CARTOGRAPHY IN A GENERAL PURPOSE RISK ANALYSIS SYSTEM SUPPORTING BLACK SWAN THEORY

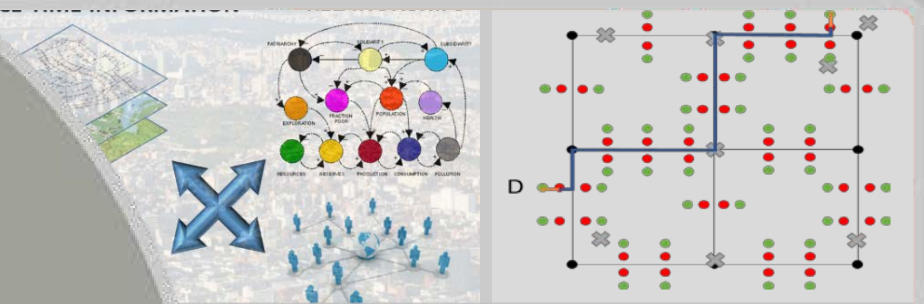
FRAMEWORK FOR INTEGRATING COMPLEX UNCERTAIN SYSTEMS (FICUS)

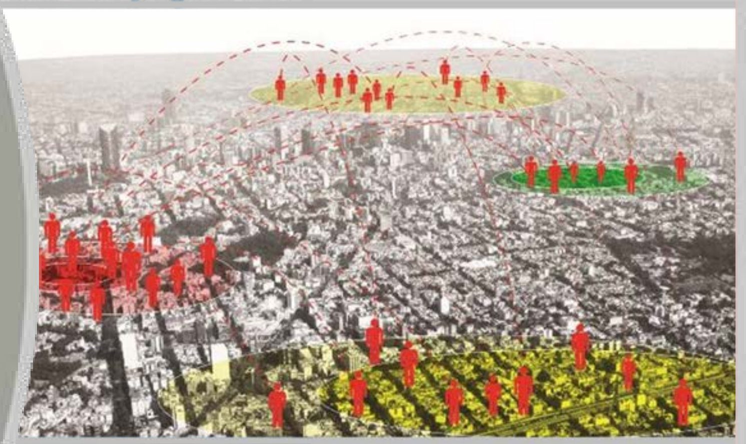
Dr. Charles Ehlschlaeger
US Army Engineer Research and Development Center, Champaign, IL

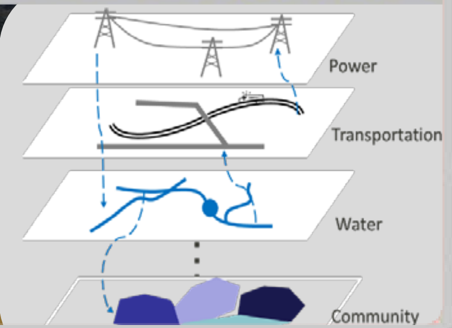
With support from Mr. Jeffrey Burkhalter, Dr. Olaf David, Dr. James Westervelt, Dr. Dawn Morrison, Dr. Yanfeng Ouyang, Dr. James Ross, Dr. Mazdak Arabi, Mr. David Patterson, Ms. Natalie Myers, Ms. Carey Baxter, Ms. Elizabeth Bastian, Dr. Yizhao Gao, Mr. Liqun Lu, Mr. Francesco Serafin, Mr. Kyle Traff, Mr. Antoine Petit, Mr. Jiang Zhoutong

26 September 2018

DISTRIBUTION STATEMENT A: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED







US Army Corps of Engineers

Colorado State University

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

ERDC
Engineer Research and Development Center

Innovative solutions for a safer, better world

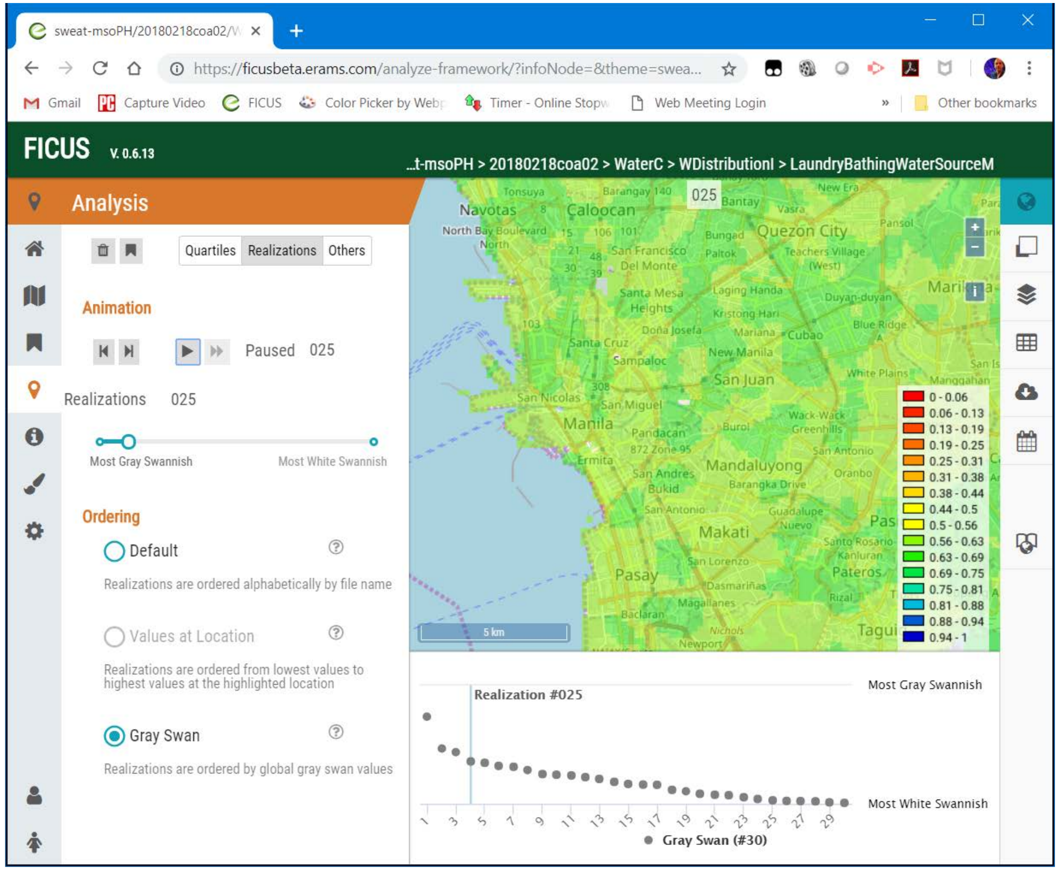
Abstract

Risk analyses for geospatial activities are constantly undermined by plans that have been created by multiple organizations for unfamiliar geographies, involving imprecise data model and uncertain model parameters. Moreover, those responsible often have little capacity to communicate about complex geo-temporal patterns with fellow information creators, analysts, and planners across the globe. Over the past decade, the University of Illinois at Urbana-Champaign (UIUC), Colorado State University (CSU), and the US Army Corp of Engineers' Engineer Research & Development Center (ERDC) has been collaborating to construct a neighborhood-scale social, infrastructural, and environmental modeling system that quantifies the uncertainty of all input data, propagates that uncertainty through tightly coupled space-time models, and visually presents uncertainty information and intuitive insights to planners and analysts. The Framework Incorporating Complex Uncertain Systems (FICUS) is a computational framework that supports all the functions of a general purpose geographic and temporal analysis system with a focus on risk analysis. FICUS minimizes the Uncertain Geographic Context Problem (UGCoP), propagates uncertainty using the Object Modeling System as its computational framework and contains interdependent urban infrastructure network models to forecast network failures. It can tightly couple models from R, Python, and NetLogo as well as most popular programming languages, and it itself is free and open source.

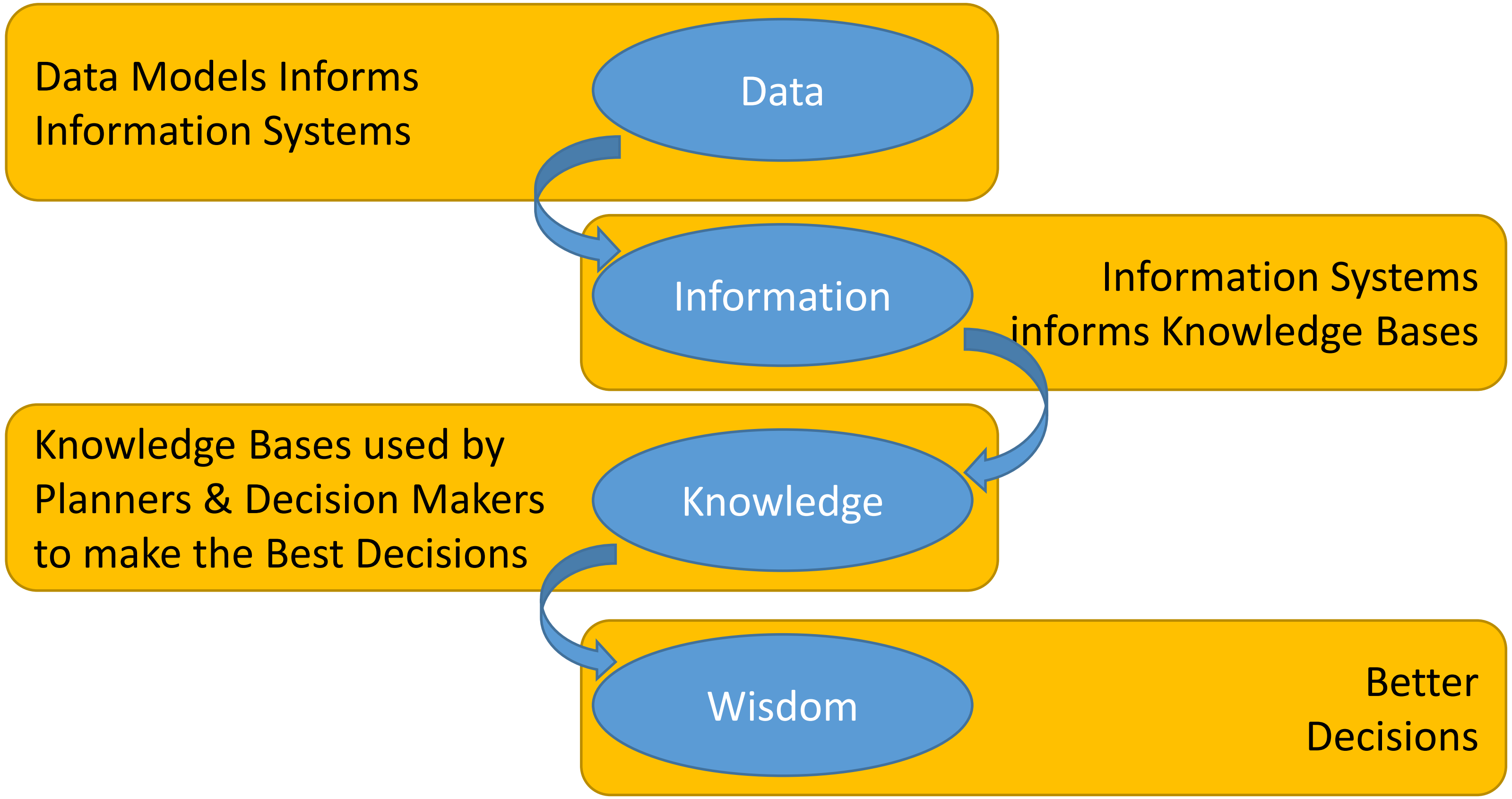
This presentation will discuss the theoretical and practical benefits of using an uncertainty quantifying, uncertainty propagating, and uncertainty visualizing geographic information system for risk analysis with a case study in the Philippines. We will focus our discussions on the techniques that minimize or can even eliminate UGCoP, calibration and validation in a multi-verse risk analysis paradigm, and cognitive issues of understanding risk using Black Swan Theory.

