IPACS MODEL

Technical Manual
v1.0

IPACS Team
Paul Forte, Alison Harper, Zehra Önen-Dumlu, Richard Wood,
Christos Vasilakis, Martin Pitt

April 2023

Table of contents

Introduction

The IPACS project – 'Improving the Flow of Patients between Acute, Community and Social Care' – was funded by Health Data Research UK. It ran from May 2020-March 2023, and was undertaken by a research team of six from Bristol, North Somerset, South Gloucestershire Integrated Care Board (BNSSG ICB) and the Universities of Bath and Exeter.

The IPACS project aimed to investigate what might constitute an 'optimal balance' of capacity along different parts of the complex care discharge pathways from acute hospital to community health and social care, and to assess how responsive total spend in the local health and social care economy might be in relation to community-based care capacity.

This guide provides details about how to parameterise and run the IPACS simulation model. It should be read in conjunction with the the previous document to understand the purpose and scope of the model.

The IPACS simulation model is a high-level computer model of the discharge-to-assess (D2A) pathways. It takes a set of input parameters (from an Excel file) and estimates potential future service outputs (occupancy, number of patients with a discharge delay, number of days waiting for discharge, total system costs) based on different configurations of parameters. The model accounts for variation in inputs and outputs, and presents results over time.

The simulation model is not intended to provide a precisely accurate set of outputs but to help structure understanding of the D2A system and available data sources affecting referrals into each pathway and locality. Comparing the effects of changing different parameters of the system (such as daily arrivals, lengths of stay, and capacity) aims to provide some insights into where best to tailor real world interventions.

To run the model, you need to have R and RStudio installed. The first time you run the model, all packages will need to be installed. The model has been annotated carefully to assist with any code adaptations. We recommend testing the code before committing any changes.

1. Overview and directory structure

- **IPACS_MODEL.Rproj** will launch RStudio in the correct directory.
- **ipacs_main_script.R**: This runs all of the functions in the model and outputs the report and data into the 'outputs' folder.
- The model functions are in the **/functions** subfolder:
  - **set_up.R**: Reads in data from model_inputs/IPACS_params.xlsx to parameterise the model, and re-formats it for the model functions to use.

    OPTIONAL:

    There is an option here to change how lognormal length-of-stay (los) parameters are calculated. A lognormal distribution is a convenient distribution for modelling healthcare los, where all los are positive (no one stays less than 0 days) and a few people have very long los, so the distribution is skewed to the right.

    The method depends on the los data parameters you have available for each pathway. If you only have the mean and median (i.e you do not know the standard deviation of the pathways los), you can use the default in ipacs_main_script est_method <- 1. *Here, the parameters are calculated in the Microsoft Excel input file for you to see. This is likely to be the simplest method. In this case, leave the sd_los column blank in the los tab of IPACS_params.xlsx*

    If you have the mean and standard deviation available, you can set ipacs_main_script sd_los <- standard deviation *and* est_method <- 2. *This is a more accurate method but the default method is acceptable and likely easier.*

    Each method is described in set_up.R.
  - **bed_functions.R**: A set of functions used by the bed-based model.
  - **bed_model.R**: Runs the simulation for P2/P3 cases.

    Outputs bed-based pathways (P2/3) raw data in the /outputs folder for report_data/.csv (the data used by the .Rmd file to create the .docx report) and stochastic_data (95% quantiles) for bed-based pathways.
  - **visit_functions.R**: A set of functions used by the visit-based model.

- **visit_model.R**: Runs the simulation for P1 cases.

Outputs visit-based pathways (P1) raw data in the /outputs folder for report_data/.csv (the data used by the .Rmd file to create the .docx report) and stochastic_data (95% quantiles) for bed-based pathways.

- **ipacs_produce_report.Rmd**: An R Markdown file that creates the 'IPACS Report' output based on the report_data/.csv produced by the models. You can adapt this file for your local needs.

- The input data for the model is stored in the **/model_inputs** subfolder.
- The **/images** subfolder contains images used in the 'IPACS Report' output.
- *OPTIONAL:* The **/testing** subfolder contains data and scripts used to test the model. It can be used in the following way to monitor code changes:
    - Keep a copy of the existing IPACS_params.xlsx for testing – you can rename.
    - The data in the /testing subfolder uses the parameters in IPACS_params.xlsx
    - Open testing_and_linting.R
    - Change the working directory
    - Check the name of all '_new' data files. They currently append '_IPACS_params'. If you have changed the name of the input file for testing, you will need to change them here. There is no need to change the name of those read in as '_test'.
    - The script will check that results are consistent. It also runs a linter which will check that your code changes are tidy.

