



ERA4CS Joint Call on Researching and Advancing Climate Services Development – Topic B  
(GRANT AGREEMENT 689029)

**European Climate Observations, Monitoring and Services initiative (2)**  
**Milestone M3.2**

**High-quality climate prediction data available to WP4**



Deliverable Title	<i>High-quality climate prediction data available to WP4</i>	
WP number	3	
Lead Beneficiary	<i>Silvia Terzago, Giulio Bongiovanni and Jost von Hardenberg (CNR)</i>	
Contributors		
Deliverable Due Date	<i>15/09/2019 - Postponed to March 2020</i>	
Nature of the Deliverable		<i>R – Report</i>
		<i>SP – Software Package</i>
	D	<i>D – Dataset</i>
		<i>O - Other</i>

Version	Date	Modified by	Comments
1.0	31/03/2020	Silvia Terzago	Submitted version
1.1	27/05/2020	Silvia Terzago	Temperature QuantMap
1.2	22/09/2020	Silvia Terzago	k parameter corrected

## **Table of contents**

<b>Executive summary</b>	<b>4</b>
<b>Seasonal forecast model bias and bias-adjustment methods</b>	<b>4</b>
<b>Methods for spatial downscaling</b>	<b>6</b>
Downscaling of surface air temperature	6
Downscaling of precipitation: a stochastic approach	8
<b>The seasonal forecasts data archive</b>	<b>10</b>
<b>The downscaling of seasonal forecasts</b>	<b>11</b>
Downscaling of air temperature	11
Downscaling of precipitation	11
<b>Evaluation of the downscaled seasonal forecasts</b>	<b>13</b>
Downscaled air temperature	13
Downscaled precipitation	15
<b>Availability of the downscaled data</b>	<b>16</b>
<b>References</b>	<b>18</b>
<b>Appendix A – Code used for temperature downscaling</b>	<b>20</b>
<b>Appendix B – Code used for precipitation downscaling</b>	<b>24</b>

---

## *Executive summary*

---

Seasonal forecast model outputs available in the C3S Climate Data Store have global coverage and spatial resolution of  $1^\circ$ . This resolution is usually too coarse compared to the scales needed in practical applications but also to the typical scales at which the latest reanalysis products (i.e. ECMWF ERA5) are available.

The overall objective of our work has been to develop a downscaling chain employing the state-of-the-art techniques to disaggregate seasonal forecast model outputs to a target resolution of  $0.25^\circ$ , comparable to that of reanalyses. The final product of this initiative is the delivery of a high-quality dataset for precipitation and temperature variables over the Mediterranean area. All techniques employed in the downscaling chain are extensively documented in literature and implemented using functions from the R package CSTools (v3.0.0. or the forthcoming release). The downscaled model outputs are made publicly available, together with the R codes used to derive them.

The present milestone provides technical information on the procedure that has been followed to disaggregate seasonal forecast model outputs, on the repository where the downscaled data are made available and on the steps to follow to access them.

### **1. Seasonal forecast model bias and bias-adjustment methods**

Imperfect physical parameterizations and errors in the model initialization are often reflected in model bias. In order to use model outputs for applications, adjustment techniques are usually employed to mitigate the discrepancies between the model data and the observations.

A range of bias adjustment methods with different degrees of complexity has been developed. Basic bias-adjustment methods correct the mean value by adding a temporally constant offset, or by applying a correction factor to the model data. This additive or multiplicative constant corresponds to the average deviation of the model time series from the observed time series over the reference period.

A generalization of this method is the quantile mapping that adjusts the quantiles of the distribution of the modelled variable based on the distribution of observations. In a comparison of bias-correction methods for hydrological impact studies, bias-correction algorithms based on quantile mapping have been found to outperform simpler bias correction methods that correct only the mean or the mean and variance of precipitation series (Cannon et al. 2015).

In our downscaling chains we adjust temperature and precipitation seasonal forecasts with the quantile mapping method as defined in Gudmundsson et al. (2012). We use ERA5 at  $0.25^\circ$  upscaled at the same resolution as the seasonal forecast data, i.e.  $1^\circ$ , as a reference dataset. The Quantile Mapping-based corrections are calculated and applied month by month using the corresponding R function `CST_QuantileMapping()` available in CSTools (see Appendix A and B)

Figure 1 shows an example of application of the quantile mapping method to ECMWF S5 seasonal forecasts of air temperature and daily precipitation. The figure displays the annual cycle of the ECMWF S5 model variables calculated by averaging all ensemble members and all hindcast years with starting date May 1<sup>st</sup> compared to the annual cycle in the ERA5 reference dataset. Temperature is underestimated during the whole season, with similar biases in all months. The model is able to capture the temporal distribution of precipitation, however it overestimates precipitation in the summer months. The application of the quantile-mapping method to ECMWF S5 seasonal forecasts allows to adjust the average model bias and to improve the agreement with the ERA5 reference annual cycle for both variables. The quantile-mapping method is applied to all the seasonal forecast models to mitigate possible biases.

