Distributed, Low-Cost, Non-Expert Fine Dust Sensing with Smartphones

To Attain the Academic Degree of

Doctor of Engineering

Faculty of Informatics
Karlsruhe Institute of Technology (KIT)

accepted

Dissertation

by

Matthias Budde

Born in Lüdenscheid, Germany

Examination Date:	January 30th, 2018
Primary Examiner:	Prof. Dr. Michael Beigl
Secondary Examiner:	Prof. Dr. Lin Zhang

Acknowledgments

This dissertation would not have been possible without the help and support of many people. First and foremost, I cannot thank my thesis adviser Prof. Michael Beigl enough for the opportunity to be part of his research team at TECO, for the freedom to explore different research directions and for his support and advice in this endeavor. Many thanks are due to my secondary adviser Prof. Lin Zhang for kindly agreeing to take on this role, and for welcoming me to work with him and his group in Beijing.

For their help, advice and support in my experiments, I would like to thank Thomas Müller, Zarko Peranić, and Claudio Laloni. Many thanks to Klaus Schäfer, Volker Ziegler, Fritz Waitz, Thomas Müller, and Jan Budde, who gave their time and expertise for proof-reading and background discussions.

Of my colleagues at TECO, my first thank you goes out to our secretary Helga Scherer and her predecessor Irina Schierholz. Without you, the office probably would have collapsed a long time ago. I would also like to thank my fellow researchers at TECO, particularly Till Riedel, Matthias Berning, Andrea Schankin and Johannes Riesterer, for the opportunity to conduct joint research with them. To all of my colleagues, and especially Markus Scholz and Erik Pescara, I extend my thanks for making TECO a great place to work.

Of my many, many students, there are several I would like to single out: Marcel Köpke and Simon Leiner, without you, this work would not have been possible. Many thanks also to Mathias Busse, Cole Bailey, Gregor Schindler, and Dominik Kogel, I am grateful for all you have done.

I would like to thank my siblings Joachim, Jan and Marie-Caroline for being a part of me and giving me the opportunity to travel back in time every now and then. I thank my parents Rita and Helmut for raising me in a family that, to this day, I have been envied for countless of times over the years. Finally, I would like to thank my own little family: My children Anna and Max, who brighten my every day, and my brilliant and beautiful wife Martina, who has been extremely supportive of me in getting to this point and who always has my back. I love you.

Abstract

This dissertation deals with the question of how fine dust can be measured using low-cost sensors and with high resolution, both spatially and temporally. For this purpose, a novel sensor system based on inexpensive off-the-shelf sensors and smartphones is presented, along with corresponding robust algorithms for signal processing, a privacy-preserving calibration scheme and research on the interaction design for participatory measurements contributed by non-expert users.

Atmospheric aerosol particles pose a serious health hazard on a global scale, which manifests itself in respiratory and cardiovascular disease and causes shortened life expectancy. In the past, air quality has been assessed on the basis of data from relatively few fixed measurement stations and been brought to a high spatial resolution using transport models, the representativeness of which for the nationwide exposure of the population remains unclear. Since it is impossible to achieve such spatial information with current static measurement networks, the trend is toward distributed measurements with large spatial resolution.

A promising approach to achieve a high spatial and temporal coverage is participatory sensing, i.e. the distributed measurement by end users with the help of their personal end devices. There are a number of challenges, in particular for air quality measurements, ranging from new sensors that are cheap and portable, over robust signal analysis and calibration algorithms, to applications that help non-experts to correctly perform measurements and contribute them while protecting their privacy.

This work focuses on the application scenario of participatory environmental sensing, in which smartphone-based sensors are used to measure the environment and usually non-experts perform the measurements in a relatively uncontrolled manner. The main contributions to this are:

1. *Systems for detecting fine dust with smartphones (low-cost sensors and novel hardware):* Based on early research on fine dust measurements with inexpensive off-the-shelf sensor technology, a novel sensor concept has been developed in which the fine dust measurement is carried out with the aid of a passive clip-on attachment for a smartphone

camera. Evaluations of the sensor performance were partially carried out against laboratory measurements with artificially generated dust and partially in real-world field evaluations in which sensors were co-located with official state-operated measurement stations.

2. *Algorithms for signal processing and evaluation:* Combinations of known OpenCV image processing methods (background subtraction, contour detection, etc.) were used for image analysis in the novel sensor designs. The resulting algorithm, in contrast to the evaluation of light-scattering sum signals, allows direct counting of particles based on individual traces of light. A second novel algorithm takes advantage of the fact that Poisson data is afflicted with signal-dependent noise, the ratio of which to the mean value of the signal is known. This makes it possible to analyze signals that are affected by systematic unknown errors based on their noise and to reconstruct the "true" signal from it.

3. *Distributed privacy-preserving calibration algorithms:* One challenge of participatory environmental measurements is the recurring need for sensor calibration. This is, on the one hand, due to the instability of (in particular) low-cost air quality sensors and, on the other, due to the problem that end users usually lack the means to perform a calibration. Existing approaches to so-called device-by-device-calibration of sensors have been applied to data from low-cost particulate matter sensors and extended by mechanisms that allow sensors to be calibrated without giving away private information (identity, location).

4. *Human-Computer-Interaction design for Participatory Sensing:* On the basis of several small exploratory studies, a taxonomy of the errors made by non-expert users when measuring environmental phenomena with smartphones was created empirically. From this, possible countermeasures were collected and classified. In a large summative study with many participants, the effect of several of these measures was evaluated by comparing four different variants of an app for participatory measurement of ambient noise. The findings form the basis for guidelines for designing efficient user interfaces for participatory sensing on handheld mobile devices.

5. *Design Patterns for Participatory Sensing Games on Mobile Devices (Gamification):* Finally, gamification of the measurement process was explored to minimize user errors by using appropriate game mechanisms. In contrast to existing work, the focus was on embedding sensing tasks in smartphone games (e.g. so-called minigames), which perform the measurement in the background once in a suitable context. For the development of this concept dubbed "Sensified Gaming", core tasks in participatory sensing have been identified and juxtaposed with Game Design Patterns that were collected from the literature.

Zusammenfassung

Diese Dissertation behandelt die Frage, wie mit kostengünstiger Sensorik Feinstäube in hoher zeitlicher und räumlicher Auflösung gemessen werden können. Dazu wird ein neues Sensorsystem auf Basis kostengünstiger off-the-shelf-Sensoren und Smartphones vorgestellt, entsprechende robuste Algorithmen zur Signalverarbeitung entwickelt und Erkenntnisse zur Interaktions-Gestaltung für die Messung durch Laien präsentiert.

Atmosphärische Aerosolpartikel stellen im globalen Maßstab ein gravierendes Problem für die menschliche Gesundheit dar, welches sich in Atemwegs- und Herz-Kreislauf-Erkrankungen äußert und eine Verkürzung der Lebenserwartung verursacht. Bisher wird Luftqualität ausschließlich anhand von Daten relativ weniger fester Messstellen beurteilt und mittels Modellen auf eine hohe räumliche Auflösung gebracht, so dass deren Repräsentativität für die flächendeckende Exposition der Bevölkerung ungeklärt bleibt. Es ist unmöglich, derartige räumliche Abbildungen mit den derzeitigen statischen Messnetzen zu bestimmen. Bei der gesundheitsbezogenen Bewertung von Schadstoffen geht der Trend daher stark zu räumlich differenzierenden Messungen.

Ein vielversprechender Ansatz um eine hohe räumliche und zeitliche Abdeckung zu erreichen ist dabei Participatory Sensing, also die verteilte Messung durch Endanwender unter Zuhilfenahme ihrer persönlichen Endgeräte. Insbesondere für Luftqualitätsmessungen ergeben sich dabei eine Reihe von Herausforderungen — von neuer Sensorik, die kostengünstig und tragbar ist, über robuste Algorithmen zur Signalauswertung und Kalibrierung bis hin zu Anwendungen, die Laien bei der korrekten Ausführung von Messungen unterstützen und ihre Privatsphäre schützen.

Diese Arbeit konzentriert sich auf das Anwendungsszenario Partizipatorischer Umweltmessungen, bei denen Smartphone-basierte Sensorik zum Messen der Umwelt eingesetzt wird und üblicherweise Laien die Messungen in relativ unkontrollierter Art und Weise ausführen. Die Hauptbeiträge hierzu sind:

1. *Systeme zum Erfassen von Feinstaub mit Smartphones (Low-cost Sensorik und neue Hardware):* Ausgehend von früher For-

schung zur Feinstaubmessung mit kostengünstiger off-the-shelf-Sensorik wurde ein Sensorkonzept entwickelt, bei dem die Feinstaub-Messung mit Hilfe eines passiven Aufsatzes auf einer Smartphone-Kamera durchgeführt wird. Zur Beurteilung der Sensorperformance wurden teilweise Labor-Messungen mit künstlich erzeugtem Staub und teilweise Feldevaluationen in Ko-Lokation mit offiziellen Messstationen des Landes durchgeführt.

2. *Algorithmen zur Signalverarbeitung und Auswertung:* Im Zuge neuer Sensordesigns werden Kombinationen bekannter OpenCV-Bildverarbeitungsalgorithmen (Background-Subtraction, Contour Detection etc.) zur Bildanalyse eingesetzt. Der resultierende Algorithmus erlaubt im Gegensatz zur Auswertung von Lichtstreuungs-Summensignalen die direkte Zählung von Partikeln anhand individueller Lichtspuren. Ein zweiter neuartiger Algorithmus nutzt aus, dass es bei solchen Prozessen ein signalabhängiges Rauschen gibt, dessen Verhältnis zum Mittelwert des Signals bekannt ist. Dadurch wird es möglich, Signale die von systematischen unbekannten Fehlern betroffen sind auf Basis ihres Rauschens zu analysieren und das „echte“ Signal zu rekonstruieren.

3. *Algorithmen zur verteilten Kalibrierung bei gleichzeitigem Schutz der Privatsphäre:* Eine Herausforderung partizipatorischer Umweltmessungen ist die wiederkehrende Notwendigkeit der Sensorkalibrierung. Dies beruht zum einen auf der Instabilität insbesondere kostengünstiger Luftqualitätssensorik und zum anderen auf der Problematik, dass Endbenutzern die Mittel für eine Kalibrierung üblicherweise fehlen. Bestehende Ansätze zur sogenannten Cross-Kalibrierung von Sensoren, die sich in Ko-Lokation mit einer Referenzstation oder anderen Sensoren befinden, wurden auf Daten günstiger Feinstaubsensorik angewendet sowie um Mechanismen erweitert, die eine Kalibrierung von Sensoren untereinander ohne Preisgabe privater Informationen (Identität, Ort) ermöglicht.

4. *Mensch-Maschine-Interaktions-Gestaltungsrichtlinien für Participatory Sensing:* Auf Basis mehrerer kleiner explorativer Nutzerstudien wurde empirisch eine Taxonomie der Fehler erstellt, die Laien beim Messen von Umweltinformationen mit Smartphones machen. Davon ausgehend

wurden mögliche Gegenmaßnahmen gesammelt und klassifiziert. In einer großen summativen Studie mit einer hohen Teilnehmerzahl wurde der Effekt verschiedener dieser Maßnahmen durch den Vergleich vier unterschiedlicher Varianten einer App zur partizipatorischen Messung von Umgebungslautstärke evaluiert. Die dabei gefundenen Erkenntnisse bilden die Basis für Richtlinien zur Gestaltung effizienter Nutzerschnittstellen für Participatory Sensing auf Mobilgeräten.

5. *Design Patterns für Participatory Sensing Games auf Mobilgeräten (Gamification):* Ein weiterer erforschter Ansatz beschäftigt sich mit der Gamifizierung des Messprozesses um Nutzerfehler durch den Einsatz geeigneter Spielmechanismen zu minimieren. Dabei wird der Messprozess z. B. in ein Smartphone-Spiel (sog. Minigame) eingebettet, das im Hintergrund bei geeignetem Kontext die Messung durchführt. Zur Entwicklung dieses „Sensified Gaming" getauften Konzepts wurden Kernaufgaben im Participatory Sensing identifiziert und mit aus der Literatur zu sammelnden Spielmechanismen (Game Design Patterns) gegenübergestellt.

Contents

List of Tables

List of Figures

"The ultimate test of a moral society is the kind of world that it leaves to its children."

— Dietrich Bonhoeffer

1 Introduction

Ubiquitous computing has become reality: the proliferation of smartphones and the emergence of the Internet of Things (IoT) have enabled pervasive technology that reveals previously unseen patterns of life. This also has lowered the bar for participation, enabling everyday users to carry, maintain, operate and even build computing equipment that senses and controls the world around them. One of the great topics of our time that is strongly affected by this development is the measurement of ambient air quality.

Figure 1.: Emerging air quality sensing paradigms promise unprecedented spatial and temporal resolution.

Atmospheric aerosol particles pose a serious health hazard on a global scale [139], which manifests itself in respiratory and cardiovascular disease and causes shortened life expectancy. As a result, societies around the globe have developed an increasing awareness and people have started to gain an interest in the levels of air pollution that they are exposed to. In the past, air quality has been assessed on the basis of data from relatively few fixed measurement stations and been brought to a high spatial resolution using transport models, the representativeness of which for the nationwide exposure of the population remains unclear [83].

The goal of this dissertation is to enable distributed mobile low-cost Particulate Matter (PM) measurements by non-expert end users. For this purpose, a novel sensor system based on inexpensive off-the-shelf sensors and smartphones is presented, along with corresponding robust algorithms for signal processing and findings on the interaction design for participatory measurements contributed by non-expert users.

1.1 CHALLENGES

Distributed ambient fine dust monitoring has the potential to deliver readings with a high temporal and spatial resolution and low latency. However, it also entails a number of challenges that need to be addressed and that cannot be viewed decoupled from one another or the specific application scenario [47]. The ones that relate to the data collection respectively to the quality of the collected data are:

- **Coverage**: In order to achieve fine-grained spatial and temporal resolution, a sufficient *Coverage* of the measurement area in space and time is required. Different sensing paradigms and collection schemes for distributed measurements have been conceived [135], ranging from Wireless Sensor Networks (WSNs) over vehicular networks to personal monitoring and Participatory Sensing (PS) [48]. These each have different constraints, advantages and drawbacks, as discussed deeper in chapter 2.

- **Instrumentation**: Closely linked to the achievement of appropriate coverage is the choice of *Instrumentation*. This of course again depends on the employed collection scheme, i.e. how and where sensors are deployed and who the sensing agents are. The larger the scale of a contemplated

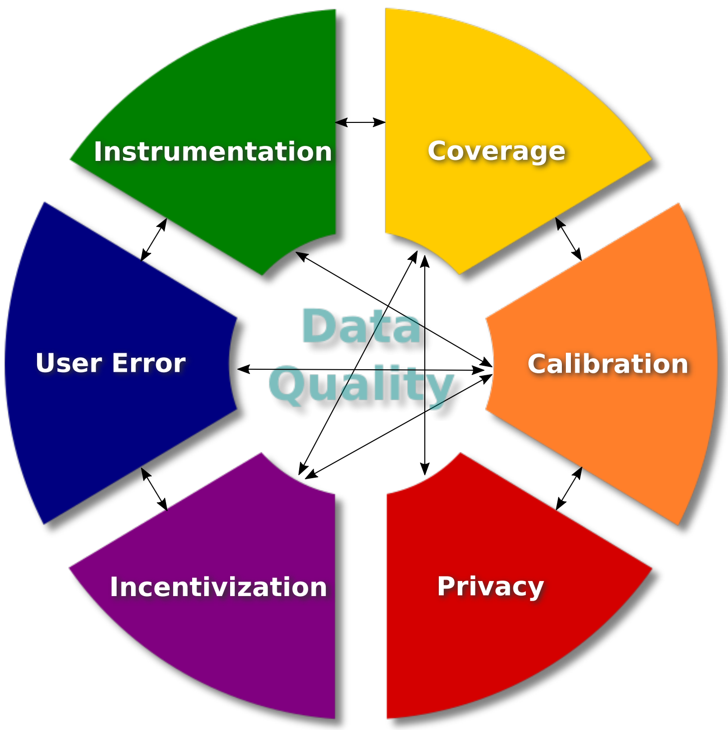


Figure 2.: Challenges for distributed low-Cost environmental sensing relating to the data collection, respectively the quality of the collected data.

measurement system, the more important becomes instrumentation that is small, low-cost and portable, yet yields data of adequate quality. Different device classes for a variety of sensing approaches to measure fine dust have been devised. They are presented in chapter 2, along with a discussion of their respective suitability for mobile distributed sensing.

- **Calibration**: Depending on both the employed instrumentation and the collection scheme, the need for suitable means of *Calibration* arises. As this may include the requirement to more or less frequently re-calibrate and maintain the devices, suitable approaches ideally can be carried out in-situ and with as little user involvement as possible. Promising approaches that have been demonstrated in other domains include self-calibration and cross-calibration between devices or against reference stations [104]. The latter may have implications for the coverage of the area, as well as for privacy, in case the collection scheme involves humans carrying and/or operating the sensors.

- **Privacy**: If humans are involved in distributed sensing, they may need to share personal data, affecting their *Privacy*. If users for instance, regularly contribute sensor data while going about their daily business, location traces or even profiles of their habits may be inadvertently revealed. Suitable mechanisms that protect sensitive data are needed, particularly in the context of rendezvous-based device-to-device calibration.

- **User Error**: Another important aspect when — especially untrained — users are involved, is that of *User Error* and the resulting implications for data quality. The more intricate the measurement procedure, the more likely it is that non-expert users may need aid in performing it. User error can either be addressed through robust sensing and/or data analysis techniques or through the appropriate design of user interfaces.

- **Incentivization**: In order for larger scale systems to be sustainable, it may be necessary to implement *Incentivization* mechanisms to incite and maintain user participation. Generally, urban sensing of air quality attracts environmentally conscious people by itself. Gamification is a possible approach to complement these intrinsically motivated participants with ones that are driven by external incentives. However, possible effects on data quality must be kept in mind when designing such measures.

Other challenges are suitable *processing* mechanisms for data fusion and analytics, as well as *visualization* — especially for appropriately displaying data with varying uncertainty, etc. While these are important aspects for real-world information systems, they do not directly relate to the quality of the data collection and are therefore not in the scope of this thesis.

1.2 CONTRIBUTIONS

In this dissertation, each of the presented challenges is addressed. While we focus on the scenario of affordable participatory fine dust sensing, all contributions except the developed instrumentation are applicable beyond, e.g. to sensing other phenomena or generally recording data with smartphones. The main contributions of this dissertation are:

1. *Systems for detecting fine dust with smartphones (low-cost sensors and novel hardware):* Early research including — to the best of our knowledge — one of the first papers comparing the data of low-cost Commercial-of-the-shelf (COTS) light-scattering wit a gauged professional reference, the first handheld platform featuring such sensors for mobile, personal PM sensing and the first evaluation involving an official state-operated reference station is presented. Based on this early work, a novel sensor concept has been developed in which the fine dust measurement is carried out with the aid of a passive clip-on attachment for a smartphone camera. Subsequent to a proof-of-concept version that demonstrates that adapting the light scattering principle to mobile phones works in this way, further design iterations of the novel sensor concept are shown, with the aim of achieving the required stability and sensitivity for measurements in a real environment. Evaluations of the sensor performance were partially carried out against laboratory measurements with artificially generated dust and partially in real-world field evaluations in which sensors were co-located with official state-operated measurement stations.

2. *Algorithms for signal processing and evaluation:* Based on naïve analyses (simple evaluation of overall image brightness), combinations of known OpenCV image processing methods (background subtraction, contour detection, etc.) were used for image analysis in the novel sensor designs. The resulting algorithm, in contrast to the evaluation of light scatter sum signals, allows direct counting of particles based on individual traces of light. As further part of the development of the novel sensor component, a second approach to sensor data evaluation was developed and successfully tested. In the method, the measurement of particulate matter is modeled as a Poisson process. The technique takes advantage of the fact that Poisson data is afflicted with signal-dependent noise, the ratio of which to the mean value of the signal is known. This makes it possible to analyze signals that are affected by systematic unknown errors based on their noise and to reconstruct the "true" signal from it.

3. *Distributed privacy-preserving calibration algorithms:* One challenge of participatory environmental measurements is the recurring need for sensor calibration. This is, on the one hand, due to the instability of (in particular) low-cost air quality sensors and, on the other, due to the problem that end users usually lack the means to perform a calibration. Existing approaches to so-called cross-calibration of sensors when co-located with a reference station or other sensors have been applied to data from low-cost particulate matter sensors and extended by mechanisms that allow sensors to be calibrated without giving away private information (identity, location). The evaluation was performed simulatively on a mobility dataset with taxi cab traces.

4. *Human Computer Interaction (HCI) design for Participatory Sensing:* On the basis of several small exploratory studies, a taxonomy of the errors made by non-expert users when measuring environmental phenomena with smartphones was created empirically. From this, possible countermeasures were collected and classified. In a large summative study with many participants, the effect of several of these measures was evaluated by comparing four different variants of an app for participatory measurement of ambient noise. The findings form the basis for guidelines for designing efficient user interfaces for participatory sensing on handheld mobile devices.

5. *Design Patterns for Participatory Sensing Games on Mobile Devices (Gamification):* Finally, gamification of the measurement process was explored to minimize user errors by using appropriate game mechanisms. In contrast to existing work, the focus was on embedding sensing tasks in smartphone games (e.g. so-called minigames), which perform the measurement in the background once in a suitable context. An extension of this approach is to integrate not only the measurement process but also other aspects of the participatory sensing system (such as ensuring high coverage, rendezvous between participants for cross-device calibration, etc.) in location-based games. For the development of this concept dubbed "Sensified Gaming", core tasks in participatory sensing have been identified and juxtaposed with Game Design Patterns that were collected from the literature. A collection of the found

design patterns was compiled and presented together with the concept and a discussion of the implications for the design.

1.3 STRUCTURE OF THIS DISSERTATION

The remainder of this dissertation is structured as follows. In chapter 2, background information on Particulate Matter (PM) and distributed sensing scenarios is provided, including an overview of existing PM measurement approaches and a discussion of their suitability for mobile and distributed fine dust sensing. Chapter 3 presents early research on the feasibility of using COTS light scattering sensors for meaningful PM measurements. In chapter 4, the adoption of that same measurement principle in form of a passive clip-on sensor for camera-smartphones is shown. Design iterations of the sensors are presented along with suitable algorithms for on-device image processing.

The remaining three chapters focus on the user-related aspects of the underlying sensing scenario: Chapter 5 discusses the application of existing node-to-node calibration algorithms to PM measurements and presents the design of extensions for these distributed calibration schemes that can protect the location privacy of the user. In chapter 6, the landscape of human error in participatory environmental sensing is explored empirically, along with a survey on measures how to avoid them and studies concerning their effects, both on error frequency as well as on user experience (UX). The concepts in chapter 7 take this approach further by embedding the sensing procedure into game environments in order to reduce user error.

Chapter 8 summarizes the findings of this dissertation and provides an outlook into possible future work. Appendix A provides a full list of peer-reviewed scientific papers published in the course of the author's PhD studies, as well as a short curriculum vitae.

PUBLICATIONS

Parts of this thesis have previously been published in scientific journals and conferences: [29]–[31], [33]–[35], [37], [39], [40], [42], [43], [45]–[47], [153], [154]. A full list of the peer-reviewed publications that were published in the course of the author's PhD studies is provided in Appendix A.

Several bachelor theses, master theses and diploma theses have been co-supervised by the author and served as basis for many results described in this thesis, most notably: [12], [66], [84], [152], [163], [175], [196].

2 Background and Related Work

This chapter first provides background information on atmospheric particles and their measurement, followed by an introduction to distributed environmental sensing scenarios and an overview of related work.

Parts of this chapter have previously been published. Government regulations on Particulate Matter (PM) have been summarized before [35] and a paper discussing distributed sensing scenarios was previously presented on the International Conference on Atmospheric Dust (DUST) [47]. Also, parts of the introduction and related work may have appeared in different previous publications [35], [37], [43].

2.1 Particulate Matter

PM, also referred to as Suspended Particulate Matter (SPM), is the sum of liquid and solid particles suspended in the air. Their origin can be both natural (e.g. dust transportation from deserts, ejection from the sea, volcanic activity, forest fires, pollen, etc.) or artificial (e.g. combustion, building, mining, road dust resuspension, tire and brake wear, etc.). Depending on the source, the chemical composition and size of the particles may vary greatly. Most PM particles are formed in the atmosphere as a result of chemical reactions between pollutants [236], as well as through coagulation, nucleation, and accumulation.

The effects of Particulate Matter on human health have been extensively studied in the past decades [59], [79], [187], [197], [202], [252]. PM particles are as small as body cells and largely invisible to the naked eye[1]. Due to that, they can travel beyond

1 Though particles can either be indirectly seen through reduced visibility (e.g. smog) or be locally visible due to high concentrations and/or strong absorption (e.g. tobacco smoke or diesel soot).

the larynx and deep into the lungs where they can damage cells or deliver toxins into the body. The results are that fine dust can be a serious health hazard, causing both acute and chronic effects, as well as damage to the environment [251]. PM presents a serious problem for human health on a global scale [139], which manifests itself in respiratory and cardiovascular disease and causes shortened life expectancy.

2.1.1 *Size Classes*

The International Organization for Standardization (ISO) has defined several size fractions to use when sampling in order to collect airborne particles [114]. The sum of all atmospheric particles is called Total Suspended Particles (TSP) or Total Suspended Particulate Matter (TSPM). The so-called *inhalable* fraction consists of all particles that are breathed in, i.e. that enter the mouth and/or nose when breathing. The part of these that travel past the larynx is the so-called *thoracic* fraction and the subset of the thoracic particles that is small enough to be transported deeply into the lungs is called the *respirable* fraction. Generally, the smaller the particles are, the deeper are they transported into the body and the larger are their effects on human health.

For the monitoring of air pollution, different classes of Particulate Matter have successively been defined, originally by the *U.S. Environmental Protection Agency (EPA)* as part of a National Ambient Air Quality Standard (NAAQS) [237] and subsequently adopted in directives of the European Commission (EC) [87] and in ISO standardization [115]. These are PM_{10}, $PM_{2.5}$, and PM_1. PM_{10} corresponds to the *thoracic convention* and $PM_{2.5}$ corresponds to the *high-risk respirable convention*, as defined in the ISO 7708:1995 standard [114]. PM_x is often colloquially defined to denote the total mass of particles with a diameter of less than or equal to x µm. The general dimensions of the currently most relevant classes PM_{10} and $PM_{2.5}$ are depicted in Figure 3 [236].

However, the above definition is simplified and not 100 % correct. The actual definition is more complex: PM_{10} (respectively $PM_{2.5}$) is actually defined as "*particulate matter which passes through a size-selective inlet with a 50 % efficiency cut-off at 10 µm aerodynamic diameter*" (respectively 2.5 µm) [87], [115]. So firstly, the definition entails the Aerodynamic Diameter (AD), rather than the physical one and secondly, there are actually particles of AD $\geq x$ in PM_x, albeit a relatively small portion. The exact

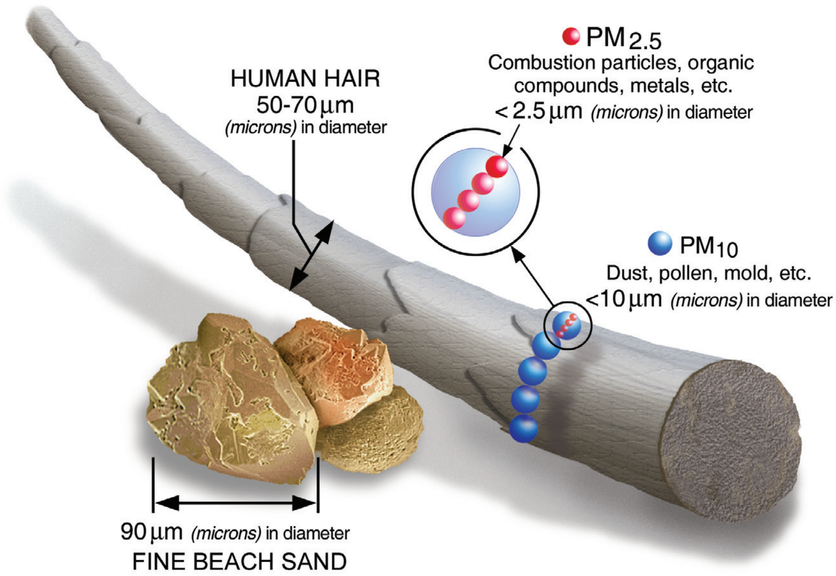


Figure 3.: Comparison of the dimensions of particles in the size classes PM_{10} and $PM_{2.5}$ (image by the EPA [236]).

measurement of Particulate Matter according to this definition is not trivial. BUTTERFIELD ET AL. discuss some issues regarding the precision [50].

Other size classes directly related to the ones mentioned above are the so-called *(inhalable) coarse particles*, which are the ones between 10 µm and 2.5 µm and can therefore also be written as $PM_{(10-2.5)}$. Particles much smaller than PM_1 are called Ultrafine Particles (UFPs). Their diameter is on the nanoscopic scale, usually below 0.1 µm. Therefore, they are sometimes also denoted as $PM_{0.1}$.

In meteorology, particles are also often classified according to their optical properties. Dark particles that heavily absorb light are called Black Carbon (BC), lighter particles that reflect or scatter light rather than absorb it are called Organic Carbon (OC).

2.1.2 *Regulations*

Due to the adverse health effects of PM pollution, more and more regulations regarding the reduction of man-made particulate matter have been set by governments around the world. Such standards usually define limits for particle matter concentrations which may not be exceeded. Today, there are usually

	Class	Maximum permitted		Tolerated exceedances
WHO	$PM_{2.5}$	10 $\frac{\mu g}{m^3}$	(annual mean)	–
		25 $\frac{\mu g}{m^3}$	(24-hour mean)	–
	PM_{10}	20 $\frac{\mu g}{m^3}$	(annual mean)	–
		50 $\frac{\mu g}{m^3}$	(24-hour mean)	–
EU	$PM_{2.5}$	25 $\frac{\mu g}{m^3}$	(annual mean)	–
	PM_{10}	40 $\frac{\mu g}{m^3}$	(annual mean)	–
		50 $\frac{\mu g}{m^3}$	(24-hour mean)	max. 35 days per year
USA	$PM_{2.5}$	15 $\frac{\mu g}{m^3}$	(annual mean)	–
		35 $\frac{\mu g}{m^3}$	(24-hour mean)	–
	PM_{10}	150 $\frac{\mu g}{m^3}$	(24-hour mean)	max. 1 day in 3 years
China	$PM_{2.5}$	35 $\frac{\mu g}{m^3}$	(annual mean)	–
		75 $\frac{\mu g}{m^3}$	(24-hour mean)	–
	PM_{10}	70 $\frac{\mu g}{m^3}$	(annual mean)	–
		150 $\frac{\mu g}{m^3}$	(24-hour mean)	–

Table 1.: Maximum permissible values for particulate matter as suggested by the WHO [227], respectively set by the EU [88], the U.S. EPA [238] and the Chinese Ministry of Environmental Protection [60].

several of such maximum permissible values for different particle size classes and observation periods.

Table 1 shows different limits for the two most commonly regulated classes (PM_{10} and $PM_{2.5}$), as defined by the World Health Organization (WHO) [227], the European Commission (EC)/European Union (EU) [88], the U.S. Environmental Protection Agency (EPA) [238], and the Chinese Ministry of Environmental Protection [60].

While the values provided by the World Health Organization (WHO) are mere recommendations, the limits of the EU, the EPA and China are compulsory. However, violations of these limits are tolerated to a certain extent as part of some of the standards (see Table 1).

The density of monitoring networks is rather sparse: Sometimes, only a single measurement station is used to determine the particulate matter load for a large urban area. However, both the individual exposure to potentially hazardous conditions as well as the susceptibility to negative health effects vary from person to person [227].

Method/Class	Type	Reading	Min. Time Resolution	Latency
Gravimetric	direct	mass	hours	days to weeks
β-Attenuation	direct	mass	minutes	minutes
TEOM	direct	mass	seconds	real-time
FBAR	direct	mass	minutes	real-time
Nephelometry	direct	aggregated (particle count, size)	seconds	real-time
OPC/Spectrometry	direct	particle count, size	seconds	real-time
Direct Imaging	direct	particle count, size	minutes	real-time
Deposition Imaging	direct	particle count, size	minutes	minutes
Aethalometry	direct	mass (BC)	minutes	real-time
Capacitive	direct	particle count, size	seconds	real-time
PAS	direct	absorption	seconds	real-time
GAM	indirect	aggregated exposure (BC)	n/a	n/a
SP2	direct	mass (BC)	seconds	real-time
LIDAR/Ceilometer	indirect	AOT	minutes	real-time
Sun Photometer	indirect	AOT	minutes	real-time
Radiometer	indirect	AOT	minutes	real-time
Spectropolarimetry	indirect	DoLP (AOT)	minutes	real-time

Table 2.: Comparison of PM measurement approaches[2].

2.1.3 *Measurement*

In order to monitor the compliance with the presented standards a variety of measurement methods have been proposed (see Table 2). Systems and techniques can be distinguished according to different characteristics. Measurement can be either *direct*, i.e. observation of effects that stem from the direct interaction with particles (such as absorption or scattering) or *indirect*, i.e. observing effects that allow for conclusions about the aerosol concentration rather than direct quantification. Sensing can either be carried out *in-situ* or *remote*. *Collecting methods* require

2 Abbreviations are defined in the running text and in the glossary.

particles to be deposited, e.g. on a filter, for a longer amount of time in order to measure them, while *online* or (near) *real-time* systems are able to determine the aerosol concentration continuously and with little or no delay.

As outcome of the measurement, instruments may deliver data on the *mass concentration*, the *particle density*, *sizes* or *composition* or other properties of the observed particles, as for instance their *radiation absorption capacity*. The former ones are more relevant for health effects and regulatory monitoring (see below) while the absorption is most interesting for climate research, as it is the decisive factor for global warming. Particle size can either be directly inferred from the signal or determined through additional techniques such as multiple instruments with different size selective inlets or other technology for size segregation, such as ionizing particles and then separating them using strong electric fields. An important factor that can interfere with particle sizing is humidity. Particles function as condensation nuclei, leading to significantly increased size readings upwards of ca. 70 % Relative Humidity (RH) [62], e.g. in optical particle measurements. Similarly, in filter-based collecting methods, humidity can accumulate in the filter material, affecting the measured mass.

Gravimetric

Filter-based gravimetric measurements are the *gold standard* in PM measurements, as they actually are the basis for the PM class definitions above. Gravimetric sampling is a collecting method, i.e. particles are deposited on a sampling filter for a certain sampling window and later weighed and possibly further analyzed. Because in order to collect enough mass, gravimetric samplers generally need to sample over longer time intervals and thus deliver average values and cannot show short-term temporal patterns.

Many official measurement stations in current air quality monitoring networks collect fine dust using a so-called High Volume Sampler (HVS) or Low Volume Sampler (LVS). If both PM_{10} and one $PM_{2.5}$ are being monitored, multiple samplers need to be employed for the measurement, one for each class with an appropriate size selective inlet[3]. Figure 4 shows the *Leckel SEQ47/50* HVS [220]. These devices feature automated

3 There are gravimetric devices capable of measuring multiple size classes using multiple impactors: The Wide Range Aerosol Classifier (WRAC) (sometimes also called *super high volume sampler*) measures a wider size range and

filter changers which collect particles for periods of 24 h on the surface of a filter element and then exchange it for a fresh one. Therefore, it can take between one and three weeks before the data from the gravimetric measurements is available, since the filters are usually collected periodically and weighed in the lab. In addition to this latency, these certified high-precision devices are large, stationary and expensive and therefore usually very sparsely deployed, typically only few stations covering large urban areas.

The gravimetric measurement approach has also been used in wearable monitoring equipment. The tiny *Personal Environmental Monitor (PEM)* [216] is designed for personal exposure tracking and reportedly offers good results, but has the same drawbacks as other gravimetric equipment: the readout is delayed and an average over longer intervals. Also, analysis is difficult for non-expert users.

Figure 4.: *Leckel SEQ47/50* High Volume Sampler (HVS) with automated filter changer.

β-Attenuation

The Beta Attenuation Monitoring (BAM) method uses the absorption of beta radiation by solids to measure PM. It exploits the fact that "*the radiation absorbed is proportional only to the mass of filtered matter and is independent of its density, chemical composition and physical or optical properties*" [142]. BAM typically uses a differential measurement approach in which a filter band collects particles from an air flow and the readings of two Geiger counter detectors are compared, one placed before and one after the flow of sampled air. Advantages of BAMs include high instrument precision and shorter averaging intervals and latency compared to gravimetric measurements [96].

can characterize different size fractions. However, it is not suitable for deployment in the general ambient environment [3].

(Micro-)Mechanical

Mechanical systems that can be used to determine PM concentrations are microbalances. In a so-called Tapered Element Ocillating Micro-Balance (TEOM), particles are deposited small conical glass tube. The frequency of the natural oscillation of the tube is changed through the additional mass of the deposited particles and since mass and frequency correlate, the particle mass can be calculated. TEOM monitors are sensitive to mechanical noise and large temperature fluctuations. In a recent report, the U.S. Environmental Protection Agency (EPA) concludes that, "*while under the correct conditions this method is reliable, it sensitivity presents complications in urban environments, where PM_{10} concentrations are of the most concern.*" [96]

In recent years, Micro Electrical Mechanical Systems (MEMS) resonators have also been proposed for PM detection. Examples are the Film Bulk Acoustic Resonator (FBAR) [80], [176] or the Surface Acoustic Waves (SAW) resonator [101], both of which use thermophoresis to force particles to deposit on the sensor element and then measure the variation in resonance frequency in real-time.

Optical

Optical particle measurement systems are based on some form of interaction between the particles and light. Depending on that interaction, optical measurement can be further distinguished into systems that measure scattered light, or measure the dampening of light, or analyze direct images. Without additional measures, optical measurements are limited to detect particles with a minimum size of ~ 100 nm.

In *light-scattering*, also known as *Nephelometry*, visible or Infrared (IR) light from a Light-Emitting Diode (LED) or a laser is emitted into a measurement chamber through which an aerosol flows. The light is scattered on the surface of the particles and then captured by a photoreceptor which is mounted at a fixed angle (typically between 90° to 120°). Simple so-called photometers that work according to this principle detect a sum signal that aggregates the light scattered by multiple particles in the measurement volume. From this signal, usually a mass concentration is derived as output, assuming a certain material composition in the sampled aerosol. This makes photometers error-prone when sampling aerosols that do not follow the underlying assumptions. Also, without additional technical measures, size discrimination of the particles is not possible.

There are very cheap commercial sensors that work according to this principle, as discussed in-depth in chapter 3.

In contrast, in *Laser Aerosol Spectrometry*, single particles pass through a laser beam or curtain, making it possible to obtain particle counts. By analyzing the shape of the signal pulses, these counts can additionally be assigned to different size bins. Therefore, devices using this technology can simultaneously measure multiple size fractions, e.g. PM_{10} and $PM_{2.5}$. They tend to be accurate and easy to use, which is why the EPA has approved one laser aerosol spectrometer for the measurement of $PM_{2.5}$ [96]. To minimize coincidence problems, i.e. that multiple particles are measured simultaneously, different approaches exist, such as accelerating the air stream to pull the particles apart.

Different technological measures exist that address the limitations of optical measurements. An often applied approach with photometers is to use impactors that limit the inlet stream to particles below a certain aerodynamic diameter. A Dynamic Mobility Analyzer (DMA) is a device that charges particles in order to separate them by size through their different mobility in an electric field, yielding a fine-grained size segregation. Condensation Particle Counters (CPCs) increase the size of particles by leading them through a saturator where they act as nucleation centers for droplets, thus enlarging them. So called Scanning Mobility Particle Sizers (SMPSs) combine several of these techniques to deliver accurate size discrimination over a wide range.

In *Aethalometry*, the sampled air is sucked through a filter on which the particles are collected. With increasing time, this creates a patch of increasing density, i.e. darkness on the filter material. Continuous or periodic measurements of the transmission of a laser beam through the collected deposited particles are performed until an upper limit for the density is reached. Then, the filter material is moved and a new collection patch is started. The increase in attenuation of the laser beam over time is proportional to the concentration of BC on the filter and thus in the ambient air. A system to measure BC with cellphones that is based on this principle was presented by Ramanathan et al. [191].

In *direct imaging*, an aerosol flows in between a light source and a camera. Passing particles occlude the light, making it possible to detect them through image analysis. The *CAMSIZER 2X* is a product that enables measurement as well as size and shape segregation of particles between 0.8 µm and 8 mm. A drawback of the system is that it takes very high concentrations for reliable measurements [194].

Another camera-based approach is *dust deposition imaging* [188]. The detection simply is periodically taking an image from a high-resolution Complementary Metal–Oxide–Semiconductor (CMOS) or Charge-Coupled Device (CCD) image sensor, which is installed at a 45° tilt. Through illumination with a uniform light source, deposited particles occlude individual pixels that can in turn be counted by differentially comparing images over time. Since it relies on gravitational sedimentation, this technique is aimed at indoor measurement as it can only measure particles in turbulence-free environments. Because of relatively low deposition rates of small particles, sampling intervals need to be long compared to other methods and only the coarse fraction can be detected. The operation principle has been successfully evaluated in the Vatican museums in Rome [189].

Acoustic

In *Photoacoustic Spectrometry (PAS)* [184], particles travel through the beam of a modulated laser, i.e. a laser with varying intensity over time, sometimes stronger, sometimes weaker. Illuminated particles absorb the energy and heat up, subsequently releasing the thermal energy into the ambient air. As the air heats up, it expands and creates a sound wave, the frequency of which is exactly the modulation frequency of the laser which was used to excite the particle. The sound wave is then detected with a very sensitive microphone and its amplitude is proportional to the absorption capacity of the particle.

Noise has also been employed for indirect estimation of particle exposure. Generalized Additive Models (GAM) built on the spectral evaluation of traffic noise have been shown to give good predictions of BC concentrations, even in different cities and cultural contexts [70].

Capacitive

An emerging paradigm for the measurement of particulate matter is capacitive detection. Carminati el al. have recently proposed this approach that makes it possible to detect the tiny changes in capacitance (in the zF range) that occur when a single microparticle enters the electric field between two electrodes. From the amplitude of the capacitance increase the particle diameter can also be estimated [55]. While still being developed further, the approach is in principle suitable for in-flow measurement of particles and has recently been reported to be able to detect particles down to 1 µm [54].

Thermal

The *Single Particle Soot Photometer (SP2)* [133] is an instrument designed specifically for the measurement of Black Carbon (BC). It measures the mass of the particles. To do this, the particles are illuminated by an extremely powerful laser (4 MW). They absorb the energy from the laser beam which heats them up so that they burn up. As they burn up, their mass is converted into thermal electromagnetic radiation. This so-called *Black Body Radiation* can be measured and is proportional to the mass of the particle.

Remote Sensing

The term *Remote Sensing* covers all methods that do not measure particles *in-situ/on-site*. Most of these techniques indirectly gather information on the *Aerosol Optical Thickness (AOT)*, which in turn can be used to make conclusions about the particle number. The gathered information often is of a qualitative nature.

So-called *Ceilometers* are *LIDAR* instruments that vertically emit short pulses of invisible laser light upwards into the sky. A receiver located next to the laser catches the backscatter from the aerosols in the atmosphere and the signal propagation delay of a series of pulses can be used to determine a vertical profile of the aerosol concentration. The main purpose of ceilometers is cloud base determination, i.e. specifying the height of clouds. However, the data on vertical visibility can also be analyzed to gain insights into the possible presence of particles in the air. Due to the relevance of cloud height for air traffic, there are large global deployments of ceilometers, i.a. installations at almost every airport.

Sun Photometers also are ground-based instruments, but contrary to ceilometers, they are pointed at the sun, measuring its radiance. They are also passive instruments, i.e. they need solar radiation to measure. Automatic versions track and follow the sun's path. The received solar radiance is affected by the earth's atmosphere and through multiple readings using appropriate spectral channels, the Aerosol Optical Thickness (AOT) can be determined and deductions about the aerosol concentrations can be made. Similar to ceilometers, there are large planet-wide deployments of sun photometers, complementing satellite instrumentation, e.g. the network AERONET [108].

Satellite-based *radiometers* can also be used to quantify the AOT. They are passive instruments that need downward atmospheric radiation to measure, i.e. they only work in the daytime. The principle is again similar: Radiometric alterations due to the optical atmospheric effects of aerosols are detected by instruments from multiple satellites comprising planetary remote sensing aerosol monitoring networks [108]. Among these approaches, *Spectropolarimetry* not only tries to gather information on the spatial distribution of atmospheric particles, but also on their microphysical properties. By observing the Degree of Linear Polarization (DoLP) under different angles and different wavelengths, information on the refractive indices of particles can be obtained for specific locations, which in turn can be used to gain information on their shape and size [132]. The *Spectropolarimeter for Planetary Exploration (SPEX)* system was actually initially developed to study the Martian atmosphere, but can also well be employed to study that of Earth.

The same approach has been applied to realize ground-based handheld remote sensing: The *iSPEX* system is an ultra-low-cost, mass-producible hardware add-on for the *iPhone* that enables crowd-sourced spectropolarimetric measurements, which aggregate multiple readings for quantitative remote aerosol sensing [217]. In this thesis, the *iSPEX* system was also used in one of the exploratory studies to gain insights into the sensing behavior of non-expert users (see chapter 6).

2.2 SENSING SCENARIOS

Particulate Matter (PM) sensing can be applied for a wide range of scenarios and applications.

A major aspect of distributed networks are the implemented deployment, collection and processing schemes, all of which affect the coverage. There are several possibilities for this, many of them established in Wireless Sensor Networks (WSNs). In WSNs however, there are often dedicated data sink nodes and direct (possibly multi-hop) communication between nodes, leading to different challenges, such as contact detection and mobility control [76]. In our approach to distributed *PM* sensing, we focus on direct data dissemination (e.g. via WiFi or mobile broadband) and centralized processing, as this facilitates low latency process-

Collection Scheme	**Advantages**	**Constraints and Drawbacks**
Static Infrastructure	no or weak power constraints; no privacy issues	low, fixed coverage; calibration intricate
Vehicular (non-personal, scheduled)	periodic calibration possible; no privacy issues	GPS needed; limited, fixed coverage; possible mobility issues
Vehicular (non-personal, unpredictable)	potentially large coverage; no privacy issues	GPS needed; possible mobility issues; possible power constraints
Vehicular (personal, unpredictable)	potentially large coverage	GPS needed; possible mobility issues; possible privacy issues; possible power constraints
Participatory Sensing	potentially large coverage; socialization of costs and benefits; supervised measurements	battery operated, hand-held equipment needed; GPS needed; possible mobility issues; possible privacy issues

Table 3.: Constraints, benefits and drawbacks of different collections schemes.

ing and hybrid approaches with heterogeneous devices as well as fusion with other data sources, e.g. meteorological data or traffic density [257], as well as calibration [104]. On the top level, collection can be differentiated into static and mobile schemes. Using statically deployed sensor nodes has the benefits that the measurement environment is well-known and there is often the possibility to power the sensor devices directly. However, to cover a large area, many nodes need to be deployed and, more importantly, maintained. This means that either (potentially more expensive) sensors with strong calibration stability need to be employed, or frequent re-calibration with high personnel cost — or again mobile agents — is necessary. The alternative to fixed deployment is the attachment of sensor devices to mobile entities. In contrast to some WSNs, in the outlined scenarios there is usually no coverage control, i.e. no control over the movement of these entities.

However, knowledge concerning their mobility patterns may exist. Mobility can be either *unpredictable*, i.e. the mobility properties (speed, direction and future positions) of the sensing agents are unknown and possibly highly dynamic. This is the case in Participatory Sensing scenarios involving end-users and/or unscheduled vehicular platforms, such as taxi cabs [260], logistics fleets or bicycles, such as the *Aeroflex* [85]. These scenarios potentially offer great coverage, but re-calibration is difficult, as co-location of sensors cannot be predicted well, if at all. However, in the case of city bikes or taxis, periodic re-calibration may be implemented at taxi stands or — if present — bicycle drop-off stations, that already have an uplink anyway (see Figure 5).

Figure 5.: Public rental bicycles in Beijing, P.R. China. City bikes with docking stands could provide a natural point for calibration, as they are typically already equipped with power and network access.

Calibration is much easier in scenarios using scheduled entities that travel along fixed paths, e.g. by deployment on public transportation infrastructure such as trains, trams or buses, which have regular and reliable routes. Example projects that employed such platforms are *OpenSense* [1] or *PMetro* [56], where PM measurement equipment was installed on the roof of trams. However, in such deployments, possible effects due to the mobility may occur. As sampling heads are usually not designed for air intake while moving, this may affect data quality. This is why, in the *Aerotram* [100] a dynamic sampling inlet was employed, which automatically adjusts the air intake according to the speed of the train.

An important challenge that arises in some collection schemes is privacy, as sensor data is only useful with accurate location data, which may in turn reveal sensitive information about individuals (personal vs. non-personal collection). Interestingly, some of the presented schemes feature a sort of "intrinsic privacy": Taxi drivers do not reveal personal information and public city bikes are not used by the same person for an extended period of time (and are tracked anyway, so no additional information is collected).

Following this general distinction, we present an overview of different use cases.

2.2.1 *Regulatory Compliance Monitoring*

The classic use case is *regulatory compliance monitoring*, but the accuracy and stability requirements are very high and thus unlikely to be satisfied by low cost instrumentation. In fact, there are very few instruments that are certified for regulatory compliance monitoring today. Nonetheless, complementing classical measurement grids with fine-grained low-cost dust sensing approaches has the potential to provide municipalities with important information, e.g. as means for finding and selecting areas in which measurements with more accurate equipment seem prudent.

2.2.2 *Personal Information*

Since the effects of both long-term and short-term exposure may vary greatly between individuals, any standard or guideline cannot completely protect each individual [227]. As a result, the need for fine-grained mobile measurements in order to monitor people at risk arises.

Personal measurement can be done based on different motivation: In *Occupational Exposure Monitoring* it is often compulsory. Similar to the kind of devices which people carry in nuclear facilities in order to measure and record their occupational exposure to radioactivity, personal particle samplers can be applied in potentially hazardous environments such as factories, chemical plants, coal mines or woodworking shops. Also, people who know of their higher-than-usual susceptibility to certain environmental conditions could use such devices as a personal warning system or exposure log. However, since such scenarios are highly health relevant, a sufficient accuracy needs to be reached.

People may also want to be sure on an informal level that they are not overexposed to high concentrations of particulate matter. Such *Personal/Life Log* or *Quatified Self (QS)* scenarios are in principle similar to the one described before. However, the focus lies more on coarse information rather than precise measurements here, much as it is the case with cheap commodity UV-meters or thermometers anyone can buy for a few dollars. Therefore, precision requirements are much lower.

2.2.3 *Urban/Participatory Sensing*

With sufficient density and appropriate networking, such individual sensing can be extended to realize *Participatory Sensing (PS)* by communities and the construction of pollution maps for entire urban areas. Urban City Sensing approaches have been proposed in the past, e.g. to create noise pollution maps of urban areas [199],[147]. Low-cost mobile PM sensors enable such approaches for fine dust sensing as well. While it is expected that the accuracy using simple devices is lower than that which can be achieved using expensive stationary equipment, mobile measurements would allow for a much higher spatial and temporal resolution. This in turn might be exploited through clever algorithms that fuse distributed multi-modal data of lower accuracy to a big picture Also, if the measurement equipment is cheap enough, such devices could e.g. allow developing countries to erect inexpensive air quality measurement grids. An important factor in so-called *Participatory Sensing (PS)* systems is the user, as PS intrinsically involves empowering citizens [49].

2.2.4 *Indoor Air Quality Monitoring*

In industry, interest is also growing regarding low-cost mobile PM sensing, especially when looking at the Asian market, where urban air quality is generally more problematic. An upcoming use case is the introduction of cheap PM sensing capabilities into air purifiers, smart items or cars to monitor the *interior air quality* or demonstrate the effectiveness of employed filtering measures. Another aspect is the monitoring of industrial equipment that is exposed to outdoor air, as part of machine health monitoring, e.g. cellphone towers.

2.2.5 *Context-aware Systems*

Another novel use case would be enabling *Context-sensitive Systems* or *Reactive Systems*, such as e.g. dynamic pollution-based traffic control or other approaches to not only monitor, but also effectively combat PM pollution. Such systems can also be conceived on a local scale, e.g. in *Smart Environments* that automatically open and close windows according to pollution levels, etc. In these systems accuracy demands vary greatly, depending on the application.

In *Activity/Situation Recognition*, actual meaningful readings may not even be required at all. Weekly et al. for instance presented a system that detected the indoor occupancy using a low-cost dust sensor [243].

2.3 introductory related work

This section only briefly covers introductory related work. Specific related work pertaining to the contributions of this dissertation is located and discussed in the respective chapters.

A recent book deals with various aspects of Participatory Sensing (PS) systems, among them a full chapter by Theunis et al. on *Sensing the Environment* [229]. In the last years several surveys in the area of next-generation air quality monitoring with different foci have been presented. Snyder et al. presented research on environmental sensing and the paradigm shift that air quality measurement is undergoing in particular [218]. The *EPA* draft of the *Roadmap for Next Generation Air Monitoring* [81] also recognizes a growing support for passive monitoring using handheld and/or wearable sensors.

In terms of technology and systems, Gozzi et al. presented a survey on the mobile monitoring of particulate matter [97]. Work on WSNs for air pollution monitoring, including a review of gas sensors, was presented by Yi et al. [256]. Carminati et al previously reviewed some of the technologies presented above for low-cost compact dust monitoring [55]. Their work includes a nice overview of emerging miniaturized (respectively miniaturizable) technologies for the pervasive monitoring of airborne particulate matter. Devices were also the focus in a review by Jovašević-Stojanović et al. [118],

An extensive report of the British *Air Quality Expert Group (AQEG)* on air quality in the United Kingdom (UK) [4] features a chapter in which different measurement approaches are discussed and compared. The U.S. Environmental Protection Agency (EPA) also discusses several measurement techniques (respectively devices) along with their principle of operation, advantages and limitations [96]. A rich source on a wide variety of information concerning handheld air quality measurements is the EPA *Air Sensor Handbook* [250]. It includes discussion on measurement use cases and according accuracy requirements.

2.4 CONCLUSION

We have presented an overview of possible approaches for the measurement of Particulate Matter (PM), along with a series of possible scenarios. On the question of how to measure fine dust using low-cost sensors with high spatio-temporal resolution, we identified the optical *light-scattering* approach as most suitable, as it is a mature technology for which cheap off-the-shelf sensors are already available. This makes it currently without alternative.

Concerning a collection scheme, we believe that hybrid approaches with many low-cost sensors and some professional reference equipment for re-calibration seem to be most promising for truly large-scale scenarios. As part of this, Participatory Sensing (PS) has the potential to reveal unprecedented information on the urban dynamics of fine dust.

In order to reach this, we argue that the presented challenges cannot be addressed separately and a combination of smart approaches on different levels has the potential to realize systems with an overall performance beyond that of the mere employed instrumentation. We expect that addressing the presented challenges in a combined, holistic approach could be most suitable for low-cost, distributed PM sensing: Motivated citizens, traversing the city they live in, operate intelligible, affordable, mobile instrumentation, the data of which is centrally combined to ensure stable re-calibration as well as high data quality and coverage.

3 Low-cost PM Sensing

In the previous chapter, we identified light-scattering as the most suitable technology for portable, low-cost Particulate Matter (PM) sensing, and a Participatory Sensing (PS) approach as well-suited collection scheme to achieve high spatio-temporal resolution measurements with low latency. This chapter presents a mobile, low-cost particulate matter sensing approach for the use in PS scenarios. It shows that meaningful readings can be achieved with handheld personal measurement devices that carry cheap commercial off-the-shelf (COTS) dust sensors. Parts of this chapter have previously been published. While since then more work on low-cost dust sensors and their performance for PM measurement has appeared, the research presented here was — to the best of our knowledge — one of the first published papers comparing the data of low-cost PM sensing with commodity sensors against professional equipment [35], the first handheld monitor featuring Commercial-of-the-shelf (COTS) dust sensors for PM measurement [33], [34], as well as the first real-world comparison measurements between such sensors and an official government operated measurement station [37].

The development of the *TECO Envboard* was joint work with Matthias Berning, Mathias Busse and Takashi Miyaki, who performed the electrical engineering design and were significantly involved in the firmware implementation. Parts of the sensor characterization and evaluation were carried out by Rayan Merched El Masri for his master's thesis [84].

3.1 Introduction

As motivated in chapter 2, a Participatory Sensing approach seems especially well-suited to realize distributed PM measurements with high spatial and temporal resolution. By providing engaged individuals with low-cost sensing devices, they can potentially quantify their individual exposure and at the same time contribute to accurate city wide estimations. However, suitable tools to track PM levels need to be identified or even developed. Such tools need to be:

- *compact*: Sensors should be small, ideally embeddable into existing ubiquitous technology like mobile phones.
- *inexpensive*: Mobile measurement solutions need to be affordable for Participatory Sensing scenarios to scale.
- *usable*: Usability is key for acceptance. Systems should ideally require as little maintenance as possible, e.g. changing filters, frequent charging, expert calibration etc.
- *accurate*: If readings are afflicted with high noise or uncertainty, the value of the data decreases. Of course, accuracy demands depend on the concrete application case at hand and trade-offs between price and data quality are inevitable. At the very least, readings need to add value to whichever overall system they are used in.
- *responsive*: Finally, in order to identify sources and to enable reactive systems, timeliness (i.e. low latency) of data is desirable.

As argued in the previous chapter, we have identified light-scattering as the currently most suitable technology for pervasive fine dust sensing, as small, portable, low-cost commodity sensors exist and are readily available.

3.2 RELATED WORK

Low-cost light-scattering dust sensors have been increasingly applied to the problem of PM measurement in the last years. Before that, e.g. the early *Personal Environmental Impact Report (PEIR)* project [145] at UCLA used an indirect mobile phone based approach to urban sensing: With the goal of sharing "*how you impact the environment and how the environment impacts you*", logged data included the $PM_{2.5}$ exposure and sensitive site impact ($PM_{2.5}$ particulate impact on sensitive sites such as schools and hospitals). The exposure was not directly measured, but instead calculated based on a variety of parameters such as the proximity to known hazardous conditions or sites, as for example freeways. While this can help people to assess their exposure, it is actually dependent on better base data.

Figure 6.: Range of measurement technology [47].

The *inAir* project [123] presented a phone-based tool to measure, visualize and share local indoor air quality through a specifically developed device in conjunction with a *Dylos* particle counter, which is stationary and has no wireless data transmission capabilities. In the *Common Sense* project [82], a similar approach was developed, featuring a custom handheld air quality monitor, but with a focus on outdoor participatory urban sensing. However, this device did not yet include a PM sensor.

Both the *OpenSense* project [2] and *da_sense* project make use of public transportation vehicles to measure air quality beyond a few fixed measurement stations. While *da_sense* proposes the integration from different sources (such as infrastructure sensors, environmental WSNs or smartphones) so far no *PM* data is integrated. *OpenSense* on the other hand integrates the *DISCmini* from *Matter Aerosol* into the mobile measurement setup which gives a fine grained resolution for the covered tracks. Still, the *DISCmini* is an expensive commercial hand-held particle monitor, far too expensive for larger scale participatory sensing scenarios. Other cheap devices, such as the *Personal Environmental Monitor (PEM)* [216] or the *UCB particle monitor* [63] are in principle suited for personal sensing, but have drawbacks of their own: The *PEM*'s gravimetric measurement reportedly offers good results, but readout is delayed and difficult for non-expert users. The *UCB monitor* is only intended for use in indoor environments. Several studies use semi-professional equipment to monitor PM_{10} levels [182], which is unsuitable for large scale Urban Sensing scenarios because of its cost.

Today, an increasing number of companies offer small, generally embeddable particulate matter sensors that could fit in a hand-held measuring device. Several of these sensors were compared in [35], the results indicating that only few of them actually seem to be suitable for the use in mobile PM measurement scenarios. The next section discusses this class of sensors in more detail. The *Sharp GP2Y1010* has been used in several dust sensing projects [71], [92], [121], [165], [177], [190], [201]. However, many of the papers focus on node design or networking aspects and none of them supply information on how they enabled accurate readings from the simple sensor — or whether they did at all. The *Shinyei PPD42NS* sensor has also been studied quite intensively in the last years, both under lab conditions [10], [150], as well as in the field [93], [110]. CHOI ET AL. very early on presented a Wireless Sensor Network (WSN) sensor node equipped, among others, with this sensor. However, their preliminary evaluation with tobacco smoke showed no remarkable detection by the sensor [61]. HOLSTIUS ET AL. [111] presented the first field evaluation of the *Shinyei PPD42NS*, which is similar to the work presented in this chapter. The sensor — respectively one of its many clones — has also been used in indoor applications such as the *PiMi airbox* [141], [261] for indoor air quality monitoring or as a sensor to detect indoor occupancy [243].

The next section presents a deeper discussion of these low-cost sensors. Related work on emerging approaches of sensing PM directly with cellphones [30], [186], [191], [217] is presented in the next chapter.

3.3 INSTRUMENTATION

For our experiments, we took a series of measurements with cheap commodity dust sensors in order to investigate their general suitability for the measurement of Particulate Matter (PM). Our goal was to observe and quantify the margin of error between our cheap sensors and a calibrated reference device and to assess for which kind of application cheap COTS dust sensors can be used, if any.

3.3.1 *Sensors Selection*

While there is a variety of stationary and handheld dust monitors commercially available, there are not many small sensors to choose from (see Table 4). All listed sensors employ the light-scattering operation principle.

