Can a simple locality index be used to improve mesoscale model forecasts?

Dissertation

Zur Erlangung des Doktorgrades der Naturwissenschaften im

Fachbereich Geowissenschaften der Universität Hamburg

vorgelegt von

Sylvia Ingrid Luise Dorothea Bohnenstengel

aus Pinneberg

Hamburg

2012

Als Dissertation angenommen vom Fachbereich Geowissenschaften der Universität Hamburg.

Auf Grund der Gutachten von Prof. Dr. K. Heinke Schlünzen
und Prof. Dr. Martin Claußen

Hamburg, den 14.12.2011

Prof. Dr. Jürgen Oßenbrügge
Leiter des Fachbereichs Geowissenschaften

Abstract

Surface fluxes couple the land surface to the atmosphere and are a source of heat, moisture and momentum for the atmosphere. They vary with land-use and atmospheric stratification and influence temperature, humidity, wind and precipitation in the atmosphere. In numerical weather and climate models land-use changes on the sub-grid scale and the characteristics are parameterised. However, in very heterogeneous areas like urban areas parameterisations can lead to unrealistic results. In this thesis a process-oriented approach is developed to evaluate two parameterisation schemes for sub-grid scale land-use effects. The model performance is evaluated for near surface atmospheric variables and the relevance of surface fluxes for the generation of precipitation in the model domain is determined.

A locality index is developed reflecting the relevance of surface fluxes for the model solution and especially the generation of precipitation in the model domain. Firstly, a precipitation analysis reveals the spatial and temporal scale of precipitation and derives uncertainty factors for the representativity of single-station precipitation amounts in the LITFASS domain in North-Eastern Germany. The index is then related to the measured precipitation. In situations with a high relevance of the surface fluxes indicated by a high index precipitation probability increases. This relation holds for two very different years and is very robust. The relation is resembled by simulated precipitation using the mesoscale transport and fluid model METRAS. However, a dependence of simulated precipitation on horizontal resolution and the parameterisation scheme is shown.

METRAS' model performance is evaluated applying the locality index for four horizontal resolutions and two parameterisation schemes namely flux aggregation with a blending height concept and parameter averaging. In very locally driven meteorological situations with high indices the surface fluxes and the horizontal resolution influence the model solution strongly and the index determines flux aggregation as the more appropriate parameterisation for the sub-grid scale surface fluxes. Parameter averaging is very resolution-dependent in contrast to flux aggregation and only resembles the results gained with flux aggregation when applying high-resolutions. In advectively driven meteorological situations the model solution depends less on the surface processes and their parameterisation. The lateral boundary conditions are more important for the model performance. Then, also parameter averaging leads to realistic model results, which is also computationally cheaper than flux aggregation. In general, the locality index can be applied in a model to choose the parameterisation scheme online and can be used to evaluate model results in a process-oriented way.

Zusammenfassung

Oberflächenflüsse koppeln die Landoberfläche mit der Atmosphäre und sind eine Quelle für Wärme, Feuchte und Impuls. Sie variieren mit Landnutzung und atmosphärischer Schichtung und verändern Temperature, Feuchte, Wind und Niederschlag in der Atmosphäre. In numerischen Wetter- und Klimamodellen ändert sich die Landnutzung subskalig and ihre Prozesse werden parameterisiert. Allerdings können Parameterisierungen für sehr heterogene Landnutzungen wie Städte zu unrealistischen Ergebnissen führen. In dieser Arbeit wird ein prozessbasierter Ansatz entwickelt, um zwei Parameterisierungen für subskalige Landnutzungseffekte zu evaluieren. Die Modelgüte wird für oberflächennahe Variable evaluiert und die Bedeutung der Oberflächenflüsse für die Entstehung von Niederschlag im Modellgebiet wird bestimmt.

Ein Lokalitätsindex wird entwickelt, der die Bedeutung der Oberflächenflüsse für die Modellösung und insbesondere die Entstehung von Niederschlag bestimmt. Zunächst werden die räumlichen und zeitlichen Skalen des Niederschlages einer Einzelmessstation anhand einer Niederschlagsanalyse bestimmt und Unsicherheitsfaktoren für die Representativität für das LITFASS Gebiet in Nord-Ost-Deutschland bestimmt. Der Index wird dann zu dem gemessenen und analysierten Niederschlag in Bezug gesetzt. In Situationen mit starkem Einfluss der Flüsse gekennzeichnet durch einen hohen Index erhöht sich die Niederschlagswahrscheinlichkeit. Diese Beziehung wird für zwei sehr unterschiedliche Jahre nachgewiesen und ist sehr robust. Die Beziehung wird ebenfalls mit simuliertem Niederschlag von dem mesoskaligen Transport- and Strömungsmodell METRAS 3.0 wiedergegeben. Allerdings ist der simulierte Niederschlag von Auflösung und Parameterisierung abhängig.

Die METRAS Modelgüte wird anhand des Lokalitätsindex für vier Auflösungen und zwei Parameterisierungen, namentlich Flussmittelung mit einem Blendhöhenverfahren und Parametermittelung, evaluiert. In sehr lokal angetriebenen Situationen mit hohem Index beeinflussen die Oberflächenflüsse und die Auflösung die Modellösung, und der Index schlägt Flussmittelung als das geeignetere Verfahren für die Parameterisierung der subskaligen Flüsse vor. Parametermittelung ist im Gegensatz zu Flussmittelung sehr auflösungsabhängig und gibt die Ergebnisse, die mit Flussmittelung erzielt wurden, nur bei hoher Auflösung wieder. In advektiven Situationen hängt die Lösung weniger von den Landoberflächenprozessen und ihrer Parametrisierung ab. Die seitlichen Grenzbedingungen sind wichtiger. Unter diesen Umständen führt auch Parametermittelung zu realistischen Ergebnissen, die rechenzeittechnisch günstiger als Flussmittelung ist. Der Lokalitatesindex kann in einem Modell online verwendet werden, um die geeignete Parameterisierung zu wählen und eignet sich um Modellergebnisse prozessbasiert zu evaluieren.

Contents

1 Introduction 1

2 Parameterisation of sub-grid scale land-use effects in numerical models 3

2.1 LITFASS measurement campaign 2003 to determine surface fluxes 4

2.2 Parameterisation of sub-grid scale surface fluxes 5

3 Sensitivity of model performance on uncertainties in initial soil temperature and humidity 15

3.1 Introduction 16

3.2 Model set-up and simulation period 17

3.3 Influence of the initialisation on the model performance for short-term forecasts 21

3.3.1 Sensitivity to surface soil temperature 21

3.3.2 Sensitivity to soil surface humidity 29

3.3.3 Sensitivity to the initialisation profile of relative humidity 31

3.4 Conclusions 31

4 Representativity of in-situ precipitation measurements – a case study for the LITFASS area in North-Eastern Germany 35

4.1 Introduction 36

4.2 Investigation area and data 38

4.3 Characteristics of precipitation events 41

4.3.1 Precipitation frequency 41

4.3.2 Duration of precipitation events 42

4.3.3 Scatter of 6 hour precipitation amounts 45

4.3.4 Spatial representativity of in-situ measurements 47

4.3.5 Temporal representativity of in-situ measurements 52

4.4 Conclusions 55

5 A locality index to investigate the influence of surface fluxes on precipitation events 58

5.1 Introduction 59

5.2 Investigation area and used data 61

5.3 Locality indices 63

5.3.1 Advection and diffusion impact: locality index I_{adv} 63

5.3.2 Diffusion impact: locality index I_{lt} 64

5.3.3 Diffusion and humidity impact: locality index I_q 65
5.3.4 Diffusion and relative humidity impact: locality index I_{rh} 65
5.3.5 Calculation of the indices 65

5.4 Locality indices and precipitation 67

5.5 Summary and Conclusions 76

6 Model performance for different sub-grid scale surface flux parameterisation schemes 79

6.1 Simulated meteorological situations 80
6.1.1 Period E1: very dry August simulation 81
6.1.2 Period E2: very humid March simulation 81
6.1.3 Period M1: humid March simulation 82
6.1.4 Period M2: humid January simulation 82
6.1.5 Period M3: very warm and dry June simulation 82
6.1.6 Period M4: very hot and very dry June simulation 82

6.2 Model evaluation method 83

6.3 Overall evaluation results 84

6.4 Evaluation of the diurnal cycle 91

6.5 Using a locality index for process oriented model evaluation 96

6.6 Conclusions 102

7 Sensitivity of mesoscale model results to parameterisation of sub-grid scale surface fluxes for a locally driven meteorological situation in the area of Berlin 104

7.1 Introduction 105

7.2 METRAS model description and urban scheme 109
7.2.1 Surface energy budget 109
7.2.2 Surface fluxes 111

7.3 Sensitivity study for different model configurations 115

7.4 Results 117
7.4.1 The spatial variability of the Bowen ratio 117
7.4.2 The diurnal cycle of the Bowen ratio 122
7.4.3 Diurnal cycle of temperature and depression of the dew point for Berlin 129
7.4.4 Model accuracy and sensitivity towards horizontal resolution 131

7.5 Conclusions 142

8 Conclusions and outlook 145

8.1 Conclusions 145

8.2 Outlook 151

Acknowledgements **152**

List of important symbols **153**

Bibliography **155**

1 Introduction

The horizontal grid resolution of atmospheric numerical models is increasing. State of the art mesoscale numerical models are currently reaching the kilometre scale and resolve the orography and the land-use characteristics more detailed. However, the landscape is often heterogeneous on spatial scales a couple of orders of magnitude smaller than the present horizontal grid resolution in numerical models. Those scales are not resolved explicitly due to the still limited computing power, although the computing power has increased significantly over the recent years. Instead, the impact of the sub-grid scale processes on the state of the boundary layer needs to be represented on the grid scale. Such so called parameterisations enable to represent the effect of sub-grid scale processes on the grid scale without actually simulating the physical mechanism explicitly but capturing its effect. In the present study the performance of sub-grid scale land-use parameterisations is investigated and their applicability range is determined.

The objective of this thesis is to derive a cost-efficient and simple locality index in order to classify meteorological situations dependent on the local impact of surface processes. This index is then intended to be used to identify the local impact of surface characteristics on the generation of precipitation. In the end it shall be used to control the surface parameterisation scheme used in models for sub-grid scale land-use heterogeneities. The rationale behind this is the assumption that the land-use plays a key role for the lowest boundary layer structure, especially in very locally driven meteorological situations, when advection is sufficiently low and the vertical coupling between the surface and the boundary layer is assumed to be strong. The situation dependent choice of the appropriate parameterisation scheme for the sub-grid scale surface fluxes aims to improve the model performance considerably. And it intends to lead to a better precipitation forecast in local meteorological situations, when the evaporation from the surface affects the generation of precipitation within the model area considerably.

The parameterisation schemes that are used in the present study are introduced in the context of other approaches in Chapter 2. Not only the parameterisation schemes for surface fluxes influence the prognostic fields like temperature and humidity in the lowest boundary layer, but the initialisation of mesoscale numerical models plays a major role for the model performance (Schlünzen and Katzfey, 2003; Ament and Simmer, 2006). The model simulations are undertaken for the very dry year 2003, when the soils in Northern Europe were already dried out during spring. This strongly affects the simulation of the whole surface energy balance and can lead to a systematic underestimation of near surface temperatures with a standard model set-up. This initialisation impact is investigated in Chapter 3.

To derive a simple locality index measurements of the area of Lindenberg in North-Eastern Germany are analysed in detail. High-resolution precipitation measurement data from Lindenberg are analysed in Chapter 4 to conclude on the characteristics of precipitation with respect to spatial and temporal representativity of in-situ measurements and their applicability for model evaluation purposes. Based on these data a locality index is developed (Chapter 5). Therefore, six hourly model results for the years 2002 and 2003 are additionally used. The two years are chosen, because they differ significantly in terms of annual precipitation amounts. 2002 was dry below average and 2003 wet above average. The locality index is then used to classify the precipitation data analysed in Chapter 4 and to identify the role surface characteristics play for the generation of precipitation (Chapter 5).

The impact of different sub-grid scale surface flux parameterisations, namely a parameter averaging approach and a flux aggregation approach with blending height concept is investigated for four different horizontal resolutions (Chapter 6) and different meteorological situations. The appropriate parameterisation scheme for sub-grid scale surface fluxes is determined dependent on the meteorological situation characterised by the locality index (Chapter 6). The locality index is then used for a process-oriented evaluation of simulated precipitation events (Section 6.5). The relationship of index and precipitation as derived from measured precipitation data is also determined for simulated precipitation.

In Chapter 7 the model performance is evaluated for the extreme case of an urban area, where small scale heterogeneities lead to very distinct sub-grid scale surface characteristics in order to test the limits of the applicability of the sub-grid scale parameterisation schemes. Overall conclusions including an outlook are drawn in Chapter 8.

This thesis partly includes submitted scientific papers or the chapters are prepared for submitting to a journal. Therefore, some parts are repeated in several chapters and the nomenclature might differ in different chapters. If this is the case this is mentioned at the beginning of a chapter.

2 Parameterisation of sub-grid scale land-use effects in numerical models

The atmospheric equations forming an atmospheric numerical model are averaged over time and space, since today's computing capacities are not sufficient to solve the atmospheric equations explicitly on the spatial and temporal scales required. Therefore, the atmospheric variables are decomposed into two parts; a first part representing the average over a time and grid volume interval and a second part representing the deviation from the time and volume average (Stull, 1988). The impact of the second terms on the grid-scale variables needs to be described in terms of simplified, mostly empirical relationships - they need to be parameterised. A parameterisation of a process does not necessarily have to describe the actual physical sub-grid scale process but its overall impact on the grid-scale variables. Parameterisation concepts within atmospheric numerical models serve several different purposes. They describe the turbulent surface fluxes of momentum, heat and moisture as well as the radiative flux divergence and cloud microphysics. Within this study only the parameterisation of the vertical turbulent surface fluxes and their dependency on the horizontal resolution are investigated.

The volume-averaged and time averaged equations for momentum, heat and moisture are simplified following the Reynolds assumption. An averaged sub-grid scale correlation term remains, which represents the impact of the sub-grid scale correlations on the averaged variables. In order to maintain the same number of equations and dependent variables the sub-grid scale correlation terms are formulated in terms of the dependent variables (Pielke, 2002). The sub-grid scale correlation terms are treated as sub-grid scale turbulent fluxes of momentum, heat and moisture and are parameterized using a first order closure in analogy to the molecular fluxes.

The orography and the land-use characteristics of the land surface like surface temperature, soil moisture, surface roughness and heterogeneity play a key role for the correct description of the atmospheric prognostic and diagnostic fields, since the land surface provides the lower boundary condition for the atmosphere. The surface roughness and orography determine the surface drag and modify the wind profile in the boundary layer via the turbulent exchange of momentum. The atmospheric thermodynamic fields are coupled to the land surface via the surface energy balance (Stull, 1988). It describes the partitioning of the net short and long wave radiative fluxes, the heat storage in the soil and turbulent exchange of heat and moisture between surface and atmosphere. The turbulent surface fluxes strongly affect the structure of the atmospheric boundary layer. Hence, their appropriate description in a numerical atmospheric model pays back on the forecast quality of the prognostic fields like temperature, humidity and wind. Also the prediction of the transport of chemicals or pollutants as well as the formation of clouds and the prediction of precipitation depends on the faithful representation of the surface fluxes, since evaporation from the surface strongly impacts the moisture content within the atmosphere.

2.1 LITFASS measurement campaign 2003 to determine surface fluxes

Extensive surface flux measurement campaigns undertaken in Lindenberg, Northern Germany e.g. LITFASS 2003 (Beyrich and Mengelkamp, 2006), underline that varying surface characteristics between different land-use classes and to a small extend minor topographic structures result in significant differences in surface temperatures of up to 10 K even on very small horizontal scales (Beyrich and Mengelkamp, 2006). Along with the temperature differences, Beyrich and Mengelkamp (2006) report remarkably large differences of the surface energy balancefluxes between the three most dominant land-use classes (forest, grassland, farmland) in the so called LITFASS area (Beyrich, 2004). But also between different types of agricultural land-use (grass, rape, maize, triticale, rye) significant differences of the surface energy fluxes are measured. When capturing such strong surface heterogeneities with an atmospheric model very small but strongly convective areas with intense upward-directed turbulent sensible heat fluxes occur next to spatially large and stably stratified areas within the very same grid box. In such situations the convectively driven sensible heat flux can dominate the grid box averaged turbulent sensible heat flux, while the mean gradient of potential temperature still indicates an overall stable stratification. Lettau (1979) reported this so-called Schmidt's paradox. Sub-grid scale surface flux parameterisations have to capture such strong heterogeneities in order to derive a representative grid-box averaged flux and further yield the correct mean temperature gradient by allowing a transport of heat "counter" the mean gradient.

One possibility to cope with those small-scale heterogeneities is to resolve the land surface explicitly by increasing the horizontal resolution of the numerical models drastically towards a couple of meters. However, this is not feasible due to the strong increase in computational costs and limitations of some parameterisations in the models. Instead, models simplify the "real world" land-use heterogeneity by defining a limited number of land-use classes, which represent the major land-use types of the region the model is applied to. The land-use classes are assumed to be homogeneous and hence are assigned with characteristic surface and soil properties. However, this simplified projection of the earth's land-use remains on the sub-grid scale. The impact of the sub-grid scale turbulent fluxes on the grid box averaged atmospheric variables still needs to be parameterised to derive grid box representative turbulent surface fluxes.

2.2 Parameterisation of sub-grid scale surface fluxes

Over homogeneous surfaces the Monin-Obhukov similarity theory is usually applied to calculate the surface fluxes dependent on the vertical gradient of the transported variable. Over surfaces with sub-grid scale heterogeneity aggregation techniques are needed to determine a mean surface flux for the model grid box that combines the influence of the sub-grid scale land-use variability. Generally, the surface characteristics are assumed to vary on a smaller scale than the atmospheric variables in most atmospheric numerical models.

When including sub-grid scale heterogeneous land-use into parameterisations for surface fluxes their applicability range needs to be considered in order to improve the overall model performance. For instance, the surface fluxes are expected to have a strong impact on the boundary layer structure in e.g. convective conditions with very weak winds. In such atmospheric conditions, hereafter referred to as locally driven meteorological situations, the boundary layer is well mixed. The information from the surface is transported high up into the vertical within the atmospheric boundary layer, while horizontal advection plays a minor role. Then, the accurate description of the turbulent exchange with the surface plays a key role for the local boundary layer structure. The impact of different parameterisation schemes on the model performance is expected to be significant. In contrast, atmospheric meteorological situations with strong winds, hereafter referred to as advective meteorological situations, also interact strongly with the surface. But under these conditions the appropriate lateral boundary conditions are probably more essential for the realistic simulation of the boundary layer structure and the model performance. Hence, in such situations a less sophisticated scheme is assumed to be the favourable one to keep computational costs down. Besides the meteorological situation also the horizontal grid resolution is expected to play a role for the model performance. With increasing resolution the land surface is resolved more explicitly and therefore the heterogeneity is resolved more precisely. This is expected to pay back on the calculation of the spatially averaged surface fluxes.

A large variety of parameterisation schemes for sub-grid scale surface fluxes have been developed over the recent years. Giorgi and Avissar (1997) give an introduction into the problems arising from the inclusion of sub-grid scale heterogeneities into surface flux parameterisations. According to them the main difficulty to cope with is the non-linear dependence of the surface fluxes on the surface layer characteristic. This means that a spatially averaged surface flux is not exactly determined by averaging the surface characteristics and then applying the flux function to the artificial but homogeneous surface characteristic value. Further, the combined grid box representative effect of two different fluxes on the main flow cannot exactly be calculated by deriving the individual effects and then simply calculating spatially averaged fluxes. Giorgi and Avissar (1997) call these deficiencies the “aggregation effect”. They further introduce the “dynamical effect” which summarizes inaccuracies due to heterogeneity induced sub-grid scale circulations like land-sea breeze circulations, which are not explicitly resolved on the grid scale. Since the present study investigates horizontal resolutions of 1 km towards 16 km the latter effect is considered to be less important than the “aggregation effect”. The “dynamical effect” needs to be accounted for by parameterisations for horizontal resolutions of several kilometres as it is used for global and regional climate or weather forecast models.

The simplest method to yield a grid box representative surface flux over heterogeneous terrain is the so called main land-use approach (Figure 2.1), sometimes also referred to as dominant land-use approach (Mölders and Raabe, 1996). Following this approach surface fluxes are calculated based on the sub-grid scale land-use class characteristics that makes up the largest fraction of the grid box. The advantage of this approach is that it is computationally cheap. On the downside this method suppresses sub-grid scale surface fluxes that contribute significantly to the grid box averaged surface flux as in the “real world” (Mölders and Raabe, 1996). Especially in very heterogeneous areas this approach is insufficient. If many sub-grid scale land-use classes are present within a grid box, the main sub-grid scale land-use class does only account for a relatively small area of the whole grid box and is not representative. For example, when a wet forest covers 55 % of a grid box and dry sand covers the rest of the grid box, the latent heat flux is significantly overestimated in this grid box, when calculating the mean surface flux based only on the surface characteristics of the forest. Mölders and Raabe (1996) further underline the importance of the horizontal grid resolution in such situations, since increasing the resolution changes the dominant sub-grid scale land-surface classes in the whole model. This leads to different mean grid box fluxes and an overall altered model result. Even the formation of clouds is affected by the altered surface conditions due to different horizontal resolutions according to Mölders and Raabe (1996). This result is not only due to the resolution and the altered surface conditions but also to the dependence of the cloud scheme itself on the resolution. Like Mölders and Raabe (1996), von Salzen et al. (1996) also Schlünzen and Katzfey (2003) compared the main land-use approach with more sophisticated parameterisation methods. They conclude that the omission of sub-grid scale heterogeneity leads to large model errors especially for temperature and dew point temperatures.

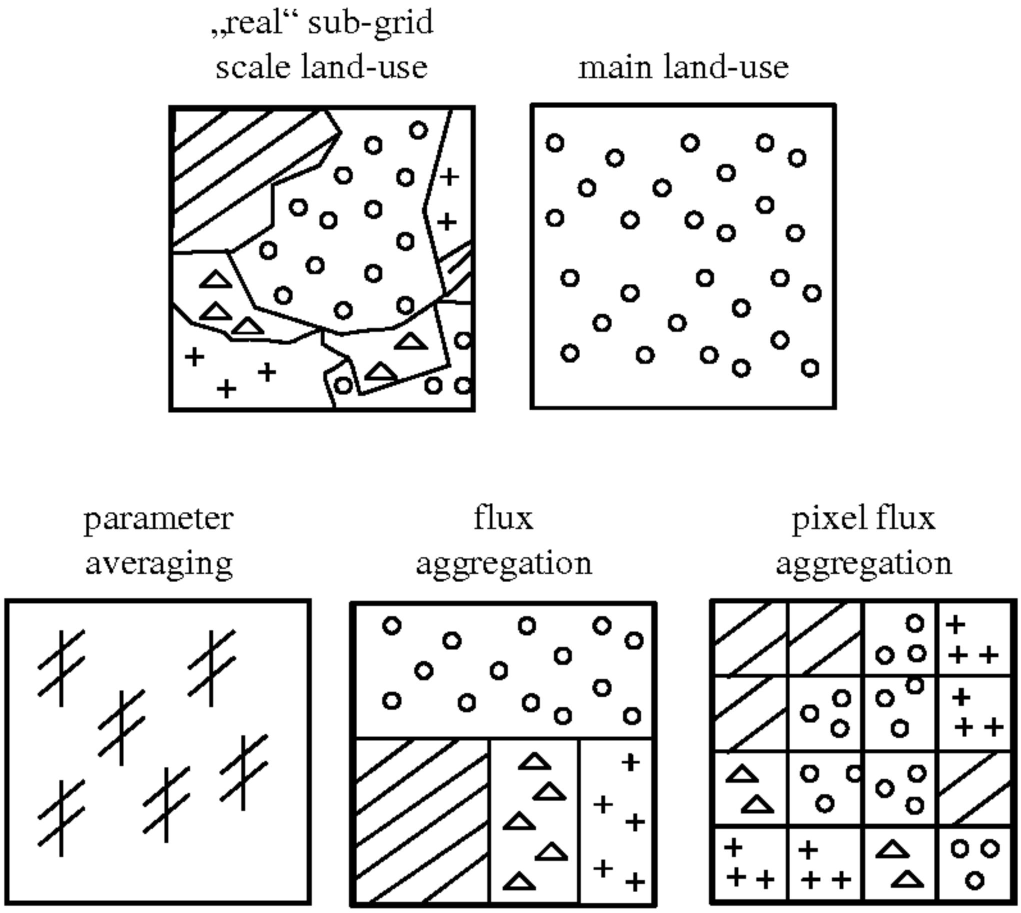


Figure 2.1: Schematic illustrating how the sub-grid scale surface land-use classes are treated within the different parameterisation schemes for sub-grid scale surface fluxes. For further details see text in this section.

More sophisticated approaches attempt to represent the overall effect of all sub-grid scale heterogeneities on the mean grid box surface flux by estimating effective surface parameters. Effective roughness lengths for momentum, heat and moisture $z_{0,eff}$, z_{0h} and z_{0q}, respectively, as well as an effective albedo A_{eff} and effective emissivity E_{eff} for each grid box in the model domain then need to be calculated. Von Salzen et al. (1996) and Schlünzen and Katzfey (2003) calculate those effective surface parameters by a fraction-weighted average of the sub-grid scale surface parameters within each grid box. The mean grid box surface flux is then derived based on the effective surface parameters. Thereby, the problem of calculating a surface flux for sub-grid scale heterogeneity is brought back to calculating the surface flux for a homogeneous but artificial land-use class that represents the overall effect of the sub-grid scale variability onto the main flow. The performance of this method strongly depends on the sub-grid scale heterogeneity. For moderately heterogeneous surfaces this method results in reliable surface fluxes, but for very heterogeneous areas the surface fluxes deviate considerably from the real mean grid box surface fluxes as Claussen (1995) showed for a single case study. Schlünzen and Katzfey (2003) underline that the parameter averaging method can produce even worse results than the main land-use approach especially for very coarse resolutions. For their study they applied two different numerical models. For the simulations using DARLAM (Division of Atmospheric Research Limited Area Model) the roughness length for momentum was computed by an area-weighted logarithmic average from high resolution land-use data prior to the model simulations. In contrast METRAS (MEsoscale TRAnsport and Stream Model) used a parameter averaging method as well as a flux aggregation method to compute the surface fluxes. However, their investigations were undertaken only for a single case study using horizontal resolutions of 18 km and 4 km and they recommend further investigations to draw general conclusions of the resolution dependence of sub-grid scale parameterisation schemes.

Heinemann and Kerschgens (2005) tested a similar approach in their paper, where the effective roughness length is calculated as a logarithmically averaged effective roughness length, while all remaining effective parameters like albedo and soil properties are calculated based on an arithmetic average. They point out that the sensible heat flux could improve from an area-averaged effective roughness length. They state that it is generally underestimated while the latent heat flux is slightly overestimated. However, in their results an improvement of the sensible heat flux results in a less accurate latent heat flux for their parameter averaging method.

Another widely accepted approach for mesoscale models is the flux aggregation approach (Claussen, 1995; von Salzen et al., 1996; Schlünzen and Katzfey, 2003), which is also referred to as mosaic approach in Avissar and Pielke (1989). It is also used in Mölders and Raabe (1996) and Mölders et al. (1996). The flux aggregation approach is referred to as tile-m approach in Heinemann and Kerschgens (2005) and as tile approach in Ament and Simmer (2006). Following this approach, the heterogeneous land surface within each surface grid box is subdivided into a limited number of homogeneous sub-grid scale land-use classes also often referred to as “tiles” or “patches” in the literature (Avissar and Pielke, 1989; Ament and Simmer, 2006). In a further simplification step similar land-use classes are grouped into spatially larger homogeneous tiles regardless of their specific location within the grid box (Figure 2.1). Hence, the impact of local advection on the scale of the sub-grid scale heterogeneity is considered to be less important than the vertical coupling of the individual tiles with the atmospheric mean grid box variables at the first atmospheric level. In contrast to the parameter averaging method the sub-grid scale surface fluxes are calculated for each of the “homogeneous” sub-grid scale land-use classes per grid box. The mean grid box averaged surface flux is then calculated as the fraction area-weighted arithmetic average of the sub-grid scale surface fluxes at the first model level (Mölders et al., 1996; Ament and Simmer, 2006) or at the so-called blending height (Schlünzen and Katzfey, 2003).

Eq. 2.1 and eq. 2.2 explain the calculation of the mean grid box averaged sensible heat flux H exemplarily, with H_i being the sensible heat flux for land-use class i, ρ the air density, C^i_h the transfer coefficient for heat, c_p the heat capacity, u_{ref} is the reference wind speed at the first model level, θ is the potential temperature, z indicates the height at model level k, $z^i_{0,h}$ is the roughness length for heat for the land-use class i, and A_i is the fraction of land-use class i per grid box. The latent heat flux and the momentum flux are calculated accordingly.

$$H_i = \rho c_p C^i_h u_{ref} \left(\overline{\theta}(z_{k=1}) - \overline{\theta}^i_{\left(z^i_{0,h}\right)} \right) \tag{2.1}$$

$$H = \sum_{i=1}^{n} A_i H_i \tag{2.2}$$

Slight modifications of the flux aggregation approach and comparison with other approaches are available in the literature, but for brevity only some are mentioned. For example, Heinemann and Kerschgens (2005) as well as Ament and Simmer (2006) slightly modify Avissar’s and Pielke’s (1989) mosaic approach. They calculate the surface and soil properties on a regular sub-grid within each atmospheric grid box with N^2 grid points instead of relocating and merging the land-use classes (eq. 2.3 and eq. 2.4). For simplicity this approach is referred to as pixel flux aggregation approach in this study regardless of how the individual authors named their approach. The main difference to eq. 2.1 is that the sub-grid scale surface fluxes are calculated at the sub-grid points j and not per land-use class i. Further, they are no longer weighted by the fractional coverage of the land-use classes per grid box but by the number of grid points of the finer sub-grid for the surface. Heinemann and Kerschgens (2005) call this pixel flux aggregation approach the "optimal mosaic" approach, while Ament and Simmer (2006) refer to it as "mosaic" approach. Both underline the advantage that different surface characteristics are taken into account even within the same land-use class. They point out that this method is optimal for future pixel-based land-use data sets, although it is computationally relatively expensive.

$$H_j = \rho c_p C_h^j u_{ref} \left(\overline{\theta}_{(Z_{k=1})} - \overline{\theta}^j_{\left(z^j_{0,h}\right)} \right) \quad (2.3)$$

$$H = \frac{1}{N^2} \sum_{j=1}^{N^2} H_j \quad (2.4)$$

Heinemann and Kerschgens (2005) compare their pixel flux aggregation approach with a flux aggregation approach ("tile-m" approach in their paper). In their study they compare the pixel flux aggregation, the flux aggregation and a parameter averaging approach offline based on a high resolution simulation with 250 m horizontal grid spacing. This high resolution simulation provides the basis for the offline tests of the various 1 km simulations with different surface flux parameterisations. For their pixel flux aggregation approach with 1 km horizontal resolution they prescribe the surface characteristics on a 250 m sub-grid with $N^2 = 16$ sub-grid scale points per 1 km x 1 km grid box by taking them offline from the 250 m simulation. This implies that a feedback mechanism from the atmosphere into the surface and soil is not considered for the coarser 1 km simulations, which might result in an underestimation of differences between the same approach and different horizontal resolutions. They report that aggregating land-use classes instead of using "pixel" sub-grid information shows comparable results to the pixel based approach; hence moderate inhomogeneities within the same land-use class are sufficiently aggregated by the computational less expensive flux aggregation approach. They further find out that the pixel flux aggregation approach and their main land-use approach are resolution-independent for their test scenarios. In contrast, Schlünzen and Katzfey (2003) as well as Mölders and Raabe (1996) conclude that the main land-use approach shows significant resolution dependence. The sub-grid scale heterogeneities are suppressed and the dominant land-use type changes dependent on the horizontal grid resolution. But in contrast to Heinemann and Kerschgens (2005) the atmosphere feeds back into the soil. Schlünzen and Katzfey (2003) further point out that their 4 km simulations deviate less from their rather coarse 18 km simulations, when applying flux aggregation rather than parameter averaging or main land-use approaches. But they also suggest that further more complex studies are necessary to draw more general conclusions.

Ament and Simmer (2006) put much effort into determining the accurate soil properties like soil moisture availability before comparing a pixel flux aggregation approach with a flux aggregation approach by running a stand-alone soil model for 2.5 years. In their study they give no preference for either parameterisation approach. But they conclude from simulations with 7 km horizontal resolution with the Lokalmodel (LM) of the German Weather Service (DWD) that the accurate soil physics is most important in order to compare parameterisation schemes. They state that its impact might be more important than the aggregation method for the fluxes.

All mentioned approaches have in common that area-averaged grid scale atmospheric variables are used for the atmospheric forcing to calculate the surface fluxes via a bulk-approach. Mölders and Raabe (1996) point out that the atmospheric forcing is then no longer in equilibrium with the underlying sub-grid scale surface, which might result in over- or underestimated surface fluxes.

The flux aggregation approach mentioned in Schlünzen and Katzfey (2003) is further extended by averaging the sub-grid scale surface fluxes at the so called blending height following the definition of Claussen (1991) rather than at the arbitrarily chosen first model level, which is usually set to 10 m in most numerical mesoscale and forecast models. Details on the implementation can be found in von Salzen et al. (1996). Briefly summarized; averaging at the blending height as defined by Claussen (1991) fulfils the criteria that the flow is in local equilibrium with the surface and is at the same time independent of local surface characteristics. Von Salzen et al. (1996) point out that no clear transition from heterogeneity towards homogeneity is detectable in their simulations, which contradicts the assumption of the blending height concept. According to them the difference in friction velocity is smaller than 6 % when using simulated blending heights of 100 m and setting the blending height to the first model level at 10 m height.

In the present study the mesoscale transport and fluid model METRAS 3.0 (Schlünzen et al. (1996a)) is used for all simulations. Within METRAS two parameterisation schemes are applied to calculate the sub-grid scale surface fluxes of momentum, sensible and latent heat. Both schemes calculate area averaged values of the scaling values friction velocity u_*, free convection velocity w_* and the scaling value for potential temperature θ_* for each grid box. METRAS accounts for 10 different land-use classes by distinguishing the roughness length, albedo, thermal diffusivity, thermal conductivity, depth of temperature wave, soil water availability and the saturation value for water content. The applied 10 land-use classes are water, mudflats, sand, mixed land-use, meadows, heath, bushes, mixed forest, coniferous forest and urban areas. The different parameterisation schemes are applied to derive the grid box averaged surface fluxes of momentum, sensible heat flux and latent heat flux representative for each grid box.

The parameter averaging scheme is the favourable one concerning computing time. It calculates a fraction weighted roughness length z_0 (eq. 2.5) from the roughness lengths z_{0i} of the sub-grid scale land-use classes for each model grid box. The resulting z_0 is an artificial homogeneous roughness length representative for the whole surface characteristics of each grid box.

$$z_0 = \sum_{i=1}^{n} f_i z_{0i} \tag{2.5}$$

The rationale behind this method is the assumption that the surface fluxes are in equilibrium with the averaged homogeneous artificial surface characteristics of the whole grid box. This assumption performs quite well for nearly homogeneous landscapes that are not too distinct in their surface characteristics. For very heterogeneous areas non-linear effects, which contribute to the grid-box averaged fluxes are not captured (Giorgi and Avissar, 1997), since the flux functions depend in a non-linear way on the surface layer characteristics. Averaging over the surface characteristic variables and applying the flux function suppresses some non-linear effects. This aggregation effect worsens model performance. Therefore, coarser horizontal resolutions are likely to perform worse than high horizontal resolutions where the surface characteristics are resolved more explicitly.

The more sophisticated scheme is a flux aggregation scheme as described by von Salzen et al. (1996), which applies a blending height concept following Claussen (1990). In more heterogeneous areas this method theoretically performs better than the parameter averaging method, since it calculates one-dimensional flux profiles for each of the sub-grid scale land-use classes independent of each other. Therefore, the sub-grid scale land-use classes itself are considered to be homogeneous. It then aggregates them fraction weighted at a certain height. In many models like METRAS or for instance the UK Met Office Unified Model this height is chosen to be the first model level at 10 m height, since advective processes become important higher up in the atmosphere. Then, it can be assumed that the sub-grid scale surface fluxes are in local equilibrium with the "homogeneous" sub-grid scale surface of the individual land-use classes. With increasing height the sub-grid scale surface fluxes are no longer in equilibrium with the homogeneous sub-grid scale land-use class they are calculated for. Instead, the surface fluxes tend to be in equilibrium with the effective surface characteristics of the whole grid box. The height, at which the flow is in equilibrium with the underlying heterogeneous surface and does no longer distinguish the effects of the local surface characteristics, is the so-called blending height (von Salzen et al., 1996, Claussen et al., 1990). The blending height l_b is a function of the characteristic length scale of the surface heterogeneities and also depends on the atmospheric stability. At this blending height the fraction weighted aggregated fluxes are assumed to be in equilibrium with the effective characteristics of the underlying heterogeneous surface. In METRAS, the sub-grid scale fluxes are calculated for each land-use class based on the specific roughness lengths for momentum, temperature and humidity. As an example the latent heat flux is given in eq. (2.6).

$$
\begin{aligned}
\rho l_{21} q_* u_* &= \rho l_{21} \sum_{i=1}^{n} f_i q_{*i} u_{*i} \\
&= \rho l_{21} \kappa^2 U(z_1) \\
&\cdot \sum_{i=1}^{n} f_i \cdot \left(q(z_1) - q(z_{0qi})\right) \cdot \left[\left(\ln\left(\frac{z_1}{z_{0i}}\right) - \Psi_m\left(\frac{z_1}{L_i}\right)\right) \cdot \left(\ln\left(\frac{z_1}{z_{0qi}}\right) - \Psi_q\left(\frac{z_1}{L_i}\right)\right)\right]^{-1}
\end{aligned}
\tag{2.6}
$$

Specific humidity is denoted by q, U(z) is the main flow at height z1, ψ_m and ψ_q are the stability functions for momentum and humidity according to Dyer (1974). The von Kárman constant κ is set equal to 0.4. z_{0qi} is the roughness length for specific humidity q for land-use class i and l_{21} is the latent heat of evaporation. The sub-grid scale surface fluxes are calculated separately for each sub-grid scale land-use class and the resulting grid box flux is the fraction weighted average of these at the first model level, which is at 10 m height. The blending height and the effective roughness length z_0 are calculated following von Salzen et al. (1996) (eq. 2.7). z_0 is the effective roughness length, L_x a characteristic length scale and f_i the fraction of land-use. This approach works reasonably well as long as the surface characteristic length scales are not too small. Then, the assumption of local equilibrium near the homogeneous surface of a sub-grid scale land-use class is fulfilled. Since METRAS is only run with a resolution down to 2 km in these sensitivity studies and since the position of each sub-grid scale land-use class is not considered but the overall fraction within each grid box is used for calculating the surface flux of each land-use class, this assumption should be fulfilled.

$$\frac{l_b}{L_x}\left(\ln\frac{l_b}{z_0}\right) = c_1\kappa$$

$$\left(\ln\frac{l_b}{z_0}\right)^{-2} = \sum_{i=1}^{N} f_i\left(\ln\frac{l_b}{z_{0,i}}\right)^{-2} \tag{2.7}$$

3 Sensitivity of model performance on uncertainties in initial soil temperature and humidity

This chapter will be submitted to the Meteorol. Z. as:

Bohnenstengel, S.I. and K.H. Schlünzen (2012): Sensitivity of the model performance on initialisation, *Meteorol. Z.*, *to be submitted*.

3.1 Introduction

Land surface processes play an important role for the development of the atmospheric boundary layer. The partitioning of the sensible and latent heat fluxes affects the interaction between the land surface and the lower atmosphere. Horizontal heterogeneities in the soil temperature and moisture fields lead to strong differences in the surface fluxes on small spatial scales. These differences can even alter the wind field and induce mesoscale circulations on longer time scales (Hess, 2008) and they have an immediate impact on the near surface temperature and moisture fields (Lam et al., 2006).

Atmospheric forecasts depend on reliable values for the initial atmospheric and soil conditions. For instance, the latent heat flux is most sensitive to changes of the soil moisture and plant conductance and is mainly a function of the difference between atmospheric humidity and soil moisture availability. An alteration of the Bowen ratio due to the initialisation of the soil parameters can result in considerable changes of the near surface temperature and dew point values and hence even drive local circulations.

Atmospheric models describe the land surface by defining a limited number of vegetation types and soils. They compile the corresponding soil and vegetation parameters such as roughness lengths for momentum, moisture and heat, albedo, emissivity, conductivity, heat capacity, etc. from the literature. Currently atmospheric models are reaching the kilometre scale and the land-use characteristics are resolved more explicitly. However, the improvements in the horizontal resolution might not enhance the overall model performance, if the initialisation of the soil parameters is not sufficient. Ament and Simmer (2006) highlight the importance of a correct soil initialisation for a reasonable forecast. They improve their initialisation of soil properties using an observation based land-surface assimilation scheme and conclude that high resolution soil moisture information is essential to simulate the land-atmosphere exchange sufficiently. However, such an assimilation scheme is not yet present in many atmospheric research models. Instead, soil properties in research models like the mesoscale transport and fluid model METRAS are initialised from measurements at single points that are then interpolated over the whole model domain. The measurements are flawed with a measurement uncertainty and are available with a coarse resolution mostly. Interpolating them onto the numerical grid of an atmospheric model leads to a large uncertainty in the initialisation. Alternatively, METRAS allows to initialise soil properties from ECMWF analysis for instance. This has the advantage that information is available on a much finer resolution. Its disadvantage is that the soil physics of such an analysis differs from the soil physics of the model being initialised. Both methods have caveats and introduce a large uncertainty into the forecast.

The aim of the present study is to determine the dependence of the performance of the regional atmospheric numerical model METRAS on the uncertainty in the initialisation of the soil temperature and soil moisture fields. Therefore, the initial soil values are varied within their uncertainty range. The sensitivity of the atmospheric variables towards the variability of the soil moisture and soil temperature is then explored for a very heterogeneous area in North-Eastern Germany. Simulations are undertaken for a very dry and warm period during summer 2003 and the simulated near surface temperature and humidity fields are compared with observations from the “Deutscher Wetterdienst” (DWD). The simulation period and the model set-up are described in Section 3.2. The sensitivity of the atmospheric model towards the initialisation of soil parameters is determined in Section 3.3. Conclusions on the sensitivity towards the uncertainty in the initialisation of soil parameters are drawn in Section 3.4.

3.2 Model set-up and simulation period

For the numerical simulations the mesoscale transport and fluid model METRAS 3.0 is used. METRAS is part of the model system M-SYS (Trukenmüller et al., 2004). METRAS has successfully been applied in a wide range of scientific questions ranging from chemistry applications (von Salzen and Schlünzen, 1999a; von Salzen and Schlünzen, 1999b ; Schlünzen and Meyer, 2007) and pollen dispersion (Schueler and Schlünzen, 2006) towards the investigation of polynyas in arctic regions (Hebbinghaus et al., 2007; Lüpkes and Schlünzen, 1996). In the following only those parts of the model are introduced that are relevant for the present study. For a more detailed METRAS model description, please refer to Schlünzen et al. (1996a) or Dierer and Schlünzen (2005).

METRAS is a non-hydrostatic prognostic mesoscale model, using terrain-following coordinates. The primitive equations are Reynolds-averaged and solved in flux form. METRAS is designed for model applications of the meso-β and meso-γ scale by employing the anelastic approximation, the Boussinesq approximation and using a constant Coriolis parameter.

METRAS 3.0 accounts for 10 different land-use classes by distinguishing the roughness length, albedo, thermal diffusivity, thermal conductivity, depth of temperature wave, soil water availability and the saturation value for water content. The 10 land-use classes applied in the present study are water, mudflats, sand, mixed land-use, meadows, heath, bushes, mixed forest, coniferous forest and urban areas. For calculating surface fluxes Monin-Obukhov similarity theory is assumed using the stability functions of Dyer (1974). The surface fluxes include sub-grid scale land-uses of the 10 classes. The fluxes are calculated using a flux aggregation scheme with blending height approach (Claussen, 1995; von Salzen et al., 1996; Schlünzen and Katzfey, 2003). Above the surface layer a first order turbulent closure is applied using a mixing-length approach for stable or nearly neutral conditions, where the mixing length is calculated following Blackadar (1962). The turbulent fluxes are calculated proportional to the local mean gradients of the transported variable. For convective conditions a counter-gradient scheme is applied (Dierer and Schlünzen, 2005). For this, a non-local counter-gradient term is added when calculating the fluxes for heat and moisture, which assures the vertical turbulent mixing even in a well mixed convective boundary layer.

The simulations are undertaken for a 400 km x 400 km domain in North-Eastern Germany with Berlin located slightly north-west of the middle of the domain. The sub-grid scale land-use in the domain is very heterogeneous while the orographic influence can be neglected, since the whole domain is very flat without any higher mountains. The land-use data are derived from the CORINE (Coordination of Information on the Environment) data set on a 30” grid. Figure 3.1 shows the main land-use in the model domain for the applied horizontal resolution of 16 km. The largest red area in Figure 3.1 indicates Berlin.

For taking into account large-scale meteorological conditions for short-term forecasts, METRAS is nested in coarser METRAS model results, an own analysis of observations (Trukenmüller et al., 2004) or into ECMWF analysis (Ries et al., 2010) as done in this paper. The ECMWF analysis data are used to derive background vertical profiles. These result from 1D METRAS, which calculates stationary and horizontally homogeneous profiles of temperature, humidity, wind speed, wind direction and pressure. Additionally, soil temperature, water temperature and soil water content are prescribed for the initialisation time. The prognostic variables like temperature, specific humidity, wind speed and wind direction are forced during the whole simulation, and thus METRAS accounts for the larger scale synoptic situation. Nine simulations are performed and all of them are started for 18 UTC of 9th August 2003. For all sensitivity tests surface pressure is set to 1019.12 hPa at the initialisation point at 53°34’41” N and 14°41’38” E. This point has a height of 5.69 m above sea level. METRAS assumes a constant basic state pressure profile and considers basic-state pressure gradients via the geostrophic wind. The parameters changed for the sensitivity studies are listed in Table 3.1 and Table 3.2. Soil temperature, the soil water content and the atmospheric vertical relative humidity profile during the initialisation are listed in Table 3.1 for all configurations.

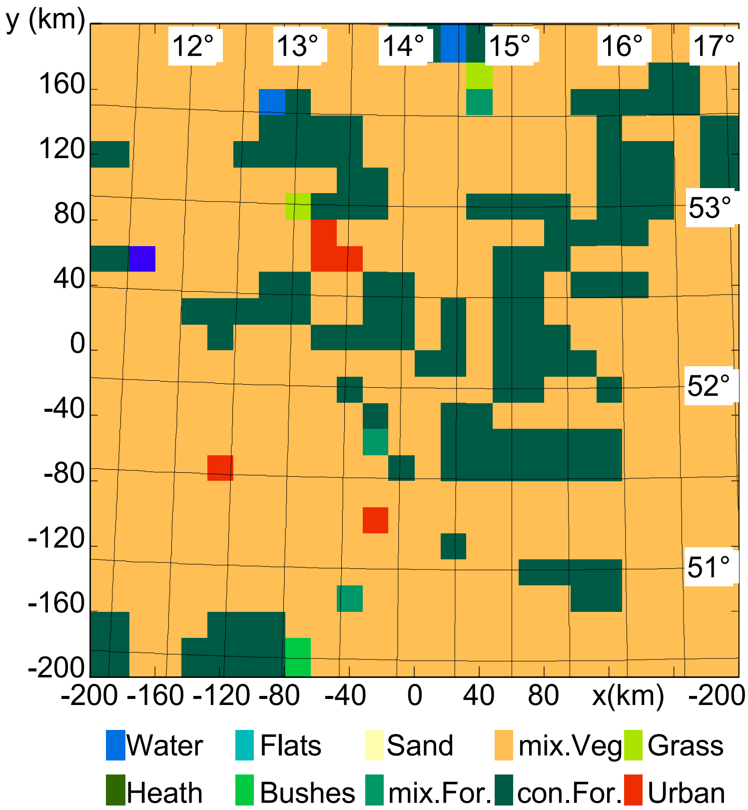


Figure 3.1: Main land-use within the whole model domain for a horizontal resolution of 16 km. Berlin is characterised by the largest red shaded area.

Case	T_{soil} and T_{water}	Soil water content	rh profile
3a	22 °C	Dry	ECMWF
3b	19 °C	Dry	ECMWF – 10 %
3c	19 °C	Dry1	ECMWF – 10 %
3d	17 °C	Dry	ECMWF
3e	17 °C	Dry1	ECMWF
3f	17 °C	Dry2	ECMWF
3g	17 °C	Dry	ECMWF – 10 %
3h	17 °C	Dry1	ECMWF – 10 %
3i	17 °C	Dry2	ECMWF – 10 %

Table 3.1: Changed parameters for sensitivity studies. T denotes temperature, rh relative humidity, ECMWF the ECMWF analysis interpolated to METRAS grid.

Soil water content	Water	Mud flats	sand	mixed land-use	meadow	heath	bushes	mixed forest	Coniferous forest	Urban
Dry	0.98	0.5	0.05	0.1	0.1	0.05	0.1	0.1	0.005	0.005
Dry1	0.98	0.5	0.01	0.01	0.01	0.05	0.1	0.1	0.005	0.005
Dry2	0.98	0.5	0.01	0.01	0.01	0.05	0.05	0.05	0.005	0.005

Table 3.2: Bulk soil water availability α at initialisation for the different cases. α describes the percentage of the maximum possible field capacity W_k available at the time of initialisation. α lies between 1 and the ratio of the bulk soil moisture content (depth of liquid water) W_s and the field capacity W_k. α is calculated as $\alpha = MIN(1,W_s/W_k)$ according to Schlünzen et al. (1996a).

