Visualization and Analysis Techniques for Urban Microclimate Data Sets

Vom Fachbereich Informatik der

Technischen Universität Kaiserslautern

zur Verleihung eines akademischen Grades

Doktor der Ingenieurswissenschaften (Dr.-Ing.)

genehmigte Dissertation

von

Kathrin Häb

Datum der wissenschaftlichen Aussprache:

17. Juli 2015

Prüfungskommission/Berichterstatter:

Prof. Dr. Jens Schmitt, TU Kaiserslautern (Vorsitz)

Prof. Dr. Hans Hagen, TU Kaiserslautern

Prof. Dr. Ariane Middel, Arizona State University

Prof. Dr. Gerik Scheuermann, Universität Leipzig

Dekan:

Prof. Dr. Klaus Schneider

D 386

To Nils, the love of my life.

Acknowledgments

Diese Dissertation wäre ohne die Menschen, die im Folgenden gelistet sind, nicht möglich gewesen. Ich möchte all diesen Menschen von ganzem Herzen danken, und zwar jedem gleichermaßen – denn jeder von ihnen hat einen wichtigen Beitrag dazu geleistet, dass ich nun meine Doktorarbeit in den Händen halten kann. Der letzte Teil der Danksagung ist in englischer Sprache verfasst, damit auch meine Kollegen aus Arizona meinen Dank verstehen.

Mein Dank gilt Herrn Prof. Dr. Hans Hagen dafür, dass er nicht an harte Grenzen zwischen Disziplinen glaubt und mir deshalb die Chance gegeben hat, als Geographin / Germanistin in seiner AG zu promovieren. Er hat sich über die Jahre immer wieder für mich stark gemacht, und ich bin außerordentlich dankbar dafür, dass er an meine Fähigkeiten glaubt. Während meiner Promotionszeit war er mir stets ein guter Lehrer, und ich bin stolz darauf, ihn als Betreuer gehabt zu haben.

Frau Prof. Dr. Ariane Middel danke ich dafür, dass sie mir sowohl auf fachlicher als auch auf emotionaler Ebene sehr viel Unterstützung gegeben hat. Sie war mir in den letzten Jahren eine kompetente und verständnisvolle Betreuerin, eine überdurchschnittlich engagierte Gastgeberin und eine sehr gute Freundin. Es war für mich eine große Freude mit ihr zusammenzuarbeiten, und unsere Kollaboration hat viele erfolgreiche Projekte und Ergebnisse hervorgebracht.

Mein Dank gilt der gesamten AG Computergrafik und HCI und AG Computational Topology dafür, dass ich mich von Anfang an in meinem Arbeitsumfeld außerordentlich wohl gefühlt habe. Viele meiner Kollegen sind mir gute Freunde geworden. In diesem Zusammenhang möchte ich auch Frau Mady Gruys hervorheben, deren Bürotür immer offen für mich war, egal bei welchem Anliegen.

Ich danke auch Herrn Prof. Dr. Gerik Scheuermann dafür, dass er so kurzfristig als Gutachter für meine Dissertation eingesprungen ist.

Ganz besonderer Dank gilt meinem Verlobten, Nils Feige, der mich nun bereits seit über 10 Jahren durch alle seitdem vorgekommenen Lebenslagen begleitet und dabei immer bedingungslos für mich da ist. Ich habe unbeschreibliches Glück, ihn an meiner Seite zu haben und kann es kaum erwarten, auch

Abstract

This dissertation focuses on the visualization of urban microclimate data sets, which describe the atmospheric impact of individual urban features. The application and adaptation of visualization and analysis concepts to enhance the insight into observational data sets used this specialized area are explored, motivated through application problems encountered during active involvement in urban microclimate research at the Arizona State University in Tempe, Arizona.

Besides two smaller projects dealing with the analysis of thermographs recorded with a hand-held device and visualization techniques used for building performance simulation results, the main focus of the work described in this document is the development of a prototypic tool for the visualization and analysis of mobile transect measurements. This observation technique involves a sensor platform mounted to a vehicle, which is then used to traverse a heterogeneous neighborhood to investigate the relationships between urban form and microclimate. The resulting data sets are among the most complex modes of in-situ observations due to their spatio-temporal dependence, their multivariate nature, but also due to the various error sources associated with moving platform observations.

The prototype enables urban climate researchers to preprocess their data, to explore a single transect in detail, and to aggregate observations from multiple traverses conducted over diverse routes for a visual delineation of climatic microenvironments. Extending traditional analysis methods, the suggested visualization tool provides techniques to relate the measured attributes to each other and to the surrounding land cover structure. In addition to that, an improved method for sensor lag correction is described, which shows the potential to increase the spatial resolution of measurements conducted with slow air temperature sensors.

In summary, the interdisciplinary approach followed in this thesis triggers contributions to geospatial visualization and visual analytics, as well as to urban climatology. The solutions developed in the course of this dissertation are meant to support domain experts in their research tasks, providing means to gain a qualitative overview over their specific data sets and to detect patterns, which can then be further analyzed using domain-specific tools and methods.

Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit Visualisierungsmethoden für Datensätze, die das städtische Mikroklima und damit den Einfluss einzelner städtebaulicher Objekte auf die Atmosphäre beschreiben. Motiviert durch die aktive Mitarbeit an Projekten zum städtischen Mikroklima an der Arizona State University in Tempe, Arizona, wird untersucht, wie Visualisierungs- und Analysekonzepte für diesen speziellen Anwendungsbereich genutzt werden können, um den Erkenntnisgewinn aus typischen Messdaten zu fördern.

Neben zwei kleineren Projekten, die sich mit der Analyse von mit tragbaren Wärmebildkameras aufgenommenen Infrarotbildern und mit der Visualisierung von Daten aus Gebäudesimulationen beschäftigen, liegt der Hauptfokus dieser Dissertation auf der Entwicklung eines prototypischen Tools für die Visualisierung und Analyse von mobilen Transekt-Messungen. Bei dieser Messmethode wird eine Sensorplattform entlang einer Route durch eine heterogene Nachbarschaft bewegt, um die räumliche Variabilität gemessener Attribute und deren Abhängigkeit von der umliegenden baulichen Form zu untersuchen. Die so gewonnenen Datensätze sind sehr komplex, da sie multivariat, fehlerbehaftet und stark an Raum und Zeit gebunden sind.

Die prototypische Software befähigt Stadtklimatologen dazu, die Daten aufzubereiten, eine einzelne Messfahrt im Detail zu erkunden, und die Beobachtungen von mehreren Fahrten über verschiedene Routen zu aggregieren, um die visuelle Abgrenzung klimatischer Mikrobiome zu erleichtern. Das Tool erweitert traditionelle Analysemethoden durch Techniken, die einerseits Beziehungen zwischen den gemessenen Attributen aufdecken und die andererseits eine Verknüpfung zwischen den Messungen und ihrem räumlichen Kontext herstellen. Darüber hinaus wird eine verbesserte Methode für die Bereinigung der aufgenommenen Daten von Sensorträgheitseffekten beschrieben, die die räumliche Auflösung von Lufttemperaturmessungen, die mit einem langsamen Sensor aufgenommen wurden, erhöhen kann.

Der interdisziplinäre Ansatz, der mit der vorliegenden Dissertation verfolgt wird, trägt sowohl zum Themenfeld der Geovisualisierung als auch zum Forschungsfeld der Stadtklimatologie bei. Die Lösungen, die im Laufe dieser Dissertation entwickelt wurden, sollen Stadtklimatologen einen qualitativ-visuellen Überblick über ihre speziellen Datensätze verschaffen, der dann mithilfe von bereichsspezifischen Analysekonzepten weiter verfeinert werden kann.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Cities are one of the major alterations of a natural landscape, transforming structural and geometrical components as well as the predominant land cover through buildings and impervious surfaces [114, 148, 155]. In addition to that, the inhabitants of a city release heat, moisture and pollutants through their activities [148, 155]. As a result, urban areas show a modification of radiative, thermal, moisture and wind flow characteristics when compared to their surrounding rural landscapes [114].

The discipline of urban climatology investigates these alterations, as well as their effect on the city dwellers [115]. The atmospheric processes related to the modifications induced by urban areas occur at different spatial scales, vertically as well as horizontally [115] (see Section 1.2.1). In this context, the urban microscale considers the atmospheric impact of surface material patches, single trees, or small-scale building and landscape configurations, all of which are important to describe the climate within the *urban canopy layer*, i.e. the area between ground and roof-level [113, 114].

The urban microscale is in the focus of the research activities of my collaboration partners at the Arizona State University (ASU) in Tempe, Arizona, USA. They analyze the physical microclimate dynamics across space and time and in dependency of urban form and design, using a variety of data modalities. The latter range from in-situ measurements within physically available neighborhoods to the simulation of microclimate variability induced by hypo-

thetical urban form and landscaping scenarios [107].

In the context of in-situ measurements, time-varying thermography with a hand-held infrared camera is utilized point-wise in space to gain an overview over the thermo-radiative environment induced by urban features such as trees. Besides, mobile transect measurements are conducted to investigate the extent of and transition between microclimates, as well as their time-varying atmospheric characteristics. These observations are retrieved by traversing different urban design configurations using a sensor platform attached to a golf cart. The resulting data sets are complex and afflicted with uncertainties, which makes their analysis difficult and time-consuming using existing tools and methods.

Motivated by the resulting domain-specific problems and as an attempt to support the described research activities, my dissertation is dedicated to the investigation of appropriate visualization and analysis techniques that can help to increase the insight into the the described in-situ observations. Providing an interactive means to explore these data sets can increase the knowledge about predominant physical interrelations between the investigated objects.

1.2   Background

1.2.1   Urban microclimate research

The importance of scale in urban climatology: Basic definitions

The main objective of urban climatology is the investigation of the interactions between the built environment and the atmosphere [115]. It operates on different spatial scales, vertically as well as horizontally [115].

In the vertical dimension, Oke [114, 113, 115] distinguishes between the *Urban Canopy Layer* (UCL), which ranges from ground to roof-level, the *Roughness Sublayer* (RSL), which ranges from ground to approximately 1.5 to 4-times the roof-level height, and the *Urban Boundary Layer* (UBL), which extends up to that height, in which the vertical atmospheric effect of the urban landscape is not noticeable anymore. The atmospheric processes in the lowest layer, the UCL, are governed by the microclimatic effects induced by individual three-dimensional urban objects, e.g. buildings or trees. The individual

impact of these features is blended at the upper border of the RSL, which is in turn dependent on a variety of synoptic and geometrical factors. Finally, the atmospheric properties above the RSL up to the upper border of the UBL are dominated by the averaged local to meso-scale impact of the "urban surface" with its characteristic roughness, thermal and moisture properties [114, 113].

In the horizontal dimension, the impacts of the urban landscape onto the atmosphere can be observed at the micro-, local, or mesoscale [113, 115]. The *microscale* ranges from less than one meter to hundreds of meters, depending on the size and configuration of individual buildings, yards, streets, small parks, or trees [113]. Consequently, the impact of these urban design elements onto the properties of the immediate atmosperic surroundings can best be observed at this scale. The combination of different microscale processes, generalized to individual "microclimates", determine the atmospheric properties and their spatial variability within the UCL [114]. The *local scale* ranges roughly from one to several kilometers, depending on the extent of more or less homogeneous building and landscape configurations [113, 148]. Observing the urban climate at a local scale means that the atmospheric impacts of entire neighborhoods or city districts are considered, averaging out the microscale impacts of the individual urban objects within these areas [113]. The *mesoscale* corresponds to the scale of an entire city and its surroundings, at an extent of several to tens of kilometers [113, 115]. At this scale, the atmospheric impact of the city as a whole as compared to its rural hinterlands can be investigated.

It has to be noted that the described scales should not be seen as strict partitions, but that they rather appear as merged within a continuum [114]. Nevertheless, as Arnfield [13] states, they constitute the morphological units based on which atmospheric processes are investigated and described, whereas scales with a larger spatial extent hierarchically aggregate all complex processes at scales with a lower spatial extent. This aggregation also results in a decreasing variability of radiative, thermal, moisture, or wind flow patterns [13]. Thus, as a result, the local climate regimes of two adjacent neighborhoods will be less distinct from each other than the microclimates induced by a grass patch with trees versus an adjacent open space with asphalt surface occuring within these neighborhoods [13].

Investigating the urban microclimate: Topics, goals, and data modalities

Although the general topic of urban climatology can be described as the study of the interrelations between the built environment and its surrounding atmosphere [115], the subtopics are many and various. They can roughly be partitioned according to the concept of horizontal and vertical scales described in Section 1.2.1. Since my dissertation focuses on visualization and analysis techniques for the urban microclimate, the topics, methods, and challenges of research activities operating within the UCL and at the microscale are briefly outlined in this Section, without claim for completeness.

In terms of the thermal regime within the UCL, the spatio-temporal characteristics of the canopy layer urban heat island (UHI) is one of the most extensively studied effects [147, 143, 13, 148]. In its most general terms, the UHI can be defined as the thermal difference between an urban area and its rural hinterland, whereas the temperatures within the city center are warmer than those in the city's surroundings [114]. The UHI intensity varies over time, and reaches its maximum during the night [75]. At daytime, shading effects induced by high-rise buildings can even reverse the effect [114, 84, 107].

Although the UHI is obviously only recognizable at the mesoscale, the canopy layer UHI is the result of the combined effects of different urban microclimates, as induced by an urban canyon. The urban canyon consists of adjacent houses and the space between these houses, and it can in simplified terms be seen as the principal unit of the urban canopy layer [114]. The distribution of radiative fluxes occuring within these units are dependent on their geometric properties [143], and the combination of the latter with different kinds of surface materials, vegetation cover, and anthropogenic activities determine the individual climatic behavior of an urban canyon in terms of moisture, air flow, air quality and temperature [114, 109].

The detailed investigation of these effects, either separate or combined, is in the focus of urban microclimatology. On a theoretical basis, the general physical processes governing the exchange of heat and moisture between the different surface elements are studied, as well as the airflow around these three-dimensional obstacles [143]. Increasing the knowledge about these processes leads to improved computational models that can be used for a variety of ap-

plications [32].

The most important application of urban microclimatology is optimization of urban design to create sustainable, comfortable, and therefore livable cities [108, 23]. For example, investigating the airflow regime within an urban canopy can help to determine pedestrian comfort, while the combined effects of urban canopy and urban boundary layer can help to predict the dispersion of pollutants [143]. Predicting the combined impact of urban form and landscaping onto radiative, thermal, moisture, and airflow conditions within the urban canopy layer can assist the design of thermally comfortable outdoor spaces, ameliorating the effects of climate change [108, 23]. Well-chosen building parameters (material, orientation, geometric design) can lead to a lower energy consumption through a decreased necessity of air condition or heating, and thus to a more sustainable urban development [108].

The data modalities used to achieve the described goals can roughly be partitioned into three groups: In-situ measurements of atmospheric variables, remote sensing, and simulation.

In-situ measurements are conducted using weather stations at fixed spatial locations within an urban area. Usually, these weather stations are placed at sites that are considered representative for the *local* climate within a more or less homogeneous neighborhood, without capturing the signal of individual urban features [109, 113, 148]. Currently only a few examples for stationary sensor networks measuring the urban microclimate exist [109]. Frequently, these measurements are conducted using small or hand-held devices which are then placed at the investigated site. In this context, thermal infrared imagery using hand-held devices can be used to complement these observations. However, wireless technology or low-cost miniature sensors are considered to increase the number of stationary microclimate measurement sites [109].

Obviously, stationary measurements can only be conducted pointwise in space. In contrast, mobile measurements using moving sensor platforms are a possibility to increase the spatial coverage of observations, enabling the investigation of the spatial variability of the urban microclimate at a high resolution (see, e.g., [73]). In a less controlled way, data crowdsourcing through novel mobile sensor technologies is increasingly used to not only measure the atmospheric state at covered locations, but to also investigate its impact on the

activities of the participating people [32].

In addition to the described methods to obtain in-situ measurements, remote sensing data are another key data modality. For example, the geometries extracted from LIDAR data sets can help to analyze small-scale radiation environments using measures such as the sky-view-factor [143]. Besides purely geometric information, detailed land cover / land use (LULC) maps are derived from remotely sensed images, revealing the spatial distribution of urban objects and surface materials (see, e.g., [46]). This information provides valuable information about the spatial context of in-situ measurements, but can also be used to inform simulation models, the third modality used in urban microclimate research.

1.2.2 Visualization and analysis techniques for atmospheric data sets

Interactive visual exploration techniques are not commonly used so far by atmospheric scientists, although sophisticated methods and toolkits have been developed by the visualization community [92, 111, 153]. In this context, Nocke et al. [111], Tominski et al. [153] as well as Ladstädter et al. [92] state that the visualization techniques used by atmopheric researchers are frequently restricted to standard techniques, including 2D diagrams such as time series graphs or scatterplots, or colored 2D maps. These (static or animated) visualizations are typically created using statistical toolkits such as MS Excel, R, or Matlab, or general purpose geographic information systems such as ArcGIS [112, 111], an observation I also made during my collaboration with urban microclimate researchers at the Arizona State University. While the resulting plots are easily understandable, they are frequently restricted to summarizing time series or scatterplots without including the spatial context in the case of 2D diagrams, or showing only one variable per image in case of spatial plots [111, 92]. This aggravates the combined visual exploration of the temporal development of investigated atmospheric parameters, their multivariate relationships, and their relation to the spatial context [111]. Thus, interesting features or patterns in the data might remain undetected [111, 92].

On the other hand, a large variety of sophisticated interactive visualization techniques for atmospheric data sets have been developed over the years, especially for global and regional scale data sets stemming from observations, remote sensing, and simulations. Application areas range from multivariate exploration of time-varying uni-modal data sets [92, 149] to the integration of observation and simulation data for verification purposes, either qualitatively [71] or quantitatively [163], to uncertainty visualization in climate and weather models [130]. Ladstädter et al. [92] demonstrate that visual exploration techniques can help to provide holistic views of the investigated climate data sets, which facilitates building hypotheses about the data. In their highly interactive tool SimVis, they incorporated techniques including interactive feature selection, brushing and linking, focus and context, and flexible (algebraic) combination of variables. Enabling scientists to find unexpected features or patterns, such techniques can complement more quantitative statistical analyses, for which, in many cases, a hypothesis is needed beforehand [92]. Helbig et al. [71] use a variety of mapping techniques to combine several data fields in one view, through which users can navigate either on a desktop computer or in a virtual reality environment. Integrating data sets from both simulation and observation data, they also support the qualitative validation of simulation results. Nocke et al. [111] created a library for a suite of visualization techniques that can be used for several analysis purposes related to climate simulation data, using both visualization techniques that are typical for the atmospheric sciences, but also innovative approaches. Tominski et al. [153] visualize global and regional climate networks within their spatial reference frame, encoding important network information using the vertices of the resulting graph and enabling data filtering to reduce visual clutter.

While there has been advancement in visualizing larger scale weather and climate data sets, examples for the visualization and analysis of urban (micro)climate data sets are rare, as also stated by Röber et al. [125]. However, several examples can be found for the visualization of urban microclimate simulations. In this context, Röber et al. [125] use Avizo Green to visualize inner city ventilation, and interactions between an individual building and its environment. Heuveline et al. [74] explore the utilization of augmented reality techniques to visualize airflow around building structures. A more integrated

approach was chosen by Wang et al. [162]. Not only limited to urban areas, the authors combine a microscale meteorological model with Google Maps / Google Earth to facilitate the manual initialization of the model in terms of the three-dimensional objects within the domain and the final visualization of the simulation results, using standard mapping techniques. The authors also demonstrate how Google Maps / Earth can be used to map observation data sets.

Even more rare are visualizations of observation data sets in an urban climate context. There are some examples for the visualization of air quality data sets, which are obviously not necessarily bound to an urban context. For example, Qu et al. [120] developed a visual analytics system for stationary air pollution measurements from Hong Kong. In their tool, they combine several visualization techniques, including a circular pixel bar chart, a parallel coordinates plot with an s-shaped axis to encode wind direction, and a complete weighted graph showing pairwise correlations between measured attributes. Their system, however, does not include the spatial context of the measurement stations. Also dedicated to air quality data, Liao et al. [97] implemented a web-based air quality analysis system for stationary measurements that comprises a map view with pie-charts encoding the distribution of certain air quality indices over a selected time interval, a parallel coordinates plot, and a time-series plot.

Recently, with the advent of mobile sensing and data crowdsourcing, a variety of web-based visualization tools have been developed that either attempt to inform the general public about the collected data sets in their city (see, e.g., [168]), or that are rather designed as a data storage, management, and display systems [87]. While providing interactive navigation of the data sets, these systems are usually limited to standard mapping techniques and do not provide multivariate exploration facilities.

1.3   Contribution

In summary, the contribution of this dissertation is the application and adaptation of visualization and exploration concepts to enhance the analysis of

important urban microclimate data modalities. The investigated approaches go beyond standard mapping techniques that are currently used in domain-specific research. Providing additional views on the complex patterns inherent in observations, the suggested techniques attempt to facilite data-driven reasoning about the relationships between urban form and the atmospheric environment it induces.

The main focus lies on the interactive visualization of data sets resulting from mobile measurements, which can be seen as one of the most complex modes of in-situ observations due to their spatio-temporal dependence, their multivariate nature, and the various error sources associated with a moving platform. Combining concepts from both visual analytics and atmospheric sciences, a prototypic tool is developed that assists urban microclimate researchers not only to explore their measurements in detail and with the explicit consideration of a sensor's spatial context [60], but also to search for multivariate patterns that can finally lead to the delineation of climatic microenvironments within a study site [61].

Within the prototype, a suite of data preprocessing techniques enables the researchers to quickly increase the reliability of their observations and to integrate them with other data sources, including high-resolution land cover maps and measurements from surrounding weather stations. In this context, research was also directed into evaluating and improving domain-specific techniques for sensor lag correction, which can enhance the spatial resolution of mobile measurements taken with a slow air temperature sensor [63].

Besides providing interactive data exploration tool for mobile transect measurements, a smaller project within this dissertation deals with another in-situ observation mode used in urban microclimatology: Thermal infrared imagery with a hand-held device. Observations are conducted to gain a spatial overview of the thermo-radiative environment at a study site, but also to investigate the time-varying progression of specific thermal features. Images recorded with hand-held devices can pose the challenge of varying perspective onto the investigated object, making the manual feature extraction necessary. Motivated by these challenges, a tool was implemented to assist accurate feature extraction and to visually summarize the results of the extraction results, combining well-known image processing and visualization techniques [58].

Finally, going beyond the visualization of the outdoor urban microclimate, currently used visualization techniques in building performance simulation results are evaluated using feedback from an architect [64]. The findings from this study were the starting point for a larger project aiming at the extension of existing exploration techniques.

1.4 Collaborations

Most of the projects described in this dissertation originated from active involvement in urban microclimate research at the Arizona State University (ASU) in Tempe, Arizona, USA. Main collaborators were Research Assistant Professor Dr. Ariane Middel (School of Geographical Sciences & Urban Planning and School of Sustainability, ASU), Assistant Professor Dr. Benjamin L. Ruddell (Fulton Schools of Engineering, ASU), and Prof. em. Dr. Anthony J. Brazel (School of Geographical Sciences & Urban Planning, ASU). The central theme of their research activities is the impact of urban design and landscaping on the surrounding microclimate. In particular, they investigate the spatio-temporal physical dynamics of microclimate in a desert city and their influence on thermal comfort and perceptions. For their research, they use in-situ measurements recorded using a sensor platform attached to a golf-cart, which is then moved through selected study sites located in the Phoenix Metropolitan Area, Arizona. Hand-held thermographic images are used to capture the impact of tree-shade on the spatial distribution of surface temperatures. While both thermographic images and mobile measurements help to examine current climatic conditions in a certain, physically available neighborhood, simulation data sets are created and analyzed to investigate a number of hypothetical urban form and landscaping scenarios in a structured way [107].

Furthermore, the tool described in Chapter 5 was implemented in collaboration with Lars Hüttenberger and Nils H. Feige.

The research described in Chapter 6 was conducted in collaboration with an architect, Eva Hagen, as well as with M.Sc. Diana Fernández Prieto, Dr. Daniel Engel, and M.Sc. Stephanie Schweitzer, at that time members of the Computer Graphics and HCI Group, Department of Computer Science,

University of Kaiserslautern. Furthermore, Dr. Inga Scheler (RHRK Kaiserslautern), and Michael Böttinger (Climate Computing Center, Hamburg) are part of this project, which is still on-going.

1.5   Structure of the dissertation

The main contribution of this dissertation is the prototypic implementation of a visualization and analysis tool for mobile urban microclimate data sets. Therefore, three chapters of this dissertation (Chapter 2 through Chapter 4) are dedicated to this topic. Chapter 2 provides an overview of the research background, data set characteristics, the framework components, and its use case for the detailed analysis of one single mobile measurement run. In Chapter 3, an improved method to correct air temperature measurements for sensor lags is portrayed, before Chapter 4 explains, how the visual delineation of climatic microenvironments can be facilitated based on the spatial aggregation of a set of mobile transect measurements.

A smaller project within this thesis focuses on the analysis of thermal infrared images for urban microclimate research. In this context, Chapter 5 describes a prototypic tool that assists the analysis of a set of time-varying thermographs.

Chapter 6 focuses on the visualization of relationships between architectural design and microclimate. In particular, the state of the art of the visualization in current building performance simulation tools is described based on a literature review and feedback from an architect.

Chapter 7 concludes the thesis.

Chapter 2

TraVis – A Protoypic Software for the Visualization and Analysis of Mobile Urban Microclimate Measurements

This chapter introduces the main project of this dissertation: The implementation of TraVis, a prototypic visualization tool for the analysis of mobile transect measurements in an urban microclimate context. TraVis is designed to support the workflow of analyzing mobile measurements by providing functionalities for data preprocessing, data representation, and data analysis. The framework complements domain-specific state-of-the-art visualization techniques, which mainly use standard mapping and timeline plots, by incorporating spatial context and multivariate relationships into the (visual) analysis. I developed TraVis in close collaboration with two urban climatologists (Ariane Middel and Benjamin L. Ruddell), who use mobile transect measurements to observe the impact of the build environment on microclimate in a desert city. They continuously evaluated the design decisions, and gave important advice about domain-specific analysis techniques.

While this chapter provides an overview of the background of the described research, the sample data set with which it was developed, and the basic system components of the tool, the following chapters describe work that extend the core visualization and analysis functionalities. In this context, Chapter 3

introduces an improved method for sensor lag correction, which is an important preprocessing step for mobile measurements conducted with slow sensors. Finally, Chapter 4 describes a purely data-driven approach for the identification of climatic microenvironments, which can handle diverse mobile measurement routes and spatial data-sparsety.

To a large part, this chapter is based on a peer-reviewed full-paper, which was published in the proceedings of PacificVis 2015 [60]. The description of the data set is taken from a paper that I recently submitted to Urban Climate [63].

2.1 Introduction

Mobile transect measurements are an important tool for urban climatology. A sensor platform is mounted to a vehicle, which is then moved along a predetermined, potentially interesting route in order to investigate the spatial variation of observed parameters. In contrast to stationary measurements, which are collected only pointwise in space, mobile measurements deliver high-resolution spatial data along a line. This is useful to examine the extend and properties of contiguous areas of similar climate, as well as the transition between them. Relating these patterns to the surrounding urban form can inform the design of sustainable and comfortable urban neighborhoods [108].

Data sets from mobile measurements are multivariate and (often-times) time-varying trajectory data. A trajectory is defined as a time-ordered sequence of spatial locations visited by an entity [7, 8, 154, 124]. Mobile measurements build on these two elementary components, although sensors mounted on the moving platform add additional environmental attributes, such as air temperature, surface temperature, relative humidity, short- and longwave radiation or air quality data, varying dynamically along the transect route. The time-varying component of the data set is introduced by repeating the transect measurements periodically. Furthermore, each sensor and therefore each trajectory attribute has a unique and dynamically varying spatial context and representative source area corresponding to the observation.

In this Chapter, I introduce a framework for the visualization and analysis of these complex movement data sets. It is structured as follows: The domain-specific background of this work is described in Section 2.2, leading to the goals

and contribution to the urban climate as well as the visualization community in Section 2.3. Related work, as seen from a visualization point-of-view, is portrayed in Section 2.4. Section 2.5 provides details on the data set, which was used for implementation and testing of the prototype. The main components of the framework are shown in Section 2.6, while Section 2.7 demonstrates, how they can be used to analyze a single mobile measurement run in detail. Finally, a conclusion is provided in Section 2.8.

2.2 Mobile measurements in urban climatology

2.2.1 State of the art

Mobile Measurements

Mobile transect measurements are frequently applied in urban climatology to investigate the spatial and temporal dynamics of atmospheric conditions within the urban canopy layer, i.e. in the space between ground and roof level. Several phenomena can be investigated using this technique. Research usually focuses on either the more general impact of urban form on microclimate [146, 22] or on the analysis of specific climatic phenomena, such as the urban heat island [73, 141, 110, 150] or park cool islands [33]. Studies also investigate health implications of the urban environment, e.g. through air quality measurements [45, 65] or through the analysis of thermal comfort [156].

Mobile transect observations frequently include air temperature and relative humidity [33, 141, 73, 156]. Depending on the research goal of the study, additional variables are derived from these observations, e.g., the dew point temperature [146, 150], or measured. For example, Vanos et al. [156] observe total incoming shortwave radiation to investigate the moderating effect of parks onto the human energy budget. In urban areas, air quality observations are also an important source of information. Thus, flexible measurement systems have been developed and used for mobile air quality measurements. The Aeroflex bike [45] is a notable example for this application.

Usually, measurements are conducted at a fixed level above ground. In

some cases, the behavior of the observations with increasing height above the surface is an integral part of the research question. For example, Chow et al. [33] used mobile measurements at four heights to derive temperature profiles over a variety of surface types to investigate the characteristics of the nocturnal park cool island effect on the Arizona State University (ASU) campus in Tempe, Arizona. Land use / land cover (LULC) surrounding the sensors is frequently included into the analysis, using various methods. Sun et al. [150] relate their observations to the normalized differenced vegetation index (NDVI) within two constant radii around the sensor locations. Heusinkveld et al. [73] use the fraction of buildings, water, and vegetation to statistically analyze the relationship between observations and urban landscape context.

Recently, larger measurement campaigns have been launched that investigate the urban climate by using sensors mounted to public transportation vehicles. These projects cover a larger area, and thus allow the analysis of the urban climate at the city-scale and with a high temporal resolution and regularity. Buttstädt *et al.* [24] use air temperature and GPS loggers attached to busses to increase the spatio-temporal coverage of air temperature data for the city of Aachen, Germany. The same team developed a prototypic multi-sensor measurement device in the URBMOBI project [88, 127], which measures air temperature, relative humidity, and solar radiation. It can also be attached to vehicles moving through a city, monitoring data to inform models [127], or to identify (un-)comfortable areas within a city [88]. The German Weather Service (DWD) mounted air temperature and relative humidity sensors to streetcars to retrieve information about spatio-temporal dynamics of the air temperature distribution in Halle, Germany, which in turn complements data retrieved from stationary measurements and simulation models [41]. While the named examples focus on the observation of the urban heat island or try to derive implications about the thermal comfort of city dwellers, other larger measurement campaigns systematically acquire data about air quality. The AERO-TRAM project [65] is one of these attempts: Here, sensors are attached to streetcars, and data along two different routes are collected continuously. Similarly, Hasenfratz *et al.* [70] equipped ten streetcars in Zurich, Switzerland, with observation systems for ultra-fine particles. Using these data, they developped a regression model that eventually results in a city-scale air pollution

map.

Visualization

In the urban climate literature, mobile transects are frequently displayed as two-dimensional plots with the spatial location on the x-axis and the observed quantities on the y-axis [150, 156, 65]. Sometimes, the x-axis is complemented by LULC information located directly under the sensor platform to facilitate the relation of measurements to spatial context [22, 146]. Another frequently used visualization is a two-dimensional spatial plot of transect measurements on a map [73, 33, 45], displaying only one variable, one measurement height, or one point in time. In TraVis, I combine the advantages of these visualization methods, displaying a user-defined number of variables as stacked ribbons on a background map, and transferring the generally two-dimensional diagrams into their native spatial context.

Recently developed measurement systems, such as URBMOBI [88, 127] or the Aeroflex bike [45], include their own automated processing and visualization pipelines. In the URBMOBI system, collected and processed data are fed into simulation models, the results of which can then be explored along with the transect measurements using the MEA user interface [106, 88, 127]. This interface provides a data integration facility, which is capable of combining information from different sources and with diverging spatial resolutions, as well as standard visualization techniques to explore the data. Thus, a user can choose an area of interest (AOI), and display the distribution of several variables within the AOI using side-by-side images. Furthermore, MEA allows the extraction of time series for individual locations, which are then displayed in a standard time-series plot [106]. For the Aeroflex bike, Elen *et al.* [45] developed a fixed data processing pipeline. The raw data collected during single runs can be visualized using a quick mapping interface called VITO SensorView, which allows plotting one variable at a time simultaneously on a map and in a time-series diagram, limited to this single run [159]. After that, the data runs through an automated data preprocessing chain, before it can be visualized using a web-based interface based on a GeoServer and Google Earth plug-in. This map is updated after completion of additional measurement runs, while repeated runs – which are conducted along the same route and at fixed points

in time – are displayed on the map using summary statistics [45]. Another example is Vizzly [87], a data storage and aggregation tool that maps aggregated measurements from both mobile and stationary sensors in a timeline plot and on a map. The tool enables interactive browsing for very large sensor data sets at different spatial and temporal resolutions, and was, e.g., used for a larger air quality measurement campaign using sensors on public transportation vehicles in Zurich, Switzerland [70]. Boufidou *et al.* [21] developed a database that is meant to facilitate the management of urban climate data sets. It can manage geometrical and material information related to the urban landscape, as well as mobile air temperature measurements that have been collected with an Arduino platform.

In contrast to the mentioned data management and visualization tools, TraVis was developed with a focus on the visualization and analysis of mobile measurement data collected during smaller campaigns. Using visualization research, the prototype enhances standard mapping techniques to provide detailed views on the multivariate and complex nature of mobile transects at the microscale.

2.2.2   Data set characteristics

Mobile transect measurements are multivariate geospatial movement data. Using the notation described in [3, 154], the latter comprise a set of spatial locations per trajectory, $S = \{s_0, ..., s_n\}$, and a time stamp t_i associated with each element of S, yielding $T = \{t_0, ..., t_n\}$, with $n > 0$ and indicating the number of sampled positions. Using T and S, further sets of dynamically changing attributes $A_{0,...,m}$ can be computed per location and time stamp, including movement direction, velocity, and acceleration [8, 154]. Similar to the data sets visualized in [154], mobile-platform transect data sets involve additional, dynamically changing attributes $A_{m,...,n}$ complementing S and T, such as air temperature, relative humidity, or surface temperature.

In addition to these properties, mobile transect measurements feature two peculiarities. Firstly, the attributes collected during mobile transect measurements are related to a certain height above ground level (a.g.l.), which is either given by the mounting height of the sensors for in-situ measurements (e.g., air temperature) or by the location of the investigated object for remotely sensed

data (e.g., surface temperature). This information is an important aspect during data analysis. It complements the 3D spatial locations of a moving entity's path in conventional trajectory data sets, which is commonly described by a set of 3D geographic coordinates (latitude, longitude, and elevation above mean sea level (a.s.l.)). Since the set of spatial locations is usually displayed on a two-dimensional map that represents the projection of the Earth's surface onto a plane, the measurement height a.g.l. introduces a vertical offset dimension in addition to the resulting set of two-dimensional coordinates on that plane.

Secondly, the analysis of the spatial context of the data differs from usual trajectory data sets. In conventional trajectory data sets, the spatial context is analyzed to facilitate reasoning about certain movement patterns. For mobile transect measurements, the trajectory is exogenously fixed based on the aim to observe relationships between multiple environmental attributes and the spatial context, utilizing spatial context to explain attributes rather than the trajectory itself. The multiple sensors in a research transect each measure a different field of view, and as such each of multiple attribute observations has a distinct and separate relationship to the spatial context.

2.2.3   Challenges during data analysis

The specific tasks during the analysis of mobile measurements depend on a central research question that is individual for each study. Nevertheless, universal tasks could be identified by the collaborators of this research, Ariane Middel and Benjamin L. Ruddell, which are also mirrored in the literature. To summarize these findings, a very general workflow is depicted in Figure 2.1.

After observation, climatologists need to preprocess mobile observations to establish precise measurements of attributes and exact relations to their spatial context. For example, this includes a time-detrending step to filter out potential impacts of changing reference weather during the course of a run [147]. This way, the variability of the attributes along a route can more reliably be traced back to the surrounding urban form and landscape. Furthermore, sensor inertia has to be taken into account. If the sensors are too slow to synchronize their readings with the conditions of the traversed microenvironments, the spatial resolution of the observations will be reduced. This error source can be avoided either through correction (see, e.g., [1]) or
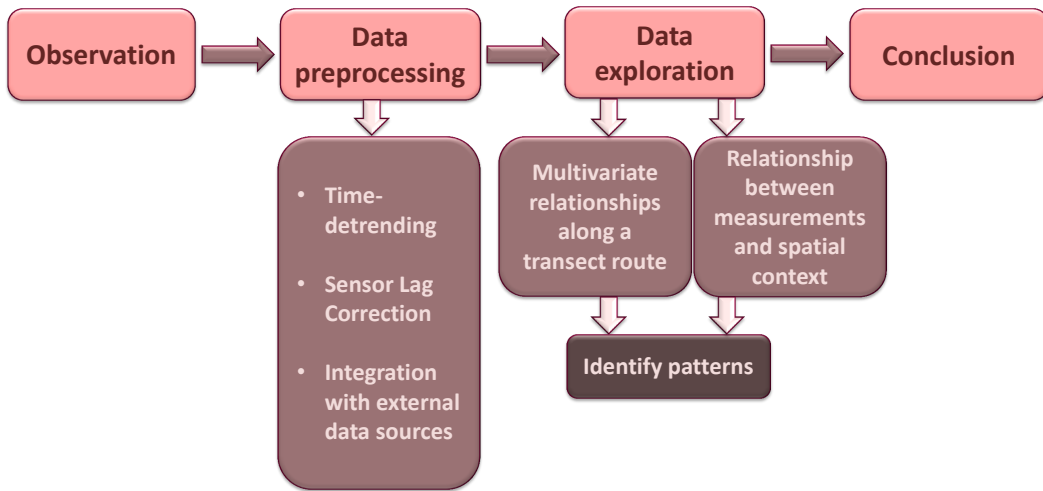
Figure 2.1: A simplified workflow for mobile measurements: From observation to conclusion.

through adapting the number of and distance between measurement points to the spatial resolution dictated by the sensor's inertia [24].

