Big Data, Little Boats

PA Water & Waste Water Technology Summit

November 2018

Bryce Aaronson (bryceaaronson@gmail.com)

PLATYPUS

Outline

PLATYPUS

- Research at the core of our identity
- Our capabilities today
- Delftland, Netherlands Case Study
  - Longitudinal - seasonal variation
  - Areas of Interest - Looking at Nitrates
  - Identifying Leaky Greenhouses
  - The impact of weather

Mission

Platypus has developed and is selling a water analytics service, Aquatical Analytics™, that utilizes cooperative robotic boats to acquire and analyze data about inland bodies of water

- Founded in 2013 by a robotics prof at Carnegie Mellon
- Offering comprises two components - Aquatical Analytics Platform™ and Aquatical Robotic Boats™ (used to perform the data acquisition)
- Robotic boats now optimized for manufacturing / sensor agnostic (array configured to any application)
- Deployments performed around the world over the past 18 months

Water, water, everywhere

- Data about inland water is sparse, ad hoc and incomplete - more judgment than data
- Cooperative, autonomous robots make data collection orders of magnitude cheaper than manual methods
- Manual data collection and sensors in fixed locations provide some information, but does not tell the whole story, even with high cost of data collection
- Robots, capable of collecting enormous amounts of data can make water management a data-driven activity
- Current status is that policy, use and management is done with incomplete data, resulting in inefficiency
- Modern AI and big data can transform that data into actionable information for actionable water management -> Aquatical Analytics

Algorithmic Background

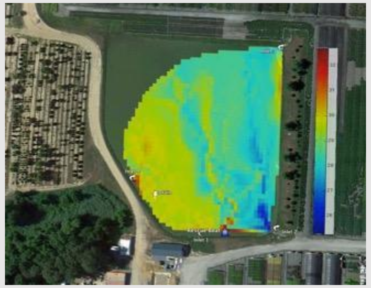
- Funding from multiple organizations for basic research

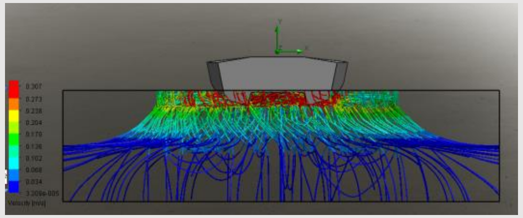
  - Information maximization

  - Energy efficiency in currents

  - Optical flow obstacle avoidance

  - Team plans

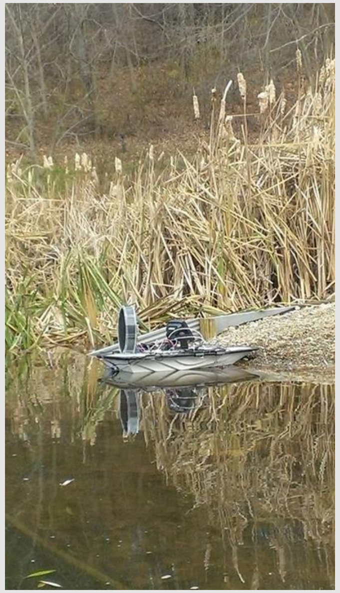




PLATYPUS

Water management

- Often combination of bathymetry and water quality
- Ad hoc data collection, clients looking to understand their water
- Working with EPA/PA-DEP on certification for permitting



THREE RIVERS WATERKEEPER®

ESTATE MANAGEMENT SERVICES

CITY OF EDINA, MINNESOTA · INCORPORATED 1888

The CITY of EDINA

PLATYPUS

Research

- Variety of uses
  - Water and robotics
- Active projects building on open source software and new sensor integration
  - Understanding how data can be useful for water management

UNIVERSITY of WASHINGTON

PRINCETON UNIVERSITY

INTCATCH 2020

PUCRS
Pontifícia Universidade Católica do Rio Grande do Sul

Water Data capabilities

- Sensors currently on boats:
  - Bathymetry, with vegetation, bottom hardness
  - Dissolved oxygen
  - Electrical conductivity
  - Temperature
  - Ph
  - Oxygen reduction potential
- Measurements typically take 10cm below the surface
  - Profiler allows for deeper measurements
- Spatially dense data
  - Exhaustive or targeted (e.g., max)

