Extracting Categorical Indicators from Epidemiological Timeseries

Srini Venkatramanan, PhD

Research Assistant Professor

Biocomplexity Institute, University of Virginia

25 June 2023

CSTE 2023 Annual Conference - Infectious Disease Forecasting Workshop

srini@virginia.edu

UNIVERSITY of VIRGINIA

BIOCOMPLEXITY INSTITUTE

https://biocomplexity.virginia.edu/

OVERVIEW

BACKGROUND

CATEGORICAL INDICATORS

TOOL INTRODUCTION

BACKGROUND

WHO WE ARE

- Biocomplexity Institute at the University of Virginia
  - Using big data and simulations to understand massively interactive systems and solve societal problems
  - 20+ years of experience of developing and analyzing infectious disease models for policy support
- Supporting VDH on COVID-19 response from April 2020
- Sustained contributions to CDC-coordinated Forecast and Scenario Modeling Hubs for COVID-19
- CSTE-funded site for Flu Forecasting (2021-22, 2022-23), Outbreak Tools development (2023-24)
- Part of CDC's Pathogen Genomics Centers of Excellence

Members of the Network Systems Science & Advanced Computing Division

Federal sponsors

https://nssac.github.io/covid-19/

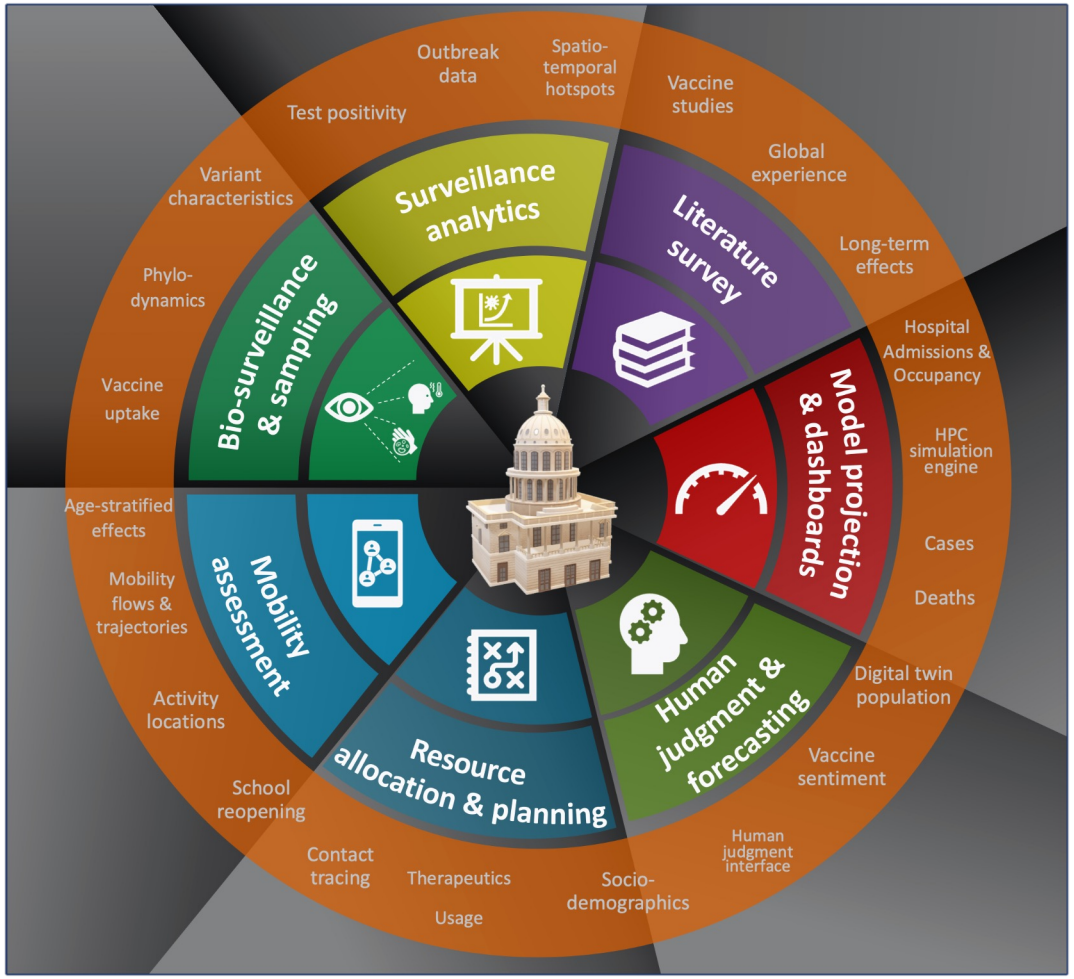
https://computational-epidemiology.org/data-tools

SUPPORTING VDH COVID-19 RESPONSE

- Multi-pronged approach to support COVID-19 response in the Commonwealth of Virginia
- 120+ weekly and biweekly briefings to policymakers and health practitioners
- Leveraging real-time datasets and HPC software workflows to build multiple classes of models
  - Situation assessment
  - Short-term forecasting
  - Medium-term projections
  - Resource allocation

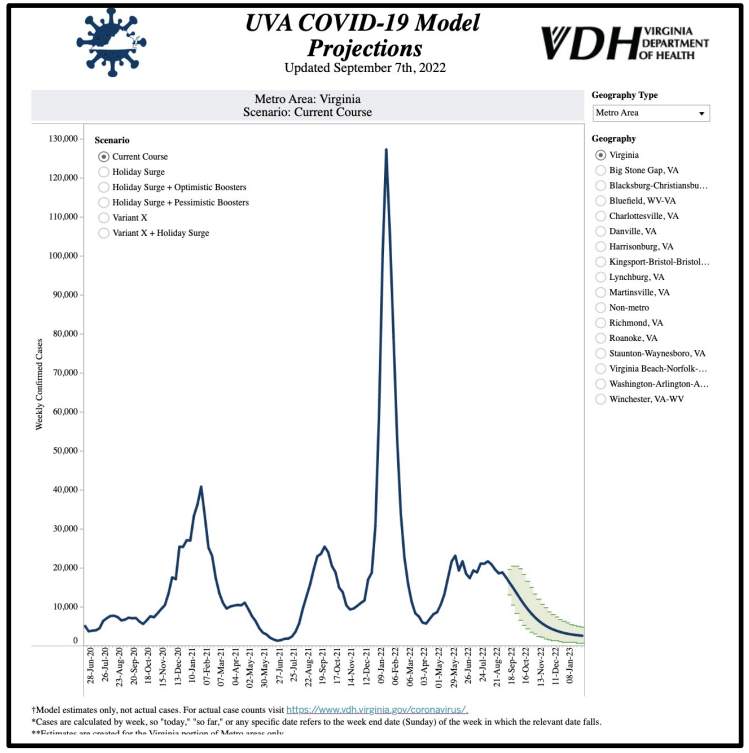
Being presented at the CSTE 2023 Annual Conference

