

RESEARCH ARTICLE

Analogue methods and ERA5: Benefits and pitfalls

Pascal Horton 

Oeschger Centre for Climate Change
Research and Institute of Geography,
University of Bern, Bern, Switzerland

Correspondence

Pascal Horton, Oeschger Centre for
Climate Change Research and Institute of
Geography, University of Bern, Bern 3012,
Switzerland.
Email: pascal.horton@giub.unibe.ch

Abstract

Perfect prognosis statistical downscaling relies on the statistical relationships established using observational data for predictands and predictors. Predictors are often retrieved from reanalyses, which are considered pseudo-observations. The impact of the choice of a reanalysis dataset on the performance of the downscaling method is usually overlooked, as global reanalyses are frequently assumed to be equivalent for the last few decades and data-rich regions such as Europe. However, it was recently shown that the reanalysis dataset can have a bigger impact on the method skill than the choice of predictor variables. Generally, reanalyses processed by more recent atmospheric models assimilate more data and perform best. This work is aimed at assessing the extent of potential gains from the use of ERA5, following its release, compared to other global reanalyses. The assessment was carried out using six variants of analogue methods, which are statistical downscaling techniques, to predict daily precipitation at 301 stations across Switzerland. ERA5 proved to be one of the best performing reanalyses across the different analogue methods. Due to data availability, we recommend using 20CR for applications starting between 1851 and 1900, CERA-20C for those between 1900 and 1950, and ERA5 for applications after 1950. However, ERA5 high spatial resolution (0.25°) turned out to be a trap for simple calibration techniques. The domains over which the predictor fields are compared need to be optimized, and high-resolution grids come along with numerous sub-optimal local solutions. An enhanced calibration procedure, thus, must be used. Besides the risk of poorly-calibrated domains, the high resolution also requires much higher computational time with no gain in skill, provided that the predictors considered are relevant at a synoptic scale. Although ERA5 should be the dataset of choice, its use at a lower resolution to predict daily precipitation should provide equivalent performance.

KEYWORDS

analogue methods, ERA5, precipitation, reanalyses, statistical downscaling, Switzerland

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Author. *International Journal of Climatology* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

1 | INTRODUCTION

Analogue methods (AMs) are statistical downscaling techniques that predict local meteorological variables, often daily precipitation, based on large-scale predictors. AMs rely on the assumption that similar synoptic situations are likely to result in similar local effects plus some variability that cannot be explained by the considered predictors (Lorenz, 1969). To account for this unexplained variability, an ensemble of analogue situations is considered. It thus provides a statistical prediction in the form of an empirical conditional distribution made from the corresponding observed predictand values. Multiple variants of the AM with different structures, approaches, and—mainly—different predictors, exist.

Applications of AMs consist of, for example, daily precipitation predictions, either in an operational forecasting context (e.g., Bontron and Obled, 2005; Hamill and Whitaker, 2006; Bliefernicht, 2010; Horton *et al.*, 2012; Marty *et al.*, 2012; Hamill *et al.*, 2015; Ben Daoud *et al.*, 2016) or for downscaling climate models (e.g., Zorita and von Storch, 1999; Wetterhall, 2005; Matulla *et al.*, 2007; Wetterhall *et al.*, 2007; Radanovics *et al.*, 2013; Chardon *et al.*, 2014; Dayon *et al.*, 2015; Raynaud *et al.*, 2016). Other applications target precipitation radar images (Panziera *et al.*, 2011; Foresti *et al.*, 2015), temperature (Kruizinga and Murphy, 1983; Delle Monache *et al.*, 2013; Caillouet *et al.*, 2016; Raynaud *et al.*, 2016), wind (Delle Monache *et al.*, 2011; Delle Monache *et al.*, 2013; Vanvyve *et al.*, 2015; Junk *et al.*, 2015a; Alessandrini *et al.*, 2015b; Junk *et al.*, 2015b), solar radiation or power production (Bessa *et al.*, 2015; Alessandrini *et al.*, 2015a; Raynaud *et al.*, 2016), and the trajectory of tropical cyclones (Sievers *et al.*, 2000; Fraedrich *et al.*, 2003). AMs are also used for seasonal forecast (Xavier and Goswami, 2007; Charles *et al.*, 2012; Wu *et al.*, 2012; Shao and Li, 2013).

AMs are most often developed in a perfect prognosis framework (Rummukainen, 1997; Maraun *et al.*, 2010), where the statistical relationship is calibrated between large-scale and local-scale observations. In this context, global reanalyses are the datasets of choice for large-scale variables as they provide multivariate gridded outputs that are physically consistent and available all around the globe (Gelaro *et al.*, 2017). Once the relationship has been established (i.e., the predictor variables selected and the other parameters calibrated), the AM is then applied in a different context, such as forecasting, climate reconstruction, or climate projections. In such contexts, precipitation is predicted using the target model output to describe the target situation for the AM and the reanalysis dataset as the archive to retrieve past analogue situations (if no long reforecast dataset is available).

Globally, two main types of reanalysis products exist: those that aim for homogeneity over a long period—starting at the beginning of the 20th century—and hence assimilate surface data only, and those that aim for accuracy over a more recent period (typically 1979) and therefore assimilate as many observations as possible, including multiple satellite products. The reanalyses accuracy depends on the quality of the model physics and the analysis process, as well as the quantity and quality of the assimilated observations (Dee *et al.*, 2011).

The nature of the application can drive the choice of a reanalysis dataset for AMs. For example, if a coverage of the 20th century is needed, the Twentieth Century Reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) (ERA-20C or CERA-20C—Poli *et al.*, 2016; Laloyaux *et al.*, 2016) or 20CR (Compo *et al.*, 2011) produced by NOAA—now available as version 3 (Slivinski *et al.*, 2019, 2021)—can be used (e.g., Kuentz *et al.*, 2015; Brigode *et al.*, 2016; Caillouet *et al.*, 2016; Bonnet *et al.*, 2017). In an operational forecasting context, with the target situation being provided by an NWP model, one should prefer a reanalysis that is produced by the same NWP model as the operational forecast to reduce inter-model biases. However, in many cases, the selection of the reanalysis is arbitrary and might be driven by a preference for the local provider.

NCEP/NCAR Reanalysis 1 (NR-1—Kalnay *et al.*, 1996; Kistler *et al.*, 2001) and NCEP/DOE Reanalysis 2 (NR-2—Kanamitsu *et al.*, 2002) were used in many AM applications (e.g., Timbal *et al.*, 2003; Altava-Ortiz *et al.*, 2006; Matulla *et al.*, 2007; Yiou *et al.*, 2014), ERA-40 (Uppala *et al.*, 2005), produced by ECMWF, has also been used substantially (e.g., Willems and Vrac, 2011; Radanovics *et al.*, 2013; Chardon *et al.*, 2014; Ben Daoud *et al.*, 2016). Its successor, ERA-Interim (ERA-INT—Dee *et al.*, 2011), has been used by Raynaud *et al.* (2016). NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA—Rienecker *et al.*, 2011) has been used by Vanvyve *et al.* (2015). The Japanese 55-year Reanalysis (JRA-55—Kobayashi *et al.*, 2015; Harada *et al.*, 2016) and its conventional-data-only JRA-55 Conventional (JRA-55C—Kobayashi *et al.*, 2014), have not been used in AMs to the author's knowledge. Newer products, such as NCEP's Climate Forecast System Reanalysis (CFSR—Saha *et al.*, 2010), MERRA version 2 (MERRA-2—Gelaro *et al.*, 2017) or ERA5 (Hersbach *et al.*, 2020) have not been used much in AMs. This is the result of the lag one often observes between the release of a new reanalysis dataset and its adoption in AMs.

Most AM applications are based on a single reanalysis dataset and its impact is overlooked. Ben Daoud *et al.* (2009)

TABLE 1 Assessed reanalysis datasets with their respective properties, sorted by type and model age

Name	Institution	Coverage	Output [°]	Model resolution and age	Input	Assimilation
20CR-2c	NOAA-CIRES	1851–2014	2 × 2	T62 (~1.88°), L28, 2008	Surface	EnKF
ERA-20C	ECMWF	1900–2010	1 × 1	TL159 (~1.13°), L9, 2012	Surface	4D-Var
CERA-20C	ECMWF	1901–2010	1 × 1	T159 (~1.13°), L91, 2016	Surface	4D-Var
NR-1	NCEP, NCAR	1948–Present	2.5 × 2.5	T62 (~1.88°), L28, 1995	Full	3D-Var
NR-2	NCEP, DOE	1979–Present	2.5 × 2.5	T62 (~1.88°), L28, 2001	Full	3D-Var
CFSR	NCEP	1979–Present	0.5 × 0.5	T382 (~0.31°), L64, 2009	Full	3D-Var
JRA-55	JMA	1958–Present	1.25 × 1.25	TL319 (~0.36°), L60, 2009	Full	4D-Var
MERRA-2	NASA GMAO	1980–Present	0.625 × 0.5	0.625° × 0.5°, L72, 2014	Full	3D-Var
ERA-INT	ECMWF	1979–2019	0.75 × 0.75	TL255 (~0.70°), L60, 2006	Full	4D-Var
ERA5	ECMWF	1950–Present*	0.25 × 0.25	T639 (~0.28°), L137, 2016	Full	4D-Var

*The data over the period 1950–1978 is only available in a preliminary version at time of writing.

compared NR-1 with ERA-40 and found no significant difference for the predictors considered. Later, Dayon *et al.* (2015) compared NR-1, MERRA, ERA-INT, and 20CR, and found that the choice of the reanalysis dataset has a non-negligible impact on the performance of AMs. This impact can even be greater than that of the choice of the predictor variables. It was concluded that the role of the reanalyses should not be underestimated. Such influence has also been observed for other statistical downscaling methods (e.g., Koukidis and Berg, 2009). Horton and Brönnimann (2019) compared 10 global reanalyses and concluded that the impact of the dataset on the results of the AM is significant. Similarly, to the work of Dayon *et al.* (2015), the impact of the reanalysis was sometimes found to be higher than that of the selected predictors.

