The Transformational Potential of Smart Cities & Society

Charlie Catlett
Senior Computer Scientist
University of Chicago and Argonne National Laboratory

Global Sustainability Summer School
Santa Fe Institute
July 2019

Collaborators:
Pete Beckman (ANL/NU)
Rajesh Sankaran (ANL)
Mike Papka (ANL/NIU)
Kathleen Cagney (UC)
Mark Potosnak (DePaul)
Doug Pancoast (SAIC)
Dan Work (Vanderbilt)
Low-Cost Sensors to Increase Spatial and Temporal Resolution
Rapidly Expanding Low-Cost Sensor Market

Particulate Matter

R² 0.65-0.80

Gasses

R² 0.85

All R² values reported in these evaluations are based either on 5-min or 1-hr average data. The Lab R² values are based on experiments conducted in the chamber, under average ambient conditions (20 degrees C and 40% RH). https://www.aqmd.gov
Chemsense (SPEC/Intel) board on AoT nodes

- Nitrogen dioxide
- Ozone
- Carbon monoxide
- Sulfur dioxide
- Hydrogen sulfide
- Respiratory irritants
- Indoor Air Quality
Array of Things air quality sensors are experimental, being evaluated using (a) EPA collocation, (b) satellite data and computational models, and (c) spatial data analytics techniques.
Need to consider and characterize different types of error

- Random error
  - Accept will be larger than reference instruments
- Systematic error
  - Zero and span
  - Trends
- Spikes
Calibration procedure

- Subtract zero current
  - Function of temperature
- Apply calibration
  - Simple span coefficient
- Remove cross-sensitivities
  - Matrix approach
- “Advanced filter”
  - If cross sensitivity removal went too far
O3 overall fit

- Spikes are lowering $r^2$, but can be removed
- During shorter comparison without NO2 spikes, $r^2$ approaches 0.80
NO2 overall fit

- Spikes are killing $r^2$
- But without spikes, there is a better correlation
Software-Defined Sensors: Observations and Adaptive Measurements
# Astonishing Advances in AI

<table>
<thead>
<tr>
<th>Sensor Input</th>
<th>Opportunity for ML inference at the edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pan-tilt-zoom Camera</td>
<td>Detection and location of fire and smoke</td>
</tr>
<tr>
<td>ETCE</td>
<td>Eddy covariance measurements, Gas Flux</td>
</tr>
<tr>
<td>Camera &amp; Mic</td>
<td>Urban wildlife and bird species detection, estimating interactions</td>
</tr>
<tr>
<td>Seismic Sensors</td>
<td>Earthquake early warning and detection</td>
</tr>
<tr>
<td>RADAR and LIDAR</td>
<td>Precipitation and aerosol rate &amp; density, storm cells</td>
</tr>
<tr>
<td>Street Camera</td>
<td>Urban gentrification, quality of life, public space perception, social cohesion &amp; neighborhoods</td>
</tr>
<tr>
<td>Street Camera &amp; Mic</td>
<td>Social interactions, crowd behavior and public space engagement</td>
</tr>
<tr>
<td>Street Camera</td>
<td>Stress, fear and violence in public spaces, cognitive impact of natural vs urban exposure</td>
</tr>
<tr>
<td>Traffic Camera</td>
<td>Urban mobility &amp; change</td>
</tr>
<tr>
<td>Microphone</td>
<td>Urban soundscape, noise impact on urban wildlife and human cognition and health</td>
</tr>
<tr>
<td>Skyfacing Camera</td>
<td>Cloud tracking and identification</td>
</tr>
</tbody>
</table>
Many Measurements Cannot be Electronically “Sensed”
Taipei Traffic Density Network (TTDN)
What does “AI-at-the-edge” Enable?

**Software-Defined Sensors**

- **Raw data**
- **Results**
- **Adaptive control**
- **Triggers**

**Autonomous Adaptive Sensing**

- Chicago Botanic Garden
- ANL/UC/NU AI Research
- ANL/UC/NU AI Research
- Lincoln Park Zoo (Urban Wildlife Initiative)
A Near Miss is Rarely Reported (so there is no data…)
Edge Vision: Cameras on Array of Things Nodes

- 8 M Pixel Camera
- Edge Processing Unit
- 5 M Pixel Camera

Field of View (FOV):
- 75 degree
- 170 degree
Software-Defined Sensors: Traffic and Pedestrian Flows
Software-Defined Sensors: Looking Up
Vehicle Brand/Model Classification

Test configuration:
- Tensorflow on Jetson TX2
- Mobilenet v2
- Stanford car dataset

