Hydrological Processes

WILEY

RESEARCH ARTICLE OPEN ACCESS

Regionalization of a Distributed Hydrology Model Using Random Forest

Siavash Pouryousefi Markhali[1] | Marie-Amélie Boucher[2] | Annie Poulin[1] | Mehrad Rahimpour Asenjan[1] | Frédéric Talbot[1] | Richard Arsenault[1]

[1]HC3-Hydrology Climate and Climate Change Laboratory, Department of Construction Engineering, École de Technologie Supérieure, Montréal, Québec, Canada | [2]Civil and Building Engineering Department, Université de Sherbrooke, Sherbrooke, Québec, Canada

Correspondence: Siavash Pouryousefi Markhali (siavash.pouryousefi-markhali@etsmtl.ca)

Received: 5 January 2025 | **Revised:** 28 January 2026 | **Accepted:** 1 February 2026

Keywords: distributed model | Random Forest | streamflow regionalisation

ABSTRACT

This study employs a Random Forest machine learning model in conjunction with a process-based hydrological model for streamflow regionalisation. The models were applied across varying spatio-temporal resolutions to test three key hypotheses: (I) the Random Forest model can better capture the nonlinear relationships between parameters and catchment descriptors compared to conventional regionalisation methods; (II) using finer temporal resolution can enhance parameter calibration, thereby improving the efficiency of the regionalisation model and (III) incorporating spatially distributed parameters can increase the model's efficiency. The results indicate that the Random Forest model outperforms conventional regionalisation methods. Furthermore, refining the temporal resolution increases model performance. For daily simulations, spatial refinement of catchment descriptors results in an approximate 10% improvement in regionalisation skill, while no discernible improvement is observed in simulations with sub-daily time steps.

1 | Introduction

Process-based distributed models have been applied across various hydrological modelling problems including flood forecasting, climate change impact assessment and analyses of hydrological processes at different spatio-temporal scales (e.g., Addor et al. 2014; Blöschl et al. 2008; Kumar et al. 2013; Rakovec et al. 2018; Thober et al. 2019; Martel et al. 2020; Valencia Giraldo et al. 2023; Talbot et al. 2025). However, for most process-based hydrological models, some parameters cannot be directly determined and therefore, the generally accepted practice is to calibrate model parameters using observed data (Fatichi et al. 2016). Still, in many locations, even in developed countries, there is a lack of observed streamflow data, or they are unreliable due to various difficulties (e.g., inaccessibility of the location, extreme weather conditions, vandalism, etc.) (Sivapalan 2003; Guo et al. 2021). The 'International Association for Hydrological Sciences' Decade on Prediction in ungauged basins' (PUB; Sivapalan 2003) and the Model Parameter Estimation Experiment (MOPEX) project (Duan et al. 2006) are examples of international large-scale initiatives on the topic of regionalisation. The objective of any regionalisation technique is to find a relationship between a model's parameters and catchment characteristics, which can then be used to estimate the parameters for ungauged catchments. This relationship can further be extrapolated to other modelling spatial units (i.e., catchments, sub-catchments, hydrological response units).

In general, regionalisation techniques have been classified into similarity-based methods (e.g., Vandewiele and Elias 1995; Randrianasolo et al. 2011; Samuel et al. 2011; Yang et al. 2018; Arsenault et al. 2019; Odusanya et al. 2022) and regression-based methods (e.g., Abdulla and Lettenmaier 1997; Wagener and Wheater 2006; He et al. 2011; Teutschbein et al. 2018; Singh and Devi 2022). The underlying assumption for similarity-based methods is that model parameters are transferable between catchments with similar physical characteristics. Proximity-based regionalisation methods make the assumption that nearby catchments share similar geophysical processes and are therefore likely similar (Arsenault and Brissette 2014, 2016). Regression-based methods rely on finding a relationship between catchment descriptors (CDs) and model parameters (Blöschl 2006; Razavi and Coulibaly 2013, 2017; Guo et al. 2021; Farfán and Cea 2023).

Regardless of the technique used for a specific regionalisation problem, a loss of modelling efficiency from calibration to regionalisation is to be expected. One concern regarding regionalisation is the a priori relationship assumed between model parameters and CDs. These assumptions are typically based on process understanding, expert knowledge and empirical evidence (Hrachowitz et al. 2013). However, such assumptions are often ambiguous. First, the link between parameters and CDs is not clear in most cases (Merz and Blöschl 2004). Second, there are different CDs which are correlated (e.g., precipitation and physiographic data) and provide a similar degree of information, resulting in equal power of prediction (Merz et al. 2020). Hence, the choice of which CDs control a specific parameter is not clear, making it difficult to constrain parameters with hydrologically reasonable transfer functions.

Machine learning (ML) techniques are capable of modelling complex multivariable relationships between predictors and predictands (e.g., Rahmani et al. 2021; Tyralis et al. 2019; Papacharalampous and Tyralis 2022; Feng et al. 2024; Shen et al. 2023; Zaerpour et al. 2021). ML offers greater flexibility compared to conventional regionalisation methods and can identify non-linear relationships between predictors and predictands through training on large datasets (Nearing et al. 2021; Lu et al. 2024). This approach eliminates the need for constraining parameters with transfer functions, a process that is both time-consuming and often uncertain. ML has primarily been applied in hydrology for prediction and benchmarking (Hsu et al. 1995; Abramowitz 2005; Best et al. 2015; Nearing et al. 2016; Kratzert et al. 2019, 2021; Martin 2023; Ma et al. 2024). Clustering techniques have also been occasionally used for catchment classification, addressing PUB through the lens of physical similarity (PS) (Papageorgaki and Nalbantis 2016; Singh et al. 2016; Kanishka and Eldho 2017; He et al. 2024). The most widely used ML algorithm in hydrology and water resources sciences is the multi-layer perceptron (Shen 2018).

Despite their academic applications, ML techniques still have limited operational use (Abrahart et al. 2012; Kirchner et al. 2006). This is because of ML models' 'black box' nature, in which the internal processes between inputs and outputs remain hidden or complex to track. Transparency and clarity are necessary for decision-making (Samek and Müller 2019; Boucher et al. 2020; Kumar et al. 2024). Water managers prefer to have the 'right answers' for the 'right reasons' to avoid unknown risks (Kirchner et al. 2006). One approach to close the gap between these two communities is to use ML techniques in parallel with hydrology models. This helps preserve the physical reality and interpretability of the modelling, while 'the underlying physical processes' that hydrology models cannot typically capture are approximated by numerical ML techniques (Kasiviswanathan and Jin 2016; Lei et al. 2024).

Random Forest (RF; Breiman 2001) is a powerful ML technique that, despite its strengths, has seen relatively limited application in hydrology (Tyralis et al. 2019). Notable examples include its use in flood and drought analyses (Anderson et al. 2018; Bachmair et al. 2016; Muñoz et al. 2018; Sultana et al. 2018; El Baida et al. 2024; Hameed et al. 2024) as well as in the study of hydrological signatures and flow regimes (Addor et al. 2018; Snelder et al. 2009; Balázs et al. 2018; Almagro et al. 2024). RF is a supervised learning method based on regression trees (Breiman 2001) and offers a degree of interpretability, as its decision-tree structure allows for traceability of the flow of information. Additionally, RF is fast, stable and resistant to overfitting (Boulesteix et al. 2012). It can handle datasets of varying sizes, even when predictors are highly correlated (Ziegler and König 2014; Tyralis et al. 2019). These characteristics make RF especially useful for regionalising hydrological model parameters, particularly in the context of computationally intensive, physics-based distributed hydrological modelling (Velásquez et al. 2022; Wang et al. 2023; Mehrvand et al. 2023).

While RF and other ML methods are increasingly used for streamflow simulation and parameter regionalisation, few studies have explored how modelling choices such as spatial resolution of parameters and temporal resolution influence the efficiency and robustness of these techniques. This study fills that gap by integrating RF into a process based hydrological framework and testing how variations in spatial and temporal discretisation affect regionalisation performance across multiple catchments by addressing the following question: How do spatio-temporal resolutions influence the performance of a RF-based regionalisation approach when coupled with a process based hydrological model for simulating streamflow in ungauged catchments? This question arises from the increasing integration of ML techniques, particularly RF, in hydrological modelling, and the need to better understand how modelling choices regarding spatial and temporal resolution affect their effectiveness.

This study utilises a distributed hydrological model to simulate streamflow in ungauged locations by applying the RF method and comparing it with conventional regionalisation approaches. Unlike many black-box applications of ML, RF is used here in a hybrid framework to estimate model parameters while preserving physical interpretability of the simulations. Additionally, we examine the impact of spatio-temporal discretisation of CDs on regionalisation efficiency, an aspect that has received limited attention in past studies. The hypotheses of this study are: (I) RF better captures the nonlinear relationships between parameters and CDs compared to conventional regionalisation methods; (II) using a finer time step provides more information for parameter calibration, thereby improving the regionalisation model's efficiency and (III) using finely distributed parameters across space improves the regionalisation model's efficiency. The structure of the paper is as follows: Section 2 introduces materials and methods including data, study area and the methodologies used for calibration of hydrological model and building the RF model. Section 3 provides the results and discussion and Section 4 gives a summary and the conclusions regarding our initial hypotheses.

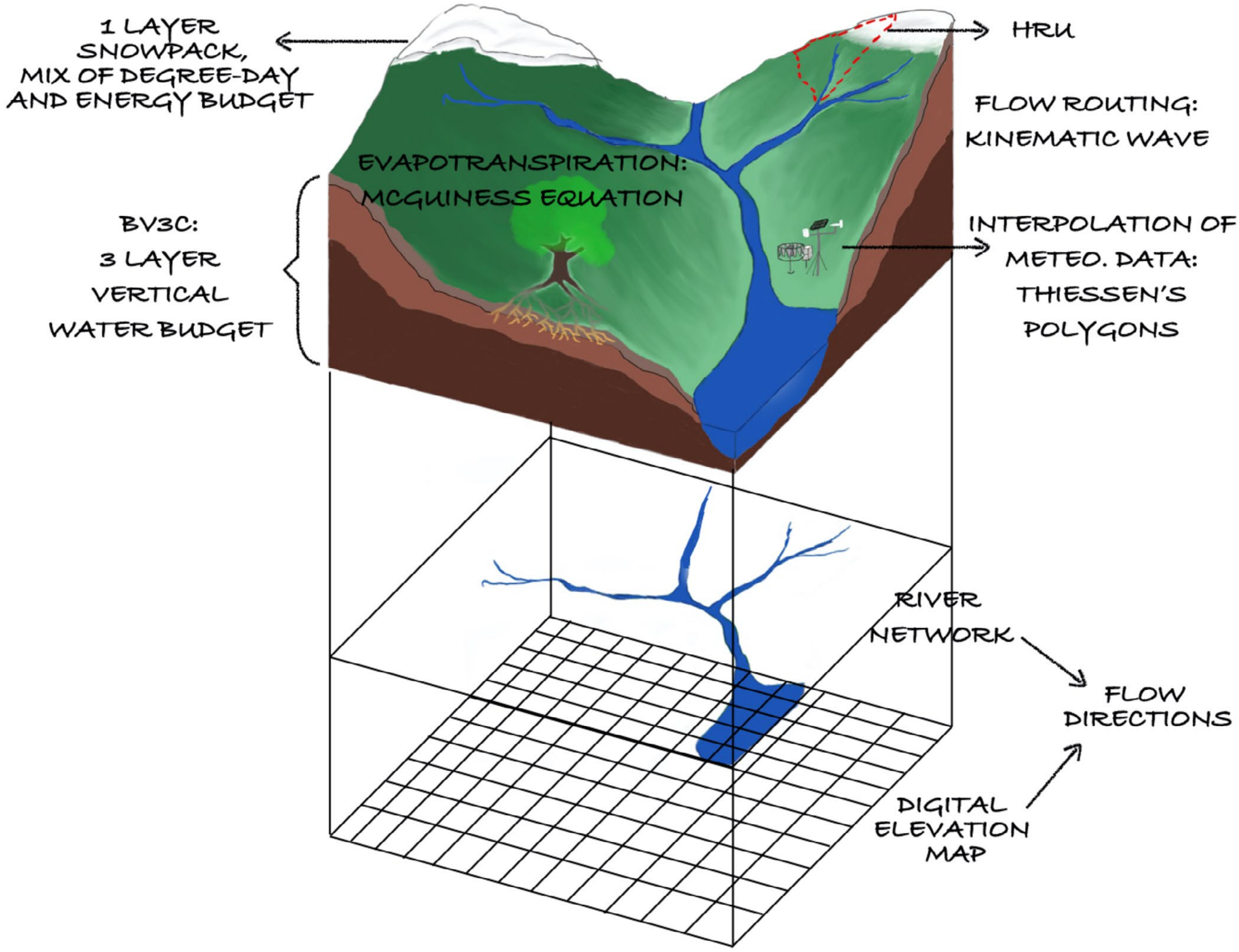

FIGURE 1 | Schematic presentation of the hydrological processes and sub-process embedded in Hydrotel.

