EPIsembleVis: A geo-visual analysis and comparison of the prediction ensembles of multiple COVID-19 models☆

Haowen Xu [*],[1], Andy Berres [*],[1], Gautam Thakur, Jibonananda Sanyal, Supriya Chinthavali

Oak Ridge National Laboratory, PO Box 2008, MS 6085, Oak Ridge, TN 37831, United States

ARTICLE INFO

ABSTRACT

We present EPIsembleVis, a web-based comparative visual analysis tool for evaluating the consistency of multiple COVID-19 prediction models. Our approach analyzes a collection of COVID-19 predictions from different epidemiological models as an ensemble and utilizes two metrics to quantify model performance. These metrics include (a) prediction uncertainty (represented as the dispersion of predictions in each ensemble) and (b) prediction error (calculated by comparing individual model predictions with the recorded data). Through an interactive visual interface, our approach provides a data-driven workflow for (a) selecting and constructing the COVID-19 model prediction ensemble based on the spatiotemporal overlap of available predictions of multiple epidemiological models, (b) quantifying the model performance using both the uncertainty of each model prediction ensemble, and the error of each ensemble member that represents individual model predictions, and (c) visualizing the spatiotemporal variability in the projection performance of individual models using a suite of novel ensemble visualization techniques, such as the data availability map, a spatiotemporal textured-tile calendar, multivariate rose chart, and time-series leaflet glyph. We demonstrate the capability of our ensemble visual interface through a case study that investigates the performance of weekly COVID-19 predictions, which are provided through the COVID-19 Forecast Hub UMass-Amherst Influenza Forecasting Center of Excellence [47] for the United States and United States Territories. The EPIsembleVis tool is implemented using open-source web technologies and adaptive system design, rendering it interoperable with Elasticsearch and Kibana for automatically ingesting COVID-19 predictions from online repositories, and it is generalizable for analyzing worldwide projections from more epidemiological models.

1. Introduction

The COrona VIrus Disease (COVID-19) epidemic began in late 2019, spreading human suffering and socio-economic turmoil around the world [48,32]. COVID-19 has posed a significant challenge to policymakers who have to decide on which mitigation strategies should be introduced (e.g., mask mandates, school closures, or lockdown), when they should be introduced, and when it is safe to lift these mandates [1]. Its infectiousness, combined with a slow onset in symptoms and the presence of a high percentage of asymptomatic individuals, have made it challenging to understand and predict the spread. Effective intervention, mitigation, and control of the epidemic require a solid understanding of the mechanism of COVID-19 that governs the pandemic's transmission, disease, and immunity [18,40], as well as precise and timely predictions of new cases and deaths. A reliable prediction needs to approach the disease from a holistic perspective that considers the interplay between the multiple variables (e.g., biophysical, social, and human) across large geographical areas [12,46].

Epidemiologists and modeling experts world-wide have risen to the challenge of developing reliable model predictions of the future COVID-

19 pandemic in terms of cases and deaths [1,22,40]. These epidemiological models are designed to clarify the extent and impact of the pandemic, providing predictions that help guide the government decisions, planning, and community preparedness in this pandemic [11,8,5].

Despite their practical value and usefulness, many of these models produce divergent and conflicting future projections [18] for some geographic areas (e.g., city, country, and state), impeding reliable decision supports for allocating resources and implementing mitigation practices. The uncertainty in these predictions often arises from the simulation process, where each prediction is computed using models based on distinct approaches (e.g., statistical and mechanistic), varied parametric assumptions (e.g., transmission and immunity), and different quality of calibration data (e.g., divergent detection of cases, reporting delays, and poor documentation) [1,18,40]. Since each model has its own unique assumptions (e.g., policies, compliance, transmission rate), constraints (e.g., geographic and temporal extent, a limited model of policy differences between areas), and predicted type of data (e.g., cases, hospitalizations, deaths, etc.), model predictions can vary depending on the amount and veracity of information about an area, but also depending on how strictly mitigation strategies are implemented, and how compliant the population is with these strategies [3].

Given the divergence in the predictions from different models, it is difficult to determine the best-performing models that produce the most reliable predictions. Analyzing the consistency of these models, which may potentially imply model performances, faces challenges due to the complex nature of these model predictions that, similar to ensemble members in many other scientific domains (e.g., climate and transportation), these model predictions are spatiotemporal, multivariate, and heterogeneous in data format and prediction coverage [24,28]. In addition, the high number of different models (over 70 for the United States alone [4]) poses new challenges, such as quantifying differences and uncertainties among multiple models, comparing their predictions against each other, and evaluating the model performance during different scenarios (e.g., in states with different population density and during different outbreak stages). All these factors make it extremely challenging to weigh models against each other at a larger geographic scale (e.g., the contiguous United States).

In this paper, we present the design and development of EPIsembleVis, an innovative visual analysis approach to evaluate the spatiotemporal variability in the consistency of COVID-19 prediction models by analyzing their results as ensembles through interactive and comparative visual exploration. The approach aims to provide data-driven insights to help epidemiologists and health care professionals explore potential factors that may affect the prediction models' consistency and select the appropriate model to support the decision-making in a time- and space-specific scenario. We develop a visual interface to provide an integrated workflow for (a) selecting and constructing the COVID-19 model prediction ensemble based on the overlapping projection availability of multiple epidemiological models in both time and space, (b) quantifying the consistency of ensemble model using both the uncertainty of each model prediction ensemble by comparing its member against each other, and the error of each ensemble by comparing its members' prediction with the recorded case information, and (c) visualize the spatiotemporal variability in the consistency of the COVID-19 model prediction ensembles (e.g., new cases and deaths) using a suite of novel ensemble visualization and user interaction techniques. The visual interface is integrated into a big-data cyberinfrastructure powered by an Elasticsearch-Kibana stack, which adopts an ontology-driven data pipeline to automate data mining, transformation, and enrichment processes. These processes aim to retrieve model prediction data and recorded case data from COVID-19 Forecast Hub [4] and provide the most updated record to the visual interface in a near-real-time fashion.

To demonstrate the capability of the EPIsembleVis, we include a case study that investigates the consistency of weekly COVID-19 predictions, which are generated at the state level from 4 popular predictive models

for the contiguous United States. By overlaying our ensemble visualizations with other spatial information, we are able to obtain some data-driven insights regarding the empirical relationship between the model uncertainty and population density. Developed with adaptive design and flexible architecture, the data-driven visual interface of EPIsembleVis is generalizable and extendable, and can support visual analysis of ensemble COVID predictions that are produced in other geographic entities (e.g., other countries and continents), from more epidemiological models, and at alternative spatial (e.g., county and city) and temporal (e.g., monthly and daily) scales.

In summary, our contributions are as follows:

1. An big-data cyberinfrastructure that adopts an ontology-based data pipeline to create a COVID-19 Forecast Ensemble datasets.
2. A visual representation that overviews necessary information for the ensemble analysis, such as the availability and predicting parameters of individual models, at a glance.
3. A textured tile-based calendar that highlights spatiotemporal patterns of differences between predictions from selected models.
4. Two glyph-based representations of model-outputs which enable the comparison between different model predictions, as well as comparison with recorded cases in both the spatial and temporal dimension simultaneously.

2. Related work

A comprehensive evaluation of the COVID-19 prediction model consistency is a complex effort that requires solid understanding and knowledge in both the epidemiological modeling, and ensemble model analysis and visualization. Based on these two types of effort, we divide our related work section into the following subsections.

2.1. COVID-19 prediction models

With the global propagation of the coronavirus disease 2019 (COVID-19), epidemiologists worldwide are rushed to develop epidemiological models for forecasting the future of the pandemic [22,1]. Despite their capability of producing quantitative projections of infections and mortality estimates, these models and their performances are usually affected by a set of model assumptions and configurations. As many aspects of the COVID-19 epidemics still remain unknown and need to be assumed [22], it would not be sufficient and reliable to rely on the forecasts of a single model that is developed based on a specific set of assumptions and calibration data with limited quality and quantity.

There are two distinct branches in epidemiological modeling: Mechanistic models [15] and Stochastic models [18]. Each type of epidemiological model has its strengths and limitations [29,11,52], and it is critical to compare results from different models to gain a better understanding of potential future behaviors.

