IPACK2017-74304:EDGE COMPUTING AND CONTEXTUAL INFORMATION FOR THE INTERNET OF THINGS SENSORS

Lavanya Turlapati ,Nigel Hinds, Dhruv Nair, Levente Klein

IBM TJ Watson Research Center,

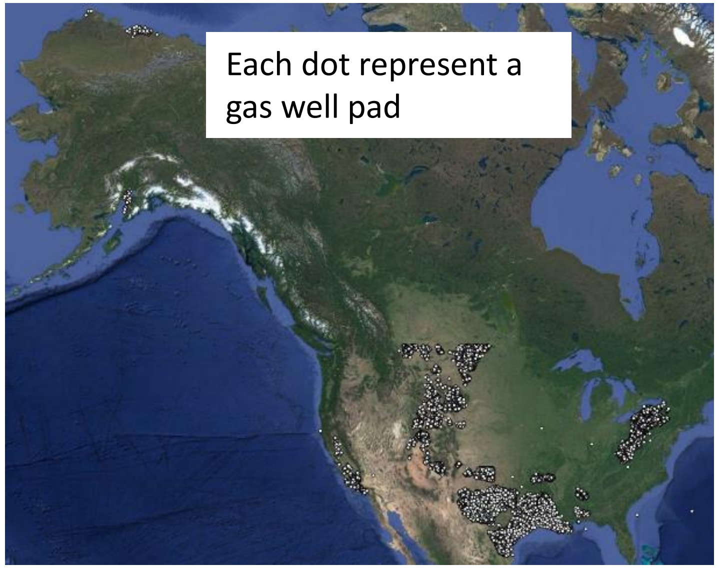
Yorktown Heights, NY 10598

Technology developed in partnership with: arpa·e awardee

Fugitive Methane Emissions in Natural Gas Processing

Methane (CH_4) is the second largest contributor to global warming after CO_2

- Greenhouse warming potential of CH_4 is 37 × greater than CO_2*



> 0.5 Million active oil and gas wells in the U.S.:

- ~30% of U.S. anthropogenic methane emissions
- ***Estimates: Leakage rate is 2-10% of total production!***

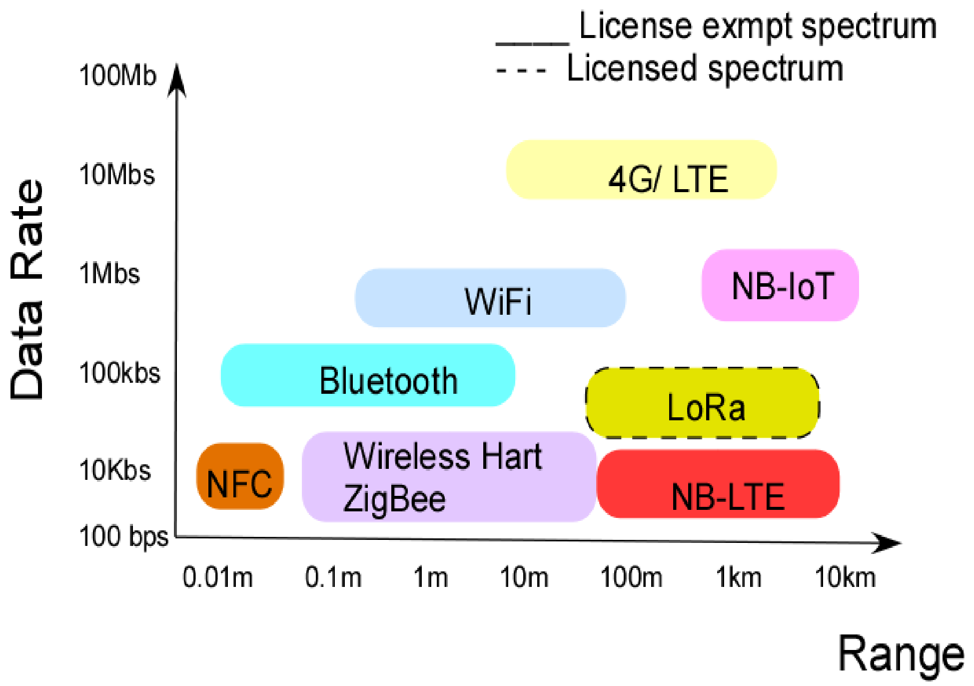
*Alvarez et. al., Proc. Nat. Acad. Sci., 109 (17), pp. 6435-6440, (2012)


Waste Water Tanks
Condensation Tanks
Well Heads

Remote sensing requirement

Power harvesting:

- ❖ solar panel
- ❖ batteries
- ❖ provide enough charge for 5 consecutive day of operation



All communication protocols needs to be time synchronized and made compatible for:

- ❖ Optimum sensor data sampling rate
- ❖ power management,
- ❖ Minimized communication bandwidth,
- ❖ computational workload
- ❖ analytics optimization on cloud/edge.

Communication	Bandwidth
loRa	2 Kbs
Celullar	10Mbs
WiFi	1Gbs

System performance

AIMS's system integrates 4 communication protocols

1. Wireless HART: motes
2. Serial RS-485: Wind sensors
3. Wide area network: LoRa
4. Satellite communication: cloud

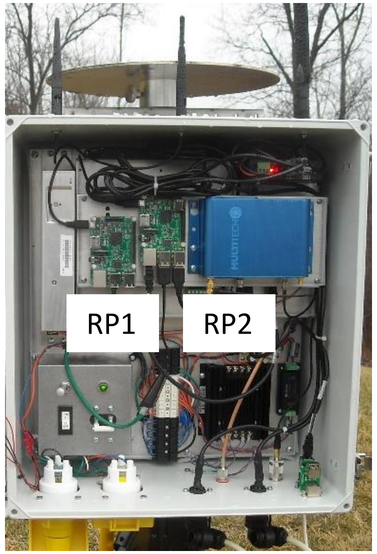
Power requirement:

1.Current ~600mA
2. Voltage 3.3-12V
3. Power ~2W

Overall strategy:
-efficient edge computation
-minimize data transmission

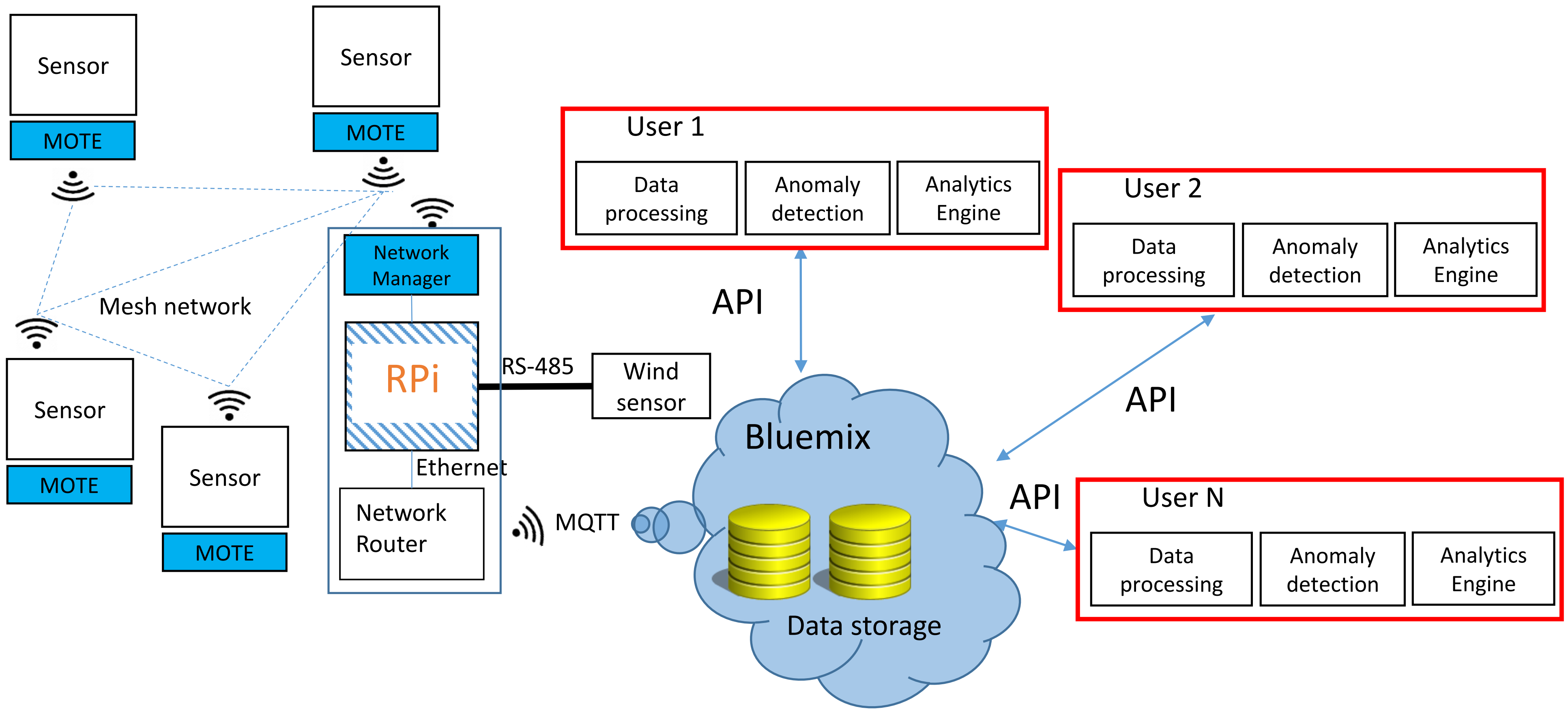
Wireless radio and sensor board

	Current(mA)	Voltage (V)	Power (mW)
Rpi-Zero	150	5	750
Figaro sensor	56	5	280
Cell modem	400	3.3	1320
GPS	20	3.3	66
Wind sensor	20	12	240
Total	646		1906



