



Research papers

Impact of catchment and climate attributes on flood generating processes and their effect on flood statistics

Svenja Fischer^a, Markus Pahlow^{b,*}, Shailesh Kumar Singh^c

^a Hydrology and Environmental Hydraulics, Wageningen University & Research, the Netherlands

^b Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand

^c National Institute of Water and Atmospheric Research, Christchurch, New Zealand

ARTICLE INFO

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Felix Frances, Associate Editor

Keywords:

Flood type
Flood generation
Flood statistics
Classification and regression tree
Change-points

ABSTRACT

A major source of uncertainty in flood statistics are the different flood generation processes. These make the assumption of homogeneous samples questionable. To overcome this issue, a framework for assessing the influence of catchment and climate attributes on flood-generating processes and their effect on flood statistics has been developed and applied to 252 catchments in New Zealand. Mean daily discharge data time series with a length ranging from 20 to 81 years were used. Flood events were classified according to their hydrograph shape. Three types were considered based on the different forcing: heavy rainfall of short duration (termed R1), moderate rainfall of medium intensity and duration (R2), and long-duration rainfall sequences of usually larger spatial extent (R3). The dominant flood type in each catchment was then linked to catchment and climate attributes. This allowed to identify the impact of each flood type on flood statistics and how the flood types have changed over time. The main drivers determining the flood type were rainfall variability and antecedent conditions. Small and steep catchments were dominated by heavy-rainfall floods of shorter duration, while flat and wet catchments were dominated by long-duration floods with larger volumes. Such information can support selection of effective flood protection and management measures.

1. Introduction

Flood events are among the most prevalent and financially challenging climate hazards, causing disruptions for lives and livelihoods worldwide. An estimated 1.81 billion people live in areas with significant flood risk (Rentschler et al., 2022). Yet not all flood events are alike: they have differing generating processes and vary in their consequences. Flood events can have generating processes such as high intensity rainfall, long-duration rainfall or snowmelt (Brunner and Fischer, 2022). Hence, also the resulting flood hydrograph differs significantly. For example, heavy-rainfall floods are usually characterised by hydrographs with a steep rising limb and a short duration. The high intensities of heavy-rainfall events often exceed the infiltration capacity of the soil, leading to fast reactions of the catchment (Singh, 1997). Long-duration rainfall floods, though, exhibit a long duration with large volume and often multiple flood peaks (Grimaldi et al., 2012). In turn, protection measures and management strategies for these floods should be selected accordingly, thereby taking into account the hydrograph shape (Brunner et al., 2017). For flash floods, a comparably small storage volume is

required, but the infrastructure must withstand the large flood peak. For long-duration-rainfall floods larger storage capacity is required. Knowledge regarding the flood-generating processes and their relation to the hydrograph shape may not only be useful for decision making in flood protection, but it is also important for obtaining the required flood statistics.

Classification of floods is commonly based on annual exceedance probabilities of their peaks estimated from probability distributions and hence depends on various assumptions and data availability. The choice of an appropriate distribution function and parameter estimation are often connected with high uncertainties (Bomers et al., 2019). In addition, limited length of data series and the stochastic characteristics of the occurrence of extreme events add further uncertainty. The different flood-generating processes make the assumption of homogeneous samples questionable.

For example, Rossi et al. (1984) and Connell and Pearson (2001) used a two-component mixture distribution to distinguish between two different kinds of floods, which is often beneficial for flood frequency analysis (Totaro et al., 2024b). Waylen and Woo (1982) distinguished

* Corresponding author.

E-mail address: markus.pahlow@canterbury.ac.nz (M. Pahlow).

between rainfall- and snowmelt-driven floods to obtain more homogeneous samples. In general, a flood typology can help to classify flood events according to their generating conditions. There exist various flood typologies with different assumptions and requirements on the data for classification. [Tarasova et al. \(2019\)](#) identified three major types of flood classification: hydro-climatic classifications with consideration of atmospheric processes, hydrological classification with consideration of the catchment state and hydrograph-based classifications with a focus on the resulting flood event. Examples of these flood classifications are provided in [Hirschboeck \(1988\)](#), [Ashley and Ashley \(2008\)](#), [Nied et al. \(2014\)](#), [Kampf and Lefsky \(2016\)](#), [Yan et al. \(2019\)](#) or [Lu et al. \(2020\)](#). All classifications have advantages and drawbacks, and their applicability should be investigated according to the task at hand and the data availability. Here, we applied a hybrid causative-hydrograph-based classification that takes into account meteorological conditions in terms of rainfall and snowmelt as well as the hydrograph shape ([Fischer et al., 2019](#)). The advantages of this approach lie in the low demand of data making it especially suitable for flood statistics and the focus on flood events rather than atmospheric conditions. The different flood types are then considered in a type-based mixture model ([Fischer and Schumann 2023](#), [Yan et al., 2023](#)), an extension to the above-mentioned two-component models that allows for a joint estimation of return periods while still preserving the information on the impact of each flood type on the flood quantile.

To understand the impact of flood types on flood statistics and their relevance for a given geographic region it is necessary to understand the linkage between the meteorological processes and the catchment state. Only the combination of unfavourable conditions of both aspects makes the occurrence of extreme floods possible ([Merz et al., 2022](#)). For example, [Pham et al. \(2020\)](#) found that hilly catchments with sparse vegetation favour the generation of flash floods. [Stein et al. \(2021\)](#) linked certain catchment attributes such as vegetation or catchment elevation to flood-generating processes. The relevance of catchment attributes for the generation of floods thereby varied with the flood type.

Therefore, to extend the causative-hydrograph perspective of the flood types, in this work, the coherence of catchment attributes and flood type was investigated by a machine learning approach. Machine learning approaches are often used to detect patterns in large data sets of a broad range of variables. In contrast to classical physical models in hydrology, which are seen as knowledge-driven as they incorporate physical knowledge on the processes in a catchment, machine-learning approaches are termed “data-driven”. This refers to the fact that a priori, no information on physical processes is given to the algorithm (an exception are physics-based machine-learning tools, see e.g. [Bhasme et al., 2022](#) or [Chen et al., 2023](#)). This allows to detect hidden patterns that may have not been expected from the beginning. Especially when dealing with the complex topic of flood generation, where many variables interact and their impact might change with the initial conditions, machine learning provides a means to detect hidden coherences. Among others, [Khosravi et al. \(2019\)](#) or [Schmidt et al. \(2020\)](#) used such approaches to model flood events based on catchment attributes. [Jiang et al. \(2022\)](#) employed machine learning to detect flood mechanisms, [Oppel and Schumann \(2020\)](#) identified dominant controls for runoff, while [Stein et al. \(2021\)](#) linked meteorological and catchment attributes to flood-generating processes. There exist numerous approaches in machine-learning that could be used to link flood types and catchment attributes, e.g. regression trees, model trees, random forests, multilayer perceptrons (MLP) or support vector machines (SVM) to name only a few. An overview is, for example, given in [Xu and Liang \(2021\)](#). In general, it can be concluded that there does not exist one, unique algorithm that is best applicable for hydrological purposes (no-free-lunch-theorem, [Zolghadr-Asli, 2023](#)). Instead, all algorithms have their pros and cons and their applicability depends on the task at hand and the available data. For example, it was shown by [Kratzert et al. \(2019\)](#) that a long-short term memory (LSTM) model can outperform classical lumped rain-runoff models and even neural networks. However, this comes with

the drawback of not being (easily) able to interpret the model structure and hence the assumed physical processes. Interpretability is an often-cited criticism of many machine-learning methods, which are often termed as black-box models ([Chakraborty et al., 2021](#)).

In the current work a Classification and Regression Tree (CART) approach has been applied. The use of CART compared to many other machine learning techniques is beneficial as it delivers hydrologically interpretable results ([Addor et al., 2018](#)). This is important for the task at hand as we aim to link catchment attributes to certain flood types and hence identify physical coherences that can be transferred to other catchments and regions as well. In this case, interpretability of the model has to be ensured which excludes many of the named machine-learning approaches. Regression trees, however, have been proven to be interpretable and hence can be used to derive rules of decision making ([Hu et al., 2017](#)).

Furthermore, when investigating flood probabilities and design floods, one should consider how the results might change in the future. A changing climate leads to changing meteorological conditions and this might result in both changes in the distribution and characteristics of the flood types ([Turkington et al., 2016](#)). Moreover, there might occur periods of few and periods with many flood events ([Fischer et al., 2023](#); [Totaro et al., 2024a](#)). Again, a differentiation of the floods according to their generating processes might reveal changes that are hidden otherwise. For example, it was found that across Europe for the time period 1961–2010 the frequency of both winter and summer long-rain events increased significantly, while that of summer snowmelt events decreased significantly ([Hundecca et al., 2020](#)). Similarly, [Burn and Whitfield \(2023\)](#) found an increased impact of rainfall for flood events in Canada, while [Fischer and Schumann \(2024\)](#) detected significant increases of heavy rainfall floods for Central Europe. A special focus in this case should be laid on the frequency of flood events. Observational records from the central United States present limited evidence of significant changes in the magnitude of flood peaks, but strong evidence points to an increasing frequency of flooding ([Mallakpour and Villarini, 2015](#)). Yet [Archfield et al. \(2016\)](#) conclude that meaningful generalisations about flood change across the United States remain elusive. Changes in flood types due to climate change may result in damage and economic cost from extreme weather events that had not been observed earlier. This was also observed for New Zealand. [Frame et al. \(2020\)](#) found that NZ\$140 million of the total NZ\$470 million in damages from the 12 worst flood events in New Zealand over the period 2007–2017 were directly attributable to climate change. Yet, it remains important to also being able to attribute these changes, i.e., link the observed changes in flood frequency and magnitude to the underlying processes.