2 | Material and Methods

2.1 | Hydrological Model

Hydrotel, a distributed, computationally intensive and partially physics-based model, is employed for this research (Fortin et al. 2001a, 2001b). The model is used by the *Direction principale des prévisions hydriques et de la cartographie* (DPPHC) as part of their flood forecasting system across the province of Quebec. For spatial inputs, the model receives GIS-based gridded data including land cover, soil type and Digital Elevation Model (DEM) raster as well as river network and lake polygons. The grids are further aggregated into multiple Relatively Homogenous Hydrological Units (RHHUs), small-scale subcatchments representing the spatial simulation units. To simulate the hydrological processes, the model offers options through various sub-models providing flexibility to the modelling practice. As Figure 1 illustrates, Hydrotel consists of six main submodels including one for the interpolation of meteorological data, vertical water budget, snow accumulation/melt, potential evapotranspiration, surface routing and channel routing. A mixture of empirical, conceptual and physical relationships constructs the governing equations to represent the processes and sub-processes. Overall, 27 parameters need to be specified. Some of these parameters can be fixed using expert knowledge or sensitivity analyses without significantly affecting model performance (Huot et al. 2019). For instance, the Manning roughness coefficient, used in surface and channel flow routing through the Kinematic Wave Equation, was not included in our calibration strategy. Instead, the Manning roughness coefficient was pre-assigned based on the land use information for each region. However, other parameters must be calibrated. The parameters calibrated in this study, along with their descriptions, are listed in Table 1.

2.2 | Study Area and Data

We selected 171 catchments from the southern and eastern parts of the province of Quebec in Canada. These catchments are selected among around 400 catchments available in the database provided by the DPPHC (https://www.cehq.gouv.qc.ca/atlas-hydroclimatique/index.htm). The database includes geographical coordinates and some physiographic information of the catchments as well as observed discharge at daily and hourly time-steps. All streamflow data were prepared for 24- and 3-h time-step. We eliminated catchments for which all hydrometric stations had been closed prior to 1990, or those with more than 40% of missing streamflow data in the calibration periods. Figure 2 demonstrates the spatial distribution of the 171 catchments across the province of Quebec.

The observed meteorological dataset was provided by the *Ministère de l'Environnement et de la Lutte contre les Changements Climatiques*, a department of the government of the province of Quebec. The dataset, called GC3h, is a gridded dataset at a 3-h time step containing precipitation as well as maximum and minimum temperature from 1990 to 2018. It is produced by kriging the measurements from over 350 ground stations on a 10 km by 10 km grid to produce the dataset over fixed grids. The fine spatio-temporal resolution of the dataset is an advantage over similar products, as the resolution of the modelling is of interest in the present study.

2.3 | Calibration

The dynamically dimensioned search algorithm (DDS-Tolson and Shoemaker 2007) is used to calibrate Hydrotel. Given that the procedure is computationally intensive, a set of 500 trials per catchment was allowed, using the Kling–Gupta Efficiency (KGE; Kling et al. 2012) as the objective function. The KGE was computed using Equation (1):

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{1}$$

where r is the linear correlation between observations and simulations, α is the variability ratio and β is the bias ratio between simulated and observed data.

The length of the calibration period varies between catchments depending on data availability. For most catchments, this period covers 10 years, anywhere between 1990 and 2017 as made available by meteorological data availability (normally 2008–2017, unless the measurement of the streamflow stopped before 2017). Yet there are few catchments for which measurements stopped before 2000 or where stations have recently opened and therefore have less than 10 years of data. The mean (median) of calibration KGE values for 24-h time-step is 0.81 (0.87) and for the 3-h time-step it is 0.71 (0.84).

2.4 | CDs

We define a set of CDs based on the existing literature (e.g., Arsenault and Brissette 2014; Merz and Blöschl 2004; Merz et al. 2020). These CDs are classified into four major groups. Table 2 presents the complete list of CDs used in this study along with their description, units and range among the 171 catchments (see Supporting Information for CDs' distributions).

2.5 | RF Method

In this study, we adopt the RF algorithm because it delivers an optimal trade-off among model interpretability, predictive accuracy on small-to-moderate datasets and computational efficiency (Tan et al. 2021; Vu et al. 2019). Our dataset including 171 basins characterised by 27 CDs, for which training a deep neural network can lead to an overfitting risk, and it offers limited scope for the complex hyperparameter tuning demanded by gradient-boosting frameworks such as XGBoost (Grinsztajn et al. 2022; Probst et al. 2019). In contrast, RF inherently resists overfitting through bootstrap aggregation, runs orders of magnitude faster than deep architectures or extensive boosting iterations, and provides transparent variable-importance diagnostics that directly support our hypothesis on catchment controls (Imani et al. 2025; Zaerpour et al. 2025).

TABLE 1 | Hydrotel's calibrated parameters and their description.

Parameter	Description (unit)	Submodel
L1	First layer thickness (m)	Vertical budget (BV3C)
L2	Second layer thickness (m)	
CR	Coefficient of recession (m/h)	
MTD	Melt threshold deciduous (°C)	Snow accumulation/melt (degree–day)
MTN	Melt threshold non-forest (°C)	
MTC	Melt threshold coniferous (°C)	
MRD	Melt rate deciduous (mm/d per °C)	
MRN	Melt rate non-forest (mm/d per °C)	
MRC	Melt rate coniferous (mm/d per °C)	
CET	Coefficient of optimisation (–)	Potential evapotranspiration (Mcguiness)
TSR	Transition from snow to rain (°C)	Interpolation (Thiessen polygons)

2.5.1 | Method Description

The RF algorithm is derived from decision trees. A decision tree model divides the input space into a number of simple segments, typically using sequential binary decision rules. Each decision splits the region into two (or more) nodes which are referred to as leaves. The number of decision rules is referred to as the *depth* of a tree. The average (in case of regression problems) or majority of votes (in case of classification problems) for each segment determines the output of the model. Decision trees are intuitive, fast and interpretable (Bishop and Nasrabadi 2006). But there are issues such as greediness (a tree getting trapped in local optimum) and arbitrariness (sensitivity over early splits and over the data points) of the algorithms that limit their applications (James et al. 2013).

An RF is a collection of decision trees and can handle both regression and classification problems. The idea is to train multiple decision trees over multiple samples of the data and take the average

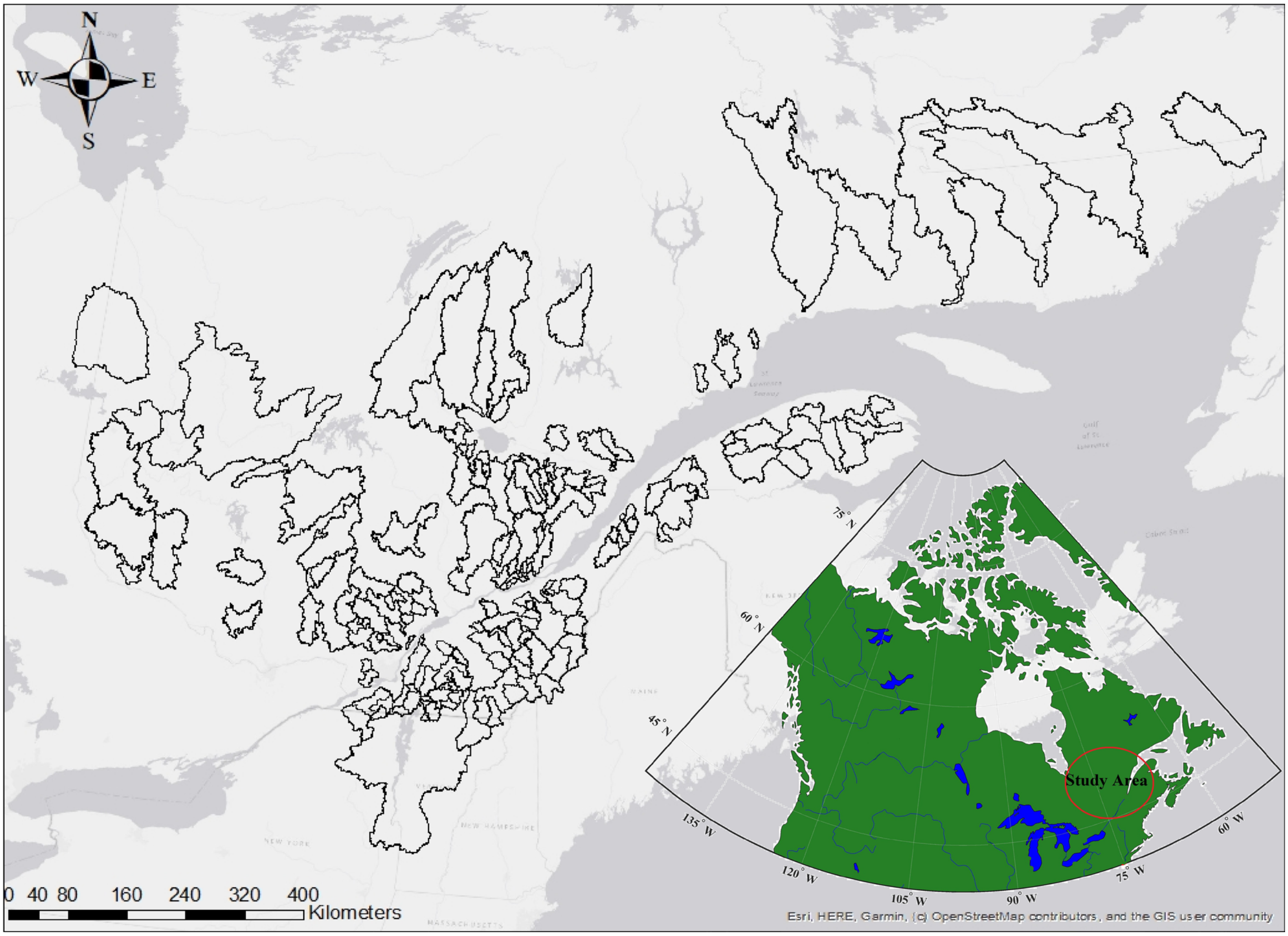

FIGURE 2 | Location of the 171 catchments used in this study, located in the southern part of the province of Quebec, in Canada. The red dots are major cities.

of that ensemble as the output. In order to create multiple samples of data, bagging, which is a combination of bootstrap and aggregation, is used (Breiman 2001). RF typically limits the maximum number of input variables corresponding to individual decision trees to maximise the variance of trees. Averaging over outputs of all trees therefore leads to a reduction of the error related to the variance of models. For this study, a regression RF model is employed to estimate the hydrological model's parameters.

One drawback of ensemble tree-based models relative to a simple decision tree is the reduction of interpretability. While the splitting procedure in decision trees is transparent and can demonstrate the effect of each split in the space of input variables, it is difficult to interpret an ensemble of trees. This, however, can be alleviated by computing the relative importance of input variables. We calculate the relative importance of input variables by recording the reduction of error corresponding with each input variable of a decision tree and average them out over all decision trees. The relative importance of an input variable, x_j, is calculated as:

$$\text{Importance}\left(x_j\right) = \frac{1}{N_{\text{trees}}} \sum_{t=1}^{N_{\text{trees}}} \sum_{s \in S_t(x_j)} \Delta\,\text{Error}_s \qquad (2)$$

where N_{trees} is the number of trees in the ensemble, $S_t\left(x_j\right)$ is the set of all splits on variable x_j in tree t and $\Delta\,\text{Error}_s$ is the reduction in error achieved by split s. As a result, a summary of the effect of input variables to approximate the model parameters can be obtained. Section 4.1 discusses the relative importance of input variables (here CDs) and the interpretability of the model.

2.5.2 | Experimental Setup for RF Regionalisation

Figure 3 schematically demonstrates the methodology for RF regionalisation. First, the hydrological model is calibrated over the selected catchments for both 24- and 3-h time-steps to obtain calibrated parameters. Afterwards, the RF regionalisation model is built using the calibrated parameters and CDs as training data. Before training the model, the catchments with poor KGE (<0.7, which accounts for 13% of catchments) in calibration are excluded to ensure the quality of the training data. The RF model is then trained using the scikit-learn library in Python, with key hyperparameters tuned using a grid search and fivefold cross-validation to minimise simulation error. The final model uses 500 trees (n_estimators = 500), a maximum tree depth of five (max_depth = 5) and a maximum features of three (max_features = 3). These values were selected as they offered the best

TABLE 2 | List of catchment descriptors (CDs), their definition, group type and range (minimum, average and maximum).