The complete directory is structured as follows:

```
|_ipacs-model
    |_IPACS_MODEL.Rproj
    |_____IPACS_main_script.R
    |_____functions
        |_____set_up.R
        |_____bed_functions.R
        |_____bed_model.R
        |_____visit_functions.R
        |_____visit_model.R
        |_____ipacs_produce_report.Rmd
    |_____model_inputs
        |_____IPACS_params.xlsx
    |_____outputs
        |_____report
            |_____.docx
        |_____report_data
            |_____ .csv files
        |_____stochastic_data
            |_____.csv files
    |_____Images
        |_____.png
    |_____Documentation
        |_____Overview.pdf
        |_____Technical User Guide.pdf (this document)
        |_____STRESS-DES reporting guidelines
```

Scenarios

The model will produce outputs for a number of different scenarios to allow comparisons between different assumptions.

The assumptions are defined by the scenarios in the Arrivals, Capacity and los (Length of Stay) tabs in IPACS_params.xls. These are prefixed with B (for Baseline), S1, S2 and so on.

The scenarios are independent of one another. Each scenario is combined with each other scenario to make a series of assumptions. Each additional scenario will cause an exponential increase in runtime. A maximum of 12 scenarios can be visualised on the output plots without appearing too crowded – this equates to the product of the number of scenarios inputted into the excel file.

For example:

- 2 arrivals scenarios * 1 capacity scenario * 2 los scenarios = 4 total scenarios
- 2 arrivals scenarios * 2 capacity scenarios * 3 los scenarios = 12 total scenarios

2. Running the model

a. Getting started:

- Ensure you have R (Version 4.2.2) and RStudio installed
- Download or clone the repository from: https://github.com/nhs-bnssg-analytics/ipacs-model into a local directory
- A sample Microsoft Excel input file is included in the **/model_inputs** subfolder to get started running the model.

The model can be run with very little interaction with the model code. The basic requirements are the following:

- Open Rstudio in the directory, or navigate to the directory in the 'Files' tab of Rstudio. Alternatively, if you launch Rstudio by opening **IPACS_MODEL.Rproj**, it will open in the correct directory.
- Open **IPACS_main_script.R**
- Set your working directory: Write getwd() in the console window and copy the output which is produced by this command to "wd <- setwd("~/set_working_directory_here") in the script window (top left window) to contain the filepath holding the scripts.
- Open the **inputs/IPACS_params.xlsx** and update parameters if you wish. Save and close the file. You may change it to a new name (all outputs will append the new name) – if so, also change it in: **input_filename <- "new_name.xlsx" in IPACS_main_script.**
- If you have not previously run the model, install the necessary packages in IPACS_main_script using install.packages("package name") or Tools>Install Packages.
- Update the number of replications (if necessary) at **nruns:**

  1. The model is *stochastic*, which means it has variables in it that can change randomly with individual probabilities. Realizations of these random variables are generated and inserted into the model at runtime. Outputs of the model are recorded, and then the process is repeated with a new set of random values. These steps are repeated until a sufficient amount of data is gathered. In the end, the distribution of the outputs shows the most

probable estimates as well as a frame of expectations regarding the likely range of values of the variables.

2. The model uses *50 replications* by default. Depending upon your computer setup, the number of days you are running the model, and the number of scenarios, this may take some time. You can reduce the number of replications to get some indicative results, but be aware that the probable estimates may differ somewhat with fewer replications. To collect summary data, you'll need a minimum of 3 replications.

- Run **IPACS_main_script** (code<Run Region<Run All OR use keyboard shortcuts *'Ctrl + Alt + R' for Windows or Linux users, or 'Cmd + Alt + R' for Mac Users*), excluding the installation of packages, and (optionally) excluding the Rmarkdown creation. This may be useful if you wish to run the model with more than 12 scenario combinations, and create your own plots from the .csv output files. To exclude the output report, comment out every line generating the report by putting "#" in front of the corresponding line of code.

b. Expected outputs:

- The model will create three sets of outputs in the /**outputs** subfolders. Each output appends the name of the input file:

1. A word document called **"report/ipacs_report_"** which contains a set of plots, one for each pathway-locality (node). Each plot contains all of the chosen scenario combinations , with four facets of output metrics:

   (i) The mean number of acute patients delayed per day

   (ii) The mean number of patients in service per day

   (iii) The mean number of days delayed

   (iv) The mean daily cost (a combined total of the cost of acute delayed discharge and the cost of the community service).