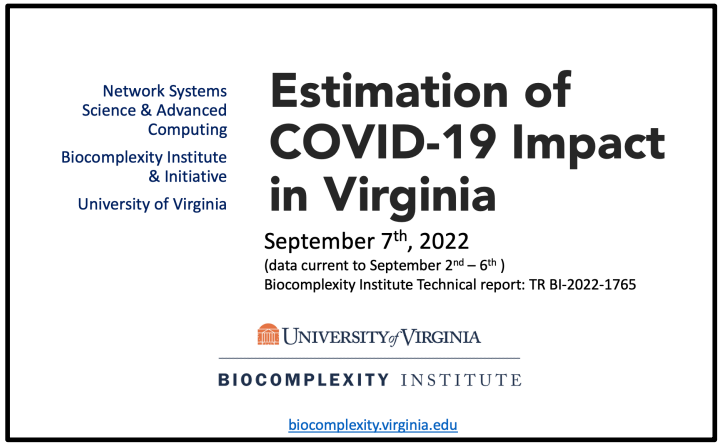
- Ongoing research threads
  - Genomic surveillance
  - Integrated Bio-surveillance
  - Human judgment systems

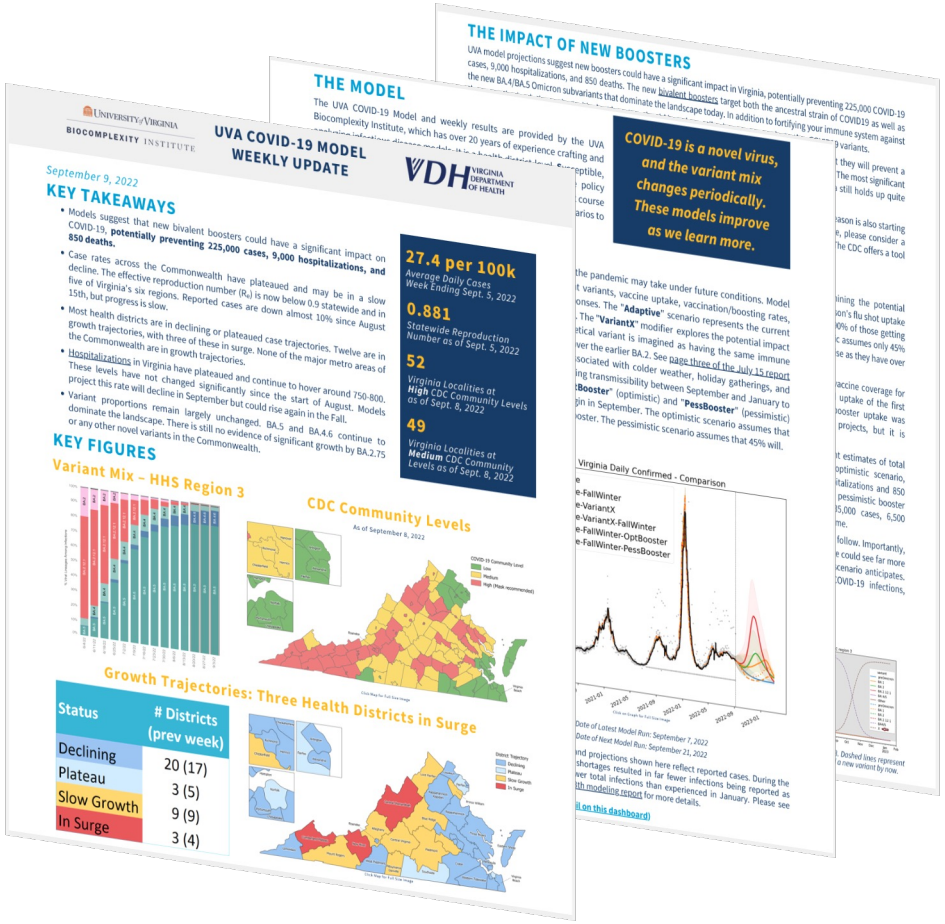


VDH VIRGINIA DEPARTMENT OF HEALTH
VIRGINIA DEPARTMENT OF EMERGENCY MANAGEMENT
VIRGINIA DEPARTMENT OF EDUCATION
VIRGINIA HOSPITAL & HEALTHCARE ASSOCIATION
An alliance of hospitals and health delivery systems

WEEKLY REPORTS AND BRIEFINGS







https://www.vdh.virginia.gov/coronavirus/covid-19-data-insights/

https://nssac.github.io/covid-19/vdh_weekly_briefings/

CSTE OUTBREAK TOOLS PROJECT

- Project started in February 2023
- Objective: Pilot test, evaluate, and deploy analytic tools with identified partners from STLT health departments

CSTE
Council of State and Territorial Epidemiologists

Funding Opportunity Announcement:
Development of forecast, analytic, and visualization tools to improve outbreak response and support public health decision-making

UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE


Ajay Sanjeevan


Aniruddha Adiga


Benjamin Hurt


Bryan Lewis


Gursharn Kaur


Madhav Marathe


Srini Venkatramanan

VDH VIRGINIA DEPARTMENT OF HEALTH

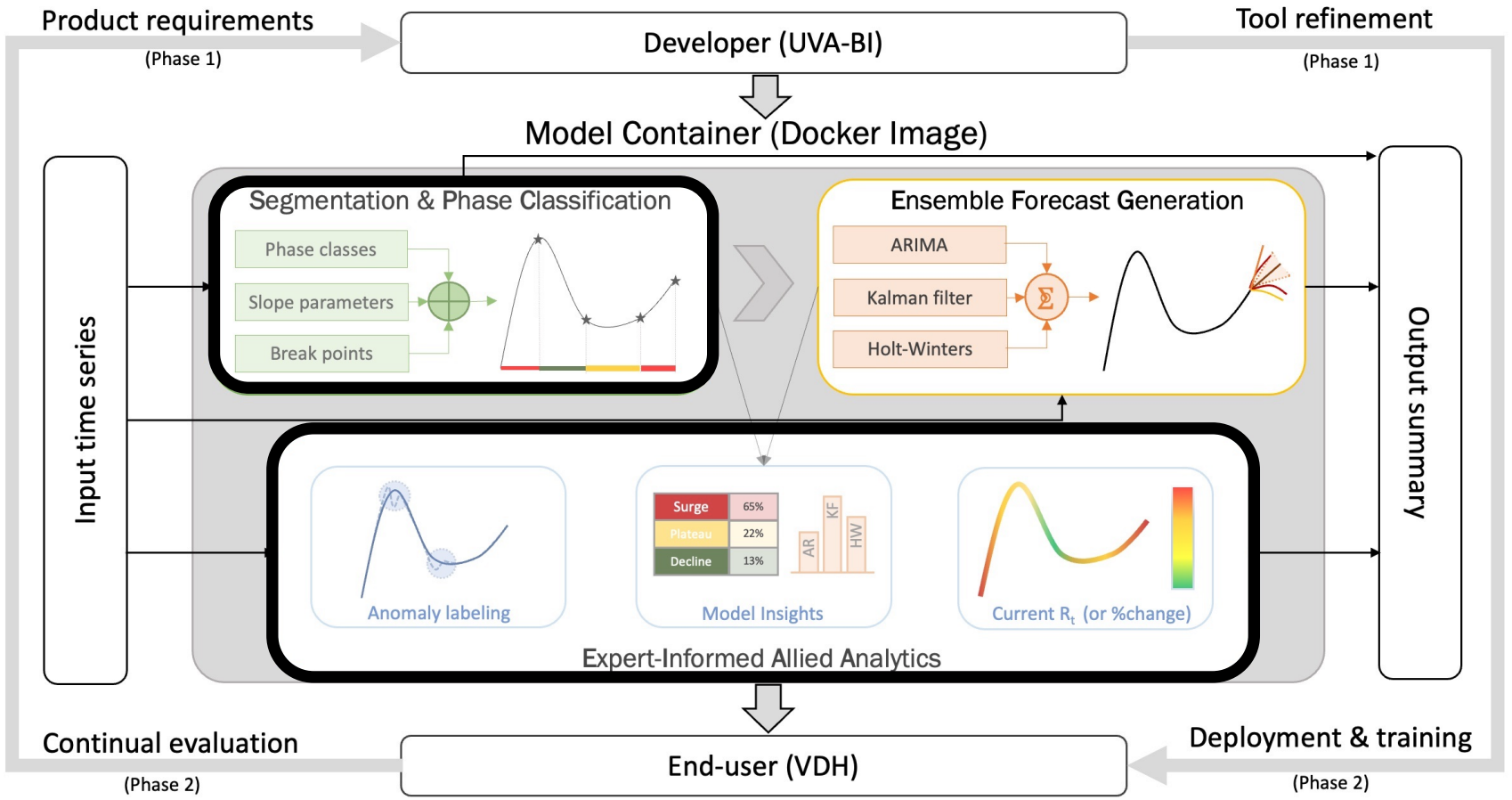
Foresight and Analytics, Office of Emergency Management