2.1.1. Prediction ensemble

Given the fact that many epidemiological models produce conflicting projections and, according to many modelers, may have varying performances for making forecasts [1,18], one approach to reduce and analyze the uncertainty in these models is through the development of prediction ensemble, through which models of different types (with different configuration and assumptions) are executed to generate multiple realizations of the same projections [49]. The prediction ensemble approach usually produces a collection of spatiotemporal outputs, each of which is generated by a single model run and is defined as an ensemble member [37,2]. Currently, ensemble data analysis and visualization are widely used in many domains, such as climate science and oceanography, to help scientists model complex systems, reduce uncertainty, and explore sensitivity to different model parameters, assumptions, and initial conditions [37,36,17,39].

Ensemble datasets are an increasingly common tool to help scientists

simulate complex systems, mitigate uncertainty, and investigate sensitivity to parameters and initial conditions. These spatiotemporal datasets are large, multidimensional, and multivariate. Due to their complexity and size, ensembles present challenges in data management, analysis, and visualization. Currently, comprehensive ensemble data analysis and visualization that evaluates the consistency for COVID-19 forecasting models in the health geography sector is still rare. The most relevant application would be the COVID-19 Forecast Hub [26,4], which focuses on the preparation and compilation of COVID-19 ensemble projections from different models, and provides basic model comparison capabilities that only focus on a single data dimension (e.g., temporal or spatial) using line charts.

2.2. Ensemble visualization

Effective ensemble visualizations help domain scientists gain a more intuitive interpretation and understanding of patterns within a complex dataset [49]. By nature, ensemble data are typically large, multivariate, and multivalued, and can be defined for multiple data dimensions [20,28,38]. Due to this complexity, ensemble data often entail multiple facets that need to be considered simultaneously during the analysis and visualization. They are challenging to explore and comprehend using a single visualization technique, which usually covers only one or two facets [49,24]. In addition, the unique member dimension in the ensemble data cannot be efficiently represented through traditional visualization techniques, posing further challenges in data analysis. Given these challenges, a variety of analytical and visualization techniques have been proposed to reduce the complexity and dimensions in ensemble datasets, characterize uncertainty, and evaluate the accuracy and reliability in the data ensemble [36,52,20,39].

One of the most recent comprehensive reviews of the ensemble visualization and visual analysis is offered by [49], who summarized a wealth of past applications that utilize combined visualization techniques (e.g., vectors, color maps, glyphs, maps, and time-series) to simultaneously cover multiple facets and dimensions of the ensemble data. Examples of these applications include the visualization of ensemble uncertainty in (a) the spatial, ensemble, and multivariate dimensions [20,35,21], (b) spatial and temporal dimensions [16,17,41], and (c) temporal and ensemble dimensions [33,25].

Despite useful developments in the ensemble visualization domain, many of these past applications were developed as either desktop-based or traditional web-based applications. They often face software engineering challenges regarding system adaptability and frequently updated new data that are available through various online repositories and cyberinfrastructure. The above-mentioned challenges limit the capability of these applications for conducting ensemble analysis on time-critical data, such as COVID-19 model projections. In addition, many of the past applications rely on the coordination of multiple visualizations in the visual interface to cover multiple dimensions of the ensemble data. This requires a significant amount of interactions between the user and the visual interface. In this setting, the visual analysis tool might not be very intuitive for non-expert users. Since the ultimate objective of analyzing COVID-19 prediction ensembles is to enable timely and accurate insights, and decision supports for preventing the further spread of the disease, it is vital to have ensemble visualization platforms which are intuitive and informative. In this regard, a single innovative visual representation, which can reveal multiple facets of the ensemble data at the same time, is preferred than multiple visualizations that require a moderate amount of users' effort for interaction and interpretation.

3. Overview

This section will provide an overview of EPIsembleVis at the example of weekly forecasts at the state level. However, these concepts translate well to other spatiotemporal granularities. We will discuss any required modifications in the corresponding subsections. At the temporal level, we focus on weekly forecasts as this allows us to bypass the stark daily variations (especially in ground truth data), and it is consistent with outputs from a large number of forecasting models, as we will discuss in Section 4. At a spatial level, we focus on states because counts are less susceptible to noise than county level, and many regulations are implemented for specific states, which provides a more consistent basis of comparison. In addition, we also aggregate to geographic regions to provide relevant geospatial visualizations at a coarser level of detail. The individual visualizations as well as their results are discussed in Section 5.

3.1. Framework and visual workflow

This subsection presents the technical details that are related to the implementation of the EPIsembleVis. The developmental efforts of the visual analysis tool can be divided into two parts: (1) data compilation and indexing on the server-side, and (2) the visual interface design on the client-side.

The data compilation and indexing are achieved through a big-data cyberinfrastructure that is powered through the combination of Elasticsearch and Kibana. The Elasticsearch is a distributed, multitenant-capable full-text search engine that is able to conduct the web mining, ingestion, and archiving of large-scale data in an automated manner. To facilitate the data retrieval from and log management of the Elasticsearch, the Kibana online data visualization platform is designed to serve as an API end-point for querying and access the tremendous amount of web-mined data gathered through the Elasticsearch. Taking advantage of this cyberinfrastructure, we developed our EPIsembleVis application as a Kibana visualization plug-in. This implementation provides an automated data provision pipeline that seamlessly connects the EPIsembleVis with the Elasticsearch, rending the visual interface flexible and generalizable. Through this connection, the visual interface is able to access and analyze the latest COVID-19-related data resources (e.g., model projections and ground truth information) from any online repositories in near-real-time [44].

The novel and customized spatiotemporal visualizations presented in this paper are implemented using D3JS JavaScript libraries that utilize both HTML5 Canvas graphics and Scalable Vector Graphics (SVG) elements to render interactive visual representations both in the interface and on the web map. The interaction between the user controls that enable the selection of variables and models and the coordination of various linked visualizations is created through the direct manipulation of the HTML Document Object Model (DOM) elements using both D3JS and JQuery JavaScript libraries.

The visual interface is developed using adaptive and interoperative web technologies and design patterns, and therefore is generalizable and scalable for conducting comparative visual analysis for prediction datasets that contain predictions from more models or in other countries. The visual interface can be readily integrated into major big-data analytical platforms, such as Grafana and Kibana, as a plugin.

3.2. Data acquisition

Our main data source is the COVID-19 Forecast Hub data [47] which is provided through the Center for Disease Control's forecasting hub [4]. The forecast hub collects data from a large number of different models which provide forecasts of cases, hospitalizations, or deaths at national level (for the United States), by state, or by county.

During the initial development of this work, we created preliminary analyses and visualizations for a mixture of public and non-public model forecasts: EpiGrid and EpiCast [11], LANL COVID-19 Cases and Deaths Forecasts [27], and IHME [19], to test the concept of the EPIsembleVis. We will focus this paper on the forecast hub data selection, since two of these models (IHME and LANL) are part of the forecast hub data, access to EpiGrid and EpiCast is manual, and EpiGrid is not routinely run for a

large number of states.

As ground truth data, we use the Johns Hopkins COVID-19 Data Repository [10,23], as well as the New York Times Data [31]. Both datasets are also available through the forecast hub [47].

As the first step of our data workflow, we pull an automatic update from the forecast GitHub [47], and optionally, other model or ground truth sources.

4. Ontology-driven data compilation

The COVID-19 Forecast Hub dataset [47] contains predictions from 70 different models at the time of writing. Each of the datasets in the collection contains predictions for weeks (58 models), days (1 model), or both (11 models). The predictions can be for cases (30 models), hospitalizations (12 models), or deaths (65 models), and each of these variables can be provided as incident data (daily new incidents of each variable), cumulative data, or both. All 70 models provide a point estimate, while some models provide additional ensembles of different quantiles.

The data is available at the following geographic levels

- National (53 models): For this subset of data, it is not clear which states and/or territories are included.
- State (63 models): The number of states included varies between 1 and 56 states. 24 models include all 50 states with or without the District of Columbia, 19 models include some or all U.S. Territories,

and 18 models include less than 50 states (12 of these models include less than half of all states).
- County (19 models): Again, there is some variation in how many counties are included. 11 models include more than 3,000 counties (50 states with D.C. have 3,141 counties), and 4 models include less than 1,000 counties.

Most model data are updated daily or weekly, with a new set of predictions for the following days/weeks. However, this does not always happen consistently that the temporal resolution of these models may change during different time periods.