For this study a very locally driven meteorological situation was chosen based on 6 hourly calculations of a locality index I_{rh} (Bohnenstengel and Schlünzen, 2012; Chapther 5 of this thesis). I_{rh} is a measure for the strength of the turbulent transport from the surface and is mainly a function of the friction velocity and the free convection velocity. Therefore, this locality index is a measure for how relevant local influences might be for a selected meteorological situation. The differences in the soil moisture and soil temperature are expected to affect the near surface thermodynamic fields. The simulation is started for 18 UTC for the 9th August 2003 and integrated for 4 days. The dry and warm period from 9th August 2003 until 13th August 2003 was nearly cloudless and dominated by the anticyclone "Michaela" which was located over central Europe. In general, the year 2003 was a very dry and hot year in Europe and the selected period is characterised by very high temperatures of above 30 °C and dew points between 2 °C and 16 °C. Low wind speeds from northerly directions allow local processes to influence the model solution. The relative humidity was highly varying between 30 % and 80 % with lower relative humidity in the northern parts of the model domain and higher relative humidity in the southern model parts, and no precipitation was measured. The period was chosen, since it was a very dry period and a good opportunity to investigate the performance of the surface flux schemes for an extreme situation, where the standard METRAS moisture approach is no longer applicable. This simulation was nudged into results derived by the newly developed pre-processor for ECMWF analysis data. A short description of the nudging method can be found in Ries et al. (2010).

3.3 Influence of the initialisation on the model performance for short-term forecasts

The impact of the different initialisation values for soil temperature, soil moisture availability and trelative humidity on the model results is investigated by evaluating the different configurations in comparison to measured data. Hit rates H are calculated following Cox et al. (1998):

$$H = \frac{100}{m} \sum_{i=1}^{m} n_i, \quad \text{with } \begin{cases} 1 \text{ for } |\text{difference}(\text{measurement}, \text{model})| < A \\ 0 \text{ for } |\text{difference}(\text{measurement}, \text{model})| \geq A \end{cases} \tag{3.1}$$

As in Schlünzen and Katzfey (2003) the desired accuracy, A, for the air temperature and dew point is set to ±2 K, for wind speed ±1 ms^{-1}, for wind direction ±30° and for pressure ±1.7 hPa. The hit rates are complemented with calculating biases for temperature, dew point, wind speed and wind direction, since the hit rate does not give an indication if the simulations are over- or underestimating the measurements.

Additionally, METRAS forecasts and DWD measurements are compared using conditional quantile plots (Wilks, 2006). Time series including ECMWF forcing data are used to gain a visual impression. DWD data for comparison are available every hour for the whole study period. For every day about 194 data pairs are compared, using data from 29 DWD stations situated in the model domain. The number of pairs differs slightly per parameter and day due to missing observations at some times.

3.3.1 Sensitivity to surface soil temperature

It is expected that the surface soil temperature and the water temperature affect the energy balance at the surface strongly and thereby influence the 2 m values of temperature and dew point, which are evaluated with the corresponding DWD measurements. Two simulations with different soil temperatures for the initialisation at 18 UTC on 9th August 2003 are compared. In case 3a 22 °C is chosen (Table 3.1) for the soil temperature and for the water temperature at the initialisation point according to the ECMWF analysis. In case 3d (Table 3.1) 17 °C is chosen for this point based on measurements in 10 cm depth at Lindenberg located in the middle of the model domain. Hit rates are calculated for temperature and dew point at 2 m height and for wind speed and wind direction at 10 m height.

Figure 3.2 shows the hit rates for all three simulation days for wind speed, wind direction, temperature and dew point for all nine configurations according to Table 3.1. The corresponding biases are shown in Figure 3.3. The red (case 3a) and magenta (case 3d) bars indicate the hit rates and biases for the two simulations testing the variability of the near surface atmospheric fields due to the uncertainty in the soil and water temperature initialisation. Hit rates and biases for wind speed and wind direction are nearly the same for case 3a and case 3d. Hit rates for both simulations seem to be independent of the soil temperature initialisation. Only small differences between configuration 3a and 3d are visible in the hit rates and biases for 2 m air temperature and during the first simulation day for the dew point. Since the soil temperature difference is homogeneously 5 °C in the beginning of the simulation horizontal gradients of air temperature, humidity and pressure are unchanged. No thermally driven circulation is generated which dominates the overall simulated pattern.

The overall performance of wind speed increases with increasing simulation time (Figure 3.2a). While the performance of wind speed results in hit rates of over 50 % for the first simulation day, the hit rates are nearly 80 % for the second simulation day and are even better during the third simulation day with over 90 %. The performance of wind direction decreases with simulation time from 70 % for the first simulation day to just above 40 % for the third simulation day. Wind speed and wind direction were both interpolated from ECMWF analysis to initialise and laterally force METRAS. METRAS slightly overestimates wind speed during the beginning of the simulations compared to the DWD measurements which are originally available with an increment of 1 kn but are converted into ms^{-1} (Figure 3.4; Figure 3.5; Figure 3.6). Figure 3.4 indicates that the simulated wind speed is nudged to higher wind speeds by the ECMWF than measured by the DWD. The corresponding conditional quantile plot (Figure 3.5) underlines that METRAS sometimes simulates higher wind speeds than the DWD measurements. However, according to the histogram in Figure 3.5 most of the wind speeds are simulated reasonably well compared to the DWD measurements. Given the small range of wind speeds between 0 and 4 ms^{-1} and the fact that high wind speeds are forecasted not very often the constant median at the upper range and its larger deviation from the 1:1 line has to be viewed with caution, since it is not statistically significant.

To understand differences in the model performance one has to understand how METRAS employs the forcing data from ECMWF analysis. In METRAS the nudging method is employed as used in several other atmospheric models. To allow the METRAS model some free development especially in the boundary layer the nudging is not used independently of height – but with smaller coefficients close to the surface. ECMWF analysis account for the Coriolis force as well as for turbulent mixing and so does METRAS, but both use different vertical resolutions and exchange coefficient formulations. Therefore, ECMWF analysis values for wind speed and direction are not used below a height of 850 hPa to avoid including these terms twice. According to Figure 3.6 the wind direction simulated by METRAS at 10 m above ground has a smaller scatter than the DWD measurements especially during the 3rd simulation day. With the in general low wind speeds this is not a large surprise, since in these conditions wind direction values are quite uncertain and locally influenced. Again, this is caused by the nudging with the ECMWF data that METRAS follows closely (Figure 3.6). With on-going simulation time METRAS shows a slightly more clockwise wind direction during the 2nd simulation day compared to the ECMWF forcing. The range of wind directions simulated by METRAS and ECMWF is smaller than the range of wind directions measured by the DWD. And METRAS is nudged to too high wind speeds by the ECMWF. This results in a decreasing performance in hit rates during the 2nd and 3rd simulation day (Figure 3.2b).

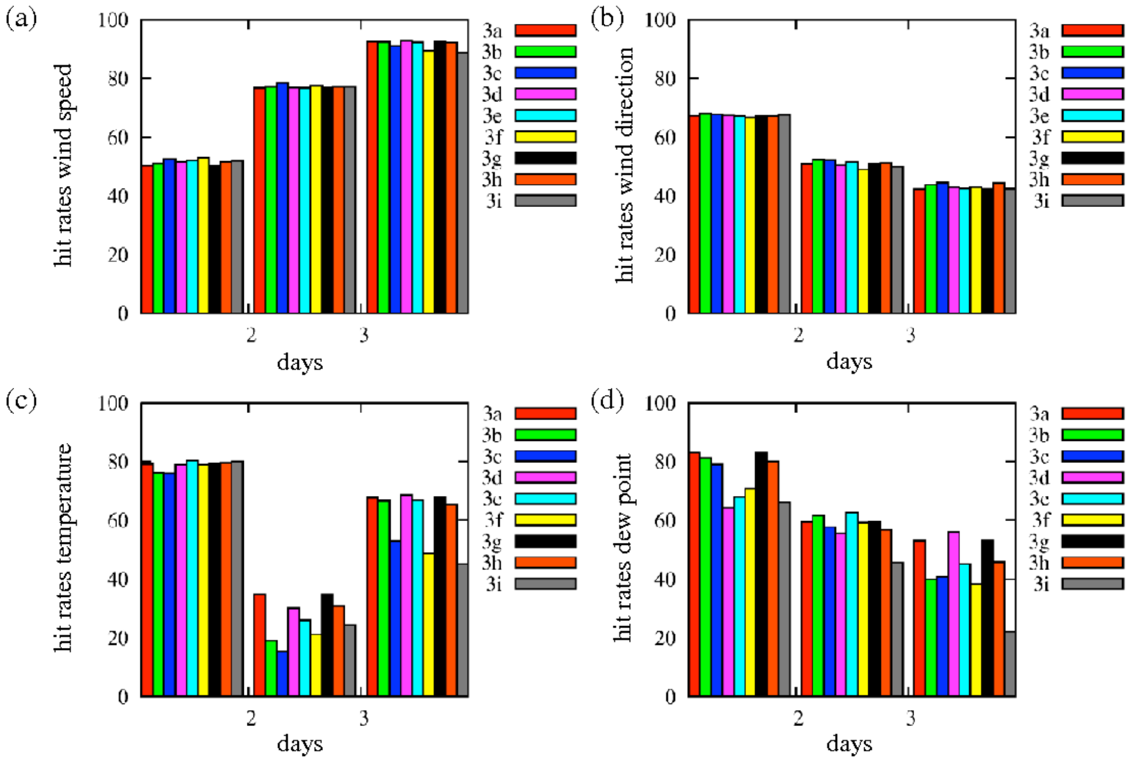


Figure 3.2: Hit rates for all configurations (Table 3.1, 3.2) for all three simulation days.

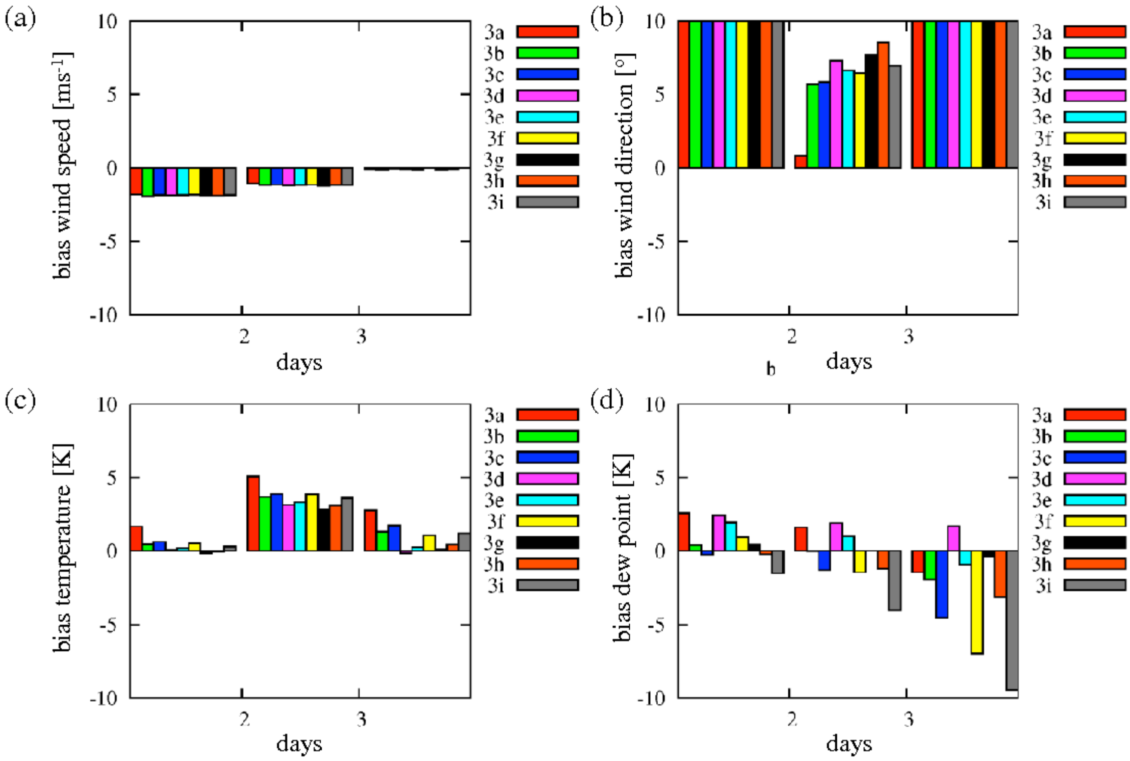


Figure 3.3: Bias for all configurations for all three-simulation days. Bias for wind direction has been cut off.

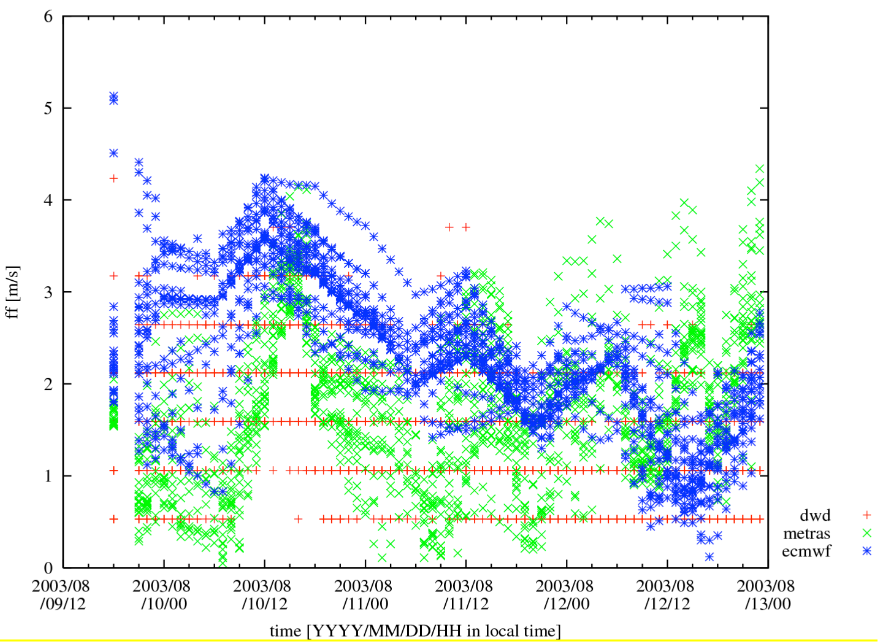


Figure 3.4: Wind speed in ms^{-1} for the whole simulation period for case 3d for METRAS, DWD and interpolated ECMWF at 10 m height using data for all measurement sites. Note that some data are plotted on top of each other and thus not visible.

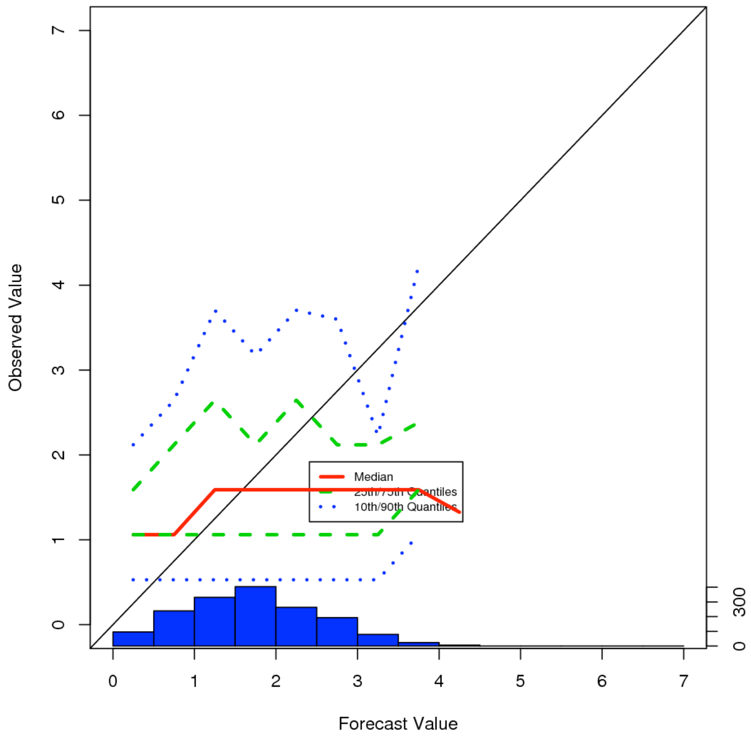


Figure 3.5: Conditional quantile plot for wind speed in ms^{-1} for case 3d. The blue bars indicate the histogram of the wind speed simulated by METRAS. The red line depicts the median of the DWD measurements for the given METRAS wind speeds. Green lines indicate the 25th and 75th percentiles and blue lines the 10th and 90th percentile. Data for all three days are used.

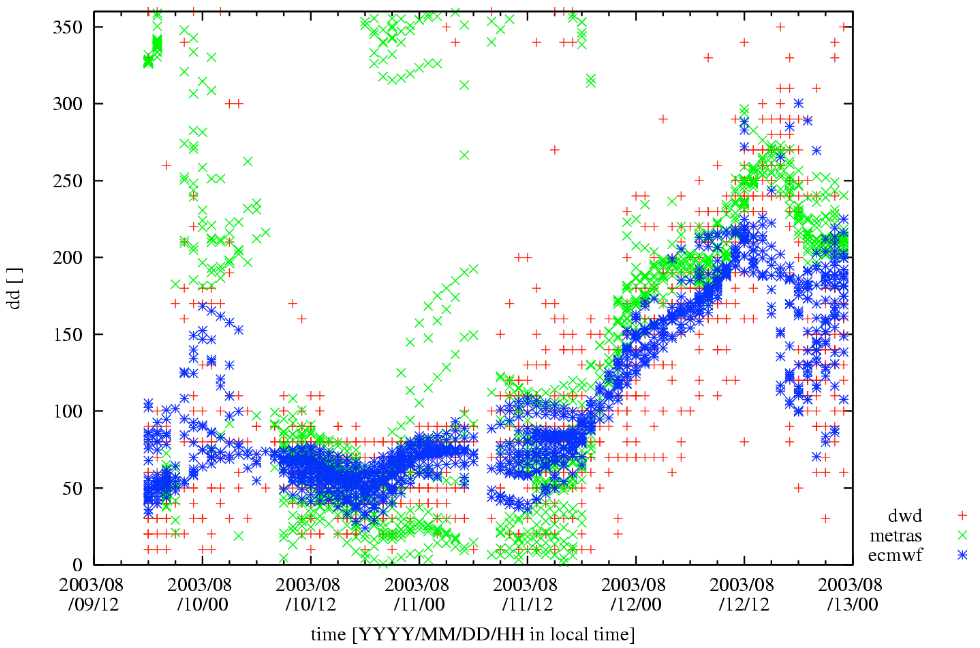


Figure 3.6: As Figure 3.4 but for wind direction in [^{o}].

The 2 m temperature is only slightly affected by the increase of soil temperature when comparing the hit rates (Figure 3.2) and allowing an accuracy range of ±2 K due to the coarse 16 km horizontal resolution. However, the conditional quantile plots for case 3a (Figure 3.7) and case 3d (Figure 3.8) reveal that the lower soil temperature (case 3d) improves the METRAS temperatures. For case 3d the median follows the 1:1 line and the percentiles are close together. The bias of air temperature in Figure 3.3 supports this. Case 3d shows a smaller positive bias than case 3a for all simulation days.

Changing the soil temperature mostly pays back for dew point temperatures when comparing the hit rates. According to Figure 3.2 the hit rates of case 3a and 3d differ strongly for the first simulation day with a decrease of over 15 % in case of lower soil temperatures (case 3d). With lower soil temperatures the dew point temperature is lower during the first simulation day. Differences between case 3a and case 3d are smaller during the second simulation day. During the third simulation day hit rates for case 3d are slightly higher. The corresponding biases show only a small difference for the first and second simulation day and indicate a positive bias for the dew point during the third simulation day of case 3d and a negative bias in dew point temperature for case 3a. The difference between the hit rates which is not reflected by a difference in the bias results from the dew points being at the upper limit of the allowed accuracy range for calculating the hit rates. Figure 3.9 shows that the dew point temperatures are at the upper limit of values the DWD measured during the whole period.

The higher soil temperature taken from ECMWF analysis results in an overall positive offset compared to DWD measurements (case 3a), while case 3d initialised from measurements in Lindenberg fits much better with METRAS model physics. In general it can be seen that METRAS overestimates low temperatures in case 3a and only slightly in case 3d. Altogether, the uncertainty in the initialisation of the soil temperature affects the accurate simulation of 2 m air and dew point temperatures. However, in both cases METRAS was too humid. One reason might be that the simulated days are part of an extremely dry weather period and that METRAS' soil was too humid. Therefore, the extremes of the daily cycle were not captured so well.

It has to be noted that the forcing with the interpolated ECMWF analysis data partly changed model results for the worse. According to Figure 3.10 ECMWF 2 m air temperatures are higher than DWD measurements. Thus, METRAS tends to improve the interpolated ECMWF analysis data.

Altogether, this sensitivity study reproduced the assumption made in the beginning that the initialisation of the surface soil temperature affects the model performance. The uncertainty range of the soil temperature at the initialisation led to a signal in METRAS performance that is larger than the measurement uncertainty given in the beginning of Section 3.3. Especially the thermodynamic values are affected by an altered surface energy balance and a thereby varied Bowen Ratio leading to changes in air temperature and dew point temperature. According to this study and previous investigations made by Schlünzen and Katzfey (2003) it has to be stressed that surface temperatures provided by the ECMWF analysis products do not fit with the METRAS soil scheme and have also led to poor results with other mesoscale models like DARLAM. Taking direct measurements of the soil temperature instead results in a much better agreement with METRAS soil physics and a more reliable model simulation.

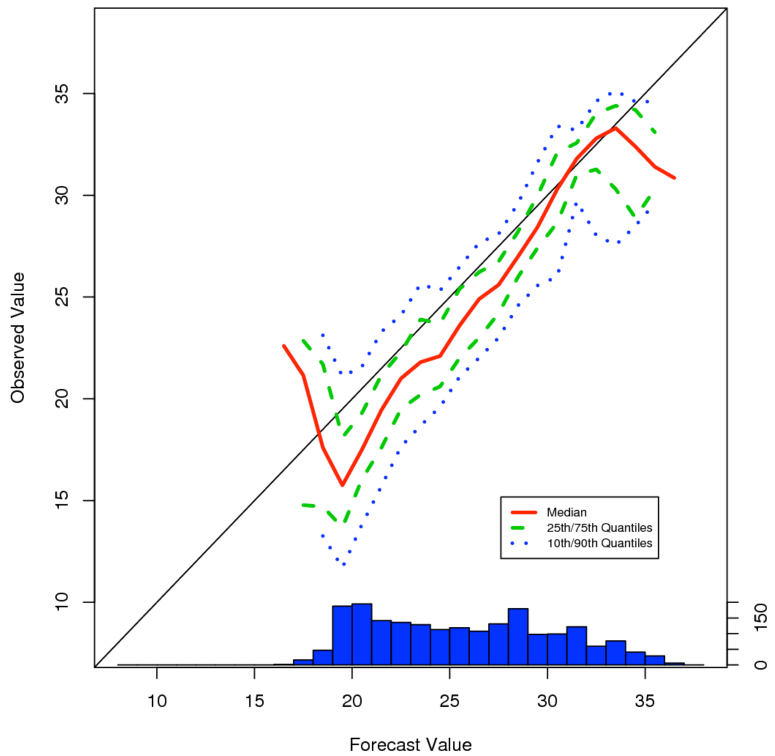


Figure 3.7: As Figure 3.5 but for 2 m temperature [^{O}C] and for case 3a.

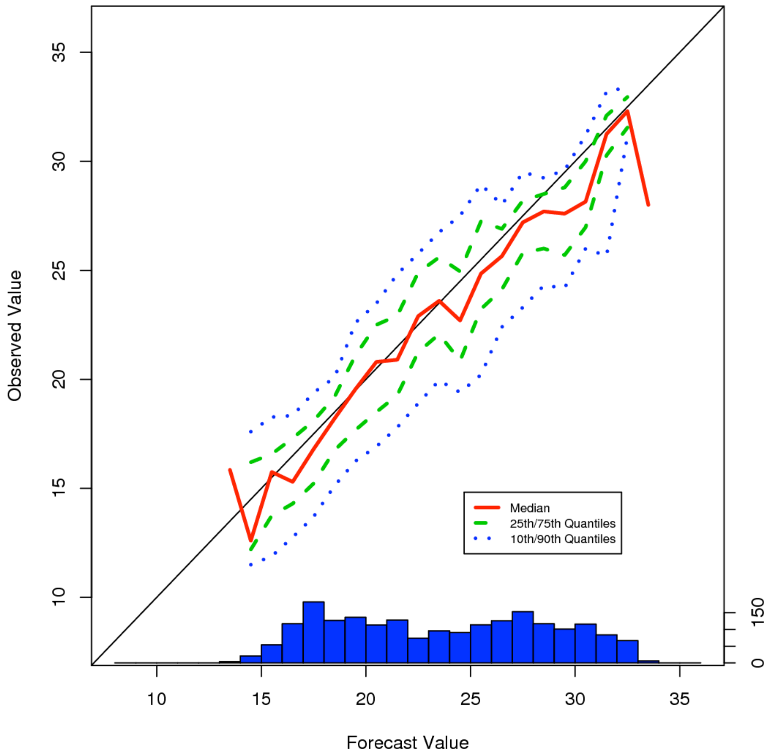


Figure 3.8: Same as Figure 3.7 but for case 3d.

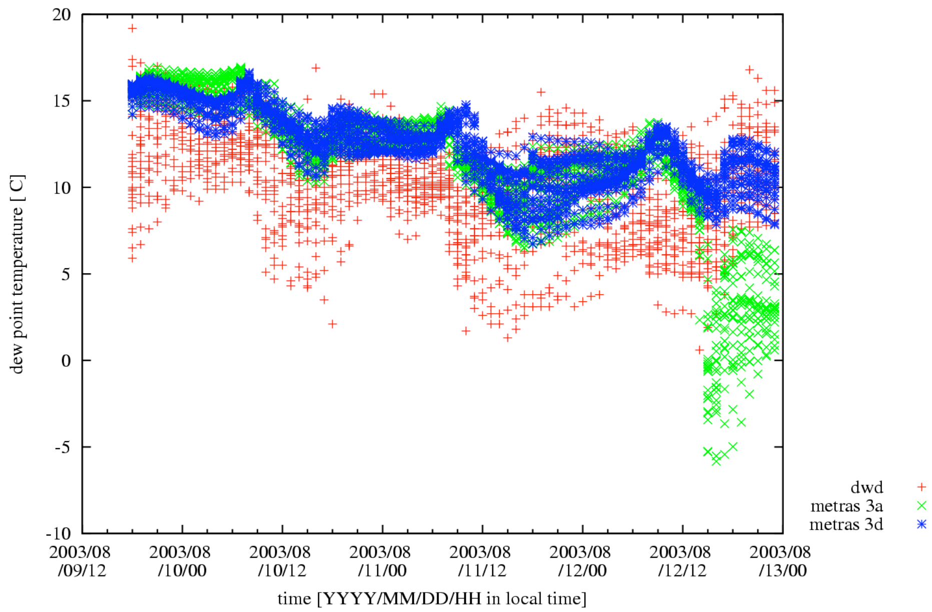


Figure 3.9: Dew point temperature for the whole simulation period for DWD, case 3a and 3d in a height of 2 m above ground.

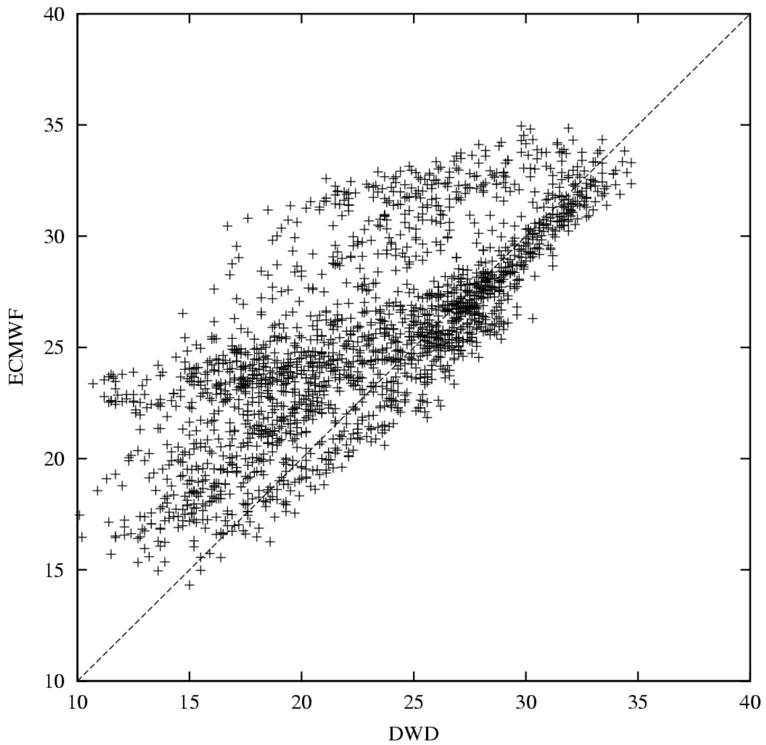


Figure 3.10: Scatter diagram of 2 m temperature [^{O}C] interpolated from ECMWF analysis and DWD measurements for the whole simulation period and all DWD sites.

3.3.2 Sensitivity to soil surface humidity

The storage and release of water in the soil depends on evaporation and thus atmospheric influences like wind speed and temperature on precipitation and on the soil and land-use type. It can be assumed that bare soil has a lower moisture capacity and a lower ability to hold and release water than vegetated surfaces. The latter might be evaporating even after a longer period of dryness and can still release energy via the latent heat flux because they use deep soil water. Thus, also the sufficient initialisation of soil moisture availability is assumed to be essential for a reliable model simulation.

To determine the impact of the uncertainty range in soil water availability configurations cases 3d-3i (Table 3.1; Table 3.2) were analysed with regards to 2 m temperature, dew point temperature and wind speed and wind direction. Decreasing the soil water availability has nearly no effect on the hit rates and biases of wind speed and wind direction (Figure 3.2a,b and Figure 3.3a,b). Changing the soil moisture does not affect the dynamic variables. For the thermodynamic values temperature and dew point temperature, changes in the hit rates between the different cases are within 5 % to 10 %. The performance of temperature stays the same with decreasing soil moisture for the 1st simulation day. During the 2nd and 3rd simulation day the model performance is decreased compared to case 3d by about 5 % to 10 % with decreasing soil moisture for cases 3e,f and compared to 3g for cases 3h,i. Comparing the hit rates for dew point temperatures day-wise shows an increase in the hit rates for the first day with decreasing soil moisture for cases 3e,f. During the second day simulation 3e shows the best results but differences due to changes in soil moisture are small. During the third simulation day the hit rates decrease strongly compared to the first simulation day and with decreasing soil moisture availability. The same behaviour is found for cases 3g-3i, whose initial profiles of relative humidity are reduced by 10 %. Given the coarse resolution of the model results that are interpolated into the location of the observations to calculate statistics like hit rates increases the uncertainty of the statistics. Therefore, it is assumed that changes in hit rates lower than 5% are within the uncertainty range of the calculation of the hit rates.

The bias (Figure 3.3) provides some more detailed information on METRAS' performance. It is calculated as average difference of METRAS simulation minus DWD measurement. With decreasing soil moisture the positive bias for air temperature increases slightly indicating that METRAS temperatures tend to increase, since more energy is channelled into the sensible heat flux. However, the biases during the 1st and 2nd simulation days are small for all cases except for cases 3a,b,c. These were initialised with higher soil temperatures which in turn increases air temperatures.

For temperature the 1st simulation day and the 3rd simulation day show the best agreement with DWD measurements in the hit rates and the bias. During the 2nd simulation day an overall overestimation for all three model configurations is visible in the bias.

Dew point temperatures tend to decrease, since the evaporation continuously dries the soil. The bias for 2 m dew point temperatures changes its sign from positive to negative with decreasing soil moisture for cases 3d,e,f and cases 3g,h,i. Largest negative biases are visible for the "Dry2" soil moisture configuration (case 3f and case 3i) especially during the third simulation day indicating that METRAS is now too dry. Overall case 3g shows the smallest bias for the thermodynamic values over the whole simulation period.

The wind speed shows no differences for the applied model configurations. METRAS underestimates the wind speed similarly for all configurations. In contrast the wind direction is affected by the model configuration and the bias differs for the various configurations. The initialisation of the soil temperature and moisture fields alters the surface energy balance. And the surface energy balance provides a strong forcing to the boundary layer structure. Therefore, the vertical mixing is affected, which explains the change in the wind direction. Also, some small local circulations might have developed and slightly changed the wind direction.

With increasing simulation time, METRAS dew point temperatures start to decrease according to the bias (Figure 3.3d). The drier the soil is initialised the higher are the resulting 2 m air temperatures and the lower are the dew point temperatures. This is due to the change in the Bowen ratio.

The presented analysis underlines that the uncertainty range in the soil water content initialisation affects the performance of the thermodynamic atmospheric variables beyond their accuracy range of ± 2 K used in the comparison with measured data. This allows us to conclude that heterogeneities in the soil fields need to be captured accurately already in the initialisation of land surface scheme in order to simulate reliable temperatures.

3.3.3 Sensitivity to the initialisation profile of relative humidity

Besides altering the available soil water content the relative humidity of the initialisation profile is now reduced by 10 % over the whole vertical profile. The corresponding test cases are cases 3g, 3h and 3i.

When comparing the hit rates for 2 m air temperature in Figure 3.2 for these cases with the corresponding ones with a 10 % higher relative humidity for the whole profile (cases 3d, 3e and 3f) the differences in the performance are negligible. The influence of the soil moisture on temperature weights much more. The same is valid when comparing the corresponding dew point temperatures (3d-3g, 3e-3h, 3f-3i). Only for the first simulation day changes occur. These become smaller during the simulation. Here, also the soil moisture content and the forcing data are much more important for the performance than the initial profile of relative humidity. The bias underlines the conclusions. For temperature the bias behaves similar for all comparable cases. The bias of the dew point temperatures shows a higher influence of the initialisation profile, but still the soil moisture content shows the larger effect on the bias.

3.4 Conclusions

In the present study the impact of altering soil input parameters within their uncertainty range was determined with regard to the near atmospheric fields like temperature, dew point and wind. To ensure a strong local signal from METRAS a very locally driven meteorological situation was chosen in order to minimize the impact of lateral forcing via the boundaries.

The meteorological situation during the whole simulation period was very dry not only in the study region but also in the whole of Europe. Since the whole spring and summer were very dry with the driest period coinciding with the study period, the soils were dried out with a soil humidity of 3 VOl.-% at 10 cm depth measured in Lindenberg (data provided by Frank Beyrich, DWD). Thus, the standard METRAS soil moisture initialisation that assumes rain every two to three days was expected to fail with too high dew point temperatures and too low temperatures. Therefore, soil water availability was decreased for the land-use classes in the model domain. In addition, the initialisation profile of relative humidity was reduced and the initial soil temperature changed based on observations and ECMWF analysis data. All tuning parameters were varied within their uncertainty range, which was determined from the differences between ECMWF analysis and observations available to initialise METRAS. During the study period measured maximum temperatures increase from 30 °C on 10th August 2003 to extremely high 35 °C on 12th August 2003 during the day. Minimum temperatures remain about the same (~ 15 °C); a relative low minimum temperature is measured on 11th August 2003 (< 10 °C). METRAS captures this slope, but overestimates the 2 m air temperatures during the 2nd simulation day due to the ECMWF forcing. The ECMWF analysis overestimated the 2 m air temperatures and thus METRAS was initialised with too high temperatures, but started correcting the air temperatures with on-going simulation time.

In general, the measurements of the DWD dew point temperatures were much more scattered than simulated by METRAS or the ECMWF. Despite assuming very dry soil conditions both, METRAS and ECMWF, simulate dew point temperatures near the upper limit of the measurements resulting in too humid situations at the beginning of the simulations. Also, the METRAS results show that the wind direction is simulated more clockwise than observed by the DWD due to the forcing with the ECMWF analysis data. It can be stated that METRAS simulated the chosen meteorological situation reasonably well, although it was an extremely hot and dry situation and despite the partially unsuitable forcing from the ECMWF analysis.

This sensitivity study reveals the impact of three tuning parameters on the model performance of the near surface values temperature, dew point, wind speed and wind direction. Horizontal homogeneous changes of initial soil moisture, soil temperature and the humidity within their uncertainty range mainly impact the thermodynamic values whereas the dynamic values change only slightly. Reducing the available soil water content has a large impact on the air temperature with differences in the hit rates of up to 20 % compared to the standard METRAS set-up (case 3a). The influence on dew point temperature is more significant than on temperature, since the latent heat flux is directly influenced by the soil moisture and controls the dew point temperature. The air temperature is only indirectly controlled: lower soil moisture enhances the sensible heat flux due to the lower latent heat flux, which can no longer transport so much energy from the soil up into the air. Here, the difference induced by reducing soil moisture is most pronounced during the third simulation day.

The influence of the soil and water temperature on the atmospheric variables was not very pronounced with regard to the hit rates, because induced changes were mostly within the ±2 K accuracy for temperature. However, the conditional quantile plots show the difference in the resulting 2 m air temperatures more clearly when comparing set-ups with soil temperatures taken from ECMWF analysis against initialising the soil from DWD measurements. The hit rates do not capture the differences that well, since the uncertainty range was set to a rather wide ±2 °C. Initialising the whole soil in the model domain from measurements at a single station and interpolating onto the whole numerical grid improved the simulation of 2 m air temperatures. When changing the soil temperature to values given by the ECMWF analysis the simulation results were strongly affected (in this case worsened with ECMWF analysis data). This is visible in the conditional quantile plots.

Reducing the initial relative humidity at all levels has nearly no impact on the dynamic values and causes nearly negligible differences for temperature. It shows some impact on dew point temperatures for the first simulation day. With on-going simulation the initial relative humidity profile is no longer visible in the model and so the solutions converge. They are much more impacted by the soil water content and the lateral boundary values.

According to the current sensitivity study it can be concluded that the initial profiles of e.g. relative humidity have a much smaller impact on model results than the prescription of the soil properties like temperature or moisture. The initial profiles are the less relevant the longer the model is integrated. The lateral boundary conditions are compared to high resolution simulations more important since the percentage of grid points impacted by the forcing is relatively large. However, despite the importance of realistic lateral boundary conditions, the sensitivity study shows that a correct description of the soil properties has a pronounced impact on the METRAS solution. And it can be stated that the surface fluxes greatly affects METRAS' solution for the 2 m thermodynamic values.

Therefore, a correct parameterisation that accounts for the sub-grid scale heterogeneity in the land surface and the soil water contents is very important for a reliable weather forecast.

4 Representativity of in-situ precipitation measurements – a case study for the LITFASS area in North-Eastern Germany

This Chapter was published as:

Sylvia I. Bohnenstengel, K.H. Schlünzen, F. Beyrich (2011a): Representativity of in-situ precipitation measurements in North-Eastern Germany. *Journal of Hydrology*. **400** (3-4): 387-395.

4.1 Introduction

Precipitation is a meteorological variable most difficult to simulate from a physical point of view, and, since it is highly variable in space and time, the evaluation of precipitation amounts from mesoscale numerical models is very challenging. When comparing simulated precipitation amounts with observed precipitation amounts often measured precipitation of single station observations is compared with simulated precipitation amounts representative for a whole grid cell of a numerical model. However, the representativity of such a single station measurement for an area average depends on the location and on the measurement time interval. Neglecting the uncertainty due to measurement errors, the representativity of a single station additionally depends on the type of precipitation: in case of convective rain the areal representativity of a point station is expected to be poorer than for stratiform rain (Joss and Germann, 2000). Besides this, topography influences the precipitation pattern and additionally reduces the representativity of a single station measurement (Buytaert et al., 2006). Since operational networks of rain gauges are often sparser than the applied model resolution the spatial-temporal representativity of a single precipitation measurement station should be known when evaluating mesoscale numerical models with these data. Rain gauges from the operational monitoring networks of the national meteorological services usually provide precipitation amounts with a temporal increment of 6 hours.

The most common approach to evaluate area-averaged precipitation with station measurements are area-to-point and point-to-area methods (Tustison et al., 2001). The area averaged precipitation amount is assigned to the centre of the model grid box and then interpolated to the locations of the gauge network (area-to-point). Then precipitation amounts can be evaluated for these locations. Alternatively, the rain gauges measurements are interpolated onto a regular grid and then area averaged values are computed (point-to-area) and compared to the forecasted precipitation amounts. Since the scale of the gauge network and the simulations are likely to differ a so called representativeness error is introduced, which is scale dependent (Tustison et al., 2001). Without any further knowledge of the spatial-temporal representativity of a single point measurement the validation of aerially simulated precipitation might lead to completely false conclusions concerning the performance of a numerical model. Marzban and Sandgathe (2009) analyse the problem of different scales of measurements and simulations further and suggest comparing precipitation fields in terms of their spatial structures with the help of variograms.

Apart from rain gauge data also radar data could be used for model validation, since they provide areal information on precipitation amounts. However, radar based information on surface precipitation amounts still have some uncertainty, since the original information is radar reflectivity at variable height above ground. These data are normally transferred to surface rain amounts by a regression model, which uses rain gauge measurements, whose spatial and temporal representativity influences the precipitation estimation (Joss and Germann, 2000; Datta et al, 2003). Thus, the knowledge on the representativity of rain gauge data is not only important for validation of numerical models but also for delivering radar-based precipitation data. Datta et al. (2003) state that the resolution of a rain gauges network does not necessarily resolve the variability of the observed precipitation systems. This leads to errors when adjusting the radar based information with the gauges information. Joss and Germann (2000) suggest an uncertainty factor of 2 for single station rain gauge based daily rain amounts in the mountainous Switzerland and stress that the uncertainty increases for shorter integration times.

An estimation of the spatial variability of single point rainfall measurements assists the validation of simulated precipitation patterns and is also of interest for deriving radar based surface precipitation estimates. Gebremichael et al. (2007) underline the importance to determine the geographical area a station derived rain fall statistics is representative for to improve the variability in numerical models and to interpret remote sensing rainfall estimates. Since rain gauge networks usually have a resolution that is too coarse to satisfactorily resolve the precipitation patterns, this paper takes the opportunity to investigate a high resolution rain-gauge network to conclude on the spatial-temporal representativity of single rain gauges.

Many studies have been carried out to investigate the variability of precipitation in space and time. van de Beek at el. (2010) investigated daily rainfall measurements from a rain gauge network in the Netherlands and they found precipitation amounts to be correlated over distances between 50 km and 150 km. Verworn and Haberlandt (2010) found precipitation amounts to be correlated over distances of 57 km for summer rain storm events for Northern Germany. However, their domain includes the mountainous Harz area, and topography is likely to affect the precipitation patterns. Further studies investigated the rainfall patterns in areas with strongly varying orography or in monsoon areas characterised by more heavy rainfall regimes. Burgueno et al. (2005) investigated daily rainfall regimes in Catalonia, Buytaert et al. (2006) applied kriging methods and a variogram analysis to rain gauge data from the very mountainous South Ecuadorian Andes. They determined a strong correlation for an inter-station distance of less than 4 km. Datta et al. (2003) underline the high variability in the rain rate, which can vary by a factor of 10 within a 10-minutes period or within a 2 km distance during the tropical rain measuring mission TRIMM. They state that rain amounts from two gauges being only 15 m apart from each other can differ by more than 10 mm h^{-1}. This is not only due to the spatial variability of precipitation, but also due to the errors occurring when using rain gauge measurements. Michelson (2004) provides detailed information on the systematic correction of gauge observations and points out various measurement errors of the bucket systems. Gebremichael et al. (2007) applied some variogram analysis and additional basic statistics on data from a rain gauge network in an area of 50 km x 75 km in Mexico. They stress that the mix of different rain fall regimes might decrease the correlation of rain amounts being measured more than 30 km apart from each other. Skok and Vrhovec (2006) investigate precipitation amounts of a rain gauge network with respect to area-averaged numerical model output: they try to determine the highest model resolution that makes the comparison of model and rain gauges' precipitation independent of the interpolation method. These authors stress that the comparison is more difficult for higher precipitation amounts and suggest that each grid box within the model should at least contain one or two rain gauges.

Many of the studies mentioned are focusing on tropical regimes or on precipitation events with orographic impacts or they are based on a coarse rain gauges network. In the present study the small scale spatial-temporal variability of precipitation amounts is investigated for a relatively flat terrain. The representativity of rain gauges is determined for different time scales and corresponding uncertainty factors are derived. This study takes advantage of a high resolution rain gauge network set up in a 25 km x 25 km domain in North-Eastern Germany, where orographically induced precipitation can be neglected. More detailed information on the data and the domain is given in Section 4.2. In Section 4.3 the character of precipitation in the investigation area is derived and the results are presented. Conclusions are drawn in Section 4.4.

4.2 Investigation area and data

The precipitation data used in this study were collected in the so-called LITFASS area around the Meteorological Observatory Lindenberg/Richard-Aßmann Observatory (MOL-RAO) of the German Meteorological Service (Deutscher Wetterdienst, DWD, e.g., Beyrich 2004, Beyrich and Mengelkamp, 2006). The LITFASS area is a 25 km x 25 km large region located in the relatively flat, north-eastern part of Germany, south-east of Berlin. The terrain height varies between 40 m above sea level in the south and 130 m above sea level in the north-eastern part. The influence of orography on precipitation can therefore be neglected. Considering the land-use it is a very heterogeneous area with forests dominating the western parts and farmland with different crops in the eastern part, each contributing to about 40-45 % of the whole land use. About 6-7 % is covered by water; settlements cover less than 4 % of the area. Land-use in the LITFASS area is illustrated in Figure 4.1 based on CORINE Land Cover data for Germany (CORINE Land Cover, 2004).

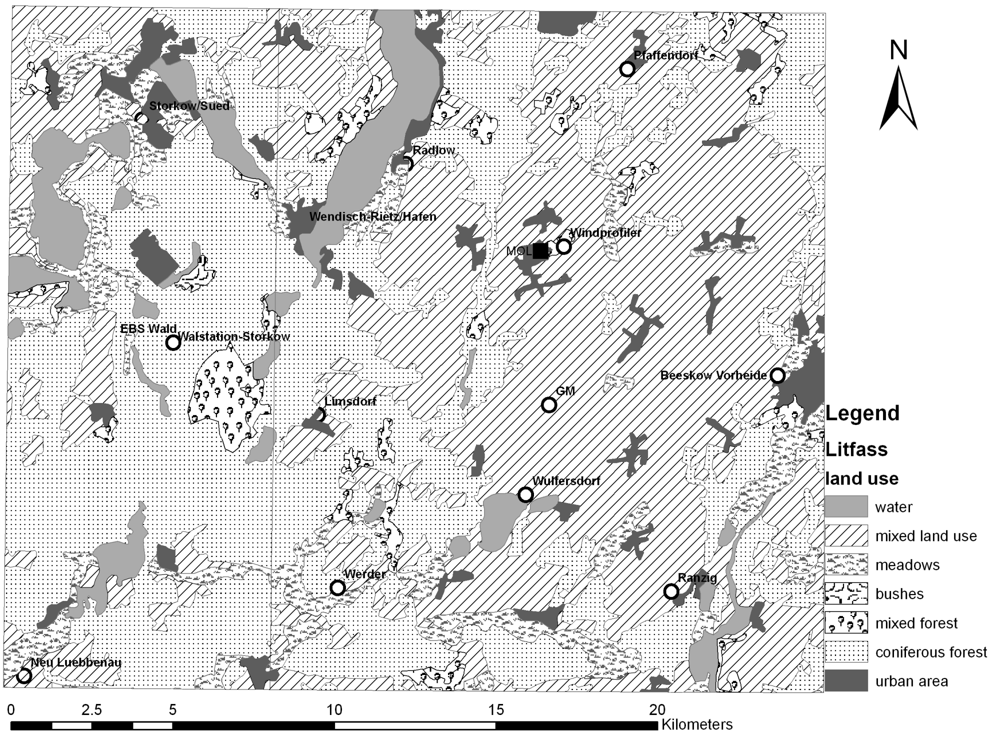


Figure 4.1: Lindenberg area. PLUVIO stations are indicated by black circles with white filling. MOL station is indicated by a black square. Land use data are based on CORINE Land Cover (2004).

Precipitation data are investigated for the year 2003 when a very dense network of rain gauges became available in the LITFASS area (Beyrich and Mengelkamp, 2006). The year 2003 was very dry (annual precipitation sum at Lindenberg 382 mm) compared to the long-term mean (563 mm). Convective situations during the summer period resulted in very local precipitation events triggered by the surface processes in the investigation area. The mean annual precipitation amount for this area typically is about 600 mm. This is based on monthly global gridded data of the "Monitoring Product" of the GPCC (Global Precipitation Climatology Centre) for a period of 30 years (1961-1990). Rudolf (2003) underlines this: according to his study the annual precipitation amount in 2003 was between 66 % and 80 % of the mean annual precipitation amount of the period 1961-1990. Figure 4.2 shows the annual cycle of the 2003 MOL data. The summer precipitation is higher than the winter values as also found for other places in Northern Germany, e.g. Hannover and Berlin (Beckmann and Buishand, 2002) or Hamburg (Schlünzen et al., 2010) as well as south-western Germany (Feldmann et al., 2008). This is typical for areas in the transition zone of maritime and continental climates. Additionally, a higher spatial variability of monthly precipitation data is found in summer compared to the winter months.

Two different types of precipitation data from the LITFASS area were used in the present study. Routine observations at six-hourly intervals (measurements at 00, 06, 12, and 18 UTC) were performed at the WMO synoptic weather station 10393 situated at the Meteorological Observatory Lindenberg (Richard-Aßmann Observatory, MOL), which is located at 52°12’31”N and 14°07’05” E at a height of 98 m above sea level (black square in Figure 4.1). Secondly, a regional precipitation measurement network consisting of 14 rain gauges (PLUVIO type) was operated during the whole year 2003 in the LITFASS area (circles in Figure 4.1). From this network precipitation amounts are available with a temporal resolution of 10 minutes. Data from these two types of measurement systems can be used to derive the representativity of a single rain gauge for an area average.