<table>
<thead>
<tr>
<th>Vehicle Brand/Model</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>bmw</td>
<td>0.99747354</td>
</tr>
<tr>
<td>audi</td>
<td>0.0018637319</td>
</tr>
<tr>
<td>ford</td>
<td>0.00043634087</td>
</tr>
<tr>
<td>dodge</td>
<td>0.00014460985</td>
</tr>
<tr>
<td>chevrolet</td>
<td>8.169599e-05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vehicle Brand/Model</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>bmw</td>
<td>0.61942875</td>
</tr>
<tr>
<td>audi</td>
<td>0.33066648</td>
</tr>
<tr>
<td>chevrolet</td>
<td>0.026881762</td>
</tr>
<tr>
<td>ford</td>
<td>0.019300539</td>
</tr>
<tr>
<td>dodge</td>
<td>0.0037224325</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vehicle Brand/Model</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>dodge</td>
<td>0.40899014</td>
</tr>
<tr>
<td>chevrolet</td>
<td>0.3156025</td>
</tr>
<tr>
<td>ford</td>
<td>0.23515542</td>
</tr>
<tr>
<td>audi</td>
<td>0.022362422</td>
</tr>
<tr>
<td>bmw</td>
<td>0.01788947</td>
</tr>
</tbody>
</table>
Transfer Learning and Data Set Training

- High definition image collection for training ML models
- Currently deploying two 4k camera-equipped Waggle nodes at the main gate of ANL
  - Collect time synched front image of cars
Pedestrian Tracker on TX2
- Deep Sort Algorithm
  - Faster RCNN for pedestrian detector
  - SORT (Simple Online and Realtime Tracking) for tracking
    - Kalman filter and Hungarian algorithm
- Model for detection: mars-small128
  - Quantized model (TensorFlow Slim)
- Training dataset: MARS (Motion Analysis and Re-identification Set) dataset
- Average processing time: 45.7 ms / frame
- Average detection time: 1.15 ms / frame
Cloud Cover Estimation and Prediction

Cloud coverage calculation script in Waggle
• Color channel based pixel distinction: sky vs. Cloud
• Over-estimate cloud coverage when color of sky is bright
AoT in Many Cities: Analyzing Diverse Urban Spaces

AoT partners in other cities are instrumenting common “venues,” (e.g., public squares, complex intersections) to enable comparative studies across diverse cultures, climates, sizes, densities, and maturity.
Thinking Further

• How might information about pollutant sources (air, noise, light) be used to better select locations for measurement?

• What new “urban activity” measures could be enabled using computer vision or hearing?

• How might an urban measurement system (like AoT) “learn” about its environment in order to detect anomalies? Are there specific aspects of “normal” (such as clothing type, movement patterns) that could serve as training wheels?
Exploring Solution (and Risk) Space with Coupled Computational Models
Imagine a City...

...that consumes 50% less total energy per person while improving economic vitality and quality of life and increasing resilience and sustainability.
Holistic Urban Modeling: Coupled Multiscale Models

- District-scale performance evaluation across energy, environment, transportation, and other facets of the urban environment.
- Characterization and optimization of building and district performance over decades.
  - Ensembles and deep learning for parameter optimization.
Theorem:
The Planner’s Problem is intractable
= cannot be solved computationally, however big one’s data!

Let a set of urban metrics $Y \in$ be functions of the social networks in the city $G$, $Y_i(G)$. Let $K$ be the total links in $G$.
The Planner considers all possible plans, which are all possible social networks $G$.
The best plan maximizes an appropriate function of the $Y_i$.

There are

$$P(K_p, K) = \frac{K_p!}{K!(K_p - K)!} = \frac{\ln K_p! - \ln K! - \ln(K_p - K)!}{\ln N}$$

plans, where $K_p$ is the number of possible social ties

$K \sim N^{1+\delta}, K_p \sim N^2$

P(N) = P_0 e^{N^{1+\delta} \ln N}$

City

P(N) = P_0 e^{1.4 \times 10^7} = P_0 10^{6 \times 10^6}$

Number of stars in the Universe: $10^{82}$

Shannon Order: All possible Chess games:

So… what is possible with today’s computing systems? In five years?
Coupled Urban Models

• Urban design at the district scale (50-100k people, 10-20k buildings, 2-20km²) is critical both in developing economies (where entire cities are being designed and built) and in the U.S. (where cities are expanding and repurposing at district scales)
  • How will a given district design perform with respect to energy? Transportation? Emissions? Economics?
    • In context of climate, economic, demographic, and other uncertainties?
    • Where factors and effects unfold over vastly different timescales, from seconds to years.

• Creating a framework in which multiple models can be coupled to provide evaluations of various scenarios to enable designers to quantify the performance envelopes of a given design.

