

How data mining helps solve aquatic invasions for social good?

Improving Management of Aquatic Invasions by Integrating Shipping Network, Ecological, and Environmental Data



Jian Xu*, Thanuka L. Wickramaratne*, Nitesh V. Chawla*, Erin K. Grey, Karsten Steinhäuser, Reuben P. Keller, John M. Drake and David M. Lodge

Risk of aquatic invasions

The unintentional transport of **invasive species** through the **global shipping network** causes substantial losses to social and economic welfare: an invisible tax that **everyone** has to pay.

Annual damage and control costs of invasive species:



Demands of management

Imposing expensive and cumbersome regulations on the shipping industry could harm the country's economy and trade relationships. A thorough understanding of ship-borne invasion risks is necessary to devise policies that are feasible, effective, and capable of bringing to bear the net long-term benefits to human welfare.

Global shipping is responsible for:



Challenges of risk assessment

Ship-borne invasions are a result of a **complex interplay** of ship traffic, ballast uptake/discharge dynamics, species survival during transport, various environmental/biological variables, etc.

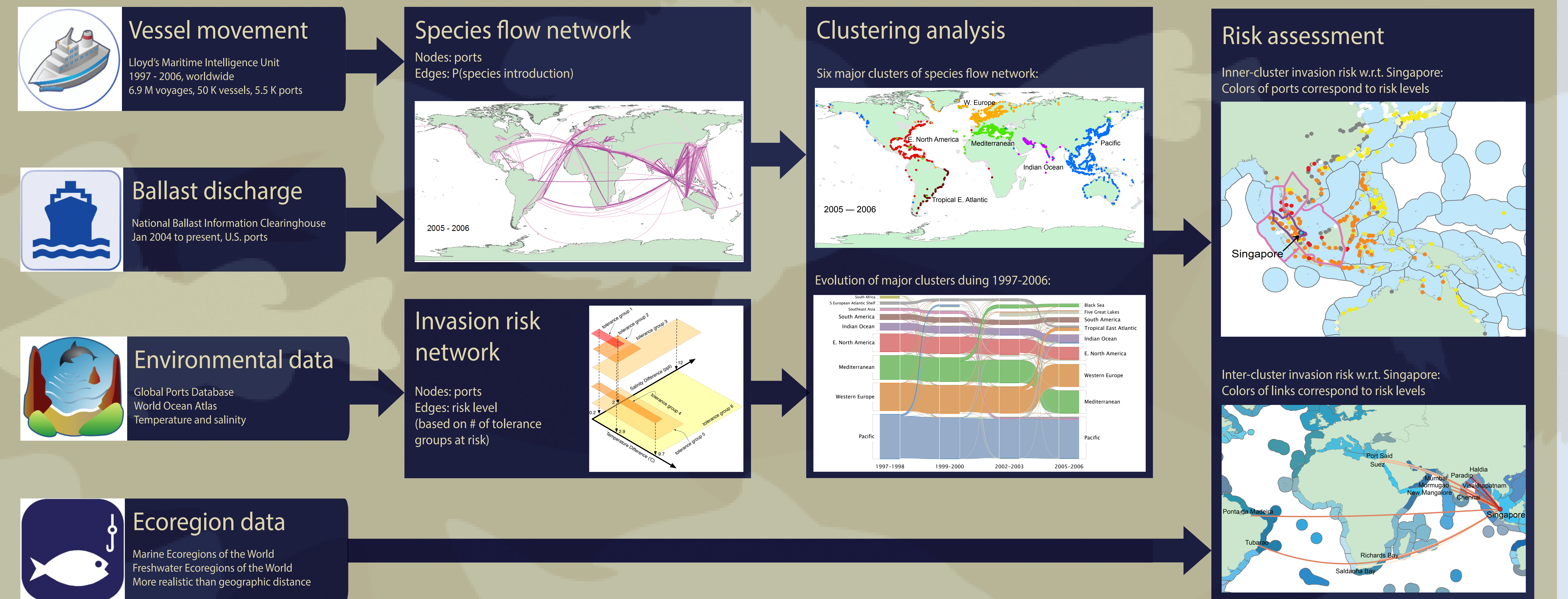
The few studies that have attempted to quantify invasion risks via probabilistic approaches have relied on multiple simplifying assumptions. Numerous difficulties associated with quantitative risk assessments (e.g. inadequate characterizations of invasion processes, lack of crucial data, large uncertainties associated with available data, etc.) have hampered the usefulness of such estimates in the task of supporting the authorities who are battling to manage invasions with limited resources.

Contributions of our work

We present here an approach for addressing the problem at hand via creative use of computational techniques and multiple data sources, thus illustrating how data mining can be used for solving crucial, yet very complex problems towards **social good**.

By experimenting on actual global shipping traffic data for years 1997-2006, we have discovered crucial knowledge on how patterns in global shipping network affect aquatic invasions, and we illustrated how such knowledge can be used to devise effective and economical invasive species management strategies.

Framework of data mining for risk assessments



Highlights of discoveries

Species can be carried from any arbitrary port to another in the world within **two intermediate stops** on average. This indicates a generally high risk of invasions in the absence of risk management.