To create a better understanding of data availability, we created an overview, which combines model updates and variable of COVID-19 data (cases, hospitalizations, deaths). The data availability map in Fig. 1 shows an overview of when models were updated, aggregated to a weekly level. Weeks are represented as columns, and models are represented as one row. For weeks in which the model was updated, the corresponding cell is colored to represent which data types are available. With this representation, it is easy to see at a glance how regularly models update and how long they have been part of this dataset. We will discuss this visualization in more detail in Section 5.1.

4.1. Data preparation and fusion

The format for model outputs in this dataset is consistent across all models, which enables us to use a single pipeline for data compilation.
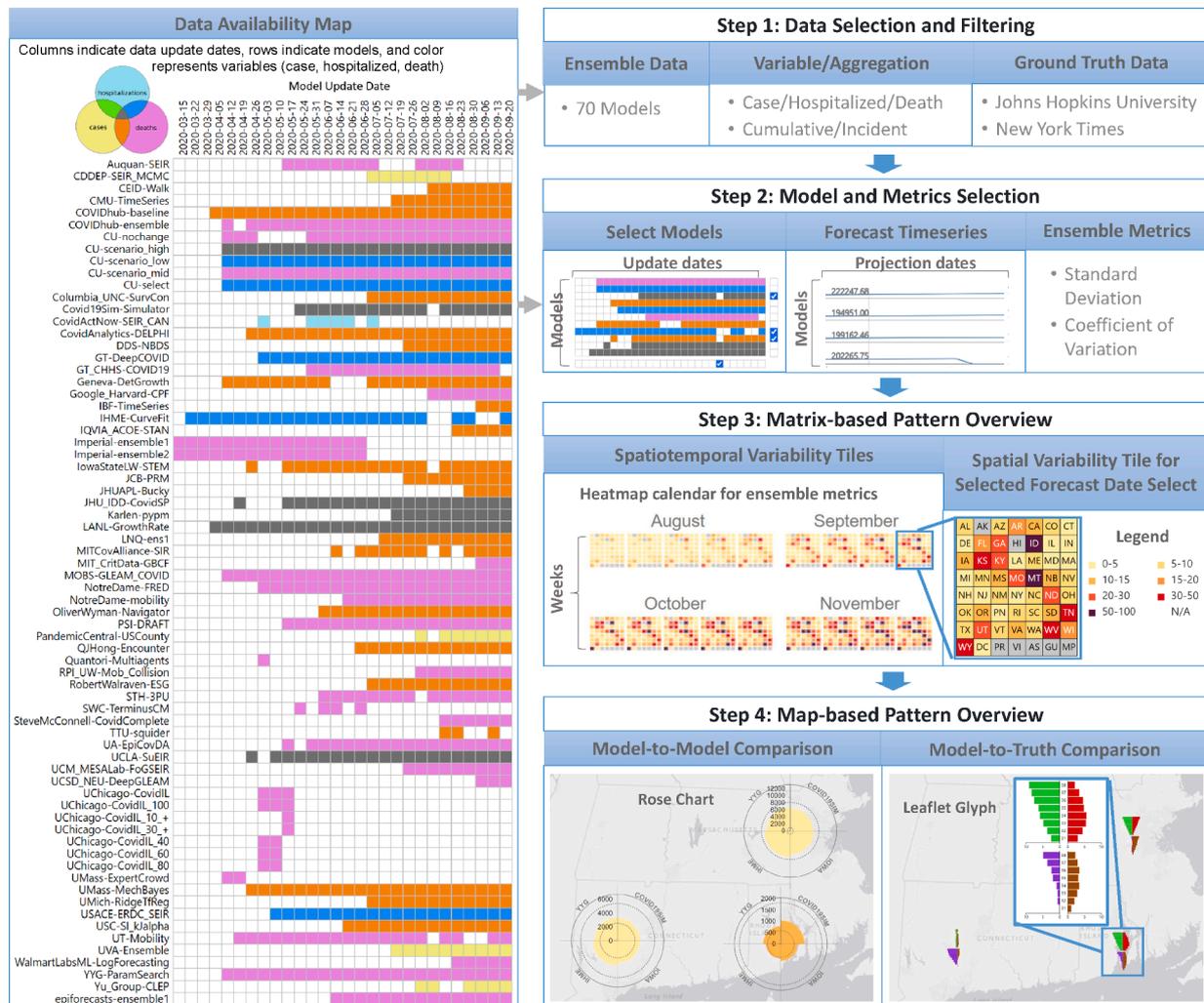

Fig. 1. The overview and general user Workflow of the visual interface.

However, due to the many options provided to accommodate different prediction parameters, it is rather complex and requires several processing steps to produce easily ingestible data for visualizations.

As the first step in data preparation, we prepare a data dictionary that holds all relevant model properties, file paths, and the number of days or weeks the models project.

We are particularly interested in comparing different models, rather than different projections by the same model. Therefore, in the next step, we filter the model outputs to contain only point estimate data, which is available from all models, instead of using the quantile data that is only provided from a few models.

Finally, we filter each model output to contain only outputs at the state level. As we are considering data for the United States and its territories, one single geographic entity is too coarse, and it can be aggregated on-the-fly. Most models that provide county data also provide state data. For the few exceptions, the county data can be aggregated up to states very easily.

4.2. Spatial and temporal aggregation

The next step of our data workflow is to develop data aggregation for each model. As the models in this dataset update on different days of the week, it produces a temporal heterogeneity in the ensemble dataset that prevents the comparison between models. To facilitate comparisons, we summarize them by week. The majority (70%) of updates happen on Sundays or Mondays. To ensure the integrity of the datasets and remain close to the conventions that are adopted by disease control professionals [7], we choose to start weeks on Sundays. As for week numbers, we assign ISO week numbers that apply to all days except Sunday, which are part of the prior week under the ISO definition but assigned to the following week for this aggregation. The data availbility map in Fig. 1 uses this aggregation to represent the frequency and time of model updates in an intuitive format, which makes it easy to compare models. It also serves as a basis for data selection for ensemble comparisons.

As part of the aggregation process, we also lay the foundation for spatial aggregations (county-to-state and state-to-region) by adding labels for regions and states to all data. The aggregation itself will be done through ElasticSearch and Kibana queries based on user interactions.

4.3. Data enrichment

Each model has a specific set of predictions. These predictions can have different combinations of aggregation type and a variable. Aggregation types are either cumulative or incident (daily update), and variables are the case, hospitalization, or death. However, none of the models has each possible combination.

To increase the comparability of models, we add missing aggregation types for each variable. To get incident data $i(t)$ from cumulative data $c(t)$ for each location, we use a simple subtraction of the previous date $(t-1)$.

$$i(t) = c(t) - c(t-1)$$

For cases in which we only have predicted incident data ($i(t)$) and want to get the cumulative data, we use the number of reported cases ($g(0)$):

$$c(t) = i(t) + g(0)$$

The resulting datasets with matching cumulative and incident data for each available variable are ingested into the ElasticSearch database.

Through a series of aggregation processes, we are able to uniform the variable type (e.g., death, case, and hospitalization), as well as spatial and temporal enumeration units of COVID-19 projections (e.g., new cases for each state and in each week) from various epidemiological models. By grouping these aggregated projections based on the same type, state, and week, we can readily construct COVID-19 model prediction ensembles, in which individual model predictions are serving as an ensemble member.

4.4. Ensemble metrics

Statistical aggregation is often used to create metrics for characterizing the ensemble data in many past studies. In this study, we use two metrics to quantify the consistency of multiple models, which include (a) the uncertainty among multiple ensemble members within a model prediction ensemble and (b) the accuracy of the prediction of an individual ensemble member compared to recorded case numbers. For the purpose of metric definitions, we use x_i to denote recorded case data, and $\widehat{x}_i$ to denote predicted data, where x_i and $\widehat{x}_i$ can be cumulative or incident data for any of the variables (case, hospitalization, death).

