Automating Historic Nutrient Delivery Model Calibration to Assess Nutrient Retention Services in Puerto Rico

Mariam Valladares-Castellanos

Abstract

Nutrient retention is an essential ecosystem service that plays a critical role in maintaining water quality by reducing nutrient runoff, particularly in ecologically sensitive regions such as Puerto Rico. This study addresses the need for accurate predictions of nutrient retention to inform effective water quality management strategies. Existing literature indicates significant shortcomings in the calibration and validation processes of ecosystem service models, notably the widely used InVEST (Integrated Valuation of Ecosystem Services and Tradeoffs), with fewer than 15% of studies incorporating robust calibration or validation efforts. This research proposes an automated calibration process for the InVEST Nutrient Delivery Ratio (NDR) model to analyze historical nutrient retention across Puerto Rican watersheds from 1977 to 2000. Utilizing Land Use Land Cover maps and relevant biophysical and topographical inputs, the study implements an automated batch calibration approach to identify optimal calibration values, focusing on the Borselli K and threshold flow accumulation parameters. By comparing model outputs against reference nutrient data, the study aims to demonstrate an example for enhancing model accuracy and processing efficiency, thereby broadening the applicability of the findings across diverse watersheds. The anticipated results will provide insights into changes in nutrient retention capacities over time, with the hypothesis that these capacities have increased due to evolving land use and watershed management practices. This research not only advances methodological innovations in ecosystem service modeling but also offers critical insights for policymakers and conservationists, enabling the development of targeted and sustainable management strategies to safeguard vital water resources in Puerto Rico.

Introduction

Nutrient retention is an ecosystem service that helps maintain water quality by reducing nutrient runoff into aquatic systems. Accurate predictions of nutrient retention services are essential for effective water quality management, especially in ecologically sensitive areas like Puerto Rico. Consequently, the precision of models predicting nutrient retention is critical for making informed decisions in environmental management.

Previous research has highlighted a significant gap in the calibration and validation processes of ecosystem service models, particularly in the most widely used one InVEST (Integrated Valuation of Ecosystem Services and Tradeoffs) (Ochoa & Urbina-Cardona, 2017). InVEST models are sensitive to the parameters governing streamflow delineation which can change considerably based on the parameters selected (Redhead et al., 2018). Figure 1 depicts an exploratory example of the variability in phosphorus estimates when using multiple parameter combinations. However, despite the variability in nutrient estimates, less than 15% of scientific publications report conducting model calibration or validation of the InVEST estimates indicating the use of default calibration parameters (Agudelo et al., 2020). For the fewer studies that do perform calibration and validation, they tend to focus on discrete watersheds with a limited time interval, potentially overlooking the complex interactions that affect ecosystem services which limits their broader applicability. While static parameter calibration can be useful, it does not meet the dynamic needs required for effective decision-making.

To gain a comprehensive understanding of historical changes in nutrient retention, a more expansive and dynamic approach to model calibration is necessary. As the accuracy of parameter values can vary across different geographic and hydrologic conditions, there is an urgent need for methods that allow for thorough evaluations of these variations to ensure model predictions are reliable. However, evaluating the accuracy of multiple parameters can be challenging and labor-intensive when done manually. This highlights the need for an automated calibration process, especially when using widely accepted ecosystem service estimation tools like InVEST across regions and periods.

This study aims to set an example of such calibration tools to examine the historical differences in nutrient retention functions across Puerto Rican watersheds from 1977 to 2000. By employing an automated batch calibration process and suitable parameter selection function, we seek to enhance model accuracy and processing efficiency, thereby informing effective watershed management and conservation strategies. The central question guiding this research is: How have nutrient retention capacities changed in Puerto Rico’s watersheds between 1977 and 2000? We hypothesize that nutrient retention capacity in these watersheds has increased during this period, driven by changes in vegetation related to land use and land cover, and evolving watershed management practices that influence the transport and accumulation of nutrients.

