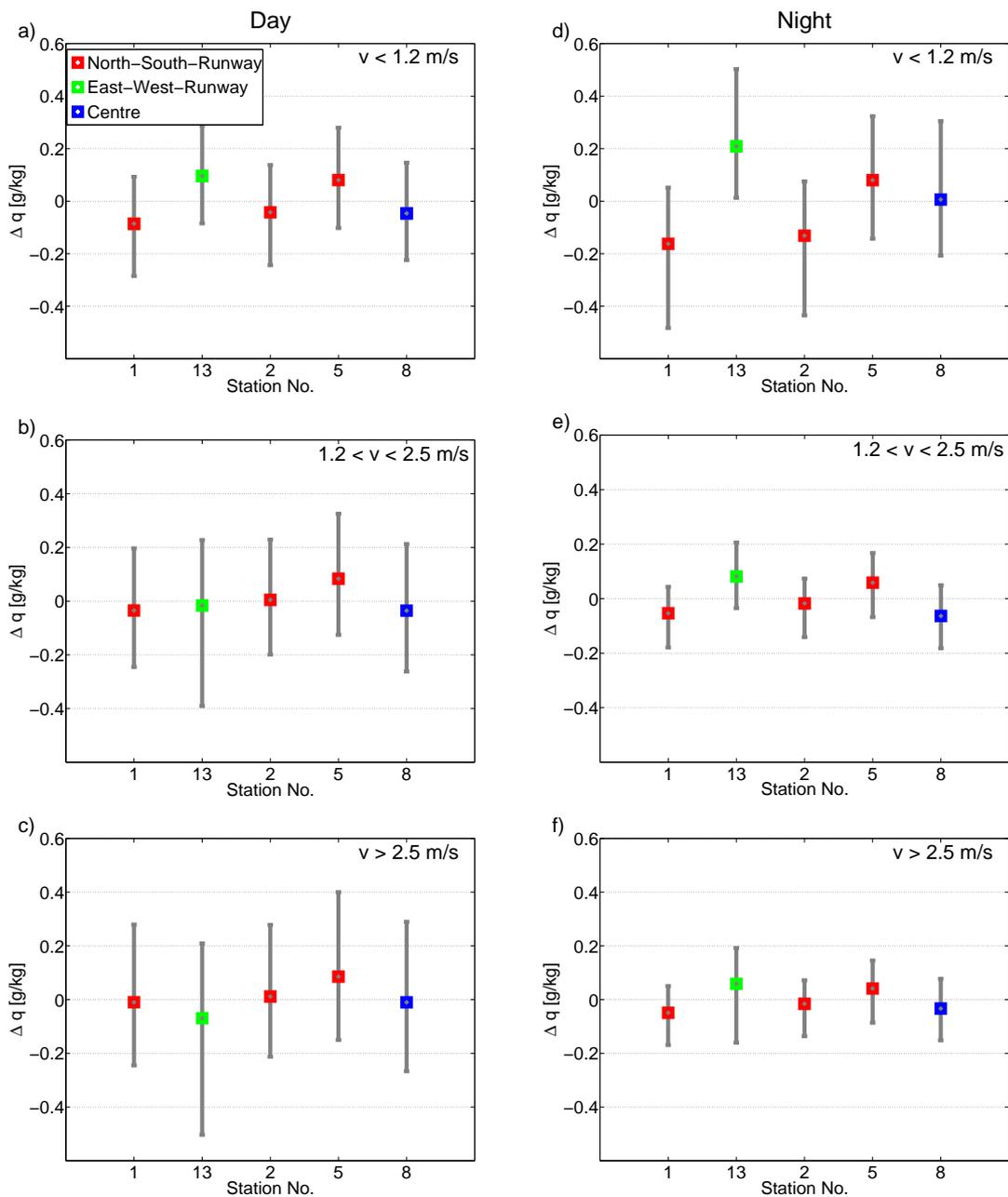


Figure 5.11: Same as Fig. 5.10, but for specific humidity.

range is two to three times the mean difference.

The variability in surface temperature decreases with increasing wind speed at night as expected. Higher wind speeds lead to a more homogeneous distribution of air temperature which, in turn, causes a more homogeneous pattern of surface temperature. During the day the variability surprisingly increases with increasing wind speed. This might be due to the fact that low wind speeds occur in the morning and in the evening (see Chapter

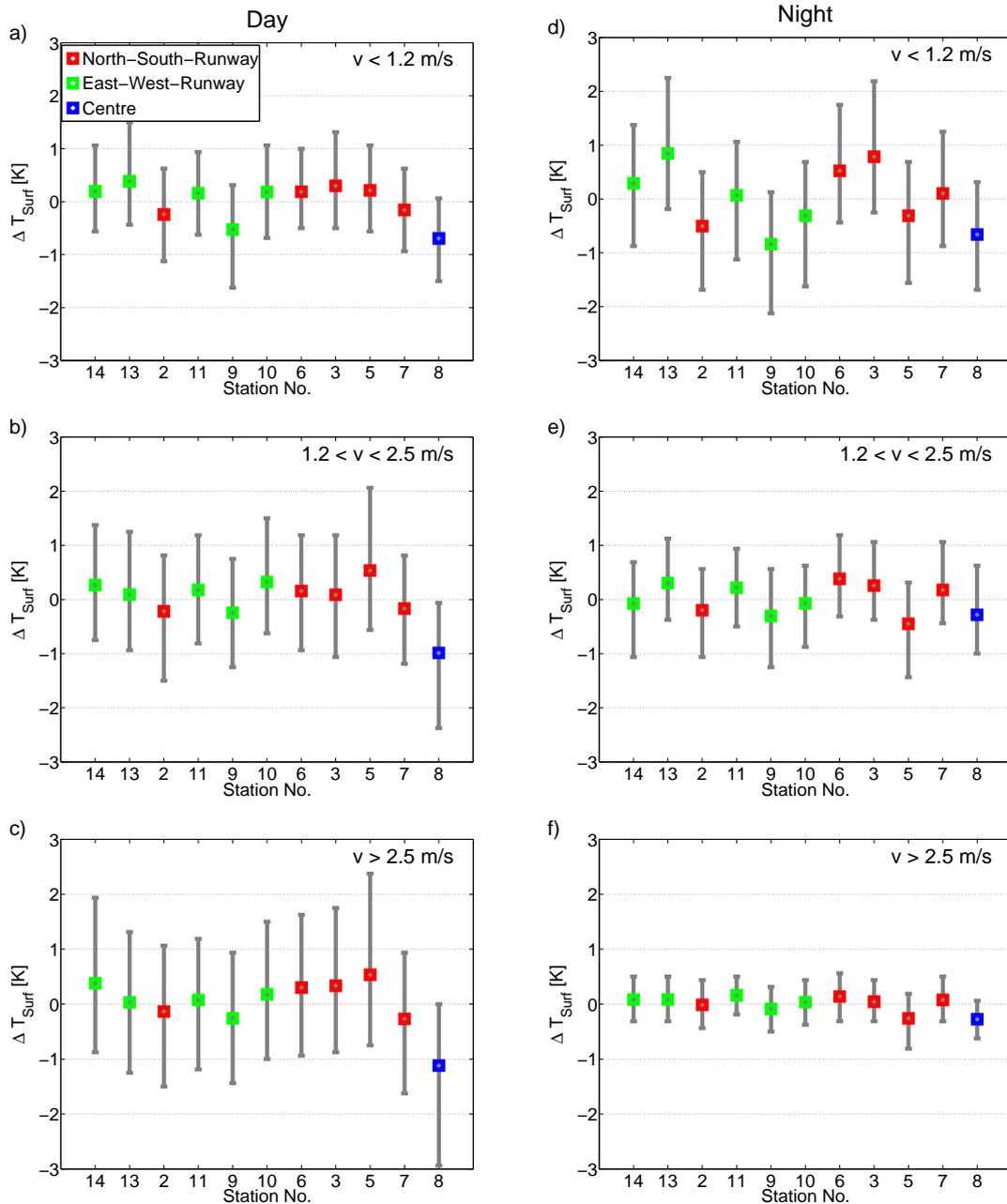


Figure 5.12: Same as Fig. 5.10, but for surface temperature.

5.2), when the influence of the sun is small. The wind speeds are higher in the afternoon when the solar radiation causes larger variabilities in surface temperature. For low wind speeds during the day in Fig. 5.12a, the mean difference ranges between -0.7 K at station 08 and 0.38 K at station 13. During the night in Fig. 5.12d, the variability is slightly higher. It is between -0.84 K at station 09 and 0.84 K at station 13. The quantile range is also larger during the night. For moderate wind speeds in Fig. 5.12b and e, the mean differences are

almost the same for day and night. They range between -0.99 K and 0.54 K. Δq_{90-10} is 1.5 times larger during the day than at night. The mean differences in surface temperature during the day for high wind speeds (Fig. 5.12c) are almost the same than for moderate wind speeds, but the quantile range is more than 2.38 K. This might be caused by quick changes in cloud cover due to high wind speeds. At night, the mean differences range between -0.28 K at station 08 and 0.16 K at station 11. The quantile range is less than 1 K. The location of the stations does not have an influence on the variability of surface temperature. Only station 08, at the centre of the airport, measured comparably low surface temperatures in all kinds of conditions. The surface at this site is distinctly colder than at all other sites, especially for moderate and high wind speeds during the day, because the soil moisture is higher at the centre of the airport. The fetch does not appear to influence the surface temperature at all.

5.4 Cluster Analysis

A detailed analysis of the factors that influence air temperature, specific humidity, surface temperature and wind speed near the surface, is given in the previous chapter. To identify stations that are exposed to similar environmental influences, cluster analyses are performed. As described in Chapter 4.4, two different kinds of cluster analysis are used: The hierarchical Average Linkage Method for one-minute and normalised one-minute data and the non-hierarchical k-means Method. After dividing the sites at the airport into groups, the mean diurnal cycle of each group is calculated to identify differences between the clusters. Finally, the error resulting from using fewer than twelve stations at the airport is calculated.

Fig. 5.13a illustrates the dendrogram of the hierarchical cluster analysis of air temperature. Stations 01 to 05 in the northern part of the airport are separated from the rest at an RMSD of 0.63 K. Station 01 is by far the coldest site and is separated from the other three stations at a RMSD of 0.52 K. The larger cluster is split at 0.53 K into two groups containing stations 06 to 09 and stations 10 to 14, respectively. The neighbouring stations 08 and 09 are most alike. Their RMSD is only 0.35 K, followed by stations 10 and 11 with a RMSD of 0.36 K.

The Average Linkage Analysis for normalised one-minute values in Fig. 5.13b gives the same results for three clusters. For normalised values, not stations 08 and 09, but stations 10 and 11, are most alike with a RMSD of 0.072 K. The neighbouring stations 06 and 07 are also quite similar. They have a RMSD of 0.074 K. Again, station 01 is separated from the cluster of the northern stations at a relatively high RMSD of 0.113 K. The reason for this separation cannot be the fact that station 01 is distinctly colder than the rest because using normalisation erases the influence of absolute temperatures.