CD	Definition	Group type	Minimum	Average	Maximum
AI	Aridity index (etpm/prm)	Meteorological	0.4	0.5	0.6
ETPM	Mean annual potential evaporation (mm)	Meteorological	403.8	528.7	640.5
PRM	Mean annual precipitation (mm)	Meteorological	886.6	1076.5	1372.5
TMM	Mean annual temperature (°C)	Meteorological	0.3	3.6	6.8
TMAX	Mean annual maximum temperature (°C)	Meteorological	1.6	4.9	8.1
TMIN	Mean annual minimum temperature (°C)	Meteorological	−1.1	2.2	5.5
WR	Water (% of surface area)	Land cover	0.0	0.0	0.1
BS	Bare soil (% of surface area)	land cover	0.0	0.0	0.2
DF	Deciduous forest (% of surface area)	Land cover	0.1	0.3	0.6
AL	Agriculture lands (% of surface area)	Land cover	0.0	0.2	0.9
CF	Coniferous forest (% of surface area)	Land cover	0.0	0.4	0.8
IR	Impermeable (% of surface area)	Land cover	0.0	0.0	0.8
BL	Bogland (% of surface area)	Land cover	0.0	0.0	0.4
WL	Wetland (% of surface area)	Land cover	0.0	0.0	0.2
SN	Sand (% of surface area)	Soil type	0.0	0.5	1.0
SL	Sandy loam (% of surface area)	Soil type	0.0	0.4	1.0
SiL	Silt loam (% of surface area)	Soil type	0.0	0.0	0.7
CL	Clay loam (% of surface area)	Soil type	0.0	0.0	1.0
SA	Surface area (km^2)	Topographical	2.8	1800.0	22113.1
DD	Drainage density (km/km^2)	Topographical	0.0	0.3	2.3
LR	Lake ratio (km/km^2)	Topographical	0.0	0.0	0.2
WTI	Wetness index	Topographical	7.1	8.6	11.0
MSLP	Mean slope	Topographical	0.5	6.8	21.3
MEL	Mean elevation (m)	Topographical	20.4	344.7	858.4
CVEL	Coefficient of variation of elevation	Topographical	0.1	0.3	0.9
MASP	Median aspect (degrees)	Topographical	90.0	186.8	305.0
LAT	Latitude (decimal degree)	Topographical	44.9	47.2	52.2
LON	Longitude (decimal degree)	Topographical	−79.3	−71.6	−57.9

trade-off between model accuracy and generalisability based on cross-validation performance. The outputs of the regionalisation model are further used as inputs for Hydrotel to assess the accuracy of parameter estimation at pseudo-ungauged locations. By pseudo-ungauged locations, we mean that even though those catchments are gauged, they are considered ungauged to act as verification basins to assess the capacity of the parameters estimated using the RF model to provide accurate streamflow simulations with Hydrotel. In the next step, the efficiency of regionalisation for both 24 and 3-h time-steps is obtained for the test.

Simulations were performed with Hydrotel using the parameter sets corresponding to different levels of discretisation derived by RF for the nested catchments (see Section 3.3). We define three levels of simulations. (1) Simulations with fully-distributed parameters (FDP), which corresponds with approximation of parameters by RF at the RHHU level; (2) simulations with semi-distributed parameters (SDP), which corresponds with approximation of parameters by RF at the sub-catchment level and (3) simulations with lumped parameters (LP), which corresponds with approximation of the parameters by RF at the catchment level. The average efficiency of these simulations across sub-catchments is further compared for 3- and 24-h time-steps to highlight the model's sensitivity to varying spatio-temporal discretisation.

Moreover, we use the regionalisation model for nested catchments. A nested catchment is a parent catchment, which is normally large, that comprises multiple smaller catchments

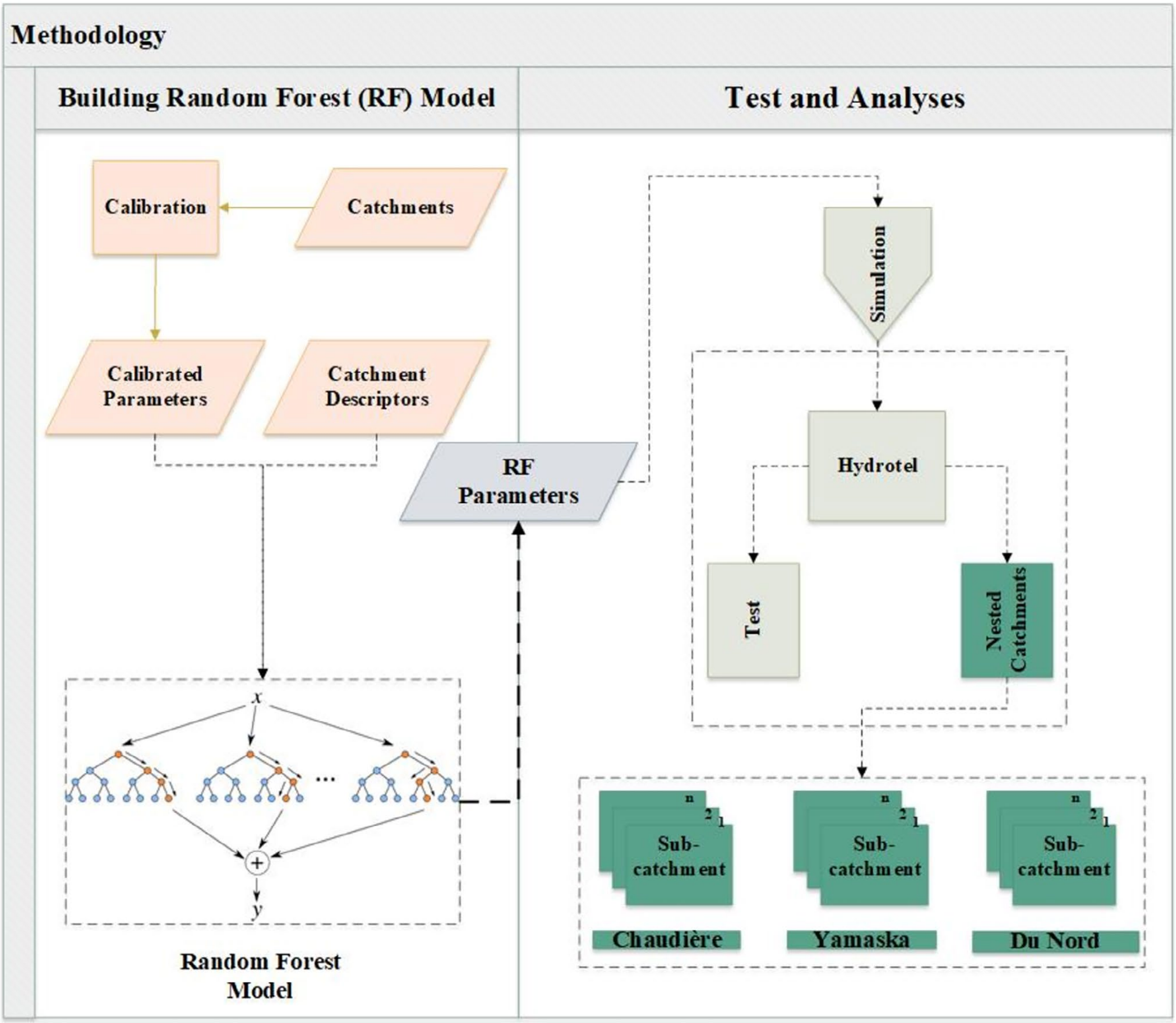

FIGURE 3 | Schematic description of the methodology for this study.

(or sub-catchments hereafter). The streamflow at different points inside the catchment is measured. Therefore, it is possible to evaluate the regionalisation technique at the internal pseudo-ungauged locations. There are three such catchments (i.e., Chaudière, Yamaska and du Nord) in the study area (Figures 4 and 5). Table 3 shows the surface area for these catchments.

2.6 | Conventional Regionalisation Schemes

2.6.1 | Multiple Linear Regression (MLR)

MLR is one of the earliest and most widely used methods for regionalising hydrological model parameters. It aims to create a statistical relationship between CDs, such as slope and land use, and hydrological model parameters calibrated across numerous sites. This relationship can then be applied to estimate parameters for ungauged catchments based on their physical characteristics. The idea behind this approach is that, ideally, the hydrological model parameters should align with the physical characteristics of the catchment.

To find the regression equation in the MLR approach, we only included catchments with a calibration KGE value above 0.7 to ensure acceptable hydrological model performance. Low-performing catchments were excluded in the regression as they would not provide meaningful information (Oudin et al. 2008; Arsenault et al. 2019). However, it is important to note that a high calibration KGE does not necessarily guarantee good transferability, as demonstrated by Garambois et al. (2015), who showed that parameter sets from poorly modelled catchments can sometimes yield better regionalisation performance. A key limitation of MLR is that different parameter sets may result in similar hydrology due to interactions between parameters, resulting in a weak regression (Oudin et al. 2008). Furthermore, MLR analyses each parameter separately, which diminishes the overall cohesion of the parameter set and may negatively impact model performance (Arsenault et al. 2019).

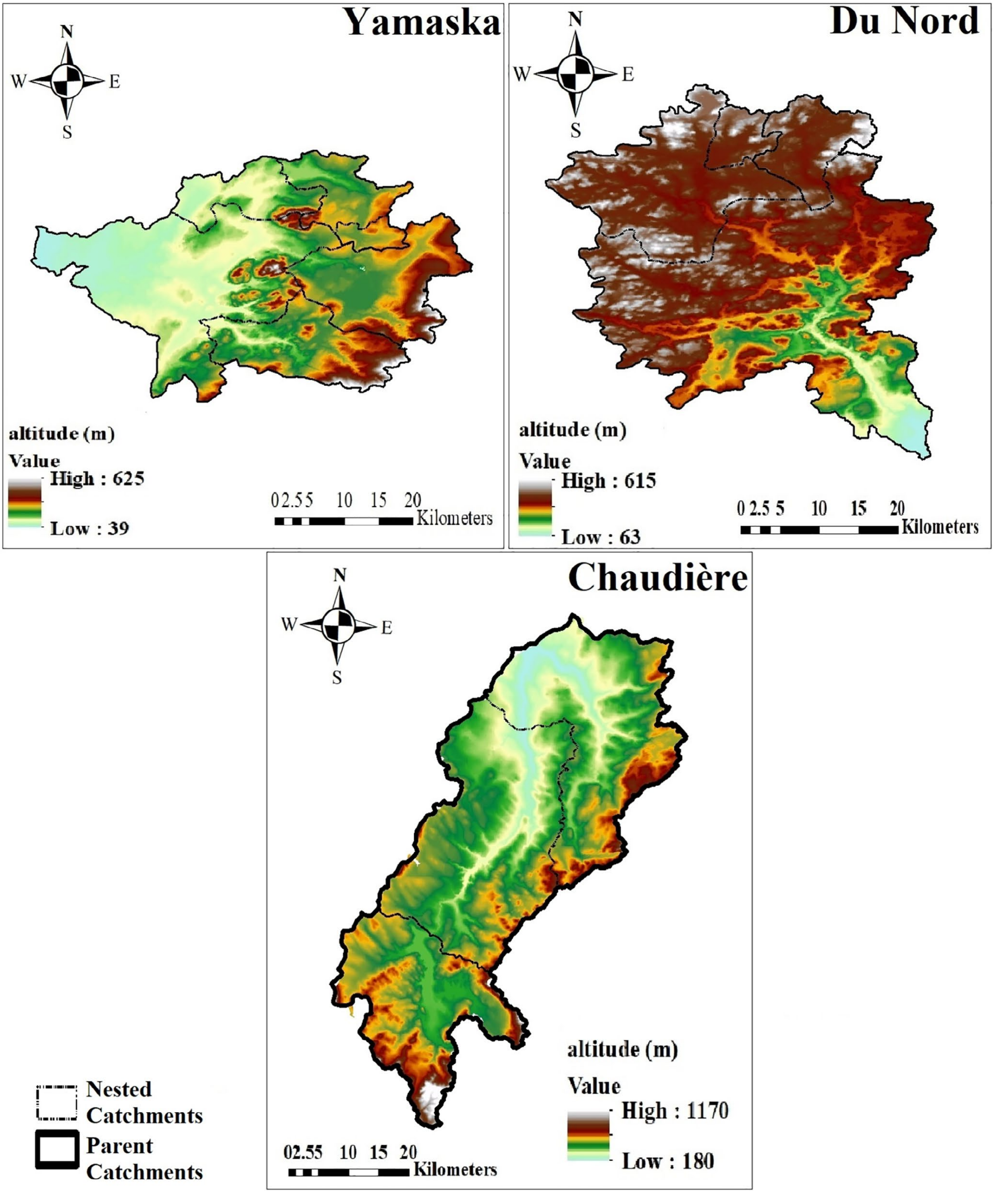

FIGURE 4 | Digital elevation models and boundaries of the nested catchments.

2.6.2 | Spatial Proximity (SP)

In the SP approach, parameters are transferred from nearby catchments based on the assumption that neighbouring catchments share similar characteristics due to comparable climate and catchment conditions (Blöschl 2006). Randrianasolo et al. (2011) demonstrated that, in France, neighbouring catchments can provide accurate predictions for ungauged basins, with at least five donor catchments required for effective model simulations. Although various versions of this method exist, we selected donor catchments using the Euclidean distance between their centroids. The success of this approach, however,

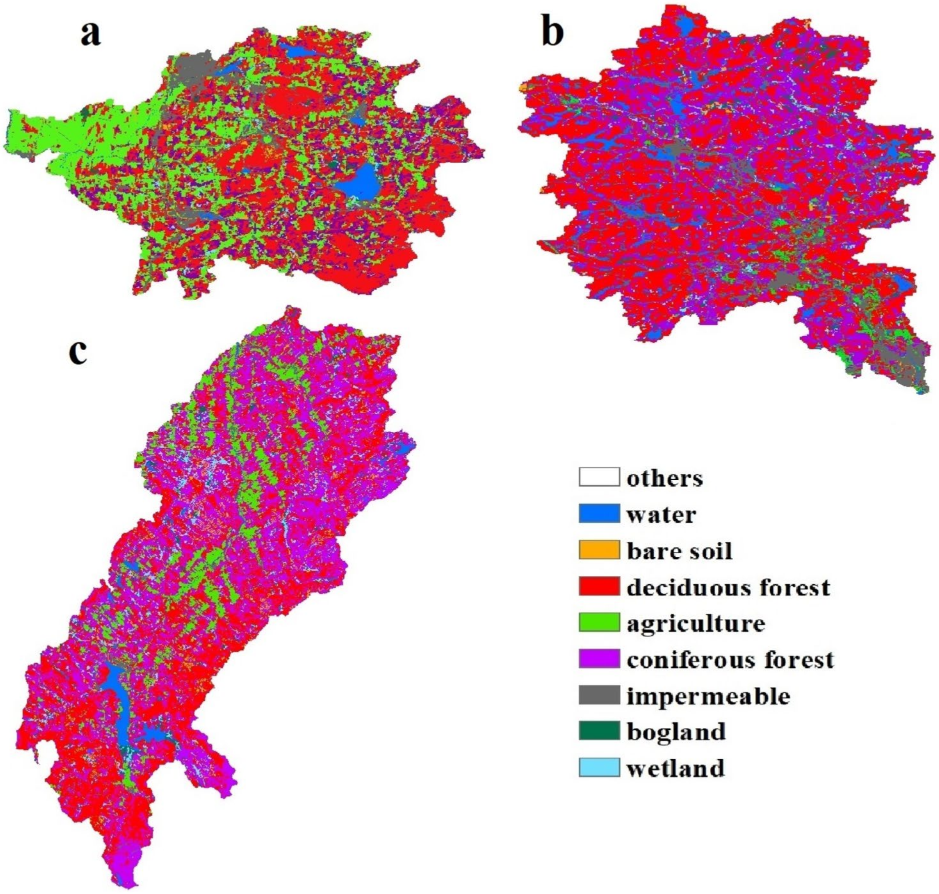

FIGURE 5 | Land use maps of the nested catchments. (a) Yamaska, (b) Du Nord and (c) Chaudière.

TABLE 3 | Nested catchments and sub-catchments surface area.

Title	ID	Area (km^2)	Description
Du Nord	40110	1163	Catchment
	40122	811	Sub-catchment
	40129	106	Sub-catchment
	40132	40.3	Sub-catchment
Yamaska	30302	1231	Catchment
	30309	131	Sub-catchment
	30314	214	Sub-catchment
	30340	235	Sub-catchment
Chaudière	30343	31.2	Sub-catchment
	30353	230.9	Sub-catchment
	23429	3085	Catchment

depends on the similarity between nearby gauged catchments and the density of the gauging network (Guo et al. 2021).

2.6.3 | PS

The PS method transfers parameters from gauged catchments that share similar attributes with the target catchment, based on the premise that catchments with similar characteristics will exhibit comparable hydrological responses. The similarity index, θ, is calculated using the normalised distance between each descriptor, as shown in Equation (3) (Burn and Boorman 1993):

$$\theta = \sum_{i=1}^{k} \frac{|\, CD_i^G - CD_i^U \,|}{\Delta\, CD_i} \tag{3}$$

where CD_i represents the ith catchment descriptor for both gauged (G) and ungauged (U) catchments and $\Delta\, CD_i$ is the range of possible values for CD_i.

Oudin et al. (2008) demonstrated that incorporating latitude and longitude as CDs enhances the performance of regionalisation, effectively combining the SP and PS approaches. This integrated method has been shown to outperform SP, PS and MLR in various studies (e.g., Parajka et al. 2013). The advantage lies in selecting donor catchments that are both similar in characteristics and geographically close. Latitude and longitude have also been used in the PS implementation in this study.

2.6.4 | Multiple Donor Averaging

For both the SP and PS methods, studies have shown that averaging from multiple donor catchments significantly improves regionalisation performance compared to using a single donor

(Oudin et al. 2008; Samuel et al. 2011). Additionally, utilising multiple donor catchments helps to effectively reduce random errors (Guo et al. 2021; Randrianasolo et al. 2011). Research suggests that using 5–6 donor catchments yields more accurate results (Guo et al. 2021; Oudin et al. 2008; Swain and Patra 2017a, 2017b), though the optimal number may vary based on the study region, regionalisation method and other factors (Guo et al. 2021). In this study, we applied the inverse distance weighting (IDW) to average parameters from five donor catchments.

3 | Results

In this section the results are presented following this order: Section 3.1 provides the RF model evaluation for 3- and 24-h time-steps and compares it with the calibration results to investigate Hypothesis I. Section 3.2 evaluates the relative importance of CDs in the determination of model parameters. Section 3.3 explores parameter transferability across spatial scales to verify/refute Hypothesis II. Section 3.3 is dedicated to multiscale hydrological simulations and sensitivity of the hydrological model to the spatial discretisation to validate Hypothesis III.

3.1 | Comparison of RF and Conventional Regionalisation Methods

To examine the first hypothesis we evaluate the performance of the RF model against conventional regionalisation methods. Figure 6 illustrates the KGE values of Hydrotel simulations using regionalised parameters, highlighting RF's significantly better performance, with a median KGE of 0.7. This is followed by MLR and SP, while PS shows the weakest results. The results emphasise RF's superior ability to capture complex relationships in hydrological data compared to classical methods. It is worth noting that the choice of the number of donor catchments plays a crucial role in determining simulation accuracy. However, conducting a thorough sensitivity analysis on the effect of donor catchment quantity is beyond the scope of this study. Despite this limitation, the overall findings strongly support RF as the most reliable approach for parameter regionalisation in this study.

3.2 | Impact of Time-Resolution on RF Regionalisation Model

Figure 7 illustrates the results of simulations carried with regionalised (RF) and calibrated (Cal) parameters for 3- and 24-h time-steps at pseudo-ungauged catchments. Subplot *a* compares the empirical cumulative distribution functions (ECDFs) of the modelling KGEs. The mean (median) of RF distributions for 24- and 3-h time-steps are 0.68 (0.7) and 0.75 (0.76), respectively, evidencing an improvement of the simulations when a finer time-step is used. The 24-h time-step training data has an advantage over the 3-h time-step for calibration in terms of efficiency: the mean (median) of distributions for 24- and 3-h time-steps are 0.88 (0.9) and 0.84 (0.85), respectively. This behaviour can be attributed to the effect of temporal aggregation. Calibration at a daily time step can smooth short-term variability and aggregate small-scale errors, thereby simplifying the optimisation

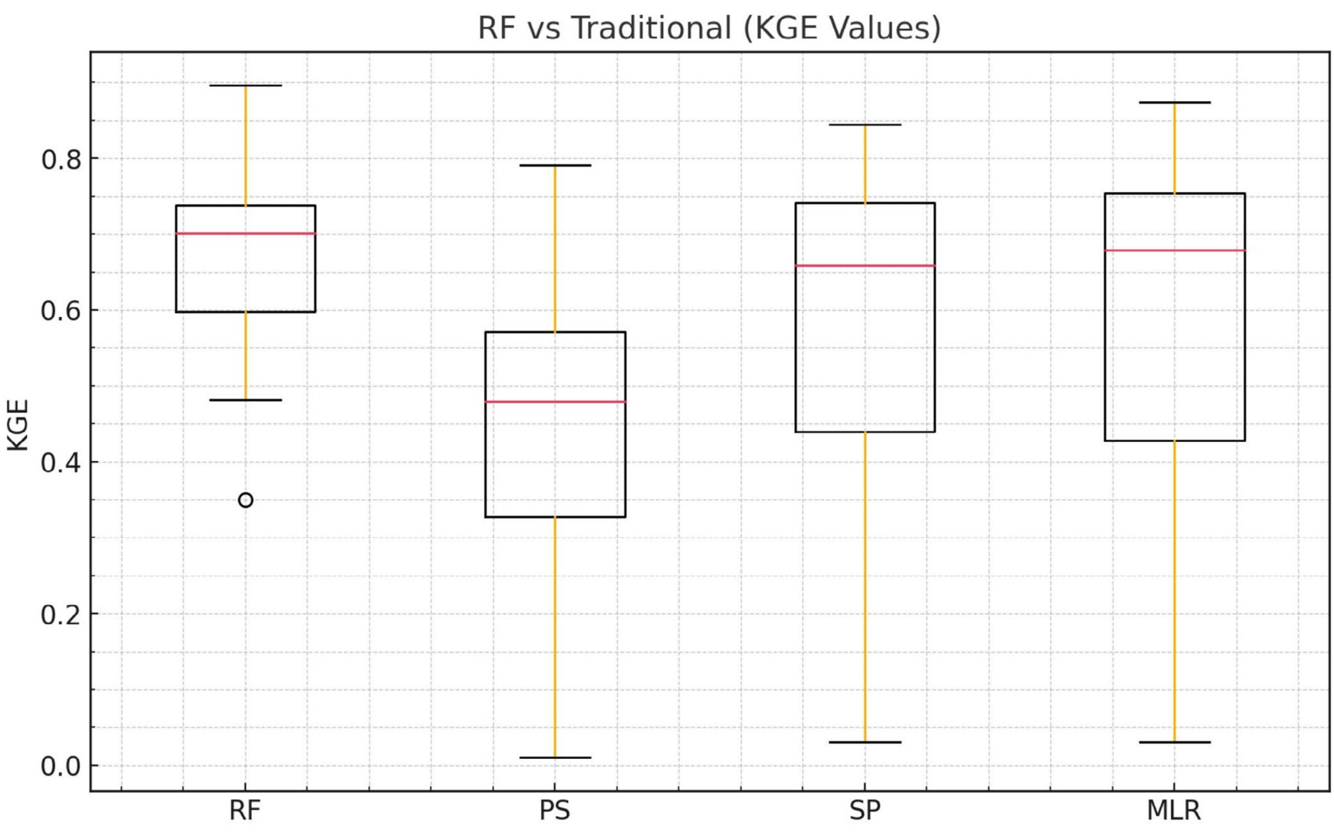

FIGURE 6 | Boxplots showing the KGE values of the regionalisation methods applied to the test's pseudo-ungauged catchments with a 24 h time step.

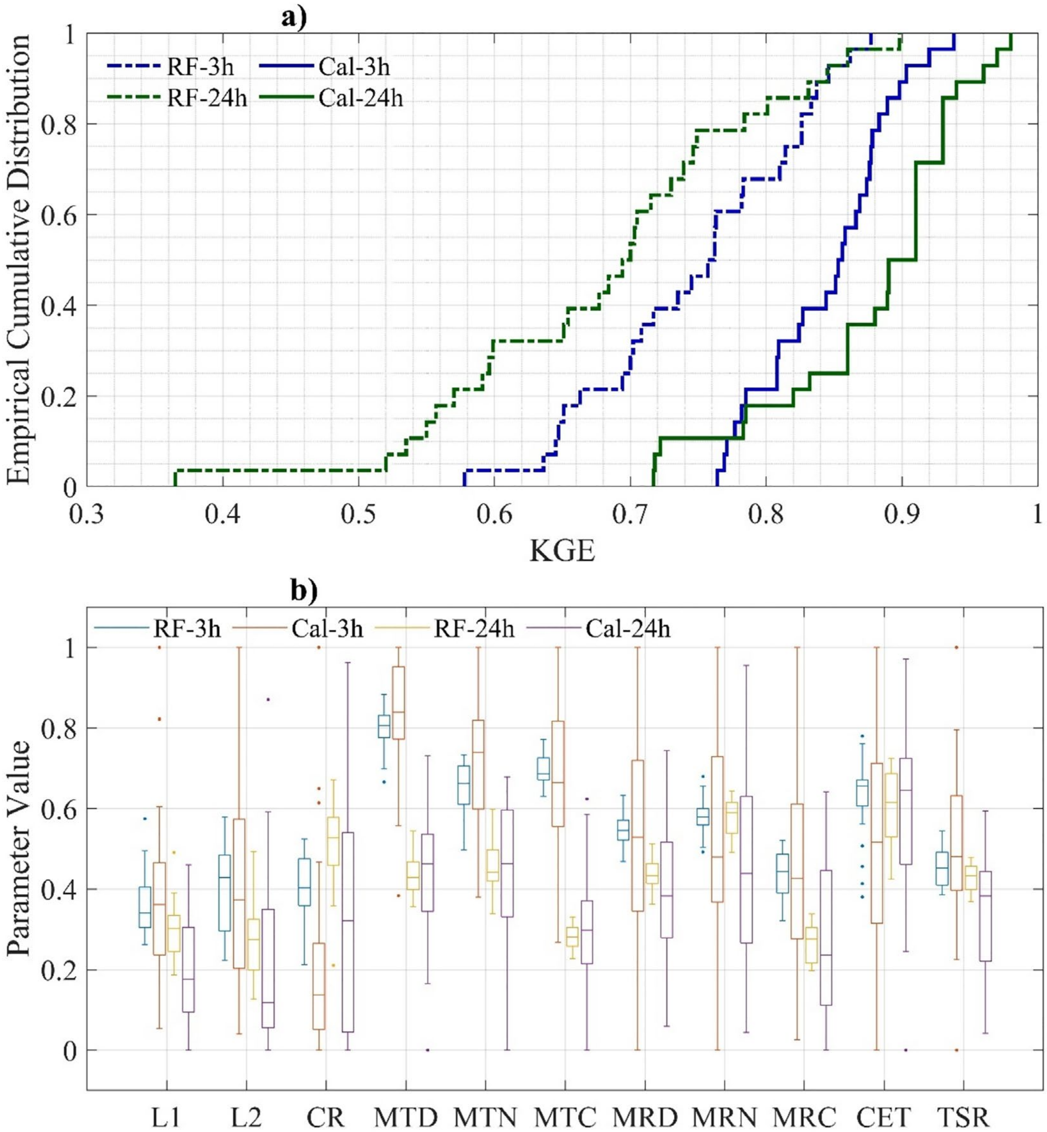

FIGURE 7 | Comparing calibration and RF-based regionalisation simulations for the test dataset (RF = Random Forest, Cal = Calibration). (a) Distribution of the regionalisation and calibration KGE for 3 and 24-h time-steps. (b) Standardised distribution of calibrated and RF-approximated parameters for 3- and 24-h time-steps, respectively.

problem and enabling higher calibration efficiency. In contrast, calibration at a 3-h time step requires the model to capture rapid and detailed hydrological processes, which increases the calibration complexity and can lead to slightly lower KGE scores (e.g., Bastola and Misra 2013). As expected from any regionalisation techniques, there is a loss of accuracy from calibration to regionalisation. However, this relative gap (Equation 3) is considerably larger for 24-h time-step simulations than for the 3-h time-step. Furthermore, the standard deviation (STD) of 24-h RF distributions is larger (STD = 0.12) than for 3-h (STD = 0.07) time-step.

Figure 7b shows the distributions of RF and calibration parameters in 3- and 24-h time-steps. Here, the RF and calibrated parameters are jointly standardised between 0 and 1 to facilitate the comparisons. As can be seen, the median of the RF and calibration distributions are in close proximity for most of the cases particularly for mixed degree-day-energy balance parameters (e.g., MTD to MRC), which demonstrates that the RF model successfully approximated the median of HYDROTEL's parameter distributions. However, the spread of calibrated parameters is larger than that of RF parameters and it seems that the RF regionalisation technique tends to systematically underestimate the spread of the parameters. Comparing the time-step of simulations, the 3-h RF parameters (i.e., RF-3h) have a better approximation in terms of the median of the corresponding calibrated parameters (i.e., Cal-3h) than that of 24-h time-step parameters for first and second soil layer thickness (i.e., L1 and L2). This might be the reason why a better performance is observed for 3-h time-step.

We acknowledge that RF shows limited skill in predicting vertical water-budget components for L1 and L2 at the 24-h scale and for CR at both 3- and 24-h scales. This may reflect the strong influence of sub-surface heterogeneity, soil texture variability, bedrock depth and preferential flow paths, which are only coarsely represented by the 27 CDs used as model inputs. While these descriptors capture broad physiographic characteristics, they may not sufficiently resolve fine-scale hydraulic properties, thereby limiting the ability of RF to robustly learn relationships for processes dominated by vertical fluxes.

3.3 | Impact of SR on RF Regionalisation Model

Figure 8 presents the average KGE for all nested catchments across different discretisation levels. Figure 8a,b depicts the mean KGE for simulations of the Chaudière sub-catchment, while panels c and d show the same for the Yamaska sub-catchments. Subplots e and f illustrate the mean KGE for the du Nord sub-catchments, all evaluated at 3- and 24-h time-steps, respectively. At the 24-h time-step, the average efficiencies demonstrate a consistent improvement as the number of distributed parameters increases (Figure 10a,c,e). The Yamaska sub-catchment (Figure 8c) exhibits the greatest improvement, with a 12% increase in efficiency, followed by the du Nord sub-catchment (6%, Figure 10e) and the Chaudière sub-catchment (4%, Figure 8a). This is expected, as Yamaska is the most heterogeneous catchment, with a clear transition from forest to agricultural lands (refer to Section 4.1 for more details). Therefore, simulations with fully distributed parameters more effectively capture this heterogeneity, resulting in improved model performance for the Yamaska catchment (see Section 4.2 for more details). On the other hand, no clear pattern is evident in the 3-h time-step simulations, even for the Yamaska catchment. This may be due to the fact that at higher temporal resolution the model already captures hydrological processes accurately, reducing the dependence on additional spatial detail to improve performance and leaving little room for further improvement by increasing the spatial discretisation of parameters.

4 | Discussion

This study utilised RF to estimate process-based distributed model parameters in ungauged basins, aiming to evaluate RF's effectiveness in regionalising parameters across various spatio-temporal scales. The results demonstrated that RF outperformed conventional approaches for regionalisation, aligning with the findings of Saadi et al. (2019). They employed a larger dataset (> 2000 catchments) to regionalise four parameters of the conceptual GR4J model, showing only a slight improvement with RF over benchmark methods. This study demonstrated a stronger performance of RF compared to other methods, highlighting RF's capacity to handle much smaller datasets effectively (Zhao et al. 2022).

The influence of temporal resolution on regionalisation has rarely been explored, particularly in process-based distributed models (Razavi and Coulibaly 2013), largely due to data scarcity and the challenge of achieving good modelling accuracy at sub-daily time steps. Contrary to previous findings that higher temporal resolution leads to greater regionalisation losses (Saadi et al. 2019; Besaw et al. 2010), this study found that parameters calibrated at sub-daily time steps reduce regionalisation

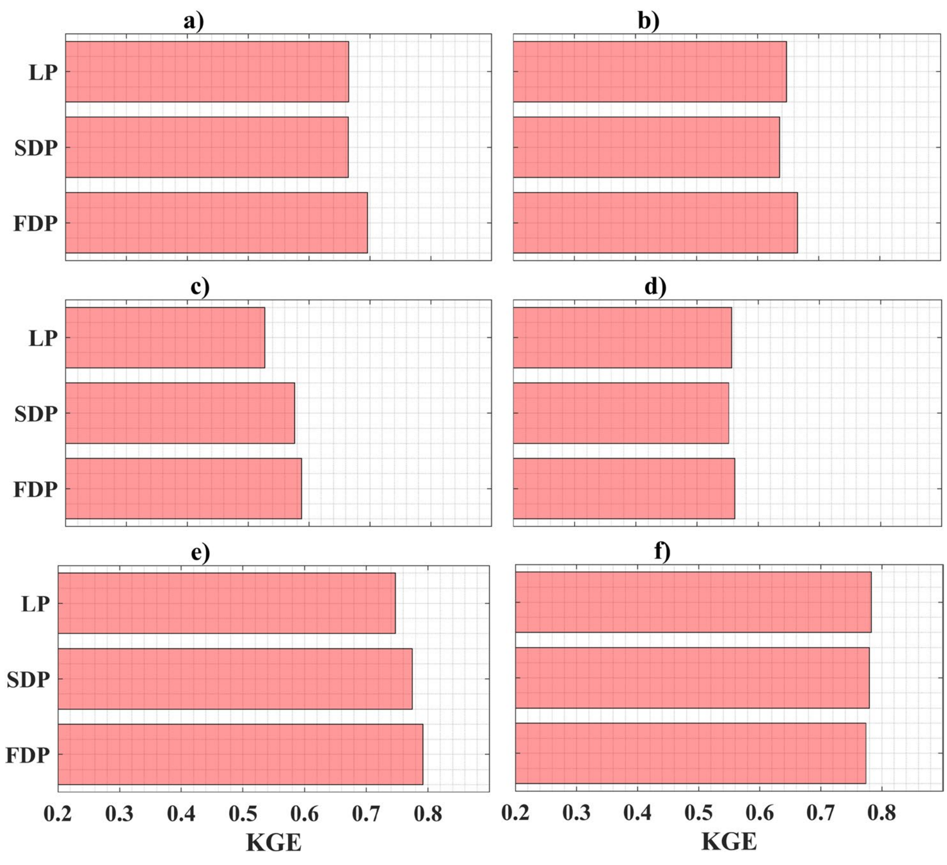

FIGURE 8 | Average efficiency of RF simulations for each nested catchment at different levels of parameter discretisation: fully distributed parameters (FDP), semi-distributed parameters (SDP) and lumped parameters (LP). (a) Chaudière-24 h, (b) Chaudière-3 h, (c) Yamaska-24 h, (d) Yamaska-3 h, (e) du Nord-24 h and (f) du Nord-3 h.

loss. This may be attributed to the fact that parameters of more process-based hydrological models are more physically meaningful and therefore more sensitive to temporal resolution (Samaniego et al. 2010, 2017).

In terms of spatial resolution, the findings of this study align with research on catchment heterogeneity (Markhali et al. 2022; Singh and Woolhiser 2002; Salvadore et al. 2015). When catchments exhibit greater heterogeneity, using a discretised parameter distribution improves regionalisation performance. In the following sections, we analyse the relative importance and spatial distribution of two selected parameters related to snow melt/accumulation and the vertical water budget, emphasising the relationship between CDs, modelling parameters and the impact of spatial discretisation on regionalisation performance.

4.1 | Relative Importance of CDs

Figure 9 demonstrates the relative importance of CDs in the estimation of two selected parameters (L1 and MTN; Table 2) for 24- and 3-h time-steps. These parameters are representative of the vertical budget and snow accumulation/melt in Hydrotel, respectively, which are major hydrological processes in snow-dominated catchments. As seen, no distinct group of features (i.e., meteorological CDs, topographic CDs, soil type CDs, land use CDs) controls parameter estimation. For both parameters, the groups of meteorological (first six, AI to TMIN) and topographic (last 10, SA to LON) CDs have similar weights. This is expected as these two groups should be correlated with one another (Merz and Blöschl 2004; Merz et al. 2020). Furthermore, land-cover features (WR to WL) are important, particularly for approximating MTN, which is involved in snow accumulation and melt. The soil type (SN to CL) shows a lesser degree of importance to determine the parameters. This might be due to the rather uniform soil texture of the study area. The most important individual features are elevation-related input variables (MEL, CVEL, MASP), mean annual precipitation and temperature (PRM, TMM) and the percentage of coniferous or deciduous forests (CF, DF).

No systematic pattern relating CDs to parameters can be found. Indeed, it is possible that there exists more than one good parameter set for each catchment, in particular when model calibration is based on a single type of observation (streamflow) which might not be enough to constrain the optimisation problem. If so, the parameter set used for a given catchment in the training of the RF model is one possibility from a sample of many other possibilities, not to mention the possible interrelations between the model parameters (Beven 2010). This could induce uncertainty in the training of the RF model. Furthermore, the multivariate nature of calibration may cause an inter-correlation of parameters (Merz and Blöschl 2004; Merz et al. 2020). Therefore, it is expected that a CD that controls a specific parameter affects

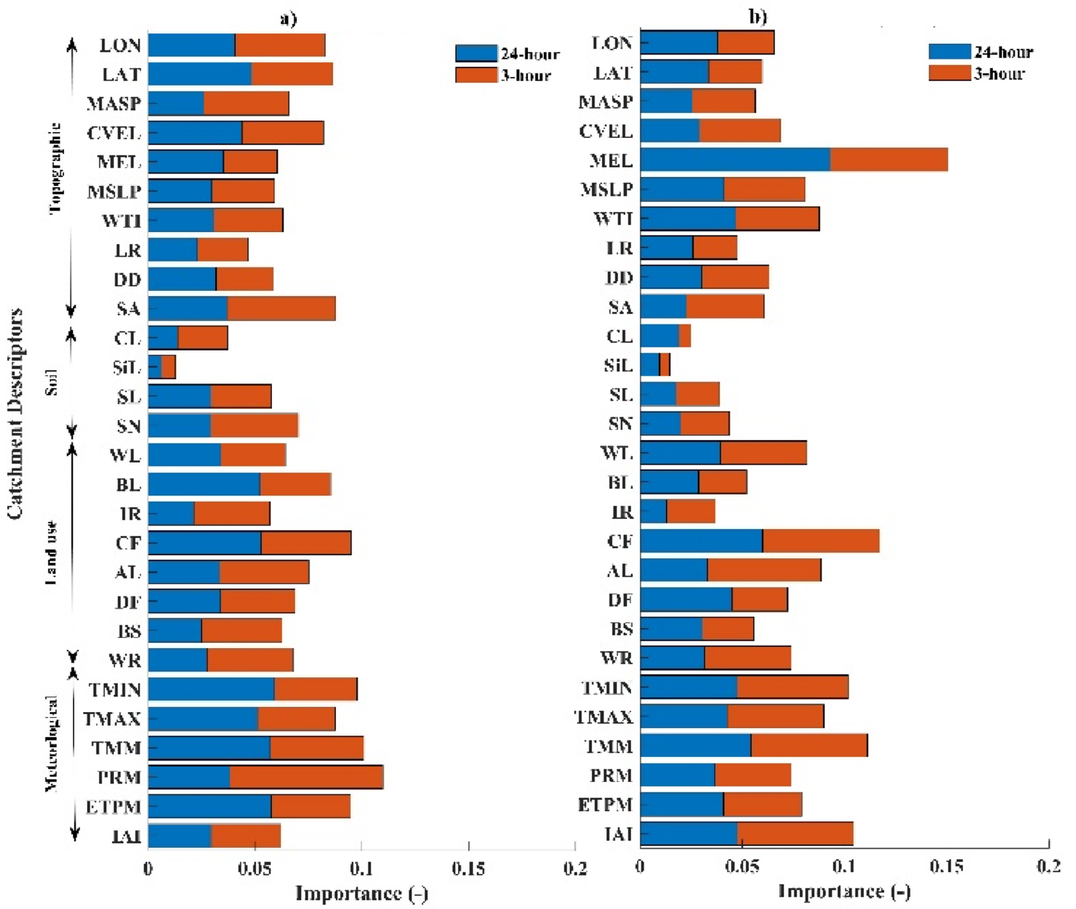

FIGURE 9 | Relative importance of predictor (catchment descriptor) features in parameters approximated by RF for simulations with 24- and 3-h time-steps. (a) Parameter (L1) and (b) parameter (MTN).

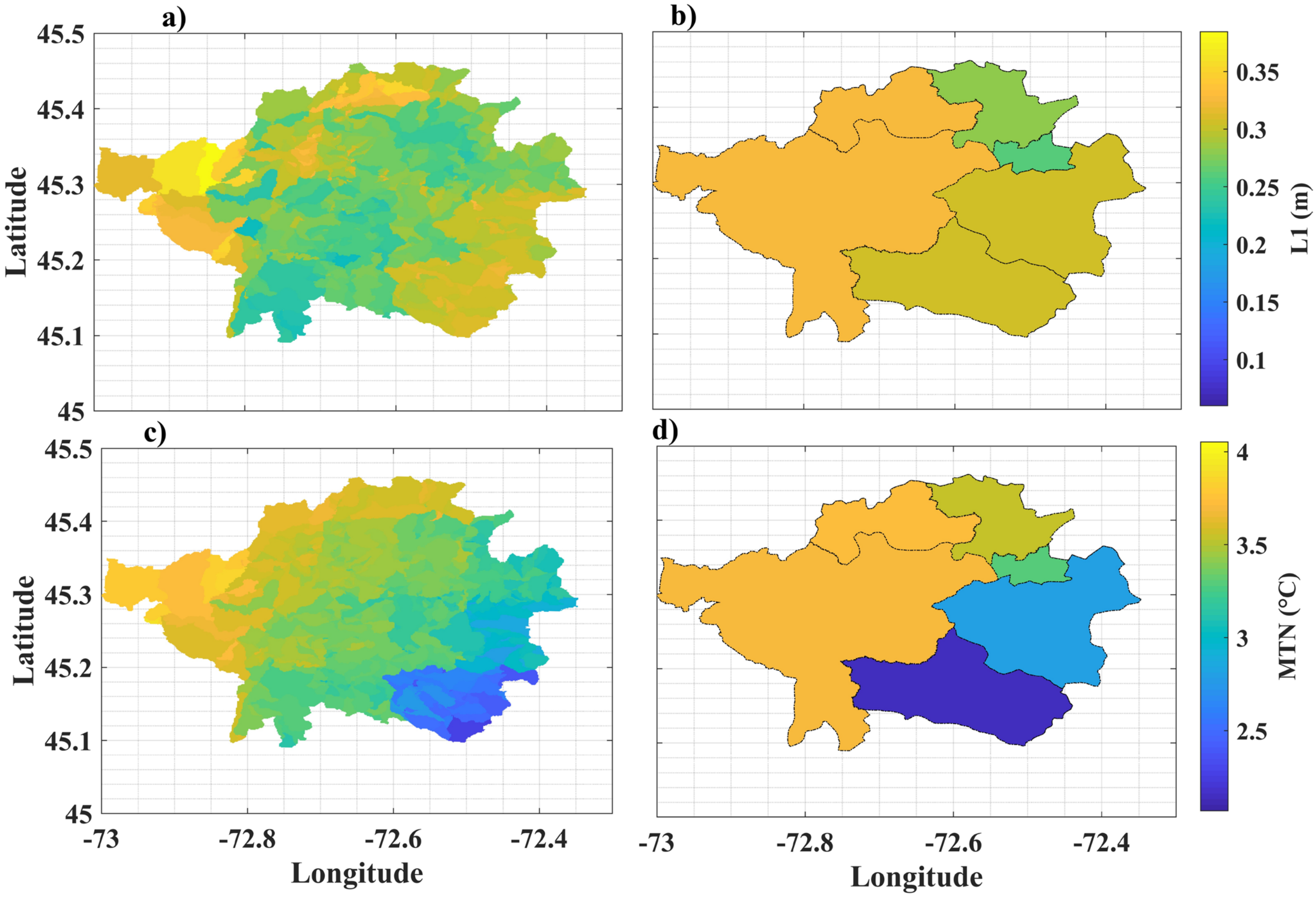

FIGURE 10 | Distribution of RF approximated parameters (L1 and MTN) in 24-h time-step across the Yamaska catchment at RHHU (sub-plots a and c) and sub-catchment (sub-plots b and d) discretisation levels. (a) L1 distribution RHHU level, (b) L1 distribution sub-catchment level, (c) MTN distribution RHHU level (d) MTN distribution sub-catchment level.

many unrelated parameters due to the correlation between parameters. These observations support the conclusion that it is difficult or improbable to constrain the parameters to an a priori to be used as a transfer function because some information may remain hidden to the system and reduce the performance of the regionalisation model.

4.2 | Impact of Parameter Distribution

Figure 10 illustrates the distribution of RF-estimated parameters at the 24-h time-step (L1 and MTN) across the Yamaska catchment at the RHHU and sub-catchment levels. Comparing Figure 10a,c with Figure 5a, there is an impact of the presence of agricultural lands on the distribution of the parameters. Notably, the distinct increase in the magnitude of L1 at the western side of the catchment (Figure 10a) matches the agricultural area in Figure 5a, showing a direct link between land cover and parameter distribution at the RHHU discretisation level. Yet the effect of agricultural land on the L1 distribution at the sub-catchment level (Figure 10b) cannot be directly observed due to the lower resolution of the simulation than that for RHHU level, which leads to the spatially averaged values for L1 over the sub-catchments. Still, a link can be observed with elevation, with lower values of L1 occurring in areas with higher elevations in the eastern part of the catchment.

A similar pattern is visible regarding the MTN spatial distribution: a smooth transition from west to east at the RHHU level (Figure 10c), in contrast with an abrupt transition of MTN in the eastern side of the catchment at the sub-catchment level (Figure 10d). It seems that such a level of parameter discretisation helped improve the representation of spatial heterogeneity leading to a better performance of the regionalisation model for Yamaska as seen in Figure 8c.

5 | Conclusion

In this paper, we focused on developing a regionalisation technique for a partially physics-based distributed hydrological model (i.e., Hydrotel) using RF. We laid out three hypotheses to investigate the efficiency of RF regionalisation technique compared to more conventional approaches, investigate the impact of temporal resolution of the modelling on the efficiency of regionalisation technique, and the added information provided by more spatially representative CDs. We designed multiple experiments to verify and/or reject the hypotheses. The summary and conclusion about each hypothesis are in the following:

1. RF shows a strong performance compared to other conventional regionalisation models (Figure 6). The first hypothesis is therefore verified for this paper.

2. Refining the time-step of the modelling shows an advantage in the calculation of the loss of the performance from calibration to regionalisation. For 3-h time-step, the calculated loss of efficiency is less than that of 24-h time-step (Figure 7). The sub-daily time-step outperforms the daily time-step, which is in accordance with the second hypothesis.
3. Figure 8 showed that using spatially refined CDs at a daily time-step could improve the performance of the modelling in a nested catchment, depending on the degree to which the catchment is heterogeneous, i.e., the improvement is higher for more heterogeneous catchments. However, a clear pattern regarding the improvement of simulations with refined CDs with a 3-h time-step has not been found. Hence, Hypothesis III cannot be verified based on the results provided.

A major limit of this study is the rather small sample size of catchments to compensate for the computational time required for calibration of distributed hydrological models. One path to continue is therefore to use a larger sample size and examine the hypotheses laid out here. Regarding the CD-parameter relationship and the importance of the scale in this context, we suggest that a more physically representative distributed model be used. Some of the hydrological processes that are represented by Hydrotel are empirical and/or conceptual.

Funding

This work was supported by PSR-SIIRI (Research Support Program, Support for International Research and Innovation Initiatives component) of the Ministry of Economy and Innovation of Quebec.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

Abdulla, F. A., and D. P. Lettenmaier. 1997. "Development of Regional Parameter Estimation Equations for a Macroscale Hydrologic Model." *Journal of Hydrology* 197, no. 1–4: 230–257. https://doi.org/10.1016/S0022-1694(96)03262-3.

Abrahart, R. J., F. Anctil, P. Coulibaly, et al. 2012. "Two Decades of Anarchy? Emerging Themes and Outstanding Challenges for Neural Network River Forecasting." *Progress in Physical Geography: Earth and Environment* 36, no. 4: 480–513. https://doi.org/10.1177/0309133312444943.

Abramowitz, G. 2005. "Towards a Benchmark for Land Surface Models." *Geophysical Research Letters* 32, no. 22. https://doi.org/10.1029/2005GL024419.

Addor, N., G. Nearing, C. Prieto, A. J. Newman, N. Le Vine, and M. P. Clark. 2018. "A Ranking of Hydrological Signatures Based on Their Predictability in Space." *Water Resources Research* 54, no. 11: 8792–8812. https://doi.org/10.1029/2018WR022606.

Addor, N., O. Rössler, N. Köplin, M. Huss, R. Weingartner, and J. Seibert. 2014. "Robust Changes and Sources of Uncertainty in the Projected Hydrological Regimes of Swiss Catchments." *Water Resources Research* 50, no. 10: 7541–7562. https://doi.org/10.1002/2014WR015549.

Almagro, A., A. A. Meira Neto, N. Vergopolan, T. Roy, P. A. Troch, and P. T. S. Oliveira. 2024. "The Drivers of Hydrologic Behavior in Brazil: Insights From a Catchment Classification." *Water Resources Research* 60, no. 8: e2024WR037212. https://doi.org/10.1029/2024WR037212.

Anderson, G. J., D. D. Lucas, and C. Bonfils. 2018. "Uncertainty Analysis of Simulations of the Turn-of-the-Century Drought in the Western United States." *Journal of Geophysical Research: Atmospheres* 123, no. 23: 13219–13237. https://doi.org/10.1029/2017JD027824.

Arsenault, R., M. Breton-Dufour, A. Poulin, G. Dallaire, and R. Romero-Lopez. 2019. "Streamflow Prediction in Ungauged Basins: Analysis of Regionalization Methods in a Hydrologically Heterogeneous Region of Mexico." *Hydrological Sciences Journal* 64, no. 11: 1297–1311. https://doi.org/10.1080/02626667.2019.1639716.

Arsenault, R., and F. P. Brissette. 2014. "Continuous Streamflow Prediction in Ungauged Basins: The Effects of Equifinality and Parameter Set Selection on Uncertainty in Regionalization Approaches." *Water Resources Research* 50, no. 7: 6135–6153. https://doi.org/10.1002/2013WR014898.

Arsenault, R., and F. Brissette. 2016. "Multi-Model Averaging for Continuous Streamflow Prediction in Ungauged Basins." *Hydrological Sciences Journal* 61, no. 13: 2443–2454. https://doi.org/10.1080/02626667.2015.1117088.

Bachmair, S., C. Svensson, J. Hannaford, L. J. Barker, and K. Stahl. 2016. "A Quantitative Analysis to Objectively Appraise Drought Indicators and Model Drought Impacts." *Hydrology and Earth System Sciences* 20, no. 7: 2589–2609.

