

Critical assessment of metrics and methods used to quantify temporal loading of rainfall events

Molly Asher^{1,2}, Mark A. Trigg¹, Cathryn E. Birch², Rasmus L.T. Henriksen³, Steven J. Böing^{3,4}, and Jonas W. Pedersen^{3,4}

¹School of Civil Engineering, University of Leeds, Leeds, UK

²Institute for Climate and Atmospheric Science, University of Leeds, Leeds, UK

³Weather research, Danish Meteorological Institute. Sankt Kjelds Plads 11, 2100 Copenhagen O, Denmark.

⁴DTU Sustain, Technical University of Denmark. Bygningstorvet Building 115, 2800 Kgs. Lyngby, Denmark

Correspondence: Molly Asher (gy17m2a@leeds.ac.uk)

Abstract. The distribution of rainfall ~~over a storm's duration, known as the~~ intensity through time at a fixed spatial location, referred to here as event temporal loading, can significantly influence hydrological and geomorphological responses, including run-off generation, urban flood risk, and soil erosion. ~~A wide range of~~ Numerous approaches have been developed to analyse rainfall event temporal loading, but these differ in how they characterise rainfall behaviour and in the aspects of storm structure they emphasise. ~~Early~~ Emerging research further suggests that climate change may alter rainfall temporal loading in complex and regionally dependent ways, underlining the importance of clear and consistent approaches to its quantification. In this study, we identify ~~52 metrics that have been~~ 48 metrics which have been previously applied to describe event temporal loading, and ~~categorise them as classification metrics, summary statistics, or intermittency~~ define a further five metrics representing aspects not fully captured in existing metrics. We calculate these metrics for 233,128 rainfall events recorded at Danish rain gauges, ~~and demonstrate that, while some metrics are robust to changes in rainfall event temporal resolution and pre-processing, others are highly sensitive. Data-driven cluster analysis further reveals how various metrics relate to one another. We use data-driven cluster analysis to reveal how the metrics relate,~~ highlighting groups of metrics that ~~may be used interchangeably~~ describe similar properties, and others that ~~describe fundamentally different properties~~ are more distinct. Based on this, we conceptualise five aspects of temporal loading (~~mass timing, peak timing, magnitude concentration, temporal concentration, and intermittency~~) ~~and recommend metrics to quantify each. Overall,~~ We demonstrate that some metrics are robust to changes in rainfall event temporal resolution and pre-processing, while others are highly sensitive. Drawing on these findings, we recommend one representative metric per aspect: the 4th quartile mass fraction (or D_{50} if a continuous measure is preferred) for mass timing; peak position ratio for peak timing; the Gini coefficient for magnitude concentration; temporal standard deviation for temporal concentration; and the ~~study provides a foundation for more deliberate and informed metric selection, helping to align research questions with appropriate representations~~ wet-dry transition rate for intermittency. Together, these recommendations provide a practical framework for deliberate metric selection and more consistent cross-study comparison of rainfall temporal loading, ~~and offering a clearer basis for cross-study comparison.~~

1 Introduction

Rainfall events vary not only in ~~their~~ total volume and duration but also in how ~~rainfall~~ intensities fluctuate as ~~the storm~~ ~~develops~~ storms develop. Storms typically ~~evolve through distinct~~ progress through phases of initiation, intensification, and dissipation, reflecting the dynamic interplay of meteorological and cloud processes. ~~In this study, we~~ These processes give rise to storms with complex space-time structure, so that rainfall evolution differs across both time and location. We use the term ‘event temporal loading’ to describe ~~this internal variability in the internal variability of~~ intensity over the course of a storm ~~. Similar concepts have been described in the literature under various terms, including the~~ at a fixed location. This Eulerian perspective is conceptually straightforward and aligns with the conventional representation of storms in hydrological modelling, such as the design storm approach, where temporal variations are condensed into a single site-based hyetograph. Event temporal loading is described in previous research under various names, including intensity profile (Dunkerley, 2021a), storm profile (Kottegoda and Kassim, 1991), rainfall temporal pattern (Wang, 2020), and intra-event rainfall variability (Todisco, 2014). While it is possible in principle to analyse temporal loading from a moving-storm, or Lagrangian, perspective, systematic methods for doing so have not yet been established.

Rainfall event temporal loading has been shown to influence hydrological and geomorphological responses across a wide range of environmental processes, however the direction and magnitude of its effects often vary between studies. For instance, flood modelling experiments have demonstrated that temporal loading alone can produce flood depth differences of up to 35% under identical rainfall volumes (Hettiarachchi et al., 2018). In landslide studies, early-peaking storms have been associated with greater infiltration and more severe slope instability (Fan et al., 2020). Soil erosion studies exemplify the diverging outcomes in temporal loading studies. Wang et al. (2016) reported greater soil loss from later-peaking storms, while Aquino et al. (2013) found early-peaking events to be more damaging. Such differences underscore that while temporal loading is clearly influential, its specific impacts depend on system characteristics, dominant processes, and methodological approaches.

Extreme rainfall events are expected to intensify in a warming climate at around 6–7% per degree Celsius, following the Clausius–Clapeyron (CC) relationship (Trenberth et al., 2003). Furthermore, for short-duration convective extremes, some studies suggest even greater rates of intensification, known as super-CC scaling (Fowler et al., 2021). This behaviour has been linked to dynamical storm changes under higher temperatures, including variations in storm speed, convective cell development, and feedbacks such as latent heat release (Lenderink and Van Meijgaard, 2009; Fowler et al., 2021). These dynamical storm processes have also been suggested to affect the temporal distribution of rainfall within storms. A growing body of work has begun to examine how event temporal loading may change in a warmer climate (Visser et al., 2023; Ghanghas et al., 2024; Asher et al., 2025a; Wasko and Sharma, 2015), but results to date show variable magnitudes and directions of change, pointing to possible regional dependencies in behaviour. Research in this area remains at an early stage and is developing actively.

Despite its demonstrated importance, rainfall event temporal loading is often ~~oversimplified or overlooked~~ simplified in impact modelling applications, particularly in design-oriented or regulatory contexts. This simplification is most commonly achieved using design storms, which are synthetic profiles used to standardise the representation of extreme events across specified return periods, and a longstanding component of engineering practice (Watt and Marsalek, 2013). Design storms

combine an estimate of total rainfall with a prescribed hyetograph, which dictates how rainfall is distributed over the course of the storm. While total rainfall depths are typically derived from locally specific statistics, the same hyetograph shape is often applied universally. While high-resolution data (e.g., 1–5 minutes) are ideal for capturing the short, intense bursts of rainfall critical for urban hydrology (Schilling, 1991; Einfalt et al., 2004; Berne and Krajewski, 2013), in reality many observational precipitation datasets are provided at lower resolutions. Rainfall impact modelling studies also frequently rely on derived or synthetic events. Here idealised or averaged storm profiles tend to prioritise capturing total volume and duration, at the expense of sub-event dynamics such as peak intensity timing or concentration. For example, symmetrical, centrally peaked intensity profiles are commonly used in flood modelling, such as the FEH design profiles in the UK (Centre for Ecology & Hydrology, 1999) and the Chicago design storm used in the US and other countries (Keifer and Chu, 1957). As a result, the full range of possible hydrological responses may not be represented, leading to underestimation of peak flows, misrepresentation of drainage performance, and failure to identify critical scenarios, compromising risk assessments, infrastructure design, and climate adaptation planning.

An alternative approach in impact modelling is continuous simulation, in which long-term rainfall time series are routed through models to generate extended sequences of runoff responses that are subsequently analysed to estimate flood magnitudes for given return periods (Boughton and Droop, 2003). This approach inherently accounts for antecedent wetness conditions and avoids many ad hoc assumptions about initial losses, but requires extensive data and substantially greater computational effort. Considerable research has examined the temporal and spatial resolution needed to capture storm dynamics relevant to urban hydrology, with several studies suggesting that sub-hourly (often 1–5 minute) data are required to resolve the short, intense bursts that drive urban response (Einfalt et al., 2004; Schilling, 1991). Radar products and stochastic rainfall generators have therefore been developed to provide such high-resolution inputs. Nonetheless, continuous simulation remains less commonly used in routine flood risk assessment and infrastructure design than event-based methods, which continue to be favoured for their simplicity, interpretability, and established role in practice (Boughton and Droop, 2003; Pathiraja et al., 2012; Formetta et al., 2018; K

Robust, interpretable metrics are therefore essential to bridge the gap between observed rainfall behaviour and the simplified forms commonly used in models. By quantifying key aspects of temporal loading, such metrics help preserve important features when rainfall events are simplified, improving the fidelity of hydrological representations. Furthermore, metrics can offer insight into why the influence of temporal loading varies across studies and contexts, by revealing which specific aspects of storm evolution are most relevant to particular hydrological responses. A Flood response, for instance, may be especially sensitive to peak intensity and its timing, whereas soil erosion and landslide initiation may respond more strongly to rainfall concentration or intermittency. Despite this, a wide variety of metrics have been proposed and used to describe event temporal loading. However, there has been with little systematic evaluation of how these metrics they relate to one another, how sensitive they are to temporal resolution and data processing choices, or whether they are broadly suitable across hydrological and climatological contexts. There is also a lack of clear terminology: metrics often target different contexts. Terminology is also inconsistent, with metrics often targeting different characteristics, such as asymmetry, peakiness, or intermittency, but these distinctions are not always explicit. The relevance of each aspect varies by discipline, and borrowing metrics across

~~domains can result in~~ rarely made explicit. Applying metrics without regard to these differences risks misplaced emphasis or misinterpretation.

This study seeks to address gaps in knowledge on how temporal loading metrics perform under different data and processing conditions, and on how metrics relate to one another. The work addresses the following research questions:

- RQ1: What key properties of rainfall event temporal loading are commonly measured, and why?
- RQ2: Which metrics are strongly correlated, suggesting they may be redundant or are suitable for use in cross-comparison of studies?
- RQ3: How sensitive are these metrics to the temporal resolution of the rainfall data?
- RQ3RQ4: How does ~~de-dimensionalisation~~ conversion of rainfall events to Dimensionless Mass Curves (DMCs) affect metric values?
- ~~RQ4: Which metrics are strongly correlated, suggesting they may be redundant or are suitable for use in cross-comparison of studies.~~

The first research question is addressed through a structured literature review that identifies existing metrics used to quantify temporal loading, followed by a synthesis of their underlying assumptions, intended applications, and data requirements (Sect. 2). The remaining questions apply data from Danish rain gauges. The second ~~and third~~ question is tackled using cluster analysis to explore similarities among metrics calculated from raw rainfall events at a 5-minute resolution (Sect. 4.1). The third and fourth questions are examined via comparative analyses of metrics calculated on rainfall events at varying temporal resolutions (Sect. 4.1) and ~~with normalisation applied after calculation of DMCs~~ (Sect. 4.2). Finally, the last research question is tackled using cluster analysis to explore similarities among metrics calculated from raw rainfall events at a 5-minute resolution (Sect. 4.1). Through this empirical analysis, the study builds toward a practical framework for metric selection, identifying five distinct aspects of rainfall temporal loading and proposing a representative metric for each, with the aim of reducing inconsistency and improving comparability across future studies.

2 Literature review

2.1 ~~Search and selection strategy~~

The literature review ~~follows~~ is conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework (Moher et al., 2010). ~~The PRISMA framework.~~ This is a standardised method of performing literature review which aims to minimise potential researcher bias, ~~and to promote transparency and reproducibility in the identification and selection of relevant studies. The framework includes four key stages: (1) identification, boolean queries used to identify potential studies, duplicated studies removed; (2) screening, abstract level review to remove irrelevant studies;~~

(3) eligibility, full-text review against pre-defined inclusion/exclusion criteria; and (4) inclusion, relevant studies included for review, and the total number of studies reported.

Boolean queries are search strings using AND, OR and NOT operators, and are constructed here by combining different terms used in the literature to describe event temporal loading, with different terms used to describe rainfall events. The full queries used to search the Scopus, Web of Science, and Google Scholar databases are outlined in [Further methodological details are provided in](#) Appendix A. The searches were performed on the 15/02/2025 and returned 903 relevant records overall (including 181 on Scopus, 409 on Web of Science, and 313 on Google Scholar). 351 duplicated records were removed, leaving 552 studies for abstract level screening. After screening of abstracts, a further 390 studies were eliminated, leaving 162 papers for full review. [main outcomes of the literature review are reported in the sections that follow.](#)

Studies were included if they met the following criteria: (a) use of search terms in the title, abstract, or keywords; (b) published in English; (c) focus on individual rainfall events with data at sub-daily temporal resolution; (d) application of a method or metric to quantify temporal loading. Exclusion criteria were: (a) not published in English; (b) inaccessible or behind paywalls (to the authors, not all included studies are open-access); (c) lack of use of sub-daily rainfall data; (d) combined spatial and temporal analysis (where the temporal loading could not be separated); (e) focus on rainfall trends over extended periods (e.g., seasonal or monthly scales); (f) no application of a metric for quantifying within-event temporal variability.

After applying these criteria, a total of 59 studies remained. We note, however, that several additional studies met all criteria except the use of a metric to quantify temporal loading. These excluded studies fall into three main groups. The first concerns the disaggregation of daily rainfall data (Pampaloni et al., 2021). The second includes stochastic event generation approaches (Cha et al., 2024; Efstratiadis et al., 2022; Kao and Govindaraju, 2008; Nguyen and Chen, 2022; Oriani et al., 2018; Pampaloni et al., 2021). The third involves applications of synthetic rainfall events with prescribed temporal patterns to investigate impacts on soil erosion (De Lima and Singh, 2002; de Lima et al., 2013; Frauenfeld and Truman, 2004; Gao et al., 2018; Parsons and Stone, 2006; Wang, 2018), landslides (Zhao et al., 2025), and herbicide runoff (Zhang et al., 1997). In these cases, event profiles are described using broad qualitative labels (e.g. linearly increasing, stepped increasing-decreasing, uniform) rather than being quantified with explicit metrics.

2.1 Domains and purposes of reviewed literature

~~Research on~~ [Research into](#) rainfall temporal loading ~~spans a range of domains concerned with~~ [is conducted across domains where](#) hydrological and geomorphological responses to precipitation. The most prominent domain is flood modelling, where studies range from rainfall-runoff models in large watersheds (Cai et al., 2024), to the hydraulic performance of urban sewer systems (Li et al., 2021; Müller et al., 2017). Other domains include soil erosion and sediment transport (Römkens et al., 2002; Liu et al., 2018), [rainfall are important, including flood and hydrological research](#), pollution modelling (Fu et al., 2021), rainfall-induced landslides (Fan et al., 2020), and climate change impact assessments on extreme rainfall characteristics (Visser et al., 2023; Hu et al., 2018).

~~Despite the range of,~~ [and soil erosion and landslide studies. Across these](#) contexts, research is generally motivated by similar questions. One common aim is to assess system sensitivity to rainfall temporal loading. ~~This is typically performed by running~~

155 ~~models with rainfall events of~~ Running models using rainfall events with fixed total volume and duration but varying temporal
~~distributions, before examining the resulting impacts on, for example, peak flow, runoff volume, or pollutant distribution allows~~
~~for impacts to be examined. Research of this nature has been performed in regards to: runoff generation in urban catchments~~
~~(Cai et al., 2024; Asher et al., 2025b), the hydraulic performance of urban sewer systems (Li et al., 2021), modelling of soil~~
~~erosion and sediment transport (Römkens et al., 2002; Liu et al., 2022; Dunkerley, 2021a; Rivera et al., 2012; Gao et al., 2024; Wang et al.~~
160 ~~, pollution wash-off (e.g. Asher et al. (2025b)). While these studies treat the system as the subject of analysis (i.e., asking how~~
~~sensitive a model or process is to rainfall structure), others assume system sensitivity and focus directly on the characterisation~~
~~(Fu et al., 2021), and rainfall-induced landslides (Fan et al., 2020)).~~

A second common aim is to characterise the temporal loading of rainfall events ~~themselves~~. This includes research developing
new or more nuanced metrics for describing internal rainfall variability (Todisco, 2014; Dunkerley, 2022), or research applying
165 existing metrics in order to group and summarise events into representative hyetographs for specific locations ~~, e.g., for China by~~
~~Wang (2020), Iran by Amin et al. (2000), Slovenia by Dolšak et al. (2016), Greece by Vantas et al. (2019), amongst numerous~~
~~others. Such work is often intended to improve development of design storms or evaluate (Wang, 2020; Amin et al., 2000; Dolšak et al., 201~~
~~. This allows assessment of~~ how well standard hyetographs reflect observed temporal patterns. ~~Finally, , and the improvement~~
~~of design storms. A third aim of research on temporal loading is to improve representation of the statistical characteristics of~~
170 ~~rainfall~~ in stochastic rainfall generation and disaggregation ~~, metrics help ensure that synthetic rainfall sequences reproduce~~
~~realistic intra-event patterns, preserving statistical characteristics critical for sub-daily modelling. Across all these purposes, the~~
~~consistent goal is to represent temporal loading more accurately in a way that enhances understanding, modelling performance,~~
~~and decision-making (Wu et al., 2006; Nguyen and Chen, 2022). A final aim is to assess the impact of climate change on~~
~~extreme rainfall characteristics (Visser et al., 2023; Hu et al., 2018; Asher et al., 2025a).~~

175 2.2 Rainfall data sources

The reviewed studies rely on two main sources of rainfall data: observational records ~~or and~~ synthetically generated events.
The majority of studies use events extracted from observed rainfall time series, typically recorded by rain gauges ~~, (e.g., Zhang~~
~~et al. (2023)), or radar , (e.g., Urgilés et al. (2024)). Events are extracted from observational records at point locations~~
on assumptions about what constitutes the start and end of a rainfall event, such as minimum intensity thresholds or inter-event
180 dry periods (Restrepo-Posada and Eagleson, 1982). ~~However~~ Additionally, event boundaries and characteristics are inevitably
shaped by the temporal resolution of the observational dataset, and so ‘observed’ rainfall events remain approximations shaped
by the resolution of the measurement system, rather than exact representations of storm behaviour.

Other studies ~~on event temporal loading~~ use synthetic rainfall events, ~~generated for controlled experimentation or scenario~~
~~testing (e.g., Cai et al. (2024), Bezak et al. (2018), Fan et al. (2020)). These synthetic events typically fall into two broad~~
185 ~~categories: idealised design storms and stochastically-generated events. Idealised storms take simple geometric forms. Stochastic~~
~~rainfall generators produce events which reproduce the statistical properties of observed rainfall while allowing control over~~
~~event magnitude and duration (Cai et al., 2024; Fan et al., 2020). These events are typically used within modelling frameworks~~
~~to explore a broader range of rainfall scenarios. While this approach enables controlled investigation of temporal effects, it~~

relies on assumptions about what constitutes a ‘representative’ event. This can influence interpretation and limit generalisation to real world rainfall. Idealised design storms are an additional form of synthetic events. These adopt simple geometric profiles, such as uniform rainfall, or triangular profiles (e.g., the rational method (Smith and Lee, 1984) or triangular hyetographs peaking early, centrally, or late in the event (e.g., Parsons and Stone, 2006), and. These profiles are used to systematically explore the influence of temporal loading while abstracting from site-specific variability. These types of events are used in the studies noted in Section ?? but not included in our review as they do not quantify event temporal loading with a metric. In these studies, the idealised profiles often form, often forming the basis for physical experiments, where pre-defined rainfall patterns are applied to laboratory set-ups simulating hillslopes or infiltration systems (Wang et al., 2023; de Lima et al., 2013; Parsons and Stone, 2006; Zhang et al., 1997). In contrast, stochastically generated events, produced using rainfall generators, aim to replicate the statistical properties of observed rainfall while allowing control over event magnitude and duration. The synthetic events are typically applied within modelling frameworks to explore a broader range of possible rainfall scenarios. While use of synthetic rainfall enables controlled investigation of temporal effects, it necessarily involves assumptions about what constitutes a “representative” event. These assumptions can affect the interpretation of results, particularly when attempting to generalise findings to real-world conditions. Idealised design storms are pivotal in a lot of the early work on understanding the influence of rainfall temporal loading, however, as temporal structure is prescribed rather than quantified using a metric they are not discussed further here.

2.3 Rainfall processing choices

Research varies not only in the type of rainfall data used, but also in how that data is processed before applying Rainfall temporal loading metrics. In some studies, metrics are may be applied directly to the a raw rainfall intensity series. This, which retains the original units and absolute values of the event, allowing metrics to reflect both timing and intensity. In contrast, other studies use double normalised representations, known as However, metrics are sometimes instead applied to dimensionless mass curves (DMCs) or Huff curves (Huff, 1967), in which both. DMCs were originally developed to characterise within-storm temporal structure independently of event magnitude and duration, and have since been widely used in rainfall analysis and design storm construction (Koutsoyiannis and Foufoula-Georgiou, 1993). They are derived by double normalising a rainfall event, scaling cumulative rainfall and event duration are scaled between 0 % and 100%. Interpolation is often applied at regular intervals (e.g., and expressing the event as a cumulative mass profile. To enable comparison across events with different durations and data resolutions, these dimensionless profiles are typically interpolated at regular fractions of storm duration (e.g. every 1% or 10% of the storm duration) to enable comparison between events), producing standardised curves of fixed length (Pan et al., 2017).

Normalised representations are widely used in studies seeking to generalise across events or derive representative storm profiles. By removing differences in scale and duration, DMCs allow researchers to isolate temporal loading and compare the shapes of events more directly across events of different magnitude and durations. In some cases, mean DMCs are computed for specific regions by averaging across many events (see Sect. 2.1). The widespread use of DMCs in the literature reflects the need for a standardised representation of rainfall temporal structure that enables direct comparison across events with

different durations, magnitudes and numbers of observations. However, the combined effects of double normalisation and interpolation also modify the internal structure of rainfall events, with implications for the behaviour and interpretation of temporal loading metrics. Metrics based on absolute values (e.g. peak intensity) lose their physical units when applied to DMCs, instead reflecting fractions of total rainfall per bin rather than mm/hr. This makes direct numerical comparison with raw equivalents meaningless, though the metrics remain defined and carry information about event shape. Metrics describing timing may become discretised, such that values are constrained to the interpolation intervals (e.g. time of peak occurring at 0, 10, 20%, etc. of event duration). As a result, conclusions drawn from metrics applied to raw rainfall and to DMCs may differ. These regional profiles are used to justify the use of particular design storms or to critique existing design standards. Other studies use DMCs as a basis for classifying events, for instance Huff curves are typically labelled according to the quartile of the Huff curve containing the most rainfall.

Despite the widespread use of DMCs, the effects of normalisation on metric outcomes are often not clearly articulated. Some studies report metrics calculated on both raw and normalised data, but many do not specify the format used at all. This, studies do not always make explicit whether temporal metrics are calculated using raw rainfall series or DMCs, nor do they consistently discuss the implications of this choice. This lack of transparency complicates comparisons across studies and makes it difficult to assess whether observed patterns reflect intrinsic features of the events or determine whether reported differences in temporal structure reflect intrinsic storm behaviour or artefacts of data processing.

2.4 Typology of metrics

The wide range of metrics identified in the reviewed studies are listed in Tables F1, ?? and ?? in Appendix F. Based on our own conceptual understanding of the metrics, and their description in the literature, we have assigned each of them to one of three metric types: classification metrics, summary statistics, and intermittency metrics. The categories and how they are delineated is described below.

2.4.1 Classification metrics

Classification metrics assign discrete labels to rainfall events. A common approach categorises events by the fraction of the storm that contains the most significant portion of rainfall, typically a third (Horner and Jens, 1942; Aquino et al., 2013; Wang et al., 2016), quarter (Huff, 1967; Amin et al., 2000; Jun et al., 2021; Harshanth et al., 2023; Azli and Rao, 2010; Amponsah et al., 2019; Wartalska and , or fifth (Villalobos Herrera et al., 2023; Ng et al., 2001; Asher et al., 2025b). While simple, the outcomes of such classifications depend heavily on implementation choices. For example, Horner and Jens (1942) introduced the original third-based classification scheme, defining events as advanced, intermediate, or delayed based on the timing of the peak intensity. By contrast, Liang et al. (2023) also use thirds but classify events based on the location of Table 1. Each metric is categorised as either a categorical or continuous metric. In cases where the same metric formulation is referred to by different names, we list both. In the remainder of the rainfall mass, identifying the third with the highest total rainfall. Further refinements to this method classify events based on the third containing a summary statistic, e.g. the centre of gravity or D_{50} (see Sect. ??). Figure 1 provides an illustrative example of some classification metrics. work, we refer to the metric only by the first name listed in Table 1.