Following these two analyses, a subdivision into three groups is suitable: The first cluster consists of stations 01 to 05, the second cluster consists of stations 06 to 09 and the third cluster consists of stations 10 to 14. To verify this subdivision, the k-means method is used for three clusters. The result is shown in Fig. 5.13c. It almost matches the outcome

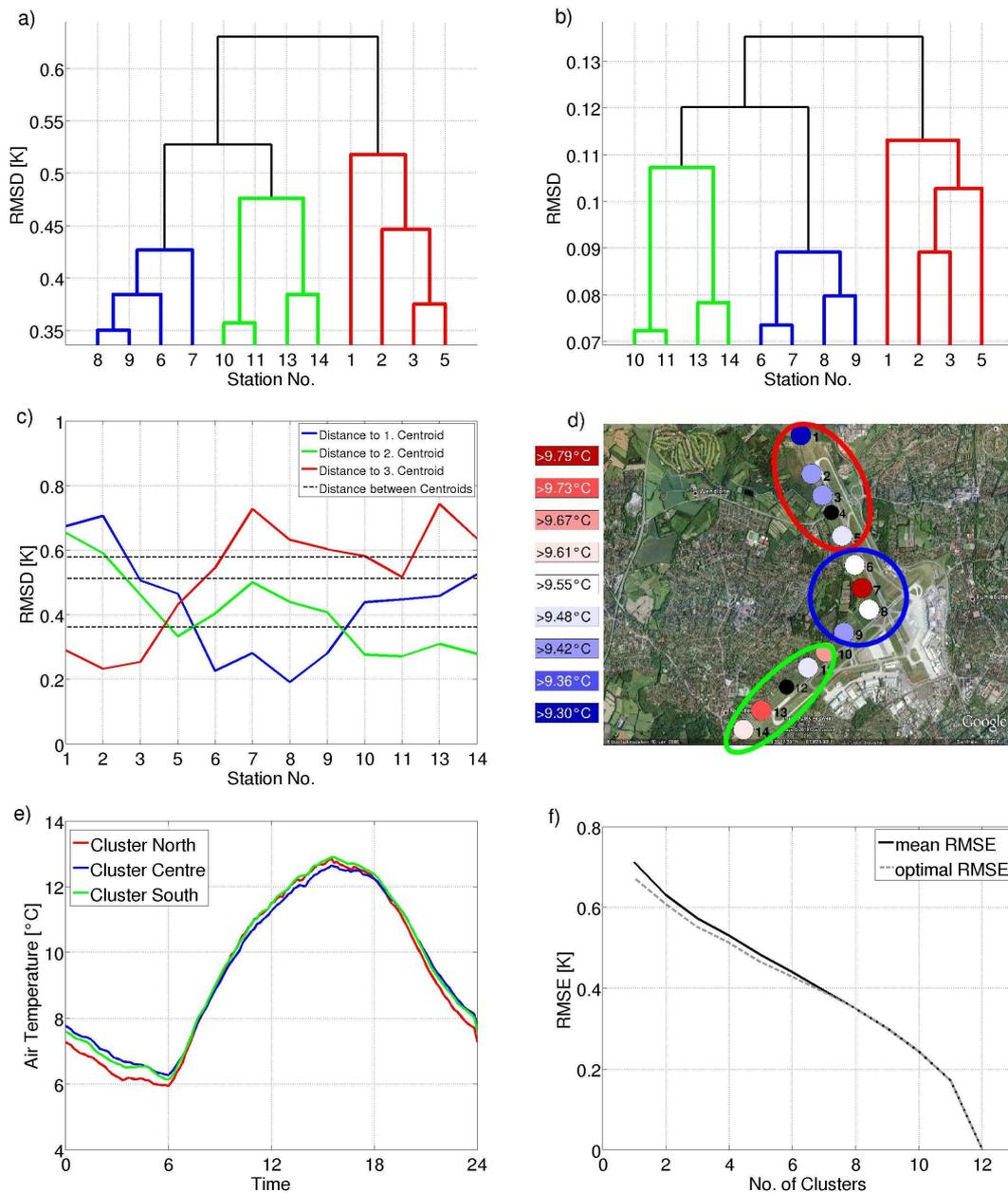


Figure 5.13: Cluster analysis for air temperature at Hamburg Airport: a) Dendrogram for one-minute data, b) dendrogram for normalised data, c) k-means analysis for one-minute data, d) clusters on the map with mean air temperature, e) mean diurnal cycle of clusters with ten minute temporal resolution, f) mean and optimal RMSE for using one to twelve stations.

of the Average Linkage Method. The only exception is station 05 that is not assigned to the group of the northern stations. With the greatest distance to the centroid of all stations (RMSD = 0.38 K), station 05 is assigned to the cluster consisting of stations 10 to 14. That is almost as large as the distance between two of the centroids. Therefore, the membership

of station 05 to this cluster is not as clear as for the other stations.

Thus, the three clusters suggested by the Average Linkage Analysis are used in the following. On the map in Fig. 5.13d, cluster *North* containing stations 01 to 05 is encircled in red, cluster *Centre* containing stations 06 to 09 is encircled in blue and cluster *South* containing stations 10 to 14 is encircled in green. This subdivision is dependent on the geographical position of the stations. Each cluster consists of neighbouring stations, suggesting local environmental influence.

To identify differences between these three groups of stations, the mean diurnal cycle of each cluster is shown in Fig. 5.13e. It is evident that nocturnal cooling is most pronounced at the stations of cluster *North* (red line). It is up to 0.5 K colder there than the other two clusters during the night due to its rural environment and low wind speeds (see Chapter 5.2). At cluster *South*, the wind speeds are also quite low during the night, but because of the urban environment nocturnal cooling is less pronounced. Cluster *Centre* is warmest by night and coldest by day, because of comparably high wind speeds. The stations in this cluster are least sheltered from the wind, because they are located in the middle of the airport.

The mean and optimal RMSE for using one to twelve stations is presented in Fig. 5.13f. If only one station were set up at the airport, the mean RMSE would have been 0.71 K. The optimal RMSE for using only the most representative station, station 06, is 0.67 K. For using one arbitrary station out of each of the three clusters *North*, *Centre* and *South*, the RMSE decreases to 0.57 K. The most representative sites are at stations 03, 08 and 10. For these three stations, the optimal RMSE is 0.55 K. Adding more stations reduces the mean and optimal RMSE by 0.045 K and 0.04 K, respectively.

For specific humidity, the cluster analyses are only done for stations 01, 02, 05, 08 and 13. The results of the Average Linkage Method for one-minute means is shown in Fig. 5.14a. The adjacent stations 01 and 02 have the greatest similarity. They form a cluster at a RMSD value of 0.16 g kg^{-1} . Station 05 is assigned to this cluster next, followed by station 08. Station 13 forms its own cluster until it is at last merged with the larger cluster at a RMSD value of 0.28 g kg^{-1} . Also in Chapter 5.2, it is evident that station 13 is different from the other stations in terms of specific humidity. The Average Linkage Cluster Analysis for normalised one-minute means of specific humidity in Fig. 5.14b gives the same results as for measured one-minute means.

Due to the fact that station 13 is clearly different from all other stations in terms of specific humidity, the k-means analysis is performed for two clusters. It provides the same two clusters (Fig. 5.14c). The distance between the centroid of the first cluster and the centroid of the second cluster (namely, station 13) is 0.25 g kg^{-1} . That is almost twice the maximum distance between the centroid of the first cluster and the stations belonging to this group and one and a half times the RMSE between the sensors during the calibration period (Chapter 3.1).

The map in Fig. 5.14d intimates that the distribution into two clusters is again based on the geographical position of the stations. Stations 01, 02, 05 and 08 form cluster *North*. They are all located along the north-south runway or at the centre of the airport. The

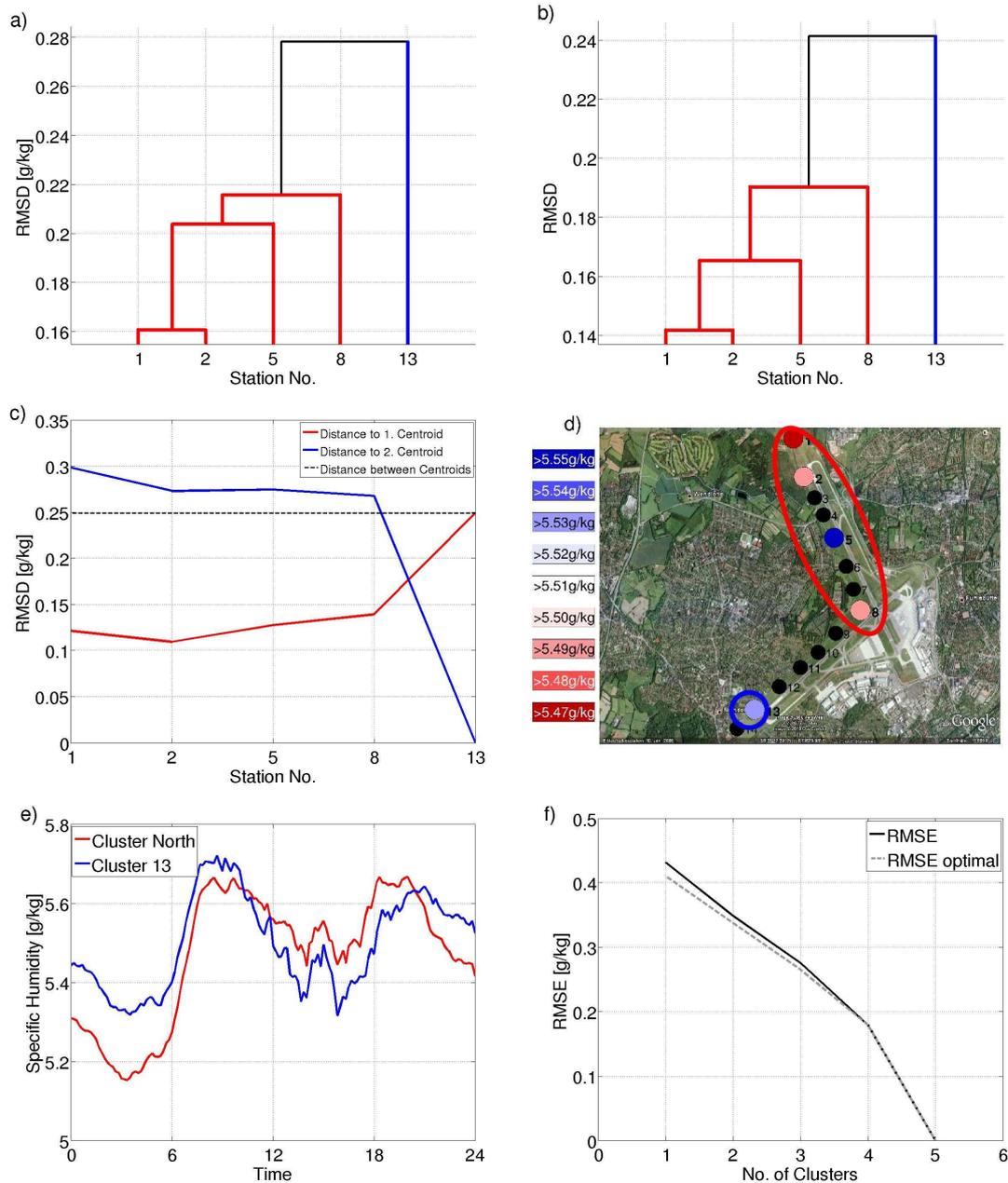


Figure 5.14: Same as Fig. 5.13, but for specific humidity.

only site with specific humidity measurements along the east-west runway constitutes its own group (cluster 13).

Fig. 5.14e presents the mean diurnal cycles of cluster *North* in red and cluster 13 in blue. The largest differences between the two clusters appear in the early morning. Cluster *North* is almost 0.2 g kg^{-1} drier than cluster 13. After the first maximum around 9 a.m., cluster 13 becomes drier than cluster *North*, but the difference between the two is not

more than 0.12 g kg^{-1} . After the second maximum in the evening, the relation is again reversed and cluster *North* becomes drier than cluster 13.

The mean and optimal RMSE for using one to five clusters is shown in Fig. 5.14f. When using only one station, the mean RMSE amounts to 0.43 g kg^{-1} , the optimal RMSE is 0.41 g kg^{-1} . Therefore, station 02 is selected as the most representative site. For two clusters, station 13 is considered as a self-contained group. The other stations form the second cluster. The mean RMSE is reduced to 0.35 g kg^{-1} . Using stations 02 and 13, the optimal RMSE amounts to 0.34 g kg^{-1} . Adding more stations reduces the RMSE by 0.08 g kg^{-1} per station.

The hierarchical cluster analysis for surface temperature in Fig. 5.15a illustrates that station 08 differs distinctly from all other stations. The other stations are divided into two groups, containing stations 02 to 07, 09 and 10 and stations 11 to 14, respectively. Stations 07 and 09 and stations 02 and 03 are most similar. In contrast to the FLUXPAT site in Chapter 4, the surface conditions are quite homogeneous at the airport with grassland at every station. Hence, this group division seems to be related to the geographical position, but is disturbed by the absolute values of the surface temperature. The coldest station by far, station 08, is separated from the rest first. Stations 07 and 09, which have similar mean values of surface temperature, are most alike.

To avoid this effect, a cluster analysis of the normalised one-minute means of surface temperature is performed and presented in Fig. 5.15b. Now the dependency for the group division on the geographical position is more evident. Stations 02 to 06 form one cluster with a maximum RMSD of 0.209 K , Stations 07 to 10 form the second cluster with a maximum RMSD of 0.184 K and station 13 and 14 merge at a RMSD of 0.195 K . Only station 11 does not fit into any of these clusters. The neighbouring stations 08 and 09 have the greatest similarity (RMSD = 0.145 K) and merge with station 07 at RMSD of 0.155 K . Stations 02 and 03 are also quite similar with RMSD of 0.16 K .

The k-means cluster analysis in Fig. 5.15c is also performed with normalised one-minute means of surface temperatures. It confirms the results of the Average Linkage Method in Fig. 5.15b for three clusters. Again, stations 02 to 06 form one group and stations 13 and 14 form another. Station 11 is assigned to the third cluster containing stations 07 to 10, but its distance to the centroid is larger than the distance between two of the centroids. This also supports the Average Linkage Method, where station 11 is assigned to a cluster at a high level of RMSD.

