V-KEMS Study Group Report

Reducing the Risk of Covid-19 Transmission on Trains

Report Authors and Contributors:

Simone Appella (University of Bath), Eldad Avital (Queen Mary University of London), Tosin Babasola (University of Bath), Thanos Bantis (Connected Places Catapult), Alexandra Brintrup (University of Cambridge), Chris Budd OBE (University of Bath), Radu Cimpeanu (University of Warwick), Philippe De Wilde (University of Kent), Jess Enright (University of Glasgow), Akshay Gupte (University of Edinburgh), Peter Hicks (OpenTrainTimes Ltd.), Andrew Lacey (Heriot-Watt University), William Lee (University of Huddersfield), Antonio Martinez-Sykora (University of Southampton), Amjad Natour (RDG), Simon Noel (QinetiQ Ltd.), Sophie Peachey (Iotics), Eileen Russell (University of Bath), Dipak Sarker (University of Brighton), Lars Schewe (University of Edinburgh), Yang Zhou (University of Bath)

Our thanks also to useful contributions from:

Jonathan Bridgewood (FirstGroup Rail), Ben Ford (Network Rail), Matt Hunt (RSSB)

WARNING: this report contains preliminary findings that have not been peer reviewed. The findings are intended to provoke further study and policy discussion and should not be treated as definitive scientific advice in response to the SARS-CoV-2 epidemic.

Whilst we expect these principles to help others formulate coherent and consistent guidelines, time has prevented any quantitative study of their effectiveness. This could be undertaken, but would require real data and time to build more detailed simulation tools. Thus, we are not able to make specific recommendations from the principles, e.g. we cannot infer that it is safe to do X if you follow principle Y.

Additionally, this report has been assembled in a short time frame, we have made every effort to ensure references and links are present. Where this is not the case, we apologies for the unintentional oversight.

Contents

1 Executive Summary

This report details the progress of a Virtual Study Group held between V-KEMS and a group of researchers from across the UK between the 11 - 13 January 2021. Three challenges were set in context by FirstGroup, RSSB and Network Rail. The ultimate aim of the session was to develop approaches which could be used to assess the risk of Covid-19 transmission on the UK rail network.

Carriage-level modelling: Section 3 outlines the discussion and progress made when considering Covid-19 transmission risk at the carriage-level scale.

1. High Level modelling.
    - Lower dimensional models can integrate a range of different aspects of the problem such as different infection mechanisms (airborne vs contact), different behaviours (breathing / talking / singing) in a simple framework.
    - They have low computational demands making them suitable for implementation as simple risk management tools, specialised to a rail carriage environment.
    - Their key present disadvantage is that they do not account for spatial variations which would enable some train carriage-specific features, however extensions can be derived and readily implemented.
    - Integration of coupling with specific air flows (be they approximated or high fidelity CFD calculations) will significantly improve the accuracy of these types of models and ensure that some of the assumptions are not too general.
2. Simplified flow modelling
    - Full CFD modelling is computationally expensive. However, a simplified model based on inviscid flow and simple sources can allow spatial aspects of the problem to be modelled.
    - This may provide a route for different carriage configurations to be modelled and allow rapid rough simulations which could be useful for "what if?" scenario testing.
    - This simple flow modelling can be combined with the high level modelling discussed above to relax the assumption of spatial uniformity.
3. Different infection modes
    - While a broad consensus has emerged that in indoor environments the greater risk is airborne transmission it would be good to assess this specifically in the case of carriages.

- If fomite infection turned out to be a significant contributor to risk then the simplified flow modelling could be used to model droplet trajectories.
- This could be used to guide other interventions such as surface agents indicating the presence of proxies for contamination.

Journey-level modelling: Section 4 explores how different modelling approaches could be used to understand, thus mitigate, possible transmission as passengers a) enter a train and b) leave a train.

1. Getting on a train
    - Modelling passenger flow onto a train in a continuous manner can be used to assess various load strategies and crowding from a platform to a train door.
    - Future work could envisage and model more realistic conditions for initial platform crowd distribution.
    - Additionally, it could be instructive to modify the equations of motion (with observed data) to better capture the characteristics of the passenger population.
2. Getting off a train
    - The simple viral load queue model shown here is worth exploring more, as queue systems are very common not only in transport, but in many setting.

Network-level modelling: In Section 5, a modelling methodology is introduced which could be developed to optimise rail capacity when managing large scale, network-level operations,

- This report is a starting point for developing a set of models to advise train operators and (in this case) HE providers to organise the return of students to campus. Similarly, it could be used to plan movements for large events, half-term holidays etc.
- For the HE example, additional data that allows us to estimate the number of students that would use a particular station is discussed. This can provide the basis of an optimisation model to adjust the arrival of students to alleviate the stress on the rail network.
- Such a model could also be used on a regional level to plan commuter traffic. It might also be of interest to investigate where additional trains might be required to deal with unusual demand.

2 Background

Passenger use of the UK's rail network fell as low as 4 % in late April 2020 (as compared with an equivalent day in 2019), and it is currently operating at around 15 % (late January 2021). The rail industry has been working hard to put plans into place to attract passengers back to rail following the end of lockdown periods; in order to instil confidence in the public, it is vital to develop methods that can quantify the risk of transmission on a train in numerous operating circumstances.

The Rail Safety and Standards Board reports that the risk of Covid-19 infection is less than 0.01 % on an average journey (this does not account for new strains of the virus and was calculated in August 2020). This is equivalent to a chance of less than 0.01 %, based on an hour-long train journey in a carriage with no social distancing or face coverings. The report also shows that the risk more than halves if passengers wear a face covering.

The current perspective of rail as unsafe, especially for the vulnerable, also disproportionately affects marginalised groups such as people with disabilities and can cause social isolation, limited access to employment and life opportunities. It is crucial that we keep inclusion in mind throughout Covid-19 recovery efforts to ensure that the changes we may see into the future enrich travel for all.

The industry looks to develop long-term resilience through innovation and collaboration which benefits and safeguards the rail industry and staff roles, meets passenger needs, reengages with customers (and new markets) as well as boosting economic recovery through infrastructure development and innovation. A project known as TRACK (Transport Risk Assessment for COVID Knowledge) is looking to create models that will quantify the level of risk faced by passengers and transport staff. This work will involve detailed simulations of the way the virus could potentially spread through airflow, from touching contaminated surfaces and being close, to someone infected with the virus.

There are currently many different operating scenarios which trains are used under, and it is not feasible nor sensible to simulate them all. Mathematical science can provide tools for how to explore such a large possibility space, and help focus the use for detailed simulations.

- Scale 1: Modelling of a carriage: controls e.g. airflow, seating design, passenger allocation to seats
- Scale 2: Modelling of a journey: static, dynamic, passenger movement, passengers getting on and off
- Scale 3: Scheduling across the country, passenger allocation to trains; High times of rail usage (e.g. start of university term); Resilience of schedule based on outbreak

3 Scale 1: Modelling of a Carriage

3.1 Challenge Description

In this section, we consider the transmission and possible actions available to reduce transmission risk at a carriage level. Transmission in a carriage could occur from surface, droplet and aerosol spread. Train operators can take mitigating actions in the form of seating design, heating ventilation & air conditioning (HVAC) airflow control, cleaning strategies, decreasing the use of lavatories and food outlets etc. An important part of the challenge here is to **complement rather than replicate** the important research under way on the Transport Risk Assessment for COVID Knowledge (TRACK) programme.

We begin in section 3.2 by sketching out a high level model illustrating the possible transmission routes and use this to focus our investigation and thus choose and develop fruitful areas to add complexity. The high level model is built around the following question: ***if a susceptible and infected person are on a train what is the chance that the susceptible person becomes infected and by what mechanism?***.

In section 3.3 we carry out a survey of the available literature focussing on what is known about the transmission of Covid-19, models of infection transmission and of aspects of human behaviour that promote transmission, most notably via speech. The specialisation of these models to a railway carriage is considered.

A key weakness of many of the transmission models discussed in the literature review is that they assume a well mixed environment. While such modelling is valuable, it does not allow us to investigate questions regarding, for instance, passenger allocation to seats. In section 3.4 we discuss how this assumption can be relaxed without resorting to numerically intensive full computational fluid dynamics modelling. A quick consideration shows that the dominant mode of particle transport which leads us to considering simple ways to estimate flow streamlines. An approach based on simplified geometry and potential flow is outlined.

While the literature reveals a growing consensus that airborne transmission dominates contact transmission this question is still not settled and it remains a possibility that this may be an important transmission mechanism in railway carriages. In section 3.5 we discuss possible interventions that could reveal potentially contaminated surfaces with the aim of alerting either passengers or cleanup crews.

3.2 Sketch of a high level model.

In order to focus the literature review and subsequent modelling a high level model of Covid-19 transmission on a train was sketched. Fig. 1 shows the key elements of the environment and the possible transmission routes. The environment in the carriage at the highest level can be separated into the airspace and surfaces. In turn the airspace is connected to the the surfaces and also the HVAC unit and the outside air. A susceptible passenger may be infected by Covid-19 either by inhalation of air containing infected droplets/aerosols or via touching a contaminated surface.

Fig. 2 shows airflows in the system. Air is inhaled and exhaled by susceptible and infected passengers, and circulated by the HVAC unit. Air from outside is exchanged with air from the carriage when doors are opened and is also exchanged with the HVAC unit.

Fig. 3 shows possible routes for contaminated aerosol transport. As aerosols persist in the air these may circulate through the HVAC (where contamination could be reduced by ultraviolet irradiation for instance).

Fig. 4 shows possible droplet infection routes. As droplets roughly follow ballistic trajectories there is less likelihood they will be drawn into the HVAC unit, but the possibility of them contaminating surfaces may be important.

Human behaviour plays in important part in the transmission of Covid-19. This suggests that it may be important to model the states of passengers during a journey, particularly whether they are wearing a face covering or not and whether they are talking, phoning or quiet. This is summarised in Fig. 15.

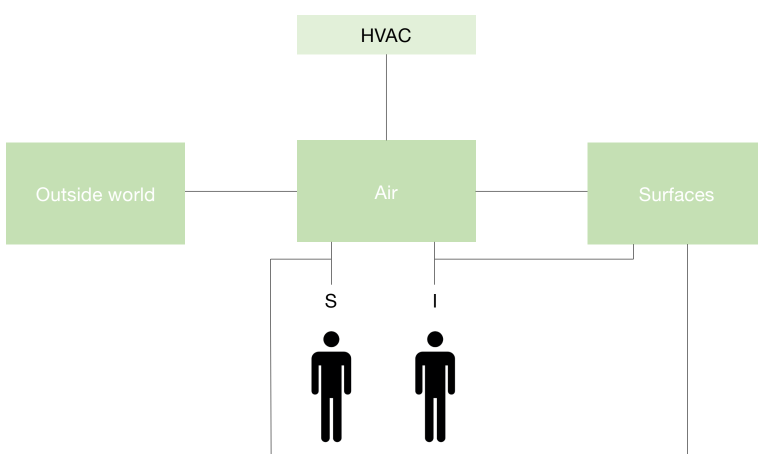


Figure 1: High level model of infection routes in a carriage.

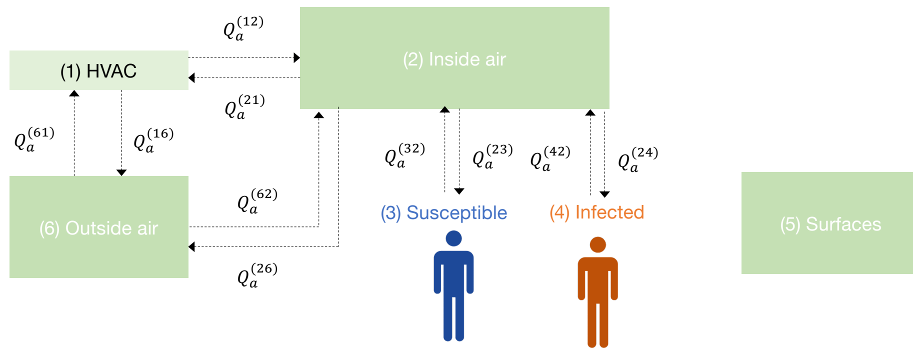


Figure 2: Airflows in a carriage.

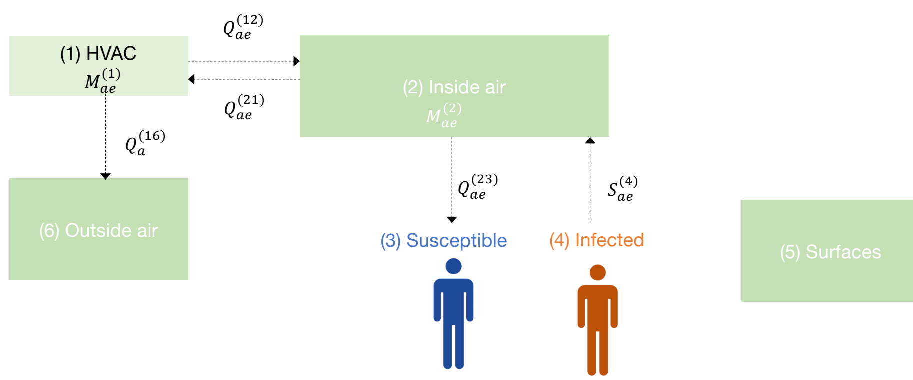


Figure 3: Transmission routes (aerosol only). $Q_a^{(12)}$ depends on ultraviolet C (UVC) in HVAC. $S_a^{(4)}$ depends on state; mask, phone etc.

3.3 Literature considerations

Before delving into the modelling specifics, a brief literature review of recent publications and preprints (see subsection 4.5 on latest information regarding transmission channels (loosely classified into three categories: 'large-drop', 'airborne' and 'contact/surface') revealed the following noteworthy findings (as extracted from Ref. (3) and not included references from original article):

- There is growing evidence that indoor airborne transmission associated with relatively small, micron-scale aerosol droplets plays a dominant role in the spread of Covid-19, especially for so-called "super-spreading events", which invariably occur indoors.

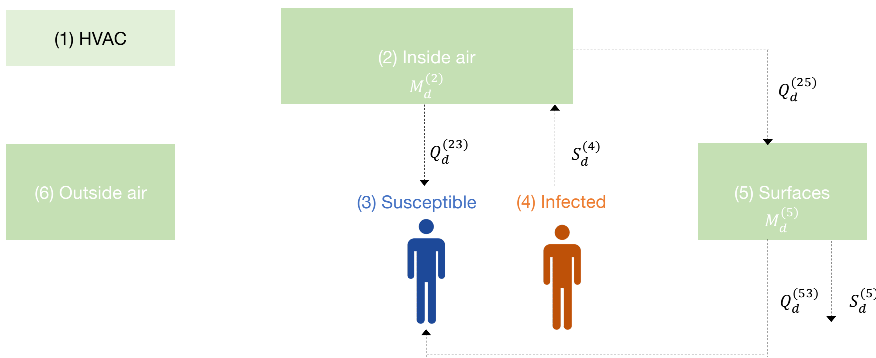


Figure 4: Transmission routes (droplets only). $S_d^{(4)}$ depends on state; mask, phone etc. $S_d^{(5)}$ depends on cleaning, surface treatment etc.

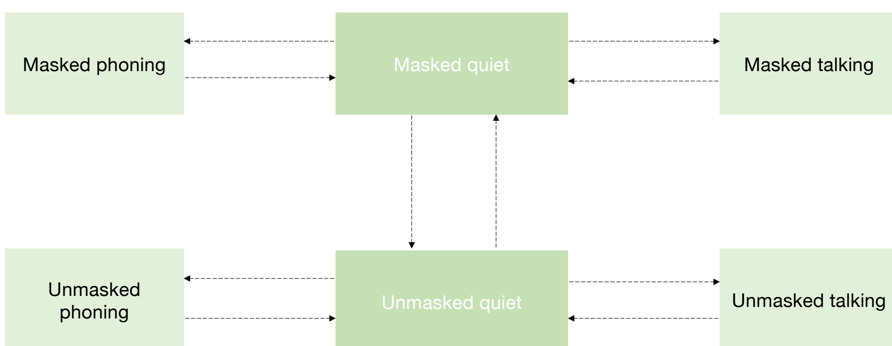


Figure 5: State model for an infected person

- Further evidence for the dominance of indoor airborne transmission has come from a recent analysis of 7324 early cases outside the Hubei Province, in 320 cities across mainland China. The authors found that all clusters of three or more cases occurred indoors, 80 % arising inside apartment homes and 34. % potentially involving public transportation. Only a single transmission was recorded outdoors.

- The evaporation time at 50 % relative humidity (RH) ranges from τ_e = 1.2 ms for r_0 = 0.5 μm to 12 s at 50 μm. These inferences are consistent with experiments demonstrating that stable respiratory aerosol distributions in the range r_{eq} < 10 μm are reached within 0.8 s of exhalation.

- Moreover, a recent experimental study of the dependence of droplet size on the infec-

> tiousness of SARS-CoV-2 virions concluded that droplets with r > 2 μm are less infectious, an inference that would underscore the importance of airborne transmission. Consistent with this finding, the viability of human influenza viruses in aerosols has been shown to be maximized at low relative humidity.

This provides an impetus to place indoor airborne modelling at the center of our modelling approach, while keeping other sources of contamination in mind (also connected to the other tasks here) given continuous updates in information from many ongoing dedicated studies. We note that relevant studies appear in the literature at a very high rate, with for example a mixed experimental-computational investigation on disease transmission inside an urban bus (18) being published during this Virtual Study Group (VSG) - with a clear message in terms of the benefits of ventilation, window and door openings to facilitate air exchanges for the purification of the enclosed environment.

3.3.1 The Wells-Riley model

The classical Wells-Riley model for indoor airborne transmission introduces the quantum of infection as an infectious dose unit to base a quantitative model on - see Ref. (15). This relates to the number of infectious airborne particles required to infect a person and has considerable variability in terms of both disease and individual immunity. A key assumption of this model is that the particles are assumed to be randomly distributed ('well-mixed' scenario) inside our confined space. Riley et al. (14) considered the intake dose of airborne pathogens in terms of the number of quanta to evaluate the probability of escaping the infection. Together with the Poisson probability distribution describing the randomly distributed discrete infectious particles in the air, the Wells–Riley equation was derived as follows:

$$P_I = \frac{C}{S} = 1 - \exp\left(-\frac{Iqpt}{Q}\right),$$

with the following relevant parameters:

- P_I is the probability of infection,
- C is the number of infection cases,
- S is the number of susceptible individuals,
- I is the number of infectors,

- p is the pulmonary ventilation rate of a person,
- q is the quanta generation rate,
- t is the exposure time interval,
- Q is the room ventilation rate with clean air.

Interestingly, q also represents one of the primary unknowns in this model and requires revisiting once other parameters are estimated and more epidemiological information is known. The other parameters offer an intuitive glimpse into infection likelihood increasing (number of infectors, exposure time) and decreasing (room ventilation rate) factors.

This approach can be interpreted as a reduction of the well-known family of SIR (or SEIR) models (see (12)). In such cases ordinary differential equation (ODE) systems describe the interplay between susceptible (S), infected (I) and recovered (R) populations. Following the supplementary material in (3), one can interpret the Wells-Riley model as effectively a reduction of the SIR model based on the assumption of slow incubation over the timescale of the event. This would be appropriate for small time windows such as a choir practice or a train journey, but the assumption no longer holds over prolonged periods. It is however a useful interpretation within our scope and time-dependent infector extensions are readily included in this formulation. In fact, in the following subsection we describe one of the most recent models which includes several relevant generalisations.

3.3.2 The Bazant-Bush generalisation

Built on the theoretical basis as the Wells-Riley model, this approach focuses on the characterisation of a concentration $C(r,t)$ of pathogen transported by drops of radius r as a function of time. This still happens within a well-mixed environment, without considering coupling or geometrical features.

The interested reader is referred to the preprint and its supplementary material for the numerous details, however some of the most important building blocks and assumptions are listed below:

1. $I(t)$ infectious individuals are assumed to exhale pathogen-laden droplets of radius r at constant rate $Q_b n_d(r) V_d(r) p_m(r) c_v(r)$, with Q_b the breathing flow rate (exhaled volume per time), $n_d(r)$ the number density of drops of radius r (or drop size distribution), $V_d(r) = \frac{4}{3}\pi r^3$ the drop volume, $0 < p_m(r) < 1$ the mask penetration factor and $c_v(r)$ the

microscopic pathogen concentration. All these variables should be extracted and constantly updated with advancements in both technology and knowledge about viral mutations.

2. $v_s(r)$ is the particle settling speed. If assuming a Stokes-type model, drops of radius $r \leq 100$ μm and density ρ_d settle in a quiescent environment (with density ρ_a and dynamic viscosity μ_a) with $v_s(r) = \frac{2}{9}(\rho_d - \rho_a)gr^2/\mu_a$, given by the balance between gravity and viscous drag. Naturally, an environment with strong background flows, large scale movement or ventilation would require an update of this simple formulation.

3. $\lambda_v(r)$ is the rate at which virions become non-infectious (or deactivated), a property which is both droplet content and environment-dependent. Non-Newtonian and/or chemical properties are likely to enter here. This may also be a viable entry point for mitigation strategies against contaminants.

4. The enclosed space is characterised by floor area A and height H, such that the control volume is $V = AH$.

5. Q represents the ventilation outflow rate, which helps construct the air exchange rate $\lambda_a = Q/V$.

6. If mechanical ventilation is present, Q_r represents the recirculation flow rate. This is a new model addition that may be helpful in including some of the ventilation-specific aspects of the carriage to a first approximation.

7. $p_f(r)$ is the probability of droplet filtration.

While comprehensive and information-rich, all of these parameters can be either extracted from the specialised literature or, in the case of some probabilities, assumed to follow typical behaviours based on either Covid-19 specific data (where known) or historical data on related diseases.

Armed with this machinery, the main equation in the Bazant-Bush model is the concentration equation

$$V\left(\frac{\partial C}{\partial t} + \lambda_v(r)C\right) = I(t)P(r) - (Q + p_f(r)Q_r + v_s(r)A)C.$$

There are at least two relevant limiting cases:

- if $\lambda_v = v_s = Q_r = 0$ (ignoring deactivation and filtration effects), the model is reduced to the previously discussed Wells-Riley model, which provides a solid validation of the methodology.
- if $\lambda_v = P = Q = Q_r = 0$, the dynamics reduces to sedimentation models for well-mixed environments, another useful link to a large body of knowledge.

With sufficiently simple initial conditions (e.g. one infected individual entering the room at time t = 0), closed form solutions solutions to this equation can be found. Otherwise, with more complex modelling ingredients, numerical solutions are needed, however readily obtained given the simplicity of the partial differential equation (PDE) to be solved.

Once the concentration is calculated, the **airborne transmission rate** is defined as

$$\beta_a(t) = Q_b \int_0^\infty C(r,t) p_m(r) c_i(r) dr,$$

where we recall Q_b to be the breathing flow rate and the usage of masks and size-dependent droplet infectivity taken into account via p_m and c_i, respectively.

The transmission rate can then be linked to the reproduction number R, which may be appropriately restricted to within the desired degree inside a carriage/journey in order to generate quantitatively-backed indoor-safety guidelines. These should be informed in a cross-disciplinary manner, however useful general starting points are provided by the authors of the study and subsequently specialised to Covid-19 transmission. An additional layer of detail is required to convert these measures to rail-specific contexts. While constructed from a high-level viewpoint and with significant reduced-order modelling in place, this framework provides a transparent and easily customisable mathematical approach towards analysing indoor transmission. A risk-management tool [1] (and associated spreadsheet [2] is provided to test the predictive capabilities of this model.

Other models of airborne transmission based on the quanta approach include the Gammaitoni-Nucci model (7) used for instance in Ref. (2).

3.3.3 Viable extensions for rail journeys

Possible generalisations for train geometries, ventilation systems and journey types:

[1] `https://indoor-covid-safety.herokuapp.com/`

[2] `https://cheme.mit.edu/wp-content/uploads/2020/11/COVID-19_Indoor_Safety_Guideline_v5.xlsx`

- the 'well-mixed' assumption is one of the weak spots of this framework. Particularly with specific ventilation systems in different carriages, ventilation outflow rates Q and recirculation flow rates Q_r can be spatially localised, leading either to a box-type model with different regions or to a generalisation for a full concentration equation on $C(r, \vec{x}, t)$, allowing for the incorporation of $\vec{x}$-dependent terms for certain desired features. The advantages are increased fidelity, while the drawbacks consist in removing some of the simplicity of closed-form formulae that was previously available. Depending on resolution, it is not anticipated however that such terms will be computationally heavy, depending on the intended usage of the tool.
- some of the modelling assumptions benefit from added complexity in view of the richer background flow. The settling and evaporation dynamics may well be affected by the details of the ventilation inside the carriage for example.
- inclusion of barriers (zero concentration volumes) for specific geometrical features, from static seating, luggage racks etc., to moving entities such as passengers (which again will require further coupling with a fluid model).
- time-dependent inlet/outlet conditions modelling carriage doors opening/closing for various conditions and durations.
- specialisation of parameters towards specific temperature/humidity conditions as implemented on various train models.
- while such a model (even with extensions) may be useful even in isolation, an improved coupling with other transmission models (e.g. surface contamination) will provide a more accurate depiction of the system-level risk factors.

There are numerous parameters involved and this will ultimately become ever more complicated with other extensions (see below), however a customisable framework with dedicated functionality for rail journeys based on the specialisation above may provide a viable mechanism for safety assessments. Depending on the desired level of sophistication and dedicated team size, it could be in principle developed on a timescale of weeks and act as a *what if?* platform for a series of relevant measures both in the short-term for Covid-19 and in the long-term for future pandemic events and other forms of risk mitigation.

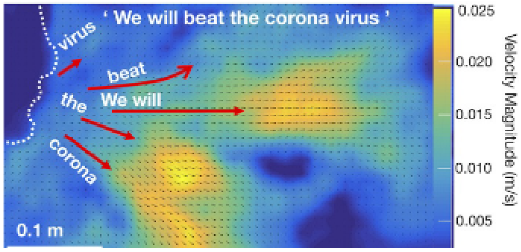
3.3.4 Notes regarding human behaviour

Many of the factors included inside the models discussed above are often concentrated into a single constant or function. It is thus easy to gloss over some nuance and a number of detailed results in recent studies that provide much needed information on some of the human-centric aspects of actions undertaken during a train journey. Below we focus on speech patterns and measurements, however recent studies e.g. on movement in confined spaces (10) are also beginning to provide additional insight into some of the finer aspects of the complex system:

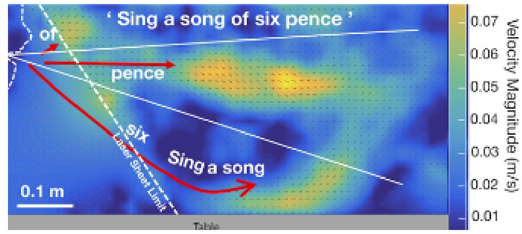
- Not all speech causes the same viral transmission levels. We also note that the stop consonants, or what are referred by linguists as plosives consonants, such as (P, B, K,...), have been demonstrated recently to produce more droplets. In these cases, the vocal tract is blocked temporarily either with the lips (P, B) or with the tongue tip (T, D) or body (K, G), so that the pressure builds up slightly and then is released rapidly, producing the characteristic burst of air of these sounds; in contrast, fricatives are produced by partial occlusion impeding but not blocking airflow from the vocal tract (1). Note the large contrast in the color scales as well in the figure below, covering an order of magnitude in measured air velocity as a result of emitted sounds. Anecdotal evidence from the authors (via a seminar in November 2020) also indicates that the usage of products such as lip balms and generally moisturisers lead to a smoother air flow around the lips and reduced ejection of droplets. Fig. 6
- The importance of the activity, from the nature of the sounds to the volume has been comprehensively discussed by other authors as well, with the key conclusion that any significant disruptions (singing, especially loudly or in a prolonged manner, or shouting) are very influential in increasing infectivity in a closed environment. While speaking does carry some risk, this is arguably of manageable in the context of the range of the studied activities. (3). Fig. 7.
- This can be further contextualised in terms of other well-known spreading events in the recent year, with striking effects. (3). Fig. 8

The above studies provide controlled evidence that minimisation of speech while in trains (via the use of mobile phones or through interaction without other passengers and staff) is desirable. Simple solutions may consist in:

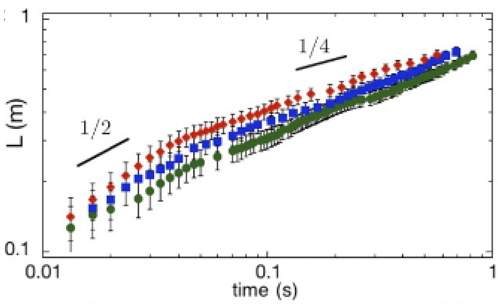
1. providing virtual (app-like) environments for customers to interact with catering services.
2. more generally, ensuring that text-/visual- rather than voice-related methods are implemented for as many services as possible.



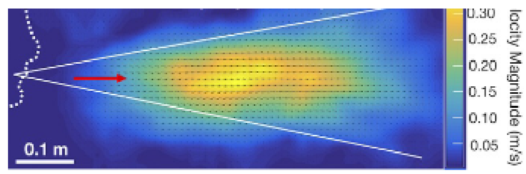
(a) "We will beat the corona virus," which is a mixture of vowels, fricatives and plosives.



(b) Sing a sing of six pence (SSPP) (25), mainly composed of the fricative 'S' except the last word that starts with 'P'.



(c) The distance traveled by the extremity of the air puff as a function of time when saying 'pence' at the end of SSPP for three different runs.



(d) 'Peter Piper picked a peck' (PPPP) (25), which is mainly composed of mainly plosives P.

Figure 6: Adapted from Ref. (1), see original paper for references. . Mean velocity field produced when speaking three different sentences. A colour code illustrates the average speeds but note that single images of the magnitude of speeds are not representative of the true instantaneous velocities.

3. extending the above to in-train announcements (recorded and transmitted via apps/text notifications) such that frequent and loud background noises do not encourage rail users to raise their volumes to get their messages across the room even if speaking.

4. minimising movement inside confined spaces as much as possible.

5. ultimately prolonged exposure in all these frameworks leads to additional risks, a situation which will of course deteriorate with the usage of inadequate mitigation strategies, from ventilation systems to inadequate face masks. If transient features such as getting on and off trains can be controlled, this may be an argument for breaking longer journeys into more legs.

Such suggested measures will contribute to a less interactive experience, however carry the advantage of minimising contacts and conditions for riskier behaviours. It should also be noted that all of the above represent modelling-based idealistic scenarios and are underpinned by assumptions of **compliance to some of the more basic actions**, such as mask-wearing

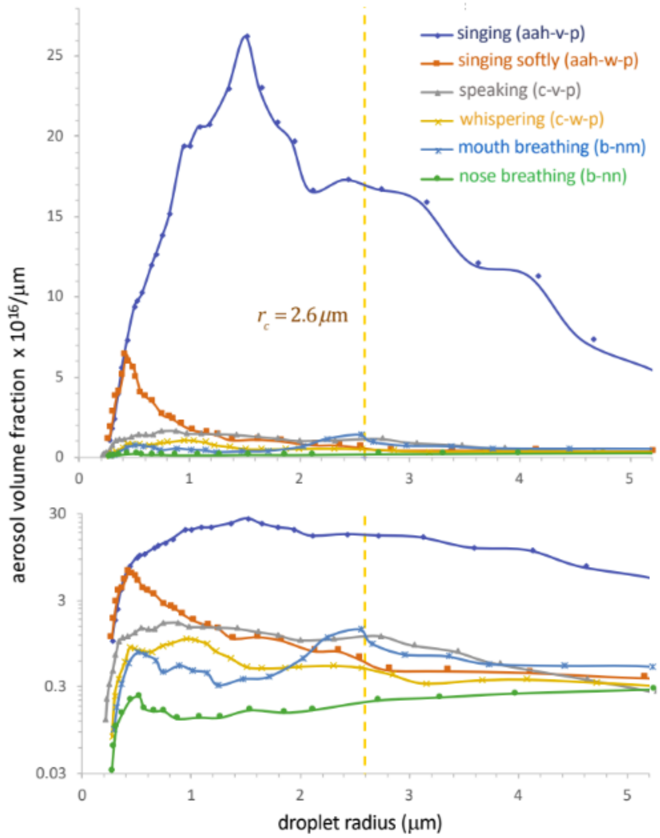


Figure 7: Adapted from Ref. (3), see original paper for references. . Model predictions for the steady-state, droplet-radius-resolved aerosol volume fraction, $\phi_s(r)$, produced by a single infectious person in a well-mixed room. The model accounts for the effects of ventilation, pathogen deactivation and droplet settling for several different types of respiration in the absence of face masks (p_m = 1). The ambient conditions are taken to be those of the Skagit Valley Chorale super-spreading incident (22, 24) (H = 4.5 m, A = 180 m^2, λ_a = 0.65 h$^-1$, r_c = 2.6 μm, λ_v = 0.3 h$^-1$, RH = 50 %). The expiratory droplet size distributions are computed from the data of Morawska et al at RH = 59.4 % (9) (see their Fig. 3) for aerosol concentration per log-diameter, using $n_d(r) = (dC/d\log D)/(r\ln 10)$. The breathing flow rate is assumed to be 0.5 m^3/h for nose and mouth breathing, 0.75 m^3/h for whispering and speaking, and 1.0 m^3/h for singing.

and distancing where applicable. Based on the literature surveyed, departures even over small timescales from some of these norms have severe consequences that far outweigh some of the optimisation aspects highlighted above. This is arguably extended to many of the tasks both in the present groups and at other granular levels - ensuring social responsibility and adherence to rules, from the simpler to the more complex is likely to generate the most beneficial outcomes at the system level.

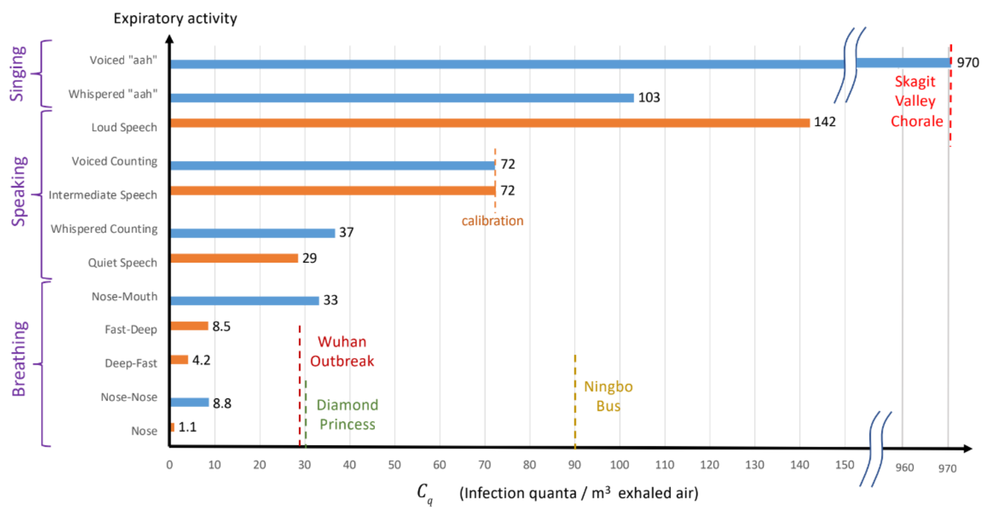


Figure 8: Adapted from Ref. (3), see original paper for references. Estimates of the "infectiousness" of exhaled air, C_q, defined as the peak concentration of COVID-19 infection quanta in the breath of an infected person, for various respiratory activities. Values are deduced from the drop-size distributions reported by Morawska et al (9) (blue bars) and Asadi et al (31) (orange bars). The only value reported in the epidemiological literature C_q = 970 quanta / m^3, was estimated (22) for the Skagit Valley Chorale super-spreading event (24). This value is rescaled by the predicted infectious aerosol volume fractions, $\phi_1(r) = \int_0^{r_C} \phi_s(r)dr$, obtained by integrating the steady-state size distributions reported in Fig. 1 for different expiratory activities (9). Aerosol volume fractions calculated for various respiratory activities from Fig. 5 of Asadi et al (31) are rescaled so that the value of C_q = 72 quanta / m^3 for "intermediate speaking" matches that inferred from Morawska et al (9) for "voiced counting". Estimates of C_q for the outbreaks during the quarantine period of the Diamond Princess (23) and the Ningbo bus journey (25), as well as the initial outbreak in Wuhan City (2, 73) are also shown (see supporting information for details)

3.4 Simplified Modelling of Airflow

As shown by the literature review very sophisticated modelling frameworks exist for modelling the spread of airborne infections in well mixed systems. However, the well mixed assumption will allow the consideration of important questions with a spatial element such as optimal seating patterns. We therefore investigate ways of relaxing the well mixed assumption and introducing airflow into our models. Section 3.4.1 considers the relative importance of convective vs diffusive motions and concludes that convective motions dominate. To model convective motions we need to estimate the air velocity field. Accurate but computationally intensive estimates of this can be obtained via CFD. Here however, in Section 3.4.2 we consider a simplified approach based on potential flow allowing airflows to be approximated at a much lower computational cost. Finally in Section 3.4.3 we briefly consider a laterally averaged model in which we integrate out variations perpendicular to the axis of carriage.

3.4.1 Convective and diffusive contributions to motion.

Once the flow field has been found particle motion can be found by post-processing. Approximating the droplets as passive scalar means we focus on the droplets nuclei of less than 20 μ m, neglecting effect of gravity (Froude number) and the time lag reaction with the flow means we assume the Stokes number St_k « 1, so we look at droplets nuclei of less than 10 μm. This is an excellent assumption for TB that is fully airborne and dominated by 5 μm particles. For Covid-19 it captures only part of the transmission. In this case we can use the convection-diffusion equation:

$$\frac{\partial c}{\partial t} + \frac{\partial (cv_i)}{\partial x_i} = \frac{\partial (D\partial c/\partial x_i)}{\partial x_i} + S - Si,$$

where S is the source and Si is a sink. Assuming steady particle motion, single diffusivity D, we get:

$$\frac{\partial (cv_i)}{\partial x_i} = D\frac{\partial^2 c}{\partial x_i^2} + S - Si$$

If we neglect convection effects, and approximate $SiasSi \sim \alpha(x_i)c$ for ventilation, we get:

$$D\frac{\partial^2 c}{\partial x_i^2} + S - \alpha c = 0$$

If we do not know where the vents of the ventilation are we can take α as a constant, depending on known air changes per hour.

Stokes-Einstein diffusion coefficient D:

$$D = \frac{k_B T_f}{3\pi/mu_f d_p}$$

where the Boltzmann constant is k_B = 1.308649 10^{-23} (J/K). The air dynamic viscosity is taken as μ_f = 18 $\times$ 10^{-6} (N s / m) for air temperature of T_f = 290 (K) yields for a droplet nucleolus of diameter $d - p$ = 10 $\times$ 10^{-6} (m), the diffusion D = 2.36 $\times$ 10^{-12} (m^2/s).

However, looking at the Schmidt number $S_c = \mu_f/(\rho D) = \times 10^{-6}/(1.2252.36 \times 10^{-12}) = 6.22 \times 10^6$, means that if we neglect diffusion for the flow field and account only for convection, we should also focus on convection for the fine droplet nuclei.

The particle's relaxation time is $\tau_p = (\rho_p d_p^2)/(18\mu_f)$. Taking the particle's density as of water (it is actually mostly salt and other minerals) ρ_p = 1000 (kg/m^3) gives τ_p = 0.0056 (s) for d_p = 10 μ m. This justifies the assumption of neglecting the time lag between the particle's velocity and the flow velocity, which can also be expressed by taking St_k « 1.

The focus of the convection effects for fine droplet nuclei is illustrated in Fig. 9 for continuous emission of 5 μ m particles as part of long speech. The particles are captured in the thermal plume emitted by the person.

3.4.2 Potential Flow

In support of higher fidelity computational fluid dynamics (CFD) being undertaken in other TRACK work packages, low fidelity fluid models have been identified as a means by which the airflow within the train carriage can be quickly characterised. With 15,000 train carriages in operation in the UK with significant variations between each design, a highly generalisable fluid model that can be quickly modified is desirable.

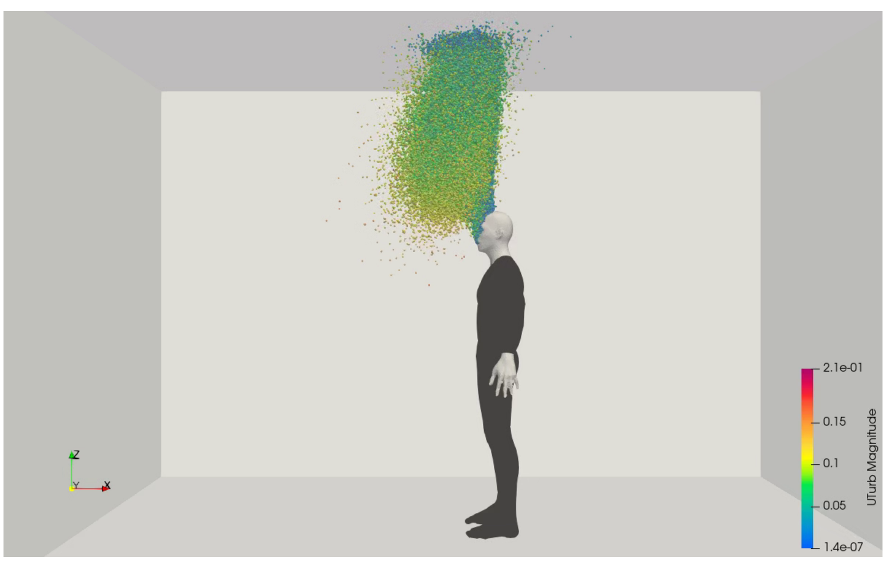


Figure 9: unsteady Reynolds-averaged Navier-Stokes (URANS) Eulerian-Lagrangian simulation of fine particles emission during continuous speech. Evaporation is not accounted as fine droplet nuclei are assumed (5 μm). Image credit: Shuo Mi, 2020

Potential flow theory considers incompressible, inviscid, irrotational flow. It is widely used within the aerospace and aviation industry as a means by which aerodynamic performance can be evaluated quickly. The advantages of this are that it allows the rapid evaluation of different geometries with virtually no downtime for computation. The theory uses several fundamental flows that can be used to build complexity - sources (entry), sinks (exit), uniform flow (global magnitude and direction), and vortices (angular motion). Using the classical method of images it is possible to represent a flow field as a superposition of fundamental elements and incorporate walls into the domain to represent boundaries through which the flow cannot travel. The analysis of a train carriage will be primarily focused on the representation of inlets and outlets as sources and sinks respectively.

In this VSG a potential flow model has been created for an example analysis of a train carriage. A two-dimensional representation of a train carriage is shown in Fig. 10, with inlets / outlets, an open door and a person breathing. It can be seen from the streamlines that the flow entering the domain is quickly pulled towards the exit, shown by the connecting lines between the inlet and outlet. In this case, the effect of the door is negligible so only a small portion of the flow leaves the domain by this means. Fig. 11 shows the same domain with the 'strengths' of the sinks representing the door being much stronger. In this case it can be seen that the flow is pulled further down from the inlet. This example highlights that the effects of different rates of air extraction, air loss due to temperature gradients across the door can be quickly evaluated in the order of minutes.

The 2D potential flow example has been extended to three dimensions for two cases shown

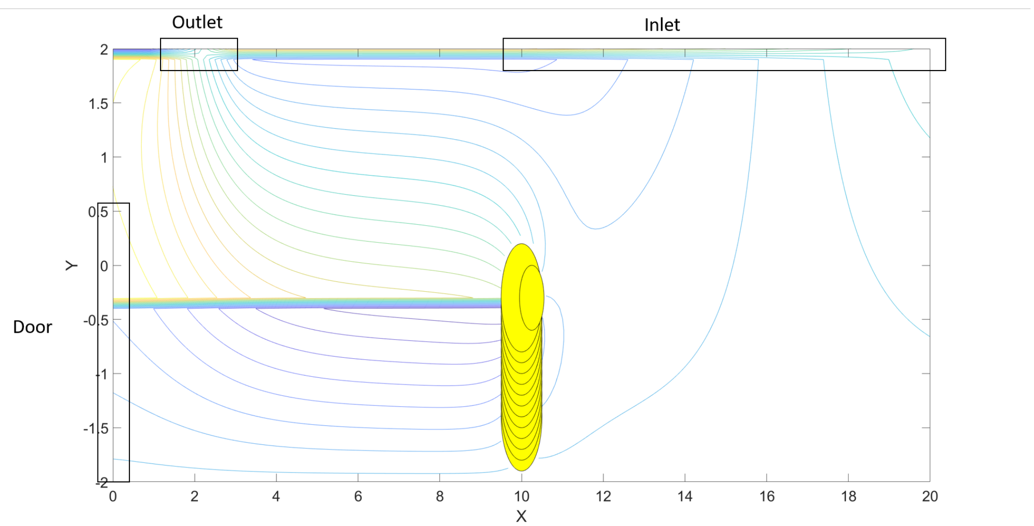


Figure 10: A two-dimensional representation of a train carriage

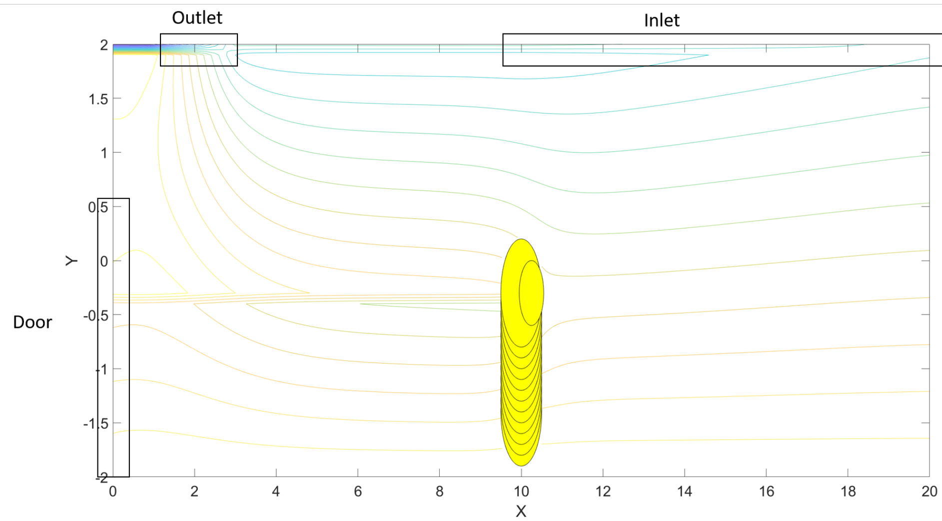


Figure 11: A two-dimensional representation of a train carriage with stronger sinks than Fig. 10

in Figs. 12 - 14, a train carriage with twin HVAC and a door at each end of the carriage, and a central HVAC with three doors on the carriage. In each case the inlet runs the length of the carriage on the roof and the yellow spots show the outlet vents. The curved lines show the streamlines originating at several points in the domain. Whilst these images are for demonstration purposes only, they highlight the regions where the inlets are dominated by the outlet, given by positive gradient on streamlines connecting to the outlet. These insights can then feed the direction of higher fidelity models or provide a simple flowfield that can feed into the methods outlined above or in Task 1.

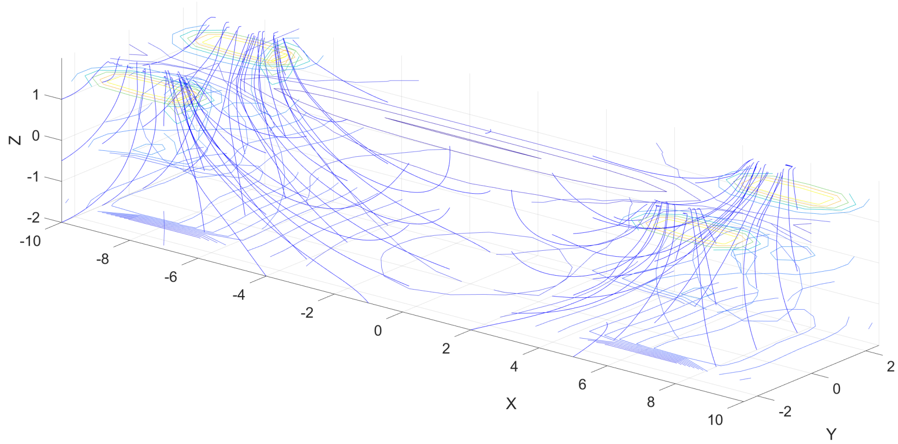


Figure 12: Streamlines and velocity contours for end HVACs with 2 doors

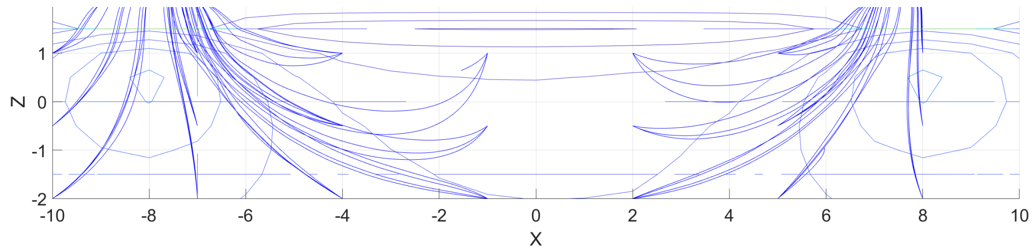


Figure 13: Streamlines and velocity contours for end HVACs with 2 doors

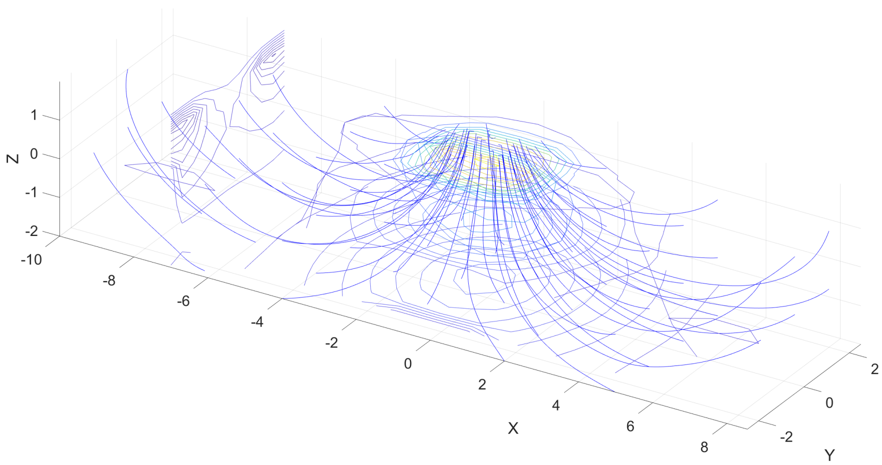


Figure 14: Streamlines and velocity contours for central HVACs with 3 doors

3.4.3 Quasi one dimensional modelling

We model the carriage as a duct having a variable cross-section A in space and time (allowing movement of people), and porous walls (e.g. windows). Quasi one-dimensional fluid dynamics equations were written for the flow continuity, momentum and concentration of pathogens C. Sources/sinks are added to represent the effects of vents, windows, pathogens generation and removal. Those pathogens sources/sinks can be approximately related to the air volume sources and also linearly to the pathogens concentration using semi-empirical constants.

$$\frac{\partial A}{\partial t} + \frac{\partial (Au)}{\partial x} = q_v$$

$$\frac{\partial (\rho A u)}{\partial t} + \frac{\partial \left[A\left(\rho u^2 + p\right)\right]}{\partial x} = p\frac{\partial A}{\partial x} + f_{\mathrm{visc}}$$

$$\frac{\partial (CA)}{\partial t} + \frac{\partial (CAu)}{\partial x} = q_C$$

3.5 Surface contamination

While the academic consensus is that aerosol and droplet transmission mechanisms of Covid-19 dominate this is still an open question and it may well be an important transmission mechanism in some environments. Thus in this section we present some ideas for representation of surface contamination.

One possibility is to incorporate active and indicative agents in the sprays and fogging agents used to clean the carriage. These may act:

1. to aid the disinfection process by using soap-based nanocarriers[3] that release "soap" and "the measuring / reagent substrate";
2. to reveal the presence of residues left on seating and handles by human contact namely:
   (a) pH - sweat (pH $\sim$ 6.3) and the mucous from the throat and nose (pH 5.5.-6.5) of COVID-carriers. Incidentally, bacterial-infected mucous or catarrh has a much higher pH 7.2-8.3 and much higher, less readily ejected viscosity;
   (b) body fluids also contain salt, protein, formic acid, metabolites, etc.
   - pH indicator dyes - litmus, universal indicator, phenol red, etc are non-toxic but could highlight via staining areas contacted by body fluids and human contact thus aiding better and safer cleaning
   - Mucous (a glycoprotein (13) sensitive “peptide” dyes (ninhydrin, copper (I/II) solutions, Folin reagent, etc), again key off amine / phenolic / carboxylate functional groups in molecules associated with the common components (mainly protein and peptide) associated with human body fluids but NOT carriage textiles and hard furnishings.
3. Highlighting contamination could serve to:
   - Improve the sophistication of CFD models but considering aerosilisation of fomites and the transmissibility of deposited viral contamination (larger drops in coughs and sneezes or smears from hands) by virtue of the vehicle it is deposited within;
   - Simply aid better compartmental clean-down by highlighting where most attention should be put into the cleaning regime.

[3]Nanocarrier / micelle components are commercial soaps such as SDS. Simple reagents are inorganic mineral salts (copper salts), dyes and pigments - all are cheap, abundant and non-toxic.

3.6 Conclusions and next steps

1. There is scope for low dimensional, low fidelity models to complement the high dimensional, high fidelity models under development as part of the TRACK project.

2. Combining these models with simple flow models will allow spatial and temporal variations to be included without a dramatic increase in (computational) complexity.

3. As these simpler models are much faster to run, they can be particularly useful for "What If?" scenario testing e.g. for estimating the effects of various interventions.

4. The implementation could take the form of a simple dashboard or spreadsheet on the vendor side (for risk management) and/or mobile phone app on the user side (for streamlining communication and reducing transmission channels).

Work to progress some of these aspects is currently in discussion to be developed further at the University of Huddersfield.

References

[1] Abkarian M., Mendez S., Xue N., Yang F. and Stone H.A., 2020. Speech can produce jet-like transport relevant to asymptomatic spreading of virus, Proceedings of the National Academy of Sciences, 117 (41) 25237-25245.

[2] Unlocking HE Spaces (2020) `https://gateway.newton.ac.uk/sites/default/files/asset/doc/2007/Unlocking%20HE%20Spaces%20July%202020_small_0.pdf`

[3] Bazant M.Z. and Bush J.W.M, 2020. Beyond Six Feet: A Guideline to Limit Indoor Airborne Transmission of COVID-19, medRxiv 2020.08.26.20182824.

[4] Bhagat, R., Davies Wykes, M., Dalziel, S., Linden, P. (2020). Effects of ventilation on the indoor spread of COVID-19. Journal of Fluid Mechanics, 903, F1. doi:10.1017/jfm.2020.720

[5] Chao C.Y.H., 2009. Review and comparison between the Wells–Riley and dose-response approaches to risk assessment of infectious respiratory diseases. Indoor Air 2010, 20: 2–16.

[6] Investigations on the propagation probability of aerosols in the passenger compartment of rail vehicles. DB Systemtechnik GmbH, competence centre aerodynamics and air conditioning (2020)

[7] Gammaitoni L, Nucci MC. Using a mathematical model to evaluate the efficacy of TB control measures. Emerg Infect Dis. 1997;3(3):335-342. doi:10.3201/eid0303.970310

[8] Gupta, J. K., Lin, C. H. Chen, Q. 2010 Characterizing exhaled airflow from breathing and talking. Indoor Air 20 (1), 31–39

[9] Hasan et al. Effect of artificial mucus properties on the characteristics of airborne bioaerosol droplets generated during simulated coughing. Journal of Non-Newtonian Fluid Mechanics. Volume 165, Issues 21–22, November 2010, Pages 1431-1441 https://doi.org/10.1016/j.jnnfm.2010.07.005

[10] Li Z. et al., 2020. Effects of space sizes on the dispersion of cough-generated droplets from a walking person. Physics of Fluids 32, 121705.

[11] Milton, D. K. 2020 A rosetta stone for understanding infectious drops and aerosols. J. Pediat. Inf. Dis. Soc. doi:10.109/pid/iaa079

[12] Noakes C., Beggs C., Sleigh P. and Kerr K, 2006. Modelling the transmission of airborne infectionsin enclosed spaces. Epidemiology & Infection 134, 1082–1091.

[13] Patw et al. Analysis of human nasal mucous glycoproteins. American Journal of Otolaryngology. Volume 5, Issue 5, September–October 1984, Pages 334-343 https://doi.org/10.1016/S0196-0709(84)80003-4

[14] Riley E.C., Murphy G. and Riley R.L., 1978. Airborne spread of measles in a suburban elementaryschool, American Journal of Epidemiology 107, 421–432.

[15] Wells W.F. et al., 1955. Airborne Contagion and Air Hygiene. An Ecological Study of Droplet Infections.

[16] WHO 2014 Infection Prevention and Control of Epidemic - and Pandemic-Prone Acute Respiratory Infections in Health Care. World Health Organization.

[17] Zhang et al. Distribution of droplet aerosols generated by mouth coughing and nose breathing in an air-conditioned room. Sustainable Cities and Society. Volume 51, November 2019, 101721. https://doi.org/10.1016/j.scs.2019.101721

[18] Zhang Z. et al., 2021. Disease transmission through expiratory aerosols on an urban bus. Physics of Fluids 33, 015116.

4 Scale 2: Modelling of a Journey

4.1 Challenge Description

In this section we consider the wider topics of disease transmission with respect to a passengers journey both in the train and at stations. A passengers journey, experience and therefore Covid-19 susceptibility will depend on many factors independent of carriage layout (as considered in Section 3.

Thinking of an individuals journey, one travels to s station → travel through a station → **enters the train from a platform** → sits on the train → **leaves the train onto a platform** → leaves the station → Travel from the station. The two parts in bold involve a lot of crowding and are potentially the most dangerous as measured by total viral loading. This section is divided into two parts; the first looks at the entering the train question, and the other at the leaving the train question.

4.2 Getting on a train

We consider a typical train and platform set up in Fig. 15 using sensible dimensions, and assuming a single entrance at the centre of the platform and spreading out according to a Gaussian distribution. Our train was assumed to be two carriages and four doors.

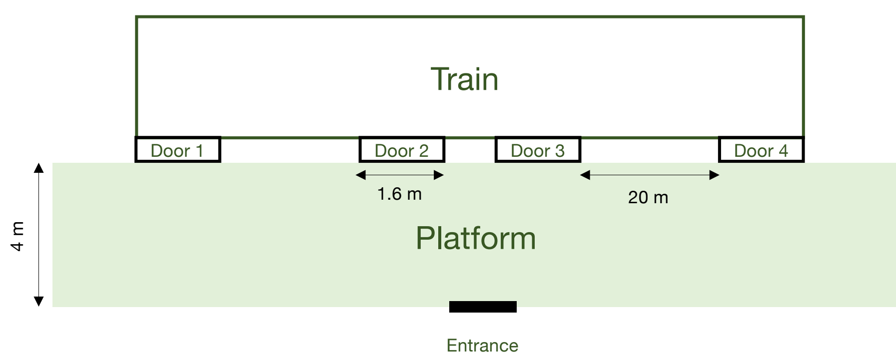


Figure 15: Platform setup - a toy model. Passengers assumed to enter the platform at the centre, then distributed according to a Gaussian distribution.

We then model the passengers motion towards the doors in a continuous framework. Passengers waiting on a railway platform for a train are initially scattered. We model their motion towards the train carriage as in a Fluid Dynamics framework. More precisely, we simulate density of passengers movement with the convection-diffusion PDE

$$\partial_t u + \nabla \cdot (\vec{v}u) = \nabla \cdot (\sigma \nabla u), \quad u \in C^1(\Omega)$$

where u is the crowd density, $\vec{v}$ is the speed, and σ is the diffusivity parameter. The initial domain Ω is represented by the platform in Fig. 16

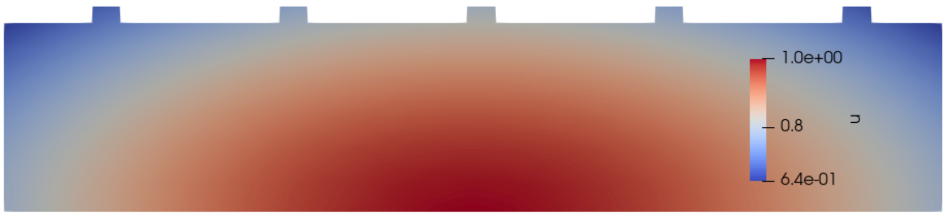


Figure 16: Initial domain

The narrowings at the top boundary represent the carriage doors. Their number, width and location are tunable parameters in order to account for different scenarios. Assumptions:

1. The passengers access to the platform from the midpoint of the bottom boundary and scatters as an exponential function:

   $$u_0 = e^{-(\beta_1(x-0.5)^2+\beta_2(y-0.8)^2)},$$

   where β_1 and β_2 characterise the level of initial spread.

2. The variable σ controls the diffusivity of the passengers. This can account for passengers that get off the train or are disoriented if do not have any reservation.

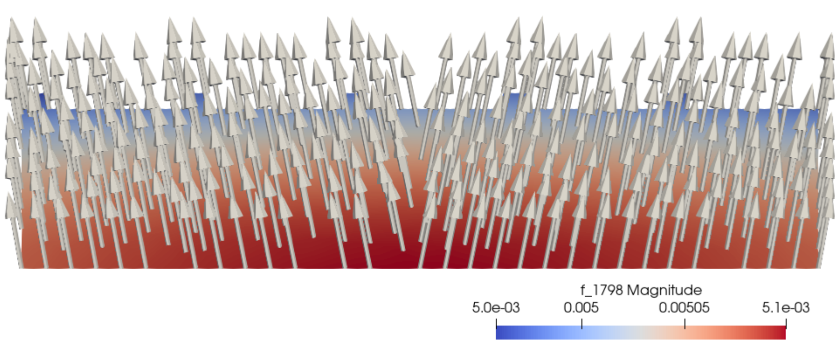


Figure 17: Velocity field for passengers.

3. The velocity $\vec{v}(x, y)$ is constant in time but dependent in space. We suppose that the passengers moves radially to avoid clustering in a specific door, Fig. 17.

   More sophisticated models can have $\vec{v}$ dependent on local conditions, e.g. crowd density u, and/or explicitly on time – people in a hurry if the train is about to go.

4.2.1 Discretization of the PDE in space

The variational formulation to be solved is:

$$\int_\Omega \left[w\frac{\partial u}{\partial t} + w\nabla\cdot(\vec{v}u) + \nabla w\cdot\sigma\nabla u\right]dx - \int_\Gamma w\frac{\partial(\sigma u)}{\partial n}dx = 0, \quad \forall w \in H^1(\Omega)$$

Alternatively we can also integrate the convective term by parts as well:

$$\int_\Omega \left[w\frac{\partial u}{\partial t} - \nabla w\cdot(\vec{v}u) + \nabla w\cdot\sigma\nabla u\right]dx + \int_\Gamma \left[w\vec{n}\cdot\vec{v}u - w\frac{\partial(\sigma u)}{\partial n}\right]dx = 0, \quad \forall w \in H^1(\Omega)$$

The boundary integrand is the test function times outward passenger flux. We also assume homogeneous zero-flux boundary conditions over all the boundaries,

The equation is discretized in space on the domain $\Omega = [0.0, 50.0] \times [0.0, 4.0]$ using the finite element method (FEM). We discretize the domain with a conforming unstructured mesh $\mathcal{T}_h$ and consider the finite element (FE) space

$$\mathbb{V} = \{u_h \in H^1(\Omega) : u_h|_K \in C^1(K) \quad \forall K \in \mathcal{T}_h\}$$

4.2.2 Discretization of the PDE in time

We discretize the previous equation in time using a semi-implicit Euler scheme. At each time step $t_n = 0, \cdots, T$, we aim to find the solution u_h^{n+1} satisfying

$$\int_\Omega \left[w_h u_h^{n+1} + \Delta t\sigma\nabla w_h\cdot\nabla u_h^{n+1}\right]dx = \int_\Omega \left[w_h u_h^n - \Delta t w_h(\vec{v}\cdot\nabla u_h^n)\right]dx, \quad \forall w_h \in \mathbb{V},$$

4.2.3 Results

We compute the density of the passengers inside each door at each time step and see how this is affected by the parameter β_1, which controls the scattering of the passengers on the platform. We can see that the more uniform the passengers on the platform, the more uniform the density of passengers waiting at each door.

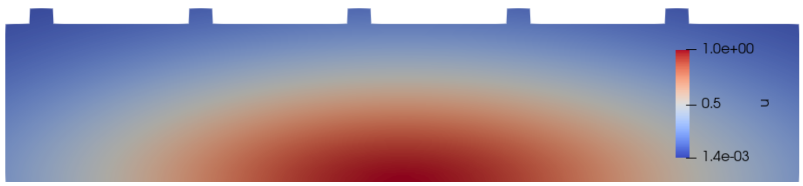


Figure 18: Initial condition ($\beta_1 = 1$)

Figure 19: Final condition ($\beta_1 = 1$)

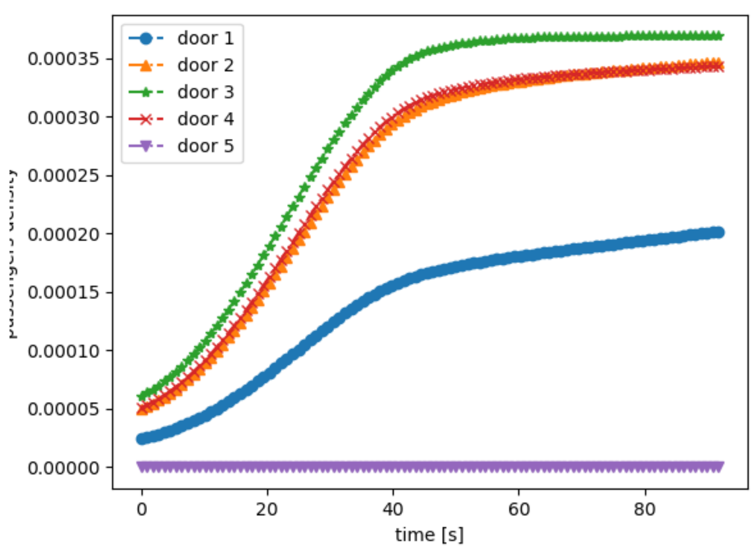


Figure 20: Density of passengers ($\beta_1 = 1$)

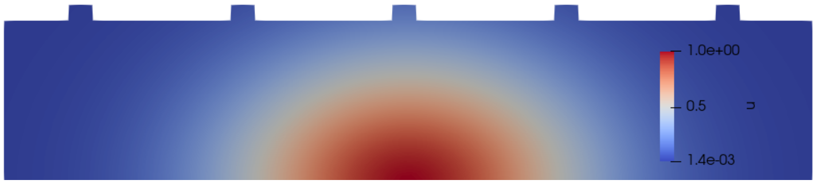


Figure 21: Initial condition ($\beta_1 = 20$)

Figure 22: Final condition ($\beta_1 = 20$)

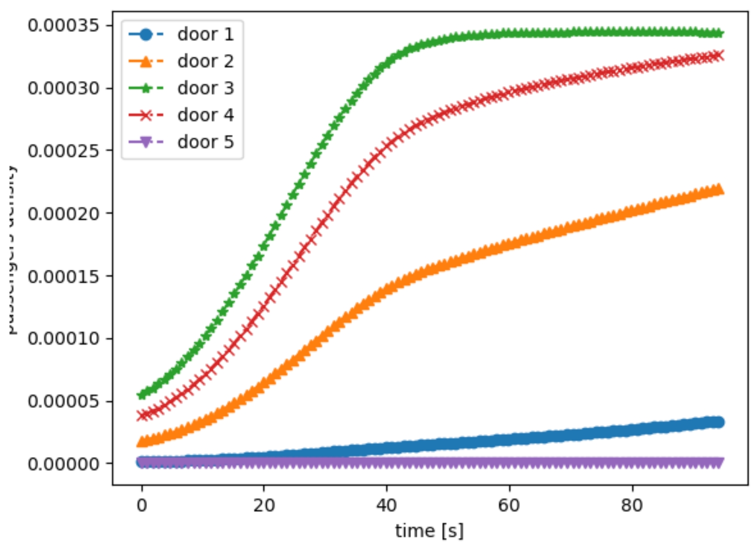


Figure 23: Density of passengers ($\beta_1 = 20$)

4.2.4 Future work

Passenger distribution

- Passengers can arrive through a platform gate unimpeded at a rate of 40 per minute [4].
- Passenger's distribution is not necessarily Gaussian nor Uniform. Passenger location on the platform is influenced by

[4] Source: Network Rail Station Capacity Planning Guidance.

  - Location of platform entrances
  - Any shelters or canopies
  - Location of CIS (customer information screen) displays
  - Knowledge of busiest carriages
  - Seat reservations in some cases

- In the future, results can be simulated for a variety of platform setups.

Equations of motion

- Walking speed to be informed by data and depend on local crowding and departure time.
- Likewise random motion passenger motion.
- Different populations to be accounted for: reservations in different carriages; people awaiting later trains; exiting passengers
- Better modelling of entry into train doors
- Coupling with infected-passenger densities to get susceptible-passenger exposure/risk.

4.3 Getting off a train

The top-level problem description is that a train is stationary at a platform, passengers exit and board, announcements are made in train and on platform. This includes the 'London Bridge situation', a penultimate station on a commuter line, where many passengers exit as well as board. Assumptions:

1. passengers only exit (commuter train has reached destination), announcements made on train. Less risky than 'London Bridge'.
2. simplification: passengers exit through a single door (left and right queue merging), announcements made on train.
3. as before, but only a single queue. Reflects mainline carriage layout.

Here the aim is to model the behaviour to see the distance and waiting time together with how long each passenger has to wait until they get to the exit. We assume that the queue has N passengers at location $X_n(t)$, $n = 0 \ldots N - 1.$, Fig. 24.

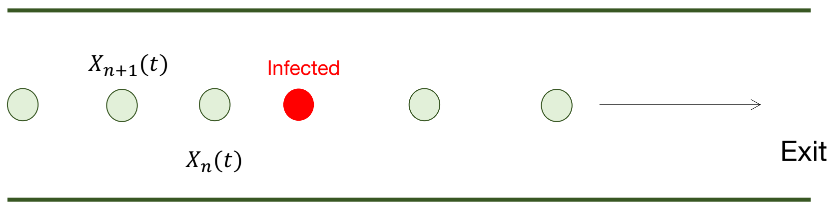


Figure 24: Diagram of queue at train door.

The passengers are ordered so that $X_{n+1} < X_n$ and the queue is assumed to move to the right so that X_n exits before X_{n+1}. Each passenger will be assumed to have a level of infection γ_n where $\gamma_n = 0$ if

1. the passenger is not infectious or
2. the passenger has left the carriage and plays no more role in the queue.

When $X_n = L$ (the length of the carriage) then the n^{th} passenger will be assumed to have left the carriage. This will take a time Δ_n.

4.3.1 Velocity based model

The model adopted here is a velocity based model, we have a function f which describes the behaviour and distance between passengers.

$$\dot{X}_{n+1} = f_{n+1}(X_n - X_{n+1})$$

$$f(r) = 0 \quad r < d_1, \quad f(r) = V, \quad r > d_2, \quad \text{linear in between.}$$

Fig. 25 shows the behaviour of this function.

- Here d_1 is a social distance (maybe depend on n)
- Take $d_2 = 2d_1$ (again depends on n).
- V is a saturation velocity which depends on announcements and dwell time.

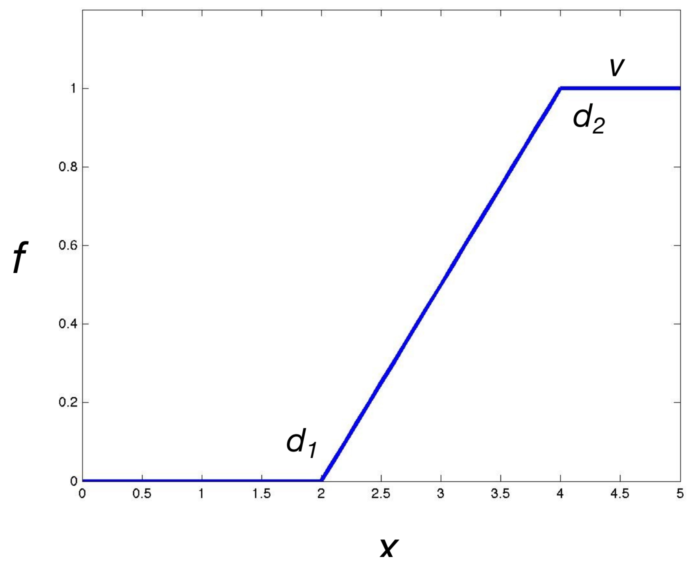


Figure 25: Function used in model which describes the behaviour and distance between passengers.

- Initially we will have everyone standing socially distant ($r > d_1$).
- Then the door opens and X_0 starts to move at velocity V.

Based on this model, we simulate 10 passengers, Fig.26. When the train gets to the destination, there is an announcement, passengers stand up. The first passenger is the blue passenger, there is no wait time for the first passenger (no queue in front of them). Passenger 2 (the curve under passenger one) has a short wait behind passenger one, and so on.

Typically a passenger;

- has a waiting time before they can move
- takes a further travel time to leave the train

Both of these lead to potential viral exposure. The simulated time for both of these events can be seen in Table. 1.

4.3.2 Viral load

Here we consider the viral load, as the passengers exit they are exposed to a viral load from both

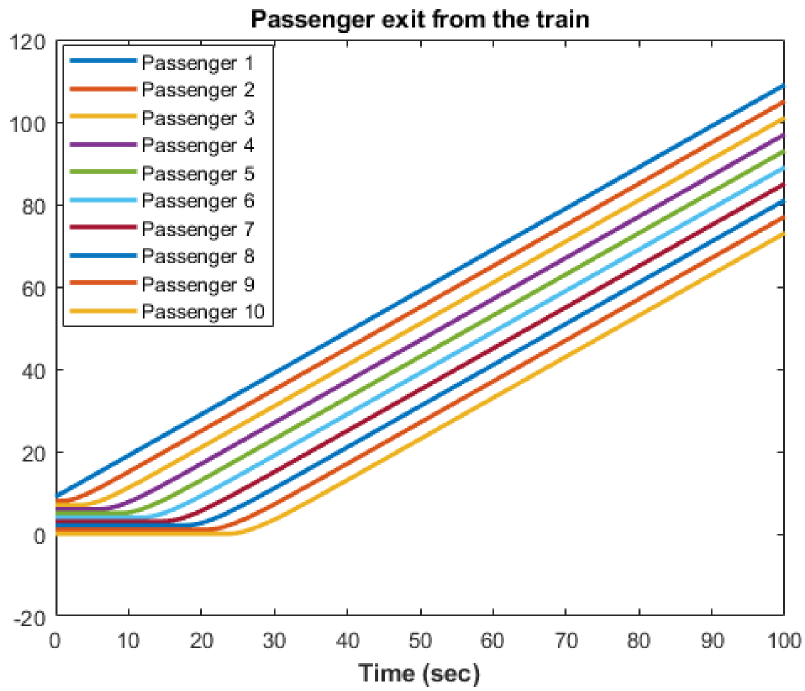


Figure 26: Simulated distance covered (in metres) by 10 passengers in the train before exit

Passenger	Waiting time (s)	Travel time (s)
1	0.45	6.04
2	0.90	10.3
3	3.45	14.2
4	6.67	18.3
5	9.24	22.4
6	12.3	26.0
7	15.3	30.1
8	17.8	34.6
9	20.8	39.0
10	23.9	42.5

Table 1: Simulated waiting and travel times for 10 passengers in a queue to leave a train.

- The the stationary passengers they walk past (Section. 3)
- The other passengers in the queue.

Viral load from stationary passengers

To estimate the viral load from the stationary passengers we assume that a proportion α of the passengers in the carriage are infectious. The number of passengers encountered is then $\alpha(L - X_n(0))$. The viral load encountered will then be

$$Y_n = \beta \Delta_n (L - X_n(0)), \quad (Y)$$

where β is a constant.

Viral load from passengers in queue

To model the viral load from other passengers in the queue we assume an inverse square law for the spread:

$$\frac{dW_n}{dt} = \theta \sum_{j \neq n} \frac{\gamma_n}{|X_n - X_j|^2}, \quad (W)$$

where θ is a constant. Note that γ_n drops to zero when the n^{th} passenger leaves the train.

The total viral load is then

$$V = \sum_n V_n \equiv \sum_n \left(Y_n + W_n(\Delta_n) \right). \quad (V)$$

How to implement

- Assume that we have code which can find $X_n(t)$
- Queue with X_0 at the front and $X_{n+1} < X_n$.
- Assume an initial stochastic distribution for γ_n. For example we might have a situation where only one (randomly chosen) passenger is infectious.
- Calculate $W_n(t)$ by solving the ODE (W) simultaneously with finding $X_n(t)$.
- Calculate $\gamma_n(t)$ setting $\gamma_n(t) = 0$ if $X_n > L$.
- Compute Δ_n
- Calculate Y_n directly from (Y)
- Find V from the sum (V).

4.4 Modelling risky behaviour in a queue

We know from previous research that the **level of risk taken** is important. We need to find definitions of risk-seeking, risk-neutral, and risk-averse in the Covid-19 context

- Risk seeking: no distance keeping, some random movements (Brownian motion), ignoring announcements related to distance keeping and direction of movement
- Risk neutral: keeps recommended distance, no random movements, follows recommended direction of movement
- Risk averse: keeps excessive distance, moves slower, follows recommended direction of movement

Simplification: we will model risk purely by distance kept in the queue. Minimal distance = high risk behaviour. Recommended distance = risk neutral. Fixed but large distance = risk averse.

4.5 Conclusions and next steps

Modelling passenger flow onto a train in a continuous manner can be used to model various load strategies. Future work could lift the Gaussian loading restriction to simulate more realistic conditions. Additionally, it could be instructive to modify the equations of motion (with observed data) to better capture the characteristics of the passenger population.

Similarly, the simple viral load modelling of a queue system is worth exploring more, as queue systems are very common not only in transport hubs, but in many setting. Both of these topics will be subject to further investigation at the University of Bath.

References

[1] Linbo Luo, Cheng Chai, Suiping Zhou, and Jianfeng Ma. 2016. Modeling Gap Seeking Behaviors for Agent-based Crowd Simulation. In Proceedings of the 29th International Conference on Computer Animation and Social Agents (CASA '16). Association for Computing Machinery, New York, NY, USA, 37–43. DOI: `https://doi.org/10.1145/2915926.2915944`

[2] Thoreau, Roselle & Holloway, Catherine & Bansal, Gurvir & Gharatya, Kuldeep & Roan, Tyng-Rong & Tyler, Nick. (2017). Train design features affecting boarding and alighting of

passengers: Train design affecting boarding and alighting. Journal of Advanced Transportation. 50. 10.1002/atr.1446.

[3] Weiss, Tomer & Litteneker, Alan & Jiang, Chenfanfu & Terzopoulos, Demetri. (2018). Position-based real-time simulation of large crowds. Computers Graphics. 78. 10.1016/j.cag.2018.10.008.

5 Scale 3: Modelling of a Network

5.1 Challenge Description

This section considers what national (network)-level impact there is from Covid-19 and in particular, how can the network cope when responding to national level events, for example, the movement of students from home to universities and vice versa. In particular, this section focuses on constructing a model, and identifying data requirements, for implementing this methodology in practice. Fig. 27 shows the proposed methodology which proposes schedule, passenger commute, and network data can be used to construct a network graph for evaluating various traveling and mitigation scenarios.

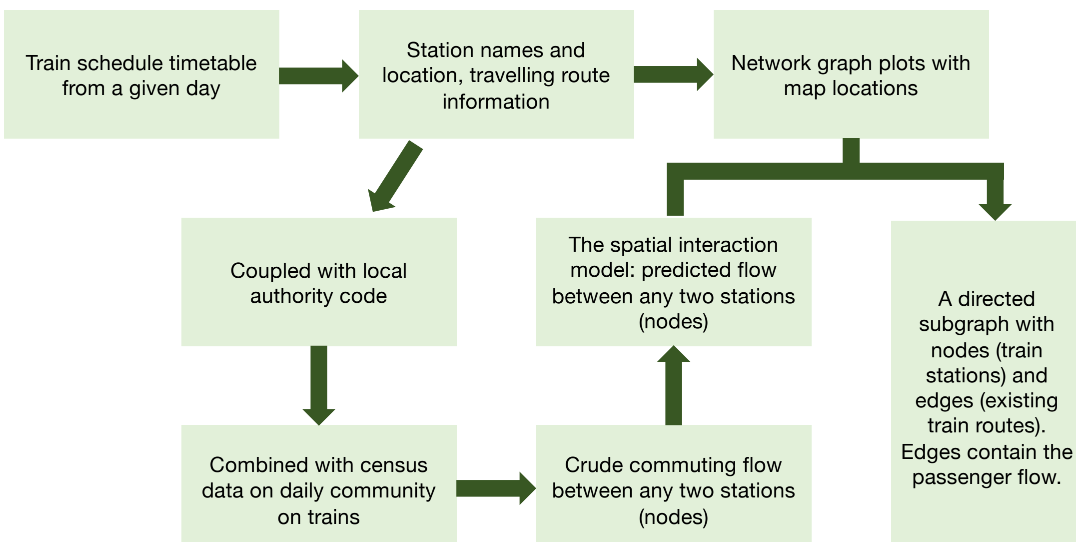


Figure 27: Challenge approach diagram

This first task is to develop a mathematical model to assess the quality of a given plan / schedule taking into account the following aspects; (i) passenger demand, (ii) trains and number of coaches of each type available, and (iii) deviation from plan - only if capacity is considered. The approach taken in this VSG is to:

- start with a small model, e.g. London to Manchester:
- student flow: estimate numbers; use parameters to control when to travel
- include coach types - what type of characteristics can we simplify.

Let $G = (N, A)$ be an undirected graph with node set N and arc set A. T is the number of time periods. Each vertex in N represents a station met by any train, and the edge between each pair of vertices means trains run from one station to another without stopping in the middle. A crude diagram of such a network can be seen in Fig. 28. We classify two types of edges in this model network:

- minimal edges - a set of continuous connections between stations not passing through a third one; e.g. trains are stopping at each stations connected by these edges
- transitive edges - connections between two stations but trains are not stopping at these stations.

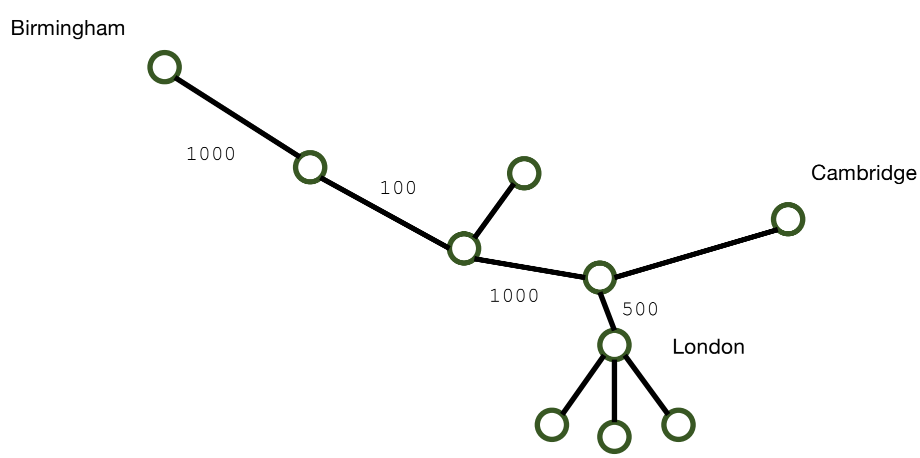


Figure 28: Example of network representation, the nodes and edges shown in this graph are not based real data.

5.2 Optimisation model

Here we propose an optimisation approach for a given scenario s, taking into account capacity C and flow f. f_{ij}^{ts} is the expected flow at time t on arc (i, j) on scenario $s \in S$. (s is related to the inflow interval). $C_{i,j}$ is the capacity for each train.

Let H be the number of different types of carriages and let $C^h,\ h \in \{1, \dots, H\}$ be the capacity of coach type h. Each train can be represented as a vector $(n_1, \dots, n_H)$, where n_h is the number of coaches of each type in the train. Note that $C_{ij} = \sum_{h=1}^{H} n_h * C_h$. We seek a solution which satisfies the below;

$$\min \max_{i,j,t,s} \{f_{ij}^{ts} - C_{ij}\}.$$

5.3 Data

Here we discuss some of the data required to parameterise our model. Table. 2 shows the train timetable on 22/12/2020 from London Northwestern Railway and West Midlands Railway. This gives a small section of the network to model, shown in Fig. 29. The timetable connects regions around London and to the Midlands around Birmingham and then extends north to Liverpool.

Table 2: Train timetable on 22/12/2020 from London Northwestern Railway and West Midlands Railway. There are 1342 train operation on this section within the day.

id	production schedule id	sequence	location code	event type	event time	advertised event time
1.74E+08	7507823	1	STRBDGT	A	22/12/2020 19:58	22/12/2020 19:58
1.74E+08	7507294	32	CREWE	A	22/12/2020 23:38	22/12/2020 23:40
1.74E+08	7507294	29	ALSAGER	D	22/12/2020 23:28	22/12/2020 23:28
1.74E+08	7507294	28	ALSAGER	A	22/12/2020 23:28	22/12/2020 23:28
1.74E+08	7507294	27	KIDSGRV	D	22/12/2020 23:24	22/12/2020 23:24
1.74E+08	7507294	26	KIDSGRV	A	22/12/2020 23:23	22/12/2020 23:24
1.74E+08	7507294	25	STOKEOT	D	22/12/2020 23:16	22/12/2020 23:16
1.74E+08	7507294	24	STOKEOT	A	22/12/2020 23:15	22/12/2020 23:15
1.74E+08	7507294	22	STONE	D	22/12/2020 23:08	22/12/2020 23:08
1.74E+08	7507294	21	STONE	A	22/12/2020 23:07	22/12/2020 23:07
1.74E+08	7507294	18	STAFFRD	D	22/12/2020 22:59	22/12/2020 22:59
1.74E+08	7507294	17	STAFFRD	A	22/12/2020 22:58	22/12/2020 22:58
1.74E+08	7507294	15	PNKRDG	D	22/12/2020 22:51	22/12/2020 22:51
1.74E+08	7507294	14	PNKRDG	A	22/12/2020 22:51	22/12/2020 22:51
1.74E+08	7507294	11	WVRMPTN	D	22/12/2020 22:43	22/12/2020 22:43

Each `schedule id` in timetable represents a unique train schedule from one beginning station, passing certain stopping stations and reaches a final destination station. So same route trains have different `id`'s with different scheduled time. Train station are labeled with codes like STRBDGT which can vary based on the operation companies. Stations listed in the timetable means there are trains stopping at them and there are no passing-by stations. Event Type column contains labels A for arrival and D for departure. In Fig. 29, we plot all the train stations from timetable Table. 2 as red dots.

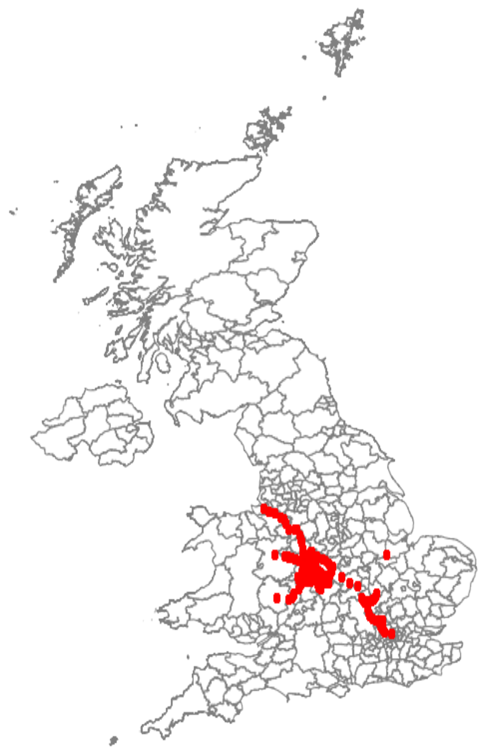


Figure 29: Sample rail network based on the train timetable in England. All nodes are train stations and presented as red dots in the graph. Edges are not drawn here because the complexity and visibility. Edges in our graph model means direct linked train stations, and trains stop at the nodes.

To reduce the complexity, two two sub-graphs are taken from this branch: (1) London to Birmingham and (2) Birmingham to Crewe. For example, take train 7507294 in Table. 2, it leaves from Birmingham New Street Rail (BHAMNWS) Station and arrives at Crewe (CRWEW) - note that the schedule in Table. 3 is in reverse order. Using these two routes, we model the daily travel demand between London and midlands regions in this paper.

Table. 3 shows the branch Birmingham (coded as BHAMNWS) at 22:22:00 and reaches Crewe (coded as CREWE) at 23:40:00 - note that the schedule in Table. 3 is in reverse order.

Table 3: Sample schedule for train 7507294. Note schedule in reverse order.

	id	production schedule id	sequence	location code	event type
1	173799567	7507294	32	CREWE	A
2	173799564	7507294	29	ALSAGER	D
3	173799563	7507294	28	ALSAGER	A
4	173799562	7507294	27	KIDSGRV	D
5	173799561	7507294	26	KIDSGRV	A
6	173799560	7507294	25	STOKEOT	D
7	173799559	7507294	24	STOKEOT	A
8	173799557	7507294	22	STONE	D
9	173799556	7507294	21	STONE	A
10	173799553	7507294	18	STAFFRD	D
11	173799552	7507294	17	STAFFRD	A
12	173799550	7507294	15	PNKRDG	D
13	173799549	7507294	14	PNKRDG	A
14	173799546	7507294	11	WVRMPTN	D
15	173799545	7507294	10	WVRMPTN	A
16	173799544	7507294	9	COSELEY	D
17	173799543	7507294	8	COSELEY	A
18	173799541	7507294	6	SNDWDUD	D
19	173799540	7507294	5	SNDWDUD	A
20	173799539	7507294	4	GALTILL	D
21	173799538	7507294	3	GALTILL	A
22	173799535	7507294	0	BHAMNWS	D

Table. 4 shows the branch to London, train 7509540 leaves from London Euston (EUSTON) at 19:15:00 to Birmingham new Street (BHAMNWS) at 19:15:00- note that the schedule in Table. 4 is in reverse order.

Table 4: Sample schedule for train 7509540. Note schedule in reverse order.

	id	production schedule id	sequence	location code	event type

821	173852328	7509540	44	BHAMNWS	A
822	173852325	7509540	41	MRSTNGR	D
823	173852324	7509540	40	MRSTNGR	A
824	173852323	7509540	39	BHAMINT	D
825	173852322	7509540	38	BHAMINT	A
826	173852321	7509540	37	HMPTNIA	D
827	173852320	7509540	36	HMPTNIA	A
828	173852319	7509540	35	BKSWELL	D
829	173852318	7509540	34	BKSWELL	A
830	173852317	7509540	33	TILEH	D
831	173852316	7509540	32	TILEH	A
832	173852315	7509540	31	CANLEY	D
833	173852314	7509540	30	CANLEY	A
834	173852313	7509540	29	COVNTRY	D
835	173852312	7509540	28	COVNTRY	A
836	173852310	7509540	26	RUGBY	D
837	173852309	7509540	25	RUGBY	A
838	173852306	7509540	22	LNGBKBY	D
839	173852305	7509540	21	LNGBKBY	A
840	173852304	7509540	20	NMPTN	D
841	173852303	7509540	19	NMPTN	A
842	173852301	7509540	17	WLVR	D
843	173852300	7509540	16	WLVR	A
844	173852299	7509540	15	MKNSCEN	D
845	173852298	7509540	14	MKNSCEN	A
846	173852297	7509540	13	BLTCHLY	D
847	173852296	7509540	12	BLTCHLY	A
848	173852295	7509540	11	LTNBZRD	D
849	173852294	7509540	10	LTNBZRD	A
850	173852284	7509540	0	EUSTON	D

In our model, we first map the location codes to stations' names and geographical locations. This way we construct our graphs based on the geographical information of nodes. In order to model passenger flow, we used census commuting data from 2011. In order to use census data, our model maps geological locations to local authority numbers. This way we can extract the census data from daily commuting (by train) in England among different cities. Based on this commuting number among cities, we use a spatial interaction model to estimate passenger flow. Details are in the next subsection.

5.4 Derivation of incoming flows

Here we show how spatial interaction models can be used to estimate flows on the passenger network and provide tools for exploring "what-if" scenarios. Spatial interaction models are based on Newtonian principles of gravity and include parameters which vary the attractiveness of node, and impedance to flow. These could be used to proxy represent the attractiveness of flow from one node to another (for example to represent how students move in response Government advice). The aim here is to estimate the expected number of people that will be using the given network. In the present work we use a doubly constrained flow model (spatial interaction model). This is defined as:

$$T_{ij} = A_i O_i B_j D_j d_{ij}^{-\beta}$$

where

$$O_i = \sum_j T_{ij}$$

$$D_j = \sum_i T_{ij}$$

and

$$A_i = \frac{1}{\sum_j B_j D_j d_{ij}^{-\beta}}$$

$$B_j = \frac{1}{\sum_i A_i O_i d_{ij}^{-\beta}}$$

In the above specification, T_{ij} represent the expected flows from an origin i to a destination j, d is a measure of separation (taken as distance here for simplicity). This model can be reparametrised to be estimated as an Poisson generalised linear model with log link function:

$$\lambda_{ij} = \exp(\mu + \alpha_i - \beta \ln d_{ij})$$

in which α_i represents the $D_j B_j$ part of the model equation.

The model outputs predicted flows accounting for a separation metric and the fact that the population flows from and to a destination have to match. This particular specification can be expanded to include any form of impedance function as well as any potential covariates that would help explain the observed variation in the flows.

The model was estimated using data from 2011 census travel to work study (8), and in that sense the predicted flows represent the situation pre-Covid-19, with RMSE = 60.46. The data are on local authority level so we assigned the predicted flows to each station belonging to the corresponding LA. However, the observed flows can be approximated using other source of data such as telecomms or ticketing data from automatic fare collection systems. Ideally we would want to fit a disaggregated model taking into account different population strata, such as students, but there is a lack of available data for this task.

5.5 Results

5.5.1 Subnetwork 1: Birmingham to Crewe

Here we show estimated flow between stations on the following network - Birmingham to Crewe:

$$\begin{bmatrix} \text{CREWE} & \text{ALSAGER} & \text{KIDSGRV} & \text{STOKEOT} & \text{STONE} & \text{STAFFRD} & \text{PNKRDG} & \text{WVRMPTN} & \text{COSELEY} & \text{SNOWDUD} & \text{GALTILL} & \text{BHAMNWS} \end{bmatrix}$$

Commuter flow based on census data in 2011:

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 94 & 94 & 94 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 94 & 94 & 94 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 354 & 354 & 354 & 15 & 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 354 & 354 & 354 & 15 & 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 555 & 555 & 555 & 45 & 45 & 441 & 441 & 0 & 0 & 0 & 0 & 0 \\ 555 & 555 & 555 & 45 & 45 & 441 & 441 & 0 & 0 & 0 & 0 & 0 \\ 555 & 555 & 555 & 45 & 45 & 441 & 441 & 0 & 0 & 0 & 0 & 0 \\ 555 & 555 & 555 & 45 & 45 & 441 & 441 & 0 & 0 & 0 & 0 & 0 \\ 24 & 24 & 24 & 0 & 0 & 36 & 36 & 175 & 175 & 175 & 175 & 0 \end{bmatrix}$$

In this matrix F, the rows (from top to bottom) are listed as from Crew to Birmingham and columns (left to right) are from Crew to Birmingham. This is in consistent with the reversed timetable order. So each entry $F_{i,j}$ represents the number of passengers getting on the train at station i and leaves at station j. For example, $F_{12,1}$ represents that there are 24 passengers who get on the train at BHAMNWS and leaves at CREWE and $F_{12,10}$ means 175 get on the train from BHAMNWS and leaves at SNOWDUD. Because it is reversed order, this F is a strict lower triangle matrix.

In the following matrix, we used the spatial interaction models to estimate a smooth passenger flow. Because the census only collects the number of people leaving from a city (assumed to be home) and reaches another city (assumed to be the working site) in 2011, we need the spatial model to give an estimate of the averaged flow. Spatial model can also be accommodated to changes in local economy, this way we have a relatively realistic number.

Commuter flow estimated by the spatial model (model data input is from the 2011 census):

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 172 & 172 & 172 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 172 & 172 & 172 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 264 & 264 & 264 & 7 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 264 & 264 & 264 & 7 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 \\ 330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 \\ 330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 \\ 330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 \\ 96 & 96 & 96 & 0 & 0 & 56 & 56 & 445 & 445 & 445 & 445 & 0 \end{bmatrix}$$

This matrix has the same structure as the previous one: each entry $F_{i,j}$ represents the number of passengers get on the train at station i and leaves at station j. The numbers are calculated using the spatial model. So the census data is adjusted and smoothed based on local development during the past 9 years.

The above two matrices shows the daily flow of passengers on the Birmingham-Crewe network. Assume we have one "train" for this flow, then the total number of commuters on this "train" linking two stations i and j, where i and j are in the network nodes:

1. original commuting census data based L_o:

$$\begin{bmatrix} 844 & 3306 & 5768 & 8230 & 106922 & 9984 & 9276 & 9348 & 9420 & 6280 & 3140 \end{bmatrix}$$

2. passenger flow estimated data L_e:

$$\begin{bmatrix} 2180 & 3067 & 3954 & 4841 & 5728 & 5834 & 5940 & 6402 & 6840 & 4576 & 2288 \end{bmatrix}$$

Here the network is in forward order:

$$\begin{bmatrix} \text{BHAMNWS} & \text{GALTILL} & \text{SNOWDUD} & \text{COSELEY} & \text{WVRMPTN} & \text{PNKRDG} & \text{STAFFRD} & \text{STONE} & \text{STOKEOT} & \text{KIDSGRV} & \text{ALSAGER} & \text{CREWE} \end{bmatrix}$$

These two arrays represent the total number of passengers are on the "train" from station i to the next i+1, i in the forward network as above. For example, first number 844 $L_o[1]$ means the amount of people on the train between BHAMNWS and GALTILL. These are the numbers on edges linking those nodes. Put into our graph model gives Figure 30:

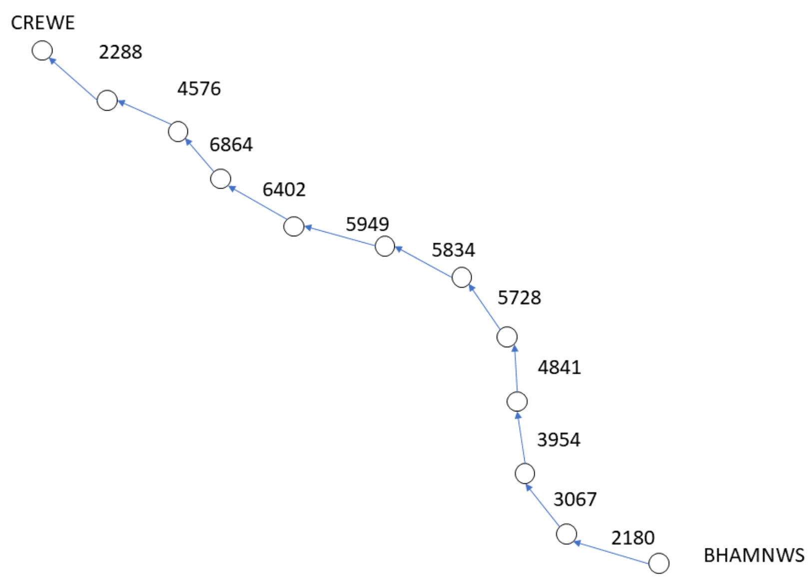


Figure 30: Graph from BHAMNWS to CREWE- daily passenger flow is estimated using our spatial model.

Therefore the maximum capacity requirement on this graph is around 6864, using the estimated date.

5.5.2 Subnetwork 2: London to Birmingham

Use our model, we can esitmate the flow from London Euston station to Birmingham station. The network is as follows - in reversed order:

['BHAMNWS' 'MRSTNGR' 'BHAMINT' 'HMPTNIA' 'BKSWELL' 'TILEH' 'CANLEY' 'COVNTRY' 'RUGBY' 'LNGBKBY' 'NMPTN' 'WLVR' 'MKNSCEN' 'BLTCHLY' 'LTNBZRD' 'EUSTON']

The flow matrix estimated by the spatial model:

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
172 & 172 & 172 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
172 & 172 & 172 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
264 & 264 & 264 & 7 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
264 & 264 & 264 & 7 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
330 & 330 & 330 & 10 & 10 & 161 & 161 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
96 & 96 & 96 & 0 & 0 & 56 & 56 & 445 & 445 & 445 & 445 & 0 & 0 & 0 & 0 & 0 \\
96 & 96 & 96 & 0 & 0 & 56 & 56 & 445 & 445 & 445 & 445 & 0 & 0 & 0 & 0 & 0 \\
75 & 75 & 75 & 1 & 1 & 55 & 55 & 208 & 208 & 208 & 208 & 70 & 70 & 0 & 0 & 0 \\
75 & 75 & 75 & 1 & 1 & 55 & 55 & 208 & 208 & 208 & 208 & 70 & 70 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 20 & 20 & 22 & 22 & 22 & 22 & 20 & 20 & 996 & 996 & 0
\end{bmatrix}
$$

Similar to Subgraph 1, this estimated flow matrix smoothed the flow based on those parameters stated before. The result is not as realistic because local economy and many other factors have not been included in the model during this VSG. But the data is still sensible for carrying out research. The link array in forward order:

$$
\begin{bmatrix}
2160 & 2473 & 2786 & 4806 & 6826 & 6830 & 6834 & 6838 & 6842 & 6762 & 6682 \\
7142 & 7602 & 5068 & 2534 & & & & & & &
\end{bmatrix}
$$

The maximum capacity needed for this route is around 7600. Combining the nodes and edges together gives the Graph 31:

By the end of the VSG, the team produced a piece of software by reading in a train timetable and census commuting data to produce subgraphs and estimated customer flow. This customer flow is calculated on the non-COVID time. However in order to understand how the nationwide network operates, we need to link related subgraphs together to do further simulation. For example, people or students who need to travel from London to Edinburgh will need to pass through several of those small subgraphs because there is no direct link between some major cities.

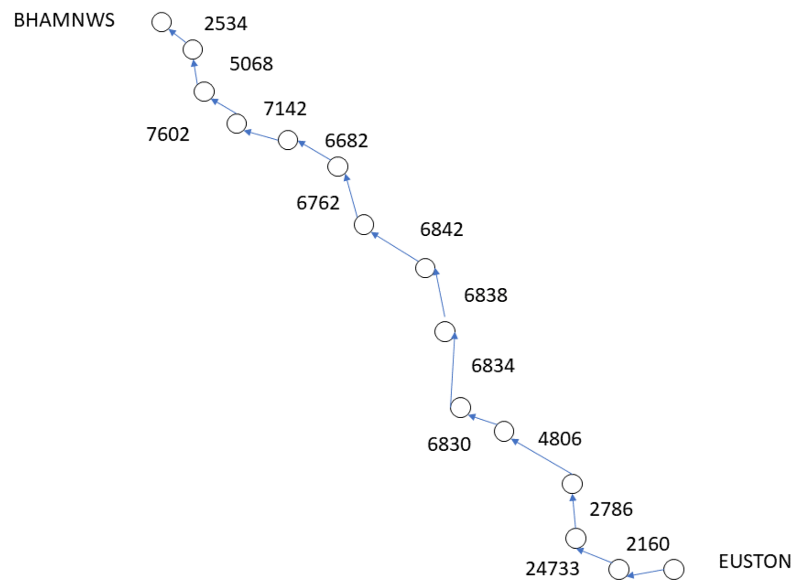


Figure 31: Graph from EUSTON to BHAMNWS- daily passenger flow is estimated using our spatial model.

For our estimation model, we can update it by considering local economy development over the past 9 years to give a higher accuracy estimation. We can combine the travel restriction rules on the train to give the COVID period travel flow. Combining those information and students or targeted passenger group travel routes, it is possible to build further optimisation model to understand effects of different travel strategies.

5.6 Future work

The ideas in this section are the starting point for developing a set of models to advise train operators and higher education (HE) providers to organise the return of students to campus. Similarly, it could be used to plan movements for large events (half-term holidays for example).

For the HE application, we would need additional data that allows us to estimate the number of students that would use a particular station (see below for possible sources). This can provide us the basis to set up an optimisation model to adjust the arrival of students to alleviate the stress on the rail network.

Such a model could also be used, on a regional level, to plan commuter traffic. It might also be of interest to investigate where additional trains might be required to deal with unusual demand.

Data

- Changes to Flow with Covid-19: While we have been working with 2011 commuting figures, rather than current figures during the pandemic, we could improve this either by using up-to-date data (perhaps estimated from ticket sales or other counts / data), or we could adapt pre-Covid-19 figures using mobility-decrease numbers, e.g. those shown in Fig. 32.

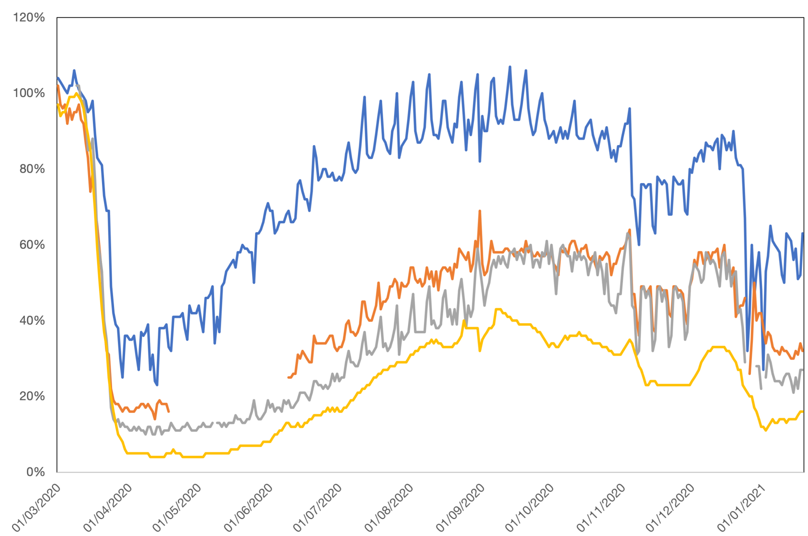


Figure 32: Changes in use of transport modes (percentage of an equivalent day or week) From (5)

- FlowCensus data: Available mainly for work and study (6)
  - Non-safeguarded at OA in England and Wales, local authority in Scotland
  - Safeguarded at OA/WPZ in Scotland from WICID
  - Available sliced by age, type of work, method of travel, hours worked, social grade, availability of car or van etc.
  - Most travel to school is within area, remainder low volume compared to travel for work

  We used an extract at 2011-local authority level, subsetting only the movements by rail.

- Student OD data: From Higher Education Statistics Authority (HESA) data (7).

- Modelling student demand: Better modelling of student demand that would allow calculation of required spread over time or extra capacity needed is possible, but requires several additional pieces of information:

- data (with appropriate permissions) from HESA on the domicile locations of students and their HE provider - ideally for this year's students, and at a reasonably fine geographic scale.
  * These data exist (can be obtained via a data request from HESA) but are not currently available for this VSG.
- indications of how many students are likely to return by HE provider, and how many have previously returned as part of the initial back-to-campus groups (medics, vets, etc). There have been some surveys that may help: for example (9) in which 53% of students said they would be either 'Likely' or 'Extremely Likely' to return to term-time accommodation even if all teaching was entirely online.
  * Limited data available - perhaps HE providers have some additional data?
- indications of how many travelling students would be likely to travel by rail if guidance suggested travel by private vehicle if possible.
  * Data availability unknown - perhaps rail providers know how many students returned to their domicile by rail?
- indications of the likelihood that students would obey differing types of guidance on return: in particular would they be willing to return on a particular day or service if they were asked to?
  * Limited data available - could model over different assumptions, or use estimated compliance with other interventions. Could vary significantly by HE provider.

5.7 Conclusions and next steps

We show here a modelling methodology which could be used to optimise capacity of rail operators when dealing with the Covid-19 pandemic. In particular, we focus on how we might advise operators on how best to safely manage the return-to-campus challenge many HE providers will need to solve. We highlight here that data availability is key, and some thoughts on the data requirements are presented. For future work, additional data and access to route-producing APIs / code is key. The essential things we need to know to proceed are:

- how many students would want to travel between particular origins and destinations?
- what routes would they use?
- does this exceed capacity under different spreading schemes?

Work to continue this modelling is underway at the University of Bath.

References

[1] Railway names data and locations. GitHub accessed Jan 2021 `https://gist.github.com/crablab/93a50eeb338646614287eddc3c2776b1`

[2] Where do HE students study? HESA data accessed Jan 2021 `https://www.hesa.ac.uk/data-and-analysis/students/where-study`

[3] Jovanović, P., N. Pavlovic, Ivan Belosevic and Sanjin Milinkovic. “A Graph Application for Design and Capacity Analysis of Railway Junctions.” (2019).

[4] Illustration of train allocation complexity `https://www.youtube.com/watch?v=gTLVbCN0QA8&feature=youtu.be`

[5] Transport use during the coronavirus (COVID-19) pandemic. GOV.UK Official Statistics `https://www.gov.uk/government/statistics/transport-use-during-the-coronavirus-covid-19-pandemic`

[6] Census Support: Flow Data `https://wicid.ukdataservice.ac.uk/`

[7] Where do HE students come from? HESA data accessed Jan 2021 `https://www.hesa.ac.uk/data-and-analysis/students/where-from`

[8] ONS Dataset: 2011 Census: Detailed characteristics on travel to work and car or van availability for local authorities in England and Wales `https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/2011censusdetailedcharacteristicsontraveltoworkandcarorvanavailabilityforlocalauthoritiesinenglan`

[9] Coronavirus and university students, 12 October to 18 October 2020, England GOV.UK Official Statistics `https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/adhocs/12449coronavirusanduniversitystudents12octoberto18october2020england`

6 Conclusions

Carriage-level

- There is scope for low dimensional, low fidelity models to complement the high dimensional, high fidelity models under development as part of the TRACK project.
- Combining these models with simple flow models will allow spatial and temporal variations to be included without a dramatic increase in (computational) complexity.
- As these simpler models are much faster to run, they can be particularly useful for "What If?" scenario testing e.g. for estimating the effects of various interventions.
- The implementation could take the form of a simple dashboard or spreadsheet on the vendor side (for risk management) and / or mobile phone app on the user side (for streamlining communication and reducing transmission channels).

Journey-level

- The loading platform model could be extended to accommodate more realistic loading conditions and passenger movement to provide insight into where crowding could occur.
- The model for queuing within a carriage could be examined further and Incorporated with other modelling work in this report to give a quantification of viral exposure during a train journey.

Network-level

- Data driven approaches could be very informative for modelling the impact of non-pharmaceutical interventions and large scale policy decisions which affect peoples travelling option.
- Integrating data from various sources, and designing network models is an attractive way to provide insight into this challenge.

7 List of Acronyms

CFD computational fluid dynamics

FE finite element

FEM finite element method

HE higher education

HESA Higher Education Statistics Authority

HVAC heating ventilation & air conditioning

ODE ordinary differential equation

PDE partial differential equation

RH relative humidity

VSG Virtual Study Group

TRACK Transport Risk Assessment for COVID Knowledge

URANS unsteady Reynolds-averaged Navier-Stokes

UVC ultraviolet C

Connecting for
Positive Change

Head Office

KTN
Suite 218, Business Design Centre
52 Upper Street
Islington
London N1 0QH

Telephone: 03333 403251
Email: enquiries@ktn-uk.org
ktn-uk.org
@KTNUK

Contact Person
Dr Matt Butchers
matt.butchers@ktn-uk.org