Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.086499

ARTICLE

AI-Powered Hydrological Modelling: Wavelet-Based Pattern Analysis and Short-Term Water Level Forecasting in a Data-Scarce River Catchment

Jack Ward[1], **Randika K. Makumbura**[1], **Hirushan Sajindra**[1], **Gul Jokhio**[1], **Bojan Đurin**[2] and **Upaka Rathnayake**[1,*]

[1]Faculty of Engineering and Design, Atlantic Technological University, Sligo, Ireland
[2]Department of Civil Engineering, University North, Varaždin, Croatia

*Corresponding Author: Upaka Rathnayake. Email: upaka.rathnayake@atu.ie

Received: 01 June 2026; Accepted: 16 September 2026; Published: 21 October 2026

ABSTRACT: Water resource management and flood defence infrastructure rely heavily on an understanding of hydrological variability and the ability to predict future water levels. Intense storm events are becoming increasingly frequent under a changing climate, making short-range water level forecasting essential for small, data-scarce catchments where traditional trend analysis methods are limited. This study investigates the use of Continuous Wavelet Transform analysis for water level pattern identification and Long Short-Term Memory neural network forecasting for short-range water level prediction in the Drumcliff River, County Sligo, Ireland, a small Atlantic-facing catchment with limited meteorological data availability. Four years of quality-controlled water level data spanning November 2021 to November 2025, comprising daily mean and 15-min interval records, were analysed. The study revealed a persistent short-period oscillatory band of 8–32 days driven by episodic storm events during autumn and winter, and a complementary long-period band of 64–256 days reflecting seasonal variability, with no significant long-term upward or downward trend observed across the study period. An extreme hydrological episode was identified in autumn 2025 with a wavelet power peak of approximately 2.2, substantially exceeding background variability. The one-day Long Short-Term Memory Algorithm (LSTM) achieved R^2 and NSE values of 0.982–0.988 with Root Mean Squared Error (RMSE) of 0.013–0.017 m, while the three-day model performed reliably at 24- and 48-h horizons but degraded substantially at 72 h, highlighting the limitation of single-input models in the absence of meteorological forcing data. Therefore, this study presents, to our knowledge, one of the first applications of the continuous wavelet transform for diagnostic multi-scale decomposition of a river water level record in a small, data-scarce Irish Atlantic catchment, providing a methodological foundation for advanced hydrological analysis and operationally deployable short-term flood early warning systems in data-scarce catchments.

KEYWORDS: Wavelet analysis; long short-term memory; water level forecasting; hydrological time series; flood early warning

1 Introduction

Rivers play a crucial role in supporting ecosystems and sustaining critical human activities such as agriculture. However, they are being increasingly influenced by climate change as a result of several different anthropogenic activities. Variations in precipitation patterns and storm frequency are evident worldwide and are particularly evident in the west of Ireland as a result of its location alongside the Atlantic Ocean. Hydrological systems are highly responsive to climatic factors and, as a result of climate change, their

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

behaviour is changing, highlighting the importance of being able to detect emerging trends and identify the drivers of these trends.

Climate change is having a drastic effect on the hydrological cycle. Changes in rainfall intensity, frequency and seasonality have been widely observed, resulting in falling groundwater levels [1], heightened evaporation and decreased precipitation [2], and increased flood and drought events [3]. In Ireland, identified an amplification of the seasonal cycle driven by increased winter precipitation and decreased summer precipitation across nine Irish catchments [4]. Bastola et al., (2011) [5] projected progressive increases in winter discharge and decreases in summer discharge over the coming century, while Sarkar Basu et al., (2022) [6] projected increases in flood quantile of up to 16% for the 100-year return period in the Dodder River Basin, Dublin. These changes produce increased hydrological variability, making river trends harder to determine, flows harder to predict, and their overall behaviour non-stationary.

The majority of statistical analyses of hydrological time series are based on assumptions that the series is homogeneous, stationary, free from trends and shifts, and non-periodic [7]. Traditional methods such as the Mann-Kendall test and linear regression largely depend on these assumptions. However, Milly et al., (2008) [8] state that stationarity is dead due to anthropogenic changes to the Earth's climate. With an ever-changing climate, the reliability of traditional stationary-based hydrological models is questionable, highlighting the increased need for analytical methods that can detect evolving patterns.

Wavelet analysis is a powerful tool for examining time series that showcase variations across multiple temporal scales, particularly when signals are nonstationary, nonlinear, or contain features that traditional methods struggle to capture. Unlike the Fourier transform, wavelet methods allow for the simultaneous localisation of information in both time and frequency, making wavelets especially suited to real-world geophysical and hydrological data, where amplitudes, frequencies, and phases often evolve with time [9]. Schulte (2016) [10] highlights that wavelet methods can extract transient trends that Fourier analysis cannot detect, while Sang (2013) [11] demonstrated that wavelet de-noising techniques are more efficient than traditional filtering methods as they simultaneously account for both temporal and frequency characteristics of hydrological signals. Applied studies have further demonstrated the utility of wavelet analysis across diverse hydrological contexts, consistently revealing patterns and dynamics undetectable through traditional methods. Continuous wavelet transforms have been used to detect temporal patterns in water quality data in mixed land use watersheds, to separate overlapping tide, wave and rainfall-runoff components from tidal river water level records, to uncover groundwater and surface water coupling from as little as one year of data, and to illustrate long-term hydrological trends in multi-decadal streamflow records that conventional statistical techniques were unable to identify [12–15]. Beyond pattern identification, wavelet analysis has also demonstrated particular value in forecasting applications, where decomposition frameworks that separate deterministic signals from noise prior to model training have been shown to substantially improve hydrological forecasting accuracy [16].

In Ireland, hydrological trend studies have primarily employed traditional methods. Murphy et al., (2013) [17] applied the Mann-Kendall and Theil-Sen tests to 43 stations across Ireland and found a lack of persistent significant trends in annual and seasonal mean flows due to high temporal variability. O'Connor et al., (2023) [18] extended trend detection to 1900 across 51 catchments and identified significant increases in Atlantic seaboard catchments in winter months. Few studies have applied wavelet analysis in an Irish hydrological context, and those that have focused on precipitation or on cross-variable relationships rather than on the diagnostic decomposition of river water levels. Moradian et al., (2023) [19] combined the discrete wavelet transform and machine learning to improve precipitation modelling in Cork, while wavelet-coherence approaches have been used to examine relationships between rainfall, stream and spring discharge in Irish catchments. There is therefore a clear lack of research in Ireland utilising wavelet analysis to

decompose river water level data, with the existing trend-based studies able to show only whether a trend exists and no in-depth work undertaken to decompose the variability and identify its underlying drivers.

The Drumcliff River in County Sligo provides an appropriate case study to address this gap. The catchment is situated in the west of Ireland along the Atlantic coast and is characterised by high rainfall variability, heavily influenced by Atlantic weather systems, and is primarily made up of agricultural and bogland. All these factors combine to produce a river that is highly responsive to large hydrological events, leaving it potentially vulnerable if these events intensify in the future. This study, therefore, applies the continuous wavelet transform to perform, to our knowledge, one of the first diagnostic multi-scale decompositions of a river water level record in a small, data-scarce catchment on Ireland's Atlantic seaboard, using four years of records from the Drumcliff River. The continuous wavelet transform and the LSTM models are applied here as complementary but independent analyses: the former to characterise the multi-scale structure of the catchment's variability, the latter to evaluate short-term forecasting feasibility. The specific contributions of this study are therefore threefold: (i) a diagnostic multi-scale decomposition of a river water level record in a small, data-scarce Irish Atlantic catchment; (ii) the detection and temporal localisation of a short-duration extreme variability episode; and (iii) an assessment of the feasibility and horizon limits of single-input LSTM forecasting in the absence of meteorological data. To this end, two Long Short-Term Memory (LSTM) machine learning models are developed to forecast one-day and three-day future water levels and to assess the feasibility of single-input forecasting in a data-scarce catchment. The primary objectives are, accordingly, to apply the continuous wavelet transform to extract transient, non-stationary patterns and quantify their statistical significance; to develop two LSTM forecasting models for one-day and three-day water level prediction; and to assess the usefulness of wavelet analysis in identifying patterns within a hydrological time series.

The remainder of this paper is organised as follows. Section 2 describes the study area, data collection and preprocessing, the continuous wavelet transform methodology, the LSTM model development, and the performance evaluation metrics. Section 3 presents the results, comprising the time series analysis, the complete-record and seasonal wavelet analyses, and the three-day and one-day LSTM forecasting models. Section 4 discusses the hydrological interpretation of the wavelet patterns, the detected extreme event, the forecasting performance and the study limitations. Section 5 concludes the paper and outlines directions for future work.

2 Materials and Methodology

The methodological framework adopted in this study is structured around three principal analytical components, as illustrated in Fig. 1. Following data screening and eligibility checks, the study proceeds along three parallel analytical pathways, including spatial analysis conducted in QGIS (version 3.34.7), encompassing catchment mapping and land use classification, time-frequency analysis using the Continuous Wavelet Transform, encompassing seasonal variability assessment and extreme event detection; and forecast modelling using LSTM, encompassing model training and performance evaluation. This structured approach ensures a comprehensive and transparent assessment of both the multi-scale hydrological behaviour of the Drumcliff River and the short-range water level forecasting capabilities in a data-scarce catchment environment.

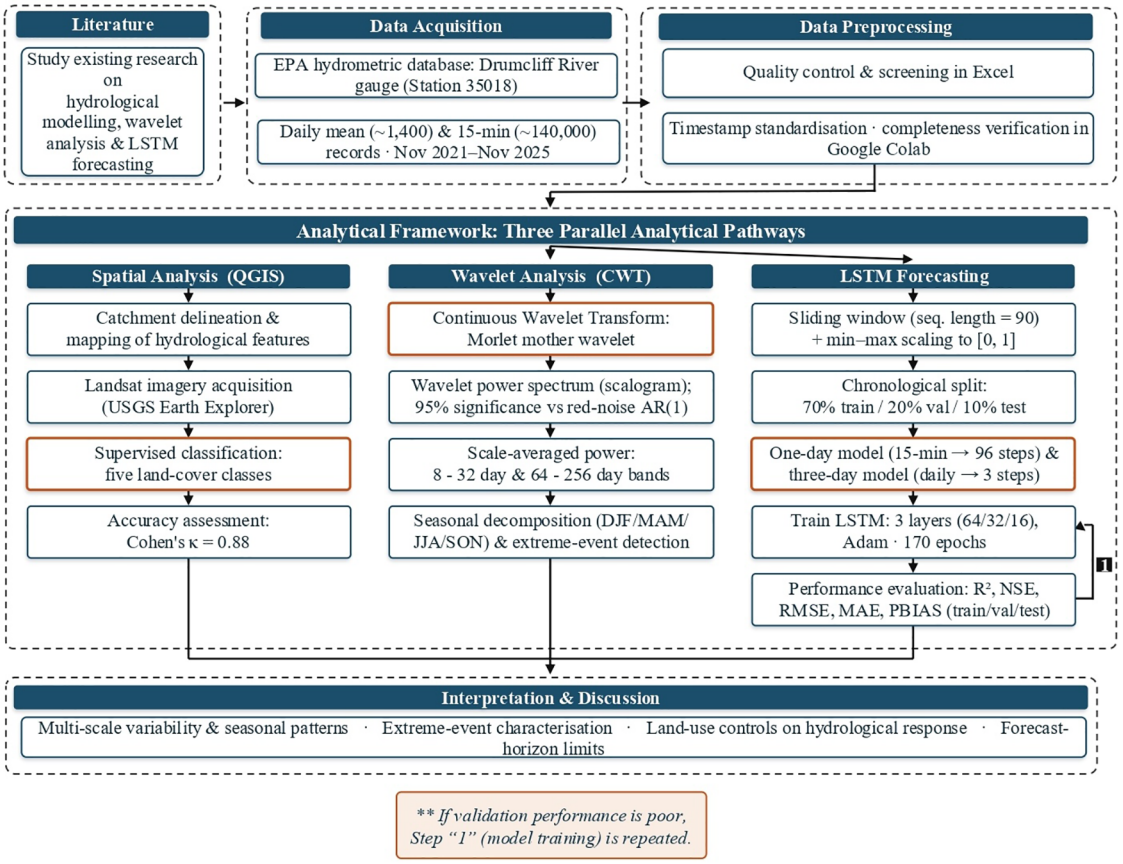
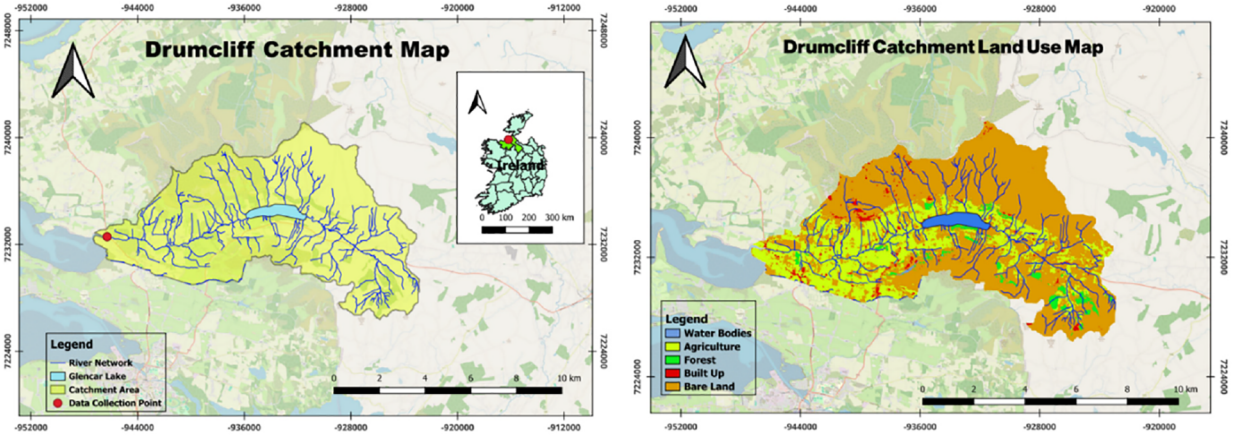


Figure 1: Methodological framework illustrating the three principal analytical pathways adopted in this study.

2.1 Study Area

The Drumcliff River is located in County Sligo in the west of Ireland. County Sligo is situated along the Atlantic coast and, due to its location, the catchment area is characterised by high rainfall variability, heavily influenced by Atlantic weather systems. The catchment is primarily made up of agricultural and bogland. The catchment covers approximately 61.29 km^2 and displays pronounced relief, ranging from near sea level at the Drumcliff estuary to around 526 m at Benbulben in the Dartry Mountains, which form its northern and eastern boundary. The river drains Glencar Lough before flowing west to Drumcliff Bay. The uplands are underlain largely by karstified Carboniferous limestone, while the lower catchment is dominated by blanket peat and poorly draining agricultural soils) [20]. The region experiences high annual rainfall driven by Atlantic frontal systems and orographic enhancement over the Dartry Mountains, with long-term averages exceeding 1200 mm at lowland stations and considerably higher totals over the uplands [21]. This combination of steep relief, high rainfall, saturated bogland and low-permeability soils produces a catchment that responds rapidly to rainfall events. The catchment map, presented in Fig. 2a, illustrates the catchment boundary and relevant hydrological features within the boundary. The land use map of the Drumcliff catchment area, presented in Fig. 2b, was constructed to quantify the dominant land cover categories and to assist with the interpretation of hydrological behaviour in the catchment area. The map was created using freely available Landsat imagery obtained from the USGS Earth Explorer platform. Classification was performed using five different land categories, including water bodies, agriculture, forest, built-up and bare land. Training samples for each class were selected (>30 for each class) through visual interpretation of

high-resolution basemap imagery, with sample sites distributed across the catchment to capture the spectral variability of each land cover type. A supervised classification was then applied to the Landsat imagery, and classification accuracy was validated against independent reference points identified from the basemaps. This assessment produced a Cohen's kappa coefficient of 0.88, indicating almost perfect agreement and confirming the reliability of the land cover map. The map indicates that the catchment area is predominantly made up of bare bog and agricultural land with scattered patches of forested areas and minimal built-up development.



(**a**) Drumcliff catchment (**b**) Land use map

Figure 2: Study area.

2.2 Data Collection

Water level data from 20 November 2021 to 30 November 2025 were extracted for the Drumcliff River and obtained through the EPA website. The model was developed using the data period from 20 November 2021 to 20 November 2025. This time frame was selected as the water level readings for this location start from September 2021, and, at the time of this study, the data points from December 2025 onwards are still unchecked on the EPA website, meaning that the data readings after this date are raw readings and have not yet been quality checked. Information regarding the Drumcliff River data collection station is presented in Table 1.

Table 1: Drumcliff river data collection station information.

Parameter	Information
Station number	35,018
Station name	DRUMCLIFF
Station status	Active
Waterbody	DRUMCLIFF
Owner	Sligo County Council
Easting	167,509
Northing	342,269
Data Provider	Environmental Protection Agency
River Basin District	Western RBD
Catchment size	61.29 km^2
Gauge datum	98.33 m

2.3 Data Preprocessing

The entire historical water level river dataset for the Drumcliff River was extracted and imported into Excel for initial cleaning and organisation. Two separate datasets were downloaded. Firstly, the daily mean water level data was extracted for the four-year timeframe. This dataset contained over 1400 water level readings. Secondly, the 15-min data readings were extracted for the four-year timeframe. This dataset contained over 140,000 water level readings. Only quality-controlled records from the EPA hydrometric database were used; unchecked raw readings from December 2025 onward were excluded, as these had not yet undergone the EPA's validation process. Each series was then screened for gauge-related inconsistencies and abnormal values by inspecting the readings against the plausible operating range of the station and by flagging any physically implausible spikes, flatlined sequences or negative values that would indicate sensor error rather than genuine hydrological variation. The datasets were also plotted and visually inspected, confirming the absence of outliers, null values or other anomalies; consequently, no records required removal or interpolation. Timestamps were standardised to a uniform interval using Excel so that the series was suitable for both wavelet analysis and machine learning forecasting. The datasets were then uploaded separately to Google Colab and checked for completeness, ensuring that all expected measurements were present with no missing or duplicated records.

2.4 Wavelet Analysis

Hydrological time series, such as rivers, often contain non-stationary signals that are undetectable using traditional trend analysis methods. These methods focus on a fixed analysis window and are limited in their use for hydrological time series. Wavelets overcome these issues by performing multi-scale decomposition and a localised time-frequency window, revealing not only if a trend exists, but exactly when it occurs. The Continuous Wavelet Transform (CWT) was applied to four years of daily mean water level time series readings from the Drumcliff River to identify dominant periodicities and their evolution over time. CWT differs from other spectral approaches as it allows for the simultaneous analysis of time and frequency, allowing transient behaviour to be identified within the time series. This is suitable for the river water level dataset as rivers have many different variables occurring at different temporal scales, such as short-term fluctuations like storm surges and longer seasonal or inter-annual events.

The Morlet wavelet was selected as the mother wavelet due to its versatility between time and frequency localisation, making it very applicable in this study. The analysis was conducted in line with the methodological framework set out by Torrence & Compo (1998) [9], in which a standardised procedure for wavelet analysis using a CWT in geophysical time series is outlined. This method allows statistically significant periodicities to be identified within the time series relative to background noise, improving the reliability of detected transient patterns. The resulting wavelet power spectrum is represented in a scalogram, showcasing the variation distribution in the time series over both time and frequency. Within the scalogram, regions of strong wavelet power emphasise times when specific frequencies prevail in the signal. Analysing the spatial patterns and consistency of these high-energy areas allows for the identification of transient events, periodic cycles, and shifts in variability over time. Smaller scales correspond to higher-frequency processes capturing short-term fluctuations, while larger scales represent lower-frequency processes associated with longer-term variability [22].

2.5 LSTM Model Development

An LSTM neural network was used to forecast water levels in the Drumcliff River. LSTM networks are a specialised form of recurrent neural network (RNN) designed to capture long-term dependencies within sequential data, making them particularly suitable for hydrological time series where temporal dependencies

and non-stationary behaviours are present. Unlike traditional statistical models, LSTMs incorporate memory cells and gating mechanisms that allow the neural network to retain or discard information over time, improving forecasting performance for complex time-dependent processes [23].

The LSTM models were developed using Python programming within Google Colab, which provides computational resources and machine learning libraries suitable for machine learning model development [24]. River water level data spanning a four-year period were used as the input time series in the models. Two separate LSTM models were developed to assess forecasting performance at different temporal resolutions (Fig. 3). The first model utilised daily mean river water level data and was trained to forecast three days ahead. The second model used higher-resolution 15-min interval water level data to produce 96 forecasted values corresponding to a one-day forecast. This dual approach allowed the evaluation of the LSTM performance across different temporal resolutions.

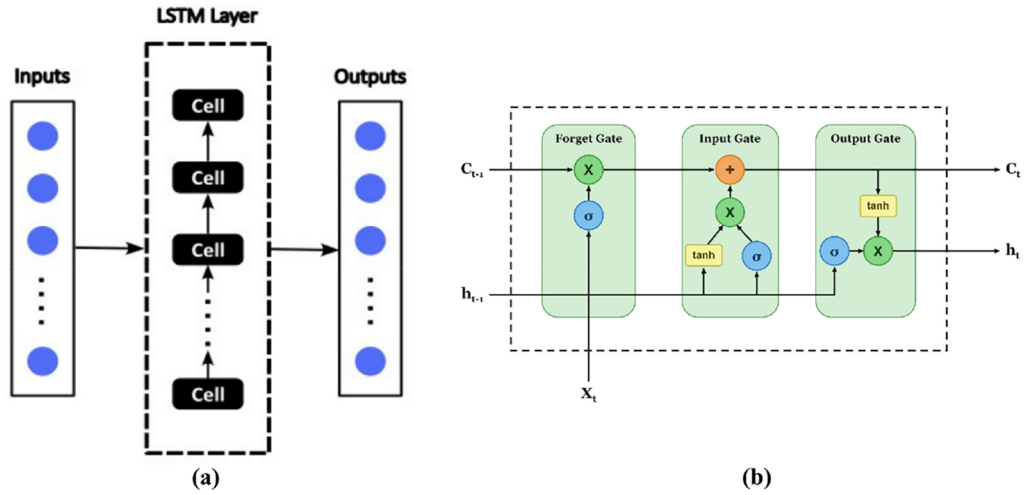


Figure 3: (**a**) Model architecture of the LSTM-based model (with one hidden layer), (**b**) LSTM memory cell, with key components [25].

Input and output sequences for both models were generated using a sliding-window approach based on the input sequence length of 90 given in Table 2. For each sample, a window of 90 consecutive past water-level observations was used as the model input, and the following value(s) as the forecast target, with the window advanced one step at a time across the record. For the three-day model, the 90-step window of daily mean water levels was mapped to the subsequent three daily values, whereas for the one-day model the 90-step window of 15-min observations was mapped to the following 96 values, corresponding to a 24-h forecast at 15-min resolution. Prior to training, the water-level series was normalised to the range $[0, 1]$ using min–max scaling, with the scaler fit on the training partition only and then applied to the validation and test partitions so that no information from the validation or test data influenced the scaling. All forecasts were inverse-transformed back to metres before the performance metrics were computed.

Prior to model training, the dataset was partitioned chronologically into 70% training, 20% validation, and 10% testing, preserving the temporal order of the series so that each model was evaluated only on data occurring after its training period. This chronological partitioning avoids the temporal information leakage that would result from randomly shuffling a time-series dataset. Several combinations of epoch numbers and network configurations were tested during model development; a configuration of 170 training epochs and the 70/20/10 split was selected on the basis of validation-set performance, with the test partition held out and used only for final evaluation. After the model was trained, forecasts were generated and compared with

the observed water level values to evaluate the effectiveness of the LSTM model's predictive capabilities. The model configuration is summarised in Table 2.

Table 2: LSTM model configuration used in the study.

Parameter	Value
Number of LSTM layers	3
Number of hidden units per layer	64, 32, 16
Dropout rate	0.2
Activation function	ReLU
Optimiser	Adam
Learning rate	0.001
Loss function	Mean Squared Error (MSE)
Batch size	32
Input sequence length	90

2.6 Model Performance Evaluation

The performance of both LSTM forecasting models was evaluated using a suite of statistical metrics. The coefficient of determination (R^2) and Nash-Sutcliffe Efficiency (NSE) were used to quantify the proportion of observed variance explained by the model. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were used to quantify the magnitude of prediction error. The Percentage Bias (PBIAS) was used to assess whether the model systematically over-predicted or under-predicted the observed water level values. These metrics were computed across all three datasets, including training, validation and testing, to ensure robust model evaluation and to confirm that the model generalises consistently beyond the training data.

3 Results

3.1 Time Series Analysis

The time series plot of the mean daily water levels in the Drumcliff River from 20 November 2021 to 20 November 2025 is illustrated in Fig. 4. Water levels in the river ranged from approximately 98.5 to 99.3 m over the four-year duration. Higher water levels frequently occur in the winter months, while lower water levels are more common during the summer months. Variability in the time series is consistent throughout the entire water level record, with rapid rises and falls. A similar seasonal structure of highs and lows is also present, characterised by seasonal peaks annually around the late autumn and early winter months. No increasing or decreasing trend is evident within the four-year record; however, this period is too short to infer any long-term hydrological or climatic trend. The overall range in annual fluctuations remains consistent across the study period.

3.2 Wavelet Analysis

3.2.1 Complete Time Series Analysis

The continuous wavelet power spectrum of the daily mean water levels for the Drumcliff River using the Morlet wavelet is depicted in Fig. 5. The scalogram displays regions of differing power across periods ranging from approximately 8 days to over 256 days. A band of consistently high power is visible at shorter periods around 8–32 days. Intermittent bursts of higher power are visible in the scalogram, occurring mainly during the winter months. Power outside the cone of influence, presented by the white dashed line, is less

reliable due to edge effects influencing the continuous wavelet transform. The white contour lines illustrated on the scalogram indicate the 95% significance level, marking areas where the observed wavelet power exceeds that expected from a red-noise background spectrum. These contours highlight the portions of the time-frequency space where river level oscillations are statistically stronger than the background variability. Wavelet power at periods longer than approximately 256 days is shown for completeness only and is not interpreted, as the four-year record (~1462 days) is too short to reliably resolve variability at these scales, and such long periods are additionally affected by edge effects within the cone of influence. All interpretation is therefore restricted to the 8–32 day and 64–256 day bands. Fig. 6 presents the scale-averaged wavelet power plot for two distinct periods across the four years of data, 8–32 days and 64–256 days. The curves in the graph illustrate how the strength of the oscillations varies over time.

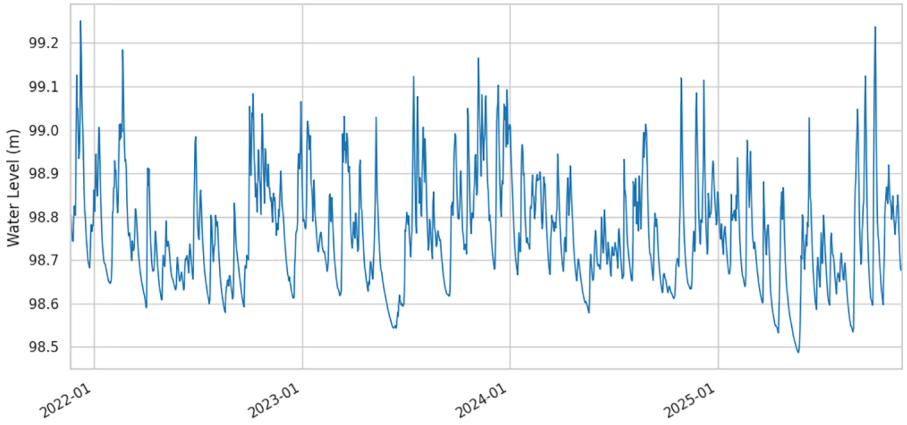


Figure 4: Daily mean water levels in the Drumcliff River, Sligo, Ireland.

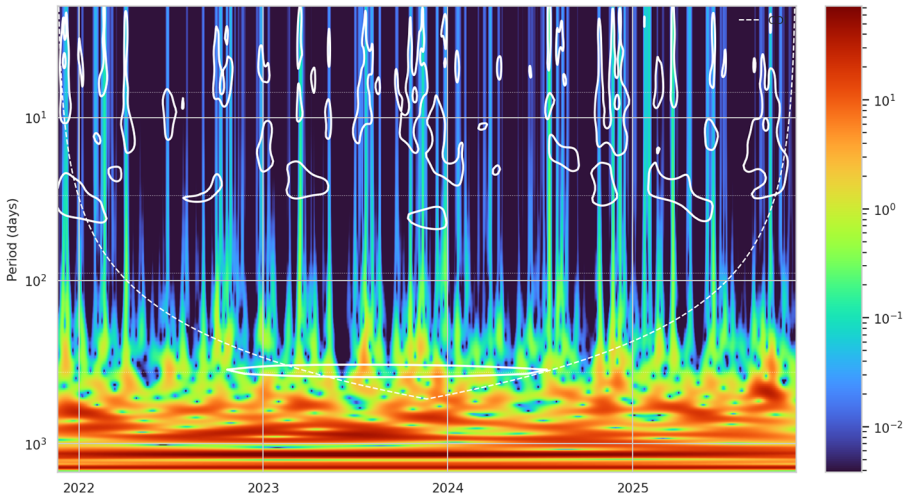


Figure 5: Continuous wavelet power spectrum (scalogram) of the Drumcliff River daily mean water level time series using the Morlet wavelet. Colour denotes wavelet power normalised by the variance of the series (dimensionless). The white dashed line denotes the cone of influence, outside which power is unreliable due to edge effects. White contour lines denote the 95% significance level relative to a red-noise (AR(1)) background spectrum.

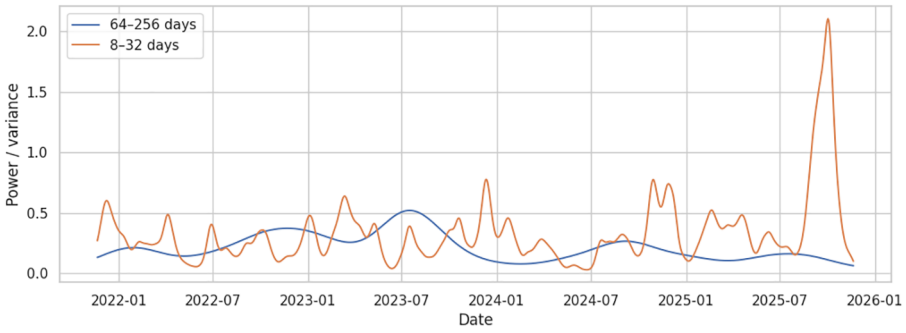


Figure 6: Scale-averaged wavelet power plot for the 8–32 day and 64–256 day period bands, Drumcliff River (2021–2025).

The short 8–32-day band exhibits variable behaviour with several short-duration peaks occurring annually throughout the time series. The high peaks occur most frequently in the winter months, with the largest amplitude across the four years occurring in October/November of 2025. The power ranges from approximately 0.1 to 0.7, with the exception of this high peak reaching 2.2. The 64–256-day band exhibits a smoother change of power over time. The power in this band ranges from approximately 0.2 to 0.5 with a gradual peak in mid–2023.

3.2.2 Seasonal Analysis

A seasonal analysis was conducted to obtain an in-depth understanding of the trends and oscillations occurring in the time series at different seasons. Fig. 7a–d presents the seasonal scale-averaged power plots for the 8–32 day and 64–256 day period bands for December-January-February (DJF), March-April-May (MAM), June-July-August (JJA) and September-October-November (SON) between 20 November 2021 and 20 November 2025 for the Drumcliff River. Each plot represents the mean power within the respective season, which allows seasonal differences to be identified more clearly than in the full series plot. The straight-line portions of the plots between each seasonal segment are due to the plotting of non-continuous, three-month periods on a continuous time axis. The 8–32 day band exhibits strong seasonal contrasts. In DJF, this higher frequency band shows several short-duration peaks each year, with the most pronounced occurring in early 2024. MAM exhibits similarly high early-season peaks, whereas JJA, the 8–32 day band, stays low throughout most of the record. There is a single isolated peak in mid–2023, followed by a rise to a consistently elevated level from mid–2024 through most of 2025, including summer periods. SON displays a broader increase in power toward the end of 2024 and then a high increase in late 2025. The 64–256-day band shows a smoother and more persistent pattern across all seasons. DJF and MAM generally exhibit lower power than JJA, which shows lower, more stable values. SON maintains a consistently low power throughout 2022–2024, before declining slightly in late 2025. The corresponding seasonal wavelet scalograms are provided in Appendix A to provide further visualisations of the results obtained.

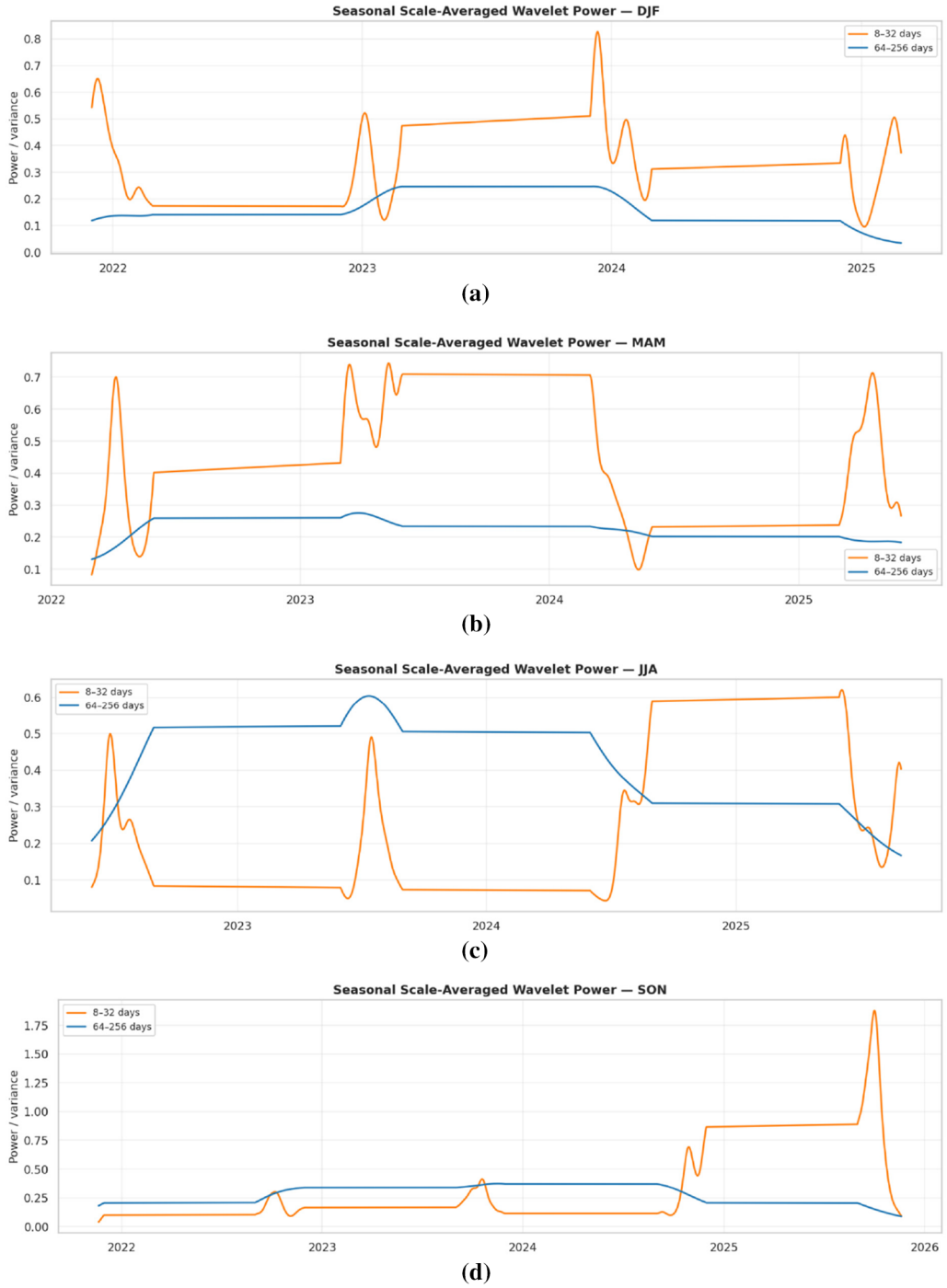


Figure 7: Seasonal scale-averaged wavelet power for the 8–32 day and 64–256 day period bands, Drumcliff River, shown separately for each meteorological season: (**a**) December–January–February; (**b**) March–April–May; (**c**) June–July–August; (**d**) September–October–November. All panels share a common time axis, period bands and graphical style. The straight segments between successive seasonal windows are plotting artefacts resulting from the display of non-continuous three-month periods on a continuous time axis and do not represent hydrological signal.

3.3 Three-Day LSTM Forecasting Model

3.3.1 Model Performance

The three-day forecasting LSTM model demonstrated consistently strong predictive performance across the training, validation, and test datasets. Table 3 outlines the training metrics obtained for the model. As outlined, the coefficient of determination (R^2) was high across all splits, staying above 0.80 for all results. Training obtained an R^2 value of 0.8766, validation was equal to 0.8462, and test obtained a value of 0.8676. The error-based metrics further improved the validity of these results. The RMSE values ranged between 0.0425 and 0.0489 m, and the MAE values ranged from 0.0263 to 0.0316 m. The Nash Sutcliffe Efficiency (NSE) value also matches the R^2 values, reinforcing that the model performs significantly better than a simple mean predictor. The PBIAS values remain near zero for all datasets, ranging between −0.0003 and −0.0072, indicating the model does not over- or under-predict water levels.

Table 3: Performance metrics for the three-day LSTM forecasting model.

Metric	Train	Validation	Testing
R^2	0.877	0.846	0.868
RMSE	0.043	0.047	0.049
MAE	0.026	0.03	0.032
NSE	0.877	0.846	0.868
PBIAS	−0.72%	−0.03%	−0.44%

The actual vs. predicted water levels time series plots for the training, validation, and test sets are illustrated in Fig. 8. The plots show that the LSTM model accurately captures both long and short-term fluctuations in the river water levels efficiently. The training set showcases predicted values closely aligned with the observed river water levels, including rapid rises and falls in the water levels. The validation set continues to accurately track the daily patterns and fluctuations in the water levels, with some slight differences visible in the peak highs and lows. Although a visible difference, particularly in the peaks, is visible in the test set, the overall amplitude and shape of the observed hydrograph are maintained. Fig. 9 illustrates the scatter plot for the training, validation and test sets. This plot outlines the strong correlation between the model outputs and the observed water levels. The data points are clustered tightly along the 1:1 reference line across all three datasets, showing that errors are generally small and the prediction errors become slightly larger at higher water level values, most visibly in the test set.

3.3.2 Forecasting Performance

Table 4 presents the three-day forecasting results, the observed water levels, and the error between the results. As expected with a multi-day forecasting model, the error increases over time. Fig. 10 illustrates the error between forecasted and actual values over the three-day forecast. The smallest error was observed on the first day with a minimal error of 0.0138 m or 1.38 cm. The error on day two increased slightly but was still relatively small at 0.0555 m or 5.55 cm. However, on the third day, the model produced its largest error, with a substantial value of 0.2195 m or 21.95 cm.

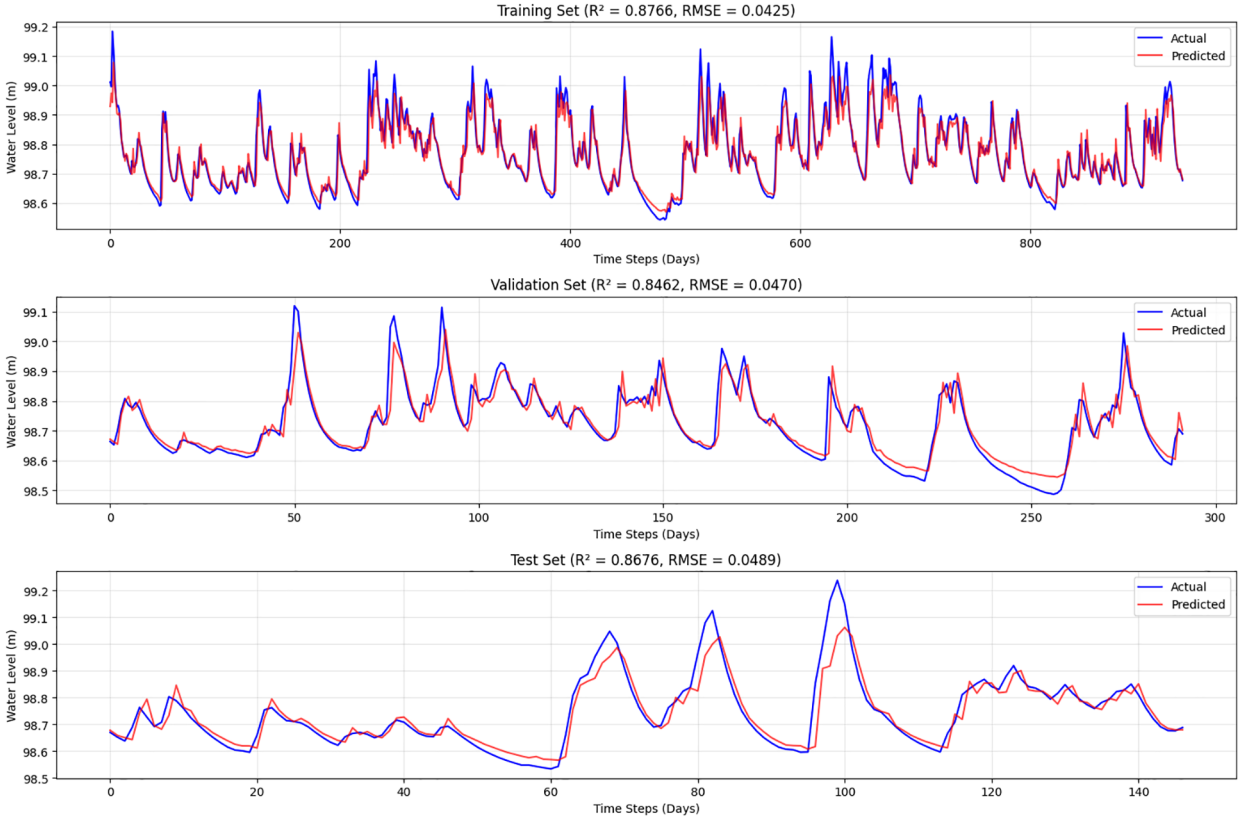


Figure 8: Time series plots of predicted vs. observed daily mean water levels for the training, validation and test partitions of the three-day LSTM forecasting model. On the *x*-axis, each time step corresponds to one day.

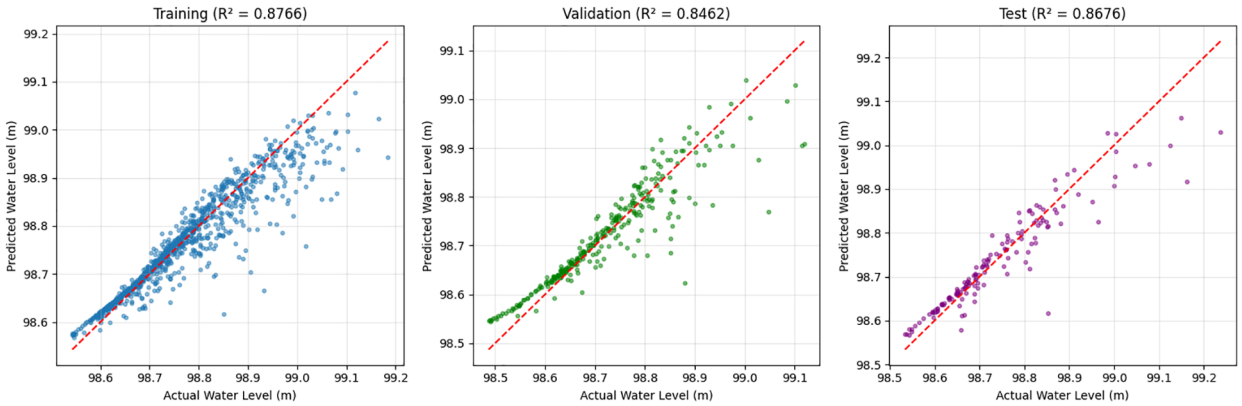


Figure 9: Scatter plots of predicted vs. observed daily mean water levels for the training, validation and test partitions of the three-day LSTM forecasting model.

Table 4: Three-day operational forecast vs. observed water level values.

Date	Forecast (m)	Actual (m)	Error (m)
21/11/2025	98.6978	98.6840	+0.0138
22/11/2025	98.7015	98.7570	−0.0555
23/11/2025	98.7024	98.9220	−0.2196

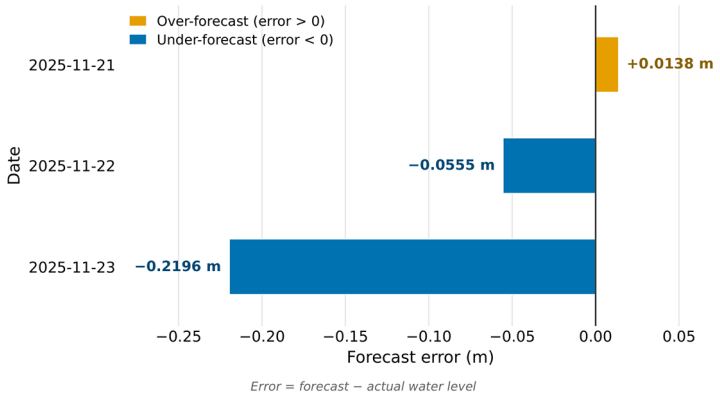


Figure 10: Error between forecasted and observed water level values over the three-day forecast window.

3.4 One-Day LSTM Forecasting Model

3.4.1 Model Performance

The one-day LSTM forecasting model demonstrated extremely strong predictive performance across the training, validation, and test datasets. Table 5 outlines the training metrics obtained for the model. The performance indicators are high across all test sets with R^2 values ranging from 0.982 to 0.988. The error-based metrics also remain low with RMSE values ranging between 0.013–0.017 m, and MAE values ranging from 0.008–0.01 m. The NSE values, which mirror the R^2 values, further reinforce the model's excellent predictive ability. The PBIAS values remain near zero, suggesting the model does not over- or under-predict across all data splits.

Table 5: Performance metrics for the one-day LSTM forecasting model.

Metric	Train	Validation	Testing
R^2	0.982	0.988	0.988
RMSE	0.017	0.013	0.015
MAE	0.01	0.008	0.009
NSE	0.982	0.988	0.988
PBIAS	−0.8%	−0.3%	−0.5%

Fig. 11 illustrates the time series plots. The plots outline close agreement between the predicted and observed water level data for each set. The training set accurately captures both small-scale fluctuations and large hydrological fluctuations present within the high resolution 15-min dataset. This high accuracy is maintained across the validation and test sets. The predicted curves track the actual values consistently,

with only small deviations and minor smoothing taking place during sharp rising peaks. Fig. 12 illustrates the scatter plot for the training, validation and test sets. These plots further support the observation made in the time series plots. The data points cluster tightly to the 1:1 reference line across all the datasets. The one-day model exhibits no visible increase in error at higher water level values, further reinforcing its ability to capture short-term hydrological variability with high precision.

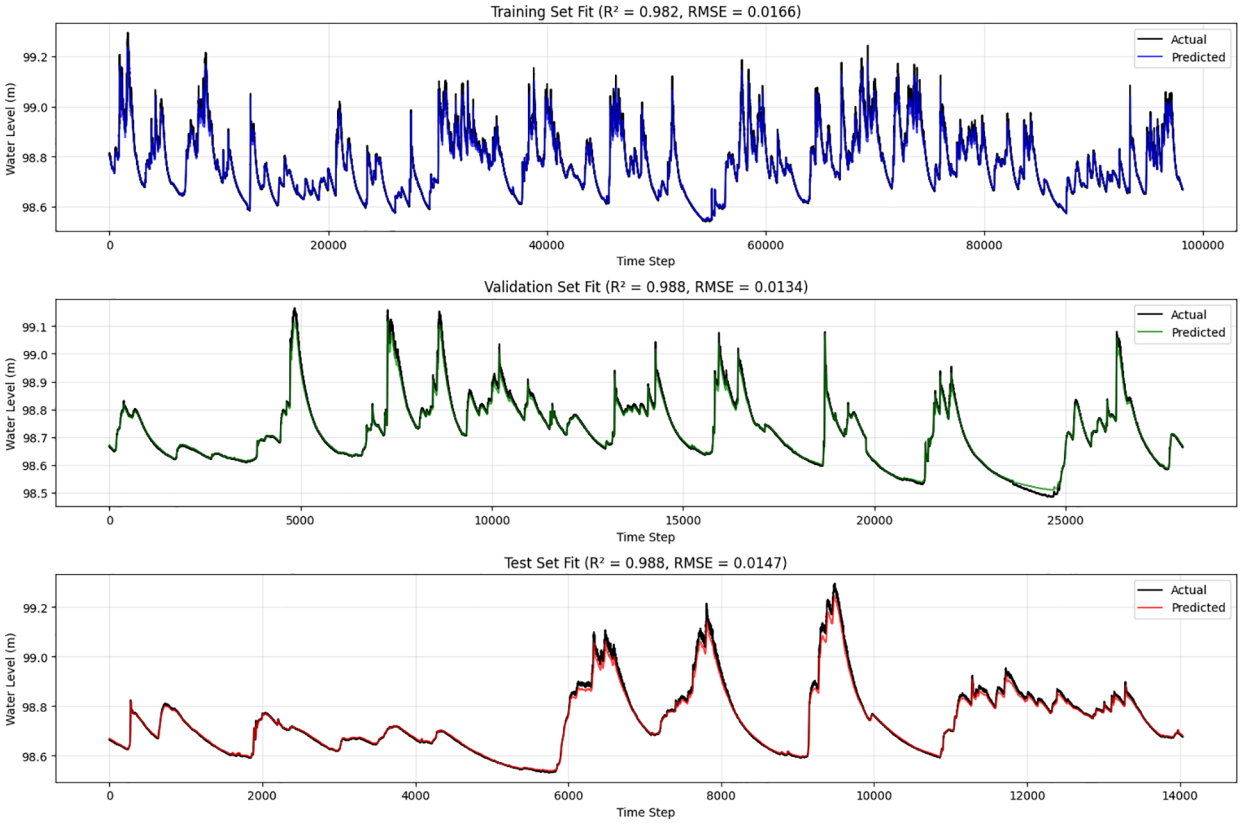


Figure 11: Time series plots of predicted vs. observed 15-min water levels for the training, validation and test partitions of the one-day LSTM forecasting model. On the *x*-axis, each time step corresponds to a 15-min interval.

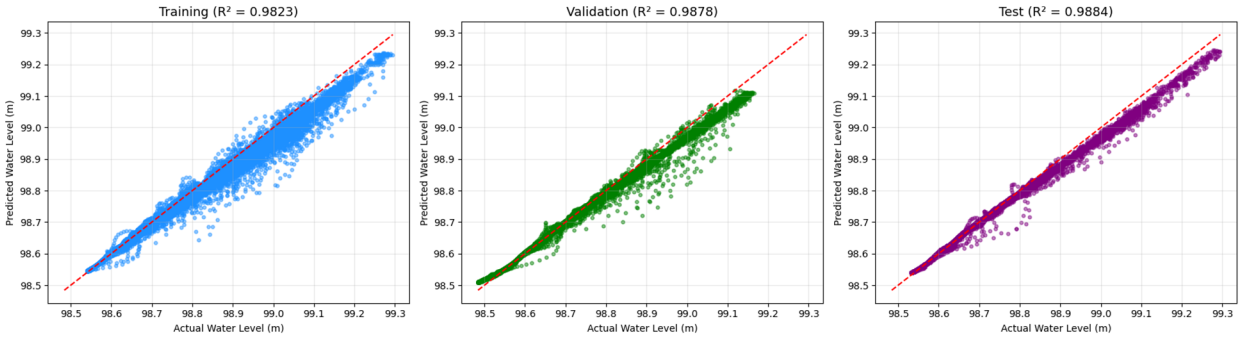
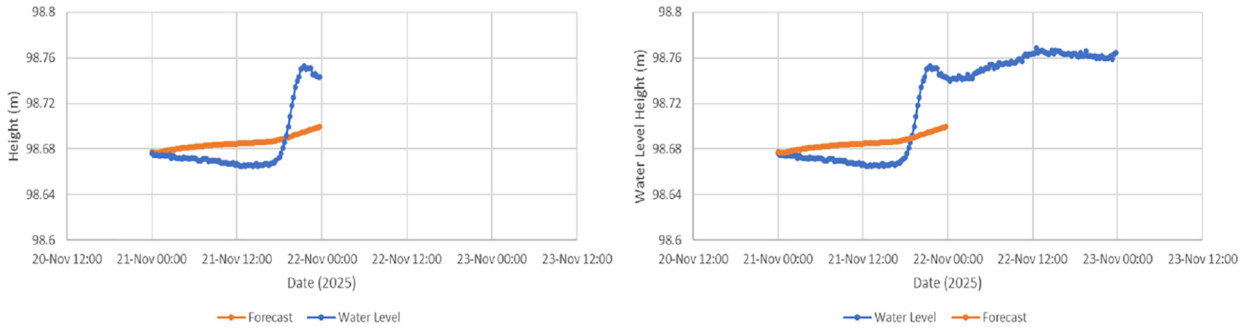


Figure 12: Scatter plots of predicted vs. observed 15-min water levels for the training, validation and test partitions of the one day LSTM forecasting model.

3.4.2 Forecasting Performance

The one-day forecasted values vs. the observed water level values are depicted in Fig. 13a. The model follows the overall trajectory of the observed values and produces a smooth forecast. However, the forecasted values remain slightly lower than the observed peak, with an error of 5–6 cm, highlighting a moderate degree of underprediction during periods of rapid hydrological water level fluctuations. Fig. 13b presents the same 24-h forecast plotted against the observed water level values for the subsequent 48-h period to further determine whether the overall trajectory of the forecasting model was accurate. Although the model was not designed to provide 48-h forecasts, this comparison offers an insight into the alignment of the 24-h forecast with the overall hydrological trajectory of the river water levels. The results show that the model follows the overall direction of the rising portion of the observed water levels, even though it continues to forecast underpredicted water levels.



(**a**) Compared for 1 day (**b**) Compared for 2 days

Figure 13: One-day LSTM forecast vs. observed 15-min water level values.

4 Discussion

4.1 Hydrological Variability and Seasonal Patterns

The time series analysis revealed a clear seasonal structure in the Drumcliff River, with elevated water levels during late autumn and winter and reduced levels during summer, reflecting the typical hydrological response of a temperate, Atlantic-facing catchment. The consistency of these seasonal fluctuations across the four-year record, with annual maximum and minimum values remaining within a narrow band of approximately 98.5–99.3 m, suggests a stable hydrological regime with no major changes in climatic or catchment conditions during the study period. This stability strengthens the reliability of both the wavelet analysis, and the LSTM forecasting models, as their training data represent a consistent and stable system.

The wavelet analysis provided a significantly deeper characterisation of this variability than the time series alone could offer. A persistent band of elevated power at 8–32 days dominated the signal throughout the record, reflecting the fortnightly to monthly timescale of episodic storm-driven hydrological responses in the catchment. This band was most intensely expressed during DJF and SON, aligning with increased rainfall, saturated catchment conditions and more frequent storm events at these times of year. These findings are consistent with Steele-Dunne et al. [4] and Bastola et al. [5], who identified an amplification of seasonal cycles characterised by increased winter precipitation and flows in Irish catchments. The complementary 64–256 day band, most prominent during summer, reflects the slowly evolving seasonal water balance governing the gradual drying and recharge of the catchment across the annual cycle.

The dominance of bare bog and agricultural land within the catchment directly governs this behaviour. Boglands, once saturated, generate near-instantaneous runoff responses to rainfall events, as their permanently high water table and low storage capacity leave little room to absorb additional rainfall, causing peatlands to act as sources of storm flow rather than as flood attenuators [26]. Compacted agricultural soils along the riparian corridor further reduce infiltration capacity and amplify surface runoff, since trampling and grazing pressure decrease soil porosity and thereby lower infiltration while increasing overland flow [27]. These land cover characteristics explain the strong short-period wavelet power consistently observed during autumn and winter and contextualise the highly responsive nature of the Drumcliff catchment to Atlantic-driven storm events.

4.2 Extreme Event Detection

A notable feature of the wavelet analysis was the identification of an exceptional hydrological episode in autumn 2025, characterised by a sharp surge in the 8–32 day power band reaching approximately 2.2, substantially exceeding the background power of 0.1–0.7 observed across the remainder of the record. This isolated high-energy burst indicates a sudden and significant hydrological event, likely as a result of an extreme rainfall event or a sequence of storms contributing to a rapid catchment response. As this event occurred in autumn, it aligns with the seasonal transition from summer into winter, with higher runoff conditions, increased soil saturation and declining evapotranspiration resulting in the catchment being more responsive to extreme events. This finding is consistent with O'Connor et al. [18], who identified significant increases in late autumn and winter flows across Atlantic seaboard catchments, including substantial upward trends in November, December and January. The clear detectability of this event within the wavelet power spectrum confirms both the physical realism of the episode and the usefulness of wavelet analysis for identifying and characterising extreme, short-duration hydrological events that traditional trend analysis methods would be unable to isolate.

4.3 LSTM Forecasting Performance

Both LSTM forecasting models demonstrated strong predictive performance within their intended forecast windows, confirming the feasibility of single-input, data-driven water level forecasting in this data-scarce catchment. The one-day model, trained on the high-resolution 15-min dataset of over 140,000 data points, achieved R^2 and NSE values of 0.982–0.988 with RMSE values of 0.013–0.017 m, capturing approximately 98% of the variance in the observed water level data. The three-day model achieved R^2 and NSE values of 0.846–0.877 with RMSE values of 0.043–0.049 m across all dataset partitions, explaining approximately 85%–88% of the observed variance using only four years of historical water level data and no meteorological inputs. The near-zero PBIAS values across both models indicate no systematic overall bias in either direction, supporting their suitability for short term forecasting.

Notwithstanding this overall lack of bias, both models displayed a tendency to underpredict peak water levels during periods of rapid hydrological rise, a well-recognised limitation of data-driven LSTM architectures trained without meteorological inputs. During rapid rainfall-driven catchment responses, the water level trajectory undergoes abrupt changes that are not predictable from prior water level history alone, as the driving forcing of precipitation is absent from the model input space. This limitation was most critically exposed in the three-day model, where forecast error increased progressively from 0.0138 m at 24 h to −0.0555 m at 48 h and −0.2196 m at 72 h. The large multi-day error reflects the model's inability to anticipate external drivers of water levels such as rainfall and soil moisture content, compounded by the restricted four-year training dataset limiting the model's exposure to a full range of multi-day hydrological behaviours.

These results indicate that the models provide their most reliable performance at the shortest lead times, with accuracy declining at longer horizons in the absence of meteorological inputs.

The forecasting performance presented herein aligns with previously reported results for deep-learning water level and discharge models in other studies. Notably, this research is distinguished by its dependence on a single input variable within a data-scarce catchment area. In the context of the lowland Tisza River, the Long Short-Term Memory (LSTM) model yielded the most precise seven-day-ahead water level forecasts among all tested methods, surpassing the operational Discrete Linear Cascade Model, as well as baseline, linear, and multilayer perceptron models [28]. Nonetheless, it consistently underestimated water levels across all forecast horizons and exhibited a delayed response to the rising limb of floods, demonstrating a propensity to underpredict rapid increases, a behaviour that closely parallels the peak underprediction observed in this study.

High short-term skill has been reported for LSTM flood-level prediction on the Yeong River, with peak accuracy within three hours [29]. Similarly, multi-station LSTM forecasts on the Red River of the North showed an R^2 decreasing from 0.999 at six hours to between 0.78 and 0.86 at one week [30], demonstrating the typical decline in model skill with longer lead times. Hybrid architectures such as Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) models incorporating satellite precipitation have maintained acceptable accuracy at a 24-h horizon in data-scarce and ungauged catchments [31], underlining the accuracy gains available when spatial rainfall information is included. This distinction is directly relevant to the present findings: a study explicitly contrasting rainfall-informed and water-level-only LSTM configurations found that direct prediction from water levels alone is more reliable when forecasted rainfall is unavailable, whereas incorporating forecasted rainfall substantially improves skill at longer lead times [32]. The short-term accuracy and the progressive multi-day error growth observed in this study are therefore fully consistent with the wider literature, and confirm that the horizon limitation documented here reflects the deliberate single-input, data-scarce design rather than a deficiency of the LSTM architecture itself.

4.4 Limitations and Further Work

This study is subject to several limitations that should be considered in the interpretation of its findings. The analysis focuses solely on the Drumcliff River, and its findings may not be representative of all Irish catchments, which may differ substantially in terms of geology, soils, land use, catchment size and climatic exposure. The four-year dataset, while consistent with similar studies in the literature, limits long-term trend detection and restricts the LSTM models' exposure to a full range of hydrological conditions; accordingly, the absence of a visible trend in the record should not be interpreted as evidence of long-term hydrological or climatic stability. Both forecasting models relied solely on historic water levels as the only input, restricting their ability to capture meteorological drivers such as precipitation and soil moisture content.

Further work could extend the study across multiple Irish rivers and longer datasets to provide a more comprehensive understanding of Ireland's overall hydrological variability. The inclusion of precipitation and soil moisture content data as additional LSTM inputs would substantially improve forecasting accuracy, particularly at longer forecast horizons and during extreme events. As the one-day forecasting model showcased strong short-range forecasting capabilities, there is potential for the model to be developed into an operational forecasting tool or early warning mechanism that could be utilised by local authorities and residents living within close proximity of the Drumcliff River.

5 Conclusion

This study investigated multi-scale hydrological variability and short-range water level forecasting in the Drumcliff River, County Sligo, Ireland, using Continuous Wavelet Transform analysis and Long Short-Term Memory neural network forecasting applied to four years of quality-controlled water level data. The Morlet-based CWT successfully revealed dominant periodicities across multiple timescales, with a persistent short-period band of 8–32 days dominating the signal during autumn and winter, reflecting the catchment's sensitivity to episodic Atlantic-driven storm events, and a complementary long-period band of 64–256 days governing summer variability through seasonal water balance dynamics. Seasonal decomposition confirmed a clear progression from event-driven hydrological variability during winter and autumn to seasonally governed variability during summer, consistent with existing Irish hydrological literature. An extreme hydrological episode identified in autumn 2025, characterised by a wavelet power peak of approximately 2.2, substantially exceeding background variability, further demonstrated the capability of wavelet analysis to detect and characterise extreme short-duration events undetectable by traditional methods.

The LSTM forecasting models confirmed that short-range water level prediction is feasible from historical water level data alone in this data-scarce catchment, with the one-day model achieving R^2 and NSE values of 0.982–0.988 and the three-day model performing adequately at 24 and 48 h horizons before degrading substantially at 72 h, highlighting the fundamental constraint of single-input models in the absence of meteorological forcing data. The dominance of saturated bogland and compacted agricultural soils within the catchment was identified as the primary land cover control governing the rapid hydrological response to storm events, directly influencing both the wavelet patterns observed and the forecasting performance of the LSTM models. This study presents, to our knowledge, one of the first applications of the continuous wavelet transform for diagnostic multi-scale decomposition of a river water level record in a small, data-scarce Irish Atlantic catchment, establishing a methodological foundation for advanced hydrological analysis across the Irish river network and providing a scientific basis for the development of operationally deployable short-term flood early warning systems in data-scarce catchments.

Acknowledgement: None.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: The authors confirm contributions to the paper as follows: Jack Ward: writing—original draft, visualization, methodology, software, formal analysis, data curation; Randika K. Makumbura: methodology, software, formal analysis, reviewing and revisions; Hirushan Sajindra: methodology, software, formal analysis, data curation, reviewing and revisions; Gul Jokhio: methodology, supervision; Bojan Đurin: methodology, supervision; Upaka Rathnayake: validation, supervision, reviewing and revisions, project administration, conceptualisation. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The datasets can be accessed using this link: https://www.epa.ie/our-services/monitoring–assessment/freshwater–marine/rivers/water-level-and-flow-data/.

Ethics Approval: Not applicable.

Conflicts of Interest: Given his role as the Guest Editor of this Special Issue, Upaka Rathnayake had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. The authors declare no other conflicts of interest.

Appendix A Seasonal Scalogram Plots

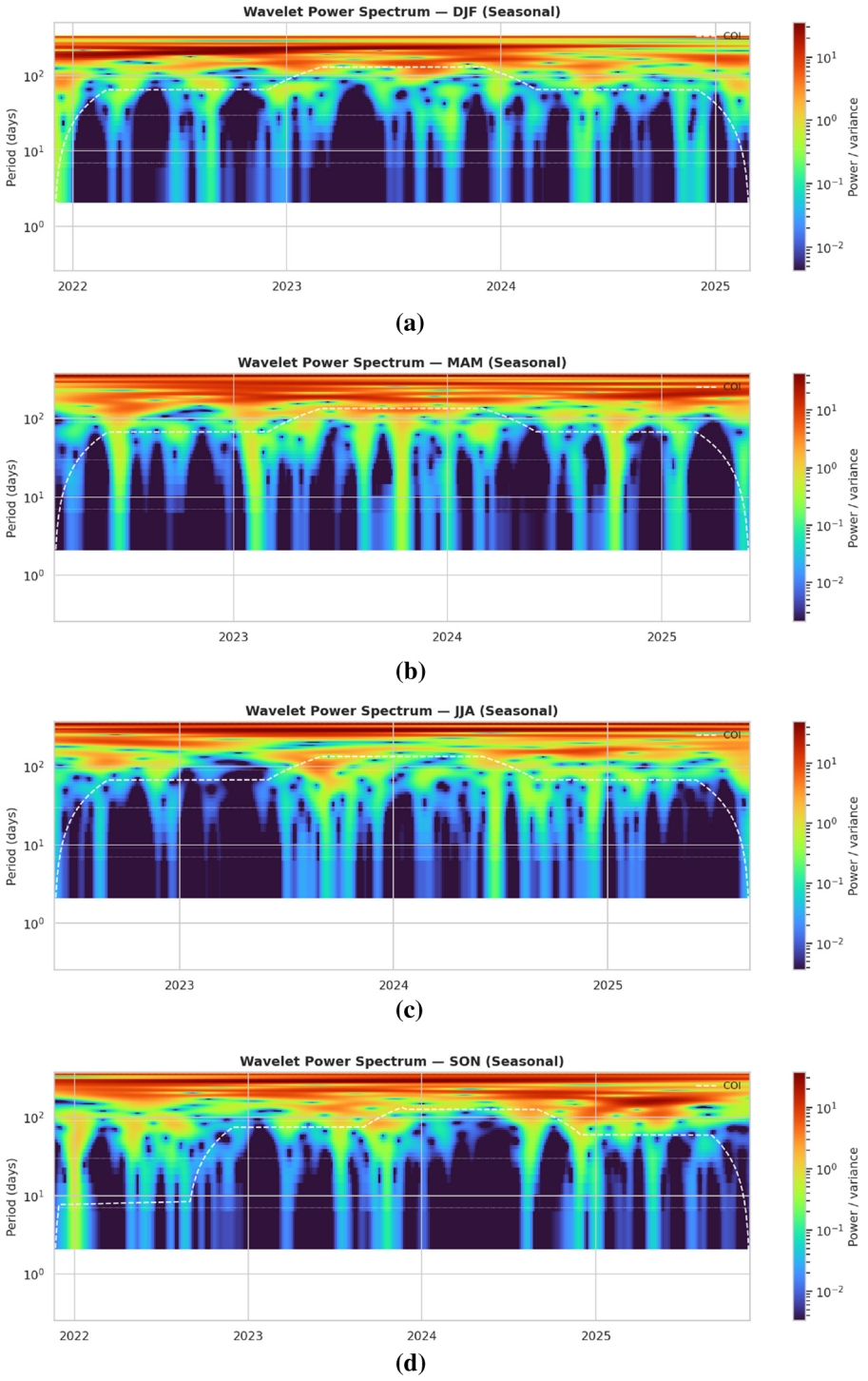


Figure A1: Seasonal wavelet power spectrums.

References

1. Yang D, Yang Y, Xia J. Hydrological cycle and water resources in a changing world: a review. Geogr Sustain. 2021;2(2):115–22. doi:10.1016/j.geosus.2021.05.003.
2. Sterly H, Borderon M, Sakdapolrak P, Adger N, Ayanlade A, Bah A, et al. Habitability for a connected, unequal and changing world. Glob Environ Change. 2025;90:102953. doi:10.1016/j.gloenvcha.2024.102953.
3. Cui L, He M, Zou Z, Yao C, Wang S, An J, et al. The influence of climate change on droughts and floods in the Yangtze River Basin from 2003 to 2020. Sensors. 2022;22(21):8178. doi:10.3390/s22218178.
4. Steele-Dunne S, Lynch P, McGrath R, Semmler T, Wang S, Hanafin J, et al. The impacts of climate change on hydrology in Ireland. J Hydrol. 2008;356(1–2):28–45. doi:10.1016/j.jhydrol.2008.03.025.
5. Bastola S, Murphy C, Sweeney J. The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. Adv Water Resour. 2011;34(5):562–76. doi:10.1016/j.advwatres.2011.01.008.
6. Sarkar Basu A, Gill LW, Pilla F, Basu B. Assessment of climate change impact on the annual maximum flood in an urban river in Dublin, Ireland. Sustainability. 2022;14(8):4670. doi:10.3390/su14084670.
7. Machiwal D, Jha MK. Time series analysis of hydrologic data for water resources planning and management: a review. J Hydrol Hydromech. 2006;54(3):237–57.
8. Milly PCD, Betancourt J, Falkenmark M, Hirsch RM, Kundzewicz ZW, Lettenmaier DP, et al. Stationarity is dead: whither water management? Science. 2008;319(5863):573–4. doi:10.1126/science.1151915.
9. Torrence C, Compo GP. A practical guide to wavelet analysis. Bull Amer Meteor Soc. 1998;79(1):61–78. doi:10.1175/1520-0477(1998)079<0061:apgtwa>2.0.co;2.
10. Schulte JA. Wavelet analysis for non-stationary, nonlinear time series. Nonlin Processes Geophys. 2016;23(4):257–67. doi:10.5194/npg-23-257-2016.
11. Sang YF. A review on the applications of wavelet transform in hydrology time series analysis. Atmos Res. 2013;122:8–15. doi:10.1016/j.atmosres.2012.11.003.
12. Bola S, Rudra R, Shukla R, Singh A, Goel P, Daggupati P, et al. Continuous wavelet analysis of water quality time series in a rapidly urbanizing mixed-land-use watershed in Ontario, Canada. Sustainability. 2025;17(19):8685. doi:10.3390/su17198685.
13. Lee M, You Y, Kim S, Kim KT, Kim HS. Decomposition of water level time series of a tidal river into tide, wave and rainfall-runoff components. Water. 2018;10(11):1568. doi:10.3390/w10111568.
14. Treviño J, Rodríguez-Rodríguez M, Montes-Vega MJ, Aguilera H, Fernández-Ayuso A, Fernández-Naranjo N. Wavelet analysis on groundwater, surface-water levels and water temperature in Doñana National Park (coastal aquifer in southwestern Spain). Water. 2023;15(4):796. doi:10.3390/w15040796.
15. Abebe SA, Qin T, Zhang X, Yan D. Wavelet transform-based trend analysis of streamflow and precipitation in Upper Blue Nile River basin. J Hydrol Reg Stud. 2022;44:101251. doi:10.1016/j.ejrh.2022.101251.
16. Sang YF. Improved wavelet modeling framework for hydrologic time series forecasting. Water Resour Manag. 2013;27(8):2807–21. doi:10.1007/s11269-013-0316-1.
17. Murphy C, Harrigan S, Hall J, Wilby RL. Climate-driven trends in mean and high flows from a network of reference stations in Ireland. Hydrol Sci J. 2013;58(4):755–72. doi:10.1080/02626667.2013.782407.
18. O'Connor P, Meresa H, Murphy C. Trends in reconstructed monthly, seasonal and annual flows for Irish catchments (1900–2016). Weather. 2023;78(9):261–7. doi:10.1002/wea.4288.
19. Moradian S, Iglesias G, Broderick C, Olbert IA. Assessing the impacts of climate change on precipitation through a hybrid method of machine learning and discrete wavelet transform techniques, case study: Cork, Ireland. J Hydrol Reg Stud. 2023;49:101523. doi:10.1016/j.ejrh.2023.101523.
20. Environmental Protection Agency. Water framework directive cycle 2 catchment assessment: Drumcliff (35_1). Washington, DC, USA: EPA Catchment Science & Management Unit; 2018.
21. Met Éireann. A climatological assessment of long-term rainfall averages and patterns in Ireland. Dublin, Ireland: Met Éireann (The Irish Meteorological Service); 2012.
22. Addison PS. The illustrated wavelet transform handbook. Boca Raton, FL, USA: CRC Press; 2017. doi:10.1201/9781315372556.

23. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735–80. doi:10.1162/neco.1997.9.8.1735.
24. Makumbura RK, Wijesundara H, Sajindra H, Rathnayake U, Kumar V, Duraibabu D, et al. QPred: a lightweight deep learning-based web pipeline for accessible and scalable streamflow forecasting. Comput Mater Contin. 2026;87(2):45. doi:10.32604/cmc.2026.075539.
25. Makumbura RK, Manatunge J, Rathnayake U. Bridging data-driven and process-based approaches for hydrological modeling in the tropics: insights from the Kelani River Basin, Sri Lanka. Results Eng. 2025;27:105975. doi:10.1016/j.rineng.2025.105975.
26. Holden J, Chapman PJ, Labadz JC. Artificial drainage of peatlands: hydrological and hydrochemical process and wetland restoration. Prog Phys Geogr Earth Environ. 2004;28(1):95–123. doi:10.1191/0309133304pp403ra.
27. Centeri C. Soil water erosion. Water. 2022;14(3):447. doi:10.3390/w14030447.
28. Vizi Z, Batki B, Rátki L, Szalánczi S, Fehérváry I, Kozák P, et al. Water level prediction using long short-term memory neural network model for a lowland river: a case study on the Tisza River, Central Europe. Environ Sci Eur. 2023;35(1):92. doi:10.1186/s12302-023-00796-3.
29. Kim HI, Jang SD, Choi H, Kim TH, Kim B. Prediction of flood level using LSTM and watershed hydrological data. J Flood Risk Manag. 2025;18(4):e70123. doi:10.1111/jfr3.70123.
30. Atashi V, Gorji HT. Enhanced flood prediction using LSTM and climate parameters: multi-station analysis of snowmelt-induced flooding in the Red River of the North. J Hydroinf. 2025;27(2):245–60. doi:10.2166/hydro.2025.236.
31. Wongchaisuwat P, Yomwilai K, Thaisiam W. A hybrid CNN-LSTM approach for multi-step discharge forecasting with satellite data. Front Water. 2026;8:1753598. doi:10.3389/frwa.2026.1753598.
32. Ji Y, Lim Y, Kim D, Sung J, Kang B. Effects of forecasted rainfall on direct and recursive LSTM-based river water level predictions. J Flood Risk Manag. 2025;18(4):e70147. doi:10.1111/jfr3.70147.