Dutch water boards

- Pilot project in an irrigation area near The Hague, Netherlands
  - 15 miles of small canals supplying water to intensive agriculture
- The need is to identify greenhouses and fields that are over-contributing to pollution in the water
  - Pilot identified two greenhouses “leaking” justifying the expense
- Aquon, an umbrella organization for ⅓ of the NL’s water boards is taking on the technology to roll it out across the country

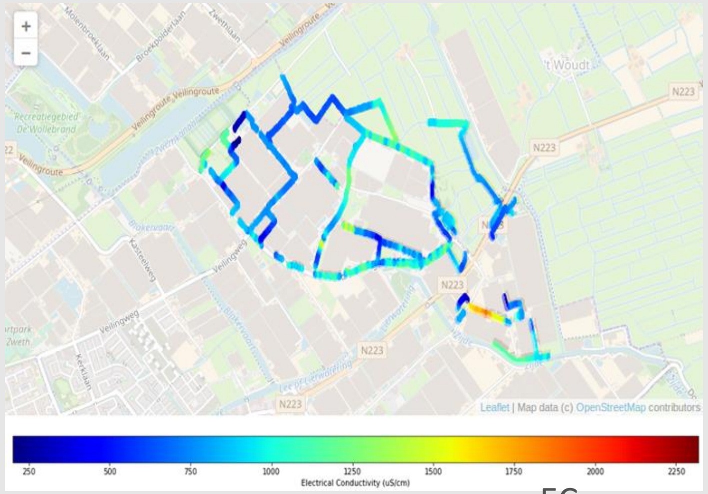
Longitudinal Collections

Data Collected over a period of time

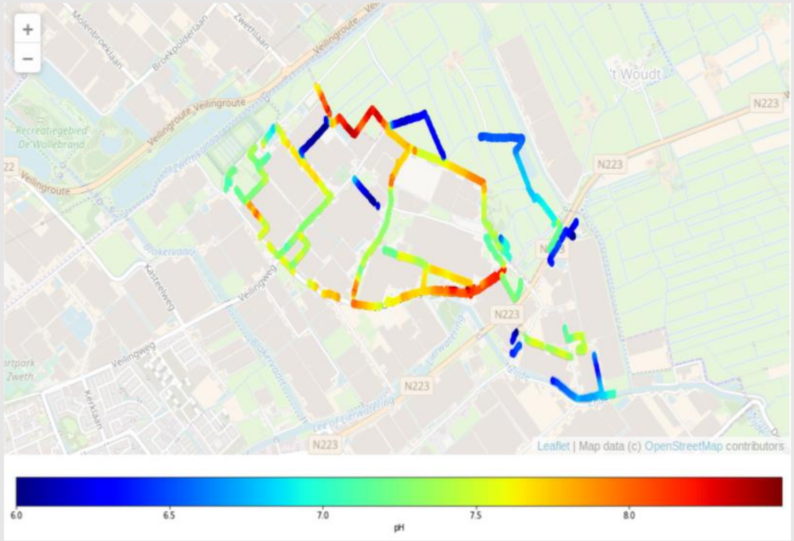
-4 Days in November 2017

- 4 Days in Spring 2018

- Over 10 Days in September 2018



EC measured across the ponder, identifying one area of surprisingly high salinity



Ph measured across the ponder, showing interesting variation and grounds for investigation