The Japanese company *Shinyei* [214] carries several relatively sophisticated particle sensing modules in the upper price range and one low-cost sensor. Their availability is good and its design has been copied many times: Two Korean sensors – the *SYhitech DSN501* and the *NIDS PS02C-PWM* – are both virtually identical to the design of the *Shinyei PPD42*.

Sensor		Air Intake	Size (mm^3)	Output	Range	Power	Retail Price
Sharp GP2Y1010/ GP2Y1012/ GP2Y1023	[208]	diffusion	46×30×18	analog (TSP)	0 – 0.5 $\frac{mg}{m^3}$	0.1 W	~ 10 $
Sharp DN7C3CA006	[209]	fan	54×54×42	$PM_{2.5}$		0.9 W	~ 30 $
SYhitech DSN501	[8]	heater	59×45×20		0 – 1.4 $\frac{mg}{m^3}$	0.45 W	~ 10 $
Shinyei PPD42NS	[213]	heater	59×45×22		0 – 800,000 $\frac{pcs}{ft^3}$	0.45 W	~ 10 $
Shinyei PPD60PV	[214]	heater	88×60×22	digital (), analog (TSP)	0 – 2,000,000 $\frac{pcs}{ft^3}$	0.7 W	~ 420 $
Shinyei AES-1	[212]	heater	90×90×23	single particles	300 – 300,000 $\frac{pcs}{ft^3}$	3.6 W	~ 1,100 $
NIDS PSX-01E	[168]	heater	59×45×20		0 – 2.0 $\frac{mg}{m^3}$	0.15 W	*n/a*
NIDS PS02C-PWM	[167]	heater	59×45×20		0 – 2.0 $\frac{mg}{m^3}$	*n/a*	*n/a*
Nova Fitness SDS011	[171]	fan	71×70×23	PM_{10}, $PM_{2.5}$		> 1 W	~ 35 $
Nova Fitness SDS018	[172]	fan	59×45×20	PM_{10}, $PM_{2.5}$		> 1 W	~ 40 $
Alphasense OPC-N2	[7]	fan	75×64×60	PM_{10}, $PM_{2.5}$, PM_1		< 5 W	~ 500 $

Table 4.: Specs of candidate dust sensors according to the data sheets.

However, while information on the *NIDS* sensors is available online [167], [168], our attempts to receive a quote for them remained unanswered. The *SYhitech* is also available under as *Apollo DSM501*. Other clones of the sensor design can be found under different names, e.g. the *STBM 271 DUST SENSOR MODULE*.

The second sensor with very good availability is the *Sharp GP2Y1010* optical dust sensor. It is mostly used in air quality equipment, such as air purifiers, and can easily be obtained in large quantities from various distributors around the world. A variant of this sensor is the *Sharp DN7C3CA006*, which is the

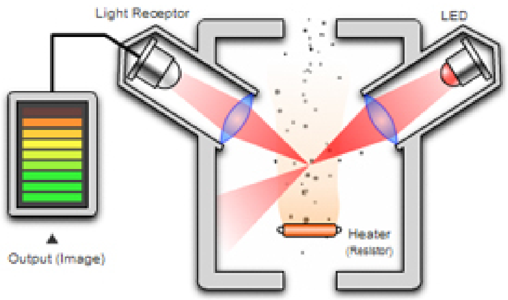


Figure 7.: Many of the sensors use a heating resistor to create an updraft. This limits the possible operation conditions, as the sensor needs to be installed having a fixed orientation (image taken from [214]).

*GP2Y1012*on the inside, but with an added fan and impactor casing.

All of the COTS dust sensors we found are principally small enough to be incorporated into a handheld device, though the larger ones could make such a device cumbersome. The sensors are also all based on the same operation principle: A light beam is emitted into a measurement chamber. When dust is present, the light is refracted by particles and the amount of scattered light is detected. All sensors except the *Sharp GP2Y1010* (and possibly the *NIDS PSX-01E*) additionally use a heating resistor to create an updraft (see Figure 7).

For the applications outlined above, the use of such a heating element has several drawbacks: First, since a current is needed to heat the resistor, the power consumption is generally higher. Second, the response time is higher, since it takes some time – usually around 30 seconds – until the resistor is heated up. Third and most important, the heating imposes strict orientation restrictions during operation. This practically prevents the use for any applications in which the device's orientation cannot be controlled. Finally, heated sensors can not be directly ventilated, because this would influence the heating. This may restrict the use in multi-sensor devices together with other environmental sensors that need an airflow.

Two sensors that were not yet available at the time this research was conducted are the *Nova Fitness SDS011* [171] and the *Alphasense OPC-N2*. We used the *Nova Fitness SDS011* in our calibration experiments in section 5.2. The *Alphasense OPC-N2* is a laser scattering sensor which reportedly performs very well, but also was de-scoped because of its significantly higher price.

We opted for the *Sharp GP2Y1010* optical dust sensor, since it best fits our scenarios' general requirements: cheap, small, low-power, and easily available for our tests.

3.3.2 *The TECO Envboard*

While our initial experiments were carried out with the *Sharp GP2Y1010* dust sensors connected to an *Arduino Mega* single-board microcontroller, for subsequent experiments and our field evaluation we designed and used the *TECO Envboard* [34], a custom AVR-based platform for Environmental Sensing. It

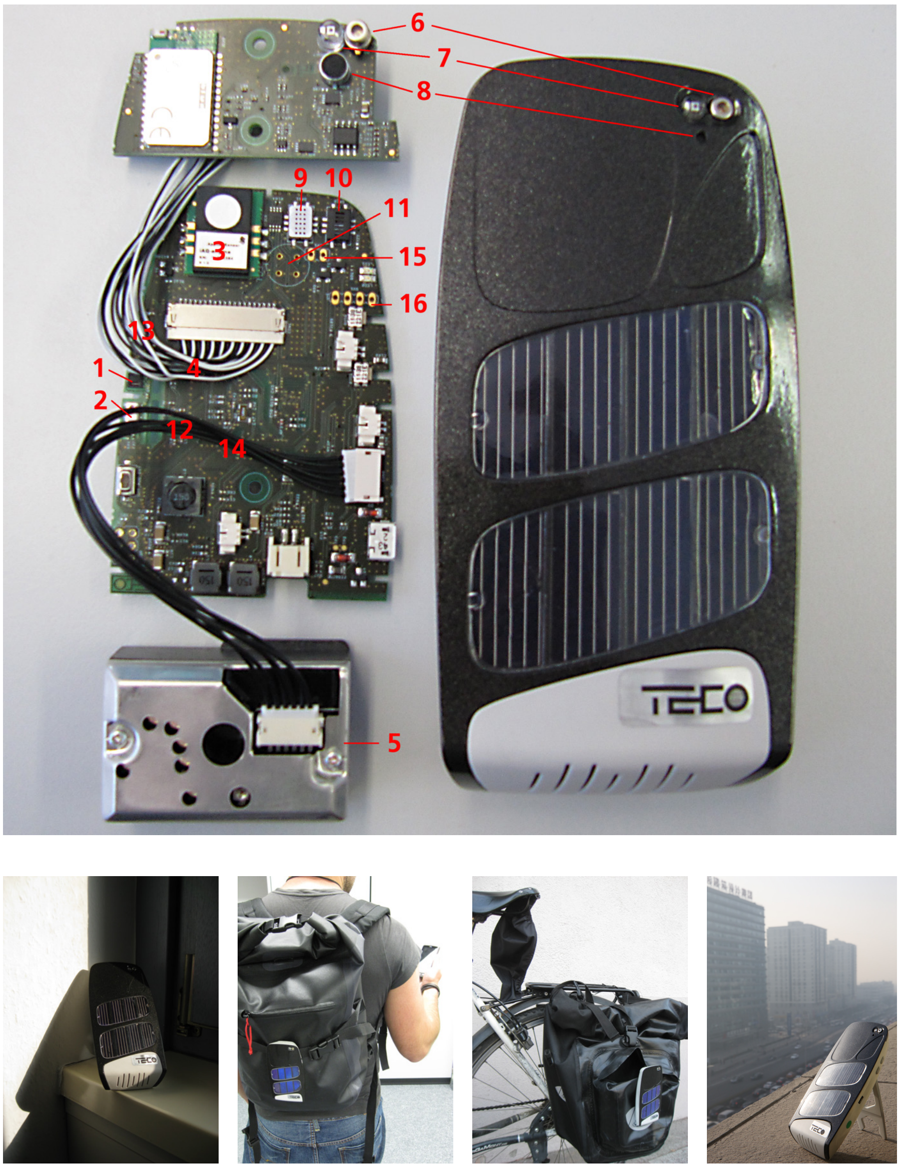


Figure 8.: The TECO Envboard [34] is a handheld multi-sensor platform for research and development.

	Sensor	Phenomenon	No.
digital	SHT-21	temperature relative humidity	1
	MPL115A	atmospheric pressure	2
	iAQ-Engine	VOC (indirect: CO_2)	3
	ADXL345	3D acceleration	4
analog	GP2Y1010	particulate matter	5
	AlGaN-TO18	UV light	6
	TEPT5700b	ambient light	7
	WM-61A	noise level (dBA)	8
	MICS 2614	O_3	9
	MICS 4514	CO, NO_x	10
optional	TGS4161	CO_2	11
	ITG-3200	3D magnetometer	12
	HMC5883	3D gyroscope	13
	MVS0608.02	motion/microvibration	14
	NTC thermistor	temperature (for compensation)	15
	GPS module	global postion	16

Table 5.: Sensors available on the TECO Envboard.

carries a variety of COTS sensors, ranging from weather sensors like temperature and humidity over gas sensors to the *GP2Y1010* for PM sensing.

The *Envboard* is powered by an integrated Lithium Ion Polymer (LiPo) battery or a standard Micro-USB connector supplying 5 V, which is also used for recharging. It can be used in a standalone fashion as well as in conjunction with a host device — e.g. an *Android* phone or a laptop computer — to which it connects via Bluetooth (BT). Sampling of the built-in sensors can either be triggered via the command API (polling), or the *Envboard* can be configured to sample periodically and send the values via BT and/or store them on the integrated microSD card for later readout.

At the heart of the Envboard is an *ATmega 2561* Microcontroller Unit (MCU), which was mainly chosen for two reasons: First, it offers a sufficient number of input pins for the variety of sensors the Envboard carries. Second, the Envboard is intended as a research and development tool that should allow for easy modification. Thus, a MCU that can be programmed using the Arduino language and IDE seemed to be a sensible choice. On the communication side, the Envboard is outfitted with a *Bluegiga WT12* Bluetooth transceiver for wireless transmission.

Alternatively, the device can be configured to communicate in the same fashion through its serial USB interface.

The protocol that is used to interface with the Envboard is based on the *Firmata* protocol[1]. It can be used to request the Envboard's status and adjust its parameter settings. This can be anything from configuring the desired mode of operation (stand-alone-measurement, periodical dissemination, polling or mixed), over setting parameters such as the system time or individual sampling intervals to (de-)activating individual sensors or adjusting calibration data.

Table 5 shows the different sensors that are incorporated into the *TECO Envboard*. These are all Commercial-of-the-shelf (COTS) components. Since we did not need them, some of the sensors (mainly the inertial sensors) were not populated in the device revision we used for our experiments. Still, the corresponding footprints and/or connectors are in place and allow quickly adding them.

The Envboard's housing protects the PCB from environmental influences. Particularly, it shields the optical dust sensor from possible outside sources of error, such as fluctuating ambient light conditions [234] or bedewing [207]. A pair of micro-fans ensures a constant air flow through the device and the dust sensor, which enables continuous measurements while reducing the risk of residual dust staying trapped in the sensor and compromising the readings.

3.3.3 *GP2Y1010 Dust Sensor*

As described above, the *Sharp GP2Y1010* employs light-scattering as its operation principle: An Infrared (IR) light beam is emitted into a measurement chamber. When dust is present, the light is refracted by particles and the amount of scattered light is detected. The measurement chamber is designed to be a light trap, so that only the refracted light falls onto the receptor (see Figure 9). While the sensor had been used in previous work and seemed promising, it was clearly not designed to provide accurate absolute readings. The *GP2Y1010* is intended for the use in air conditioners and air purifiers, its default detection granularity is limited to the coarse distinction between *house dust*, *cigarette smoke*, and *no dust* [207]. Although its data sheet shows an exemplary relationship between the dust density and

1 `https://github.com/firmata/protocol`

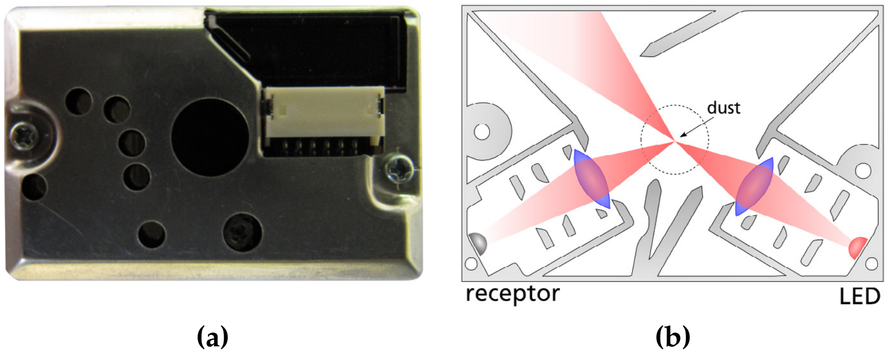


Figure 9.: The *Sharp GP2Y1010* dust sensor (a), and the structure of its light trap on the inside and visualization of its operation principle (b).

the sensor's output voltage, it states that the graphs are "*just for reference and are not for guarantee*" [208].

In order to test its performance, we conducted a series of parallel measurements with the *GP2Y1010* and a high-accuracy laser photometer as a reference device, the *TSI DustTrak DRX 8533 Aerosol Monitor (DRX 8533)* [232]. After applying an approximation of the curve from the data sheet to the readings of the sensor and comparing it to the measurements of the *DRX8533* reference device, we can see that the sensor output is very noisy and the curves do not match (see Figure 10). This,

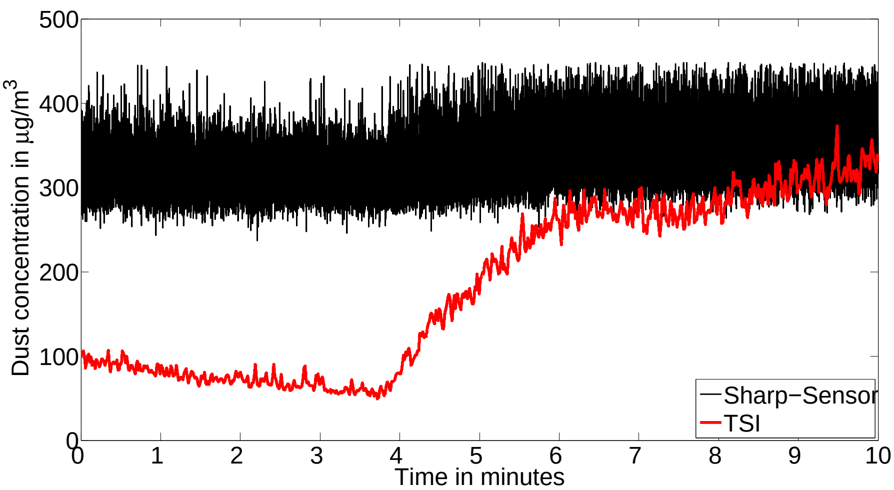


Figure 10.: Raw readings of the *GP2Y1010*, computed according to the exemplary reference curve in the datasheet [208], vs. those of the *DRX8533*.

along with the fact that different specimen of the sensor displayed strongly varying output levels, lead to experiments with signal processing and calibration.

3.3.4 *Accuracy Improvements*

We started developing our refinements by investigating the performance of the *Sharp GP2Y1010* and its ability to measure the particulate matter concentration in the air using the setup described in [35]. All sensors were used as they were delivered, using their unmodified factory sensitivity settings. We sampled the sensors at the maximum possible frequency according to the LED pulse width and waiting times documented in the data sheet [208], which resulted in a sampling rate of ~100 Hz. The *DRX 8533* reference monitor sampled at its maximum rate of 1 Hz, calibrating it according to the manual [232] prior to each measurement run. We neither used impactors nor filters to keep our samples clean from coarse dust.

Noise Reduction

The first step towards de-noising the sensor output was eliminating the outliers and thereby smoothing the output. Since our reference device was sampled at 1 Hz, we also sliced the *GP2Y1010* readings into windows of 1 s length and calculated the median over the 100 samples. The results are shown in Figure 11. A correlation between the *Sharp GP2Y1010* output (upper curve) and the *DRX8533* measurements (lower curve) becomes more clearly visible. As the particulate matter concentration decreases from about $100\frac{\mu g}{m^3}$ to $50\frac{\mu g}{m^3}$ within the first four minutes, the *GP2Y1010* output shows a similar tendency and decreases as well, albeit only slightly. The increase in dust concentration between the fourth and the sixth minute is also reflected in the sensor's readings. As a second process step to further reduce the noise, we applied a moving average filter with a window size of 60 s (i.e 60 data points) on the data. We separated the noise reduction into two steps, because the first one can be easily carried out in the sensing device before logging or transmitting the data. By this, we can achieve data reduction without losing significant information. The second step for further smoothing can either be carried out on the device or on a back-end system. By adjusting the window size, a trade-off between accuracy and timeliness can be made.

Calibration

Using these improvements, we attempted to calibrate the *Sharp GP2Y1010* by mapping its output to the corresponding partic-

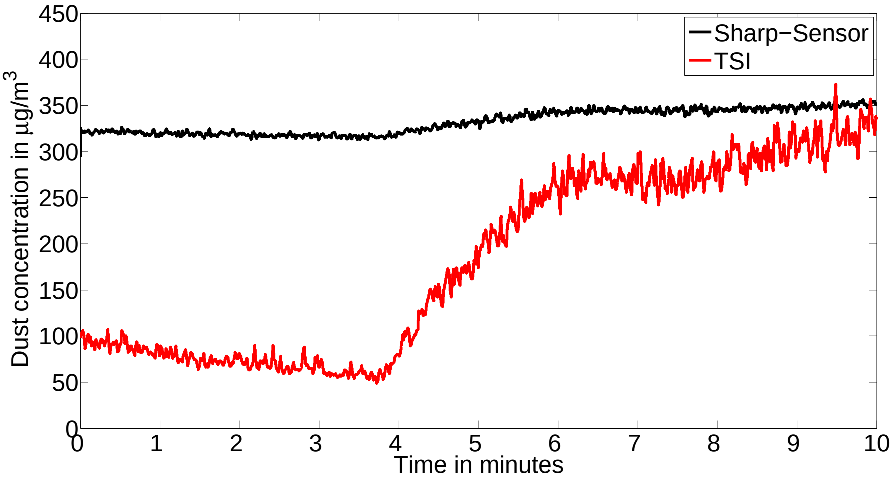


Figure 11.: De-noised sensor output by averaging (median) over 1 s-windows.

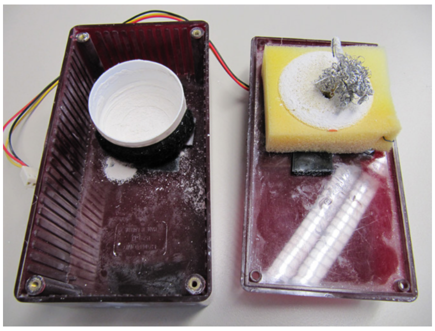
ulate matter concentration, in order to later allow the direct calculation of the dust concentration in the air. The sensor does not feature different channels or any other means to distinguish between particles of different sizes. Instead, we derived different calibration coefficients for PM_{10} and $PM_{2.5}$ respectively. To have a broad spectrum of dust concentrations for calibration, we built a self-made dust dispenser (see Figure 12). It basically consists of a box and fan that is connected to a small bale of steel wool (a). When the fan is turned on, the steel abrades chalk inside of the box and blows it into the outer containment (b). A filter sheet is used to prevent too much dust being dispensed at once. In the full calibration setup, the air flows through the dispenser, then into the box containing the *Envboards* and finally through the *DRX8533* (c). This dispenser makes it possible to quickly generate high dust concentrations which will decay slowly after turning off the dispenser. By alternating dispensing and ventilation phases, we enabled readings over the full spectrum of the sensor. For the actual calibration of the sensors we performed measurements over 18 hours, again sampling the *GP2Y1010* at 100 Hz and the *DRX8533* at 1 Hz. The dust dispenser was set to be turned on for 15 minutes once an hour. This lead to a repeated sequence of rising and falling dust concentrations, allowing the sensors to repeatedly measure different concentrations levels.

We first applied the two de-noising steps that were described in the previous section. The second step was also applied to the readings of the *DRX8533*. Based on this data, we calculated a linear scale factor a and offset b between the two curves as coefficients for the raw readings x to calculate the concentration $\rho(x)$:

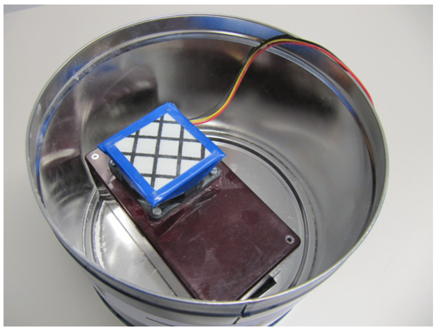
$$\rho(x) = a \cdot x + b$$

The results of these steps are depicted in Figure 13, once after the first de-noising step (a) and once after the subsequent smoothing of both curves (b). The graph's ordinate represents the time (in min) and plotted on the y-axis are the readings of the *GP2Y1010* (10-bit ADC-values, black curve), respectively the PM_{10} values measured by the reference device (in $\frac{\mu g}{m^3}$, red curve). These figures clearly show that it is possible to align the readings of both devices by linear calibration coefficients.

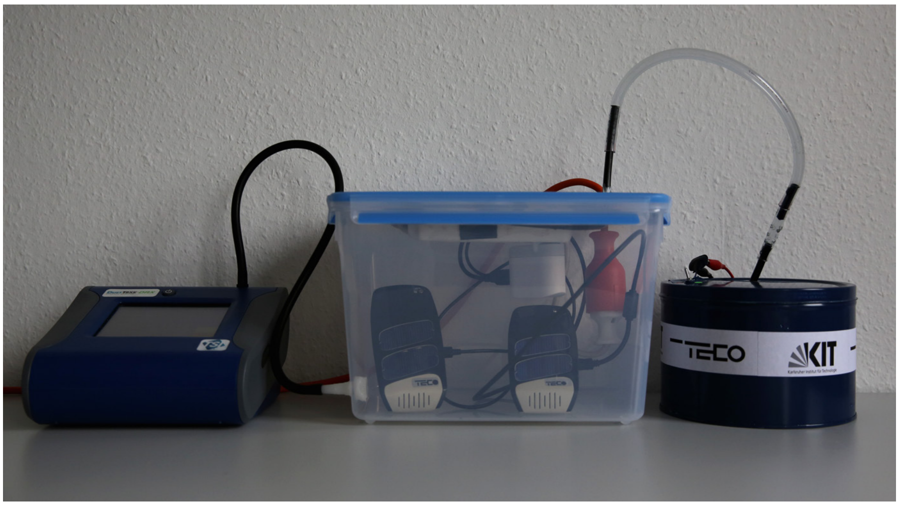
However, when applying the calibration data on consecutive measurements, we encountered new problems: We discovered that the offset of the sensor seemed to "'jump around" between different measurement runs, i.e. the sensor baseline



(a)



(b)



(c)

Figure 12.: Calibration setup: dust dispenser box with chalk reservoir and steel wool (a), outer containment (b), and complete setup with Envboards (c).

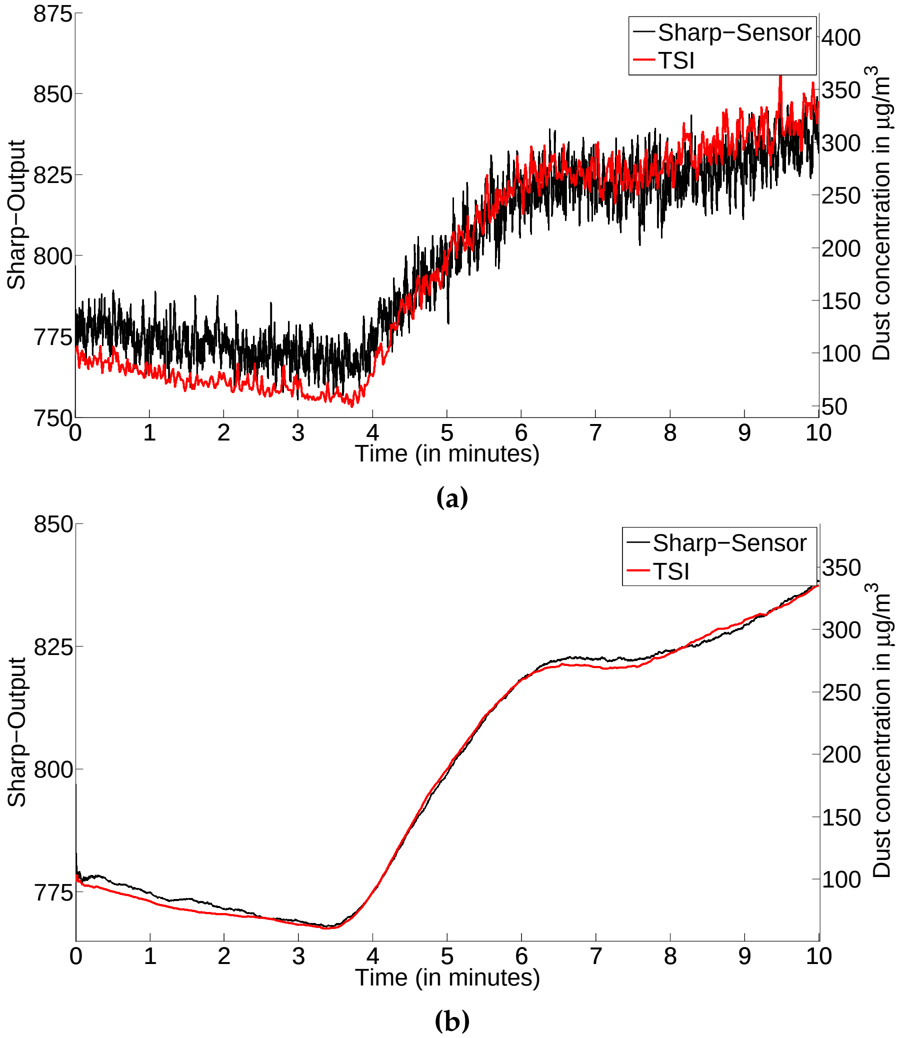
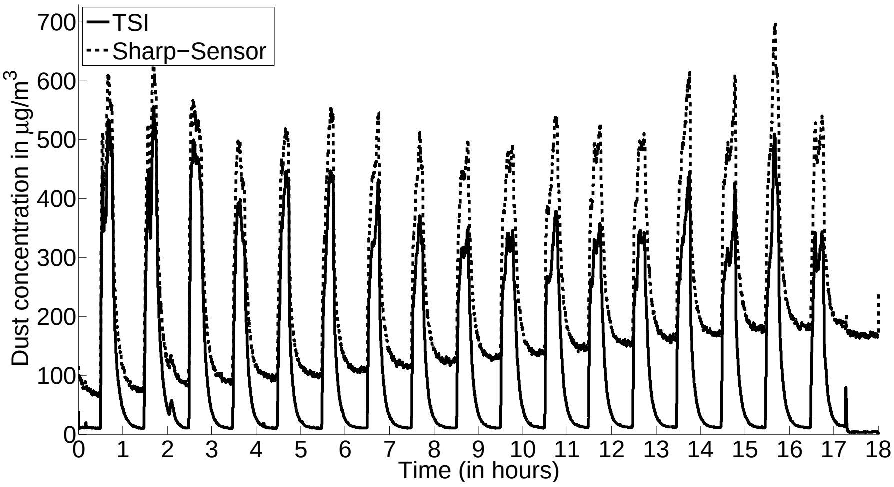
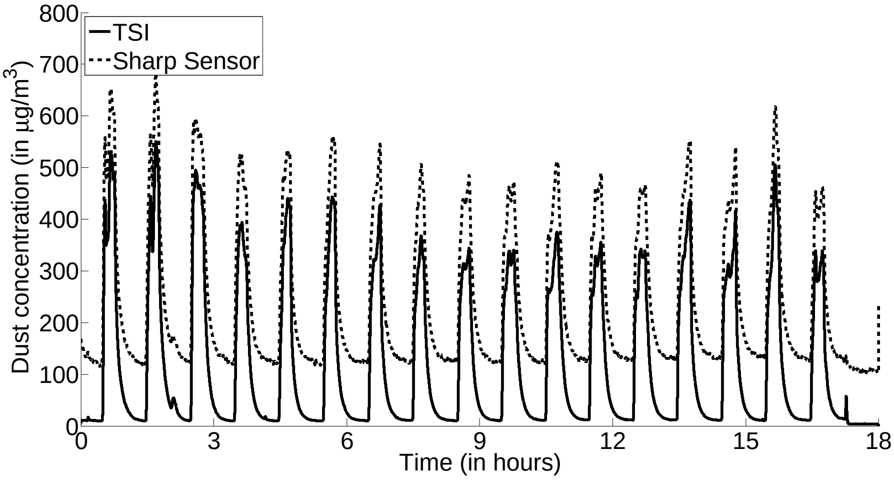


Figure 13.: Processing by de-noising and linear calibration: (a) slicing into 1 s windows, (b) smoothening through moving average filter with 60 s window.

de-calibrated. Also, the sensors displayed a significant drift over time. Both effects can be observed in Figure 14. The graph shows an 18-hour sampling session with the dust dispensing pattern described above. We applied the coefficients derived from a previous calibration run. In order to quantify the drift, we examined several sensors over multiple measurement runs. We found that the drifting behavior exhibited was nearly linear with time and very similar for multiple passes. Thus, we were able to reduce the drift by simple relative baseline manipulation. We introduced a separate calibration step for each sensor to determine its time-dependent drift factor k. Using this, we



(a)



(b)

Figure 14.: (a) Drift when applying the calibration on a second 18 h-measurement and (b) compensation through simple relative baseline manipulation.

adjusted our calculation of a and b. This lead to the following new formula for calculating the concentration ρ:

$$\begin{aligned}\hat{x}(t) &= x - k \cdot t \\ \rho(x,t) &= a \cdot \hat{x}(t) + b \\ &= a \cdot (x - k \cdot t) + b\end{aligned}$$

Figure 14 (b) shows the result. Still, we saw further room for improvement. In order to tackle this, we examined the effects of other parameters on the *GP2Y1010* output.

Device/Sensor		Detection Method	Rate	Range	Price
Envboard [34]					
Sharp GP2Y1010	[208]	light (IR) scattering	1 Hz†	~0 – 500 $\frac{\mu g}{m^3}$	$ ~10‡
DustTrak DRX 5833	[232]	light (laser) scattering	1 Hz	1 – 150,000 $\frac{\mu g}{m^3}$	$ ~9,000
State-operated Measurement Station					
Grimm EDM 180	[98]	light (laser) scattering	~ 10 min	0.1 – 6,000 $\frac{\mu g}{m^3}$	$ ~35,000
Leckel SEQ47/50	[220]	gravimetric	24 h-mean	*n/a*⋆	$ ~20,000
Leckel SEQ47/50		gravimetric	24 h-mean	*n/a*⋆	$ ~20,000

† Using the de-noising steps presented in this work. The maximum raw sampling rate is ~100 Hz.
‡ Cost of the analogue sensor. Additional costs for the data logger platform.
⋆ Gravimetric measurements do not have an upper bound except their total filter capacity.

Table 6.: Comparison of measurement equipment in the different evaluation settings.

Sensor Fusion

At this point, we switched to using the *TECO Envboard* sensor platform [33], since there is a documented temperature dependency of the *GP2Y1010* [207]. We analyzed the readings of the *Envboard*'s internal *Sensirion SHT21* digital temperature and humidity sensor. There is a very strong relationship between the readings of the two sensors. To correct for this, we again devised a linear compensation[2] as a function of the temperature T according to measurements taken at a reference temperature T_0 of 20 °C. We introduced another calibration step after the drift compensation and before calculating the scale factor and offset, again leading to a revised formula for calculating a and b, respectively ρ:

$$
\begin{aligned}
\hat{\hat{x}}(T) &= \hat{x}(t) + \alpha_T \cdot \Delta \\
\rho(x,t,T) &= a \cdot \hat{\hat{x}}(T) + b \\
&= a \cdot (x - k \cdot t + \alpha_T \cdot \Delta T) + b
\end{aligned}
$$

Overall, the combination of these steps greatly improved accuracy of the readings. However, the already observed effect of the offset de-calibration could not completely be eliminated by this. While the scale factor a could be accurately derived from the calibration process, several independent influences lead to

2 We expect this formula to perform less well in extreme temperatures and aim at replacing the linear correction in the future.

a shift in the offset b. The base line of the sensor output shifts not only with varying temperature, but also depending on other factors, such as changes in the measurement frequency (even though within specification). As a result, we neither sampled the sensor irregularly nor changed the fixed sampling frequency between measurement passes. We also observed shifting base levels depending on the subset of sensors that we sampled. While this may be due to device peculiarities, we decided to use a fixed set of sensors for all our consecutive measurements. Even so, we kept encountering changes in the offset between measurement runs. The offsets seemed to change randomly every time the sensors are turned off and on, even after trying to remove any residual charge. Therefore, we decided to make use of additional information we may have in Participatory Sensing scenarios to combat this problem.

On-the-fly Calibration Correction

While all previous improvement steps took place on the device level, Participatory Sensing scenarios have the potential to further improve measurement accuracy by sharing information across devices. This can be as simple as averaging readings from co-located devices to reduce measurement errors. More sophisticated approaches may take the shape of the actual data, dispersion models, calibration age, device type, etc. into account when correcting values as well. An example for the application of instant calibration of low-cost gas sensors, either in each other's vicinity or even multi-hop, was presented in [104]. We propose to use the data from co-located sensors to eliminate the problem of offset de-calibration that the *GP2Y1010* described above. In order to do this, we used measurements from a co-located reference point to correct the calibration of the hand-held devices. A reference point can either be a high-precision professional measurement station or another device which has a high confidence that it is correctly calibrated. The device that carries the *GP2Y1010* sensor then uses the reference values to correct its bias. As we only intend to correct changing offsets, only very few measurements have to be transmitted from the reference device to achieve notable improvement. We show the potential improvement by simulation in the next section of this paper. More sophisticated device-by-device calibration techniques, including privacy-protecting measures, are presented and discussed in chapter 5.

3.4 EVALUATION

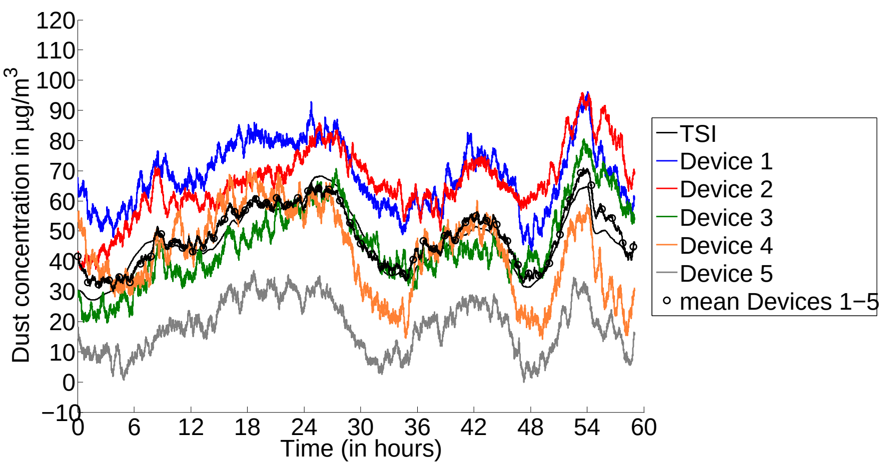
Aside from the hours of measurements we made throughout the process of improving the sensors' accuracy, we conducted two longer measurement sessions in order to evaluate the performance of our system under operating conditions: Firstly, we did a controlled indoor evaluation of the calibration. Secondly, we co-located the sensor platforms with official state-owned measurement stations. Thirdly, we simulated on-the-fly calibration correction for all evaluation runs and discuss the possible improvements. In addition to our sensor boards and the reference device, we obtained the data from the officially approved measurement equipment that is used in the state's monitoring stations. Table 6 shows an overview of the measurement equipment that was used in the test. It is noteworthy that the *GP2Y1010* dust sensor costs only a fraction of the reference

Figure 15.: Setup of the indoor lab evaluation: Six *Envboards* and the *DRX8533* as reference.

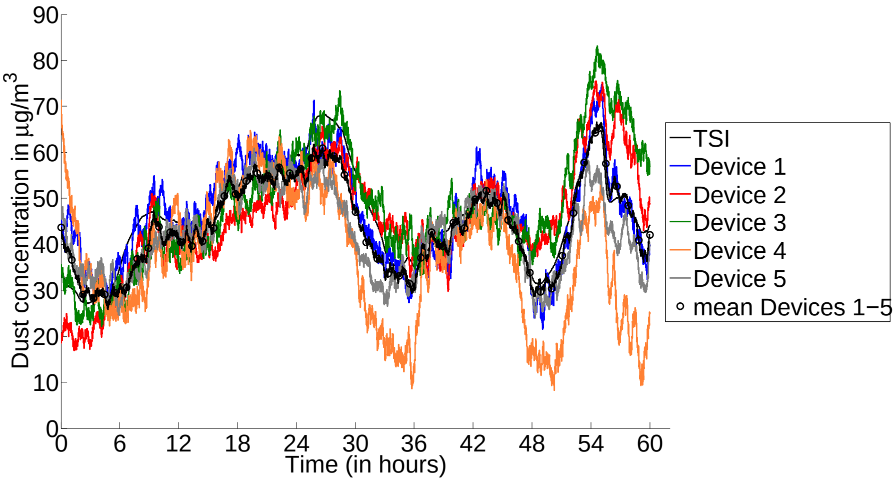
devices. This section shows how well our improved readings compare to the accuracy of the professional equipment.

3.4.1 *Lab Evaluation (Indoor)*

The first session was an indoor evaluation of our processing steps. In contrast to the prior calibration, our sensor platforms were only co-located with the reference meter, but not sampling the exact same air flow (see Figure 15). We measured the indoor particulate matter concentrations using six *TECO Envboards* and the *DRX8533*, which was only sampled every fourth second, since the maximum sampling frequency is limited by the internal logging space (18 h at 1 Hz) and we intended to validate our



(a)



(b)

Figure 16.: Indoor $PM_{2.5}$ evaluation: (a) calibrated *GP2Y1010* sensors against the *DRX8533* reference, (b) values after on-the-fly correction.

refinements over a longer period of time (three days). The measurements of the $PM_{2.5}$-concentration are shown in Figure 16 (a).

As expected, it is clearly visible that the readings from the calibrated hand-held devices show a strong correlation to those of the reference device, the scale factor calibration was successful. However, it can also be observed that the problem of offset decalibration persisted. The *DRX8533* measured an average of

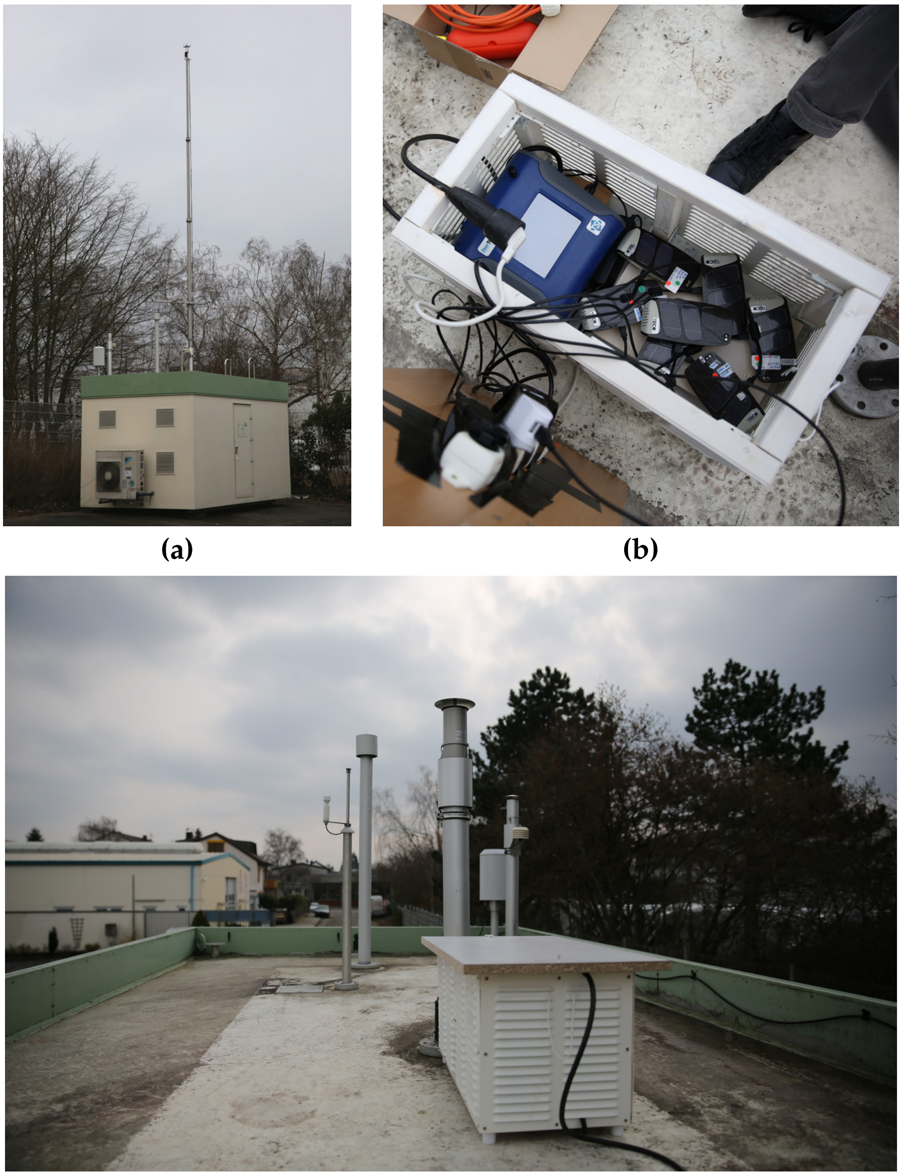


Figure 17.: Field evaluation: state-operated measurement station (a), equipment in weather protection box (b), and installment on rooftop (c).

$46.8\frac{\mu g}{m^3}$ over the 60 hours, the *Envboards* measured averages between $18.5\frac{\mu g}{m^3}$ and $68.5\frac{\mu g}{m^3}$. This can be already considered to be very accurate in light of the intended use of the *Sharp GP2Y1010*. Further calculations lead to even better results: By simply taking the mean of the five devices, we arrive at a very close match to the values measured by the *DRX8533*. However, this is not generalizable and only limited trust can be put into the values of a single device.

This is why we continued to simulate the on-the-fly calibration correction we presented earlier. Figure 16 (b) shows the dust concentrations measured by the hand-held devices after applying the on-the-fly calibration step. We randomly selected three consecutive data points from the reference device and "transmitted" them to the mobile devices, which in turn "calculated" the difference between the locally measured values and the reference value in order adjust their offset accordingly. This notably improved the accuracy of the devices. Similar to $PM_{2.5}$, the PM_{10} curves of the hand-held devices show the same general behavior. Without on-the-fly calibration, the offsets were a little larger, and the simple mean did not fit as well. After simulating on-the-fly calibration, the gain was comparable to the $PM_{2.5}$-case.

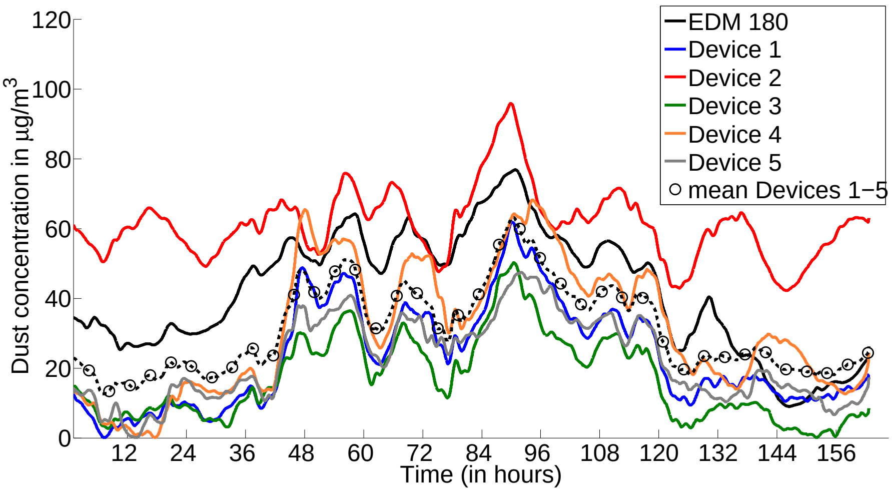
3.4.2 *Field Evaluation (Outdoor)*

For our field evaluation, we co-located several *Envboards* with an official state-owned station that measures different types of background pollution. Our measurements took place in the late Winter 2012/13. We used the same, unaltered devices as in the lab evaluation, the only difference being that we placed them inside a small, well ventilated box in order to shield them from rain and snow (see Figure 17). We added the *DRX8533* as well. This setup was then placed on the rooftop of the measurement station, next to the air inlets of the other samplers, and logged for seven days continuously. After retrieving our setup, we compared the data of the official measurements to our own.

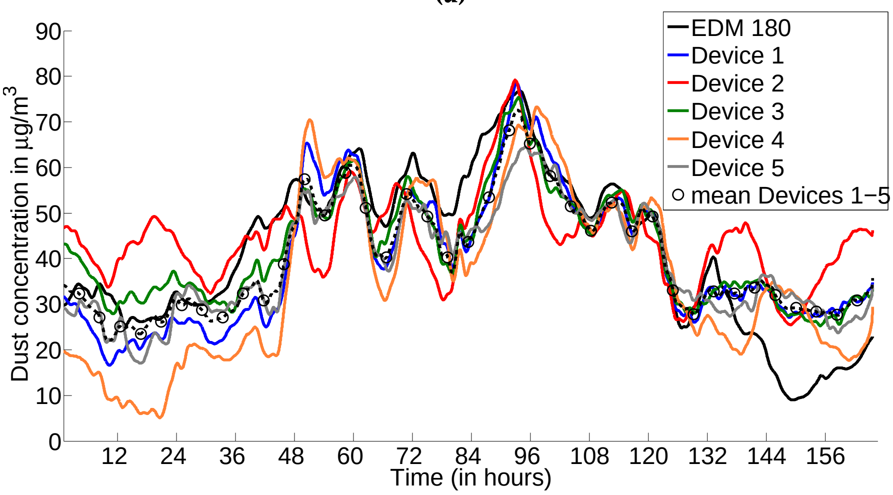
The state uses three measurement devices at the station, one optical and two gravimetric (for details, see Table 6). The *Grimm Technologies Model EDM 180 PM Monitor* is a laser scattering aerosol meter that has "*the European Equivalence Approval for PM_{10} and $PM_{2.5}$ as well as the US-EPA Approval for $PM_{2.5}$*" [98]. It measures the PM_{10}, $PM_{2.5}$ and PM_1 levels at a maximum frequency of ten samples per minute. Usually, the state is not interested in such a high temporal resolution, so that only 15 or 30 minute averages are recorded. Their main aim is to be able to release timely readings of the 24 h-means before the gravimetric measurements are analyzed in the lab. The gravimetric readings in the station are gathered by a pair of *Leckel SEQ47/50* High Volume Sampler (HVS) [220], one for PM_{10} and one for $PM_{2.5}$ measurements. It takes between one and three weeks before the data from the gravimetric measurements is available, since the filters are periodically collected and weighed in the lab. The resulting data is then also used to perform a backwards correction of the time series data from the *EDM180*, since experience has shown that even the certified optical measurements show a deviation of within $\pm 10\%$ accuracy. However, since we expressed our interest in data with a higher temporal resolution, the state supplied us with 1 min-averages of the sampled values.

The $PM_{2.5}$-concentration over the seven days is shown in Figure 18, PM_{10} in Figure 19. The devices show the same phenomenon as in the indoor experiment. The *GP2Y1010* is able to detect the changes of the dust concentrations but the values have a constant offset to reference values. Additionally, some inaccuracies regarding the scale factor are also visible. The transfer of the indoor calibration coefficients to the outdoor scenario did not work as smoothly as we had hoped. One explanation for the observed deviation could be that the temperature outside was as low as −5 °C, much lower than we went when characterizing our sensors' temperature dependency. We assume that the simple linear correction we used is inadequate at "more extreme" temperatures.

Aside from the continuous measurements, we also looked at 24 h-means of each device, as this is the quantity that is currently relevant for regulatory purposes. The results are shown in Figure 20 and Figure 21 respectively. We can see that on-the-fly-calibration achieves improvement in the outdoor scenario as well. This is especially true for the 24 h-means which can be brought down to a very small error, both individually or when averaging over multiple devices.



(a)



(b)

Figure 18.: Outdoor $PM_{2.5}$ evaluation: (a) calibrated *GP2Y1010* sensors against the *DRX8533* reference, (b) values after on-the-fly correction.

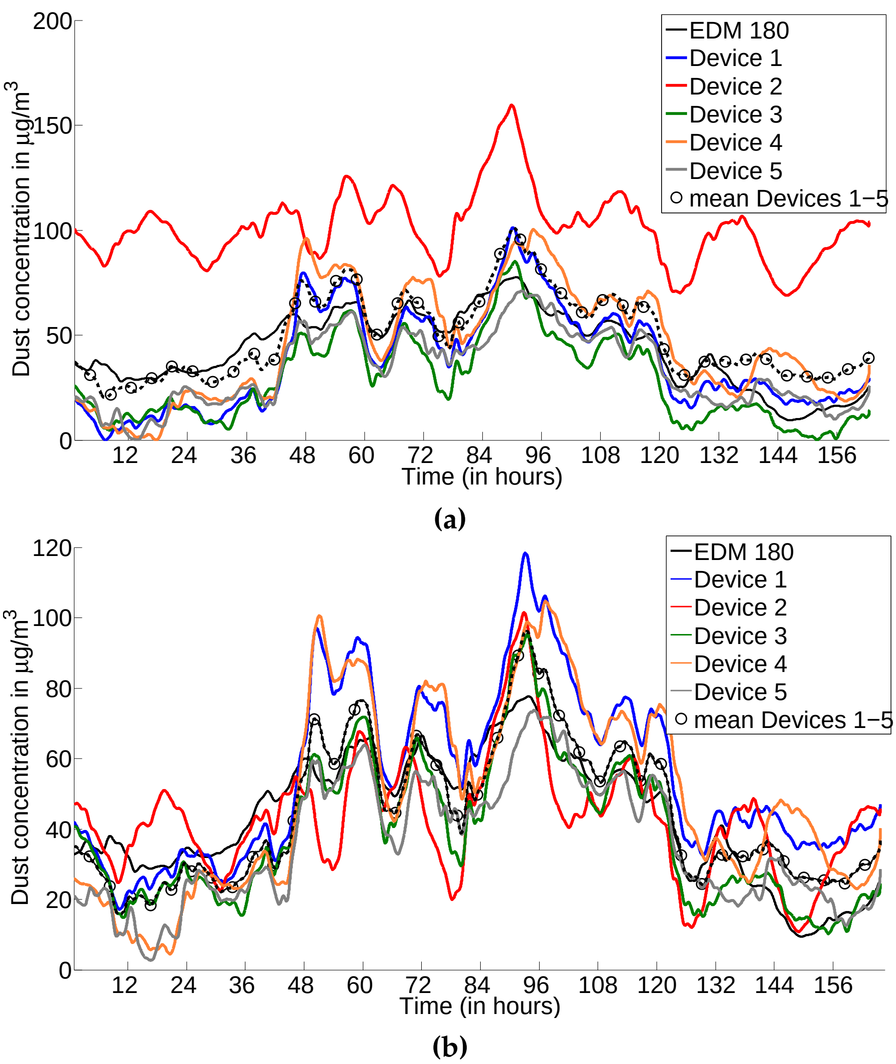


Figure 19.: Outdoor PM_{10} evaluation: (a) calibrated *GP2Y1010* sensors against the *DRX8533* reference, (b) values after on-the-fly correction.

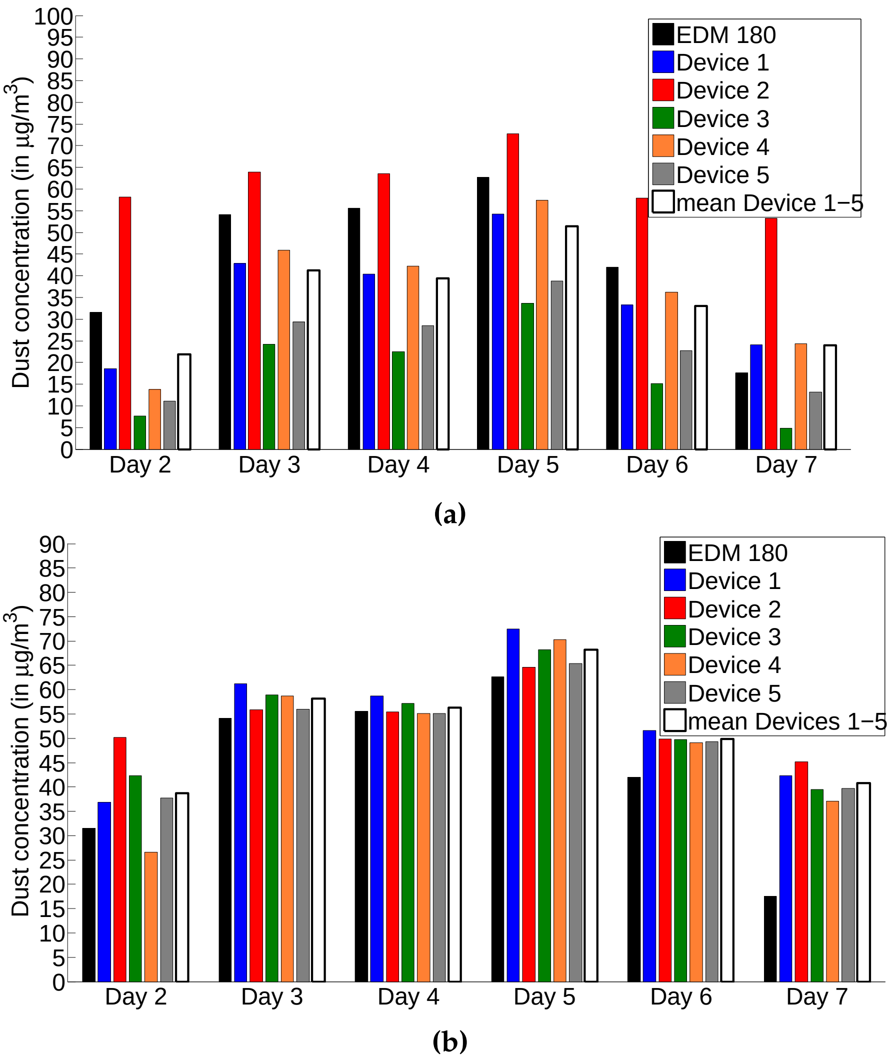
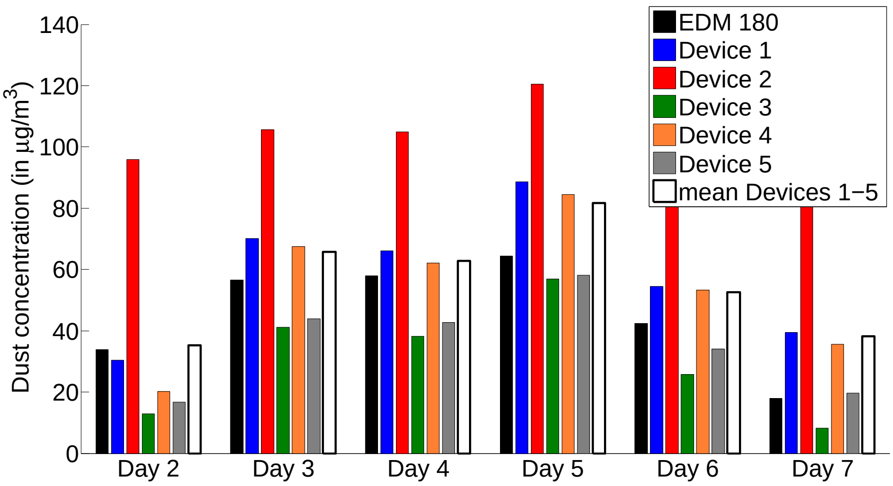
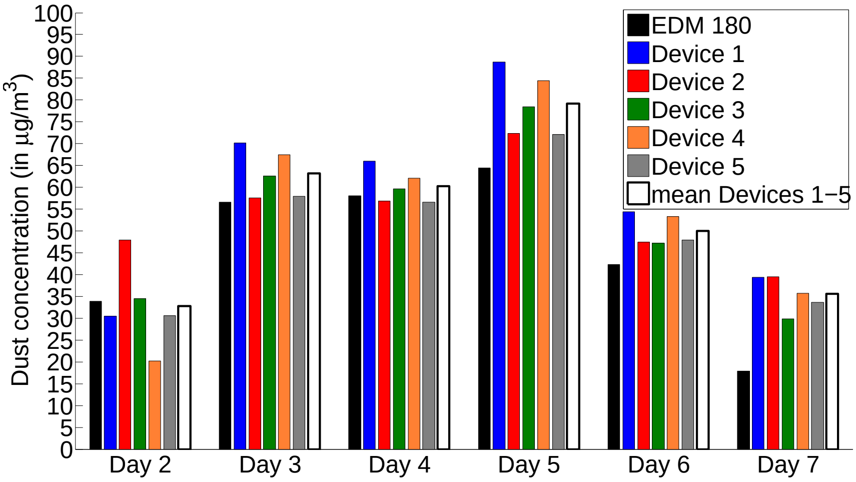


Figure 20.: Comparison of 24 h-means for $PM_{2.5}$ before (a) and after (b) applying the on-the-fly calibration.



(a)



(b)

Figure 21.: Comparison of 24 h-means for PM_{10} before (a) and after (b) applying the on-the-fly calibration.

3.5 CONCLUSION

In this chapter, we have presented past research on the use of low-cost PM sensing technology in Participatory Sensing scenarios. We investigated a cheap commercial off-the-shelf (COTS) dust sensor, the *Sharp GP2Y1010* in terms of its accuracy and presented several calibration, processing and sensor-fusion steps, that lead to meaningful readings from the sensor. We showed, that in a Participatory Sensing scenario, devices equipped with the sensor can use information from co-located devices in order to stabilize and improve their readings. We conducted a series of experiments to juxtapose the performance of a gauged high-accuracy measurement device and the *GP2Y1010*, that show good performance in lab situations and practically relevant results in a realistic setting.

4 Novel Sensing

In this chapter, this work discusses ways of measuring Particulate Matter (PM) *directly* with mobile phones. So, in contrast to the approach described in chapter 3, no additional standalone sensor or dedicated device that needs to be paired to the phone should be involved. To this end, a method of retrofitting a sensor to a camera phone without the need for electrical modifications is presented, in which the flash and camera of the phone are used as light source and receptor of an optical dust sensor respectively. Several design iterations are presented along with two different algorithmic approaches to process the recorded camera images.

Parts of this chapter have been previously published. The proof-of-concept version of the camera phone-based dust sensing approach has been implemented by Pierre Barbera for his bachelor's thesis [12] and the concept has been published in the proceedings of the *International Symposium on Wearable Computers (ISWC)* [30]. Some of the preliminary considerations for the custom 3D printed sensor design were presented on the *International Conference on Atmospheric Dust (DUST)* [40]. The Poisson Particle Detection (PPD) algorithm was developed in cooperation with Marcel Köpke and presented on the *ISWC* as well [39].

4.1 Related Work

As related work on handheld fine dust monitoring has already been covered in section 3.2, this section only discusses work related to internal smartphone sensors or clip-on approaches that enable direct measurement with phones. Related work on the algorithmic approaches is presented below in section 4.5.1.

A system to measure Black Carbon (BC) with cellphones was presented by Ramanathan et al. [191]. The *Aethalometry*-based approach involves BC aerosol collection on a quartz filter, the coloration of which is than captured by the phone's camera and transmitted to an analytics component for real-time evaluation.

Another interesting approach to measuring atmospheric dust in Participatory Sensing scenarios has been presented by the *Air Visibility Monitoring* [186] respectively the *iSPEX* [217] projects. In the former, people use their camera phones to take pictures of the sky and upload them to a central database. There, from the image luminance, the location and phone sensor data (e.g. orientation), the particle pollution is estimated. Cloudy skies and indoor environments present clear limitations to this approach. The *iSPEX* system makes use of a passive spectropolarimetric clip-on module for the iPhone. It has been successfully used in single-day measurement campaigns on a Participatory Sensing scale.

