Mathematical models to support Victoria's COVID-19 response: a blunt instrument to a complex problem

Michael J. Lydeamore[1,2]

[1]Victorian Department of Health, Government of Victoria, Melbourne, Australia

[2]Department of Infectious Diseases, The Alfred and Central Clinical School, Monash University, Melbourne, Australia

July 23, 2021

Throughout Victoria's COVID-19 response, a suite of mathematical and statistical models have been used to understand the spread and subsequent control of the pandemic. High-profile mathematical outputs, such as case forecasts, give a good picture of the general epidemic activity in a given region. However, when case numbers are small, forecasts can be very sensitive. Moreso, when there is only a handful of cases, more detailed factors such as geographic distribution of cases, or at-risk industries, are possibly more informative.

As part of my role in Victoria's COVID-19 response, my team applied a number of statistical techniques to case data. These techniques have varying levels of sophistication, but one of the most usedwas Diggle's space-time K-function [2]. This model is relatively unrefined, but the advantage of that is that very little information is required to compute it: just the date of infection of cases, and their geographical location. Both of these are collected almost immediately once the case is notified, meaning that this function can be calculated regularly and quickly, two factors that are critical in informing epidemic response.

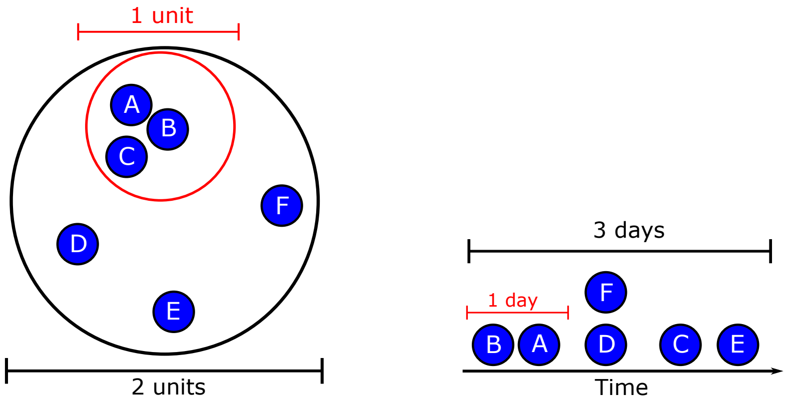
The D_0 function can be interpreted as the proportional increase in case events at a given space-time arising from interactions at that space-time. In an infectious diseases context, this is a proxy measure of disease contagiousness. The lower this increase, the stronger the indication of successful intervention measures.

To calculate the D_0 function, we start with the K-function[1]. The K-function is defined as the cumulative number of expected case events, K, as a function of the (straight-line) distance from an arbitrarily selected case [7]. That is,

$$K_d(s) = N^{-1} \sum_i \sum_{j \neq i} I[d_{ij} < s], \tag{1}$$

where N is the *total* number of cases, d_ij is the distance between case i and case j, and I is an indicator function. Eq (1) is sometimes termed the 'spatial' K-function, but by

[1]For some reason, spatial science has some of the most non-transparent function naming



(a) (b)

Figure 1: Graphical illustration of the K-function in (a) space and (b) time.

swapping out the d_{ij} term, it is possible to calculate a K-function across any attribute. For COVID-19, we used *time*, denoted $K_t(t)$, as well as space. The time between two cases was the number of days between their onset dates (which we assumed were a proxy for infection dates) as opposed to the notification dates, for which there was sometimes a long delay. Figure 1 gives an illustration of calculating the K-function in (a) space and (b) time. For the arbitrarily chosen case – case B here – the K-function in space at 1 unit, $K_d^B(1) = 2/6$, and at 2 units, $K_d^B(2) = 5/6$. Comparatively, in time, $K_t^B(1) = 1/6$, and $K_t^B(3) = 5/6$. To estimate the K-function in it's entirety, we would repeat this process for each of the other cases A–F.

So far, we have considered space and time completely separately. As the last example has shown, often the clustering of cases in space and time can be different. If space and time were completely independent, then the space-time K-function, denoted $K(s,t)$, would be the product of the space and time K-functions. That is,

$$K(s,t) = K_d(s) \times K_t(t).$$

However, that is rarely the case, particularly in infectious diseases, where disease spreads from one individual to another. In the example visualised in Figure 1, $K(1,1) = 1/36$ as cases A and B are within 1 unit and 1 day of each other [2]. However, $K_d(1) \times K_t(1) = 2/6 \times 1/6 = 2/36$ [3]. Because of this dependence between space and time, we are interested in estimating how many times greater $K(s,t)$ is compared to the product of $K_d(s)$ and $K_t(t)$. Thus, we arrive at the definition of D_0 function,

$$D_0(s,t) = \frac{K(s,t)}{K_d(s)K_t(t)}. \tag{2}$$

The actual value of the D_0 function is not particularly important. Rather, it is how this function changes over time or geographically that is the most important. The absence of space-time interaction (i.e. a relatively flat D_0 function) is a sign of control success.

From March until the end of 2020, we estimated the D_0 function as part of the routine reporting framework. It was regularly one of the fastest measures to show control success,

[2] Note that we have N^2 here as we are considering the two dimensions

[3] Note that we have labelled the cases here for the purposes of the example, but when actually calculating these functions the cases are considered unlabelled.

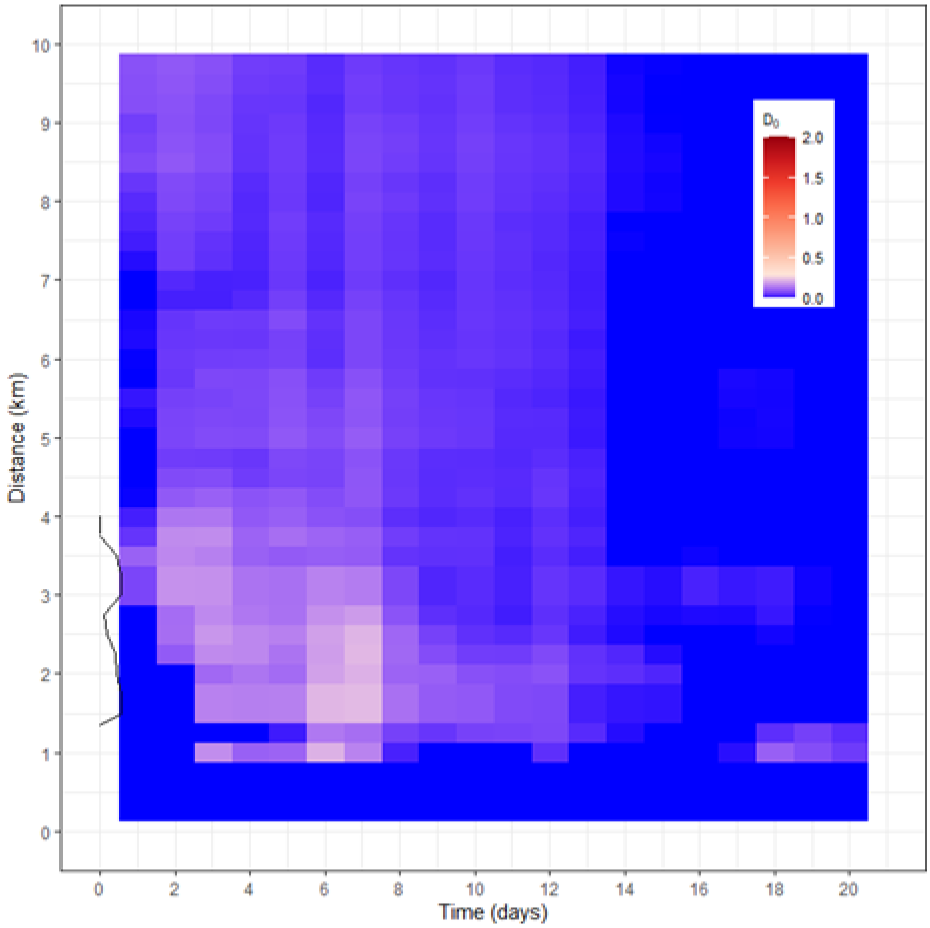


Figure 2: Estimate of the D0 function based on COVID-19 data up to June 20, 2020 in Victoria. The contagiousness estimated by this function is very low, despite signs of increasing case numbers.

but perhaps one of the most valuable moments was the reporting in mid-June. The period from May through to the start of June was one of relative calm, with 218 cases diagnosed over the 31 days, an average of roughly 7 notifications per 24 hours. A similar story was true at the start of June, with 114 community acquired or unknown source cases over the first 20 days: an average of under 6, relatively the same as what we had so far. In our routine report, using data up to June 20, 2020, something seemed amiss with the D_0 function, reproduced in Figure 2. What was once a nicely clustered mass near 0 days and 0 kilometres seemed to have spread out, particularly over distance. There was, unusually, relatively little clustering to be seen.

After digesting this figure, we got together with the epidemiology team who had been compiling notes on all the cases and their contacts across Victoria. When we pooled all the evidence together – statistical models, case notes, forecasts of incoming numbers, genomic information, geographic risk profiles, and so much more – we arrived at a hypothesis we hoped would not be true: infection had been scatter-gunned across the greater Melbourne area. The more we looked, the more it seemed like it could be true. In the 10 days that followed, a further 369 community acquired or unknown cases would be notified, more than 6 times as many cases per day than that of the previous 3 weeks. Victoria's second wave had arrived.

The D_0 function is not a new concept to infectious diseases. Although not typically applied in human diseases (as thankfully we don't have many pandemics), it has been used in Rift Valley fever [6], highly pathogenic avian influenza [4, 5] as well as a handful of others [1, 3, 8]. The technique is well-known, and in an environment where time is of the essence, it's quick compute time was proven valuable. COVID-19 was, and is, an infection that we knew little about, and the ability to apply a tried and tested model when it was needed most meant that our ability to respond was as strong as it could be.

There is no crystal ball when it comes to pandemic predictions. No matter how complex the model we develop and apply, no one can ever accurately predict the future. Much of my and my team's role was to synthesise the information from these relatively unsophisticated models and communicate them to the people who needed to know. In the example discussed here, it was the discussion of information with our epidemiologists that led to the conclusion, not a piece of data or a model. No-one knows when exactly the peak of infections will be, but between the epidemiology and the modelled data, we can come together and give an idea of whether we're likely to see increases or if the control measures being applied are working.

Victoria has since reached a state of elimination for COVID-19, along with the rest of Australia. A feat shared only by a few globally. Here, we have seen one example of a response coming together to solve an issue, but it is far from the only example. If there's one thing to take away from Victoria's COVID-19 response, it's that the pieces of the puzzle are always stronger together.

Acknowledgements: The author would like to thank Mark A. Stevenson and Kira Leeb for their comments on this article, as well as the Victorian Department of Health Analytics Team (COVID-19 Intelligence) for their work throughout the pandemic.

References

[1] E. Delmelle, I. Casas, J. H. Rojas, and A. Varela. Spatio-Temporal Patterns of Dengue Fever in Cali, Colombia. *International Journal of Applied Geospatial Research (IJAGR)*, 4(4):58–75, Oct. 2013.

[2] P. Diggle, A. Chetwynd, R. Häggkvist, and S. Morris. Second-order analysis of space-time clustering. *Statistical Methods in Medical Research*, 4(2):124–136, June 1995.

[3] A. C. Gatrell, T. C. Bailey, P. J. Diggle, and B. S. Rowlingson. Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. *Transactions of the Institute of British Geographers*, 21(1):256–274, 1996.

[4] C. Guinat, G. Nicolas, T. Vergne, A. Bronner, B. Durand, A. Courcoul, M. Gilbert, J.-L. Guérin, and M. C. Paul. Spatio-temporal patterns of highly pathogenic avian influenza virus subtype H5N8 spread, France, 2016 to 2017. *Eurosurveillance*, 23(26):1700791, June 2018.

[5] L. Loth, L. T. Pham, and M. A. Stevenson. Spatio-temporal distribution of outbreaks of highly pathogenic avian influenza virus subtype H5N1 in Vietnam, 2015– 2018. *Transboundary and Emerging Diseases*, 68(1):13–20, 2019.

[6] R. Métras, T. Porphyre, D. U. Pfeiffer, A. Kemp, P. N. Thompson, L. M. Collins, and R. G. White. Exploratory Space-Time Analyses of Rift Valley Fever in South Africa in 2008–2011. *PLOS Neglected Tropical Diseases*, 6(8):e1808, Aug. 2012.

[7] B. D. Ripley. The Second-Order Analysis of Stationary Point Processes. *Journal of Applied Probability*, 13(2):255–266, 1976.

[8] J. W. Wilesmith, M. A. Stevenson, C. B. King, and R. S. Morris. Spatio-temporal epidemiology of foot-and-mouth disease in two counties of Great Britain in 2001. *Preventive Veterinary Medicine*, 61(3):157–170, Nov. 2003.