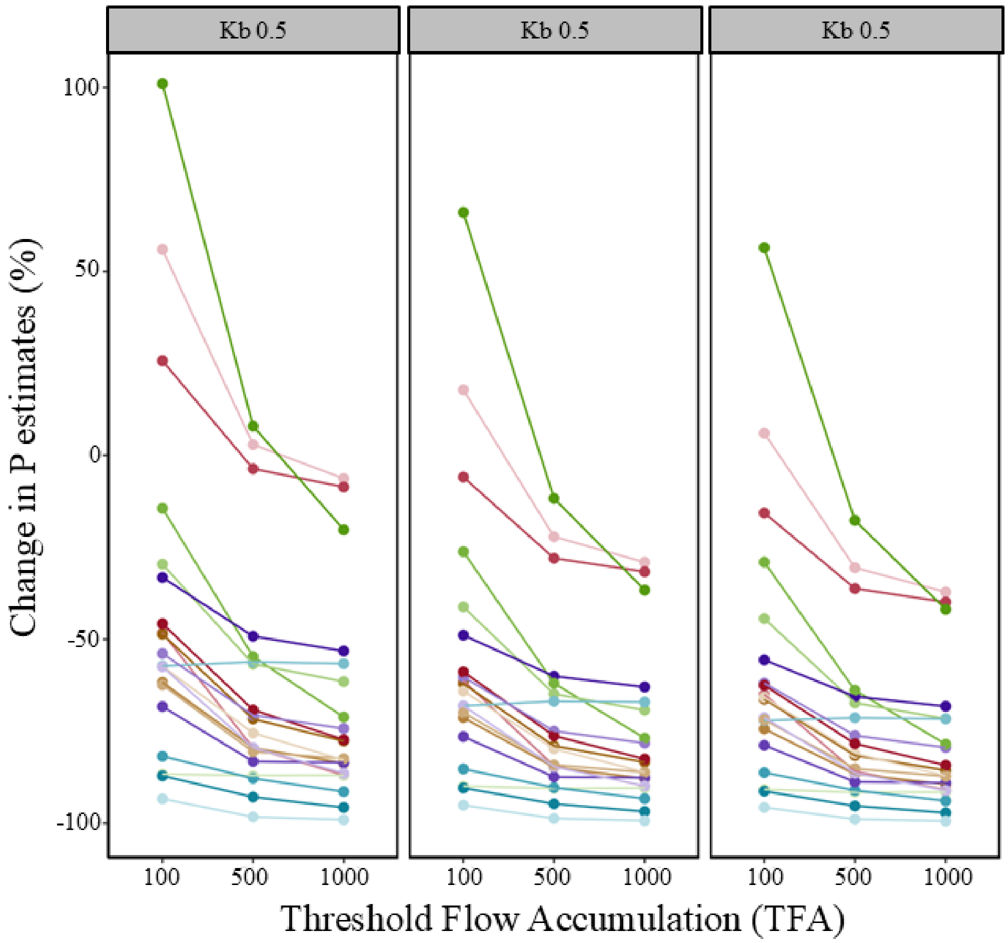

Figure 1. Examples of variations in InVEST nutrient export estimates for phosphorus across twenty watersheds in Puerto Rico which illustrate the impact of altering two model parameters: the threshold flow accumulation and the Borselli K.

Methods

The study aims to demonstrate an automated calibration process for nutrient retention estimates derived from the InVEST Nutrient Delivery Ratio (NDR) model Python package (natcap.invest). The workflow of the model construction, calibration, and suitable parameter selection is represented in Figure 2.