Illustration of some classification metrics for an example rainfall event: (a) the raw rainfall data, and classification of the event based on (b) *3rd with peak* (Aquino et al., 2013; Fatone et al., 2021; Pinheiro et al., 2018; Wang et al., 2016), (c) *5th with most rainfall* (Villalobos Herrera et al., 2023), (d) *3rd with D_{50}* (Visser et al., 2023).

More complex classification schemes compare the dimensionless mass curve of an event to that of a uniform storm. For instance, Terranova and Iaquina (2011) assign a four-digit Binary Shape Code (BSC) based on

Table 1: Overview of metrics based on literature review

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>3rd/4th/5th with peak</u>	<u>Finds the fraction of event with the peak rainfall</u>	<u>Categorical</u>
<u>3rd/4th/5th with most</u>	<u>Finds the fraction of event with the most rainfall</u>	<u>Categorical</u>
<u>3rd with D_{50}</u>	<u>Finds the third in which 50% of cumulative rainfall is reached (see D_{50} definition later in Table 1)</u>	<u>Categorical</u>
<u>3rd with CoG</u>	<u>Finds the third containing the centre of gravity of rainfall</u>	<u>Categorical</u>
<u>Binary Shape Code</u>	<u>Measures whether cumulative rainfall in each quartile exceeds that of a uniform distribution. Kottegoda and Kassim (1991) use the crossing properties concept to classify events with a number based on, and uses this to assign a 4-digit code</u>	<u>Categorical</u>

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>Crossing properties</u>	<u>Measures</u> the number of times the event DMC crosses the uniform DMC ,and <u>a letter, which describes whether it exhibits a generally</u> and whether the event <u>has generally an</u> increasing or decreasing intensity profile. Overall, classification metrics simplify complex event structures into interpretable categories, and provide a tractable way to group storms by structural type. This is particularly useful in design and modelling contexts, where efficiency and consistent processes are important. However, their categorical nature may mask finer differences between events and may be sensitive to how the storm duration is divided. 2.4.1 Summary statistics Summary statistics are continuous metrics that quantify variation in rainfall intensity over time. We broadly group them into three types based on the way they are described in the literature: peakiness indicators, asymmetry indicators, and concentration indicators. Peakiness indicators capture how sharp or pronounced the intensity of peaks are relative to the rest of the storm. Metrics such as <i>maximum intensity</i>, <i>mean intensity</i>, and assigns a number and a letter based on this	<u>Categorical</u>
<u>I30 (the maximum rainfall depth accumulated in 30 minutes)</u>	<u>Identifies the max rainfall depth within</u> any 30-minute period)provide <u>baseline measures of peak magnitude and concentration. Other common examples include the <i>standard deviation</i>, <i>coefficient of variation</i>, <i>classical skewness</i>, and <i>classical kurtosis</i>, borrowed from general statistics, which reflect variability and the prominence of extremes across the distribution (Brommer et al., 2013; Yang et al., 2015; Kimura et al., 2014; García-Bartual and Andrés-Doménech, 2017; . A number of metrics specifically compare the peak with the mean. These include the <i>peak-to-mean ratio</i> (also called</u>	<u>Continuous</u>
<u>Classical skewness [(C) skewness]</u>	<u>Quantifies the asymmetry of the rainfall intensity value distribution. Positive values indicate many small intensities and a few large ones, while negative values indicate the reverse. Does not account for the temporal ordering of intensities. This is referred to as ‘Classical skewness’ to distinguish from ‘Temporal skewness’ which is introduced in Table 2</u>	<u>Continuous</u>

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>Classical kurtosis [(C) kurtosis]</u>	Quantifies the tailedness of the rainfall intensity irregularity (Wartalska et al., 2020)), and the relative amplitude , defined as the range divided by the mean (Gao et al., 2024; Liu et al., 2022). As the minimum intensity is often 0 mm/hr, and otherwise very close to 0, these two metrics are often equivalent. Similarly, value distribution . Positive values indicate a distribution with heavier tails and more extreme values relative to the variance, while negative values indicate lighter tails and fewer extremes. Does not account for the temporal ordering of intensities. This is referred to as ‘Classical kurtosis’ to distinguish from ‘Temporal kurtosis’ which is introduced in Table 2	<u>Continuous</u>

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
$m2$ Classical standard deviation [(C) std] compares the amount of rainfall in the highest intensity time step to the total event volume (Wartalska et al., 2020) , while zone-based metrics, such as the mean intensity within HZ (periods exceeding the event average), quantify how extreme the most active parts of the storm are relative to the overall distribution (Liu et al., 2022). Illustration of several summary statistics for an example rainfall event, including $m1$, $m2$, $m4$, $m5$ (Wartalska and Kotowski, 2020) , $T25$, $T50$, $T75$ (Knighton and Walter, 2016) , centre of mass and time to peak (Jun et al., 2021; Knighton and Walter, 2016) Lateral asymmetry indicators quantify whether rainfall is concentrated toward the beginning or end of an event. As with classification metrics, some indicators focus on the timing	Quantifies the dispersion of rainfall intensity values around the mean intensity. A higher value indicates greater variability in intensity magnitudes (i.e., larger deviations from the mean), while a lower value indicates more uniform intensities throughout the event. Does not account for the temporal ordering of intensities. This is referred to as ‘Classical std’ to distinguish from ‘Temporal std’ which is introduced in Table 2	Continuous

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>Coefficient of variation (CV)</u>	<u>Measures the variation of intensity values around the mean intensity by calculating the classical standard deviation divided by the mean</u>	<u>Continuous</u>
<u>Peak-mean ratio or Rainfall intensity irregularity (n_i)</u>	<u>Calculates the maximum intensity divided by mean intensity</u>	<u>Continuous</u>
time-to-peak <u>Relative amplitude and the peak-position-ratio</u> (also called the coefficient of peak-rainfall intensity) describe how early or late the maximum intensity occurs (Fu et al., 2021; Wartalska and Kotowski, 2020; Cai et al., 2024; Wartalska et al., 2020; Knighton and Walter, 2016)	<u>Calculates the range (maximum - minimum intensity) divided by the mean intensity</u>	<u>Continuous</u>
Skew_p <u>Time to peak</u> is also anchored to-	<u>Represents the time to peak in minutes</u>	<u>Continuous</u>
<u>Peak position ratio or Coefficient peak rainfall intensity</u>	<u>Calculates the time to peak divided by duration</u>	<u>Continuous</u>
<u>Skew_p</u>	<u>Measures the peak timing, but measures the relative position of the peak within the rainfall duration (Oh et al., 2024). Other metrics capture asymmetry across the entire rainfall mass. The-, and is therefore anchored to the peak timing. Takes a value from -0.25 to 0.25. Negative values indicate front-loading, positive values indicate back-loading</u>	<u>Continuous</u>
<u>T25 / T50 or D₅₀ / T75</u>	<u>Calculates the proportion of event duration elapsed when a specified percentage (e.g., 25%, 50%, 75%) of total rainfall has occurred</u>	<u>Continuous</u>

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>m1</u> metric compares the	<u>Calculates the relative</u> volume of rainfall before versus after the peak (Wartalska et al., 2020), while the <i>event loading index</i> similarly contrasts pre- and post-peak rainfall (Ghanghas et al., 2024). Percentile timing indicators such as the <i>centre of gravity</i> , D_{50} (Visser et al., 2023), and $T_{25\%}$, $T_{50\%}$, and $T_{75\%}$ (Knighton and Walter, 2016) describe how rainfall accumulates over time. Müller et al. (2017) introduce <i>asymmetry of dependence</i> to describe the degree to which an event timeseries is reversible. Mass distribution indicators	<u>Continuous</u>
<u>m2</u>	<u>Calculates the percentage of rainfall in the highest-intensity time step divided by rainfall in whole event</u>	<u>Continuous</u>

Metric	Meaning	Type
$m3$ $m4$ $m5$ express the proportion of rainfall in the	<u>Calculates the percentage rainfall in</u> first 33% , $m4$ 30% , and $m5$ 50% of the event duration (Wartalska et al., 2020), while Liu et al. (2022) calculate the proportion of rainfall in each quarter of the event (<i>fraction Q1</i>, etc) . A number of these asymmetry-based metrics are visualised in Figure 2. Concentration indicators describe how rainfall is clustered within an event. A key example is the <i>Precipitation Concentration Index (PCI)</i> , originally developed to assess monthly and seasonal rainfall distributions (Oliver, 1980) but also applied to individual events (He et al., 2022). Higher <i>PCI</i> values indicate a greater share of rainfall concentrated in fewer time steps. The <i>Temporal Concentration Index (TCI)</i> was introduced as a scale-independent alternative that could be used to compare across events of different durations (Long et al., 2021). It quantifies the deviation of an event 's temporal loading from that of a uniform rainfall event with the same duration and volume. Another related measure is the normalised root-mean-square error around the peak ($NRMSE_p$), which assesses how tightly rainfall is clustered near the peak time step (Oh et al., 2024). The <i>Gini coefficient</i>, originally developed to measure income inequality, has also been adapted to assess the unevenness of rainfall distribution within an event (Lu et al., 2025). Applied to rainfall, it reflects how much cumulative rainfall deviates from a perfectly even distribution. For highly concentrated events , the accompanying <i>Lorenz asymmetry coefficient</i> describes whether this arises from a few extreme high-intensity bursts or from many moderately intense intervals, both of which could lead to high concentration. Zone-based indicators, including the <i>% of rainfall in high-intensity zone (HIZ)</i>, <i>time in high-intensity zone (HIZ)/ low-intensity zone (LIZ)</i>, indicate how long the storms stays intense and how much rainfall is concentrated in these intense bursts (Liu et al., 2022). 2.4.1 Intermittency metrics Intermittency metrics are designed to capture the presence and frequency of dry periods within rainfall events. These metrics are particularly relevant in contexts where wet-dry fluctuations affect hydrological processes such as infiltration, soil saturation, and flash flood generation. Although fewer in number than other metric types, intermittency measures add an important dimension to the characterisation of temporal loading by highlighting fluctuations that may not be captured by peak-based or cumulative statistics. The event-dry ratio calculates the proportion of time steps within an event that record no rainfall as a proportion of the total number of recorded intervals in the	Subcategory Detail Number of studies <u>Continuous</u>

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>Event loading index</u>	Characterise rainfall profiles in an area <u>Classifies the temporal variability (captured by the standardised temporal heterogeneity, STH) of a storm by comparing it against that of a symmetrised version constructed by mirroring the rising limb around the peak</u>	23 <u>Continuous</u>
(+) 2-5		
<u>Asymmetry of dependence</u>	<u>Classifies the degree of reversibility in a timeseries using ranked-based statistics (copulas)</u>	Flooding/hydrology <u>42</u> <u>Continuous</u>
<u>Precipitation Concentration Index (PCI)</u>	<u>Compares sum of squared intensities to square of total rainfall. Higher values indicate more rainfall concentrated in fewer steps</u>	Landslide—1 <u>Continuous</u>
<u>Temporal Concentration Index (TCI)</u>	<u>Measures how strongly rain is clustered in time around each time-step, selects value for the time step around which rainfall is most tightly clustered. Higher values indicate stronger temporal concentration</u>	Pollution—2 <u>Continuous</u>
<u>NRMSE_p</u>	Soil erosion 13 (1)2-5 Climatic study 5 Stochastic event-generation 3 Evaluating satellite data <u>Quantifies the concentration of rainfall near the peak by calculating the RMSE from the peak, normalised by the rainfall total. Values range from 0 (uniform distribution) to 1 (all rainfall in a single time step)</u>	<u>Continuous</u>
<u>Gini coefficient</u>	Metric development <u>Measures the inequality of the rainfall distribution. Higher values indicate greater concentration of rainfall in fewer intervals</u>	4 <u>Continuous</u>
<u>Lorenz asymmetry coefficient</u>	Classification <u>Quantifies the skewness of the rainfall intensity distribution relative to the mean, distinguishing whether the concentration arises from a few extreme high-intensity bursts or from many moderately intense intervals, even if overall inequality is similar</u>	33 <u>Continuous</u>
	<u>The High Intensity Zone (HIZ) is defined as the period of the event of above-average intensity; The LIZ is the period of the event below average intensity</u>	Summary stats 25-

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>Mean intensity in HIZ</u>	<u>Intermittency</u> Measures the average intensity in the HIZ	<u>2-Continuous</u>
<u>% time in LIZ / HIZ</u>	<u>Real</u> Measures the proportion of time in the LIZ and HIZ	<u>47</u> <u>Continuous</u>
<u>% rain in HIZ</u>	<u>Synthetic</u> Measures the proportion of rainfall in the HIZ. More rainfall in HIZ indicates a more concentrated event	<u>12</u> <u>Continuous</u>
<u>Frac. in Q1 / Q2 / Q3 / Q4</u>	<u>Raw-rainfall-events</u> Measures the percentage of total rainfall in each quarter	<u>24</u> <u>Continuous</u>
<u>Event dry ratio or Intermittency fraction</u>	<u>Dimensionless Mass Curves</u> Quantifies the proportion of event duration spent at 0 mm/hr	<u>28</u> <u>Continuous</u>
<u>Unclear 7-</u>		

Figure C1 presents the number of

2.4.1 Categorical metrics

Those metrics flagged as categorical in Table 1 assign rainfall events to discrete structural classes. Some complex schemes exist, such as the Binary Shape Code (Terranova and Iaquina, 2011) and Crossing Properties (Kottegoda and Kassim, 1991), both of which compare an event's DMC to a uniform reference profile. Most approaches, however, divide an event into equal fractions (e.g. thirds, quarters or fifths) and identify which fraction contains a defining feature such as the peak intensity or the largest rainfall volume.

While simple, the outcomes of these classification metrics depend heavily on implementation choices. For example, Horner and Jens (1994) introduced the original third-based classification scheme, defining events as advanced, intermediate, or delayed based on the timing of peak intensity. By contrast, Liang et al. (2023) also use thirds but classify events based on the location of rainfall mass, identifying the third with the highest total rainfall. Figure 1 illustrates how these conceptually similar approaches can assign different structural labels to the same event. The *3rd with peak* metric classifies the event as centred, since maximum intensity occurs in the middle third, while both *3rd with most* and *3rd with D₅₀* identify it as front-loaded, since the majority of cumulative rainfall mass falls within the first third. This has direct implications for the comparability of studies using categorical metrics. A study using *3rd with peak* might report that most events are centred, while a study using *3rd with most* applied to the same dataset might report that most events are front-loaded. Without careful documentation of which metric

was used and what it measures, such differences in reported event structure could be mistaken for genuine differences between study catchments or climates, when in fact they simply reflect differences in metric definition.

Additionally, by reducing a continuous temporal profile to a discrete label, categorical metrics inevitably mask within-class variability. Despite this, they offer a tractable way to group storms by type, reducing complex temporal profiles into interpretable structural classes. This is particularly valuable in design and modelling contexts where consistent, repeatable classification procedures are important.

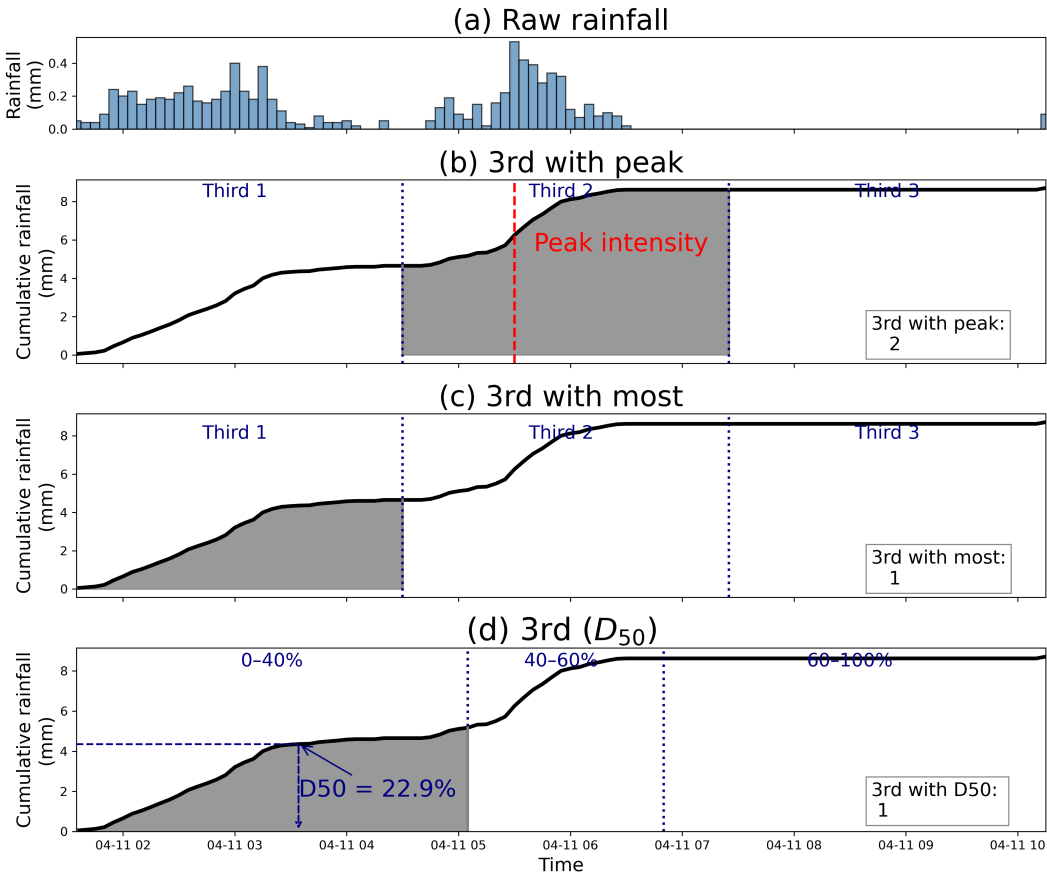


Figure 1. Illustration of some categorical metrics for an example rainfall event: (a) the raw rainfall data, and classification of the event based on (b) *3rd with peak*, (c) *3rd with most*, (d) *3rd with D_{50}* (NB: these are not even thirds)

2.4.2 Continuous metrics

While categorical metrics reduce temporal structure to a discrete label, continuous metrics quantify it using numerical indices. They describe structural properties such as the prominence of intense bursts, the timing of rainfall mass, the degree of concentration into short intervals, and the presence of intermittency. Many continuous metrics aim to quantify similar conceptual

properties, such as whether rainfall is early- or late-loaded, but differ in their mathematical construction and therefore in what they actually capture. The distinction between peak-based and mass-based definitions, already seen for the categorical metrics, reappears here. Some metrics define loading in relation to the timing of peak intensity (e.g. *peak position ratio*), while others use the timing of the bulk of rainfall mass (e.g. D_{50}). In the example event shown in Figure 2, these perspectives lead to similar interpretations, but this need not always be the case. An event with a short intense burst near the beginning but most of its rainfall occurring later would be classified as early-loaded by a peak-based metric and late-loaded by a mass-based one.

This tension between peak- and mass-based definitions, which run across both categorical and continuous metrics, reflect a broader point: rainfall temporal structure is multidimensional, and no single metric captures it fully. Metric choice therefore shapes interpretation, and different metrics can produce substantively different characterisations of the same event.

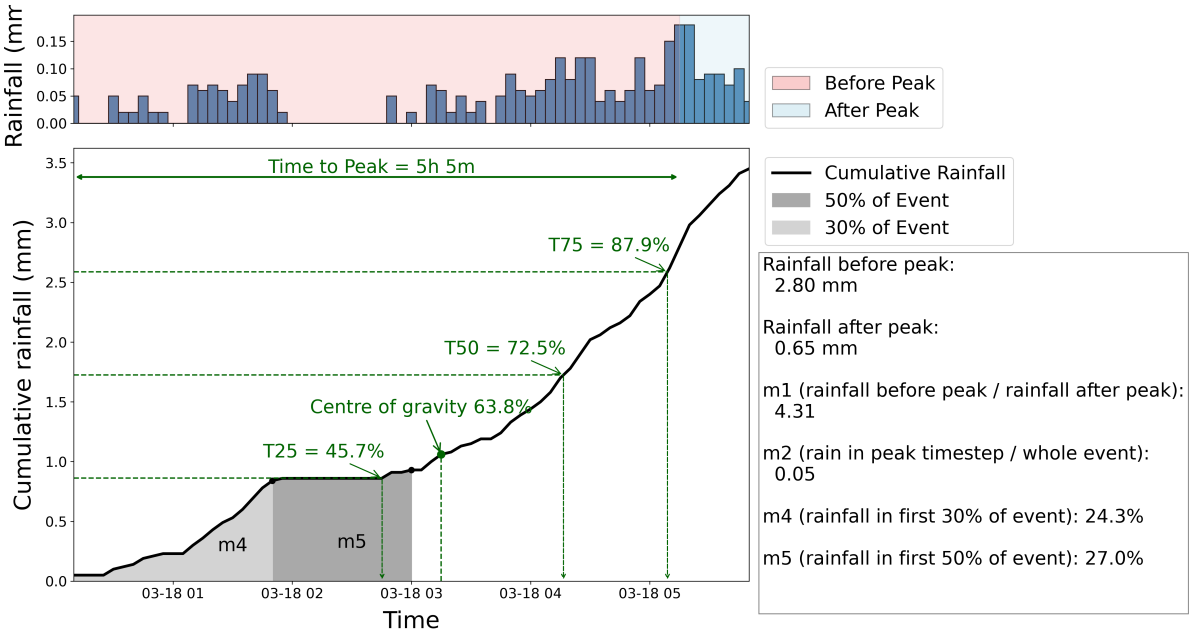


Figure 2. Illustration of several continuous metrics for an example rainfall event, including $m1$, $m2$, $m4$, $m5$ (Wartalska and Kotowski, 2020), $T25$, $T50$, $T75$ (Knighton and Walter, 2016), *centre of gravity* and *time to peak* (Jun et al., 2021; Knighton and Walter, 2016)

2.5 Summary of reviewed studies and metrics

A full overview of the reviewed studies applying each metric and highlights the geographical focus of those studies. The figure shows that most metrics have only been applied in one or two papers, with is provided in Table B1 (Appendix B). These studies span diverse domains, data sources, processing approaches, and metric types, which reflects the breadth of contexts in which temporal loading is studied. Table ?? (Appendix C) summarises the distribution of these characteristics across studies. Despite this diversity, a small number dominating of metrics dominate the literature. Classification metrics, which have been in use for

several decades, are the most prevalent. Among these, two interpretations As shown in Figure C1 (Appendix C), most metrics appear in only one or two studies. Notable exceptions are the 4th with most metric, which forms the basis of Huff quartiles (Huff, 1967), and the 4th with most and 4th 3rd with peak, first introduced in 1967, have been applied extensively across different regions. The 3rd with peak, derived from the earlier work of Horner and Jens (1947), is also frequently used, with a particular concentration of studies focusing on soil erosivity in Brazil metric, which underlies the advanced, intermediate, and delayed classification of Horner and Jens (1942). Both have been widely adopted and represent the earliest formal attempts to categorise storm temporal structure. This uneven distribution of metric use, combined with the sensitivity of even the most widely adopted metrics to methodological choices, points to the need for more deliberate and consistent metric selection, which is addressed in the subsequent sections.

~~Distribution of metrics across the reviewed studies, showing both the number of studies in which each metric was applied and the countries from which these studies were drawn.~~

3 Methods

3.1 Rainfall data

This research uses rainfall events extracted from Danish rain gauge data at 1 minute resolution. Figure 3 shows the national Danish rain gauge network, which is maintained and operated by the Danish Meteorological Institute (DMI). The individual gauges in the network are mainly owned by DMI and the Danish water utilities organised through the ‘Water Pollution Committee of The Society of Danish Engineers’ (abbreviated ‘SVK’ in Danish for ‘Spildevandskomiteen’). DMI’s gauges are weighing rain gauges of the brands Geonor and OTT Pluvio². The measurement resolution of Geonor gauges are 0.1 mm, while the resolution of the Pluvio² gauges started at 0.1 and was updated during the time period focused on in this study to 0.01 mm. SVK’s gauges are tipping bucket gauges by Rimco with a measurement resolution of 0.2 mm. Both gauge types record data with a temporal resolution of one minute. Figure 3 (left) shows that all DMI’s gauges have been operational for less than 15 years, while there is a much larger spread in time series lengths for the SVK gauges with the oldest ones providing continuous time series from 1979 to the present. Figure 3 (right) shows the spatial locations of the gauges with DMI gauges being relatively evenly spread across the country, while SVK gauges cluster around the major cities.

