Uncertainty Visualization Study Group

Report presented to

Professor Veronica Bowman (Defence Science Technology Laboratory)

The Study Group with Industry was hosted by:

Knowledge Transfer Network

1 Summary and Key Learnings

This report documents the outputs of a three day Study Group hosted at the Business Design Centre, London. This Study Group explored the state of the art in visualization of uncertainty when applied to disease forecasting scenarios. This Section summarizes the outputs and linkages between each groups whilst a full account of the Study Group work can be found in Sections 2 to 8.

There is a rich literature on graphics for uncertainty, and a growing amount of research on the interaction with decision makers. Whilst emphasizing the importance of context, David Spiegelhater in Ref. [32] provides thoughts on what should be considered when constructing visualizations which are underpinned by probabilities, these can be found in Section. 2.

The premise of this work was two-fold. Firstly, how visual representations might be designed which are intuitive to interpret, leading the decision maker to make the best decision from the data. Secondly, to be able to convey the presence and effect of various uncertainties which are inherent in computational models, Ref. [2]. The particular application for this work is in providing decision support tools for those dealing with epidemiological models in disaster scenarios. This application forms the basis of the CrystalCast project:

CrystalCast

The aim of CrystalCast is to bring these disparate models and data streams together in a seamless, rigorous mathematical framework supported by a fully interactive, state of the art, visualization system. [...] CrystalCast will utilize an innovative visualization process developed to empower non-experts to make decisions based on uncertain information. This technique, visualization via 'frequency trees', was initially developed to enable decisions about medical treatment under uncertainty. However, their ease of use and flexible generation along with their applicability to decision makers, makes them an ideal basis for CrystalCast's Graphical User Interface (GUI). (See Fig. 1 for mock-up interface.)

Four scenarios involving uncertainty interpretation were explored with the aid of data provided by the Defence Science Technology Laboratory (Dstl) and Public Health England (PHE). These topics were:

1. The output needs to reflect the assumptions which make up the predictions.
2. Decision aids that include summaries of temporal predictions
3. Spatio-temporal uncertainty visualization

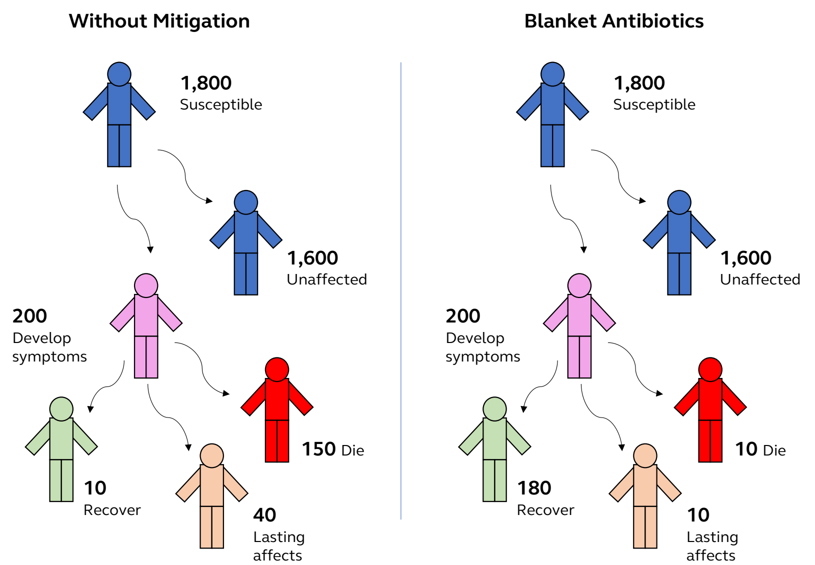
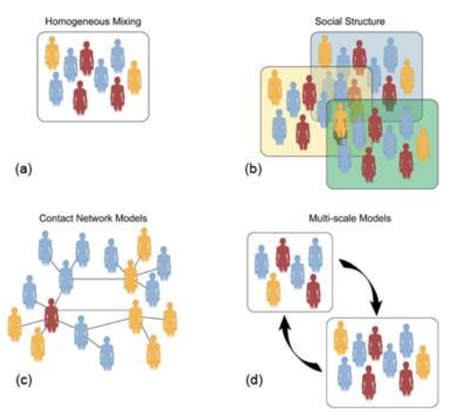


Figure 1: Initial Mock-Up of a frequency tree for epidemiological decision-making

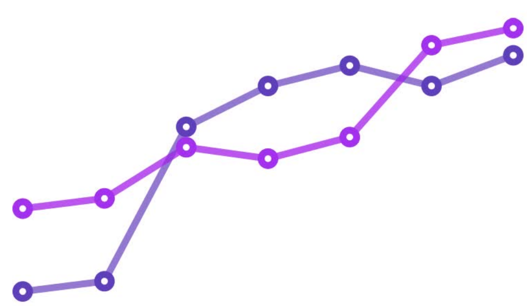
4. Visualizing uncertainty in the numbers in 'at risk' groups due to multiple model predictions

Reflecting the assumptions which make up the model predictions.

The team explored the literature for insightful graphics and suggestions developed to visualize model assumptions. Various representations were discussed which covered static glyphs which highlight the nature of the model assumptions underpinning the predictions (see Fig. 2a), dynamic animations which could be part of the model output (for example



(a) Example of visualizing model assumptions.



(b) Animated line chart rendering.

Figure 2: Ideas of how model assumptions could be encoded into visualizations.

could the way in which the plot is animated encode some information related to the model assumptions? (see Fig. 2b)).

Two assumptions in the disease models were worked up in detail; model complexity and population mixing assumptions. The former involved a graphical representation of the detail of the model being employed. The latter focuses on interactive techniques where model assumptions can be deduced from the behavior of the data and the effect of assumption can be tested.

Decision aids that include summaries of temporal predictions.

In this challenge the temporal nature of disease predictions were explored. The problem was broken down into a two-stage process (for analysts and decision maker separately) and interfaces were designed for both. Special attention was given to the kinds of questions both groups might be asking of the visualization, and the pitfalls ill-conceived communication strategies might fall into.

In Fig. 3 we show snapshot and cumulative scenarios at important milestones; end of week 1, after three months etc, these would be customizable for the decision maker. The shading of the icons represent some level of confidence in the given prediction.

Spatio-temporal uncertainty visualization In this group approaches in how uncertainty can be understood when there is a spatial component to the prediction are explored. The group

Figure 3: Snapshot and cumulative plots.

developed three approaches aimed at visualizing spatial uncertainty of disease predictions;

- Firstly an interactive map with impact / uncertainty was developed to allow the user to explore areas according to a quadrant map of low to high for both impact and uncertainty.
- Borrowing from engineering design, parallel coordinates were used to visualize high dimensional spaces where decision makers might explore explicit and implicit risk factors for certain geographical locations.
- Finally, a tile map was developed which linked geographical maps to tile maps (containing impact and uncertainty data) to a ranked list.

The temporal evolution of disease spread was not considered by this group, but clear overlaps and synergies can be made with the temporal group.

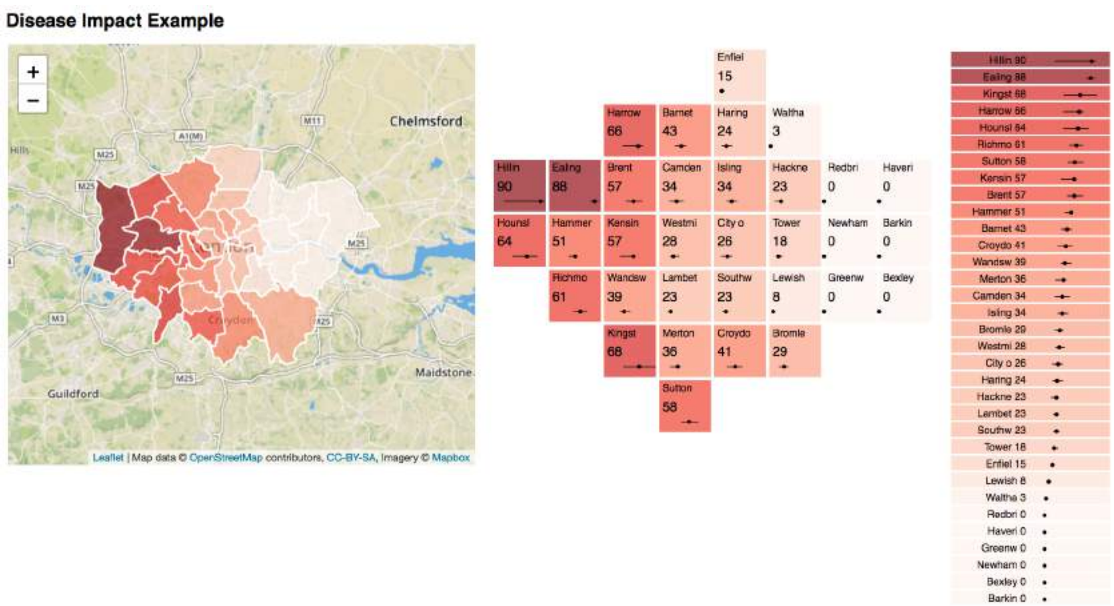


Figure 4: Tile map example showing linked map, tile map, and impact list.

Visualizing uncertainty in the numbers in 'at risk' groups due to multiple model predictions

This group explored the concept of texture to encode more information into the graphic, in particular textures to represent infection strength *and* model agreement. It was decided that alternative color scales would best encode this type of information. Fig. 5 shows a concept diagram of this (green / red color scale indicating model prediction and the dark blue / blue scale the degree of model disagreement.)

This encoding idea is translatable across into other visualization approaches, for example,

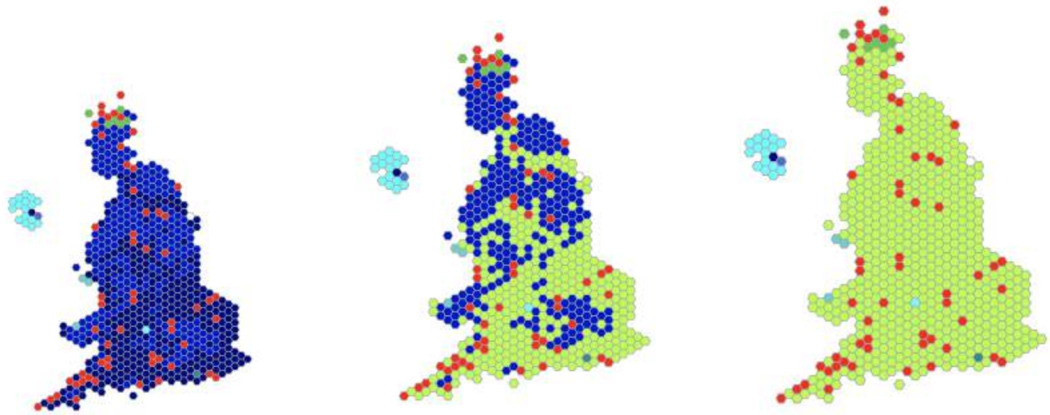


Figure 5: Example of encoding model uncertainty into visualization

one might consider the icons in Figs. 1, 3 and 4 following a related color coding for highlighting model discrepancy.

Taken together, the work achieved in this Study Group reviews existing approaches and takes inspiration from this to create implementable solutions via a context driven methodology which draws upon expertise from the statistical, epidemiological, psychological, data visualization, and design communities in the United Kingdom (UK). The authors and sponsors are grateful to all those who gave their time and effort in completing this task.

Contents

2 Overview of Uncertainty Visualisation

> *Most visualization techniques have been designed on the assumption that the data to be represented are free from uncertainty. Yet this is rarely the case. Recently the visualization community has risen to the challenge of incorporating an indication of uncertainty into visual representations [...] We may encounter error bars on graphs, but we rarely see the equivalent on contour maps or isosurfaces. Indeed the very crispness of an isosurface gives an impression of confidence that is frankly often an illusion. This is a major issue when visualizations are used in decision making - such as planning evacuations based on a visualization of the predicted hurricane path. Indeed as computation power and capability continue to increase we see a rise in ensemble computing, where many simulations of a phenomenon are carried out for different initial conditions, or different settings of unknown parameters - leading not to a unique data value, but to a set of values - so-called multivalue data.*
>
> Ken Brodlie *et al* in *A Review of Uncertianty Visualisation* [6]

The literature is full of exciting visualization techniques, but according to Ref. [4] *"despite this array of tools, the standard approaches to graphical display remain those based on relatively simple point and line drawings, such as histograms, box plots, bar charts and scatterplots, supplemented occasionally by colour filling"*.

Why is uncertainty visualization difficult? Is it simply because it adds another dimension to data to visualize? Brodlie *et al* reviewed the situation in Ref. [6], quoting directly from his paper, the reasons for the difficulties are:

1. **Uncertainty is complex** - "Uncertainty, by its very nature, is a difficult subject [...] even the terminology is often unhelpful."
2. **Uncertainty information is presented in different ways**- "as a Probability Density Function (PDF), as multivalue data, as bounded data, and other representations."
3. **Uncertainty propagates**- "When we calculate with uncertain data, we propagate the uncertainty [...] We need to understand how to propagate uncertainty in the data, through to uncertainty in [an] image."
4. **Uncertainty adds a dimension to the visualisation**- "We have enough problems

visualising exact 3D or higher dimensional data without introducing another dimension for uncertainty."

5. **Uncertainty tends to dominate certainty**- "In most natural visual representations of uncertainty, the greatest emphasis is placed on data of greatest uncertainty."

6. **Uncertainty adds another discipline**- "Some of the best visualizations have been created by multidisciplinary teams, bringing together domain scientists, numerical analysts, visualization scientists and artists. There is a further discipline to be added now: statistics."

According to Spiegelhalter *et al* in Ref. [32] *"there are few reproducible experimental findings for assessing best practice in visualizing uncertainty. Instead, reviewers have emphasized how graphics can be adapted to the aims of the communicator, stressing the importance of the context of the communication exercise and the needs and capabilities of the audiences.[...] Despite the burgeoning interest in infographics, there is limited experimental evidence on how different types of visualizations are processed and understood, although the effectiveness of some graphics clearly depends on the relative numeracy of an audience. Fortunately, it is increasingly easy to present data in the form of interactive visualizations and in multiple types of representation that can be adjusted to user needs and capabilities.*

What is the best way to visualize probabilistic uncertainty? Ref. [32]

The most suitable choice of visualization to illustrate uncertainty depends closely on the objectives of the presenter, the context of the communication, and the audience. Visschers et al. Ref. [35] concluded that the "task at hand may determine which graph is most appropriate to present probability information" and it is "not possible to formulate recommendations about graph types and layouts." Nonetheless, if we aim to encourage understanding rather than to just persuade, certain broad conclusions can be drawn, which hold regardless of the audience.

- Use multiple formats, because no single representation suits all members of an audience.
- Illuminate graphics with words and numbers.
- Design graphics to allow part-to-whole comparisons, and choose an appropriate scale, possibly with magnification for small probabilities.
- To avoid framing bias, provide percentages or frequencies both with and without the outcome, using frequencies with a clearly defined denominator of constant size.
- Helpful narrative labels are important. Compare magnitudes through tick marks, and clearly label comparators and differences.
- Use narratives, images, and metaphors that are sufficiently vivid to gain and retain attention, but which do not arouse undue emotion. It is important to be aware of affective responses.
- Assume low numeracy of a general public audience and adopt a less-is-more approach by reducing the need for inferences, making clear and explicit comparisons, and providing optional additional detail.
- Interactivity and animations provide opportunities for adapting graphics to user needs and capabilities.
- Acknowledge the limitations of the information conveyed in its quality and relevance. The visualization may communicate only a restricted part of a whole picture.
- Avoid chart junk, such as three-dimensional bar charts, and obvious manipulation through misleading use of area to represent magnitude.
- Most important, assess the needs of the audience, experiment, and test and iterate toward a final design.

3 Visualizing Uncertain Disease Forecasts

> *Ensemble data sets are an increasingly common tool to help scientists simulate complex systems, mitigate uncertainty, and investigate sensitivity to parameters and initial conditions. These data sets are large, multidimensional, multivariate and multivalued over both space and time. Due to their complexity and size, ensembles provide challenges in data management, analysis, and visualization.*
>
> K. Potter *et al* in Ref. [29]

3.1 Problem Statement

There are many models and techniques for disease forecasting over a huge range of spatial and temporal scales, using a wealth of potential data streams. This results in many different predictions of disease spread and effect on a population, with no clear indication of the accuracy of any one outcome. In combination with a wealth of possible intervention strategies, this leaves the decision maker information rich, but potentially unable to sift this information and provide a succinct evidence base for key decisions.

Visualizing uncertainty in a way that is easily understood and which facilitates immediate action is a complex and application specific problem. There have been a number of research efforts in the recent past looking at this problem, particularly addressing how uncertainty should be presented for a decision maker. Visualizing uncertainty from an ensemble of epidemiological models poses a number of particular challenges:

- The output needs to reflect the assumptions which make up the predictions.
- The output must enable high-level decisions as well as detailed analysis of specific model outputs.
- The output is defined over a spatio-temporal region, meaning there are multiple dimensions to be presented.
- The output must be visualized over a number of different modeling scales, from high level approximations to detailed simulation models and reflect the associated uncertainties.

3.2 Decision Making Chain

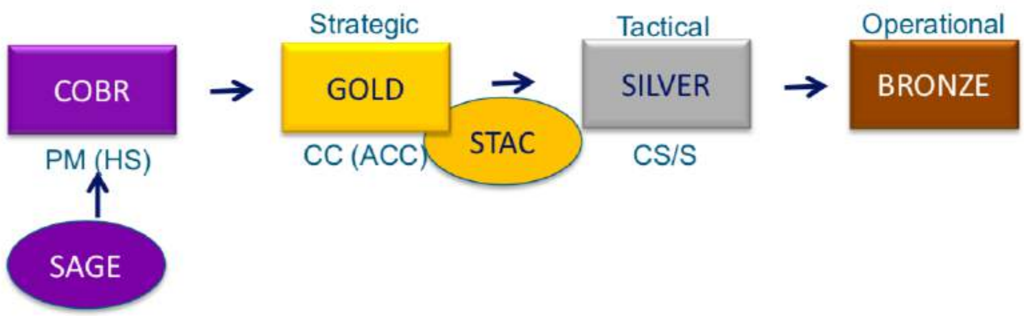


Figure 6: There are three levels of command headed up by COBR

Various assumptions need to be made regarding the decision making process. It is assumed that the analysts (modelers) are experts in the modeling process but are not (always) aware of the contextual details of how the model output will be used in operation - thus it is important to give the decision maker freedom to explore the model outputs to support an unknown decision making process. For the purpose of the Study Group, the command chain shown in Fig. 6 was assumed Ref. [5]

3.3 CrystalCast

The aim of CrystalCast is to bring these disparate models and data streams together in a seamless, rigorous mathematical framework supported by a fully interactive, state of the art, visualization system.

CrystalCast will incorporate information from all available disease forecast models and data sources into a single, integrated output, empowering the user to make clear, evidence based decisions. The system will provide the ability to interrogate, manipulate and, crucially, to mathematically optimize, the underlying models and potential courses of action.

CrystalCast will be a Bio-Surveillance Ecosystem (BSVE) application, which can synthesize data from BSVE data streams and analyze outputs from existing BSVE Disease Forecast Model (DFM) applications. It will also include an internal toolbox of additional DFMs .

The system will comprise of three modules, an internal Disease Forecast Toolbox (DFT), an Ensemble Ensemble Uncertainty Quantification and Optimization (EUQO) framework and a Visualisation Suite (VS). Each module will run stand-alone, as well as in combination, providing a unique, powerful decision support tool and a single point of reference for disease forecasting.

- DFT - CrystalCast will incorporate a DFM suite provided by PHE, the authoritative disease forecast providers within UK Government. This includes source estimation models for non-person-to-person infections that project expected case numbers, simulation models for person-to-person spread infections and, potentially, vector borne infections. Agreed pathogen-specific models will be included with credible intervention strategies and will be extended for similar pathogens where possible.
- EUQO - The world-class teams from the University of Southampton and Dstl will incorporate research into emulation methodology with novel approaches to provide the ability to run Uncertainty Quantification (UQ) across ensembles of disparate model runs. The uncertainties in these ensembles will be explored, characterized and synthesized for unified analysis of the current disease picture. This information will be combined with real-time optimization techniques to recommend optimal potential courses of action.
- VS - CrystalCast will provide a fully interactive, state of the art, visualization suite using frequency trees as the backbone. This will provide 'at a glance interpretation' alongside the ability to drill down into specific details of the highly complex and uncertain outcomes. The visualization suite will provide a full audit trail of all modeling queries for later analysis and be compatible with the BSVE.

In the base year, a prototype stand-alone system will be developed, combining the DFT with data and models drawn from the BSVE and a basic visualization system. In option year one, CrystalCast will be developed to plug into the BSVE as an app and link with its current data streams, it will incorporate initial EUQO capability. In option year two, CrystalCast will incorporate a fully interactive visualization system with the full EUQO suite. The app will be demonstrated on a Defense Threat Reduction Agency (DTRA) approved set of pathogens such as *B. anthracis*, Ebola and Zika.

Visualization Visualizing uncertainty in a way that is easily understood and immediately actionable is a complex and application specific problem. There have been a number of research efforts in the recent past looking at this problem, particularly addressing how uncertainty should be presented for a decision maker and defining a process for developing a visualization system for a bespoke application. Visualizing uncertainty from an ensemble of epidemiological models poses a number of particular challenges:

- The output is defined over a spatio-temporal region, one of the most difficult visualization situations.
- The output must be visualized over a number of different modeling scales, from high level approximations to detailed agent based models.

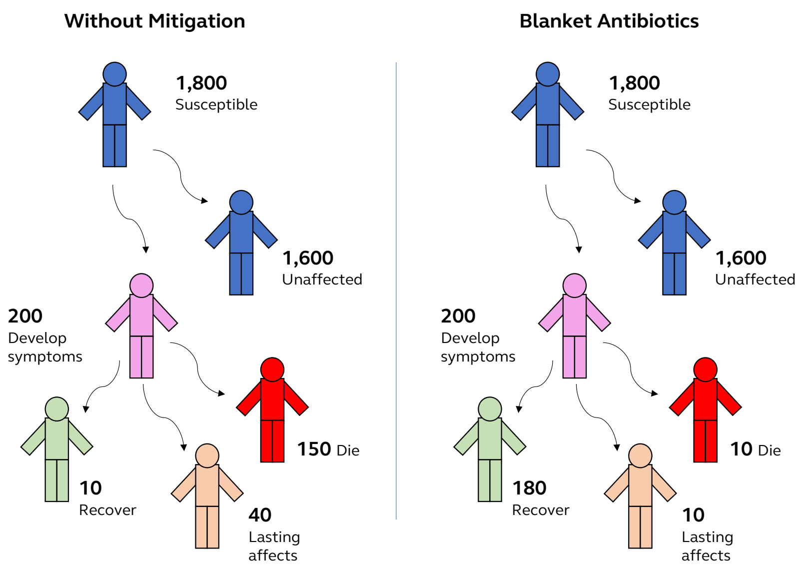


Figure 7: Initial Mock-Up of a frequency tree for epidemiological decision-making

- The output must enable high-level decisions as well as detailed analysis of specific model outputs.

In order to address these challenges, CrystalCast will utilize an innovative visualization process developed to empower non-experts to make decisions based on uncertain information. This technique, visualization via 'frequency trees', was initially developed to enable decisions about medical treatment under uncertainty. However, their ease of use and flexible generation along with their applicability to decision makers, makes them an ideal basis for CrystalCast's GUI.

Frequency trees will provide a high-level overview of both potential courses of action and their associated uncertainty, as well as the hierarchy of outputs from increasingly complex model types. They will be interactive, so that each node can be selected individually to display more detailed visualizations of the information underpinning that node. In this way, a seamless transition can be made from high-level information to detailed visualizations of model input / output for an expert analyst.

The detailed visualization mode will allow interaction with all underpinning model parameters. Each parameter will be analyzed and a variety of graphical plotting techniques, drawn from standard practice and cutting edge research, will be provided. Where appropriate, bespoke graphical outputs will be developed so that the high dimensional output is easily understandable. Where multiple fidelities of information are available, all outputs will be made interactive so that the user can drill down into the presented information providing flexible displays. These displays will be recorded in a user story, so that views that are required often can be created instantly by the user without further interaction.

In addition, the GUI will allow full interactivity with the underlying model runs, the optimization process and the uncertainty analysis, as well as providing the ability to easily document the analysis undertaken in the current session and the information produced to make a particular decision. In this way, any user will be empowered to easily construct an output that effectively documents an evidence-based decision and that includes all modeling assumptions and analysis decisions. This output will be output to the BSVE for incorporation into searches and analysis done on the wider BSVE system.

4 Methodology

Dstl and PHE wished to understand the state-of-the-art in interactive visualization approaches to enable evidence based decision making within the CrystalCast project; requiring users to be able to access the underlying model data, assumptions etc from within the constructed GUI. Four scenarios were developed and sample data was created to facilitate the discussion for each; these scenarios were:

1. Communicating the assumptions underlying model predictions to help the decision maker understand the utility of the outputs.
2. Decision aids that include data summaries and temporal predictions
3. Spatio-temporal visualization of hazard areas and populations at risk
4. Visualizing uncertainty in the numbers in 'at risk' groups due to multiple model predictions

The Knowledge Transfer Network (KTN) sourced experts from the fields of statistics, scientific visualization, knowledge co-production, user-centered design and others to join in the Study Group. These experts (mainly from UK research institutions) were facilitated by context experts from Dstl and PHE - a full list of participants can be found in Section 4.2.

4.1 Study Group with Industry

The format of the Study Group will be following the highly successful European Study Groups with Industry in a condensed, three-day, format. Problem posers presented their challenges on the morning of the first day to the group. Researcher participants ask questions and then choose which group they may be able to help with. The group is given three days to construct tasks, work on them and then feed back on the final day. This report captures that effort.