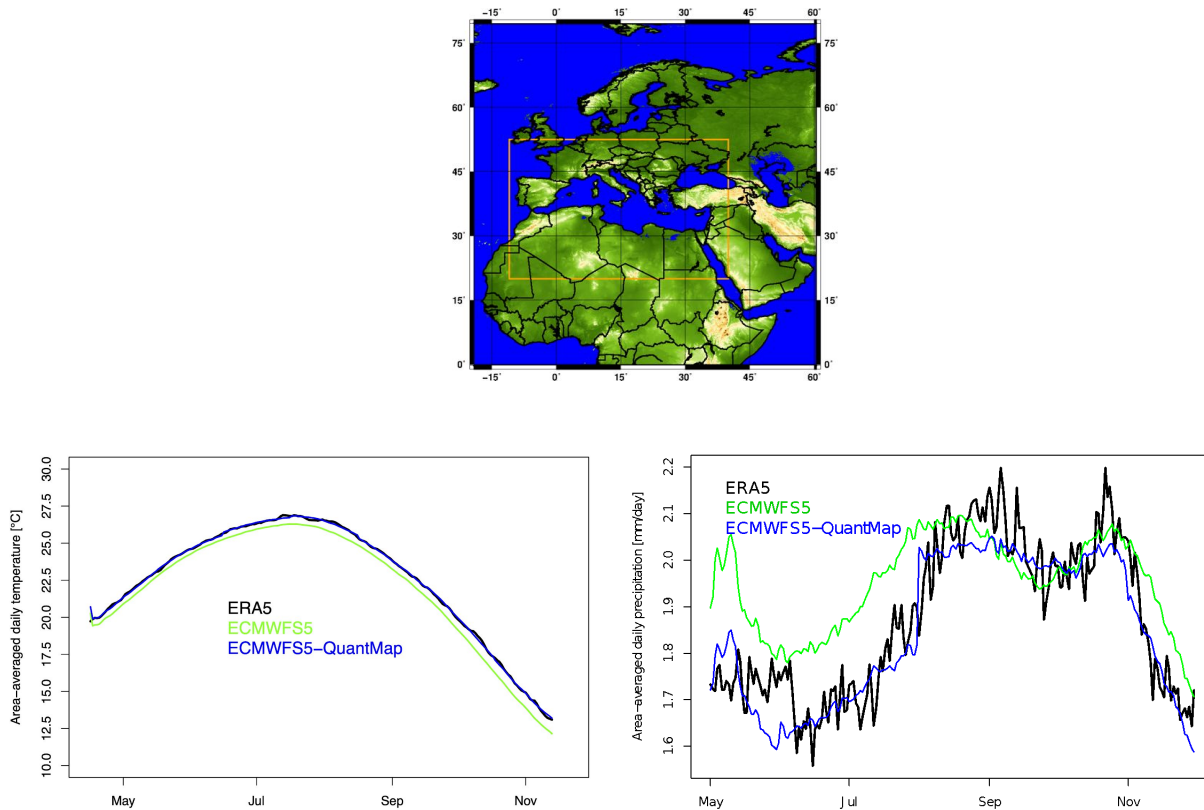


Figure 1 (top) The orange box represents the domain for which we produced downscaled temperature and precipitation data (11°W,40°E; 20°N,52.5°N) identified in the text as the “Mediterranean” domain. To avoid border effects the spatial downscaling of precipitation is initially performed over the “full” domain shown in the plot (19°W,60.5°E; 0°N,79.5°N) and then only the gridpoints included in the Mediterranean domain are retained; (bottom left) Daily air temperature averaged over the Mediterranean domain, (bottom right) daily precipitation averaged over the full domain for the ECMWF S5 model (green), the ECMWF S5 model after the application of the quantile mapping (blue) and the ERA5 reference datasets upscaled at the model resolution. Start date May 1<sup>st</sup>, hindcast years 1993-2016 for temperature and 1993-2018 for precipitation, respectively.

## 2. Methods for spatial downscaling

In order to bridge the gap between the scales of climate models and the scales needed for practical applications, downscaling methods can be employed. The typical spatial resolution of seasonal forecast models is on the order of 1° (about 100 km) both in longitude and in latitude. The spatial scales required for applications depend on the specific study but they can vary in a range between tens of km down to hundreds of meters. In the following we will briefly present two methods that can be applied to downscale two of the main climatic variables, surface air temperature and precipitation. These methods are documented in literature and we suggest to refer to the references reported in the text for further details.

### 2.1 Downscaling of surface air temperature

In coarse-scale surface air temperature datasets, such as those provided by seasonal forecast models, one of the main features which influences the temperature at the sub-grid scale is the topography. In fact, in mountain areas elevational gradients cause high spatial variability of air temperature over short distances. A simple method to downscale surface air temperature data consists of taking into account the effects of the sub-grid topography.

The rate at which air cools with elevation may vary in the range between  $-9.8^{\circ}\text{C}/\text{km}$  (dry-air adiabatic lapse rate) for unsaturated air and  $-4^{\circ}\text{C}/\text{km}$  (saturated adiabatic lapse rate) for saturated air (Dodson and Marks 1997). The average temperature lapse rates show strong variability in relation to the climatic zone, with highest values reached in summer over tropical deserts and lowest negative values in winter over Siberia, Canada, and polar regions (Barry et al., 2008).

A study focusing on the Austrian-Italian Alps and considering more than 600 meteorological stations representative of a wide elevation range (from 100 m a.s.l. to above 2000 m a.s.l.) showed yearly lapse rates ranging from  $-5.4^{\circ}$  to  $-5.8^{\circ}\text{C}/\text{km}$  (Rolland et al. 2003). However, in absence of more accurate lapse rate estimates based on measurements along elevational transects, average temperature gradients of  $-5.5^{\circ}$  (Angot 1892),  $-6.0^{\circ}$  (Dodson and Marks 1997), or  $-6.5^{\circ}\text{C}/(1000\text{ m})-1$  (Barry and Chorley 1987) are often used in applications.