Data quality control for all gauges in the network are performed manually by DMI’s climatology department (DMI, 2025). Data points ruled non-trustworthy by manual quality control are excluded from this study.

3.2 Rainfall event extraction and pre-processing

For each rain gauge, independent rainfall events are extracted over the full period of available data. The rainfall time series is first aggregated to 5-minute resolution. This resolution is deemed to best preserve temporal detail, while minimising noise due to the measurement resolution of the rain gauge data. This noise arises because both tipping-bucket gauges and weighing

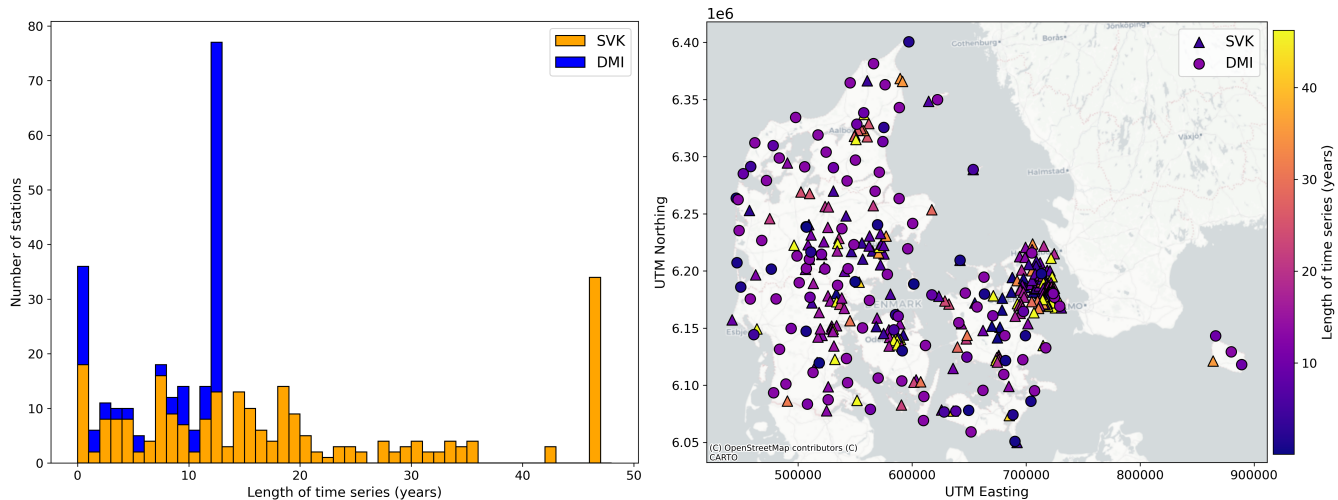


Figure 3. A stacked histogram of the length of the time series for rain gauges (left), and a map of Denmark showing the locations of each gauge, including indications of gauge ownership and time series length (right). Base map: © OpenStreetMap contributors 2025. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

gauges record rainfall in discrete increments, when the buckets tip or when the accumulated weight flips a decimal, which leads to many small, artificial peaks in the 1-minute time series.

To ensure event independence, we extract events using a minimum inter-event time (MIT) threshold (Restrepo-Posada and Eagleson, 1982; Molina-Sanchis et al., 2016). An ‘event’ thus constitutes any rainfall separated by at least 11 hours of rain-free conditions, following practice in several Danish hydrological studies (Gregersen et al., 2013; Thomassen et al., 2023). This approach ensures that each event begins and ends with non-zero rainfall. The MIT-parameter-choice of MIT has been shown to play an important role in determining both the number and properties of rainfall events selected (Dunkerley, 2008). For temporal loading, which is interested in what happens around the peak, the way in which the edges of the event are defined is particularly important, but in this research we do not investigate its influence further identified (Dunkerley, 2008). In this study, event definition is treated as a fixed preprocessing choice rather than a variable of investigation, reflecting a deliberate scoping decision. While the delineation of event boundaries can influence the timing and relative prominence of peak intensity, and hence derived temporal loading metrics, its effects have been examined in several previous studies, e.g. Dunkerley (2008, 2010, 2015); Wang et al. (2019); Freitas et al. (2020); Molina-Sanchis et al. (2016); Haile et al. (2011); Medina-Cobo et al. (2012). In contrast, this study focuses on methodological choices that have received less systematic attention, namely the selection and interpretation of temporal loading metrics and their sensitivity to rainfall representation and aggregation.

We define and analyse entire rainfall events, rather than only the most intense ‘burst’ periods, capturing the complete temporal evolution of each storm. Events with less than 4 mm of total rainfall are excluded to remove very light events unlikely to be hydrologically significant. This process produces a dataset of 233,128 rainfall events observed between 1979 and 2025.

350 Coarser-resolution versions of each event are also derived at 10-, 30-, and 60-minute resolutions. These are generated based on the start and end times of the 5-minute event, so for a given coarser resolution, the first timestamp equal to or before the 5-minute event start is taken as the new start, and the last timestamp equal to or later than the event end is taken as the new end. As a result, the duration of events at coarser resolutions may be equal or longer in the duration than their 5-minute counterparts.

3.3 Dimensionless mass curve generation

355 Dimensionless mass curves (DMCs) are generated from all 5-minute resolution rainfall events. A DMC represents the ~~cumulative~~ distribution of rainfall within an event, scaled such that both time and accumulated rainfall range from 0% to 100%. Each DMC is interpolated to have 10 equally spaced time points. Importantly, while DMCs are typically defined as cumulative profiles, in this analysis we derive a double normalised incremental representation by converting each interpolated DMC back into incremental rainfall. This version reflects rainfall intensity as a proportion of total event mass, distributed over relative time.

360 All temporal loading metrics in this study are applied to these 10-point, double normalised incremental series rather than to cumulative curves. This preserves consistency with how metrics are typically applied to raw rainfall intensities, while enabling direct comparison across events in a dimensionless domain. We additionally apply all metrics to events that have been double normalised but retain their original number of time steps, isolating the effect of normalisation from that of aggregation to 10 points.

365 3.4 Metric computation and post-processing

Rainfall temporal loading metrics are implemented in Python based on the definitions provided in the original publications. The full codebase for rainfall event extraction and metric calculation is openly available at our GitHub repository (<https://github.com/masher92/MetricEvaluation/>), providing a consistent and reproducible framework for future applications of these metrics. In some cases, the literature lacks sufficient detail on the precise application of metrics (e.g., with respect to temporal

370 resolution, handling of zero values, normalisation procedures). Where ambiguity exists, we make reasonable interpretations of the provided formulations. Two of the ~~identified metrics—~~metrics identified in Table 1, the Binary Shape Code and the event's Crossing Properties~~—~~, are not straightforward numeric indicators and are thus deemed outwith the scope of this study.

In addition to the literature-identified metrics, we ~~compute~~calculate a set of ~~statistical moments~~additional metrics on each event's intensity profile~~to help interpret our~~. Firstly, we include the mean and maximum event intensity to aid interpretation

375 of the clustering results. ~~These include classical—~~Although these do not describe temporal loading directly, they can help identify cases where other metrics primarily reflect overall magnitude rather than temporal structure. We then calculate several temporal moments, which ~~are calculated on the raw sequence of intensities without respect to their temporal order. These include the classical mean, which is the mathematical average of the intensity values; the classical standard deviation, which is the root-mean-square deviation of intensities from their mean, i.e. the typical spread or variability of rainfall rates (previously~~

380 ~~applied to rainfall events by García-Bartual and Andrés-Doménech (2017)); the classical skewness, which measures the asymmetry of the intensity distribution around the classical mean; and the classical kurtosis, which measures tail heaviness or peakiness of the intensity distribution around that centre~~describe the overall distribution of rainfall intensity through time. These provide a

compact and physically interpretable summary of how rainfall mass is positioned, spread, and skewed within an event. Despite their general applicability as distributional descriptors, such metrics have not previously been applied in the context of rainfall temporal structure. Finally, we introduce an author-defined metric, termed the wet-dry transition rate, which quantifies the frequency of wet–dry transitions within an event. This was included to test an event-based representation of intermittency that aligns closely with the intuitive, dictionary definition of the term (i.e., alternation between wet and dry periods). While previous metrics have been proposed to quantify intermittency—for example, the internal intermittency measure of Schleiss (2018), which calculates the fraction of dry sub-intervals, and the burstiness and memory metrics of Schleiss and Smith (2016) and Dey (2023), which operate on interarrival times over long, multi-event time series—these approaches either measure dry fraction rather than alternation, or require long series to provide stable estimates. The wet-dry transition rate provides a simple, interpretable, event-based alternative that is directly applicable at the scale of individual rainfall events.

We also calculate temporal moments which consider the distribution of intensity values in respect to time. These include the temporal mean (referred to here as the centre of gravity), which weights

Table 2: Overview of metrics defined by authors

Metric	Meaning	Type
<i>Centre of Gravity (CoG)</i>	Equivalent to the temporal mean. Calculates the rainfall-weighted mean time of the event by multiplying each time step by its rainfall intensity to calculate the average time at which rainfall occurs during the event, normalised, summing these values, and dividing by the total rainfall. This is then normalised to fall between 0 and 1 ; the temporal standard deviation, which measures the	Continuous
<i>Temporal kurtosis [(T) kurtosis]</i>	Quantifies how sharply the rain is clustered around the temporal mean	Continuous
<i>Temporal skewness [(T) skewness]</i>	Measures the asymmetry of rainfall in time around the temporal mean	Continuous

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
<u>Temporal standard deviation [(T) std]</u>	<u>Measures the</u> spread of rainfall in time around this-<u>the</u> temporal mean, indicating how temporally concentrated or dispersed the rainfall is over the event duration ;the temporal skewness, which measures the asymmetry of rainfall in time around this temporal mean; and the temporal kurtosis, which quantifies how sharply the rain is clustered around that centre. The full list of metrics applied, including those from the literature alongside additional statistical and temporal moments, is presented in Table ??. The full codebase for rainfall event extraction and metric calculation is openly available at our GitHub repository , providing a consistent and reproducible framework for future applications of these metrics. List of metrics within each typology applied in this researchMetric type	Metrics <u>Continuous</u>
Classification		
<u>Wet-dry transition rate</u>	3rd / 4th / 5th with peak; 3rd / 4th / 5th with most; 3rd with D_{50}; 3rd with CoG. Summary statistics <u>Measures the normalised (by event length) frequency of switches between a wet and dry state</u>	<u>Continuous</u>

<u>Metric</u>	<u>Meaning</u>	<u>Type</u>
Max intensity ; I30; Cv; (C) skewness; (C) kurtosis; mean intensity; (C) std; (T) skewness; (T) kurtosis; (T) std; Centre of gravity; Peak position ratio; time to peak; Peak mean ratio; Relative amplitude; Skew_p; NRMSE_p; Gini coefficient; Lorenz asymmetry coefficient; Event loading index; Fraction of rainfall in Q1 / Q2 / Q3 / Q4; m1; m2; m3; m4; m5; T25; T50; T75; PCI; TCI; Mean intensity HIZ; % time in LIZ / HIZ; % Rain in HIZ. Intermittency metrics	<u>Identifies the maximum rainfall intensity within the event</u>	<u>Continuous</u>
<u>Mean intensity</u>	Intermittency fraction; Event dry ratio. <u>Identifies the mean rainfall intensity within the event</u>	<u>Continuous</u>

395 **3.5 Testing metric overlap and complementarity (RQ2)**

To identify metrics which are strongly correlated we apply agglomerative hierarchical clustering based on pairwise metric similarity across all events at 5-minute resolution (Jain et al., 1999). Similarity between metrics is quantified using the absolute Spearman rank correlation coefficient, and clustering is performed using average linkage. While clustering directly on metric vectors using Euclidean distance was considered, this would have prioritised numerical proximity over functional similarity.
 This is problematic because some metrics which have very high negative correlation, and likely describe the same feature, are very distant in Euclidean space. Prior to clustering, each metric is assessed for skewness, and those with an absolute skewness ≤ 1 are retained in their original form. For variables with higher skewness, a logarithmic transformation ($\log_{1p}$) is applied if all values are positive, and otherwise, the Yeo-Johnson transformation (Yeo and Johnson, 2000), which supports zero and

negative values, is used. After transformation, all variables are MinMax scaled to common [0,1] range. To inform the number of clusters, we compute silhouette scores for a range of candidate cluster numbers. Silhouette scores quantify how appropriate a cluster number is based on how well each metric fits within its assigned cluster, based upon its similarity to members of its own cluster compared with its dissimilarity to those in other neighbouring clusters (Rousseeuw, 1987). This analysis suggests twelve clusters to be the optimal number (Figure D1, Appendix D).

3.6 Testing sensitivity to temporal aggregation (RQ2RQ3)

To assess the sensitivity of each metric to temporal aggregation of rainfall data, we compare metric values computed on rainfall events derived from data aggregated at different temporal resolutions. We treat metrics computed on events from 5-minute rainfall data as the reference ('truth') and compare them to values obtained from coarser aggregations at 10-, 30-, and 60-minute intervals.

For each metric and resolution pair, we quantify both numerical and ranking sensitivity. These two complementary measures allow us to distinguish between metrics which exhibit substantial changes in raw values (numerical sensitivity), those which show shifts in the relative ordering of events (ranking sensitivity), and those which display both or neither aspect of sensitivity.

For continuous metrics, the numerical sensitivity is assessed using the symmetrical median absolute percentage error (sMAPE) (Makridakis and Hibon, 1995). sMAPE measures the the absolute percentage error between the metric values at 5 minute and at coarser resolutions, and is calculated as:

$$sMAPE = \frac{100}{n} \sum_{i=1}^n \frac{|x_i - y_i|}{(|x_i| + |y_i|)/2} \quad (1)$$

where x_i and y_i are the original and transformed metric values for event i , respectively.

The ranking sensitivity is assessed using Spearman's rank correlation coefficient (ρ), which quantifies the degree to which the relative ordering of events changes due to temporal aggregation, and is calculated as:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (2)$$

where n is the number of paired observations, $d_i = R(x_i) - R(y_i)$ is the difference between the ranks of the i^{th} pair of observations.

For categorical metrics, the % of events in a different category from 5 minute at each coarser resolution is used for numerical sensitivity, and Kendall's τ quantifies the ranking sensitivity (Kendall, 1938) by measuring the association between two ranked variables. It is defined as:

$$\tau = \frac{P - Q}{\sqrt{(P + Q + T)(P + Q + U)}} \quad (3)$$

where P / Q are the number of concordant / discordant pairs, and T is the number of ties only in x , and U is the number of ties only in y . If a tie occurs in both x and y for the same pair, it is not counted in either T or U .

We also visually explore how each metric responds to aggregation by plotting distributions across resolutions (using histograms for continuous metrics, and bar plots for categorical metrics).

435 We note that comparing ~~metrics across different data types, particularly continuous summary statistics and discrete categories,~~
continuous and categorical metrics introduces interpretive limitations. When calculating sMAPE scores and plotting metric distributions we opted to not transform or scale metrics. This was to preserve the physical interpretability of each metric, allowing percentage errors to reflect changes in real units (e.g., minutes or intensity proportions) rather than abstracted or standardised values. However, although sMAPE expresses the average percentage difference between two representations (e.g., at different
440 resolutions), the meaning of a given percentage depends on the nature and scale of the metric. For instance, a 70% sMAPE for a bounded percentile-based timing measure like D_{50} does not carry the same meaning as a 70% sMAPE for a scale-dependent metric such as standard deviation. This presents a trade-off, where preserving metric-specific meaning improves interpretability within each metric, but complicates comparative analysis across them. We recognise that we are operating within these constraints, but by interpreting sMAPE scores in conjunction with visual plots, rank correlations, and physical understanding
445 of the metric's behaviour, we can still make valuable assessments of metric behaviour at different resolutions and processing options.

3.7 Testing sensitivity to ~~normalisation~~-DMC construction (RQ3RQ4)

To evaluate the robustness of rainfall metrics to ~~dimensionless mass curve (DMC)~~-DMC conversion, we compare values calculated on raw 5-minute rainfall events and DMC-transformed rainfall time series. Note that, $I30$ cannot be calculated, as
450 DMCs have no concept of 30-minutes. Likewise the *Frac. in Q1 / 2 / 3 / 4* can only be calculated on events with n timesteps divisible by 4, and so are not calculated either. Following the methods for temporal aggregation sensitivity testing (Sect. 3.6), we calculate numerical and ranking sensitivity of continuous metrics using sMAPE and Spearman's ρ , and for categorical metrics using % differences in categorisation and Kendall's τ . Metric value distributions are also compared using histograms for continuous metrics, and bar plots for categorical metrics.

455 3.8 ~~Testing metric redundancy and complementarity~~ (RQ4)

4 Results and discussion: Quantitative metric analysis

~~To identify metrics which are strongly correlated and which may therefore be considered redundant or interchangeable, we apply agglomerative hierarchical clustering based on pairwise metric similarity across all events (Jain et al., 1999). Similarity between metrics is quantified using the absolute Spearmanrank correlation coefficient, and clustering is performed using average linkage. While clustering directly on metric vectors using Euclidean distance was considered, this would have prioritised numerical proximity over functional similarity. This is problematic because some metrics which have very high negative correlation,~~
460

4.1 Metric overlap and complementarity (RQ2)

The results of the hierarchical cluster analysis suggest that several well-defined groups of metrics exist. Within each cluster, metrics rank rainfall events similarly. On the dendrogram in Figure 4.(a), the shorter the horizontal distance between metrics or groups of metrics, the greater the similarity in rankings. Figure 4.(b) shows the matrix of Spearman's rank correlation scores upon which the clustering is based. Negative correlations arise where metrics respond to the same underlying asymmetry but with reversed directionality — for instance, a front-loaded event will have a lower D_{50} (proportion of event elapsed when 50% of cumulative rainfall has fallen), but a higher $m3$ (proportion of rainfall in the first third).

The clusters that emerge from this analysis form the empirical basis for defining five distinct aspects of rainfall temporal loading, each capturing a different dimension of storm structure. As described in the following sections, four aspects map closely onto the clusters, while the fifth — intermittency — emerges from a combination of its isolated position in the dendrogram and its prominent treatment in the literature. Two clusters are treated as contributing to a single aspect, magnitude concentration, given their shared conceptual basis.

Cluster 1 is the largest, containing fifteen metrics all describing what we term mass timing — the time at which most rainfall falls during an event. The metrics span both categorical and continuous types, yet exhibit very high Spearman's rank correlation coefficients throughout. This suggests that coarser categorical metrics, such as *3rd with most*, rank events similarly to more detailed continuous alternatives such as the *centre of gravity*, despite their differences in construction.

The second cluster contains eight metrics describing a related but distinct aspect: peak timing, the time at which the peak rainfall intensity occurs. Correlation coefficients are again high despite the mixture of categorical and continuous metrics. While most metrics in this group explicitly quantify the timing of the peak time step, interestingly, the metric $m1$, representing the ratio of rainfall before versus after the peak, is also placed within this cluster. Although $m1$ is seemingly focused on the balance of rainfall mass distribution rather than peak timing, its position within this cluster suggests that, in practice, it is strongly influenced by peak placement.

~~Clusters 3 and 4 likely describe the same feature, are very distant in Euclidean space. Prior to clustering, each metric is assessed for skewness, and both relate to what we term magnitude concentration — the degree to which rainfall is unevenly distributed across timesteps in terms of intensity. Although they form distinct clusters, both describe the contrast between high- and those with an absolute skewness ≤ 1 are retained in their original form. For variables with higher skewness, a logarithmic transformation ($\log_{10}$) is applied if all values are positive, and otherwise, the Yeo-Johnson transformation (Yeo and Johnson, 2000), which supports zero and negative values, is used. After transformation, all variables are MinMax scaled to common 0-1 range low-intensity periods, and are treated here as related parts of the same underlying aspect.~~

Cluster 3 emphasises distributional imbalance — how unequally rainfall is shared across timesteps, rather than the size of peaks in absolute terms. *Peak-mean ratio*, *relative amplitude*, and the *coefficient of variation* form a subcluster reflecting mean-based contrasts between intense and weak timesteps. *classical skewness* and *classical kurtosis* characterise the shape of the intensity distribution. As negative kurtosis is rare in rainfall events (as this would involve lots of high intensity timesteps, accompanied by few low intensity), *classical skewness* and *classical kurtosis* generally co-vary, tending to be high in events

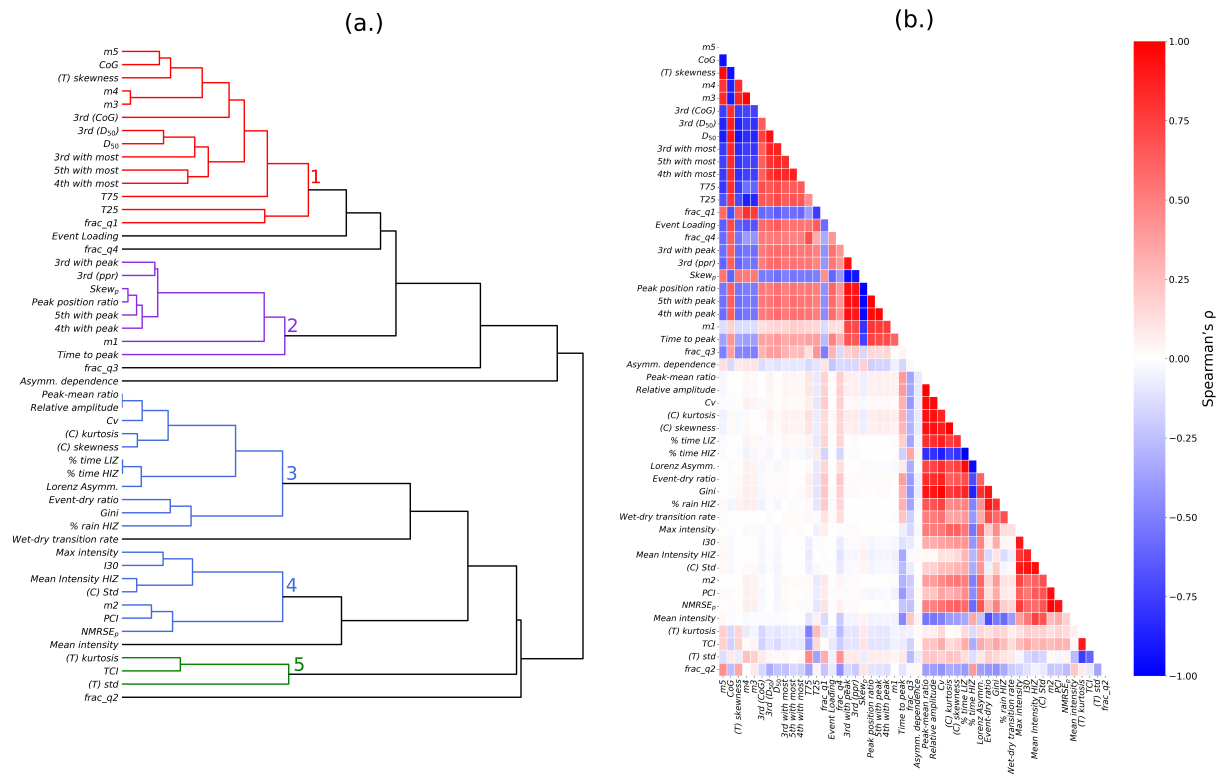


Figure 4. (a) Dendrogram visualising results of hierarchical cluster analysis of the pairwise similarity of metrics across all events. Similarity is quantified using the absolute value of Spearman's ρ , and clustering is performed using agglomerative hierarchical clustering with average linkage and 12 clusters. Clusters with more than one member are highlighted in colour and marked with a cluster number. (b) The full Spearman correlation matrix upon which clustering is based

where rainfall is concentrated into few timesteps. A further subcluster tightly links the *% time in LIZ* with its complement, *% time in HIZ*, reflecting their inverse relationship, with *Lorenz asymmetry* joining nearby as it also quantifies imbalance in the rainfall distribution between wetter and drier periods. In a separate branch, *event dry ratio*, *Gini coefficient*, and *% of rainfall in the HIZ* cluster together, collectively describing how rainfall volume concentrates into a few timesteps, with events scoring high on one metric tending to do so on the others as well.