4.4.1. Ensemble uncertainty

Given the multivariate nature of each model prediction ensemble, we utilize two statistical aggregates to quantify the prediction uncertainty within each ensemble. These statistical aggregates include (1) the Standard Deviation (STD) and the (2) Coefficient of Variation (CV) of each ensemble, and are calculated using the predictions produced from individual ensemble members. The STD is able to characterize the absolute variations among the predictions $\widehat{x}_i$ from multiple ensemble members, therefore is suitable for quantifying the uncertainty of ensembles with a similar magnitude of predictions. For an ensemble with N models for a model run date τ and a prediction date t, we get:

$$\text{STD}(\tau, t) = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (\widehat{x}_i(\tau, t) - \text{mean}(\widehat{x}(\tau, t)))^2}$$

In contrast, the CV is calculated by dividing the ensemble's STD by its mean, thus provides a normalized characterization of the relative variation (in percentage) among multiple ensemble members. In this regard, the CV allows the comparison of the uncertainty between two ensembles with different magnitude of predictions, such as comparing the death prediction for Illinois (with the magnitude of ten thousand) with that for West Virginia (with the magnitude of hundreds).

$$\text{CV}(\tau, t) = \frac{\text{STD}(\tau, t)}{\frac{1}{N} \sum_{i=1}^{N} \widehat{x}_i(\tau, t)}$$

In essence, consistent model ensembles are expected to have consensus projections from its members, therefore they should have relatively low STD and CV.

4.4.2. Prediction error

To determine the accuracy of individual models, we compute the relative error E between model prediction ($\widehat{x}$) and recorded cases (x), i.e. $E(\tau, t) = \frac{\widehat{x}(\tau, t) - x(\tau, t)}{x(\tau, t)}$. This metric provides us a sense of how accurate each model predicts, meanwhile quantifies the accuracy using an easy-to-interpret numeric parameter with linear scaling. For example, if there are 100 recorded cases and the model predicts 80 cases, $E = 0.2$. For a prediction of 200, $E = -2$.

5. Visualization

In this section, we discuss visualizations of different aspects of the data, including the range of model run dates (dates on which models were updated), forecast dates (created in a model run), and forecast data itself. We present visualizations which compare different model outputs within an ensemble with each other, and with recorded case data.

5.1. Data availability map

The data availability map represents the dates of model runs (τ) for

all models, as well as the type(s) of predictions each model contributes, as introduced in Section 4. Its name is inspired by heatmaps, which are a common visualization tool for matrix-like data. As seen in Fig. 1, it is easy to see at a glance which data have gaps in their updates, and how consistently they have been updated. For easier navigation, we offer options that allow users to sort the list of models by model name and prediction type. As a colormap, we have chosen a trivariate color scheme, which can be represented as a Venn diagram. Colors of models which only provide one type of prediction is kept light (yellow, pink, cyan), colors for models with two types of predictions are darker (blue, green, red), and models with all three predictions are darkest (gray).

In addition to serving as an overview visualization, the data availability map is also a key element to navigating the data. Users can sort the list of models by dataset name or by the combination of variables in the model. This helps to find a specific model by name, and it enables the identification of models with similar variables (e.g., case, death, hospitalization). On the right side of the data availability map, an array of checkboxes serve as the model selection tool. Similarly, the checkboxes that are placed below the data availability map serve as date selection for the model run.

5.2. Model selection

For the purpose of this paper, we compare four models. Unless noted otherwise, all Figures display data from these models. The choice of model is based on their long predictive windows, the large overlap between the different models' prediction dates, and diversity in approach (both mechanistic and statistical models are represented in this sample). The selected models are listed below by the name provided in the Forecast hub [4].

- **Covid19Sim-Simulator** Chhatwal et al. [6]: This is a compartment model that uses the SEIR compartments [15] with continuous-time progression. It uses state-specific inputs from JHU [23] and The COVID Tracking Project [45].
- **IHME-CurveFit** [30,19]: This model utilized non-linear mixed-effects curve-fitting to predict death rates [23] based on the ratio of reported COVID-19 deaths (compared to a baseline death rate, and models health service utilization as a function of deaths (based on hospital capacity and utilization data)).
- **IowaStateLW-STEM** [50,51]: This is a non-parametric spatiotemporal model for disease transmission to study COVID-19 spread at the county level. It uses the New York Times COVID-19 dataset [31] as well as information from health department webpages about county-level infections and deaths to predict cases and deaths.
- **YYG-ParamSearch** [14]: This model uses a hyperparameter optimization approach, which minimizes the error between reported deaths and model predictions. It includes some fixed variables based on literature (e.g., latency and infectious period), as well as optimized variables (e.g., mortality rate, initial and post-lockdown R_0).

In addition to reported deaths from JHU [23], YYG-ParamSearch includes data about individual state-by-state reopenings.

5.3. Spatiotemporal variability tiles

The purpose of spatiotemporal variability tiles is to show a high-level overview of how model predictions and ensemble metrics (color) evolve over time (tiles) for different states (sub-tiles). This visualization had its inception as a calendar view for ensembles of daily prediction data for EpiGrid [11], IHME [30], and LANL Growth Rate Model [34], seen in Fig. 2a. Each day is represented as a textured tile that displays data for each state. The tiles are arranged in a calendar shape to provide an intuitive yet compact representation of variation over time. To reflect the predominant model prediction resolution (weekly) in this ensemble of models, we modified this calendar view to display a truncated view that contains one tile for each week of data. This is demonstrated in Fig. 2b for an ensemble of the Covid19Sim-Simulator [6] and IowaStateLW-STEM [50,51] data provided in the Forecast Hub collection.

When a user clicks on a tile in the calendar view, a detailed version of the tile is displayed, as depicted in Fig. 2c. This selection is also used to choose the date for rose charts (Section 5.5). Each tile is a 7×8 matrix, which contains one sub-tile per state, including the District of Columbia and United States Territories. While the small version only uses color to represent each state, the large version is labeled with each entity's postal abbreviation. If a single model is selected, the color reflects the values of the chosen model and variable (e.g., cumulative cases). However, if multiple models are selected, it shows the selected ensemble metric for the ensemble of chosen models. We use a colormap with a scale from light yellow to dark red, with missing data displayed in gray.

To learn more about the exact numbers for each state, a user can click its sub-tile. This will display the exact value for each model on the selected date, as well as the ensemble metric that overviews each selected model. The user can also highlight the corresponding rose chart or leaflet glyph on the map by hovering over a state.

5.4. Leaflet glyph

The purpose of this visualization is to allow users to simultaneously compare the predictions from multiple models (shown in different colors) in an ensemble with the recorded data as a time-series (vertical axis). Through an interactive user workflow that is developed using the level of detail technique, the user is able to view both an overview of the model consistency across a large geographic area, as well as detailed deviations between individual models and the recorded data. The details of the workflow are as the flowing:

Step 1: select models for ensemble comparison; Step 2: present overview of the spatiotemporal comparisons between individual ensemble members (case predictions) against recorded data in the web-map using leaflet-glyph visual representation; Step 3: present detailed leaflet-glyph representation when the user select a specific state. The
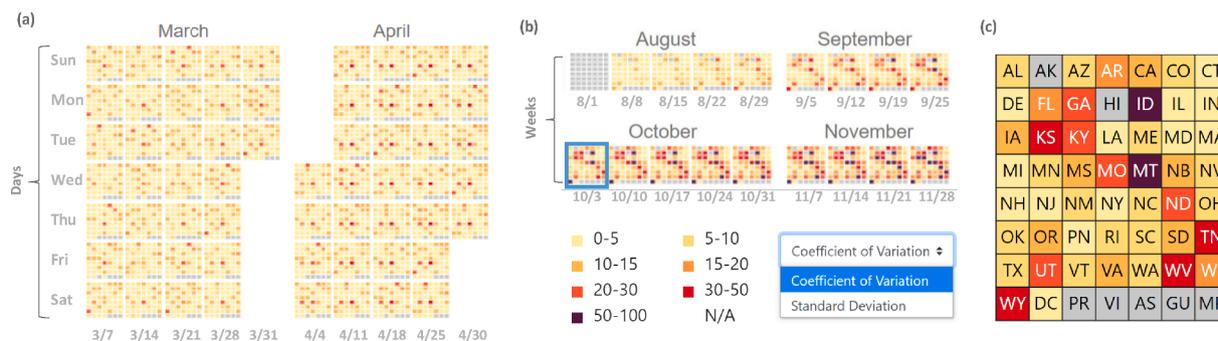

Fig. 2. Spatiotemporal tiles in a calendar view for daily predictions (a), a weekly view for weekly predictions (b), and a detailed view of a single tile (c).

layout for leaflet glyphs mirrors that of the rose charts: we display each state's leaflet glyph of the selected models on the map, and we provide a close-up, more detailed version of the glyph next to the map when the user hovers over one of the glyphs.