Besides these data-immanent correction procedures, the observations also need to be integrated with external data sources that complement the information provided by the mobile measurements. This information can, e.g., consist of maps, revealing the spatial context of the traversed route. Also data from stationary observations can be useful either as a base for comparisons with mobile measurements (e.g., [24]) or to retrieve additional information about the weather conditions during a measurement run (e.g., [150]).

After preprocessing, the data can be analyzed in terms of the individual research goal. However, a common goal of mobile measurements is the investigation of the spatial variability of observed attributes (because otherwise, stationary measurement would suffice), and to explore the reasons for this variability (e.g. in terms of a varying landscape). Patterns identified during exploration can then lead to a conclusion.

2.3   Research goals and contribution

The overarching goal of this project is the development of a tool that assists urban microclimate researchers during data preparation, transformation, exploration, and analysis of the relationship between multiple mobile platform observations and spatial context. As such, this research contributes to both the discipline of urban climatology and visualization.

2.3.1   Contribution to the urban climate community

The contribution of this project to the urban climate community is an interactive visualization and analysis tool, which enables researchers to explore their mobile measurements in detail, to put it into relation with the spatial context, and to detect potentially relevant patterns in the data based on clustering and aggregation techniques. To the best of my knowledge, to date no tool exists for the comprehensive analysis of the collected data, although mobile transect measurements are a common observation technique in urban climatology.

While the tool includes standard preprocessing techniques such as time-detrending, research has also been directed towards the evaluation and improvement of existing approaches for sensor lag correction (compare Chapter 3). The stacking-based visualization approach combines the advantages of the frequently used uni-variate geospatial mapping with the evenly frequently used multivariate x-y-diagrams, enabling the simultaneous inspection of a set of measured attributes in their spatial context. At the same time, the estimated sensor-specific source area can be used to relate the measurements to the land cover fractions that theoretically impact them. Finally, I implemented a visualization technique for spatially aggregated mobile measurements that integrates data from diverging traverse routes. The visualization does not just use simple aggregation measures, but it provides a detailed representation of urban microenvironments, their characteristics, their extends, and transitions between them.

2.3.2 Contribution to the visualization community

From a visualization perspective, this project adds to the research field of geospatial trajectory visualization: I developed an interactive visualization for *multivariate* trajectory attribute data, providing a detailed view of their dynamic behavior during a mobile transect measurement run. In addition, the dynamic spatial context is not limited to visualizing a background-map, but its impact onto the trajectory attributes can interactively be explored employing the meteorological concept of a *source area*, which is a notion for the spatial field of view of this class of sensors [136]. An interactive clustering interface enables the user to classify transect segments according to coherent patterns of multivariate relationships either between atmospheric attributes alone, or between atmospheric attributes and their associated land use fractions within the sensor-specific source area. Finally, a glyph-based visualization technique for aggregated multivariate trajectory data was developed and added to the framework (Chapter 4). As a result, the prototype comprises capabilities to qualitatively and quantitatively explore spatial and multivariate interdependencies within the data set under investigation, and structures the underlying space according to what has been measured on top of it.

2.4 Related work

Since mobile transect data sets share several characteristics with conventional trajectory data sets, the related work for this research mainly consists of visual analytics for movement data, which will be described in this subsection.

During the last decades, research has been conducted on the visualization of and interaction with large geospatial trajectory data sets. Standard trajectory data sets are multidimensional, and include space, time, and other attributes associated with movement (speed, direction, acceleration), moving entities (age, gender, vessel type), or – as in the case considered for this research – contextual environmental attributes that change over the course of a trajectory [160]. Depending on the research goal and the underlying application scenario, studies concentrate on one or more of these dimensions.

Studies focusing on the spatial dimension of movement data frequently cluster, filter, or aggregate trajectories to facilitate reasoning about explanatory

patterns. Rinzivillo *et al.* [124] apply progressive clustering to a data set describing car movement in Milan, where simple distance functions based on start, destination, a set of waypoints, temporal distances between those points or a combination of these aspects can be used successively. Andrienko *et al.* [6] describe an iterative clustering approach on a trajectory data set, which relies on the initial definition of a "classifier" resulting from clustering a subset of the data set. This classifier is then applied to the remaining data, while it can also be interactively refined if necessary. Andrienko and Andrienko [7] use accumulations of significant points on trajectories (starting points, destinations, stops, turns) to partition the underlying space. Movement occurring between these areas is aggregated and visualized using a flow map. Krüger *et al.* [91] developed "TrajectoryLenses", an intuitive spatial filter interface for a large trajectory data set. Using three different kinds of TrajectoryLenses, a user can investigate the set of trajectories with common starting points, end points, intermediate route points, or a combination of these. A temporal filter can also be applied to the data by selecting an interval on a hierarchical time slider.

In addition, studies include spatial semantics and a visualization of how trajectory patterns change over time. One example is described by Krüger *et al.* [90], where information about places of interest in direct adjacency to the trip destinations are streamed from a location-based application called Foursquare. A time-line view reveals which places were visited when and how often. Growth ring maps were introduced by Bak *et al.* [17]. Whenever an entity visits a specific location, the ring around that place grows. Spatial semantics and temporal information are included by color-coding the rings. In a study by Wang *et al.* [164], spatial semantics occur as an underlying road network. Trajectory data recorded by taxis are matched with this network to investigate the spatial distribution of traffic jams in Beijing, and the result is visualized as a 2D map. The average speed of the taxis is encoded as additional attribute through color-coding. A time-line view shows the temporal distribution of the traffic jams, while a graph projection view displays the interrelations between traffic jams over space and time. Lundblad *et al.* [98] developed SWIM, a tool designed to help shipping companies in route planning. In SWIM, the trajectories of ships are set into context with weather

forecast data along their routes, using linked views to explore the atmospheric spatial context of one or more ships.

The simultaneous analysis of space, time, and multiple attributes has also been addressed in recent studies. Hurter *et al.* [80] describe a visual query system to select and analyze data subsets in several dimensions. The query results can be spread out over several displays to efficiently analyze subsets of the data, while still having the remaining data at hand. The visualization relies on 2D-scatterplots, i.e., attributes are paired without spatial context.

Some approaches spatially plot multiple trajectory attributes. Lampe *et al.* [94] use quantitative differences between kernel density estimate plots to compare movement, taking into account different time intervals, or different attributes either associated with the moving entities or the environment. Interactive exploration of the data set is facilitated by multiple coordinated views that are connected to each other via brushing and linking. Scheepens *et al.* [134] develop a density map, in which several movement density fields can be combined by three different aggregation techniques. The final density is encoded using a height map, while the initial density fields are color-coded for differentiation in the final image. Later, this visualization was extended [133]; the kernel for the density estimation can now be dynamically adapted to attributes of interest, and the combination of single density fields was designed to be more flexible.

Attempts have been made to encode space, time, and additional attributes of *individual* trajectories in detail. Visualizing trajectories on this level is certainly not always advantageous. It can lead to enormous clutter, and the overview is lost, unless route variations are low or the number of trajectories is small. Both conditions may be met for microclimate mobile-platform transect data sets, which makes studies dealing with such detailed visualizations valuable for my research.

Ware *et al.* [166] developed "TrackPlot", a tool to visualize whale tracks. The authors encode a whale's trip through 3D-space as a ribbon, while the movement direction is shown by arrows on the ribbon's surface. The ribbon is wound around a central axis, highlighting the animal's roll behavior, and glyphs are attached to show certain swimming techniques. Vrotsou *et al.* [160] simplify trajectories into movement episodes and display the result-

ing regionalized track. Closely related to the transect segmentation capability of my solution, they also provide the opportunity to classify these movement episodes according to similar multivariate relationships. However, in contrast to my implementation of a detailed mobile transect visualization, they do not show the details of multivariate relationships on the map.

The study most closely related to the visualization display described in this chapter is the framework introduced by Tominski *et al.* [154]. This approach combines time, space, and one attribute in a hybrid 2D/3D visualization display. The authors stack individual trajectories as color-coded ribbons along the z-axis, ordered by a user-defined attribute. The temporal dimension can be included into the visualization by *temporally* ordering the ribbons along the z-axis. Additionally, the temporal variation of attributes is investigated by using a "time lens", which is a circular display that can be moved to a position of interest. The visualization technique used to represent a single transect run relies on several design decisions presented in [154]. Although the authors refer to the clustering capabilities of their framework to include multivariate relationships into the analysis, a limitation is that only one trajectory attribute can be visualized in detail per display instance. I overcome this limitation by using the third dimension for the synchronous display of multiple attributes of interest that belong to the same trajectory (by the cost of not showing multiple transects at once), while additionally using clustering techniques to summarize multivariate relationships. Furthermore, I enable the user to interactively and quantitatively explore the spatial context of the trajectories.

2.5 Sample data set

2.5.1 Study site

The data sets used to conduct this study were recorded by Benjamin L. Ruddell, Ariane Middel and their team at Power Ranch, a master-planned community in Gilbert, Arizona, USA (33.27 N, -111.69 W, 406 m a.s.l.). It is located in the southeastern part of the Phoenix Metropolitan Area (Figure 2.2), which in turn is situated in the northeastern Sonoran Desert. The climate in Phoenix can be described as hot and semiarid, with mean daily minimum temperatures

ranging from 7.1°C in December to 28.1°C in July, and mean daily maximum temperatures ranging from 19.6°C in December to 41.0°C in July. The average annual precipitation is 207.7 mm ([171]).

Power Ranch mainly consists of low-rise residential buildings arranged in

Figure 2.2: Location of the study area (background image: Google Earth).

single-family home neighborhoods, lined up along asphalt roads and concrete sidewalks. All homes are surrounded by yards, which frequently feature a swimming pool and have landscaping types ranging from desert-like xeriscaping to irrigation-intensive mesiscaping. The space between neighborhoods is occupied by parks, a large portion of which is covered by grass. An artificial lake is located in the center and in the northern part of Power Ranch, respectively (Figure 2.4).

2.5.2 Instrumentation

Four air temperature sensors are attached to the measurement platform. Two of these sensors are Resistance Temperature Detectors (RTD) used in the Rotronic HC2S3 combined temperature and relative humidity probe (purchased via Campbell Scientific). The RTDs are platinum type and have a resistance of 100 Ω (Pt100 RTD) [28]. Their rated measurement range lies between -40°C and $+60$°C, while their precision is ± 0.1°C at 23°C [28].
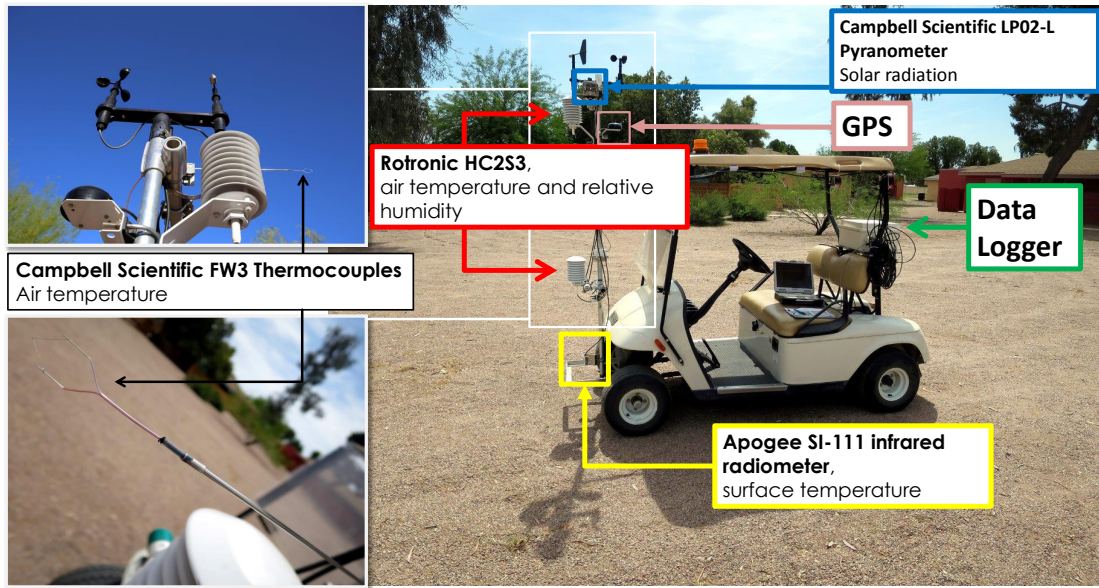
Figure 2.3: Measurement platform and sensor setup.

The other two sensors are Campbell Scientific FW3 thermocouples, which are Fine-Wire Thermocouples (FWT) of Type E (Chromel-Constantan) with a diameter of $0.0762\ mm$ $(0.003\ in)$ [26]. In accordance with the thermistor located onboard the CR1000 data logger, they have a precision of $\pm 0.1°$C in a range between $0°$C and $40°$C [27]. A typical time constant for a thermocouple of this diameter is $0.525\ s$ in still air (linearly interpolated from table 2 in [116] [1]).

In addition to the air temperature sensors, relative humidity is measured at $1\ m$ and $2\ m$ height using two relative humidity probes of Rotronic's HC2S3, i.e. the Rotronic Hygromer IN-1 with a precision of $\pm 0.8\%$ relative humidity at $23°$C [28]. The time constant reported in the specifications ranges between $22\ s$ and $30\ s$, dependent on the utilized filter [28].

The sensors are mounted on a mast attached to the front of a golf cart (Figure 2.3). One group of sensors consisting of both one HC2S3 combined temperature and relative humidity probe and one FW3 thermocouples is located at each of the two measurement heights ($1\ m$ and $2\ m$). The HC2S3 sensors are each housed in naturally ventilated M.R. Young 10-plate radiation

---

[1]The response times for a fine gage bare wire thermocouple of Type J (Iron-Constantan) are listed in this table. However, Campbell Scientific Customer Support referred to this table for a time constant estimate for the FW3 Type E thermocouple.

shields.

Air temperature and relative humidity sensors have been calibrated by the manufacturer. However, to make sure that sensors at equal heights measure the same air temperature, the domain-experts recorded a reference time series over 4.5 h at night with the platform remaining at the same place. The comparison between the average air temperature sampled by the FWT and the RTD mounted at 2 m showed a difference of 0.04°C, while the air temperature sampled with the FWT and the RTD mounted at 1 m deviated by 0.044 °C. Since this difference is negligible, the sensors are still calibrated well enough.

Surface temperature is measured using an Apogee SI-111 infrared radiometer. This radiometer has a spectral range of 8 μm to 14 μm, a 22° half angle field of view, and an operating environment of -55°C to 80°C [10]. The radiometer needs 0.6 s to respond to 95 % of an impulse change [10]. In addition to that, solar radiation is measured using a Campbell Scientific LP02-L Pyranometer [29], while wind speed and direction are observed using a cup anemometer and a wind vane.

Meteorological data were recorded at 1 Hz frequency and logged with a Campbell Scientific CR1000 data logger, which is mounted to the back side of the golf cart. In addition to the meteorological data, the platform velocity vector and the position of the golf-cart is tracked using a GPS device.

2.5.3 Data collection

The data were recorded during 21 transect runs (Figure 2.4) conducted on four different days and three different times of year (May, September, and February). Each transect covers areas with buildings and open spaces, thus representing a variety of land covers. The approximate physical scale of expected microclimate variation ranged between approx. 10 m (moving the golf cart over an intersection while driving between two arrays of evenly spaced houses before and after this intersection) and approx. 400 m (crossing the park in the center of the study area in east-west direction). Microclimate heterogeneity is maximized by driving the transect platform through intermittent shade of trees located next to the route and by periodically stopping the platform in selected microenvironments. This was extensively done during 17 of the 21 transect runs, with the cart on average being stopped for 65.5 % of the
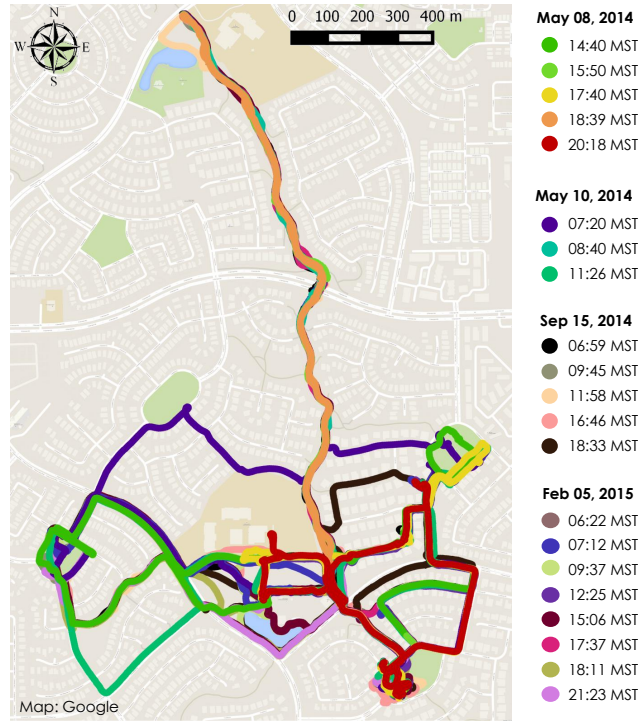
Figure 2.4: Transect runs in Power Ranch (background map: Google).

entire transect duration. While not stopping, the platform was moved at an average speed of 3.51 m/s. The remaining three runs are "path runs", i.e. the cart is not stopped as often as during the tree runs (on average only 18.3 % of the entire transect duration). For the path runs, the platform is moved at an average speed of 4.88 m/s when not stopping. The measurements were carried out between 14:40 MST and 21:22 MST on May 08, 2014 (five runs), between 07:20 MST and 12:57 MST on May 10, 2014 (three runs), between 06:59 MST and 19:36 MST on September 15, 2014 (five runs), and between 06:22 MST and 22:14 MST on February 05 (eight runs). The duration of the transect runs ranged between 16 and 144.5 minutes, which is long enough for the diurnal cycle to affect measurements. Weather conditions during the runs are listed in Table 2.1; they are mostly warm, stable, and clear.

2.6   System components

My tool combines important data transformation and contextualization steps, a visual interface, and interactive analysis capabilities to facilitate the explo-

Date and time	Average air temperature [°C]	Average wind speed [m/s]	Average relative humidity [%]	Cloud cover
May 08, 2014 14:40 MST – 21:22 MST	26.40	1.06	18.17	thinly scattered to scattered in early afternoon
May 10, 2014 07:20 MST – 12:57 MST	27.21	1.38	19.90	clear
September 15, 2014 06:59 MST – 19:36 MST	36.02	1.31	29.54	thinly scattered to scattered starting from noon
February 05, 2015 06:22 MST – 22:14 MST	19.58	0.50	50.30	thinly scattered until late afternoon, then clear

Table 2.1: Reference weather during the mobile measurements, averaged over the time span in which all measurements of a day took place (reference meteorology from MesoWest 2014 [105]: Cloud cover from KIWA Phoenix-Mesa Gateway, temperature and wind data is the average of the data from four nearby stations surrounding the study site, i.e. AU340 Gilbert, D2495 Gilbert, AT202 Gilbert, and SRP31 Rittenhouse).

ration of mobile-platform transect data. An overview over the system components is given in this Section, while another important feature – the spatial aggregation of a set of mobile measurements – will be described in Chapter 4. The functionalities of the prototype can be steered using the main user interface shown in Figure 2.5, as will also be described in context with the description of the framework components. The prototype was implemented using C++, Qt 4, and openGL 4.

2.6.1 Database setup

To enable queries and real-time interaction with the data, transect data files are loaded into a relational database (MySQL 5.6, Oracle Corporation). Each transect run is stored as a table in the data base. The results of the preprocessing steps are added as additional columns so that each preprocessing step has to be performed only once, while all intermediate and final results of the analysis are stored. This way, TraVis can also only be used as a data
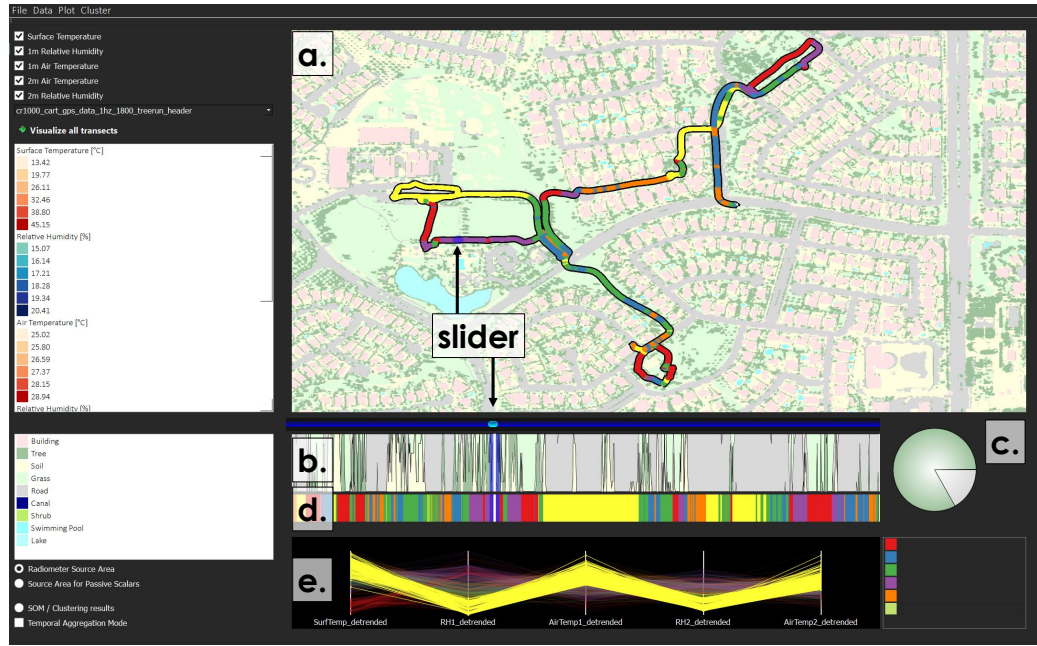
Figure 2.5: The graphical user interface. (a) The Map View. Plan-view overlay of a transect on the LULC layer, visualizing a segmentation result (see also Section 2.6.4; background map: [46]). (b) The Fraction Plot, showing the LULC fractions within the surface temperature infrared radiometer's field of view. (c) The pie-chart that complements the fraction plot. (d) Classification result based on surface temperature and land use fractions in the infrared radiometer's source area. (e) Parallel coordinates show the semantical meaning of the classes, which have been determined during the segmentation process.

preprocessing tool, and, as a trivial extension of the current implementation, preprocessing results can be easily be exported for further processing in other environments.

Range, mean, and standard deviation of the attributes are provided through a database summary functionality (Figure 2.6).

2.6.2 Data preprocessing

Several data transformation steps are required to correct the spatial and temporal position of the transect observations for an accurate visualization. First, the time series recorded by the sensors during the transect is linearly time-
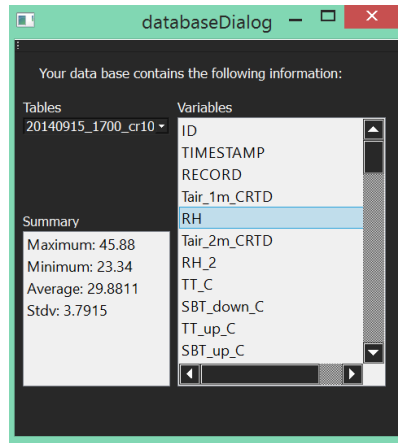
Figure 2.6: The database summary functionality.

detrended to estimate each observation's value at the transect's starting time. Reference conditions can change significantly during the time elapsed while the transect route is traversed, contaminating the observed relationship between spatial context and environmental attributes along the transect route [147]. Time-detrending is performed individually for each measured time series by using a linear best fit time-trend to adjust all temperatures to the baseline observed at the beginning of each transect.

Data transformation also includes Geographical Information System (GIS) georeferencing, projecting the attribute observations into the coordinate system of the spatial context maps. In the sample data set, the GPS positions are given as geographical coordinates and projected into a UTM conformal projection to store all data in a Cartesian reference space. The GIS coordinate transformations are carried out using the open-source GDAL library [53].

After projection, I extract the LULC type directly underneath the sensor platform location from a high-resolution LULC map [46] and store them as additional attribute in the corresponding tables, i.e. each measurement tuple is complemented with the LULC information. This is the immediate spatial context of the mobile platform itself, and is important for the visualization strategy, as will be explained in the next Section.

Finally, a filtering step is conducted before transect data are rendered to avoid visual artifacts. The tuples corresponding to temporary pauses of the mobile platform's movement are extracted. In this case, the filter is applied

to tuples in which the platform moves slower than its average speed. Then, the extracted tuples are grouped according to their sequence on the trajectory. For each connected component, the central vertex of the resulting trajectory segment is used as a spatial representative, and is associated with attributes that have been averaged over all members of the particular trajectory segment. Note that the results of this preprocessing step are not stored in the data base, but calculated each time a transect run is queried. This way, potentially important information is not lost.

2.6.3 Data representation

Similar to [154], my visualization relies on a hybrid 2D / 3D representation of the data in its geographical reference space using wall-layers stacked in z-direction on a two-dimensional land use / land cover (LULC) map. In contrast to [154], I do not use the third dimension to visualize the dynamics of one specific attribute on *different* trajectories. I rather use the individual wall-layers to display the values associated with different environmental attributes of the *same* trajectory, approximating the intuitive notion of measurement height above the ground that is inherent in the data.

The spatial distribution of attribute values along the route can be interactively explored in the map view (Fig. 2.5a). The spatial context is currently represented by a high-resolution 2D LULC map, given by a georeferenced image file that is loaded into the map widget using the GDAL library [53]. Each pixel of this image is assigned a LULC class, and represents a spatial domain of 1 m^2 [46]. The displayed map is automatically scaled to the bounding box of the entire data set, while a visual buffering offset is added to each side. The LULC classes are by default semantically color-coded to facilitate quick perception of the spatial context, but the user may specify an individually preferred color scheme on-the-fly. The colors of the background map are rendered in low saturation to visually emphasize the stacked wall.

Similar to the visualization design described in [154], I use a hybrid 2D/3D representation of the trajectory data considered in the given application scenario (Fig. 2.7). In accordance with the semantics inherent in mobile transect data, the third dimension can be used to encode the measurement height of the the multiple environmental attributes, stacked in order of the offset height

Feb 05, 2015   09:37 MST
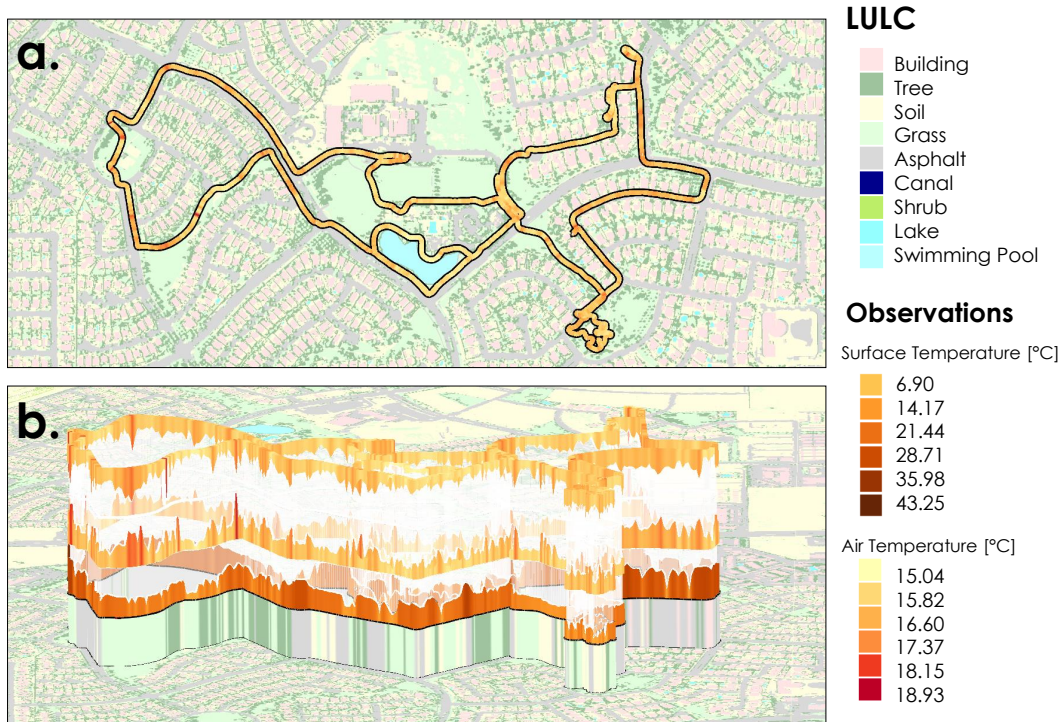
Figure 2.7: A wall as seen from (a) bird's view, and (b) from the side (facing north). Surface temperature, 1 m air temperature, and 2 m air temperature are stacked upon each other (from bottom to top). The lowest layer shows the LULC class of the particular route nodes. Note that some relationships between 1 m and 2 m air temperatures can be detected: The peaks in 1 m are more intensive than those in 2 m height (background map: [46]).

above ground of each sensor. For the sample data set described in Section 2.5, this results in the following order: The surface temperature is rendered in the lowest layer, followed by 1 m relative humidity, 1 m air temperature, 2 m relative humidity, and finally 2 m air temperature as top layer.

The user selects a single transect run from a drop-down menu to the left side of the widget that provides all transects in the database as options. Then, the user can dynamically select a combination of environmental attributes and their sequence of display in the wall. It is also possible to only examine the value distribution of a single variable. Further layers can subsequently be added to or removed from the visualization display, enabling the user to choose another stacking order than the semantically meaningful default described

above. User-defined adjacency of two specific layers in the wall is desirable to ease direct comparison of the two attributes they represent. Figure 2.8 shows
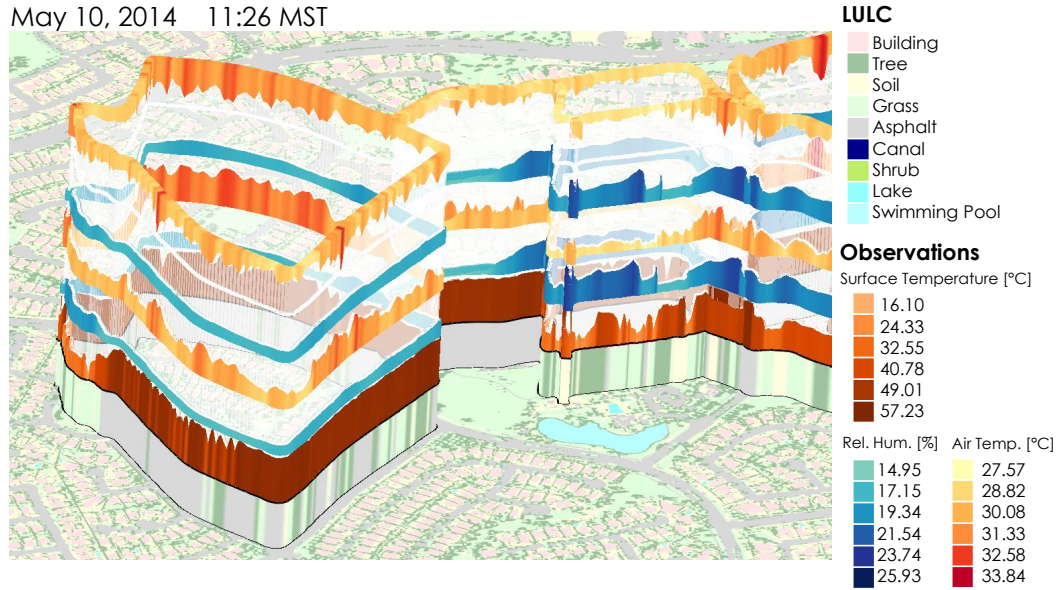
Figure 2.8:   A wall in close up view, showing a segment of a transect run included in the sample data set: surface temperature, 1 m relative humidity, 1 m air temperature, 2 m relative humidity, and 2 m air temperature (from bottom to top). The lowest layer shows the LULC class of the particular route nodes. Note the visual pattern of association between LULC, temperature, and humidity (background map: [46]).

the wall-design in a close-up view, including five environmental attributes in the sample data set. The attributes are visualized as ribbons stacked in the vertical wall that winds along the transect route. They are built around the underlying trajectories, which are created by connecting the points given by the GPS (the "nodes") with line segments. Using the concept of offset curves [77], I duplicate the trajectories on either side of the initial line to construct the bottom part of each wall layer. In contrast, the vertical thickness of an attribute's ribbon is not homogeneous over the course of the route, but is instead proportional to the measured values at each node. Between nodes the height is linearly interpolated to create a sense of spatial continuity.

A known drawback of 3D visualization is the difficulty to perceive and compare magnitudes [37]. Relying on the height of the wall-layers as the only indicator for the value of an attribute at a certain location is thus insufficient.

Therefore, the ribbons are additionally color-coded so that both the intensity of color and thickness of the ribbon are proportionate to the normalized value of the attribute. The color scale for the values depicted along the ribbons is predefined through a set of sequential colors in the ColorBrewer [67]. Lighter colors are associated with lower values and darker colors with higher values, following the cartographic convention of *"Dark equals more"* [67]. The user may specify an individually preferred color scheme for each ribbon. Per default, red is selected for temperature and blue for humidity attributes.

The semi-transparent surface between the ribbons serves three functions. First, it enhances the perception of the height and color of the individual ribbon. To ensure that this effect also applies to the uppermost ribbon, the top ribbon is inverted. Second, the surface visually connects individual wall-layers and enhances the perception of which points on the ribbons share a spatial location on the map. Third, the gaps that arise between the ribbons reduce occlusion because the semi-transparency provides a view behind the wall. However, a certain amount of occlusion is still possible, so the view can be rotated in three dimensions and translated or scaled to change the perspective.

An issue that cannot be solved using rotation, translation, and scaling is the occlusion of the immediate spatial context directly under the mobile platform transect by the wall itself. To avoid losing this important context information, a LULC class ribbon is added as the lowest layer of the wall, repeating the color-coding of the immediate spatial context.

In the prototype, it is only possible to stack five observed attributes upon each other, and these variables are based on the sample data set described in Section 2.5. This implementation was chosen to make variable-specific operations possible, which, e.g., include the combined normalization of air temperature and relative humidity measured at two different heights.

2.6.4   Data analysis

Analyzing the relationship between land use/land cover and micro-climate

Theoretical considerations

In micrometeorology, the field of view of an atmospheric sensor can be

described using the *source area* [136] concept. Sensor height a.g.l., wind speed and direction, and atmospheric factors determine the shape and position of the *source weight function*, which describes the relative contribution of the different upwind locations to the sampled value [157, 136]. The source area is the projection of the isopleths of this source weight function to a 2D plane [136].

For the surface temperature, which is in this case measured using an infrared radiometer, the source area is a circular area that is the field of view of the instrument [137]. The radius r of this disc can easily be determined using the following equation [137, 113]:

$$r = \frac{h}{\sqrt{\frac{1}{F} - 1}} \tag{2.1}$$

with h quantifying the mounting height of the radiometer in $[m]$. F is the dimensionless view factor that describes the ratio between a circular area on the ground and its surrounding annulus (limited by the horizon for a radiometer with a field of view of $180°$) [137]. In my implementation, F is constantly set to 0.95, which means that the elements located within the resulting disc contribute to 95 % of the measured surface temperature signal [113].

For radiometers with a smaller field of view, as is the case for the radiometer used to record the sample data set, the source area can also be determined using simple geometric principles [11]. The current version of TraVis uses Equation 2.1 to determine a radiometer's source area. For future work, however, it is planned to also include the more exact radiometer source area based on the field of view half angle input by a user based on the principles described in [11].

The computation of a source area for passive scalars such as air temperature is more complex, because the heat fluxes determining this quantity underly turbulent transport mechanisms [137]. These, in turn, are dependent on several atmospheric properties that are prevalent at the time the quantity was measured. The land surfaces impacting the air temperature are located upwind of the sensor. Thus, the orientation of the source area's semimajor axis generally varies with the predominant wind direction at the spatio-temporal location of sampling. However, other factors also affect the source area; this has been a research area over the last 30 years, applying methods such as Large
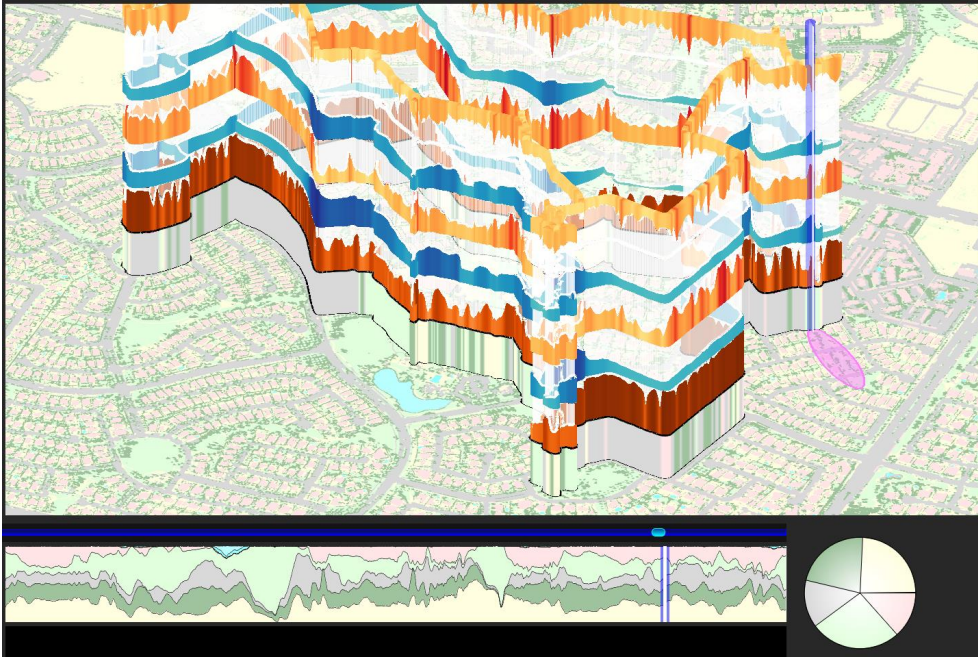
Figure 2.9: The slider can be used to interactively explore the spatial context of the data. The source area is attached to the bottom of the slider, while the fraction plot and the pie chart underneath the map display the composition of LULC fractions within the source area. The ellipsoid exemplifies the source area for air temperature measured at a height of 1 m (background map: [46]).