Cluster 4 shifts emphasis towards the absolute strength of intense phases. Its first subcluster groups *maximum intensity* and *I30* (the maximum intensity in any 30-minute period), both capturing the magnitude of the single wettest short-duration period. The second subcluster links *mean intensity in the HIZ* (timesteps above the event's mean intensity) with *classical standard deviation*, both reflecting how rainfall is distributed relative to the event mean. The third subcluster contains *m2* (percentage of total rainfall in the highest-intensity timestep), *PCI*, and *NRMSE_p*. These metrics explicitly compare rainfall in some timesteps with rainfall in other, directly quantifying how concentrated rainfall is within an event. They also share a potential sensitivity

to event length, tending to return higher values when an intense burst is embedded in a short event than when it appears in a longer event with more low-intensity timesteps.

510 The fifth cluster groups *temporal kurtosis*, *temporal standard deviation*, and the *TCI*, defining a further distinct aspect: temporal concentration. Unlike magnitude-based concentration metrics, these only return high values when high-intensity timesteps are tightly grouped in time. Their very low correlation with metrics in other clusters indicates that this dimension of temporal loading is largely independent from the others. An event can have strongly concentrated intensity without that intensity being temporally clustered, and vice versa. This distinction is particularly relevant for applications such as flash-flood
515 risk and soil erosion modelling, where closely clustered high-intensity periods can greatly amplify impacts compared to if high intensity timesteps were spread across an event.

The remaining seven clusters each contain only a single metric, reflecting a range of different situations. *Asymmetry of dependence* characterises how rainfall intensities evolve over time, identifying whether events intensify and decay at similar rates — a feature potentially relevant for understanding storm generation and event evolution. The *event loading index* measures
520 the deviation in temporal variability between an observed event and its time-reversed mirror, capturing structural irregularities not represented by other metrics. Both describe properties that are conceptually interesting and may have value in specific applications, but they do not correspond cleanly to any of the four aspects identified above and are not sufficiently general to warrant inclusion as a core component of the framework. The fractional metrics, *fraction of rainfall in Q2*, *Q3*, and *Q4*, are isolated not because they are uniquely informative, but because individually they offer limited descriptive power, with
525 their information more completely conveyed by mass timing metrics such as *D₅₀* or *4th with most*. ~~To inform the number of clusters, we compute silhouette scores for a range of candidate cluster numbers. Silhouette scores quantify how appropriate a cluster number is based on how well each metric fits within its assigned cluster, based upon its similarity to members of its own cluster compared with its dissimilarity to those in other neighbouring clusters (Rousseeuw, 1987). This analysis suggests twelve clusters to be the optimal number.~~ *Mean intensity* was included as a diagnostic to test whether any other metrics are
530 implicitly driven by overall event magnitude rather than temporal structure; its isolation confirms this is not the case.

The wet-dry transition rate is also isolated in the cluster analysis, but unlike the metrics discussed above, we argue for this metric to form the basis of a fifth aspect of temporal loading: intermittency. Intermittency is frequently discussed in the literature, especially in the context of soil erosion (Dunkerley, 2015, 2023), and dry intervals within events have been shown to influence infiltration capacity, such that more intermittent rainfall can be tolerated at higher intensities during wet intervals
535 (Dunkerley, 2021b; Aryal et al., 2007). However, whether such processes are occurring depends on the actual arrangement of wet and dry periods within an event, not simply on the proportion of dry time. Despite this, intermittency has primarily been measured using the *event dry ratio*, which captures only the proportion of event time that is dry. This creates a conceptual inconsistency: while the literature emphasises the importance of the frequency and sequencing of wet and dry periods, the metrics used typically reduce this to a single duration-based quantity. Two events can share the same dry-time proportion while
540 differing markedly in how fragmented the rainfall is. An event with a single prolonged dry interval is qualitatively different from one in which the same total dry duration is distributed across many short alternating periods. The clustering results confirm this: the *event dry ratio* groups with measures of magnitude concentration, reflecting the fact that a high proportion

of dry timesteps implies rainfall is compressed into fewer wet periods, irrespective of how those periods are arranged in time. In contrast, the wet-dry transition rate remains isolated, suggesting that intermittency — when defined in terms of wet-dry alternation frequency — represents a genuinely distinct dimension of event structure not captured by commonly used metrics. Together with the four aspects identified through the cluster analysis, intermittency forms the basis of the conceptual framework developed in the remainder of this study.

5 Results and discussion: Quantitative metric analysis

4.1 Sensitivity to temporal aggregation (RQ2RQ3)

The sensitivity of each temporal loading metric to changes in the temporal resolution of the rainfall data is examined in Figure 5. Metrics located in the bottom-right corner of their plot have exhibit high robustness (low numerical and ranking sensitivity), whereas those trending upwards and to the left show increasing sensitivity to aggregation. These results are supported by Figure 6, which shows histograms of metric values across all events, with overlain distributions for each temporal resolution. Robustness varies substantially, pointing to meaningful patterns in behaviour.

Metrics that describe the peak intensity of rainfall events consistently display a low numerical robustness to temporal aggregation in Figure 5. This includes metrics such as Figure 6 complements this by illustrating how full metric distributions shift across temporal resolutions. Results are discussed with reference to the five aspects of temporal loading identified in Section 4.1.

Mass timing metrics, including temporal skewness, centre of gravity, and percentile-based measures such as peak-mean ratio D_{50} and relative amplitude T_{75} , which compare the magnitude of the peak to the overall average intensity. As the temporal resolution becomes coarser, these metrics are distorted because, while the event mean remains relatively stable, the peak intensity is often smoothed, causing the relative peak intensity to decrease. This effect is visible in Figure 6, where distributions for the relative amplitude, peak-mean ratio and maximum intensity shift towards lower values at coarser resolutions. Low numerical robustness is also seen in Figure 5 for m_2 , the % of rainfall in the whole event found in the peak time-step. However, in this case, the distribution in Figure 6 shifts in the opposite direction, towards higher values. This reflects coarser aggregations having larger time-steps, meaning that the highest intensity interval captures a larger fraction of rainfall. While these peak intensity-based metrics exhibit high numerical sensitivity to temporal aggregation, their ranking sensitivity remains comparatively low, as evidenced by Spearman's ρ values exceeding 0.8 for all metrics discussed. This indicates that although absolute values change with resolution, the relative ordering of events is largely preserved, suggesting that such metrics may still be useful for comparative analyses, even when temporal resolution varies. are among the most robust to temporal aggregation. These metrics summarise the cumulative distribution of rainfall mass rather than individual intensity extremes, and their numerical values and event rankings remain tightly clustered across resolutions. Categorical metrics derived from these quantities, such as 3rd with D_{50} and 3rd with CoG , are correspondingly stable, though metrics classifying events by the fraction containing the bulk of rainfall show moderate sensitivity — substantially less, however, than those defined by peak timing.

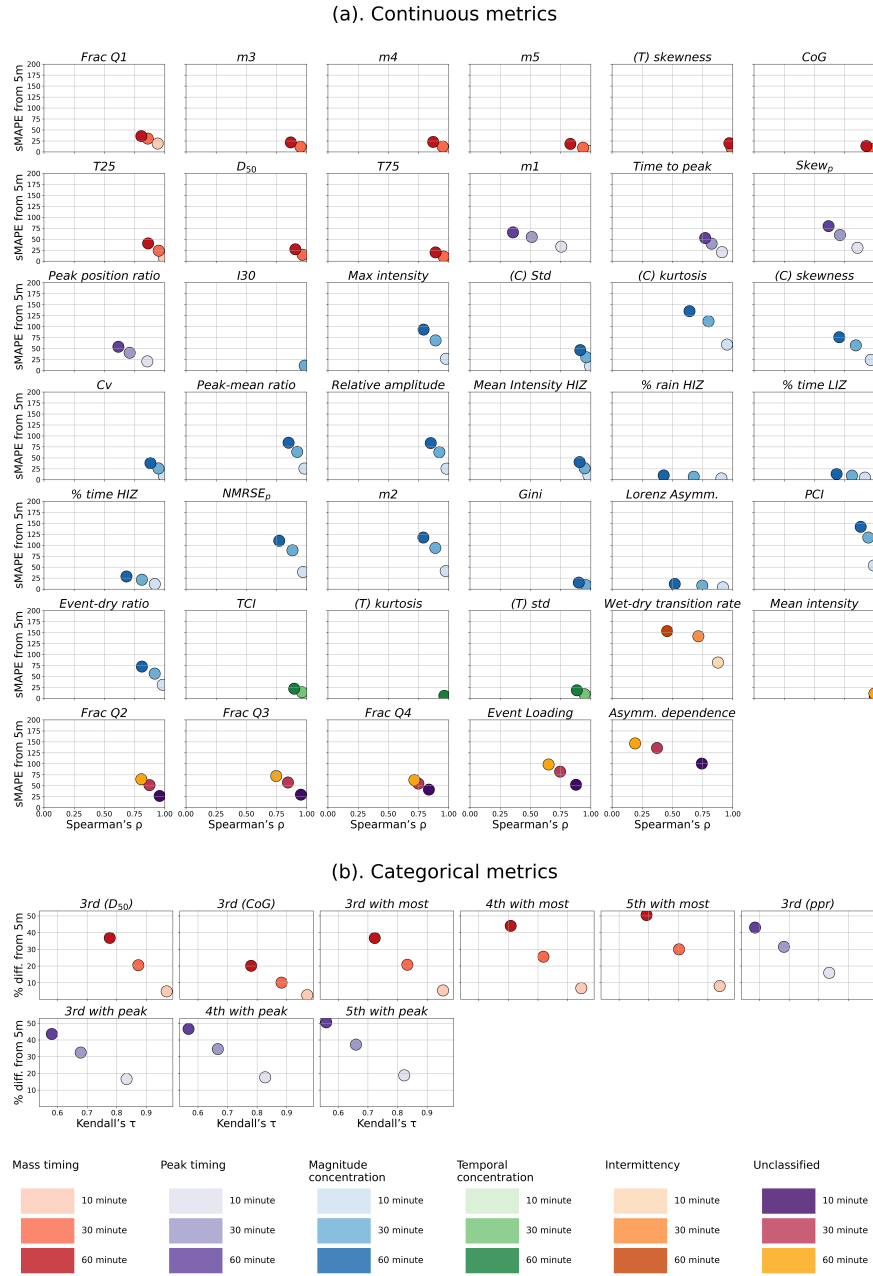


Figure 5. Sensitivity of metrics to temporal aggregation. For each metric (separate subplot), points represent comparisons between the metric calculated at coarser resolutions (10-, 30-, 60-minutes) and the reference 5-minute data. For (a). continuous metrics (~~summary statistics and intermittency metrics~~), the x-axis shows Spearman's rank correlation coefficient (ρ) and y-axis symmetrical mean absolute percentage error (sMAPE) relative to the 5-minute baseline. For (b). categorical metrics, the x-axis shows Kendall's tau, and y-axis shows the percent disagreement with 5-minute baseline.

Metrics that describe the timing of Peak timing metrics show considerably higher sensitivity. Continuous measures describing the position of the peak, such as *time to peak* and *peak-position-ratio*, or that are anchored to the timing of the peak, such as *m1*, *event loading index* and *skew_p*, are also sensitive to resolution. In contrast to the peak intensity metrics, however, these timing-based metrics show both display high numerical and ranking sensitivity. This is expected, as temporal instability under aggregation. Temporal aggregation not only smooths peak intensities but can also displace the perceived timing of the peak. In particular, where the aggregation process assigns aggregated values, particularly where aggregated values are assigned to the end of time intervals, the peak can be effectively shifted later in time leading to systematic shifts at coarser resolutions. Figure 6 shows broader, more variable distributions for these metrics under coarser resolutions, reflecting increased distortion. A similar pattern holds for categorical metrics. Those that classify events based on the fraction containing the peak rainfall (*3rd with peak*, broaden under aggregation and event rankings are frequently reordered. Categorical metrics based on peak position exhibit a similar pattern. Although this behaviour is conceptually expected, the magnitude and consistency of ranking instability across a large event sample cannot be inferred a priori and therefore warrants explicit quantification.

Within the magnitude concentration group, sensitivity varies depending on metric construction. Metrics explicitly dependent on peak intensity, including *4th with peak-mean ratio*, *5th with peak relative amplitude* exhibit greater sensitivity to resolution changes than those based on the fraction containing the bulk of rainfall (*3rd with most*, *4th with most*, *5th with most*). This suggests that metrics tied closely to the peak are generally more resolution-dependent, while metrics reflecting broader event structure tend to be more robust.

The statistical moments, and *maximum intensity*, show strong numerical sensitivity, with values decreasing at coarser resolutions due to peak smoothing. Conversely, *m2*, defined as the percentage of rainfall in the peak timestep, increases with aggregation as longer intervals allow the highest-intensity step to capture a larger fraction of total rainfall. In both cases, ranking sensitivity remains comparatively low (Spearman's $\rho > 0.8$), suggesting that relative event ordering is largely preserved even where absolute values shift. Such metrics may therefore remain useful for comparative analyses across datasets of differing resolution, though absolute values should not be interpreted independently of it. *classical skewness* and *classical kurtosis* both show high numerical and ranking sensitivity to temporal aggregation in Figure 5. As seen in Figure 6, their distributions shift leftward at coarser resolutions, suggesting that rainfall events appear more symmetrical and less heavy-tailed. This reflects the mathematical sensitivity of these moments to extreme values and their positions within the distribution, features that are especially affected, reflecting their mathematical dependence on extreme values directly modified by aggregation. Smoothing sharp, intense peaks reduces the extremity of the tails (lowering kurtosis), while averaging over longer intervals can flatten sharp rises or falls in intensity, reducing asymmetry in the distribution of intensity values (lowering skewness). Ranking sensitivity is also high, as aggregation can reorder events by altering the prominence of subtle asymmetries or outliers.

Metrics that split Metrics partitioning events into high- and low-intensity zones (HIZ/LIZ) show moderate ranking sensitivity, but fairly low numerical sensitivity (Figure 5). The distribution of exhibit more moderate responses, with limited numerical change but some ranking instability. *% time in HIZ* *Event dry ratio* shifts right at coarser resolutions, while for *% time in LIZ* it shifts left. A higher proportion of time above the mean intensity at coarser resolutions suggests that fewer time steps fall below

the mean when extremes are smoothed (and the mean is lower). The *% rainfall in HIZ* shows less change across resolutions. Its distribution is already heavily skewed toward high values, and aggregation slightly lowers the upper tail, suggesting fewer extreme values, but does not substantially alter the metric's central tendency. The *mean intensity in HIZ* is relatively robust across resolutions also shows moderate sensitivity. An event with a high dry ratio is likely to contain at least one substantial contiguous dry period, which would need to be entirely absorbed by coarsening before the metric is meaningfully altered.

The highest robustness in Figure 5 is observed for the time-based moments, including Temporal concentration metrics are among the most robust in the analysis. Temporal kurtosis, temporal skewness standard deviation, temporal kurtosis, temporal std, and centre of gravity, and percentile timing metrics like D_{50} and T_{75} . These summarise the overall timing and distribution of rainfall across the event, which is less affected by resolution changes than short-lived peaks. Their the TCI integrate information across the full event duration and are less dependent on the precise magnitude or location of individual peaks. Both numerical values and event rankings remain comparatively stable across resolutions, and distributions in Figure 6 remain tightly clustered and largely unchanged across resolutions 6 show only minor shifts, indicating that broader patterns of temporal clustering are largely preserved under aggregation.

The intermittency metric exhibits both numerical and ranking sensitivity to aggregation, and we argue this reflects a fundamental incompatibility between the metric and coarsely aggregated data. At coarser resolutions, intermittent dry periods within events are obscured by aggregation, rendering intermittency effectively unmeasurable. Treating coarse-resolution estimates as meaningful risks drawing incorrect conclusions in applications that depend on accurate characterisation of within-event dry periods, including soil erosion modelling, where dry intervals influence infiltration recovery (Dunkerley, 2019, 2021b), and detection of long-term changes in rainfall intensity. If intermittency increases over time but goes undetected due to coarse temporal aggregation, apparent stability in hourly or daily totals could mask a real increase in wet-period intensity (Schleiss, 2018). Categorical metrics based on these values, such as the 3rd with D_{50} , or the 3rd with CoG, also show the greatest stability among categorical approaches.

Together, Figures 5 and 6 indicate that metrics capturing the cumulative structure and timing of rainfall events can be used with confidence even when only coarse-resolution data are available. Overall, while some responses to aggregation, particularly for peak-dependent metrics, are intuitive, the results demonstrate that the magnitude and ranking implications are highly metric-specific and cannot always be reliably anticipated without explicit testing. These findings provide a practical reference for metric selection under data constraints. Where only coarse-resolution data are available. By contrast, those reliant on the timing or magnitude of short-lived intensity bursts should be applied with caution in such cases, mass timing and temporal concentration metrics offer the most reliable characterisation, peak timing and magnitude concentration metrics should be interpreted with caution, and the intermittency metric should be considered unreliable.

4.2 Sensitivity to normalisation DMC construction (RQ3 RQ4)

Sensitivity to conversion of raw 5-minute rainfall into double-normalised, 10-step representations (DMCs) varies substantially across temporal loading metrics, as shown in Figure 7 and Figure 8. As with temporal aggregation (Section 4.1), metrics that describe the overall distribution of rainfall over time show the least sensitivity to normalisation. In Figure 7a, these robust

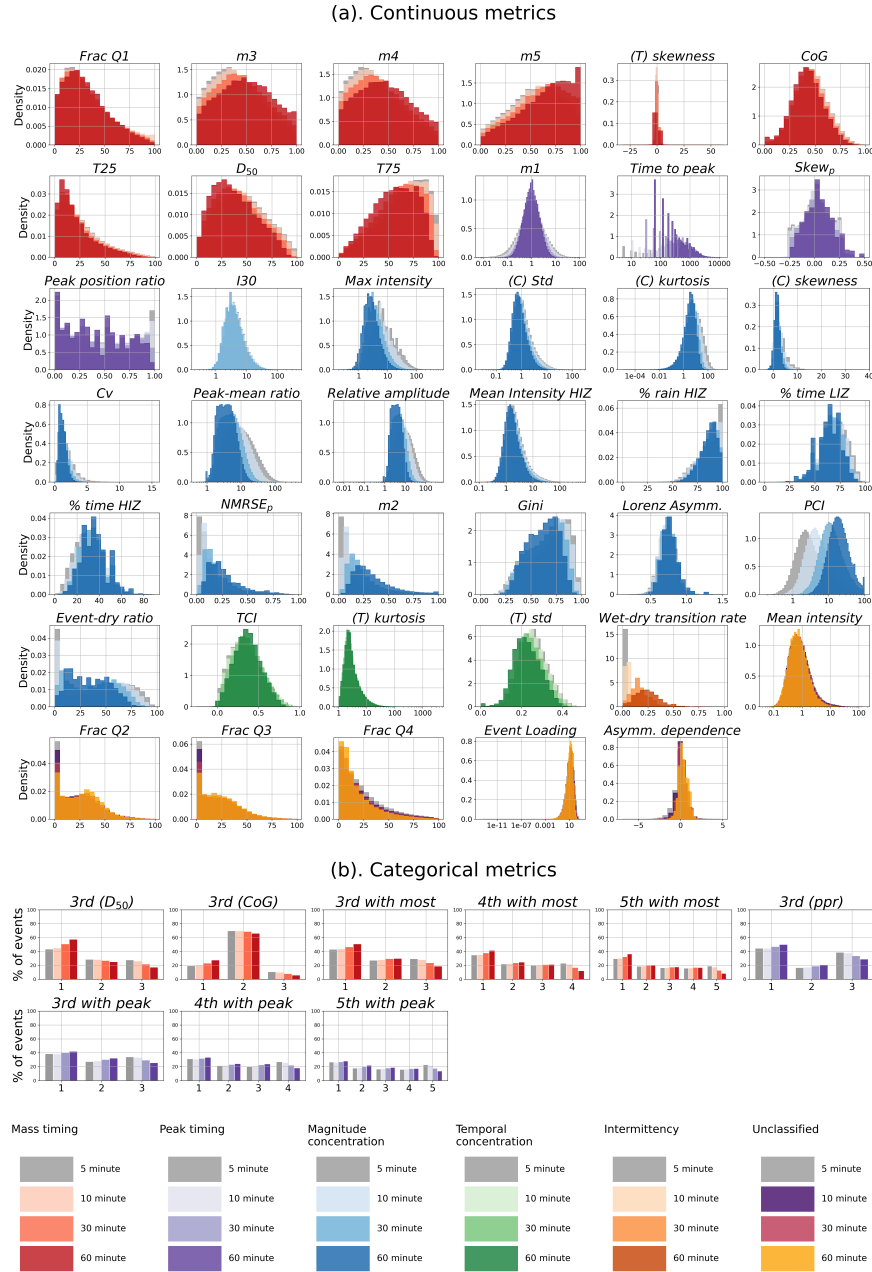


Figure 6. Sensitivity of metric distributions to temporal aggregation. Each subplot shows how the distribution of a given metric changes when calculated on rainfall data at different temporal resolutions (5-, 10-, 30-, 60-minutes). For (a). continuous metrics, this is shown using overlaid histograms. For (b). categorical metrics, grouped bar charts display the relative frequency of each category across resolutions.

metrics cluster in the bottom-right and are shown in more detail in the inset. They include mass distribution indicators ($m3$, $m4$, $m5$), percentile timing metrics. Results are discussed with reference to the five aspects of temporal loading identified in Section 4.1. The effects of DMC transformation reflect both de-dimensionalisation, with both time and intensity rescaled to a $[0,1]$ domain, and also resolution coarsening through interpolation. Appendix E presents supplementary sensitivity analysis using double-normalised events that retain their original temporal resolution (i.e., without interpolation to 10 time steps).

Mass timing metrics are among the most robust to DMC conversion. Percentile timing measures ($T25$, $D50$, $T75$), and moment-based metrics (centre of gravity, temporal skewness, and), and mass distribution indicators (temporal kurtosis). The value distributions for these metrics in Figure 8 also remain $m3$, $m4$, $m5$ all show limited numerical and ranking sensitivity, with value distributions remaining largely unchanged after DMC conversion (Figure 8). These metrics are inherently robust to both interpolation and loss of dimensionality because they reflect the shape and balance of the event over its duration, rather than relying on specific peaks or absolute magnitudes. Their values, and are either unitless by construction or meaningful when expressed as proportions of total event time or mass, making them more stable across transformations. Categorical metrics based on these values, such as the derived from these quantities, including 3rd with $D50$, or 3rd with CoG, as well as and those recording the fraction with most rainfall (3rd / 4th / 5th with most), also show the greatest stability among categorical approaches (Figure 7, Figure 8) show corresponding stability.

By contrast, several metrics that take continuous values on raw rainfall series become discretised when applied to DMCs interpolated to only 10 time steps. These include Peak timing metrics show considerably higher sensitivity to DMC conversion. Metrics such as % time in LIZ / HIZ, peak position ratio, event dry ratio, intermittency, peak position ratio, centre of gravity, skew_p, and classical skewness $m1$. These depend on the proportion of event time spent above or below certain thresholds or in specific zones, and so are limited in the DMC are affected both by the interpolation-induced smoothing of sharp transitions and by discretisation arising from the 10-step representation, which constrains values to multiples of 10%, constraining meaningful variation in range and granularity and limits meaningful variation across events. This is very apparent in the stepped histograms of these metrics. The stepped histograms visible in Figure 8a. Similarly, the interpolation used in the DMC conversion introduces a form of resolution coarsening similar to temporal aggregation. (a) illustrate this clearly. As with coarser time steps, this process smooths sharp transitions and reduces variability, disproportionately affecting metrics that are sensitive to peak timing and magnitude, such as peak-mean ratio, relative amplitude, peak position ratio, $m1$, $m2$, and the event loading index. temporal aggregation, peak displacement under interpolation can reorder event rankings, and this sensitivity is particularly pronounced for metrics anchored to the precise timing of peak intensity.

