

# Location Based Sensing for Health Diagnoses: A Prototype for Personalized Pollen Profiling

Klaus Böhm<sup>1</sup>, Nikolai Bock<sup>1</sup> and Torsten Sehlinger<sup>2</sup>

<sup>1</sup>i3mainz, Mainz · klaus.boehm@hs-mainz.de

<sup>2</sup>Bluestone Technology GmbH

Full paper double blind review

## Abstract

Personalized information and forecasts of environmentally triggered medical symptoms are becoming more and more critical, since the numbers of patients suffering from environmentally induced symptoms is still increasing (e.g. allergy or asthma patients). This project aimed at developing a solution to enable gathering personalized pollen exposure information and mapping this information with geospatial data from the personal habitat. This enables personal forecasts based on geospatial information and, when distributed to other suffering individuals, gives them the possibility of avoiding harmful environments.

Within an interdisciplinary team of researchers from the medical, engineering, and geoinformatics professions, a system was developed to gather, qualify, and quantify aeroallergen particles (pollen), monitor additional climate parameters, and map this information, together with geospatial information, to the daily habitat of a single individual.

## 1 Introduction

Environmental conditions have significant influence on the personal well-being. Moreover, the correlation to individual diseases is evident. A very prominent disease is pollen allergy leading, e.g. to hay fever and asthma. Today's modern industrial civilizations face an increasing number of people suffering from allergy symptoms. The number of affected citizens is already 30 – 40 % (on average) of the total population (PAWANKAR et al. 2011, CECCHI 2013).

Although the individual pollen exposure plays a significant role for the anamnesis, there are currently no means to determine it. As of today, this exposure rate is extrapolated by pollen data gathered by a very limited number of pollen traps spread around the country. The situation in Germany is fortunate and a lot of pollen traps are installed compared to other countries. Still, the extrapolated pollen exposure is not precise at all. Since pollen traps in Germany are situated at least 80 kilometers apart, the extrapolation neither includes the influences of the flora surrounding the subject in its daily habitat, nor sufficiently depicts geographic and urban conditions. Expensive stationary pollen traps, e.g. the Burkard Pollen Sampler, don't provide an adequate solution for this situation. In this paper we describe our approach to the demand described above, focusing on a new general location based sensing approach, and how the application incorporates the location relevant information on the level of individual pollen exposure. After a description of the requirements and our general

approach, including a brief description of all components of the solution, we present a typical use case. Then our approach for calculating the location reference of individual pollen exposure is outlined. Finally we describe our plans for geospatial analysis based on the new datasets.

## 2 Approach

### 2.1 General Requirements

Currently no similar hardware/software system is available. The general requirements, which have been considered during our research and development, are the following:

- Tracking, and thus knowing an individual's pollen exposure throughout the day, as well as when and where the exposure took place.
- Tracking, and thus knowing an individual's symptoms, including time and location thereof, medication usage and/or allergen immunotherapy background.
- Tracking, and thus knowing an individual's environmental factors like temperature, humidity, sun exposure, being indoors or outside, etc.

### 2.2 Technical Approach for “Personal Pollen Profiling Solution”

To analyse the pollen exposure of an individual it is necessary to collect the pollen, associated with other environmental data, within the individual's daily life habitat, and record its allergy medication usage and allergy symptoms promptly, as well as qualify and quantify the pollen/data collected. Therefore, the Personal Pollen Profiling solution consists of a wearable Pollen sampler, equipped with a test cartridge, an analysing laboratory device, called the Pollen Scanner, a mobile/web application for real-time anamnesis called Allergy Diary, and a centralized data system to store, further analyse, and condition the data gathered by the collector and the scanner, called centralized infrastructure.

The **personal pollen sampler** is the basis for providing the new unique dataset; the **centralized infrastructure** and the **pollen diary** are described in more detail.

### 2.3 Personalized Pollen Sampler

The personal pollen sampler is a small portable device, which can be attached to a belt, a backpack, or a handbag. The pollen sampler actively sucks in ambient air through a miniature fan. The resulting air stream is distorted and guided onto the test stripe in the test cartridge, where particles (like pollen or fungal spores) will adhere to the test stripes' surface. The pollen collector is also equipped with GPS and sensors for temperature, relative humidity, and barometric pressure.

This newly developed device is available as a prototype, demonstrating all specified functionalities and the required mobility in one unique unit. The performance of the device has been successfully evaluated in comparison to stationary Burkard or Hirst pollen traps (SEHLINGER et al. 2013).

Figure 1 and 2 describe the general use of the pollen sampler.



**Fig. 1:**  
Daily Usage of Pollen Sampler



- Pollen Collector is logging environmental data of personal habitat
  - Temperature
  - Humidity
  - Barometric Pressure
  - UV-A, UV-B Radiation
- Logging also includes GPS-Location
- Data is automatically transferred to Central Server when WiFi or USB connection available

**Fig. 2:**  
Environmental Data Logging

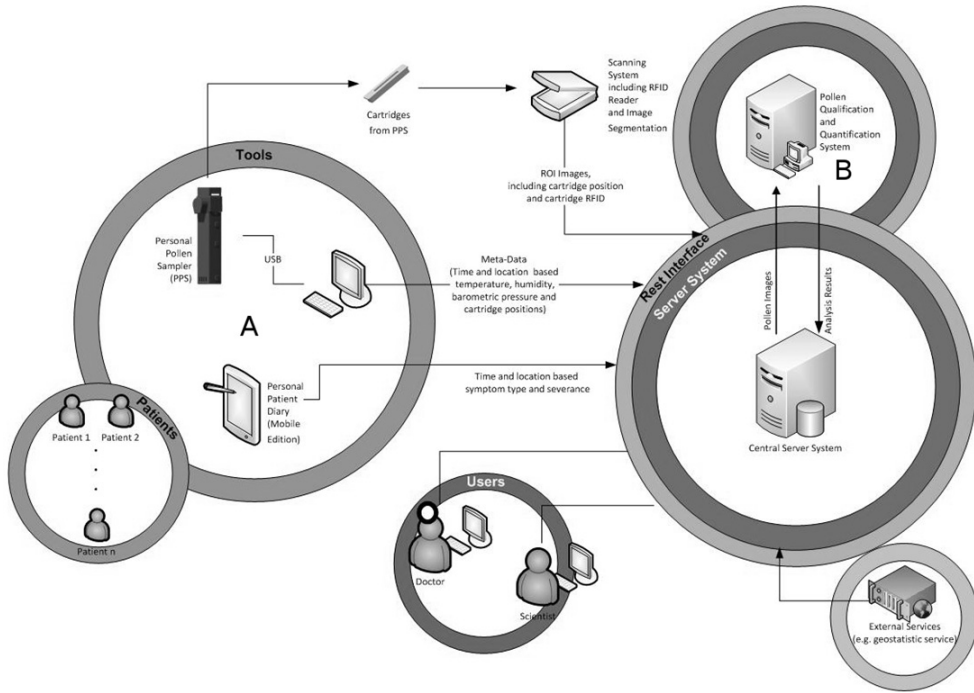
## 2.4 Allergy Self-Anamnesis Application (Allergy Diary)

The Allergy Diary is a mobile application, able to run on any smartphone. This application supports the clinical trials subject and/or patient, keeping track of his or her symptoms, providing possibilities to select symptoms, severance, and allergy medication taken. The data gathered through the diary is hosted within an instance of the centralized infrastructure. The data to be collected through the allergy diary is defined by a generally established selection of symptoms (see VOUKANTISIS et al. 2013), and in addition by individual requirements for specific clinical studies.

## 2.5 Centralized Infrastructure and Workflow

This chapter describes the centralized infrastructure and the overall workflow (Figure 3).

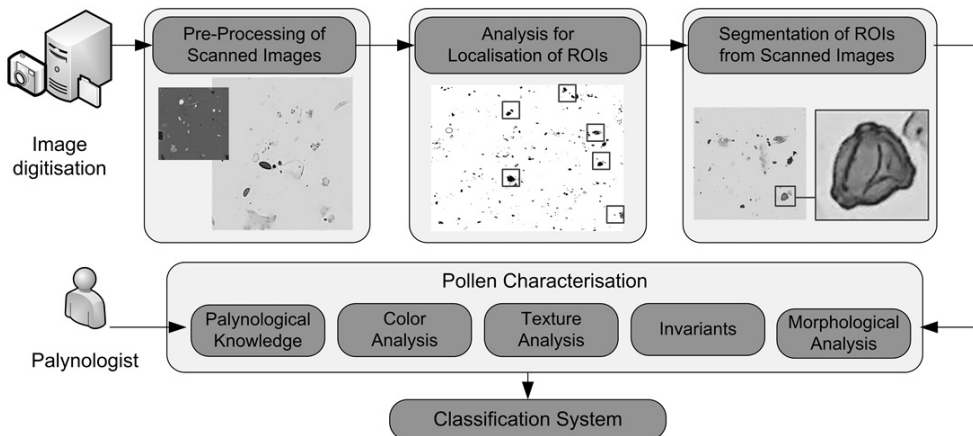
The patient or user wears the personal pollen sampler (PPS) and operates the personal allergy diary (Fig. 3, circle A). Then data of the diary is stored directly in the centralized server system. The metadata from the PPS is uploaded via a dedicated program to the centralized server system. The classification of the pollen requires intensive processing and is therefore performed offline. This process, shown in the small circle (Fig. 3, circle B), is described briefly in Section 2.5.1. The results of the classification are also stored in the centralized server system. The data is finally analysed by various user groups such as doctors, which support the clinical studies, scientists, and the patient/user themselves.



**Fig. 3:** Overview of the centralized infrastructure

### 2.5.1 Automatic Pollen Detection

The automatic pollen detection (in the Figure 3 in the upper right corner) is based on advanced image processing. It includes various steps shown in Figure 4. The first step is scanning the exposed stripes, which is a time and data intensive task. For the scanning, the pollen stripes have to be cut into various slices, which can be handled on standard path-



**Fig. 4:** Pollen detection system

logy scanners. Those slices are scanned in various layers and the images are then used in the analysis and the semi-automatic classification process. The process is further described in LOZANO-VEGA et al. 2013 and LOZANO-VEGA et al. 2014.

The results of the analysis process are the quantified and classified pollen, including the coordinate of the pollen on the test stripe.

### 3 Challenge: Location Reference of Individual Pollen

In the following we focus on the approach of calculating and maintaining the location reference of the individual pollen.

The difficulty of this task is mainly caused by the asynchronous nature of the data handling and the complex classification process. During the pollen collection process, two kinds of data are gathered: the pollen, adhering to the stripe and the metadata (a record consisting of geolocation, temperature humidity and barometric pressure). The pollen adheres at a certain location of the stripe, leading to the obvious question of how to correlate the recorded metadata with the classified pollen. The following basic conditions have to be considered.

- The test stripe moves with a constant speed  $v$  when the collector is activated.  
→  $v$  is a known constant value
- The metadata is recorded at a fixed frequency.
- The pollen collector can be switched off and on, stopping the metadata recording and the movement of the test stripe.  
→ stints are the time periods between turning the device on and off
- The internal clock is used for recording the on/off status.
- The test stripe is subdivided into slices of equal sizes for the scanning process.  
→ the width of the slices is constant and known.
- The data provided for every detected pollen by the segmentation and classification process is: the specific information about the pollen kind, the unique stripe identifier, the slice number and the X-, Y- coordinate on the slice

Figure 5 shows an abstract scheme for the location calculation.

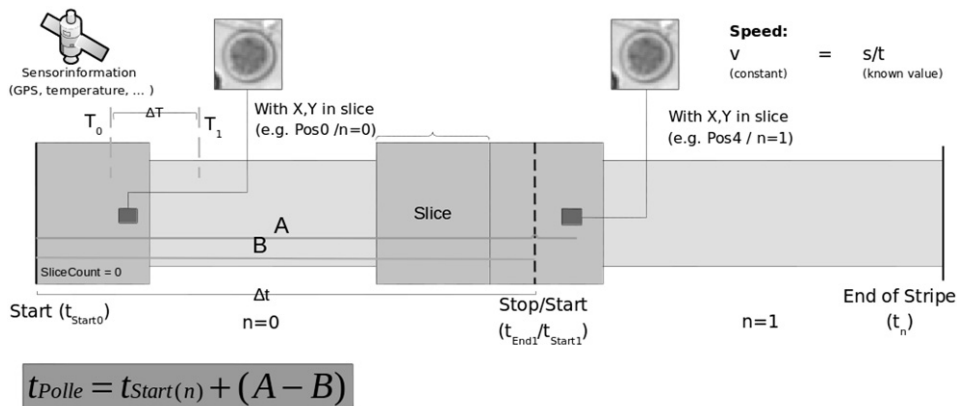


Fig. 5: Calculation of individual pollen location

In order to calculate the geodetic coordinate of the pollen we first have to determine the exact time  $t_{Polle}$  when the pollen adhered:

**Step 1:** Calculate the distance of the pollen in relation to the origin of the test stripe

$$dist = SliceCount * \Delta x + X$$

- SliceCount is the number of slices on which pollen adhered to, starting at 0
- $\Delta x$  is the width of one slice
- $X$  is the position (x-value) of the pollen in the slice coordinate system.

**Step 2:** Calculate  $A$ , the corresponding time for the distance  $dist$ , which will be used as a conditioning variable to determine the stint (on/off status of the device) in which the pollen adhered. This is realized by dividing the distance from step 1 by the speed  $v$ .

$$A = \frac{dist}{v}$$

**Step 3:** Determine the stint in which the pollen adhered

- $\Delta t_n$  is the time range for one stint
- For each set of all stints
  - $t_{sum(n)} = t_{sum(n-1)} + \Delta t_n$
  - if  $t_{sum(n)} > A$  // the stint was found
    - break
  - end loop
- The pollen adhered in stint ( $n$ )
- The reference time now is  $t_{Start(n)}$

**Step 4:** Calculate  $B$ , the time at which the collector was active before the relevant stint with the starting time at  $t_{Start(n)}$

- Calculate  $B$  as the sum of  $\Delta t$  from  $j=0$  to  $n-1$   $t_{End(n-1)}$

$$B = \sum_{j=0}^{n-1} (t_{End(j)} - t_{Start(j)})$$

**Step 5:** Calculate  $t_{Polle}$  the time when pollen adhered.

At the end, the pollen time can be calculated by adding the difference of  $A$  and  $B$  to the start time of stint  $n$ .

$$t_{Polle} = t_{Start(n)} + (A - B)$$

After calculating  $t_{Polle}$  for the individual pollen, the correlation to the metadata is possible. There are different options for determining the correct metadata record, as the calculated time might not fit with the recording times of the metadata:

- Using the previous record
- Using the following record
- Using the temporally nearest observation
- Interpolation between the previous and following observation

The option to be used at the end depends on the recording frequency, which is currently being determined during field tests. Independent of this open fact, the result of the calculation provides the required dataset. For every pollen the following information can be provided:

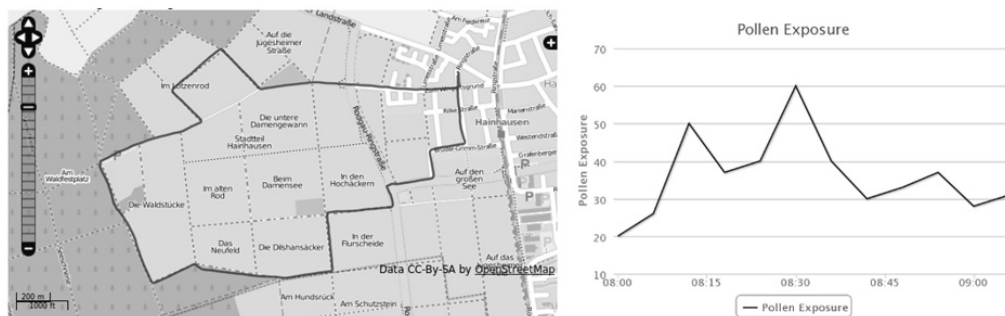
Pollen type; **longitude, latitude**; temperature; humidity; pressure; time

## 4 Geospatial Analysis of the Data

The availability of this unique dataset allows new types of analysis. In our research, we focus on two objectives for which we developed conceptual approaches and prototypes: the global perspective and a personal perspective.