Justin Crow


Alex Telionis

PROPOSED FRAMEWORK

Strengthening **P**ublic **H**ealth **I**nformatics using **N**e**X**t-generation tools

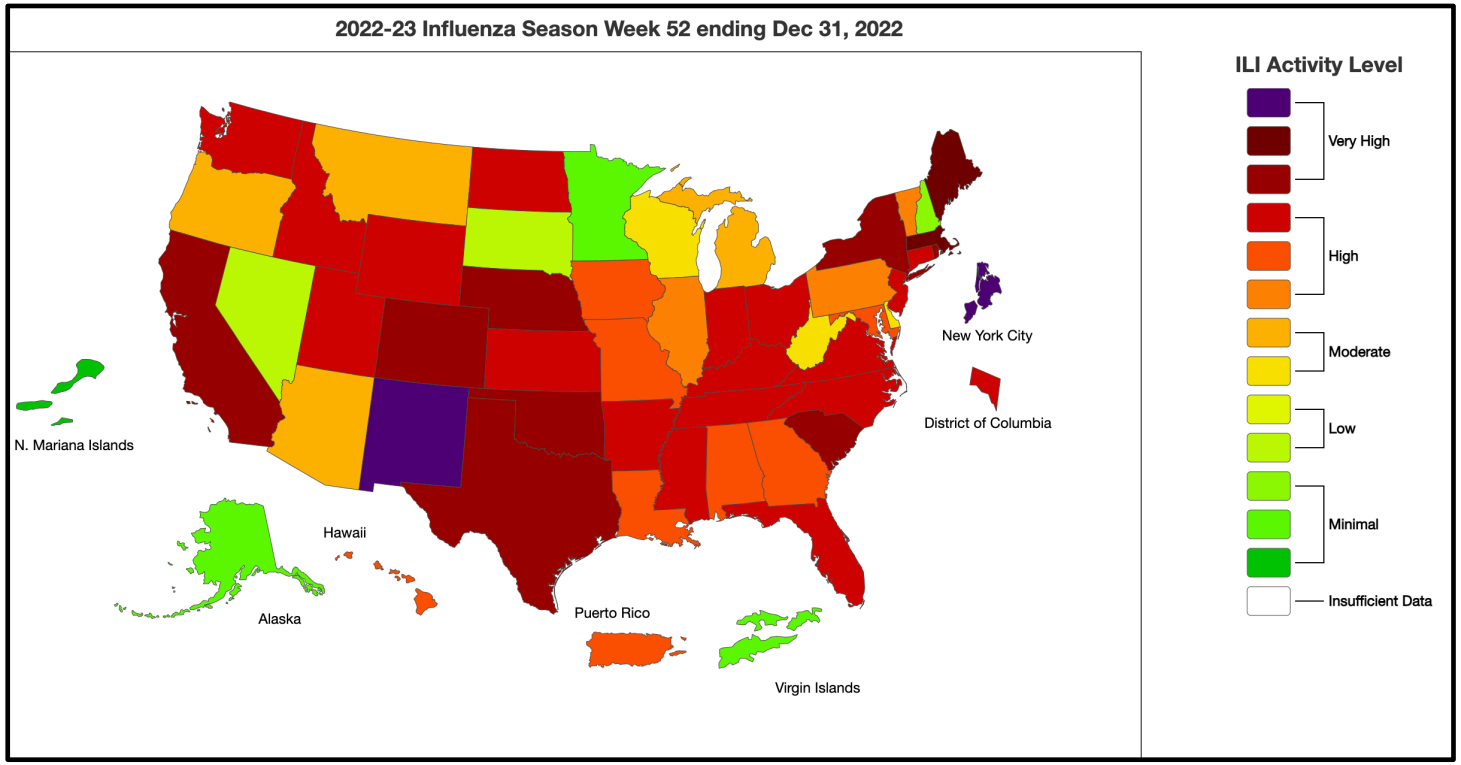


CATEGORICAL INDICATORS

CATEGORICAL INDICATORS FOR EPIDEMICS

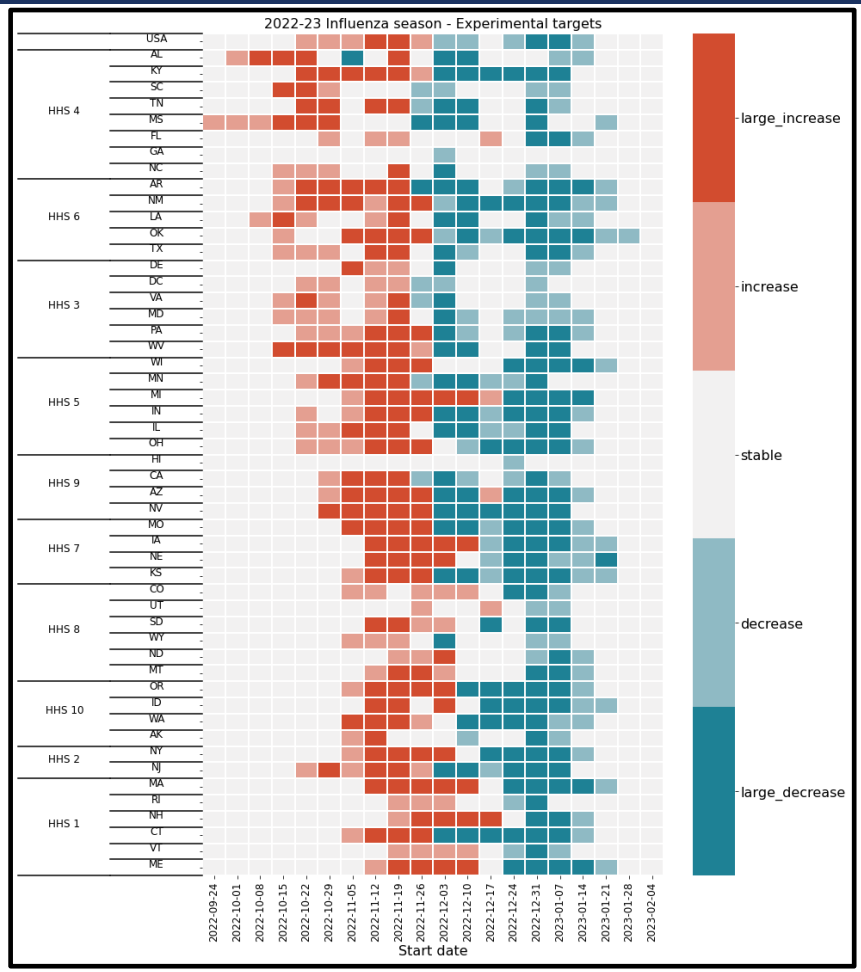
- What do we mean by categorical indicators?
  - Indicators of current level of activity (e.g., High, Medium, Low)
  - Indicators of current trend in the activity (e.g., Surging, Plateau, Declining)
  - Could be based on the raw time series or other derived quantities (e.g., R_t)
- Why categorize?
  - Useful summary of epidemic activity for situation assessment and prognosis
  - Valuable for decision makers as well as public communication
- Challenges
  - Various definitions and lack of standardization
  - Sensitivity to anomalies and dynamic range
- Opportunities
  - Correlating across multiple signals
  - Aggregating across scales