Balázs, B., T. Bíró, G. Dyke, S. K. Singh, and S. Szabó. 2018. "Extracting Water-Related Features Using Reflectance Data and Principal Component Analysis of Landsat Images." *Hydrological Sciences Journal* 63, no. 2: 269–284. https://doi.org/10.1080/02626667.2018.1425802.

Bastola, S., and V. Misra. 2013. "Sensitivity of Hydrological Simulations of Southeastern United States Watersheds to Temporal Aggregation of Rainfall." *Journal of Hydrometeorology* 14: 1334–1344. https://doi.org/10.1175/JHM-D-12-096.1.

Besaw, L. E., M. R. Donna, R. B. Paul, and R. H. William. 2010. "Advances in Ungauged Streamflow Prediction Using Artificial Neural Networks." *Journal of Hydrology* 386, no. 1–4: 27–37. https://doi.org/10.1016/j.jhydrol.2010.02.037.

Best, M. J., G. Abramowitz, H. R. Johnson, et al. 2015. "The Plumbing of Land Surface Models: Benchmarking Model Performance." *Journal of Hydrometeorology* 16, no. 3: 1425–1442. https://doi.org/10.1175/JHM-D-14-0158.1.

Beven, K. 2010. "A Manifesto for the Equifinality Thesis." *Journal of Hydrology* 320, no. 1–2: 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007.

Bishop, C. M., and N. M. Nasrabadi. 2006. *Pattern Recognition and Machine Learning*. Springer.

Blöschl, G. 2006. "Rainfall-Runoff Modeling of Ungauged Catchments." In *Encyclopaedia of Hydrological Sciences*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470848944.hsa140.

Blöschl, G., C. Reszler, and J. Komma. 2008. "A Spatially Distributed Flash Flood Forecasting Model." *Environmental Modelling & Software* 23, no. 4: 464–478. https://doi.org/10.1016/j.envsoft.2007.06.010.

Boucher, M.-A., J. Quilty, and J. Adamowski. 2020. "Data Assimilation for Streamflow Forecasting Using Extreme Learning Machines and Multilayer Perceptrons." *Water Resources Research* 56, no. 6: e2019WR026226. https://doi.org/10.1029/2019WR026226.

Boulesteix, A.-L., J. Silke, K. Jochen, and R. K. Inke. 2012. "Overview of Random Forest Methodology and Practical Guidance With Emphasis on Computational Biology and Bioinformatics." *WIREs Data Mining and Knowledge Discovery* 2, no. 6: 493–507. https://doi.org/10.1002/widm.1072.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45, no. 1: 5–32. https://doi.org/10.1023/A:1010933404324.

Burn, D. H., and D. B. Boorman. 1993. "Estimation of Hydrological Parameters at Ungauged Catchments." *Journal of Hydrology* 143, no. 3–4: 429–454. https://doi.org/10.1016/0022-1694(93)90203-L.

Duan, Q., J. Schaake, V. Andréassian, et al. 2006. "Model Parameter Estimation Experiment (MOPEX): An Overview of Science Strategy and Major Results From the Second and Third Workshops." *Journal of Hydrology* 320, no. 1–2: 3–17. https://doi.org/10.1016/j.jhydrol.2005.07.031.

El Baida, M., F. Boushaba, M. Chourak, M. Hosni, H. Sabar, and T. Zahaf. 2024. "Flood Risk Decomposed: Optimized Machine Learning Hazard Mapping and Multi-Criteria Vulnerability Analysis in the City of Zaio, Morocco." *Journal of African Earth Sciences* 220: 105431. https://doi.org/10.1016/j.jafrearsci.2024.105431.

Farfán, J. F., and L. Cea. 2023. "Regional Streamflow Prediction in Northwest Spain: A Comparative Analysis of Regionalisation Schemes." *Journal of Hydrology: Regional Studies* 47: 101427. https://doi.org/10.1016/j.ejrh.2023.101427.

Fatichi, S., E. R. Vivoni, F. L. Ogden, et al. 2016. "An Overview of Current Applications, Challenges, and Future Trends in Distributed Process-Based Models in Hydrology." *Journal of Hydrology* 537: 45–60. https://doi.org/10.1016/j.jhydrol.2016.03.026.

Feng, D., H. Beck, J. de Bruijn, et al. 2024. "Deep Dive Into Hydrologic Simulations at Global Scale: Harnessing the Power of Deep Learning and Physics-Informed Differentiable Models (δHBV-globe1.0-hydroDL)." *Geoscientific Model Development* 17, no. 18: 7181–7198. https://doi.org/10.5194/gmd-17-7181-2024.

Fortin, J.-P., R. Turcotte, S. Massicotte, R. Moussa, J. Fitzback, and J.-P. Villeneuve. 2001a. "Distributed Watershed Model Compatible With Remote Sensing and GIS Data. I: Description of Model." *Journal of Hydrologic Engineering* 6, no. 2: 91–99. https://doi.org/10.1061/(ASCE)1084-0699(2001)6:2(91).

Fortin, J.-P., R. Turcotte, S. Massicotte, R. Moussa, J. Fitzback, and J.-P. Villeneuve. 2001b. "Distributed Watershed Model Compatible With Remote Sensing and GIS Data. II: Application to Chaudière Watershed." *Journal of Hydrologic Engineering* 6, no. 2: 100–108. https://doi.org/10.1061/(ASCE)1084-0699(2001)6:2(100).

Garambois, P. A., H. Roux, K. Larnier, D. Labat, and D. Dartus. 2015. "Parameter Regionalization for a Process-Oriented Distributed Model Dedicated to Flash Floods." *Journal of Hydrology* 525: 383–399. https://doi.org/10.1016/j.jhydrol.2015.03.052.

Grinsztajn, L., E. Oyallon, and G. Varoquaux. 2022. "Why Do Tree-Based Models Still Outperform Deep Learning on Tabular Data?" *Journal of Machine Learning Research* 23, no. 206: 1–63.

Guo, Y., Y. Zhang, L. Zhang, and Z. Wang. 2021. "Regionalization of Hydrological Modeling for Predicting Streamflow in Ungauged Catchments: A Comprehensive Review." *WIREs Water* 8, no. 1: e1487. https://doi.org/10.1002/wat2.1487.

Hameed, M. M., S. F. Mohd Razali, W. H. M. Wan Mohtar, and Z. M. Yaseen. 2024. "Examining Optimized Machine Learning Models for Accurate Multi-Month Drought Forecasting: A Representative Case Study in the USA." *Environmental Science and Pollution Research* 31, no. 39: 52060–52085. https://doi.org/10.1007/s11356-024-34500-6.

He, M., S. Jiang, L. Ren, et al. 2024. "Streamflow Prediction in Ungauged Catchments Through Use of Catchment Classification and Deep Learning." *Journal of Hydrology* 639: 131638. https://doi.org/10.1016/j.jhydrol.2024.131638.

He, Y., A. Bárdossy, and E. Zehe. 2011. "A Review of Regionalisation for Continuous Streamflow Simulation." *Hydrology and Earth System Sciences* 15, no. 11: 3539–3553. https://doi.org/10.5194/hess-15-3539-2011.

Hrachowitz, M., H. H. G. Savenije, G. Blöschl, et al. 2013. "A Decade of Predictions in Ungauged Basins (PUB)—A Review." *Hydrological Sciences Journal* 58, no. 6: 1198–1255. https://doi.org/10.1080/02626667.2013.803183.

Hsu, K. l., H. V. Gupta, and S. Sorooshian. 1995. "Artificial Neural Network Modeling of the Rainfall-Runoff Process." *Water Resources Research* 31, no. 10: 2517–2530. https://doi.org/10.1029/95WR01955.

Huot, P.-L., A. Poulin, C. Audet, and S. Alarie. 2019. "A Hybrid Optimization Approach for Efficient Calibration of Computationally Intensive Hydrological Models." *Hydrological Sciences Journal* 64, no. 10: 1204–1222. https://doi.org/10.1080/02626667.2019.1624922.

Imani, A., J. D. Smith, M. K. Lee, and Q. Zhang. 2025. "Random Forest Modelling of Catchment Controls: Resisting Overfitting, Variable Importance Diagnostics, and Hydrological Prediction." *Hydrological Processes* 39, no. 4: 512–528. https://doi.org/10.1002/hyp.12345.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. *An Introduction to Statistical Learning*. Springer.

Kanishka, G., and T. Eldho. 2017. "Watershed Classification Using Isomap Technique and Hydrometeorological Attributes." *Journal of Hydrologic Engineering* 22, no. 10: 04017040. https://doi.org/10.1061/(asce)he.1943-5584.0001562.

Kasiviswanathan, S. P., and H. Jin. 2016. "Efficient Private Empirical Risk Minimization for High-Dimensional Learning." In *Proceedings of the 33rd International Conference on Machine Learning. Proceedings of Machine Learning Research*, vol. 48, 488–497. PMLR (Proceedings of Machine Learning Research). https://proceedings.mlr.press/v48/kasiviswanathan16.html.

Kirchner, K., K.-H. Tölle, and J. Krieter. 2006. "Optimisation of the Decision Tree Technique Applied to Simulated Sow Herd Datasets." *Computers and Electronics in Agriculture* 50, no. 1: 15–24. https://doi.org/10.1016/j.compag.2005.07.002.

Kling, H., H. V. Gupta, and S. Y. Yang. 2012. "Improving a Novel Efficiency Criterion for Hydrological Models." *Hydrological Processes* 26, no. 12: 2799–2805. https://doi.org/10.1002/hyp.8597.

Kratzert, F., D. Klotz, S. Hochreiter, and G. S. Nearing. 2021. "A Note on Leveraging Synergy in Multiple Meteorological Datasets With Deep Learning for Rainfall-Runoff Modeling." *Hydrology and Earth System Sciences* 25: 2685–2703.

Kratzert, F., D. Klotz, G. Shalev, G. Klambauer, S. Hochreiter, and G. Nearing. 2019. "Towards Learning Universal, Regional, and Local Hydrological Behaviors via Machine Learning Applied to Large-Sample Datasets." *Hydrology and Earth System Sciences* 23, no. 12: 5089–5110. https://doi.org/10.5194/hess-23-5089-2019.

Kumar, R., A. Kalnawat, A. M. Pawar, et al. 2024. "Transparency in Algorithmic Decision-Making: Interpretable Models for Ethical Accountability." *E3S Web of Conferences* 491: 02041. https://doi.org/10.1051/e3sconf/202449102041.

Kumar, R., L. Samaniego, and S. Attinger. 2013. "Implications of Distributed Hydrologic Model Parameterization on Water Fluxes at Multiple Scales and Locations." *Water Resources Research* 49, no. 1: 360–379. https://doi.org/10.1029/2012WR012195.

Lei, H., H. Li, and W. Hu. 2024. "Enhancing the Streamflow Simulation of a Process-Based Hydrological Model Using Machine Learning and Multi-Source Data." *Ecological Informatics* 82: 102755. https://doi.org/10.1016/j.ecoinf.2024.102755.

Lu, X., L. Sun, Y. Zhang, et al. 2024. "Predicting cd Accumulation in Crops and Identifying Nonlinear Effects of Multiple Environmental Factors Based on Machine Learning Models." *Science of the Total Environment* 951: 175787. https://doi.org/10.1016/j.scitotenv.2024.175787.

Ma, K., D. He, S. Liu, X. Ji, Y. Li, and H. Jiang. 2024. "Novel Time-Lag Informed Deep Learning Framework for Enhanced Streamflow

Prediction and Flood Early Warning in Large-Scale Catchments." *Journal of Hydrology* 631: 130841. https://doi.org/10.1016/j.jhydrol.2024.130841.

Markhali, S. P., A. Poulin, and M.-A. Boucher. 2022. "Spatio-Temporal Discretization Uncertainty of Distributed Hydrological Models." *Hydrological Processes* 36, no. 6: e14635. https://doi.org/10.1002/hyp.14635.

Martel, J.-L., F. Brissette, and A. Poulin. 2020. "Impact of the Spatial Density of Weather Stations on the Performance of Distributed and Lumped Hydrological Models." *Canadian Water Resources Journal = Revue Canadienne Des Ressources Hydriques* 45, no. 2: 158–171. https://doi.org/10.5194/gmd-12-2501-2019.

Martin, N. 2023. "Dynamic Assimilation of Deep Learning Predictions to a Process-Based Water Budget." *Hydrology* 10, no. 6: 129. https://doi.org/10.3390/hydrology10060129.

Mehrvand, S., M.-A. Boucher, K. Kornelsen, and A. Amani. 2023. "Comparing Three Machine Learning Algorithms With Existing Methods for Natural Streamflow Estimation." *Hydrological Sciences Journal* 69, no. 1: 79–94. https://doi.org/10.1080/02626667.2023.2273402.

Merz, R., and G. Blöschl. 2004. "Regionalisation of Catchment Model Parameters." *Journal of Hydrology* 287, no. 1–4: 95–123. https://doi.org/10.1016/j.jhydrol.2003.09.028.