**Global perspective:** Following the assumption that there will be a wide distribution of the personal pollen sampler, a large amount of “personal sensors” is contributing to a common dataset potentially covering a large geographic area (similar NETATMO). This data can provide significantly better results compared to the extrapolated pollen distribution based on data of 50 stationary pollen traps (the number given for Germany). For geographical and geostatistical visualization of these pollen concentrations, a flexible system (see BOCK & BÖHM 2013) has been developed and different types of geo statistical visualizations like Barnes Surface, Kriging, or IDW have been applied (ALBA et al. 2006). This enables much more accurate and reliable pollen forecasts to be used for medical, phenological or farming purposes.

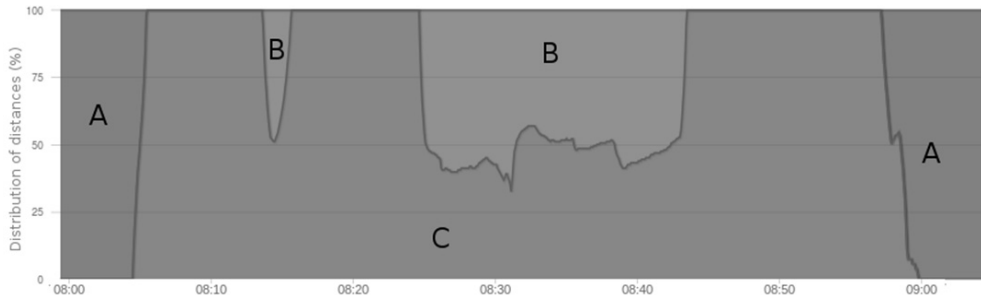
**Personal perspective:** Here the focus is on the geospatial visualization of the individual pollen information. Therefore we apply the combined map-track-chart view (Figure 6).



**Fig. 6:** Track-Chart-View

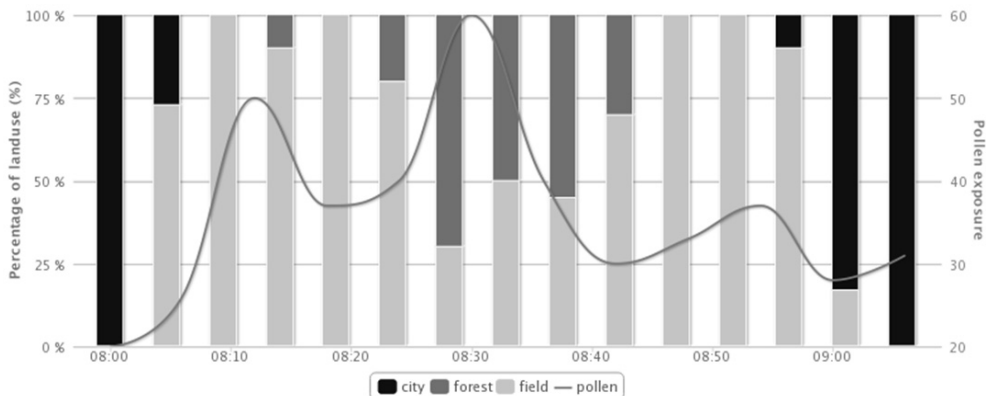
Furthermore, this allows analysing the correlation between land-use and the individual pollen exposure (KÜMMEL 2012). This is relevant for answering the question of whether there is a correlation between your daily environment in terms of land-use and your pollen exposure. For this analysis, we use the OpenStreetMap dataset, which provides land-use information on the level of polygons. We developed various techniques to calculate the dominant land-use along the users track, e.g. buffer or shortest distance for land-use-intersection.

For visualising these calculations, we use various diagrams, e.g. the shortest distance to a land-use type (Fig. 7), which describes the land-use type in a given distance along the track. E.g. in the beginning the user is in an urban environment (A), followed by pure country side, and then partially field (C) and forest (B) environment.



**Fig. 7:** Distances to land uses in a 50m buffer (adapted from KÜMMEL 2012)

Alternatively, Fig. 8 shows the percentage of land-use including an overlay of simulated pollen exposure information as a simple line diagram, which visualizes the correlation in a test scenario.



**Fig. 8:** Percentage of land uses in a 50m buffer (adapted from KÜMMEL 2012)

## 5 Conclusion/Outlook

In our research we have addressed various fields including development of a completely new device and new test stripes, new segmentation and classification methods, new specialized infrastructure and interfaces for different stakeholders.