FRAMEWORK FOR INTEGRATING COMPLEX Uncertain SYSTEMS (FICUS)

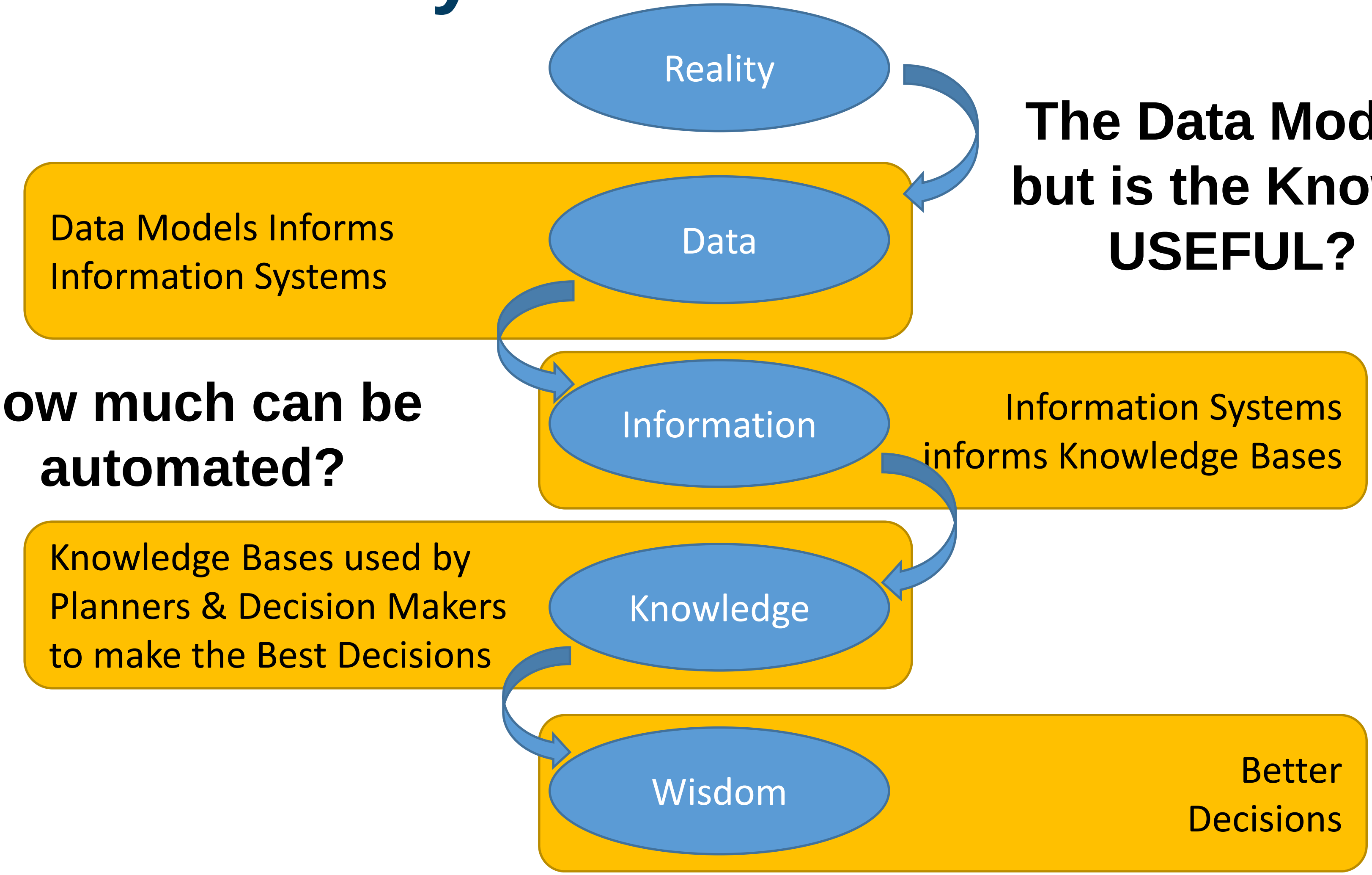
- The Framework Incorporating Complex Uncertain Systems (FICUS) is a computational framework, currently in development at ERDC, UIUC, and CSU, that supports all the functions of a general purpose geographic and temporal analysis system with a focus on risk analysis.



The Data to Wisdom Continuum



The Reality to Decision Continuum

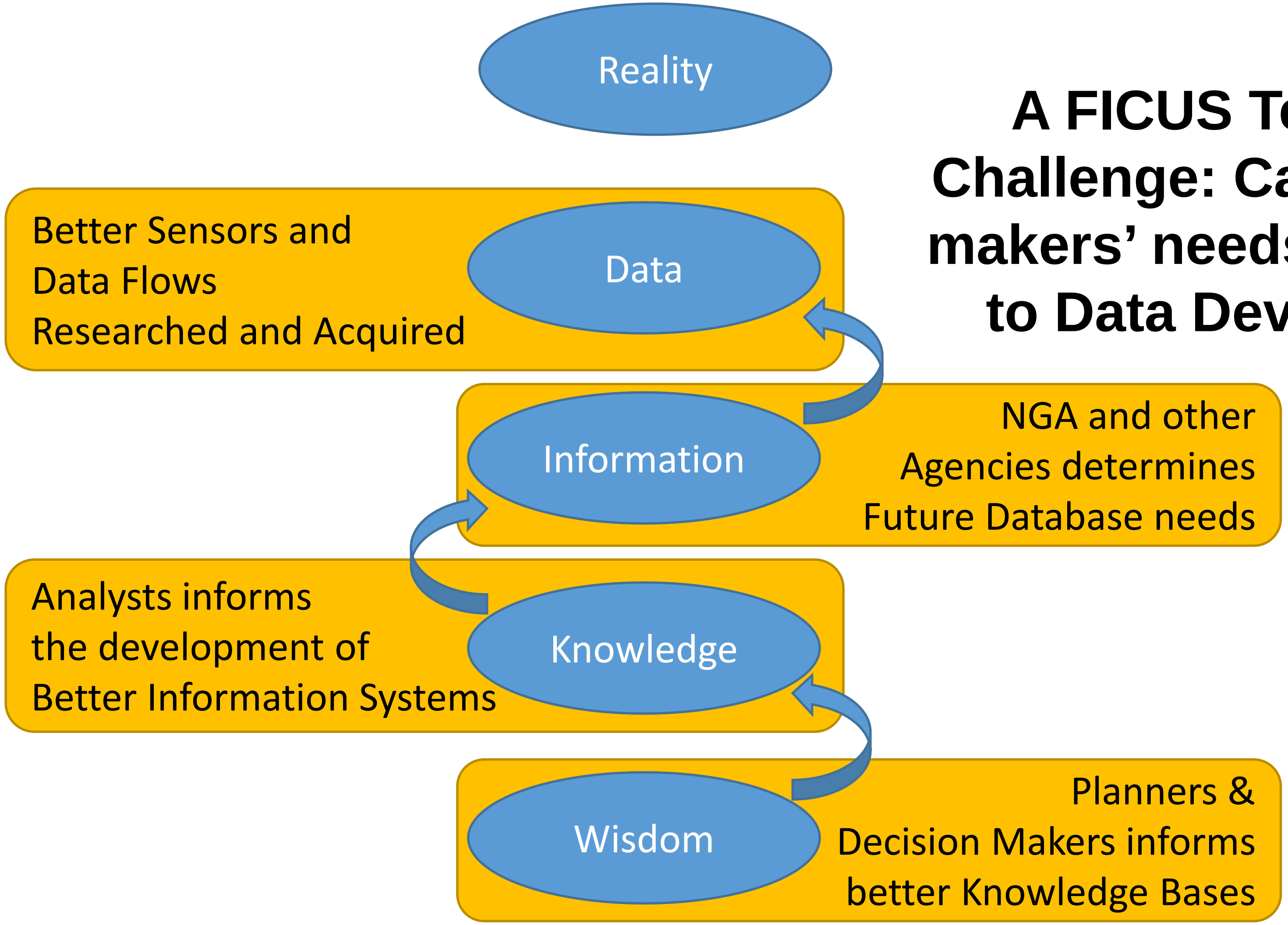


The Data Model IS wrong, but is the Knowledge Base USEFUL? (George Box)

How much can be automated?

A FICUS Theoretical Challenge: Communicating uncertainty to Decision Makers

Sensor, Info System, or Analysis Improvement



A FICUS Technical Challenge: Can decision makers' needs be known to Data Developers?

Framework for Incorporating Complex Uncertainty Systems (FICUS)

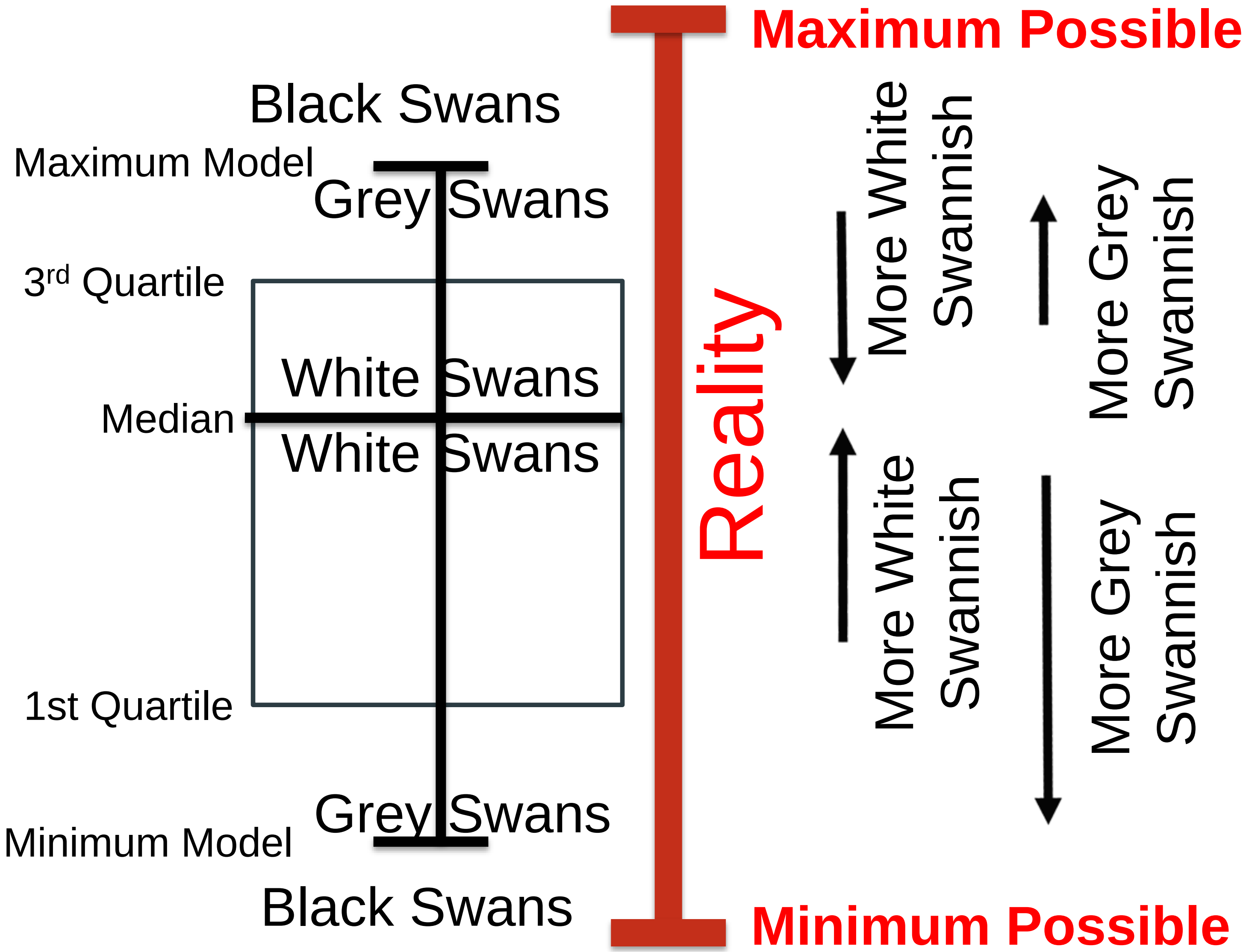
- ▶ **Uncertainty Quantification (UQ) of Data Model**
  - Box Plot Data Mapping of all variables for Black Swan Theory analysis
  - Geotemporal map multiverse of plausible realities
- ▶ **Uncertainty Propagation (UP) through Applications**
- ▶ **Uncertainty Visualization (UV) of Decision Supporting Data/Maps Connected to Inputs**

All three needed for Risk Analysis in a distributed Reality to Decision Continuum

Nassim Taleb’s Black Swan Theory, Plus Grey Swans

- Black Swan Events: **Unpredictable highly improbable events**
- White Swan Events: **Expected highly probable events**
- Grey Swan Events: Predictable highly probable event… (Walker 2010: earliest we’ve found), others argue GSEs are **predictable less probable events**, with the later definition currently most often used
- The `Grey Swannish’ of an event: The more unlikely a predictable event is, the more Grey Swannish that event is (My definition, Google wasn’t able to find another use of the term Grey Swanish)

