Analysis and Visualization of Urban Data

Claudio T. Silva
Tandon School of Engineering
Center for Data Science
Center for Urban Science + Progress
Courant Institute for Mathematical Sciences
New York University

Joint work with Juliana Freire, Huy Vo, Harish Doraiswamy, Carlos Dietrich, Fernando Chirigati, Theo Damoulas, Nivan Ferreira, Masayo Otta, Kien Pham, Jorge Poco, Luciano Barbosa, Marcos Vieira, Marcos Lage, Joao Comba, Luis Gustavo Nonato, Luc Wilson, Heidi Werner, Muchan Park, Jonathas Costa, and many others

Funded by grants/gifts from NSF, NASA, DOE, Moore and Sloan Foundations, MLB.com, DARPA, AT&T, NVIDIA, and IBM

NYU TANDON SCHOOL OF ENGINEERING

VIDA VISUALIZATION IMAGING AND DATA ANALYSIS CENTER

Why do visualization?

- Pictures help us think
  - substitute perception for cognition
  - free up "working memory"

[Munzner, "Visualization Chapter" in Shirley and Marshner's textbook]

Set A		Set B		Set C		Set D	
X	Y	X	Y	X	Y	X	Y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.11	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Summary Statistics

$u_X = 9.0 \sigma_X = 3.317$

$u_Y = 7.5 \sigma_Y = 2.03$

Linear Regression

$Y^2 = 3 + 0.5\ X$

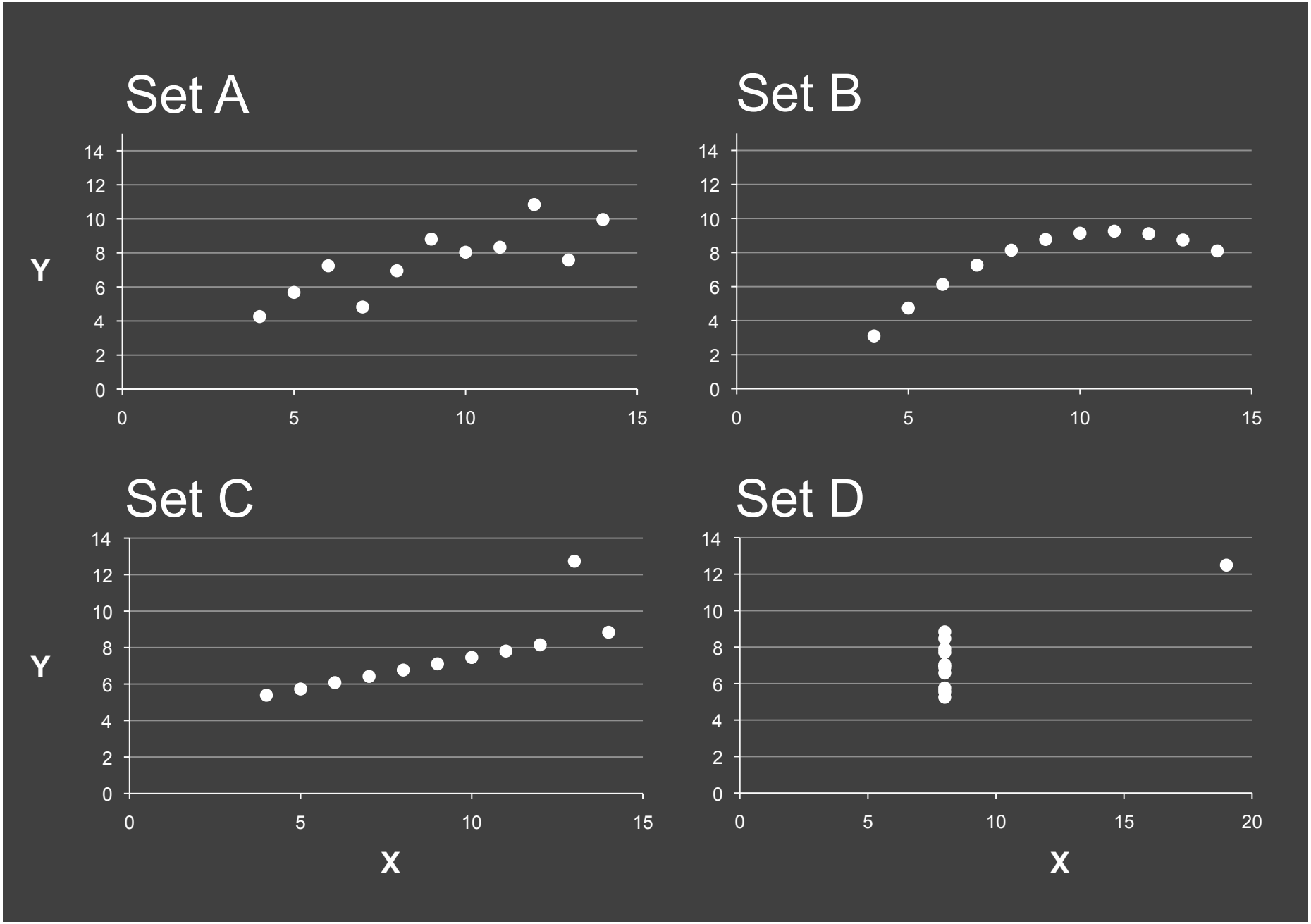
$R^2 = 0.67$

Anscombe 1973

a brief history of

Data Visualization

Jeffrey Heer Stanford University



a brief history of

Data Visualization

Jeffrey Heer Stanford University

Visual Analytics

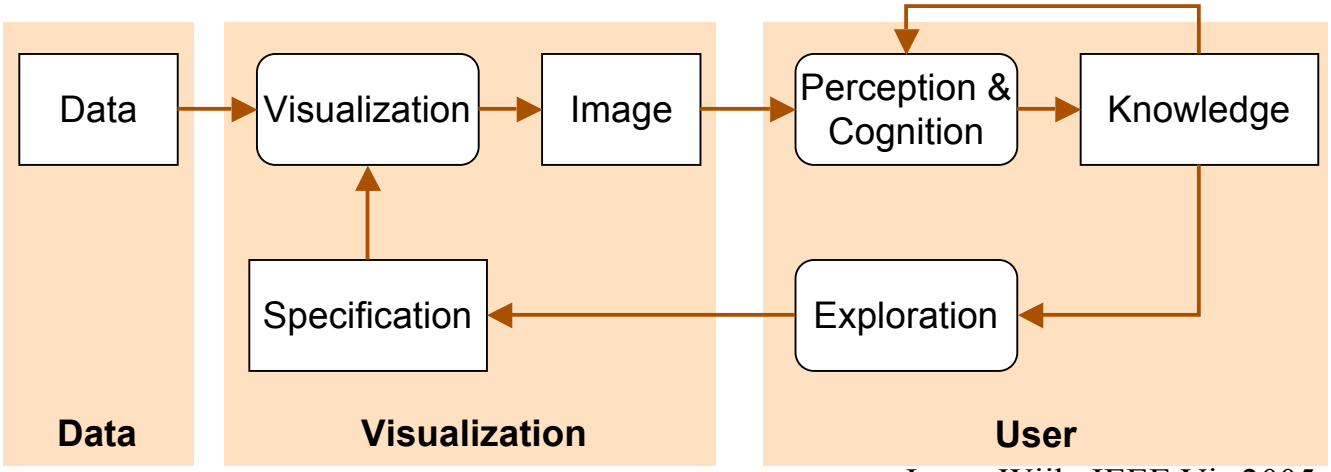
"the science of analytical reasoning facilitated by visual interactive interfaces." - [Thomas and Cook, 2005]

Interdisciplinary:

- Visualization (both InfoVis and SciVis)
- Interactive data analysis
- Data management
- Analytic reasoning
  - Machine learning, cognitive science, etc.

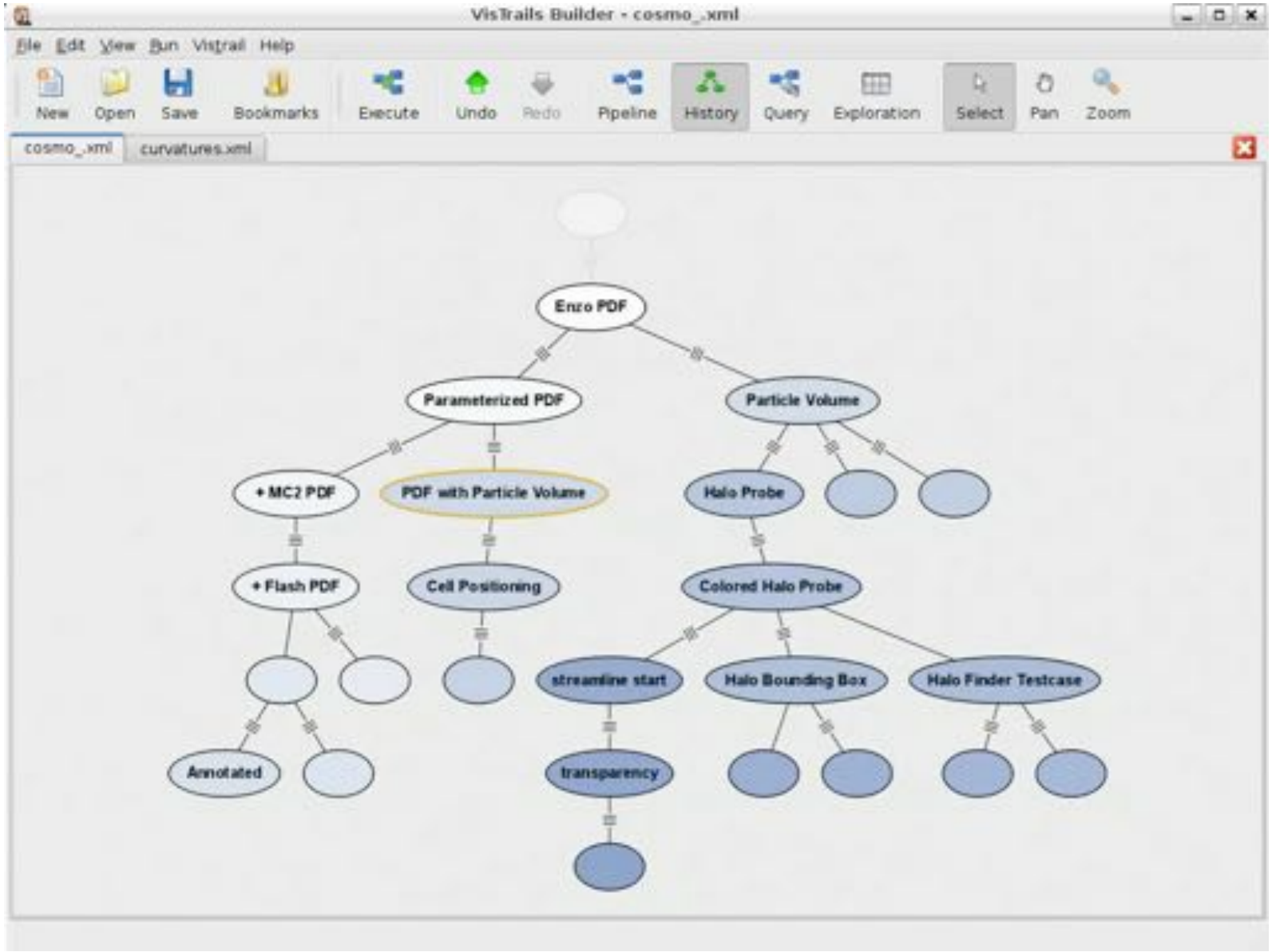
Data Exploration through Visualization

- Insightful visualizations help analyze and validate various hypothesis
- But creating a visualization is a complex, iterative process



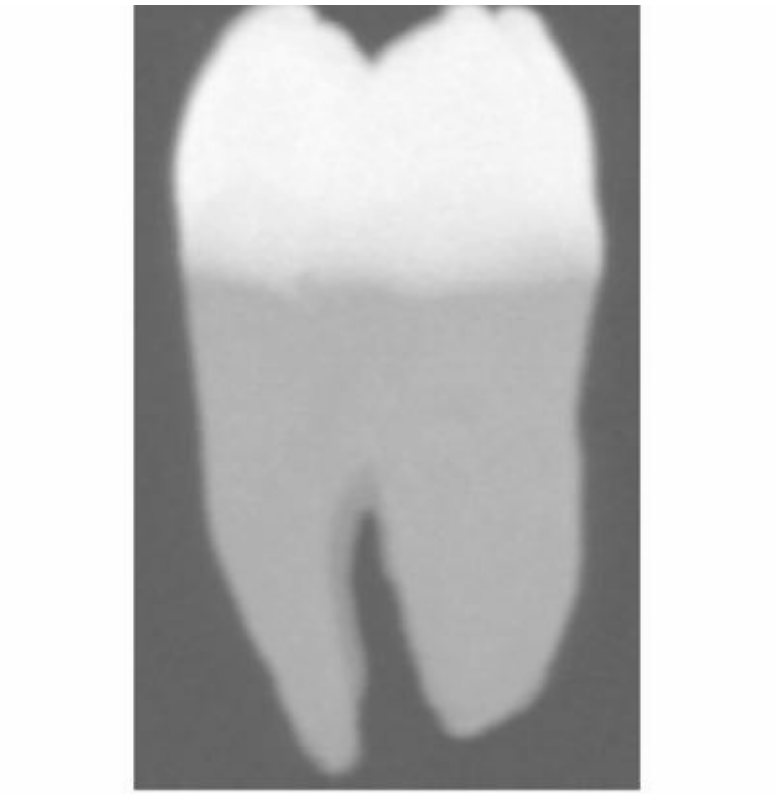
J. van Wijk, IEEE Vis 2005

VisTrails Freire et al, 2006



Anderson et al, 2008

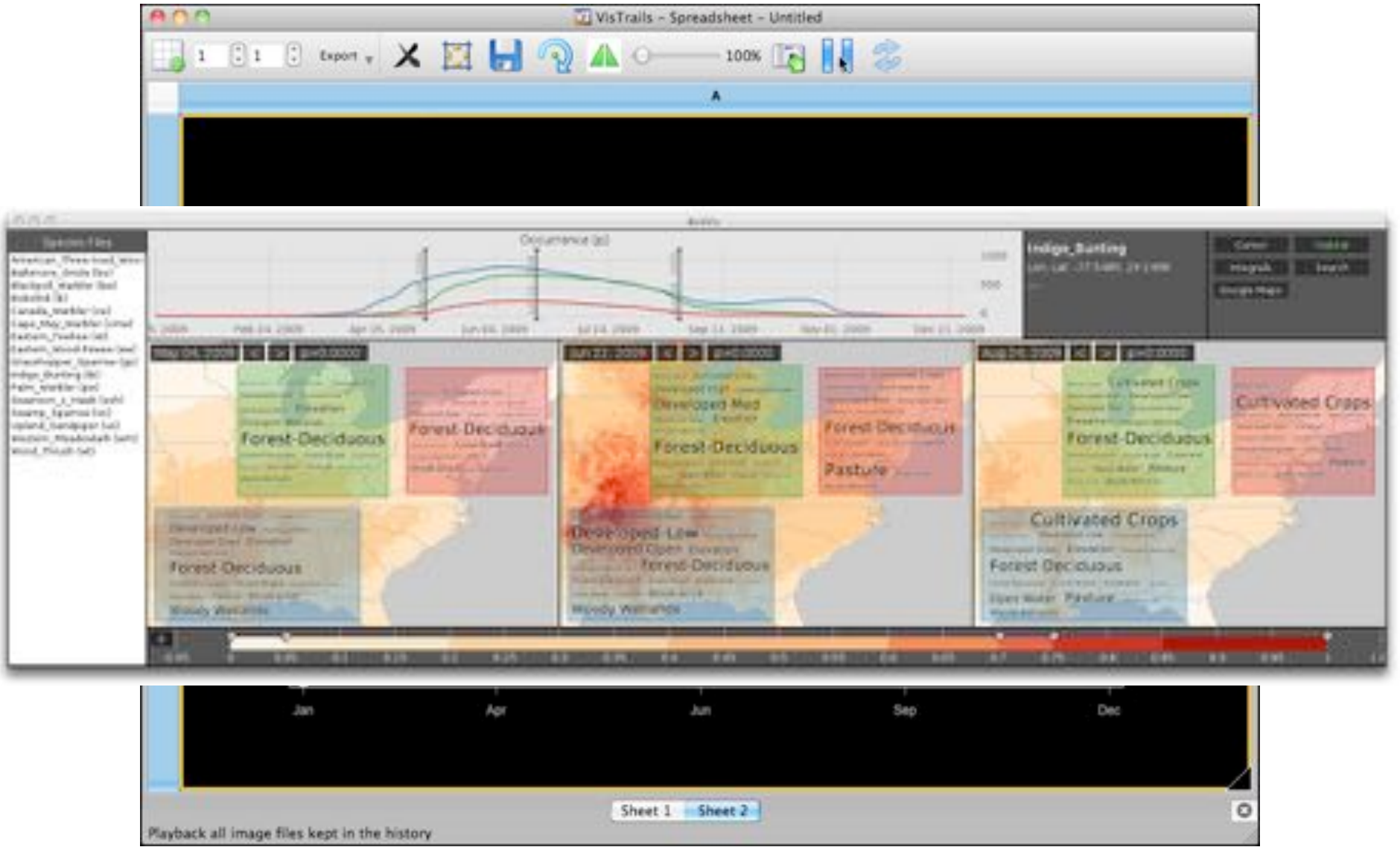
Visualization is a quickly evolving field!



Maximum Intensity Projection (MIP)

Full Volume Rendering

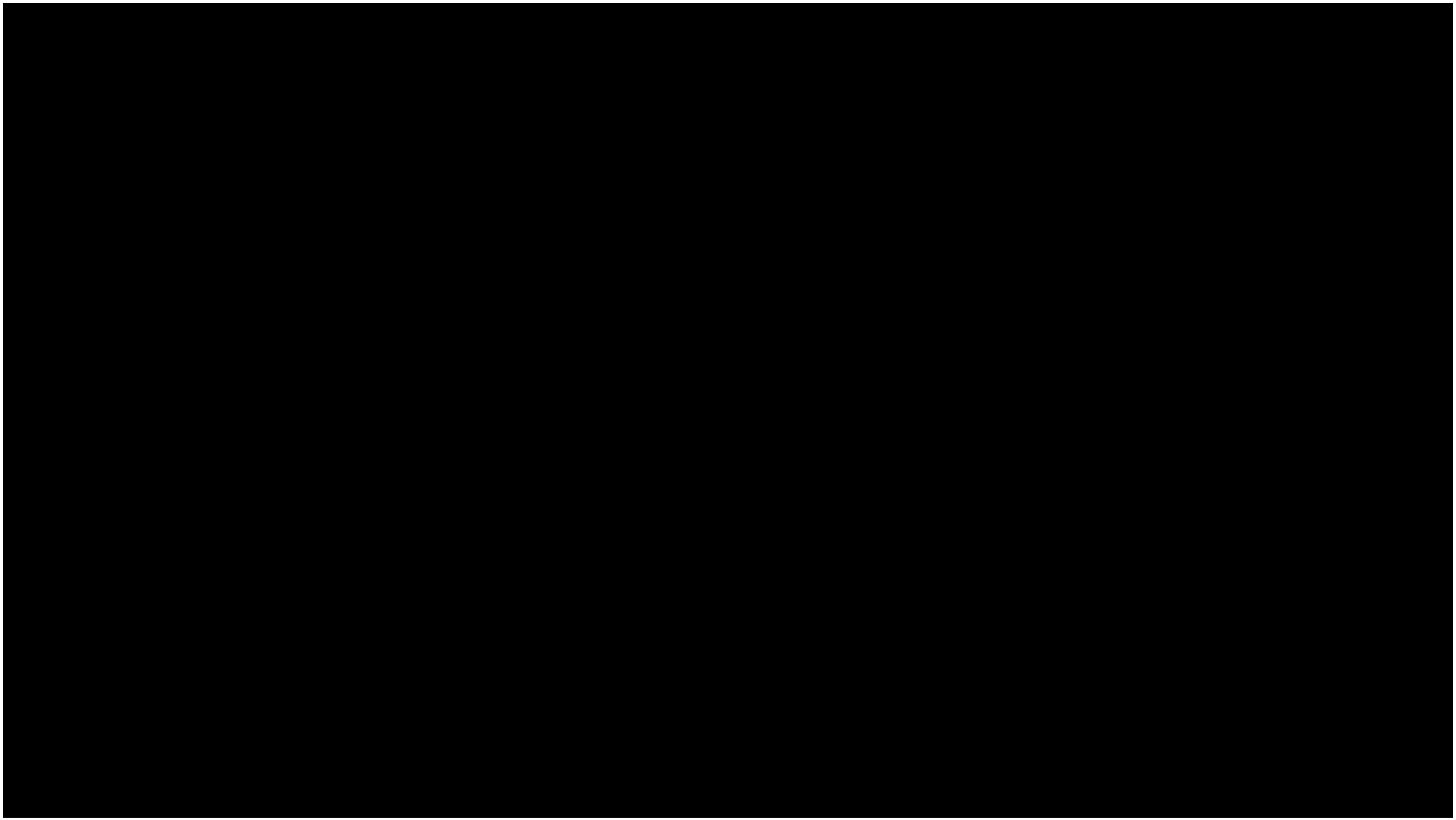
BirdVIS



[D. Fink et al., Cornell Lab of Ornithology; N. Ferreira et al., U. of Utah]