EXAMPLES FROM INFLUENZA SURVEILLANCE



ILINet State Activity Indicator Map

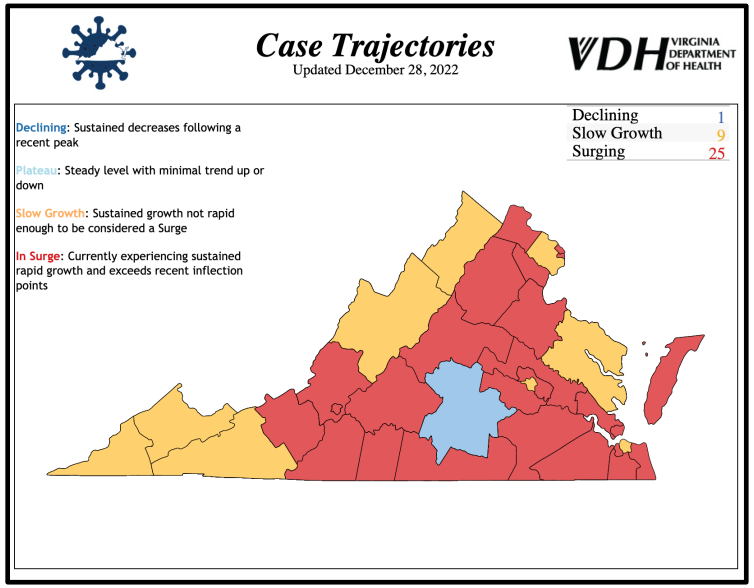
Source: https://gis.cdc.gov/grasp/fluview/main.html



HHS hospitalizations trend indicators (FluSight 2022-23)

https://github.com/cdcepi/Flusight-forecast-data/tree/master/data-experimental

EXAMPLE FROM COVID-19: CASE SURVEILLANCE

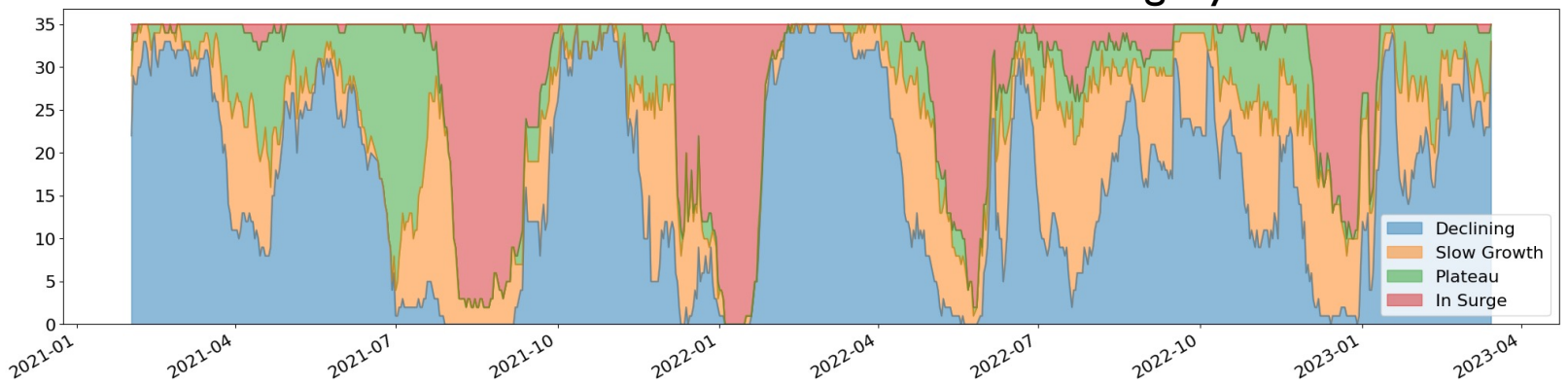


https://www.vdh.virginia.gov/coronavirus/see-the-numbers/covid-19-modeling/district-trajectory-map/

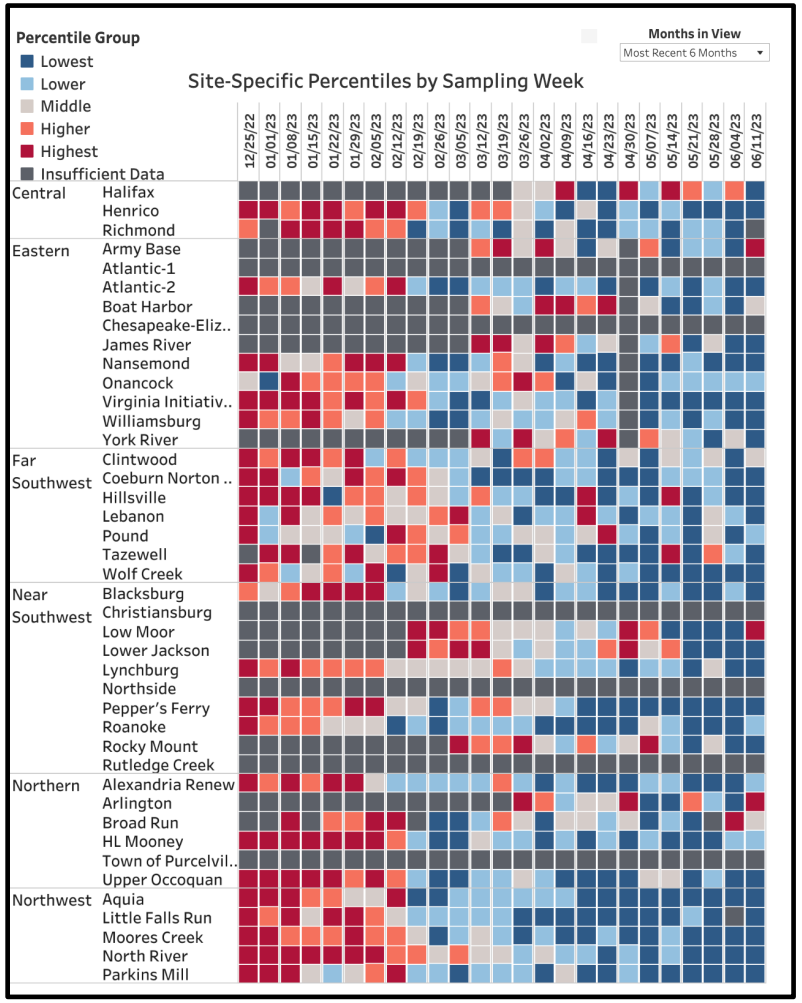
Customized slope thresholds for category definitions