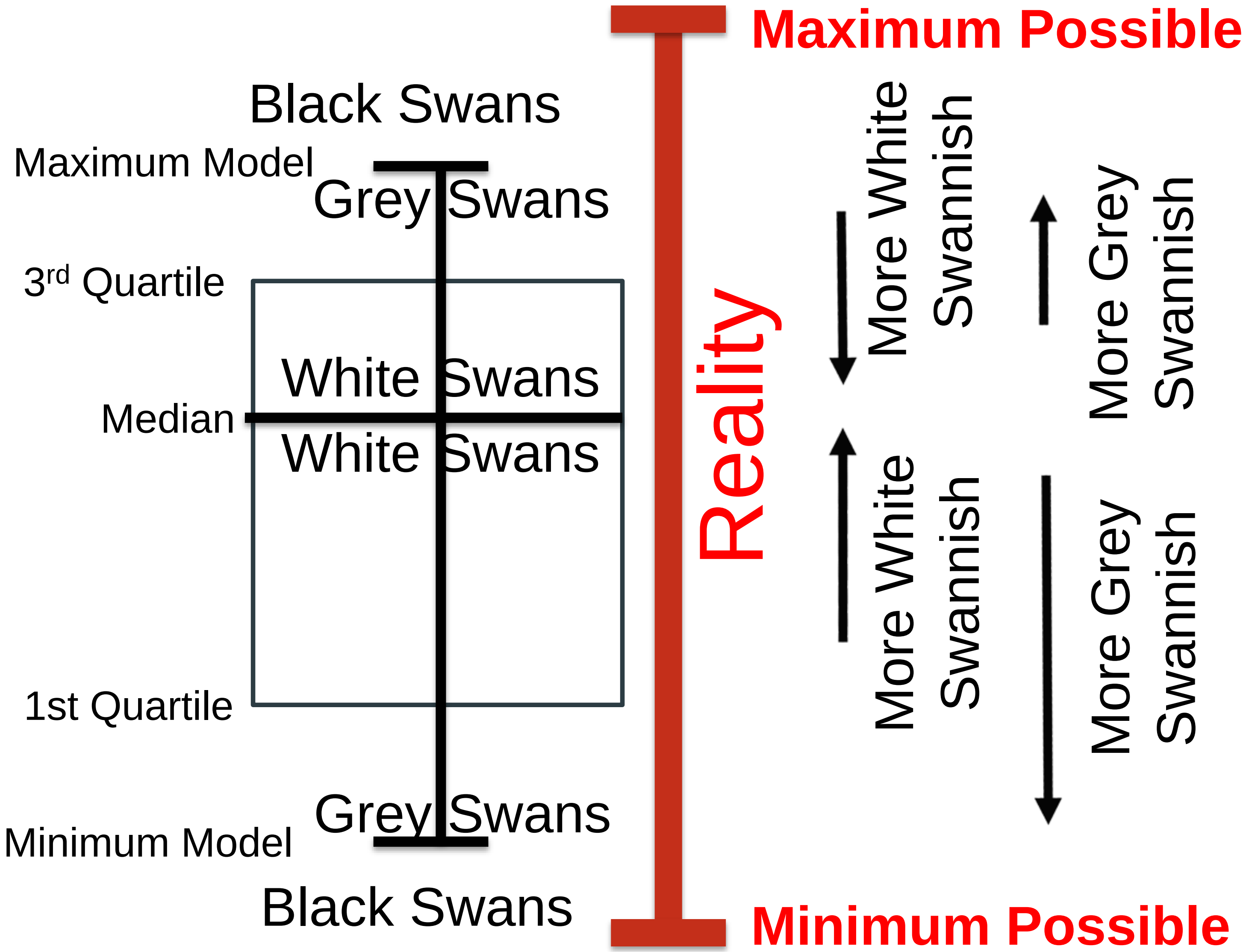
UQ: Black Swan Theory Data Model



Data model box plot provides error estimates:

All data should be represented as quantiles, with quartiles providing an elegant way to measure wrongness, bias, data model arrogance, and data model meekness

UQ: Wrong Black Swan Theory Data Model

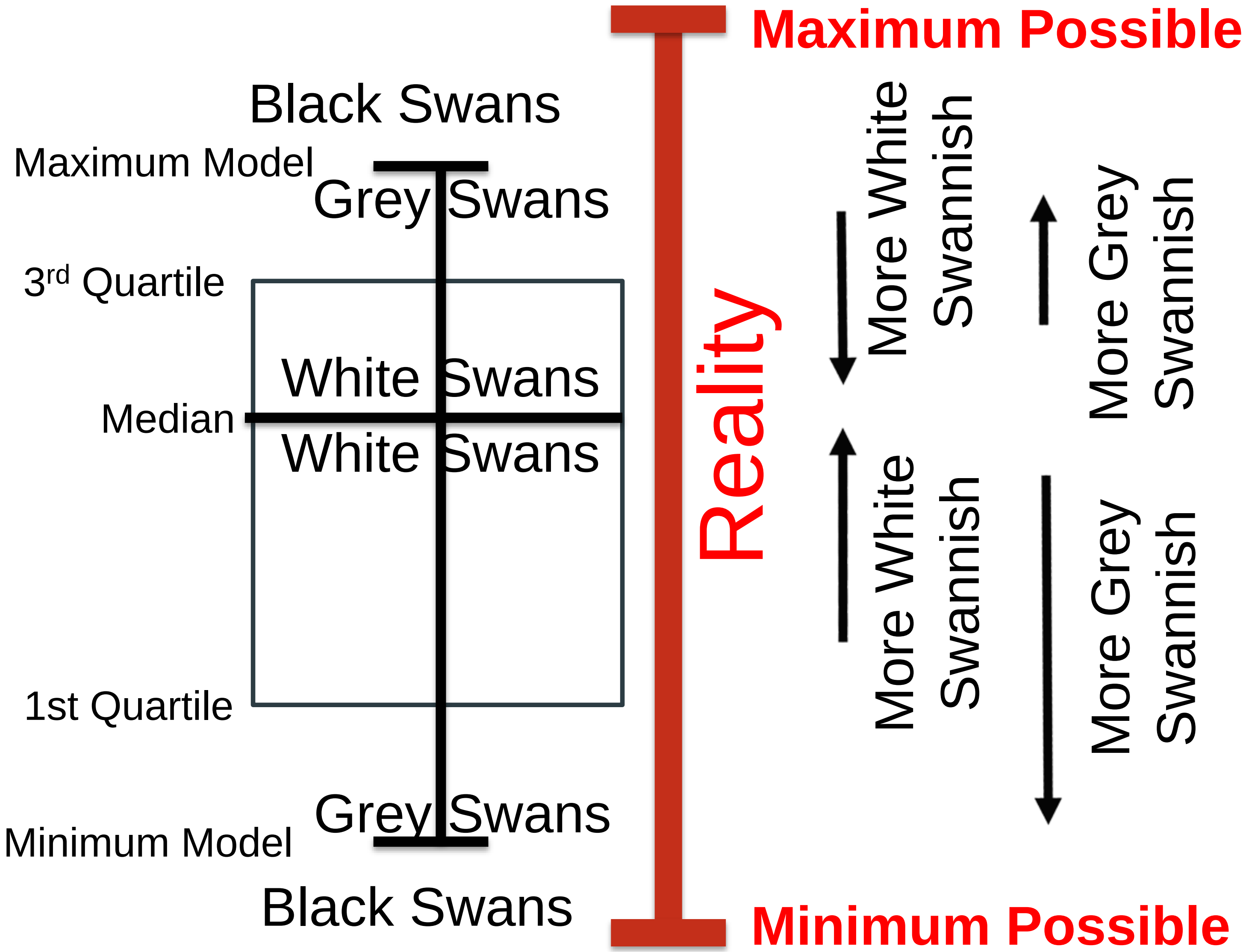


Wrong Model: When ANY ground truth is higher than maximum or lower than minimum

Must be adjusted or replaced immediately

Time to eliminate the George Box quote “All models are wrong but some are useful” as an excuse for wrong models trying to represent reality

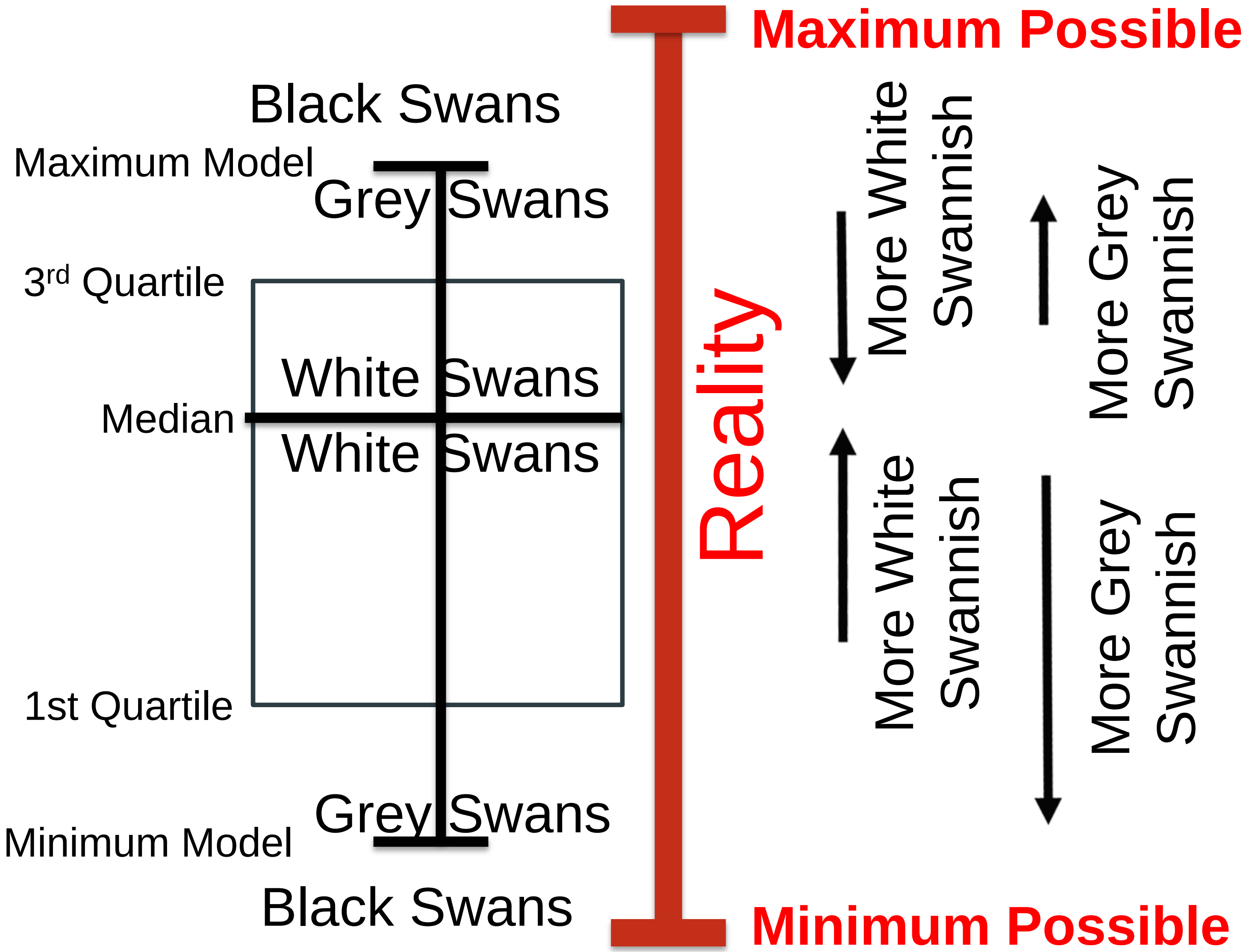
UQ: Biased Black Swan Theory Data Model



Biased Model: When statistical tests of ground truth are confident that median value (or quartiles) are biased

Must be adjusted or replaced immediately

UQ: Arrogant or Meek Black Swan Theory Data Model

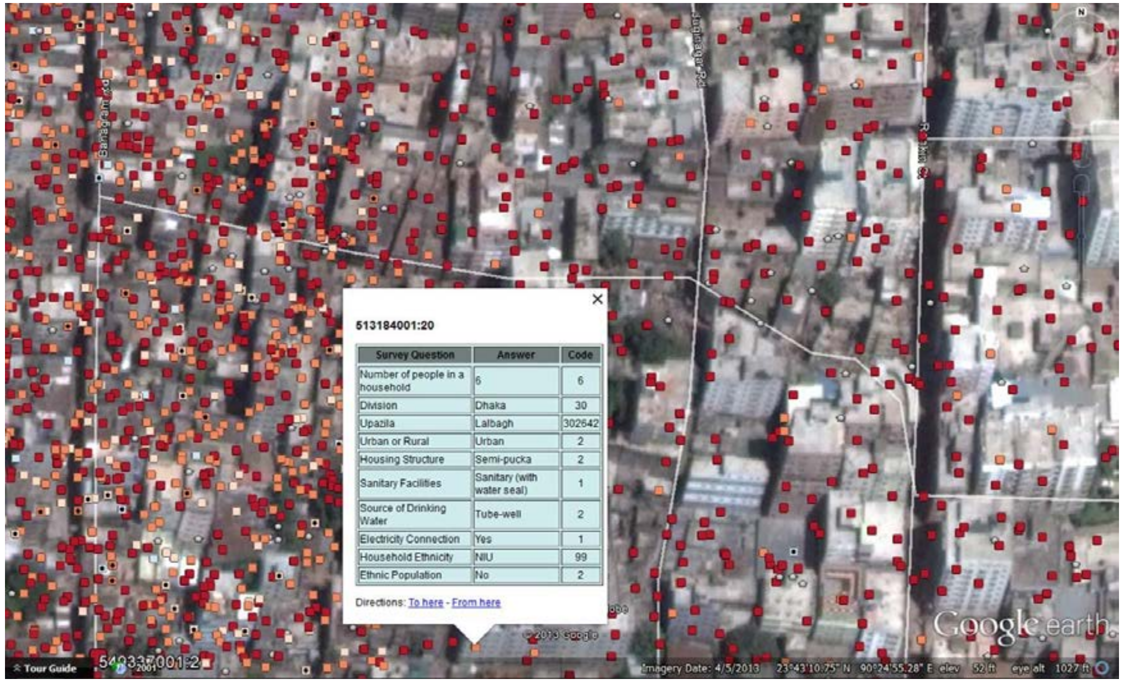


Arrogant or Meek: When statistical tests of ground truth are confident that too many or too few data samples are outside of Inter-Quartile Range

Tools in Data Model should adjust parameters to eliminate model arrogance or meekness

