GEOG 365
Introduction to GIS Programming

Instructor: Weiming Hu

Time: Tu Th 12:05 PM - 1:20 PM

Email: weiming@psu.edu

Place: 208 Walker Building

PennState

Outline

- Personal Introductions
- Course Introduction
- Syllabus

PennState

Personal Introductions

- 4th year Ph.D. candidate
- Computational Algorithm, Remote Sensing, GIScience, and HPC
- Guitar, low-cost devices

PennState

Personal Introductions

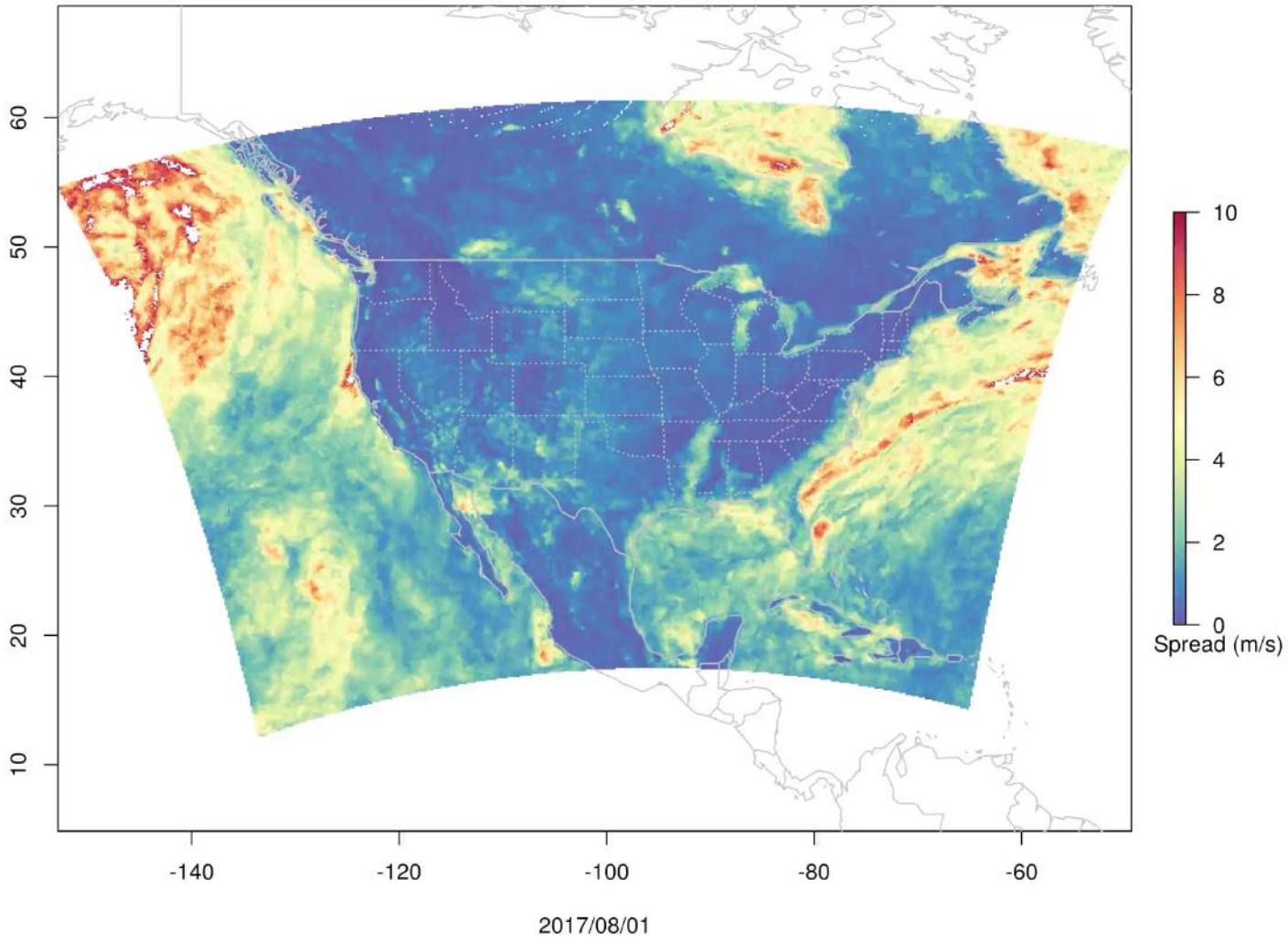
Wind speed ensemble prediction

- NCEP WRF-NAM and Analog Ensemble
- 11-km spatial resolution
- Daily forecasts with a 3-hour interval
- 1 year of ensemble forecasts for surface wind speed
- About 7,000 core hours on the supercomputer Cheyenne

PennState

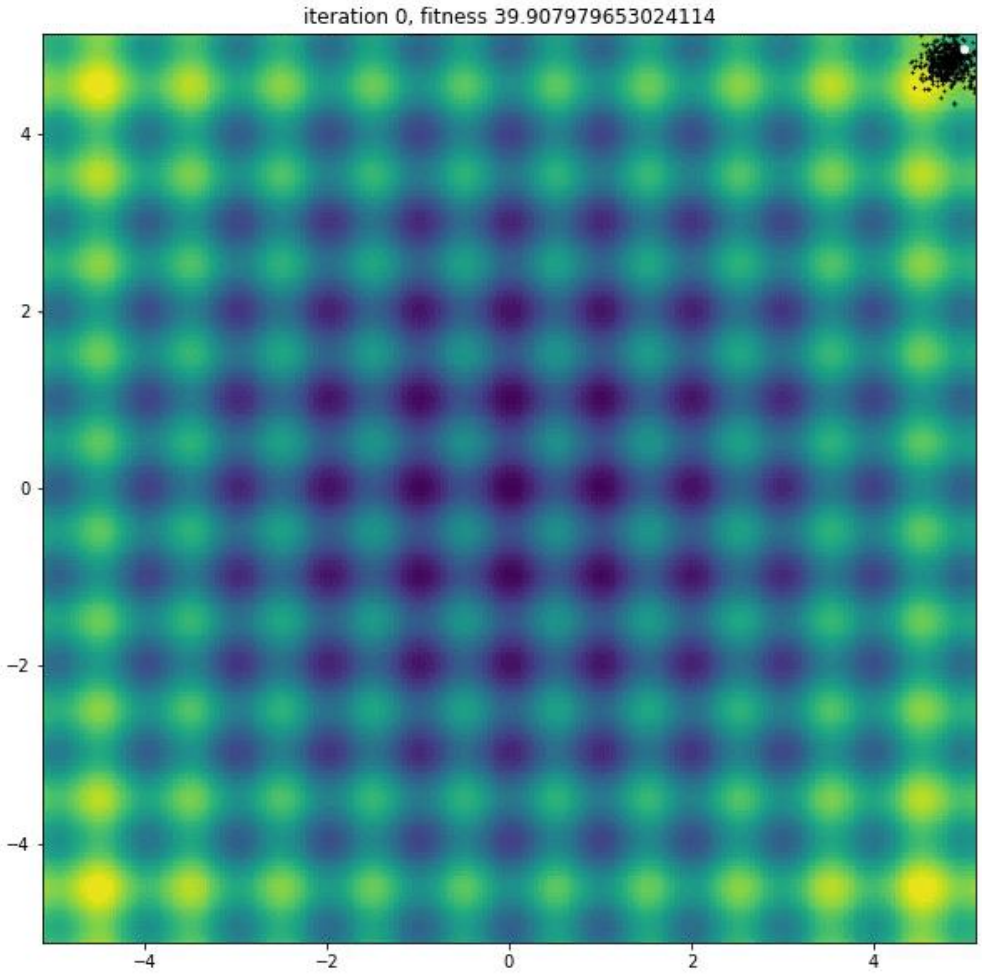
Weiming Hu, Guido Cervone. "PAnEn: An Efficient Parallel Implementation ofAnalog Ensemble (AnEn)". Earth Science Informatics (under review).