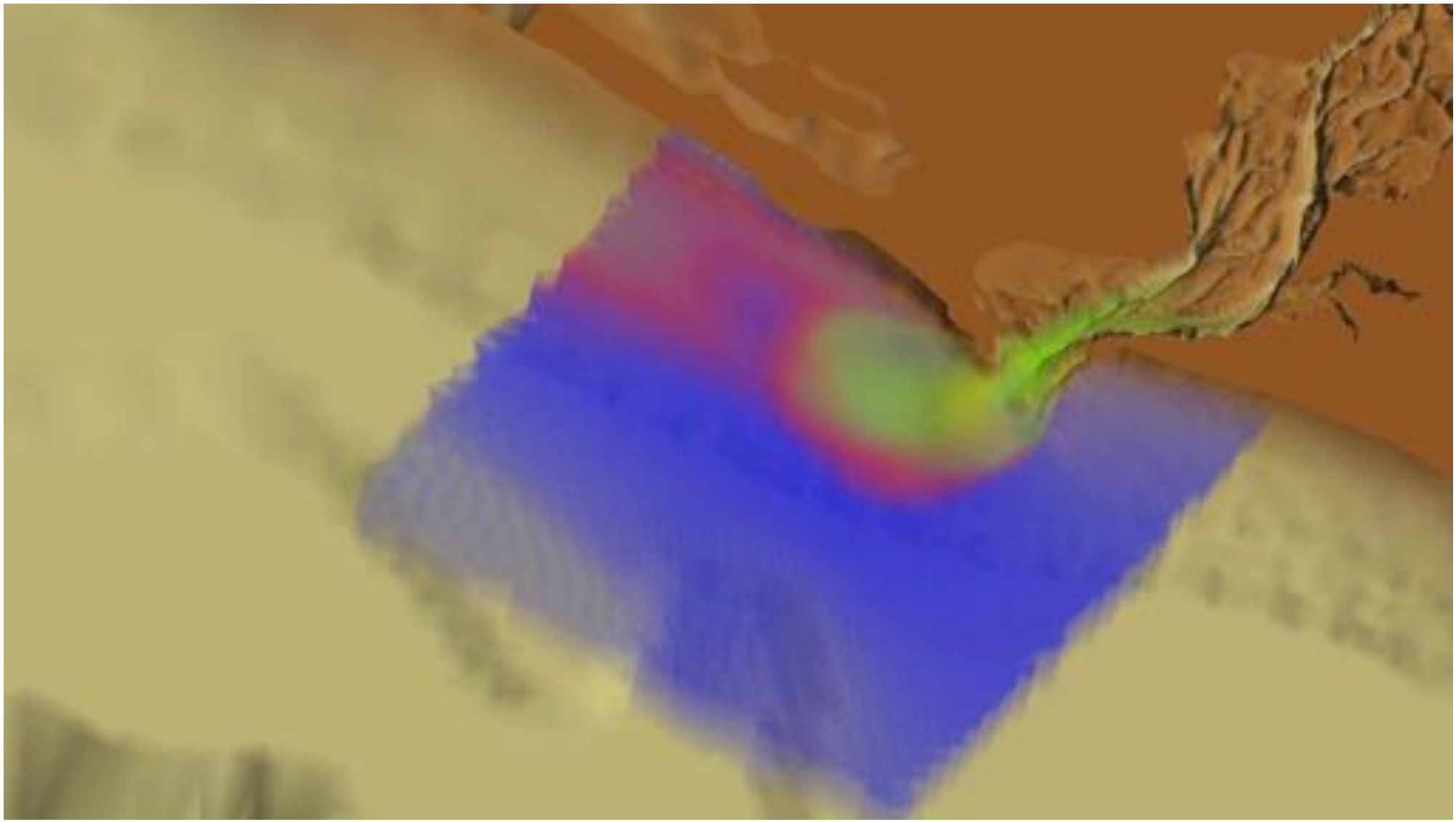
BirdVis

Ferreira et al, 2011

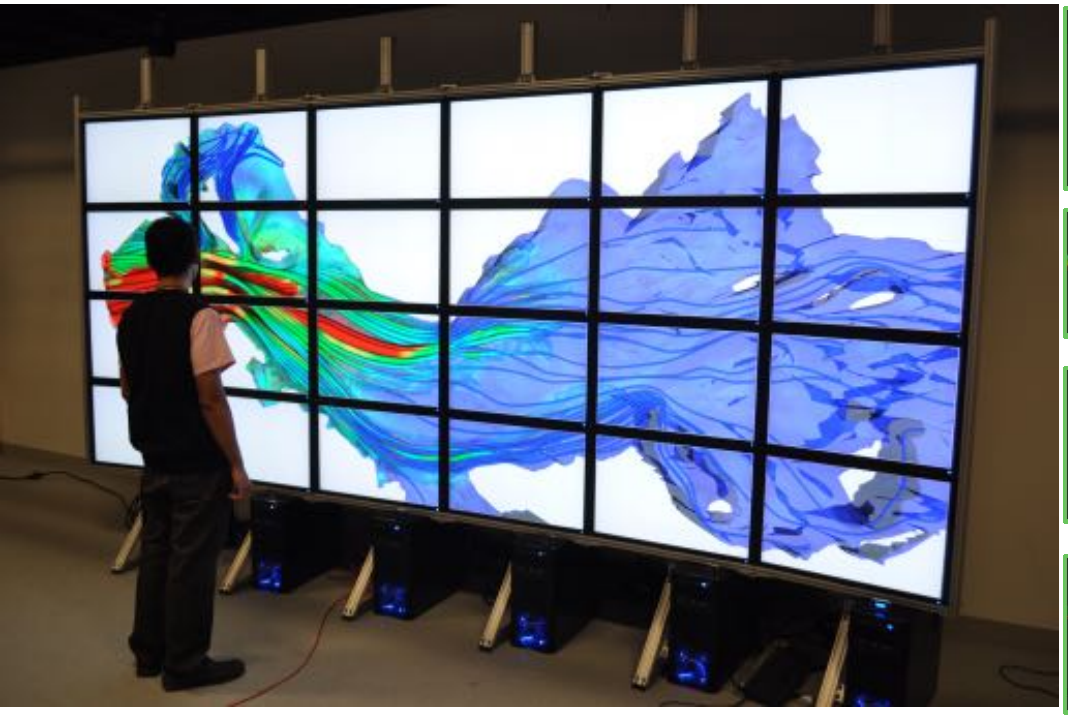


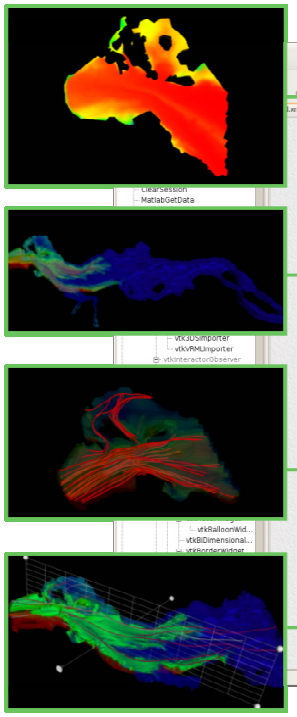
Freshwater Plume

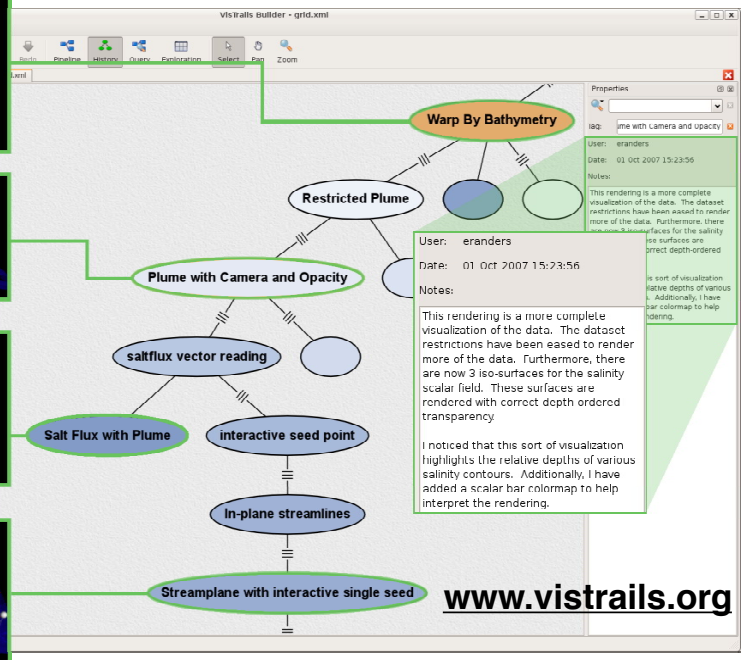
Jimenez et al, 2003

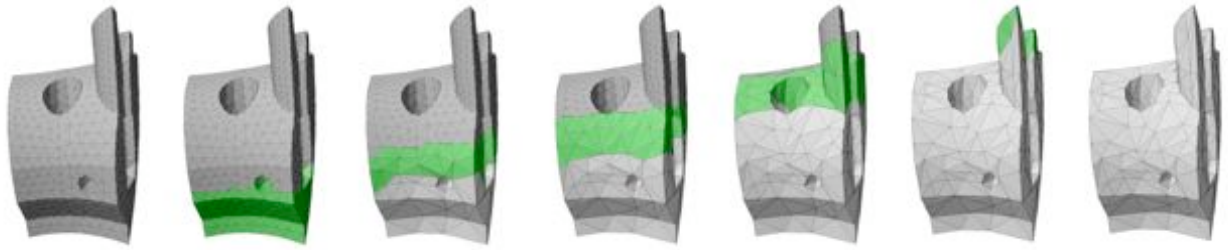


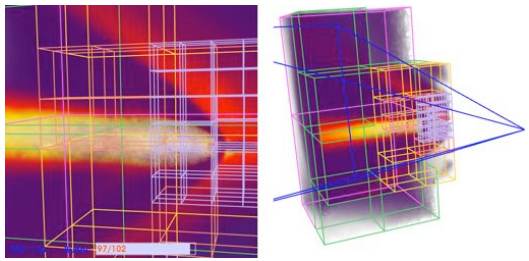
3D detail of the freshwater plume during the ebb tide, showing a sharp density front (right side of the image).



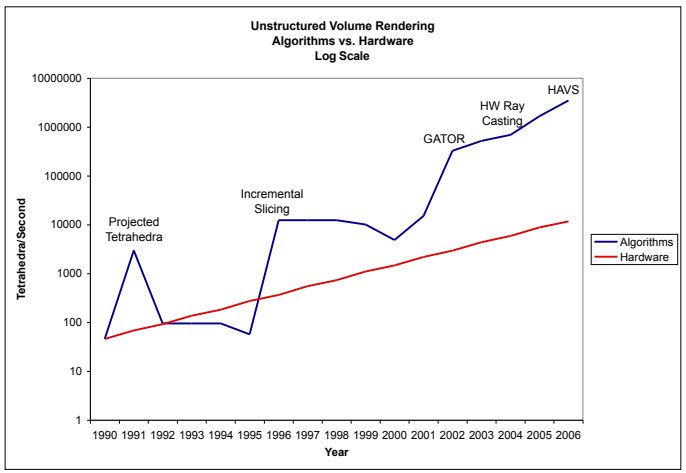


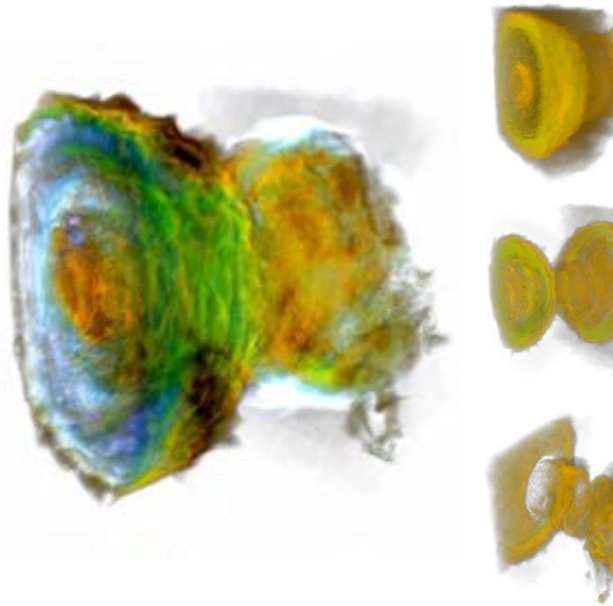

vistrails Builder - grid.xml
Warp By Bathymetry
Restricted Plume
Plume with Camera and Opacity
saltflux vector reading
Salt Flux with Plume
interactive seed point
In-plane streamlines
Streamplane with interactive single seed
User: eranders
Date: 01 Oct 2007 15:23:56
Notes:
This rendering is a more complete visualization of the data. The dataset restrictions have been eased to render more of the data. Furthermore, there are now 3 iso-surfaces for the salinity scalar field. These surfaces are rendered with correct depth ordered transparency.
I noticed that this sort of visualization highlights the relative depths of various salinity contours. Additionally, I have added a scalar bar colormap to help interpret the rendering.
www.vistrails.org






SDK 10 Update: New Stencil Routed K-Buffer Sample
DEVELOPER ZONE
SDK 10 Update: New Stencil Routed K-Buffer Sample
Could Your Game Run 35% Faster?
NVPerfHUD 4


Unstructured Volume Rendering
Algorithms vs. Hardware
Log Scale
Tetrahedra/Second
Year
Projected Tetrahedra
Incremental Slicing
GATOR
HW Ray Casting
HAVS
Algorithms
Hardware



Urban Data Analysis

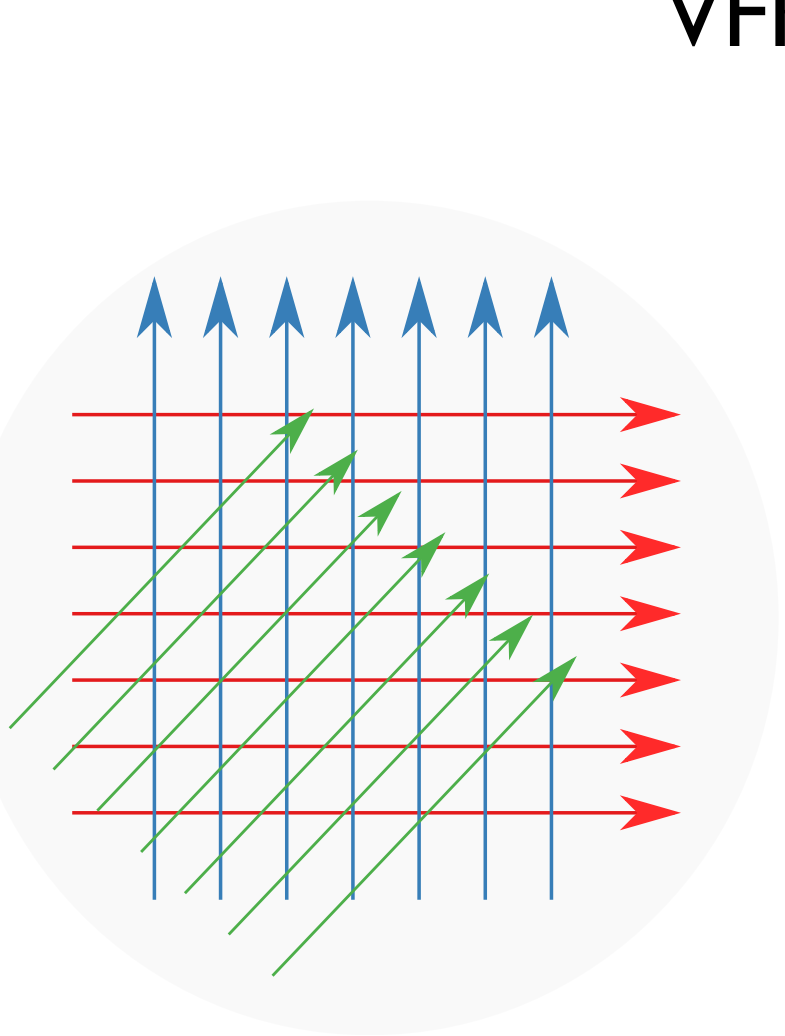
Joint work with **Juliana Freire, Huy Vo, Harish Doraiswamy,**
Fernando Chirigati, Theo Damoulas, Nivan Ferreira, Masayo Otta, Kien Pham, Jorge Poco, Luciano Barbosa, Marcos Vieira, Marcos Lage, Joao Comba, Luis Gustavo Nonato, Luc Wilson, Heidi Werner, Muchan Park, Jonathas Costa, and many others

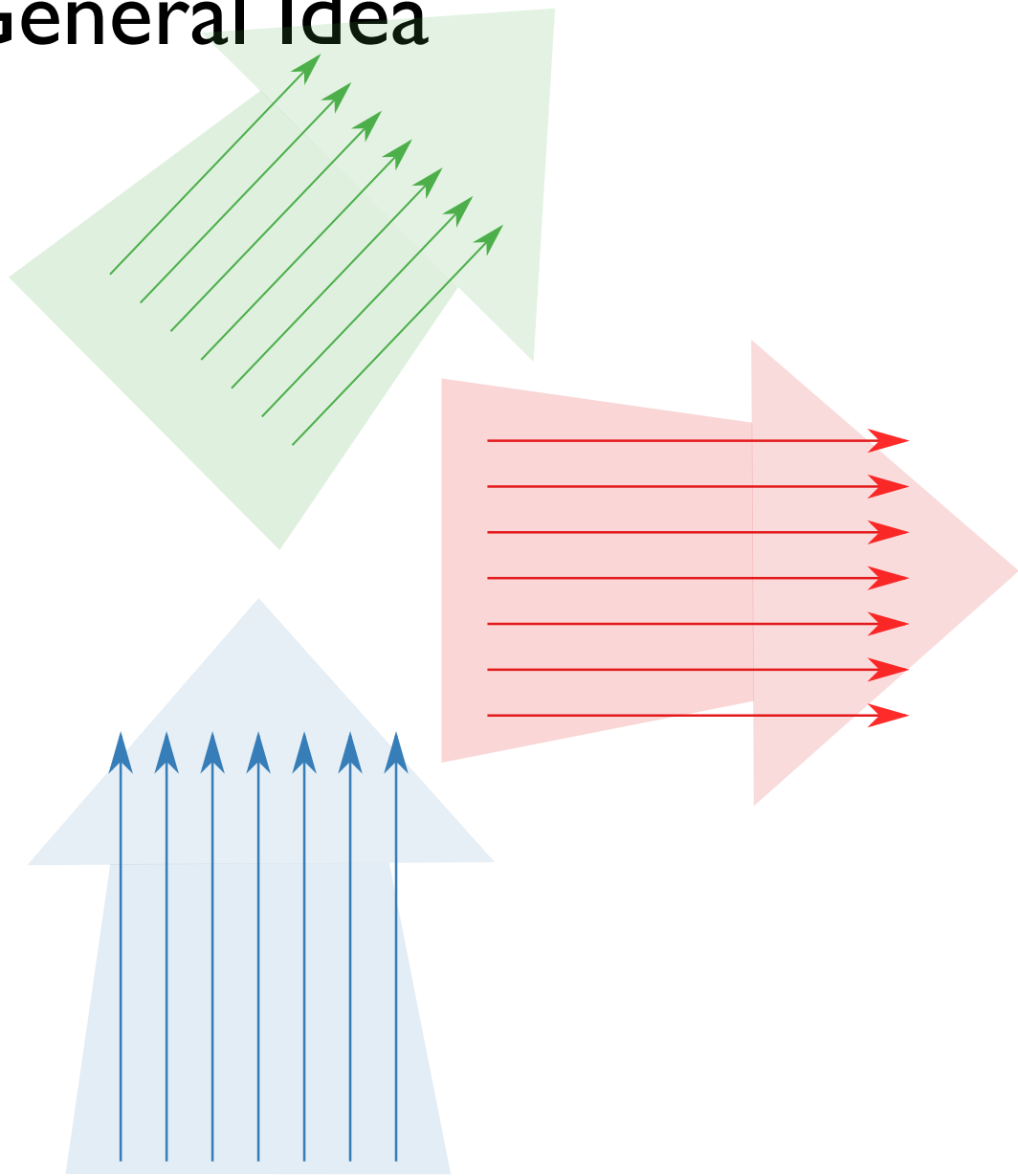
Vector Field K-Means

"Vector Field k-Means: Clustering Trajectories by Fitting Multiple Vector Fields," N Ferreira, JT Klosowski, C Scheidegger, C Silva, Computer Graphics Forum, 2013. ***Best Paper Honorable Mention***.

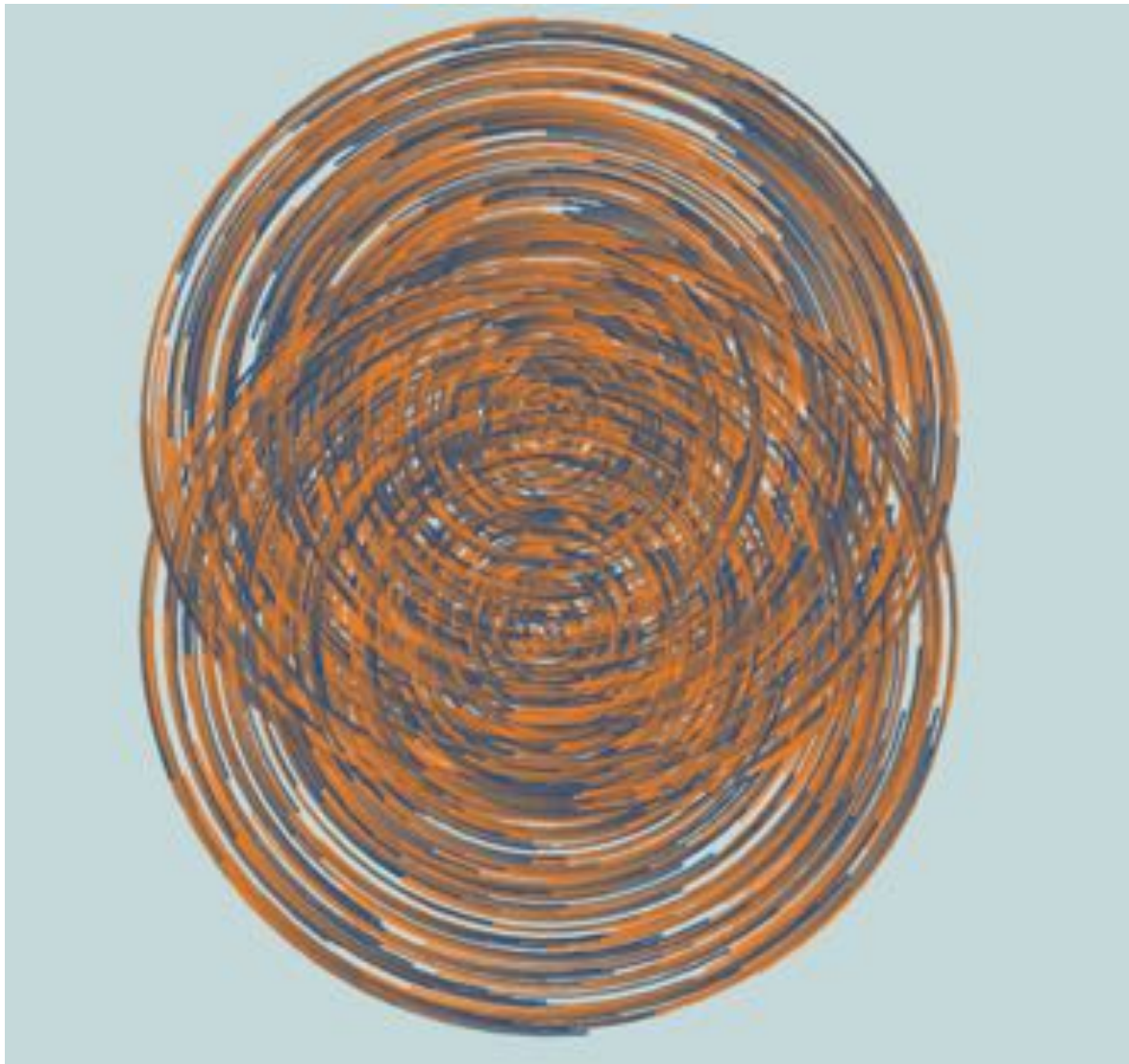
Trajectory Analysis

VFKM: General Idea

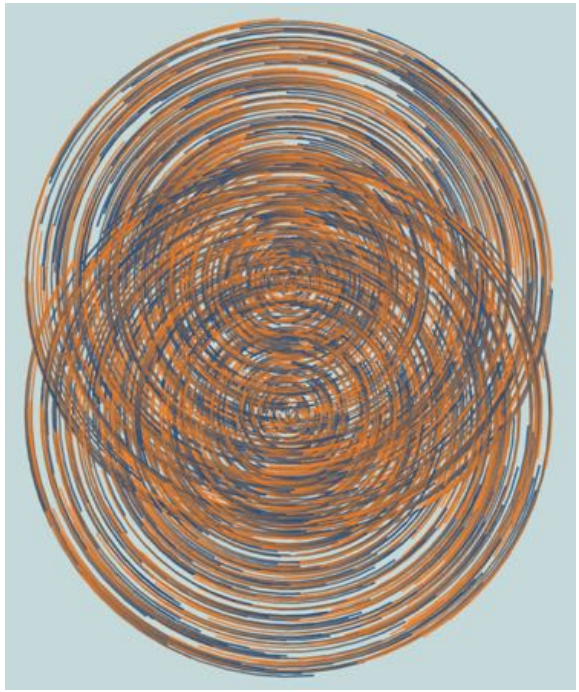


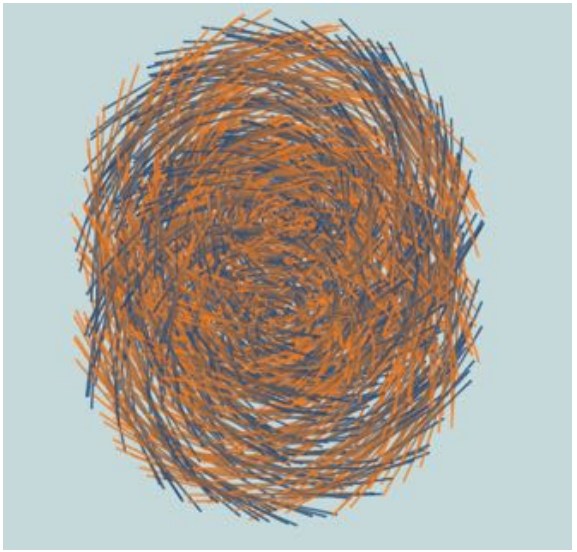


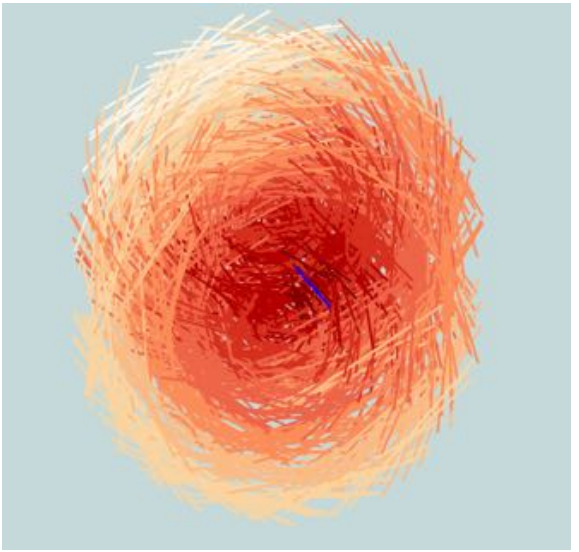
Trajectories with "missing data"