A simple method to downscale temperature data can exploit available information on lapse rate to correct the large scale temperature fields for the elevation mismatch between the coarse model orography and the fine-scale orography of the target grid. In addition to the coarse scale dataset to be downscaled, the only required information is i) an environmental lapse rate and ii) the fine-scale orography. The fine-scale orography can be obtained, for example, by conservatively remapping a km-resolution digital elevation model (DEM) to the target grid. So the fine scale temperature field will be given by

$$t(x, t) = [T(X, t)]_x + L \cdot (z(x) - \langle z(x) \rangle_X)$$

where  $T(X)$  is the coarse-scale temperature field,  $[\dots]_x$  is an interpolation operator to the fine scale  $x$ ,  $\langle \dots \rangle_X$  is an averaging operator at the coarse scale  $X$ ,  $z(x)$  is the DEM orography at the target resolution and  $L$  is the chosen environmental lapse rate. For interpolation two methods have been explored and implemented, providing similar results, a linear interpolation operator, or a nearest-neighbour interpolation, followed by a smoothing step using a flat circular kernel (with equal weights) with a given radius (typically half a pixel size of the coarse grid can be used). For averaging,  $\langle \dots \rangle_X$ , the same smoothing is used. This method is particularly suited if no fine-scale detailed climatology at the target resolution is available.

A second possible method is to exploit information from a reference fine-scale temperature climatology (if available, an example would be the EOBS or ERA5 climatologies). In this case a simple bias correction method can be applied to the coarse data interpolated to the fine scale:

$$t(x, t) = [T(X, t)]_x + \overline{e(x, t)} - \overline{[T(X, t)]_x}$$

where all symbols have the meaning defined above,  $\overline{e(x, t)}$  is a fine-scale reference climatology, and the overbar denotes climatological averaging in time. Notice that this transformation also guarantees that the resulting climatology is identical to the reference climatology:  $\overline{t(x, t)} = \overline{e(x, t)}$

In our application the goal is to downscale seasonal forecast data from 1° to 0.25° spatial resolution. To test our downscaling methods we set up a “perfect model experiment” in which ERA5 daily air temperature data at 0.25° spatial resolution have been upscaled to 1° and then downscaled back to 0.25° with the two different methods presented above, i.e. one based on the lapse rate and the other based on the reference climatology. The results of the two different downscaling procedures have been evaluated against the original ERA5 fields at 0.25° resolution. As an example, Table 1 reports the root mean square error (in K) between the downscaled daily data and the reference ERA5 climatology in the Alpine region (4-19°E, 43-49°N). This test shows a significant improvement in the agreement of the downscaled data with the reference, particularly when a bias correction to a reference climatology is used, that both the nearest neighbour method with smoothing and the bilinear interpolation provide similar results and that the exact choice of the lapse rate has only a small impact.

	Climatology	Lapse rate $L=-5.5$ K/Km	Lapse rate $L=-6.5$ K/Km	Only interpolation
Nearest neighbour	0.520	1.079	1.082	1.399
Bilinear	0.527	1.086	1.083	1.427

Table 1. Root mean square error (in K) between ERA5 data at 0.25° and ERA5 data at 1° downscaled with different methods, compared for the months from May to November in the years 1993-2012, in the Alpine region.

Within the MEDSCOPE project the downscaling methods described above have been implemented into the CSTools R package (Nuria Perez-Zanon et al., 2020), and specifically in the CST\_RFTemp function.

## 2.2 Downscaling of precipitation: a stochastic approach

Stochastic downscaling methods have been identified as particularly suitable for precipitation data (Maraun et al., 2010). We employed the RainFARM stochastic precipitation downscaling method (Rebora et al., 2006; D'Onofrio et al., 2014), recently extended to be used in complex orography (Terzago et al., 2018). Within the MEDSCOPE project the RainFARM method has been implemented in the CSTools R package (Nuria Perez-Zanon et al., 2020).



The stochastic precipitation downscaling method RainFARM has been conceived to downscale large-scale spatio-temporal precipitation fields by producing ensembles of stochastic realizations at finer spatial resolution, which preserve the large-scale statistical properties of the original field and with realistic spatial and temporal correlation structures. The RainFARM method can extrapolate the small-scale correlation structure of precipitation from the spatial power spectrum of the large-scale field and, in principle, it does not require further fields besides the original precipitation to downscale. It does not require, in fact, observational data for calibration as other statistical methods typically do. In presence of a reference fine-scale precipitation dataset, the small-scale correlation structure can be derived from it and then employed to generate the downscaled fields.

An example of the spatial power spectrum of ERA5 daily precipitation at  $0.25^\circ$  over the full domain shown in Figure 1 (period 1993-2018, month of January) is shown in Figure 2a. The spatial power spectrum is approximately isotropic in space. The section of the spectrum shows an approximate power-law behavior, i.e. linear behavior in log-log coordinates as in Figure 2d,e.

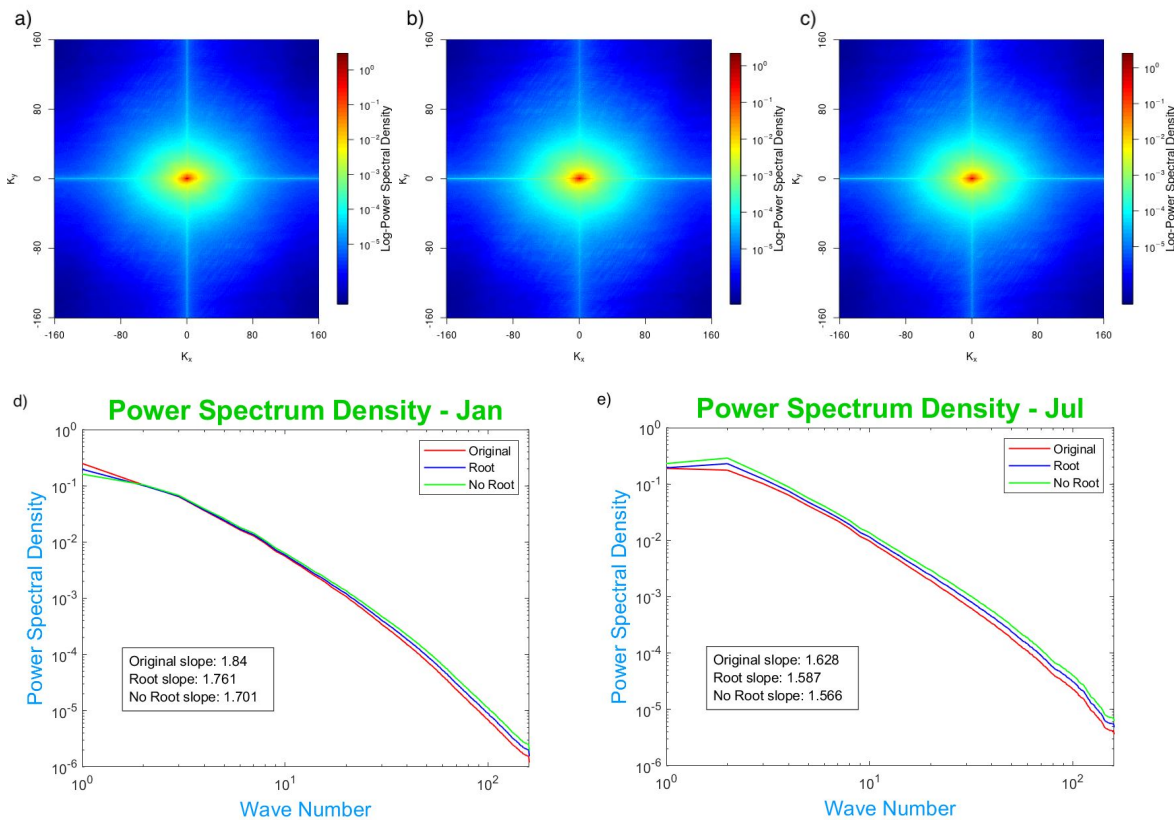


Figure 2. Example of spatial power spectrum (logarithm of the power spectral density) of the ERA5 reanalysis at  $0.25^\circ$  over the full domain of Figure 1, period 1993-2018, month of January, obtained using a) ERA5 original data; b) ERA5 data weighted for the square root of the cosine of the latitude; c) ERA5 data weighted for the cosine of the latitude; (d,e) Examples of power spectral density plotted in log-log coordinates for a winter month (January) and a summer month (July) respectively. Coloured lines correspond to the different ERA5 data used: original data (red), data weighted for the square root of the cosine of the latitude (blue); ERA5 data weighted for the cosine of the latitude (green).

The RainFARM method explained above does not take into account orographic effects at scales smaller than those resolved by the original precipitation field to downscale. The fine-scale



distribution of precipitation in the downscaled fields is not affected by the fine-scale orography, so the long-term climatology at individual grid points could differ significantly from observations in areas with complex topography and particularly in mountain regions. Applications for which the small-scale hydrological balance is important, such as studies of impacts on snow cover or water resources in small mountain basins, require the representation of orographic features. A simple modification has been recently implemented in the RainFARM method to include the orographic effects into the downscaled fields (Terzago et al., 2018). This method exploits information on the fine-scale precipitation climatology, usually available from spatial interpolation of observations from a dense network or from high-resolution model simulations. This reference climatology is used to derive corrective weights which are applied to the stochastic precipitation realizations generated by RainFARM. This method has been demonstrated to reproduce a more realistic long-term precipitation distribution at the fine scales (Terzago et al., 2018).

We show here the RainFARM stochastic downscaling method in a “perfect-model” experiment using ERA5 daily precipitation reanalysis at  $0.25^\circ$  spatial resolution over the full domain shown in Figure 1. ERA5 precipitation fields at  $0.25^\circ$  resolution are first upscaled by aggregating them to a coarser resolution of  $1^\circ$  and then downscaled with RainFARM to the initial resolution of  $0.25^\circ$  (enabling the orographic correction). Spectral slopes are derived from the ERA5  $0.25^\circ$  fields for each month separately (Table 2) and five downscaling realizations are produced. The agreement between the downscaled data and the original  $0.25^\circ$  resolution data provides information on the skill of the downscaling method.

Figure 3 summarizes the results of the perfect model experiment and shows separately the months from May to November (corresponding to the seasonal forecast starting date May 1<sup>st</sup>) and from November to May (corresponding to the seasonal forecast starting date November 1<sup>st</sup>). For each starting date, each of the 26 seasons over the period 1993-2018 and eventually each downscaling realization a Probability Density Function (PDF) of the precipitation field is calculated. The plot shows the PDFs of ERA5 precipitation data downscaled from  $1^\circ$  to  $0.25^\circ$  spatial resolution (cyan line) compared to the original data at  $1^\circ$  (black line) and the original ERA5 data at  $0.25^\circ$  resolution (blue line). The lines represent the 5-95<sup>th</sup> percentile range of the PDF ensembles. The PDF of the downscaled dataset is in very good agreement with the PDF of the original fine-scale dataset, indicating that the method is able to reproduce the precipitation distribution with a high level of accuracy as well as the high-precipitation values, which were not captured by the upscaled ERA5 data (black line). No remarkable differences are found considering the two different periods May-November of November-May. This exercise demonstrates that the RainFARM stochastic downscaling method successfully reconstructs the distribution of precipitation at the small scales.