Personal Introductions



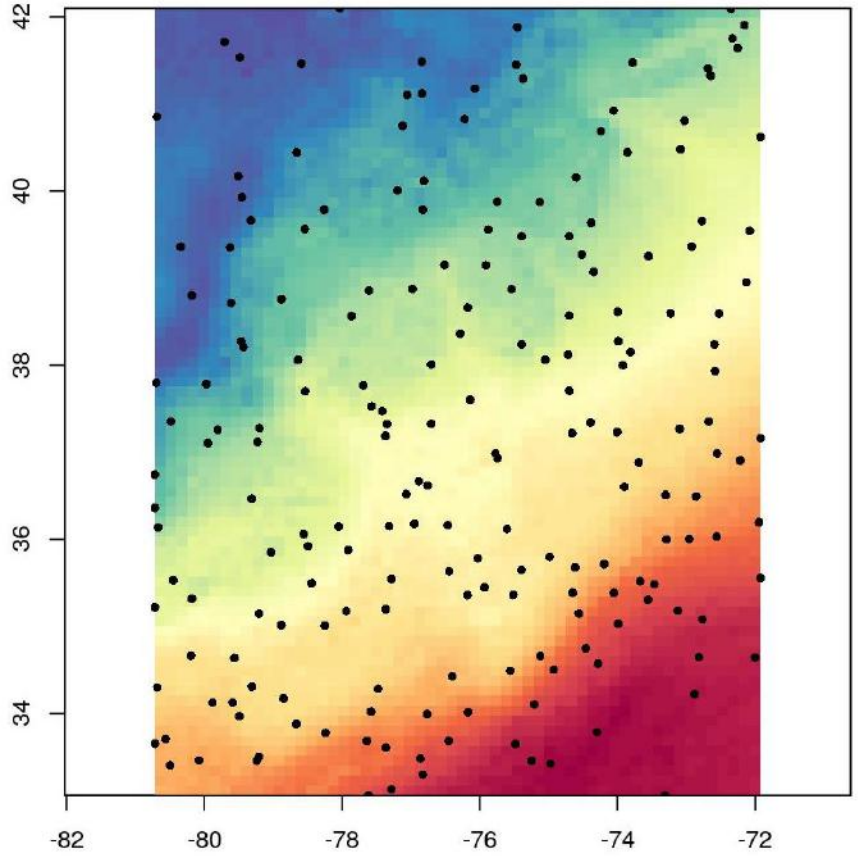
PennState

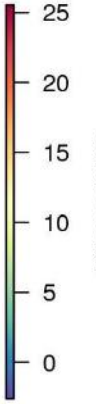
Personal Introductions

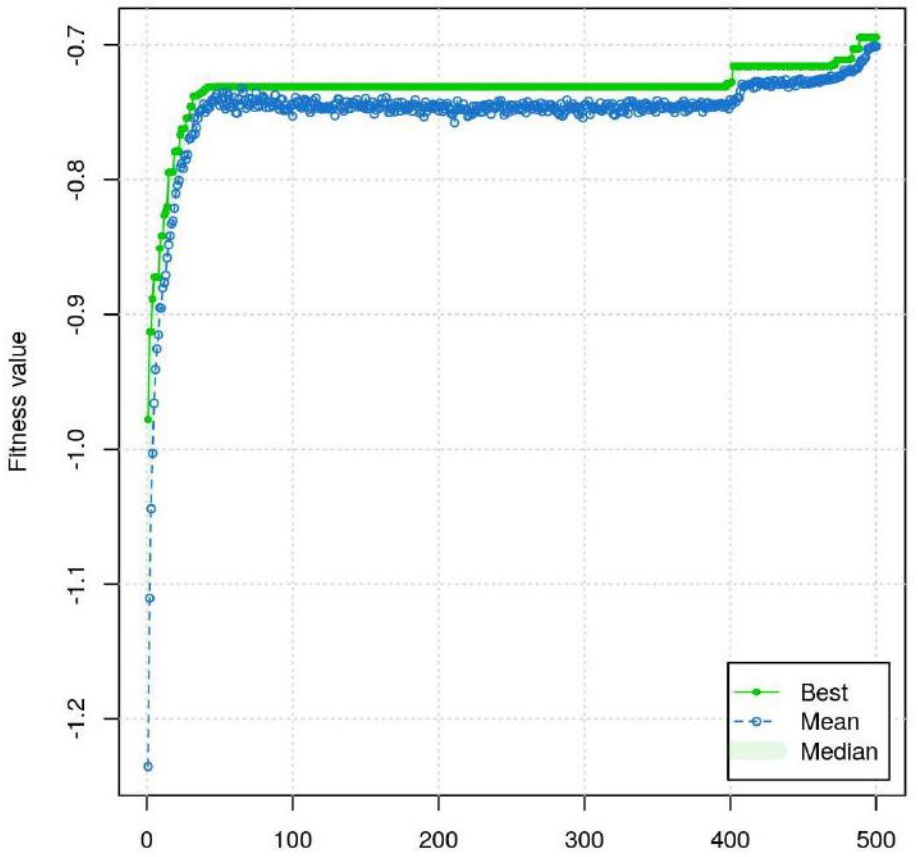


Evolving Solutions With Genetic Algorithms https://medium.com/@LeonFedden/evolving-solutions-with-genetic-algorithms-e3122473a35f