A division into three clusters based on the k-means method is chosen for surface temperature at the airport. The clustering based on the geographical position seems to be plausible, since the surface conditions are homogeneous. On the map in Fig. 5.15d, cluster *North*, consisting of stations 02 to 06, is encircled in red, cluster *Centre*, consisting of stations 07 to 11, is encircled in blue and cluster *South*, consisting of stations 13 and 14, is encircled in green.

The mean diurnal cycles of the clusters in Fig. 5.15e emphasise that it is difficult to identify different features for the three clusters. During the day, cluster *North* and *South* demonstrate very similar surface temperatures and cluster *Centre* is coldest. In the af-

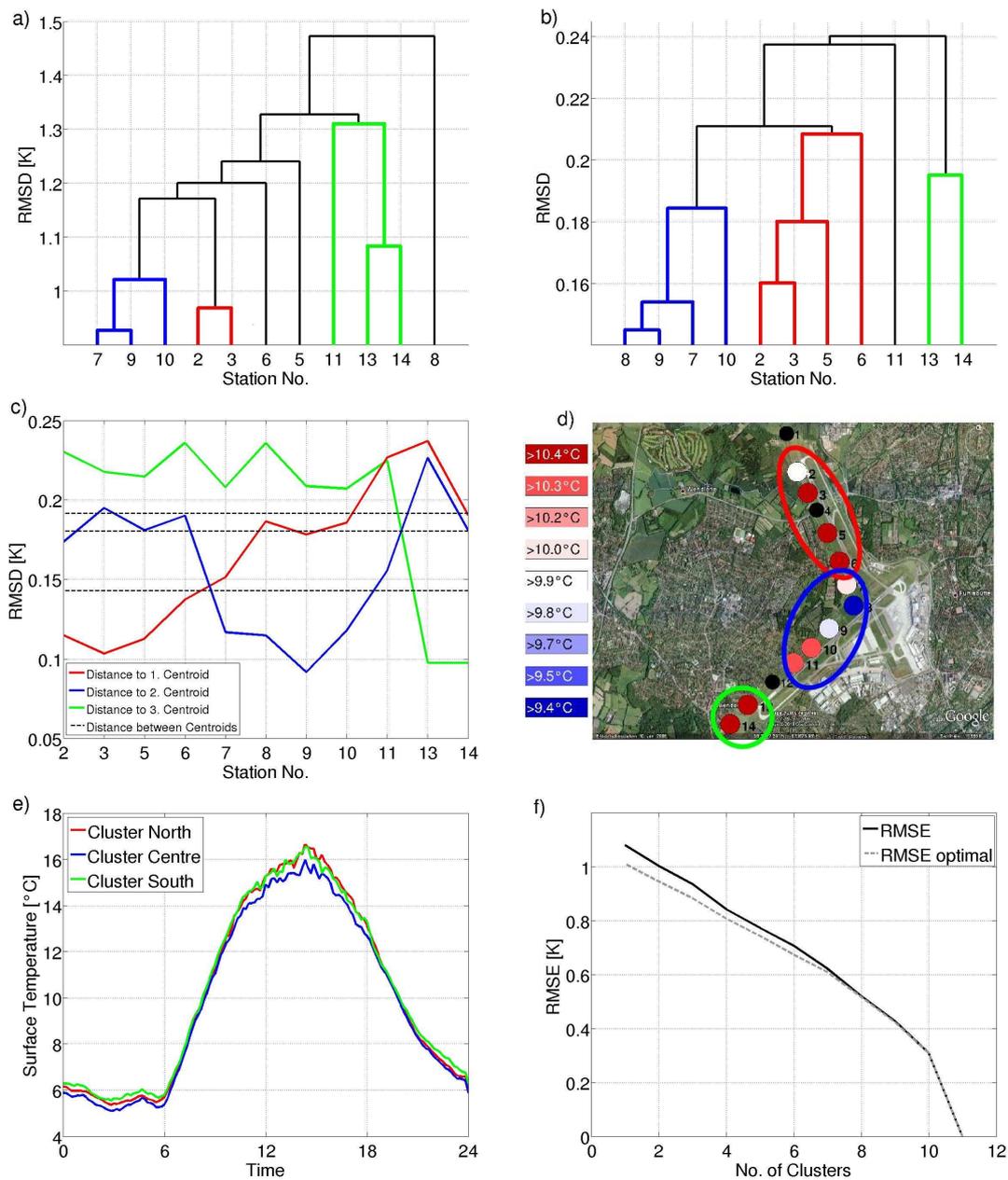


Figure 5.15: Same as Fig. 5.13, but for surface temperature.

noon, it is up to 0.8 K colder than the other two clusters. This is also evident for the air temperature in Fig. 5.13e. Another relationship between surface and air temperature is that cluster *North* is colder than cluster *South* in the night. In contrast, cluster *Centre*, which is warmest for nocturnal air temperature, shows the coldest surface temperatures. The deep soil at the centre of the airport might still be colder from the winter because of relatively high soil moisture.

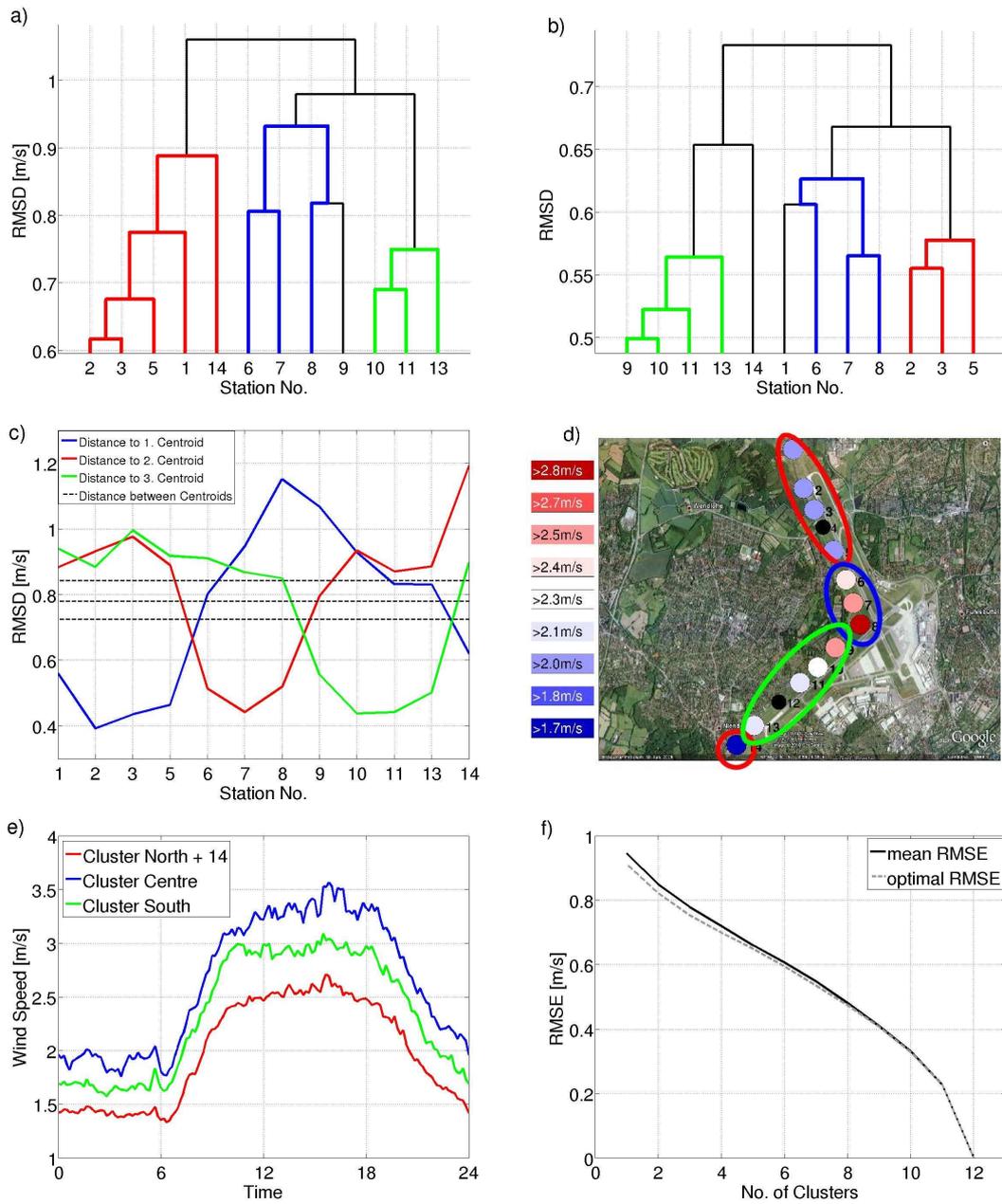


Figure 5.16: Same as Fig. 5.13, but for wind speed.

In Fig. 5.15f, the mean and optimal RMSE for using one to eleven stations is shown. The mean RMSE for only one station is 1.08 K, the optimal RMSE for using only station 07 is 1.01 K. Having one station in each of the three clusters *North*, *Centre* and *South* reduces the mean RMSE to 0.94 K. Picking the most representative stations 05, 11 and 13 or 14 for each cluster, the optimal RMSE amounts to 0.88 K. Adding more stations reduces the RMSEs by 0.08 K per station.

Fig. 5.16a shows the results of the Average Linkage Analysis for one-minute wind speed measurements. As for air temperature, humidity and surface temperature, the division into clusters is based on the geographical position of the stations. The adjacent stations 02 and 03 are most alike. They merge at a RMSD of 0.62 m s^{-1} . At a RMSD of 0.67 m s^{-1} , station 05 is added to this group. All three stations are located at the northern part of the north-south runway and are shielded from the wind on the eastern and western side. Therefore, they display similar behaviour. The northernmost station, station 01, is also assigned to this cluster. This station is additionally shielded on the northern side. A second cluster contains stations 06 and 07 and stations 08 and 09, which are merged at $\text{RMSD} = 0.94 \text{ m s}^{-1}$. These four stations are located at the centre of the airport, where the wind speeds are highest (see Chapter 5.2). Another group contains stations 10 to 13, which stand along the east-west runway and are shielded from the wind on the northern and southern side. Only station 14 does not fit into the concept of geographical clustering. It is also added to the cluster of stations 01 to 05 at a RMSD of 0.89 m s^{-1} . A possible reason is that the site at station 14 is surrounded by trees not only on the northern and southern side, but also on the western side.

The hierarchical cluster analysis of normalised one-minute means of wind speed in Fig. 5.16b also indicates a division based on the geographical position of the stations, but it differs from Fig. 5.16a. Stations 09 and 14 are added to the group of stations 10 to 13, so that all stations along the east-west runway are in one group. Station 01 in the northernmost part of the airport forms one group with the central stations 06 to 08. The third cluster contains stations 02 to 05.

According to the results of the Average Linkage Method, a subdivision into three clusters seems reasonable. In Fig. 5.16c, the results of the k-means analysis for three clusters is presented. It differs slightly from the results of the Average Linkage Method in Fig. 5.16a. Again, stations 01 to 05 form one group together with station 14. Station 09 is added to the cluster of stations 10 to 13 as for the Average Linkage Method for normalised values and not to the cluster of stations 06 to 08. This group division seems to be explicable because of the environment and wind shielding. Therefore, it is used in the following.

On the map in Fig. 5.16d, the three groups are marked in different colours: Cluster *North+14*, consisting of stations 01 to 05 and station 14, is circled in red. The blue framed stations 06 to 08 belong to cluster *Centre*. Cluster *South* is circled in green and contains stations 09 to 13.

All three clusters have the same diurnal cycles with a maximum in the afternoon and a minimum in the early morning (see Fig. 5.16e). Cluster *North+14* has the lowest wind speeds all day. It ranges from 1.4 m s^{-1} to 2.6 m s^{-1} . At cluster *South*, the wind speed is higher than at cluster *North+14*. The minimum in the morning is 1.6 m s^{-1} , the maximum in the afternoon reaches wind speeds up to 3 m s^{-1} . The highest wind speeds throughout the day occur at cluster *Centre* in the middle of the airport. Here the wind speeds are between 1.8 m s^{-1} and 3.6 m s^{-1} .

