Hydrology and
Earth System
Sciences

Open Access

EGU

# Forecasting green roof detention performance by temporal downscaling of precipitation time-series projections

**Vincent Pons**[1,2]**, Rasmus Benestad**[3]**, Edvard Sivertsen**[4]**, Tone Merete Muthanna**[1]**, and Jean-Luc Bertrand-Krajewski**[2]

[1]Department of Civil and Environmental Engineering, The Norwegian University of Science and Technology,
Trondheim, 7031, Norway
[2]Univ Lyon, INSA Lyon, DEEP, EA7429, 11 rue de la Physique, 69621, Villeurbanne cedex, France
[3]Norwegian Meteorological Institute, Oslo, Norway
[4]SINTEF AS, S.P. Andersens veg 3, 7465 Trondheim, Norway

**Correspondence:** Vincent Pons (vincent.pons@ntnu.no)

**Abstract.** A strategy to evaluate the suitability of different multiplicative random cascades to produce rainfall time series, taking into account climate change, inputs for green infrastructures models. The multiplicative random cascades reproduce a (multi)fractal distribution of precipitation through an iterative and multiplicative random process. In the current study, the initial model, a flexible cascade that deviates from multifractal scale invariance, was improved with (i) a temperature dependency and (ii) an additional function to reproduce the temporal structure of rainfall. The structure of the models with depth and temperature dependency was found to be applicable in eight locations studied across Norway and France. The resulting time series from both reference period and projection based on RCP 8.5 were applied to two green roofs with different properties. The different models led to a slight change in the performance of green roofs, but this was not significant compared to the range of outcomes due to ensemble uncertainty in climate modelling and the stochastic uncertainty due to the nature of the process. The hydrological dampening effect of the green infrastructure was found to decrease in most of the Norwegian cities due to an increase in precipitation, especially Bergen (Norway), while slightly increasing in Marseille (France) due to decrease in rainfall event frequency.

## 1 Introduction

Hydrologic performance of stormwater green infrastructure (GI) is usually divided between retention and detention. Retention refers to water stored, infiltrated, or evapotranspirated. Actual evapotranspiration can be estimated from a water balance including potential evapotranspiration, accumulated precipitation, a soil moisture evaluation function, and a crop factor (Johannessen et al., 2017; Oudin et al., 2005). The temporal resolution for modelled evapotranspiration process for green infrastructure is typically daily (Stovin et al., 2013) or hourly (Kristvik et al., 2019). Detention refers to water temporarily stored in the GI before being discharged into a downstream stormwater network. The process temporal resolution is typically minutes. Consequently, modelling GI detention performance requires higher-resolution data to estimate its outflow (Schilling, 1991). Therefore, both high-resolution climate data and projections at subdaily and subhourly scales are needed in order to model GI and to estimate their potential as a climate change adaptation measure.

In Norway and most of the European countries, precipitation has been measured with tipping buckets in numerous cities from years to decades. Moreover, climate projections at daily resolution for future precipitation and temperature from the EURO-CORDEX project are available at $1 \times 1$ km spatial resolution in Norway (Dyrrdal et al., 2018) and $12 \times 12$ km resolution in France (Jacob et al., 2014). Consequently, the use of such data by urban hydrologists to assess the resilience of GI solutions to face climate change is conditioned by the possibility to downscale them to a subhourly resolution.

Downscaling includes two families of methods: dynamical downscaling and statistical downscaling (Benestad, 2016). Dynamical downscaling methods use physically based equations and are usually computationally expensive specially to generate high-resolution data. Statistical downscaling consists in improving the resolution of data based on statistical properties observed from a lower-resolution dataset. The computational cost is lower; therefore, statistical methods might still be used to fill the gap in the next decades until the computational power is sufficient for routine dynamical downscaling. In addition, the use of a stochastic approach is necessary due to the current limitation in parametrization of small-scale processes (below the truncation scale) and the lack of coupling between resolved and parameterized scales (Sanchez et al., 2016).

Statistical downscaling has already been extensively used to temporally downscale data for various temporal resolutions, usually hourly or daily data. Three popular methods can be mentioned: (i) the method of fragment, (ii) the method based on point process theory, and (iii) the method of multiplicative random cascades. The method of fragment (Li et al., 2018; Lu et al., 2015) is a resampling method based on $k$-nearest neighbours (Kalra and Ahmad, 2011), which has been applied to derive hourly data from daily data. It can be accurate and effective due to its resampling nature, but it requires a large dataset, and by its design it cannot ensure extrapolation from observed data. Therefore, it might not be suitable to downscale climate projections. Methods based on point process theory have been used (Glasbey et al., 1995; Onof et al., 2000). The main principle is to generate storm occurrences and then describe them based on rain cells and statistical distribution based on Poisson point process. Multiplicative random cascades (MRCs) consist of using successively random cascades to split data in $N$ data of finer resolution ($N = 2$ in most of the cases). It is a very popular method that deserves further investigation (Gaur and Lacasse, 2018; Rupp et al., 2012; Thober et al., 2014). They were originally based on the hypothesis of multifractal scale invariance (Schertzer and Lovejoy, 1987) and were further developed by Gupta and Waymire (1993) and Olsson (1998). While multifractal scale invariance continues to be studied (Gires et al., 2020), several studies noted a deviation from that behaviour which led to the use of more flexible models (Koutsoyiannis and Langousis, 2011; Veneziano et al., 2006). Multiplicative random cascades can be divided between canonical and micro-canonical types. The canonical one ensures conservation on average, while the micro-canonical one ensures exact conservation. The parameters of the canonical MRC are often calibrated by fitting between observed and simulated non-centred moments of depths or intensity through the timescale (Paschalis et al., 2012). The principle of micro-canonical MRC is usually based on reverse cascades: studying how the data are split and then reproducing the properties of the weights distribution depending on different quantities. The influences of timescale, rainfall intensity (Paschalis et al.,

2012; Rupp et al., 2009) or season (McIntyre et al., 2016) have been extensively studied. Lombardo et al. (2012) suggested that the commonly used MRC suffers from conceptual weaknesses due to the non-stationary process of autocorrelation and proposed a method to improve the model. More recently, Bürger et al. (2014, 2019) suggested to include a temperature dependency in MRC models to make them more robust. This also enables them to be used with projections.

Green infrastructures, due to their retention and detention capacities, are seen as a promising solution to manage storm water and cope with climate change, especially in cities where urbanization increases. Among green infrastructures, green roofs are especially suitable for dense urban centres. They are designed to retain day-to-day rain by evapotranspiration and attenuate major rainfall events (Stovin, 2010). Depending on their characteristics they can also help to detain extreme rainfall (Hamouz et al., 2020). Due to the timescale of their detention process, and their sensitivity to initial water content at the beginning of a rainfall event, they are suitable for evaluating downscaled time series. Moreover, it is especially relevant to evaluate their detention performance by the end of the century under a scenario such as RCP 8.5 (Thorndahl and Andersen, 2021). The results could be used to evaluate, at strategic level, their potential in mitigating storm water in order to make robust decision (Walker et al., 2013).

While downscaling models have been used to model the performance of green infrastructure under current climate (Stovin et al., 2017) or applied to intensity–duration–frequency (IDF) curves to do an event-based simulation of local stormwater measures (Kristvik et al., 2019), none has been developed to produce future high-resolution time series as input for green infrastructure models. The aim of this research is to evaluate different MRC downscaling models and their potential to produce input time series to predict the performance of stormwater green infrastructure, for the case of green roofs. In order to achieve this aim, different parts are detailed in the paper: (i) the development of a general structure of MRC, (ii) the improvement of the MRC structure by adding a temperature dependency, (iii) the addition of an ordering function to improve the temporal structure of the produced rainfall time series, (iv) the evaluation of the capability to reproduce the performance of GI based on observed data, and finally (v) the analysis of a possible shift in performance of GI at the end of the century.

## 2 Methods

### 2.1 Meteorological data

Time series of precipitation and temperature from six locations in Norway and two in France, representing four different climates (Table 1) according to the Köppen–Geiger classification (Peel et al., 2007), were used to apply the downscaling method. In Norway, the precipitation was measured

by 0.2 mm plumatic tipping rain gauge manufactured by Kongsberg Våpenfabrikk. The rain gauges were not heated and thus did not operate in cold temperature. They were successively replaced by Lambrecht 1518H3 (measuring tip of 0.1 mm) in the 1990s and 2000s. The stations were operated by the Norwegian Water Resources and Energy Directorate (NVE) and the Norwegian Meteorological Institute (MET). The data were quality checked by the Norwegian Meteorological institute (MET) (Lutz et al., 2020). In Lyon and Marseille, precipitation was measured by 0.2 mm Précis-Mécanique tipping bucket rain gauges. Ten climate projections (temperature and precipitation) on daily resolution with the RCP 8.5 for the period from 2071 to 2099 for Norwegian cities were available online https://nedlasting.nve.no/klimadata/kss (last access: 26 May 2022; Dyrrdal et al., 2018). For Lyon and Marseille (France), 12 climate projections at daily resolution were available for the same period and RCP (2071 to 2099, RCP 8.5) from http://www.drias-climat.fr/ (last access: 25 May 2022). The RCP 8.5, a scenario with a high gas-emission baseline leading to a radiative forcing of $8.5 \, \mathrm{W \, m^{-2}}$ in 2100, and the end of the century were chosen to test the methods since it was the scenario and period that deviate the most from the current climate among the available data in both countries. In practice, it is relevant to evaluate GI performance at the end of the century, but their design could be based on a different period depending on their lifetime.

## 2.2 Downscaling models and workflow

### 2.2.1 Data aggregation and processing

The historical data were aggregated two by two from 1 min resolution (and accordingly 6 min) to more than 1 d resolution in order to capture a part of the uncertainty linked to the estimation of the parameter of the models. The aggregation was done for each possible time steps: all multiples of 2 smaller than 1500 min (as there are 1440 min in a day). During the process of aggregation, both the weights, Eq. (1), and the rainfall continuity indicator, Eq. (2), measuring the proportion of high weight on the side of the highest neighbouring depth were computed. Given $j \in [1 \ldots 750]$ a timescale in minutes, $i \in [0 \ldots N_{2j}]$ a time step with $N_{2j}$ the number of time steps at scale $2j$, and $d_{i,2j}$ a rainfall depth, the weight $w_{i,2j}$, and the indicator $S_{i,2j}$ of the side of the neighbour were calculated according to

$$w_{i,2j} = \frac{\min(d_{2i,j}, d_{2i+1,j})}{d_{2i,j} + d_{2i+1,j}} \in [0; 0.5] \tag{1}$$

$$S_{i,2j} = \begin{cases} 0, & \text{if } (d_{i-1,2j} = d_{i+1,2j}) \cup (d_{2i,j} = d_{2i+1,j}) \\ 1, & \text{if } (d_{2i,j} > d_{2i+1,j} \cap d_{i-1,2j} > d_{i+1,2j}) \\ & \cup (d_{2i,j} < d_{2i+1,j} \cap d_{i-1,2j} < d_{i+1,2j}) \\ 2, & \text{if } (d_{2i,j} > d_{2i+1,j} \cap d_{i-1,2j} < d_{i+1,2j}) \\ & \cup (d_{2i,j} < d_{2i+1,j} \cap d_{i-1,2j} > d_{i+1,2j}). \end{cases} \tag{2}$$

### 2.2.2 Downscaling process

The MRC downscaling process consists of transforming daily rainfall depths to rainfall depths at lower timescale, e.g. 1 min, by means of successive distribution of the depth of a parent time steps between its two children time steps. The process is repeated by iteration until the desired timescale is reached. Figure 1 describes the downscaling process. In practice, the downscaling started at 1440 min (1 d) time step with eight iterations to reach a time step of 5.625 min. The results were interpolated and scaled to a 6 min time step for comparison with observed data. The final time step of 6 min was chosen based on the resolution of original datasets in Lyon and Marseille. Three steps are necessary to downscale a parent time step to two children time steps. The occurrence of a zero weight, i.e. the probability to assign all the water from the parent time step to only one of the children time steps (Fig. 1, centre left), is tested. This property is especially important and acknowledged by other studies. If a zero weight does not occur, a non-zero weight $w_{i,2j} \in ]0, 0.5]$ is generated from a probability distribution (Eq. 3b). It distributes the depth from the parent time step between the two children time steps, as illustrated in Fig. 1, centre right. Finally, the weights $w_{i,2j}$ and $1 - w_{i,2j}$ have to be assigned to the children time steps. The occurrence of $S_W$ (Eq. 4), i.e. allocating the highest weight to the children with the neighbour with the highest depth, is tested (Fig. 1, bottom).