Personal Introductions





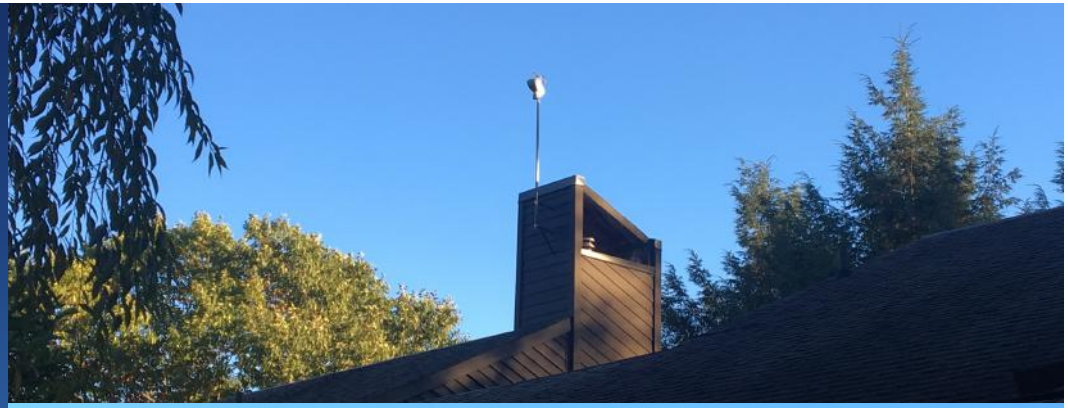


https://www.youtube.com/watch?v=iU1TLynEJIU

PennState

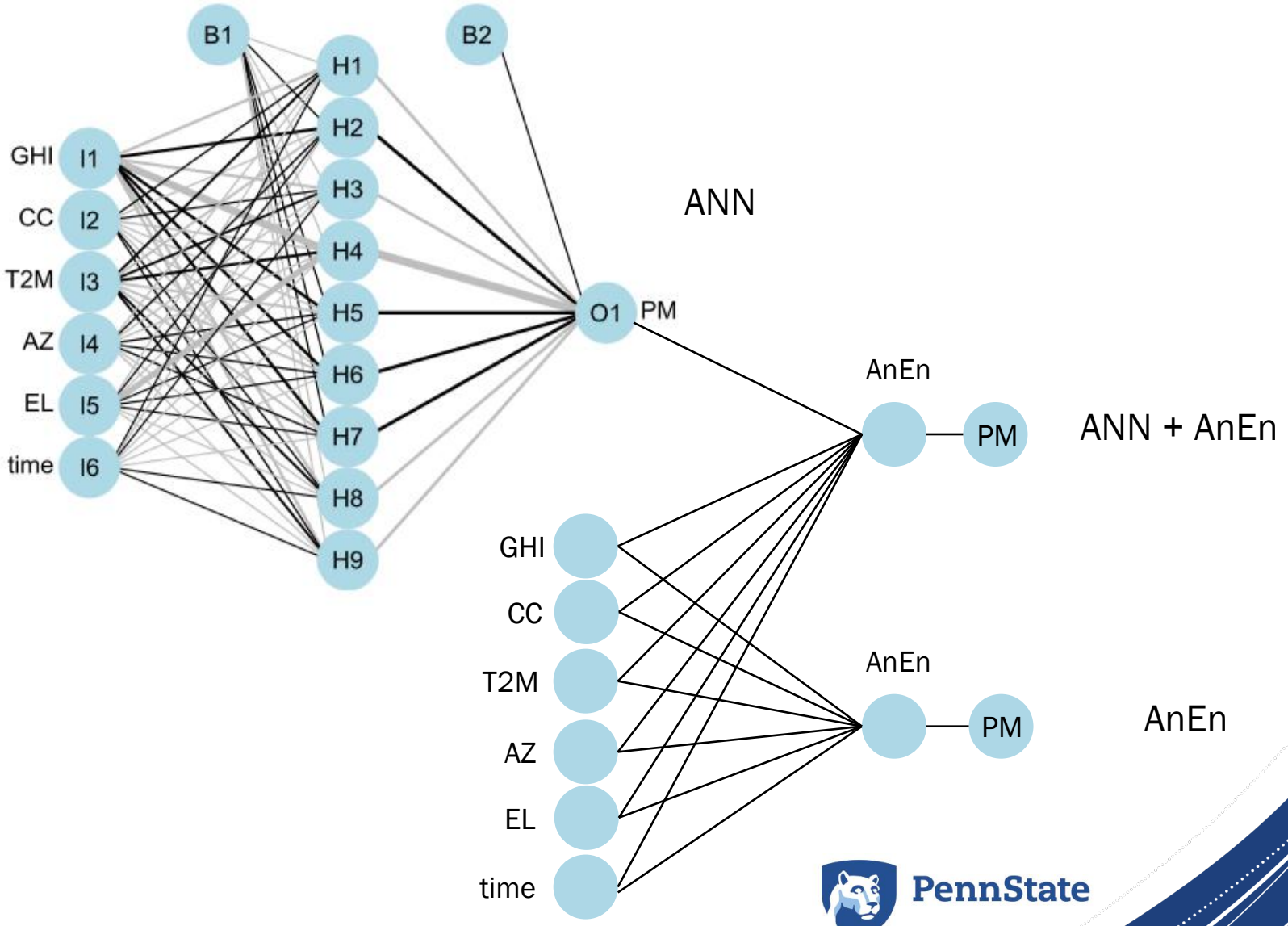
Hu, Weiming, and Guido Cervone. "Dynamically Optimized Unstructured Grid (DOUG) for Analog Ensemble of numerical weather predictions using evolutionary algorithms." *Computers & Geosciences*, (2019).