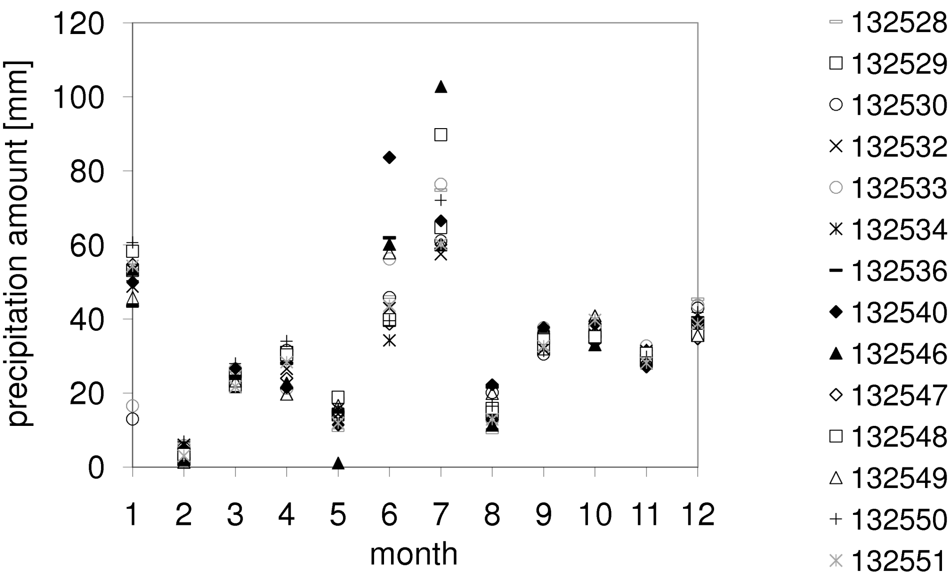


Figure 4.2: Monthly precipitation amounts for all 14 PLUVIO stations in the LITFASS domain for year 2003.

The precipitation amounts from both systems, the MOL single rain gauge and the PLUVIO network, have been recalculated to cover the same time intervals. The MOL data of precipitation sums for 12 hours and 6 hours (according to the WMO synoptic observation rules) were reconstructed by subtracting the amounts of the (first) 6 hour measurement intervals (00 UTC and 12 UTC) from the 12 hour data reported at 06 UTC, and 18 UTC, respectively. When 6 hour data were not available the 6-hourly data were reconstructed by halving the amounts from the 12 hour periods (MOL_6). The 10-minutes precipitation values of the PLUVIO sensors were integrated per station for the same 6-hour periods and then averaged for the LITFASS area resulting in PLUVIO_av6. Additionally, 10-minute arithmetic area averages were calculated considering all stations resulting in the data set PLUVIO_av.

4.3 Characteristics of precipitation events

4.3.1 Precipitation frequency

The histogram (Figure 4.3) derived from hourly precipitation amounts of the 14 PLUVIO stations underlines the high frequency of very low precipitation amounts within the LITFASS domain during the year 2003. While precipitation amounts of 0.1 mm per hour occur nearly 800 times, amounts of 1 mm only occurred about 50 times in 2003. Higher amounts of 3 mm and more were rarely measured. Hence, the precipitation distribution is strongly skewed over this area.

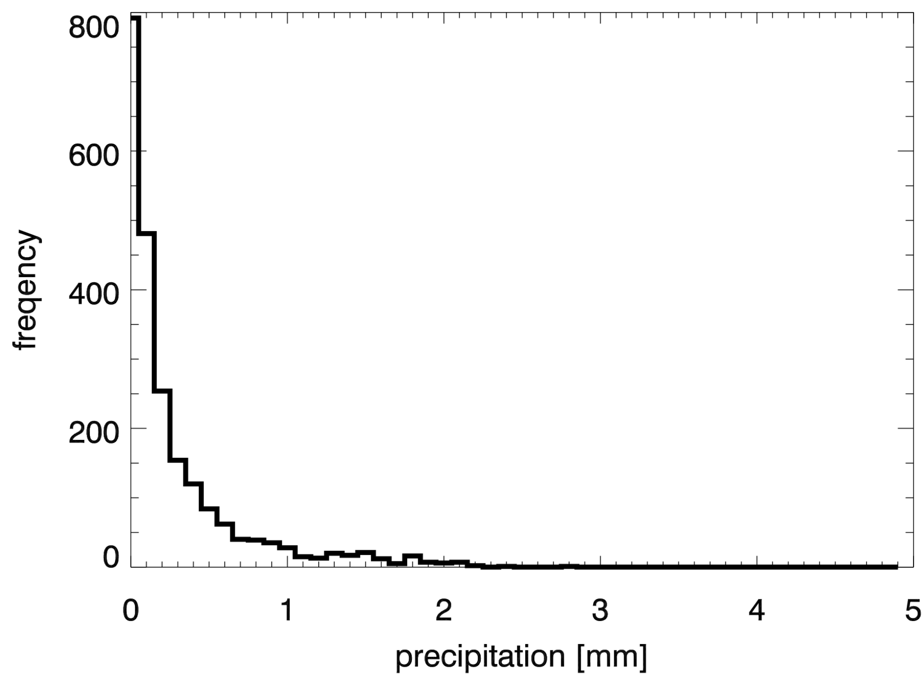


Figure 4.3: Histogram of 2003 hourly precipitation amounts measured by the 14 PLUVIO stations within the Lindenberg area.

From PLUVIO_av the area averaged diurnal cycle of precipitation can be derived for 2003 (Figure 4.4). Two peaks are visible in the mean values, one in the late night and one in the late afternoon. Medians are much lower than mean values, underlining the typically skewed distribution of precipitation. Very frequent events with low precipitation amounts and rare events with intense precipitation characterize the precipitation distribution in this area. The difference between median and mean is largest for the late afternoon. This suggests that more or less rare but extreme precipitation events seem to occur more often in the convectively dominated late afternoon than during times with frontal precipitation. The peaks in the late night might be connected with frontal precipitation. This explanation for the two peaks was also given by Ines Langer (FU Berlin, personal communication and Cubasch et al., 2006) for the station Berlin Dahlem, where she investigated routine observations of 13 years. The convective peak observed in the afternoon was also found in radar images investigated by Walther and Bennartz (2006), who distinguished the precipitation characteristics within the whole Baltic area into frontal and convective precipitation.

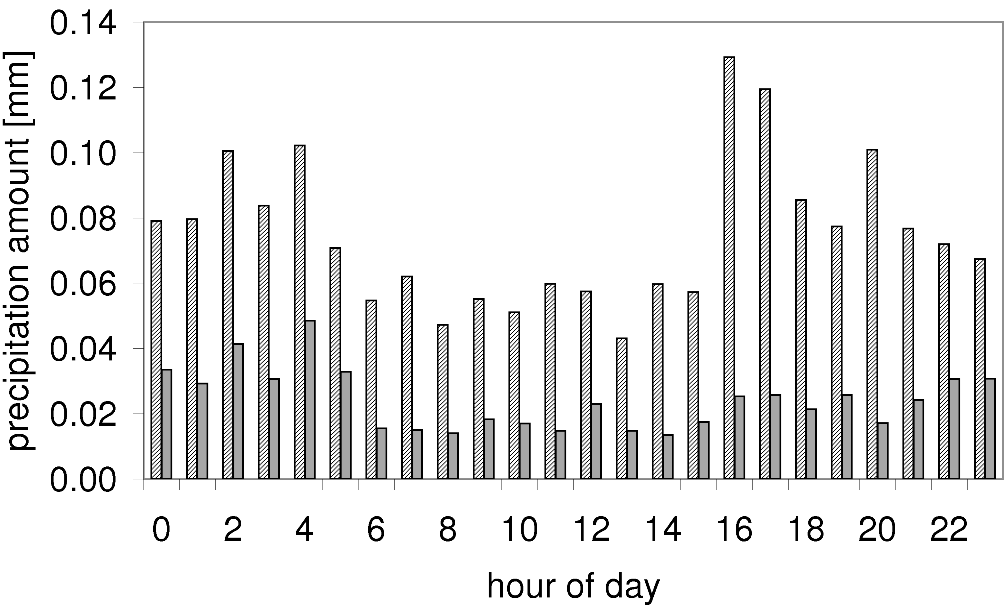


Figure 4.4: Hourly values of precipitation for the LITFASS area during the year 2003, derived from the PLUVIO_av data set. The hourly values are averages over the whole year, resulting in medians (grey bars) and mean values (dashed bars). Dates without precipitation at any rain gauge within the whole network were not considered for calculating the median and mean.

4.3.2 Duration of precipitation events

The PLUVIO_av precipitation data were additionally used to determine the duration of precipitation events. Assuming that a precipitation event is characterized by temporally connected 10-minutes precipitation amounts, which may be interrupted by dry events of less than one hour, precipitation events were constructed. The beginning of a dry episode lasting one hour or more was used to define the end of one precipitation event. If precipitation occurred again some time later this was defined as the beginning of a new precipitation event. By this method, the single 10-minutes PLUVIO_av amounts are merged into events. The percentage of precipitation events lasting less than one hour is 60 % (Figure 4.5). Events of more than one and up to two hours make up 16 % of the precipitation events; longer lasting precipitation events are even less frequent. Most of the precipitation events last up to three hours (84 % in total). 94 % of the annual precipitation events last less than 6 hours, which seems to be the maximum impact time for a precipitating system in this region. The remaining 6 % belong to rare events, some of them last longer than a whole day. These findings on the character of precipitation within the Lindenberg area fit well with the findings from Weusthoff and Hauf (2008) analysing 5-minute radar composites over the whole of Germany. They showed that about 80 % of the post-frontal precipitation cells only have a lifetime of up to 35 minutes with an exponential decay of the lifetime.

We calculated the annually averaged hourly precipitation amount dependent on the duration of the events (black crosses in 4.5). A clear trend is not visible. However, hourly precipitation amounts increase with increasing duration up to 6 hours. For longer durations the hourly precipitation amount is highly variable – partly due to the low number of events. While hourly precipitation amounts increase with increasing duration of the precipitation event for events of up to 5 hours, longer lasting precipitation events show average hourly amounts between 0.23 and 0.95 mm hour^{-1}. As mentioned earlier, 60 % of all precipitation events last up to one hour. However, the corresponding precipitation amount covers as little as 1.5 % of the annual precipitation sum (Figure 4.5). The maximum contribution to the annual precipitation amount is connected with events lasting between 2 and 5 hours, contributing to 46 % of the annual precipitation. About 52 % of the annual precipitation is covered by events with durations of up to 6 hours. The 6 % of situations that last more than 6 hours contribute to 48 % of the annual precipitation. These rare events contribute in nearly the same way to the annual precipitation as the very frequent events lasting only up to six hours and are highly variable in intensity. The two most intense hourly precipitation amounts correspond to events lasting 13 and 14 hours, which occurred in July and in November 2003. Some seven and eight hour events occurred randomly over the year. Two events lasting 17 hours occurred in March and in December 2003. In the LITFASS area very few long-lasting events contribute with nearly the same quantity to the annual precipitation as the 94 % of all shorter precipitation events lasting up to 6 hours.

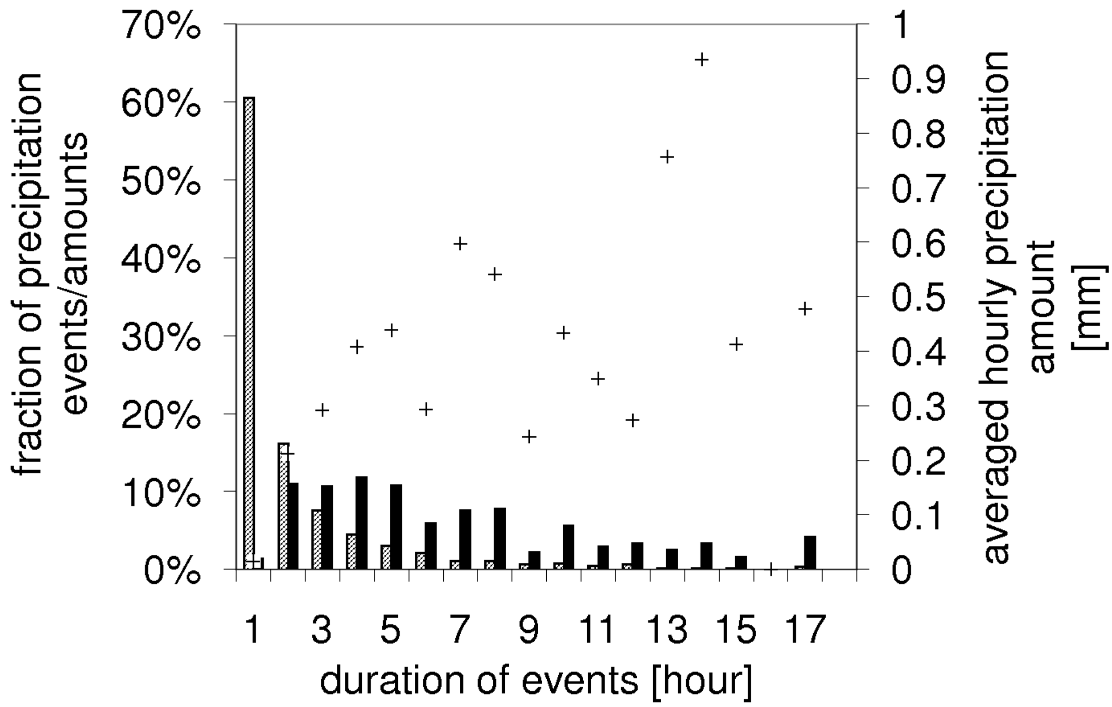
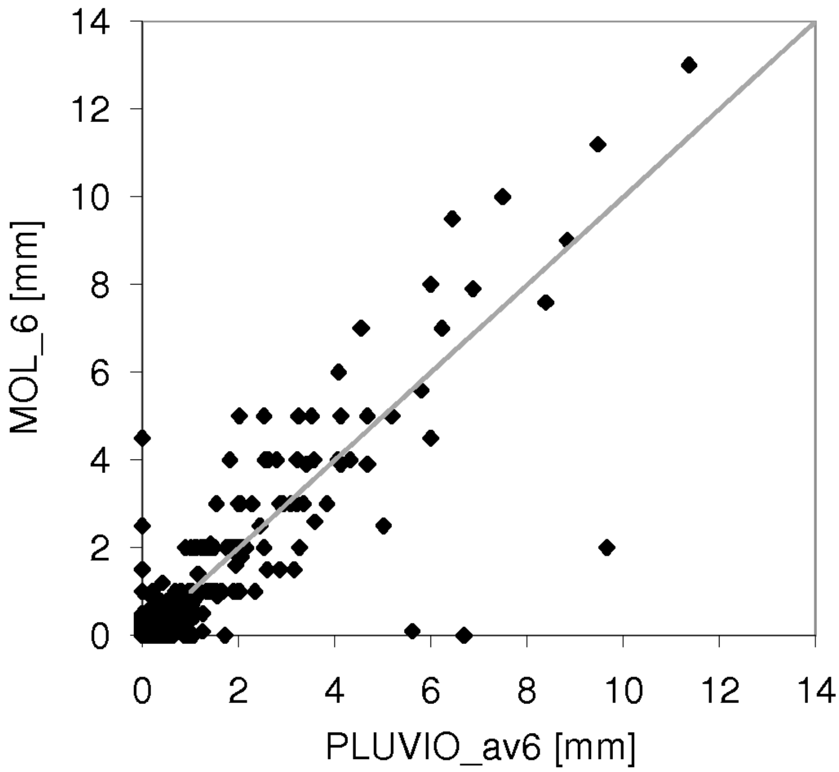


Figure 4.5: Fraction of precipitation events (dashed bars) and fraction of precipitation amounts (black bars) dependent on the duration of the event based on PLUVIO_av data for 2003. The averaged hourly precipitation amount is given by the black crosses.

4.3.3 Scatter of 6 hour precipitation amounts

The six-hour integral values of precipitation are given in Figure 4.6a and Figure 4.6b for the routine observational data MOL_6 and for PLUVIO_av6, since this is the time increment most often provided by the routine observations. Due to spatial averaging, PLUVIO_av6 clearly shows more often lower precipitation values than MOL_6, which instead shows more zero precipitation amounts (Figure 4.6b). There are also some values of MOL_6 precipitation data that are much larger than the PLUVIO_av6 data: in 6 % out of all cases MOL_6 collected precipitation when PLUVIO_av6 did not. The scatter is a result of the spatial and temporal variability of precipitation. The spatially averaged PLUVIO_av6 precipitation amounts have less variability than the MOL_6 point measurements, since the spatial averaging reduces the variability to some degree. This is dependent on the scale of the spatial domain and variability of the original precipitation field. As can be seen from Figure 4.6a, differences between PLUVIO_av6 and MOL_6 exist for all amounts of precipitation, not only for small values (Figure 4.6b). The precipitation is too local to collect the same amount within 6 hours at one site (MOL_6) as collected – in the average – with several samplers in an area of 25 km x 25 km (PLUVIO_av6).

(a)



(b)

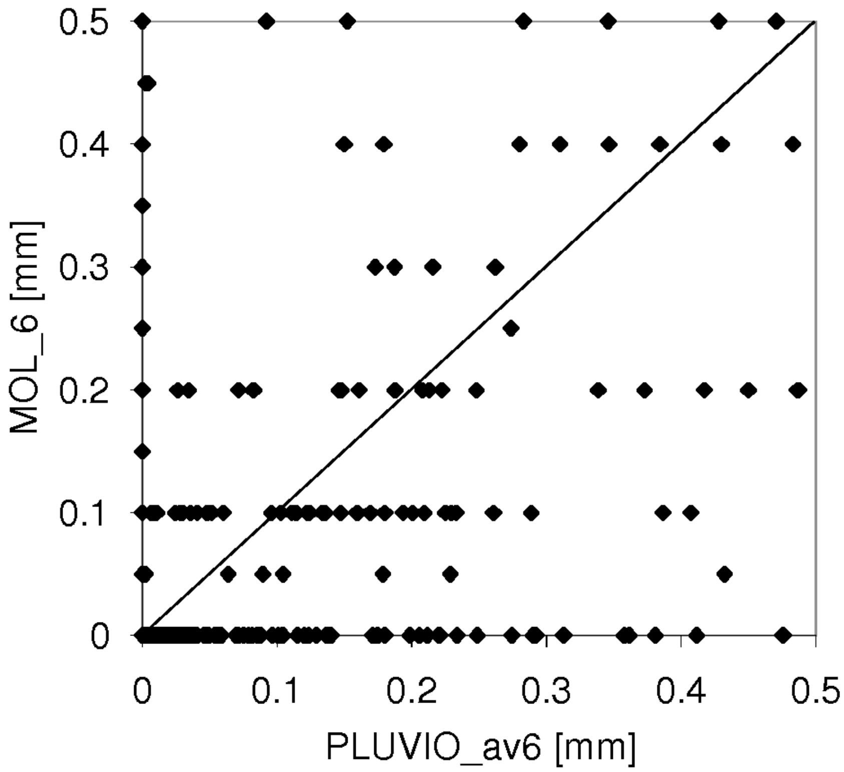


Figure 4.6: Six-hour precipitation integrals from the routine observations MOL_6 and the PLUVIO_av6 data for year 2003 (a). Enlargement of the scatter diagram for low precipitation amounts (b).

4.3.4 Spatial representativity of in-situ measurements

From the previous sections it is obvious that a large spatial-temporal variability for in-situ measurements of precipitation exists. It can be expected that the spatial representativity depends, amongst others, on the integration time of the measurements. Therefore, an experimental spatial semi-variogram analysis of the PLUVIO data set is performed for three characteristic time scales: hourly, daily and weekly precipitation amounts are calculated for each of the 14 PLUVIO stations from the 10-minutes values to simulate different integration times of rain gauges and to study the impact of sampling duration on the spatial representativity.

The experimental spatial semi-variogram, in the following abbreviated as variogram, is an estimator for the variability of geostatistical data in terms of their interstation distance. It assumes that stations close together tend to behave more similar than stations far apart from each other (Pohlmann, 1993). The variogram is only a function of the distance between station pairs and does not consider their actual location. This assumption fails for mountainous areas, where nearby stations separated by a hill might be totally uncorrelated and stations further away from each other but on the same side of a hill might behave more similar. Since the study is carried out in relatively flat terrain, the location of the stations can be neglected. Due to the small domain size isotropy is assumed as well. To derive the variogram each of the 14 stations of the PLUVIO network is paired with each station leading to 92 measurement pairs with distances ranging from 1 km up to 25 km. These measurement pairs are then grouped into distance classes with a bin width of 1 km. The variograms are calculated for each distance class for the three time intervals of hourly, daily and weekly precipitation amounts, respectively. The resulting so called “climatological” variogram (Grimes and Pardo-Iguzquiza, 2009), which is representative for the whole year 2003, is then calculated as the average of the individual variograms following eq (4.1) based on Skoien et al. (2003) and Grimes and Pardo-Iguzquiza (2010).

$$\gamma(d) = \frac{1}{T}\sum_{j=1}^{T}\frac{\sum_{i=1}^{N_j(d)}\left[Z(x_i,t_j)-Z(x_i+d,t_j)\right]^2}{\sigma_j^2 * 2N_j(d)} \tag{4.1}$$

$Z(x_i, t_j)$ and $Z(x_i+d, t_j)$ are the amounts of precipitation in mm measured at time t_j at station x_i and at station x_i+d, which is located at the distance d from x_i. The resulting variogram (Figure 4.7) is the average of the variograms for each time interval t_j with T in equation (4.1) indicating the number of spatial semi-variograms over which was averaged. $N_j(d)$ is the number of measurement pairs during each time interval t_j with a separation distance of d, where precipitation was measured above a given threshold of 0.1mm at least at one of the two stations. The index i runs over all time intervals of the whole year 2003. The variogram for each time interval t_j is normalized by the spatial variance σ^2_j of the precipitation for that time interval t_j, which is derived from the PLUVIO data. The resulting "climatological" variogram is representative for the variance of the precipitation of the whole year 2003. Seasonal influences are not taken into account when computing the "climatological" variogram. A time-lag between the measurements is not accounted for, since the domain is rather small.

The offset, called nugget, at a station distance d of 0 m describes the random noise of stations nearby and describes the variability at scales smaller than the sampling size. Theoretically the nugget is zero for a station distance d of 0 m. Due to the occurrence of different types of precipitation within the data set for 2003 the upper limit of the variogram exceeds 1 even though the individual variograms are normalized by the spatial variance for each individual variogram.

The variogram indicates the strength of the decorrelation of precipitation amounts of stations with interstation distance d. The variogram reaches the so called sill, where the distribution levels out, at a certain range. Measurement stations with a distance larger than the range are assumed to be uncorrelated. A more detailed description of the variogram analysis is beyond the scope of this paper, but can be found in Pohlmann (1993), Haberlandt (2007), Skoien et al. (2003) and Grimes and Iguzquiza (2010).

Figure 4.7 shows three variograms for the spatial representativity of the PLUVIO precipitation amounts based on hourly, daily and weekly measuring intervals, respectively. Additionally, exponential fitting functions are plotted for the three experimental variograms. All three variograms have in common a nugget of 0.5 indicating the variability at distances smaller than the sampling distance. The slope for the weekly variogram at distances up to 10 km shows a smaller gradient than the slopes for the daily and hourly variograms. All three variograms do not show a distinct sill, which would clearly indicate that the maximum level of variability is reached. The range, over which precipitation amounts are increasing, is similar for all measurement times. The range is about 14 km for weekly measurement intervals and about 11 km for hourly and daily measurement intervals following the exponential fitting functions. Longer integration times seem to result in a slightly larger areal representativity of a single station. However, the large scatter makes it difficult to determine the exact range. But it can be concluded that weekly precipitation amounts do not seem to be representative for the whole LITFASS area (25 km x 25 km) when analysing an annual variogram. The same conclusion holds for hourly and daily values. However, they are correlated for distances of the order of 10 km.

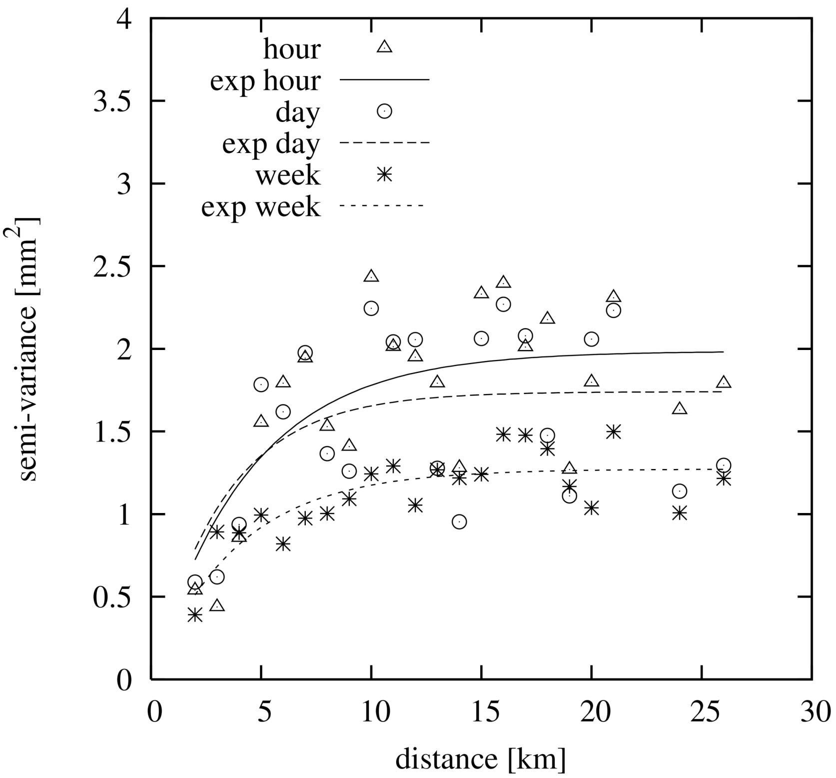


Figure 4.7: Experimental semi-variogram based on the 14 PLUVIO stations for threshold values greater than 0.1 mm for (a) hourly data, (b) daily data and (c) weekly data. The semi-variograms are fitted with an exponential function above a threshold of 0.1 mm.

The variability increases strongly from almost 0.5 to 1.7 according to the exponential fitting function. Note that these values are normalized by the spatial variance of the precipitation field and, therefore, they do not indicate the absolute variability of the precipitation field for hourly values. However, the sill is accompanied by a large scatter of the variance values: beyond an inter-station distance of 11 km the variogram values vary between 1.3 and 2.5 for different distance classes (Figure 4.7). This is caused by calculating variograms for different types of precipitation. Convective rain results in a much larger variability in the measurement field than stratiform rain for instance and has a smaller spatial scale indicated by a shorter range of the variogram. Additionally, measurement errors and some random noise affect the variability. Since the wind direction was dominantly from south-west and west during 2003 with a secondary maximum for north-east to east winds (Beyrich et al., 2004), there might be less variability and a higher spatial representativity for south-west to north-east oriented station pairs. The variogram solely depends on the distance and due to the limited amount of data no attempt is made to include the orientation of the correlated pairs of stations. The variograms underline that the degree to which the variability in the variogram is reduced depends on the measurement time scale. Increasing the measurement time interval from hourly to daily and weekly precipitation amounts reduces the spatial variability of the variogram considerably. The variogram values for daily precipitation amounts range between 0.55 and 2.3 (Figure 4.7). The variability for weekly precipitation amounts is reduced to values between 0.4 and 1.4 (Figure 4.7). Compared to hourly amounts the width of the scatter is a third for daily and weekly values. Additionally, the correlation is stronger for longer measurement times like daily and weekly precipitation amounts.

The range of the three variograms determined in the present study is about 5 times shorter than the range of 57 km found by Verworn and Haberlandt (2010) for summer precipitation in Northern Germany. Verworn and Haberlandt's (2010) rain gauges network is coarser than the network of the present study and covers a considerably larger area including the mountainous Harz region, where precipitation patterns are additionally impacted by orographically induced phenomena. The present study is based on a dense rain gauges network over a small domain in flat terrain a few 100 km east of the area considered by Verworn and Haberlandt (2010). Here the continental influence is stronger and orography effects are small. Further, the precipitation data sets differ considerably in their precipitation intensity. While Verworn and Haberlandt (2010) base their variogram analysis on flood events with large rainfall amounts, the present data set includes all precipitation events. Longer lasting rain events are rare in the Lindenberg area in 2003 (Section 4.3.2). It is therefore likely that the results from Verworn and Haberlandt (2010) are representative for the more extreme precipitation events while our analysis represents all events. Additionally, the domain size for the present study is limited to 25 km x 25 km. Therefore, the exponential fit does not capture precipitation patterns beyond that range.

The short range of the order of 10 km is accompanied by a relatively large scatter for the hourly and daily variograms around the sill, which indicates a large variability in the present data set even on small scales. This is consistent with findings from van de Beek et al. (2010) who demonstrate a large variability in the fits for variograms for summer precipitation in the Netherlands based on daily precipitation amounts. However, the analysis for Lindenberg demonstrates a reduced variability when increasing the integration time from hours to a day and a week, the range of the variograms, however, does not increase with integration time.

Since precipitation amounts seem to be correlated over distances of the order of 10 km, the correlation of occurrence of precipitation is investigated in terms of the critical success index CSI.

$$CSI(d) = \frac{1}{T}\sum_{j=1}^{t}\sum_{i=1}^{Nj(d)} \frac{A(x_i, x_i + d, t_j)}{A(x_i, x_i + d, t_j) + B(x_i, x_i + d, t_j) + C(x_i, x_i + d, t_j)} \qquad (4.2)$$

The stations are paired similar to the variogram analysis. $A(x_i,x_i+d,t_j)$ counts the number of events, when both stations measure precipitation, $B(x_i,x_i+d,t_j)$ counts the number of events where rain occurs only at station x_i and not at station x_i+d and $C(x_i,x_i+d,t_j)$ counts the number of events where no rain occurs at x but rain occurs at x_i+d for time interval t_j. The CSI(d) for the whole year 2003 is then calculated as the average over the individual spatial CSI(d, t_j) with T being the number of spatial CSI(d, t_j). Note that no-rain events at both stations are not considered. The CSI ranges from 0 to 1 with 1 indicating rain at both stations all the time. 0 indicates rain at one station while the other station in distance d measures no rain and 0.5 indicates rain at one station half the time which is not accompanied by rain at the other station (Gebremichael et al., 2007).

Figure 4.8 demonstrates how the CSI's variability is decreasing with increasing integration times. The CSI does show little dependence on station distance for any of the three integration time intervals. Hourly CSI values lie between 0.7 and 0.9 with some minimum in the 5 km to 15 km range. In contrast, daily CSI values lie between 0.8 and 0.9. Weekly CSI values approach a value of 1 indicating a perfect correlation (Figure 4.8). The spread of the CSI decreases more than the spread of the variogram with increasing integration time. This indicates that the areal representativity of the occurrence of rain events is larger than that of the rain amounts. In other words: if a single station within the PLUVIO area measures precipitation, there is a high probability that all sites in the investigation area might experience some rain within a day, but the amount can remain different within that period.

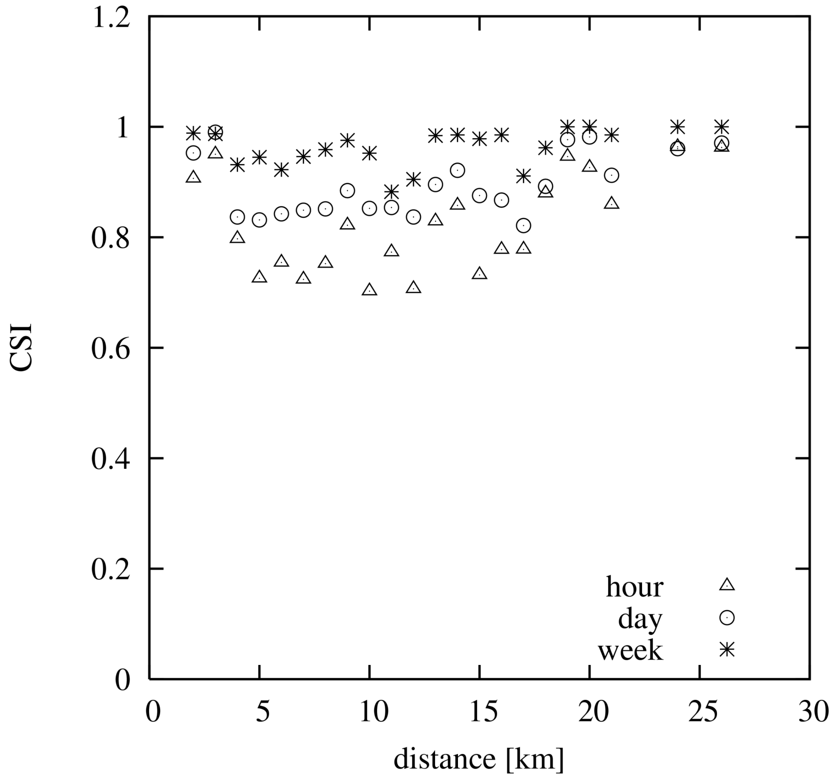


Figure 4.8: Critical success index (CSI) for hourly (cross), daily (triangle) and weekly (circle) precipitation values.

This is an important result for verifying precipitation amounts simulated by mesoscale numerical models for instance. When only comparing the occurrence of precipitation the CSI suggests a large spatial representativity of a single gauge measurement and might be used for scales even larger than 25 km and a small representativity error can be assumed. However, verifying precipitation amounts on small scales is challenging since the previous analysis suggests a large variability on spatial scales of about 10 km. The variogram analysis suggest large representativity errors when scaling the single station measurements up to area averaged values (point-to area) or interpolating the area-averaged simulated values into the locations of the gauges (area-to-point). This analysis strongly recommends using the scale information of the variograms for evaluation purposes. On top of identifying the spatial variability the variogram analysis can form the basis of a kriging algorithm to predict the precipitation amounts at any location in the domain accounting for the scale issues.

4.3.5 Temporal representativity of in-situ measurements

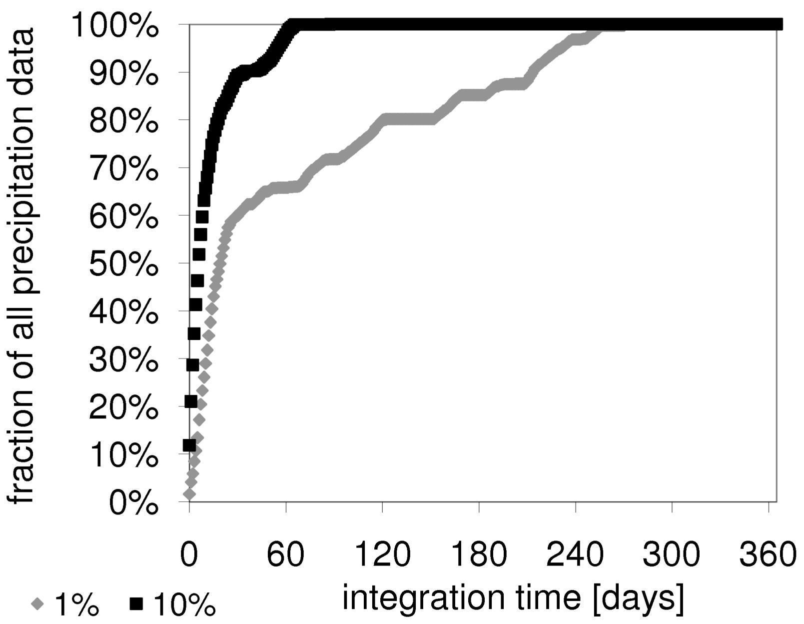
The previous section showed that a reliable spatial representativity of single station precipitation measurements partly depends on the integration time of the measurements. For evaluating simulated area-averaged precipitation amounts routine observations from synoptic weather stations are often used. Here, the opportunity is taken to estimate the spatial-temporal representativity of such a station by comparing MOL_6 data with the area-averaged PLUVIO_av6 data that are used as a substitute for an area averaged precipitation amount. The two precipitation data sets MOL_6 and PLUVIO_av6 of 2003 are each integrated in time to investigate the dependence of differences in precipitation amounts between the single station data MOL_6 and the area average PLUVIO_av6 on integration time. The integration is started every six hours at T_0 throughout the whole year 2003 and ends at time T_{R_α} when (eq. 4.3) is fulfilled.

$$\left| \frac{\int_{T_0}^{T_{R\alpha}} \mathrm{MOL\_6}(t)\,dt - \int_{T_0}^{T_{R\alpha}} \mathrm{PLUVIO\_av6}(t)\,dt}{\max\left(\int_{T_0}^{T_{R\alpha}} \mathrm{MOL\_6}(t)\,dt, \int_{T_0}^{T_{R\alpha}} \mathrm{PLUVIO\_av6}(t)\,dt \right)} \right| \leq \alpha\% \tag{4.3}$$

$\alpha = \left\{ \begin{matrix} 10 \\ 1 \end{matrix} \quad \text{for} \quad \begin{matrix} 10\% \\ 1\% \end{matrix} \right\}$ denotes the relative difference in integrated precipitation amounts. T_{R_α} is the integration time for which the relative differences of the two integrated precipitation amounts are below 10 % (T_{R10}) or 1 % (T_{R1}). In other words, the integration time scale is calculated, for which a single station measurement matches the area-averaged measured precipitation within an allowed uncertainty range of either 10 % or 1 %.

Figure 4.9a and Figure 4.9b show the accumulated fraction of precipitation data differing less than 10 % or 1 % as function of integration times. For an integration time of 30 days the fraction is nearly 90 % and 60 %, respectively (Figure 4.9b). This means, monthly MOL single station precipitation measurements are representative for the LITFASS area averaged precipitation amounts in 90 % of all cases with precipitation when accepting 10 % difference in the precipitation amounts. When only accepting 1 % difference in precipitation amounts between the single point measurement and the area average, an integration time of 214 days is needed to yield a representativity of 90 % between the single MOL station and the area averaged PLUVIO amounts. A representativity of 100 % is achieved for integration times of more than 88 days (10 % deviation) and 271 days (1 % deviation). For longer integration times of nearly two months 95 % of single station and area average precipitation values agree within 10 %. When allowing a deviation of 10% between the area averaged values PLUVIO_av6 and the in-situ data MOL_6, 90%, 55% and 12% of the monthly, weekly and daily integral values are within the 10% deviation.

(a)



(b)

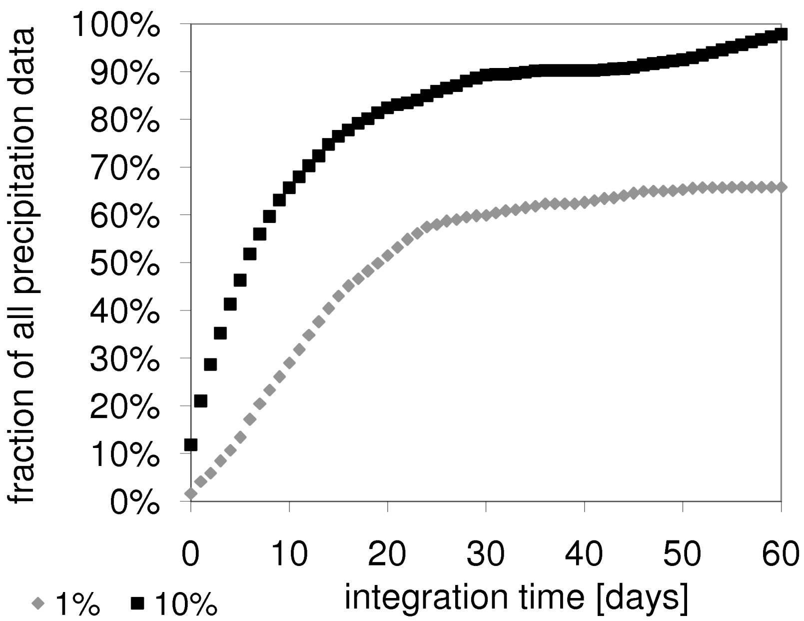


Figure 4.9: Percentage of observed precipitation amounts PLUVIO_av6 and MOL_6 differing less than 10 % (black dots) or 1 % (grey dots) for different integration times for (a) up to 360 integration days and (b) up to 60 integration days.

The annual precipitation values of both data sets agree very well. The PLUVIO_av6 data give an annual precipitation amount of 384 mm in 2003, the reconstructed 6-hour data series MOL_6 also results in 384 mm. This agreement underlines that annual precipitation values from single stations are at least in the study region representative for an area of 25 km x 25 km. The factor within which daily, weekly and monthly integrated precipitation amounts measured by the single MOL station are representative for the PLUVIO area average was derived. Therefore, the deviation of MOL amounts from PLUVIO amounts was calculated as follows: |MOL – PLUVIO|/PLUVIO for daily, weekly and monthly integrated precipitation amounts. The integration started every 6 hours over the whole year of 2003. The resulting deviations were rank ordered to determine the 95^{th} percentile. The factor |MOL-PLUVIO|/PLUVIO was derived from the deviation for the 95^{th} percentile representing the upper limit of deviation for 95 % of the data. Comparing results for monthly, weekly and daily integral values of the two datasets, 95 % of the data are within a factor of 1.4, 2 and 3.3, respectively. This can be interpreted as an uncertainty factor for the precipitation amount measured at a single station. For the region investigated it can be concluded that single station daily precipitation amounts are representing precipitation of an area of 25 km x 25 km within a factor of 3.3 and are thus very uncertain. This has to be taken into account when comparing simulated precipitation values with single station data. The comparison would be much more reliable when using weekly or monthly data.

This analysis shows again the well known spatial variability of precipitation. Since the integrated precipitation amounts from both data sets differ notably for short integration periods of a few days to a few weeks, single station data like MOL_6 cannot be taken as representative for an area of 25 km x 25 km. However, when investigating time periods of 1 month and more at least 90 % of the single station MOL data agree within 10 % with the average precipitation amount of the PLUVIO area. From the results it can be concluded that measured and simulated precipitation values should only be compared when using several weeks' integral values, or when area-averaged precipitation values are available. As an alternative approach the large uncertainties in the data need to be considered in the evaluation.

4.4 Conclusions

The overall aim of this paper is to determine the temporal and spatial scales a single station rainfall measurement is representative for. Two data sets were available for the relatively flat 25 km x 25 km large LITFASS area for the whole year 2003. A network consisting of 14 registrating rain gauges provides 10 minutes averaged precipitation amounts and is analysed in comparison to routine station rain gauge data available at 6 hours time resolution. The precipitation in the LITFASS area is characterized by a skewed distribution with very frequent low precipitation amounts. 94 % of all precipitation events last up to 6 hours and contribute to about half of the annual precipitation amount.

Application of an experimental semi-variogram analysis and the computation of the critical success index CSI show a strong spatial and temporal representativity for the occurrence of precipitation. The occurrence of precipitation detected at a single station was nearly always representative for a distance of up to 25 km – even for short time scales like an hour the correlation values were still high. Increasing the time scale increased the representativity significantly.

This study provides new insights into the spatial representativity of a single rain gauge within a high resolution gauge network over flat terrain. The spatial representativity of the precipitation amounts depends only slightly on the temporal scale within the domain size of 25 km. Hourly, daily and weekly precipitation amounts are correlated over distances of the order of 10 km. In-situ measurements of precipitation amounts with a temporal increment of an hour are rarely correlated over distances larger than 11 km. And even weekly precipitation amounts were hardly correlated for a distance of more than 14 km. However, the overall variability is reduced considerably for larger integration times.

While the length scales of representativity determined here are considerably smaller than the length scales determined from coarser resolving measurement networks for the Netherlands and the mountainous Harz region in Northern Germany, the findings are consistent with the conclusions drawn by Buytaert et al. (2006) who determined a correlation length scale of up to 4 km for daily precipitation amounts.

For longer integration times the uncertainty of the single precipitation sums decreases from a factor of 3.3 for daily values, to a factor of 2 for weekly amounts, and to a factor of 1.4 for monthly amounts when used for an area of 25 km x 25 km. 95 % of the data are within these before mentioned factors. Joss and Germann (2000) suggest a factor of 2 for single station precipitation amounts but it is not clear for how large an area. Thus, based on our findings at least monthly precipitations amounts should be used for validating regional climate models, if in-situ data are used. For the validation of hourly precipitation amounts predicted by weather forecast models with a horizontal resolution within the kilometre range a high density rain gauge network is necessary or variograms should be used for interpolating gauge data to capture the scale dependence of the representativity and the difference in scale of the measurements and the area averaged forecasts. Otherwise only the occurrence of precipitation can be validated against rain gauges measurements and radar derived products should be given preference.

In this study the high-resolution PLUVIO data were only investigated for the very dry year 2003 and a small domain. The study was carried out for a relatively flat terrain in North-Eastern Germany. Nevertheless, it can be expected that the main findings are not restricted to this area but are valid for areas with similar precipitation characteristics. The precipitation characteristics with very often low precipitation amounts was also found for the Hannover area of Germany based on radar measurements (Weusthoff and Hauf, 2008 and personal communication). To investigate the small scale differences in precipitation patterns an even finer spatial resolution of the measurement network would be desirable. It also has to be mentioned that the MOL data were compared with an area-average of the PLUVIO data that was simply derived by averaging all precipitation amounts, independent of the site positioning.

5 A locality index to investigate the influence of surface fluxes on precipitation events

This chapter will be submitted to Quarterly Journal of the Royal Meteorological Society:

Bohnenstengel, S.I. and K.H. Schlünzen (2012): A locality index to investigate the influence of surface fluxes on precipitation events, *Q. J. R. Meteorol. Soc.*, to be submitted.

5.1 Introduction

Forecasting precipitation is very challenging, since many physical and chemical atmospheric processes are involved in precipitation generation. For global models the role of the land surface-atmosphere interaction and the significance of evapotranspiration for the generation of precipitation have widely been studied. In global models evaporation is the source of water in the atmosphere since no lateral boundary conditions exist. In case of limited area regional models the impact of land surface characteristics on the generation of precipitation within the model domain is partly depending on the extension of the domain. In case of short-term forecasts the precipitation might be totally driven by advection. In these situations the forcing at the lateral boundaries is much more important than the accurate description of the surface processes. However, in locally driven situations like convective conditions, evaporation might still play a major role for the generation of precipitation within the model domain. Then, a good initialisation of the soil moisture and a correct description of the surface fluxes are essential for a reliable precipitation forecast.

The nature of precipitation has been investigated in various ways, having as a common approach the attempt to classify precipitation situations and the corresponding relevant processes. Classifications are also used in cloud parameterisations for models (e.g. Arakawa and Schubert, 1974) or for comparing model results and measurements (Raible et al., 1999). Several methods exist to classify precipitation situations, but all of them are rather complex. Houghton (1950, 1968) classified precipitation events with respect to the nature of clouds and subdivided precipitation events into stratiform and convective ones. Huschke (1959) determined the origin of precipitation by associating precipitation amounts to cloud types. Houze (1993) distinguished between stratiform and convective precipitation by using the vertical wind from 4-D radar imagery as a criterion. Rice and Holmberg (1973) determined the contribution of convective precipitation events to total precipitation on the basis of monthly or annual precipitation data by introducing a threshold value. Studies by Llasat (2001) extended this approach and describe and classify not only monthly and annual precipitation amounts but also the convective character of precipitation for shorter precipitation periods as a time series. These last two methods use predefined threshold values for precipitation, by which the fraction of convective precipitation events is extracted from the precipitation time series. These thresholds can be defined in a more or less complicated way. One approach is to define lower and upper values for precipitation intensity as thresholds and to consider all precipitation amounts within these thresholds to be of convective character. This method was verified for a specific geographic region (Watson et al. 1982) but is not generalised.

Baldwin and Lakshmivarahan (2002) focus on the spatial structure of precipitation. They developed an automated classification technique based on the spatial structure of precipitation. For the classification they use a generalized method of moments estimation technique to determine the parameters of the gamma distribution fitting the observed distribution of rainfall. Hierarchical clustering analysis is performed with these parameters to classify precipitation into convective and non-convective.

The mentioned classification methods use either long time series of data or information on the spatial structure of precipitation or clouds. Their online calculation in numerical models is time consuming and – when long time series are needed – not possible. No classification exists that relates the precipitation at a specific site to the meteorological situation observed a few hours before the precipitation event. Once established, such an index could be used to forecast precipitation in the coming few hours. It can also be applied to determine the precipitation probability dependent on the meteorological situation and thus might help to develop and utilize surface flux parameterisations tailored to improve precipitation forecast.

In this paper we develop indices to detect and describe those meteorological situations most easily, in which the surface fluxes are relevant for precipitation in a limited area model domain. These indices shall be easy to apply in models and later help to intelligently determine the complexity of a surface flux parameterisation scheme to be used in a model. In a first step the indices are derived from measured data. Observed meteorological situations are classified objectively with respect to the strength of the turbulent exchange between the surface and the boundary layer and their link to precipitation. The resulting classification shall give insight into the relevance of surface fluxes for later precipitation and may also be used in numerical models for surface flux parameterisation control. Therefore, the classification shall be kept as simple as possible and its online calculation in a model shall be possible.

Four different indices are introduced and calculated in this paper to classify meteorological situations (Section 5.3). They are determined from data for years 2002 (summer precipitation above average) and 2003 (summer precipitation below average) and calculated using routine observational data (Section 5.2) taken at Lindenberg (Germany) and results of a 1D numerical model (Section 5.3.5). To investigate the relevance of the indices for forecasting precipitation probability and precipitation amounts, the calculated indices are related to precipitation data that have been analysed for the present domain (Section 5.4). Conclusions are drawn in Section 5.5.

5.2 Investigation area and used data

The precipitation data are collected in the area of Lindenberg (Figure 5.1) in a 25 km x 25 km large area hereafter called "Lindenberg column" around the Meteorological Observatory Lindenberg/Richard-Aßmann Observatory (MOL-RAO) of the German Meteorological Service, located in the relatively flat, north-eastern part of Germany, south-east of Berlin. The area has the advantage that the influence of topography effects on precipitation is small. The terrain height in the Lindenberg area varies between 40 m above sea level in the south and 130 m above sea level in the north-eastern part. The land-use is quite heterogeneous with forests dominating the western parts and farmland with different crops in the eastern part, each contributing to about 40-45 % of the whole land-use. About 6-7 % of the area is covered by water; settlements are found in less than 4 % of the area. Land-use information in Figure 5.1 is based on CORINE Land Cover data for Germany (CORINE Land Cover, 2004).

