Comparison and Evaluation of Statistical Rainfall Disaggregation and High-Resolution Dynamical Downscaling over Complex Terrain

B. POSCHLOD,^a Ø. HODNEBROG,^b R. R. WOOD,^a K. ALTERSKJÆR,^b R. LUDWIG,^a
G. MYHRE,^b AND J. SILLMANN^b

^a *Department of Geography, Ludwig-Maximilians-Universität München, Munich, Germany*

^b *Center for International Climate and Environmental Research—Oslo, Oslo, Norway*

(Manuscript received 27 June 2018, in final form 30 October 2018)

ABSTRACT

Representative methods of statistical disaggregation and dynamical downscaling are compared in terms of their ability to disaggregate precipitation data into hourly resolution in an urban area with complex terrain. The nonparametric statistical Method of Fragments (MoF) uses hourly data from rain gauges to split the daily data at the location of interest into hourly fragments. The high-resolution, convection-permitting Weather Research and Forecasting (WRF) regional climate model is driven by reanalysis data. The MoF can reconstruct the variance, dry proportion, wet hours per month, number and length of wet spells per rainy day, timing of the maximum rainfall burst, and intensities of extreme precipitation with errors of less than 10%. However, the MoF cannot capture the spatial coherence and temporal interday connectivity of precipitation events due to the random elements involved in the algorithm. Otherwise, the statistical method is well suited for filling gaps in subdaily historical records. The WRF Model is able to reproduce dry proportion, lag-1 autocorrelation, wet hours per month, number and length of wet spells per rainy day, spatial correlation, and 6- and 12-h intensities of extreme precipitation with errors of 10% or less. The WRF approach tends to underestimate peak rainfall of 1- and 3-h aggregates but can be used where no observations are available or when areal precipitation data are needed.

1. Introduction

There is a broad range of applications for temporally high-resolution precipitation data, such as the design of urban water infrastructure (Bruni et al. 2015; Ochoa-Rodriguez et al. 2015), the simulation of runoff in catchments with short response time (Bennett et al. 2016; Reynolds et al. 2017), and the understanding and modeling of meteorological short-duration extreme events (Sillmann et al. 2017). In Norway, heavy precipitation events cause damage to the infrastructure by triggering floods, landslides, and avalanches (Dyrrdal et al. 2018; Heikkilä et al. 2011). In urban areas, such as Oslo, short intense rainfall can lead to urban flooding if the drainage system is not sufficient (Arnbjerg-Nielsen et al. 2013; Hanssen-Bauer et al. 2009).

The urban hydrological response system is very sensitive to small temporal scales of 1 h or less due to a typically high degree of soil sealing and therefore imperviousness of the ground (Bruni et al. 2015; Ochoa-Rodriguez et al. 2015). Hence, it is important that the temporal disaggregation methods can preserve the high-resolution rainfall intensities.

In general, the measurement stations of hourly precipitation data are distributed quite sparsely and unevenly, with the existing gauges often delivering only short time series including more or fewer data gaps. Precipitation data originating from climate models, both global climate models (GCMs) and conventional regional climate models (RCMs), lack the spatial resolution to sufficiently represent temporally high-resolution precipitation fields in complex terrain. Therefore, there is a strong need to fill the data gaps of hourly observations and to disaggregate coarser-resolution precipitation data from observations, reanalysis, and climate models.

In this paper, two different approaches are evaluated to address these problems. The nonparametrical Method

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-18-0132.s1>.

Corresponding author: Benjamin Poschlod, benjamin.poschlod@campus.lmu.de

FIG. 1. Location and elevation of the rain gauges in Oslo. Blue-marked stations are to be disaggregated, and red-marked stations contribute to the MoF database only [DEM in 10-m resolution by [Geonorge \(2018\)](#) and coastline by [EEA \(2018\)](#)]. The gray dashed lines in the overview map denote the outer 15-km and inner 3-km model domain of the WRF Model setup.

of Fragments (MoF; [Sharma and Srikanthan 2006](#); see also [Li et al. 2018](#); [Pui et al. 2012](#); [Westra et al. 2012](#)) uses statistical distributions from (nearby) high-resolution data to split temporally coarse-resolution data into fine-resolution fragments. The second approach features the high-resolution Weather Research and Forecasting (WRF) Model ([Skamarock and Klemp 2008](#)), driven by reanalysis data and set up in 3-km resolution.

The objectives of this paper are 1) to draw a comparison between a statistical precipitation disaggregation method and a high-resolution physically based climate model in terms of reproducing the characteristics of hourly precipitation and 2) to evaluate the ability of the MoF and the WRF Model to capture the intensity of precipitation extreme events.

This comparison is not meant to be competitive, as both techniques differ in their scope of application as well as the degree of their computational cost and data-intensive setup, but it highlights the advantages and limitations of each method in terms of providing high-resolution precipitation information in a complex-terrain urban environment.