Personal Introductions



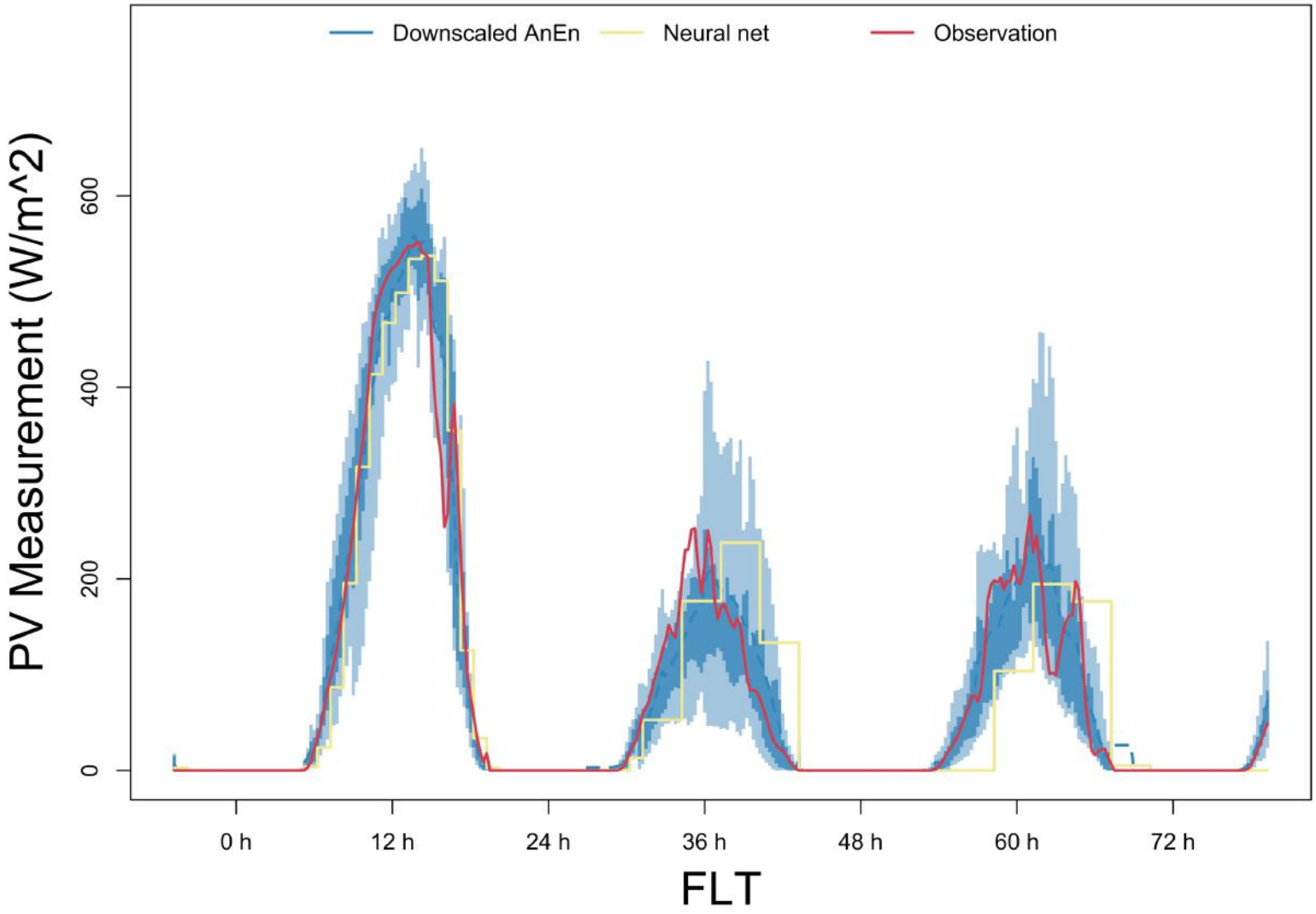
PennState

Personal Introductions



PennState

Personal Introductions



Personal Introductions

- Parallel Analog Ensemble
  - https://weiming-hu.github.io/AnalogsEnsemble/
- MoreParallelR
  - https://weiming-hu.github.io/MoreParallelR/

PennState

Personal Introductions

- Majors
- Expectations and interest
- Past experiences with GIS software
- Past experiences with R or any programming language

PennState

Course Introduction

What is GIS?

Geographic Information Systems (GISs) are computer assisted systems for the capture, storage, retrieval, analysis and display of spatial data.

PennState

Clarke, Keith C. "Advances in geographic information systems." *Computers, environment and urban systems* 10.3-4 (1986): 175-184.

Course Introduction

What is GIS?

- Geographic
  - Implies that locations of the data are known, or can be calculated, in terms of geographic coordinates;
- Information
  - Implies that data in a GIS are organized to yield useful knowledge, often as colored maps and images, but also as statistical graphics, tables, and various on-screen responses to interactive queries;
- System
  - Implies that a GIS is made up from several inter-related and linked components with different functions. Thus, GIS has functional capabilities for data capture, input, manipulation, transformation, visualization, combinations, query, analysis, modelling, and output.

PennState

Sumanta Das, An Introduction to Geographic Information System
https://www.slideshare.net/sumantagargibhattacharyadas/geographic-information-system-29590419?next_slideshow=3

What is missing?

Course Introduction

Problem 1: The missing New Zealand

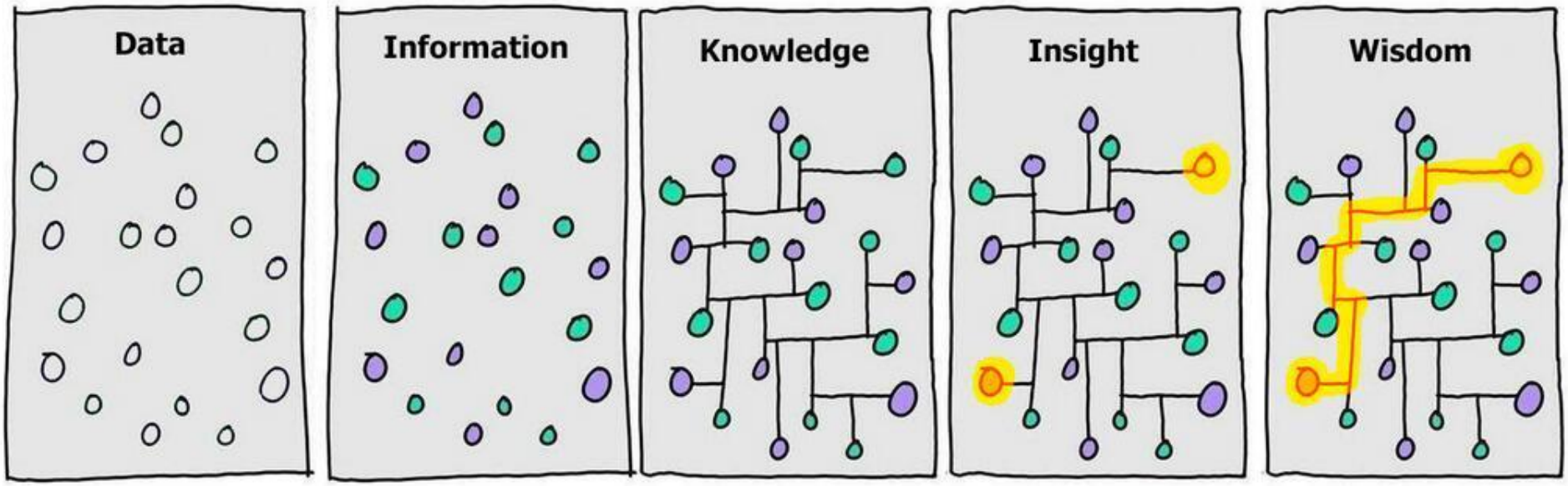
The country that keeps getting left off maps, BBC, https://www.bbc.com/news/av/magazine-41905040/the-country-that-keeps-getting-left-off-maps

PennState

IKEA Map Poster Omits New Zealand, https://www.maproomblog.com/2019/02/ikea-map-poster-omits-new-zealand/, The Map Room
New Zealand map from Last Week Tonight, https://www.hbo.com/last-week-tonight-with-john-oliver/2019/91-episode-150