Trajectory	Description	Weekly Case Rate Slope (per 100k)	Weekly Hosp Rate Slope (per 100k)
Declining	Sustained decreases following a recent peak	slope < -0.88/day	slope < -0.07/day
Plateau	Steady level with minimal trend up or down	-0.88/day < slope < 0.42/day	-0.07/day < slope < 0.07/day
Slow Growth	Sustained growth not rapid enough to be considered a Surge	0.42/day < slope < 2.45/day	0.07/day < slope < 0.21/day
In Surge	Currently experiencing sustained rapid and significant growth	2.45/day < slope	0.21/day < slope

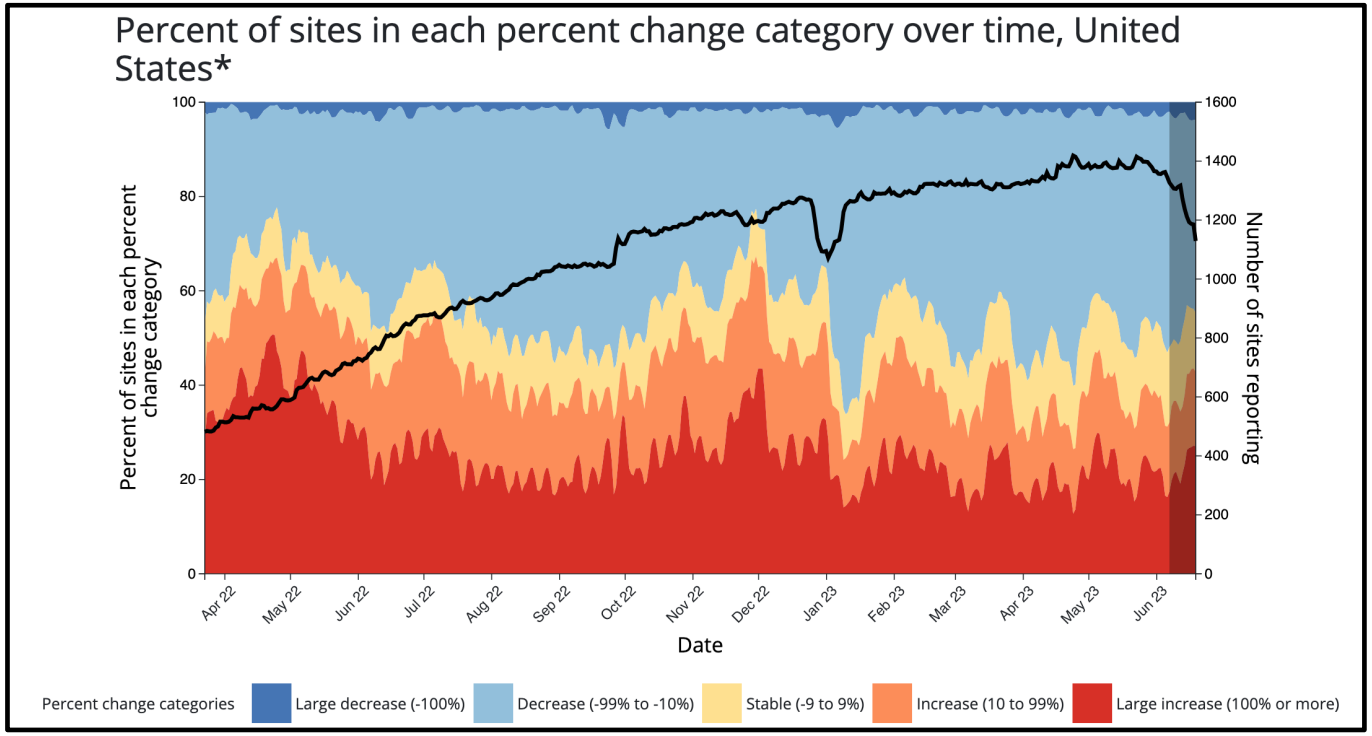
Number of health districts in each category



EXAMPLES FROM COVID-19: WASTEWATER SURVEILLANCE



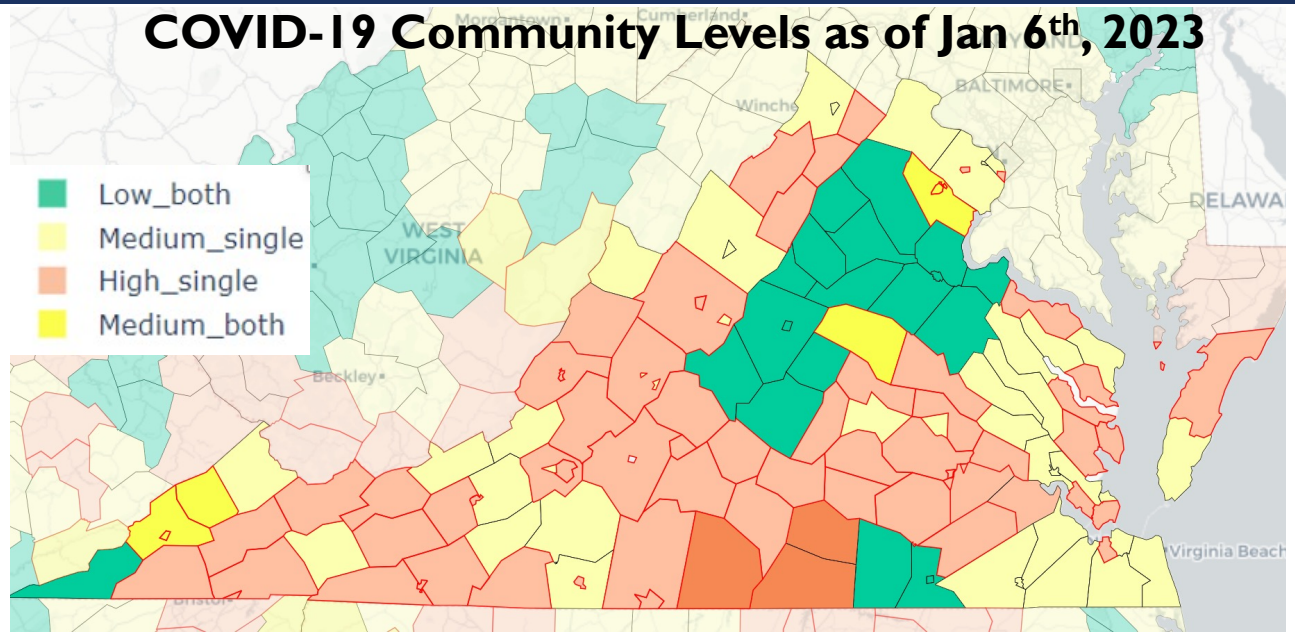
NATIONAL™ WASTEWATER SURVEILLANCE SYSTEM



https://covid.cdc.gov/covid-data-tracker/#wastewater-surveillance

https://www.vdh.virginia.gov/coronavirus/see-the-numbers/covid-19-data-insights/sars-cov-2-in-wastewater/