Sensors - November 2017

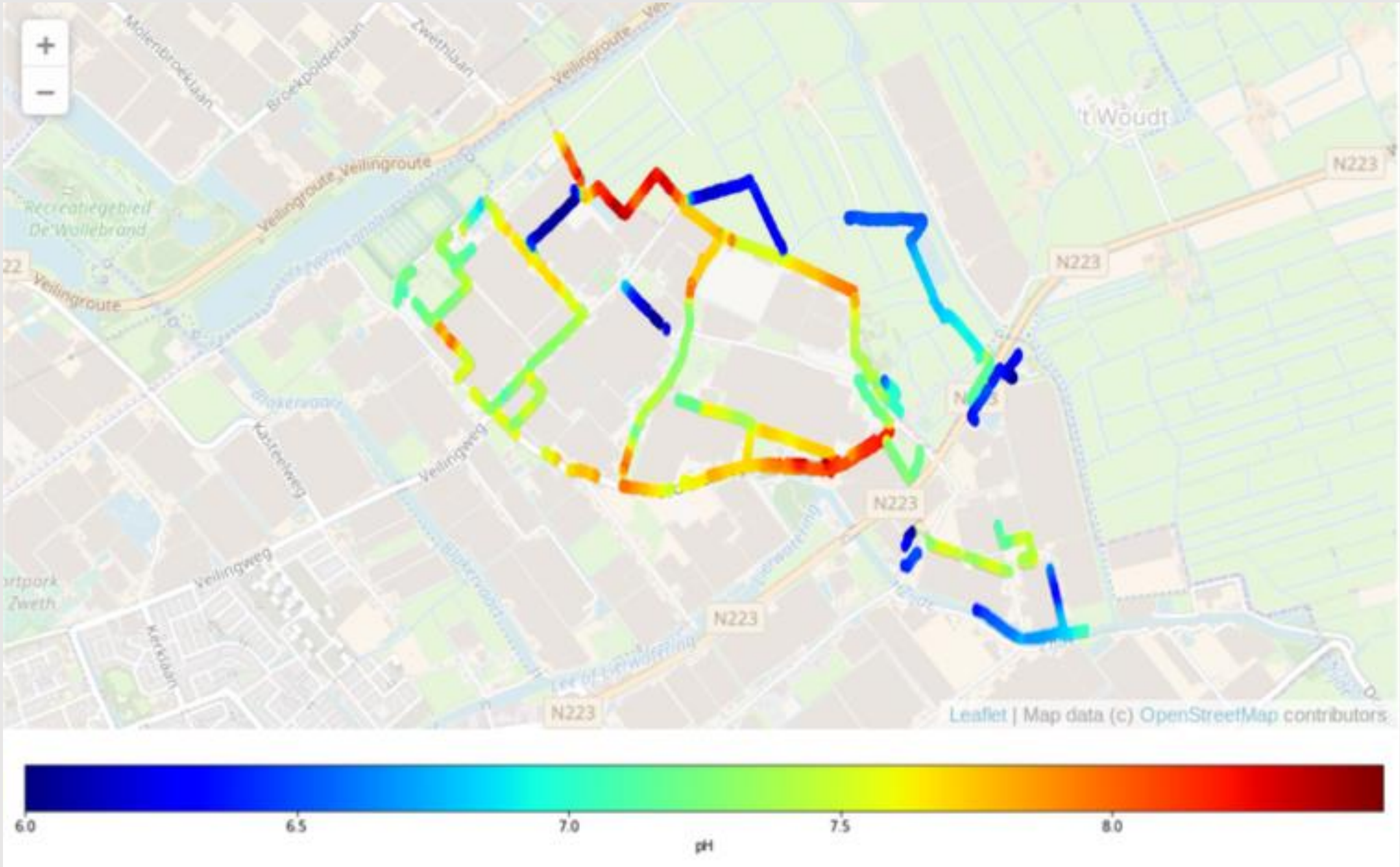
- Atlas Scientific pH
- Meter Group ES-2 - Electrical Conductivity & Temperature Sensor

- Both have similar response times, temperature ranges.
- Different Calibration Requirements