UQ: Population Attributes at Atomic Scale

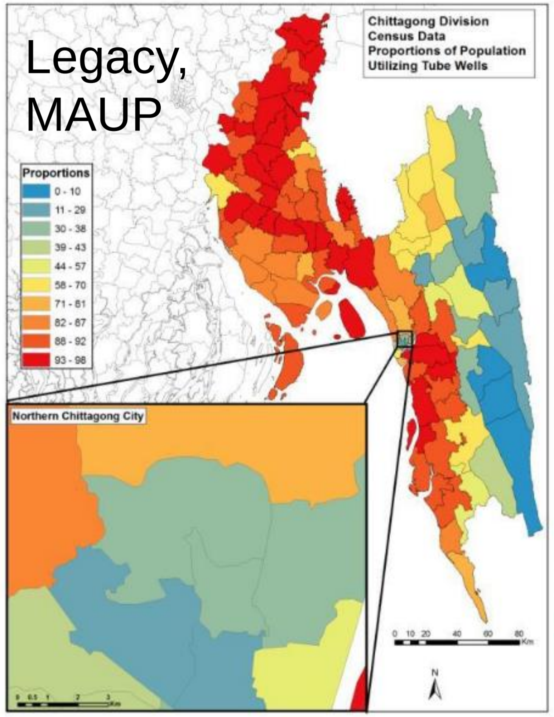
- Multiple realizations (multi-verse) of simulated geotemporal located populations of individual households based on microdata or survey cases. I. e., IPUMS, USAID M-DHS
- Household and person attributes fitted to 1st and 2nd order properties of uncertainty based on census enumerations, infrastructure maps, ground truth information
- Attribute covariance `built in' due to complete survey cases as data structure
- Population simulator designed to allow all input variables, including forecasts, to be stochastic ranges of possible values

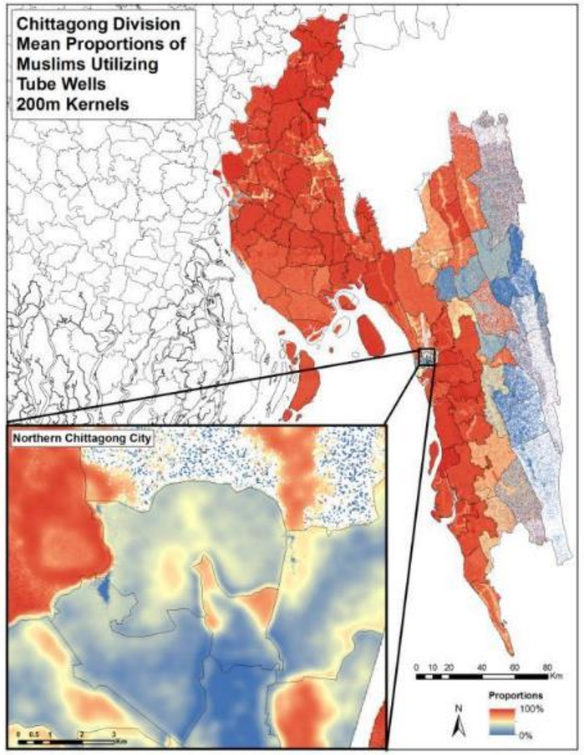


Demographic Analysis

- Stochastic Monte-Carlo demographic simulation
- Summary statistics maps are generated to provide error and uncertainty estimates
- Tools to develop hierarchical framework analysis with uncertainty

Generates space-time uncertainty quantified information about population at neighborhood scale

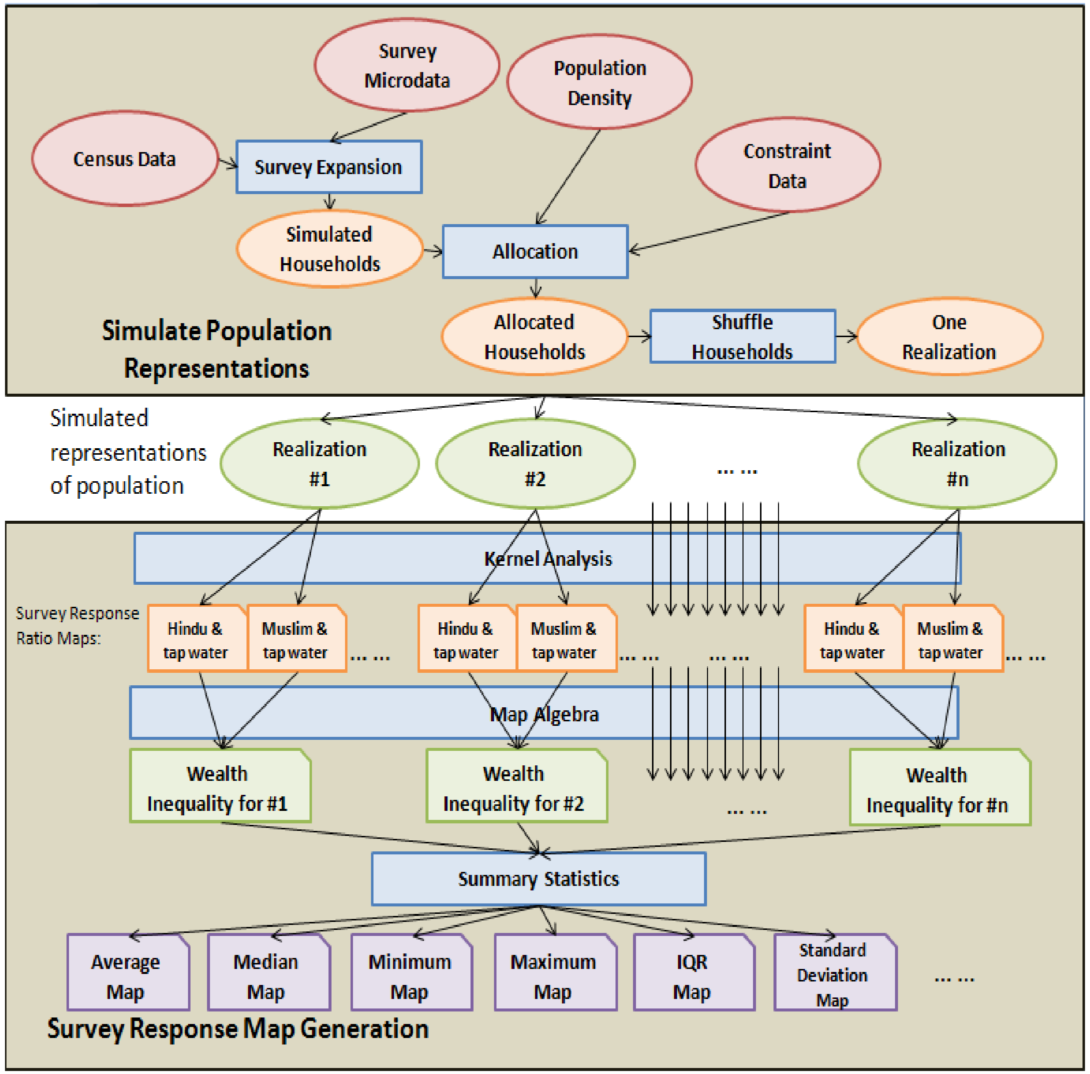




Analysis at atomic scale, then filtered

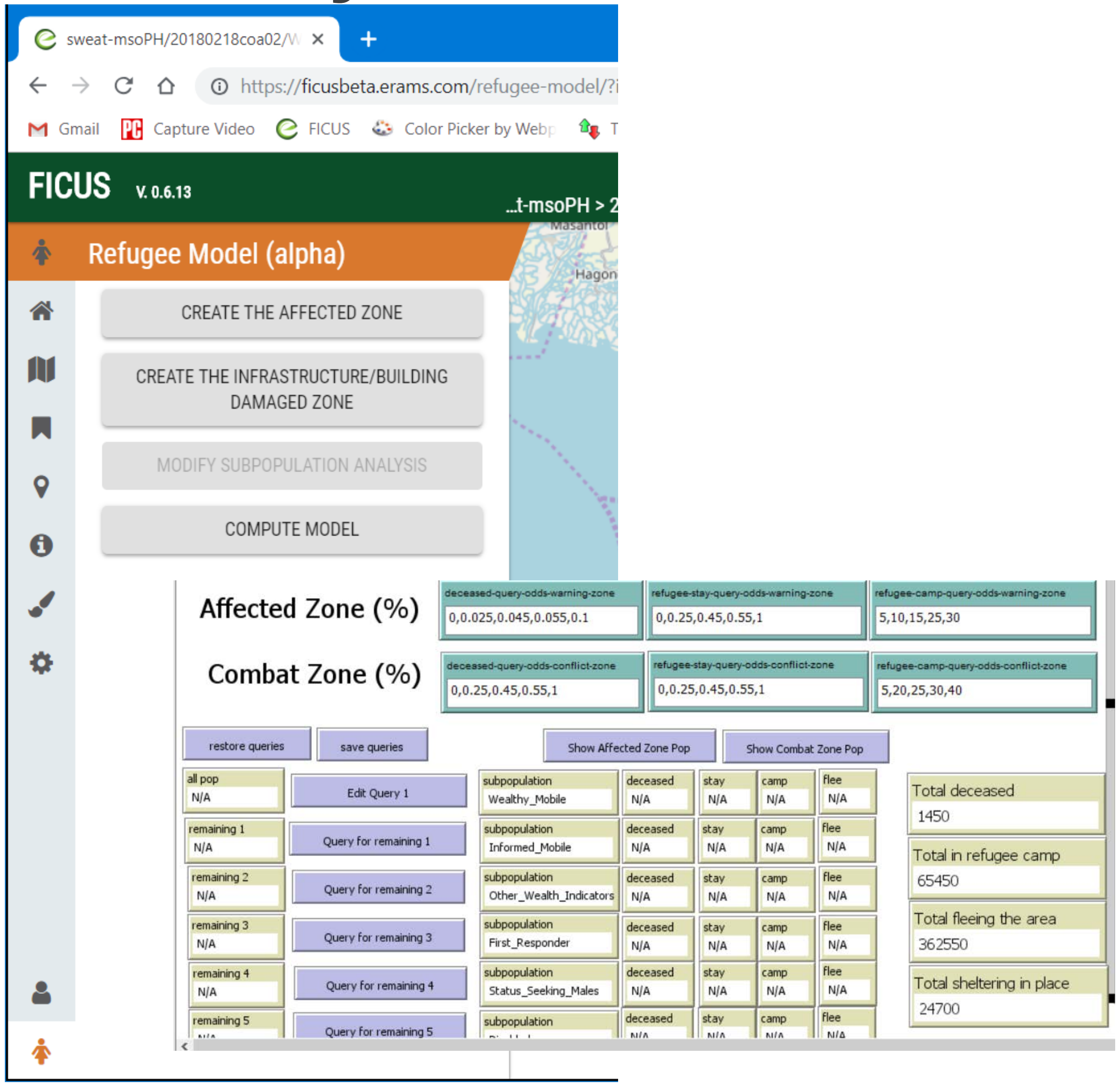
UQ: Simple GeoAnalytics with Population Attributes

- Population-centric multi-verse of individuals allows analyses of attribute combinations without creating modifiable areal unit problem (MAUP) issues
- Wealth inequity analysis example using traditional demographics analysis techniques with a map algebra :
  - Where is neighbourhood scale wealth inequity between Muslims and Hindus in Bangladesh (Ehlschlaeger et al. 2016)?
  - Wealth metrics by household
  - Kernel analysis generates wealth metrics of subpopulations by neighbourhood
  - Map algebra to compare different subpopulations
  - Summary statistics maps are generated to provide Black Swan Theory box plot error estimates