However, the The effects of DMC transformation extend beyond temporal smoothing. Unlike temporal aggregation, DMC conversion involves full conversion on magnitude concentration metrics are more varied. Metrics relying on relative contrasts between timesteps, such as peak-mean ratio, relative amplitude and $m2$, are affected by interpolation-induced smoothing in a manner similar to temporal aggregation. However, de-dimensionalisation, with both time and intensity rescaled to a $[0,1]$ domain. This has more severe consequences for metrics that rely on real physical quantities. For instance, mean intensity becomes meaningless, as the total intensity is fixed to 1, and so the mean intensity is 0.1 in all cases (Figure 8a). Similarly, metrics that rely introduces additional consequences not present under aggregation alone. Metrics that depend on absolute intensity

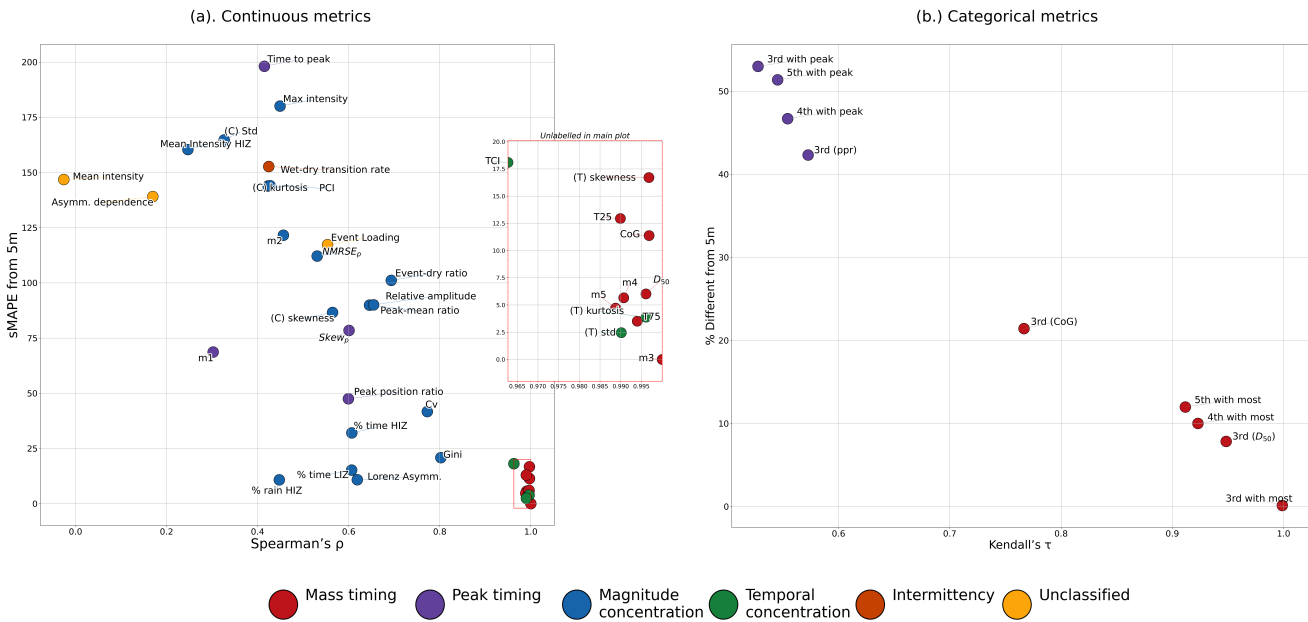


Figure 7. Sensitivity of metrics to dimensionless mass curve representation. For each metric, points represent comparisons between the metric calculated with 5-minute raw data and with a 10-point, DMCs calculated on 5-minute data. For (a), continuous metrics (including summary statistics and intermittency metrics), the x-axis shows Spearman's rank correlation coefficient (ρ) and y-axis the symmetrical mean absolute percentage error (sMAPE). For (b), categorical metrics, the x-axis shows Kendall's tau and the y-axis the percent disagreement between the pairs of metric values.

680 contrasts, such as ~~the classical standard deviation or and coefficient of variation~~, lose interpretability when all intensities are normalised. ~~In these cases, to a [0,1] domain, as~~ relative differences are compressed and the metric no longer reflects the original signal.

~~This distinction is particularly important when DMCs are used to generate synthetic events or inform design storms. In these cases, dimensionless curves are often re-scaled using external information about storm depth and duration to reconstruct plausible, physically realistic events. Metrics that remain interpretable in the dimensionless domain are therefore better suited to characterising DMC shapes or classifying rainfall types, while those that depend~~ Mean intensity becomes entirely uninformative, returning a fixed value of 0.1 in all cases (Figure 8.(a)). Metrics describing distributional imbalance in relative terms, such as the Gini coefficient, are more stable, as their construction is less dependent on physical units or fine-grained temporal variation may require adaptation or caution when applied to normalised forms.

690 ~~Taken together, these findings underscore the importance of considering input representation when comparing rainfall metrics across studies, particularly when simplifications like DMCs are involved. Metrics that are robust to DMC transformation may be more appropriate for comparative analysis or pattern identification, while others may require adjusted interpretation when removed from their physical context. Temporal concentration metrics show moderate sensitivity to DMC conversion. The~~

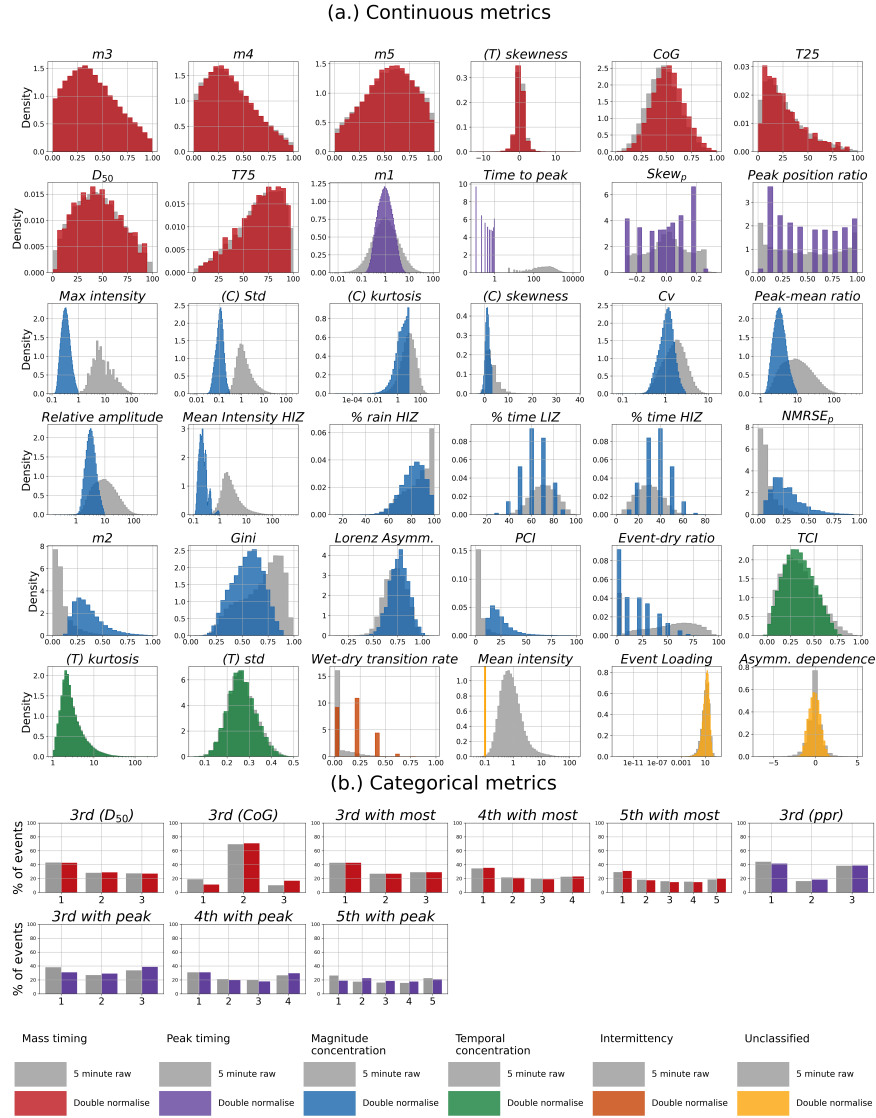


Figure 8. Sensitivity of metric distributions to [double-normalisation-in-calculation](#) on DMCs. Each subplot shows how the distribution of a given metric changes when calculated on raw rainfall data at 5-minute temporal resolution and on a double-normalised version of the event, interpolated to ten data points. For (a). continuous metrics, distributions are shown using overlaid histograms. Note that gaps in some histograms arise from the discretisation of values when metrics are computed on DMCs—certain metric values are simply not possible, so some histogram bins remain empty. For (b). categorical metrics, grouped bar charts show the relative frequency of each category across resolutions.

695 interpolation process smooths sharp transitions between timesteps, which can reduce the apparent clustering of high-intensity periods and affect metrics such as *temporal kurtosis* and *temporal standard deviation*. However, because these metrics integrate information across the full event rather than depending on individual peaks, they remain more stable than peak timing metrics, and broader patterns of temporal clustering are largely preserved.

4.3 Metric redundancy and complementarity (RQ4)

700 The conclusions drawn above for the aggregation sensitivity of the intermittency metric apply equally to the DMC representation. Interpolation to ten time steps inevitably obscures within-event dry periods, and intermittency derived from DMC events should be treated with the same caution as coarse-resolution estimates. As with aggregation, treating DMC-based intermittency estimates as meaningful risks masking real variability in wet-dry switching behaviour, and we recommend that intermittency is excluded from any analysis based solely on DMC representations.

705 Figure 4a plots a dendrogram which visualises the results of hierarchical cluster analysis based on the similarity between each pair of temporal loading metrics across the rainfall event dataset. The horizontal distance between each metric and where it joins with another branch represents similarity, with items merging at shorter distances being more similar. Twelve clusters are used, which was indicated in silhouette score analysis to be the optimal number to increase in-group cohesion, and reduce between-group separation (see Figure D1, Appendix ??). On the dendrogram, the clusters containing more than one metric are coloured and labelled with a cluster number. The resulting clusters reveal groups of metrics that behave similarly across events and may therefore capture overlapping or redundant structural information. Overall, these findings reinforce the practical guidance offered by the temporal aggregation analysis. Mass timing and temporal concentration metrics are the most transferable across input representations, while peak timing and magnitude concentration metrics require particular caution when applied to DMC data, and intermittency metrics become meaningless. This distinction is especially relevant when DMCs are used to classify rainfall types or inform design storms, where metric interpretability in the dimensionless domain should be an explicit consideration in metric selection.

715 Cluster 1 is the largest cluster, with fifteen metrics, and relates to the timing of mass distribution. This includes classification metrics (*3rd with most*, *4th with most*, *5th with most*, *3rd with D_{50}* and *3rd with CoG*), as well as percentile-based metrics (*T25*, *T50*, *T75*), mass distribution indicators (*m4*, *m3*, *m5*), *temporal skewness*, the *centre of gravity*, and the *fraction of rainfall in Q1*. This cluster contains metrics with the strongest (positive or negative) correlations, as can be seen in Figure 4b which plots the Spearman's correlation matrix. Negative correlations arise where metrics respond to the same underlying asymmetry but with reversed directionality. For instance, a front-loaded event will have a lower D_{50} , but a higher *m3* (proportion of rainfall in the first third). The high correlations between metrics in the cluster suggests that these metrics can be used interchangeably to rank the asymmetry in rainfall mass distribution in a set of rainfall events.

4.3 Recommendations

725 ~~Cluster~~ Based on the findings of the literature review and the data-driven clustering exercise, we recommend breaking event temporal loading into five components: (1) mass timing, (2) contains a grouping of eight metrics clearly related to the timing

of the peak timestep. Metrics in this peak timing cluster include several classification metrics (e.g. *3rd with peak*, *4th with peak*) peak timing, (3) magnitude concentration, (4) temporal concentration, and (5) intermittency. Table 3 summarises each component, lists the metrics associated with it, *5th with peak*, and *3rd (D_{50})*), as well as summary statistics that explicitly quantify peak position (e.g., *peak-position-ratio*, *time-to-peak*, and *skew_p*). Interestingly, the metric *mI*, representing the ratio of rainfall before versus after the peak, is also placed within this cluster. Although *mI* is seemingly focused on the balance of rainfall mass distribution rather than peak timing, its position within this cluster suggests that, in practice, it is strongly influenced by peak placement. Cluster 2 appearing adjacent to, but distinct from, cluster 1 implies that while both peak and mass timing relate to the temporal distribution of rainfall, they are not completely interchangeable concepts. This highlights the importance of establishing the aspect of temporal loading most relevant to a specific application before selecting metrics to utilise.

(a) Dendrogram visualising results of hierarchical cluster analysis of the pairwise similarity of metrics across all events. Similarity is quantified using the absolute value of Spearman's ρ , and clustering is performed using agglomerative hierarchical clustering with average linkage and 12 clusters. Clusters with more than one member are highlighted in red and marked with a cluster number. (b) The full Spearman correlation matrix upon which clustering is based

Unlike the first two clusters, which both capture aspects of when rainfall occurs within the event, whether in terms of the distribution of mass or the relative timing of the peak, the metrics in the third, fourth and fifth clusters are not concerned with sequencing in time. Instead, they describe the magnitude of the peaks and how rainfall is concentrated among timesteps and identifies a recommended metric for practical use. The recommended metrics are: the 4th with most rainfall and D_{50} (mass timing); the peak-position-ratio (peak timing); Gini coefficient (magnitude concentration); and for some metrics, how closely such high-intensity timesteps occur; Temporal standard deviation (temporal concentration), without distinguishing whether these bursts fall early, mid, or late in the event.

The third and fourth clusters contain metrics relating to the peak magnitude and magnitude concentration. In cluster 3, the first of the three subclusters identified on the dendrogram is formed by *maximum intensity* and *I30* (the maximum intensity in any 30-minute period). These both capture the magnitude of the single wettest short-duration period without reference to; and the rest of the event. The second subcluster groups the *mean intensity in the HHZ* (timesteps above the event's mean intensity) with *classical standard deviation* (the variation of rainfall intensities around the mean intensity). Both metrics emphasise how rainfall is distributed relative to the event mean, reflecting the degree to which rainfall is concentrated into particularly intense phases as opposed to being spread more evenly across the event. The third subcluster contains *wet-dry transition rate* (intermittency). Recommended metrics are selected on the basis of two criteria. Where a component showed high sensitivity to temporal aggregation or DMC construction, we recommend the metric that proved most robust across resolutions and representations. Where a component showed consistently low sensitivity, we prioritise comparability with existing literature and recommend the simplest and most widely used metric. An exception is mass timing, for which we recommend two metrics — one continuous and one categorical. D_{50} is recommended as the more precise and informative measure. However, the categorical *m24th with most* (percentage of total rainfall in the highest-intensity timestep), *PCI*, and *NRMSE_p*. These metrics explicitly compare rainfall in some timesteps with rainfall in others, directly quantifying how concentrated rainfall is within an

event. They also share a potential sensitivity to event length, tending to return higher values when an intense burst is embedded in a short event than when it appears in a longer event with more low-intensity timesteps. Taken together, the three subclusters demonstrate that cluster 3 relates to metrics of storm magnitude and concentration. The first subcluster isolates the absolute peak intensity, while the second and third reflect different ways of capturing how rainfall is concentrated across timesteps. Events with very high peaks often also display strong contrasts between high and low-intensity periods, which is consistent with the behaviour of concentration-focused metrics such as *mean intensity in the HIZ*, *(C) std*, *m2*, *PCI*, and *NRMSE_p*, metric is so widely used in the literature and performs sufficiently well in sensitivity testing that we also recommend its continued use, to maintain comparability with existing studies. The equation for each of the suggested metrics are detailed in Equations 4, 6, 7, 8, 9, and 12.

Cluster 4 relates to magnitude concentration, but focuses more on metrics that reflect distributional imbalance, showing how rainfall is unevenly shared across timesteps rather than simply reflecting the size of absolute peaks. The first subcluster captures how rainfall peaks dominate relative to the event average. *Peak-mean ratio* Conceptually, D_{50} is defined as the percentage of event duration at which 50% of the total event rainfall volume has accumulated. Given a precipitation rate, $r(t)$, and *relative amplitude* form a tight group, both measuring peak intensity relative to the mean, while their association with the *coefficient of variation (CV)* reflects the shared emphasis on mean-based ratios. Although *CV* compares the mean to the standard deviation rather than the peak directly, the standard deviation rises sharply as peaks intensify, highlighting the contrast between intense and weaker timesteps. The second subcluster contains the moment-based descriptors *(C) kurtosis* and *(C) skewness*, which characterise the shape of the rainfall distribution. As negative kurtosis is rare in rainfall events (as this would involve lots of high-intensity timesteps, accompanied by few low-intensity), skewness and kurtosis generally co-vary. Both metrics tend to be high in events with rainfall concentrated in a few timesteps, and low in more uniform events. Another subcluster tightly links the *% time in LIZ* with its complement, *% time in HIZ*, reflecting their inverse relationship, with *Lorenz asymmetry* joining nearby as it also quantifies imbalance in the rainfall distribution between wetter and drier periods. In a separate branch, *event-dry ratio*, *a precipitation accumulation over time*, *Gini coefficient*, and *% of rainfall in the HIZ* cluster together, collectively describing how rainfall volume concentrates into a few timesteps, with events scoring high on one metric tending to do so on the others as well. $P(t) = \int_0^t r(t)dt$, D_{50} is given by

$$D_{50} = \frac{t_{50}}{t_{end}} \quad (4)$$

where t_{end} is the total event duration, and t_{50} satisfies $P(t_{50}) = \frac{1}{2}P(t_{end})$.

The fifth cluster in Figure 4.a groups *temporal kurtosis*, *temporal standard deviation*, and the *TCl*, all of which quantify the temporal concentration of rainfall. Unlike magnitude-based concentration metrics, these measures only return high values when high-intensity timesteps are tightly grouped in time. Their very low correlation with other metrics indicates that this dimension of temporal loading is largely independent from others. This property is particularly relevant for applications such as flash-flood risk assessment and soil erosion modelling, where closely clustered high-intensity periods can greatly amplify impacts compared to if high-intensity timesteps were spread across an event.

795 The remaining seven ‘clusters’ each contain only a single metric. In some cases this suggests they provide unique information about rainfall events. The *asymmetry of dependence* characterises how rainfall intensities evolve over time, identifying whether events increase and decrease in intensity at similar rates, a feature that may be relevant for understanding event evolution or storm-generation. In Figure 4a, this metric is shown to split off very early from the other metrics, highlighting its unique position. The *event loading index* measures the deviation in temporal variability between an observed rainfall event and a version of that event but mirrored in *4th with most* is calculated by splitting a rainfall event into four segments of equal duration, and defining this set as $q = 1, 2, 3, 4$. If the total number of time steps is not divisible by four, then the time steps are rescaled accordingly and the precipitation is linearly interpolated to the new time steps. We then define the total sum of precipitation in each quartile as:

$$805 \quad P_q = \sum_{i \in T_q} p_i, \quad q = 1, 2, 3, 4 \quad (5)$$

where T_q is the set of time steps in each segment, and $2nd$ clusters. Likewise, *intermittency* captures the frequency with which rainfall starts and stops during an event, a feature not reported by other metrics. In contrast, the isolated clustering of p_i is the precipitation accumulation in the fractional metrics, *fraction rainfall in Q2, Q3, and Q4*, more likely reflects their limited descriptive power when considered individually. Unless one quarter dominates the rainfall distribution, these metrics tend not to strongly distinguish between different event types or structures. Similarly, *i'th time step*. Finally, *mean intensity* *4th with most* does not intuitively convey any information about temporal loading. It is included in this analysis to test whether any other metrics are implicitly driven by the total event scale. This appears to not be the case, which is a useful conclusion. These results highlight the importance of selecting a diverse set of metrics when analysing rainfall temporal loading. While many metrics are highly interrelated, a subset offer complementary insights that may enhance interpretation or predictive performance when used together. becomes:

$$815 \quad Q_{\max} = \arg \max_{q \in \{1, 2, 3, 4\}} P_q \quad (6)$$

5 Recommendations

The *Gini coefficient* is provided by

$$820 \quad Gini = \frac{1}{2n^2\mu} \sum_{i=1}^n \sum_{j=1}^n |p_i - p_j| \quad (7)$$

where n is the total number of time steps in the event, and μ is the mean precipitation intensity.

A central contribution of this paper is the argument that analyses of rainfall temporal loading must begin with a clear definition of the aspect(s) most relevant to the research question or application. Rather than prescribing a single universal

metric, we provide a conceptual framework to guide deliberate, defensible, and replicable metric selection. Based on the findings of the literature review and the data-driven clustering exercise, we recommend breaking event temporal loading into five components: The peak position ratio is given as the time step where the peak, t_{peak} , is located over the total duration of the event, t_{end} :

$$\text{PeakPositionRatio} = \frac{t_{peak}}{t_{end}} \quad (8)$$

The temporal standard deviation is defined as

$$\sigma_t = \frac{1}{t_{end}} \sqrt{\frac{\sum_{i=1}^n (t_i - t_\mu)^2 p_i}{\sum_{i=1}^n p_i}} \quad (9)$$

where t_μ is the centre of gravity and t_i is the time accumulation during the event at time step i :

$$t_\mu = \frac{\sum_{i=1}^n t_i p_i}{\sum_{i=1}^n p_i} \quad (10)$$

$$t_i = (i + 1) \Delta t \quad (11)$$

The wet-dry transition rate is the proportion of event duration where a switch between wet and dry states occur:

$$\text{Wet - dry transition rate} = \frac{1}{n} \sum_{i=1}^{n-1} \mathbf{1}(s_i \neq s_{i+1}) \quad (12)$$

where $\mathbf{1}(\cdot)$ is an indicator function (1 = true, 0 = false), and $s_i = \mathbf{1}(p_i > 0)$ is a binary wet/dry indicator.

Our analysis shows that many metrics are sensitive to temporal resolution or DMC transformations. For peak timing, magnitude concentration, and intermittency, no metrics are fully robust, reflecting the inherent variability of peaks and the smoothing introduced by temporal aggregation. We highlight that this means that studies using these metrics across different resolutions should interpret comparisons with caution. For these components, we recommend use of metrics which have been more commonly applied in the past and which are conceptually simple. Temporal concentration and mass timing metrics are generally more robust. For these components, our recommended metrics also consider simplicity and prior usage in the literature, in addition to robustness. While we highlight the ‘robust’ metrics in Table 3, we note that Tables F1, ?? and ?? in Appendix F, provides full metric-specific summaries of robustness.

4.1 Study limitations

5 Summary and conclusions

Table 3. Five recommended components of rainfall event temporal loading. Each component is accompanied by its definition and a list of all reviewed metrics which are best described by this component, with the associated metrics found to be listed, and a recommended metric identified. Listed metrics robust at different to temporal resolutions aggregation and after DMC normalisation processing are highlighted in red. A recommended metric for use studying each component is also provided.