$$u_{0,i,2j} \sim \mathcal{U}([0, 1]),$$

$$\begin{aligned} &\text{if } u_{0,i,2j} < P\big(W = 0 | S_{\text{time}} = 2j, \ D = d_{i,2j}, \\ &\quad T = T_{i,2j}\big), \text{ then } w_{i,2j} = 0 \end{aligned} \tag{3a}$$

$$\text{else, } w_{i,2j} \sim \mathcal{N}_{[0,\frac{1}{2}]}\left(\frac{1}{2}, \ \sigma(S_{\text{time}} = 2j)^2\right) \tag{3b}$$

$$P(S_W = 1), \ \text{with} \ S_W \in \{0; 1\} \tag{4}$$

### 2.2.3 Downscaling models conceptualization and calibration

Based on the observed data, six different MRC models were developed. Different mathematical expressions and probability distributions, detailed in Appendix A, were defined to represent Eqs. (3a), (3b) and (4), depending on the hypothesis inherent to the later described models (Table 2). The models consist in three generators: a zero-weight generator, a non-zero-weight generator and a stochastic element permutation generator (SEP generator). Each of the zero-weight and non-zero-weight generators (Eqs. 3a and 3b) were considered to vary with timescale (indicated with S in the model naming). The letter I in the nomenclature indicates a depth/Intensity for the zero-weight generator (Eqs. A2a–A2d). The temperature dependency for the zero-weight generator (Eqs. A3a–A3e) was indicated by the letter T in the nomenclature. The

**Table 1.** Locations and input data for current and future climate. The climate column gives the Köppen–Geiger classification for climate. Observed days is the number of observed days with data. YearPr is the annual precipitation in millimetre, YearWt the annual number of wet days (> 1 mm). YearTe is the mean annual temperature; for these quantities the 5th, 50th, and 95th percentiles are given as indicators.

| Location | Observed days | Climate | Latitude | Period | YearPr | YearWt | YearTe |
|---|---|---|---|---|---|---|---|
| Bergen – Sandsli, MET 50 480, NVE no. 56.1. | 6150 | Cfb | 60.4 | Obs: RCP 8.5: | 1505, 2081, 2504 2240, 3012, 4009 | 153, 189, 218 169, 201, 238 | −2.1, 8.0, 17.8 2.2, 10.3, 19.3 |
| Bodø – Skivika, MET 82 310, NVE no. 165.11. | 7204 | Dfc | 67.3 | Obs: RCP 8.5: | 643, 991, 1858 1150, 1600, 2139 | 114, 152, 266 147, 178, 214 | −3.9, 5.4, 15.4 −0.3, 8.1, 18.2 |
| Lyon (France), 6 min time step | 7671 | Cfb | 45.7 | Obs: RCP 8.5: | 706, 865, 1161 550, 830, 1187 | 80, 97, 114 77, 105, 135 | 0.5, 12.8, 24.3 3.9, 15.9, 29.9 |
| Hamar – Hamar II (Disen), MET 12 290 | 4011 | Dfb | 60.8 | Obs: RCP 8.5: | 406, 546, 659 508, 689, 861 | 70, 92, 105 88, 110, 134 | −9.8, 5.7, 18.6 −5.3, 8.1, 21.0 |
| Kristiansand – Sømskleiva, MET 39 150, NVE no. 21.49. | 5219 | Cfb | 58.1 | Obs: RCP 8.5: | 1155, 1512, 1868 1258, 1662, 2099 | 120, 137, 161 115, 142, 167 | −0.9, 9.7, 18.6 0.3, 10.0, 20.5 |
| Kristiansund – Karihola, MET 64 300, NVE no. 110.1. | 8664 | Cfb | 63.1 | Obs: RCP 8.5: | 714, 1094, 2521 1440, 2051, 2829 | 131, 167, 226 153, 192, 230 | −0.7, 7.8, 16.2 1.9, 9.6, 18.0 |
| Marseille (France), 6 min time step | 7305 | Csa | 43.3 | Obs: RCP 8.5: | 310, 533, 840 250, 492, 767 | 35, 49, 64 33, 50, 68 | 4.1, 15.2, 26.2 7.3, 18.0, 30.6 |
| Trondheim – Risvollan, MET 68 230, NVE 123.38. | 10 722 | Dfc | 63.4 | Obs: RCP 8.5: | 669, 965, 1256 853, 1176, 1599 | 117, 150, 173 133, 163, 196 | −6.1, 6.3, 17.9 −1.4, 8.6, 19.2 |

temperature dependency was added in an attempt to improve the robustness of the model toward climate change under the hypothesis that the change in rainfall pattern would be correlated to the shift in temperature and that the existing observed datasets already carry the necessary information for calibration. In the models $MRC_{S-SEP}$, $MRC_{SI-SEP}$ and $MRC_{SIT-SEP}$ (Table 2) the weights generated were permuted stochastically depending on the neighbour (indicated with SEP, Eqs. 4 and A5), while the $MRC_S$, $MRC_{SI}$, and $MRC_{SIT}$ model considered equal probability (0.5) to permute the two children weights.

Similarly to Rupp et al. (2009), the generators of the MRC models include timescale dependency through analytical formulas. In practice it means that there is a single set of parameters per generator and not a different set at each cascade step. It ensured a relatively parsimonious number of parameters compared to other recent works on microcanonical MRC where a set per cascade step is often used (e.g. 12 to 36 total parameters by Bürger et al., 2019 or from 6 to 224 parameters per cascade step by Müller-Thomy, 2020). It should

be noted that based on dataset analysis and as advised by Serinaldi (2010), despite the fact that universal and canonical MRC represent the most parsimonious approach parameterwise, their microcanonical counterpart was preferred. This choice of model that deviates from the hypothesis of multifractal scale invariance was supported by several other studies (Koutsoyiannis and Langousis, 2011; Veneziano et al., 2006). The number of parameters can be lower with the use of universal MRCs which were excluded from this study due to lack of flexibility (Serinaldi, 2010). It also allows the model to be used with any desired initial resolution lower than 1500 min. Homogeneity of the resolution in the input datasets was not required for calibration and data processing (i.e. the model can be calibrated using multiple datasets with different resolutions between 1 min and 1 d). The parameters of each generator of MRC models and each location required calibration. A single-step calibration, based on the processed data, was sufficient for generators with only timescale dependency. A multiple-step calibration with data manipulation was necessary for generators with depth/inten-
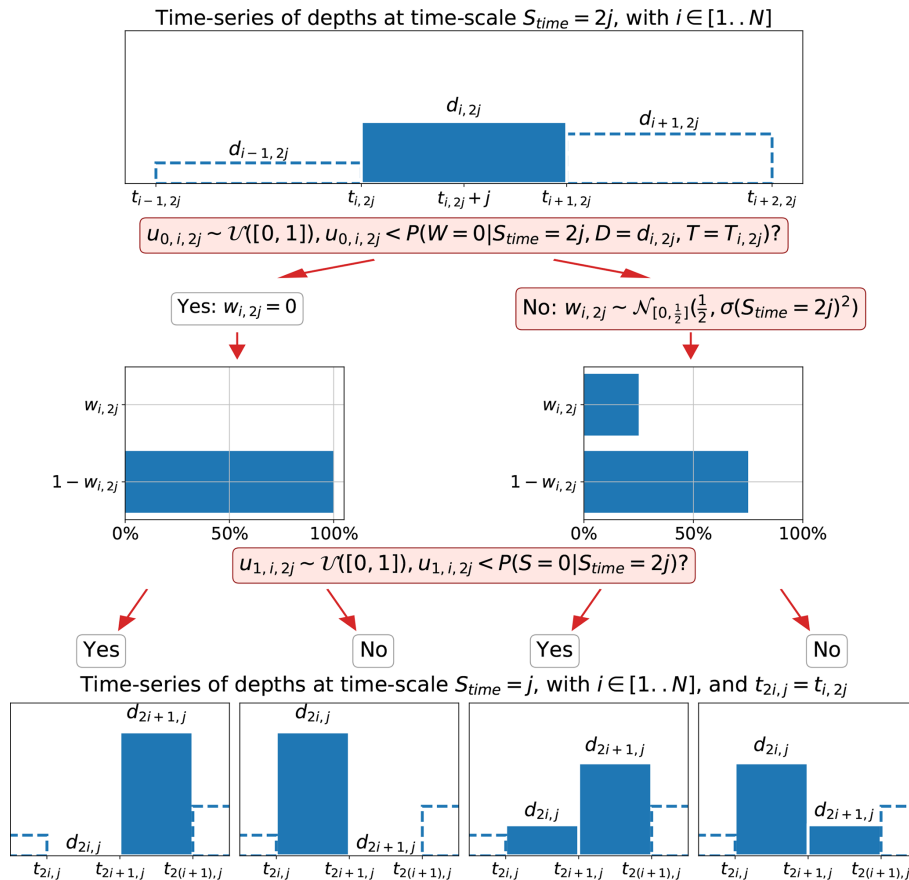
**Figure 1.** Workflow for downscaling to transfer a depth from time step $T$ to time step $\frac{T}{2}$. The red boxes involve the generation of a random number. The process starts with 1440 min time step to reach 5.625 min. An interpolation is then done to reach a 6 min time step.

**Table 2.** Nomenclature of the models and various quantities taken into account by each model depending on the process considered; $S$ is the timescale, $D$ the rainfall depth/intensity, $T$ the temperature and N the close neighbour.

| Model | $P(W=0)$ | | | | $P(W\|W \neq 0)$ | | | | $P(S_W = 1)$ | | | | Number of |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $S$ | $I$ | $T$ | $N$ | $S$ | $I$ | $T$ | $N$ | $S$ | $I$ | $T$ | $N$ | parameters |
| MRC$_S$ | × | | | | × | | | | | | | | 8 |
| MRC$_{S\text{-SEP}}$ | × | | | | × | | | | × | | × | | 13 |
| MRC$_{SI}$ | × | × | | | × | | | | | | | | 14 |
| MRC$_{SI\text{-SEP}}$ | × | × | | | × | | | | × | | × | | 19 |
| MRC$_{SIT}$ | × | × | × | | × | | | | | | | | 13 |
| MRC$_{SIT\text{-SEP}}$ | × | × | × | | × | | | | × | | × | | 18 |

sity, temperature dependency, and for the non-zero-weight generator. This choice was motivated by the development of the model through data analysis and conceptualization of the model. Especially, the steps and calibration of the three generators were chosen to avoid compensation between processes using a bottom-up approach (i.e. starting from local properties and then add dependencies to progressively upscale the model). Additionally, the choice of regression over timescale was used to avoid parameter sets that lead to the correct distribution of precipitation intensities without temporal consistency. Later studies can further improve the procedure to make it more easily calibrated. The optimizations were based on non-linear least squares in the standard library scipy.optimize with default parameters in Python (Virtanen et al., 2020).

– The parameters of zero-weight generator with only timescale dependency (Eq. A1) followed a single-step calibration against observed zero-weight proportions by non-linear least squares.

– The parameters of the zero-weight generator with timescale and depth dependency (Eqs. A2a–A2d) followed a two-step calibration: (i) for each timescale, the proportion of zero-weight depending on depth was evaluated using a weighted running window to compensate for rare occurrence of extreme depths. The proportion of zero weight depending on depth was then fitted to a function (Eq. A2a). (ii) The functions modelling the parameters depending on timescale were then calibrated by least squares (parameters of Eqs. A2b, A2c and A2d).

– The parameters of zero-weight generator with timescale, depth and temperature dependency (Eqs. A3a–A3e) followed a similar calibration procedure. (i) Using running windows of temperature, the proportion of zero-weight depending on depth was fitted by least squares for different temperatures (Eq. A3a). (ii) Given a timescale the parameters depending on temperature were fitted to a Gaussian function (Eq. A3b). (iii) The parameters of the Gaussian function depending on timescale were then fitted to set of functions by least squares (Eqs. A3c, A3d and A3e).