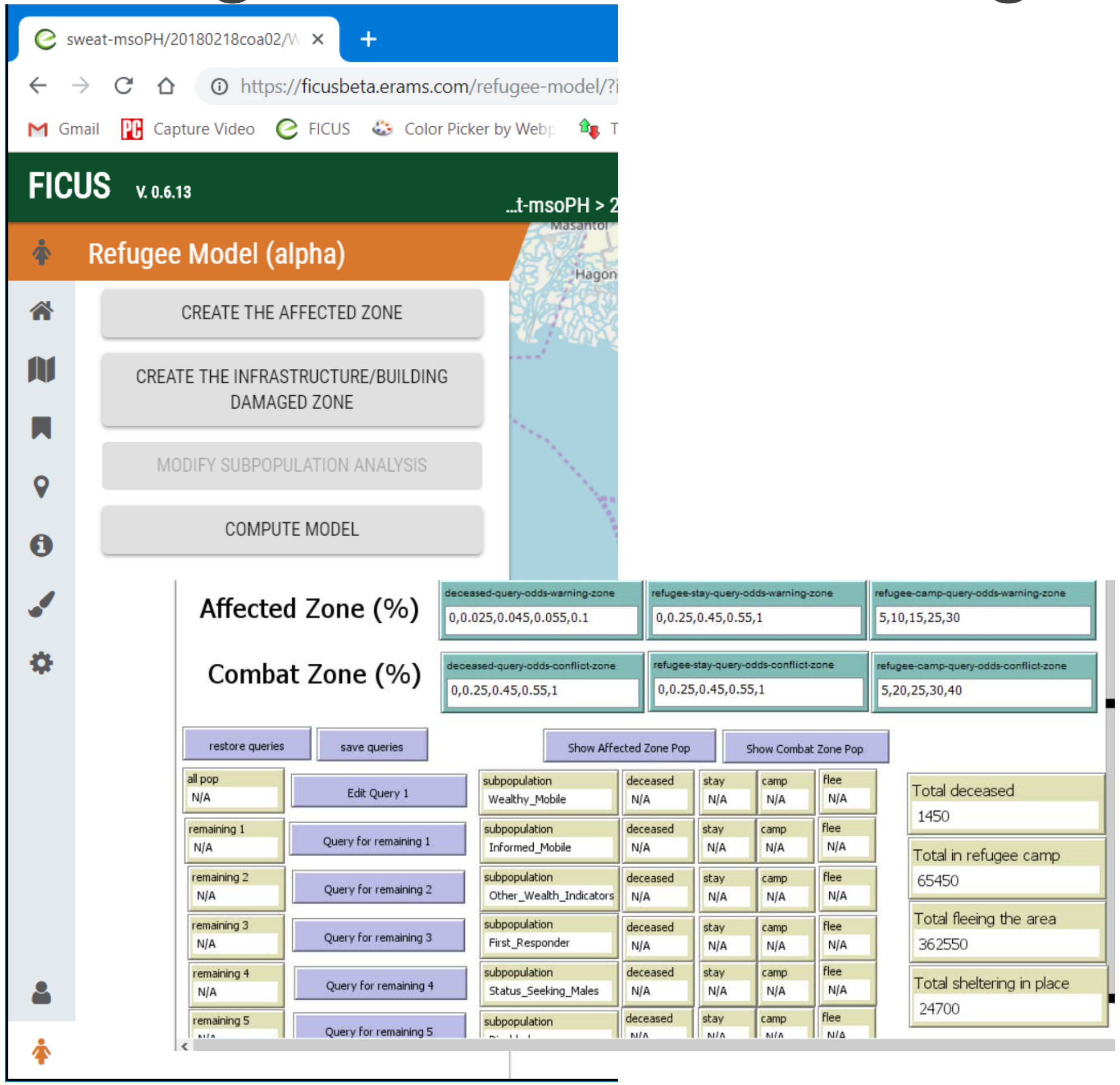
UQ: Population Analysis

- Example application: Examine potential impacts of Battle of Marawi, May-Oct 2017, on population in Mindanao, Philippines
- Classify population across cohorts: flee to family/friends, shelter-in-place, move to IDP camp, become casualties
  - Identify locations and subpopulations based on open source information regarding siege
  - Estimate impacts based on range of possible outcomes per subpopulation. For example:
    - What is the likelihood that `non Muslim adult males' from the affected zone will become casualties? Is broken down into:
      - What is the minimum likelihood that `non Muslim adult males' from the affected zone will become casualties?
      - What is the maximum likelihood that `non Muslim adult males' from the affected zone will become casualties?
      - What is the median likelihood that `non Muslim adult males' from the affected zone will become casualties?
      - ....



UQ: Population Analysis with Agent Based Modeling

- Simple polygons or probability surfaces are least effective, maximizing UGCoP
- Filter functions aggregating attributes into proportions early in an analysis have reduced UGCoP compared to polygons
- Agent Based Modeling incorporating trafficability analysis, supply/demand resource allocation and other system of systems analysis tools minimize UGCoP



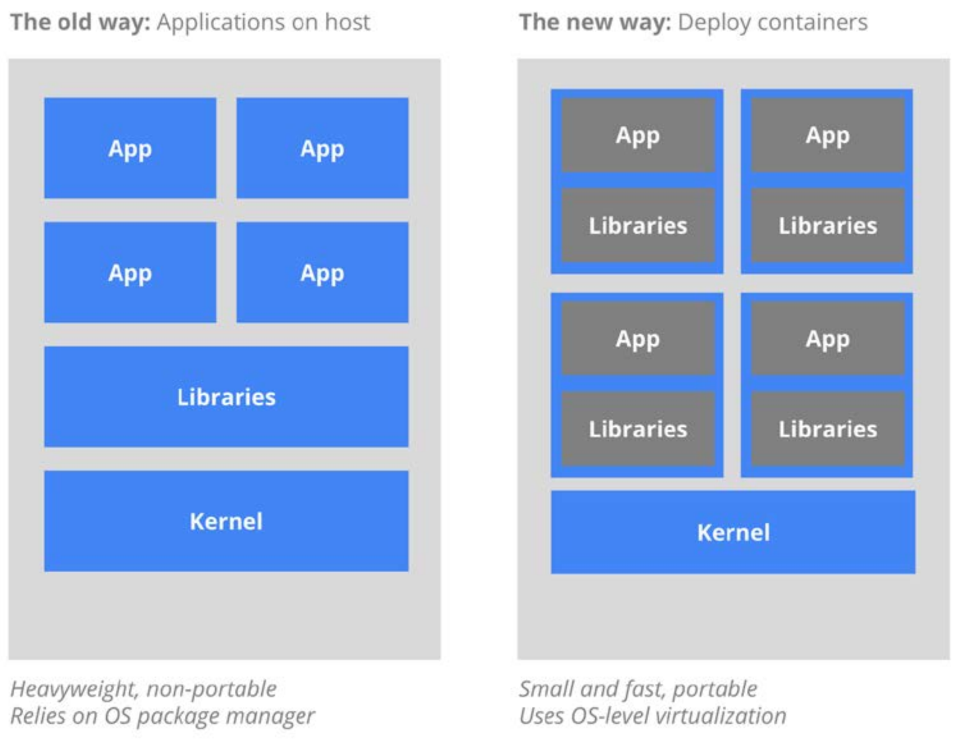
Framework for Incorporating Complex Uncertainty Systems (FICUS)

- Uncertainty Quantification (UQ) of Data Model
- Uncertainty Propagation (UP) through Applications
- Uncertainty Visualization (UV) of Decision Supporting Data/Maps Connected to Inputs

All three needed for Risk Analysis in a distributed Reality to Decision Continuum

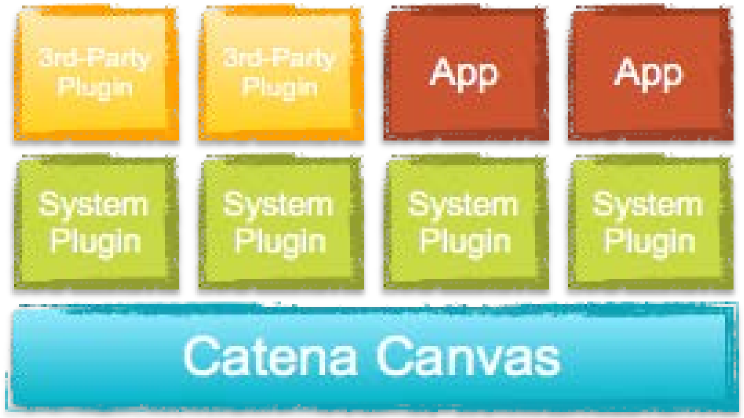
Object Modeling System Computational Framework

- FICUS can easily embed software models and apps written in Java, C/C++, Python, NetLogo, R, Fortran, including geo-temporal analysis tools using OMS
- Apps are packaged in a container with everything the application needs to run
  - No need to install custom software stacks
  - Apps can be run anywhere - Windows, Mac OSX and Linux
  - Apps can be run on local computers and production servers alike
    - Easy to move from development to production
- Containers can be tightly coupled using the Object Modeling System allowing models written in supported languages to operate at compile speeds

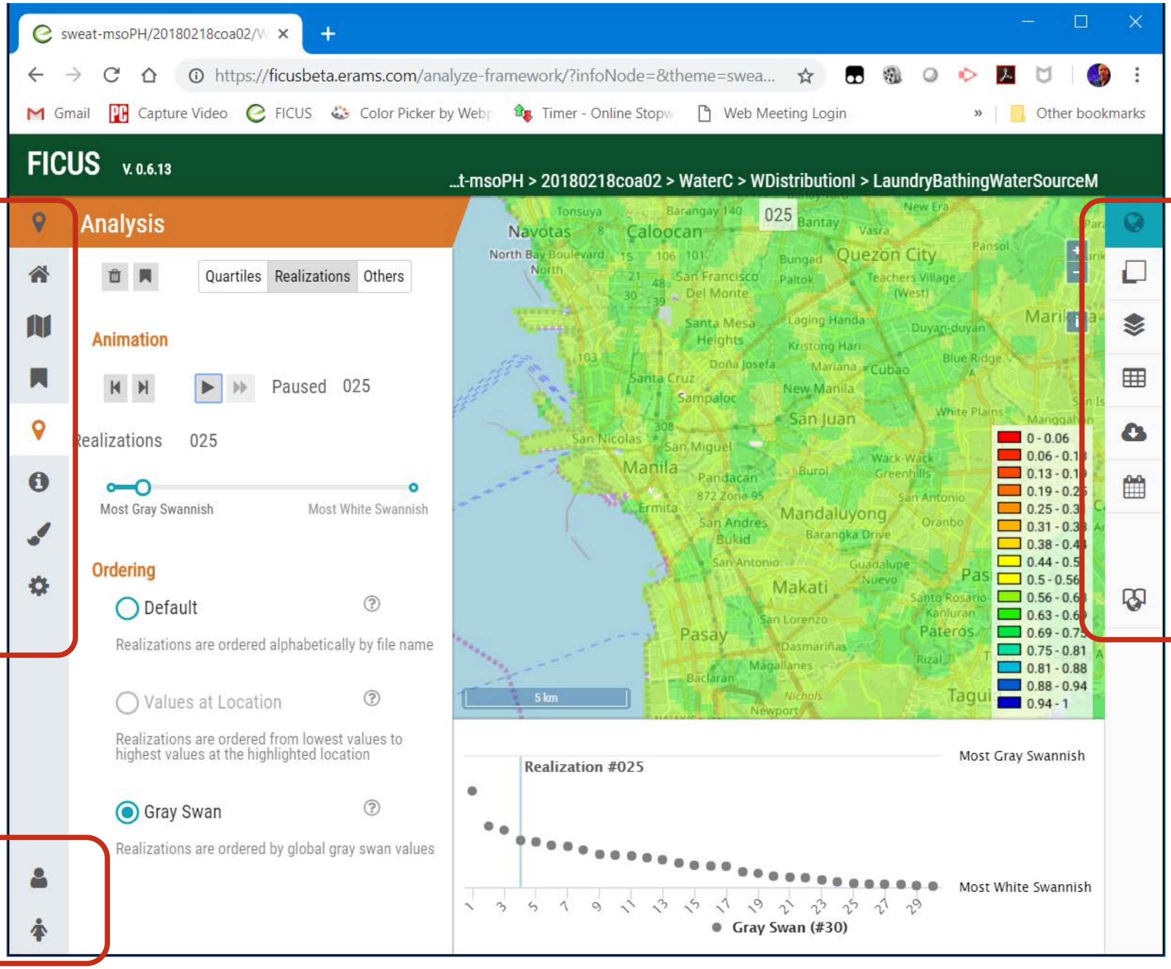


FICUS uses Catena

- Catena is a software framework & libraries for geospatial applications developed at CSU
  - Lightweight core
  - Plugin-based architecture
  - Focus on the app, not the framework



FICUS Cartography, Geotemporal UV



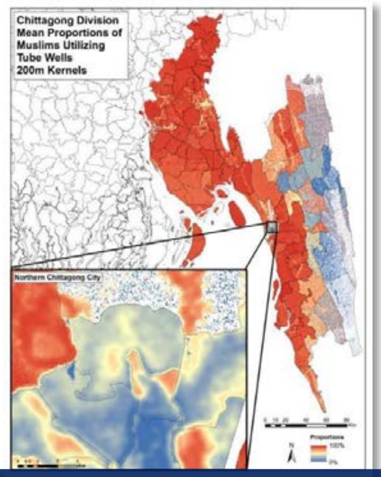
Catena: general purpose GIS w/out Uncertainty

Special Purpose Apps from Any Software, embedded w/in FICUS for UQ/UP/UV

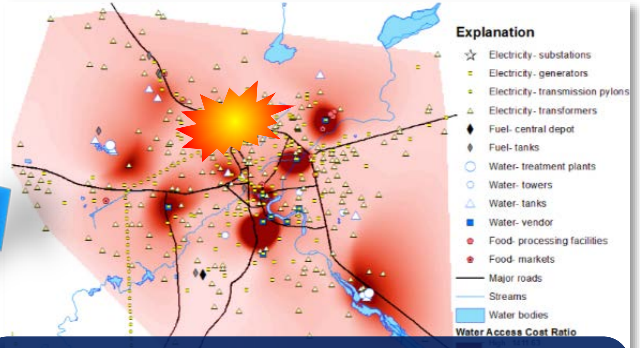
FICUS Component Goal (in development)

- FICUS will soon easily embed any OMS component into the FICUS UQ/UP/UV environment by wrapping the OMS component into a component that allows uncertainty propagation
- FICUS Components will allow stochastic adjustment of all OMS component outputs to correct Gray Swan wrongness, bias, arrogance, or meekness
- FICUS Components will allow for immediate uncertainty visualization of all inputs and outputs `linked and brushed' in the FICUS-UI to improve model communication