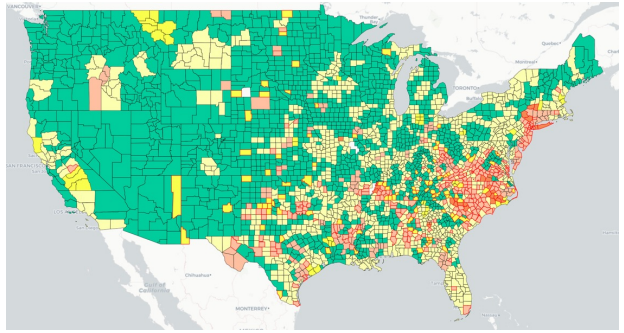
MULTIVARIATE CATEGORY DEFINITIONS

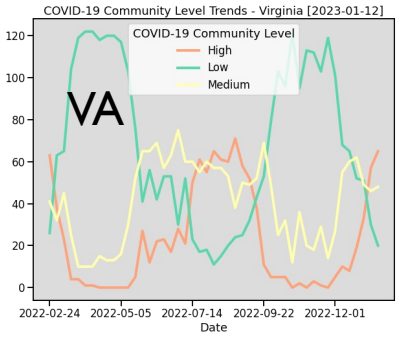


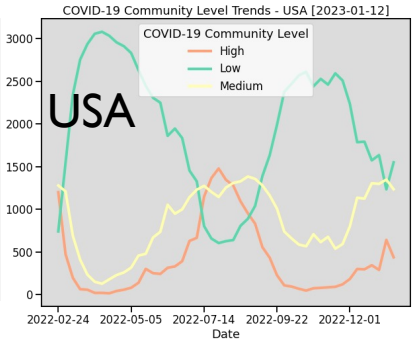
Red outline indicates county had 200 or more cases per 100k in last week

Pale color indicates either beds or occupancy set the level for this county

Dark color indicates both beds and occupancy set the level for this county





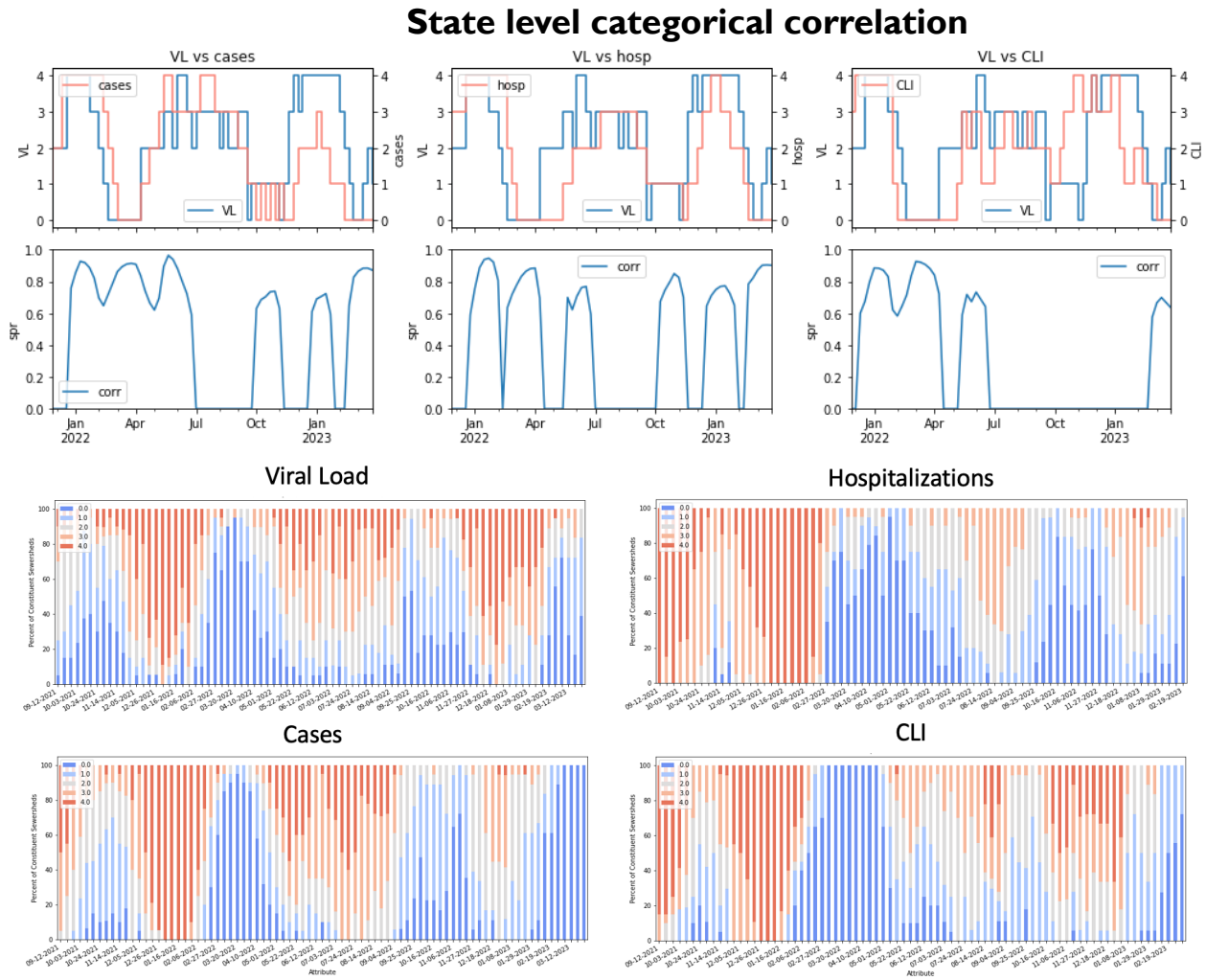


New COVID-19 Cases Per 100,000 people in the past 7 days	Indicators	Low	Medium	High
Fewer than 200	New COVID-19 admissions per 100,000 population (7-day total)	<10.0	10.0-19.9	≥20.0
	Percent of staffed inpatient beds occupied by COVID-19 patients (7-day average)	<10.0%	10.0-14.9%	≥15.0%
200 or more	New COVID-19 admissions per 100,000 population (7-day total)	NA	<10.0	≥10.0
	Percent of staffed inpatient beds occupied by COVID-19 patients (7-day average)	NA	<10.0%	≥10.0%

