Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: www.tandfonline.com/journals/thsj20

Controls on lag time in Philippine catchments identified using rainfall–runoff modelling and a generalized additive model (GAM)

Pamela Louise M. Tolentino, Martin D. Hurst, Richard D. Williams, Trevor B. Hoey & Richard J. Boothroyd

To cite this article: Pamela Louise M. Tolentino, Martin D. Hurst, Richard D. Williams, Trevor B. Hoey & Richard J. Boothroyd (16 Mar 2026): Controls on lag time in Philippine catchments identified using rainfall–runoff modelling and a generalized additive model (GAM), Hydrological Sciences Journal, DOI: 10.1080/02626667.2025.2608147

To link to this article: https://doi.org/10.1080/02626667.2025.2608147

Published online: 16 Mar 2026.

HYDROLOGICAL SCIENCES JOURNAL
https://doi.org/10.1080/02626667.2025.2608147

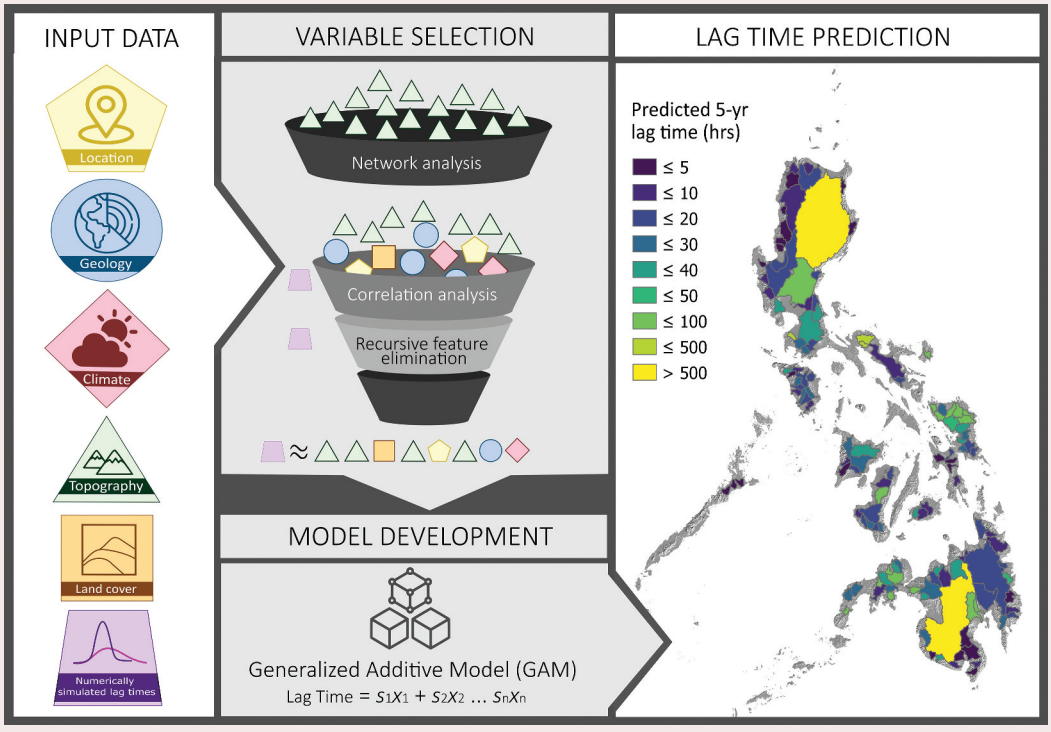
Controls on lag time in Philippine catchments identified using rainfall–runoff modelling and a generalized additive model (GAM)

Pamela Louise M. Tolentino [a], Martin D. Hurst [a], Richard D. Williams [a], Trevor B. Hoey [b] and Richard J. Boothroyd [c]

[a]School of Geographical and Earth Science, University of Glasgow, Glasgow, UK; [b]Department of Civil and Environmental Engineering, Brunel University London, London, UK; [c]Department of Geography and Planning, University of Liverpool, Liverpool, UK

ABSTRACT

Understanding the controls upon lag time, can improve river and flood management decision-making. This study investigates the relative importance of catchment characteristics in explaining lag time variability across the Philippines. Numerically simulated 5-year return period lag times for 291 catchments were analysed using a generalized additive model (GAM) to capture non-linear relationships with location, geology, climate, topography, and land use. The 5-year return period is representative of moderate flood response, as lag time varies little across return periods. Correlation analysis and recursive feature elimination guided variable selection, while bootstrapping assessed model stability and uncertainty. Ten significant controls on lag time were identified, with relief ratio, land cover index, and catchment area most influential. The GAM achieved an R^2 of 0.77 and explained 84% of deviance. Land cover emerged as the only anthropogenically modifiable control, highlighting a key management lever. National hydrological observations are needed to further support model calibration.



ARTICLE HISTORY
Received 13 May 2025
Accepted 17 November 2025

EDITOR
A. Fiori

ASSOCIATE EDITOR
A. Petroselli

KEYWORDS
lag time; general additive model; natural flood management; tropical catchments; rainfall–runoff modelling; Philippines

1 Introduction

Lag time, which is influenced by factors such as climate, catchment morphology and soil characteristics, reflects the response time of river catchments to hydrometeorological events, offering valuable insights for effective river management (Seyam and Othman 2014). For catchments of comparable size, short lag times indicate rapid flow accumulation resulting in faster time to peak, which increases the probability of flash floods, during

CONTACT Pamela Louise M. Tolentino Pammie.Tolentino@glasgow.ac.uk School of Geographical and Earth Science, University of Glasgow, Glasgow G12 8QQ, UK

Supplemental data for this article can be accessed online at http://dx.doi.org/10.5525/gla.researchdata.2197

and after intense rainfall. Conversely, longer lag times result from delayed arrival of water from upstream areas, providing more time for flood warning and lower peak discharge. Understanding lag time is particularly important in tropical rivers which typically exhibit high runoff values regardless of catchment area due to high and seasonally concentrated rainfall (Syvitski *et al.* 2014). Many tropical countries are also particularly susceptible to damaging floods; 7 out of 10 countries with the highest percentage of population exposed to high fluvial flood risk are in the tropics, including the Philippines, Indonesia, and Bangladesh, with communities living in low-lying river basins most at risk (Rentschler *et al.* 2022, World Risk Report 2024). This highlights the urgent need to better understand hydrological response in tropical catchments, as it is also predicted that climate change will cause flooding to increase further in tropical regions including South and Southeast Asia due to more intense precipitation (Eccles *et al.* 2019). Managing flood risk thus needs to be informed by hydrological analysis and modelling to inform decision making on appropriate mitigation measures. Specifically, discharge predictions in ungauged rivers can be improved by incorporating information about controls on lag time (Yuan and Forshay 2021).

Lag times are primarily controlled by the duration and intensity of rainfall but are also influenced by catchment characteristics such as topography, geology, climate, land use and land cover (LULC), and soil properties. The topographic properties of a catchment, such as catchment size, topographic relief, and drainage density, affect both water movement pathways within a catchment and travel time (Mahala 2019, Bernini *et al.* 2021). Geology also has a major influence on a catchment's hydrological behaviour (Vannier *et al.* 2016), because bedrock type influences soil type and determines sub-surface flow paths and storage capacity (Sayama *et al.* 2011). Climate and land use changes both can have profound impacts on hydrological processes (Guzha *et al.* 2018, Clerici *et al.* 2019, Afonso de Oliveira Serrão *et al.* 2022), including the timing and magnitude of streamflow, changes in evaporation and transpiration rates, and modifications to sediment transport dynamics. For example, deforestation often leads to increased runoff and decreased infiltration rates, producing higher peak flows (Hou *et al.* 2023), and urbanization decreases lag time (Shao *et al.* 2020, Black *et al.* 2021, Xu *et al.* 2024). Soil properties also play a role by affecting infiltration rates and hydraulic conductivity, further influencing runoff dynamics and lag time (Merz and Plate 1997, Haga *et al.* 2005). The influence of the catchment structure varies depending on the characteristics of each catchment (Fenicia *et al.* 2014) and these characteristics can also be used to investigate hydrological similarities between catchments (Sawicz *et al.* 2011).

Changes in lag time have been used to evaluate the effectiveness of natural flood management (NFM) strategies (Nicholson *et al.* 2012, Kay *et al.* 2019, Peskett *et al.* 2021) which are an ensemble of approaches to *slow the flow*, the guiding principle of NFM (Shuttleworth *et al.* 2019). Understanding the factors that influence catchment response provides valuable insights into planning and designing NFM strategies (Peskett *et al.* 2023). In the Philippines, lag time has proven useful in flood forecasting and disaster risk management. For example, in the Pampanga catchment, observations of tropical cyclones indicate lag times are shorter compared to other regions in the country (Macalalad *et al.* 2021). We suggest that efforts to mitigate the impacts of floods in this catchment could include measures to increase lag time. However, there has been no comprehensive national-scale analysis of lag time in the Philippines. Such analysis could be achieved through studying the data acquired from a national network of continuous rainfall and river discharge monitoring. Although the installation of a Philippine hydrometric monitoring network has begun (Blanco *et al.* 2015), national-scale data collection and the associated rating of water level gauges are in their infancy. Thus, consistent national-scale observational hydrological datasets are not available. Despite this, short-term rainfall and discharge observations have been used to calibrate rainfall–runoff models across the Philippines, as part of the DREAM national flood mapping exercise (Blanco *et al.* 2015). These numerically modelled data thus provide the only feasible resource for investigating controls on lag time, which could inform flood risk management strategies. The 5-year return period was chosen because it is the national baseline for flood hazard mapping in the Philippines (Blanco *et al.* 2015, Lagmay *et al.* 2017). Moreover, sensitivity analyses of rainfall–runoff models show that lag time estimates change little across return periods (e.g. 5–100 years) (Paringit and Pascua 2017). Focusing on the 5-year case therefore provides a robust and representative measure of hydrological response, while ensuring consistency across catchments.

Many strategies have been explored to predict discharge in ungauged catchments, including the use of hydrological modelling and statistical techniques (e.g. Bourgin *et al.* 2015, Dasgupta *et al.* 2024, He *et al.* 2024). Statistical methods in predicting lag time include regional frequency analysis using generalized additive models (GAMs; Chebana *et al.* 2014, Durocher *et al.* 2015), machine and deep learning (Rasheed *et al.* 2022) and time series analysis (Melone *et al.* 2002). These methods can help identify the relationship between predictor variables and lag time, allowing for more accurate predictions of discharge in ungauged catchments. In the UK, the *Flood Estimation Handbook* (FEH) provides a standard national reference in rainfall–runoff analysis and flood frequency estimation, calibrated using extensive hydro-meteorological datasets from gauging stations throughout the UK (UK Centre for Ecology and Hydrology 2024). In the US, the Natural Resources Conservation Service (NRCS) similarly developed an empirical approach based on flow length, slope, and potential retention (National Resources Conservation Service 2008). These methods have been widely applied in mid-latitude regions, and the application of these methods in tropical catchments could considerably underestimate lag time and overestimate runoff volume (Sultan *et al.* 2022). Therefore, it is important to understand lag time in tropical catchments, considering their distinct hydrometeorological and geomorphological conditions.

The aim of the study is to determine a national-scale relationship between catchment characteristics and their hydrological response in terms of lag time in the Philippines. Using lag times corresponding to a 5-year return period for 291 catchments, we applied a GAM to identify non-linear

relationships together with data on catchments' topography, geology, land cover, climate, and location. The model was then extended to predict lag times for 128 of the largest catchments, without prior estimates, for which there were previously no lag time estimates, providing new insights into flood response across the country.

This represents the first national-scale analysis of lag time in the Philippines. Unlike existing empirical methods calibrated in temperate regions, our approach uses a GAM to capture non-linear relationships between multiple catchment characteristics and lag time. To our knowledge, it is the first predictive equation developed specifically for tropical catchments using GAM. The equation provides a new tool for predicting lag time in ungauged catchments, improving national-scale flood risk management in data-scarce regions.

2 Study area

The Philippines comprises more than 7000 islands with a total area of about 300 000 km^2 situated between latitudes 4°N and 22°N, and between longitudes 116°E and 127°E. The archipelago lies in the typhoon belt which poses unique challenges in terms of water resource management (Rola *et al.* 2015) and disaster risk (World Risk Report 2022, 2023). Four general geological groups occur across the archipelago: metamorphic gneiss and schist ophiolites and ophiolitic rocks; magmatic rocks and active volcanic arcs; and sedimentary basins (Aurelio and Peña 2010). The most dominant land cover in the Philippines is crop land (43% – about 127 000 km^2), followed by forest (23%) and then shrubs (20%) (National Mapping and Resource Information Authority 2022). The country has 421 principal river basins, 18 of which are major basins of more than 1000 km^2 (Environmental Management Bureau 2006). Within this broader context, our study focuses on 291 catchments modelled as part of the DREAM and Phil-LiDAR national flood mapping programmes (Blanco *et al.* 2015, Cadiz 2018). These catchments span a wide range of sizes, elevations, and climatic settings across Luzon, Visayas, and Mindanao, making them collectively representative of the hydrological diversity of the Philippines.

The Philippines is frequently impacted by devastating flood events that have severe effects on communities, infrastructure, and the economy (Abon *et al.* 2011, Anticamara and Go 2017, Skoufias *et al.* 2020, World Risk Report 2022, 2023, 2024, Tolentino *et al.* 2024, Boothroyd *et al.* 2025, Quick *et al.* 2025). One of the attempts to assist the country in disaster risk reduction and management was the launch of Project NOAH (Nationwide Operational Assessment Hazards) by the Department of Science and Technology (DOST) in 2012 (UP NOAH 2024). One of its components was the Disaster Risk and Exposure Assessment for Mitigation (DREAM) Program, which was extended by the Phil-LIDAR 1 Program (Blanco *et al.* 2015) whose main task was to generate hydraulic models and flood maps (Lagmay *et al.* 2017). The hydraulic models used for national-scale flood mapping required input hydrographs that were derived from rainfall–runoff modelling using the Hydrological Engineering Center – Hydrologic Modeling System (HEC-HMS) software (Bartles *et al.* 2022) and the rainfall intensity-duration curve of the national meteorological and hydrological agency, DOST-PAGASA (Philippine Atmospheric, Geophysical and Astronomical Services Administration). Through Project NOAH and its associated projects, 291 catchments have been modelled (Blanco *et al.* 2015, Cadiz 2018). The location (Fig. 1) and modelled 5 year lag time of each river were collated from digital reports on the modelling of each catchment. The 291 catchments capture variability in land use (cropland, forest, shrubland), mean annual rainfall, and underlying geology, all of which are known to influence lag time. Summary statistics and distributions of these characteristics are provided in the Data and methods section, and the geographical distribution is shown in Fig. 1.

3 Data and methods

3.1 Overview

In this section, we describe the workflow employed to model the statistical relationship between lag time and catchment characteristics (Fig. 2). Outflow point locations from the Project NOAH rainfall–runoff modelling (Section 3.1.1) served as the basis for delineating catchments. These catchment areas were used to derive 49 topographic properties (Section 3.1.2), such as drainage area, stream length, catchment-averaged slope, elevation, shape and form factors, using TopoToolbox v. 2 (Schwanghart and Scherler 2014). National-scale geology (Section 3.1.3) and LULC (Section 3.1.4) datasets were reclassified into fewer classes to limit the number of variables. Proportions of each class in each catchment were then derived.