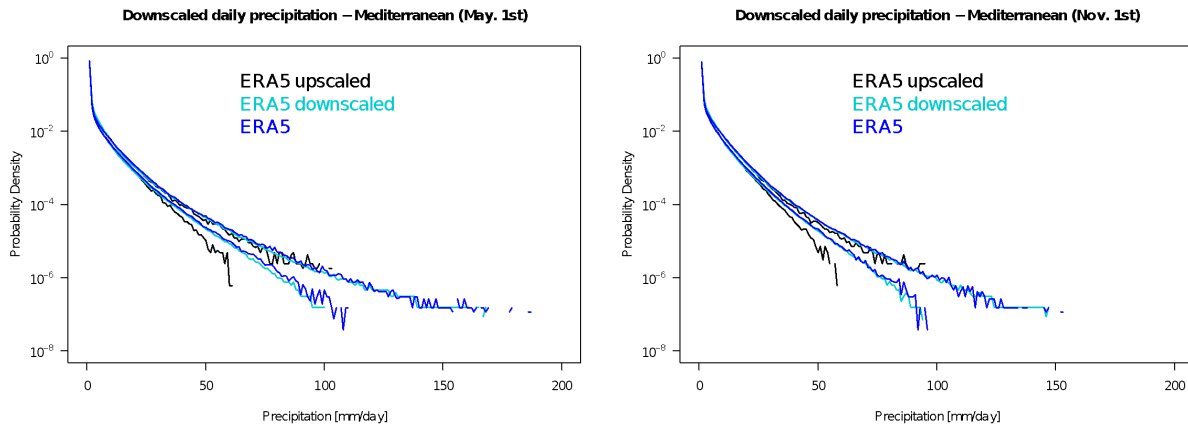


Figure 3 The "perfect model" experiment: Probability density function (PDF) of ERA5 daily precipitation upscaled at  $1^\circ$  spatial resolution, of ERA5 downscaled from  $1^\circ$  to  $0.25^\circ$  compared to the original ERA5 data at  $0.25^\circ$  resolution. The lines represent the 5-95th percentile range of the PDF ensembles.

### 3. The seasonal forecasts data archive

The downscaling chains described above have been applied to seasonal forecasts data. We considered seasonal forecasts of air temperature (tas) and daily precipitation (prlr) from ECMWF S5 and MFS6 models available on Copernicus Climate Change Service (C3S) and downloaded to the CINECA archive available to all MEDSCOPE project participants. Original seasonal forecast data are stored in the following folders:

```
/CINECA01/home/DRES_MEDSCOPE/BSC/C3S/ecmwfS5/daily/prlr
/CINECA01/home/DRES_MEDSCOPE/BSC/C3S/mfs6/daily/prlr

/CINECA01/home/DRES_MEDSCOPE/BSC/C3S/ecmwfS5/6hourly/tas
/CINECA01/home/DRES_MEDSCOPE/BSC/C3S/mfs6/6hourly/tas
```

The archive stores seasonal forecasts with starting dates May 1st and November 1st and 7-months lead-time, covering the periods May-November and November-May, respectively, for all hindcasts years in the period 1993-2018.

The seasonal forecast data, originally on  $1^\circ$  horizontal resolution grid have been downscaled to  $0.25^\circ$  over the Mediterranean domain (Figure 1).

## 4. The downscaling of seasonal forecasts

### 4.1 Downscaling of air temperature

Within the MEDSCOPE project we downscaled seasonal forecast air temperature datasets at  $1^\circ$  spatial resolution to  $0.25^\circ$  over the Mediterranean area by applying the spatial downscaling method described in Section 2.1. In particular, having determined in that section that the method based on a reference climatology provides better results over areas with complex topography compared to a lapse-rate based method, we applied a correction using this method, with a fine-scale reference climatology provided by the ERA5 reanalysis (Hersbach and Dee, 2016; DOI: [10.24381/cds.adbb2d47](https://doi.org/10.24381/cds.adbb2d47)) at  $0.25^\circ$  spatial resolution.

### 4.2 Downscaling of precipitation

We downscaled seasonal forecast precipitation datasets at  $1^\circ$  spatial resolution to  $0.25^\circ$  over the Mediterranean area by applying the RainFARM orographic downscaling method. We used the ERA5 reanalysis at  $0.25^\circ$  spatial resolution as a reference dataset to derive:

- the spatial correlation structure of precipitation at the small scales, i.e. the slope of the power spectrum, for each month
- the reference monthly precipitation climatology necessary for generating the orographic weights.

We employed ERA5 for two reasons: i) it has a global coverage, including ocean areas commonly unavailable in observation-based datasets and key in our study, ii) it is defined at the target resolution of our downscaling exercise, i.e.  $0.25^\circ$

The RainFARM downscaling method has been originally developed and tested on spatial domains on the order of 100 Km. The applications over larger domains, such as the Mediterranean area on the order of several 1000 Km, requires a preliminar applicability study.

As the characteristics of precipitation can be very different within the Mediterranean region (with Western regions heavily exposed to Atlantic air masses, central Europe characterized by the interplay of continental and maritime air masses, mountain ranges subjected to orographic precipitation, North-Africa regions characterized by dry conditions, etc ...) we tested the characteristics of ERA5 precipitation in terms of spatial power spectrum over different sub-regions of the Mediterranean domain (Figure 4, left).

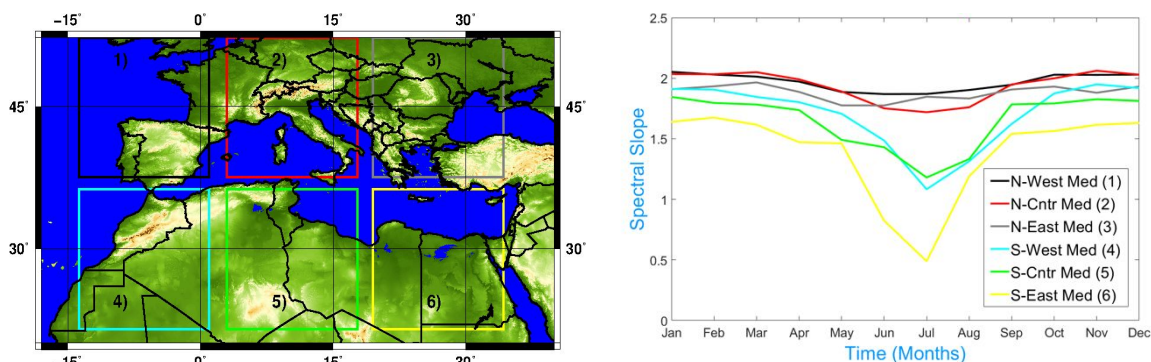


Figure 4 (left) Sub-domains of the Mediterranean region for which we investigated the characteristics of the ERA5 daily precipitation fields in terms of spatial power spectrum; (right) logarithmic slopes of the spatial power spectra of ERA5 daily precipitation, for each month and each sub-domain.