– The non-zero-weight generator consisted of a truncated normal distribution on [0, 0.5] with $\mu = 0.5$ (Eq. 3b) and a function $\sigma$ depending on timescale (Eq. A4). It was chosen against more commonly used beta distributions (McIntyre et al., 2016) after a goodness-of-fit test was applied to the historical data. The calibration was done in two steps. (i) $\sigma$ was evaluated by non-linear least squares for each timescale. (ii) The parameters of Eq. (A4) were calibrated against the evaluated $\sigma$ depending on timescale by least squares.

– The parameters of the SEP generator (Eq. A5) followed a single-step calibration by least squares with processed proportion of high weight on the side of highest neighbour depending on timescale.

## 2.3 Green Infrastructure modelling

In order to quantify the influence of rainfall input in green roof performance estimation, two green roofs located in Trondheim were modelled. They were selected due to data availability and the contrast of their behaviours: (i) a typical extensive green roof (E-Green roof) with sedum vegetation, 30 mm of substrate, and 10 mm "eggbox" drainage layer (Hamouz and Muthanna, 2019), and (ii) a detention-based extensive green roof (D-Green roof) with sedum vegetation, 30 mm of substrate, and 100 mm of lightweight clay aggregates (Hamouz et al., 2020). The model (Eqs. 5a–5c) was a simple reservoir model with smoothed linear function (Eq. 5c) for the outflow, Oudin's model for potential evapotranspiration (PET, Eq. 5b) and a soil moisture evaluation

function to estimate actual evapotranspiration (AET) (Johannessen et al., 2017).

$$WC_i = WC_{i-1} + P_{i-1} - Q_{i-1} \\ - WC_{i-1} \times PET_i \times C \tag{5a}$$

$$PET_i = \begin{cases} 0, & \text{if } T_i <= 5\,°C \\ \frac{R_a}{\lambda \rho} \times 0.01 \times (T_{mean} + 5), & \text{if } T_i > 5\,°C \end{cases} \tag{5b}$$

$$Q_i = \begin{cases} \frac{S_K}{1+\exp(-\frac{4 \times K}{S_K} \times (WC_i - WC_K - \frac{S_K - 1}{2 \times K}))}, \\ \qquad \text{if } WC_i > WC_K + \frac{S_K - 1}{2 \times K} \\ \\ K \times (WC_i - WC_K) + \frac{1}{2}, \\ \qquad \text{else} \end{cases} \tag{5c}$$

$WC_i$ is the water content (mm) at time $t_i$. $P_i$ is the precipitation (mm min$^{-1}$). The discharge $Q_i$ (mm min$^{-1}$) is based on the empirical (Eq. 5c). The temperature $T_{mean}$ is in degrees Celsius, and the extra-terrestrial radiation $R_a$ is derived from the latitude and the Julian day. The constant $\frac{1}{\lambda \rho} \approx 0.408$ depends on latent heat and volumetric mass of water. The factor $C$ is a calibrated factor depending on the maximum storage and the crop factor. The smoothed linear function (Eq. 5c) has three parameters: $K$ the conductivity slope, $S_K$ the smoothing factor and $WC_K$ the starting delay. The model was developed based on data from extreme tests with artificial precipitation (Hamouz et al., 2020) by establishing a relationship between water content and runoff. The outflow depending on water content was used as input for calibration of the parameters of the discharge function using a Bayesian calibration with DREAM setup (Laloy and Vrugt, 2012). It should be noted that the model remains limited as it lumps processes and neglects dynamical effect, i.e. the wetting of the aggregates and substrate and the spatial distribution of water content within the roof (Hamouz et al., 2020). The D-Green roof's model was tested against measured discharge with, as input, a rainfall series of 2.5 months from 10 July to 25 September 2018 and a 1-month series from 5 September to 5 October 2019. The E-Green roof's model was tested against measured discharge with a rainfall series from April 2017 to September 2017 as input. Snow periods were mostly excluded for the evaluation.

## 2.4 Evaluation of the downscaling models

For each location, the observed precipitations were aggregated to daily resolution and downscaled to obtain 200 time series of 6 min time step. They were used to model the extensive and detention-based extensive green roofs in parallel. It should be noted that irrigation needs and snow periods were neglected since the primary objective of the study was to evaluate the produced time series. There were 10 projections available in Norway for the RCP 8.5 and 12 in France with the EURO-CORDEX project. Each projected time series was downscaled 20 times (200 simulations for Norwe-

gian locations, and 240 simulations for French locations) to capture the following: (i) the variability between the projections and (ii) the variability due to the nature of the downscaling model. The number of simulations per location and per period was chosen to ensure reasonably low simulation time and represent the stochastic uncertainty inherent to the downscaling process. The stability of the percentile estimator with 200 simulations was verified against 1000 simulations in one model and one location to validate the choice.

To evaluate the performance of the downscaling model and the projected performance of green roofs, different indicators were used:

- The lag-1 autocorrelation depending on timescale was evaluated. It was chosen to assess the temporal structure of the produced time series. The autocorrelation depending on lag time for timescale 6, 48 and 180 min were used for an in-depth analysis.

- The survival distribution of precipitation and discharge from both roofs were assessed at 6 min time steps. This approach is similar to the use of flow duration curves recently applied to green roofs by Johannessen et al. (2018). The exceedance probabilities were presented with a log axis to account for extreme probabilities. The median, 5th and 95th percentiles of the downscaled time series were represented. The survival distribution of discharge from the roofs with downscaled time series compared to the distribution based on observed data indicates the applicability of the downscaled time series as an input for green infrastructure modelling.

- Along with the survival distribution, a performance indicator derived from the Kolmogorov–Smirnov (KS) distance was used. The KS distance was indeed not relevant for the survival distributions where the extreme probabilities are of prime importance. The authors did not find a standard indicator for such cases in the literature; therefore the following indicator, which penalizes more errors for extreme probabilities, was developed:

$$KS_{rel} = \max \left( \frac{Distrib_{Sim,median} - Distrib_{Obs}}{Distrib_{Obs}} \right). \quad (6)$$

- Three different discharge thresholds were used to report exceedance frequency on different operating modes: $1\,L\,s^{-1}\,ha^{-1}$ for small events, $10\,L\,s^{-1}\,ha^{-1}$ for major events and $100\,L\,s^{-1}\,ha^{-1}$ for extreme events. Those thresholds were chosen in common for all roofs to facilitate comparison. They represent a compromise to have the same operating modes for each location even if the occurrence of those modes differs due to different climate conditions. Small event durations were counted in days per year, major events in hours per year and extreme events in minutes per year.

- The distribution of dry periods and the retention fraction were computed. They are not expected to be affected by the downscaling process since the dry periods affecting the roofs can be observed on daily resolution, and the retention fraction can be estimated with conceptual models using daily time-step data. However, they provide additional information to analyse the behaviour of the roofs.

## 2.5 Hybrid event-based downscaling

In order to assess the applicability of downscaled time series to predict the future performance of green infrastructure, the methods were compared to the current recommended practice in the locations: the use of an event-based design method based on IDF curves with a climate factor (CF) (Kristvik et al., 2019). In particular, the variational method (Alfieri et al., 2008) is applied. It consists, given a return period, in considering the constant-intensity rainfall leading to the highest discharge. It should be noted that the comparison intended to follow the recommended design method and not to follow the guidelines of a specific city since they can differ in terms of regulation. For instance, in Trondheim a threshold for maximum discharge has to be fulfilled (Trondheim Kommune, 2015), while in Lyon the first 15 mm of a 20-year return period has to be retained, and beyond those 15 mm a threshold is set for maximum discharge from the parcel (Greater Lyon council, 2020). The longest available time series, which originated from Trondheim, was the most adequate for this example. For 2-, 5- and 10-year return period rainfall and runoff events, three approaches were compared: (i) peak runoff of runoff events based on an observed precipitation time series (reference), (ii) peak runoff of rainfall events based on variational method, the IDF curves and with and without climate factor (typical design approach), and (iii) a hybrid approach based on downscaling $10^5$ rainfall events with a daily depth based on the return period curves with and without climate factors. This last approach used the $MRC_{SIT\text{-}SEP}$ model, the initial water content was set to the most probable value based on analysis of a long time series. According to the current recommendation in Norway for Trondheim municipality, a climate factor of 1.4 was applied (Dyrrdal and Førland, 2019).

## 3 Result and discussion

### 3.1 Green infrastructure model

The parameterized empirical reservoir model was applied to the extensive green roof and the detention-based extensive green roof. The performance was evaluated both on the time series and individual events extracted from the time series. The criteria were (i) Nash–Sutcliffe efficiency (NSE) indicator on time series for both discharge and water content, (ii) NSE for rainfall events defined with a minimum inter

events time of 6 h to analyse further the behaviour of the model, and (iii) the volumetric error on the time series to account for model retention evaluation. The water content was estimated directly from discharge measurement using the empirical curve. The performance was as follows:

- NSE > 0.8 for both discharge and water content for the extensive green roof. For the three most intense events the NSE ranged from 0.9 to 0.75. The water balance error was found to be 2.1 %.

- NSE > 0.94 for both discharge and water content for the detention-based extensive roof. For the three most intense events the NSE ranged from 0.96 to 0.85. The water balance error was found to be 5 %.

The conceptual limitation of the model can be seen in Fig. 2 at the beginning of the events of the testing period. It suggests that short events with low intensity are not reproduced well by the model as it cannot represent the delay induced by the wetting of the different layers of the roofs. Since the objectives of this study involve the use of a simple model to reproduce the behaviour of two roofs, the model was not further improved.

## 3.2 Analysis of climates properties

Figure 3 presents the zero-weight proportion depending on timescale, depth and temperature for two different datasets (Bodø and Hamar). In Fig. 3a the proportion of zero-weight decreases with increasing timescale for Bodø. In Hamar the proportion decreases until 45 min and increases for higher timescales. Based on this observation, two types of datasets were identified in terms of zero-weight occurrence. For data from Bodø, Bergen, Kristiansund and Trondheim, the proportion of weights that equalled zero decreased with increasing timescale. For the data from Hamar, Lyon, Marseille and Kristiansand, the proportion decreased until 45 min timescale and increased afterwards (Fig. B1a). Given a timescale, the proportion of weight equal to zero was not uniform depending on the weights (e.g. Bodø and Hamar Fig. 3b with a timescale of 48 min). Therefore, the monotony or non-monotony of the proportion of weights equalling to zero depending on timescale can be explained by different distributions of depth in the observed data. The proportion depended on depth, which is consistent with previous work (Rupp et al., 2009). It should be noted that a high proportion of zero-weight is linked to shorter and more intense rainfall events. It could explain why the proportion is higher in Lyon than in Bergen (cf. Appendix B, Fig. B1a).

In Fig. 3b, the zero-weights proportion decreases with increasing depth for the case of Bodø. In the case of Hamar, it increases for depth higher than 2 mm. Figure 3c and d show that a temperature dependency may explain this behaviour. In Bodø, the proportion depending on depth gives similar results for different ranges of temperature at 48 min

resolution (Fig. 3c). On the contrary, in Hamar, the subsets with lower temperature lead to a lower proportion of weights being equal to zero, compared to subsets with higher temperature (Fig. 3d). Moreover, the higher depths were observed in subsets with higher temperature. The increase observed in Hamar can be explained by the distribution of observed values. It is consistent with the observation of different temporal distributions of rainfall for different temperature ranges such as convective rains (Berg et al., 2013; Zhang et al., 2013). If, given a depth of 10 mm at resolution of 48 min, the probability to have a weight equal to zero is higher, then there is a higher probability to have an intense rainfall. The non-homogeneity of observed datasets and the shift in temperature with climate change might lead to inconsistency in time series produced by the downscaling methods that exclude depth and/or temperature dependency. The 48 min timescale was chosen to exemplify these properties. The same properties can be observed for different timescales, but the magnitude differs and tends to lower with higher timescale (Fig. B1b, c and d). Developing a model without temperature dependency might prevent comparability of parameters between locations and does not necessarily lead to parameter parsimonious models. Moreover, a model such as MRC$_{SI}$ can result in overfitting when used with datasets like Hamar. The functions necessary to represent the behaviour without considering the temperature dependency are more complex and less explanatory. Based on this analysis it was possible to add the temperature dependency and conceptualize a more explanatory model (MRC$_{SIT}$, with Eqs. A3a–A3e) with more robust results for the influence of climate change. This underlies the hypothesis that the information about the correlation between rainfall and precipitation will be expressed in the same way through those variables in the future.