4.2 Attendees

First Name	Surname	Job Title, Institution	Skills	A	B	C	D
Helen	Adams	Senior Software Engineer Riskaware	Programming support.	1	1	1	1
Hazel	Biggs	Defence & Security Knowledge Transfer Network	Significant experience in law enforcement both from a crime prevention and community safety aspect in addition to 25 years of forensic investigative experience. Ex Home Office working as Chief of Staff to the HO Chief Scientific Advisor and the Head of CAST working with Ministers and the Permanent Secretary to address HO S&T priorities.	1	1	1	1
Veronica	Bowman	Principal Statistician Dstl	Specialises in uncertainty calculation and communication and Bayesian inference	1	1	1	1
Matt	Butchers	Knowledge Transfer Manager Knowledge Transfer Network	N/A	0	0	0	0
Min	Chen	Professor of Scientific Visualization University of Oxford	Visual analyticsInformation theory	1	1	1	1
Xin	Chen	Research Student Cranfield University	Uncertainty quantification. Data visualisation. Sensitivity analysis	0	0	1	1
Haeran	Cho	Lecturer in Statistics University of Bristol	My research centres around sparse modelling and estimation of large datasets including the detection of (possibly) multiple change-points in the stochastic properties of (possibly high-dimensional) data.	1	0	1	0
Sacha	Darwin	Principal Consultant Riskaware	Software engineering (C++/Java/JS)	1	1	1	1
David	Demeritt	Professor of Geography King's College London	I am a social scientist interested in the communication and use of science and risk information for policymaking	0	1	1	1
Brian	Dixon	Lecturer and Researcher Glasgow School of Art	I have a background in information design.	1	1	0	1
Soufiane	El Fassi	Ph.D Student Cranfield University	Uncertainty quantification and management in engineering design	0	0	1	1
James	England	Principal Consultant Riskaware	Mathematical ModellingC++/Java/Web DeveloperVisual Design for UX	1	1	0	0
Delina	Evans	Deisgn Knowledge Transfer Network	Human-centered design	1	1	1	1
Thomas	Finnie	Senior Geospatial Modeller Public Health England	Project Lead	0	1	0	0
Virginia	Foot	Principal Scientist Dstl	Part of Dstl team	0	0	0	1
Alex	Forrester	Associate Professor University of Southampton	Surrogate Modelling. Optimisation. Leading 'visualisation of uncertainty due to model selection' group	0	0	1	0

First Name	Surname	Job Title, Institution	Skills	A	B	C	D
Sam	Grainger	Postdoctoral Research Fellow University of Leeds	User-centred design. User testing. Uncertainty communication. Water Resources. Climate Change. Developing countries	0	0	0	1
Ian	Hall	Reader Manchester University	Epidemic Modelling, mathematical Statistics	0	0	0	1
David	Haw	Research Associate Imperial College London	High resolution spatial modelling; multi-annual modelling; analysis & forecasting (using WHO fluNet data)	0	0	1	1
Nicolas	Holliman	Professor of Visualization Newcastle University	Data VisualizationPython, R, Power Bi, Tableau, Blender, Cloud (Azure)	1	1	1	1
Samuel	Jackson	Senior Research Assistant	Modelling and statistics	0	0	1	0
Timoleon	Kipouros	Senior Research Associate University of Cambridge	Multi-dimensional data visualisation and interactive analysis.	1	1	1	1
Dick	Lacey	Ex-Chief Scientist Home Office Dstl	Speaker	0	0	0	0
Duncan	Lee	Reader in Statistics University of Glasgow	Spatio-temporal disease modelling and software development.	0	1	0	0
Fiona	Lethbridge	Senior Press Office Science Media Centre	Speaker	0	0	0	0
Polina	Levontin	Researcher Imperial College London	Shiny apps (web based interactive dashboards) and some previous use of information design applied to fisheries science	0	0	1	1
Graham	McNeill	Researcher, Oxford Internet Institute University of Oxford	Broad background in data science and machine learning. Specific expertise in JavaScript and web-based visualisation of spatio-temporal data.	0	1	0	0
Noel	Nelson	Senior Dispersion Scientist Met Office	Expertise in atmospheric dispersion modelling and experience in synthesising science for policy developers.	0	0	0	1
Toby	Pilditch	Research Associate University College London	Cognitive Psychology. More precisely, evidential and probabilistic reasoning, belief-updating, complex system simulation, cognitive biases, and decision making. Some of the tools I use include Bayesian Network models, Agent-Based Models, various statistical models, as well as expertise in experimental design (and related concepts).	1	1	0	1

First Name	Surname	Job Title, Institution	Skills	A	B	C	D
Adrian	Pratt	Senior Mathematical Modeller Public Health England	Infectious disease modeller. R programmer.	1	0	0	0
Jonathan	Roberts	Professor Bangor University	Broad knowledge of visualisation and visualisation design.	1	0	0	0
Pranay	Seshadri	Postdoctoral Fellow University of Cambridge	Dimensional reduction and visualisation	1	1	1	0
Daniel	Silk	Statistician Dstl	Statistical inference, modelling and emulation	0	1	1	1
Simon	Smith	Project Manager Dstl	CrystalCast PM	1	0	0	0
Hannah	Williams	Mathematical Modeller Public Health England		1	0	0	0
Anthony	Wilson	Group Leader Pirbright	Expertise in disease ecology (of vector-borne diseases), entomology, and modelling.	1	1	1	1
Kai	Xu	Associate Professor Middlesex University	My expertise is data visualisation. The information about some of the tools I developed recently is online here: http://vis4sense.github.io/	0	0	0	1

(A) Multi-level temporal
(B) Spatio-temporal
(C) Model selection
(D) Assumptions

5 Challenge 1: Visualisation of Underpinning Model Assumptions

5.1 Overview of Problem

Introduction to Theme: Mathematical and statistical models are attempts to sufficiently describe reality in order to understand complex behavior or predict future events. Such models enable theoretical conjectures to be investigated, but require approximations and estimates to be made. These approximations may be roughly classified into two types. Firstly, those that remove 'unnecessary' complexity in a system to give a mathematical representation. These are often introduced in numerical (computer-based) solutions of certain mathematical models or when physical effects may be neglected. Crucially these assumptions can be tested and justified by researchers. The second type is for mathematical convenience, made so that a system can be analyzed or due to key data gaps. However, with little evidence base for these approximations, the impact of making them is uncertain.

In emerging research areas where experiments are costly or ethically challenging the second type of approximations are often unavoidable. Models may be developed for any physical, chemical or biological system. In this workshop we are specifically focused on epidemiological models of disease spread. Because this topic is predicated on human decision making and behavior as well as complex mechanisms, assumptions are both numerical and convenience based.

The simplest model for disease spread in a community is the so-called Susceptible, Infected, Removed (SIR) model. This model is a foundational example for mathematical epidemiology and has five key assumptions made in its derivation.

1. That the community is **"closed"**. That is that there is no birth, death, or movement of individuals into or out of the community.
2. That the community can be **"compartmentalized"**. That is that each individual can be precisely allocated to one of three states of infection (susceptible, infectious and removed).
3. That flow between disease states is **"deterministic"**. Whilst we talk of individuals for intuitive model construction, when turned into a mathematical model the model predicts the fraction of the population in each state and so is a representative measure of disease in community. Moreover, each time the model is run the same answer will be returned, there is no randomness or uncertainty in the result.
4. Transition from infectious state to removed state is **"memory-less"**. In other words, the

people stop being infectious at a constant rate and Markovian (how long a person has been infectious in the model has no bearing on whether they will lose infectiousness in the next hour/day).

5. Human contacts are governed by the **"mass action"** (or homogeneous mixing) approximation. Infection happens between each susceptible and infectious individual in the population at a constant rate.

These assumptions give a set of equations from which three critical concepts emerge: the **reproduction number** (a threshold value that if above one an epidemic will occur and if below one then it will not occur); **herd immunity** (the critical fraction of the community that requires vaccination to mean that the reproduction number is below one) and the final attack ratio (given a reproduction number what fraction of the community will have been infectious at some stage). From these models we can observe that epidemics fade out because of a lack of infectious cases rather than a lack of susceptible ones. However, these high level insights need to be checked in different settings and if multiple interventions are going to be deployed, or interventions have logistic constraints, then more sophisticated modeling may be necessary.

These assumptions may be relaxed in a number of ways.

- **Closed:** we may allow for birth and natural death, and even aging of the population by compartmentalising the community into age groups. We may break the community into smaller neighborhoods to allow movement between and allow for external travel (say international flights bringing disease into the country).
- **Compartmentalized:** we may allow for additional compartments to allow for interventions (vaccination, quarantine) as well as travelers (see above). We can also allow for additional disease states – so that people can be incubating the disease, asymptomatic, infectious with different transmissibility rates.
- **Deterministic:** probability theory may be applied and the transition between states can be stochastic random events. This is critical in early and late stages of epidemics, and for small outbreaks, especially if we want to retain the concept of an individual.
- **Memory-less:** a disease state may be split into smaller partitions. Mathematically this retains the memory-less transition between one sub-stage and the next in the disease progression but reduces the probability of rapid or slow residency in a specific disease state. Alternative model constructs (agent-based simulation) would allow other disease progression models.
- **Mass-action:** contacts may be viewed on static or dynamic networks (with assumed connection typically being taken as random, all people connected, or just nearest

neighbors on network), and infection may be a random event. Furthermore different contact rates in different age groups and neighborhoods can add heterogeneity. It is very hard to collect representative human contact data, though the recent BBC pandemic documentary may provide some useful insights.

Expectation and output: A simple model may have large output uncertainty while a complex model may have large input uncertainty. However, the assumptions made impact on the decisions we might want to make so they need to be efficiently communicated to the decision maker. At present no simple means of communicating these assumptions is available and so this workshop aims to develop a method of visually representing the assumptions underpinning a model output.

Material format: The data from Challenges 2 and 3 will be made available, example images from Challenges 2 and 3 can be seen below. Please reference Challenges 2 and 3 descriptions if you would like further information.

5.2 Group Members

Min Chen	University of Oxford	min.chen@oerc.ox.ac.uk
Soufiane El Fassi	Cranfield University	s.el-fassi@cranfield.ac.uk
Virginia Foot	Dstl	vefoot@dstl.gov.uk
Ian Hall	PHE / University of Manchester	ian.hall@manchester.ac.uk
David Haw	Imperial College London	dhaw@imperial.ac.uk
Polina Levontin	Imperial College London	levontin@hotmail.com
Kai Xu	Middlesex University	k.xu@mdx.ac.uk

The presentation given by the group at the Study Group can be found here:

https://youtu.be/0wKXdxEdwj0

5.3 Approaches and Progress

This group approached the challenge by firstly reviewing relevant cases and examples from other disciplines where model structure had been visualized. Secondly, the group focused on how one might consider visualizing the assumption of model complexity and population mixing.

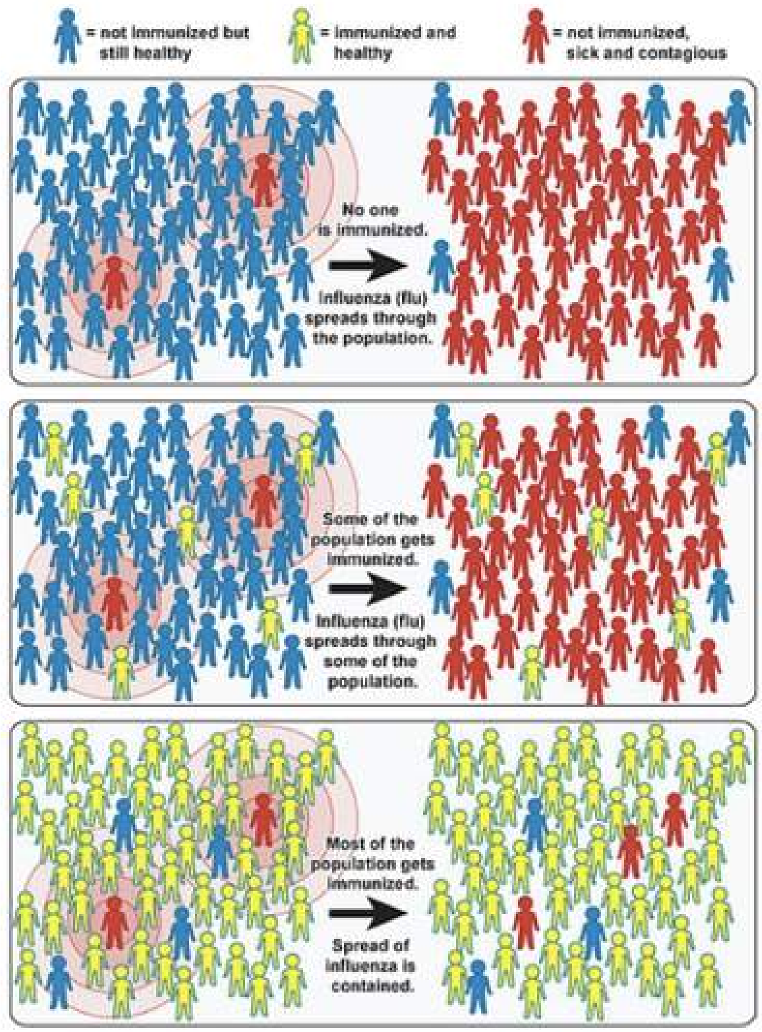
Review of Illustrative Examples: Interactive Examples

- Best looking visualization of showing epidemic evolving over time: [1]
- A dashboard with modeling examples but the inputs are very simply presented in ui, the idea is to replace a certain number of assumptions with more descriptive pictograms [2]
- Gif showing a spread of cases, one possibility is to use gif type pictograms to describe model assumptions. [3]
- Spread Gif, another example of showing spread, this time, plume like - could be a concept for indicating that the model is spatial, even if results are not displayed spatially. [4]
- Animation - moving between classes. A possible concept for gif showing an assumption of modeling infection spread among age groups - young infecting old, etc. [5]

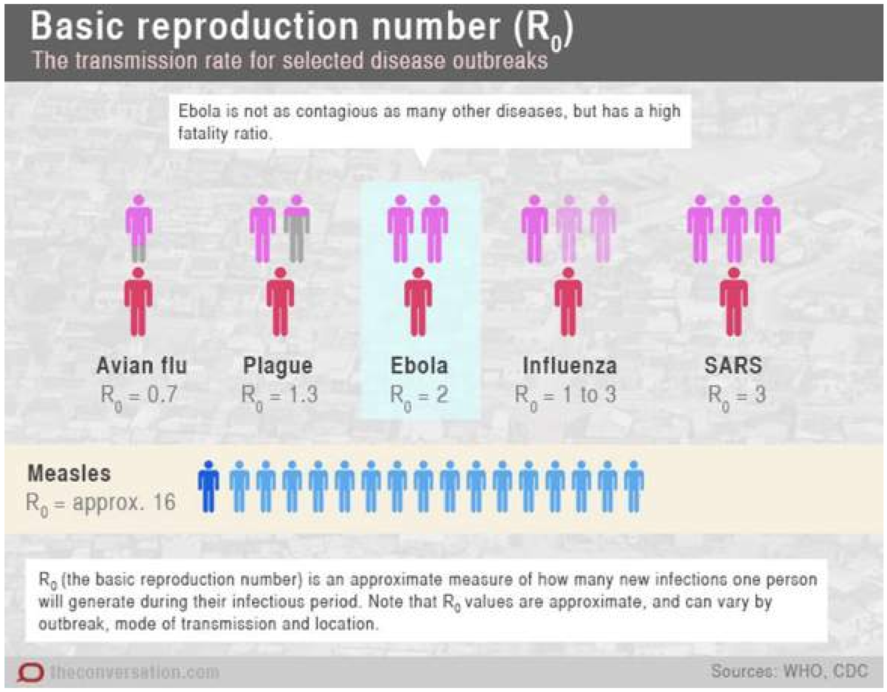
[1] https://www.nytimes.com/interactive/2014/11/04/health/visuals-ebola-model.html
[2] https://www.khanacademy.org/science/health-and-medicine/current-issues-in-health-and-medicine/ebola-outbreak/p/modelling-an-epidemic
[3] https://commons.wikimedia.org/wiki/File:Diabetes_County_level_estimates_2004-2009.gif
[4] https://io9.gizmodo.com/map-shows-how-quickly-a-zombie-outbreak-would-spread-fr-1692010744
[5] https://www.nytimes.com/interactive/2018/03/19/upshot/race-class-white-and-black-men.html

Review of Illustrative Examples: Non-interactive Examples

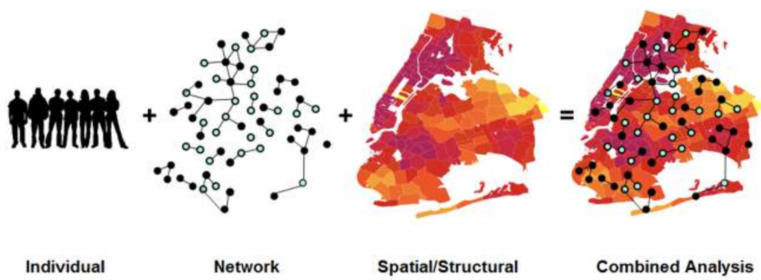
Visualizing Interventions



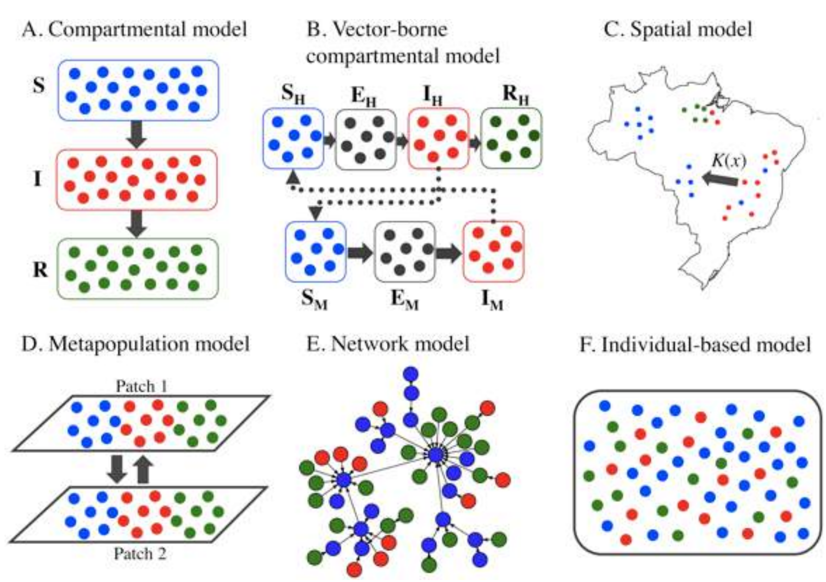
Visualising R_0 if it is a major source of uncertainty



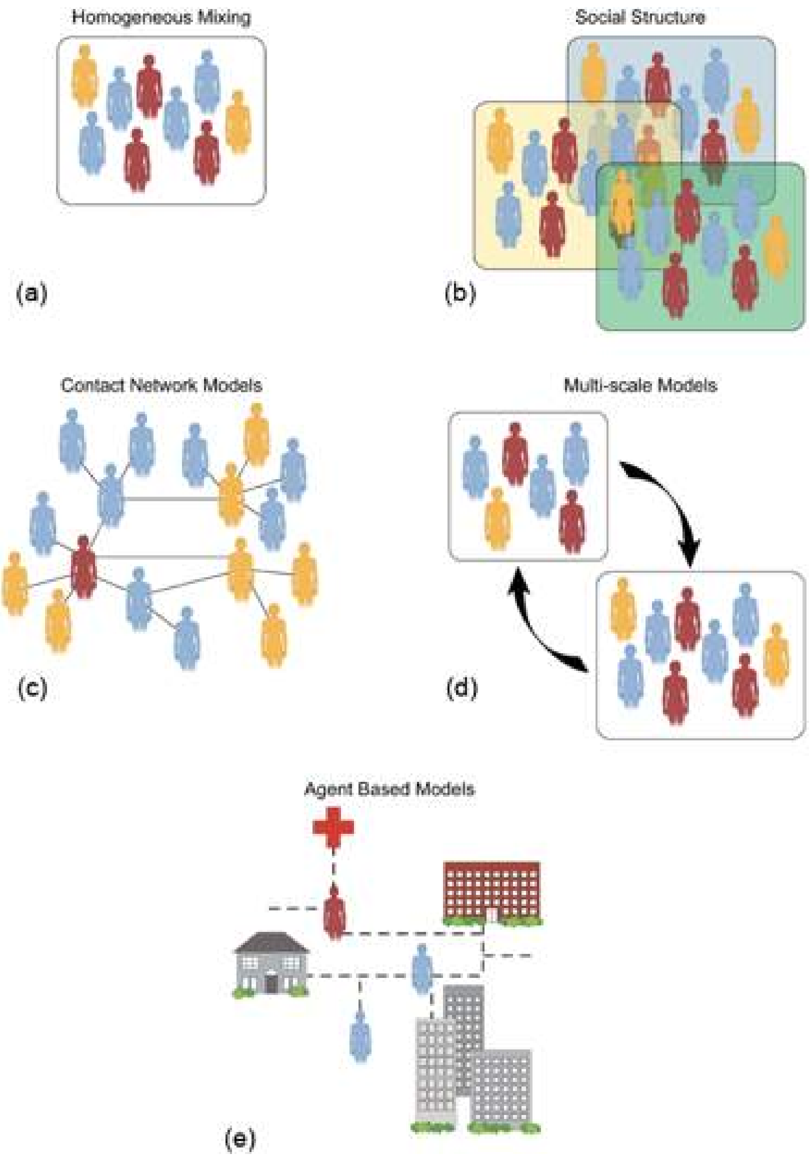
Visualizing modeling approach which combines agent-based, social and spatial modelling:



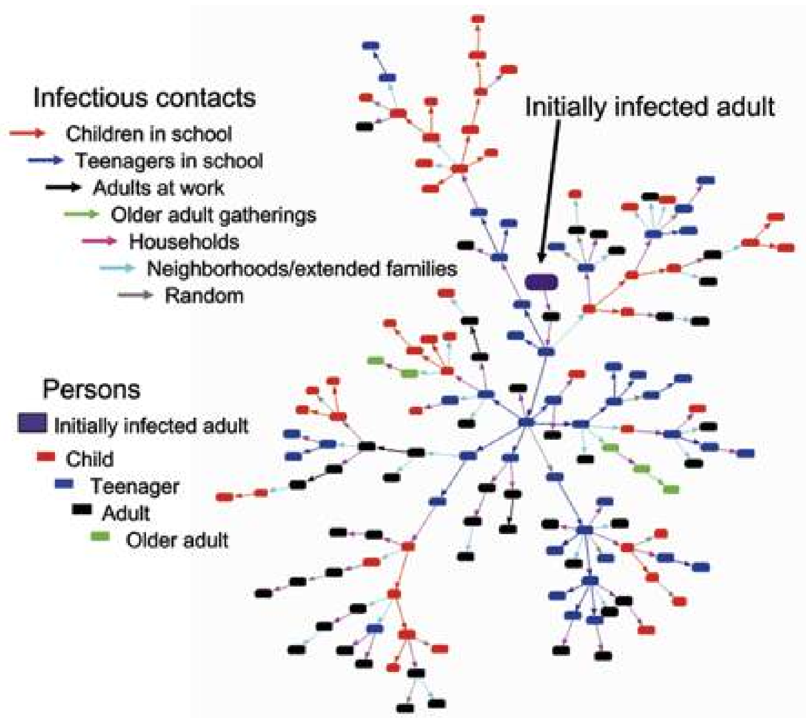
Examples of icons representing different modeling assumptions



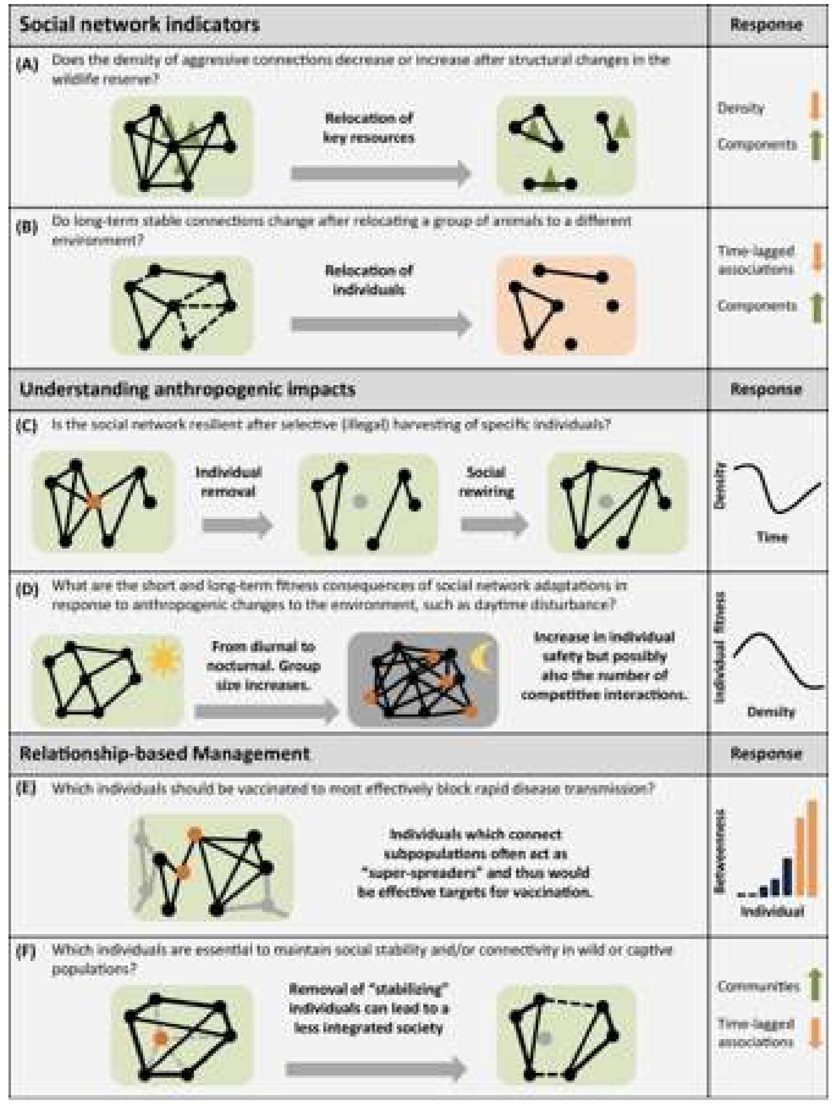
Another example to visualize modeling assumptions. This would be a good starting point for visualizing different assumptions about mixing. Suggested improvements to this visualization: use the same total number of icons for each modeling scenario. Use age type icons (shape = age). Color to indicate a state of illness especially if the sick person is in the centre of a connected network (test with audiences if color in the network matter to how this information is perceived?).