Figure 4 (right panel) reports the slopes of the spatial power spectrum for each month and in each sub-domain shown in the left panel. As expected, some variability is found among different sub-domains. North-Mediterranean sub-domains are characterized by small differences in the monthly precipitation spectra and their logarithmic slopes. On the contrary, South-Mediterranean sub-domains are subjected to larger differences in monthly precipitation spectra and we obtain flatter power spectra in summer months compared to the rest of the year (logarithmic slopes comparable to or lower than 1).

Since for the downscaling we need unique monthly slopes for the whole Mediterranean region, we decided to employ the slopes found for the North-Center sub-domain as they are comparable to those found for most sub-domains (N-West and N-East sub-domains in all month of the year, and South sub-domains in autumn, winter and spring months). To better reproduce the slope of the monthly spectra at the small spatial scales we fit the values corresponding to wavenumbers  $\geq 5$ , a range where all the monthly power spectra show approximately a power-law behavior. The values of the spectral slopes are reported in Table 2. Southern sub-domains have generally flatter power spectra compared to the Northern sub-domains in summer months. Please note that our choice of unique monthly slopes for the Mediterranean region may result in a sub-optimal representation of the downscaled precipitation in North Africa in summer months (JJA). We recommend to evaluate fine-scale precipitation produced by RainFARM against observations before using the data for applications.

Table 2 *Logarithmic slopes of the spatial power spectrum of ERA5 precipitation at  $0.25^\circ$  over the North-Central Mediterranean region ( $2.5^\circ\text{E} - 18.25^\circ\text{E}$ ;  $37.5^\circ\text{N} - 53.25^\circ\text{N}$ , Figure 4) over the period 19930101-20181231. Slopes are calculated considering wavenumbers  $\geq 5$  where the spectra follow approximately a power-law behavior.*

Logarithmic slopes of the ERA5 daily precipitation spatial power spectrum											
Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
2.45	2.43	2.43	2.36	2.22	2.05	1.98	2.06	2.37	2.47	2.52	2.46

As mentioned before, the domain of study is large and up to now we have considered all gridpoints to have the same weight despite higher latitude gridpoints represent smaller areas compared to lower latitude gridpoints. We test the impact of this approximation in the following way. We derive a “modified” ERA5 dataset in which at each gridpoint the original ERA5 precipitation is “weighted” (multiplied) by a factor that is proportional to i) the square root of the cosine of the latitude ii) the cosine of the latitude. In this way latitude grid cells closer to the equator are weighted more than those poleward. The weights are also normalized so that they average to 1 over the domain.

We evaluated the spectra of the original ERA5 precipitation field and that of the two “modified” ERA5 fields, where the precipitation data are area-weighted as explained above. No remarkable difference is found among the power spectra (Figure 2d,e) so we employed the original ERA5 data in the downscaling chain.

The Mediterranean region includes several mountain ranges including the Alps, Pyrenees, Sierra Nevada, Atlas, etc ..., which affects the spatial distribution of precipitation. We take into account the effects of complex topography on fine scale precipitation by applying the RainFARM downscaling algorithm optimized for mountain regions (Terzago et al., 2018). Orographic weights are derived on a monthly basis from ERA5 precipitation dataset at  $0.25^\circ$ .

One of the advantages of stochastic downscaling methods is that they can generate ensembles of possible realizations of the downscaled fields, allowing us to evaluate the uncertainty associated to fine-scale fields. We fully exploit the potential of the RainFARM stochastic downscaling method by generating ensembles of 5 downscaling realizations for each seasonal forecast dataset.

## 5. Evaluation of the downscaled seasonal forecasts

### 5.1. Downscaled air temperature

The bias adjusted and downscaled daily air temperature fields at  $0.25^\circ$  spatial resolution were evaluated against ERA5 temperature reanalysis at the same resolution. The comparison is performed in terms of root mean square error (RMSE) in time between the downscaled monthly-average air temperature and the corresponding ERA5 values (Figure 5, left panels). To highlight the added value of the downscaling method employed with respect to a simpler linear interpolation of the coarse-scale field to the fine-scale grid, we also show the RMSE of the seasonal forecasts linearly interpolated to the fine scale with respect to the ERA5 reference (Figure 5, right panels). The RMSE is smaller for the downscaled fields than for the interpolated fields in all models and all seasons, demonstrating an improvement of the downscaling method compared to a simpler linear interpolation.

We finally quantified the effect of applying (or not) the quantile mapping adjustment to the original forecasts (“Raw”) before downscaling (Table 3). The application of the bias adjustment allows to remarkably reduce the RMSE compared to ERA5 reference data and the combination of the quantile-mapping and the downscaling shows the lowest RMSE for all models and seasons.