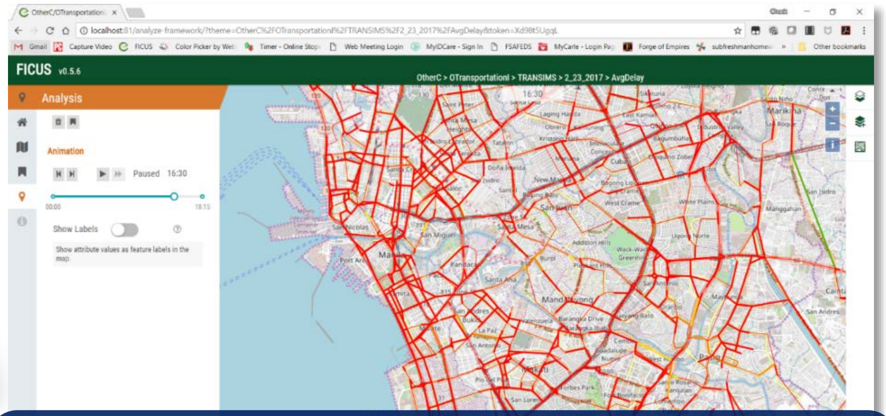
Uncertainty Propagation



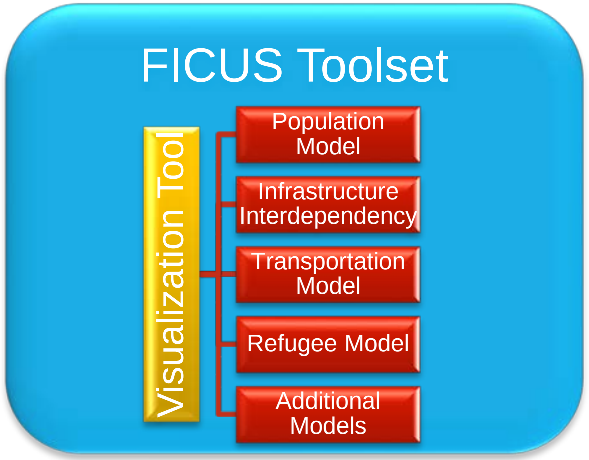
Represent neighborhood population attributes



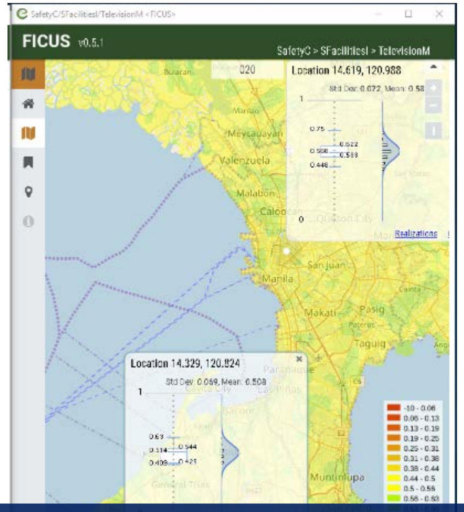
Analyze cascading effects of infrastructure damage (on people and other systems)



See dynamic changes in transportation flow based on damage and change in resource demand



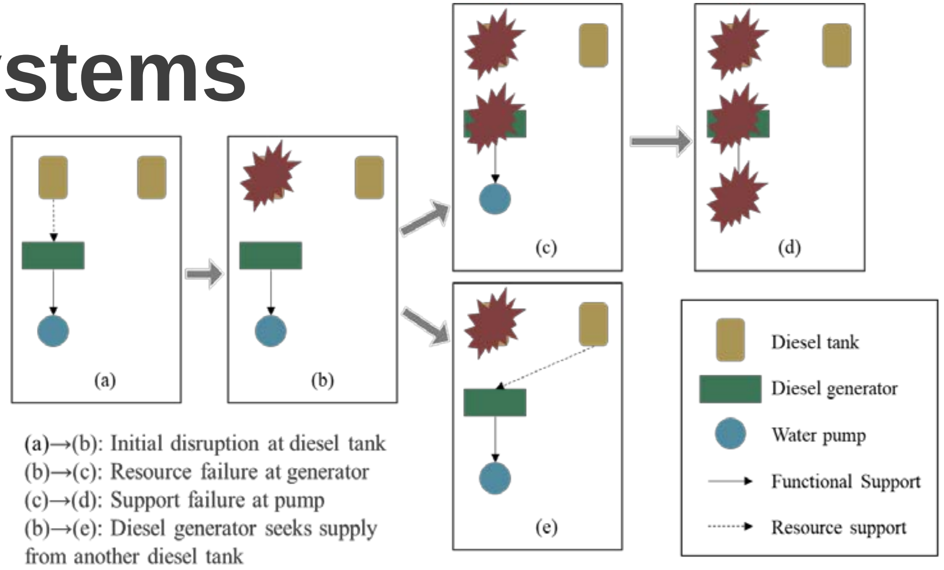
Filter model outputs through risk & decision framework

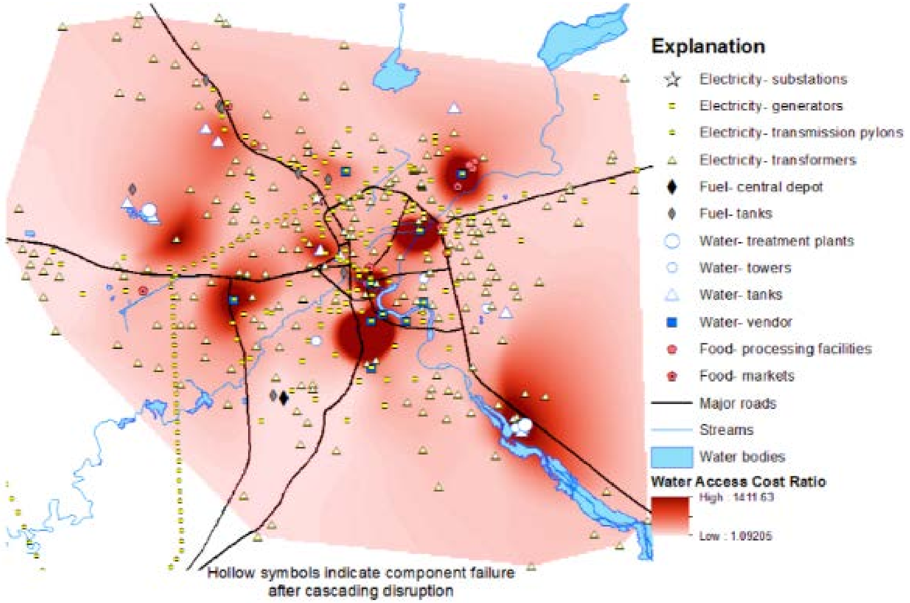


View analysis results through a web-service

Infrastructure System of Systems

- UIUC Civil Engineering leads effort
- Network infrastructure interdependency model
- Considers connections between different infrastructure types
- Connects ‘communities’ to infrastructure resources (proximity)
- Considers how change in resource capacity impacts population

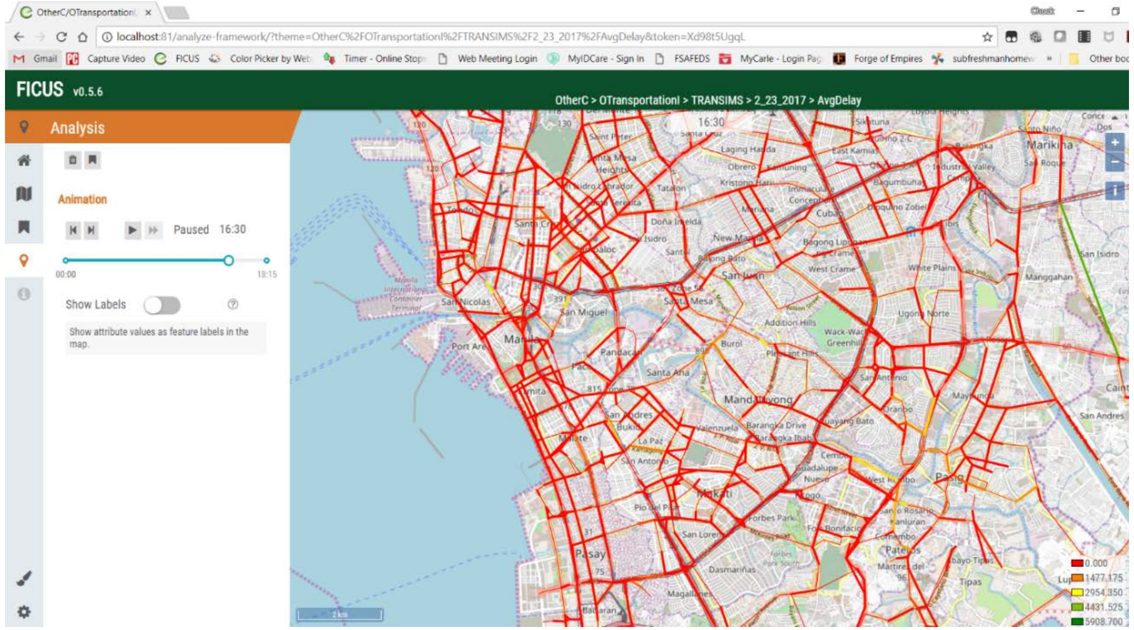




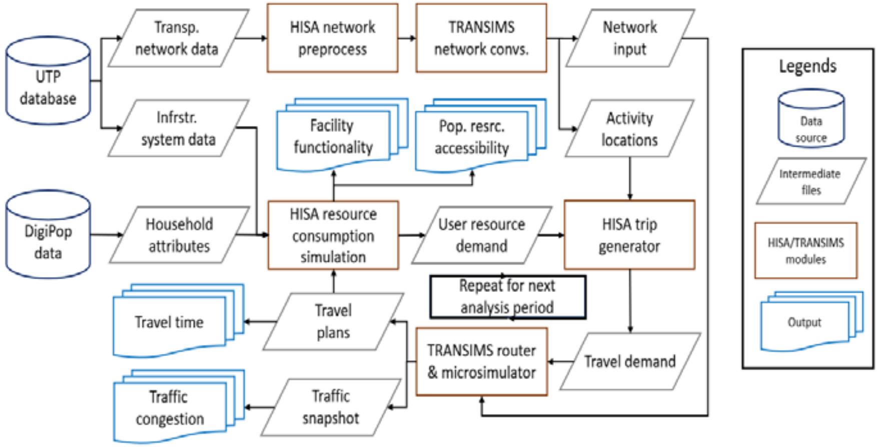
Relationship model that expresses the connection between infrastructure & population

Tight-coupled System of Systems Modeling

- Link demographic, infrastructure interdependency, and transportation models
- Using road network and population data, generates a set of activity locations for origins/destinations
- Enable near real-time movement analysis in the transportation network
- Determines impact of infrastructure failure on access to resources, and dynamic changes in traffic

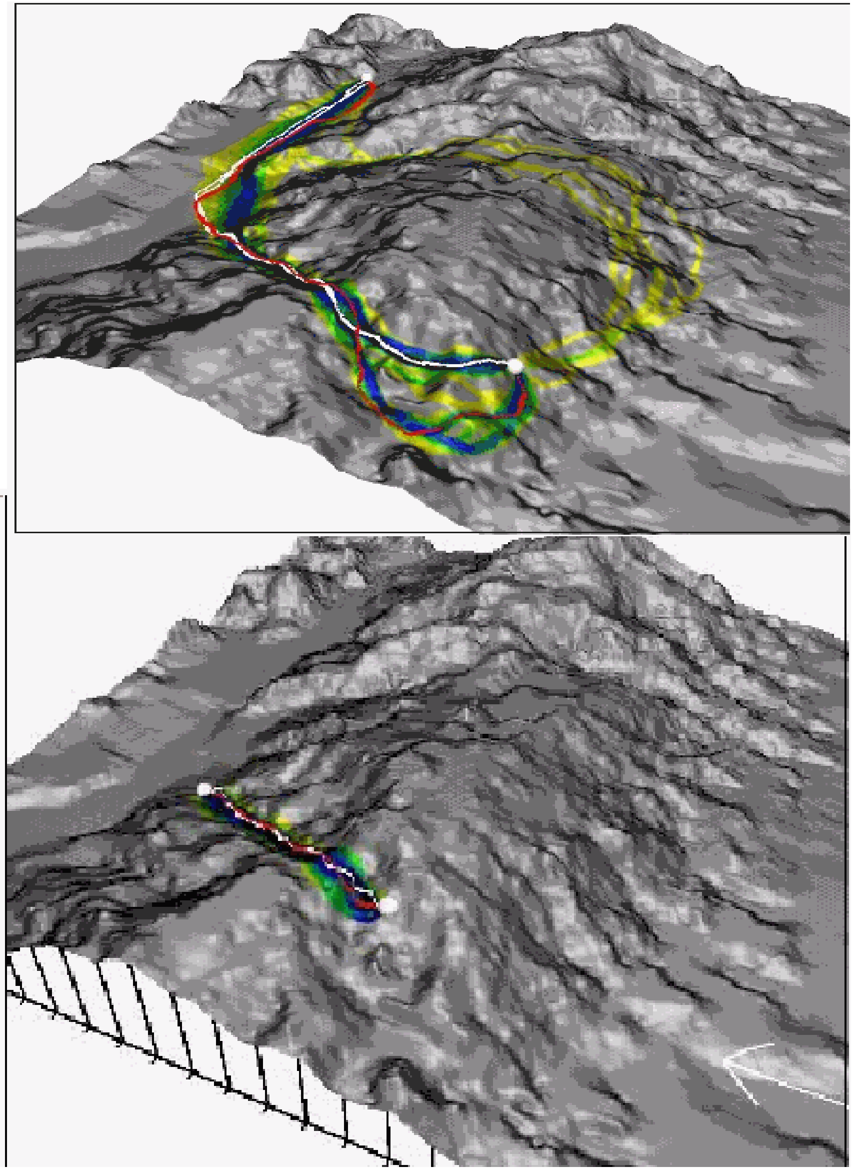


Develop dynamic transportation simulation influenced by resource capacities & population attributes