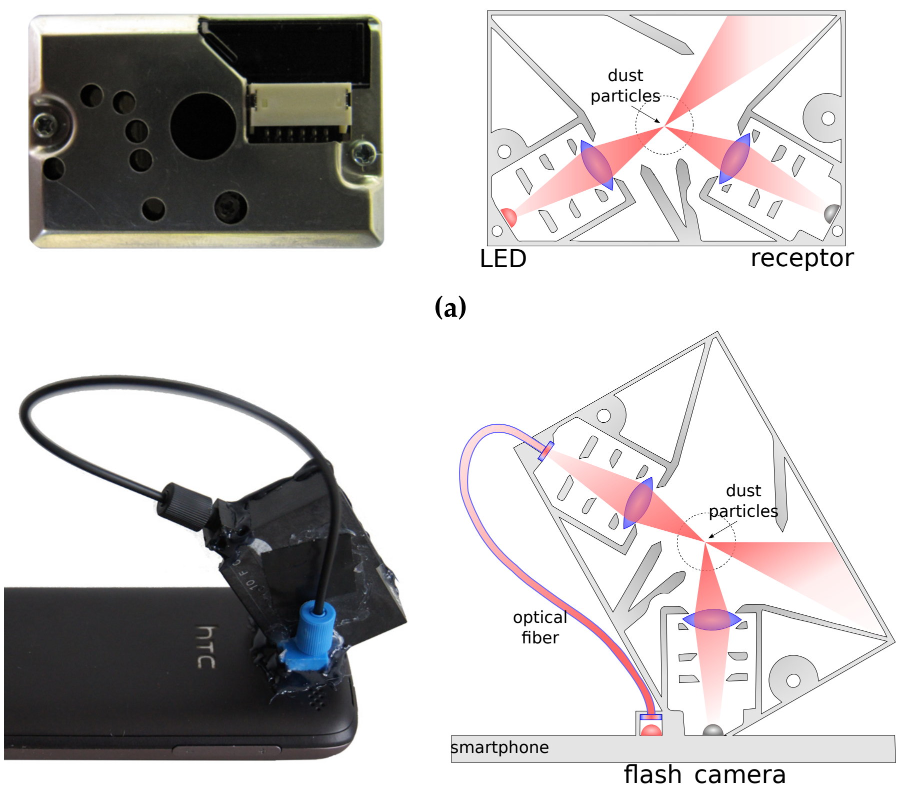
Other efforts to enable particulate matter sensing with smartphones have been made in the past, among them approaches that aim at developing small sensors that can actually be integrated into the casing of a smartphone: Carminati et al. presented the design of a capacitive particle sensor that has the potential to be micro-fabricated and embedded into phones [53]. In a different approach, Doering et al. enabled direct measurement of the mass concentration of particles with an air-microfluidic Micro Electrical Mechanical Systems (MEMS) design [80]. These developments both show interesting approaches as well as promising performance. However, there may be one issue with the general approach of miniaturizing sensors that far. This is not a question of detection principle rather than purely one of statistics: The smaller the detector volume, i.e. the amount of air that can be sampled at a time, the more consecutive measurements need to be taken in order to make a statistically reliable statement concerning the mean concentration. This has implications for air flow and measurement frequency requirements. We discuss this a bit deeper in section 4.3.1

Our own idea of enabling direct smartphone-based PM measurement is to use the flash and camera of a smartphone as active components of a clip-on light-scattering fine dust sensor (see Figure 22). Other work has shown that a phone's camera and flash can be leveraged beyond taking photos, e.g. to measure physiological parameters [205], such as the heart rate [179] or Arterial Blood Oxygenation ($SpO_2\%$) levels [134], or for fluorescence-based measurements with disposable optical sensor chips [174].

4.2 PROOF OF CONCEPT

As described in the previous chapter, experiments with the *Sharp GP2Y1010* dust sensor had shown that meaningful PM measurements can be made with cheap Commercial-of-the-shelf (COTS) dust sensors. In order to remove the need for external platforms that communicate their measurements to a host, e.g. via Universal Serial Bus (USB) or Bluetooth (BT) and thereby reduce the added cost arising for an MCU platform as well as the additional maintenance effort being imposed on the user by needing to carry, charge and operate a second device. Our alternative is to retrofit camera smartphones with an exchangeable dust sensor, e.g. attached to or embedded into the back cover (see Figure 22).

This way, the sensor is very easy to install and can e.g. be exchanged again for a regular back shell when it is not needed.



(a)

(b)

Figure 22.: (a) *Sharp GP2Y1010* dust sensor and operation principle, and (b) prototypical implementation with modified emitter-receptor configuration embedded in the back cover of an otherwise unaltered phone.

For our modifications, we used an unaltered *HTC Desire* smartphone running *Android 2.3.3*. We removed the photodiode and LED from a *Sharp GP2Y1010* dust sensor and attached the disassembled sensor's light trap and lenses to the back cover of the smartphone. This was done so that the phone's camera replaced the original photodiode. A short piece of optical fiber was then used to convey the light from the phone's LED flash to the correct position of the light trap, replacing the sensor's original LED (Figure 22 b). Both components were optically isolated from each other. A simple application was used to record a series of still images or video frames for the duration of the measurement. We compared our prototype against a gauged reference device, a *TSI DustTrak DRX 8533 Aerosol Monitor (DRX 8533)* aerosol monitor. The captured images were analyzed and as transfer function, the overall light intensity was summed up as a measure for the concentration level (see Figure 23). Another way to look at it is, that first, a grayscale histogram of the image is built and then the amount of pixels n_i for each of the 256 intensity levels i is counted. The output is the weighted sum $\Sigma_{i=0}^{255} = i \cdot n_i$, thus giving each pixel a weight according to its intensity.

```
input  : an RGB bitmap img of size w × h
output : a single brightness value corresponding to the dust
         concentration

brightness ← 0;
for i ← 0 to h − 1 do
    for j ← 0 to w − 1 do
        pixel ← (img[i, j][0] + img[i, j][1] + img[i, j][2])/3;   // RGB to gray

        brightness ← brightness + pixel;                          // sum up
    end
    brightness ← brightness/(w × h);               // normalize to [0,255]
end
return brightness;
```

Figure 23.: Naïve Brightness Algorithm

Figure 24 and Figure 25 show the results of our initial tests. The readings of our retrofit PM sensor corresponded very well to those of our reference device, albeit in this first prototype only down to concentration levels of ~10 mg m^{-3}. While this was enough to detect smoke or large concentrations of coarser dust, it was not yet sufficient for the detection of typical fine dust concentrations.

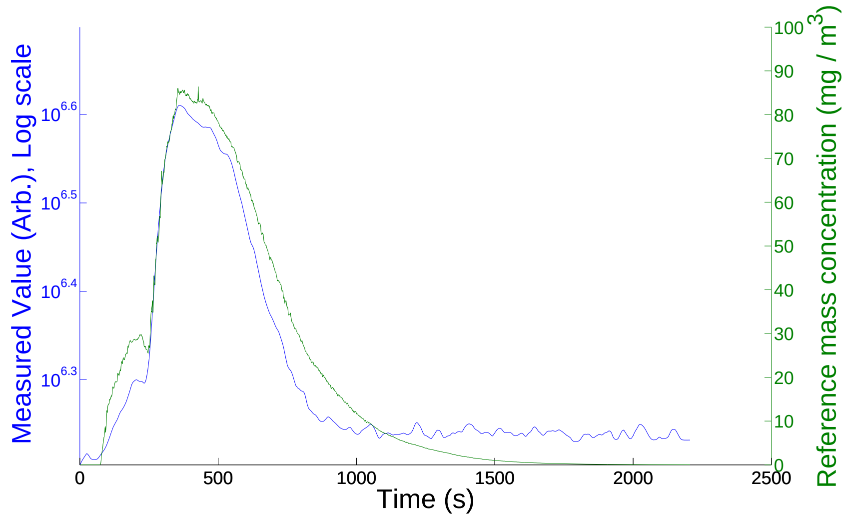


Figure 24.: Example particulate matter concentration measured with the proof-of-concept prototype, producing accurate readings down to ~10 mg m^{-3}.

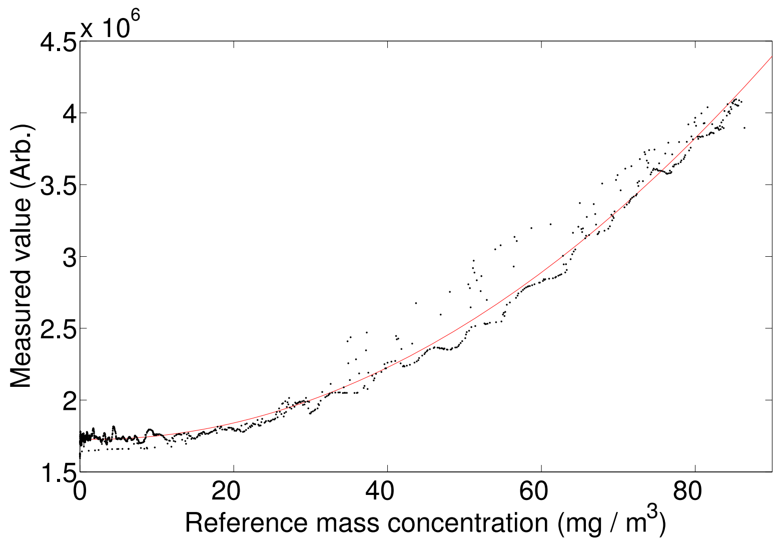


Figure 25.: Output values of our prototypical smartphone dosimeter vs. those of the reference device *TSI DustTrak DRX 8533*.

Still, these preliminary experiments have demonstrated the feasibility of the approach of retrofitting camera phones with dust sensing capabilities. The presented design was tested and shows good initial qualitative performance. However, the evaluation also showed that the sensitivity of the sensor needed to be improved to detect much lower concentrations. Experiments with active version, in which the optical fiber was replaced by an LED for testing purposes, revealed that the passive sensor design of the proof-of-concept prototype suffered from insufficient light intensity levels in the measurement chamber, probably due to the simple coupling between flash and optical fiber. Other aspects that needed improvement were reducing the sensor's form factor to make it less bulky.

In order to address these issues, the next section discusses design constraints, limitations and the benefits and drawbacks of different design variants.

4.3 DESIGN CONSIDERATIONS

This section first presents some theoretical considerations regarding the field of tension between different design parameters of the sensor. This includes lower bounds concerning sensible choices for the dimensions of the hardware add-on as well as a discussion of active vs. passive versions of the sensor.

4.3.1 *Estimations and theoretical limitations*

For the following considerations, we assume that the detector is operated in an environment with a constant particle concentration during the measurement time frame T. Of course, one can imagine fast intense events perturbing the particle concentration. However, in the majority of scenarios, mean concentrations of fine dust can be assumed to be constant over intervals in the range of minutes. For near-real-time particulate matter sensing this would be an acceptable temporal resolution. We argue that the few cases in which these measurement constraints are not met are more or less negligible since our design aims at distributed sensing scenarios with many dense and possibly redundant individual measurements.

As particles move through space due to diffusion and natural convection, they also travel through the detector.[1] Equivalently to the environment moving through the detector we can imagine the detector moving through the environment, with each picture taken by the camera representing a different section of the environment.

Particulate matter is a discrete measurement variable, meaning that dust is composed of discrete particles, entering the measuring chamber at discrete points in time. The pure act of counting these signals within a given volume is a stochastic

1 The particle exchange in the detector can be actively performed using a fan attached to the detector, as also discussed in the next section.

process. In a given picture i, the number of particles counted in that picture is defined as n_i. The mean counting rate $\overline{n}(t_i)$ at the time t_i when the i-th picture frame is shot, is defined as the mean of n_i over the A last frames shot:

$$\overline{n}(t_i) = \frac{1}{A} \sum_{j=i-(A-1)}^{i} n_j \tag{1}$$

Such processes obey a Poisson probability distribution if the expected mean counting rate $\overline{n}(t_i)$ is constant over time.[2]
For such Poisson processes the mean number $\overline{n}$ of measured particles in A last taken picture frames is correlated to the standard deviation σ_n of the measured signal:

$$\sigma_n = \sqrt{\overline{n}} \tag{2}$$

This gives constructive limitations to the measurement device if a certain amount of statistical error is not to be exceeded. An indicator for the theoretically achievable precision and repeatability is given by the Relative Standard Deviation (RSD) or Coefficient of Variation (CV), denoted c_v:

$$c_v = \frac{\sigma_n}{\overline{n}} = \frac{1}{\overline{n}} \tag{3}$$

With a bigger $\overline{n}$ the coefficient of variation improves. So in general, one wants to capture and detect as many particles as possible within a single measurement.

For a certain mean concentration ρ_i, a detector volume V, a mean particle diameter D, and a mean particle density ρ, the theoretical coefficient of variation (respectively the amount of error) for a single measurement is:

$$c_v = \frac{1}{\sqrt{\overline{n}}} = \sqrt{\frac{\rho \cdot \frac{4}{3} \cdot \pi \cdot \left(\frac{D}{2}\right)^3}{\rho_i \cdot V}} \tag{4}$$

As an example, for a mean concentration of $\rho_i = 10.0\,\mu\text{g}\,\text{m}^{-3}$ (typical background concentration in industrialized countries), a measurement volume of $V = 1000\,\text{mm}^3$ (10 mm edge length), a mean particle diameter of $D = 10\,\mu\text{m}$ and a particle density of $\rho = 2.5\,\text{g}\,\text{cm}^{-3}$ (which is in the range of common particulate matter solids), the relative statistical error would be $E \approx 1140\,\%$.

2 The counting rate can be assumed to be constant, as we are in an environment of constant particle concentration.

Thus a single detector of such characteristic length scale cannot measure such concentrations reliably. However, by performing several independent measurements within the time interval T, in which the counting rate $\overline{n}$ is assumed to be constant, we can treat the individual readings as a single big measurement and reach a higher effective counting rate $\overline{n}_e = f \cdot T \cdot \overline{n}$ with f being the measurement rate and T being the sampling duration. The CV rating improves significantly:

$$c_{v,e} = \frac{1}{\sqrt{f \cdot T \cdot \overline{n}}} = \frac{1}{\sqrt{f \cdot T}} \cdot c_v \tag{5}$$

If we consider the calculation for the example above again, a number of $k \approx f \cdot T = 52349$ independent measurements which e.g. correspond to a time period of $T \approx 8.72$ min and a measurement rate of $f = 100$ Hz have to be performed to lower the c_v

Coefficient of Variation ($c_{v,e}$)	Sampling frequency (f)	Sampling duration (T)
5 %	30 Hz	~29.0 min
5 %	60 Hz	~14.5 min
10 %	30 Hz	~ 7.3 min
10 %	60 Hz	~ 3.6 min
20 %	30 Hz	~ 1.8 min
20 %	60 Hz	~50.0 s

Table 7.: Theoretical estimations for different sampling times and frequencies for the measurement of particles of 10 µm diameter in a 1000 mm^3 detector volume. $c_{v,e}$ denotes the relative statistical error that can be achieved.

rating to 5 %. Table 7 and Figure 26 show different combinations of these parameters for various sampling frequencies (current smartphone cameras easily reach 30 Hz to 60 Hz, some models 120 Hz and more) and measurement duration.

This shows that theoretically meaningful readings with a small sensor of 1 cm edge length can be carried out, given sufficiently high sampling frequencies and measurement intervals.

4.3.2 *Active vs. Passive Design*

As discussed above, two versions of the sensor add-on were prototyped: A passive version that uses only the internal cam-

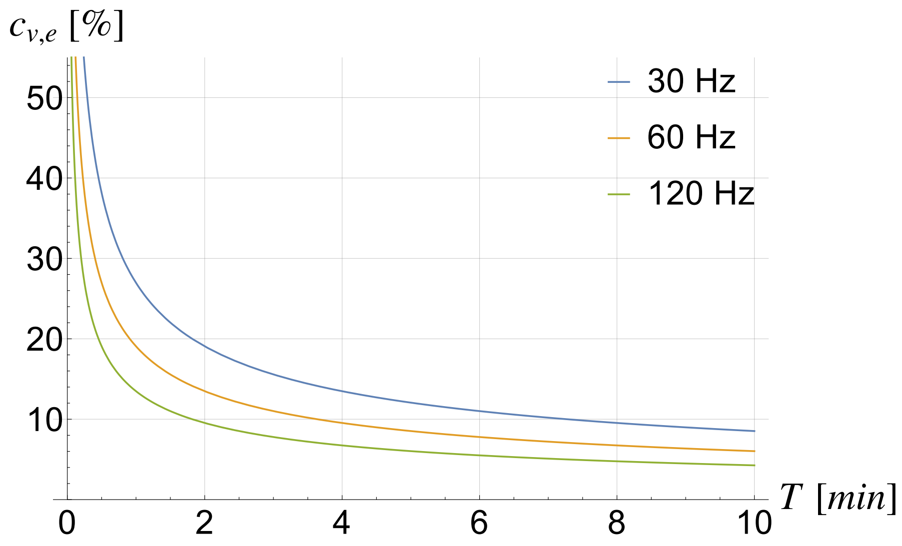


Figure 26.: The relative statistical error $c_{v,e}$ over time for different sampling frequencies.

era and LED flash and an active one with an external LED. The passive version in principle has the benefit that it can be realized at extremely low cost. The active version on the other hand could possibly be realized in a way that it fits to different phones without modification, as it does not need to be adapted to the geometry of camera and flash of a particular phone model. From a design perspective, the active version is more challenging, as the need to power external electronics arises. Another component that may require some sort of energy source is a micro fan.

In order to supply the LED used in our prototype, 20 mA@3.2 V are required in rated operation, which corresponds to a power of $P = U \cdot I = 64$ mW. Adding a micro fan for ventilation would additionally require upwards of 35 mA@2 V = 70 mW. If we were to use standard (non-rechargeable) batteries, options are either tubular batteries or some sort of coin cell(s). With a standard CR2032 coin cell (230 mA h@3 V), we could operate our setup for ~1.7 h, two would bring up to a maximum of ~3.5 h. AAA batteries, which would notably increase the size of the setup, come at up to 1100 mA h@1.2 V, which would mean two batteries would be needed to power the setup for up to ~8.5 h.

Standard Lithium Ion Polymer (LiPo) rechargeable batteries with a higher energy density and suitable dimensions (e.g. from Adafruit[3], 34 × 62 × 5 mm) can provide up to 1200 mA h@3.7 V, providing energy for up to ~9 h of operation. Among energy

3 https://www.adafruit.com/product/258

harvesting approaches (e.g. converting solar, thermal, kinetic energy, etc.), solar panels are currently the only option that can potentially deliver the amount of energy needed for operation of our setup (or, in combination with batteries, to recharge them) while also fitting our size constraints. Small cells[4] (35 × 22 mm) with an efficiency of 22 % can deliver up to ~100 mW at maximum power point. Integrating two or three onto the surface of the sensor could be a realistic option that could potentially remove the requirement to charge the sensor (at the cost of additional parts and engineering complexity).

A different approach would be to power the external components through the phone. One option to do this is pulling power from the mobile phone's audio interface (i.e. microphone jack), as proposed by Kuo et al. [131]. Relevant to the question whether this is feasible or not is primarily the power that can be drawn. In their work, they describe that under optimum laboratory conditions (a perfectly adapted load), they were able to reach a current of 66 mA@250 mV, i.e. 15.8 mW. Even without the loss that converting this voltage to suitable levels, the approach of powering the LED directly through the headphone jack, let alone an additional fan, is not feasible.

Another option to draw power from the smartphone would be *USB On-the-go (OTG)*. This has the benefit of being in principle able to supply sufficient power, theoretically up to 500 mA@5 V. However, according to the USB 2.0 OTG specification [240], devices acting as OTG power source must only provide a minimum current of 8 mA@4.4 V to 5.25 V, anything beyond is allowed based on negotiation, but not guaranteed. Realistically, this means that for handheld portable devices 100 mA @ 5 V are a commonly accepted maximum for external loads [225]. A drawback is that this approach could possibly require introducing additional electronics on the sensor side in order to authenticate to the phone and negotiate the power supply. However, as added bonus, this would also allow communication over the same connection, which would eliminate the need for the user to actively switch on the sensor module. The communication could also be realized in a different manner, e.g. by controlling the sensor via Bluetooth. Again, this would mean additional overhead, such as adding a suitable communication module and by that further increasing the complexity of the module.

In summary, the active design of the sensor is more flexible, but also much more intricate than the passive one. Of the

4 http://www.digikey.com/short/39vczc

options for powering an external sensor module, the microphone jack approach can be discarded as infeasible and using non-rechargeable batteries could lead to an acceptance problem (producing too much waste). The other approaches have their pros and cons, ranging from simple solutions requiring more user interaction and maintenance (rechargeable batteries) to more sophisticated solutions that involve external electronics, making them more expensive in terms of design and unit costs. We argue that the passive solution, provided it can properly be ventilated, is the most elegant one. When designing the active sensor, the approaches using USB OTG or rechargeable LiPo batteries (possibly recharged via small solar panels) are options we intend to explore further.

4.4 hardware design

This section presents the design iterations of the clip-on light-scattering sensor. Two different approaches are presented, based on different optical designs.

4.4.1 *Optical Design*

In order to count particles using a smartphone camera, we used two different optical design approaches:

Brightness-based approach

This simple approach is basically the direct transfer of the light-scattering (*Nephelometry*) approach to a smartphone, as described above in section 4.2. The only difference is that instead of a photodiode, the smartphone camera is the light detector. As the camera is located at the focal point of the lens ($k \approx f$), all rays are collimated on the camera chip. So the camera image is not a visual representation of the inside of the detector volume. Instead, the overall brightness of the camera image is proportional to the scattered light. As with other photometers, the brightness represents a sum signal (count, size) from multiple particles in the measurement volume.

Magnifier-based approach

In order to make use of the information that is potentially available due to the fact that the detector features a 2-dimensional (2D) matrix sensor (camera) instead of a single pixel (photodi-

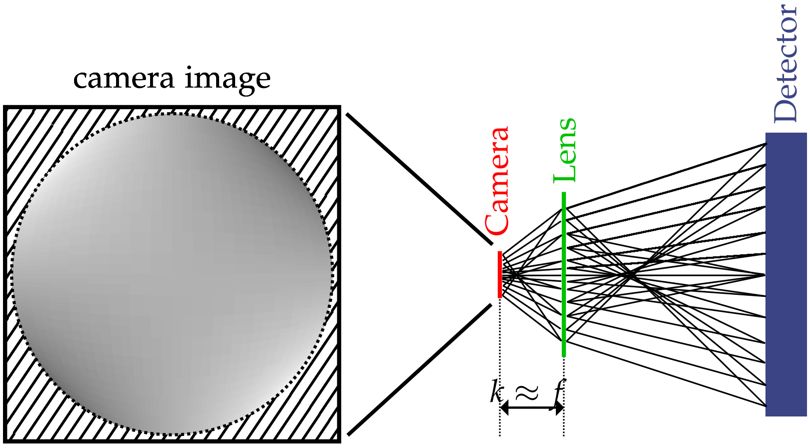


Figure 27.: Sketch of the brightness-based sensor principle: The optical system is a direct translation of the light-scattering principle, the only difference being that instead of a single photodiode, the receptor is the surface of the smartphone's camera.

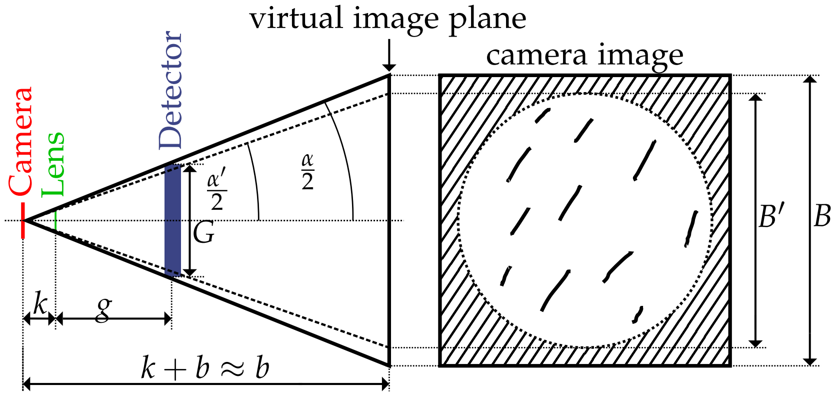
ode), we adjusted the optical system to actually *see*[5] how the particles move. For that, a large magnification is needed. We propose a magnifier approach, which creates a virtual image of the dust particles and only requires the use of a single lens.

If the distance g between the lens and the detector volume is sufficiently small compared to the lenses focal length f, the magnifying lens creates a virtual image of the particles: The camera sees the particles on the virtual image plane. These are magnified by the factor

$$V = \frac{b}{g}. \tag{6}$$

Also, the virtual image is always in focus. As long as the distance k between the camera and the magnifying lens is sufficiently smaller than f, k is irrelevant to the image focus. But k does have an impact on the ratio of the image that the camera sees through the lens: If the lens is close to the camera, it fills the whole camera image. When k increases, the lens only takes up a part of the image and the environment of the lens is visible at the edges. When the total width of the image is B, we call the diameter of the image *through* the lens $B' =: p \cdot B$. If p is to small, the entire detector volume is not visible anymore. We

5 Actually, we just see the scatter traces of the particles, the particles themselves can not be seen with the employed magnification. But since this distinction is not relevant as (for now) our goal is counting them, we will continue to speak of particles.



B : image size	b : image distance
B': eff. image size	k : camera distance
G: detector size	α : aperture angle
g : detector distance	α': effective aperture angle

Figure 28.: Sketch of the magnifier-based sensor principle: This detector design allows virtual images of the scatter traces that are created by individual particles traveling through the measurement volume.

want p to be very close to 1, so we choose k as small as possible.

$$k + b \approx b \qquad \text{and} \qquad k + g \approx g \tag{7}$$

From Figure 28, we get:

$$\frac{G}{g} = \frac{B}{b} \tag{8}$$

And, using Equation 7:

$$\tan\left(\frac{\alpha'}{2}\right) = \frac{B'}{2b} = \frac{B}{2pb} \underset{(8)}{=} \frac{G}{2pg} \tag{9}$$

$$\Updownarrow$$

$$g = \frac{G}{2p \cdot \tan(\frac{\alpha'}{2})} \tag{10}$$

The lens formula for virtual images is:

$$\frac{1}{f} = \frac{1}{g} - \frac{1}{b} \tag{11}$$

If we combine Equation 10 and Equation 11, we get:

$$f = \frac{G \cdot b}{2 \cdot p \cdot b \cdot \tan(\frac{\alpha'}{2}) - G} \tag{12}$$

Now, b can be determined:

$$b = \frac{G \cdot f}{2 \cdot p \cdot f \cdot \tan(\frac{\alpha'}{2}) - G} \tag{13}$$

Summing up, there are three free parameters: The distance between the camera and the magnifying lens k should be chosen as small as possible. In order get an optimal magnification (see Equation 6), the distance g between the magnifying lens and the detector should be as small as possible. On the other hand, construction limitations apply: g must be big enough so that the lens can actually be inserted into the measurement chamber and the lens does not intrude into the actual detector volume. The focal length f of the magnifying lens has a very small impact on the magnification V. However, it must be bigger than k and g, as this was our fundamental assumption.

4.4.2 *Design Iterations*

With the above considerations, we continued the iterative hardware design. This section shows three sensor generations.

1st generation sensor

The original proof-of-concept version of the clip-on light-scattering sensor [30] as already presented above in section 4.2 is shown in Figure 29. That version basically consisted of an original *Sharp GP2Y1010* dust sensor that was attached to the back of a smartphone so that the phone's camera replaced the original photodiode receptor. The light of the phone's LED flash was rerouted to the position of the LED in the original sensor using an optical fiber devkit. This prototype demonstrated the general feasibility of the clip-on light-scattering approach, but did not yet achieve a sensitivity suitable for realistic applications (see section 4.2 above). Active versions using external LEDs were tested for comparison in both of the early generations.

Figure 29.: *HTC Desire* phone with 1^{st} generation (proof-of-concept) prototype built from a *Sharp GP2Y1010* dust sensor and an optical fiber devkit.

2^{nd} generation sensor

Subsequent versions, starting with the second sensor generation (a.k.a. *MobileDust*), were 3D printed for rapid evaluation and featured lenses and mirrors as additional optical elements. We printed the sensors from black polylactide (PLA) or Acrylonitrile Butadiene Styrene (ABS).

The second generation sensor was built on top of a *Nexus 4* camera smartphone. Its light trap basically was a scaled-down version of the one in the *GP2Y1010*. In order to increase the amount of light that is emitted into the measurement chamber, we added a collimator lens to improve the optical coupling of the LED flash's light into the optical fiber (see Figure 30). We used a semi-spherical lens for this, as camera phones' LED flashes are designed to emit diffuse light. Without a collating lens this made the coupling to the optical fiber very ineffective.

When comparing the passive version to an active one that – instead of the optical fiber – featured an externally powered white LED, we observed that the principle was sound, but the passive version still failed to produce a sufficient light intensity within the measurement chamber. As a result, also in this generation the active version clearly outperformed the passive one. Therefore, we switched to a mirror-based layout in the third generation.

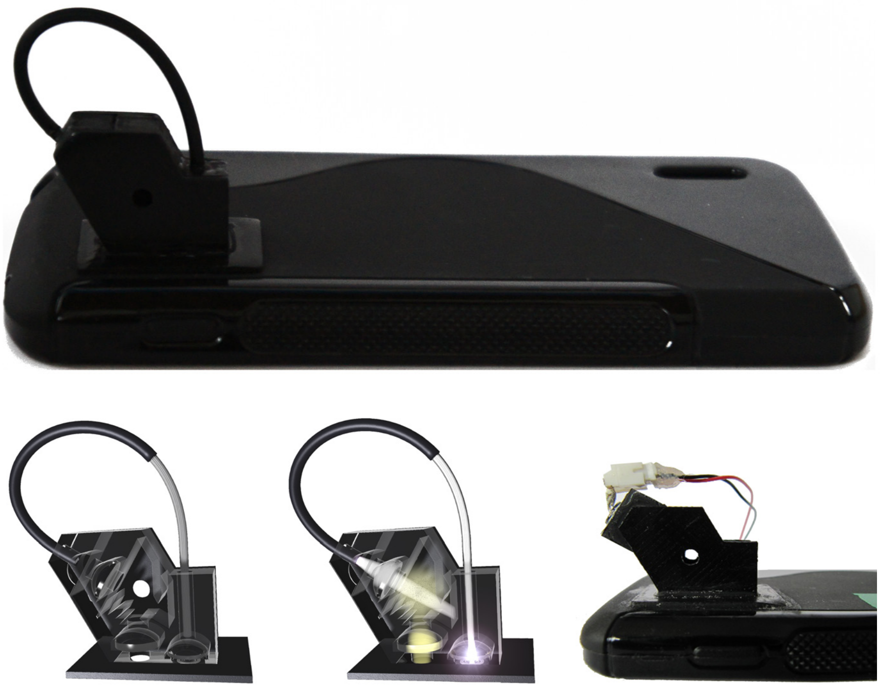


Figure 30.: *Nexus 4* phone with 2nd generation sensor prototype (a.k.a. *MobileDust*). The *GP2Y1010*'s light trap was scaled down and a collimator lens was added to improve the optical coupling between the LED flash and the optical fiber. An active version with an externally powered LED was evaluated in parallel (bottom right).

3rd generation sensor

In the third prototype generation (a.k.a. *FeinPhone*), the optical fiber approach was discontinued. Instead, a mirror was used to illuminate the measurement chamber with the LED flash and a custom sensor casing was prototyped in 3D-printing. At the same time, the employed lenses were changed to use the magnifying optical approach as described above. A prototype was designed for a *Galaxy S6* smartphone (see Figure 31).

Throughout all sensor generations, we kept true to the strategy of designing an active and a passive version in parallel, as there are certain advantages and drawbacks to both designs: The active versions need some sort of power supply for the LED and possibly involve additional interaction (turn on / off) by the user. This could also cause secondary effects, such as possible limitations of the runtime, or negative effects (especially for

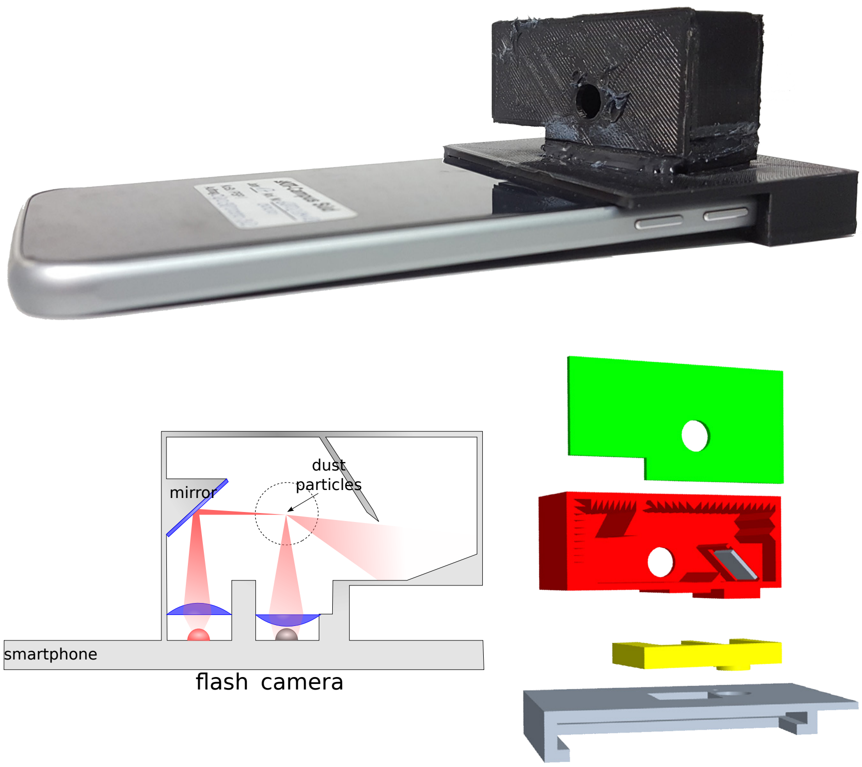


Figure 31.: *Galaxy S6* phone with 3rd generation sensor prototype (a.k.a. *FeinPhone*). Instead of an optical fiber, a custom light trap design with a mirror to relay the light from the LED flash was used.

environmentally conscious users) if consumables, e.g. external batteries, are frequently required. On the other hand, the active version can potentially be attached to a wider range of phones without individualized design.

The biggest advantage of the passive version is simply that: it is passive. This makes it ultra-low-cost and the control of the whole measurement can be implemented in software on any phone. A drawback is that the layout of the camera and the flash is model dependent, so the physical sensor design has to be adopted for different phone types according to their geometry. Proper ventilation of the measurement chamber to ensure that individual measurements are actually independent may also be an issue. An approach to make a fan-less measurement system could be to vent the detector volume by moving the phone and use the phone's inertial sensors to estimate the flow rate through the measurement chamber.

4.5 ALGORITHM DESIGN

In this section, two different algorithms for the evaluation of the images that are recorded with the presented hardware designs are introduced.

As air quality sensing is undergoing a paradigm shift towards the inclusion of low-cost pollution sensors [218], increasingly being set up, operated and maintained by novice end users, a wide range of issues appear. When compared to standardized procedures, e.g. as described in *DIN1319*, most handling requirements to ensure results of high validity are typically not fulfilled:

- Correct and fixed placement of sensors — in mobile and wearable scenarios, sensors are often placed opportunistically,
- Periodic or constant expert calibration of sensor — not feasible in a wearable or mobile device
- Standardized measurement process — difficult to reach due to inexperienced users
- Controlled environment conditions — very difficult due to mobility of the user or movement of worn sensor

Thus, the significance of such measurements is seemingly low and errors restrict the credibility and therewith the use of the gathered information. There are two types of errors in measurement readings: *statistical* error and *systematic* error. *Statistical error* refers to a deviation between multiple measurements of the same phenomenon, e.g. due to sensor noise and/or the statistical nature of the sensing process. This error is often distributed Gaussian between the individual readings and can thus be averaged out.

In contrast, *systematic error* means that any measurement differs from the actual value in the same way, or in other words: The measurement system is de-calibrated. Systematic errors can stem from a number of sources:

- *Low-cost sensors* may be susceptible to systematic cross-sensitivity, e.g. caused by temperature dependencies of electro-chemical sensors, cameras or photodiodes [37]
- *Sensor aging* effects, especially in low-cost environmental sensors, can introduce both drifts over time as well as sudden offsets. Causes for sensor drift may be e.g. dirt deposition within the sensor, while abrupt changes might be caused by degradation, e.g. pixel defects in image sensors.

- *Limited parameter control* may be an issue if an existing system's sensors are re-purposed as environmental sensors. Such systems may not allow deep hardware control, e.g. mobile or wearable sound recording or camera devices such as smartphones and smart watches may not allow full control of the recording parameters. As a result, automatic gain or sensitivity adjustments may occur unsolicited or even unnoticed, potentially causing de-calibration.

- *Novice/untrained users*: In Participatory Sensing, anyone can participate. Especially when measurement requires proper device handling and/or involves an assembly procedure, the changing integrity of the sensing system is another source for errors. For example in a camera-based sensing task, a user could inadvertently put a smudge on the camera lens, which then creates an offset in subsequent readings.

In principle, all of these systematic errors can be quantified and removed from the data. However, it is often not feasible to do this in-situ without recalibration and/or a reference device. The last of the four presented sources of errors is of particular interest, as technology continues to improve and may eventually mitigate some of the other problems. The users' capabilities however do not change fundamentally. While certain amount of training/proper instructions/suitable interfaces will always be required, novice non-expert users – being only human – will also always make mistakes in handling the equipment. Approaches to deal with this on the Human Computer Interaction (HCI) level are discussed in-depth in chapter 6.

In this section, we present a simple signal reconstruction scheme for the monitoring of certain environmental phenomena that is robust against the presented errors by reconstructing the true signal solely from the Poisson noise of the erroneous signal. While the signal itself may be skewed or distorted, its noise is a relative property, rather than an absolute one. The unique aspect about this approach is, that intensive characterization of systematic errors in order to remove them from the signal are not necessary: Our approach is robust against both static and dynamic baseline shifts (offset and drift), as well to a certain extent against cross-sensitivity.

We demonstrate the feasibility of our approach for different phenomena: Aside from the use case of fine dust sensing with camera smartphones, we selected a dataset from low-cost gas sensing[6].

4.5.1 *Related Work*

There is a number of different purposes for which noise is analyzed: One is fingerprinting sensors to uniquely identify them, which has been demonstrated for cameras [58] and accelerometers [75]. In image processing, Poisson noise is usually seen as something undesirable that should be to be removed from an image [22]. However, the noise in question is the pixel noise (also called *shot noise*) of a camera sensor element. In contrast to the *sensor noise* which these approaches deal with, we analyze *sensor data noise*, i.e. the Poisson noise of the signal variation. This is an approach that is also used in Statistical Physics for applications such as analyzing properties of solid state bodies.

Kalman filters can be used for state estimation of noisy signals, e.g. for sensor drift correction. Yet, they typically require a priori knowledge about the noise characteristics, i.e. the models used in the filter need to match the physical situation. While Adaptive Kalman Filters can to a certain extent learn these models, they still require readings from other sensors and/or periodic access to ground truth to somehow rate the quality of the current estimation. An approach to reconstruct signals from noisy data or incomplete information is *Compressed Sensing*. Haupt et al. [106] showed that signal information can be obtained from several random Fourier projections. Furthermore, Model et al. [159] demonstrated the use of multi-sensor arrays to reduce data noise of spatially separated signal sources. A more general approach addresses data that is corrupted by some sort of Poisson noise, in particular image noise [137]. However, all these attempts have the assumption in common, that the noisy data does actually represent the signal and is not altered by some sort of systematic error or drift over time. As motivated above, systematic error can also often be attributed to improper handling by untrained users. Interestingly, while research on mobile, low-cost and participatory sensing recognizes the need to ensure credible readings from cheap sensors [135], the focus is seldomly placed on the effects that are caused by novice users. An intuitive approach is to either train participants or try and determine their skill level or reputation beforehand [230]

6 Dataset from the *OpenSense* project [104], [140] (`http://www.opensense.ethz.ch/`) courtesy of David Hasenfratz.

and/or select them accordingly [193]. However, this again requires some kind of ground truth determined by expert users or a series of campaigns, making it an intricate option. We discuss approaches to instruct or guide the user towards correctly performing environmental measurements in chapter 6. To the best of our knowledge, using signal processing to mitigate certain problems that are likely to occur through improper handling of measurement equipment is a novel approach.

So, while some of the related work can help to deal with certain types of errors, none of the approaches are able to compensate for all of the presented sources of systematic errors without additional knowledge or sensor data. Our approach uses only raw data and is applicable for environmental sensing of any phenomena that can be modeled as Poisson processes.

Aside from the already mentioned cases of measuring particulate matter or gas concentrations, a more exotic example is *RadioActivity*, a measurement app that turns a mobile phone into a Geiger counter [128].

4.5.2 *Poisson Particle Detection (PPD)*

The basic idea of our approach is very simple: Whichever environmental phenomenon we wish to measure — as long as it can be modeled as particles — a way to look at the underlying scenario is as follows: We wish to observe a changing concentration of an environmental parameter in a certain measurement volume. Because we are only observing a small measurement chamber we have a certain chance that e.g. a particle is present during our measurement or that it is not. But, as we are conducting multiple measurements which reflect different sectors of the environment (see section 4.3.1), we are basically looking at a series of counting experiments that are conducted in parallel.[7] This is — generally speaking — the process of counting uniformly distributed events in a spacial volume. This is by definition a *Spatial Poisson Process (SPP)* or *Poisson Point Process (PPP)*. A vivid example for a Poisson process is observing raindrops on the tiles of a rooftop [117].

We want to reconstruct the original signal (e.g. the number of particles in the environment) from the signal we measure. But we can not deduce the original signal directly from our

7 For this, we assume that our measurement rate f is much higher than the change rate of the environmental phenomenon. We show that this assumption holds in our validation.

measurement, because we face two systematic measurement errors:

- A *Multiplicative Error*, which is caused by the mere fact that not every particle in the detector will be seen (*detected*) by the system. We will eliminate this error by the calibration of the sensor.
- An *Additive Error*. In the case of considering the image brightness as a measurement for the number of particles in the detector, this can be caused by de-calibration, the camera's automatic brightness adjustment or dirt deposition inside the detector over time. To eliminate this error, we will conduct a noise analysis as described below.

The number of observed occurrences in a Poisson process fluctuates with a standard deviation of $\sigma_n = \sqrt{\overline{n}}$ around its mean $\overline{n}$, or in other words, there is signal dependent noise. From this noise, we can directly calculate the mean concentration of the signal. This of course only works reasonably well if the signal fluctuations σ_n are greater than the sensor background noise, since the magnitude of the signal is always higher than that of the noise. If the sensor background noise is as high as the signal noise, the sensor is simply bad and signal reconstruction is impossible.

As noise is a relative property, the inferred concentration is

input : raw environmental time series $X = \{X_t : t \in T\}$, window size w_{tf}
output: reconstructed signal $Y = \{Y_t : t \in T\}$

$X_S \leftarrow$ `avg(`X`);` // moving average with Gaussian window
$s(t) \leftarrow$ `spline(`X_S`);` // spline interpolation
$N \leftarrow \{N_t : N_t = X_t - s(t)\}$; // isolate noise
forall $I_t = [t - \frac{w_{tf}}{2}, t + \frac{w_{tf}}{2}] \subset T$ **do**
  $Y_t \leftarrow$ `stdev(`$(N|_{I_t})$`);` // std. dev. of N on interval I_t
end
return Y;

Figure 32.: Poisson Particle Detection (PPD): Signal reconstruction from Poisson Noise. The input is a time series of brightness values. To process the data of the 2$^{\text{nd}}$ generation sensor (*MobileDust*), the Naïve Brightness algorithm (see Figure 23) must be applied to the time series of images in order to get the brightness time series.

unaffected by the additive error. This allows to derive the actual signal values if the noise-to-signal-dependency is known. This statement is generally true, even for non-Poisson processes. In case of a Poisson process however, the dependency is known: The noise will behave like the square-root of the signal. Since we focus on environmental sensing and our data is Poisson distributed, we focus on the special case for brevity.

The algorithm to reconstruct the signal has four steps, as shown in Figure 32. We first apply a simple moving average to the noisy data to reduce it to its mean values. Then, a spline interpolation is constructed on the smoothed data in order to determine the mean value on any point within the measurement series. After that, the actual noise can be extracted by subtracting the mean from the raw data. Finally, to obtain

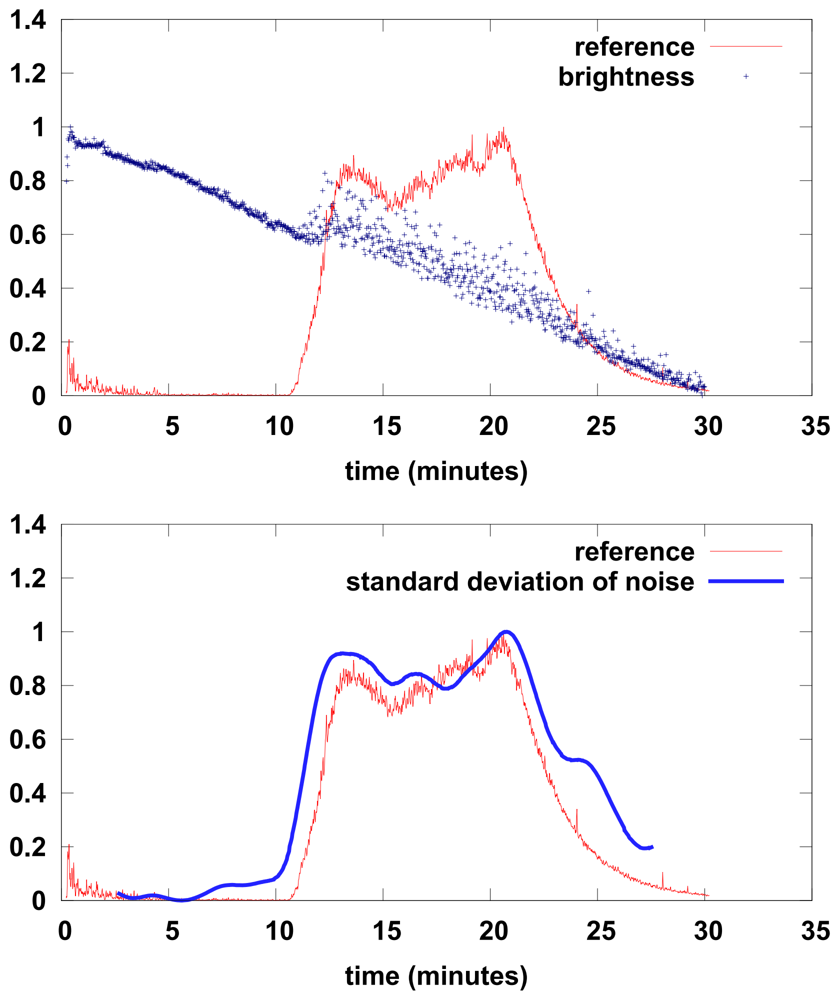


Figure 33.: Brightness signal vs. particle concentration (all normalized) for measurements from the *MobileDust* data-set affected with dynamic baseline shifts.

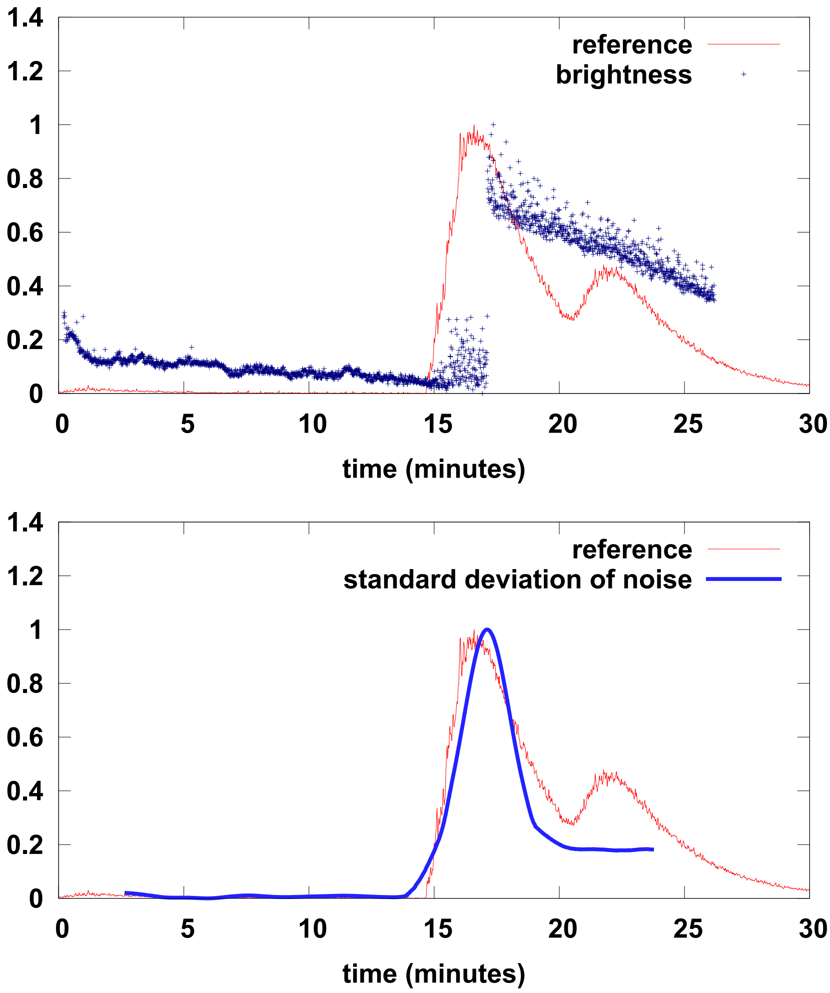


Figure 34.: Brightness signal vs. particle concentration (all normalized) for measurements from the *MobileDust* dataset affected with a sudden jump in offset.

a measure of fluctuation, we calculate the standard deviation of that noise on several time intervals T_i, which is then linearly correlated to the square-root of the signal mean values corrected for drifts. Whenever a summation is done on parts of the data (e.g. averaging) we weigh this sum with a Gaussian window to minimize boundary effects of convolution.

Evaluation

We show the validity of our approach on two separate real-world datasets: To demonstrate the feasibility for different phenomena, we selected data from PM sensing with removable dust sensors on camera-phones [30] as well as low-cost Ozone (O_3) sensing.

The gas dataset was recorded in the *OpenSense* project [104], [140].

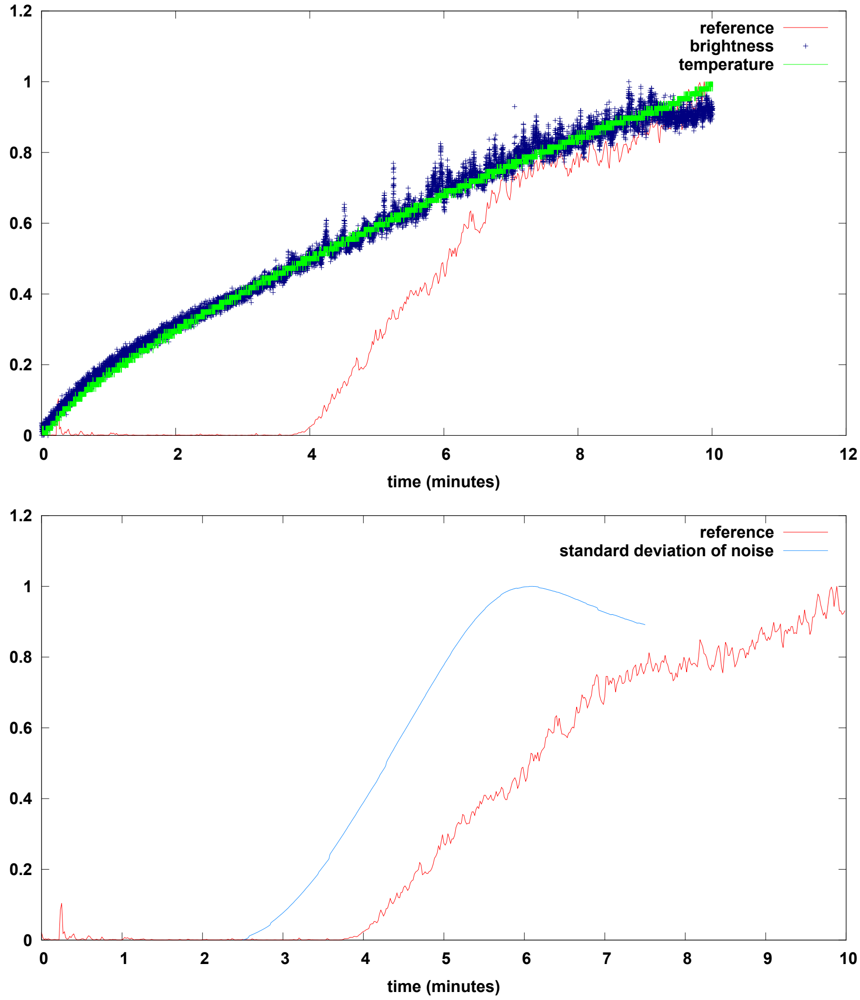


Figure 35.: Brightness signal vs. particle concentration (all normalized) for measurements from the *MobileDust* dataset affected with systematic cross-sensitivity, in this example a strong temperature dependency.

The PM data set was recorded in a lab setting with the second generation (*MobileDust*) prototype (see section 4.4.2 above) attached to a *Google Nexus 4* smartphone. Image resolution was 3264×2448 and sampling frequencies ranged from 0.5 Hz to 1.5 Hz, depending on the settings for the individual measurement run. The phone was placed in a container so that the reference device was exposed to the same air flow as the smartphone. From the images, we calculated the accumulated brightness per picture as feature (as described in the Naïve Brightness Algorithm in Figure 23 at the beginning of this chap-

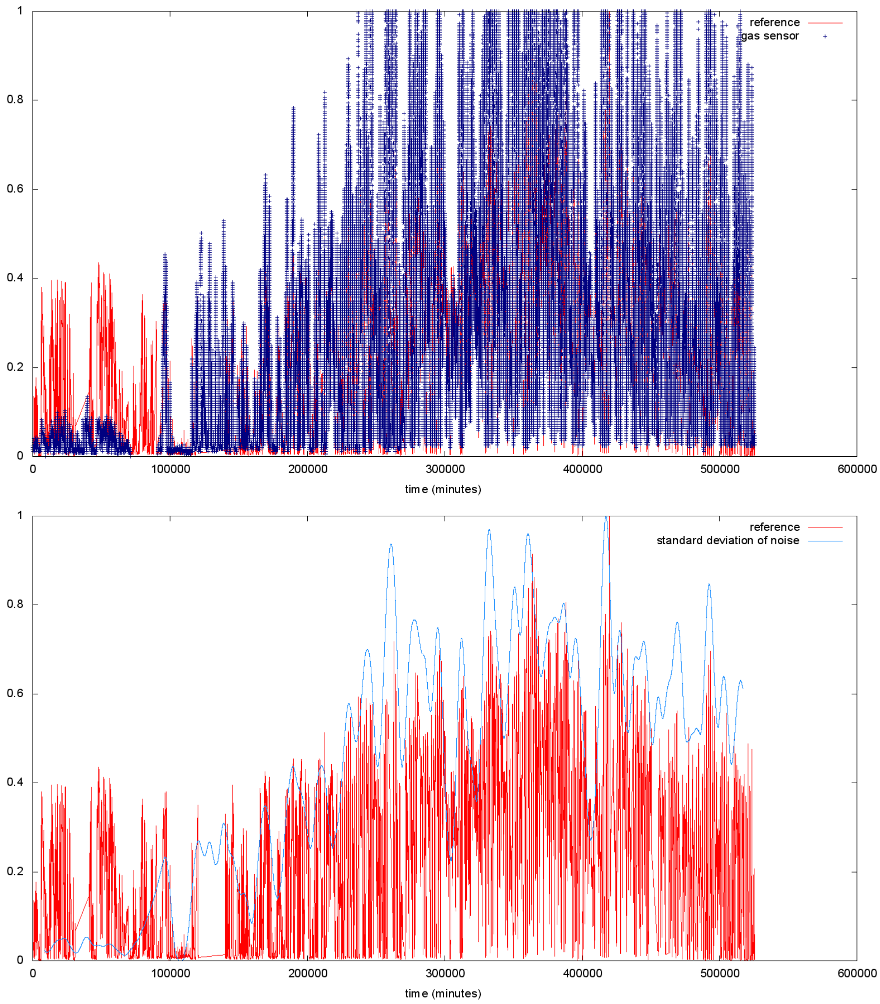


Figure 36.: Raw data and derived signal from the *OpenSense* dataset, showing the successful signal reconstruction from the noise.

ter). The brightness correlates with the amount of particles inside the detector volume. As window size for the algorithm, we empirically selected 130 to 150 samples. Three illustrative examples from the data set are shown in Figure 33, Figure 34 and Figure 35. The plots show that even though the original brightness signal is affected by a dynamic baseline shift, sudden offsets or systematic cross-sensitivity (in this example a strong temperature dependency), the true signal can be reconstructed without additional information.

The *OpenSense* dataset contains one full year of O_3 measurements, recorded with the low-cost *MiCS-OZ-47* sensor (one reading every 60 seconds) as well as by a fixed station from the national air pollution monitoring network (*NABEL*) in Zürich [140] (10-minute mean values). Figure 36 shows an envelope of the original reference signal as signal reconstruction for this dataset.

Windowing

In order to achieve good results as shown before, it is important to choose an appropriate window size w. If w is chosen too small, the values Y_t will still be subject to the random distribution of data values around its mean. If w is chosen too large, the reconstructed signal will be flattened due to this big averaging window. This effect is illustrated in Figure 37.

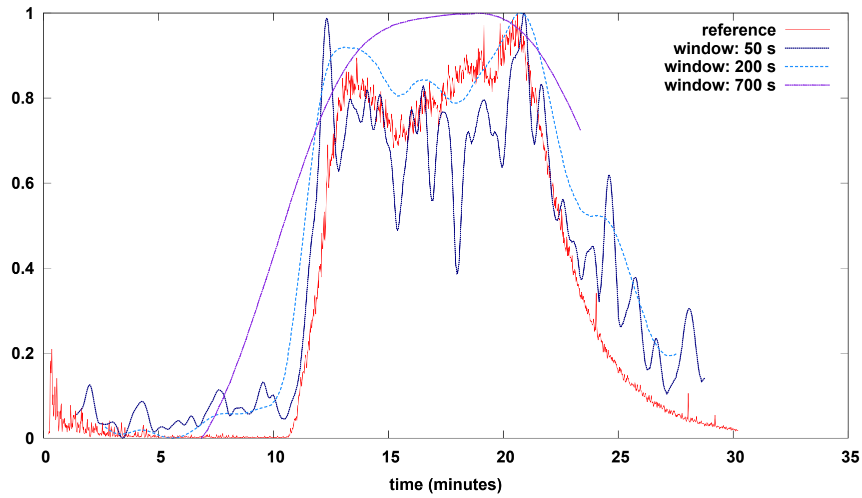


Figure 37.: The window size w needs to be chosen appropriately. For the *MobileDust* dataset, we empirically determined a window size of $130 - 150$ samples to be optimal.

Calibration Stability

In order to evaluate the calibration stability, i.e. whether it is possible to determine a fixed set of parameters to calibrate the noise to map to absolute concentration values, we analyzed a series of twelve different measurements from the *MobileDust* data set.

All data was recorded using the same smartphone (and therefore the same camera), but in independent sensing runs over the course of one year with long pauses between them. Thus, data was subject to possible systematic errors due to varying temperature, slightly different experimental setup and even changed methods of data collection: The first few data sets were

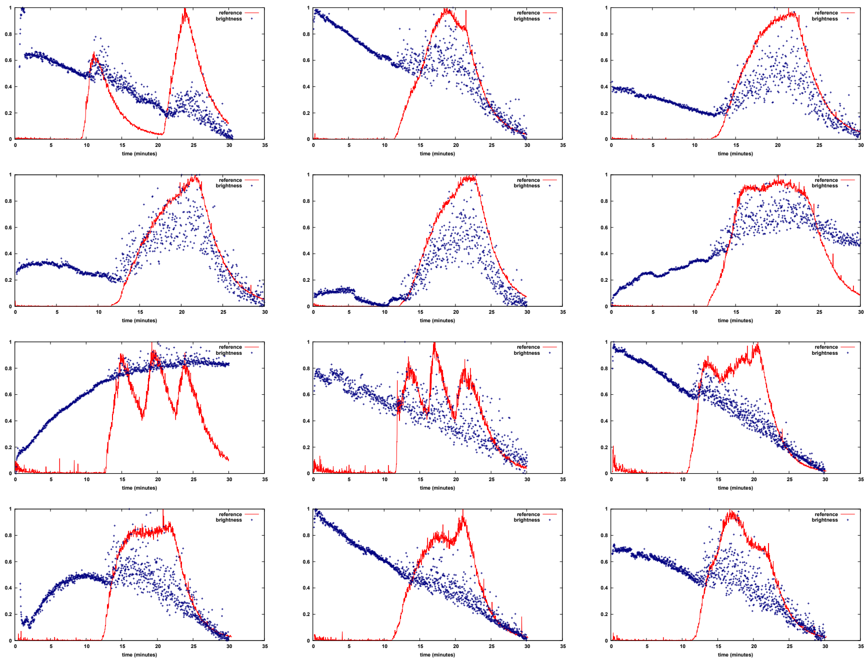


Figure 38.: Calibration stability of the *MobileDust* dataset: Raw data from the twelve measurements was completely different and looks uncorrelated. Still, the Poisson algorithm recognized the concentration changes.

recorded with an older version of the sensing app, which heavily used automatic white balance and exposure compensation. Between sessions, the experimental setup was disassembled and the clip-on sensor fully detached from the phone. There were no restrictions regarding the placement of the dust sensing smartphone inside the measurement chamber. Measurements were taken at diverse daytimes and seasons, with possible influence of ambient light and temperature on the mean brightness of the pictures. The batteries that powered the light source of the dust sensor were not changed or checked throughout the measurement series, therefore resulting in a possible systematic brightness drift. Additionally, it is possible that dirt or imperfect assembly may have had influences on the recorded images.