The mean and the optimal RMSE using one to twelve stations is shown in Fig. 5.16f. If only one station is used, the mean RMSE amounts to 0.94 m s^{-1} . The most representative site in terms of wind speed is station 02. Using this station reduces the RMSE to

0.91 m s^{-1} . Having one station in each of the three clusters *North+14*, *Centre* and *South* lowers the mean RMSE to 0.79 m s^{-1} and the optimal RMSE to 0.78 m s^{-1} . The most representative stations are number 02, 08 and 11. Adding more stations reduces the RMSE by 0.06 m s^{-1} per station.

5.5 Comparison with DWD Data

A climate reference station is operated by the DWD at the airport. It is located in the middle of the airport 100 m from station 08 and observes air temperature with a PT100, relative humidity with a HMP45 (campbellsci.com, 2011c) and wind speed with an anemometer made by *Thies* (thiesclima.com, 2011). Specific humidity is calculated from air temperature, relative humidity and pressure measurements. Wind speed measurements are taken at a height of 10 m at both ends of the runways and at a third station in the eastern part of the airport. Depending on which runway is used for starting and landing, wind speeds are taken from one of these sites. However, the 10 m wind speed is expected to be relatively homogeneous for the whole airport. Station 08 is located in the least sheltered area. Therefore, the conversion of the 10 m wind speed to 2 m wind speed is most reliable at this station. To make these observations comparable to the network's wind speed, the following formula based on roughness length z_0 is used to calculate the wind speed at a height of z m above ground under the assumption of neutral stratification:

$$v(z) = v_* \cdot \ln\left(\frac{z}{z_0}\right), \quad (5.1)$$

with the friction velocity v_* .

This leads to an equation for wind speeds at heights of 2 and 10 m:

$$\frac{v(2m)}{v(10m)} = \frac{\ln\left(\frac{2m}{z_0}\right)}{\ln\left(\frac{10m}{z_0}\right)}. \quad (5.2)$$

For wind speed at a height of 2 m, the following equations applies:

$$v(2m) = v(10m) \cdot \frac{\ln\left(\frac{2m}{z_0}\right)}{\ln\left(\frac{10m}{z_0}\right)}. \quad (5.3)$$

The reference station is located on grassland. Pielke (1984) gives a roughness length for grasslands with short to long blades of grass between 0.3 cm and 10 cm. In the following, a roughness length of 1 cm is used.

The DWD reference station gives ten-minute means of air temperature, humidity and wind speed. In this chapter, the measurements of the DWD station are first compared to ten-minute means of station 08. The distributions of air temperature, specific humidity and wind speed measurements are analysed in terms of mean value, standard deviation, skewness and kurtosis. Furthermore the DWD station is compared to ten-minute means of all network stations to address the question: Is the DWD station representative for the whole airport area?

Comparison to network station 08

The left column of Fig. 5.17 displays the scatterplots of air temperature, specific humidity and wind speed of the reference station and network station 08. The right column contains distributions of these meteorological conditions for both stations. To compare both distributions, the first four statistical moments are calculated:

- **Mean:** The mean μ is the expected value of a given distribution X of N elements X_i :

$$\mu = E(X) = \frac{\sum_{i=1}^N X_i}{N}. \quad (5.4)$$

- **Standard deviation:** The standard deviation σ is a measure of the variation from the mean:

$$\sigma = \sqrt{E(X - \mu)^2}. \quad (5.5)$$

- **Skewness:** The skewness s is a measure of the asymmetry of the data around the mean. If the skewness is positive the data is spread out more to the right of the mean than to the left and vice versa for negative skewness. The skewness of a distribution is defined as:

$$s = \frac{E(X - \mu)^3}{\sigma^3}. \quad (5.6)$$

- **Kurtosis:** The kurtosis k is a measure of the peakedness of a distribution. The kurtosis of the normal distribution is 3. Distributions that are more influenced by extreme outliers than the normal distribution have a kurtosis greater than 3; distributions that are less influenced by outliers have a kurtosis less than 3. The kurtosis is given by:

$$k = \frac{E(X - \mu)^4}{\sigma^4}. \quad (5.7)$$

The comparison of air temperature in Fig. 5.17a indicates that the measurements of station 08 match the reference station very well. The network station slightly underestimates the reference: the mean difference m is -0.11 K. The maximum deviation amounts to -1.8 K. For 93% of the data points, the difference between both stations is less than the network sensor's accuracy of ± 0.3 K. Even the RMSE of 0.29 K is within the accuracy. Both stations are very well correlated (correlation = 0.997).

The distributions of the air temperature measured by station 08 and by the reference station are illustrated in Fig. 5.17b in green and blue, respectively. Both have a similar shape. The distribution of station 08 is slightly shifted to the left to lower temperatures compared to the distribution of the reference station. The mean of station 08 amounts to 9.59°C, the mean of the reference is 9.7°C. Both distributions have almost the same standard deviation with 4.05°C and 4°C for station 08 and the reference station, respectively. Also skewness and kurtosis are similar for both distributions. Therefore, despite the small offset, the distributions of air temperature measured by station 08 and by the reference station are in very good agreement.

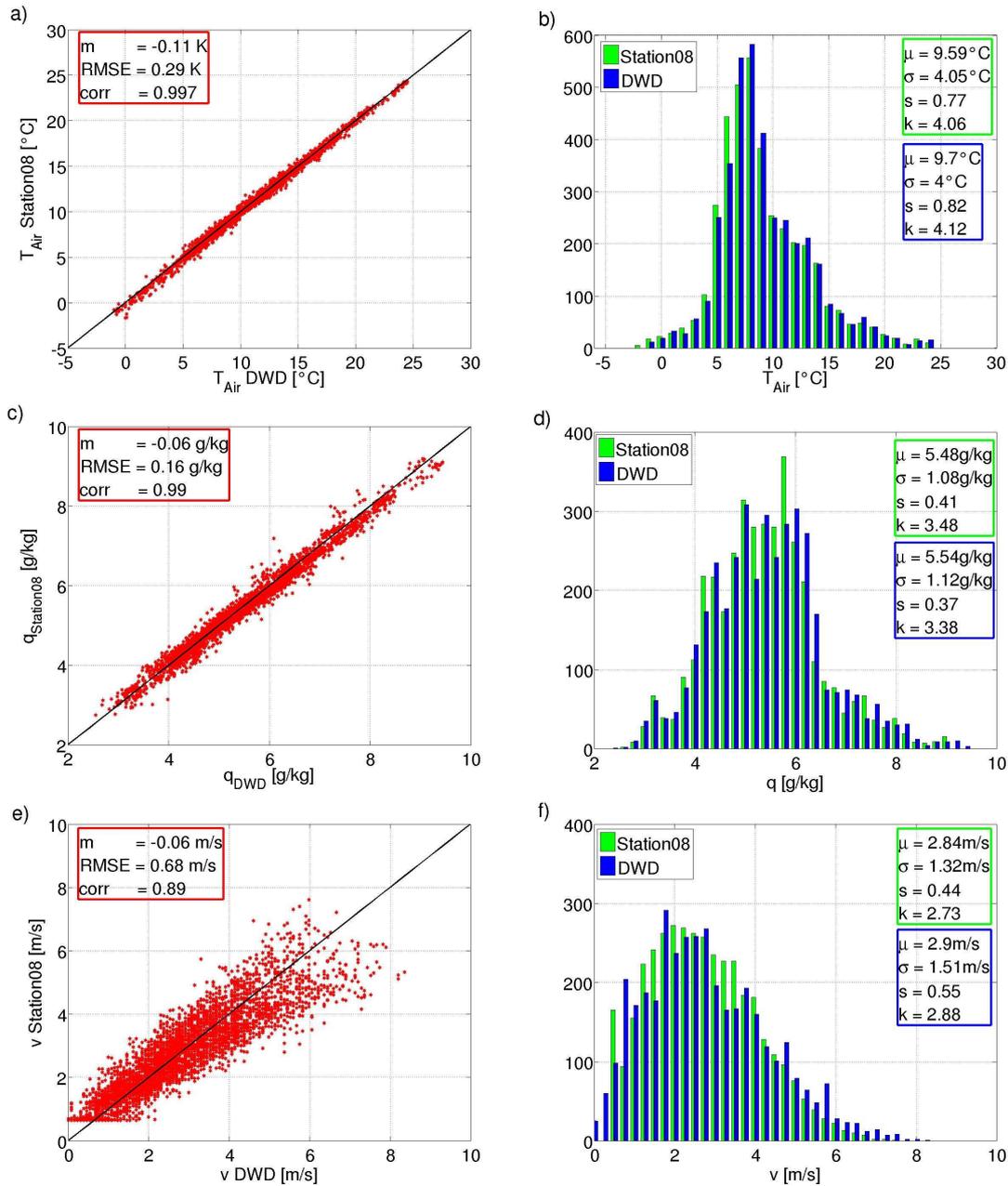


Figure 5.17: Comparison of measured ten-minute means and the distribution function of network station 08 in green and the DWD station in blue for air temperature (a and b), specific humidity (c and d) and wind speed (e and f).

For specific humidity in Fig. 5.17c, the data observed by station 08 fits the DWD data well. In 90% of the cases, the difference is less than $0.25 g kg^{-1}$; in 99% of the cases, it is less than $0.5 g kg^{-1}$. The maximum difference between station 08 and the reference amounts to $1.1 g kg^{-1}$. Station 08 underestimates the reference station in particular for specific humidities more than $7 g kg^{-1}$. This leads to a mean difference of $-0.06 g kg^{-1}$.

Reference and station 08 are very well correlated. The correlation coefficient amounts to 0.99.

In Fig. 5.17d, the distributions of specific humidity for station 08 and the DWD station and their first four statistical moments are given. Both have similar shapes, but the maximum specific humidity for the reference station is 9.4 g kg^{-1} , for station 08 it is only 9.0 g kg^{-1} . The distribution is wider for the DWD station. This is also evident from the statistical moments: The mean of the reference, 5.54 g kg^{-1} , is slightly higher than for station 08 (5.48 g kg^{-1}). Because of the wider range of specific humidity, the standard deviation is larger for the DWD station. It amounts to 1.12 g kg^{-1} . For station 08, it is 1.08 g kg^{-1} . The skewness is positive for both stations, meaning that both distributions are more spread to the right side of the mean to higher values of specific humidity. Although the reference station gives higher values of specific humidity, the skewness is marginally larger for station 08. That might be due to the sharp maximum in the distribution at 5.8 g kg^{-1} that does not exist for the reference station. Both distributions have a kurtosis of more than 3 ($k = 3.48$ for station 08 and $k = 3.38$ for the DWD station), so that they are influenced by outliers more than a normal distribution. Overall, the distributions for both stations match very well.

It is evident from Fig. 5.17e that the anemometer used by the network station cannot measure wind speeds less than 0.64 m s^{-1} . The wind sensor of the DWD station is able to measure also lower wind speeds. The data points are arranged symmetrically to the black diagonal, which represents where the measurements of both stations match. Therefore, the mean difference between station 08 and the DWD station is only 0.06 m s^{-1} . That is within the accuracy of 0.1 m s^{-1} of the network sensor. Ninety percent of the differences are less than 0.75 m s^{-1} , more than 99% are less than 1.5 m s^{-1} . For higher wind speeds, the spread around this diagonal gets larger. The maximum deviation is 3.14 m s^{-1} when the reference station gives wind speeds of more than 7 m s^{-1} . This part of the spread is due to the interpolation from 10 to 2 m height. The correlation of 0.89 is quite high although the reference data is interpolated from 10 m wind speeds.

The distributions of wind speed for the reference and station 08 in Fig. 5.17f are similar. The mean wind speed for station 08 is 2.84 m s^{-1} , for the reference station it is 2.9 m s^{-1} . The standard deviation is higher for the DWD station because it measures a wider range of wind speeds from 0 m s^{-1} to 8.25 m s^{-1} . The range of station 08 is only from 0.64 m s^{-1} to 7.5 m s^{-1} . As expected for wind speed measurements, both distributions are skewed right to higher wind speeds with skewness of 0.44 and 0.55 for station 08 and the reference station, respectively. The skewness for the reference is higher because of distinctly more data points for wind speeds of 4.75 m s^{-1} and 5.75 m s^{-1} , while station 08 has more data points around the mean. The kurtosis is less than 3 for both distributions. Station 08 is less influenced by outliers with a kurtosis of 2.73 than is the reference station with a kurtosis of 2.88. Although the wind speeds at the 2 m reference height are calculated from 10 m observations, the distributions for station 08 and the DWD station are in good agreement.

Comparison to all network stations

As described above, the distribution of ten-minute means of air temperature, specific

humidity and wind speed at station 08 and the DWD reference station are comparable. Now the reference station is compared to all stations of the sensor network. In Fig. 5.18, the mean differences between the network stations and the reference are illustrated as coloured lines and the grey area represents the 10% and 90% quantiles. The correlation between the time series of the network stations and the reference is represented by the green line.