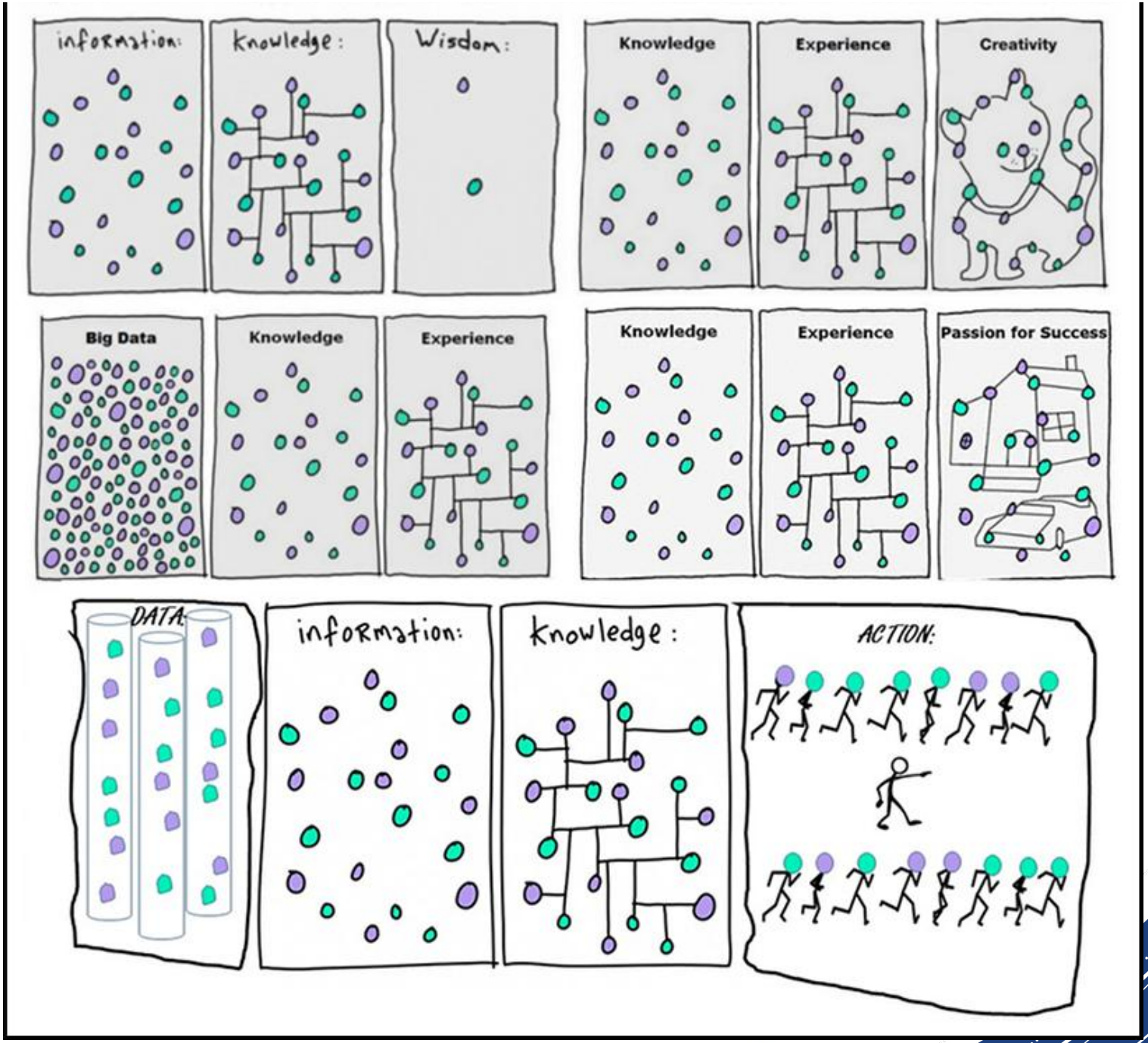
Course Introduction

Problem 2: Data? Information? Knowledge?



PennState

Difference Between Information and Knowledge, https://keydifferences.com/difference-between-information-and-knowledge.html
Data To Wisdom Via Information, Knowledge & Insight, https://www.theifactory.com/news/gaining-wisdom-from-data/


information:
knowledge:
Wisdom:
Knowledge
Experience
Creativity
Big Data
Knowledge
Experience
Knowledge
Experience
Passion for Success
DATA:
information:
knowledge:
ACTION:

Course Introduction

Problem 2: Data? Information? Knowledge?

Any real-world examples that follows the data-information-knowledge path?

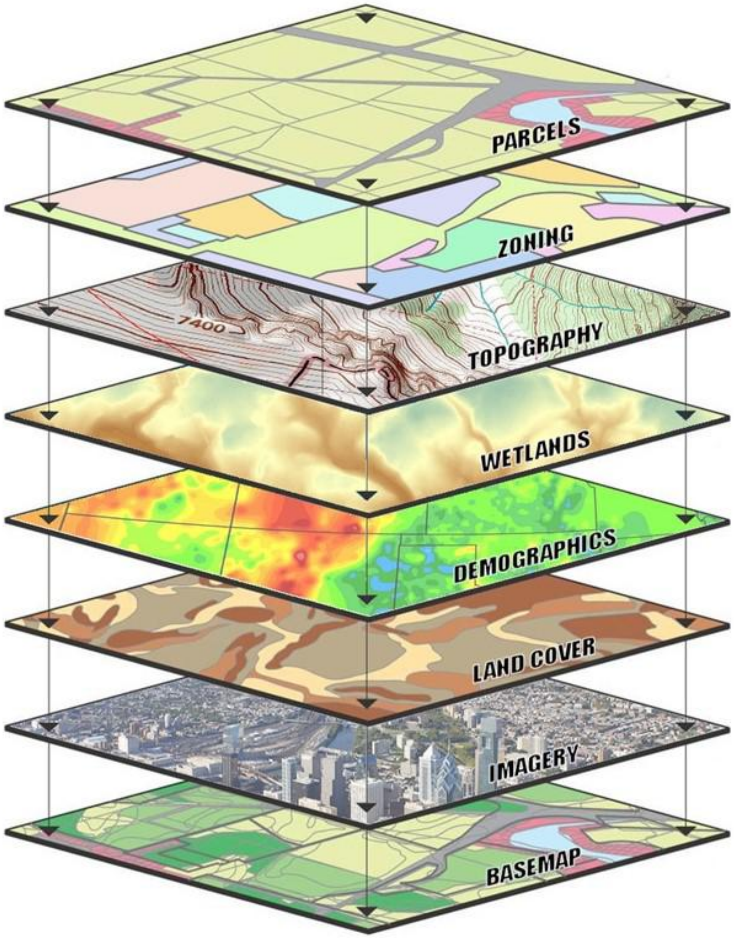
Course Introduction

GIS components

Land Infotech, http://www.land-infotech.com/

Course Introduction

Data integration



GIS DATA LAYERS

Many different types of data can be integrated into a GIS and represented as a map layer.

Examples can include: streets, parcels, zoning, flood zones, client locations, competition, shopping centers, office parks, demographics, etc.