## 3.3 Evaluation of the downscaling methods

An overview of the performance of the downscaling and green roof models in Bergen is presented on Fig. 4. All the downscaling models performed similarly in terms of dry-period distribution and slightly underestimated the dry periods in observed data (Fig. 4b). The dry periods were directly linked to the zero-weight probability. In green infrastructure modelling, the length of the dry periods influences the retention performance as it can lead to water stress hindering evapotranspiration. However, dry periods leading to water stress can be also evaluated with daily time-step series (there is no need for minute time-step series). Therefore, dry periods longer than the initial daily resolution are not significantly affected by downscaling.

The distribution of precipitation (Fig. 4a) was properly reproduced by MRC$_{SI}$, MRC$_{SI-SEP}$, MRC$_{SIT}$, and MRC$_{SIT-SEP}$ (KS$_{rel}$ = 1 in this case, indicating that the maximum distance has the same order of magnitude in data and model results) models, while MRC$_S$ and MRC$_{S-SEP}$ underestimated
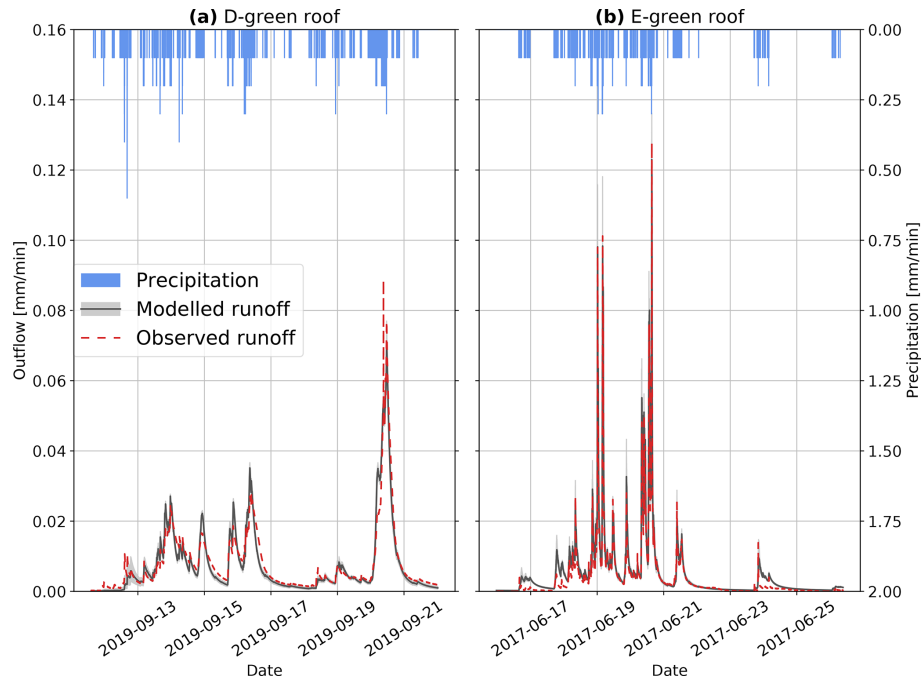
**Figure 2.** Testing of the green roof's reservoir model. Observed and modelled runoff of the detention-based extensive green roof (D) model on 10-day period **(a)** and extensive green roof (E) for a period of 8 d **(b)** in Trondheim.
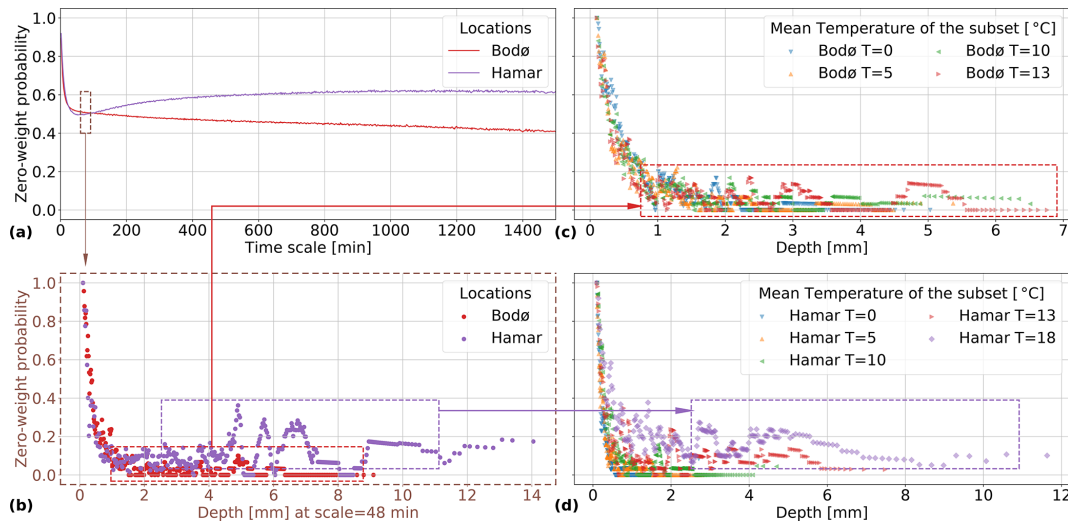


**Figure 3.** Dependency of the probability to have a weight equal to zero on timescale **(a)**, rainfall depth **(b)**, and temperature **(c, d)** for datasets observed in Bodø and Hamar. Panels **(b)**, **(c)**, and **(d)** are based on data at 48 min resolution.

low precipitation and overestimated high precipitation depths ($KS_{rel} = 10^2$ meaning that the maximum distance reached 2 orders of magnitude). This was expected as the time steps with high depth have higher probability to not be split in the observed data. This is not the case for $MRC_S$ and $MRC_{S\text{-}SEP}$ models, for which probability is uniformly distributed. In Bergen, the observed precipitations were within the range of 90 % coverage interval for $MRC_{SI}$, $MRC_{SI\text{-}SEP}$, $MRC_{SIT}$, and $MRC_{SIT\text{-}SEP}$. For the four later mentioned models, the

discharge of the D-Green roof was underestimated by 1 order of magnitude with a $KS_{rel}$ of $1.7 \times 10^1$ (Fig. 4c), due to the behaviour of the roof with rare high discharge. The hyetographs produced by downscaling probably tend to generate less favourable hyetographs for this roof. Although the discharge of the E-Green roof did not fall in the 90 % coverage interval, it can be considered as slightly underestimated since the magnitude is similar with a $KS_{rel}$ of 2.0 (Fig. 4d). However, this was not the case for all locations, as in Hamar

the most extreme precipitation tended to be underestimated, while the discharge from both roofs had the same order of magnitude as the observed data but tended to be overestimated. These findings could suggest inconsistency in the temporal structure of rainfall. This hypothesis can be confirmed by the autocorrelation (Fig. 5) being overestimated at 6 min time step.

The autocorrelation was underestimated by $MRC_S$ and $MRC_{S-SEP}$ models. The use of the rainfall continuity indicator increased the lag-1 autocorrelation for all models but did not improve the overall performances. The models $MRC_{SI}$, $MRC_{SI-SEP}$, $MRC_{SIT}$, and $MRC_{SIT-SEP}$ underestimated the lag-1 autocorrelation between 48 and 300 min timescales, but an in-depth analysis with different lags at 48 and 180 min timescale shows that despite that underestimation for lag-1 the general behaviour of the observed time series is reproduced. Similar observations were done for other locations.

To evaluate the produced time series it is necessary to compare the discharge with observed time series to the discharge with downscaled time series. For most of the locations, the predicted range of precipitation or discharge deviated for lowest probabilities from the values obtained with observed time series: (i) when the precipitation range matched with the observed distribution, the discharge tended to be overestimated; (ii) when the precipitation was underestimated, the discharge with observed data tends to lie in the range obtained from downscaled time series. While the downscaled time series suffer from some limitation when compared to results obtained from the observed time series, the raw discharge time series might as well not be suitable for robust decision making in green infrastructure implementation as it does not represent the natural variation of performance of green infrastructure.

In order to evaluate the potential of discharge from downscaled time series to approach the range of performance linked to natural variability, a 3-year moving window was used on precipitation and discharge time series resulting from observed precipitation. The resulting 5th and 95th percentiles of the annual duration exceeding 1, 10 and $100\,L\,s^{-1}\,ha^{-1}$ are presented in Fig. 6 to evaluate the time series in different operating modes of the roofs. It is compared to the stochastic variability (5th and 95th percentiles) from the six models. Each horizontal line in Fig. 6 represents the range between 5th and 95th percentiles for the threshold and model considered. The different thresholds represent the discharge for respectively small events, for major events and extreme events. In Fig. 4, the thresholds correspond to 0.006, 0.06 and $0.6\,mm\,min^{-1}$. A good estimate is defined by a complete or partial overlap between the observed natural variability and the stochastic variability range; the order of magnitude of the estimates should be similar. For instance, in Bergen (first column), the observed range of the E-Green roof higher than $100\,L\,s^{-1}\,ha^{-1}$ (third row) is predicted, based on observed input, from 4 to 10 min; the $MRC_S$ model provided values around 200 min; it is not a good es-

timate as there is no overlap and the order of magnitude varies; the $MRC_{SI}$ model resulted in a range from 10 to 20 min. It is a good estimate as the ranges are overlapping, and the orders of magnitude are similar. The $MRC_S$ and $MRC_{S-SEP}$ models tend to underestimate the order of magnitude of the range of exceedance frequencies of the small events ($1\,L\,s^{-1}\,ha^{-1}$) (in Bergen, Hamar and Marseille) but tend to overestimate major ($10\,L\,s^{-1}\,ha^{-1}$) (Hamar) and extreme events ($100\,L\,s^{-1}\,ha^{-1}$) (Bergen, Bodø, Hamar and Marseille). The other models gave mostly good estimates for each of the thresholds (Figs. 6, C1). In Marseille, the models $MRC_{SI}$, $MRC_{SI-SEP}$, $MRC_{SIT}$, and $MRC_{SIT-SEP}$ tended to underestimate the higher bound of the extreme event precipitation with values lower than $50\,min\,yr^{-1}$, whereas the observed time series led to a maximum of $90\,min\,yr^{-1}$. However, those models kept the order of magnitude, while $MRC_S$ and $MRC_{S-SEP}$ models estimated it higher than $10^2$ min. The same behaviour was observed with Hamar (Fig. 6) and Lyon datasets (Appendix C, Fig. C1). This suggests that the models performed worse for dryer locations, possibly due to the calibration procedure since fewer wet days are available for calibration. $MRC_{SI}$ and $MRC_{SIT}$ performed similarly, but due to its structure, $MRC_{SI}$ may overfit to the calibration data. It could result in an inaccurate prediction in the case of significant temperature shift between the calibration and prediction datasets. To conclude, $MRC_S$ and $MRC_{S-SEP}$ lead to overestimation of the natural variability, while $MRC_{SI}$, $MRC_{SI-SEP}$, $MRC_{SIT}$, and $MRC_{SIT-SEP}$ give more accurate estimates.