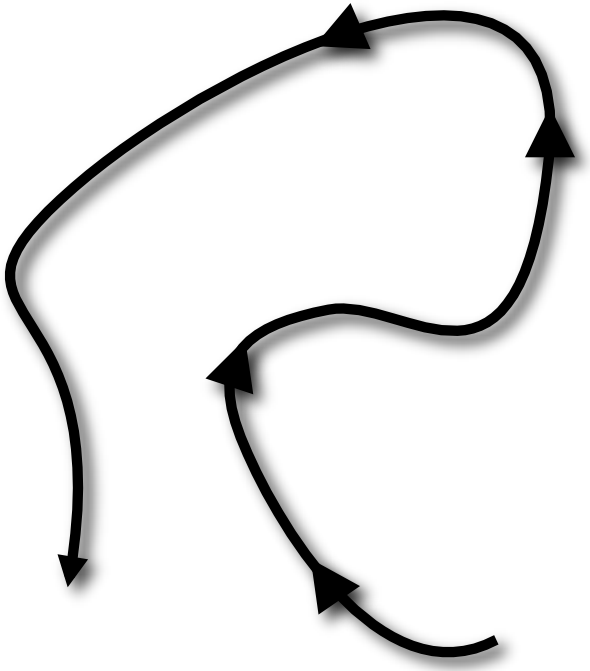
TraClus Results

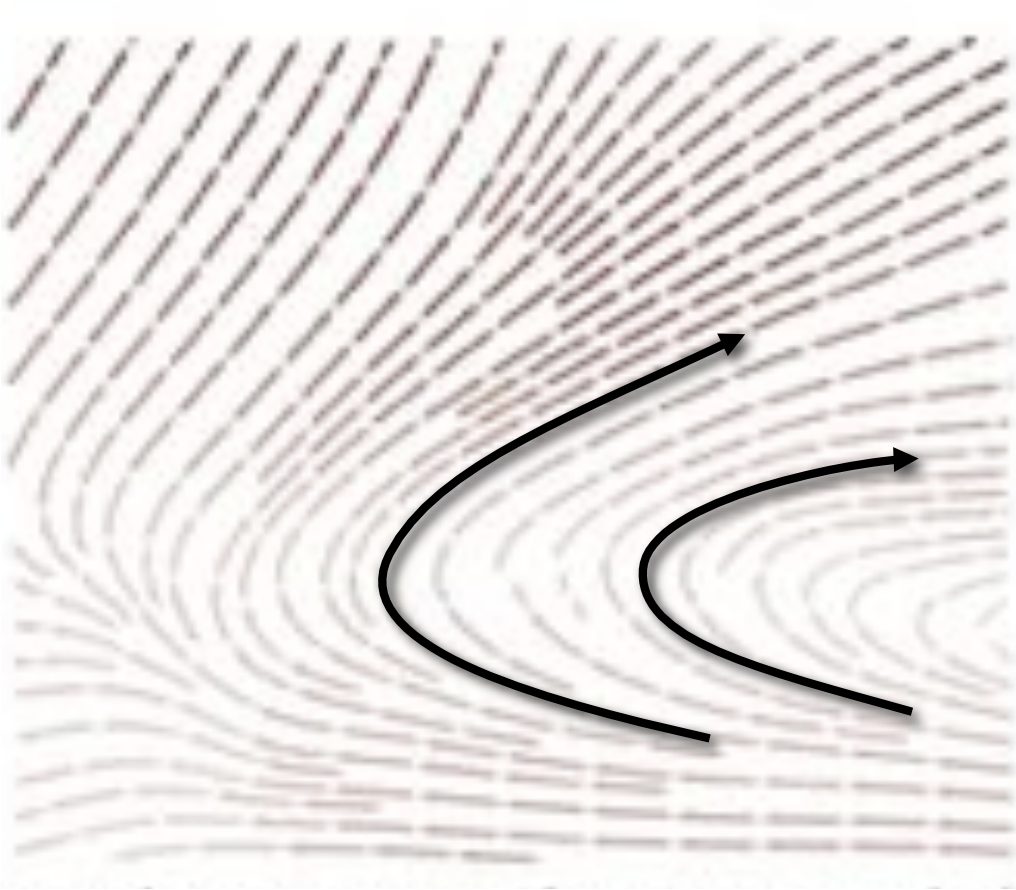






Observation: trajectories have tangent vectors





Vector Field k-Means

- Use vector field for clustering
- Natural way to encode direction and speed
- Handle partial trajectories through "trajectory modeling"
- Pattern "summaries" for free

Problem Formalization

Derive **k** nice vector fields and an **assignment** of trajectories to one of the vector fields such that each vf "*approximates*" the trajectories assigned to it

$$E'(X,\alpha) = \int_{t_0}^{t_1} ||X(\alpha(t)) - \alpha'(t)||^2 dt$$

Problem Formalization

Derive **k** nice vector fields and an **assignment** of trajectories to one of the vector fields such that each vf "*approximates*" the trajectories assigned to it

$$E'(X, \{\alpha_i\}) = \sum_{i=t_0}^{t_1} \int_{t_0}^{t_1} ||X(\alpha_i(t)) - \alpha_i'(t)||^2 dt$$

Problem Formalization

Derive **k** ~~nice~~ vector fields and an **assignment** of ~~trajectories~~ to one of the vector fields ~~such that each vf~~ "*approximates*" the trajectories assigned to it

$$E'(X, \{\alpha_i\}) = \sum_{i=t_0}^{t_1} \int_{t_0}^{t_1} ||X(\alpha_i(t)) - \alpha_i'(t)||^2 dt$$

$$E(X, \{\alpha_i\}) = ||\Delta X||^2 + \lambda E'(X, \{\alpha_i\})$$

Problem Formalization

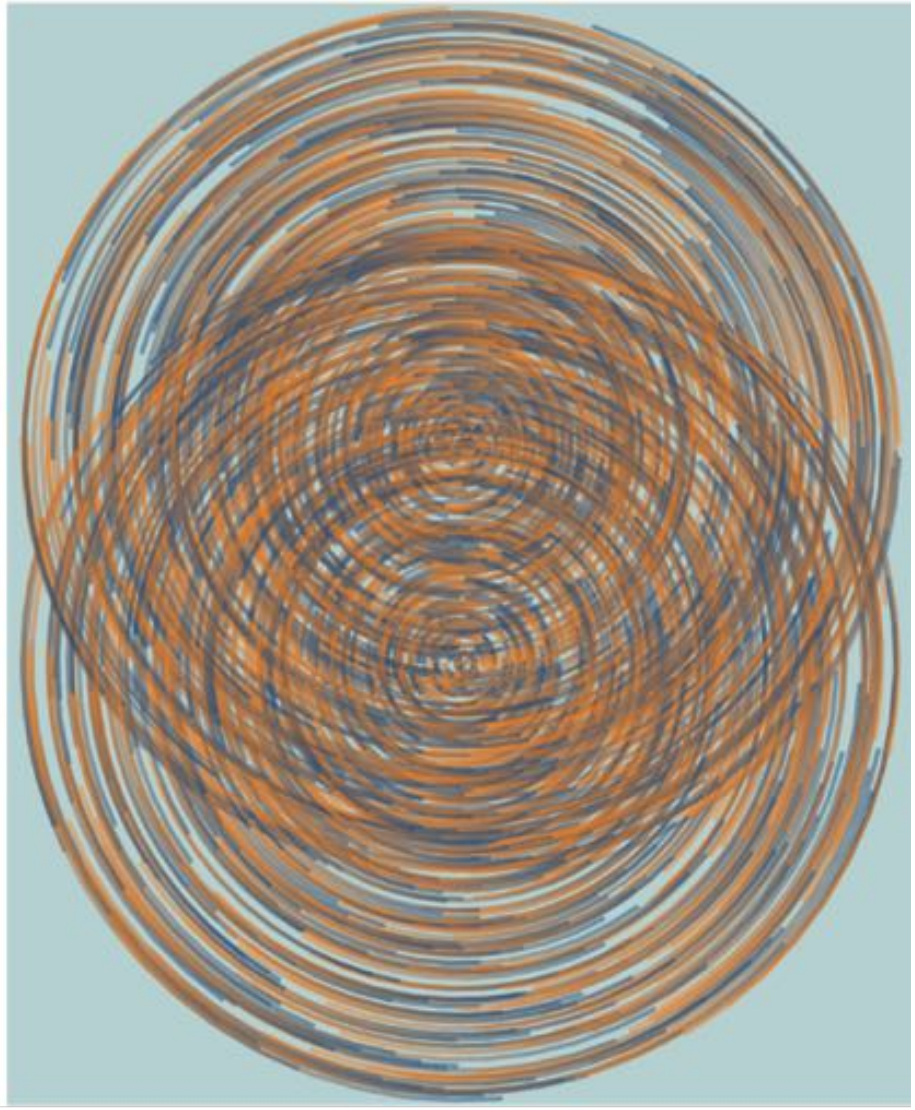
Derive **k nice vector fields** and an **assignment** of trajectories to one of the vector fields such that each vf "*approximates*" the trajectories assigned to it

$$E'(X, \{\alpha_i\}) = \sum_{i=t_0}^{t_1} \int_{t_0}^{t_1} ||X(\alpha_i(t)) - \alpha_i'(t)||^2 dt$$

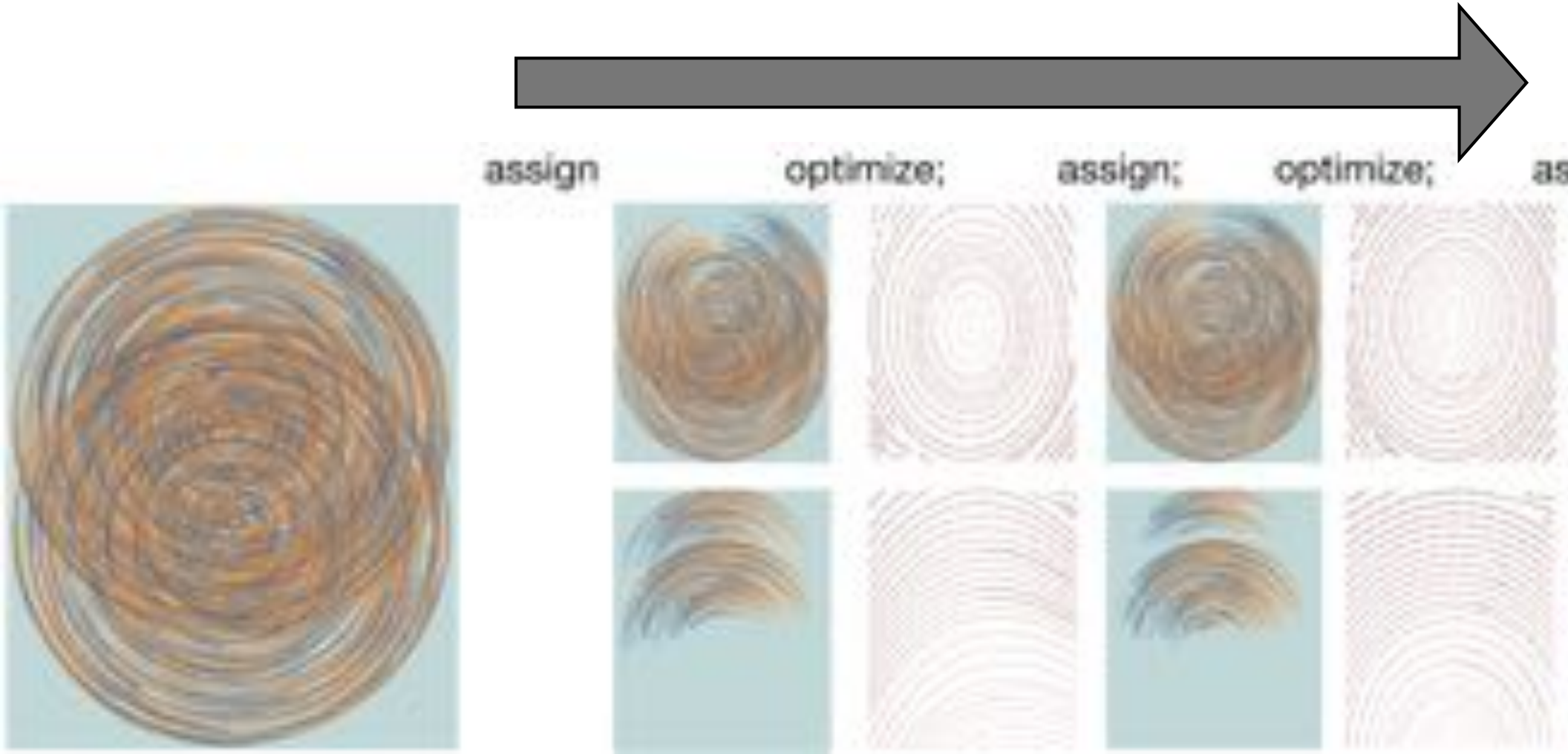
$$E(\{X_j\}, \{\alpha_i\}) = \sum_{j=1}^{k} ||\Delta X_j||^2 + \lambda E'(X_j, \{\alpha_i\}_{i \in \Phi^{-1}(j)})$$

Synthetic Dataset

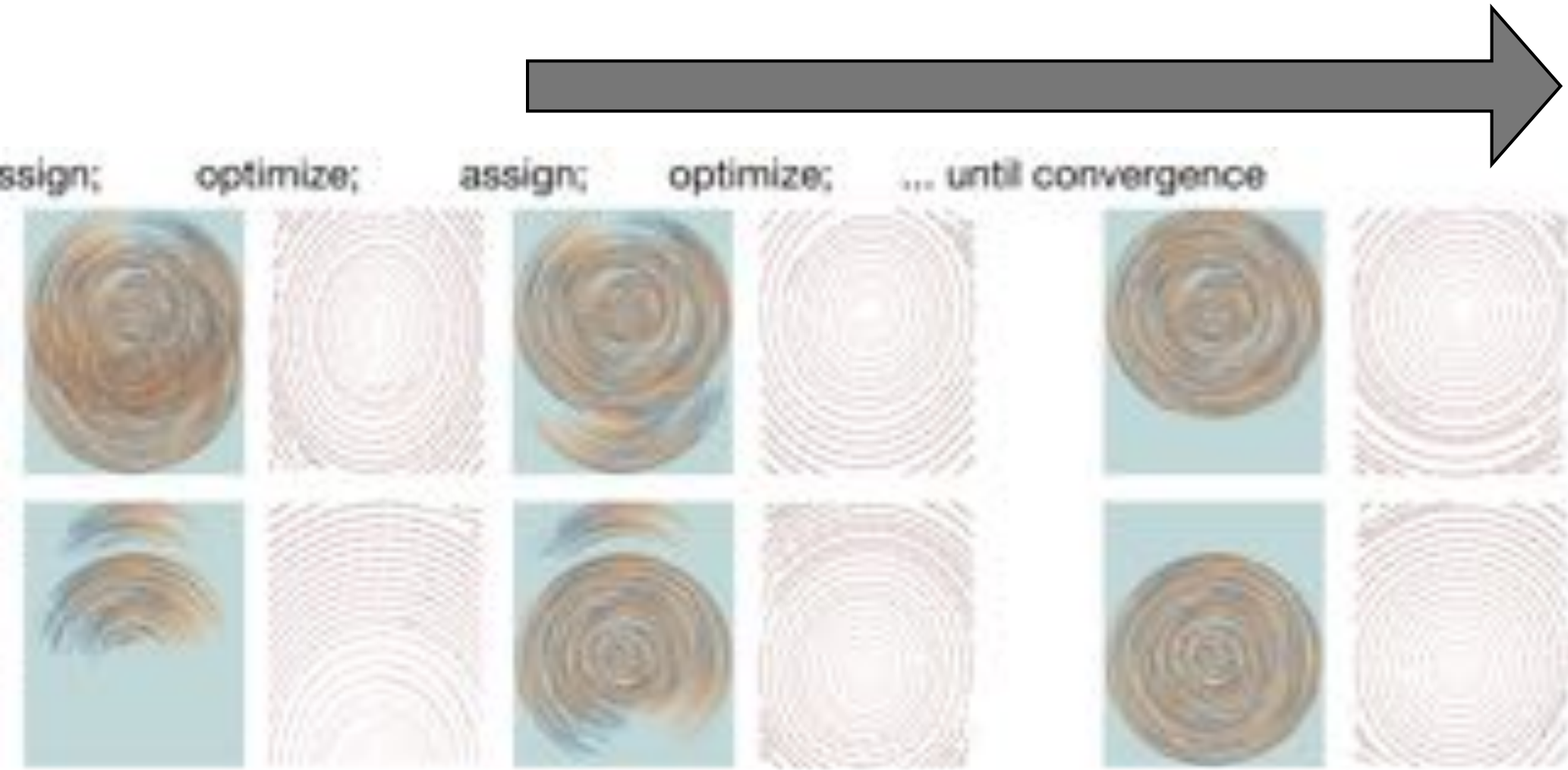
- 2000 Trajectories
- Extracted from two center patterns
- Colors Represent Direction: From Blue to Orange



Algorithm Execution



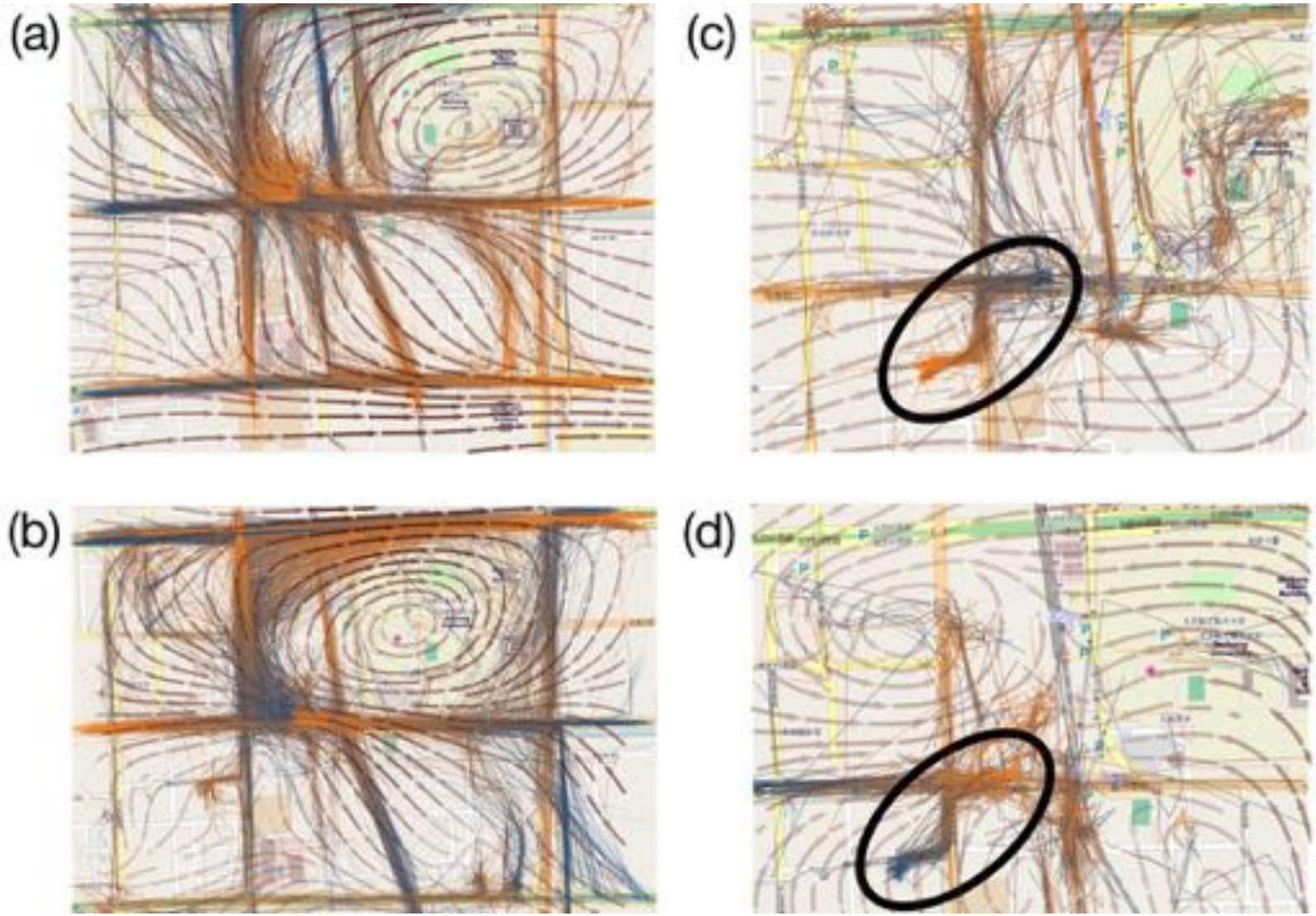
Algorithm Execution



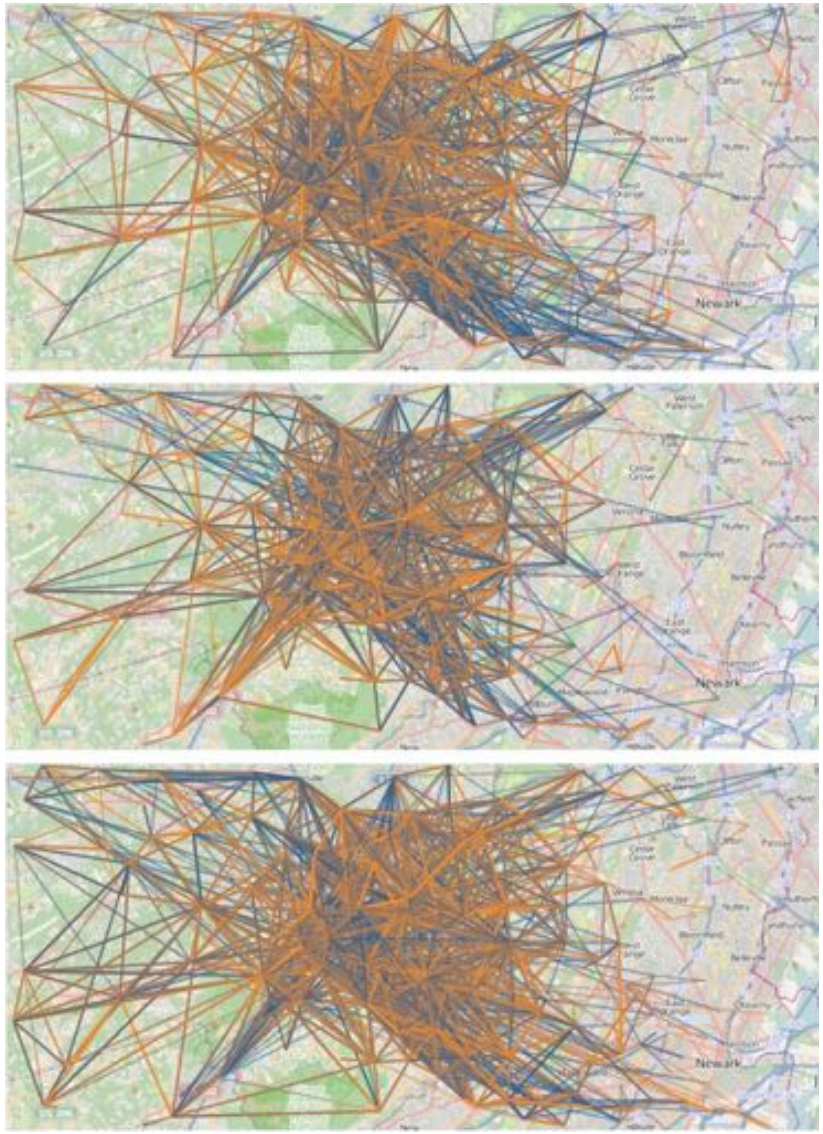
Geolife Dataset

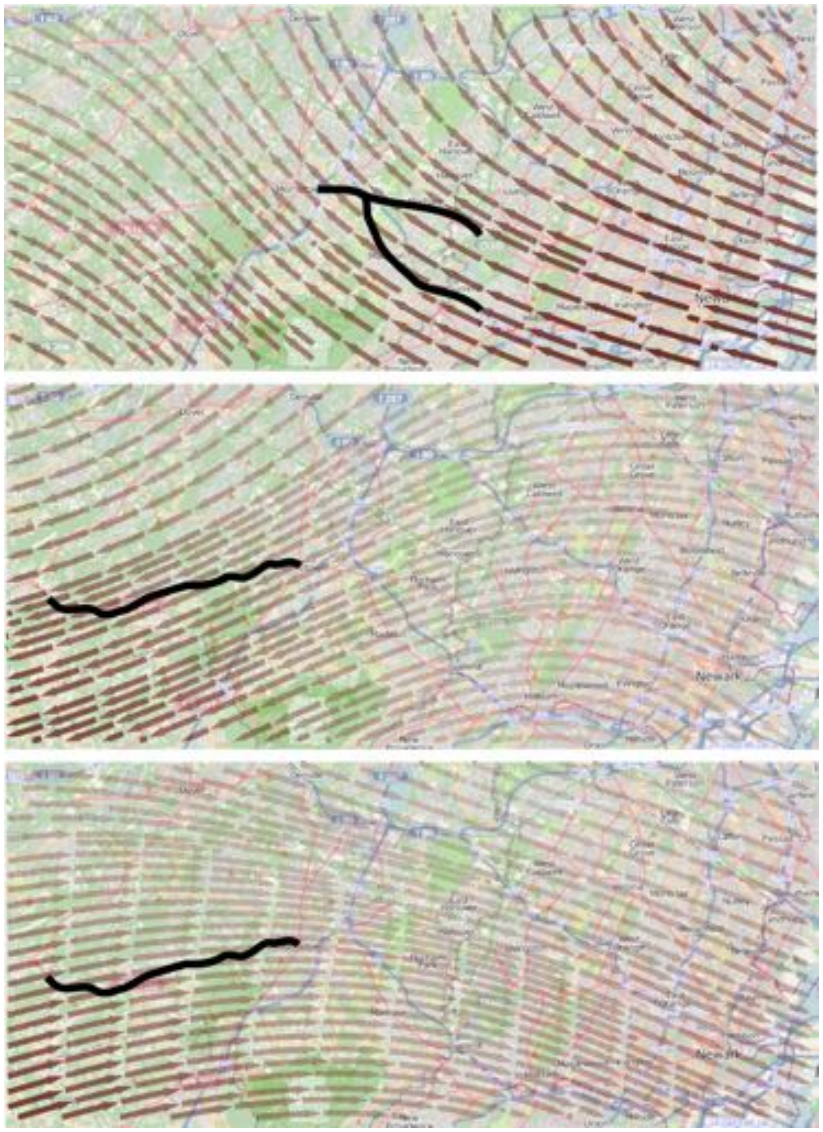
- GPS tracks of MSR employees
- 12833 trajectories

