Scottish COVID-19 Response Consortium

Monday, 27th July 2020

Scottish COVID-19 Response Consortium

Monday, 27th July 2020

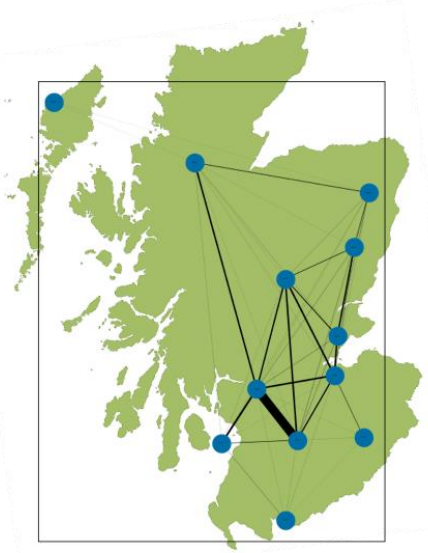
Network Simulation

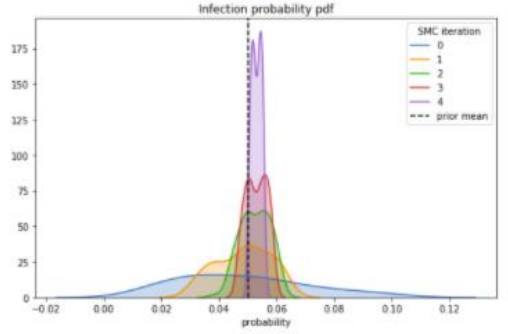
https://github.com/ScottishCovidResponse/simple_network_sim

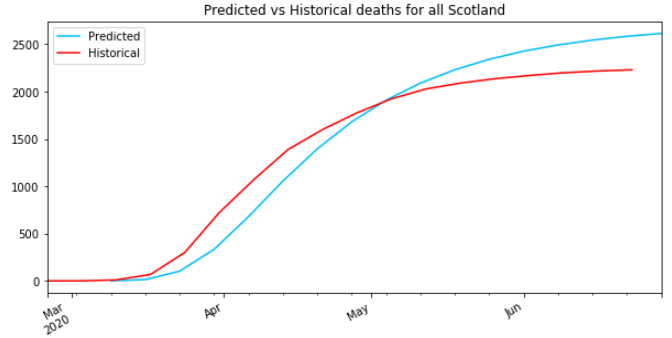
Jess Enright (Glasgow)
Bob Turner (Sheffield)
Rafael de Almeida (Man Group)
Chris Hughes (Man Group)
Fedor Gorokhovik (Man Group)

Will Pettersen (Glasgow)
Anthony O'Hare (Stirling)
Leighton Pritchard (Strathclyde)
Alex Konovalov (St Andrews)
Jason Dykes (City)
Jo Wood (City)

Language	Python
Spatial scale	Data dependent – currently Health Board or IZ
Model type	Spatially-embedded network of age-structured populations, deterministic or stochastic, compartmental model
Features	- Movement flows from data or simulated - Age-class mixing from input matrix - Local and long-distance spread - Dynamic movement and contact damping
Mode	Forward simulation, daily timesteps Basic inference implemented – ABC-SMC







Key questions

- Control related to age class or movements
- Importance of network flows, impact of changes to flows

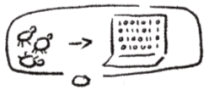
Network Simulation

https://github.com/ScottishCovidResponse/simple_network_sim

Significant achievements over the RAMP period

Software engineering	Many major improvements in design, automated testing, documentation, reproducibility, and more!
Modelling functionality	- Choice to focus on network-of-populations - Generalisation to arbitrary regions given data - Generalised compartmental model - Time-varying activity damping - Stochastic model operation - Generating realistic behaviour - Enormously improved software choices!
Data Interface	- Move to inputs from files instead of hard-coded or simulated - Using real data as input - Data Pipeline API integration
Inference and Sensitivity	- Linear and non-linear sensitivity analysis - Implementation of ABC-SMC for basic inference

Contact Tracing

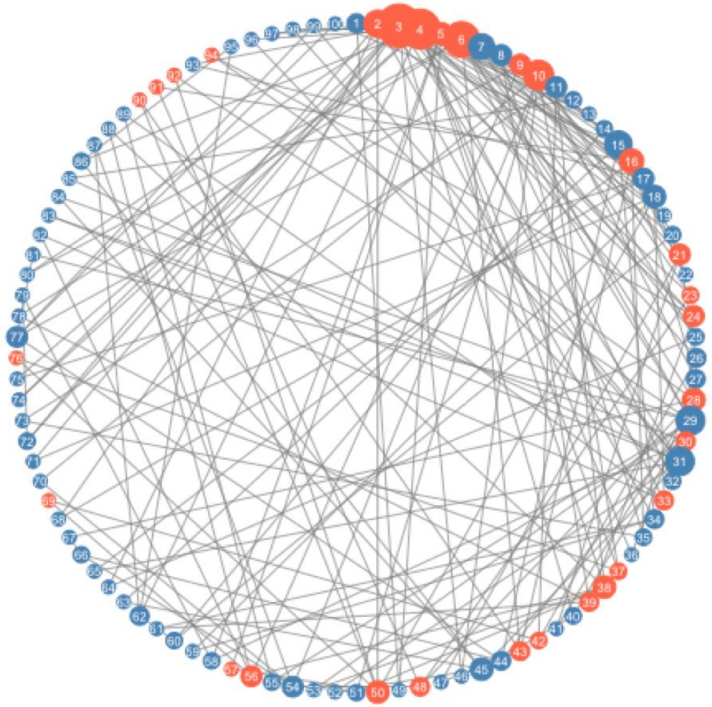


Sibylle Mohr, Louise Matthews
(EPIC / Glasgow)

Sam Brett, Vino Mano (Man Group),
John Nonweiler (UKAEA), **Ed Townsend**

https://github.com/ScottishCovidResponse/Contact-Tracing-Model

Language	Java
Spatial scale	No explicit spatial scale but a form of 'location' is retained
Model type	Individual-based stochastic network model • Age-structured individuals • Weighted contacts
Key Feature	Flexible contact tracing scenarios
Mode	Forward simulation, user-defined timesteps (quarter of day)

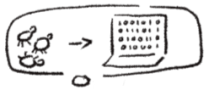


Key questions

What are the optimal contact tracing protocols?

- Benefits of blanket vs tailored actions and advice for alertees?
- How can we reduce the burden of contact tracing whilst maintaining high levels of efficacy
  - Targeted alerts / rapid testing / release of contacts from isolation
- How much more effective does the tracing become if tracing of secondary contacts is allowed?
- What is the impact of basing alerting on a positive test rather than symptoms?

Contact Tracing: Key features


Sibylle Mohr, Louise Matthews
(EPIC / Glasgow)

Sam Brett, **Vino Mano** (Man Group),
John Nonweiler (UKAEA), Ed Townsend

Explore contact-tracing policies

Isolation measures

- Blanket 14-day isolation vs. dependent on the time of contact (0-14 days)

Alerts / testing

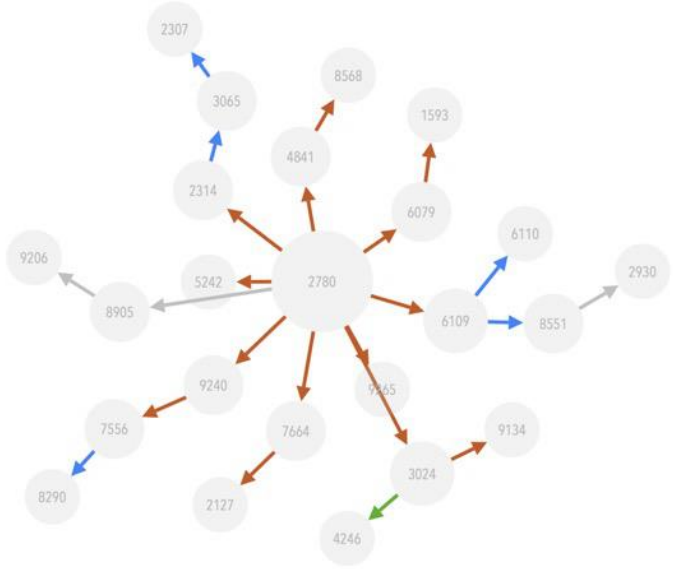
- targeting of alerts at individuals most likely to be infected
- testing of individuals to release their contacts from isolation

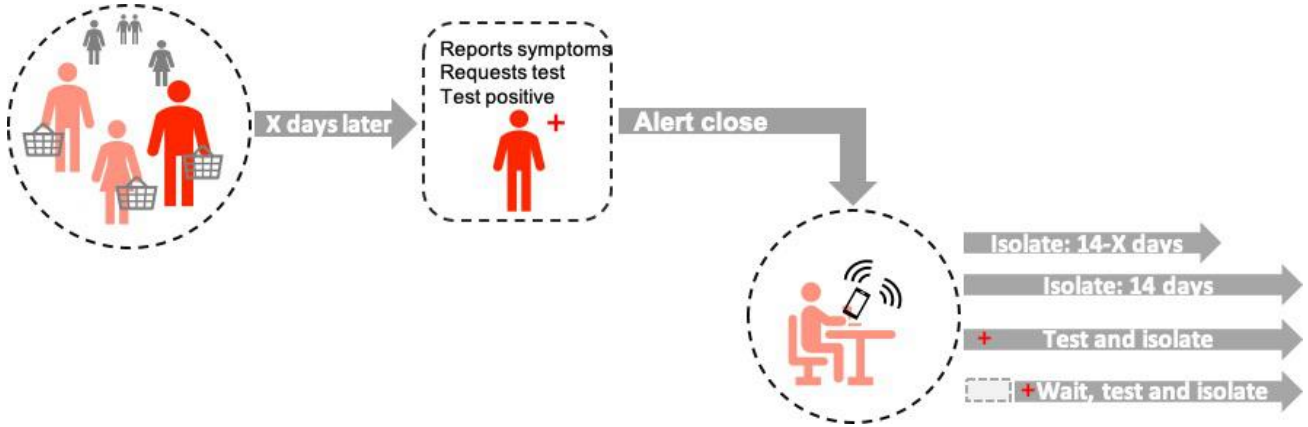
Different 'alerting' scenarios

- Alerting based on symptomatics
- Alerting based on positive test

Contact tracing

- Direct contacts only
- Also secondary contacts (contacts of direct contacts)





Contact Tracing Team

https://github.com/ScottishCovidResponse/Contact-Tracing-Model

RSE core team	Sam Brett Vino Mano (Man Group) John Nonweiler (UKAEA) Edward Townsend
Inference / model assessment	Rikiya Takahashi Fedor Gorokhovik (Man Group)
Visualisation team	Cagatay Turkay (Warwick) Dan Archambault (Swansea) Kai Xu (Middlesex)
Contact tracing options	Louise Matthews (UoG) Jamie Prentice (UoG)

Major RAMP Achievements

Software development

- Code clean-up; codebase tested / integrated / optimised; re-factoring; added software tests / unit tests...
- Documentation, verification, integration into the SCRC environment...

Model functionality

- Multiple, sophisticated contact-tracing scenarios / secondary contacts / global lockdown policies / testing scenarios / test efficacy
- Model runs on various contact data, together with list of age-structured individuals; Random infection / targeted seeding
- Complex statistics output (R calculations; traced / isolated / infected...)

Inference

- Sensitivity analysis for model parameters / policy parameters
- Comparison between CTM model and ODE model

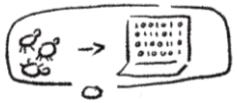
Visualisation

- Multiple static and dynamic visualisation of different output data
  - Network data / compartmental outputs / infection maps

And so much more!!