## 3.4 Assessment of green roof future performance

All six models were used to assess the performance of green roofs for future climate as illustrated for Bergen in Fig. 7. It was nevertheless acknowledged that $MRC_S$ and $MRC_{S-SEP}$ models gave less accurate estimates. The four model $MRC_{SI}$, $MRC_{SI-SEP}$, $MRC_{SIT}$, and $MRC_{SIT-SEP}$ lead to similar results in Bergen (Fig. 7). The difference in estimates between the models with coherence indicator ($MRC_{SI-SEP}$, $MRC_{SIT-SEP}$) and without ($MRC_{SI-SEP}$, $MRC_{SIT-SEP}$) was negligible in comparison to the stochastic uncertainty inherent to the models and the variability linked to the different projections available under RCP 8.5 (Fig. 7). In Bergen, according to the projections, the performance of the two solutions is likely to lead to worse performance: under the current climate, the $100\,L\,s^{-1}\,ha^{-1}$ exceedance was lower than 1 min for the D-Green roof; according to the $MRC_{SIT-SEP}$ model it might reach between 5 and 19 min in future climate. It suggested a shift in the order or magnitude from $10^0$ to more than $10^1$ min. Similarly, the E-Green roof might have a $100\,L\,s^{-1}\,ha^{-1}$ exceedance shift from $10^1$ to $10^2$ min. It means that the threshold would regularly be reached. As illustrated by Figs. 8 and C2, the performance shift depends highly on the location. While the $100\,L\,s^{-1}\,ha^{-1}$ exceedance of the green roofs was likely to get worse in Bergen, it was found to stay stable despite a small increase in Bodø

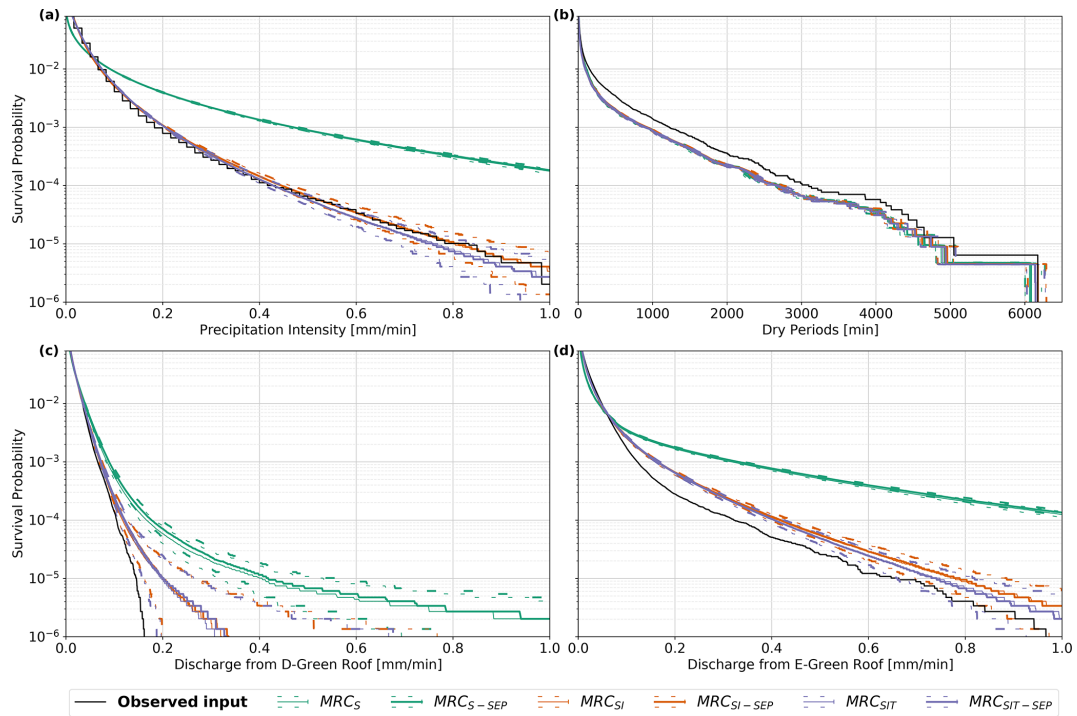**Figure 4.** Model performance with data from Bergen current climate for $MRC_S$, $MRC_{S\text{-}SEP}$, $MRC_{SI}$, $MRC_{SI\text{-}SEP}$, $MRC_{SIT}$, and $MRC_{SIT\text{-}SEP}$ with a range from 5th to 95th percentiles. Observed input represents the fine-resolution observed time series or simulation using this time series as input.
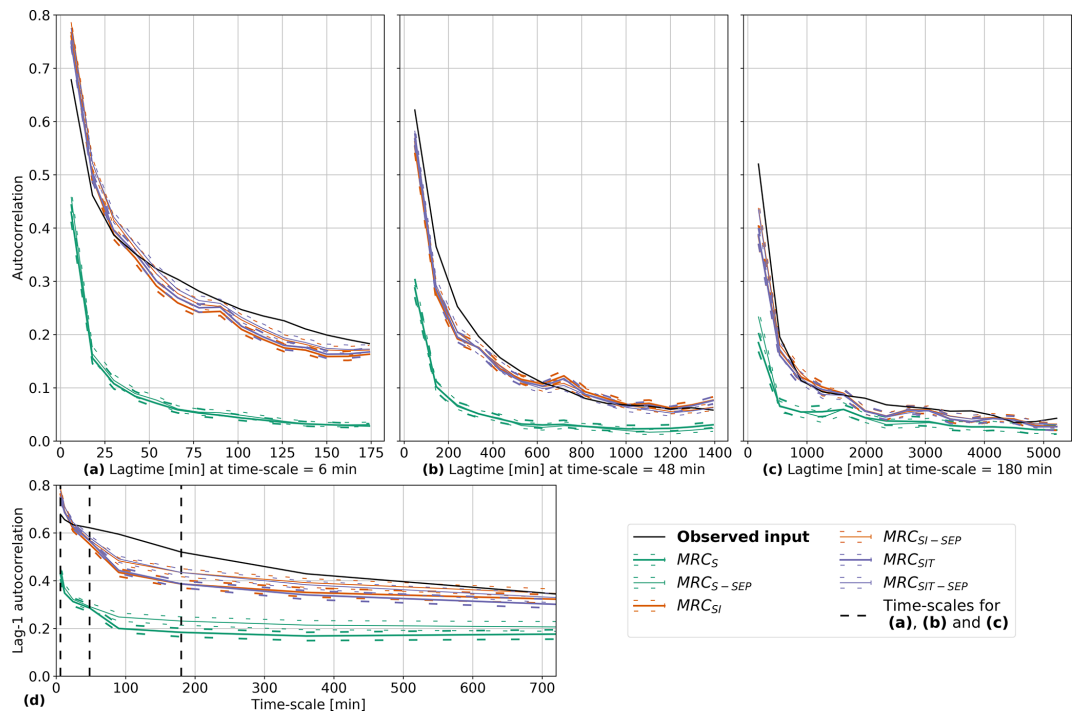


**Figure 5.** Autocorrelation with data from Bergen current climate for $MRC_S$, $MRC_{S\text{-}SEP}$, $MRC_{SI}$, $MRC_{SI\text{-}SEP}$, $MRC_{SIT}$, and $MRC_{SIT\text{-}SEP}$ with a range from 5th to 95th percentiles. Autocorrelation with different lags for 6, 48, and 180 min timescales, and lag-1 autocorrelation depending on timescale. They are compared to the observed input which represents the fine-resolution observed time series.
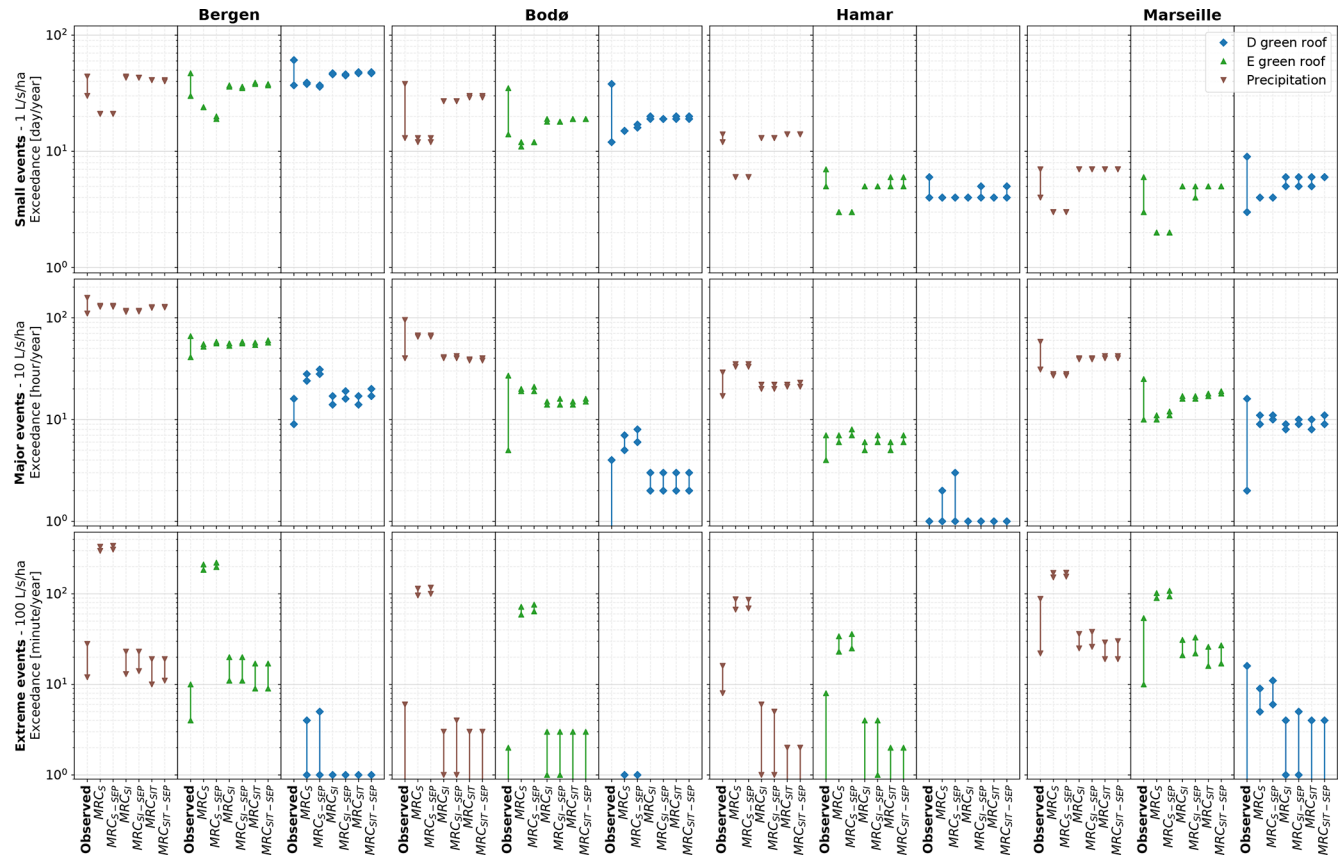
**Figure 6.** Performance of the downscaled time series in Bergen, Bodø, Hamar, and Marseille; exceedance frequency for small events, major events, and extreme events. The stochastic variability linked to the downscaled time series is evaluated with the 5th to 95th percentiles. *Observed* represents the fine-resolution observed time series or simulation using this time series as input. The 5th to 95th percentiles were estimated with a 3-year moving window. Due to log axis, occurrences lower than $10^0$ are not visible.

and to improve in Hamar and Marseille. The increase of exceedance frequency in the Norwegian cities was due to an increase in precipitation (Table 1). However, the increase in temperature led to an increase in potential evapotranspiration and therefore might have attenuated or even counterbalanced the effect of rainfall increase by lowering the initial water content in the roofs at the beginning of a rainfall event. The Table 3 shows that the retention fraction was likely to decrease in Bergen, Bodø, Hamar, Kristiansand and Kristiansund. It was found to increase more significantly in Lyon, Marseille and slightly in Trondheim. The models with temperature dependency performed similarly to the model with only depth dependency in most of the locations. However, in Lyon and Marseille, the $100\,\mathrm{L\,s^{-1}\,ha^{-1}}$ exceedance or precipitation predicted differed from 16–27 to 21–50 min (14–30 to 14–43 in Marseille). This suggests that some locations are more sensitive than others to temperature-dependent patterns. The models $MRC_{SI}$, $MRC_{SI\text{-}SEP}$, $MRC_{SIT}$, and $MRC_{SIT\text{-}SEP}$ allow the evaluation of shift in performance for the different roofs using exceedance range.