2. Study area and data

Located north of the Oslofjord, the city of Oslo (59°55'N, 10°45'E) is characterized by a humid continental climate (Köppen–Geiger climate classification Dfb) with an annual mean temperature of 6°C ([Benestad 2011](#); [Kotteck et al. 2006](#)). The average

annual precipitation of 800 mm follows a seasonal cycle with more rainfall during summer ([Tjelta and Mamen 2014](#)). Despite having less annual precipitation than western Norway, the highest minute-to-hour rainfall intensities in Norway can be measured in the region around the Oslofjord ([Hanssen-Bauer et al. 2009](#)). This is due to southeastern Norway being exposed to mixed-type precipitation systems, such as isolated convective showers, stratiform frontal systems, and embedded convective cells within frontal systems ([Dyrrdal et al. 2016](#)).

The spatial distribution of the rain gauges (see [Fig. 1](#)) includes elevations from 12 to 200 m above sea level in a complex terrain with mountainous slopes and a coastline toward the Oslofjord. The gauges are tipping-bucket pluviometers with a resolution of 0.1 mm. Covering the period from 2000 to 2017, the data are provided by the Norwegian Meteorological Institute ([NMI 2018](#)) and have been aggregated from minute to hourly resolution. Over this 18-yr period, no significant trends were observed at the daily and subdaily time scale.

The blue-marked stations (see [Fig. 1](#), [Table 1](#)) are the 10 locations to be disaggregated from daily to hourly precipitation, whereas the red-marked stations only contribute to the MoF database because of limited record length and big data gaps. As the high precipitation intensities and convective extreme events predominantly occur during the summer season, this study focuses on the period from April to September.

TABLE 1. Data coverage and observed and WRF Model precipitation sum in summer for the rain gauges. Missing data have been flagged as missing or qualified as “slightly uncertain” and less certain. Days with missing data are not considered in the WRF Model precipitation sums either.

Station No.	Station name	Latitude (°N)	Longitude (°E)	Time period	Missing data April–September	Precipitation sum April–September	WRF Model precipitation
18920	Besserud	59.957	10.704	2001–17	27%	422 mm	401 mm
18701	Blindern	59.942	10.720	2000–17	6%	468 mm	438 mm
18420	Disen	59.946	10.790	2000–17	3%	517 mm	502 mm
18269	Haugenstua	59.954	10.905	2004–17	5%	528 mm	486 mm
18320	Hausmannsgt	59.931	10.755	2000–12	15%	393 mm	384 mm
18210	Hovin	59.923	10.804	2000–17	15%	433 mm	404 mm
18020	Lambertseter	59.878	10.786	2000–17	7%	434 mm	406 mm
18980	Lilleaker	59.920	10.634	2004–17	7%	465 mm	457 mm
17980	Ljabruveien	59.843	10.824	2000–17	11%	425 mm	387 mm
18270	Vestli	59.967	10.933	2000–17	11%	462 mm	462 mm

3. Methods

a. Method of Fragments

The MoF is a temporal disaggregation method that was developed to disaggregate streamflow data but was adapted to precipitation data (Sharma and Srikanthan 2006). It is applied using temporally coarser-resolution precipitation data of the location of interest (LOI) as well as fine-resolution precipitation data of rain gauges nearby. In this paper, every station of the 10 blue-marked stations (see Fig. 1) is disaggregated using the hourly precipitation data of the respective remaining blue- and red-marked stations in a leave-one-out cross-validation scheme. Subdaily data of the LOI are not used. The applied MoF algorithm can be described as follows:

Step 1: Obtain a daily rainfall time series $R_{\text{LOI}}(d)$ at the LOI, with d denoting the day.

Step 2: Group $R_{\text{LOI}}(d)$ into four classes, depending on the precipitation of $R_{\text{LOI}}(d - 1)$ and $R_{\text{LOI}}(d + 1)$:

$$\begin{aligned}
 \text{Class 1: } & R_{\text{LOI}}(d - 1) = 0 \wedge R_{\text{LOI}}(d + 1) = 0 \\
 \text{Class 2: } & R_{\text{LOI}}(d - 1) > 0 \wedge R_{\text{LOI}}(d + 1) = 0 \\
 \text{Class 3: } & R_{\text{LOI}}(d - 1) = 0 \wedge R_{\text{LOI}}(d + 1) > 0 \\
 \text{Class 4: } & R_{\text{LOI}}(d - 1) > 0 \wedge R_{\text{LOI}}(d + 1) > 0.
 \end{aligned} \quad (1)$$

The $R_{\text{LOI}}(d)$ with missing data for $R_{\text{LOI}}(d \pm 1)$ is assigned to every possible class.

Group the daily precipitation $R_S(d)$ at the nearby stations S into these classes as well. The splitting of $R_{\text{LOI}}(d)$ and $R_S(d)$ into these classes is done to account for the interday connectivity of rainfall events (Li et al. 2018). This restriction ensures that only days with the same previous-day and next-day wetness state are available for selection. This constraint has also been used by other studies featuring the MoF independently of the region (Li et al. 2018;

Pui et al. 2012; Sharma and Srikanthan 2006; Westra et al. 2012).

Step 3: To take into account the seasonality of precipitation, a moving window of only ± 15 days around the day of year (doy) of $R_{\text{LOI}}(d)$ is considered, with doym being the corresponding day of year for d . Create a look-up-table (LUT) for every $R_{\text{LOI}}(d)$ using $\{R_S(\text{doy} - w), R_S[\text{doy} - (w - 1)], \dots, R_S(\text{doy}), \dots, R_S[\text{doy} + (w - 1)], R_S(\text{doy} + w)\}$ of every year of historical records of all stations S with corresponding classes only, where w denotes the window size.