Data from: CDC Data Tracker Portal

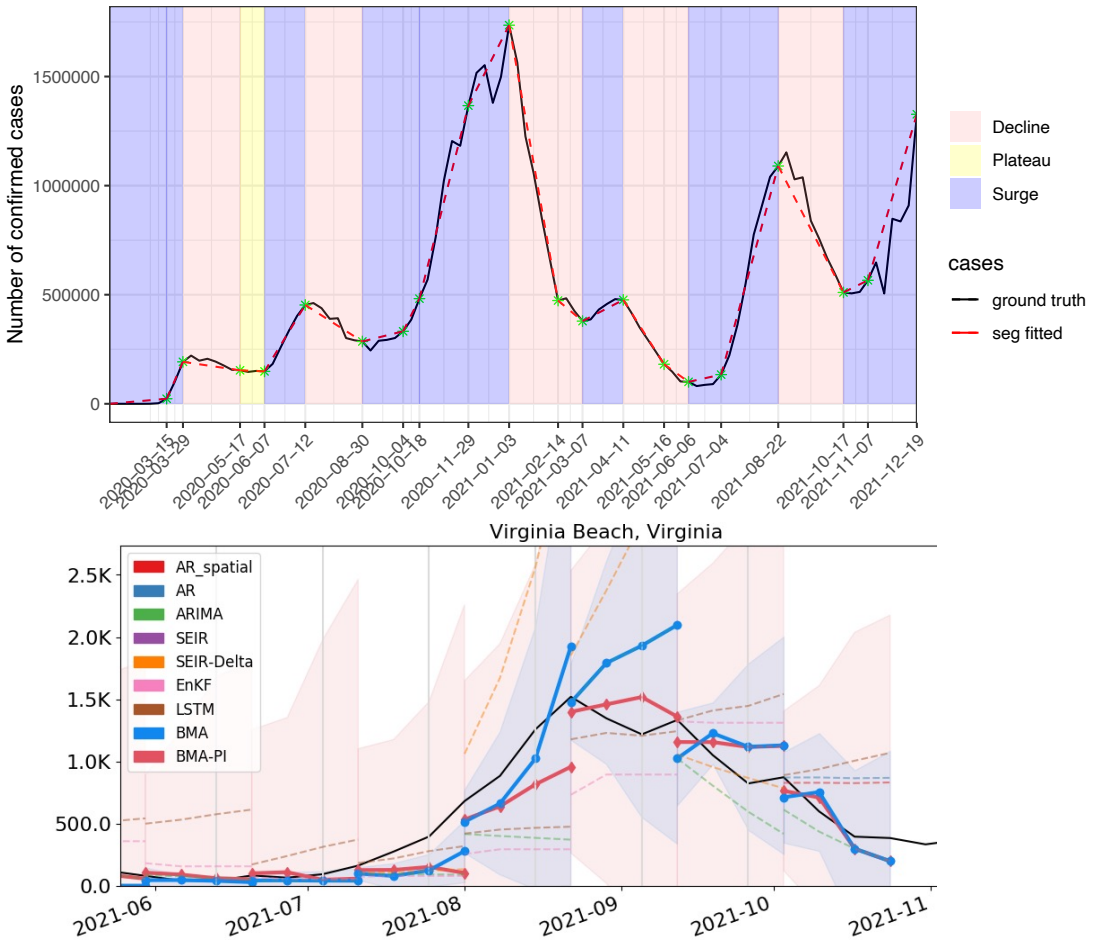
MULTI-TIME SERIES ANALYTICS: COVID-19 IN VIRGINIA

- **Source data:**
  - WW viral load by sewershed
  - Cases mapped to sewershed catchment (from line list)
  - CLI% available per health district
  - Hospitalizations by health district (from state level)
- **Categorization:**
  - Obtain quintiles from entire time series (per signal, location)
  - Categorize to level 0 through 4
- **Multi-timeseries analytics:**
  - Pairwise Spearman correlation
  - % locations at various quintiles by signal



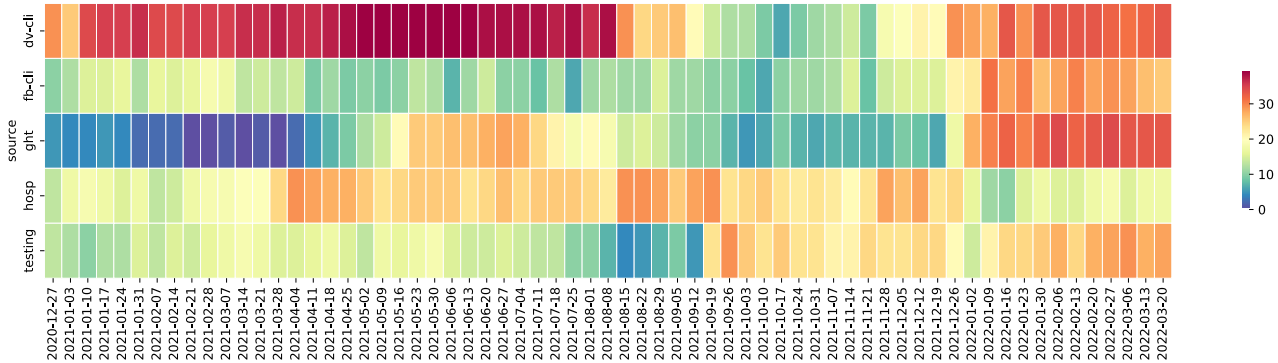
BEYOND ANALYTICS: PHASE-INFORMED FORECASTING

Phase-informed model ensembles

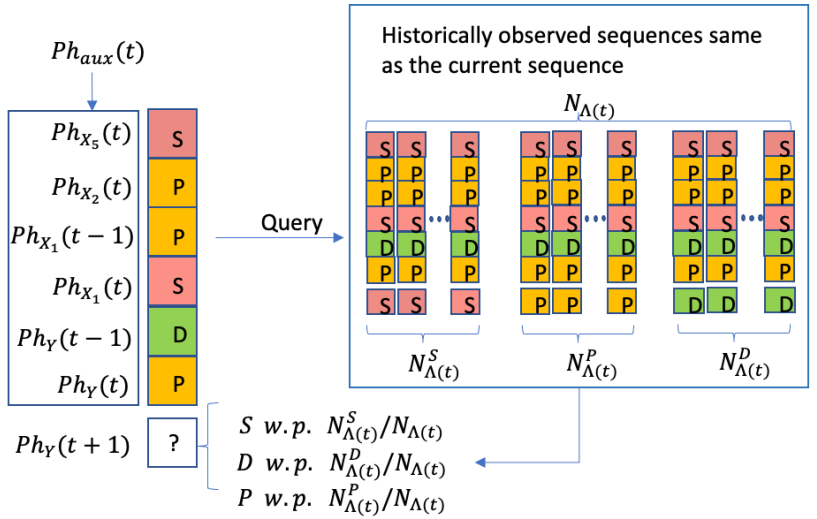


Adiga et al., Phase-Informed Bayesian Ensemble Models Improve Performance of COVID-19 Forecasts, Annual Conference on Innovative Applications of Artificial Intelligence (IAAI), 2023.

Phase prediction using leading indicators



# of states for which an individual signal was identified as a leading indicator



Adiga et al., Enhancing COVID-19 Ensemble Forecasting Model Performance Using Auxiliary Data Sources, IEEE International Conference on Big Data, 2022.

TOOL INTRODUCTION

TAXONOMY OF CATEGORIZATION RULES

- Level-based categories
  - Using equal-width intervals for the data range (e.g., ILINet Activity Indicator)
  - Using quantiles of the data to inform category bins (e.g., VDH COVID-19 wastewater surveillance)
- Trend-based categories
  - Rate change over a fixed length window (HHS hospitalization trends)
  - Rate change over automatically detected segments (VDH case trajectories)
  - Percent change in activity over a fixed length window (CDC NWSS dashboard)
- Combined rules
  - Multi-signal-based categories (e.g., CDC Community Levels)