When these layers are drawn on top of one another, undetected spatial trends and relationships often emerge. This allows us to gain insight about relevant characteristics of a location.

PennState

Ontario County, New York, https://www.co.ontario.ny.us/1176/GIS-Program-Background

Course Introduction

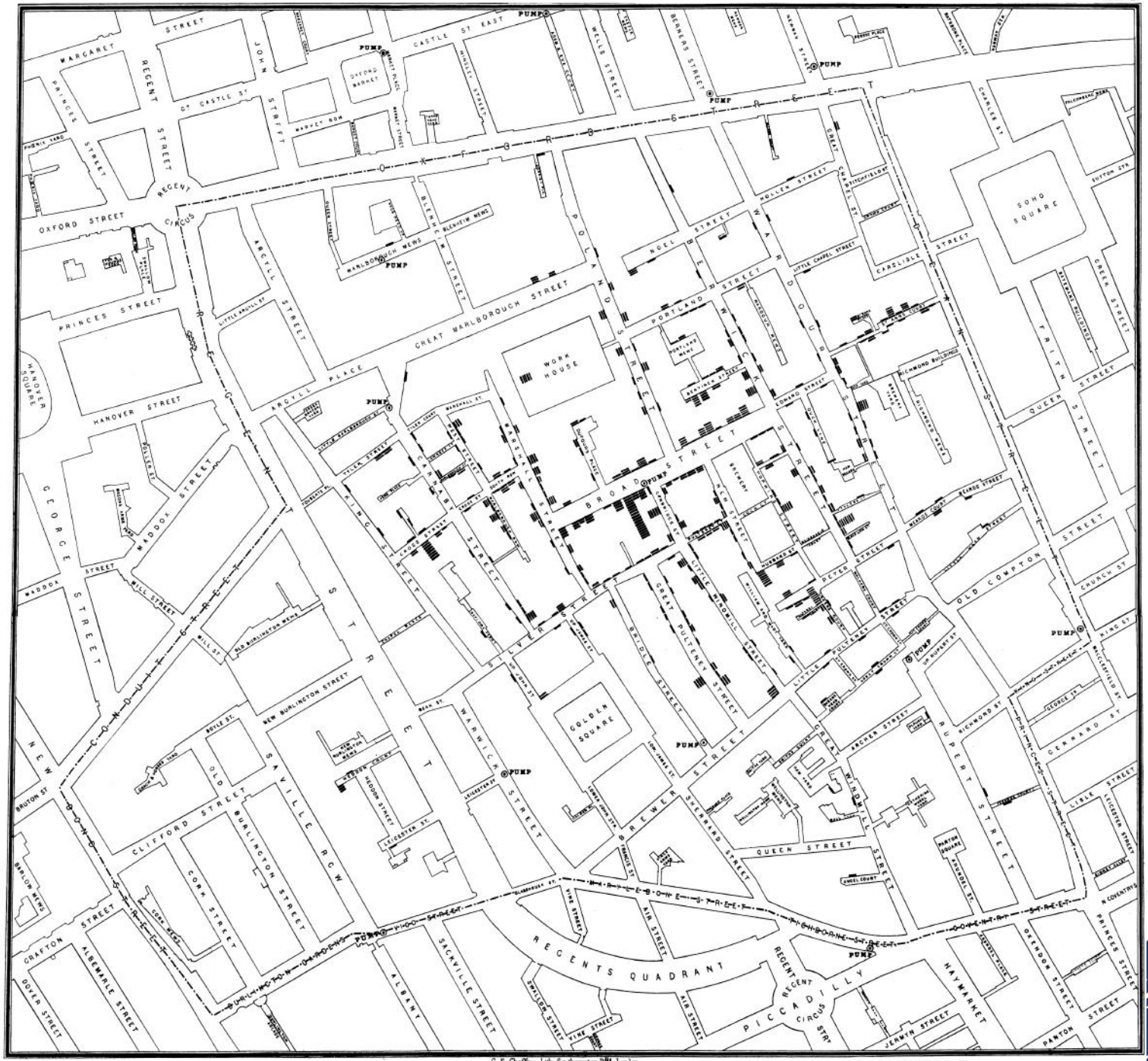
John Snow's Cholera Map

John Snow's well known cholera map is often cited as one of the earliest known examples of using geographic inquiry to understand a health epidemic. Starting on August 31, 1854, an outbreak of cholera hit the London district called Soho. Over the course of three days, 127 people died from the disease and by September 10, over 500 had died.

Physician John Snow was able to narrow down his suspicion of the source of the disease to a water pump on Broad Street. Snow mapped out the locations of the local water wells and cholera deaths to back his argument about the contaminated pump, convincing the local council to remove the handle to prevent its use.

PennState

Caitlin Dempsey, John Snow's Cholera Map using GIS Data, https://www.gislounge.com/john-snows-cholera-map-gis-data/
Simon Rogers , John Snow's data journalism: the cholera map that changed the world, https://www.theguardian.com/news/datablog/2013/mar/15/john-snow-cholera-map#data



C.F. Cheffins, Lith. Southampton Bldgs. London.

SCALE 30 INCHES TO A MILE.

Course Introduction

John Snow's Cholera Map

Contemporary critics:

- It was believed that bad airs originating from a 1665 plague burial site on which housing had been built and under which new sewers had been laid might be a source of cholera contagion;

- It was believed the sewers themselves, as independent conduits of foul odors generated by humane waste sites, might be complicit.

PennState

Koch, Tom, and Kenneth Denike. "Crediting his critics' concerns: Remaking John Snow's map of Broad Street cholera, 1854." *Social science & medicine* 69.8 (2009): 1246-1251.

Course Introduction

John Snow's Cholera Map

Can we do better with modern GIS?

- Overlay analysis
  - Analysis with the burial site and sewer line data shows lower mortality than the observed.