Generally, more recent products that assimilate more data and come with a higher spatial resolution result in better predictions in AMs (Horton and Brönnimann, 2019). However, this is not systematically the case as, for example, JRA-55, which has a lower spatial resolution (1.25°), was found to be one of the best options for more complex AMs. Moreover, the choice of a reanalysis can be driven by the period of interest. As there was no better dataset overall, a table has been provided by Horton and Brönnimann (2019) for guidance in the selection of a reanalysis for a given predictor and period. Some characteristics of the datasets were also assessed, such as the spatial resolution, archive length, and use of ensemble members. For most AMs and datasets, there was a gain in skill with increasing resolution up to 1°; below this, however, the gains were almost null. The use of long archives for reanalyses with extended data improved the performance to some extent by providing potentially better analogue situations, particularly for less frequent situations. However, former times were more data-scarce and data might be inhomogeneous over long periods, which can, in turn, have a negative effect on the skill of AMs

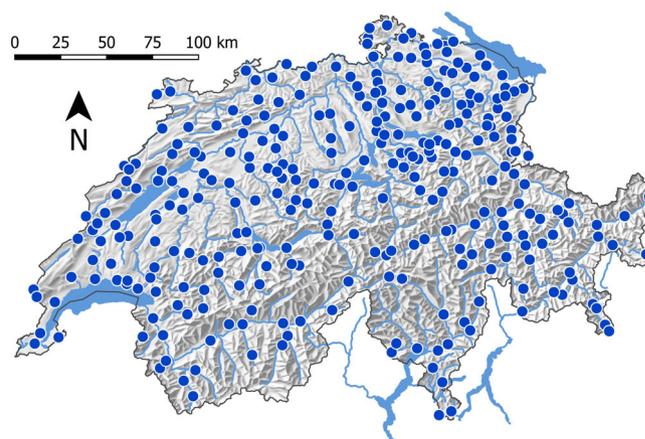


FIGURE 1 Map of the 301 precipitation stations with good data coverage of the period 1981–2010. Background map: © SwissTopo. Modified from Horton and Brönnimann (2019) [Colour figure can be viewed at wileyonlinelibrary.com]

exploiting these archives (Horton and Brönnimann, 2019). Finally, the use of ensemble members showed limited benefits and that great caution is required for their use in methods with multiple levels of analogy.

This work is a follow-up to Horton and Brönnimann (2019) and focuses on ERA5 by comparing its impact on six AMs using nine other major reanalyses at 301 stations in Switzerland. The study of the characteristics of the reanalyses as provided in Horton and Brönnimann (2019) is not repeated here as the conclusions are likely identical. The optimal parameters, such as the number of analogues and the size of the comparison domain, were here compared across reanalyses and discussed. This provides an original insight into the role of the spatial resolution and reanalysis quality. The benefits and potential pitfalls of ERA5's high spatial resolution are also assessed.

The article is structured as follows: Section 2 describes the reanalyses compared in this work, provides some information on the precipitation dataset used and the study period, details the assessed AMs and their calibration. Section 3 shows the results. It starts with an analysis of the role of the spatial resolution and the related pitfalls for the calibration procedure (Section 3.1). A comparison of the AM parameters calibrated for the different reanalyses is then provided (Section 3.2). Finally, it shows how ERA5 compares to the other reanalyses (Section 3.3). The conclusions are provided in Section 4.

2 | DATA AND METHODS

2.1 | Reanalyses

The present work aims at comparing ERA5 (Hersbach *et al.*, 2020) with other global reanalyses, whose characteristics are provided in Table 1, thus complementing the work of Horton and Brönnimann (2019). The first three reanalyses in Table 1 are surface-input (Fujiwara *et al.*, 2017) products that assimilate surface data only but cover an extended period, typically from the start of 20th century. NOAA produced the Twentieth Century Reanalysis version 2c, 20CR-2c—Compo *et al.*, 2011), which only assimilates surface pressure data and uses

observed monthly sea-surface temperatures and sea-ice distributions as boundary conditions. A more recent version, 20CRv3 (Slivinski *et al.*, 2019, 2021), has been released since the comparison presented here was performed (and is thus not part of this work). This third version solves issues found in version 2c and was produced using an upgraded data assimilation system as well as a newer forecast model. The ECMWF developed two products using surface input only. The first is ERA-20C (Poli *et al.*, 2016), which assimilates marine wind observations and is forced by sea surface temperature, sea ice cover, atmospheric composition changes, and solar forcing. The second is CERA-20C (Laloyaux *et al.*, 2018), which has an additional coupling to the ocean and was produced by a more recent version of the Integrated Forecasting System (IFS) model.

The other reanalyses are full-input products that assimilate all available data, including satellite data (Fujiwara *et al.*, 2017). NCEP/NCAR Reanalysis 1 (NR-1—Kalnay *et al.*, 1996; Kistler *et al.*, 2001) was the first global reanalysis followed by NCEP/DOE Reanalysis 2 (NR-2—Kanamitsu *et al.*, 2002), which fixed some identified problems. The Climate Forecast System Reanalysis (CFSR – Saha *et al.*, 2010) is the most recent one produced by the NCEP. The Japanese 55-year Reanalysis (JRA-55—Kobayashi *et al.*, 2015; Harada *et al.*, 2016) is produced by the Japan Meteorological Agency (JMA). JRA-55 Conventional (JRA-55C—Kobayashi *et al.*, 2014),

TABLE 2 Analogue methods considered in this work, listed by increasing complexity

Method	P0	L1	L2	L3	Reference
2Z	PC	Z1000@12 hr Z500@24 hr			Bontron (2004)
4Z	PC	Z1000@06 hr Z1000@30 hr Z700@24 hr Z500@12 hr			Horton <i>et al.</i> (2018)
2Z-2MI	PC	Z1000@12 hr Z500@24 hr	MI850@12 + 24 hr		Bontron (2004)
4Z-2MI	PC	Z1000@30 hr Z850@12 hr Z700@24 hr Z400@12 hr	MI700@24 hr MI600@12 hr		Horton <i>et al.</i> (2018)
PT-2Z-4MI	T925@36 hr T600@12 hr	Z1000@12 hr Z500@24 hr	MI925@12 + 24 hr MI700@12 + 24 hr		BenDaoud (2016)
PT-2Z-4W-4MI	T925@36 hr T600@12 hr	Z1000@12 hr Z500@24 hr	W850@06-24 hr	MI925@12 + 24 hr MI700@12 + 24 hr	Ben Daoud <i>et al.</i> (2016)

Note: The analogy criterion is S1 for SLP and Z and RMSE for the other variables. P0, preselection (PC: ± 60 days around the target date); L1, L2, and L3, subsequent levels of analogy. Z, geopotential height; T, air temperature; W, vertical velocity; MI, moisture index (product of the relative humidity at the given pressure level and the total water column).

which uses conventional data only, was not considered in this work as it provides similar results to JRA-55 (Horton and Brönnimann, 2019). NASA's Global Modelling and Assimilation Office (GMAO) released the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2—Gelaro *et al.*, 2017), which is an improvement on the first MERRA reanalysis Rienecker *et al.*, 2011). Finally, the predecessor of ERA5, ERA-Interim (ERA-INT—Dee *et al.*, 2011), was also considered.

ERA5 (Hersbach *et al.*, 2020) is meant to replace ERA-Interim. It benefits from multiple improvements to IFS in terms of model physics, core dynamics, and data assimilation (Hersbach *et al.*, 2020). It provides more outputs with higher temporal (hourly) and spatial (0.25°) resolutions. It has the highest spatial and temporal resolution to date. The use of a 10-member ensemble of data assimilation at a lower spatial and temporal resolution allows an estimation of its uncertainty. ERA5 assimilates significantly more data than ERA-Interim, such as data from ground-based radar and new satellite sensors, and uses improved observation operators, thus allowing better comparison of model outputs with observations.

Additionally, it indirectly benefits from improvements in historical observations of both conventional and satellite data (Hersbach *et al.*, 2020). ERA5 provides, among others, more consistent sea surface temperature and sea ice concentration, an improved representation of tropical cyclones, a better balance of evaporation and precipitation, and improved soil moisture. ERA5 is also more suitable for climate analyses as it relies on appropriate radiative forcing and boundary conditions (e.g., changes in greenhouse gases, aerosols, SST, and sea ice).

2.2 | Precipitation dataset

The local variables to be predicted here are daily precipitation totals (06:00 h UTC to 06:00 h UTC the following day) at 301 stations of the MeteoSwiss network in Switzerland (Figure 1), with good coverage of the 1981–2010 period. The 30-year precipitation dataset was divided into a calibration period (CP) and an independent validation period (VP) which was evenly distributed over the entire series (1 year out of every 5; a total of

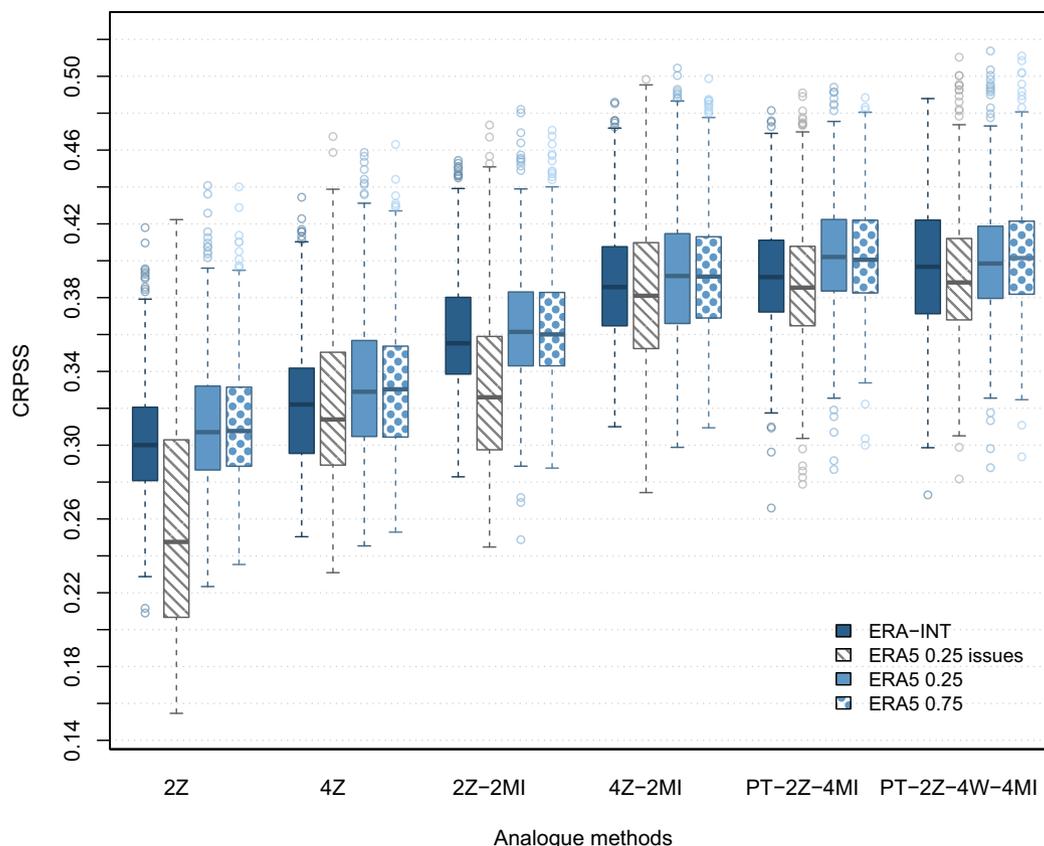


FIGURE 2 Comparison of ERA-INT and ERA5 at 0.25° and 0.75° (dotted) as well as a failed calibration of AMs with ERA5 (striped; named ‘ERA5 0.25 issues’; see explanation in text). The performance is given by the CRPSS for all stations and all AMs on the VP. The boxes show the 25th, 50th, and 75th percentiles. The whiskers extend to the most extreme data point, which is no more than 1.5 times the interquartile range [Colour figure can be viewed at wileyonlinelibrary.com]

6 years). The archive period (AP), in which the analogue dates are being retrieved, is the same as the CP—also with the VP being excluded—with an additional exclusion of ± 30 days around the target date. The results below are always shown for the VP.

2.3 | Analogue methods

Most AM variants from the analysis of Horton and Brönnimann (2019) were considered here (Table 2). These AMs have different degrees of complexity, but they all start



FIGURE 3 Calibrated number of analogues for the different methods and reanalyses, and for all stations. The reanalyses are ordered by increasing spatial resolution. The triangles represent the mean values of all stations [Colour figure can be viewed at wileyonlinelibrary.com]

with a preselection to cope with seasonality. A commonly used preselection is based on calendar dates (PC: preselection on a calendar-basis in Table 2) as candidate situations are extracted from the archive for a period of 120 days centred around the target date. To allow for a more dynamic

approach, Ben Daoud *et al.* (2016) based this preselection on the similarity of air temperature (T) at 925 and 600 hPa at the nearest grid point. A resulting mixing of spring and autumn situations, which is undesirable, is discussed in Caillouet *et al.* (2016).

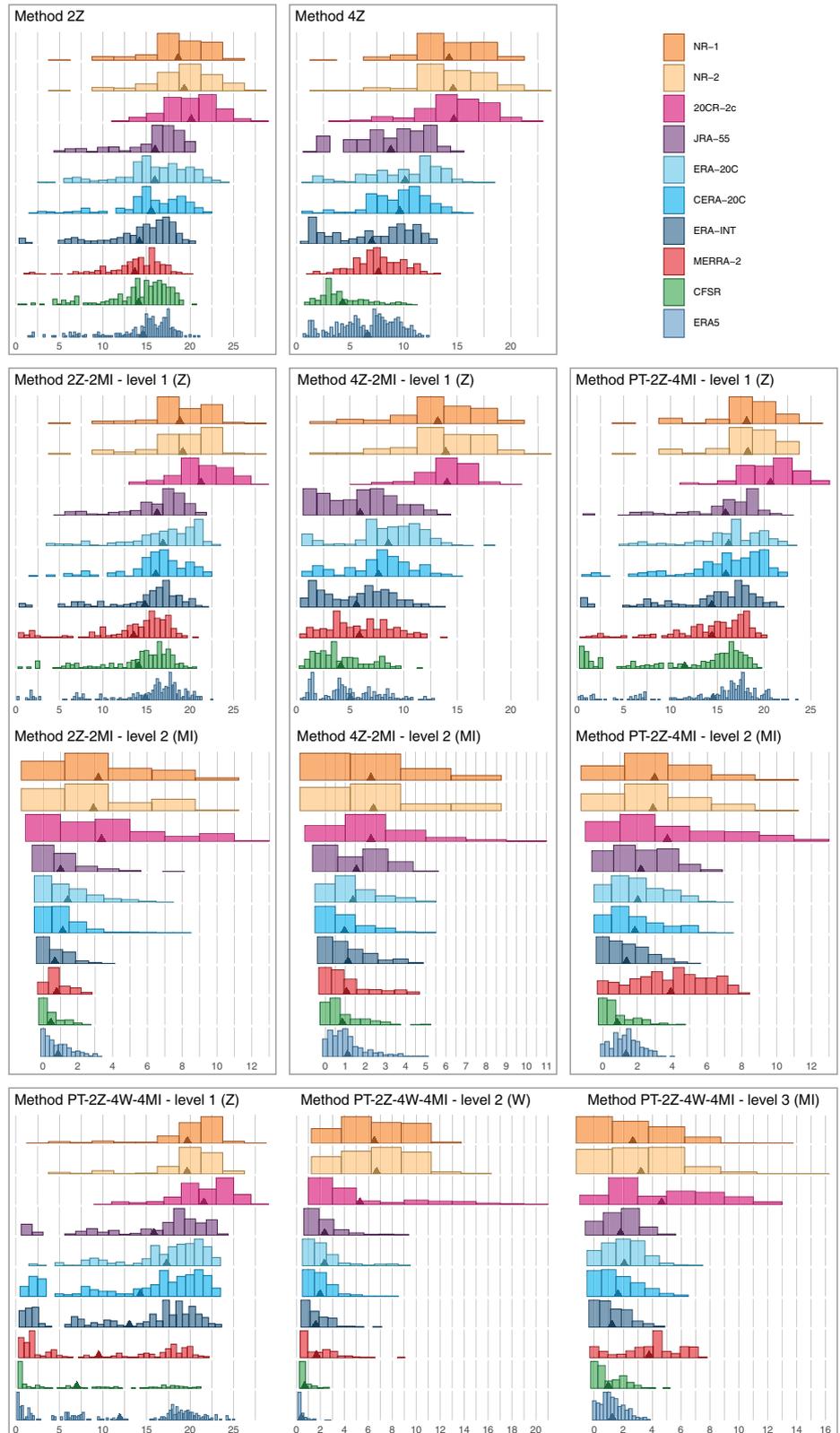


FIGURE 4 Calibrated zonal extents (in degrees) of the spatial windows for the different methods and reanalyses, and for all stations. Same conventions as Figure 3 [Colour figure can be viewed at wileyonlinelibrary.com]

All considered AMs have a first level of analogy based on the atmospheric circulation, using the geopotential height (Z) at different pressure levels and time as predictor. This analogy is quantified by the S1 criterion (Equation (1), Teweles and Wobus, 1954; Brown

et al., 2012), which is a comparison of gradients over a selected (calibrated) domain. The S1 criterion allows comparing the atmospheric circulation instead of the absolute values of the respective fields. The S1 values processed for different levels/hours are averaged into

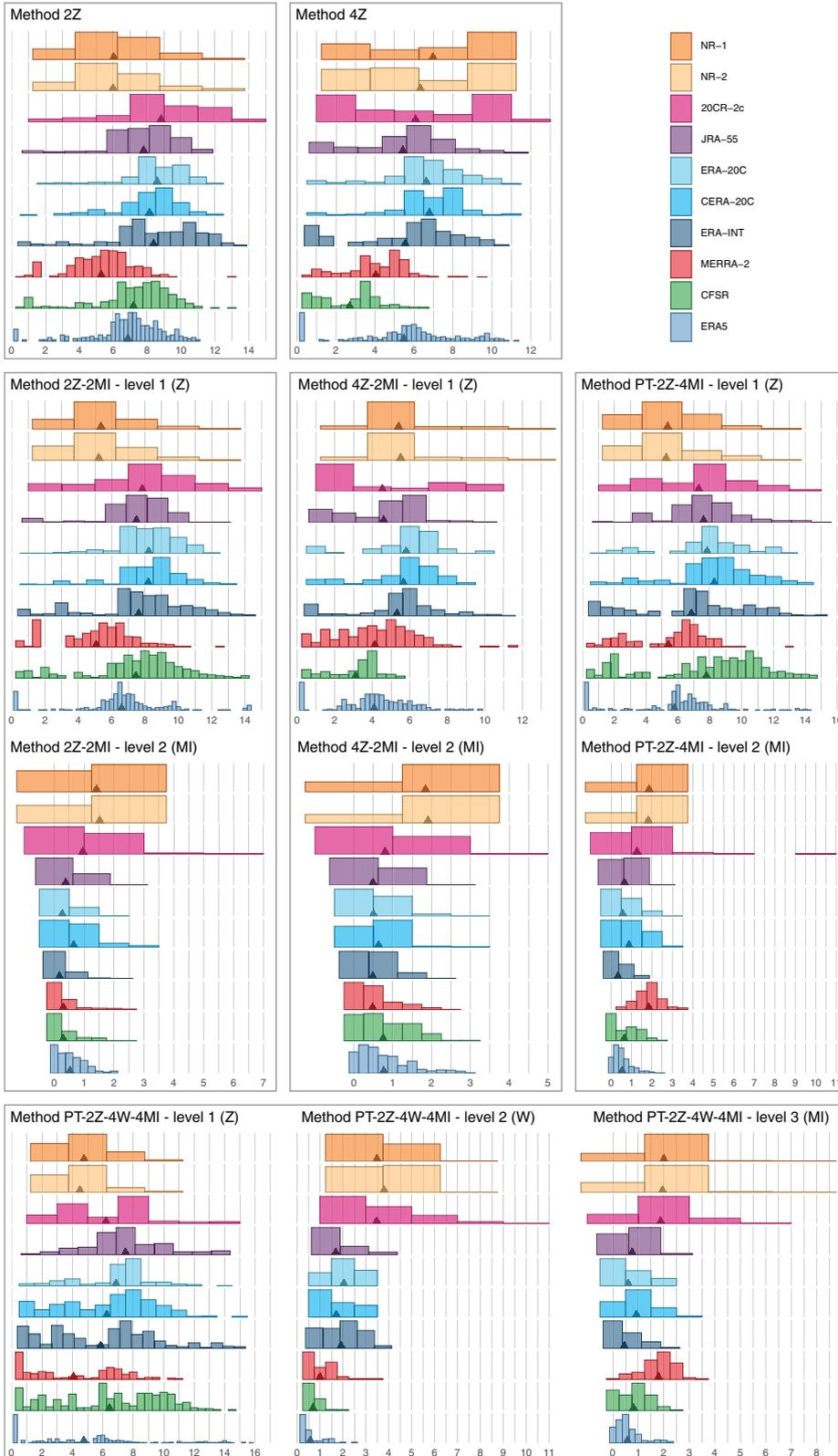


FIGURE 5 Calibrated meridional extents (in degrees) of the spatial windows for the different methods and reanalyses, and for all stations. Same conventions as Figure 3 [Colour figure can be viewed at wileyonlinelibrary.com]

a single value. More advanced approaches introduce a weighting between these components (Horton *et al.*, 2017a).

$$S1 = 100 \frac{\sum_i |\Delta\hat{z}_i - \Delta z_i|}{\sum_i \max\{|\Delta\hat{z}_i|; |\Delta z_i|\}} \quad (1)$$

where $\Delta\hat{z}_i$ is the geopotential height gradient between the i -th pair of points for the target day and Δz_i is the corresponding observed geopotential height gradient for the candidate situation. The smaller the value $S1$ is, the more similar are the pressure fields.

The simplest method, 2Z, is based on the analogy of the atmospheric circulation alone, using the geopotential height at 1,000 and 500 hPa (Bontron, 2004). The most similar N_1 dates with the lowest values of $S1$ are selected as analogues to the target date. The daily precipitation measured on these N_1 dates then provides the empirical conditional distribution, considered as the probabilistic prediction for the target date.

The second method, 4Z, is also based only on geopotential heights, but it uses four combinations of pressure levels and temporal windows (Horton *et al.*, 2018). The method considered here is a simplified version of a more sophisticated method elaborated by genetic algorithms. It was found that using four geopotential heights provides more information than using two (Horton *et al.*, 2018).

The next three methods (2Z-2MI, 4Z-2MI, and PT-2Z-4MI, Table 2) add a second level of analogy based on moisture variables to subsample from the first level. The predictor for this second level of analogy is a moisture index (MI). It is the product of the total precipitable water (TPW) and the relative humidity (RH) and was introduced by Bontron (2004). This predictor is assessed using the root mean square error (RMSE) criterion. The 2Z-2MI method (Bontron, 2004) uses this index at 850 hPa for two different hours (+12 and +24 h). The 4Z-2MI method (Horton *et al.*, 2018), derived from optimizations with genetic algorithms, uses the MI at 600 and 700 hPa, with a change in the selected levels of the geopotential height compared to 4Z. The 4Z-2MI method (Ben Daoud *et al.*, 2016) considers the MI at 700 and 925 hPa, both at +12 and +24 h.

Finally, the most complex method, PT-2Z-4W-4MI, also named Stepwise Analogue Downscaling Method for Hydrology (SANDHY) (Ben Daoud *et al.*, 2016; Caillouet *et al.*, 2016), adds an intermediate level of analogy before the moisture predictors, based on the vertical velocity (W) at 850 hPa. It was primarily developed for large and relatively flat/lowland catchments in France (Saône, Seine).

2.4 | Calibration of the methods

The Continuous Ranked Probability Score (CRPS, Brown, 1974; Matheson and Winkler, 1976; Hersbach, 2000) is often used as the performance metric for AMs as it allows evaluation of the predicted cumulative distribution provided by the analogues with regard to the single observed value on the target date. Its skill score expression, the Continuous Ranked Probability Skill Score (CRPSS), was used here, using the climatological distribution of precipitation at each station as the reference. It allows a better comparison among stations as it takes into account the differences in climatology to some extent. The CRPSS is expressed as follows (Bradley and Schwartz, 2011):

$$CRPSS = 1 - \frac{\overline{CRPS}}{\overline{CRPS}_{clim}} \quad (2)$$

where $CRPS_{clim}$ is the CRPS value for the climatological distribution. A better prediction is characterized by an increase in CRPSS.

The methods were calibrated using the semi-automatic sequential procedure, which was developed by Bontron (2004) and also used in Radanovics *et al.* (2013) and Ben Daoud *et al.* (2016). It allows calibrating the spatial windows on which the predictors are compared (specific to each level of analogy) and the number of analogue dates for the different levels of analogy. The procedure starts with the calculation of the relevance map, which involves assessing the skill of all points of the predictor grid within a predefined domain. The spatial windows are then calibrated by incremental increases starting from the data point(s) with the highest skill. Increases are assessed in the four directions and the most skillful move is retained. The procedure stops when increasing the domain does not improve the skill. The number of analogues is calibrated jointly for all levels of analogy once the spatial windows are calibrated by assessing multiple combinations. A detailed description of this procedure can be found in Horton (2019a).

After the initial calibration of all AMs for ERA5, some issues were identified with the spatial windows (shown and discussed in Section 3.1). These were due to some limitations of the calibration technique related to the high-resolution grid. The calibration procedure was then repeated using an improved version of the semi-automatic technique, named classic+ (Horton, 2019a). This allows additional moves in the calibration of the spatial windows and thus allows escaping local solutions to some extent. The additional moves are: (1) an increase in two simultaneous directions, (2) a decrease in one or

two simultaneous directions, (3) expansion or contraction in every direction, (4) a shift of the window (without resizing) in eight directions including diagonals, and (5) all the moves described above but with a factor of 2, 3, or more (3 was considered the maximum here). These additional moves can help avoid local optima but do not guarantee optimal spatial windows. Another advantage of the classic+ calibration is that it does not assess all grid points when processing the relevance map (identification of the most skilled point), which substantially saves time.

The AtmoSwing software (Horton, 2019a) was used to perform these calibrations. The calculations were performed on an HPC cluster at the University of Bern. The calibration was done independently for every station, method, and reanalysis. It means that the spatial windows and number of analogues differ from one reanalysis to the other, and that these changes will impact the selection of the analogue dates. The a priori selection of common spatial windows would not have been possible due to differences in grid resolutions (e.g., MERRA-2 does not overlap easily with the other grids) and it would not have been fair for reanalyses for which optimal windows differ from the pre-selected ones. With independent calibrations, all spatial windows should be optimal for the reanalysis at hand so that the comparison between reanalyses is more accurate.

3 | RESULTS

The results are provided for the VP (independent VP, Section 2.2). All reanalyses were used at their highest available spatial resolution and at a 6-hourly time step. The results for reanalyses other than ERA5 were taken from the work of Horton and Brönnimann (2019).

3.1 | High spatial resolution: pitfalls and impact on performance

Although the calibrations were performed on an HPC cluster, using the classic calibration ended up being highly time-consuming for ERA5 (Section 2.4). The reason is that the process involves calculating a relevance map for each level of analogy, which requires assessments of every unitary cell of the predictor grid and the successive extensions of the spatial window in all four directions. On a high-resolution grid, these procedures imply numerous assessments.

The first results of the classic calibration showed that ERA5 (greyed bars in Figure 2) underperformed ERA-INT. An analysis of the resulting parameters revealed that the spatial windows for Z were significantly smaller

for ERA5 as compared to the other reanalyses. For example, for 2Z, the mean zonal and meridional extents of the spatial windows were less than half of those for ERA-INT. In general, while using ERA5, the spatial windows for Z were very small whereas the number of analogues was higher than that for ERA-INT. The reason for such non-optimal parameters is due to the classic calibration being stuck in local solutions during the procedure of the spatial window expansion, as none of the increments in the four directions resulted in a gain in skill. This issue, which was caused by the high resolution of ERA5, did not occur in a significant way for the other reanalyses. High resolution generates a noisier response surface in terms of skill, with local solutions that can trap the classic calibration. The problem here is not the lack of fidelity of the high-resolution input, but limitations of the calibration procedure. To address this issue, the classic+ calibration (Section 2.4) was then used. It allows additional moves in the growth of the spatial windows, including simultaneous increments in multiple grid points to get out of local solutions. The classic+ calibration also ended up being faster as it does not assess every grid point when calculating the relevance map. This approach still uses the full resolution of ERA5, but improves the calibration of the spatial windows. Global optimization techniques, such as genetic algorithms (Horton *et al.*, 2017a), would have overcome this issue more efficiently. However, carrying out this comparison with genetic algorithms would have been computationally intensive and too time-consuming. With the advent of high-resolution datasets, one should be aware of the potential pitfalls of calibration techniques that were developed on low-resolution datasets.

Horton and Brönnimann (2019) showed that below 1° , the spatial resolution of the reanalyses generally has no significant impact on the performance of AMs. It is because, for example, the geopotential at 500 hPa has a half-autocorrelation distance of about 1,000 km (Thiébaux, 1985) and thus, the high resolution does not bring new information. It might, however, be different for moisture variables. ERA5 has the highest spatial resolution (0.25°) among global reanalyses. An additional calibration was performed using the same resolution as ERA-INT (0.75°) to quantify the contribution of the resolution to the gain in skill. Figure 2 shows that degrading the resolution of ERA5 to the levels of ERA-INT (named ERA5 0.75 in Figure 2) did not significantly impact the score. Although a higher model resolution is likely to have contributed to the improvements of ERA5 as compared to ERA-INT, the high output resolution has no significant impact on the AMs' performance for the prediction of daily precipitation. We expected to see some gains for moisture variables, but they did not materialize. The gains

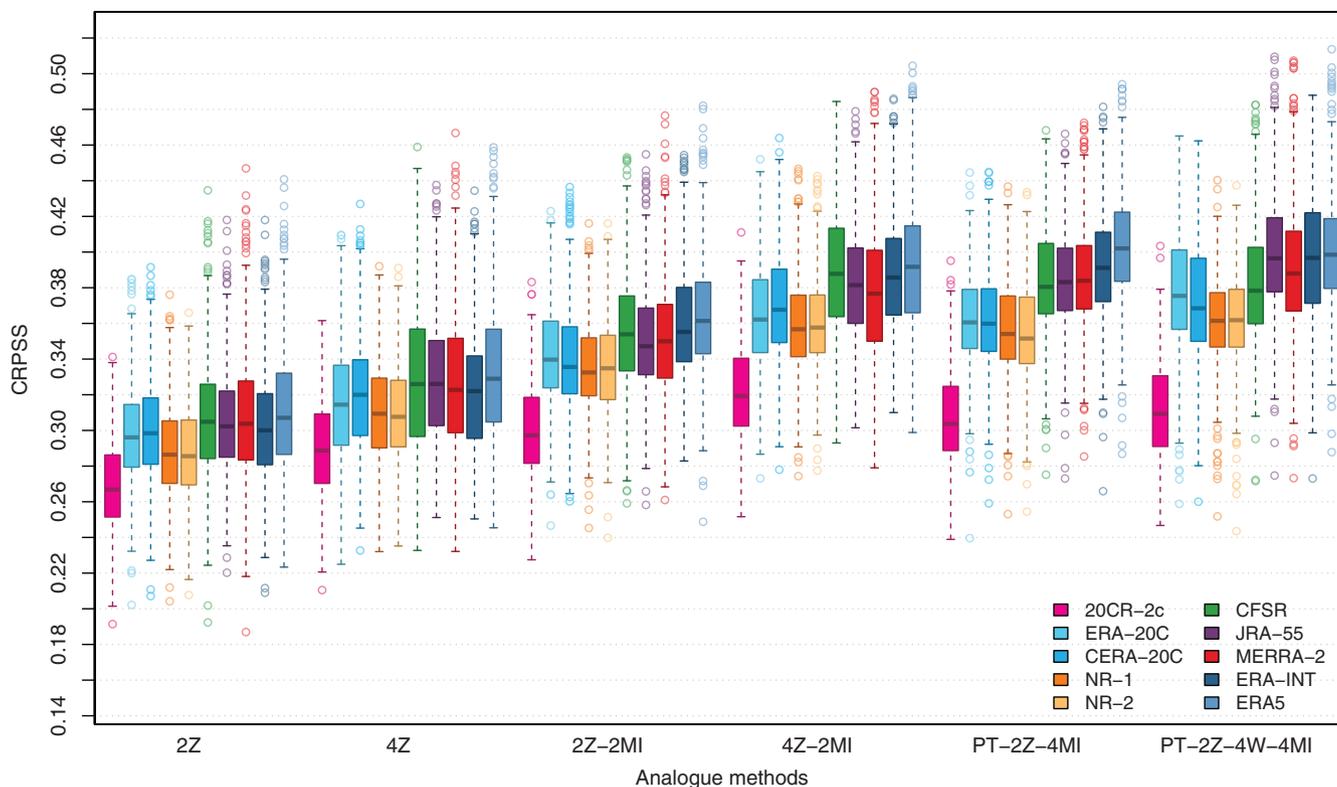


FIGURE 6 CRPSS for all stations and all considered AMs and reanalysis datasets on the VP. A higher CRPSS means better performance. The parameters of the AMs were calibrated for every station, dataset, and method. The boxes show the 25th, 50th, and 75th percentiles. The whiskers extend to the most extreme data point, which is no more than 1.5 times the interquartile range [Colour figure can be viewed at wileyonlinelibrary.com]

relative to ERA-INT are thus likely due to improvements in the forecast model, assimilation scheme, and assimilated data. While this conclusion is likely valid for predictands principally driven by large-scale predictors, it cannot be transposed to all applications. Indeed, when the phenomenon at hand is mainly driven by mesoscale predictors such as lightning activity, the output spatial resolution does matter (unpublished results).

3.2 | Comparing resulting parameters across reanalyses

Every AM has been calibrated independently for each reanalysis to be optimal for each configuration. In consequence, the resulting parameters differ from one setup to the other. The parameters that were calibrated here are the number of analogues (Figure 3) and the spatial windows on which the predictor fields are compared (Figure 4 for the zonal extent and Figure 5 the meridional extent). Substantial differences can be observed among the resulting parameters.

The number of analogues varies non-homogeneously across methods and datasets (Figure 3). For methods with

a single level of analogy (2Z and 4Z), the three reanalyses with the lowest spatial resolution and the poorest prediction skill (Section 3.3) are characterized by a higher number of analogues. The other datasets show similar distribution. A higher number of analogues can be related to the size of the spatial window and thus the spatial resolution, but it was also shown to be linked to the skill of the prediction and thereby the quality of the dataset (Horton *et al.*, 2017b). However, this cannot be transposed to all levels of analogy of more complex methods as interactions between these levels result in more complex behaviour. For example, while there is no clear trend for the number of analogues for the atmospheric circulation (first level of analogy) in 2Z-2MI and 4Z-2MI, PT-2Z-4MI shows an inverse trend with more skilled datasets being related to a higher number of analogues on the first level of analogy (Z). This method emphasizes more moisture predictors, with four predictors being used in the second level of analogy. A likely reason for such an inverted tendency is better moisture variables. It is important to remember here that the number of analogues at each level of analogy is optimized jointly at the end of the calibration procedure by assessing multiple combinations. Thus, datasets with

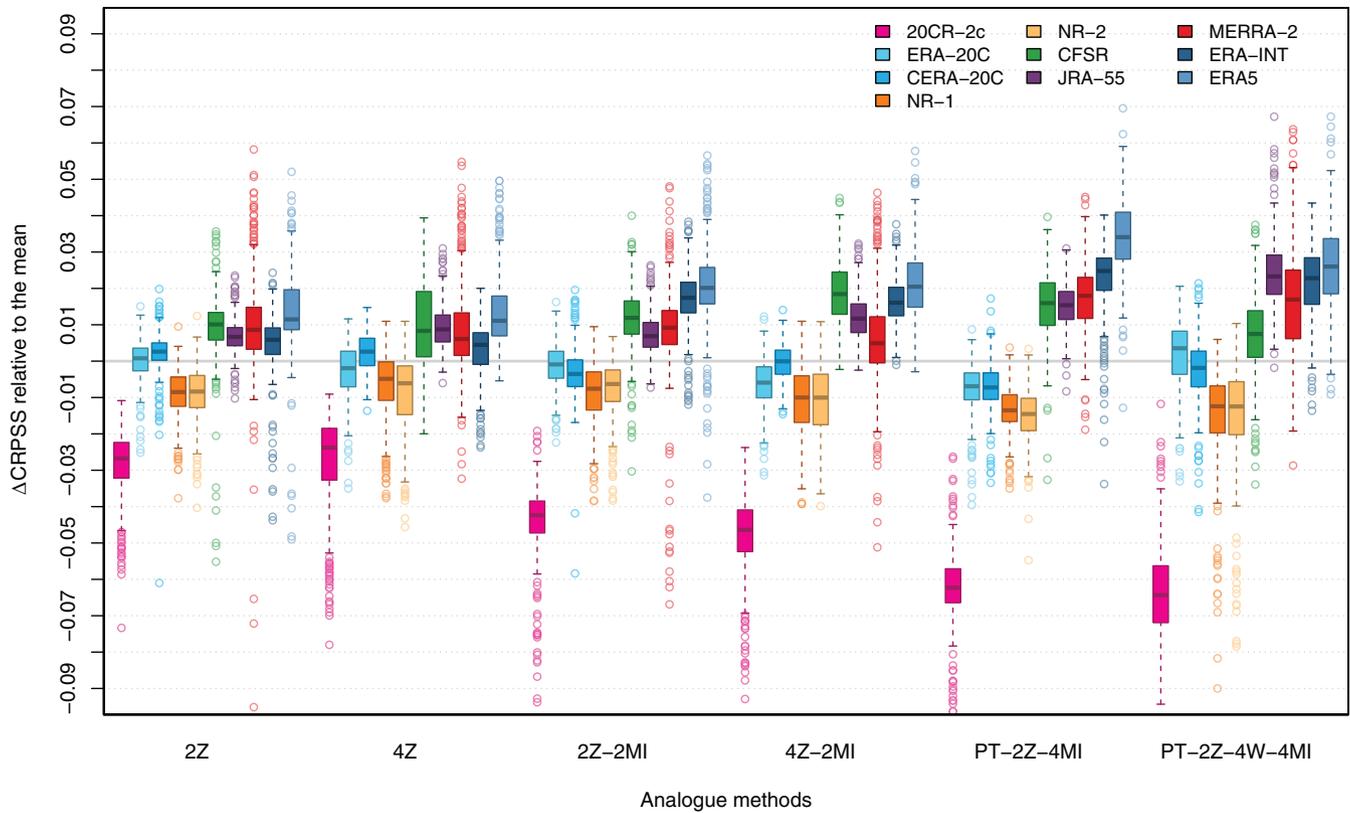


FIGURE 7 Impact of the reanalysis dataset on performance, isolated by processing the improvement in CRPSS for one dataset compared to the mean performance on all datasets, per station and method. Note that the methods cannot be compared here, only the datasets. Same conventions as Figure 6 [Colour figure can be viewed at wileyonlinelibrary.com]

better moisture variables in the second level of analogy benefit from discriminating the analogue candidates by making more use of these rather than of circulation variables, and thus keep a higher number of analogues from the first level of analogy. The same pattern can be observed with PT-2Z-4W-4MI, the most complex method, for Z and even more so for W. These selections of larger samples of analogue dates based on circulation variables correspond to datasets with higher spatial resolution and better skill. However, this happens only when the method has multiple levels of analogy. This suggests that parameters might not be optimal when transferred from one reanalysis to the other, and that one should be careful when manually adapting parameters since the direction of change can be counter-intuitive. Moreover, some distributions are highly spread out, which makes the generalization of parameter values more difficult. However, the number of analogue distributions for the last level of analogy are unimodal and less spread out than for prior levels of analogy.

The zonal extents of the spatial windows (Figure 4) also show substantial differences among datasets depending on the AM. For circulation-only AMs (2Z and 4Z), the zonal extent tends to decrease with an

increase in the spatial resolution of the reanalysis and even more so when four geopotential fields are considered (4Z). Further, all spatial windows tend to be substantially smaller for 4Z than for 2Z, with mean values halved for some datasets. Unlike the number of analogues, the spatial windows were calibrated sequentially for each level of analogy without posterior re-calibrations. This results in distributions of the zonal extent of geopotential height fields that are relatively similar between the first level of analogy of more complex methods and AMs based on Z only. Some variations can be observed, which are due to different initial numbers of analogues at the start of the calibration procedure for 2Z-2MI and 4Z-2MI as well as different preselection of the candidate analogues for PT-2Z-4MI and PT-2Z-4W-4MI (based on temperature instead of a calendar preselection). This similarity of the spatial windows among AMs does not hold anymore when using a global optimization technique that optimizes all parameters of the different levels of analogy jointly, since spatial windows of the circulation variables are reduced when moisture variables are used on a second level of analogy (Horton *et al.*, 2018). Similar zonal extents among the methods, in this case, do not correspond to an optimal but to a

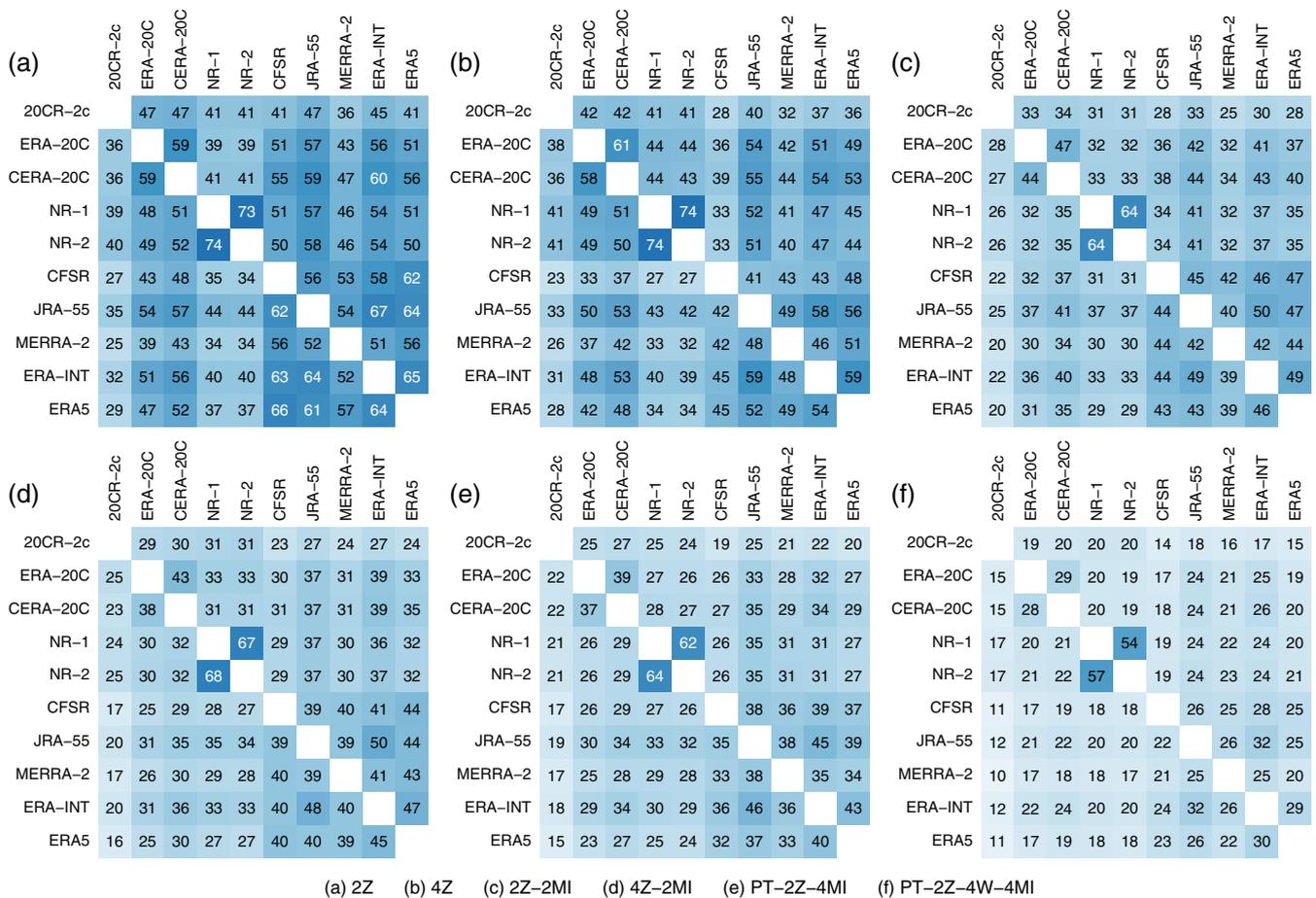


FIGURE 8 Percentage of identical analogue dates for different AMs, selected when using the reanalyses in columns that were also found when using the reanalyses in rows. The values were averaged for all stations on the VP [Colour figure can be viewed at wileyonlinelibrary.com]

limitation of the sequential calibration procedure that can only be overcome by a global optimization technique.

Moisture variables and vertical velocity are local variables that are characterized by smaller spatial windows (Ben Daoud *et al.*, 2016; Horton *et al.*, 2018). Their zonal extent decreases with the increase of the spatial resolution of the reanalysis, except for MERRA-2, when four moisture variables are used (the reason has not been investigated). Additionally, the zonal extents of W are noticeably smaller than those of moisture variables for reanalyses with high resolution—CFSR and ERA5. Finally, while the distributions of the zonal extents are mostly unimodal for W and MI, higher resolution datasets show multimodal distributions for Z. Here also, the differences in the optimal spatial windows for the different reanalyses raise caution about the transferability of parameters among datasets. Horton (2019a) showed that the spatial windows are better off being too large rather than too small. It is thus less critical to transfer larger spatial windows from a low-resolution dataset to a

dataset with higher spatial resolution rather than the opposite.

The meridional extents of the spatial windows (Figure 5) show similarity with the zonal extents in that MI and W are characterized by smaller extents than Z, with W being assessed on very small windows for CFSR and ERA5, while MERRA-2 diverges from the trend when four moisture variables are used. In contrast to the zonal extents, the meridional extents for Z do not show a clear trend with the spatial resolution. Also, when using more geopotential heights (from 2Z to 4Z), the meridional extent is not substantially reduced for most reanalyses which is contrary to the zonal extent. The meridional extents show distributions that differ among datasets with no unique optimal range for Z.

The frequency of spatial windows for Z with very small zonal or meridional extents is relatively high, even more for higher resolution datasets. However, for most stations and reanalyses, these small values do not co-occur that is, they represent narrow and elongated spatial windows rather than small squared domains (not

shown). Furthermore, when one dimension is small (zonal or meridional), the other can be located in the upper part of the distribution. For example, the spatial windows with a minimal meridional extent for 4Z have a zonal extent mostly between 7° and 10° for ERA5. This is even more pronounced for PT-2Z-4W-4MI where almost all the meridional extents above 8° correspond to very small zonal extents. However, the opposite is not true for this method as the very small meridional extents correspond to zonal extents between 5° and 8° that is, it is below the average. Three datasets (CERA-20C, ERA-INT, and MERRA-2) show relatively small extents jointly in both directions for Z in PT-2Z-4W-4MI. The parameters for these methods were calibrated using the classic calibration instead of the classic+ version (Section 2.4) and they may represent local non-optimal solutions.

As for the thin zonally or meridionally elongated spatial windows, we cannot exclude that they result from limitations of the classic/classic+ calibration techniques. However, it is interesting to note that the use of a global optimization technique (genetic algorithms) also resulted in some narrow and zonally elongated spatial windows (Horton *et al.*, 2018). Additionally, there were spatial patterns among the corresponding station locations, showing some consistency with the different climatic regions in Switzerland. Here, depending on the reanalysis, the zonally elongated spatial windows are dominant for stations located in the southern part of Switzerland, where southerly circulation is responsible for the most extreme precipitation events. Meridionally elongated spatial windows are prevalent in regions where westerlies are associated with heavy rainfall. While some optimal spatial windows for processing the gradients of the geopotential height are relatively large and rectangular, the relevance of narrow and elongated spatial windows perpendicular to the flux associated with heavy precipitation cannot be discarded.

3.3 | How ERA5 compares to other reanalyses

The skill scores of the different methods and reanalyses are shown in Figure 6. As discussed in Horton and Brönnimann (2019), there are clear differences among not only the methods, as more complex AMs generally perform better, but also the reanalyses, in favour of modern full-input reanalyses. ERA5 is here systematically one of the best performing reanalyses. When isolating the impact of the reanalysis by removing the influence of the methods and the stations (Figure 7), the positive contribution of ERA5 to the skill score is even more clear. ERA5 shows superior skills for 2Z and relatively similar skills as CFSR, which has the second-highest spatial

resolution, for 4Z. Overall, the impact of the reanalyses tends to increase with the method complexity. While ERA-INT was already performing well for complex methods, ERA5 showed even more skill. Thus, it surpassed all other reanalyses for 2Z-2MI and showed the largest difference for the method PT-2Z-4MI. Like for 4Z, it is on par with CFSR for the method 4Z-2MI. Interestingly, JRA-55 showed similar skills as ERA5 for the most complex method, PT-2Z-4W-4MI, even though JRA-55 has an output resolution that is five times lower (1.25°) than ERA5.

ERA5 was selected as the best reanalysis for about 50% of the stations on average, with a minimum of 34.9% for 4Z and a maximum of 78.7% for PT-2Z-4MI. When considering the first two ranks, ERA5 was selected for 75% of the stations on average, with a minimum of 61.8% for PT-2Z-4W-4MI and a maximum of 91.7% for PT-2Z-4MI. There were no specific spatial patterns as to which stations performed best with ERA5, as was the case with the other reanalyses in Horton and Brönnimann (2019). Even though ERA5 did not systematically supersede ERA-INT for all stations, its performance is overall superior and, hence, we would recommend using it in AMs. Moreover, with ERA-INT being discontinued, ERA5 is the obvious choice in most applications.

The use of predictors from different datasets with distinct spatial windows results in a different selection of analogue dates. The percentage of identical analogue dates for the different AMs over the VP is shown in Figure 8. This percentage decreases with the complexity of the AM and ranges from a maximum of 74% between NR-1 and NR-2 for the 2Z method, to a minimum of 10% between MERRA-2 and 20CR-2c for the PT-2Z-4W-4MI method. If all the reanalyses were compared over the same spatial windows, these numbers would likely be higher.

As it was expected, ERA5 shared some similarities in the selection of analogue dates with ERA-INT, but not to an extent that is significantly different from the other modern full-input reanalyses. Indeed, the percentage of shared analogue dates is of the same order of magnitude as with CFSR and JRA-55, which are the two reanalyses produced by models with a high spatial resolution (Table 1) but fewer vertical levels. It is worth noting that JRA-55 presents strong similarities with ERA5 in terms of the selected analogue dates for Switzerland, which is consistent with their relatively similar performance, mainly for PT-2Z-4W-4MI.

4 | CONCLUSIONS

The choice of a reanalysis has a significant impact on the performance of AMs in the perfect prognosis context, as

has already been shown in the work of Dayon *et al.* (2015) and Horton and Brönnimann (2019) for the European region. The impact of the dataset on AM skills can be higher than that of the choice of the predictor variables. Horton and Brönnimann (2019) concluded that, as a general rule, the latest version of the full-input reanalyses should be used as they contribute to better skill. The later release of ERA5 then raised questions of its potential benefits over ERA-INT and the extent to which the high spatial resolution could be informative for statistical downscaling.

This work focused on comparing ERA5 to the other main global reanalyses and showed that it should be the dataset of choice for all tested AMs, since it often surpassed the other reanalyses. ERA5 was selected as the best dataset by about 50% of the stations and was one of the top two reanalyses for 75% of them. Even though ERA5 might not systematically be the best reanalysis for all stations, it has constantly featured among the best and thus should be used for AMs in Switzerland and across Europe. However, for certain applications, it would be advisable to use a selection of varied reanalyses to account for their relative uncertainties, rather than a single one.

ERA5 is provided at high spatial (0.25°) and temporal (1 hr) resolutions. Although AMs are computationally light approaches, their calibration over several decades can be time-consuming. The time consumption increases with a higher spatial resolution of the predictor grids. The high resolution even turned into a trap for the calibration technique that ended up stuck in local solutions, resulting in poor skills. When using high-resolution predictor grids, one should ensure that the calibration technique used can satisfactorily optimize the spatial windows. The classic+ calibration (Horton, 2019a), in this case, was able to successfully calibrate the spatial windows, thanks to additional moves that allow going out of local solutions.

We also showed that the optimal spatial windows and number of analogues differ across reanalyses. The number of analogues is often related to the size of the spatial window and thereby the spatial resolution as well as the quality of the dataset (or skill of the method). A lower number of analogues in the last level of analogy is usually associated with a better skill of the reanalysis. However, the opposite can be observed for intermediate levels of analogy, where a higher number of analogues can be selected from circulation variables to leverage better moisture variables in a subsequent level of analogy. Thus, one should take care when manually selecting the number of analogues for the best solution can be counter-intuitive.

The spatial windows were shown to vary substantially among reanalyses. Spatial windows generally tend to

decrease with an increase in the spatial resolution of the reanalysis. Narrow and elongated spatial windows were sometimes observed for the geopotential height at some stations when using datasets with a high spatial resolution. These spatial windows differ from the traditional ones with a rather square shape. Although a limitation of the calibration procedure cannot be excluded, the orientation of these windows perpendicular to the flux associated with heavy precipitation for the given stations suggests a relevance that cannot be discarded either. Overall, the differences in the optimal spatial windows and the number of analogues across reanalyses raise caution about the transferability of parameters among datasets.

Horton and Brönnimann (2019) showed that using reanalyses at a spatial resolution inferior to 1° did not significantly improve the performance score. ERA5 was here used at its original resolution as well as that of ERA-INT (0.75°). The reduction in spatial resolution did not impact the performance of AMs for daily precipitation, which is mainly driven by synoptic-scale processes. This conclusion was shown to be not valid for predictands driven by mesoscale drivers such as the lightning activity. It also implies that the output resolution does not explain the gains in skill compared to ERA-INT, which are likely due to improvements in the forecast model, the assimilation scheme, and the assimilated data.

To conclude, ERA5 is recommended for the prediction of daily precipitation with AMs in Europe. It should ideally be used along with other reanalyses in a multi-model approach. ERA5 was shown to be systematically among the best options for each AM and thus, there is no need for a predictor-conditioned selection as in Horton and Brönnimann (2019). We recommend using 20CR for applications starting between 1851 and 1900, CERA-20C for those between 1900 and 1950, and ERA5 for applications after 1950. However, for predictands primarily driven by synoptic-scale processes, using ERA5 at its full resolution is not justified and might even be counterproductive if the calibration method does not handle it well.

ACKNOWLEDGEMENTS

Precipitation time series were provided by MeteoSwiss. The NCEP/NCAR, NCEP/DOE, and 20CR-2c were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, at <http://www.esrl.noaa.gov/psd/>. Support for the Twentieth Century Reanalysis Project dataset is provided by the US Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. The CFSR and JRA-55 were

obtained from the CISL Research Data Archive (<http://rda.ucar.edu/>) at NCAR in Boulder, Colorado, and the NCAR is supported by grants from the National Science Foundation. The Climate Forecast System Reanalysis (CFSR) project is carried out by the Environmental Modeling Center (EMC), National Centers for Environmental Prediction (NCEP). The Japanese 55-year Reanalysis (JRA-55) project is carried out by the Japan Meteorological Agency (JMA). The MERRA-2 was obtained from the Goddard Earth Sciences Data and Information Services Center, Greenbelt, Maryland, from their website at <http://disc.sci.gsfc.nasa.gov/mdisc>. ERA-interim, ERA-20C, and CERA-20C were obtained from the ECMWF Data Server at <http://apps.ecmwf.int/datasets/>. ERA5 was obtained from the C3S climate data store (CDS) at <https://cds.climate.copernicus.eu>. Calculations were performed on UBELIX (<http://www.id.unibe.ch/hpc>), the HPC cluster at the University of Bern. Thanks to C. Obled for correcting the manuscript.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

All calculations were performed with the open-source AtmoSwing software v2.1.1 (Horton, 2019b). The resulting files were processed using AtmoSwing R-toolbox v1.2.0 (Horton and Burkart, 2018). The resulting analogue dates for every combination of station, reanalysis, and analogue method have been published. These archives also contain different files: the parameter files used in AtmoSwing for the calibration, the resulting calibrated parameters, and the files listing all assessed parameter sets. These files are available for ERA5 (Horton, 2021) and the other reanalyses (see references in Horton and Brönnimann, 2019).

ORCID

Pascal Horton  <https://orcid.org/0000-0003-0466-0359>

REFERENCES

- Alessandrini, S., Delle Monache, L., Sperati, S. and Cervone, G. (2015a) An analog ensemble for short-term probabilistic solar power forecast. *Applied Energy*, 157, 95–110.
- Alessandrini, S., Delle Monache, L., Sperati, S. and Nissen, J.N. (2015b) A novel application of an analog ensemble for short-term wind power forecasting. *Renewable Energy*, 76, 768–781.
- Altava-Ortiz, V., Barrera, A., Llasat, M.C., Prat, M.A., Gibergans-Báguena, J. and Barnolas, M. (2006) Application of the MM5 and the analogous method to heavy rainfall event, the case of 16–18 October 2003 in Catalonia (NE Spain). *Advances in Geosciences*, 7, 313–319.
- Ben Daoud, A., Sauquet, E., Bontron, G., Obled, C. and Lang, M. (2016) Daily quantitative precipitation forecasts based on the analogue method: improvements and application to a French large river basin. *Atmospheric Research*, 169, 147–159.
- Ben Daoud, A., Sauquet, E., Lang, M., Obled, C. and Bontron, G. (2009) La prévision des précipitations par recherche d'analogues: état de l'art et perspectives. *La Houille Blanche*, 95, 60–65.
- Bessa, R., Trindade, A., Silva, C.S. and Miranda, V. (2015) Probabilistic solar power forecasting in smart grids using distributed information. *International Journal of Electrical Power & Energy Systems*, 72, 16–23.
- Bliefernicht, J. (2010) *Probability forecasts of daily areal precipitation for small river basins*. Ph.D. thesis, Universität Stuttgart.
- Bonnet, R., Boé, J., Dayon, G. and Martin, E. (2017) 20th-century hydro-meteorological reconstructions to study the multidecadal variations of the water cycle over France. *Water Resources Research*, 53, 1–17.
- Bontron, G. (2004) *Prévision quantitative des précipitations: Adaptation probabiliste par recherche d'analogues. Utilisation des Réanalyses NCEP/NCAR et application aux précipitations du Sud-Est de la France*. Ph.D. Thesis, Institut National Polytechnique de Grenoble.
- Bontron, G. and Obled, C. (2005) L'adaptation probabiliste des prévisions météorologiques pour la prévision hydrologique. *La Houille Blanche*, 1, 23–28.
- Bradley, A.A. and Schwartz, S.S. (2011) Summary verification measures and their interpretation for ensemble forecasts. *Monthly Weather Review*, 139, 3075–3089.
- Brigode, P., Brissette, F., Nicault, A., Perreault, L., Kuentz, A., Mathevet, T. and Gailhard, J. (2016) Stream flow variability over the 1881–2011 period in northern Québec: comparison of hydrological reconstructions based on tree rings and geopotential height field reanalysis. *Climate of the Past*, 12, 1785–1804.
- Brown, B.G., Gilleland, E. and Ebert, E.E. (2012) Forecasts of spatial fields. In: Jolliffe, I.T. and Stephenson, D.B. (Eds.) (chap. 6) *Forecast Verification*. Chichester, West Sussex, UK: Wiley, pp. 95–117.
- Brown, T. (1974) *Admissible scoring systems for continuous distributions*. Technical Report URL: <http://eric.ed.gov/?id=ED135799>.
- Caillouet, L., Vidal, J.-P., Sauquet, E. and Graff, B. (2016) Probabilistic precipitation and temperature downscaling of the twentieth century reanalysis over France. *Climate of the Past*, 12, 635–662.
- Chardon, J., Hingray, B., Favre, A.-C., Autin, P., Gailhard, J., Zin, I. and Obled, C. (2014) Spatial similarity and transferability of analog dates for precipitation downscaling over France. *Journal of Climate*, 27, 5056–5074.
- Charles, A., Timbal, B., Fernandez, E. and Hendon, H. (2012) Analog downscaling of seasonal rainfall forecasts in the Murray Darling Basin. *Monthly Weather Review*, 141, 1099–1117.
- Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R.I., Grant, A.N., Groisman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, Ø., Ross, T.F., Trigo, R.M., Wang, X.L., Woodruff, S.D. and Worley, S.J. (2011) The twentieth century reanalysis

- project. *Quarterly Journal of the Royal Meteorological Society*, 137, 1–28.
- Dayon, G., Boé, J. and Martin, E. (2015) Transferability in the future climate of a statistical downscaling method for precipitation in France. *Journal of Geophysical Research: Atmospheres*, 120, 1023–1043.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E. V., Isaksen, I., Kallberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N. and Vitart, F. (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.
- Delle Monache, L., Eckel, F.A., Rife, D.L., Nagarajan, B. and Searight, K. (2013) Probabilistic weather prediction with an analog ensemble. *Monthly Weather Review*, 141, 3498–3516.
- Delle Monache, L., Nipen, T., Liu, Y., Roux, G. and Stull, R. (2011) Kalman filter and analog schemes to postprocess numerical weather predictions. *Monthly Weather Review*, 139, 3554–3570.
- Foresti, L., Panziera, L., Mandapaka, P.V., Germann, U. and Seed, A. (2015) Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall. *Meteorological Applications*, 22, 141–155.
- Fraedrich, K., Raible, C.C. and Sielmann, F. (2003) Analog ensemble forecasts of tropical cyclone tracks in the Australian region. *Weather and Forecasting*, 18, 3–11.
- Fujiwara, M., Wright, J.S., Manney, G.L., Gray, L.J., Anstey, J., Birner, T., Davis, S., Gerber, E.P., Lynn Harvey, V., Hegglin, M. I., Homeyer, C.R., Knox, J.A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B.M., Santee, M.L., Tegtmeier, S., Chabrillat, S., Tan, D.G., Jackson, D.R., Polavarapu, S., Compo, G.P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J.S. and Zou, C.Z. (2017) Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmospheric Chemistry and Physics*, 17, 1417–1452.
- Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M. and Zhao, B. (2017) The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30, 5419–5454.
- Hamill, T. and Whitaker, J. (2006) Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application. *Monthly Weather Review*, 134, 3209–3229.
- Hamill, T.M., Scheuerer, M. and Bates, G.T. (2015) Analog probabilistic precipitation forecasts using GEFS reforecasts and climatology-calibrated precipitation analyses. *Monthly Weather Review*, 143, 3300–3309.
- Harada, Y., Kamahori, H., Kobayashi, C., Endo, H., Kobayashi, S. and Ota, Y. (2016) The JRA-55 reanalysis: representation of atmospheric circulation and climate variability. *Journal of the Meteorological Society of Japan*, 94, 269–302.
- Hersbach, H. (2000) Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15, 559–570.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.N. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.
- Horton, P. (2019a) AtmoSwing: analog technique model for statistical weather forecastING and downscaling (v2.1.0). *Geoscientific Model Development*, 12, 2915–2940.
- Horton, P. (2019b) AtmoSwing v2.1.1.
- Horton, P. (2021) Analog dates for ERA5 at 301 precipitation stations in Switzerland (1981–2010). Mendeley Data, v1.
- Horton, P. and Brönnimann, S. (2019) Impact of global atmospheric reanalyses on statistical precipitation downscaling. *Climate Dynamics*, 52, 5189–5211.
- Horton, P. and Burkart, K. (2018) AtmoSwing R tools.
- Horton, P., Jaboyedoff, M., Metzger, R., Obled, C. and Marty, R. (2012) Spatial relationship between the atmospheric circulation and the precipitation measured in the western Swiss Alps by means of the analogue method. *Natural Hazards and Earth System Sciences*, 12, 777–784.
- Horton, P., Jaboyedoff, M. and Obled, C. (2017a) Global optimization of an analog method by means of genetic algorithms. *Monthly Weather Review*, 145, 1275–1294.
- Horton, P., Jaboyedoff, M. and Obled, C. (2018) Using genetic algorithms to optimize the analogue method for precipitation prediction in the Swiss Alps. *Journal of Hydrology*, 556, 1220–1231.
- Horton, P., Obled, C. and Jaboyedoff, M. (2017b) The analogue method for precipitation prediction: finding better analogue situations at a sub-daily time step. *Hydrology and Earth System Sciences*, 21, 3307–3323.
- Junk, C., Delle Monache, L. and Alessandrini, S. (2015a) Analog-based ensemble model output statistics. *Monthly Weather Review*, 143, 2909–2917.
- Junk, C., Delle Monache, L., Alessandrini, S., Cervone, G. and von Bremen, L. (2015b) Predictor-weighting strategies for probabilistic wind power forecasting with an analog ensemble. *Meteorologische Zeitschrift*, 24, 361–379.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. (1996) The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437–471.
- Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J.J., Fiorino, M. and Potter, G.L. (2002) NCEP–DOE AMIP-II reanalysis (R-2). *Bulletin of the American Meteorological Society*, 83, 1631–1643.

- Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., Van Den Dool, H., Jenne, R. and Fiorino, M. (2001) The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, 82, 247–267.
- Kobayashi, C., Endo, H., Ota, Y., Kobayashi, S., Onoda, H., Harada, Y., Onogi, K. and Kamahori, H. (2014) Preliminary results of the JRA-55C, an atmospheric reanalysis assimilating conventional observations only. *Solaia*, 10, 78–82.
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K. and Takahashi, K. (2015) The JRA-55 reanalysis: general specifications 540 and basic characteristics. *Journal of the Meteorological Society of Japan. Series II*, 93, 5–48.
- Koukidis, E. and Berg, A. (2009) Sensitivity of the Statistical Down-Scaling Model (SDSM) to reanalysis products. *Atmosphere-Ocean*, 47, 1–18.
- Kruizinga, S. and Murphy, A. (1983) Use of an analogue procedure to formulate objective probabilistic temperature forecasts in The Netherlands. *Monthly Weather Review*, 111, 2244–2254.
- Kuentz, A., Mathevet, T., Gailhard, J. and Hingray, B. (2015) Building long-term and high spatio-temporal resolution precipitation and air temperature reanalyses by mixing local observations and global atmospheric reanalyses: the ANATEM model. *Hydrology and Earth System Sciences*, 19, 2717–2736.
- Laloyaux, P., Balmaseda, M., Dee, D., Mogenssen, K. and Janssen, P. (2016) A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 142, 65–78.
- Laloyaux, P., de Boissesson, E., Balmaseda, M., Bidlot, J.R., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D., Haimberger, L., Hersbach, H., Kosaka, Y., Martin, M., Poli, P., Rayner, N., Rustemeier, E. and Schepers, D. (2018) CERA-20C: a coupled reanalysis of the twentieth century. *Journal of Advances in Modeling Earth Systems*, 10, 1172–1195.
- Lorenz, E. (1969) Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*, 26, 636–646.
- Maraun, D., Wetterhall, F., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W., Sauter, T., Themeßl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M. and Thiele-Eich, I. (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, 48, 1–34.
- Marty, R., Zin, I., Obled, C., Bontron, G. and Djerboua, A. (2012) Toward real-time daily PQPF by an analog sorting approach: application to flash-flood catchments. *Journal of Applied Meteorology and Climatology*, 51, 505–520.
- Matheson, J. and Winkler, R. (1976) Scoring rules for continuous probability distributions. *Management Science*, 22, 1087–1096.
- Matulla, C., Zhang, X., Wang, X.L., Wang, J., Zorita, E., Wagner, S. and von Storch, H. (2007) Influence of similarity measures on the performance of the analog method for downscaling daily precipitation. *Climate Dynamics*, 30, 133–144.
- Panziera, L., Germann, U., Gabella, M. and Mandapaka, P.V. (2011) NORA-nowcasting of orographic rainfall by means of analogues. *Quarterly Journal of the Royal Meteorological Society*, 137, 2106–2123.
- Poli, P., Hersbach, H., Dee, D.P., Berrisford, P., Simmons, A.J., Vitart, F., Laloyaux, P., Tan, D.G.H., Peubey, C., Thépaut, J.N., Trémolet, Y., Hólm, E.V., Bonavita, M., Isaksen, L. and Fisher, M. (2016) ERA-20C: an atmospheric reanalysis of the twentieth century. *Journal of Climate*, 29, 4083–4097.
- Radanovics, S., Vidal, J.-P., Sauquet, E., Ben Daoud, A. and Bontron, G. (2013) Optimising predictor domains for spatially coherent precipitation downscaling. *Hydrology and Earth System Sciences*, 17, 4189–4208.
- Raynaud, D., Hingray, B., Zin, I., Anquetin, S., Debionne, S. and Vautard, R. (2016) Atmospheric analogues for physically consistent scenarios of surface weather in Europe and Maghreb. *International Journal of Climatology*, 37, 2160–2176.
- Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M. and Woollen, J. (2011) MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, 24, 3624–3648.
- Rummukainen, M. (1997) Methods for statistical downscaling of GCM simulations. Technical Report 80. <http://agris.fao.org/agris-search/search/display.do?f=2012/OV/OV201205379005379.xml;SE19970167745>
- Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.T., Chuang, H.Y., Juang, H.M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Van Den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G. and Goldberg, M. (2010) The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91, 1015–1057.
- Shao, Q. and Li, M. (2013) An improved statistical analogue downscaling procedure for seasonal precipitation forecast. *Stochastic Environmental Research and Risk Assessment*, 27, 819–830.
- Sievers, O., Fraedrich, K. and Raible, C.C. (2000) Self-adapting analogue ensemble predictions of tropical cyclone tracks. *Weather and Forecasting*, 15, 623–629.
- Slivinski, L.C., Compo, G.P., Sardeshmukh, P.D., Whitaker, J.S., McColl, C., Allan, R.J., Brohan, P., Yin, X., Smith, C.A., Spencer, L.J., Vose, R.S., Rohrer, M., Conroy, R.P., Schuster, D.C., Kennedy, J.J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T.A., Domínguez-Castro, F., Freeman, J.E., Gergis, J., Hawkins, E., Jones, P.D., Kubota, H., Lee, T.C., Lorrey, A.M., Luterbacher, J., Mock, C.J., Przybylak, R. K., Pudmenzky, C., Slonosky, V.C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., Wood, K. and Wyszyński, P. (2021) An evaluation of the performance of the twentieth century reanalysis version 3. *Journal of Climate*, 34, 1417–1438.
- Slivinski, L.C., Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Giese, B.S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L.J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Domínguez-Castro, F., Freeman, J.E.,

- Gergis, J., Hawkins, E., Jones, P.D., Jourdain, S., Kaplan, A., Kubota, H., Blancq, F.L., Lee, T.C., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C.J., Moore, G.W., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V.C., Smith, C.A., Tinz, B., Trewin, B., Valente, M.A., Wang, X.L., Wilkinson, C., Wood, K. and Wyszyński, P. (2019) Towards a more reliable historical reanalysis: improvements for version 3 of the twentieth century reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 145, 2876–2908.
- Teweles, S. and Wobus, H.B. (1954) Verification of prognostic charts. *Bulletin of the American Meteorological Society*, 35, 455–463.
- Thiébaux, H.J. (1985) On approximations to geopotential and wind-field correlation structures. *Tellus A*, 37 A, 126–131.
- Timbal, B., Dufour, A. and Mcavane, B. (2003) An estimate of future climate change for western France using a statistical downscaling technique. *Climate Dynamics*, 20, 807–823.
- Uppala, S.M., Kallberg, P.W., Simmons, A.J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A. E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P. and Woollen, J. (2005) The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131, 2961–3012.
- Vanvyve, E., Delle Monache, L., Monaghan, A.J. and Pinto, J.O. (2015) Wind resource estimates with an analog ensemble approach. *Renewable Energy*, 74, 761–773.
- Wetterhall, F. (2005) *Statistical downscaling of precipitation from large-scale atmospheric circulation: comparison of methods and climate regions*. Ph.D. thesis, Uppsala University.
- Wetterhall, F., Halldin, S. and Xu, C.-Y. (2007) Seasonality properties of four statistical-downscaling methods in Central Sweden. *Theoretical and Applied Climatology*, 87, 123–137.
- Willems, P. and Vrac, M. (2011) Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change. *Journal of Hydrology*, 402, 193–205.
- Wu, W., Liu, Y., Ge, M., Rostkier-Edelstein, D., Descombes, G., Kunin, P., Warner, T., Swerdlin, S., Givati, A., Hopson, T. and Yates, D. (2012) Statistical downscaling of climate forecast system seasonal predictions for the Southeastern Mediterranean. *Atmospheric Research*, 118, 346–356.
- Xavier, P.K. and Goswami, B.N. (2007) An analog method for real-time forecasting of summer monsoon subseasonal variability. *Monthly Weather Review*, 135, 4149–4160.
- Yiou, P., Boichu, M., Vautard, R., Vrac, M., Jourdain, S., Garnier, E., Fluteau, F. and Menut, L. (2014) Ensemble meteorological reconstruction using circulation analogues of 1781–1785. *Climate of the Past*, 10, 797–809.
- Zorita, E. and von Storch, H. (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of Climate*, 12, 2474–2489.

How to cite this article: Horton, P. (2021). Analogue methods and ERA5: Benefits and pitfalls. *International Journal of Climatology*, 1–19. <https://doi.org/10.1002/joc.7484>