## 3.5 Design perspectives

The potential of downscaling models to improve the current practices was investigated. Figure 9 presents results based on continuous simulation, on the variational method and on the hybrid approach with downscaled events. It shows that the variational method underestimated the peak runoff with observed data, and the distribution from the hybrid approach covered them. It suggests that the variational method might not be enough conservative when compared to peak runoff from runoff events instead of rainfall events. Even if the results from the hybrid event-based downscaling lead to realistic distribution based on probable rainfall events, the downscaling models might need a different calibration or conceptualization to be optimized specifically for extreme events. Moreover the initial water content for the events remains a limitation of this method. The observed peaks show a range of possible outcomes, which highlights the limitations of the variational method with a single estimate, whereas the hybrid downscaling-event based method, leading to a range of probable outcomes, gave promising results that can lead to more robust design and decision making. Due to its characteris-
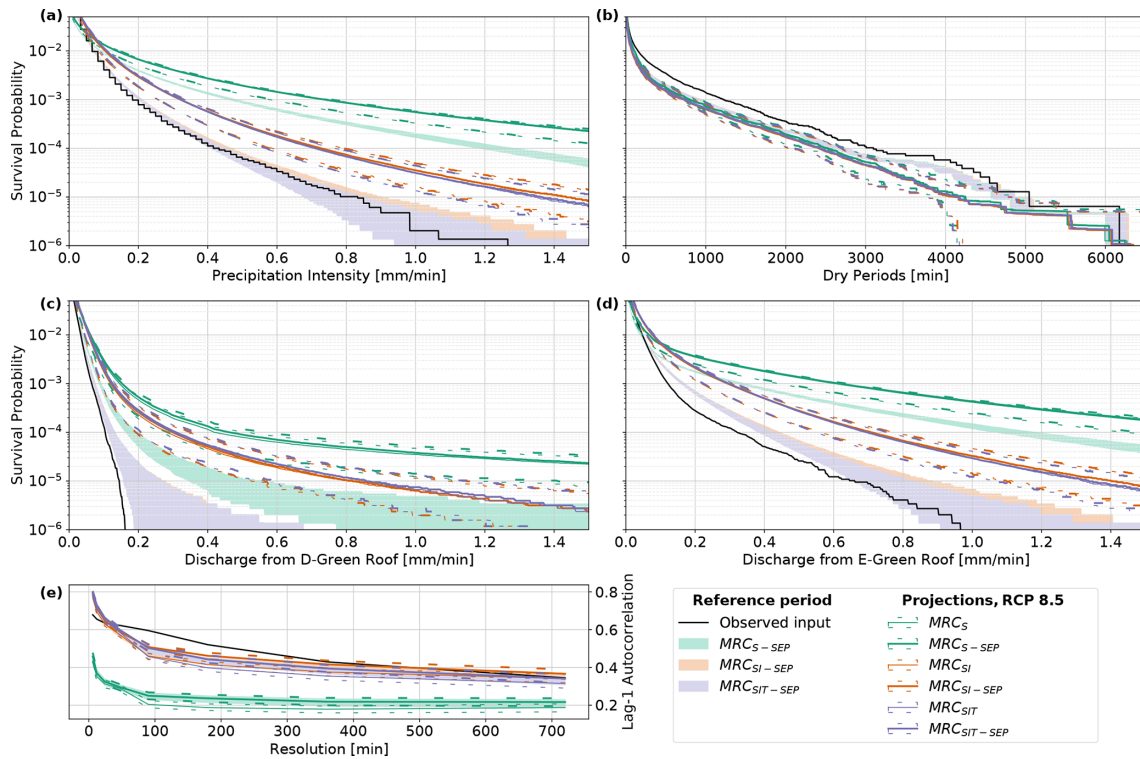
**Figure 7.** Comparison between performance under current climate and future climate in Bergen for the MRC$_S$, MRC$_{S\text{-SEP}}$, MRC$_{SI}$, MRC$_{SI\text{-SEP}}$, MRC$_{SIT}$, and MRC$_{SIT\text{-SEP}}$ with a range from the 5th to 95th percentiles. They are compared to observed input which represents the fine-resolution observed time series or simulation using this time series as input.

**Table 3.** Retention fraction in the different locations defined as the sum of outflow divided by the sum of precipitation.

| Location | Bergen | | Bodø | | Lyon | | Hamar | |
|---|---|---|---|---|---|---|---|---|
| Period | Observed | Projected | Observed | Projected | Observed | Projected | Observed | Projected |
| D-Green roof | 0.20 | 0.17 | 0.21 | 0.20 | 0.43 | 0.47 | 0.47 | 0.40 |
| E-Green roof | 0.19 | 0.16 | 0.21 | 0.20 | 0.39 | 0.44 | 0.44 | 0.38 |

| Location | Kristiansand | | Kristiansund | | Trondheim | | Marseille | |
|---|---|---|---|---|---|---|---|---|
| Period | Observed | Projected | Observed | Projected | Observed | Projected | Observed | Projected |
| D-Green roof | 0.24 | 0.22 | 0.25 | 0.20 | 0.27 | 0.30 | 0.41 | 0.47 |
| E-Green roof | 0.22 | 0.20 | 0.24 | 0.20 | 0.26 | 0.29 | 0.36 | 0.42 |

tics, the shift in performance between current climate and future climate is higher for the E-Green roof than for the D-Green roof. This is due to the detention layer in the D-Green roof which is not saturated by a 10-year return period event (Hamouz et al., 2020).

## 4 Conclusions

In this study, multiplicative random cascades models with different variable dependency were developed. They were based on a study of timescale, depth, and temperature characteristics of the datasets to ensure a consistent structure in the

view to apply them to daily-resolution climate projections. The applicability of the synthetic time series to be used as input for performance modelling of green infrastructure was evaluated. They were used to predict the shift in runoff exceedance under a future climate.

Six downscaling models were developed: two models with only timescale dependency (MRC$_S$ and MRC$_{S\text{-SEP}}$), two models with timescale and depth dependency (MRC$_{SI}$ and MRC$_{SI\text{-SEP}}$), and two models with timescale, depth and temperature dependency (MRC$_{SIT}$ and MRC$_{SIT\text{-SEP}}$). The models MRC$_{S\text{-SEP}}$, MRC$_{SI\text{-SEP}}$ and MRC$_{SIT\text{-SEP}}$ include a rainfall continuity property with the intention of improving the
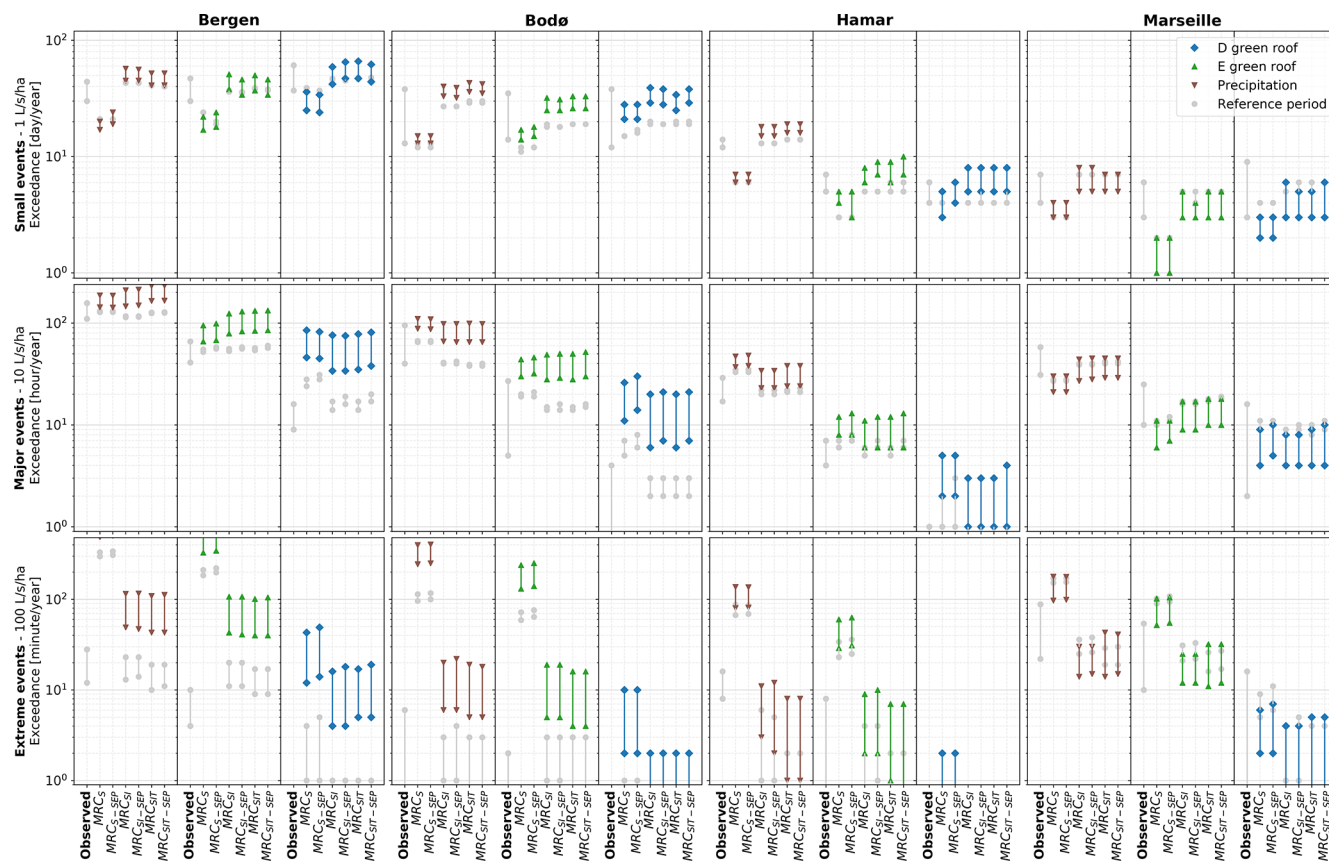
**Figure 8.** Future performance of green roofs (D and E) in Bergen, Bodø, Hamar and Marseille; exceedance frequency for small events, major events and extreme events. The stochastic variability linked to the downscaled time series is evaluated with the 5th to 95th percentiles. *Observed* represents the fine-resolution observed time series or simulation using this time series as input. The 5th to 95th percentiles were estimated with a 3-year moving window. Due to log axis, occurrences lower than $10^0$ are not visible.

temporal structure of the rainfall. The parametrization of the models ensures the continuity of the different properties modelled and a low number of parameters.

The $MRC_S$ and $MRC_{S-SEP}$ were not sufficient to predict the future performance of green infrastructure as they lead to an overestimation of runoff; the $MRC_{SI}$, $MRC_{SI-SEP}$, $MRC_{SIT}$, and $MRC_{SIT-SEP}$ lead to better performance: it was possible to predict runoff exceedance frequency with similar order of magnitude as the estimate of the natural variability of performance based on observed time series. The structure of the $MRC_{SI}$ and $MRC_{SI-SEP}$ models makes them more vulnerable to overfitting than $MRC_{SIT}$ and $MRC_{SIT-SEP}$, which makes them less reliable for future performance estimates. However, the differences between them were negligible compared to the variability linked to the different outcome of climate models, the variability inherent to the model and its accuracy. The $MRC_{S-SEP}$, $MRC_{SI-SEP}$ and $MRC_{SIT-SEP}$ add an equation to improve the temporal structure of downscaled rainfall. The models predicted higher runoff from the detention-based extensive green roof, which is consistent

with their properties; however the change in performance was not significant compared to stochastic uncertainty.

Using the RCP 8.5, the different downscaling and the green roof models suggest that the performance shift due to climate change highly depends on the location. The runoff exceedance is likely to increase in Bergen while slightly decreasing in Lyon and Marseille and keeping the same order of magnitude in the other locations. The results were compared to one of the current practices: the use of the variational method with a climate factor. It highlighted the limitation of this practice that provides a singular estimate and underestimates the observed peaks. A hybrid method using downscaling on extreme events led to promising results by estimating a distribution of performance of peak runoff.