GeoLife

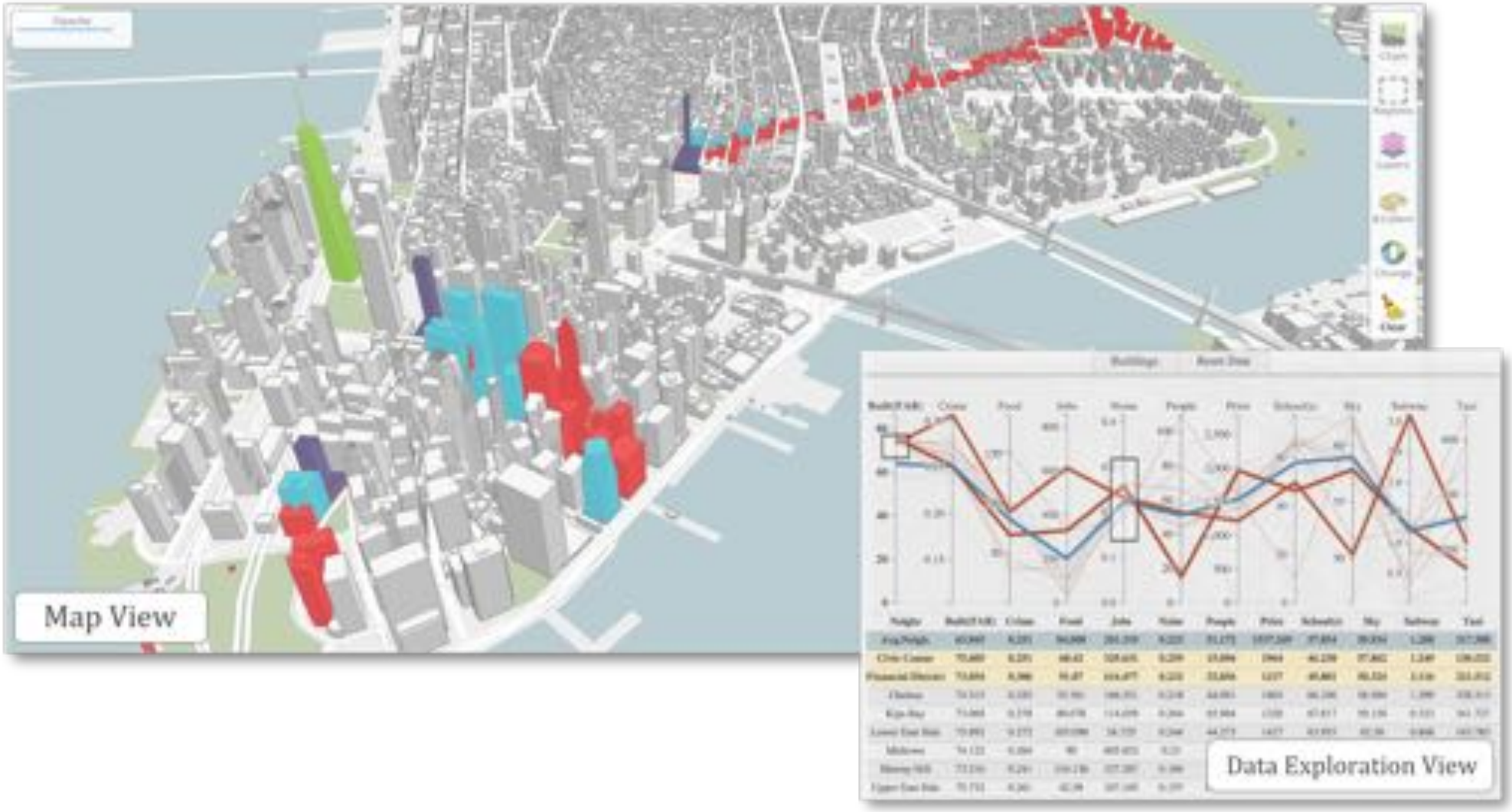


CDR Dataset / AT&T cell network





Urbane: first-of-a-kind UrbanGIS



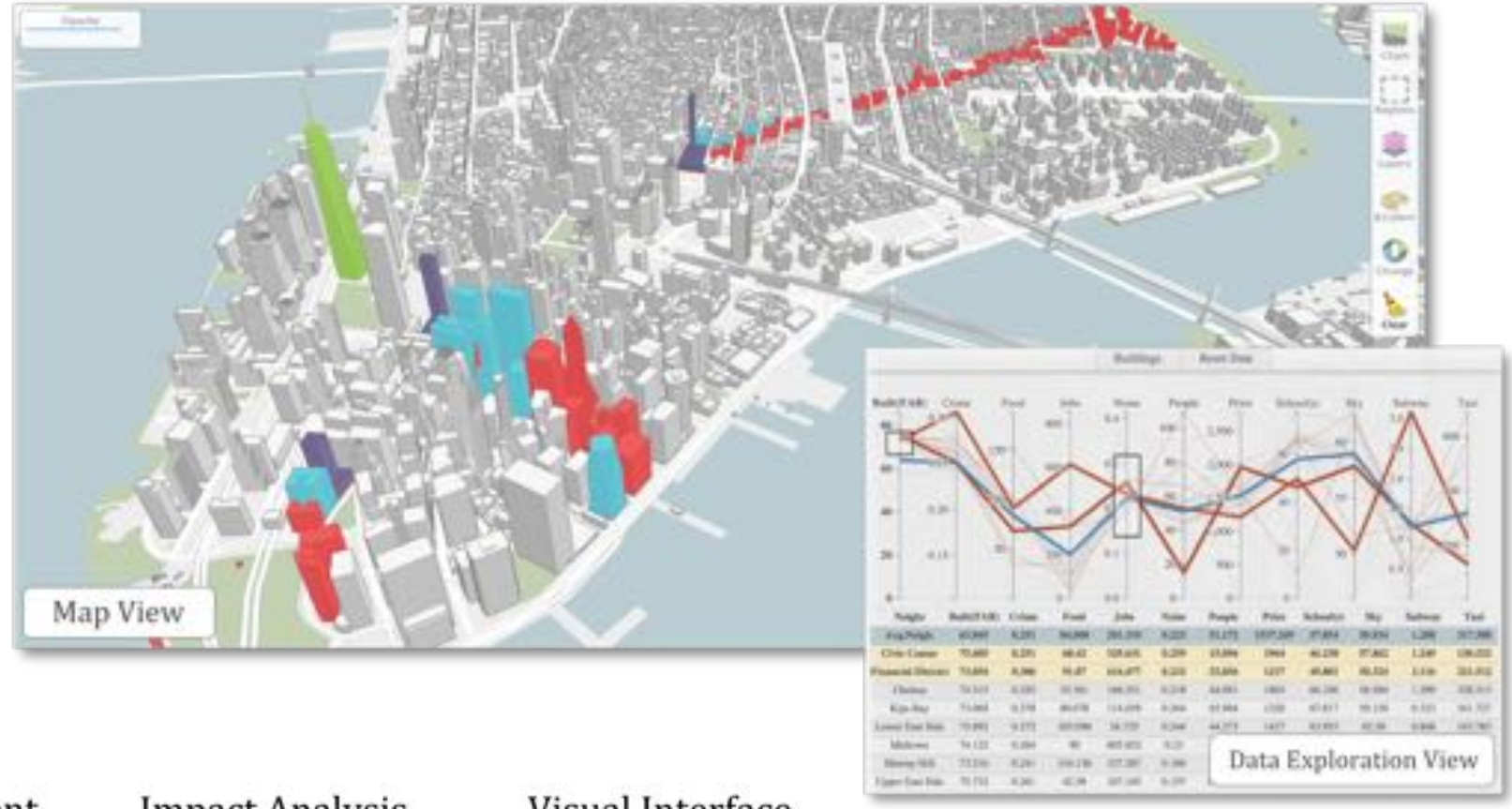
URBANE:

A 3D Framework to Support Data Driven Decision Making in Urban Developments

IEEE VAST 2015
Submission ID: 268

Harish Doraiswamy[1], Nivan Ferreira[1], Huy Vo[1], Claudio Silva[1], Marcos Lange[2], Muchan Park[3], Heidi Werner[3], Luc Wilson[3]

New York University[1], Universidade Federal Fluminense[2], Kohn Pederson Fox Associates PC[3]



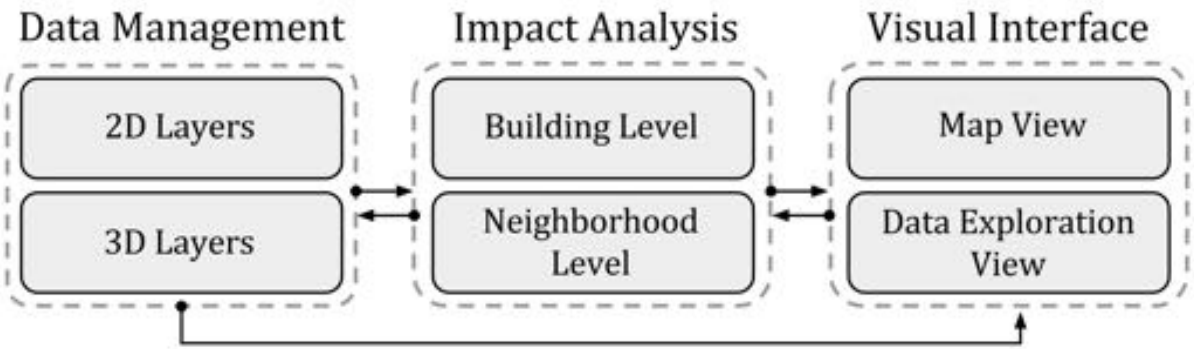


Fig. 2. The different components of Urbane and how they interact. The data management component supports the use of both 2D and 3D data layers. The impact analysis component enables the assessment of how new buildings affect their surroundings. The visual interface component supports the exploration of the data layers in the system.

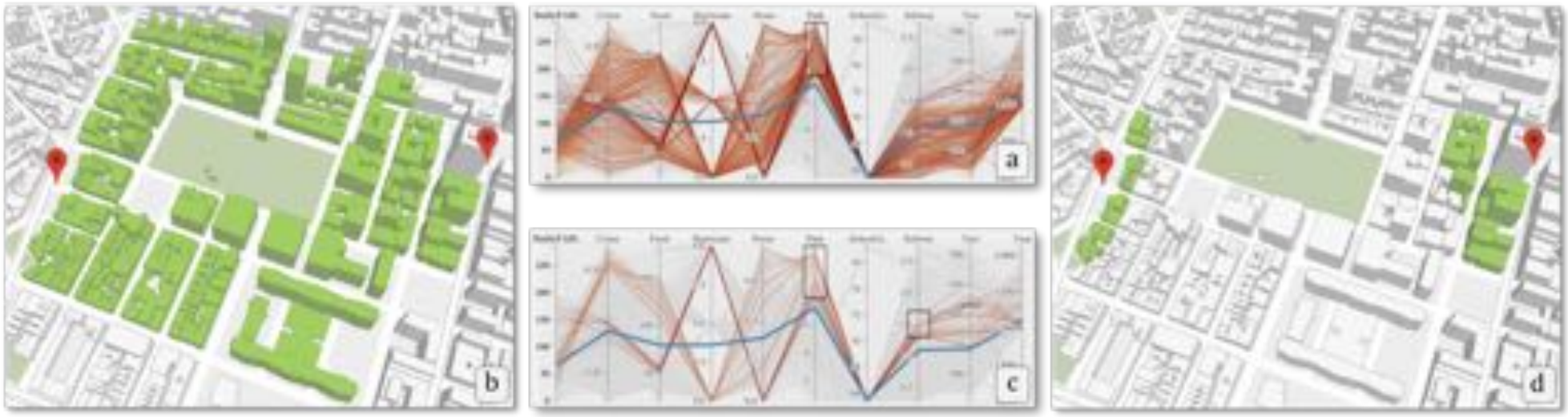


Fig. 5. Exploring the city at multiple scales using the data exploration view. **(a)** The user first selects only buildings close to a park using the PCC (the value corresponds to the area of park space weighted by distance). **(b)** The buildings satisfying the constraints are highlighted in the map view. **(c)** The user now selects only those sites having a high density of subway near them. **(d)** The buildings remaining after this filter is applied.

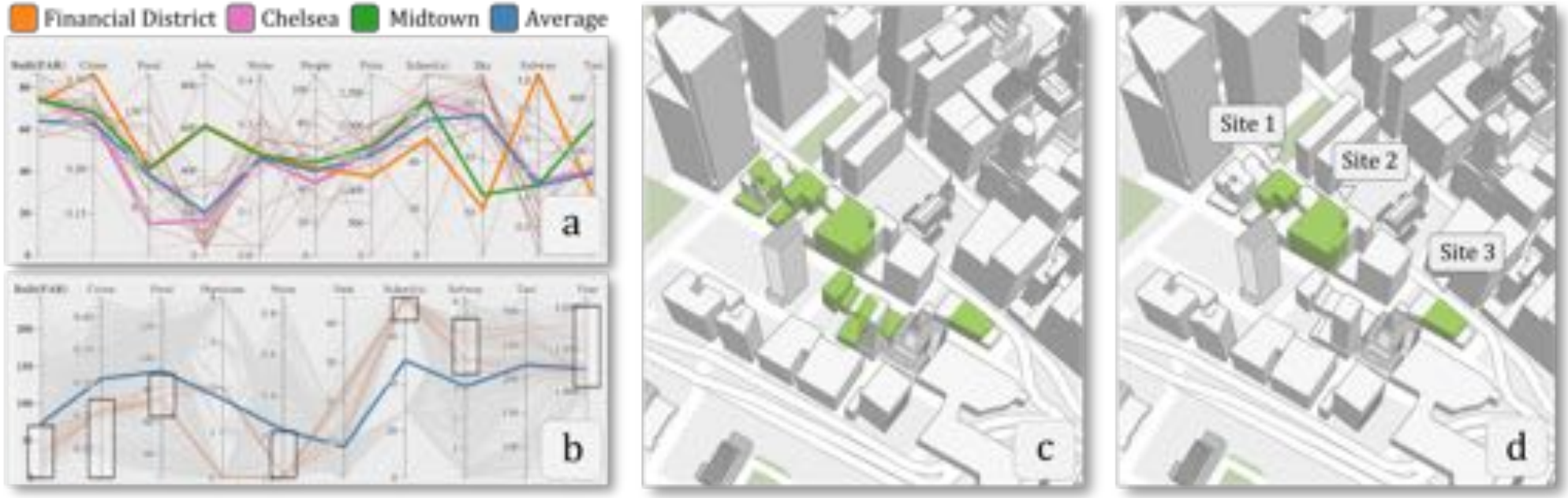


Fig. 6. Using Urbane to identify development sites in Financial District. The data exploration view of Urbane is used to study the characteristics of Financial District with respect to other neighborhoods in Manhattan (a). This is then used to filter (b) and identify potential development sites (c). Further filtering based on the site properties isolates three sites (d) that have high development potential.

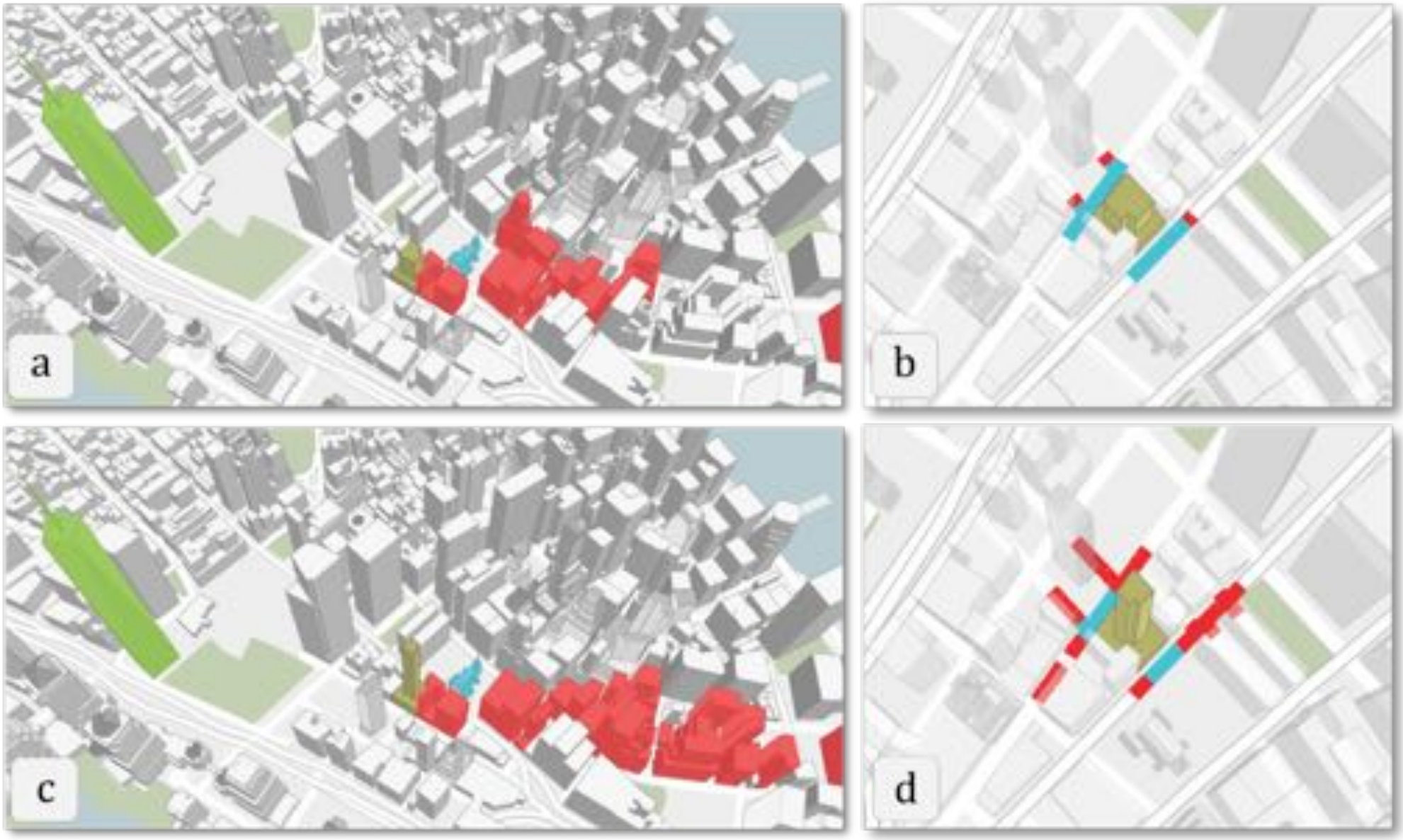


Fig. 7. Understanding the impact of different building designs. The view (a) and sky exposure (b) impact when using a 80×120 floor plate vs. the impact when using a 65×65 floor plate (c & d) for the proposed buildings.

Topology-based Catalogue Exploration Framework for Identifying View-Enhanced Tower Designs

SIGGRAPH Asia
Submission ID: 338

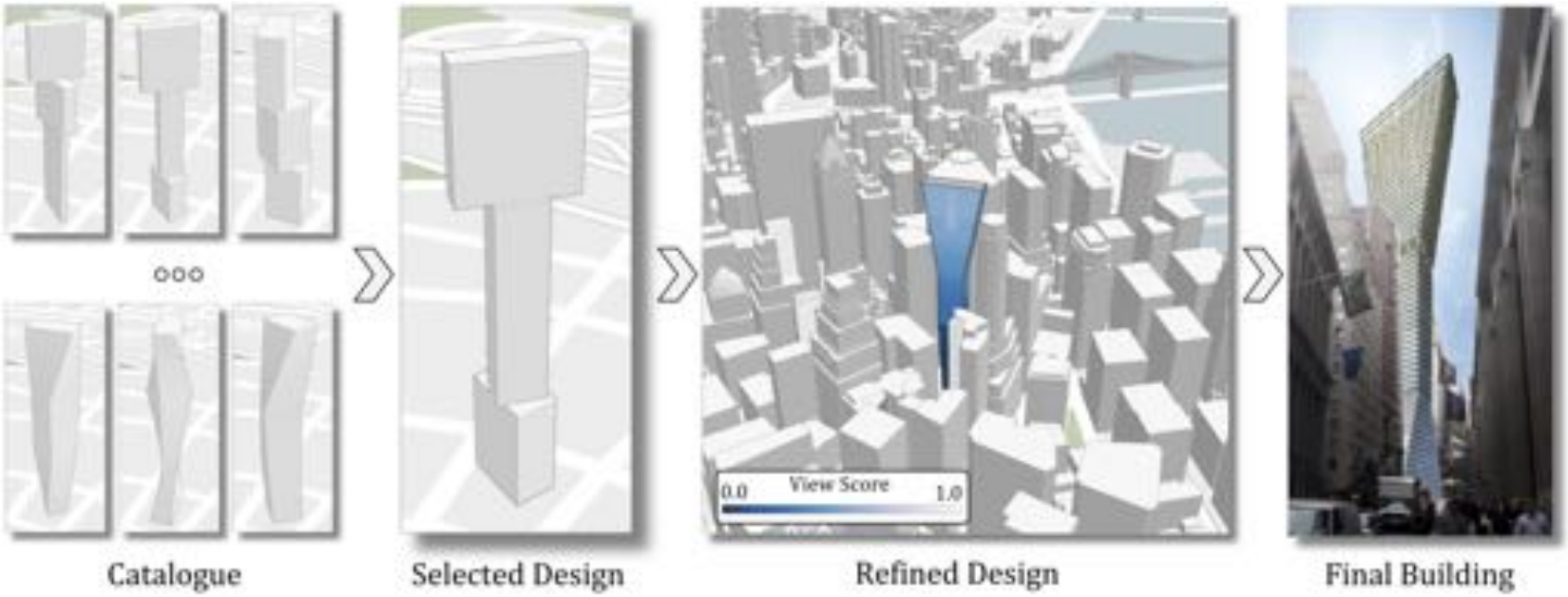
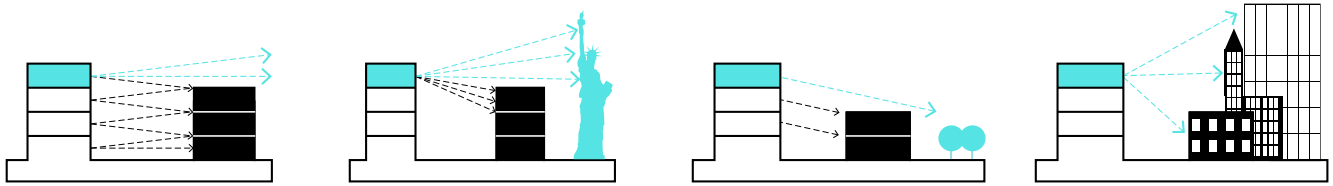


Figure 1: *The topology-based catalogue framework identifies building designs that have good performance with respect to outward views. In their conceptual design phase, architects use the catalogue to identify a few high performing designs, which are then further refined to satisfy client objectives and government regulations into a final building design.*