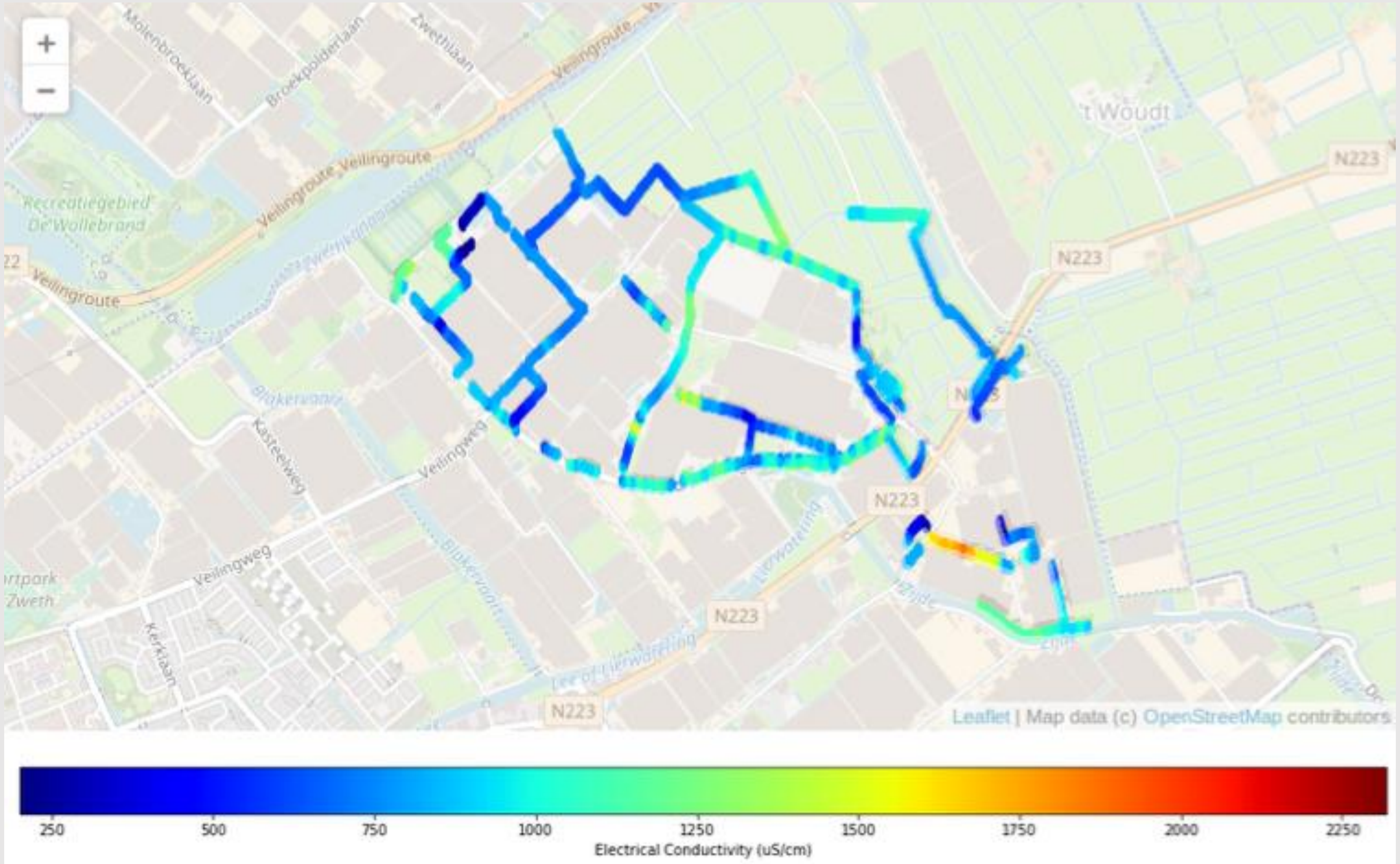
https://www.atlas-scientific.com/_files/_datasheets/_probe/pH_probe.pdf

