

(Network Diffusion Library [3]), a python package we designed to model diffusive phenomena.

NDlib represents a multi-level solution to simulate the spread of epidemics. We organised the library in three incremental modules: the NDlib core library (written in Python), a remote RESTful experiment server accessible through API calls and, finally, a web-oriented visual interface.

At the core of our tool, is NDlib, a Python package built upon the network facilities offered by NetworkX and iGraph. NDlib models diffusive phenomena as discrete-time agent-based processes. It implements several epidemics and opinion dynamics models (around 20) and allows its users to design novel ones. To allow the user to easily analyse simulation evolution, NDlib also exposes a set of visual facilities. The initial set of infected nodes has a relevant impact on how a diffusive process unfolds. For this reason, with a single line of code, NDlib allows execution to occur several times in parallel: the simulation of the same model over a given network topology while varying the initial infection status. All the diffusion models implemented in NDlib extend the same template, which makes it easy to extend the NDlib library.

As with classical analytical tools, the simulation facilities offered by NDlib are specifically designed for users who

want to run experiments on their local machine. However, in some scenarios—e.g., due to limited computational resources or to other needs arising—it may be convenient to separate the machine on which the definition of the experiment is made from the one that actually executes the simulation. To satisfy such needs, we developed a RESTful service, NDlib-REST, that builds upon NDlib with an experiment server, queryable through API calls.

Finally, upon the NDlib-REST service is a visualisation platform. NDlib-Viz ([L1], Figure 1) aims to allow non-technicians to design, configure and run epidemic simulations, thus removing the barriers introduced by the usual requirements of programming language knowledge. Indeed, apart from the usual research-oriented audience, we developed NDlib-Viz to support students and facilitate teachers to introduce epidemic models: in order to better support the educational nature of the proposed platform and collect user feedback, we currently employ NDlib-Viz in a Social Analysis course of the Computer Science Master Degree at the University of Pisa, Italy. The platform itself is a web application: it can be executed on a local as well as on a remote NDlib-REST installation.

The proposed library aims to provide a reasonable trade-off between ease of use and efficiency. NDlib is released as free software under a BSD-2-Clause

license and the NDlib project is hosted on GitHub ([L2]), documented on ReadTheDocs ([L3]) and shipped through Pypi ([L4]). For advertising purposes, we released a promo video hosted on YouTube ([L5]).

Finally, NDlib is also indexed in the SoBigData catalogue ([L6]). SoBigData is an EU H2020 project that promotes the integration of already existing software infrastructures.

Links:

[L1] <https://goo.gl/tYi48o>

[L2] <https://kwz.me/h2M>,
<https://kwz.me/h2Q>

[L3] <http://ndlib.readthedocs.io/>

[L4] <https://pypi.python.org/pypi/ndlib/>

[L5] <https://youtu.be/tYHNOuKJwbE>

[L6] <http://www.sobigdata.eu>

References:

[1] L. Ahrenberg, et al.: *NepidemiX*. Technical report, 2016.

[2] S. Jenness, S.M. Goodreau, M. Morris: “Epimodel: Mathematical modeling of infectious disease”, package version 1.3.0., 2017. <http://www.epimodel.org>

[3] G. Rossetti, et al.: “NDlib: a python library to model and analyze diffusion processes over complex networks”, *International Journal of Data Science and Analytics*, 5(1), 61-79.

Please contact:

Letizia Milli, University of Pisa, Italy
milli@di.unipi.it

Predicting the Spread of COVID-19 through Marine Ecological Niche Models

by Gianpaolo Coro (ISTI-CNR)

Researchers from ISTI-CNR (Italy) used marine models, designed to monitor species habitats and invasions, to identify the countries with the highest risk of COVID-19 spread due to climatic and human factors. The model correctly identified most locations where large outbreaks were recorded, independent of population density and dynamics, and is a valuable source of information for smaller-scale population models.

The correlation between the reduction of the spread of COVID-19 and season change was foreseen at the beginning of the 2020 pandemic and has been observed during the Northern Hemisphere summer. The models that forecast this correlation ranged from dynamic systems (that constantly mon-

itor the infection's spatio-temporal spread) and general models that aim to identify global similarities between areas strongly affected by the virus. However, while dynamic systems are still struggling to produce new recommendations to reduce the spread of infection, general models are informing

scientists and decision-makers about large-scale trends and ideal conditions for infection. These models can match patterns at a large temporal and spatial resolution and can provide information about the general conditions that support the virus's persistence, independently of social dynamics. Patterns

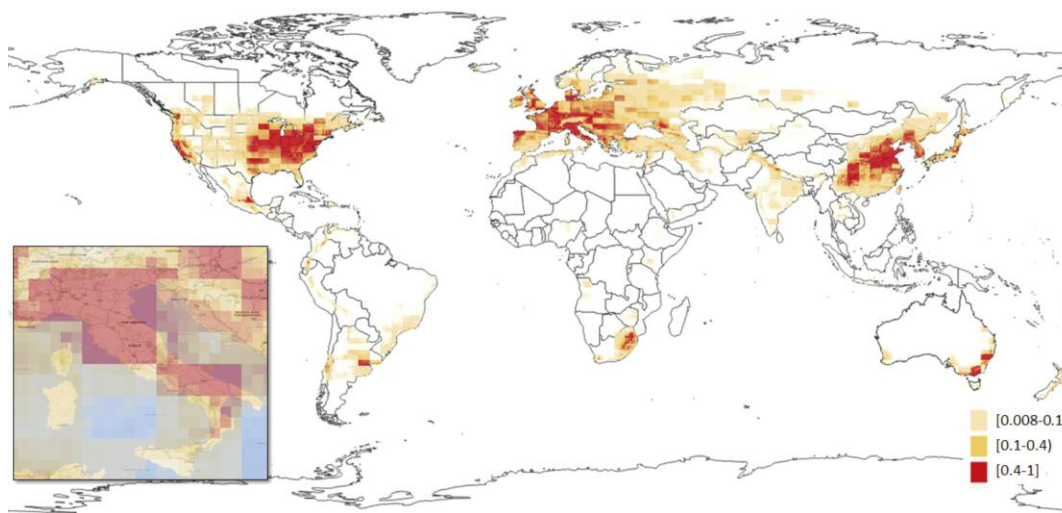


Figure 1: Global-scale probability distribution of the COVID-19 infection rate produced by Coro's (2020) model, with Italy magnified at the lower-left hand side (situation of March 2020).

related to temperature variation have been evident in some parts of the world, and have been used (and sometimes abused) by political authorities to modify regulations.

Among the general models used to explore correlations, ecological niche models (ENMs) have been extensively and effectively used to predict the spread of viruses, including SARS-CoV-2. Generally, an ENM estimates the presence of a particular species in a geographical area by correlating species-specific occurrence records in its native habitat with environmental parameters. The species' niche is the space within a hypervolume of numerical vectors – corresponding to environmental parameter ranges – that is correlated with the species' presence and that fosters the species' persistence in an area. Once a model has estimated the species' ecological niche, it can then project it across a new geographical area to identify other locations that provide ideal conditions for the species.

While COVID-19 is capable of persisting in a range of environments, some populated regions experience lower infection rates than others. Studies using ENMs have pinpointed very interesting similarities between global regions with the highest infection rates. One of these used a Maximum Entropy-based Ecological Niche Model [1] to estimate the COVID-19 infection rate (disease spreading speed) globally at a half-degree spatial resolution. The model identified geographical areas that favour higher infection rates due to their particular geophysical (surface air temperature, precipitation and elevation)

and anthropogenic characteristics (CO₂ and population density). Relying solely on data from Italian provinces that reported high infection rates in March 2020, the model was then used to predict global regions that were likely to experience high infection rates. Surprisingly, the model highlighted strong similarities between the climatic and anthropogenic characteristics of locations where disease outbreaks have really occurred, e.g. the Hubei province in China, the Western United States, and Europe, and correctly foresaw a high risk of fast disease increment in 77% of the countries that have then reported significant outbreaks. This experiment has highlighted that certain parameter combinations, such as a temperature range of 11 to 12 °C, moderate humidity, and a high level of air pollution, are strongly correlated with a high infection rate. Once above a low threshold, this correlation is also independent of population density. The model also indicates that in some regions (e.g. Brazil, Ecuador and Peru), the high infection rate cannot be explained in terms of the identified environmental factors, with population dynamics possibly being more important than climatic conditions. Overall, this kind of ENM can inform fine-grained dynamic models about which areas have the conditions to be a catalyst of the infection.

One interesting aspect of the ENM [1] is that the model comes from the marine science domain, where it had been used to predict the spread of invasive species [2] and the potential distribution of rare species [3]. The model was directly reused to predict COVID-19 infection

rates because it was based on findable, accessible, interoperable, and reusable (FAIR) data and on an Open Science-oriented methodology that standardised the input, output, and the processing, and thus enhanced the reusability of the model.

The challenge now is to introduce these models into larger decision support systems to increase their prediction accuracy and resolution, and to estimate the risk for individuals to end up in intensive care, based on background data about their environmental and health conditions.

Link:

[L1] Suitability Map of COVID-19 Virus Spread – Experimental Data and Results <https://kwz.me/h2g>

References:

- [1] G. Coro: "A global-scale ecological niche model to predict SARS-CoV-2 coronavirus infection rate", *Ecological Modelling*, 431, 109187, 2020.
- [2] G. Coro, et al.: "Forecasting the ongoing invasion of *Lagocephalus sceleratus* in the Mediterranean Sea", *Ecological Modelling*, 371, 37-49, 2018.
- [3] G. Coro, C. Magliozzi, A. Ellenbroek, P. Pagano: "Improving data quality to build a robust distribution model for *Architeuthis dux*", *Ecological modelling*, 305, 29-39, 2015.

Please contact:

Gianpaolo Coro
ISTI-CNR, Italy
gianpaolo.coro@isti.cnr.it