Eddy Simulation or Lagrangian stochastic particle dispersion models [157].

Since the source area needs to be estimated for each node on each trajectory, the estimation technique must be computationally affordable. Therefore, I rely on Schmid's simple parametric model to determine a source area's shape and size [136, 1]. Five parameters are required to compute the shape parameters for the resulting ellipsoid, with the orientation of the semimajor axis being parallel to the wind direction: measurement height, Obhukov length (a measure for the atmospheric stability), surface roughness length, friction velocity, and the standard deviation of lateral wind speeds. These parameters cannot be directly computed from the sample data set, but have to be determined using external information. For the prototype described in this dissertation, the orientation of the source area was estimated using averaged wind direction values retrieved from four weather stations surrounding the study area [105]. The other necessary parameters were selected such that the final ellipsoid ap-

proximates a realistic size.

The estimation of the source area for passive scalars is still ongoing work. TraVis currently only computes the source area for sensors mounted in $1\ m$ height for the case of a stable atmosphere. To retrieve more reliable results, it is necessary to add the capability to evaluate multiple types of spatial context such as the height map of the urban environment's features. Based on the resulting addition of geometrical information, the estimation of the source areas for passive scalars needs to be improved and evaluated thoroughly. While currently only wind direction data from surrounding weather stations is used to estimate the orientation of the source area, more data sources need to be integrated into the framework to retrieve information about the atmospheric stability and approximate turbulence characteristics at each measurement location. To refine the estimation of the passive scalar source area's shape, additional source area models can be added that incorporate height maps and canopy structure.

Visualization

Depicting all source areas for each point on the route simultaneously would result in numerous overlapping areas, making it difficult to distinguish the source areas of individual points on the route. In my framework, the source areas for distinct sensors can be interactively explored by means of a slider, which can be moved along the transect route. The user can choose which source area is displayed by using the radiobuttons in the graphical user interface (Fig. 2.5). The slider consists of two parts: The visualization of the source area at the bottom, and a semi-transparent cylindrical bar. The latter allows for the unambiguous attribution of the slider elements to a particular spatial location traversed by the mobile platform. To ensure that the bar does not occlude the measurement values at the selected location, it is rendered transparently. A thin white plane is added to the bar, which indicates the exact slider position (Fig. 2.9).

Two different measurement field of views are computed within TraVis as described above: The radiometer source area and the source area for passive scalar fluxes (e.g., air temperature). In the sample data set described in

Section 2.5, the radiometer is mounted at a height of approximately 0.3 m, resulting in a disc radius of 1.28 m. Consequently, the circle that represents this source area is so small that it disappears underneath the wall, whose width on the map is broader than this radius. Furthermore, depending on the predominant wind direction, the source area for passive scalars might overlap with the transect route, so that parts of the resulting ellipsoid are occluded by the wall. Nevertheless, the fractions of the LULC classes within the source area can be detected in the LULC composition display located underneath the map, and the adjacent pie-chart that provides an alternative representation of the fractions. The map, the fraction plot, and the pie-chart are linked: While the slider indicating the position on the transect is repeated in the fraction plot, the pie-chart adapts dynamically to the values associated with the slider's position on the map and in the fraction plot. To ensure a smooth visual transition of the pie chart between different segments, the LULC classes are always allocated in the same order on the pie chart.

The source area for passive scalar fluxes is constructed using the shape parameters given by Schmid's parametric model [136], which include the upwind distance from the sensor to the near end of the ellipsoid, the distance to the far end, the distance to the point where the ellipsoid has the maximum width, and the length of the corresponding semi-minor axis. I use rational cubic Beziér curves to create conic sections that are attached to each other to construct the egg-shaped function [48]. The resulting shape is rendered semitransparent, so that the LULC underneath this shape can still be recognized.

Segmenting a transect according to similar multivariate behavior

The functionality of the framework is complemented by a clustering interface, which can be used to classify the transect into segments of comparable multi-attribute patterns – similar to [160]. While the above described visualization and analysis capabilities facilitate the qualitative reasoning on the values and their interdependencies, the classification provides quantitative support.

The clustering is based on a combination of a self-organizing map (SOM) [89] and the k-means clustering algorithm. This combination has been demonstrated to reliably structure the results of the SOM into a lower number of
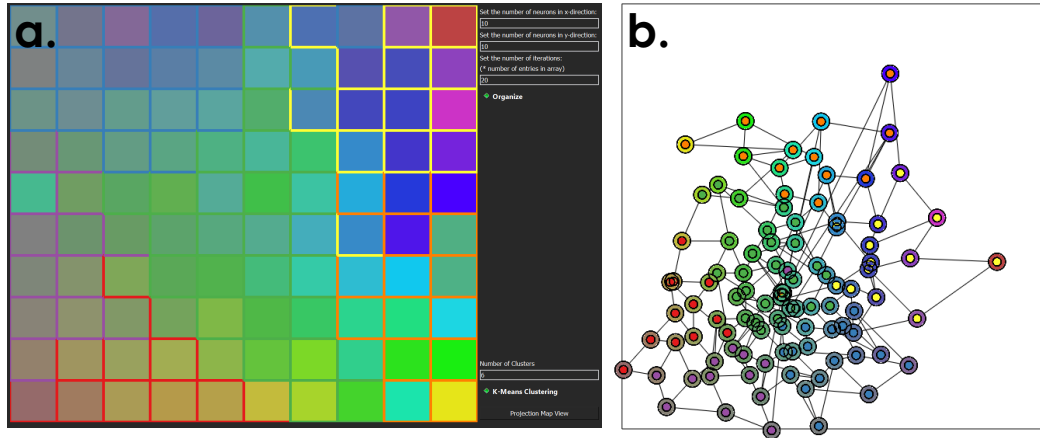
Figure 2.10: (a) The cluster interface, which is linked to the main window depicted in Figure 2.5. In the background, the weight vectors associated with the individual self-organizing map (SOM) nodes are visualized. They are colored using Sammon's mapping [128] of the weight vectors into a 2D color space [131, 9, 160]. The borders of the nodes represent each node's cluster membership after the k-means clustering has been applied. (b) The results of Sammon's mapping [128] of the neurons into a 2D space [131, 158]. The example is based on segmenting a mobile measurement run conducted on February 05, 2015, at 09:37 MST, based on surface temperature and air temperature in 1 m and 2 m height.

clusters, which is particularly useful if the number of nodes in the SOM is large [158]. Furthermore, it has also been successfully used in climate research, e.g. to visualize areas of similar temporal snowfall variation [161]. A drawback is the computation time, which does not allow for real-time interaction if a large number of SOM nodes and SOM training iterations is used. Another disadvantage of this approach lies in the dependency of the k-means clustering algorithm on the number and location of the initial cluster centroids [158].

My implementation of the SOM roughly follows the description provided in Kohonen's paper [89]. The weight vectors associated with the SOM's neurons are initialized randomly using the Boost.Random library (Boost 1.55) [100]. The input, based on which the SOM is trained, is a vector consisting of the

normalized values of a user-defined set of variables, ordered according to the sequence in which these variables are chosen. The variables are normalized over an entire transect run according to the variable-specific maximum and minimum, while variables describing the same physical property are normalized together to facilitate comparisons.

To train the SOM, a set of N (with $N =$ the number of observation points along the route) input vectors is randomly drawn from the set of input vectors during each epoch [89, 158]. For each input vector, the best-matching neuron (BMN) on the SOM is determined based on Euclidian distance [89]. Then, the weights of all neurons within a neighborhood of a radius r around the BMN are updated according to coefficients given by an exponential kernel function, which is centered around the BMN, which decreases with increasing distance, and which regulates the impact of the input vector on the weights of the surrounding neurons. Additionally, a linearly decreasing adaptation coefficient regulates this impact over time [158, 89]. r also decreases linearly over time, starting with the SOM's longest side as an initial value. The clustering procedure can be steered via an interface (Fig. 2.10a), on which several algorithm parameters can be determined by the user. First of all, the size of the SOM in x- and y-direction can be defined, as well as the number of epochs. After the algorithm has finished, the user can select the number of cluster centroids for the k-means clustering algorithm. While the initial positions of the cluster centroids are determined randomly, a user can refine the number of cluster centers after visually inspecting the results of the SOM [158], or by considering hypotheses about the expected number of different microenvironments along the route. The final weights SOM nodes are colored using Sammon's mapping [128] of the nodes into a two-dimensional color space to visualize the similarity of the nodes weight vectors, and hence, to support the task of visual inspection [131, 9, 160]. Additionally, the results of the mapping procedure can be investigated in an additional window [158, 131] (Fig. 2.10b).

The quality of the partition can be explored using a parallel coordinates plot in the main window (Fig. 2.5, [160, 57]), which shows the clustering results as applied to the selected transect run. In this plot, classes are color-coded with a qualitative color-scheme [67] and can be brushed. If the user is not satisfied with the current partition, the number of cluster centroids for the k-means

clustering can be interactively refined.

The user can select the attributes to be clustered, since clustering some combinations of attributes might not lead to meaningful results. For example, there is no physical relationship between the infrared radiometer's source area and the sampled air temperature because the thermometer employed to observe the latter has a different field of view.

On the main window, the results of the classification are visualized in a dedicated widget (Fig. 2.5d), which is visually connected to the slider described above, so that the link between cluster membership, spatial context, and attribute values at that point on the transect can be interactively explored. Additionally, the clustering results are added to the wall as topmost layer.

2.7   Use case

The visualization was tested using the mobile transect data set recorded on May 08, 2014, between 14:40 and 15:50 MST. In order to investigate whether the observations along the transect route can be classified into distinct segments of similar variable behavior, the clustering capability of the framework is used. In the illustrated case, the normalized values for air temperature at $1\ m$ and $2\ m$, relative humidity at $1\ m$ and $2\ m$, and surface temperature is utilized as input for the SOM. The SOM is setup with 1600 ($40\ x\ 40$) neurons, which are trained during $N * 20$ iterations (N is the number of nodes on the selected transect run). The mobile transect data set is then partitioned into 6 classes (Fig. 2.11a), which is also the number of the predominant LULC classes at the study site: soil, grass, tree, asphalt, building, and water (swimming pool and lake).

The classification of the points on the transect route can be inspected by stacking the cluster membership as the topmost layer upon the wall. The cluster membership can now be examined in a standard 2D spatial plot (Fig. 2.11b). Comparing the SOM / k-means results with the land use patches surrounding the corresponding transect node, a visual correlation between the traversed patch and the cluster membership can be observed. To investigate this further, the map can be rotated to visually explore the meaning of the different clusters (Fig. 2.11c/d), which can also be inspected using the parallel
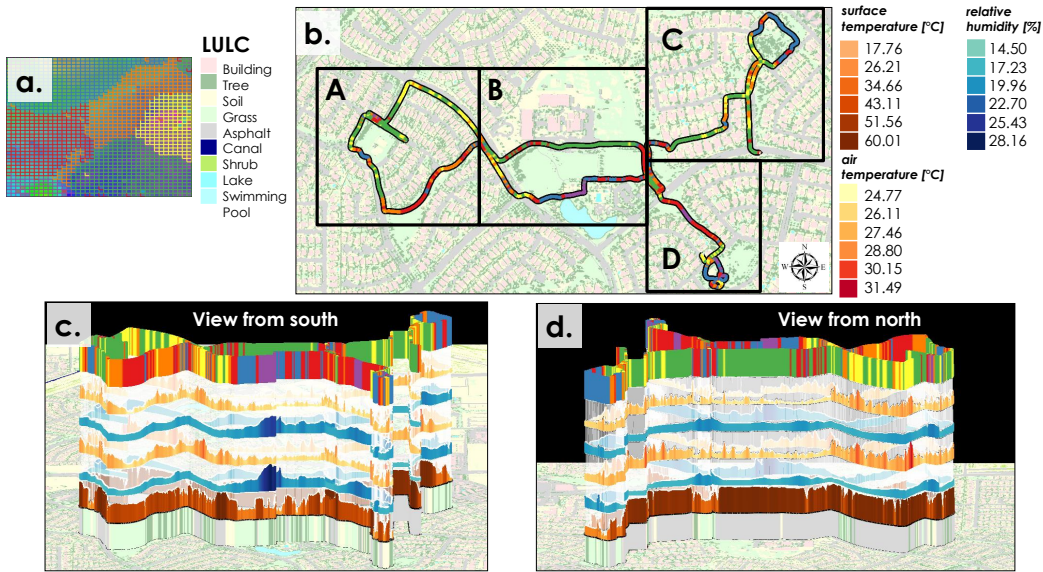
Figure 2.11: A use case for TraVis. In this example, the classification of a transect route into similar microenvironmental segments is demonstrated, based on a combined SOM and k-means clustering algorithm [158, 161] over all measured quantities. This clustering is based purely on associations between (in order of stacking from surface) surface temperature, air temperature, and relative humidity at 1 m and 2 m heights a.g.l.. (a) displays the visualization of the clustering output. (b) shows the distribution of the cluster members along the transect route from the birds-view perspective. Note the correlations between patches of land use and cluster membership. (c) and (d) visualize the meaning of the different clusters in terms of multivariate value distribution in the map, with (c) showing the wall from the south, (d) from the north (background map: [46]).

coordinates plot (Fig. 2.12). Since the rotation reveals the value distribution of all selected attributes, the clusters can be related to the multivariate value distribution shown on the stacked ribbons (also compare Fig. 2.12). Using this technique, it can be observed that the seemingly isolated segments belonging to cluster 2 near the lake in part B and in the south-eastern parks (part D). This cluster corresponds to exceptionally high humidity (Fig. 2.11c, Fig. 2.12), presumably owing to the lake's open water surface evaporation and increased evapotranspiration in the park in part D.
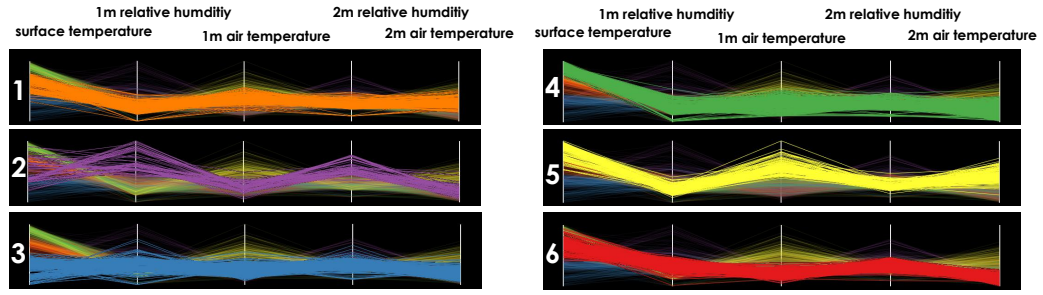
Figure 2.12: Semantical meaning of the clusters resulting from the segmentation described in Section 2.7.

Interestingly, park areas with grass LULC are associated with different microenvironmental clusters. Cluster 3 occurs near the lake in part B (Fig. 2.11b), in the north-eastern park (part C), in the park in the southern part of D, as well as during small segments in the parks in part A. It is associated with a typical value distribution for vegetated areas, especially in terms of the low surface temperature. However, park areas are not consistently represented by cluster 3. Thus, the northern part of the park in part B largely belongs to cluster 4, which is defined by high surface temperatures (Fig. 2.11b, d, Fig. 2.12). The same holds for the northern part of the park in the western part of A. The two park stretches in the southern part A and in the center of part D also show a different cluster distribution. This can be explained by a difference in relative humidity, which is higher in the park stretch in part D (purple part, Fig. 2.11c).

Cluster 5 most consistently occurs in the area between two arrays of buildings. This cluster is defined by high surface temperatures due to the asphalt cover, high air temperatures, and low relative humidity. There are, however, certain disruptions of this pattern, mostly by transect segments related to cluster 4. In these segments, the air temperature is lower, while surface temperature and relative humidity are similar to cluster 5.

2.8 Conclusion

A prototypic tool for the visualization and analysis of mobile environmental transect measurements is introduced based on contributions from the disci-

plines of visualization and urban climatology. Each measured attribute is visualized as a three-dimensional ribbon, whereas height and color emphasize the values measured at a particular location. Several ribbons can be vertically stacked in a wall to allow for the comparison of the value distribution between the attributes. A slider can be used to interactively investigate the composition of the estimated *source area* at each measurement location, while the LULC composition within this field of view is encoded by a pie-chart and a fraction plot. Complemented by a clustering capability, the framework combines analysis capabilities for the exploration of the spatial context and the multivariate relationships between the measured attributes.

The framework has been tested using data from a mobile transect measurement campaign in Gilbert, Arizona, USA. Multivariate relationships between measured attributes could easily be detected, physically meaningful clusters determined, and connections between land use and observations drawn.

The framework is planned to be continously extended and evaluated in the future, based on ideas arising through collaborations with end-users.

Chapter 3

Sensor Lag Correction for mobile air temperature measurements in an urban microclimate context

In this chapter, I describe an optimized procedure for sensor lag correction, which can be used to correct air temperature measurements taken with inert ("slow") sensors to resemble the measurements of faster sensors. The goal is to increase the spatial resolution of mobile measurements conducted with slower sensors in order to accurately sample the extent of, and transition between, urban microenvironments. While the correction procedure itself is based on previously known techniques, I determine three parameters, which are empirically optimized using 42 time series representing mobile measurements at two heights, different times of day, different seasons, and during different platform speeds. The optimization is conducted by comparing the agreement between corrected, "slow" air temperature measurements to simultaneously sampled "fast" air temperature measurements, which are assumed as a ground truth. The optimized correction procedure delivers robust results for nearly all investigated test cases, i.e. the alignment of the time series sampled by the "slow" sensor with the ground-truth could be improved.

This chapter is based on an article that I recently submitted to the journal "Urban Climate" [63].

3.1   Introduction

In all applications related to mobile measurements, the utilized sensors need to accurately sample the microclimate in a dynamically changing physical environment. The combined effects of sensor inertia, air speed, sampling frequency, and platform velocity determine the spatial resolution of a mobile transect data set [1].

Sensor inertia is dependent on the time constant τ_{63}, which is defined as the time a sensor needs to adapt to 63% of an impulse change [51, 1]. For air temperature sensors, this impulse is a change in the ambient temperature, which can occur frequently during a mobile transect run. With larger time constants, the temperature curve recorded by a sensor is smoothed because local minima and maxima cannot be resolved [1, 101]. Mayer et al. (2009) [101] and Foken (2006) [51] refer to this error as the *dynamical error*. Mayer et al. (2009) [101] also state that *lag-times* in some cases complement this error. Lag-times can cause a linear shift of the time series in addition to the dynamical error [101].

In this study, I apply the approach for sensor lag correction reported by Achberger and Bärring (1999) [1] to an urban microclimate setting, using moving average as a low-pass filter, and adding a linear shift of the corrected time series. The correction procedure is carried out for air temperature, which is one of the most important variables for urban heat island [147] and micrometeorology studies. Air temperatures at two different heights were measured using Resistance Temperature Detectors (RTD) and Fine-Wire Thermocouples (FWT) simultaneously while moving a a sensor platform through a residential area in the Phoenix Metropolitan area. The estimated physical microclimate scale along the transect route is 10 m at its minimum, based on the scale of variation of the urban canopy structure and land cover. This means, that the sampling frequency of the measurement system needs to be faster than 0.5 Hz for an average platform speed of 5 m/s to resolve the meteorological impact of these microenvironments. Although observations are logged with a frequency of 1 Hz and although the FWT sensors have a time constant below one second, the spatial resolution of the RTD is notably restricted by its time constant [1], which is 46.21 s at wind speeds approaching the platform speed. In simple

mathematical terms, this relationship can be written as

$$r^* = v\tau \tag{3.1}$$

where r^* is the spatial resolution $[m]$, v the platform speed $[ms^{-1}]$, and τ the time constant of the sensor $[s]$. The time constant is dependent on the ventilation of the sensor [44], which is, in the following, assumed to be equal to the average platform speed of 3.77 ms^{-1}. The ventilation of the sensor due to ambient wind speeds is neglected, since surrounding weather stations indicate that they are slower than the average platform velocity (Table 2.1).

The goal of the research described in this Chapter is to determine how measurements with a relatively slow sensor can be corrected in order to estimate high-resolution observations in an urban microscale environment. For the correction procedure, I determined three degrees of freedom: the size of the moving average window, the form of the time-derivative of the impulse response function, and the amount of linear shift. I empirically optimized the choice for these three parameters by comparing the agreement between the corrected RTD temperature to the FWT measurements, which I assumed as a ground truth. Based on the results, the implications for the spatial resolution of mobile measurements taken at different platform velocities are discussed.

The remainder of this chapter is structured as follows: Section 3.2 describes the related work of this study. Section 3.3 provides an overview of the methods used for this research. In Section 3.4, the results of the sensor lag correction procedure are evaluated, followed by a discussion about the spatial scale sensors can cover during mobile measurements. Finally, I draw conclusions in Section 3.5.

3.2 Related work

Sensor inertia is generally an important factor when a sensor is moved through space. Studies on sensor inertia have thus been carried out in the context of radiosonde or airborne temperature measurements, in which the sensor moves at very high air speed and velocity [102, 82, 99, 52]. Frequently, it is assumed that the sampled temperature is a convolution between the true temperature

and some transfer function, e.g. the time-derivative of the impulse response function using the time constants of the sensor and its housing [102, 82] or solely the time constant of the sensor [1]. The transfer function can then be interpreted as a filter for the true temperature, i.e. the convolution of the true temperature and the transfer function yields the measured temperature. Thus, *deconvolution* of the measured temperature results in the true temperature [102, 82, 52].

Studies on sensor lag correction in a mobile ground-based, micrometeorological context and the verification of the approaches from radiosonde or aircraft measurements for this setting are rare [101]. Mayer et al. [101] successfully demonstrate the applicability of the deconvolution approaches by Inverarity [82] and McCarthy [102] to vertical measurements conducted with an elevator tower. They found good agreement between the corrected elevator measurements and the stationary measurements conducted at different heights on a nearby mast. In addition, they developed a simpler correction scheme for two time constants that relies on the subsequent application of the weighted sensor response function for the time constants of the probe and its housing, achieving similar good results for their simpler correction procedure [101].

Hübner et al. [79] and Achberger and Bärring [1] investigate correction procedures for horizontally moving sensors during mobile transects in a micrometeorological context. Hübner et al. [79] installed a garden railway system perpendicular to a forest edge and examine how the correction procedure using a linear adjustment of the impulse response function can reproduce the step-change in air temperature at the transition between forest and clearing. Achberger and Bärring [1] use a correction filter based on the combination of a low-pass-filter and the time-derivative of the impulse response function, similar to McCarthy [102] and Inverarity [82]. To verify their approach, they compare the corrected air temperature to the values recorded by a nearby stationary sonic anemometer. They are able to bring the power spectra of the two air temperature time series into accordance, although the different spatial positions of the two sensors pose restrictions on the direct comparability of the two data sets.

3.3 Methodology

3.3.1 Determining the time constant of the HC2S3 Pt100 RTD

I followed the procedure described in [101] and [19] to determine the time constant for the HC2S3 Pt100 RTD. The ambient air temperature was measured using a reference sensor, while the target sensor was heated up by applying an external heat source. After the heat source was removed, the sensor was placed into a naturally ventilated 10-plate radiation shield to determine the time constant for the entire measurement system. Target sensor air temperatures were recorded at a frequency of 1 Hz, until they reached the value measured by the reference sensor. This procedure was carried out while exposing the RTD and the radiation shield to still air, as well as to wind speeds ranging from approx. 3.2 m/s up to approx. 7 m/s. During the duration of the experiments, the specific humidity of the surrounding air remained nearly constant, averaging at 8.39 g/kg. This value is approximately in the center of the humidity levels that were predominant during the May transect runs (on average 4.53 g/kg), during the September transect runs (on average 11.9 g/kg), and during the February transect runs (on average 6.64 g/kg). It can thus be assumed that the time constant is representative for the data sets under consideration in this study.

In accordance to [101], the time constant from the experimental data was computed by first determining the normalized response using the following expression:

$$T_{norm}(t) = \frac{|T_{target}(t) - T_{target}(0)|}{T_{target}(0) - T_{ref}} \tag{3.2}$$

where t is the time, $T_{target}(t)$ is the time series sampled by the heated up sensor, $T_{target}(0)$ is the maximum temperature of the target sensor at the beginning of each experiment, and T_{ref} is the reference temperature, measured by the reference sensor and averaged over the duration of each experiment. Equation 3.2 yields values around zero for $T_{norm}(0)$, while approching one for $T_{norm}(t_{max})$. To finally determine τ_{63}, the exponential impulse response function was fitted to the normalized response curve [101, 19]. The impulse

response function is given by [101, 51]

$$f(t) = (1 - e^{\frac{-t}{\tau}}) \tag{3.3}$$

Results implied that the time constant for the entire measurement system remains fairly constant for higher wind speeds ($> 3.2\ m/s$) and averages at $\tau_{63,mov} = 46.21\ s$ with a standard deviation of 3.7 s. For still air, the time constant is $\tau_{63,still} = 172.67\ s$ with a standard deviation of 0.08 s.

3.3.2 Correction strategy

Retrieving the true air temperature spectrum from measurements

The relationship between the true temperature and the temperature measured by a sensor can be described by the convolution integral

$$T_m(t) = \int_{-\infty}^{t} T_a(t')h(t')dt' \tag{3.4}$$

where T_m is the measured temperature and T_a the true temperature [1, 101, 102, 82]. h is a transfer function that expresses the sensor response to a unit impulse change and can be described using the time-derivative of the impulse response function, i.e.

$$h(t) = \frac{1}{\tau}e^{\frac{-t}{\tau}} \tag{3.5}$$

where t is the time in seconds [1] ([82] and [102] use the same equation with two time constants). Using the relationship described in Equation 3.4, the true temperature can be obtained by de-convoluting the measured temperature [1, 101, 102, 82]. McCarthy (1973) [102], Inverarity [82] and Mayer et al. [101] describe sensor lag correction procedures for measurement systems with two time constants. This is particularly relevant if the housing of the sensor introduces considerable time lags in addition to those of the sensor itself. In these cases, the transfer function consists of a linear combination of twice the expression in 3.5, one of which uses the time constant of the sensor, while the other one uses the time constant of the housing [101, 102, 82].

In this study, I adapt the approach described by Achberger and Bärring [1], who model the response of the sensor to an impulse-like temperature change

of 1°C using Equation 3.5. Since I determined the time constant of the entire measurement system, i.e. for the combined effects air temperature sensor and radiation shield, this approach is also appropriate for the data set considered in this research. Achberger and Bärring [1] use the Fourier transforms of the measured temperature series and Equation 3.5 to ease the convolution, so that Equation 3.4 becomes

$$G(f) = F(f)H(f) \tag{3.6}$$

where $G(f)$ is the Fourier transform of the measured air temperature time series, $F(f)$ the Fourier transform of the true temperature, and $H(f)$ the Fourier transform of Equation 3.5. The true temperature spectrum can now simply be obtained by solving Equation 3.6 for $F(f)$ [1].

The choice for τ in Equation 3.5 is one of the parameters, which I yield to optimize in the described approach. The most natural choice for a moving sensor platform would be $\tau_{63,mov}$, although possible factors lengthening the time constant, such as reduced ventilation while the platform stops, could motivate the selection of a larger value.

Low-pass filter and linear shift

The deconvolution procedure as described above also amplifies noise [1, 82]. Thus, it is important to use a low-pass filter to avoid this unwanted amplification. Achberger and Bärring (1999) [1] choose a cut-off frequency according to a change of slope in the power spectrum of the measured air temperature, since higher frequencies can be interpreted as noise ([1] after Kaimal and Finnigan). However, they do not specify which particular low-pass filter they use. Inverarity [82] apply a single-pole recursive filter to their data.

Similarly to McCarthy [102], but in another context, the goal is to leave major trends in the data as they are, while filtering out short-term temperature fluctuations. Thus, I follow his approach and smooth the RTD time series by utilizing a simple two-sided unweighted moving average as low-pass filter

as given by

$$
T_{m,filt}(t) = \begin{cases} \dfrac{1}{l_{win}} \displaystyle\sum_{i=t-\frac{l_{win}-1}{2}}^{t+\frac{l_{win}-1}{2}} T_m(i) & \text{for } l_{win} \bmod 2 = 1 \\[4ex] \dfrac{1}{l_{win}} \displaystyle\sum_{i=t-(\frac{l_{win}}{2}-1)}^{t+\frac{l_{win}}{2}} T_m(i) & \text{otherwise} \end{cases} \tag{3.7}
$$

and as implemented in the R "zoo" package ([175], function "rollmean"; note that even numbers are also possible as input). In signal processing terms, the moving average filter does not cut-off higher frequencies, but continuously attenuates their power up to the Nyquist frequency (Figure 3.1b).

The size of the moving average window is another parameter that I attempt to optimize in this study. This parameter is dependent on the amount of noise that is apparent in the data, which is in turn dependent on the sensor, the platform speed, and the physical microclimate variation experienced by the sensor. During the measurements, the platform moved at a maximal average speed of 5.23 m/s (averaged over all speeds of a transect run, excluding the instances when the cart stopped) over potentially climatologically relevant patches with an extent of minimally 10 m. Following Equation 3.1, there cannot be any physically motivated variation occuring at a time interval that is smaller than approx. 2 s, which I thus chose as a lower bound for the moving average window size. The upper bound can be derived in analogy to the lower bound: the sensor is moved with a minimal average speed of 2.41 m/s (averaged over all speeds of a transect run, excluding the instances when the cart stopped) over patches with an extend of maximal 400 m, leading to an upper bound of approx. 166 s. An intuitive and theoretically natural choice for the moving average window size would be dependent on $\tau_{63,mov}$ because this value takes the sensor's spatial resolution into account.

For the FWT sensor, I smoothed the data using a moving average window size of 2 s because of its time constant of $\tau_{63} = 0.525$ s ($\tau_{95} = 3\tau_{63} = 1.575$ s). Following Equation 3.1, the low τ_{63} implies that the sensor is theoretically able to resolve spatial air temperature variations of all patches with a spatial extend larger than 2.75 m (8.24 m to resolve 95% of the resulting impulse change), while moving at a speed of 5.23 m/s. Since the smallest patch has an extent of
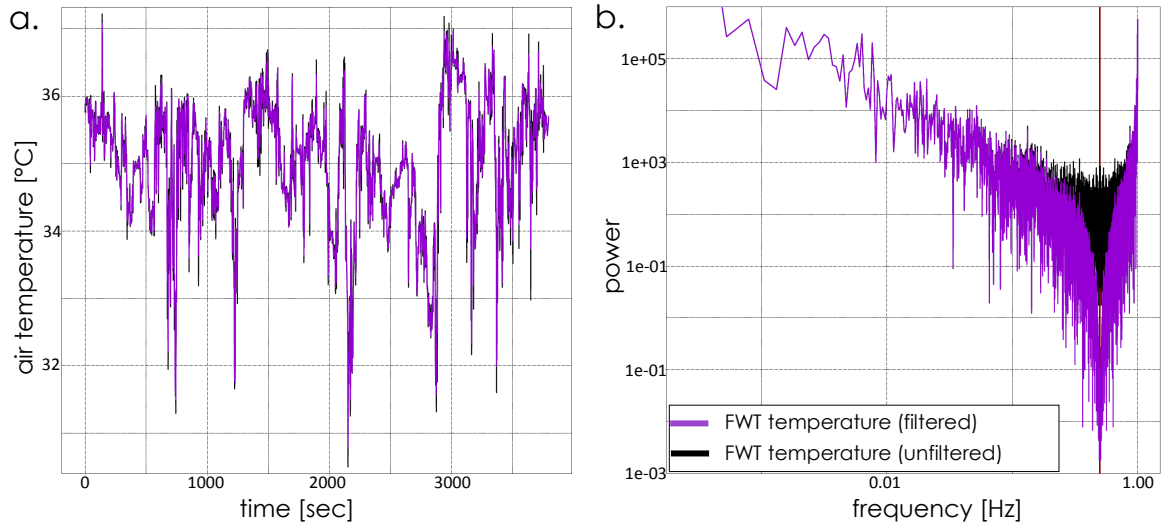
Figure 3.1: Time-detrended 1m FWT air temperature measurements (filtered and unfiltered) in the time and frequency domain. The time series was recorded on September 15, 2014, between 18:33 MST and 19:36 MST. (a) shows the time series, before and after the moving average has been applied. (b) displays these time series in the frequency domain. The red line indicates the Nyquist frequency.

10 m, large FWT-measured air temperature variation occurring within a time interval of less than 2 s is improbable.

The application of the deconvolution described in Section 3.3.2 on the filtered RTD temperature leads to results, in which major features of the variability of the "ground-truth" FWT sensor can be reproduced. However, in most cases, there is a linear shift between the corrected RTD and the raw (but time-detrended) FWT time series. This linear shift causes the local maxima and minima in the corrected RTD time series to appear later than in the FWT series (*lag-time* [101]). Choosing a multiple of $\tau_{63,mov}$ can attenuate this effect, but on the other hand, it can also intensify the amplitude between local minima and maxima in the time series, causing larger-than-necessary deviations from the "ground-truth". Therefore, the amount of linear shift to balance out an optimal choice for τ in Eq. 3.5 is the third parameter that I attempt to optimize in this study.

Optimizing the parameter choice

Three parameters were optimized in this study: the choice for the moving average window size, the choice for τ in Equation 3.5, and the choice for the amount of linear shift. For each of these parameters, I iterated through all possible combinations of the elements included in the following sets (all values are based on $\tau_{63,mov}$):

- Moving average window size: $\{2, 11, 23, 34, 46, 57, 69, 92, 115, 138, 161, 166\}$

- τ in Equation 3.5: $\{0.5\tau_{63}, \tau_{63}, \tau_{86}, \tau_{95}, \tau_{99}\}$

- Amount of linear shift: $\{n \in \mathbb{N}^0 | n \leq 46\}$

I applied each of these parameter combinations to correct the HC2S3 Pt100 RTD measurements according to the procedure described in the previous Sections. For each correction result, I then computed four error measures between the corrected RTD time series and the FWT measurements, whereas the latter have been filtered using a moving average window size of 2 s. In particular, I computed the mean absolute error (MAE), the maximal absolute error (AE_{max}), the root mean squared error ($RMSE$) and the index of agreement d as described by Willmott [170, 169]. The index of agreement is a dimensionless and easily interpretable value that indicates the accurateness of a prediction when compared to an observation. An index of agreement of $d = 1$ denotes perfect prediction of the observed values by a model, while $d = 0$ means disagreement [169]. The best combination of values should minimize MAE, AE_{max}, and $RMSE$, while d should be maximized. To optimize across all data sets, I averaged each error measure over all data sets per parameter configuration (Fig. 3.2). Then, I sorted the resulting table according to the error measures. While the optimal parameter configuration should maximize d, the other error measures ($RMSE$, MAE, and AE_{max}) should be minimized. Therefore, I computed an aggregate attribute by multiplying the error measures as follows:

$$\Delta_{aggr} = (1 - d) \cdot RMSE \cdot MAE \cdot AE_{max} \tag{3.8}$$

The table is sorted according to Δ_{aggr}, and the first entry in the sorted table is assumed to be the best possible value combination.
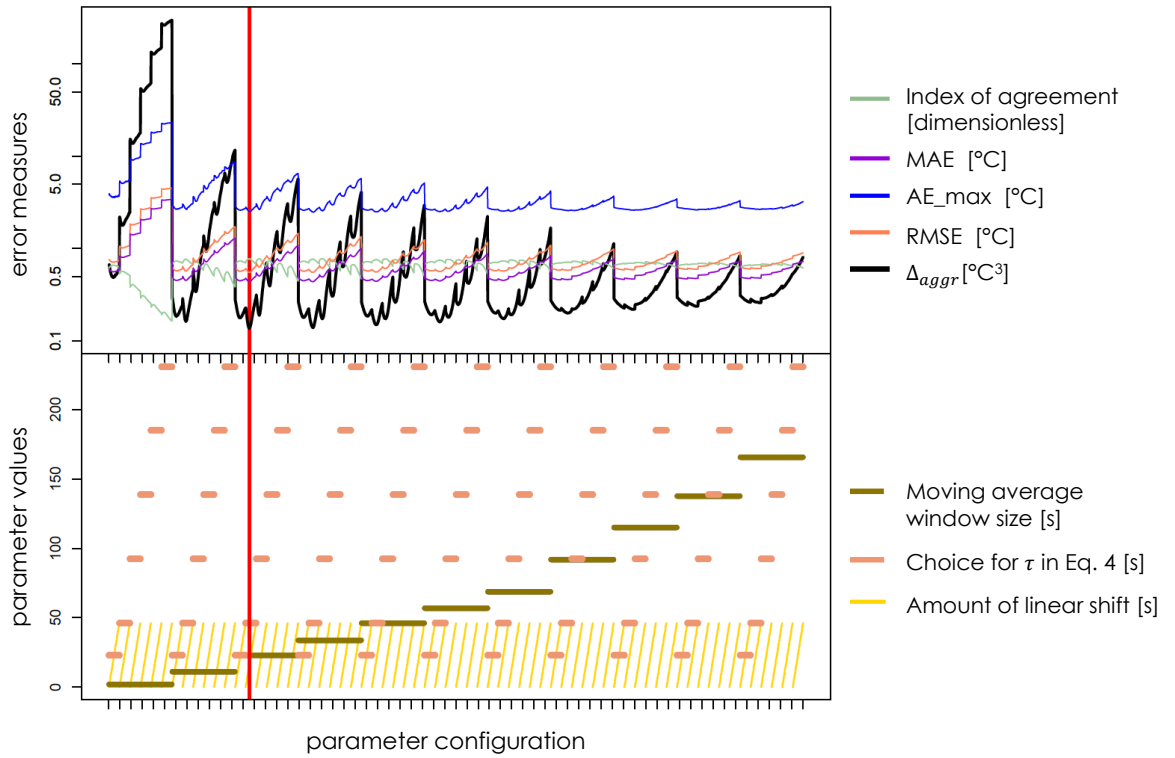
Figure 3.2: *Upper image*: Results for the parameter optimization experiment, averaged over all available data sets. The results correspond to the parameter configurations shown on the *lower image*: All possible combinations of varying values for the three to-be-optimized parameters have been examined. The red line indicates the best result in terms of Δ_{aggr}.