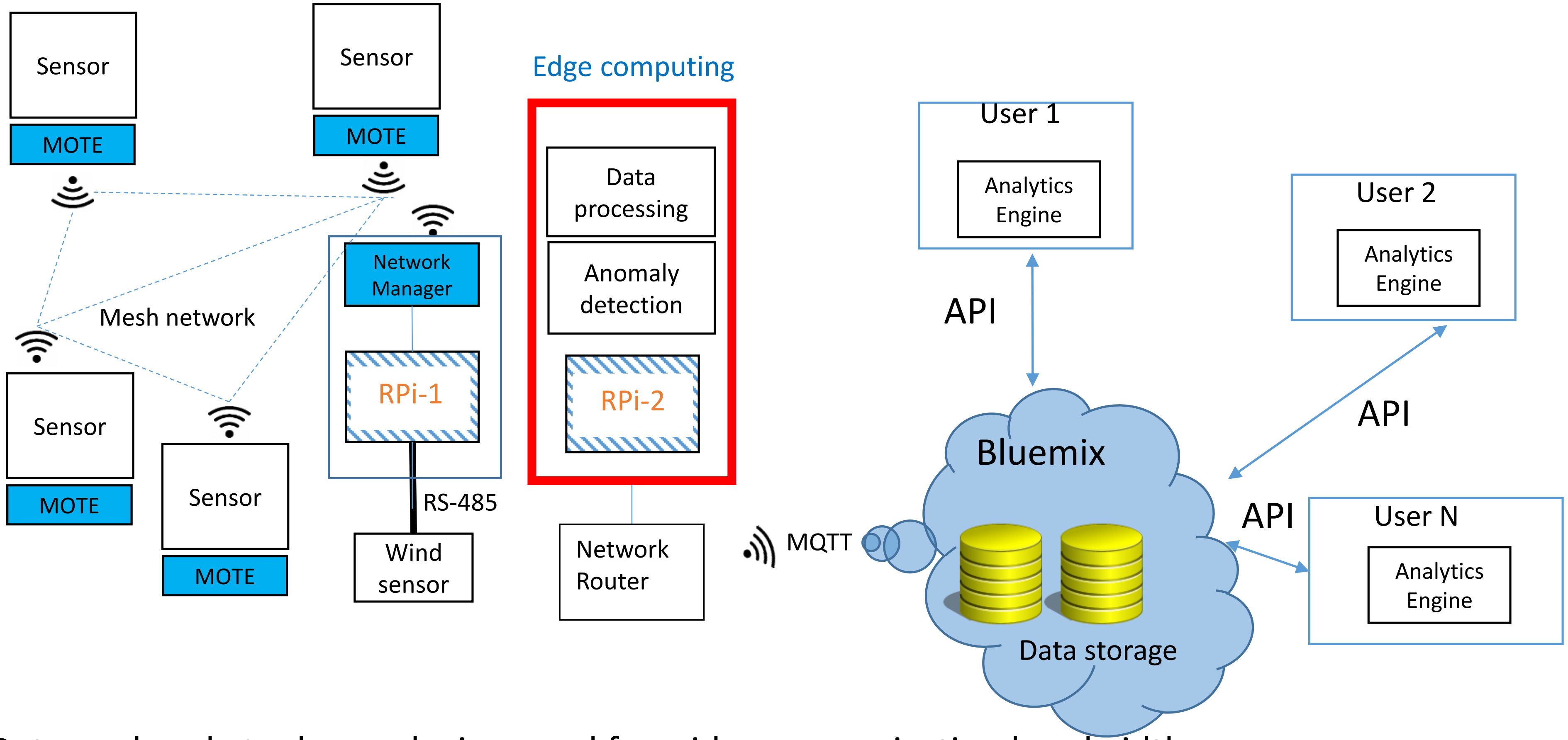
Remote gateway

Current implementation



- Require high communication bandwidth
- Similar computation carried out multiple time by different users
- Good for exploratory work but is not scalable

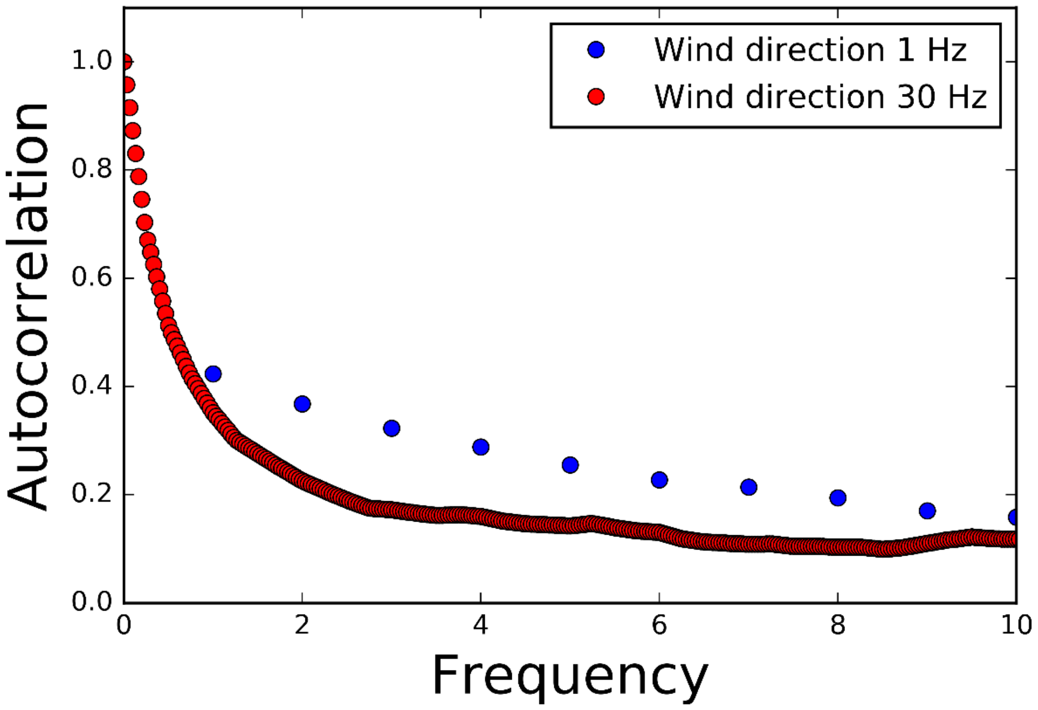
Edge computing

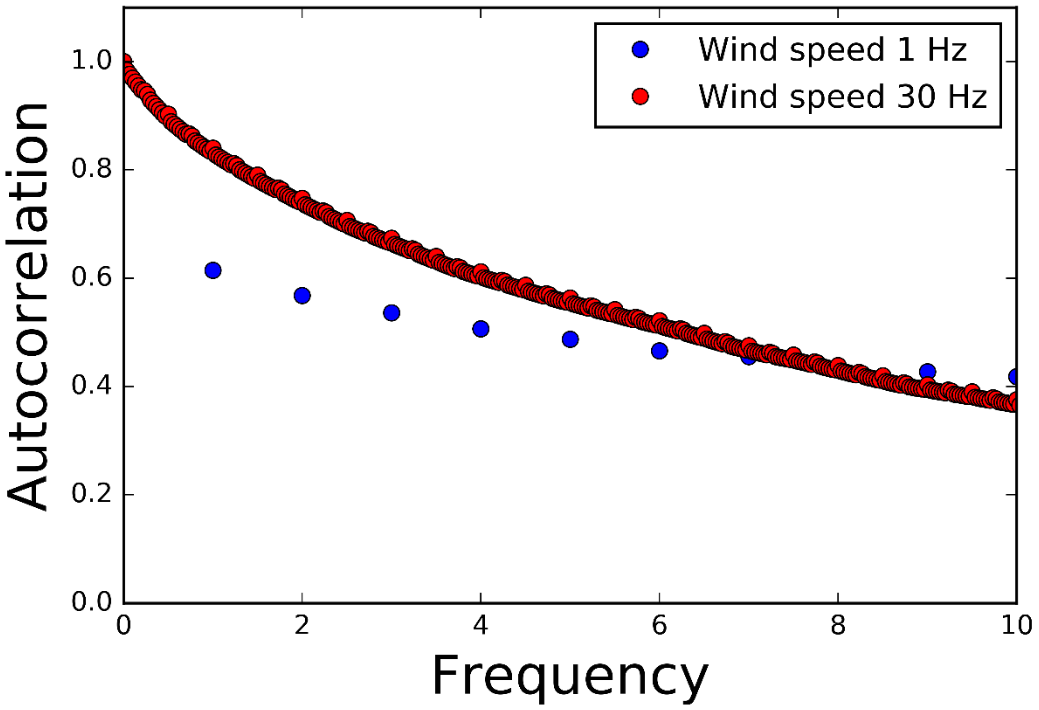


Data analyzed at edge- reducing need for wide communication bandwidth
Uniform data analysis and interpretation
Less storage required

Data acquisition rates

Data acquisition rate is dependent on the stability of the wind and local turbulance.