A variable selection procedure was used to identify variables which could help explain lag time. The 49 topographic variables were subjected to semi-automated cross-correlation based grouping. One variable per group was then selected as a potential predictor variable (Section 3.2.1). The variables were then subjected to feature elimination (Section 3.2.2). We chose to fit a GAM because it can capture the non-linear relationships between the selected variables and lag time, providing a flexible framework for developing the final predictive model. Finally, a bootstrapping analysis, which involves repeatedly fitting GAMs to a resampled subset of the data, was conducted to ensure the reliability and stability of this model.

3.2 Input data preparation

3.2.1 Lag time data

The 291 lag time estimates (Fig. 1) are from national-scale flood hazard mapping research initiatives (DREAM and Phil-LIDAR projects) of the University of the Philippines (UP) Training Center for Applied Geodesy and Photogrammetry (Blanco *et al.* 2015). Lag time, or time to peak, is defined in this study as the time difference (in hours) between the centroid of effective rainfall and the peak of the discharge hydrograph. Values were obtained from calibrated HEC-HMS rainfall–runoff model outputs, as reported in DREAM and Phil-LiDAR studies (e.g. table 31 in the Bacarra River project report; Paringit and Pascua 2017), where models were validated against observed streamflow. In total, 291 lag time estimates (Fig. 1) were assembled from these national projects for use in this analysis (UP DREAM Program 2024).

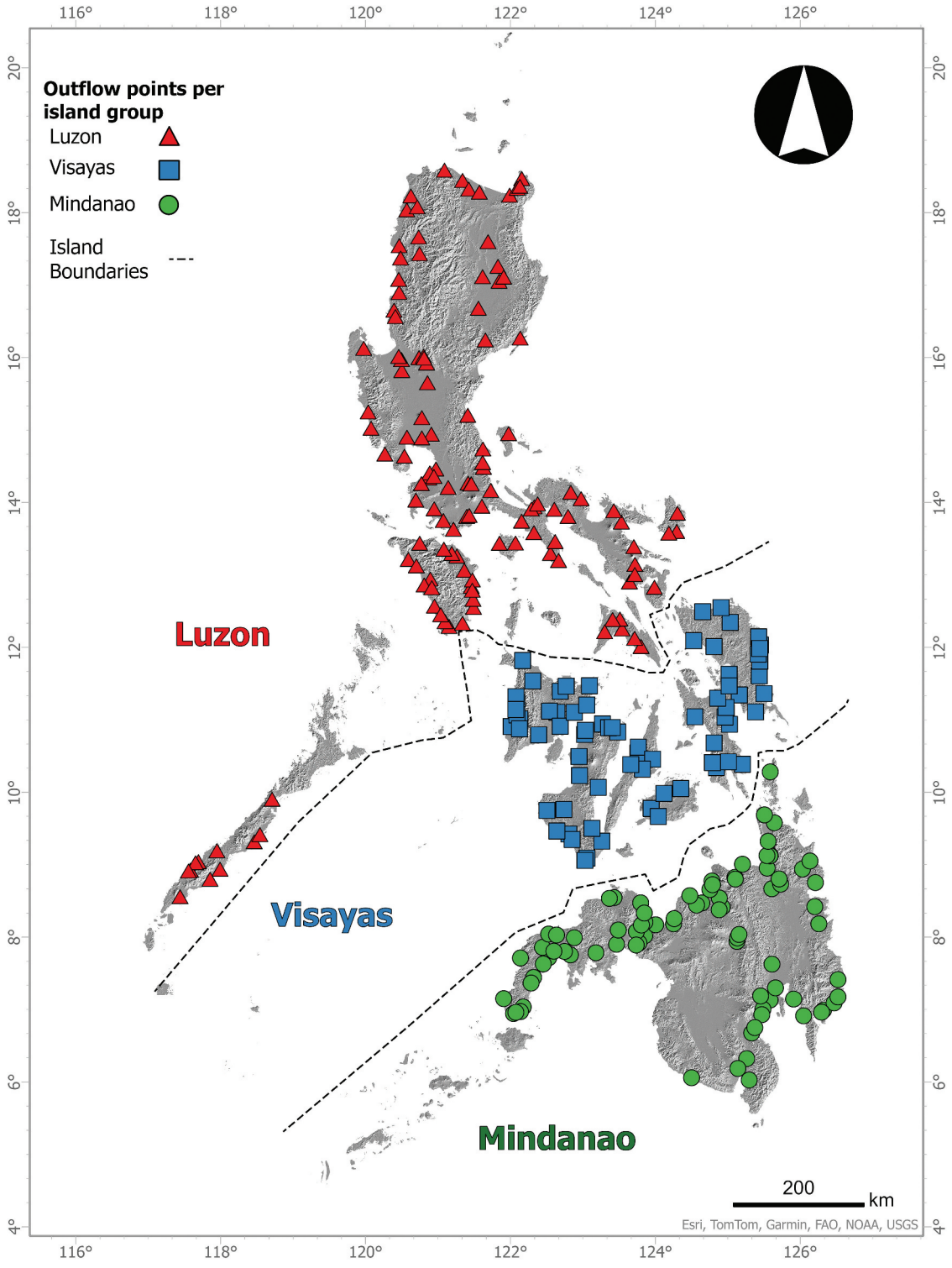


Figure 1. Locations of outflow points of 291 catchments in the Philippines, across the country's three island groups. At each outflow point, lag time was numerically modelled through rainfall–runoff modelling that was undertaken during Project NOAH (Blanco *et al.* 2015, Cadiz 2018). Background is a hillshaded digital elevation model (DEM).

Figure 3 shows the log-log plot of the lag time against catchment area compiled from these reports. Whilst it would be advantageous to use observational data, regional- or national-scale hydrological observations are not available and thus numerically simulated data are the only means to improve hydrological understanding in this data-scarce region.

Input data for the UP DREAM hydrological models included a 10 m resolution digital elevation model (DEM) (Grafil and Castro 2014), soil classification data (Bureau of Soils and Water Management 2004), LULC data (National Mapping and Resource Information Authority 2022), and rainfall data (PAGASA 2024). The

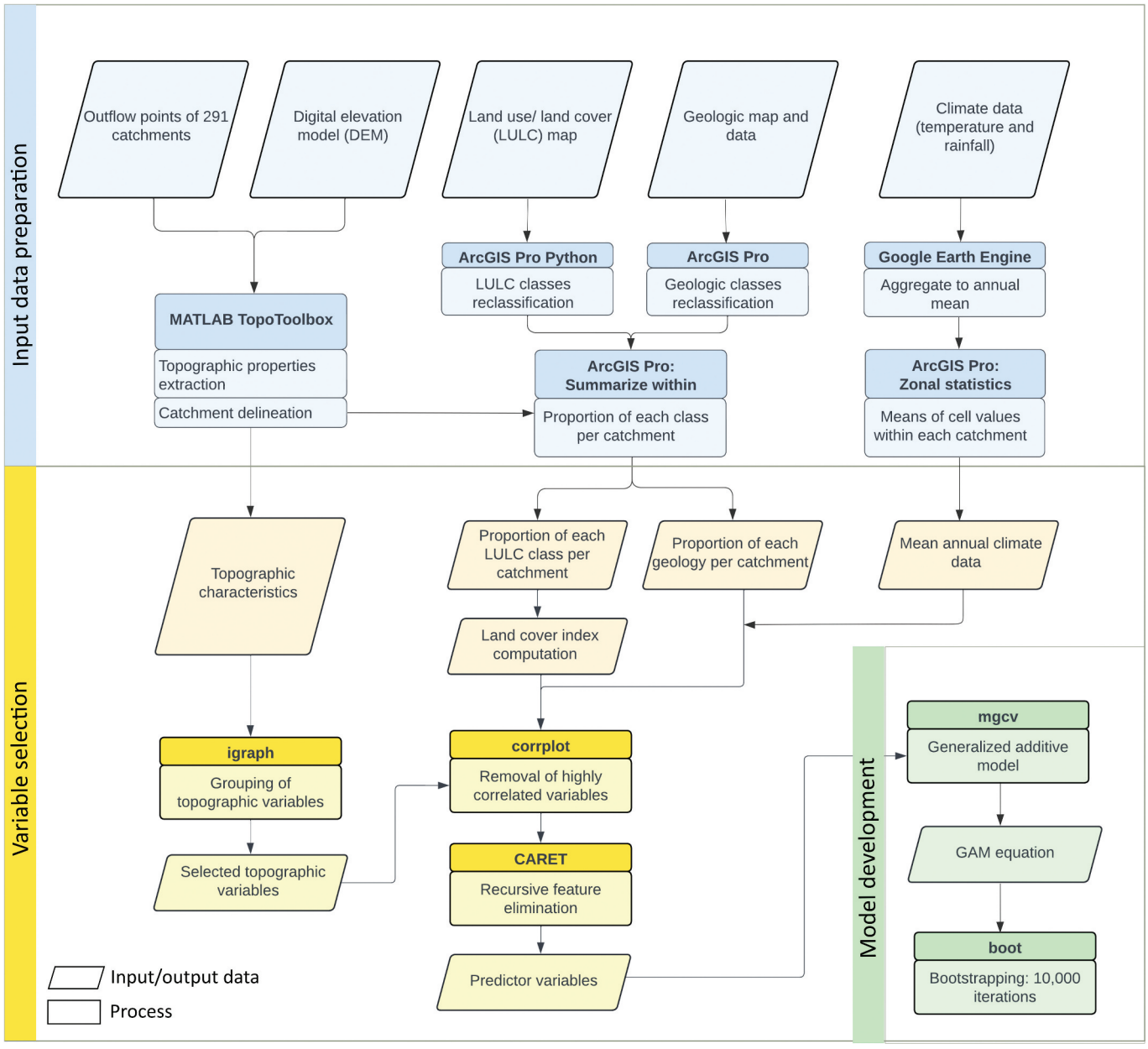


Figure 2. Workflow for extracting input data for variable selection and model development. Outflow points are from LiDAR Survey and Flood Mapping technical reports (UP DREAM Program 2024); geological and LULC geospatial datasets are from the Mines and Geosciences Bureau (2021) and National Mapping and Resource Information Authority (2022), respectively. Temperature data is from the ERA-5 Land reanalysis dataset (Muñoz Sabater 2019) and rainfall data is from the Global Precipitation Measurement (GPM) international satellite mission (Huffman *et al.* 2019). The upper and more densely shaded box of each process contains the software (i.e. MATLAB, ArcGIS Pro, Google Earth Engine) and R packages (i.e. igraph, corrplot, CARET, mgcv, boot) used.

DEM was used for catchment delineation and characterization to obtain input parameters such as catchment area, flow length (i.e. longest stream path within the catchment) and average catchment slope. Meanwhile, the soil and land cover data were used to obtain roughness value estimates for each catchment. An initial lag time was required to initialize the HEC-HMS models and was estimated using a modified SCS lag time equation (Equation 1) calibrated for the Philippines:

$$T_L = \frac{L^{0.8}(S+1)^{0.7}}{560Y^{0.5}} \quad (1)$$

where T_L is lag time (h), L is hydraulic length (m), Y is average catchment slope (m/m), and S is maximum potential retention (mm), which is computed using:

$$S = \left(\frac{1000}{CN}\right) \quad (2)$$

where *CN* is the average curve number, a hydrological parameter used to estimate direct runoff or infiltration from rainfall events, of the catchment derived from the Philippine National Curve Number Map (Lagmay *et al.* 2017). This value was only used for model initialization; the lag times analysed in this study were those produced by the calibrated and validated HEC-HMS models.

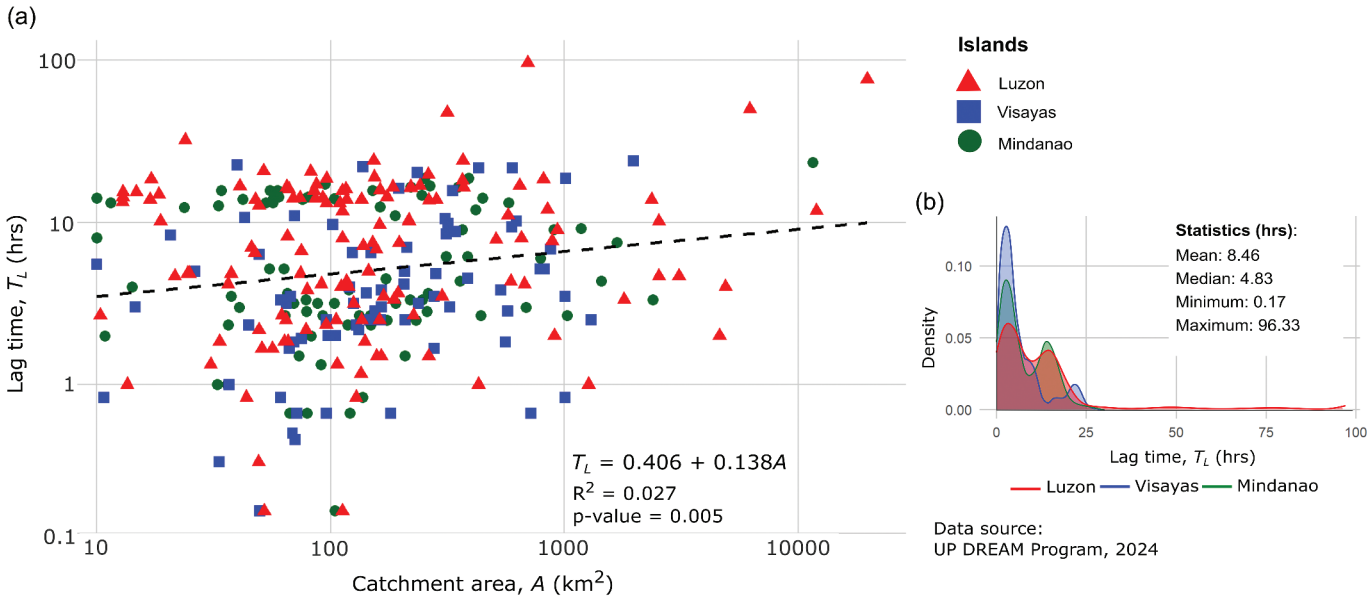


Figure 3. Lag time data from the UP DREAM Program. (a) Scatter plot of lag time vs. catchment area. Data points are colour-coded by island, with Luzon represented by red triangles, Visayas by blue squares, and Mindanao by green circles. (b) Probability density function plots of lag time, with density estimates colour-coded for each island, red for Luzon, blue for Visayas and green for Mindanao. The summary statistics use the data from all islands combined.

The HEC-HMS models were calibrated for each catchment using historical discharge data and recent field discharge measurements. Model performance was then evaluated by comparing the computed values with actual discharge measurements. Examples of model outputs from the Bacarra catchment in Luzon (Paringit and Pascua 2017) and the Imbang catchment in Visayas (Paringit and Sinogaya 2017) and the results of the efficiency tests are shown in Tables 1 and 2, respectively.

3.2.2 Topographic properties

Topographic properties were derived from a nationwide DEM with a 5 m spatial resolution (Grafil and Castro 2014). A workflow using Topotoolbox v. 2 (Schwanghart and Scherler 2014) was applied to extract topographic properties within each catchment. The workflow had previously been developed for a national-scale geodatabase of the rivers and catchments in the Philippines (Boothroyd *et al.* 2023). To make the analysis computationally feasible, the 5 m resolution DEM was resampled to 10 m. The drainage area threshold for defining the stream network was set to 10 km^2. The coordinates of the outflow points of the 291 rivers (Section 3.1.1) were adjusted and snapped to the nearest stream to ensure accurate delineation of the contributing catchment areas. These corrected outflow points were used for catchment delineation, which were subsequently used to summarize additional properties within each catchment.

3.2.3 Geology

The Philippines is situated in one of the most tectonically active regions in the world due to its position at the convergence of the Eurasian, Philippine Sea and Indo-Australian Plates, resulting in a diverse and complex geology (Aurelio and Peña 2010). Geological maps and data were obtained from the Mines and Geosciences Bureau (Mines and Geosciences Bureau 2021), including lithologic units and descriptions that had previously been compiled (Aurelio and Peña 2010). To

Table 2. Results of the model efficiency tests for the Bacarra and Imbang models using several error indices: root mean square error (RMSE); coefficient of determination (R^2); observation standard deviation ratio (RSR); positive percent bias (PBIAS); and Nash-Sutcliffe efficiency (NSE).

Index	Bacarra HMS model	Imbang HMS model	Optimal value
RMSE (h)	1.40	0.20	0.00
R^2	0.89	0.98	1.00
NSE	0.86	0.94	1.00
RSR	0.38	0.24	0.00
PBIAS	−0.98	−1.60	0.00

The optimal values for each index are indicated for comparison, indicating generally good results for both models.

Table 1. Lag time values obtained from HEC-HMS modelling of: (a) the Bacarra catchment (574 km^2) using the Laoag RIDF data (Paringit and Pascua 2017); and (b) the Imbang catchment (163 km^2) using the Iloilo RIDF data (Paringit and Sinogaya 2017).

	Bacarra catchment			Imbang catchment		
RIDF return period (years)	Peak rainfall (mm)	Peak outflow (m^3/s)	Lag time (h)	Peak rainfall (mm)	Peak outflow (m^3/s)	Lag time (h)
5*	31.4	1309	11.00	28.7	101.2	5.50
10	37.2	1729	10.83	33.9	129.4	5.50
25	44.5	2294	10.83	40.5	166.6	5.33
50	50.0	2730	10.00	45.4	194.9	5.33
100	55.3	3178	10.40	50.3	223.8	5.33

*Return period used in GAM analysis.

enable national-scale analysis, lithology was reclassified to a fourfold classification: igneous, clastic sedimentary, limestone, and metamorphic. The reclassification from the initial 32 lithologic descriptions, which were summarized into four geological groups in Section 2 (metamorphic rocks, ophiolites and ophiolitic rocks, magmatic rocks and active volcanic arcs, and sedimentary basins), streamlines the analysis while preserving the hydrologically relevant geological characteristics. Catchment polygons (Section 3.2.2) were used to extract the proportion of each lithology in each catchment. For context, of the total area of all catchments in this study, 47.9% is classified as igneous, 41.3% as sedimentary, 8.4% as limestone, and 2.1% as metamorphic, while 0.2% was unclassified.

3.2.4 Land use and land cover data

LULC data were sourced from the National Mapping and Resource Information Authority (NAMRIA) of the Philippines (Santos 2018). The digital land cover classification used Landsat 8 satellite data with 30 m resolution. Accuracy assessment was undertaken by NAMRIA using reference data from Google Earth, topographic maps, ground validation and IfSAR data to determine the coastline. A classification accuracy of 93.9% using 7330 points across all 12 LULC categories was reported in the confusion matrix (Santos 2018). For the purpose of our analysis, the original 12 categories from the 2015 LULC data were reclassified into eight simplified categories: 1 – built-up area, 2 – crop, 3 – forest, 4 – grassland, 5 – inland water, 6 – open/barren, 7 – shrubs, and 8 – flooded vegetation. Catchment polygons delineated in Section 3.1.2 were then used to extract the proportion of each LULC category within each catchment.

In early iterations of lag time GAM development, there was autocorrelation between land use classes, requiring some further aggregation (Brown *et al.* 2012). To overcome this, while still capturing the influence of land use on lag time, we derived a land cover index (Lc_i) for each catchment as the sum of the proportions of each LULC category (P_{Lj}) multiplied by their hydrological effectiveness as captured via Manning's roughness coefficient (n_j), where the subscript j refers to each LULC category:

$$Lc_i = \sum_{j=1}^{j=12} P_{Lj} \times n_j \quad (3)$$

The values of Manning's n used for each LULC are given in Table S1. The land cover index aims to capture the combined hydrological influence of all different land use classes within a catchment in influencing runoff generation without introducing bias towards individual LULC types, or losing information by discounting/combining particular categories.

3.2.5 Climate data

The climate data used in this study are mean annual temperature (MAT) (°C) and mean annual rainfall (MAR) (mm). For temperature, the ERA5-Land reanalysis dataset was used, known for its high-resolution atmospheric and land-surface data (Muñoz Sabater 2019). Rainfall data was sourced from the Global Precipitation Measurement (GPM) international satellite mission (Huffman *et al.* 2019). The ERA5-Land reanalysis and GPM datasets had temporal resolutions of 1.0 h and 0.5 h, respectively. MAT and MAR were calculated from years 2000 to 2024 within each of the delineated catchments. The zonal statistics tool in ArcGIS Pro was then used to calculate the means of the cell values within each catchment area, providing MAT and MAR values specific to each catchment.

3.3 Variable selection

3.3.1 Variable grouping according to correlation

The *igraph* package in R (Csárdi *et al.* 2024) was used to group the 49 variables (Table 3) based on their pairwise correlation. Variables with a correlation greater than 0.8 ($p < .0001$) were clustered automatically, yielding seven groups. From each cluster, one representative variable was selected based on their ease of extraction, on common practices, and frequent use in hydrological and catchment analyses to be carried forward for regression analysis and model development.

3.3.2 Correlation matrix

A correlation matrix of 14 variables including lag time, seven topographic properties from the variable grouping, four geological categories, one land cover index, two location

Table 3. Results of network analysis to group topographic properties derived from TopoToolbox based on their shared features or attributes.

Automated grouping	Topographic properties in the group	Group description	Chosen variable (units)
1	Perimeter, basin length (Euclidean), basin length (constrained), relative perimeter, total stream length, mean stream length, Stream number, trunk stream length, trunk Euclidean distance, drainage texture, area	Size of the basin	Catchment area (km^2)
2	Form factor (constrained), form factor (Euclidean), elongation ratio (Euclidean), elongation ratio (constrained), circularity index	Catchment form	Form factor (dimensionless)
3	Compactness index 2 (alternative method), compactness index	Basin compactness	Compactness index (dimensionless)
4	Unit shape factor (Euclidean), unit shape factor (constrained)	Catchment shape	Unit shape (dimensionless)
5	Relief ratio (Euclidean), relative relief ratio, relief ratio (constrained), Melton ruggedness number	Basin relief	Relief ratio
6	Maximum elevation, relief, 10% elevation, 50% elevation, 90% elevation, standard deviation elevation, mean elevation, mean slope, maximum slope, 10% slope, 50% slope, 90% slope, standard deviation slope, mean stream slope, median stream slope, standard deviation stream slope, ruggedness number	Elevation and slope	Mean slope (degrees)
7	Infiltration number, drainage density, drainage frequency, drainage intensity	Runoff capacity	Drainage density (km. km^{-2})

The seven chosen variables are listed.

Table 4. Catchment characteristics considered for variable selection.

Type	Catchment characteristic variable name	Definition	Source
Topography	Catchment area	A	Horton (1932)
Topography	Form factor	$A/(L)^2$	Horton (1932)
Topography	Compactness index	$0.28P/A^{0.5}$	Gravelius (1914)
Topography	Shape factor	$L/A^{0.5}$	Horton (1932)
Topography	Relief ratio	R/L	Schumm (1956)
Topography	Mean slope	S	Horton (1945)
Topography	Drainage density	SL/A	Horton (1945)
Geology	Igneous	Proportion of the area of igneous lithologies (km^2)	Mines and Geosciences Bureau (2021)
Geology	Sedimentary	Proportion of the area of igneous lithologies (km^2)	Mines and Geosciences Bureau (2021)
Geology	Metamorphic	Proportion of the area of metamorphic lithologies (km^2)	Mines and Geosciences Bureau (2021)
Geology	Limestone	Proportion of the area of limestone lithologies (km^2)	Mines and Geosciences Bureau (2021)
Location	Latitude	Position of the outflow point	UP DREAM Program (2024)
Location	Longitude	Position of the outflow point	UP DREAM Program (2024)
Climate	Mean annual temperature (MAT)	Mean annual temperature from 2000 to 2024 (°C)	Muñoz Sabater (2019)
Climate	Mean annual rainfall (MAT)	Mean annual rainfall from 2000 to 2024 (mm)	Huffman *et al.* (2019)
Land cover	Land cover index	Equation (3)	National Mapping and Resource Information Authority (2022)

Legend: A = catchment area; P = perimeter; L = catchment length (longest path); R = relief; SL = total stream length (length of all the streams).

parameters (latitude and longitude) and two climate parameters (MAR and MAT) (Table 4) was generated using the *corrplot* package in R (Wei and Simko 2021). For variable pairs with a correlation >0.7, a common (albeit arbitrary) threshold for identifying multicollinearity, the variable with a higher mean absolute correlation to other variables was removed to reduce redundancy and improve the robustness of the model. The resulting set of variables then used for recursive feature elimination (RFE).

3.3.3 Recursive feature elimination

Automated feature (variable) selection was undertaken using RFE from the *caret*(*Classification And REgression Training*) program in R (Kuhn 2008). RFE utilizes a machine learning algorithm to identify optimal variables (Metzger 2004). This process iteratively creates models to evaluate the performance of variables, ranking them based on their elimination order. In datasets with N (number) features, RFE can potentially explore 2 × N combinations. Model complexity is assessed based on the number of features, as a higher number can lead to issues such as multicollinearity and overfitting, where the model becomes excessively tailored to the training data, capturing noise rather than the underlying pattern. Therefore, determining the optimal number of features is crucial. RFE employs repeated cross-validation to ensure the robustness of the selected features, ultimately producing the final set of features based on their predictive power for the lag time.

3.4 Model development

3.4.1 Generalized additive model (GAM)

A GAM is a nonparametric regression allowing non-linear relationships between the response variable, lag time, and key predictors from RFE. It assumes that the response variable (Z) can be modelled by a sum of arbitrary functions of the predictor variables (x_n) rather than linear combinations of variables (Equation 3):

$$Z = s_1x_1 + s_2x_2 + \ldots + s_nx_n \quad (4)$$

where s_n is a spline, a smooth, nonparametric function. GAMs work well in fitting wiggly data because there is no single polynomial required to fit all of the data. To derive a GAM to predict lag time as a function of catchment characteristics, the R package *mgcv* (mixed GAM computational vehicle) and the function *gam* were used (Wood 2017).

From the variable selection (Section 3.2), variables were finalized using an iterative backwards selection procedure based on p values, where the single term with the highest non-significant p value (>.05) was sequentially dropped from the model and then re-fitted until all terms were significant. The degree of smoothness or wiggliness (Wood 2017) was checked to avoid overfitting. This is controlled using penalized regression determined automatically in *mgcv* through generalized cross-validation (GCV) or restricted maximum likelihood (REML).

The flexibility, or "wiggliness," of the model can be controlled by the basis dimension denoted by k. The k parameter determines the number of basis functions (e.g. polynomials, splines) used to create each smooth term in the model. A k of zero results in a straight line and small values of k result in a less wiggly fit, while larger k produces a model that is more flexible but can be wiggly and risks overfitting. For each smooth term, k sets the maximum number of effective degrees of freedom (edf). To ensure that k is appropriate, model validation is essential. We used the *gam.check* function, which runs a simulation-based check that assesses the *edf*, k-index, and p value. Here, the p values test the null hypothesis that k is sufficiently large to model the relationship between the predictor and response variables. A low p value (often when $k < 1$) suggests that k might be too small, especially if the *edf* is close to k. We adjusted the k value until all p values were not significant, indicating that the smooth terms are appropriately sized as described above.

The *mgcViz* package in R (Fasiolo *et al.* 2023) was used to visualize and diagnose GAM fitted using *mgcv*. The function *appraise* was used to produce diagnostic plots to assess the goodness of fit.

3.4.2 Bootstrapping

To assess model stability, robustness and uncertainty, an iterative bootstrapping resampling technique was implemented (10 000 iterations) using the R package *boot* (Canty and Ripley 2024). Each iteration involved randomly sampling 80% of the original dataset without replacement. Subsequently, the GAM equation was fitted to each sample of the data. Relevant statistics were extracted from the model summary and reported here. These included: (a) R^2, the proportion of variance in lag time that is explained by the predictor variables; (b) deviance explained, which is a measure of the difference between the observed lag time values and the values predicted by the mo (c) GCV, which is a method used to estimate effective degrees of freedom and helps prevent overfitting (lower GCV score is generally better but it should be assessed alongside other model diagnostics); and (d) significance of each explanatory variable individually, which assesses whether the inclusion of that variable in the model significantly improves the model's fit compared to a model without the variable. By applying bootstrapping across 10 000 iterations, we were able to evaluate the stability of these metrics and ensure that the model performs robustly even when small portions of the data are excluded.

4 Results

4.1 Variable selection

4.1.1 Reduction of topographic properties

Seven clusters of topographic variables were automatically identified from the initial 49 variables (see Table S2 for the definition of each property), based on their correlation using network analyses. Each group was assigned a name for general description based on the type of variables in each cluster (Table 3). These categories were: 1 – size of the basin, 2 – basin form, 3 – basin compactness, 4 – basin shape, 5 – basin relief, 6 – elevation and slope, and 7 – runoff capacity. Within each category, one property was selected as representative for inclusion, effectively reducing the original 49 topographic properties to seven. The variable chosen was based on which variables are commonly reported or included in hydrological or catchment analyses rather than driven by a statistical metric such as their correlation, but we note that the grouping was correlation-driven in order to avoid complications of autocorrelation.

4.1.2 Correlation of topographic properties, LULC and geology

All 16 resultant variables (Table 4) were subjected to further correlation analysis to test for multicollinearity.

Two pairs of variables were highly correlated (>0.9) to each other: shape factor and form factor; and areas of igneous and sedimentary lithologies. The variables shape factor and sedimentary lithology area have lower correlations than form factor and igneous lithology area, respectively, with lag time and thus were removed.

4.1.3 Recursive feature elimination (RFE)

Figure 4 shows the importance of the 14 different catchment characteristic variables in predicting lag time, as determined by the RFE. The variable importance of each feature/variable is ranked, with relief ratio identified as the most important in predicting lag time, followed by the other variables in descending order of significance. Metamorphic rock area is shown as the least important variable of those analysed (Fig. 4). RFE also identified 14 as the number of features to select to carry forward. Figure 5 shows a scatter plot matrix to visually explore the relationships between these catchment variables. Half of the variables have a positive correlation with lag time while the other half show negative correlations. There are also pairs of variables that exhibit strong correlations, indicating that changes in one variable are closely associated with changes in another. For example, the land cover index and mean slope are positively correlated, suggesting that specific land cover types tend to occur on particular slopes (Fig. 5).

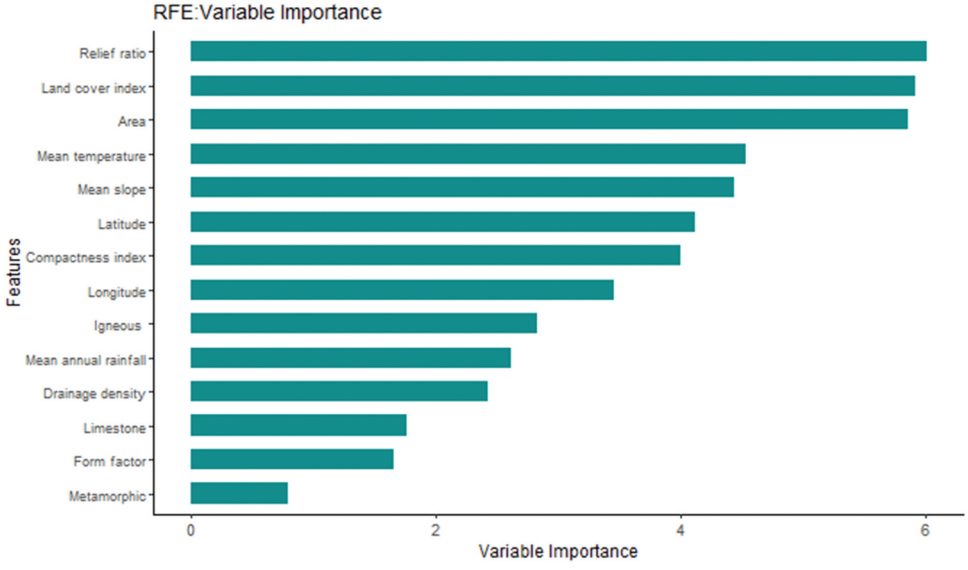


Figure 4. Variable importance for predicting lag time using recursive feature elimination (RFE). This bar plot illustrates the relative importance of variables in predicting lag time, as determined by the RFE algorithm. Each bar represents the importance of a variable, with longer bars indicating greater importance.

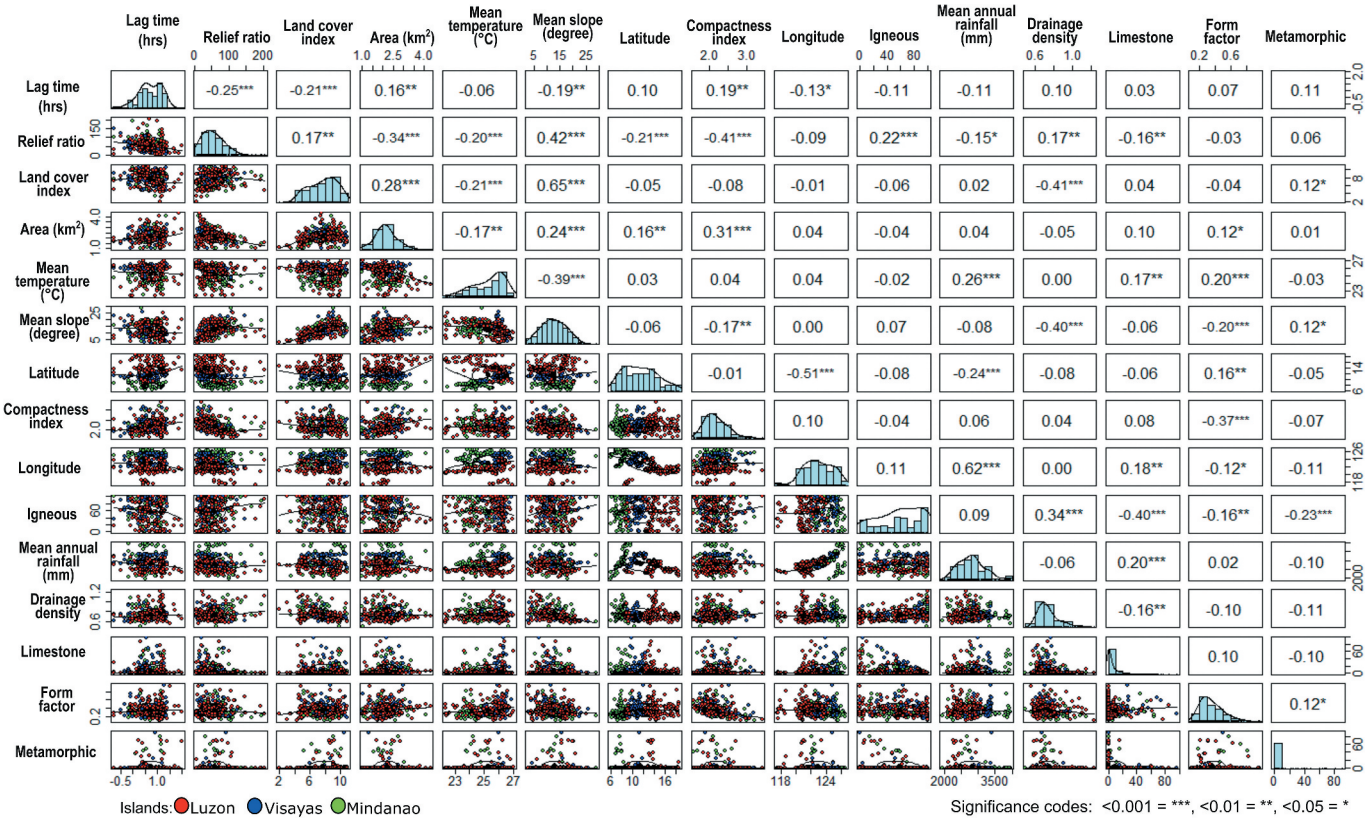


Figure 5. Scatter plot matrix of important variables resulting from RFE and GAM. Each panel in the matrix displays the relationship between pairs of variables, including lag time, relief ratio, area, land cover index, MAT, mean slope, latitude, compactness index, longitude, igneous, MAR, drainage density, limestone, form factor and metamorphic (from left to right, top to bottom). The diagonal panels show histograms of each variable. The scatter plots provide insights into potential associations between the variables. Correlations (R) are shown in the top right of the figure, with significant indicated by *for $p < .05$; **for $p < .01$; ***for $p < .001$. Units are displayed on the left and right y-axes. The points are colour-coded based on island group – red for Luzon, blue for Visayas, green for Mindanao.

Similarly, a strong correlation was observed between longitude and MAR (Fig. 5), which may correspond to the boundaries of different climate types. This suggests that geographical location, defined by longitude, is a key factor influencing rainfall patterns.

4.2 GAM results

After performing variable selection through correlation and RFE to identify the most significant variables, we modelled these 14 variables using GAM. During this process, the model was further refined by eliminating non-significant variables, resulting in a final model with 10 predictors (Table 5).

4.2.1 Final model

The final GAM equation with an R^2 of 0.77 and deviance explained of 0.84 is:

$$\begin{aligned} \text{Lagtime} = {} & s(reliefratio, k = 4) + s(landcoverindex, k = 4) \\ & + s(log(area), k = 10) + s(meanslope, k = 10) \\ & + s(latitude, k = 12) + s(compactnessindex\_2, k = 16) \\ & + s(longitude, k = 10) + s(Igneous, k = 16) \\ & + s(rainfall, k = 11) + s(formfactor, k = 13) \end{aligned} \tag{5}$$

where lag time is the response variable, and s(., k) represents the smooth functions for each of the 10 predictor variables with specified number of basis functions k, respectively. Table 5

Table 5. Summary of the model specifications including predictors, significance, and overall model fit measures.

Number	Variables	Significance level	k	R^2	GCV	Deviance explained
1	Relief ratio	***	4	0.77	29.88	83.7%
2	Land cover index	***	4			
3	Catchment area	***	10			
4	Mean slope	***	10			
5	Latitude	***	12			
6	Compactness index	***	16			
7	Longitude	**	10			
8	Igneous lithology area	***	16			
9	Mean annual rainfall	***	11			
10	Form factor	***	13			

Significance codes: ***, <0.001; **, <0.01; *, <0.05.

summarizes the predictors and their significance and the overall goodness-of-fit measures of the final model (i.e. R^2, deviance explained). The GCV score of 29.88 suggests a reasonably good fit, but its adequacy depends on the scale of the response variable and comparison to alternative models. The final model used variables identified using RFE selection (Fig.4). Figure 6 (a) shows the individual effect of each variable. Notably, some variables exhibit a clearer effect on lag time. Confidence intervals around the partial effects are generally narrow across the central range of observed predictor values, indicating robust model fit. Intervals widen at the extremes, particularly for large catchments and steep slopes, where fewer observations are available. This reflects greater uncertainty at the margins of the data and should be considered when interpreting predictions outside the main data range. For example, lag time generally increases as the relief ratio decreases, indicating an inverse relationship between these variables while lag time increases with catchment area. Figure 6(b) shows the residuals of the

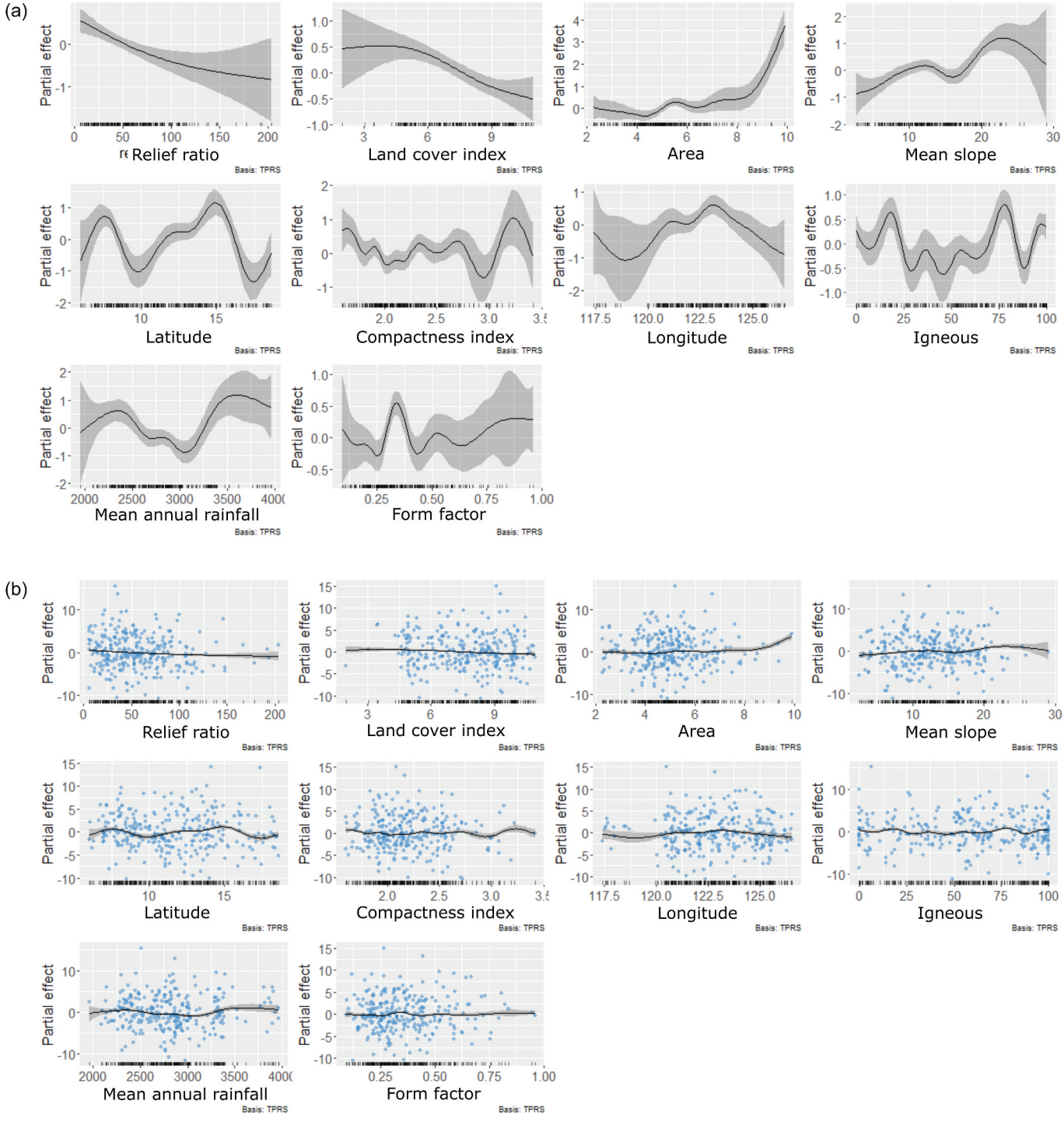


Figure 6. (a) Partial effects of the covariates in final GAM (see Table 4). The plots are centred on the mean and the shaded area represents the 95% confidence interval. (b) Residual plots of each covariate. In (a) each rug mark (small vertical line along the x-axis) represents an individual data point. In (b) data points are shown by both rug marks and blue dots.

Table 6. Summary of the final set of predictor variables based on Pearson's correlation and GAM.

Predictor variables	Expected relationship to lag time	Scatter plot matrix: Pearson's correlation	Scatter plot significance	GAM significance
Relief ratio	(-) The relief ratio measures the difference in height between the highest and lowest points in the basin. Higher relief ratio results in shorter lag time (Altaf *et al*. 2013).	–	***	***
Land cover index	(+) Land cover index (land cover × roughness) factors in roughness which should increase lag time (Bond *et al*. 2020).	–	*	***
Catchment area	(+) As the catchment area gets bigger, the lag time is expected to be longer (Garzon *et al*. 2023).	+	***	***
Mean slope	(-) This is a measure of steepness. Higher (steeper) slope causes water to go faster downhill resulting in shorter lag time (Obeidat *et al*. 2021, Garzon *et al*. 2023).	–	***	***
Latitude	Catchment properties and rainfall may both vary systematically with latitude, but these effects will be captured by other variables and relationships with latitude result from autocorrelation.	+	*	***
Compactness index	(+) A measure of 1 indicates that the catchment is near circular in shape. Higher value means more compact/elongated which results in longer lag time (Garzon *et al*. 2023).	+	**	***
Longitude	There is no expected relationship as this is just a location parameter.	–	**	**
Igneous	(-) Igneous rocks are generally more resistant to erosion. Porosity is limited in fractures thus most of the water becomes runoff. A higher percentage of igneous results in shorter lag time (Hale and McDonnell 2016).	–	Not significant	***
Mean annual rainfall (MAR)	(-) A high amount of rainfall often leads to shorter lag time.	–	Not significant	***
Form factor	(-) A value of 1 denotes a circular shape while 0 or lower value denotes an elongated shape. Higher value then indicates a shorter lag time.	+	Not significant	***

Significance codes: ***, <0.001; **, <0.01; *, <0.05.
+, positive relationship; -, negative relationship.

Table 7. Summary of catchment properties included in the GAM equation for the six Panay Island catchments.

	Panay catchments					
Catchment characteristics	Sibalom	Cagaranan	Tigum	Jalaur	Aklan	Panay
Relief ratio	35.09	59.03	39.55	31.82	32.03	18.46
Land cover index	6.3	7.04	4.5	4.51	7.12	4.62
Area (km^2)	628	298	412	1687	891	1996
Mean slope (°)	18.95	24.9	12.3	10.16	22.97	11.84
Latitude (dd)	10.81	11.02	10.75	10.78	11.73	11.48
Compactness index	2.43	2.32	2.46	2.62	2.6	2.74
Longitude (dd)	121.95	122.04	122.58	122.65	122.37	122.83
Igneous (%)	62.85	60.05	35.41	35.64	68.68	71.7
Mean annual rainfall (mm)	2877.38	3049.58	2542.61	2470.47	3185.92	2907.96
Form factor	0.27	0.23	0.23	0.41	0.18	0.42
Predicted lag time (h)	3.99	6.03	10.31	24.25	26.83	32.12

model, which represent the differences between observed and predicted values.

Out of 56 different catchment characteristics, only 10 were retained to be included in the model, which resulted from variable selection and GAM (Table 6). These 10 variables are relief ratio, land cover index, area, mean slope, latitude, compactness index, longitude, igneous, MAR, and form factor. The significance of each variable from correlation analysis (Section 4.1.2) and GAM (Section 4.2.1) is also presented in Table 6.

The high number of variables needed to explain the variance shows the complexity in predicting lag time (Supplementary Tables S2 and S3). For example, 65% of the iterations with $R^2 > 0.77$ (best fit model) did not have any non-significant variables, with all 10 variables have p values < .05 (Supplementary Table S3). Latitude is significant in 100% of all iterations (Supplementary Table S2), showing its strong relationship with lag time, as also shown in Fig. 5.

The final GAM was also evaluated through a series of diagnostic plots (Fig. 7). The Q-Q residuals plot indicates deviations from normality, especially at the tails (Fig. 7(a)). There is also clustering and a potential trend where residuals change in a non-random manner as the linear predictor values increase, suggesting that there is a pattern or some structure that the model has not fully captured (Fig. 7(b)). The histogram of residuals shows an approximately symmetric distribution centred on zero, with a slight positive skewness value of 0.28. Additionally, the histogram shows the presence of some outliers (Fig. 7(c)). In terms of the accuracy of the model's predictions, most points are clustered around the lower values, with a few points deviating from the main cluster suggesting that the model captures the lower range well and might have issues with fitting higher values (Fig. 7(d)).

4.2.2 Bootstrapping

Bootstrap analysis was conducted to evaluate the robustness and the stability of the GAM, particularly given the limited size of the dataset. Reanalysis was conducted with 10 000 iterations by randomly selecting 80% of the data from the original dataset without replacement (Fig. 8; Table S3 and S4). The mean R^2 value across the bootstrap iterations was 0.77, with a median of 0.79, while the R^2 value for the best-fit model was 0.77. The mean deviance explained across bootstrap iterations was 0.83, with a median of 0.85, while the explained deviance of the best-fit model is 0.84. These results suggest that the bootstrap

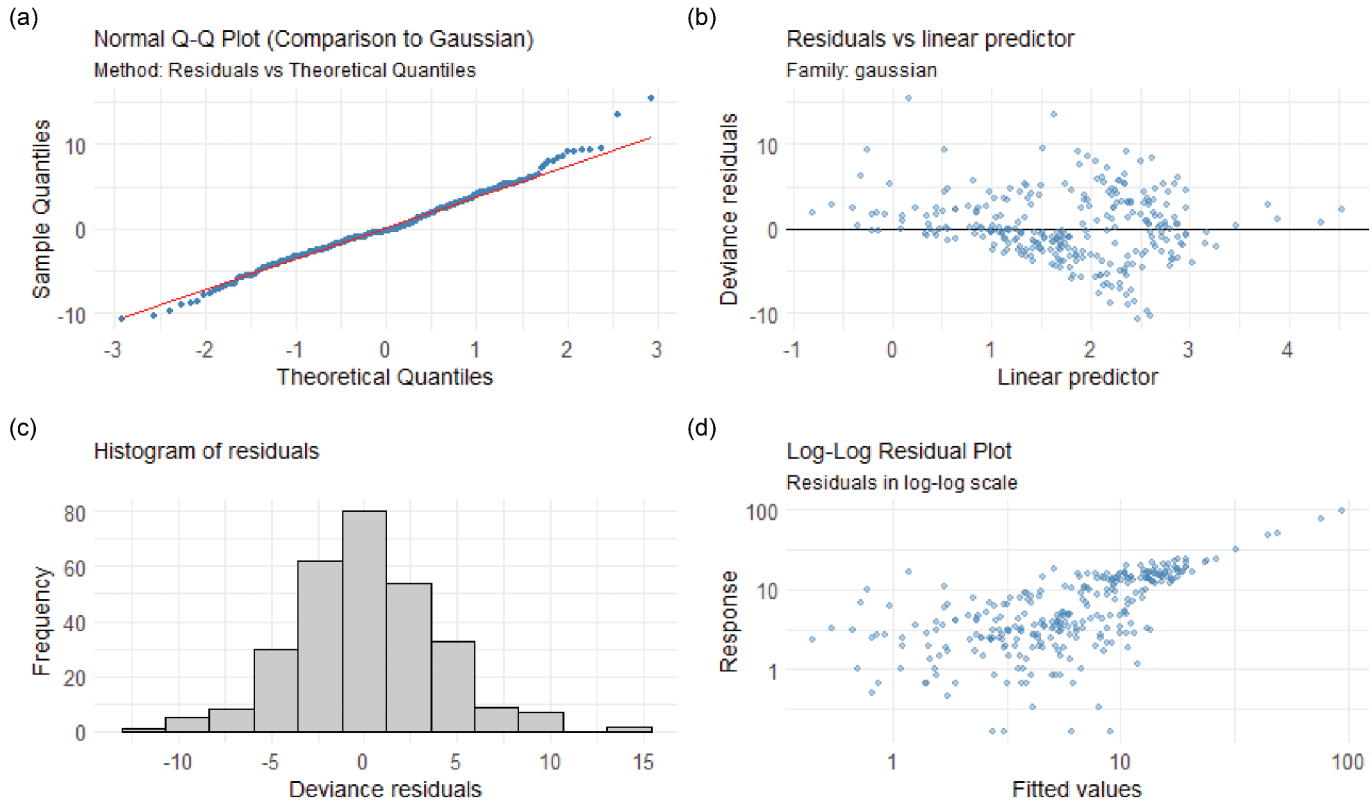


Figure 7. Model diagnostic plots of the fitted GAM. (a) Normal Q-Q residuals plot comparing residuals to a Gaussian distribution; (b) residuals vs. linear predictor plot to assess the model fit under the assumption of normally distributed; (c) histogram of residuals to check for normality; and (d) observed vs. fitted values for lag time ((in hours) on a log-log scale), used to assess the model's predictive accuracy across the different magnitudes of lag time.

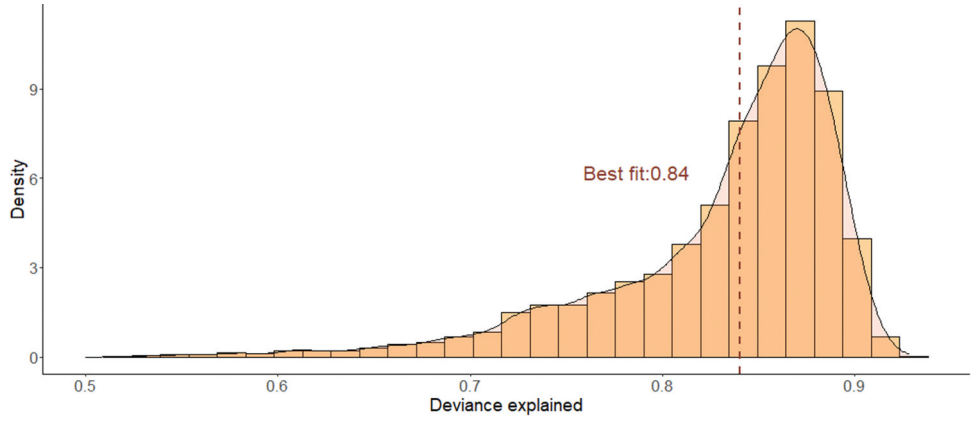


Figure 8. Histogram showing the distribution of deviance explained by the bootstrap results from 10 000 iterations in GAM analysis. The vertical line shows the best-fit model (using all the data): mean: 0.83, median: 0.85, best fit: 0.84. Deviance explained is the metric used in assessing the fit of the models that do not assume a normal distribution of errors.

distribution of R^2 and deviance explained closely resemble the values obtained from the best-fit model, indicating the stability and reliability of the model estimates.

5 Discussion

5.1 Relationship of lag time and catchment characteristics

Machine learning algorithms for variable selection and a GAM approach were used to assess the non-linear controls of catchment characteristics on lag time variability in 291 catchments in the Philippines. Due to its flexibility, GAM is becoming widely used in hydrological studies such as in concentration–discharge relationships (von Brömssen *et al.* 2023), regional frequency analysis (Chebana *et al.* 2014, Jung *et al.* 2024), and reservoir management (Auffray *et al.* 2023, Brunner and Naveau 2023). Additionally, GAM has an interpretability advantage, where the contribution of each independent variable to prediction is clearly encoded (Larsen 2015). Our final model includes 10 variables significant in predicting lag time, which is supported by previous studies (Table 6).

Specifically, variables such as drainage area, land cover index, and compactness index generally have positive relationships with lag time (Bond *et al.* 2020, Garzon *et al.* 2023). In contrast, those that have a negative relationship include relief ratio, mean slope, the area of igneous lithology and form factor (Altaf *et al.* 2013, Hale and McDonnell 2016, Obeidat *et al.* 2021, Garzon *et al.* 2023, Raja Shekar and Mathew 2024). The final model was rigorously tested, and its reliability is shown by the high overall R^2 (Tables S2 and S3). The high R^2 and deviance explained indicates that the model, with all variables included, effectively explains a significant proportion of the variance in lag time and underpins the importance of considering the interplay between multiple factors to accurately understand controls on, and predict lag time, and highlights the robustness of the model in capturing the complex dynamics of catchments in the Philippines.

5.2 Evaluation of land cover index and its impact on lag time determination

The land cover index did not exhibit the expected positive relationship with lag time. Higher values of the land cover index, indicating a greater proportion of LULC with high Manning's roughness, were anticipated to increase lag time by delaying runoff entering the river channels. However, the observed negative relationship between land cover index and lag time suggests that the index alone is insufficient for capturing the complexity of runoff responses, hence we conducted sensitivity analysis on land cover index (Section 5.4). Further, catchments with similar land cover index could have different proportions of LULC (Fig. 9), highlighting a limitation of using a simplistic land cover index. Furthermore, the spatial configuration and distribution of each land cover type across catchments, which affects hydrological response (Fiener *et al.* 2011, Azuka and Igué 2020), is not captured by a simplistic land cover index. Focusing a land cover index solely on the combination of land cover and associated hydraulic roughness may overlook critical components of the hydrological system, such as hydraulic conductivity, surface permeability, and surface storage, which all vary with different land cover types and soil properties (Boorman *et al.* 1995, Archer *et al.* 2013, Vogel 2019, Vereecken *et al.* 2022). However, the lack of a consistent national-scale soil dataset prevents a comprehensive analysis of these variables across catchments. As a result, incorporating soil data into this study was not feasible. Incorporating local and regional scales, along with detailed mapping of catchment properties, into land cover indices offers a fuller understanding of hydrological systems. To advance this, future research should explore spatial regression techniques using observational data. These models account for the spatial distribution of land uses, enhancing predictions of lag, and pinpointing critical areas for targeted data collection.

Supporting these observations, studies conducted across various temporal scales employing different techniques have consistently shown that LULC changes, such as deforestation and urbanization, significantly affect hydrological processes such as increased surface runoff and sediment yield (Abdulkareem *et al.* 2018, Younis and Ammar 2018, Afonso de Oliveira Serrão *et al.* 2022, Waikhom *et al.* 2023). These changes all tend to reduce lag time in catchments. Conversely, reforestation programmes that promote ecological reconstruction, such as the Grain-for-Green programme that converts cultivated land to forest, lead to reduced runoff and soil erosion (Deng *et al.* 2012). These findings underscore the importance of considering the spatial distribution of land cover types in hydrological studies and in the planning and implementation of NFM measures.

5.3 Improved data collection and utilization

Improving the determination of lag time is a critical aspect of hydrological research, particularly in countries like the Philippines that are often regarded as suffering from data scarcity (Necesito *et al.* 2023, Dasallas *et al.* 2024). Despite these challenges, data-driven numerical modelling can significantly improve our understanding of hydrological processes until such time as observational networks are more comprehensively developed. While the use of HEC-HMS model lag time outputs in our analysis ensures consistency, it also presents potential limitations due to overlapping input data such as the DEM and LULC. Although this overlap may introduce bias, it is important to note that differences in how these data are extracted and utilized in our analysis vs. the HEC-HMS modelling can lead to variations in the results. To ensure the robustness of this study, we employed bootstrapping and sensitivity analyses, demonstrating the usability and integrity of our approach even in the presence of overlapping input data.

This study highlights the importance of analysing available data collected by government agencies, research institutions or non-governmental organizations to extract meaningful insights, that can guide targeted data collection efforts, prioritization of future research, and resource allocation. Data quality is crucial in assessing hydrological responses. Lag time data used in this study are from hydrological models which are validated in single-event scenarios. However, the recent installation of 179 water level monitoring stations with raingauges (Tandem) nationwide by DOST-ASTI (https://philsensors.asti.dost.gov.ph/), in 2022, marks a significant advancement in data collection. As these stations continue to gather data over time, this observational data should be used to refine and update lag time determinations, improving their accuracy and reliability. This progress can motivate funders and local government to continue supporting the operational and maintenance costs of these gauging stations.

Additionally, this highlights that data collection should be viewed not as an end in itself but as a vital input for analyses that deepen our understanding of catchment hydrology, crucial for effective management. Open access to hydrological data, such as discharge and streamflow, is essential for democratizing information access, making it available to all, regardless of financial and institutional barriers. Equally important is the inclusion of comprehensive metadata, as it provides critical context which is key in the usability and reliability of the data. Moreover, open access to all information facilitates broader scientific inquiry by allowing researchers from diverse fields to explore the data, fostering collaboration across disciplines to address complex challenges like flooding.

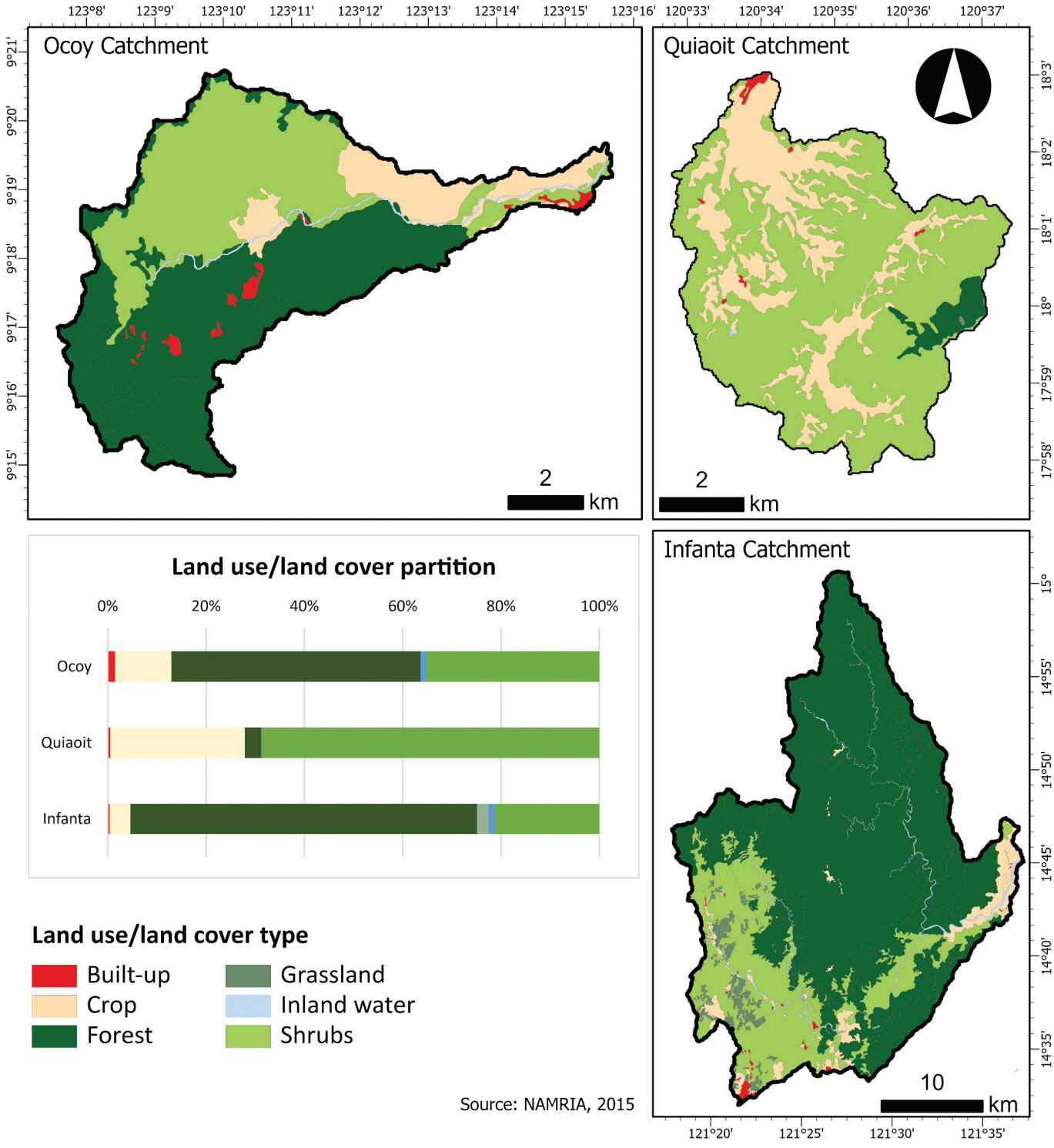


Figure 9. LULC maps and bar plots of three catchments with similar land cover index (9.95). The maps show the different distributions of each LULC type in each catchment while the bar plots show the different proportions of each LULC type per catchment.

5.4 Predicting lag time to inform flood management

Predicting lag time and applying the findings across various catchments can greatly enhance flood management strategies. We expanded our analysis by applying the final GAM equation to estimate lag time for 128 large catchments that previously lacked such data. This equation integrates topographic properties, geology, LULC, location, and climate variables, as detailed in Sections 3.1.3, 3.1.4, and 3.1.5. Examining the catchments in Panay Island reveals numerous factors affecting lag time (Fig. 10; Table 7). For instance, although the Cagaranan catchment has the highest relief ratio, the smallest area, and the steepest slope, it does not exhibit the shortest lag time. Similarly, while the Tigum and Jalaur catchments have similar land cover indices, MAR, and

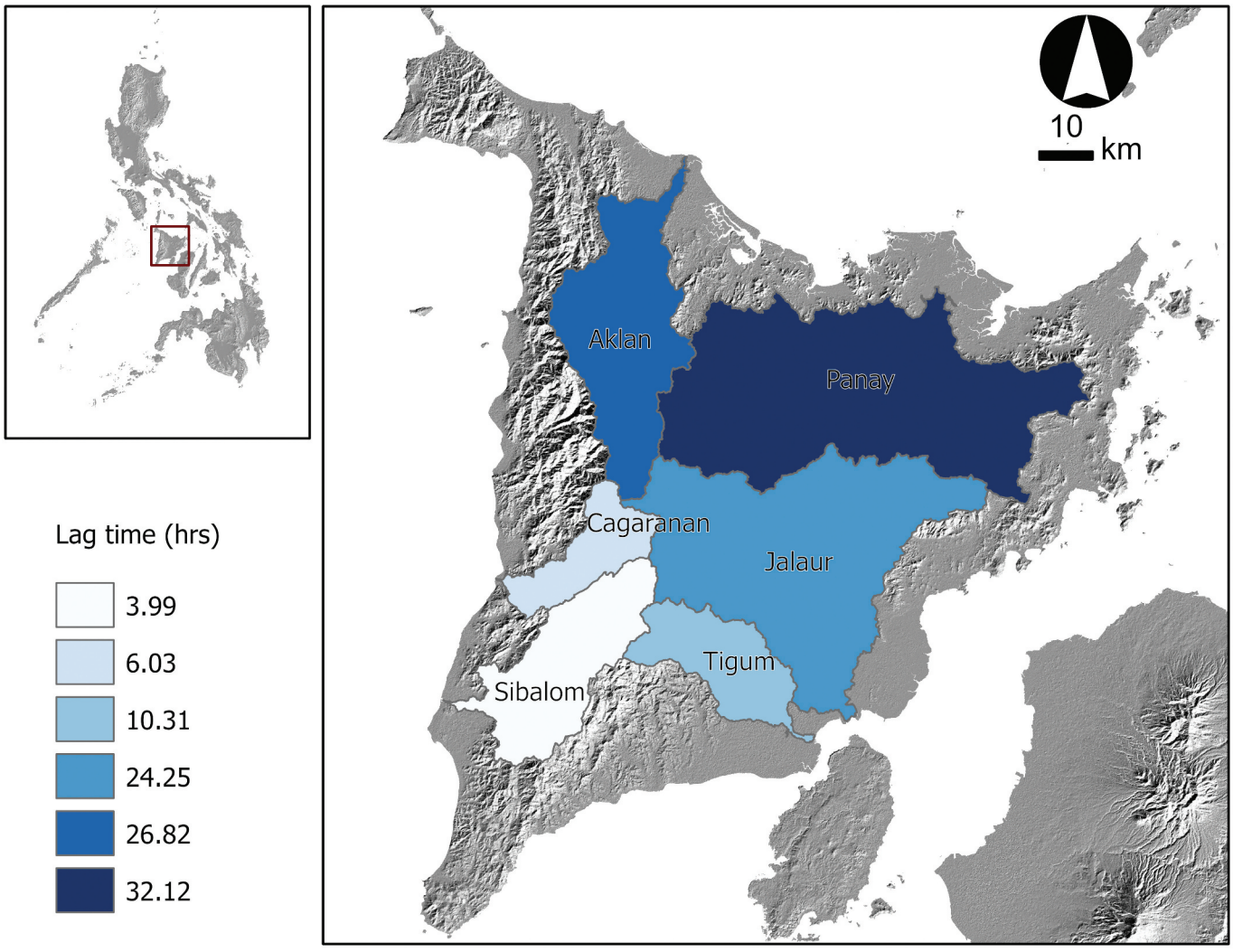


Figure 10. Map of catchments in Panay Island in Visayas (Fig. 1) highlighting variations in lag time.

slope, Jalaur has more than double the predicted lag time of Tigum. These observations suggest that lag time is influenced by a complex interplay of factors, highlighting the need for a more nuanced understanding of the hydrological processes at work in these catchments.

Furthermore, including latitude and longitude in our model allows us to predict lag times for any location within the catchment, extending even to its sub-catchments. While this study initially used lag times modelled at catchment outlets, our approach now allows for predictions at critical points within a catchment, such as different tributaries having different lag times (Pattison *et al.* 2014). This has important implications for flood management, as it helps identify "problematic" areas of the catchment where targeted interventions can be implemented, leading to more precise flood management strategies and better prioritization of management actions.

Among the variables in the GAM equation, only the land cover index can be practically adjusted to increase lag time. As reported in Section 5.2, the sensitivity analysis indicates that lag time increases with higher land cover indices, peaking when the index is between 3 and 4. Beyond this range, however, lag time begins to decrease (Table S5; Fig. 11). For example, increasing the land cover index from 1 to 2 results in a 3.9% increase in lag time, while increasing the index from 6 to 7 can lead to a decrease of up to 12.1%. To illustrate this, we can compare two catchments, Rio Grande de Mindanao (18 513 km^2) and Cagayan (27 684 km^2). Despite the larger size of Cagayan, Rio Grande de Mindanao has nearly four times the lag time when both catchments have similar land cover indices, demonstrating how adjustments to land cover can significantly impact flood management outcomes (Fig. 11).

5.5 Implications for natural flood management

Lag time elongation is a key strategy in flood risk management (FRM), aimed at reducing peak flow and slowing water passage through a catchment. In recent years, there has been growing recognition of the importance of lag time in NFM, particularly for temperate climates (Lane 2017, Burgess-Gamble *et al.* 2018). While understanding and managing lag time is crucial in both temperate and tropical rivers, where the goal is often to stagger water delivery from sub-catchments to the main trunk, the higher runoff potential in tropical rivers, due to their generally wetter climate, adds an

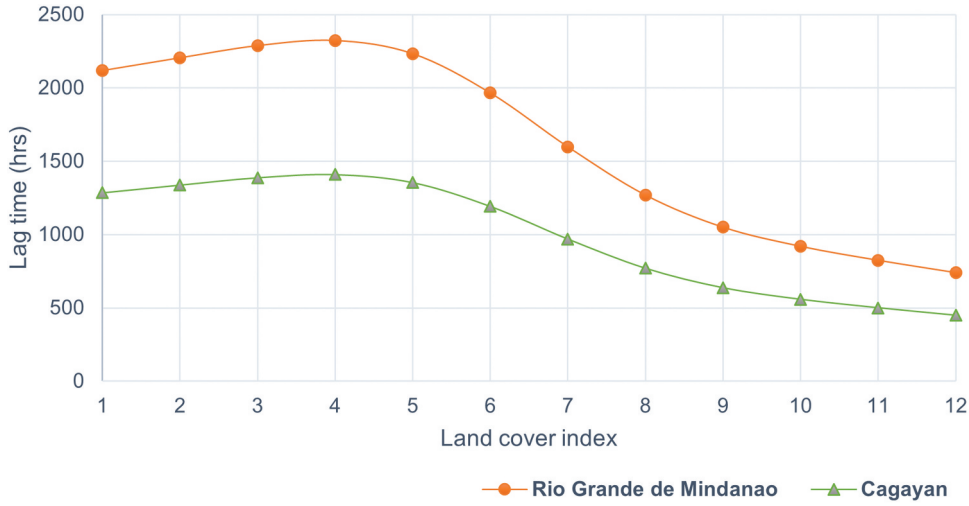


Figure 11. Sensitivity analysis results showing the effect of different land cover indices on lag time for catchments of different areas (Rio Grande de Mindanao, 18 513 km^2; Cagayan, 27 684 km^2).

additional layer of complexity (Tolentino *et al.* 2025). This complexity is expected to intensify with climate change (Seneviratne *et al.* 2012, Tolentino *et al.* 2016). Current and forecast increases in exposure to flood risk warrants the shift to a more integrated approach to FRM, including NBS. Land management, such as reforestation, has also been identified as an effective intervention for flood risk as it reduces stream discharge by intercepting rainfall, effectively increasing lag time (Gao and Yu 2017, Pamungkas *et al.* 2017). In the Philippines, the National Greening Program (NGP), initiated in 2011 and expanded under the Enhanced NGP, has reforested more than a million hectares with the aim of restoring forest cover and reducing flood risk (Juanico 2025). The programme specifically targets upland and denuded slopes, which are major contributors to downstream flooding, and employs watershed forestry approaches such as silviculture, agroforestry, and riparian buffer restoration. While systematic evidence of impacts on lag time is still limited, reforestation is expected to enhance infiltration and reduce surface runoff, thereby lengthening lag time. Continued hydrological monitoring is needed to confirm long-term effects. Small-scale runoff attenuation features, i.e. structures placed in-channel or on floodplains to slow down water (Quinn *et al.* 2022), can also contribute to lag time elongation, but their effectiveness may be more context-specific. For instance, recent work from Nepal highlights the importance of spatial targeting in flood management measures, suggesting that effectiveness of interventions may vary depending on the specific scale and hydrological context (Pearson *et al.* 2022, Sharma *et al.* 2023). Moreover, the integration of these strategies can create a more resilient approach to FRM. Accurate prediction of lag time becomes increasingly important in this context, as it enables a more holistic approach that not only mitigates flood risks but also plays a crucial role in controlling the transport of pollutants in Philippine catchments (Gabriel *et al.* 2023). In addition to playing an important role in understanding flood processes (Barbero *et al.* 2022), lag times also affect ecological and ecosystem dynamics (Sakan *et al.* 2021, Gao *et al.* 2022), through influencing the transport of nutrients (Sapač *et al.* 2020), sediment, and pollutants such as metals, pharmaceuticals, plastics and PFAS (Burgess-Gamble *et al.* 2018, van Emmerik *et al.* 2022, Zhang *et al.* 2022, Byrne *et al.* 2024). Species distributions and abundance are also affected by lag time as different organisms have varying tolerance to water flow rates, and their ability to disperse or colonize new habitats depends on flow conditions (Bunn and Arthington 2002).

The identification of catchments with shorter lag times facilitates targeted flood management actions such as the installation of gauging stations, flood hazard mapping, early flood warning systems, and land use planning (Fig. 12). This approach presents another application of a national-scale geodatabase for inter-catchment comparisons (Boothroyd *et al.* 2023), revealing patters in lag time variability across different regions. While land use planning presents the most viable solution for managing lag time within catchments, it is important to recognize the influence of other factors such as topography and geology. Given the inevitability of water-related challenges, it is essential to adopt strategies like NFM that emphasize making space for water. Strategic planning of land use activities such as urban development, agriculture, and forestry can mitigate the impact of land use changes on hydrological processes, leading to increased lag time. This, in turn, is beneficial for slowing water flow and enhancing the effectiveness of NFM strategies.

Beyond these practical applications, the modelling approach itself offers advantages over conventional lag-time estimation methods. Empirical equations such as those from the NRCS or the UK Flood Estimation Handbook were developed in temperate regions and often underestimate lag times in the tropics (Sultan *et al.* 2022). In contrast, the GAM approach captures non-linear relationships, accommodates diverse predictor types (topographic, climatic, geological, and land cover), and provides interpretable partial effects for each variable. Because the model was developed using Philippine catchment data, it avoids the biases associated with transferring formulas from other contexts and provides a flexible and reliable tool for national-scale applications in data-scarce tropical settings.

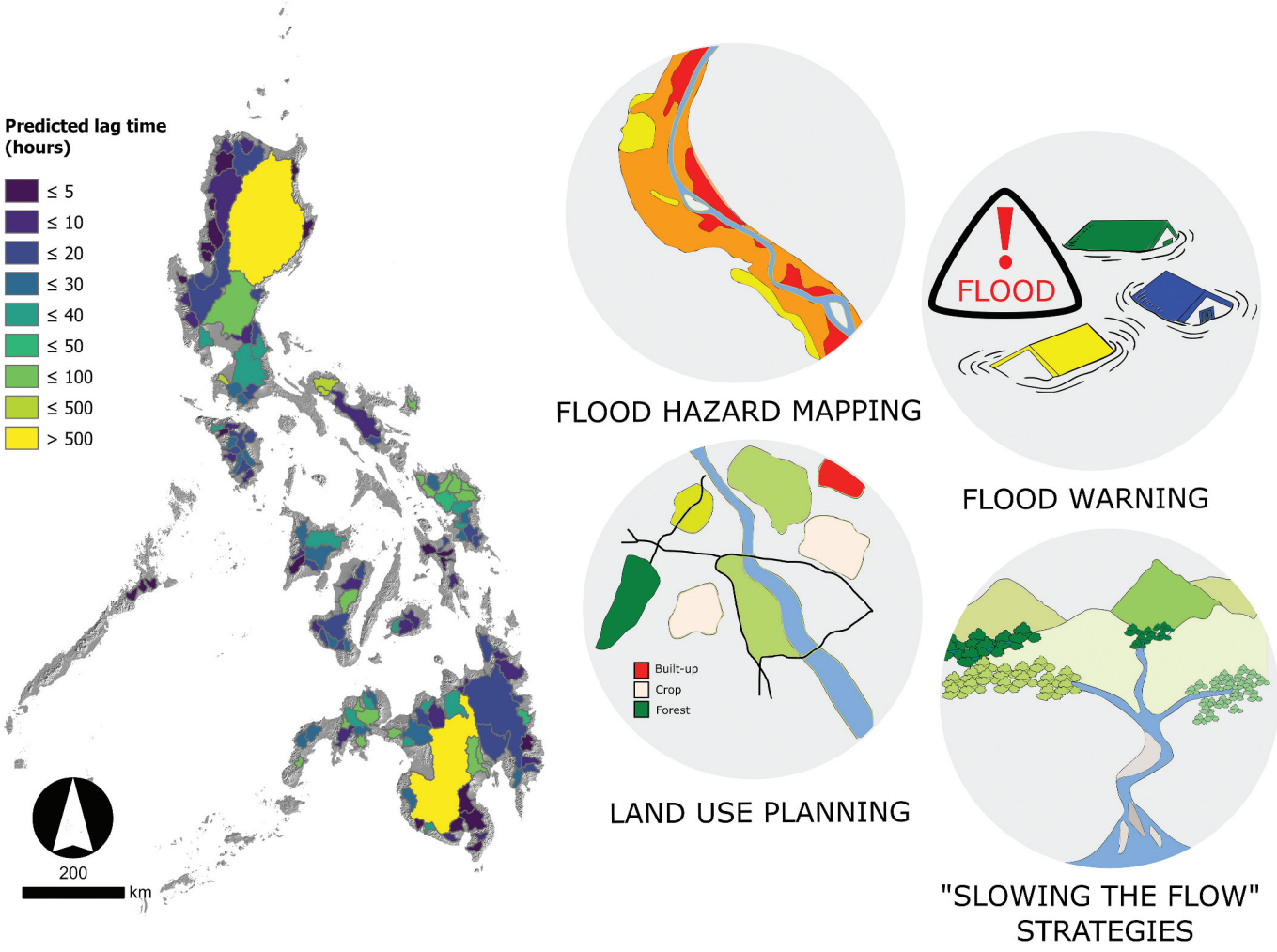


Figure 12. Predicted lag times using the GAM equation developed for Philippine catchments available in a national-scale geodatabase. Knowing the lag time can direct targeted actions such as flood hazard mapping, flood warning, land use planning, and natural flood management strategies to address flood risk.

6 Conclusion

This study developed a GAM-based predictive equation for lag time across 291 Philippine catchments, achieving strong performance (R^2 = 0.77). Relief ratio, land cover index, and catchment area were identified as the most influential predictors. Practical applications of these findings include prioritizing flood monitoring and early warning systems in catchments with short lag times, informing land use planning and reforestation strategies to lengthen lag time where feasible, and integrating lag-time predictions into flood hazard mapping at national scale. These recommendations align with nature-based solutions (NBS) for flood risk management, where reforestation, sustainable land management, and runoff attenuation features can complement engineered interventions. Importantly, the relationship with land cover is not straightforward to constrain, yet land cover is the only modelled variable that can be actively managed, highlighting the need for further work to parameterize and assess how the spatial distribution of land use influences runoff generation. Future research should investigate lag times under multiple return periods, incorporate soil and high-resolution land cover datasets, and examine climate change impacts on lag time dynamics. Strengthening national hydrological observation networks remains essential for further validation and refinement of predictive models.

Author contributions

CRediT: **Pamela Louise M. Tolentino** – Conceptualization, Data curation, Validation, Visualization, Investigation, Formal analysis, Methodology, Writing – original draft; **Martin D. Hurst** – Conceptualization, Methodology, Supervision, Writing – review & editing; **Richard D. Williams** – Conceptualization, Methodology, Supervision, Writing – review & editing; **Trevor B. Hoey** – Conceptualization, Writing – review & editing; **Richard J. Boothroyd** – Data curation, Writing – review & editing

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Natural Environment Research Council [NE/S003312]; British Council Philippines [Newton-Agham PhD Scholarship Programme]; Philippine Council for Industry, Energy, and Emerging Technology Research and Development ; and Science

Education Institute, Department of Science and Technology, Republic of the Philippines.

ORCID

Pamela Louise M. Tolentino http://orcid.org/0000-0002-1803-9734
Martin D. Hurst http://orcid.org/0000-0002-9822-076X
Richard D. Williams http://orcid.org/0000-0001-6067-1947
Trevor B. Hoey http://orcid.org/0000-0003-0734-6218
Richard J. Boothroyd http://orcid.org/0000-0001-9742-4229

Data availability

Input data and R codes are available the University of Glasgow Enlighten data repository https://researchdata.gla.ac.uk/.

References

Abdulkareem, J.H., *et al.*, 2018. Relationship between design floods and land use land cover (LULC) changes in a tropical complex catchment. *Arabian Journal of Geosciences*, 11 (14). doi:10.1007/s12517-018-3702-4

Abon, C.C., David, C.P.C., and Pellejera, N.E.B., 2011. Reconstructing the Tropical Storm Ketsana flood event in Marikina River, Philippines. *Hydrology and Earth System Sciences*, 15 (4), 1283–1289. doi:10.5194/hess-15-1283-2011

Afonso de Oliveira Serrão, E., *et al.*, 2022. Impacts of land use and land cover changes on hydrological processes and sediment yield determined using the SWAT model. *International Journal of Sediment Research*, 37 (1), 54–69. doi:10.1016/j.ijsrc.2021.04.002

Altaf, F., Meraj, G., and Romshoo, S.A., 2013. Morphometric analysis to infer hydrological behaviour of Lidder Watershed, Western Himalaya, India. *Geography Journal*, 2013, e178021. doi:10.1155/2013/178021

Anticamara, J.A. and Go, K.T.B., 2017. Impacts of super-typhoon Yolanda on Philippine reefs and communities. *Regional Environmental Change*, 17 (3), 703–713. doi:10.1007/s10113-016-1062-8

Archer, N.A.L., *et al.*, 2013. Soil characteristics and landcover relationships on soil hydraulic conductivity at a hillslope scale: a view towards local flood management. *Journal of Hydrology*, 497, 208–222. doi:10.1016/j.jhydrol.2013.05.043

Auffray, M., *et al.*, 2023. Reservoirs regulated by small dams have a similar warming effect than lakes on the summer thermal regime of streams. *Science of the Total Environment*, 869, 161445. doi:10.1016/j.scitotenv.2023.161445

Aurelio, M. and Peña, R., 2010. *Geology of the Philippines*. 2nd ed. Mines and Geosciences Bureau - Department of Environment and Natural Resources. Available from: https://scholar.google.com/scholar_lookup?hl=en&publication_year=2010&pages=132-133&author=M.+A.+Aurelio&author=R.+E.+Pe%C3%B1a&title=Geology+of+the+Philippines

Azuka, C.V. and Igué, A.M., 2020. Surface runoff as influenced by slope position and land use in the Koupendri catchment of northwest Benin: field observation and model validation. *Hydrological Sciences Journal*, 65 (6), 995–1004. doi:10.1080/02626667.2020.1729360

Barbero, G., *et al.*, 2022. 2D hydrodynamic approach supporting evaluations of hydrological response in small watersheds: implications for lag time estimation. *Journal of Hydrology*, 610, 127870. doi:10.1016/j.jhydrol.2022.127870

Bartles, M., *et al.*, 2022. *HEC-HMS users manual*. 19 January. Available from: https://www.hec.usace.army.mil/confluence/hmsdocs/hmsum/latest/report-documentation-page

Bernini, A., *et al.*, 2021. Evaluation of gully erosion susceptibility using a maximum entropy model in the Upper Mkhomazi River Basin in South Africa. *ISPRS International Journal of Geo-Information*, 10 (11), Article 11. doi:10.3390/ijgi10110729

Black, A., *et al.*, 2021. Natural flood management, lag time and catchment scale: results from an empirical nested catchment study. *Journal of Flood Risk Management*, 14 (3), e12717. doi:10.1111/jfr3.12717

Blanco, A.C., *et al.*, 2015. The Phil-Lidar 2 program: national resource inventory of the Philippines using lidar and other remotely sensed data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7-W3, 1123–1127. 36th International Symposium on Remote Sensing of Environment (Volume XL-7/W3) 11–15 May 2015, Berlin, Germany. doi:10.5194/isprsarchives-XL-7-W3-1123-2015

Bond, S., *et al.*, 2020. Seasonal vegetation and management influence overland flow velocity and roughness in upland grasslands. *Hydrological Processes*, 34 (18), 3777–3791. doi:10.1002/hyp.13842

Boorman, D.B., Hollis, J.M., and Lilly, A., 1995. *Hydrology of soil types: a hydrologically-based classification of the soils of the United Kingdom*. Inst. of Hydrology.

Boothroyd, R., *et al.*, 2023. National-scale geodatabase of catchment characteristics in the Philippines for river management applications. *PLOS ONE*, 18 (3), e0281933. doi:10.1371/journal.pone.0281933

Boothroyd, R. *et al.*, 2025. Big data show idiosyncratic patterns and rates of geomorphic river mobility. *Nature Communications*, 16 (1), 3263. doi:10.1038/s41467-025-58427-9

Bourgin, F., *et al.*, 2015. Transferring global uncertainty estimates from gauged to ungauged catchments. *Hydrology and Earth System Sciences*, 19 (5), 2535–2546. doi:10.5194/hess-19-2535-2015

Brown, S., *et al.*, 2012. Assessment of spatiotemporal varying relationships between rainfall, land cover and surface water area using geographically weighted regression. *Environmental Modeling and Assessment*, 17 (3), 241–254. doi:10.1007/s10666-011-9289-8

Brunner, M.I. and Naveau, P., 2023. Spatial variability in Alpine reservoir regulation: deriving reservoir operations from streamflow using generalized additive models. *Hydrology and Earth System Sciences*, 27 (3), 673–687. doi:10.5194/hess-27-673-2023

Bunn, S.E. and Arthington, A.H., 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, 30 (4), 492–507. doi:10.1007/s00267-002-2737-0

Bureau of Soils and Water Management, 2004. *Soil data set* [Data set]. Available from: https://www.bswm.da.gov.ph/process/sw-info-system/

Burgess-Gamble, L., *et al.*, 2018. *Working with natural processes evidence directory*. February. Available from: https://assets.publishing.service.gov.uk/media/6036c5468fa8f5480a5386e9/Working_with_natural_processes_evidence_directory.pdf

Byrne, P., *et al.*, 2024. PFAS river export analysis highlights the urgent need for catchment-scale mass loading data. *Environmental Science & Technology Letters*, 11 (3), 266–272. doi:10.1021/acs.estlett.4c00017

Cadiz, N.R., 2018. UP NOAH in building resilient Philippines; multi-hazard and risk mapping for the future. *Procedia Engineering*, 212, 1018–1025. doi:10.1016/j.proeng.2018.01.131

Canty, A. and Ripley, B., 2024. *boot: bootstrap functions*. Available from: https://cran.r-project.org/web/packages/boot/boot.pdf

Chebana, F., *et al.*, 2014. Regional frequency analysis at ungauged sites with the generalized additive model. *Journal of Hydrometeorology*, 15 (6), 2418–2428. doi:10.1175/JHM-D-14-0060.1

Clerici, N., *et al.*, 2019. Spatio-temporal and cumulative effects of land use-land cover and climate change on two ecosystem services in the Colombian Andes. *Science of the Total Environment*, 685, 1181–1192. doi:10.1016/j.scitotenv.2019.06.275

Csárdi, G., *et al.*, 2024. *igraph for R: r interface of the igraph library for graph theory and network analysis* (Version v2.0.2) [Computer software]. Zenodo. doi:10.5281/zenodo.10681749

Dasallas, L., An, H., and Lee, S., 2024. Providing solutions for data scarcity in urban flood modeling through sensitivity analysis and DEM modifications. *Journal of Hydroinformatics*, 26 (2), 459–479. doi:10.2166/hydro.2024.173

Dasgupta, R., *et al.*, 2024. Revisit hydrological modeling in ungauged catchments comparing regionalization, satellite observations, and machine learning approaches. *HydroResearch*, 7, 15–31. doi:10.1016/j.hydres.2023.11.001

Deng, L., Shangguan, Z., and Li, R., 2012. Effects of the grain-for-green program on soil erosion in China. *International Journal of Sediment Research*, 27 (1), 120–127. doi:10.1016/S1001-6279(12)60021-3

Durocher, M., Chebana, F., and Ouarda, T.B.M.J., 2015. A nonlinear approach to regional flood frequency analysis using projection pursuit regression. *Journal of Hydrometeorology*, 16 (4), 1561–1574. doi:10.1175/JHM-D-14-0227.1

Eccles, R., Zhang, H., and Hamilton, D., 2019. A review of the effects of climate change on riverine flooding in subtropical and tropical regions. *Journal of Water and Climate Change*, 10 (4), 687–707. doi:10.2166/wcc.2019.175

Environmental Management Bureau, 2006. *National water quality status report 2001–2005*. Department of Environment and Natural Resource. Available from: https://emb.gov.ph/wp-content/uploads/2019/08/NWQSR-2001-to-2005.pdf

Fasiolo, M., *et al*., 2023. *mgcViz: visualisations for generalized additive models* (Version 0.1.11) [Computer software]. Available from: https://cran.r-project.org/web/packages/mgcViz/index.html

Fenicia, F., *et al*., 2014. Catchment properties, function, and conceptual model representation: is there a correspondence? *Hydrological Processes*, 28 (4), 2451–2467. doi:10.1002/hyp.9726

Fiener, P., Auerswald, K., and Van Oost, K., 2011. Spatio-temporal patterns in land use and management affecting surface runoff response of agricultural catchments—A review. *Earth-Science Reviews*, 106 (1–2), 92–104. doi:10.1016/j.earscirev.2011.01.004

Gabriel, A.D., *et al*., 2023. Riverine microplastic pollution: insights from Cagayan de Oro River, Philippines. *International Journal of Environmental Research & Public Health*, 20 (12), Article 12. doi:10.3390/ijerph20126132

Gao, Q. and Yu, M., 2017. Reforestation-induced changes of landscape composition and configuration modulate freshwater supply and flooding risk of tropical watersheds. *PLOS ONE*, 12 (7), 1–14. doi:10.1371/journal.pone.0181315

Gao, Y., *et al*., 2022. Analyzing the critical locations in response of constructed and planned dams on the Mekong River Basin for environmental integrity. *Environmental Research Communications*, 4 (10), 101001. doi:10.1088/2515-7620/ac9459

Garzon, L.F.L., *et al*., 2023. Exploring the effects of catchment morphometry on overland flow response to extreme rainfall using a 2D hydraulic-hydrological model (IBER). *Journal of Hydrology*, 627, 130405. doi:10.1016/j.jhydrol.2023.130405

Grafil, L.B. and Castro, O.T., 2014. *Acquisition of IfSAR for the production of nationwide DEM and ORI for the Philippines under the unified mapping project*. Infomapper. Available from: https://www.namria.gov.ph/jdownloads/Info_Mapper/21_im_2014.pdf

Gravelius, H., 1914. *Morphometry of drainage basins*. Elsevier. Available from https://scholar.google.com/scholar_lookup?&title=Flusskunde.%20Goschen%20Verlagshan%20dlung%20Berlin&publication_year=1914&author=Gravelius%2CH

Guzha, A.C., *et al*., 2018. Impacts of land use and land cover change on surface runoff, discharge and low flows: evidence from East Africa. *Journal of Hydrology: Regional Studies*, 15, 49–67. doi:10.1016/j.ejrh.2017.11.005

Haga, H., *et al*., 2005. Flow paths, rainfall properties, and antecedent soil moisture controlling lags to peak discharge in a granitic unchanneled catchment. *Water Resources Research*, 41 (12). doi:10.1029/2005WR004236

Hale, V.C. and McDonnell, J.J., 2016. Effect of bedrock permeability on stream base flow mean transit time scaling relations: 1. A multiscale catchment intercomparison. *Water Resources Research*, 52 (2), 1358–1374. doi:10.1002/2014WR016124

He, M., *et al*., 2024. Streamflow prediction in ungauged catchments through use of catchment classification and deep learning. *Journal of Hydrology*, 639, 131638. doi:10.1016/j.jhydrol.2024.131638

Horton, R., 1932. Drainange basin characteristics. *American Geophysical Union*, 13, 350–361.

Horton, R., 1945. *Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology*. Bulletin of the Geological Society of America.

Hou, Y., *et al*., 2023. A global synthesis of hydrological sensitivities to deforestation and forestation. *Forest Ecology and Management*, 529, 120718. doi:10.1016/j.foreco.2022.120718

Huffman, G.J., *et al*., 2019. *GPM IMERG final precipitation L3 half hourly 0.1 degree x 0.1 degree V06* [Data set]. Greenbelt, MD: Goddard Earth Sciences Data and Information Services Center (GES DISC). doi:10.5067/GPM/IMERG/3B-HH/06

Juanico, M., 2025. National greening program: boosting forest cover for flood prevention. *INQUIRER.Net*, 28 August. Available from: https://opinion.inquirer.net/185659/national-greening-program-boosting-forest-cover-for-flood-prevention

Jung, K., *et al*., 2024. Assessment of non-linear models based on regional frequency analysis for estimation of hydrological quantiles at ungauged sites in South Korea. *Journal of Hydrology: Regional Studies*, 52, 101713. doi:10.1016/j.ejrh.2024.101713

Kay, A.L., *et al*., 2019. An assessment of the potential for natural flood management to offset climate change impacts. *Environmental Research Letters*, 14 (4), 044017. doi:10.1088/1748-9326/aafdbe

Kuhn, M., 2008. Building predictive models in R using the caret package. *Journal of Statistical Software*, 28 (5), 1–26. doi:10.18637/jss.v028.i05

Lagmay, A.M.F., *et al*., 2017. Disseminating near-real-time hazards information and flood maps in the Philippines through Web-GIS. *Journal of Environmental Sciences*, 59, 13–23. doi:10.1016/j.jes.2017.03.014

Lane, S.N., 2017. Natural flood management. *WIREs Water*, 4 (3), e1211. doi:10.1002/wat2.1211

Larsen, K., 2015. *GAM: the predictive modeling silver bullet | stitch fix technology – multithreaded*. 30 July. Available from: https://multithreaded.stitchfix.com/blog/2015/07/30/gam/

Macalalad, R.V., *et al*., 2021. Hydrological response of the Pampanga River Basin in the Philippines to intense tropical cyclone rainfall. *Journal of Hydrometeorology*, 22 (4), 781–794. doi:10.1175/JHM-D-20-0184.1

Mahala, A., 2019. The significance of morphometric analysis to understand the hydrological and morphological characteristics in two different morpho-climatic settings. *Applied Water Science*, 10 (1), 33. doi:10.1007/s13201-019-1118-2

Melone, F., Corradini, C., and Singh, V.P., 2002. Lag prediction in ungauged basins: an investigation through actual data of the upper Tiber River valley. *Hydrological Processes*, 16 (5), 1085–1094. doi:10.1002/hyp.313

Merz, B. and Plate, E.J., 1997. An analysis of the effects of spatial variability of soil and soil moisture on runoff. *Water Resources Research*, 33 (12), 2909–2922. doi:10.1029/97WR02204

Metzger, R.C., 2004. 7—debugging strategies. *In*: R.C. Metzger, ed. *Debugging by thinking*. Amsterdam, The Netherlands: Digital Press, 189–199. doi:10.1016/B978-155558307-1/50007-0

Mines and Geosciences Bureau, 2021. *MGB public portal*. Available from: http://databaseportal.mgb.gov.ph/#/public/geological-maps

Muñoz Sabater, J., 2019. *ERA5-land monthly averaged data from 1950 to present* [Data set]. Bonn, Germany: Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi:10.24381/cds.68d2bb30

National Mapping and Resource Information Authority, 2022. *2010, 2015, 2020 landcover_Philippines_PRS92* [Data set]. Taguig City, Philippines.

National Resources Conservation Service, 2008. *Chapter 15: time of concentration*. Washington, DC: United States Department of Agriculture.

Necesito, I.V., *et al*., 2023. Deep learning-based univariate prediction of daily rainfall: application to a flood-prone, data-deficient country. *Atmosphere*, 14 (4), Article 4. doi:10.3390/atmos14040632

Nicholson, A.R., *et al*., 2012. Runoff attenuation features: a sustainable flood mitigation strategy in the Belford catchment, UK. *Area*, 44 (4), 463–469. doi:10.1111/j.1475-4762.2012.01099.x

Obeidat, M., Awawdeh, M., and Al-Hantouli, F., 2021. Morphometric analysis and prioritisation of watersheds for flood risk management in Wadi Easal Basin (WEB), Jordan, using geospatial technologies. *Journal of Flood Risk Management*, 14 (2), e12711. doi:10.1111/jfr3.12711

PAGASA, 2024. *Climate of the Philippines*. Available from: https://www.pagasa.dost.gov.ph/information/climate-philippines

Pamungkas, A., Bekessy, S., and Lane, R., 2017. Adaptations assessment on the impacts of flooding under current condition and climate change

scenario, case study: Centini Village, Indonesia. *Tataloka*, 19 (3), 163. doi:10.14710/tataloka.19.3.163-174

Paringit, E. and Pascua, C., 2017. *LiDAR surveys and flood mapping of Bacarra river*. Quezon City, the Philippines: University of the Philippines - Diliman.

Paringit, E. and Sinogaya, J.R., Eds, 2017. *LiDAR surveys and flood mapping of Imbang river*. Quezon City: University of the Philippines Training Center on Applied Geodesy and Photogrammetry.

Pattison, I., *et al*., 2014. The role of tributary relative timing and sequencing in controlling large floods. *Water Resources Research*, 50 (7), 5444–5458. doi:10.1002/2013WR014067

Pearson, C.J., *et al*., 2022. Identification of floodwater source areas in Nepal using SCIMAP-flood. *Journal of Flood Risk Management*, 15 (4), e12840. doi:10.1111/jfr3.12840

Peskett, L.M., *et al*., 2021. Tracers reveal limited influence of plantation forests on surface runoff in a UK natural flood management catchment. *Journal of Hydrology: Regional Studies*, 36 (December 2020), 100834. doi:10.1016/j.ejrh.2021.100834

Peskett, L.M., *et al*., 2023. Land cover influence on catchment scale subsurface water storage investigated by multiple methods: implications for UK natural flood management. *Journal of Hydrology: Regional Studies*, 47, 101398. doi:10.1016/j.ejrh.2023.101398

Quick, L., *et al*., 2025. Author correction: confined and mined: anthropogenic river modification as a driver of flood risk change. *NPJ Natural Hazards*, 2 (1), 1. doi:10.1038/s44304-025-00066-7

Quinn, P.F., *et al*., 2022. The role of runoff attenuation features (RAFs) in natural flood management. *Water*, 14 (23), Article 23. doi:10.3390/w14233807

Raja Shekar, P. and Mathew, A., 2024. Morphometric analysis of watersheds: a comprehensive review of data sources, quality, and geospatial techniques. *Watershed Ecology and the Environment*, 6, 13–25. doi:10.1016/j.wsee.2023.12.001

Rasheed, Z., *et al*., 2022. Advancing flood warning procedures in ungauged basins with machine learning. *Journal of Hydrology*, 609, 127736. doi:10.1016/j.jhydrol.2022.127736

Rentschler, J., Salhab, M., and Jafino, B.A., 2022. Flood exposure and poverty in 188 countries. *Nature Communications*, 13 (1), Article 1. doi:10.1038/s41467-022-30727-4

Rola, A., *et al*., 2015. Challenges of water governance in the Philippines. *Philippine Journal of Science*, 144 (2), 197–208.

Sakan, S., *et al*., 2021. Evaluation of element mobility in river sediment using different single extraction procedures and assessment of probabilistic ecological risk. *Water*, 13 (10), Article 10. doi:10.3390/w13101411

Santos, R.N., 2018. *National mapping efforts: the Philippines*. International Regional Science Meeting, 30 May. Available from: https://lcluc.umd.edu/sites/default/files/Rizaldia.pdf

Sapač, K., *et al*., 2020. Lag times as indicators of hydrological mechanisms responsible for NO3-N flushing in a forested headwater catchment. *Water*, 12 (4), Article 4. doi:10.3390/w12041092

Sawicz, K., *et al*., 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrology and Earth System Sciences*, 15 (9), 2895–2911. doi:10.5194/hess-15-2895-2011

Sayama, T., *et al*., 2011. How much water can a watershed store? *Hydrological Processes*, 25 (25), 3899–3908. doi:10.1002/hyp.8288

Schumm, S., 1956. Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. *GSA Bulletin*, 67 (5), 597–646. doi:10.1130/0016-7606(1956)67%255B597:EODSAS%255D2.0.CO;2

Schwanghart, W. and Scherler, D., 2014. Short communication: TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences. *Earth Surface Dynamics*, 2 (1), 1–7. doi:10.5194/esurf-2-1-2014

Seneviratne, S., *et al*., 2012. Changes in climate extremes and their impacts on the natural physical environment. *In*: S.K.A. Field, *et al*., eds. *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Vol. 9781107025. Cambridge, UK: Cambridge University Press, 109–230. doi:10.1017/CBO9781139177245.006

Seyam, M. and Othman, F., 2014. The influence of accurate lag time estimation on the performance of stream flow data-driven based models. *Water Resources Management*, 28 (9), 2583–2597. doi:10.1007/s11269-014-0628-9

Shao, M., *et al*., 2020. Quantifying the effects of urbanization on floods in a changing environment to promote water security—a case study of two adjacent basins in Texas. *Journal of Hydrology*, 589, 125154. doi:10.1016/j.jhydrol.2020.125154

Sharma, A.P., Fu, X., and Kattel, G.R., 2023. Is there a progressive flood risk management in Nepal? A synthesis based on the perspective of a half-century (1971–2020) flood outlook. *Natural Hazards*, 118 (2), 903–923. doi:10.1007/s11069-023-06035-5

Shuttleworth, E.L., *et al*., 2019. Restoration of blanket peat moorland delays stormflow from hillslopes and reduces peak discharge. *Journal of Hydrology X*, 2, 100006. doi:10.1016/j.hydroa.2018.100006

Skoufias, E., *et al*., 2020. Identifying the vulnerable to poverty from natural disasters: the case of typhoons in the Philippines. *Economics of Disasters and Climate Change*, 4 (1), 45–82. doi:10.1007/s41885-020-00059-y

Sultan, D., *et al*., 2022. Evaluation of lag time and time of concentration estimation methods in small tropical watersheds in Ethiopia. *Journal of Hydrology: Regional Studies*, 40, 101025. doi:10.1016/j.ejrh.2022.101025

Syvitski, J.P.M., *et al*., 2014. How important and different are tropical rivers? — an overview. *Geomorphology*, 227, 5–17. doi:10.1016/j.geomorph.2014.02.029

Tolentino, P.L.M., *et al*., 2016. Projected impact of climate change on hydrological regimes in the Philippines. *PLOS ONE*, 11 (10), e0163941. doi:10.1371/journal.pone.0163941

Tolentino, P.L.M., Williams, R.D., and Hurst, M.D., 2024. Cascading consequences of structural interventions in a tropical wandering gravel-bed river. *River Research and Applications*, n/a (n/a). doi:10.1002/rra.4362

Tolentino, P.L.M., Williams, R.D., and Hurst, M.D., 2025. Natural flood risk management in tropical Southeast Asia: prospects in the biodiverse archipelagic nation of the Philippines. *WIREs Water*, 12 (1), e70000. doi:10.1002/wat2.70000

UK Centre for Ecology and Hydrology, 2024. *Flood estimation handbook (FEH) | UK centre for ecology & hydrology*. Available from: https://www.ceh.ac.uk/our-science/projects/flood-estimation-handbook

UP DREAM Program, 2024. *UP PHIL-LiDAR 1 program technical reports » UP DREAM program*. Available from: https://dream.upd.edu.ph/up-phil-lidar-1-program-technical-reports/

UP NOAH, 2024. *NOAH - nationwide operational assessment of hazards*. Available from: https://noah.up.edu.ph/

van Emmerik, T., *et al*., 2022. Rivers as plastic reservoirs. *Frontiers in Water*, 3. doi:10.3389/frwa.2021.786936

Vannier, O., Anquetin, S., and Braud, I., 2016. Investigating the role of geology in the hydrological response of Mediterranean catchments prone to flash-floods: regional modelling study and process understanding. *Journal of Hydrology*, 541, 158–172. doi:10.1016/j.jhydrol.2016.04.001

Vereecken, H., *et al*., 2022. Soil hydrology in the Earth system. *Nature Reviews Earth & Environment*, 3 (9), 573–587. doi:10.1038/s43017-022-00324-6

Vogel, H.-J., 2019. Scale issues in soil hydrology. *Vadose Zone Journal*, 18 (1), 190001. doi:10.2136/vzj2019.01.0001

von Brömssen, C., *et al*., 2023. Modeling complex concentration-discharge relationships with generalized additive models. *Environmental Modeling and Assessment*, 28 (6), 925–937. doi:10.1007/s10666-023-09915-z

Waikhom, S.I., *et al*., 2023. Impact assessment of land use/land cover changes on surface runoff characteristics in the Shetrunji River Basin using the SWAT model. *Water Practice & Technology*, 18 (5), 1221–1232. doi:10.2166/wpt.2023.071

Wei, T. and Simko, V., 2021. *R package 'corrplot': visualization of a correlation matrix* (Version 0.92) [R]. Available from: https://github.com/taiyun/corrplot

Wood, S.N., 2017. *Generalized additive models: an introduction with R*. 2nd ed. New York: CRC Press/Taylor & Francis Group.

World Risk Report, 2022. *WorldRiskReport 2022—focus: digitalization—world | ReliefWeb*. Available from: https://reliefweb.int/report/world/worldriskreport-2022-focus-digitalization

World Risk Report, 2023. *The WorldRiskReport 2023 – disaster risk and diversity—world | ReliefWeb*. Available from: https://reliefweb.int/report/world/worldriskreport-2023-disaster-risk-and-diversity

World Risk Report, 2024. *World risk report 2024: focus: multiple crises*. Available from: https://weltrisikobericht.de/worldriskreport/

Xu, P., *et al*., 2024. A dynamic von Mises-based model to evaluate the impact of urbanization and climate change on flood timing in Yangtze and Huaihe River Basins, China. *Journal of Hydrology*, 634, 131120. doi:10.1016/j.jhydrol.2024.131120

Younis, S.M.Z. and Ammar, A., 2018. Quantification of impact of changes in land use-land cover on hydrology in the upper Indus Basin, Pakistan. *The Egyptian Journal of Remote Sensing and Space Science*, 21 (3), 255–263. doi:10.1016/j.ejrs.2017.11.001

Yuan, L. and Forshay, K.J., 2021. Enhanced streamflow prediction with SWAT using support vector regression for spatial calibration: a case study in the Illinois River watershed, U.S. *PLOS ONE*, 16 (4), e0248489. doi:10.1371/journal.pone.0248489

Zhang, J., *et al*., 2022. Transport and partitioning of metals in river networks of a plain area with sedimentary resuspension and implications for downstream lakes. *Environmental Pollution*, 294, 118668. doi:10.1016/j.envpol.2021.118668