Final correction procedure

The final method follows six steps (similar to [1, 102, 82]):

1. Application of a two-sided unweighted moving average filter with a total window size of $0.5 \cdot \tau_{63,mov} = 23$ to the raw, but time-detrended RTD data

2. Setup of the correction filter resulting from Equation 3.5 using $\tau_{63,mov} = 46.21$ s, starting at $t = 0$.

3. Fast Fourier Transform of both Equation 3.5 (yielding $H(f)$) and the filtered RTD data (yielding $G(f)$)

4. Division of $G(f)$ by $H(f)$ to retrieve the true temperature spectrum

5. Inverse Fourier transform of the true temperature spectrum, yielding the corrected RTD data without shift

6. Linear time shift of the corrected RTD data by $0.7 \cdot \tau_{63,mov} \approx 16$ s backwards in time

For steps 1 and 6 the input parameter needs to be an integer. For these steps, $\tau_{63,mov}$ was rounded to the closest integer. The algorithm was implemented using R 3.1.1 [121].

3.4   Results and discussion

3.4.1   Validation of the correction procedure

To verify my results, the corrected RTD temperature time series were compared to the respective time-detrended and filtered FWT time series sampled at the same height. The FWT measurements were used as a ground-truth because of their small time constant. Figure 3.3 shows the verification results for a total of 42 test cases consisting of $1\ m$ and $2\ m$ air temperature data for all 21 runs. In nearly all cases, the index of agreement d between the corrected HC2S3 Pt100 RTD time series and the ground-truth $d_{CORR,FWT}$ exceeds the index of agreement between the uncorrected HC2S3 Pt100 RTD time series and the ground-truth $d_{UNCORR,FWT}$ by at least 0.09, thus indicating an improvement. However, in two cases (February 05, 2015, 07:12 MST, $1\ m$ and $2\ m$ height, Fig. 3.3, bottom), the correction procedure worsens the agreement between HC2S3 Pt100 RTD measurements and the ground-truth by 0.006 on average, which means that the uncorrected HC2S3 Pt100 RTD measurements have already been in line with the ground-truth (Fig. 3.4b). Therefore, results imply that the correction procedure only marginally affects the air temperature time series in cases where the error between measurement and ground-truth is small. The RMSE between ground-truth and RTD measurements ranges between 0.25 °C and 1.35 °C, while staying below 1 °C for all cases except for the time series recorded at $1\ m$ height on February 05, 2015, 12:25 MST. The same holds for the mean absolute error, which ranges between 0.18 °C and 1.19 °C. d is also lowest for the time series recorded at $1\ m$ height on February
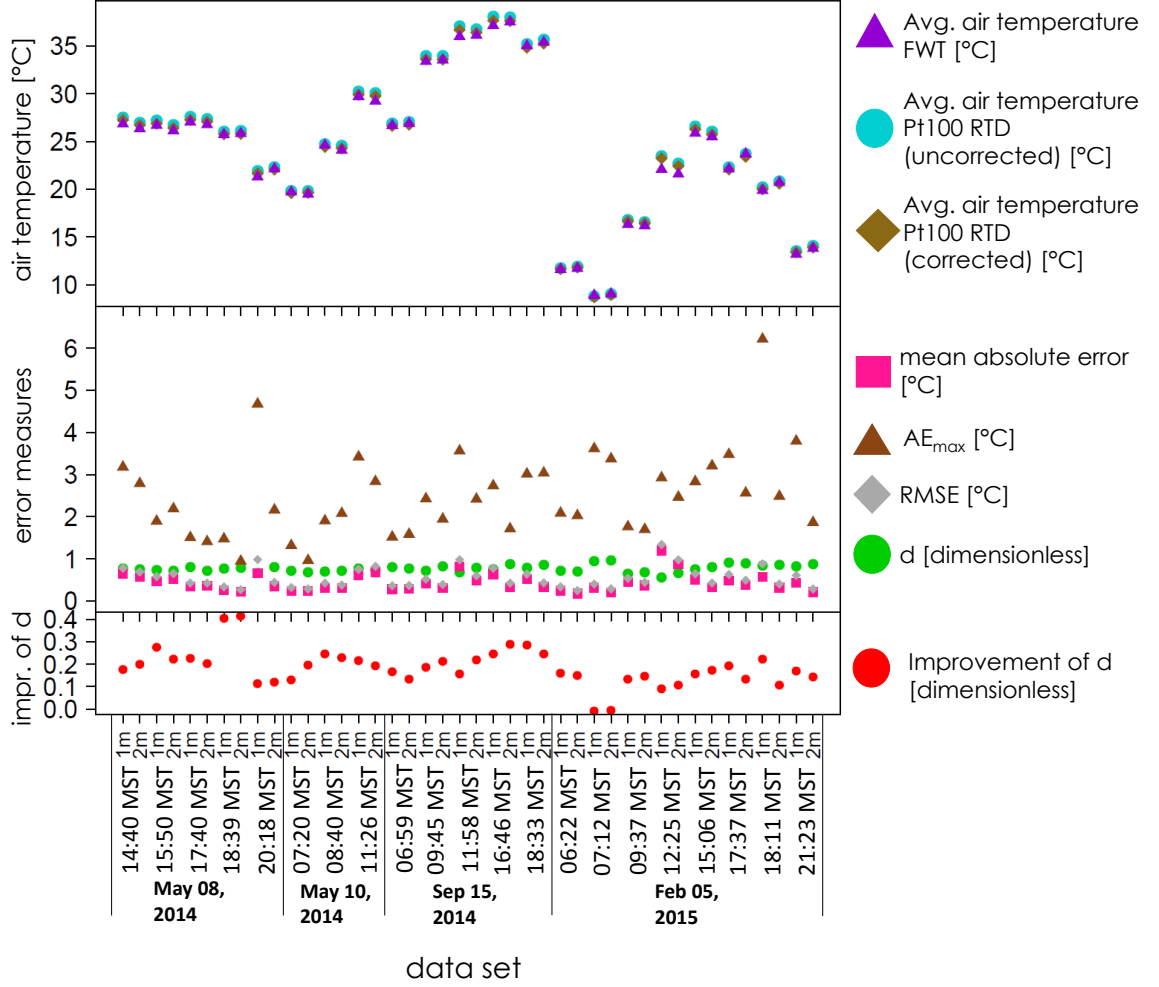
Figure 3.3: Data sets, their summaries and error analysis. *Top*: Average air temperatures per data set for each time series (filtered FWT, uncorrected HC2S3 Pt100 RTD measurements, and corrected HC2S3 Pt100 RTD measurements). All time series were time-detrended prior to correction. *Center*: Error between filtered FWT measurements ("ground-truth") and corrected HC2S3 Pt100 RTD measurements. *Bottom*: The improvement of the index of agreement d, which is determined by subtracting $d_{UNCORR,FWT}$ from $d_{CORR,FWT}$.
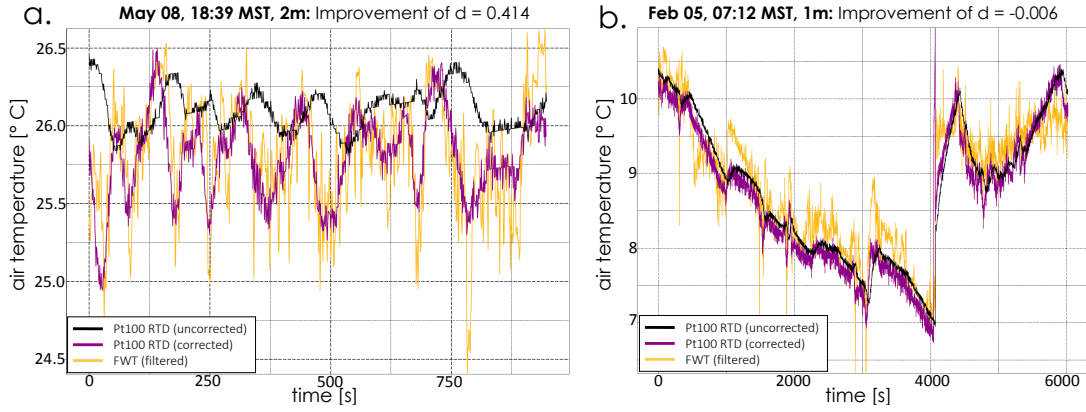
Figure 3.4: (a) Best and (b) worst correction result from a set of 42 test cases in terms of the improvement of d. The best and worst cases are determined by the difference of the index of agreement between measured and uncorrected data and d between measured and corrected data ($d_{CORR,FWT} - d_{UNCORR,FWT}$). All time series were time-detrended prior to correction.

05, 2015, 12:25 MST, indicating an agreement of 0.56 between corrected RTD measurements and ground-truth. Although these error measures do not imply a good result for this particular time series, the agreement between ground-truth and HC2S3 Pt100 RTD is still improved by the correction. On February 05, 18:11 MST, the corrected 1 m air temperatures show the highest deviation (6.21 °C) between the corrected RTD temperature and the thermocouple readings in terms of the maximal absolute error (Fig. 3.3, center plot). However, RMSE and MAE stay below 1 °C, with d adding up to 0.84 for this data set. The high AE_{max} value is thus caused by fast fluctuations in the FWT temperature time series, which cannot be reconstructed due to the difference between the time constants of the RTDs and the FWTs, respectively.

The improvement of agreement can also be observed on the scatterplots shown in Figure 3.5, where the corrected HC2S3 Pt100 RTD time series align better with the 1:1 slope line, which would represent a perfect match to the FWT time series. This is also reflected in the overall RMSE, which is lower between FWT time series and corrected HC2S3 Pt100 RTD data. In general, the corrected HC2S3 Pt100 RTD measurements show larger local amplitudes, and the local maxima and minima are time-shifted to resemble the FWT time series (Figure 3.4a). The average air temperature is altered only marginally

by correction (Fig. 3.3, top), although the values measured by the FWT sensor are slightly lower than the corrected and uncorrected HC2S3 Pt100 RTD measurements.

In general, the algorithm performs in a robust way, even though the time
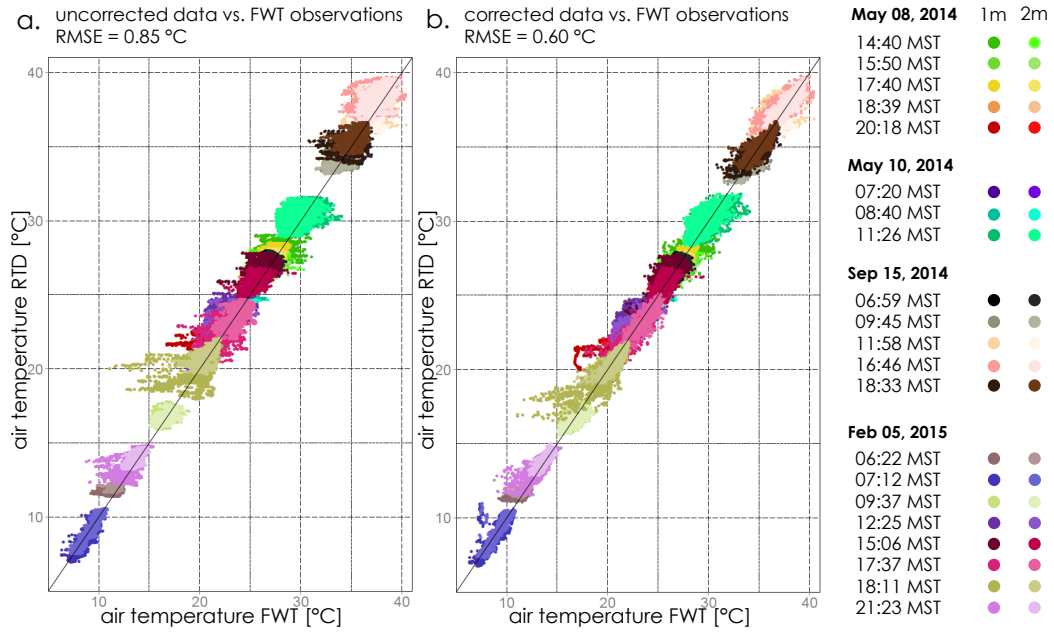
Figure 3.5: Correction results. The scatter plots display the alignment of the time-detrended and filtered FWT observations to the (a) uncorrected (but time-detrended) and (b) corrected RTD observations. The 1:1 slope lines shown on both scatterplots indicate perfect alignment. Different colors correspond to different time series, in correspondence to Figure 2.4.

constant was kept unchanged during all computations. This is an approximation, since the time constant is known to increase with decreasing ventilation, which occurs during decreasing platform velocities (e.g., Downs et al. [44], who determined the time constant for a HC2S3 Pt100 RTD for different air speeds). For the majority of time series available in the sample data set, the sensor platform stopped for smaller time intervals, which should theoretically lead to an increased τ_{63}. Despite this inhomogeneous velocity distribution, the correction procedure leads to reliable results.

3.4.2   Limitations and further evaluation

There are limitations to the generalizability of the correction procedure. Although I achieved stable results for 42 time series representing measurements at two heights, different times of day, different seasons, and during different platform speeds, I was only able to quantitatively validate the approach for one sensor setup.

Furthermore, the FWT readings, which I used as ground-truth in this study, do not represent the true temperature of the traversed environments because their response time is affected by the thermocouple time constant. However, with a time constant τ_{63} of approximately 0.6 sec in still air, the FWT readings should theoretically represent the surrounding microclimate with a fairly small spatial resolution finer than that represented by this microenvironment.

For further evaluation, but without quantitative validation, I also tested the correction procedure for a time series sampled by another air temperature sensor: the Campbell Scientific HMP45C combined temperature and relative humidity probe. The time series for which the algorithm was tested, was recorded on June 29, 2011, 11:11 MST, and was time-detrended as described in Section 2.4. During this transect run, the average speed of the sensor platform was 3.49 m/s (averaged over all instances with a speed $>$ 0 m/s, which was the case for 54.09 % of all samples).

The HMP45C includes a platinum-type RTD with a resistance of 1000 Ω (Pt1000 RTD), and an accuracy of 0.2 °C at 20 °C [25]. I experimentally determined τ_{63} for the HMP45C Pt1000 RTD using the approach described in Section 2.3, although we additionally conducted experiments to measure the time the *cooled* sensor needed to adapt to surrounding air temperatures. Postprocessing revealed a time constant of 128.15 s for wind speeds $>$ 2.45 m/s. Based on these values, the correction procedure for the HMP45C yields a parameter setup with a moving average window size of $0.5\tau_{63,mov} \approx 64$s, a τ of $\tau_{63,mov} = 128.15$ in Eq. 3.5, and a linear time shift by $0.7\tau_{63,mov} \approx 90$s. It could be observed that the local maximum and minimum air temperatures are amplified, which is one of the goals of the correction procedure.

From Table 3.1, uncorrected "slow" sensor observations, such as the HC2S3 Pt100 RTD or HMP45C Pt1000 RTD are not able to capture microclimate variability at the typical built environment spatial scales which are on the order of

10 m. This is true even at relatively low platform velocities (or wind speeds) of 1 m/s. A "slow" sensor mounted on a near-stationary platform under "calm" wind conditions of roughly 0.1 m/s can sample the 10 m scale microclimate, but this defeats the intent of a mobile and spatially distributed observation strategy. However, fine-wire thermocouples, or other sensors with a factory rated response time τ_{63} faster than 1 Hz, can adequately sample microclimate variations at 10 m spatial scales even for platform velocities approaching typical urban vehicle traffic velocities of approximately 10 m/s. As demonstrated

v [m/s]	r^*: Effective Spatial Resolution [m], $\quad r^* = v\tau$					
	Campbell Scientific FW3 Thermocouples *		Rotronic HC2S3 (Pt100 RTD) **		Campbell Scientific HMP45C (Pt1000 RTD) ***	
	$\tau_{63} = 0.6s$	$\tau_{95} = 1.8s$	$\tau_{63} = 46.21s$	$\tau_{95} = 138.63s$	$\tau_{63} = 128.15s$	$\tau_{95} = 384.44s$
0.1	0.06	0.18	**4.62**	**13.86**	**12.82**	**38.44**
1	0.6	1.8	**46.21**	**138.63**	**128.15**	**384.44**
5	3.0	9.0	231.05	693.15	640.75	1922.2
10	6.0	18.0	462.10	1386.30	1281.50	3844.40

*   typical τ_{63} for this sensor (not experimentally determined)
** τ_{63} experimentally determined using airspeeds greater than 3.2 m/s
*** τ_{63} experimentally determined using airspeeds greater than 2.45 m/s
Bold: because v is lower than the sensor's rated airspeed, τ and therefore also r^* is underestimated

Table 3.1: Comparison of the spatial resolution at three different movement speeds, given by the time constants of different sensors.

in Section 3.4.1, the correction procedure described in this Chapter can help to approximate observations with a high spatial resolution (as provided by the FWT), even though the original data was recorded with a slower sensor (in this case the HC2S3 Pt100 RTD). An argument for a simple and optimized method for this correction were presented, and the optimal parameters of correction were laid out in the previous Sections. It has to be noted, however, that this approximation can only succeed, if the slower sensor captured a signal from

the microclimate patch it traversed.

3.5 Conclusion

In the research described in this chapter, I experimentally optimized a set of three parameters used in a procedure to correct air temperatures recorded with "slow" sensors for sensor inertia: Moving average window size, transfer function setup, and linear time shift. The correction procedure was mainly based on the algorithm described by Achberger and Bärring [1]. The linear time shift was added to the procedure as a final step to balance the effect of the transfer function on the intensification of local air temperature maxima and minima. The correction was carried out with a set of parameter configurations on time series measured with a HC2S3 Pt100 RTD. The optimal setting was determined by evaluating the error between the corrected time series and "ground-truth" FWT observations, taken simultaneously and at the same location.

Although the time constant was assumed to be constant during all platform speeds, the optimized algorithm delivered robust results for nearly all of the investigated test cases. Thus, the corrections enhanced nearly all time series measured with slow sensors in terms of their alignment with a ground-truth. However, the correction procedure only works in cases, in which the slower sensor recorded a signal representing a microscale land use patch. This capability is limited by the time constant of a sensor.

Although the sample data set covers different times of day, different seasons, and different measurement heights, I could only quantify results for one particular sensor setup. Therefore, further research is necessary to validate the suggested approach for other measurement settings, including other sensors.

The significance of this work's result is that measurements from "slow" sensors can be corrected to more closely resemble the measurements of faster sensors, and to approximate the effective spatial resolution of faster sensors. In general, the performance of a sensor in terms of spatial resolution is an important consideration when observing the microclimate within an urban area or an otherwise spatially complex terrain.

Chapter 4

Visualization of Climatic Microenvironments based on the Spatial Aggregation of a Set of Mobile Transect Measurements

In this chapter, I describe an approach to aggregate mobile measurements taken over diverse routes without losing important information about potential uncertainties resulting from a small set of samples corresponding to certain locations within a study site. The presented approach extends currently used spatial aggregation techniques for mobile transect measurements by visually encoding the number of transects traversing a location, the predominant wind direction while conducting each transect, and the aggregated multivariate relationships at this location. This visualization can give insight into the uncertainty related to an aggregated value, and it facilitates the relation of the aggregated sample to upwind LULC. Furthermore, an automated approach for the comparison of the aggregated data over the entire study site is described, which facilitates identifying complex patterns in the data and, as a result, to visually delineate potential climatic microenvironments.

This chapter resembles and extends a peer-reviewed short paper that was published in the proceedings of the "Workshop on Visualization in Environ-

mental Sciences (EnvirVis) 2015 [61].

4.1   Introduction

Mobile transect measurements are frequently used in urban climatology to gain insight into the spatial variation of atmospheric properties. Based on the resulting observations, conclusions can be drawn about the impact of urban form on the surrounding climate [33, 73, 150]. However, a single transect is not sufficient to generalize such findings because it only represents a single meteorological background condition. Mobile measurements have to be repeated, so that the samples are representative for a given area in a variety of situations [24]. This results in a spatially dependent, multivariate, time-varying data set, which is difficult to analyze.

Spatial aggregation techniques can be used to reduce the complexity of a mobile measurement data set and to facilitate reasoning about coherent segments with comparable atmospheric conditions. This is only possible if the observation routes are spatially identical, or if at least a sufficiently large number of measurements is available per spatial compartment [70]. To guarantee for this, mobile transect measurements are frequently repeated over the same set of routes. For larger campaigns using public transportation vehicles (e.g., [24, 70]), this is necessarily the case. However, also mobile measurements conducted with more flexible sensor transportation modes attempt to repeat their initial traverses [45, 95]. Then, a variable of interest can be aggregated over a predefined set of waypoints [24, 45], road segments [95], or a regular grid [70].

But this ideal case is not given for every application of mobile measurements. An example are data crowdsourcing scenarios, in which the routes are dependent on the individual movement of the participants. Examples for the latter case can be found in [69] and [168] for air quality measurements. In [69], the authors map their data using average values over a regular grid, regardless of the number of measurements taken in one of the compartments. In [168], a web-based tool was developed to inform the general public about individual measurement results along a route or in an area. In both cases, only one variable is mapped at a time. But also in specialized and structured mobile measurement campaigns varying routes are taken, either to research different

climatic phenomena, or to simply increase the spatial coverage of the collected data, so that a larger number of potential climatic microenvironments are represented in the samples. Both is the case for the sample data set used for the development of TraVis, which was collected by Ariane Middel, Benjamin L. Ruddell, and their team.

The approach described in this Chapter is based on the traditional approach of spatially aggregating multiple transect runs over a regular grid. Multivariate relationships are classified separately for each transect run using a combination of the self-organizing map (SOM) [89] and the k-means clustering algorithm [161, 158]. Thus, per grid cell, a summary about the multivariate behavior of all observations taken within this spatial compartment is gained. To visually encode this information, I designed a glyph that represents the number of transect routes crossing the grid cell, their respective cluster membership at this location, and the wind direction that has been predominant during data collection in the field. The approach is implemented into TraVis [60], whose core functionalities have already been described in Chapter 2.

This chapter is structured as follows: First, related work is described in Section 4.2. Then, the user-steered aggregation techniques are introduced in Section 4.3. Section 4.4 provides a validation of the applied approach, while a use case demonstrates its functionality in Section 4.5. A conclusion is drawn in Section 4.6.

4.2 Related work

Andrienko and Andrienko [4] describe a general framework for the aggregation of movement data sets. They distinguish two different views on movement: The *trajectory-oriented view*, which focuses on the movement of single entities, and the *situation-oriented view*, which focuses on the state of the entities at one or more points in time. For each of these two views, they give recommendations about applicable aggregation techniques, which are again based on two different views of space: The *space-centered view*, focusing on the space in which the movement occurs, and the *entity-centered view*, focusing on the movement itself. According to this theoretical framework, the analysis pur-

pose followed in this research corresponds to the category *space-centered* and *situation-oriented*, since the relationship between trajectory attributes at certain points in time and space is of predominant interest. Thus, following [4] and in line with the research goal of this project, each mobile transect measurement run is treated as a sequence of independent measurement events.

The traffic-oriented view on a car data set in Milano [3] is closely related to my visualization approach. The authors suggest to partition the space using a regular grid and to aggregate the data over each grid cell. The aggregation results are visualized using either small multiples showing the frequency of car traverses per grid cell and time step, or – similar to my solution – using radial glyphs that encode the traffic intensity per movement direction. Scheepens *et al.* [134, 133] combine multiple density maps to provide a comprehensive overview over multivariate movement data sets. To enhance this visualization, they add pie-charts as glyphs at certain locations, which summarize the attributes used to create the underlying density maps [135]. In a later study [132], they extended these pie charts to summarize local density, static attributes of moving objects, their heading, and the proportion of stationary objects within a spatial compartment – all related to a certain point in time. Thus, again related to the work presented in this chapter, their radial glyph displays the density as well as related attributes of objects traversing a region. Andrienko and Andrienko [7] aggregate trajectory data over data-inherent compartments, which they extract from massive movement data sets by clustering locations of characteristic movement. Similar to my approach, they also provide the opportunity to adapt the aggregation scale. In a follow-up study, Andrienko *et al.* [5] extend this work to allow for complex, overlapping spatial compartments, as well as for a more flexible cluster input to extract spatial events. The spatial events lead to so-called relevant places, which are determined using a metric that combines a user-defined set of attributes. Sun *et al.* [151] spatially aggregate taxi trips per road segment and visualize the temporal variability of analyzed attributes as temporal plots within their spatial context.

In some cases, the view on trajectories is already given by the nature of the visualized data. This is the case for episodic movement data sets, where spatial positions are only logged at certain locations or at certain points in time and information about continuous trajectories is lost [9]. Visualization

and exploration techniques applied to this special kind of movement data sets frequently analyze spatial situations, either at certain points in time or over an entire observation period. In this context, Andrienko *et al.* [9] classify spatial situations by clustering feature vectors representing presence counts per sample location, and additionally visualize the flow between these points. Wang *et al.* provide a tool to explore traffic volume at and between transportation cells [165]. Bak *et al.* [17] aggregate their episodic movement data set using *growth-ring maps.*

4.3   Implementation

4.3.1   Workflow overview

A natural approach to spatially aggregate a set of trajectories is to partition the space into compartments and to summarize data collected for each of these compartments [3, 132]. Since the transect routes vary, the number of mobile measurement runs traversing a grid cell also varies. Thus, an appropriate aggregation and visualization technique has to take this asymmetry into account.

I developed a user-steered workflow that attempts to solve the issue described above. The workflow leads to a visualization that supports the identification of climatic microenvironments, while still showing the uncertainties associated with the computation. Figure 4.1 summarizes this workflow, while the implementation of its components will be explained in detail in the following Sections. The visualization of climatic microenvironments based on the spatial aggregation of a set of mobile measurement runs extends the core functionalities of TraVis, which are described in Chapter 2.

In summary, the workflow is based on the selection of a single, representative transect run, which is then spatially aggregated over a regular grid. The user can select the representative transect run and the grid cell size, while TraVis uses this grid to average the data collected within each of the resulting grid cells. Then, the averaged values are normalized for the representative transect run.

The user can now setup a self-organizing map, which is trained based on
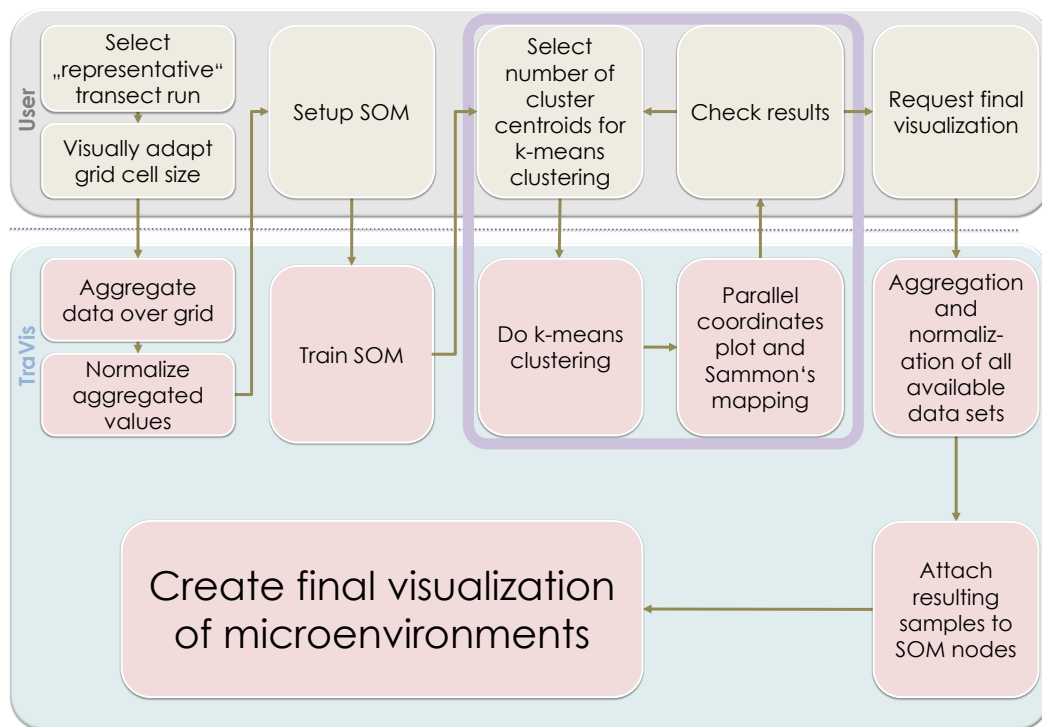
Figure 4.1: Overview over the user-steered workflow to visualize microenvironments based on the spatial aggregation of mobile measurements.

the aggregated, representative transect run. Then, the user can select the number of cluster centroids for the k-means clustering algorithm, which is applied to classify the neurons of the SOM. In a visual feedback loop, the user can evaluate the outcome of the algorithm and refine the number of cluster centroids, if necessary. Once the user is satisfied with the provided partition, s/he triggers the request for the final visualization, which is then automatically generated by TraVis. Finally, the user can navigate through this visualization and build hypotheses about the distribution and characteristics of the final climatic microenvironments.

4.3.2 Spatial aggregation of multivariate mobile transect measurements

In my solution, following the traditional approach, a regular grid is spanned over the bounding box of all transect routes. Since spatial scale plays an important role during the analysis, the size of the grid cells can be adapted to

the scale under investigation using a slider (Fig. 4.2). This scale can also be chosen according to the spatial resolution of the sensors that have been used for the mobile measurements (see also Chapter 3, [24]).

The observations, which are processed as described below, are averaged separately for each transect run and each variable over each grid cell. Furthermore, to guarantee comparability of data sampled at different points in time and data represented in different units, all aggregated samples belonging to one transect run and one variable are scaled to the interval $[0,1]$ based on their individual value range. While this technique leads to a loss of absolute values, statements about general multivariate relationships are facilitated [115]. As

Figure 4.2: Aggregating data over a regular grid (background map: [46]).

already mentioned in Section 4.3.1, the workflow starts with the selection of a single, "representative" transect run, which is then aggregated and normalized over the user-defined grid. Then, the processed data is used to train a self-organizing map (SOM) based on the implementation described in Chapter 2, Section 2.6.4. The user can, again, define the setup of the self-organizing map and select the number of cluster centroids for the k-means clustering, which is used to used to partition the set of neurons. Finally, the clustered SOM is applied to the aggregated and normalized representative transect run, leading to a summary measure for multivariate behavior at the grid cells covered by this run [160].

Once the partition of the exemplary transect run is finished, the clustering results are applied successively to all other transect runs available in the data set. For this procedure, the neurons of the SOM are used as "classifiers": The vectors associated with each traversed grid cell are associated with the k-means cluster membership of the closest neuron in terms of the Euclidian distance.

On the one hand, the disadvantage of this approach lies in the potential sensitivity of the result on the selection of the first, "representative" transect run. Theoretically, it has to be chosen based on the number of distinct multivariate value combinations, which can then be classified appropriately and detected in subsequently added runs. On the other hand, the SOM theoretically decreases this sensitivity: If the number of neurons is large, the number of classifiers is also large. Furthermore, a large number of neurons can also result in classifiers, which have not or only partially been adapted to the initial input. Thus, the neurons do not necessarily strictly resemble the multivariate relationships available in the representative transect run, but can additionally provide classifiers that might match a sample points from another transect run perfectly. This issue will be discussed in detail in Section 4.4.

4.3.3 Visualization approach and glyph design

To visualize this the aggregated, normalized, and classified data from all transect runs on a map, a radial glyph is assigned to each grid cell that has been crossed at least once. This is related to work done by [135, 132, 3], but depending on the given application context, my glyph was carefully designed to encode the...

- R1: ...number of transect runs traversing the grid cell.

- R2: ...cluster membership for each transect traversing it.

- R3: ...predominant wind direction during the time the transect has been conducted.

- R4: ...grid cell size over which the data has been aggregated.

The number of clusters (R1) has to be included into the glyph to enhance the perception of uncertainty associated with the visualization. If only one sample
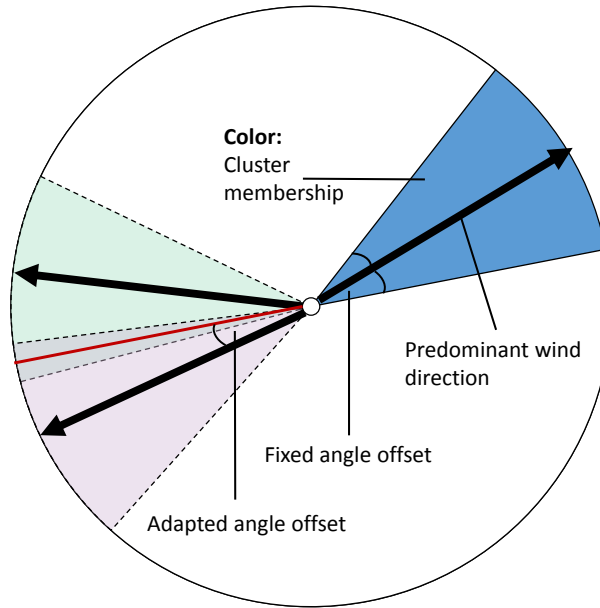
Figure 4.3: Glyph design.

is responsible for the appearance of a grid cell, the reliability of conclusions drawn from this representation is reduced. Encoding the cluster membership for each transect traversing a grid cell is also mandatory for the visualization (R2) because it reveals information about multivariate relationships found at this location. The predominant wind direction for each transect traversing the grid cell (R3) can give further hints about the relationship between the values measured at this point and the LULC upwind. This holds especially true for atmospheric attributes, whose spatial distribution is dependent on mixing processes [136]. Finally, the grid cell size (R4) determines the spatial resolution of the resulting aggregation.

Figure 4.3 shows the glyph design in a schematic way. A circular layout is used, since this design can easily encode the predominant wind direction at sampling time (R3) by the orientation of segments. These segments are created by applying a fixed offset angle left and right of the vector pointing into the wind direction, which is in my implementation given by the predominant wind direction measured at four weather stations surrounding the study site [105, 60] and therefore not necessarily represents the local wind conditions at a

particular sample location on a transect. The number of transect runs traversing the grid cell is encoded by the number of segments arranged around an inner circle, fulfilling R1. Color coding these segments according to the cluster membership fulfills R2. The grid cell size (R4) is proportional to the radius of the circle, which also prevents spatially adjacent glyphs from overlapping.

If two segments overlap due to similar wind directions at sampling time, the border between these two segments is moved to the half-angle between the two respective wind directions. This way, both segments are still visible, although it has to be noted that this solution impacts the perception of the wind direction. However, the approximate wind direction, and therefore also the approximate relation to the upwind LULC, can still be visually determined.

Figure 4.4 exemplifies the resulting glyph-based visualization. This particular example was created using a transect run recorded in Power Ranch on September 15, 2014, starting at 11:58 MST, as a representative run. The spatial aggregation was conducted for surface temperature, relative humidity in 1 m, air temperature in 1 m, relative humidity in 2 m, and air temperature in 2 m, using a grid cell size of 40 m, while an SOM of 20 x 20 neurons was trained over $N \cdot 10$ iterations and partitioned into 6 classes using the k-means clustering algorithm.

4.3.4   Glyph comparison

The visualization shown on Figure 4.4 can reveal a first insight into potential microenvironments within a study area, and details about the multivariate value distribution can be explored by zooming into locations of interest. However, the display lacks a proper guidance on which glyphs share certain properties, and a structured comparison of glyphs that are located further apart from each other can become difficult. Therefore, I decided to color the traversed grid cells according to glyph similarity in addition to showing the glyphs at their location. This way, the distribution of similar grids is summarized, while the advantages of the glyph-based visualization described in the previous section are still kept. The work described in this Section is still ongoing. The results are preliminary, and although they look promising, the procedure needs further evaluation in future work.

Grouping the glyphs according to similarity is not trivial. This is due to
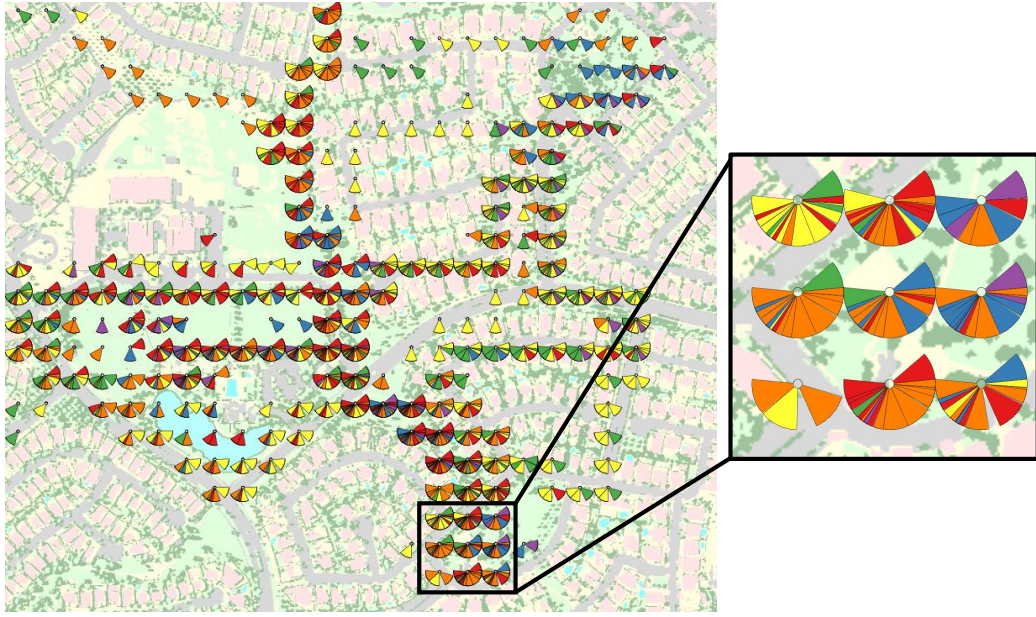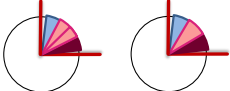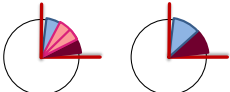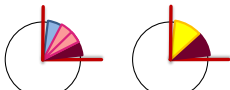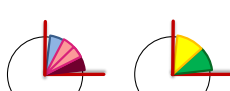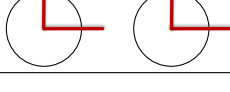
Figure 4.4: An example for the glyph-based visualization of climatic microen-vironments, showing a part of the covered study area. A close-up of the final glyphs is shown to the right (background map: [46]).

the nominal character of a glyph's appearance, which is given by the cluster memberships of the individual segments, i.e. through their assignment to a certain category of multivariate relationships (e.g., a segment can belong to the category "high surface temperatures and low humidities"). Another problem arises because the number of segments per glyph are different from grid cell to grid cell. Finally, the predominant wind direction per segment and, hence, the distribution of cluster memberships around a 360° circle adds to the complexity of the problem.

