



Editorial: Data Mining and Methods for Early Detection, Horizon Scanning, Modelling, and Risk Assessment of Invasive Species

Aristides Moustakas^{1*} and Stelios Katsanevakis^{2*}

¹ Institute for Applied Data Analytics, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei, ² Department of Marine Sciences, University of the Aegean, Mytilene, Greece

Keywords: data mining, alien species, risk assessment, data analytics, ecology, biological invasions, computational modeling, methods development

Editorial on the Research Topic

Data Mining and Methods for Early Detection, Horizon Scanning, Modelling, and Risk Assessment of Invasive Species

Alien species are taxa introduced to areas beyond their natural distribution by human activities, overcoming bio-geographical barriers [1]. Accelerating movement of humans, animals and goods are driving an increasing rate of biological invasions [2, 3]. Through a multitude of mechanisms of introduction, broadly classified into five categories—release, escape, contaminant, stowaway, and corridor—many thousands of species have moved into new regions [2]. In Europe alone, ~15,000 alien species have been recorded, with an increasing trend in new introductions [4]. For most taxonomic groups, the global rate of new introductions is increasing with no sign of saturation in the accumulation of alien species [5].

Many alien species have become invasive with substantial impacts on biological diversity, human health and ecosystem services. Impacts of invasive species range from reduction in individual fitness of native species to population declines, local extinctions, changes in community composition, effects on entire ecosystem processes, and wider ecosystem function, health problems in humans or even deaths, and substantial economic losses [6–11]. The ecological impacts of invasive species are so severe that they are considered as one of the major drivers of global biodiversity loss [12, 13].

The Convention on Biological Biodiversity (CBD) recognized the need for the “compilation and dissemination of information on alien species that threaten ecosystems, habitats, or species to be used in the context of any prevention, introduction and mitigation activities,” and calls for “further research on the impact of alien invasive species on biological diversity” [14]. Indeed, research efforts to better understand the factors controlling alien species introduction, initial dispersal, establishment success, distribution, abundance, spatio-temporal dynamics, and invasiveness have been intensified. In the 1990s and 2000s, research on human-mediated invasions has exploded, with an exponential increase of publications dealing with or mentioning biological invasions (**Figure 1**). In addition, many peer-reviewed journals specifically targeting biological invasions have been established (e.g., “Biological Invasions,” “Aquatic Invasions,” “Management of Biological Invasions,” “BioInvasions Records”), a large number of dedicated information systems and online databases have been created [15], and a number of international initiatives advancing research and collaboration in the field have been promoted [16].

OPEN ACCESS

Edited and reviewed by:

Xuan Zhu,
Monash University, Australia

*Correspondence:

Aristides Moustakas
arismoustakas@gmail.com
Stelios Katsanevakis
stelios@katsanevakis.com

Specialty section:

This article was submitted to
Environmental Informatics,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

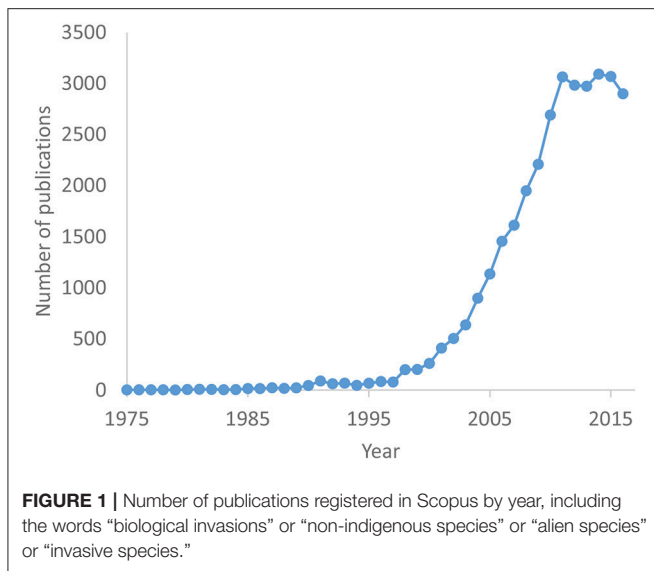
Received: 05 February 2018

Accepted: 05 March 2018

Published: 21 March 2018

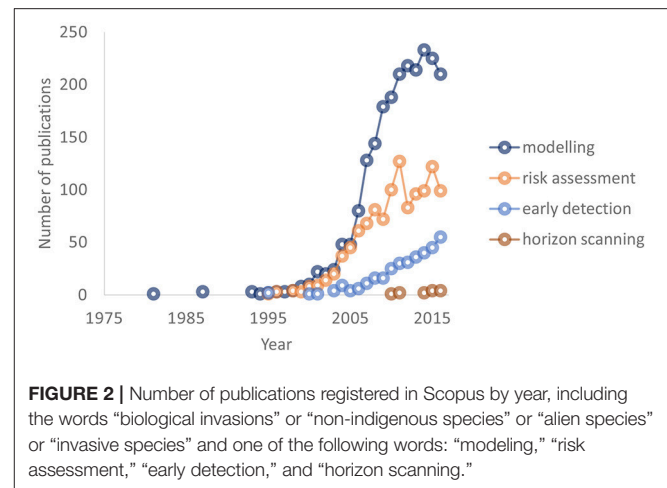
Citation:

Moustakas A and Katsanevakis S
(2018) Editorial: Data Mining and
Methods for Early Detection, Horizon
Scanning, Modelling, and Risk
Assessment of Invasive Species.
Front. Appl. Math. Stat. 4:5.
doi: 10.3389/fams.2018.00005



Despite the intensification of research efforts, there are still substantial gaps in our knowledge on the dynamics of biological invasions and the accumulation of alien species across regions and taxa. Understanding the spread and successful introduction of invasive species is highly complex as biological, social, geographic, economic, and climatic factors influence the way an invasive species is introduced and determine the options available for its eventual detection and control. This in turn calls for an interdisciplinary approach as traditional ecological approaches need to be combined with data mining and data analytics [17]. Data availability has often been the limiting factor in predictive modeling [18, 19]. With the rapid development of smart sensors, social networks, digital maps, and remotely-sensed imagery, spatio-temporal data are more ubiquitous and richer than ever before [20]. The availability of such large datasets (big data) poses great challenges in data analysis [21]. In addition, increased availability of computing power facilitates the use of computationally-intensive methods for the analysis of such data [22]. Thus new methods are needed to understand biological invasions and to efficiently use new technologies and handle big datasets. The current Research Topic in Environmental Informatics aims to address this interesting and cutting-edge theme. Specifically, the special issue focuses on data analysis and novel methods for early detection, horizon scanning, modeling, and risk assessment of invasive species. These themes have been rapidly developing or are expected to develop in the near future (Figure 2). Ten contributions were finally accepted after the peer review process.

Data in alien species inventories are often scattered in numerous disconnected formats and databases that lack interoperability entangling their analysis. In their contribution, Groom et al. describe seven ways that data on alien species can be made more accessible and useful. These include data management plans, interoperability of information sources, metadata documentation, existing data formulation, use of common data semantics, data availability, and long-term data preservation.



Spatially explicit high resolution data are elementary for assessing the impacts of alien species and mitigation; however, often these data are unavailable. Environmental explanatory covariates available at the resolution of the alien species, as well as at finer resolutions, can be used to infer alien species presences at finer resolutions using random forests, a machine learning technique, as suggested by Daliakopoulos et al.

Building and managing large datasets of alien species is elementary for their monitoring and management. The geodatabase of the European Alien Species Information Network (EASIN), an initiative of the European Commission, is presented by Deriu et al. The key feature of EASIN is that it aggregates, integrates, and harmonizes spatio-temporal data of alien species in Europe, through standardized processes, supporting both research and policies on biological invasions.

Land use change is the major contributor to the introduction and spread of alien species, and forest lodging and road constructions increase the exposure to infectious diseases. In addition, climatic changes can facilitate alien species and disease spread. The spatio-temporal spread of Hemorrhagic fever with renal syndrome (HFRS) is studied using powerful statistical methods and core areas in China, and the association of climatic cycles are quantified by He et al. demonstrating the applicability of modern space-time modeling and mapping techniques to assess biological invasions.

Islands are undergoing strong pressure from alien species due to their restricted size and increased human impact. Hypotheses of invasibility were explicitly tested by Bjarnason et al. using a detailed statistical framework, and it was concluded that areas of high species richness have greater invasibility. A negative relationship between alien and endemic species richness was also recorded potentially providing inverse learning patterns between alien and endemic species.

The fact that closely related species share a higher chance of becoming infected or attacked by pests allows the identification of taxa with different degrees of vulnerability. Using publicly available information about pest-host interactions and their geographic distributions, in combination with host phylogenetic reconstructions, Robles-Fernández and Lira-Noriega estimated a

pest-host interaction or infection index in geographic space as spatially explicit tool for risk assessment.

Species distribution models are a commonly used tool for the analysis of alien species. However, they strongly rely on species' presence/absence and confirmed absences are hard to quantify. Environmental DNA is a novel technique that presents higher detectability and accuracy in comparison to conventional sampling techniques, and can effectively differentiate between species' presence/absence, improving species distribution models, as argued by Muha et al.

Calibrated Individual based models (IBMs) are useful tools for investigating the dynamics of alien species. Samson et al. highlight how early-stage data-limited IBMs can play an important role in defining key research priorities of invasive species in order that subsequent models can provide robust insight into potential management interventions.

Ensemble methods are meta-algorithms that combine several techniques into a unique predictive model to minimize variance; they are increasingly employed in science including a recent application to draw up the habitats' maps for mosquito alien species [23]. However, the development of these models should not be done in a "black box" mode and it should be accompanied by a set of in-depth analysis regarding key training and operation decision points, thereby promoting a reproducible evaluation of the proposed methodology, as suggested by Demertzis et al.

Uncertainty is an attribute of (lack of) information and it is a vital component in other scientific disciplines; however, it is usually missing from marine invasion science studies. The final contribution in this Research Topic, by Katsanevakis and Moustakas, argues that without assessing uncertainty, it is hard

to envisage future improvements. In addition, if scientists are perceived by the public either to overstate their findings in order to receive high visibility or to downplay the uncertainty of their findings, society is likely to lose confidence in the outputs of invasion science.

We believe that this Research Topic provides useful new insights for advancing the field of invasion science, by combining knowledge and techniques from multiple disciplines. Science initially went from the phase of "homo-universalis" to high specialization potentially missing the inter-connection with other scientific disciplines. While this may have been a necessity in the history of science, in the big data and globalization era the necessity and feasibility of a more integrated data-based but combining practical and theoretical knowledge approach appears to be more feasible than ever before.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

We thank all contributors for submitting their work in this Research Topic, the reviewers for their time and effort, the editorial office of Frontiers for coordination and support, and the Associate Editors Xuan Zhu, Yang Liu, and Juergen Piltz for handling the editorial as well as other contributions that AM and SK were coauthoring.

REFERENCES

- Blackburn TM, Pyšek P, Bacher S, Carlton JT, Duncan RP, Jarošík V, et al. A proposed unified framework for biological invasions. *Trends Ecol Evol.* (2011) **26**:333–9. doi: 10.1016/j.tree.2011.03.023
- Essl F, Bacher S, Blackburn TM, Booy O, Brundu G, Brunel S, et al. Crossing frontiers in tackling pathways of biological invasions. *Bioscience* (2015) **65**:769–82. doi: 10.1093/biosci/biv082
- Canning-Clode J. *Biological Invasions in Changing Ecosystems: Vectors, Ecological Impacts, Management and Predictions*. Warsaw; Berlin: Walter de Gruyter GmbH & Co KG (2015).
- Katsanevakis S, Deriu I, D'Amico F, Nunes AL, Sanchez SP, Crocetta F, et al. European alien species information network (EASIN): supporting European policies and scientific research. *Manage Biol Invas.* (2015) **6**:147–57. doi: 10.3391/mbi.2015.6.2.05
- Seebens H, Blackburn TM, Dyer EE, Genovesi P, Hulme PE, Jeschke JM, et al. No saturation in the accumulation of alien species worldwide. *Nat Commun.* (2017) **8**:14435. doi: 10.1038/ncomms14435
- Vilà M, Basnou C, Pyšek P, Josefsson M, Genovesi P, Gollasch S, et al. How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. *Front Ecol Environ.* (2010) **8**:135–44. doi: 10.2307/20696461
- Blackburn TM, Essl F, Evans T, Hulme PE, Jeschke JM, Kühn I, et al. A unified classification of alien species based on the magnitude of their environmental impacts. *PLOS Biol.* (2014) **12**:e1001850. doi: 10.1371/journal.pbio.1001850
- Katsanevakis S, Wallentinus I, Zenetos A, Leppäkoski E, Çınar ME, Öztürk B, et al. Impacts of marine invasive alien species on ecosystem services and biodiversity: a pan-European review. *Aquat Invas.* (2014) **9**:391–423. doi: 10.3391/ai.2014.9.4.01
- Mazza G, Tricarico E, Genovesi P, Gherardi F. Biological invaders are threats to human health: an overview. (2014) *Ethol Ecol Evol.* **26**:112–29. doi: 10.1080/03949370.2013.863225
- Bellard C, Cassey P, Blackburn TM. Alien species as a driver of recent extinctions. *Biol Lett.* (2016) **12**:20150623. doi: 10.1098/rsbl.2015.0623
- Ojaveer H, Galil BS, Campbell ML, Carlton JT, Canning-Clode J, Cook EJ, et al. Classification of non-indigenous species based on their impacts: considerations for application in marine management. *PLoS Biol.* (2015) **13**:e1002130. doi: 10.1371/journal.pbio.1002130
- CDB. *Convention on Biological Diversity, Programme of Work on Invasive Alien Species*. Available online at: <https://www.cbd.int/invasive> (2016).
- MEA (Millennium Ecosystem Assessment). *Ecosystems and Human Wellbeing: Biodiversity Synthesis*. Washington, DC: World Resources Institute (2005). p. 86.
- CDB. Convention on biological diversity, interim guiding principles. In: *Conference of the Parties Decision V/8 Alien Species That Threaten Ecosystems, Habitats or Species*. Available online at: <http://www.cbd.int/decision/cop/default.shtml?id=7150> (2000).
- Katsanevakis S, Roy HE. Alien species related information systems and information management. *Manage Biol Invas.* (2015) **6**:115–17. doi: 10.3391/mbi.2015.6.2.01
- Lucy FE, Roy H, Simpson A, Carlton JT, Hanson JM, Magellan K, et al. INVASIVESNET towards an international association for open knowledge on invasive alien species. *Manage Biol Invas.* (2016) **7**:131–9. doi: 10.3391/mbi.2016.7.2.01
- Moustakas A, Voutsela A, Katsanevakis S. Sampling alien species inside and outside protected areas: does it matter? *SciTot Environ.* (2018) **625**:194–8. doi: 10.1016/j.scitotenv.2017.12.198

18. Evans MR, Moustakas A. A comparison between data requirements and availability for calibrating predictive ecological models for lowland UK woodlands: learning new tricks from old trees. *Ecol Evol.* (2016) **6**:4812–22. doi: 10.1002/ece3.2217
19. Evans MR, Benton TG, Grimm V, Lessells CM, O'Malley MA, Moustakas A, et al. Data availability and model complexity, generality, and utility: a reply to Loneragan. *Trends Ecol Evol.* (2014) **29**:302–3. doi: 10.1016/j.tree.2014.03.004
20. Moustakas A. Spatio-temporal data mining in ecological and veterinary epidemiology. *Stoch Environ Res Risk Assess.* (2017) **31**:829–34. doi: 10.1007/s00477-016-1374-8
21. Moustakas A, Evans MR. A big-data spatial, temporal and network analysis of bovine tuberculosis between wildlife (badgers) and cattle. *Stoch Environ Res Risk Assess.* (2017) **31**:315–28. doi: 10.1007/s00477-016-1311-x
22. Moustakas A, Evans M. Coupling models of cattle and farms with models of badgers for predicting the dynamics of bovine tuberculosis (TB). *Stoch Environ Res Risk Assess.* (2015) **29**:623–35. doi: 10.1007/s00477-014-1016-y
23. Cunze S, Koch LK, Kochmann J, Klimpel S. *Aedes albopictus* and *Aedes japonicus* - two invasive mosquito species with different temperature niches in Europe. *Parasit Vect.* (2016) **9**:573. doi: 10.1186/s13071-016-1853-2

Conflict of Interest Statement: 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.

Copyright © 2018 Moustakas and Katsanevakis. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.