Step 4: Sort the entries of the LUT by the difference of daily precipitation $D = |R_{\text{LOI}}(d) - R_S(\text{doy})|$ and select the k entries with smallest D , with $k = \sqrt{n}$ and n = number of entries in the LUT. If $n < 10$, k is set to n .

Step 5: If there are entries with $D = 0$, choose one of these entries randomly. Else, assign the probability $P(k)$ (Lall and Sharma 1996) to every entry, where

$$P(k) = \frac{1}{D(k)} \cdot \frac{1}{\sum_{i=1}^k \frac{1}{D(i)}}, \quad (2)$$

and choose one of the k entries randomly according to $P(k)$.

Step 6: Obtain the subdaily precipitation for the chosen station and day. Form the fragment vector $\mathbf{F}(d) = [f_1(d), f_2(d), \dots, f_{23}(d), f_{24}(d)]$, with

$$f_h(d) = \frac{R(h)}{R(d)}, \quad (3)$$

as ratio of the hourly precipitation $R(h)$ and the daily precipitation sum $R(d)$. The subdaily precipitation $H_{\text{LOI}}(d)$ is calculated with

TABLE 2. Physical options of the WRF Model setup.

Physical option	Chosen setup	Source
Radiation	CAM radiation scheme	Collins et al. (2004)
Convection	Grell–Freitas ensemble scheme (only used in the 15-km domain)	Grell and Freitas (2014)
Shallow convection	University of Washington scheme	Park and Bretherton (2009)
Boundary layer	YSU PBL scheme	Hong et al. (2006)
Microphysics	WSM 6-class graupel scheme	Hong and Lim (2006)
Surface processes	Revised MM5 Monin–Obukhov scheme	Jimenez et al. (2012)
Land surface	Unified Noah land surface model	Tewari et al. (2004)
Vertical levels	50	
Spectral nudging	Spectral nudging of temperature, horizontal winds, and geopotential height in the outer (15 km × 15 km) domain	

$$H_{\text{LOI}}(d) = \mathbf{F}(d) \times R_{\text{LOI}}(d). \quad (4)$$

As the algorithm is nondeterministic because of the probabilistic selection in step 5, the results of the individual MoF runs differ. For computational time reasons, 100 runs are calculated. Furthermore, the calculated statistics always relate to the median (arithmetic mean of the upper and lower median) of the 100 runs.

b. WRF Model

The WRF Model was originally developed for regional simulation and forecasting of weather, but it can also be used for regional climate modeling driven by GCM simulations. Released in 2000, WRF is by now the most-used atmospheric model (Powers et al. 2017). Flato et al. (2013) declare in the IPCC Fifth Assessment Report with *high confidence* that dynamical downscaling via RCMs is able to add value to GCM simulations regarding mesoscale phenomena and extreme events, especially in regions with complex terrain. Tabari et al. (2016) claim that dynamical downscaling is expected to be a possible way to decrease the systematic biases and narrow the gap between coarse GCM outputs and the need for fine-resolution precipitation in hydrological and water engineering studies. The WRF Model has been successfully applied in Scandinavia by Heikkilä et al. (2011; 30 and 10 km) and Mayer et al. (2015; 8 km), proposing that a higher resolution would improve the performance.

In this study, the WRF-ARW 3.8.1 model is set up with a spatial resolution of 3 km with a model domain of 480 km (north–south) × 510 km (west–east; see Fig. 1) nested in a 15-km grid with a model domain of 2400 km (north–south) × 2550 km (west–east). It is driven by NCEP FNL reanalysis data (1° horizontal resolution; NCEP 2000) as initial and boundary conditions in 6-h resolution. The WRF simulation was divided into 18 one-year time slices, where each of the years were initialized on 1 December the previous year, allowing for

4 months of spinup. While the 15-km domain uses the convection scheme by Grell and Freitas (2014), convection is assumed to be sufficiently resolved at resolutions up to 4 km (Prein et al. 2015; Tabari et al. 2016). The detailed model setup is documented in Table 2 and has been optimized to minimize precipitation bias over Scandinavia. Therefore, several 1-yr test runs were carried out with different combinations of cumulus, microphysics, and radiation schemes. Results were compared against observations, and the setup that gave the lowest bias was used.

The comparison of results and observations uses the nearest grid point of the WRF Model data to the rain gauges. Despite the spatial resolution of the model setup being very high (3 km), the difference of the elevation of the model grid point to the digital elevation model (DEM) is governed by the mountainous terrain. The error reaches up to 56-m overestimation in Haugenstua and 45-m underestimation in Blindern (see Fig. S1 in the online supplemental material).

However, it has to be kept in mind that the data from rain gauges are point measurements. Due to the high spatial variability of precipitation events, it is unknown how representative these point measurements are for the surrounding grid cell area, especially in complex terrain (Cristiano et al. 2017).