In Fig. 5.18a, the comparison for air temperature between the network stations and the DWD station is given. Only three network stations, namely 07, 10 and 13, are warmer than the DWD station on average. Station 10 has the lowest mean difference of all network stations with 0.01 K. The farthest station, station 01, has the largest mean difference to the DWD station with -0.39 K. Stations 01 to 05 in the northern part of the airport show large negative differences compared to the reference in the middle of the airport because of strong nocturnal cooling in the rural environment. The 10% quantile is largest for the northernmost station, station 01, ($q_{10} = -1.14$ K). The quantile range is largest for station 02: The 10% quantile is -1.13 K, the 90% quantile is 0.64 K. That leads to $\Delta q_{90-10} = 1.77$ K. In Chapter 5.4, station 07 was identified as most representative of the airport area in terms of air temperature. Together with station 08, it has the lowest quantile range with 0.61 K. The mean differences to the reference station are also similar with 0.1 K and -0.11 K for stations 07 and 08, respectively. At the western end of the airport, the mean differences are less than 0.2 K, because the urban environment inhibits nocturnal cooling at the site. The quantile ranges are between 0.98 K and 1.37 K.

The correlation between the network stations and the reference station is more than 0.98 because of the strong diurnal cycle of air temperature. It is clearly lower at both ends of the runways (correlation < 0.994), than at the middle of the airport. For the central stations 06 to 09, the correlation is more than 0.996.

Overall, the reference station at the centre of the airport represents air temperature of the neighbouring stations 06 to 09 well. Therefore, it is a good representative of the airport site because station 07 was identified as most representative for all network stations. But the local features at the end of the runways should be confirmed by additional stations.

The analysis for specific humidity is given in Fig. 5.18b. The five stations which give reliable data have mean differences to the reference station of less than 0.08 g kg^{-1} . The largest difference occurs with the farthest station, station 01, the smallest at station 13. Surprisingly, station 08 next to the DWD station has the second largest mean difference of 0.06 g kg^{-1} , but the quantile range is smallest. Station 01 at the northern end of the airport has the largest negative deviation with a 10% quantile of -0.31 g kg^{-1} . It has also the largest quantile range of 0.51 g kg^{-1} . The most positive deviation shows station 05, which is located close to a water reservoir and, therefore, is wetter than the rest with $q_{10} = -0.17 \text{ g kg}^{-1}$ and $q_{90} = 0.28 \text{ g kg}^{-1}$. Station 02 was identified as the most representative site of the airport in Chapter 5.4. It is not in good agreement with the reference station having a mean difference of -0.05 g kg^{-1} and Δq_{90-10} of 0.48 g kg^{-1} .

The diurnal cycle of specific humidity leads to relatively high correlations of more than 0.97 between the network stations and the reference. It is highest for the neighbouring

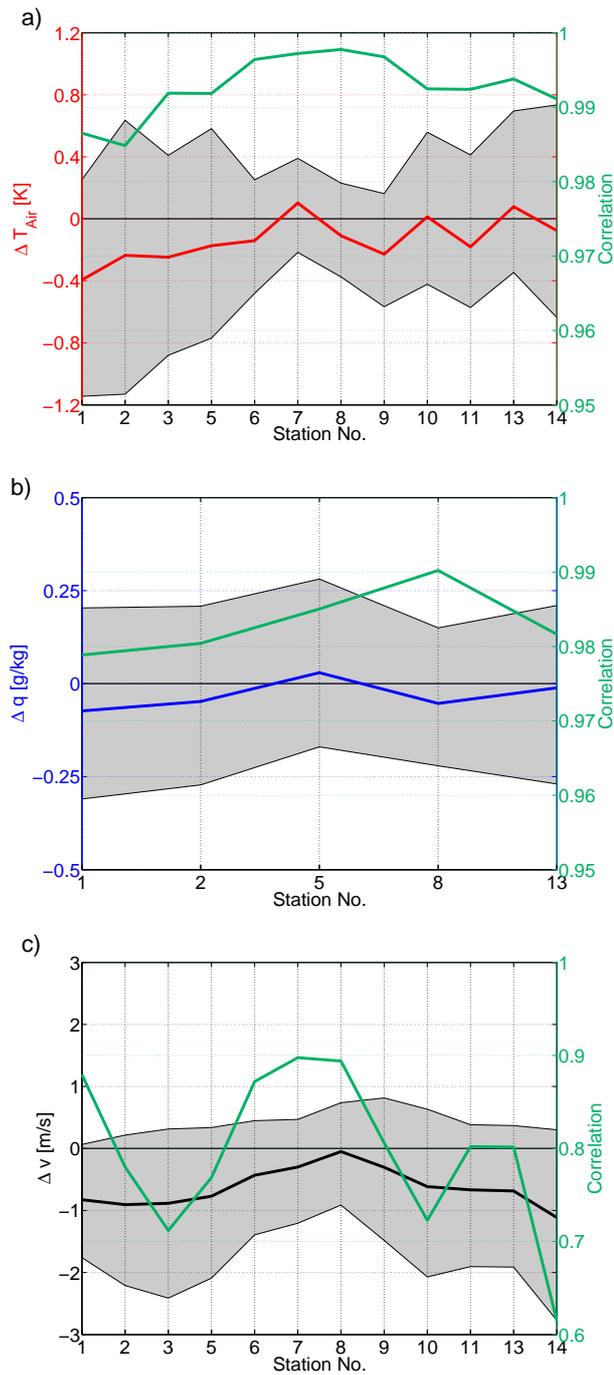


Figure 5.18: Mean difference between network stations (thick red, blue and black line) and DWD station, 10% and 90% quantile (grey area) and correlation between network stations and DWD station (green line) for ten minute means of (a) air temperature, (b) specific humidity and (c) wind speed.

station 08 with 0.99 and decreases with increasing distance to the reference station to 0.979 at station 01.

Despite the offset of 0.06 g kg^{-1} , the reference station corresponds best to station 08. However, this site is not the most representative of the airport area as stated in Chapter 5.4. Therefore, it might be useful to measure the humidity at different sites at the airport.

Wind speeds are highest at the centre of the airport (see Chapter 5.2). In Fig. 5.18c the mean difference between the network stations and the reference in the middle of the airport is negative for all sites. The interpolation of wind speeds from 10 to 2 m height might have induced this bias. The wind speed measurements at station 08 are closest to the reference with a mean difference of -0.06 m s^{-1} . The closer to the ends of the runways the stations are located, the lower is the wind speed and, hence, the larger is the mean difference. Station 14, in the most western part of the airport, differs most from the DWD station with a mean difference of -1.11 m s^{-1} . The quantile range also increases with increasing distance to the airport centre: For station 08, it is only 1.65 m s^{-1} ; for station 14, it is 3.07 m s^{-1} . Exceptions are stations 01 and 02 at the very northern end of the airport. For these two stations, the mean differences and quantile ranges decrease with increasing distance to the centre.

The correlation between the time series of wind speed for the network stations and the DWD station is larger than 0.6. Stations 07 and 08 are best correlated with the reference station with a correlation coefficient of 0.9 and 0.89, respectively. The correlation decreases with increasing distance to the centre to 0.72 for station 03 and 10. As for the percentile range, the correlation also increases for stations 01 and 02. Station 14 and the DWD station are least correlated. Because of the interpolation and the fact that the DWD observes wind speeds at three different locations at the airport it cannot be stated, whether another location would be better suited.

Chapter 6

Comparison and Conclusion

This study demonstrates that a wireless sensor network is an appropriate tool to examine small scale variabilities of atmospheric conditions near the surface. The measurements of the *SensorScope* network are not highly accurate but consistent. The mean difference between the network sensors during the calibration is 0.04 K for air temperature, 0.004 g kg⁻¹ for specific humidity, 0.13 K for surface temperature and 0.002 m s⁻¹ for wind speed. Therefore, an analysis of spatial and temporal variabilities of the atmospheric conditions is feasible.

The network has been deployed at a heterogeneous and at a homogeneous site. Air temperature, humidity, surface temperature and wind speed are observed. With the selection of the two completely different sites, the environmental influence, e.g. of surface type and land use, can be examined. A direct comparison between both sites is difficult because the campaigns took place at different times of the year and in different areas.

The first campaign took place in August 2009 in western Germany. The site was a heterogeneous area, which covered different land uses, settlement and a small river. The inter-station distance was at least 140 m, but not more than 480 m, and the transect was 2.3 km long. The second campaign took place in April/May 2010 at the airport in Hamburg, which was a very homogeneous area with only grassland. The site was much larger with 5.5 km and the inter-station distance varied between 260 m and 710 m. Because of the time of the year, it was much colder during the second campaign than during the first. The mean air temperature was 19.4°C at the FLUXPAT site and 9.6°C at the airport site. The range of daily mean temperatures at the network stations was of the same order of magnitude with approximately 10 K. Despite more rain events at the airport (six rain events compared to four during the FLUXPAT campaign), the atmosphere was drier during the airport campaign with a mean specific humidity of 5.5 g kg⁻¹, because it was colder. At the FLUXPAT site, the mean specific humidity amounted to 9.5 g kg⁻¹. The area at the airport was less sheltered from the wind than the area in western Germany. The mean wind speed at the airport was 2.26 m s⁻¹, for the FLUXPAT campaign it was only 1.65 m s⁻¹. Nevertheless, an attempt will be made to investigate the influence of heterogeneity of the surface on the meteorological conditions near the ground by answering the question raised in Chapter 1.

How large are the variabilities in air temperature, specific humidity, surface temperature and wind speed for both campaigns?

The variabilities in air temperature, specific humidity and surface temperature are up to five times higher at the heterogeneous FLUXPAT site than at the homogeneous airport site. For wind speed, the variability is in the same order of magnitude for both sites.

At Hamburg Airport, the mean air temperature at a height of 2 m ranges between 9.3°C and 9.8°C. This results in a difference of 0.5 K across a distance of 2.7 km. The variability in mean air temperature during the FLUXPAT campaign is distinctly higher: The difference between the warmest and coldest station is 0.8 K across a distance of only 0.8 km. This difference is most likely caused by formation of cold air in the river valley. Bodine et al. (2009) found differences in 2 m air temperature for single values of up to 4 K across a distance of 100 m and a height difference of 6 m.

For specific humidity, the influence of the heterogeneous surface is even more evident: The mean difference between the wettest and the driest station is 0.08 g kg⁻¹ across a distance of approximately 1.8 km at the airport. At the heterogeneous site, the variations in specific humidity are five times larger. The largest differences at the FLUXPAT site occur in the afternoon and evening. This is in agreement with findings from Fabry (2005). He investigated the variability of humidity on the mesoscale. The variability in humidity was smallest around sunrise. During the afternoon and shortly after sunset, it was twice as large.

The deviation of surface temperature is almost uniform at Hamburg Airport. The mean difference across a distance of 1.9 km amounts to 1 K. During the FLUXPAT campaign it is more than three times larger. For wind speeds, the heterogeneity of the surface does not play a significant role. For both sites, the mean difference between the stations with the highest and the lowest wind speeds is 1.1 m s⁻¹.

This study indicates that the heterogeneity of the catchment has the clearest influence on surface temperature and specific humidity. Air temperature is also affected by the diversity of the surroundings. Wind speed at a height of 2 m is not influenced by variations in surface type.

What causes the variabilities in the meteorological conditions close to the surface? Are there different explaining factors for both campaigns?

The variabilities in wind speed, specific humidity and air temperature are most likely influenced by the same factors at both sites: Wind speed is influenced by the surrounding obstacles, specific humidity is driven by nearby waters and wind speed, and air temperature is affected by the geographical position and the wind speed. However, surface temperature is mainly influenced by the heterogeneous land use at the FLUXPAT site. At the airport, the most important explaining factors are soil moisture and wind speed.

The variations in wind speed are in the same range for the homogeneous and the heterogeneous site. At both areas, the lowest wind speeds occur at stations that are surrounded by hedge banks and other obstacles which slow down the wind. Stations on open fields at the FLUXPAT site and at the centre of the airport observed the highest wind speeds. Therefore, the explaining factors for variability in wind speed are the same for both sites.

The explaining factors for air temperature are the geographical position and the wind speed at both sites, but the influence of these factors is larger in the heterogeneous area. For both sites, the variability in air temperature is twice as large for wind speeds less than 1.2 m s^{-1} than for wind speeds more than 1.2 m s^{-1} . The influence of wind speed on air temperature variability is particularly pronounced at night. This is in good agreement with Kawashima and Ishida (1992), who found a clear dependency of air temperature patterns on wind speed in summer during the night. However, during the day, the quantile range is at the homogeneous site half the quantile range of the heterogeneous site. At night, the heterogeneity of the land surface does not have an influence on the variability. The river in the FLUXPAT catchment is an important explaining factor for the variability in air temperature, too. Stations close to the river are distinctly colder than stations far away from the river.