The precipitation data used in this study are measured by the DWD. A high resolution precipitation network consisting of 14 PLUVIO stations was set up in the Lindenberg column providing 10 minutes averaged precipitation amounts throughout the whole year 2003 (Bohnenstengel et al., 2011a). For the present study 6-hourly area-averaged precipitation amounts were calculated from the 14 PLUVIO stations following Bohnenstengel et al. (2011a) and will be referred to as PLUVIO_av6 in the following. In addition, sounding data and 6 hour integrated precipitation amounts from a routine station operated by the DWD at the Lindenberg Meteorological Observatory (MOL) (Figure 5.1) are available for the very dry year 2003 and the very wet year 2002. The MOL data will be referred to as MOL_6 in the following. All precipitation data have been extensively analysed by Bohnenstengel et al. (2011a). Data of the two very extreme years 2002 and 2003 are analysed to investigate, if the relationship between a new locality index and the investigated precipitation is robust. Figure 5.1 illustrates the distribution of the PLUVIO stations and the position of the routine station at MOL.

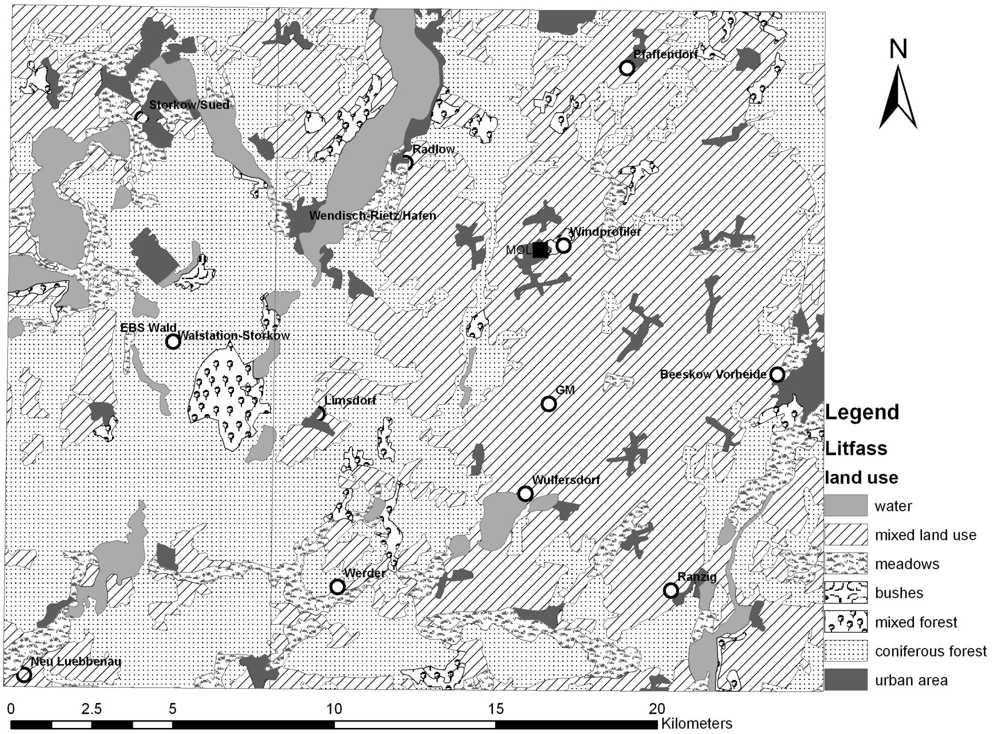


Figure 5.1: Lindenberg area. PLUVIO stations are indicated by black circles with white filling. MOL station is indicated by a black square. Land use data are based on CORINE Land Cover (2004).

Before deriving indices to determine the relevance of surface fluxes for precipitation (Section 5.3), the character of precipitation in this area has to be known. Detailed information on this can be found in Bohnenstengel et al. (2011a). Here, a summary of the main results is given. The distribution of precipitation is much skewed with frequent low precipitation amounts and few large precipitation amounts. Bohnenstengel et al. (2011a) determined the characteristic time scale of a precipitation episode to be in the range of up to 6 hours, since 94 % of all precipitation events in 2003 were within this range. These 94 % of all situations cover about 50 % of the annual precipitation. The rare and randomly occurring longer lasting precipitation events (6 % of all events) contribute to the other half of the annual precipitation. The amount of hourly, daily and even weekly precipitation amounts was found to be highly spatially varying even on small scales of a few kilometres. Thus it cannot be expected to find a simple relation of e.g. surface fluxes and precipitation amounts. The occurrence of precipitation is less spatially varying than the precipitation amounts. Therefore, a relation to surface fluxes might be possible. Emphasis is laid on the impact of the surface fluxes for the generation of precipitation by investigating the occurrence of precipitation and not the amount.

5.3 Locality indices

Precipitation is not an instantaneously developing atmospheric phenomenon but depends on several atmospheric processes with different time scales. The origin of precipitation may be several hours before the actual precipitation event; precipitation is a phenomenon with a history. Therefore, we have to consider the meteorological situation several hours before precipitation occurs and link this situation to the following precipitation events. Four approaches to characterise the meteorological situation were tried and corresponding non-dimensional locality indices defined.

Since clouds cannot develop without a sufficient amount of moisture in the atmosphere, evaporation from the surface, the source of water in the atmosphere, is of key relevance. The faster moisture is transported from the surface and reaches the condensation level, the higher is the probability that it influences precipitation within the next few hours. This is the case for more convective situations but also for advective situations with large wind speeds, in which the dynamically induced turbulence also impacts precipitation by transporting moisture from the surface into the upper boundary layer. The faster this vertical transport takes place, the larger is the probability for the local evaporation to influence precipitation in the limited model domain.

Two indices (I_{adv}, I_{lt}) are derived by comparing the characteristic time scales of relevant transport processes, where surface fluxes as a measure for evaporation are considered in all suggested indices. In a further step the transported moisture is included resulting in indices with absolute moisture (I_q) and with the impact of relative humidity (I_{rh}). Cloud microphysical processes are neglected, since their inclusion would need the use of a complex model, which contradicts the approach for simplicity of the index and of its calculation.

5.3.1 Advection and diffusion impact: locality index I_{adv}

For defining the locality index I_{adv} (eq. 5.1) it is assumed that the path of a parcel through the lower atmosphere originating at the surface mainly depends on two dynamical processes: horizontal advection and vertical diffusion. The relation of the characteristic times for advection, T_{adv}, and diffusion, T_{diff}, can be used to characterise the impact of horizontal advection by wind $\langle U \rangle$ and vertical diffusion on the path the parcel takes on its way towards the condensation level.

$$I_{adv} = \frac{T_{adv}}{T_{diff}} \tag{5.1}$$

In a convective situation vertical diffusion in the atmospheric boundary layer is very much influenced by the surface fluxes (e.g. Lüpkes and Schlünzen, 1996). These can well be taken as a measure for the vertical mixing. For convective situations the characteristic scaling velocity is the free convection velocity w_*, scaling height is the inversion height z_I (Businger, 1973). Within a stable, neutral and slightly unstable atmospheric stratification the vertical fluxes at the surface are characterised by the friction velocity u_*. The intensity of the vertical mixing in the atmospheric boundary layer in these cases also depends on the stratification above. However, to keep the calculation of the index simple we leave out this dependency and only consider u_* at the surface. Thus, the surface fluxes are used for all stratifications. They are characterised by s_*, the maximum of the scaling parameters u_* and w_*. The characteristic times in eq. (5.1) are calculated by:

$$T_{adv} = \frac{L_x}{< U >} \tag{5.2}$$

$$T_{diff} = \frac{L_z}{s_*}, \text{with}\, s_* = \max(u_*, w_*) \tag{5.3}$$

$$w_* = \left(-\frac{g}{\theta_v} u_* \theta_{v*} z_i \right)^{1/3} \tag{5.4}$$

L_x denotes a characteristic length scale in horizontal direction and L_z the length scale in vertical direction. <U> is the column averaged wind speed of the lowest 300 m of the planetary boundary layer and is derived from the sounding data taken at MOL. z_I is the simulated mixing layer height derived from the inversion of the potential temperature gradient, θ_{v*} is the scaling value for the virtual potential temperature θ_v and g the acceleration due to gravity. For L_x 150 km are arbitrarily chosen according to the size of a very small mesoscale model domain and L_z is set to 300 m, since a more accurate calculation of the condensation level as a vertical length scale would require the simulation of the whole thermodynamics, which jars against the intended simplicity of the index.

5.3.2 Diffusion impact: locality index I_{lt}

For defining I_{lt} it is assumed that the turbulent diffusion at the surface only contributes to precipitation, when its characteristic time scale is shorter than the characteristic time scale of a precipitation period T_{lt}. T_{lt} is set to 6 hours according to the findings of Bohnenstengel et al. (2011a), that 94 % of all precipitation events last up to 6 hours. The use of six hours is also of advantage, since it allows using routine observational precipitation data that are only available every six hours. Thus, I_{lt} is defined as relation of the characteristic time scales of a precipitation episode, T_{lt}, and the characteristic time scale of turbulent diffusion T_{diff}.

$$I_{lt} = \frac{T_{lt}}{T_{diff}} \tag{5.5}$$

T_{diff} is calculated as given in eq. (5.3) using again L_z = 300 m. When T_{lt} is set to 6 hours, I_{lt} results into $I_{lt} = 72\ s_*$ [s m^{-1}]. I_{lt} is only changed in value by s_* and is thereby completely driven in amplitude by the surface fluxes.

5.3.3 Diffusion and humidity impact: locality index I_q

The previously introduced non-dimensional indices I_{adv} and I_{lt} only account for the transport mechanism of moisture but do not consider moisture itself. They might be insufficient for the very dry meteorological situation of 2003. Therefore, I_{lt} is combined with the absolute humidity q derived from surface observations at MOL:

$$I_q = I_{lt} * q \tag{5.6}$$

With this index meteorological situations with intense turbulent exchange are weighted by the amount of available moisture.

5.3.4 Diffusion and relative humidity impact: locality index I_{rh}

Like Index I_q the index I_{rh} includes moisture. I_{rh} uses relative humidity taken from the MOL surface observations as weighting factor and is thereby a better measure for the saturation of the atmosphere. It should be a good indicator for the probability of precipitation.

$$I_{rh} = I_{lt} * RH \tag{5.7}$$

5.3.5 Calculation of the indices

Surface fluxes and atmospheric variables that are needed to calculate the indices are no measured quantities but derived from results of a 1D numerical model. The model is driven by routine observational data taken at MOL.

The 1D version of the 3D atmospheric mesoscale transport and fluid model METRAS (Schlünzen, 1990) is applied to the Lindenberg area. The 1D METRAS solves the equations for the horizontal wind, potential temperature and humidity in flux form. Considered in these calculations are Coriolis force and vertical mixing as well as surface exchange processes. For calculating the vertical exchange coefficient a mixing length approach is applied for stable, neutral and slightly unstable stratification and a counter gradient scheme is used for convective situations (Lüpkes and Schlünzen, 1996). These counter gradient fluxes are mainly driven by the surface heat fluxes and depend on boundary layer height. To account for the average impact of local surface characteristics on the flow, the heterogeneous landscape is considered in the 1D model. The surface characteristics are described by considering the fractional cover f_i of nine land-use classes in the Lindenberg area. A flux aggregation method with blending height approach (von Salzen et al., 1996) is applied to calculate area averages of the scaling values friction velocity u_* (eq. 5.8), free convection velocity w_* (eq. 5.10) and the scaling value for the virtual potential temperature θ_{v*} (eq. 5.11).

$$u_* = \sqrt{\sum_{i=1}^{9} f_i \left(u_{*i}\right)^2} \tag{5.8}$$

$$u_{*} = \kappa U(1) \frac{1}{\ln\left(\frac{z_1}{z_{0i}}\right) - \psi_m\left(\frac{z_1}{L_i}\right)} \tag{5.9}$$

$$w_* = \left(-\frac{g}{\theta_v}\sum_{i=1}^{9}\left(f_i u_{*i} \theta_{v*i} H_i\right)\right)^{1/3} \tag{5.10}$$

$$\theta_{v*} = \frac{1}{u_*}\sum_{i=1}^{9} f_i u_{*i} \theta_{v*i} \tag{5.11}$$

$$\theta_{v*i} = \frac{C_h^j}{\sqrt{C_m^j}}\left(\overline{\theta}_v(z_1) - \overline{\theta}_{v_i}(z_{0i})\right) \tag{5.12}$$

$$\theta_{vi} = T_{vi}\left(\frac{p_0}{p}\right)^{\frac{R}{c_p}} \tag{5.13}$$

$$T_v = \left[1 + \left(\frac{R_1^1}{R} - 1\right) q_1^1\right] T \tag{5.14}$$

C^j_m and C^j_h are transfer coefficients for momentum and heat, L is the Monin-Obukhov length and z_1=10 m designates the height of the first vertical model level, i describes the land-use-class and f_i its fractional cover. z_{0i} is the roughness length of each land-use class i and U(1) is the velocity at height z_1. The stability function for momentum Ψ_m is calculated according to Dyer (1974). T_v is the virtual temperature, p is the pressure, p_0 is set to 1000 hPa, R is the gas constant for dry air, R_1^1 the gas constant for water vapour, q_1^1 the specific humidity and T is the real temperature.

This method to calculate surface fluxes has been shown to deliver quite resolution independent results in 3D model simulations performed for the region around Lindenberg (Schlünzen and Katzfey, 2003). The parameterisation ensures that the 1D profiles not only represent one specific land-use characteristic but also include the heterogeneity of the land-use found in the Lindenberg area.

For model initialisation profiles of upper air wind, potential temperature and specific humidity are taken from the soundings at MOL. Both precipitation data sets PLUVIO_av6 and MOL_6 provide precipitation data with a time interval of 6 hours at 00, 06, 12 and 18 UTC. In consistence with available 6 hourly precipitation data (Bohnenstengel et al., 2011a), 1D METRAS is integrated for the Lindenberg area every 6 hours for the two years 2002 and 2003. The 1D model is integrated until the wind profiles are balanced. The temperature and humidity profiles are kept constant during this integration to ensure conservation of the thermal stratification. After balanced wind profiles are achieved temperature and humidity are calculated for 10 more time steps (about 10 minutes). This ensures that effects of local physical processes like convection are present in the model results in accordance with the measurements and also ensures that observed stratifications are kept during the integration. The resulting values for $<U>$, w_*, u_* and θ_* are used to calculate the four locality indices.

5.4 Locality indices and precipitation

The values of s_* are used for all index calculations and are the same for each index. The I_{lt} values and I_{adv} values are in the same order of magnitude, but I_{lt} values are more variable than I_{adv} values (Figure 5.2a). Values for I_{adv} and I_{lt} are between 10 and 100, for I_{rh} between 0 and 90 and for I_q between 0 and 12 (Figure 5.2b). No diurnal, seasonal or annual cycles are visible (Figure 5.2). Both years show a similar spread of the index values (not shown), despite the difference in total precipitation between the two years.

In Figure 5.3 the fraction of meteorological situations with precipitation in the six hours following each calculated index value are given for different values of I_{adv} (Figure 5.3a), I_{lt} (Figure 5.3b), I_q (Figure 5.3c) and I_{rh} (Figure 5.3d). The figure displays the fraction of all situations within the specified index interval (e.g. for I_{lt} = 30-39) that are connected with precipitation in the following 6 hours for years 2002 and 2003 using routine observational data at MOL and, in addition, integrated 6 hourly area averaged PLUVIO_av6 data for 2003.

For I_{adv} no dependence of precipitation occurrence and meteorological situation can be established. While there seem to be fewer meteorological situations with precipitation for decreasing I_{adv} when using the PLUVIO data (grey bars), the relation is somewhat reversed when using routine observations from MOL (dashed bars). Since the index relation needs to be independent of the precipitation measurement device, the index I_{adv} is not of help to classify meteorological situations with respect to precipitation probability.

For I_{lt} the fraction of meteorological situations with precipitation increases with increasing I_{lt} for both years 2002 and 2003 (Figure 5.3b). This relation is independent of the precipitation data set used. It is relevant that the relation is similar for both years, even though the total precipitation amounts differ considerably. 2002 was very wet (annual precipitation sum at Lindenberg 718 mm) and 2003 very dry (annual precipitation sum at Lindenberg 382 mm) compared to the long-term mean (563 mm). The percentage of "precipitating situations" per class is higher for the spatially averaged PLUVIO data, which underlines the local character of precipitation. It can be stated that meteorological situations with high I_{lt} values, representing situations with intense vertical mixing and a noticeable impact of surface conditions, have a higher probability for precipitation in the following six hours than situations with low I_{lt} values.

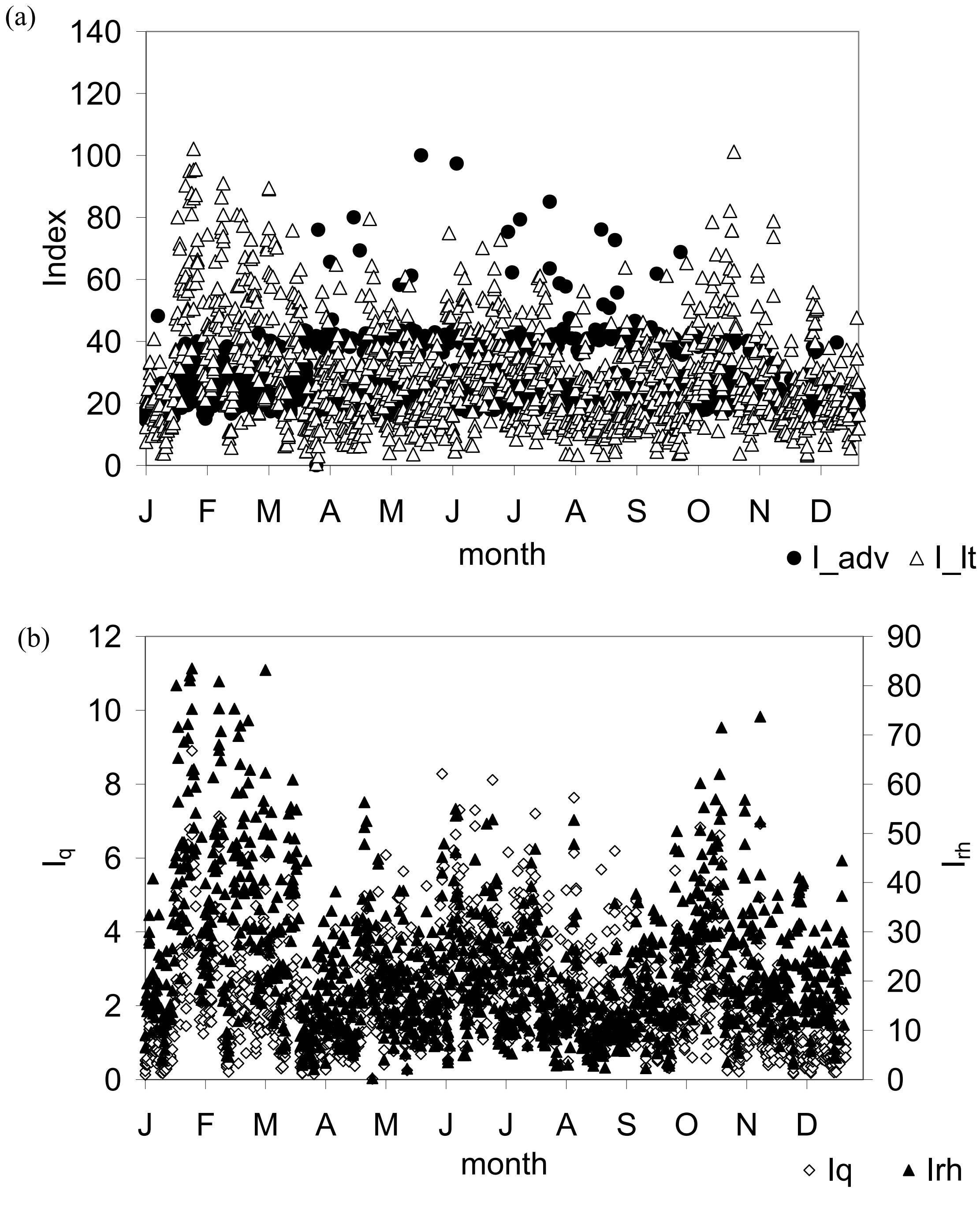


Figure 5.2: (a) Locality indices I_{adv} (dots) and I_{lt} (triangles) and (b) locality indices I_q (rhombi) and I_{rh} (triangles) for year 2002.

When including the absolute moisture q of the surface layer into the index (I_q), the increase of precipitation probability with increasing index value is less well defined as for I_{lt} (Figure 5.3c). The tendency is still there but precipitation probability does hardly increase for values larger than I_q=4.5. When using index I_{rh} and considering the relative humidity instead of the amount of moisture, a stronger dependency of precipitation occurrence on the index values (I_{rh}) is achieved (Figure 5.3d). I_{rh} shows higher precipitation probabilities than I_{lt} with increasing index values especially for MOL data. The weighting of high I_{lt} situations with low relative humidity values leads to lower values of I_{rh} compared to I_{lt}, and high relative humidity values have the opposite effect. This effect improves the relation of the index to precipitation probability. Thus, the index I_{rh} describes an increased probability for precipitation with increasing index values. As can also be seen, with I_{rh} values larger than 50, more than 50 % of the meteorological situations will produce precipitation in the following 6 hours.

To gain deeper insight into possible threshold values for the indices with respect to precipitation probability, the accumulated numbers of meteorological situations above a specific threshold value are calculated. The percentage of meteorological situations above the values of I_{adv}, I_{lt}, I_q and I_{rh} are displayed in Figure 5.4 and Figure 5.5. For all indices the 2 % most seldom situations with largest indices are not taken into account. Meteorological situations with high I_{adv} (Figure 5.4a) and high I_{lt} values (Figure 5.4b) represent situations with very large impact of the surface fluxes; these situations are rare in both years (lines with squares and rhombi).

Precipitation probability shows no relation to I_{adv} (Figure 5.4a) as already found from Figure 5.3. The precipitation probability is similar for all I_{adv} threshold values (solid and dashed lines). Therefore, I_{adv} is not a good measure to classify meteorological situations with respect to precipitation probability.

About half of all simulated situations have I_{lt} values above a threshold $I_{lt} \geq 30$ (Figure 5.4b). This amount is rapidly decreasing for higher threshold values: Only 30 % of all situations have values $I_{lt} \geq 40$ and only about 10 % of the situations have values $I_{lt} \geq 50$. These high I_{lt} values are rare, but connected with a higher probability for precipitation (solid and dashed lines) than lower I_{lt} threshold values. The higher the I_{lt} threshold value, the more probable it will rain in the following six hours.

Accumulated frequency values for index I_q (Figure 5.5a) show that about 50 % of all meteorological situations can be related to I_q values larger than 2 (dotted lines). Only about 10 % of all situations have values above 4. Higher I_q values are rare but connected with higher probabilities for precipitation than low I_q values. Maximum probabilities are 60 % for $I_q \geq 5$, the same probability is found $I_{lt} \geq 50$ (Figure 5.4b). However, probabilities for I_q values related to MOL are lower than for the higher I_{lt} and I_q values. Thus, I_{lt} is a better measure for precipitation probability than I_q. This is even more evident for the class wise distribution (Figure 5.3), where I_{lt} data show a much clearer trend for high values and for both precipitation data sets.

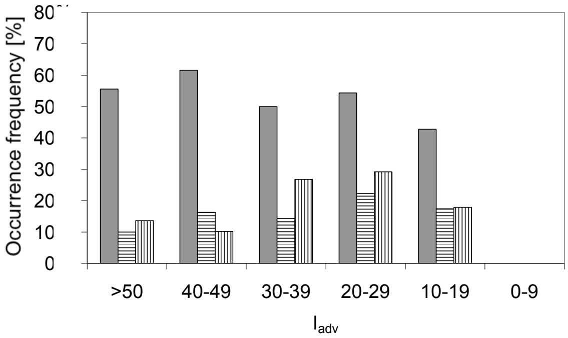
Index I_{rh} (Figure 5.5b) shows the same general behaviour as I_{lt} (Figure 5.4b). However, the probability for precipitation is somewhat larger for the higher index values of I_{rh}. 50 % of all meteorological situations have I_{rh} values of 18 and above. 20 % of the situations are captured by an index $I_{rh} \geq 30$. Differences to I_{lt} are feasible for the rare high index situations, where I_{rh} captures some five percentages more of the precipitating situations. Thus, I_{rh} seems to be the most appropriate measure for precipitation probability.

All indices displayed in Figure 5.4 and Figure 5.5 underline the findings of Bohnenstengel et al. (2011a) that precipitation amounts measured by a single station are not representative for a larger area. The probability for rain is about 10 % higher in case of using the area averaged PLUVIO data (solid line), compared to the routine observations at MOL. Between the two years 2002 (very wet) and 2003 (very dry) the probability for precipitation above specific threshold values is remarkably similar when comparing the two routine observation data sets MOL. Differences only occur for very high thresholds that are connected with very rare situations.

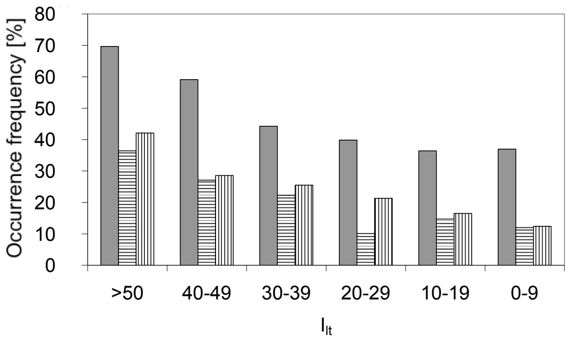
A relation of index values to precipitation amounts cannot be established. For index I_q, which takes the amount of moisture into account a small dependence between amount and I_q is found (Figure 5.6). This relation, however, differs for the more frequent situations of indices between I_q=1 and I_q=4. All other indices also show little dependence (Figures not shown). This is most probably a result of the locality of precipitation and it can be expected since the short-term precipitation amounts are very local. As Bohnenstengel at al. (2011a) have shown precipitation amounts have to be integrated for much longer time scales to be representative for even such a small area as the Lindenberg column. For instance, daily integrated precipitation amounts from a single station are only representative for the Lindenberg column within a factor of 3.3 (Bohnenstengel et al. (2011a). Precipitation amounts are very local and not representative for a larger area. Therefore, the non-existing relation of indices and precipitation amounts is not surprising.

The results suggest that precipitation probability is indeed related to the meteorological situation prevailing the precipitation event as characterised by I_{lt} and I_{rh}, and that this relation is slightly influenced by the actual relative humidity of the atmosphere. The relationship between I_{rh} (I_{lt}) and the probability for precipitation is very robust, since it is found for two quite different years. I_{rh} and I_{lt} are both adequate indices to classify meteorological situations and are related to precipitation probability. Index I_{adv} shows no relation, while index I_q shows a less robust relation than I_{rh} and I_{lt}. The absolute humidity is of course very important for the precipitation amount but is of no use in our context looking at short-term events. The inclusion of relative humidity of the lower atmosphere into the index I_{rh} shows some improvements in the relationship of the index to the probability of precipitation when compared to index I_{lt}.

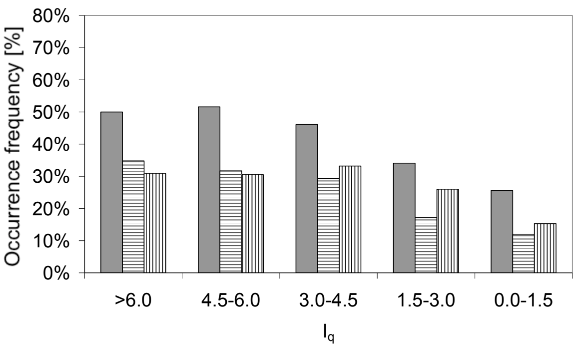
(a)



(b)



(c)



(d)

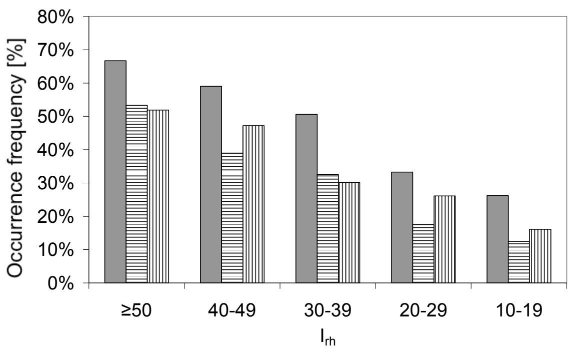
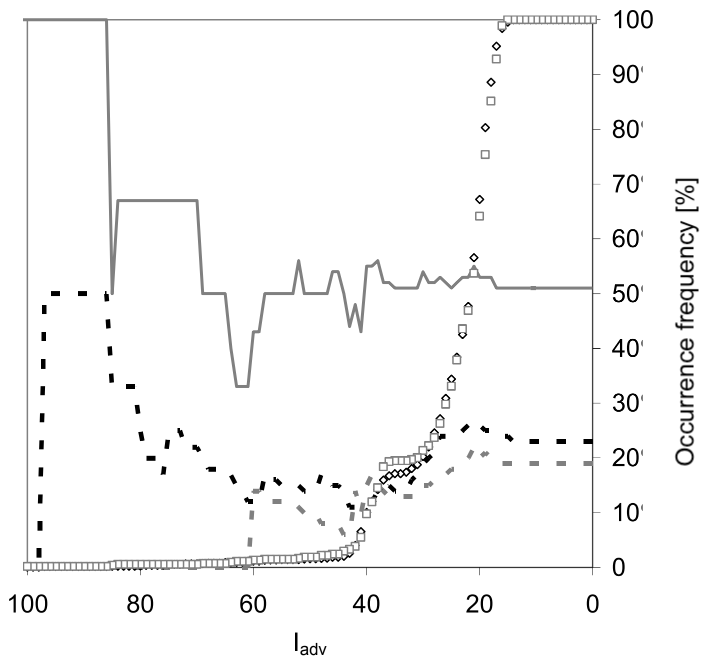


Figure 5.3: Fraction of meteorological situations with precipitation per index class (a) I_{adv}, (b) I_{lt}, (c) I_q and (d) $I_{rh.}$ Results are for PLUVIO measuring network in 2003 (grey bars) and single routine observations at MOL in 2003 (horizontally dashed bars) and 2002 (vertically dashed bars).

(a)



(b)

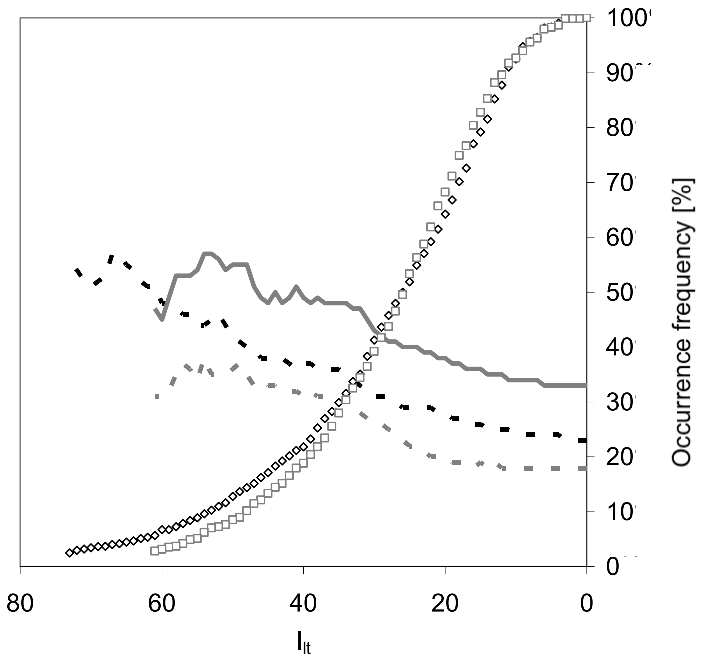
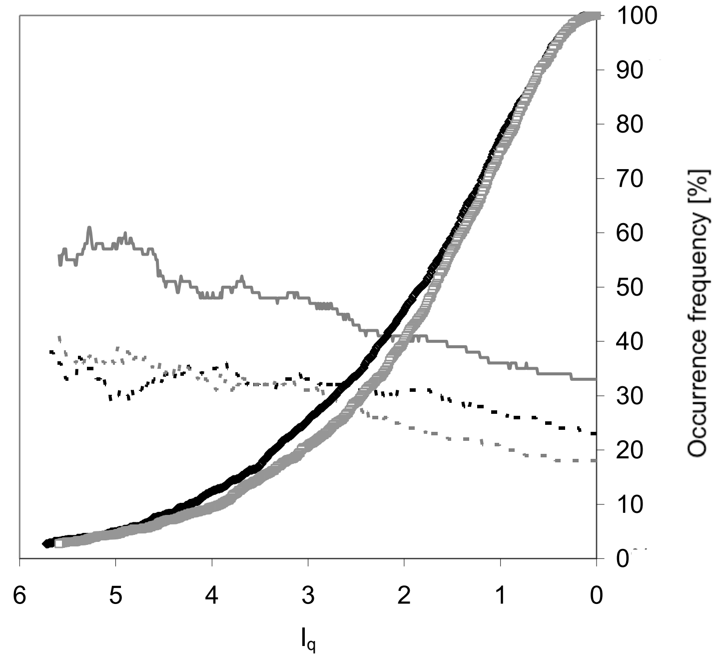


Figure 5.4: Accumulated relative frequency of meteorological situations above index threshold values (rhombi and squares) for I_{adv} (a) and I_{lt} (b) and the fractions of meteorological situations with precipitation compared to the total number of all situations above the index threshold value (solid and dashed lines). The grey solid line corresponds to PLUVIO data; the dashed lines to routine observations at MOL. Black lines correspond to year 2002 and grey lines to year 2003.

(a)



(b)

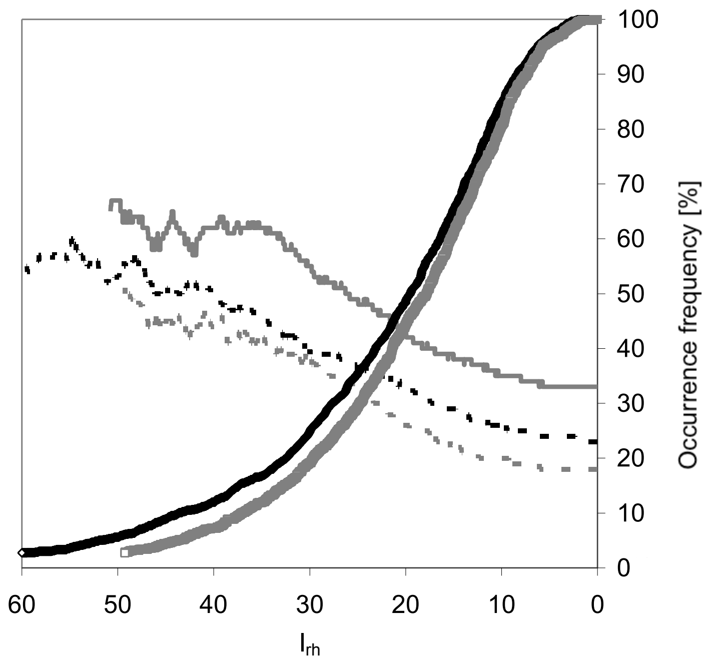


Figure 5.5: Like Figure 5.4 but for I_q (a) and I_{rh} (b). Accumulated relative frequency of meteorological situations above index threshold values (rhombi and squares appear as thick). Grey thin solid line corresponds to PLUVIO, the dashed lines to MOL. Black lines correspond to 2002 and grey lines to 2003.

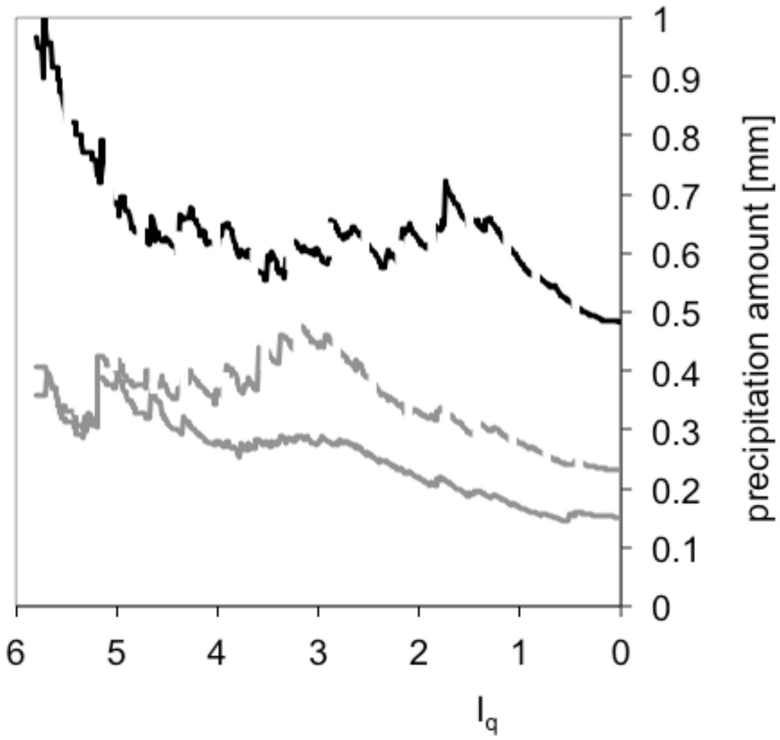


Figure 5.6: Precipitation amount for I_q values. The black dashed line corresponds to MOL data for 2002, the grey dashed line to MOL data for 2003 and the fractional grey lines to PLUVIO data for 2003.

5.5 Summary and Conclusions

We introduced four easy to calculate indices for classifying meteorological situations with the aim to find indices that are related to precipitation. All indices are based on characteristic time scales of atmospheric processes and I_q and I_{rh} additionally include humidity measures (specific and relative humidity). To calculate the index values, the necessary flux parameters were simulated with the 1D version of the numerical mesoscale transport and fluid model METRAS for the Lindenberg area every six hours. The model was initialised with routine observations and soundings. The index values were related to precipitation in the six hours following the initialisation. Three data sets were used: the six-hour precipitation amounts measured at the single station MOL in years 2002 and 2003 and the Lindenberg area average 6-hourly precipitation data calculated from the 14 instruments of the PLUVIO measuring network (Beyrich, 2004). This is spatially distributed covering an area of 25 km x 25 km around MOL.

None of the four indices introduced shows a clear relationship to precipitation amounts in the following six hours. This is a result of the locality of the intensity of precipitation events mentioned in Bohnenstengel et al. (2011a). Index I_{adv} also shows no relation to precipitation probability. In contrast, I_{lt} can be used as a measure for the probability of precipitation within the following six hours; the higher the I_{lt} value and thus the impact of surface fluxes, the larger is the probability for rain. Very high index situations (I_{lt} > 40) are rare, but are connected with a high probability for precipitation (probability > 30 % at a single site, probability > 49 % for a 25 km x 25 km area).

Moisture was considered in indices I_q and I_{rh}. While I_q shows no improvement for precipitation probability, I_{rh} shows an improved relation especially for the precipitation probability at mid-level and high index values. The advantage of I_{rh} compared to I_{lt} is mostly small, but for high index values the results are more independent of the investigated year. Thus, both indices can be used; I_{lt} is somewhat simpler and easier to calculate than I_{rh}. But I_{rh} might be of advantage especially in dry regions (here: year 2003). Both indices can be calculated from single station data if some information on vertical stability is available. We used sounding data and a 1D numerical model for this purpose.

The indices I_{lt} and I_{rh} have been developed to characterise meteorological situations by mainly comparing characteristic time scales of the processes involved. They filter meteorological situations with regard to the impact of local surface conditions and the intensity of vertical turbulent diffusion. When investigating a long time series it is easy to use the indices for developing frequency distributions of meteorological situations with respect to precipitation probability. The higher the I_{lt} or I_{rh} values and thus the local impact of surface conditions, the higher is the probability for precipitation at the site within the next 6 hours. The situations with high I_{lt} or I_{rh} values are either connected with intense convection or with strong surface wind conditions resulting in a larger friction velocity, which does also enhance the vertical turbulent exchange and thus affects the generation of precipitation by surface fluxes. Using I_{rh} increases the calculated probability for precipitation compared to using I_{lt}.

It was intended to develop a very simple index for classifying meteorological situations with regard to precipitation. The assumptions used for simplification like horizontal homogeneity are not valid for strong heterogeneities in land-use like land/water or horizontal warm/cold contrasts. For these cases the indices probably do not work. Also the representativity of the index is not given for a longer time period. It is evident, that in these cases further physical processes not captured by the index contribute to precipitation. Therefore, only using the index for estimating precipitation probability will lead to false alarms. However, for the investigated up to six hours period the relation between I_{lt} (I_{rh}) and precipitation probability can clearly be seen. It can be applied as a simple to derive tool for determining precipitation probability and, furthermore, can also be used to evaluate model results for situations with different precipitation probability or even be used to decide, in which meteorological situations the use of a more complex method for calculating the surface fluxes is necessary.

6 Model performance for different sub-grid scale surface flux parameterisation schemes

Parts of this Chapter have been published as:

Bohnenstengel S., Schlünzen K.H. (2009): Performance of different sub-grid scale surface flux parameterizations for urban and rural areas. In: Baklanov A., S. Grimmond, A. Mahura, M. Athanassiadou (Eds), Urbanisation of meteorological and air quality models, Springer Publishers, 165p., ISBN 978-3-642-00297-7

The importance of the surface fluxes for the boundary layer structure is different for different meteorological situations. In order to improve model forecasts the applicability range of the parameterisation schemes for surface fluxes within the kilometre scale needs to be determined with regard to the meteorological situation and horizontal grid resolution. Consequently, the locality indices I_{rh} and I_{lt} were developed as a tool to characterise meteorological situations into locally driven and more advectively driven situations (Chapter 5). In the following, a process-oriented evaluation is undertaken for a number of model simulations using overall values for I_{rh} and I_{lt}, which indicate the importance of the surface fluxes for the simulated day. Studies are carried out for 6 different meteorological situations with respect to the horizontal resolution and the parameterisation scheme for sub-grid scale surface fluxes.

6.1 Simulated meteorological situations

Six different meteorological situations were chosen from year 2003. They differ in the six hourly locality index values calculated for Lindenberg, which is situated in the middle of the selected model domain as described in Section 5.2. The situations were selected to ensure that situations with different impact of the local surface characteristics on the boundary layer structure are simulated. The corresponding locality index values are summarized in Table 6.1. Two of the periods were nudged into the ECMWF analysis data (Ries et al., 2010), while the four other periods were driven by forcing data derived from an own resolution-dependent analysis (Gao, 2001), which is based on soundings and water temperatures (British Atmospheric Data Centre) and observational data from the DWD. In the following the abbreviations according to Table 6.1 will be used for the different simulation periods. The following description of the meteorological situations is based on information from the "Berliner Wetterkarte e.V.".

Abbreviation	Initialisation day	Integration time [days]	Grid sizes, [km]	Forcing	I_{rh}	I_{lt}	Conditions
E1	09.08.2003	4	16, 8, 4, 2	ECMWF	12	40	very dry
E2	10.03.2003	4	16, 8, 4, 2	ECMWF	30	60	very humid
M1	04.03.2003	3	16, 8, 4	Analysis	9	10	humid
M2	04.01.2003	2	16, 8, 4	Analysis	18	20	humid
M3	10.06.2003	2	16, 8, 4	Analysis	18	30	very warm & dry
M4	03.06.2003	3	16, 8, 4	Analysis	16	40	very hot & very dry

Table 6.1: Summary of simulation periods.

6.1.1 Period E1: very dry August simulation

This period was chosen based on the six hourly calculations of the locality index I_{lt} (Chapter 5), which indicated that this situation is a very local but dry situation with 30 - 60 % precipitation probability. I_{rh} resulted in a value of only 12, which suggests a 15 – 30% precipitation probability (Figure 5.3d). The I_{lt} of 40 indicates a very locally driven meteorological situation. This simulation is started for 18 UTC on the 9th August 2003 and integrated for 4 days. The dry and warm period from 9th August 2003 until 13th August 2003 was nearly cloudless and dominated by the anticyclone "Michaela" which was located over central Europe. The year 2003 was a very dry and hot year in Europe and the selected period is characterised by very high temperatures of about 30 °C and dew points between 2 °C and 16 °C. Low wind speeds from northerly directions allow local processes to influence the model solution. No precipitation was measured in the model domain. This period was chosen, because it was a very dry period and a good opportunity to investigate the performance of the surface flux schemes for an extreme situation, where the standard METRAS moisture approach assumes soil conditions that are likely overestimating the soil moisture for this case. This simulation was nudged into ECMWF analysis data (Ries et al., 2010).

6.1.2 Period E2: very humid March simulation

This period is characterised by large locality indices I_{lt} and I_{rh} indicating a higher precipitation probability (30 – 50 %) and a strong influence of the surface fluxes. Simulation E2 is initialised for 18 UTC on the 10th March 2003 and integrated for 4 days. During this period Southern Germany was dominated by the high pressure system "Kerstin" located in the Mediterranean region, while Northern Germany was dominated by the low "Gordian" situated west of Scandinavia with lowest pressure values of 985 hPa in the centre and pressure values between 1010 hPa and 1020 hPa over the model domain. During the 11th March 2003 a warm front over Northern Germany resulted in intense precipitation amounts of up to 8 mm in 12 hours in the model domain. Since wind speeds were considerably low during this day it is expected that the surface fluxes affect the precipitation in this area. The very wet period with a relative humidity over 90 % was characterised by temperatures between 0°C and 5°C and increasing wind speeds during the 12th March 2003 and 13th March 2003. The most interesting day for our evaluation purpose during this period is the 11th March 2003 where the precipitation occurred. This simulation was nudged into ECMWF analysis data.

6.1.3 Period M1: humid March simulation

The simulation was initialised for 18 UTC on the 4th March 2003 and integrated for 3 days. The focus lay on the first and second simulation days which were dominated by the high pressure system "Jutta" located over Belarus and leading to south and south-easterly wind directions with low wind speeds over the model domain. Accordingly, temperatures were relatively low between -3°C and 5°C with strongly varying relative humidity between 50 % and 90 %. On the 5th March 2003 a very weak low with pressure values of 1025 hPa was situated over the border between Northern Germany and Poland. Some precipitation was observed in a small band over Northern Germany due to a cold front occlusion over the Berlin region. According to the locality index analysis (Chapter 5) this situation was more advection dominated with a low precipitation probability. This simulation was nudged into results derived from an own analysis based on DWD routine observations and soundings from the BADC.

6.1.4 Period M2: humid January simulation

This simulation was started for 18 UTC on the 4th January 2003 and integrated for 2 days. Whole Northern Germany was dominated by cold Arctic air masses leading to temperatures below -10°C and relatively high pressure values of about 1010 hPa and 1020 hPa on the 5th January 2003. Later, the occlusion of the low "Dino" let to some snow in Northern Germany. According to the index I_{rh} the situation was relatively local and slightly humid, thus some impact of the surface on the near surface boundary layer structure can be expected. Similar to M2, this simulation was driven by an own analysis.

6.1.5 Period M3: very warm and dry June simulation

The simulation was started for 18 UTC on the 10th June 2003 and integrated for 2 days. Temperatures in Northern Germany were very warm during this period with values of up to 32 °C. The air was mainly relatively dry during the simulation period except for some small precipitation amounts of up to 0.8 mm in 3 hours observed at the station Berlin Dahlem (Berliner Wetterkarte e.V.) during the night from the 10th June 2003 to the 11th June 2003. Pressure was relatively high between 1015 hPa and 1020 hPa (Berliner Wetterkarte e.V.). The whole situation was very local with low wind speeds and relative humidity values between 50 % and 90 %, and according to the German Meteorological Service some very local thunderstorms developed.

6.1.6 Period M4: very hot and very dry June simulation

The simulation was started for 18 UTC on the 3rd June 2003 and integrated for 3 days. The whole period was characterised by the high pressure system "Zoe" over the Baltic leading to high pressure values over Northern Germany. Temperatures were very high with up to 34 °C in the Berlin area (Berliner Wetterkarte e.V.). According to the “Berliner Wetterkarte e.V.” the vertical temperature gradient in the lower boundary was super-adiabatic over Northern Germany during the 4^{th} June 2003, which should result in a strong impact of the surface fluxes on the whole boundary layer and is very interesting for the process oriented evaluation of the model performance. In some parts of Germany very intense precipitation events occurred which were very local, but according to the measurement systems of the DWD in Northern Germany nearly no precipitation was detected in the model area.

All situations were simulated with the flux aggregation scheme with blending height approach as well as with the parameter averaging scheme (Section 2.2) and at least three different horizontal resolutions (Table 6.1). In the following the different six configurations for the four “M-cases” and the eight configurations for the two “E-cases” of the model will be abbreviated with e.g. FL4 for the flux aggregation scheme with 4 km resolution or e.g. PA8 for the configuration using parameter averaging with 8 km resolution.

6.2 Model evaluation method

In the following the model performance for all six situations is evaluated by calculating hit rates. Hit rates were already successfully applied e.g. in Schlünzen and Katzfey (2003). The hit rates are calculated for temperature, dew point temperature, wind speed and wind direction based on the different configurations per simulation period and routine observations from the 27 DWD stations in the model domain to conclude on the overall model performance. Therefore, the model results are interpolated into the locations of the DWD stations at 2 m height above the surface for the thermodynamic values and at 10 m height for the dynamic values applying Monin-Obhukov similarity theory. The hit rates are then calculated following eq. (6.1). A in eq. 6.1 accounts for the accepted uncertainty of the data. Following values suggested by Cox et al. (1998) A is set to be ±2 K for the air temperature and dew point temperature in the present analysis. . For wind speed A is set ±2 ms^{-1}, which is double the value used by Schlünzen and Katzfey (2003). For wind direction ± 30° is used.

$$H = \frac{100}{m} \sum_{i=1}^{m} n_i, \quad \text{with } n_i = \begin{cases} 1 \text{ for } |\text{difference}(\text{measurement,model})| < A \\ 0 \text{ for } |\text{difference}(\text{measurement,model})| \geq A \end{cases} \qquad (6.1)$$

A hit rate of 100 % indicates an excellent model performance with 100 % of the simulated model results being within the desired accuracy range A, while a hit rate of 0 % indicates that none of the model results matched the observations within the accuracy range A. The hit rate does not give a hint on the bias of the model results. However, a big advantage is that hit rates do not take into account the distribution of the simulated values e.g. a Gaussian distribution. A further advantage is that hit rates allow to compare very different meteorological situations and model set-ups. They are applicable to evaluate the model performance while taking into account measurement uncertainties or uncertainties due to assumptions for the interpolation of the variables to the DWD stations. The uncertainty of calculated hit rates is set to ±5 %, since the METRAS data were interpolated to the DWD locations. Therefore, changes within hit rates of 5 % are indicated by black lines in the following figures and are considered to be negligible. The hit rates for the simulations are sorted by their corresponding locality indices to evaluate the performance of the configurations based on the impact of the surface fluxes, where high index values indicate a strong influence of the surface characteristics.