2. **report_data/bed_output_.csv**

   report_data/visit_output_.csv

   These two files contain the *mean data* used to generate the plots in the report. Re-running will overwrite an existing file using the same name, so rename files if you wish to keep them. The data is presented in 'wide' format with a column for each scenario, and each mean metric.

   For example, column P2_B_Bcap_Blos_S1Arr__occ refers to pathway P2, locality B, Baseline capacity, Baseline los, Scenario 1 arrivals, metric = occupancy.

Column P3_NS_S1Cap_S1los_S1Arr__niq refers to pathway P3, locality NS, Scenario 1 capacity, Scenario 1 los, Scenario 1 arrivals, metric = number in queue (i.e. delayed discharges).

3. **stochastic_data/stochastic_bed_output_.csv**

   stochastic_data/stochastic_visit_output_.csv

   These two files contain *stochastic output data*, rather than only the mean results across runs. The distribution of the outputs shows the most probable estimates as well as a frame of expectations regarding what ranges of values the variables are more or less likely to fall in. It is presented in 'long' format for easy plotting if required. Re-running will overwrite an existing file on the same date, so rename files if they need to be kept.

   Column 'node' is the scenario combination, e.g. P1_B_Bcap_Blos_BArr. Each scenario combination has data per day and date of the model runtime in the form of quantiles plus the median and mean. Column 'measure' is the output metrics: niq (number in queue in acute); occ (number in occupancy in community service); wait (days waiting in acute hospital); cost (total combined system cost of delayed discharges and community service provision).

3. Parameterisation

Model parameters can be changed for the baseline case and experimental scenarios in model_inputs/IPACS_params.xlsx (or create and name a new file – ensure you change the name of input_filename <- "new_name.xlsx" in IPACS_main_script.R if you wish to parameterise from a new file).

The simulation model requires aggregate-level data to run, and this is read entirely off the .xlsx workbook, which is structured with five tabs.

Each tab provides a field called '**node**'. The node is a linked combination of the pathway (P1, 2 or 3) and the locality (if multiple locality pathways are required). Ensure that there is a '**node**' in each tab, and that it is formatted P1/2/3_locality. e.g.P1_LocA. Each node represents a service to be simulated, with a baseline case and any additional experimental scenarios.

Tabs '**arrivals'**, '**capacity**' and '**los'** can be used to introduce scenario experimentation. Each of these include a column 'scenario' with will require a baseline (e.g. BArr represents baseline arrivals, that is, the current system state) for each locality and each date (where relevant). Where a scenario is tested, for each node (and each date where relevant) there will be a second scenario value (e.g. S1Arr represents scenario 1 arrivals). The values can be adjusted as required for each scenario (for example, arrivals+10%) for each node.

A maximum of *12 scenario combinations* can be outputted into the markdown report, including the baseline, for the outputs to be digestible. Scenario combinations will include ALL COMBINATIONS of scenarios added into the Microsoft Excel input file. This means that for each of the variable parameters, two scenarios (including baseline) can be inputted. One variable may have a second scenario (for example you may wish to investigate arrivals +/- 10%). One suggestion is to use two scenario combinations (including the baseline) for each variable parameter (arrivals, capacity, los) plus adding an 'infinite' capacity scenario.

However, any number of *scenario combinations* can be outputted to the .csv files generated by the models (optionally comment out the lines of code which call the markdown script, at the end of IPACS_model_script.R).

Mean **arrivals** (Figure 1) are a daily input to the model, and have a 'date' field which will correspond to the model run time. Arrival rates are a 'Poisson' process, which follows a discrete probability distribution. These are inputted daily for each node as a mean daily arrival rate for baseline arrivals (representing the current system state) and for any arrivals scenarios you wish to investigate.

There are several ways arrivals scenarios might be managed. For example:

* You could simply use mean current referral activity and project this forward for the desired time period;

* You could use mean current referral activity per day of week and project this forward over time.

* Apply a 'growth factor', for example a winter uplift if that is part of your model run time horizon and you reasonably expect demand to change;

* Supply other estimates as available, such as forecasted demand, which may be based on forecasted acute arrivals adjusted (for example) for acute los, proportion of complex discharges, and proportion entering each pathway.

* For each locality, the overall numbers can be varied across pathways, for example to experiment with a scenario of reducing referrals into bedded pathways, and increasing referrals into P1-type pathways.

* In all cases the data input process is the same with the number of rows of data depending on the time horizon to be modelled, as well as the node (number of pathways and number of localities) and scenarios (baseline +/- other scenario):

node	arrivals	scenario	date	day
P1_B	10.48	BArr	2022-11-16	3
P1_NS	6.93	BArr	2022-11-16	3
P1_SG	8.04	BArr	2022-11-16	3
P2_B	4.59	BArr	2022-11-16	3
P2_NS	3.89	BArr	2022-11-16	3
P2_SG	3.37	BArr	2022-11-16	3
P3_B	2	BArr	2022-11-16	3
P3_NS	1.44	BArr	2022-11-16	3
P3_SG	3.44	BArr	2022-11-16	3
P1_B	13.28	S1Arr	2022-11-16	3
P1_NS	9.53	S1Arr	2022-11-16	3
P1_SG	11.55	S1Arr	2022-11-16	3
P2_B	1.9	S1Arr	2022-11-16	3

Figure 1: 'arrivals' tab for entering daily arrival rate per node and scenario

Capacity (Figure 2) is the current capacity in terms of beds (for P2/3 pathways) and number of patients accommodated (for P1 pathway). It is entered per node (and scenario, if relevant). For an 'infinite capacity' scenario, enter a very large number (e.g., 2000) to estimate how much capacity is required to have no acute patient discharges delayed.

node	scenario	capacity
P1_B	BCap	92
P1_B	S1Cap	2000
P1_NS	BCap	71
P1_NS	S1Cap	2000
P1_SG	BCap	71
P1_SG	S1Cap	2000

Figure 2: 'capacity' tab

Length of stay (los) (Figure 3) is inputted as the mean, and median for each node. The distributions and distribution parameters are calculated from these inputs and do not need to be adjusted. If you don't have the standard deviation (sd_los) please leave this column empty.

*[Please see note in Section 1. re **set_up.R** There is an option in this script to change how the lognormal parameters are calculated].*

Additionally, for P1 pathway, two additional, optional parameters are provided. These are 'IVR' and 'FVR', which correspond to 'initial visit rate' and 'final visit rate'. These are used for calculating P1 visits-based capacity based on time-varying capacity requirements, assuming that patient visit requirements decrease across their service duration. The default settings are a mean IVR of 4, and a mean FVR of 1. Note that the

upper limits are truncated in the model code, and the standard deviation is 0.5.

node	median	sd_los	IVR	FVR	scenario	mean_los	los_params	los_dist	mu	sigma
P1_B	9		4	1	Blos	18.07	2.197225 , 1.180702	lnorm	2.197225	1.180702
P1_B	5		4	1	S1los	10	1.609438 , 1.17741	lnorm	1.609438	1.17741
P1_B	7		4	1	S2los	14.04	1.94591 , 1.179831	lnorm	1.94591	1.179831
P1_NS	5		4	1	Blos	18.7	1.609438 , 1.624245	lnorm	1.609438	1.624245
P1_NS	2.7		4	1	S1los	10	0.993252 , 1.618229	lnorm	0.993252	1.618229
P1_NS	3.8		4	1	S2los	14.35	1.335001 , 1.630183	lnorm	1.335001	1.630183
P1_SG	5		4	1	Blos	12.08	1.609438 , 1.328242	lnorm	1.609438	1.328242
P1_SG	4.1		4	1	S1los	10	1.410987 , 1.335364	lnorm	1.410987	1.335364
P1_SG	4.6		4	1	S2los	11.04	1.526056 , 1.32323	lnorm	1.526056	1.32323
P2_B	34				Blos	42.79	3.526361 , 0.67815	lnorm	3.526361	0.67815
P2_B	16.7				S1los	21	2.815409 , 0.676925	lnorm	2.815409	0.676925
P2_B	25.3				S2los	31.895	3.230804 , 0.680654	lnorm	3.230804	0.680654

Figure 3: 'los' tab with mu/sigma parameter calculations

Tab **'initial conditions'** (Figure 4) initialises the model with the starting system state for 'delayed transfers of care' (dtoc [usually now referred to as 'no criteria to reside']), which are acute beds that have had a navigation decision; and 'pathway occupancy' (occ), which are community beds/caseload capacity for each node.

node	measure	value
P1_B	dtoc	39
P2_B	dtoc	31
P3_B	dtoc	43
P1_NS	dtoc	48
P2_NS	dtoc	15
P3_NS	dtoc	12
P1_SG	dtoc	42
P2_SG	dtoc	12
P3_SG	dtoc	11
P1_B	occ	92
P2_B	occ	97

Figure 4: 'initial conditions' tab

Tab **'costs'** (Figure 5) provides daily indicative costs for one day of acute care, and one day of community care per pathway node. This enables an estimate of the total system cost of different scenario combinations, as a combination of the cost of acute delays and the cost of community service provision. The **costs** provided in the input file are derived as per Table 1.

node	community_cost	acute_dtoc
P1_B	125	346
P2_B	150	346
P3_B	164	346
P1_NS	125	346
P2_NS	150	346
P3_NS	164	346
P1_SG	125	346
P2_SG	150	346
P3_SG	164	346

Figure 5: 'costs' tab

Table 1: Derivation of costs provided in the input file

	Average cost of weekly service	Relative Cost ratios	Source of costs
P1	£875	5	2017/18 NHS reference costs
P2	£1,050	6	National Audit of Intermediate Care 2017/18
P3	£1,150	7	System costs from Bristol care system
Acute	£2422	14	2017/18 NHS reference costs

4. Tailoring the model to your locality

The IPACS model is licensed with a GPL-3 You are welcome to download, share, modify or re-use the models/code as required. You should re-use the same license.

A summary of required data parameters:

- Numbers of patients already in each of the P1-P3 pathways (per locality if relevant)
  [in .xlsx: initial conditions tab, measure = occ]
- Current available capacity (ie. number of beds (P2, P3); home care caseload capacity (P1))
  [in .xlsx: capacity tab]
- Average length of stay for a patient in each complex care pathway and each locality
- [in .xlsx: los tab. Enter median and mean los]

- The number of patients currently delayed in acute beds who have been assigned to a D2A pathway but cannot be discharged into it because there is no capacity
  [in .xlsx: initial conditions tab, measure = nctr / dtoc]
- The number of patients of referred for each pathway and locality (I.e. activity demand)
  [in .xlsx: arrivals tab, required daily inputs for each scenario]

Note that if exact values for any of these parameters are not readily available then you can always enter an 'expert guess', or a national average figure to use as a starting point and simply revise this as more specific or accurate local data become available. Any changes will be accounted for the next time the model is run.

5. Validation and model scope

The simulation model is not intended to provide an accurate set of outputs.

It provides a high-level overview of the system structure as a computer model, and when run over time, it will simulate the flow of patients through the system in increments of one day. The model makes a number of assumptions, for example:

- All patients referred into a pathway will enter the pathway – in reality some patients may enter a different pathway, or not enter a pathway at all.
- If there is capacity in a pathway and patients queuing, they will enter the pathway – in reality, patients may be delayed for a variety of reasons unrelated to community capacity.
- Capacity in the model is static over the model runtime – in reality, flexing of capacity occurs, for example spot-purchasing of beds, or reduced workforce impacting available capacity.
- Mean lengths of stay are stable over the model runtime – in reality, mean lengths of stay in intermediate care pathways vary for a number of reasons, such as changes in social care capacity, involving families in care, and flexing of capacity.
- We recommend using the model in the first instance to help you to structure your understanding of your own system and available data sources (which may come from different organisations, and require working with analysts across organisations). One way of gaining some confidence in your understanding of what the model can and can not tell you, is to run the model using historical arrivals and compare the outputs with historical outputs.
- Figure 6 is an example from BNNSSG where the model is run using historical arrivals data and the outputs (delayed acute discharges)

are compared with historical delayed acute discharges over the same time period.

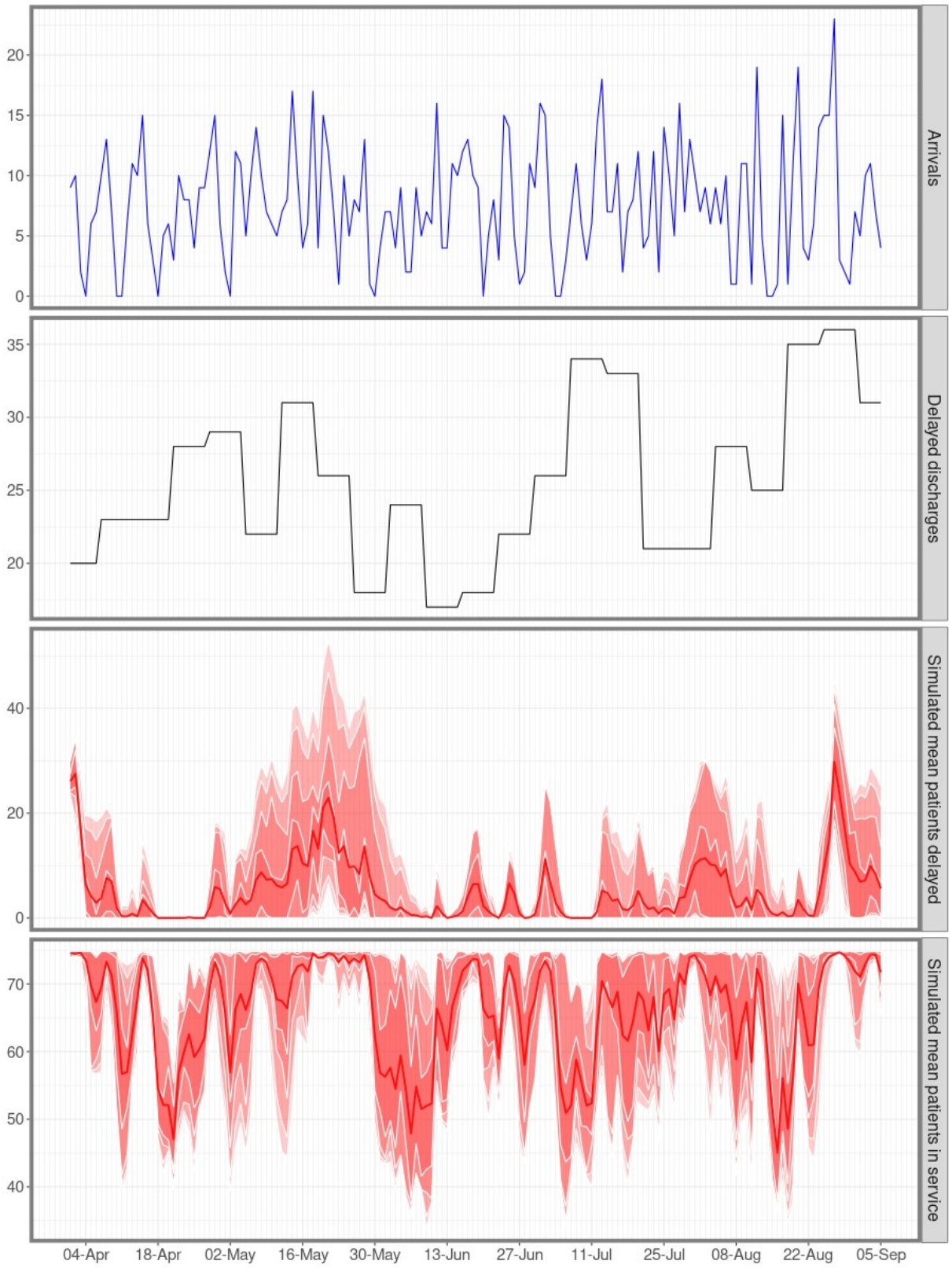

Figure 6: Sample validation plot

The plot in Figure 6 shows historical arrivals from April to September 2022 for one pathway locality in the top facet. The second facet is historical delayed discharges over the same time period (the data for this was presented as a weekly summary). The third and fourth facets are simulation outputs: simulated delayed discharges and simulated patients in service (both plotted using stochastic outputs).

Some consideration needs to be given to interpreting these outputs.

We know that the service operates at 100% occupancy much of the time, and we see that in the simulated outputs.

We can compare actual historical delayed discharges with the simulation outputs. The overall pattern matches reasonably well. In July, actual delayed discharges rose significantly. From exploring the data made available, we were able to see that LoS increased significantly at that time (well above the mean used to run the model) and then was brought back down quite low in July (well below the mean used to run the model). These can account for some of the discrepancies.

There are likely to be others that can be considered when attempting to understand how this information may be useful. Once an understanding and confidence is gained in the baseline case, comparing the effects of different interventions on the baseline case, and between interventions, provides information which can confidently be used to support planning decisions for resourcing pathways.

In summary:

- The IPACS model enables planning and resourcing of community capacity pathways to support decision across the acute-community NHS interface.
- It supports system-wide thinking and encourages collaborative working across services for strategic planning.
- The model is free and open source to use, re-use, adapt and share as required.