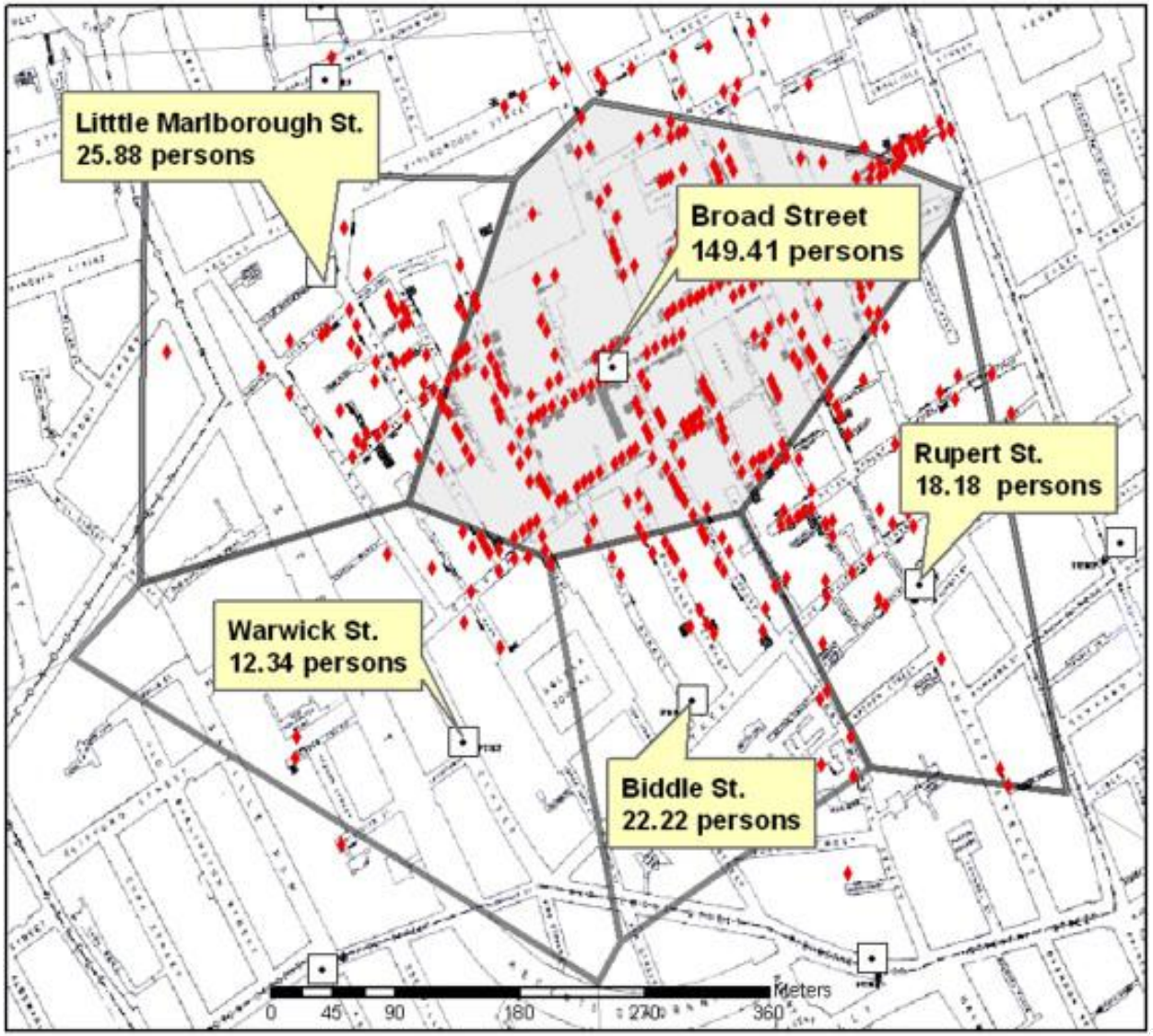
- Dirichlet tessellation and detailed quantification
  - 93.6% of all deaths mapped by Snow occurred in the central five water service areas.

PennState

Koch, Tom, and Kenneth Denike. "Crediting his critics' concerns: Remaking John Snow's map of Broad Street cholera, 1854." *Social science & medicine* 69.8 (2009): 1246-1251.

Remaking Snow's 1855 Map

Cholera Mortality per 1,000 persons



Koch, Tom, and Kenneth Denike. "Crediting his critics' concerns: Remaking John Snow's map of Broad Street cholera, 1854." *Social science & medicine* 69.8 (2009): 1246-1251.

Course Introduction

Why do we need programming when various GIS software exists?

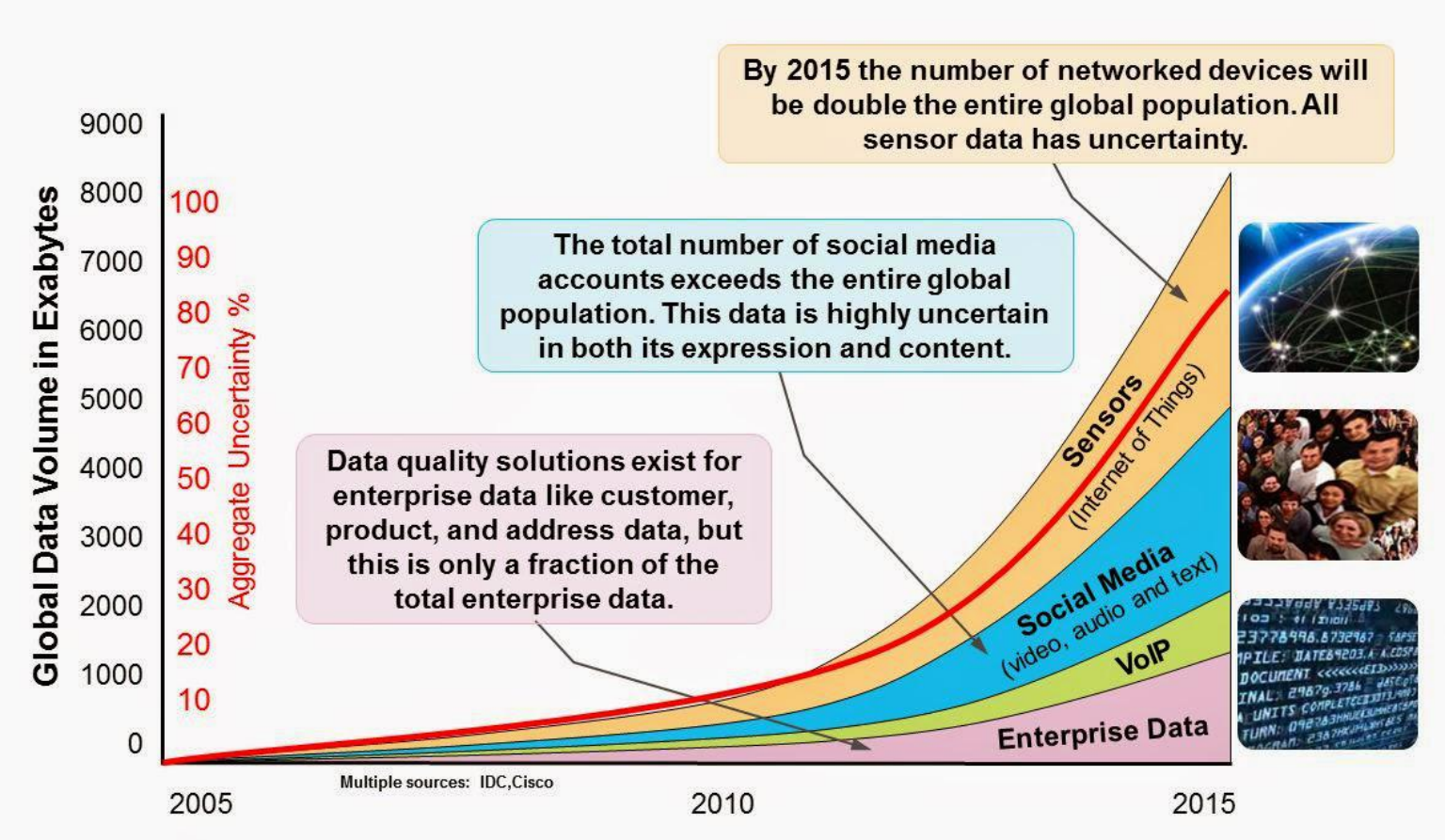
In John Snow's case, he was limited in data and efficient analytical tools.