Therefore, the first step in comparing the glyphs to each other is finding a way to properly describe their properties. In my implementation, this description relies on the segmentation of the glyphs into eight sectors with a constant size of 45°. Per sector, the number of glyph segments is stored along with their individual cluster membership.

Based on this information, all glyphs in the display are categorized. Glyphs can only fall into the same category, if they are completely equal in terms of the above explained description: In each sector, they must share the same amount of glyph segments, and those glyph segments must have the same distribution

Case	Description	Graphical summary	Points (per sector)	Points (entire glyph)
1	- contain the same amount of segments - distribution of cluster memberships is equal		15	w*15
2	- do not contain the same amount of segments - the same set of different cluster memberships is available in both glyphs.		12	w*12
3	- do not contain the same amount of segments - the set of cluster memberships glyph 1 is a subset of the cluster memberships in glyph 2.		10	w*10
4	- may or may not contain the same amount of segments - the set of cluster memberships in glyph 1 is different from that in glyph 2, but there is a percentage of agreement		$10p_a$**	$w*10p_a$
5	- may or may not contain the same amount of segments - there is no agreement between glyph 1 and glyph 2 in terms of the cluster memberships		0.01	0.01
6	One sector contains segments, while the other does not.		0.01	--
7	Both sectors do not contain any segments.		0.01	--

** $p_a = \frac{2*\#segments\_agr}{\#segments_1 + \#segments_2}$ w* = weight factor

$\#segments\_agr$ = number of segments with agreeing cluster membership
(1 in the example given in the graphical summary of case 4)

Table 4.1: Metric applied for glyph comparison.

of cluster memberships. As a result, each glyph is assigned the resulting category, which is again defined by the description explained above.

The actual computation of similarity is conducted between these categories. To determine the amount of similarity, a "distance" between the glyph categories has to be defined, which is done based on the metric depicted in Table 4.1. Using this metric, the glyphs are compared, while the similarity "points", as shown in the two rightmost colums of Table 4.1, are added up.

The final distance between two glyphs, Δ_{glyph}, is then given by

$$\Delta_{glyph} = \frac{1}{\sum Points} \tag{4.1}$$

with $\Delta_{glyph} = 0$ for glyphs of the same category.

The comparison is conducted in two steps. First, the entire glyphs are compared to each other, independent from the above described subdivision of the glyphs into sectors. The points received for the cases depicted in Table 4.1 are added up and multiplied with a weight factor, which determines the importance of the overall similarity of the glyphs. In my implementation, this weight factor is constantly set to 10. Then, in a second step, the two compared glyphs receive additional similarity points for agreements in corresponding sectors. The latter comparison is conducted sector-wise, and points are added to the overall similarity score for each sector, in which one of the descriptions noted in Table 4.1 holds.

Both comparison steps are necessary, as Figure 4.5 demonstrates in a simplified way. Using only a sector-wise comparison would result in a large distance between glyphs showing the same overall distribution of cluster memberships, but don't agree in their appearance in corresponding sectors (Fig. 4.5, upper row). On the other hand, using a purely glyph-dependent comparison, i.e. not taking wind directions into account, would render glyphs with a diverging circular setup similar – they would share the same distance to each other. Adding points for corresponding segments in certain sectors decreases the distance between the regarded glyphs (Fig. 4.5, lower row).

Once the pairwise comparison of all glyphs is finished, the grid cells can be colored. Sammon's mapping [128] of the distances between glyphs into a two-dimensional color space is used to determine the background color of the grid cells. This way, grid cells containing similar glyphs will be assigned similar colors [160, 9, 131]. The algorithm is initialized using the normalized two-dimensional spatial coordinates of the glyphs' center points to take the spatial proximity of the grid cells and their corresponding glyphs into account. Iteratively, the two-dimensional coordinates are adapted to resemble the distance field given by the set of Δ_{glyph} [128]. An example for the final visualization is depicted in Figure 4.6, using the same setup as for the visualization in Figure 4.4.
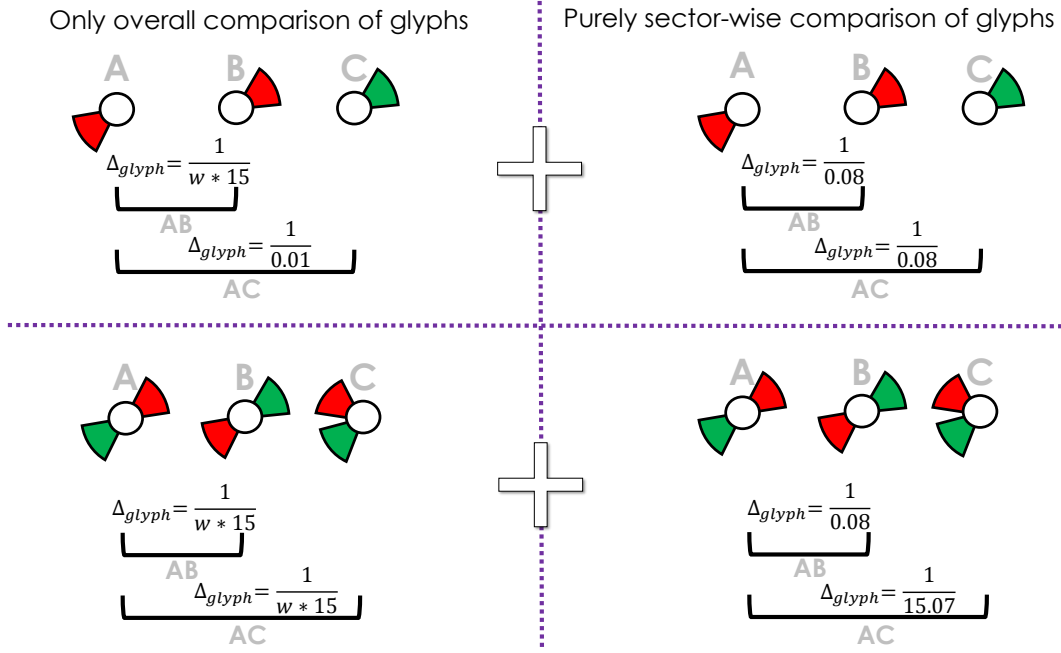
Figure 4.5: Using either only an overall glyph comparison (left column) or a purely sector-wise glyph comparison (right column) can both lead to unintuitive results. Therefore, the glyphs are compared as a whole *and* sector-wise. The resulting points are added up to $\sum_{Points}$.

Figure 4.6 reveals cases where the background color is not as different as expected. For example, the background colors of grid cells with only green cluster memberships are very similar to those with only yellow clusters, although they should be very distant to each other according to the above described metric. The same holds for background colors of grid cells with only orange colors, which share a similar background color with grid cells containing only red segments on their corresponding glyphs. On the other hand, the background color seems too distinct in other cases. One example for this is shown in Figure 4.6a, where two grid cells containing very similar glyphs receive red in one case and dark green in the other.

Besides that, the overall results look promising, as can be observed by qualitatively analyzing the example shown in Figure 4.6. Grid cells containing glyphs with predominantly yellow segments consistently appear in a gray-green color, while grid cells containing glyphs with predominantly orange segments are consistently colored in different shades of blue. Glyphs with predominantly blue segments appear in front of a dark-purple background, and grid cells host-
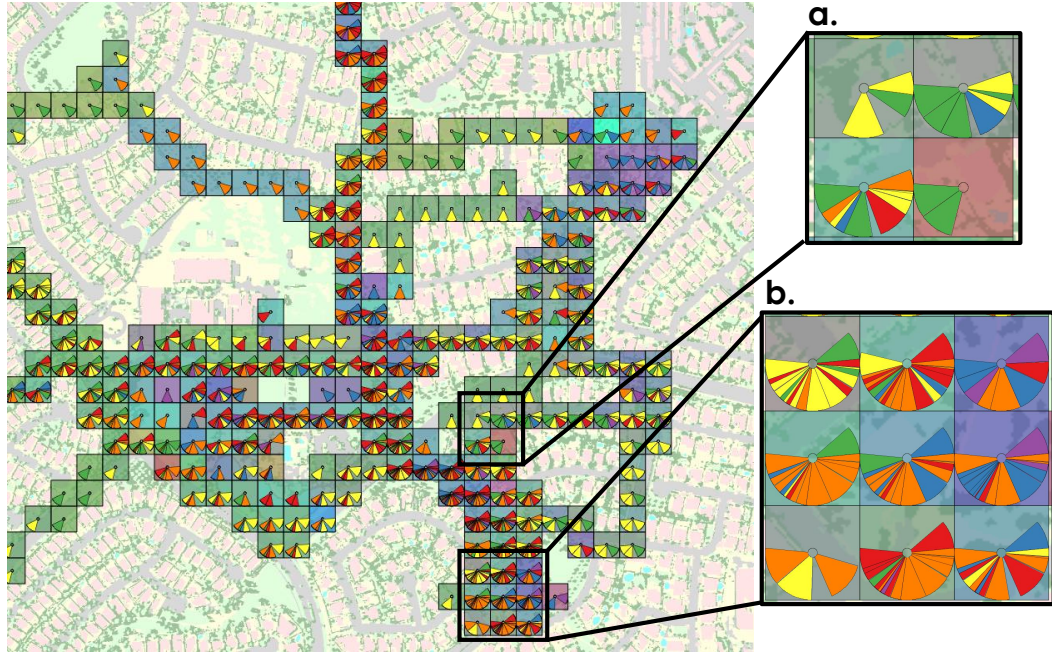
Figure 4.6: Coloring the background of each grid cell according to the similarity of the glyphs associated with each grid cell. (a) An example for a case, in which the background colors of similar glpyhs become too distinct. (b) An example, where the described procedure works well (background map: [46]).

ing glyphs with green segments appear in a green (but are very similar to those with yellow clusters). The close-up displayed in Figure 4.6b shows that different combinations of cluster memberships lead to different colors, but similar glyphs receive similar background colors. Thus, coloring the background using the metric described in this chapter can be useful to detect patterns of similar multivariate relationships, i.e. patterns of potential microenvironments, within the investigated area.

4.4 Validation

4.4.1 Classification of sample locations

The disadvantage of the aggregation approach lies in the potential sensitivity of the cluster structure on the selection of the first, "representative" transect run. Theoretically, it has to be chosen based on the number of distinct mul-

Variables (time-detrended)	Grid cell size	SOM height	SOM width	Number of SOM training iterations	Number of k-means cluster centroids
• Surface temperature • Air temperature 1m • Relative humidity 1m • Air temperature 2m • Relative humidity 2m	30	20	20	10 * N (N = number of samples after aggregation)	6

Table 4.2: Experimental setup for the quantitative sensitivity analysis.

tivariate value combinations, which can then be classified appropriately and detected in subsequently added runs. To estimate the consequences of the currently chosen initialization method, the sample points belonging to each of the k clusters are analyzed in terms of their mean value (Fig. 4.7). This way, an idea about the final distinctiveness of the clusters can be gained.

Iterating through a set of 21 different mobile measurement runs, each of them was applied once to initialize the algorithm. The other necessary parameters, i.e., the choice of input variables, the grid size, the SOM size, the number of SOM training iterations, and the number of k-means cluster centroids was kept constant (Table 4.2) to receive comparable results. The aggregated and normalized samples and their respective cluster membership were written to two .csv-files, whereas one of these files represents the aggregated, normalized, and clustered data of the "representative" mobile transect run, and the other file contains the samples of all runs at the end of the procedure.

Figure 4.7 shows the results of this experiment. Each compartment of this plot in the upper row shows the averaged parallel coordinates plot resulting from clustering the "representative" run, which is indicated below each column of the plot. The compartments in the lower row show, how the average values of each cluster changed, after the SOM trained using the "representative" transect run has been applied to all other mobile transect runs in the data set. The objective of this analysis was to see, to which extend the average values of the clusters change, when all other transect runs are applied to the initial SOM. Comparing the upper compartments of the plot with their corresponding lower compartments reveals that the main cluster structures are kept in a consistent way. For example, in the first column, the blue cluster meant "high surface

temperatures, low humidity in 1 m and 2 m, and high air temperatures in 1 m and 2 m" for the "representative" transect run chosen for the initialization (upper row). This structure could be kept after all other transect runs have been applied to this clustering result (lower row).

However, there are exceptions to this consistency. For example, if the mobile transect run started at 06:22 MST on February 05, 2015 is chosen for initialization, the cluster structure changes for the clusters colored in dark blue and in purple. The average air temperatures in 1 m and 2 m belonging to the dark cluster decrease considerably, and the dark blue cluster structure that has been prevalent for the initial mobile transect run is disturbed. In turn, the average surface temperature belonging to the purple cluster increases in a way that disturbes the purple cluster structure belonging to the initial mobile transect run shown in the corresponding plot in the upper row. As a result, choosing this mobile transect run as a representative run can decrease the reliability of the final visualization.

Overall, although major cluster structures are kept, results imply that not all mobile transect runs are equally well suited to initialize the self-organizing map. Furthermore, Figure 4.7 reveals that the distribution of average values belonging to different clusters changes depending on the choice of the initial transect run, sometimes more, sometimes less. Further investigation on this issue was conducted, as will be described in the following subsection.
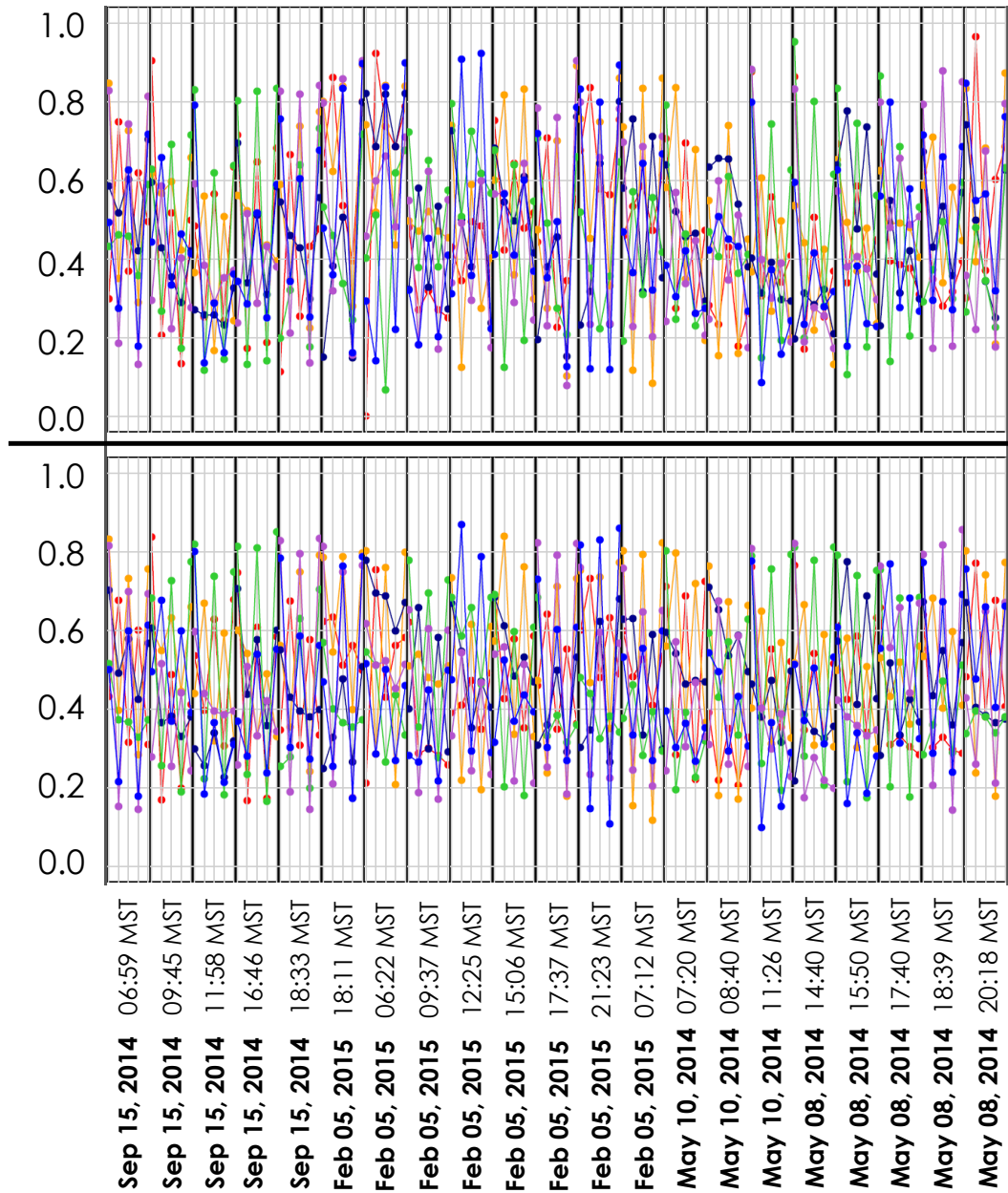
Figure 4.7: Averaged parallel coordinates plot. *Upper row:* Averaged parallel coordinates for the clustered "representative" transect run. *Lower row:* Averaged parallel coordinates, after all other runs have been attached to these clusters. Each averaged parallel coordinates plot represents (from left to right): Surface temperature, relative humidity 1 m, air temperature 1 m, relative humidity 2 m, and air temperature 2 m. Different colors represent different clusters, and the dots represent the average value of all members belonging to this cluster.

4.4.2   Sensitivity to selecting a representative transect run

To determine the sensitivity of the final visualization to the selection of the "representative" mobile transect run, the metric described in Section 4.3.4 was used as a basis to compare results created with different initializations. The experimental setup was the same as described in Section 4.4.1. $\sum_{Points}$ was computed between glyph categories of subsequent grid cells and exported as .csv-files. If the glyphs in subsequent grid cells belonged to the same category ($\Delta_{glyph} = 0$), they were assigned $\sum_{Points} = 400$.

Two error measures were computed between all possible combinations of two data sets: The mean absolute error (MAE), and the index of agreement (d, [169, 170]). Figure 4.8 shows the results of these comparisons. Overall, the dispersion of the error measures is small over the entire set of different combinations. The index of agreement does not become lower than 0.5, i.e. there are no cases, in which there is absolutely no agreement between two different initializations. The mean absolute error ranges between 40.52 and 73.76, averaging at 56.51, which can be regarded as relatively low when compared to a maximum possible range of 0 and 399.91. However, it has to be noted that the maximum possible deviation of 399.91 occurs in each of the pairwise comparisons, i.e. for each combination, the maximum absolute error is 399.91. Transferred to the notion used in Section 4.3, this means that two neighboring glyphs are, in at least one neighboring pair of grid cells, equal when initializing the algorithm with mobile transect run A, while they are completely different when initializing the algorithm with mobile transect run B. On the other hand, the average MAE never approaches this maximum value, i.e. the cases, in which this extreme deviation occurs, is not very frequent. As an average, this extreme deviation occurs in approx. 7% of all investigated neighboring grid cells.

In general, these results show that the described procedure for the spatial aggregation of mobile transect measurements is, to a certain degree, sensitive to the choice of the initial, "representative" mobile measurement run. Deviations are, however, not extremely high, so that the applied initialization method can still be regarded as valid. Nevertheless, in future work, research
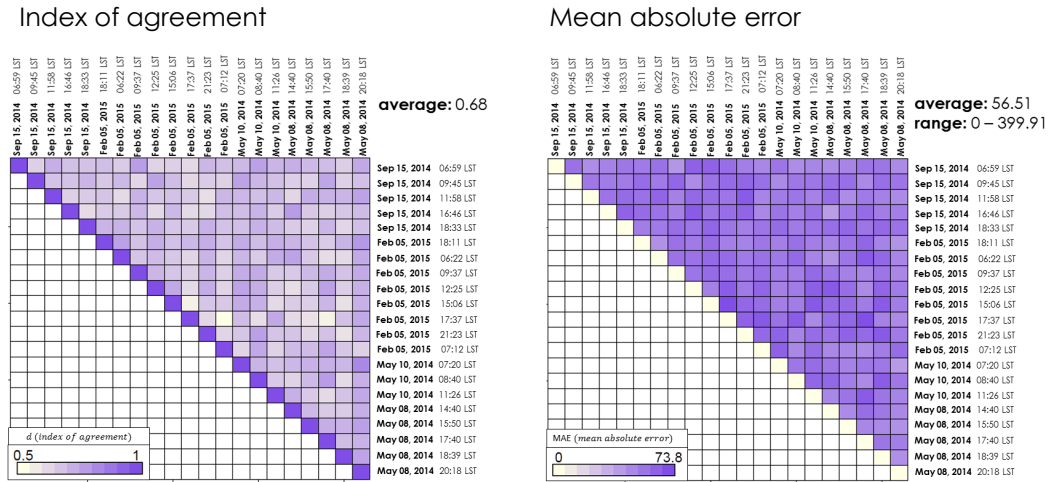
Figure 4.8: Validation results in terms of the index of agreement (d) and the mean absolute error (MAE).

on alternative initialization methods has to be conducted in order to find an approach that leads to more consistent, and therefore more reliable, results.

4.5 Use Case

The spatial aggregation was tested using the mobile transect measurement data set described in Section 2.5. For analysis, five variables were considered: Surface temperature, 1 and 2 m air temperature, and 1 and 2 m relative humidity. The objective of the use case was to find regions within the study area, which are comparable in terms of their multivariate relationship at all points in time.

The run conducted on September 15, 2014, at 0700 MST was chosen as representative transect run. The mobile platform traverses a longer route, covering a potentially large number of different multivariate value configurations. The grid cell size is chosen to be 30 m. For the SOM, a field of 20 x 20 neurons is used, which is trained run over $N \cdot 10$ iterations (N is the number of grid cells traversed by the exemplary transect run). Then, the SOM is clustered using k-means clustering over 6 cluster centroids. Using the parallel
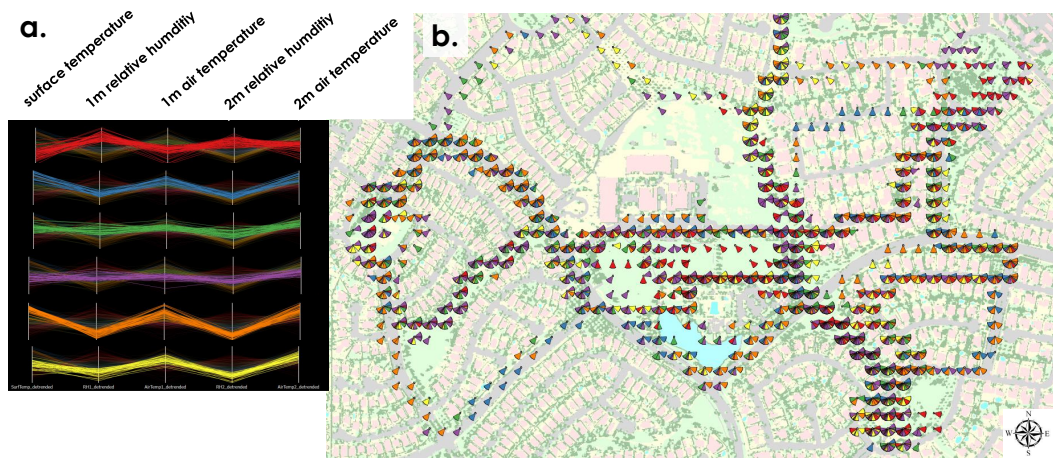
Figure 4.9: a. Brushed parallel coordinates plots to visualize the meaning of the clusters. b. The glyphs combining the entire set of mobile transect measurements on a map (background map: [46]).

coordinates plot, it can be observed that the data was well-partitioned into distinct classes of multivariate relationships (Fig. 4.9a).

Then, the clustering results are used to classify the spatially aggregated data belonging to all other transect runs. In the resulting visualization (Fig. 4.9b), several patterns can be identified: First, it is obvious that the wind direction is never northerly. Second, orange and blue classes, associated with high surface temperatures, low humidities, and high air temperatures, appear frequently over asphalted areas and between arrays of houses, as would be expected for this kind of environment. Third, the red cluster, associated with low surface temperatures, high humidities and relatively low air temperatures, and the purple cluster, which is similar to the red cluster, but with lower humidities, can predominantly be found in parks.

To highlight these patterns, the background of the grid cells are colored based on the procedure described in Section 4.3.4 (Fig. 4.10). The parks now clearly stand out: In contrast to the other, more impervious areas of the study area, they appear in pink and purple colors. Furthermore, another type of microenvironment becomes obvious, which is colored in light blue. It frequently

Figure 4.10: Coloring the background of the grid cells to highlight potential climatic microenvironments in the study area (background map: [46]).

appears near the transition of park areas and building-dominated areas. The glyphs hosted by these grid cells show combinations of all clusters, except for the red cluster, which never appears in the light blue grid cells. In the area south of the lake, a certain distribution of different cluster membership according to wind directions can be observed. While orange clusters are predominant for wind coming from easterly directions on the eastern side of the lake, the blue clusters frequently point towards the lake. While the overall structure of the blue and orange clusters are similar, they only share the high surface temperatures as a common baseline, while air temperatures are higher and humidities are lower in the orange cluster. This leads to the hypothesis that the road south of the lake, colored in light blue in Figure 4.10, corresponds to a transition area between a park- and lake dominated microclimate to a microclimate that is mostly influenced by impervious surfaces.

This statement, however, has to stay an hypothesis because the current implementation of TraVis only uses the average wind direction of four surrounding weather stations to approximate the air flow situation during a certain sampling time. The high density of flow obstacles in the study area can alter the wind direction at a certain sampling point, i.e. the approximation of the

wind direction might not be valid in all cases. Nevertheless, an obvious pattern could be detected, which is to be investigated further in a closer analysis.

4.6   Conclusion

In this Chapter, I describe a visualization approach to visually identify climatic microenvironments within a study site based on a number of mobile transect measurements. I partition the space using a regular grid before I aggregate the data associated with each grid cell by classifying it according to multivariate relationships and visualize it using radial glyphs. The glyph design enables the synchronous visualization of (a) the number of transect runs that contributed to the glyph, (b) the predominant wind direction at recording time, (c) multivariate relationships, and (d) the grid cell size.

To visually structure the display, glyph similarity is highlighted by coloring the background of the grid cells according to a metric that takes the glyph structure into account. This method can also increase the scalability of the described approach because the glyph design does not support the display of a large amount of mobile transect measurement runs.

In summary, the described approach supports forming hypotheses about the impact of urban design on microclimate, while also taking local data sparseness into account.

Chapter 5

Visualizing the Temporal Development Of Thermo-Radiative Features on Ground-Based Thermographs

Thermal infrared images frequently complement urban microclimate observations because they can be used to gain insight into the spatial distribution of surface temperatures or other thermal properties of surrounding materials. While the previous chapters focus on the visualization and analysis of mobile measurements, this chapter deals with the analysis of thermal infrared images.

The contents of this chapter were motivated by my collaboration in an urban microclimate project at the Arizona State University, Tempe, Arizona. Ariane Middel, Anthony J. Brazel, and their team investigated the impact of trees on the microclimate in a residential neighborhood in Mesa, Arizona. Thermal infrared imagery was part of this project to investigate the shape of shade cast by different tree species and its local impact on surface temperatures over the course of a day. One problem during the analysis of these images was the accurate delineation of the tree shade on the low-resolution images and the extraction of corresponding surface temperatures, which was considered difficult and error-prone using existing tools. Furthermore, the observations covered a variety of trees during each measurement cycle. Hence, the perspective of the images onto an individual tree varied from cycle to cycle, because initial

recording position and perspective could not be completely resembled. This hindered the registration of the images and made manual feature extraction necessary.

In this chapter, I present a tool that can help to approach the visualization and analysis of thermal features (as, e.g., introduced by tree shade) of time-varying images recorded with infrared cameras. Motivated by the problems described above, the approach should work on low-resolution images with a varying image perspective between the individual frames and guide the user in manually (and therefore reliably) extracting the features of interest. To adress this requirement, I combined the well-known active contours ("snakes") [85] and their extension, the gradient vector flow (GVF) [172], with typical graph-based visualization techniques for time-varying features. The behavior of a thermal feature over time is encoded by the nodes of the graph, while an animation highlights the features' movement on the original images. The tool was implemented in collaboration with Nils Feige (Computer Graphics and HCI Group, University of Kaiserslautern) and Lars Hüttenberger (Computational Topology Group, University of Kaiserslautern).

This Chapter is based on an article that was published in Environmental Earth Sciences [58].

5.1   Introduction

In urban microclimate research, ground-based thermography has various objectives. In some cases, it is used as an additional source of information besides point-wise measurement techniques for air and surface temperature, humidity, wind direction and speed or global solar radiation [144, 76, 55, 72]. In these studies, the infrared images provide additional insight into the spatial dimension of the thermo-radiative environment within the respective study area. Another motivation for microscale thermography is to increase the resolution of the information given by larger scale TIR images as, e.g., done by Hartz et al. [68]. Here, the smaller scale thermographs are linked to images recorded by satellites or scanners attached to airplanes or helicopters.

However, some studies primarily focus on ground-based stationary or hand-held thermography [138, 35, 78]. The goal is to gain insight into the thermal

behavior of construction elements and surface types over a certain time period. In general, the time-varying surface temperature patterns on the monitored components comprise a set of significant thermal features, such as hot spots or cold areas, whose temporal development can play a substantial role in understanding the complex feedbacks within the thermoradiative system of the investigated scene. In order to retrieve relevant information for this analysis, several TIR images have to be recorded over the course of a day, a month, a year or an even longer time period. This procedure is referred to as *time-sequential thermography* (TST) [34, 78]. If single static images are recorded, each of the individual thermographs can only represent a snapshot of a spatio-temporally continuous process, i.e., visual gaps arise naturally between two subsequent time steps. By closing these gaps and simultaneously highlighting the temporal development, the insight gained from the set of thermal images can increase.

Those questions are strongly connected to basic research in the visualization domain, since appropriate visualization techniques can significantly improve the insight into possible relations between the built environment and its thermo-radiative behavior. Therefore, I present a visualization system that is optimized for a feature-based analysis of a time series of infrared images. In this context and for the work described in this Chapter, I define a *thermal feature* as an area on a captured surface element, whose thermo-radiative signal diverges significantly from that of its surrounding, e.g., hot spots or cold areas emerging on a street or a building wall. Based on this concept, the tool desribed in this Chapter facilitates the identification of salient thermal features, which are highlighted in each snapshot using an edge detector. For every time step, the user can select features of interest by roughly surrounding them with a polygon. Each of these polygons initializes an active contour, a so-called "snake" [85]. This yields a shape descriptor for every single feature, because the snake converges to corresponding edges and delivers a parameterized closed curve. Since an automatic registration of the selected features from image to image can be error-prone, especially if the time difference between two subsequent images is very large, this task is delegated to the user to guarantee reliable results.

Once this information is entered into the system, the tool automatically

creates a visual summary of the features' temporal development. This summary consists of a directed graph whose vertices correspond to the identified features and which encodes topological events as well as each feature's size and summary statistics. In addition, an animated sequence visualizes the spatiotemporal development of selected hot spots or cold areas on the background of the linearly interpolated original TIR images. This final animation shows topological changes and approximates the time-dependent shape development of each feature.

In the research described in this Chapter, I attempted to create a tool that enables the user to process even difficult data sets, i.e., sets of images with low resolution, noise, and large time differences between consecutive frames. This Chapter is structured as follows: Interdisciplinary related work is described in Section 5.2. Then, in Section 5.3, the implementation details of the presented tool are laid out. In Section 5.4, the capabilities of the tool are demonstrated based on two use cases. The limitations of the current implementation are described in Section 5.5, before I draw a conclusion in Section 5.6.

5.2 Related Work

5.2.1 Analysis techniques for ground-based time-sequential thermography in urban microclimate research

Ground-based stationary or hand-held thermography in urban microclimate research is, amongst other things, applied to complement point-wise measurements or larger scale thermographs retrieved from air- or spaceborne sensors. It can also be used as the main source of information to investigate the thermal behavior of diverse surface types. In the latter case, time-sequential thermography is also used to gain insight into the temporal and spatial thermoradiative fluctuations of the investigated materials. Various examples for such research activities can be found in the urban climate literature [34, 78, 104, 138, 35, 103]. All of these studies face the difficulty of analyzing a more or less extensive time series of thermal infrared images. Since the applied analysis techniques are important for the motivation of the proposed tool, several examples are described in this section.

Meier *et al.* [104] recorded TIR images at a 1-min-interval over 48 h to determine persistence effects in upward long-wave radiation flux density patterns. Their resulting image set comprised 2,880 thermographs which were analyzed with regard to the temporal stability of the patterns by merging the thermographs to a sequence of 3- and 24-h summary images. While examining the resulting set of images, the authors made an effort to highlight salient flux density patterns by modifying the contrast of the aggregated images.

Another approach to analyze TIR images in an urban microclimate context is described by Chudnovsky *et al.* [35]. In this study, the authors investigated the thermal behavior of various urban fabrics over the course of a day. They utilize a high-resolution video thermal radiometer mounted on a pole to capture a scene comprising common urban surface types. Over 24 h, the camera continuously recorded the thermal scene for 10 min each hour. To analyze the resulting data set, the authors surrounded each urban object with a polygon and computed the average surface temperature and its standard deviation for the resulting set of image segments. Although this approach reliably summarizes the results for each object, this method does not account for the temporal development of interesting thermal patterns possibly occurring on larger surface elements.

A similar approach was utilized by Hoyano *et al.* [78]. Here, the authors chose two buildings with significant differences in material characteristics and recorded the buildings' surface temperatures for one summer and one winter day. For each day and each building, three complete thermographs were created, which the authors then combined into one single multi-temporal thermograph for a better comparison. Then, they enclosed the different building parts in polygons and computed the sensible heat flux for each of the identified segments. Although this technique fulfills the requirement of creating a comprehensive visualization for a time series of images, it does not seem very intuitive.

However, considering the studies described above, it can be stated that the progression of identified features might be worthy of closer investigation, e.g., to complement traditional polygon-based aggregations. These features do not necessarily have to represent surface temperature patterns, but can also consist of sophisticated radiation flux measures. In studies in which TIR mea-

surementswere conducted with a temporal resolution of a minute or even less, and where no significant gaps arise, the data set to be analyzed is very large. In these cases, the features can occur on a set of summarized spatio-temporal maps (e.g., [104]).

5.2.2 Analysis techniques for thermography in civil engineering and cultural heritage protection

In the research field of cultural heritage protection, infrared thermography is utilized to study the thermal behavior of monument materials, to assess the quality of the applied conservation method, to monitor the development of the investigated artifacts, and to document their state [140]. Similarly, the purpose of thermal infrared imagery in building façade monitoring is to detect heat loss, insulation lacks, air leakage or moisture sources – problems that usually result in an abnormal temperature distribution on the investigated object's surface [18, 123]. Since the thermography analysis tool described in this Chapter is in principle adaptable to such tasks, contributions in these domains are important for my work. Here, I describe two examples.

Lerma *et al.* [96] compared two different analysis techniques for the exploration of thermal images acquired from a historic building's façade at different times of day. To determine moisture-prone areas, they created (a) a simple pixel-wise temperature difference map and (b) conducted a principal component analysis (PCA). Findings showed that simple temperature differences cannot represent moisture problems on a façade accurately, since temperatures are sensitive to shadows and image noise. On the contrary, PCA delivered reliable results in mapping moisture problems.

Multitemporal image pairs of digital photographs and TIR images were utilized by Sidiropoulou-Velidou *et al.* [140] to uncover alterations on the façade of the Cathedral of Valencia. The authors collected four image pairs over the course of 1 day. In order to create a comprehensive map of possible façade alterations, they combined the results from four analysis methods: PCA, supervised and unsupervised classification, pseudo color images, and temperature measurements of selected image spots.

Although the temporal component of a sequence of thermal infrared images

is taken into account in the presented studies, a time-dependent visualization is not provided. In fact, the multitemporal images are mostly aggregated to one single thermograph instead. Hence, the temporal development of a hot spot or cold area or the thermal development of a structure element over time cannot be monitored – an aspect which is covered by the tool described in this Chapter.

5.2.3 Multi-temporal image processing and visualization of time-dependent data

Multi-temporal image processing is a substantial research topic in a variety of fields, such as remote sensing [31, 12], medical imaging [47], or video processing [66, 176]. In these domains, numerous methods have been developed to detect, track and describe changing shapes of objects. Related objectives can be found in the visualization literature, where feature detection and development in a time-varying data set is the prevalent motivation (see, e.g., [167, 142, 122]).