### Downscaling

### Bilinear interpolation

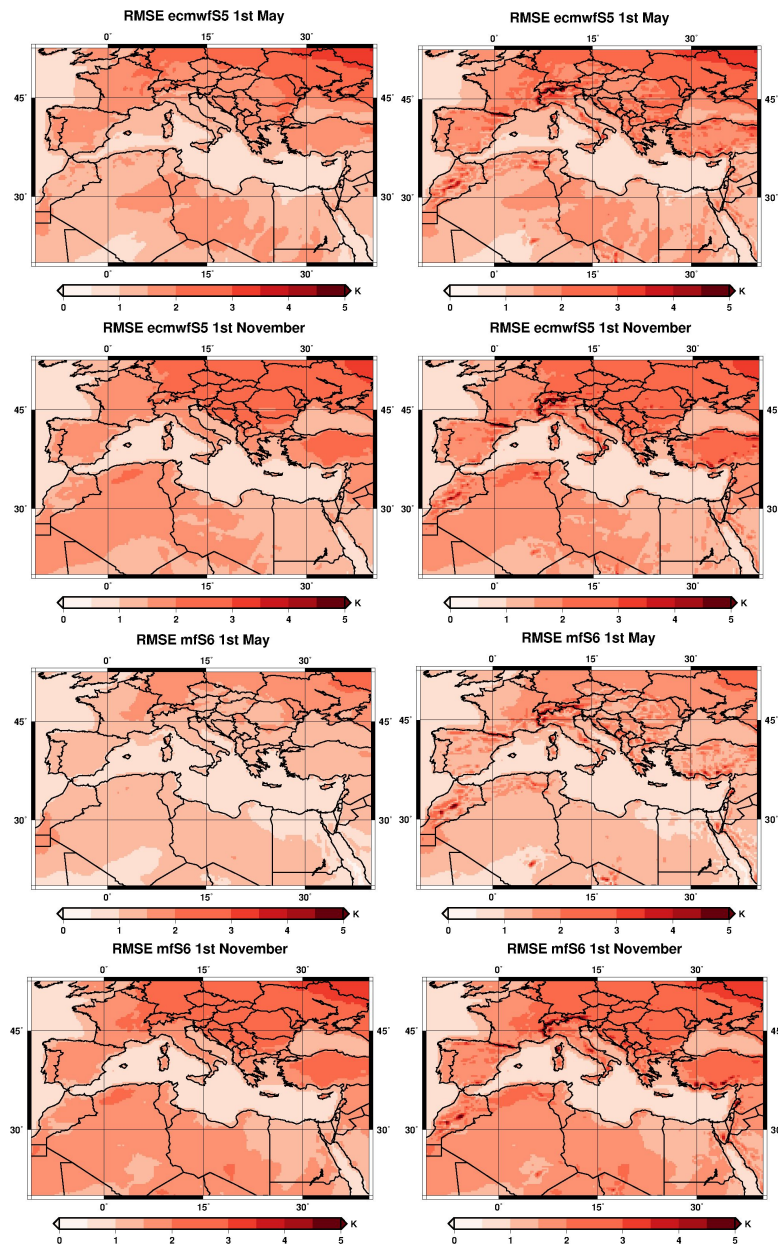


Figure 5 Root mean square error between (left) the downscaled (right) the linear interpolated monthly-average air temperature seasonal forecasts and the reference ERA5 reanalysis, for ECMWFS5 and MFS6 models and for May 1<sup>st</sup> and November 1<sup>st</sup> starting date.



Table 3. Root mean square error between the monthly-average air temperature seasonal forecasts (first column) linearly interpolated, (second column) downscaled, (third column) bias-adjusted and linearly interpolated, (fourth column) bias-adjusted and downscaled to the target 0.25° grid and the reference ERA5 reanalysis, for ECMWFS5 and MFS6 models and for May 1<sup>st</sup> and November 1<sup>st</sup> starting dates.

	RMSE [°C]			
	Raw + Bilinear Interpolation	Raw + Downscaling	QuantMap + Bil. Interpolation	QuantMap + Downscaling
ECMWFS5 - May	2.15	1.78	1.52	1.44
ECMWFS5 - Nov	2.08	1.56	1.59	1.50
MFS6 - May	1.93	1.42	1.24	1.13
MFS6 - Nov	2.51	1.99	1.67	1.58

## 5.2. Downscaled precipitation

The downscaled precipitation fields at 0.25° spatial resolution produced with the RainFARM method were evaluated against ERA5 precipitation reanalysis at the same resolution. As done for the “perfect-case experiment” the comparison is performed in terms of probability density functions.

Clearly we have to compare ensembles of PDFs: in the case of seasonal forecast precipitation data we obtained one PDF for each hindcast year, each ensemble member and each stochastic realization of the downscaled field; in the case of ERA5 we obtained one PDF for each hindcast year. From each PDF ensemble we compute and compare the 5-95<sup>th</sup> percentile range.

Figure 6 shows for each model, ECMWFS5 and MFS6, and each starting date, May 1<sup>st</sup> and November 1<sup>st</sup>, the 5-95<sup>th</sup> percentile ranges of the PDF ensembles generated from:

- The seasonal forecast data at 1° spatial resolution
- The seasonal forecast data downscaled to 0.25° resolution
- The reference ERA5 data at 0.25° resolution

Seasonal forecast data at 1° resolution are unable to represent heavy-precipitation events particularly in summer. The application of the RainFARM downscaling allows to reconstruct the tails of the PDFs of the seasonal forecasts in a more realistic way, so that the PDFs of the downscaled data are similar to those from ERA5 reference, in all seasons and for all models. We can conclude that the application of the RainFARM method allows to obtain fine-scale precipitation fields with similar characteristics as those found in the reference data.