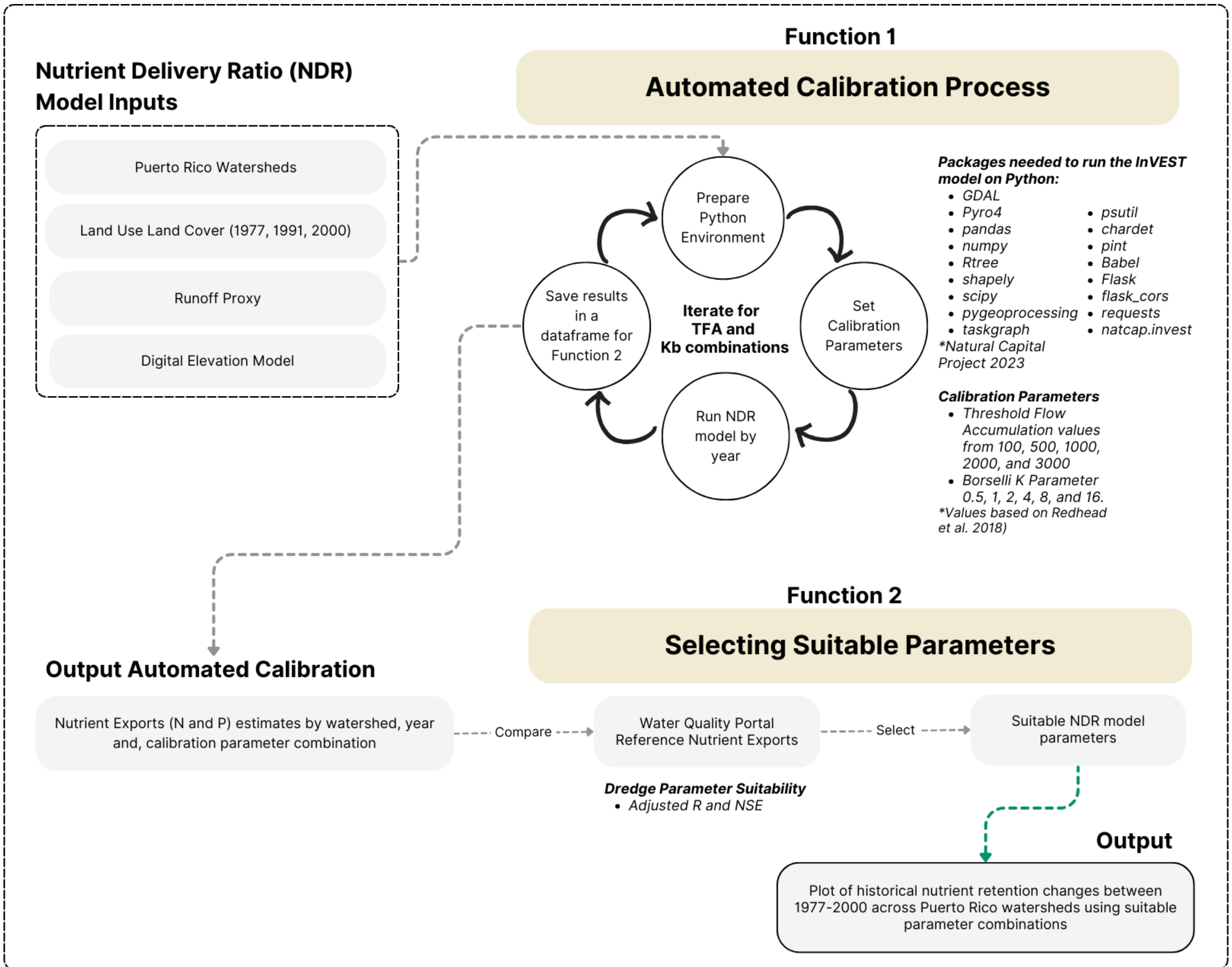

Figure 2. Workflow illustrating the process proposed in this study for establishing an automated calibration function to execute the natcap.invest NDR model, along with a secondary function for selecting optimal parameter combinations when comparing estimates to reference observations.

Estimating Nutrient Retention Services

The Nutrient Delivery Ratio (NDR) model utilizes a mass balance to assess total nitrogen and total phosphorus exports from uplands drainage areas into the streams. The model computes nutrient loads from non-point sources. By incorporating landscape, elevation, and runoff features, the NDR model calculates the fraction of nutrients that drain into streams. This fraction is termed the Nutrient Delivery Ratio (Perrine et al., 2015). The output from the model are the nutrient exports per watershed per year.

Model Parameterization

In this study, the inputs for the natcap.invest NDR model will include 30-meter resolution Land Use Land Cover (LULC) maps from the years 1977, 1991, and 2000, sourced from USDA-IITF files (Helmer et al., 2002). These maps will be reclassified into seven primary categories based on the Anderson Classification system using ArcGIS Pro 3.0, as outlined by Beck et al. (2013). The second input will consist of biophysical parameters adapted from Smith et al. (2017), which are essential for establishing baseline nutrient values for each LULC class. The third input, a 30-meter Digital Elevation Model (DEM), will be acquired from the Advanced Spaceborne Thermal Emission and Reflection Radiometer ASTER (2019) product. Additionally, the runoff proxy layer will utilize the InVEST Urban Flood Mitigation 43 mm storm layer from De Jesus Crespo et al. (2019), representing an average rainfall event for the island. Finally, watersheds will be delineated using the USGS Stream Stats v4.16.1 tool (Ries Iii et al., 2008), with the resulting watershed shapefiles employed to summarize the model estimates.

Function 1 Automated Calibration Process

An automated batch calibration function will be built in Python to execute the natcap.invest NDR model for each year under consideration (1977, 1991, 2000), selecting for each iteration a new set of calibration parameters across watersheds in Puerto Rico. Calibration in this context will focus on evaluating the model's relative performance in response to changes in parameters. The model estimates resulted from each watershed-year-parameter iteration will be stored in a data frame for statistical comparison against observed nutrient loads.

For this study, two calibration parameters will be tested, the Borselli K parameter (Kb) and the threshold flow accumulation (TFA). Kb represents the relationship between the watershed stream connectivity and the fraction of nutrients that could potentially reach the system, while TFA indicates the minimum number of pixels required to define a stream. Based on preliminary analyses three TFA values will be tested: 100, 1000, and 10,000; and 0.5, 1, 2, 4, 8, and 16 for Kb (Anjinho et al., 2022; Redhead et al., 2018).

Function 2 Selecting Suitable Model Parameters

Given the extensive range of potential watershed-year-parameter combinations and to facilitate the management of statistical models, another python function will be built, a dredge function. The function will be applied to compare the model estimates against reference nutrient observations to identify the optimal predictive parameter combination. Suitable parameter combinations for each site/year will be selected based on the highest adjusted R^2 values (Barton, 2015; Grueber et al., 2011).

The reference nutrient exports for total nitrogen, total phosphorus, and streamflow, spanning from January 1, 1950, to December 31, 2020, will be acquired from the Water Quality Portal (WQP) (National Water Quality Monitoring Council, 2020). Any gaps in the reference data will be addressed using a random forest machine learning method, wherein 70% of observed data will train the random forest algorithm, while the remaining 30% will be used for validation. Reference nutrient exports will be calculated by multiplying daily nutrient concentrations by streamflow. The loads will be aggregated in ten-year blocks to estimate the mean nutrient load per year of interest (1977, 1991, and 2000).

Finally, after selecting the suitable NDR parameter combinations, the difference in total nutrient retention estimates across watersheds will be summarized to report the changes in NDR between years.

Relevance and Applicability of the Study

This study addresses a critical gap in the efficiency and automation of ecosystem service model calibration. Streamlining the calibration process allows for the simultaneous analysis of historical changes across multiple watersheds, providing a comprehensive understanding of nutrient dynamics on a broader scale. This methodological innovation not only enhances the precision of model outputs but also significantly contributes to the scientific discourse on nutrient retention and water quality management. The insights derived from this research will be instrumental for policymakers and conservationists, enabling them to develop targeted management strategies that are both effective and sustainable, ultimately leading to improved ecological outcomes and the preservation of vital water resources in Puerto Rico.

Timeline for completion

Note: Weeks described in the timeline begin October 28th and end December 6th, 2024.

References

Advanced Spaceborne Thermal Emission and Reflection Radiometer ASTER. (2019). *Global digital elevation model version 3 (GDEM 003)*. https://asterweb.jpl.nasa.gov/gdem.asp.

Agudelo, C. A. R., Bustos, S. L. H., & Moreno, C. A. P. (2020). Modeling interactions among multiple ecosystem services. A critical review. *Ecological Modelling*, *429*, 109103. https://doi.org/https://doi.org/10.1016/j.ecolmodel.2020.109103

Anjinho, P. d. S., Barbosa, M. A. G. A., & Mauad, F. F. (2022). Evaluation of InVEST's Water Ecosystem Service Models in a Brazilian Subtropical Basin. *Water*.

Barton, K. (2015). Package 'mumin'. *Version*, *1*(18), 439.

Beck, H. E., Bruijnzeel, L. A., van Dijk, A. I. J. M., McVicar, T. R., Scatena, F. N., & Schellekens, J. (2013). The impact of forest regeneration on streamflow in 12 mesoscale humid tropical catchments. *Hydrol. Earth Syst. Sci.*, *17*(7), 2613-2635. https://doi.org/10.5194/hess-17-2613-2013

De Jesus Crespo, R., Wu, J., Myer, M., Yee, S., & Fulford, R. (2019). Flood protection ecosystem services in the coast of Puerto Rico: Associations between extreme weather, flood hazard mitigation and gastrointestinal illness. *Sci Total Environ*, *676*, 343-355. https://doi.org/10.1016/j.scitotenv.2019.04.287

Grueber, C. E., Nakagawa, S., Laws, R. J., & Jamieson, I. G. (2011). Multimodel inference in ecology and evolution: challenges and solutions. *Journal of evolutionary biology*, *24*(4), 699-711. https://doi.org/https://doi.org/10.1111/j.1420-9101.2010.02210.x

Helmer, E., Ramos, O., Del, T., López, M., Quiñones, M., Diaz, A., & López-Marrero, T. (2002). Mapping the Forest Type and Land Cover of Puerto Rico, a Component of the Caribbean Biodiversity Hotspot. *Caribbean Journal of Science*, *38*.

National Water Quality Monitoring Council. (2020). *Water Quality Portal* (2000-2022. https://doi.org/https://doi.org/10.5066/P9QRKUVJ.

Ochoa, V., & Urbina-Cardona, N. (2017). Tools for spatially modeling ecosystem services: Publication trends, conceptual reflections and future challenges. *Ecosystem Services*, *26*, 155-169. https://doi.org/https://doi.org/10.1016/j.ecoser.2017.06.011

Perrine, H., Chaplin-Kramer, R., Sim, S., & Mueller, C. (2015). A new approach to modeling the sediment retention service (InVEST 3.0): Case study of the Cape Fear catchment, North Carolina, USA. *Science of the Total Environment*, *524-525*, 166-177. https://doi.org/https://doi.org/10.1016/j.scitotenv.2015.04.027

Redhead, J. W., May, L., Oliver, T. H., Hamel, P., Sharp, R., & Bullock, J. M. (2018). National scale evaluation of the InVEST nutrient retention model in the United Kingdom. *Science of the Total Environment*, *610-611*, 666-677. https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.08.092

Ries Iii, K. G., Guthrie, J. D., Rea, A. H., Steeves, P. A., & Stewart, D. W. (2008). *StreamStats: A water resources web application* [Report](2008-3067). (Fact Sheet, Issue. U. S. G. Survey. https://pubs.usgs.gov/publication/fs20083067

Smith, A., Yee, S. H., Russell, M. J., Awkerman, J. A., & Fisher, W. S. (2017). Linking ecosystem service supply to stakeholder concerns on both land and sea: An example from Guánica Bay watershed, Puerto Rico. *Ecological indicators*, *74*, 371-383. https://doi.org/https://doi.org/10.1016/j.ecolind.2016.11.036