frontiers | Frontiers in Applied Mathematics and Statistics

TYPE Editorial
PUBLISHED 24 January 2023
DOI 10.3389/fams.2023.1129749

OPEN ACCESS

EDITED AND REVIEWED BY
Tim Rogers,
University of Bath, United Kingdom

*CORRESPONDENCE
Jacques Demongeot
✉ Jacques.Demongeot@univ-grenoble-alpes.fr

SPECIALTY SECTION
This article was submitted to
Mathematical Biology,
a section of the journal
Frontiers in Applied Mathematics and Statistics

RECEIVED 22 December 2022
ACCEPTED 13 January 2023
PUBLISHED 24 January 2023

CITATION
Demongeot J and Magal P (2023) Editorial: Data driven modeling in mathematical biology. *Front. Appl. Math. Stat.* 9:1129749. doi: 10.3389/fams.2023.1129749

Editorial: Data driven modeling in mathematical biology

Jacques Demongeot[1]* and Pierre Magal[2]

[1]Laboratory AGEIS EA 7407, Team Tools for e-Gnosis Medical, Faculty of Medicine, University Grenoble Alpes (UGA), La Tronche, France, [2]Individual Based Modeling, UMR CNRS 5251, University Bordeaux, Talence, France

KEYWORDS

data driven modeling, mathematical biology, epidemiology, mathematical medicine, theoretical biology

Editorial on the Research Topic
Data driven modeling in mathematical biology

The present data overflow (sometimes a data tsunami), which we can see in hospitals or biological laboratories, results from the multiplication of the investigation tools. It is pushing both the conceiver of medical instruments, the experimentalist, the medical doctor, and/or the expert in biomedical modeling to follow a dual approach: (i) to build a model of the observed reality accounting for the acquired data, and (ii) to select crucial variables to be acquired in the framework of a model (conceptual or formal) of this reality. This dual attitude is representative of the present effort done by the scientific and medical community to advance rationally toward the new frontiers of modeling and instrumentation for increasing biomedical knowledge.

It is clear that the crucial variables likely to explain the fundamental mechanisms of a new pathology observed through epidemiologic surveillance, clinical practice, and biomedical devices, allowing a specific measure (excluding redundant and non-relevant observables) and centering medical and R&D activity on a fast and precise *in situ* acquisition.

During the COVID-19 pandemic, two tools were at the extremities of the spectrum for the data acquisition methods. First, the databases like Worldometer® stored daily and worldwide new cases and deaths caused by the disease. Secondly, the quantitative PCR of the viral genome proved that it is possible to obtain specific statistics and signals devoted to manifesting a given known or new medical pathology.

In the Research Topic entitled "Data-Driven Modeling in Mathematical Biology," the sixth papers are representative of the present data/model dual approach:

(1) The first concerns "*A model of vaccination for dengue in the Philippines 2016–2018*" by Magal et al.. It describes a mathematical model of the dengue epidemic in the Philippines and analyses the impact of children's vaccination in 2016–2017. Reported case and mortality data are used to analyze this vaccination program quantitatively by comparing the epidemic outcomes of different vaccination policies.
(2) The second paper deals with "*Modeling of COVID-19 pandemic vis-à-vis some socioeconomic factors*" and has been written by Oshinubi et al.. It studies the impact of the COVID-19 epidemic on the socioeconomic status of countries worldwide. They establish a mathematical relationship between Theil and Gini indices, then investigate the relationship between epidemiological data and socioeconomic determinants, using several machine learning and deep learning methods. Their results show a clear difference between the first and the second wave of the pandemic, confirming the impact of the real dynamics of the epidemic's spread in several countries and the means by which it was mitigated.

(3) The third article consists of a "*Model-based analysis of myocardial strains in left bundle branch block*" by Taconné et al.. They propose a system-level model of the cardiovascular system is proposed, integrating:

(1) The cardiac electrical system.
(2) Right and left atria.
(3) A multi-segment representation of the RVs and LVs.
(4) The systemic and pulmonary circulations.

After a sensitivity analysis step, they identify specifically for each patient model parameters, showing that strain morphologies are related to both electrical conduction delay and heterogeneity of contractile levels within the myocardium.

(4) The fourth contribution describes "*A flexible smoother adapted to censored data with outliers and its application to SARS-CoV-2 monitoring in wastewater*" and is written by Courbariaux et al.. It concerns a sentinel network, Obépine, designed to monitor SARS-CoV-2 viral load in wastewaters arriving at wastewater treatment plants (WWTPs) in France as an indirect macro-epidemiological parameter. The authors propose a method based on an auto-regressive model adapted to censored data with outliers, and the resulting smoothed signal shows a good correlation with other epidemiological indicators and is currently used by Obépine to provide an estimate of virus circulation over the watershed.

(5) The fifth paper presents "*A step forward to formalize tailored to problem specificity mathematical transforms*" by Glaría et al.. Linear functional analysis, historically founded by Fourier and Legendre, played a significant role in providing a unified vision of mathematical transformations between vector spaces. They extend the classical linear functional analysis by Fourier and Legendre when the basis of the vector space used in the model is built Tailored to the Problem Specificity (TPS) and not from the convenience or effectiveness of mathematical calculations. The proper use of computational intelligence tools allows complexity algorithm optimization, encouraging the search for a general TPS methodology.

(6) The last article, "*Estimation of some epidemiological parameters with the COVID-19 data of Mayotte*" by Manou-Abi et al.. They study some statistical methods to fit the probability distribution, which underlines the serial interval distribution of COVID-19 on a given set of data collected on the viral shedding in patients with laboratory-confirmed. The best-fit model of the non-negative serial interval distribution is given by a mixture of two Gamma distributions with different shapes and rates. Using some regression and forecasting methods, they provide learning models of hospitalized, intensive care, and death cases over a given period.

All the contributions to the Research Topic on "Data-Driven Modeling in Mathematical Biology" show the same concern about building models well adapted to the data, themselves acquired according to a process filtered by the constraints of the modeling. The ideal is obviously to obtain the optimal model, that is, the most retro-predictive of the acquired data and the most capable of forecasting the future evolution of the observed biomedical phenomenon while using the most appropriate and economical means of investigation. That is to focus on the minimal set of the most explanatory variables of the mechanisms at the origin of the studied disease or the observed biological process.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note