In Tom Koch's case, "The counting and analysis for these calculations took about four days of on and off deskwork."

Can we do better?

Course Introduction

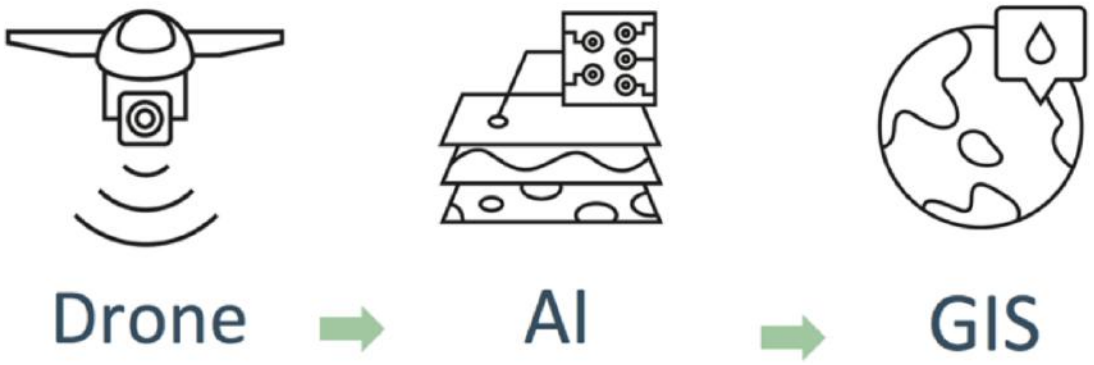
The exponential growth of data



Is Big Data Objective, Truthful, and Credible? https://john-poppelaars.blogspot.com/2015/01/is-big-data-objective-truthful-and.html

Course Introduction

Automation means easier, faster, more accurate

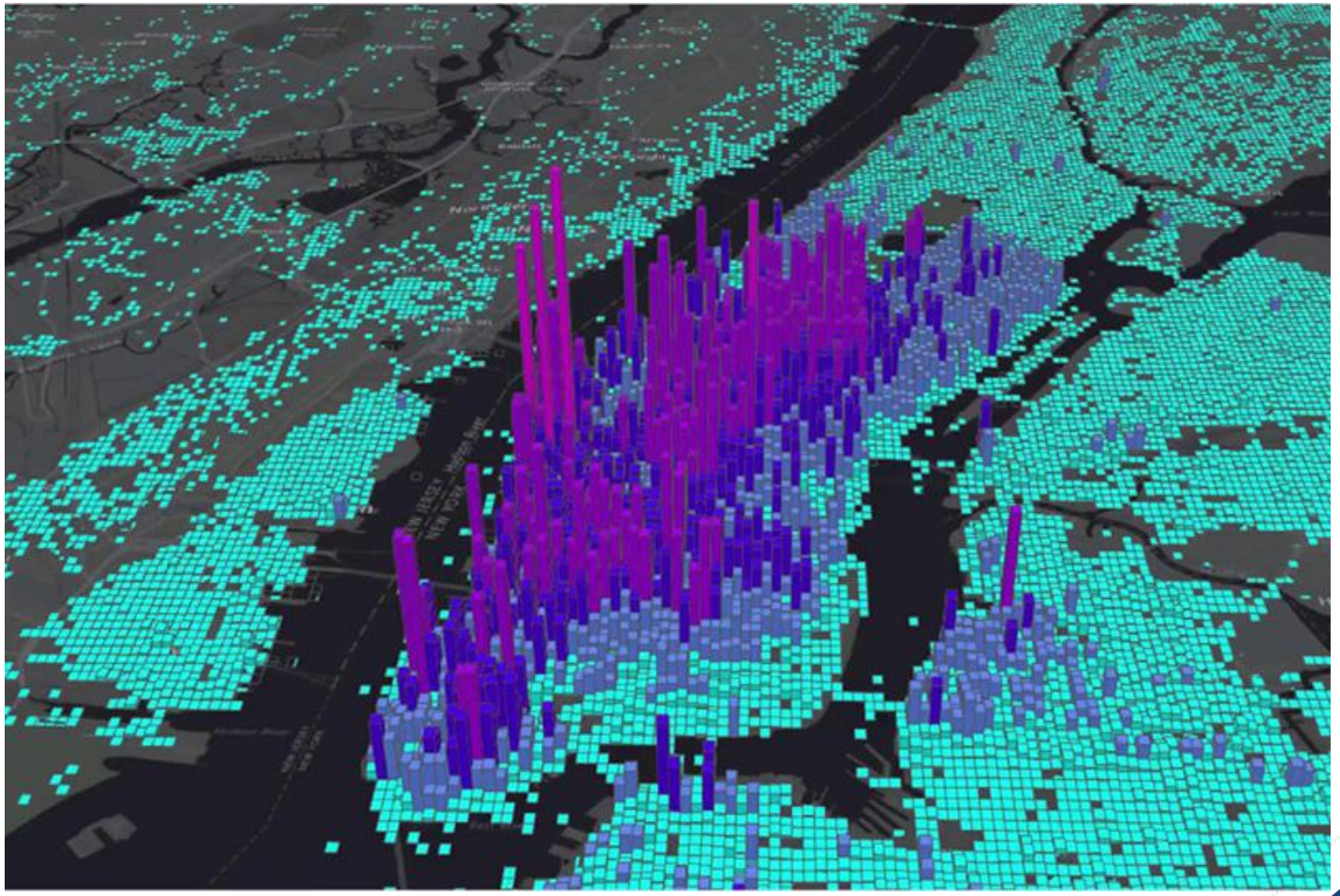


PennState

How drones are revolutionizing GIS, https://www.dronitech.com/how-drones-are-revolutionizing-gis/

Course Introduction

Automation means easier, faster, more accurate



PennState

5 case studies for merging BIG Data Analytics with GIS, http://www.qualitylifemag.com/case-studies-big-data-analytics-gis/

Course Introduction

Why R?

- R is a multi-platform, open source language and environment for statistical computing and graphics;
- R has a rich collection of packages for geographic prepresentation and analysis;
- Flexibility to other software and programming languauges;

But,

- R is slow;
- R is memory-intensive;
- R code can be messy.

Syllabus

The latest version of syllabus can be found at

- Overleaf (http://bit.ly/geog365)
- Canvas

PennState