Figure 38 shows a thumbnail overview of the twelve measurements, each showing a comparison of measured reference and sensor values. There is no obvious correlation visible. Furthermore it seems that the qualitative characteristics differ between some measurement series. Nevertheless, our analysis yielded

	data01	data02	data03	data04	data05	data06
A	19214.2	16515.3	19424.1	16026.0	15847.0	18162.2
ΔA	115.7	57.4	62.5	49.4	48.2	84.5
	data07	data08	data09	data10	data11	data12
A	13451.2	12934.0	15961.0	16424.4	15958.4	17059.5
ΔA	71.3	67.4	40.6	37.9	55.5	57.9

Table 8.: Calibration Parameters derived by LMS fit for the individual measurements. The relative error between parameters is only 11.4%.

that the correlation behaves like the square-root of reference, as expected for a Poisson process. The exact model we used is

$$y = A \cdot \sqrt{x} + B \tag{14}$$

with $B = 800$ due to background noise of the sensor. A Least Mean Squares (LMS) fit of this model yields the parameters as shown in Table 8, ΔA being the statistical error of the fit.

The mean relative statistical error is only $\bar{\Delta A} = 0.38\%$. This is a measure of correctness of the model. Therefore the claim of Poisson distributed data values holds with high probability. Furthermore the standard deviation of the fit parameter A from its mean $\bar{A} = 16414.8$ is $\sigma_A = 1874.4$. So we have relative error of about 11.4% in adjusting the fit parameters between different measurement series, which is a measure for stability of the calibration approach.

Discussion

While the proposed method is clearly able to compensate for systematic errors, there are some constrains regarding its application. If the data set has already been smoothed, e.g. due to some commercial and/or inaccessible sensor setup, the time resolution will decrease. This is because our approach is based on signal noise, and a smoothing process removes that noise. As common in signal analysis, bigger processing intervals will eventually compensate for "bad data" but this will always go at the expense of time resolution and in-time accessibility.

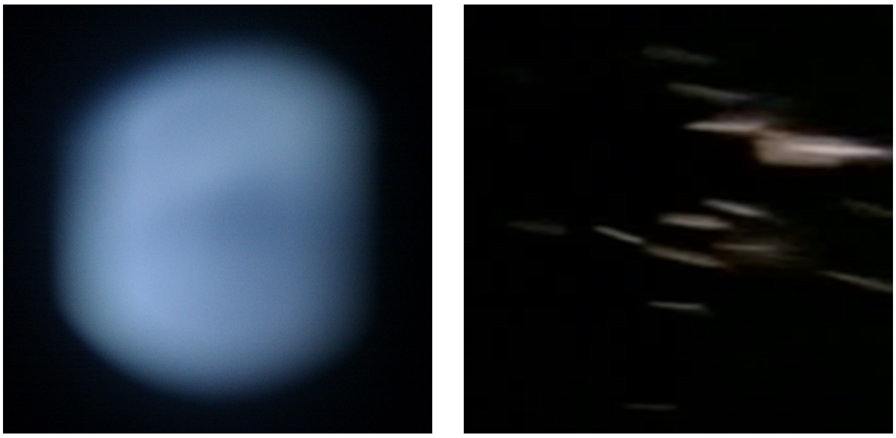
Also, there is a restriction to drift compensation. If systematic errors change the signal too rapidly, that is with high frequency, they cannot be distinguished from signal noise. In a sense, it is then possible to regard the systematic error approximately as pure statistical fluctuations. If these fluctuations overcome the signal noise by means of magnitude, it is difficult to apply the noise extraction, since the noise-to-signal-dependency will be blurred. This is not a practical limitation though, as poor data with huge background noise will always be problematic, no matter which method of analysis is used. Applicability regarding drift from cross-sensitivity may be limited in case that both measured phenomena are Poisson distributed. If a gas sensor is e.g. sensitive to two different gases, our approach may not be able to reconstruct the signal, depending on the individual magnitudes.

As already stated before, there is one intrinsic disadvantage to our approach: By measuring the signal noise, we deliberately sacrifice part of the Signal-to-Noise Ratio (SNR), because signal noise is by magnitude always smaller than the signal itself. Our method thus is a trade-off between SNR and stability. On the one hand, it will never yield as good results as proper characterization and subsequent removal of the systematic error, but on the other hand, no additional information is needed in order to account for de-calibration and drift of (almost) any kind.

4.5.3 *Contour Detection Particle Counting (CDPC)*

Starting with the *FeinPhone* sensor (third generation prototype, see section 4.4.2 above), the optical layout was changed so that the recorded images no longer display one large blob, but instead capture individual scatter traces from particles, as shown in Figure 39. The rationale behind this approach is the following: Firstly, it is a more direct way of assessing the particle count, compared to deriving it from the signal of a single sensor. Secondly, it concentrates the scattered light from individual particles to a smaller area, so that the sensitivity of the sensor may be improved. And lastly, it has the potential to reveal further information, for instance on the size spectrum of the detected particles.

In order to detect individual particles from the *FeinPhone* sensor data, the *Open Source Computer Vision (OpenCV)* library [26] was used. Since the images from our custom *FeinPhone* sensors all were affected by background illumination due to imperfect light trap design and the manual assembly of the sensors, we combined several standard *OpenCV* algorithms to



(a) (b)

Figure 39.: Changed signal between 2^{nd} and 3^{rd} sensor generation: scatter blob (a) vs. scatter traces (b), for illustrative purposes both at high particle concentrations.

isolate the particle traces. The complete Contour Detection Particle Counting (CDPC) algorithm is shown in Figure 40.

As preprocessing step, the first five seconds of the video were cut off, since the first frames in the beginning may be affected with noise due to automatic camera adjustments, such as setting the configured focus length, etc. Subsequently, we converted the RGB video to grayscale. The first major step is processing the video with a background subtraction algorithm. We used one of the standard background/foreground segmentation methods available in *OpenCV*: The *MOG2* subtractor is based on Gaussian Mixture Models (GMM) [262], [263] in order to determine the static background pixels. It features shadow detection, which was disabled since we did not need it.

The so-called learning rate $r_{learn} \in [0, 1]$ is a parameter of the *MOG2* algorithm that specifies how fast the background model is learned, respectively updated. A value for of 0 means that the background model is not updated at all, assigning a rate of $r_{learn} = 1$ means that every frame, the background model is completely reinitialized from the previous frame. Since we have an almost static background, we only varied the learning rates close to 0. The results are discussed in the next section.

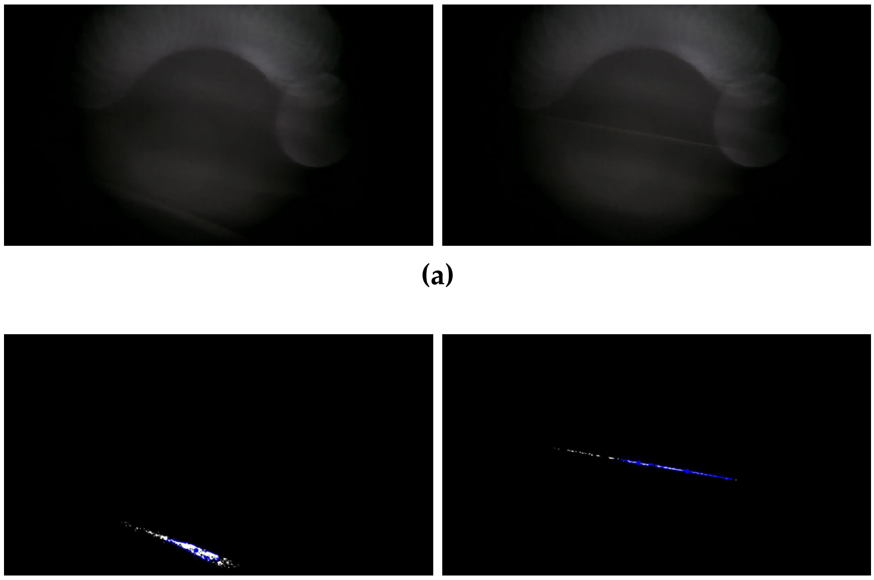
Subsequent to the background subtraction, another second was cut from the beginning of the video, as the first few frames may still display some artifacts until the background model is learned. The next step of the algorithm applies a Gaussian blur filter. This is a common step in image processing before contour or edge detection, as it reduces the noise in the image, thus

input : time series of RBG bitmap images (video) $vid = \{img_t : t \in T\}$, of size $w \times h$ and framerate r_{fps},
MOG2 learning rate r_{learn},
number of standard deviations for Gaussian blur $\langle \sigma_{blur} \rangle$,
threshold for binarization $\theta_{b/w}$,
threshold for contour detection $\theta_{contour}$

output: time series of extracted counts corresponding to the dust concentration $Y = \{Y_t : t \in T\}$

$vid \leftarrow \{img_t \in vid | t > 5 \cdot r_{fps}\}$; `// cut off 5 sec`
forall $img_t \in vid$ **do**
  $img_t \leftarrow$ `cv.cvtColor(`img_t, *COLOR_BGR2GRAY*`);` `// RGB to gray`
  `cv.BackgroundSubtractorMOG2.apply(`img_t, r_{learn}`);` `// bg`
end
$vid \leftarrow \{img_t \in vid | t > r_{fps}\}$; `// cut off 1 sec`
forall $img_t \in vid$ **do**
  `cv.GaussianBlur(`$img_t, \langle \sigma_{blur} \rangle$`);` `// blur`
  `cv.Threshold(`$img_t, \theta_{b/w}, 255,$ *THRESH_BINARY*`);` `// binarize`
  $contours_t \leftarrow$ `cv.findContours(`img_t`);` `// contour detection`
  forall $c_i \in contours_t$ **do**
    if `size(`c_i`)` $\geqslant \theta_{contour}$; `// threshold and count`
    then $Y_t \leftarrow Y_t + 1$;
  end
end
return Y;

Figure 40.: Contour Detection Particle Counting (CDPC) algorithm based on *OpenCV*.



(a)

(b)

Figure 41.: Contour Detection Particle Counting: The original recordings (a) undergo background subtraction, blur and binarization before a contour detection algorithm isolates continuous patches, of which all with an area exceeding a preset threshold are counted (b).

preventing the contour detection from falsely identifying noise as edges.

The next image processing step is to detect contours in the blurred images. For this, the `findContours()` function of *OpenCV* was used. It was configured to use `RETR_EXTERNAL` as contour retrieval mode, meaning that it did not return hierarchical contours, i.e. contours within other contours.

Finally all contours are counted in each frame if their area exceeds a certain threshold. This parameter can be used to tweak the algorithm in order to differentiate between actual particles and possible miscounting due to remaining noise. Figure 41 shows two examples of the original images from our evaluation (a) and the respective images after application of the CDPC algorithm (b).

Evaluation

In order to evaluate the smartphone-based PM sensing approach, we compared several *FeinPhone* prototypes with reference measurements in a lab setting at the *World Calibration Center for Aerosol Physics (WCCAP or GAW-WCCAP)*. The WCCAP is a facility at the Leibniz Institute for Tropospheric Research e.V.

(TROPOS) that is operated in cooperation with the Umweltbundesamt (Federal Environmental Agency of Germany) (UBA) and the World Meteorological Organization (WMO). It is base-

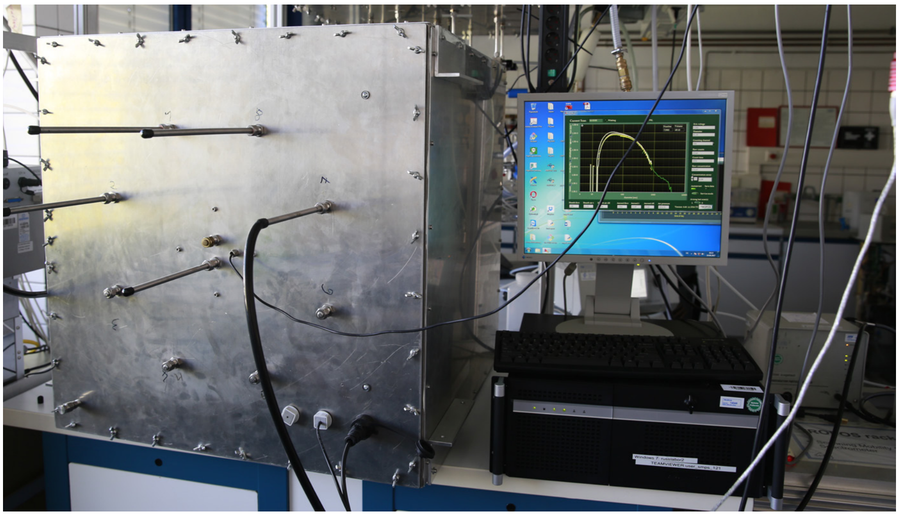


Figure 42.: Evaluation Setup: 5 *FeinPhone* prototypes were placed inside an aluminum measurement chamber (left), into which a varying concentration of polydisperse particles was injected. An SMPS and a APS 3321 (right) were used for reference measurements.

funded by the UBA and conducts calibrations of physical aerosol measurement instruments as well as environmental and work place measurements of aerosols.

The experiment setup is depicted in Figure 42. Through the air inlet of an otherwise airtight aluminum chamber, into which 5 *FeinPhone* prototypes were placed, Ammonium sulfate ($(NH_4)_2SO_4$) was injected in order to create a rising concentration of polydisperse particles inside the chamber. The air outlet was connected to two reference devices: A *TSI Aerodynamic Particle Sizer (APS) Spectrometer Model 3321 (APS 3321)* [231] and an Scanning Mobility Particle Sizer (SMPS) that was custom made at the WCCAP [246]. The combination of these two devices enabled the measurement of 92 aerodynamic size channels between 10 nm and 20 µm. Time resolution was one reading every 4.5 minutes. The SMPS samples the different channels in a time multiplex fashion, i.e. the channels are sampled one after another over the course of 4.5 minutes. Therefore, that is also the time by which individual readings may be off or in which dynamics can not be captured correctly.

From these channels, we calculated the three size classes PM_{10}, $PM_{2.5}$ and PM_1 in the following way: We assumed the

Sensor	Fan	ISO	Shutter time	Framerate	Fokus	Resolution
B001	yes	400	33.33 ms	30	10 (inf)	1920×1080
B002	yes	200	33.33 ms	30	10 (inf)	1920×1080
B003	no	400	33.33 ms	30	10 (inf)	1920×1080
B004	no	200	33.33 ms	30	10 (inf)	1920×1080
B005	yes	400	33.33 ms	30	10 (inf)	1920×1080

Table 9.: Device configuration in CDPC evaluation.

ammonium sulfate particles to be spherical and homogeneous with a density ρ of $1.7\,\mathrm{g\,m^{-3}}$. With this, the size spectrum can be converted to a geometric diameter $d_{geometric} = d_{aerodynamic} / \sqrt{\rho}$. This can be used to calculate a volume-size-spectrum. Converted back to aerodynamic diameter, the respective size channels are summed up and multiplied with the density ρ to get a mass concentration for the three PM_x size classes.

The five sensors (dubbed *B001* to *B005*) were 3D printed and manually assembled, all of them identically constructed as described in section 4.4.2. Each of them was attached to a *Samsung Galaxy S6* smartphone with at least 32 GB internal storage capacity running Android version 7.0. In terms of video recording options, they were configured differently to test different sensitivity (ISO) settings. Also, three of the sensors were each vented using a *SUNON UF383-100* microfan running at a voltage of 1.8 V and two were not vented in order to compare the detection performance. The settings for the five sensors are listed in Table 9.

The recorded videos were processed with the Contour Detection Particle Counting (CDPC) algorithm described in section 4.5.3. An algorithm parameter sweep was performed in which the following parameters were varied:

- MOG2 learning rates r_{learn} of 0.1, 0.01, and 0.001
- number of standard deviations for Gaussian blur $\langle \sigma_{blur} \rangle$ of 5 and 9
- threshold for binarization $\theta_{b/w}$ of 10, 50 and 100
- threshold for contour detection $\theta_{contour}$ of 10, 50, 100, and 500

Initially, the results seemed disappointing. Sensors *B002* and *B004* did not yield useful scatter traces, probably due to the lower ISO sensitivity. While for each of the other three sensors

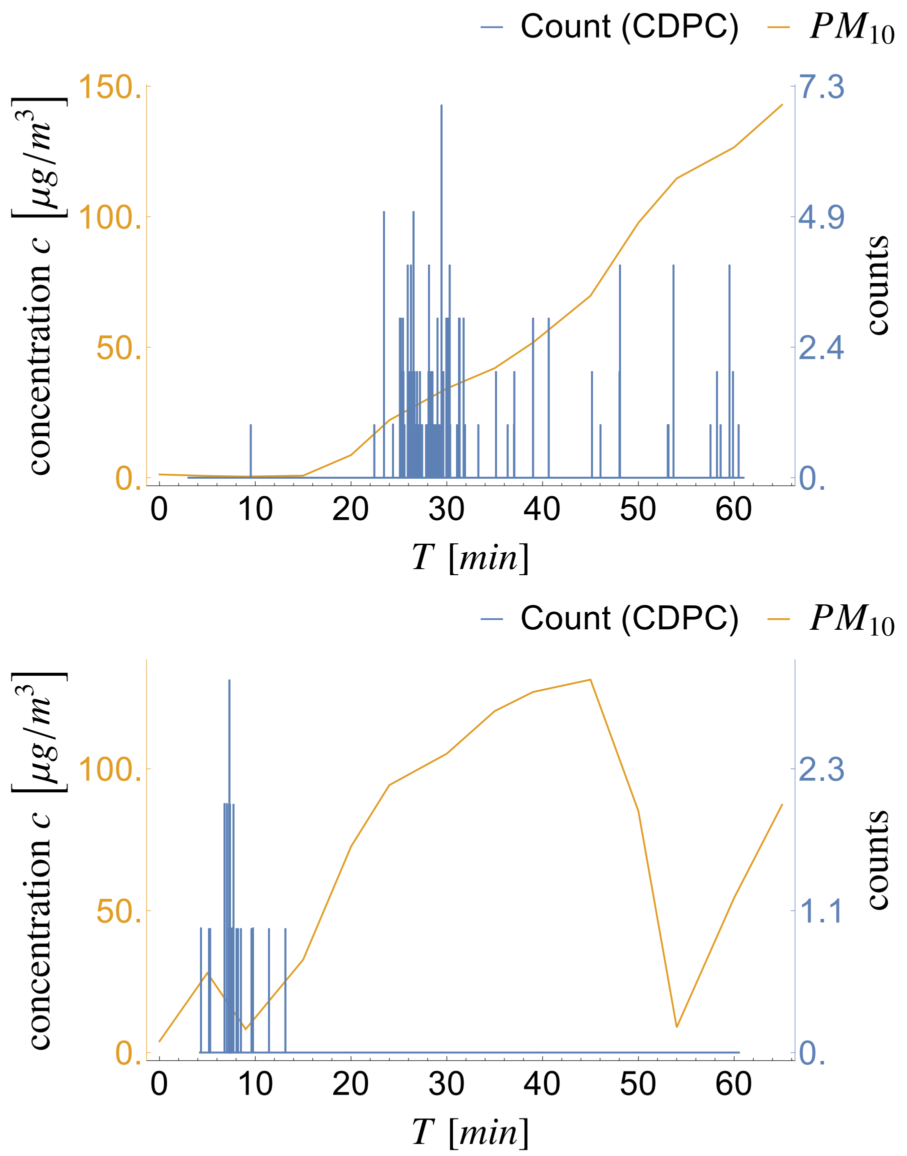
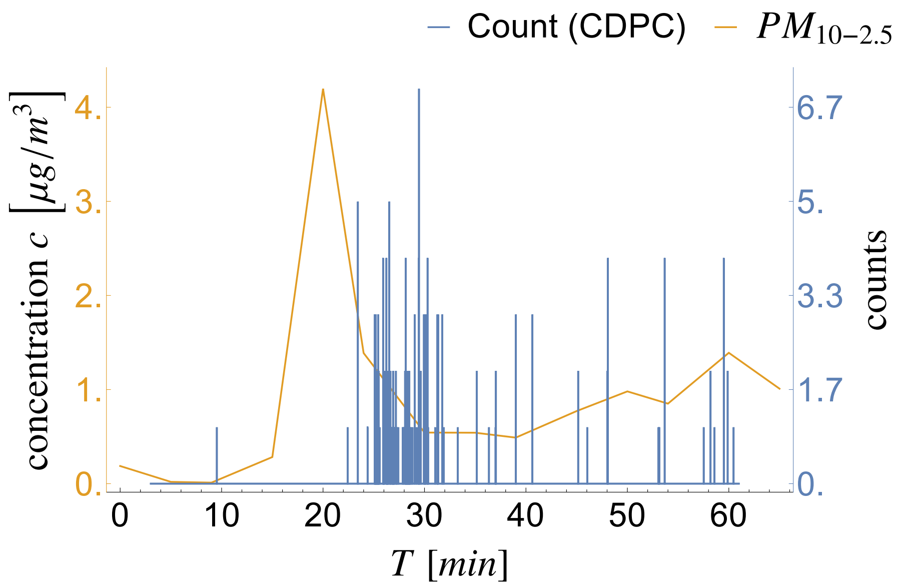


Figure 43.: Particle counts obtained from the CDPC algorithm vs. reference signal (PM_{10} size fraction) for sensors B005 (top) and B001 (bottom). CDPC parameters: r_{learn}: 0.1, $\langle\sigma_{blur}\rangle$: 9, $\theta_{b/w}$: 50.0, $\theta_{contour}$: 100.0.

the videos showed clear scatter traces that were visible to the naked eye, initially we could not find a clear correlation with either of the three size classes PM_{10}, $PM_{2.5}$ or PM_1. Instead, we observed a phase of high particle counts at the beginning of each of our measurement runs, that then faded in spite of the fact that the concentration levels continued to rise (see Figure 43).

Looking for patterns of explanation for these observations, we theorized that with our imperfectly constructed measurement chambers, we may only be able to see the scatter traces from



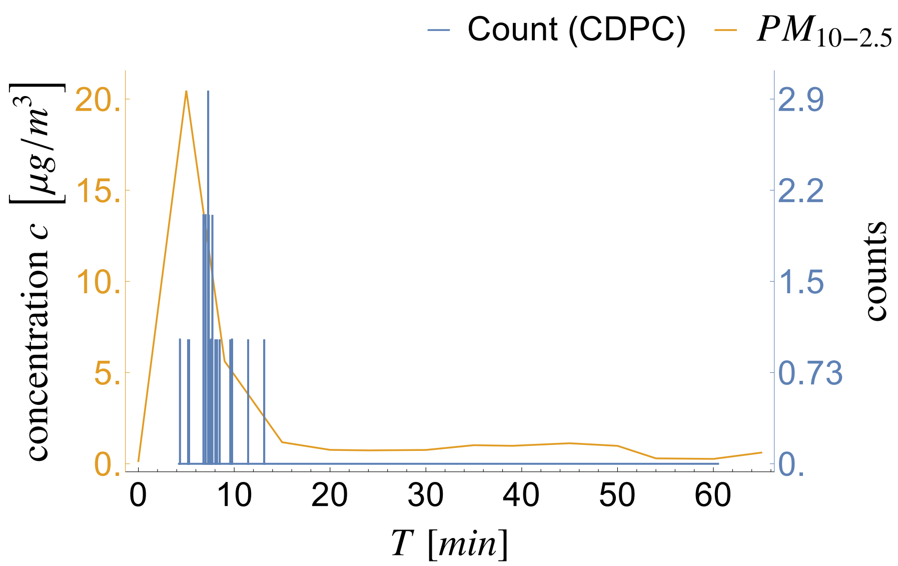


Figure 44.: Particle counts obtained from the CDPC algorithm vs. reference signal ($PM_{(10-2.5)}$ size fraction) for sensors B005 (top) and B001 (bottom). CDPC parameters: r_{learn}: 0.1, $\langle\sigma_{blur}\rangle$: 9, $\theta_{b/w}$: 50.0, $\theta_{contour}$: 100.0.

the fraction of the larger particles of the size spectrum. This assumption was supported by the fact that the scatter traces were most dense during ca. 15 minutes after injecting the ammonium sulfate into the chamber. The larger particles settle faster than the smaller ones and therefore could not be detected anymore afterwards.

In order to test this hypothesis, we calculated the fraction of Inhalable Coarse Particles as $PM_{(10-2.5)} = PM_{10} - PM_{2.5}$. This revealed an excellent qualitative correlation for the three sensors

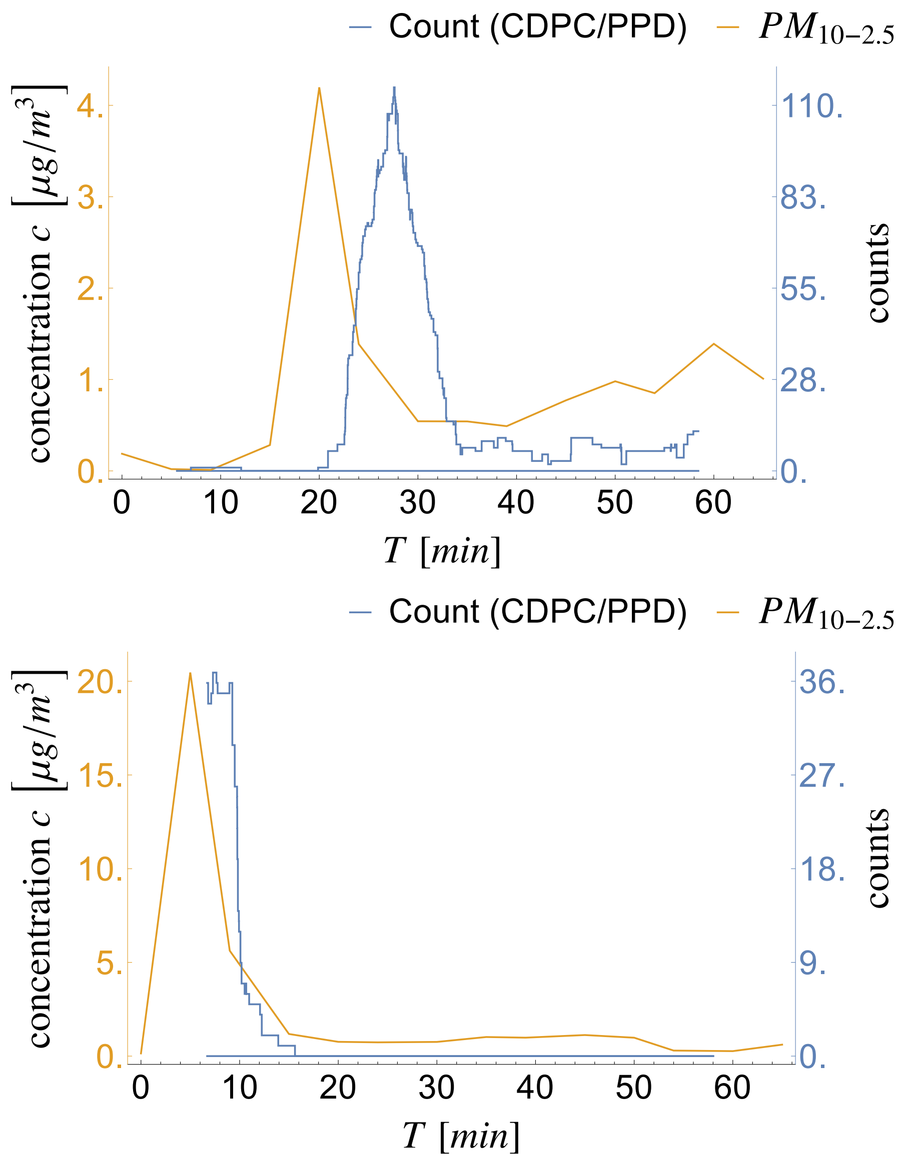


Figure 45.: Combined approach: The results from the particle counting (CDPC) algorithm are subsequently piped through the Poisson Particle Detection. The graphs show very good qualitative agreement with the $PM_{(10-2.5)}$ size fraction of the reference for sensors B005 (top) and B001 (bottom).

with high ISO settings (*B001*, *B003* and *B005*). Figure 44 shows the reference for $PM_{(10-2.5)}$ and the particle counts obtained from analyzing the recorded videos using the Contour Detection Particle Counting (CDPC) algorithm. Figure 45 shows the same, but after additionally feeding the output of the CDPC algortihm into the Poisson Particle Detection (PPD).

4.5.4 *Discussion*

This section discusses some of the limitations and possible future improvements of the prototype sensor.

Ventilated vs. Unventilated

Another aspect of our evaluation concerns the sampling of the air, i.e. whether the measurement chamber was actively vented by a fan or whether we relied on diffusion to transport particles into the sensor. Both the ventilated (Figure 45 below) and the unventialated version (Figure 46) of our sensor performed similarly.

However, whether this result can be transferred to real-life measurements remains unclear and needs to properly evaluated. In the lab environment, the aluminum measurement box itself was ventilated, possibly facilitating quicker air exchange also in the passive sensors.

Detection Size Limit

In our evaluation, the Contour Detection Particle Counting (CDPC) algorithm was only successful in detecting coarser particles. After these had settled, no individual counts were detected anymore. Outdoors this maybe different. As naturally occurring turbulence keeps the coarser size fraction suspended for a longer period of time (or resorbs particles), this limitation may be less relevant.

On the other hand, the detection size limit may be a hardware issue. Due to the imperfect measurement chambers that were a result of rapid prototyping the sensor, we may have lost information. All of our sensors exhibited different backgrounds due to 3D printing and manufacturing of the sensor modules (see Figure 47).

The background subtraction step of the CDPC algorithm turned out to remove some faint particle traces that could be identified by humans in the original recordings are subtracted along with the background image. A possibility would be to fine-tune the algorithm to each individual sensor case. However, since the background images vary, this would require a calibration step for each sensor. A much more sensible approach would therefore be to first perfect the light trap design, so that no background image is visible in the absence of particles. This can be achieved with a thorough optical design, which was not in the focus in this thesis. The construction material of

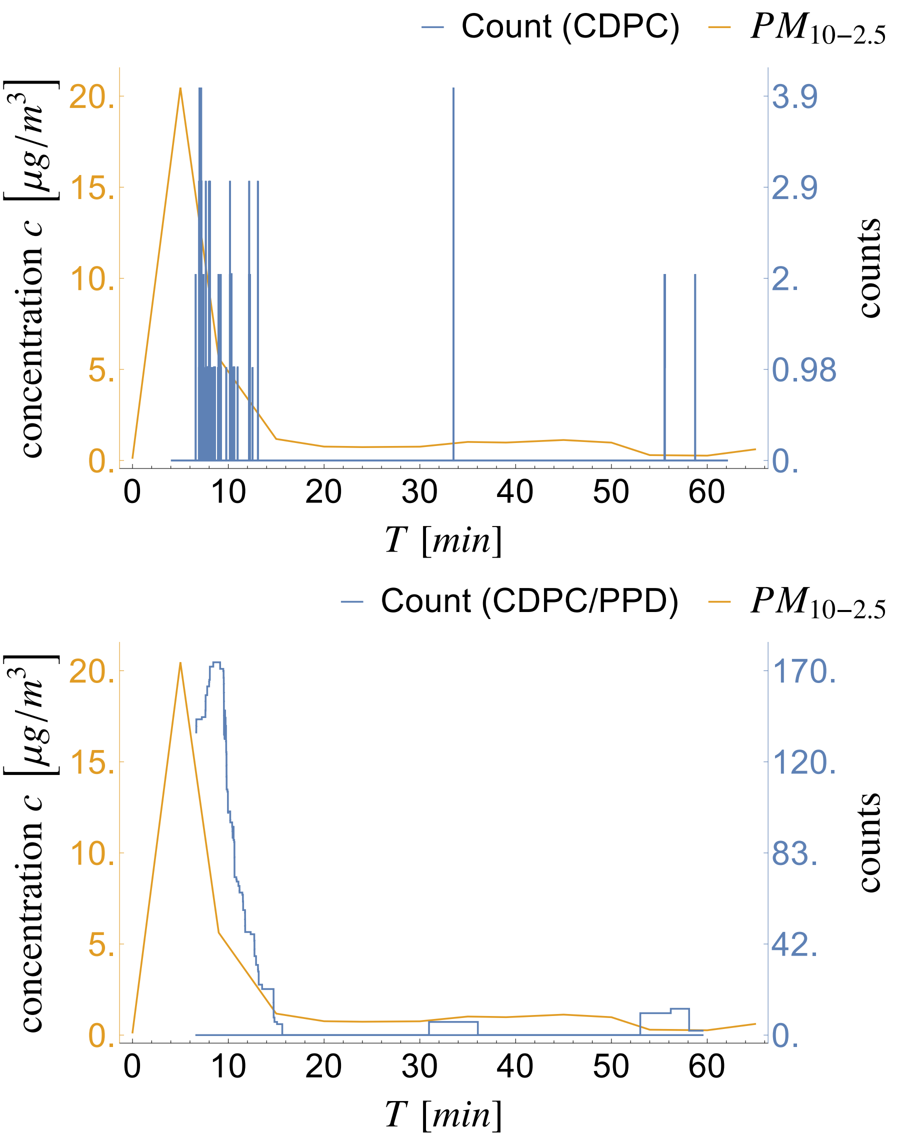
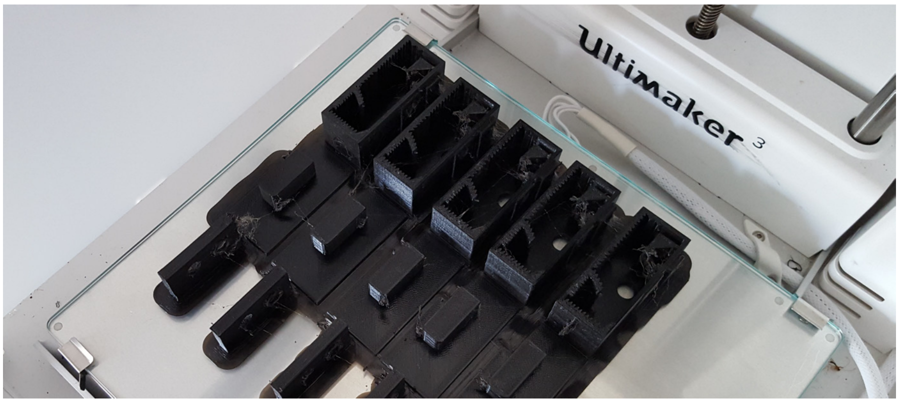
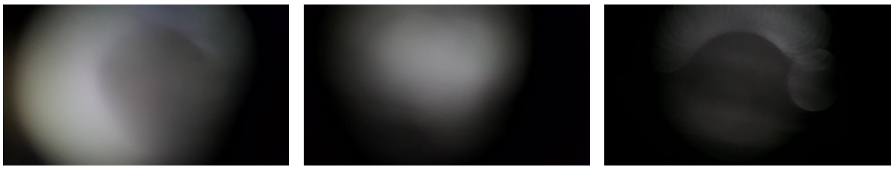


Figure 46.: Results from the unventilated sensor (B003). In out lab experiments, it did not make a big difference, whether the sensor was ventilated or not (cmp. Figure 45).

the chamber and its surface can also be improved from the 3D printed versions. Additionally, the inside of the light trap could be coated with absorbing paint.



(a)



(b)

Figure 47.: Due to rapid prototyping with a 3D printer (a) and the manual assembly of the sensor hardware, the measurement chambers were not 100 % identical, resulting in different background illumination (b).

4.6 CONCLUSION

In this chapter, we presented a novel means of measuring particulate matter with smartphones. In an iterative design process, we adapted the principle of low-cost light-scattering fine dust measurements into a passive clip-on sensor for the use with the camera and flash of a smartphone. This includes two algorithms for signal processing, one of which exploits the characteristics of Poisson processes to reconstruct the "true" signal from data afflicted with unknown systematic noise, accounting for the natural instability of mobile and wearable measurement setups for end-user environmental sensing. We have confirmed the principle of operation in a series of studies and reached excellent qualitative agreement with a professional reference device when measuring coarse inhalable particles at realistic concentrations.

5 Networked Sensing

Both chapter 3 and chapter 4 focused on enabling low-cost Particulate Matter (PM) measurements. While we demonstrated that remarkable results can in principle be achieved with cheap technology, we also saw that this often entails efforts to calibrate the sensors in some respect. This chapter deals with device-by-device calibration of mobile air quality sensors. The focus is placed on field calibrations that are either performed against reference stations or among two mobile nodes when they are co-located, i.e. having a so-called rendezvous. Existing calibration techniques are applied to data from low-cost laser-scattering PM sensors and the performance is discussed along with possible practical limitations. Subsequently, a novel privacy-preserving calibration scheme is presented and evaluated based on simulated ozone measurements and real-world taxicab mobility traces.

Parts of this chapter have previously been published. The privacy preserving Peer-to-Peer (P2P) calibration has originally been published in the proceedings of the *2nd EAI International Conference on IoT in Urban Space (Urb-IoT 2016)* [153]. An extended version is scheduled to be appear in the *EAI Transactions on the Internet of Things* [154]. The private calibration scheme was developed jointly with Jan-Frederic Markert, Gregor Schindler and Markus Klug.

5.1 Related work

While research on sensor calibration often addresses the same issues, different terminology is used, some of it interchangeable, some entailing different meanings.

Generally, there are two different ways of calibrating a device:

- against another device in close proximity, typically one with a higher reliability

- by exposing the device to a defined, often artificially created condition for calibration, usually in a lab environment

As we are concerned with distributed sensing systems, we focus on device-by-device calibration, sometimes also known as comparison calibration [143] or cross-calibration. Also, device-by-device calibration can either be carried out under factory or laboratory conditions or in the field. Other terms can be found for the distinction between calibration that is carried out by sensors autonomously (*auto* calibration [233]) in contrast to calibration that requires manual interaction (*manual* calibration [52]).

Regarding calibration with or without the existence of reliable, ground truth measurements, many terms can be found in the literature. So-called *blind* methods achieve calibration gain without ground truth reference data [11]. In line with this, TAN ET. AL use the term *semi-blind* when referring to calibration with partial ground truth data and *non-blind* for calibration with full ground truth data [224]. In addition to this, the terminology *off-line* respectively *on-line* [90] can also be found when referring to calibration with or without ground truth data, the latter is sometimes also called *in-place calibration* [51]. The terminology *multi-hop* respectively *single-hop* [104] does not make a distinction regarding the availability of ground truth, but rather the quality of the reference. Single-hop calibration means calibrating directly against reliable reference sensors, while multi-hop calibration makes use of sensors whose calibration was propagated through several nodes. Another term is virtual in-situ calibration[258]. Throughout this work we use the terminology of blind calibration and add the qualifier multi-hop where applicable.

Regarding the level on which the calibration is carried out, i.e. whether on the sensing devices themselves or on a central instance that collects all data, also different terms are used. Some authors refer to *device-level* respectively *system-level* [223] calibration, while the terms *micro* respectively *macro* calibration are also found [244]. Calibration schemes without a central instance that use only P2P communication are referred to as distributed calibration [158]. In addition, *local calibration* is also used in some works for calibration on the sensing devices themselves [223]. Throughout this work we use the terminology of device respectively system-level calibration as they are self-explanatory.

A lot of research has been done related either to mobile participatory sensing and its privacy implications, or to the calibration of dynamic sensor systems. However, there is only very limited work combining the two. To the best of our knowledge, *PPCS* [248] is the only privacy-preserving calibration mechanism presented so far. *PPCS* is a MIX-network-based pseudonymization scheme for mobile sensor systems with server-client architecture. It uses so-called *non-blind* calibration, i.e., relies on high quality ground truth reference data. Therefore, *PPCS* can not easily be applied to multi-hop settings with rendezvous-based calibrations. It thus is not so well-suited for end-user participatory sensing. The same applies to a slightly different version of *PPCS*, which was published under the name *PRICAPS* [247]. Another system, *TAPAS* [120], presents approaches to privately select participants for collection tasks – a technique that is not applicable to calibration.

In privacy preservation, *Proximity Testing* can be used to privately and "*continuously report all events of mobile users being within the distance of each other*" [215]. This can be used for the task of finding sensors that have measured the same phenomenon at approximately the same time and location. While private one-to-one matchings can reliably be done with pairwise exchanged keys, one-to-many matchings with proximity tests against an unknown number of strange users fail due to the bad scalability, especially regarding key exchange and pairwise distance calculation. Instead, spatial generalization as proposed in [241] can be applied. A *MIX network* [57] is a way to ensure that the network traffic and the corresponding devices are unlinkable.

Our privacy-preserving approach combines blind multi-hop calibration with Private Proximity Testing and a MIX network to build a privacy preserving rendezvous-based calibration scheme for participatory sensing scenarios. It is presented in detail in section 5.3.

5.2 MULTI-HOP CALIBRATION

As we have seen in our experiments with low-cost dust sensors described in chapter 3, Commercial-of-the-shelf (COTS) dust sensors may require calibration, as they may display inconsistencies amongst each other from the manufacturing procedure or may deviate over time. Therefore, such sensors do not only need to be calibrated before deployment, but may also need to be re-calibrated in certain intervals. For many of the scenarios outlined in section 2.2, this step needs to be performed in-situ, as collection and re-deployment is not a feasible option, especially when devices are operated by citizens rather than trained technical staff.

In order to deal with calibration issues in volatile low-cost sensing scenarios, multi-hop device-by-device calibration algorithms, in which sensors calibrate each other, have been proposed. In this section, we apply this approach to the calibration of low-cost PM sensors.

5.2.1 *Data*

Different calibration schemes were tested on real data from a series of *Nova Fitness SDS011* laser scattering sensors. As with the evaluation of our clip-on sensor for smartphones before, we conducted our comparison measurements at the World Calibration Center for Aerosol Physics (WCCAP). The setup in which this data was recorded is exactly the same as depicted in Figure 42 in section 4.5.3.

Figure 48.: *Nova Fitness SDS011* laser-scattering sensor used in our experiments (image by Ubahnverleih[1]).

We placed 17 *Nova Fitness SDS011* sensors (see Figure 48) — 14 of which delivered data — in an airtight aluminum chamber and subsequently let them sample ambient air (from outside)

1 `https://commons.wikimedia.org/wiki/File:Feinstaubsensor_SDS011.jpg`, made available under the *Creative Commons CC0 1.0 Universal Public Domain Dedication*.

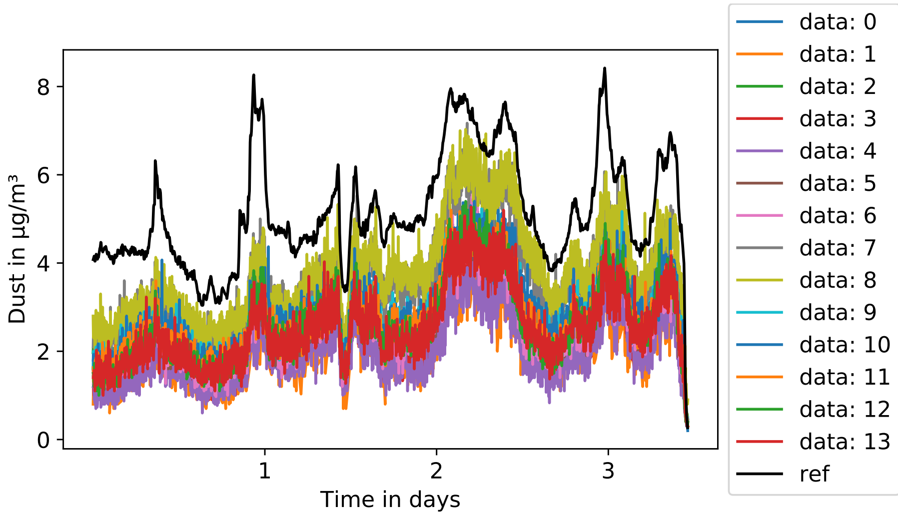


Figure 49.: Raw SDS011 sensor data vs. $PM_{2.5}$ reference data.

for ≈4 days. From two reference devices connected to the air outlet of the chamber, we calculated the three size classes PM_{10}, $PM_{2.5}$ and PM_1 (see section 4.5.3 for details).

The readings of the *Nova Fitness SDS011* all were similar and already initially showed generally good agreement with the reference, as depicted in Figure 49. This data set was then used to evaluate different calibration strategies in a simulated manner: Although in reality stationary, we simulated mobility by defining intervals in which sensors were assumed to rendezvous. The same approach has been applied by Hasenfratz et al. [104], on which our work is built.

5.2.2 *Calibration*

In this section, the results from simulating different calibrations are presented, based on the real sensor data as described above. We employed the calibration methods from Hasenfratz et al., who proposed a multi-hop calibration scheme for mobile sensors, in which the sensors utilize each others measurements from rendezvous in order to improve the calibration *on-the-fly* [104]. We investigated the calibration between sensors and a reference station as well as blind multi-hop calibration, in which sensors can calibrate each other without having access to ground truth.

In *forward calibration*, the calibration parameters are calculated based on the measurements from a certain time slot (T_{tup}). The calibration parameters are then used for future measurements

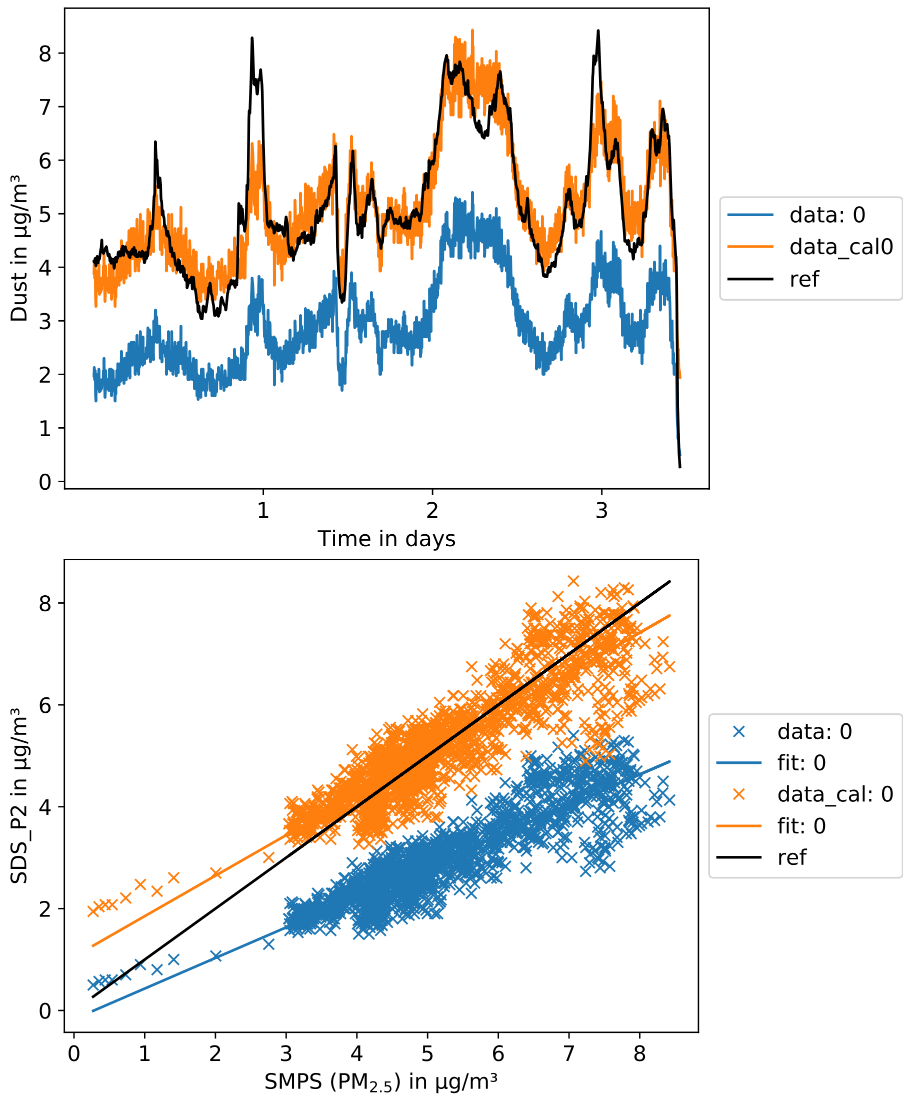


Figure 50.: The sensor is calibrated once. For the calibration all data is used. The parameters are calculated to $a = 1.324$ and $b = 1.281$. The RMSE decreases from $RMSE_{data} = 2.318$ to $RMSE_{data_{c}al} = 0.568$.

until the next calibration takes place. If these new parameters are also used to re-evaluate the previous forward calibration, the scheme is called *backward calibration*, which is used to "*improve the measurements accuracy if the sensor characteristics significantly differ during time period*" [104].

In the case of the *Nova Fitness SDS011* data, the sensors did not exhibit significant drift during the 4 days of measurement, but noticeable offsets among each other and relative to the reference (see Figure 49). To calibrate the sensors, generally a linear function with $f(x) = a \cdot x + b$ is used. The parameter a is

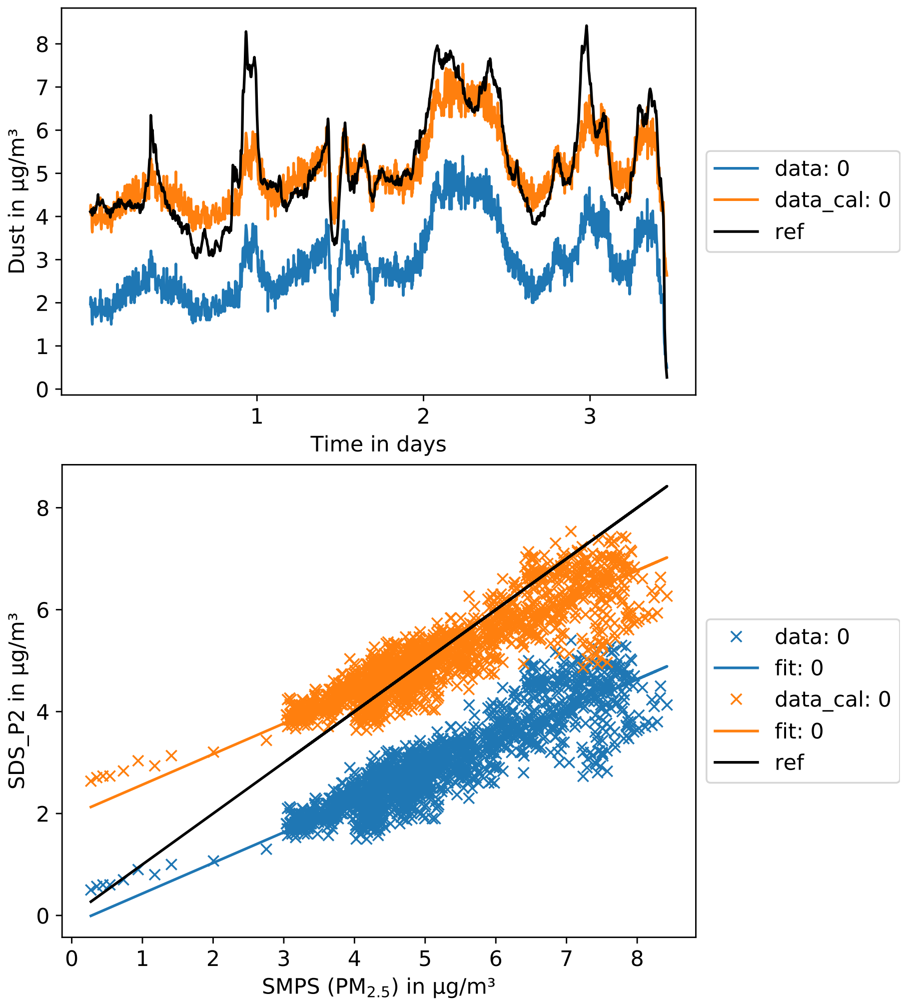


Figure 51.: The sensor is once calibrated right at the beginning with forward and backward calibration. The parameter a,b is calculated with the least-squares (LS) method to $a = 1$ and $b = 2.1355$.

the slope to widen or compress the distance between minimums and maximums and b is an offset. In a real case scenario a calibration is possible if a sensor passes a reference station or another sensor. Then, some amount of measurements are recorded in the same spatial and temporal window. This tuple of data is used for the calibration.

In cases where all data were to cluster around one value, a slope of ≈ 0 would be calculated and the data would only be shifted. As this leads to bad calibration, we used a threshold for the slope in order to avoid this. For slopes smaller than the threshold (0.7), the slope is set to 1 and only an offset ($f(x) = x + b$) is used for the calibration.

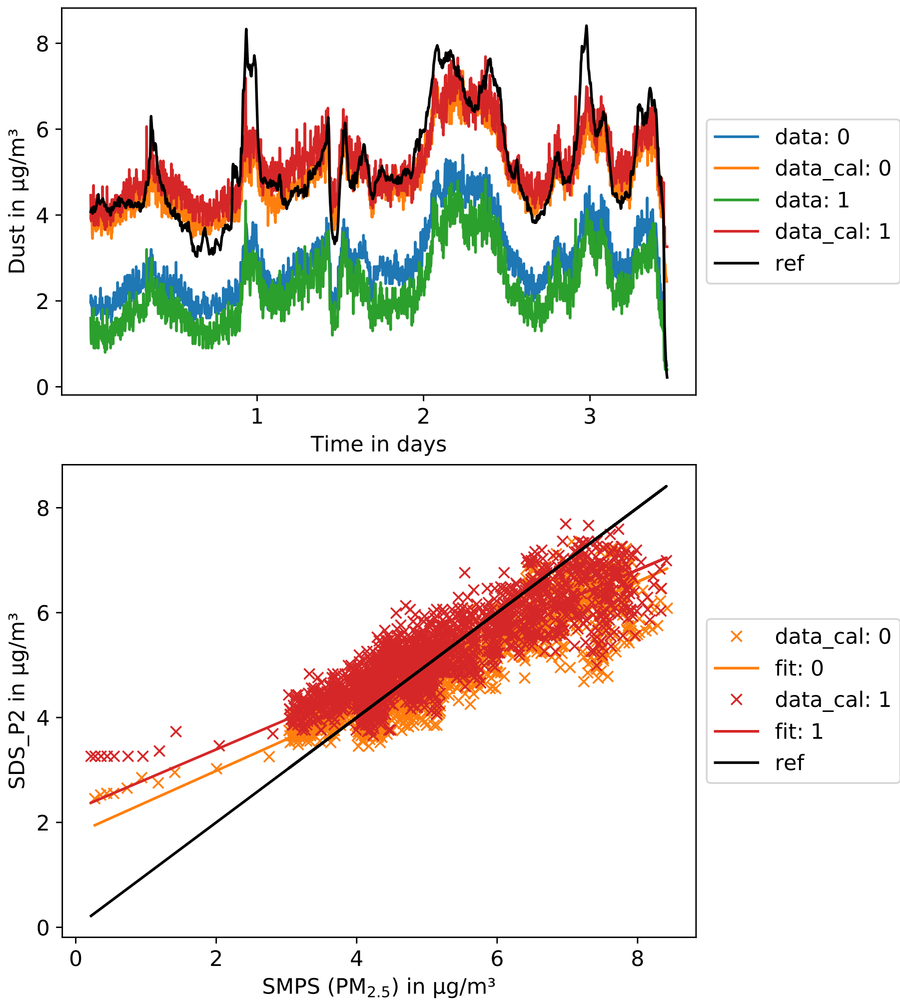


Figure 52.: Sensor 1 is calibrated with the reference at the beginning. After half of the time the second sensor (sensor 0) is calibrated with sensor 1 as reference. The slopes in both cases are $a = 1$ and the offset parameters are $b_1 = 2.861$ and $b_0 = 1.953$ with a $\mathrm{T}_{tup} = 10$ min.

In the following, different calibration scenarios are are shown. In the first case (see Figure 50) the data is calibrated once and all data is used for the calibration. Therefore, this can be considered a baseline. A linear calibration ($f(x) = a \cdot x + b$) is used. The parameters are calculated to $a = 1.324$ and $b = 1.281$. The Root Mean Squared Error (RMSE) decreases from $RMSE_{data} = 2.318$ to $RMSE_{data_cal} = 0.568$. For the original data and the calibrated the correlation coefficient to the reference is $r = 0.8917$ and the R^2-score is $R^2 = 0.795$.

In Figure 51 only one sensor is calibrated once, right in the beginning. The data used for the calibration is taken from a 10-minute-window ($T_{tup} = 10\,min$). With an measurement

Sensor	RMSE_{data}	RMSE_{cal}	R^2_{data}	R^2_{cal}
Sensor 0	2.318	0.719	0.795	0.690
Sensor 1	2.907	0.716	0.748	0.718
Sensor 2	3.263	0.902	0.684	0.528

Table 10.: RMSE before and after the calibration and R^2 after the calibration values of the Multi-hop Calibration

frequency of one reading every 2 minutes this results in ≈5 data points. The calibration is calculated through the least-squares (LS) method. The parameters a and b are calculated to $a = 1$ and $b = 2.135$. As parameter a is exactly 1, it can be assumed that only an offset calibration was used. The RMSE decreases from $RMSE_{data} = 2.318$ to $RMSE_{data_cal} = 0.638$.

In the next scenario (see Figure 52), there is one sensor (sensor 1) calibrated with the reference at the beginning. After half of the time the second sensor (sensor 0) "enters" the system and is also calibrated but with the first sensor as reference (2-hop calibration). The slopes in both cases are $a = 1$ and the offset parameters are $b_1 = 2.861$ and $b_0 = 1.953$ with a $\text{T}_{tup} = 10$ min. The RMSE for sensor 0 decreases from $RMSE_{data} = 2.318$ to $RMSE_{data_cal} = 0.689$. This example shows that even another sensor can be used to calibrate an uncalibrated sensor

The fourth method is again the multi-hop calibration as proposed by Hasenfratz et al. [104] (this time with more than 2 hops). It is used if two sensors that are already calibrated are at the same time at the same location. Then the sensor with the older calibration time is calibrated with an assumed "true value" ($y_{assumption}$). This value is calculated through:

$$y_{assumption} = \frac{t_{c1} \cdot p_2 + t_{c2} \cdot p_1}{t_{c1} + t_{c2}} \tag{15}$$

with p_i the calibrated measurement of the sensor i and t_{ci} the time since the last calibration of the sensor. Therefore in this scenario (see Figure 53), there are three sensors which calibrate either with the reference or with an other sensor. Every 180 minutes one sensor calibration takes place. The T_{tup} is 10 minutes.

The results of the multi-hop calibration are shown in Table 10.

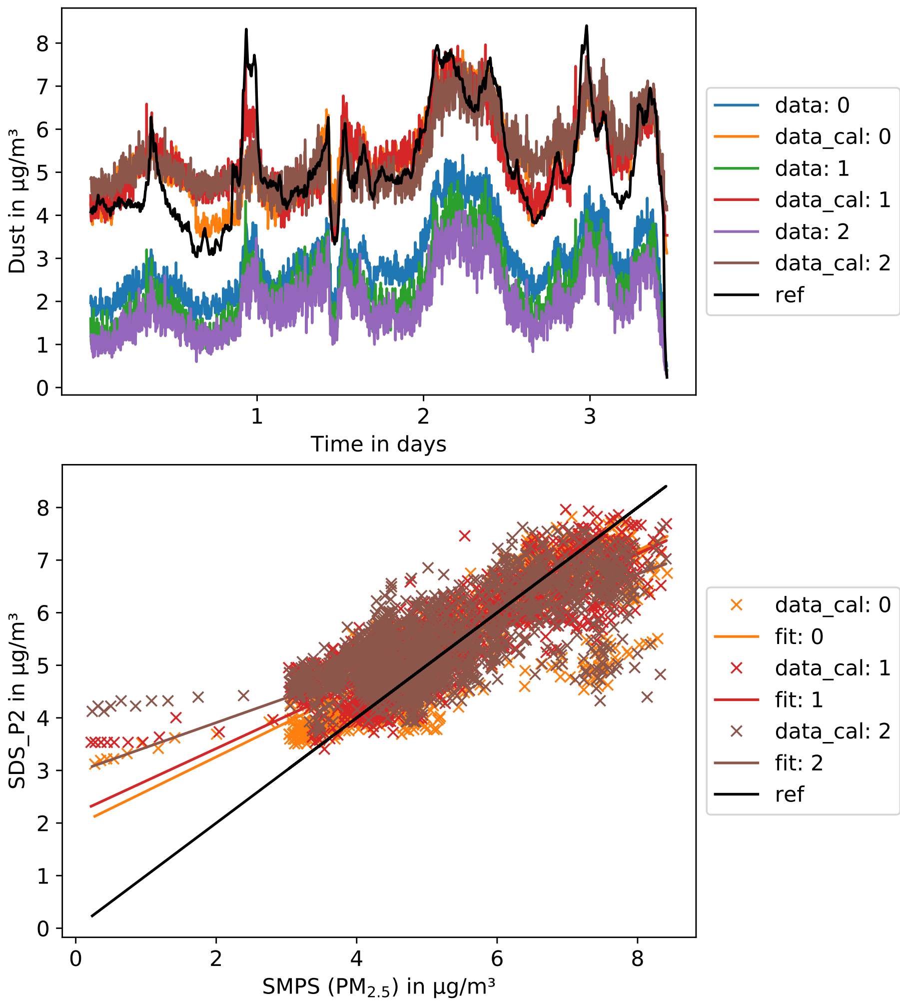


Figure 53.: For three sensors the multi-hop calibration is used. The sensors calibrate either with the reference or with another sensor every 180 minutes. The tuple time is 10 minutes, this leads to an offset calibration in the most cases.

5.2.3 *Discussion*

To compare the above presented calibration methods, the results are listed in Table 11. The data displayed for each calibration method was obtained using only one sensor (sensor 0).