A way to communicate features of modeling social contact

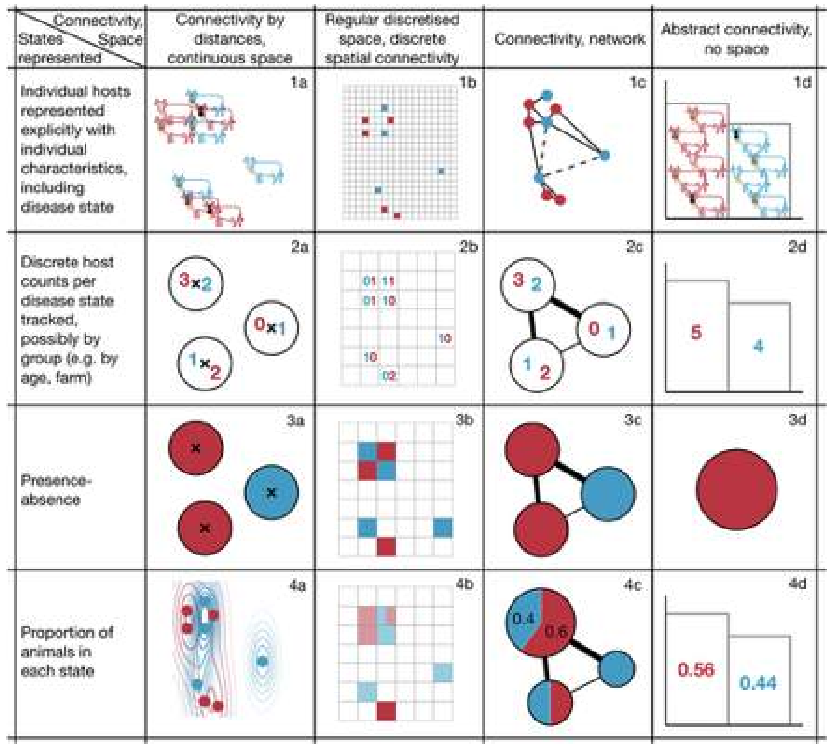


An example of linking some features of modeling to outcomes

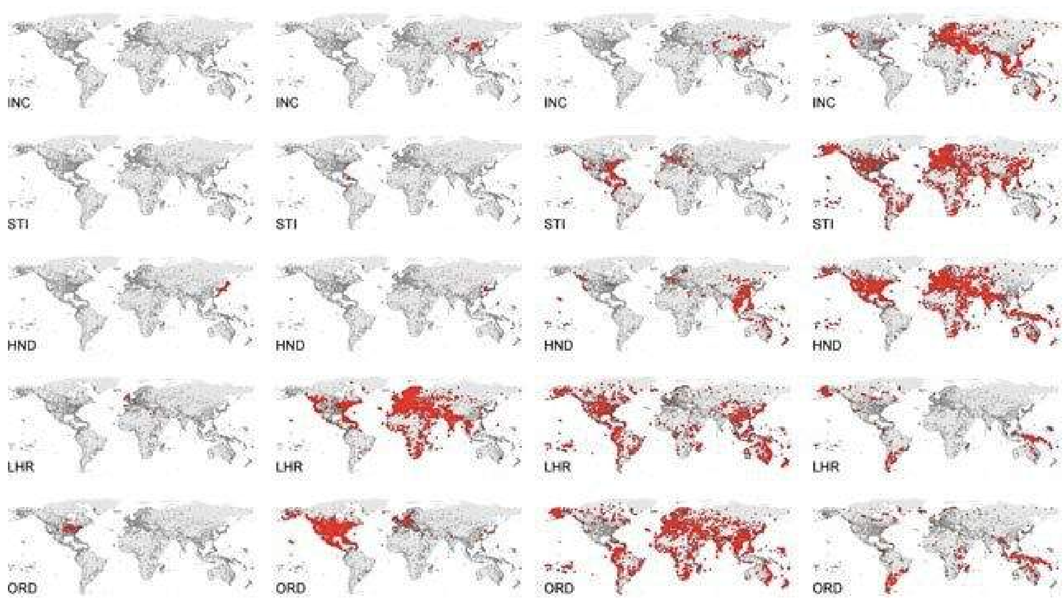


Trends in Ecology & Evolution

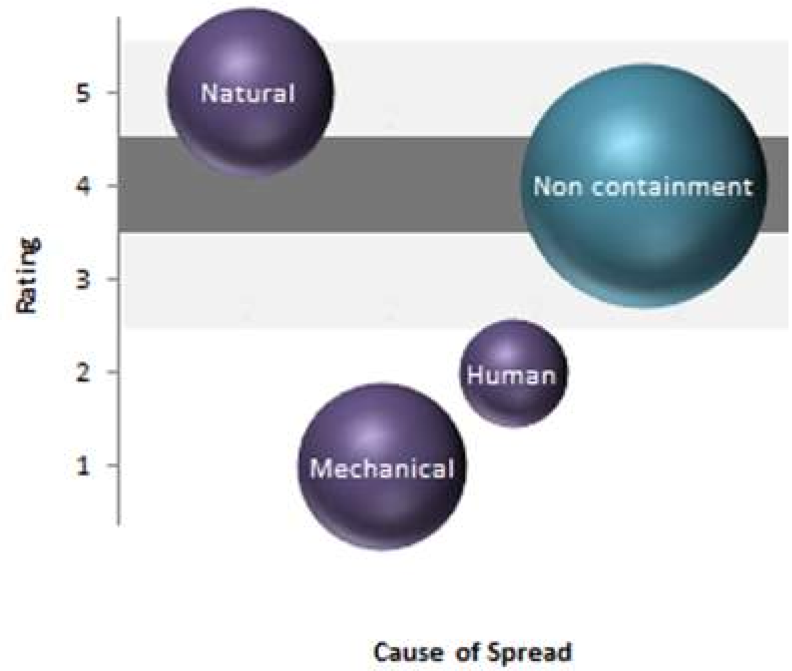
Another way to communicate assumptions



Use of labels and image position to communicate modeling results under different assumptions



Putting the assumption on the X axis and some measure of risk on the Y axis with size representing another dimension of modeling results



A way to represent analyst's uncertainty about assumptions in the model, this is a visualization that is part of elicitation from experts involved in modeling. Basically for a list of assumptions/uncertainty each modeler is asked about importance in terms of changing modeling results, the degree to which the current model already takes each source of uncertainty into account, and the size depicts the potential to reduce uncertainty if more data is collected. This is another way to give a sense to the decision maker of how the analysts feels about the model in terms of its key assumptions and uncertainties

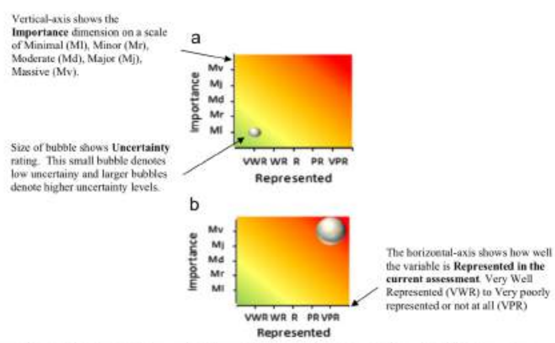


Fig. 1. Visualizing responses by individual experts: (a) example of *lowest* risk extreme characterized by being of minimal Importance, very well represented in the assessment and very low knowledge uncertainty. The Bubble consequently occupies the low risk bottom-left green zone and (b) *highest risk* extreme caused by massive Importance, very poorly Represented in the Assessment, and very high Knowledge uncertainty. The Bubble occupies the upper-right red zone indicating a high priority variable.

Comparing models using Negative Space idea. Basically the columns correspond to different models or cases (for example if we want to compare models of known flu and models of something less well understood), the rows are features of the simulation that contribute to uncertainty: data, knowledge that enables the formulation of structural hypotheses, the bottom part summarizes how much each model was interrogated and tested prior to use (the purple are questions concerning common model validation tests, can be changed to suit epidemiology models) and the last couple of rows are about the common sources of uncertainties and it is a basic check list (e.g. was sensitivity to uncertainty about assumptions in transmission matrix between age groups tested, and so on - create a list that is relevant)

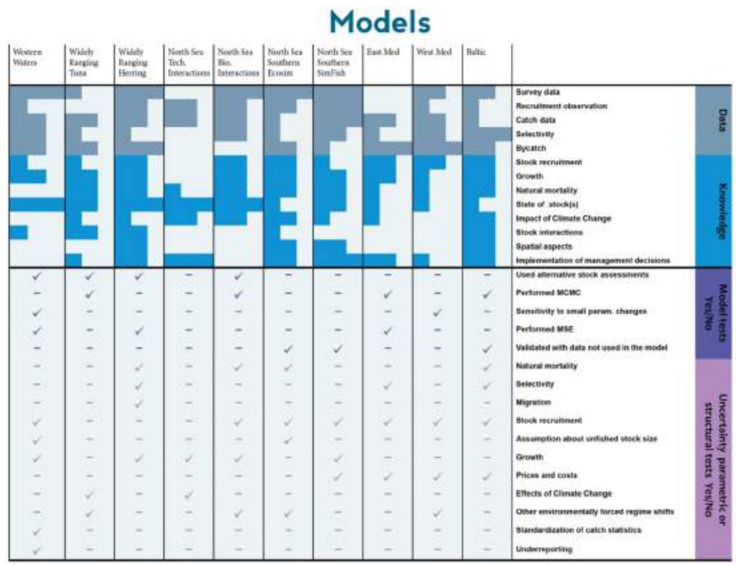
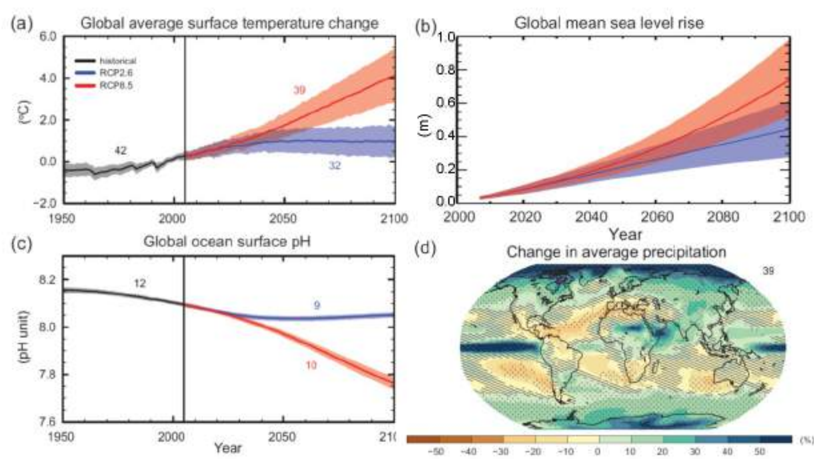
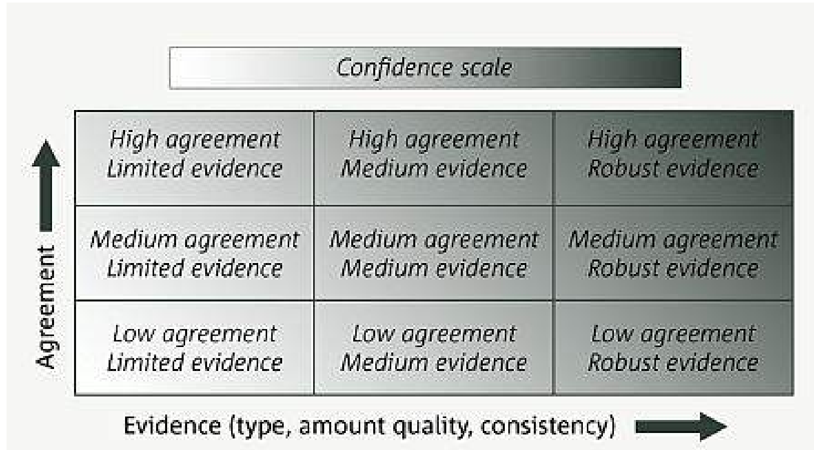


Fig. 6. Model comparisons. Western Mediterranean model relative to other case studies (West Med column).

Combining maps and graphs or different type of visualization in one image; this an example of a layout more than anything else, it might be useful to show graphs describing model output in various



Confidence in the modeling result, after IPCC. Better if it is color coded. A way to show how much confidence a modeler has in a particular set of results so as that their uncertainty can be communicated to decision-makers



Looking at different assumptions, interventions, and criteria for comparison. For example, the gray row can represent some unacceptable level of risk, the sub-columns can be different criteria (i.e. instead of low SSB stock spawning biomass it can be risk or running out of hospital beds, instead of unsustainable F [fishing mortality] it can be shortage of vaccines, etc), then the grid can be different model assumptions and interventions, it allows to quickly see which model assumptions and interventions result in prediction of unacceptable outcomes

Page 14 of 19 T. R. Carruthers et al.

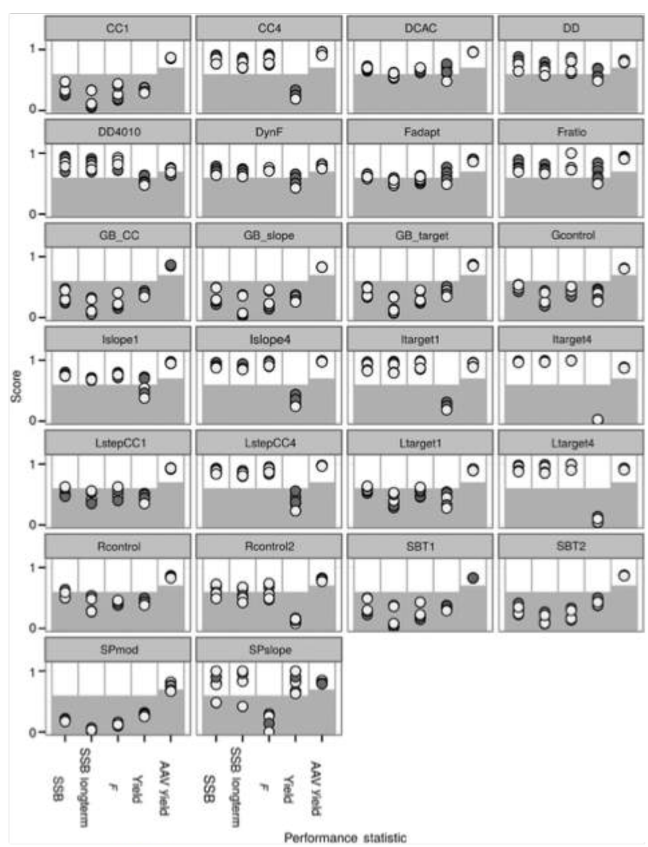
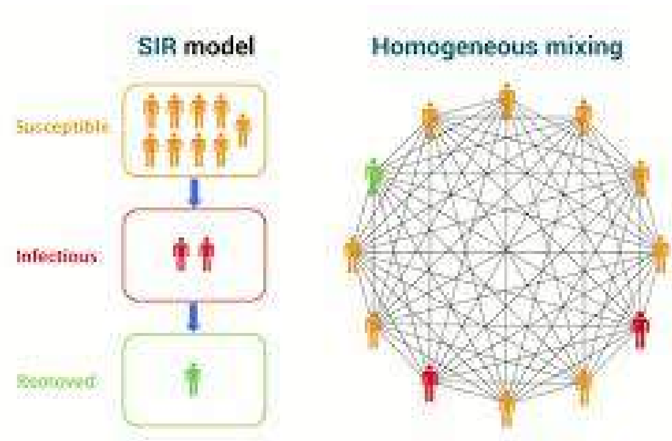
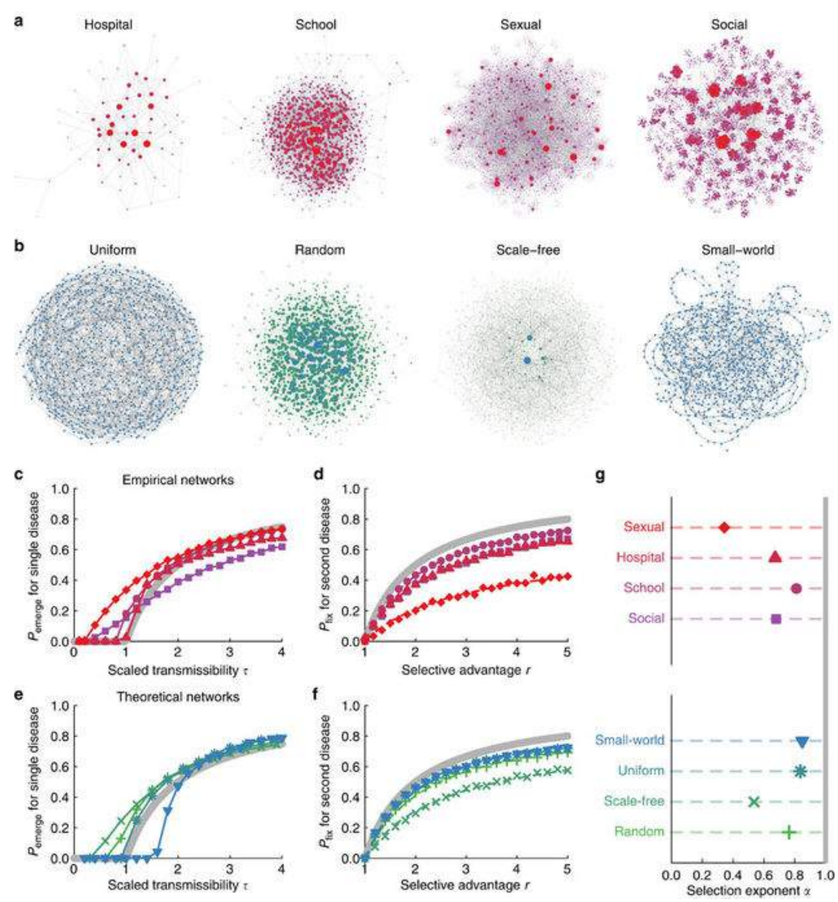


Figure 6. Choice plots (Kell, 2015) summarizing the performance of MPs given the herring life history. Each point represents performance of the MP for a given operating model (data quality and level of recruitment autocorrelation). The score is the frequency of simulations for which a performance goal was achieved. For example an SSB point at 0.91 indicates that 91% of simulations succeeded in keeping SSB above the target level of 50% of *MSY* levels. White areas represent regions that exceed the target performance level. White and grey points represent simulations with low and high recruitment autocorrelation, respectively.

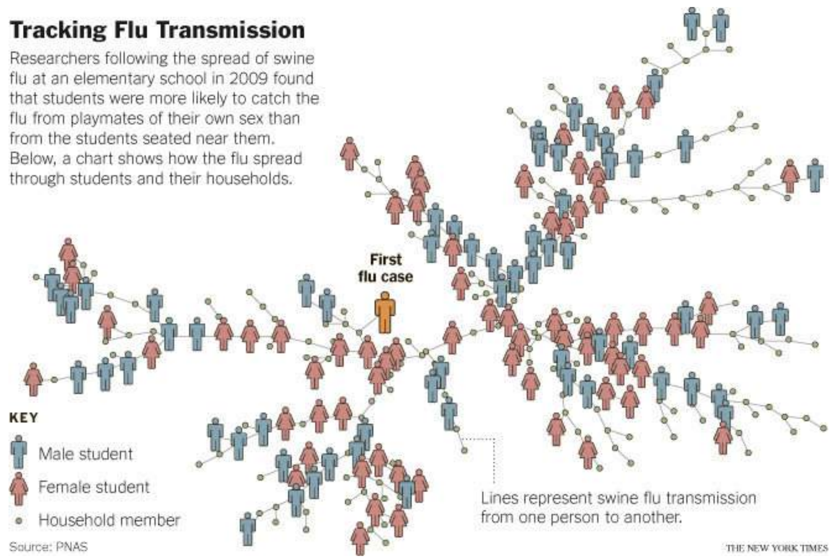
SIR and mixing: can expand the flow chart to include more complex cases. We discussed using a fixed number of icons in different variations to show different model assumptions



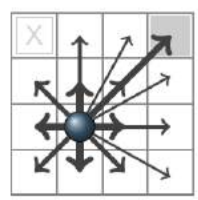
This is an example of combining icons / pictograms depicting assumptions in the model (color coded), here the top red row represents environment or a type of social network for transmission (and corresponds to data) and the blue / green row represents the way this data is modeled mathematically. Below are the comparisons in terms of deterministic graph corresponding to empirical network vs. mathematical abstraction models of the same thing



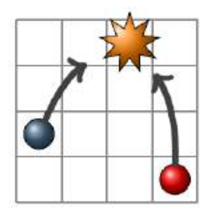
Transmission, a concept for illustrating some specific assumptions about transmission pattern (here it is showing that the risk of transmission is gender dependent)



A schematic way to show how probability of infection is modeled in space represented by a grid

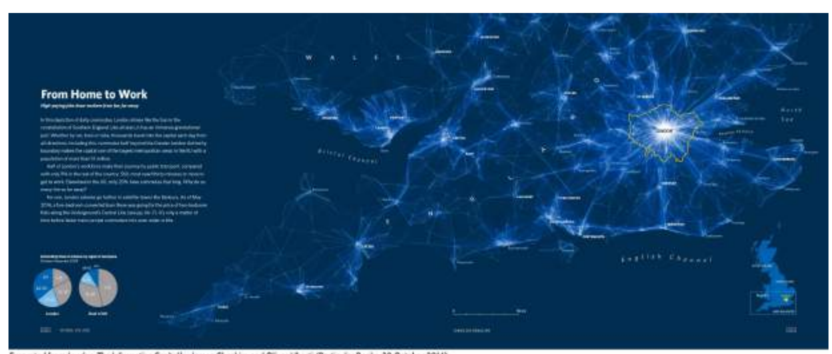


Spatial Kernel



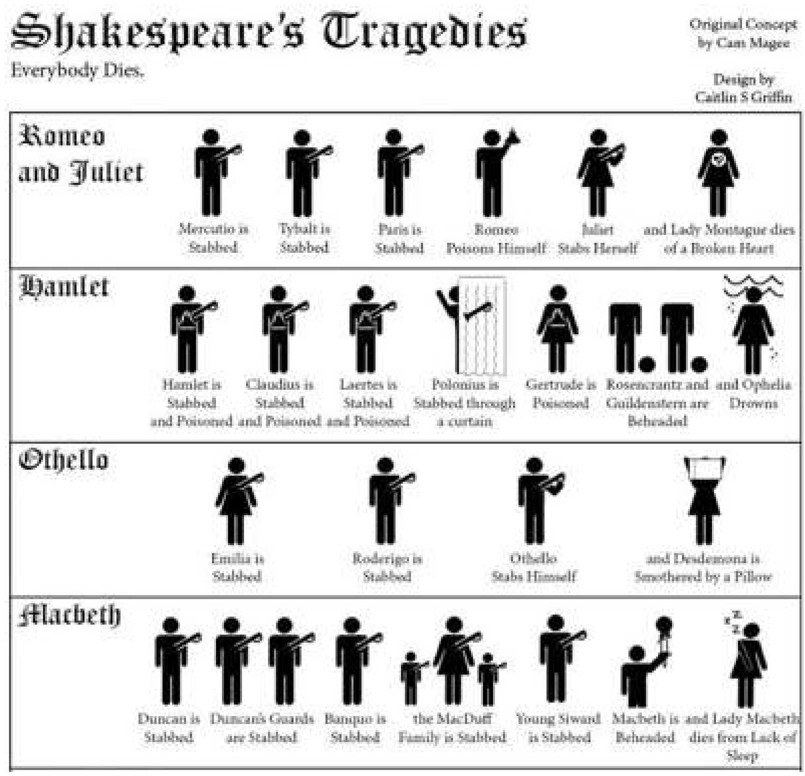
Force of Infection

UK network - a concept for showing how the geographical likelihood of spread is modeled if for instance commuting data is used, this could be a basis for drawing an icon saying: 'in this model we accounted for commuter behavior'

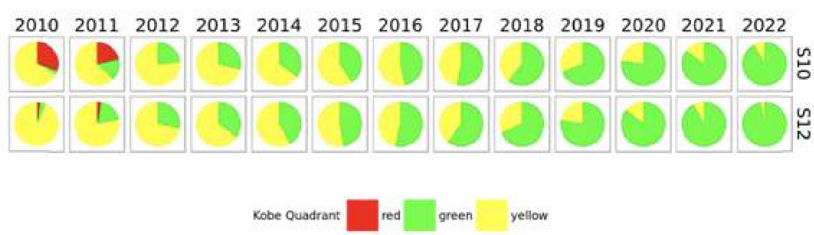


Excerpted from *London: The Information Capital* by James Cheshire and Oliver Uberti (Particular Books, 30 October 2014)

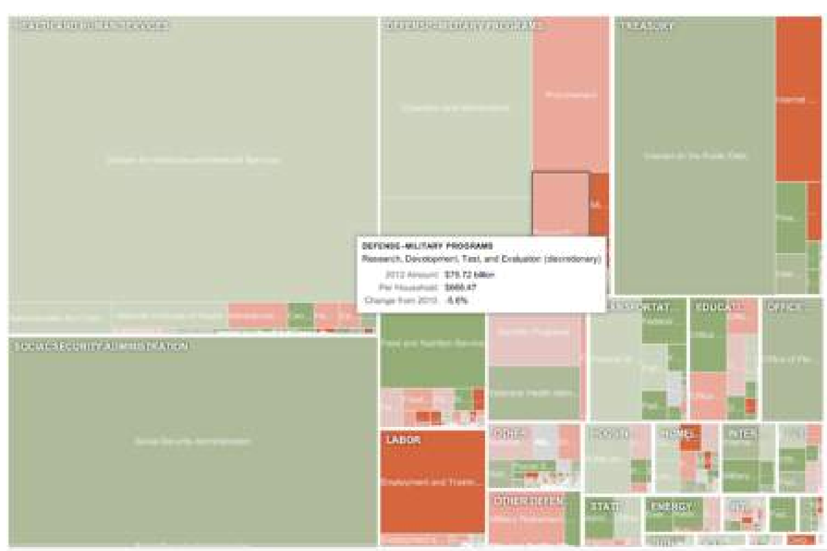
Representing outcomes under different narratives - the basic concept here is using icon arrays to quickly sum up the differences between results based on different interventions, or models arranged as rows in a grid



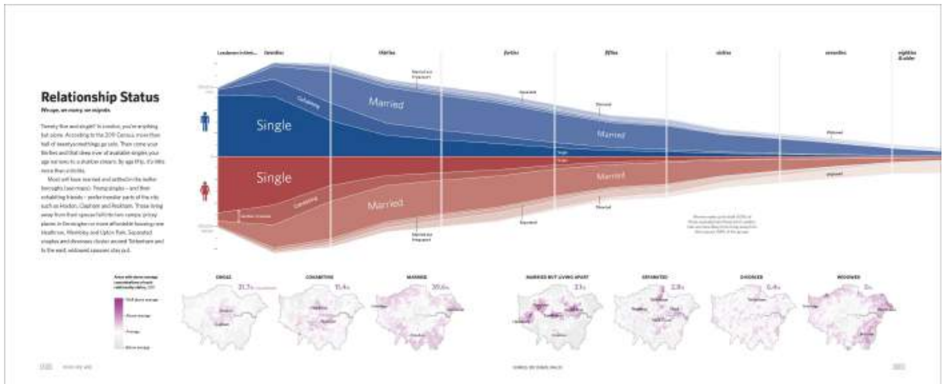
Pie charts over time - another way to compare dynamic behavior under different models, or interventions



A concept for representing relative importance of assumptions in a particular scenario, where the total area is the assumption space and the area of sub-rectangles could be interpreted in terms of sensitivity of model results to different assumptions. For example, if school closures are considered, assumptions about age and social parameters might be shown as largest and draw attention to the importance of examining their implications. These can be pre-calculated based on simulated data for most typical scenarios.