Category	Component Description	All metrics	Recommendation	Reasons
Mass timing	Describes when the majority of rainfall mass falls during the event. From the start (low values), through the middle, to the end (high values).	3rd / 4th / 5th with most, 3rd with D_{50} , 3rd with CoG, Centre of gravity, D_{50} , T25/75, m1, m3, m4, m5, Event loading index, (T) skewness (t) skewness, Asymmetry of dependence, Frac. in Q1 / 2 / 3 / 4	4th with most (categorical) D_{50} (continuous)	Of the robust metrics, these ones have been most commonly applied previously. Many metrics are robust to both temporal aggregation and DMC construction. D_{50} is recommended as the most precise and informative continuous measure. The 4th with most is additionally recommended due to its widespread use in the literature, supporting comparability with existing studies.
Peak timing	Describes when the peak intensity occurs during the event. From the start (low values), through the middle, to the end (high values).	3rd / 4th / 5th with peak, 3rd ppr, time-to-peak, peak position ratio, peak position ratio, Skew _p , m1	Peak position ratio	No peak timing metrics are robust, so use in cross-study to temporal aggregation or DMC construction, so all should be used in cross-resolution comparison with care. Peak position ratio is the recommended as the simplest and most widely applied, easily understood option.
Magnitude concentration	Describes the strength of intense phases, contrasts between high- and low-intensity periods, or overall statistical inequality. Not affected by whether timesteps concentrating rainfall are close together in time, or where in time they occur.	Max intensity, I30, PCI, (C _c) skewness, (C _c) kurtosis, (C _c) std, CV, Mean intensity in HIZ, % time in LIZ / HIZ, % Rain in HIZ, Gini Coefficient, Lorenz Asymmetry Coefficient, NRMSE _p , Peak-mean ratio, Relative amplitude, m2, event dry ratio	Gini coefficient	Gini coefficient. Most metrics in this category are sensitive to both temporal aggregation and DMC construction. The Gini coefficient is the only metric robust to both temporal resolution and DMC transformation, and is therefore recommended despite being less commonly applied than alternatives such as peak-mean ratio.
Temporal concentration	Describes how narrowly rainfall is clustered around a point in time, regardless of when in time that cluster appears.	TCI, T (kurtosis), T (std), TCI (T) std, TCI (T) kurtosis, (t) std		All metrics are robust. While in this category are robust to both temporal aggregation and DMC construction. (t) standard deviation is recommended over TCI has been used previously in one study, we recommend that T (std) is a simpler metric to measure the same as it is simpler to compute and interpret, and directly related to a well-established statistical

850 This study provides a comprehensive review and empirical analysis of rainfall event temporal loading metrics, drawing on literature from multiple domains including flood modelling, soil erosion and sediment transport, pollution modelling, landslide prediction, and climate change impact assessment. We identify 52 metrics that have been used to describe the temporal distribution of rainfall within storms, and apply these to over It is important to reflect on the representativeness of the rainfall data used in this study. Our analysis draws on the high-quality, dense network of gauges across Denmark, providing over
855 233,000 events observed by a large network of rain gauges. This represents the first large-scale, systematic evaluation of how such metrics behave across real-world 000 well-observed rainfall events.

We begin by organising the metrics with a typology that groups metrics into three broad categories: (i) classification metrics, which assign events to discrete categories; (ii) summary statistics, which generate continuous descriptors of temporal loading; and (iii) intermittency metrics, which quantify the frequency and spacing of dry intervals within an event. This typology is
860 constructed based primarily on metric form (e.g. categorical versus continuous) and purpose. We then perform a data-driven hierarchical cluster analysis that reveals further groups of metrics that transcend these categories. Notably, two well-defined groups of metrics are identified, including one capturing the timing of peak intensity and another describing the timing of rainfall mass distribution. Each includes both summary statistics and classification metrics, suggesting that despite differences in format, these metrics capture similar underlying aspects of temporal loading and tend to rank events similarly, indicating
865 a degree of interchangeability for comparative purposes. However, we note that this does not necessarily imply that metric choice within these clusters will not materially alter overall findings. For instance, an analysis based on the *3rd with most rainfall* may conclude differently on whether a geographical area generally has a tendency towards This scale of dataset allows for robust evaluation of metric behaviour under real storm conditions. However, Denmark's temperate maritime climate is not fully representative of the range of rainfall regimes globally. In particular, strongly convective extremes typical of tropical
870 and subtropical climates are less frequent, and storm dynamics may differ from those in arid or mountainous regions. As such, while our findings on the behaviour, sensitivity, and clustering of metrics are expected to generalise across contexts, the specific distribution of rainfall events and the relative prevalence of different temporal loading structures may not. We therefore recommend that future work apply similar analyses in a wider set of climatic settings, both to test the robustness of our methodological conclusions and to build a more globally representative basis for metric selection.

875 This analysis also has limited ability to assess whether metric choice would materially alter substantive conclusions about rainfall characteristics in a given context. The cluster analysis shows that many metrics rank events in a similar order, but does not consider how those rankings are converted into interpretable outcomes. In practice, this often involves applying thresholds or grouping events—for example, classifying events as front- or back-loaded using D_{50} . Although different metrics may agree on which events are more or less front-loaded, they do not necessarily assign events to the same categories, particularly for
880 events that sit near the classification boundaries. As a result, the same set of events could support different conclusions—for instance, one metric indicating a predominance of front-loaded events, while another suggests a more balanced distribution.

This issue extends to aggregated analyses. Differences in how events are classified can propagate into summary statistics, such as the proportion of front-loaded events ,than if D_{50} is the chosen metric. Exploring how variations in absolute metric values within correlated groups may alter or comparisons between catchments and climates. The key limitation is therefore

885 not disagreement in relative ordering, but in how that ordering is translated into categorical or aggregated interpretations.
Quantifying the extent to which this influences substantive conclusions was beyond the scope of this study, but ~~we highlight it~~
~~as a valuable focus for future research~~represents an important direction for future work.

The cluster analysis also reveals several looser groupings and a number of isolated metrics. This reflects a broader issue that
many metrics used under the banner of ‘temporal loading’ in fact represent quite different concepts. This can lead to metrics
890 ~~being borrowed across domains without clear understanding of whether the metric is actually appropriate for the intended~~
~~application. To address this, we propose a conceptual framework that distils our findings into five aspects of temporal loading:~~
~~(1) mass timing, (2) peak timing, (3) magnitude concentration, (4) temporal concentration, and (5) intermittency. The choice~~
~~of these five aspects is motivated by a combination of statistical evidence, existing practice, and conceptual need. Clarifying~~
~~and standardising these concepts represents a critical step toward improving comparability and interpretability across studies.~~

895 5 Summary and conclusions

We note that comparing metrics across different data types, particularly continuous summary statistics and discrete categories,
introduces interpretive limitations. When calculating sMAPE scores and plotting metric distributions we opted to not transform
or scale metrics. This was to preserve the physical interpretability of each metric, allowing percentage errors to reflect changes
in real units (e.g., minutes or intensity proportions) rather than abstracted or standardised values. However, although sMAPE
900 ~~expresses the average percentage difference between two representations (e.g., at different resolutions), the meaning of a~~
~~given percentage depends on the nature and scale of the metric. For instance, a 70% sMAPE for a bounded percentile-based~~
~~timing measure like D_{50} does not carry the same meaning as a 70% sMAPE for a scale-dependent metric such as standard~~
~~deviation. This presents a trade-off, where preserving metric-specific meaning improves interpretability within each metric, but~~
~~complicates comparative analysis across them. We recognise that we are operating within these constraints, but by interpreting~~
905 sMAPE scores in conjunction with visual plots, rank correlations, and physical understanding of Rainfall event temporal
loading, the distribution of rainfall intensity over time at a fixed location, plays an important role in driving hydrological and
geomorphological responses such as run-off, urban flooding, and soil erosion. Yet despite its significance, a coherent, consistent
approach to its quantification has previously been lacking. This study provides a comprehensive, interdisciplinary review of
rainfall event temporal loading metrics, identifying 48 metrics previously applied across scientific domains and introducing five
910 additional metrics to represent aspects not fully covered by existing approaches. Considerable inconsistency in terminology
and framing was found across the literature, making it difficult to determine what different studies are measuring, what aspect
of storm structure they intend to capture, or whether seemingly similar analyses are directly comparable. The five-component
conceptual framework proposed here provides a structured way to distinguish between the metric’s behaviour, we still make
valuable assessments of metric behaviour at different resolutions and processing options. key properties of temporal loading
915 and clarify the purpose of different metrics.

Our empirical analysis of whether metrics report consistent results when applied to rainfall events at different temporal
resolutions and when converted to normalised, dimensionless representations goes some way towards addressing the broader

lack of methodological transparency in event temporal loading research. We show that some metrics, especially those based on the peak timing or magnitude, are highly sensitive to temporal resolution. All 53 metrics were calculated across 233,128 rainfall events from Danish rain gauge records, and cluster analysis revealed which metrics capture overlapping information and which describe distinct aspects of storm structure. This provided an empirical basis for the suggested five-component conceptual framework and metric recommendations.

Metrics were also evaluated under different data and processing conditions. Sensitivity to temporal resolution was found to vary substantially between metrics: those describing mass timing and temporal concentration are relatively stable under aggregation, whereas metrics capturing the overall precipitation distribution tend to be more robust. Similar patterns apply for double-normalised events, where the reduced number of time steps constrains accurate representation of peaks. Additionally, the de-dimensionalisation process leads to discretisation of some metric values and renders those based on absolute values uninterpretable. While these findings provide practical guidance for selecting appropriate metrics in contexts with limited or pre-processed data, a number of methodological uncertainties remain. In particular, it is often unclear whether metrics should be applied to the full rainfall event or to only the most intense burst. Additionally, the treatment of events with multiple peaks, or with no clear peak, remains insufficiently explored. Many metrics assume events have a single dominant peak, but the implications of including or excluding multi-peaked or uniform events in analyses are not well understood. Some research, e.g., Nguyen and Chen (2022), attempts to separate uniform rainfall events, using the Schutz Index, and to treat these differently. We recommend future research investigate this further, and consider filtering approaches based on peak structure, including the potential exclusion of events with multiple or indistinct peaks, to ensure appropriate and meaningful application of temporal loading-related to peak timing and intermittency are highly sensitive. The behaviour of magnitude concentration metrics was more variable, and dependent upon their specific formulation. In addition, transforming events to dimensionless mass curves alters the interpretability of some metrics, while leaving others largely unaffected. Given that rainfall data from gauges, radar, and reanalysis products are typically available at differing levels of temporal aggregation, these findings highlight that robustness to data representation is a key consideration. Metric selection should therefore be aligned not only with the aspect of storm structure of interest, but also with the characteristics and limitations of the available data.

Finally, it is important to reflect on the representativeness of the rainfall data used in this study. Our analysis draws on the high-quality, dense network of gauges across Denmark, providing over 233,000 well-observed rainfall events. This scale of dataset allows for robust evaluation of metric behaviour under real storm conditions. However, Denmark's temperate maritime climate is not fully representative of the range of rainfall regimes globally. In particular, strongly convective extremes typical of tropical and subtropical climates are less frequent. Overall, this work provides a clearer understanding of what temporal loading metrics measure, how they relate to one another, and storm dynamics may differ from those in arid or mountainous regions. As such, while our findings on the behaviour, sensitivity, and clustering of metrics are expected to generalise across contexts, the specific distribution of rainfall events and the relative prevalence of different temporal loading structures may not. We therefore recommend that future work apply similar analyses in a wider set of climatic settings, both to test the robustness of our methodological conclusions and to build a more globally representative basis for metric selection. how they behave

under different data conditions. It offers a foundation for more explicit, consistent, and reproducible characterisation of rainfall temporal loading, and supports clearer cross-study comparison going forward.

Code availability. Code is available at <https://github.com/masher92/MetricEvaluation/>

Appendix A: Literature review: search and selection strategy

The literature review follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework (Moher et al., 2010). The PRISMA framework is a standardised method of performing literature review which aims to minimise potential researcher bias, and to promote transparency and reproducibility in the identification and selection of relevant studies. The framework includes four key stages: (1) identification, boolean queries used to identify potential studies, duplicated studies removed; (2) screening, abstract level review to remove irrelevant studies; (3) eligibility, full-text review against pre-defined inclusion/exclusion criteria; and (4) inclusion, relevant studies included for review, and the total number of studies reported.

Boolean queries are search strings using AND, OR and NOT operators, and are constructed here by combining different terms used in the literature to describe event temporal loading, with different terms used to describe rainfall events. The full queries used to search the Scopus, Web of Science, and Google Scholar databases are outlined below:

- Scopus supports proximity operators (e.g., W/*n* requires two words to appear within *n* words of each other; Pre/*n* requires one word to precede the other within *n* words), a feature not available in the other databases. As such, for Scopus the search term was: [((("rainfall event" OR "precipitation event" OR "rain storm" OR "rainstorm" OR "storm event" OR (("precipitation" OR "rain" OR "rainfall" OR "storm") PRE/3 "event")) OR (("precipitation varia*" OR "rainfall varia*") W/1 "intra-event")) W/5 ("rainfall profile" OR "temporal pattern" OR "temporal distribution" OR "event loading" OR "temporal profile" OR "temporal loading" OR "intensity profile" OR (("precipitation pattern" OR "rainfall pattern") W/3 "temporal")))] ~~and-for-~~
- For Web of Science and Google Scholar, it was: ("rainfall event" OR "precipitation event" OR "rain storm" OR "rainstorm" OR "storm event" OR ("precipitation" OR "rain" OR "rainfall" OR "storm") AND "event" OR ("precipitation varia*" OR "rainfall varia*") AND "intra-event") AND ("rainfall profile" OR "temporal pattern" OR "temporal distribution" OR "event loading" OR "temporal profile" OR "temporal loading" OR "intensity profile" OR ("precipitation pattern" OR "rainfall pattern") AND "temporal")].

The searches were performed on the 15/02/2025 and returned 903 relevant records overall (including 181 on Scopus, 409 on Web of Science, and 313 on Google Scholar). 351 duplicated records were removed, leaving 552 studies for abstract level screening. After screening of abstracts, a further 390 studies were eliminated, leaving 162 papers for full review.

Studies were included if they met the following criteria: (a) use of search terms in the title, abstract, or keywords; (b) published in English; (c) focus on individual rainfall events with data at sub-daily temporal resolution; (d) application of a method or metric to quantify temporal loading. Exclusion criteria were: (a) not published in English; (b) inaccessible or behind paywalls (to the authors, not all included studies are open-access); (c) lack of use of sub-daily rainfall data; (d) combined spatial and temporal analysis (where the temporal loading could not be separated); (e) focus on rainfall trends over extended periods (e.g., seasonal or monthly scales); (f) no application of a metric for quantifying within-event temporal variability.

After applying these criteria, a total of 59 studies remained. We note, however, that several additional studies met all criteria except the use of a metric to quantify temporal loading. These excluded studies fall into three main groups. The first concerns the disaggregation of daily rainfall data (Pampaloni et al., 2021). The second includes stochastic event generation approaches (Cha et al., 2024; Efstratiadis et al., 2022; Kao and Govindaraju, 2008; Nguyen and Chen, 2022; Oriani et al., 2018; Pampaloni et al., 2021). The third involves applications of synthetic rainfall events with prescribed temporal patterns to investigate impacts on soil erosion (De Lima and Singh, 2002; de Lima et al., 2013; Frauenfeld and Truman, 2004; Gao et al., 2018; Parsons and Stone, 2006; Wang et al., 2020), landslides (Zhao et al., 2025), and herbicide runoff (Zhang et al., 1997). In these cases, event profiles are described using broad qualitative labels (e.g. linearly increasing, stepped increasing–decreasing, uniform) rather than being quantified with explicit metrics.

Appendix B: Selected papers

Table B1. Summary of reviewed papers

Authors	Metric type	Metric	Study purpose	Rainfall	Processing
Amin et al. (2000)	Classification	4th with most	Characterising typical profiles in an area	Real	DMC
Aquino et al. (2013)	Classification	3rd with peak	Impact of storm temporal pattern on soil erosion	Real	Raw
Asher et al. (2025b)	Classification	5th with most	Impact of storm temporal pattern on flood risk	Synthetic	DMC
Azli and Rao (2010)	Classification	4th with most	Characterising typical profiles in an area	Real	DMC
Back and Rodrigues (2021)	Classification	4th with peak	Characterising typical profiles in an area	Real	Unclear
Bezak et al. (2018)	Classification	4th with most	Impact of storm temporal pattern on hydraulic model outcomes	Synthetic	DMC
Brommer et al. (2013)	Summary — stats Continuous	mean, (C) std, (C) skewness, (C) kurtosis	Characterising typical profiles in an area	Real	Raw
Cai et al. (2024)	Summary — stats Continuous	Peak-position-ratio, Peak-position-ratio, PCI	Impact of storm temporal pattern on peak runoff generation	Synthetic	Raw
Chen et al. (2015)	Classification	4th with most	Characterising typical profiles in an area	Real	Unclear
de Assunção Montenegro et al. (2018)	Classification	3rd with peak	Impact of storm temporal pattern on soil moisture patterns	Real	Raw
de Andrade et al. (2020)	Classification, Summary — stats Continuous	130, 3rd with peak, 130	Impact of storm temporal pattern on soil erosion	Real	Raw
Dolšák et al. (2016)	Classification	4th with most, BSC	Characterising typical profiles in an area	Real	DMC
Dunkerley (2015)	Continuous	Intermittency fraction	Developing methods for describing temporal profile	Real	Raw
Dunkerley (2021b)	Intermittency Continuous	Intermittency fraction	Developing methods for describing temporal profile	Real	Raw
Fan et al. (2020)	Summary — stats Continuous	(C) skewness, (C) kurtosis	Impact of storm temporal pattern on landslides	Synthetic	Raw

Authors	Metric type	Metric	Study purpose	Rainfall	Processing
Fatone et al. (2021)	Classification	3rd with peak	Impact of storm temporal pattern on hydrograph outflow	Real	DMC
Fu et al. (2021)	Summary stats Continuous	Coefficient peak rainfall intensity, 3rd with ppr	Impact of typhoon temporal pattern on pollution generation	Real	Raw
Gao et al. (2024)	Summary stats Continuous	Relative amplitude	Impact of storm temporal pattern on soil erosion	Real	Raw
García-Bartual and Andrés-Doménech (2017)	Summary stats Continuous	(C) kurtosis, (C) standard deviation	Characterising typical profiles in an area	Real	Raw
Garcia-Guzman and Aranda-Oliver (1993)	Classification	4th with most	Characterising typical profiles in an area	Real	DMC
Ghanghas et al. (2024)	Summary stats Continuous	Event Loading Index	Impact of climate on temporal pattern	Real	Raw
He et al. (2022)	Summary stats Continuous	PCI	Characterising typical profiles in an area	Synthetic	Raw
He et al. (2024)	Summary stats Continuous	(C) skewness, (C) kurtosis	Impact of storm temporal pattern on slope instability	Synthetic	Raw
Horner and Jens (1942)	Classification	Third 3rd with peak	Impact of storm temporal pattern on surface runoff	Real	Unclear
Hu et al. (2018)	Summary stats Continuous	Gini coefficient, Lorenz coefficient	Changes to temporal inequality of daily precipitation over last 50 years	Synthetic	Raw
Huff (1967)	Classification	4th with most	Characterising typical profiles in an area	Real	DMC
Jeon et al. (2017)	Classification	4th with most	Impact of storm temporal pattern on water quality from storm runoff	Synthetic	DMC
Jun et al. (2021)	Summary stats Classification Continuous	4th with most, time to peak	Characterising typical profiles in an area	Real	Unclear
Knighton and Walter (2016)	Summary stats Continuous	T25 / 50 / 75, time to peak	Stochastic event generation	Real	DMC
Li et al. (2023)	Summary stats Continuous	Peak-mean ratio, peak-position-ratio 53 peak-position-ratio position ratio	Evaluating satellite precipitation products	Real	Raw
Liang et al. (2023)	Classification	3rd with most	Impact of storm temporal pattern on	Synthetic	DMC

Li et al. (2021)	Classification	4th with most	Characterising typical profiles in an area; Impact of storm temporal pattern on sewer overflows	Real	Raw
Liu et al. (2022)	Summary — stats Continuous	Relative amplitude, % rain HIZ, % time in HIZ / LIZ	Impact of storm temporal pattern on runoff and soil loss	Real	Raw
Long et al. (2021)	Summary — stats Continuous	TCI	Impact of climate on temporal (and spatial) pattern	Real	Unclear
Loveridge et al. (2015)	Classification	3rd with D50	Characterising typical profiles in an area	Real	DMC
Lu and Qin (2020)	Classification	4th with most	Impact of storm temporal pattern on urban drainage	Synthetic	DMC
Müller et al. (2017)	Summary — stats Continuous	Asymmetry of dependence	Impact of storm temporal patterns on sewer flow simulations	Real	Raw
Nel (2007)	Classification	4th with peak	Characterising typical profiles in an area; Impact of storm temporal pattern on runoff and soil erosion	Real	DMC
Oh et al. (2024)	Summary — stats Continuous	$NRMSE_p$ and $Skew_p$	Impact of storm temporal pattern on peak flood discharge	Real	Unclear
<u>Pandey et al. (2025)</u>	<u>Continuous</u>	<u>Intermittency fraction</u>	<u>Impact of climate on temporal pattern</u>	<u>Real</u>	<u>Raw</u>
Pinheiro et al. (2018)	Classification	3rd with peak	Characterising typical profiles in an area; Impact of storm temporal pattern on soil erosion	Real	DMC
Pohle et al. (2018)	Summary — stats; Intermittency Continuous	Event dry ratio, Frac. in $Q1 / Q2 / Q3 / Q4$	Stochastic event generation	Synthetic	Raw
Terranova and Iaquina (2011)	Classification	BSC	Characterising typical profiles in an area	Real	DMC
Tian et al. (2017)	Summary — stats Continuous	CV	Stochastic event generation	Synthetic	Raw
Todisco (2014)	Classification	CV, Run number	Developing methods for describing temporal profile; Impact of storm temporal pattern on soil erosion	Real	Raw

Serna and Taipe (2019)	Classification	3rd with peak	Characterising typical profiles in an area	Real	DMC
Schiff (1943)	Classification	3rd with peak	Impact of storm temporal pattern on infiltration and runoff	Real	DMC
Varga et al. (2009)	Classification	3rd with CoG	Developing methods for describing temporal profile	Real	DMC
Vantas et al. (2019)	Classification	4th with most	Developing methods for describing temporal profile	Real	Raw
Villalobos Herrera et al. (2023)	Classification	5th with most	Characterising typical profiles in an area	Real	DMC
Visser et al. (2023)	Summary — stats <u>Continuous</u>	D_{50}	Impact of climate on temporal pattern	Real	DMC
Wang et al. (2016)	Classification	3rd with peak	Impact of storm temporal pattern on soil erosive loss	Real	DMC
Wang et al. (2018)	Classification	3rd with ppr	Impact of storm temporal pattern on soil water transport	Real	DMC
Yang et al. (2015)	Summary — stats <u>Continuous</u>	Coefficient of variability	Assessing performance of data assimilation on events with different temporal loading characteristics	Real	Raw
Yang et al. (2022)	Summary — stats <u>Continuous</u>	Peak-position-ratio <u>position ratio</u>	Characterising typical profiles in an area Regional patterns in temporal profile	Real	Raw
Yang et al. (2024)	Classification	5th with peak	Impact of storm temporal pattern on flash flooding	Real	Unclear
Yin et al. (2016)	Classification	4th with peak	Characterising typical profiles in an area	Real	DMC
Wartalska et al. (2020)	Summary — stats <u>Continuous</u>	$m1, m2, m3, m4, m5,$ <u>peak-position-ratio</u> <u>peak position ratio</u>	Characterising typical profiles in an area	Real	DMC
Wartalska and Kotowski (2020)	Summary — stats <u>Continuous</u>	$m4, n_i$	Characterising typical profiles in an area; Impact of future climate on typical profiles.	Real	DMC
Williams-Sether (2004)	Classification	4th with most	Characterising typical profiles in an area	Real	DMC

Appendix C: [Summary of reviewed literature and metrics](#)

Appendix D: [Selected metrics](#)

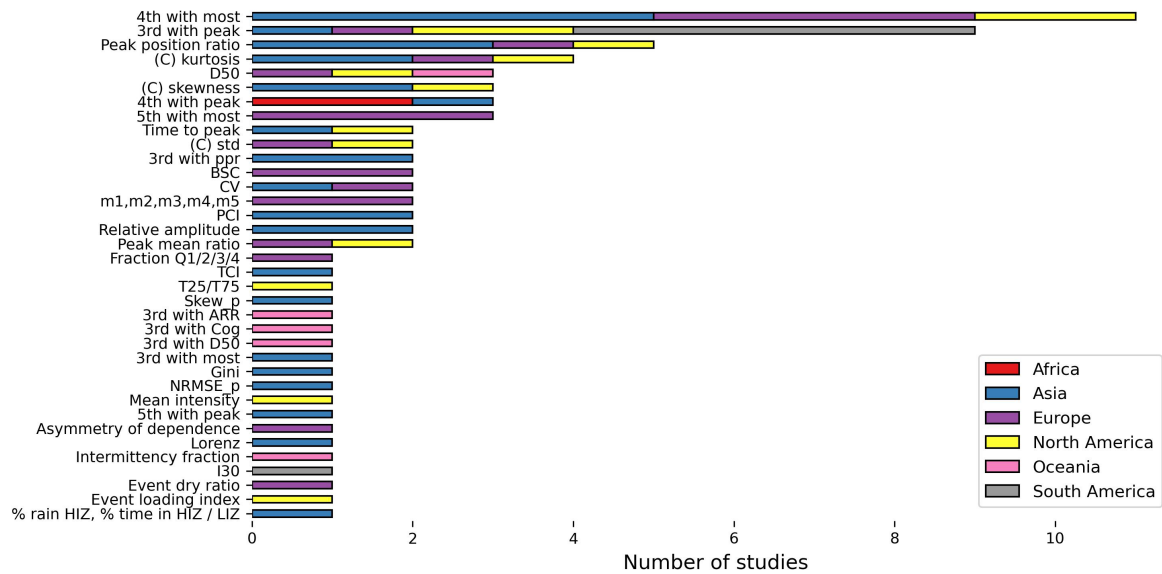


Figure C1. [Summary-Distribution of reviewed-metrics](#) ([Part across the reviewed studies, showing both the number of studies in which each metric was applied and the countries from which these studies were drawn.](#)

Appendix D: Silhouette score

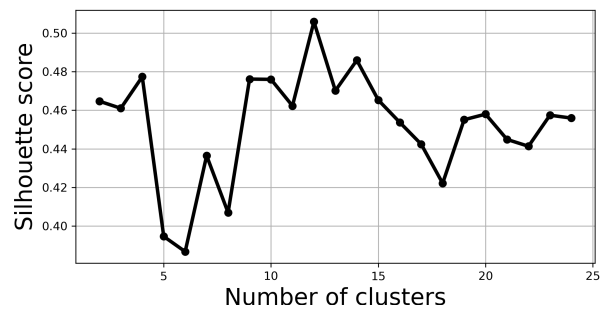


Figure D1. Silhouette scores for different numbers of clusters

1000 **Appendix E: Repeat of experiments in Section 4.2 using double normalised events**

Figures E1 and E2 present supplementary sensitivity analyses that isolate the contribution of de-dimensionalisation to the overall DMC sensitivity reported in Section 4.2, by evaluating sensitivity against double-normalised events that retain their original temporal resolution (i.e., without interpolation to 10 time steps).