4. Results and discussion

To compare observed and modeled precipitation, appropriate measures have to be used for validation, which are adapted to the spatiotemporal character of precipitation data (Koutsoyiannis 2003; Pui et al. 2012). As the MoF algorithm makes use of the daily precipitation sum at the LOI, the validation measures aim at subdaily time scales only. The focus is set on the following performances: 1) reproducing the standard validation statistics, 2) rebuilding the amount of wet hours per month as well as the number and length of intraday wet spells, 3) restoring the spatial correlation between each station, and 4) reconstructing the timing

TABLE 3. Standard validation statistics of observed and disaggregated precipitation for 1-, 3-, 6-, and 12-h time scales. The statistics corresponding to the MoF are calculated from the median of the respective 100 runs. A “dry” proportion is here defined as an average rainfall intensity below 0.1 mm h^{-1} , $0.2 \text{ mm (3 h)}^{-1}$, $0.4 \text{ mm (6 h)}^{-1}$, and $0.6 \text{ mm (12 h)}^{-1}$.

	Besserud	Blindern	Disen	Haugenstua	Hausmannsgt	Hovin	Lambertseter	Lilleaker	Ljabruveien	Vestli	Oslo (mean)	
Variance												
Obs	0.480	0.398	0.396	0.431	0.384	0.417	0.341	0.387	0.342	0.296	0.387	
WRF	0.378	0.297	0.336	0.322	0.274	0.293	0.269	0.319	0.278	0.339	0.310	1 h
MoF	0.510	0.388	0.402	0.440	0.323	0.417	0.321	0.378	0.339	0.332	0.385	
Obs	2.604	2.061	2.102	2.346	1.903	2.166	1.787	2.047	1.789	1.715	2.052	
WRF	2.056	1.582	1.793	1.722	1.422	1.546	1.395	1.637	1.426	1.824	1.640	3 h
MoF	2.723	2.017	2.118	2.348	1.747	2.156	1.723	2.021	1.776	1.769	2.040	
Obs	7.105	5.566	5.749	6.474	4.924	5.730	4.854	5.465	4.956	4.689	5.551	
WRF	5.821	4.401	5.032	4.748	3.851	4.265	3.864	4.494	3.975	4.990	4.544	6 h
MoF	7.426	5.396	5.734	6.504	4.709	5.797	4.667	5.527	4.890	4.865	5.551	
Obs	19.214	13.903	14.689	17.404	12.447	14.606	12.518	14.790	12.905	12.258	14.473	
WRF	15.850	11.499	13.450	12.303	10.152	11.182	9.676	12.084	10.031	13.180	11.941	12 h
MoF	19.282	13.892	14.698	16.974	12.192	14.785	12.297	14.413	12.641	12.487	14.366	
Lag-1 autocorrelation												
Obs	0.453	0.417	0.456	0.505	0.392	0.438	0.444	0.451	0.445	0.512	0.451	
WRF	0.457	0.447	0.441	0.442	0.419	0.442	0.435	0.459	0.428	0.439	0.441	1 h
MoF	0.449	0.429	0.424	0.453	0.440	0.411	0.430	0.452	0.428	0.450	0.436	
Obs	0.355	0.335	0.368	0.421	0.316	0.336	0.355	0.347	0.368	0.367	0.357	
WRF	0.407	0.409	0.410	0.397	0.379	0.407	0.399	0.397	0.404	0.382	0.399	3 h
MoF	0.341	0.307	0.324	0.340	0.327	0.318	0.316	0.345	0.331	0.331	0.328	
Obs	0.343	0.262	0.293	0.350	0.277	0.277	0.289	0.326	0.297	0.305	0.302	
WRF	0.322	0.308	0.326	0.323	0.310	0.324	0.300	0.316	0.304	0.307	0.314	6 h
MoF	0.275	0.242	0.259	0.269	0.267	0.257	0.264	0.278	0.249	0.248	0.261	
Obs	0.255	0.219	0.248	0.259	0.228	0.232	0.218	0.221	0.214	0.260	0.235	
WRF	0.225	0.237	0.237	0.263	0.218	0.246	0.246	0.233	0.242	0.241	0.239	12 h
MoF	0.216	0.188	0.209	0.219	0.212	0.200	0.205	0.214	0.195	0.197	0.195	
Dry proportion												
Obs	0.881	0.890	0.881	0.880	0.901	0.891	0.884	0.892	0.883	0.890	0.887	
WRF	0.880	0.893	0.886	0.885	0.895	0.892	0.899	0.886	0.897	0.886	0.890	1 h
MoF	0.885	0.891	0.884	0.883	0.895	0.891	0.892	0.892	0.889	0.892	0.890	
Obs	0.846	0.853	0.849	0.842	0.867	0.855	0.857	0.859	0.850	0.833	0.851	
WRF	0.831	0.847	0.838	0.836	0.848	0.845	0.853	0.838	0.853	0.839	0.843	3 h
MoF	0.849	0.854	0.847	0.847	0.861	0.855	0.858	0.856	0.853	0.855	0.853	
Obs	0.815	0.824	0.818	0.813	0.835	0.824	0.829	0.828	0.820	0.813	0.822	
WRF	0.797	0.814	0.804	0.801	0.817	0.813	0.818	0.805	0.821	0.803	0.809	6 h
MoF	0.818	0.823	0.817	0.814	0.835	0.823	0.831	0.825	0.823	0.824	0.823	
Obs	0.745	0.755	0.750	0.744	0.772	0.757	0.767	0.762	0.755	0.753	0.756	
WRF	0.725	0.747	0.737	0.733	0.751	0.744	0.750	0.739	0.753	0.735	0.741	12 h
MoF	0.748	0.755	0.748	0.748	0.770	0.757	0.768	0.761	0.755	0.757	0.757	

of maximum rainfall bursts within the day and intensities of extreme precipitation.