In the field of remote sensing, multi-temporal image processing is used to detect changes in land cover features on various scales. Occasionally, active contours ("snakes") or a variation thereof are applied for this task. One example can be found in the work of Ardila *et al.* [12]. Here, the authors use a localized active contours method to detect changes in urban tree crowns on a time series of highresolution aerial images, which are initialized by fitting an ellipse to each tree crown. A similar study applies active contours to a time series of satellite images to evaluate their ability to detect and model changing object contours [31]. In the case of moving objects, the authors assumed a monotonous evolution of the contour position and used multi-snakes to model their shape development. Although the results in the example given are reliable, it cannot be applied to all sample cases investigated in this Chapter, since the contours of the features in TIR images might be subject to complex movement as well as splitting and merging.

In general, tracking complex feature development is a common problem in the analysis of large data sets, mostly stemming from simulations, and is consequently recurrently addressed in the visualization community. The first step

here is to extract features from a given data set according to a problem-specific criterion, e.g., via isolines, isosurfaces or other thresholding techniques based on scalar quantities derived from the data set [167]. Related topological concepts such as Reeb graphs or contour trees can additionally be applied to a scalar field to retrieve information about the connectivity of the features [142]. After initial definition, the features can be tracked over a number of time steps to analyze their temporal behavior, which can involve continuation, creation, dissipation, bifurcation, and amalgamation [129]. Numerous techniques have been proposed to solve this problem, which can roughly be classified into pixel-based methods and feature-based methods [122]. While pixel-based methods rely on the offset of the features from frame to frame, which can, for example, be retrieved via optical flow, feature-based methods use measures such as overlap or attribute correspondence to identify matching features in subsequent time steps [122]. However, the successful application of such concepts assumes that the features can be clearly identified over time, which is not necessarily given in TIR image data sets. In a worst case scenario, the thermographs have a low spatial and temporal resolution and are noisy, which makes it difficult to automate the process of identifying features in one frame and finding their unique correspondence in the next. Nevertheless, the notions are still helpful for the developed tool in terms of classifying time-dependent feature events.

5.3 Implementation

The main purpose of the tool is to assist researchers in analyzing their set of multitemporal TIR images (or their time series of preprocessed summary images) and to help them gain insight into the thermo-radiative behavior of the investigated scene. This analysis is feature-based, i.e., the output of the system is a dynamic visualization of the temporal development of hot spots, cold areas or other objects of interest. In this section, the main aspects of the tool are described in detail, while a graphical summary of the processing pipeline is provided in Fig. 5.1.

The tool is implemented using C++, Qt4, openGL 1 and openCV.
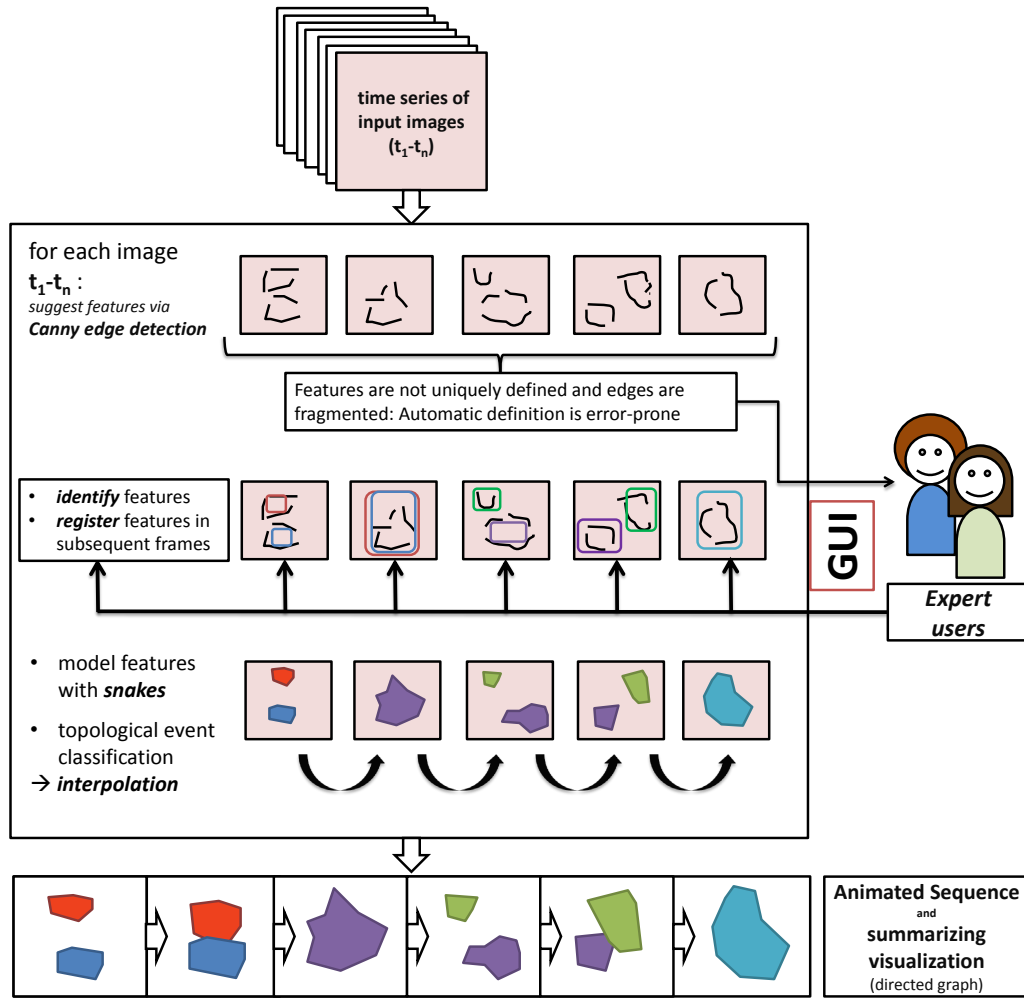
Figure 5.1: Steps in the method from a set of static images to a visual summary of the temporal feature development.

5.3.1 Finding significant thermo-radiative features

The tool described in this Chapter processes a time series of TIR images or sophisticated preprocessing results, e.g., flux density maps as proposed by Meier *et al.* [104]. They consist of a (usually relatively low) number of pixels. Each image pixel represents a specific data value such as radiative temperature or other preprocessing results, which I will simply refer to as "temperatures" in the following.

Initial feature acquisition

First, images in the input time series are smoothed with a Gaussian filter to reduce image noise. Then, the tool highlights the most significant feature edges for each time step. I assume that the major features are surrounded by large temperature gradients and therefore apply an edge detector to determine their location. The Canny edge detector [30] is chosen for this task, since this algorithm is known to be accurate and simple [174]. Furthermore, I found that it delivers satisfactory results for the purpose of this work, as shown in Figure 5.2. In this TIR image, the tree shade in the center of the image can easily be perceived, as well as the shade of the tree in the right part of the scene. The problem with Canny edge detection, however, is that the edges are fragmented in most cases, i.e., they do not form closed curves which could then be regarded as feature *contours*. One possibility to solve this issue would be to utilize segmentation algorithms which partition the input images into closed regions [174]. Although this would lead to unique segments in the thermograph and could therefore help to automatize the process of defining features, it has three major disadvantages. First, image segmentation is an error-prone procedure, since over- or under-segmentations can occur easily [152]. Second, the identification of regions in one image with corresponding regions in the next picture is complicated, because segmentation criteria applied in one frame might not lead to consistent results in the following frame. Of course, the criteria for each image time step could be adapted dynamically, but this implies that the regions in subsequent frames are not defined by the same criteria, which eliminates the basis for a reliable identification over time. Third, automated image segmentation neglects the user's individual research needs. The possibly large number of predefined segments might not be able to highlight what the user is actually interested in and could lead to unnecessary clutter.

These considerations led to the conclusion that the best solution to this problem is to delegate the task of closing the fragmented edges to the user. Based on the Canny edge detector results that are overlaid over the original TIR images, a number of regions can be identified according to the specific research focus. Using the mouse, the user can surround each feature with a polygon, which then serves as initialization for an active contour (a "snake").
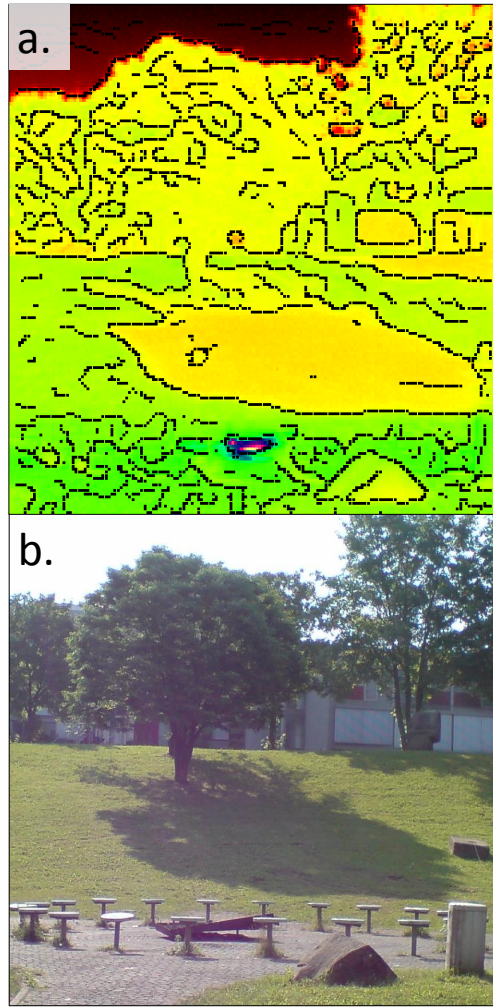
Figure 5.2: a) Edge detection of significant thermal-radiative features in a TIR image using the Canny edge detector. b) Corresponding digital photograph (provided for context)

Parameterizing the shape of selected features

Starting from the initial set of polygons specified by the user, the shape of the corresponding features is modeled using *snakes* [85]. A snake is a parameterized curve that minimizes the energy functional given as

$$E^* = \int_0^1 E_{int}(x(s)) + E_{img}(x(s)) + E_{constr}(x(s))ds \qquad (5.1)$$

where $x(s) = (x(s), y(s))$ is a parametrical description of the snake's position in the image plane. $E_{int}(x(s))$ represents the curve's internal energy and can

be written as [85]

$$E_{int} = 0.5[\alpha|x'(s)|^2 + \beta|x''(s)|^2].$$ (5.2)

Here, the first partial derivative with respect to s expresses the curve's surface energy which can be controlled by the coefficient α, while the second partial derivative with respect to s regulates the curve's bending energy. The latter can be manipulated through β [85]. In my tool, both α and β are user-defined, meaning the convergence behavior of the snake can be controlled to a certain degree. To give inexperienced users an initial guess for the value range of these parameters, the default is set to the common values of $\alpha = 0.6$ and $\beta = 0.5$. $E_{img}(x(s))$ expresses the image forces acting on the snake and is derived from the temperature values in the TIR image $I(x,y)$. It can be used to create force fields leading to the individually desired snake behavior by placing force minima at meaningful positions in the image, e.g., at edges [85]. $E_{constr}(x(s))$ are additional constraints that determine whether the curve is attracted or repulsed by these positions. Once initialized, the snake moves through the image force field and converges at a location that minimizes both internal and external energies and therefore minimizes E^* [85].

However, this traditional active contour model only delivers appropriate results if the snake is initialized near a meaningful position due to the original creation of the image force field. If the snake is supposed to move to salient image edges (as is the case in the attempted analysis), $E_{img}(x(s))$ is originally derived by computing

$$E_{img}(x(s)) = -|\nabla I(x,y)|^2.$$ (5.3)

This does not lead to an extended force field, and a snake can thus easily get stuck in areas with minimal forces [172]. In addition, traditional snakes have the disadvantage that a convergence to contours with deep and narrow concavities is not possible, because the image force field within the concavity only points towards the edges and not into the concavity itself [172].

To overcome these issues, the *gradient vector flow* (GVF) proposed by Xu and Prince [172] is used to compute all external forces acting upon the snake's movement. It is initialized by a binary edge map $b(x,y)$ of the image that can be derived through the Canny edge detector. Based on this input, the GVF is

a vector field $\boldsymbol{v}(x,y) = (u(x,y), v(x,y))$ that minimizes

$$E_{GVF} = \int \int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla b|^2 |v - \nabla b| dx dy \qquad (5.4)$$

where μ compensates for image noise, while u_x, u_y, v_x, v_y are the partial derivatives of the vector components in x- and y-direction, respectively [172]. Originating at the edges, $\boldsymbol{v}(x,y)$ diffuses the force field over an extended area of the image, while the edges themselves reach minimal values. Thus, the snakes converge to the edges even if they are not initialized in their direct adjacency. Because $\boldsymbol{v}(x,y)$ is naturally smooth, the snakes can also proceed into edge concavities within this force field [172].

To compute the snake convergence, the internal energy of the active contour and the external energy of the gradient vector flow field have to be balanced. Consequently, Equations 5.2 and 5.4 have to be solved and combined. To do so, the implementation of the GVF was geared to the modus operandi described by Ivins and Porrill [83], while replacing the original external force field with the GVF, which were in turn computed as proposed by Xu and Prince [172]. In a first step, a minimal energy condition for the internal energy of the snake is deduced by means of the Euler-Lagrange equation. Ivins and Porrill [83] present an iterative method that yields a system of $2N$ equations with N being the number of points on a snake. The resulting matrix is a cyclic symmetric pentadiagonal banded matrix M [83]. Since the components of M are constants derived from the snake's internal energy and remain constant during the iteration cycle, it is sufficient to invert M only once, i.e., before the first iteration. To implement this inversion, Cholesky decomposition is used. Now, the convergence of the snake in the force field is computed as follows:

$$u_{t+1} = M^{-1}(u_t + \delta t \frac{\partial P}{\partial u_t}) \qquad (5.5)$$

Here, u_t are the positions of the snake points at iteration step t, while $\frac{\partial P}{\partial u_t}$ corresponds to the external force field at u_t. The latter can be paraphrased to be the direction of the force position u.

Overall, the Canny edge detector results combined with the GVF snakes work well for the identification of meaningful features on TIR images. An example is depicted in Figure 5.3. Here, the shape of the shade in the center of the TIR image was roughly sketched with a polygon (Fig. 5.3a), the initial snake matches the previously detected shade edges perfectly (Fig. 5.3c).
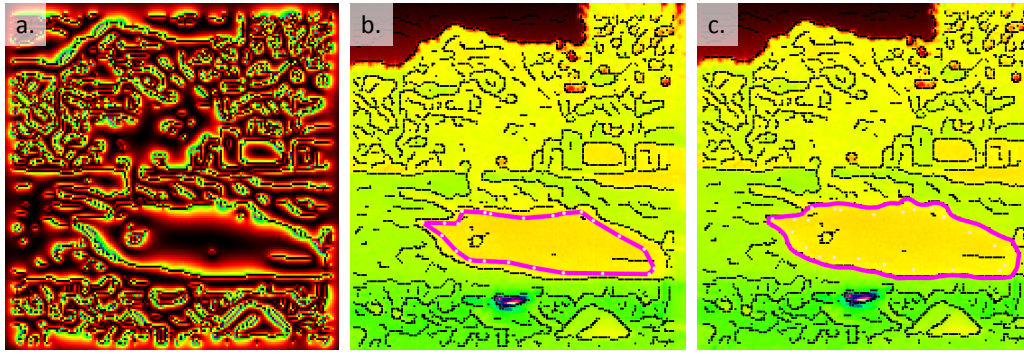
Figure 5.3: The convergence of an active contour to salient image edges. a) GVF field computed as proposed by Xu and Prince [172], b) initial polygon, and c) converged snake. For the corresponding digital photograph, see Figure 5.2b.

5.3.2 Summarizing temporal feature development

Once the thermo-radiative features were identified in each TIR image, their temporal development can now be summarized. Therefore, the corresponding features in two subsequent frames need to be registered, such that they can be transformed into each other. Although active contours can be applied to this task [31, 85], errors can occur if the difference between the individual TIR images is very large. This issue is demonstrated in Figure 5.4, where the converged snake in time step 1 (Fig. 5.4a) surrounds a hot spot in the center of the street displayed. Mapped to the image representing time step 2, the snake partially degenerates to a line, since the contour converged to the wrong edge. While a human user easily perceives the correspondence between the smaller warm area on the first image and its larger pendant on the second one, it is very difficult and error-prone to automatize this process. Therefore, the task of finding corresponding thermo-radiative features on subsequent image frames is controlled by the user to guarantee reliable results.

To summarize the temporal development of thermal features over a set of images, a topological classification of feature events is applied based on the categories described by Samtaney *et al.* [129]: continuation, creation, dissipation, bifurcation, and amalgamation. The progression of these events is visualized and internally stored as a directed graph, as this is a well-established technique (see, e.g., [142, 122, 167]). Here, each vertex represents the state of the
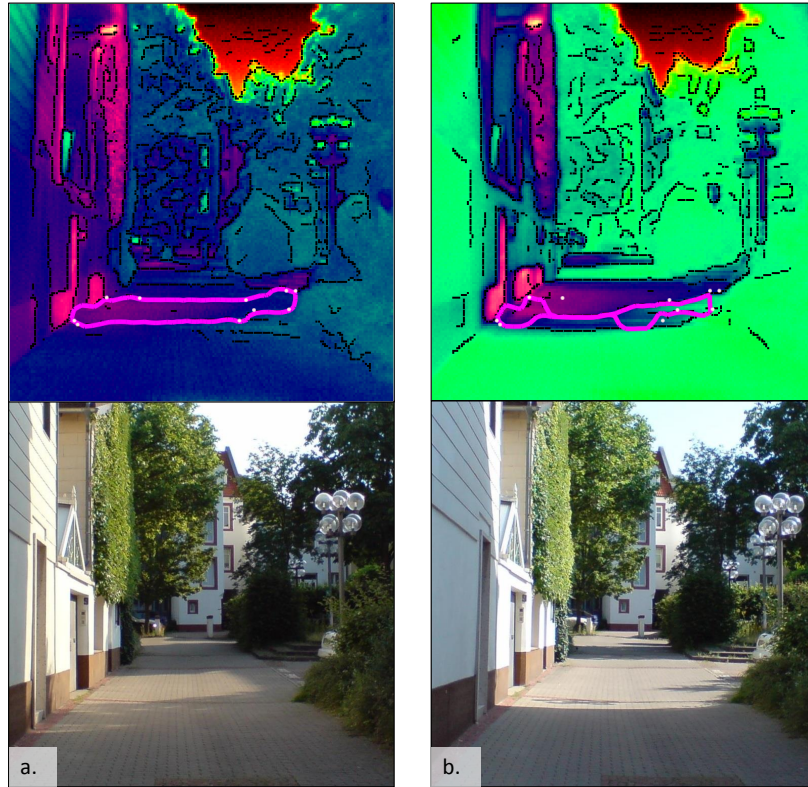
Figure 5.4: If the time difference between two subsequent TIR images is very large, the snake may converge erroneously. The images on a) have been collected at 8:17 CEST, while those on b) were taken at 10:01 CEST. Digital photographs are provided for context.

respective feature in an image frame. The vertices of the graph were used to visually encode summarizing information about the selected thermal features. Hence, the size of a vertex corresponds to the size of the corresponding feature. In detail, the sizes of all features identified in the image series are normalized, and the radius of each vertex scales with the normalized value assigned to each feature. The color of a vertex represents the average temperature of the selected region. To account for a feature's internal temperature variability, the standard deviation of the values enclosed by a snake is encoded by the color of another circle that is drawn around each vertex.

The directed graph also serves as a basis for the interpolation of the features' shape, location and topology over the time series of thermographs. The interpolation itself is conducted between the vertices along each path with splines. As a result, the interpolating curve provides a smooth transition between the

images and is, hence, well suited for the followed purpose: the animation of the features' temporal development. To increase the comprehensibility of the rather abstract graph-based representation, a slider was added that visually and interactively connects the animation and the directed graph. As a result, the user can simultaneously skim through the animation and the graph. In the latter, a vertical line indicates the time step currently shown in the animation. Figure 5.5 shows the graph layout as it is implemented in the described tool. The color scale for the vertices was created using a color scheme from ColorBrewer [67].
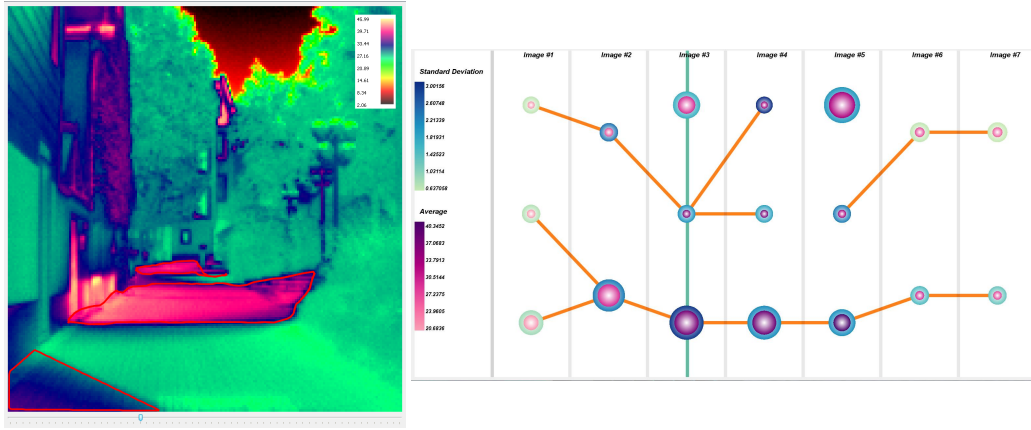
Figure 5.5: A directed graph is used to visually summarize the temporal behavior of thermal features over time. Each vertex corresponds to one thermal feature identified in the respective frame. The color for the inner circle of each vertex encodes the average temperature for the thermal feature it represents. The outer circle's color is used to visualize the standard deviation of the measured temperatures within this feature. Finally, the size of each vertex represents the thermal feature's size in the image.

5.3.3 Animating feature development

To animate the feature development over time, the following interpolation procedure was applied: In a first step, all possible paths through the directed graph described in the previous section are determined via depth-first search, whereas each path corresponds to the temporal evolution of one snake. For every vertex, the current state of the passing snakes is internally represented

as a discrete set of points $(x_{i,t}, y_{i,t})$ at a time step t. These points are stored in a list, ordered according to the sequence of points the user selects to create a polygon that initializes the desired active contour. As the user does this for every frame while being free in how to do this, it can occur that the arrangement of the points changes from image to image, even if the features have been declared as correspondent to each other. While transforming the first feature into its subsequent pendant, a reversed order of points from clockwise to counter-clockwise can lead to irritating artifacts in the animation. Therefore, the next step after the initial setup of paths and curves is to adjust the order of points on the curve to always be oriented clockwise.

Once the orientation of the individual snakes matches, the points on corresponding contours have to be identified to approximate a smooth transformation between them. Therefore, I consistently defined that the first point in each of the curves' list corresponds to the uppermost point on the curve. As each curve consists of a fixed number of vertices, the correspondence between the remaining points belonging to two snakes is given by their index in a vertex list. Between corresponding points on both curves, *natural cubic splines* are used for interpolation.

The described approach is a straightforward procedure to transform one snake into its subsequent pendant, if the curves undergo the topological event of *continuation*. Although the procedure during the other topological events is based on this algorithm, I implemented slight variations for the other events:

- *Creation*: the center point of the to-be-created curve is extended until it matches the shape of the final feature.

- *Dissipation*: the curve is contracted to its center point, which is then deleted.

- *Amalgamation*: the initial two curves are both mapped onto the contour of the final feature. Merging occurs as soon as the snakes begin to overlap. Affected points on the curves are deleted.

- *Bifurcation*: the initial curve is mapped onto the two resulting feature contours. Analogous to the amalgamation process, the points leaving overlapping sections of two curves are shown successively.

5.3.4 The graphical user interface
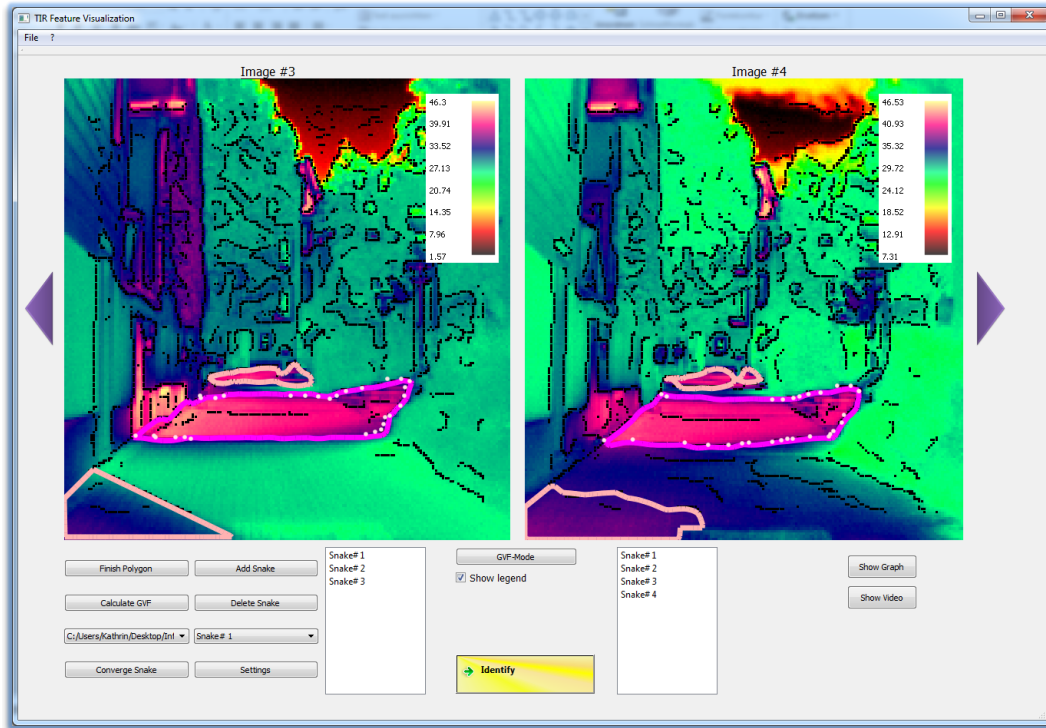
Figure 5.6: The graphical user interface.

The graphical user interface (Figure 5.6) was designed in a simple and ordered way:

- The working space

- The animation output

- The graph-view mode

In the *working space*, the expert can edit features in every time frame. To assist the user, two subsequent TIR images with their highlighted edges are always shown at a time, while scrolling through the whole time series of thermographs is possible with the arrow-buttons to the left and right side of the window. In the TIR images, even small spatial temperature variations can be perceived easily, since a large number of colors is used to depict temperatures by choosing the values located on a helical path through the RGB color space. A legend to the upper right side of the image canvases shows the attribution

of temperature values to colors. Since this information can occlude potentially interesting parts of the image, it can also be hidden if necessary.

The user can now select a thermo-radiative feature by drawing a polygon in its adjacency after clicking the "Add snake" button. It is not necessary to accurately draw a closed polygon, since "Finish polygon" completes it by adding the last missing edge. If all features are marked this way, the "Calculate GVF" button can be clicked to compute the gradient vector flow field, as described in Section 5.3.1. On demand, the resulting GVF field can be shown on the canvas by means of the "GVF-mode" button. With "Converge snake", a selected polygon is automatically transformed such that it matches the nearby edges. If the user is not satisfied with the resulting contour, they can adjust the snake parameters in the "Settings". All created feature contours on a thermograph are listed in the white boxes underneath the according TIR images. If two features should be identified with each other, the user selects the corresponding features in the list and clicks the "Identify" button. Analogously, an asymmetric number of features can be selected for both lists, which corresponds to amalgamation, bifurcation, creation or dissipation, respectively.

Once the user has edited all images that way, the interpolation results can be viewed in the animation output, which is triggered with the "Show video" button as soon as all images are loaded into the memory. A new window opens, in which the animation can be dynamically steered with a slider. This way, the expert has the option to repeat possibly interesting sequences. To render a smooth transition of the TIR images in the background of the highlighted feature development, the images are blended linearly. The animation is stored as a video file. This video also serves as a backup for the work conducted within the tool and can be used for presentation purposes rather than the slider view, since the latter may stutter if the image data become very large.

Finally, the *graph-view* mode shows a depiction of the topological event graph (Fig. 5.5b) on which the animation is based. Being a comprehensive and simple visualization, it can give the user a summarizing view of the features' temporal development.

5.4 Use Cases

5.4.1 Displaying the temporal development of sunny areas on a small street

Data set

The tool described in this Chapter was tested against a time series of seven infrared images recorded with a FLIR i60 thermal camera on July 9, 2013. The sample thermographs have a resolution of $180x180$ pixels and were collected outdoors to show the thermoradiative environment in a street in Kaiserslautern, Germany. The utilized camera operates at a long wave length with a spectral range of 7.5-13 μm, which makes it suitable for such analyses: It is insensitive to reflections of solar radiation and can record small spatial temperature changes [96]. The set of thermographs represents varying thermal situations over the course of a day, since they were taken within varying time intervals between 08:17 and 21:26 CEST.

The images were recorded in a hand-held manner, i.e., without tripod, such that rotation, translation and scaling occur between the image frames. Due to the trees in the background of the scene, the images become partially noisy. Finally, the low resolution of the thermal camera aggravates the analysis and automatic registration of the images.

Results

The preprocessed TIR images were entered into the tool and a realistic set of thermal features in each image was selected by drawing polygons onto the canvas. Of predominant interest was the temporal thermo-radiative behavior of the street, which covers a big part of the image's lower half. In the chosen example, the development of shaded and sunny areas on the street was tracked, since they serve well as unique thermal features and can be identified easily due to clear weather conditions on the measurement day. The suggested feature edges, which were computed using the Canny edge detector as described in Section 5.3.1, matched the borders between shaded and sunny areas very well in most of the cases. However, due to image noise, some additional edge fragments occurred in parts of the TIR images, but they did not disturb the

overall perception of feature contours.

After setting up the polygons according to the investigation focus, the gradient vector flow field is computed and the active contours are triggered to converge to the edges. Satisfactory results could be obtained with the snake parameters set to $\alpha = 0.6$ and $\beta = 0.5$. However, the edges surrounding interesting features were severely fragmented in two cases, such that the snakes did not converge properly – they partially degenerated to lines. The tool described in this Chapter handles such situations as follows: the user is not forced to use active contours to let the initial polygons converge to the edges. If the snake does not produce acceptable results, the user can leave the polygon unchanged. This way, the expert can choose features which are not covered by the Canny edge detector.

The number of initial polygons per TIR image varied between two and four, and the resulting graph can be found in Figure 5.5. During the registration of features in consecutive images, each topological event was covered at least once. In the resulting animation, the temporal development of the street's hot spots and cold areas could be observed, since the snakes only *highlight* these features while the original temperature distribution of the surrounding is still shown in the background at any time. With regards to content, the animation shows the shape development of the most dominant sunlit and shaded areas on the investigated street section. It could easily be seen that the central part of the street section is influenced by an extended hot spot, which moves from the left to the right side of the street, while first increasing and then decreasing in size. This feature seems to dominate the whole thermal situation on the investigated surface element. In the rear part of the street, a smaller hot spot develops slowly and splits in the middle of the day to disappear in the evening. Near this location, but more towards the camera, a cold spot develops in the afternoon. In the foreground, a significant warm area develops on the street surface. Later, a larger cold area develops in the late afternoon due to the shade cast by adjacent buildings. After that, this significant cold area dissolves again by assimilating to the surrounding surface temperatures.

5.4.2 Displaying the temporal development of tree shade

Data set

The data set used for the second use case was recorded on June 21, 2012, in a neighborhood in Mesa, Arizona, by Ariane Middel, Anthony J. Brazel, and their team. The image series consists of 17 thermographs, which were recorded between 0600 MST and 2200 MST and which capture the development of the shade cast by a blue palo verde (*parkinsonia florida*). The TIR images were recorded using a FLIR infrared camera, and the images have a resolution of $60x60$ pixels.

As was also the case for the data set described in Section 5.4.1, the images were recorded in a hand-held manner, and rotation, translation and scaling occur between the individual frames.

Results

The main interest of this analysis was the difference between surface temperatures in sunny areas and surface temperatures under the canopy of the blue palo verde during the early morning, daytime, and after sunset. Therefore, the tool was used to visualize the thermal development of a small sample patch outside of the tree's influence area as a reference condition and as a comparison to the temporal and (as far as possible) shape-wise development of the area under the canopy of the focused tree.

After loading the images into the tool, the edges of the shade were automatically detected and displayed on top of the images. Since the resolution of the images is extremely low, the edges were fragmented, but they could still give an initial idea about the borders of the shade in the lower part of the images (Fig. 5.7). Furthermore, due to the limited field of view of the utilized infrared camera, the shade was in some cases not fully represented on the TIR images. These two properties led to the snakes not converging properly, and the polygons were chosen to be the better choice for delineation.

Figure 5.8 shows the resulting graph. The diurnal course of surface temperatures can be observed, as well as the size of the tree's area of influence (as far as it was available on the infrared images, which was not always the case due to the limited field of view of the camera). Obviously, and not surprisingly,
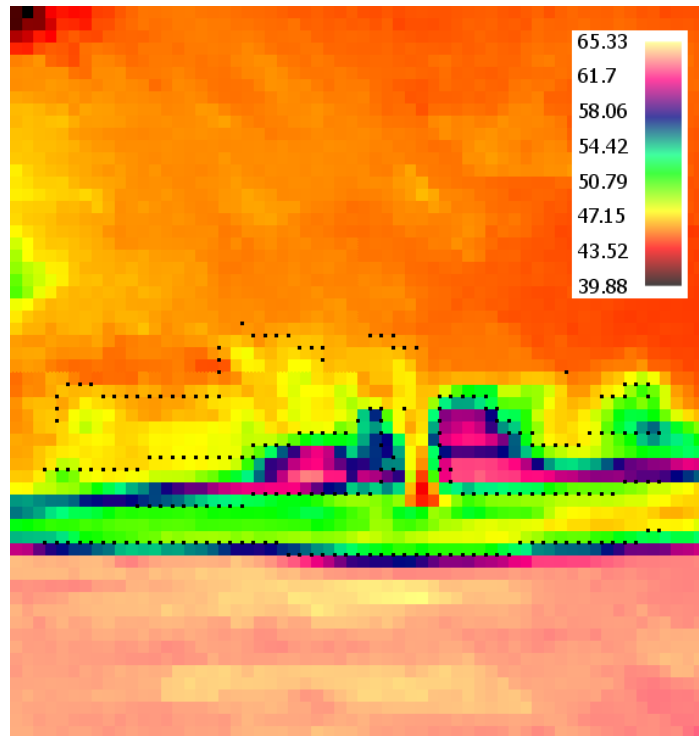
Figure 5.7: The edges of the shade were fragmented, but could still be recognized.

the surface temperatures outside of the tree's shade are higher during daytime than those inside the shaded area. For the given use case, the topological events are not of interest, since they just include continuation.

Figure 5.9 shows the time series of images along with the delineated areas, whose average surface temperature and the corresponding standard deviation are shown on Figure 5.8. It can be observed that the translation between the images becomes considerably large. Hence, the implemented animation would not be an advantegeous visualization method for the given use case.

5.5 Discussion

Although the tool is well suited for the visual analysis of the temporal development of thermo-radiative features, it still has several drawbacks.

One major flaw is certainly the high amount of user involvement in the processing pipeline, which can make the analysis of the time series of ther-
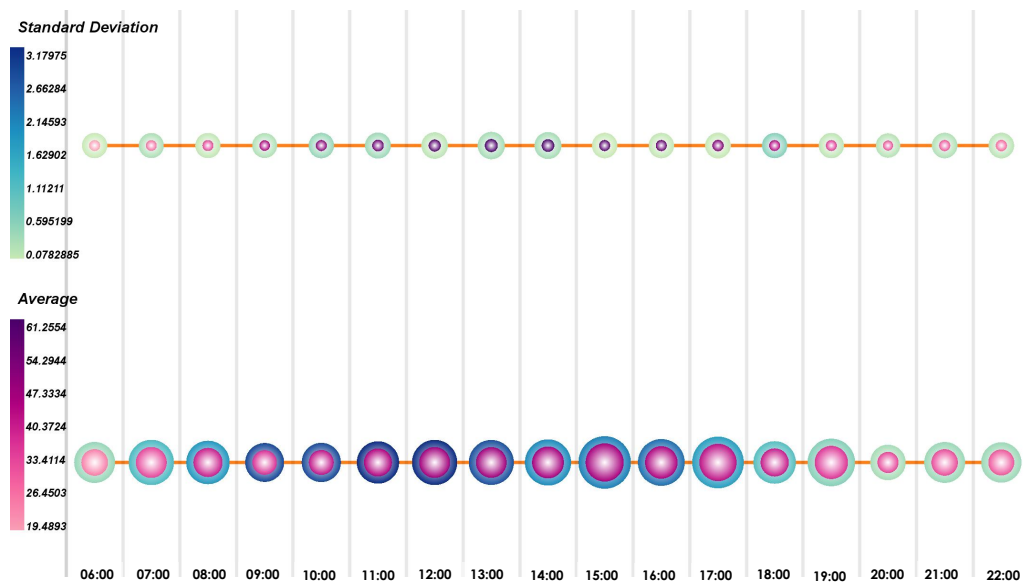
Figure 5.8: The graph that corresponds to the delineation of the surface area underneath the tree's canopy, corresponding to the image series in Figure 5.9. The graph represents the diurnal curse of the surface temperatures (times of day were added to the image for clarity, they are not generated by the tool).

mal images time-consuming. However, the interaction with the tool can lead to satisfactory results, since the user steers the main part of the processing procedure and is therefore able to adapt the feature selection to her or his individual research interests and to the quality of the images.

For some image configurations, the snakes are not the most advantageous delineation technique, as was shown in the use case described in Section 5.4.2. In this use case, this was mainly due to the fact that the feature of interest was not always fully captured by the infrared camera. Despite this limitation, it can be shown that the approach in principle still works for features, which are fully available on a low-resolution image. This is demonstrated in Figure 5.10. Furthermore, the tool described in this Chapter can still handle such cases by just leaving the choice to the user, whether the initial polygon should converge to the fragmented edges or not.