1005 For the majority of metrics, retaining the original temporal resolution substantially reduces sensitivity, with Spearman's ρ approaching 1 ~~:Classification~~ and sMAPE approaching zero. This confirms that interpolation to ten data points is the dominant driver of sensitivity to DMC construction for most metrics, and that double normalisation alone has limited effect on ranking or value distributions across events.

1010 A notable exception is the cluster of metrics, including *maximum intensity*, *mean intensity*, *time to peak*, *standard deviation*, and *mean intensity within the HIZ*, which exhibit sMAPE values approaching 200. These metrics are defined in terms of absolute magnitudes, meaning that their values depend directly on the scale of the rainfall intensities. When applied to double normalised representations constrained to the range [0,1], this magnitude information is removed, and the resulting values are no longer directly comparable to those derived from the raw data. As a result, the difference between raw and normalised values is dominated by this scale mismatch, causing sMAPE to approach its upper bound. While variability is not eliminated under normalisation, it is substantially compressed relative to the raw values, which may span several orders of magnitude.

1015 This compression further inflates sMAPE, as differences are evaluated relative to a much smaller range. These results should therefore be interpreted as evidence of incompatibility between the metric definition and the normalised representation, rather than as an indication of genuine sensitivity.

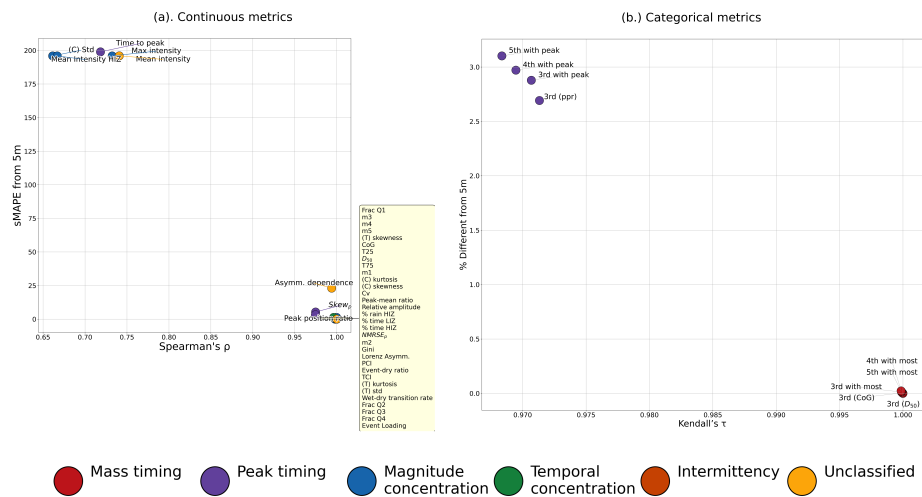


Figure E1. Sensitivity of metrics to calculation on double normalised events. For each metric, points represent comparisons between the metric calculated with 5-minute raw data and with a double normalised version of the event calculated on 5-minute data. For (a), continuous metrics the x-axis shows Spearman's rank correlation coefficient (ρ) and y-axis the symmetrical mean absolute percentage error (sMAPE). For (b), categorical metrics, the x-axis shows Kendall's tau and the y-axis the percent disagreement between the pairs of metric values.

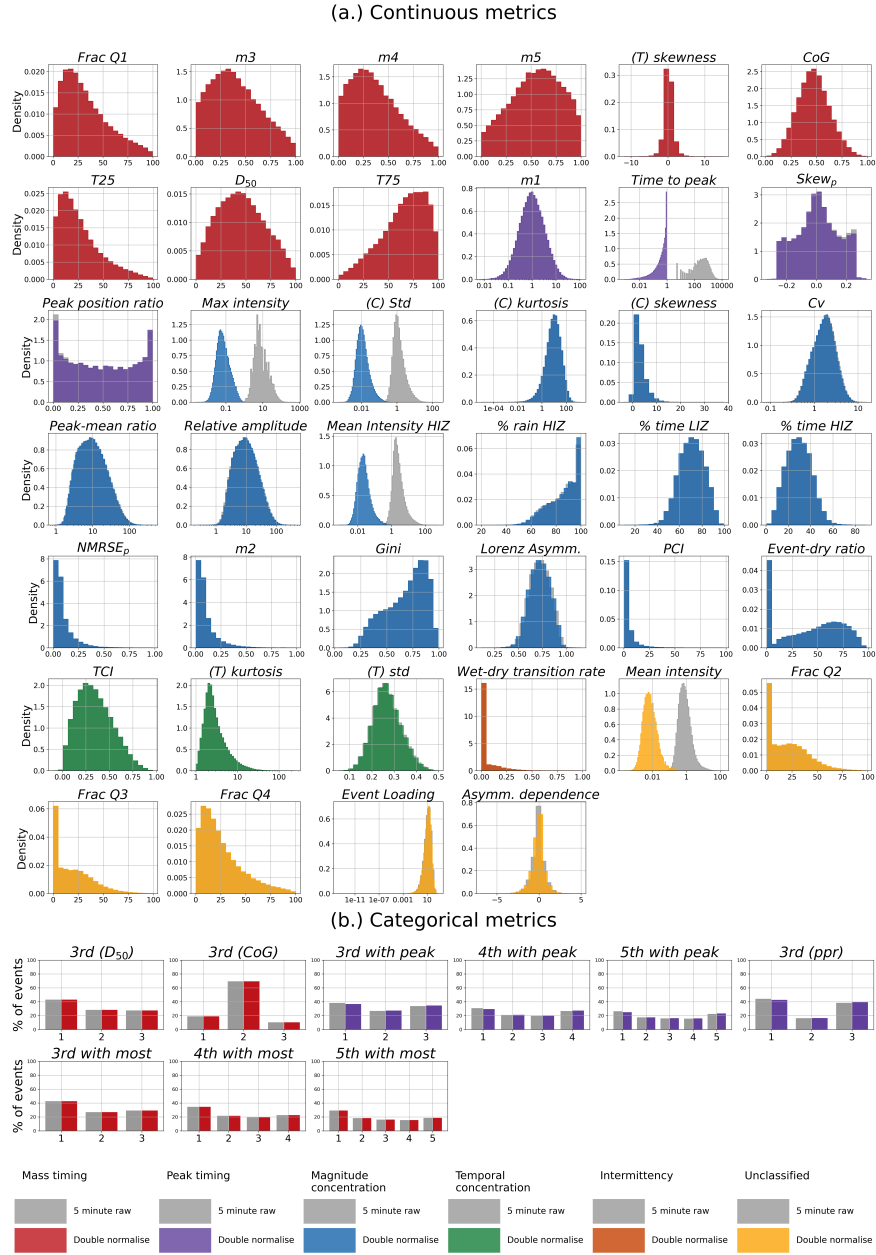


Figure E2. Sensitivity of metric distributions to calculation on double normalised events. Each subplot shows how the distribution of a given metric changes when calculated on raw rainfall data at 5-minute temporal resolution and on a double-normalised version of the event calculated on 5-minute data. For (a), continuous metrics, distributions are shown using overlaid histograms. For (b), categorical metrics, grouped bar charts show the relative frequency of each category across resolutions.

Appendix F: Summary of experimental results for selected metrics

Table F1: Summary of experimental results for investigated metrics

Metric	Meaning <u>Classification</u>	Focus	Robust to:
		<u>Temporal resolution</u>	Temp-res DMC transformation
3rd/4th/5th with peak rainfall Fraction of event with peak rainfall Peak timing 3rd/4th/5th with most rainfall Fraction of event with most rainfall Asymmetry ✓✓ 3rd with D₅₀ Third in which 50% of rainfall is reached Asymmetry ✓✓ 3rd with CoG Third containing centre of gravity of rainfall Asymmetry ✓✓ Max intensity <u>Frac. in Q1</u>	Maximum rainfall intensity <u>Mass timing</u>	Peakiness ×	×
Mean intensity <u>m3 / m4 / m5</u>	Mean rainfall intensity <u>Mass timing</u>	Peakiness ✓	×✓
130 <u>Temporal skewness [(T) skewness]</u>	Max rainfall depth in any 30-minute period <u>Mass timing</u>	Peakiness ✓	×✓
Classical <u>Centre of gravity (CoG) std</u>	Standard deviation: variation of rainfall intensity values around mean intensity <u>Mass timing</u>	Peakiness ✓	×✓

Summary of reviewed metrics (Part 2: Summary statistics) Metric	MeaningClassification	Focus	Robust to:
		Temporal resolution	Temp-res DMC transformation
Peakiness indicators			
Cv <u>T_{25} / T_{50} and D_{50} / T_{75}</u>	Coefficient of variation (std/mean): standardised variation of intensity values around mean intensity. <u>Mass timing</u>	Peakiness ✓	✗ ✓
Classical (C) <u>skewness</u> <u>ml</u>	Asymmetry of intensity value distribution. + skewness indicates most values small, some big, - skewness indicates most values big, some small. <u>Mass timing</u>	Peakiness ✗	✗
Classical (C) <u>kurtosis</u> <u>3rd with D_{50}</u>	Tailedness of intensity value distribution. + kurtosis indicates sharp, intense bursts concentrated in short periods, - kurtosis reflects rainfall spread more evenly over time without extreme peaks. <u>Mass timing</u>	Peakiness ✓	✗ ✗ ✓
Peak-mean ratio <u>3rd with CoG</u> ± Rainfall intensity irregularity (n_i)	Peak-to-mean ratio: maximum intensity divided by mean intensity <u>Mass timing</u>	Peakiness ✓	✗ ✗ ✓

Summary of reviewed metrics (Part 2: Summary statistics) Metric	Meaning	Classification	Focus	Robust to:
			<u>Temporal resolution</u>	Temp-res DMC transformation
Peakiness indicators				
<i>Relative amplitude (R_{am})</i> <u>3rd/4th/5th with most</u>	Range divided by mean	Peakiness ✓	✗ ✗ ✓	
<i>m2</i> <u>Time to peak</u>	% rainfall in highest intensity time step / rainfall in whole event	Peak timing	Peakiness ✗	✗
<i>Mean intensity in HIZ</i> <u>Skew_p</u>	Average intensity during above-average intensity periods	Peak timing	Peakiness ✗	✓ ✗
Asymmetry indicators <u>Peak position ratio / Time to peak</u> <u>Coefficient-peak-rainfall-intensity</u>	Time to peak in minutes	Peak timing	✗	✗
<i>Peak-position-ratio</i> <u>3rd with ppr</u> <u>coefficient-peak-rainfall-intensity</u>	Time to peak divided by duration	Peak timing	✗	✗
<i>Skew_p</i> <u>3rd/4th/5th with peak</u>	Value from -0.25 to 0.25 measuring relative position of peak within rainfall duration. indicates front-loading, + indicates back-loading	Peak timing	✗	✗
<i>T25 / T50 and D₅₀ / T75</i> <u>I30</u>	Percent of event passed when X% of rainfall happened	Magnitude <u>concentration</u>	Asymmetry of rainfall ✓	✓ ✗

Summary of reviewed metrics (Part 2: Summary statistics) Metric	Meaning Classification	Focus	Robust to:
		Temporal resolution	Temp-res DMC transformation
Peakiness indicators			
m3 / m4 / m5 <u>Max intensity</u>	% rainfall in first 33% + 30% + 50% of event duration <u>Magnitude concentration</u>	Asymmetry of rainfall	✓✓x
Metric <u>Classical standard deviation [(C) std]</u> Temp-res DMC transformation	Meaning <u>Magnitude concentration</u>	Focus ✓	Robust to: x
m1 <u>Classical kurtosis [(C) kurtosis]</u>	Relative volume of rainfall before versus after the peak <u>Magnitude concentration</u>	Asymmetry of rainfall	x
Centre of gravity <u>Classical skewness [(CoG) skewness]</u>	Timing of rainfall mass centre, snapped to nearest time step <u>Magnitude concentration</u>	Asymmetry of rainfall	✓✓x
Centre <u>Coefficient of gravity (interpolated) (CoG interp.)</u> <u>variation</u>	Fractional timing of rainfall mass centre across event duration <u>Magnitude concentration</u>	Asymmetry of rainfall	✓x

Summary of reviewed metrics (Part 2: Summary statistics) Metric	Meaning Classification	Focus	Robust to:
		Temporal resolution	Temp-res DMC transformation
Peakiness indicators			
Event loading index Peak-mean ratio / Rainfall intensity irregularity (n_i)	Comparison of temporal variability (captured by the standardised temporal heterogeneity, STH) of a storm against that of a symmetrised version constructed by mirroring the rising limb around the peak. Magnitude concentration	Asymmetry of rainfall ×	×
Asymmetry of dependence Relative amplitude	Describes the degree of reversibility in a timeseries using ranked-based statistics (copulas) Magnitude concentration	Asymmetry of rainfall ×	×
Temporal (T) skewness Mean intensity in HIZ	Skewness of rainfall distribution considered over time more front-loaded, + more back-loaded Magnitude concentration	Asymmetry of rainfall ✓	✓ ×

Summary of reviewed metrics (Part 2: Summary statistics)		Meaning	Classification	Focus	Robust to:
Metric				Temporal resolution	Temp-res DMC transformation
Peakiness indicators					
Concentration indicators	Precipitation			×	×
Concentration Index (PCI)	% rain in HIZ	Compares sum of squared intensities to square of total rainfall. Higher values indicate more rainfall concentrated in fewer steps. Magnitude concentration			
Temporal Concentration Index (TCI)	% time in LIZ / HIZ	Measures how strongly rain is clustered in time around each time-step; selects value for time step around which rainfall is most tightly clustered. A higher TCI means a more tightly concentrated “core” of high rain intensity. Temporal Magnitude concentration		✓ ×	✓ ×
NRMSE _p		Quantifies concentration of rainfall near the peak (RMSE from peak, normalised by rainfall total). 0 = uniform distribution, 1 = all rainfall in single time step. Magnitude concentration		Temporal Concentration ×	×

Summary of reviewed metrics (Part 2: Summary statistics)		Meaning <u>Classification</u>	Focus	Robust to:
Metric			<u>Temporal resolution</u>	Temp-res DMC transformation
Peakiness indicators				
Gini coefficient <u>m2</u>		Measures inequality of rainfall distribution (higher values indicate greater concentration of rainfall in fewer intervals) <u>Magnitude concentration</u>	×	×
<u>Gini coefficient</u>		Magnitude concentration	✓	✓
Lorenz asymmetry coefficient		Quantifies skewness of rainfall intensity distribution relative to the mean, distinguishing whether the concentration arises from a few extreme high-intensity bursts or from many moderately intense intervals, even if overall inequality is similar. <u>Magnitude concentration</u>	×	×

Summary of reviewed metrics (Part 2: Summary statistics)		Meaning <u>Classification</u>	Focus	Robust to:
Metric			<u>Temporal resolution</u>	Temp-res DMC transformation
Peakiness indicators				
Metric <u>Precipitation Concentration Index (PCI)</u>	Meaning Focus Robust to: Temp-res DMC transformation % time in LIZ / HIZ Time spent below / above mean intensity. Indicates how long the storm stays intense for. Magnitude concentration		×	×
% rain in HIZ <u>Event dry ratio</u>	Proportion of rainfall in above-average intensity periods. More rainfall in HIZ indicates a more concentrated event. Magnitude concentration		×	×
Frac. in Q1 / Q2 / Q3 / Q4 <u>Temporal Concentration Index (TCI)</u>	% total rainfall in each quarter — Magnitude <u>Temporal</u> concentration		×✓	×✓
Temporal <u>kurtosis</u> [(T) kurtosis]	Whether rainfall is very concentrated in time around the centre of gravity, or spread out in the tails Temporal concentration		✓	✓

Summary of reviewed metrics (Part 2: Summary statistics)	Meaning	Classification	Focus	Robust to:
Metric			Temporal resolution	Temp-res DMC transformation
Peakiness indicators				
Temporal <u>standard deviation</u> [(T) std]			✓	✓
	How spread out rainfall is in time around centre of gravity, weighted by heaviness of rain	Temporal concentration		
Summary of reviewed metrics (Part 3: Intermittency metrics)	Meaning	Intermittency	Focus ✗	Robust to: ✗
Metric <u>Wet-dry transition rate</u>				
<u>Event loading index</u>		Unclassified	✗	Temp-res DMC transformation ✗
<u>Event-dry ratio</u> <u>Asymmetry of dependence</u>				×
	Proportion of event duration spent at 0 mm/hr	Intermittency	×	
	Unclassified			
<u>Intermittency fraction</u> <u>Mean intensity</u>				×
	Frequency of wet/dry transitions in the event	Intermittency	✓	
	Unclassified			
<u>Frac. in Q2 / Q3 / Q4</u>		Unclassified	×	✗
Appendix G: Silhouette score Silhouette scores for different numbers of clusters				

1020 *Author contributions.* Conceptualisation and methodology was by MA and JP, Funding acquisition and supervision by MT and CB. Data
curation was by MA, JP and RH. Formal analysis was by MA and RH. Investigation and visualisation was by MA. The original manuscript
was written by MA, with review and editing by JP, MT, CB and SB.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to thank the Water Pollution Committee of The Society of Danish Engineers for allowing us to use their
data for this study.
1025 We also extend thanks to Lawrence Jackson for his guidance on the statistical analysis of metrics performed in this study.

References

- Alavinia, M., Saleh, F. N., and Asadi, H.: Effects of rainfall patterns on runoff and rainfall-induced erosion, *International Journal of Sediment Research*, 34, 270–278, 2019.
- Amin, S., Sepaskhah, A. R., and Keshmiripour, B.: Rainfall intensity distribution patterns in climatological conditions of four stations in Fars province, Iran, *Iranian Journal of Science and Technology*, 24, 89–98, 2000.
- Amponsah, A. O., Daraio, J. A., and Khan, A. A.: Implications of climatic variations in temporal precipitation patterns for the development of design storms in Newfoundland and Labrador, *Canadian Journal of Civil Engineering*, 46, 1128–1141, 2019.
- Aquino, R. F., Silva, M. L. N., Freitas, D. A. F. d., Curi, N., and Avanzi, J. C.: Soil losses from typic Cambisols and Red Latosol as related to three erosive rainfall patterns, *Revista Brasileira de Ciência do Solo*, 37, 213–220, 2013.
- 1035 Aryal, R., Furumai, H., Nakajima, F., and Jinadasa, H.: The role of inter-event time definition and recovery of initial/depression loss for the accuracy in quantitative simulations of highway runoff, *Urban Water Journal*, 4, 53–58, 2007.
- Asher, M., Pedersen, J. W., Böing, S., Birch, C., Trigg, M., and Kendon, E.: Extreme rainfall and temporal loading in Great Britain: Analysis of present and future trends using a convection-permitting climate model, *Journal of Hydrology: Regional Studies*, 62, 102 750, <https://doi.org/10.1016/j.ejrh.2025.102750>, 2025a.
- 1040 Asher, M., Trigg, M., Böing, S., and Birch, C.: The Sensitivity of Urban Pluvial Flooding to the Temporal Distribution of Rainfall Within Design Storms, *Journal of Flood Risk Management*, 18, e70 097, 2025b.
- Azli, M. and Rao, A. R.: Development of Huff curves for peninsular Malaysia, *Journal of Hydrology*, 388, 77–84, 2010.
- Back, A. J. and Rodrigues, M. L. G.: Characterization of temporal rainfall distribution in Florianópolis, Santa Catarina, Brazil, *Revista Brasileira de Climatologia*, 28, 201–219, 2021.
- 1045 Berne, A. and Krajewski, W. F.: Radar for hydrology: Unfulfilled promise or unrecognized potential?, *Advances in Water Resources*, 51, 357–366, 2013.
- Bezak, N., Šraj, M., Rusjan, S., and Mikoš, M.: Impact of the rainfall duration and temporal rainfall distribution defined using the Huff curves on the hydraulic flood modelling results, *Geosciences*, 8, 69, 2018.
- Boughton, W. and Droop, O.: Continuous simulation for design flood estimation—a review, *Environmental Modelling & Software*, 18, 309–318, 2003.
- 1050 Brommer, D. M., Cervený, R. S., and Balling Jr, R. C.: An analysis of intra-event precipitation variability for the United States (1980–2010), *Physical Geography*, 34, 456–470, 2013.
- Cai, Y., Yang, Y., Tan, Q., Dai, C., Zhu, Z., and Zhang, X.: Impacts of temporal/spatial rainfall heterogeneities on peak runoff distribution and intensities for an urban river basin of south China, *River*, 3, 24–37, 2024.
- 1055 Centre for Ecology & Hydrology: The Flood Estimation Handbook (five volumes). Volume 3: Statistical procedures for flood frequency estimation, Centre for Ecology & Hydrology, Wallingford, UK, 1999.
- Cha, H., Baik, J., Lee, J., Na, W., Bateni, S. M., and Jun, C.: Generation of rainfall scenarios based on rainfall transition probability to determine temporal distribution of independent rainstorms, *Stochastic Environmental Research and Risk Assessment*, pp. 1–19, 2024.
- Chen, Z., Yin, L., Chen, X., Wei, S., and Zhu, Z.: Research on the characteristics of urban rainstorm pattern in the humid area of Southern China: a case study of Guangzhou City., *International Journal of Climatology*, 35, 2015.
- 1060

- de Andrade, A. F., Machado, R. L., Rezende, C. C., Ferreira, E. A., de Águas, M. A., Correchel, V., and de Carvalho, D. F.: Precipitation Patterns and Rainfall Erosivity Return Period Under Savanna Conditions in Formosa, Goiás, Brazil, *Journal of Agricultural Studies*, 8, 554–569, 2020.
- de Assunção Montenegro, A. A., dos Santos Souza, T. E. M., de Souza, E. R., and Montenegro, S. M. G. L.: Temporal dynamics of soil moisture and rainfall erosivity in a tropical volcanic archipelago, *Journal of Hydrology*, 563, 737–749, 2018.
- De Lima, J. and Singh, V.: The influence of the pattern of moving rainstorms on overland flow, *Advances in Water Resources*, 25, 817–828, 2002.
- de Lima, J. L., Carvalho, S. C., and P de Lima, M. I.: Rainfall simulator experiments on the importance of when rainfall burst occurs during storm events on runoff and soil loss, *Zeitschrift fur Geomorphologie Supplement*, 57, 91–109, 2013.
- 1070 Dey, P.: On the Structure of the Intermittency of Rainfall, *Water Resources Management*, 37, 1461–1472, 2023.
- DMI: Climate Data — DMI Open Data, https://opendatadocs.dmi.govcloud.dk/en/Data/Climate_Data, accessed: 2025-06-22, 2025.
- Dolšák, D., Bezak, N., and Šraj, M.: Temporal characteristics of rainfall events under three climate types in Slovenia, *Journal of Hydrology*, 541, 1395–1405, 2016.
- Dunkerley, D.: Identifying individual rain events from pluviograph records: a review with analysis of data from an Australian dryland site, *Hydrological Processes: An International Journal*, 22, 5024–5036, 2008.
- 1075 Dunkerley, D.: Intra-event intermittency of rainfall: An analysis of the metrics of rain and no-rain periods, *Hydrological Processes*, 29, 3294–3305, 2015.
- Dunkerley, D.: The importance of incorporating rain intensity profiles in rainfall simulation studies of infiltration, runoff production, soil erosion, and related landsurface processes, *Journal of Hydrology*, 603, 126 834, 2021a.
- 1080 Dunkerley, D.: Intermittency of rainfall at sub-daily timescales: New quantitative indices based on the number, duration, and sequencing of interruptions to rainfall, *Atmospheric Research*, 253, 105 475, 2021b.
- Dunkerley, D.: Huff quartile classification of rainfall intensity profiles ('storm patterns'): A modified approach employing an intensity threshold, *Catena*, 216, 106 371, 2022.
- Dunkerley, D.: Rainfall intensity and intermittency over timescales of minutes: Measurement using low sampling-rate acoustic recording, *Atmospheric Research*, 290, 106 814, 2023.
- 1085 Dunkerley, D. L.: How do the rain rates of sub-event intervals such as the maximum 5-and 15-min rates (I5 or I30) relate to the properties of the enclosing rainfall event?, *Hydrological Processes*, 24, 2425–2439, 2010.
- Dunkerley, D. L.: Rainfall intensity bursts and the erosion of soils: An analysis highlighting the need for high temporal resolution rainfall data for research under current and future climates, *Earth Surface Dynamics*, 7, 345–360, 2019.
- 1090 Efstratiadis, A., Dimas, P., Pouliasis, G., Tsoukalas, I., Kossieris, P., Bellos, V., Sakki, G.-K., Makropoulos, C., and Michas, S.: Revisiting flood hazard assessment practices under a hybrid stochastic simulation framework, *Water*, 14, 457, 2022.
- Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirmbach, M., Vaes, G., and Vieux, B.: Towards a roadmap for use of radar rainfall data in urban drainage, *Journal of Hydrology*, 299, 186–202, 2004.
- Fan, L., Lehmann, P., Zheng, C., and Or, D.: Rainfall intensity temporal patterns affect shallow landslide triggering and hazard evolution, *Geophysical Research Letters*, 47, e2019GL085 994, 2020.
- 1095 Fatone, F., Szeląg, B., Kiczko, A., Majerek, D., Majewska, M., Drewnowski, J., and Łagód, G.: Advanced sensitivity analysis of the impact of the temporal distribution and intensity of rainfall on hydrograph parameters in urban catchments, *Hydrology and Earth System Sciences*, 25, 5493–5516, 2021.