The water-bodies mainly influence the specific humidity at both sites. In the FLUXPAT catchment, stations close to the river are wettest. At the airport, station 05 is wettest due to the neighbouring water reservoir. At the heterogeneous site, the quantile range is up to four times larger than at the homogeneous site during the day, but only two times larger during the night. The explaining factor for both sites is the wind speed. The quantile range is two times larger at wind speeds less than 1.2 m s^{-1} than at wind speeds more than 1.2 m s^{-1} .

Variations in surface temperature are not influenced by the same factors at both catchments. The differences between stations at the homogeneous airport site are caused mainly by the geographical position of the stations, wind speed and soil moisture. The coldest surface temperatures are observed at the centre of the airport. This is the area with highest wind speeds and lowest soil moisture. At the end of the runways, where the wind speed is distinctly lower, the surface temperatures are highest. During the day, the quantile ranges are largest with up to 3.2 K. The variability decreases with decreasing wind speed. At night, it is vice versa. In the heterogeneous area, systematic effects caused by different land uses dominate the variability in surface temperature. Stations that are located on surfaces where the soil is visible observe the highest surface temperatures, while stations on grassland are coldest. Similar to the homogeneous site, the variability is smaller during the night than during the day. The influence of wind speed is also the same as for the airport: During the day, the variability increases with increasing wind speeds up to maximum of 7.3 K. At night, the variability increases with decreasing wind speed.

Which sites are effectively redundant and need to be observed only once? How large is the resulting error if only parts of the network were used?

Setting up a dense network of stations is not always feasible due to costs or the catchment environment. To investigate the error evolving from using a smaller set of stations and which sites are redundant, cluster analyses are performed. Results demonstrate that for air temperature observations at a height of 2 m, the stations at the heterogeneous FLUXPAT site can be clustered according to their geographical position and, therefore, their distance to the river. Having one arbitrary station within a radius of 110 m from the

river, one at least 110 m but not more than 600 m away from the river and one outside the 600 m radius around the river leads to a RMSE of 0.7 K compared to using the full network. If only one instead of all 13 stations is used, the RMSE amounts to 0.86 K.

The clustering at the homogeneous airport site is done depending on the geographical position of the stations. Three clusters can be formed containing stations in the north, at the centre and in the south of the airport, respectively. The RMSE for using one station out of each cluster or for using only one station for the whole site is 20% smaller than for the FLUXPAT site.

For specific humidity, both sites are hardly comparable, because humidity is observed at 13 stations during the FLUXPAT campaign, but only five stations provide reliable humidity data at the airport. At the heterogeneous site, the 13 stations are clustered based on their distance to the river and their geographical position. All three stations on the western side of the river differ from all other sites. Each builds its own cluster. The stations on the eastern side of the river can be divided into two groups: Stations closer than 600 m to the river and stations farther than 600 m from the river. A set-up with one station in each of the two clusters and the three stations west of the river provides a RMSE of 0.5 g kg^{-1} . Using only one station instead of the whole network leads to a RMSE of 0.7 g kg^{-1} .

At the homogeneous catchment, the mean and optimal RMSE for having only one instead of five stations amounts to 0.4 g kg^{-1} . The clustering is dependent on the position of the stations. All stations along the north-south runway build one cluster, the station at the east-west runway builds another. Having one station at both runways reduces the RMSE to 0.35 g kg^{-1} .

The clustering for surface temperature at the heterogeneous site is due to the land use. Three clusters are suggested. The first group contains all stations that are located on grassland. The sites in the second cluster are all harvested and the bare soil is visible. In the third group, harvested corn still lies on the ground. A set-up with one station at each land use type leads to a RMSE of 1.0 K. Representing the whole catchment site with only a single arbitrary station induces a RMSE of 1.4 K.

The land use at the airport site is grassland for all stations. Therefore, the RMSE for using one station instead of all eleven is 22% smaller than for the heterogeneous site. The stations are clustered according to their geographical positions. They are divided into three groups containing the stations in the northern part, at the centre and in the southern part of the airport. This reduces the RMSE to 0.9 K for a random station out of each cluster.

It has already been pointed out that the heterogeneity of the surface does not influence the wind speed at a height of 2 m. Hence, the clustering for both sites depends on the surroundings. Three groups are formed for both sites. The first group consists of stations that are surrounded by trees and hedge banks and observe low wind speeds. The second group contains stations with moderate wind speeds, which are less sheltered from the wind than the first group. The stations in the third group are exposed most to the wind and provide the highest wind speeds. In the heterogeneous catchment, the RMSE for observing one site in each of these groups is 0.7 m s^{-1} . For the homogeneous airport, it is even 18% larger. Also, using only one station leads to slightly higher RMSEs for the homogeneous site.

The DWD operates a climate reference station at the centre of the airport in Hamburg close to station 08. It has been shown that the DWD measurements of air temperature and specific humidity are in good agreement with the observations of station 08. Additionally wind measurements are taken at the ends of both runways and in the east of the airport. Although the wind speed data has to be interpolated from a height of 10 m to 2 m, it is also in good agreement with station 08. The air temperature, specific humidity and wind measurements are compared to the observations of the full sensor network to answer the following question:

Are the observations at the centre of the airport representative for the whole airport area?

The climate reference station represents the centre of the airport very well, but for observing the meteorological conditions at the whole airport, additional stations are needed. For air temperature, the reference station represents the network stations 06 to 09 at the centre of the airport very well. The mean differences between network and reference station is smaller than the sensor uncertainty of 0.3 K and the correlation is more than 0.996. For local features, e.g. the formation of cold air in the northern part of the airport, additional stations at the end of both runways should be installed.

The DWD station corresponds best to network station 08 in terms of specific humidity. They have a correlation coefficient of 0.99. However, station 02 was identified as the most representative station of the airport site. Therefore, specific humidity should be observed at several locations at the airport.

Station 02 is also the most representative of the airport area with respect to wind speed, but the measurements of the DWD are again closest to station 08 with a correlation coefficient of 0.89 and a mean difference of -0.06 m s^{-1} . All other stations underestimate the reference measurements more. Because of the fact that the reference measurements are made at three different locations and interpolated from wind speed at a height of 10 m, it is impossible to make a statement about the quality of the location.

In this study, it has also been proven that a wireless sensor network is a useful tool to validate simulations of microscale models. The simulation of air temperature and specific humidity of the mesoscale atmospheric model FOOT3DK for level 01 (ground level) and level 02 (10 m height) are compared to the network's measurements of the FLUXPAT campaign at a height of 1.5 m and the questions asked in Chapter 1 are addressed.

How accurate are the simulations of air temperature and humidity of the FOOT3DK model? Is it possible to simulate small scale variabilities with this model?

The FOOT3DK model is only able to simulate either temperature or humidity satisfactorily at level 01 dependent on the percentage of vegetation in the grid box and the LAI. The simulations at level 02 match the observations very well.

The FOOT3DK simulations of air temperature for August 18th, 2009, at level 02 are in good agreement with the observations, although there is a height difference of 8.5 m. The mean difference between measurements and model does not exceed 0.4 K and the RMSE is not more than 2.7 K. As expected, the diurnal cycle is less pronounced in the model, because the atmosphere is more thoroughly mixed at 10 m than at the measuring

height of 1.5 m. For level 01 close to the surface, the differences between simulations and observations do not exceed 3.5 K during the night for all stations. During the day, however, the simulations match the measurements only at sites with a high percentage of vegetation (> 60%) within the grid box and, therefore, a high leaf area index of more than 1.8. At stations with less than 30% vegetation and a LAI less than 1.1, the model overestimates the measured temperature by up to 8.3 K.

For specific humidity, it is vice versa. The simulated specific humidity during the day at level 01 shows a diurnal cycle similar to the observations for sites with low vegetation cover and LAI. At sites with high vegetation cover and LAI, the model calculates a diurnal cycle for specific humidity with only one maximum. This leads to overestimations up to 7.2 g kg^{-1} compared to the observations. The RMSE reaches values up to 3.4 g kg^{-1} . At night, the simulations are in good agreement with the measurements for all stations. The simulations at level 02 show the same diurnal cycle as the observations. Again it is less pronounced because of the mixing at a height of 10 m. With mean differences not more than 0.3 g kg^{-1} and RMSEs not more than 0.9 g kg^{-1} , the simulations at level 02 fit the network measurements comparably well.

Another possible source of error is the soil moisture, that is set to values between 18% and 23% for each grid box depending on the percentage of vegetation. The measurements of the network show that sites with high vegetation cover observe soil moisture within the given range, but at sites with low vegetation cover the soil moisture is underestimated.

To simulate sensible and latent heat flux many models apply the so-called mosaic approach. This approach considers heterogeneity in vegetation, but neglects it with respect to the atmospheric forcing. To investigate the influence of heterogeneity in forcing and vegetation, the TERRA model is driven with both heterogeneous and homogeneous forcing and vegetation.

How important are heterogeneous forcing and vegetation for the simulation of sensible and latent heat fluxes?

Compared to the run with heterogeneous forcing and vegetation, the simulation with homogeneous forcing and vegetation underestimates the sensible heat flux by 9 W m^{-2} on average. For single stations, the difference ranges between -27 W m^{-2} and 16 W m^{-2} . The mean RMSE between the homogeneous and the heterogeneous run is 33 W m^{-2} . For latent heat, the homogeneous simulation overestimates the heterogeneous run by 16 W m^{-2} (between -0.5 W m^{-2} and 32 W m^{-2} for single stations) and the mean RMSE amounts to 66 W m^{-2} . Using heterogeneous vegetation as in the mosaic approach leads to more uniform results. Although the mean differences and RMSEs are in the same range as for the completely homogeneous run, the differences for single stations are only in a range of 30 W m^{-2} for sensible and 10 W m^{-2} for latent heat flux. Using heterogeneous forcing but homogeneous vegetation clearly reduces the RMSE by 36% for sensible heat flux. For latent heat flux the RMSE is reduced by 12%. This corresponds to Schomburg et al. (2010). They found a clear reduction of the error in heat flux simulations by using a downscaling scheme for the atmospheric variables.

Chapter 7

Outlook

For future campaigns, the communication and reliability of the wireless sensor network needs to be improved. That would allow for long term measurements to prove the findings of this study. Also an investigation of seasonal dependency of near surface atmospheric variability on land use would be interesting. Deploying two sensor networks to observe a homogeneous and a heterogeneous area simultaneously would be reasonable. This would enable a direct investigation of the influence of heterogeneous surface cover on atmospheric conditions.

For the FOOT3DK model, it has been shown that wireless sensor networks are helpful in validating model simulations because they allow for more accurate estimates of grid box means. A sensor network observes atmospheric conditions, like air temperature, humidity, surface temperature, wind, solar radiation and precipitation in a high spatial and temporal resolution. With these precise datasets, an independent validation of micro- and mesoscale models, which consider the surface energy balance, is feasible.

Information about the small scale variability of near surface conditions is becoming increasingly important for models. Recent turbulence parametrisation scheme applications are introduced, that consider the turbulent potential energy in addition to the turbulent kinetic energy (Mauritsen et al., 2007). The turbulent potential energy depends on temperature variabilities, which are now directly observable by the sensor network. Many models use mosaic or tiling approaches, because resolving the surface heterogeneity by refining the model grid is computationally too expensive. In these approaches, the assumption is made that atmospheric variability can be neglected. Only the land surface needs to be resolved at a higher resolution. Ament and Simmer (2006) investigate details of such parameterisations. The study examines whether it is sufficient to subdivide the atmospheric grid boxes into fractions of different land uses as is done for the tiling approach (Avissar and Pielke, 1989) or if the explicit subgrid of the mosaic approach (Seth et al., 1994) is necessary. Schomburg et al. (2010) suggest considering the atmospheric variability through a statistical scheme.

Since this study reveals that surface temperature variability is driven by the land cover type, the tiling approach is supported. Future analyses and applications of the sensor network will provide new observational guidance for the development of parameterisations

for exchange processes of heterogeneous land surfaces.

The observations in this study point out some curiosities which could not be completely explained. One example is the behaviour of the specific humidity in the night. At stations where formations of cold air occur at night, the specific humidity is surprisingly lower than at warmer sites. Another undeclared issue in this study is the increasing variability in specific humidity with increasing wind speed. It would be interesting to find out whether models can reproduce these features. This might then provide an explanation.