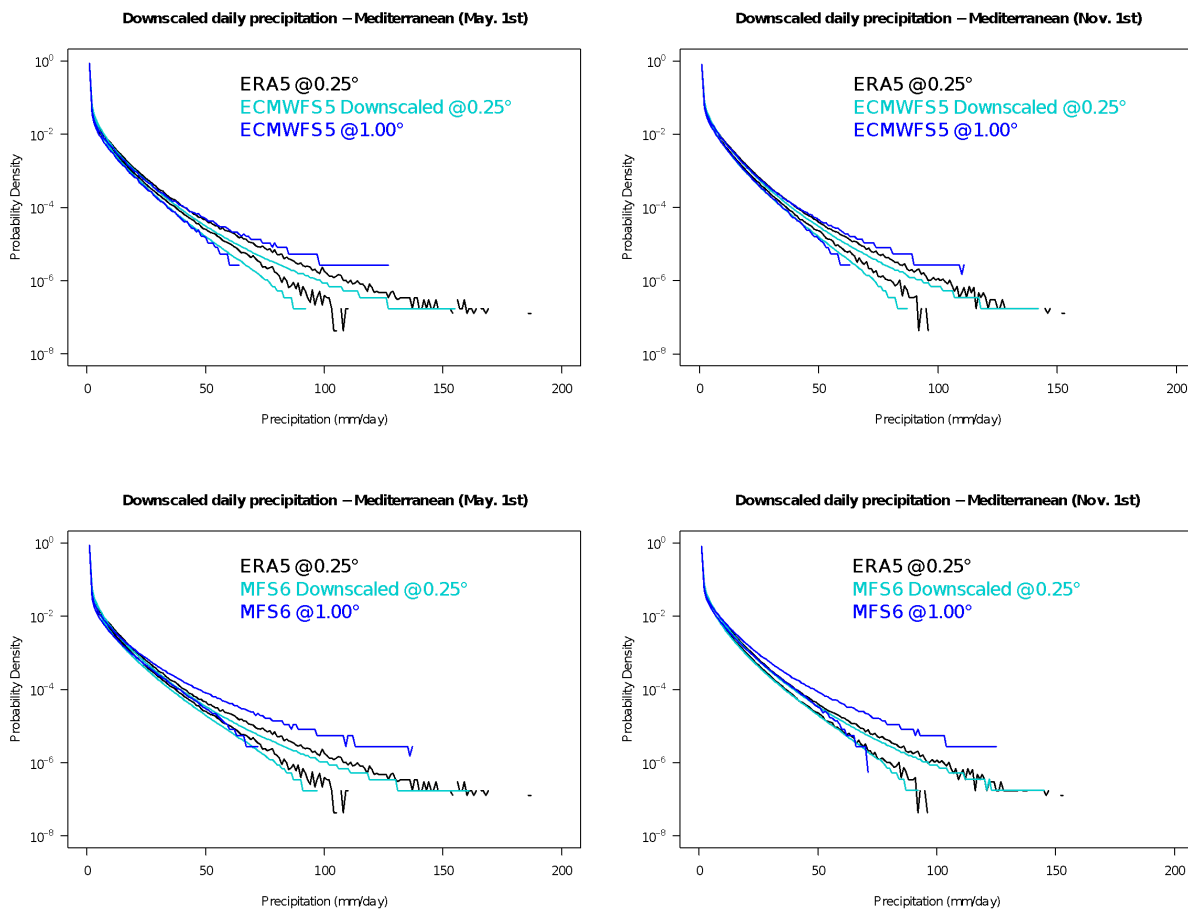


Figure 6. Probability density function (PDF) of daily precipitation from: seasonal forecasts at  $1^\circ$  spatial resolution, seasonal forecasts downscaled from  $1^\circ$  to  $0.25^\circ$ , ERA5 reanalysis at  $0.25^\circ$  resolution. The lines represent the 5-95th range of the PDF ensembles. The plots show the results for different models (ECMWF S5, MFS6) and different starting dates (May 1<sup>st</sup> and November 1<sup>st</sup>).

## 6. Availability of the downscaled data

The downscaled seasonal forecasts of air temperature and precipitation are accessible from the CINECA archive in the following folders:

C- /CINECA01/home/DRES\_MEDSCOPE/CNR-ISAC/downscaling/prlr/ecmwfS5  
C- /CINECA01/home/DRES\_MEDSCOPE/CNR-ISAC/downscaling/prlr/mfs6

C- /CINECA01/home/DRES\_MEDSCOPE/CNR-ISAC/downscaling/tas/ecmwfS5  
C- /CINECA01/home/DRES\_MEDSCOPE/CNR-ISAC/downscaling/tas/mfs6



Detailed information on how to access the CINECA server and download the data is available in this document:  
<https://docs.google.com/document/d/1kFLXtGmR7RpaV4l9NTwDTsWDcuZOH1Y2VaGxRxWLiZ0/edit?usp=sharing>

The R scripts to derive the downscaled data are provided in Appendix A (air temperature) and Appendix B (precipitation).



## References

- Angot, A. (1892). Sur la décroissance de la température de l'air avec la hauteur. C. R. Acad. Sci., **115**, 1272–1273.
- Barry, R. G. and R. J. Chorley (1987). Atmosphere, Weather and Climate. 5th ed. Routledge, 448 pp.
- Barry, R. G. (2008). Mountain weather and climate. Cambridge University Press.
- Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes?. Journal of Climate, 28(17), 6938-6959.
- Dodson, R., and Marks, D. (1997). Daily air temperature interpolated at high spatial resolution over a large mountainous region. Climate Research, 1-20.
- D'Onofrio, D., Palazzi, E., von Hardenberg, J., Provenzale, A., and Calmanti, S. (2014). Stochastic rainfall downscaling of climate models. Journal of Hydrometeorology, 15(2), 830-843.
- Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Skaugen, T. E. (2012). Technical Note: Downscaling RCM precipitation to the station scale using quantile mapping—a comparison of methods. Hydrol. Earth Syst. Sci. Discuss, 9(5), 6185-6201.
- Maraun, D., et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003, doi:10.1029/2009RG000314.
- Nuria Perez-Zanon et al. (2020) CStools: Assessing Skill of Climate Forecasts on Seasonal-to-Decadal Timescales. R package version 3.0.0 <https://CRAN.R-project.org/package=CStools>
- