Municipality Zwolle : Senshagen
Copernicus Hackathon Report

team Dieliene
- members: Mahabir, Elaine, Diego, Harman, Yuan, Sehan

The Challenge

The Municipality of Zwolle is interested in exploring future possibilities and verifying results from the SensHagen citizen science pilot project. Specific interest goes to investigating if:
(1) the data measured by citizens can be used as ground truth for satellite data; and
(2) if the pilot can be scaled up to cover the whole municipality using satellite data.

Proposed Solution

After analysing the data currently available from the SensHagen pilot and Copernicus satellites, our team proposes to create an AI model, using both the sensor and satellite data, which will provide ongoing measurements and the capability to predict future air quality within the city of Zwolle.

Using the in situ air quality sensor data collected in Stadshagen since 2018 as ground truth for data collected by the Sentinel 5P satellite over the same period, the model will be able to then extrapolate and predict air quality using the Sentinel data alone to scale up coverage to the whole municipality.

Although additional data collected via ongoing citizen science projects around the city would serve to improve both the accuracy of the model as well as output resolution and would certainly be "nice to have", deep learning would enable the estimation of values based on the Stadshagen data for a larger region without requiring additional sensor data. Furthermore, we have noted that additional existing and past air quality sensor data[1] has been collected in other regions of the Netherlands and could therefore also be used to train the model.

With increasingly better insight into the local air quality conditions of Zwolle, our proposed model can be a useful monitoring and prediction tool to support local policy decision-making that will have long-term future impacts on the health of the citizens of Zwolle taking into account ever-increasing climate change.

In particular, our AI model will:
1. give insight into the spatial and temporal fluctuations of measured emissions;
2. be able to predict which areas are most prone to hazardous values;
3. function with a sparse in situ sensor network.

Moreover, the user will get real time information in a few clicks through web and mobile apps.

[1] Street monitoring stations of the Air Quality Monitoring Network [www.luchtmeetnet.nl] and data collected for the Environmental Health Atlas (https://www.atlasleefomgeving.nl/en).

Our proof of concept

For our proof of concept, we have chosen to use the data variable, nitrogen dioxide (NO2). Although the NO2 data collected by the Copernicus satellites are less accurate than the in situ NO2 sensor data collected for Stadshagen, our model can train the satellite data with the latter to improve its utility.

We will first use the 2 data-products for the creation of the AI-model. The AI model will use neural networks to conduct spatial and temporal prediction of NO2 values within Stadshagen, supported by data visualisation techniques using mobile and web applications. Specifically, the 2 data-products are:

1. local NO2 data collected from the SensHagen-RIVM sensors,
2. NO2 data from Sentinel 5-P

The following steps will be used for generating the AI model:

Step 1(prototyped): Use Senshagen data to make predictions for the Stadshagen pilot region
Step 2: Combine Stadshagen and Sentinel 5P data to make estimations for city of Zwolle
Step 3: Combining step 1 and 2 to make future predictions for the whole of Zwolle

Once our AI model is validated for Zwolle, the same approach can be used to model air quality for other cities across the Netherlands and elsewhere around the world. Results are shown for a prototyped version of step 1 in the product development section further down below.

Technical Specifications and Requirements

Although training a neural network requires a lot of computational power, a relatively small area enhances the workability and computational time.

Stadshagen is an ideal area with respect to:
- result interpretation
- computational time
- size sufficient input data

The area of Stadshagen has the right spatial resolution of 5 km2. Training data from the 19 RIVM-sensors with hourly NO2 values for almost 2-years is an amount commonly used in AI-tutorials.

The data required as inputs into our model already exist and as they have been collected via Senshagen sensors and Copernicus satellites, they are freely available. Additional open data sources[2] for air quality in The Netherlands have also been identified and can therefore also be used in the future.

The air quality data from Sentinel 5P, with specifications below, will be retrieved from Sentinel-hub. In this hub, products are easily configured e.g. atmospheric correction, cloud coverage, area of interest, time-span. Even indices can be created and applied on the data. The only issues so far encountered have been that the data-platform for the sensor-data is in the dutch language and that the sentinel-hub is not intuitive for downloading data, but these are easily overcome.

[2] see note 1.

Data Set Specifications

Temporal coverage:	Since 30-Apr-2018
Spatial coverage:	~2600 km swath. Full daily surface coverage of radiance and reflectance measurements for latitudes > 7° and < -7°, and better than 95 % coverage for latitudes in the interval [-7°, 7°].
Spatial resolution:	3.5 x 7.0 Km (across x along track), at beginning of mission 3.5 x 5.5 Km (across x along track), since 6 August 2019
Processor Version:	01.xx.xx, with x=any digit between 0 and 9

A local system is sufficient for the initial draft version of the model. But for future production, outsourced services like Sobloo (integrated Copernicus data) or AWS should be used.

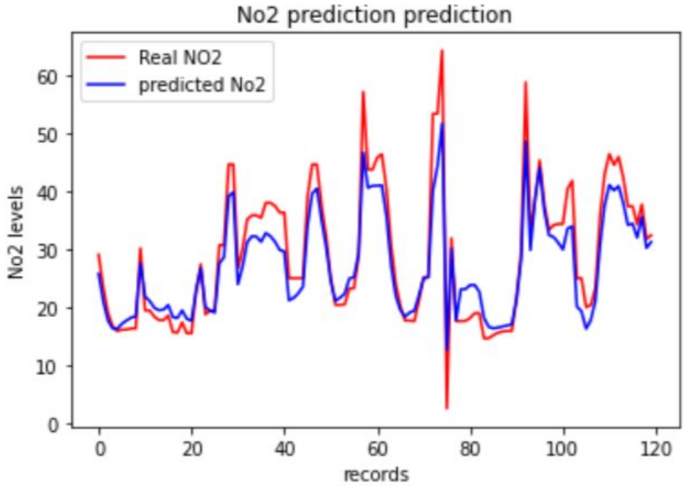
Tools and Technology - Python,Pycharm, Google Colab ,GDAL. and Soblo infrastructure. The method processing time is based on the volume of data and configuration of software and hardware.

Product Development

Backend development :

- A Temporal deep learning model to produce results for step 1 mentioned in the proof of concepts section of this report
- CNN deep learning model to produce results for step 2
- A combination of step 1 and step 2 to produce predictions for the whole region of zwolle

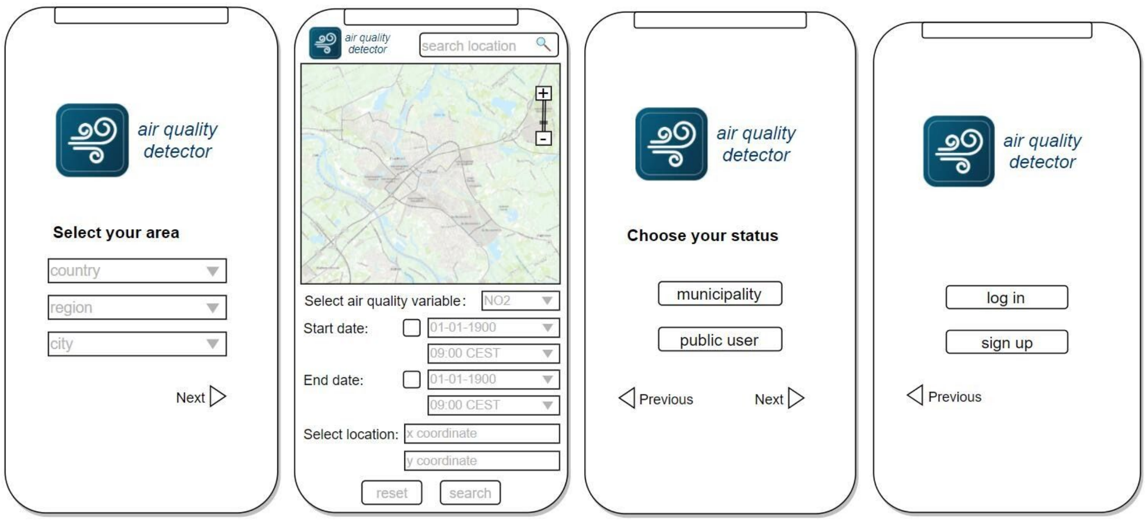
Step 1 of this process was prototyped and the results are displayed below.



Img 1.) The image above shows that the he model is able to predict results for upto 120 readings in advance with an MSE loss of 0.004. Note: the sensors record data every hour for NO2. therefore, 120 readings would equate to 5 days worth of hourly data.

Frontend development :

- Develop mobile app with different levels of access for government and public users as shown below:



Img 2.) The graphics above show a prototyped version of the UI for the mobile app we envisioned.

Why is our proposed solution innovative?

There are many organisations measuring and modelling air pollution using sensors, but no start-ups have so far been identified that propose combining sensor and satellite data to predict future air quality using artificial intelligence.

Due to the growing impacts from climate change, there is strong market demand around the world for air pollution modeling, which is currently only met by sensor data. There are many cities around the world that are not yet measuring air quality, which is a basic requirement for becoming climate-smart, something all cities will eventually need to become in the not-too-distant future. Modelling air quality based primarily on existing satellite data would offer considerable cost savings in terms of implementation and maintenance of sensors, while also offering increased versatility. Although satellite data is, in general, less accurate than using a sensor array, combining satellite data with sensor data would increase the output resolution, while also reducing the need for as many sensors.

Although currently, our model would be limited by the data resolution of the Sentinel 5P satellite, we are exploring the capability to improve the resolution of the Sentinel satellite images using AI, which would then improve the data resolution of our air quality monitoring & prediction models for cities located almost anywhere[3] around the world.

Added Societal, Economic & Market Value

Knowledge is everything. Being able to reliably predict the future is even better. The key to future climate-smart cities is not just implementing green measures, but also the ability to monitor and evaluate the effectiveness of such measures in order to decide locally which would be the most suitable initiatives to adapt, given the reality of finite budgets and long term consequences. Our tool would aid cities around the world by enabling them to properly evaluate their own local needs based on real data and potentially save considerable costs by choosing to implement policy decisions and urban planning measures to improve air quality by addressing their own particular problem points and not those of other cities. In addition, with the growing concern of the general public over climate change, both citizens and industry can also benefit from increased knowledge of the evolution of local air quality levels so that they can be empowered to take responsibility for their own actions while mobilising together to mitigate known sources of excessive pollution in their locales.

Reducing air pollution has a number of benefits to society, including improving the health of the population and reducing damage to crops, forests, ecology, building and other materials. A 2017 European Commission report[4] specifically assessed the costs, benefits and economic impacts of the latest clean air strategies and their implications on innovation and competitiveness. In addition to standard quantification of premature mortality and morbidity and their economic valuation, the study addressed work days lost due to the exposure to air pollution. For the EU28, costs associated with lost working days fall from €18.5 billion/year in 2005 to €8.5 billion/year by 2030 under the revised policy scenarios. Furthermore, quantification of the benefits of the revised emission ceilings indicated that they will exceed costs by a large margin. Taking a conservative position on the valuation of mortality led to benefit:cost ratios in excess of 14, and taking a less conservative

[3] Due to its wide swath of 108° (approximately 2 600 km on the ground), the TROPOMI/SENTINEL-5P instrument provides a full daily surface coverage of radiance and reflectance measurements for latitudes > 7° and < -7°, and better than 95 % coverage for latitudes in the interval [-7°, 7°].

[4] https://ec.europa.eu/environment/air/pdf/clean_air_outlook_economic_impact_report.pdf

position pushed the ratio above 50. A strong domestic market – due to domestic environmental and energy regulatory policies – emerged as an important success factor in all case studies.

Furthermore, a 2019 OECD report[5] (The economic cost of air pollution: Evidence from Europe) provides the first evidence that air pollution actually causes economy-wide reductions in market economic activity based on data for Europe. The analysis combined satellite-based measures of air pollution with statistics on regional economic activity throughout the European Union over the period 2000-15. Similar to the EC report, these results suggest that public policies to reduce air pollution may contribute positively to economic growth. Indeed, the large economic benefits from pollution reduction uncovered in the study compare with relatively small abatement costs. Thus, more stringent air quality regulations could be warranted based solely on economic grounds, even ignoring the large benefits in terms of avoided mortality.

Value for Municipality of Zwolle:

Examples of how our tool can help local policy decisions include:

- It would provide insight into how urban planning influences the nearby air quality conditions, for example by identifying:
  - which factors contribute most to the air quality in Zwolle as a whole, but also for each neighbourhood
  - which streets have or will have the worst air quality
  - when the highest concentrations occur (eg. during a traffic jam, during the summer, during new year, rush hour, parents bringing their children to school, etc.)
  - the biggest pollution sources (eg. building sites, ships, factories, roadways, etc)

- It would provide insight into the effectiveness of ongoing or past initiatives to improve air quality at the local level
- It would help prioritise which neighbourhoods should to be targeted for mitigation measures (not only due to poorer air quality, but for example, variables such as average age of local residents could be factored in to address the level of priority)
- It could be used in tandem with the planning of future citizen science projects (eg. identifying which air quality variable and in which locations sensors would be useful to gather more in situ data eg. PM2.5 data was not collected within the SensHagen pilot, but is considered one of the most)

Other potential uses of the tool could be for monitoring, regulatory, or enforcement purposes:

- By looking forward, the tool's predictive capability can be helpful for the government to make decisions to limit or target the actions of particular pollution sources (such as certain types of industrial production) over others
- By backtracking the air pollution data, we can not only predict the future state, but can also:
  - identify originating sources of abnormal or excessive levels air pollution
  - quantify the contribution of each industry for the city's current levels of air pollution

[5] https://www.oecd-ilibrary.org/economics/the-economic-cost-of-air-pollution-evidence-from-europe_56119490-en

If the Municipality of Zwolle chooses to make the output of our proposed tool publicly available, other Stakeholders for our proposed roll-out for Zwolle could include:

- ★ Dutch governmental agencies with responsibility for air quality (eg. KNMI, RVO, lenW)
  - they could learn from the data collected in Zwolle and apply lessons learned to the rest of the territory on a policy-level
- ★ Province of Overijssel
- ★ Other municipalities in the Netherlands (and elsewhere)
  - could learn from Zwolle how to become more climate-smart
- ★ researchers and universities around the world
  - data from Zwolle could be used to aid current scientific research on the impact of climate change and air pollution at a more granular level, showing variances over time within different city zones
- ★ industry
  - individual companies and industry associations would be able to view how much they contribute to local air pollution and take actions to reduce this
- ★ citizens of Zwolle
  - poor air quality is dangerous for human health, especially for those already with compromised respiratory systems - knowledge of local air pollution levels will allow citizens to take personal actions to avoid areas of poor air quality on particular days or times of day, and even consider when choosing a new neighbourhood to live in
  - ultimately, it could motivate entire neighbourhoods, such as Stadshagen, to mobilise against excessive pollution sources and engage in community projects that help improve their local air quality

Next steps (future Product Development):

•Implement step 2 and 3 of the deep learning model

•Add more data and variables in the AI model to get better accuracy

•Develop highly scalable AI model to support new places

•Interactive Dashboard for the stakeholder which supports decision making

After our proof of concept, the NO2 model can be extended further to any additional areas where the citizen science project has been extended and eventually to the rest of Zwolle. Moreover,, the model can be extended thematically by incorporation of other sensor and satellite data (and external inputs such as meteorological forecasts) to incorporate other air quality variables (eg. sulfur dioxide, ozone, fine particulates, etc). This will eventually provide the Municipality of Zwolle with a detailed predictive model of the air-quality for all of Zwolle to help it become a leading climate-smart city within the Netherlands and beyond.

Once the (data-driven) AI model is completed, data assimilation methods could be integrated to backtrack the air pollution sources and to improve the model to achieve better performance. For example, 4D-Var method -- a typical data assimilation approach, is widely used in backtracking problems and parameterization. It is very powerful in determining the initial values of models, so possible pollution sources can be worked out based on the time series of the measurements as well as the updated satellite data. Meanwhile, the data assimilation method can help to improve the model itself by optimizing its parameters by comparing the real-time measurements and the model results. In this way, the model can improve itself at each measured time and gradually output a more accurate prediction.