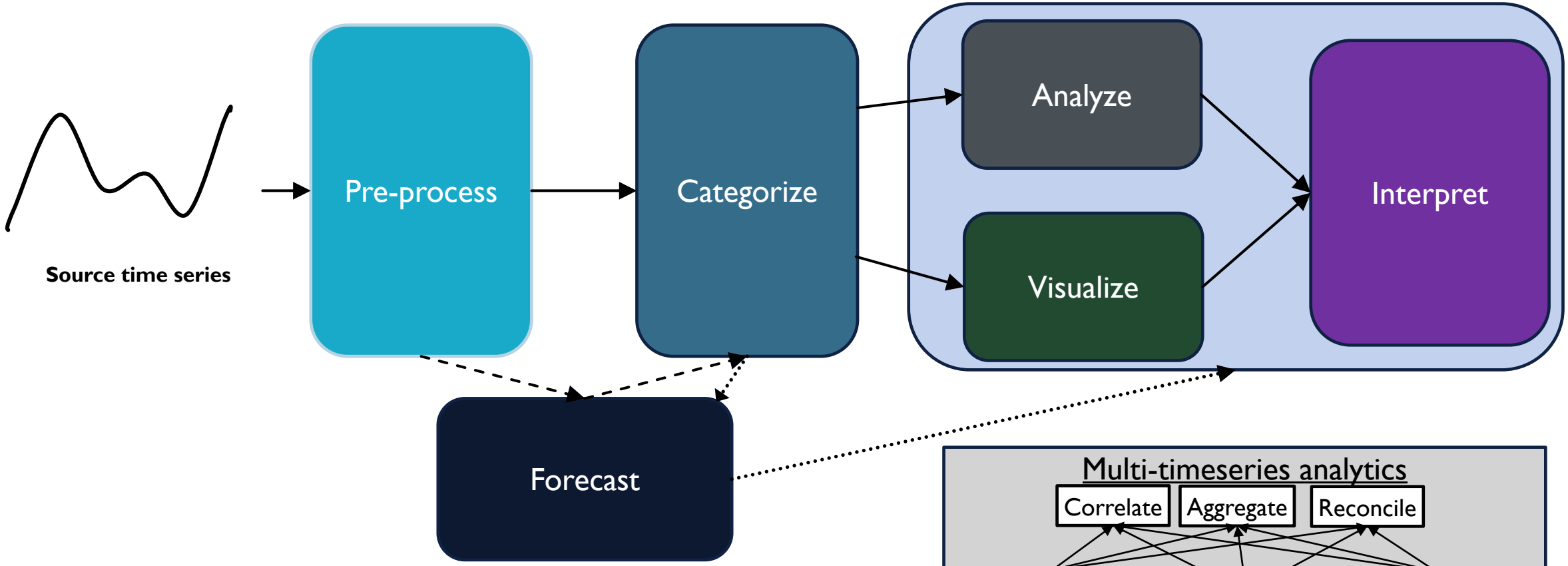
How to decide on the appropriate categorization for a given signal or region?

PEpiTA: **P**hase-based **Epi***demic* **T***imeseries* **A***nalyzer*

Phase-based Epidemic Time-series Analyzer (PEpiTA)

- Upload or choose a publicly-available time series for analysis
- Perform standard preprocessing (e.g., interpolation, smoothing)
- Choose among pre-defined rules for category extraction
- Customize analytical parameters such as number of bins, range, trend windows
- Download visualizations and analytical summaries

PEpiTA - WORKFLOW



Phase-based **Epi**demic **T**imeseries **A**nalyzer (PEpiTA)

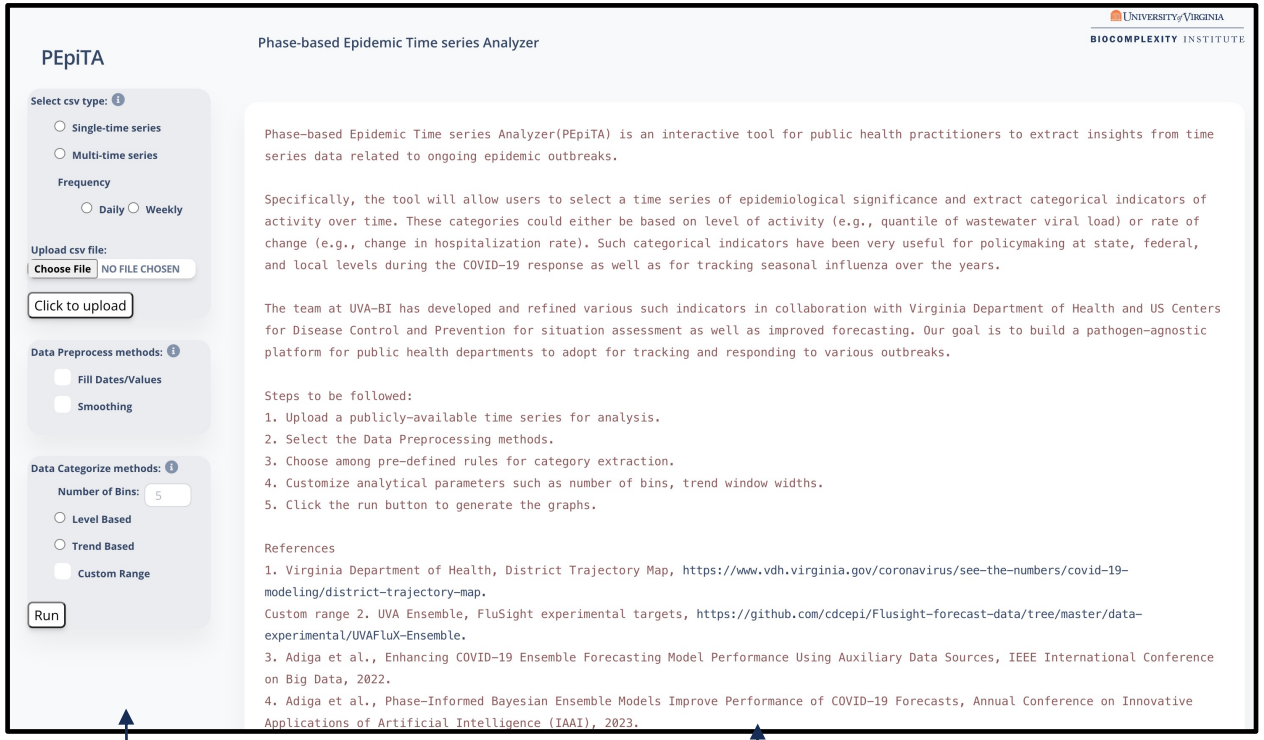
PEpiTA - INTERFACE

- Browser-based interface using Django framework
- Codebase in Python (using pandas, matplotlib)
- Setup for headless mode; Exploring installation options

https://github.com/NSSAC/PEpiTA

```
/Users/vsriniv/Documents/Research/Github/PEpiTA
(base) vsriniv@Srinivasans-MacBook-Pro PEpiTA % cd webdesign
(base) vsriniv@Srinivasans-MacBook-Pro webdesign % python manage.py runserver
Watching for file changes with StatReloader
Performing system checks...

System check identified no issues (0 silenced).
June 23, 2023 - 18:15:37
Django version 4.2.1, using settings 'webdesign.settings'
Starting development server at http://127.0.0.1:8000/
Quit the server with CONTROL-C.
```



Input pane

Landing page

PEpiTA - INTERFACE

Input file properties

Preprocessing method

Categorization method

Visual Summary

Analytical Summary

ONGOING AND FUTURE WORK

- Incorporating stakeholder feedback from VDH and other STLT agencies (like you!)
- Expanding categorization rules and making them *FAIR*

  Findable, **A**ccessible, **I**nteroperable, and **R**eusable
- Interoperability with other analytical tools (e.g., epiprocess, EpiNow2)
- Improved analytics and interactive visualizations
- Nowcasting and forecasting capabilities; Advanced multi-time series workflows

If you're interested in testing and providing feedback, please reach out to srini@virginia.edu