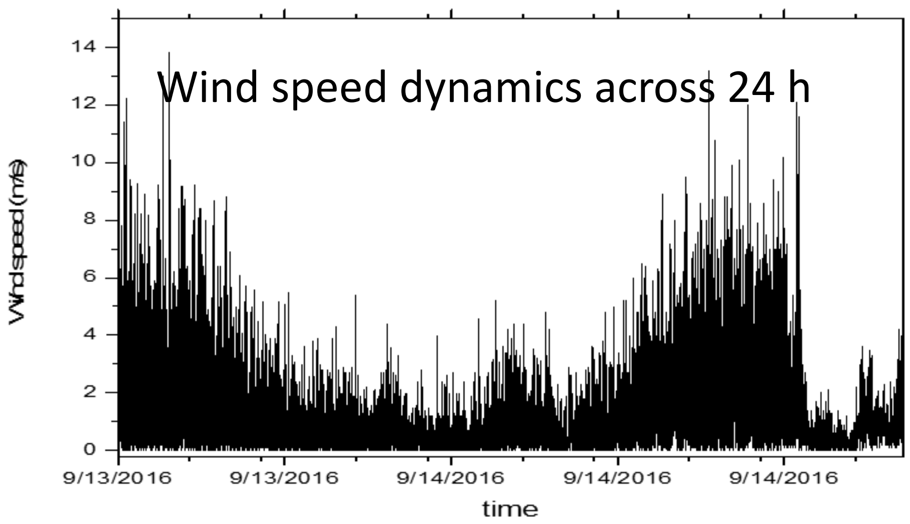




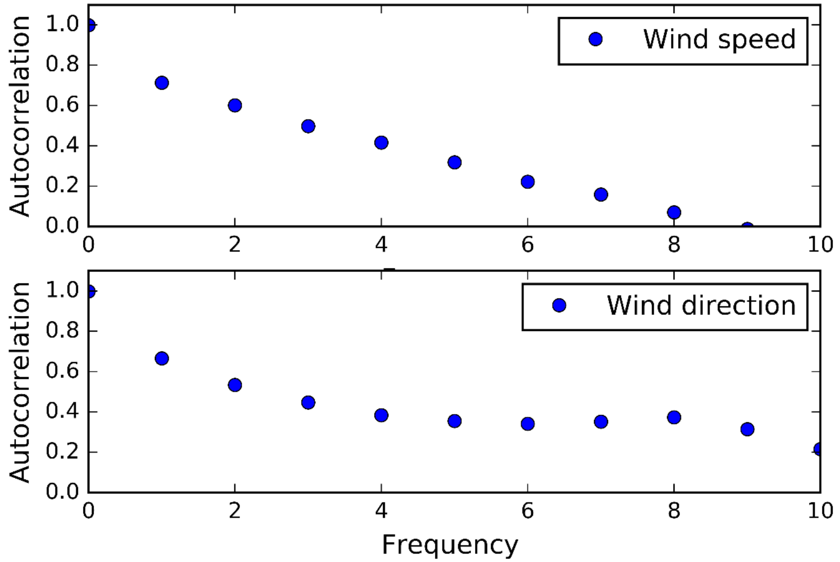
- Data sampling rate driven by autocorrelation of the wind speed/direction.
- Wind speed more stable than wind direction
- Depending on the data sampling rate-autocorrelation can be slightly different.

Spatial-temporal analytics

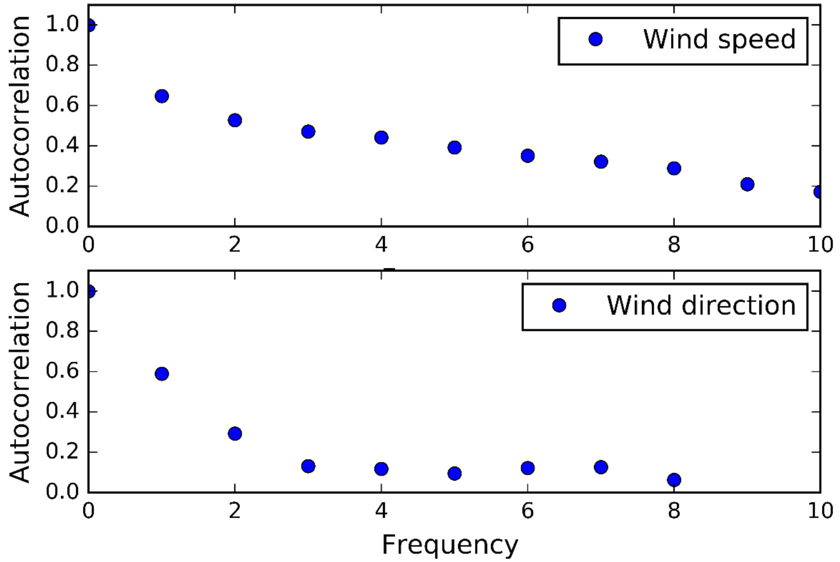
- Wind data has a strong geospatial component
  - -geographical location dependence
  - -daily and seasonal dependence
- Gas leaks may have temporal dependence
- Analytics needs to be adaptable to accommodate dynamic behavior
- The data sampling rate will need to be “cognitive”, recognize the environment and adjust the sampling rate