The rainfall totals of each location are reproduced well by the WRF Model with a slight underestimation of 5% on average (see Table 1). The highest deviations of 13% occur in August.

a. Standard validation statistics

The variance of hourly precipitation, the lag-1 autocorrelation, and the dry proportion are important statistical measures in order to provide a functional evaluation of performance for rainfall disaggregation (Socolofsky et al. 2001; see Table 3).

The MoF performs well at reproducing the monthly variance of hourly precipitation with an average mean absolute error (MAE) across all stations of 9.2% (Fig. 2a). Also, the range of occurring variances across all months is represented adequately. While being able to reproduce the variance in April, May, and September, the WRF Model underestimates the variances in June–August, leading to an average MAE of 24.5%. Since the calculation of the variance is sensible to extreme values, the underestimation by the WRF Model is mainly governed by the underestimation of convective high precipitation events, which is shown in section 4d.

FIG. 2. (a) Average monthly variance of hourly precipitation, (b) average of wet hours per month, (c) number, and (d) length of intraday wet spells per rainy day, aggregated for all stations in Oslo. The range of variation across all stations is presented as box plots. The boxes denote the interquartile range (IQR), which equals the six middle stations of 10 in total.

The lag-1 autocorrelation represents the dependence of the precipitation at two consecutive times. It is estimated by the autocorrelation coefficient L_1 ,

$$L_1 = \frac{\sum_{t=1}^{N-1} [R(t) - \bar{R}][R(t+1) - \bar{R}]}{\sum_{t=1}^N [R(t) - \bar{R}]^2} \quad (5)$$

with $R(t)$ as rainfall at time t of a series of length N with mean precipitation $\bar{R}$. Both methods perform well at reconstructing this statistical measure at the 1-h time scale with an average MAE across all stations of 5.8% (WRF) and 6.0% (MoF).

However, the performance decreases for 3-h aggregates with still acceptable MAEs of 13.2% (WRF) and 8.4% (MoF). At the aggregated time scales of 6 and 12 h the WRF Model outperforms the MoF with errors of 8.4% (WRF 6 h) and 7.6% (WRF 12 h) compared to 13.6% (MoF 6 h) and 12.7% (MoF 12 h). For all temporal aggregations the MoF generally underestimates the lag-1 autocorrelation. This systematic underestimation was also reported by Li et al. (2018) for the application in Singapore and China. This can be explained by the MoF being unable to preserve the interday connectivity of precipitation events. On average, 11.7 precipitation events per 6-month period occurring overnight from one day to another are

observed. These events are defined as raining in the last hour of the day and in the first hour of the following day. The MoF can reconstruct 5.1 of these events leading to a reduced autocorrelation coefficient L_1 , whereas the WRF Model can reproduce 11.5 events per 6-month period. For the aggregated time scales, this effect has a greater impact on the value of L_1 due to the higher differences of rainfall at consecutive times entering in Eq. (5).

To evaluate the reproduction of dry and wet times, the dry proportion is calculated. It is important for the further application of the disaggregated data that they show no bias having too many dry or wet times. “Dry” is defined as less than 0.1 mm h^{-1} mean rainfall intensity due to the minimal resolution of the rain gauges. This is necessary to provide consistency between the observational and WRF modeled datasets. For the aggregated periods of 3, 6, and 12 h the threshold value cannot be derived from scaling down the daily threshold for drizzle of 1 mm day^{-1} (Sun et al. 2006) linearly, as the daily threshold does not account for drizzle for every hour of the 24 h. Therefore, adapted thresholds of 0.2 mm (3 h), 0.4 mm (6 h), and 0.6 mm (12 h) are introduced. The MoF is able to reproduce these dry proportions with a MAE of less than 0.5% for all time scales. The MAEs of the WRF Model amount to 0.7% (1 h), 1.2% (3 h), 1.6% (6 h), and 1.9% (12 h). Hence, both models can preserve

the percentage of wet and dry periods at the respective time scales.

b. Wet hours per month, and number and length of intraday wet spells

The amount of wet hours per month is considered to assess the ability of both methods to capture the temporal and site-specific variety of wet and dry periods. The threshold value of 0.1 mm h^{-1} is again applied to distinguish between wet and dry. The MoF is capable of estimating the monthly sum of wet hours with an MAE of 4.9% (see Fig. 2b) and of representing the range of minima and maxima of all stations for each month. The WRF Model also performs satisfyingly with an error of 9.6%, but overestimates the amount of wet hours in April and underestimates it in August.

The number and length of intraday wet spells is of great importance for the soil moisture conditions and infiltration rates. Therefore, disaggregated precipitation data preserving these characteristics are crucial for hydrological modeling and subsequent estimation of design floods (Haberlandt et al. 2008). An intraday wet spell is defined by consecutive hours of precipitation $\geq 0.1 \text{ mm h}^{-1}$ within a day.

The number of intraday wet spells per rainy day (see Fig. 2c) is reconstructed adequately with an MAE across all stations and months of 4.5% (WRF) and 3.2% (MoF). Both methods underestimate the maxima in August and September, which are caused by just one station differing significantly during these two months (Vestli). The corresponding durations of the intraday wet spells vary across the whole summer season, with shorter, often convective events during June–August and longer spells in April, which are more frequently caused by stratiform frontal systems.