6.3 Overall evaluation results

The meteorological situation and the horizontal resolution are assumed to play a key role for the performance of the sub-grid scale surface flux parameterisation scheme. Figure 6.1 summarizes the overall hit rates averaging the hit rates of temperature, dew point temperature, wind speed and wind direction over all configurations per simulation period resulting in an overall hit rate for the whole situation (red bars). The overall performance does not vary much for the different simulation periods. When comparing the hit rates averaged over temperature and dew point temperature (blue bars) with the hit rates averaged over wind speed and wind direction (green bars) the simulations forced by the own analysis show a better performance for the dynamic values than for the thermodynamic values. The simulations forced by the ECMWF analysis data show the opposite behaviour with the best hit rates for the thermodynamic values for the E1 simulation.

In order to determine the sensitivity of METRAS results towards the parameterisation method, the horizontal resolution and the meteorological situation, the hit rates are compared to the theoretically worst case with parameter averaging and a very coarse resolution of 16 km (PA16) for each configuration individually. The hit rates are calculated separately per configuration for each simulation period for temperature, dew point temperature, wind speed and wind direction. Then the difference of the hit rates between each configuration of a simulation period and the corresponding PA16 result is calculated. Figure 6.2 and Figure 6.3 summarize the differences of the hit rates, whereby positive values indicate that the chosen configuration performs better than parameter averaging with 16 km resolution (PA16). Negative values describe a decrease in model performance when increasing the resolution or applying flux aggregation instead of parameter averaging.

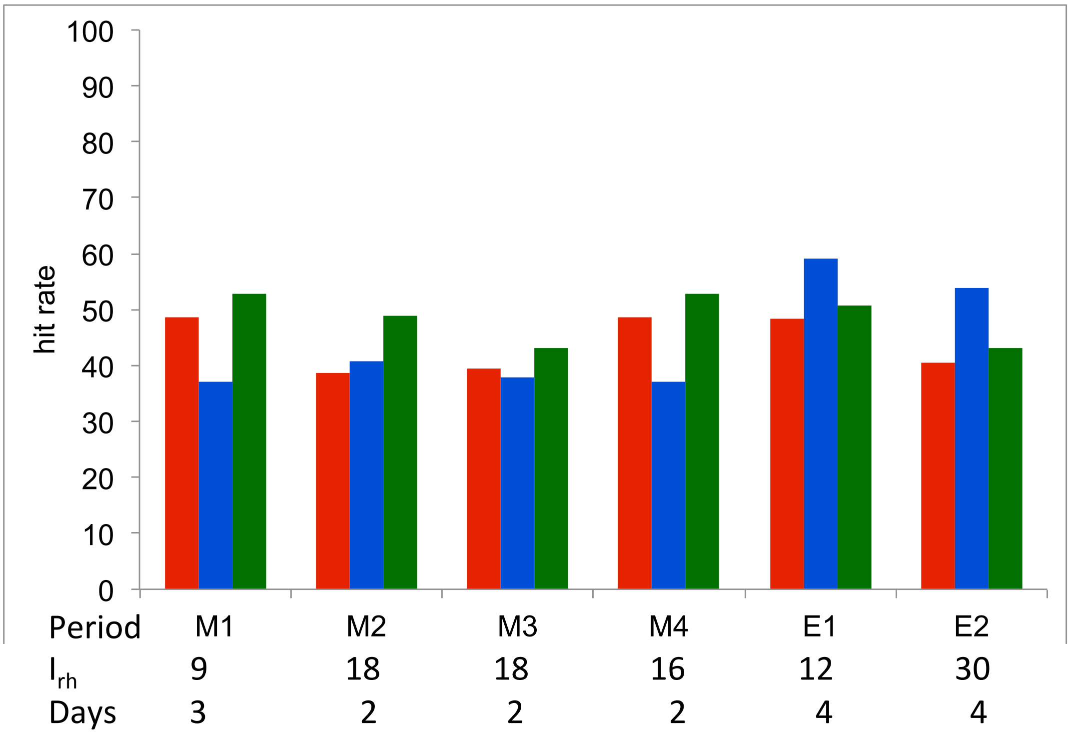


Figure 6.1: Hit rates are averaged over all variables and configurations for each simulation period (red bars), averaged only for thermodynamic values (blue bars) and only for the dynamic values (green bars) for the different simulations indicated by the locality index I_{rh}. I_{rh} is the locality index value representative for the length of each simulated period as indicated by days.

Figure 6.2a shows the differences of the hit rates for the case M1 with an overall low locality index of $I_{rh} = 9$ and $I_{lt} = 10$. The differences of the hit rates are below 20 % and mostly within the uncertainty range of ± 5 %. For wind speed ff, temperature te and dew point temperature td increasing the resolution or applying flux aggregation improves the model results compared to PA16. Differences in the range of +10 % to +15 % occur for flux aggregation with 8 km (ff, te, td) or 4 km (ff) resolution.

For situation M2 (Figure 6.2b) with a doubled locality index ($I_{rh} = 18$ and $I_{lt} = 20$) differences are still relatively small, but here increasing the resolution for parameter averaging does only slightly increase the model performance for 8 km resolution for wind speed and dew point, but not for wind direction. For temperature the improvement is within the uncertainty range. For 4 km resolution the performance for wind direction and temperature is worsened – and within the uncertainty range for wind speed and dew point. Hence, increasing the model resolution does not necessarily improve the model results. Applying parameter averaging instead of flux aggregation worsens model results for wind direction and temperature, but not for the dew point temperature and wind speed (8km resolution). Most changes are within the uncertainty range of ±5 % and therefore negligible.

For situation M3 (Figure 6.2c, I_{rh} = 18 and I_{lt} = 30) the sensitivity towards parameterisation and resolution is mostly within the uncertainty range of ±5 %. Differences within the parameter averaging scheme are negligible (except dew point temperature), but with a slight increase of the performance for higher resolutions. Flux aggregation performs better than parameter averaging except for temperature, where flux aggregation with 4 km resolution has a 10 % lower hit rate than PA16.

With increasing locality index I_{lt} = 40 (I_{rh} = 16), the sensitivity of the model results towards the method to include the sub-grid scale surface fluxes increases. For case M4 (Figure 6.3a) increasing the resolution towards 4 km improves the hit rates for parameter averaging except for wind direction. The dew point temperature benefits most from increasing the resolution in the case of parameter averaging. While the difference between PA16 and PA8 is negligible PA4 shows an enhancement of the model performance of 10 %. This improvement is even larger for dew point temperature when using flux aggregation. Increasing the resolution or using flux aggregation results in higher hit rates for all variables (except wind direction). For wind direction differences are mostly within 5 % except for PA4 and FL4, where the performance worsened slightly. The hit rates for dew point temperature with flux aggregation are less resolution dependent than for parameter averaging. They show a slightly larger improvement of the hit rate for all resolutions than parameter averaging with 4 km resolution does.

Situation E1, with the same locality index of I_{lt} = 40 as M4 but a lower I_{rh} = 12 shows large impacts of METRAS towards the parameterisation scheme and resolution. While the thermodynamic values overall improve for increasing the resolution or applying flux aggregation, the dynamic values are simulated worse with FL4 in case of wind direction and FL2 in case of wind speed. The simulation of temperature benefits most from applying flux aggregation, although the hit rates drops for the 2 km resolution case compared to the coarser resolutions. Differences in dew point temperature are much smaller and most improved for 2 km. Since this case was relatively dry, the impact of the latent heat flux on the dew point temperature is smaller than the impact of the sensible heat flux on the temperature.

The sensitivity of model performance on resolution and parameterisation is largest for the case E2 with the highest locality index (I_{lt} = 60 and I_{rh} = 30) indicating a meteorological situation that is very much driven by the surface fluxes. This case was very humid and simulated precipitation within the model domain. The sensible and latent heat flux were both very important for the simulation of the boundary layer structure, and the impact of the surface fluxes was very strong. In this case none of the schemes is clearly superior. The wind direction is the only variable, which consistently benefits from increasing the resolution and applying flux aggregation. The wind speed is less well simulated with increasing resolution. There are nearly no differences in the performance between FL2 and PA2, FL4 and PA4 as well as FL8 and PA8. Only for 16 km resolution parameter averaging shows considerably better results than flux aggregation (except DD). In general, PA16 shows the best results for wind speed. In case of the thermodynamic values, only FL4 and PA4 show an improvement of the performance for temperature as well as dew point temperature. All other configurations perform less well than PA16. With great caution it is suggested that in this case changes are not only resolution and parameterisation dependent, but a further unknown process might affect the solution.

The local impact of the surface fluxes on the boundary layer structure varied for all case studies. The conclusion can be drawn that the sensitivity of the model performance on resolution and parameterisation increases for higher locality indices I_{lt}. These are situations where the local impact of the surface fluxes plays a key role for the boundary layer structure. A good parameterisation for the surface fluxes is essential for a faithful representation of the meteorological processes. For locally driven situations the differences between parameter averaging and flux aggregation are most pronounced. Especially, parameter averaging tends to benefit from higher resolutions. The land-use is resolved more explicitly, and the averaged surface characteristics tend to describe real sub-grid scale impact significantly more realistically. The flux aggregation scheme seems to be less resolution-dependent, except for 2 km where the model sometimes performed significantly different from the coarser resolutions with flux aggregation. This parameterisation scheme and the prescribed surface characteristics were designed for coarser resolutions. For instance, vegetation was assumed to be present in an urban tile. Therefore, the prescribed land surface characteristics might have to be adjusted for higher resolutions. For instance, in case of the urban tile less vegetation can be assumed to be present, since the urban area is resolved more explicitly. When still using the surface parameters designed for coarser resolutions with higher resolutions, the Bowen ratio will be affected and lead to an overestimation of 2 m temperatures, while at the same time the dew point temperature is underestimated or vice versa. This behaviour is reflected in the performance for flux aggregation especially for the cases with a stronger local impact of the surface fluxes. For these simulations the sensitivity of the thermodynamic values is increased compared to the cases with lower impact of the surface fluxes.

The sensitivity of the thermodynamic values which are directly dependent on the surface fluxes like sensible and latent heat flux is larger than for the dynamic values. The dynamic values show more consistency in their behaviour. A main outcome of the presented analysis is that the variability for the thermodynamic values is larger than for the dynamic values. A second outcome is that flux aggregation tends to be the superior scheme to parameter averaging for the presented case studies. Thirdly, it was shown that in case of flux aggregation increasing the resolution does not necessarily improve the model results. Flux aggregation with a blending height concept is in general relatively resolution-independent. Fourthly, high index I_{lt} situations show a larger sensitivity towards the parameterisation and the resolution, since the surface characteristics have a bigger impact on the lower boundary layer structure. It makes a significant difference especially for the 2 m thermodynamic values whether a surface process is well parameterised or not. Consequently, I_{lt} and I_{rh} were applied to evaluate the model performance in a process oriented way by indicating the importance of the surface fluxes for the simulated temperature and wind variables. The results suggest that differences between the various configurations are relatively small for I_{lt} values below 40. Those situations are likely to have a considerable advective impact and are less locally driven. For low index situations the forecasted temperature and wind values are relatively independent of the parameterisation scheme, since the surface fluxes have less impact on the model solution. Such a distinction is not possible for I_{rh}. From these case studies it is concluded that I_{rh} is not an appropriate parameter for determining the best parameterisation of the surface fluxes in a model online. However, I_{lt} is applicable to determine the appropriate surface flux parameterisation scheme dependent on the process. Also, in case of dry situations it outperforms I_{rh}, since I_{rh} underestimates the locality of the meteorological situation once the relative humidity is very low. Still, this conclusion needs to be based on more simulations. However, in combination with I_{lt} I_{rh} is a good indicator on the likelihood of locally driven precipitation events.

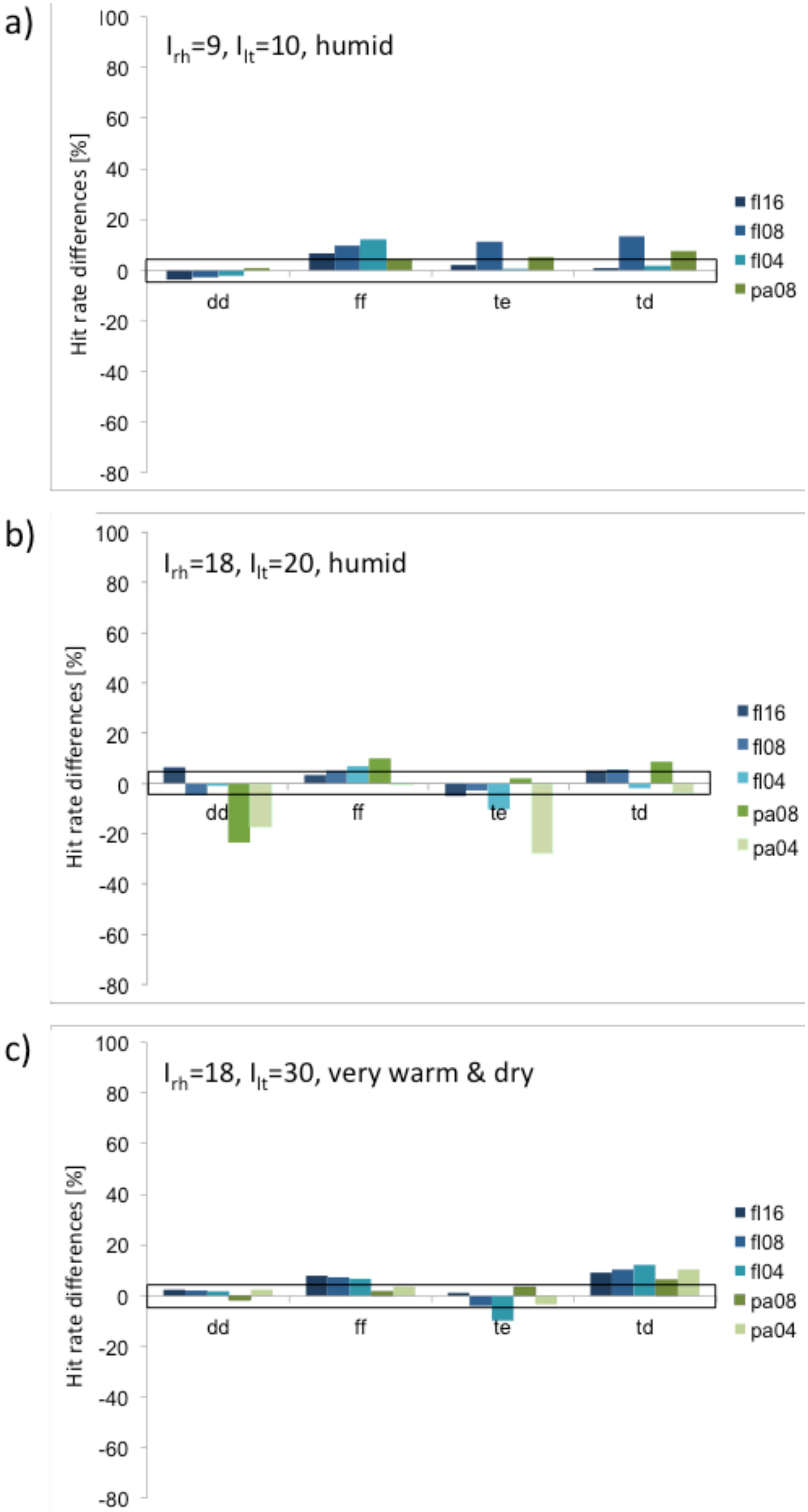


Figure 6.2: Differences in hit rates compared to parameter averaging with 16 km resolution for (a) M1, (b) M2, (c) M3. dd stands for wind direction, ff for wind speed, te for temperature and td for dew point temperature.

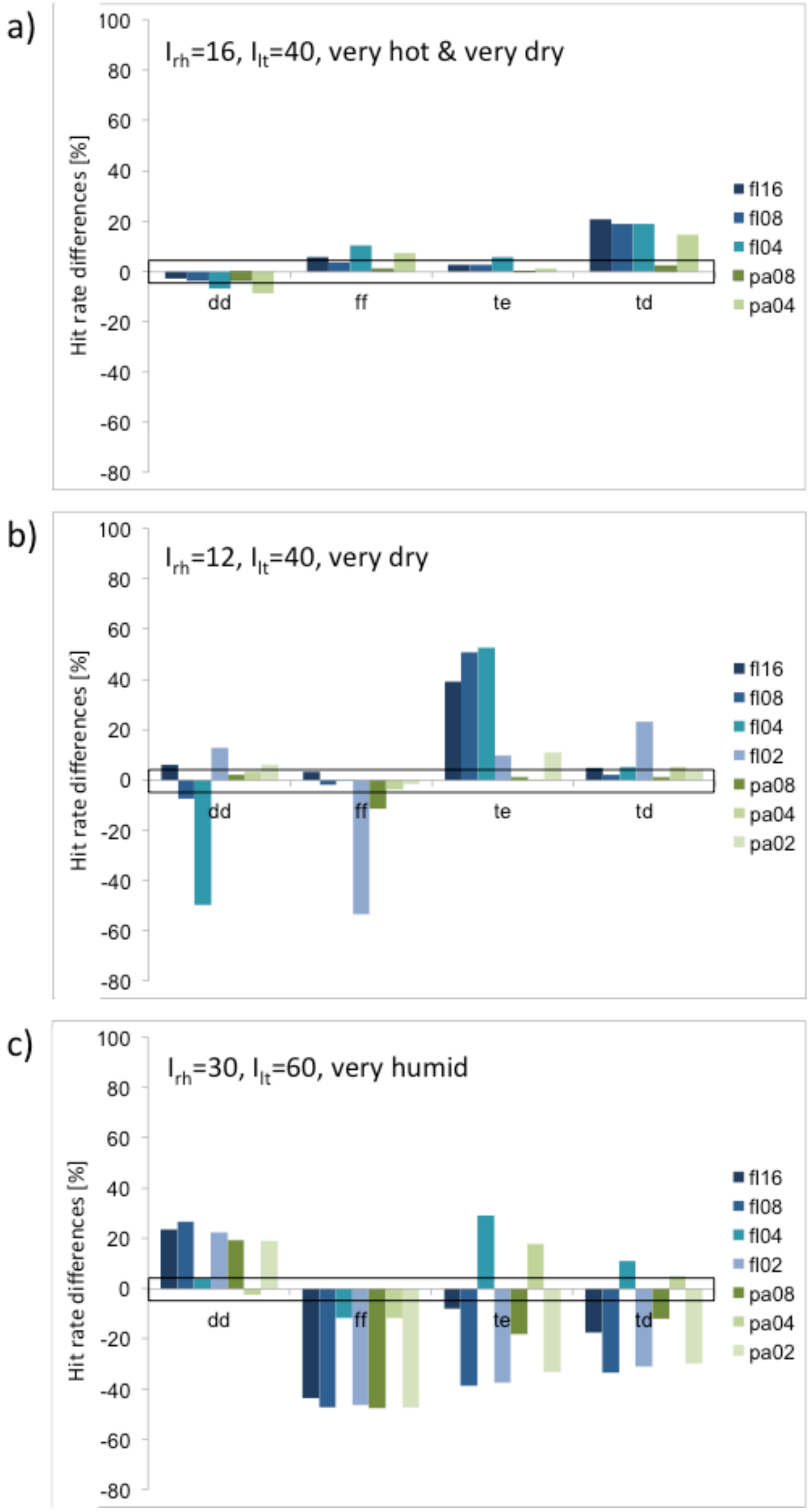


Figure 6.3: Same as Figure 6.2 but for (a) M4, (b) E1, (c) E2.

6.4 Evaluation of the diurnal cycle

For the previous investigations every single simulation was described by the averaged locality index based on the 6 hourly locality indices I_{lt} and I_{rh} derived in Chapter 5. The hit rates were determined for the individual variables of temperature, dew point, wind speed and wind direction. In the following the emphasis is laid on the performance of the thermodynamic values.

For a better understanding of METRAS' sensitivity towards the influence of sub-grid scale surface fluxes the simulation of the diurnal cycle is investigated in more detail. According to the locality index I_{lt} and I_{rh} case E1 and even more E2 are very locally driven meteorological situations. Therefore, the surface fluxes play a key role for the realistic simulation of the screen level temperatures in the surface layer and it is expected that the model solution is sensitive towards the parameterisation scheme for surface fluxes. In the following, case E1 was chosen for analysing the diurnal cycle of temperature for the eight different configurations. The diurnal amplitude of temperatures was large for case E1 and the surface characteristics strongly influenced the model solution. Unlike case E2 case E1 was not influenced by precipitation. Hence, case E1 has the advantage that the influence of the resolution-dependence of the microphysics scheme is small, since clouds affecting the surface energy balance are negligible in this case. Differences in the diurnal cycle of the temperature are expected to be caused by the surface parameterisation scheme. Since this case was very dry and the sensible heat flux much larger than the latent heat flux, the focus lies now on the simulation of the 2 m temperature only.

Increasing the horizontal resolution results in very heterogeneous model answers for the different configurations as shown in Figure 6.3b. When comparing the simulated diurnal cycles of spatially averaged temperatures for the eight configurations of case E1 with observations from the DWD (Figure 6.4) this wide range of answers is visible as well. The DWD measurements show a strong diurnal variation of the temperature with the highest temperatures occurring around 2 UTC and lowest temperatures at about 3 UTC in the early morning. All configurations resemble this timing relatively well. However, they all underestimate the maximum temperature. When comparing the performance of flux aggregation and parameter averaging independent of the resolution, flux aggregation produces significantly better results. The results are very close together. In contrast, parameter averaging tends to be very resolution-dependent. The maximum temperature ranges from 23 °C (PA16) up to 28 °C (PA2). Set ups with coarser resolution underestimate the day time temperatures by up to 8 °C. The diurnal amplitude is damped compared to the DWD measurements. From Figure 6.4 it is evident that parameter averaging tends to benefit most from increasing the horizontal resolution: the parameter averaging method tends to converge towards the solution of the flux aggregation scheme. FL2 and PA2 only differ by 1 °C at most during the afternoon while PA16 simulates temperatures about 6 °C lower than FL16. The increase of the horizontal resolution has the positive side effect that the averaged surface characteristics describe the heterogeneity of the surface more realistically, since they are resolved more explicitly. Parameter averaging calculates a surface flux per grid box based on the averaged surface temperature (Section 2.2). This can lead to large errors in the grid box averaged surface fluxes. The averaged homogeneous but artificial surface characteristics deviate less from the "real" surface parameters once the sub-grid scale land-use classes become grid resolved for smaller grids. When comparing flux aggregation with parameter averaging for similar resolutions, the offset in simulated averaged temperatures between both schemes reduces with increasing resolution. Hence, differences in the performance between both schemes become negligible for higher resolutions.

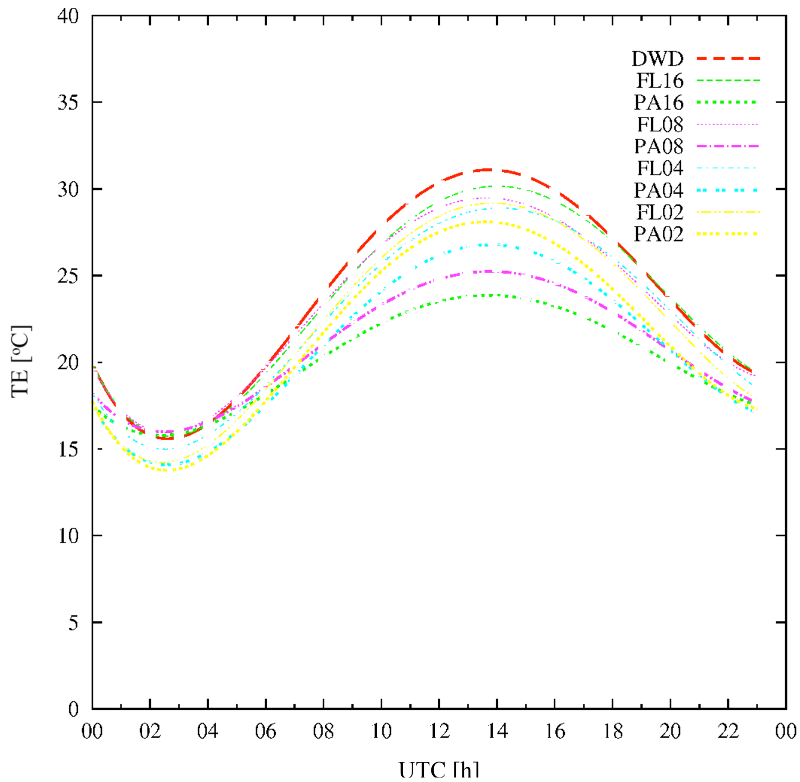


Figure 6.4: Diurnal cycle of 2 m temperature for case E1 averaged over all DWD stations and all four simulation days for measured DWD temperatures and simulated temperatures with METRAS applying flux aggregation and parameter averaging with different resolutions.

According to Figure 6.4 the importance of a well parameterized sensible heat flux for the realistic simulation of the near surface temperatures is strongly reflected by the sen-

sitivity of the diurnal temperature cycle towards the parameterisation of the surface fluxes with regard to the horizontal resolution. Dew point temperature, wind speed and wind direction are less strongly affected by the parameterisation of the surface fluxes.
The comparison of the diurnal cycles illustrated the model performance of the different configurations for case study E1 averaged over all DWD stations within the model domain. In the following, corresponding pairs of DWD measurements and simulations are compared in scatter plots (Figure 6.5 and Figure 6.6) to gain further insight into the dependence of the model sensitivity on the meteorological situation. Following the conclusions from Section 6.3 that I_{lt} is a more meaningful indicator than I_{rh} to determine the importance of the surface fluxes especially in case of very dry situations, corresponding pairs of DWD temperature measurements and simulated temperatures are colour-coded by I_{lt}. E1 is described by an averaged I_{lt} value of 40 indicating a very locally driven meteorological situation with a strong influence of the surface fluxes on the 2 m thermodynamic values. However, the spatial-temporal pattern of I_{lt} values over the whole model simulations ranges from 0 up to 60, since the situations locally differ considerably.

The simulations with flux aggregation (Figure 6.5a, Figure 6.5c, Figure 6.6a and Figure 6.6c) show a tendency for simulating higher locality indices than the simulations applying parameter averaging (Figure 6.5b, Figure 6.5d, Figure 6.6b and Figure 6.6d). For parameter averaging low index situations with an I_{lt} below 40 dominate, while for flux aggregation situations with an I_{lt} higher than 30 are simulated more often. With increasing resolution high index situations occur more often for parameter averaging, which reflects that more grid points are associated with stronger surface fluxes. On the one hand this is due to the fact that more grid points are used for the high resolution simulations. On the other hand the land-use is resolved more explicitly leading to more accurate land-use characteristics and the effect that areas with intense surface fluxes are no longer underestimated by averaging the surface characteristics. For 16 km and 8 km low temperatures are significantly overestimated and high temperatures are underestimated with parameter averaging (Figure 6.5b,d). This was already visible in the averaged diurnal cycle for the 2 m temperatures (Figure 6.4). For flux aggregation differences between the four configurations are much smaller, but flux aggregation shows a tendency to overestimate the lowest temperatures for all configurations.

Low I_{lt} values are distributed over the whole simulated temperature range in case of both parameterisation schemes. The scatter plots underline further that the highest index values are not necessarily connected with the highest temperatures. For flux aggregation two clusters occur for high I_{lt} values. One cluster is connected with low temperatures and the second cluster with higher temperatures (black dots in Figures 6.5a, Figure 6.5c,

Figure 6.6a and Figure 6.6c). The low temperatures being associated with a high I_{lt} index for flux aggregation are clearly associated with evening transitions, night-time values and morning transitions. The second cluster with higher temperatures occurs during day time. Parameter averaging does not show a splitting into two clusters. Instead, black dots indicating high I_{lt} values are in the same place as the cluster for high temperatures for flux aggregation. Since the index values are dependent on the maximum of the friction velocity and the scaling value for the virtual potential temperature the two clusters suggest two situations. The day time cluster describes situations that are very locally driven. The local surface fluxes are intense due to a strong vertical temperature gradient. The night-time cluster is connected with a large friction velocity when the atmosphere is stably stratified. The different behaviour of the flux aggregation and the parameter averaging scheme suggests that the flux aggregation scheme resolves local circulations at night, which the parameter averaging scheme does not capture. For both parameterisation schemes highest I_{lt} values during day time are not necessarily connected with the highest temperatures. But since I_{lt} values are dependent on the maximum of friction velocity and the convective velocity scale and the scaling values for the temperature, they are connected with the largest surface fluxes. Since surface temperatures and surface fluxes peak earlier than 2 m air temperatures within the diurnal cycle this behaviour looks reasonable.

When comparing flux aggregation with parameter averaging differences occur especially for coarse resolutions, when parameter averaging tends to underestimate the observed temperatures especially for low I_{lt} values. This behaviour was already visible in the diurnal cycle of temperatures. With increasing resolution temperatures are simulated more realistically with parameter averaging. Also high index situations seem to occur more often when applying parameter averaging with higher resolutions than coarse resolutions. This looks reasonable since the land-use is resolved more realistically and strong but very local surface fluxes are less smoothed by averaging the surface characteristics.

From this detailed study of case E1, it is evident that flux aggregation with blending height concept is the preferable parameterisation scheme. Its resolution-dependence is small compared to the resolution-dependence of the parameter averaging scheme. The surface energy balance is calculated for each sub-grid scale tile individually. Consequently, surface heterogeneities are taken into account more realistically even for the coarse control run with 16 km resolution (FL16). Parameter averaging does only seem to produce reasonably realistic results when the averaged effective surface parameters are not differing too much from the original land-use characteristics. Parameter averaging does only work for very homogeneous surfaces or where one land-use class clearly dominates the grid box. This is the case for very high resolutions where the sub-grid scale heterogeneity is small, since it is resolved more explicitly. Flux aggregation is especially preferable in low index situations for coarse resolutions where parameter averaging shows a tendency to overestimate low temperatures and underestimate high temperature.

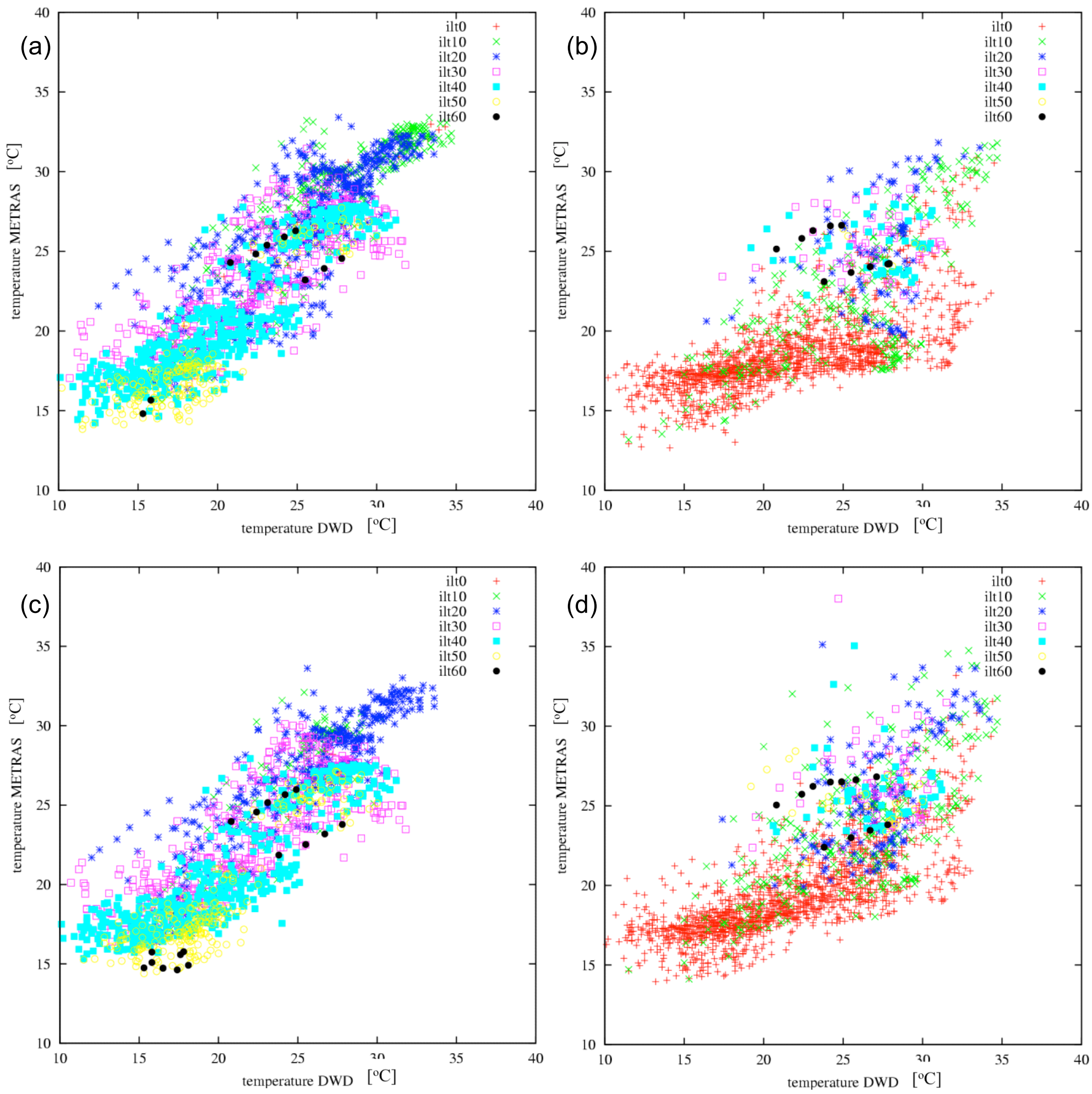


Figure 6.5: Scatter plots of hourly simulated (METRAS) and observed (DWD) 2 m temperature for simulation E1 for (a) FL16, (b) PA16, (c) FL8 and (d) PA8 are colour-coded with I_{lt} index values.

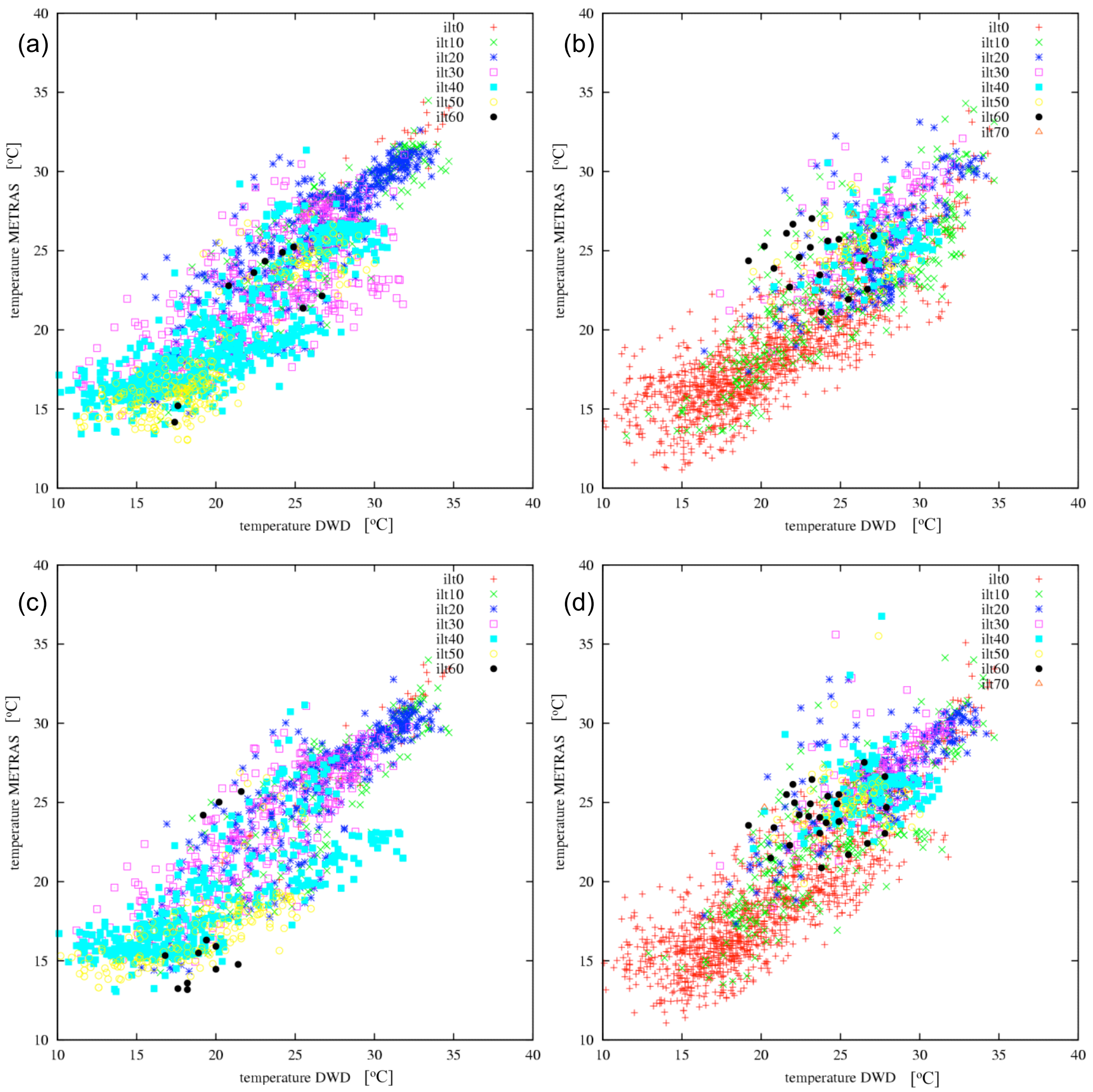


Figure 6.6: Like Figure 6.5 but for (a) FL4, (b) PA4, (c) FL2 and (d) PA2.

6.5 Using a locality index for process oriented model evaluation

So far the locality indices I_{lt} and to a certain extend I_{rh} were successfully applied for a process-oriented model evaluation relating the impact of surface fluxes to model performance for the near surface atmospheric variables. In Chapter 5.4 it was further shown that I_{lt} and I_{rh} can serve as an indicator for precipitation probability. However, this relationship was derived using observed precipitation and not tested for modelled precipitation yet. In the following simulation E2 is used to determine if I_{rh} can also be used to indicate simulated precipitation probability while indicating the importance of surface fluxes on the model solution e.g if precipitation occurring was influenced by local processes.

The locality indices are derived from simulations using METRAS 3D and 8 different configurations. The locality index I_{rh} is calculated at each grid point and for each model output time step for the meteorological situation E2 (Table 6.1). E2 was the only simulation with considerable precipitation. A frequency distribution for I_{rh} is calculated for each of the 8 configurations and compared with the frequency distribution of I_{rh} derived from observations at Lindenberg station. In case the simulated frequency distributions resemble the measured ones, I_{lt} and I_{rh} can probably be applied to verify the impact of the surface fluxes on prognostic variables like temperature, humidity, wind and on precipitation online in a model.

I_{rh} was calculated at each grid point in the model domain for each model output time step. Each I_{rh} index is then related to the forecasted precipitation within the following 6 hours within a radius of 10 km around each grid point. The chosen radius is based on the analysis in (Bohnenstengel et al, 2011a) which showed that precipitation events are representative for an area of up to 25 km x 25 km. The procedure to relate I_{rh} with the occurrence of precipitation within the following 6 hours is similar to the procedure in Section 5.4, where it was already done but for observed precipitation.

Figure 6.7 and Figure 6.8 summarize the accumulated frequency distribution of meteorological situations characterised by I_{rh} for the eight configurations of case E2. In consistence with the analysis of the observed I_{rh} values (Section 5.4), the most extreme values were not taken into account. Therefore, the 99.9 % percentile of I_{rh} data was defined as the upper threshold for the simulated I_{rh} values for each of the 8 configurations. The upper threshold value for I_{rh} varies slightly between 55 and 77 for the different configurations. This range of values is slightly larger than the range of 50-60 determined for year 2002 and 2003 from observations at Lindenberg. In contrast to the measurement based I_{rh} values for Lindenberg the simulations provide I_{rh} values for the whole model domain at each model output time step. Therefore, the number of I_{rh} values above the I_{rh} threshold value for each configuration is still large in contrast to the observed values analysed for Lindenberg (Section 5.4).

The I_{rh} values are indicated by the black symbols in Figure 6.7 and Figure 6.8. All configurations resemble the same distribution of I_{rh} values as the measurement based I_{rh} for Lindenberg (Section 5.4, Figure 5.5). The slope of the measurement based I_{rh} distributions for 2002 and 2003 is resembled by the accumulated distribution of the simulated

I_{rh} values for case E2 regardless of the set-up chosen. The similarity of the frequency distributions of the I_{rh} values between the simulations and the measurement based values is promising in that way that the use of I_{rh} for the process-oriented evaluation of simulations is justified. Therefore, I_{rh} is used in a second step to evaluate the relationship of I_{rh} and precipitation probability in the model simulations.

The observations in the Lindenberg area indicate that precipitation probability increases for more locally driven meteorological situations (Section 5.4). Higher I_{rh} values indicate an increased impact of the local surface fluxes on the generation of precipitation within a radius of 10 km and also indicate an increased precipitation probability for high I_{rh} values. The grey symbols in Figure 6.7 and Figure 6.8 show the precipitation probability. The relationship found for the measurement based I_{rh} values is resembled by some of the set-ups for case E2. However, precipitation probability is high for low I_{rh} situations already with values of up to 50 % for low I_{rh} when using high horizontal resolution. FL16 shows for low I_{rh} values the lowest precipitation probability with values around 22 % and PA4 the highest precipitation probability with values around 53 %. Precipitation is resolved better when using high resolution than low resolution, which explains the increase of precipitation probability for low I_{rh} values with increasing resolution. Altogether, the slope indicating the increase of precipitation probability with increasing I_{rh} is not as pronounced as in the measured analysis for Lindenberg for any of the configurations. However, the maximum precipitation probability occurring for the upper threshold I_{rh} values varies mostly between 50 and 75 depending on the set-up. This is very similar to the measurement based values of around 60 % for I_{rh} values of 50 to 60.

In general, the coarser resolutions show a slight increase in precipitation probability followed by a decrease for the highest I_{rh} values (FL16, FL8, PA8). Only PA16 shows a strong increase of precipitation probability for high I_{rh} values after a dip. For higher resolutions the precipitation probability increases with increasing I_{rh} value and the relationship derived from the Lindenberg data is resembled. The slope is steeper for FL4, FL2 and PA2 and precipitation probabilities reach 75 % in case of FL4 for instance. Only PA4 does not resemble the relationship from the measurement based data.

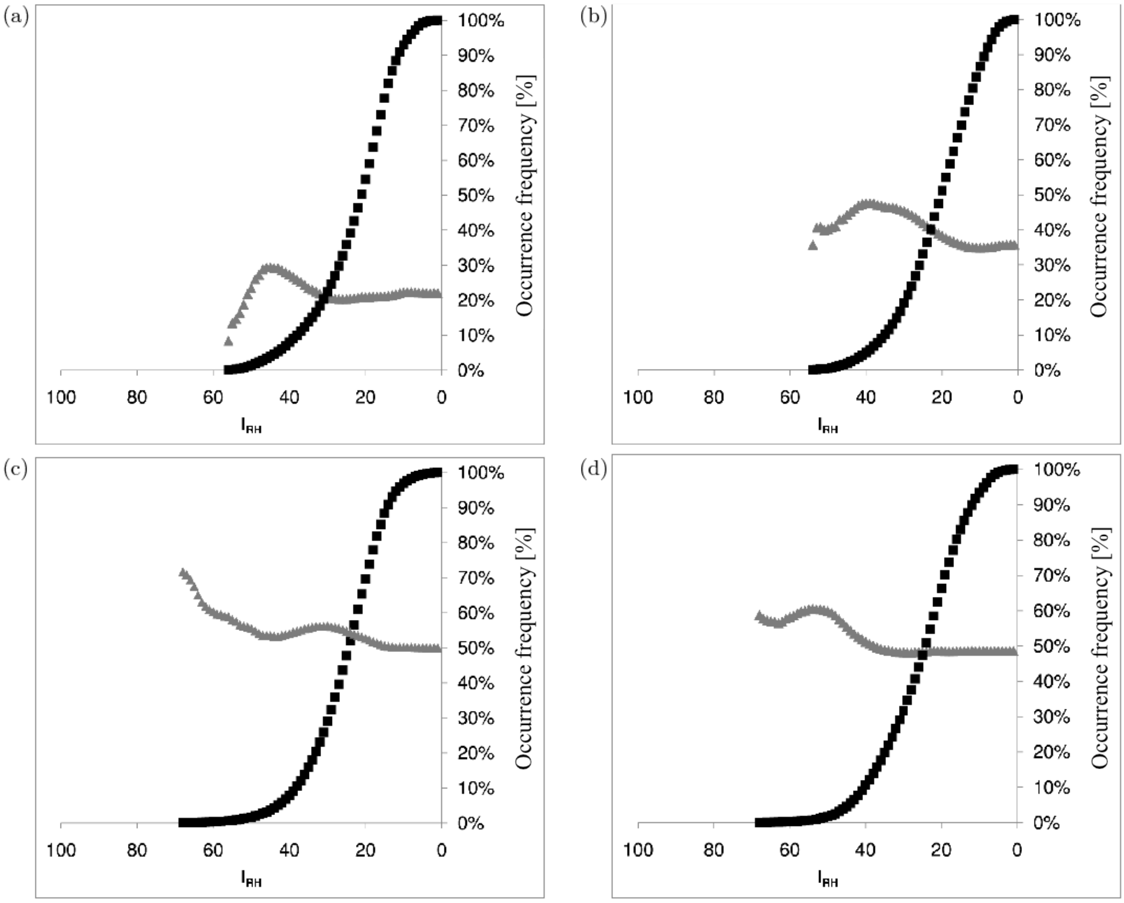


Figure 6.7: Black squares depict the accumulated relative frequency of simulated meteorological situations I_{rh} for configuration (a) FL16, (b) FL8, (c) FL4 and (d) FL2. Grey triangles depict the fraction of meteorological situations with precipitation in the following 6 hours compared to the total number of all situations above the corresponding I_{rh} value.

In general, higher resolution configurations like FL4 (Figure 6.7c) and PA2 (Figure 6.8d) capture the observed increase of the precipitation probability for larger I_{rh} more accurately than the coarse resolution configurations like FL16 (Figure 6.7a). The configurations with high horizontal resolution resolve local heterogeneities better than the coarse ones. And precipitation occurs often in the neighbouring grid box. This is not captured well with coarse resolution where the grid resolution is smaller than the radius of 10 km chosen for the analysis.

Increasing the horizontal resolution improves the relationship for flux aggregation. For instance, using a horizontal resolution of 4 km for flux aggregation FL4, (Figure 6.7c) alters the relationship between I_{rh} and precipitation probability compared to 8 km horizontal resolution FL8 (Figure 6.7b) and FL16 (Figure 6.7a). The precipitation probability is increasing from 50 % to 70 % for I_{rh} values larger than 20. In contrast, precipitation probability decreases after reaching a local maximum for FL16 and FL8. For parameter averaging a general improvement in the relationship is not found for increasing resolution. While PA16 and PA2 show an increasing precipitation probability with I_{rh}, PA8 and PA4 do not show a clear relationship.

The results show that the surface flux scheme does have an impact on the relationship between I_{rh} and precipitation probability. Only in case of 2 km resolution both schemes, flux aggregation and parameter averaging, show a similar increase in precipitation probability with increasing I_{rh} like in the measurement based analysis. When applying coarser resolution the results in the relationship between I_{rh} and precipitation probability start to differ more. The largest differences occur for 16 km resolution between flux aggregation and parameter averaging, although the configurations only differ with regard to the surface flux parameterisation scheme. While the simulation applying flux aggregation shows a local maximum for I_{rh} values around 50 (Figure 6.7a), parameter averaging (Figure 6.8a) shows a similar relationship as found in Section 5.4 for the observed precipitation measurements with increasing precipitation probability for higher I_{rh} values. For 8 km both schemes indicate an increase in precipitation probability with I_{rh} followed by a decrease for higher I_{rh} values. And FL4 shows the measurement based relationship, while PA4 does not. The differences between the parameterisation schemes become smaller for higher resolution and the I_{rh} precipitation probability relationship more pronounced. Both approaches show a precipitation probability of 0.5 for low I_{rh} values between 0 and 40. For both set-ups the precipitation probability then similarly increases to 60.

The relationship between I_{rh} and precipitation probability derived from measurements in Section 5.4 cannot be reproduced with the coarse 16 km resolution by flux aggregation and also PA16 shows a dip before the increase in precipitation probability. One reason is the less well simulated precipitation due to the coarse resolution. However, when increasing the horizontal resolution flux aggregation shows an increase of precipitation probability with increasing I_{rh}. For parameter averaging the increase in precipitation probability is less well pronounced than for flux aggregation. But both schemes show a similar behaviour for 2 km resolution and resemble the relationship found for Lindenberg. This indicates that only when the surface processes and the precipitation processes are parameterised well, the relationship found from the measurement based analysis in Lindenberg holds.