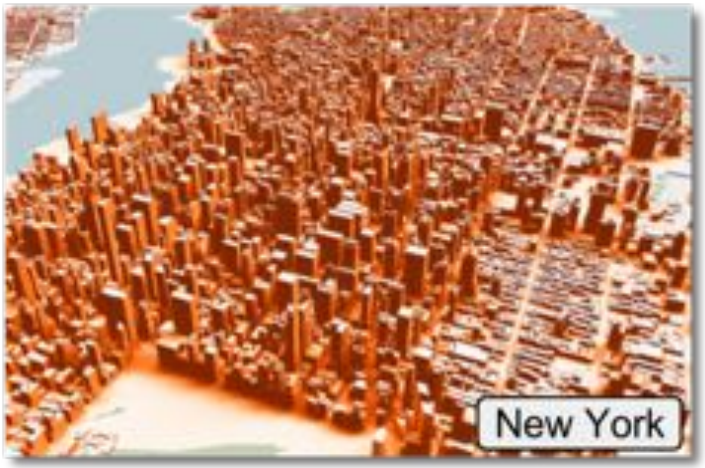


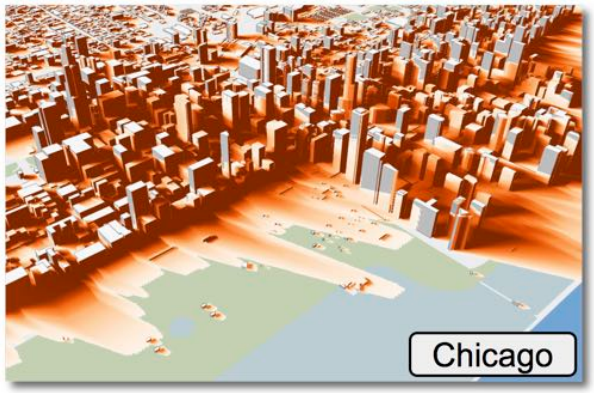
(a) Unobstructed View (b) Landmark Building View (c) Landscape View (d) Building Variation View

Figure 3: *Illustration of the different view types that are used to evaluate the views from a building.*

Figure 9: *The views obtainable from different height zones (20 floors each) from the two sites in Financial district and Midtown. The buildings highlighted in red indicate the reach from these zones. The inset shows the typical view from a representative window.*

Mapping the Shadows of a City



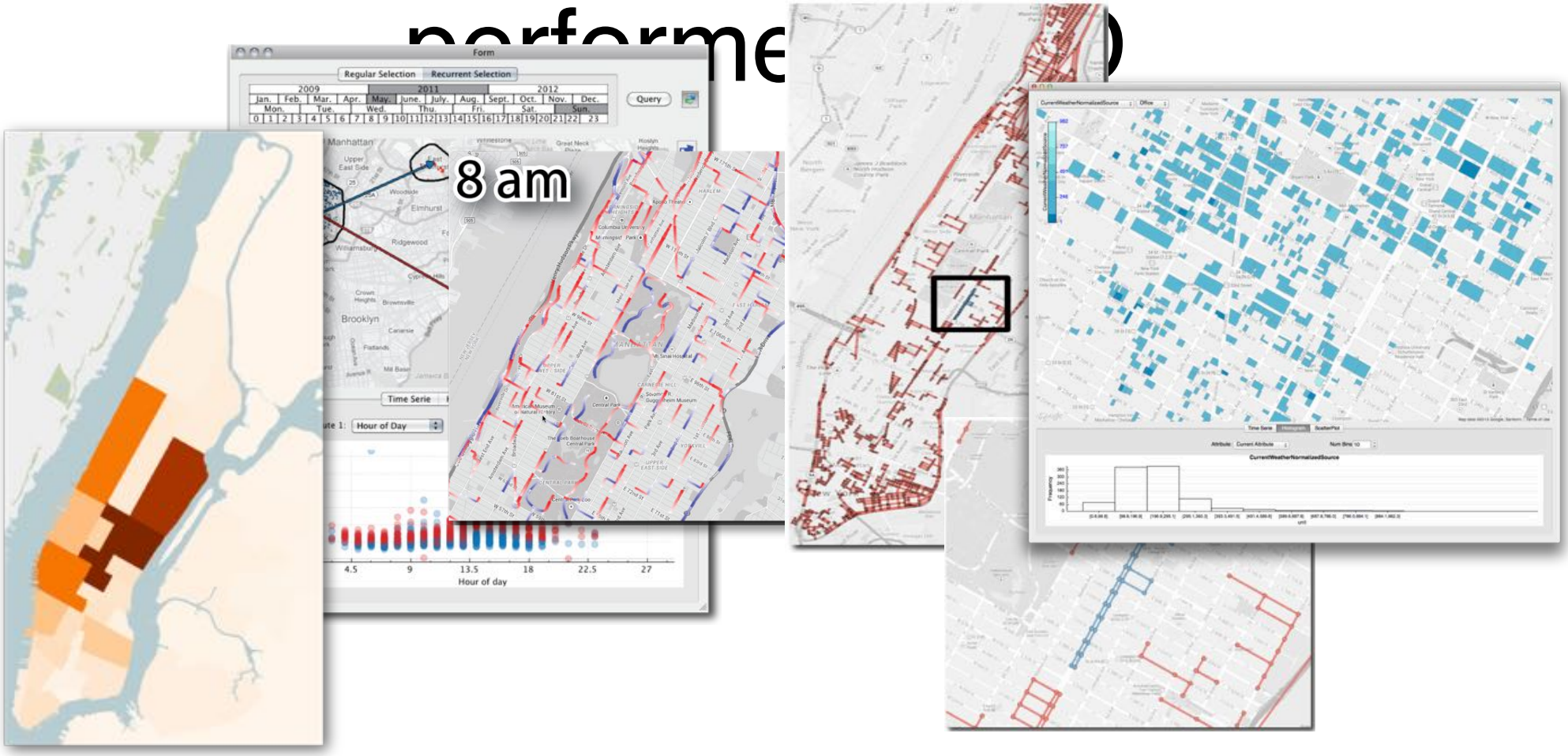


Harish Doraiswamy

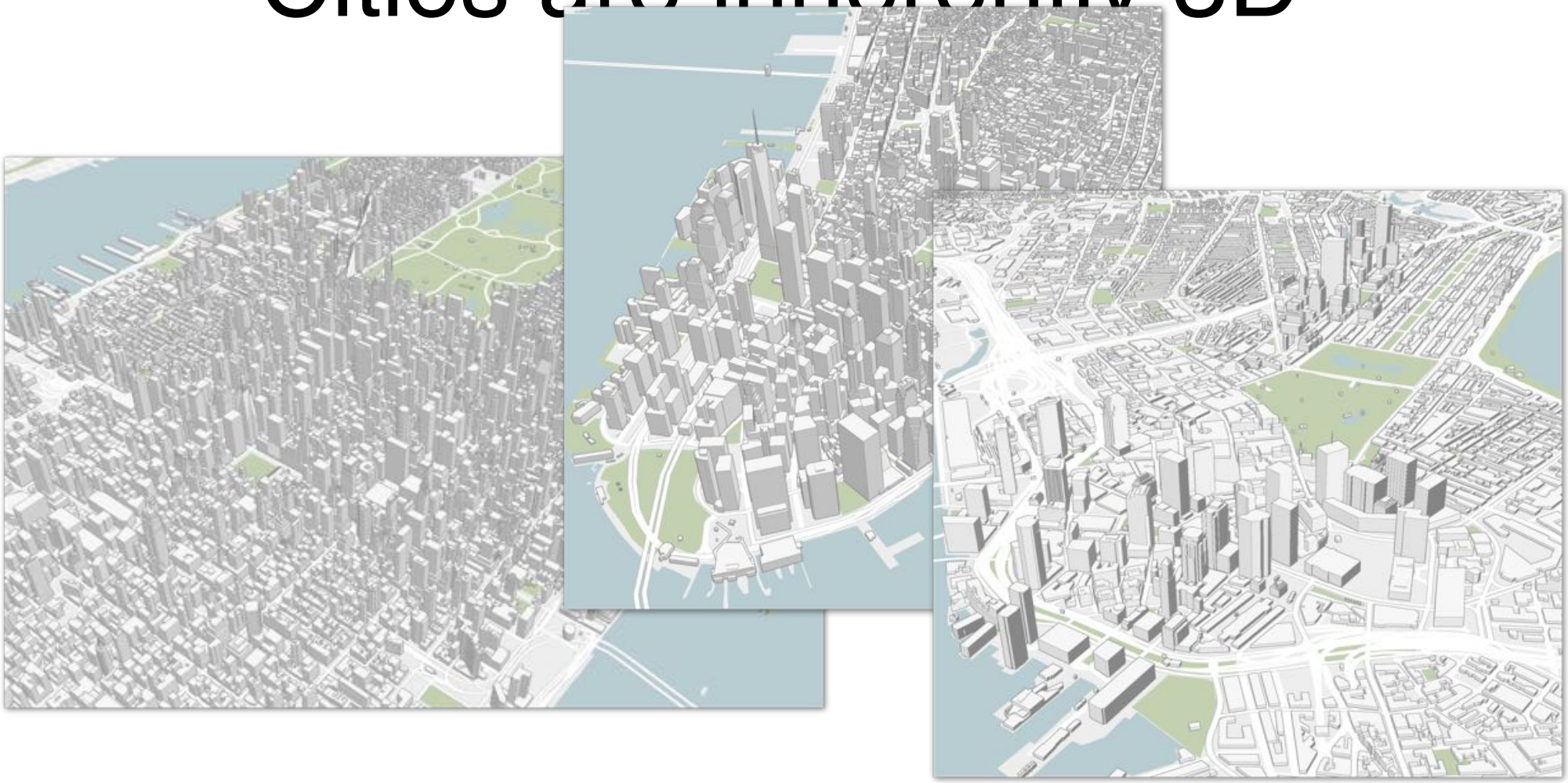
in collaboration with

Fabio Miranda, Marcos Lage, Luc Wilson, Mondrian Hsieh, and Claudio Silva

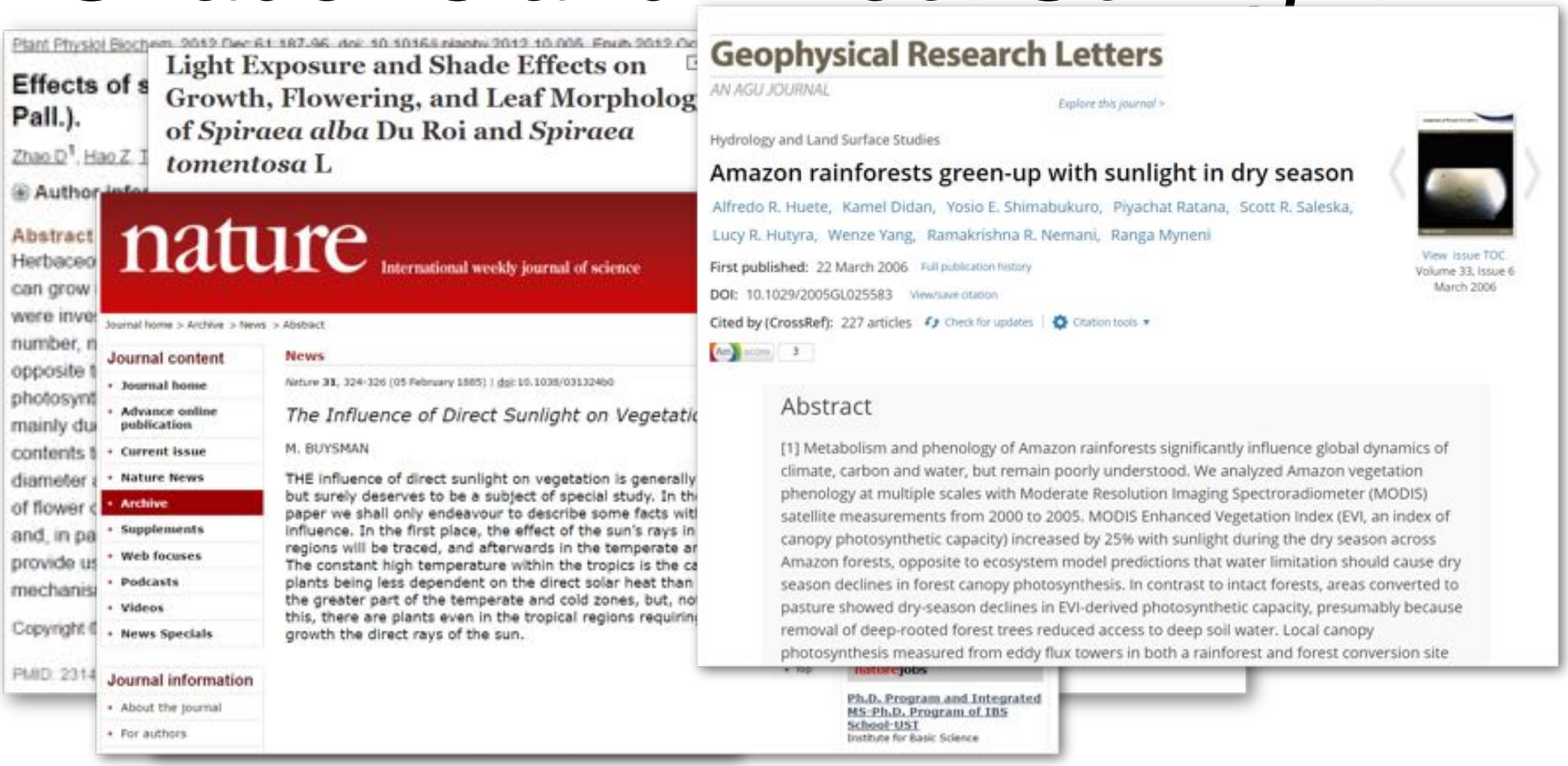
Urban Analysis traditionally



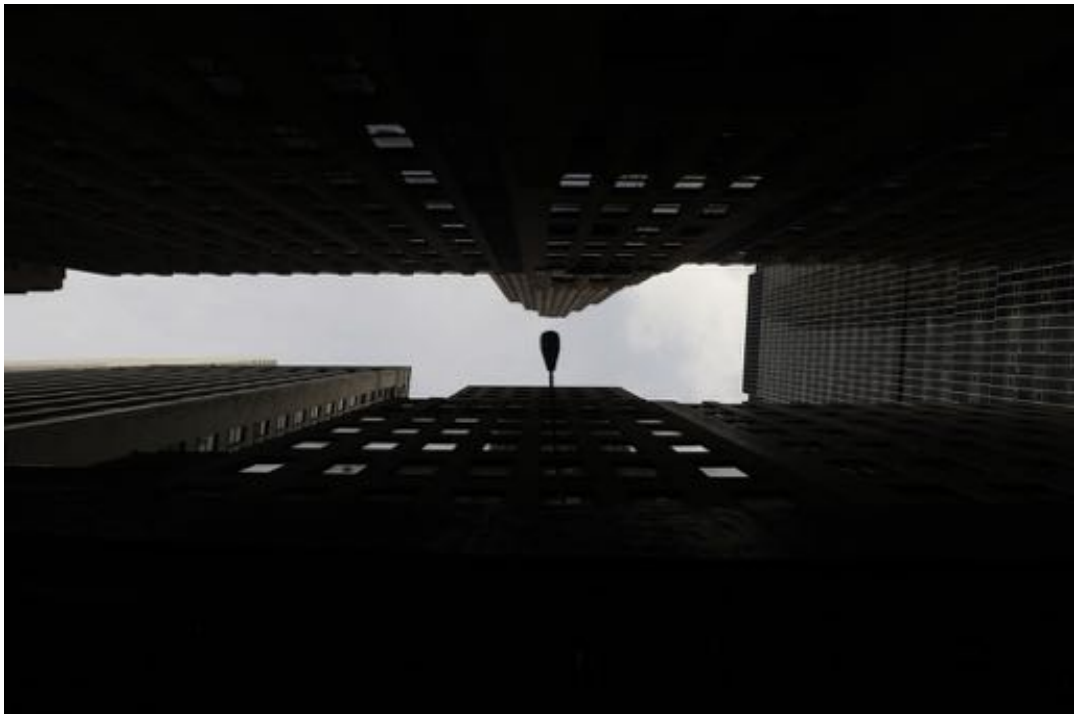
Cities are inherently 3D



Shadows and Direct Sunlight



Shadows and Direct Sunlight



[Source: New York Times]

Shadows and Direct Sunlight

SHADOWS

CHAPTER 8

Within urban environments, the structures constituting the city's built fabric constantly c
ate vicinity. As the city develops and redevelops, the extent and duration of the shadows
cess continues, direct sunlight exposure becomes an increasingly scarce resource for peo
focuses on the interaction between proposed new and altered structures and the shad
space, historic and cultural resources, and natural areas.

Sunlight and shadows affect people and their use of open space all day long and throughc
fects vary by season. Sunlight can entice outdoor activities, support vegetation, and enl
such as stained glass windows and carved detail on historic structures. Conversely, shadow
and sustainability of natural features and the architectural significance of built features.

The purpose of this chapter is to assess whether new structures may cast shadows o
accessible resources or other resources of concern such as natural resources, and to asses
pact. Potential mitigation strategies and alternatives are also presented and should be ex
verse shadow impacts are identified. Because of the sunlight-sensitive nature of many ope
al resources, and natural resources, this chapter is closely linked to the data and anal
Space," Chapter 9, "Historic and Cultural Resources," and Chapter 11, "Natural Resources."

The majority of projects subject to CEQR do not require a detailed shadow analysis. Sectic
of analysis to screen most projects for the purpose of assessing shadow impacts. As with
under CEQR, it is important for an applicant to work closely with the lead agency during
view process. The lead agency may determine that it is appropriate to consult or coor
technical agencies for a particular project. The New York City Department of City Planni
for information, technical review, and recommendations relating to shadows. With regar
City Landmarks Preservation Commission (LPC), the New York City Department of Environ

City Environmental Quality Review
Manual - NYC.gov

Ancient lights

LAW

WRITTEN BY: The Editors of Encyclopaedia Britannica
LAST UPDATED: 9-12-2008 See Article History

Ancient lights, in English property law, the right of a building or house owner to the light received from and through his windows. Windows used for light by an owner for 20 years or more could not be obstructed by the erection of an edifice or by any other act by an adjacent landowner. This rule of law originated in England in 1663, based on the theory that a landowner acquired an easement to the light by virtue of his use of the windows for that purpose for the statutory length of time. The doctrine did not acquire wide acceptance by courts in the United States.

Ancient lights signs below windows in Clerkenwell, London.
Mike Newman

Boston developer wants exception to shadow law

Tall order is sought for tower pr

Donna Goodison Monday, April 10, 2017

Credit: COURTESY RENDERING

TOWER OF POTENTIAL: City officials say an exemption from two new development is necessary to bring in $153 million. (ARTIST

Menu Business & Tech

Shadow debate shows how d

By Shirley Leung | GLOBE STAFF APRIL 25, 2017

It was as close as we would get to a mea culpa from city planning and development czar Br Golden on Shadowgate.

"Look, this isn't the way we wanted it," Golden told me after being grilled by the Boston City Council Monday about changing two-decade-old laws that limit shadows on the Public Garden and Boston Common.

Menu Opinion Subscribe Members Sign In

The absurdity of the Boston Common shadow debate

The Winthrop Square Ga

By Dante Ramos

I DON'T KNOW HOW Common is ridiculous

At earlier points in ou

Boston City Council votes in favor of changing 'shadow' law

Council backs shade

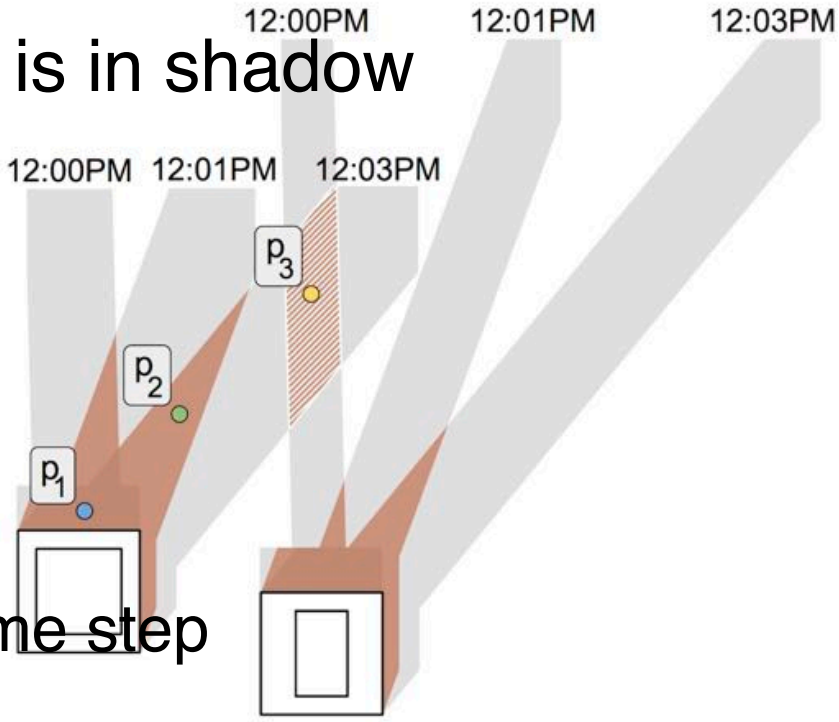
Dan Atkinson Thursday, April 27, 2017

Credit: Stuart Cahill

The Winthrop Square parking garage at 240 Devonshire St. Wednesday, April 26, 2017. (Staff photo by Stuart Cahill)

Shadow Accumulation

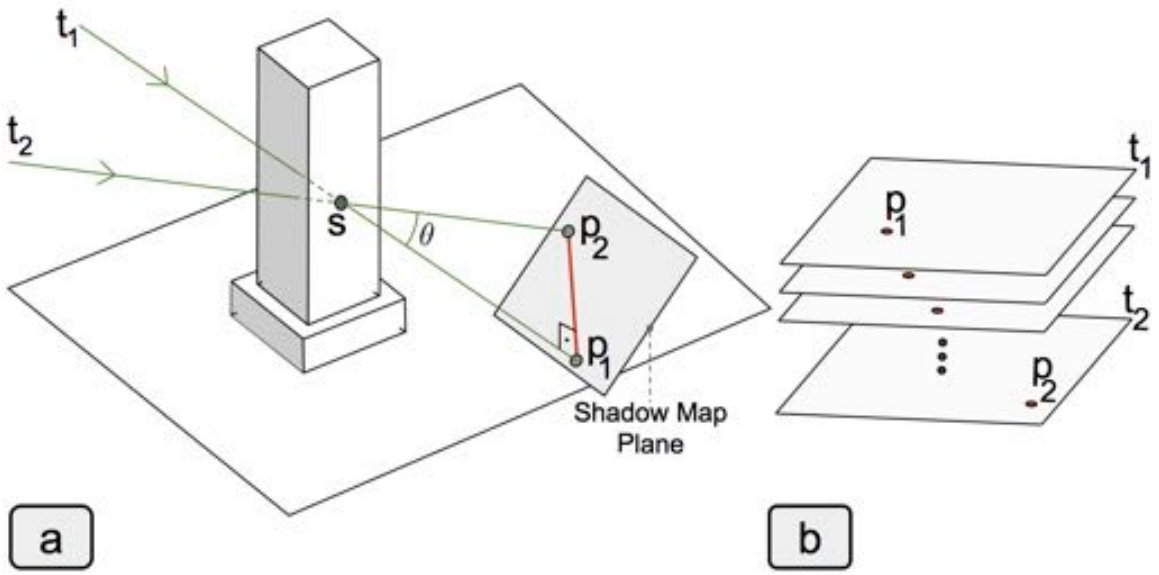
- Amount of time a given point is in shadow
- Depends on the sun position (day of the year)
- > 10,000 buildings
- Existing techniques
  - Focus on fixed light source
  - Compute shadow for every time step
  - Expensive