The MoF can reproduce the duration with a MAE of 3.4% (Fig. 2d), whereas the WRF Model tends to exaggerate the seasonal cycle, underestimating the duration in June–August and overestimating the duration in April and September (MAE: 7.2%).

c. Spatial correlation

The preservation of the spatial correlation of each rain gauge to its surrounding rain gauges is assessed by calculating the Kendall rank coefficient τ of all pairs of stations for every month, as this measure does not rely on any assumptions on the distribution or the linearity of correlation.

The values for τ for the station pairs of observed hourly precipitation range from 0.40 to 0.86 (see Fig. 3). The MoF fails to reconstruct this spatial correlation, underestimating it systematically with a range from 0.14 to 0.52. The WRF Model performs better at reproducing

FIG. 3. Intersite Kendall rank correlation between the hourly precipitation of all pairs of stations for every month.

the spatial correlation, but generally overestimates τ by 10% on average, ranging from 0.60 to 0.92.

This can be explained by the neighboring conditions of the grid cells leading to a too high similarity between the cells and therefore too high correlation. Additionally, the simulation of 9-km^2 areal average precipitation lowers the differences between the sites compared to the point observations (Gjegersen et al. 2013).

d. Timing of maximum rainfall bursts and exceedance probability of extreme rainfall intensities

The MoF is able to reconstruct the timing of maximum rainfall bursts within a day very well for the mean of all sites, as the daily cycle is quite similar at every location (Fig. S3). The MAE for the median of the 100 runs for every single station and month amounts to 9.0%. The MoF reproduces the monthly cycle with the interval of 12–17 h (see Fig. 4a) increasing from April to June and decreasing from June to September due to convective precipitation occurring more frequently in the afternoon. The WRF Model is capable of rebuilding the characteristics of the monthly cycle in the interval of 12–17 h, but generally underestimates the percentage of maximum rainfall bursts during 6–11 h while overestimating it during 18–23 h (MAE: 15.5%). To set the calculated MAEs into relation, they can be compared to a uniform distribution of timings for every site (a quarter of maximum bursts for each 6-h period), which would result in a MAE of 13.3%. Therefore, the MoF can clearly add value, whereas the WRF cannot reproduce the overall timing.

In the following section, extreme events with an empirical annual exceedance probability (AEP) < 1 to occur

FIG. 4. (a) Mean percentage of all sites for the timing of the maximum rainfall burst per day within the four 6-h periods TIME1–TIME4. The gray bars represent the observational data, and the green lines represent the WRF Model data. The 100 runs of the MoF are shown as box plots, with the boxes denoting the IQR and the red crosses denoting outliers, which are outside the whisker length of 1.5 IQR. (b) Empirical annual exceedance probability for extreme rainfall intensities at different time scales to occur at any of the 10 sites. Dashed lines for the MoF represent the 5% and 95% quantiles of the 100 runs, and the solid line denotes the median of the 100 runs.

at any of the 10 sites are studied. The Weibull plotting position formula (Weibull 1939) is applied to estimate the AEP:

$$\text{AEP}(m) = \frac{m}{N+1}, \quad (6)$$

with m denoting the rank of the sorted precipitation intensities and N denoting the number of years in the respective precipitation record of each site (see Fig. 4b).

The median of the 100 MoF runs can reconstruct the intensities of extreme events across all time scales very well. This is caused by the ability of the MoF algorithm to produce hourly intensities at the LOI, which are above any observed hourly intensity at the rain gauges contributing to the LUT.

The WRF Model delivers satisfying results at the longer time scales of 6 and 12 h with 10% and 9% underestimation but generally underestimates the hourly and 3-h intensities by 17% and 14%. Gregersen et al. (2013) explained the underestimation of the magnitude of extreme events partially with the inherent

deviations between areal, gridded data and point observations.

The highest intensities of up to 40 mm h^{-1} being observed in June 2014 in Blindern during a very small-scale convective event cannot be reproduced by the WRF Model. This is due to the coarser-resolution reanalysis data not forcing the RCM to reproduce the event and the convection-permitting WRF Model not being able to simulate these high intensities with this forcing.

5. Conclusions

Generally, the MoF is able to reproduce the variance, dry proportion, wet hours and spells, and timing of the maximum rainfall burst as well as the intensities of extreme events across all time scales. However, the lag-1 autocorrelation and spatial correlation between the precipitation at the locations could not be restored.

The rainfall totals simulated by the WRF Model show very little bias and can capture the variety among the different rain gauges. Overall, the WRF Model can

reproduce the lag-1 autocorrelation, dry proportion, the amount of wet hours, number and length of intraday wet spells, and spatial correlation. The partial underestimation of variance is mainly governed by the 9%–17% underestimation of precipitation extremes, which can be partially referred to the difference of grid-based simulations to point measurements.

The performance of the MoF shows high suitability for filling data gaps of subdaily historical records and also disaggregation of data from rain gauges measuring only daily precipitation. The preservation of hourly intensities of heavy precipitation events makes the MoF especially relevant for flood modeling of hydrological systems with short response time. The MoF can also be set up in minute resolution, which shows high potential for applications within urban hydrology. However, it requires a sufficient amount of high-resolution rainfall observation stations and relies on historical data, restricting its scope of application regarding future scenarios.