- Data API for Java is taking shape; new documentation / inferences / visualisation...

Spatial Simulation



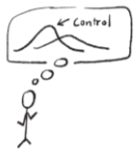
Claire Harris (EPIC / BioSS)

James Cook (UKAEA)

https://github.com/ScottishCovidResponse/Simulation.jl

Language	Julia
Spatial scale	1km^2 grid, Scotland / UK
Model type	Explicitly spatial, stochastic, compartmental model
Features	- Airborne and environmental spread of virus - Local spread vs. long-distance commuting - Impact of climate / environment on transmission - Demographic structure, population density - Inference and visualisation frameworks
Mode	Forward simulation, daily timesteps

RSE team: Alex Robson, Bella Wu, Eric Perim Martins and Sean Lovett (Invenia)



Epidemiology: Glenn Marion and Iain McKendrick (BioSS), Louise Matthews and Richard Reeve (Glasgow)

Inference: Ben Swallow and Simon Babayan (Glasgow)

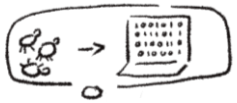
Visualisation: Aidan Slingsby (London City)

Data: Sonia Mitchell and Ciaran McMonagle (Glasgow), Stephen Catterall (BioSS) and Jeremy Walton (Met Office)

Key questions

- What effect does long term pollution have on susceptibility / severity of COVID-19?
- What is the effect of weather/environment on the virus?
- What are the relative contributions of direct/indirect transmission?

Spatial Simulation



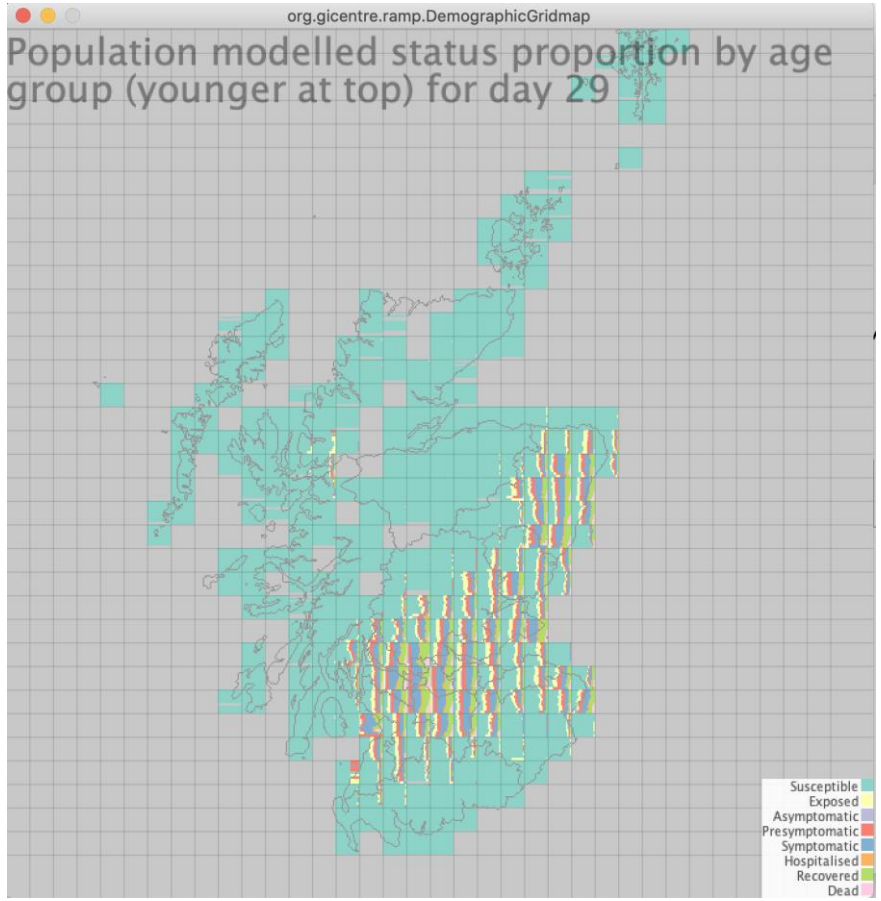
Claire Harris (EPIC / BioSS)

James Cook (UKAEA)

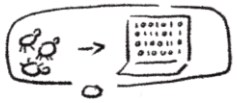
https://github.com/ScottishCovidResponse/Simulation.jl

Major developments

- Codebase:
  - Integrated, tested, documented, optimised
  - Python wrapper for data pipeline
  - Fully deterministic version of code
- Epidemiology:
  - Mostly everything!
- Data:
  - Data flowing from the pipeline
  - Demographics/Epi params/Weather/*Pollution*
- Inference:
  - ABC-SMC interface in R that can recover initial parameters
- Visualisation:
  - Fully interactive Java visualisation, with multiple modes, scaling and spatiotemporal dynamics



Individual-based Simulation



Paul Bessell (Edinburgh),

John Nonweiler (UKAEA),
Blair Archibald (Glasgow),
Bob Taylor

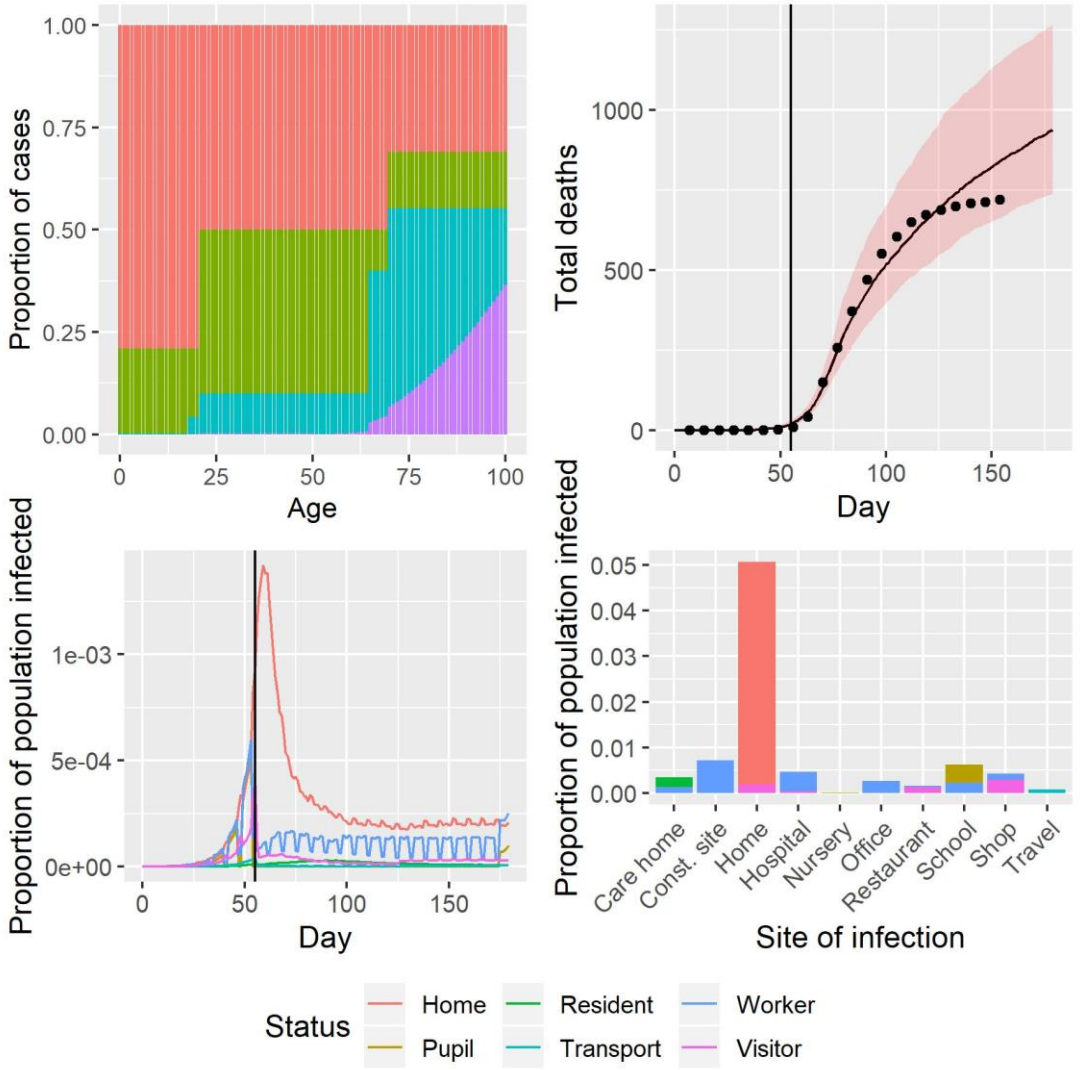
https://github.com/ScottishCovidResponse/Covid_Simulation_Model

Language	Java
Spatial scale	Defined area (eg Lothian HB, Orkney)
Model type	Individual-based stochastic simulation
Features	- Individual hourly movements and mixing - Implicit network structure - Full age and household structure - Mixing and transmission in distinct community compartments - Health boards and care homes
Mode	Forward simulation, hour time steps
Fitting	Relies on parameters and visual comparison to data
Data	Most key parameters are now evidenced, but would be good to complete the list

Model purposes

- Detailed simulation of different communities
- Completely flexible reopening/reclosing of businesses and compartments of society
- Comparison between different communities / areas
- Scenarios for opening schools
- Flexible addition of new societal compartments (could add theatres, cinemas etc)
- Shielding populations
- Strategies for roll-out of vaccination or new tests
- Role of superspreaders
- Informing other models with network structure

Individual-based Simulation



Major RAMP Achievements

Software development

- Simple terribly written code has been transformed into a very efficient and complete epidemiological model
- Complete testing environment with the model
- 1 run of a 200 day simulation of 910,000 population on DiRaC is ~22 mins
- Full software documentation

Model functionality

- Fully age structured disease progression
- Age structured household structure and implicit contact network
- Fully age-structured non-Covid hospital utilisation
- Disease transmission during travel
- Build-your-own lockdown release scenarios

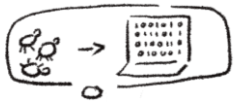
Fitting

- (For Lothian at least) Model broadly fits the mortality, hospitalisation and R values)
- Will run it through Thibaud's model

Further steps

- Write up the epidemiological documentation
- Get better contact parameter data

ABC-smc



Thibaud Porphyre (Edinburgh)

Peter Fox (UKAEA)

https://github.com/ScottishCovidResponse/Covid19_EERAModel

Language	C++
Spatial scale	Flexible
Model type	Age-structured compartmental stochastic model
Features	- Different transmission behaviour between community and HCW - Impact of bed capacity on hospitalisation rate
Mode	Inference – ABC-smc Forward simulation

Model purposes

- Inference of key epidemiological parameters
- Impact of scales (national vs regional vs health board)
- Prediction of community level infection
- Quantification of number of primary infection
- Optimization lifting movement restrictions strategies