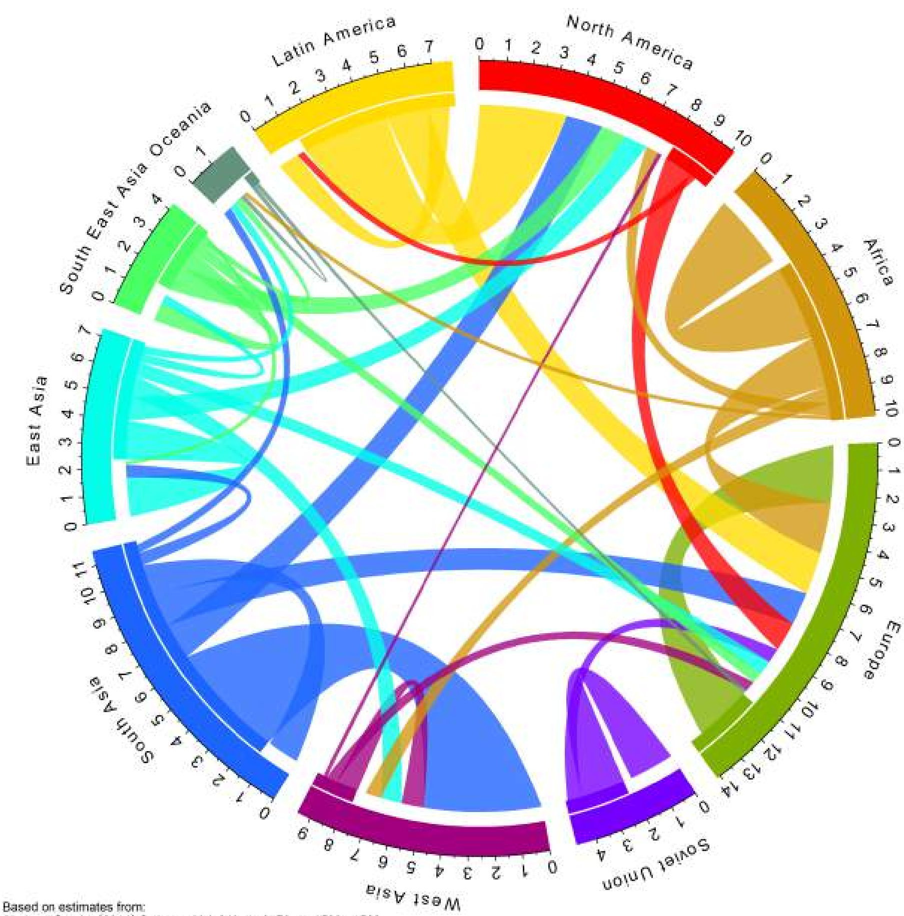


A possible way for visualizing results replacing sick/dead for of male/female and having time on the X axis, and the combination with a spatial information below could be useful



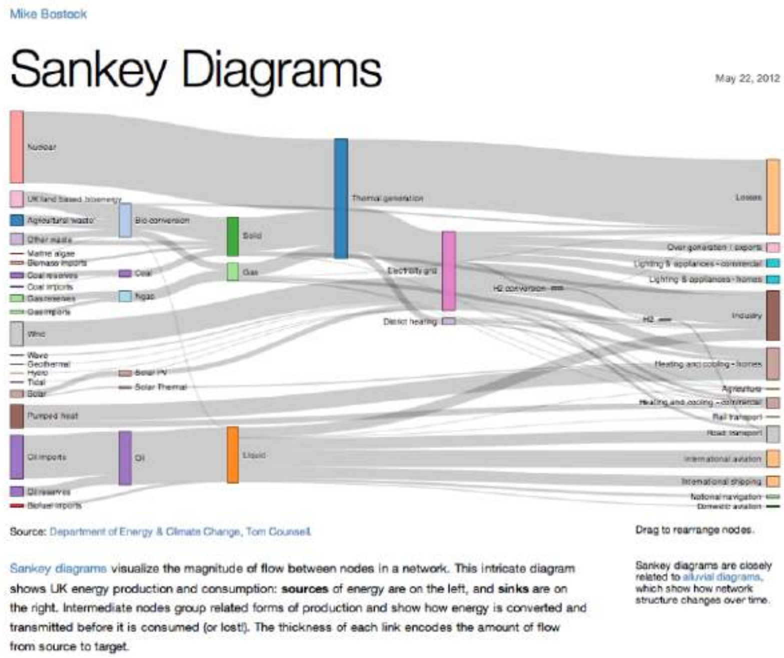
Excerpted from *London: The Information Capital* by James Cheshire and Oliver Uberti (Particular Books, 30 October 2014)

A visual concept for showing how transmission is modeled, replacing countries with age groups of spatial groups or another kind of grouping (risk groups)

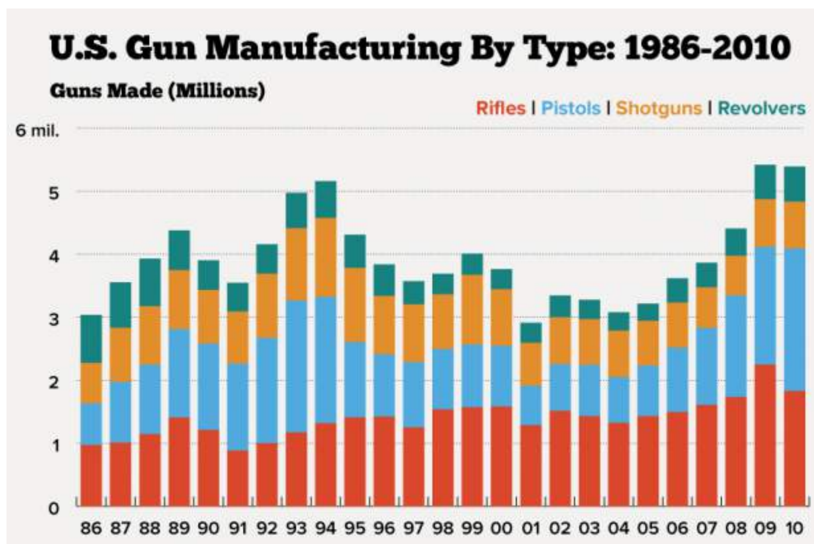


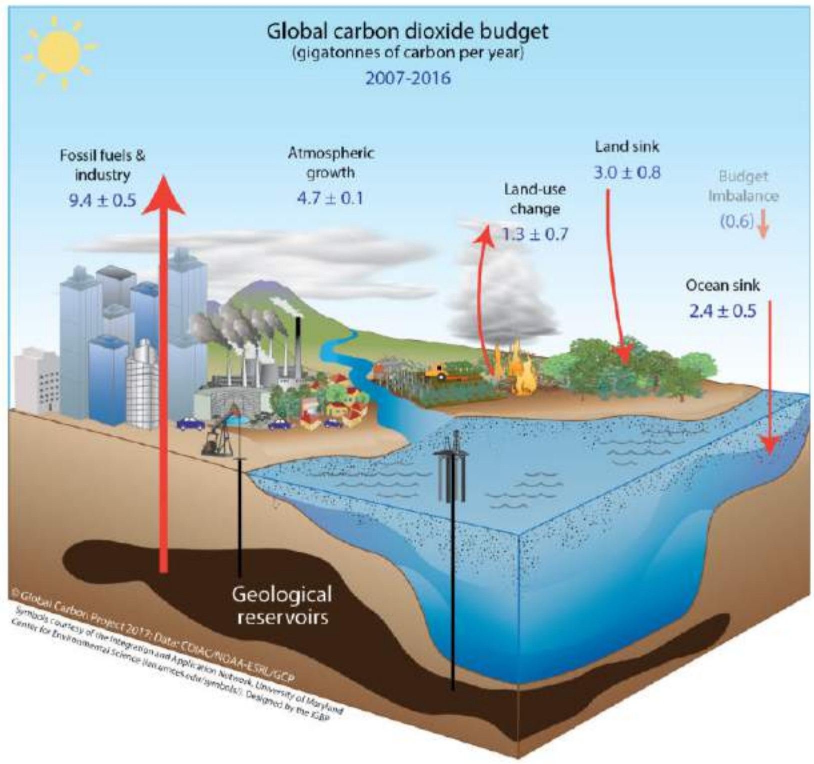
Based on estimates from:
Abel and Sander (2014) *Science* Vol. 343 no. 6178 pp. 1520 – 1522

A more complicated concept for showing how transmission between groups can be impacted by an intervention such as school closures in the middle



Use stack bar chart to see where infections are coming from over time

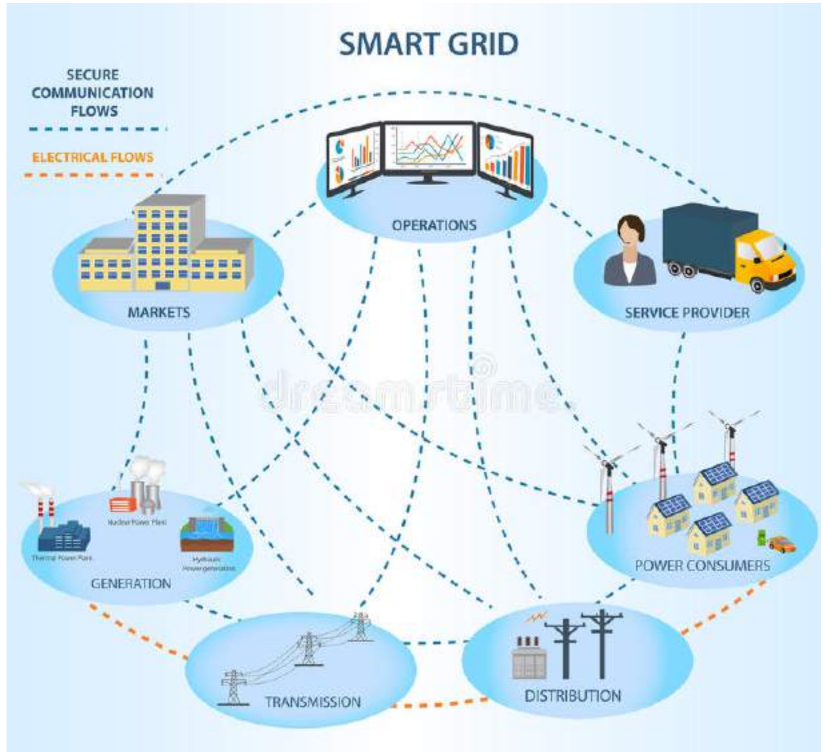




A way of presenting a set of model assumption as a picture with interchangeable elements such as buildings or no buildings, wind or no wind, moving icons of people or static, people of different ages or homogeneous population. A concept:

We also discussed a possibility of having pictograms for different assumptions that can be highlighted or greyed out or simply exclude and it can be one alternative visualization for a total set of assumptions. A concept is below, but the icons can be variations of the visual ideas above. Further, the icons for modeling assumptions need not be arranged as a grid but as a tree or as a conceptual model graph.

Icons arranged as a conceptual model:



5.4 Application to challenge

Using the learnings from the literature search above, the group drew in two ideas for further development. The first approach focuses on providing a pictorial representation of the underlying model structure to the user. The purpose of this being to remove the potential of a black box effect so that users are able to interrogate the structure of the mathematical model in a non-mathematical way.

If we consider a basic SIR model, then we might represent our flow diagram as the lefthand image in Fig. 8. However, this hides details from the user on the possibility of more complicated terms being introduced; for example a vaccination term V, an exposed term E, numerous susceptible states S_i etc.

A more complete way to visualize the structure of the model, and modeling complexity might be shown on the right hand side of Fig. 8. This more complicated graphic is equivalent to the left hand if those terms not considered are grayed out. One might consider a glyph which could represent such a model structure.

If we begin to include these higher modeling assumptions, which again could be symbolized as a glyph, a possible flow structure might look as the below

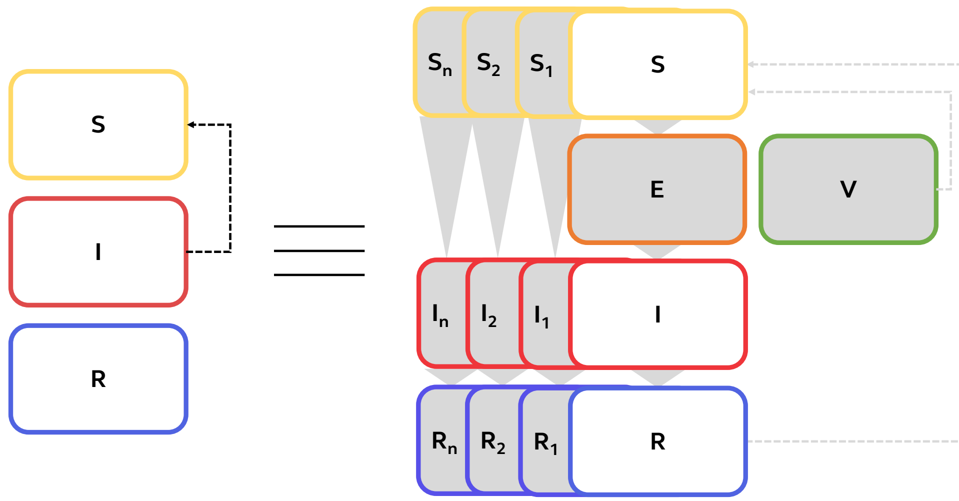


Figure 8: (left) SIR flowchart description (right) more complexity of SIR model.

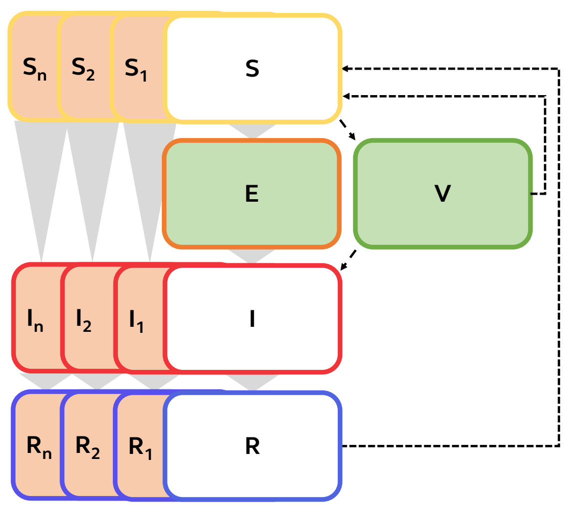


Figure 9: SIR flowchart description with enabled further complexity.

The assumption of how contact is made determines where and who may get infected; knowledge of this assumption is therefore vitally important for a decision maker. As has been suggested above, glyphs could be used to statically represent the mixing assumption. We might also consider how animated rendering of the graph might convey information on the assumptions which underpin it. With reference to Section. 8, care should be exercised when considering how many 'multiples' of model assumptions we have as this would limit the amount of bespoke representations we are able to create.

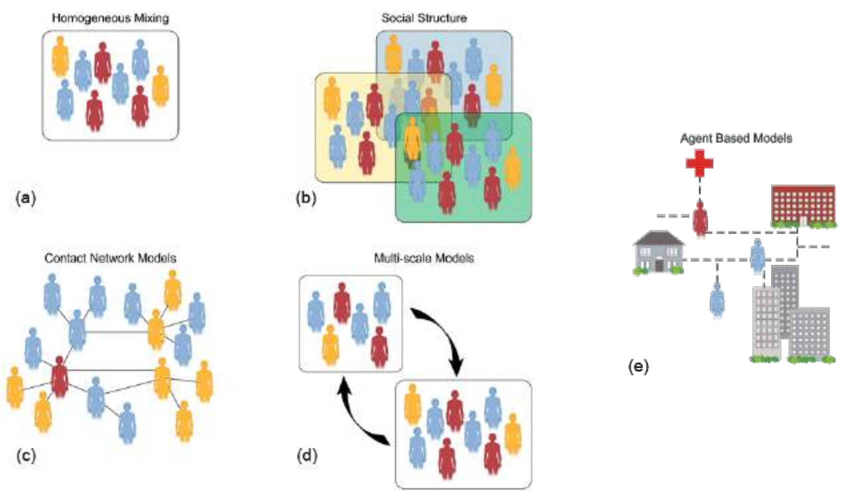


Figure 10: This schematic shows increasingly structured and complex ways to model the spread of a disease in a population. Each color represents a disease stage: susceptible, infected, or recovered. (a) Homogeneous mixing models treat each individual at a given disease stage in the same way as the average of all others at that disease stage. Structure can be added by decomposing populations into distinct demographic groups (b), where the spread of a disease will depend on the particular group. (c) In a contact network model, links connect individuals while multiscale models (d) involve connections between subpopulations. (e) Agent-based models can track individuals in very large populations on the order of cities the size of Chicago. From Ref. [31] and references therein

Fig. **??** shows ideas from the literature on how this could be done. This would be a good starting point for visualizing different assumptions about mixing. Suggested improvements to this visualization could be to use the same total number of icons for each modeling scenario. Use age type icons (shape = age). Color to indicate a state of illness especially if the sick person is in the centre of a connected network (test with audiences if color in the

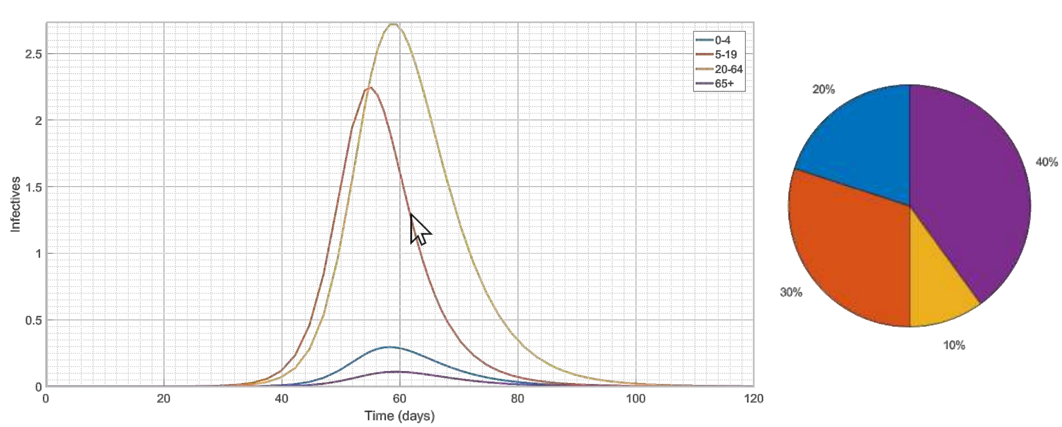


Figure 11: SIR infection curves and decomposition chart for WAIFW.

network matter to how this information is perceived?).

We now consider how the user might be able to interrogate the model predictions to determine some information on both the kind of population mixing assumptions, and also the effect this has on the output. Fig. 11 shows four infection curves from four different age groups. The fact that these curves are different suggests some assumption on infection rate has been made based on the age.

Fig. 11 shows that by selecting one of the age curves, a summary pie-chart appears. This chart shows the decomposition of the force of infection / WAIFW. Examination of this decomposition provides information to the decision maker on what assumptions are being made on the age-mixing parameters of the model.

We can see from the decomposition in Fig. 11 that the majority of the infections for the age range 5 - 19 are acquired from the over 65 age range. The user can then use their judgment as to whether that assumption is valid for the case in hand.

We can continue to explore the effect these assumptions have on the model prediction by

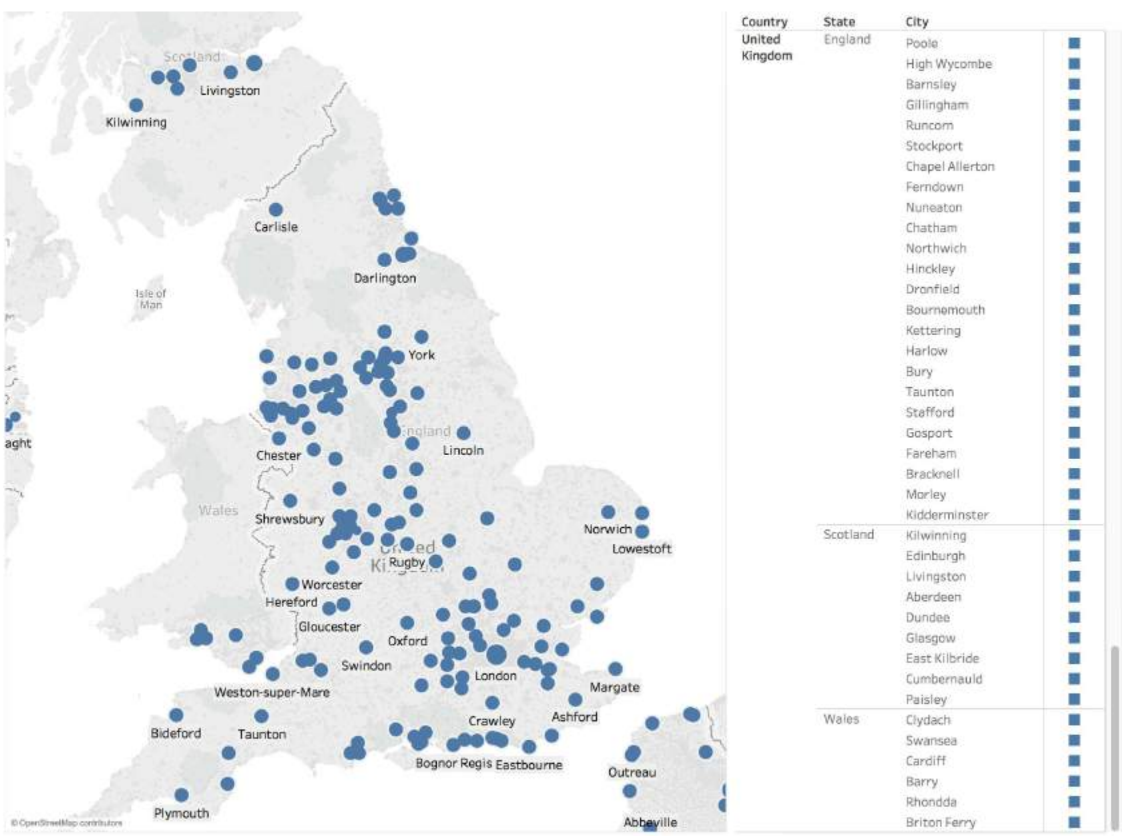


Figure 12: Spatial breakdown of WAIFW using mass action assumption

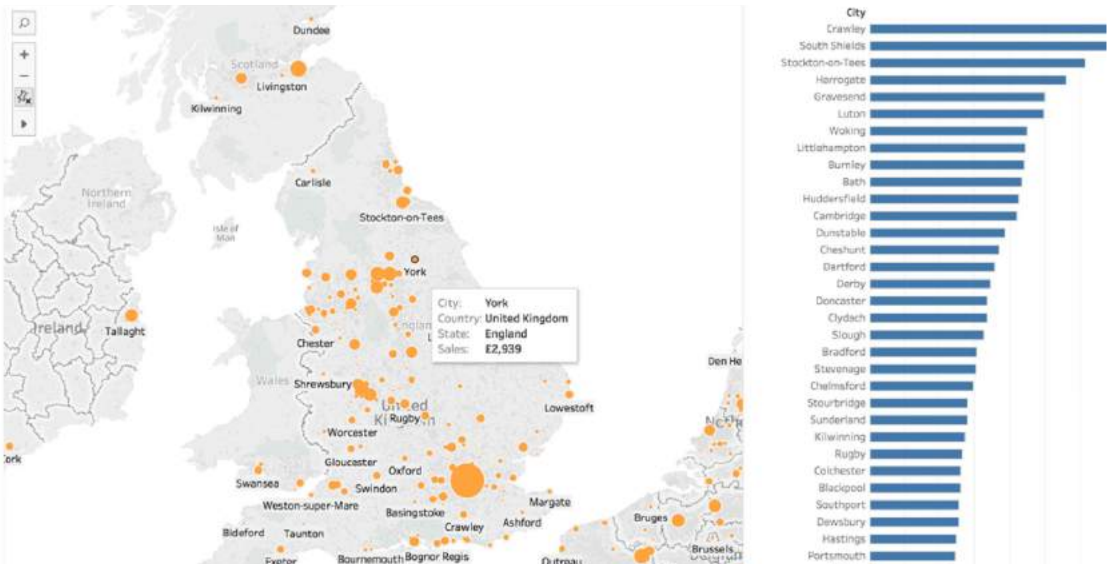


Figure 13: Spatially explicit breakdown, showing the WAIFW contribution from across the UK.

exploring the spatial domain. We see in Fig. 12 that the result of the mass action assumption is that the spatial distribution of the WAIFW term is homogeneous. This assumption is may not be realistic or helpful when considering how towns and cities differ in the infection rate.

In Fig.13 we show what a the breakdown would look like if we had a spatially varying WAIFW term. If you choose *York* for example, then the contribution to infection from each of the constituent areas can be seen This is an intuitive way to visualize the type and effect of the mixing assumption.