The models performed well in the eight locations and four different climates. The use of a more advanced calibration procedure with Bayesian methods should improve the results. Similarly, a sensitivity analysis could improve the parametrization, especially for the models with depth and temperature dependency in order to fix non-behavioural parameters. The current study does not include irrigation, and
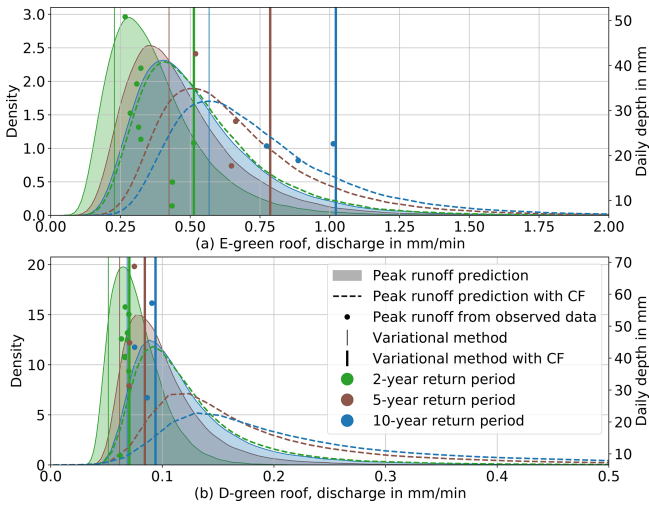
**Figure 9.** Performance depending on the return period in Trondheim for the extensive green roof (**a**) and the detention-based extensive green roof (**b**). The transparent coloured area is the distribution based on the hybrid event-based downscaling under current climate; the dotted line denotes the use of CF; the points represent the peaks runoff of runoff events from observed precipitation; the vertical lines the results found based on the variational method (VM): 2-, 5- and 10-year return periods are displayed.

snow modelling a study centred on green infrastructure modelling is therefore needed to extend the results. In order to be applied in practice on event-based simulation for design perspectives, the downscaling models need to be improved with a calibration procedure developed for extreme events and not on the complete spectrum of observation as in the current study.

## Appendix A: Generators description

### A1 Zero-weight generator with only timescale dependency

$$
\begin{aligned}
\text{ZeroGen}_S(S_{\text{time}}) = {} & a_{14} \times \log(S_{\text{time}})^4 a_{13} \times \log(S_{\text{time}})^3 \\
& + a_{12} \times \log(S_{\text{time}})^2 \\
& + a_{11} \times \log(S_{\text{time}}) + a_{10}
\end{aligned}
\tag{A1}
$$

### A2 Zero-weight generator with both timescale and depth dependency

$$
\text{ZeroGen}_{SI}(d_{i,2j}, S_{\text{time}} = 2j) =
$$

$$
\begin{cases}
\frac{1}{1+d_{i,2j}-P_3(S_{\text{time}})} f_0(S_{\text{time}}) \\
\quad + \left(1 - \frac{1}{1+d_{i,2j}-P_3(S_{\text{time}})} f_1(S_{\text{time}})\right), \\
\quad \text{if } d_{i,2j} > P_3(S_{\text{time}}) \\[2ex]
1, \quad \text{else}
\end{cases}
\tag{A2a}
$$

$$
f_0(S_{\text{time}}) = \frac{S_{\text{time}}^{a_{00}-1} \times (1 + S_{\text{time}})^{-a_{00}-a_{01}}}{a_{02}} + a_{03}
\tag{A2b}
$$

$$
\begin{aligned}
P_3(S_{\text{time}}) = {} & b_{13} \times S_{\text{time}}^3 + b_{12} \times S_{\text{time}}^2 \\
& + b_{11} \times S_{\text{time}} + b_{10}
\end{aligned}
\tag{A2c}
$$

$$
f_1(S_{\text{time}}) =
\begin{cases}
A \times \left(1 - \frac{1}{1+\exp(-\frac{4\times B}{A} \times (WC_i - C - \frac{A}{2\times B}))}\right), \\
\quad \text{if } S_{\text{time}} > C - \frac{A}{2\times B} \\[2ex]
B \times (S_{\text{time}} - C), \quad \text{else}
\end{cases}
\tag{A2d}
$$

### A3 Zero-weight generator with timescale, depth and temperature dependency

$$
\text{ZeroGen}_{SIT}(d_{i,2j}, T_{i,2j}, S_{\text{time}} = 2j) =
$$

$$
\frac{1}{1 + d_{i,2j}}^{\text{gauss}(T_{i,2j}, S_{\text{time}})}
\tag{A3a}
$$

$$
\text{gauss}(T_{i,2j}, S_{\text{time}}) =
$$

$$
A_0(S_{\text{time}}) \times \exp\left(\frac{(T_{i,2j} - \mu_T(S_{\text{time}}))^2}{2 \times \sigma_T(S_{\text{time}})^2}\right)
\tag{A3b}
$$

$$
\begin{aligned}
\mu_T(S_{\text{time}}) = {} & a_{14} \times S_{\text{time}}^4 a_{13} \times S_{\text{time}}^3 + a_{12} \times S_{\text{time}}^2 \\
& + a_{11} \times S_{\text{time}} + a_{10}
\end{aligned}
\tag{A3c}
$$

$$
A_0(S_{\text{time}}) = \frac{b_0}{(1 + S_{\text{time}})^{b_1}}
\tag{A3d}
$$

$$
\sigma_T(S_{\text{time}}) = \frac{c_0}{(c_2 + S_{\text{time}})^{c_1}}
\tag{A3e}
$$

### A4 Non-zero-weight generator

It consists in a truncated normal distribution described by Eq. (3b). The function $\sigma$ depends on timescale:

$$
\text{NonZeroGen}_S(S_{\text{time}}) = a_{10} \times (S_{\text{time}})^{\frac{1}{a_{11}}} + a_{12}.
\tag{A4}
$$

### A5 SEP generator

The stochastic element permutation follows a function generating the threshold to be compared to a uniformly generated random number depending on timescale:

$$
\begin{aligned}
\text{SEPGen}_S(S_{\text{time}}) = {} & a_{14} \times \log(S_{\text{time}})^4 a_{13} \times \log(S_{\text{time}})^3 \\
& + a_{12} \times \log(S_{\text{time}})^2 \\
& + a_{11} \times \log(S_{\text{time}}) + a_{10}.
\end{aligned}
\tag{A5}
$$
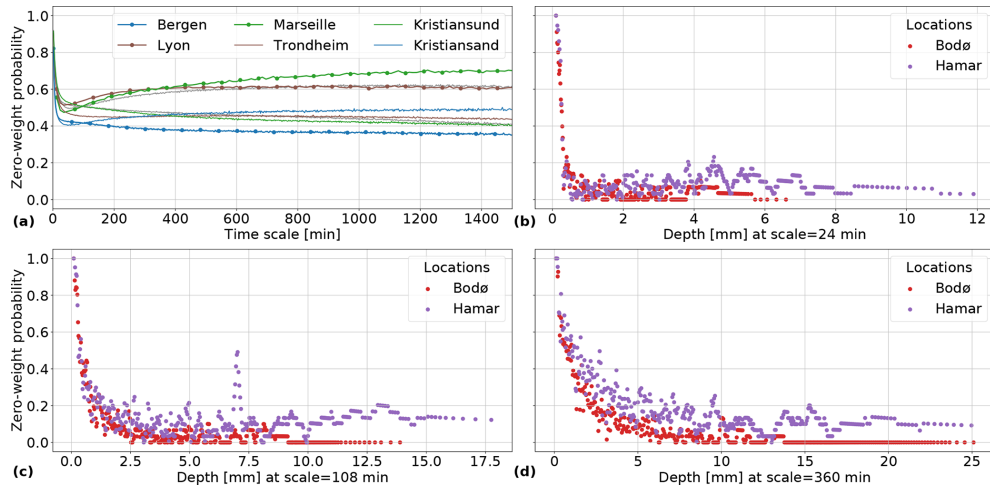
## Appendix B: Data analysis



**Figure B1.** Zero-weight probability depending on timescale for Bergen, Lyon, Marseille, Trondheim, Kristiansund and Kristiansand **(a)**. Zero-weight probability depending on the rainfall depth for different timescale: 24 min **(b)**, 108 min **(c)** and 360 min **(d)** for Bodø and Hamar.
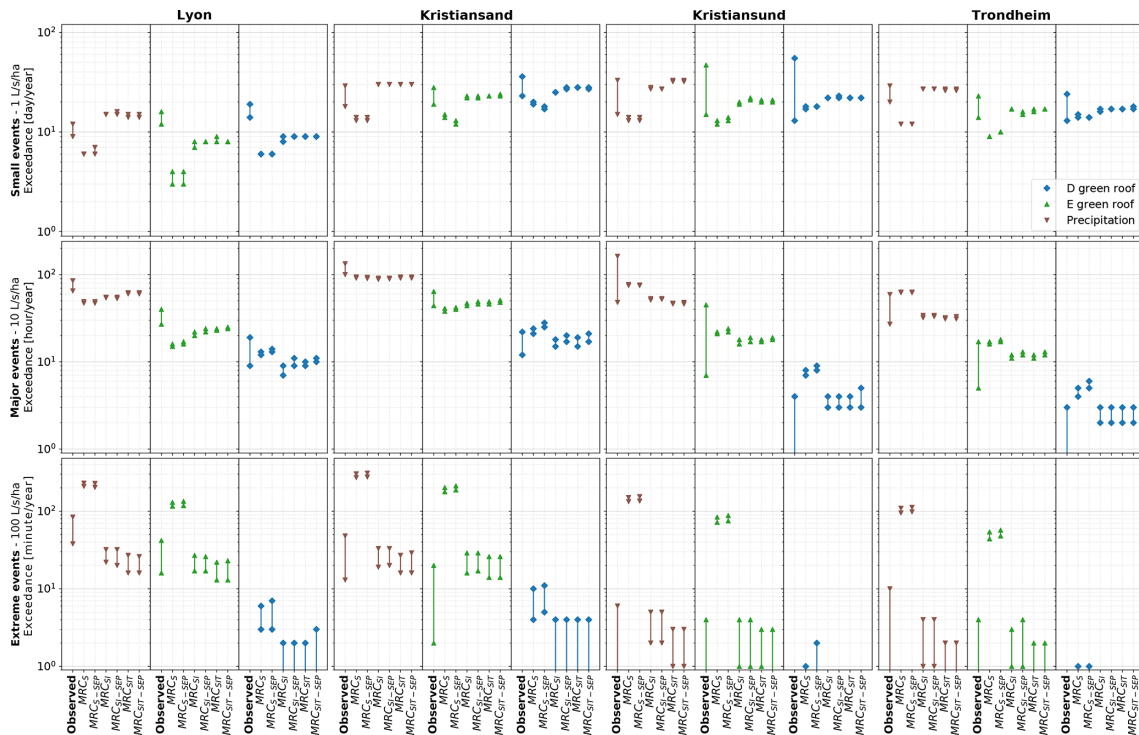
## Appendix C: Other locations performance



**Figure C1.** Performance of the downscaled time series in Lyon, Kristiansand, Kristiansund and Trondheim; exceedance frequency for small events, major events and extreme events. The stochastic variability linked to the downscaled time series is evaluated with the 5th to 95th percentiles. *Observed* represents the fine-resolution observed time series or simulation using this time series as input; the 5th to 95th percentiles were estimated with a 3-year moving window. Due to log axis, occurrences lower than $10^0$ are not visible.

**Figure C2.** Future performance of green roofs (D and E) in Lyon, Kristiansand, Kristiansund and Trondheim; exceedance frequency for small events, major events and extreme events. The stochastic variability linked to the downscaled time series is evaluated with the 5th to 95th percentiles. *Observed* represents the fine-resolution observed time series or simulation using this time series as input. The 5th to 95th percentiles were estimated with a 3-year moving window. Due to log axis, occurrences lower than $10^0$ are not visible.