In the final animation, the temporal behavior of the highlighted thermo-radiative features is shown in the foreground of the smoothly blended original TIR images. Although the animation itself could be improved by a better identification of corresponding points on two subsequent snakes, the develop-
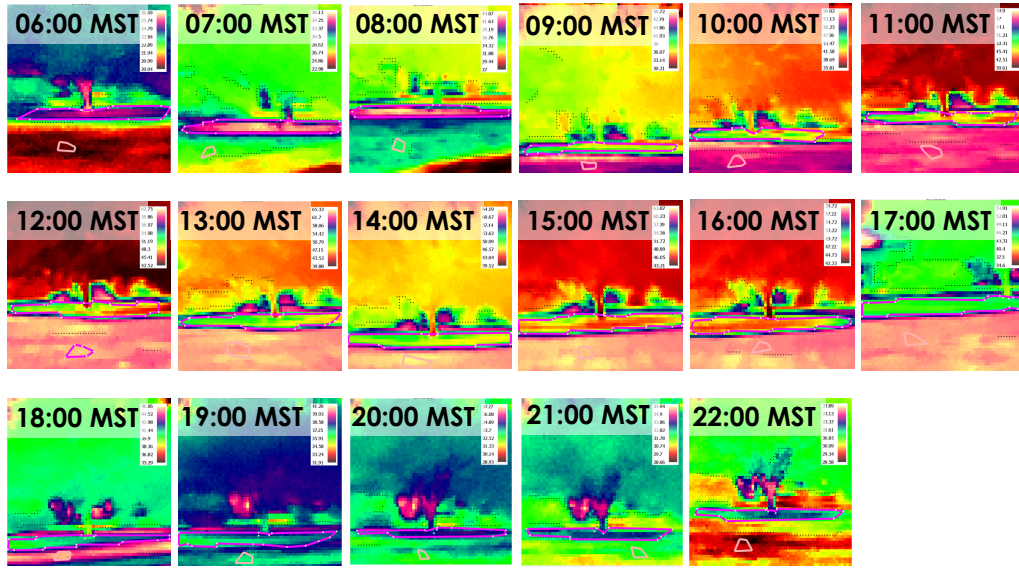
Figure 5.9: The image series, on which the graph in Figure 5.8 is based (times of day were added to the image for clarity, they are not generated by the tool). The translation between the images becomes large, which does not make an animation reasonable.

ment of the surrounding temperatures provides a holistic view for the user, who can follow the temperature profile of the monitored surface element simultaneously. On the other hand, it is disadvantageous that the analysis is currently only conducted on the infrared channel. This can lead to uncertainties in the analysis, since object boundaries might not be sufficiently salient. Therefore, the described tool needs a functionality to register a TIR image with a digital photograph taken of the same scene and at the same time to fuse the information from both channels. With a reliable registration, it also becomes possible to map the animated thermal features onto the photograph to provide spatial context. This is inspired by commercial TIR image analysis tools, such as, e.g., the FLUKE SmartView [50] or the FLIR Reporter software [49], in which the TIR images can be overlaid over the according digital photographs. In contrast to the functions provided in these programs, our tool aggregates temporal developments in the images supported by a focused animation and a directed graph by means of the methodology described in this Chapter.

A shortcoming of our tool is that common preprocessing steps for TIR images, such as atmospheric correction, still have to be carried out in external
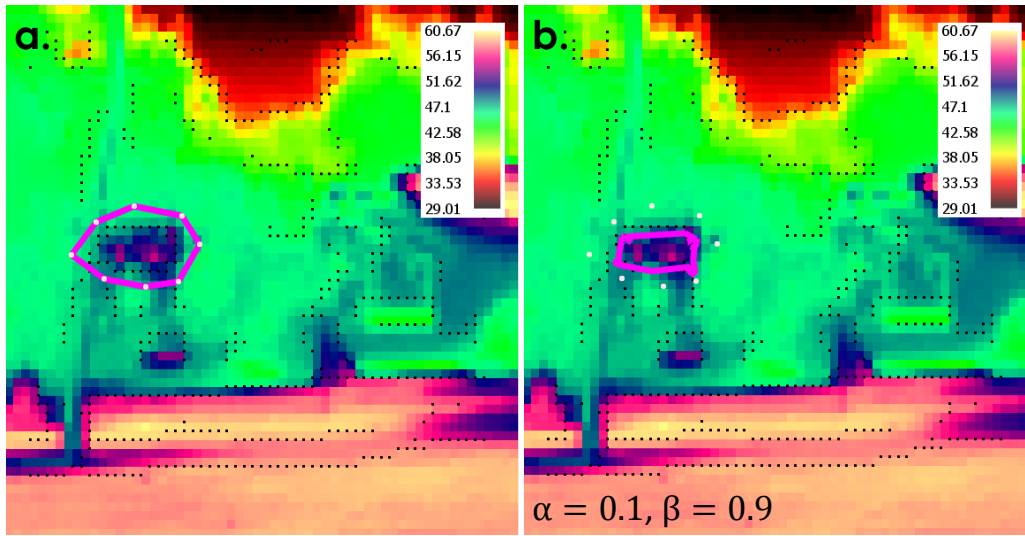
Figure 5.10: Even on ultra-low-resolution images, features can be delineated using snakes and GVF [85, 172], if they are salient enough. (a) shows the initial polygon, and (b) the converged snake. The image has a resolution of $60x60$ pixels.

software.

5.6 Conclusion

In this Chapter, I presented a tool that was developed for the visual analysis of a time series of TIR images. It was designed to assist expert users while investigating the thermal behavior of a monitored scene over the course of a certain time span. The analysis is conducted in a feature-based manner, i.e. salient hot spots, cold areas or other objects of interest are shown in their temporal development. To account for the users' individual research interest and for "difficult" data sets, they are involved in the main part of the processing pipeline by selecting and registering corresponding thermal features. The resulting visualization is two-fold: On the one hand, it includes a graph which serves as a summarizing visualization of the progression of topological events the features undergo. On the other hand, an animation dynamically shows the features' movement and deformation. Although the full variety of included capabilities is not necessarily needed for each application scenario, the tool can

still be used to generate a graph based on delineated features. This flexibility, together with the applied steering concept, guarantees for satisfactory results that can help to summarize a selection of thermo-radiative feature behaviors within a captured scene.

Chapter 6

Visualization in Building Performance Simulation Tools: Microclimate and Architecture

The research described in this Chapter narrows the scale down to single buildings or rooms. It looks at the currently used visualization techniques of microclimate data sets from the architectural perspective. In particular, this Chapter provides an initial and brief view into the state of the art in visualizing building performance simulation (BPS) results.

The capabilities and the usability of BPS tools has been addressed in the literature before. For example, Attia [14] and Attia *et al.* [15] conducted an extensive survey to identify the reasons of architects and engineers to use certain tools, but they do not specifically focus on the employed visualization techniques. Similarly, Pilgrim *et al.* [117] surveyed users of BPS tools about their satisfaction during all stages of the simulation process, from input to output analysis, but their findings only represent the viewpoint of HVAC engineers and structural engineers. With a clearer focus on data representation, Srivastav et al. [145] classify currently used techniques for the visualization of BPS results based on generalized tasks and data dimensionality.

In contrast, but also in extension to examples above, this Chapter analyzes tasks and requirements for the visualization capabilities of BPS tools using feedback from an architect, and evaluates state-of-the-art visualization strategies used for building performance simulation results based on the qual-

ity aspects derived before. Transferring the problem to the visualization discipline, it discusses the potential of utilizing other data representation and exploration techniques to meet the identified prerequisites. The work in this Chapter marked the starting point of a larger project, in which innovative visualization techniques for BPS results are developed in collaboration with Eva Hagen, an architect. It was conducted in collaboration with M.Sc. Diana Fernández Prieto, Dr.-Ing. Daniel Engel, Stephanie Schweitzer (at that time all affiliated with the Computer Graphics and HCI Group, University of Kaiserslautern), Dr.-Ing. Inga Scheler (Regional Computing Center, University of Kaiserslautern), and Michael Böttinger (German Climate Computing Center, Hamburg). My part in this project consisted of a literature research and the aggregation of the findings to a short-paper, which was published in the proceedings of PacificVis 2014 [64] and which this Chapter replicates.

6.1   Introduction
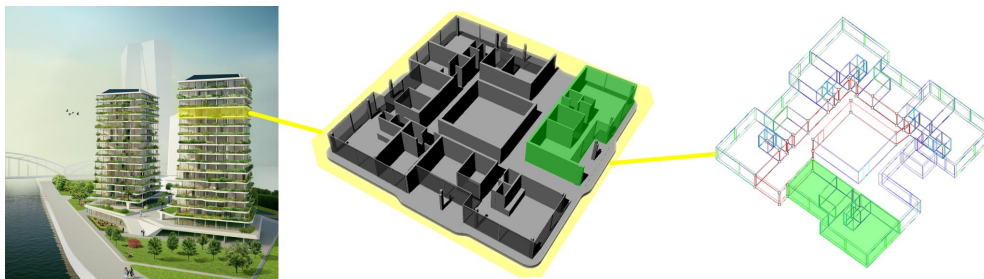
Figure 6.1:   The award-winning building design by Eva Hagen, part of which was used as a testbed for state-of-the-art BPS tools (marked in green).

While developing a design concept for a building, an architect has to consider several constraints: Administrative restrictions, client wishes, the behavior of potential users, the urban or rural environment, minimal greenhouse-gas emissions and optimized energy consumption have to be integrated into the final construction plan. Figure 6.1 (left) displays a complete example for a sustainable building design. Its features are intended to optimize the quality of life of the structure's inhabitants, while concurrently reducing the overall energy use. In order to quantify the effectiveness of this design approach, high-level commercial building performance simulation (BPS) tools can be used. This

kind of software is intended to support architects, engineers and researchers while creating or analyzing sustainable and occupant-friendly building design options (see, e.g., the product flyers of [40] and [81]). Hence, depending on their focus, they predict certain performance aspects of a building, such as energy use, indoor air quality or natural ventilation [36].

Despite the continuous improvement of BPS tools in the last decades [15], it has been reported in recent literature that the acceptance of such software in the architectural practice is relatively low (see, e.g., [39, 38, 145]), although their potential is indeed acknowledged [139]. This finding was also confirmed by the domain expert who was interviewed concerning this topic. One possible reason for this can be found in the significant gap between tool developers and engineers on the one hand and architects and designers on the other hand [38, 20, 39, 14]. This hints at the necessity to provide insightful visualizations of the generated data, which can help architects to understand the effects of their designs in terms of various performance aspects [145, 38]. More specifically, Srivastav et al. [145] point out the flaws in BPS visualization techniques, especially in the representation of spatiotemporal data stemming from CFD or daylight analyses. They therefore claim an investigation and improvement of existing visualization methods.

In this Chapter, state-of-the-art visualization approaches that are currently used for the exploration of building performance simulation results are exemplarily reviewed, without a claim for completeness. With feedback from an architect, the tasks and requirements that appropriate visualization techniques should ideally meet are identified. These quality criteria are then used to systematically evaluate domain-specific visualization strategies in both research and commercial tools. For this purpose, two exemplary BPS systems are tested against a part of the award-winning building design as visually outlined in Figure 6.1. The Chapter ends in an open discussion about the potential of other approaches and describe how basic research in visualization can help to close the gap between currently applied techniques and domain-specific needs.

This Chapter is structured as follows: In Section 6.2, the tasks and requirements for an appropriate visualization design are laid out. After that, the state of the art in both domain-specific visualization research as well as in commercial tools is assessed in Section 6.3. The findings of the previous two

Sections are discussed in Section 6.3, giving recommendations about potential solutions. Finally, the Chapter is concluded in Section 6.5.

6.2 Tasks and Requirements

6.2.1 Architectural workflow

A typical design process for a building consists of several stages. First, the architect has to conduct a site analysis. In this phase, important boundary conditions such as location, environmental factors, administrative constraints, and potential needs of the building users are assessed. With this knowledge at hand, a preliminary building plan can be developed, which is then refined and digitized. Iteratively, the building is optimized until a detailed construction plan is parameterized and realistically rendered (see Fig. 6.1 (left) for an example). After that, the architect builds a non-digital model which can be presented to the stakeholders.

In times of climate change, it also becomes particularly important to consider the reduction of energy consumption and greenhouse gas emissions as a major design factor. Hence, an initial design idea has to be evaluated not only in terms of function, aesthetics, and the surrounding environment, but also with respect to its environmental impact, energy performance, and indoor air quality. This yields an iterative process including the following tasks:

- *explore* the environment around a building and its impact on the surroundings
- *adapt* the design to a changing climate
- *construct* an initial digital model
- *interact* with the current building model to refine it iteratively
- *analyze* the impact of a particular design on energy use, thermal comfort or thermodynamic responses
- *optimize* a building design to meet all constraints
- *present* a refined design to stakeholders and promote its advantages, also in terms of performance.

6.2.2   Visualization requirements

Although the capabilities of building performance software increase continuously [15], three major reasons hinder the application of these tools in the design process: the difficulty to create a building model suited for performance simulations, the challenge of accurately representing complex geometries within the simulation, and the time needed to analyze the simulation results [56]. The last reason hints at a lack of appropriate and understandable visualization techniques [145, 56]. As a consequence, the question arises of how to improve the visualization techniques that are currently applied in the BPS domain. Based on a literature review and suggestions made by a domain expert, seven aspects can be identified:

Spatial context Since most data generated by building performance simulations is dependent on shape and space of a building, it is necessary that the visualization establishes a link to the current design [38, 15]. This also helps to uncover thermal relationships between different zones of the buildings [56].

Visual Feedback Architects need to be provided with a possibility to gain insight into the impact of single design decisions on the performance of the building [38, 145, 43, 14]. Therefore, it is necessary that an immediate visual feedback is established between design modifications and their specific effects. Ideally, it should be possible to change the design on-the-fly while examining the visualization, and to receive direct response with low interaction times.

Comparison of design alternatives Strongly related to visual feedback is the necessity to provide an environment, in which the effects of a set of related design settings that diverge in the values of certain building *parameters* can be compared to each other. This helps to contrast alternatives and supports the architect in finding an optimal solution [93, 56, 20, 15].

Spatiotemporal analysis In order to reveal spatiotemporal performance patterns within a specific building design, a holistic visualization has to be provided that incorporates the temporal dimension of the generated three-dimensional data set [145].

Suggestions to optimize the design Suggestions that lead the overall design concept into a sustainable direction can be helpful for an architect. Hence, optimization algorithms can be used to highlight possible improvements (see, e.g., [139]) or critical areas [93].

Insight into the thermodynamic system Arising from the knowledge gap between tool developers, engineers and physicists on the one hand and architects and designers on the other hand, it is necessary that the visualization of building performance simulation results also provides insight into the thermodynamic system established in a building context [38, 15].

Suitability for presentations Since the presentation of models and their advantages to various stakeholders plays an important role in the iterative refinement process of a building design, an appropriate visualization environment should also deliver a possibility to communicate simulation results in a visually pleasing and intuitive way [145, 43].

6.3   State of the art

In this section, state-of-the-art visualization strategies used for BPS results are exemplified by considering domain-specific research as well as commercial tools. Each technique described is evaluated by relating it to the requirements defined above.

6.3.1   Domain-specific research

For scalar simulation results, Bosworth and Pratt [20] propose a statistical approach to provide insight into the impact of design variations in parametric design. The authors conduct an ANOVA analysis to quantify the response of particular design parameters on a targeted performance variable. The results are then displayed using parallel coordinates and a voxel plot. In a third window, the CAD model of a selected design variant is shown via brushing and linking. This approach is well suited for parametric design studies and establishes a link between quantitative analyses and qualitative design options. Based on a simulated parameter, the visualization can give hints about an optimal design in terms of the performance variable under investigation. Related to this is the automatic file-sorting system proposed by Dondeti and Reinhart [42]. It supports the architect in managing even a large amount of different design scenarios. Performance measures of single- or cross-project variants can be compared quickly via bar charts, but no visual relationship to the particular building design is provided. This hinders quick insight into the effect of certain

design decisions.

Visualizations of spatiotemporal data are frequently provided in daylight analyses. Lagios et al. [93] couple the Rhinoceros/Grasshopper environment with a daylight simulation tool to facilitate the comparison between a large set of design scenarios in terms of daylight performance. They visualize the results using an animation that shows the amount of daylight within the spatial context of the building. Furthermore, critically overlit areas are highlighted, facilitating optimization [93]. However, although animations are well-suited for presentations, they aggravate a quantitative comparison of design alternatives [2]. A detailed spatiotemporal visualization framework for daylight analyses is provided by Glaser and Ubbelohde [54]. Via brushing and linking, a 2D temporal plot of daylight hours over the course of a year is connected to a low-resolution spatial distribution display of daylight within a building level. Although this gives a summarizing overview over the daylight performance of a certain design, no hints at possible optimizations are given. Furthermore, a link to the whole 3D building model is missing.

In current research, there are also efforts to automatically optimize a given design scenario. An example for this can be found in [139]. Here, the authors provide a framework to automatically and iteratively generate designs that optimize performance aspects such as solar gain on a roof, energy use or daylighting. The optimal design is visualized in Rhinoceros, such that the architect can keep control over the result. However, only one performance aspect can be targeted at a time, which makes it difficult to balance a variety of demands.

Prazeres and Clark [119, 118] present a flexible system that provides an integrated view on several aspects investigated during building simulations. They combine multiple perspectives on the data, including spatial plots, barcharts, sounds or animations that help the user to investigate a large set of performance aspects and to quickly compare different design alternatives. Although the data can be explored interactively, the system does not seem to enable direct design modifications. In contrast to this, Greenberg et al. [56] describe a framework that incorporates several modules such as a CAD platform, a building energy model and a building simulation. All of these modules have interfaces that can be steered by the user to enable quick visual feedback

for any changes in the building design. The visualization within the proposed system provides several views on the generated data which include temporal energy use, the current building design, weather data or heat loss/gain on particular surfaces. They acknowledge the importance of spatial context and are currently investigating according visualization methods, e.g. to show thermal relationships between parts of a building or to compare the performance of different design alternatives. However, it is not completely clear, how exactly spatiotemporal data from thermal simulations are displayed.

Part of the described domain-specific visualization research is conducted by members of architecture departments. This shows that there is a need to improve existing visualization strategies to adapt it to the typical tasks in this application area. Despite of the described efforts, there is still potential for improvement, especially in terms of exploration, interaction and spatiotemporal analysis of the generated data.

6.3.2 Visualization in commercial tools

BPS tools have the goal to support decision-making in terms of an energy-efficient configuration of a building's design and material parameters [40, 81] and, thus, to support a subset of the tasks in the architectural workflow. Among architects, the top-ranked tools are IES-VE [81], Ecotect [16] and DesignBuilder [40] because of their user-friendliness and data exchange capabilities, inter alia [14]. In order to exemplarily investigate current visualization options for spatial data such as air temperature or wind speed and direction, DesignBuilder (version 3.2.0.070) [40] and IES-VE (version 2013.0.0.2) [81] were chosen, since both of them incorporate the possibility to run a CFD analysis. Although this option mainly targets engineers, it is hypothesized that architects are also interested in the thermodynamic system that establishes in and around their designed building, since it is strongly related to thermal comfort and other performance aspects. As a basis to evaluate both commercial tools, an individual apartment is extracted from an award-winning building model: The state-of-the-art mixed-use residential tower designed by Eva Hagen (Fig. 6.1), which is the third prize winner of the building industry award of Rheinland-Pfalz in 2012. Table 6.1 summarizes the divergence between the above defined requirements and the capabilities of both tools.

Wind Speed	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.54	m/s
Air Temp.								20.97	21.61	22.26	22.91	23.55	°C
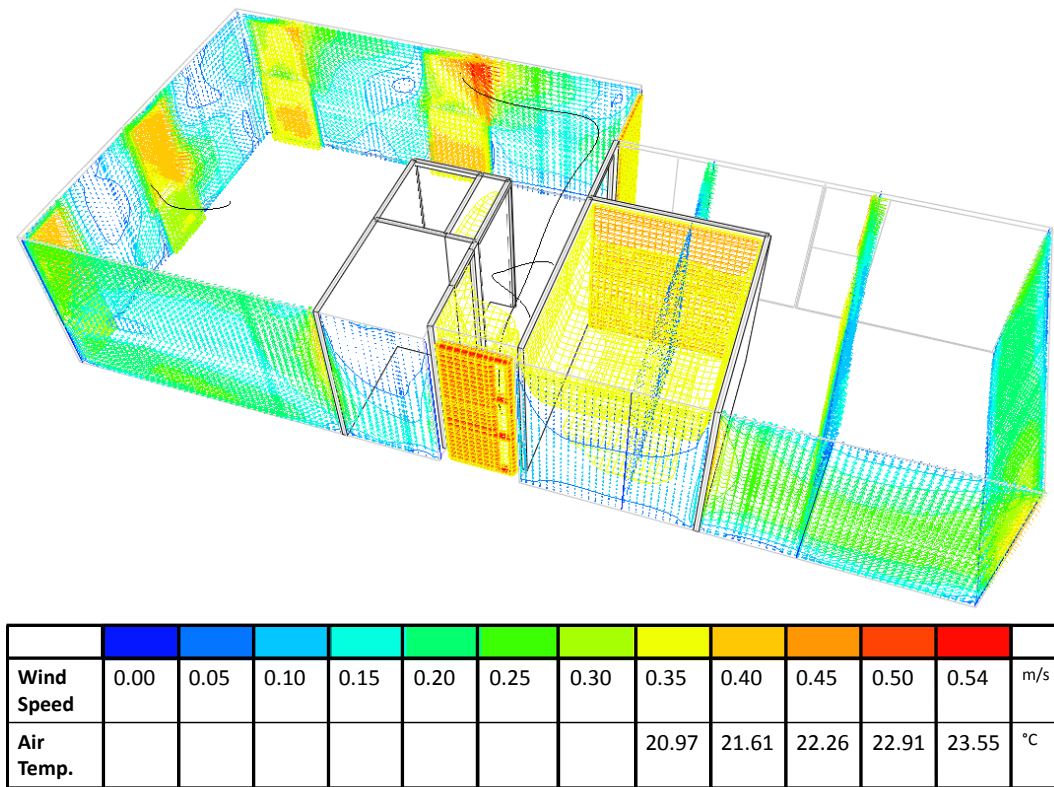
Figure 6.2: A combined visualization example (created with IES-VE): Isosurfaces for air temperature (gridded surfaces), cutting planes with color-coded vector glyphs for air flow and contour lines for wind speed, and tracked particles.

For the CFD results, the visualization options are very similar in both tools. In the spatial context of the imported CAD model, a user can choose between cutting planes, isosurfaces, and arrow-shaped vector glyphs (Fig. 6.2). In addition, the IES-VE package provides a particle tracking option. The simulation results for the interior are displayed within the investigated building components, whereas features such as walls or windows are by default shown in a wireframe mode. Texture-based rendering for these building features can also be applied, as it is advantageous for the representation of exterior CFD results.

The cutting planes, which can be placed into the simulation domain, show the local temperature and airflow situation for one snapshot at a time. They represent the interior circulation components if the slices are dispersed equally throughout the area. However, a large number of slices within the domain lead to clutter and occlusion. If an insufficient amount of slices is chosen,

Requirement	DesignBuilder	IES-VE
Spatial context	Available	Available
Visual feedback	Iteratively available	Iteratively available
Comparison of design alternatives	**Available** via a parametric design option	**Available** via a parametric design option
Spatiotemporal analysis	**Iteratively available** by selecting one snapshot after the other (CFD). **Available** for solar analysis via animation.	**Available** via particle-tracking and animation (solar analysis); **iteratively available** by selecting one snapshot after the other (CFD).
Suggestions to optimize the design	**Partially available** (e.g., highlighting of overlit / underlit areas)	**Partially available** (e.g., highlighting of visual discomfort areas caused by glare (Radiance))
Insight into the thermodynamic system	**Partially available** with basic visualization techniques	**Partially available** with basic visualization techniques
Suitability for presentations	**Available** with basic visualization techniques	**Available** with basic visualization techniques

Table 6.1: Capabilities of DesignBuilder / IES-VE as opposed to the requirements resulting from an architectural workflow.

important thermal or fluid features may be skipped accidentally. In both DesignBuilder and IES-VE, this concern can partially be eliminated by using isosurfaces to display the continuous distribution of certain parameter levels. Since the surfaces are designed to be transparent and color-coded, several layers of isosurfaces can be shown at once (Fig. 6.3). Clutter problems can be tackled by reducing the number of isosurface thresholds, but the user does not know, which of the available values would provide the most insight. Thus, the problem of accidentally missing important features is still existent. In IES-VE, the possibility of tracking particles through an interior building component is a further attempt of creating a certain amount of continuity within the visualization. Air movement combined with another scalar variable at focused locations within the simulation domain can be shown comprehensively with this method. On the other hand, separate launching of each particle can be time-consuming, and missing important features is consequently more likely to occur here than in the other visualization options.

The basic visualization techniques applied for the representation of the CFD-results fulfill the requirement to show the data within their original spatial context. However, both tools lack the possibility to explore the data according to other relevant aspects. It is thus not possible to conduct a quantitative spatial comparison of the CFD results between different climate/weather scenarios or design variants by means of a comprehensive visualization interface. This makes an iterative refinement of a specific design difficult.

Air Temp.	16.45	17.10	17.74	18.39	19.03	19.68	20.32	20.97	21.61	22.26	22.91	23.55	°C
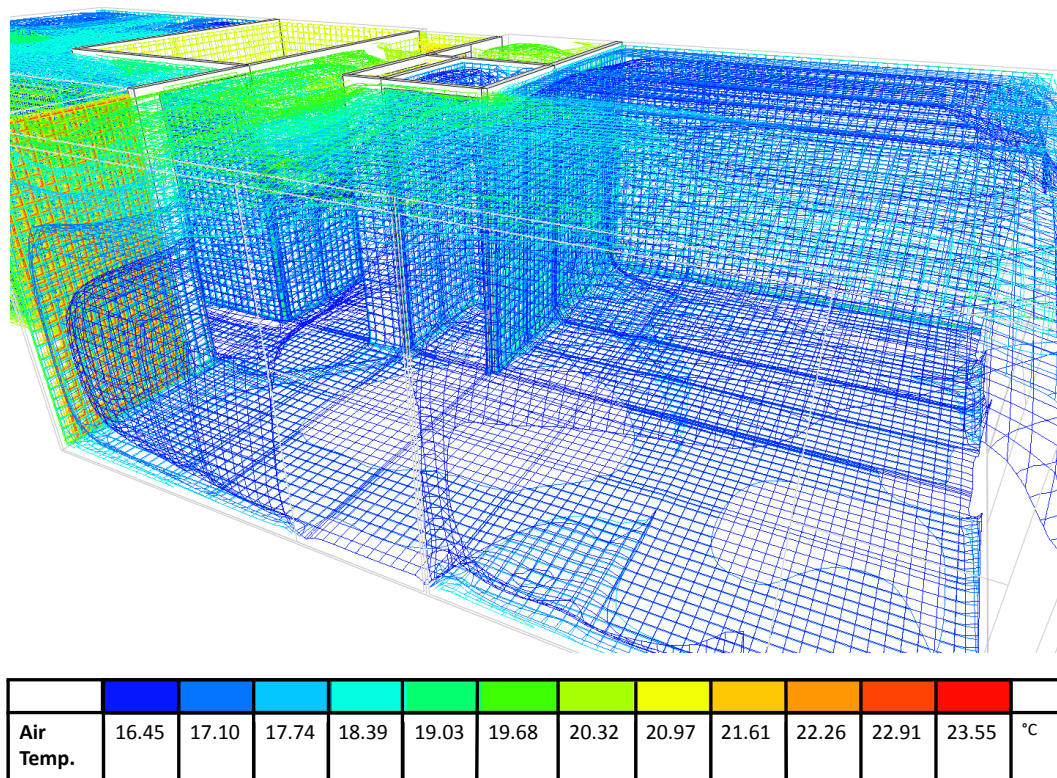
Figure 6.3: Isosurfaces for air temperature (created with IES-VE).

6.4 Discussion: Visualization techniques with application potential for BPS

This Section discusses how certain visualization strategies can be used to meet the requirements outlined above.

Spatial context / visual feedback / suitability for presentations
Space and shape are considered extremely important in architectural design

[139]. For the final visualization of BPS results, it is therefore advantageous to embed the visual representation of the generated scalar- or vector fields into a realistically rendered context of the original CAD building model (similar to [93] or, e.g., DesignBuilder). This way, computed values can directly be related to building features and materials. This enhances the comprehension about the impact of specific design decisions on the thermodynamic system that establishes within and around a building [38]. At the same time, it helps to communicate these findings in an intuitive way. Immediate visual feedback between design changes and according simulation result adaptation can even enhance the required comprehension of the effects resulting from design changes [145, 14].

Comparison of design alternatives In order to analyze the outcomes of simulations for different design variants, *multirun* or *ensemble* visualization techniques can be used. These are frequently based on statistical representations, as they result from clustering, dimensionality reduction or simple summary statistics [86]. However, spatial context is still important and should be accessible on demand, while statistical summaries simultaneously support the architect to gain an overview about the impact of varying a certain design parameter. It is therefore advantageous to provide support for different interaction tasks (see [173] for examples) as well as views that combine both spatial and abstract views on the generated data (e.g. multiple coordinated views, see [126]). This approach is also partly present in state-of-the-art visualizations that have been specifically tailored to parametric design (see, e.g., [20] for brushing and linking of scalar data). Besides performance assessments for different *designs*, these views can also be used to compare the behaviour of one specific building model under different climate change scenarios.

Spatiotemporal analysis Srivastav et al. [145] identify the need to incorporate the temporal dimension into the visual representation of spatial building performance simulation results to comprehensively analyze performance *patterns*. There exists a large variety of visualization approaches for multitemporal data representations and analyses, with common strategies being animation, interactive visualization or aggregated, still images [86, 2]. Each of these categories can be useful for the discussed application field. Thus, animations are especially suited for presentation purposes [2]. Interactive visualizations,

e.g. using brushing and linking, can help to provide an overview over temporal developments as well as a detailed view on specific, maybe critical, time steps. Aggregated, still images can help to represent interesting temporal patterns within a spatial domain. For CFD analyses in the BPS context, flow visualization techniques can be used to facilitate the exploration of interior or exterior air flow patterns.

Suggestions to optimize the design Research in the BPS domain already attempted to focus on the automated optimization of building designs (see, e.g., [139]) or to visualize critical areas within a planned structure [93] via highlighting techniques. For the latter, feature-based visualization techniques for scalar- or vector fields can be applied to solve this task, with features being critical areas within the three-dimensional space of a building design. The problem, however, is the definition and parameterization of critical areas, which is strongly dependent on the specific design intention and the performance question asked. Therefore, a high level of flexibility and interaction has to be added to a potential visualization framework.

Insight into the thermodynamic system can only be achieved by showing the connections between the simulated variables. Multivariate analysis can help to achieve this goal, since it is concerned with showing the interrelations between multiple computed fields [86]. For the visualization of BPS results, addressed to an architect or designer, multivariate analysis techniques that reduce the number of variables to meaningful measures in terms of thermal comfort (which is already adressed in state-of-the-art analysis tools) or energy use or that provide a spatial view on correlations are considered most valuable.

6.5   Conclusion

In this Chapter, visualization techniques used for BPS results are reviewed and related to the tasks and requirements derived from the architectural workflow. None of the reviewed visualization strategies combine all of the identified necessities, neither in the examined commercial tools nor in recent domain-specific visualization research. Therefore, the potential of other visualization approaches to close the gap between currently used techniques and the prerequisites that make BPS valuable in practice are discussed.

The preliminary investigation described in this Chapter indicates opportunities for improvement in current visualization techniques used in building performance simulations, especially for the analysis of three-dimensional scalar or vector fields stemming from more engineer-oriented CFD analyses. In the reviewed BPS solutions, only basic visualization techniques are applied so far. Although these are partially useful for the practical, architectural workflow, future directions in visualization should focus on creating an explorative environment in which design *experiments* can be conducted by an architect, as claimed by [38]. A future-oriented building model should also be analyzed in terms of its performance under different climate change scenarios. These aspects imply a high amount of interaction with both designs and data under fast and intuitive change-response times, while highlighting critical or otherwise interesting areas in the spatial domain. Thus, a high potential is seen in combining multiple coordinated views (including interactive links to the spatial dimension of a design solution) with feature-based visualization. A larger project, for which this survey was conducted, is already addressing these options.

Chapter 7

Conclusion

The overarching goal of the work described in this dissertation was the application and adaptation of visualization and analysis techniques to urban microclimate data sets, including observational data from thermal infrared imagery and mobile measurements. Furthermore, going beyond the urban microclimate research context, the state of the art in the visualization of building performance simulation results was assessed.

Thermal infrared imagery with a hand-held device can, in some cases, be challenging to analyze, especially if the images are time-varying, have a low resolution, and are sampled from varying viewpoints. To approach these challenges, a tool was implemented to assist manual feature extraction and to visually summarize the results of the extraction results, combining well-known image processing and visualization techniques. In two use cases, the flexibility of the tool in terms of image resolution and analysis goal could be demonstrated.

Relating the exploration of (urban) microclimate data to the architectural domain, the state of the art in the visualization of building performance simulation results was assessed in another project. This investigation showed that currently available visualization techniques do not necessarily meet all requirements needed to support architects during the planning process of sustainable buildings. This finding served as a starting point for a larger project aiming at the extension of existing exploration techniques, which is currently conducted in the Computer Graphics and HCI Group.

Besides these two smaller contributions, the main project of this thesis deals

with supporting the interactive exploration of mobile transect measurements by developing a tool that includes facilities for data preprocessing, visualization, and analysis. Combining geovisual analytics and meteorology concepts, I implemented a prototypic software that supports the investigation of multivariate relationships and places the spatially dependent observations into context with a sensor's field of view.

While the tool can be used to analyze one transect run in detail, it also provides a workflow for the spatial aggregation of multiple transect runs over diverse routes. In contrast to domain-specific aggregation measures, which usually include statistical summary values such as the average, I designed a glyph-based visualization that provides additional information about multivariate relationships within a spatial aggregation department, while also including information about meteorological background conditions at observation time and the summary's reliability. Coloring the background of the aggregation compartments according to the similarity of the glyphs could enhance the structure in the final visualization, highlighting climatic microenvironments according to a user-defined set of observation variables.

Staying in the context of mobile measurements, I described an improved method for the correction of air temperature measurements sampled with slow sensors. Building on previously known techniques, I determined and evaluated suitable parameter choices by comparing the corrected time series to data sampled with a faster air temperature sensor. In nearly all cases, the agreement between these time series could be improved through correction, meaning that the spatial resolution of the measurements could be increased through the described correction procedure.

I developed TraVis in close collaboration with domain experts at the Arizona State University, who continously evaluated the design decisions made and the concepts used. I also presented components of the tool at international conferences for Geography [59] and Meteorology [62], where I received positive feedback from other scientists working in application-related fields. This demonstrates that the described prototype not only contributes to the field of geospatial visual analytics, but also to the application domain it was designed for.

Bibliography

[1] C. Achberger and L. Bärring. Correction of surface air temperature measurements from a mobile platform. *Agricultural and Forest Meteorology*, 98–99(0):227–238, 1999.

[2] W. Aigner, S. Miksch, W. Mueller, H. Schumann, and C. Tominski. Visualizing time-oriented data – a systematic view. *Computers & Graphics*, 31(3):401 – 409, 2007.

[3] G. Andrienko and N. Andrienko. Spatio-temporal aggregation for visual analysis of movements. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 51–58, Oct 2008.

[4] G. Andrienko and N. Andrienko. A general framework for using aggregation in visual exploration of movement data. *The Cartographic Journal*, 47(1):22–40, 2010.

[5] G. Andrienko, N. Andrienko, C. Hurter, S. Rinzivillo, and S. Wrobel. From movement tracks through events to places: Extracting and characterizing significant places from mobility data. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on*, pages 161–170, Oct 2011.

[6] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi, and F. Giannotti. Interactive visual clustering of large collections of trajectories. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on*, pages 3–10, Oct 2009.

[7] N. Andrienko and G. Andrienko. Spatial generalization and aggregation of massive movement data. *Visualization and Computer Graphics, IEEE Transactions on*, 17(2):205–219, Feb 2011.

136

[8] N. Andrienko, G. Andrienko, N. Pelekis, and S. Spaccapietra. Basic concepts of movement data. In F. Giannotti and D. Pedreschi, editors, *Mobility, Data Mining and Privacy*, pages 15–38. Springer Berlin Heidelberg, 2008.

[9] N. Andrienko, G. Andrienko, H. Stange, T. Liebig, and D. Hecker. Visual analytics for understanding spatial situations from episodic movement data. *KI - Künstliche Intelligenz*, 26(3):241–251, 2012.

[10] Apogee Instruments, Inc. Infrared radiometers: SI-100 series. Online, 2013. `http://www.apogeeinstruments.co.uk/content/infrared-sensor-specs.pdf`, Last accessed: 05/09/2015.

[11] Apogee Instruments, Inc. and A. Del Toro. Field of view. Online, 2015. `http://www.apogeeinstruments.co.uk/field-of-view/`, Last accessed: 05/09/2015.

[12] J. P. Ardila, W. Bijker, V. A. Tolpekin, and A. Stein. Multitemporal change detection of urban trees using localized region-based active contours in VHR images. *Remote Sensing of Environment*, 124(0):413 – 426, 2012.

[13] A. J. Arnfield. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. *International Journal of Climatology*, 23(1):1–26, 2003.

[14] S. Attia. Building performance simulation tools: Selection criteria and user survey. research based report. Université Catholique de Louvain: Louvain La Neuve, Belgium, 2010.

[15] S. Attia, L. Beltrán, A. D. Herde, and J. Hensen. 'architect friendly': A comparison of ten different building performance simulation tools. In *Proceedings of the Eleventh International IBPSA Conference, July 27-30, 2009, Glasgow, Scotland*, 2009.

[16] Autodesk, Inc. Autodesk Ecotect Analysis. `http://www.autodesk.de/adsk/servlet/pc/\index?siteID=403786&id=15073595`. Accessed: 2013-10-29.

[17] P. Bak, F. Mansmann, H. Janetzko, and D. Keim. Spatiotemporal analysis of sensor logs using growth ring maps. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):913–920, Nov 2009.