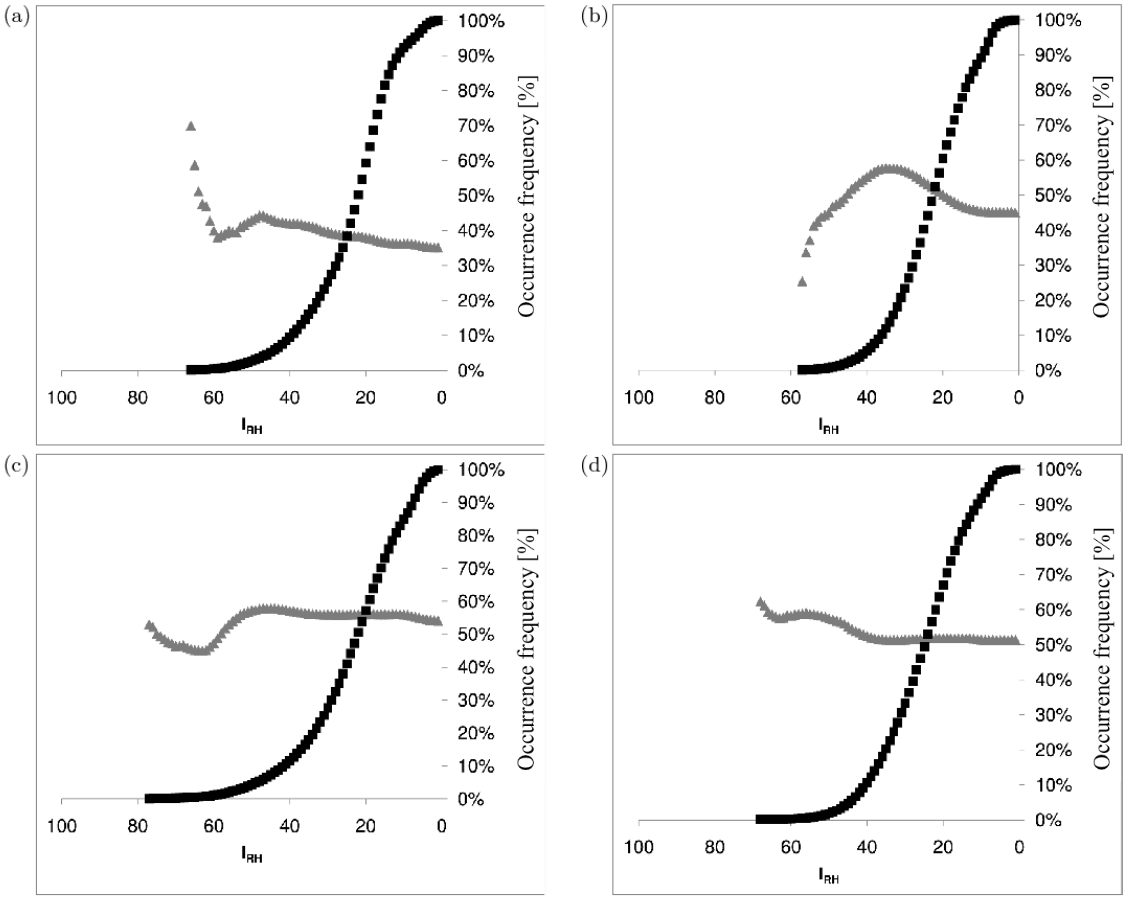


Figure 6.8: Same as Figure 6.7 but for (a) PA16, (b) PA8, (c) PA4, (d) PA2.

The relationship between I_{rh} and precipitation probability derived from observations looks very robust, since it could be established for two very different years (Section 5.4). The relationship is reflected for modelled precipitation with both parameterisation schemes, when using high horizontal resolutions of a few kilometres (< 4 km). From the modelled results it is visible that both parameterisation schemes lead to slightly different results when applying the same resolution. Hence, the surface layer characteristics are simulated slightly different, which leads to differences in the near surface thermodynamic and dynamic fields and affects precipitation. For 2 km resolution the relationship between precipitation probability and I_{rh} has a similar distribution for flux aggregation and parameter averaging. This indicates that differences between both parameterisation schemes tend to vanish for higher resolution. This behaviour underlines the necessity to determine the applicability range of the surface flux parameterisation schemes with respect to the meteorological situation as well as the resolution.

The results show that differences caused by resolution and surface flux parameterisation scheme do not only occur for very locally driven meteorological situations (high I_{rh} values) but also at less locally driven sites. The parameterisation schemes seem to affect precipitation and in the following the temperature and moisture fields. Since the I_{rh} index is based on surface layer scaling variables, which trigger the sensible and latent heat flux and hence the turbulent exchange between the surface and the boundary layer, I_{rh} also estimates the impact of the surface on the lower boundary layer thermodynamic and dynamic fields as well as the generation of precipitation. Therefore, I_{rh} is assumed to be a reasonable tool to identify the dependence of the model performance on the resolution, parameterisation scheme and especially the meteorological situation itself.

6.6 Conclusions

Results of 6 different meteorological situations most of these run in six or eight different configurations were evaluated against DWD measurements. The evaluation took into account the importance of the parameterised local surface processes by indexing the situations with locality indices I_{lt} and I_{rh}. A more detailed analysis was undertaken for two of the six case studies. E1 as a very dry situation was chosen to test the ability of I_{lt} to evaluate model results based on the strength of the parameterised surface fluxes (Section 6.4). The locality index I_{lt} was extended to include relative humidity leading to I_{rh} and applied to test the relationship between precipitation probability and locality of the meteorological situation as derived from measurement based indices (Section 5.4) to predict precipitation probability in the model simulation for the very wet situation E2 (Section 6.5).

The surface fluxes have a large impact on the thermodynamic values in case of locally driven situations indicated by I_{lt} values above 40. The thermodynamic values are directly influenced by the surface heat fluxes. Wind speed and direction are more affected by the surface drag, which is not that dependent on the surface energy balance and less heterogeneous.

For the presented case studies flux aggregation showed only a small dependence on the horizontal grid resolution, whereas parameter averaging performed better for higher resolutions. For coarse resolutions parameter averaging does not show a reasonable model performance and flux aggregation should be used instead. For high resolutions of the order of 2 km parameter averaging shows similar results as flux aggregation. It might then become the favourable scheme, since it is less computationally expensive. However, coarser resolution of the order of 4 km with flux aggregation showed similar results as flux aggregation with 2 km. Instead of increasing the resolution further, coarser model simulations with flux aggregation might be sufficient for some studies for the sake of computing time.

A more detailed analysis was undertaken for case study E1. The different configurations of E1 were evaluated by classifying results by the local influence of the surface fluxes using I_{lt}. In general, for flux aggregation coarse and hence computationally less expensive configurations resulted in similar results as when applying high resolutions, especially for overall low index situations. In contrast, parameter averaging does only produce reliable results for situations dominated by advection, where the influence of the surface fluxes is small. For high index situations parameter averaging does only produce reliable results for highly resolved surface heterogeneities. In fact, the formerly sub-grid scale land use becomes grid-scale now and is therefore resolved more explicitly. The results suggest that I_{lt} allows evaluating model results based on the influence of the surface fluxes in a process oriented way. Furthermore, the results of this case study suggest that the locality index I_{lt} can be used to determine the best parameterisation scheme for the surface fluxes dependent on the meteorological situation. However, the results also suggest that this makes only sense for the thermodynamic part. Based on this outcome, I_{lt} is suggested for use as an indicator to determine the best surface flux parameterisation scheme online during model simulations.

The applicability of I_{lt} to evaluate or influence model results during the ongoing simulation was tested further in Section 6.5 when relative humidity was included into I_{lt} resulting in I_{rh}. The results underlined the conclusion drawn in Chapter 5 that I_{rh} can be used as a predictor for precipitation probability within the following 6 hours within a radius of 10 km. I_{rh} showed the same frequency distribution as the I_{rh} values derived from observed data in Section 5.4. A precipitation probability distribution was found for higher horizontal resolutions that is similar to the measurement based one in Section 5.4. It suggests a higher precipitation probability for more locally driven meteorological situations.

7 Sensitivity of mesoscale model results to parameterisation of sub-grid scale surface fluxes for a locally driven meteorological situation in the area of Berlin

Parts of this Chapter have been published as:

Bohnenstengel S., Schlünzen K.H. (2009): Performance of different sub-grid scale surface flux parameterizations for urban and rural areas. In: Baklanov A., S. Grimmond, A. Mahura, M. Athanassiadou (Eds), Urbanisation of meteorological and air quality models, Springer Publishers, 165p., ISBN 978-3-642-00297-7

7.1 Introduction

More than half of the world's population lives in urban areas and this percentage still increases. In terms of human comfort coinciding with an increased mortality rate during heat waves, an accurate prediction of for instance urban temperature fields is very important. Besides a more accurate prediction of the thermodynamic and dynamic fields in urban areas a better representation of urban effects also pays back on e.g. air pollution modelling. Then the near surface conditions are most important for an accurate prediction, since the dispersion modelling strongly depends on the processes in the lowest model levels. With increasing model resolution towards the kilometre scale urban areas are much better resolved now than years ago. But due to the very heterogeneous surface structure urban effects are still not resolved in detail but remain on the sub-grid scale. Numerical models capture their impact on the local weather and climate by accounting for the effects which distinguish urban areas from rural ones. Parameterisations for the urban surface energy balance have been developed with many levels of complexity over the recent years.

The main urban features that have to be captured by any mesoscale model can be split into dynamic and thermodynamic features. From the perspective of the city scale an urban canopy consists of a higher roughness compared to rural areas and the inhomogeneities in the horizontal and vertical are larger. Energy from the main flow approaching the urban areas is transferred into turbulent motion. In the urban canopy layer the Reynolds stress is largest at the roof level reducing the wind speed and consequently increasing the turbulent exchange in the urban boundary layer. One way to describe the interaction of the urban canopy layer with the atmospheric layer is to describe the turbulent transport out of the urban canopy layer into the urban boundary layer in terms of a bulk approach, where the lowest model level should at least be above the top of the buildings. Unfortunately, Monin-Obukhov similarity theory is not applicable within the urban canopy layer, where the wind profile does no longer follow a logarithmic but an exponential law up to a height of about 2.5 of the average building height (Roth, 2000). Since METRAS uses a higher local roughness length to describe the frictional impact of urban areas on the flow, the mean wind profile is still assumed to be logarithmic according to Monin-Obukhov similarity theory in contrast to observations made in urban areas. This is justified, since METRAS, as is used here, does not intend to simulate the wind or temperature within urban street canyons but merely the effect of urban areas on the wind and temperature field above the urban canopy layer.

The humidity budget in urban areas differs from rural budgets. The altered urban surfaces are usually dryer than rural surfaces due to the larger amount of sealed surfaces and the lower ability of the soil materials to store water. On the other hand, irrigated gardens and parks are a source of humidity during dry periods in summer.

The building geometry affects the radiation budget by shading and trapping of radiation within the street canyons. Additionally, the urban area exposes a larger surface to the atmosphere due to the elevated surfaces dependent on the geometry. And the material properties of the urban fabric have a higher capacity to store heat than rural areas (Harman, 2003). This increases the thermal inertia of the urban canopy compared to the rural areas. Urban areas store more heat than rural areas. The large urban thermal inertia causes a phase delay of the sensible heat flux in urban areas compared to rural areas. The large urban thermal inertia also dampens the amplitude of the sensible heat flux and as a consequence also the diurnal cycle of the surface and air temperatures (Porson et al., 2009) The large urban storage term maintains higher urban surface temperatures than rural surface temperatures during the night and drives a positive sensible heat flux in urban areas. This then causes urban screen level temperatures to be warmer at night than in the rural surrounding (Bohnenstengel et al., 2011b). Altogether, the larger thermal inertia favours the development of a nocturnal urban heat island (UHI), especially under very locally driven meteorological situations with low wind speeds and clear skies. Additionally, the altered urban surface energy balance results in an overall larger bulk Bowen ratio compared to rural areas (Schlünzen et al., 2010). While the larger thermal inertia dampens the amplitude of the sensible heat flux and causes a phase delay, the sealed surfaces increase the sensible heat flux on the expense of the latent heat flux in urban areas. During daytime the decreased evapotranspiration enhances the sensible heat flux and leads to enhanced screen level temperatures.

Since the urban surface energy balance determines the sensible and latent heat fluxes, the mean profiles of temperature and moisture in urban areas are affected. Both are linked to the dynamically driven urban processes via their dependence of the wind profile on the atmospheric stability and of the fluxes on wind speed. This even pays back on the depth of the urban boundary layer, which is considerably deeper at night than its rural counterpart (Bohnenstengel et al., 2011b, Hunt et al., 2011). During the night the sensible heat flux feeds into a much smaller volume than during daytime. Therefore, the surface fluxes can have a much bigger impact during the night and their accurate simulation is essential for the faithful simulation of the near surface variables.

While microscale models resolve urban areas with high accuracy (e.g. Bohnenstengel et al. 2004), urban areas are currently represented by very different levels of complexity in mesoscale models. Since regional simulations are applied on the mesoscale range, the flow around urban canyons and houses is not resolved yet, but the buildings effects are parameterised for model resolutions in the range of kilometres. More sophisticated attempts to parameterize urban areas can be found in the literature. Harman and Belcher (2006) describe the urban geometry in a simplified 2D form, but use a very sophisticated surface energy balance scheme based on a resistance network approach, which has been applied to simulate the London urban heat island (Bohnenstengel et al., 2011b). Dupont and Mestayer (2006), Martilli et al. (2002) and Masson (2000) use similar approaches. All their attempts have in common a two-dimensional description of the urban building morphology. Martilli et al. (2002) add more complexity by using multiple levels within the urban canopy and by calculating the wind within the canopy via a drag approach. In this study we apply a much simpler, less input demanding and more cost-efficient approach by only enhancing the roughness length and adjusting relevant surface parameters like Albedo values for urban areas. This approach is supported by a study undertaken by Sue Grimmond showing that more complex urban parameterisation do not necessarily improve model results (Grimmond et al., 2010, 2011).

The rural to urban transition regions and the urban areas itself are very heterogeneous in terms of their surface characteristics like roughness, soil humidity, temperature, heat storage capacity and vegetation. Due to the complexity of urban areas, most of the heterogeneous land-use in urban areas is treated as sub-grid scale. Therefore, aggregation schemes are used to derive the grid box representative bulk surface fluxes, which interact with the main flow at the lowest atmospheric model level and determine the vertical profiles of the prognostic variables like temperature. Most of the aggregation methods have in common that they show good results for more or less homogenous areas or heterogeneous areas with no extreme roughness changes or very different surface fluxes within one grid cell. Giorgi and Avissar (1997) mention the so called “aggregation effect”, which affects the representativeness of the grid box averaged fluxes, since the system acts highly non-linear. This is especially the case for urban areas. Here irrigated gardens or parks may result in a low Bowen ratio with an intense heat release due to the latent heat flux next to sealed surfaces made of concrete with a very high Bowen ratio. Both extremes may occur within the same grid box. Both fluxes, however, need to be represented by some kind of grid box averaged flux. Vertical turbulent exchange and flow field react sensitive towards the parameterisation applied for sub-grid scale surface fluxes of heat and momentum. In this paper we investigate two types of sub-grid scale parameterisations for surface fluxes. The parameter averaging method calculates the grid representative surface fluxes dependent on grid box averaged surface parameters. This method has the advantage to be very cost-efficient. The more sophisticated flux scheme is a flux aggregation scheme as described by von Salzen et al. (1996), which applies a blending height concept following Claussen (1990). For the latter method surface fluxes are calculated separately for each sub-grid scale land-use class and the resulting grid box flux is the fraction weighted average of these. Similar methods are derived by Heinemann and Kerschgens (2005) as well as Ament and Simmer (2006).

The model performance also depends on the resolution applied (Schlünzen and Katzfey, 2003 and Chapter 6). Mesoscale models and their parameterisations are mostly designed for coarser resolutions. Their performance is validated for rural areas and specific meteorological situations but rarely for urban areas. One problem is that measurements in urban areas are usually not of large spatial representativeness and thus not comparable with model results. The lowest model level is usually assumed to be within the inertial sub-layer, where the Monin-Obukhov similarity theory approach used in the models to describe the surface fluxes is valid. However, routine measurements taken in the urban environment are located within the urban canopy, where the Monin-Obukhov similarity theory approach used in numerical models is not valid due to the horizontal inhomogeneity of the fluxes.

The aim of the paper is firstly to determine if simple urban parameterisation schemes can faithfully represent urban processes on the mesoscale. Further, the two parameterisation schemes for averaging the surface fluxes are tested for the case of the very heterogeneous urban of Berlin (Bohnenstengel and Schlünzen, 2009; Schlünzen et al., 2011). For this purpose, sensitivity studies are carried out for the urban area Berlin. Berlin is located in North-Eastern Germany, a region with relatively flat but very heterogeneous land-use. Thus, orographically induced effects are negligible. The simulations are undertaken with the mesoscale transport and fluid model METRAS 3.0, whose development is coordinated by the Meteorological Institute, University of Hamburg. METRAS is run with different resolutions of 16 km, 8 km, 4 km and 2 km with two alternative parameterisations of the sub-grid scale land-use effects. The flux aggregation scheme with blending height concept and a parameter averaging scheme are tested for a very dry situation in the year 2003. The results are evaluated with routine observations of the "Deutscher Wetterdienst" (DWD), since flux measurements were not available for the Berlin area.

The paper first details the relevant model aspects of METRAS (Section 7.2). The domain and the model set-up are described in Section 7.3. Model results for a very locally driven summer day are discussed in Section 7.4 and conclusions on the applicability of aggregation schemes dependent on the model resolution are drawn in Section 7.5.

7.2 METRAS model description and urban scheme

For this sensitivity study the mesoscale transport and fluid model METRAS is used. METRAS is a non-hydrostatic prognostic mesoscale model using terrain-following coordinates. It is based on the primitive equations and solves the equations in flux form. Designed for model applications of the meso-beta and meso-gamma scales METRAS assumes the anelastic approximation, the Boussinesq approximation and sets the Coriolis parameter constant in the current version. For taking into account large-scale meteorological conditions METRAS is nested into ECMWF analysis. Detailed information on the model can be found in Schlünzen (1990) and Schlünzen et al. (1996a) and Dierer and Schlünzen (2005). Here, only those parts of METRAS relevant for the current sensitivity study are briefly described.

7.2.1 Surface energy budget

The surface energy balance in METRAS 3.0 consists of the direct and diffuse solar radiation S with Albedo A, the long wave radiation terms (L_{in} and L_{out}), the sensible (Q_H) and latent (Q_E) surface turbulent heat fluxes and a ground heat flux (Q_S). The METRAS version 3.0 used for this study does not account for an anthropogenic heat flux term.

$$(1-A)S_{in} + L_{in} + L_{out} + Q_H + Q_E + Q_s = 0 \tag{7.1}$$

The heating and cooling rates at the surface and at each grid point in the atmosphere due to absorption and reflection of radiation by water vapour and liquid water are calculated with a two-stream approximation according to Bakan (1995). A small inconsistency with the sub-grid scale land-use scheme does occur here, since the short wave radiation is calculated accounting for the Albedo of the sub-grid scale land-use classes, whilst the emitted long wave radiation is based on the effective surface temperature of the whole grid box and not on the sub-grid scale temperature of the individual sub-grid scale land-use classes. However, even for a sub-grid scale surface temperature difference of 5 K, the resulting heterogeneity in emitted long wave radiation would be relatively small. To give an example: according to the Stefan Boltzmann law $\varepsilon\sigma\left(278.0^4 - 273.0^4\right) \approx 23.7\mathrm{Wm}^{-2}$ with $\varepsilon=1$ or even smaller with lower emissivity values, which is about 7 % of the maximum emitted long wave radiation in this case. In case of 10 K difference, the heterogeneity is of the order of 13 %, which is not negligible but still only about 1/4 of the influence of heterogeneity on turbulent surface fluxes, which can be up to 40 % in suburban areas according to Schmid et al. (1991).

The soil is coupled conductively to the atmosphere by calculating the surface temperature at the surface following Tiedtke and Geleyn (1975) and Deardorff (1978) applying a force-restore method. For the heat exchange with the soil (Q_S) the sub-grid scale parameters of thermal diffusivity κ_s and thermal conductivity ν_s are considered (eq. 7.2) for each land-use class and the conduction of heat is calculated following the diffusion equation (7.3). Values for the soil properties like κ_s, ν_s are chosen according to Table 5 in Schlünzen at al. (1996b). The surface temperature T_s is calculated for each sub-grid scale land-use class following eq. (7.4). The calculation of T_s takes into account a changing soil temperature by calculating $\overline{T}(-h_\theta)$ as the lower boundary condition following Deardorff (1978). h_θ is the depth of the temperature wave within the soil and t is the local time. In case of a cloud free sky the emissivity ε is set to 0.22 following de Jong (1973). The Stefan Boltzmann constant σ is $5.67{*}10^{-8}$. ρ_0 is the density of air at the surface, θ_*, q_* and u_* are the scaling values for potential temperature, specific humidity and the friction velocity, respectively. l_{21} is the latent heat of vaporisation of water, Z(t) is the zenith angle and μ is calculated to be 0.75 (1-A) with A being the Albedo over land. Note that overbars denote a grid box representative value. $\overline{T}(-h_\theta)$ is kept constant for short-term forecasts.

$$Q_s = \nu_s \left(\frac{\partial \overline{T}}{\partial z} \right)_s \tag{7.2}$$

$$\frac{\partial \overline{T}}{\partial t} = \frac{\partial}{\partial z} \left(\kappa_s \frac{\partial \overline{T}}{\partial z} \right) \tag{7.3}$$

$$\frac{\partial \overline{T}_s}{\partial t} = \frac{2\pi\pi_s}{\nu_s h} \left(\mu I_\infty \cos Z(t) - \varepsilon\sigma \overline{T}_s^{\,4} + c_p \rho_0 \theta_* u_* + l_{21} \rho_0 q_* u_* - \sqrt{\pi} \nu_s \frac{\overline{T}_s - \overline{T}(-h_\theta)}{h_\theta} \right) \tag{7.4}$$

$$h = \sqrt{\kappa_s \tau_1}, \text{with } \tau_1 = 86400 \text{ s.} \tag{7.5}$$

7.2.2 Surface fluxes

For calculating surface fluxes Monin-Obukhov similarity theory is assumed using the stability functions of Dyer (1974). Above the surface layer a first order closure is applied as described in eq. (7.7) - (7.9) when using boundary following coordinates. A mixing-length approach is used for stable or nearly neutral conditions, where the mixing length l_n (eq. 7.6a) is calculated following Blackadar (1962) with the maximum mixing length λ usually calculated following eq. (7.6) but here being set to 200 m. κ is the Von Kármán constant. The turbulent fluxes are proportional to the local mean gradients of the transported variable and the counter gradient terms Γ_θ and Γ_q in eq. (7.8) and eq. (7.9) are zero for these stratifications. A counter-gradient scheme is applied (Lüpkes and Schlünzen, 1996) in case of convective situations in the flux functions for heat (eq. 7.8) and moisture (eq. 7.9). Γ_θ and Γ_q are the counter gradient terms for potential temperature θ and specific humidity q. This assures vertical turbulent mixing even in a well-mixed boundary layer. The equations are given in terrain-following coordinates. K denotes the turbulent exchange coefficient, i and j the horizontal direction of the coordinate. Overbars indicate mean values and dots terrain following coordinates in horizontal direction ($\dot{x}^1, \dot{x}^2$) and surface normal direction ($\dot{x}^3$).

$$l_n = \frac{\kappa z}{1 + \frac{\kappa z}{\lambda}} \tag{7.6a}$$

$$\lambda = 0.007 \frac{u_*}{f} \tag{7.6b}$$

$$-\rho_0 \overline{u'_i u'_j} = \rho_0 K_{ij} \left\{ \frac{\partial \bar{u}_i}{\partial \dot{x}^k} \frac{\partial \dot{x}^k}{\partial x^j} + \frac{\partial \bar{u}_j}{\partial \dot{x}^k} \frac{\partial \dot{x}^k}{\partial x^i} \right\} \tag{7.7}$$

$$-\rho_0 \overline{w'\theta'} = \rho_0 K_{vert,\theta} \left\{ \frac{\partial \bar{\theta}}{\partial \dot{x}^3} \frac{\partial \dot{x}^3}{\partial z} - \Gamma_\theta \right\} \tag{7.8}$$

$$-\rho_0 \overline{w'q'} = \rho_0 K_{vert,q} \left\{ \frac{\partial \bar{q}}{\partial \dot{x}^3} \frac{\partial \dot{x}^3}{\partial z} - \Gamma_q \right\} \tag{7.9}$$

METRAS version 3.0 accounts for 10 different land-use classes by distinguishing the roughness length, Albedo, thermal diffusivity, thermal conductivity, depth of temperature wave, soil water availability and the saturation value for water content. These values are set according to Table 5 in Schlünzen et al. (1996b). The applied 10 land-use classes are water, mudflats, sand, mixed land-use, meadows, heath, bushes, mixed forest, coniferous forest and urban areas. Two different parameterisation schemes are applied to derive the grid box averaged surface fluxes of momentum, sensible heat flux and latent heat flux representative for each grid box. Both schemes calculate area averaged values of the scaling values friction velocity u_*, free convection velocity w_* and the scaling value for potential temperature (θ_*) and for humidity (q_*) for each grid box.

The parameter averaging scheme is the favourable one concerning computing time. It calculates fraction weighted average parameters. For the roughness length z_0 this is given in eq. (7.10). The roughness length is calculated from the different z_{0i} of the sub-grid scale land-use classes. The resulting $\langle z_0 \rangle$ is an artificial homogeneous roughness length used as being representative for the whole surface characteristics of each grid box.

$$\langle z_0 \rangle = \sum_{i=1}^{n} f_i z_{0i} \tag{7.10}$$

The rationale behind this method is the assumption that the surface fluxes are in equilibrium with the averaged homogeneous artificial surface characteristics of the whole grid box. This assumption performs quite well for nearly homogeneous landscapes that are not too distinct in their surface characteristics. For very heterogeneous areas non-linear effects, which contribute to the grid-box averaged fluxes are not captured (Giorgi and Avissar, 1997), since the flux functions depend in a non-linear way on the surface layer characteristics. While some of this effect is relatively easy to capture in the grid average $\langle z_0 \rangle$ by e.g. averaging the $\ln(z_{0i})$ terms, a simple approach is not available for κ_s or ν_θ. For simplicity the linear parameter averaging is used for all parameters. By this approach an averaging over the surface characteristic variables is done. Applying the flux function then suppresses some non-linear effects. This aggregation effect worsens model performance in some cases.

The non linearity of the fluxes is more accurately described with a flux aggregation method. In METRAS an averaging based on the blending height concept (Claussen, 1990) is applied (von Salzen et al., 1996). In heterogeneous areas this method theoretically performs better than the parameter averaging method, since it calculates surface fluxes for each of the sub-grid scale land-use classes independent of each other. The fluxes are then aggregated at a certain height weighted by the fractional land-use. The sub-grid scale land-use class itself is considered to be in itself homogeneous. Then it can be assumed that the sub-grid scale surface fluxes are in local equilibrium with the "homogeneous" sub-grid scale surface of the individual land-use classes. With increasing height the sub-grid scale surface fluxes are less and less in equilibrium with the underlying sub-grid scale land-use class they are calculated for. Instead, the surface fluxes tend to be in equilibrium with the effective surface characteristics of the whole grid box. At the height at which the flow is in equilibrium with the underlying heterogeneous surface the model does no longer distinguish the effects of the local surface characteristics; this height is the so-called blending height (von Salzen et al., 1996, Claussen, 1990). Above the blending height the fraction weighted aggregated fluxes are assumed to be in equilibrium with the effective characteristics of the underlying heterogeneous surface. The blending height is a function of the characteristic length scale of the surface heterogeneities and also depends on the atmospheric stability. In METRAS the blending height is calculated iteratively. However, the surface fluxes are not averaged at the blending height but at the first vertical atmospheric model level at 10 m, since advective processes become more important further away from the surface.

In METRAS, the sub-grid scale fluxes are calculated for each land-use class based on the specific roughness lengths for momentum, temperature and humidity. As an example the latent heat flux is given in eq. (7.11).

$$\begin{aligned}\rho l_{21} q_* u_* &= \rho l_{21} \sum_{i=1}^{n} f_i \, q_{*i} \, u_{*i} \\ &= \rho l_{21} \kappa^2 U(z_1) \\ &\cdot \sum_{i=1}^{n} f_i \cdot \left(q(z_1) - q(z_{0qi})\right) \cdot \left[\left(\ln\left(\frac{z_1}{z_{0i}}\right) - \Psi_m\left(\frac{z_1}{L_i}\right)\right) \cdot \left(\ln\left(\frac{z_1}{z_{0qi}}\right) - \Psi_q\left(\frac{z_1}{L_i}\right)\right)\right]^{-1}\end{aligned} \tag{7.11}$$

Specific humidity is denoted by q, U(z) is the main flow at height z_1, ψ_m and ψ_q are the stability functions for momentum and humidity according to Dyer (1974). The von Kárman constant κ is set equal to 0.4. z_{0qi} is the roughness length for specific humidity q for land-use class i and l_{21} is the latent heat of evaporation. The blending height l_b is calculated following eq. (7.12). f_i ($0 \le f_i \le 1$) is the fraction of land-use i in the grid cell.

$$\frac{l_b}{L_x}\left(\ln\frac{l_b}{z_0}\right) = c_1 \kappa \tag{7.12a}$$

$$\left(\ln\frac{l_b}{z_0}\right)^{-2} = \sum_{j=1}^{N} f_j \left(\ln\frac{l_b}{z_{0j}}\right)^{-2} \tag{7.12b}$$

z_0 is the effective roughness length, L_x a characteristic length scale for the surface heterogeneity. The surface energy balance is calculated for each land-use class separately leading to different surface temperatures $T_{s,i}$ for the different land-use classes. This leads to different stability regimes for the different land-use classes. The same method is used to calculate the scaling value for specific humidity q_*.

This approach works reasonably well as long as the surface characteristic length scales are not too small. For small-scale heterogeneities, the assumption of local equilibrium within a sub-grid scale land-use class is not fulfilled. METRAS is used with a resolution down to 2 km in the present sensitivity studies with input land-use data of a resolution of 30" (~1 km). Hence, the assumption of local equilibrium should be fulfilled in the set-ups used in this study, since sub-grid scale heterogeneities below a resolution of 30" are not captured by the land-use data.

The representation of the urban areas is kept simple by adjusting the roughness length for momentum to 0.7 m. The soil properties mentioned earlier are adjusted to represent the effect of semi-sealed surfaces like the typically mixed sealed and build-up areas with surrounding green areas typically found in Northern German towns and cities. These properties are listed in Table 2 in Schlünzen and Katzfey (2003). The roughness length for heat $z_{0\theta}$ and moisture z_{0q} is set to 1/10 of the roughness length for momentum for vegetated surfaces. For the urban areas it is calculated based on the roughness Reynolds number $Re=u_* z_0/\nu$ and the Prandtl number Pr (eq 7.13). z_{0q} is calculated similar to $z_{0\theta}$.

$$\frac{z_0}{z_{0\theta}} = \min\left(\exp\left(\kappa\left(7.3\mathrm{Re}^{1/4}\sqrt{\mathrm{Pr}} - 5\right)\right), 50000\right), \text{ with} \qquad (7.13)$$
$$\mathrm{Pr} = 0.71$$

Even for the simulations with a high horizontal resolution of 2 km cities are not completely sealed. This is the reason why the land-use characterising variables of the urban area account for a certain amount of vegetation to represent the impact of green in Northern German cities. The need to incorporate the impacts of gardens into the urban land-use tile depends on the resolution of the land-use data itself. And it depends on the dataset, if gardens are grouped into vegetated land-use classes or not. For the CORINE land-use data (CORINE Land Cover, 2004) they are grouped into the urban land-use class. Therefore, the urban land-use parameters represent the inclusions of some vegetation in the urban tile.

The surface energy balance is solved as for the other land-use classes balancing the net radiation consisting of incoming and outgoing short wave radiation as well as incoming and outgoing long wave radiation with the ground heat flux, the sensible heat flux and the latent heat flux. No additional anthropogenic heat flux is introduced in the present studies.

7.3 Sensitivity study for different model configurations

The simulations are undertaken for a 400 km x 400 km domain in North-Eastern Germany. A large heterogeneity of the sub-grid scale land-use is found in this area. However, orography changes are small, since the whole domain is relatively flat without high mountains. Berlin is situated about in the middle of the domain; this allows studying the local solution of METRAS away from the lateral boundaries. The land-use data are derived from the CORINE Land Cover data set for Germany on a 30” grid. Figure 7.1 shows the main land-use in the model domain for a horizontal resolution of 2 km.

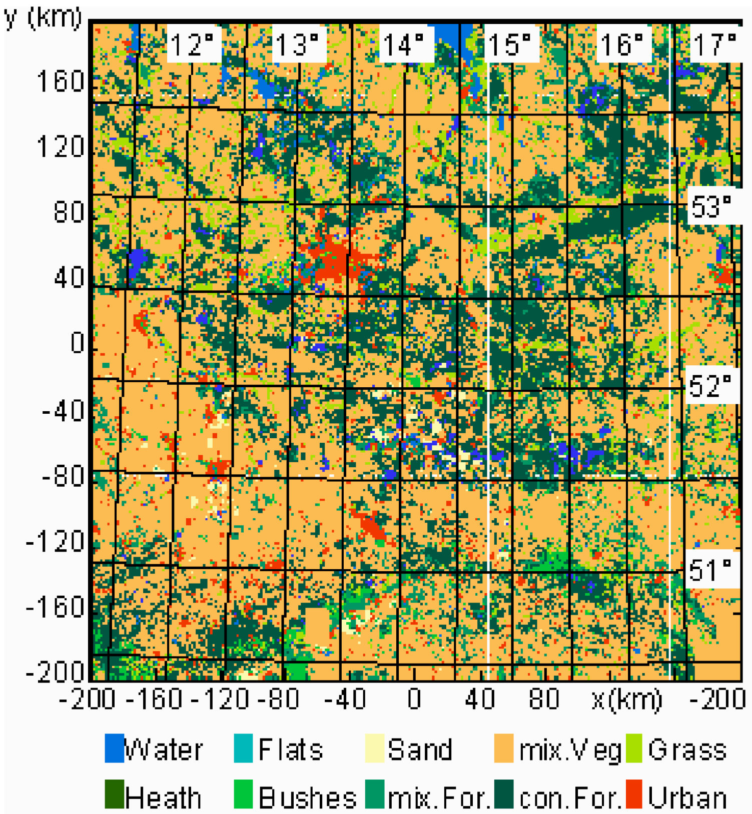


Figure 7.1: Main land-use within the whole model domain for 2 km horizontal resolution. Berlin is characterised by the large red shaded area.

The sensitivity study is carried out for a summer's period in the very dry and warm year 2003. The sensitivity simulations are started from the 9^{th} August 2003 and are integrated for 4 days. The situation is dominated by light winds ensuring more or less local meteorological situations, where the surface processes are most important for the local weather to ensure ideal conditions to test the performance of the parameterisation schemes. METRAS is run applying the flux aggregation scheme with blending height as well as the parameter averaging scheme. For both schemes resolutions of 16 km, 8 km, 4 km and 2 km are used to investigate the resolution-dependence of the model results obtained with the two schemes.

The dry and warm period from 9^{th} August 2003 until 13^{th} August 2003 was nearly cloudless and dominated by the anticyclone "Michaela" located over central Europe. The year 2003 was all in all a very dry and hot year in Europe, and the selected period is characterised by very high temperatures of about 30 °C and dew points between 2 °C and 16 °C. Low wind speeds from northerly directions allow local processes to influence the model solution. The relative humidity was highly varying between 30 % and 80 % with lower relative humidity in the northern parts of the model domain and higher relative humidity in the southern part of the model domain; no precipitation was measured. This period was chosen, because it was a very dry period and a good opportunity to investigate the performance of the surface flux schemes for an extreme situation, where the standard METRAS moisture approach is no longer applicable.

An intercomparison between the results obtained with the different model configurations is undertaken by investigating the Bowen ratio and the diurnal cycle of temperature, dew point and wind averaged over all urban grid points. Further, some statistics are derived to estimate the sensitivity of METRAS towards resolution and aggregation scheme. Additionally, model results are evaluated by calculating hit rates based on DWD routine observations. These fulfil the criteria for weather prediction stations. The hit rates are calculated by interpolating the model results horizontally and vertically towards the geographic location of the measurement stations. Temperature and dew point are measured at 2 m height above ground level and wind speed and direction are measured at 10 m height. All parameters were available with a temporal resolution of 1 hour. It has to be kept in mind that the spatial representativity of the city measurement stations might be poor (Oke, 2006a; Oke, 2006b) and the scale they are representative for differs from the simulated values at the 1^{st} model level. However, it is likely that the air in the city is well mixed from the flow around the buildings so that the measurements are representative for a larger area similar to the grid resolution of the model

(Bohnenstengel et al., 2011b). The urban measurements are more influenced by local effects due to the stronger heterogeneity than more rural stations. In addition, only grid averaged values are compared, not the values calculated for a sub-grid land-use class.

7.4 Results

The model results show that the adjustments towards higher roughness lengths z_0 for urban land-use types and the appropriate description of soil parameters like Albedo, thermal diffusivity, thermal conductivity, depth of water wave, soil water availability and saturation value for water content already lead to increased surface temperatures and a higher Bowen ratio (Figure 7.2) for the Berlin area in the model without taking into account an additional anthropogenic heat flux. One might argue, that the anthropogenic heat flux is already prescribed by altered surface properties, since the increased heat storage capacity reflects parts of the anthropogenic heat impact on the local urban climate. However, the main factor in the urban energy balance is the increased storage term that causes a phase delay of the sensible heat flux compared to rural surrounding especially around the evening transition and the altered net short wave radiation during the day. These are obviously affected by the adapted surface properties leading to a higher storage of energy during the day. The urban surface Albedo is lower reflecting the net effect of shadowing and trapping of the incoming solar within the street canyons.

7.4.1 The spatial variability of the Bowen ratio

Due to the adjusted surface parameters for urban areas the partitioning of the surface sensible and latent heat fluxes is altered and results in higher Bowen ratios for Berlin (Figure 7.2). Such an increase of the sensible heat flux at the expense of the latent heat flux is also found for other European urban areas and confirmed by studies using even more sophisticated urban parameterisations (Bohnenstengel et al., 2011b).

Grid averaged values of the Bowen ratio are compared for all 8 configurations for a snapshot in time at 1500 h local time. The intensity and the spatial extend of the urban Bowen ratio varies between all 8 configurations (Figure 7.2). In general, all simulations show an increased Bowen ratio for Berlin compared to the rural surrounding. The only exception is parameter averaging with 16 km resolution: the Bowen ratio differs only slightly between Berlin and the rural areas. The spatial extend and the intensity of high Bowen ratios differs considerably between the two parameterisation schemes. Figure 7.2 demonstrates that the flux aggregation method (left column) results in a higher urban Bowen ratio for Berlin than the parameter averaging method (right column). These differences between the two schemes are most pronounced for coarse resolutions of 16 km and 8 km (Figure 7.2a,b,e,f), where areas with a high Bowen ratio cover the whole of Berlin when applying flux aggregation, while parameter averaging shows hardly any urban signal in the Bowen ratio.

Both parameterisations demonstrate a resolution-dependence of the intensity of the Bowen ratio. But in general, the results obtained with coarse resolution resemble the high resolution results reasonably well. The resolution dependence in the flux aggregation scheme can be explained with the increasing vegetation fractions in some grid boxes when applying coarser resolutions. Parameter averaging calculates grid box averaged fluxes in contrast to flux aggregation, which calculates sub-grid surface fluxes for each land-use class. But, to ensure a fair comparison between the two schemes only the grid box averaged values are compared here. The fraction of the sub-grid scale land-use classes per grid box changes when applying different horizontal resolutions. While the sub-grid scale fluxes stay roughly the same when changing the resolution, this is not the case for the grid box averaged fluxes and the Bowen ratio. The increasing fraction of vegetated areas when using coarser resolution leads to a lower grid box averaged Bowen ratio compared to high resolutions when the entire grid box is merely covered by urban land-use. Figure 7.3 and 7.4 demonstrate the impact of increasing resolution on the fraction of urban areas in a grid box (Figure 7.3 and Figure 7.4). According to Figure 7.3 the relative number of grid boxes where the urban fraction exceeds 50 % increases with horizontal resolution. The horizontal 2D plots of urban fraction per grid box in Figure 7.4 demonstrate that Berlin is resolved more “explicitly” with increasing resolution. The shape and structure of the whole city is described more precisely.

Results obtained with parameter averaging demonstrate stronger resolution dependence than flux aggregation. Increasing the resolution pays back most for the parameter averaging method. However, for parameter averaging the strong resolution dependence of the Bowen ratio can not only be explained by the vegetation dominating the urban surface fluxes when decreasing the resolution. The Bowen ratio has values of up to 25 for high resolutions of 2 km and 4 km (Figure 7.2g,h). Parameter averaging calculates artificial grid box averaged land-use parameters and smoothes the influence of the urban land-use class. Simulations applying a coarser resolution of 8 km, therefore, suppress the spatial variability and do not capture the impact of the large sub-grid scale fluxes on the grid box averaged fluxes realistically. Therefore, parameter averaging only leads to more realistic results when resolving the land-use more explicitly. For instance, when comparing flux aggregation with 8 km (Figure 7.2b) and parameter averaging with 8 km (Figure 7.2f) against the results obtained with 2 km resolution (Figure 7.2d,h), the flux aggregation method resembles the shape of the increased bowen ratio received with higher resolutions better than parameter averaging. Parameter averaging results in much lower Bowen ratios similar to the rural areas and only captures higher values of up to 8 in small areas of Berlin.

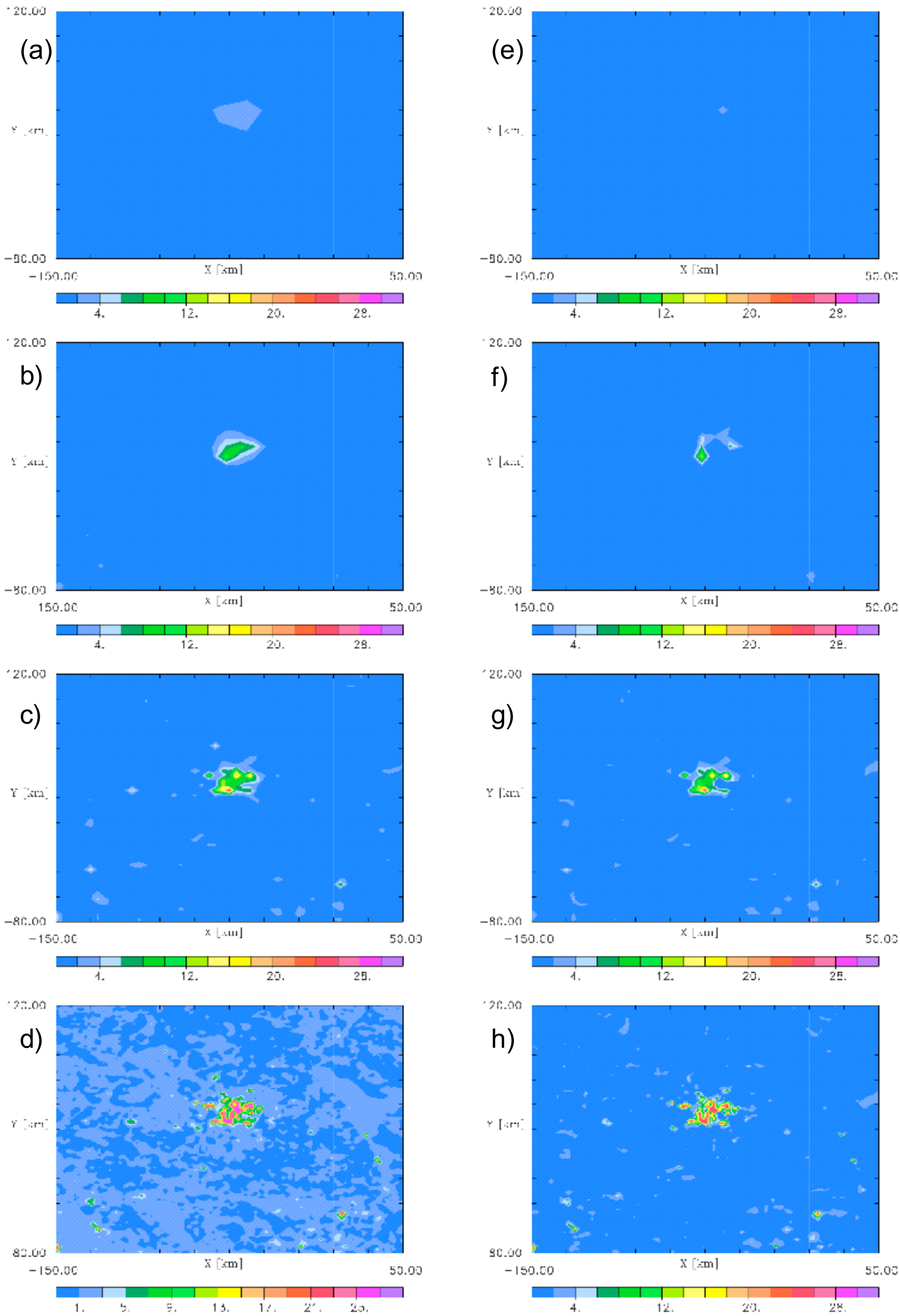


Figure 7.2: Bowen ratio for 1500 h local time of the first simulation day. Shown is a part of the whole model area simulated with (a) FL16, (b) FL8, (c) FL4, (d) FL2, (e) PA16, (f) PA8, (g) PA4 and (h) PA2.

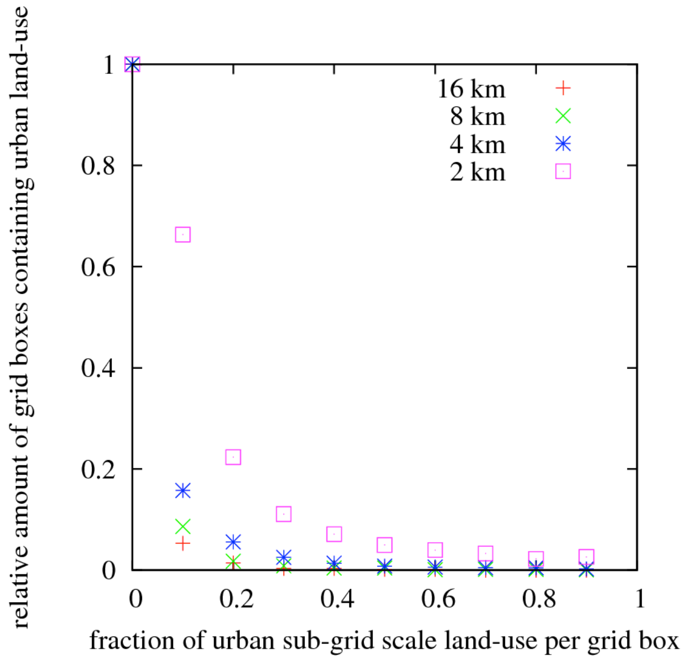


Figure 7.3: Fraction of grid boxes with sub-grid scale urban land-use in the grid box above a certain threshold value (x-axis) out of the total amount of grid boxes containing urban sub-grid scale land-use. This shows how increasing the resolution resolves urban sub-grid scale land-use more explicitly.

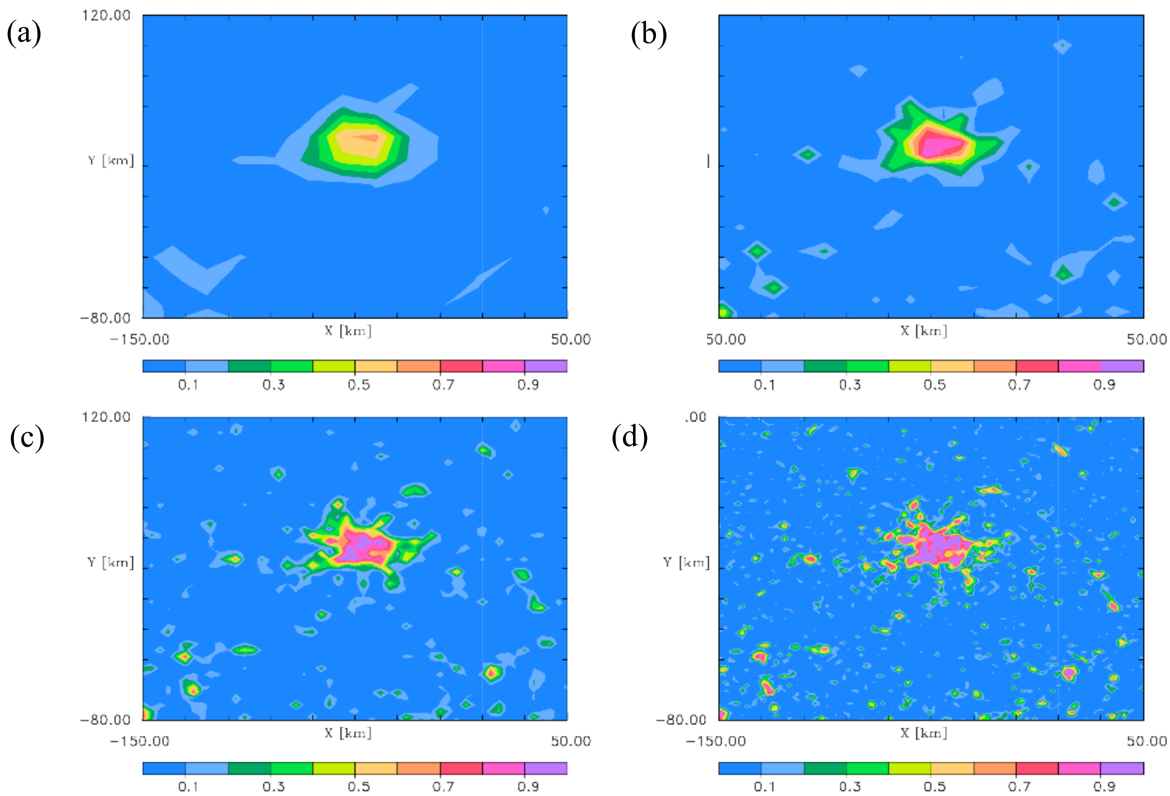


Figure 7.4: Fraction of urban areas within each grid box for (a) 16 km, (b) 8 km, (c) 4 km and (d) 2 km resolution.