Shadow Accumulation

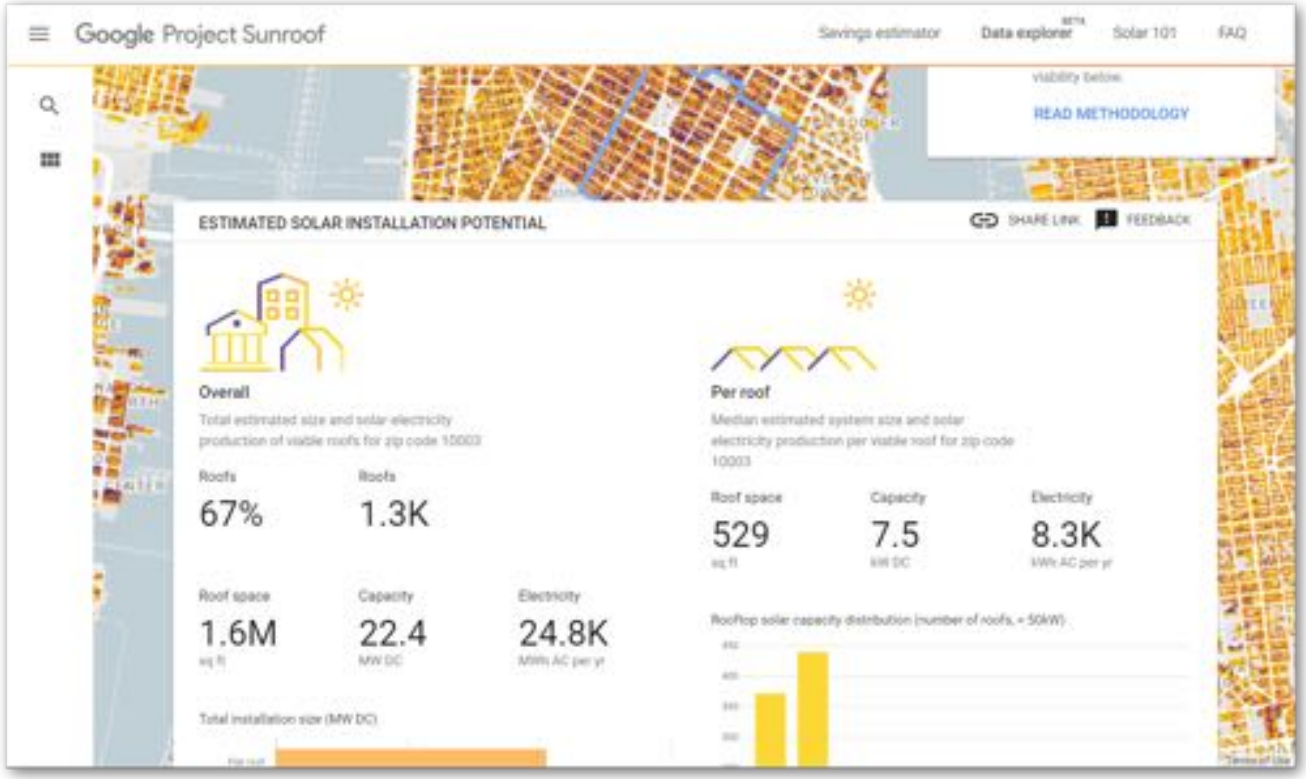
Efficient

70X Speedup



Shadow Accrual Maps

Shadows and Direct Sunlight




The New York Times
Mapping the Shadows of New York City: Every Building, Every Block
Amount of Day in Shadow

Urban Pulse:
Capturing the Rhythm of Cities

Fabio Miranda, **Harish Doraiswamy**, Marcos Lage, Kai Zhao,
Bruno Gonçalves, Luc Wilson, Mondrian Hsieh, Claudio Silva

NYU

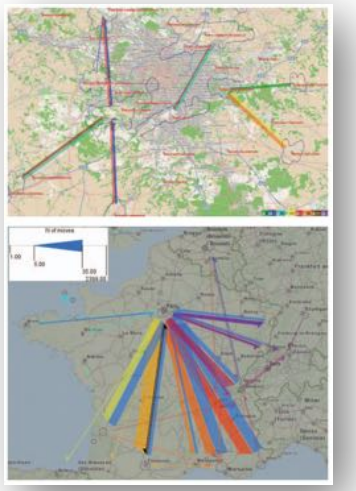
Instituto de Computação

KPF

Understanding Urban Data

- Analysis of vehicular / traffic movement

  [Andrienko et al. 2013]
  [Poco et al. 2015]



[Andrienko et al. 2013]

[Poco et al. 2015]

Understanding Urban Data

- Human mobility

  [Veloso et al. 2011]
  [Liang et al. 2012]
  [Peng et al. 2012]

- Identifying patterns / events in the da

  [Doraiswamy et al. 2014]

[Veloso et al. 2011]

[Doraiswamy et al. 2014]

Understanding Urban Data

- Gentrification and Real Estate

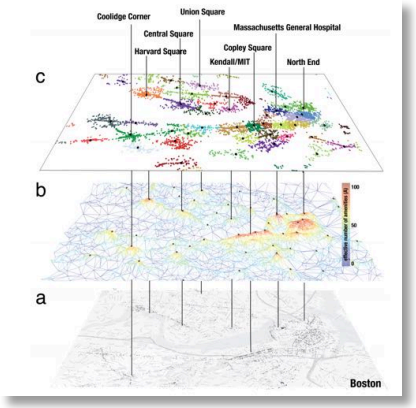
  [Hidalgo and Castañer 2015]
  [Sun et al. 2013]

- Air Pollution

  [Qu et al. 2007]

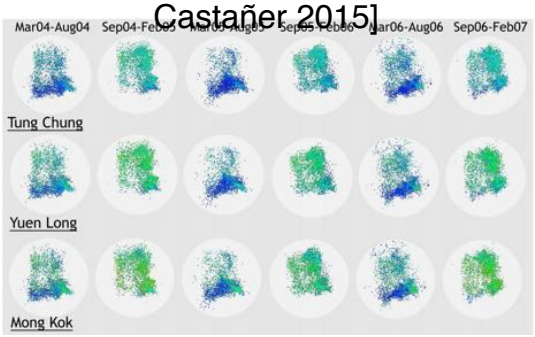
- Public Utility Service

  [Zhang



[Hifalgo and Castañer 2015]

[Zhang et al. 2014]



[Qu et al. 2007]

Objective

How to analyze / compare different properties of a city?

- How do cities behave during different times?
  - Summer vs. Winter
  - Weekdays vs. Weekends
- Data sets about different cultural communities in a city
  - What patterns do the different communities follow?
  - How do these patterns compare?

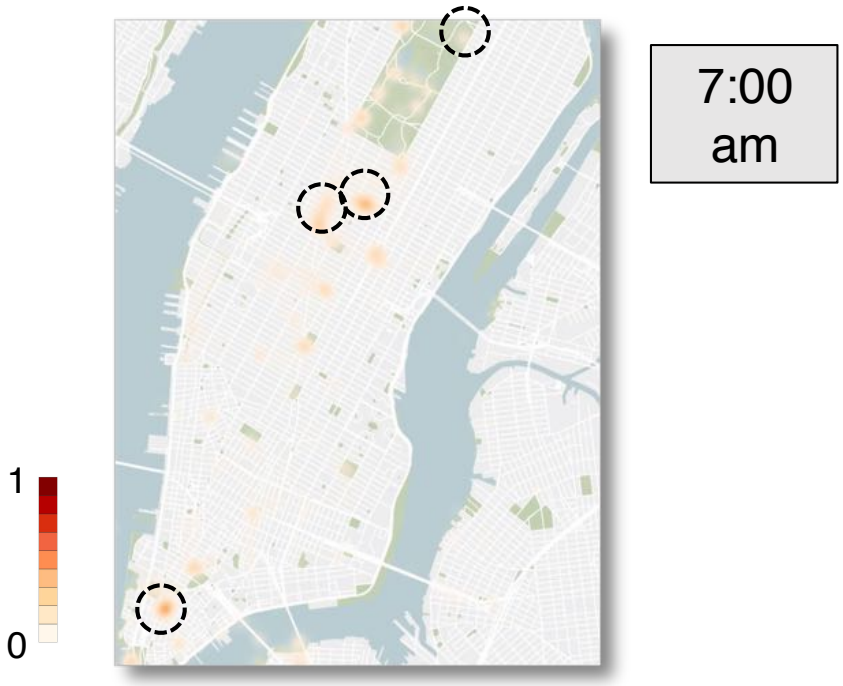
Objective

How to compare cities?

- Design of public spaces
  - Understand what works / doesn't work in one city
  - Use this to improve design in another city
- Understand properties of neighborhoods
  - Compare "activity" between neighborhoods with similar properties
  - Compare properties between neighborhoods with similar "activity"

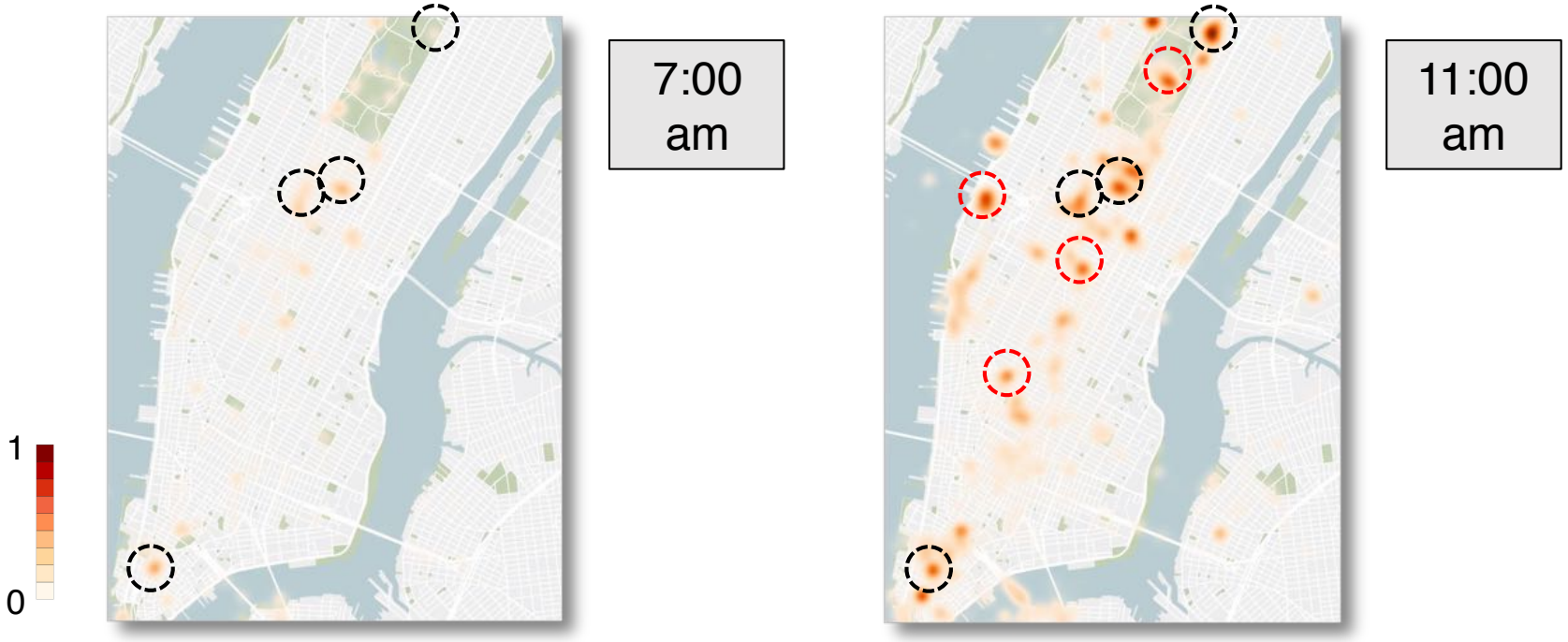
Urban Pulse

- Flickr activity in New York City



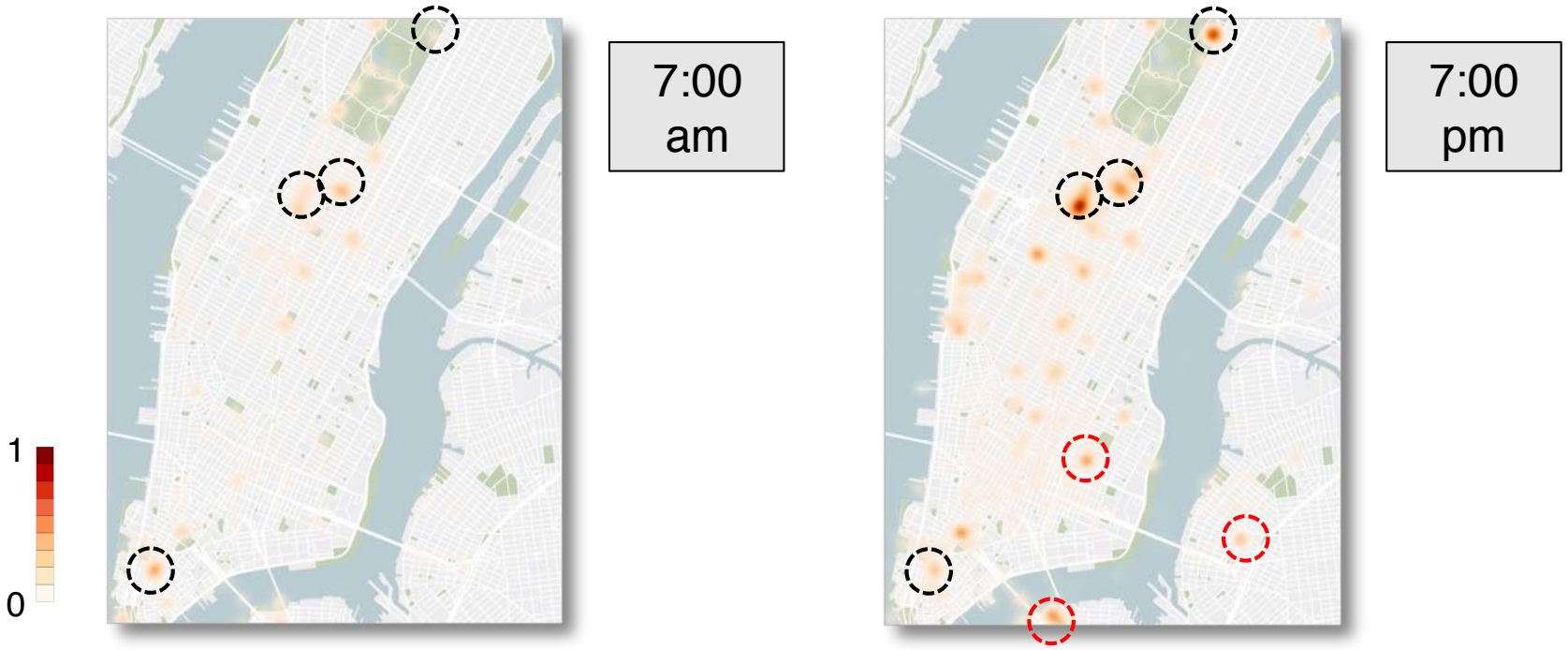
Urban Pulse

- Flickr activity in New York City



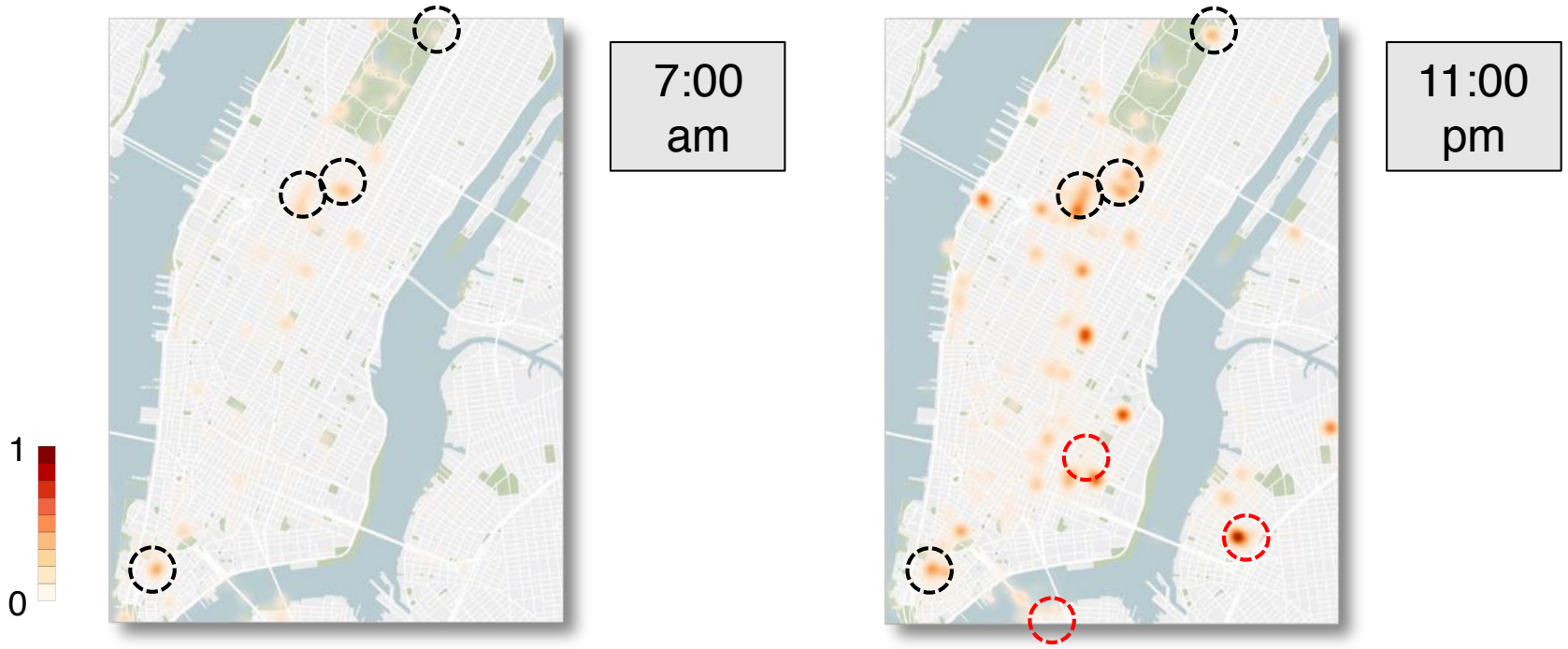
Urban Pulse

- Flickr activity in New York City



Urban Pulse

- Flickr activity in New York City



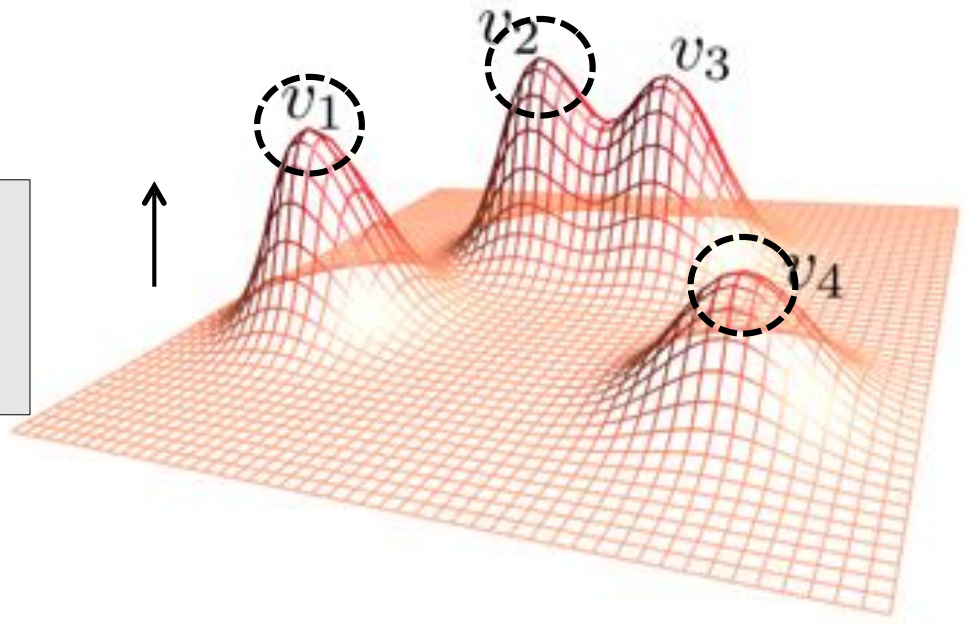
Urban Pulse: Desiderata

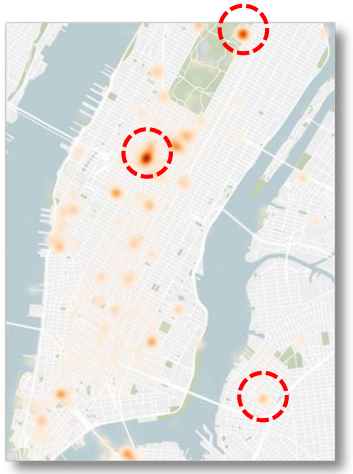
- Capture locations where the pulse is “interesting”
- Quantify the pulse
  - Track “activity”
- Temporal resolutions

1. Identify Locations
2. Quantify Pulse

Step 1: Identify Pulse Locations

Maxima

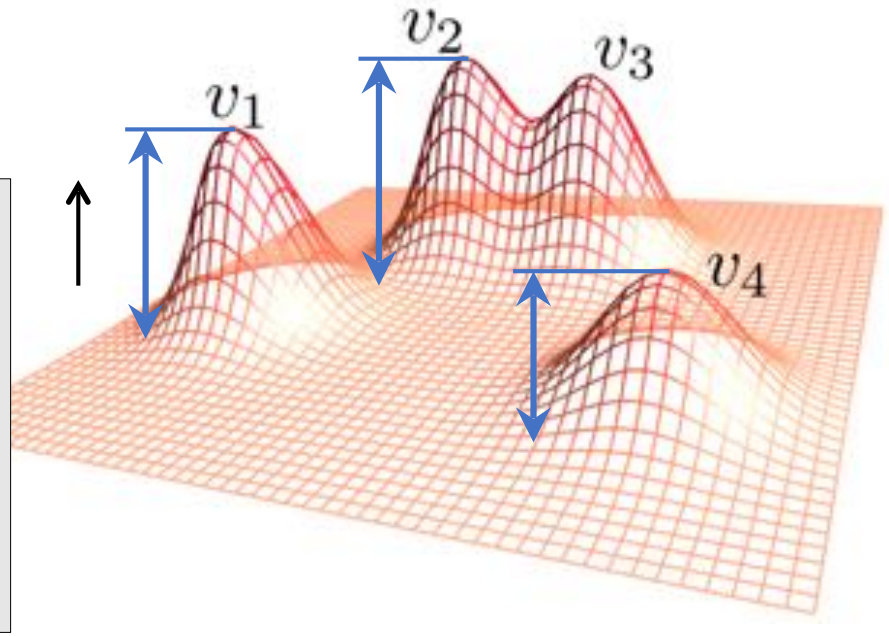


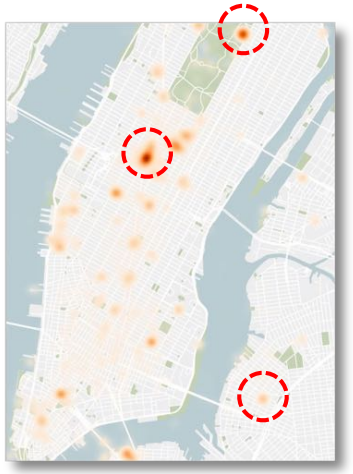


1. Identify Locations
2. Quantify Pulse

Step 1: Identify Pulse Locations

Topological Persistence





1. Identify Locations
2. Quantify Pulse

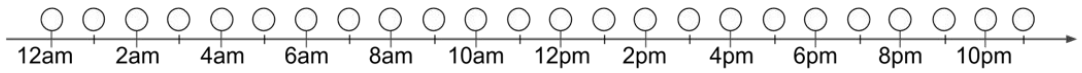
Step 1: Identify Pulse Locations

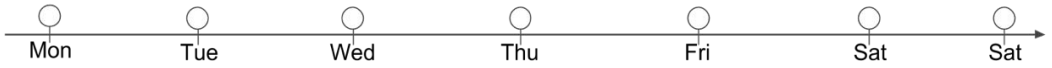
- Set of scalar functions over time
- Identify all maxima
- Location of **prominent** pulses
  - is a high persistent maxima in at least 1 time step

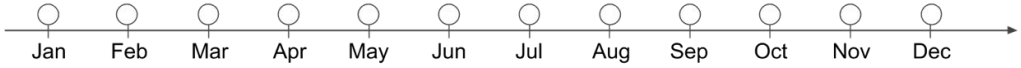
Handling Temporal Resolutions

- Assume functions are defined along 3 resolutions

Time of Day
Day of Week
Month of Year







1. Identify Locations
2. Quantify Pulse

Step 1: Identify Pulse Locations

- Set of scalar functions over time
- Identify all maxima
- Location of **prominent** pulses
  - is a high persistent maxima in at least 1 time step
  - is a high persistent maxima in at least 1 resolution

1. Identify Locations
2. Quantify Pulse

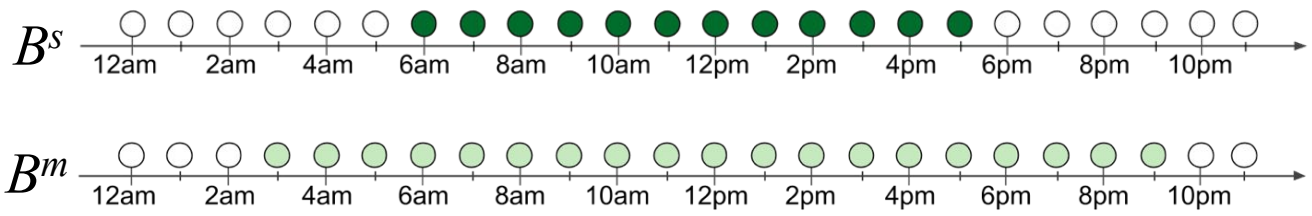
Step 2: Quantifying Pulse

- 3 **Beats** to quantify the pulse at each location
- Significant Beats
  - Is the location a high persistent maximum?