Each leaflet glyph is built up as follows: The vertical axis displays weekly predictions, labeled with the week number of the model prediction date. For the data presented here, there are 8 weeks of overlap between model predictions and recorded data. On the horizontal axis, we display the prediction error of all model predictions. In these graphics, we display $\text{abs}(E)$, however, a modification which applies a texture to columns with negative values is under development. Models are aligned in pairs to the left and right of a central axis, and each model is shown in a different color. The user can choose one of several colormaps to differentiate between models.

Fig. 5 displays some examples of the glyphs at full level-of-detail, which will be discussed in Section 6.

We provide two different views of these glyphs: a local view and a global view. With the local view (Fig. 3a), users can compare different model predictions of deaths for a single state. The axes for all sections of a single leaflet glyph are scaled to be identical, but each state has independent axes. In this example, one can see that Covid19Sim (green) and IHME (purple) have very little difference from the recorded deaths, whereas the Iowa model has the biggest difference within this ensemble. With the global view (Fig. 3b), the leaflet glyphs for all states share the same axis. This enables users to compare how accurate model predictions are for different states. The leaflets for some states, such as most states in the Northeast, vanish to almost a line. The states which stand out with much higher discrepancies are predominantly sparsely populated states.

5.5. Rose charts

The purpose of this visualization is to compare different models in an ensemble with each other, both in a quantitative way (size of petals representing model predicted values) and a qualitative view (color representing disagreement). Rose charts are a well-known glyph-based visualization technique for comparing different data, in which each variable (here: model prediction) is displayed as a petal in a polar coordinate view. For each state, we display a rose chart of the selected models on the map. When the user hovers over one of the charts, a close-up, more detailed version of the chart is displayed next to the map. The user can choose one of the metrics (coefficient of variation or standard deviation) as a display option. The date for this step is selected using the temporal variability tiles.

Each rose chart is built up as follows: Each model is represented as a section of a circle (or *petal*), and the abbreviated model name is displayed around the perimeter of the close-up version. The radius of the petal represents the model's predicted value. Concentric circles aid in reading the numbers. An example of this can be seen in Fig. 4 The color of the slices represents the agreement between models. Light yellow indicates agreement (low coefficient of variation or standard deviation), and red indicates disagreement between models (high coefficient of variation or standard deviation). When rendered on the map, we reduce the level of detail to just petals, crosshairs, and one concentric circle to minimize visual clutter.

For example, a set of models with high agreements will look like a light yellow circle (e.g., Pennsylvania in both projections). A set of models with strong disagreement will have distinctly differently-sized, red slices (e.g., West Virginia in the 4 week projection). In between these extremes lies a range of darker yellows and oranges (e.g., West Virginia in the 8 week projection) with increasingly mismatched slices the closer the color gets to red.

5.6. Limitations

The limitations that are associated with the rose chart and leaflet glyph include (1) resulting visual clutter that may hinder the overview and visual exploration of the pattern when visualizing sub-county level COVID-19 predictions, and (2) providing a limited capability for visualizing a large number of model projections, which exceed the maximum number of screen pixels.

6. Case study and results

The following case study aims to demonstrate the capabilities and effectiveness of our visual analysis for (a) revealing the spatiotemporal variability in the consistency of COVID-19 multi-model prediction ensembles, and (b) providing data-driven insights that enable users to explore potential contributing factors that may affect the performance of model prediction ensembles.

In this case study, we analyze and compare the weekly predictions from four COVID-19 prediction models across the contiguous United States. These models include COVID-19 Simulator [9], IHME-CurveFit [13], IowaStateLW-STEM [43], and YYG-ParamSearch [42], and their prediction results are aggregated to state level. We justify our selection of models using the following reasons that (a) all 4 models provide weekly death predictions, and (b) their prediction availability (spatial and temporal coverage) has the maximum overlap (13 weeks in total across the contiguous United States).

From an ensemble uncertainty perspective, we can observe empirical relationships between the prediction ensemble's uncertainty and the geographic distribution and density of the population (detailed in Fig. 4) from the rose chart that visualizes the dispersion in each ensemble prediction. We chose the NASA night-light map as a base layer for this visualization as it gives a good indication of population density (i.e., densely populated areas are lit up, whereas sparsely populated areas remain dark). As one can see in this Figure, most of the states with high disagreements between models (high uncertainty) are sparsely populated. Furthermore, one can see that between the earlier prediction (Fig. 4a) and the later prediction (Fig. 4b), the uncertainty rises, as indicated by the darker coloring of most rose charts. This effect is particularly strong in West Virginia (circled blue). More spatial data layers (e.g., regional mobility and the implementation of mitigation strategy) can be readily integrated into the web map (developed using the Leaflet map engine) to provide additional data-driven insights that are associated with other social-economic aspects.

From a model consistency perspective, we visualize the time-series of model prediction error by comparing each prediction against the recorded data from corresponding calendar weeks using the leaflet glyph (as depicted in Fig. 5). This time range is determined by the availability in both the model prediction and recorded data. Another interesting observation is that the leaflet pattern for Texas (triangles pointing up) is very different from its surrounding states. This pattern indicates the prediction errors of all 4 models are highest at the beginning of the time range (week 31) and gradually decrease as weeks elapse. On the contrary, the patterns of Missouri and Arkansas are in the shape of a downward-pointing triangle, indicating that the model prediction errors in these states gradually increase over time (from week 31 to 38).

To enable comparisons between different states, we set the leaflet glyph to visualize global differences, the maximum in the x-axis of each glyph is the same. In this setting, we can observe that the 4-model predictions have a very small error in Louisiana, as seen by the leaflet's almost linear shape. Oklahoma and Missouri have relatively low errors, as indicated by their slim leaflets, whereas the errors are high for most models in Arkansas, Kansas, and Texas.

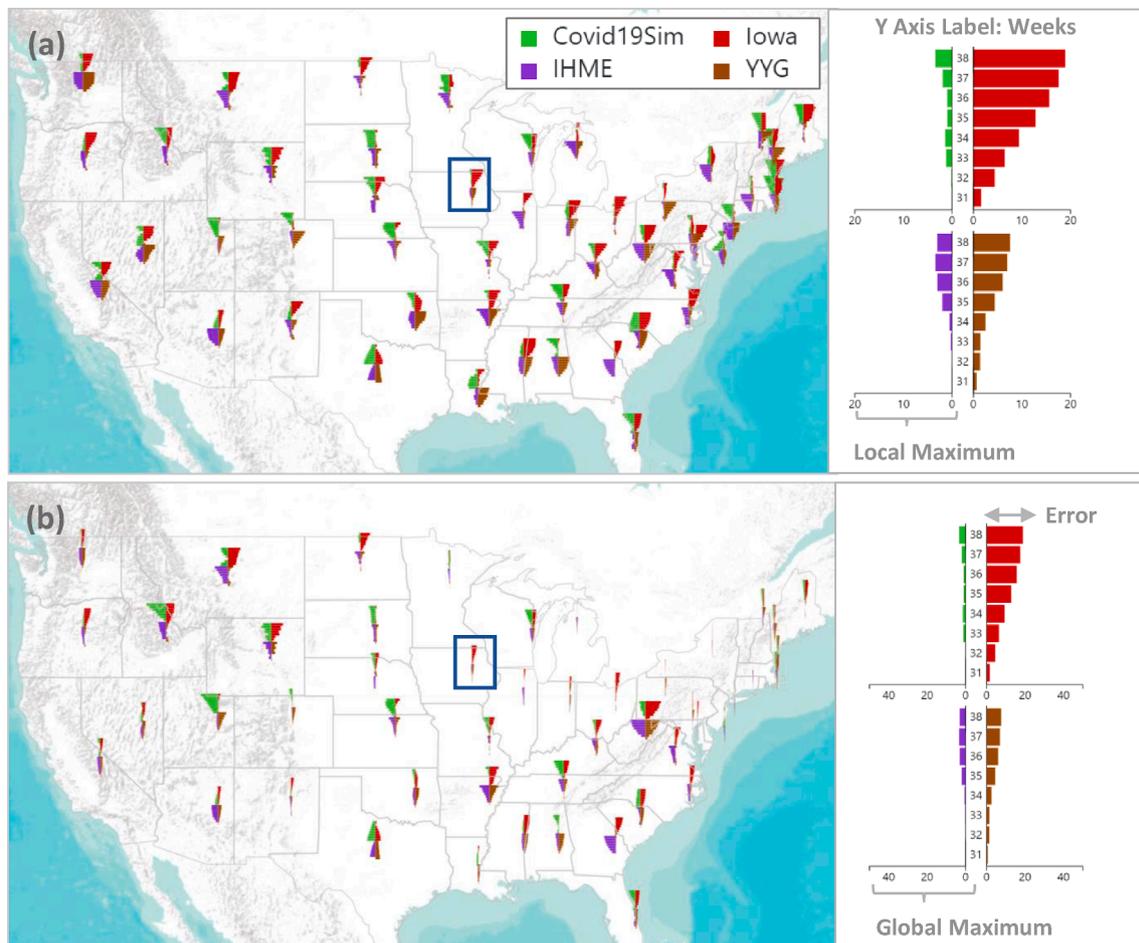There is no model that performs outstandingly in all states, but on the