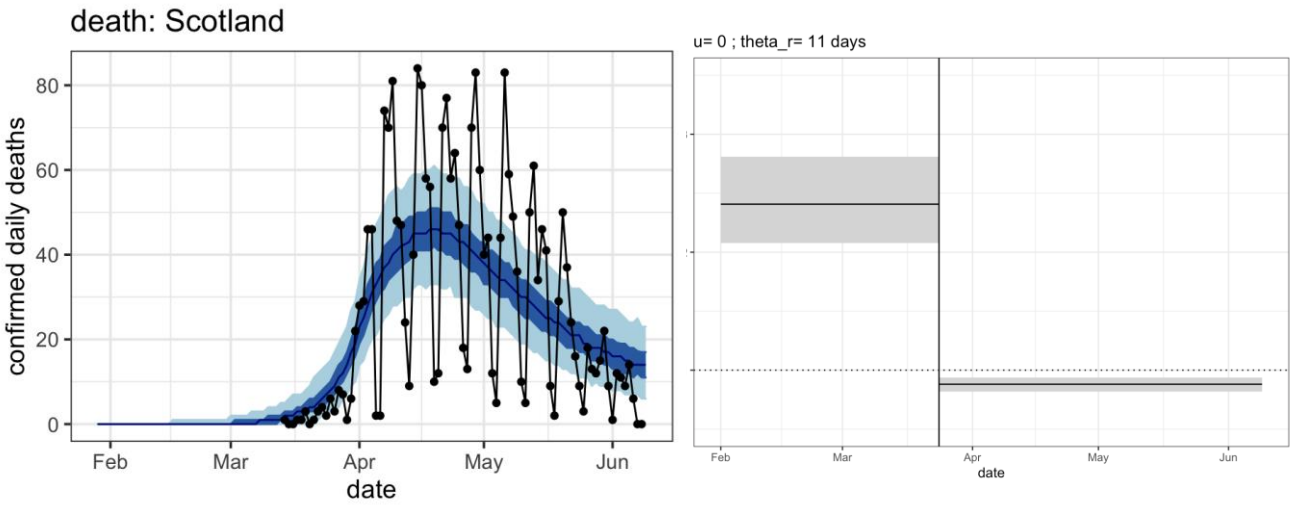
RSE team: Sanket Gadgil, Kristian Zarębski, Qingfeng Xia (UKAEA), Andy Bennett

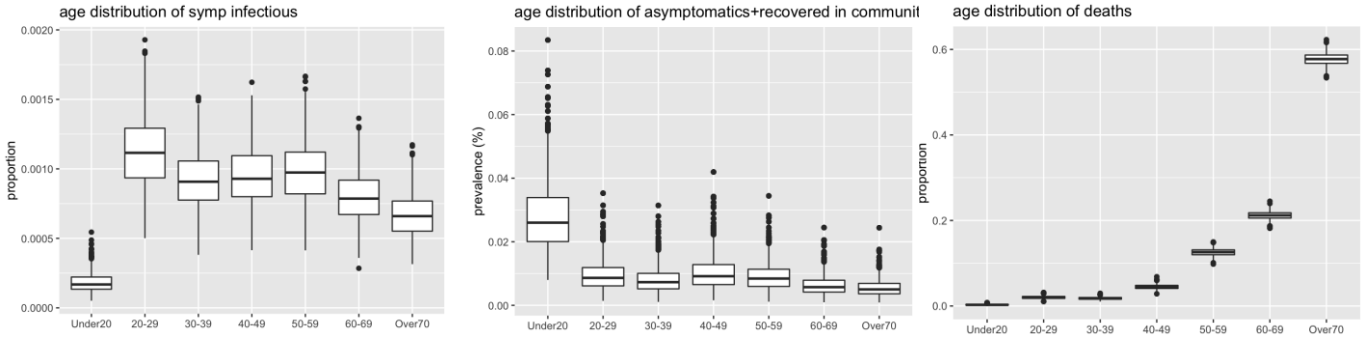
Epidemiology: Mark Bronsvoort, Paul Bessell (Edinburgh), Louise Matthews (Glasgow)

Uncertainty quantification: Ben Swallow (Glasgow), Ian Vernon (Durham), and the rest of #UQ team

Data: Stella Mazeri (Edinburgh)

ABC-smc





Major RAMP Achievements

Software development

- Simple terribly written code has been transformed into a very efficient and complete epidemiological model
- Complete testing environment with the model
- Full software documentation
- Full epidemiological documentation
- Sensitivity analysis

Model functionality

- Both inference and forward prediction implemented
- Fully age structured disease progression
- Role of HCW in the case reporting

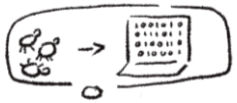
Fitting

- Model broadly fits the mortality, hospitalisation and R values for Scotland
- Inference framework validated

Further steps

- Uncertainty quantification
- Coding and assessing lockdown strategies
- Fit on all health board in UK

BEEPmbp



Chris Pooley (BioSS / EPIC)
Glenn Marion (BioSS/ EPIC)

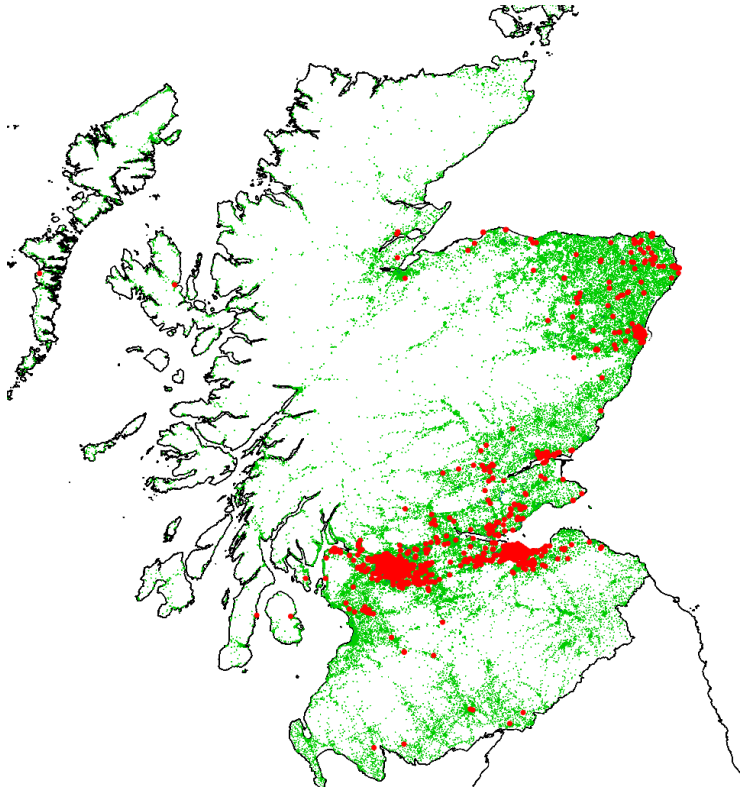
https://github.com/ScottishCovidResponse/CoronaPMCMC

Ian Hinder (Manchester)
Robin Williams (Bristol)

Data: Richard Bailey (Roslin)
Stephen Caterall (BioSS)

Model: Glenn Marion (BioSS)
Andrea Doeschl-Wilson (Edinburgh)

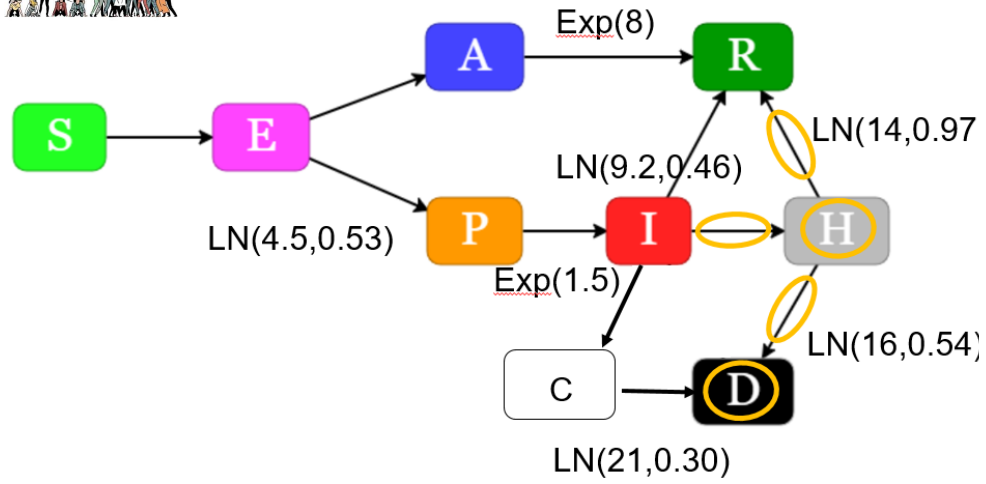
Language	C++
Spatial scale	Middle Layer Super Output Areas (MSOA) or output areas (OA), Scotland
Model type	Individual-based, spatial, stochastic compartmental model
Features	- Individual risk factors, e.g. age, gender - Area risk factors, e.g. population density - Time variation in transmission - Calculation of model evidence (ME) - Optimized parallel implementation
Mode	Inference using multi-temperature MCMC with model based proposals

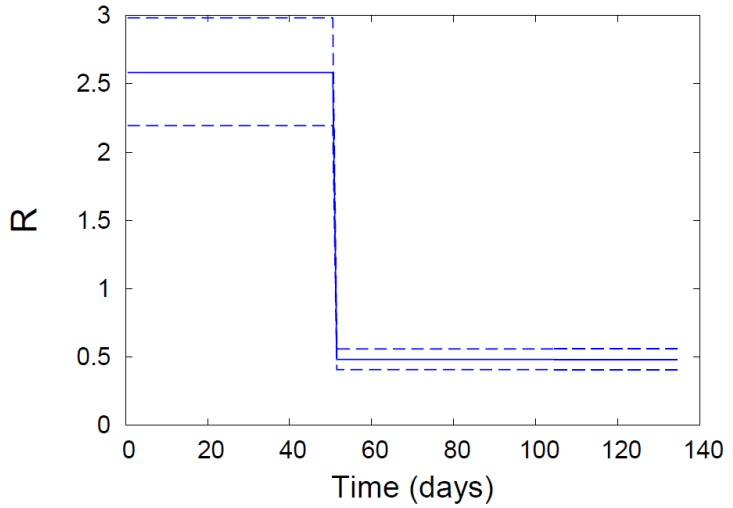


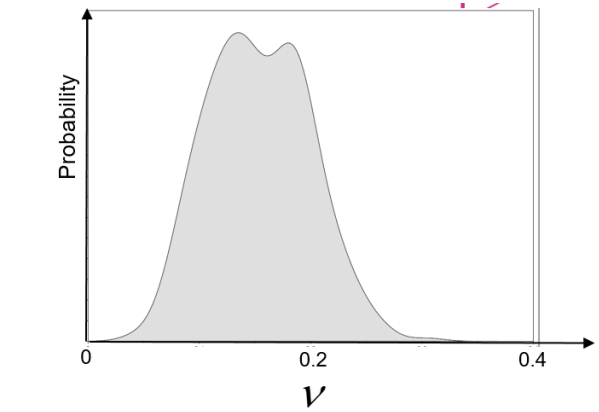
Key questions

- How is *R* changing as a function of time?
- What determines the spatial spread of disease?
- What are the effects of government policies?

BEEPmbp







$$R_a = \overline{R}\left(\frac{d_a}{\overline{d}}\right)^{\nu}$$

$\nu = 0$ Frequency dependent

$\nu = 1$ Density dependent

Model parameter inference

- Data: National hospital admissions /populations and regional death data.
- Distributions for certain transitions.

Variation in R

- Captured time variation in disease transmission before/after lockdown.

Population density

- Higher disease transmission in areas of high population density

To Do…

- Comprehensive validation.
- Results for age structured populations with sex.
- Incorporation into data pipeline.
- Different area factors, e.g. deprivation, pollution.
- Sensitivity analysis.
- Use model evidence to identify best available model for Covid-19 in Scotland.

Open Epidemiology

Requirements

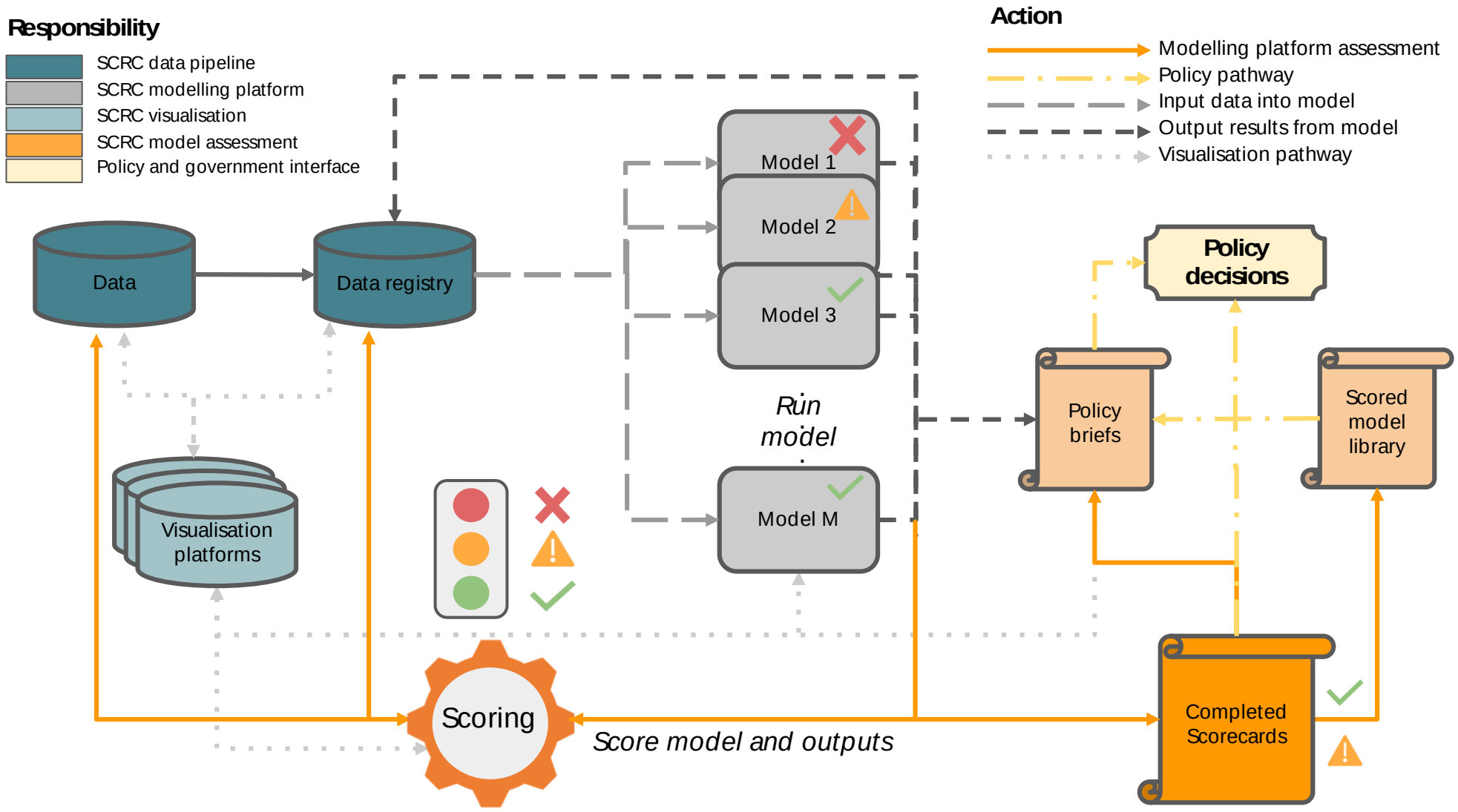
- Reproducible
- Traceable
- Validated
- Open source
- Open data

Open Epidemiology

https://github.com/ScottishCovidResponse

Requirements

- Reproducible
- Traceable
- Validated
- Open source
- Open data



RSE

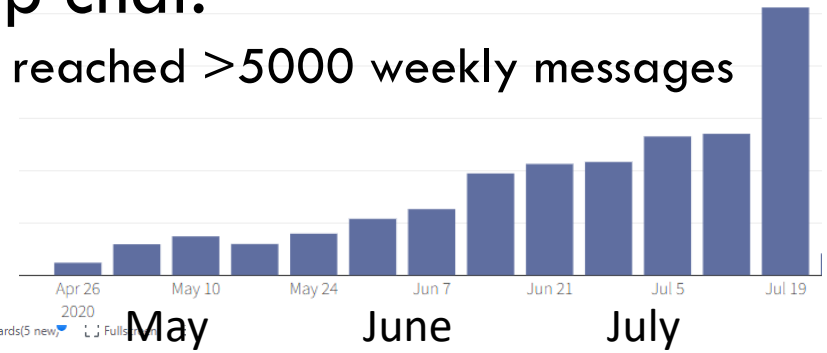
Alys Brett (UKAEA)

Team processes: Lead RSEs
Collaboration infrastructure: Andrew Lahiff, Jonathan Hollocombe, Nathan Cummings (UKAEA). Collaboration: everyone!

Enabling software teamwork

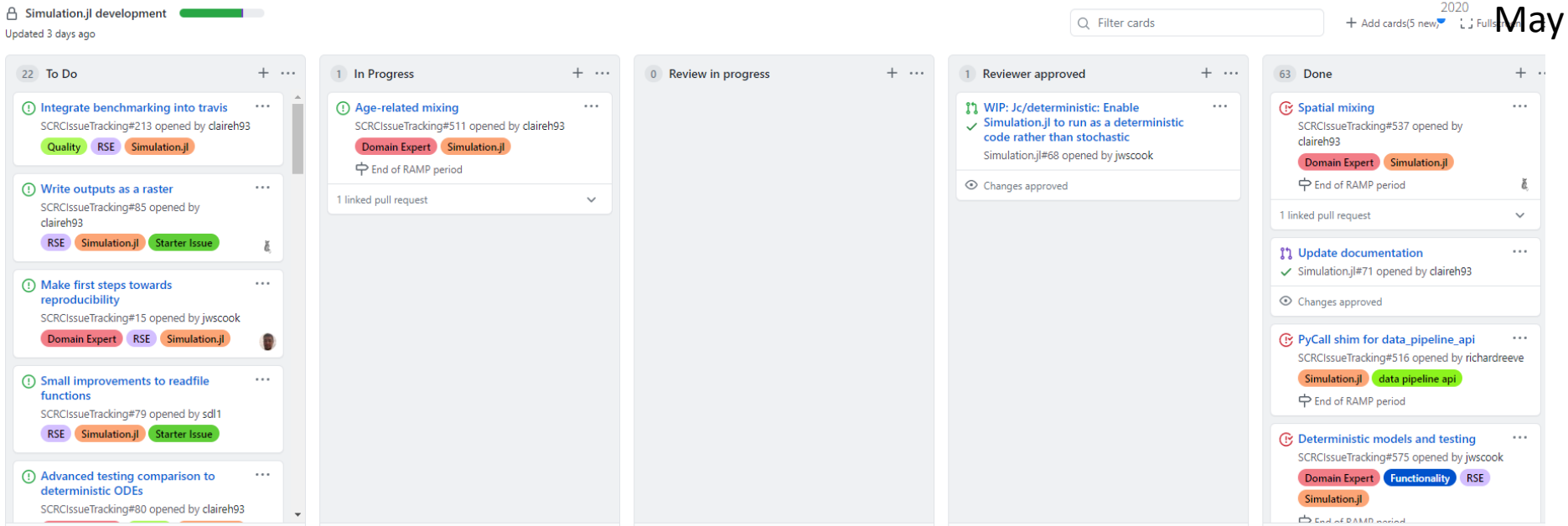
Zulip chat:

Now reached >5000 weekly messages



Github

33 repositories, 14 teams, 13 projects

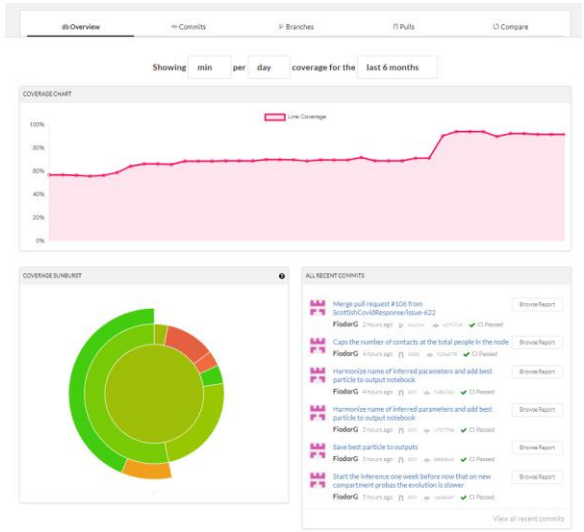


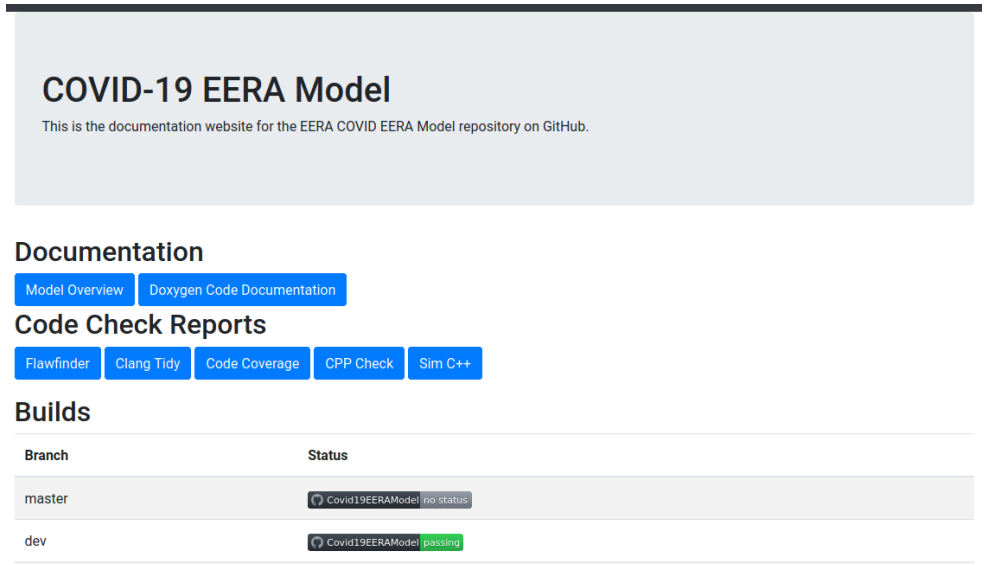
RSE

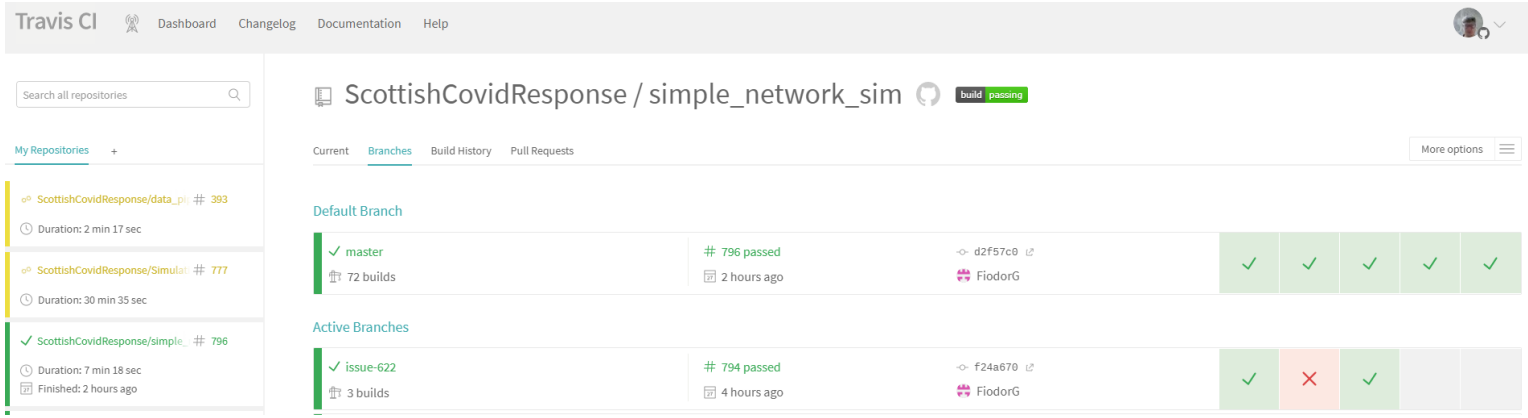
Alys Brett (UKAEA)