1. Identify Locations
2. Quantify Pulse

Step 2: Quantifying Pulse

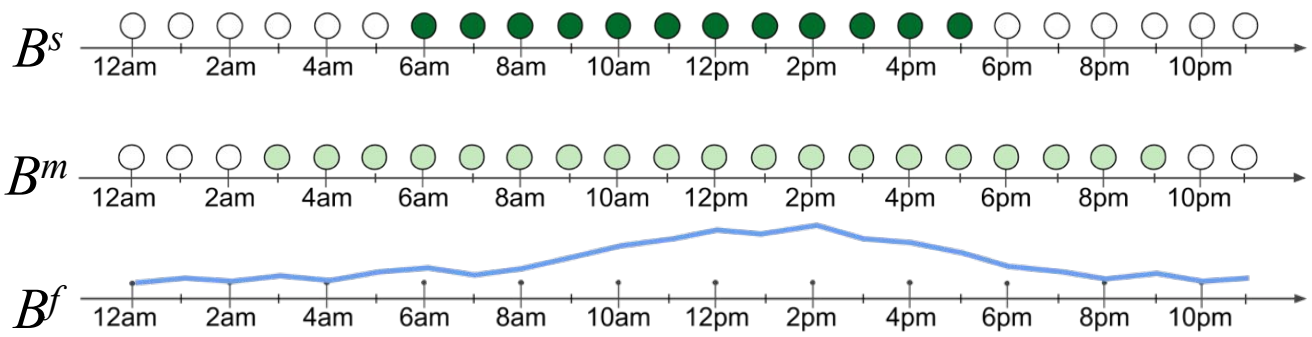
- 3 **Beats** to quantify the pulse at each location
- Maxima Beats
  - Is the location a maximum?



1. Identify Locations
2. Quantify Pulse

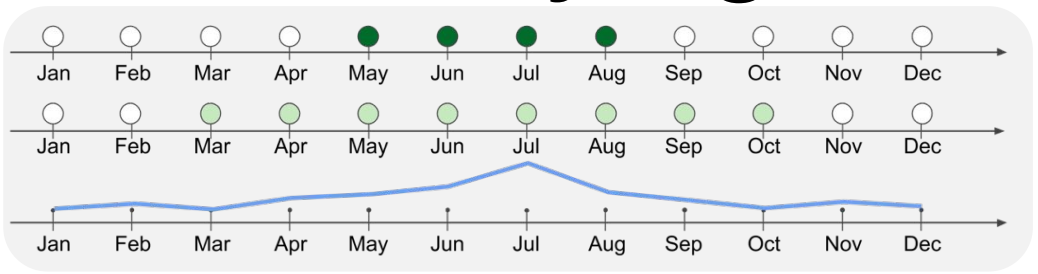
Step 2: Quantifying Pulse

- 3 **Beats** to quantify the pulse at each location
- Function Beats B^f
  - Variation of the function values

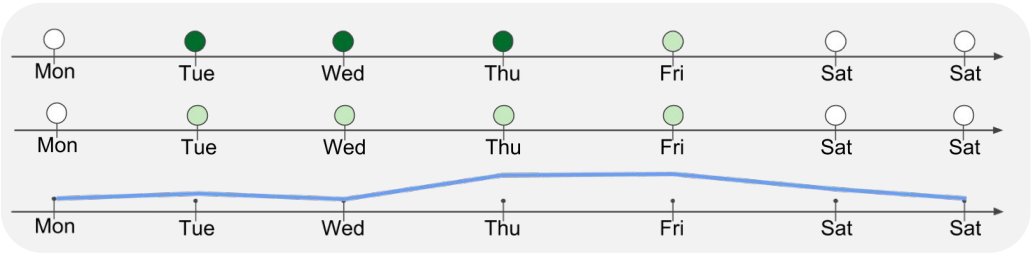


1. Identify Locations
2. Quantify Pulse

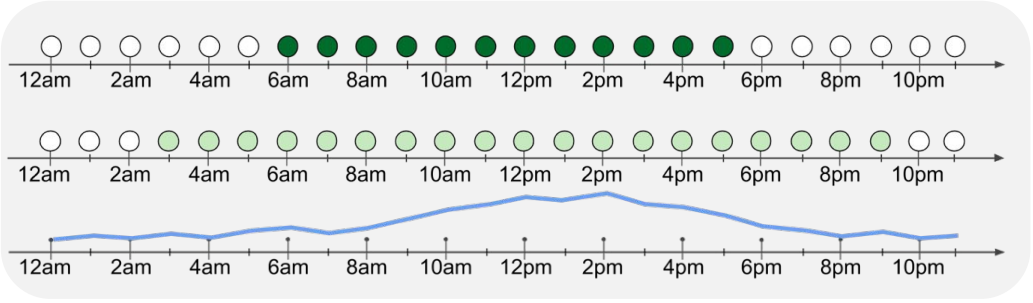
Step 2: Quantifying Pulse



Month of Year



Day of Week

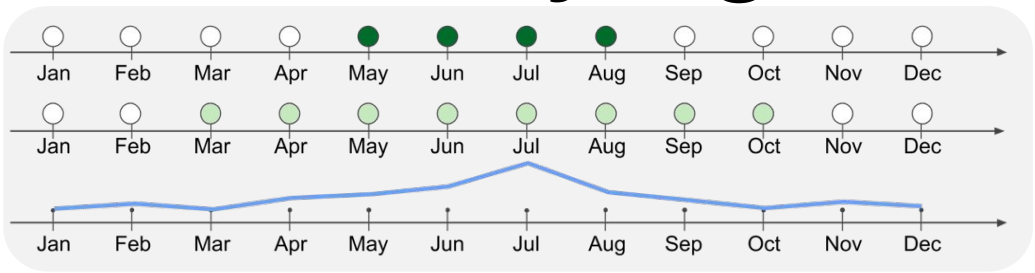


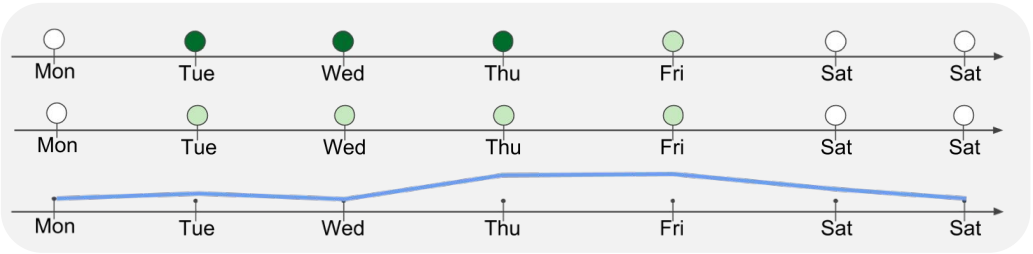
Time of Day

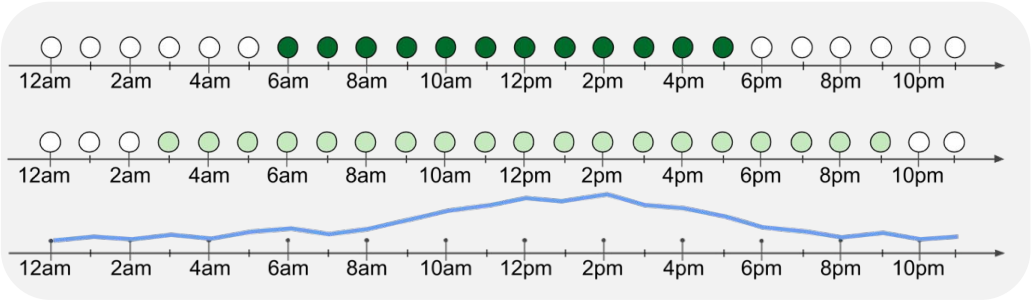
1. Identify Locations
2. Quantify Pulse

Step 2: Quantifying Pulse

B_1
B_2
B_3
B_4
B_5
B_6
B_7
B_8
B_9







Signature

Data Obliviousness

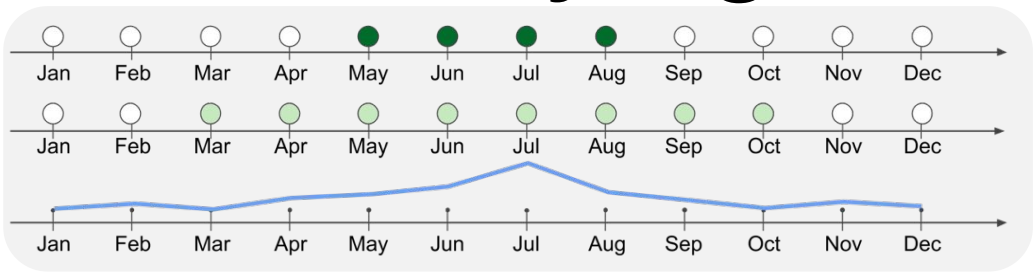
Rank

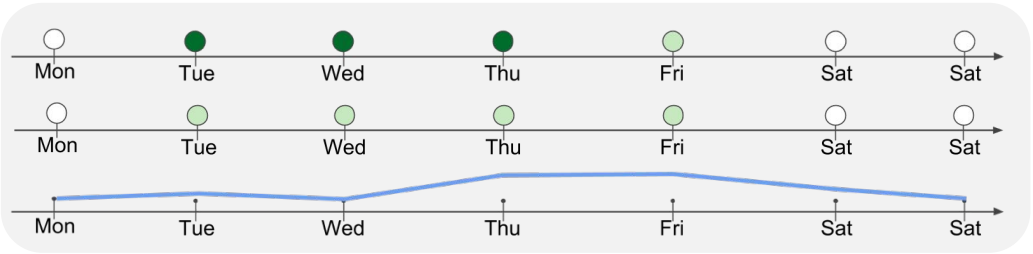
1. Identify Locations

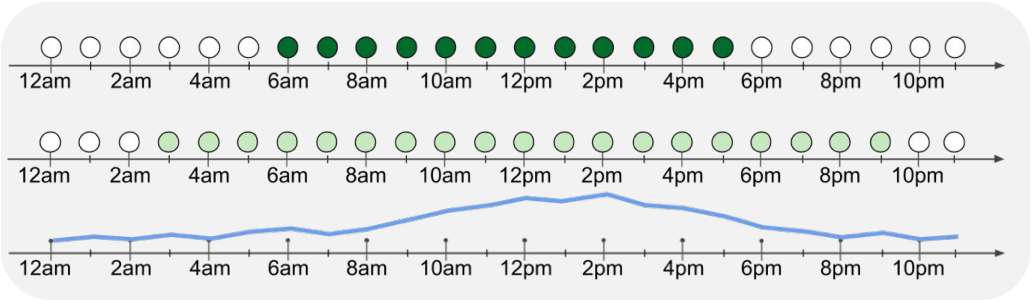
2. Quantify Pulse

Step 2: Quantifying Pulse

B_1
B_2
B_3
B_4
B_5
B_6
B_7
B_8
B_9





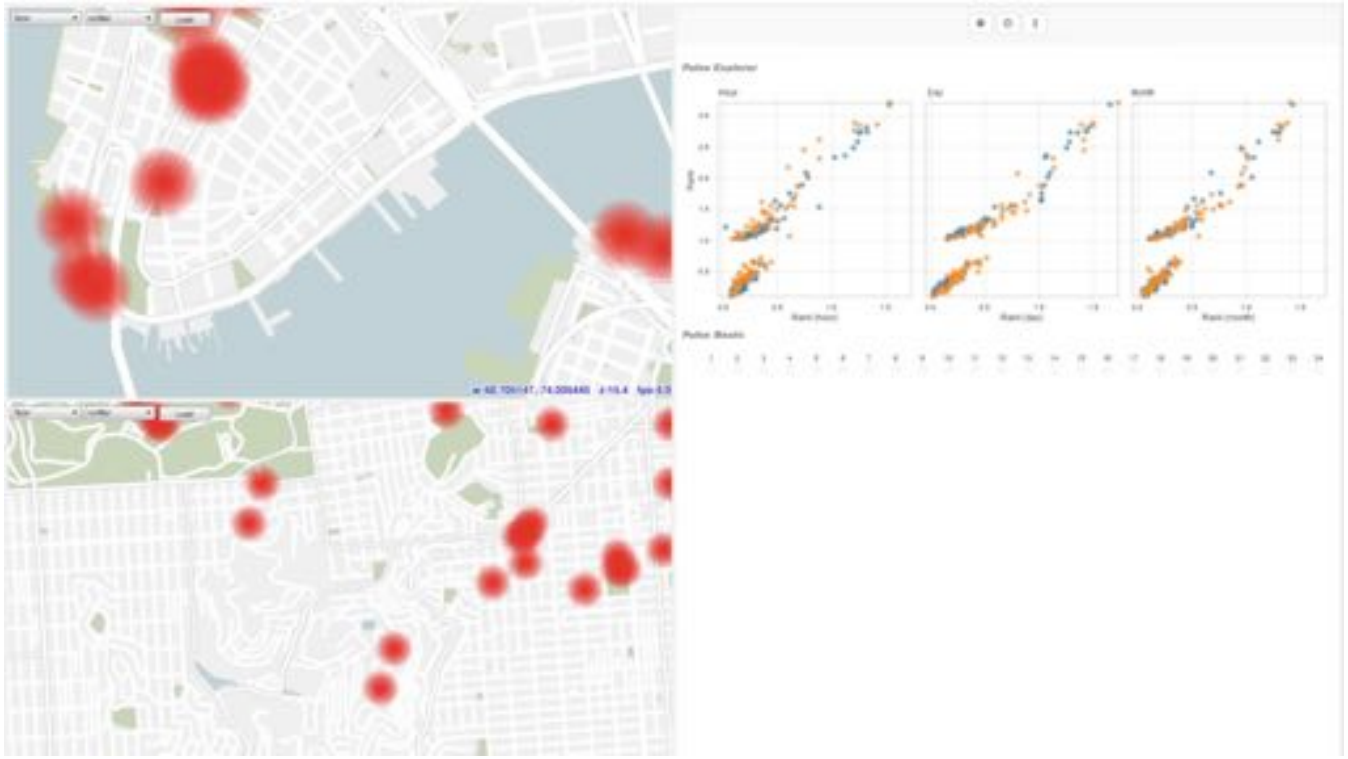


Signature

Data Oblivious

Compare

Urban Pulse Interface



Use Case

- Provided the interface to domain experts
- Architects from Kohn Pedersen Fox
  - Urban planning
- Human behavioral expert
  - Try to understand the cohabitation between cultural communities
  - Twitter as proxy for cultural communities

Use Case: Understanding Public Spaces

Rockefeller Center

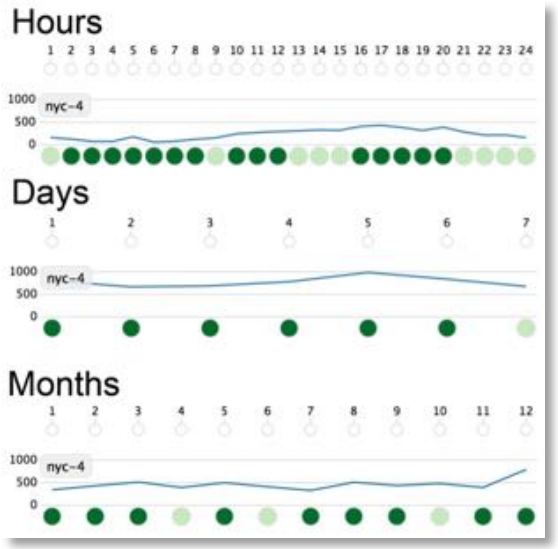
Union Square

Bryant Park

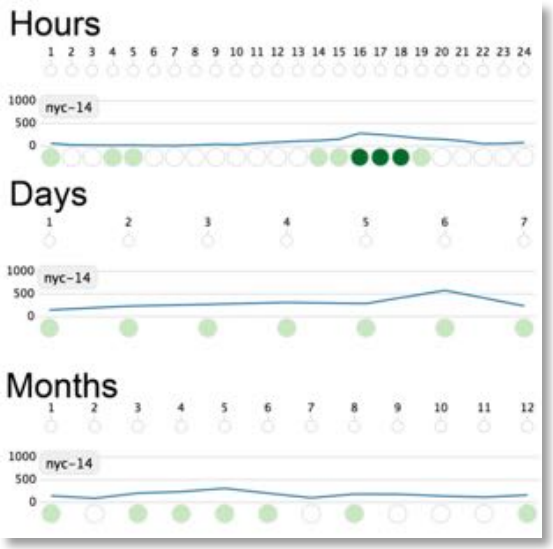
- Typically classified together as being similar

Use Case: Understanding Public Spaces

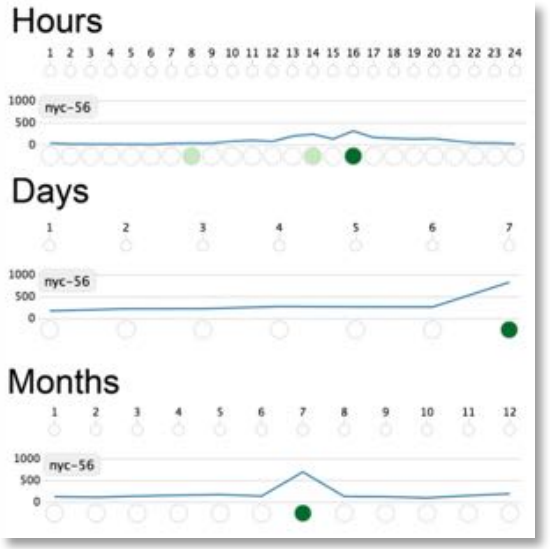
Rockefeller Center



Union Square

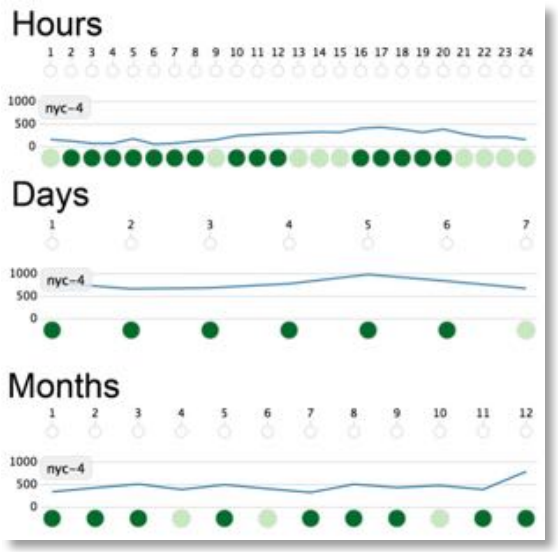


Bryant Park



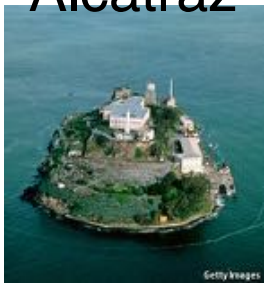
Use Case: Understanding Public Spaces

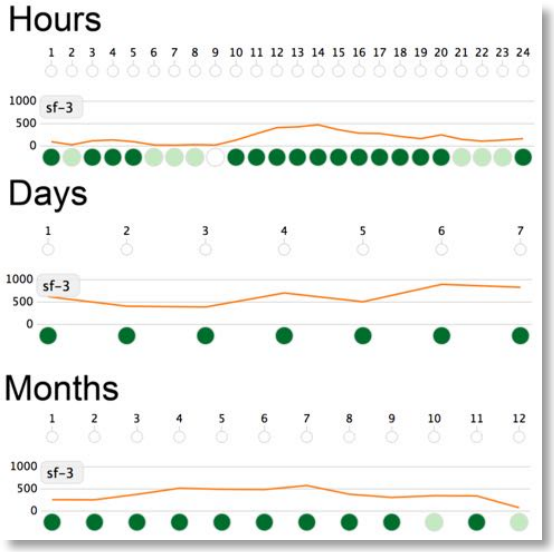
Rockefeller Center



San Francisco

Alcatraz





Conclusion

- Define the notion of **Urban Pulse**
  - Uses the topology of the data to characterize the activity within a city
- Signature for locations
  - Data oblivious
- Compare locations within cities
- Compare locations across cities
- Used by experts
  - Urban planning
  - Cohabitation of cultural communities

Urban Pulse:
Capturing the Rhythm of Cities

The Economist | World politics | Business & finance | Economics | Science & technology | Culture

Urban planning

Listen to the music of the traffic in the city

Places, like people, have pulses—if only you know how to measure them

Oct 22nd 2016 | From the print edition

Timekeeper | Like 221 | Tweet

THE Rockefeller Centre sprawls across 89,000 square metres of midtown

VIS 2015

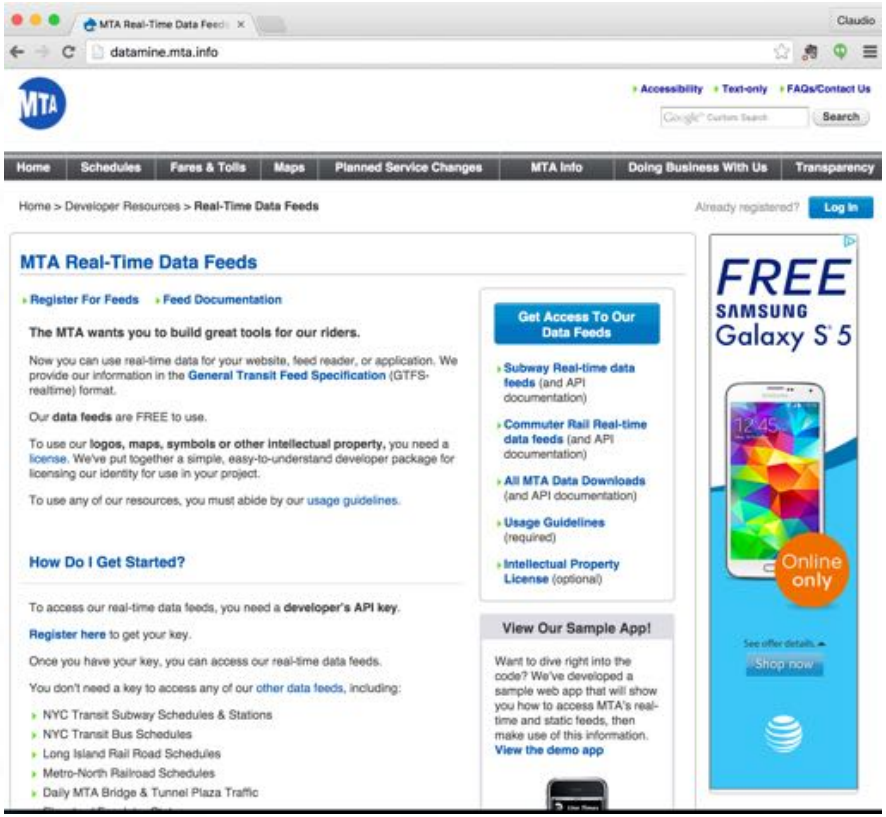
VAST • INFOVIS • SCIVIS

Visually Exploring Transportation Schedules

Cesar Palomo, Zhan Guo, Claudio Silva, and Juliana Freire
New York University

New York City MTA and the GTFS feed

- Data standard: General Transit Feed Specification (GTFS)
- Every 30 second it is possible to get a “status” report on the state of a large portion of the MTA subway and bus infrastructure
- Open API
  - short-term availability; if you do not save it, it is gone



NYU TANDON SCHOOL OF ENGINEERING

VIS 2015

Early work: MTA viewer

- Fall 2012; before the GFTS feeds were active
- Initial ideas centered around a ma… rep… a…
- Showing "delay":

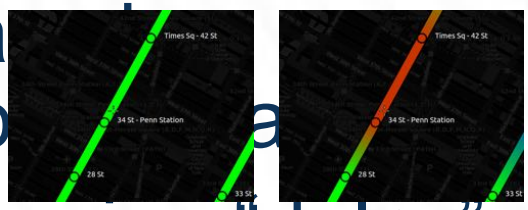


Fig. 8. Subway delay near Time Square at different time on Sep 5 2012 (a) From 1 pm to 2 pm no delay is observed; (b) Significant delay is perceived from 2 pm to 3 pm.

MTA Viewer: a Real-time Visualization System for NYC Subway Lines

Bowen Yu and Cesar Palomo

Fig. 1. Overview of the interface of MTA Viewer. In the middle is the map rendering of the subway system. The charts for data aggregation and comparison by stop and time are on the left, while the control panel of the system is on the right.