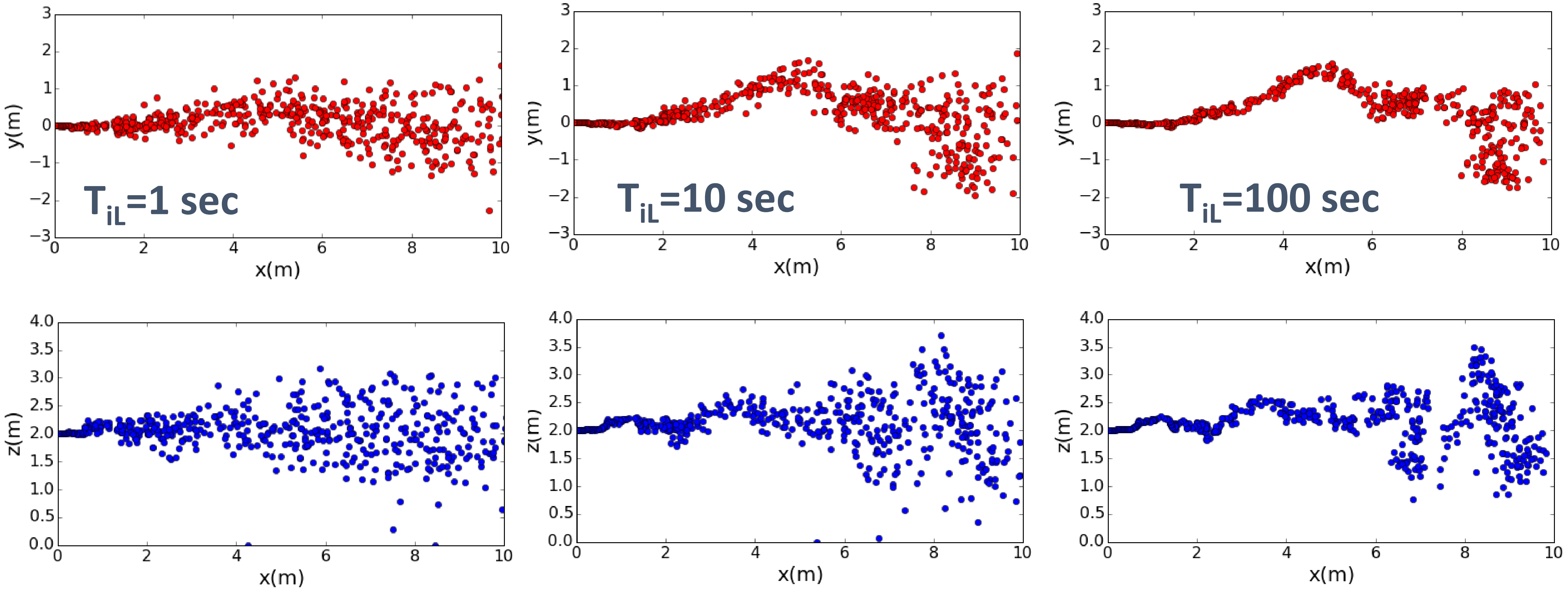
Yorktown Heights, NY July 2017



Dallas, TX July 2017



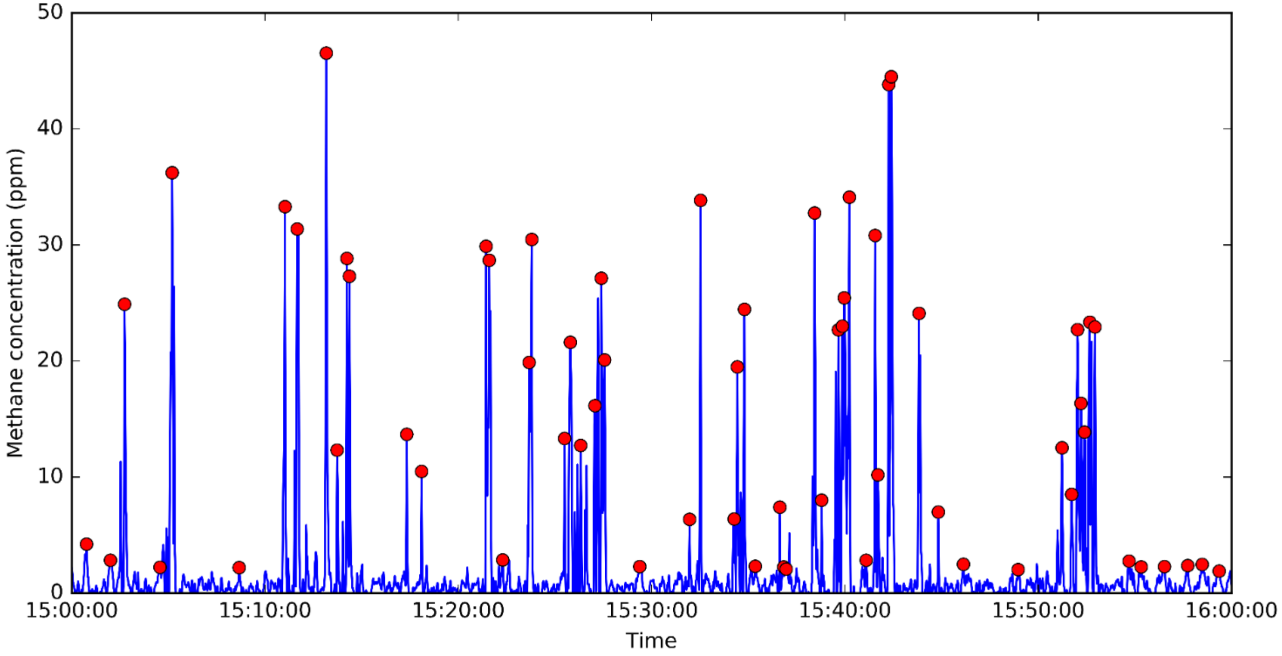
Instantaneous Plume: Lagrangian Integral Time Scale Effect



T_{IL} velocity auto-correlation time

As T_{IL} increases, instantaneous snapshot shows a more wispy plume and sensor hits more unlikely unless sensor directly downwind from leak

Peak detection algorithm



The information in chemical plume detection is carried by

1. peak height,
2. peak width and
3. timestamp

Additional information:

1. background methane level

Peak detection algorithm:

-wavelet convolution

-derivatives crossing zero

-maxima preceded by a delta

<u>Computationally efficient</u>

<u>Perform well with a given signal</u>

Implemented in Python

Run on a buffered dataset with 10 min of data acquired at 1 Hz frequency

Reduce data size by 99% compared to all data acquired

Wind data processing

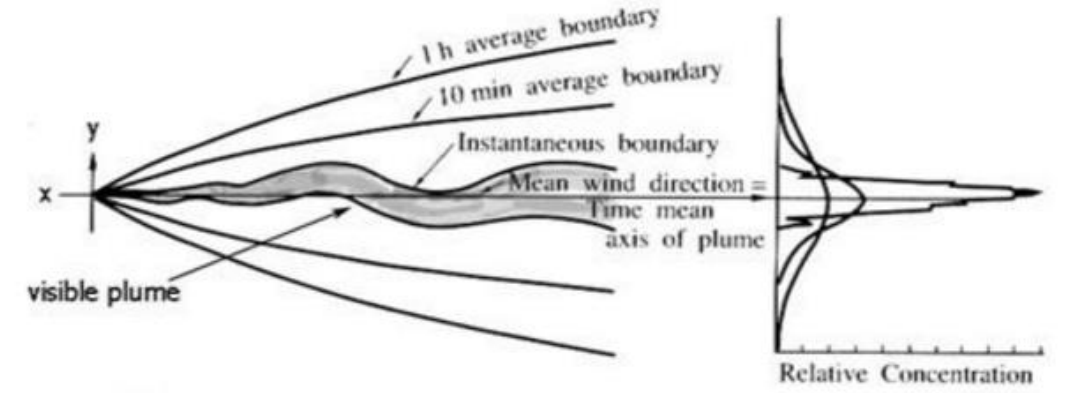
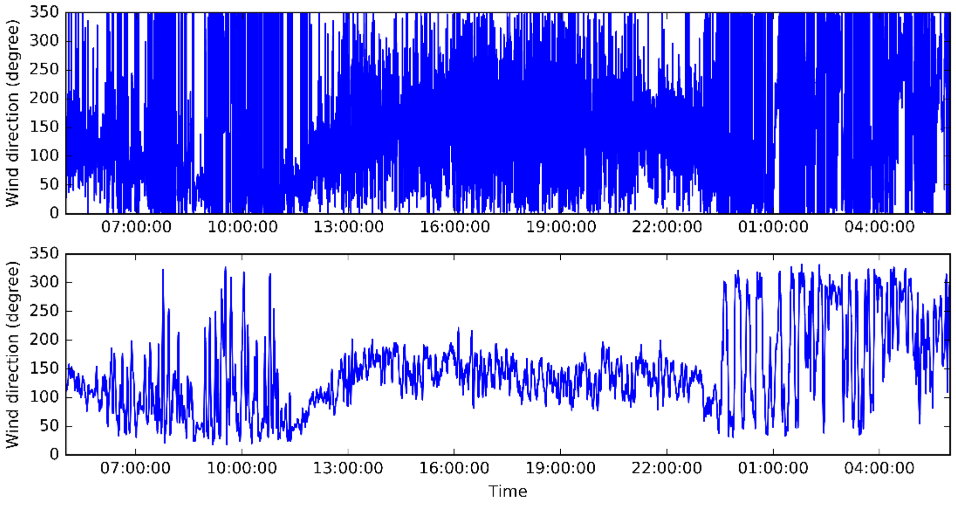
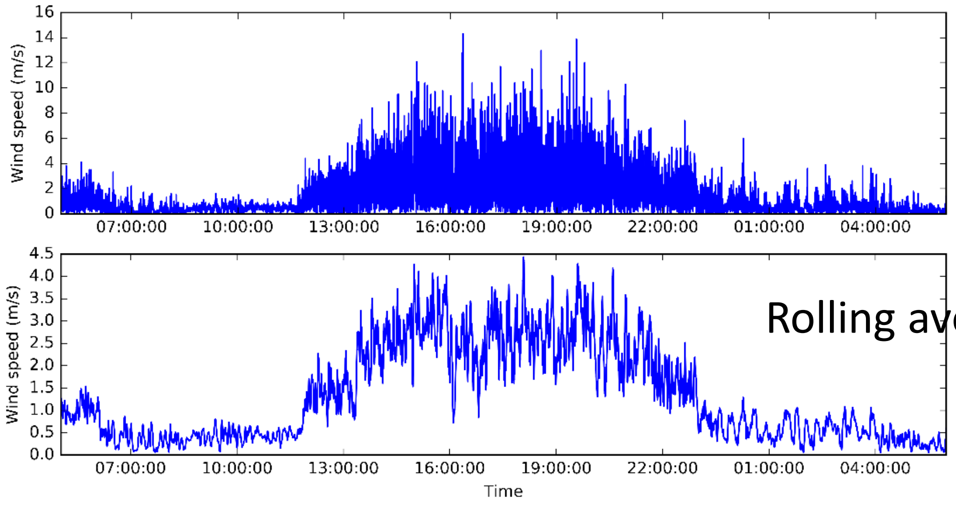


FIGURE 2.9 Plume boundaries and concentration distributions of a plume at different averaging times. (adapted from Seinfeld and Pandis)

- Instantaneous plume may be wispy and narrow leading to sporadic sensor readings
- On time averaging a Gaussian plume can be obtained
- Analytics strategy to exploit both: instantaneous and averaged plumes