7.4.2 The diurnal cycle of the Bowen ratio

The grid box averaged sensible and latent heat fluxes are area averaged for Berlin and the rural surrounding, respectively. The Berlin area is defined as the area in and around Berlin, where the urban sub-grid scale land-use is larger than 10% per grid box. Figure 7.4 shows the sub-grid scale urban land-use fractions in and around Berlin. The area-averaged fluxes are calculated from the grid box mean fluxes and not from just averaging the sub-grid scale fluxes for the urban sub-grid scale land-use tiles. Hence, the area-averaged fluxes also account for non-urban land-use within the Berlin area for instance. Then the Bowen ratios for the Berlin area and the rural area are calculated. The resulting diurnal cycles of the Bowen ratio for Berlin and the rural area are plotted for all configurations in Figures 7.5 and Figure 7.6. The diurnal cycles demonstrate that the turbulent surface fluxes and hence the whole surface energy balance are sensitive to the aggregation scheme and the resolution. The urban-rural difference in the surface fluxes is additionally shown by the ratio of Berlin's Bowen ratio BU and the rural surrounding's Bowen ratio BR based on spatial averages for every output time step resulting in BUR = BU/BR. Figure 7.7 and Figure 7.8 show the diurnal variation of BUR. The urban and rural Bowen ratios hardly differ when applying 16 km resolution. However, the Bowen ratio is slightly larger for the flux aggregation scheme than the parameter averaging scheme, especially in the afternoon (Figure 7.7a,b). For 8 km resolution flux aggregation simulates distinctly higher Bowen ratios for the urban area during the day and much lower values during the night, while parameter averaging simulates smaller differences between urban and rural Bowen ratios during day and night (Figure 7.5c,d). The influence of the urban surface energy balance shows up in the area averaged diurnal cycle when using the flux aggregation scheme. This is underlined by the ratio of the urban and rural Bowen ratio as represented by BUR in Figure 7.7. BUR indicates slightly higher values for flux aggregation than for parameter averaging when comparing 8 km resolution. The Bowen ratio for urban and rural areas becomes more distinct with the larger values for urban areas considered for higher resolution. Higher resolutions simulate larger differences in the partitioning of the turbulent surface fluxes between Berlin and the rural surrounding more faithfully. With increasing resolution the Bowen ratio for Berlin exceeds the one for the rural surrounding more and more (Figure 7.7, 7.8). For the 4 km and 2 km resolutions both schemes show an enhancement by a factor of about 3 and 5 during daytime, respectively. Differences between the parameterisation schemes vanish for the daytime maxima of the Bowen ratios when reaching 4 km and 2 km resolution but are maintained during the night. The sub-grid scale land-use is resolved more explicitly for higher resolutions and accounts more accurately for the heterogeneities on the sub-grid scale. The BUR values (Figure 7.8) show that both parameterisation schemes capture the daytime differences between the Bowen ratios for urban and rural areas. The flux aggregation and parameter averaging schemes converge to similar values for high resolutions. However, night-time values differ considerably. Large differences in BUR during night-time often result from dividing sensible and latent heat fluxes being nearly zero but with an order of magnitude difference. Apart from this the BUR differences also have a physical explanation. For instance, Figure 7.6c indicates that the urban Bowen ratio is much larger than the rural one during daytime. The reason is that the sensible heat flux in the urban area and rural area is very similar during daytime, but it is mainly the latent heat flux that causes the daytime difference. The latent heat flux in the urban area is supressed by the urban surfaces resulting in a higher urban Bowen ratio. Towards the evening transition the urban sensible heat flux becomes more important. The urban sensible heat flux is delayed compared to the rural one due to the larger thermal inertia in the urban environment and stays positive longer. This phase shift causes a larger positive urban Bowen ratio compared to the rural one. While the urban Bowen ratio has values of around 4 at 18 h, the rural one is already slightly negative in Figure 7.6c. This phase shift in the sensible heat flux leading to a larger urban Bowen ratio than compared to the rural one is also reflected in the increasing BUR values in Figure 7.8c during the afternoon and towards 18 h. The large negative values from 18 h onwards are caused by the urban Bowen ratio being still positive because of the positive sensible heat flux, while at the same time the rural Bowen ratio is already negative due to the negative sensible heat flux. The rural area has a lower thermal inertia and becomes stably stratified earlier than the urban area. Later into the evening and night BUR values become positive again (Figure 7.8c), since the urban Bowen ratio also becomes negative due to a change in the sign of the sensible heat flux. However, the urban Bowen ratio becomes in the early night far more negative than the rural Bowen ratio (Figure 7.6). The urban sensible heat flux converges towards the rural one, but is slightly less negative than the rural one. At the same time the latent heat flux, which is positive with values of about 1-3 Wm^{-2} in the rural area is smaller in the urban area due to the sealing of the surfaces. This then causes a larger negative Bowen ratio in the urban area compared to the rural one.

Independent of the parameterisation scheme Bowen ratio differences increase for higher resolutions, since the dryer urban area is resolved much more explicitly. Differences between the parameterisation schemes occur, since averaging the surface parameters and then applying the flux function does not provide a physical background as the fraction weighted flux aggregation scheme for averaging the sub-grid scale surface fluxes does. This leads to a completely different urban energy balance in case of parameter averaging, especially for the coarser resolutions. The parameter averaging scheme calculates an unrealistic artificial but homogeneous land-use class for each grid box. When applying the flux function to this value non-linear contributions of the sub-grid scale fluxes to the grid box averaged flux, as it would have been observed in nature are not captured. The more heterogeneous the sub-grid scale land-use variations in roughness or e.g. soil humidity are, the less representative is the grid box averaged flux.

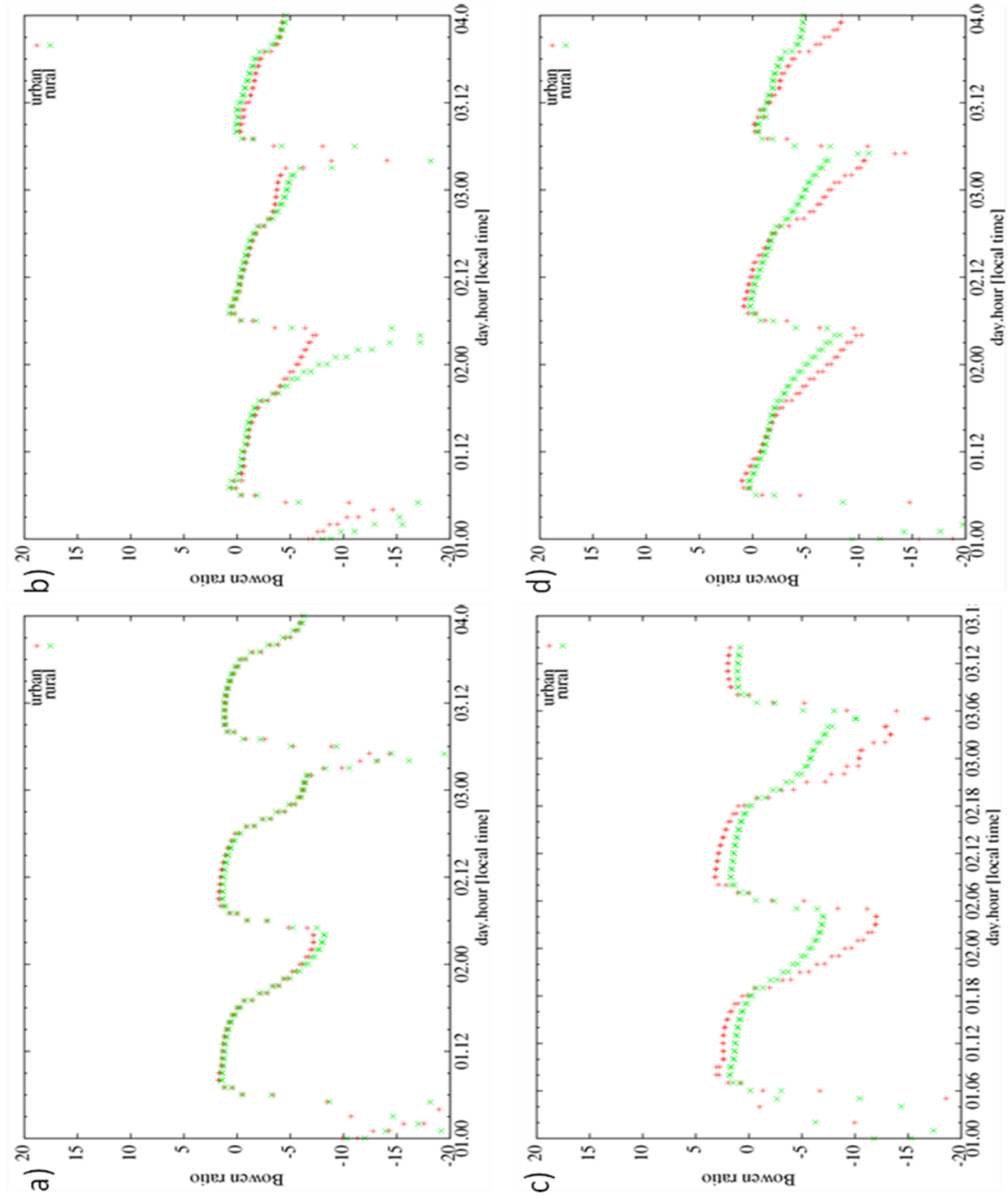


Figure 7.5: Area averaged Bowen ratio for Berlin (red crosses) and rural surrounding (green crosses) for (a) flux aggregation with 16 km resolution, (b) parameter averaging with 16 km, (c) flux aggregation with 8 km and (d) parameter averaging with 8 km resolution.

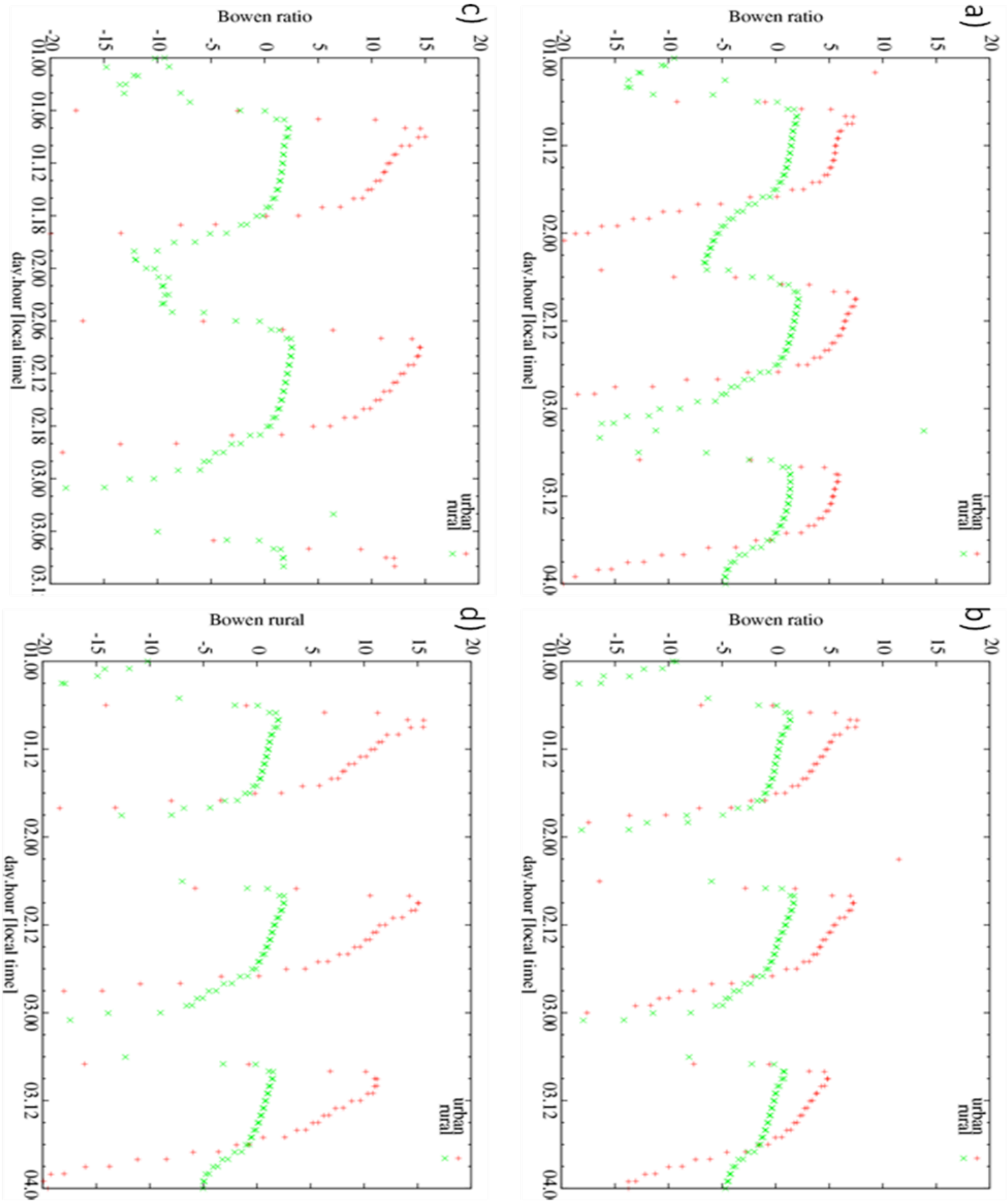


Figure 7.6: As Figure 7.5 but for (a) flux aggregation with 4 km, (b) parameter averaging with 4 km, (c) flux aggregation with 2 km and (d) parameter averaging with 2 km resolution.

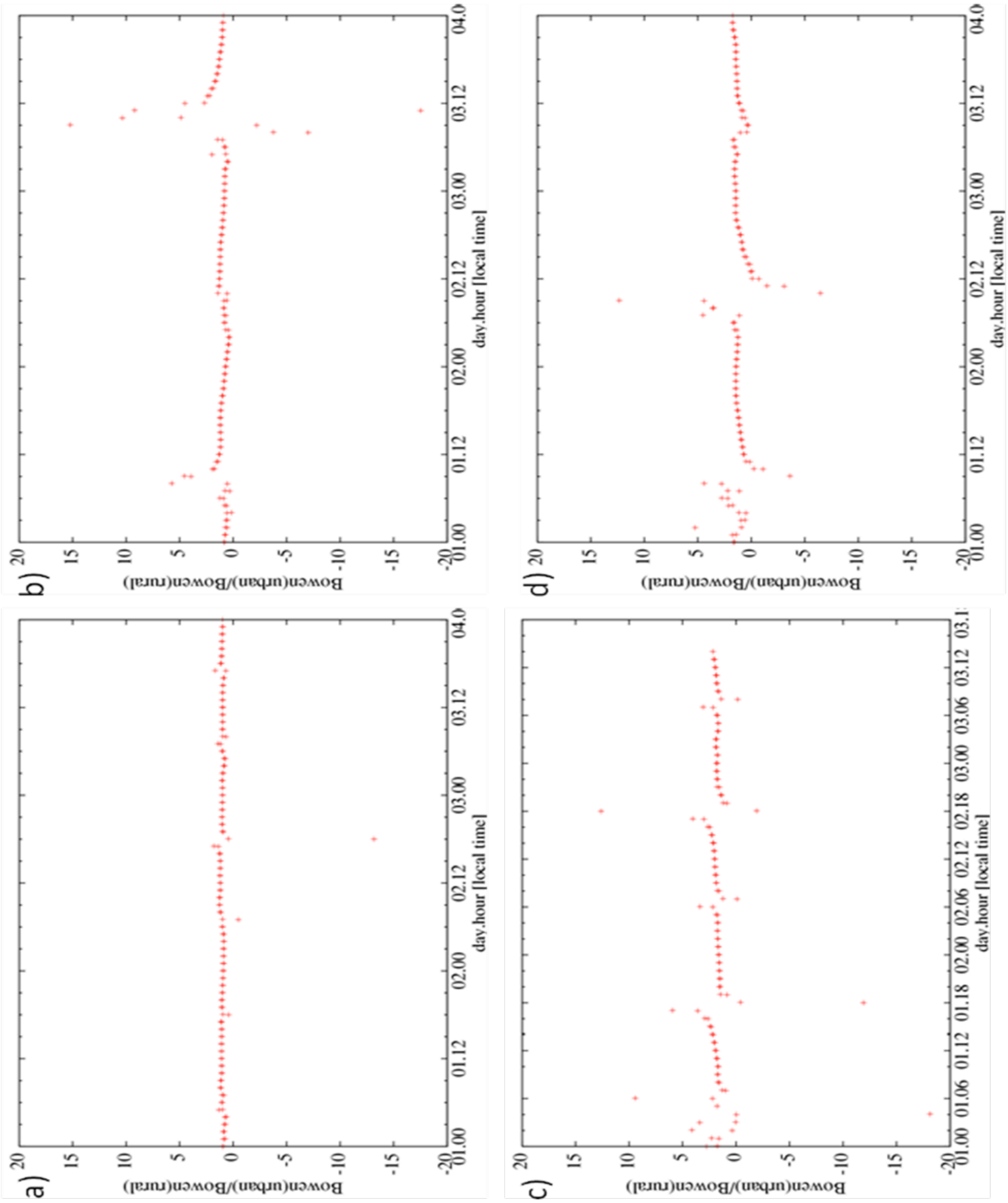


Figure 7.7: Diurnal cycle of the spatially averaged BUR value relating the Berlin Bowen ratio with the rural Bowen ratio for (a) flux aggregation with 16 km resolution, (b) parameter averaging with 16 km, (c) flux aggregation with 8 km and (d) and parameter averaging with 8 km resolution.

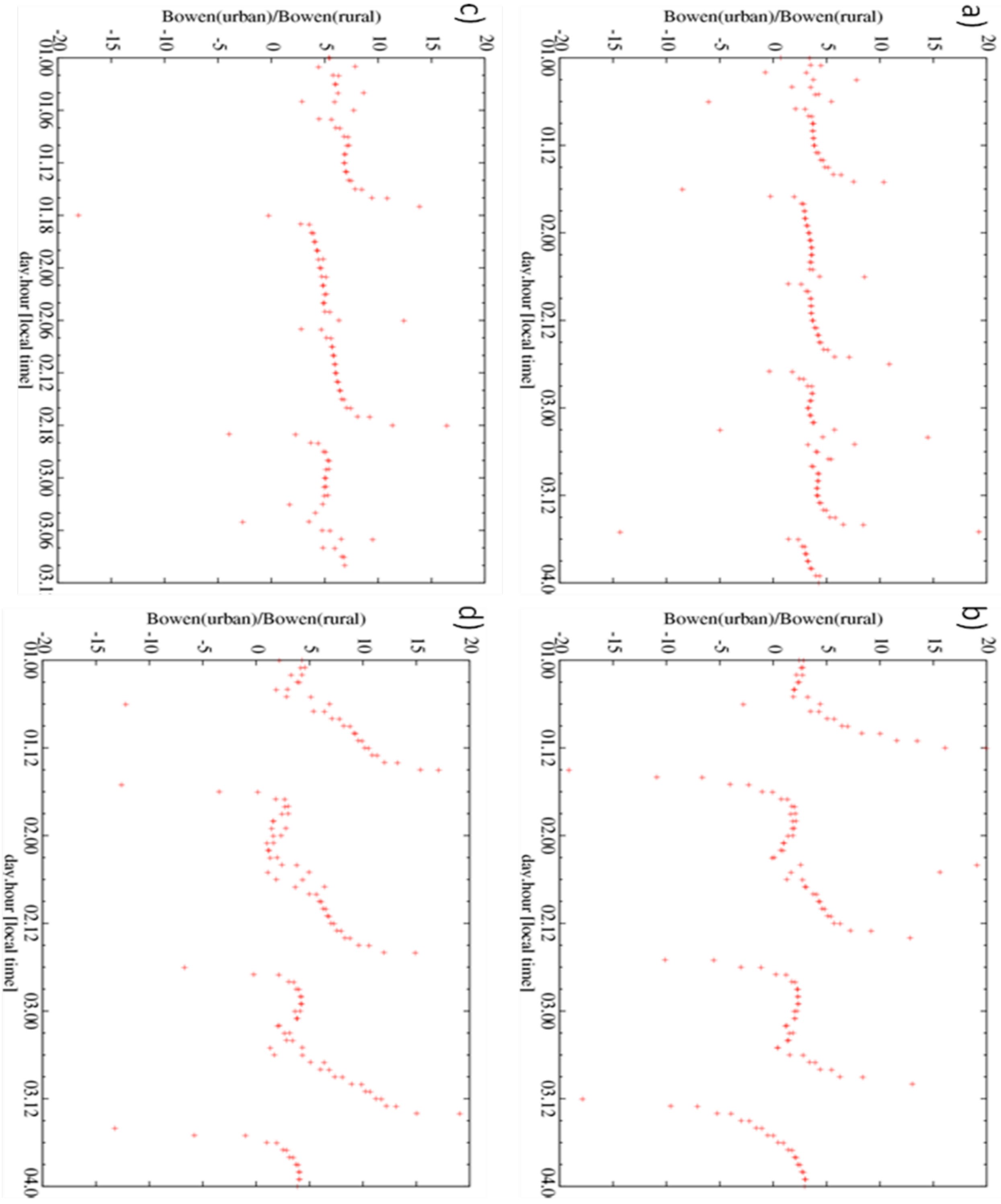


Figure 7.8: As Figure 7.7, but (a) flux aggregation with 4 km, (b) parameter averaging with 4 km, (c) flux aggregation with 2 km and (d) parameter averaging with 2 km resolution.

7.4.3 Diurnal cycle of temperature and depression of the dew point for Berlin

Figure 7.9 and Figure 7.10 compare the diurnal cycle of the depression of the dew point at 2 m height for rural and urban areas as simulated by METRAS with the measurements. The simulated temperatures are interpolated into the locations of the DWD measurements at 2 m height using Monin-Obhukov similarity theory. The amplitude of the diurnal cycle in the urban area (dashed line) is damped compared to the rural diurnal cycle (straight line). The rural diurnal cycle of the depression of the dew point reaches values around 18 K during day time and 5.5 K during night-time. In contrast, the urban diurnal cycle reaches slightly lower day time values around 17 K but higher night-time values around 8 K. The urban and rural differences are largest during the night according to the measurements (DWD).

Figure 7.9 and Figure 7.10 highlight the dependency of both parameterisation schemes on the horizontal resolution and reflect the values of the measured data. The flux aggregation scheme is nearly resolution independent during the day and reflects the values of the measured data. During the night differences between the 4 resolutions reach 4 K. Compared to the measurements the flux aggregation scheme overestimates the dryness in the rural surrounding around 22 h, underestimates it at the same time for the urban areas and underestimates the dryness everywhere in the early morning hours. One reason for this might be that the parameters for the soil characteristics are based on assumptions for coarser model resolutions like 8 km. For such resolutions it is assumed that bushes and gardens still cover a certain fraction within areas classified as urban. They are therefore taken into account when describing the typical urban parameters. Thus, it is assumed that an urban area is not purely urban and made of e.g. concrete but contains vegetation to a certain extend. With increasing resolution, the vegetated fraction per grid cell decreases in urban areas. The sub-grid scale land-use is resolved more explicitly. In this case urban areas are more purely described by urban building materials. However, the chosen surface parameters are independent of the resolution in this study. Bushes and evaporation are therefore likely to be slightly overestimated in highly resolved simulations of the urban areas.

The parameter averaging scheme shows a strong dependence of model results on the model resolution, in contrast to the flux aggregation scheme with blending height concept. The values are relatively far away from measured data. The simulations with coarser model resolution do not capture the whole range of “dryness”-values in the diurnal cycle in the depression of the dew point at all. The 16 km run clearly underestimates the dryness for Berlin over the whole diurnal range. However, the solution converges towards the solution of the flux aggregation scheme with increasing resolution, but does never capture the maximum difference around high noon. When applying 2 km horizontal resolution it shows a solution quite similar to the flux aggregation scheme, however, it underestimates the dryness. Parameter averaging performs much better for higher resolutions when the heterogeneity is more explicitly resolved. Then the parameterisation scheme does no longer average over too different sub-grid scale land-use classes when applying higher horizontal resolution. It no longer results in a too smooth artificial homogeneous land-use, which is not capable to represent strong surface heterogeneities. For higher resolutions the parameter averaging scheme simulates the net effect of strong differences in sub-grid scale surface fluxes on the grid box averaged flux more realistically and similar to the flux aggregation scheme.

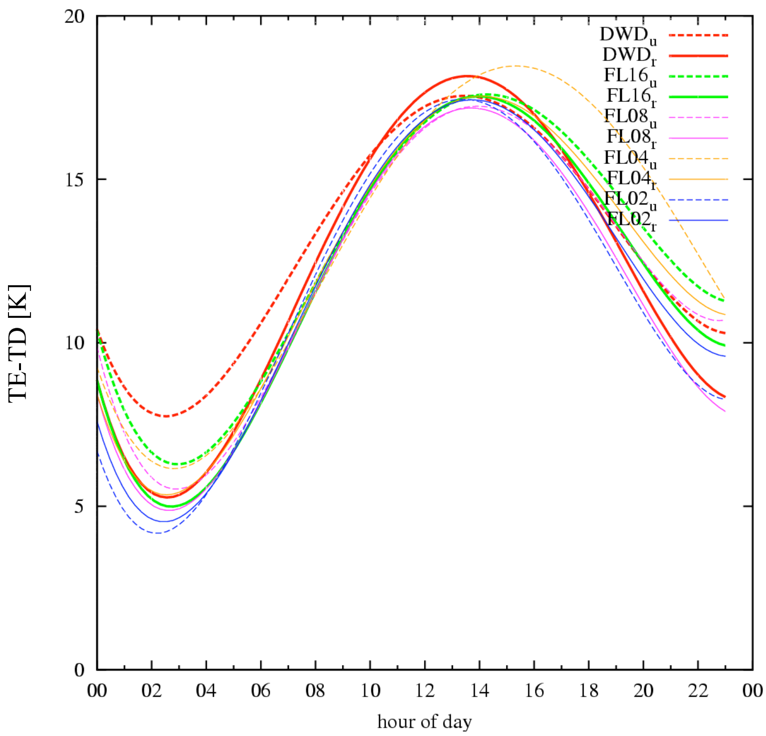


Figure 7.9: Averaged difference of temperature and dew point temperature for urban and rural grid points at the DWD locations within the model domain as simulated by the flux aggregation scheme with blending height concept for different resolutions. Measurements are indicated in red. Urban areas are indicted by dashed lines, solid lines show rural points. Results were averaged over 4 days.

All simulations result in smaller differences between urban and rural areas than the measurements indicate. The measurements indicate about 1 K difference in the depression of the dew point between rural and urban areas during the day with a slightly wetter urban area than rural. During the night the difference is larger with 3 K caused by the urban dryness island. The simulated urban-rural difference of dew point depression is about 2 K at night and nearly not distinguishable during day. Hence, the simulations are too humid during the night for Berlin. Only the flux aggregation scheme with 2 km resolution shows higher values for the depression in the dew point for the rural surrounding than for the Berlin area.

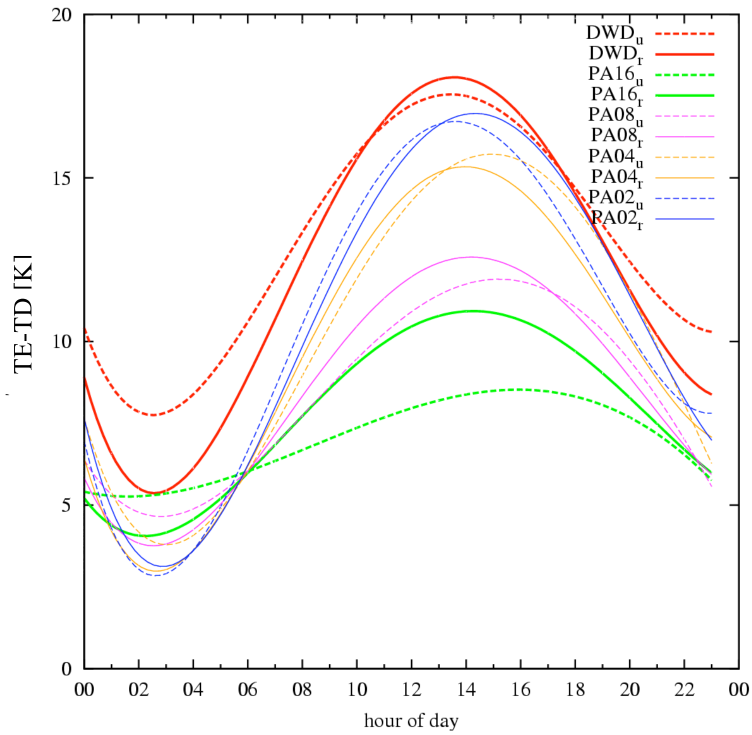


Figure 7.10: Same like Figure 7.9 but for parameter averaging.

7.4.4 Model accuracy and sensitivity towards horizontal resolution

Hit rates are calculated for temperature, dew point, wind speed and wind direction as well as pressure and specific humidity to evaluate the model results with a model independent measure (Schlünzen and Katzfey, 2003; Trukenmüller et al. 2004). Additional statistical measures like correlation coefficient, slope, root mean square error and mean absolute error are calculated based only on the model results to determine the dependency resulting from horizontal resolution and the averaging scheme.

The hit rates are derived from routine measurements taken by the “Deutscher Wetterdienst” (DWD). 27 DWD stations with hourly measurements are located in the model domain. For the Berlin area only 4 stations with hourly measurements were available for the evaluation. The model results of air temperature, dew point, wind speed and wind direction are interpolated to the locations of the DWD stations at 2 m height for temperature and humidity and 10 m height for wind above the surface applying Monin-Obhukov similarity theory. The hit rates are then calculated following eq. (7.14) with a desired accuracy A of ±2 K for the air temperature and dew point, for wind speed A is ±2 ms^{-1} and for wind direction ±30 °, for pressure ±1.7 hPa and for specific humidity it is ±3 gkg^{-1}.

$$H = \frac{100}{m} \sum_{i=1}^{m} n_i, \quad \text{with } n_i = \begin{cases} 1 \text{ for } |\text{difference}(\text{measurement,model})| < A \\ 0 \text{ for } |\text{difference}(\text{measurement,model})| \geq A \end{cases} \tag{7.14}$$

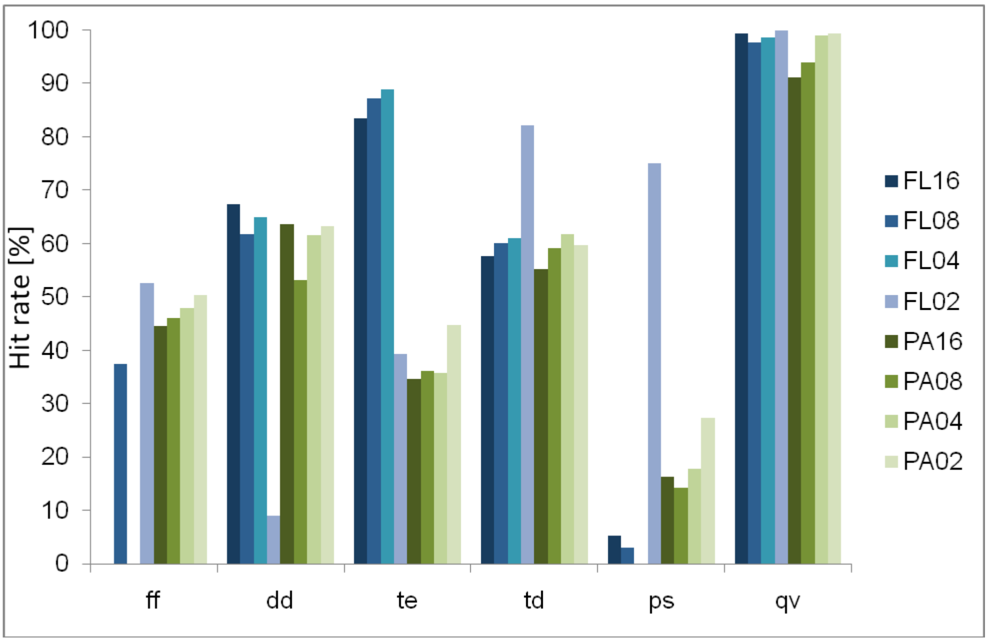
The hit rates are summarized for the whole model domain (Figure 7.11a) and for stations in and around Berlin only (station numbers 10381, 10382, 10384, 10389) (Figure 7.11b). For both domains a large variability in terms of forecast quality between the evaluated variables is visible. In general, the flux aggregation scheme performs similar or better than the parameter averaging scheme for all resolutions and variables except for wind speed in case of flux aggregation with resolutions 16 km (FL16), 8 km (FL8) and 4 km (FL4) for the whole model domain. For the urban areas the wind speed is not simulated well at all, but worst for 8 km flux aggregation (FL8) and 2km flux aggregation (FL2). The largest difference between flux aggregation and parameter averaging occurs for the thermodynamic values. The hit rates for the temperature in case of flux aggregation are nearly doubled compared to parameter averaging. This is mainly caused by the stronger coupling between the surface fluxes and the thermodynamic properties. This behaviour is expected from the dependency of the diurnal cycle of the depression of the dew point temperature from the model resolution for parameter averaging (Section 7.4.3). For dew point and specific humidity the differences in hit rates between the two schemes are smaller than for temperature. The dew point is simulated with 2 km flux aggregation, but the hit rate for temperature is much lower than for the other resolutions applying flux aggregation.

Figure 7.12 shows the difference in hit rates between all stations in the model domain (Figure 7.11a) and the urban (Figure 7.11b) stations only. The black horizontal lines in Figure 7.12 indicate the 5 % uncertainty range due to the interpolation of the simulated variables to the DWD locations. For wind speed the hit rates perform consistently better when comparing the stations of the whole domain than the urban stations only. For temperature all high resolution configuration with 4 km and 2 km perform better when comparing all stations than only comparing urban areas. However, it is not clearly detectable if METRAS performs better when comparing all stations in the domain, which are mostly rural stations, or when only comparing urban stations against measure ments. No clear trend is visible for any variable. With only 4 urban stations available to calculate hit rates for the urban area the number of data points for the urban area is probably too limited to form a robust data set. Therefore, these conclusions are not definite. Further simulations are needed to test the performance for a more robust urban data set.

Figure 7.13(a) and Figure 7.13(b) compare the hit rates of each configuration with the theoretically worst configuration with parameter averaging and 16 km resolution (PA16), respectively. Here, a clear trend is visible. The performance for all parameters (except wind) improves a lot from using flux aggregation or from increasing the resolution in case of parameter averaging. This is most evident for the whole model domain (Figure 7.13a) and especially for the thermodynamic values. The same trend is also visible for the Berlin area, although the hit rates for this area are not as robust as for the whole model domain, since they are based on less data points. Here more sensitivity studies are needed to derive reliable results.

The diurnal cycles of temperature, pressure, wind speed and wind direction averaged over the measurement sites of the whole model domain and over the whole simulation period are summarized in Figure 7.14. Additionally, the forcing values used from the ECMWF analysis and the DWD measurements are included. For temperature the forcing overestimates the measured temperatures especially at night-time. Still METRAS captures the diurnal cycle realistically and especially the flux aggregation scheme performs well and nearly independent of the resolution. The flux aggregation only slightly overestimates the night-time temperatures by 2 K for the coarser 8 km and 16 km resolutions. In contrast the parameter averaging scheme gains much from applying higher resolutions (Figure 7.14), since the sub-grid scale land-use is resolved more explicitly. The parameter averaging scheme behaves similar to the flux aggregation scheme for the night-time temperatures but shows large differences during the day. For coarse resolutions parameter averaging underestimates screen temperatures by up to 6 K. Both schemes show a phase shift compared to the measurements despite correcting for the time shift between local time and UTC. They peak about 1 hour earlier than the observations indicate although the ECMWF forcing peaks much later than the measurements. The phase shift of the parameter averaging scheme is resolution dependent. It is reduced for parameter averaging with increasing resolution.

(a)



(b)

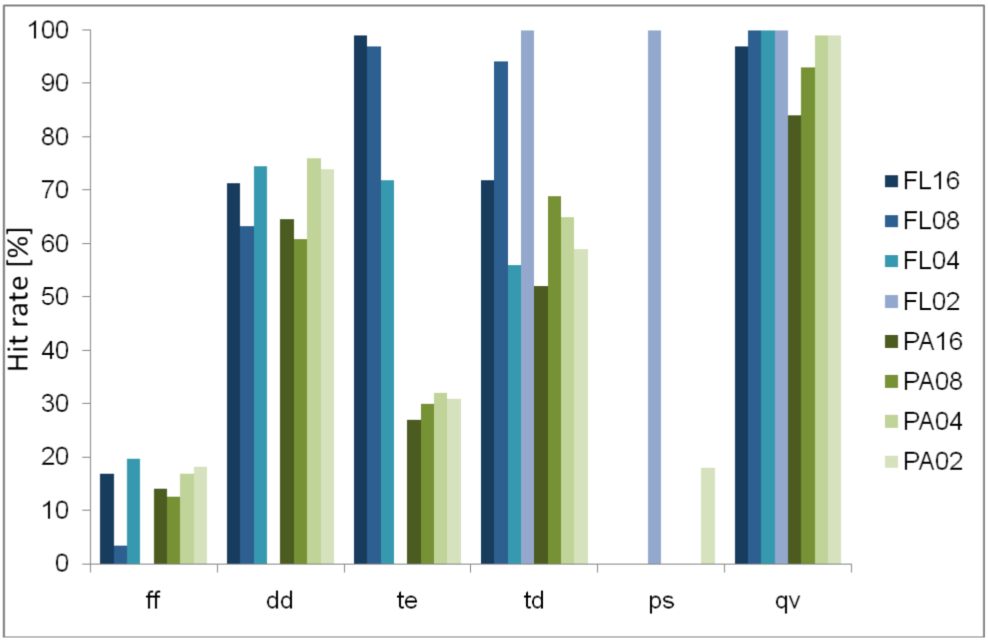


Figure 7.11: Hit rates for (a) whole model domain and (b) Berlin area. Blue bars indicate flux aggregation scheme, green colours indicate parameter averaging scheme and numbers stand for the resolution applied. The hit rates are displayed for wind speed (ff), wind direction (dd), temperature (te), dew point temperature (td), pressure (ps) and specific humidity (qv), separately.

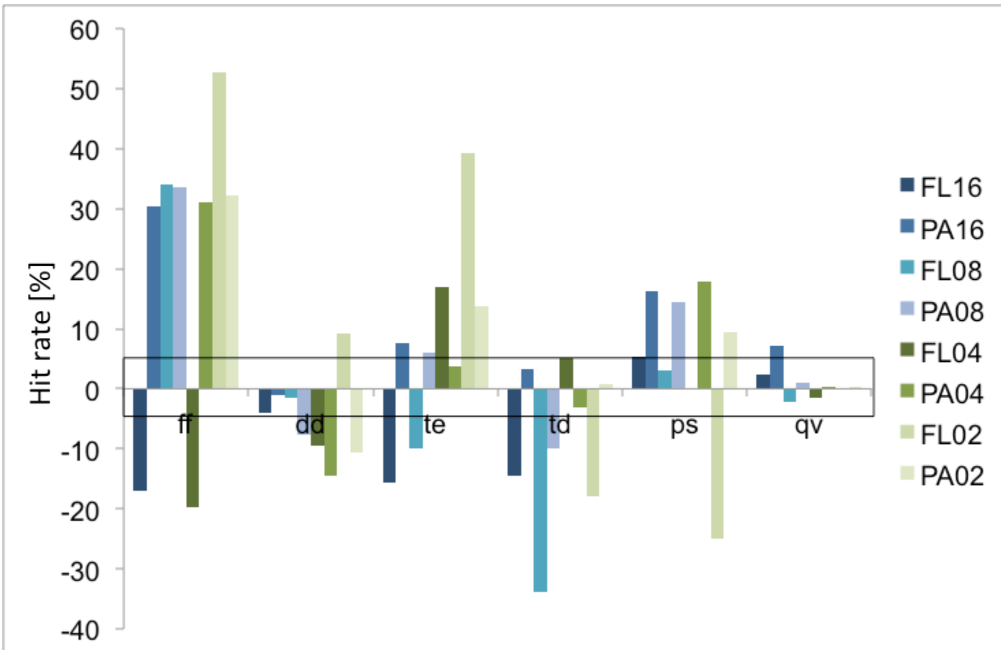
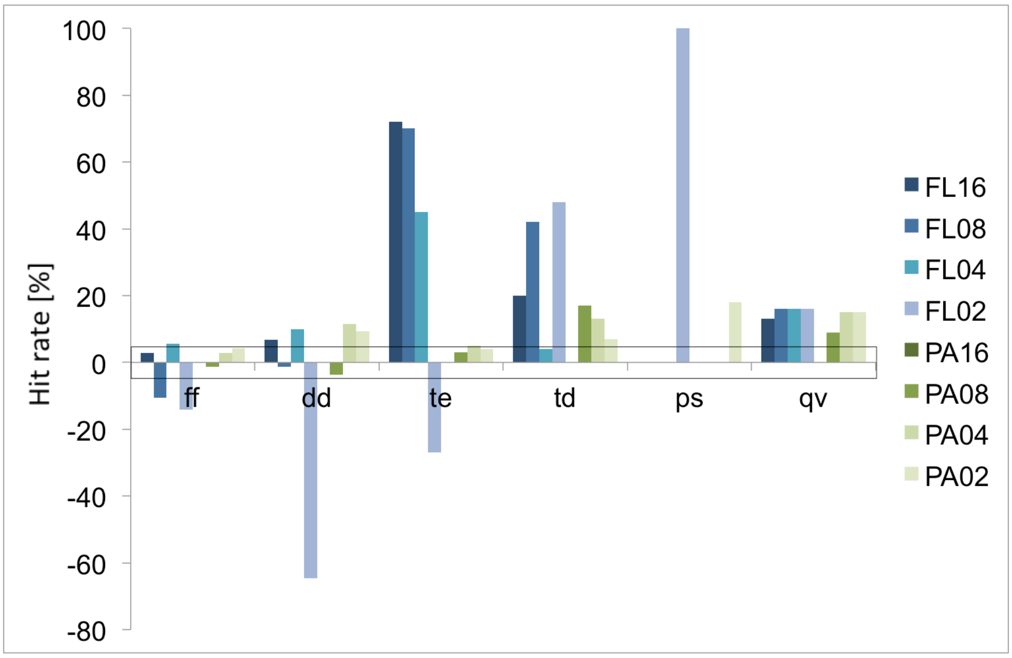


Figure 7.12: Absolute differences in hit rates between whole model area and Berlin area. Positive bars indicate that hit rates of the whole model area are better than those of the Berlin area. Negative bars indicate that hit rates for Berlin are better than hit rates for the whole model area. The black rectangle indicates the inaccuracy region of 5 %.

Differences between the two schemes are more pronounced during daytime. This might be caused by the very locally driven meteorological situation. The simulations with coarse resolution and parameter averaging underestimate the turbulent transport from the surface into the boundary layer. This worsens for coarser resolutions and the more heterogeneous the sub-grid scale land-use is. The parameter averaging method does not account for the local influences from the surface when calculating the surface fluxes based on an artificial land-use. In contrast, the flux aggregation method calculates the surface energy balance for each sub-grid scale land-use class separately. Consequently, the flux aggregation scheme accounts better for local effects. The flux aggregation method does result in very similar differences between the averaged temperatures for the four resolutions for day and night-times. Since the boundary layer is much deeper and well mixed during daytime larger differences in the sensible heat flux are needed to result into the same temperature differences during the day as during the night. The night-time minimum temperatures between midnight and 6 am in the morning show an offset between the coarse and high resolution simulations independent of the averaging scheme applied. This difference likely occurs since the boundary layer is much shallower during night-time and hence the temperature in the boundary layer has to change more in order to accommodate differences in the surface fluxes as a consequence of changes in the resolution and hence the sub-grid scale land-use.

(a)



(b)

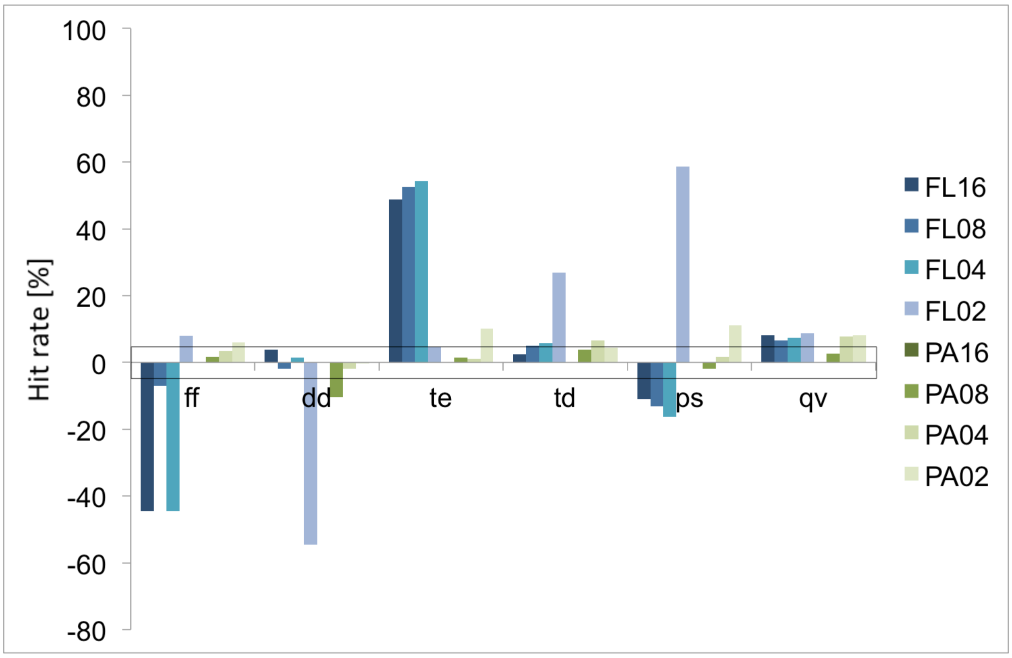


Figure 7.13: Absolute differences of hit rates compared to the theoretically worst configuration with parameter averaging and 16 km resolution. Positive values indicate an improvement of the model performance; negative values indicate a worsening of the model performance. Differences are shown for the whole model area (a) and the Berlin area (b).

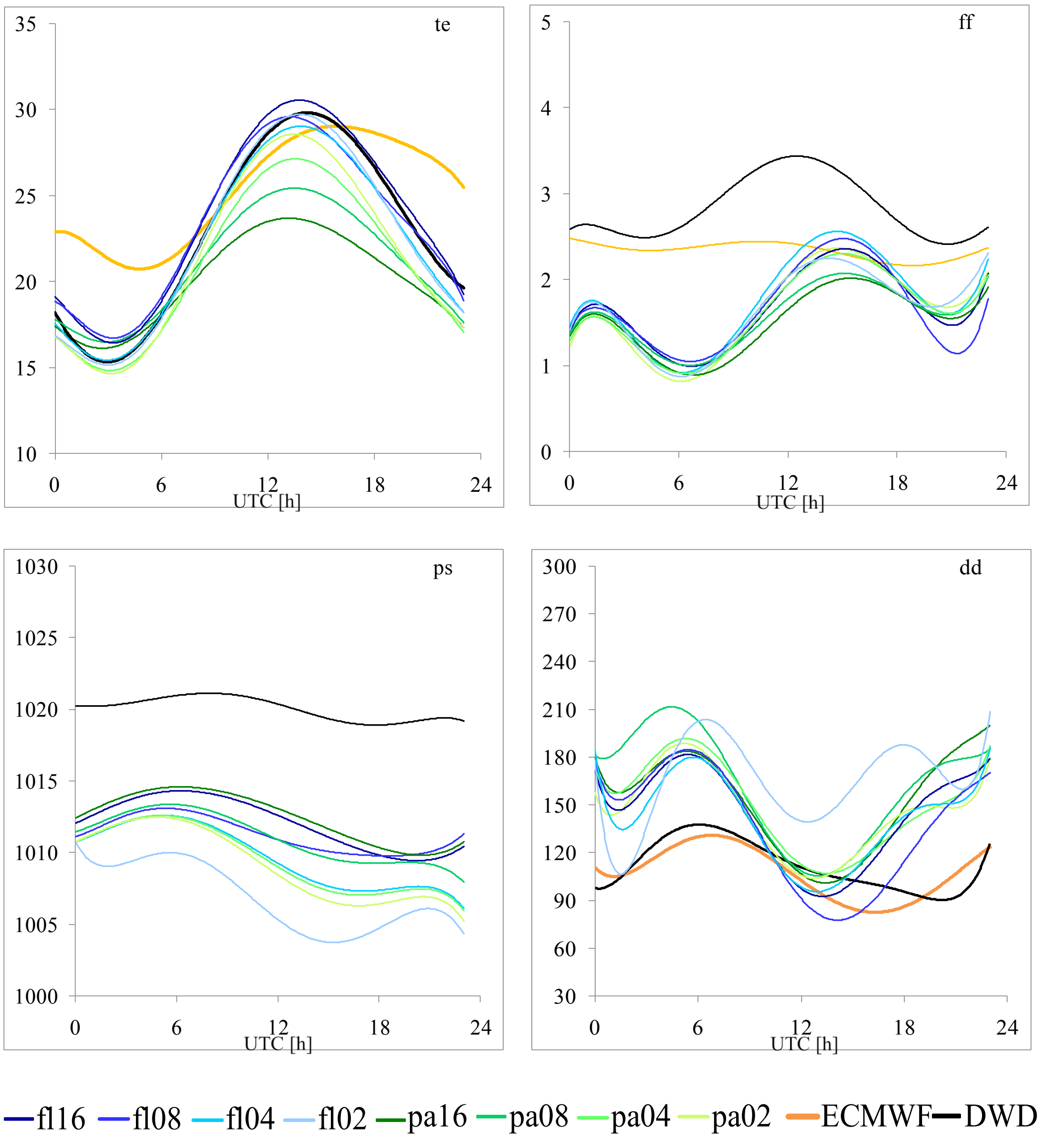


Figure 7.14: Temperature (te), pressure (ps), wind speed (ff) and wind direction (dd) averaged over all measurement sites with time in UTC. Black lines show diurnal cycle from measurements and orange lines show the ECMWF forcing data. Note that pressure was not forced in METRAS but initialised at a single point from ECMWF data.

