

Species distribution models are inappropriate for COVID-19

Species distribution models are a powerful tool for ecological inference, but not every use is biologically justified. Applying these tools to the COVID-19 pandemic is unlikely to yield new insights, and could mislead policymakers at a critical moment.

Colin J. Carlson, Joseph D. Chipperfield, Blas M. Benito, Richard J. Telford and Robert B. O'Hara

Since the turn of the century, species distribution models (SDMs) have become a foundational tool in an ecologist's toolkit. At their core, SDMs are a statistical tool to make the concept of 'habitat' tractable with high-dimensional digital data; through regression or machine learning, a naturalist can identify the signal of ecological constraints on where species are found. Those signals tell us how to partition a complex n -dimensional space into habitat suitability: for every free-living animal and plant, some places are suitable, and others are not.

Increasingly, SDMs have found their way into medical geography. Many epidemiological questions have ecological answers. For example, most emerging diseases originate in a wild animal, and many are driven by ecological processes like deforestation, warming temperatures or biological invasions^{1–3}. SDMs can help a virologist or veterinarian determine which areas have measurable infection risk, and which are safely outside the habitat of the parasite or its hosts. That lets public health experts answer critical questions: how many people or livestock are at risk, and how many are safe? What determines the spatial pattern of risk? Will it change if the world changes?

By adapting SDM workflows for the mapping of infectious diseases, ecologists have made a massive contribution to public health^{4–6}. Most pathogens are as poorly mapped as any given plant or animal, and sampling is often biased in similar ways, making it hard to develop accurate risk maps. This is particularly true for zoonotic and vector-borne diseases — those that come from animals, or are spread among humans by them — a category of infections that are disproportionately neglected by global health surveillance⁷. By harnessing ecological predictors with high confidence, and identifying mechanisms connecting ecology to disease, ecologists have helped fill surveillance gaps that otherwise limit public health research. A model can tell

researchers where soil is too acidic for the spores of the anthrax bacterium⁸; where cold temperatures prevent flaviviruses such as dengue or Zika from replicating inside *Aedes* mosquitoes^{9,10}; or even which undersampled forests in the Democratic Republic of Congo contain the bats (of still uncertain identity) that are hosts for Ebola viruses¹¹.

It should be unsurprising then that during an epidemic, ecologists would be eager to apply SDM methods as part of an outbreak response. A shift towards open science, rapid turnaround of interdisciplinary work, and preprint servers have helped ecologists find a seat at the table. During the 2016 Pan-American epidemic of Zika virus, three teams produced rapid-turnaround projections of the full possible spread of transmission^{10,12,13}. Though these models differed greatly in their projections, they became an important and influential part of outbreak response, especially as the United States planned for eventual transmission in the Gulf Coast¹⁴. A similar conversation is unfolding now in response to the COVID-19 pandemic, with multiple studies applying an SDM framework to try and understand the role climate plays in SARS-CoV-2 transmission. These studies have raised a basic question for SDM experts: does a respiratory virus have an ecological niche, and if so, can it be mapped?

We caution that SDM methods are not appropriate for all pathogens. Their biological meaning and methodological relevance increases with the level of environmental control over a pathogen or parasite's life cycle. Metazoan parasites with a free-living stage, such as ticks or tapeworms, have the most readily conceptualized ecological niche, while microparasites (bacteria, viruses and prions) are a greater challenge. For a vector-borne virus, geographic limits of transmission might be set by the strength of temperature effects on a mosquito biting rate¹⁵, or of relative humidity on tick desiccation¹⁶. For a virus with multiple transmission routes, this

can become complicated. SDM approaches can identify where mosquitoes cannot transmit Zika virus, but that risk map will always bear an asterisk: sexual transmission of the virus can happen anywhere, including in places far outside the niche of the mosquito¹⁷. SDMs cannot capture the possible extent of that transmission from occurrence data. They are simply the wrong approach.

This does not mean SDMs cannot be run. Machine-learning methods will almost always be able to pull statistically significant signals from predictors and outcomes with spatial structure, even when the predictors have no causal link to outcomes¹⁸. Even if there are relationships, SDMs are never the right approach if every area is intrinsically suitable habitat. As an example, there are long-term indirect effects of weather through socioeconomic risk factors on HIV transmission in Africa¹⁹. A determined ecologist could generate an SDM from a geolocated database of HIV cases, and with any given methodology, generate a risk map with high accuracy that shows Africa has more suitable climates for HIV transmission, or that the Arctic and Antarctic are too cold for risk. The mistake in the approach is not investigating climate–disease relationships, but using the wrong tool — one assuming that the underlying truth is a binary split in risk, rather than a multi-factoral, epidemiological mess. However, misapplying the tool will lead to the wrong conclusions about climate.

Direct transmission — communication of a virus from an infected host to an uninfected one, without a long-term environmental stage²⁰ — bypasses the relevance of 'unsuitable habitat' as an ecological concept. As virologists currently understand the transmission of SARS-CoV-2 and other directly transmitted viruses, there are no absolute climatic limits to their transmission. Once an outbreak starts spreading among humans, there is no unsuitable habitat.

This is not to say that climate drivers are not relevant to processes like aerosol transmission, but instead, that their signal is a partial one, confounded by human behaviour and the microclimate of built environments. Influenza offers a useful comparison point: temperature and relative humidity are strong drivers of seasonal influenza outbreak severity, duration and geography²¹. These influences are still poorly understood and hotly contested, and exploring them has required specialized epidemiological models^{22,23}, usually ones that examine incidence (case totals and severity over time and space) rather than geographic occurrence (binary presence/absence). The SDM approach, as a tool for partitioning landscapes into at-risk and not-at-risk, generates the wrong answer to a trivial question: influenza remains transmissible anywhere in the world, between any two people.

The same is true of SARS-CoV-2. There are probably environmental influences on the virus that might look, to a biologist, like a niche. Like other respiratory viruses, it is possible that relative humidity has an effect on transmission. Like other viruses with a lipid envelope, SARS-CoV-2 is able to persist in the environment for different durations on different surfaces²⁴. These lines of reasoning have been cited by SDM studies as a potential underpinning to a climate–COVID-19 relationship, but few ecological approaches are built to explore microclimate at the right scale: contact tracing data from China indicate that over 99% of known transmission occurred indoors²⁵. Even if long-term aerosol or surface persistence exists as a brief environmental stage influenced by microclimate, the absence of suitable microclimate is not prohibitive: two humans in contact will always be able to spread the virus. An SDM approach, which splits the world into suitable or unsuitable outdoor climates for transmission, will always miss the ubiquity of both human contact and human built environments, and correspondingly identify false patterns.

There may well be an impact of microclimate on viral transmission, as there is for other respiratory viruses, but this finding has been misinterpreted (wilfully or not) by policymakers as rationale to lift social-distancing policies by summer. Potential for even greater harm exists, if governments in Africa or Latin America decide to act on SDM-based studies finding that climate could protect tropical countries from transmission.

The stakes are high, and SDM approaches are the wrong tool for the

job. Epidemiological work using models that track transmission dynamics within populations will likely better answer the same questions within the next few months. In the interim, it is critical that ecologists do not contribute to misunderstanding of risk²⁶. We suggest that scientists can point the public to the recently released National Academies panel's guidance²⁷, a resource which captures scientific consensus on a few key points. In particular, a true climate–COVID-19 relationship may exist, but in the pandemic phase of transmission, the effect of weather will be minimal compared to human contact and outbreak response measures (pandemic influenza, as a useful comparison, has displayed little of the climate-driven patterns that seasonal influenza does), and developing policy based on potential protective effects of weather could lead to disaster.

That is not to say there is no place for species distribution modelling in coronavirus research. SDMs have been helpful for coronaviruses and other directly transmitted zoonoses, mainly in accounting for where virus is detected in reservoirs, or where spillover into humans occurs^{28,29}. These are processes with a much clearer set of drivers. For example, the density of intermediate livestock hosts (poultry for influenza, camels for MERS-CoV), or seasonal vegetation and precipitation drivers explaining the distribution and health of waterfowl. For SARS-CoV-2, a novel virus originating in a (presumed) single spillover event with an unknown geographic location³⁰, these approaches are not applicable yet. But serological data have indicated that SARS-like coronaviruses have spilled over into human populations several times, and in conjunction with virological data, these may be enough to start developing risk maps using SDM frameworks. In the meantime, the virus is still circulating somewhere in a wildlife host. Advancing basic biogeography and bioinformatics in Asia, and having high-confidence, high-resolution maps of wildlife species distributions, will be a key first step to mapping the risk of future coronavirus outbreaks. □

Colin J. Carlson ¹✉, Joseph D. Chipperfield ², Blas M. Benito³, Richard J. Telford ⁴ and Robert B. O'Hara ⁵

¹Center for Global Health Science and Security, Georgetown University Medical Center, Washington, DC, USA. ²Norwegian Institute for Nature Research, Bergen, Norway. ³Department of Ecology & Multidisciplinary Institute for Environment Studies “Ramon Margalef”, University of Alicante, Alicante,

Spain. ⁴Department of Biological Sciences, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway. ⁵Department of Mathematical Sciences and Centre for Biodiversity Dynamics, Norwegian University of Science and Technology, Trondheim, Norway.

✉e-mail: colin.carlson@georgetown.edu

Published online: 6 May 2020
<https://doi.org/10.1038/s41559-020-1212-8>

References

- Jones, K. E. et al. *Nature* **451**, 990–993 (2008).
- Smith, K. F. et al. *J. R. Soc. Interface* **11**, 20140950 (2014).
- Allen, T. et al. *Nat. Commun* **8**, 1124 (2017).
- Hay, S. I. et al. *Phil. Trans. R. Soc. B* **368**, 20120250 (2013).
- Kraemer, M. U. et al. *Trends Parasitol.* **32**, 19–29 (2016).
- Johnson, E. E., Escobar, L. E. & Zambrana-Torrel, C. *Trends Ecol. Evol.* **34**, 655–668 (2019).
- Maudlin, I., Eisler, M. C. & Welburn, S. C. *Phil. Trans. R. Soc. B* **364**, 2777–2787 (2009).
- Carlson, C. J. et al. *Nat. Microbiol.* **4**, 1337–1343 (2019).
- Bhatt, S. et al. *Nature* **496**, 504–507 (2013).
- Messina, J. P. et al. *eLife* **5**, e15272 (2016).
- Pigott, D. M. et al. *eLife* **3**, e04395 (2014).
- Carlson, C. J., Dougherty, E. R. & Getz, W. *PLoS Negl. Trop. Dis.* **10**, e004968 (2016).
- Samy, A. M., Thomas, S. M., Wahed, A. A. E., Cohoon, K. P. & Peterson, A. T. *Mem. Instit. Oswaldo Cruz* **111**, 559–560 (2016).
- Carlson, C. J., Dougherty, E., Boots, M., Getz, W. & Ryan, S. J. *Sci. Rep.* **8**, 4921 (2018).
- Shapiro, L. L., Whitehead, S. A. & Thomas, M. B. *PLoS Biol.* **15**, e2003489 (2017).
- Rodgers, S. E., Zolnik, C. P. & Mather, T. N. *J. Med. Entomol.* **44**, 372–375 (2007).
- Foy, B. D. et al. *Emerg. Infect. Dis.* **17**, 880–882 (2011).
- Fourcade, Y., Besnard, A. G. & Secondi, J. *Glob. Ecol. Biogeogr.* **27**, 245–256 (2018).
- Burke, M., Gong, E. & Jones, K. *Econ. J.* **125**, 1157–1189 (2015).
- Dougherty, E. R., Seidel, D. P., Carlson, C. J., Spiegel, O. & Getz, W. M. *Ecol. Lett.* **21**, 588–604 (2018).
- Lowen, A. C. & Steel, J. J. *J. Virol.* **88**, 7692–7695 (2014).
- Marr, L. C., Tang, J. W., Van Mullekom, J. & Lakdawala, S. S. *J. R. Soc. Interface* **16**, 20180298 (2019).
- Park, J.-E. et al. *Influenza Other Respir. Viruses* **14**, 11–18 (2020).
- van Doremalen, N. et al. *N. Engl. J. Med.* **382**, 1564–1567 (2020).
- Qian, H. et al. Preprint at *medRxiv* <https://doi.org/10.1101/2020.04.04.20053058> (2020).
- Gog, J. R. *Nat. Rev. Phys.* <https://doi.org/10.1038/s42254-020-0175-7> (2020).
- National Academies of Sciences, Engineering, and Medicine *Rapid Expert Consultation on SARS-CoV-2 Survival in Relation to Temperature and Humidity and Potential for Seasonality for the COVID-19 Pandemic (April 7, 2020)* (The National Academies Press, 2020); <https://doi.org/10.17226/25771>
- Zhu, G. & Peterson, A. T. Potential geographic distribution of the novel avian-origin influenza A (H7N9) virus. *PLoS ONE* **9**, e93390 (2014).
- Reeves, T., Samy, A. M. & Peterson, A. T. *BMC Res. Notes* **8**, 801 (2015).
- Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. *Nat. Med.* **26**, 450–452 (2020).

Acknowledgements

We thank B. Muscarella for helpful comments.

Author contributions

C.J.C., J.D.C., B.M.B., R.J.T. and R.B.O. all contributed to the writing of the manuscript.

Competing interests

The authors declare no competing interests.