Real time data is characterized by turbulence



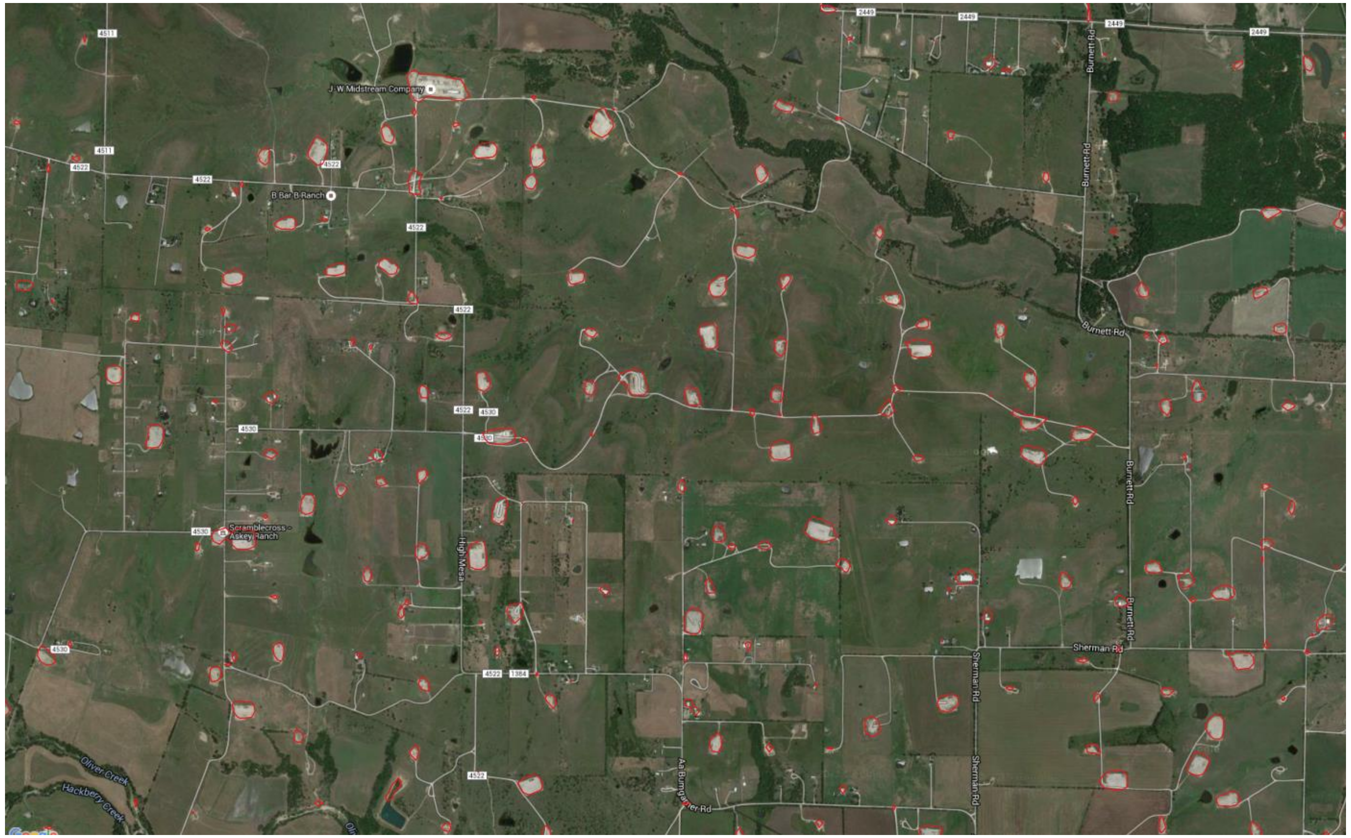


Rolling average 10 min

Classification using Multi-Mode Feature Recognition

Conventional:
- Single set of satellite images
- Empirical feature engineering.

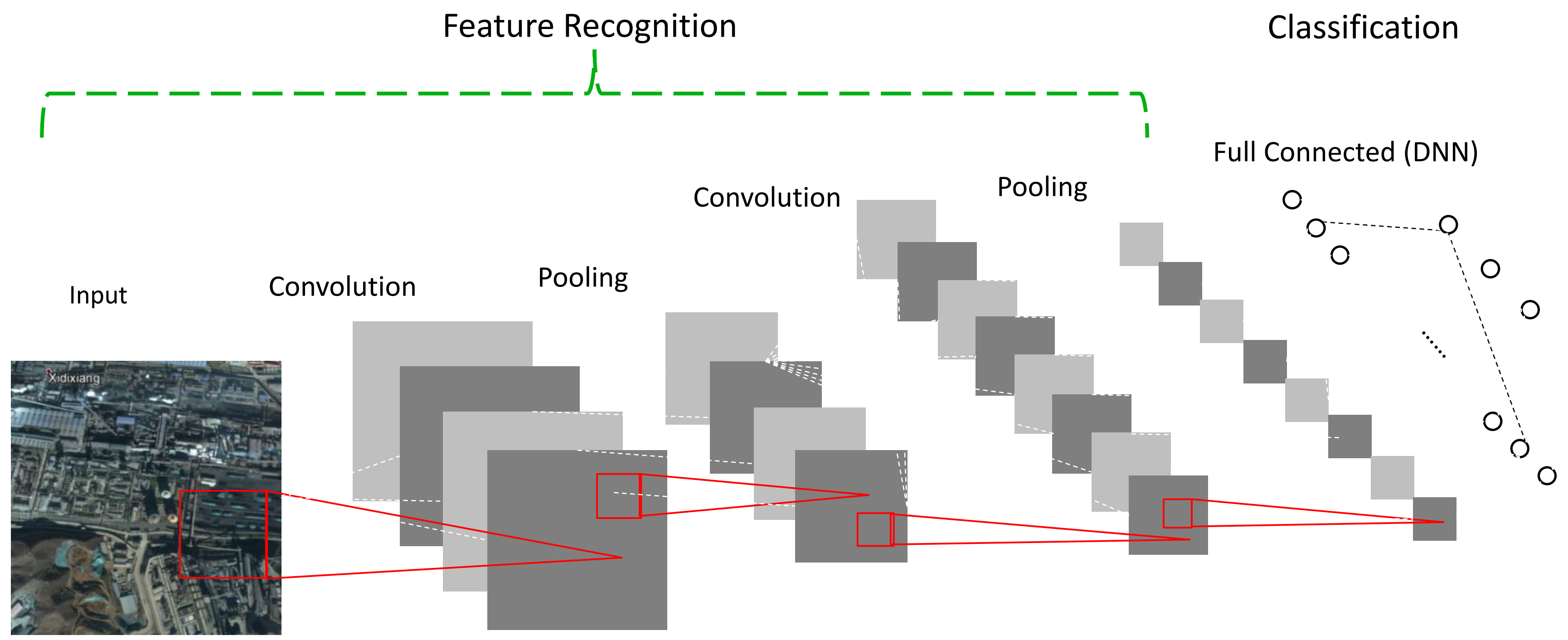
Shale gas sites found in Texas



• Multi-Mode Feature Cognition
  - Methane abs@ 1.65 µm, 2.3 µm
  - Shape, Heat, Road Connectivity …