[18] C. Balaras and A. Argiriou. Infrared thermography for building diagnostics. *Energy and Buildings*, 34(2):171 – 183, 2002.

[19] K. Blennow and P. Persson. Modelling local-scale frost variations using mobile temperature measurements with a GIS. *Agricultural and Forest Meteorology*, 89(1):59 – 71, 1998.

[20] D. E. Bosworth and K. B. Pratt. New methods for the construction and interpretation of high dimensional parametric building energy models. In *Proceedings of the Fifth National Conference of IBPSA-USA, August 1-3, 2012, Madison, Wisconsin*, 2012.

[21] E. Boufidou, T. J. F. Commandeur, S. B. Nedkov, and S. Zlatanova. Measure the climate, model the city. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-4/C21:59–66, 2011.

[22] A. J. Brazel and D. M. Johnson. Land use effects on temperature and humidity in the Salt River Valley, Arizona. *Urban Forestry & Urban Greening*, 15(2):54–61, 1980.

[23] R. D. Brown. Ameliorating the effects of climate change: Modifying microclimates through design. *Landscape and Urban Planning*, 100(4):372 – 374, 2011. Landscape and Urban Planning at 100.

[24] M. Buttstädt, T. Sachsen, G. Ketzler, H. Merbitz, and C. Schneider. A new approach for highly resolved air temperature measurements in urban areas. *Atmospheric Measurement Techniques Discussions*, 4(1):1001–1019, 2011.

[25] Campbell Scientific. HMP45C temperature and relative humidity probe brochure. Online, 2011. `https://s.campbellsci.com/documents/us/product-brochures/b_hmp45c.pdf`, Last accessed: 05/03/2015.

[26] Campbell Scientific. FW05, FW1, and FW3 Type E fine wire thermocouples brochure. Online, 2012. `http://s.campbellsci.com/documents/us/product-brochures/b_fw05-fw3.pdf`, Last accessed: 08/11/2014.

[27] Campbell Scientific. CR1000 measurement and control system: Operators manual. Online, 2013. `http://s.campbellsci.com/documents/us/manuals/cr1000.pdf`, Last accessed: 08/11/2014.

[28] Campbell Scientific. HC2S3 temperature and RH sensor brochure. Online, 2014. `http://s.campbellsci.com/documents/us/product-brochures/b_hc2s3.pdf`, Last accessed: 08/11/2014.

[29] Campbell Scientific. LP02 solar radiation sensor. Online, 2015. `https://s.campbellsci.com/documents/us/product-brochures/b_lp02.pdf`, Last accessed: 05/09/2015.

[30] J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679–698, 1986.

[31] Y. Chambenoit, E. Trouve, N. Classeau, J.-P. Rudant, and P. Bolon. Different fusion strategies to detect geographical objects by active contours in multitemporal SAR images. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International*, volume 1, pages –100, 2004.

[32] F. Chen, R. Bornstein, S. Grimmond, J. Li, X. Liang, A. Martilli, S. Miao, J. Voogt, and Y. Wang. Research priorities in observing and modeling urban weather and climate. *Bulletin of the American Meteorological Society*, 93(11):1725–1728, 2012.

[33] W. T. Chow, R. L. Pope, C. A. Martin, and A. J. Brazel. Observing and modeling the nocturnal park cool island of an arid city: horizontal and vertical impacts. *Theoretical and Applied Climatology*, 103(1-2):197–211, 2011.

[34] A. Christen, F. Meier, and D. Scherer. High-frequency fluctuations of surface temperatures in an urban environment. *Theoretical and Applied Climatology*, 108(1-2):301–324, 2012.

[35] A. Chudnovsky, E. Ben-Dor, and H. Saaroni. Diurnal thermal behavior of selected urban objects using remote sensing measurements. *Energy and Buildings*, 36(11):1063 – 1074, 2004.

[36] D. B. Crawley, J. W. Hand, M. Kummert, and B. T. Griffith. Contrasting the capabilities of building energy performance simulation programs. *Building and Environment*, 43(4):661 – 673, 2008. Part Special: Building Performance Simulation.

[37] T. N. Dang, L. Wilkinson, and A. Anand. Stacking graphic elements to avoid over-plotting. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1044–1052, Nov 2010.

[38] C. B. de Souza. Contrasting paradigms of design thinking: The building thermal simulation tool user vs. the building designer. *Automation in Construction*, 22(0):112 – 122, 2012. Planning Future Cities-Selected papers from the 2010 CAAD Conference.

[39] C. B. de Souza. Studies into the use of building thermal physics to inform design decision making. *Automation in Construction*, 30(0):81 – 93, 2013.

[40] DesignBuilder Software Ltd. DesignBuilder, version 3.2.0.070. `http://www.designbuilder.co.uk/`. Accessed: 2013-10-29.

[41] Deutscher Wetterdienst. Der Deutsche Wetterdienst ist dem Klima in Halle (Saale) auf der Spur... Online, 2015. `http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=P27200165321293012986287&T176000465321293013118855gsbDocumentPath=Navigation%2FOeffentlichkeit%2FKlima__Umwelt%2FKlima__Zukunft%2FKlimaproj__Stadt%2FProjekt__Halle%2FHalle__Start__node.html%3F__nnn%3Dtrue`, Last accessed: 05/07/2015.

[42] K. Dondeti and C. F. Reinhart. A 'Picasa' for bps – An interactive data organization and visualization system for building performance simulations. In *Proceedings of the 12th Conference of IBPSA, November 14-16, 2011, Sydney, Australia*, 2011.

[43] M. R. Donn. A survey of users of thermal simulation programs. In *IBPSA Proceedings: Building Simulation Conference, September 8-10, 1997, Prague, Czech Republic*, 1997.

[44] M. J. Downs, D. H. Ferriss, and R. E. Ward. Improving the accuracy of the temperature measurement of gases by correction for the response delays in the thermal sensors. *Measurement Science and Technology*, 1(8):717, 1990.

[45] B. Elen, J. Peters, M. V. Poppel, N. Bleux, J. Theunis, M. Reggente, and A. Standaert. The Aeroflex: A bicycle for mobile air quality measurements. *Sensors*, 13(1):221–240, 2012.

[46] Environmental Remote Sensing and Geoinformatics Lab and CAP LTER. 4 band NAIP land classification of Central Arizona, 2012. Arizona State University.

[47] Z. Fang, T. Möller, G. Hamarneh, and A. Celler. Visualization and exploration of time-varying medical image data sets. In *Proceedings of Graphics Interface 2007*, GI '07, pages 281–288, New York, NY, USA, 2007. ACM.

[48] G. Farin. *Curves and Surfaces for CAGD: A Practical Guide*. Academic Press, Inc., London, UK, 4th edition, 1997.

[49] FLIR® Systems. FLIR® Reporter, 2011. `http://www.flir.com/cs/emea/de/view/?id=42405`, Last accessed: 09/02/2013.

[50] FLUKE® Corporation. FLUKE® SmartView®, 2011. `http://www.fluke.com/fluke/usen/accessories/software/fluke-smartview-ir.htm?PID=56169`, Last accessed: 09/02/2013.

[51] T. Foken. *Angewandte Meteorologie: Mikrometeorologische Methoden*. Springer, Berlin/Heidelberg, 2006.

[52] S. Foster and P. Chan. Improving the wind and temperature measurements of an airborne meteorological measuring system. *Journal of Zhejiang University SCIENCE A*, 13(10):723–746, 2012.

[53] GDAL Development Team. *GDAL - Geospatial Data Abstraction Library, Version 10.0.1.* Open Source Geospatial Foundation, 2013.

[54] D. C. Glaser and M. S. Ubbelohde. Visualization for time dependent building building simulation. In *Proceedings of the Seventh Conference of IBPSA, August 13-15, 2001, Rio de Janeiro, Brazil*, 2001.

[55] J. S. Golden, J. Carlson, K. E. Kaloush, and P. Phelan. A comparative study of the thermal and radiative impacts of photovoltaic canopies on pavement surface temperatures. *Solar Energy*, 81(7):872 – 883, 2007.

[56] D. Greenberg, K. Pratt, B. Hencey, N. Jones, L. Schumann, J. Dobbs, Z. Dong, D. Bosworth, and B. Walter. Sustain: An experimental test bed for building energy simulation. *Energy and Buildings*, 58(0):44 – 57, 2013.

[57] D. Guo, J. Chen, A. M. MacEachren, and K. Liao. A visualization system for space-time and multivariate patterns (VIS-STAMP). *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1461–1474, Nov. 2006.

[58] K. Häb, N. H. Feige, L. S. Hüttenberger, A. Middel, and H. Hagen. Visualizing the temporal development of thermo-radiative features on ground-based thermographs. *Environmental Earth Sciences*, 72(10):3781–3793, 2014.

[59] K. Häb, A. Middel, and Ruddell. Classifying urban microclimate patches based on multitemporal mobile transect observations. In *Annual Meeting of the Association of American Geographers, Chicago 2015*, April 2015. Abstract and Talk.

[60] K. Häb, A. Middel, B. Ruddell, and H. Hagen. TraVis - a visualization framework for mobile transect data sets in an urban microclimate context. In *Visualization Symposium (PacificVis), 2015 IEEE Pacific*, pages 167–174, April 2015.

[61] K. Häb, A. Middel, B. L. Ruddell, and H. Hagen. Spatial Aggregation of Mobile Transect Measurements for the Identification of Climatic Microenvironments. In A. Middel, K. Rink, and G. H. Weber, editors,

Workshop on Visualisation in Environmental Sciences (EnvirVis), pages 19–23. The Eurographics Association, 2015.

[62] K. Häb, A. Middel, B. L. Ruddell, and H. Hagen. Visual analytics and microclimate analysis: A use case for a visualization tool developed for mobile measurements. In *31st Conference on Environmental Information Processing Technologies at the 95th Annual Meeting of the American Meteorological Society, Phoenix 2015*, April 2015. Abstract and Talk.

[63] K. Häb, B. L. Ruddell, and A. Middel. Sensor lag correction for mobile urban microclimate measurements. *Urban Climate*, 2015. Under rewiev.

[64] K. Häb, S. Schweitzer, D. Fernandez Prieto, E. Hagen, D. Engel, M. Böttinger, and I. Scheler. Visualization of building performance simulation results: State-of-the-art and future directions. In *Pacific Visualization Symposium (PacificVis), 2014 IEEE*, pages 311–315, March 2014.

[65] R. Hagemann, U. Corsmeier, C. Kottmeier, R. Rinke, A. Wieser, and B. Vogel. Spatial variability of particle number concentrations and NOx in the Karlsruhe (Germany) area obtained with the mobile laboratory 'AERO-TRAM'. *Atmospheric Environment*, 94(0):341 – 352, 2014.

[66] J. Han and B. Bhanu. Fusion of color and infrared video for moving human detection. *Pattern Recognition*, 40(6):1771 – 1784, 2007.

[67] M. Harrower and C. Brewer. ColorBrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, Jun 2003.

[68] D. Hartz, L. Prashad, B. Hedquist, J. Golden, and A. Brazel. Linking satellite images and hand-held infrared thermography to observed neighborhood climate conditions. *Remote Sensing of Environment*, 104(2):190 – 200, 2006.

[69] D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele. Participatory air pollution monitoring using smartphones. In *Mobile Sensing: From Smartphones and Wearables to Big Data*, Beijing, China, Apr 2012. ACM.

[70] D. Hasenfratz, O. Saukh, C. Walser, C. Hueglin, M. Fierz, and L. Thiele. Pushing the spatio-temporal resolution limit of urban air pollution maps. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*, pages 69–77, March 2014.

[71] C. Helbig, H.-S. Bauer, K. Rink, V. Wulfmeyer, M. Frank, and O. Kolditz. Concept and workflow for 3d visualization of atmospheric data in a virtual reality environment for analytical approaches. *Environmental Earth Sciences*, 72(10):3767–3780, 2014.

[72] S. Henninger. Preventing a mass panic by micrometeorological measurements. *8th International Conference on Urban Climates*, August 2012. Dublin, Ireland.

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

[74] V. Heuveline, S. Ritterbusch, and S. Ronnas. Augmented reality for urban simulation visualization. *Preprint Series of the Engineering Mathematics and Computing Lab*, 0(16), 2011.

[75] J. Hidalgo, V. Masson, A. Baklanov, G. Pigeon, and L. Gimeno. Advances in urban climate modeling. *Annals of the New York Academy of Sciences*, 1146(1):354–374, 2008.

[76] W. N. Hien, T. P. Yok, and C. Yu. Study of thermal performance of extensive rooftop greenery systems in the tropical climate. *Building and Environment*, 42(1):25 – 54, 2007.

[77] J. Hoschek. Spline approximation of offset curves. *Computer Aided Geometric Design*, 5(1):33 – 40, 1988.

[78] A. Hoyano, K. Asano, and T. Kanamaru. Analysis of the sensible heat flux from the exterior surface of buildings using time sequential thermography. *Atmospheric Environment*, 33(24–25):3941–3951, 1999.

[79] J. Hübner, J. Olesch, H. Falke, F. X. Meixner, and T. Foken. A horizontal mobile measuring system for atmospheric quantities. *Atmospheric Measurement Techniques Discussions*, 7(5):4551–4588, 2014.

[80] C. Hurter, B. Tissoires, and S. Conversy. FromDaDy: Spreading aircraft trajectories across views to support iterative queries. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1017–1024, Nov 2009.

[81] Integrated Environmental Solutions Ltd. Integrated Environmental Solutions Virtual Environment (IES-VE), version 2013.0.0.2. `http://www.iesve.com/`. Accessed: 2013-10-29.

[82] G. W. Inverarity. Correcting airborne temperature data for lags introduced by instruments with two-time-constant responses. *Journal of Atmospheric and Oceanic Technology*, 17(2):176–184, 2000.

[83] J. Ivins and J. Porrill. Everything you always wanted to know about snakes. *AIVRU Technical Memo 86, July 1993 (Revised June 1995; March 2000)*, 2000.

[84] E. Johansson. Influence of urban geometry on outdoor thermal comfort in a hot dry climate: A study in fez, morocco. *Building and Environment*, 41(10):1326 – 1338, 2006.

[85] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[86] J. Kehrer and H. Hauser. Visualization and visual analysis of multi-faceted scientific data: A survey. *Visualization and Computer Graphics, IEEE Transactions on*, 19(3):495–513, 2013.

[87] M. Keller, J. Beutel, O. Saukh, and L. Thiele. Visualizing large sensor network data sets in space and time with vizzly. In *LCN Workshops*, pages 925–933, 2012.

[88] L. Klok, P. van der Mark, E. Nieuwkoop, and the URBMOBI team. Urbmobi – a mobile measurement device for urban environmental monitoring. In *International Conference: Deltas in Times of Climate Change II,*

24-26 September, Rotterdam, the Netherlands, September 2014. Poster P048.

[89] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, Sep 1990.

[90] R. Krüger, D. Thom, and T. Ertl. Visual analysis of movement behavior using web data for context enrichment. In *Pacific Visualization Symposium (PacificVis), 2014 IEEE*, pages 193–200, March 2014.

[91] R. Krüger, D. Thom, M. Wörner, H. Bosch, and T. Ertl. TrajectoryLenses – a set-based filtering and exploration technique for long-term trajectory data. *Computer Graphics Forum*, 32(3):451–460, 2013.

[92] F. Ladstädter, A. K. Steiner, B. C. Lackner, B. Pirscher, G. Kirchengast, J. Kehrer, H. Hauser, P. Muigg, and H. Doleisch. Exploration of climate data using interactive visualization. *Journal of Atmospheric and Oceanic Technology*, 27(4):667–679, 2010.

[93] K. Lagios, J. Niemasz, and C. F. Reinhart. Animated building performance simulation (ABPS) – linking Rhinoceros/Grasshopper with Radiance/Daysim. In *Proceedings of the Fourth National Conference of IBPSA-USA, August 11-13, 2010, New York City, New York*, 2010.

[94] O. D. Lampe, J. Kehrer, and H. Hauser. Visual analysis of multivariate movement data using interactive difference views. In R. Koch, A. Kolb, and C. Rezk-Salama, editors, *VMV 2010: Vision, Modeling & Visualization*, pages 315–322, Siegen, Germany, 2010. Eurographics Association.

[95] F. Leconte, J. Bouyer, R. Claverie, and M. Pétrissans. Using local climate zone scheme for UHI assessment: Evaluation of the method using mobile measurements. *Building and Environment*, 83(0):39 – 49, 2015. Special Issue: Climate adaptation in cities.

[96] J. L. Lerma, M. Cabrelles, and C. Portalés. Multitemporal thermal analysis to detect moisture on a building facade. *Construction and Building Materials*, 25(5):2190 – 2197, 2011.

[97] Z. Liao, Y. Peng, Y. Li, X. Liang, and Y. Zhao. A web-based visual analytics system for air quality monitoring data. In *Geoinformatics (GeoInformatics), 2014 22nd International Conference on*, pages 1–6, June 2014.

[98] P. Lundblad, O. Eurenius, and T. Heldring. Interactive visualization of weather and ship data. In *Information Visualisation, 2009 13th International Conference*, pages 379–386, July 2009.

[99] A. Mahesh, P. von Walden, and S. G. Warren. Radiosonde temperature measurements in strong inversions: correction for thermal lag based on an experiment at the south pole. *Journal of Atmospheric and Oceanic Technology*, 14(1):45–53, 1997.

[100] J. Maurer and S. Watanabe. Boost.random, 2013. `http://www.boost.org/doc/libs/1_55_0/doc/html/boost_random.html`, Last accessed: 05/10/2015.

[101] J.-C. Mayer, K. Hens, U. Rummel, F. X. Meixner, and T. Foken. Moving measurement platforms – specific challenges and corrections. *Meteorologische Zeitschrift*, 18(5):477–488, 2009.

[102] J. McCarthy. A method for correcting airborne temperature data for sensor response time. *Journal of Applied Meteorology*, 12(1):211–214, 1973.

[103] F. Meier and D. Scherer. Spatial and temporal variability of urban tree canopy temperature during summer 2010 in Berlin, Germany. *Theoretical and Applied Climatology*, 110(3):373–384, 2012.

[104] F. Meier, D. Scherer, and J. Richters. Determination of persistence effects in spatio-temporal patterns of upward long-wave radiation flux density from an urban courtyard by means of time-sequential thermography. *Remote Sensing of Environment*, 114(1):21 – 34, 2010.

[105] MesoWest / University of Utah. MesoWest data. Online, 2014. `http://mesowest.utah.edu/index.html`, Last accessed: 08/14/2014.

[106] Meteorological Environmental Earth Observation. The Multi-sensor Evolution Analysis (MEA). Online, 2015. `https://mea.eo.esa.int/`, Last accessed: 08/05/2015.

[107] A. Middel, K. Häb, A. J. Brazel, C. A. Martin, and S. Guhathakurta. Impact of urban form and design on mid-afternoon microclimate in Phoenix Local Climate Zones. *Landscape and Urban Planning*, 122(0):16 – 28, 2014.

[108] G. Mills. Progress toward sustainable settlements: a role for urban climatology. *Theoretical and Applied Climatology*, 84(1-3):69–76, 2006.

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

[110] D. J. Murphy, M. H. Hall, C. A. S. Hall, G. M. Heisler, S. V. Stehman, and C. Anselmi-Molina. The relationship between land cover and the urban heat island in northeastern Puerto Rico. *International Journal of Climatology*, 31(8):1222–1239, 2011.

[111] T. Nocke, M. Flechsig, and U. Bohm. Visual exploration and evaluation of climate-related simulation data. In *Simulation Conference, 2007 Winter*, pages 703–711, Dec 2007.

[112] T. Nocke, T. Sterzel, M. Böttinger, and M. Wrobel. Visualization of climate and climate change data: An overview. *Digital earth summit on geoinformatics*, pages 226–232, 2008.

[113] T. Oke. *Initial Guidance to Obtain Representative Meteorological Observations at Urban Sites*. Instruments and Observing Methods 81, WMO/TD 1250. World Meteorological Organization, 2006.

[114] T. R. Oke. *Boundary layer climates*. Routledge, Abingdon and New York, 2 edition, 1987.

[115] T. R. Oke. Towards better scientific communication in urban climate. *Theoretical and Applied Climatology*, 84(1-3):179–190, 2006.

[116] Omega. Unsheathed fine gage microtemp thermocouples J, K, T, E, R & S. Online, 2014. `http://www.omega.com/pptst/IRCO_CHAL_P13R_P10R.html`, Last accessed: 08/11/2014.

[117] M. Pilgrim, N. Bouchlaghem, D. Loveday, and M. Holmes. Towards the efficient use of simulation in building performance analysis: a user survey. *Building Services Engineering Research and Technology*, 24(3):149–162, 2003.

[118] L. Prazeres and J. A. Clarke. Communicating building simulation outputs to users. In *Proceedings of the Eighth International IBPSA Conference, August 11-14, 2005, Eindhoven, Netherlands*, 2003.

[119] L. Prazeres and J. A. Clarke. Qualitative analysis on the usefulness of perceptualization techniques in communicating building simulation outputs. In *Proceedings of the Ninth International IBPSA Conference, August 15-18, 2005, Montreal, Canada*, 2005.

[120] H. Qu, W.-Y. Chan, A. Xu, K.-L. Chung, K.-H. Lau, and P. Guo. Visual analysis of the air pollution problem in Hong Kong. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1408–1415, Nov 2007.

[121] R Core Team. R: A language and environment for statistical computing, 2014. R Foundation for Statistical Computing, Vienna, Austria.

[122] F. Reinders, F. H. Post, and H. J. Spoelder. Visualization of time-dependent data with feature tracking and event detection. *The Visual Computer*, 17(1):55–71, 2001.

[123] S. Ribarić, D. Marčetić, and D. S. Vedrina. A knowledge-based system for the non-destructive diagnostics of façade isolation using the information fusion of visual and IR images. *Expert Systems with Applications*, 36(2, Part 2):3812 – 3823, 2009.

[124] S. Rinzivillo, D. Pedreschi, M. Nanni, F. Giannotti, N. Andrienko, and G. Andrienko. Visually driven analysis of movement data by progressive clustering. *Information Visualization*, 7(3-4):225–239, 2008.

[125] N. Röber, M. Salim, A. Gierisch, M. Böttinger, and H. Schlünzen. Visualization of urban micro-climate simulations. In *Proceedings of EnvirVis (Visualization in Environmental Sciences) 2013*, June 2013.

[126] J. Roberts. State of the art: Coordinated multiple views in exploratory visualization. In *Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV '07. Fifth International Conference on*, pages 61–71, 2007.

[127] T. Sachsen, G. Ketzler, I. Maras, C. Schneider, and the URBMOBI team. Urbmobi – a mobile measurement device for urban environmental monitoring. In *EFUF 2014, 3.6-7.6.2014, Lausanne, Switzerland*, June 2014. Poster.

[128] J. Sammon. A nonlinear mapping for data structure analysis. *Computers, IEEE Transactions on*, C-18(5):401–409, May 1969.

[129] R. Samtaney, D. Silver, N. Zabusky, and J. Cao. Visualizing features and tracking their evolution. *Computer*, 27(7):20–27, 1994.

[130] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. Moorhead. Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1421–1430, Nov 2010.

[131] P. Sarlin and S. Ronnqvist. Cluster coloring of the self-organizing map: An information visualization perspective. In *Information Visualisation (IV), 2013 17th International Conference*, pages 532–538, July 2013.

[132] R. Scheepens, H. van de Wetering, and J. van Wijk. Non-overlapping aggregated multivariate glyphs for moving objects. In *Pacific Visualization Symposium (PacificVis), 2014 IEEE*, pages 17–24, March 2014.

[133] R. Scheepens, N. Willems, H. van de Wetering, G. Andrienko, N. Andrienko, and J. van Wijk. Composite density maps for multivariate trajectories. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):2518–2527, Dec 2011.

150

[134] R. Scheepens, N. Willems, H. Van De Wetering, and J. van Wijk. Interactive visualization of multivariate trajectory data with density maps. In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, pages 147–154, March 2011.

[135] R. Scheepens, N. Willems, H. van de Wetering, and J. van Wijk. Interactive density maps for moving objects. *Computer Graphics and Applications, IEEE*, 32(1):56–66, Jan 2012.

[136] H. Schmid. Source areas for scalars and scalar fluxes. *Boundary-Layer Meteorology*, 67(3):293–318, 1994.

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

[138] J. F. Sham, S. A. Memon, Tommy, and Y. Lo. Application of continuous surface temperature monitoring technique for investigation of nocturnal sensible heat release characteristics by building fabrics in Hong Kong. *Energy and Buildings*, 58(0):1 – 10, 2013.

[139] X. Shi and W. Yang. Performance-driven architectural design and optimization technique from a perspective of architects. *Automation in Construction*, 32(0):125 – 135, 2013.

[140] D. Sidiropoulou-Velidou, A. Georgopoulos, and J. Lerma. Exploitation of thermal imagery for the detection of pathologies in monuments. In M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, and R. Caffo, editors, *Progress in Cultural Heritage Preservation*, volume 7616 of *Lecture Notes in Computer Science*, pages 97–108. Springer Berlin Heidelberg, 2012.

[141] M. Sofer and O. Potchter. The urban heat island of a city in an arid zone: the case of Eilat, Israel. *Theoretical and Applied Climatology*, 85(1-2):81–88, 2006.

[142] B.-S. Sohn and C. Bajaj. Time-varying contour topology. *Visualization and Computer Graphics, IEEE Transactions on*, 12(1):14–25, 2006.

[143] C. Souch and S. Grimmond. Applied climatology: urban climate. *Progress in Physical Geography*, 30(2):270–279, 2006.

[144] A. Speak, J. Rothwell, S. Lindley, and C. Smith. Reduction of the urban cooling effects of an intensive green roof due to vegetation damage. *Urban Climate*, 3(0):40 – 55, 2013.

[145] S. Srivastav, S. Lannon, D. K. Alexander, and P. Jones. A review and comparison of data visualization techniques used in building design and in building simulation. In *Proceedings of the Eleventh International IBPSA Conference, July 27-30, 2009, Glasgow, Scotland*, 2009.

[146] L. B. Stabler, C. A. Martin, and A. J. Brazel. Microclimates in a desert city were related to land use and vegetation index. *Urban Forestry & Urban Greening*, 3(3–4):137–147, 2005.

[147] I. D. Stewart. A systematic review and scientific critique of methodology in modern urban heat island literature. *International Journal of Climatology*, 31(2):200–217, 2011.

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

[149] J. Sukharev, C. Wang, K.-L. Ma, and A. T. Wittenberg. Correlation study of time-varying multivariate climate data sets. In *IEEE Pacific Visualization Symposium, 2009. PacificVis' 09.*, pages 161–168. IEEE, 2009.

[150] C.-Y. Sun, A. Brazel, W. Chow, B. Hedquist, and L. Prashad. Desert heat island study in winter by mobile transect and remote sensing techniques. *Theoretical and Applied Climatology*, 98(3-4):323–335, 2009.

[151] G. Sun, Y. Liu, W. Wu, R. Liang, and H. Qu. Embedding temporal display into maps for occlusion-free visualization of spatio-temporal data. In *Pacific Visualization Symposium (PacificVis), 2014 IEEE*, pages 185–192, March 2014.

[152] A. Toet, M. Hogervorst, S. Nikolov, J. Lewis, T. Dixon, D. Bull, and C. Canagarajah. Towards cognitive image fusion. *Information Fusion*, 11(2):95 – 113, 2010.

[153] C. Tominski, J. Donges, and T. Nocke. Information visualization in climate research. In *Information Visualisation (IV), 2011 15th International Conference on*, pages 298–305, July 2011.

[154] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko. Stacking-based visualization of trajectory attribute data. *Visualization and Computer Graphics, IEEE Transactions on*, 18(12):2565–2574, Dec 2012.

[155] J. Unger. Intra-urban relationship between surface geometry and urban heat island: review and new approach. *Climate Research*, 27(0):253 – 264, 2004.

[156] J. K. Vanos, J. S. Warland, T. J. Gillespie, G. A. Slater, R. D. Brown, and N. A. Kenny. Human energy budget modeling in urban parks in Toronto and applications to emergency heat stress preparedness. *Journal of Applied Meteorology and Climatology*, 51(9):1639–1653, 2012.

[157] T. Vesala, N. Kljun, U. Rannik, J. Rinne, A. Sogachev, T. Markkanen, K. Sabelfeld, T. Foken, and M. Leclerc. Flux and concentration footprint modelling: State of the art. *Environmental Pollution*, 152(3):653 – 666, 2008.

[158] J. Vesanto and E. Alhoniemi. Clustering of the self-organizing map. *Neural Networks, IEEE Transactions on*, 11(3):586–600, May 2000.

[159] VITO. VITO SensorView (demo). Online, 2015. `http://sensorview.vito.be/demo_aww/`, Last accessed: 08/05/2015.

[160] K. Vrotsou, H. Janetzko, C. Navarra, G. Fuchs, D. Spretke, F. Mansmann, N. Andrienko, and G. Andrienko. SimpliFly: A methodology for simplification and thematic enhancement of trajectories. *Visualization and Computer Graphics, IEEE Transactions on*, 21(1):107–121, Jan 2015.

[161] N. Wang, T. W. Biggs, and A. Skupin. Visualizing gridded time series data with self organizing maps: An application to multi-year snow dynamics in the northern hemisphere. *Computers, Environment and Urban Systems*, 39(0):107 – 120, 2013.

[162] Y. Wang, G. Huynh, and C. Williamson. Integration of Google Maps/Earth with microscale meteorology models and data visualization. *Computers & Geosciences*, 61(0):23 – 31, 2013.

[163] Y. H. Wang, C. R. Fan, J. Zhang, T. Niu, S. Zhang, and J. R. Jiang. Forecast verification and visualization based on gaussian mixture model co-estimation. *Computer Graphics Forum*, pages 1–12, 2014.

[164] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. Van De Wetering. Visual traffic jam analysis based on trajectory data. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2159–2168, Dec 2013.

[165] Z. Wang, T. Ye, M. Lu, X. Yuan, H. Qu, J. Yuan, and Q. Wu. Visual exploration of sparse traffic trajectory data. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):1813–1822, Dec 2014.

[166] C. Ware, R. Arsenault, M. Plumlee, and D. Wiley. Visualizing the underwater behavior of humpback whales. *IEEE Computer Graphics and Applications*, 26(4):14–18, 2006.

[167] W. Widanagamaachchi, C. Christensen, P.-T. Bremer, and V. Pascucci. Interactive exploration of large-scale time-varying data using dynamic tracking graphs. In *Large Data Analysis and Visualization (LDAV), 2012 IEEE Symposium on*, pages 9–17, 2012.

[168] W. Willett, P. Aoki, N. Kumar, S. Subramanian, and A. Woodruff. Common sense community: Scaffolding mobile sensing and analysis for novice users. In P. Floréen, A. Krüger, and M. Spasojevic, editors, *Pervasive Computing*, volume 6030 of *Lecture Notes in Computer Science*, pages 301–318. Springer Berlin Heidelberg, 2010.

[169] C. J. Willmott. On the validation of models. *Physical Geography*, 2(2):184–194, 1981.

[170] C. J. Willmott. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society*, 63(11):1309–1313, 1982.

[171] World Meteorological Organization. World weather information service: Phoenix, Arizona, United States of America. Online, 2014. `http://worldweather.wmo.int/en/city.html?cityId=806`, Last accessed: 09/03/2014.

[172] C. Xu and J. Prince. Snakes, shapes, and gradient vector flow. *Image Processing, IEEE Transactions on*, 7(3):359–369, 1998.

[173] J. S. Yi, Y. ah Kang, J. Stasko, and J. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1224–1231, 2007.

[174] A. Yilmaz, O. Javed, and M. Shah. Object tracking: A survey. *ACM Comput. Surv.*, 38(4), Dec. 2006.

[175] A. Zeileis, G. Grothendieck, J. A. Ryan, and F. Andrews. S3 infrastructure for regular and irregular time series (z's ordered observations), 2014. `http://cran.r-project.org/web/packages/zoo/index.html`, Last accessed: 11/05/2014.

[176] Y. Zhang, X. Zhang, S. Maybank, and R. Yu. An IR and visible image sequence automatic registration method based on optical flow. *Machine Vision and Applications*, 24(5):947–958, 2013.

Appendix A

Curriculum Vitae

Personal Data

Name	Kathrin Häb
Place of birth	Daun, Germany

Education

January 2012 - July 2015	Ph.D. student at the Computer Graphics and HCI Group, Department of Computer Science, University of Kaiserslautern, Germany (including qualification studies for Ph.D. admission)
October 2005 - March 2011	Study of German and Geography at the University of Trier, Germany; Degree: Erstes Staatsexamen für das Lehramt an Gymnasien
September 1996 - March 2005	Secondary education at the Thomas-Morus-Gymnasium Daun, Germany; Degree: Abitur

Work Experience

August 2011 - November 2011	Supply teacher at the Albert-Schweitzer-Gymnasium, Kaiserslautern, Germany
October 2009 - November 2011	Research assistant at the Environmental Meteorology Group, Department of Geography, University of Trier, Germany
August 2009 - June 2011	Pedagogical assistant at the Humboldt-Gymnasium Trier, Germany

Appendix B

Schriftenverzeichnis

Journal publications

[1] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell, Hans Hagen. Visual Identification of Climatic Microenvironments based on Spatial Aggregation. To be submitted to *IEEE Transactions on Visualization and Computer Graphics, SI on Environmental Visualization, 2015.*

[2] **Kathrin Häb**, Benjamin L. Ruddell, Ariane Middel. Sensor Lag Correction for Mobile Urban Microclimate Measurements. *Urban Climate*, 2015. Under review.

[3] **Kathrin Häb**, Nils H. Feige, Lars Hüttenberger, Ariane Middel, Hans Hagen. Visualizing the Temporal Development of Thermo-Radiative Features on Ground-Based Thermographs. *Environmental Earth Sciences* **72**(10): 3781-3793, 2014.

[4] Ariane Middel, **Kathrin Häb**, Anthony J. Brazel, Chris A. Martin, Subhrajit Guhathakurta. Impact of Urban Form and Design on Mid-Afternoon Microclimate in Phoenix Local Climate Zones. *Landscape and Urban Planning* **122**(0): 16-28, 2014.

Peer-reviewed conference papers

[5] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell, Hans Hagen. Spatial Aggregation of Mobile Transect Measurements for the Identification of Climatic Microenvironments. In Ariane Middel, Karsten Rink, Gunther H. Weber, editors, *Workshop on Visualization in Environmental Sciences (EnvirVis)*: 19-23. The Eurographics Association, 2015.

[6] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell, Hans Hagen. TraVis – A Visualization Framework for Mobile Transect Data Sets in an Urban Microclimate Context. *IEEE Pacific Visualization Symposium (PacificVis)*: 167-174, April 2015.

[7] **Kathrin Häb**, Stephanie Schweitzer, Diana Fernández-Prieto, Eva Hagen, Daniel Engel, Michael Böttinger, Inga Scheler. Visualization of Building Performance Simulation Results: State-of-the-Art and Future Directions. *IEEE Pacific Visualization Symposium (PacificVis)*: 311-315, March 2014.

[8] **Kathrin Häb**, Ariane Middel, Hans Hagen. Using K-Means Clustering for a Spatial Analysis of Multivariate and Time-Varying Microclimate Data. In Olaf Kolditz, Karsten Rink, Gerik Scheuermann, editors, *Workshop on Visualization in Environmental Sciences (EnvirVis)*: 13-17, Eurographics Association, 2014.

[9] Ariane Middel, **Kathrin Häb**, Anthony J. Brazel, Chris A. Martin, Subhrajit Guhathakurta. Urban Form, Landscape Design, and Microclimate in Phoenix, Arizona. In *Proceedings ICUC8 – 8th International Conference on Urban Climate*, Dublin, Ireland, 2012.

Posters and talks

[10] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell. Classifying Urban Microclimate Patches Based on Multitemporal Mobile Transect Observations. Talk at the *Annual Meeting of the Association of American Geographers (AAG)*, Chicago, Illinois, USA, April 2015.

[11] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell. Sensor Lag Correction for Mobile Air Temperature Measurements in an Urban Microclimate Context. Poster at the *17th CAP LTER All Scientists Meeting Poster Symposium*, Scottsdale, Arizona, USA, January 2015.

[12] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell, Hans Hagen. Visual Analytics and Microclimate Analysis: A use case for a visualization tool developed for mobile measurements. Talk at the *31st Conference on Environmental Information Processing Technologies, 95th Annual Meeting of the American Meteorological Society*, Phoenix, Arizona, USA, January 2015.

[13] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell, Anthony J. Brazel, Chris A. Martin. Source Area Computation for Microclimate Measurements within the Urban Canopy Layer. Talk at the *Annual Meeting of the Association of American Geographers (AAG)*, Tampa, Florida, USA, April 2014.

[14] Ariane Middel, **Kathrin Häb**, Anthony J. Brazel, Chris A. Martin, Benjamin L. Ruddell. Linking Shading Patterns of Trees in Phoenix, AZ to thermal comfort. Talk and Poster at the *11th Symposium on the Urban Environment, 94th Annual Meeting of the American Meteorological Society*, Atlanta, Georgia, USA, February 2014.

[15] **Kathrin Häb**, Ariane Middel, Benjamin L. Ruddell. Visualizing Urban Microclimate Transect Measurements. Poster at the *16th Annual CAP LTER All Scientists Meeting Poster Symposium*, Scottsdale, Arizona, USA, January 2014.

[16] Ariane Middel, **Kathrin Häb**, Anthony J. Brazel, Chris A. Martin. Understanding the Physical Dynamics of Microclimate: Ongoing Research Projects. Poster at the *16th Annual CAP LTER All Scientists Meeting Poster Symposium*, Scottsdale, Arizona, USA, January 2014.

[17] Ariane Middel, **Kathrin Häb**, Anthony J. Brazel, Chris A. Martin, Subhrajit Guhathakurta. Impact of Urban Form and Design on Mid-Afternoon Microclimate in Phoenix neighborhoods. Poster at the *15th Annual CAP LTER All Scientists Meeting Poster Symposium*, Scottsdale, Arizona, USA, January 2013.