In the current work, a change of frequency of flood types was investigated for New Zealand. For this purpose, the methodology proposed by [Fischer and Schumann \(2023\)](#) was applied. By differentiating between flood types, one is able to infer the flood-generating processes and hence the cause of change. The focus on the frequency allows for consideration of the temporal distribution of such events.

The overall aim of this study is to determine the impact of catchment and climate attributes on flood generating processes and their effect on flood statistics. To achieve this the following objectives were pursued:

- i. Spatially differentiated classification of flood events into flood types;
- ii. Linkage of the dominating flood types to attributes (e.g. catchment size, climate, etc.);
- iii. Analysis of the impact of flood types on flood statistics;
- iv. Determination and assessment of temporal changes of the frequency of flood types.

The novelty of this approach lies in the consideration of flood types in combination with flood statistics and machine-learning. By making use of a flood event classification that takes into account the flood-generating processes, not only homogeneity is ensured in flood

frequency analyses (which is crucial but in practice often ignored), but the physical processes that are identified by linking the flood type and the catchment attributes that are relevant for this flood type can be interpreted much easier. Many current studies assume that the same processes are valid for all types of floods, which is questionable and might hide certain changes and coherences (for example, an increase of frequency of rainfall-driven floods and a synchronous decrease of snow-driven floods could be interpreted as no change when the floods are not differentiated regarding their generation). The proposed approach based on flood types, instead, allows for a differentiation of physical processes and hence a more precise attribution of extremes and changes observed within a region. It allows to include physical knowledge in classical stochastic approaches and hence extends both classical approaches that aim to link catchment attributes and floods as well as classical flood statistics. The application to New Zealand offers a unique opportunity to consider one of the most hydrologically diverse regions in the world that is crucially impacted by floods.

2. Study area, data and methods

2.1. Study area and data

In this work, daily discharge data from 252 gauged catchments across New Zealand with a minimum record length of 20 years were used (Fig. 1). The data were obtained from the National Institute of Water and Atmospheric Research’s (NIWA) national database, alongside data supplied by regional councils. Only sites that were not affected by large engineering projects such as dams, diversions or substantial abstractions were included (Booker and Woods, 2014). The catchments range between order 1 and order 7 of the Strahler classification, with the majority of catchments being order 5. The catchment area varies from 0.3 to 6600 km² (see Fig. S1 in the supplementary material).

The catchments in New Zealand represent a very wide range of hydrological conditions with different topography, geology, soil type, land cover and land use (Mosley, 1981). However, the majority of the

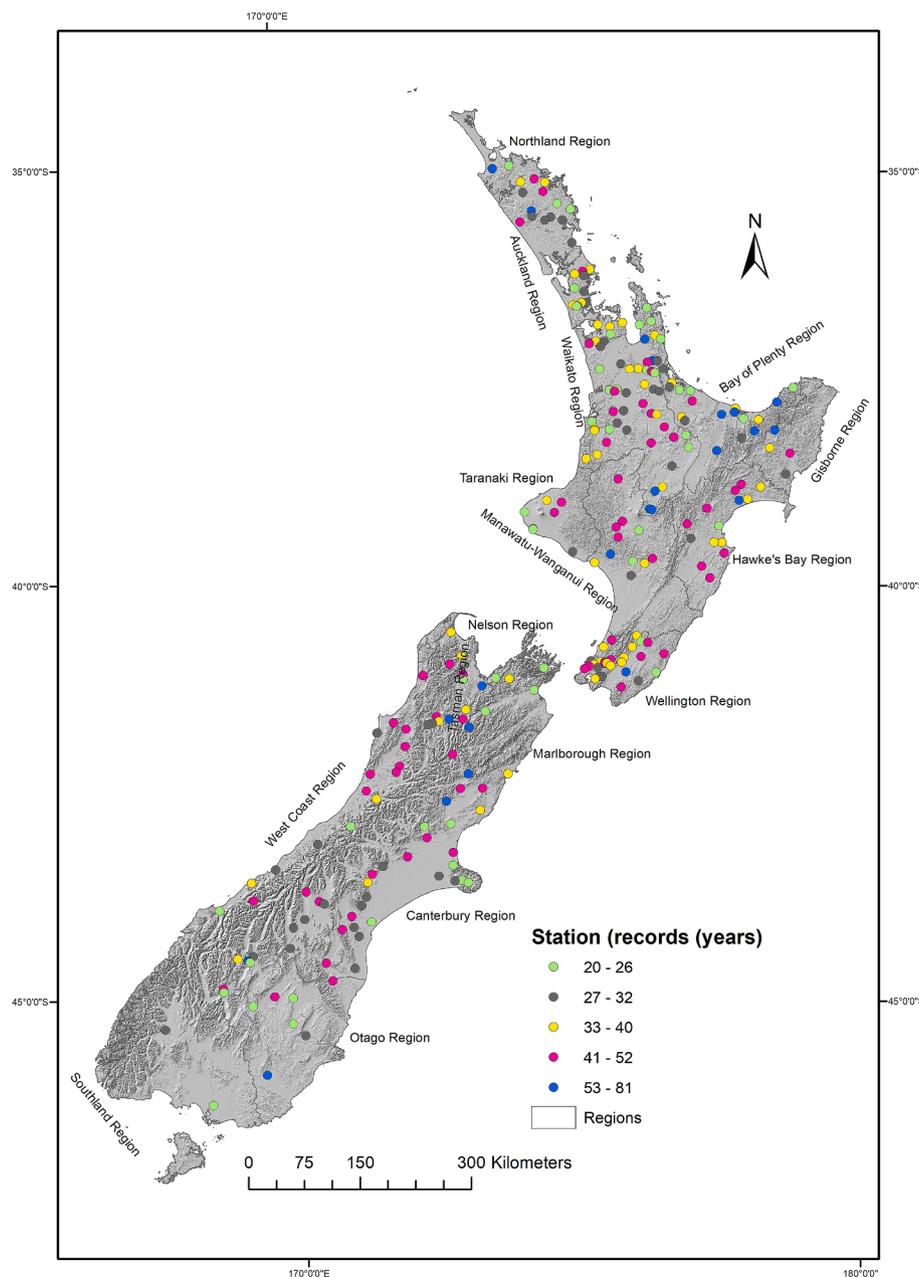


Fig. 1. Locations of the 252 flow gauging station used in this study and the lengths of the respective discharge time series.

catchments are rural, with only minor urbanised areas. The main land uses are dairy cattle farming, sheep farming and forestry. The length of the mean daily discharge data time series ranges from 20 to 81 years (between 1930 and 2011), where at least the period 1991 until 2011 is covered for all stations. See Booker and Woods (2014) and McMillan et al. (2016) for more details regarding these gauging stations. Daily

temperature and precipitation for the same time period were retrieved from New Zealand’s Virtual Climate Station Network (Tait et al., 2006). The digital river network was taken from the River Environment Classification (REC) (Snelder and Biggs 2002). Land cover and soils data were retrieved from the Land Cover Database and Land Resource Inventory (Newsome et al., 2000). Additional catchment attributes were

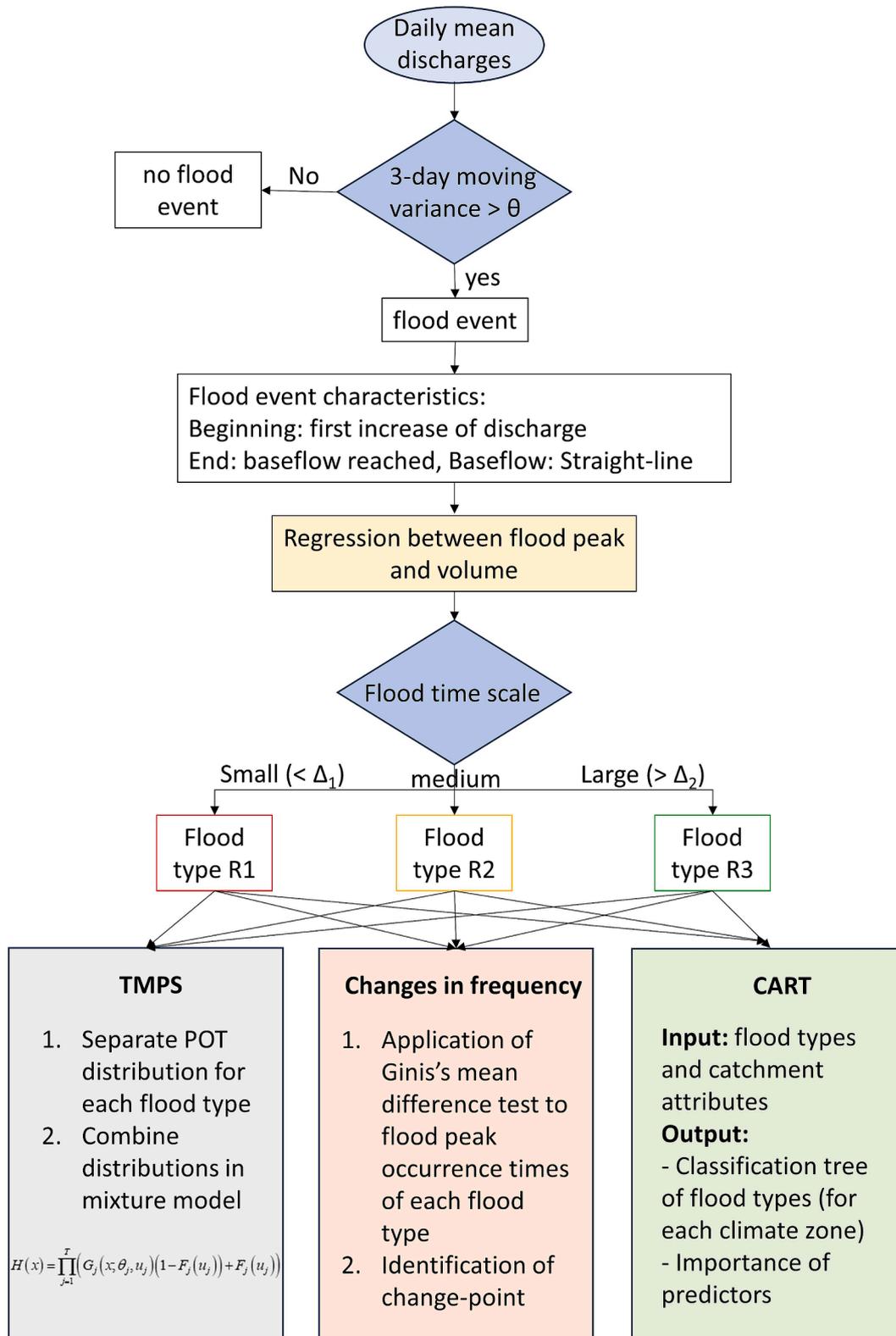


Fig. 2. Flow chart of the proposed methodological framework.

obtained from a range of sources (Booker and Snelder 2012; Booker and Woods 2014; Snelder et al., 2013). Overall, the following fourteen attributes were utilised (maps of selected attributes are provided in the supplementary material): Baseflow index (BFI) (ratio of long-term mean baseflow to total streamflow), rain days >10 mm, annual rainfall variability (coefficient of variation of annual catchment rainfall) (Figure S2), climate (categories warm extremely wet, warm wet, warm dry, cool extremely wet, cool wet, cool dry), source of flow (mountain, hill, and low elevation, lake), runoff from area of catchment with slope >30 %, drainage density (total length of streams per catchment area), catchment specific flow (flow divided by the catchment area) (Figure S3), catchment area (Figure S1), catchment elevation, upstream average slope (Figure S4), geology category (spatially dominant geology from alluvium, hard sedimentary, soft sedimentary, volcanic basic, volcanic acidic, plutonic), particle size of rocks (reflecting flow paths and infiltration capacity) and surface hardness (induration).

3. Methodology

The methodological framework consists of several steps that are introduced in this section. They are all based on the separation and classification of flood events. An overview over the methodological framework is given in Fig. 2.

3.1. Event separation and classification

To allow for an application of flood type classification, first the separation of flood events was carried out. Here, the automated event separation proposed by Fischer et al. (2021) has been employed, using daily discharge data and a variance-based threshold to define beginning and end of a flood event. This separation algorithm takes into account the changing dynamics of flood events compared to non-flood discharge and is capable to separate multiple-peaked flood events (Fischer et al., 2021). More precisely, the three-day moving variance is considered. As soon as this variance exceeds the long-term mean plus 0.5 times the long-term standard deviation of the variance (analogous to the 2-sigma rule), a flood event is detected. The beginning of the flood event is then defined as the day before the detected variance exceedance where the discharge increases for the first time. The end is defined as the day where the baseflow, i.e. the discharge at the beginning of the event, is reached again or a criterion for baseflow increase is met. As a result, the flood volume and further flood properties such as baseflow (determined with a straight-line method between beginning and end of the flood event), the flood peak(s) and flood duration were obtained. Furthermore, the associated event precipitation amount was estimated by application of a change-point test that estimates the beginning of the related rainfall event as that point where the cumulated rainfall changes most in terms of slope (Fischer et al., 2021). The end of the associated rainfall event was assumed to be equal to the end of the flood event. A flood typology was applied to the events to derive information on the underlying meteorological causes of the floods. To achieve this, the typology proposed by Fischer et al. (2019) was utilised. This typology has a low data requirement and can be applied in an automated way. It is mainly based on the hydrograph shape and makes use of the flood time scale (Gaál et al., 2012), i.e. the relation between flood volume and flood peak as a measure of the spread of the hydrograph. The time-scale ordered flood events are then split into three subgroups. The optimal split is determined by using an optimisation of the coefficient of determination of the linear regression between direct flood peak and volume (i.e., peak and volume minus baseflow) for each subgroup separately. This optimal split then defines the three flood types, one belonging to each subgroup. This simple procedure, which only requires separated flood events, makes the classification suitable for large data sets with different observation periods. In this work, three flood types are considered: R1, R2 and R3 (Fig. 3). Though neither rainfall nor catchment attributes are directly used for the classification, the resulting flood types are nonetheless

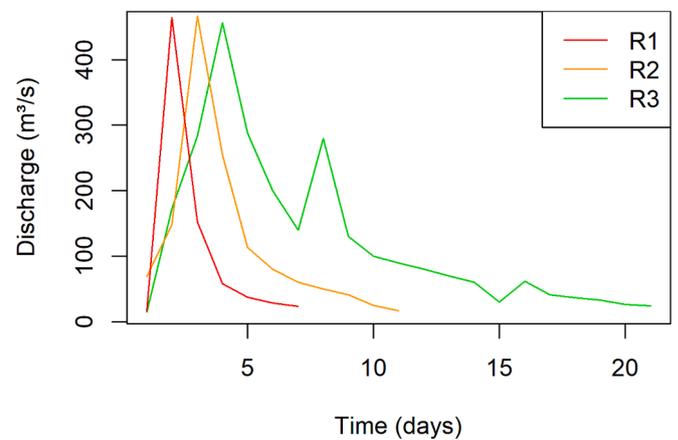


Fig. 3. Example hydrographs with similar peaks for the flood types considered in this study recorded at station 9,010,876 (Hutt River at Taita Gorge).

related to these, as the hydrograph shape is mostly defined by the atmospheric and catchment conditions (Brunner and Fischer, 2022). R1 refers to floods usually associated with heavy rainfall. These floods are characterised by a small volume, short duration but often large flood peaks. R2 events are usually associated with moderate rainfall events of medium intensity and duration. R3 floods are related to long-duration sequences of rainfall of usually large spatial extent. These floods often have multiple peaks as the wet soil quickly generates flood peaks whenever a sequence of rainfall starts. Snowmelt-related flood types as in Fischer et al. (2019) or Brunner and Fischer (2022) are not considered in this study, as the role of snowmelt for the catchments under consideration is rather small. For example, Kerr (2013) has shown that the contribution of snowmelt induced river flows is merely 3.3 % across the South Island of New Zealand. Hence the likelihood that the threshold of 20 % contribution of snowmelt (as defined by Fischer et al. (2019)) to define a snowmelt-induced flood event) would be exceeded, is low.

3.2. Statistical mixture model for flood types

To obtain flood statistics that take into account the different flood types separately but also deliver flood frequency analyses for their joint distribution, the mixture model proposed by Fischer (2018) was applied. This so-called type-based mixture model for partial duration series (TMPS) was developed for considering different flood types and their impact on the joint mixture model. Each flood type $j = 1, \dots, 3$ was considered as separate sample X_{1j}, \dots, X_{n_j} . To each of these samples, a threshold u_j was applied to obtain the peak-over-threshold (POT) sample $X_{i,j} > u_{j,i} = 1, \dots, n_j$. The threshold u_j was chosen as three times the weighted mean annual discharge, where the weights are chosen according to the frequency of the respective flood type in each month. The Generalized Pareto distribution was fitted to these POT-samples using L-moments:

$$G_j(x; \theta_j = (\kappa_j, \beta_j), u_j) = 1 - \left(1 + \kappa_j \left(\frac{x - u_j}{\beta_j} \right) \right)^{-\frac{1}{\kappa_j}} \quad (1)$$

for a shape parameter $\kappa_j \neq 0$ and scale-parameter $\beta_j > 0$ with support $x \geq u_j$. The POT approach can be generalised to an annual distribution \tilde{G}_j for each flood type by

$$\tilde{G}_j(x) = \sum_{k=0}^{\infty} P_j(l=k) (G_j(x, \theta_j, u_j))^k \quad (2)$$

where $P_j(l=k)$ is the probability that the annual number l of flood peaks of type j above the threshold u_j is equal to k and can be described by the Poisson distribution with parameter λ_j (Cunnane, 1973; Stedinger et al.,

1993)

$$P_j(l = k) = \frac{\lambda_j^k}{k!} e^{-\lambda_j} \quad (3)$$

Additionally, the annual joint distribution of all flood types can be obtained by

$$H(x) = \prod_{j=1}^T (G_j(x; \theta_j, u_j) (1 - F_j(u_j)) + F_j(u_j)) \quad (4)$$

where F_j is the Generalized Extreme Value (GEV) distribution fitted to all floods of type j and describes the probability of non-exceedance of the threshold u_j .

The consideration of type-based statistics allows for an analysis of the impact of each flood type on the joint distribution. By evaluating which flood type leads to the largest flood quantile for a given return period, one can investigate which flood type contributes most to the joint mixture model.

The focus here is on large floods, hence the 1 % AEP flood was chosen as benchmark. The 1 % AEP flood is an often-used quantile, e.g., for design floods, for regionalisation evaluation or as validation criteria for hydrological models. Here, it serves as a characteristic large flood. More precisely, it is being investigated which flood type has the largest impact on the TMPS model for the 1 % AEP flood, i.e., is closest to the TMPS estimation of this flood quantile. This flood type then is referred to as the 'dominating flood type'. With this consideration, it is possible to define the flood type which potentially leads to the largest flood peaks and also dominates the mixed distribution function for large quantiles. As shown by Fischer (2018) or Fischer and Schumann (2023), mostly one flood type tends to dominate the right tail of the mixture distribution. If this flood type is known, together with an estimation of the flood peak, the flood event in terms of volume and duration can be reconstructed by using the flood timescale that was used to define the flood types. Additionally, also the corresponding flood hydrograph can be estimated (Fischer and Schumann, 2023). All this information is essential for the design of adequate flood protection measures. The dominating flood type was estimated by using the flood-type-specific distributions introduced above. It was compared to the TMPS model estimation of the 1 % AEP flood to determine the deviation from the type-specific distribution to the mixture distribution.

3.3. Changes of frequency

The approach by Fischer et al. (2019) and Fischer and Schumann (2023, 2024) has been employed to each flood type to test whether there is a change in the frequency of occurrence of flood events of the respective flood type over time. For this purpose, a change point test for the variance based on Gini's Mean Difference was applied to the time series of time differences between two flood events of the same flood type (in days). The test statistic for the change-point test based on Gini's Mean Difference is defined as

$$G = \frac{1}{2\sigma\sqrt{n}} \max_{1 \leq k \leq n-1} |g_{1,k} - g_{k+1,n}| \quad (5)$$

with σ being the sample standard deviation and

$$g_{k,n} = \frac{2}{(n-k+1)(n-k)} \sum_{k \leq i < j \leq n} |X_i - X_j| \quad (6)$$

The change point was obtained by taking the argmax of G and the tendency of change, i.e., whether there occurred more or less events, was estimated by comparing the mean number of events per year before and after the change point.

3.4. Flood types and catchment attributes

While the flood type provides information on the meteorological and catchment state right before the flood event, it does not include the general properties of a catchment, e.g., the shape or the land use. However, flood types and especially the flood hydrographs are impacted by these factors, too. Therefore, it has been analysed how the flood types in New Zealand are linked to catchment attributes. Specifically, the dominating flood type was linked to the fourteen catchment attributes introduced in Section 2.1.

Importance of the different attributes was calculated as the increase in mean square prediction error when each predictor is randomly permuted and is based on the out of bag (OOB) samples. Importance represents the contribution to accuracy of independent predictions for each predictor and is equivalent to the increase in error when a term is dropped from a linear model (Singh et al., 2019).

In a second step, a Classification and Regression Trees (CART) algorithm was used for the prediction of the dominating flood type. The training set comprised 70 % of the data to predict the remaining 30 % in the test sample of all catchments. The sampling was repeated 1000 times to estimate uncertainty in the prediction. Input data were all fourteen catchment attributes used in this study. The use of the CART, in contrast to many other machine learning techniques for classification and prediction, makes the resulting classification tree reproducible and the decisions can be checked for hydrological meaningfulness. This makes this algorithm preferable compared to many other classification methods, especially since we aim to understand the different relationships between catchment attributes and flood types.

4. Results and discussion

4.1. Spatial distribution of the different flood types and type-based flood statistics

Fig. 4 illustrates the prevalent flood types in New Zealand, indicating that 12 %, 47 %, and 41 % of the 252 catchments are predominantly characterised by flood types R1, R2, and R3, respectively. This distribution aligns with the results reported by Stein et al. (2020). In their comprehensive global study, Stein et al. (2020) employed an event-based classification to analyse the processes responsible for generating river floods. The prevalent dominant flood generating process, which the authors define as the process occurring most often in the time series (based on an analysis of at least 20 years of data), for the 21 catchments studied across New Zealand was classified as excess rainfall due to saturated conditions. In the current study, the Wellington and Auckland regions are mainly dominated by flood type R2, whereas the West Coast and Manawatu-Wanganui region are dominated by flood type R3. For the eastern part of the Nelson region, flood type R1 is dominant. A mix of types R2 and R3 was found for the Hawkes Bay, Gisborne and Bay of Plenty regions. For the remaining regions, there is a mixture of dominant flood types. This can have different reasons. On the one hand, the dominating flood type might not be distinct, i.e. several flood types could contribute similarly to the joint distribution, though only the one with the highest proportion is shown here. On the other hand, the dominating flood type might be also related to the catchment size. Hence, in one region, different flood types can dominate depending on the catchment size. For example, one would expect R1 floods to be less relevant in large catchments, as they are usually related to heavy rainfall which contributes less to large areas due to its small spatial extension. This will be investigated later on.

Fig. 5 shows the design floods estimated using the TMPS model for different return periods. Two stations were selected, one located at the Hutt River on the North Island in the Wellington region (catchment size 655 km²) and the other at the Buller River in the West Coast region of the South Island (6350 km²). For the first, the TMPS distribution clearly displays a heavy tail which is caused both by the heavy-rainfall floods

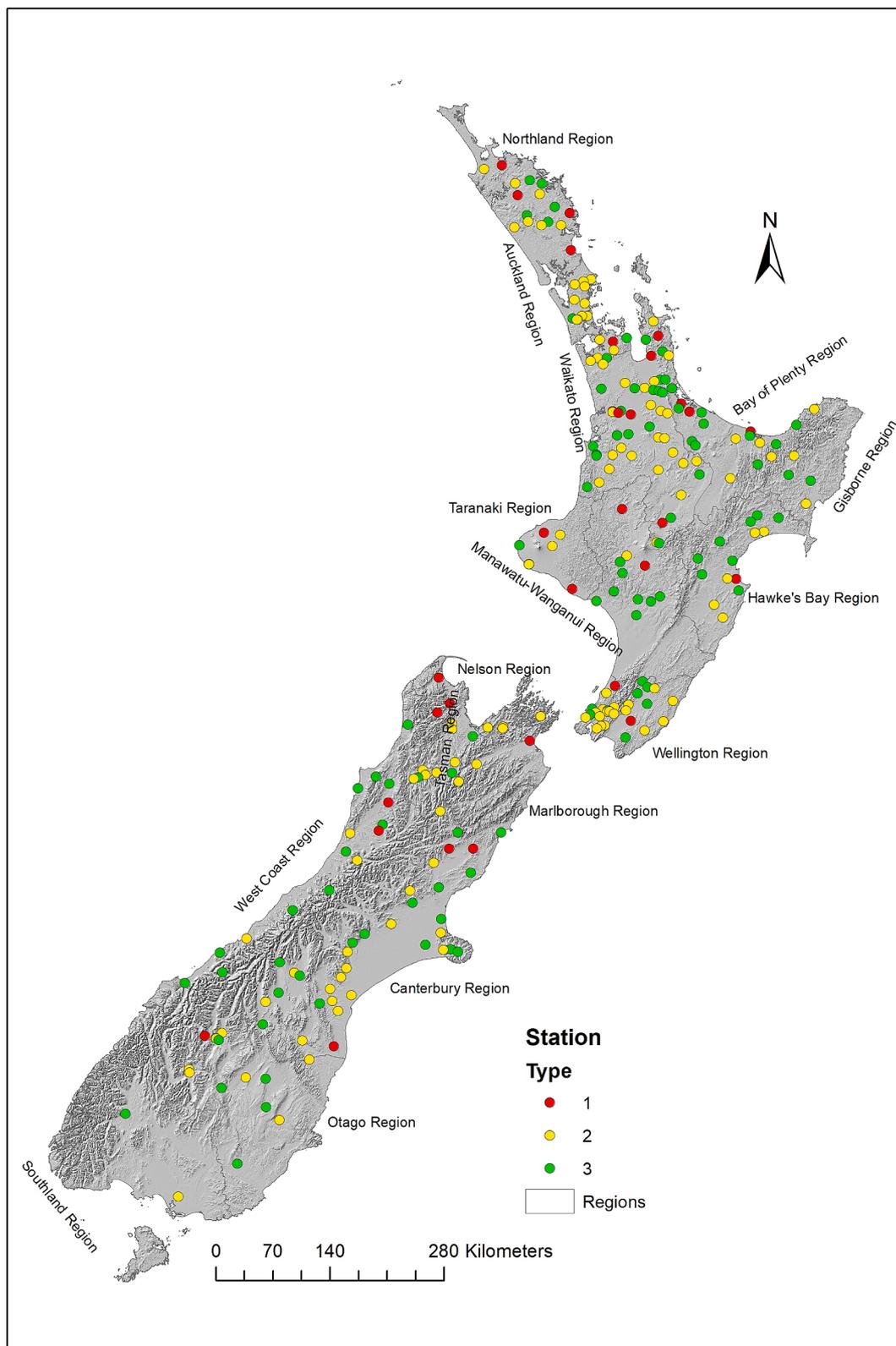


Fig. 4. Spatial distribution of dominant flood types R1, R2 and R3 identified across New Zealand.

(R1) and the sequence-of-rain floods (R3), which have a similar distribution shape. The R1 flood events result in the largest design flood quantiles. Yet due their low frequency these do not have a strong impact on the TMPS model. The R3 flood events dominate for this catchment, i. e., the TMPS model is clearly impacted by these events. R2 floods only play a minor role and only for small return periods. For the Buller River,

the picture changes. Here, no heavy tail can be detected for any of the type-specific distributions nor the TMPS model. R1 flood events and R2 flood events have a similar distribution, both causing the most extreme floods in this catchment. The TMPS model is clearly impacted by both flood types, though R2 floods are more frequent here and dominate the TMPS model.

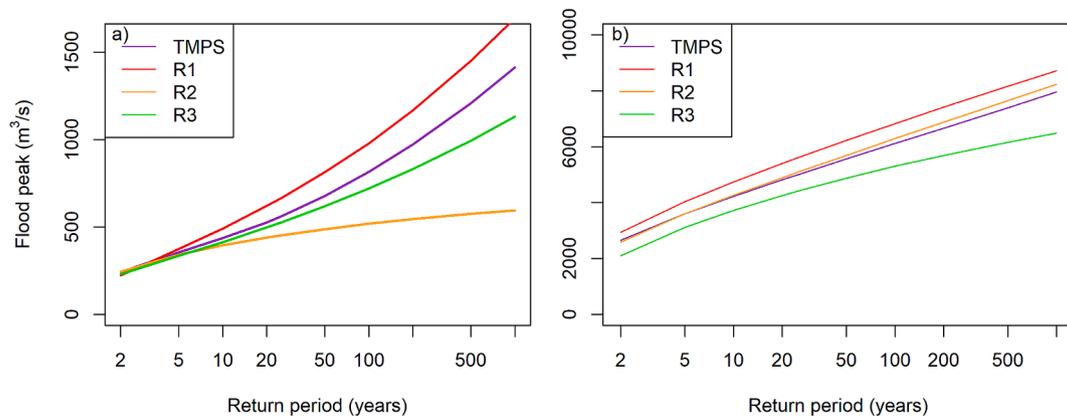


Fig. 5. Estimated design floods for different return periods for the TMPS model and the type-specific distributions at a) station 9,010,876 (Hutt River at Taita Gorge) and b) station 12,012,479 (Buller River at Te Kuha).

4.2. Change-points

The change-point analysis revealed that for the investigated time period there is an overall tendency for an increase of the frequency of occurrence of heavy rainfall floods (R1) on the North Island, while the opposite is observed for the South Island, i.e., R1 floods occur less frequently (Fig. 6). The R2 flood type was observed less frequently in the Nelson, Marlborough and Canterbury regions of the South Island. R3 floods occurrence was found to decrease for the northern part of the South Island (Nelson and Tasman regions). The same holds true for the eastern part of the South Island (Marlborough, Canterbury and Otago regions), while the R3 flood frequency increased on the West Coast region (Fig. 6). These two regions of the South Island are separated by the Southern Alps mountain range (see Fig. 4). The Southern Alps act as a barrier to the prevailing westerly airstream which is both deflected by them and forced to ascend. This causes rain which is often heavy and prolonged in the West Coast region (Macara 2016), with an average annual rainfall of 4 m (Griffiths et al., 2020). In the Gisborne and Hawkes Bay regions, the occurrence of R3 flood decreased, while the frequency of R2 floods has increased. An increase in particular in R1 and R3 flood frequency was found for the Northland region, while in the adjacent Auckland region R3 flood occurrence decreased.

Queen et al. (2023) studied spatiotemporal trends in near-natural New Zealand river flow for records of at least 50 years length,

including the years 1969–2019. The authors identified that winter streamflow has significantly increased, while spring streamflow has increased on the West Coast region of the South Island. This seems to link well with our finding that R3 floods, i.e., floods related to long-duration sequences of rainfall, have increased in this region.

4.3. Classification and regression trees (CART)

CART has been employed to predict the dominant flood type based on catchment attributes. The predictor power stayed constant at ~50 % for the mean for 1,000 runs with a 70 % training dataset. The R1 flood type was found to be linked to high elevation and small catchments, while the R2 flood type was identified for medium size catchments with steeper gradients and high rainfall variability (Fig. 7). BFI is used to differentiate between R2 and R3 events, as one would expect R3 in case of high baseflow. Large flatter catchments at low elevation with high baseflow and large amounts of rainfall with small variability were linked to the R3 flood type.

Additionally, R1 is dominant in small catchments with high river density, which is also expected as otherwise the local heavy rainfall events simply vanish in space.

In a second step, the catchments have been divided by climate zones and the CART analyses were repeated. The performance was similar for all zones and still around 50 %. Results for the climate zone ‘warm-dry’

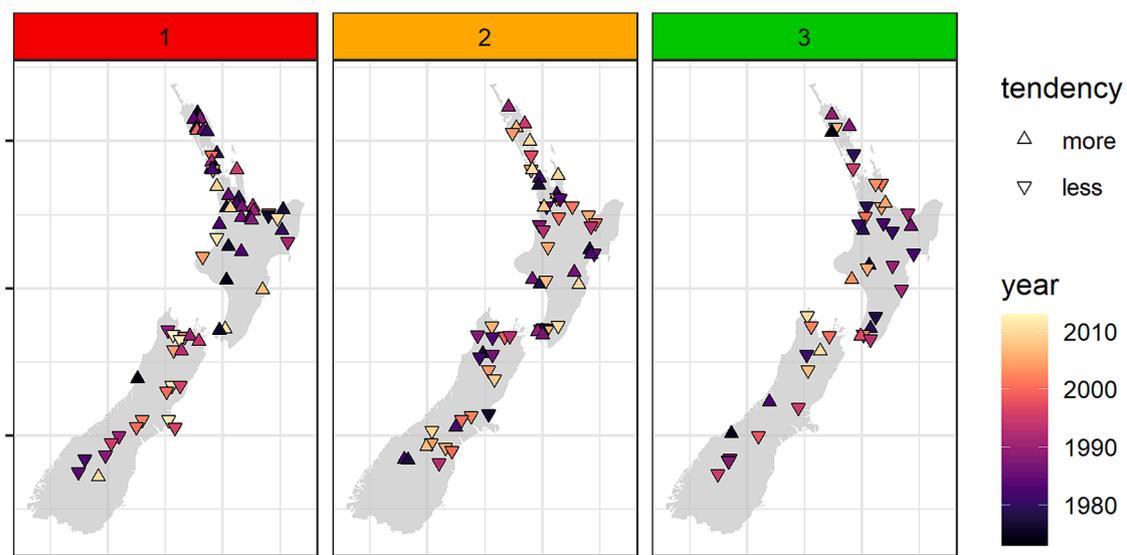


Fig. 6. Change points identified per flood type and tendency of frequency.

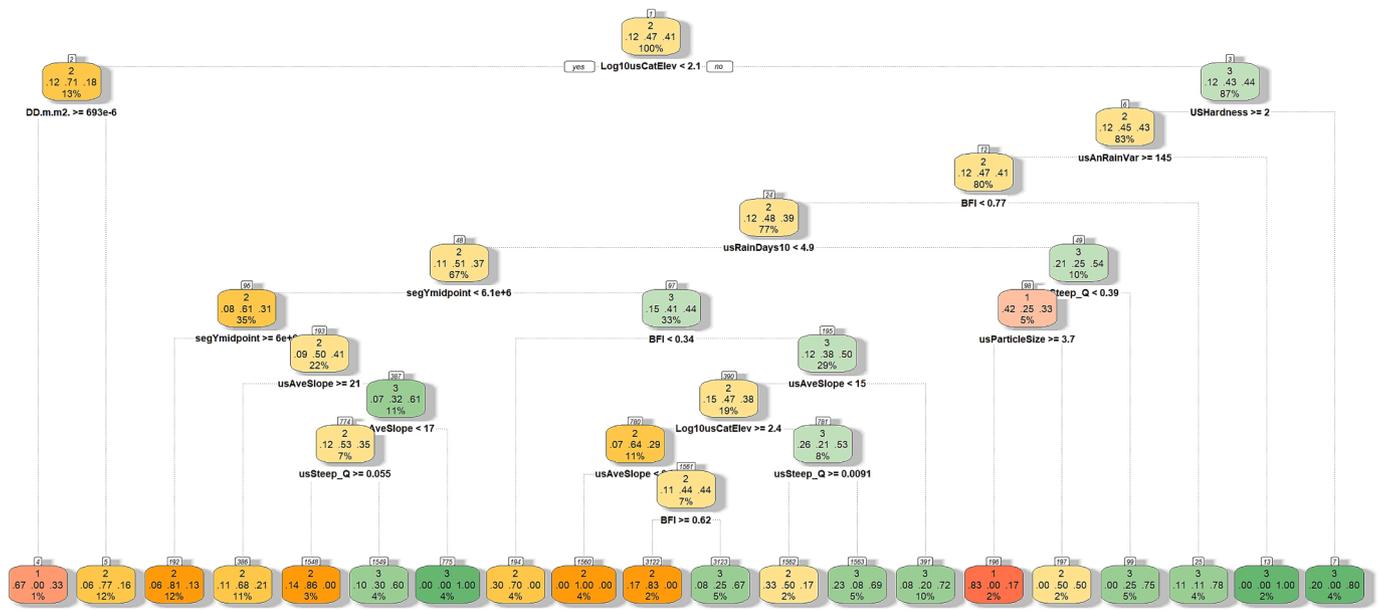


Fig. 7. Overall CART for all climate zones. Each node shows the predicted flood type, the predicted probability of each flood type and the percentage of observations in the node. The colour refers to the most probable flood type, the shade of the colour to the probability.

could not be considered as there were not enough catchments in this class to run a CART. The climate zone ‘cool-dry’ had flood type R3 as dominant flood type with high BFI (Fig. S5 in the supplementary material). For the climate zone ‘cool-extremely wet’, the catchment attributes rainfall variability and rainfall amount were most influential regarding the occurrence of either the R1 or the R2 flood type (Fig. S6). These attributes play a major role for this climate zone since the extremely wet conditions favour long-lasting rainfall.

Rainfall variability played a major role in separating R3 and R2 flood types for the ‘cool-wet’ climate zone (Fig. S7). Smaller and higher elevation catchments tend to have more floods with steep hydrographs and short duration, corresponding to flood types R1 and R2, i.e., heavy rainfall and long-duration rainfall. Finally, the drainage density can be used as an indicator in these catchments to further conclude on the dominating flood type. Slope, elevation and catchment size were found to be the key catchment attributes separating R1 floods and the

remaining flood types R2 and R3 for the climate zone ‘warm-wet’ (Fig. 8). The distinction between R2 and R3 flood types is mainly due to baseflow, but particle size plays an important role as well. This seems to correlate with volcanic geology in the Auckland region and south of the Bay of Plenty region that mainly show the R2 flood type (Fig. 4).

4.4. Relationship with catchment attributes

The fourteen predictors to explain the flood types were assessed in terms of their importance. Fig. 9 shows the relative importance of each of the predictors. As stated in Section 3.4, importance of the different attributes was calculated as the increase in mean square prediction error when each predictor is randomly permuted. It can be interpreted as the increase in error of the estimation when this particular predictor (i.e. catchment attribute) is left out of the estimation. The higher the importance, the greater this error would be.

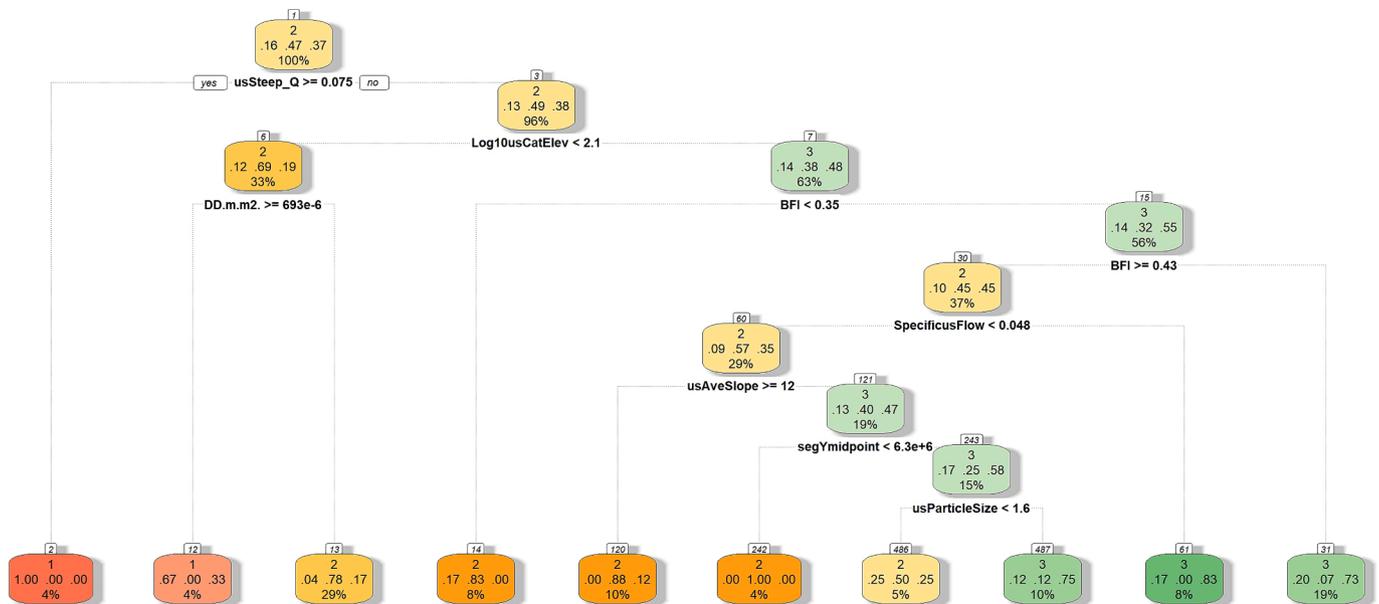


Fig. 8. CART for the climate zone ‘warm-wet’.

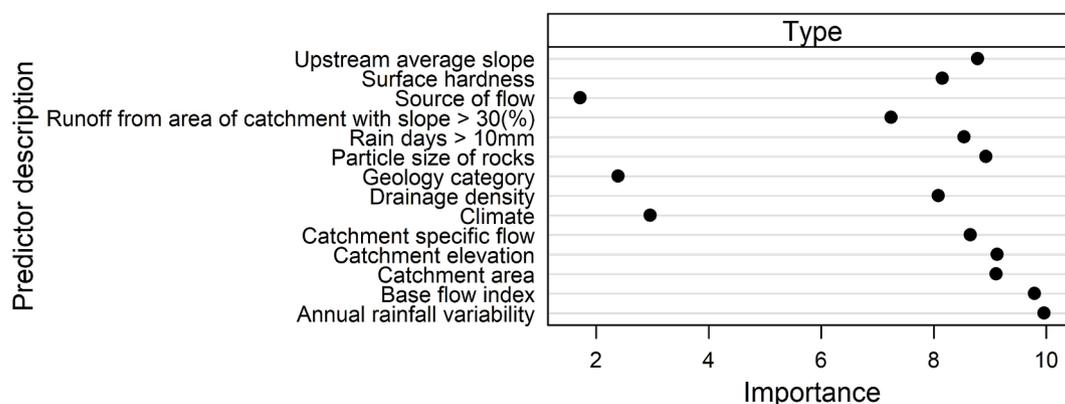


Fig. 9. Relative importance of the predictors.

Most influential are annual rainfall variability and the base flow index. The latter is linked to antecedent conditions. Next in terms of importance follow catchment elevation and catchment area. It was found that climate, geology and source of flow have comparably low importance for explaining the flood type.

Since the catchment attribute geology is non-numeric, the linkage to the different flood types R1, R2 and R3 has been investigated separately. The width of the colour bars in Fig. 10 represents the respective proportion of that catchment attribute. Geology was assigned to a catchment according to the largest proportion in the catchment. It must be acknowledged that the drawback of this assumption is that such a link merely represents one attribute even if more attributes are present, plus a potential change of the attribute over time is not accounted for. R1 was the least common flood type for any of the geology types considered. The geology types that cover the largest area, hard sedimentary, soft sedimentary and volcanic acidic, show similar frequency of flood types R2 and R3. R3 was found to be the dominating flood type for plutonic and volcanic basic geology, while it was the sole flood type for the category miscellaneous.

5. Conclusions

The focus of the current study was on flood types which potentially generate large floods that can cause a threat to human life and damage infrastructure. Knowledge regarding the type of these floods can support tailored design of flood protection and management measures since the flood types differ significantly in their flood volume and duration. For New Zealand and the time period studied, the main drivers that determine the flood type were rainfall variability and antecedent conditions. This finding suggests that in order to create an effective early-warning system investment in a dense rain and soil moisture sensing network is sensible. Furthermore, small and steep catchments were found to be dominated by heavy rainfall floods R1, while flat and wet catchments are dominated by long duration floods with large volumes, i.e., flood type R3. The results can be used in practice to improve the flood prediction and flood protection measures. Flood forecasting should take into account that different atmospheric and catchment states result in different flood types. Thus, a different parameterisation of the forecasting models might be meaningful. By taking into account the different flood types and thus different hydrograph shapes, flood protection measures can be directly adapted for the given situation, e.g. by making space in the reservoir when an R3 flood event with high volume is expected or by accounting for large flood peaks when an R1 flood event is predicted. Examples for how to incorporate flood types in design hydrograph estimation are given in Fischer and Schumann (2023) for Europe and could be adapted for New Zealand.

In this study, heavy rainfall floods were found to occur more frequently on the North Island and less frequently on the South Island in

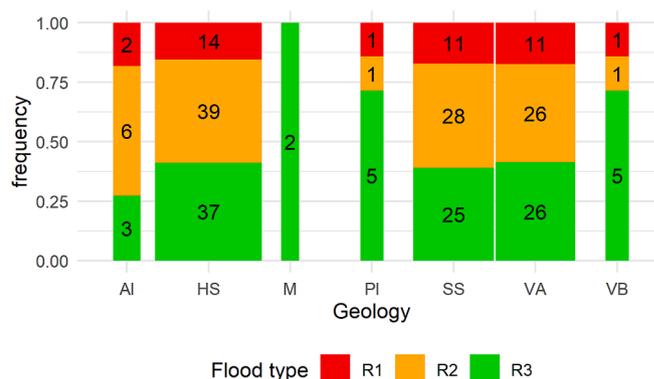


Fig. 10. Occurrence of flood types for different geology: (AI: Alluvium; HS: Hard Sedimentary; M: Miscellaneous; PI: Plutonic; SS: Soft Sedimentary; VA: Volcanic Acidic; VB: Volcanic Basic). The widths of the bars represent the number of catchments in the respective category, given also as numbers for each bar.

the last decades. Future work should identify trends under climate change scenarios for the North and South Island, or sub-regions thereof, to provide additional information to select appropriate measures to increase flood resilience.

CRediT authorship contribution statement

Svenja Fischer: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Markus Pahlow:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Shailesh Kumar Singh:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank the Associate Editor, the reviewer Lei Yan and two anonymous reviewers for providing constructive comments that helped to improve this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jhydrol.2024.132361>.

Data availability

The authors do not have permission to share data.

References

- Addor, N., Nearing, G., Prieto, C., Newman, A.J., Le Vine, N., Clark, M.P., 2018. A ranking of hydrological signatures based on their predictability in space. *Water Resour. Res.* 54, 8792–8812. <https://doi.org/10.1029/2018WR022606>.
- Archfield, S.A., Hirsch, R.M., Viglione, A., Blöschl, G., 2016. Fragmented patterns of flood change across the United States. *Geophys. Res. Lett.* 43, 10232–10239. <https://doi.org/10.1002/2016GL070590>.
- Ashley, S.T., Ashley, W.S., 2008. The storm morphology of deadly flooding events in the United States. *Int. J. Climatol.* 28, 493–503. <https://doi.org/10.1002/joc.1554>.
- Bhasme, P., Vagadiya, J., Bhatia, U., 2022. Enhancing predictive skills in physically-consistent way: physics informed machine learning for hydrological processes. *J. Hydrol.* 615, 128618. <https://doi.org/10.1016/j.jhydrol.2022.128618>.
- Bomers, A., Schielen, R.M.J., Hulscher, S.J.M.H., 2019. Decreasing uncertainty in flood frequency analyses by including historic flood events in an efficient bootstrap approach. *Nat. Hazards Earth Syst. Sci.* 19, 1895–1908. <https://doi.org/10.5194/nhess-19-1895-2019>.
- Booker, D.J., Snelder, T.H., 2012. Comparing methods for estimating flow duration curves at ungauged sites. *J. Hydrol.* 434–435, 78–94. <https://doi.org/10.1016/j.jhydrol.2012.02.031>.
- Booker, D.J., Woods, R.A., 2014. Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. *J. Hydrol.* 508, 227–239. <https://doi.org/10.1016/j.jhydrol.2013.11.007>.
- Brunner, M.I., Fischer, S., 2022. Snow-influenced floods are more strongly connected in space than purely rainfall-driven floods. *Environ. Res. Lett.* 17, 104038. <https://doi.org/10.1088/1748-9326/ac948f>.
- Brunner, M.I., Viviroli, D., Sikorska, A.E., Vannier, O., Favre, A.-C., Seibert, J., 2017. Flood type specific construction of synthetic design hydrographs. *Water Resour. Res.* 53, 1390–1406. <https://doi.org/10.1002/2016WR019535>.
- Burn, D.H., Whitfield, P.H., 2023. Climate related changes to flood regimes show an increasing rainfall influence. *J. Hydrol.* 617, 129075. <https://doi.org/10.1016/j.jhydrol.2023.129075>.
- Chakraborty, D., Bağaçoğlu, H., Winterle, J., 2021. Interpretable vs. noninterpretable machine learning models for data-driven hydro-climatological process modeling. *Expert Syst. Appl.* 170, 114498. <https://doi.org/10.1016/j.eswa.2020.114498>.
- Chen, C., Zhang, H., Shi, W., Zhang, W., Xue, Y., 2023. A novel paradigm for integrating physics-based numerical and machine learning models: A case study of eco-hydrological model. *Environ. Model. Softw.* 163, 105669. <https://doi.org/10.1016/j.envsoft.2023.105669>.
- Connell, R., Pearson, C.P., 2001. Two-component extreme value distribution applied to Canterbury annual maximum flood peaks. *J. Hydrol. (New Zealand)* 40 (2), 105–127.
- Cunnane, C., 1973. A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. *J. Hydrol.* 18, 257–271. [https://doi.org/10.1016/0022-1694\(73\)90051-6](https://doi.org/10.1016/0022-1694(73)90051-6).
- Fischer, S., 2018. A seasonal mixed-POT model to estimate high flood quantiles from different event types and seasons. *J. Appl. Stat.* 45, 2831–2847. <https://doi.org/10.1080/02664763.2018.1441385>.
- Fischer, S., Schumann, A.H., 2023. *Type-based flood statistics – An interlink between stochastic and deterministic flood hydrology*. Springer, Cham, Switzerland. <https://doi.org/10.1007/978-3-031-32711-7>.
- Fischer, S., Schumann, A., Bühler, P., 2019. Timescale-based flood typing to estimate temporal changes in flood frequencies. *Hydrol. Sci. J.* 64, 1867–1892. <https://doi.org/10.1080/02626667.2019.1679376>.
- Fischer, S., Schumann, A., Bühler, P., 2021. A statistics-based flood event separation. *J. Hydrol. X* 10, 100070. <https://doi.org/10.1016/j.jhydroa.2020.100070>.
- Fischer, S., Lun, D., Schumann, A.H., Blöschl, G., 2023. Detecting flood-type-specific flood-rich and flood-poor periods in peaks-over-threshold series with application to Bavaria (Germany). *Stoch. Env. Res. Risk A.* 37 (4), 1395–1413. <https://doi.org/10.1007/s00477-022-02350-8>.
- Fischer, S., Schumann, A.H., 2024. Temporal changes in the frequency of flood types and their impact on flood statistics. *J. Hydrol. X* 22, 100171. <https://doi.org/10.1016/j.jhydroa.2024.100171>.
- Frame, D.J., Rosier, S.M., Noy, I., Harrington, L.J., Carey-Smith, T., Sparrow, S.N., Stone, D.A., Dean, S.M., 2020. Climate change attribution and the economic costs of extreme weather events: a study on damages from extreme rainfall and drought. *Clim. Change* 162, 781–797. <https://doi.org/10.1007/s10584-020-02729-y>.
- Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., Blöschl, G., 2012. Flood timescales: understanding the interplay of climate and catchment processes through comparative hydrology. *Water Resour. Res.* 48, 1–21. <https://doi.org/10.1029/2011WR011509>.
- Griffiths, G.A., Singh, S.K., Mc Kerchar, A.I., 2020. Flood frequency estimation in New Zealand using a region of influence approach and statistical depth functions. *J. Hydrol.* 589, 125187. <https://doi.org/10.1016/j.jhydrol.2020.125187>.
- Grimaldi, S., Petroselli, A., Serinaldi, F., 2012. Design hydrograph estimation in small and ungauged watersheds: continuous simulation method versus event-based approach. *Hydrol. Process.* 26 (20), 3124–3134. <https://doi.org/10.1002/hyp.8384>.
- Hirschboeck, K.K., 1988. *Flood hydroclimatology*, Chapter 2 in Baker, V.R., Kochel, R.C. and Patton, P.C., eds., *Flood Geomorphology*, John Wiley & Sons, 27–49.
- Hu, Y., Quinn, C.J., Cai, X., Garfinkle, N.W., 2017. Combining human and machine intelligence to derive agents' behavioral rules for groundwater irrigation. *Adv. Water Resour.* 109, 29–40. <https://doi.org/10.1016/j.advwatres.2017.08.009>.
- Hundecha, Y., Parajka, J., Viglione, A., 2020. Assessment of past flood changes across Europe based on flood-generating processes. *Hydrol. Sci. J.* 65, 1830–1847. <https://doi.org/10.1080/02626667.2020.1782413>.
- Jiang, S., Zheng, Y., Wang, C., Babovic, V., 2022. Uncovering flooding mechanisms across the contiguous United States through interpretive deep learning on representative catchments. *Water Resour. Res.* 58. <https://doi.org/10.1029/2021WR030185> e2021WR030185.
- Kampf, S.K., Lefsky, M.A., 2016. Transition of dominant peak flow source from snowmelt to rainfall along the Colorado front range: Historical patterns, trends, and lessons from the 2013 Colorado front range floods. *Water Resour. Res.* 52 (1), 407–422. <https://doi.org/10.1002/2015WR017784>.
- Kerr, T., 2013. The contribution of snowmelt to the rivers of the South Island, New Zealand. *J. Hydrol. (new Zealand)* 52 (2), 61–82.
- Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.B., Gróf, G., Ho, H.L., Hong, H., Chapi, K.I., Prakash, I., 2019. A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. *J. Hydrol.* 573, 311–323. <https://doi.org/10.1016/j.jhydrol.2019.03.073>.
- Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., Nearing, G., 2019. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrol. Earth Syst. Sci.* 23, 5089–5110. <https://doi.org/10.5194/hess-23-5089-2019>.
- Lu, W., Lei, H., Yang, W., Yang, J., Yang, D., 2020. Comparison of floods driven by tropical cyclones and monsoons in the southeastern coastal region of China. *J. Hydrometeorol.* 21, 1589–1603. <https://doi.org/10.1175/JHM-D-20-0002.1>.
- Macara, G.R., 2016. The climate and weather of West Coast. *National Institute of Water and Atmospheric Research (NIWA) Sci. Technol. Series* 72, 40 p.
- Mallakpour, I., Villarini, G., 2015. The changing nature of flooding across the central United States. *Nat. Clim. Chang.* 5, 250–254. <https://doi.org/10.1038/nclimate2516>.
- McMillan, H.K., Booker, D.J., Cattoën, C., 2016. Validation of a national hydrological model. *J. Hydrol.* 541, 800–815. <https://doi.org/10.1016/j.jhydrol.2016.07.043>.
- Merz, B., Basso, S., Fischer, S., Lun, D., Blöschl, G., Merz, R., Guse, B., Viglione, A., Vorogushyn, S., Macdonald, E., Wietzke, L., Schumann, A., 2022. Understanding heavy tails of flood peak distributions. *Water Resour. Res.* 58. <https://doi.org/10.1029/2021WR030506> e2021WR030506.
- Mosley, M.P., 1981. Delimitation of New Zealand hydrologic regions. *J. Hydrol. (NZ)* 49, 173–192.
- Newsome, P.F.J., Wilde, R.H., Willoughby, E.J., 2000. Land resource information system spatial data layers. Landcare Research, Technical report, Palmerston North, New Zealand, 74 p.
- Nied, M., Pardowitz, T., Nissen, K., Ulbrich, U., Hundecha, Y., Merz, B., 2014. On the relationship between hydro-meteorological patterns and flood types. *J. Hydrol.* 519, 3249–3262. <https://doi.org/10.1016/j.jhydrol.2014.09.089>.
- Oppel, H., Schumann, A.H., 2020. Machine learning based identification of dominant controls on runoff dynamics. *Hydrol. Process.* 34 (11), 2450–2465. <https://doi.org/10.1002/hyp.13740>.
- Pham, B.T., Phong, T.V., Nguyen, H.D., Qi, C., Al-Ansari, N., Amini, A., Ho, L.S., Tuyen, T.T., Yen, H.P.H., Ly, H.-B., Prakash, I., Bui, D.T., 2020. A comparative study of Kernel Logistic Regression, radial basis function classifier, multinomial naive bayes, and logistic model tree for flash flood susceptibility mapping. *Water* 12, 239. <https://doi.org/10.3390/w12010239>.
- Queen, L.E., Dean, S., Stone, D., Henderson, R., Renwick, J., 2023. Spatiotemporal trends in near-natural New Zealand river flow. *J. Hydrometeorol.* 24, 241–255. <https://doi.org/10.1175/JHM-D-22-0037.1>.
- Rentschler, J., Salhab, M., Jafino, B.A., 2022. Flood exposure and poverty in 188 countries. *Nat. Commun.* 13, 3527. <https://doi.org/10.1038/s41467-022-30727-4>.
- Rossi, F., Fiorentino, M., Versace, P., 1984. Two-component extreme value distribution for flood frequency analysis. *Water Resour. Res.* 20 (7), 847–856. <https://doi.org/10.1029/WR020i007p00847>.
- Schmidt, L., Heße, F., Attinger, S., Kumar, R., 2020. Challenges in applying machine learning models for hydrological inference: a case study for flooding events across Germany. *Water Resour. Res.* 56 (5). <https://doi.org/10.1029/2019WR025924> e2019WR025924.
- Singh, V.P., 1997. Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrol. Process.* 11 (12), 1649–1669. [https://doi.org/10.1002/\(SICI\)1099-1085\(19971015\)11:12<1649::AID-HYP495>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1099-1085(19971015)11:12<1649::AID-HYP495>3.0.CO;2-1).
- Singh, S.K., Pahlow, M., Booker, D.J., Shankar, U., Chamorro, A., 2019. Towards baseflow index characterization at national scale in New Zealand. *J. Hydrol.* 568, 646–657. <https://doi.org/10.1016/j.jhydrol.2018.11.025>.
- Snelder, T.H., Biggs, B.J.F., 2002. Multi-scale river environment classification for water resources management. *J. Am. Water Resour. Assoc.* 38, 1225–1240. <https://doi.org/10.1111/j.1752-1688.2002.tb04344.x>.
- Snelder, T.H., Detry, T., Lamouroux, N., Larned, S.T., Sauquet, E., Pella, H., Catalogne, C., 2013. Regionalization of patterns of flow intermittence from gauging station records. *Hydrol. Earth Syst. Sci.* 17, 2685–2699. <https://doi.org/10.5194/hess-17-2685-2013>.

- Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. Chapter 18 in: David R. Maidment (editor): Handbook of Hydrology. McGraw-Hill, New York.
- Stein, L., Pianosi, F., Woods, R., 2020. Event-based classification for global study of river flood generating processes. *Hydrol. Process.* 34 (7), 1514–1529. <https://doi.org/10.1002/hyp.13678>.
- Stein, L., Clark, M.P., Knoben, W.J.M., Pianosi, F., Woods, R.A., 2021. How do climate and catchment attributes influence flood generating processes? a large-sample study for 671 catchments across the contiguous USA. *Water Resour. Res.* 57. <https://doi.org/10.1029/2020WR028300> e2020WR028300.
- Tait, A., Henderson, R., Turner, R., Zheng, X., 2006. Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface. *Int. J. Climatol.* 26 (14), 2097–2115.
- Tarasova, L., Merz, R., Kiss, A., Basso, S., Blöschl, G., Merz, B., Viglione, A., Plötner, S., Guse, B., Schumann, A., Fischer, S., Ahrens, B., Anwar, F., Bárdossy, A., Bühler, P., Haberlandt, U., Kreibich, H., Krug, A., Lun, D., Müller-Thomy, H., Pidoto, R., Primo, C., Seidel, J., Vorogushyn, S., Wietzke, L., 2019. Causative classification of river flood events. *WIREs Water* 6, e1353.
- Totaro, V., Gioia, A., Kuczera, G., Iacobellis, V., 2024a. Modelling multidecadal variability in flood frequency using the two-component extreme value distribution. *Stoch. Env. Res. Risk A.* 1–18. <https://doi.org/10.1007/s00477-024-02673-8>.
- Totaro, V., Kuczera, G., Iacobellis, V., 2024b. Goodness-of-fit, identifiability and extrapolation: Can the Two-Component Extreme Value distribution be used in at-site flood frequency analysis? *J. Hydrol.* 640, 131590. <https://doi.org/10.1016/j.jhydrol.2024.131590>.
- Turkington, T., Breinl, K., Ettemaa, J., Alkema, D., Jettena, V., 2016. A new flood type classification method for use in climate change impact. *Weather Clim. Extremes* 14, 1–16. <https://doi.org/10.1016/j.wace.2016.10.001>.
- Waylen, P.R., Woo, M., 1982. Prediction of annual floods generated by mixed processes. *Water Resour. Res.* 18, 1283–1286. <https://doi.org/10.1029/WR018i004p01283>.
- Xu, T., Liang, F., 2021. Machine learning for hydrologic sciences: an introductory overview. *Wiley Interdiscip. Rev. Water* 8 (5). <https://doi.org/10.1002/wat2.1533> e1533.
- Yan, L., Xiong, L., Ruan, G., Xu, C.Y., Yan, P., Liu, P., 2019. Reducing uncertainty of design floods of two-component mixture distributions by utilizing flood timescale to classify flood types in seasonally snow covered region. *J. Hydrol.* 574, 588–608. <https://doi.org/10.1016/j.jhydrol.2019.04.056>.
- Yan, L., Zhang, L., Xiong, L., Yan, P., Jiang, C., Xu, W., Xiong, B., Yu, K., Ma, Q., Xu, C.Y., 2023. Flood frequency analysis using mixture distributions in light of prior flood type classification in Norway. *Remote Sens. (Basel)* 15 (2), 401. <https://doi.org/10.3390/rs15020401>.
- Zolghadr-Asli, B., 2023. No-free-lunch-theorem: a page taken from the computational intelligence for water resources planning and management. *Environ. Sci. Pollut. Res.* 30, 57212–57218. <https://doi.org/10.1007/s11356-023-26300-1>.