Merz, R., L. Tarasova, and S. Basso. 2020. "Parameter's Controls of Distributed Catchment Models—How Much Information Is in Conventional Catchment Descriptors?" *Water Resources Research* 56, no. 2: e2019WR026008. https://doi.org/10.1029/2019WR026008.

Muñoz, P., J. Orellana-Alvear, P. Willems, and R. Célleri. 2018. "Flash-Flood Forecasting in an Andean Mountain Catchment—Development of a Step-Wise Methodology Based on the Random Forest Algorithm." *Water* 10, no. 11: 1519. https://doi.org/10.3390/w10111519.

Nearing, G. S., F. Kratzert, A. K. Sampson, et al. 2021. "What Role Does Hydrological Science Play in the Age of Machine Learning?" *Water Resources Research* 57, no. 3: e2020WR028091. https://doi.org/10.1029/2020WR028091.

Nearing, G. S., D. M. Mocko, C. D. Peters-Lidard, S. V. Kumar, and Y. Xia. 2016. "Benchmarking NLDAS-2 Soil Moisture and Evapotranspiration to Separate Uncertainty Contributions." *Journal of Hydrometeorology* 17, no. 3: 745–759.

Odusanya, A. E., K. Schulz, and B. Mehdi-Schulz. 2022. "Using a Regionalisation Approach to Evaluate Streamflow Simulated by an Ecohydrological Model Calibrated With Global Land Surface Evaporation From Remote Sensing." *Journal of Hydrology: Regional Studies* 40: 101042. https://doi.org/10.1016/j.ejrh.2022.101042.

Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine. 2008. "Spatial Proximity, Physical Similarity, Regression and Ungaged Catchments: A Comparison of Regionalization Approaches Based on 913 French Catchments." *Water Resources Research* 44, no. 3: W03413. https://doi.org/10.1029/2007WR006240.

Papacharalampous, G., and H. Tyralis. 2022. "A Review of Machine Learning Concepts and Methods for Addressing Challenges in Probabilistic Hydrological Post-Processing and Forecasting." *Frontiers in Water* 4: 961954. https://doi.org/10.3389/frwa.2022.961954.

Papageorgaki, I., and I. Nalbantis. 2016. "Classification of Drainage Basins Based on Readily Available Information." *Water Resources Management* 30, no. 15: 5559–5574. https://doi.org/10.1007/s11269-016-1410-y.

Parajka, J., A. Viglione, M. Rogger, et al. 2013. "Comparative Assessment of Predictions in Ungauged Basins—Part 1: Runoff Hydrograph Studies." *Hydrology and Earth System Sciences* 17, no. 5: 1783–1795. https://doi.org/10.5194/hess-17-1783-2013.

Probst, P., M. N. Wright, and A.-L. Boulesteix. 2019. "Hyperparameters and Tuning Strategies for Random Forest." *WIREs Data Mining and Knowledge Discovery* 9, no. 3: e1301.

Rahmani, F., C. Shen, S. Oliver, K. Lawson, and A. Appling. 2021. "Deep Learning Approaches for Improving Prediction of Daily Stream Temperature in Data-Scarce, Unmonitored, and Dammed Basins." *Hydrological Processes*.

Rakovec, O., R. Kumar, J. Mai, et al. 2018. "Multiscale and Multivariate Evaluation of Water Fluxes and States Over European River Basins." *Journal of Hydrometeorology* 19, no. 11: 287–307. https://doi.org/10.1175/JHM-D-15-0054.1.

Randrianasolo, A., M. Ramos, and V. Andréassian. 2011. "Hydrological Ensemble Forecasting at Ungauged Basins: Using Neighbour Catchments for Model Setup and Updating." *Advances in Geosciences* 29: 1–11. https://doi.org/10.5194/adgeo-29-1-2011.

Razavi, T., and P. Coulibaly. 2013. "Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods." *Journal of Hydrologic Engineering* 18, no. 8: 958–975. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000690.

Razavi, T., and P. Coulibaly. 2017. "An Evaluation of Regionalization and Watershed Classification Schemes for Continuous Daily Streamflow Prediction in Ungauged Watersheds." *Canadian Water Resources Journal = Revue Canadienne Des Ressources Hydriques* 42, no. 1: 2–20. https://doi.org/10.1080/07011784.2016.1184590.

Saadi, M., L. Oudin, and P. Ribstein. 2019. "Random Forest Ability in Regionalizing Hourly Hydrological Model Parameters." *Water* 11, no. 8: 1540. https://doi.org/10.3390/w11081540.

Salvadore, E., J. Bronders, and O. Batelaan. 2015. "Hydrological Modelling of Urbanized Catchments: A Review and Future Directions." *Journal of Hydrology* 529 (octobre): 62–81. https://doi.org/10.1016/j.jhydrol.2015.06.028.

Samaniego, L., R. Kumar, and S. Attinger. 2010. "Multiscale Parameter Regionalization of a Grid-Based Hydrologic Model at the Mesoscale." *Water Resources Research* 46, no. 5: W05523. https://doi.org/10.1029/2008WR007327.

Samaniego, L., R. Kumar, S. Thober, et al. 2017. "Toward Seamless Hydrologic Predictions Across Spatial Scales." *Hydrology and Earth System Sciences* 21, no. 9: 4323–4346. https://doi.org/10.5194/hess-21-4323-2017.

Samek, W., and K.-R. Müller. 2019. "Towards Explainable Artificial Intelligence." In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, edited by W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, and K.-R. Müller, 5–22. Springer International Publishing. https://doi.org/10.1007/978-3-030-28954-6_1.

Samuel, J., P. Coulibaly, and R. A. Metcalfe. 2011. "Estimation of Continuous Streamflow in Ontario Ungauged Basins: Comparison of Regionalization Methods." *Journal of Hydrologic Engineering* 16, no. 5: 447–459. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000338.

Shen, C. 2018. "A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists." *Water Resources Research* 54, no. 11: 8558–8593. https://doi.org/10.1029/2018WR022643.

Shen, C., A. P. Appling, P. Gentine, et al. 2023. "Differentiable Modelling to Unify Machine Learning and Physical Models for Geosciences." *Nature Reviews Earth and Environment* 4, no. 8: 552–567. https://doi.org/10.1038/s43017-023-00450-9.

Singh, N. M., and T. T. Devi. 2022. "Regionalization Methods in Ungauged Catchments for Flow Prediction: Review and Its Recent Developments." *Arabian Journal of Geosciences* 15, no. 11: 1019. https://doi.org/10.1007/s12517-022-10287-z.

Singh, S. K., H. McMillan, A. Bárdossy, and C. Fateh. 2016. "Nonparametric Catchment Clustering Using the Data Depth Function."

Hydrological Sciences Journal 61, no. 15: 2649–2667. https://doi.org/10.1080/02626667.2016.1168927.

Singh, V. P., and D. A. Woolhiser. 2002. "Mathematical Modeling of Watershed Hydrology." *Journal of Hydrologic Engineering* 7, no. 4: 270–292.

Sivapalan, M. 2003. "Prediction in Ungauged Basins: A Grand Challenge for Theoretical Hydrology." *Hydrological Processes* 17, no. 15: 3163–3170. https://doi.org/10.1002/hyp.5155.

Snelder, T. H., N. Lamouroux, J. R. Leathwick, H. Pella, E. Sauquet, and U. Shankar. 2009. "Predictive Mapping of the Natural Flow Regimes of France." *Journal of Hydrology* 373, no. 1–2: 57–67. https://doi.org/10.1016/j.jhydrol.2009.04.011.

Sultana, Z., T. Sieg, P. Kellermann, M. Müller, and H. Kreibich. 2018. "Assessment of Business Interruption of Flood-Affected Companies Using Random Forests." *Water* 10, no. 8: 1049. https://doi.org/10.3390/w10081049.

Swain, J. B., and K. C. Patra. 2017a. "Streamflow Estimation in Ungauged Catchments Using Regional Flow Duration Curve: Comparative Study." *Journal of Hydrologic Engineering* 22, no. 7: 04017010. https://doi.org/10.1002/wat2.1487.

Swain, J. B., and K. C. Patra. 2017b. "Streamflow Estimation in Ungauged Catchments Using Regionalization Techniques." *Journal of Hydrology* 554: 420–433. https://doi.org/10.1016/j.jhydrol.2017.08.054.

Talbot, F., J. D. Sylvain, G. Drolet, et al. 2025. "Assessing Hydroclimatic Impacts of Climate Change in Snowy Catchments Using a Physically Based Hydrological Model." *Journal of Hydrology: Regional Studies* 59: 102453.

Tan, C., F. Sun, L. Xu, and H. Liu. 2021. "A Benchmarking Study of Tree-Based and Deep Learning Methods for Structured Data." *Neurocomputing* 463: 96–107.

Teutschbein, C., T. Grabs, H. Laudon, R. H. Karlsen, and K. Bishop. 2018. "Simulating Streamflow in Ungauged Basins Under a Changing Climate: The Importance of Landscape Characteristics." *Journal of Hydrology* 561: 160–178. https://doi.org/10.1016/j.jhydrol.2018.03.060.

Thober, S., M. Cuntz, M. Kelbling, R. Kumar, J. Mai, and L. Samaniego. 2019. "The Multiscale Routing Model mRM v1. 0: Simple River Routing at Resolutions From 1 to 50 km." *Geoscientific Model Development* 12, no. 6: 2501–2521. https://doi.org/10.5194/gmd-12-2501-2019.

Tolson, B. A., and C. A. Shoemaker. 2007. "Dynamically Dimensioned Search Algorithm for Computationally Efficient Watershed Model Calibration." *Water Resources Research* 43, no. 1: W01413. https://doi.org/10.1029/2005WR004723.

Tyralis, H., G. Papacharalampous, and A. Langousis. 2019. "A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources." *Water* 11, no. 5: 910. https://doi.org/10.3390/w11050910.

Valencia Giraldo, M. D. C., S. Ricard, and F. Anctil. 2023. "Assessment of the Potential Hydrological Impacts of Climate Change in Quebec—Canada, A Refined Neutral Approach." *Water* 15, no. 3: 584. https://doi.org/10.3390/w15030584.

Vandewiele, G., and A. Elias. 1995. "Monthly Water Balance of Ungauged Catchments Obtained by Geographical Regionalization." *Journal of Hydrology* 170, no. 1–4: 277–291. https://doi.org/10.1016/0022-1694(95)02681-E.

Velásquez, N., R. Mantilla, W. Krajewski, F. Quintero, and A. D. L. Zanchetta. 2022. "Identification and Regionalization of Streamflow Routing Parameters Using Machine Learning for the HLM Hydrological Model in Iowa." *Journal of Advances in Modeling Earth Systems* 14, no. 7: e2021MS002855. https://doi.org/10.1029/2021MS002855.

Vu, H. N., R. C. Deo, and Z. M. Yaseen. 2019. "Comparative Assessment of Machine Learning Algorithms for Daily Streamflow Simulation." *Water Resources Research* 55, no. 7: 5535–5554.

Wagener, T., and H. S. Wheater. 2006. "Parameter Estimation and Regionalization for Continuous Rainfall-Runoff Models Including Uncertainty." *Journal of Hydrology* 320, no. 1–2: 132–154. https://doi.org/10.1016/j.jhydrol.2005.07.015.

Wang, W., Y. Zhao, Y. Tu, R. Dong, Q. Ma, and C. Liu. 2023. "Research on Parameter Regionalization of Distributed Hydrological Model Based on Machine Learning." *Water* 15, no. 3: 518. https://doi.org/10.3390/w15030518.

Yang, X., J. Magnusson, J. Rizzi, and C.-Y. Xu. 2018. "Runoff Prediction in Ungauged Catchments in Norway: Comparison of Regionalization Approaches." *Hydrology Research* 49, no. 2: 487–505. https://doi.org/10.2166/nh.2017.071.

Zaerpour, M., A. Imani, S. Fatichi, and P. Molnar. 2025. "Agriculture's Impact on Water–Energy Balance Varies Across Climates." *Proceedings of the National Academy of Sciences* 122, no. 12: e2410521122. https://doi.org/10.1073/pnas.2410521122.

Zaerpour, M., S. M. Papalexiou, and A. Nazemi. 2021. "Informing Stochastic Streamflow Generation by Large-Scale Climate Indices at Single and Multiple Sites." *Advances in Water Resources* 156: 104037. https://doi.org/10.1016/j.advwatres.2021.104037.

Zhao, X., Y. Li, X. Wang, and Y. Zhang. 2022. "Research on the Application of Random Forest in Small-Sample Data Analysis." *Journal of Applied Statistics* 49, no. 5: 1173–1192.

Ziegler, A., and I. R. König. 2014. "Mining Data With Random Forests: Current Options for Real-World Applications." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 4, no. 1: 55–63. https://doi.org/10.1002/widm.1114.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** hyp70437-sup-0001-Figures.docx.