6 Challenge 2: Decision Aids That Include Data Summaries and Temporal Predictions

Introduction to Theme: The spread of a transmissible disease (such as influenza) is hard to predict, not least because there are stochastic elements to disease transmission within a population. Computer simulations are often presented showing one possible prediction of the future, however, multiple plausible predictions exist. Effective and comprehensive communication to decision makers is dependent upon illustrating the fullest possible set of information garnered from simulations.

A classic model for describing a transmissible disease may group populations by: those that are uninfected but susceptible to disease; those that are infected (and transmissible); and those that are "removed" for instance, through recovery, acquired immunity, or death. In this example we consider a model that accounts for those who are susceptible, infected, and removed, but we further sub-categorize those removed by whether they recovered with no complication, recovered with complication, or died. Full model equations, parameters, and assumptions are not provided in this workshop so that the focus is on the visualization of the results of such a model.

A number of existing approaches are taken to visualize model outputs such as these. Frequency trees (as illustrated in Fig. 14 (left) have been used to communicate information

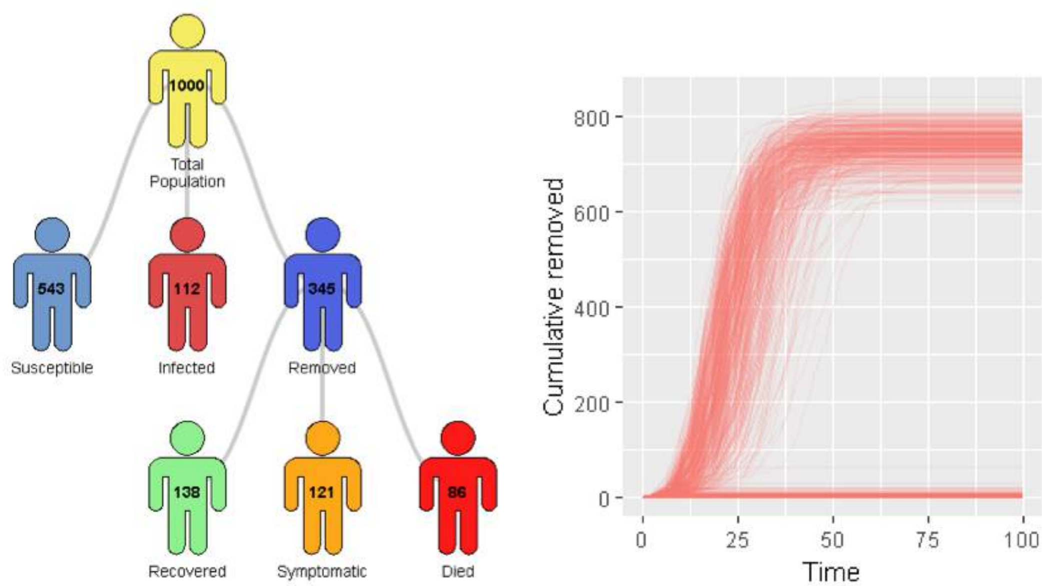


Figure 14: (Left) Frequency tree showing the distribution of the population across the different disease states that are captured by the model. (Right) Cloud plot showing multiple simulations of the same model across time for the removed population.

but not uncertainty. Cloud plots (shown in Fig. 14 (right) in which multiple simulations of the same model are plotted on top of each other with some level of transparency, demonstrate the uncertainty in solutions over time, but can look too 'busy', and may not be effective for problems with multiple variables. Each figure has current value, but novel, improved methods for illustrating the uncertainty of these model outputs are sought.

Expectation of output: The primary output sought is innovative methods to visualize the model outputs such as the one presented here. It is the *concepts* demonstrated in sufficient detail that are required. Prototype computer code would be beneficial, though hand drawn illustrations with any appropriate supplementary material are acceptable.

Material format: A comma separated variable file is provided that contains 1000 realizations of the model described above [6]. The full list of variables are described below.

Variable name	**Description**
Run	A unique identifier indexing which simulation run the data belongs to (numbered 1 to 1000).
Time	A discrete time-point ranging from 0 to 100 (inclusive)
Population	Total population
Susceptible	Susceptible population size
Infected	Infected population size
Removed_total	Removed population size (Total, regardless of intervention)
Removed_recovered	Removed population subcategorised by those that *recover* (in the *absence* of intervention)
Removed_symptoms	Removed population subcategorised by those that have *lasting symptoms* (in the *absence* of intervention)
Removed_die	Removed population subcategorised by those that *die* (in the *absence* of intervention)
Removed_recovered_interventions	Removed population subcategorised by those that *recover* (in the *presence* of intervention)
Removed_symptoms_intervention	Removed population subcategorised by those that have *lasting symptoms* (in the *presence* of intervention)
Removed_die_intervention	Removed population subcategorised by those that *die* (in the presence of intervention)

Additionally, two files are provided to visualize the model outputs using frequency trees (as

[6] https://github.com/mattbktn/uncertainty-visualisation-disease/tree/master/Group2_TemporalUncertainty

demonstrated in Fig. 14). It is up to participants if these form part of any future visualizations. The file "people.html" can be used to view a .csv file. This viewer is not supported in Microsoft Internet Explorer; we suggest Google Chrome is used.

6.1 Group Members

Haeran Cho	University of Bristol	haeran.cho@bristol.ac.uk
Brian Dixon	Glasgow School Art	b.dixon@gsa.ac.uk
James England	Riskaware	james.england@riskaware.co.uk
Toby Pilditch	University College London	tdpilditch@gmail.com
Adrian Pratt	Public Health England	adrian.pratt@phe.gov.uk
Anthony Wilson	Pirbright Institute	anthony.wilson@pirbright.ac.uk

The presentation given by the group at the Study Group can be found here:

https://youtu.be/aoQXl0Ifht4

6.2 Approaches and Progress

The spread of disease is hard to predict, and many plausible computer simulations exist. The communication to decision makers is dependent on illustrating the fullest possible set of information from simulations, visualizing what they need and not what they do not need. Existing approaches include frequency tree (Fig. 14) - this however does not include uncertainty, and importantly for this challenge, they are static in time. Cloud plots are also used, but these tend to look 'busy'. The three tasks explored by the group were:

1. **What questions do decision makers (and to a lesser extent, analysts) need to know?** Decision makers tend to be are busy people and not mathematicians (how might they deal with box plots?) they do however have contextual knowledge which isn't always available to the analysts.
2. **How can the visualizations be misunderstood?** - are they taking too long to get to the conclusion? or worse get to the wrong conclusion.
3. **How can we visualize temporal uncertainty?**

6.2.1 Identifying Questions

The initial task for the group was to explore what are the questions which will be asked of the visualization; this was divided into two types; non-intervention specific questions; and intervention-specific questions.

Non-Intervention specific

What will be the impact?; what is the type of impact (e.g. cases, deaths, hospitalizations)? what different ways to present impact (snapshot, cumulative, peak)? **When** will be the impact?; when does it peak / when is it worst, when does it return to normal? What's the best worst-case scenario? (but also other reasonable scenarios). How **confident** are we in the numbers presented?

Intervention specific:

What is the best case scenario (if we "overreact")? What if we do this (these) intervention(s)? What is the societal costs to the intervention? (adverse events) How quickly do we need to act?

6.3 Avoiding Potential Bias in Presentation Formats

In human judgment and decision making, bias can be considered any form of systematic deviation from optimal. Typically, such deviations are attributable to decision-irrelevant factors. For the present case, these can be considered factors influencing judgments and decisions that stem from the way the information is presented, rather than the information itself. Useful references [14, 18, 19, 34, 32]

The Framing, Order, Complexity (and Uncertainty), Search (FOCUS) acronym below provides a curated list of known biasing factors, and where possible, suggest solutions.

- **Frame:** How outcomes are phrased can influence perception of likelihood / decision preference. For instance, if highlighting a forecast of an intervention on 10,000 people, framing it as "4,000 lives are saved", vs "6,000 lives are lost" can lead to decision makers being more risk averse in the latter case. Caution should be used when interpreting uncertainty perceptions in natural language.

  Solution: Where possible, avoiding unnecessary descriptions of outcomes that carry consequences, but otherwise presenting both sides of the frame, e.g. "X leads to [4,000 lives saved and 6,000 lives lost]".

- **Order:** If presenting information sequentially the order in which information is

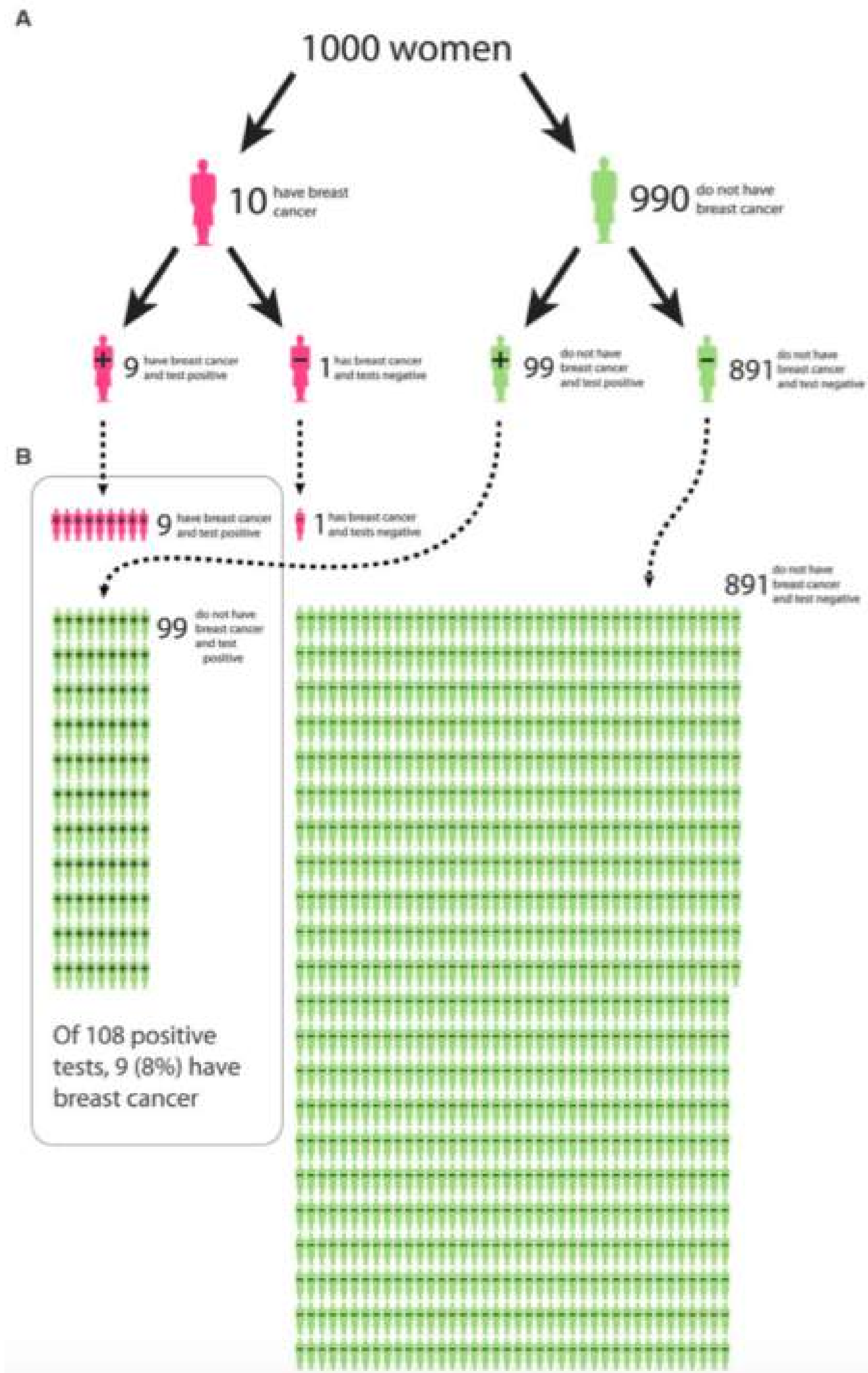


Figure 15: Taken from Ref. [32]. Visualizations of the predictive accuracy of a screening test. (A) Tree diagram showing the consequences for 1000 women attending mammography screening from a population with 1 % prevalence of the disease, when the screening test correctly classifies 90 % of women with cancer and 90 % of women without cancer. Although nearly all the women with cancer are detected, they are greatly outnumbered by false-positive tests arising from those without cancer. (B) Icon array of the same information, which shows explicitly that out of 108 positive tests, only 9 (8 %) would be expected to reveal breast cancer.

presented can lead to systematic over / under-weighting.

- When making judgments step-by-step (during process), the first-presented information has a disproportionate influence (otherwise known as a primacy effect).
- When making judgments at the end-of-sequence, the last-presented information also can have a disproportionate influence (otherwise known as a recency effect).

it Solution: Avoid sequential presentation wherever possible. If unavoidable, be wary that information in the middle-order is the most likely to be overlooked.

- **Complexity and Uncertainty** People are generally poor at probabilistic reasoning in two main ways: oversimplifying interactions (e.g. naively averaging estimates or focusing on min/max values – notable in cases of complexity) and under-adjusting given new information (notable in cases of uncertainty).

  Solutions: Presenting information in natural frequency formats where possible (rather than probabilities), and highlighting exemplar "paths" representing possible interaction outcomes and updates.

  Representing probabilities with graphics; Natural frequencies appear to be superior to percentages in improving understanding of biomedical screening tests [15, 26], owing to the cognitive effort required for interpreting conditional probabilities. Data from the U.K. Breast Cancer Screening Programme Ref. [10] are used in Fig. 15 to illustrate the outcomes of a mammography test on a population with a 1 % prevalence of breast cancer. The test is positive for around 90 % of women with cancer, but it is also positive for around 10 % of women without cancer. The issue Ref. [11] is to communicate the probability that a woman who tests positive actually has breast cancer: The fact that this probability is only 8 % is generally unintuitive. This is a notoriously tricky problem, but understanding can be greatly improved Ref. [15, 27] by supplementing the icon array with a tree diagram. Audiences respond well to multiple types of display of the same information, and the composite diagram in Fig. 15 not only allows part-to-whole comparisons, but also shows how the probabilities are worked out. [32]

- **Search:** When exploration of the information space is left to the user, the type of information and length of search can be influenced by *a priori* / initial expectations or favored hypotheses of the user. More precisely, research has shown searchers tend to seek out confirmatory (expectation/hypothesis consistent) information, and further, can terminate their searches prematurely.

  Solution: Ideally, avoid user searches. If absolutely necessary, caution users against the influence of expectation and early hypotheses - encourage falsification / alternative hypothesis consideration.

6.3.1 What ways are there to visualise uncertainty?

The group discussed and reviewed methods for encoding uncertainty information into visualizations by using, position, opacity (see Fig. **??**), color (see Fig. 16 (left)), size and shape (see Fig. **??**), orientation, texture (see Section. 8), movement (see Fig. 16 (right)), edge blurring, and others. Insight was drawn on how accurately visual encodings of time-series data are interpreted from Ref. [17]. Which ones of these schemes work best for temporal uncertainty in our challenge?

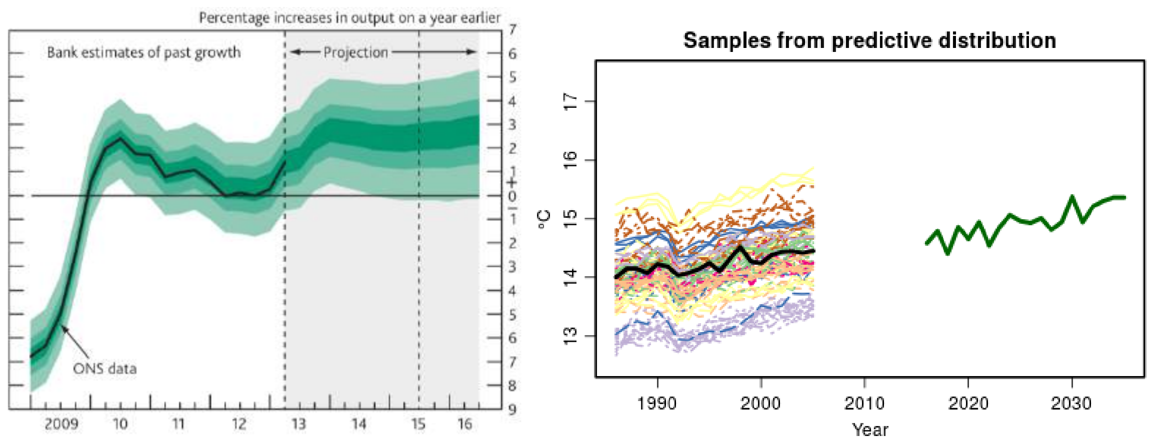


Figure 16: (Left) 2013 projections for GDP based on market interest rate expectations (Right) Annual mean global temperatures, 2016 - 2035 under specified emissions scenario. (a) all 32 models, (b) shows percentiles of uncertainty distributions and (c) shows intervals of specified probability. Data from Mariana Demetriou, created with assistance of Prof. Richard Chandler at University College London.

6.3.2 Overarching approach

Based upon feedback from analysts and their knowledge of the decision makers involved and given the nature of the problem it was noted that often the wider contextual information is not available to the analyst. Thus it was decided that there need be a two-step approach to visualization; one for analyst and one for decision makers. The first step is seen by analysts, they explore based on their expert knowledge and pick which summary plots are taken to step two, for the policy makers. The temporal information can be summarized in plots for decision makers (Fig. 17), these would be customizable by analysts but with recommendations (default report format - likely they'll need info abut peak incidences, halfway points, after a week, etc).

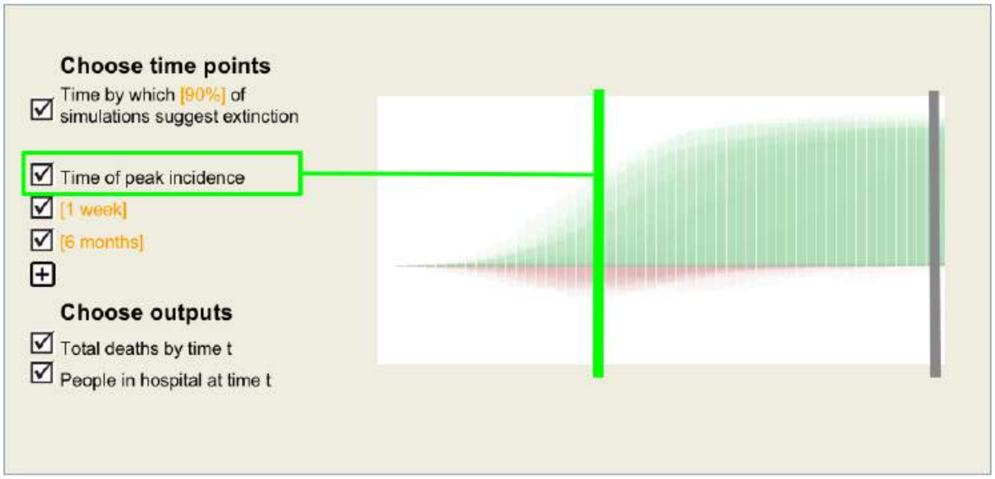


Figure 17: Summary information for temporal predictions.

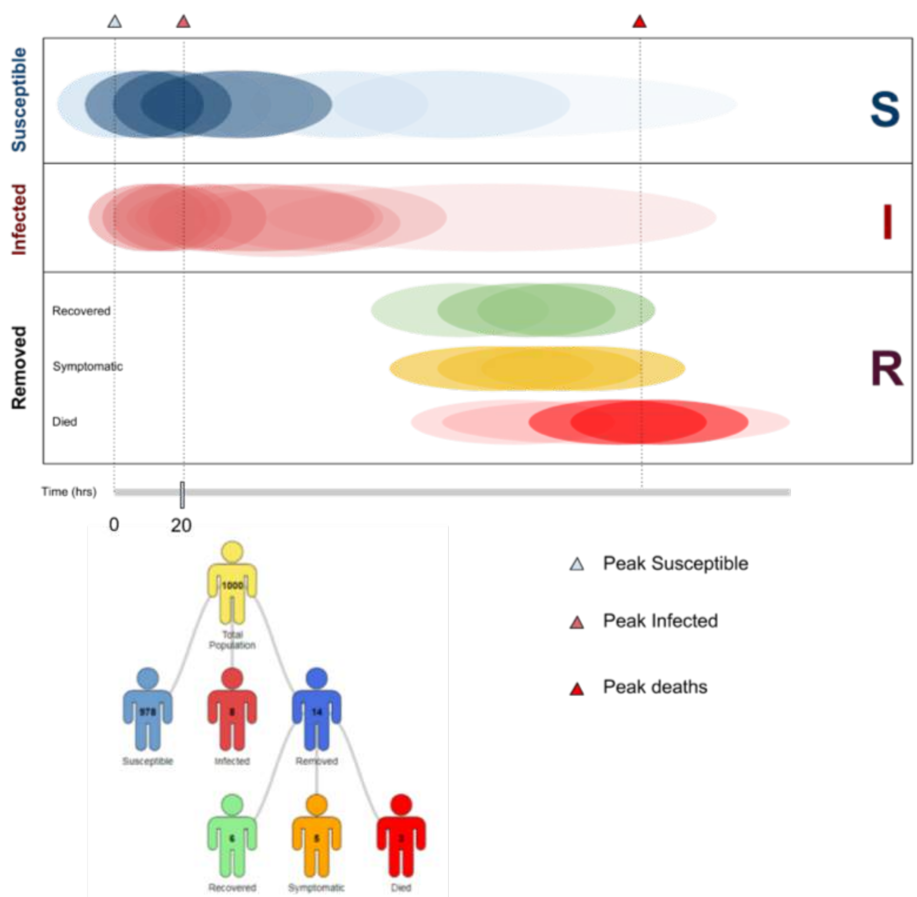


Figure 18: Mock up interface for analysts.

6.3.3 Analyst View - Initial Ideas

The mock-up in Fig. 18 would be for analysts only. A time slider could be used to explore the predictions in time. The analyst could take snapshots and export data for presentation to decision makers. The uncertainty in these plots are encoded by; color, multiple runs overlaid (opacity), number of people with time (natural frequencies), and size and position.

These visualization mock-ups went through various iterations, it was felt that decision

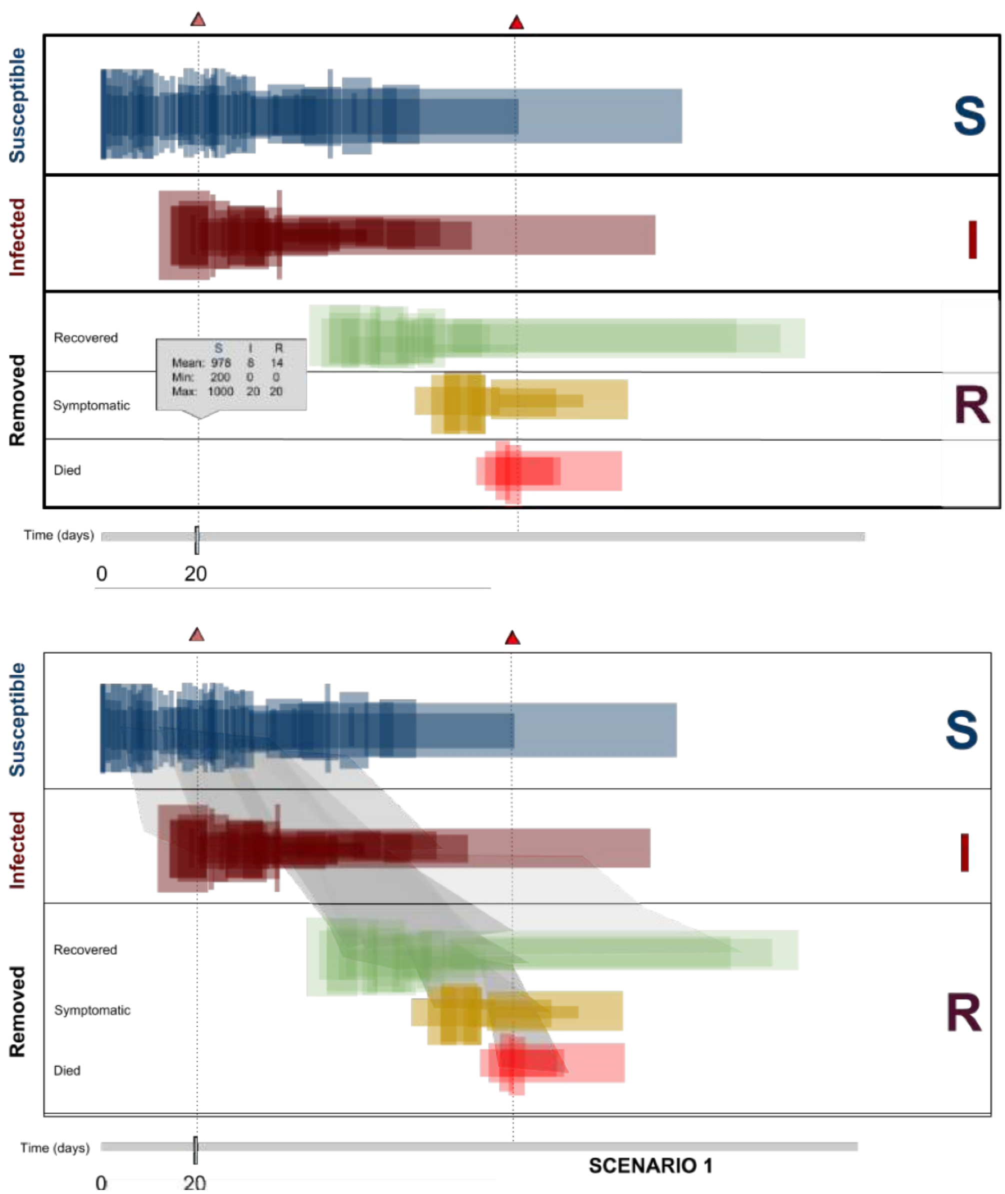


Figure 19

makers did not really require information on how many people are susceptible, and more interested in how many are infected; if 100,000's of people are not infected this is not a number we tend to care about.

Visualizations were refined (Fig. 19) by carefully reviewing comparative studies of error rates and completion times for different visualization types and different tasks, using a lighter color value to represent uncertain regions - or error bars for judging durations and temporal bounds, and gradient plots - using fading color or transparency - for judging probability values, see Ref. [17].

6.3.4 Decision-maker View

A visualization strategy for a decision maker combines the ideas of natural frequencies and scenario building. Temporal information is encoded via customizable milestones which the decision maker can easily assess based on two (or more) intervention strategies, Fig. 20.

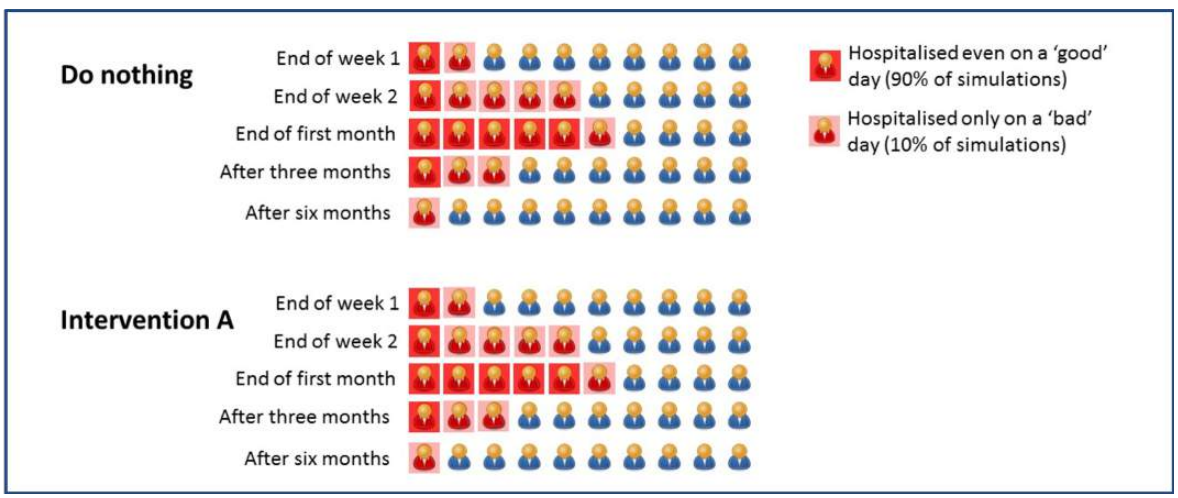


Figure 20: Red indicates infection, blue icons are not infected. Shading of icons demonstartes confidence in that prediction.

However, in an outbreak the number of people who are infected tends to be small in comparison to the total population. In such a case Fig. 20 would be dominated by blue glyphs. This leads us to Fig. 21, where visualizations are per 1,000 people. At a glance the decision maker can interrogate this plot and ask *What does intervention A do?*, the plot would reveal,*at peak incidence, we are likely to have 5 out of 1,000 in hospital at the end of the the first month. If we intervene we are likely only to have have 2 out of 1,000 but there is a 10 percent chance of 5 out of 1,000 on a 'bad' day in hospital. Number of deaths after 6 months is likely to be half if we intervene.*

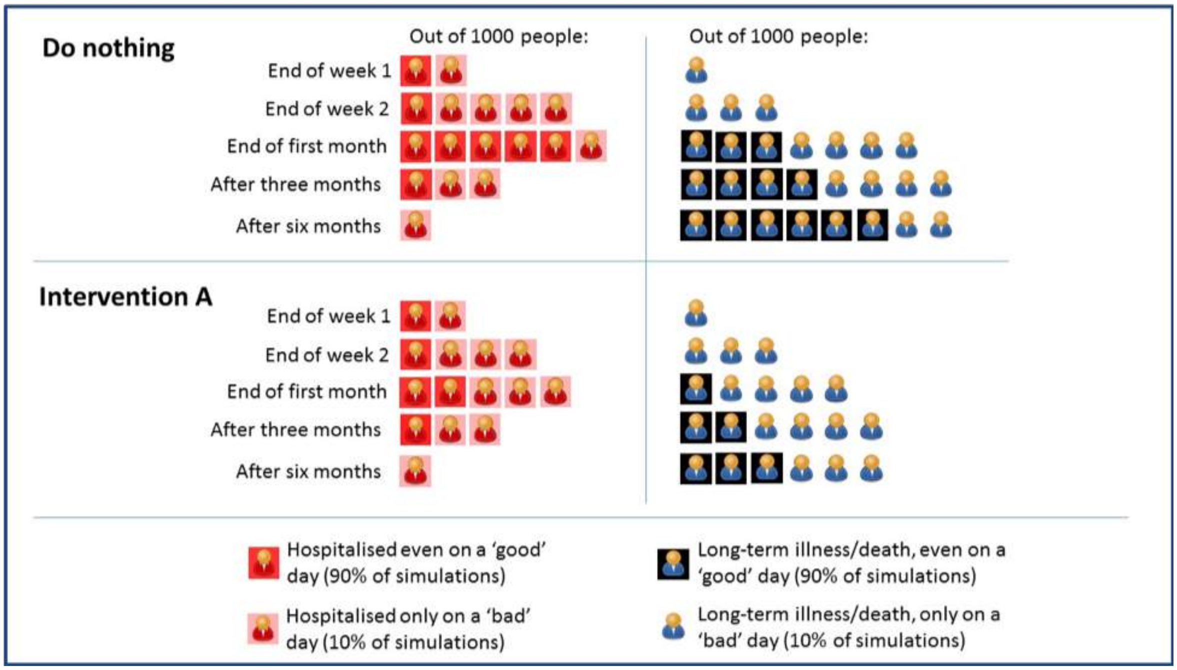


Figure 21: Hospital, illness and death likelihoods with and without intervention. Each glyphs represent 1,000 people.

In Fig. 22 we include running totals as well as snapshots. Snapshots may have more use when planning hospital capacity etc, and cumulative plots help show the wider picture of intervention strategies.

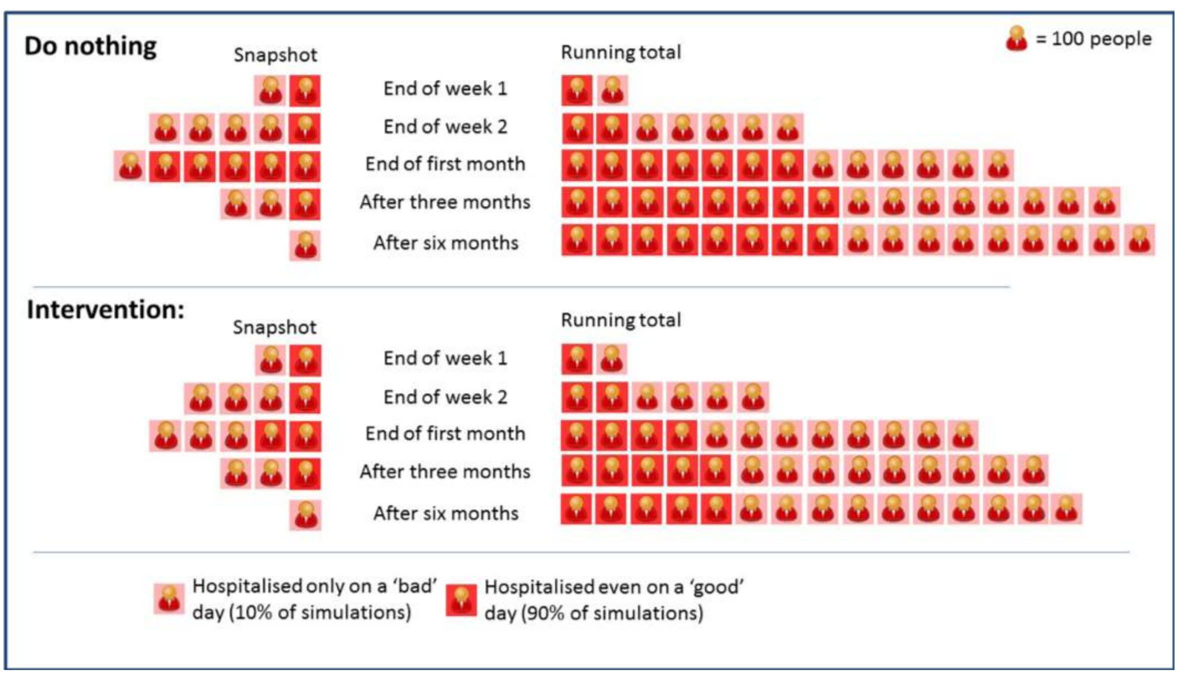


Figure 22: Snapshot and cumulative plots.

A note on Multi-modality.

Caution was advised when dealing with models which may produce multimodal predictions, in which case it does not make sense to average over. For example, if you have 50 percent of your simulations predicting zero people will die and the other 50 percent predicting 100,000's of people die, taking the average is not a helpful approach. So need to have some sensible way of representing trends. Fig. 23 shows SIR-curves which shows the density of color encoding the two modes of model output - the darker the color the more agreement.

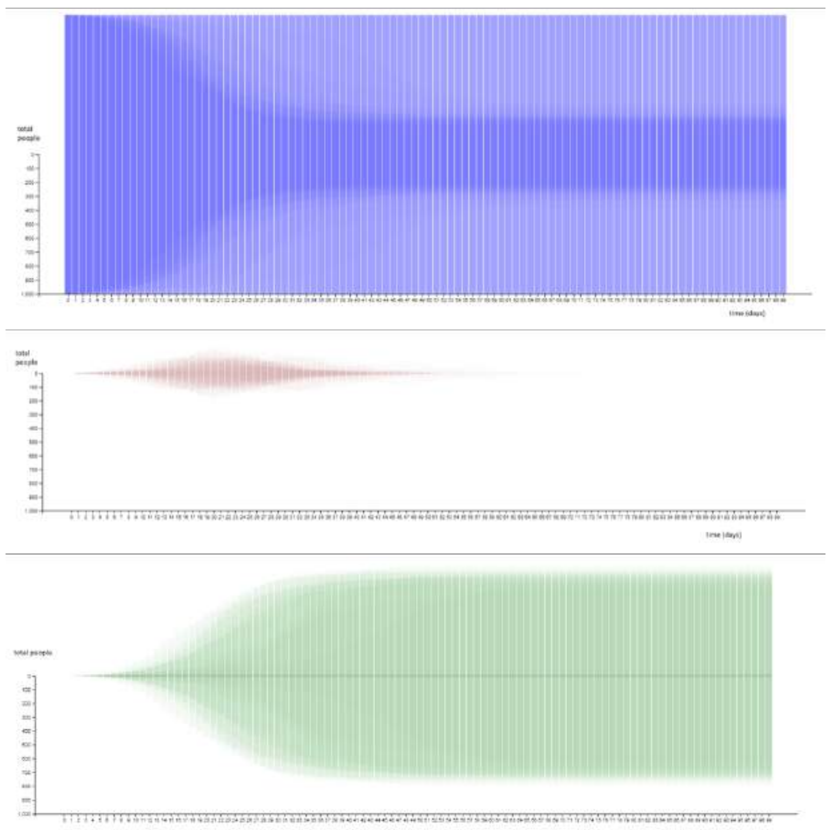


Figure 23: (Top) Susceptible (Middle) Infected (Bottom) Removed

Conclusions

The work on this challenge set out by exploring the questions a decision maker is likely to ask. We also set out the pitfalls to avoid when considering any communication approach. We considered how the building blocks of plots could be used in novel ways, and generated novel visualization outputs that adhere to most of the requirements for decision makers. In addition, we also generated novel, more complex plots which display the fullest extent of information for analysts to draw from.

There is no one "best" way to summaries data and encode uncertainty, this will depend heavily on the simulation results (e.g. multi-modality). We suggest that human supervision will be required (but can provide a decision tree). It is probably a good idea to "ground-truth" the top-level visualization approaches on some actual decision-makers. It is felt that civil service fast trackers provide a good source of expert knowledge to test these out on.

7 Challenge 3: Spatio-temporal Uncertainty Visualisation

Introduction to Theme: Often the output of modeling systems is not a single prediction but a mathematical function that describes a range of outcomes and the chance of particular outcomes occurring; a probability distribution. This is particularly true when considering problems involving geographical space. It may be possible to predict that an event may occur but the exact location is impossible to determine. This situation is often complicated further as we may not be able to work directly with the distribution and only be able to take samples from it.

For example consider an imaginary, highly complex, computer model of the processes of storm cloud creation, movement, and electrostatic discharge. Given the correct starting conditions (a day in British summer) the model would be able to simulate a storm cloud from creation through to the points in time and space where lightning strikes occur. However, such a process is driven by many chance events and each time the model is run, different places are struck at different times. This does not mean that such a model is valueless in telling us where lightning strikes occur, just that it is more difficult to directly understand. In such instances it is normal practice to run the model many times and obtain results that show the variability of such a storm model. In this context each run is called a realization. Visualization of a set of realizations produced in such a manner is tricky and the subject of this work group.

The actual problem set for this group is based around a complex plume and disease

Figure 24: (Left) Contours of the plume from the plume model.(Right) Four different realizations of the location and number of people falling ill from the particles from the plume in the left-hand image.

model combination. The plume model projects how particles of a disease-causing organism spread when it has been released into the air (Fig. 24 (left)). The disease model simulates the response of a human body when such a disease-causing organism is breathed in. Does the person become ill or not? This model, in concept, is run for each individual exposed to the disease. As in the storm model, the actual people that fall ill is subject to a range of complex factors and so each time that the model is run a slightly different result is produced (Fig. 24 (right)).

For the purposes of this workshop, so that our effort is concentrated on the visualization of such model output, we have imposed some limitations:

1. We are only able to take the results of a series of runs of these models and not look at the detail of the mathematical probability
2. The range of conditions either model considers have been restricted. For example:
   (a) The plume model only uses a subset of all possible weather conditions
   (b) The disease model is not set to mimic any known disease (but if it helps the disease behaves a bit like Legionella or Q fever, a similar bacterial disease that infects the lung when breathed in)
3. The population exposed is static (ie the people do not move and are exposed at home)

Expectation of output: We are primarily looking for innovative **methods to visualize** model outputs and systems such as those presented here. As such we are looking for concepts demonstrated in sufficient detail that they may be turned into usable tools to communicate the results to decision makers and the wider public.

Prototype computer code would be beneficial, though hand drawn illustrations with any appropriate supplementary material would be acceptable.

Material format: The plume model has been run 250 times and has produced 250 separate realizations of possible courses that the disease particles could have taken [7]. The disease model has been run to produce 4 realizations for all people exposed for each of these plumes.

There are two GeoJSON files for each plume realization. The first, ending in "_grid_contour.json" is a file containing polygons with the contours of the plume. The second contains point data. Ignoring the release location point, each point represents the centre of a 100 x 100m square and has the following data attached:

[7] https://github.com/mattbktn/uncertainty-visualisation-disease/tree/master/Group3_SpatialUncertainty

1. mean: the mean number of disease particles in the air of that square
2. variance: the variance on the number of particles within the square
3. population: the number people living in the square
4. infected_DM_0 to infected_DM_3: the number of people who were found to be infected in that square by the disease model (DM) on each of its four runs.

The primary material is presented as a zip file containing GeoJSON files. GeoJSON is a standard, non-proprietary, format for geographical data. It may be opened in a wide range of computer mapping systems (GIS) and many computer languages have libraries that are capable of reading such files. More information may be found at http://geojson.org/.

To assist in your initial viewing of the data, within the leaflet folder you will find an html file that is capable of producing a quick look at these files using online mapping technology. Open the file in your browser and then use the browse... button to open a file extracted from the zip.

7.1 Group Members

Sacha Darwin	Riskaware	sacha.darwin@riskaware.co.uk
Thomas Finnie	Public Health England	tomas.finnie@phe.gov.uk
Sam Grainger	University of Leeds	s.grainger@leeds.ac.uk
Nick Holliman	Newcastle University	Nick.Holliman@newcastle.ac.uk
Timos Kipouros	University of Cambridge	tk291@eng.cam.ac.uk
Duncan Lee	University of Glasgow	Duncan.Lee@glasgow.ac.uk
Graham McNeil	University of Oxford	graham.mcneill@oii.ox.ac.uk
Noel Nelson	Pirbright / Met Office	noel.nelson@pirbright.ac.uk
Dan Silk	Dstl	dssilk@dstl.gov.uk

The presentation given by the group at the Study Group can be found here:

https://youtu.be/tlu1bC3lVvo

7.2 Approaches and Progress

The group explored three approaches

1. A PowerBI visualization of choices between impact and certainty for epidemic simulation spatial outputs,

2. visualization and analysis using interactive parallel coordinates and change propagation, and

3. a tile map approach.

7.2.1 A PowerBI Visualisation of Choices Between Impact and Certainty for Epidemic Simulation Spatial Outputs.

Online interactive version published using PowerBI is here: https://bit.ly/2lM265X

PowerBI solutions do not require any coding and with some training a commander or aide may be able to sketch their own visualizations, if for example they were used to producing powerpoint presentations. The PowerBI approach opens up the range of potential people who could produce visualizations of outputs. It would always be wise to check any such visualizations before operational use. Other similar commercial tools include Tableau, Spotfire and Qlic

Using two tree maps to select using those areas that match impact and uncertainty criteria co-selected by the two tree maps. The impact is measured as the mean number of infections per 100 x 100m^2 cell over four run outputs shown in deciles. The uncertainty is measured by the CV - shown in deciles. The CV is a standardized measure of dispersion of

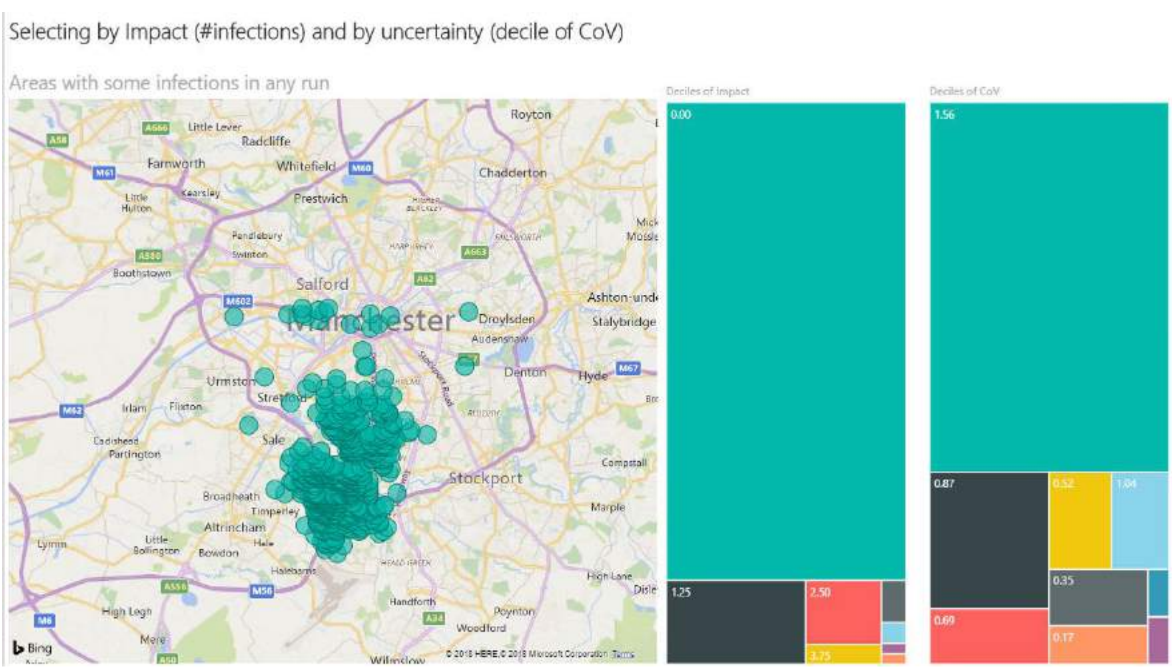


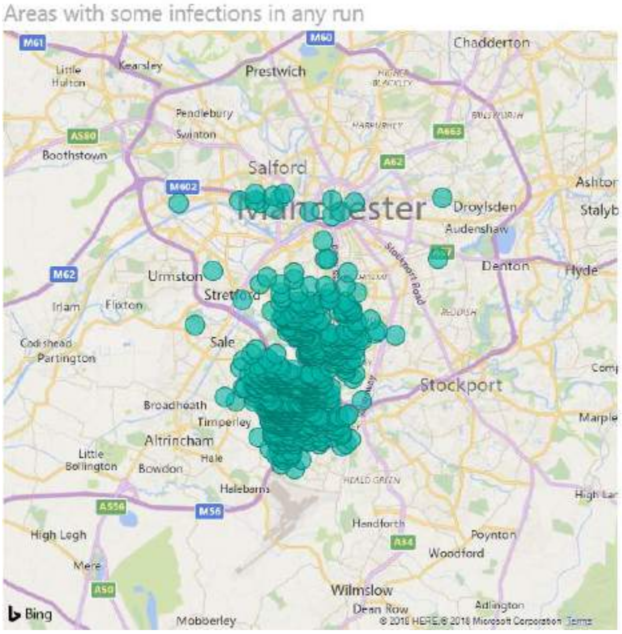
Figure 25: Selecting by impact (μ) and by uncertianty (decile of CV)

a probability distribution or frequency distribution, and defined as the ratio of the standard deviation to the mean (σ/μ). This is shown in Fig. 25

We also plot in Fig. 26 again the impact measured as mean infections per 100 x 100m^2 cell over four run outputs - shown in quartiles and an uncertainty estimate measured by the Index of Dispersion (IoD) (σ^2/μ) - shown in quartiles. The IoD, like the CV is a normalized measure of dispersion of a probability distribution, it is used to quantify whether a set of observed occurrences are clustered compared to a standard statistical model.

Selecting by Impact (quartile of infections) and by Uncertainty (quartile of IoD)

Areas with some infections in any run



Uncertainty

Impact (Mean deaths) ▲	0.00	0.71	1.42	2.13
0.00	0.44	0.77	1.56	2.31
3.13	0.30	0.86	1.73	
6.25	0.27	1.08		
9.38	0.42	1.39		

Figure 26: Sketch in Power BI of selection of affected areas by two variables in an impact / risk matrix.

Informed to some extent by this process for making decisions in conditions of high uncertainty in Ref. [9] and by classical Strength, Weakness, Opportunities, Threats (SWOT) analysis charts, we chose a four way split of certainty and impact to highlight options in the spatial simulation outputs.

Selecting by binary split of the measures of impact and certainty with and without ranked list (Figs. 27 and 28 respectively). The uncertainty here is quantified by the IoD (σ^2/μ), which is similar to the CV (σ/μ).

As this visual requires both geographic and statistical outputs it would be worth trying to visualize it also in Tableau, as it provides alternative options for geographic mapping and probability distribution visualization.

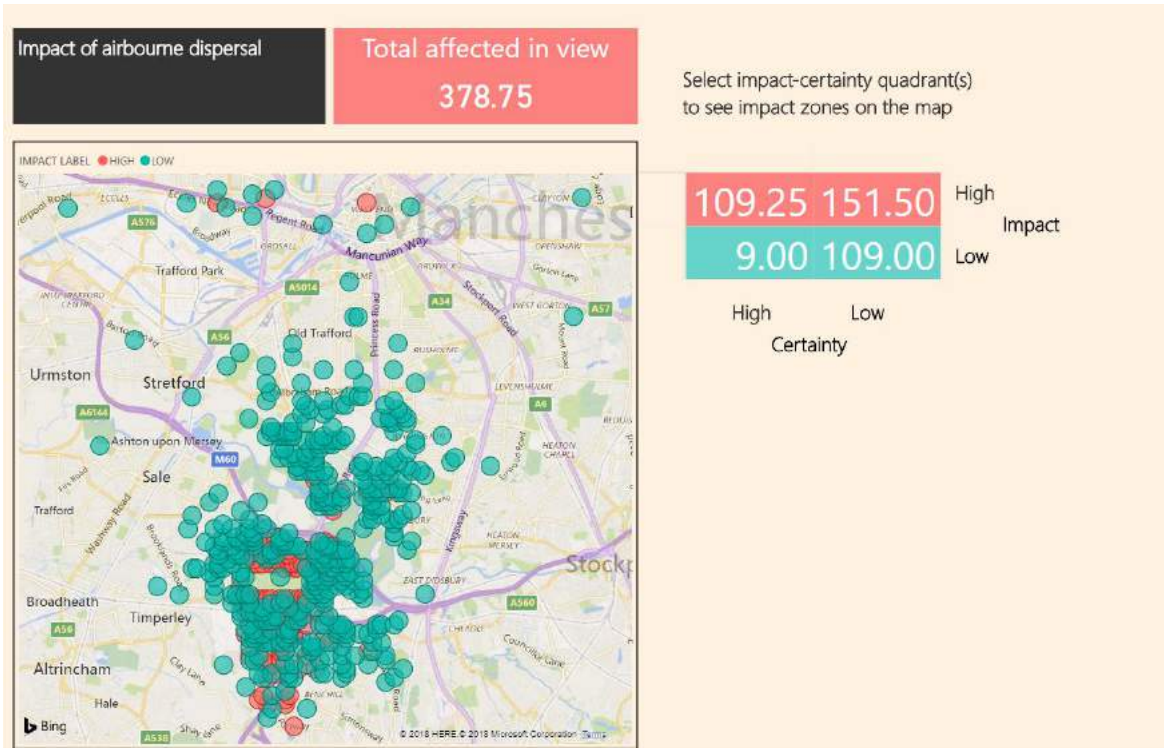


Figure 27: Online interactive version published using PowerBI is here: https://bit.ly/2lM265X

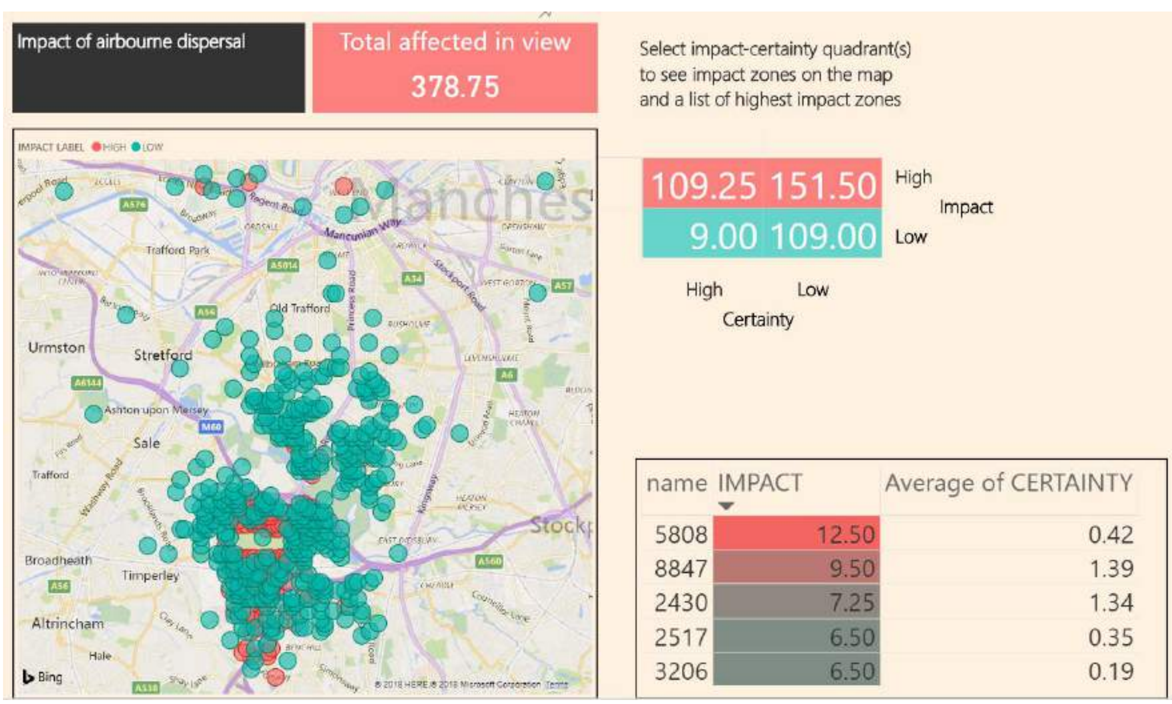


Figure 28: Online interactive version published using PowerBI is here: https://bit.ly/2lM265X

As is common when visualizing high dimensional data, it can be convenient to summaries complex information in a very fast visual cue. Fig. 29 shows quick visualizations (in R) of all 250 simulation outputs supplied, arranged by estimated wind speed / direction and impact. As one might expect in a city, low wind speed often leads to the highest impact on the population.

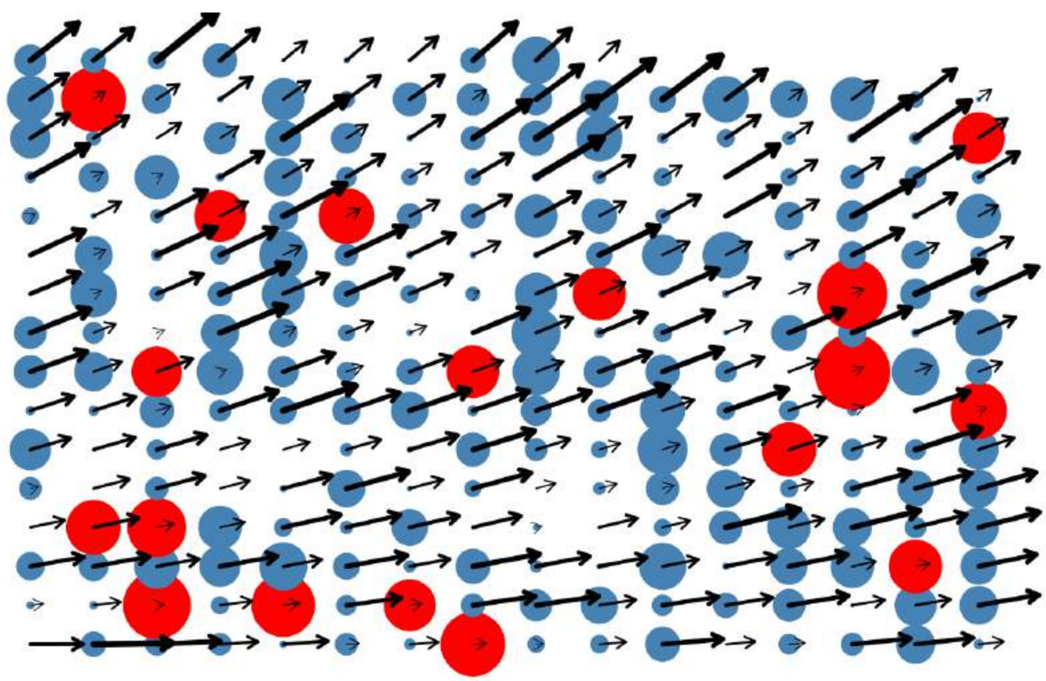


Figure 29: KEY: Arrows: Wind speed + direction. Discs: High certainty impact. Wind direction changes from W to SW across and up the plot. Red is highest quartile of impact. Visualisation provided post workshop by Prof. Nick Holliman, Digital Institute, Newcastle University, UK

7.2.2 Visualization and Analysis Using Interactive Parallel Coordinated and Change Propagation

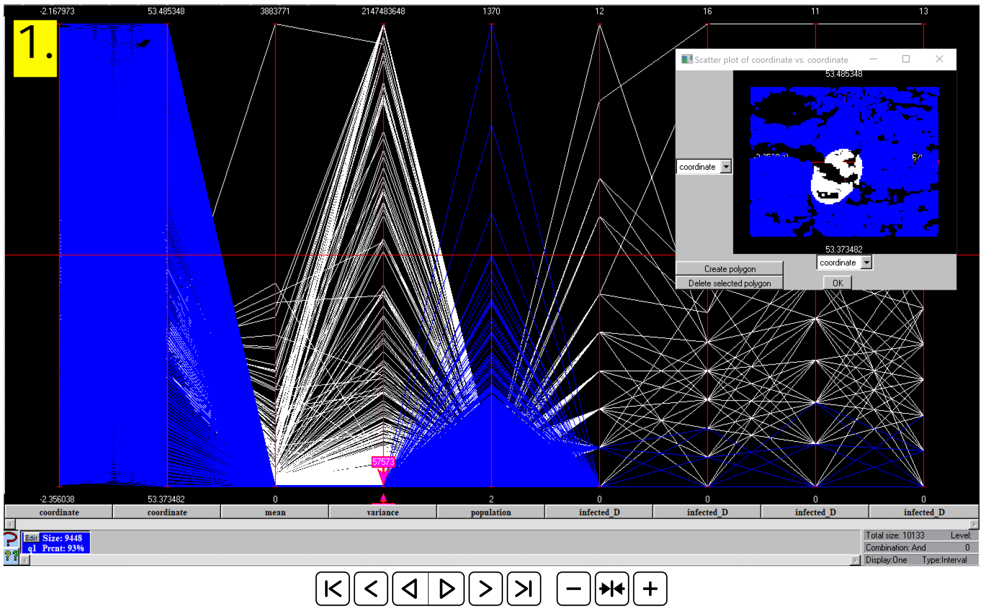


Figure 30: Use control to scroll through seven frames (only when viewed in Adobe™ reader)

Images in Fig. 30 can be viewed with Adobe™ Reader, and can be scrolled using the control bar underneath.

- Image 1: Representation of data set 000 (plume 0) in Parallel Coordinates and selection for low variance (blue). In the scatter plot there are the x-y coordinates
- Image 2: Selection for high values of all models DM0-3 and the corresponding x-y coordinates (location in green) for plume 0.
- Image 3: Representation of data set 001 (plume 1) in Parallel Coordinates and selection for low variance (blue). In the scatter plot there are the x-y coordinates.
- Image 4: Representation of data sets 000-004 (plumes 0-4) in Parallel Coordinates and selection for low variance (blue). In the scatter plot there are the x-y coordinates. (10 dimensions in total including the number of plume).
- Image 5: Considering 5 plumes (0-4) and selecting simultaneously the high values of affected population from all 4 models (DM0-3). This means the confidence level is high.
- Image 6: Considering 5 plumes (0-4) and selecting the high values of affected population from model 0 only. This selection though is subjective, and hence, the decision maker has the opportunity to define / express what 'high value' means. Or

actually vary interactively this value in the plot and see how that affects the result in terms of affected areas.

- Image 7: Considering 5 plumes (0-4) and overlaying the two selections from Figures 5 and 6. The green area is the one with more confidence and the blue is the additional area due to the prediction of DM0 only. The decision maker can now interact with the data and explore scenarios considering only one model at a time, or combination of 2 and 3, and hence develop the understanding how the uncertainty in predicting the most affected regions varies.

This type of interactive visualization can be enhanced by adding to each x-y coordinate the variance to the prediction of number of infections between the 4 different models and the 250 plumes (1,000 potential values for each x-y coordinate).

Change Propagation Analysis: We can now collect the up to 1000 instances of affected prediction for each of the 100 x 100m^2 geographical location and develop a distribution of the particles. Then use this information, together with any other factors, such as population size, population density, population age, etc, in order to estimate the likelihood of the affection to this map location. We can then build a DSM model to represent these likelihood values for each map location, as well as the impact to the neighborhood map location(s). A sample of such model is shown in Fig. 31 (left).

We can then simulate the propagation of the impact and reveal the implicit (indirect) connections between the map locations. This is shown in Fig. 31 (right).

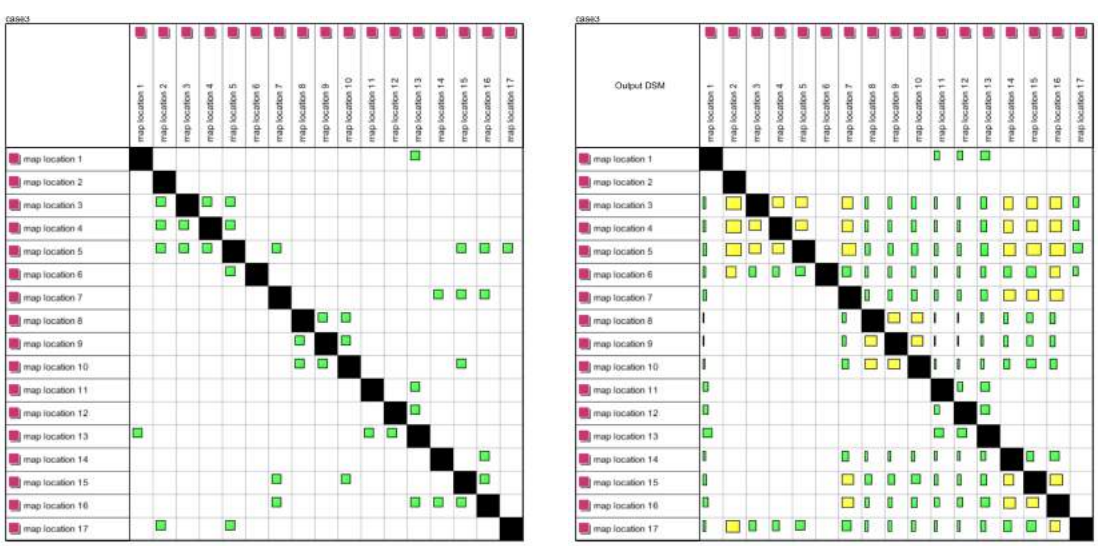


Figure 31: (Left) The representation of direct connectivity between 17 map locations in a DSM model. (Right) The output DSM model after the simulation of the CPM

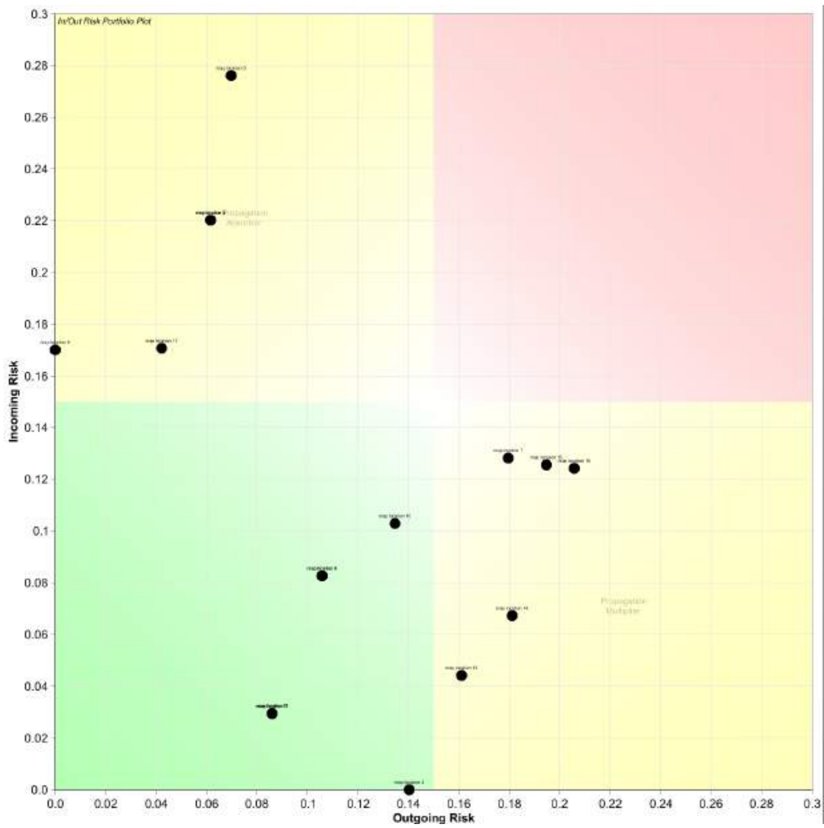


Figure 32: The In / Out Risk portfolio plot.

The square connections now represent risk and the size represents the magnitude. We can see that some indirect map locations are revealed to exhibit high risk. Another way to visualize the results is in an In/Out Risk portfolio plot, as shown in Fig. 32.

We can now identify the risk multipliers, or the most influential map locations, as well as the risk absorbers. Useful references [20, 23, 7]

7.2.3 Tile Map Approach

There is a live example of this approach here https://gjmcn.github.io/london-disease-impact-example/ (use Google Chrome) and a GitHub repository https://github.com/gjmcn/london-disease-impact-example.The example is an unfinished demo.

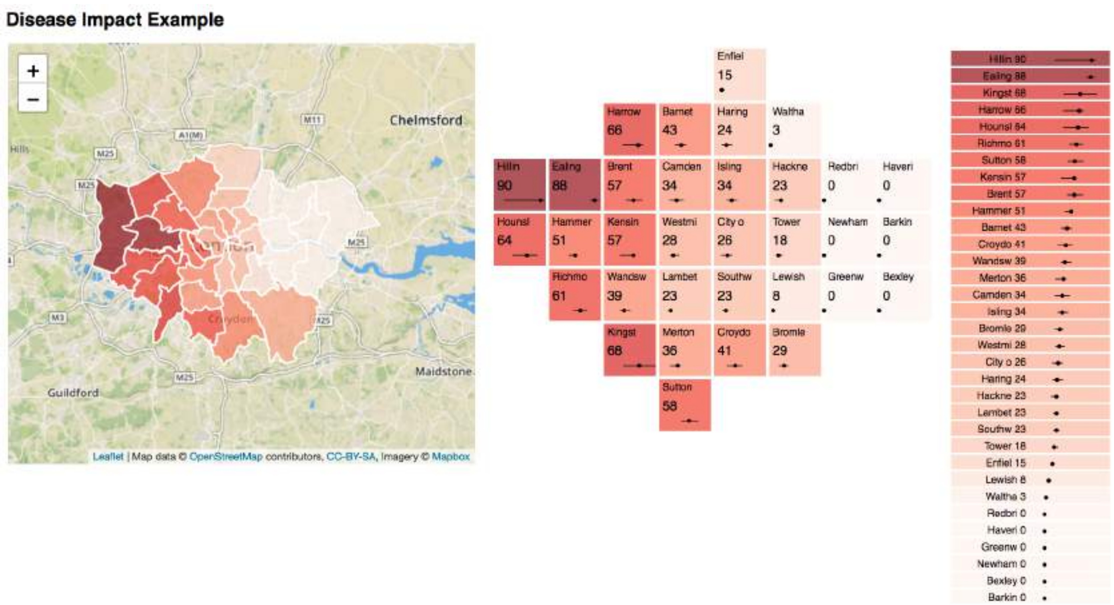


Figure 33: Tile map example showing linked map, tile map, and impact list.

The Tile Map demo uses toy data which is generated each time the example is loaded - use `Ctrl-R` to see different 'scenarios'. The range for the two boroughs with the highest median impact is always tweaked: the highest is given a large range, the second highest a small range. For the example shown in Fig. 33, *Hillingdon* has the high median impact (90) and a large range (indicated by the length of the horizontal line in the tile) and the second highest *Ealing* with 80 has the smallest range.

There are various missing features of this demo, e.g. all three components of the visualization should be linked and have tooltips (currently the map has tooltips; the tiles and ranges are linked on hover). There should be an axis above the list of ranges on the right.

- In cases where there are many regions (e.g. at postcode level), we would not show all the info on the tiles nor all the ranges on the right. When zoom-in, could show the tile details and the ranges for the visible regions.
- Other relevant info could be shown on the tiles — not just related to the model, but

e.g. if there is a hospital in the borough.

- We could consider multiple scenarios simultaneously or compare prediction for do nothing/intervention by showing e.g. multiple linked maps or multiple-linked tile maps or multiple-linked ranges (i.e. only use one component of the visualization).
- We could animate or have time slider for temporal data.
- The range diagrams could show multi-modal data — there would be multiple dots and segments.
- Could show a worst case scenario by displaying the 95 th percentile for each tile, but this may be misleading. Specifically, the decision maker might think the sum of the numbers is the overall 95th percentile; this would be an overestimate since (probably) none of the simulated plumes will be in the worst 5 percent for all boroughs.

8 Challenge 4: Visualizing Uncertainty in the Numbers in 'at risk' Groups due to Multiple Model Predictions

Introduction to Theme: Epidemiological modeling is a complex process requiring expert judgment, models and information. Currently, there are many ways to predict disease spread; however, it is not possible to synthesize these predictions into a single holistic picture for the end user to interpret consequences.

UQ and visualization for epidemiological models presents a number of challenges, which have only been partially addressed in the literature. Of interest to this group is the variety of models available to predict the outputs for a particular disease scenario. If we aim to provide the ability to quantify uncertainty across ensembles of disparate model runs we will need to be able to effectively visualize the uncertainties associated with the selection of individual or multiple models. Where do models agree on low/high uncertainty, where do they disagree, and how can this be presented to aid decision making?

The availability of this ensemble of different, scientifically-plausible models provides the opportunity for more realistic UQ via approximation of the uncertainty in the science itself. Such ensemble UQ has typically involved simulations at different fidelities or resolutions. In disease modeling, we additionally have parallel ensembles of alternative simulators that, a priori, are assumed to be of (roughly) equal fidelity but of varying complexity and underpinned by varying assumptions. This case has been given much less consideration.

The choice of model is a difficult one due to the rapid pace of change in the spread of a disease, lack of data or risk aversion on model selection. In some situations, multiple models may be appropriate and the output should be visualized over a number of different modeling scales (from high level approximations to detailed simulations) and must reflect the associated uncertainties.

We are taking the spread of a 'flu-like' disease as an example. The data from two example models are available to our group:

1. An 'expensive' disease forecast model (hours to run per simulation)
2. A 'cheap' disease forecast model (minutes to run per simulation)

These are stochastic models (each run of the model will yield a different answer) that are affected by the uncertainty that we are looking to visualize (see, e.g. Fig. 34. Each model has been run 1000 times.

Expectation of output: We are primarily looking for innovative ways to visualize the

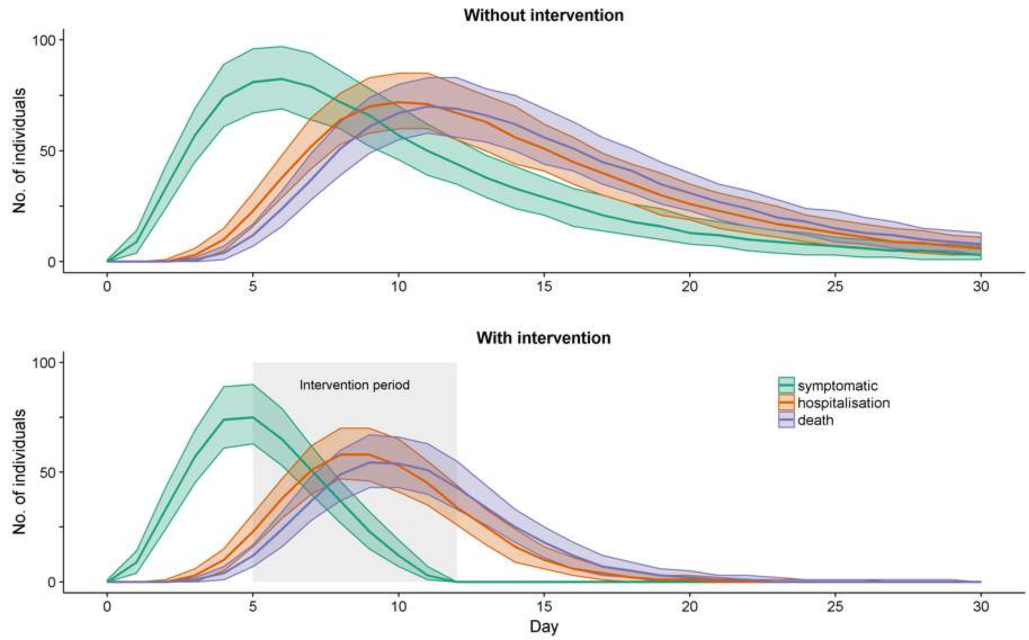


Figure 34: Example output from a single disease forecast model, with a simple representation of uncertainty.

uncertainty resulting from the model selection. For instance, if we have two equally valid model outputs which differ in their estimates, how do we visualize this? We are also interested in whether the uncertainty in both models can be combined in an image to give a better illustration of uncertainty than that provided by just one model. Is it possible to visualize the data to best inform decisions such as to re-run model(s) and/or make interventions?

Material format: Two comma separated variable files are provided that contain 1000 realizations of the models described above. The full list of variables is described below[8][9].

Variable name	Description
Run	A unique identifier indexing which simulation run the data belongs to (numbered 1 to 1000).
Time	A discrete time-point ranging from 0 to 100 (inclusive)
Population *	Total population
Susceptible *	Susceptible population size
Infected *	Infected population size
Removed_total *	Removed population size (Total, regardless of intervention)

[8] In the expensive model, these variables will have four columns representing spatial areas.

[9] https://github.com/mattbktn/uncertainty-visualisation-disease/tree/master/Group4_ModelUncertainty

8.1 Group Members

Helen Adams	Risk Aware	helen.adams@riskaware.co.uk
Xin Chen	Cranfield University	xin.chen@cranfield.ac.uk
Alex Forrester	University of Southampton	Alexander.Forrester@soton.ac.uk
Samuel Jackson	University of Southampton	s.e.jackson@soton.ac.uk
Jonathan C. Roberts	Bangor University	j.c.roberts@bangor.ac.uk
Pranay Seshadri	University of Cambridge	ps583@cam.ac.uk
Hannah Williams	Public Health England	Hannah.Williams@phe.gov.uk

The presentation given by the group at the Study Group can be found here:

https://youtu.be/w3ExDTunp0s

8.2 Approaches and Progress

The Problem - Numbers of People in Risk Groups

The first question addressed by the group was in understanding where the different model runs come from? Why would we have multiple model predictions? These could come from many sources; for example different *runs* of the model, several different types of *models*, different *coding approaches* from different research groups.

Next the group discussed the scale of the different "multiples", as to talk about visualization we first need to understand the scale of the multiples; for example for things which could have only two or three different values we might consider using symbols to distinctify them, for things which might have potentially thousands of multiples we might have to consider schemes which could have thousands of variations; color scales for example. Table. 1 shows quantities and possible visualization schemes.

The Questions That The End-Users Want

As with other groups, this group discussed how best to "pose" questions the visualizations

Table 1: Sizes of Multiples

Data	Quantity	How Could We Visualize? (e.g.)
Models	$\sim$ 5 models	
Patches	1 to 10 k	Different maps
Runs	1000 +	Different colored lines on chart
Interventions	5 to 10	Shapes

are designed to address. Modelers tend to talk / think about peak, average, gradient, envelope, mean of infected people, etc. however end-users want more "natural worded" questions, e.g. When to treat? What happens when we add an intervention? When does it become an epidemic? How certain is X? What happens when half-term hits? Where are the risks? How bad will it be? What do I do about it?

Strategies for Showing 'Multiples'

The group discussed different strategies for showing multiples; juxtaposition (showing different objects separately), superposition (overlaying objects in the same space, like we do in Geographic Information Systems (GIS) with layered information) and explicit encoding of relationships, shown in Fig. 35, see Ref. [16]

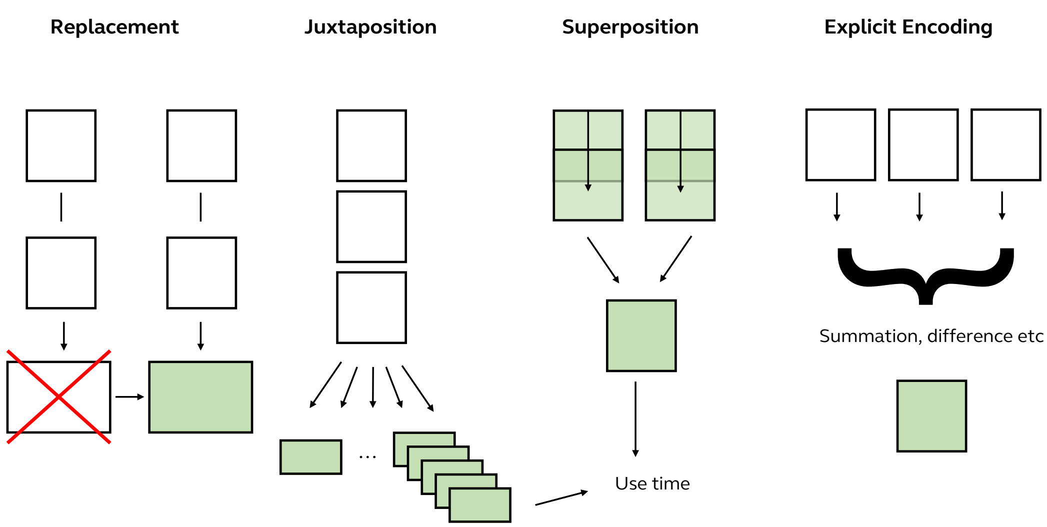
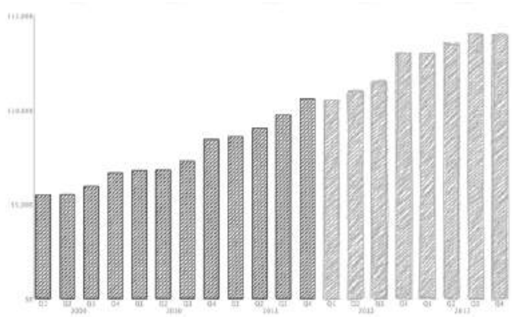


Figure 35: Strategies for visualizing multiple datasets.