Software Quality

- Improving code clarity and software structure
- Automated testing and code analysis







RSE

Alys Brett (UKAEA)

Improving reproducibility and trust

- Data pipeline integration
- Managing configuration, randomness, build, release run processes

Software Checklist

https://github.com/ScottishCovidResponse/SCRC-internal/blob/master/software_checklist.md

- Framework for evaluating software engineering aspects of model readiness
- Completed for current status of all models

RSE

Alys Brett (UKAEA)

Simple Network Sim

Bob Turner (Sheffield)

- **Synchronised docs and code** using docstrings & typehints to making it easy to get into the codebase
- **Extensive automation of unit and regression tests and code-checking through CI** - lets us focus on adding features and fixing bugs
- Adopted **industry standard Python data science tools** `pandas` and `numpy` - peer-assured variable-types and functions
- **Tight control of execution using `conda` virtual environment** to help ensure reproducibility
- **Great multidisciplinary working** and input to current code from a range of contributors: academics, quantitative researchers, postdocs, engineers

Simulation.jl

James Cook (UKAEA)

- **Software teamwork practices (first multi-developer team)**
  - Issue boards & labelling, contribution and review guidelines etc
- **Automated testing**
  - Travis & Appveyor for CI
  - Codecov for test coverage
- **Extensive testing of new epidemiology code**
- **Reproducibility improvements**
  - "Deterministic hijacking" of the stochastic model to allow regression testing
  - Run scripts, TOML files for code & dependancy versions
- **Self-contained example run scripts** serve as end to end tests
- **Integration of data pipeline**

RSE

Alys Brett (UKAEA)

Contact Tracing Model and Covid Simulation Model

John Nonweiler (UKAEA)
Sam Brett

- Many **model refinements and new features**
- Initial **code clean-up** then incremental **improvements to design & implementation**
- **Automated tests and CI**
  - 90% unit test coverage, regression tests, basic automated epidemiological/model-validation tests
- **Documentation** for user, developers & to understand algorithms
- **Improved reproducibility and trustworthy results**
  - refactoring RNG usage
  - moving parameters to an input file, adding logging of git hash and inputs
  - R scripts for validation, version numbering and release processes
- Enabled **parallel runs** of models on DiRAC HPC
- Implementing the Java **data pipeline**
  - Close to end-to-end model run (using CSV instead of HDF5 files initially)

RSE

Alys Brett (UKAEA)

EERA Model

Peter Fox (UKAEA)

New functionality: Multiple model structures, forward prediction / parameter inference modes

Reproducibility: Parameterised config of random number generation, Logging inputs, outputs, versions

Software engineering and code quality

- Extensive **refactoring and architecture improvements**
- CMake **build system**
- Unit and regression **test frameworks**
- **CI automation** of: tests, documentation generation (Doxygen), code coverage, static analysis
- Created **project dashboard** https://scottishcovidresponse.github.io/Covid19_EERAModel

BEEPmbp

Ian Hinder (Manchester)

• Development of C++ interface to Python **data pipeline** code, and integration with model (in progress)

• Fixed configuration in header files replaced with **TOML configuration file**

• Storing code **provenance** information and making runs **reproducible**

• Improved **code structure,** i.e. making it more modular

• Implemented CI (**automated testing**), as well as **regression testing** (flag whether a code change affects the numerical results)

• Added **automated source code analysis** to catch bugs - found several

RSE – External Models

Alys Brett (UKAEA)

Imperial CovidSim Model

Peter Hill (York)

Integration with SCRC data pipeline

- Via a Python wrapper layer using pipeline API
- No need to touch original code
- Can treat chosen version as a black box whilst verifying inputs & outputs

LSHTM Model

Kristian Zarebski (UKAEA)

- **Rationalised parameter initialisation**
  - external config at start of run
- **Faster runs** with single scenario selected via parameter
- **Integration with the SCRC data pipeline**
  - R frontend for python to interact with pipeline API
  - Inputs converted to HDF5 & TOML - now moving to data pipeline locations
- **Collaboration with Swansea RSEs**

Data

Ruth Dundas (UoG), Iain McKendrick (BioSS)

Initial specification of Data Registry: with Aaron Reeves (SRUC)

Specification of data and metadata workflows

Literature review: procedure and associated data extraction tool

Dataset assessment: 'intrinsic quality', ‘representativeness’ and ‘fitness for purpose’ assessment procedure

Meta-analysis and distribution analysis: Jim Lewsey, Houra Haghpanahan, Ryan Field, Janine Illian, Caitlin Fox, Mahi Siddika, Megan Laxton, Martje Rave

Data brokering: Richard Bailey, Ciaran McMonagle

Specialist Dataset expertise: Ruth, Ciaran, Jeremy Walton, Denise Brown, Stephen Catterall, Amira Elayouty

Data source and provenance IDing and wrangling: Sonia Mitchell, Ruth, Ciaran, Denise

Data team: Richard, Ciaran, Sonia, Ryan, Jeremy Gribble, Hannah Jevans, Nardus Mollentze, Denise

Data pipeline

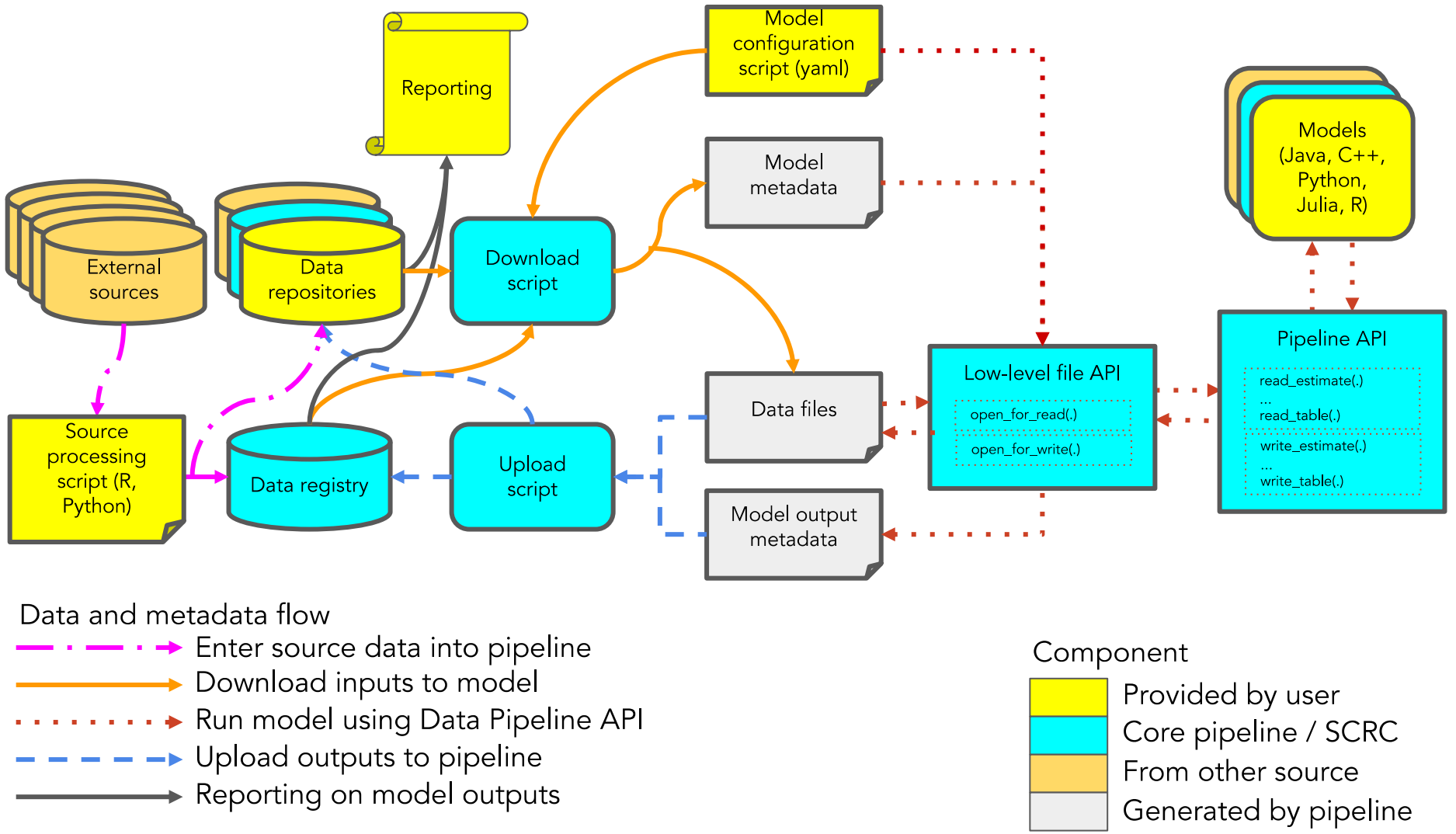
https://github.com/ScottishCovidResponse

Requirements

- Reproducible
- Traceable
- Open source
- Open data

API

- C/C++
- Java
- Python
- Julia
- R



Data pipeline for input

https://github.com/ScottishCovidResponse/SCRCdataAPI

https://github.com/ScottishCovidResponse/SCRCdata

Sonia Mitchell
Ciaran McMonagle
Nardus Mollentze
Ryan Field

SCIENCE

> 31,000 papers uploaded to the registry

8 data products uploaded to the registry

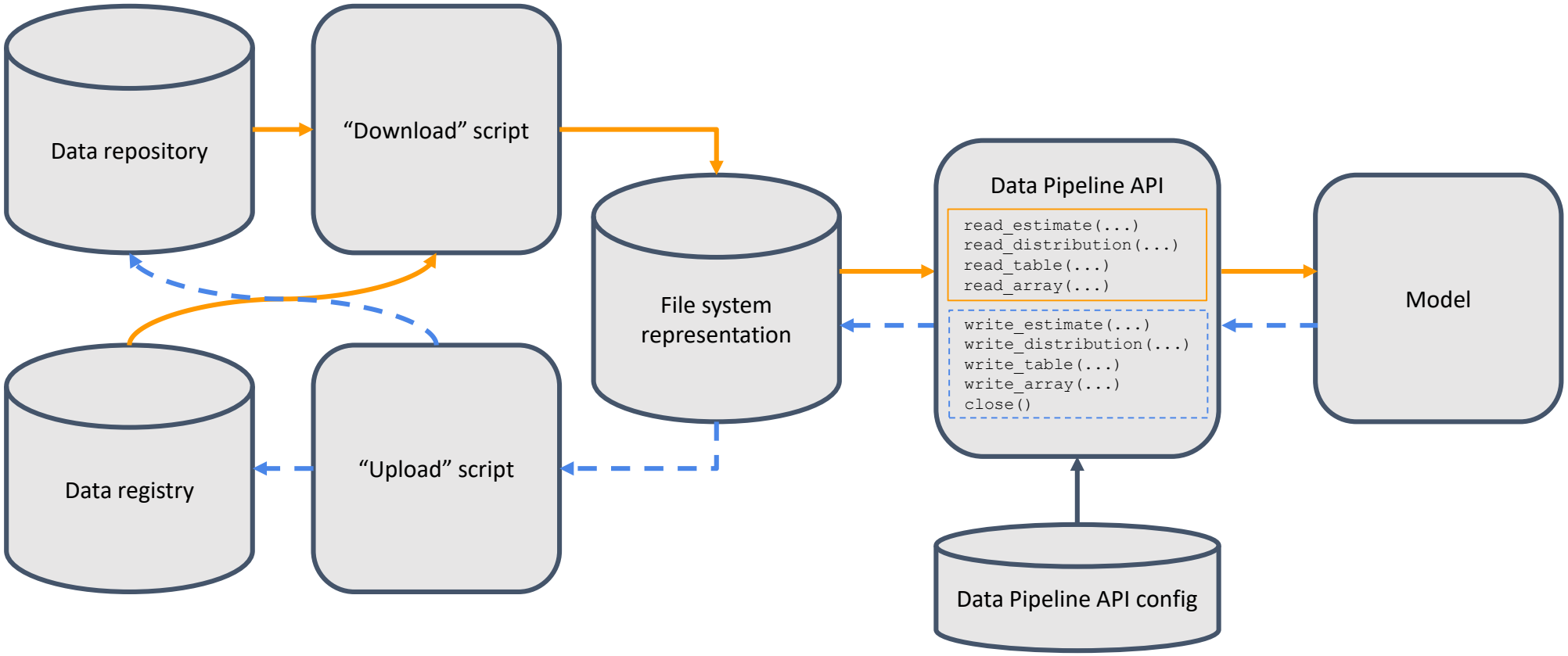
4 upload templates

Working on processing additional data products, generating upload templates, and checking robustness of code

Data pipeline for modelling

Richard Blackwell, Chris Hughes, Vino Mano, Rafael de Almeida (Man Group),
Ian Hinder (Manchester), Kristian Zarębski (UKAEA), Alex Robson,
Sean Lovett (Invenia)

https://github.com/ScottishCovidResponse/data_pipeline_api



Data registry

Jonathan Hollocombe and Andrew Lahiff (UKAEA)

https://github.com/ScottishCovidResponse/data-registry

- Available at data.scrc.uk
- Provides web-page view of stored data and information about registry API.
- API can be seen at data.scrc.uk/api
- More info available at data.scrc.uk/docs

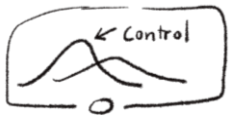
SCRC data registration and management system Links Logged in as: jholloc logout

Home

Data Products

Search

Namespace	Name	Version
EERA	fixed-parameters/K	0.1.0
EERA	fixed-parameters/T_hos	0.1.0
EERA	fixed-parameters/T_inf	0.1.0
EERA	fixed-parameters/T_lat	0.1.0
EERA	fixed-parameters/T_rec	0.1.0

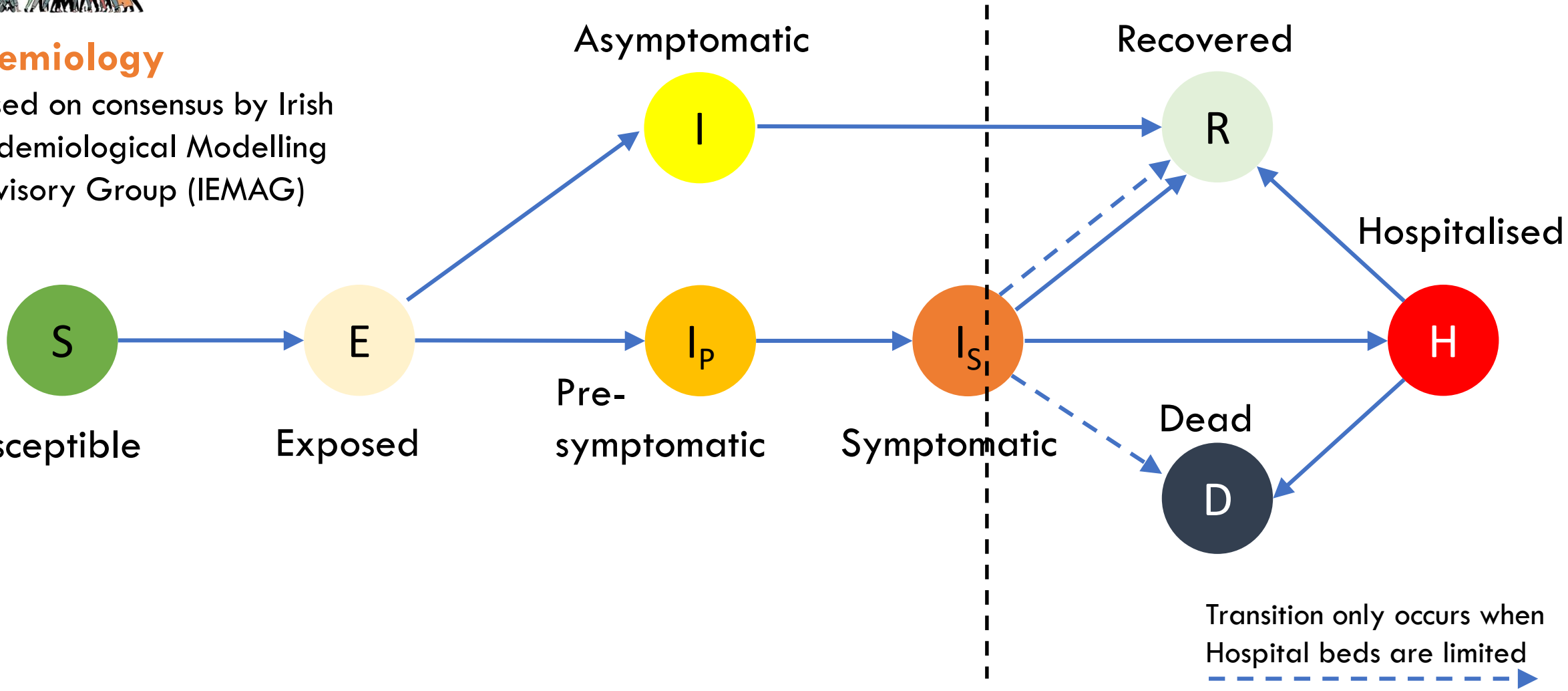


Epidemiology

Louise Matthews (Glasgow / EPIC)

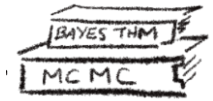
Epidemiology

- Based on consensus by Irish Epidemiological Modelling Advisory Group (IEMAG)



https://www.gov.ie/en/publication/dc5711-irish-epidemiology-modelling-advisory-group-to-nphet-technical-notes

Inference and model assessment



Glenn Marion (BioSS / EPIC)

Inference: quantify epidemic and effect of interventions

- focus on simulation-based methods (model agnostic + parallelisable)
- But also data-augmentation MCMC methods (BEEPmbp)
- And emulation for Uncertainty Quantification and model fitting

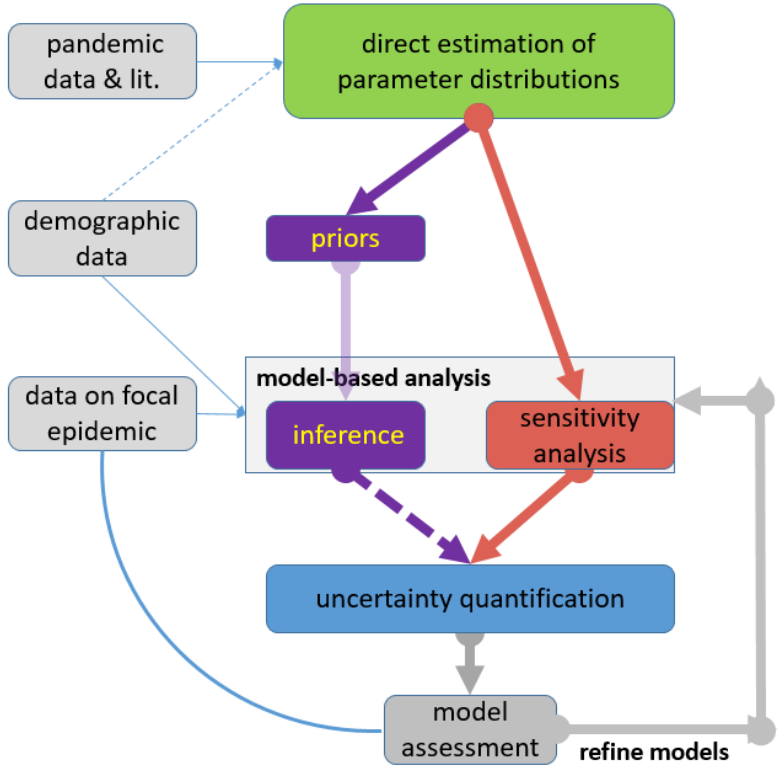
Model assessment: better models

A key issue moving forwards

- statistical fit (evidence, residuals, DIC etc.) and predictions
- model misspecification
- **validation** within and btw model/code-base

Now in more detail from

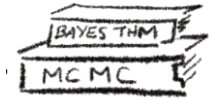
- ❑ Ben Swallow (UoG)
- ❑ Rikiya Takahashi and Fedor Gorokhovik (Man Group)



Inference* + SA > quantify uncertainty > model assessment > better models

*focus on Bayesian inference

Inference and model assessment



Ben Swallow (UoG)

Inference: quantify unknown epidemiological parameters

- Again focus on simulation-based methods (easily(?) parallelisable)
- Uses an ABC-SMC algorithm to estimate 6 parameters on 1km^2 grid of Scotland
- **quantify uncertainty** in model output and parameters
- Flexible to varying data outputs and thresholds

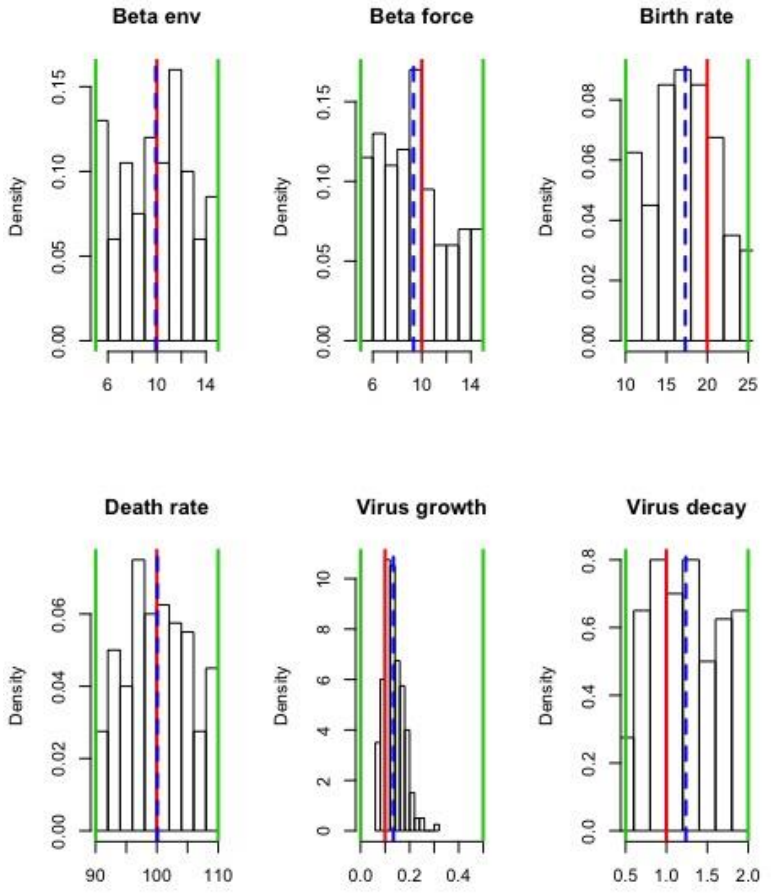
Model testing

- Using data from model simulations at known parameter values we are able to retrieve the true values
- Stochastic model so repeated runs at the same parameter values
- Tested and seems all parameters retrieved using either deaths or symptomatic cases

Issues/open questions

- Link to data pipeline
- Incorporate priors beyond uniform
- Consolidate multiple compartment summaries

Bayesian inference: simulated data



Red solid line = true parameter value
Blue dashed line = posterior mean

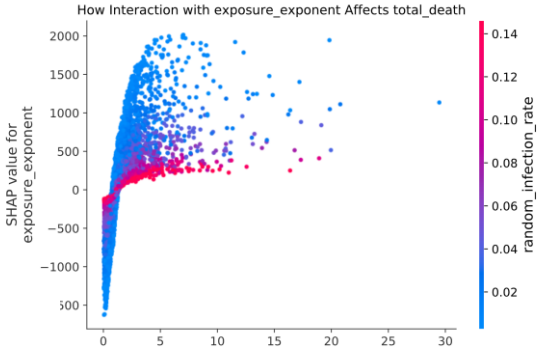
Inference and model assessment

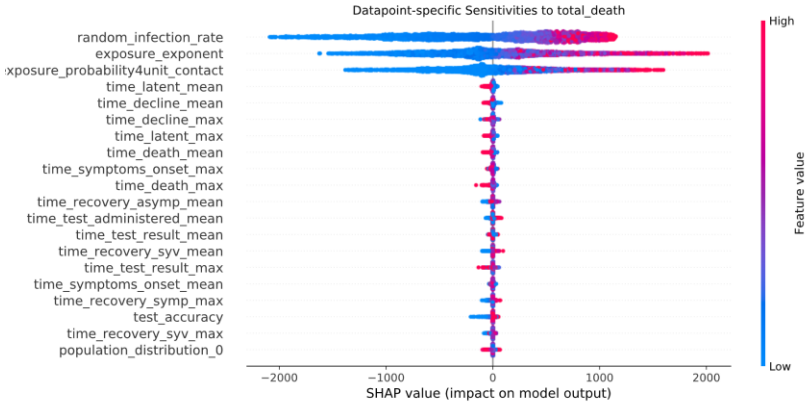
Rikiya Takahashi and
Fedor Gorokhovik (Man Group)

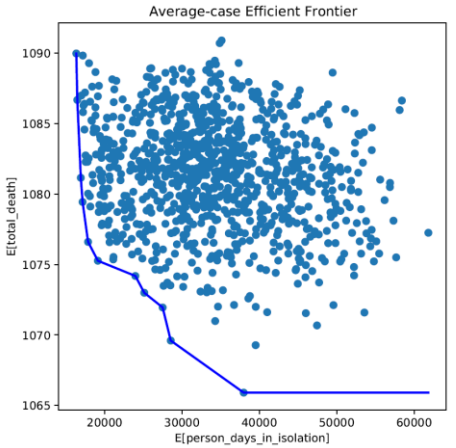
Sensitivity Analysis

- Reparametrisation of exposure-related disease parameters in Contact Tracing Model
- Non-linear regression and SHAP analysis for total number of death and peak number of severely symptomatic persons
  - SHAP: SHapley Additive exPlanations (Lundberg & Lee, 2017)
- Efficient frontier in optimisation of the policy parameters under uncertainty

Open Issues Real-data-based posterior in policy optimiser, distributional policy parameter recommendations instead of point estimates



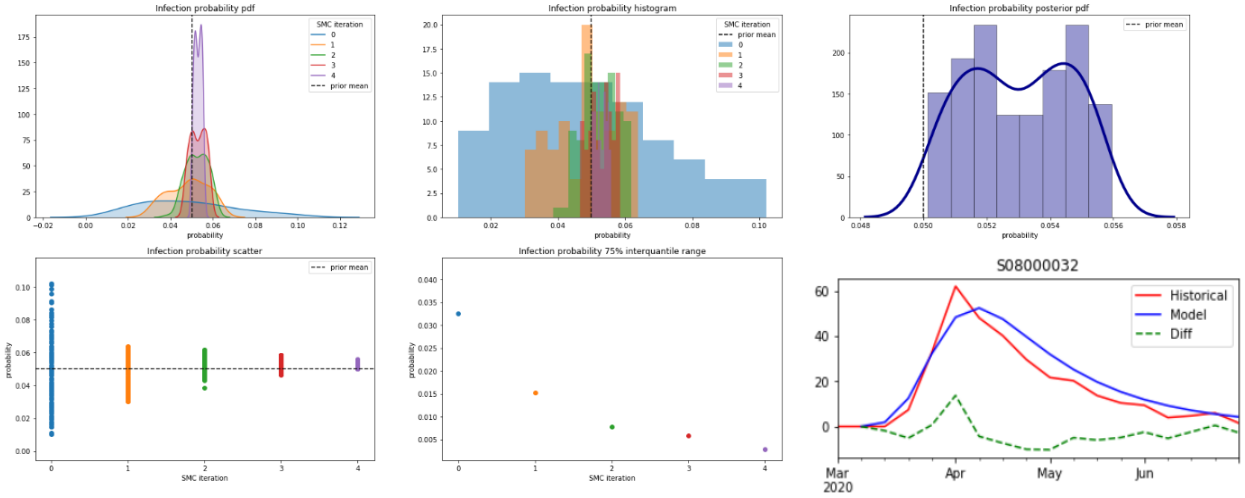




Inference

- ABC-SMC used to infer infection probability and initial infections in Simple network sim (contact multipliers to come)
- To be used in conjunction to stochastic mode to have more realistic posterior std dev
- Fitted on historical death per healthboard in Scotland
- Other parameters sources from external sources to be as credible as possible
- Wide priors chosen from Beta, Lognormal distribution

Open Issues Improved rejection mechanism, parallelization of inference



S. M. Lundberg & Su-In Lee, A Unified Approach to Interpreting Model Predictions, NeuIPS 2017.

VIS

27 VIS/VA volunteers, plus Min Chen (Oxford U.)

https://github.com/ScottishCovidResponse/rampvis-ui

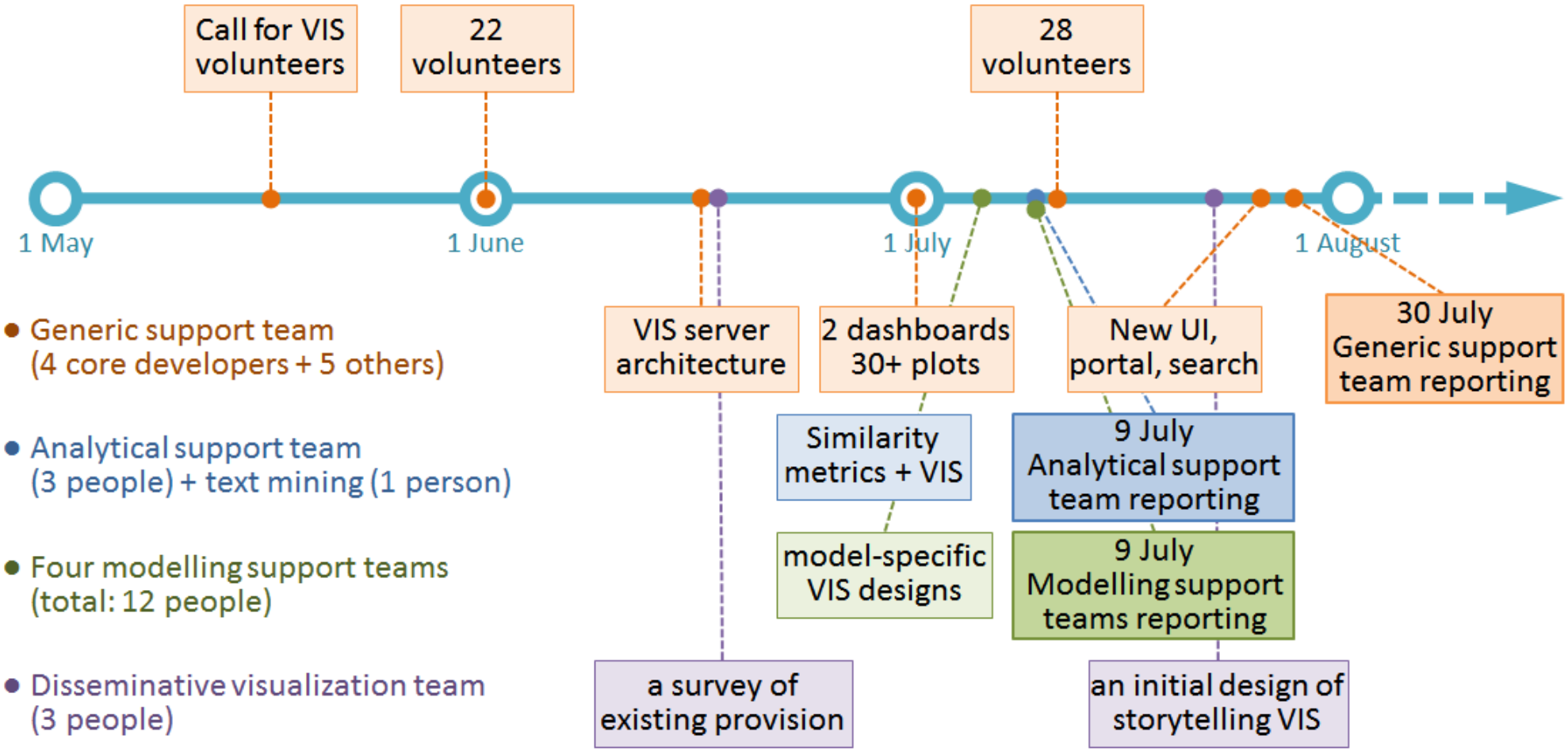
rampvis-api

scrc-vis-analytival

scrc-vis-modelling

Dr. Alfie Abdul-Rahman, *KCL*
Dr. Daniel Archambault, *Swansea*
Dr. Benjamin Bach, *Edinburgh*
Dr. Rita Borgo, *KCL*
Professor Min Chen, *Oxford*
Professor Jason Dykes, *City*
Dr. Hui Fang, *Loughborough*
Elif Firat, *Nottingham*
Dr. Radu Jianu, *City*
Professor Nigel W. John, *Chester*
Dr. Saiful Khan, *Horus Security*
Dr. Robert S Laramee, *Nottingham*
Dr. Stella Mazeri, *Edinburgh*
Benjamin Nash
Dr. Phong H. Nguyen, *Red Sift Ltd.*
Dr. Helen C. Purchase, *Glasgow*
Dr. Dylan Rees, *Swansea*
Dr. Panagiotis Ritsos, *Bangor*
Professor Jonathan Roberts
James Scott-Brown, *Edinburgh*
Dr. Aidan Slingsby, *City*
Dr. William Teahan, *Bangor*
Dr. Thomas Torsney-Weir, *Swansea*
Dr. Cagatay Turkay, *Warwick*
Dr. Franck Vidal, *Bangor*
Qiru Wang, *Nottingham*
Tianci Wen, *Bangor*
Professor Jo Wood, *City*
Dr. Kai Xu, *Middlesex*

- Generic support team (4 core developers + 5 others)
- Analytical support team (3 people) + text mining (1 person)
- Four modelling support teams (total: 12 people)
- Disseminative visualization team (3 people)



VIS

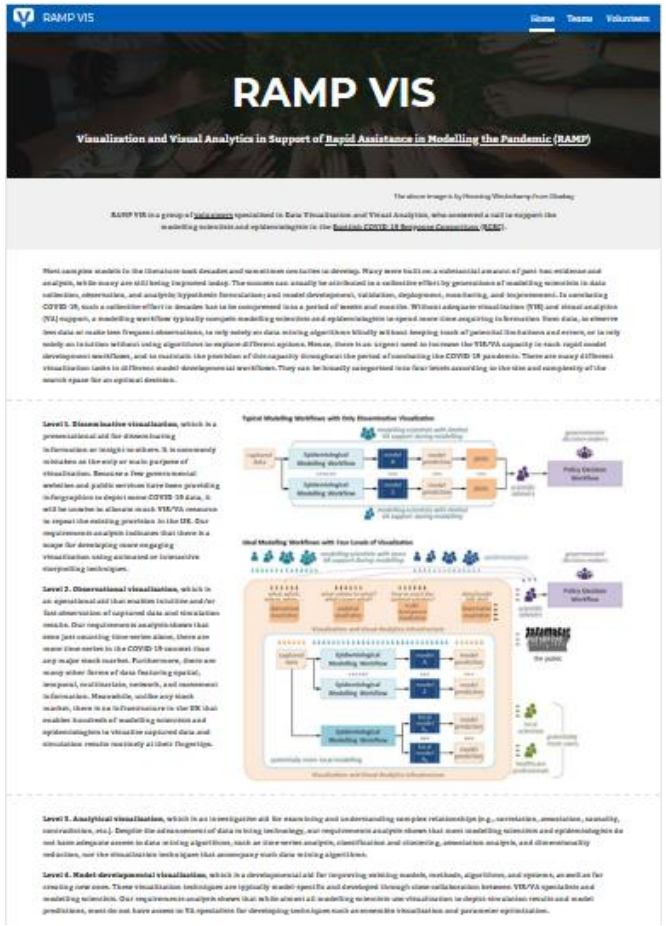
27 VIS/VA volunteers, plus Min Chen (Oxford U.)

https://github.com/ScottishCovidResponse/rampvis-ui

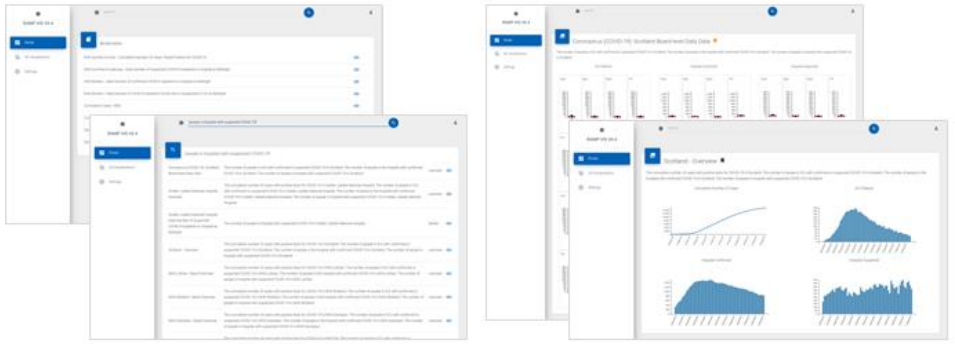
rampvis-api

scrc-vis-analytival

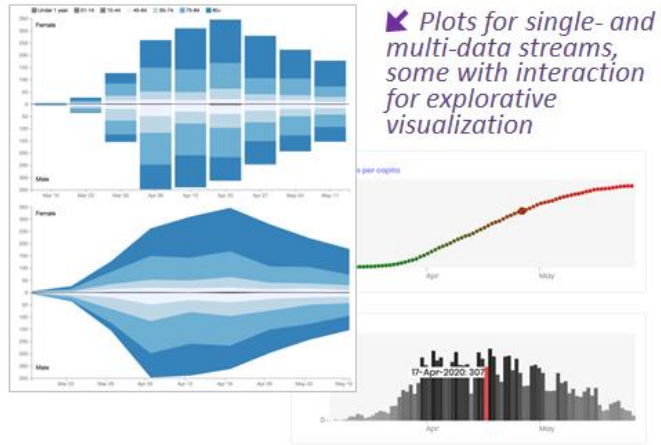
scrc-vis-modelling



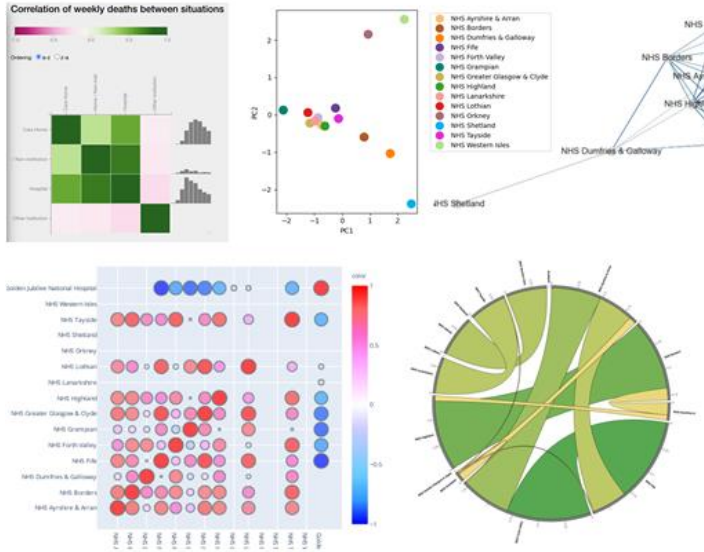
https://sites.google.com/view/rampvis/



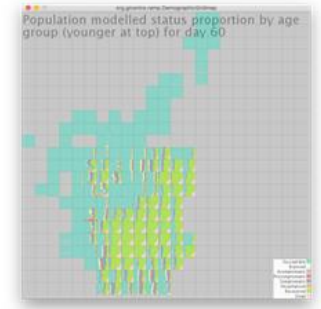
The RAMP VIS interface with personal portals, a search facility, dashboards.



Plots for single- and multi-data streams, some with interaction for explorative visualization



Different analytical visualizations depicting similarity measures computed using different analytical algorithms



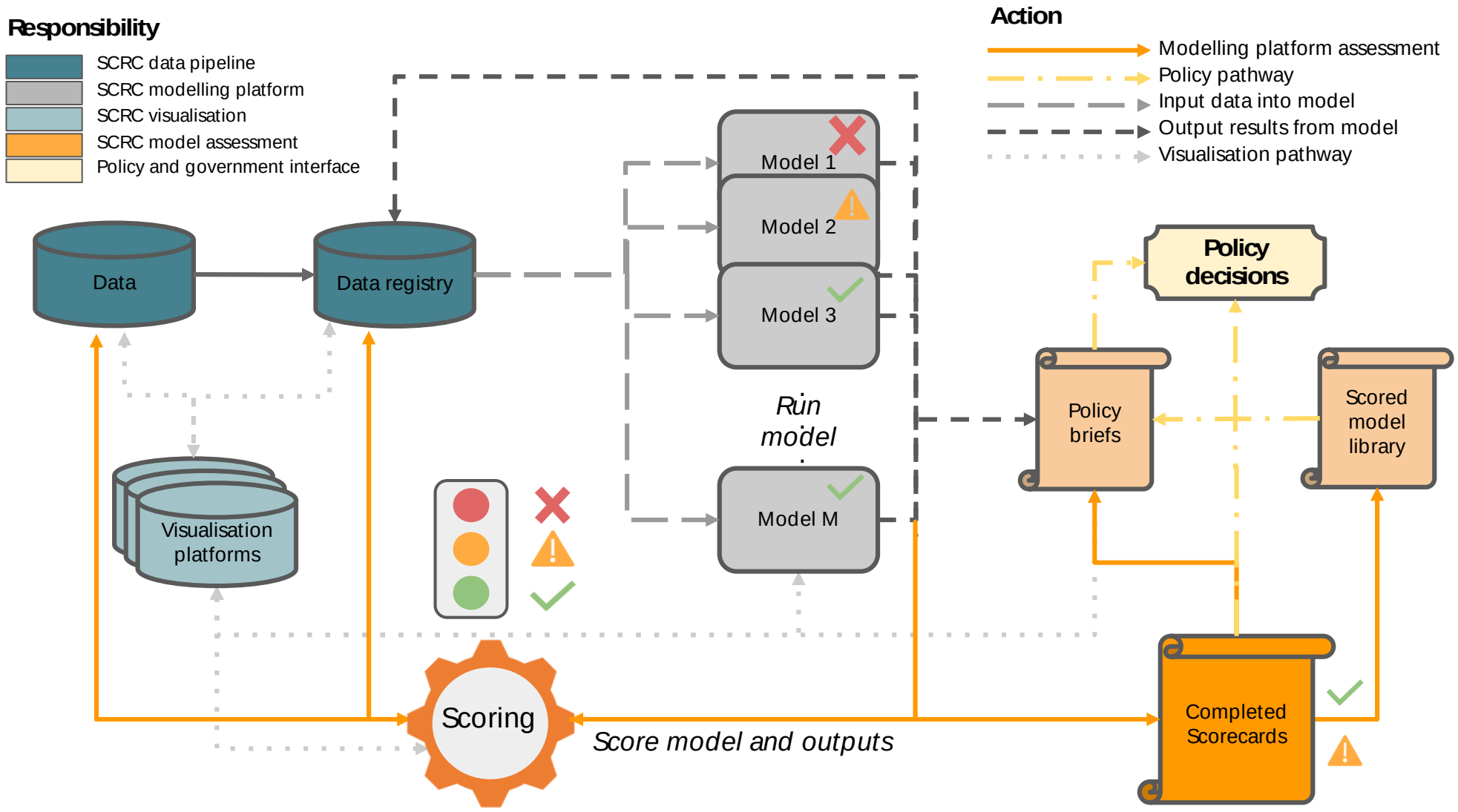
Model-specific visualizations

Open Epidemiology

https://github.com/ScottishCovidResponse

Requirements

- Reproducible
- Traceable
- Validated
- Open source
- Open data



Administration

Lynn Esson (Glasgow) and Yasmin Abdalla (Edinburgh)

Rachel O'Brien (Glasgow)
Mark Meenan (Glasgow)

Requirements

- Extreme patience

https://github.com/ScottishCovidResponse

Universities of Glasgow, Edinburgh, SRUC, Sheffield, Manchester,
St Andrews, Stirling, Strathclyde, Heriot-Watt, Southampton,
Cambridge, Oxford, Swansea, Bangor, Nottingham,
Middlesex, City, U. of London, Warwick, Chester,
King's College London, Loughborough, EPIC,
Biomathematics and Statistics Scotland,
UK Atomic Energy Authority, Met Office,
Software Sustainability Institute,
Man Group, Invenia,
Horus Security,
Red Sift Ltd,
and others