• Rather than a handful of “hero” simulations, the need is for hundreds (thousands) of simulations to examine performance properties across many parameters and conditions.
Multiscale Coupled Urban Systems

Vehicle emissions, heat

Weather

Building emissions, heat

Vehicle mix, driving habits

Response times

Wind, pressure, temperature, moisture, radiation

Building Demand

Building Mix, Pricing

Environment & Infrastructure

Longitudinal Measurements

Population and Economics

Municipal Data Sources

Sensor Networks

Census, Social Sources, Mobility…

• TRANSIMS

• NEK5000
  • WRF

• CommuterSIM, ChiSIM

• EnergyPlus

NEK5000 (P. Fischer, M. Min, A. Obabko)
Test Area

- 3 km$^3$
  (760 acres)
- 2,000 buildings
- 16 km riverfront
- 20 km$^3$
- 20k buildings
Challenge Problem: District-Scale Urban Design Evaluation

City of Chicago building shapefiles with North Branch Area highlighted.

FY18 test region (20-buildings)

Current Array of Things measurement nodes (more to come in late 2018) shown. ANL-Exelon CRADA also provides 15min resolution energy usage for two years at zip+4 spatial groupings.
Coupled Microclimate-Building Models

Boundary Conditions:
- Outside surface temperature

Building data
- Meteorological data
- Systems: lighting, HVAC, etc.

Building zones:
- Thermal load

Air and heat flow

Urban atmosphere:
- Outdoor air condition

Mass flow

Data samplers

EnergPlus
- Node

Nek5000
- Mesh

Rajeev Jain (ANL), Xuan Luo (LBNL), Aleks Obabko, Gökhan Sever (ANL), Martin Berzins (Utah)


Predict average and peak energy usage over 10 yrs

For each of 4 possible redevelopment plans (transit, building, and green infrastructure designs):

- predict energy usage and demand for 20,000 buildings over 10 years for 4 climate change scenarios (from E3SM).

Swift/T-based hub

On today’s Petascale machines each scenario takes tens of minutes.
Identifying “Vulnerable” Buildings

• How will 1000s of buildings perform during more frequent extreme weather events in 2040?
• What retrofit strategies will address these performance issues?
• How can cities modify building codes and create incentives programs to “shape” the city (building heights, orientation, spacing, green spaces, etc., e.g. to optimize for more frequent extreme weather events?
• What energy demand load profiles will they present to distribution grids?
New district/campus development with zero-net energy (ZNE) or zero-net carbon (ZNC) goal

For each of \( n \) candidate redevelopment plans:
1. Predict energy demand over \( m \) years for \( p \) IPCC climate change scenarios
2. Predict peak electricity demand under \( q \) scenarios of heat waves
3. Optimize the balance of energy efficiency and on-site renewable energy generation towards the ZNE/ZNC goal, e.g., evaluating energy infrastructure options
4. Evaluate impact of the \( n \) building mix/diversity and density on energy demand
5. Evaluate impact of urban morphology on energy demand
6. Feasibility study of \( q \) options for district energy systems considering loads diversity and density
7. Quantify interdependency between buildings and urban microclimate to guide UHI policy and mitigation
8. Evaluate impacts of EV and electrification on energy demand
9. Risk assessment and mitigation of indoor environment under extreme weather events (e.g., heat waves, cold waves, wildfire-generated plumes)
10. Evaluate energy resilience considering deployment of renewable energy, energy storage, EV, and advanced control strategies.
11. Evaluate changing future demographics and their transportation choices.
12. Evaluate transportation infrastructure changes such as adding dedicated transit corridors.

How would you evaluate these 2500 permutations?
HPC and AI to Optimize Urban Design

Plan

Data

Forecast

Evaluate

Optimize
Vision (FY20) – Machine Learning for Optimization

EMEWS framework
Thinking Further

• What sort of parameterization might make sense for a district in order to facilitate approaches such as parameter optimization?

• Can we model the economic evolution of a district, to the degree that we might bound the risk of a district design being economically sustainable?

• Given a neighborhood area or district, is there a minimal set of economic assets required for sustainability? E.g., “if we have at least 7 of the following 12 key businesses operating within the area then we are probably OK?”
Other Topics
Education and Youth

• Eight week curriculum has trained over 600 students in four years.
• 5-day professional development workshop tested in July 2018.
• “School of Things” in multiple Chicago in (2018-2019).

STEM+C Education

Image: Horsch Gallery

Images: Jose Osorio, Chicago Tribune
AoT partners with local research institutions who have formed teams with their cities to identify one or more policy or science questions for which an AoT installation makes sense. Partner installations are typically 4-10 nodes. These nodes are “turnkey” with management and data services provided by the University of Chicago.