Heteroskedastic Uncertainty Verification Problem

- Heteroskedasticity: when the variability of a variable is unequal across the range (or geography)
- Heteroskedastic Uncertainty Verification Problem (HUVP) is the notion that verification for fitness for use of a data model to a geographic computational model is difficult because certainty of specific use cases will not guarantee certainty in other circumstances
- Simple Example (on right): Shortest Path Uncertainty Analysis using UQ 3' DEM data has uncertain results (top) or certain results (bottom) depending on model parameters even using the same dataset
- Complex Demographic Example: Microdata or surveys will seldom provide the appropriate stratification of random samples for important demographic variables for specific applications (urban/rural in USAID's M-DHS for example). Ergo, some applications will provide fit-for-use results from the same survey with other application results unfit-for-use
- Dr. Gerard Heuvelink has argued since the 1980's that Monte Carlo Simulation is the only technique that `can' provide uncertainty analysis, which FICUS uses as it's core.
- Computational Modelers need to adopt a "Manifesto for Antifragile Models" similar to Computer Science's Manifesto for Antifragile Software, antifragilesoftwaremanifesto.org



Antifragile ~~Software~~ Geo-temporal Analysis Manifesto

If you replace `software' with `geo-temporal models', we are nearly there…

1. The customer: The main priority of Antifragility is so the satisfaction of the customer needs, delivering an antifragile system of systems. The system of systems includes all USG data collection flowing through geo-temporal computational models connected to decision makers with meta-system feedback loops to identify weaknesses in the analytic framework.
2. The context: We welcome changing scenarios where unexpected events (Black Swans) are the real paradigm shifting entities, or used to improve the analytic framework. We also welcome identifying wrongness, arrogance, and bias in our models to minimize poor decisions.
3. … 12.

Because individual models can only be calibrated and validated on unit test data available at development:

If we want to tightly couple multiple models, designed by different researchers, into systems with multiple feedback loops, we need a Framework for Integrating Complex Systems that can handle the Uncertainty so that the Analytic Framework is Antifragile.

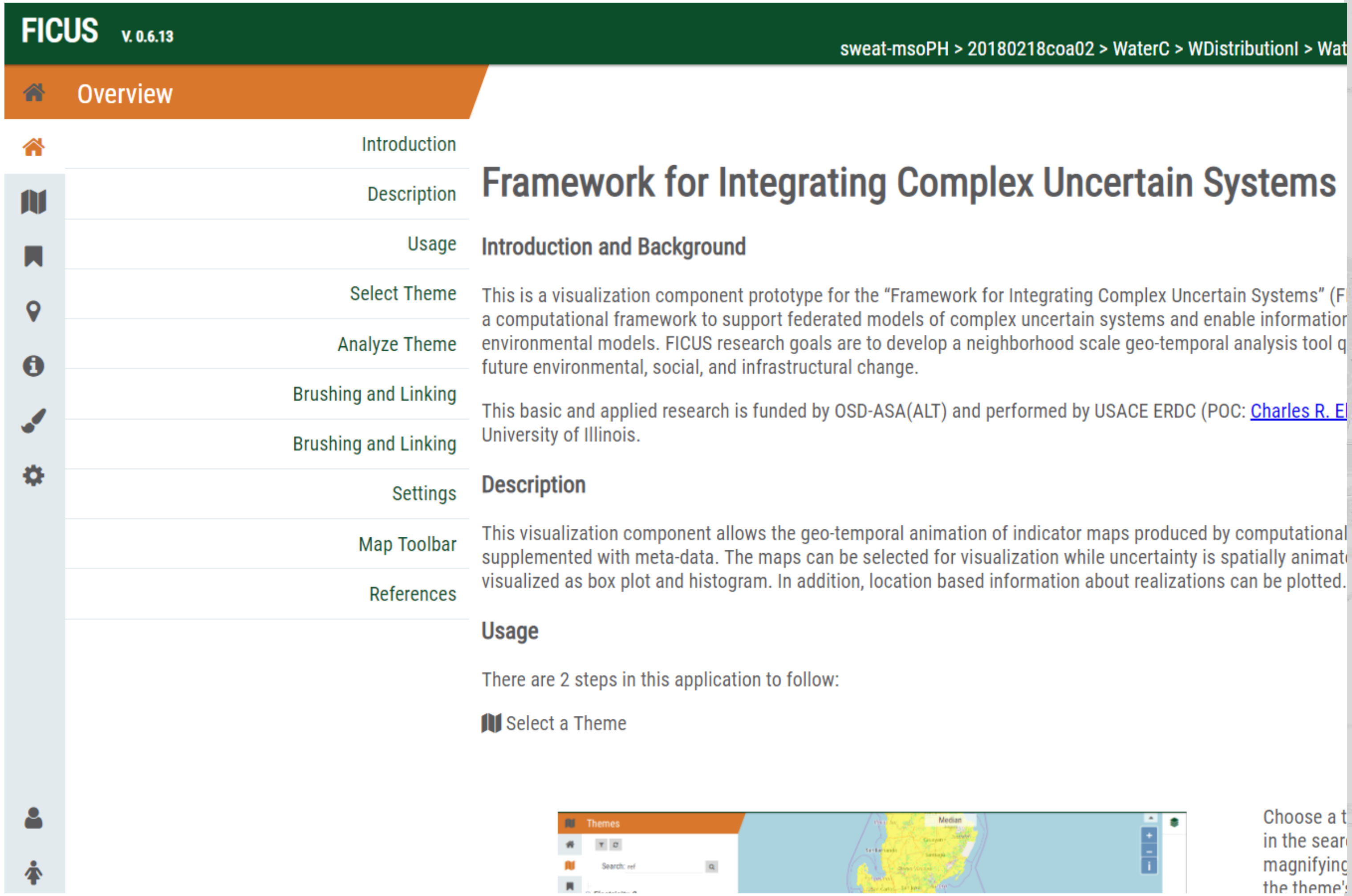
Framework for Incorporating Complex Uncertainty Systems (FICUS)

- Uncertainty Quantification (UQ) of Data Model
- Uncertainty Propagation (UP) through Applications
- Uncertainty Visualization (UV) of Decision Supporting Data/Maps Connected to Inputs

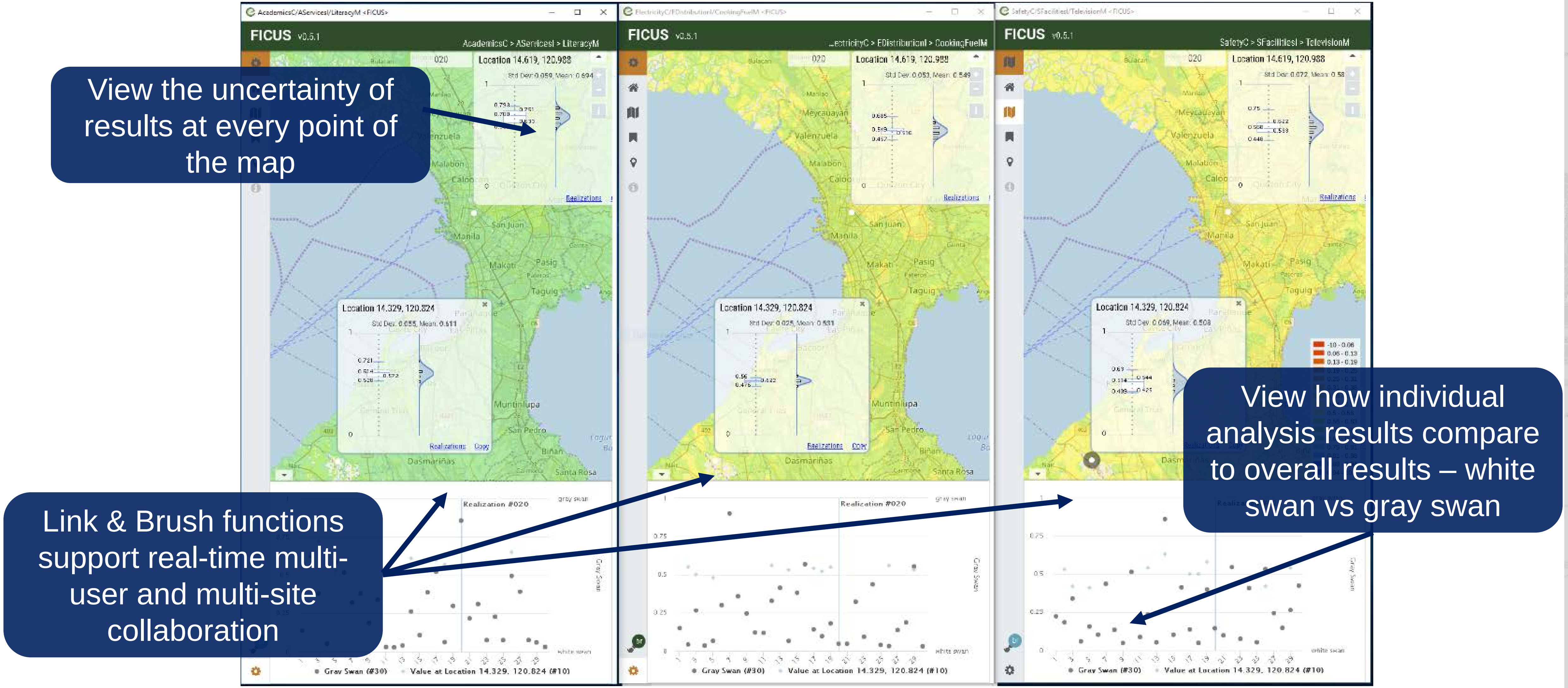
All three needed for Risk Analysis in a distributed Reality to Decision Continuum

Play w/ FICUS Yourself

- In a hurry, advance two slides and go to ficusbeta.erams.com for a walkthrough of the cartographic components to the FICUS User Interface (or ficus.erams.com for the latest stable version)

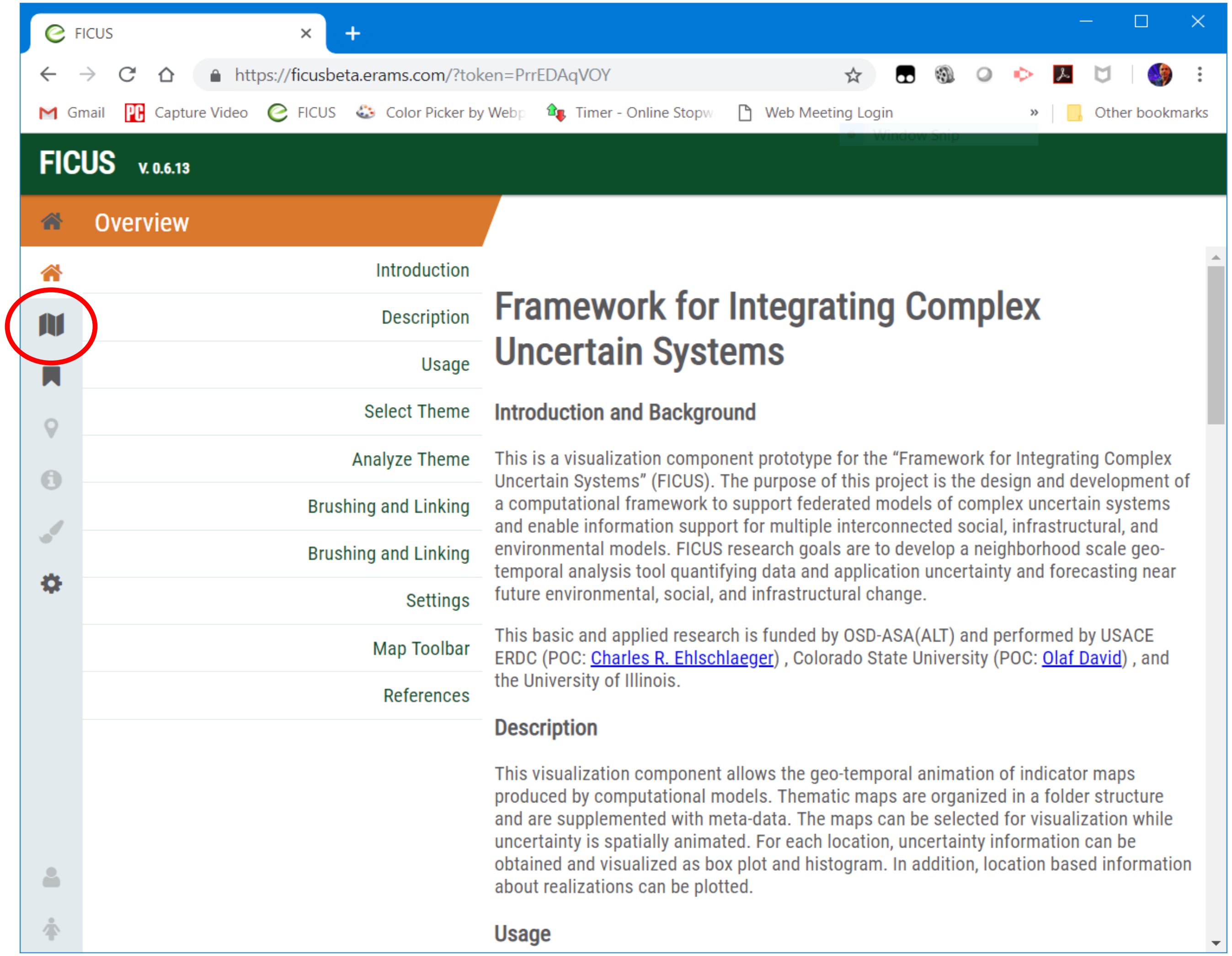


FICUS Visualization Interface



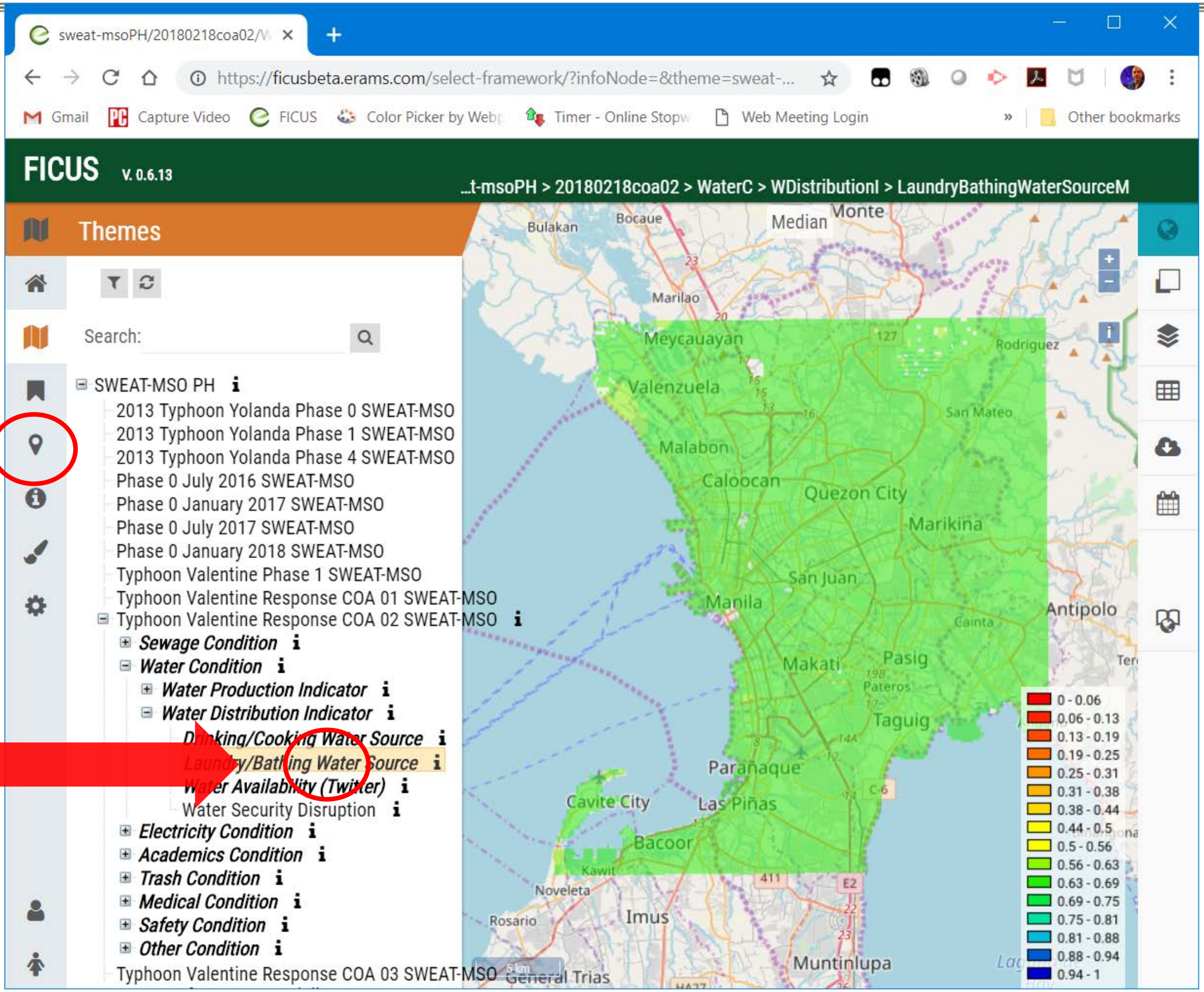
Play w/ FICUS Yourself

- ficusbeta.erams.com



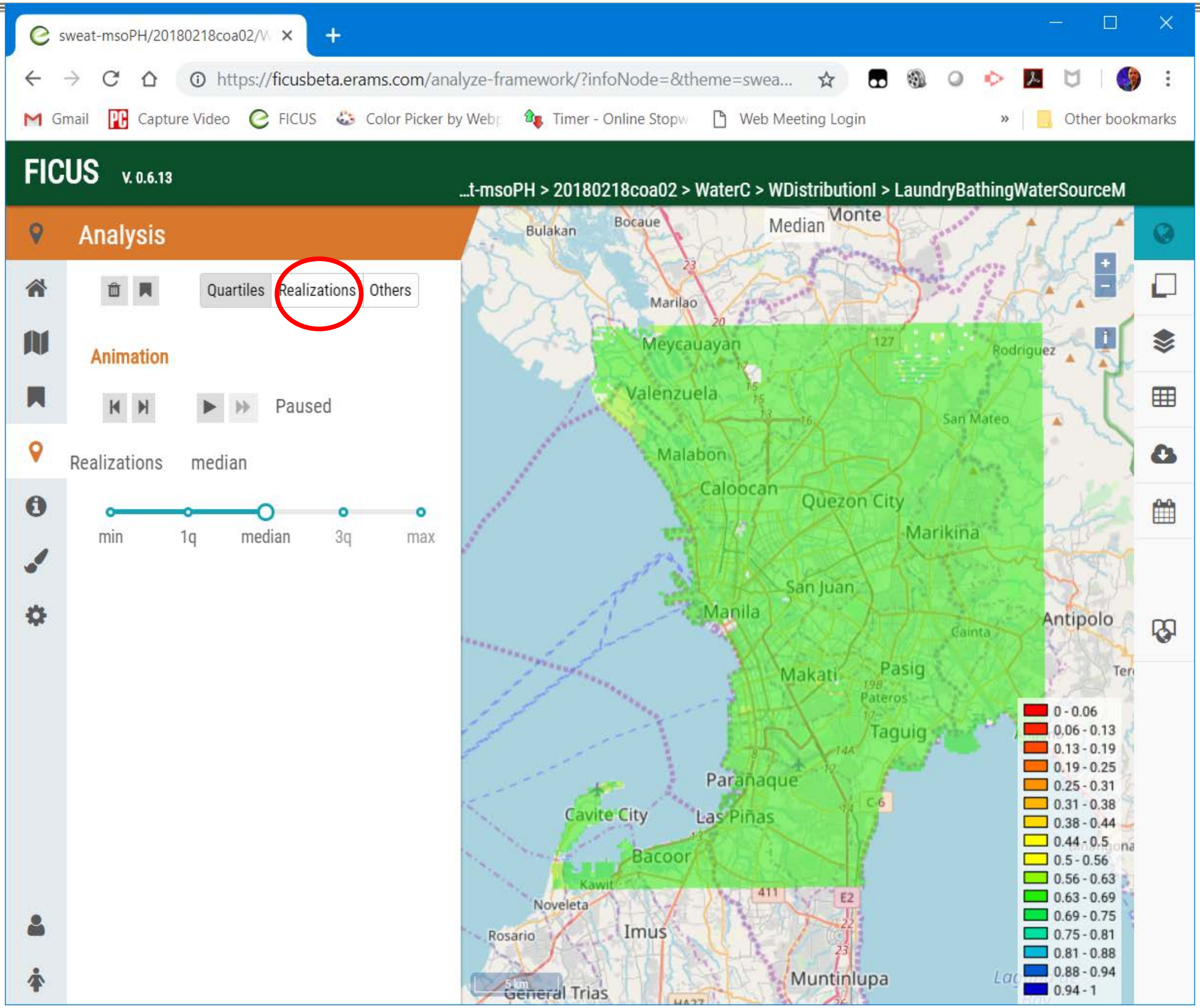
Play w/ FICUS Yourself 2

Then click

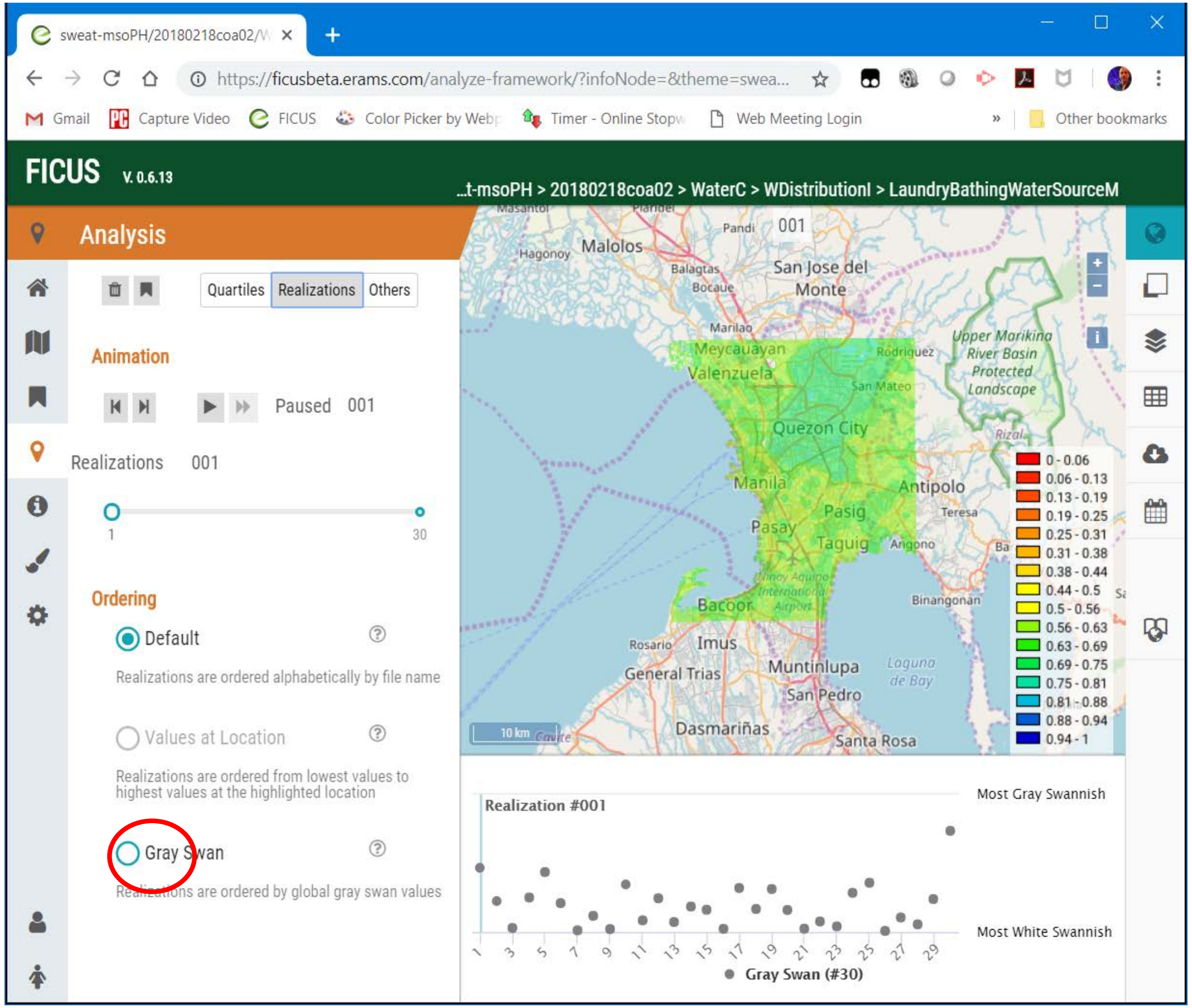


First click

Play w/ FICUS Yourself 3



Play w/ FICUS Yourself 4



Closing Remarks

- UQ/UP/UV at atomic scale eliminates Modifiable Areal Unit Problem
- Eliminating Uncertain Geographic Context Problem (UGCoP) requires application-specific options not found in legacy GIS
- UQ/UP/UV reduces Heteroskedastic Uncertainty Verification Problem issues
- UQ/UP/UV techniques provide connection between decision making knowledge gaps and improved data collection, necessary for true risk analysis
- Distributed Geotemporal Risk Analysis representing complex systems requires novel data models, application development methods, and cartographic techniques not available in current GIS
- Contact Dr. Ehlschlaeger for employment opportunities, or sabbaticals
- Contact Dr. Ehlschlaeger and Dr. David for collaborative research opportunities

POC:
Dr. Charles Ehlschlaeger,
charles.r.ehlschlaeger@usace.army.mil
Cell: 217-418-8942
Office: 217-373-7287

POC:
Dr. Olaf David
Olaf.David@colostate.edu

QUESTIONS?