http://library.metergroup.com/Manuals/13890_ES-2_Web.pdf

Overall pH - November 2017

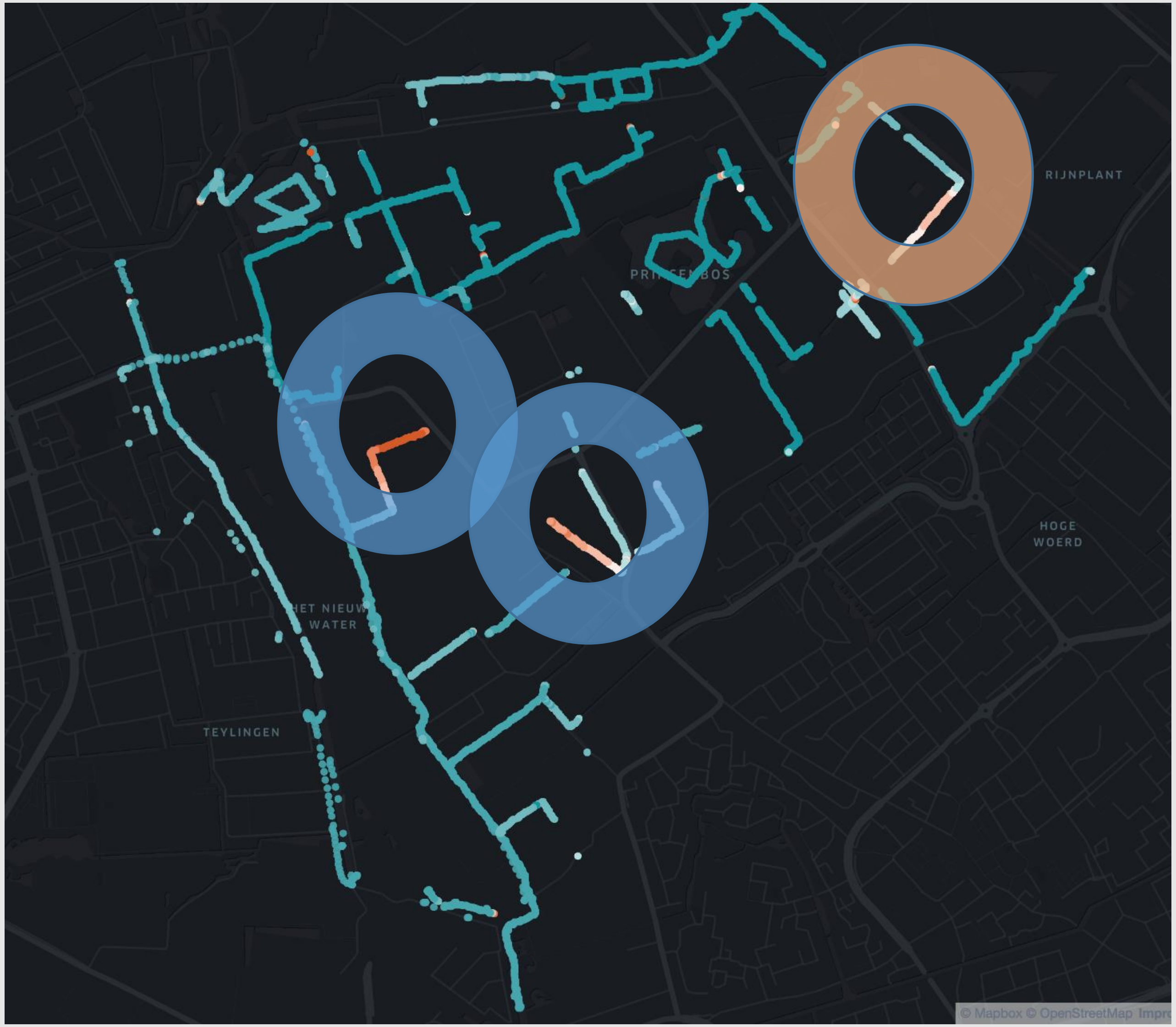


Overall EC - 2017



September 2018 Collection

- Used In-Situ AquaTroll 600 and sensors for all data collection
- Data collection took place in two separate times, separated by a week
  - Due to need to get a new sensor to replace a damaged one
- NO3 data shows overall good water quality, with three areas of high (> 30 mg/L) levels
  - Areas circled on adjacent map
  - Orange circled area is a little surprising not in an obvious collection point



Week to Week differences

- Data was distinctly different from the first data collect to the second data collect
- Area covered with the yellow polygon was first week, rest in second week

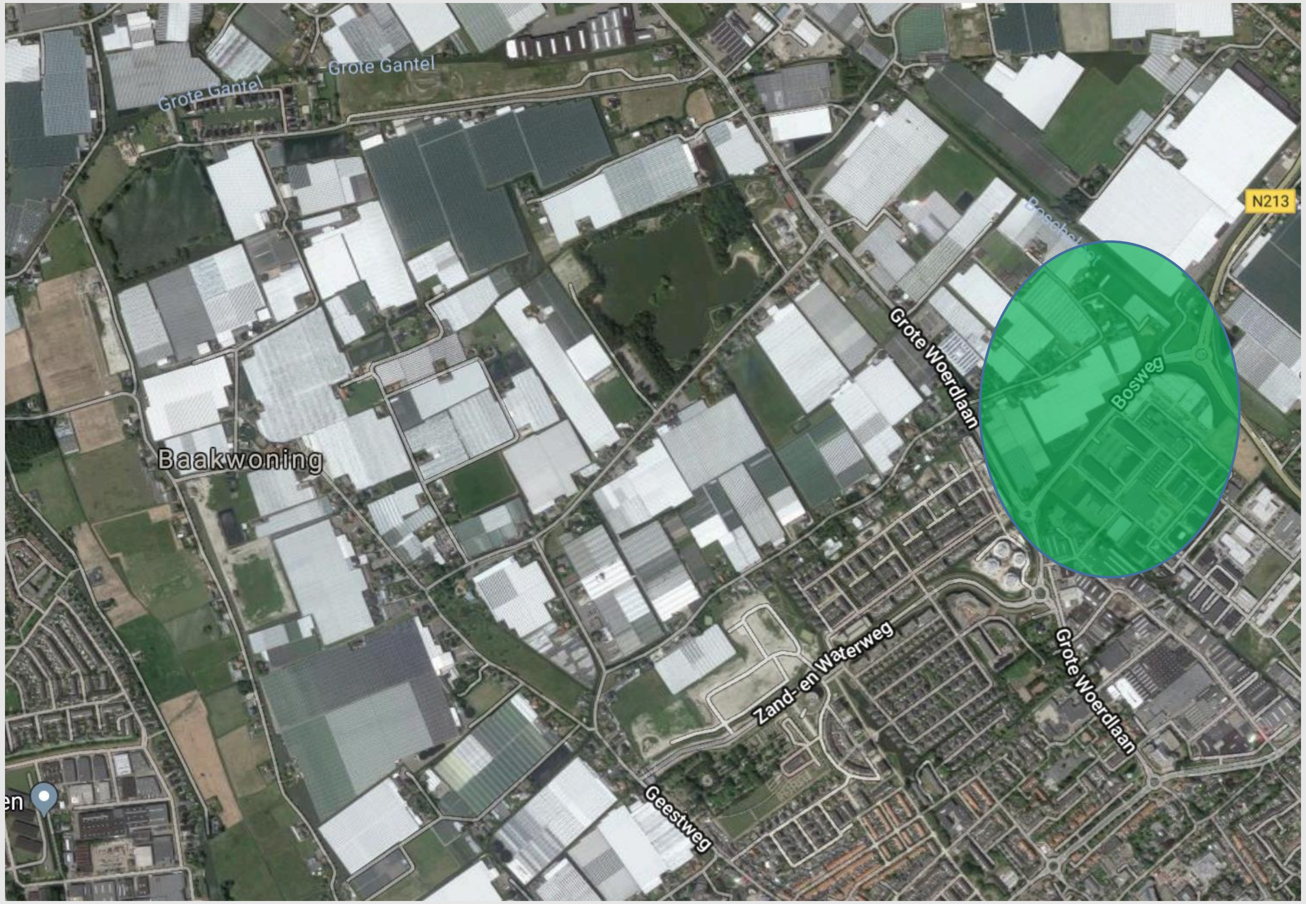
	NO3	pH	RDO	EC
First	3.73	8.29	10.8	445
Second	12.7	8.03	8.6	622

Alternative color scaling

- This alternative color coding makes the difference clearer, with the second week nearly all in orange or red versus the first week in blue
- However, notice that in absolute terms the difference is relatively small
  - This color coding shows relatively clean and healthy water in both blue and orange
  - All the NO3 data in blue on the **previous** page is considered below safe drinking levels by the US federal government