Appendix A

Regression Coefficients

Table A.1: Regression coefficients for air temperature correction

Sensor No.	a	b [°C]
01	-0.001292	-0.0547
02	-0.001782	0.0175
03	-0.001533	-0.3072
04	-0.001459	-0.2420
05	-0.001570	-0.0852
06	-0.001523	0.0134
07	-0.001221	-0.0769
08	-0.001506	0.2149
09	-0.001446	0.0476
10	-0.001502	0.2076
11	-0.001375	-0.0138
12	-0.001552	0.0780
13	-0.001460	-0.1320
14	-0.001554	-0.0438
15	-0.001661	-0.0895
16	-0.002182	0.1548
17	-0.001890	0.1941
18	-0.001537	-0.1282
19	-0.001584	-0.0217
20	-0.001688	0.0975

Table A.2: *Regression coefficients for relative humidity correction*

Sensor No.	a	b	c	d [%]
01	$1.30 \cdot 10^{-4}$	$-2.22 \cdot 10^{-2}$	0.9362	-4.9065
02	$-8.92 \cdot 10^{-5}$	$2.01 \cdot 10^{-2}$	-1.5237	38.1441
03	$-8.03 \cdot 10^{-5}$	$1.84 \cdot 10^{-2}$	-1.4377	37.3855
04	$-7.56 \cdot 10^{-5}$	$1.77 \cdot 10^{-2}$	-1.4062	37.1946
05	$-8.04 \cdot 10^{-5}$	$1.85 \cdot 10^{-2}$	-1.4255	36.9903
06	$-8.13 \cdot 10^{-5}$	$1.88 \cdot 10^{-2}$	-1.4537	37.6744
07	$-6.43 \cdot 10^{-5}$	$1.51 \cdot 10^{-2}$	-1.1907	31.4288
08	$-8.94 \cdot 10^{-5}$	$2.02 \cdot 10^{-2}$	-1.5355	39.1185
09	$-3.84 \cdot 10^{-5}$	$1.03 \cdot 10^{-2}$	-0.9179	26.1817
10	$-1.07 \cdot 10^{-4}$	$2.29 \cdot 10^{-2}$	-1.6639	40.1178
11	$-1.83 \cdot 10^{-4}$	$3.62 \cdot 10^{-2}$	-2.4641	56.4975
12	$-7.55 \cdot 10^{-5}$	$1.77 \cdot 10^{-2}$	-1.3587	34.5732
13	$-8.76 \cdot 10^{-5}$	$1.96 \cdot 10^{-2}$	-1.5062	38.5740
14	$-1.04 \cdot 10^{-4}$	$2.28 \cdot 10^{-2}$	-1.6814	41.2810
15	$-9.55 \cdot 10^{-5}$	$2.15 \cdot 10^{-2}$	-1.6126	39.9552
16	$-1.22 \cdot 10^{-4}$	$2.63 \cdot 10^{-2}$	-1.8502	42.7502
17	$-1.20 \cdot 10^{-4}$	$2.53 \cdot 10^{-2}$	-1.7711	40.8398
18	$-1.03 \cdot 10^{-4}$	$2.26 \cdot 10^{-2}$	-1.6291	38.9308
19	$-9.99 \cdot 10^{-5}$	$2.17 \cdot 10^{-2}$	-1.5765	38.1903
20	$-1.33 \cdot 10^{-4}$	$2.73 \cdot 10^{-2}$	-1.8997	43.8097

Table A.3: *Regression coefficients for wind speed correction*

Sensor No.	a	b [m s^{-1}]
01	0.9655	0.4695
03	0.9658	0.4884
05	0.9743	0.4823
06	0.9776	0.4427
07	0.9571	0.5851
08	0.9227	0.6006
09	0.9738	0.4603
10	0.9241	0.6418
11	0.9514	0.5822
12	0.9251	0.6611
13	0.9443	0.5452
14	0.9658	0.4703
15	0.9179	0.6954
16	0.9868	0.4470
17	0.9672	0.4244
18	0.9515	0.5773
19	0.9688	0.5287
20	0.9259	0.6159

Table A.4: *Regression coefficients for solar radiation correction*

Sensor No.	a	b [W m^{-2}]
01	1.0239	4.4163
02	0.9887	4.5248
04	1.0457	4.9712
05	1.0487	2.7778
06	0.9922	7.8823
07	1.0019	6.3329
08	1.0181	0.7744
09	0.9946	5.9536

Table A.5: *Regression coefficients for rain rate correction*

Sensor No.	c
01	1.0095
02	0.9843
03	1.0361
04	1.0095
05	0.9843
07	1.1249
08	1.1249
09	1.0361

Table A.6: *Bias for soil moisture and temperature correction*

Sensor No.	$\Delta(\text{Moisture})$ [%]	$\Delta(\text{Temperature})$ [°C]
01	0.0107	0.1615
02	0.5283	0.0809
03	1.5341	-0.1410
04	-1.4123	-0.0557
05	-2.3391	0.1238
07	-1.8024	0.2144
08	0.3818	0.1497
09	1.9664	-0.3671
10	1.1331	-0.1666

Appendix B

FLUXPAT2009

Table B.1: Sensors of the reference stations made by Campbell Scientific

Sensor Name	Reference
HMP45C Temperature and Humidity Probe	campbellsci.com (2011c)
IR120 Infra-red Remote Temperature Sensor	campbellsci.com (2011d)
WindSonic Ultrasonic Wind Sensor	campbellsci.com (2011e)
52203 Tipping Bucket Raingauge	campbellsci.com (2011a)
CS106 Barometric Pressure Sensor	campbellsci.com (2011b)

Table B.2: Leaf area index, plant cover, roughness length, root depth, minimum and maximum stomatal resistance and albedo for the four different vegetation types used in the TERRA model

	No Vegetation	Soil Visible	Soil Covered	Grassland
LAI_{Min}	1	1	1.5	2
LAI_{Max}	1	1	1.5	2
$PlCov_{Min}$	0	0.05	0.3	0.95
$PlCov_{Max}$	0	0.05	0.3	0.95
z_0 [m]	0.005	0.01	0.02	0.15
$RootDepth$ [m]	–	0.1	0.1	0.25
$RStom_{Min}$ [s m ⁻¹]	150	150	150	150
$RStom_{Max}$ [s m ⁻¹]	3000	3000	3000	3000
$VegAlb$	0.15	0.15	0.15	0.15

Appendix C

Technical Problems

During the two measurement campaigns, I experienced multiple technical problems. The availability of the loggers and the data is shown in Table C.1. During the first campaign within the FLUXPAT2009 project, battery failure occurred at two stations (logger 1092 and

Table C.1: Availability of loggers during the two experiments during FLUXPAT2009 and at Hamburg Airport in 2010: Logger was available (green checkmark), logger was available available but not the whole time (blue checkmark), logger was not used (black dash). Checkmarks in parantheses stand for months where the not all station sensors are available.

Logger No.	FLUXPAT2009	Hamburg Airport 2010
1085	—	(✓)
1086	—	—
1087	✓	(✓)
1088	—	✓
1089	✓	✓
1090	(✓)	(✓)
1091	✓	(✓)
1092	✓	—
1093	✓	(✓)
1094	✓	(✓)
1095	✓	(✓)
1096	✓	(✓)
1097	✓	—
1098	✓	—
1099	✓	(✓)
1100	✓	✓
1101	✓	✓
1102	✓	(✓)
1186	—	—

logger 1102). At station 03 (logger 1090), bugs blocked the tipping bucket so that the precipitation measurements were erroneous and the wind vane always gave wind from northward direction. During the campaign at the airport in Hamburg, the station with loggers 1094 and 1099 had battery problems and broke down ahead of schedule. At 9 out of the 14 stations, the humidity measurements were unusable due to corrosion of the sensors. At station 01 (logger 1090), the surface temperature measurements were erroneous. After May 20th, 2010, communication between the stations broke down. All attempts to keep the network running again failed, although the distance between the stations was clearly smaller than the given maximum communication range. Therefore, in October 2010 the network was rearranged and all 14 stations were set up along the north-south runway. Thus, the inter-station distance was reduced by half. The availability of loggers and measurements within this new set up is listed in Table C.2. Despite the smaller distances between the stations, communication was still unsteady and it became impossible to continue running the whole network. In August 2010, the defective temperature and humidity sensors were replaced by new ones; but until the end of the campaign in November 2011, no more than 9 stations operated simultaneously. The variety of technical problems limited the period of usable measurements to approximately seven weeks in this study.

Table C.2: Availability of loggers at Hamburg Airport between October 2010 and November 2011: Logger was available (green checkmark), logger was available available but not the whole time (blue checkmark), logger was not available (red cross), logger was not used (black dash). Checkmarks in parentheses stand for months where the not all station sensors are available.

Logger No.	2010			2011										
	10	11	12	01	02	03	04	05	06	07	08	09	10	11
1085	(✓)	x	x	(✓)	(✓)	(✓)	—	(✓)	(✓)	(✓)	✓	x	x	x
1086	—	—	—	—	—	—	—	—	—	—	—	—	—	—
1087	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	✓	✓	✓
1088	(✓)	(✓)	x	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	x	x	x
1089	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	✓	✓	✓
1090	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	x	x	✓	x	x	x
1091	(✓)	x	x	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	✓	✓	x
1092	—	—	—	—	—	—	—	x	x	x	—	—	—	—
1093	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	(✓)	(✓)	(✓)
1094	(✓)	(✓)	x	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)
1095	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	—	—	—	—	—	—
1096	(✓)	(✓)	(✓)	—	—	(✓)	(✓)	(✓)	(✓)	x	(✓)	(✓)	(✓)	(✓)
1097	—	—	—	—	—	—	—	—	—	—	—	—	—	—
1098	—	—	—	—	—	—	—	—	—	—	—	—	—	—
1099	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	x	✓	✓	x	x
1100	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	(✓)	✓	✓
1101	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	✓	✓	✓
1102	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	✓	✓	✓
1186	—	—	—	—	—	—	—	—	—	—	✓	x	x	✓

Bibliography

- Ament, F., 2006: *Energy and Moisture Exchange Processes over Heterogeneous Landsurfaces in a Weather Prediction Model*. Dissertation, University of Bonn, urn:nbn:de:hbz:5N-08348.
- Ament, F. and C. Simmer, 2006: Improved Representation of Land-Surface Heterogeneity in a Non-Hydrostatic Numerical Weather Prediction Model. *Bound.-Layer Meteorol.*, **121**, 153–174.
- Avissar, R. and R. A. Pielke, 1989: A Parameterization of Heterogeneous Land Surfaces for Atmospheric Numerical Models and its Impact on Regional Meteorology. *Mon. Wea. Rev.*, **117**, 2113–2136.
- Barrenetxea, G., F. Ingelrest, G. Schaefer, and M. Vetterli, 2008: Wireless Sensor Networks for Environmental Monitoring: The SensorScope Experience. In: *Proc. of International Zurich Seminar on Communications*, pp. 98–101. Zurich, Switzerland.
- Bodine, D., P. M. Klein, S. C. Arms, and A. Shapiro, 2009: Variability of Surface Air Temperature over Gently Sloped Terrain. *J. Appl. Meteorol. Climatol.*, **48**, 1117–1141.
- Brücher, W., M. Sogalla, and M. J. Kerschgens, 2003: FOOT3DK-Handbuch: Modell, Tools und Klimatologiemodul. *Institut für Geophysik und Meteorologie der Universität zu Köln*.
- campbellsci.com, 2011a: 52203 Tipping Bucket Raingauge. [Available online at ftp://ftp.campbellsci.com/pub/csl/outgoing/uk/manuals/52202_52203.pdf].
- campbellsci.com, 2011b: CS106 Barometric Pressure Sensor. [Available online at http://www.campbellsci.com/documents/product-brochures/b_cs106.pdf].
- campbellsci.com, 2011c: HMP45 Humidity and Temperature Sensor. [Available online at <http://www.campbellsci.com/hmp45c-1>].
- campbellsci.com, 2011d: IR100/IR120 Infra-Red Remote Temperature Sensor. [Available online at ftp://ftp.campbellsci.com/pub/csl/outgoing/uk/manuals/ir100_ir120.pdf].
- campbellsci.com, 2011e: WindSonic Ultrasonic Wind Sensor. [Available online at <http://www.campbellsci.com/windsonic1>].
- Colton, D. E., 1976: Numerical Simulation of Orographically Induced Precipitation Distribution for Use in Hydrologic Analysis. *J. Appl. Meteorol.*, **15**, 1241–1251.