NYU TANDON SCHOOL OF ENGINEERING

VIS 2015

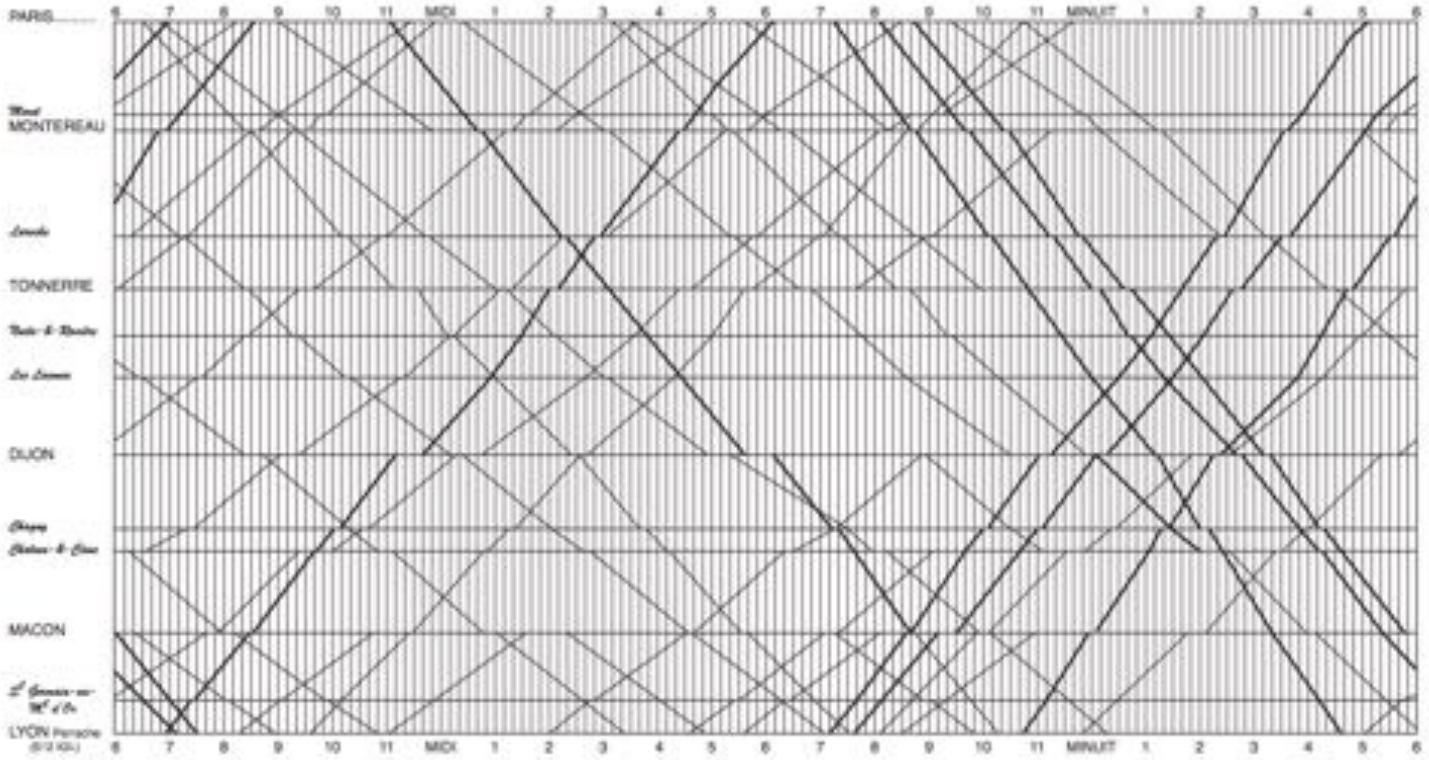
What questions were we interested in?

- T1: Compare planned timetables against real service
- T2: Characterize speed profile at different route segments
- T3: Assess delay, wait time and reliability at the station level
- T4: Study the interplay of different attributes

NYU TANDON SCHOOL OF ENGINEERING

VIS 2015

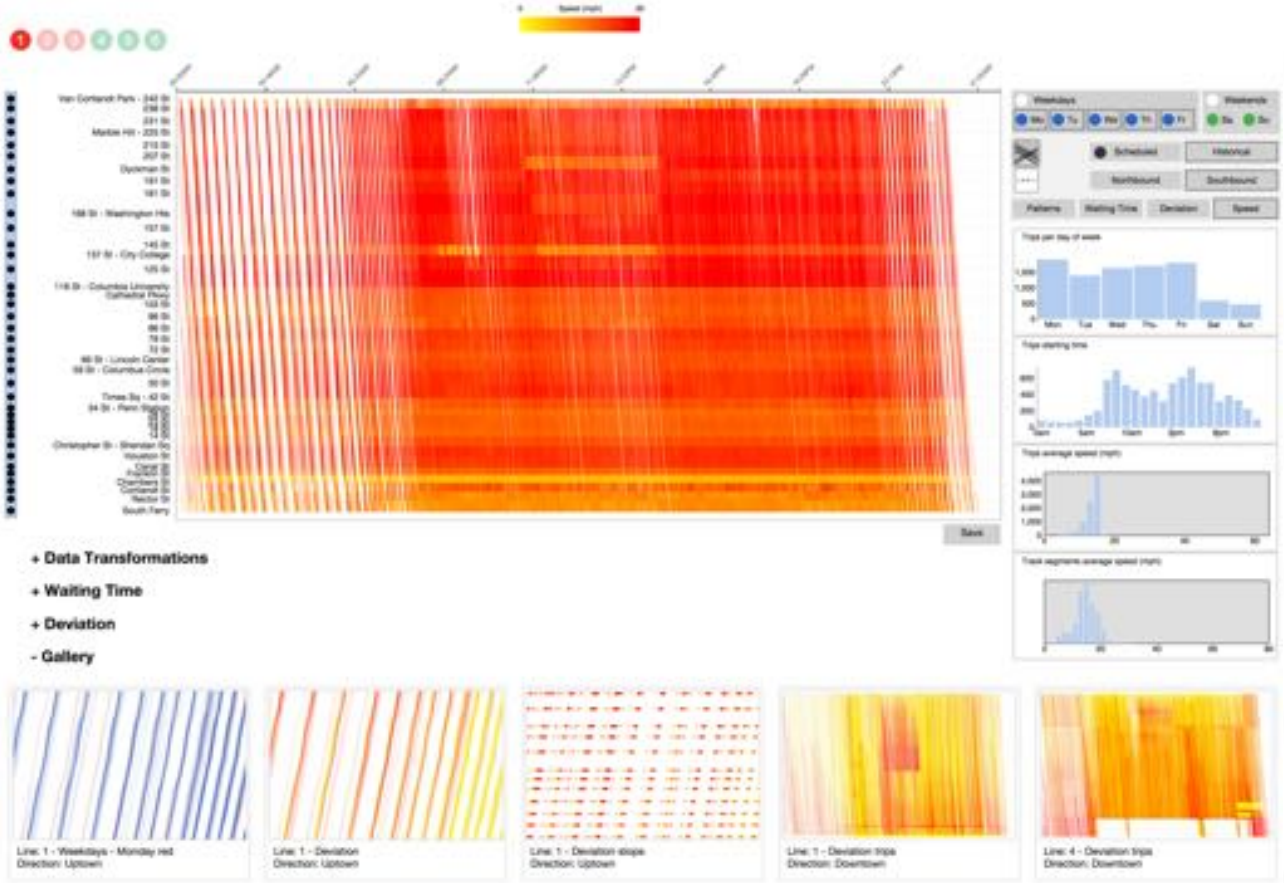
Overhaul: Building on Marey Graph Design



NYU TANDON SCHOOL OF ENGINEERING

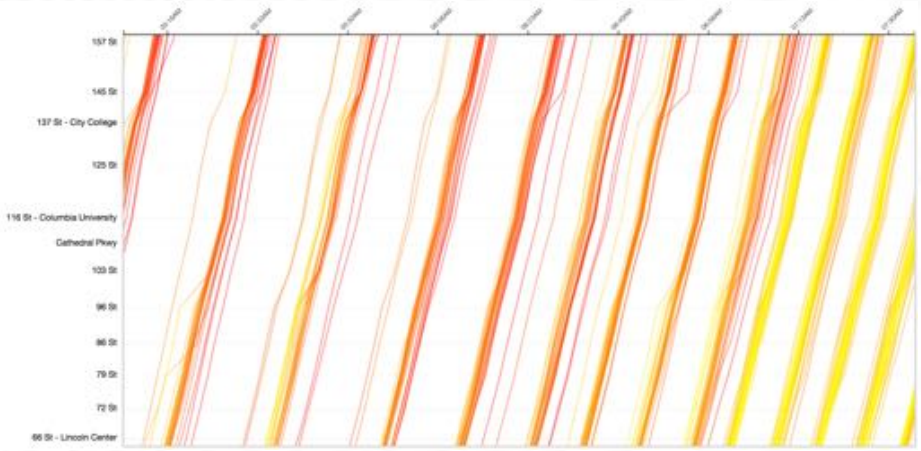
VIS 2015

2nd Try: Building on Marey Graph Design

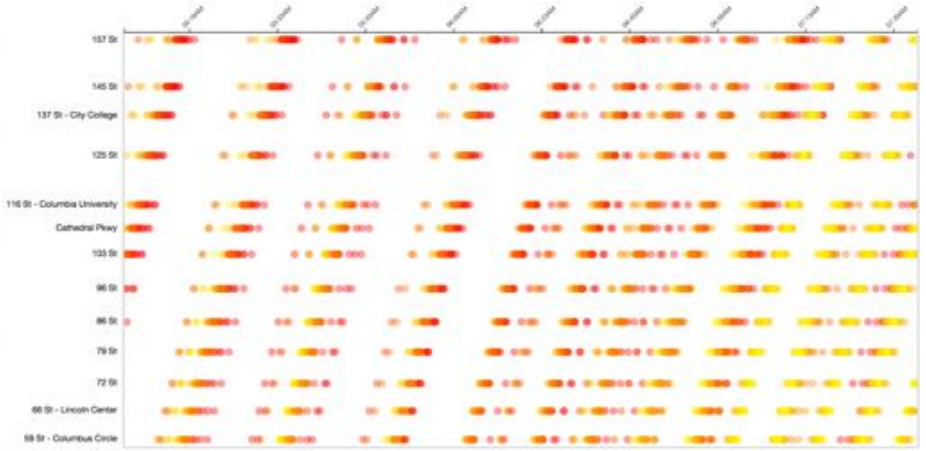


- web-based tool
- uses D3.js and WebGL
- cluttering was a problem

Visual Elements



(b) Trips Explorer, delay view

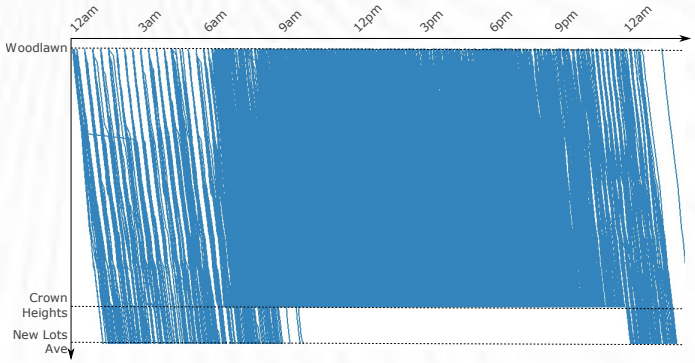


(d) Stops Explorer, delay view

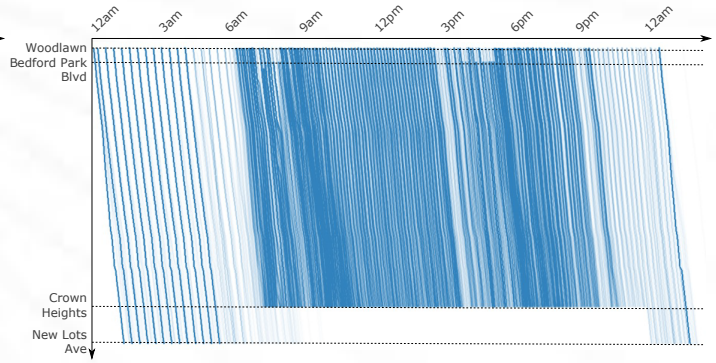
NYU TANDON SCHOOL OF ENGINEERING

VIS 2015

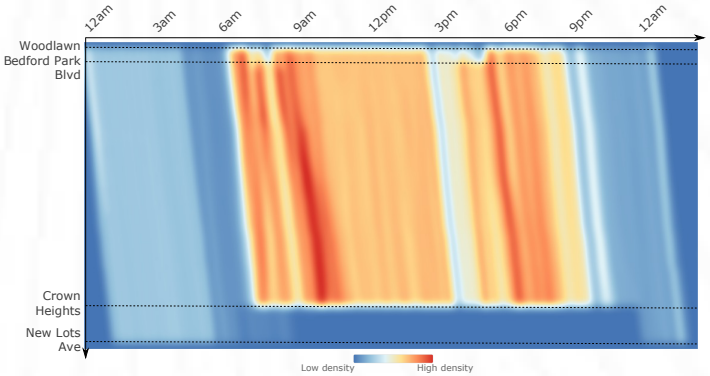
TR-EX: adding user-controlled KDE



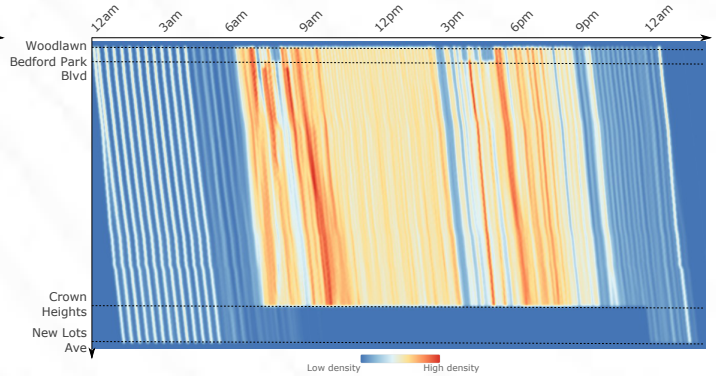
(a) Original Marey graph: each trip is represented by a line with full opacity



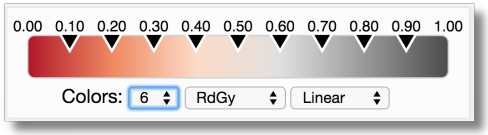
(b) Additive blending of trips with constant transparency



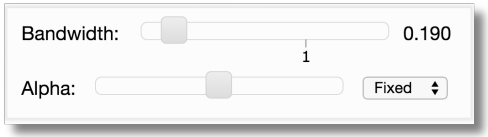
(c) Using KDE for lines: wide bandwidth reveals coarser details



(d) Using of KDE for lines: narrow bandwidth reveals finer details



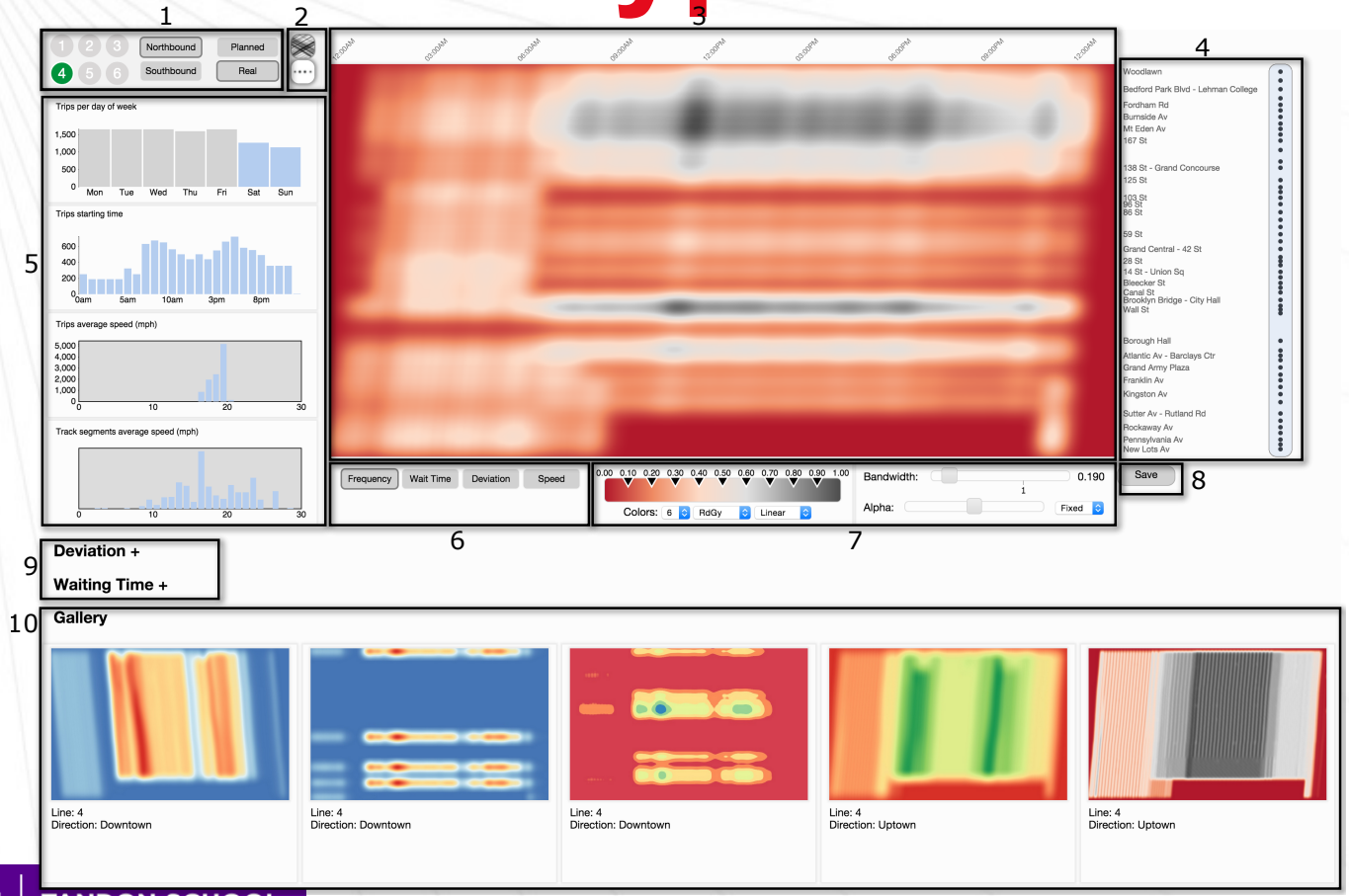
(a) Editable color scale



NYU TANDON SCHOOL OF ENGINEERING

VIS 2015

TR-EX Prototype



NYU TANDON SCHOOL OF ENGINEERING

VIS 2015
VAST • INFOVIS • SCIVIS

Speed Visualization

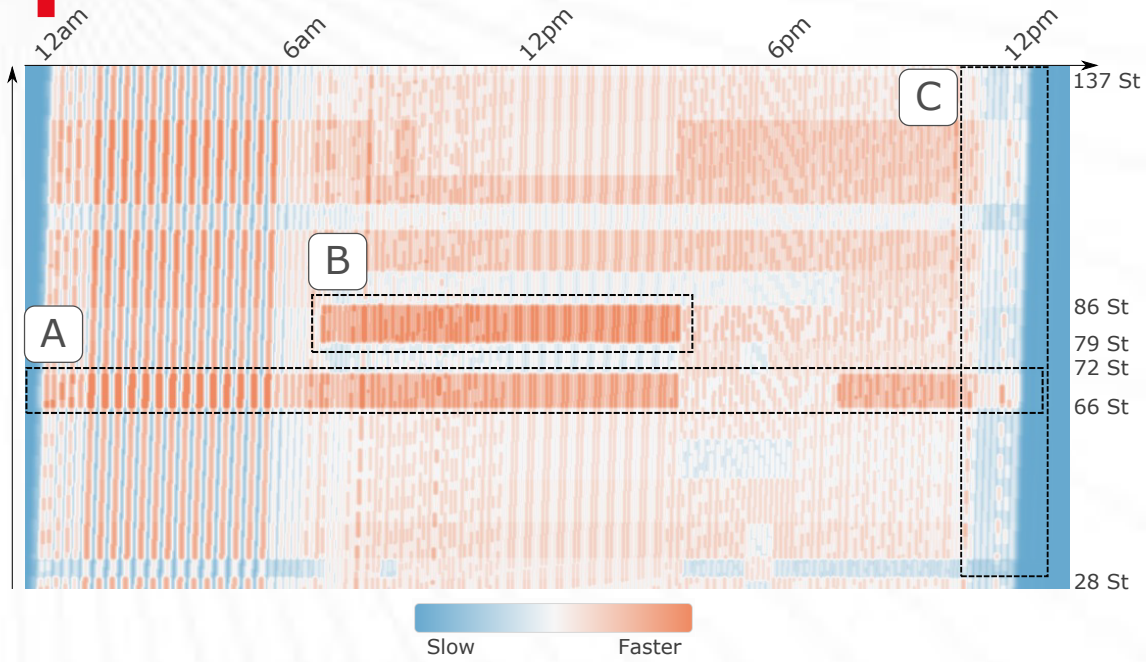


Fig. 5: Speed visualization with Trips Explorer for uptown trips in subway line 1. Region A shows that the speed for trips between 66th St and 72th St stations is mostly constant, except during peak hours, when vehicles run slower. Localized regions depict vehicles running faster than usual (region B) or slower during late nights (region C).

Wait Time

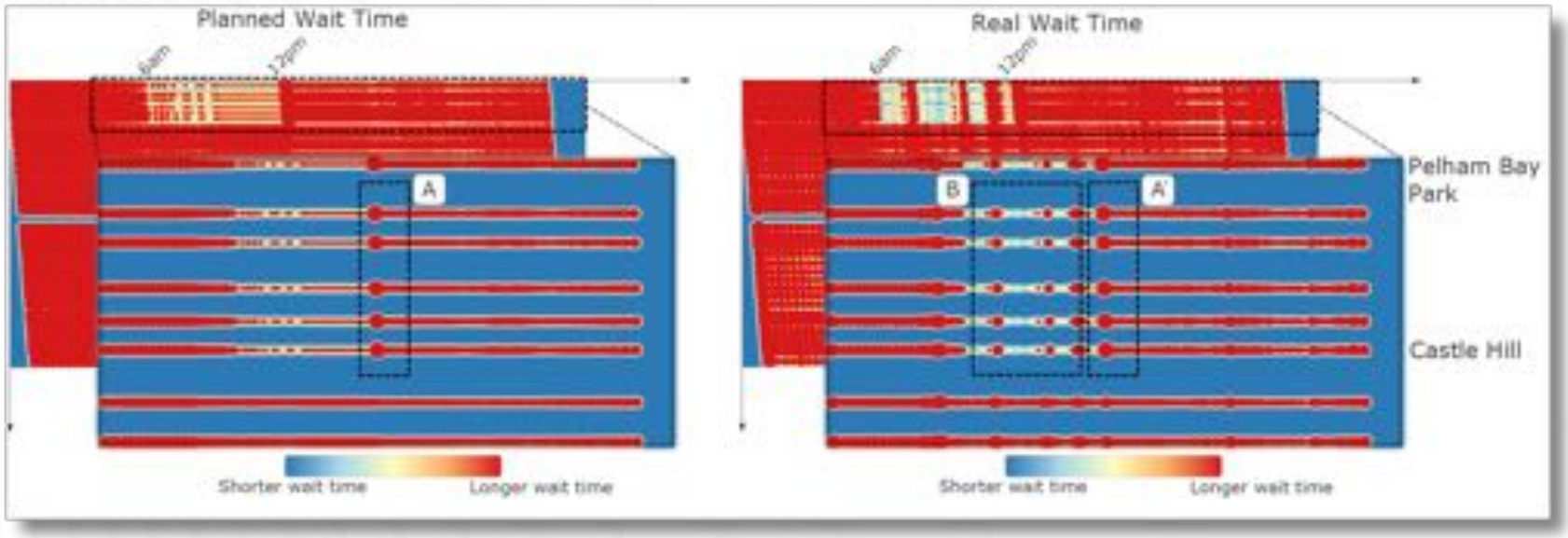


Fig. 9: Comparing planned (left) and observed (right) wait times at stations for subway line 6 toward Brooklyn Bridge - City Hall on weekdays. The visualizations show a substantial divergence between the planned and actual wait times in the highlighted region (top). Zooming into this region, we can see that in the planned service, stops are equally spaced from each other, while in the real service trips are concentrated in three main clusters: time: around 8am, 10am and 11am. This leads to a considerable increase in observed wait time at later times, as shown in region *B* of the real service. The behavior of the real service stabilizes after region *A'*, where it is similar to the planned schedule.

Thank you!

csilva@nyu.edu