- Formetta, G., Prosdocimi, I., Stewart, E., and Bell, V.: Estimating the index flood with continuous hydrological models: an application in Great Britain, *Hydrology Research*, 49, 123–133, 2018.
- Fowler, H. J., Lenderink, G., Prein, A. F., Westra, S., Allan, R. P., Ban, N., Barbero, R., Berg, P., Blenkinsop, S., Do, H. X., et al.: Anthropogenic intensification of short-duration rainfall extremes, *Nature Reviews Earth & Environment*, 2, 107–122, 2021.
- Frauenfeld, B. and Truman, C.: Variable rainfall intensity effects on runoff and interrill erosion from two coastal plain ultisols in Georgia, *Soil Science*, 169, 143–154, 2004.
- Freitas, E. d. S., Coelho, V. H. R., Xuan, Y., de CD Melo, D., Gadelha, A. N., Santos, E. A., Galvão, C. d. O., Ramos Filho, G. M., Barbosa, L. R., Huffman, G. J., et al.: The performance of the IMERG satellite-based product in identifying sub-daily rainfall events and their properties, *Journal of Hydrology*, 589, 125 128, 2020.
- Fu, X., Liu, J., Mei, C., Luan, Q., Wang, H., Shao, W., Sun, P., and Huo, Y.: Effect of typhoon rainstorm patterns on the spatio-temporal distribution of non-point source pollution in a coastal urbanized watershed, *Journal of Cleaner Production*, 292, 126 098, 2021.
- Gao, G., Liang, Y., Liu, J., Dunkerley, D., and Fu, B.: A modified RUSLE model to simulate soil erosion under different ecological restoration types in the loess hilly area, *International Soil and Water Conservation Research*, 12, 258–266, 2024.
- Gao, J., Kirkby, M., and Holden, J.: The effect of interactions between rainfall patterns and land-cover change on flood peaks in upland peatlands, *Journal of Hydrology*, 567, 546–559, 2018.
- García-Bartual, R. and Andrés-Doménech, I.: A two-parameter design storm for Mediterranean convective rainfall, *Hydrology and Earth System Sciences*, 21, 2377–2387, 2017.
- Garcia-Guzman, A. and Aranda-Oliver, E.: A stochastic model of dimensionless hyetograph, *Water Resources Research*, 29, 2363–2370, 1993.
- Ghanghas, A., Sharma, A., and Merwade, V.: Unveiling the evolution of extreme rainfall storm structure across space and time in a warming climate, *Earth’s Future*, 12, e2024EF004 675, 2024.
- Gholami, L., Khaledi Darvishan, A., Spalevic, V., Cerdà, A., and Kavian, A.: Effect of storm pattern on soil erosion in damaged rangeland: Field rainfall simulation approach, *Journal of Mountain Science*, 18, 706–715, <https://doi.org/10.1007/s11629-020-6157-1>, 2021.
- Gregersen, I. B., Sørup, H. J. D., Madsen, H., Rosbjerg, D., Mikkelsen, P. S., and Arnbjerg-Nielsen, K.: Assessing future climatic changes of rainfall extremes at small spatio-temporal scales, *Climatic change*, 118, 783–797, 2013.
- Haile, A., Rientjes, T., Habib, E., Jetten, V., and Gebremichael, M.: Rain event properties at the source of the Blue Nile River, *Hydrology and Earth System Sciences*, 15, 1023–1034, 2011.
- Harshanth, R., Dauji, S., and Srivastava, P.: Development of site-specific time distribution pattern of rainfall for Tarapur, India, *Journal of Earth System Science*, 132, 77, 2023.
- He, S., Wang, Z., Wang, D., Liao, W., Wu, X., and Lai, C.: Spatiotemporal variability of event-based rainstorm: The perspective of rainfall pattern and concentration, *International Journal of Climatology*, 42, 6258–6276, 2022.
- He, Z., Huang, Y., Li, Y., and Li, X.: Probabilistic fragility assessment of slopes considering uncertainty associated with temporal patterns of rainfall intensity, *Computers and Geotechnics*, 173, 106 534, 2024.
- Hettiarachchi, S., Wasko, C., and Sharma, A.: Increase in flood risk resulting from climate change in a developed urban watershed—the role of storm temporal patterns, *Hydrology and Earth System Sciences*, 22, 2041–2056, 2018.
- Horner, W. and Jens, S.: Surface runoff determination from rainfall without using coefficients, *Transactions of the American Society of Civil Engineers*, 107, 1039–1075, 1942.

- Hu, S., Feng, F., Liu, W., and She, D.: Changes in temporal inequality and persistence of precipitation over China during the period 1961–2013, *Hydrology Research*, 49, 1283–1291, 2018.
- Huff, F. A.: Time distribution of rainfall in heavy storms, *Water resources research*, 3, 1007–1019, 1967.
- Jain, A. K., Murty, M. N., and Flynn, P. J.: Data clustering: a review, *ACM computing surveys (CSUR)*, 31, 264–323, 1999.
- 1140 Jeon, D. J., Ki, S. J., Baek, S.-S., Cha, Y., Cho, K. H., Yoon, K.-S., Shin, H. S., and Kim, J. H.: Assessing the efficiency of aggregate low impact development (LID) at a small urbanized sub-catchment under different storm scenarios, *Desalination and Water Treatment*, 86, 1–8, 2017.
- Jun, C., Qin, X., Chen, M., and Seo, H.: Investigating event-based temporal patterns of design rainfall in a tropical region, *Hydrological Sciences Journal*, 66, 1986–1996, 2021.
- 1145 Kao, S.-C. and Govindaraju, R. S.: Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas, *Water Resources Research*, 44, 2008.
- Keifer, C. J. and Chu, H. H.: Synthetic storm pattern for drainage design, *Journal of the hydraulics division*, 83, 1332–1, 1957.
- Kendall, M. G.: A new measure of rank correlation, *Biometrika*, 30, 81–93, 1938.
- Kimura, N., Tai, A., Chiang, S., Wei, H.-P., Su, Y.-F., Cheng, C.-T., and Kitoh, A.: Hydrological flood simulation using a design hyetograph
 1150 created from extreme weather data of a high-resolution atmospheric general circulation model, *Water*, 6, 345–366, 2014.
- Klamkowski, K., Treder, W., and Tryngiel-Gač, A.: The effect of rainfall intensity and floor management practices on changes in soil water status in an apple orchard, pp. 539–544, 2012.
- Knighton, J. O. and Walter, M. T.: Critical rainfall statistics for predicting watershed flood responses: rethinking the design storm concept, *Hydrological processes*, 30, 3788–3803, 2016.
- 1155 Kottegoda, N. and Kassim, A.: Classification of storm profiles using crossing properties, *Journal of Hydrology*, 127, 37–53, 1991.
- Koutsoyiannis, D. and Foufoula-Georgiou, E.: A scaling model of a storm hyetograph, *Water Resources Research*, 29, 2345–2361, 1993.
- Krvavica, N. and Rubinić, J.: Evaluation of design storms and critical rainfall durations for flood prediction in partially urbanized catchments, *Water*, 12, 2044, 2020.
- Lenderink, G. and Van Meijgaard, E.: Unexpected rise in extreme precipitation caused by a shift in rain type?, *Nature Geoscience*, 2, 373–373,
 1160 2009.
- Li, R., Guilloteau, C., Kirstetter, P.-E., and Foufoula-Georgiou, E.: How well does the IMERG satellite precipitation product capture the timing of precipitation events?, *Journal of Hydrology*, 620, 129 563, 2023.
- Li, T., Lee, G., and Kim, G.: Case study of urban flood inundation—impact of temporal variability in rainfall events, *Water*, 13, 3438, 2021.
- Liang, Y., Gao, G., Liu, J., Dunkerley, D., and Fu, B.: Runoff and soil loss responses of restoration vegetation under natural rainfall patterns
 1165 in the Loess Plateau of China: The role of rainfall intensity fluctuation, *Catena*, 225, 107 013, 2023.
- Liu, J., Liang, Y., Gao, G., Dunkerley, D., and Fu, B.: Quantifying the effects of rainfall intensity fluctuation on runoff and soil loss: From indicators to models, *Journal of Hydrology*, 607, 127 494, 2022.
- Long, K., Wang, D., Wang, G., Zhu, J., Wang, S., and Xie, S.: Higher temperature enhances spatiotemporal concentration of rainfall, *Journal of Hydrometeorology*, 22, 3159–3169, 2021.
- 1170 Loukas, A. and Quick, M. C.: Spatial and temporal distribution of storm precipitation in southwestern British Columbia, *Journal of hydrology*, 174, 37–56, 1996.
- Loveridge, M., Babister, M., and Retallick, M.: Australian Rainfall and Runoff Revision Project 3: Temporal Patterns of Rainfall, Engineers Australia, Barton, 2015.

- Lu, W. and Qin, X.: Integrated framework for assessing climate change impact on extreme rainfall and the urban drainage system, *Hydrology Research*, 51, 77–89, 2020.
- Lu, X., Wen, T., Zhang, L., and Zhang, Q.: Impact of Spatiotemporal Rainfall Distribution and Underlying Surface Changes on Flood Processes in Meijiang River Basin, China, *Water*, 17, 466, 2025.
- Makridakis, S. and Hibon, M.: Evaluating accuracy (or error) measures, 1995.
- Medina-Cobo, M., García-Marín, A., Estévez, J., and Ayuso-Muñoz, J.: The identification of an appropriate Minimum Inter-event Time (MIT) based on multifractal characterization of rainfall data series, *Hydrological Processes*, 30, 3507–3517, 2016.
- Meier, C. I., Moraga, J. S., Pranzini, G., and Molnar, P.: Describing the interannual variability of precipitation with the derived distribution approach: effects of record length and resolution, *Hydrology and Earth System Sciences*, 20, 4177–4190, 2016.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Group, P., et al.: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *International journal of surgery*, 8, 336–341, 2010.
- Molina-Sanchis, I., Lázaro, R., Arnau-Rosalén, E., and Calvo-Cases, A.: Rainfall timing and runoff: The influence of the criterion for rain event separation, *Journal of Hydrology and Hydromechanics*, 64, 226, 2016.
- Müller, T., Schütze, M., and Bárdossy, A.: Temporal asymmetry in precipitation time series and its influence on flow simulations in combined sewer systems, *Advances in Water Resources*, 107, 56–64, 2017.
- Nel, W.: Intra-storm attributes of extreme storm events in the Drakensberg, South Africa, *Physical Geography*, 28, 158–169, 2007.
- Ng, C. W., Wang, B., and Tung, Y.-K.: Three-dimensional numerical investigations of groundwater responses in an unsaturated slope subjected to various rainfall patterns, *Canadian Geotechnical Journal*, 38, 1049–1062, 2001.
- Nguyen, D. T. and Chen, S.-T.: Generating continuous rainfall time series with high temporal resolution by using a stochastic rainfall generator with a copula and modified huff rainfall curves, *Water*, 14, 2123, 2022.
- Oh, B., Kim, J., and Hwang, S.: Influence of Rainfall Patterns on Rainfall–Runoff Processes: Indices for the Quantification of Temporal Distribution of Rainfall., *Water* (20734441), 16, 2024.
- Oliver, J. E.: Monthly precipitation distribution: a comparative index, *The Professional Geographer*, 32, 300–309, 1980.
- Oriani, F., Mehrotra, R., Mariethoz, G., Straubhaar, J., Sharma, A., and Renard, P.: Simulating rainfall time-series: how to account for statistical variability at multiple scales?, *Stochastic Environmental Research and Risk Assessment*, 32, 321–340, 2018.
- Pampaloni, M., Sordo-Ward, A., Bianucci, P., Gabriel-Martin, I., Caporali, E., and Garrote, L.: A stochastic procedure for temporal disaggregation of daily rainfall data in SuDS design, *Water*, 13, 403, 2021.
- Pan, C., Wang, X., Liu, L., Huang, H., and Wang, D.: Improvement to the huff curve for design storms and urban flooding simulations in Guangzhou, China, *Water*, 9, 411, 2017.
- Pandey, R. S., Carrassi, A., Porcù, F., and Adirosi, E.: High resolution analysis of rainfall intermittency at mid-latitude using disdrometer decade-long data, Available at SSRN 5489686, 2025.
- Parsons, A. and Stone, P.: Effects of intra-storm variations in rainfall intensity on interrill runoff and erosion, *Catena*, 67, 68–78, 2006.
- Pathiraja, S., Westra, S., and Sharma, A.: Why continuous simulation? The role of antecedent moisture in design flood estimation, *Water Resources Research*, 48, 2012.
- Pinheiro, A. G., Souza, T. E. d. S., Montenegro, S. M., Montenegro, A. A. d. A., and Guerra, S.: Rainfall pattern and erosion potential in the physiographic regions of the state of Pernambuco, Brazil, *Revista Brasileira de Engenharia Agrícola e Ambiental*, 22, 849–853, 2018.
- Pohle, I., Niebisch, M., Müller, H., Schümborg, S., Zha, T., Maurer, T., and Hinz, C.: Coupling Poisson rectangular pulse and multiplicative microcanonical random cascade models to generate sub-daily precipitation timeseries, *Journal of Hydrology*, 562, 50–70, 2018.

- Restrepo-Posada, P. J. and Eagleson, P. S.: Identification of independent rainstorms, *Journal of Hydrology*, 55, 303–319, 1982.
- Rivera, D., Arumí, J. L., and Lillo-Saavedredra, M.: Spatio-Temporal Patterns in Soil Water Content Time Series: Influence of the Time Series Length and Precipitation Events, p. 1, 2012.
- 1215 Römken, M. J., Helming, K., and Prasad, S.: Soil erosion under different rainfall intensities, surface roughness, and soil water regimes, *Catena*, 46, 103–123, 2002.
- Rousseeuw, Peter, J.: A graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics*, 20, 53–65, 1987.
- Schiff, L.: Classes and pattern of rainfall with reference to surface-runoff, *Eos, Transactions American Geophysical Union*, 24, 438–452, 1220 1943.
- Schilling, W.: Rainfall data for urban hydrology: what do we need?, *Atmospheric Research*, 27, 5–21, 1991.
- Schleiss, M.: How intermittency affects the rate at which rainfall extremes respond to changes in temperature, *Earth System Dynamics*, 9, 955–968, 2018.
- Schleiss, M. and Smith, J. A.: Two simple metrics for quantifying rainfall intermittency: The burstiness and memory of interamount times, 1225 *Journal of Hydrometeorology*, 17, 421–436, 2016.
- Serna, J. R. V. and Taipe, C. L. R.: Determination of storm profiles in the Central Andes of Peru, in: E-proceedings of the 38th IAHR World Congress, 2019.
- Smith, A. A. and Lee, K.-B.: The rational method revisited, *Canadian Journal of Civil Engineering*, 11, 854–862, 1984.
- Terranova, O. and Iaquina, P.: Temporal properties of rainfall events in Calabria (southern Italy), *Natural hazards and earth system sciences*, 1230 11, 751–757, 2011.
- Thomassen, E. D., Arnbjerg-Nielsen, K., Sørup, H. J. D., Langen, P. L., Olsson, J., Pedersen, R. A., and Christensen, O. B.: Spatial and temporal characteristics of extreme rainfall: Added benefits with sub-kilometre-resolution climate model simulations?, *Quarterly Journal of the Royal Meteorological Society*, 149, 1913–1931, 2023.
- Tian, J., Liu, J., Yan, D., Li, C., and Yu, F.: Numerical rainfall simulation with different spatial and temporal evenness by using a WRF 1235 multiphysics ensemble, *Natural Hazards and Earth System Sciences*, 17, 563–579, 2017.
- Todisco, F.: The internal structure of erosive and non-erosive storm events for interpretation of erosive processes and rainfall simulation, *Journal of Hydrology*, 519, 3651–3663, 2014.
- Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, *Bulletin of the American Meteorological Society*, 84, 1205–1218, 2003.
- 1240 Tsubo, M., Walker, S., and Hensley, M.: Quantifying risk for water harvesting under semi-arid conditions: Part I. Rainfall intensity generation, *Agricultural Water Management*, 76, 77–93, 2005.
- Urgilés, G., Céleri, R., Bendix, J., and Orellana-Alvear, J.: Identification of spatio-temporal patterns in extreme rainfall events in the Tropical Andes: A clustering analysis approach, *Meteorological Applications*, 31, e70005, 2024.
- Vantas, K., Sidiropoulos, E., and Vafeiadis, M.: Optimal temporal distribution curves for the classification of heavy precipitation using 1245 hierarchical clustering on principal components., 2019.
- Varga, C., Ball, J. E., and Babister, M.: A hydroinformatic approach to development of design temporal patterns of rainfall, IAHS-AISH Publication, 2009.
- Villalobos Herrera, R., Blenkinsop, S., Guerreiro, S. B., Dale, M., Faulkner, D., and Fowler, H. J.: Towards new design rainfall profiles for the United Kingdom, *Journal of Flood Risk Management*, p. e12958, 2023.

- 1250 Visser, J. B., Wasko, C., Sharma, A., and Nathan, R.: Changing storm temporal patterns with increasing temperatures across Australia, *Journal of Climate*, 36, 6247–6259, 2023.
- Wang, F.: Temporal pattern analysis of local rainstorm events in China during the flood season based on time series clustering, *Water*, 12, 725, 2020.
- Wang, G., Sun, W., Xue, B., Kiem, A., et al.: Stratification response of soil water content during rainfall events under different rainfall
1255 patterns., *Hydrological processes*, 32, 2018.
- Wang, H., Chen, R., Leung, A. K., and Garg, A.: Hydrological responses to early-peak rainfall in unsaturated rooted soils, *Heliyon*, 9, 2023.
- Wang, W., Yin, S., Xie, Y., Liu, B., and Liu, Y.: Effects of four storm patterns on soil loss from five soils under natural rainfall, *Catena*, 141, 56–65, 2016.
- Wang, W., Yin, S., Xie, Y., and Nearing, M. A.: Minimum inter-event times for rainfall in the eastern monsoon region of China, *Transactions*
1260 *of the ASABE*, 62, 9–18, 2019.
- Wartalska, K. and Kotowski, A.: Model hyetographs of short-term rainfall for Wrocław in the perspective of 2050, *Atmosphere*, 11, 663, 2020.
- Wartalska, K., Kaźmierczak, B., Nowakowska, M., and Kotowski, A.: Analysis of hyetographs for drainage system modeling, *Water*, 12, 149, 2020.
- 1265 Wasko, C. and Sharma, A.: Steeper temporal distribution of rain intensity at higher temperatures within Australian storms, *Nature Geoscience*, 8, 527–529, 2015.
- Watt, E. and Marsalek, J.: Critical review of the evolution of the design storm event concept, *Canadian Journal of Civil Engineering*, 40, 105–113, 2013.
- Williams-Sether, T.: Empirical, dimensionless, cumulative-rainfall hyetographs developed from 1959-86 storm data for selected small wa-
1270 tersheds in Texas, 2004, US Geological Survey, 2004.
- Wu, S.-J., Yang, J.-C., and Tung, Y.-K.: Identification and stochastic generation of representative rainfall temporal patterns in Hong Kong territory, *Stochastic environmental research and risk assessment*, 20, 171–183, 2006.
- Yang, J., Duan, K., Wu, J., Qin, X., Shi, P., Liu, H., Xie, X., Zhang, X., and Sun, J.: Effect of data assimilation using WRF-3DVAR for heavy rain prediction on the northeastern edge of the Tibetan Plateau, *Advances in Meteorology*, 2015, 294 589, 2015.
- 1275 Yang, J., Xiang, Y., Xu, X., and Sun, J.: Design hyetograph for short-duration rainstorm in Jiangsu, *Atmosphere*, 13, 899, 2022.
- Yang, P., Xu, Z., Yan, X., and Wang, X.: Establishing a rainfall dual-threshold for flash flood early warning considering rainfall patterns in mountainous catchment, China, *Natural Hazards*, 120, 6657–6684, 2024.
- Yeo, I.-K. and Johnson, R. A.: A new family of power transformations to improve normality or symmetry, *Biometrika*, 87, 954–959, 2000.
- Yin, S.-q., Xie, Y., Nearing, M. A., Guo, W.-l., and Zhu, Z.-y.: Intra-storm temporal patterns of rainfall in China using Huff curves, *Transac-*
1280 *tions of the ASABE*, 59, 1619–1632, 2016.
- Zhang, X., Norton, L., and Hickman, M.: Rain pattern and soil moisture content effects on atrazine and metolachlor losses in runoff, *Tech. rep.*, Wiley Online Library, 1997.
- Zhang, Y., Liang, K., and Liu, C.: Time distribution pattern and spatial heterogeneity of hourly scale event-based extreme precipitation in China, *Journal of Hydrology*, 622, 129 712, 2023.
- 1285 Zhao, B., Marin, R. J., Luo, W., Yu, Z., and Yuan, L.: Rainfall thresholds for shallow landslides considering rainfall temporal patterns, *Bulletin of Engineering Geology and the Environment*, 84, 1–13, 2025.