Low NO3 near shopping

- We would be tempted to conclude a calibration error on the second week that produced consistently higher numbers
- However, the area marked in green got some of the lowest NO3 numbers
- This area is on the edge of the irrigation area near a shopping district
- We do not have specific information that this area does have less fertilizer runoff, but it would be consistent with the type of area
- The presence of some low values during the second data collect reduces the probability that a simple calibration error that only inflated values was to blame



High sensor correlation

- The second data collect had higher NO3, higher EC, lower pH and lower DO
- The relative values correlate imperfectly, as would be expected
  - Some correlation, but not perfect correlation
  - Correlation is consistent across the data collections
    - This need not be the case, depending on what caused the change in the water, but the fact that the correlation is the same reduces the probability of sensor error
- The EC sensor is never calibrated and the optical DO sensor is rarely calibrated
- If this was only a calibration or sensor error problem, the uncalibrated sensors would not change in the same way as the calibrated sensors

NO3

EC

pH

DO

Concerns and Explanations

- Since data collected at different times was statistically different we need to determine whether this was a data collection problem or a change in the water
- A data collection problem might be due to:
  - A fault in the sensor
  - Mis-calibration of the sensor
  - Incorrect operation
  - Data processing error due to the many small segments that needed to be assembled
- Changes in the water might be due to:
  - Weather or other factors influencing processes in the water
  - Water coming into the system or leaving it

Calibration and Collection Process

- The sensors were calibrated according to manufacturer instructions
- A test after calibration with a text solution showed the sensors to be reporting accurate values
- The data was collected by two technicians that have collected data before
  - They reported no unusual circumstances or robot or sensor behavior that would cause problems with the data
  - There was a change in the phone on the robot during the second data collect, however Platypus technicians can see no reason why this would cause a problem with the data
- The internal statistical properties of the data were consistent during both data collects, e.g., no spikes, approximately the same noise levels and distributions of values (though around different means)

Conclusions

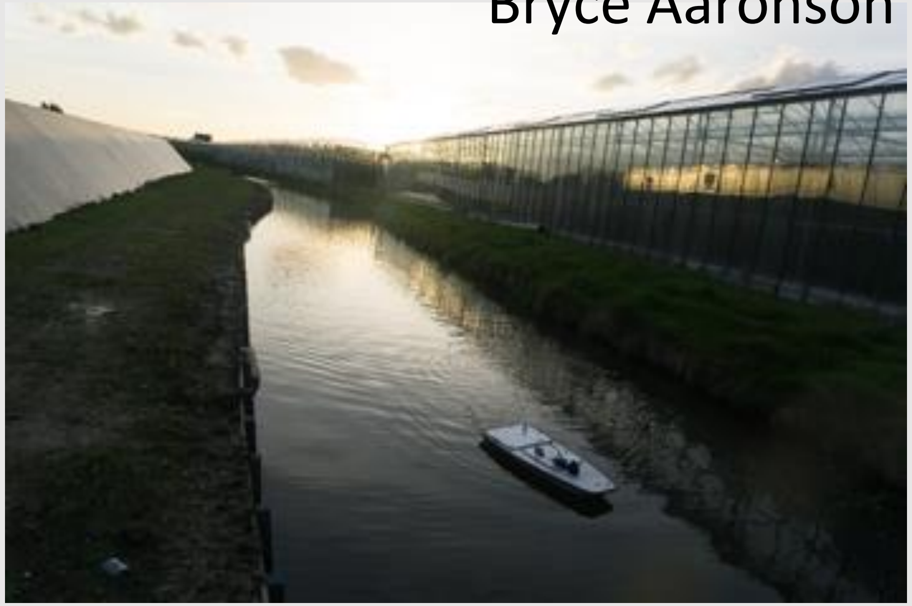
- Three problematic areas identified
- Concerns about data differences between data collection efforts
  - Conclusion is that the water changed over the week
- Focus on relative numbers is important for trouble spots
- Repeatedly covering the same area at different times can show how the water changes
- These early runs serve as the beginning of a long term Netherlands study

Big Data, Little Boats

Questions?

Paul Scerri (CEO) - Paul@senseplatypus.com

Bryce Aaronson (Tech) - bryceaaronson@gmail.com)



PLATYPUS