The WRF Model has not only proven that it adds value to the reanalysis data, but also has the ability to generate hourly-resolution precipitation data, reproducing the main characteristics of the observations. The biggest advantage over statistical models is that the WRF Model can be applied in regions without observations. Due to its capability of producing areal precipitation data the WRF Model can serve as a data source for spatially distributed models.

Acknowledgments. The work has been carried out within the SUPER project (250573) and the HYPRE project (243942), both funded by the Research Council of Norway and the ClimEx project funded by the Bavarian State Ministry for the Environment and Consumer Protection. The work has been supported by the European Communities 7th Framework Programme Funding under Grant Agreement 603629-ENV-2013-6.2.1-GLOBAQUA. We acknowledge the precipitation data by the Norwegian Meteorological Institute and the reanalysis data by the National Centers for Environmental Prediction (NCEP; data sources in references).

REFERENCES

- Arnbjerg-Nielsen, K., P. Willems, J. Olsson, S. Beecham, A. Pathirana, I. B. Gregersen, H. Madsen, and V. T. V. Nguyen, 2013: Impacts of climate change on rainfall extremes and urban drainage systems: a review. *Water Sci. Technol.*, **68**, 16–28, <https://doi.org/10.2166/wst.2013.251>.
- Benestad, R. E., 2011: Updated temperature and precipitation scenarios for Norwegian climate regions. Met.no Rep. 16/2011, 111 pp., <https://www.met.no/publikasjoner/met-report/met-report-2011>.
- Bennett, J. C., D. E. Robertson, P. G. D. Ward, H. A. Prasantha Hapuarachchi, and Q. J. Wang, 2016: Calibrating hourly rainfall-runoff models with daily forcings for streamflow forecasting applications in meso-scale catchments. *Environ. Modell. Software*, **76**, 20–36, <https://doi.org/10.1016/j.envsoft.2015.11.006>.
- Bruni, G., R. Reinoso, N. C. van de Giesen, F. H. L. R. Clemens, and J. A. E. ten Veldhuis, 2015: On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrol. Earth Syst. Sci.*, **19**, 691–709, <https://doi.org/10.5194/hess-19-691-2015>.
- Collins, W. D., and Coauthors, 2004: Description of the NCAR Community Atmosphere Model (CAM 3.0). NCAR Tech. Note NCAR/TN-464+STR, 214 pp., <https://doi.org/10.5065/D63N21CH>.
- Cristiano, E., M.-C. ten Veldhuis, and N. van de Giesen, 2017: Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – A review. *Hydrol. Earth Syst. Sci.*, **21**, 3859–3878, <https://doi.org/10.5194/hess-21-3859-2017>.
- Dyrddal, A. V., T. Skaugen, F. Stordal, and E. J. Førland, 2016: Estimating extreme areal precipitation in Norway from a gridded dataset. *Hydrol. Sci. J.*, **61**, 483–494, <https://doi.org/10.1080/02626667.2014.947289>.
- , F. Stordal, and C. Lussana, 2018: 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>.
- EEA, 2018: Europe coastline shapefile. European Environment Agency, accessed 27 April 2018, <https://www.eea.europa.eu/data-and-maps/data/eea-coastline-for-analysis-1/gis-data/europe-coastline-shapefile>.
- Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.
- Geonorge, 2018: DTM 10 Terrengmodell (UTM32). Accessed 26 April 2018, <https://kartkatalog.geonorge.no/metadata/kartverket/dtm-10-terrengmodell-utm32/fd851873-f363-46f9-9fc6-bb1b403575df>.
- Gregersen, I., H. Sørup, H. Madsen, D. Rosbjerg, P. Mikkelsen, and K. Arnbjerg-Nielsen, 2013: Assessing future climatic changes of rainfall extremes at small spatio-temporal scales. *Climatic Change*, **118**, 783–797, <https://doi.org/10.1007/s10584-012-0669-0>.
- Grell, G. A., and S. R. Freitas, 2014: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, **14**, 5233–5250, <https://doi.org/10.5194/acp-14-5233-2014>.
- Haberlandt, U., A.-D. Ebner von Eschenbach, and I. Buchwald, 2008: A space-time hybrid hourly rainfall model for derived flood frequency analysis. *Hydrol. Earth Syst. Sci.*, **12**, 1353–1367, <https://doi.org/10.5194/hess-12-1353-2008>.
- Hanssen-Bauer, I., and Coauthors, 2009: Klima i Norge 2100. Bakgrunnsmateriale til NOU klimatilpassing. NCCS Rep. 2/2015, 204 pp.
- Heikkilä, U., A. Sandvik, and A. Sorteberg, 2011: Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model. *Climate Dyn.*, **37**, 1551–1564, <https://doi.org/10.1007/s00382-010-0928-6>.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42** (2), 129–151.
- , Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Jimenez, P. A., J. Dudhia, J. F. Gonzalez-Rouco, J. Navarro, J. P. Montavez, and E. Garcia-Bustamante, 2012: A revised scheme

- for the WRF surface layer formulation. *Mon. Wea. Rev.*, **140**, 898–918, <https://doi.org/10.1175/MWR-D-11-00056.1>.
- Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 2006: World Map of the Köppen-Geiger climate classification updated. *Meteor. Z.*, **15**, 259–263, <https://doi.org/10.1127/0941-2948/2006/0130>.
- Koutsoyiannis, D., 2003: Rainfall disaggregation methods: Theory and applications. *Proc. Workshop on Statistical and Mathematical Methods for Hydrological Analysis*, Rome, Italy, University of Rome, 1–23, <http://itia.ntua.gr/getfile/570/1/documents/2003RainDisag.pdf>.
- Lall, U., and A. Sharma, 1996: A nearest neighbor bootstrap for time series resampling. *Water Resour. Res.*, **32**, 679–693, <https://doi.org/10.1029/95WR02966>.
- Li, X., and Coauthors, 2018: 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>.
- Mayer, S., C. Fox Maule, S. Sobolowski, O. B. Christensen, H. J. D. Sørup, M. A. Sunyer, K. Arnbjerg-Nielsen, and I. Barstad, 2015: Identifying added value in high-resolution climate simulations over Scandinavia. *Tellus*, **67A**, 24941, <https://doi.org/10.3402/tellusa.v67.24941>.
- NCEP, 2000: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, accessed 22 February 2018, <https://doi.org/10.5065/D6M043C6>.
- NMI, 2018: eKlima. Norwegian Meteorological Institute, accessed 12 February 2018, <http://eKlima.met.no>.
- Ochoa-Rodriguez, S., and Coauthors, 2015: Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *J. Hydrol.*, **531**, 389–407, <https://doi.org/10.1016/j.jhydrol.2015.05.035>.
- Park, S., and C. S. Bretherton, 2009: The University of Washington shallow convection and moist turbulence schemes and their impact on climate simulations with the Community Atmosphere Model. *J. Climate*, **22**, 3449–3469, <https://doi.org/10.1175/2008JCLI2557.1>.
- Powers, J. G., J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia, D. O. Gill, J. L. Coen, and D. J. Gochis, 2017: The Weather Research and Forecasting (WRF) Model: Overview, system efforts, and future directions. *Bull. Amer. Meteor. Soc.*, **98**, 1717–1737, <https://doi.org/10.1175/BAMS-D-15-00308.1>.
- Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361, <https://doi.org/10.1002/2014RG000475>.
- Pui, A., A. Sharma, R. Mehrotra, B. Sivakumar, and E. Jeremiah, 2012: A comparison of alternatives for daily to sub-daily rainfall disaggregation. *J. Hydrol.*, **470–471**, 138–157, <https://doi.org/10.1016/j.jhydrol.2012.08.041>.
- Reynolds, J. E., S. Halldin, C. Y. Xu, J. Seibert, and A. Kauffeldt, 2017: Sub-daily runoff predictions using parameters calibrated on the basis of data with a daily temporal resolution. *J. Hydrol.*, **550**, 399–411, <https://doi.org/10.1016/j.jhydrol.2017.05.012>.
- Sharma, A., and S. Srikanthan, 2006: Continuous rainfall simulation: A nonparametric alternative. *30th Hydrology & Water Resources Symp.: Past, Present & Future*, Launceston, Australia, Institution of Engineers, 86–91.
- Sillmann, J., and Coauthors, 2017: Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities. *Wea. Climate Extremes*, **18**, 65–74, <https://doi.org/10.1016/j.wace.2017.10.003>.
- Skamarock, W., and J. Klemp, 2008: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J. Comput. Phys.*, **227**, 3465–3485, <https://doi.org/10.1016/j.jcp.2007.01.037>.
- Socolofsky, S., E. E. Adams, and D. Entekhabi, 2001: Disaggregation of daily rainfall for continuous watershed modeling. *J. Hydrol. Eng.*, **6**, 300–309, [https://doi.org/10.1061/\(ASCE\)1084-0699\(2001\)6:4\(300\)](https://doi.org/10.1061/(ASCE)1084-0699(2001)6:4(300)).
- Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2006: How often does it rain? *J. Climate*, **19**, 916–934, <https://doi.org/10.1175/JCLI3672.1>.
- Tabari, H., and Coauthors, 2016: Local impact analysis of climate change on precipitation extremes: Are high-resolution climate models needed for realistic simulations? *Hydrol. Earth Syst. Sci.*, **20**, 3843–3857, <https://doi.org/10.5194/hess-20-3843-2016>.
- Tewari, M., and Coauthors, 2004: Implementation and verification of the unified NOAA land surface model in the WRF Model. *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 14.2a, https://ams.confex.com/ams/84Annual/techprogram/paper_69061.htm.
- Tjelta, T., and J. Mamen, 2014: Climate trends and variability of rain rate derived from long-term measurements in Norway. *Radio Sci.*, **49**, 788–797, <https://doi.org/10.1002/2014RS005477>.
- Weibull, W., 1939: A statistical theory of the strength of materials. *Ing. Vetensk. Akad. Handl.*, **151**, 1–45.
- Westra, S., R. Mehrotra, A. Sharma, and R. Srikanthan, 2012: Continuous rainfall simulation: 1. A regionalized subdaily disaggregation approach. *Water Resour. Res.*, **48**, W01535, <https://doi.org/10.1029/2011WR010489>.