- davisnet.com, 2011a: Anemometer. [Available online at http://www.davisnet.com/product_documents/weather/spec_sheets/6410_spec_Rev_E.pdf].
- davisnet.com, 2011b: Rain Collector. [Available online at http://www.davisnet.com/product_documents/weather/manuals/Rain_Collector_II_Rev_B.pdf].
- davisnet.com, 2011c: Solar Radiation Sensor. [Available online at http://www.davisnet.com/product_documents/weather/spec_sheets/6450_SS.pdf].
- decagon.com, 2011: EC-TM Probe. [Available online at <http://www.decagon.com/products/sensors/soil-moisture-sensors/5tm-soil-moisture-and-temperature>].
- Dickinson, R. E., 1984: Modeling evapotranspiration for the three-dimensional global climate models. In *Climate Processes and Climate Sensitivity. Geophysical Monograph 29, Maurice Ewing Volume 5*, pp. 58–72.
- Dingman, S. L., 2002: *Physical Hydrology - Second Edition*. Prentice Hall.
- Doms, G., J. Förstner, E. Heise, H.-J. Herzog, M. Raschendorfer, R. Schrodin, T. Reinhardt, and G. Vogel, 2005: *A Description of the Non-Hydrostatic Regional Model LM Part II: Physical Parameterization*. Deutscher Wetterdienst, Offenbach.
- Fabry, F., 2005: The Spatial Variability of Moisture in the Boundary Layer and its Effect on Convection Initiation: Project-Long Characterization. *Mon. Wea. Rev.*, **134**, 79–91.
- Gustavson, T., M. Karlsson, J. Bogren, and S. Lindquist, 1998: Development of Temperature Patterns during Clear Nights. *J. Appl. Meteorol. Climatol.*, **37**, 559–571.
- Heise, E., B. Ritter, and R. Schrodin, 2006: *Operational Implementation of the Multilayer Soil Model - Technical Report No. 9*. Consortium for Small Scale Modelling (COSMO).
- Hubbard, K. G., 1994: Spatial Variability of Daily Weather Variables in the High Plains of the USA. *Agric. For. Meteorol.*, **68**, 29–41.
- Hunt, E. D., J. B. Basara, and C. R. Morgan, 2007: Significant Inversions and Rapid In Situ Cooling at a Well-Sited Oklahoma Mesonet Station. *J. Appl. Meteorol. Climatol.*, **46**, 353–367.
- Ingelrest, F., G. Barrenetxea, G. Schaefer, M. Vetterli, O. Couach, and M. Parlange, 2010: SensorScope: Application-Specific Sensor Network for Environmental Monitoring. *ACM Transactions on Sensor Networks*, **6**.
- Jacobsen, I. and E. Heise, 1982: A New Economic Method for the Computation of the Surface Temperature in Numerical Models. *Contributions to Atmospheric Physics*, **55**, 128–141.
- Kawashima, S. and T. Ishida, 1992: Effects of Regional Temperature, Wind Speed and Soil Wetness on Spatial Structure of Surface Air Temperature. *Theo. Appl. Climatol.*, **46**, 153–161.

- kippzonen.com, 2011: CM11 Pyranometer. [Available online at <http://www.kippzonen.com/?product/1371/CMP+11.aspx>].
- Koyama, C. N., W. Korres, P. Fiener, and K. Schneider, 2010: Variability of Surface Soil Moisture Observed from Multitemporal C-Band Synthetic Aperture Radar and Field Data. *Vadose Zone J.*, **9**, 1014–1024.
- Lengfeld, K. and F. Ament, 2011: Observing Local Scale Variability of Near Surface Temperature and Humidity Using a Wireless Sensor Network. *accepted in J. Appl. Meteorol. Climatol.*.
- Louis, J.-F., 1979: A Parametric Model of Vertical Eddy Fluxes in the Atmosphere. *Bound.-Layer Meteorol.*, **17**, 187–202.
- Mahrt, L., 2006: Variation of Surface Air Temperature in Complex Terrain. *J. Appl. Meteorol. Climatol.*, **45**, 1481–1493.
- Mauritsen, T., G. Svensson, S. S. Zilitinkevich, I. Esau, and B. Grisogono, 2007: A Total Turbulent Energy Closure Model for Neutrally and Stably Stratified Atmospheric Boundary Layers. *J. Atmos. Sci.*, **64**, 4113–4126.
- Mellor, G. L. and T. Yamada, 1982: Development of a Turbulence Closure Model for Geophysical Fluid Problems. *Rev. Geophys.*, **20**, 851–875.
- Nadeau, D. F., W. Brutsaert, M. B. Parlange, E. Bou-Zeid, G. Barrenetxea, O. Couach, M.-O. Boldi, J. S. Selker, and M. Vetterli, 2009: Estimation of Urban Sensible Heat Flux Using a Dense Wireless Network of Observations. *Environ. Fluid Mech.*, **9**, 635–653.
- Noilhan, J. and S. Planton, 1989: A Simple Parametrization of Land Surface Processes for Meteorological Models. *Mon. Wea. Rev.*, **117**, 536–548.
- Orlanski, I., 1975: A Rational Subdivision of Scales for Atmospheric Processes. *Bull. Amer. Meteorol. Soc.*, **56**, 527–530.
- Pielke, R. A., 1974: A Comparison of Three-Dimensional and Two-Dimensional Numerical Predictions of Sea Breezes. *J. Atmos. Sci.*, **31**, 1577–1585.
- Pielke, R. A., 1984: *Mesoscale Meteorological Modeling*. Academic Press, 2nd edition.
- Polastre, J., R. Szewczyk, A. Mainwaring, D. Culler, and J. Anderson, 2004: Analysis of Wireless Sensor Networks for Habitat Monitoring. In: C. Raghavendra, K. M. Sivalingam, and T. Znati, eds., *Wireless Sensor Networks*, pp. 399–423. Springer-Verlag.
- Rijtema, P. E., 1969: Soil Moisture Forecasting. Instituut voor Cultuurtechniek en Waterhuishouding, Wageningen.
- Schomburg, A., V. Venema, R. Lindau, F. Ament, and C. Simmer, 2010: A Downscaling Scheme for Atmospheric Variables to Drive Soil-Vegetation-Atmosphere Transfer Models. *Tellus Series B-Chem. and Phys. Meteor.*, **62**, 242–258.

- Schrodin, R. and E. Heise, 2001: *The Multi-Layer Version of the DWD Soil Model TERRA LM - Technical Report No. 2*. Consortium for Small Scale Modelling (COSMO).
- semtech.com, 2011: Semtech XE1205 Radio Transceiver. [Available online at <http://www.semtech.com/images/datasheet/xe1205.pdf>].
- sensirion.com, 2011: SHT75 - Digital Humidity Sensor (RH & T). [Available online at http://www.sensirion.com/en/01_humidity_sensors/06_humidity_sensor_sht75.htm].
- Seth, A., F. Giorgi, and R. E. Dickinson, 1994: Simulating Fluxes from Heterogeneous Land Surfaces - Explicit Subgrid Method Employing the Biosphere-Atmosphere Transfer Scheme (BATS). *J. Geophys. Res.- Atmospheres*, **99**, 18651–18667.
- Shapiro, A., P. M. Klein, S. C. Arms, D. Bodine, and M. Carney, 2009: The Lake Thunderbird Micronet Project. *Bull. Amer. Meteorol. Soc.*, **90**, 811–823.
- Swinbank, W. C., 1963: Long Wave Radiation from Clear Skies. *Quarterly J. Royal Meteorol. Soc.*, **89**, 339–348.
- thiesclima.com, 2011: Thies Clima Anemometer. [Available online at <http://www.thiesclima.com/wind.html>].
- ti.com, 2011: Texas Instruments MSP430 Ultra-Low-Power Microcontrollers. [Available online at <http://www.ti.com/lit/sg/slab034u/slab034u.pdf>].
- tinynode.com, 2011: [Available online at <http://www.tinynode.com>].
- Werner-Allen, G., K. Lorincz, J. Johnson, J. Lees, and M. Welsh, 2006a: Fidelity and Yield in a Volcano Monitoring Sensor Network. In: *Proc. of 7th USENIX Symposium on Operating Systems Design and Implementation*, pp. 381–396. Seattle, WA, U.S.A.
- Werner-Allen, G., K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, and M. Welsh, 2006b: Deploying a Wireless Sensor Network on an Active Volcano. *IEEE Internet Computing*, **10**, 18–25.
- wetter3.de, 2011: [Available online at <http://www.wetter3.de>].
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 2nd edition.
- wintron.com, 2011: Selection Guide: KT19-Series II- Infrared Radiation Pyrometer. [Available online at <http://www.wintron.com/infrared/kt19iip/kt19iip.html>].
- Xu, M., J. Chen, and Y. Qi, 2002: Growing Season Temperature and Soil Moisture along a 10 km Transect Across Forested Landscape. *Climate Res.*, **22**, 57–72.
- Zemel, Z. and J. Lomas, 1976: An Objective Method for Assessing Representativeness of a Station Network Measuring Minimum Temperature Near the Ground. *Bound.-Layer Meteorol.*, **10**, 3–14.

Zhang, P., C. M. Sadler, T. Liu, I. Fischhoff, M. Martonosi, S. A. Lyon, and D. I. Rubenstein, 2005: Habitat Monitoring with ZebraNet: Design and Experiences. In: N. Bulusu and S. Jha, eds., *Wireless Sensor Networks: A Systems Perspective*, pp. 235–257. Artech House.

zytemp.com, 2011: Infrared Module Product - TN901. [Available online at <http://www.zytemp.com/products/tn901.asp>].

Acknowledgement

An dieser Stelle möchte ich mich bei allen bedanken, die mich unterstützt haben und zu dieser Dissertation beigetragen haben.

Ein besonderer Dank gilt Prof. Dr. Felix Ament, der mir die Arbeit an diesem interessanten Thema ermöglicht hat. Ich bedanke mich für die gute Betreuung, für viele hilfreiche Ideen und für die durch (fast) nichts zu trübende gute Laune.

Prof. Dr. Bernd Leitl danke ich für die zeitaufwendige Zweitkorrektur dieser Arbeit.

Ein Teil, der in dieser Arbeit verwendeten Daten, wurde im Rahmen des *Transregional Collaborative Research Centre (SFB/TR) 32* Projektes *Patterns in Soil-Vegetation-Atmosphere-Scheme: Monitoring, Modeling and Data Assimilation* gewonnen. Ich danke den Organisatoren für die Möglichkeit an diesem Projekt teilzunehmen. Insbesondere haben mir Dr. Jan Schween und Dr. Alexander Graf bei der Planung und Durchführung meiner Messkampagne sehr geholfen. Außerdem bedanke ich mich beim Meteorologischen Institut der Universität Bonn für die Bereitstellung der langwelligen Strahlungsdaten.

Ich bedanke mich bei der Flughafen Hamburg GmbH, insbesondere bei Jürgen Wächtler, sowie bei Joachim Horn und Ulrike von Barga vom DWD, die die Messkampagne am Hamburger Flughafen ermöglicht haben. Beim Deutschen Wetterdienst bedanke ich mich für die Versorgung mit Daten von der Klimareferenzstation am Hamburger Flughafen.

Dem Meteorologischen Institut der Universität Köln und insbesondere Dr. Stefan Zacharias möchte ich für die Bereitstellung der FOOT3DK Simulationen danken. Bei Dr. Mark Reyers bedanke ich mich für die Beantwortung meiner Fragen zu diesem Modell.

Vielen Dank an Rainer Senke und Werner Biselli für die Hilfe bei technischen Problemen und an Ingo Lange für die Bereitstellung der Wettermast-Daten und die Hilfe bei der Sensorkalibrierung.

Ein großer Dank geht an alle aktuellen und ehemaligen Mitglieder der Arbeitsgruppe *Synthese von Beobachtungen und Modellen* für die angenehme Arbeitsatmosphäre, Hilfe bei wissenschaftlichen und programmiertechnischen Fragen, Unterstützung bei Aufbau und Wartung des Sensornetzwerks und für Gespräche über "Löwenbabies" nach Feierabend.

Bei Peter Hoffmann bedanke ich mich für das richtige Maß an wissenschaftlichen und weltlichen Diskussionen während der Tee- und Mittagspausen.

Ein ganz liebes Dankeschön an Nicole Feiertag für die Diskussion und Lösung jedes noch so kleinen Problems und für die einmalige Erfahrung wie eine "Banane" zu fliegen.

Vielen Dank auch an das wandelnde Englischwörterbuch Dr. David Flagg für das Korrekturlesen dieser Arbeit und für die geduldige Beantwortung meiner Grammatikfragen.

Zum Schluss noch ein ganz besonderer Dank an meine Eltern, die mich immer moralisch (und finanziell) unterstützt haben.

Affirmation

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, Januar, 2012

(Katharina Lengfeld)