Calibration	$\mathbf{RMSE}_{data}$	$\mathbf{RMSE}_{cal}$	$\mathbf{R}^2_{data}$	$\mathbf{R}^2_{cal}$	$\mathbf{r}_{data}$	$\mathbf{r}_{cal}$
over all data	2.318	0.568	0.795	0.795	0.8917	0.8917
with reference	2.318	0.638	0.795	0.795	0.8917	0.8917
2-hop	2.318	0.689	0.795	0.795	0.8917	0.8917
multi-hop	2.318	0.719	0.795	0.690	0.8917	0.8309

Table 11.: RMSE, R^2 and the correlation coefficient r to the reference before and after the calibration and R^2 after the calibration values of the Multi-hop Calibration

Because the sensors are very stable, it is enough to calibrate the sensor once (considering the overall measurement time of only four days). A linear calibration with all data leads — as expected — to the best result. Calibrating with another sensor leads to a small loss in accuracy. In the case of stable sensors, the multi-hop calibration is not as good as expected, because the sensors calibrate each other many times. This leads to a loss in accuracy compared to the once calibrated sensor. In all these scenarios, except the calibration over all data, in most cases an offset calibration is used because the tuple times are quite small and the introduced slope threshold of 0.7 is not reached.

To show the result without the threshold for the slope, the multi-hop calibration for the three sensors is shown in Figure 54. The calibration also takes place every 180 minutes but with a tuple time of 120 minutes.

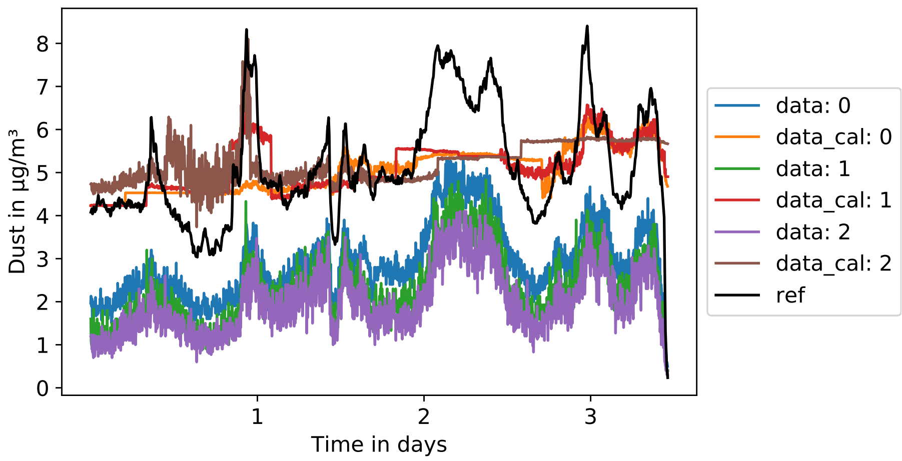


Figure 54.: The sensors are multi-hop, linearly calibrated with the reference every 180 minutes and with and tuple time of 120 minutes.

As can clearly be seen, often the slope is ≈ 0 and thus the measurement results in a fixed value. Depending on the variety of the measured variable, very long tuple times can be needed. In this data even a tuple time of two hours is not as good as basis for a simple offset shift. This is displayed in Figure 55: Because the tuple time is 360 minutes the calibrations only done every 720 minutes. Moreover only one sensor is used to calibrate with the reference. Still, we can see that before day one, there is very little dynamics and therefore not all calibration parameters are chosen very well. If sensors were to calibrate each other with such data as basis, this would lead to huge errors.

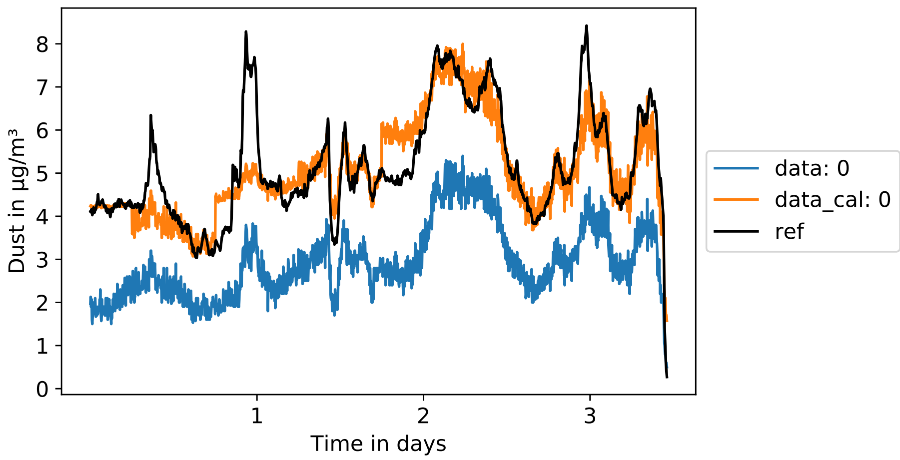


Figure 55.: The sensor is linearly calibrated $f(x) = a \cdot x + b$ with a reference every 720 minutes with an tuple time of 360 minutes.

The sensors used for this experiment are very stable. One single offset calibration in four days was even enough to achieve a good result. Thus the multi-hop calibration was not required in this time period. Considering other cheap sensors, this could be more important. A multi-hop calibration would bring more accuracy if the sensors drift is higher. In this case the results are better the more sensors are used. If the sensors exchange their data, it is important that the privacy of all participants is respected. The privacy aspect of such sensors is regarded in the next section.

5.3 privacy-preserving calibration

In the previous section we have seen that rendezvous-based blind calibration is a possible strategy to compensate for systematic error and to prevent quality loss in Participatory Sensing (PS) scenarios with low-cost sensors. However, the proximity-based data exchange approach generally entails privacy issues: Partial traces might be identified based on location information, such as frequently visited places or velocity, network characteristics (e.g. latency), or others.

This section presents a privacy-protecting calibration scheme for participatory environmental sensing. Collaborative blind macro-calibration is combined with several privacy preserving measures, such as private proximity testing [215], personalized exclusion zones, spatial generalization [241], pseudonymization and MIX networks [57].

5.3.1 *Preliminary Assumptions*

The definition of privacy in this work is to "guarantee that participants maintain control over the release of their sensitive information" [64].

ATTACKER MODEL Attackers can be administrators, participants as well as external entities. The attackers' role is either passive or active: passive attackers may eavesdrop on communication, while active attackers might also compromise servers and communication. Their motivation is assumed to be either malicious or honest-but-curious. Attackers' objectives can be rather general, e.g. desiring the traces themselves, or more specific, e.g. being interested in the location of a certain person at a specific time. Furthermore, the attackers can enhance their capabilities by utilizing additional information, e.g. publicly available address information from yellow pages or frequently visited places found on social media.

TRUST MODEL The participants trust the devices' soft- and hardware to correctly implement the scheme. Moreover, they trust the system administrator for choosing reasonable privacy-affecting parameters. The network provider is also trusted, as it already knows the nodes' approximate location.

Further, the server is not taken as honest or benevolent. Positioning services such as the Global Positioning System (GPS) are assumed to utilize passive client applications and thus need not to be trusted.

5.3.2 *Approach*

Our privacy-protecting collaborative blind macro-calibration method can be decomposed into the following separate parts:

1. Sensing
2. Proximity Testing
3. Calibration
4. Upload

STEP 1: SENSING The first step naturally is the process of the sensors taking measurements. Each reading consists of essentially three data entries: location, time and the measurement itself.

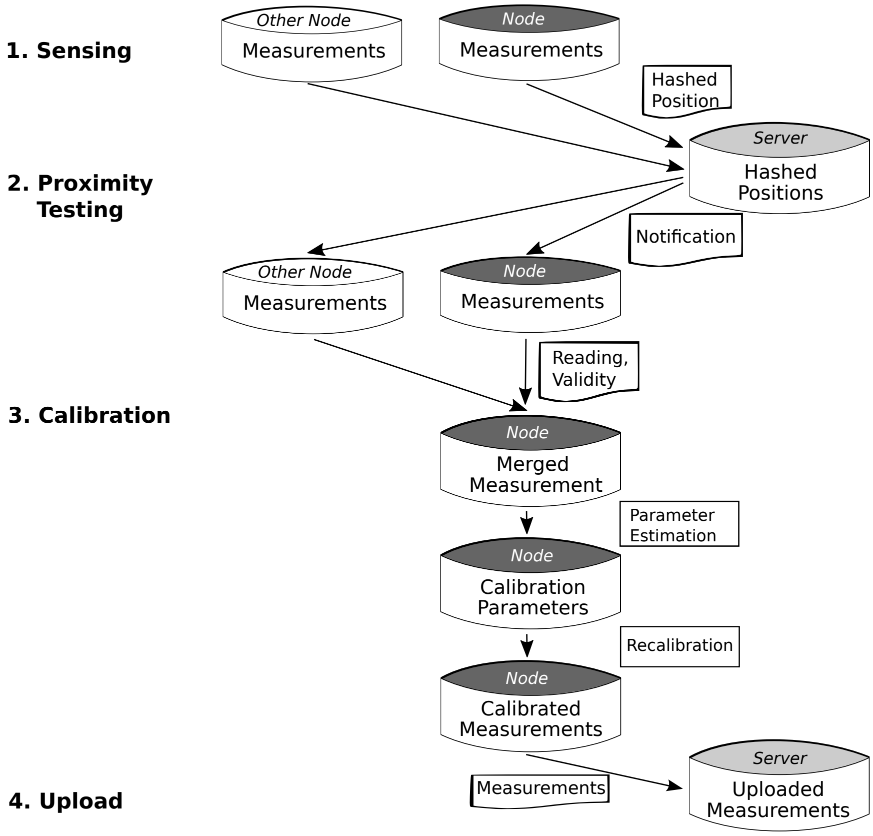


Figure 56.: Calibration pipeline: Rendezvous are identified by the server (1) and the sensor nodes' respective measurements and validities are exchanged (2). The node with the lower validity merges these into calibration tuples, estimates the calibration parameters by linear regression and recalibrates its measurement parameters accordingly (3). Finally, the measurements are uploaded to the server (4).

Measurements with low-cost sensors typically deviate to a certain degree from the ground truth. This measurement error is composed of two parts: (1) The statistical error, caused by random hardware noise or inaccuracies in the measurement apparatus, as well as the statistical nature of the measurement process; (2) the systematic error, depending on multiple factors such as the sensed phenomena and the environment. With low-cost sensors, the systematic error may increase with time, e.g. due to *sensor aging* or other causes [39]. Some sensors, e.g. electro-chemical gas sensors, are more susceptible to this kind of sensor drift than others.

In order to represent the reliability of a sensor's measurement, we introduce the *validity* (v) as a meta attribute. The validity ranges between 1 and 0, with an initial value of 1 representing a status of perfect calibration. As the sensors' systematic error increases continuously due to sensor aging, the validity decreases monotonically. The daily decrease depends on a global parameter *dailyValidityLoss* and is calculated as follows:

$$v(t+1) = v(t) * (1 - dailyValidityLoss) \tag{16}$$

Accordingly, the half-life (hl) of the validity can be calculated as

$$hl(dailyValidityLoss) = ln(2)/dailyValidityLoss \tag{17}$$

For an exemplary half-life of five days, a daily validity loss of 0.138 would be suitable.

While the statistical error is random, calibration is required to estimate the systematic measurement error and subsequently minimize the measurements' deviation from the actual values. The systematic error approximation is based on rendezvous and uses the fact that two spatially and temporally close sensors should measure the same value for a phenomenon. Depending on the homogeneity of the phenomenon different values for the temporal and spatial closeness are required. These so called rendezvous are determined through Private Proximity Testing via a server.

STEP 2: PROXIMITY TESTING While pairwise distance comparison against a proximity threshold suffices for proximity testing [166], a private implementation requires the exchange of private keys between each pair of nodes and pairwise operations hamper scalability. This complexity is not manageable in a large-scale network of mutually strange nodes. Instead, we utilize a reduction of proximity testing to equality testing via spatial generalization similar to [241]. The positions are mapped to cells via a globally deterministic function and the resulting cells are compared for equality. Additionally, from the privacy perspective, this coarsens the location and reduces the detail of the released personal information.

The grid characteristics have an impact on the quality of the proximity detection. The basic grid form is a composition of distinct rectangular cells. In order to better approximate a circular neighborhood, multiple mutually offset hexagonal grids can be utilized [166]. Furthermore, the size of the grid cells impact the neighborhood relation: the larger the cells, the greater the rendezvous neighborhood, and the less detailed the released personal information.

The temporal and geographic sampling position is first discretized to the corresponding cell in the grid. As the discretization also involves the temporal dimension, the same geographic position will be in a different cell regularly, preventing frequency analysis attacks to infer population density of certain locations. The distinct cell identifier is then mapped with a cryptographic hash function, making it impossible to recognize the original cell. Depending on the grid, an appropriate hash length needs to be chosen in order to prevent conflicting hashes. Finally, the hash value is uploaded to the server along with a pseudonym in order to query for rendezvous.

The rendezvous detection is done centrally on the server. For every newly uploaded query, the server checks for matches in the set of already uploaded queries. If a match is found, a data exchange between the co-located nodes represented by the pseudonyms is established.

The exchanged data includes the measurements and the respective validity. As the validity is sensor- and thus person-specific, this can have privacy implications. The probability of validities to be relatively distinct is rising with the decrease of the measurement density in the corresponding cells. As a countermeasure, the validity is discretized according to a global discretization step before exchange. In order to protect the participants privacy during the exchange, a secure communication channel is established via asymmetric encryption. The discretized validity and the calibrated measurement are then sent to the respective rendezvous partner.

The presence of ground truth sensors (i.e. reference stations) in the system is not required, but can improve the calibration performance. Reference stations act as regular nodes, except that they do not move and exhibit no measurement error (constant validity of 1). Their measurements along with the respective cell hashes are also accessed by the server. In case of rendezvous, the server performs the data exchange on behalf of the stations.

In order to be able to protect the participants' privacy also in low density areas, we added tailored sensing in the form of personalized exclusion zones to our scheme. The participants can set up so called sensitive locations, for instance their home or workplace. Subsequently, entries that are located within a given radius of such a sensitive location are discarded. In such areas, subsequent measurements might otherwise be linkable to a trace utilizing prior knowledge on mobility patterns, such as speed and frequent whereabouts.

STEP 3: CALIBRATION The computation of the calibration parameters and the calibration application is done locally on the nodes to reduce possible privacy implications as the server could link successive characteristic parameters to re-identify participants.

To perform a calibration, two prerequisites have to be met: (1) At least one of the participating sensors possesses a sufficient validity (*rendezvousValidityThreshold*). This ensures that a calibration actually results in an accuracy improvement. (2) There was no recalibration based on a rendezvous that happened later. As a result of network latencies, the server e.g. might recognize a rendezvous before the data on a different rendezvous that actually happened before that one is processed. If a preceding rendezvous is recognized later, it is therefore discarded as outdated.

The two calibrated measurements retrieved from the rendezvous are merged in order to estimate the unknown ground truth. The validities representing the measurement's reliability are utilized as weights. Thus, the estimation is calculated as a validity-weighted arithmetic mean of the two measurements. The estimation and the rendezvous time constitute the so-called calibration tuple: $\{estimate, \ time\}$.

After that, the most recent calibration tuples are merged with the sensor's respective uncalibrated measurements $\{r\}$. The number of chosen calibration tuples depends on a global parameter (*calibrationWindowSize*), and its choice has great impact on the calibration performance: while a higher value yields a more solid calculation basis for regression, the chance of considering already outdated measurements increases.

For the calibration parameters' calculation, different regression methods can be applied depending on the characteristics of the systematic error. For a systematic error best described as polynomial of first order depending on ground truth, linear regression with the method of least squares is utilized for error approximation. However, when the data range is below a threshold (*minimumDataRange*), linear regression can lead to poor results. In this case, we model the systematic error as constant and disregard the present dependency on the ground truth. In both cases, the calibration parameters are updated after error approximation and the following measurements are calibrated accordingly.

In order to account for the calibration gain, the validity is updated after the calibration. While the sensor with the higher validity keeps his validity, the other rendezvous partner adopts the higher value.

Additionally, a so called validity boost is applied, slightly increasing the validity for both. The boost accounts for the calibration gain that not results from calibrating with more valid sensors, but from the fact that rendezvous among uncalibrated nodes still yield positive effects when accumulated for many sensors with different errors. The validity boost, parameterized by a global parameter *validityBoost*, is applied by the following function:

$$v' = \frac{v + validityBoost}{1 + validityBoost} \tag{18}$$

This function ensures that the validity never exceeds 1.

STEP 4: UPLOAD We implemented different measures to ensure privacy in the data upload step: The participants' privacy with respect to network communication is protected through the use of a MIX network and dynamic pseudonyms. There are different types of MIX networks, that exhibit e.g. different latency characteristics. We assume that a suitable implementation as in [78] is realized.

Pseudonyms are freshly chosen for every communication, in order to prevent any linking. The pseudonym length and the decentral generation mechanism are chosen in a way to prevent pseudonym collusion, which depends on the size of the area to be monitored, the number of nodes and the communication frequency.

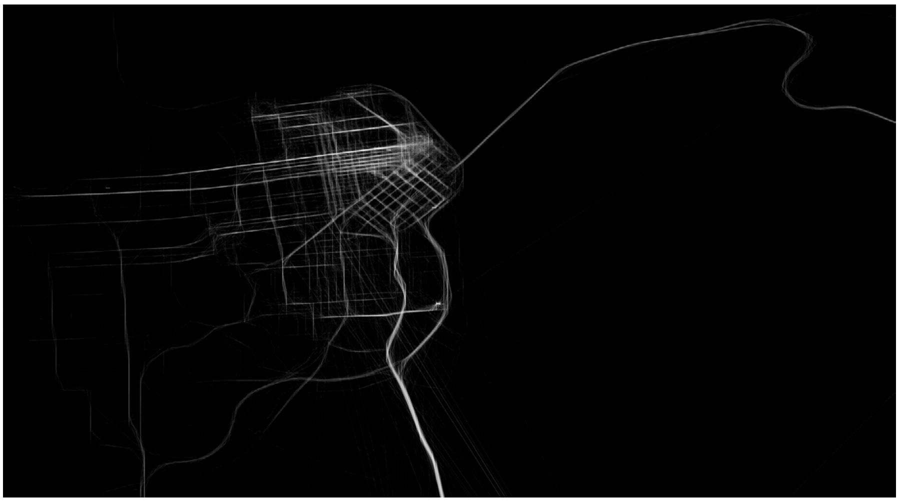


Figure 57.: For the evaluation we used taxicab location traces from the *epfl/mobility* data set at CRAWDAD (image courtesy of Stamen Design)[2].

Additionally, to prevent attacks based on the upload time, uploads are globally limited to certain points in time defined by a periodic interval. Finally, the uploaded data consists of the calibrated measurement and the respective time and location. There is no need for an identifier, as the calibration is finished with the upload.

5.3.3 *Evaluation*

We evaluated our scheme by combining simulated ozone measurements with real-world taxicab mobility traces: For the location traces of the simulated mobile nodes we use data from the *epfl/mobility* data set at CRAWDAD [185]. The data set contains real-world GPS traces of 537 taxicabs tracked while serving in San Francisco, USA (see Figure 57). As the data set is limited to 22 days, so is our simulation time. The measurement frequency results from the respective GPS logging frequency and amounts to once per minute on average.

For the simulation of the ground truth ozone distribution, we use data from a noise generator based on a free implementation of the *OpenSimplex* noise generation algorithm [219]. The three-dimensional noise has a continuous gradient in all dimensions and nearly no artifacts. We assume ozone to be homogeneous in the order of 30 minutes in time respectively 100 meters in space, in line with [103]. Its amplitude ranges between 0 and

2 https://stamen.com

Number of mobile nodes	50
Number of reference stations	0
Spatial grid form	basic (quadratic)
Spatial cell length	100 m
Temporal cell length	30 min
Calibration window size	10
Minimum data range	35
Daily validity loss	0.13
Validity boost	0.00003
Validity discretization step	0.0003

Table 12.: Scheme parameters of simulation setup.

140 ppb (parts per billion) as common ozone concentrations range between 0 ppb and 70 ppb [104] and EU regulations state 90 ppb as information threshold and 120 ppb as alert threshold [226].

In line with e.g. [52], we model the measurement error as the sum of two separate components: The statistical error is modeled with a Gaussian distribution: $n \sim N(0, \sigma^2)$ ppb. Its variance is chosen at the beginning of each day individually for each sensor: $\sigma \sim N(1, 3)$ ppb.

The systematic error b is modeled as a function of the measured value as well as the sensor age. Based on the literature [104], [254] the systematic error linearly depends on the ground truth and increases with time:

$$b(gt, t) = b_0(t) + b_1 * gt \tag{19}$$

where the coefficients are determined by uniform distributions:

$$b_0 \sim U(-9 - \frac{d}{5}, 9 + \frac{d}{5})\ ppb \tag{20}$$

$$b_1 \sim U(-0.2, 0.2)\ ppb \tag{21}$$

By introducing a temporal dependency for b_1, sensor aging is incorporated. The coefficients are updated on the beginning of each day, thus t denotes the past full days since deployment. For the systematic error model to be more realistic, the parameters are interpolated between two subsequent days in order to obtain a continuous function of time.

The simulation setup regarding the scheme parameters is shown in Table 12.

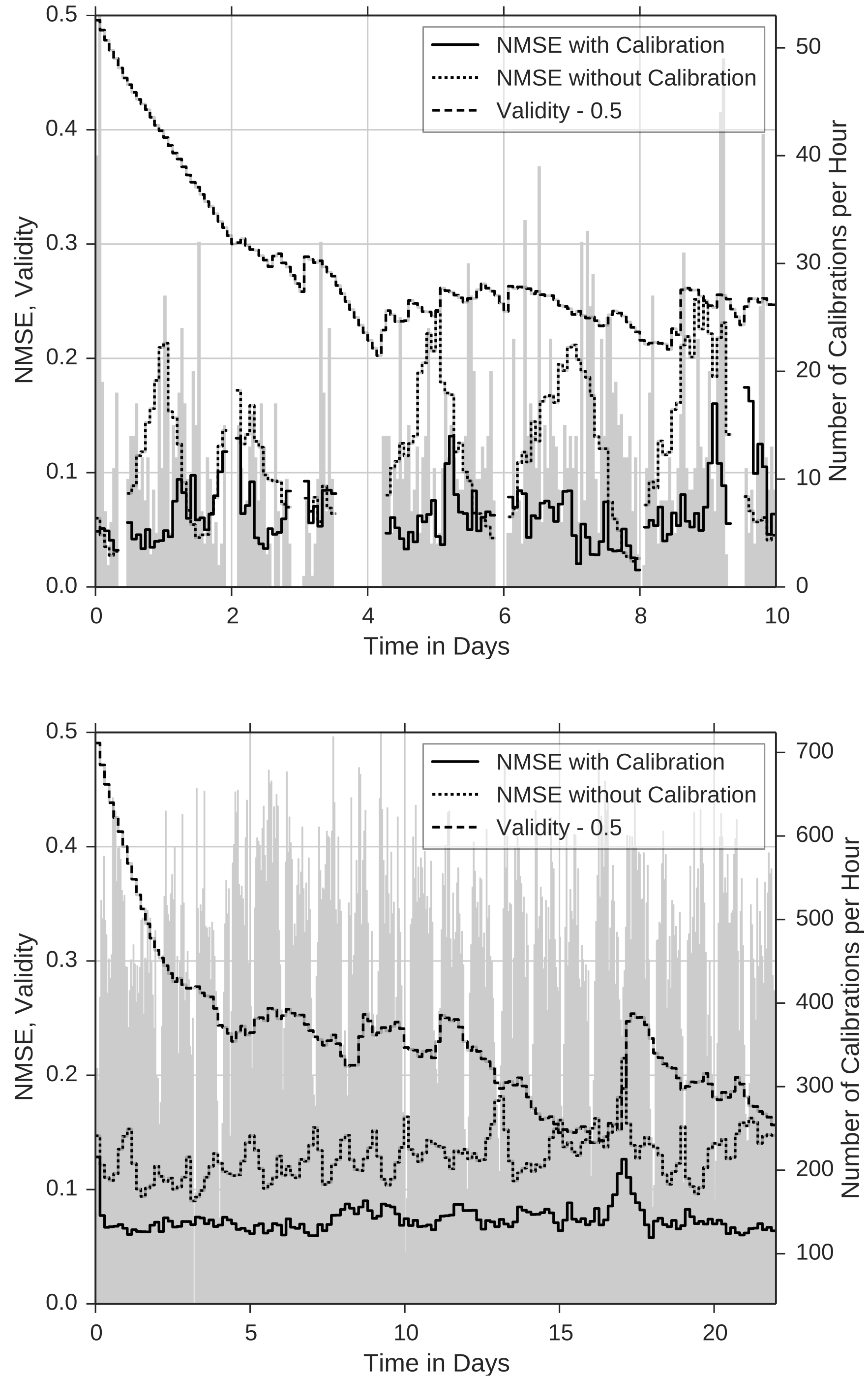


Figure 58.: NMSE with and without calibration and validity (note numerical shift). Gray bars in the background represent number of calibrations per hour. (a) Exemplary course of a single node for 10 days. (b) Course averaged over all 50 nodes for 22 days.

CALIBRATION GAIN The calibration gain, a measure for the effectiveness of a calibration, is computed as the ratio between the difference of the Normalized Mean Squared Error (NMSE) between uncalibrated and calibrated measurements, normalized by the uncalibrated NMSE:

$$calibrationGain = \frac{NMSE_{uncalib} - NMSE_{calib}}{NMSE_{uncalib}} \tag{22}$$

The mean squared error is a standard metric to quantify measurement errors [200]. The NMSE, the mean squared error normalized by the ground truth, is calculated as follows:

$$NMSE = \frac{1}{n}\sum_{i}^{n}\frac{(m_i - gt_i)^2}{gt_i^2} \tag{23}$$

summing over all nodes at all time steps.

For the sake of representation NMSE and validity over time in Figure 58 were created by temporally binning the data with 150 bins, hence the angular course.

In Figure 58 (top) we see a calibration course of a single exemplary node. The NMSE of calibrated measurements (solid line) in comparison to uncalibrated measurements (dotted line) is improved at nearly every point in time.

Figure 58 (bottom) shows the calibration course averaged over all nodes of the same simulation. The NMSE of calibrated measurements increases much slower and remains nearly constant despite sensor aging. Generally, the quality of the calibrated measurements is significantly better than the uncalibrated measurement. While the calibrated NMSE ranges around 0.8, the uncalibrated NMSE fluctuates around 1.4, yielding a calibration gain of 65%.

The periodicity of the uncalibrated error can be explained by the systematic error model, which interpolates between daily chosen parameters. Remarkably, this periodicity vanishes in the calibrated error course, indicating that the remaining error is for the most part of statistical origin.

SYSTEM LIFE TIME In Figure 58 (top) the exponential validity loss is best recognizable at times where no calibrations are present, especially at day 4. This loss is slowed when calibration processes are happening. At areas of high calibration density, depicted by the gray shaded bars, the validity stabilizes, increases or even jumps due to the implemented validity boosts.

The global validity in Figure 58 (bottom) drops with advancing time, as the validity boosts are not able to handle the global loss. Still, in times with a high number of calibrations, the validity rises again as the boosts dominate. The graph shows the trend that the validity ranges between 0.65 and 0.75 from day five on, with highs and lows. If the global validity drops under a specific threshold, it is assumed that the system is not able to recover itself and it stops yielding reasonable data. This marks the end of the system's life.

The expectable system life time without calibration is determined by the validity threshold and half-life. If a critical node

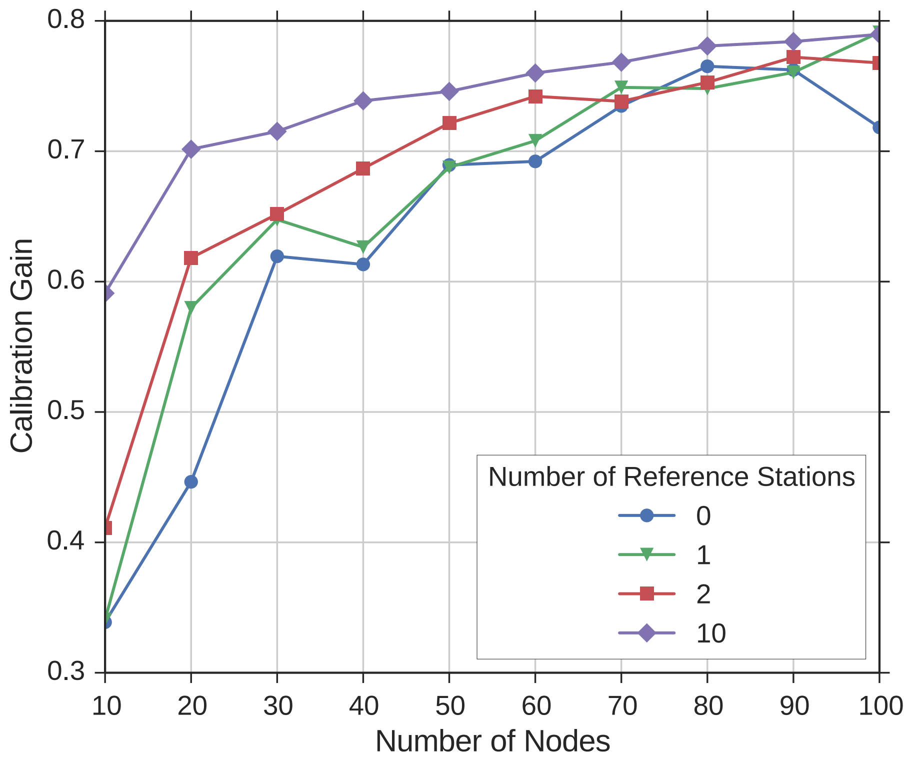


Figure 59.: Competitiveness of pure rendezvous-based calibration is shown by impact of reference stations.

density and subsequently a sufficient number of calibrations is reached, the life time is significantly prolonged. With a sufficient amount of reference stations, this could enable a hypothetically infinite system life time. The requirements for such a state are highly dependent on the data set and the validity configuration boost.

REFERENCE STATIONS The impact of reference stations is shown in Figure 59. In order to achieve reliable results, multiple simulations are fused in the diagram. The reference stations

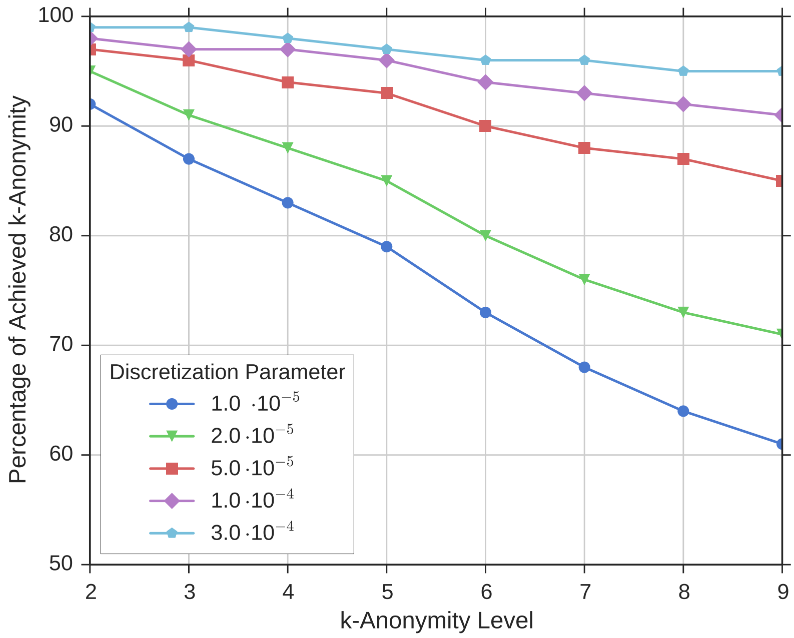


Figure 60.: K-anonymity in dependency of validity discretization, depicting the decrease in anonymity with smaller discretization steps.

were placed strategically at the most frequented locations. It is obvious that the deployment of more reference stations results in better calibration gains. However, the difference compared to a setting without reference stations diminishes with an increasing number of nodes, resulting from the utilization of rendezvous among imperfect nodes. This shows that our rendezvous-based approach performs best when deployed in a greater scale, and that the accuracy can compete with reference stations.

IDENTIFICATION VIA RENDEZVOUS The risk of trace reconstruction via rendezvous increases with low measurement density and high validity diversity. The validity diversity can be measured with k-anonymity [221]. Here, k represents the number of validities that are discretized to the same value: the lower the density and the higher the diversity, the lower k.

Figure 60 shows the percentage of achieved k-anonymity for different discretization steps. It can be seen that the smaller the discretization steps, the lower the percentage of achieved k-anonymity and consequently the higher the potential privacy risk. As the validity is only utilized as a weight, even a discretization with the largest of the tested steps is not expected to impair the calibration performance significantly. Thus, given a reasonable validity discretization, a successful privacy protection with respect to trace reconstruction via rendezvous is feasible.

5.3.4 *Discussion*

In the course of our evaluation we decided on certain parameters (proximity thresholds, grid size etc.) for our simulation. While these assumptions were certainly not made arbitrarily and are in line with previous research, we would like to discuss in this section whether our scheme generalizes to other settings or what needs to be adjusted when applying it to different scenarios.

Our spatio-temporal parameters were chosen as previous research suggested for ozone [103] (see above). This choice of course is dependent on the phenomenon (i.e. environmental parameter) that is sensed, respectively its homogeneity and dispersion behavior. Different pollutants or phenomena dictate a different choice of parameters. The same is true for the general environment: A city with street canyons may call for other proximity thresholds than an open area in nature.

An important prerequisite in this context is that the sensing system somehow should ensure that the same phenomenon is actually being measured in the first place and that measurement takes place under the same circumstances. If, for example, one sensor is used to measure the temperature in open sunlight and another in the shadow, that means they actually are not measuring the same parameter and of course this makes the readings incomparable. Another example would be the usage of air quality sensors in greatly different sampling contexts (e.g. standing vs. riding a motorcycle), in which the difference in speed could lead to an invalid air sample in the latter case. However, such problems need to be addressed at a different level, such as training of participants or outlier detection. Colocation and proper handling of measurement equipment could also be incentivized through the use of game elements [42].

Finally, the actual mobility patterns that are likely to be exhibited in the sensing scenario may differ from the ones used in our simulation. We used taxicab traces as basis because they reflect the movement of real people through a real urban environment. On the other hand, the authors are aware that if sensors were actually deployed on the taxicabs, privacy problems would probably be secondary. Still, the general properties of the mobility data should be realistic for the underlying scenario: Everyday people traversing the public spaces they live in.

All of these are aspects that need to be taken into account, both when designing a Participatory Environmental Sensing application and when determining the parameters for using our scheme to calibrate sensors within them. Nevertheless, we do not see that any of this would invalidate the general applicability of our scheme to different environmental sensing scenarios.

5.4 conclusion

In this chapter, we first showed the application of different calibration schemes from the literature on real-world data from laser-scattering fine dust sensors. Subsequently, we presented a novel privacy-protecting calibration scheme for participatory environmental sensing that combines collaborative blind macro-calibration with Private Proximity Testing, personalized exclusion zones, spatial generalization, pseudonymization and MIX networks. This enables the calibration of low-cost sensors based on rendezvous and the exchange of measurements between them. We evaluated our scheme on 22 days of simulated data, which combines real-world mobility traces with modeled calibration errors. The results show that our method is capable of achieving significant calibration gain even without reliable reference stations present and protecting the users' location privacy at the same time.

6 Human Factors

This chapter deals with the user in Participatory Environmental Sensing. So far, this thesis focused on the technology that enables low-cost, mobile, and distributed dust sensing. However, in Participatory Sensing, the human is actually an important part of the sensing architecture, with potentially large effects on data quality.

While the topic of this thesis is mobile Particulate Matter (PM) measurement, some of the studies in this chapter entail the measurement of other phenomena. Some studies were built on the scenario of noise pollution sensing with smartphones, since no additional hardware is needed for this type of application and this greatly facilitated conducting user studies at a larger scale.

A version of this chapter has been previously published in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)* [45]. The exploratory study on audio recording and annotation was conducted by Marcel Danz for his bachelor's thesis [66], pre-studies and the implementation of the app for the field study have been carried out by Julien Hoffmann . Parts of the statistical analyses have been conducted by Andrea Schankin.

6.1 introduction

Mobile and wearable devices – being always on, always with the user and context-sensitive – present a perfect platform for so-called Participatory Sensing [49]. Projects are highly diverse, ranging from plant observation over data processing (e.g. classifying and labeling data) to sensing environmental phenomena. An extensive survey of Participatory Sensing was presented by Christin et al. [64]. In contrast to the potential of such systems stand the many sources of systematic error that may affect data quality in mobile and wearable Participatory Sensing

[39]. That is why, as Bonney et al. point out, "*Despite the wealth of information emerging from citizen science projects, the practice is not universally accepted as a valid method of scientific investigation. [...] At the same time, opportunities to use citizen science to achieve positive outcomes for science and society are going unrealized*" [23]. This goes to the extreme that data from volunteers is considered undesirable by experts or policy makers and may even be prohibited for official use [95]. On the other hand, research shows that laypersons can collect data of comparable quality to experts, if properly familiarized with the task [206]. The problem is that non-experts are typically not, and thus cannot ensure standardized sensing processes. They may be:

- *Untrained:* unfamiliar with the way the sensing process is intended to be performed,
- *Overwhelmed:* uncomprehending or unable to recall the correct measurement procedure,
- *Inattentive:* not paying attention to all details of the process (esp. likely if participation is extrinsically driven, e.g. through monetary incentives or gamification),
- *Digital immigrants:* not digital natives, i.e. have little or no experience with mobile or wearable technology, or even
- *Malicious:* deliberately trying to influence the measurement process or "to play" the system.

We argue that, in the design of Participatory Sensing tools and applications, the user is still mostly regarded as someone who needs to be motivated and should ideally have a good time using it. Meanwhile it is neglected that in fact often she or he is also an important part of the technical sensing architecture, directly affecting the quality of the generated data or performed task. Harding et al. recently recognized that "*[...] this application domain is poorly understood by most system designers who focus almost exclusively on empowering citizens rather than considering the needs of both citizens and civic authorities and establishing trusted relationships between these stakeholders*" [102]. By designing for both adequate data quality and intelligibility, this trust relationship between users and authorities is strengthened.

This chapter and its contribution are divided into two parts: In the first part (section 6.3), we present a series of empirical field observations that we conducted to explore the variance in behavior that non-experts display in different Participatory Sensing settings. We categorize our observations, focusing the perspective of our analysis on the correct execution of the respective sensing process. Subsequently, we gather and categorize mechanisms that can be employed to prevent or mitigate this kind of adverse behavior (section 6.4). The collected knowledge can be used to guide the development of systems that help non-experts to perform measurement tasks more uniformly and to prevent certain mistakes, thereby increasing data quality.

In the second part (section 6.5), we present a large field study of an exemplary Participatory Sensing application. Four different designs ('app flavors') are compared to validate the effectiveness of the collected measures, discuss their interplay with user experience and illustrate the importance of making measures understandable to the user.

6.2 RELATED WORK

This section gives an overview of generally related work. An in-depth discussion of measures that can be employed to increase data quality is discussed in section 6.4 below.

As already shortly mentioned above, HARDING ET AL. recently recognized that the "*perceived value of civic crowdsourcing applications has remained low*' and that the design space is yet poorly understood [102]. However, their work focuses on engagement and the important trust relationship between different stakeholders, whereas this paper addresses the relationship between non-expert user behavior and data quality. Also centering on motivational aspects, specifically regarding online citizen science platforms, is the work by YADAV ET AL. [255].

Sensr [122] is one of the rare systems to guide the design of Participatory Sensing tools and deployments. It is a framework for authoring mobile data-collection tools for citizen science. The focus lies on facilitating the process of creating mobile applications for people without technical skills, e.g. through a visual programming environment. As such, it addresses novice programmers more than novice users. Their work also includes a short overview on approaches to ensure data quality. Other than that, individual authors occasionally include discussion on possible non-expert user errors in their work. KLAKEGG ET AL. for instance presented a mobile sensing system that enables end-users to perform portable Near Infrared Spectroscopy (NIRS) [127]. As specific challenges, the authors list ensuring the correct distance and angle to the measured object, an even surface, and environmental factors like avoiding stray light or interference. In our own previous work, we included an overview of different sources of systematic error in mobile sensing [39], one of them being non-expert users. Alagarai Sampath et al. researched how improving the presentation of a task to crowdworkers affects their performance [5], and Dey et al. presented a tool to support building intelligible context-aware systems by exposing the application logic to the user [74]. We included exploiting knowledge about the sensing context as a promising approach to mitigating some non-expert user errors in our discussions below. To the best of our knowledge, no work has yet comprehensively explored the dimensions of non-expert user errors in mobile sensing for citizen science.

Norman very early presented high-level design rules for computer systems based on human error [169]. While his guidelines are still applicable today and also valid for the design of Participatory Sensing, they are very generic. Among other methods, he proposes to "*Use analyses of people's performance in a variety of situations – but especially their errors – to construct an analysis of the appropriate form of human-machine interface that would optimize performance and minimize [. . .] error*". This is what this chapter explores for Participatory Sensing.

6.3 EMPIRICAL DESIGN SPACE EXPLORATION

In the beginning we had little more than the idea that people's sensing behavior is likely to be diverse, considering the fact that the concept of Participatory Sensing emphasizes distributed sensing by everyday users with their personal mobile devices in the public sphere [49], and scenarios generally aim at a large scale. In order to gather information about how much variance people display in their sensing behavior and to what extent "naïve" users possess an intuitive knowledge of different sensing tasks, we ran a series of small exploratory field studies that also serve as a baseline for the subsequent research.

Methodologically, we opted against methods like interviews since procedural knowledge is difficult to express verbally [89]. Instead, all studies were run in the field: We conducted three measurement studies and one assembly study, each representative for environmental Participatory Sensing. Subjects were merely put into the context ("*Imagine [. . .] How would you do that?*"), as the natural event of measuring with mobile devices is too rare for true ethnographic observation. No action options or restrictions were specified. The behavior was observed (live and partially additionally in video recordings) by an instructor present throughout the trial run. As participants for the measurement studies, passers-by were approached in urban public spaces (such as a park near a university campus, in the street, etc.). Volunteers were not paid.

To adequately explore the design space and capture the peculiarities of different Participatory Sensing tasks, we selected four different use cases for our empirical field research studies: The first two both required the users to record an audio signal with a smartphone, but differed in the source of the signal. For *Noise Level Monitoring*, participants were asked to take audio measurements with the goal of capturing the outdoor ambient noise level. *Audio Recording and Data Annotation* required users to generate an audio signal themselves, record it and finally annotate the recorded data with a ground-truth label. In the third use case, *Participatory Air Quality Sensing*, participants were asked to use the *iSPEX* camera clip-on module [217] for fine dust measurements. The first two settings both entail tasks that are seemingly simple and that we expected people to have a certain intuition for, even without explicit instructions on "correct" sensing behavior. The third use case was selected to observe behavior in presence of a more complex sensing task and according instructions. A fourth study covered the scenario of *Grassroots Sensing with DIY Hardware*. In this setting, we observed participants while assembling a Do-it-yourself (DIY) kit of a sensor station for citizen science air quality monitoring.

The focus of observation in the exploratory studies was the variance in exhibited behavior, specifically the kind that may have an adverse effect on data quality. We refer to this kind of behavior as *human error* in this paper, following NORMAN [169]: Human error both covers *mistakes* (errors in the intention) and *slips* (errors in carrying out the intention). We also explicitly include "mistakes" that are beyond the direct control of the user or caused because the user had incorrect or incomplete information on the task. However, we would like to stress that we do not imply that in these cases the user is to blame. Still, from a technical perspective, they remain errors. An overview of observed human errors is shown and discussed at the end of this section in Table 13, after the presentation of the exploratory studies.

6.3.1 *Exploratory Study 1: Noise Level Monitoring*

The underlying scenario for this study is smartphone-based noise pollution sensing. Multiple authors have built phone-based sensing systems for this use case in the past [119], [149], [164], [192], [203]. We selected this application case because it represents the task of measuring an environmental phenomenon with the internal sensors of a standard smartphone. The idea was to get an insight into the varying behavior that we expected to be exhibited for instance by standard users who download a crowdsensing app from some app store and just intuitively start recording. This use case is the same that also was explored in the final field study in part II of this work.

Figure 61.: Exploratory Study 1: We investigated non-expert user behavior for the use-case of *Noise Level Monitoring* with smartphones.

Participants and Task

Seven participants (four men and three women, ages ranging from 21 to 26 years) were asked to record audio samples representing environmental noise levels using the default *Apple iPhone* audio-recording app (see Figure 61). All participants were approached in the public sphere of a major city. First, they were given a short introduction into the concept of participatory noise pollution maps. Subsequently, participants were instructed to use the phone to make an audio recording that is representative of the ambient noise pollution level. The only specification on how to do this was the abstract instruction to do it in a way that they felt would yield the highest possible data quality. They were asked to complete a single recording and notify the observing instructor when they thought that they had successfully completed the task. By design, subjects were *not* instructed on correct or incorrect ways to perform the task, in order to not artificially narrow down the range of possible behavior.

Observations

We expected to observe human error in the measurement procedure, and participants' behavior indeed showed both large diversity and scale. As a baseline on what constitutes an error, we adopted the best practices for noise level monitoring from the PDF user guide of the *NoiseTube* project [149]. Of the seven participants, six moved the phone around while recording, causing audible wind-noise in the recording. Six participants held the device too close to their face, breathing into the microphone. Two participants inadvertently covered the microphone with their finger while holding the phone, leading either to muffled recordings or loud scratching noises. In the absence of a timing instruction, the measurement time participants deemed representative of the ambient noise pollution level varied greatly: The duration of the recordings ranged from a few seconds up to more than three minutes. Several users actively tracked noise sources, such as cars passing by, and one of them even tried to close in on noise sources by approaching a group of people that were talking and attempted to record the sound of their conversation from a close distance. One participant proceeded to record ambient noise levels without going outside first. Other observed influences were users audibly walking around, scratching themselves loudly and even absurd behavior like talking or whistling (out of boredom) while recording.

6.3.2 *Exploratory Study 2: Audio Recording and Annotation*

In the second empirical study, we explored a scenario involving data collection and annotation. This differs from exploratory study 1 in two key aspects: The scenario dealt with additionally performing an activity in contrast to just measuring, as well as sampling an object rather than an environmental phenomenon. This is a more complex setting which also presents more room for human error.

Participants and Task

Thirty-one students (22 men and 9 women, aged between 19 and 32 years) volunteered for this study. We simulated a data collection and labeling task in the following way: Participants were approached on a university campus and asked to shake a *Kinder Surprise* egg and record the produced sound. *Kinder Surprise* is a chocolate egg that contains a small toy inside, which may either be a collectible figure or something that requires assembly. "Experts" claim to be able to determine the type of surprise by the sound of the egg when shaking it. For the use case of building an automated classifier to detect the content of the egg, participants were handed an egg and asked to recording the shaking sound with their smartphone. The specific instruction was to shake the egg and to use their smartphone to record the sound for approximately five seconds and with as little as background noise as possible. Actually opening the egg and assigning a label to the recording was not part of the study.

Observations

While recording, only three participants shook the egg near the microphone, all others shook it somewhere behind or in front of the phone or at one of the sides (left or right). This observation points to an incorrect mental model of sound recording with a smartphone. Participants probably either related the sound recording function to the camera or assumed that the distance between sound source and microphone is irrelevant. Some participants additionally produced unwanted noise by loosely worn watches or bracelets on the arm they used to shake the egg.

Regarding the activity, grip and shaking technique have a large influence and determine the quality of the generated sound. For later classification, it is important that the procedure is ideally performed in the same defined way. The exploratory study

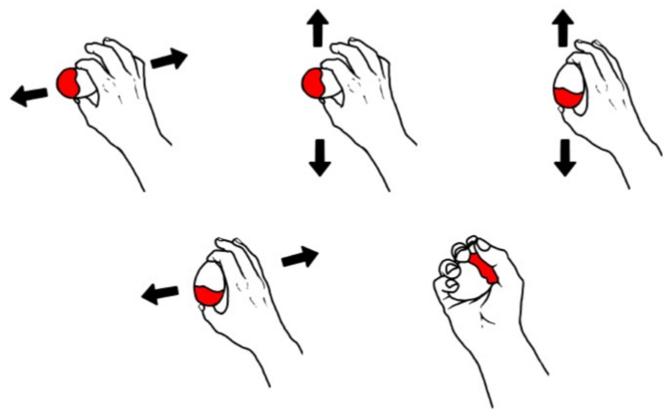


Figure 62.: Exploratory Study 2: In *Audio Recording and Annotation*, we observed a wide range of grip and shake variations.

shows that, without further instruction, participants hold and shake the egg very differently (see Figure 62b). Most participants held the egg at the long sides and shook it to the peaks (N=13) or perpendicular to the peak axis (N=6). Others held the egg at the peaks and shook it either to the peaks (N=6) or perpendicular to the peak axis (N=2). Finally, some participants held the egg inside their closed hand (N=4). Half of the participants held the egg with their dominant hand and the smartphone with the other hand, the other half did it the other way around. Although this is probably not relevant concerning the data quality, it may be of interest in other use cases and generally from the interaction design perspective.

Overall, the exploratory study shows high variance between participants, indicating that the quality of the recorded audio signal can not be guaranteed without more specific instruction.

6.3.3 *Exploratory Study 3: Participatory Air Quality Sensing*

The third exploratory study dealt with smartphone-based air quality sensing. Participants were given the *iSPEX* system [217] to measure the fine dust levels in the atmosphere. *iSPEX* is a passive spectropolarimetric add-on for smartphones that uses their camera to determine the levels of atmospheric particles by analyzing the polarization of the light when pointing the sensor add-on at a patch of blue sky.

The measurement process itself is quite intricate. Therefore, the *iSPEX* app includes in-app instructions and mechanisms to guide the user to correct measurements: The principle measurement process is explained both on a one-page paper manual as well as in a tutorial inside the app. Since it is required that the user orients himself so that the sun is in his back, the app calculates the position of the sun for the time of day and the user's location and uses a compass to point him or her in the right direction. The app also tries to detect whether the hardware module is installed and triggers an alert and prohibits measurement if it is not. For a correct measurement, the user slowly needs to raise his or her arm upwards. The app prompts the user in real-time to do this and plays a sound after correct execution to indicate success. All of these features were present in the exploratory study, as we used the standard app available from the Apple App Store. Because of these measures, sometimes users made errors during the study before the app detected and prevented them and eventually enabled them to correctly follow the measurement procedure. As we are interested in exploring the range of possible behavior, we still included these errors in our recordings and discussion at the end of this section.

We selected this use case because it represents a smartphone-based measurement task for which users are likely to have no intuition at all and that requires additional unfamiliar hardware. Other examples for complex sensing tasks tasks like this are e.g. smartphone-based portable near infrared spectroscopy (NIRS) [127] or light-scattering particle measurements with cameraphones [30].

Participants and Task

In this study, ten pedestrians (seven men and three women, aged 19 to 28, were recruited in a inner-city park close to a European university campus. For the study, they were given an *iPhone 5s* with the most recent app version installed (last updated: Oct, 2015) as well as an *iSPEX* module, complete in box, including the sensor module, an adapter for the *iPhone 5s* and a quick manual (see Figure 63). The purpose of measuring fine dust using a camera smartphone was shortly introduced to them and they were then asked to use the phone (respectively the app) and the sensor add-on to perform measurements. No specific instructions, e.g. on how to install the add-on were given. All participants owned a smartphone themselves, three of them an *iPhone* and the rest *Android* phones. While the *iSPEX* system only exists for the *iPhone* and thus there was no alternative, our observations do not suggest any impact of providing participants with an unfamiliar phone.

Figure 63.: Exploratory study 3: *Participatory Air Quality Sensing* using the *iSPEX* clip-on module [217] for the *iPhone 5s*.

Observations

Six participants immediately hit the "start measurement" button after starting the app, without having installed the *iSPEX* add-on or reading the instructions. Three participants then still left the module in the box at first and tried to take measurements with the app without the hardware add-on being installed, one of them even over and over again for several minutes. Conversely, one user carefully read the included paper manual, installed the module correctly and then started performing the correct procedure, but without having pressed the "start measurement" button first.

All participants encountered problems when attaching the *iSPEX* module to the phone, as they first had to attach the separately packed *iPhone 5s* adapter, and it took a lot of time for them to get it to fit right. Still, for some of them, the app even then wrongly kept displaying the alert "*add-on missing*", preventing them from taking a measurement. One user installed the add-on module facing the wrong way. After receiving an alert, the user corrected this. Another user at first installed the hardware add-on at the bottom side of the iPhone over the microphone, because he read in the manual that he should enable the sound on his phone and therefore thought that the microphone or speaker was used to make the measurements. After reading the in-app tutorial, the user corrected his mistake. The subjects generally had trouble understanding what kind of arm movement was expected from them and that they should continue to raise their arm until the phone was over their head.

Apart from the above errors that were eventually prevented by the mechanisms of the *iSPEX* app, we observed several errors that the app was not able to catch: Of the ten participants, seven tried to take measurements even though no sufficiently large patch of blue sky was visible, five were too close to trees or buildings for a valid measurement. Three participants measured while being seated, without orienting themselves away from the sun.

Overall, a noteworthy observation is that even though the app successfully prevented many types of misbehavior quite reliably by interrupting the measurement attempt, it did not inform the user concerning the reason for the interruption. As a result, three of the ten participants eventually became frustrated and aborted the measurement attempt. As main problem with the use of the app they spontaneously reported the lack of specific feedback (participant #7: "*The app does not tell me what I am doing wrong!*"). Regardless of the source of errors, alerts almost exclusively contained the message "*add-on missing*".

6.3.4 *Exploratory Study 4: Grassroots Sensing with DIY Hardware*

The underlying scenario for this study is grassroots environmental monitoring. Around the world, we increasingly witness examples of sensing campaigns that are driven by activists. *Hackspaces* and *Fab Labs*, along with according project descriptions that are widely available over the Internet, have enabled citizens to build and operate sensor stations who could not have done this before. A real-world example for this is the DIY fine dust sensor by the so-called *OK Lab* of the *Open Knowledge Foundation (OK) Germany* in Stuttgart, Germany, a nonprofit organization that advocates open knowledge, open data, transparency, and civil participation. The *OK Lab* provides an online manual[1] that explains the assembly and installment of a sensor station. They also operate and maintain a server to which measurements can be uploaded and an online platform which visualizes the data.

Participants and Task

In this exploratory study, no instructions were given at all. Nine participants (five men and four women, ages ranging from 21 to 57 years) agreed to being observed while assembling the DIY

1 Assembly instructions for the do-it-yourself (DIY) fine dust sensor node: `http://luftdaten.info/feinstaubsensor-bauen/`.

sensing kit of the *OK Lab*. All participants worked at a local newspaper, most of them as journalists, one as technical staff. They chose to build the DIY kit on their own accord. Prior to the observation, they had as a group ordered all parts necessary for the assembly, as listed on the project's website: 10 pcs. each of the WiFi enabled *NodeMCU ESP8266* board, the *SDS011* dust sensor, and the *DHT22* temperature/humidity sensor, as well as some wires, a USB power supply, plastic tubing and a piece of flexible hose. Subsequently, they chose to assemble the individual devices in a group session (see Figure 64a), as explained on the project's website and a Frequently Asked Questions (FAQ) video.

Assembling the sensor station required connecting seven wires to the appropriate pins, connecting the flexible hose to the air inlet of the sensor, flashing the firmware onto the NodeMCU from a shell, and finally installing the resulting system into the plastic tubing (see Figure 64b).

Observations

When assembling the sensor, each of the participants worked by himself on one sensor kit. Three of the participants at least partially failed to connect the wires to the correct sockets. In four of the assembled sensors, individual wires had slipped out of the sockets. Two of the participants were not able to build the kit by themselves and eventually asked others in the group for help. Three participants complained the manual being unstructured or even missing steps. To verify that they had successfully assembled the DIY kit, participants connected the completed kits to a power outlet and listened whether the dust sensor's fan started to make a light noise. However, this test did not prevent further error: One participant did not succeed in flashing the firmware without noticing, leaving the stock firmware on the device. Inserting the finalized sensor into the plastic tubing housing went smoothly for all but one participant, who experienced this as being difficult. None of the participants shortened the piece of flexible hose that was connected to the air inlet of the optical dust sensor, which might enable stray light to be reflected into the measurement chamber. When trying to register the sensors in the local WiFi, three of the sensors did not advertise their Service Set Identifier (SSID) as described in the instructions and as needed to finalize the configuration. Of the five sensors that were successfully registered, only two included all of the sensor data in their communication. The

(a)

(b)

Figure 64.: Exploratory Study 4: Nine participants (a) assembled a do-it-yourself (DIY) sensor station for air quality monitoring and (b) prepared it for installation for the use case of *Grassroots Sensing with DIY Hardware*.

other three transmitted empty values for either the dust sensor or the temperature sensor, probably due to bad connections or cable breaks.

Since some of the observed errors (and the underlying causes) are not directly evident to the user, after our study, we discussed our findings with an organizer of the *OK Lab*. He reported that they perform regular sanity checks at the back-end, especially on newly registered sensors, as they also observed missing data or that the connectivity of sensors sometimes varies. This apparently happens mostly either because people install them in an inappropriate place outside of the range of their own WiFi or when users do not maintain proper operating conditions, e.g. by switching off their WiFi over night to save energy. To be able to give the users feedback on this, the *OK Lab* has started to transmit the WiFi signal strength along with the sensor data. While the *SDS011* dust sensor comes pre-calibrated and user calibration of the DIY station is not intended in the project, long-term data on stability of the sensor is not yet available and (re-)calibration or sensor replacement may be required [47].

6.3.5 *Analysis of Observed Human Error*

The exploratory studies revealed many ways in which participants exhibited human error (both slips and mistakes [169]) in the measurement process, even in seemingly elementary tasks. We collected all our observations in Table 13. The table includes behavior which arguably may not strictly be erroneous per se, but which varied strongly between participants, suggesting a potentially significant effect on the resulting measurement.

Subsequently, we grouped similar instances of behavior into more abstract types of errors in smartphone-based mobile measurements and finally defined six dimensions of human error to form a taxonomy (see Figure 65).

Hardware

This aspect both concerns the employed smartphone (or other personal mobile devices) as well as potentially any other hardware, active or passive, that may be required for the sensing task. In general, one can assume that users will be most comfortable with their own device and that unfamiliar platforms and especially add-ons and external devices are more likely to promote erroneous behavior. This is especially likely if they include intricate assembly and/or maintenance.

Device Handling

This dimension regards the handling of the employed device(s). Requirements may range from virtually non-existing for robust

	Type	Example Behavior	Study
Hardware	Faulty installation	clip-on module missing	3
		clip-on module incorrectly installed	3
	Incomplete / wrong assembly	incorrectly built DIY kit	4
	Undocumented / surplus parts	polarization foil	3
		incomplete manual	4
	Missing maintenance	operating conditions not ensured	4
	No / faulty calibration	un- or decalibrated sensors	4
	Faulty device association	sensing device not paired, loss of data	4
Device handling	Wrong orientation	user turns around own axis while measuring	1
		microphone pointing in the wrong direction	1
		moving device in wrong angle	3
	Wrong height or distance	microphone too close to face	1
		arm not extended	1
	Unwanted device movement	user noisily moves device around	1
		user shakes device instead of (or along with) egg	2
	Covering sensor	user has finger on microphone	1
User Activity	Generating noise	talking, whistling	1
		coughing, clearing one's throat, breathing noisily	1
		scratching, clapping	1
		unwanted noise (watches rattling)	2
	Unwanted user movement	walking around	1
	Phone use	making a call or texting	1
Measurement/ Observation	Wrong sample properties	recording too short	1
	Wrong amount of samples	only one measurement instead of two	3
	Not actually sensing / recording	user forgot to press recording button	3
	False or no annotation	wrong label assigned by mistake	2
		no position data, as GPS disabled / no signal	1
Object/ Phenomenon	Wrong object handling	wrong grip or shake	2
	Wrong object/phenomenon	no blue skies visible	3
	Wrong alignment	shaking egg nowhere near microphone	2
		user follows noise source and tries to get very close	1
Environment/ Context	Inappropriate weather	noisy wind in audio recording	1
		no blue skies visible	3
	Indoors instead of outdoors (or v.v.)	attempting to measure ambient noise levels indoors	1
	Environmental disturbances	stray light entering sensor	4
	Wrong time and/or place	measuring at wrong location or point in time	4

Table 13.: Non-expert behavior from exploratory studies (study 1: *Noise Level Monitoring*, study 2: *Audio Recording and Annotation*, study 3: *Participatory Air Quality Sensing*, study 4: *Grassroots Sensing with DIY Hardware*).

tasks (e.g. recording and submitting textual observations) to tightly constrained procedures that the user is unfamiliar with and/or need to be followed precisely in order to collect meaningful data.

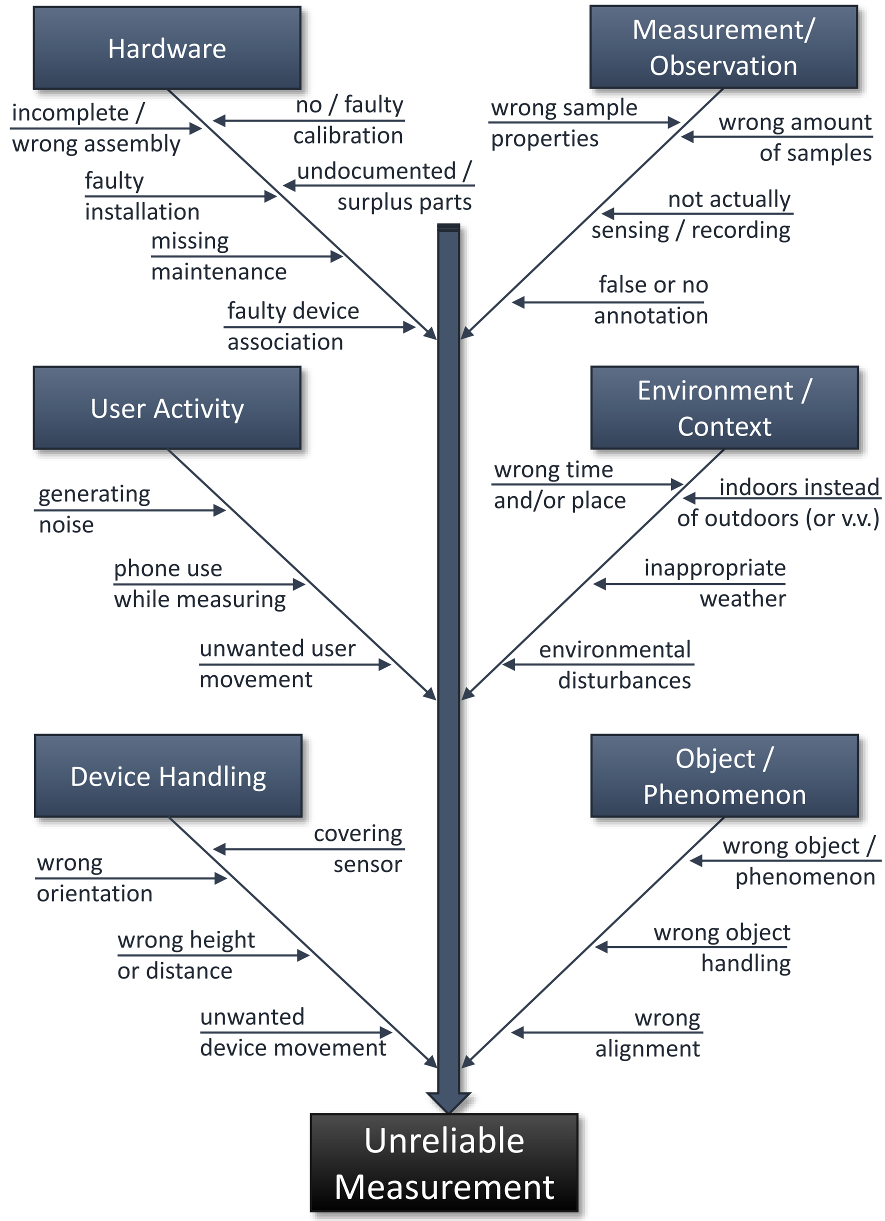


Figure 65.: Ishikawa diagram of the identified dimensions through which users may affect the quality of the measurement result.

User Activity

The next dimension concerns any behavior of the user that is unrelated to the measurement process and may still affect it, potentially reducing data quality. Mostly, this covers unwanted physical activity and the like. As with device handling, there may be tight constraints regarding this dimension or none at all, depending on the sensing task.

Measurement/Observation

This dimension concerns requirements regarding the recorded observation. Such requirements for high quality data may range from completely free observations to tightly defined constraints, e.g. regarding sample size, annotation requirements, synchronization of different readings etc. The fewer constraints are defined, the more diverse data will probably be collected across participants. Generally, this will likely make comparison and/or data fusion more difficult.

Object/Phenomenon

This dimension covers any requirements regarding the phenomenon or an object that is at the center of the observation. Constraints may range from basically none to looking at precisely defined aspect of a specific object. The boundaries between this dimension and the handling of the device may overlap, e.g. when a certain alignment between device and object is required.

Environment/Context

The last dimension concerns the environmental context[2] of the user. Measurements may be robust to external factors and be allowed anywhere and anytime or again, tightly constrained and well-defined.

6.4 ENHANCING DATA QUALITY

After having analyzed and compiled ways in which participants' behavior may adversely affect data quality in Participatory Sensing, we look at ways to prevent the undesired behavior or mitigate its effects in this section. Research on mobile sensing recognizes the need to ensure viable readings from low-cost sensors [135], even though the focus is seldomly placed on the effects that are caused by non-expert users. Many general approaches exist in the literature, some of which are applicable to typical Participatory Sensing scenarios and some of which are not. Whether or not a technique is suitable or not depends

2 We are aware that in Ubiquitous Computing, the term *Context* by itself usually includes aspects that are already covered in other dimensions here, e.g. activity. Context has i.a. been defined as "*any information that can be used to characterize the situation of an entity*" [73]. In contrast, environmental context here is meant more narrow.

on a variety of aspects, such as the specifics of the task at hand, the scale of the deployment, etc. This section discusses classes of possible countermeasures gathered from literature review. Table 14 summarizes the results.

Participant Selection is an approach that has been used in different fields to identify and separate suitable personnel from such that is unfit for a task. Wickens et al. review different methods of identifying people who are likely to perform successfully along with different measures of ability, albeit with a focus on assigning people to jobs [245]. However, by definition, Participatory Sensing addresses everyday users, which may make pre-selection an undesirable step. Additionally, at larger scale, screening may become prohibitively expensive.

The most intuitive approach to ensure that users perform a task correctly is *training* [144], [206], [210]. Thelen et al. reported that "*numerous studies have demonstrated that volunteers can successfully perform basic data collection tasks when given a half day or more of practical field training*." [228]. This highlights the biggest drawback of training sessions: A lot of resources (experts, facilities, etc.) are needed and the approach does not scale. Slightly different forms of training that do not require the user to keep a mental model of the process are *instructions* (e.g. manuals or tutorials). The key difference to training is that instructions are typically given in writing or otherwise fixed form (video, etc.) which is used to make the non-experts to understand the measurement process. Understanding is defined as the ability to hold and process all elements that define the measurement process simultaneously in working memory [222]. However, working memory is extremely limited in capacity [157] and in duration [183], in particular for novel information that needs to be processed in a novel way [222]. That is, people might fail to understand or completely recall new material if it is sufficiently complex, as may be the case in Participatory Sensing. As we have shown in our exploratory studies, even a seemingly easy task like recording an audio signal involves a complex measurement process for the user. Also, pure manuals are of little help, as people tend to not read them [173], especially if they do not encounter problems, as would be the case in a badly but successfully performed measurement process. On the other hand, instructions can be given much "closer" to the actual task (spatially and temporally). In shorter form and *in-situ*, instructions provide an advantageous approach, up to providing a step-by-step walk-through.

Another popular approach are *reputation* systems [113]. There are different flavors, ranging from picking users based on their reputation or skill level [193] (cmp. *selection* above) over assessing it beforehand [230] to building it through data analytics. However, this again requires some kind of ground truth determined by expert users or a series of campaigns, making it an intricate option. In Participatory Sensing systems, individual readings can often not be re-evaluated and the classification of them as being correct or wrong after-the-fact is often infeasible, making reputation levels difficult to build. Additionally, "ranking users can backfire" [126], influence the participants' motivation, and paradoxically lead to the best performing participants quitting, as they would feel they had "won the game".

Verification of data entries is another approach to increase data quality. Gardliner et al. [94] differentiate between verified and direct citizen science. Entry verification can either be approached automatically by using some sort of computational recognition or simple sanity tests. The advantage of this sort of verification is, that it can be performed in real-time and the user can be prompted before leaving the area, as proposed by Burke et al. (*"Did you really just see 40 diesel trucks go by in five minutes?"*) [49]. In community-based data validation [249], instead of revisiting their own data, participants verify data from their peers. Another form of verification are *expert reviews* [122] of data. They have the disadvantage that data has already been collected and can only be discarded, as the analysis takes place after-the-fact.

Computational approaches are diverse. The simplest ones are of a statistical nature: *redundancy* [122] and/or *repetition* [206] both lead to multiple instances of the same data which can then in turn be processed, e.g. to remove outliers. These approaches only work, if the underlying assumption holds that the overall error is non-systematic, i.e. people will on average perform the task correctly. However, as we have seen in our exploratory studies, there are certain errors which the majority of people tend to make. More sophisticated approaches like *outlier/anomaly detection* or *Bayes filtering* take the structure of the data into account. The drawbacks of filtering out anomalies is that the smoothing makes the approach less suitable for highly dynamic phenomena. If only few data points are available or no model can be constructed, filtering is also not applicable.

A different way of computationally addressing procedural errors is *context recognition*. Mechanisms may be as simple as detecting whether the GPS receiver is turned on or the acceleration sensors of a device pick up movement when there should be none, to integrating full-fledged activity recognition. A robust way to deal with different types of error afflicted data is signal *reconstruction* from noise [39]. However, it is not generally applicable, as the measured phenomenon must be modelable as particles, among other constraints.

An interesting approach is *data design*, i.e. using HCI methods not only to design interfaces, but also to assess the needs of data consumers to collect reliable, standardized and overall more useful data [124]. However, this works for observations that are reported in a free form (e.g. textual), but not so much for pure sensor data. Additionally, models have been developed to exchange, revise and merge structured offline data, e.g. from contributions that are accomplished via paper [211].

Finally, one of the most universal mechanisms is *feedback*. Since we assume that the user actually is interested in collecting and submitting high-quality data[3], it is important to make the measurement process as transparent as possible. Feedback (e.g. on the correct execution of a step, etc.) can greatly contribute to this understanding.

Some of the discussed approaches can be combined with *Gamification* techniques. In this way, the location-based game *GeoSnake* [156] has been used to boost verification rates, the game *PhotoCity* gamifies training [235] and it has been proposed to use game contexts to ensure correct execution of a sensing task [42]. This approach is discussed in-depth in chapter 7.

3 We disregard malicious users here, as we are convinced that someone determined to willingly submitting false data will find a way to do so.

Measure	Advantages	Disadvantages
Participant Selection	domain expertise / prior knowledge	usually requires thorough analysis, high resource cost; selection success hard to verify
Training	best / covers everything; trainer can assess success	high resource cost, experts needed continuously
Instruction (manual)	clear	mentally demanding; passive access
Instruction (in-situ)	very clear; temporally close	requires some sort of display
Reputation	good for sorting out single users; helps against malicious intent	can de-motivate users; may be hard to build
Verification	ensures data quality	infeasible for many tasks; after-the-fact;
Expert Reviews	ensures data quality	infeasible for many tasks; experts needed continuously; after-the-fact
Redundancy	very simple	only robust against statistical error, not systematic; may be difficult to achieve; bad readings still contribute
Outlier/ Anomaly Detection	eliminates implausible readings	prevents capturing "true" anomalies
Bayes Filter	adapts to available data	needs basic models and multiple readings
Repetition	simple	only robust against statistical error, not systematic; needs to be triggered somehow
Context Recognition	potentially fine grained control over measurements	only certain class of errors; maybe technically difficult
Reconstruction	extremely robust	only applicable in special cases
Data Design	provides structure	mostly for textual data
Feedback	supports user in verifying correct procedure himself; almost always possible	may overwhelm or frustrate user if not carefully balanced
Gamification	motivates; increases hedonic quality; may enhance measurement frequency	may distract from sensing task: can demotivate intrinsically motivated participants

Table 14.: Overview of possible measures to improve the data quality in mobile non-expert sensing.

6.5 FIELD STUDY

The previous observation studies revealed a surprising diversity in user behavior, which is likely to yield a high variance in data quality. Thus, it is important not only to unify the measurement process per se but also to guide the user behavior, in particular in Participatory Sensing. As already discussed in section 6.4, there are two main strategies to achieve this goal. First, users could be trained or instructed to show the required behavior or, second, the correct user behavior could be supported by technical measures. At the same time, user experience is important in order to keep the user motivated to participate in citizen science.

We argue that implementing purely technical measures to prevent certain errors may increase data quality but at the same time may have adverse effects on the general user experience. Conversely, focusing purely sensing on ease of use and an understandable process may still result in users keeping to make certain kinds of errors. We hypothesized that the combination of both will lead to substantially better systems. To test this hypothesis, four functional variants (flavors) of a mobile sound recording app were developed. The four app flavors mainly differed in (a) the way participants were instructed how to record an audio signal properly and (b) the technical measures supporting correct recording (see below).

6.5.1 *Participants*

A total of 535 passing-by pedestrians were recruited to volunteer in the study (opportunity sample; 209 women, 321 men; 5 missing values, mean age: 30 years, age range: 18-76 years). The experiment was conducted face-to-face in field (e.g. on the street or in parks). The overall education level of the participants was rather high. About one third had a university degree. About half of the participants ($N = 267$) had a technical or natural scientific background, the others had a commercial ($N = 80$), medical ($N = 24$), juristic ($N = 7$), pedagogic ($N = 53$), administrative ($N = 15$), manual ($N = 22$) or artistic ($N = 18$) education. Forty-nine participants worked in other branches ($N = 41$) or did not indicate their line of business ($N = 8$).

All participants were well familiar with using mobile phones, most of them (96.6%) also using a smartphone. Almost all participants had used their smartphone to call someone (94.2%) or had sent messages (93.6%) before. Most participants had also in the past sent emails (76.6%) or taken pictures (81.7%). More important in the context of the study, only fewer than half of the participants had used their smartphone to record videos (47.5%), speech (26.0%) or music (16.4%). None of the participants were experts in Participatory Sensing, only 18 of them (3.3%) had already participated in a previous Participatory Sensing study. Overall, the sample was heterogeneous with regard to age and education and representative for app users.

All participants gave their written informed consent and did not receive any compensation for their participation. The study was carried out in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

6.5.2 *Material, Study Design, and Procedure*

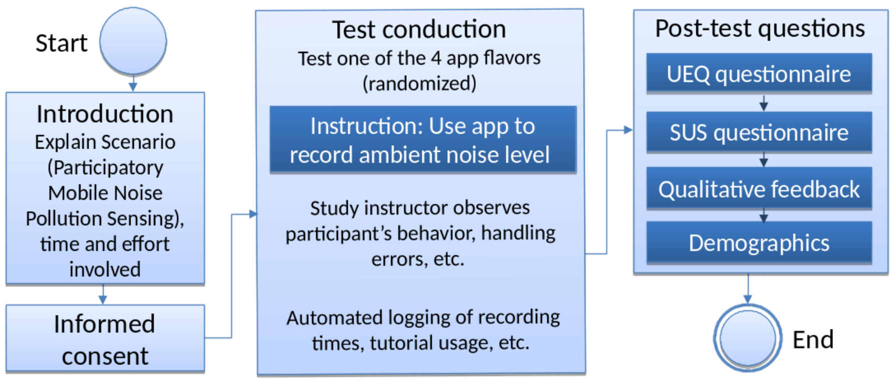


Figure 66.: Session structure of the field study.

Procedure

At the beginning of the experiment, participants were introduced to the scenario. They were told to imagine that they were part of a community that tries to build a noise map of their city. For that purpose, they would use an app on a smartphone to record the ambient noise level of their environment with the goal to achieve data quality as high as possible. Participants were then asked to record ambient noise levels with a quality as good as possible. They were given one of the recording apps, which they could use as long as they wanted and undertake so many trials until they felt that they recorded a good signal. No further instruction about the measurement process or the handling of the app was given. After completing the task, they filled in the User Experience Questionnaire (UEQ) and the System Usability Scale (SUS) and gave qualitative feedback to evaluate the app. In addition, they answered some questions about their smartphone usage behavior as well as a few questions concerning demographics and their habits regarding technology use. The whole experiment took about 15-20 minutes per participant. An overview of the session structure is shown in Figure 66.

App Flavors

Four functional variants (flavors) of a mobile sound recording app were developed (see screenshots in Figure 67). The four app flavors mainly differed in (a) the way participants were instructed how to record an audio signal properly and (b) the technical measures supporting correct recording (see Table 15). App 1 (*Basic*) was a simple one-button app that only allowed starting and stopping a recording. In app 2 (*Basic+*) a short tutorial at the beginning provided detailed instructions of how to avoid erroneous behavior while recording an audio signal and how to use the app. The presented best practices included, for example, instructions to point the microphone of their smartphone away from their body and avoid shaking it.

	Basic	*Basic+*	*Premium*	*Premium+*
Animated Button	×	×	×	×
Time Display	×	×	×	×
Button Vibration	×	×	×	×
In-App-Tutorial		×		×
In-App-Manual		×		×
Flipped Interface			×	×
Check: GPS on?			×	×
Check: Orientation?			×	×
Check: No shaking?			×	×
Check: Rec. time?			×	×
Alerts			×	×
	baseline	*instruction*	*technology*	*both*

Table 15.: The four different experimental conditions, i.e. app flavors: A simple 1-button recording app (*Basic*) serves as baseline for our study. The *Premium* flavor features multiple measures to improve data quality, including sensor-based verification of parts of the sensing context. *Basic+* and *Premium+* additionally display an in-app tutorial (see also Figure 67).

(a) *Basic* / *Basic+* (b) *Premium* / *Premium+*

(c) Tutorial: best practices (d) Tutorial: app manual

Figure 67.: Screenshots: The simple *Basic* flavors only feature an animated recording button and a time display (a). In *Premium* flavors, recording is disabled unless certain constraints are met (b). Additionally, the *Basic+* and *Premium+* versions each feature an in-app-tutorial explaining best sensing practices (c) and the app (d).

App 3 (*Premium*) helped avoiding common errors by providing feedback such as status indicators, e.g. when no GPS signal was available or error messages when the user shook their smartphone too strongly (see Figure 67). By flipping the orientation of the display upside-down, the user was forced to rotate the smartphone so that the microphone pointed away from the body. In app 4 (*Premium+*) the same short tutorial was included as in app 2. Aside from the sound sample, all app flavors automatically recorded certain events (tutorial usage, recording times, etc.) in a logfile for evaluation.

Study Design

The study consisted of four experimental conditions, i.e. the four app flavors. Test conditions were evenly assigned to participants in a between-subject design, i.e. each participant used only one of the four app flavors (Basic: $N = 123$; Basic+: $N = 137$; Premium: $N = 130$; Premium+: $N = 145$).

Collected Data

We captured (i) the number and types of errors that users made with different variants of the citizen science app, as well as (ii) the user experience while performing the sensing task. User errors were recorded in categories by the study instructor in an observation protocol, following the previously identified error dimensions (cmp. Figure 65 above). Erroneous behavior was defined by the same best practices as in our exploratory studies. User experience and usability were measured with the German versions of the User Experience Questionnaire (UEQ) [136] and the System Usability Scale (SUS) [28]. The SUS is a standardized questionnaire with ten short questions that primarily measures the usability aspect of a product. The UEQ is supposed to measure user experience in a wider scope and consists of 26 bipolar items that are assigned to the six scales *Attractivenes*, *Perspicuity*, *Efficiency*, *Dependability*, *Stimulation*, and *Novelty*. In addition, qualitative feedback was collected.

6.5.3 *Data Analysis*

Data was analyzed by comparing only those app flavors with each other which were of interest regarding our study goals: (i) To analyze the effect of instruction (i.e. in-app tutorial) alone, *Basic* and *Basic+* were compared; (ii) the effect of technical measures alone (without explanation), was evaluated by comparing *Basic* and *Premium*; (iii) to explore the difference between the effects of instruction and technical measures, *Basic+* (instruction only) and *Premium* (technical measures only) were compared; and (iv) *Basic* and *Premium+* were compared to show the complementary effect of instruction and technical measures.

Statistically, four independent t tests were computed, separately for usability and user behavior data.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1 X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad S_{X_1 X_2} = \sqrt{\frac{(n_1 - 1)S_{X_1}^2 \cdot (n_2 - 1)S_{X_2}^2}{(n_1 - 1) + (n_2 - 1)}} \tag{24}$$

Because of multiple comparisons, p values were adjusted according to BONFERRONI. For this correction, the alpha level (usually $p < .05$) is divided by the total number of pairwise comparisons. Then each of the p values are compared to that shrunken value of alpha. To improve the readability of the text, we did not reduce the level of significance but report Bonferroni adjusted p values that allow a direct comparison to a level of significance of $p < .05$. This is common practice, for example, by statistics software such as SPSS. SPSS multiplies each of the actual p values by the total number of possible pairs. That is, a test can be considered statistically significant if the reported $p < .05$. The descriptive results are presented in Figure 68 (user behavior during the measurement process) and Table 16 (usability and user experience).

Qualitative data was analyzed with a content analysis. We started with small clusters of semantically related comments which were further grouped on the basis on the categories of the DIN EN ISO 9241/110 (software ergonomics). Some clusters were related to very specific aspects of the app (e.g. the tutorial). On the most abstract level, we grouped the clusters based on pragmatic quality (perspicuity, efficiency, and dependability), hedonic quality (novelty and stimulation), and attractiveness in order to compare them to the results measured with the UEQ. It is important to note that the apps were developed with the objective to observe the measurement process, i.e. the functionality of the app was limited[4]. This was also noticed

4 Specifically, after sensing, there was no visualization of the measurement result in a map, as dB value, or the like. This was by design, as we focused

by the participants. While these comments were omitted from analyses, they do explain the low ratings for all four app flavors in some of the categories of the UEQ.

6.5.4 *Study Results*

The study results are reported sorted by (i) the effect of instruction alone (*Basic* vs. *Basic+*); (ii) the effect of technical measures alone (*Basic* vs. *Premium*); (iii) the difference between the effects of instruction and technical measures (*Basic+* (instruction alone) vs. *Premium* (technical measures alone)); and (iv) the complementary effect of instruction and technical measures (*Basic* vs. *Premium+*). A summary is given at the end of this section.

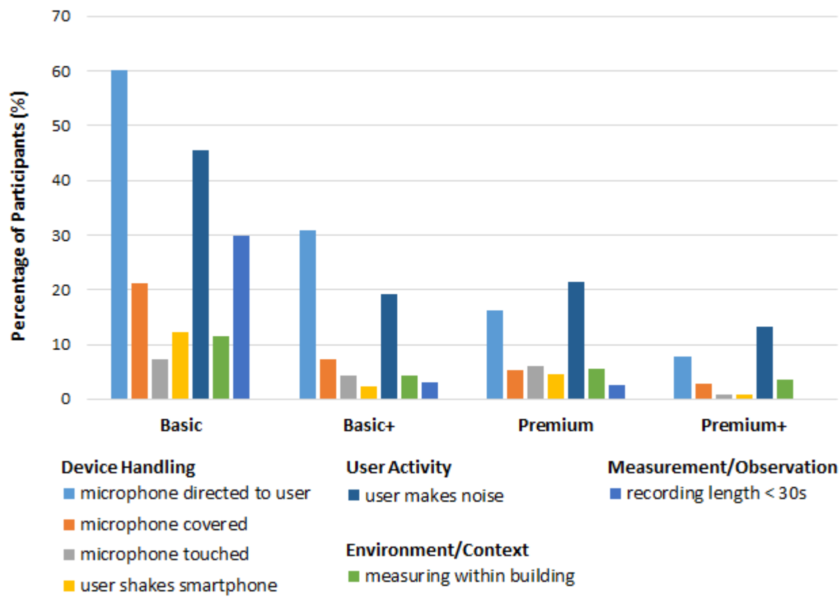


Figure 68.: Measurement errors observed in the four different test conditions (i.e. app flavors).

Effect of Instruction Only

Adding a tutorial (*Basic+*) to a simple one-button app (*Basic*) improved the measurement process in terms of errors made. In particular, the microphone was less frequently directed to the user, $t(258) = 4.845$, $p < .001$, or being covered, $t(258) = 3.280$, $p = .008$; users shook the smartphone less often, $t(258) = 3.224$,

on measurement error and did not want to solicit qualitative feedback on the visualization in this study.

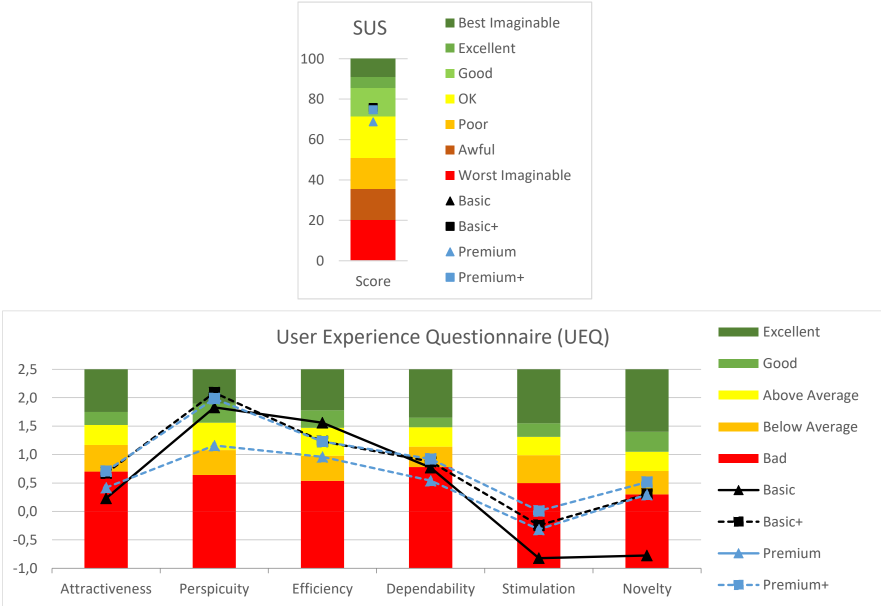


Figure 69.: Results from the SUS and UEQ questionnaires.

$p = .008$; the required recording duration was achieved more frequently, $t(243) = 6.126$, $p < .001$; and the user produced less interfering noise, $t(258) = 4.781$, $p < .001$. However, despite the general reduction in errors, users still made some errors.

Adding a tutorial (*Basic+*) that described the measurement process did not change the usability of the app. The SUS score remained the same, $t(246) = 0.066$, $p = 1.000$, as well as perspicuity, $t(258) = 2.225$, $p = .112$, efficiency, $t(258) = 1.840$, $p = .268$, and dependability, $t(258) = 1.259$, $p = .836$, as measurement of pragmatic aspects of user experience. The perceived hedonic quality increased, i.e. stimulation, $t(258) = 4.764$, $p < .001$, and novelty, $t(258) = 7.466$, $p < .001$. Overall, an app with a tutorial was perceived as being more attractive than without it, $t(258) = 3.713$, $p < .001$.

The qualitative data allowed a more detailed understanding of these results. With regard to the pragmatic quality, most users mentioned that both one-button apps were easy and fast to use and it was easy to understand how to use it. The observed change in hedonic quality was also reflected in the qualitative data. While 14 users mentioned that using the basic app was boring or uninteresting, this number reduced to 5 when adding a tutorial. More important, 15 participants even rated the app as being innovative or that they felt fun while recording the audio signal. The low value in attractiveness was probably due to the

	SUS	Att	Per	Eff	Dep	Sti	Nov	
Basic	76	0.2	1.8	1.5	0.8	-0.9	-0.9	*(baseline)*
Basic+	76	0.7	2.1	1.3	0.9	-0.1	0.3	*(instruction)*
Premium	**69**	0.4	**1.2**	1.0	0.5	-0.2	0.3	*(technology)*
Premium+	75	0.7	2.0	1.2	1.0	0.1	0.6	*(both)*

Table 16.: Results from the SUS and UEQ questionnaires (Att = attractiveness; Per = perspicuity; Eff = efficiency; Dep = dependability; Sti = stimulation; Nov = novelty). While technical measures alone (*Premium*) reduce user errors in a similar way as instructions alone (*Basic+*, see Figure 68), the perceived usability is lower.

overall look of the app. Although most participants liked the clear and simple design, it also appeared empty to them.

Effect of Technical Measures Only

The technical implementations were also able to improve user behavior: The microphone was less frequently directed to the user, $t(250) = 8.227$; $p < .001$, or covered, $t(250) = 3.785$, $p < .001$. The recording duration was less frequently below 30 sec, $t(234) = 6.143$, $p < .001$; and the user produced less interfering noise, $t(250) = 4.110$, $p < .001$. In this case, the usability decreased: The SUS score was significantly lower, $t(242) = 3.022$, $p = .012$, and also two of the UEQ scales that are supposed to measure the perceived pragmatic quality of a product showed reduced values, i.e. perspicuity, $t(250) = 4.151$, $p < .001$, and efficiency, $t(250) = 4.526$, $p < .001$; dependability did not change, $t(250) = 1.855$, $p = .264$.

The qualitative data show that participants liked the responsive design. They mentioned that the technical measures helped them avoiding measurement errors. However, without an explanation, not all participants did understand how to use the technical measures correctly. Although participants were able to complete the task, the task itself (e.g., the goal or the action steps) often remained unclear.

Difference Between Instruction and Technical Measures Only

The behavior when using technical measures to guide the measurement process was comparable to that observed when a tutorial was provided, all $t(264) < 1.0$, $p = 1.0$, except for the direction of the microphone. Here, the technical implementation

	Variable	*Basic* vs. *Basic+* Change	*Basic* vs. *Basic+* p value	*Basic* vs. *Premium* Change	*Basic* vs. *Premium* p value
User Experience	SUS	0.1	1.0	-6.4	.012
	UEQ Attractiveness	0.5	<.001	0.3	.192
	UEQ Perspicuity	0.3	.112	-0.6	<.001
	UEQ Efficiency	-0.2	.268	-0.6	<.001
	UEQ Dependability	0.2	.836	0.2	.264
	UEQ Stimulation	0.8	<.001	0.6	<.001
	UEQ Novelty	1.2	<.001	1.2	<.001
User Behavior	Microphone to user	28.8	<.001	44.7	<.001
	Microphone covered	13.8	.008	15.7	<.001
	Microphone touched	2.9	1.0	1.1	1.0
	Shaking	10.0	.008	7.5	.128
	Length < 30 sec	24.1	<.001	23.0	<.001
	Inside building	7.0	.156	5.9	.376
	User makes noise	26.6	<.001	23.8	<.001

	Variable	*Basic* vs. *Premium+* Change	*Basic* vs. *Premium+* p value	*Basic+* vs. *Premium* Change	*Basic+* vs. *Premium* p value
User Experience	SUS	0.6	1.0	6.5	.004
	UEQ Attractiveness	0.6	<.001	-0.3	.208
	UEQ Perspicuity	0.1	1.0	-0.9	<.001
	UEQ Efficiency	-0.3	.064	-0.3	.020
	UEQ Dependability	0.2	.348	-0.4	.004
	UEQ Stimulation	0.9	<.001	-0.1	1.0
	UEQ Novelty	1.4	<.001	0.0	1.0
User Behavior	Microphone to user	52.5	<.001	15.9	.008
	Microphone covered	18.3	<.001	1.9	1.0
	Microphone touched	6.6	.032	-1.8	1.0
	Shaking	11.5	<.001	-2.5	1.0
	Length < 30 sec	27.1	<.001	-1.0	1.0
	Inside building	7.9	.068	-1.1	1.0
	User makes noise	32.2	<.001	-2.7	1.0

Table 17.: Statistical study results. The *Change* value represents the observed difference between the two conditions, whereby negative values represent an aggravation relative to the *Basic* app, whereas positive values represent improvements. The change is statistically significant if $p < .05$. Because of multiple comparisons, p values were adjusted according to Bonferroni.

worked more effectively, $t(264) = 3.088$, $p = .008$. However, the usability was much better when providing a tutorial, all $t(264) > 2.829$, $p < .021$, while the perceived hedonic quality was similar, all $t(264) < 1$, $p = 1.0$.

The qualitative data show that the tutorial enhanced the comprehension of the task (i.e. the measurement process per se), whereas the technical measures mainly help the users to handle the smartphone correctly.

Complementary Effect of Technical Measures and Instruction

The previous comparisons showed that the technical implementation and the instruction led to a similar user behavior, but the usability was reduced dramatically when technical features were implemented without explaining them to the user. The users made fewer errors in all categories, all $t(265) > 2.690$, $p < .033$, except for the location, $t(265) = 2.417$, $p = .068$. This might be a statistical artifact of the study, as about 80% of the participants were already outside of a building when asked to participate in the study. Because we recorded the initial location and the location of the actual measurement, we were able to assess how many participants changed their location. When using the *Basic* app, 10 of 24 participants (41.7%) left the building, whereas 24 of 29 participants (82.8%) left the building when using the *Premium+* app. This difference was statistically significant, $\chi^2(1, N = 53) = 9.642$, $p = .002$.

When combining the positive effects of the technical measures with those of the instruction, compared to the *Basic* app, the usability did not significantly decrease, measured with the SUS, $t(252) < 1.0$, $p = 1.0$, and the three scales of the UEQ that are supposed to measure the pragmatic aspects of a product, i.e. perspicuity, $t(265) < 1.0$, $p = 1.0$, efficiency, $t(265) = 2.417$, $p = .064$, and dependability, $t(265) = 1.724$, $p = .348$. However, the hedonic aspects, i.e. novelty, $t(265) = 9.704$, $p < .001$, and stimulation, $t(265) = 6.000$, $p < .001$, increased.

In the qualitative data two clusters were observed that grouped together likes and dislikes of (i) the technical support of the user (which was described as responsive behavior of the app), and (ii) the tutorial of the *Premium+* app. Overall, the responsive behavior was perceived as very positive. None of the users commented not to understand the function of the technical support. The *Basic* and the *Premium+* app were described as easy and fast to use. However, only when a tutorial was provided, the participants reported to comprehend the task. They even reported that they felt that they could not make any mistakes. However, it should be noted that the hedonic quality remained low and the task itself was still experienced as being boring.

Summary of the results

The study showed that (i) technical measures as well as instructions reduced observed error rates, (ii) technical measures without explanation reduced the perceived usability and user experience, and (iii) technical measures and instructions nicely complement each other. The instruction enhanced the comprehension of the task (i.e. the measurement process per se), whereas the technical measures mainly help the users to handle the smartphone correctly.

6.6 DISCUSSION AND LESSONS

In this section we discuss our previously presented findings and summarize takeaway messages.

6.6.1 *Empirical Taxonomy*

In the course of empirically building our taxonomy of human error in Participatory Sensing and Citizen Science (as summarized above in Figure 65) we had many discussions concerning the comprehensiveness of the collection, its value to others as well as the selection of the use cases for building it.

We are of course aware that there are many areas of mobile sensing and current devices' capabilities afford for the collection of multiple types of data, some more prone to user error than others. In order to build an abstract characterization of problems, we attempted to select specific but representative use cases that entail different aspects of Participatory Sensing. On an abstract level, many applications fit the topics of the four selected use cases (sensing phenomena, handling objects, annotating data, using additional unfamiliar hardware/devices and assembling equipment). As such, we are convinced that the taxonomy will be a useful resource for designers and researchers concerned with improving non-expert data collection from mobile devices.

Regarding the comprehensiveness of the taxonomy, we are not aware of analytical work on errors in Participatory Sensing that could have served as a baseline for comparison with our empirical approach. We have however recently become aware of a characterization of all factors that make up uncertainty in measurement processes in industrial testing, authored by the *German Association of the Automotive Industry (VDA)* [242]. They distinguish influences between pertaining to the measurement system (measurement standard, mounting fixture, measuring equipment and measurement parameters) or the measurement process (environment, object, methods, and operator). Our own taxonomy can be mapped well to most of these aspects, indicating a high degree of completeness. The most notable difference is that in industrial processes, the human operator has no power over the process or the system, while in mobile sensing, the user directly or indirectly influences all of its aspects.

6.6.2 *Balancing Data Quality and Usability*

The results of our study show that technical measures alone can already help to significantly reduce human error in Participatory Sensing. However, an interesting finding is that built-in automated mechanisms for improving data quality may be detrimental to the user experience. While technical measures alone and instructions alone each seem to have a roughly similar effect in terms of error reduction, technical measures without explanation notably reduce the perceived usability of the app, potentially frustrating the users.

Instructions vs. technical measures

Taken at face value, this seems to suggest that citizen science mobile sensing app designers may be better off focusing on making clear instructions (and embedding them into the app) before designing mechanisms to nudge user behavior towards accurate data collection. Looking closer, it is not that simple. While quantitatively, both technical measures and instructions reduce error by a similar amount, they address different kinds of adverse behavior. Obviously, some errors are not preventable using technical measures: not making noise in sound sensing is a good example. While e.g. algorithms for voice activity detection have existed for a long time [112], it is nigh impossible to automatically decide if the detected speech is part of the ambient noise or an artifact of improper measurement procedure. On the other hand, we can simply instruct participants to not talk while recording (and hope that they do). Another difficulty in using technical measures was illustrated in our empirical study with the *iSPEX* system (see Exploratory Study 3 above). Mechanisms that automatically verify certain aspects of the sensing context may be designed poorly or too restrictive. When the user is certain he has done everything correctly and the system insists that this is not the case, this results in frustration. Approaches to improve this include either imposing less strict constraints or making the app more intelligible, e.g. by making it possible for the user to better comprehend and maybe even override automatic decisions, thus providing control over the context-aware application to the user [14].

Instructions on the other hand have their drawbacks as well. Designing adequate instructions is a complex task [253], and on top of that, people tend to not read manuals [173] and/or skip tutorials. A downside of written instructions also is readability, particularly outdoors: one of the participant in Exploratory Study 1 performed their measurements in direct sunlight, having a lot of trouble reading the screen and moving around a lot as a result. Varying levels of literacy may also need to be considered when designing tutorials in certain countries. Solutions to these problems may include using different kinds of instructions (icons, videos) or displays (audio instructions, vibration, etc.).

Remaining Uncertainty

A finding that one may overlook in the face of the observed improvements is that even in the best of our four cases, a significant amount of participants still exhibit erroneous behavior: More than 12% of the participants made noise while recording despite being instructed not to and roughly 8% of the users still attempted to sense with the interface of the app being upside-down and after being explicitly prompted to flip their phone. This suggests that even if considerable effort is placed into reducing human error, a certain amount of afflicted data with questionable quality may always have to be expected. Of course, this can not simply be generalized. An important aspect in that is to not only think about types of errors and the frequency of their occurrence, but also rate them regarding their severity in terms on their effect on the quality of the measurement. The qualitative feedback that participants gave in our field study also provided some insights into why some errors still might have remained. The requirement to measure over a longer time period, e.g. more than 30 seconds as in the current study, was perceived as being boring and uninteresting. This phase may have contributed to the occurrence of users making noise. Here, some strategy to reduce boredom, such as embedding the sensing task in a game context [42], might work to mitigate this type of erroneous behavior.

Recurring Users / Long-term Behavior

In our study, we sampled every participant exactly once, so in that sense, we have shown that the implemented measures can effectively mitigate errors caused by people who have not performed the task before. This is important as it can significantly lower the threshold for newcomers and infrequent participants. But what about recurring participants? Without further measures, it can be both argued that performing a task repeatedly may either increase data quality or not. On the one hand, revisiting data collection activities should reduce slips, i.e. errors in carrying out the intention [169], over time. Mistakes (errors in the intention) on the other hand are more likely to be repeated.

Concerning the effect of instructions and technical measures, we would argue that technical measures are likely to maintain their positive effect over time and maybe even combat slips that would have otherwise occurred, e.g. due to decreasing motivation. Regarding instructions, it is likely that in the long run, people will read the tutorial less carefully or not at all anymore. However, the instructions do not only have a teaching character but also serve as a reminder. Not revisiting them may result in an increasing amount of mistakes in the long run, as people may forget individual steps or mix them up. This also relates to certain design choices that may be important for the long-term experience: Should tutorials be skippable? If so, would people skip them already the first time without reading them? If not, will people be annoyed by them in the long run? One approach could be to reshape the tutorial over time to slowly transform from a tutorial to a shorter reminder.

Gamification approaches, in which participants are extrinsically motivated through the use of game elements, may also have a positive effect on long-term motivation. We present the concept of embedding sensing tasks into games in order to incentivize correct sensing in chapter 7 of this work.

6.6.3 *Stakeholders*

Both for grassroots movements, like the campaign of the *OK Lab* described in Exploratory Study 4, and "top-down" approaches, often driven by experts, it is important to incorporate the interests of all relevant parties. If e.g. activists gather information concerning environmental stress, it is crucial to talk to civic authorities early on, because in the end they will judge whether they accept the data quality as being adequate and make decisions concerning data use. Conversely, organizers need to keep the interests of participants in mind, as people do not want to be instrumentalized and reduced to data collection tools.

However, there is not only a range of stakeholders (participants, organizers, researchers, authorities,...), in our discussions so far we have also seen different positions within these groups. For example, we have heard of municipalities that are strongly interested in the possibilities that distributed low-cost sensing may offer and that actively work towards integrating such approaches into their current monitoring networks. On the other hand, there are civic authorities that are either wary, or even actively work against crowdsensing projects, possibly out of fear that the gained information may result in financial burden. But even among participants, we have seen both citizens that are e.g. interested in improving air quality and those that are opposed, because they feel that the pollution does not really affect them, but measures to combat it probably would, like bans on motorized traffic. Things get even more complicated if extrinsic motivation to participate is present, like gamification or monetary incentives, because the primary goal of the participant may not be data collection anymore.

In the end, we believe that project coordinators need to work closely with volunteers and system designers, and ultimately successful systems will have to involve a creative collaboration between the variety of actors and stakeholders, which also addresses non-technical aspects like social, cultural, and political issues [170]. Human-centered design, starting with observations on how participants err in real-world situations can be used to build software systems for high-quality Participatory Sensing in an incrementalist approach, that in turn can help to establish mutual trust among stakeholders.

6.6.4 *Takeaway Lessons*

Gardliner et al. [94] argue, that while the risk of low data accuracy is present in citizen science, the cost-effectiveness of crowd-sourced science compared to the conventional approach outweighs the risk, if properly handled. We summarize our key findings regarding the proper handling of risks from our studies and discussions, respectively the recommendations for building Participatory Sensing systems in the following list:

- Be aware of the diversity of human error and the effect it may have on data quality.
- Analyze people's errors in order to adopt user interfaces that help to prevent them.
- Address different classes of error with appropriate measures (e.g. instructions and/or technology).
- Design for intelligibility: It's not just what technology can do, but also how a user perceives it.
- Do not overreach, too strict constraints will frustrate the user.
- Involve stakeholders early on to balance required data quality and necessary complexity.

While we encourage designers to shift the perspective towards the correct execution of the measurement process, data quality should not be the only goal. Rather, the focus on data quality should complement proven user-centered design processes. Our study shows that technical measures and instructions nicely complement each other. Our findings highlight the criticality of balancing technological features and their perceived ease of use, a fact which both practitioners and researchers need to be aware of.

Given the problems inherent in accurately predicting system performance in real-world environments, conducting small exploratory studies has proven to be an easy way to collect erroneous behavior specific to the task at hand. The error dimensions presented in this work should provide a useful starting point for system designers developing interfaces and interaction in a way that minimizes the occurrence of human error and thus leads to more uniform and overall better data quality in Participatory Sensing. This is the theme of this paper.

6.7 Conclusion

This chapter focused on the interplay between non-expert user behavior and data quality in Participatory Sensing. To foster a deeper understanding of citizen science tools, it explores the design space of mobile citizen science sensing tools and applications, with the focus on human error. We have presented an empirical taxonomy of errors exhibited in non-expert smartphone-based sensing, based on four small exploratory studies. A large field study that compares instructions and technical measures to address these errors shows that technical measures without explanation notably reduce the perceived usability and the combination of technology and instructions achieves a significant reduction in observed error rates while not affecting the user experience negatively.

7 Sensified Gaming

The previous chapter concerned itself with the errors non-expert users may make when taking environmental measurements with their phones. In this chapter, we employ gamification techniques for the same purpose. Because Participatory Sensing (PS) scenarios often require a critical mass of users, applying gamification to different areas in order to increase user engagement has been proposed. However, existing attempts often default to the standard points, badges, and leaderboards and fail to recognize the potential of exploiting game design elements beyond creating user engagement. We propose not to think of Gamified Participatory Sensing when designing such systems, but rather of *Sensified Gaming*. To this end, this chapter presents a collection of design patterns and game mechanics that can be used to identify or design suitable games, into which participatory sensing tasks can be embedded. Eventually, we present a mobile minigame that opportunistically performs ambient noise level measurements during gameplay.

Parts of this chapter have previously been published and presented on the *2016 International Conference on Advances in Computer Entertainment Technology (ACE)* [42]. The initial collection of Game Design Patterns for Gameful Environmental Sensing was composed by Rikard Johann Öxler for his diploma thesis [175], and reviewed jointly with Jussi Holopainen for publication. The game *SpaceMaze* was partially designed and fully implemented by Jan Felix Rohe for his bachelor's thesis [196].

7.1 Introduction

In Participatory Sensing [48], users with personal mobile devices (e.g. smartphones) collaboratively collect information at different locations and upload it for a joint cause. Applications cover a wide range, including urban issue tracking [125], real-time monitoring of road congestion, weather conditions, air quality

[82] or noise pollution [148]. Of the many tasks and systems that exist, practically all depend on a certain level of participation and user engagement to function. Gamification has been proposed as an approach to provide incentives for participation [9], [25], [155], [204], [239]. However, often attempts at gamifying such systems are executed half-heartedly or fail to recognize the potential of gamification techniques beyond standard PBLs (points, badges, leaderboards). In this paper we propose the concept of *Sensified Gaming* as an alternative way of thinking as opposed to gamified Participatory Sensing: The idea is – rather than gamifying a task – to think of designing a game in the first place, that secondly also is suitable to support a sensing task. Depending on the application case, Participatory Sensing has various requirements (e.g. moving around, being outside, visiting certain locations of interest, etc.). This paper focuses on the research question what the crucial elements to create *Sensified Gaming* are, i.e. how to support participatory sensing in games. For this, we (1) identify core tasks from the field of participatory sensing and (2) collect and map game mechanics that are suitable to embed these tasks, presenting a set of design patterns.

7.2 PARTICIPATORY SENSING

Participatory Sensing is one of many similar concepts that overlap to a certain degree: *Citizen Science*, *Volunteer Monitoring*, *Crowd Sensing*, *Citizen Observatories*, *Amateur Science*, *Community Science*, just to name a few. They all have in common that a group of (often untrained) people collaboratively works on (parts of) a joint task. The tasks itself and the platforms that are used differ between the individual concepts. Participatory Sensing, as defined by Burke et al. [48], emphasizes distributed sensing done by everyday users with the personal mobile devices they carry and control in the public sphere. In this work, we focus on such settings: environmental sensing with smartphone sensors (or other personal devices connected to them), in the real-world (e.g. a city) by ordinary (i.e. non-expert) people. Still, the presented core tasks generalize to a wide range of applications.

7.2.1 *Core Tasks*

We have identified four core tasks as requirements environmental sensing:

- *Coverage*
- *Touch POI*
- *Rendezvous*
- *(Correct) Sensing*

Coverage

Since the goal of Participatory Sensing generally speaking is to crowdsource a task to people in a public space, achieving suitable coverage of that area is an obvious requirement. This is especially true for applications in which a map is constructed from individual measurements or observations, like in environmental sensing. Here, coverage is meant both in time and in space, i.e. ascertaining that sufficient measurements are recorded across the area of interest continuously over time.

Touch POI

This core task subsumes activities that require going to a point-of-interest (POI), i.e. a specific location (or one of a set of locations). The reasons for this can be diverse. In scenarios with low-cost sensors, especially air quality sensing, the need for regular (re-)calibration of sensors is present [47]. Calibration can be carried out by co-locating a sensor with a high-precision reference device or station whose readings are used for calibration. Scenarios with user-generated reports may also require measurements at certain locations to verify data points or complement automatic data cleaning approaches [36]. Another need to travel to a certain location may be data offloading [138]. Especially in data heavy sensing tasks, e.g. if high-volume data like video or high-frequency data of many sensors is recorded, the need to move data off the participant's device may arise regularly. This may require offloading traffic to a WiFi network in case the participant does not have a data plan or wireless service reception is bad. The same is true for situations with no connectivity in which the collected data needs to be uploaded within a certain time to be of value.

Rendezvous

Calibration can not only be carried out against high-precision data (so-called *ground truth*), but also against other already calibrated devices. Such an approach was e.g. presented for low-cost gas sensors by HASENFRATZ ET AL. [104]. They proposed a multi-hop calibration scheme in which the sensors exchange measurements collected during a rendezvous in order to improve their individual calibration "on-the-fly". Other similar approaches exist [254], among them one that additionally incorporates privacy protecting measures [153]. Other reasons for co-locating participants may be the desire to collect redundant readings, the need to collaboratively collect readings (e.g. for verification), or again for data offloading.

(Correct) Sensing

In participatory sensing with low-cost sensors, the potential for systematic error that leads to low-quality or even useless data is high [39]. As usually non-experts perform the tasks, ensuring the correct execution of the measurement process is important. Aspects of a correct sensing procedure with smartphones include correct body posture, device orientation, environmental constraints (good weather, being outside, remaining motionless, etc.), sufficient measurement duration, data annotation as well as the correct sequence of the steps.

7.3 SENSIFIED GAMING

We argue that gamification can do much more for Participatory Environmental Sensing applications than 'merely' provide incentives for participation. Different mechanics can be used to support the presented core tasks. As it is important that the various mechanics are not looked at individually but rather in the context of their interplay, we encourage the notion of "sensifying a game" rather than "gamifying a task". This section attempts to more closely define the term *Sensified Gaming* and place it on the continuum of existing nomenclature.

Mobile Games are (video) games that are played on mobile devices, as phones, tablets, smartwatches, and the like. As we focus on support for smartphone-based sensing applications, Sensified Gaming naturally (but not necessarily) entails Mobile Gaming. At the same time, since sensing clearly pertains to the real world, we touch the field of *Pervasive Games* [146]. Depending on the definition, Pervasive Games can narrowly be seen as a combination of game reality and physical reality within the gameplay or more broadly as games that have "*one or more salient features that expand the contractual magic circle of play spatially, temporally, or socially*" [161]. A deep discussion including a classification has been presented by HINSKE ET AL. [107].

One sub-class that is certainly closely related are *Location-based Games*.

The distinction between gamified applications and games is often made according to the underlying design goals [72], [151]: In *Gameful Design*, game-like thinking is applied to a design process without the actual inclusion of game elements. *Gamification*, as defined by DETERDING ET AL. is "*the use of game-elements in non-game contexts*", explicitly excluding full-fledged games [72]. *Serious Games*, in contrast, are "*full-fledged games for non-entertainment purposes*". More fine grained distinctions of Serious Games exist, again characterized along the difference in design goals. MARCZEWSKI [151] sub-categorizes them into *Teaching Games / Games For Learning*, *Simulators*, *Meaningful Games / Games For Good* and *Purposeful Games*. Of these, none perfectly accommodates our notion of Sensified Gaming. While the term Purposeful Games covers it best, there are notable differences, e.g. that sensing itself does not directly affect the real world, as the definition of purposeful games entails. Also, the player does not really need to know that playing the game achieves something in the real world – at least from a classification point of view[1]. Finally, *(Full-fledged) Games* are the ones that are designed purely for entertainment.

One could argue that while defining these classes according to the design goal makes sense, thinking too much about the task when designing a serious game may result in a badly designed, unenjoyable or shallow game. Another issue of designing such a purposeful game with the task being the first thing in mind is that the resulting game almost automatically will be tailored to the people who are expected to work on the task anyway. The fear that a badly designed system may be counter-productive was e.g. expressed by the developers of the citizen science game *Floracaching*, who "*. . . don't want glitches in the technology to de-motivate this important group of contributors, potentially preventing them from submitting future data*." [25]. In contrast, the target user group of Sensified Gaming is gamers that don't need to have any motivation regarding the underlying purpose. While technically, the ultimate design goal of *Sensified Gaming* is in fact supporting a Participatory Sensing task, we argue that this should not make a difference and propose to still put entertainment first in the design process. Overall, we introduce the term *Sensified Gaming* as a simpler way of saying "*Digital Purposeful Pervasive Mobile*

1 Whether it is right to deceive or coerce a participant into working on a task without knowing it is a question of its own.

Games for Participatory Sensing that are designed to be full-fledged games."

7.3.1 *Gamified Participatory Sensing*

There are some success stories of lightly gamified Participatory Sensing systems, but all with the focus on motivation/engagement and – with one notable exception – not taking into account additionally supporting any of the core tasks. UEYAMA ET AL. for instance present a system of badges, points and leaderboards (PBL) to supplement monetary incentives, the core goal being to reduce the cost spent on monetary incentives [239]. This work was continued by ARAKAWA ET AL. [9]. PBLs are also the core of the gamification design for the participatory noise pollution monitoring system *NoiseMap*, whose authors could show an increase of recorded measurements in a comparison of different prototypes with varying degree of gamification [204]. In this app, the sensing task is still clearly central and it is not intended to be a full-fledged game. The exception mentioned above also is in the field of noise pollution sensing: *NoiseBattle* and *NoiseQuest* [155] are two game prototypes addressing different player types (the types *Killer* and *Explorer* according to Bartle [13]). Both games are designed not only to encourage participation but also to increase coverage in the game area by making players explicitly explore the area respectively battle for control of cells in a grid by measuring ambient noise with smartphones. Unfortunately, the games seem to have been research prototypes that were never publicly released. A notable commercial project that incorporates players collecting information in urban environments is the pervasive game *Ingress*[2], which managed to attract an enormous player base of several million people around the world. While this manuscript was undergoing review, the game *Pokémon GO*[3] was released by the same company, quickly surpassing *Ingress*'s success and being played massively all over the world, players having walked a total of 4.6 billion kilometers so far [129]. This shows the great potential of such location-based games and indicates that they will become of interest to a broader audience.

To the best of our knowledge, game design elements have today not been used further to support the presented core tasks. While FLATA ET AL. presented so-called *calibration games* [91],

2 `https://www.ingress.com/`
3 `http://www.pokemongo.com/`

their notion of calibration does not relate to the calibration of mobile environmental sensors, but rather to the adjustment of input devices such as eye-trackers.

7.4 GAME DESIGN PATTERNS

Design patterns [6] were introduced within the discipline of architecture for easy knowledge transfer between professionals and non-specialists. These patterns encode design practices as problem-solution pairs with accompanying information and interrelate to form hierarchies or nets. This design patterns concept has spread from architecture to several other areas, e.g. programming [86] and interaction design [24], [65], [86].

Design patterns are an example of explicitly creating a design language [195], as a way of understanding a design discipline through the relevant elements and materials, how these elements can be structured, and in which situations specific elements and choices of structures are appropriate. Specifically, they let those involved in the process consciously consider and discuss what the implications of design choices would be for the final design. Design patterns are not complete design languages in themselves since they focus on the basic elements and do not describe the steps of a design process.

The idea to use design patterns for game design was introduced by KREIMEIER [130] in the context of computer games and has since then been generalized to all types of games [20], [21], resulting in a pattern collection of nearly 300 gameplay design patterns [19]. These patterns differ from the original structure by replacing the problem-solution pair with a causes-consequences pair describing how the pattern can occur in a game design and how it can affect the gameplay and player experiences. The reason for the change was because the patterns are intended to support both the design and analysis of games and also to allow the use of specific patterns as design goals. In 2009, Björk started a wiki to collect more patterns, and has up to this point assembled an extensive (and still growing) gameplay design patterns collection [18]. The selection of patterns in this work is largely based on that collection.

7.4.1 *Methodology*

After having identified the core tasks, we surveyed literature and online sources for game design elements that are fitting to build (or identify) games suitable for Sensified Gaming. Björk's wiki [18] is with currently 536 entries the by far largest collection and was therefore the main source for the mechanics presented in this work. Other pattern collections [19], [67], [109] mostly contained subsets of the set of mechanics found in this wiki. A small number of patterns was added from these sources, as well as individual ones that we did not find in any collection but rather came up with ourselves in the process of discussing the core tasks and the concept.

The pattern analysis was conducted by three researchers who were familiar with the patterns approach, the pattern collections used in the analysis, and the principles of participatory sensing. One of the researchers did the initial selection of relevant patterns based on whether the pattern could substantially support at least one of the core tasks. The selection was then reviewed by the two other researchers, suggesting additional patterns to be included. Patterns that are generally suitable for all types of games but that were classified by us as not being especially relevant regarding the core tasks were left out to provide a more condensed collection. This includes patterns such as tension or the "usual suspects" points, badges, leaderboards and achievements. However, this does not mean that we think other patterns should not be used in Sensified Gaming, but rather that they do not indicate specific suitability. Finally the patterns were categorized with input from all three researchers following the principles of thematic analysis [27].

7.4.2 *Pattern Collection*

All of the selected mechanics are shown in Table 18, divided into categories. The table features the name of each design element[4] and one column for each of the core tasks ('×' indicates special suitability for the respective task). In addition to these, we added a column for *Sustained Engagement* to the table, which is used to denote mechanics that especially encourage "replayability" (rather then generally motivate, which would basically be any game mechanic). This section elaborates on the elements in our collection and their categorization.

4 For the sake of better readability, the textual description of the mechanics were omitted in the table. It can be found in the running text of this chapter and in Björk's wiki [18].

	Mechanic	Coverage	Touch POI	Rendezvous	(Correct) Sensing	Sustained Engagement
Prerequisites	Game Servers	×	×	×		
	Mediated Gameplay	×	×	×	×	
	Dedicated Game Facilitators	×	×	×	×	
	Hybrid Gameplay Spaces					×
	Persistent Game Worlds	×				
Modes	Pervasive Gameplay	×				
	Asynchronous Gameplay	×	×			
	Attention Demanding Gameplay			×		
	Casual Gameplay	×				
	Lull Periods	×	×		×	
	Massively Single-Player Online Games	×				
	Real-Time Games			×	×	
Exploration / Expansion	Area Control	×				×
	Capture	×				×
	Expansion	×				×
	Game World Exploration	×				×
	Fog of War	×				×
	Free-roam / Open World	×				×
	Configurable Gameplay Areas	×				×
	Physical Navigation	×			×	×
	Artifact-Location Proximity	×	×			
	Real Life Activities Affect Game State	×	×	×		×

(continues on next page)

Table 18.: Design patterns for Sensified Gaming, mostly selected from Björk's wiki [18] according to their suitability to support the core tasks in participatory environmental sensing.

	Mechanic	Coverage	Touch POI	Rendezvous	(Correct) Sensing	Sustained Engagement
	(continued from previous page)					
Location / Proximity	Player-Location Proximity	×	×			
	Player-Artifact Proximity	×	×			
	Bases		×	×		
	Game Items	×				
	Pick-Ups	×				
	Resource Locations	×		×		
	Point of Interest Indications	×	×	×		
	Events Timed to the Real World	×		×	×	×
	Geo-fencing	×	×	×		×
	Location-Fixed Abilities		×	×		
Social	Instances			×		
	Massively Multiplayer Online Games	×		×		×
	Multiplayer Games			×		×
	Synchronous Gameplay			×		
	Late Arriving Players			×		×
	Player-Player Proximity		×	×		×
	Common Experiences			×		×
	Game Element Trading			×		×
	Mutual Goals			×		×
	Symbiotic Player Relations			×		×
	Collaborative Actions			×		
	Cooperation			×		
	Privileged Abilities/Orthog. Differentiation			×		
	Team Strategy Identification			×		
	(continues on next page)					

Table 18.: Design patterns for Sensified Gaming, mostly selected from Björk's wiki [18] according to their suitability to support the core tasks in participatory environmental sensing.

	Mechanic	**Coverage**	**Touch POI**	**Rendezvous**	**(Correct) Sensing**	**Sustained Engagement**
	(continued from previous page)					
Sensing	Feelies				×	×
	Unlocking				×	×
	Mimetic Interfaces/Physical Enactment				×	
	Minigames				×	
	Tutorials				×	
	Reputation				×	
Engagement	Predetermined Story Structures	×	×	×	×	
	Alarms			×	×	
	Dyn. Diff. Adjustm./Challeng. Gameplay					×
	Notifications			×		×
	Replayability	×				×
	Ubiquitous Gameplay	×				×
To consider	Extra-Game Input	×			×	
	Inaccessible Areas	×	×		×	
	Player Physical Prowess					×
	Unmediated Social Interaction			×		
	Extra-Game Consequences				×	

Table 18.: Design patterns for Sensified Gaming, mostly selected from Björk's wiki [18] according to their suitability to support the core tasks in participatory environmental sensing.

Prerequisites

There are some design elements that are important, if not crucial to sensified games. Since typically a central instance that coordinates players is needed, having *Game Servers* is more or less mandatory. Since the system presents the game state and controls the interactions with other players before a possible meeting in the real world, we usually have *Mediated Gameplay*. *Dedicated Game Facilitators* are responsible for this mediation.

They keep track of the game state and guide the players through the game world, e.g. by *Game Element Insertion*, controlling non-player characters or by giving information to the players. As the measurements take place in the real-world, *Hybrid Gameplay Spaces* are a direct consequence. *Persistent Game Worlds* are not a necessity, but can enable a deeper and more complex interaction by enabling players to play asynchronously.

Modes

There are different modes of gameplay that can be used. Some of them are conflicting, so they can not be used at the same time but it is possible to switch between them in different parts or stages of a game. A possibility that e.g. lets players embed playing into other everyday life activities is some sort of *Pervasive Gameplay* or *Casual Gameplay* with many *Lull Periods*. When someone is waiting for the bus they can just start the game to kill some time. *Attention Demanding Gameplay* is fitting for more thrilling games (or phases of a game) when players want to devote themselves more. *Massively Single-Player Online Games* with *Asynchronous Gameplay* could be used if users just play by themselves instead of with others. Further modes are presented in the category *Social* below.

Exploration / Expansion

This category mostly covers mechanics that motivate players to move around the world (both game and physical) and are therefore especially suitable to support the core task of reaching coverage. Possibilities to *free-roam* around the world and/or increasingly discover it through *Game World Exploration* promote player movement, e.g. to discover locations of other game elements. Mechanics such as *Fog of War* can be used in different ways to motivate travel to certain areas: If discovered areas on a map are e.g. covered again after a certain time, regular movement around the (entire) world is stimulated. Techniques like *Area Control* (for regions) or *Capture* (for items) also promote player movement, but do so by addressing the desire to increase the zone of influence. This can either happen continuously or during an explicit *Expansion* phase in a game. Most of these elements also encourage players to come back to play the game, as they address a sense of accomplishment. A bit of an exception to this are *Inaccessible Areas*, which are described further below.

Physical Navigation can help to increase coverage and – depending on the sensing task – may also be used to support correct sensing. *Artifact-Location Proximity* is a mechanism that can be used when, for example, sensors are not embedded in personal smartphones but rather in separate devices or real-world items that are not carried by the player continuously. An example are environmental sensors that are built into rentable public city bikes [47], as mentioned in the discussion of sensing scenarios in chapter 2. A game element could entail moving the bikes in the real world so that they then take measurements at different locations even after the player has moved on.

Location-based / Proximity

Instead of motivating to explore and roam the world, attempts can also be made to guide or lure players to specific locations. The mechanics in this category are mainly suitable for *touching POI* by encouraging players to travel to certain places. Indirectly, by that many mechanisms also support player *rendezvous*, as (randomly or deliberately) guiding people to certain locations in parallel greatly includes the probability of co-location. In both cases, an effect of this can be increased coverage, provided the locations in question are spread around the world accordingly. Basis to attract the user to special places are the *Player-Location Proximity* and *Player-Artifact Proximity*. Special places in the game can e.g. be *Bases* or locations of *Game Items* (*Pick-Ups* or *Resource Locations*), which in turn can in the real world be locations that feature ground truth reference stations for calibration or WiFi hotspots for data offloading. Such places can be made even more interesting to players if they feature *Location-Fixed Abilities*. *Point of Interest Indications* can lead players to special locations. To also increase temporal coverage, *Events Timed to the Real World* can be used. A possibility would be to e.g. only allow certain actions at specific points in time, under certain weather conditions etc. *Geo-fencing* is a way to trigger certain actions when a player crosses a certain perimeter, a mechanic that can be used in various ways for the core tasks, e.g. to alert players that they are close to a location of interest or entering an area of *Attention Demanding Gameplay* (see *Modes* above).

Social

Multiplayer Games form the principle basis for games with social interaction, by enabling *Synchronous Gameplay* for multiple players. Especially *Massively Multiplayer Online Games* seem suitable to support interaction in pervasive games as they increase the chance of finding other players in the vicinity. Supporting *Late*

Arriving Players is almost a necessity, as players should be able to independently start, join or leave a running game. *Instances* can help to reduce technical requirements like server load and can facilitate the formation of closer social groups.

Player-Player Proximity is an element that can be used for encouraging calibration rendezvous. In order to bring users together at the same time and place, *Common Experiences* can be used. Players can e.g. be brought together by *Mutual Goals*, *Collaborative Actions*, that increase a sense of community or *Game Element Trading* that requires real-world proximity. Also, elements that support *Team Strategy Identification*, like *Symbiotic Player Relations*, e.g. by *Orthogonal Differentiation*, can be used to bring players together.

(Correct) Sensing

An interesting and until now relatively unexplored area is the use of game design elements to ensure the sensing task being triggered and carried out correctly. Whether and how mechanics can be used strongly depends on the concrete sensing application and the employed sensor(s). If for example sensing requires using external sensors, gadgets, or smartphone extensions such as e.g. clip-on air quality sensors for smartphones [30], [217], they could also be shipped with the game and act as *Feelies* to increases the game experience. To reduce cost and improve data quality, it would also be a possibility to give out sensors only after *Unlocking* to the best performing players as an in-game reward. The players with the best coverage, social interaction and most accurate simulated measurements unlock the sensor, which they then receive as hardware. Another approach could be to offer sensors of different quality and cost as tangible game items that can be bought with real-world currency. A related example for this is the *Pokémon GO Plus* wristband[5], that acts as a physical accessory to the digital game. Such wearables could easily also house sensing capabilities, act as tangible or *Mimetic Interfaces* or real-world items that e.g. boost game stats. Smartphone sensors such as accelerometer, light sensor, proximity sensor, etc. can provide information on the way the player handles the measurement device or on his physical activities, which can in turn be used to monitor if the user handles the device correctly. Another possibility to e.g. ensure a certain device orientation is the use of *Minigames*. The sensing itself could be a small *Minigame*, like balancing a virtual marble on

5 `http://www.pokemongo.com/en-us/pokemon-go-plus/`

the screen to steady the phone. If desired, in-game *Tutorials* can be used to enable the players to learn how to use the sensors correctly. In contrast to manuals, players benefit from getting immediate feedback. Tutorials can take different forms, they could e.g. also be embedded in training missions. If the game behavior that involves sensing has an effect on *Reputation* (either of the players themselves or their game avatars), data quality can possibly be increased. Aside from comparing the quality of the measured data to that of other players, the game could also try to verify that the players are real persons. In both cases players will probably be more eager to make more accurate measurements.

Engagement

This section covers mechanics that provide the player with reasons to play the game (more often), i.e. to sustain engagement. *Replayability* is very important for these types of games. To be able to collect a lot of data, it is important to keep up the interest of the player. *Predetermined Story Structures*, such as Adventures or Quests, can be used to keep the game from getting boring by constantly supplying new content. As an added benefit they also can be used for all of the core tasks: to bring players together or to certain places, to increase coverage or to task them with a certain procedure. Through *Dynamic Difficulty Adjustment* players can be kept in flow and steadily have *Challenging Gameplay* which would provide an incentive to collect even more sensor data. Since players carry their smartphones with them most of the time anyway, *Ubiquitous Gameplay* should be possible. If the player is running the game anyway, she can be alerted to phases of *Attention Demanding Gameplay* and "pulled into to the action" by *Alarms*. If not, *Notifications* may remind them that the game still exists.

To consider

There are also some game elements which are not purely beneficial (or even intended) but should be kept in mind (see also *Discussion* below): As players move in the real world, they may encounter *Inaccessible Areas*. Private property should not be entered and the players thus not prompted to do so. Places that require an admission fee (e.g. a zoo) should maybe also be excluded. In addition, environmental measurements should typically only take place outdoors. If the game allows being played inside of buildings, it should possibly recognize this and hence not record sensor data or dismiss the recordings after the fact. Another thing to bear in mind is possibly varying *Player Physical Prowess*. Games should ideally be accessible to anyone and players with poorer physical abilities should not be demotivated by being tasked with something they can not compete in with other players. *Unmediated Social Interaction* should also be considered. As players encounter each other during the game, they may talk to each other, befriend each other, etc. Possible effects should be contemplated. The same is true for *Extra-Game Input* or *Extra-Game Consequences*. If the players know that the measurements could have consequences in the real world (e.g. air quality data is used for automatic traffic control), it could make them try to manipulate the data and play the real-world effect rather than the game, to for instance deliberately close off a street for traffic.

7.5 DISCUSSION

While this work attempts to provide building blocks for *Sensified Gaming*, there are of course general design considerations to be followed. There has been more than a decade of research on the design of pervasive games alone, and many lessons can be learned from previous work, like believable story-telling [99], wisely choosing technology platforms, carefully balancing single- and multi-player content of the game, and offering sufficiently diverse possibilities in the game to the players [178]. It should also be kept in mind that e.g. external sensing devices should fit the overall design of the game, as this can influence the players' perception and attitude towards them [116]. Another important aspect is that while supporting one of the core tasks, some of the mechanics may have adverse effects regarding another core task or (sustained) engagement.

An important aspect that should not be underestimated concerns the ethical issues connected to *Extra-Game Consequences*. Some vivid examples for this were encountered by people playing the recently released augmented reality app-based game *Pokémon GO*: There have been reports of people being robbed after being lured into a trap by muggers specifically targeting players of the game [259], as well as numerous cases of injuries and dangerous behavior. Also, inserting competitive game elements in a pervasive game could excite unwanted real-world interaction between players. Situations in which players could be tempted to compensate in-game inferiority by somehow engaging opponents in the real-world should not occur. It may be prudent to design games without shared resources that players compete for. But not only undesired interaction between rivaling players may be an issue, unforeseen problems involving people outside the game may also occur: As an example, a man attacked a *Pokémon GO* player, slashing him across the face, as he apparently thought the gamer was video-recording him on his phone [198]. While this surely is an extreme example, it pays to consider how the sensing procedure could be seen and possibly misinterpreted by bystanders. Such effects are possibly strongest and extremely hard to foresee if the game is played without bounds, i.e. by anyone, at any time and in any place. In contrast, so-called *event games* constrain the game environment to a certain playing time, game area, player group and/or limited hardware, allowing the game organizer to exercise more powerful control over the game [160]. Many ethical issues such as the use of public places and different aspects of privacy and security are discussed at length by Montola et al. in a report on Ethics of Pervasive Gaming in the *IPerG* project[6], which can also be read as a guideline document for reflecting individual game designs from the ethical point of view [162].

Overall, many things have to be considered to create games that are fun to play, deliver meaningful data and do not place players in harm's way. We would like to stress that simply selecting mechanisms from the list and combining them together is not what this work proposes as a design practice. As mentioned before, design patterns do not describe the steps of a design process. Rather, the selection and the accompanying discussion in this work can serve as a tool to facilitate building new games or identifying existing ones that could be suitable to be 'sensified'. Games are hard to design and good games even harder, that is why there are game designers. We believe that for the process to work best, game designers and sensing experts should ideally work hand-in-hand to successfully realize the concept of *Sensified Gaming*.

7.6 real-world example: spacemaze

This section presents the design and evaluation of *SpaceMaze*, a sensified game for noise level monitoring. Similar to the audio recording app flavors that were presented in chapter 6, we created a mobile minigame with the purpose of guiding the

6 http://iperg.sics.se/index.php

user to performing a correct sensor procedure. *SpaceMaze* was implemented by Jan Felix Rohe as part of his bachelor's thesis [196].

7.6.1 *Design*

The idea behind *SpaceMaze* is similar to that of the *Premium* app flavor from the user study in section 6.5: The phone's internal sensors can be used to determine whether the sensing context is correct and allow or deny the recording of ambient audio levels accordingly. As requirements of the sensing scenario, we therefore also have the same constraints (as derived from the best practices for noise level monitoring from the PDF user guide of the *NoiseTube* project [149]). When recording, ...

- the device should be upside down, so that the microphone points away from the user.
- the device should be held very still, no shaking or moving too much.
- a recording must at least 30 seconds long to be valid.
- the user should not talk or make other noise.
- the user should be outside.
- no phone usage, e.g. typing, chatting or any other device functions for the duration of the recording.

The latter two aspects are actually hard to control. On a

Figure 70.: Game Elements used in Space Maze (left-to-right): dark green finish line, light green checkpoint, blue moving obstacle, yellow pickup, player space ship, red player-killing obstacle, and warp point

purely technical level, it is not possible to ensure that the player is quiet during the game. We mapped the remaining requirements to possible game design patterns. As a result, *SpaceMaze* features the following visible game elements (see Figure 70) corresponding to their appropriate game design patterns.

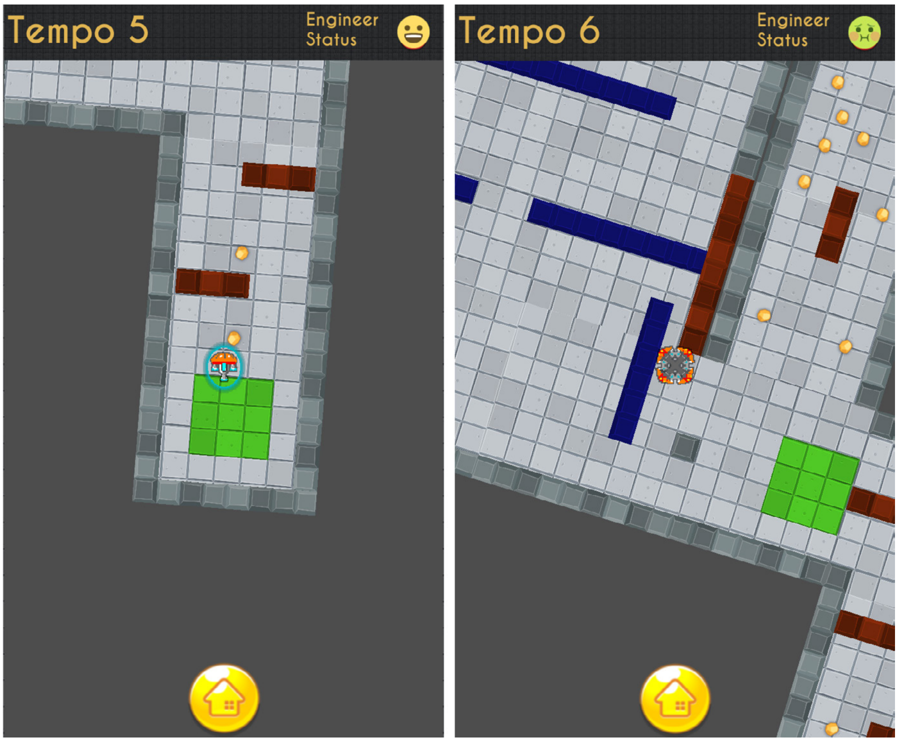


Figure 71.: Screenshots from the gameplay of *SpaceMaze*.

In *SpaceMaze*, the players need to steer a spaceship through a labyrinth (see Figure 71). The ship always moves and cannot be stopped. A shield protects the ship from damage, preventing it from being destroyed on collisions with blue obstacles or the level walls. Red obstacles will always destroy the ship, regardless of its shield's status. Upon destruction, the ship is respawned instantaneously at the last checkpoint. Yellow orbs are placed throughout the level as pickups. When collected, they increase the ship's movement speed, potentially allowing the player to clear the level faster. The total time upon finishing a level directly corresponds to the points earned.

Aside from the general gameplay, some of the elements were specifically selected to guide the device handling during the measurement process (see Table 19): Tho most central design pattern is *Mimetic Interfaces* [18]. By controlling the ship via the phone's inertial sensors (accelerometer and gyroscope), well-defined constraints are placed on the possible device motion. Mimetic interfaces are again used in conjunction with the *warp points* game element. In order to beam the spaceship to the next stage of the level, the user needs to rotate the phone by 180°. By that, the correct orientation of the phone for a measurement is

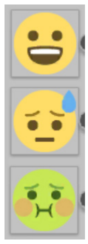
Task/ Constraint	Design Pattern(s)	Game Element	
Avoid shaking or movement; no phone usage	*Mimetic interfaces; Player physical prowess*	The spaceship is controlled by rotating the device ever so slightly. A transmission of 1:12 makes sure that the phone movement is minimized: 1° of rotation of the phone causes 12° of rotation in-game.	
Avoid shaking or movement; no phone usage	*Mimetic interfaces; Game items that protect from damage*	In the game's status bar, the status of the ship's engineer is shown. If the device is shaken or held in the wrong orientation, the engineer get's sick and the ship's shields power down. Without shields, the ship will explode when touching anything, e.g. the level walls.	
Device should be held upside-down	*Mimetic interfaces*	Warp-points transport the ship from one part of the level to another. To activate them, the user needs to turn the phone by 180°.	
Record for at least 30 s	*Real-time games;*	The ship is always moving forward and can not be stopped. The camera is locked on the ship, keeping it centered in the display. The continuous gameplay facilitates longer periods in which recordings can be made.	
Record for at least 30 s	*Pre-defined goals*	The levels have a linear structure, i.e. the user knows a-priori that he simply needs to play until reaching the end of the level, giving ample time for recording in the background.	

Table 19.: Selected game design patterns and their effect concerning the constraints of the measurement task.

ensured. Mechanisms to ensure a certain length of recordings are *Real-time games* and *Pre-defined goals*.

7.6.2 *Evaluation*

We conducted a user study in order to explore whether the design of *SpaceMaze* indeed has a positive effect on the performance of the measurement process, and if so, how this relates to the non-gamified app prototypes that we studied before [45], as presented in section 6.5. We recruited a total of 17 study

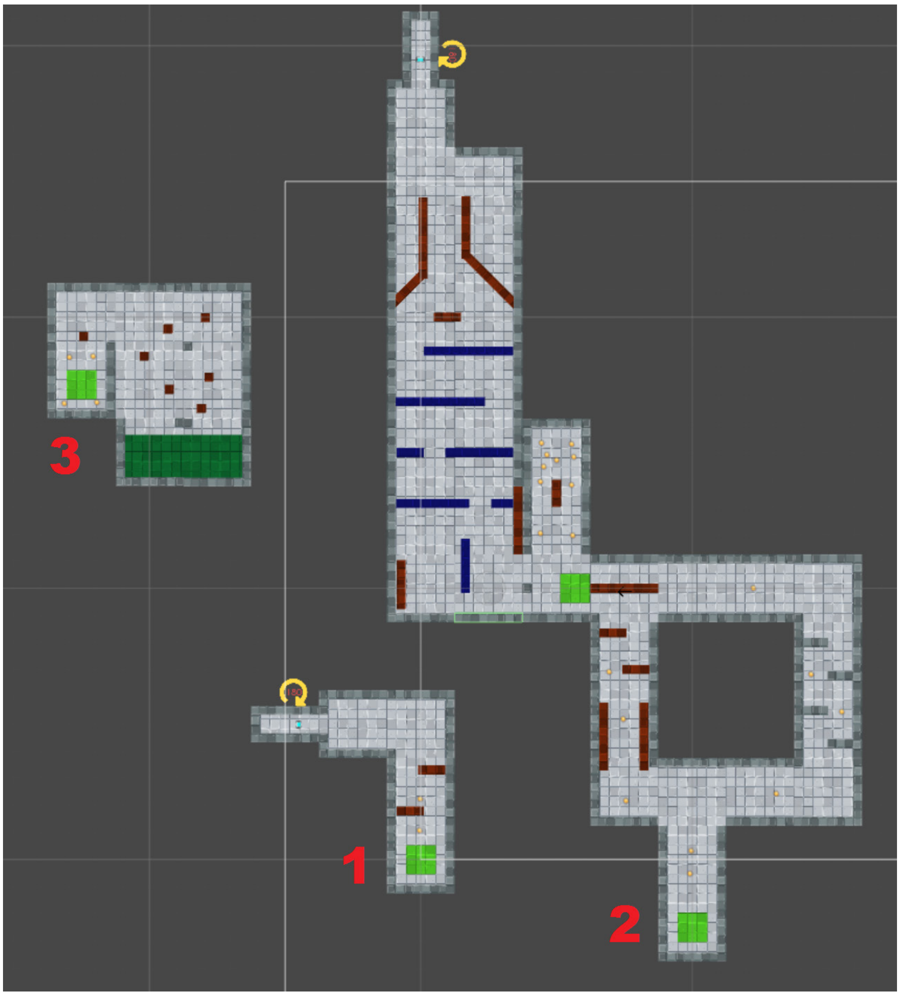


Figure 72.: Map of the first level of *SpaceMaze*: Players start in part 1 and are beamed to part 2 using a warp-point. This larger area is where the phone opportunistically performs the ambient audio sensing in the background. After having cleared the obstacle course the ship is eventually warped to part 3 of the level, leaving the user with a normally oriented phone on level completion.

participants (eleven male, six female), aged from 18 to 27 (mean age: 22). We approached passing pedestrians inside a city park and on a university campus and asked them if they would volunteer to participate (sample of opportunity, response rate at ca. 2/3). All participants were accustomed to mobile phone usage, stating they used their device daily. Additionally, each of them had played games on their mobile device in the past, many of them (N=9) still playing at least once a week. All participants gave their written informed consent and did not receive any compensation for their participation.

After having introduced the scenario, participants played the tutorial level once to familiarize themselves with the principle of the game. Then, they were asked to play one full level until they completed it or gave up. No further instructions were given. The study instructor observed participants and recorded all user errors (cmp. section 6.5). Besides that, the game recorded automatically logged gameplay data. After having finished, participants were asked to fill in the User Experience Questionnaire (UEQ) [136], the System Usability Scale (SUS) [28] and provide written qualitative feedback.

Study Results

All participants were able to generate valid noise recordings (mean duration: 85 s) during gameplay sessions. Almost all players (N=16) completed the main level, one gave up. Still, that participant still created a valid recording of 31 seconds. *SpaceMaze* reached a SUS score of 77, indicating an overall good usability, comparable to that of the non-gamified apps from chapter 6. The mimetic interface pattern was well-received, N=8 participants stating that they liked the unconventional motion control.

While all of the participants created a valid recording according to the game log, the observer log showed some erroneous behavior. Most prominently (N=11), participants made noise, specifically by talking. This was often prompted by events inside the game, such as collisions leading to the destruction of the spaceship. One participant (N=1) disrupted the recording by covering the device's microphone during play and one participant (N=1) tapped the device screen every time his ship crashed, seemingly attempting to accelerate the respawning of the ship.

Figure 73 shows the amount of errors made while playing *SpaceMaze* in comparison to those made with the *Basic* and *Premium* versions of the noise recording app presented in chapter 6.

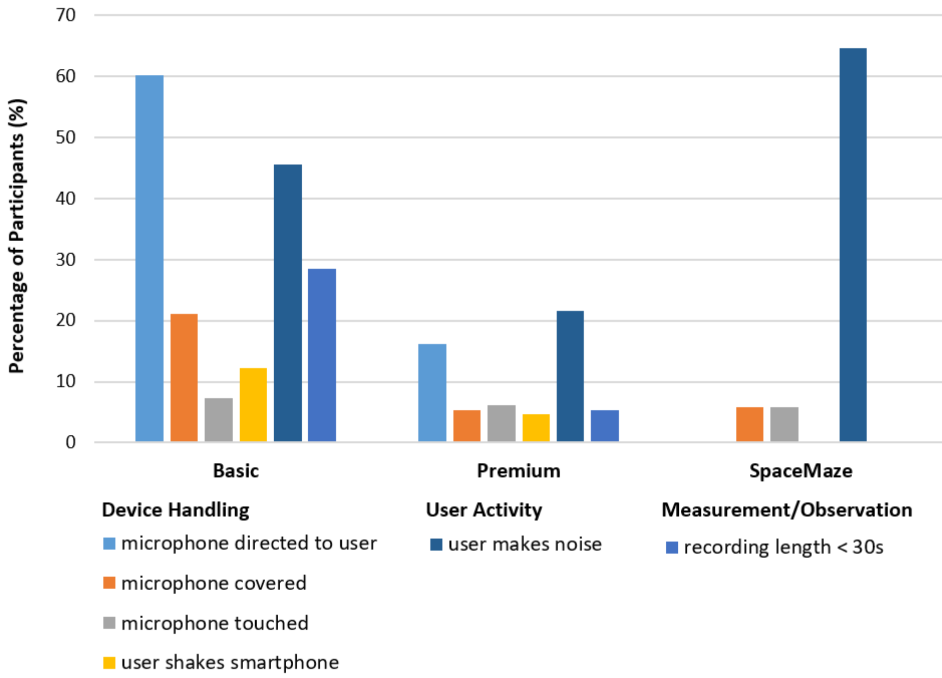


Figure 73.: Measurement errors observed during the sensing process in the mobile minigame *SpaceMaze* as well as the two applications *Basic* and *Premium* (cmp. section 6.5).

The *Basic* version is a simple 1-button recording app, whereas the *Premium* version features technical measures to guide the measurement process.

Since *SpaceMaze* opportunistically selects the longest possible recording interval that occurred during gameplay, naturally the error of too short recordings is eliminated entirely. For the same reason, errors on the account of device not being held upside-down did not occur. The handling errors that could not be automatically prevented are the touching or covering of the microphone. Quantitatively, they seem to occur at comparable levels to those of the *Premium* app, but a larger sample size would be needed to confirm this. The one error that occurred disproportionately often was user-made noise: Two thirds of the participants (N=11) talked or otherwise made sounds during gameplay.

7.6.3 *Discussion*

We believe that two factors mainly contribute to the fact that a large percentage of users made noise during recording. First, while we mentioned that the applications was a game which also happened to record noise levels during gameplay, participants were instructed to focus on playing the game. Therefore, we believe that they foremost perceived the application as a game and not as a measurement tool.

Second, participants were only asked to play each the tutorial as well as the main level exactly once. Since most of the observed instances of making noises were sounds of frustration or surprise upon death or when struggling with an obstacle, their frequency may be lower if players are a bit more familiar with the game. Also, we believe that the presence of a study instructor may have actually acted as a catalyst: If there were no-one to hear the people vent their frustration, they might not have expressed it out loud. However, this would need to be confirmed in a future study. Either way, a technical approach to deal with this may be to remove the parts recorded directly after player death from the analysis.

A final point worth mentioning is that of design effort. While we have shown that it is clearly possible to aid the correct execution of sensing tasks through the use of games, designers should balance the effort that goes into the creation of a game with the expected benefits regarding the quality of the collected data. For smaller studies, other approaches like e.g. training (see section 6.4 for an overview) may be more sensible that creating a game. On a larger scale, such as the PS scenarios that are the basis for this work, the effort may become less significant in perspective.

7.7 CONCLUSION

In this chapter, we presented the notion of *Sensified Gaming*, which proposes not to think of gamifying Participatory Sensing applications but rather embedding Participatory Sensing tasks into games that can support them. We highlighted the potential of exploiting game design elements beyond creating user engagement and presented a collection of game design elements that can be used to identify or design suitable games. For this, we identified four core tasks from participatory environmental sensing and sensor networks research, reviewed hundreds of game design elements from different collections and mapped our selection of 63 game design patterns to the core tasks.

As an example to evaluate this novel concept, we built the game *SpaceMaze* that opportunistically performs ambient noise level measurements during gameplay. The evaluation of its performance in terms of error reduction shows the validity of our approach.

8 Conclusion and Outlook

This chapter concludes this dissertation on distributed, low-cost, non-expert fine dust sensing with smartphones. After a summary of the research and its contributions, we provide an outlook on future work.

8.1 Summary

The paradigm of computing is changing. Computers are increasingly pervading our daily lives and the environments we live in. Through the ubiquity of smartphones and the rise of the Internet of Things (IoT), more and more sensor-equipped, always-connected devices surround us, that have the potential to let us see the world in an unprecedented manner. Driven by this, the paradigm of environmental sensing is changing as well.

This dissertation deals with the question of how Particulate Matter (PM) can be measured using low-cost sensors and with high resolution, both spatially and temporally. To this end, we have approached several challenges, ranging from the identification of suitable sensing scenarios, over the design of low-cost instrumentation to addressing human factors and incentivizing correct sensing behavior. These are all aspects that relate to the data collection, respectively its quality, in Participatory Sensing (PS). Addressing these challenges together in a wholistic way is important to satisfy the dependencies between them.

The main contributions of this dissertation are:

- We showed that meaningful Particulate Matter (PM) measurements are possible with low-cost Commercial-of-the-shelf (COTS) dust sensors.

- We presented the design of the first handheld environmental monitor to include a COTS PM sensor for mobile and participatory environmental sensing.

- We successfully adapted the light-scattering measurement principle of low-cost sensors to camera smartphones using a self-designed passive external hardware add-on.

- We extended the existing light-scattering principle by changing both the optical layout as well as the image analysis algorithms so that counting individual particles at realistic concentration levels becomes possible. Since the sensor is still based on light-scattering, this approach has the potential to allow simultaneous size segregation and particle counting in software in the future.

- We presented a robust signal processing algorithm that can be used to reconstruct the "true" signal from signals that are heavily afflicted by systematic error. The only prerequisite of the algorithm is that the phenomenon can be modeled ad a Poisson process. We show that this makes it applicable to other phenomena besides PM by applying it to data from low cost gas sensors.

- We extended existing multi-hop calibration algorithms with several measures to protect the identity and location privacy of the participants.

- We empirically created a taxonomy of the errors made by non-expert users when measuring environmental phenomena with smartphones and surveyed possible approaches to mitigate or avoid them.

- We conducted a large field study with > 500 participants in which we compared the effect of instructions vs. technical measures to address user error in smartphone-based noise level sensing. Our results show that technical measures without explanation notably reduce the perceived usability and the combination of technology and instructions achieves a significant reduction in observed error rates while not affecting the user experience negatively.

- We introduced the notion of *Sensified Gaming* as a way of supporting smartphone-based environmental sensing in mobile games. We presented a set of design patterns for the design of such games and demonstrated the feasibility through the development and evaluation of a mobile minigame.

8.2 OUTLOOK

We have addressed a number of challenges that present themselves when attempting to perform Particulate Matter (PM) measurements using low-cost sensors and with high spatio-temporal resolution. Still, bringing it into reality entails a number of open challenges, several of which we attempt to approach in the recently started project *SmartAQnet*. Parts of this section have previously been published as part of the project outline [43].

8.2.1 *Extending Existing Networks*

While it has become clear that air quality monitoring will change fundamentally in the future [218], it is unlikely — and also counterproductive — that existing networks will be replaced completely by distributed measurement networks. Instead, in order to pursue this new generation of air quality monitoring consistently, the existing measurement stations should be supplemented with a narrow, heterogeneous sensing network, a part of which could also be Participatory Sensing (PS) activities.

A sensible combination of technology requires both the integration of data from existing high-end devices and low-cost sensors, as well as possibly the development of new devices in between. The challenges in the development of such new sensor systems include weighing low investment costs and mobility against high precision, long-term stability and a high tolerance to environmental influences (e.g. temperature, pressure and humidity).

In order to integrate a wide range of measurements with instruments of various price and quality classes, aspects such as network capability and smartphone connection must be addressed as well. This is e.g. a prerequisite for the real-world validation of distributed calibration approaches like the one presented in this work.

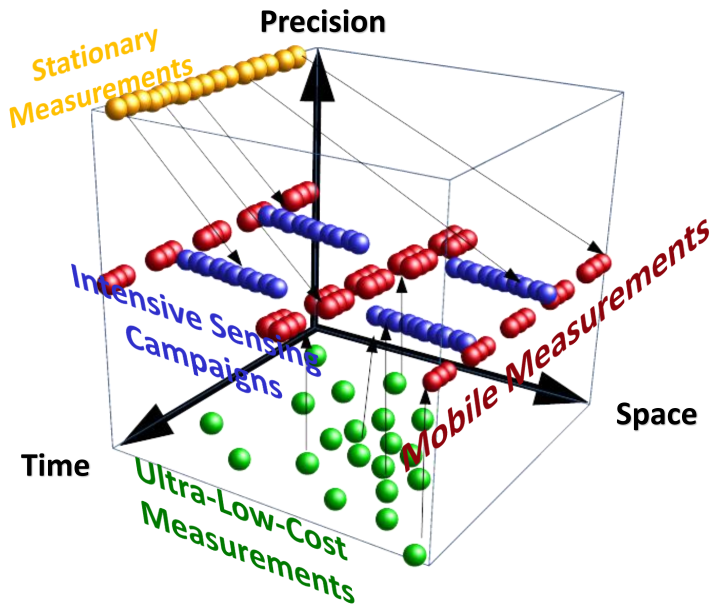


Figure 74.: One challenge for future air quality monitoring networks is the integration of heterogeneous measurement technology into one joint sensing grid [43].

8.2.2 *From Mobile Sensing to Big Data Analytics*

Researching under which conditions low-cost sensor technology (operated by non-experts) can be used to generate valuable large high-resolution data sets is not a question that can be answered with classical evaluation of individual sensors, e.g. through comparison measurements. Instead, it is expected that the fusion with other data will enable us to better understand relative influences from weather, traffic or building situations. This requires appropriate big data analytics for quality improvement and model validation. Research questions include novel algorithms, e.g. for central calibration of distributed sensors, verification of data sources, source appointment, etc. Figure 75 shows the data architecture of project *SmartAQnet*, which implements a complete Internet of Things (IoT) stack using state-of-the-art SmartData technology [43]. The underlying software architecture is a so-called kappa architecture, in which live data as well as historical data can be integrated continuously from constantly growing data sources.

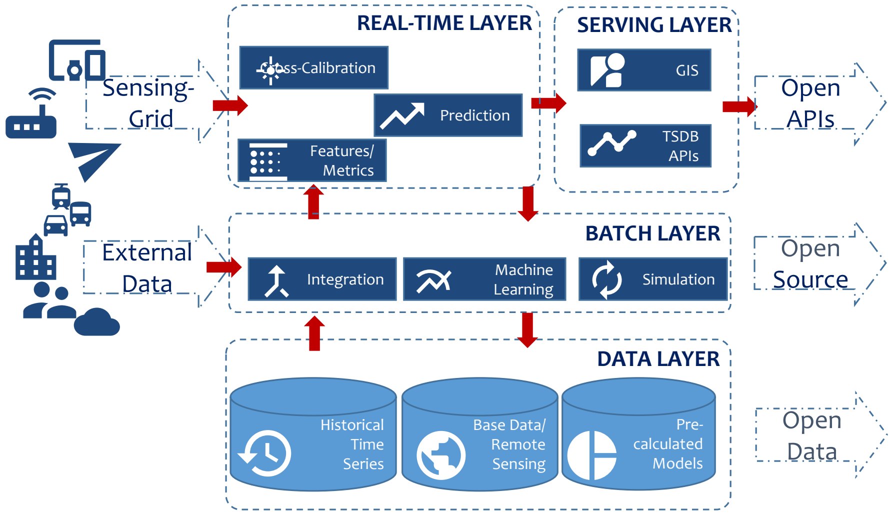


Figure 75.: Architecture of the *SmartAQnet* data management system [43].

8.2.3 *Closing the Loop*

Ultimately, the challenge lies in the development of a wholistic overall system for the recording, visualization and prediction of the spatial distribution of air pollutants in urban atmospheres that is relevant to the citizens. Important aspects concerning the sustainability of such a system is an open and participatory approach, which requires awareness and active participation ([49]). To ensure long-term recording and provision of the data, citizens should be able to record data and feed it into the systems, e.g. via their smartphones, which in turn could for instance provide more accurate air quality information for the user's location or movement route. Challenges that have not been in the focus of this work include the development of appropriate visualization schemes that are capable of conveying additional information (such as the measurement uncertainty), as well as the design of novel applications and services based on the obtained data.

A Own Publications

This appendix lists all peer-reviewed publications that were published in the course of the author's PhD studies. Aside from the work presented in-depth in this dissertation, this also includes individual publications on side-projects as well as contributions to colleagues' research.

Most notable among these is the design and evaluation of a usable security system that provides a WiFi access point, the signal of which is physically constrained to the surface of a 2D waveguide sheet, which can be used for usable WiFi authentication [41], [44].

Furthermore, contributions to measures that aim at establishing research-oriented teaching at a much earlier stage than in traditional university teaching were made [38].

Other activities in the course of the author's doctoral research include contributions to colleagues' work on mobile activity recognition [15]–[17], [32], [105], urban analytics [36], [68], [69] and vibro-tactile feedback [77], [180], [181].

full list of publications

2017

M. Beigl, M. Berning, and M. Budde, "Experiences and Failures from Two Decades in Embedded System Design," in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 2017. doi: 10.1145/3123024.3124393.

M. Budde, A. Exler, T. Riedel, M. Beigl, and A. Schankin, "Lessons from Failures in Designing and Conducting Experimental Studies – a Brief Anecdotal Tutorial," in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Sym-*

posium on Wearable Computers, 2017. DOI: 10.1145/3123024.3124392.

M. Budde, T. Riedel, M. Beigl, K. Schäfer, S. Emeis, J. Cyrys, J. Schnelle-Kreis, A. Philipp, V. Ziegler, H. Grimm, and T. Gratza, "SmartAQnet: Remote and In-Situ Sensing of Urban Air Quality," in *Proc. SPIE 10424, Remote Sensing of Clouds and the Atmosphere XXII, 104240C*, 2017. DOI: 10.1117/12.2282698.

M. Budde, A. Schankin, J. Hoffmann, M. Danz, T. Riedel, and M. Beigl, "Participatory Sensing or Participatory Nonsense? — Mitigating the Effect of Human Error on Data Quality in Citizen Science," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 1, no. 3, 2017. DOI: 10.1145/3131900.

V. Diener, M. Beigl, M. Budde, and E. Pescara, "VibrationCap: Studying Vibrotactile Localization on the Human Head with an Unobtrusive Wearable Tactile Display," in *21st International Symposium on Wearable Computers (ISWC 2017)*, 2017, pp. 82–89. DOI: 10.1145/3123021.3123047.

J.-F. Markert, M. Budde, G. Schindler, M. Klug, and M. Beigl, "Privacy-Preserving Collaborative Blind Macro-Calibration of Environmental Sensors in Participatory Sensing," *EAI Endorsed Transactions on the Internet of Things*, vol. 3, 10 2017, To appear.

L. Müller, M. Budde, N. Weibel, E. A. Spencer, M. Beigl, and D. Norman, "Learning from Failure: Designing for Complex Sociotechnical Systems," in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 2017, pp. 988–991. DOI: 10.1145/3123024.3124460.

E. Pescara, A. Wolpert, M. Budde, A. Schankin, and M. Beigl, "LifeTact - Utilizing Smartwatches as Tactile Heartbeat Displays in Video Games," in *16th International Conference on Mobile and Ubiquitous Multimedia (MUM 2017)*, 2017. DOI: 10.1145/3152832.3152863.

2016

M. Budde, "Towards Distributed, Low-Cost, Non-Expert Fine Dust Sensing with Smartphones," in *Doctoral Colloquium @ 6th International Conference on the Internet of Things (IoT 2016)*, 2016.

M. Budde and M. Beigl, "Advances in Smartphone-based Fine Dust Sensing," in *II International Conference on Atmospheric Dust – DUST 2016*, ser. Scientific Research Abstracts, vol. 5, 2016, p. 23.

M. Budde, S. Grebing, E. Burger, M. Kramer, B. Beckert, M. Beigl, and R. Reussner, "Praxis der Forschung: Eine Lehrveranstaltung des forschungsnahen Lehrens und Lernens in der Informatik am KIT," *Neues Handbuch Hochschullehre*, no. 74, pp. 55–79, 2016.

M. Budde, M. Köpke, and M. Beigl, "Design of a Light-scattering Particle Sensor for Citizen Science Air Quality Monitoring with Smartphones: Tradeoffs and Experiences," *ProScience*, vol. 3, no. 2nd International Conference on Atmospheric Dust – DUST2016, pp. 13–20, 2016. DOI: 10.14644/dust.2016.003. [Online]. Available: http://www.scientevents.com/proscience/download/design-of-a-light-scattering-particle-sensor-for-citizen-science-air-quality-monitoring-with-smartphones-tradeoffs-and-experiences/?wpdmdl=279.

M. Budde, R. Öxler, M. Beigl, and J. Holopainen, "Sensified Gaming – Design Patterns and Game Design Elements for Gameful Environmental Sensing," in *13th International Conference on Advances in Computer Entertainment Technology (ACE2016)*, ACM, 2016. DOI: 10.1145/3001773.3001832.

J. De Melo Borges, M. Budde, O. Peters, T. Riedel, and M. Beigl, "Towards Two-Tier Citizen Sensing," in *2nd IEEE International Smart Cities Conference (ISC2-2016)*, 2016. DOI: 10.1109/ISC2.2016.758077.

J. De Melo Borges, M. Budde, O. Peters, T. Riedel, A. Schankin, and M. Beigl, "EstaVis: A Real-World Interactive Platform for Crowdsourced Visual Urban Analytics," *Proceedings of the Second International Conference on IoT in Urban Space - UrbIoT'16*, vol. To Appear, 2016.

J.-F. Markert, M. Budde, G. Schindler, M. Klug, and M. Beigl, "Private Rendezvous-based Calibration of Low-Cost Sensors for Participatory Environmental Sensing," in *2nd EAI International Conference on IoT in Urban Space (UrbIoT'16)*, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2962754.

E. Pescara, M. Beigl, and M. Budde, "RüttelFlug – A Wrist-Worn Sensing Device for Tactile Vertical Velocity Perception in 3D-Space," in *2016 ACM International Symposium on Wearable Computers (ISWC'16)*, ACM, 2016, pp. 172–175. DOI: 10.1145/2971763.2971795.

2015

M. Berning, M. Budde, T. Riedel, and M. Beigl, "bPart - A Small and Versatile Bluetooth Low Energy Sensor Platform for Mobile Sensing," in *13th International Conference on Mobile Systems, Applications and Services (MobiSys'15)*, 2015. DOI: 10.1145/2742647.2745903.

M. Budde, M. Köpke, and M. Beigl, "Robust In-situ Data Reconstruction from Poisson Noise for Low-cost, Mobile, Non-expert Environmental Sensing," in *19th International Symposium on Wearable Computers (ISWC'15)*, 2015. DOI: 10.1145/2802083.2808406.

M. Scholz, L. Kohout, M. Horne, M. Budde, M. Beigl, and M. Youssef, "Device-Free Radio-based Low Overhead Identification of Subject Classes," in *2nd Workshop on Physical Analytics (WPA-15), co-located with ACM Mobisys*, 2015.

2014

M. Budde, J. De Melo Borges, S. Tomov, T. Riedel, and M. Beigl, "Improving Participatory Urban Infrastructure Monitoring through Spatio-Temporal Analytics," in *3rd ACM SIGKDD International Workshop on Urban Computing (UrbComp'14)*, 2014.

——, "Leveraging Spatio-Temporal Clustering for Participatory Urban Infrastructure Monitoring," in *The First International Conference on IoT in Urban Space (UrbIoT'14)*, 2014.

M. Budde, T. Riedel, M. Köpke, M. Berning, and M. Beigl, "A Comparative Study to Evaluate the Usability of Context-based Wi-Fi Access Mechanisms," in *16th International Conference on Human-Computer Interaction (HCI International 2014)*, 2014, ISBN: 978-3-319-07445-0. DOI: 10.1007/978-3-319-07446-7_44.

M. Budde, L. Zhang, and M. Beigl, "Challenges and Approaches for Low-Cost Particulate Matter Sensing in Smart Cities," in

I International Conference on Atmospheric Dust – DUST 2014, ser. Scientific Research Abstracts, vol. 3, 2014, p. 55.

——, "Distributed, low-cost particulate matter sensing: scenarios, challenges, approaches," *ProScience*, ProScience Conference Proceedings, vol. 1, no. First International Conference on Atmospheric Dust (DUST 2014), pp. 230–236, 2014, ISSN: 2283-5954. DOI: 10.14644/dust.2014.038.

2013

M. Budde, P. Barbera, R. El Masri, T. Riedel, and M. Beigl, "Retrofitting Smartphones to be Used as Particulate Matter Dosimeters," in *17th International Symposium on Wearable Computers (ISWC'13)*, 2013, pp. 139–140, ISBN: 978-1-4503-2127-3. DOI: 10.1145/2493988.2494342.

M. Budde, M. Berning, C. Baumgärtner, F. Kinn, T. Kopf, S. Ochs, F. Reiche, T. Riedel, and M. Beigl, "Point&Control - Interaction in Smart Environments: You Only Click Twice," in *International Joint Conference on Pervasive and Ubiquitous Computing (Ubicomp'13), Adjunct Proceedings*, 2013, pp. 303–306, ISBN: 978-1-4503-2215-7. DOI: 10.1145/2494091.2494184. [Online]. Available: http://www.youtube.com/watch?v=RqlCYBIUMos.

M. Budde, R. El Masri, T. Riedel, and M. Beigl, "Enabling Low-Cost Particulate Matter Measurement for Participatory Sensing Scenarios," in *12th International Conference on Mobile and Ubiquitous Multimedia (MUM 2013)*, 2013. DOI: 10.1145/2541831.2541859.

M. Budde, M. Köpke, M. Berning, T. Riedel, and M. Beigl, "Using a 2DST Waveguide for Usable, Physically Constrained Out-of-Band Wi-Fi Authentication," in *International Joint Conference on Pervasive and Ubiquitous Computing (Ubicomp'13)*, 2013, pp. 221–224, ISBN: 978-1-4503-1770-2. DOI: 10.1145/2493432.2494264.

Y. Ding, P. Goncalves Da Silva, M. A. Neumann, M. Budde, L. Zhang, and M. Beigl, "A Control Loop Approach for Integrating The Future Decentralized Power Markets and Grids," in *4th IEEE International Conference on Smart Grid Communications (SmartGridComm 2013)*, 2013. [Online]. Available: http://www.youtube.com/watch?v=bFx0e6EybL4.

M. Hauber, A. Bachmann, M. Budde, and M. Beigl, "jActivity: Supporting Mobile Web Developers with HTML5/JavaScript based Human Activity Recognition," in *12th International Conference on Mobile and Ubiquitous Multimedia (MUM 2013)*, 2013. DOI: 10.1145/2541831.2541873.

2012

M. Budde, M. Berning, M. Busse, T. Miyaki, and M. Beigl, "Handheld Particulate Matter Measurements with COTS Sensors," in *10th International Conference on Pervasive Computing (Pervasive 2012)*, 2012.

——, "The TECO Envboard: a Mobile Sensor Platform for Accurate Urban Sensing - and More," in *9th International Conference on Networked Sensing Systems*, IEEE, 2012, pp. 1–2, ISBN: 978-1-4673-1784-9. DOI: 10.1109/INSS.2012.6240573.

M. Budde, M. Busse, and M. Beigl, "Investigating the Use of Commodity Dust Sensors for the Embedded Measurement of Particulate Matter," in *9th International Conference on Networked Sensing Systems (INSS 2012)*, IEEE, 2012, pp. 1–4, ISBN: 978-1-4673-1784-9. DOI: 10.1109/INSS.2012.6240545.

2011

M. Berchtold, H. Günther, M. Budde, and M. Beigl, "Scheduling for a Modular Activity Recognition System to Reduce Energy Consumption on SmartPhones," in *24th International Conference on Architecture of Computing Systems (ARCS 2011), Second Workshop on Context-Systems Design, Evaluation and Optimisation (CoSDEO 2011)*, VDE Verlag, 2011. [Online]. Available: http://www.tu-braunschweig.de/Medien-DB/iti/cosdeo2011.pdf.

M. Budde, M. Berchtold, and M. Beigl, "Activity Recognition on Mobile Phones - Why do we need it and how can it be done?" In *9th International Conference on Pervasive Computing (Pervasive 2011)*, Jun. 2011.

Y. Ding, N. Namatame, T. Riedel, T. Miyaki, M. Budde, and M. Beigl, "SmartTecO: Context-Based Ambient Sensing and Monitoring for Optimizing Energy Consumption," in *International Conference on Autonomic Computing (Poster presentation)*, Jun. 2011, pp. 169–170, ISBN: 978-1-4503-0607-2. DOI: 10.1145/1998582.1998612.

S. Sigg, M. Budde, Y. Ji, and M. Beigl, "Entropy of Audio Fingerprints for Unobtrusive Device Authentication," in *The 7th International and Interdisciplinary Conference on Modeling and Using Context*, M. Beigl, H. Christiansen, T. R. Roth-Berghofer, A. Kofod-Petersen, K. R. Coventry, and H. R. Schmidtke, Eds., ser. Lecture Notes in Computer Science, vol. 6967, Karlsruhe, Germay: Springer Berlin Heidelberg, 2011, pp. 296–299, ISBN: 978-3-642-24278-6. DOI: 10.1007/978-3-642-24279-3. [Online]. Available: http://www.springerlink.com/content/d751363263v73k04/.

2010

M. Berchtold, M. Budde, D. Gordon, H. R. Schmidtke, and M. Beigl, "ActiServ : Activity Recognition Service for Mobile Phones," in *14th IEEE International Symposium on Wearable Computers (ISWC 2010)*, vol. i, Oct. 2010, pp. 1–8, ISBN: 978-1-4244-9046-2. DOI: 10.1109/ISWC.2010.5665868.

M. Berchtold, M. Budde, H. R. Schmidtke, and M. Beigl, "An Extensible Modular Recognition Concept that Makes Activity Recognition Practical," in *33rd Annual German Conference on Advances in Artificial Intelligence (KI 2010)*, Springer Berlin / Heidelberg, Sep. 2010, pp. 400–409.

M. Scholz, L. Ramirez, S. Denef, M. Betz, T. Dyrks, P. Scholl, M. Busse, M. Stoetzer, M. Budde, M. Berning, D. Shishkova, T. Riedel, and M. Beigl, "A MVC Prototype for the landmarke Firefighter Navigation System," in *International Internet of Things Conference (IOT 2010)*, 2010.

List of Symbols

A	The number of frames used to calculate $\overline{n}$. 59
D	Mean particle diameter. 59
E	Relative statistical error. 59
$PM_{(10-2.5)}$	(Inhalable) Coarse Particles, approximately $PM_{10} - PM_1$. 11, 88–90
PM_x	Particulate matter which passes through a size-selective inlet with a 50 % efficiency cut-off at x µm aerodynamic diameter. xxi, 10–12, 14, 16, 17, 28, 29, 31, 38, 39, 45–52, 86, 87, 89, 90, 99, 203
T	Sampling duration. 58, 60, 76, 83
V	Volume of the measurement chamber. 59
$\langle\sigma_{blur}\rangle$	Number of standard deviations used in Gaussian blur filter. 83, 87, 88
$\overline{n}_e$	The effective $\overline{n}$ when combining multiple measurements. 60
$\overline{n}$	Mean number of counted particles in the chamber. 58–60, 74, 203
ρ_i	Mass concentration of dust particle in the athmospheric air. 59
ρ	Particle density. 59, 86
σ_n	Standard deviation of the counted particles. 59, 74
$\theta_{b/w}$	Binarization threshold. 83, 87, 88
$\theta_{contour}$	Threshold for contour detection, i.e. minimum area at which patches are counted as contour. 83, 87, 88
c_v	Coefficient of variation. 59, 60
$c_{v,e}$	Effective coefficient of variation using multiple measurements. 60, 61
$d_{aerodynamic}$	Aerodynamic particle diameter. 86
$d_{geometric}$	Geometric particle diameter. 86
f	Measurement rate. 60, 73
h	Image height. 83
n_i	Number of particles seen in the i-th picture frame. 58, 59

r_{fps}	Video framerate. 83
r_{learn}	Learning rate of MOG2 background subtraction algorithm. 83, 84, 87, 88
t	Time. 83
w_{tf}	Window size for Poisson Particle Detection (PPD). 74
w	Image width. 83
B	image size. 65
B'	effective image size. 65
G	detector size. 65
α	aperture angle. 65
α'	effective aperture angle. 65
b	image distance. 65
f	the focal length of a lens. 63
g	detector distance. 65
k	camera distance. 64, 65

Acronyms and Abbreviations

2D	2-dimensional. 63
3D	3-dimensional. 34, 92
ABS	Acrylonitrile Butadiene Styrene. 67
ACE	International Conference on Advances in Computer Entertainment Technology. 161
AD	Aerodynamic Diameter. 10, 86, 203
ADC	Analog-to-Digital Converter. 39
AERONET	Aerosol Robotic Network. 20
AOT	Aerosol Optical Thickness. 13, 19, 20
API	Application Programming Interface. 34
APS	Aerodynamic Particle Sizer. 86, 205
APS 3321	TSI Aerodynamic Particle Sizer (APS) Spectrometer Model 3321. 85, 86
AQEG	Air Quality Expert Group. 25
AVR	Alf and Vegard's RISC processor (MCU architecture, commonly believed to be named after its developers). 33
BAM	Beta Attenuation Monitoring. 15
BC	Black Carbon. 11, 17–19, 53
BGR	Default color model in OpenCV, which is RGB with reversed channel order. 83
BT	Bluetooth. 34, 55
CCD	Charge-Coupled Device. 18
CDPC	Contour Detection Particle Counting. 83–85, 87, 88, 90, 92
CMOS	Complementary Metal–Oxide–Semiconductor. 18
CO	Carbon monoxide. 34
CO_2	Carbon dioxide. 34
COTS	Commercial-of-the-shelf. 5, 7, 27, 30, 32, 34, 35, 55, 97, 187, 188

CPC Condensation Particle Counter. 17
CR2032 Lithium Manganese Dioxide ($LiMnO_2$) coin cell battery. 61
CV Coefficient of Variation. 59, 60

dBA Decibel A-weighting. 34
DIN Deutsches Institut für Normung (German Institute for Standardization). 70, 148
DIY Do-it-yourself. 125, 132
DMA Dynamic Mobility Analyzer. 17
DN7C3CA006 Sharp DN7C3CA006 PM2.5 Sensor module. 31
DoLP Degree of Linear Polarization. 13, 20
DRX 8533 TSI DustTrak DRX 8533 Aerosol Monitor. 36, 37, 56
DSM501 Samyoung Hitech DSM 501 Dust Sensor. 31
DUST International Conference on Atmospheric Dust. 9, 53

EC European Commission. 10, 12
EDM180 Grimm Technologies Model EDM 180 PM Monitor. 42, 47
EPA U.S. Environmental Protection Agency. 10, 12, 16, 17, 25, 26
EU European Union. 12

FAQ Frequently Asked Questions. 133
FBAR Film Bulk Acoustic Resonator. 13, 16

GAM Generalized Additive Models. 13, 18
GAW Global Atmosphere Watch Programme. 85
GMM Gaussian Mixture Models. 84
GP2Y1010 Sharp GP2Y1010 Dust Sensor. 30, 31, 33–37, 42, 47, 48, 55, 56, 66, 67
GPS Global Positioning System. 21, 34, 107, 113

HCI Human Computer Interaction. 6, 71
HVS High Volume Sampler. 14, 15, 48

IDE Integrated Development Environment. 34
IMWUT Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 121

IoT	Internet of Things. 1, 187, 190
IR	Infrared. 16, 35, 42
ISO	International Organization for Standardization. 10, 86, 87, 148
ISWC	International Symposium on Wearable Computers. 53
LASER	Light Amplification by Stimulated Emission of Radiation. 16, 17, 42
LED	Light-Emitting Diode. 16, 56, 60, 68
LIDAR	Portmanteau of the words "light" and "radar", also considered short for Light Detection and Ranging or for Light Imaging Detection and Ranging. 13, 19
$LiMnO_2$	Lithium Manganese Dioxide. 206
LiPo	Lithium Ion Polymer. 34, 61, 63
LMS	Least Mean Squares. 81
LS	Least Squares. 101, 102
LVS	Low Volume Sampler. 14
MCU	Microcontroller Unit. 34, 55
MEMS	Micro Electrical Mechanical Systems. 16, 54
NAAQS	National Ambient Air Quality Standard. 10
$(NH_4)_2SO_4$	Ammonium sulfate. 86
NIRS	Near Infrared Spectroscopy. 123
NMSE	Normalized Mean Squared Error. 116
NO_x	Mono-nitrogen oxides. 34
O_3	Ozone. 34, 76, 78
OC	Organic Carbon. 11
OK	Open Knowledge Foundation. 132
OPC	Optical Particle Counter. 13
OPC-N2	Alphasense OPC-N2 Optical Particle Counter. 31, 32
OpenCV	Open Source Computer Vision. 83, 85
OTG	USB On-the-go. 62, 63
P2P	Peer-to-Peer. 95, 96
PAS	Photoacoustic Spectrometry. 13, 18

PBLs	Points, Badges, Leaderboards. 162, 166
PCB	Printed Circuit Board. 35
PDF	Portable Document Format. 127, 178
PEIR	Personal Environmental Impact Report. 28
PEM	Personal Environmental Monitor. 15, 29
PLA	Polylactic Acid or Polylactide. 67
PM	Particulate Matter. 2, 7, 9, 10, 14, 15, 20, 24, 26–30, 34, 48, 53–56, 76, 85, 95, 98, 121, 187–189
PPD	Poisson Particle Detection. 53, 74, 90, 204
PPD42NS	Shinyei Particle Sensor Model PPD42NS. 30
PPP	Poisson Point Process. 73
PS	Participatory Sensing. 2, 24–27, 106, 161, 184, 187, 189
QS	Quatified Self. 24
RGB	Additive color model based on the colors Red, Green and Blue. 56, 83
RH	Relative Humidity. 14, 34
RMSE	Root Mean Squared Error. 102, 103
RSD	Relative Standard Deviation. 59
SAW	Surface Acoustic Waves. 16
SD	Secure Digital. 34
SDS011	Nova Fitness SDS011 Dust Sensor. 31, 32, 98–100
SDS018	Nova Fitness SDS018 Dust Sensor. 31
SEQ4750	Leckel SEQ47/50 High Volume Sampler. 14, 15
SMPS	Scanning Mobility Particle Sizer. 17, 85, 86
SNR	Signal-to-Noise Ratio. 82
SP2	Single Particle Soot Photometer. 13, 19
SPEX	Spectropolarimeter for Planetary Exploration. 20
SPM	Suspended Particulate Matter. 9
$SpO_2\%$	Arterial Blood Oxygenation. 54
SPP	Spatial Poisson Process. 73
SPSS	Statistical Package for the Social Sciences. 148
SSID	Service Set Identifier. 134
SUS	System Usability Scale. 145, 147, 182
TEOM	Tapered Element Ocillating Micro-Balance. 13, 16

TROPOS	Leibniz Institute for Tropospheric Research e.V.. 85
TSP	Total Suspended Particles. 10, 31
TSPM	Total Suspended Particulate Matter. 10
UBA	Umweltbundesamt (Federal Environmental Agency of Germany). 85, 86
UEQ	User Experience Questionnaire. 145, 147, 151, 182
UFPs	Ultrafine Particles. 11
UK	United Kingdom. 25
USB	Universal Serial Bus. 34, 55, 62
UV	Ultraviolet. 24, 34
UX	User Experience. 7
VDA	Verband der Automobilindustrie (German Association of the Automotive Industry). 155
VOC	Volatile Organic Compounds. 34
WCCAP	World Calibration Center for Aerosol Physics. 85, 86, 98
WHO	World Health Organization. 12
WiFi	IEEE 802.11x protocol family. 21, 133
WMO	World Meteorological Organization. 85
WRAC	Wide Range Aerosol Classifier. 14
WSN	Wireless Sensor Network. 2, 20, 25, 29, 30

Bibliography

[1] K. Aberer, S. Sathe, D. Chakraborty, A. Martinoli, G. Barrenetxea, B. Faltings, and L. Thiele, "OpenSense: Open Community Driven Sensing of Environment," in *Proceedings of the ACM SIGSPATIAL International Workshop on GeoStreaming*, ser. IWGS '10, San Jose, California: ACM, 2010, pp. 39–42, ISBN: 978-1-4503-0431-3. DOI: 10.1145/1878500.1878509.

[2] ——, "OpenSense: open community driven sensing of environment," in *ACM SIGSPATIAL Workshop (IWGS '10)*, San Jose, California, 2010, ISBN: 978-1-4503-0431-3. DOI: 10.1145/1878500.1878509. [Online]. Available: http://doi.acm.org/10.1145/1878500.1878509.

[3] Air Quality Expert Group (AQEG), "Methods for monitoring particulate concentrations," in *Particulate Matter in the United Kingdom*. 2005, pp. 125–154. [Online]. Available: https://uk-air.defra.gov.uk/assets/documents/reports/aqeg/ch5.pdf (visited on 11/02/2017).

[4] ——, *Particulate Matter in the United Kingdom*. 2005.

[5] H. Alagarai Sampath, R. Rajeshuni, and B. Indurkhya, "Cognitively Inspired Task Design to Improve User Performance on Crowdsourcing Platforms," in *32nd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '14, Toronto, Ontario, Canada: ACM, 2014, pp. 3665–3674, ISBN: 978-1-4503-2473-1.

[6] C. Alexander, S. Ishikawa, and M. Silverstein, *A Pattern Language: Towns, Buildings, Construction*. Oxford University Press, 1977, ISBN: 0195019199.

[7] Alphasense Ltd. (2015). Alphasense User Manual OPC-N2 Optical Particle Counter. 072-0300 Issue 3, [Online]. Available: https://nettigo.pl/attachments/398 (visited on 11/15/2017).

[8] Apollo Electronics Co., Ltd. (2011). Apollo DSM501 dust sensor module, [Online]. Available: http://www.apollounion.com/Upload/DownFiles/DSM501%

20Technical%20Specifications.pdf (visited on 11/21/2017).

[9] Y. Arakawa and Y. Matsuda, "Gamification Mechanism for Enhancing a Participatory Urban Sensing: Survey and Practical Results," *Journal of Information Processing*, vol. 24, no. 1, pp. 31–38, 2016, ISSN: 1882-6652.

[10] E. Austin, I. Novosselov, E. Seto, and M. G. Yost, "Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor," *PLOS ONE*, vol. 10, no. 9, pp. 1–17, Sep. 2015. DOI: 10.1371/journal.pone.0137789. [Online]. Available: https://doi.org/10.1371/journal.pone.0137789.

[11] L. Balzano and R. Nowak, "Blind calibration of sensor networks," in *IPSN'07*, ACM, 2007, pp. 79–88.

[12] P. Barbera, "Physical Augmentation of Camera Phones for Optical Dust Sensing," Bachelor's Thesis, Karlsruhe Insititute of Technology (KIT), 2013.

[13] R. Bartle, "Hearts, clubs, diamonds, spades: Players who suit MUDs," *Journal of MUD research*, vol. 1, no. 1, p. 19, 1996.

[14] V. Bellotti and K. Edwards, "Intelligibility and accountability: human considerations in context-aware systems," *Human–Computer Interaction*, vol. 16, no. 2-4, pp. 193–212, 2001.

[15] M. Berchtold, M. Budde, D. Gordon, H. R. Schmidtke, and M. Beigl, "ActiServ : Activity Recognition Service for Mobile Phones," in *14th IEEE International Symposium on Wearable Computers (ISWC 2010)*, vol. i, Oct. 2010, pp. 1–8, ISBN: 978-1-4244-9046-2. DOI: 10.1109/ISWC.2010.5665868.

[16] M. Berchtold, M. Budde, H. R. Schmidtke, and M. Beigl, "An Extensible Modular Recognition Concept that Makes Activity Recognition Practical," in *33rd Annual German Conference on Advances in Artificial Intelligence (KI 2010)*, Springer Berlin / Heidelberg, Sep. 2010, pp. 400–409.

[17] M. Berchtold, H. Günther, M. Budde, and M. Beigl, "Scheduling for a Modular Activity Recognition System to Reduce Energy Consumption on SmartPhones," in *24th International Conference on Architecture of Computing Systems (ARCS 2011), Second Workshop on Context-Systems Design, Evaluation and Optimisation (CoSDEO 2011)*, VDE

Verlag, 2011. [Online]. Available: http://www.tu-braunschweig.de/Medien-DB/iti/cosdeo2011.pdf.

[18] S. Björk, *Gameplay Design Pattern Collection*, Accessed on May 4th, 2016, 2009. [Online]. Available: http://www.gameplaydesignpatterns.org/ (visited on 11/21/2017).

[19] S. Björk and J. Holopainen, *Patterns in Game Design*. Boston, MA: Charles River Media, 2004.

[20] ——, "Games and Design Patterns," in *The Game Design Reader: A Rules of Play Anthology*, K. Salen and E. Zimmerman, Eds., MIT Press, 2005.

[21] S. Björk, S. Lundgren, and J. Holopainen, "Game Design Patterns," in *Level Up*, M. Copier and J. Raessens, Eds., Wordware Publishing, 2003, pp. 180–193, ISBN: 1584503548.

[22] R. Boie and I. Cox, "An Analysis of Camera Noise," *Pattern Analysis and Machine Intelligence*, vol. 14, no. 6, 1992, ISSN: 0162-8828. DOI: http://doi.ieeecomputersociety.org/10.1109/34.141557.

[23] R. Bonney, J. L. Shirk, T. B. Phillips, A. Wiggins, H. L. Ballard, A. J. Miller-Rushing, J. K. Parrish, *et al.*, "Next steps for citizen science," 6178, vol. 343, American Association for The Advancement of Science, Washington, USA, Mar. 28, 2014. DOI: 10.1126/science.1251554.

[24] J. O. Borchers, "A pattern approach to interaction design," *Ai & Society*, vol. 15, no. 4, pp. 359–376, 2001.

[25] A. Bowser, D. Hansen, Y. He, C. Boston, M. Reid, L. Gunnell, and J. Preece, "Using gamification to inspire new citizen science volunteers," in *Proceedings of the First International Conference on Gameful Design, Research, and Applications*, ACM, 2013, pp. 18–25.

[26] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, Jan. 15, 2008. [Online]. Available: http://www.drdobbs.com/open-source/the-opencv-library/184404319 (visited on 11/06/2017).

[27] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[28] J. Brooke, "SUS – A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, 1996.

[29] M. Budde, "Towards Distributed, Low-Cost, Non-Expert Fine Dust Sensing with Smartphones," in *Doctoral Colloquium @ 6th International Conference on the Internet of Things (IoT 2016)*, 2016.

[30] M. Budde, P. Barbera, R. El Masri, T. Riedel, and M. Beigl, "Retrofitting Smartphones to be Used as Particulate Matter Dosimeters," in *17th International Symposium on Wearable Computers (ISWC'13)*, 2013, pp. 139–140, ISBN: 978-1-4503-2127-3. DOI: 10.1145/2493988.2494342.

[31] M. Budde and M. Beigl, "Advances in Smartphone-based Fine Dust Sensing," in *II International Conference on Atmospheric Dust – DUST 2016*, ser. Scientific Research Abstracts, vol. 5, 2016, p. 23.

[32] M. Budde, M. Berchtold, and M. Beigl, "Activity Recognition on Mobile Phones - Why do we need it and how can it be done?" In *9th International Conference on Pervasive Computing (Pervasive 2011)*, Jun. 2011.

[33] M. Budde, M. Berning, M. Busse, T. Miyaki, and M. Beigl, "Handheld Particulate Matter Measurements with COTS Sensors," in *10th International Conference on Pervasive Computing (Pervasive 2012)*, 2012.

[34] ——, "The TECO Envboard: a Mobile Sensor Platform for Accurate Urban Sensing - and More," in *9th International Conference on Networked Sensing Systems*, IEEE, 2012, pp. 1–2, ISBN: 978-1-4673-1784-9. DOI: 10.1109/INSS.2012.6240573.

[35] M. Budde, M. Busse, and M. Beigl, "Investigating the Use of Commodity Dust Sensors for the Embedded Measurement of Particulate Matter," in *9th International Conference on Networked Sensing Systems (INSS 2012)*, IEEE, 2012, pp. 1–4, ISBN: 978-1-4673-1784-9. DOI: 10.1109/INSS.2012.6240545.

[36] M. Budde, J. De Melo Borges, S. Tomov, T. Riedel, and M. Beigl, "Leveraging Spatio-Temporal Clustering for Participatory Urban Infrastructure Monitoring," in *The First International Conference on IoT in Urban Space (UrbIoT'14)*, 2014.

[37] M. Budde, R. El Masri, T. Riedel, and M. Beigl, "Enabling Low-Cost Particulate Matter Measurement for Participatory Sensing Scenarios," in *12th International Conference*

on Mobile and Ubiquitous Multimedia (MUM 2013), 2013. DOI: 10.1145/2541831.2541859.

[38] M. Budde, S. Grebing, E. Burger, M. Kramer, B. Beckert, M. Beigl, and R. Reussner, "Praxis der Forschung: Eine Lehrveranstaltung des forschungsnahen Lehrens und Lernens in der Informatik am KIT," *Neues Handbuch Hochschullehre*, no. 74, pp. 55–79, 2016.

[39] M. Budde, M. Köpke, and M. Beigl, "Robust In-situ Data Reconstruction from Poisson Noise for Low-cost, Mobile, Non-expert Environmental Sensing," in *19th International Symposium on Wearable Computers (ISWC'15)*, 2015. DOI: 10.1145/2802083.2808406.

[40] ——, "Design of a Light-scattering Particle Sensor for Citizen Science Air Quality Monitoring with Smartphones: Tradeoffs and Experiences," *ProScience*, vol. 3, no. 2nd International Conference on Atmospheric Dust – DUST2016, pp. 13–20, 2016. DOI: 10.14644/dust.2016.003. [Online]. Available: http://www.scientevents.com/proscience/download/design-of-a-light-scattering-particle-sensor-for-citizen-science-air-quality-monitoring-with-smartphones-tradeoffs-and-experiences/?wpdmdl=279.

[41] M. Budde, M. Köpke, M. Berning, T. Riedel, and M. Beigl, "Using a 2DST Waveguide for Usable, Physically Constrained Out-of-Band Wi-Fi Authentication," in *International Joint Conference on Pervasive and Ubiquitous Computing (Ubicomp'13)*, 2013, pp. 221–224, ISBN: 978-1-4503-1770-2. DOI: 10.1145/2493432.2494264.

[42] M. Budde, R. Öxler, M. Beigl, and J. Holopainen, "Sensified Gaming – Design Patterns and Game Design Elements for Gameful Environmental Sensing," in *13th International Conference on Advances in Computer Entertainment Technology (ACE2016)*, ACM, 2016. DOI: 10.1145/3001773.3001832.

[43] M. Budde, T. Riedel, M. Beigl, K. Schäfer, S. Emeis, J. Cyrys, J. Schnelle-Kreis, A. Philipp, V. Ziegler, H. Grimm, and T. Gratza, "SmartAQnet: Remote and In-Situ Sensing of Urban Air Quality," in *Proc. SPIE 10424, Remote Sensing of Clouds and the Atmosphere XXII, 104240C*, 2017. DOI: 10.1117/12.2282698.

[44] M. Budde, T. Riedel, M. Köpke, M. Berning, and M. Beigl, "A Comparative Study to Evaluate the Usability of Context-based Wi-Fi Access Mechanisms," in *16th International Conference on Human-Computer Interaction (HCI International 2014)*, 2014, ISBN: 978-3-319-07445-0. DOI: 10.1007/978-3-319-07446-7_44.

[45] M. Budde, A. Schankin, J. Hoffmann, M. Danz, T. Riedel, and M. Beigl, "Participatory Sensing or Participatory Nonsense? — Mitigating the Effect of Human Error on Data Quality in Citizen Science," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 1, no. 3, 2017. DOI: 10.1145/3131900.

[46] M. Budde, L. Zhang, and M. Beigl, "Challenges and Approaches for Low-Cost Particulate Matter Sensing in Smart Cities," in *I International Conference on Atmospheric Dust – DUST 2014*, ser. Scientific Research Abstracts, vol. 3, 2014, p. 55.

[47] ——, "Distributed, low-cost particulate matter sensing: scenarios, challenges, approaches," *ProScience*, ProScience Conference Proceedings, vol. 1, no. First International Conference on Atmospheric Dust (DUST 2014), pp. 230–236, 2014, ISSN: 2283-5954. DOI: 10.14644/dust.2014.038.

[48] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory Sensing," in *Workshop on World-Sensor-Web (WSW'06)*, 2006.

[49] J. A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," *Center for Embedded Network Sensing*, 2006.

[50] D. M. Butterfield and P. Quincey, "Measurement science issues relating to PM10 and PM2.5 airborne particles," National Physical Laboratory (NPL), NPL Report AS 15, 2007. [Online]. Available: http://publications.npl.co.uk/npl_web/pdf/as15.pdf (visited on 10/30/2017).

[51] V. L. Bychkovskiy, "Distributed In-Place Calibration in Sensor Networks," Master's thesis, University of California, 2003.

[52] V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak, "A Collaborative Approach to In-Place Sensor Calibration," in *IPSN'03*, 2003.

[53] M. Carminati, L. Pedalà, E. Bianchi, F. Nason, G. Dubini, L. Cortelezzi, G. Ferrari, and M. Sampietro, "Capacitive detection of micrometric airborne particulate matter for solid-state personal air quality monitors," *Sensors and Actuators A: Physical*, vol. 219, no. Supplement C, pp. 80–87, 2014, ISSN: 0924-4247. DOI: `https://doi.org/10.1016/j.sna.2014.09.003`. [Online]. Available: `http://www.sciencedirect.com/science/article/pii/S0924424714003987`.

[54] M. Carminati, P. Ciccarella, M. Sampietro, and G. Ferrari, "Single-Chip CMOS Capacitive Sensor for Ubiquitous Dust Detection and Granulometry with Sub-micrometric Resolution," in *Sensors: Proceedings of the Third National Conference on Sensors, February 23-25, 2016, Rome, Italy*, B. Andò, F. Baldini, C. Di Natale, G. Marrazza, and P. Siciliano, Eds. Cham: Springer International Publishing, 2018, pp. 8–18, ISBN: 978-3-319-55077-0. DOI: `10.1007/978-3-319-55077-0_2`.

[55] M. Carminati, G. Ferrari, and M. Sampietro, "Emerging miniaturized technologies for airborne particulate matter pervasive monitoring," *Measurement*, vol. 101, pp. 250–256, 2017.

[56] S. Castellini, B. Moroni, and D. Cappelletti, "PMetro: Measurement of urban aerosols on a mobile platform," *Measurement*, vol. 49, no. Supplement C, pp. 99–106, 2014, ISSN: 0263-2241. DOI: `https://doi.org/10.1016/j.measurement.2013.11.045`. [Online]. Available: `http://www.sciencedirect.com/science/article/pii/S0263224113005988`.

[57] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–90, Jan. 1981.

[58] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Determining Image Origin and Integrity Using Sensor Noise," *Information Forensics and Security*, vol. 3, no. 1, 2008.

[59] Y. Chen, A. Ebenstein, M. Greenstone, and H. Li, "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy," *Proceedings of the National Academy of Sciences*, vol. 110, no. 32, pp. 12 936–12 941, 2013.

[60] China Environmental Science Press, "Ambient Air Quality Standards," National Standard of the Peoples Republic of China GB 3095, 2012. [Online]. Available: http://kjs.mep.gov.cn/hjbhbz/bzwb/dqhjbh/dqhjzlbz/201203/W020120410330232398521.pdf (visited on 11/30/2017).

[61] S. Choi, N. Kim, H. Cha, and R. Ha, "Micro sensor node for air pollutant monitoring: Hardware and software issues," *Sensors*, vol. 9, no. 10, pp. 7970–7987, 2009.

[62] J. C. Chow and J. G. Watson. (1998). Guideline on Speciated Particulate Monitoring. N. Frank and J. Homolya, Eds. Draft 3, [Online]. Available: https://www3.epa.gov/ttnamti1/files/ambient/pm25/spec/drispec.pdf (visited on 11/05/2017).

[63] Z. Chowdhury, R. D. Edwards, M. Johnson, K. Naumoff Shields, T. Allen, E. Canuz, and K. R. Smith, "An inexpensive light-scattering particle monitor: field validation," *J. Environ. Monit.*, vol. 9, 10 2007. DOI: 10.1039/B709329M. [Online]. Available: http://dx.doi.org/10.1039/B709329M.

[64] D. Christin, A. Reinhardt, S. S. Kanhere, and M. Hollick, "A survey on privacy in mobile participatory sensing applications," *Journal of Systems and Software*, vol. 84, no. 11, pp. 1928–1946, 2011, Mobile Applications: Status and Trends, ISSN: 0164-1212. DOI: http://dx.doi.org/10.1016/j.jss.2011.06.073. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0164121211001701.

[65] C. Crumlish and E. Malone, *Designing social interfaces: Principles, patterns, and practices for improving the user experience*. O'Reilly Media, Inc., 2009.

[66] M. Danz, "Studies and Design of User Interaction for an Acoustical Object Recognition Program via Smartphone," Bachelor's Thesis, Karlsruhe Insititute of Technology (KIT), 2015.

[67] O. Davidsson, J. Peitz, and S. Björk, "Game design patterns for mobile games," *Project report to Nokia Research Center, Finland*, 2004.

[68] J. De Melo Borges, M. Budde, O. Peters, T. Riedel, and M. Beigl, "Towards Two-Tier Citizen Sensing," in *2nd IEEE International Smart Cities Conference (ISC2-2016)*, 2016. DOI: 10.1109/ISC2.2016.758077.

[69] J. De Melo Borges, M. Budde, O. Peters, T. Riedel, A. Schankin, and M. Beigl, "EstaVis: A Real-World Interactive Platform for Crowdsourced Visual Urban Analytics," *Proceedings of the Second International Conference on IoT in Urban Space - Urb-IoT'16*, vol. To Appear, 2016.

[70] L. Dekoninck, D. Botteldooren, L. I. Panis, S. Hankey, G. Jain, S. Karthik, and J. Marshall, "Applicability of a noise-based model to estimate in-traffic exposure to black carbon and particle number concentrations in different cultures," *Environment International*, vol. 74, pp. 89–98, 2015. DOI: 10.1016/j.envint.2014.10.002.

[71] P. P. Deng, "Dust Sensing Project," University of Melbourne, Report, 2010.

[72] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining gamification," in *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, ACM, 2011, pp. 9–15.

[73] A. K. Dey, "Understanding and Using Context," *Personal and Ubiquitous Computing*, vol. 5, no. 1, pp. 4–7, 2001, ISSN: 1617-4909.

[74] A. K. Dey and A. Newberger, "Support for Context-aware Intelligibility and Control," in *CHI '09*, ACM, 2009, pp. 859–868.

[75] S. Dey, N. Roy, W. Xu, R. R. Choudhury, and S. Nelakuditi, "Accelprint: Imperfections of accelerometers make smartphones trackable," in *Network and Distributed System Security Symposium*, 2014.

[76] M. Di Francesco, S. K. Das, and G. Anastasi, "Data Collection in Wireless Sensor Networks with Mobile Elements: A Survey," *ACM Trans. Sen. Netw.*, vol. 8, no. 1, 7:1–7:31, Aug. 2011, ISSN: 1550-4859. DOI: 10.1145/1993042.1993049.

[77] V. Diener, M. Beigl, M. Budde, and E. Pescara, "VibrationCap: Studying Vibrotactile Localization on the Human Head with an Unobtrusive Wearable Tactile Display," in *21st International Symposium on Wearable Computers (ISWC 2017)*, 2017, pp. 82–89. DOI: `10.1145/3123021.3123047`.

[78] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," Tech. Rep. 2004.

[79] D. W. Dockery, C. A. Pope, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris Jr, and F. E. Speizer, "An association between air pollution and mortality in six US cities," *New England Journal of Medicine*, vol. 329, no. 24, pp. 1753–1759, 1993.

[80] F. L. Doering, I. Paprotny, and R. M. White, "MEMS air-microfluidic sensor for portable monitoring of airborne particulates," in *2012 Solid-State Sensors, Actuators and Microsystems Workshop, Hilton Head 2012*, Transducer Research Foundation, 2012.

[81] "DRAFT Roadmap for Next Generation Air Monitoring," U.S. Environmental Protection Agency, Tech. Rep., Mar. 2013, DRAFT 3/8/13.

[82] P. Dutta, P. M. Aoki, N. Kumar, A. Mainwaring, C. Myers, W. Willett, and A. Woodruff, "Common sense: participatory urban sensing using a network of handheld air quality monitors," in *Proceedings of the 7th ACM conference on embedded networked sensor systems (SenSys '09)*, ACM, Berkeley, California, 2009, pp. 349–350, ISBN: 978-1-60558-519-2. DOI: `10.1145/1644038.1644095`.

[83] J. Duyzer, D. van den Hout, P. Zandveld, and S. van Ratingen, "Representativeness of air quality monitoring networks," *Atmospheric Environment*, vol. 104, pp. 88–101, 2015. DOI: `10.1016/j.atmosenv.2014.12.067`.

[84] R. M. El Masri, "Evaluierung von Low-Cost Staubmessungen für Urban Sensing Anwendungen," Master's thesis, Karlsruhe Insititute of Technology (KIT), 2013.

[85] B. Elen, J. Peters, M. Van Poppel, N. Bleux, J. Theunis, M. Reggente, and A. Standaert, "The Aeroflex: a bicycle for mobile air quality measurements," *Sensors*, vol. 13, no. 1, pp. 221–240, 2013.

[86] T. Erickson, "Lingua Francas for design: sacred places and pattern languages," in *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, ACM, 2000, pp. 357–368.

[87] European Commission. (1999). COUNCIL DIRECTIVE 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air, [Online]. Available: `http://www.irceline.be/~celinair/documents/EU_guidelines/so2_nox_pm10_pb_en.pdf` (visited on 10/31/2017).

[88] ——, (2008). DIRECTIVE 2008/50/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21 May 2008 on ambient air quality and cleaner air for Europe, [Online]. Available: `http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0050&from=EN` (visited on 10/31/2017).

[89] J. S. B. Evans, "The knowledge elicitation problem: a psychological perspective," *Behaviour & Information Technology*, vol. 7, no. 2, 1988.

[90] J. Feng, S. Megerian, and M. Potkonjak, "Model-based calibration for sensor networks," in *IEEE Sensors*, vol. 2, ieeexplore.ieee.org, Oct. 2003, 737–742 Vol.2.

[91] D. R. Flatla, C. Gutwin, L. E. Nacke, S. Bateman, and R. L. Mandryk, "Calibration games: making calibration tasks enjoyable by adding motivating game elements," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, ACM, 2011, pp. 403–412.

[92] Y. Fu and B. Hallberg, "A Personal Environment Monitoring System for Pulmonary Disease Management," in *iCBBE'10*. DOI: `10.1109/ICBBE.2010.5516510`.

[93] M. Gao, J. Cao, and E. Seto, "A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China," *Environmental Pollution*, vol. 199, no. Supplement C, pp. 56–65, 2015, ISSN: 0269-7491. DOI: `https://doi.org/10.1016/j.envpol.2015.01.013`. [Online]. Available: `http://www.sciencedirect.com/science/article/pii/S0269749115000160`.

[94] M. M. Gardiner, L. L. Allee, P. M. Brown, J. E. Losey, H. E. Roy, and R. R. Smyth, "Lessons from lady beetles: accuracy of monitoring data from US and UK citizen-science programs," *Frontiers in Ecology and the Environment*, vol. 10, no. 9, pp. 471–476, 2012.

[95] S. K. Gibb, "Volunteers Against Pollution," *Chemical & Engineering News (C&EN)*, vol. 93, no. 36, Sep. 2015.

[96] J. H. Gilliam and E. S. Hall, "Reference and Equivalent Methods Used to Measure National Ambient Air Quality Standards (NAAQS) Criteria Air Pollutants - Volume I," U.S. Environmental Protection Agency, EPA/600/R-16/139, 2016. [Online]. Available: `https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=321491` (visited on 11/03/2017).

[97] F. Gozzi, G. Della Ventura, and A. Marcelli, "Mobile monitoring of particulate matter: state of art and perspectives," *Atmospheric Pollution Research*, vol. 7, no. 2, pp. 228–234, 2016.

[98] Grimm Technologies, *Environmental Dust Monitors – EDM 180*. [Online]. Available: `http://wiki.grimm-aerosol.de/images/5/56/Datasheet_180-PLUS_EDM_ENG_V2p0.pdf` (visited on 11/21/2017).

[99] A. Gustafsson, J. Bichard, L. Brunnberg, O. Juhlin, and M. Combetto, "Believable environments: generating interactive storytelling in vast location-based pervasive games," in *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*, ACM, 2006, p. 24.

[100] R. Hagemann, U. Corsmeier, C. Kottmeier, R. Rinke, A. Wieser, and B. Vogel, "Spatial variability of particle number concentrations and NO x in the Karlsruhe (Germany) area obtained with the mobile laboratory 'AERO-TRAM'," *Atmospheric environment*, vol. 94, pp. 341–352, 2014.

[101] W.-c. Hao, J.-l. Liu, M.-h. Liu, and S.-t. He, "Development of a new surface acoustic wave based PM 2.5 monitor," in *Piezoelectricity, Acoustic Waves, and Device Applications (SPAWDA), 2014 Symposium on*, IEEE, 2014, pp. 52–55. DOI: `10.1109/SPAWDA.2014.6998524`.

[102] M. Harding, B. Knowles, N. Davies, and M. Rouncefield, "HCI, Civic Engagement & Trust," in *33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15, Seoul, Republic of Korea: ACM, 2015, pp. 2833–2842, ISBN: 978-1-4503-3145-6.

[103] D. Hasenfratz, "Enabling Large-Scale Urban Air Quality Monitoring with Mobile Sensor Nodes," Master's thesis, ETH Zürich, 2015. DOI: `10.3929/ethz-a-010361120`. [Online]. Available: `https://www.research-collection.ethz.ch/handle/20.500.11850/100110`.

[104] D. Hasenfratz, O. Saukh, and L. Thiele, "On-the-Fly Calibration of Low-Cost Gas Sensors," in *Wireless Sensor Networks (EWSN'12)*, ser. LNCS, vol. 7158, 2012, ISBN: 978-3-642-28168-6. DOI: `10.1007/978-3-642-28169-3_15`.

[105] M. Hauber, A. Bachmann, M. Budde, and M. Beigl, "jActivity: Supporting Mobile Web Developers with HTML5/JavaScript based Human Activity Recognition," in *12th International Conference on Mobile and Ubiquitous Multimedia (MUM 2013)*, 2013. DOI: `10.1145/2541831.2541873`.

[106] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *Information Theory*, 2006.

[107] S. Hinske, M. Lampe, C. Magerkurth, and C. Röcker, "Classifying pervasive games: on pervasive computing and mixed reality," *Concepts and technologies for Pervasive Games-A Reader for Pervasive Gaming Research*, vol. 1, p. 20, 2007.

[108] B. Holben, T. Eck, I. Slutsker, D. Tanré, J. Buis, A. Setzer, E. Vermote, J. Reagan, Y. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov, "AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization," *Remote Sensing of Environment*, vol. 66, no. 1, pp. 1–16, 1998, ISSN: 0034-4257. DOI: `https://doi.org/10.1016/S0034-4257(98)00031-5`. [Online]. Available: `http://www.sciencedirect.com/science/article/pii/S0034425798000315`.

[109] J. Holopainen, H. Korhonen, E. Ollila, V. Nenonen, S. Björk, J. Peitz, and O. Davidsson, "IPerG – Integrated Project on Pervasive Gaming, Work Package WP5: Design and Evaluation Deliverable D5.10: Design Kit: Massively

Multiplayer Mobile Phone Games," Public Project Report, Mar. 19, 2008.

[110] D. Holstius, "Monitoring Particulate Matter with Commodity Hardware," Doctoral Dissertation, University of California, Berkeley, 2014.

[111] D. M. Holstius, A. Pillarisetti, K. Smith, and E. Seto, "Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California," *Atmospheric Measurement Techniques*, vol. 7, no. 4, pp. 1121–1131, 2014.

[112] J. D. Hoyt and H. Wechsler, "Detection of human speech in structured noise," in *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, IEEE, vol. 2, 1994, pp. II–237.

[113] K. L. Huang, S. S. Kanhere, and W. Hu, "Are You Contributing Trustworthy Data?: The Case for a Reputation System in Participatory Sensing," in *13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, ser. MSWIM '10, New York, NY, USA: ACM, 2010, pp. 14–22.

[114] International Standardization Organization. (1995). ISO 7708:1995 Air quality – Particle size fraction definitions for health-related sampling. Last reviewed and confirmed in 2008, [Online]. Available: `https://www.iso.org/standard/14534.html` (visited on 10/30/2017).

[115] ——, (2009). ISO 23210:2009 Stationary source emissions – Determination of PM10/PM2,5 mass concentration in flue gas – Measurement at low concentrations by use of impactors. Last reviewed and confirmed in 2015, [Online]. Available: `https://www.iso.org/standard/53379.html` (visited on 10/30/2017).

[116] H.-C. Jetter, S. Gallacher, V. Kalnikaite, and Y. Rogers, "Suspicious boxes and friendly aliens: exploring the physical design of urban sensing technology," in *Proceedings of the First International Conference on IoT in Urban Space*, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014, pp. 68–73.

[117] J. Joss and A. Waldvogel, "Raindrop Size Distribution and Sampling Size Errors," *Atmospheric Sciences*, 3 1969. DOI: 10.1175/1520-0469(1969)026<0566:RSDASS>2.0.CO;2.

[118] M. Jovašević-Stojanović, A. Bartonova, D. Topalović, I. Lazović, B. Pokrić, and Z. Ristovski, "On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter," *Environmental Pollution*, vol. 206, pp. 696–704, 2015.

[119] E. Kanjo, "Noisespy: A real-time mobile phone platform for urban noise monitoring and mapping," *Mobile Networks and Applications*, vol. 15, no. 4, pp. 562–574, 2010.

[120] L. Kazemi and C. Shahabi, "TAPAS: Trustworthy privacy-aware participatory sensing," *Knowledge and information systems*, vol. 37, no. 1, pp. 105–128, 2013.

[121] M. I. Khadem and V. Sgarciu, "Dust Monitoring Systems," in *Proceedings of the ICSNC'11*, 2011, pp. 68–71, ISBN: 978-1-61208-166-3.

[122] S. Kim, J. Mankoff, and E. Paulos, "Sensr: Evaluating a Flexible Framework for Authoring Mobile Data-collection Tools for Citizen Science," in *Computer Supported Cooperative Work*, ser. CSCW '13, San Antonio, Texas, USA: ACM, 2013, ISBN: 978-1-4503-1331-5.

[123] S. Kim and E. Paulos, "inAir: Measuring and Visualizing Indoor Air Quality," in *Proceedings of the 11th International Conference on Ubiquitous Computing*, ser. UbiComp '09, Orlando, Florida, USA: ACM, 2009, pp. 81–84, ISBN: 978-1-60558-431-7. DOI: 10.1145/1620545.1620557.

[124] S. Kim, C. Robson, T. Zimmerman, J. Pierce, and E. M. Haber, "Creek Watch: Pairing Usefulness and Usability for Successful Citizen Science," in *SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11, ACM, 2011, pp. 2125–2134, ISBN: 978-1-4503-0228-9.

[125] S. F. King and P. Brown, "Fix my street or else: using the internet to voice local public service concerns," in *Proceedings of the 1st international conference on Theory and practice of electronic governance*, ACM, 2007, pp. 72–80.

[126] E. Kintisch, *How to Grow Your Own Army of Citizen Scientists*, Feb. 2011.

[127] S. Klakegg, C. Luo, J. Goncalves, S. Hosio, and V. Kostakos, "Instrumenting Smartphones with Portable NIRS," in *Adj. Proceedings UbiComp'16, Workshop in Ubiquitous Mobile Instrumentation (UbiMI)*, 2016.

[128] R.-D. Klein. (). RadioAcivity – RadioactivityCounter for mobile phones. Accessed April 17th, 2014., [Online]. Available: http://www.hotray-info.de/html/radioactivity.html (visited on 11/21/2017).

[129] J. Koetsier. (2016). Augmented Exercise: People Playing Pokémon Go Have Burned 340 Billion Calories, [Online]. Available: https://www.forbes.com/sites/johnkoetsier/2016/09/08/augmented-exercise-people-playing-pokemon-go-have-burned-340-billion-calories/ (visited on 11/21/2017).

[130] B. Kreimeier, *The Case for Design Patterns*, 2002. [Online]. Available: http://www.gamasutra.com/view/feature/132649/the%7B%5C_%7Dcase%7B%5C_%7Dfor%7B%5C_%7Dgame%7B%5C_%7Ddesign%7B%5C_%7Dpatterns.php.

[131] Y.-S. Kuo, S. Verma, T. Schmid, and P. Dutta, "Hijacking Power and Bandwidth from the Mobile Phone's Audio Interface," in *Proceedings of the First ACM Symposium on Computing for Development*, ser. ACM DEV '10, London, United Kingdom: ACM, 2010, 24:1–24:10, ISBN: 978-1-4503-0473-3. DOI: 10.1145/1926180.1926210. [Online]. Available: http://doi.acm.org/10.1145/1926180.1926210.

[132] E. Laan, D. Stam, R. Hoogeveen, G. Bazalgette Courreges-Lacoste, and E. Boom, "SPEX – A SPECTROPOLARIMETER FOR PLANETARY EXPLORATION ONBOARD THE EXOMARS ORBITER," in *4th International Interplanetary Probe Workshop (IPPW4)*, 2006.

[133] M. Laborde, P. Mertes, P. Zieger, J. Dommen, U. Baltensperger, and M. Gysel, "Sensitivity of the Single Particle Soot Photometer to different black carbon types," *Atmospheric Measurement Techniques*, vol. 5, no. 5, p. 1031, 2012. DOI: 10.5194/amt-5-1031-2012.

[134] F. Lamonaca, D. L. Carní, D. Grimaldi, A. Nastro, M. Riccio, and V. Spagnolo, "Blood oxygen saturation measurement by smartphone camera," in *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2015. DOI: 10.1109/MeMeA.2015.7145228.

[135] N. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. Campbell, "A survey of mobile phone sensing,"

IEEE Communications Magazine, vol. 48, no. 9, 2010, ISSN: 0163-6804.

[136] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in *Symposium of the Austrian HCI and Usability Engineering Group*, Springer, 2008, pp. 63–76.

[137] T. Le, R. Chartrand, and T. J. Asaki, "A variational approach to reconstructing images corrupted by Poisson noise," *Mathematical Imaging and Vision*, vol. 27, no. 3, 2007.

[138] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile data offloading: How much can WiFi deliver?" *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 2, pp. 536–550, 2013.

[139] J. Lelieveld, J. Evans, M. Fnais, D. Giannadaki, and A. Pozzer, "The contribution of outdoor air pollution sources to premature mortality on a global scale," *Nature*, vol. 525, no. 7569, pp. 367–371, 2015. DOI: 10.1038/nature15371.

[140] J. J. Li, B. Faltings, O. Saukh, D. Hasenfratz, and J. Beutel, "Sensing the air we breathe-the opensense zurich dataset," in *AAAI: Artificial Intelligence*, 2012.

[141] L. Li, Y. Zheng, and L. Zhang, "Demonstration abstract: PiMi air box: a cost-effective sensor for participatory indoor quality monitoring," in *Proceedings of the 13th international symposium on Information processing in sensor networks*, IEEE Press, 2014, pp. 327–328.

[142] A. Liberti, "Modern methods for air pollution monitoring," *Pure and Applied Chemistry*, vol. 44, no. 3, pp. 519–534, 1975.

[143] X.-D. Lu and D. L. Trumper, "Self-Calibration of On-Axis Rotary Encoders," *CIRP Annals - Manufacturing Technology*, vol. 56, no. 1, pp. 499–504, 2007.

[144] K. Luther, S. Counts, K. B. Stecher, A. Hoff, and P. Johns, "Pathfinder: An Online Collaboration Environment for Citizen Scientists," in *Human Factors in Computing Systems*, ser. CHI '09, Boston, MA, USA: ACM, 2009, pp. 239–248, ISBN: 978-1-60558-246-7.

[145] M. Mun, et al., "PEIR, the personal environmental impact report, as a platform for participatory sensing systems research," in *MobiSys'09*, 2009.

[146] C. Magerkurth, A. D. Cheok, R. L. Mandryk, and T. Nilsen, “Pervasive Games: Bringing Computer Entertainment Back to the Real World,” *Comput. Entertain.*, vol. 3, no. 3, pp. 4–4, Jul. 2005, ISSN: 1544-3574.

[147] N. Maisonneuve, M. Stevens, M. E. Niessen, and L. Steels, “NoiseTube: Measuring and mapping noise pollution with mobile phones,” in *Information Technologies in Environmental Engineering*, 2009, ISBN: 978-3-540-88351-7.

[148] N. Maisonneuve, M. Stevens, M. E. Niessen, P. Hanappe, and L. Steels, “Citizen noise pollution monitoring,” in *Digital Government Research (dg.o'09)*, 2009, pp. 96–103.

[149] N. Maisonneuve, M. Stevens, and B. Ochab, “Participatory noise pollution monitoring using mobile phones,” in *Information Polity*, vol. 15, 2010, pp. 51–71. DOI: 10.3233/IP-2010-0200.

[150] A. Manikonda, N. Zíková, P. K. Hopke, and A. R. Ferro, “Laboratory assessment of low-cost PM monitors,” *Journal of Aerosol Science*, vol. 102, pp. 29–40, 2016.

[151] A. Marczewski. (Sep. 24, 2013). Serious Games: Too Broad a Term to be Meaningful, [Online]. Available: https://www.gamified.uk/2013/09/24/serious-games-too-broad-a-term-to-be-meaningful/ (visited on 11/21/2017).

[152] J. F. Markert, “Design Process for Privacy-Aware Distributed Sensor Calibration,” Master's thesis, Karlsruhe Insititute of Technology (KIT), 2016.

[153] J.-F. Markert, M. Budde, G. Schindler, M. Klug, and M. Beigl, “Private Rendezvous-based Calibration of Low-Cost Sensors for Participatory Environmental Sensing,” in *2nd EAI International Conference on IoT in Urban Space (UrbIoT'16)*, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2962754.

[154] J.-F. Markert, M. Budde, G. Schindler, M. Klug, and M. Beigl, “Privacy-Preserving Collaborative Blind Macro-Calibration of Environmental Sensors in Participatory Sensing,” *EAI Endorsed Transactions on the Internet of Things*, vol. 3, 10 2017, To appear.

[155] I. G. Martí, L. E. Rodríguez, M. Benedito, S. Trilles, A. Beltrán, L. Díaz, and J. Huerta, "Mobile application for noise pollution monitoring through gamification techniques," in *Entertainment Computing-ICEC 2012*, Springer, 2012, pp. 562–571.

[156] S. Matyas, P. Kiefer, C. Schlieder, and S. Kleyer, "Wisdom about the Crowd: Assuring Geospatial Data Quality Collected in Location-Based Games," in *International Conference on Entertainment Computing (ICEC 2011)*, 2011, pp. 331–336.

[157] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review*, vol. 91, pp. 81–97, 1956.

[158] E. Miluzzo, N. D. Lane, A. T. Campbell, and R. Olfati-Saber, "CaliBree: A Self-calibration System for Mobile Sensor Networks," in *Distributed Computing in Sensor Systems*, ser. Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2008, pp. 314–331.

[159] D. Model and M. Zibulevsky, "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing*, vol. 86, no. 3, pp. 624–638, 2006, ISSN: 0165-1684. DOI: http://dx.doi.org/10.1016/j.sigpro.2005.05.033. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168405002252.

[160] M. Montola, "A ludological view on the pervasive mixed-reality game research paradigm," *Personal and Ubiquitous Computing*, vol. 15, no. 1, pp. 3–12, 2011, ISSN: 1617-4917.

[161] M. Montola, J. Stenros, and A. Waern, *Pervasive games: theory and design*. Morgan Kaufmann Publishers Inc., 2009.

[162] M. Montola, A. Waern, J. Kuittinen, and J. Stenros, "IPerG – Integrated Project on Pervasive Gaming, Work Package WP5: Design and Evaluation Deliverable D5.5: Ethics of Pervasive Gaming," Public Project Report, Oct. 13, 2016, Release date: October 13 2006.

[163] F. Münchbach, "Algorithms for particlulate matter concentration measurement with consumer-cameras," Bachelor's Thesis, Karlsruhe Insititute of Technology (KIT), 2015.

[164] L. A. Muratori, P. Salomoni, and G. Pau, "Feeling the pack: Strategies for an optimal participatory system to sense and recognize noise pollution," in *2011 IEEE International Conference on Consumer Electronics -Berlin (ICCE-Berlin)*, Sep. 2011, pp. 17–21.

[165] C. Nafis, *Air Quality Monitoring*, 2012. [Online]. Available: http://www.howmuchsnow.com/arduino/airquality/ (visited on 11/21/2017).

[166] A. Narayanan, N. Thiagarajan, M. Lakhani, M. Hamburg, and D. Boneh, "Location Privacy via Private Proximity Testing," in *NDSS*, 2011.

[167] NIDS Sensor Technology. (2004). Specifications: PS02C-PWM. P-PS02C-K01, [Online]. Available: https://www.icbanq.com/data/ICBShop/board/%EC%A0%9C%ED%92%88%EC%82%AC%EC%96%91%EC%84%9C_PS02C-PWM_.pdf (visited on 11/21/2017).

[168] ——, (2005). Specifications: PSX-01E, [Online]. Available: http://114.200.199.54/pdf/H-PSX-01E.pdf (visited on 11/30/2017).

[169] D. A. Norman, "Design Rules Based on Analyses of Human Error," *Commun. ACM*, vol. 26, no. 4, pp. 254–258, Apr. 1983, ISSN: 0001-0782.

[170] D. A. Norman and P. J. Stappers, "DesignX: Complex Sociotechnical Systems," *She Ji*, vol. 1, pp. 83–106, 2 2015.

[171] Nova Fitness Co., Ltd. (2015). SDS011 sensor. Version: V1.3, [Online]. Available: https://nettigo.pl/attachments/398 (visited on 11/15/2017).

[172] ——, (2015). SDS018 sensor. Version: V1.5, [Online]. Available: http://ecksteinimg.de/Datasheet/SDS018%20Laser%20PM2.5%20Product%20Spec%20V1.5.pdf (visited on 11/15/2017).

[173] D. Novick and K. Ward, "Why Don't People Read the Manual?" In *Int. Conference on Design of Communication*, ser. SIGDOC '06, Myrtle Beach, SC, USA: ACM, 2006, ISBN: 1-59593-523-1.

[174] S. O'Driscoll, B. D. MacCraith, and C. S. Burke, "A novel camera phone-based platform for quantitative fluorescence sensing," *Anal. Methods*, vol. 5, pp. 1904–1908, 8 2013. DOI: 10.1039/C3AY40116B. [Online]. Available: http://dx.doi.org/10.1039/C3AY40116B.

[175] R. J. Öxler, "Game Design Elemente für Ortsbasierte Participatory Sensing Spiele," Diploma Thesis, Karlsruhe Insititute of Technology (KIT), 2017.

[176] I. Paprotny, F. Doering, P. A. Solomon, R. M. White, and L. A. Gundel, "Microfabricated air-microfluidic sensor for personal monitoring of airborne particulate matter: Design, fabrication, and experimental results," *Sensors and Actuators A: Physical*, vol. 201, pp. 506–516, 2013. DOI: 10.1016/j.sna.2012.12.026.

[177] (). Particlecamp, [Online]. Available: http://particlecamp.org/blog/ (visited on 11/21/2017).

[178] J. Peitz, H. Saarenpää, and S. Björk, "Insectopia: exploring pervasive games through technology already pervasively available," in *Proceedings of the international conference on Advances in computer entertainment technology*, ACM, 2007, pp. 107–114.

[179] P. Pelegris, K. Banitsas, T. Orbach, and K. Marias, "A Novel Method to Detect Heart Beat Rate Using a Mobile Phone," *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, 2010, ISSN: 1557-170X. [Online]. Available: http://www.biomedsearch.com/nih/novel-method-to-detect-Heart/21096290.html.

[180] E. Pescara, M. Beigl, and M. Budde, "RüttelFlug – A Wrist-Worn Sensing Device for Tactile Vertical Velocity Perception in 3D-Space," in *2016 ACM International Symposium on Wearable Computers (ISWC'16)*, ACM, 2016, pp. 172–175. DOI: 10.1145/2971763.2971795.

[181] E. Pescara, A. Wolpert, M. Budde, A. Schankin, and M. Beigl, "LifeTact - Utilizing Smartwatches as Tactile Heartbeat Displays in Video Games," in *16th International Conference on Mobile and Ubiquitous Multimedia (MUM 2017)*, 2017. DOI: 10.1145/3152832.3152863.

[182] J. Peters, J. Theunis, M. Van Poppel, and P. Berghmans, "Monitoring PM10 and Ultrafine Particles in Urban Environments Using Mobile Measurement," in *AAQR*, 2013.

[183] L. Peterson and M. Peterson, "Short-term retention of individual verbal items," *Journal of Experimental Psychology*, vol. 58, 1959.

[184] A. Petzold and R. Niessner, "Photoacoustic soot sensor for in-situ black carbon monitoring," *Applied Physics B: Lasers and Optics*, vol. 63, no. 2, pp. 191–197, 1996.

[185] M. Piorkowski, N. Sarafijanovic-Djukic, and M. Grossglauser, *The epfl/mobility dataset (v. 2009-02-24)*, 2009. [Online]. Available: `https://crawdad.org/epfl/mobility/20090224/` (visited on 11/21/2017).

[186] S. Poduri, A. Nimkar, and G. S. Sukhatme, "Visibility Monitoring using Mobile Phones," Department of Computer Science, USC, Tech. Rep., 2010.

[187] C. A. Pope III, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, and G. D. Thurston, "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution," *Jama*, vol. 287, no. 9, pp. 1132–1141, 2002.

[188] A. Proietti, F. Leccese, M. Caciotta, F. Morresi, U. Santamaria, and C. Malomo, "A new dusts sensor for cultural heritage applications based on image processing," *Sensors*, vol. 14, no. 6, pp. 9813–9832, 2014.

[189] A. Proietti, M. Panella, F. Leccese, and E. Svezia, "Dust detection and analysis in museum environment based on pattern recognition," *Measurement*, vol. 66, pp. 62–72, 2015.

[190] *Public Lab: Air Column Monitor*, 2012. [Online]. Available: `https://publiclab.org/wiki/air-column-monitor` (visited on 11/21/2017).

[191] N. Ramanathan, M. Lukac, T. Ahmed, A. Kar, P. Praveen, T. Honles, I. Leong, I. Rehman, J. Schauer, and V. Ramanathan, "A cellphone based system for large-scale monitoring of black carbon," *Atmospheric Environment*, vol. 45, no. 26, pp. 4481–4487, 2011, ISSN: 1352-2310. DOI: `http://dx.doi.org/10.1016/j.atmosenv.2011.05.030`.

[192] R. K. Rana, C. T. Chou, S. S. Kanhere, N. Bulusu, and W. Hu, "Ear-phone: an end-to-end participatory urban noise mapping system," in *9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, 2010, pp. 105–116.

[193] S. Reddy, D. Estrin, and M. Srivastava, "Recruitment Framework for Participatory Sensing Data Collections," English, in *Pervasive Computing*, ser. LNCS, vol. 6030, 2010, ISBN: 978-3-642-12653-6.

[194] Retsch Technology. (2017). CAMSIZER X2 Product Information, [Online]. Available: https://www.retsch-technology.de/de/api/?action=product_pdf&productId=399&id=2296509&L=6&print_info=1&print_image=1&print_examples=1&print_advantages=1&print_features=1&print_principle=1 (visited on 11/18/2017).

[195] J. Rheinfrank and S. Evenson, "Design languages," *Bringing design to software*, 1996.

[196] J. F. Rohe, "Incentivising High Quality Mobile Sensing Using Mini-Games," Bachelor's Thesis, Karlsruhe Insitute of Technology (KIT), 2017.

[197] R. Rückerl, A. Schneider, R. Hampel, S. Breitner, J. Cyrys, U. Kraus, J. Gu, J. Soentgen, W. Koenig, and A. Peters, "Association of novel metrics of particulate matter with vascular markers of inflammation and coagulation in susceptible populations–results from a panel study," *Environmental research*, vol. 150, pp. 337–347, 2016.

[198] R. Salonga, *San Jose Pokémon Go player tells of being slashed in face by suspicious bystander*, The Mercury News, Aug. 2016. [Online]. Available: http://www.mercurynews.com/2016/08/03/san-jose-pokmon-go-player-tells-of-being-slashed-in-face-by-suspicious-bystander/ (visited on 11/21/2017).

[199] S. Santini, B. Ostermaier, and R. Adelmann, "On the use of sensor nodes and mobile phones for the assessment of noise pollution levels in urban environments," in *INSS'09*, 2009, ISBN: 978-1-4244-6313-8.

[200] O. Saukh, D. Hasenfratz, and L. Thiele, "Reducing Multi-Hop Calibration Errors in Large-Scale Mobile Sensor Networks," in *IPSN'15*, ACM, 2015.

[201] M. Schroyer, *DustDuino*, 2013. [Online]. Available: http://www.mentalmunition.com/2013/05/dustduino-plan-to-crowdsource.html (visited on 11/21/2017).

[202] J. Schwartz, D. W. Dockery, and L. M. Neas, "Is daily mortality associated specifically with fine particles?" *Journal of the Air & Waste Management Association*, vol. 46, no. 10, pp. 927–939, 1996.

[203] I. Schweizer, R. Bärtl, A. Schulz, F. Probst, and M. Mühläuser, "NoiseMap – real-time participatory noise maps," in *Proc. 2nd Int'l Workshop on Sensing Applications on Mobile Phones (PhoneSense'11)*, 2011.

[204] I. Schweizer, C. Meurisch, J. Gedeon, R. Bärtl, and M. Mühlhäuser, "Noisemap: Multi-tier Incentive Mechanisms for Participative Urban Sensing," in *Proceedings of the Third International Workshop on Sensing Applications on Mobile Phones*, ser. PhoneSense '12, Toronto, Ontario, Canada: ACM, 2012, 9:1–9:5, ISBN: 978-1-4503-1778-8.

[205] C. G. Scully, J. Lee, J. Meyer, A. M. Gorbach, D. Granquist-Fraser, Y. Mendelson, and K. H. Chon, "Physiological Parameter Monitoring from Optical Recordings With a Mobile Phone," *IEEE Trans. Biomed. Engineering*, vol. 59, no. 2, pp. 303–306, 2012. [Online]. Available: `http://dblp.uni-trier.de/db/journals/tbe/tbe59.html#ScullyLMGGMC12`.

[206] L. See, A. Comber, C. Salk, S. Fritz, M. van der Velde, C. Perger, C. Schill, I. McCallum, F. Kraxner, and M. Obersteiner, "Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts," *PLoS ONE*, vol. 8, no. 7, 2013.

[207] SHARP Corporation. (). Application note of Sharp dust sensor (GP2Y1010AU), [Online]. Available: `http://www.beck-elektronik.de/fileadmin/templates/beck_folder/opto/sensor/sharp/an-gp2y1010au.pdf` (visited on 11/21/2017).

[208] ——, (2006). GP2Y1010 optical dust sensor. Sheet No.: E4-A01501EN, [Online]. Available: `http://sharp-world.com/products/device/lineup/data/pdf/datasheet/gp2y1010au_e.pdf` (visited on 11/04/2017).

[209] ——, (2014). Device Specification for PM2.5 Sensor module MODEL NO. DN7C3CA006. SPEC No.: ESH-14601B, [Online]. Available: `https://www.digchip.com/datasheets/parts/datasheet/424/DN7C3CA006-pdf.php` (visited on 11/16/2017).

[210] S. A. Sheppard and L. Terveen, "Quality is a verb: the operationalization of data quality in a citizen science community," in *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 2011, pp. 29–38.

[211] S. A. Sheppard, A. Wiggins, and L. Terveen, "Capturing Quality: Retaining Provenance for Curated Volunteer Monitoring Data," in *17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ser. CSCW '14, ACM, 2014, pp. 1234–1245.

[212] Shinyei. (2010). Aerosol Sensor Model AES-1 Model AES-4, [Online]. Available: `http://c1170156.r56.cf3.rackcdn.com/UK_SHN_AES-1_DS.pdf` (visited on 11/16/2017).

[213] ——, (2010). Particle Sensor Model PPD 42NS, [Online]. Available: `http://wiki.timelab.org/images/f/f9/PPD42NS.pdf` (visited on 11/16/2017).

[214] (Oct. 26, 2013). Shinyei – Products – Particle Sensor, [Online]. Available: `https://web.archive.org/web/20131026091620/http://sca-shinyei.com/particlesensor`.

[215] L. Šikšnys, J. R. Thomsen, S. Šaltenis, M. L. Yiu, and O. Andersen, "A Location Privacy Aware Friend Locator," in *Advances in Spatial and Temporal Databases*, ser. Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, pp. 405–410.

[216] SKC inc. (). Personal Environmental Monitor for Measurement of PM10 and PM2.5 in Indoor Air (data sheet). Publication 1367 Rev 1601, [Online]. Available: `http://www.skcinc.com/instructions/1367.pdf` (visited on 11/21/2017).

[217] F. Snik, J. H. Rietjens, A. Apituley, H. Volten, B. Mijling, A. Di Noia, S. Heikamp, R. C. Heinsbroek, O. P. Hasekamp, J. M. Smit, *et al.*, "Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters," *Geophysical Research Letters*, vol. 41, no. 20, pp. 7351–7358, 2014. DOI: `10.1002/2014GL061462`.

[218] E. G. Snyder, T. H. Watkins, P. A. Solomon, E. D. Thoma, R. W. Williams, G. S. W. Hagler, D. Shelow, D. A. Hindin, V. J. Kilaru, and P. W. Preuss, "The Changing Paradigm of Air Pollution Monitoring," *Environmental Science & Technology*, vol. 47, no. 20, pp. 11 369–11 377, 2013. DOI: `10.1021/es4022602`.

[219] K. Spencer, *OpenSimplexNoise.java*. [Online]. Available: `http://gist.github.com/KdotJPG/b1270127455a94ac5d19` (visited on 11/29/2017).

[220] Sven Leckel Ingenieurbüro GmbH. (). Sequential Sampler SEQ47/50, [Online]. Available: http://www.leckel.de/index.php?option=com_docman&task=doc_download&gid=20 (visited on 11/21/2017).

[221] L. Sweeney, "k-Anonymity: A Model for Protecting Privacy," *Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, pp. 557–570, 2002.

[222] J. Sweller, "Visualisation and instructional design," in *International Workshop on Dynamic Visualizations and Learning*, 2002.

[223] R. Tan, G. Xing, Z. Yuan, X. Liu, and J. Yao, "System-Level Calibration for Fusion-Based Wireless Sensor Networks," in *Real-Time Systems Symposium (RTSS), 2010 IEEE 31st*, ieeexplore.ieee.org, Nov. 2010, pp. 215–224.

[224] ——, "System-level Calibration for Data Fusion in Wireless Sensor Networks," *ACM Trans. Sen. Netw.*, vol. 9, no. 3, 28:1–28:27, Jun. 2013.

[225] Texas Instruments, *Battery chargers in USB OTG devices*, Retrieved: November 2016, Jun. 2010. [Online]. Available: http://www.ti.com/lit/wp/sszy001/sszy001.pdf (visited on 11/21/2017).

[226] The European Parliament and the Council of the European Union, *Directive 2008/50/EC*, May 2008.

[227] The World Health Organization. (2005). WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide – Global update 2005 – Summary of risk assessment, [Online]. Available: http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf (visited on 10/31/2017).

[228] B. A. Thelen and R. K. Thiet, "Cultivating connection: Incorporating meaningful citizen science into Cape Cod National Seashore's estuarine research and monitoring programs," *Park Science*, vol. 25, no. 1, 2008.

[229] J. Theunis, M. Stevens, and D. Botteldooren, "Sensing the Environment," in *Participatory Sensing, Opinions and Collective Awareness*, V. Loreto, M. Haklay, A. Hotho, V. D. Servedio, G. Stumme, J. Theunis, and F. Tria, Eds. Cham: Springer International Publishing, 2017, pp. 21–46, ISBN: 978-3-319-25658-0. DOI: 10.1007/978-3-319-25658-0_2. [Online]. Available: https://doi.org/10.1007/978-3-319-25658-0_2.

[230] A. Truskinger, H. Yang, J. Wimmer, J. Zhang, I. Williamson, and P. Roe, "Large Scale Participatory Acoustic Sensor Data Analysis: Tools and Reputation Models to Enhance Effectiveness," in *E-Science*, 2011.

[231] TSI. (2017). Aerodynamic Particle Sizer Model 3321. P/N 1930087 Rev F, [Online]. Available: `http://www.tsi.com/uploadedFiles/_Site_Root/Products/Literature/Spec_Sheets/3321.pdf` (visited on 11/13/2017).

[232] TSI Incorporated, *DustTrak DRX Aerosol Monitor Model 8533/8534/8533 EP – Operation and Service Manual*, 2013. [Online]. Available: `http://www.tsi.com/uploadedFiles/_Site_Root/Products/Literature/Manuals/8533-8534-DustTrak_DRX-6001898-web.pdf` (visited on 11/21/2017).

[233] W. Tsujita, H. Ishida, and T. Moriizumi, "Dynamic gas sensor network for air pollution monitoring and its auto-calibration," in *IEEE Sensors*, ieeexplore.ieee.org, Oct. 2004, 56–59 vol.1.

[234] A. S. Tucker, "DUSTY – The Intellivac," University of Florida, Intelligent Machine Design Laboratory, Report, 2005. [Online]. Available: `http://mil.ufl.edu/5666/papers/IMDL_Report_Spring_05/skipp-trevor/dusty.pdf` (visited on 11/21/2017).

[235] K. Tuite, N. Snavely, D.-y. Hsiao, N. Tabing, and Z. Popovic, "PhotoCity: Training Experts at Large-scale Image Acquisition Through a Competitive Game," in *SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11, ACM, 2011, ISBN: 978-1-4503-0228-9.

[236] U.S. Environmental Protection Agency. (). Particulate Matter (PM) Pollution, [Online]. Available: `https://www.epa.gov/pm-pollution/particulate-matter-pm-basics` (visited on 11/21/2017).

[237] ——, (1987). Revisions to the national ambient air quality standards for particulate matter. Federal Register 24634, [Online]. Available: `https://www3.epa.gov/ttn/naaqs/standards/pm/previous/pm-1987-final-52fr24634.pdf` (visited on 10/30/2017).

[238] ——, (2006). National Ambient Air Quality Standards for Particulate Matter. Federal Register 61144, [Online]. Available: https://www.gpo.gov/fdsys/pkg/FR-2006-10-17/pdf/06-8477.pdf (visited on 11/21/2017).

[239] Y. Ueyama, M. Tamai, Y. Arakawa, and K. Yasumoto, "Gamification-based incentive mechanism for participatory sensing," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on*, IEEE, 2014, pp. 98–103.

[240] USB Implementers Forum Inc., *On-The-Go Supplement to the USB 2.0 Specification*, Rev 1.0, Dec. 18, 2001. [Online]. Available: http://www.usb.org/developers/onthego/otg1_0.pdf (visited on 11/21/2017).

[241] M. Ushida, Y. Yamaoka, K. Itoh, and H. Tsuda, "New Privacy-Preserving Method for Matching Location Data," in *IMIS'14*, 2014, pp. 594–599.

[242] VDA-QMC, "Messsystem und Messprozess sind zweierlei," *QZ*, vol. 56, no. 5, 2011.

[243] K. Weekly, D. Rim, L. Zhang, A. M. Bayen, W. W. Nazaroff, and C. J. Spanos, "Low-cost coarse airborne particulate matter sensing for indoor occupancy detection," in *Automation Science and Engineering (CASE), 2013 IEEE International Conference on*, ieeexplore.ieee.org, Aug. 2013, pp. 32–37.

[244] K. Whitehouse and D. Culler, "Calibration as parameter estimation in sensor networks," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, New York, NY, USA: ACM, Sep. 2002, pp. 59–67.

[245] C. D. Wickens, J. Lee, Y. D. Liu, and S. Gordon-Becker, *Introduction to Human Factors Engineering: Pearson New International Edition*. Pearson Higher Ed, 2014, Second Edition.

[246] A. Wiedensohler, W. Birmili, A. Nowak, A. Sonntag, K. Weinhold, M. Merkel, B. Wehner, T. Tuch, S. Pfeifer, M. Fiebig, A. M. Fjäraa, E. Asmi, K. Sellegri, R. Depuy, H. Venzac, P. Villani, P. Laj, P. Aalto, J. A. Ogren, E. Swietlicki, P. Williams, P. Roldin, P. Quincey, C. Hüglin, R. Fierz-Schmidhauser, M. Gysel, E. Weingartner, F. Riccobono, S. Santos, C. Grüning, K. Faloon, D. Beddows,

R. Harrison, C. Monahan, S. G. Jennings, C. D. O'Dowd, A. Marinoni, H.-G. Horn, L. Keck, J. Jiang, J. Scheckman, P. H. McMurry, Z. Deng, C. S. Zhao, M. Moerman, B. Henzing, G. de Leeuw, G. Löschau, and S. Bastian, "Mobility particle size spectrometers: harmonization of technical standards and data structure to facilitate high quality long-term observations of atmospheric particle number size distributions," *Atmospheric Measurement Techniques*, vol. 5, no. 3, pp. 657–685, 2012. DOI: 10.5194/amt-5-657-2012. [Online]. Available: https://www.atmos-meas-tech.net/5/657/2012/ (visited on 11/21/2017).

[247] K. Wiesner and F. Dorfmeister, "PRICAPS: A System for Privacy-Preserving Calibration in Participatory Sensing Networks," 2014.

[248] K. Wiesner, F. Dorfmeister, and C. Linnhoff-Popien, "Privacy-Preserving Calibration for Participatory Sensing," in *Mobiquitous'13, Rev. Sel. Papers*, 2014.

[249] A. Wiggins and Y. He, "Community-based Data Validation Practices in Citizen Science," in *19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, ser. CSCW '16, ACM, 2016, pp. 1548–1559, ISBN: 978-1-4503-3592-8.

[250] R. Williams, V. Kilaru, E. Snyder, A. Kaufman, T. Dye, A. Rutter, A. Russell, and H. Hafner, "Air Sensor Guidebook," U.S. Environmental Protection Agency, Report, 2014, EPA/600/R-14/159 (NTIS PB2015-100610). [Online]. Available: https://www.epa.gov/air-sensor-toolbox/how-use-air-sensors-air-sensor-guidebook (visited on 11/28/2017).

[251] World Health Organisation (WHO). (2014). 7 million premature deaths annually linked to air pollution, [Online]. Available: http://www.who.int/mediacentre/news/releases/2014/air-pollution (visited on 11/21/2017).

[252] World Health Organization (WHO), "Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide," World Health Organization (WHO), Report on a WHO Working Group, 2003. [Online]. Available: http://apps.who.int/iris/bitstream/10665/107478/1/E79097.pdf (visited on 11/21/2017).

[253] P. Wright, ""The instructions clearly state. . . " can't people read?" *Applied Ergonomics*, vol. 12, pp. 131–141, 3 Sep. 1981.

[254] Y. Xiang, L. Bai, R. Piedrahita, R. P. Dick, Q. Lv, M. Hannigan, and L. Shang, "Collaborative Calibration and Sensor Placement for Mobile Sensor Networks," in *IPSN'12*, ACM, 2012, pp. 73–84.

[255] P. Yadav and J. Darlington, "Design Guidelines for the User-Centred Collaborative Citizen Science Platforms," *arXiv preprint arXiv:1605.00910*, 2016.

[256] W. Y. Yi, K. M. Lo, T. Mak, K. S. Leung, Y. Leung, and M. L. Meng, "A survey of wireless sensor network based air pollution monitoring systems," *Sensors*, vol. 15, no. 12, pp. 31 392–31 427, 2015.

[257] X. Yu, W. Zhang, L. Zhang, V. O. Li, J. Yuan, and I. You, "Understanding urban dynamics based on pervasive sensing: An experimental study on traffic density and air pollution," *Mathematical and Computer Modelling*, vol. 58, no. 5, pp. 1328–1339, 2013, The Measurement of Undesirable Outputs: Models Development and Empirical Analyses and Advances in mobile, ubiquitous and cognitive computing, ISSN: 0895-7177. DOI: `https://doi.org/10.1016/j.mcm.2013.01.002`.

[258] Y. Yu and H. Li, "Virtual in-situ calibration method in building systems," *Autom. Constr.*, vol. 59, pp. 59–67, Nov. 2015.

[259] A. Yuhas, *Pokémon Go: armed robbers use mobile game to lure players into trap*, 2016. [Online]. Available: `https://www.theguardian.com/technology/2016/jul/10/pokemon-go-armed-robbers-dead-body` (visited on 11/21/2017).

[260] W. Zhang, B. Zhu, L. Zhang, J. Yuan, and I. You, "Exploring Urban Dynamics Based on Pervasive Sensing: Correlation Analysis of Traffic Density and Air Quality," in *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on*, ieeexplore.ieee.org, Jul. 2012, pp. 9–16.

[261] Y. Zheng, L. Li, and L. Zhang, "PiMi air community:: getting fresher indoor air by sharing data and knowhows," in *Proceedings of the 13th international symposium*

on Information processing in sensor networks, IEEE Press, 2014, pp. 283–284.

[262] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004)*, IEEE, vol. 2, 2004, pp. 28–31.

[263] Z. Zivkovic and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern recognition letters*, vol. 27, no. 7, pp. 773–780, 2006.