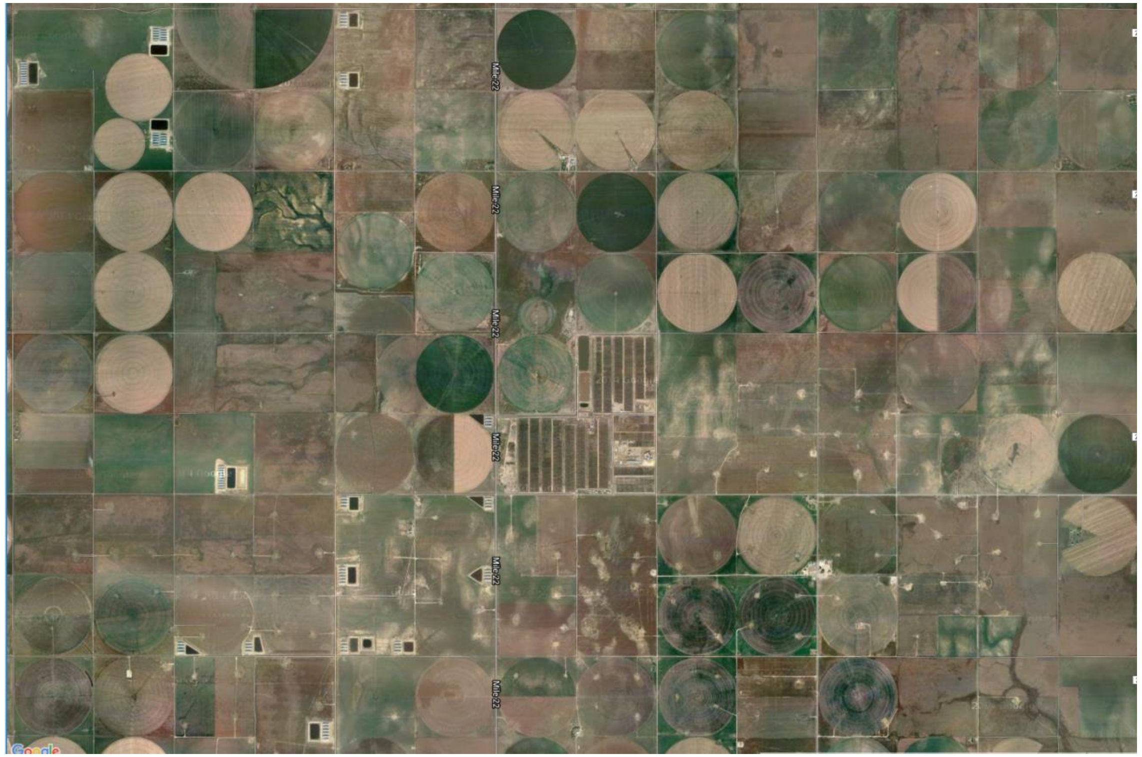
• Deep Learning to extract high-order hierarchical features

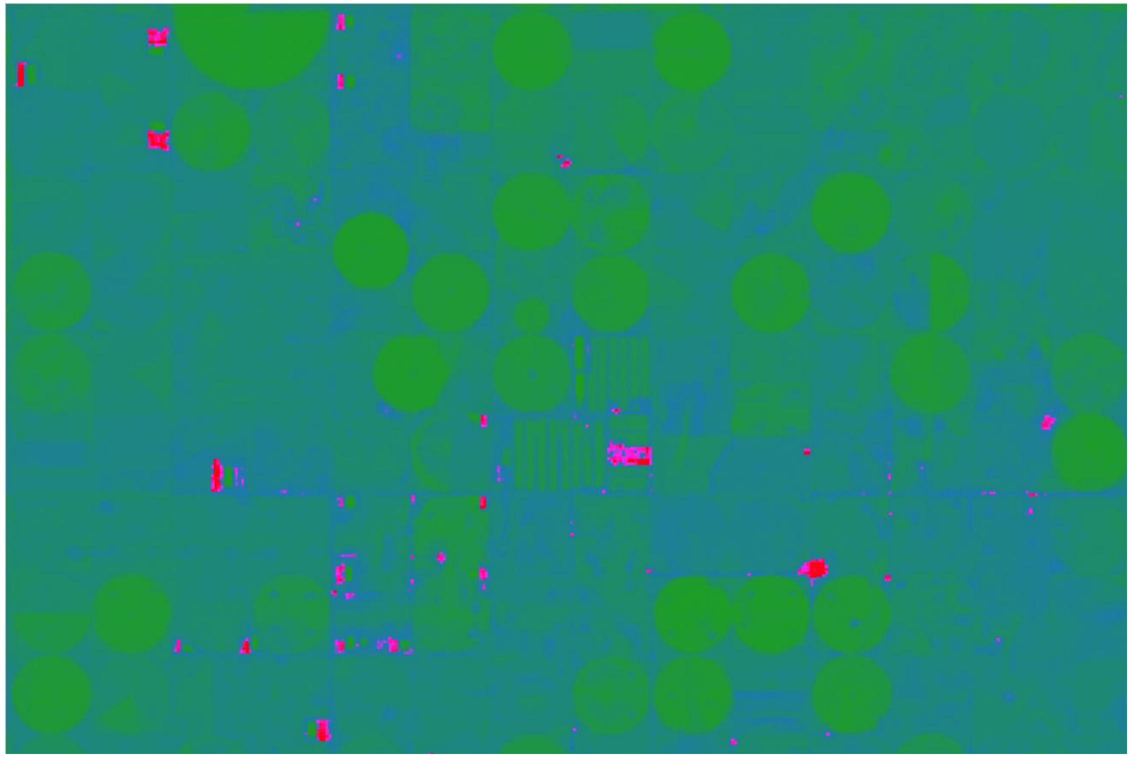
Deep Learning – Extraction of Hierarchical Features



Classification using Multi-Mode Feature Recognition

Kansas, Identification of Livestocks





Conclusions

- Data strategy is driven by industrial applications where signal integrity determines the analytics output.
- Edge computing can reduce data size by orders of magnitude making IoT solution more amenable for remote applications where data bandwidth and connectivity is an issue.
- Contextual data can enable Internet of Things applications to be automated