Fig. 3. Through the visual interface, we compare the local (a) and global (b) views of leaflet glyphs for visualizing the accuracy of different models across the states.

selected model run[2], YYG performs well in most states, with the exception of Texas (where its predictions are off by an almost identical factor each week), and Arkansas, where Covid19Sim is the only model that performs well. The Iowa and IHME models both perform well in Kansas, Oklahoma, and Louisiana, and worst in Missouri, Arkansas, and Texas.

Our selection of models aims to provide a showcase of potential insights that can be generated through our visual analysis. As the GIS-based visual interface allows users to select a different combination of models that provide predictions for other variables (e.g., cases and hospitalizations) and have different levels of spatial (e.g., county and city) and temporal (e.g., daily and monthly) aggregations, more data-driven inferences can be generated during user interaction with the interface.

7. Conclusion

In this paper, we presented an innovative web-based tool, debuted as the EPIsembleVis, for conducting a comparative visual analysis on the consistency of COVID-19 ensemble predictions. By analyzing individual model projections as ensemble members, the EPIsembleVis is devised to (a) quantify the consistency of the prediction ensemble using metrics based on statistical aggregates, which include the coefficient of variation in individual ensembles (the dispersion in the predictions from different

ensemble members) and the prediction error (by comparing the model predictions against the recorded data), and (b) allows users to overview and explore in details the spatiotemporal variability of the ensemble predictions (e.g., similarities and dissimilarities of the ensemble members) at each spatial and temporal aggregation (e.g., weekly predictions in each state across the contiguous United States) using a suite of novel visualization techniques. EPIsembleVis was developed based on an automated data provisioning workflow powered through a ElasticSearch-Kibana stack, and it can automate the compilation of an ensemble dataset using public-available COVID-19 predictions from a variety of epidemiological models in a near-real-time fashion. This setting also makes our approach generalizable and scalable to analyze COVID-19 predictions produced from more models and in other geographic areas (e.g., country and continent). The tool was developed with open-source web technologies and adaptive system designs that make the system light-weight, low-cost, and interoperable with major online data analytical platforms, such as Kibana.

The visual analysis approach presented in this paper aims to enable heuristic explorations of the complex patterns in COVID-19 prediction ensemble datasets and serves as a pilot study to guide future investigative efforts through data-driven insights, which can help epidemiologists improve the performance and consistency of COVID-19 prediction models, as well as identifying the best performing models for certain scenarios (e.g., geographic areas and outbreak stages).

We summarize our experiences and findings from the approach as the following:

---

[2] Please note that all statements are made for this particular selection of model runs, and these statements are not universally true for the different models.
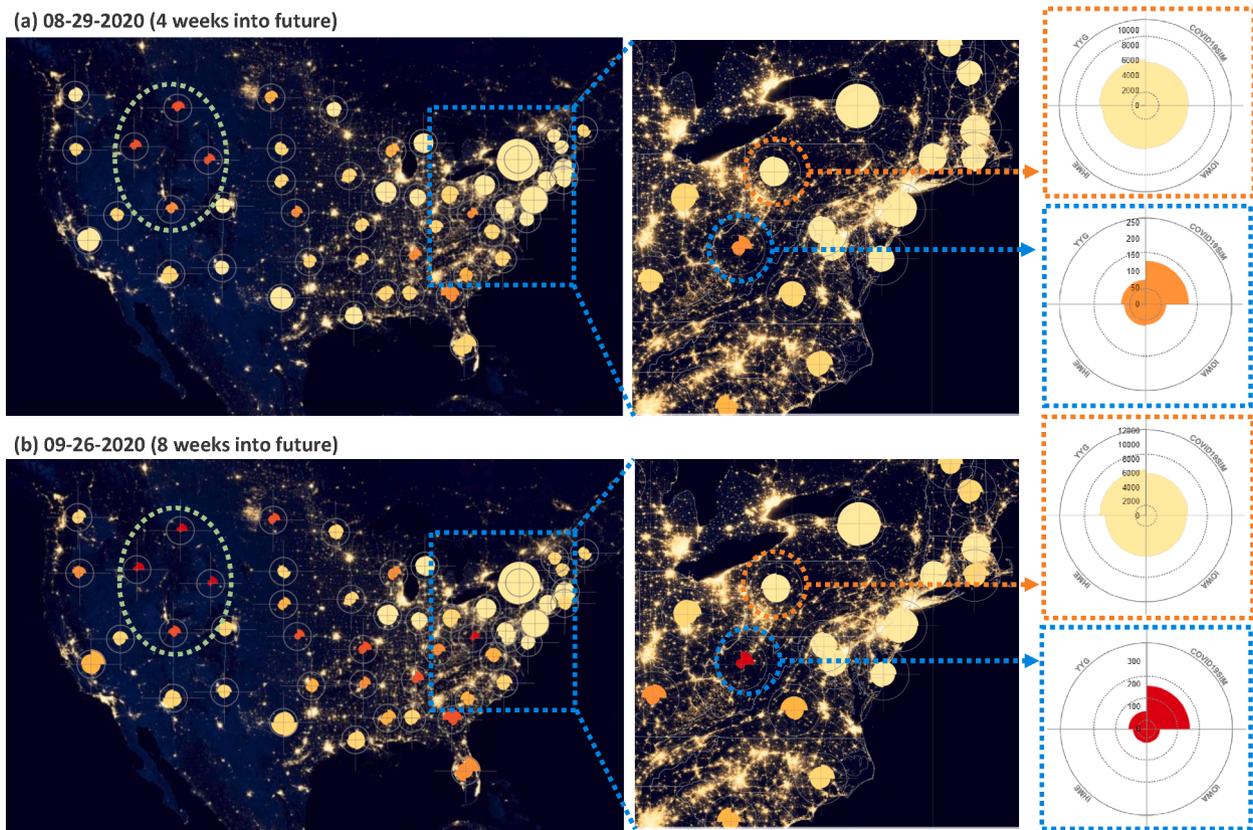
Fig. 4. Through the overlay of the NASA night-light map, the rose chart reveals that the overall disagreement between models is less for the 4-week-into-future projection. Prediction ensembles in populated areas usually have a lower coefficient of variation (indicated by light yellow colors and well-balanced petals) compared with the ensembles in sparsely populated areas. Examples of these sparsely populated areas include (a) Montana, Idaho, and Wyoming (highlighted by the green dashed ellipsoid), and (b) West Virginia (circled in blue) in the Eastern area (highlighted with the blue dashed square). Ensemble predictions in these states present high dispersion as indicated by their rose charts through the dark red color-coding and unbalanced petal sizes. Close-ups of the rose charts for Pennsylvania (orange) and West Virginia (blue) are shown on the right-hand side.

- For a comparative study, metadata visualizations, such as the data availability matrix presented in this paper, can help users identify applicable data for comparisons.
- Our approach provides a visual workflow that aims to foster a better understanding of the ensemble data itself and its associated data operations (e.g., data enrichment and different types of aggregations). The interactive workflow serves as a medium for users to interact with the data directly to extract useful information along with all steps of the workflow.
- Data-driven insights derived through the visual analysis, such as the empirical relationship between the COVID-19 prediction uncertainty and population density, can be used to generate new hypotheses for guiding future modeling and investigation efforts. Based on the spatiotemporal patterns acquired from our case study, we would propose a multivariate analysis to further quantify the relationship between the ensemble prediction performance and state-level demographic characteristics.