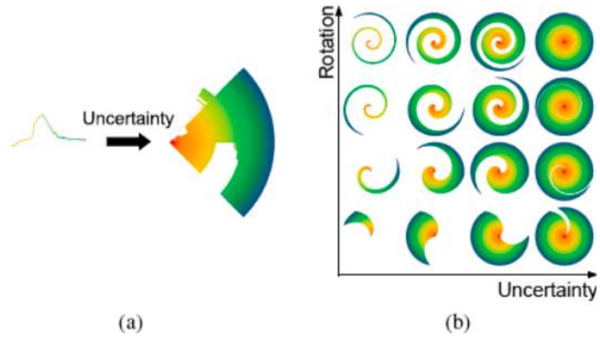
Visual Methods of Encoding Uncertainty (and many variables)

Next the group discussed different ways to visualize the different models and uncertainty therein. Examples of these strategies are shown on the following page.

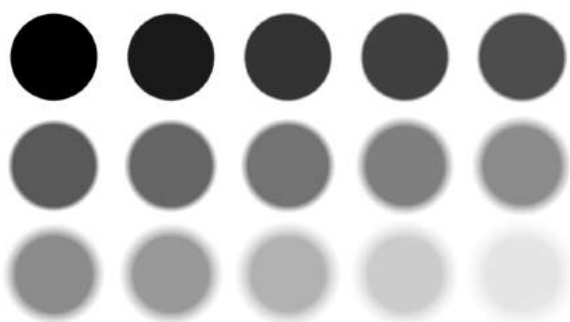
(i) Sketchy Rendering: Sketchy rendering naturally gives the impression that things are uncertain. See Ref. [38]



(ii) Using Glyphs: See Ref. [36]. Image from Ref. [24]

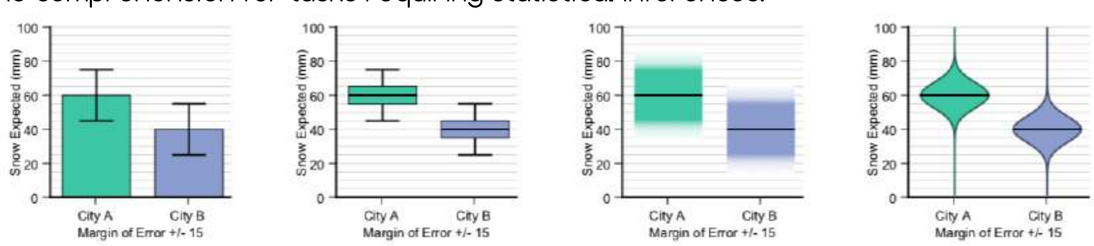


(iii) Transparency, color, size etc.: The more uncertain an estimate is, the more difficult it is to see. You can achieve this effect a number of ways, such as with transparency, color scale, or blurriness, Ref. [13]



(iv) Small multiples and Sparklines: A sparkline is a small intense, simple, word-sized graphic with typographic resolution. Ref. [33]. Example below from Ref. [21]

(v) Error bars: Four encodings for mean and error evaluated in Ref.[8]. Each priorities a different aspect of mean and uncertainty, and results in different patterns of judgment and comprehension for tasks requiring statistical inferences.



(a) **Bar chart** with error bars: the height of the bars encodes the sample mean, and the whiskers encode a 95% t-confidence interval.

(b) **Modified box plot**: The whiskers are the 95% t-confidence interval, the box is a 50% t-confidence interval.

(c) **Gradient plot**: the transparency of the colored region corresponds to the cumulative density function of a t-distribution.

(d) **Violin plot**: the width of the colored region corresponds to the probability density function of a t-distribution.

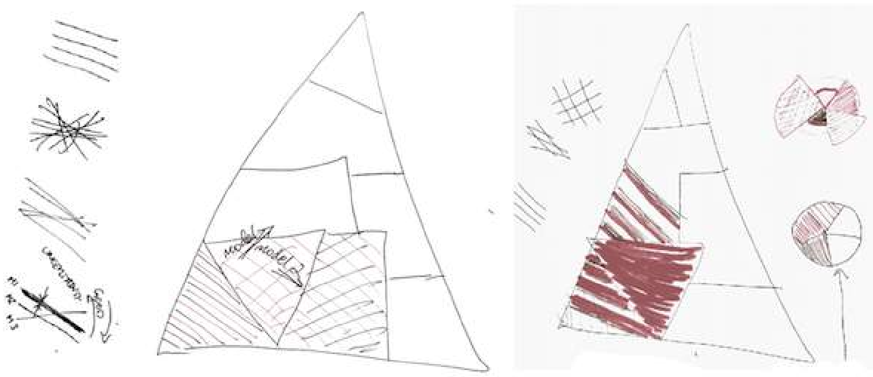


Figure 36: Sketches demonstrating how to encode uncertainty and disease prediction through 'texture'.

Geographical Areas (Patch Models)

The group started to consider how the patches and information we wish them to convey could be visualized, ideally we would like to display the difference in the *models* as well as the difference in *values*. As there are many more variables we may wish to encode into the patches, the group considered whether pop-out information boxes could be used when zooming into geographical regions.

The group started to consider the idea of *texture*; for example if someone is infected can we have one type of line and if they are not infected, we could have another type. Additionally, let us consider model discrepancy, if we have ten models - all of which agree - all the lines could go in the same direction, if there is a significant degree of disagreement, the lines could go in a different direction. Concepts for this are shown in Fig. 36. Useful reference in Ref. [30]

Tesselated View (Hexagons)

The group then considered different shapes on a UK map, e.g., hexagon, such as used by the Office of National Statistics (ONS) (for example Ref. [28]) and Fig. 37. These equally sized hexagons could then be selected and explored in more detail. It is, however, difficult to mix all data and information into these small hexagons, so the group came up with the idea of using colour instead of texture (cross hatching).

If multiple models agree then the hexagon will show the *outcome* (red if there are high number of predicted infections and green if there is a low number of predicted infections). If the models disagree we then switch to a different color scale (light blue if there is mild disagree moving towards dark blue if there is strong disagreement between the models), this is shown in Fig. 38. The user could then zoom in on those troublesome areas for more detailed information. The advantages of this approach are that the view is clean and

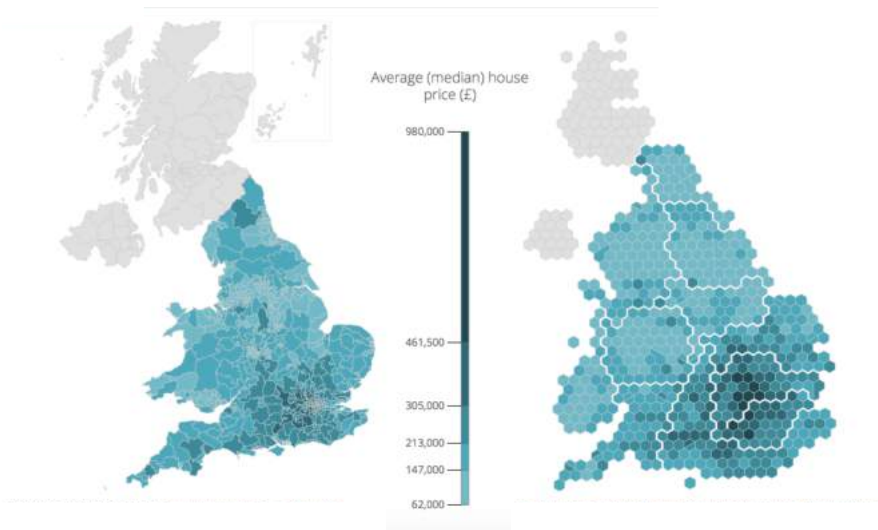


Figure 37: Example of tessellated hexagon view of UK taken from the ONS

compact, it gives a good overview. The disadvantages include the need to parameterize the level of disagreement.

How can we tell if our models are agreeing or not? An approach is shown in Fig. 38. If , for example, we have a thousand runs from a patch model, we would sample from that, take two draws, compare and calculate the distance between them. The difference out of the total population will be our percentage of the population by which those models differed. The process could be bootstrapped thousands of times to get a confidence. In Fig. 39 the blue line is the 95 percent confidence our models have 10 percent difference, a 20 percent difference and so on.

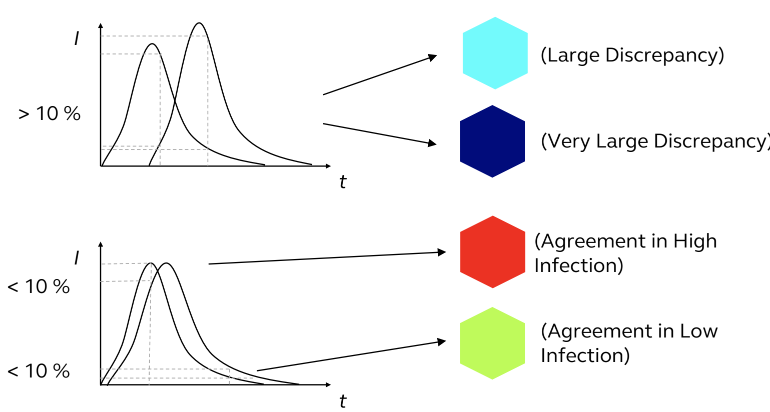


Figure 38: Example of how to assess model disagreement

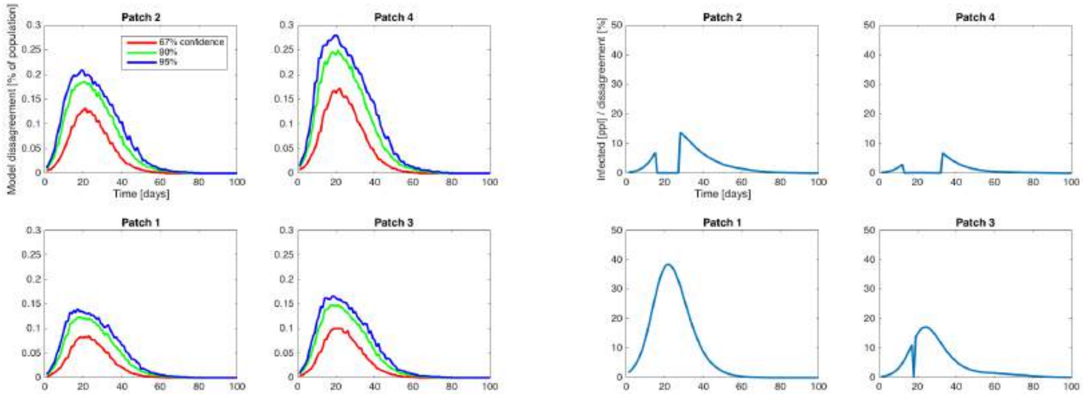


Figure 39: Infection curves and disagreement plots

The ultimate aim is to pick the threshold of where we say our models disagree so we can move our red to green colour scale to our blue (agree / disagreement) colour scale.

The right-hand plot of Fig. 39 shows the infection rate in patches 1 - 4. In patch 1, all the models agree, so we plot out the mean prediction for the models and this would be plotted on the red / green color scheme. But if the disagreement goes above a set limit 10 percent, say, as featured in patches 2, 3 and 4, we plot on the blue color scheme.

Fig. 40 we present three demonstrations of the scheme above. In the left-hand image, we see there are large amounts of differences across the map, and a few areas with bad infection on which the models agree. In the middle image, many of the models agree (except much of the north and around London). In the right-hand image, almost all of the models are in agreement and areas of low and high infection are plotted across the UK.

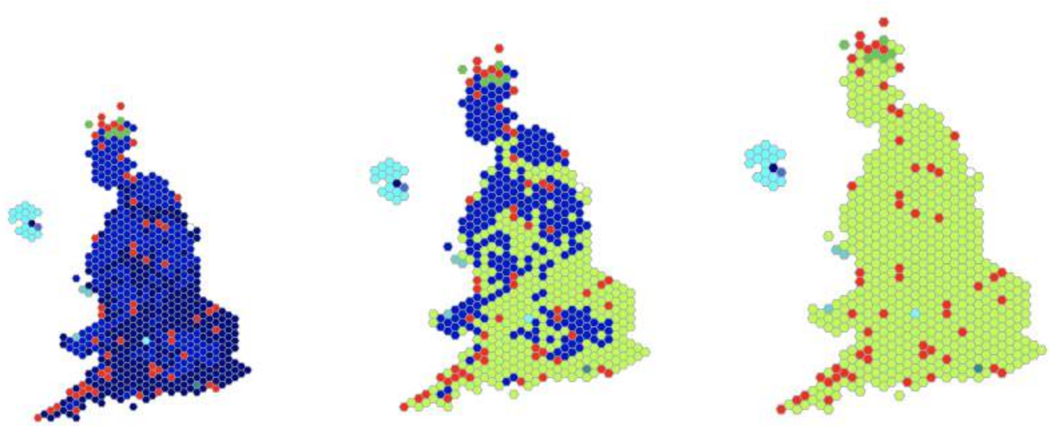


Figure 40: Sample images (using D3.js)

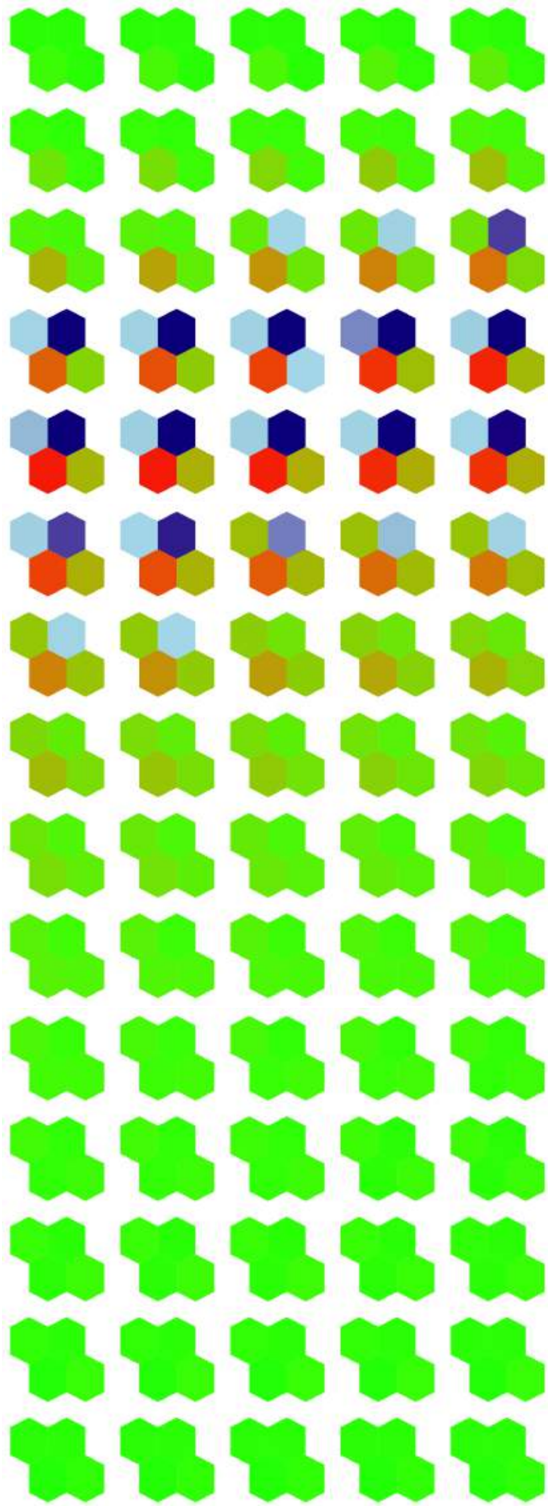


Figure 41: Demonstration of approach for four patches, evolving through time

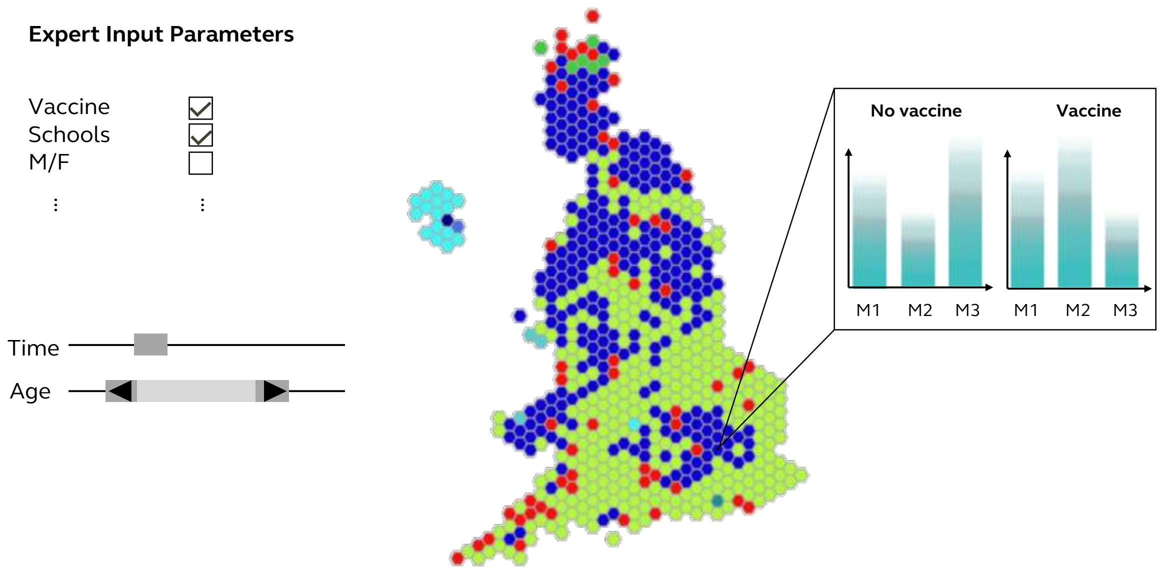


Figure 42: Mock up GUI for above approach

In Fig. 41 we show just four patches (i.e. not the whole UK) time stepping through the model. At $t = 0$, all models for all patches are in agreement with not many infected. As we move to rows 4 - 6, we see varying levels of disagreement on the infected number, such that the upper patches begin to turn blue; we know then there is a modeling error. In this scenario, the decision maker will not then make a decision based on how many are infected, but rather zoom in on the blue hexagons and see what is happening with the underlying models.

We show in Fig. 42 a sketch of an interface which has a map, and customizable drop downs of models and interventions, and a time slider. If there is a blue box (indicating model disagreement) you could interrogate it by clicking on it to show bar charts which tell you of your uncertainty in models 1, 2, 3 etc. This view could also highlight your modeling assumptions so that you can start making decisions about which assumptions are appropriate, eliminate poorly behaving models, simulate more data etc.

9 List of Acronyms

BSVE Bio-Surveillance Ecosystem

COBR Cabinet Office Briefing Room

CPM Change Propagation Method

CV Cofficient of Variation

DFM Disease Forecast Model

DFT Disease Forecast Toolbox

DSM Disease Structure Matrix

Dstl Defence Science Technology Laboratory

DTRA Defense Threat Reduction Agency

EUQO Ensemble Uncertainty Quantification and Optimization

FOCUS Framing, Order, Complexity (and Uncertainty), Search

GIS Geographic Information Systems

GUI Graphical User Interface

IoD Index of Dispersion

KTN Knowledge Transfer Network

ONS Office of National Statistics

PDF Probability Density Function

PHE Public Health England

SIR Susceptible, Infected, Removed

SWOT Strength, Weakness, Opportunities, Threats

UK United Kingdom

UQ Uncertainty Quantification

VS Visualisation Suite

WAIFW Who Acquires Infection from Who?

10 References

[1] Bank of England: Inflation Report (2013) Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2013/august-2013.pdf?la=en&hash=FE0AAD6C68F5E20AB94ACD521CE25FDA0CB9ED0E

[2] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/6825 modelling-blackett-review.pdf Council for Science and Technology (2017)

[3] Bonneau GP. et al. (2014) *Overview and State-of-the-Art of Uncertainty Visualization.* In: Hansen C., Chen M., Johnson C., Kaufman A., Hagen H. (eds) *Scientific Visualization. Mathematics and Visualization.* Springer, London

[4] Adrian W. Bowman. Graphics for uncertainty. Available http://www.stats.gla.ac.uk/%7Eadrian/papers/graphics-for-uncertainty- paper.pdf, 2015

[5] https://www.liverpool.ac.uk/media/livacuk/instituteofriskanduncertainty/events/LiverpoolVBowman24041 2017

[6] Ken Brodlie et al. A Review of Uncertainty in Data Visualisation, University of Leeds. Available https://pdfs.semanticscholar.org/27bf/388d5dcfa6281d1ae67306b1e9f43372acdc.pdf, 2010

[7] Clarkson, P. J., Simon, C. S. and Eckert, C. M. (2004). Predicting Change Propagation in Complex Design. ASME Journal of Mechanical Design, 126(5), 788-797.

[8] Correll, M. Gleicher, M. *Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error* (2014) IEEE Transactions on Visualization and Computer Graphics

[9] https://hbr.org/1997/11/strategy-under-uncertainty

[10] Breast Screening—Performance of Mammography: Cancer Research UK (available at http://info. cancerresearchuk.org/cancerstats/types/breast/screening/mammography/).

[11] J. G. Dolan, S. Iadarola, BMC Med. Inform. Decis. Mak. 8, 14 (2008).

[12] EPSRC, Dstl. Defining a Research Strategy for Uncertainty. Available https://www.epsrc.ac.uk/newsevents/pubs/defining-a-research-strategy-for-uncertainty, 2016

[13] http://flowingdata.com/2018/01/08/visualizing-the-uncertainty-in-data/

[14] Gigerenzer, G., & Hoffrage, U. (1995). *How to improve Bayesian reasoning without instruction: frequency formats.* Psychological Review, 102(4), 684.

[15] G.Gigerenzer, W.Gaissmaier, E.Kurz-Milcke, L.M.Schwartz, S. Woloshin, Psychol. Sci. Public Interest 8, 53 (2007).

[16] Gleicher M. et al (2011)/ *Visual Comparison for Information Visualisation*. Information Visualisation - Special Issue on State of the Field and New Research Directions, (Vol. 10, Issue 4) https://doi.org/10.1177/1473871611416549

[17] T. Gschwandtnei, M. Bögl, P. Federico and S. Miksch, "Visual Encodings of Temporal Uncertainty: A Comparative User Study," in IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1, pp. 539-548, Jan. 31 2016. doi: 10.1109/TVCG.2015.2467752

[18] Hahn, U., & Harris, A. J. (2014). *What does it mean to be biased: Motivated reasoning and rationality*. In Psychology of Learning and Motivation (Vol. 61, pp. 41-102). Academic Press.

[19] Hogarth, R. M., & Einhorn, H. J. (1992). /it Order Effects in Belief Updating: The Belief-adjustment Model. Cognitive Psychology, 24(1), 1-55.

[20] Inselberg, A. (2009) Parallel Coordinates: Visual Multidimensional Geometry and Its Applications. Springer-Verlag, New York.

[21] http://www.jamiebicknell.com/20-Jun-2014/Sparklines-via-PHP

[22] Visualising Uncertainty in Geographic Data using Hierarchical Spatial Data Structures. J. Kardos, A. Moore, and G. Benwell http://www.geocomputation.org/2003/Abstracts/Kardos_Abs.pdf

[23] Kipouros, T., Inselberg, A., Parks, G. T. and Savill, A. M. (2013). Parallel Coordinates in computational engineering design, in 9th AIAA Multi-Disciplinary Design Optimization Specialist Conference, Boston, Massachusetts. AIAA-2013-1750.

[24] P. Leube, D. Weiskopf, W. Nowak and M. Hlawatsch, "Flow Radar Glyphs - Static Visualization of Unsteady Flow with Uncertainty," in IEEE Transactions on Visualization & Computer Graphics, vol. 17, no. , pp. 1949-1958, (2011). doi:10.1109/TVCG.2011.203

[25] MacEachren, A. M., C. A. Brewer and L. W. Pickle (1998). "Visualizing georeferenced data: representing reliability of health statistics." Environment and Planning A 30: p1547-1561.

[26] B. J. McNeil, S. G. Pauker, H. C. Sox Jr., A. Tversky, N. Engl. J. Med. 306, 1259 (1982).

[27] E. Kurz-Milcke, G. Gigerenzer, L. Martignon, Ann. N.Y. Acad. Sci. 1128, 18 (2008).

[28] https://www.ons.gov.uk/

[29] K. Potter et al., "Ensemble-Vis: A Framework for the Statistical Visualization of Ensemble Data," 2009 IEEE International Conference on Data Mining Workshops, Miami, FL, 2009, pp. 233-240. doi: 10.1109/ICDMW.2009.55. URL: https://ieeexplore.ieee.org/document/5360497/

[30] Stuart J. Rose, Pak Chung Wong, "DriftWeed: a visual metaphor for interactive analysis of multivariate data", Proc. SPIE 3960, Visual Data Exploration and Analysis VII, (28 February 2000); doi: 10.1117/12.378887; https://doi.org/10.1117/12.378887

[31] I. B. Schwartz, and L. B. Shaw, "Rewiring for adaptation". Physics **3**, 17 (2010)

[32] Visualising Uncertainty About the Future. David Spiegelhalter *et al. Science* **333**, 1393 (2011); DOI: 10.1126/science.1191181

[33] Beautiful Evidence, Edward R. Tufte. Graphics Press, (2006)

[34] Tversky, A., & Kahneman, D. (1974). *Judgment under Uncertainty: Heuristics and Biases.* Science, 185(4157), 1124-1131.

[35] V. H. M. Visschers, R. M. Meertens, W. W. F. Passchier, N. N. K. de Vries, Risk Anal. 29, 267 (2009).

[36] Craig, M. Wittenbrink, Alex, T. Pang, and Suresh K. Lodha.

[37] Glyphs for visualizing uncertainty in vector fields. IEEE Transactions on Visualization and Computer Graphics, 2:266–279, (1996)

[38] Wood, J., Isenberg, P., Isenberg, T., Dykes, J., Boukhelifa, N. & Slingsby, A. (2012). Sketchy rendering for information visualization. IEEE Transactions on Visualization and Computer Graphics, 18(12), pp. 2749-2758. doi: 10.1109/TVCG.2012.262

[39] Zekâi, S. (2017) *Innovative Trend Methodologies in Science and Engineering*, edn 1, Springer International Publishing