## Appendix D: Table of abbreviations

| Abbreviation | Meaning |
|---|---|
| GI | Green infrastructures |
| MRC | Multiplicative random cascades |
| IDF curves | Intensity–duration–frequency curves |
| NVE | Norwegian Water Resources and Energy Directorate |
| MET | Norwegian Meteorological Institute |
| PET | Potential evapotranspiration |
| AET | Actual evapotranspiration |
| E-Green roof | Extensive green roof |
| D-Green roof | Detention-based extensive green roof |
| RCP 8.5 | Representative Concentration Pathway scenario with a 8.5 W m$^{-2}$ radiative forcing in 2100 |
| NSE | Nash–Sutcliffe efficiency |
| VM | Variational method |
| $S_{i,2j}$ | Rainfall continuity indicator at time step $i$ and timescale $2j$ |
| $d_{i,2j}$ | Depth [mm] at time step $i$ and timescale $2j$ |
| $w_{i,2j}$ | Minimum weight at time step $i$ and timescale $2j$ from aggregation of depths $D_{2i,j}$ and $D_{2i+1,j}$ |

| | |
|---|---|
| $S_W$ | Random discrete variable of neighbour of the highest weight |
| CF | Climate factor |

# References

Alfieri, L., Laio, F., and Claps, P.: A simulation experiment for optimal design hyetograph selection, Hydrol. Process., 22, 813–820, https://doi.org/10.1002/hyp.6646, 2008.

Benestad, R.: Downscaling Climate Information, in: Oxford Research Encyclopedia of Climate Science, Oxford University Press, https://doi.org/10.1093/acrefore/9780190228620.013.27, 2016.

Berg, P., Moseley, C., and Haerter, J. O.: Strong increase in convective precipitation in response to higher temperatures, Nat. Geosci., 6, 181–185, https://doi.org/10.1038/ngeo1731, 2013.

Bürger, G., Heistermann, M., and Bronstert, A.: Towards subdaily rainfall disaggregation via clausius-clapeyron, J. Hydrometeorol., 15, 1303–1311, https://doi.org/10.1175/JHM-D-13-0161.1, 2014.

Bürger, G., Pfister, A., and Bronstert, A.: Temperature-driven rise in extreme sub-hourly rainfall, J. Climate, 32, 7597–7609, https://doi.org/10.1175/JCLI-D-19-0136.1, 2019.

Dyrrdal, A. V. and Førland, E. J.: Klimapåslag for korttidsnedbør-Anbefalte verdier for Norge, NCCS report 5, https://klimaservicesenter.no/kss/rapporter/rapporter-og-publikasjoner_2 (last access: 26 May 2022), 2019.

Dyrrdal, A. V., Stordal, F., and Lussana, C.: Evaluation of summer precipitation from EURO-CORDEX fine-scale RCM simulations over Norway, Int. J. Climatol., 38, 1661–1677, https://doi.org/10.1002/joc.5287, 2018.

Gaur, A. and Lacasse, M.: Multisite multivariate disaggregation of climate parameters using multiplicative random cascades, Urban Climate, https://doi.org/10.1016/j.uclim.2018.08.010, 2018.

Gires, A., Tchiguirinskaia, I., and Schertzer, D.: Blunt extension of discrete universal multifractal cascades: development and application to downscaling, Hydrolog. Sci. J., 65, 1204–1220, https://doi.org/10.1080/02626667.2020.1736297, 2020.

Glasbey, C. A., Cooper, G., and McGechan, M. B.: Disaggregation of daily rainfall by conditional simulation from a point-process model, J. Hydrol., 165, 1–9, https://doi.org/10.1016/0022-1694(94)02598-6, 1995.

Greater Lyon council: Règlement du service public de l'assainissement collectif [Stormwater and wastewater service regulation], City guideline, Greater Lyon, https://www.grandlyon.com/services/gestion-des-eaux-de-pluie.html (last access: 10 November /2021), 2020.

Gupta, V. K. and Waymire, E. C.: A Statistical Analysis of Mesoscale Rainfall as a Random Cascade, J. Appl. Meteorol. Clim., 32, 251–267, https://doi.org/10.1175/1520-0450(1993)032<0251:ASAOMR>2.0.CO;2, 1993.

Hamouz, V. and Muthanna, T. M.: Hydrological modelling of green and grey roofs in cold climate with the SWMM model, J. Environ. Manage., 249, 109350, https://doi.org/10.1016/j.jenvman.2019.109350, 2019.

Hamouz, V., Pons, V., Sivertsen, E., Raspati, G. S., Bertrand-Krajewski, J.-L., and Muthanna, T. M.: Detention-based green roofs for stormwater management under extreme precipitation due to climate change, Blue-Green Systems, 2, 250–266, https://doi.org/10.2166/bgs.2020.101, 2020.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: Erratum to: EURO-CORDEX: new high-resolution climate change projections for European impact research, Reg. Environ. Change, 14, 579–581, https://doi.org/10.1007/s10113-014-0587-y, 2014.

Johannessen, B. G., Hanslin, H. M., and Muthanna, T. M.: Green roof performance potential in cold and wet regions, Ecol. Eng., 106, 436–447, https://doi.org/10.1016/j.ecoleng.2017.06.011, 2017.

Johannessen, B. G., Muthanna, T. M., and Braskerud, B. C.: Detention and Retention Behavior of Four Extensive Green Roofs in Three Nordic Climate Zones, Water, 10, 671, https://doi.org/10.3390/w10060671, 2018.

Kalra, A. and Ahmad, S.: Evaluating changes and estimating seasonal precipitation for the Colorado River Basin using a stochastic nonparametric disaggregation technique, Water Resour. Res., 47, W05555, https://doi.org/10.1029/2010WR009118, 2011.

Koutsoyiannis, D. and Langousis, A.: 2.02 – Precipitation, in: Treatise on Water Science, edited by: Wilderer, P., 27–77, Elsevier, Oxford, https://doi.org/10.1016/B978-0-444-53199-5.00027-0, 2011.

Kristvik, E., Johannessen, B. G., and Muthanna, T. M.: Temporal Downscaling of IDF Curves Applied to Future Performance of Local Stormwater Measures, Sustainability, 11, 1231, https://doi.org/10.3390/su11051231, 2019.

Laloy, E. and Vrugt, J. A.: High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing, Water Resour. Res., 48, W01526, https://doi.org/10.1029/2011WR010608, 2012.

Li, X., Meshgi, A., Wang, X., Zhang, J., Tay, S. H. X., Pijcke, G., Manocha, N., Ong, M., Nguyen, M. T., and Babovic, V.: Three resampling approaches based on method of fragments for daily-to-subdaily precipitation disaggregation, Int. J. Climatol., 38, e1119–e1138, https://doi.org/10.1002/joc.5438, 2018.

Lombardo, F., Volpi, E., and Koutsoyiannis, D.: Rainfall downscaling in time: theoretical and empirical comparison between multifractal and Hurst-Kolmogorov discrete random cascades, Hydrolog. Sci. J., 57, 1052–1066, https://doi.org/10.1080/02626667.2012.695872, 2012.

Lu, Y., Qin, X. S., and Mandapaka, P. V.: A combined weather generator and K-nearest-neighbour approach for assessing climate change impact on regional rainfall extremes, Int. J. Climatol., 35, 4493–4508, https://doi.org/10.1002/joc.4301, 2015.

Lutz, J., Grinde, L., and Dyrrdal, A. V.: Estimating rainfall design values for the City of Oslo, Norway-comparison of methods and quantification of uncertainty, Water (Switzerland), 12, 1735, https://doi.org/10.3390/W12061735, 2020.

McIntyre, N., Shi, M., and Onof, C.: Incorporating parameter dependencies into temporal downscaling of extreme rainfall using a random cascade approach, J. Hydrol., 542, 896–912, https://doi.org/10.1016/j.jhydrol.2016.09.057, 2016.

Müller-Thomy, H.: Temporal rainfall disaggregation using a micro-canonical cascade model: possibilities to improve the autocorrelation, Hydrol. Earth Syst. Sci., 24, 169–188, https://doi.org/10.5194/hess-24-169-2020, 2020.

Olsson, J.: Evaluation of a scaling cascade model for temporal rain- fall disaggregation, Hydrol. Earth Syst. Sci., 2, 19–30, https://doi.org/10.5194/hess-2-19-1998, 1998.

Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheater, H. S., and Isham, V.: Rainfall modelling using poisson-cluster processes: A review of developments, Stoch. Env. Res. Risk A., 14, 384–411, https://doi.org/10.1007/s004770000043, 2000.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, J. Hydrol., 303, 290–306, https://doi.org/10.1016/j.jhydrol.2004.08.026, 2005.

Paschalis, A., Molnar, P., and Burlando, P.: Temporal dependence structure in weights in a multiplicative cascade model for precipitation, Water Resour. Res., 48, W01501, https://doi.org/10.1029/2011WR010679, 2012.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci., 11, 1633–1644, https://doi.org/10.5194/hess-11-1633-2007, 2007.

Rupp, D. E., Keim, R. F., Ossiander, M., Brugnach, M., and Selker, J. S.: Time scale and intensity dependency in multiplicative cascades for temporal rainfall disaggregation, Water Resour. Res., 45, W07409, https://doi.org/10.1029/2008WR007321, 2009.

Rupp, D. E., Licznar, P., Adamowski, W., and Leśniewski, M.: Multiplicative cascade models for fine spatial downscaling of rainfall: parameterization with rain gauge data, Hydrol. Earth Syst. Sci., 16, 671–684, https://doi.org/10.5194/hess-16-671-2012, 2012.

Sanchez, C., Williams, K. D., and Collins, M.: Improved stochastic physics schemes for global weather and climate models, Q. J. Roy. Meteor. Soc., 142, 147–159, https://doi.org/10.1002/qj.2640, 2016.

Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, J. Geophys. Res.-Atmos., 92, 9693–9714, https://doi.org/10.1029/JD092iD08p09693, 1987.

Schilling, W.: Rainfall data for urban hydrology: what do we need?, Atmos. Res., 27, 5–21, https://doi.org/10.1016/0169-8095(91)90003-F, 1991.

Serinaldi, F.: Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models, Nonlin. Processes Geophys., 17, 697–714, https://doi.org/10.5194/npg-17-697-2010, 2010.

Stovin, V.: The potential of green roofs to manage Urban Stormwater, Water Environ. J., 24, 192–199, https://doi.org/10.1111/j.1747-6593.2009.00174.x, 2010.

Stovin, V., Poë, S., and Berretta, C.: A modelling study of long term green roof retention performance, J. Environ. Manage., 131, 206–215, https://doi.org/10.1016/j.jenvman.2013.09.026, 2013.

Stovin, V., Vesuviano, G., and De-Ville, S.: Defining green roof detention performance, Urban Water J., 14, 574–588, https://doi.org/10.1080/1573062X.2015.1049279, 2017.

Thober, S., Mai, J., Zink, M., and Samaniego, L.: Stochastic temporal disaggregation of monthly precipitation for regional gridded data sets, Water Resour. Res., 50, 8714–8735, https://doi.org/10.1002/2014WR015930, 2014.

Thorndahl, S. and Andersen, C. B.: CLIMACS: A method for stochastic generation of continuous climate projected point rainfall for urban drainage design, J. Hydrol., 602, 126776, https://doi.org/10.1016/j.jhydrol.2021.126776, 2021.

Trondheim Kommune: VA-norm. Vedlegg 5. Beregning av overvannsmengde Dimensjonering av ledning og fordrøyningsvolum [Water and Wastewater Norm. Attachment 5. Calculation of stormwater flows. Design of pipes and detention basins], City guideline, The City of Trondheim, https://www.va-norm.no/trondheim/ (last access: 10 November 2021), 2015.

Veneziano, D., Furcolo, P., and Iacobellis, V.: Imperfect scaling of time and space–time rainfall, J. Hydrol., 322, 105–119, https://doi.org/10.1016/j.jhydrol.2005.02.044, 2006.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson,

E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nat. Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Walker, W. E., Haasnoot, M., and Kwakkel, J. H.: Adapt or Perish: A Review of Planning Approaches for Adaptation under Deep Uncertainty, Sustainability, 5, 955–979, https://doi.org/10.3390/su5030955, 2013.

Zhang, Q., Li, J., Singh, V. P., and Xiao, M.: Spatio-temporal relations between temperature and precipitation regimes: Implications for temperature-induced changes in the hydrological cycle, Global Planet. Change, 111, 57–76, https://doi.org/10.1016/j.gloplacha.2013.08.012, 2013.