The visual interface was presented and well-received at the Centers for Disease Control and Prevention (CDC) headquarters. During the presentation, we prepared a user survey for the presentation audiences to evaluate the usefulness of our visual interface. The survey was not designed to assess the usability and utility of the visual interface using formal protocols and cognitive walk-through defined by the visualization and visual analytics communities. Instead, we used the survey to collect qualitative feedback on the usefulness of the visual interface and user-friendliness of the visual representation from the perspective of healthcare professionals. Most audiences of the presentation commented

positively on the usefulness of the visual interface and offered new visions and use cases for exploring the consistency of COVID-19 model predictions. Based on the feedback, we have identified our future work as the following: (1) applying more advanced metrics to provide a more in-depth and comprehensive characterization of each prediction ensemble's uncertainty and error. Examples of these metrics include the root mean squared error (RMSE) and the coefficient of variance of the root mean square error (RMSE CV), which are often used to evaluate the deviation of model prediction from reality. Different quantiles of each ensemble can also be incorporated into the analysis to provide a more detailed characterization of the variability in the individual model predictions; (2) adapting the leaflet glyph to visualize negative values when visualization the ensemble error through different shades of color-coding within the same hue; (3) conducting a case study using weekly projections that are produced at the county level across the United States. With higher spatial resolution, the county-level evaluation on modeling performances can provide more practical insights for supporting COVID-19 mitigation strategies, and (4) developing a heuristic data analysis using a combination of unsupervised machine learning and multivariate visualization techniques to explore potential factors (e.g., state, demographic, land-use/land-cover and mobility attributes) that can affect the uncertainty of prediction ensemble, as well as the prediction accuracy of individual models. We also plan to conduct a formal usability and utility test of the visual interface by inviting professionals from both the visualization and epidemiology communities. Ultimately, the future work aims to supplement the existing evaluation of the ensemble prediction uncertainty and model prediction accuracy, with the additional capabilities to provide a more straightforward (semi-
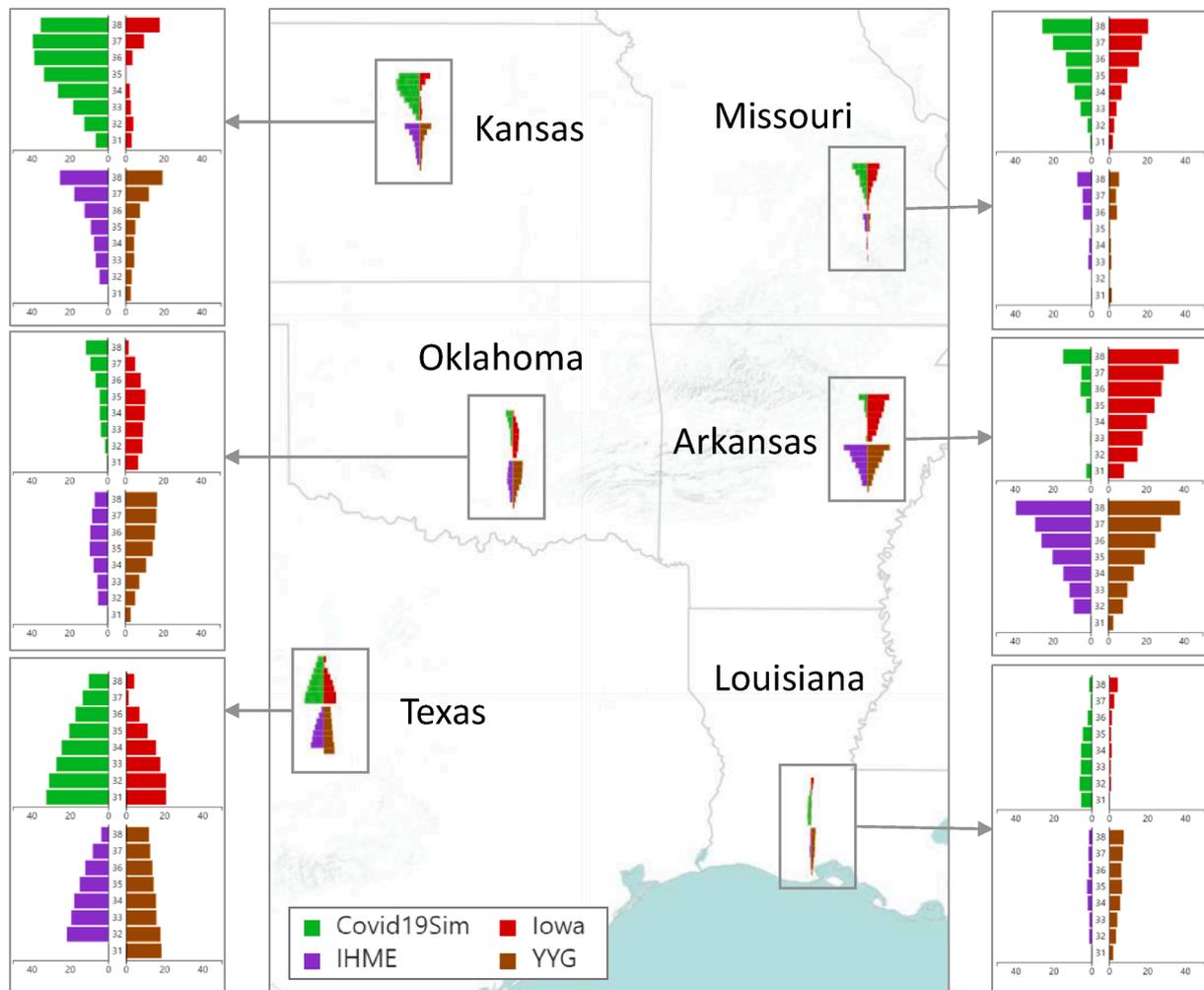
Fig. 5. Through the leaflet glyph, we are able to overview and compare the time-series of the prediction error from 4 models across 6 states. The prediction error is estimated by comparing the model prediction with the NYT recorded data.

automated) answer questions, such as which predictive model is best for a given scenario, and under which conditions (mitigation strategies, growth rate, population density, total population, etc).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D. Adam, Special report: The simulations driving the world's response to covid-19, Nature 580 (2020) 316.

[2] U.D. Bordoloi, D.L. Kao, H.W. Shen, Visualization techniques for spatial probability density function data, Data Sci. J. 3 (2004) 153–162.

[3] CDC, Community mitigation measures, 2020. https://www.cdc.gov/coronavirus /2019-ncov/global-covid-19/community-mitigation-measures.html.

[4] Center for Disease Control and Prevention, Interpretation of forecasts of new cases, 2020. online. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecastin g-us.html.

[5] T. Chakraborty, I. Ghosh, Real-time forecasts and risk assessment of novel coronavirus (covid-19) cases: A data-driven analysis, Chaos Solitons Fract. (2020) 109850.

[6] J. Chhatwal, T. Ayer, B. Linas, O. Dalgic, P. Mueller, M. Adee, M.A. Ladd, J. Xiao, Covid-19 simulator, 2020. online. https://covid19sim.org/.

[7] COVID, I., Modeling covid-19 scenarios for the united states, Nature Medicine.

[8] I. COVID, C.J. Murray, et al., Forecasting covid-19 impact on hospital bed-days, icu-days, ventilator-days and deaths by us state in the next 4 months, MedRxiv (2020).

[9] covid19sim, Policy simulator methodology, 2020. https://www.covid19sim.org /images/docs/Policy_simulator_methodology_20201001.pdf.

[10] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track covid-19 in real time, Lancet. Infect. Dis 20 (2020) 533–534, https://doi.org/10.1016/ S1473-3099(20)30120-1.

[11] P. Fenimore, B. McMahon, N. Hengartner, T. Germann, J. Mourant, A suite of mechanistic epidemiological decision support tools, Online J. Public Health Informat. 10 (2018).

[12] I. Franch-Pardo, B.M. Napoletano, F. Rosete-Verges, L. Billa, Spatial analysis and gis in the study of covid-19. a review, Sci. Total Environ. (2020) 140033.

[13] J. Friedman, P. Liu, E. Gakidou, I. COVID, M.C. Team, Predictive performance of international covid-19 mortality forecasting models, medRxiv (2020).

[14] Y. Gu, Covid-19 prediction model, 2020. online. https://covid19-projections.com/.