In this paper we outlined our successful approach of putting each individually collected and qualified pollen in its spatial/temporal context. This enables the personalized correlation of pollen exposure and individual symptoms. Moreover, it allows land-use correlations.



All developments have reached a prototype status, which allows further test cases towards clinical studies. The next steps concentrate on testing and validating the prototypes, including the validation of the collecting device in lab environments (SEHLINGER et al. 2015). With respect to the geospatial analysis components that were described in this paper, large scale tests will allow the validation of the visualization components with real life data.

On the architecture level, the concept of the mobile and personalized pollen sampler, and the associated individual measurements, should be further investigated in the context of “People as Sensors” (see RESCH 2013). This would enable the combination with other sensor information from different sources, and run even more complex data mining and prediction analysis for the evaluation of healthy environments (see also RESCH 2011).

## Acknowledgements

We would like to thank the Central Innovation Program SME of the Federal Ministry of Economics and Technology of Germany for funding the project 3P-GM – Personalized Pollen Profiling and Geospatial Mapping, which provided the financial support for the research. Additional special thanks goes to Kristian Kümmel for his excellent bachelor thesis combining GPS-tracks with land uses of the ambiance.

## References

- ALBA, F., NIETO-LUGILDE, D., COMTOIS, P., DIAZ DE LA GUARDIA, C., DE LINARES, C. & RUIZ, L. (2006), Airborne-pollen map for *Lea europaea* L. in eastern Andalusia (Spain) using GIS: Estimation models. *Aerobiologia*, 22 (2), 107-116.
- BOCK, N. & BÖHM, K. (2013), Flexibles System zur geostatistischen Visualisierung FOSSGIS 2013. <http://www.fossgis.de/konferenz/2013/programm/events/601.de.html>.
- CECCHI L. (2013), Introduction. In: SOFIEV, M. & BERGMANN, K.-C. (Eds), Allergenic pollen – A Review of Production, Release, Distribution and Health Impacts. Springer, Dordrecht/Heidelberg/New York/London, 1-7.
- KÜMME, K. (2012), Visualisierung von GPS-Trackingdaten mit OSM unter Berücksichtigung von Umgebungsinformationen (unpublished Bachelor thesis).
- LOZANO-VEGA, G., BENEZETH, Y., UHLER, M., BOOCHS, F. & MARZANI, F. (2013), Sketch of an automatic image based pollen detection system. DGPF Tagungsband 21/2012.
- LOZANO-VEGA, G., BENEZETH, Y., MARZANI, F. & BOOCHS, F. (2014), Analysis of Relevant Features for Pollen Classification. In *Artificial Intelligence Applications and Innovations*, Springer, Berlin/Heidelberg, 395-404.
- NETATMO, Netatmo Weathermap. <https://www.netatmo.com/weathermap>.
- PAWANKA, R., CANONICA, G. W., HOLGATE, S. T., LOCKEY, R. F. (2011), WAO White Book on Allergy 2011-2012: Execute Summary.
- RESCH, B., BRITTER, R., OUTRAM, C., XIAOJI, R. & RATTI, C. (2011), Standardised geosensor webs for integrated urban air quality monitoring. *Environmental Monitoring*, 2011.
- RESCH, B. (2013), People as Sensors and Collective Sensing-Contextual Observations Completing Geo-Sensor Network Measurements, *Progress in Location-Based Services – Lecture Notes in Geoinformation and Cartography*, 2013, 391-406.

- SEHLINGER, T., BOEHM, K., GOERGEN, F. & BERGMANN, K. (2013), Measuring Individual Pollen Exposure. *Journal of Allergy and Clinical Immunology*, 131 (2), Supplement, Page AB79.
- SEHLINGER T., BERGMANN K-C., GOERGEN F., ZUBERBIER T. (2015), „A novel Mobile Chamber for Allergen Exposure Tests“. EACCI Annual Meeting 2015, Abstract (submitted).
- VOUKANTIS, D., KARATZAS, K., JAEGER, S., BERGER, U. & SMITH, M. (2013), Analysis and forecasting of airborne pollen-induced symptoms with the aid of computational intelligence methods. *Aerobiologia*, 29 (2), 175-185.