Wind speed and wind direction are simulated with hit rates between 40 % and 50 % for the whole model domain (Figure 7.11a), however FL8 shows 0 %. For Berlin the simulation of the wind speed is even worse but the wind direction is captured much better with hit rates of up to 75 % (Figure 7.11b). According to Figure 7.14 the DWD measured low wind speeds between 2.7 and 3.4 ms^{-1} on average with larger wind speeds around noon. The ECMWF forcing in contrast captures the night-time values well but underestimates the daytime values and has nearly no visible diurnal cycle (Figure 7.14). METRAS underestimates the measurements over the whole simulation period by about 0.7 ms^{-1} at night and 1.5 ms^{-1} in the morning. However, it captures the shape of the diurnal cycle with increasing wind speeds in the afternoon well peaking about an hour later than the DWD. On average, the parameter averaging scheme tends to simulate lower wind speeds than the flux aggregation scheme and is more dependent on the resolution. The drop in hit rates for the wind speed in the Berlin area can be due to the underestimation of the urban roughness lengths (Figure 7.13b). The increase in wind speed with increasing resolution for parameter averaging underlines this assumption. The better resolved sub-grid scale roughness is resolved more realistically. This leads to more accurate local roughness lengths and an increased friction velocity enhancing the mixing.

While the DWD and the ECMWF show wind directions between 90° and 140° (Figure 7.14), METRAS simulates too southerly winds in the morning. During noon it captures the wind direction reasonably well and tends to more southerly direction in the evening and during night-time hours. All METRAS simulations show a much stronger diurnal cycle than the DWD and ECMWF with the 2 km flux aggregation scheme showing the largest deviations from the measurements.

Pressure was not forced with ECMWF analysis, but only one value was taken from the ECMWF to initialise the lowest point in the METRAS model domain. This method might lead to a pressure offset but does not necessarily worsen the model results in general. Differences between the pressure values can arise, because of height differences in the orography of different resolution. While the DWD measurements show a nearly constant pressure of 1020 hPa averaged over all sites, METRAS clearly underestimates the pressure. The differences between the various configurations are quite large with up to 8 hPa. The linear interpolation of the model results in the horizontal and the hydrostatic assumption for the vertical interpolation towards the location of the DWD measurements result in some uncertainties of the interpolated model results, which is most pronounced in the pressure fields.

Increasing the resolution for each of both parameterisation methods mostly pays back for the less well performing parameter averaging method, since sub-grid scale surface characteristics are resolved more explicitly by higher resolutions. This leads to more realistic averaged surface characteristics. In contrast, for the flux aggregation method with blending height concept differences in hit rates due to changes in resolution are mostly within 5 %, which can be considered as negligible. Exceptions are found for some of the urban measurement sites.

Overall, the thermodynamic values are more sensitive to the parameterisation and resolution applied than the dynamic values in rural as well as in urban areas. They are directly affected by the sensible and latent heat fluxes. The dynamic values show a smaller dependence on resolution for rural and urban areas for parameter averaging than for the flux aggregation approach. When comparing the hit rates between the rural areas and Berlin for all parameters it is found that METRAS performs better for rural than for urban areas, although METRAS captures some of the urban phenomenon's associated with an urban heat island like the enhanced dryness for example.

Table 7.1 summarizes some statistics for the spatially averaged sensible heat flux over the Berlin area. The statistics are calculated between all model configurations to gain information about the sensitivity of METRAS to the parameterisation scheme and the horizontal resolution and hence their impact on the grid box averaged sensible heat flux. Therefore, the correlation coefficient (eq. 7.15), slope of the regression s_{xy} (eq.7.16), mean absolute error MAE (eq.7.17) and the root mean square error RMSE (eq. 7.18) are calculated in the following for each combination of model configurations.

$$r_{xy} = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} \tag{7.15}$$

$$s_{xy} = \frac{\sigma_x}{\sigma_y} r_{xy} \tag{7.16}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - y_i\right| \tag{7.17}$$

$$\mathrm{RMSE} = \sqrt{\mathrm{mse}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \tag{7.18}$$

In general the correlation coefficients show a very close agreement between all configurations with correlation coefficients between 0.93 and 1 indicating a perfect correlation. Slightly lower correlations (0.93) are found between FL4 and FL8, PA4 and FL8 (0.96) as well as FL4 and FL8 (0.95). In general the high correlation coefficients underline that the different model results are somewhat linearly related. The corresponding slope gives some more information about the direction of the differences. Following Table 7.1 the configurations applying parameter averaging result in larger slopes when comparing with all other configurations. In contrast, the slope between configurations with flux aggregation and different resolution varies only between 0.89 and 0.95. This underlines that configurations with parameter averaging and flux aggregation differ much in terms of the sensible heat flux. Increasing the resolution for parameter averaging shows a convergence towards the results gained with flux aggregation, while flux aggregation is less sensitive towards the resolution applied.

The RMSE will always be equal or larger than the MAE. Together they indicate the variance in the deviations in the two compared data sets. The larger the difference between RMSE and MAE the larger is the variance in the individual deviations of the two data sets that are compared. The MAE is less sensitive towards large values than the RMSE and indicates the averaged difference between two configurations. If the difference between RMSE and MAE is small it is unlikely that large deviations have occurred between two configurations, but usually some variation in the deviation exists. In Table 7.1 the MAE shows largest differences between those configurations with coarse resolutions, especially when comparing with PA16 (parameter averaging with 16 km resolution) and PA8 (parameter averaging with 8 km resolution). The RMSE is about 50 % larger compared to the MAE for the coarser resolutions indicating some very large difference to be likely, while they are least likely for PA2. In general, differences are larger when comparing PA coarse resolution cases (30 to 40). The differences become smaller when comparing FL cases with different resolutions (below 20). This underlines the findings from Chapter 6 that the resolution has a minor impact on the flux aggregation than on the parameter averaging scheme.

corr	FL16	PA16	FL8	PA8	FL4	PA4	FL2	PA2
FL16	1	0.98	0.99	0.98	0.98	0.99	0.97	0.97
PA16		1	0.98	0.97	0.95	0.97	0.96	0.97
FL8			1	0.96	0.93	0.95	0.98	0.98
PA8				1	0.96	0.96	0.96	0.99
FL4					1	0.99	1	1
PA4						1	0.99	1
FL2							1	1
PA2								1

MAE	FL16	PA16	FL8	PA8	FL4	PA4	FL2	PA2
FL16	0	26.04	4.67	20.34	9.26	7.28	9.54	2.36
PA16		0	25.99	7.61	25.8	24.68	21.73	1.04
FL8			0	19.67	9.94	9.04	6.2	1.93
PA8				0	20.31	18.9	16.78	0.75
FL4					0	4.12	3.35	0.42
PA4						0	4.04	0.46
FL2							0	0.32
PA2								0

Slope	FL16	PA16	FL8	PA8	FL4	PA4	FL2	PA2
FL16	1	0.38	0.95	0.52	0.92	0.92	0.89	0.64
PA16		1	2.5	1.35	2.36	2.39	2.31	0.98
FL8			1	0.54	0.94	0.95	0.94	0.68
PA8				1	1.71	1.71	1.69	1.01
FL4					1	0.99	0.99	0.93
PA4						1	0.99	0.96
FL2							1	1.04
PA2								1

RMSE	FL16	PA16	FL8	PA8	FL4	PA4	FL2	PA2
FL16	0	43.36	7.14	33.74	13.36	10.53	14.49	3.04
PA16		0	43.29	12.16	42.55	41.05	38.81	1.39
FL8			0	31.64	20.44	18.25	10.42	2.5
PA8				0	33.3	31.98	30.67	0.91
FL4					0	6.28	5.24	0.6
PA4						0	7.13	0.54
FL2							0	0.38
PA2								0

Table 7.1: Correlation coefficient (corr), Mean absolute error MAE, slope s_{xy} and root mean square error RMSE for all model configurations for the Berlin area.

7.5 Conclusions

The impact of the parameterisation scheme for calculating surface fluxes on the model performance was tested. Additionally, the influence of the horizontal resolution on the model performance was included in this paper. Strong sub-grid scale heterogeneities pose a challenge for surface flux aggregation schemes. Especially, urban areas are very heterogeneous. Therefore, emphasis was laid on how METRAS performs for the area of Berlin in Northern Germany. The urban sub-grid scale heterogeneity is represented by large roughness differences and e.g. sealed surfaces next to vegetation. A very locally driven meteorological situation from 9th August until 12th August 2003 was simulated to ensure local processes to dominate the model solution. The presented sensitivity study was carried out using the non-hydrostatic mesoscale transport and fluid model METRAS applying a flux aggregation scheme with blending height concept and a parameter averaging method and using 4 different horizontal resolutions (16 km, 8 km, 4 km, 2 km). The urban areas were represented with adjusted surface parameters but not a very sophisticated urban scheme, since METRAS attempts to represent the overall city-wide effects changed by the surface processes within urban areas without explicitly simulating the urban surface energy balance in every detail.

Even though no sophisticated attempt is made to describe the urban processes, METRAS is capable of capturing essential parts of the urban features. The surface fluxes are altered by the appropriate choice of surface characteristic variables. This already results in an urban heat island in the later afternoon and early evening hours under calm wind conditions in this very locally driven meteorological situation. The horizontal distribution of the urban heat island is well correlated with the fraction of urban land-use. The centre of the urban heat island is slightly shifted westerly due to the easterly wind direction.

Overall, METRAS simulates the difference between the urban and the rural Bowen ratio reasonably well due to the altered surface parameters with the latent heat flux being lower and the sensible heat flux being enhanced in Berlin. This results in higher near surface temperatures for Berlin. The simulation of the Bowen ratio clearly demonstrates the dependence of the parameter averaging scheme on the resolution, while the results using the flux aggregation scheme show little dependence on resolution.

The study highlights the variability in the model performance when using either a parameter averaging scheme or a more sophisticated approach using a flux aggregation scheme with a blending height technique. The flux aggregation approach accounts better for the intense heterogeneity of urban areas. The urban-rural transition regions, where very heterogeneous sub-grid scale surface characteristics have to be combined to yield representative grid scale fluxes are represented more realistically with the flux aggregation scheme. A strong dependence of the parameter averaging scheme on the grid resolution reveals its lack to account for very heterogeneous surface characteristics. While parameter averaging smoothes the model results too drastically for coarse resolutions it converges towards the flux aggregation's results for higher resolutions with comparable hit rates for the prognostic variables. Differences are largest near the surface, where the influence of the averaging scheme is strongest. For instance, a strong resolution dependence of the diurnal cycle of the Bowen ratio is shown in case of the parameter averaging method. The flux aggregation hardly shows significant differences of the spatially averaged thermodynamic values. It usually performs better than the parameter averaging scheme and is nearly resolution independent. In contrast, temperature predictions benefit most from higher horizontal resolutions when using parameter averaging. The sub-grid scale land-use is then resolved more explicitly and the averaged surface characteristics reflect the physical impact on the surface fluxes more realistically. The dynamic parameters like wind speed and direction as well as pressure are significantly less resolution dependent than the thermodynamic parameters.

The calculation of hit rates for temperature, dew point, wind speed, wind direction, pressure and specific humidity underlines the superiority of the flux aggregation scheme with blending height approach as an aggregation scheme for sub-grid scale surface fluxes. Temperatures and dew point temperatures reveal a strong sensitivity to the parameterisation scheme and the resolution applied. The associated diurnal cycles underline the strong resolution dependence of the parameter averaging scheme in particular during convective daytime conditions. Then the fluxes are very intense but are drastically underestimated for coarse resolutions. The diurnal cycles for temperature and depression of the dew point as simulated by the parameter averaging scheme converge for higher resolutions towards the solution of the flux aggregation scheme, since the sub-grid scale heterogeneities are then resolved more explicitly resulting in more realistic grid box averaged surface characteristics.

The rmse, mean absolute error, slope and correlation calculated for the various configurations pinpoint the large variability of the parameter averaging scheme due to resolution issues and clearly underline the capability of the flux aggregation scheme to capture the minima and maxima in the diurnal cycle of the sensible heat flux independent of the resolution. Since the sensible heat flux drives the temperature in local meteorological situations its reliable calculation is essential for a good temperature forecast.

The simulation of the dynamic variables showed a clear underestimation of the wind speed for Berlin as well as for the rural surrounding, and the simulated wind was from too southerly directions. The wind speeds themselves were very low indeed and the accuracy could not be expected to be high under such conditions. In general the dynamic variables showed a much lower dependence on the averaging scheme and the resolution than the thermodynamic values. Also, pressure was clearly underestimated, but since this value was not forced by the ECMWF analysis, but was initialised at a single point in the model domain a better initialisation might have improved its hit rates significantly, since the pressure hardly changed during the simulation period.

This study was restricted to a very dry and locally driven situation to enable METRAS to simulate a local meteorological solution not driven by the lateral model boundaries. Since the soils in Northern Germany were drastically dried out during this period of 2003 it remains open how the model would capture the urban-rural differences for nearly saturated soils. But it can be expected that the differences would be even more intense for higher resolutions where the vegetated fraction within “urban” grid boxes is minimal and hence the evaporation is drastically reduced compared to the rural surrounding.

As a follow-up for future work the results suggest to include a more sophisticated urban surface energy balance scheme to increase the thermal inertia of the urban canopy layer to enable METRAS to simulate the phase shift of ground and sensible heat flux between urban and rural areas better. Consequently this would enable us to simulate the urban heat island more realistically also later into the night and might also result in slightly unstable and well mixed boundary layers in urban areas. In this case the highly varying turbulent fluxes at the sub-grid level would benefit from the more sophisticated flux aggregation scheme. Since it is less dependent on the horizontal resolution a more expensive urban parameterisation could be applied while taking advantage of the good performance of flux aggregation for coarser resolutions.

.

8 Conclusions and outlook

The objective of this study was to investigate the applicability range of sub-grid scale surface flux parameterisation schemes applying a mesoscale numerical atmospheric model. METRAS' model performance was determined with respect to the horizontal grid resolution and the simulated meteorological situation applying a parameter averaging scheme and a flux aggregation scheme with blending height approach. The main improvement of this study compared to previous studies is the development and application of a locality index to evaluate the parameterisation schemes based on a process-oriented approach. The rationale behind this method is the assumption that the importance of surface fluxes for the atmospheric variables varies with the meteorological situation and the horizontal resolution. Under very locally driven meteorological situations a well parameterised surface flux is assumed to enhance the model performance significantly. In contrast, the parameterisation has a minor effect on the model performance for advectively driven situations, where the lateral boundary conditions are more important. Therefore, the locality index is calculated from surface layer scaling variables and is applied to characterise meteorological situations in terms of the impact of the surface characteristics on the atmospheric prognostic variables like e.g. temperature or precipitation probability. One advantage is that the locality index is an objective measure for the strength of the local impact of the surface fluxes at every time-step and it enables METRAS to select the most appropriate parameterisation scheme for sub-grid scale surface fluxes online in order to improve the model performance. A further benefit of applying the locality index is that it gives a hint on the precipitation probability within the model domain and the likelihood that precipitation is altered or even generated by the surface fluxes in the model domain.

8.1 Conclusions

In order to access the model performance of surface flux parameterisation schemes the sensitivity of METRAS towards the initialisation especially of the soil properties needed to be determined at first. Therefore soil moisture, soil temperature and the vertical profile of relative humidity were varied within their uncertainty range for a rather coarse resolution of 16 km.

- It turned out that the changes paid back mostly on the thermodynamic values whereas the dynamic values were affected only slightly. Especially the dew point temperature was considerably affected by slight changes of the soil properties, e.g. a reduction of the available soil water content showed a considerable change in hit rates of up to 20 %.
- Further, it was underlined that the impact of changing the soil parameters and especially the soil moisture was more significant for the 2 m dew point temperature than for the 2 m temperature, since the latent heat flux directly controls the dew point temperature. The temperature is only indirectly controlled by lower soil moistures via an enhanced sensible heat flux due to the lower latent heat flux, which can no longer transport so much energy from the soil up into the air.
- Altogether, the sensitivity of the model towards the correct initialisation of the soil properties is strong and might even out-weight the impact of sub-grid scale parameterisation schemes.

However, a correct initialisation of the soil moisture within the whole model domain is very challenging. Ament and Simmer (2006) for example received good results by running a soil moisture analysis forced by measurements. Importantly, this enabled them to correct the soil moisture in case of precipitation events and their results show some improvement of the surface fluxes due to the corrected soil moisture. In the present study no such system was available for METRAS, but the soil moisture was set according to measurements at the initialisation point in the model domain. Then it was recalculated for the remaining grid points. This method showed better results than taking a higher resolution data set for the soil moisture content from the ECMWF analysis. The ECMWF analysis caused a considerable decrease in the model performance. In order to detect the influence of the surface flux parameterisation schemes, care had to be taken to ensure that although the horizontal grid resolution changed, the initialisation and the forcing data at the boundaries were equal for all simulations.

The characteristics of the precipitation in the model domain were determined with respect to the importance of surface properties to investigate the possibility for an improved precipitation forecast via a better representation of the surface fluxes:

- According to measurements from the high resolution precipitation network within the LITFASS domain and additional routine observations from the DWD, the precipitation in the model domain is very patchy and highly varying on spatial as well as temporal scales.

- The hourly amounts of precipitation at different stations are correlated for distances of the order of 10 km. However, the occurrence of precipitation is well correlated over distances of up to 25 km when taking into account at least hourly precipitation amounts.
- For longer temporal scales the correlation between more distant stations within the LITFASS domain improves. Due to the large observed variability of precipitation amounts even on very small scales, it is concluded that the precipitation amounts as predicted by numerical models can only be evaluated with a very large uncertainty, whereas the occurrence of precipitation with a numerical model can more reliably be evaluated. Daily precipitation amounts are reliable with an uncertainty factor of 3.3 for a 25 km x 25 km area. Weekly and monthly precipitation amounts are representative for an area of 25 km x 25 km with uncertainty factors of 2 and 1.4, respectively.
- Very locally driven meteorological situations do not frequently occur, but those rare situations are connected with a very high precipitation probability. Hence, the surface fluxes are expected to have an impact on the precipitation. Those rare but very convective situations account for nearly 50 % of the annual precipitation amount within this study. It was shown that the precipitation probability increases with increasing locality index. In very locally driven meteorological situations the precipitation probability is highest and it can be concluded that the surface fluxes strongly influence the generation of precipitation. Hence, a good parameterisation of the surface fluxes would not only lead to a better forecast of the thermodynamic and dynamic fields but might also enhance precipitation forecast.
- The relationship found between the locality index and precipitation probability turned out to be very robust, since it holds for two very different years in terms of precipitation amounts and soil moisture availability. Therefore, the index is applicable as a predictor for precipitation events.
- Further, the established relationship between locality index and precipitation probability does also hold for simulated precipitation. The tendency for higher precipitation probability alongside high index values underlines the importance and influence of surface fluxes on precipitation.
- The relationship between locality index and precipitation probability is sensitive towards the parameterisation scheme and the horizontal resolution. Besides the fact that precipitation is less likely to be simulated for very coarse resolutions, the comparison of both parameterisation schemes showed a significant dependency of the model solution on the way the sub-grid scale surface fluxes are included.

Based on these findings the applicability range of the sub-grid scale parameterisation schemes was determined for different meteorological situations and resolutions by applying the locality index as an objective tool to estimate the importance of the surface fluxes for the performance of the simulation. Six case studies were chosen from the 6 hourly locality index values from the year 2003 with varying influence of the surface fluxes at the Lindenberg grid point within the model domain. The assumption was made that METRAS shows a lower sensitivity towards the parameterisation schemes for less locally driven situations and a significant sensitivity was expected from the previous investigations (Chapter 3, Chapter 4 and Chapter 5) for very locally driven meteorological situations. In the following the key findings are summarized:

- The variability of the model performance for different set-ups is negligible for advectively driven meteorological situations, where the surface processes do not contribute significantly to the model solution. Hence, the parameterisation scheme for the inclusion of the sub-grid scale effects plays a minor role.
- In contrast, very locally driven meteorological situations showed strongly varying model answers for the different model set-ups. In such cases the parameterisations are very sensitive towards the strength of the surface fluxes, and well-parameterised surface processes are essential for a reliable model forecast.
- In such very locally driven situations the thermodynamic values were strongly dependent on the parameterisation scheme. The appropriate inclusion of the surface characteristics plays a major role for the correct model solution. The dynamic parameters were less significantly sensitive towards the parameterisation scheme and the resolution. Hence, major improvements of the model performance can be achieved especially for the thermodynamic values by choosing the appropriate sub-grid scale surface flux parameterisation.
- Increasing the resolution did not necessarily improve the model performance. Coarse simulations with flux aggregation resulted in comparable but more cost-efficient model answers. Only in case of parameter averaging higher resolutions resulted in better model results, since the land-use was resolved more explicitly leading to smaller sub-grid scale variability. Then, the parameter averaging approach starts to fulfil the underlying assumption of land-use homogeneity and results in more reliable model answers.
- In general, flux aggregation is the preferable scheme: in very locally driven situations, where the surface fluxes were important, flux aggregation showed the best results independent of the model resolution, while parameter averaging tended to produce better results with increasing resolution. This result is along the lines with the findings made by Schlünzen and Katzfey (2003).

It was clearly shown that parameter averaging gains most from increasing the resolution, while flux aggregation was nearly resolution-independent. Since the parameterisation schemes show the largest deficiencies in very heterogeneous terrain, the investigations focused on the Berlin area to detect deficiencies of the parameterisation schemes. Again, flux aggregation is clearly the superior scheme by being nearly resolution-independent. Further, parameter averaging gained most from increasing the resolution, since the land-use was resolved more explicitly leading to less sub-grid scale heterogeneities in the model grid boxes. It was also underlined that METRAS 3.0 captures most of the urban characteristics, although no sophisticated urban scheme is applied. The urban area is simulated dryer, which is most evident in the difference of Bowen ratios between Berlin and the rural surrounding leading to a warmer Berlin area into the evening hours with a slight urban heat island. The urban sensible heat flux showed a slight phase delay around the evening transition compared to the rural one, which resembles results from Bohnenstengel et al. (2011b) using a more sophisticated urban scheme in the Met Office Unified Model for London.

The present study underlines some of the findings of Ament and Simmer (2006), who claimed that correctly assimilated soil properties are essential for getting an accurate model forecast. It further underlines earlier preliminary findings from Schlünzen and Katzfey (2003) stating that flux aggregation is preferable to the parameter averaging method. But the present study goes further and presents a process-oriented evaluation to conclude on the importance of the surface fluxes for the modelled process itself. The developed locality index was successfully introduced to determine the applicability range of the surface flux parameterisation schemes taking into account the meteorological situation and the horizontal grid resolution and even give a hint on precipitation probability. The locality index enables us to generalise evaluation results from numerical models with regard to the simulated meteorological situation and the importance of the parameterised processes when comparing evaluation results across different models or a single model for very different meteorological situations.

The findings have to be viewed carefully, since for instance changing the horizontal resolution has indeed a direct influence on the representation of the surface and hence the surface fluxes in the model, but further side-effects also need be taken into account. Changing the resolution alters surface fluxes indirectly via the resolution dependence of the microphysics scheme, which might be considered in further more idealized studies. The simulation and distribution as well as the coverage and shape of clouds change due to the resolution-dependence of the clouds scheme. These differences of the microphysics also alter the surface fluxes considerably by shading or precipitation. Therefore, it is very difficult to estimate the pure direct influence of the surface flux parameterisation scheme on the model performance especially in cloudy situations, where several processes affect the model solution.

The flux aggregation scheme turned out to be the superior scheme for calculating the sub-grid scale surface fluxes. In the present study the surface fluxes were averaged at the first model level at 10 m height above the surface. It remains open to which extend the correct averaging at the blending height would change model results. In case of an underestimation or overestimation of the blending height surface fluxes are weighted incorrectly and hence lead to an under- or overestimation of the overall grid box averaged surface flux. Also, advective processes become more important higher up in the atmosphere.

The applicability range of the parameterisation schemes is derived from 6 very different case studies during the warm year 2003. From these results some recommendations can be drawn: In general the flux aggregation method is preferable to the parameter averaging method. Especially, in very locally driven situations the model performance benefits from a correctly represented grid box averaged surface flux at the cost of the computational efficiency. The costs of the more expensive flux aggregation scheme can be offset by the fact that the scheme is nearly resolution independent and a coarser resolution might result in similar model answers as a high resolution simulation. In rather advectively dominated situations, where the surface fluxes play a minor role for the atmosphere a parameter averaging scheme is suitable and has the advantage to be less costly at least given the current resolution of the input data. Also, for relatively homogeneous land-use the parameter averaging scheme should lead to sufficient model results. In very heterogeneous areas the high–resolution parameter averaging scheme and the flux aggregation scheme tend to lead to similar results. The parameter averaging scheme might then be the favourable one when considering computing time. Applying an index to conclude on the best scheme is advisable to improve model results and cut down computing costs at the same time, since the meteorological situation changes very quickly within the model domain and over time. Hence, flux aggregation might be the appropriate scheme at some point but parameter averaging might be advisable at a later time at the same grid point during a simulation. However, from the findings it can be concluded that the thermodynamic values and hence the numerical weather forecast benefit most from the possibility to adjust the surface flux parameterisation schemes. Other model purposes like pollution forecast are less dependent on the thermodynamics but the dynamics, which are not affected significantly. Hence, it also depends on the purpose of the model run, if the present scheme is useful.

8.2 Outlook

For drawing more general conclusions idealized studies might be the appropriate way to limit the influence of other resolution-dependent processes like clouds that also alter the model solution. Furthermore, the results are based on simulations for a very heterogeneous but flat terrain; further studies are needed to evaluate the topographic impact on the parameterisation schemes, since strong heterogeneities in the terrain might be more important than the treatment of the sub-grid scale surface fluxes. All findings in this study are valid for very heterogeneous land-use on small scales. Consequently, further studies are needed to explore the applicability range of the locality index in less patchy regions or areas with hillier terrain, especially with respect to its ability to predict the likelihood of precipitation. It is expected that this relationship works in similar terrain. Since, Weusthoff and Hauf (2008) showed a similar precipitation pattern in the area of Hannover as was found in Lindenberg it can be expected that the locality index is at least applicable in similar environments. Adjustments are expected to be needed in completely different regions like deserts.

The results of the sensitivity study showed, that higher resolutions did not necessarily improve the model performance. With the current tendency to increase model resolutions to resolve the atmosphere with more detail and highly detailed surface land cover data, the land-use categories should be reconsidered and their prescribed surface and soil properties need adjustment, since urban areas appear to be more sealed in current high resolution data sets, where it has no longer to be taken into account, that e.g. trees and gardens exist in urban areas, when those will be resolved explicitly.

To improve the parameterisation schemes further, a validation of the simulated fluxes with observed fluxes would give more insight into deficiencies. Here it has to be considered that single point measurements might not be representative for a larger area and hence be inappropriate for comparison with area-averaged representative fluxes as modelled by a numerical model.

Acknowledgements

This work was partly supported by the German Science Foundation (DFG) through SPP 1167 (grant SCHL499-2) and the Cluster of Excellence 'CliSAP' (EXC177) and partly supported by the Max-Planck-Institute for Meteorology. Data from routine observations were kindly provided by the German Meteorological Service (DWD) and the British Atmospheric Data Centre. Information on meteorological situations was provided by “Berliner Wetterkarte e.V.” The author is grateful to Petra Dereszynski and Sieghard H. Richter (DWD / MOL-RAO) who operate the Pluvio network in the LITFASS area. Thanks for discussions go to Prof. Dr. Jürgen Fischer, Priv.-Doz. Dr. Peter Nevir, Dr. Eberhard Reimer, Dr. Ines Langer and Dr. Andi Walther (all FU Berlin). Thanks for helpful discussions on variogram analysis go to Dr. David Grimes (University of Reading, UK). The author wishes to thank Peter Kirschner (University of Hamburg) for the preparation of Figure 4.1 and the two anonymous reviewers for their helpful comments on Chapter 4 (Bohnenstengel et al, 2012). Chapter 4 was reprinted from Bohnenstengel, S.I. and K.H. Schlünzen (2012): A locality index to investigate the influence of surface fluxes on precipitation events, Q. J. R. Meteorol. Soc. 400 (3-4): 387-395 with permission from Elsevier. Thanks also go to Dipl.-Met. Günter Meschkat for discussions when coding parts of the ECMWF nudging routines.

My biggest thanks go to my supervisor Prof. Dr. K.H. Schlünzen for her continuous support throughout this study and the very helpful and interesting discussions we had. I would also like to thank Prof. Dr. M. Claussen for being my second supervisor and for supporting this work especially during my last year in Hamburg. Special thanks also go to Dr. Frank Beyrich for very helpful discussions on the Lindenberg data and to Dipl.-Met. Patrick Kalb-Rottmann for extracting the Lindenberg precipitation data.

Many thanks go to the METRAS/MITRAS group and the windtunnel group for a great atmosphere to work in and for helpful discussions. Special thanks here go to Dr. Guido Schroeder and Dipl.-Met. David Grawe for sharing the office in a nice atmosphere and to Dr. Merike Fiedler (Schultz), Dr. Elke Meyer, Dr. Hinnerk Ries, Dr. Ursula Bungert and Dipl.-Met Frauke Pascheke.

Thanks also go to the DKRZ for providing computing time and the CIS group, especially to Dipl.-Ing Rainer Senke, for support regarding IT issues. Many thanks also to Volker Reinhard for IT help, especially when I moved to the UK.

Last but not least I wish to thank my friends, family and Peter Strobel for their great support during these years.

List of important symbols

α	bulk soil water availability
c_p	heat capacity
C_h	transfer coefficient for heat
C_m	transfer coefficient for momentum
CSI	critical success index
f	coriolis parameter
g	gravity constant
$\gamma(d)$	variogram for distance d
Γ_θ	counter gradient term for potential temperature
Γ_q	counter gradient term for specific humidity
H	sensible heat flux
H_j	sensible heat flux for sub-grid scale land-use j
h_θ	depth of temperature wave within the soil
I_{adv}	locality index for advection and diffusion impact
I_{lt}	locality index for diffusion impact
I_{rh}	locality index including relative humidity
I_q	locality index including specific humidity
i	index
j	index
κ	von Kárman constant
l_b	blending height
L_z	characteristic length scale in vertical direction
L	Monin-Obukhov-length
l_{21}	latent heat of evaporation
L_x	characteristic length scale
MOL_6	six hourly integrated precipitation amount at MOL
p	pressure
PLUVIO_6av	area averaged 6 hourly integrated precipitation amounts from PLUVIO network
q	specific humidity
q*	scaling value for specific humidity
RH	relative humidity
R	gas constant for dry air

R^1_1	gas constant for water vapour
Re	Reynolds number
ρ	air density
s*	maximum of scaling parameters u_* and w_*
θ	potential temperature
T_{diff}	time scale for diffusion
T_{lt}	time scale for precipitation episode
t	time
T	real temperature
T_{adv}	time scale for advection
T_s	surface temperature
Tv	virtual temperature
θ_*	scaling value for potential temperature
θv	virtual potential temperature
U	main flow in x-direction
<U>	column averaged wind speed over lower 300m
u_{ref}	reference wind speed at first model level
u_*	friction velocity
ψ_m	stability function for momentum
ψ_q	stability function for humidity
w_*	free convection velocity
W_s	depth of liquid water
W_k	field capacity
Z(t)	zenith angle
Pr	Prandtl number
z	height
Z	precipitation amount
z_0	roughness length for momentum
$z_{0\theta}$	roughness length for heat
z_{0q}	roughness length for humidity

Bibliography

Ament, F. and C. Simmer (2006): Improved representation of land-surface heterogeneity in a non-hydrostatic numerical weather prediction model. *Boundary- Layer Meteorol.* **121**: 153-174.

Arakawa, A., and W.H. Schubert (1974): Interaction of a cumulus ensemble with the large-scale environment. *Part I. J. Atmos. Sci.*, **31**: 674-674.

Avissar, R. and R.A. Pielke (1989): A parameterisation of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.* **117**: 2113-2136.

Bakan, S. (1995): A simple radiation parameterisation for use in mesoscale models. *Interner Bericht*. *Max-Planck-Institut für Meteorologie, Hamburg*. pp 37.

Baldwin, M. E., and S. Lakshmivarahan (2002): Rainfall classification using histogram analysis: An example of data mining in meteorology, Intelligent Engineering Systems Through Artificial Neural Networks, 12, C.H. Dagli, A.L. Buczak, J. Ghosh, M.J. Embrechts, O. Ersoy, S.W. Kercel, Eds., ASME Press, 429-434.

Beckmann, B.-R. and T.A. Buishhand (2002): Statistical downscaling relationships for precipitation in the Netherlands and North Germany. *Int. J. Climatol.* **22**: 15–32, DOI: 10.1002/joc.718.

Berliner Wetterkarte e.V., Eingetragener Verein zur Förderung der meteorologischen Wissenschaft c/o Institut für Meteorologie der Freien Universität Berlin, C.-H.-Becker-Weg 6-10, D - 12165 Berlin.

Beyrich, F. (2004): Verdunstung über einer heterogenen Landoberfläche: Das LITFASS-2003 Experiment – Ein Bericht. *DWD – Forschung und Entwicklung: Arbeitsergebnisse*. **79**: 100pp.

Beyrich, F. and H.-T. Mengelkamp (2006): Evaporation over a heterogeneous land surface: EVA-GRIPS and the LITFASS-2003 experiment-an overview. *Boundary-Layer Meteorol.* **121**: 5-32.

Blackadar, A. K. (1962): A vertical distribution of wind and turbulent exchange in a neutral atmosphere. *J. Geophys. Res.* **67**(8): 3095–3102.

Bohnenstengel, S.I., Schlünzen, K.H. and F. Beyrich (2011a): Representativity of in-situ precipitation measurements – a case study for the LITFASS area in North-Eastern Germany. Journal of Hydrology. **400** (3-4): 387-395.

Bohnenstengel, S.I. and K.H. Schlünzen (2012): A locality index to investigate the influence of surface fluxes on precipitation events, *Q. J. R. Meteorol. Soc.* To be submitted.

Bohnenstengel, S.I., Clark, P. and S.E. Belcher (2011b): Simulations of the London urban heat island. *Q. J. R. Meteorol. Soc*. **137**: 1625–1640.

Bohnenstengel S., Schlünzen K.H. (2009): Urbanisation of meteorological and air quality models. Baklanov A., S. Grimmond, A. Mahura, M. Athanassiadou (Eds), Springer Publishers, 165p.

Bohnenstengel, S., Schlünzen, K.H. and D. Grawe (2004): Influence of thermal effects on street canyon circulations, *Meteorol. Z.* **13**(5): 381-382.

Burgueno, A., Martinez, M. D., Lana, X. and C. Serra (2005): Statistical distributions of the daily rainfall regime in Catalonia (northeastern Spain) for the years 1950-2000, *Int. J. Climatol*. **25** (10): 1381–1403.

Businger, J.A. (1973): A note on free convection. *Boundary-Layer Meteorol.* **4**: 323-326.

Buytaert, W., Celleri, R., Willems, P., Bievre, B. de and G. Wyseure (2006): Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*. **329** (3-4): 413–421.

Claussen, M. (1995): Flux aggregation at large scales: on the limits of validity of the concept of blending height: Atmospheric and Hydrological Processes at the Soil-Vegetation-Atmosphere Interface. *Journal of Hydrology*. **166**(3-4): 371–382.

Claussen, M. (1991): Estimation of areally-averaged surface fluxes, *Boundary-Layer Meteorol.* **54**: 387-410.

Claussen, M. (1990): Area-Averaging of Surface Fluxes in a Neutrally Stratified, Horizontally Inhomogeneous Atmospheric Boundary-Layer. *Atmospheric Environment*. **24** (6): 1349–1360.

CORINE Land Cover (2004): Federal Environmental Agency, *German Aerospace Centre*.

Cox, R., Bauer, B. L. and T. Smith (1998): Mesoscale model intercomparison. Bull. Am. Meteorol. Soc. **79**: 265–283.

Cubasch, U., Nevir, P., Reimer, E., Claußnitzer, A. and I. Langer (2006): Berlin Progress report of the first phase of the project: Statistical-Dynamical Methods for Scale Dependent Model Evaluation and short term Precipitation Forecasting (STAMPF), *Institut für Meteorologie, Freie Universität, Germany*.

Datta, S., Jones, W. L., Roy, B. and A. Tokay (2003): Spatial variability of surface rainfall as observed from TRMM field campaign data. *J. Appl. Meteor. Climatol*. **42** (5): 598–610.

de Jong, B. (1973): Net radiation received by a horizontal surface at the earth. *Delft University Press*. 99p.

Deardorff, J.W. (1978): Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation. *J. Geophys. Res*. **83**: 1889-1903.

Dierer, S. and K.H. Schlünzen (2005): Influence Parameters for a polar mesocyclone development. *Meteorol. Z*. **14**: 781–792.

Dupont, S. and P.G. Mestayer (2006): Parameterisation of the Urban Energy Budget with the Submesoscale Soil Model. *J. Appl. Meteor. Climatol.*. **45**(12): 1744–1765.

Dyer, A. (1974): A review of flux-profile relationship. *Boundary Layer Meteorol*. **7**: 362–372.

Feldmann, H., Frueh, B., Schaedler, G., Panitz, H.-J., Keuler, K., Jacob D. and P. Lorenz (2008): Evaluation of the precipitation for south-western Germany from high resolution simulations with regional climate models. *Meteorol. Z.* **17**(4). 455-465.

Gao, S. (2001): On the application of Kalman filters and nudging methods in atmospheric data assimilation. *Dissertation. University of Qingdao, PR China*. pp 115.

Gebremichael, M., Vivoni, E. R., Watts, C. J. and J. C. Rodriguez (2007): Submesoscale spatiotemporal variability of North American monsoon rainfall over complex terrain. *J. Climate*. **20**(9): 1751–1773.

Giorgi, F. and R. Avissar (1997): Representation of heterogeneity effects in earth system modeling: experience from land surface modelling. *Reviews of Geophysics*. **35**(4): 413–438.

Grimes, I.F. and E. Pardo-Iguzquiza, (2010): Geostatistical Analysis of Rainfall. *Geographical Analysis*. **42**: 136-160.

Grimmond C.S.B., M. Blackett, M.J. Best, J.-J. Baik, S.E. Belcher, J. Beringer, S.I. Bohnenstengel, I. Calmet, F. Chen, A. Coutts, A. Dandou, K. Fortuniak, M.L. Gouvea, R. Hamdi, M. Hendry, M. Kanda, T. Kawai, Y. Kawamoto, H. Kondo, E.S. Krayenhoff, S.-H. Lee, T. Loridan, A. Martilli, V. Masson S. Miao, K. Oleson, R. Ooka, G. Pigeon, A. Porson, Y.-H. Ryu, F. Salamanca, G.-J. Steeneveld, M. Tombrou, J.A. Voogt, D. Young and N. Zhang (2011): Initial Results from Phase 2 of the International Urban Energy Balance Comparison Project. *Int. J. Climatol.* **31**: 244–272 DOI: 10.1002/joc.2227.

Grimmond C.S.B., M. Blackett, M. Best, J. Barlow, J.J. Baik, S. Belcher, S.I. Bohnenstengel, I. Calmet, F. Chen, A. Dandou, K. Fortuniak, M.L. Gouvea, R. Hamdi, M. Hendry, T. Kawai, Y. Kawamoto, H. Kondo, E.S. Krayenhoff, S.H. Lee, T. Loridan, A. Martilli, V. Masson, S. Miao, K. Oleson, G. Pigeon, A. Porson, Y.H. Ryu, F. Salamanca, G.J. Steeneveld, M. Tombrou, J. Voogt, D. Young and N. Zhang (2010): The International Urban Energy Balance Models Comparison Project: First results from Phase 1. *J. Appl. Meteor. Climatol.*. **49**: 1268-1292. doi: 10.1175/2010JAMC2354.1.

Haberlandt (2007): Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *Journal of Hydrology*. **332**: 144-157.

Harman, I.N. and S.E. Belcher (2006): The surface energy balance and boundary layer over urban street canyons. *Q. J. R. Meteorol. Soc.* **132**: 2749–2768.

Harman I.N. (2003): The energy balance of urban areas, Dissertation, Department of Meteorology. *University of Reading*. 157pp.

Hebbinghaus, H., Schlünzen, K.H. and S. Dierer (2007): Sensitivity studies on vortex development over a polynya. *Theor. Appl. Climatol*. **88**(1): 1–16.

Hess, R.L.M.a.W.W. (2008): Evaluation of the variational soil moisture assimilation scheme at Deutscher Wetterdienst. *Q. J. R. Meteorol. Soc.*. **134**: 1499–1512.

Heinemann G. and M. Kerschgens (2005): Comparison of methods for area-averaged surface energy fluxes over heterogeneous land surfaces using high resolution nonhydrostatic simulations. *Int J. Climatol.* **25**: 379–403.

Houghton, H.G. (1968): On precipitation mechanisms and their artificial modification. *J. Appl. Meteorol*. **7**: 851-859.

Houghton, H.G. (1950): A preliminary quantitative analysis of precipitation mechanisms. *Journal of Meteorology*. **7**: 363-369.

Houze, R.A. Jr. (1993): Cloud dynamics. *Academic Press*, *New York*. 570 pp.

Hunt, J.C.R., Bohnenstengel, S.I., Belcher , S.E. and Y. Timoshkina (2011): Implications of Climate change for expanding cities worldwide, *Urban design & planning journal.* Accepted.

Huschke, R.E. (1959): Glossary of Meteorology. American Meteorological Society, Boston, MA. 638 pp.

Joss, J. and U. Germann (2000): Solutions and problems when applying qualitative and quantitative information from weather radar. *Phys. Chem. Earth*, **25**(10-12): 837–841.

Lam, J. S. L., Lau, A. K. H. and J. C. H. Fung (2006): Application of refined land-use categories for high resolution mesoscale atmospheric modelling. *Boundary-Layer Meteorol*. **119**: 263–288.

Langer, I. (2005): Institut für Meteorologie, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany (personal communication).

Lettau, H. H. (1979): Wind and temperature profile prediction for diabatic surface layers including strong inversion cases. *Boundary-Layer Meteorol*. **17**(4): 443–464.

Llasat, M.C. (2001): An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of Spain. *Int. J. Climatol.* **21**: 1385-1400.

Lüpkes, C., and K.H. Schlünzen (1996): Modelling the Arctic convective boundary-layer with different turbulence parameterisations. *Boundary-Layer Meteorol.* **79**: 107-130.

Martilli A, Clappier A, Rotach MW (2002): An urban surface exchange parameterisation for mesoscale models. *Boundary-Layer Meteorol*. **104**: 261–304.

Marzban, C. and S. Sandgathe (2009): Verification with variograms. *Weather and Forecasting*. **24**: 1102-1120.

Michelson, D. B. (2004): Systematic correction of precipitation gauge observations using analyzed meteorological variables. *Journal of Hydrology*. **290**: 161–177.

Mölders, N. and A. Raabe (1996): Numerical investigations on the influence of sub-grid-scale surface heterogeneity on evapotranspiration and cloud processes. *J. Appl. Meteorol.* **35**: 782–795.

Mölders, N., Raabe A. and G. Tetzlaff (1996): A comparison of two strategies on land surface heterogeneity used in a mesoscale β meteorological model. *Tellus*. **48A**: 733–749.

Oke, T.R. (2006a): Towards better scientific communication in urban climate. *Theor. Appl. Climatol.* **84**: 179–190.

Oke, T.R. (2006b): Instruments and Observing Methods Report No. 81, Initial Guidance to obtain representative meteorological observations at urban sites, WMO/TD-No.

1250. World Meteorological Organization: Canada.

Pielke, S.R. (2002): Mesoscale meteorological modelling. *International Geophysics Series*. **78**: pp. 676.

Pohlmann, H. (1993): Geostatistical modelling of environmental data. *CATENA*. **20**: 191-198.

Porson A., Harman, I.N., Bohnenstengel, S.I. and S.E Belcher (2009): How many facets are needed to represent a surface energy balance of an urban area? *Boundary-Layer Meteorol*. **132**: 107-128.

Raible, C.C., Bischof, G., Fraedrich, K., and E. Kirk (1999): Statistical single-station short-term forecasting of temperature and probability of precipitation: area interpolation and NWP combination. *Weather and Forecasting*. **14**: 203-214.

Ries, H., Schlünzen, K.H., Brümmer, B., Claussen, M. and G. Müller (2010): Impact of surface parameter uncertainties on the development of a trough in the Fram Strait region. *Tellus A*, **62**: 377–392.

Rice P.L., and N.R. Holmberg (1973): Cumulative time statistics of surface-point rainfall rates. *IEEE Transactions on Communications COM-21*. **10**: 1131-1136.

Roth, M. (2000): Review of atmospheric turbulence over cities. *Q. J. R. Meteorol. Soc*. **126**: 941–990.

Rudolf, B., (2003) Hydroklimatologische Einordnung der in Europa extremen Jahre 2002 und 2003, *Klimastatusbericht 2003*, Deutscher Wetterdienst, Offenbach am Main, Germany.

Schlünzen K.H., Grawe D., Bohnenstengel S.I., Schlüter I., Koppmann R. (2011): Joint modelling of obstacle induced and mesoscale changes – current limits and challenges. *J. Wind Eng. Ind. Aerodyn*. **99**: 217–225.

Schlünzen K.H., Hoffmann P., Rosenhagen G. and W. Riecke (2010): Long-term changes and regional differences in temperature and precipitation in the metropolitan area of Hamburg. *Int. J. Climatol*. **30**: 1121-1136.

Schlünzen, K.H. and E. M. I. Meyer (2007): Impacts of meteorological situations and chemical reactions on daily dry deposition of nitrogen into the Southern North Sea, *Atmospheric Environment*. **41**(2): 289–302.

Schlünzen, K.H. and J.J. Katzfey (2003): Relevance of sub-grid scale land-use effects for mesoscale models. *Tellus*. **55A**: 232–246.

Schlünzen K.H., Bigalke K., Lüpkes C., Niemeier U. and K. von Salzen (1996a): Concept and realisation of the mesoscale transport- and fluid-model 'METRAS', *Meteorologisches Institut, Univerität Hamburg, METRAS Techn. Rep.* **5**: 156pp.

Schlünzen K.H., Bigalke K., Lüpkes C., Niemeier U. and K. von Salzen (1996b): Hints for using the mesoscale model 'METRAS', *Meteorologisches Institut, Universität Hamburg. METRAS Techn. Rep.* **6**: pp 79.

Schlünzen K.H., Bigalke K., Niemeier U. and K. von Salzen (1994): The mesoscale transport- and fluid-model 'METRAS' -- model concept, model realisation, *Meteorologisches Institut, Universität Hamburg, METRAS Techn. Rep.* **1**: pp 150.

Schlünzen, K.H. (1990): Numerical studies on the inland penetration of sea breeze fronts at a coastline with tidally flooded mudflats. *Beitr. Phys. Atmosph.* **63**: 243-256.

Schmid, H. P., Cleugh, H. A., Grimmond, C. and T. Oke (1991): Spatial variability of energy fluxes in suburban terrain. *Boundary-Layer Meteorology*. **54**(3): 249–276.

Schueler, S. and K.H. Schlünzen (2006): Modeling of oak pollen dispersal on the landscape level with a mesoscale atmospheric model. *Environmental Modeling and Assessment*. **11**(3): 179–194.

Skoien, J.O. and G. Bloeschl (2003): Characteristic space scales and time scales in hydrology. *Wat. Resour. Res.* **39**: 1304.

Skok, G. and T.Vrhovec (2006): Considerations for interpolating rain gauge precipitation onto a regular grid. *Meteorol. Z.* **5**(5): 545–550.

Stull, R. B. (1988): An Introduction to Boundary Layer Meteorology, *Kluwer Academic*. 666 pp.

Tiedtke, M., and J.F. Geleyn (1975): The DWD general circulation model description of its main feateres. *Beitr. Phys. Atmosph.* **48**: 255-277.

Trukenmüller A., Grawe D. and K.H. Schlünzen (2004): A model system for the assessment of ambient air quality conforming to EC directives. *Meteorol. Z.* **13**(5): 387-394.

Tustison, B., Harris, D. and E. Foufoula-Georgiou (2001): Scale issues in verification of precipitation forecasts. *J. Geophys. Res.* **106**: 775-784.

van de Beek, C.Z., Leijnse, H., Torfs, P.J.J.F. and R. Uijlenhoet (2010): Climatology of daily rainfall semivariance in The Netherlands. *Hydrol. Earth Syst. Sci. Discusss.* **7**(2): 2085-2120.

Verworn, A. and U. Haberlandt (2010): Spatial interpolation of hourly rainfall – effect of additional infroation, variogram interference and storm properties. *Hydrol. Earth Syst. Sci. Discuss*. **7**(4): 6407-6446.

von Salzen K. and Schlünzen K.H. (1999a): A prognostic physico-chemical model of secondary and marine inorganic multicomponent aerosols I. Models description. *Atmos. Environm.* **33**: 567-576.

von Salzen K. and Schlünzen K.H. (1999b): A prognostic physico-chemical model of secondary and marine inorganic multicomponent aerosols: II. model tests. *Atmos. Environ*. **33**: 1543-1552.

von Salzen, K., Claussen, M., and K.H. Schlünzen (1996): Application of the concept of blending height to the calculation of surface fluxes in a mesoscale model. *Meteorol. Z.* **5:** 60-66.

Walther, A. and R. Bennartz (2006): Radar-based precipitation type analysis in the Baltic area. *Tellus*. **58A**: 331-343.

Watson, P.A., Gunes, M., Potter, B.A., Sathiaseelan, V., and D.J. Leitas, 1982: Development of a climatic map of rainfall attenuation for Europe. Final report of the European space agency, ESTEC CONTR. No. 4162/79/NL. Postgraduate School of Electrical and Electronic Engineering, University of Bradford, UK, 134 pp.

Weusthoff, T. and T. Hauf (2008): The life cycle of convective-shower cells under post-frontal conditions. *Q. J. R. Meteorol. Soc.* **134**(633): 841-857.

Wilks, D. S. (2006): Statistical Methods in the Atmospheric Sciences Academic Press. Burlington, MA., USA, 627pp.