[15] H.W. Hethcote, The mathematics of infectious diseases, SIAM Rev. 42 (2000) 599–653.

[16] M. Hlawatsch, P. Leube, W. Nowak, D. Weiskopf, Flow radar glyphs–static visualization of unsteady flow with uncertainty, IEEE Trans. Visual Comput. Graphics 17 (2011) 1949–1958.

[17] T. Höllt, A. Magdy, G. Chen, G. Gopalakrishnan, I. Hoteit, C.D. Hansen, M. Hadwiger, Visual analysis of uncertainties in ocean forecasts for planning and operation of off-shore structures, in: 2013 IEEE Pacific Visualization Symposium (PacificVis), IEEE, 2013, pp. 185–192.

[18] I. Holmdahl, C. Buckee, Wrong but useful–what covid-19 epidemiologic models can and cannot tell us, N. Engl. J. Med. (2020).

[19] IHME COVID-19 Forecasting Team, Covid-19 estimate downloads, 2020. online. http://www.healthdata.org/covid/data-downloads.

[20] M. Jarema, I. Demir, J. Kehrer, R. Westermann, Comparative visual analysis of vector field ensembles, in: 2015 IEEE Conference on Visual Analytics Science and Technology (VAST), IEEE, 2015, pp. 81–88.

[21] M. Jarema, J. Kehrer, R. Westermann, Comparative visual analysis of transport variability in flow ensembles, 2016.

[22] N.P. Jewell, J.A. Lewnard, B.L. Jewell, Predictive mathematical models of the covid-19 pandemic: Underlying principles and value of projections, Jama 323 (2020) 1893–1894.

[23] Johns Hopkins University, Covid-19 data repository by the center for systems science and engineering (csse) at johns hopkins university, 2020. https://github.com/CSSEGISandData/COVID-19.

[24] J. Kehrer, H. Hauser, Visualization and visual analysis of multifaceted scientific data: A survey, IEEE Trans. Visualiz. Comput. Graph. 19 (2012) 495–513.

[25] P. Köthur, C. Witt, M. Sips, N. Marwan, S. Schinkel, D. Dransch, Visual analytics for correlation-based comparison of time series ensembles, in: Computer Graphics Forum, Wiley Online Library, 2015, pp. 411–420.

[26] T.R. Lab, Covid-19 forecast hub, 2020. https://github.com/reichlab/covid19-forecast-hub.

[27] Los Alamos National Laboratory, Interpretation of forecasts of new cases. online, , 2020. https://covid-19.bsvgateway.org/.

[28] A.L. Love, A. Pang, D.L. Kao, Visualizing spatial multivalue data, IEEE Comput. Graphics Appl. 25 (2005) 69–79.

[29] F.S. Lu, A. Nguyen, N. Link, M. Santillana, Estimating the prevalence of covid-19 in the united states: three complementary approaches, 2020.

[30] C.J. Murray, Forecasting covid-19 impact on hospital bed-days, icu-days, ventilator-days and deaths by us state in the next 4 months, medRxiv (2020). doi: 10.1101/2020.03.27.20043752, arXiv:https://www.medrxiv.org/content/early/2020/03/30/2020.03.27.20043752.full.pdf.

[31] New York Times, Coronavirus (covid-19) data in the United States, 2020. online. URL https://github.com/nytimes/covid-19-data.

[32] M. Nicola, Z. Alsafi, C. Sohrabi, A. Kerwan, A. Al-Jabir, C. Iosifidis, M. Agha, R. Agha, The socio-economic implications of the coronavirus pandemic (covid-19): A review, Int. J. Surg. (Lond. Engl.) 78 (2020) 185.

[33] H. Obermaier, K. Bensema, K.I. Joy, Visual trends analysis in time-varying ensembles, IEEE Trans. Visualiz. Comput. Graph. 22 (2015) 2331–2342.

[34] D. Osthus, S. Del Valle, C. Manore, B. Weaver, L. Castro, C. Shelley, M.M. Smith, J. Spencer, G. Fairchild, T. Pitts, D. Gerts, L. Dauelsberg, A. Daughton, M. Gorris, B. Hornbein, D. Israel, N. Parikh, D. Shutt, A. Ziemann, LANL COVID-19 Cases and Deaths Forecasts, 2020. https://covid-19.bsvgateway.org/.

[35] T. Pfaffelmoser, R. Westermann, Correlation visualization for structural uncertainty analysis, Int. J. Uncertainty Quantif. 3 (2013).

[36] K. Potter, J. Kniss, R. Riesenfeld, C.R. Johnson, Visualizing summary statistics and uncertainty, in: Computer Graphics Forum, Wiley Online Library, 2010. pp. 823–832.

[37] K. Potter, A. Wilson, P.T. Bremer, D. Williams, C. Doutriaux, V. Pascucci, C. R. Johnson, Ensemble-vis: A framework for the statistical visualization of ensemble data, in: 2009 IEEE International Conference on Data Mining Workshops, IEEE, 2009, pp. 233–240.

[38] K.C. Potter, C.R. Johnson, V. Pascucci, Y. Zheng, P.T. Bremer, A Flexible Approach for the Statistical Visualization of Ensemble Data. Technical Report. Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), 2009b.

[39] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, R. Moorhead, Noodles: A tool for visualization of numerical weather model ensemble uncertainty, IEEE Trans. Visual Comput. Graphics 16 (2010) 1421–1430.

[40] G.R. Shinde, A.B. Kalamkar, P.N. Mahalle, N. Dey, J. Chaki, A.E. Hassanien, Forecasting models for coronavirus disease (covid-19): A survey of the state-of-the-art, SN Comput. Sci. 1 (2020) 1–15.

[41] Q. Shu, H. Guo, J. Liang, L. Che, J. Liu, X. Yuan, Ensemblegraph: Interactive visual analysis of spatiotemporal behaviors in ensemble simulation data, in: 2016 IEEE Pacific Visualization Symposium (PacificVis), IEEE, 2016, pp. 56–63.

[42] A. Srivastava, T. Xu, V.K. Prasanna, Fast and accurate forecasting of covid-19 deaths using the sikj alpha model, , 2020. arXiv preprint arXiv:2007.05180.

[43] K.S. Taylor, J.W. Taylor, A comparison of aggregation methods for probabilistic forecasts of covid-19 mortality in the United States, 2020. arXiv preprint arXiv:2007.11103.

[44] G. Thakur, K. Sparks, A. Berres, V. Tansakul, S. Chinthavali, M. Whitehead, E. Schmidt, H. Xu, J. Fan, D. Spears, et al., Covid-19 joint pandemic modeling and analysis platform, in: Proceedings of the 1st ACM SIGSPATIAL International Workshop on Modeling and Understanding the Spread of COVID-19, 2020, pp. 43–52.

[45] The Atlantic, The COVID Tracking Project, 2020. covidtracking.com.

[46] B.L. Turner, Contested identities: Human-environment geography and disciplinary implications in a restructuring academy, Ann. Assoc. Am. Geogr. 92 (2002) 52–74.

[47] UMass-Amherst Influenza Forecasting Center of Excellence. Covid-19 forecast hub, 2020. online. https://github.com/reichlab/covid19-forecast-hub/.

[48] United Nations, The social impact of covid-19, 2020. online. https://www.un.org/development/desa/dspd/2020/04/social-impact-of-covid-19/.

[49] J. Wang, S. Hazarika, C. Li, H.W. Shen, Visualization and visual analysis of ensemble data: A survey, IEEE Trans. Visualiz. Comput. Graph. 25 (2018) 2853–2872.

[50] L. Wang, G. Wang, L. Gao, X. Li, S. Yu, M. Kim, Y. Wang, Z. Gu, Spatiotemporal dynamics, nowcasting and forecasting of covid-19 in the United States, 2020a. arXiv:2004.14103.

[51] L. Wang, G. Wang, L. Gao, X. Li, S. Yu, M. Kim, Y. Wang, Z. Gu, Spatiotemporal epidemic modeling, 2020b. online. https://covid19.stat.iastate.edu.

[52] D.M. Weinberger, T. Cohen, F.W. Crawford, F. Mostashari, D. Olson, V.E. Pitzer, N. G. Reich, M. Russi, L. Simonsen, A. Watkins, et al., Estimating the early death toll of covid-19 in the united states, bioRxiv (2020).