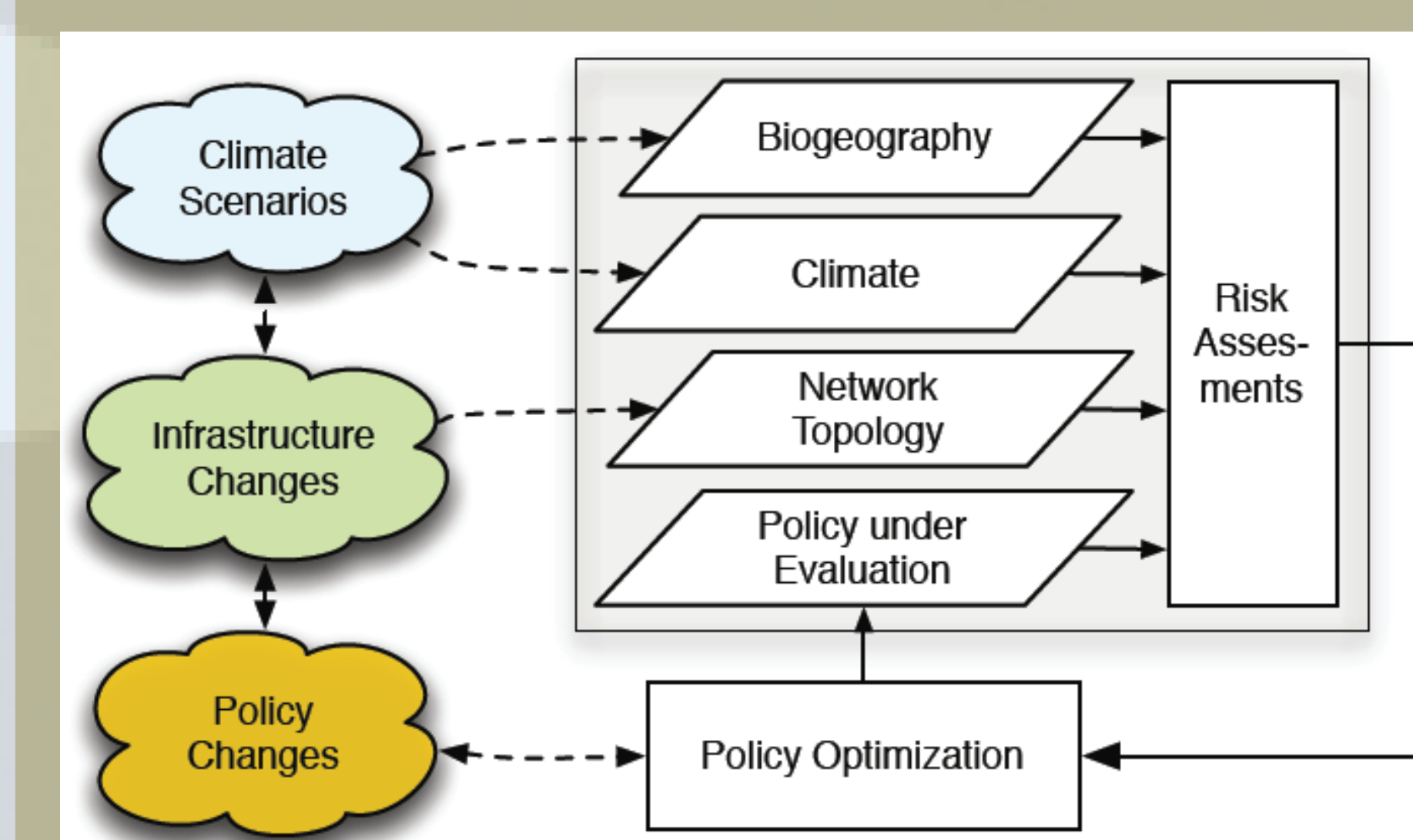
Six major clusters continue to exist over the duration studied. The vast majority of ports continue to remain in the same cluster. This provides a solid foundation for devising management strategies targeting clusters and inter-cluster connections.

A few ports contribute to lots of inter-cluster species flow. For example, **Singapore** alone contributes to 26% of total inter-cluster flow from/to Pacific cluster which contains 818 ports. By taking precautions on top 20% of most-connected ports, translocations of species will be at least twice as difficult.

Vessels such as **container carriers** travel frequently between clusters, while **barge ships** etc. usually stay within clusters.

Deployment of framework

Use of discovered knowledge in a potential deployed setting for invasion risk assessment with respect to changing climate, policy and infrastructure:



Future work

We are preparing to launch a detailed project looking at numerous aspects of invasive species. This project was recently funded by US National Science Foundation (NSF) for 2015-2019 for USD 2M.

Under this project, we are planning to develop elaborate methods to generate species distributions, mathematical modeling of invasions, evaluation of the impact of global climate change on invasions, etc.

Furthermore, we are also going to implement a Management Transition Board, where our research findings will be effectively communicated with the authorities and their feedback will be utilized to drive our research to be both effective and useful to the society.

Contact us

This work is based on research supported by the Notre Dame Office of Research via the Environmental Change Initiative (ECI), the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053, the U.S. Air Force Office of Scientific Research (AFOSR) and the Defense Advanced Research Projects Agency (DARPA) grant #FA9550-12-1-0405, NOAA CSCOR and EPA GLRI grants. We are grateful to the anonymous reviewers for their constructive comments.

* Jian Xu, Thanuka L. Wickramaratne, Nitesh V. Chawla are with Interdisciplinary Center for Network Science and Applications (iCeNSA @ Notre Dame)
For latest research progress, please visit:

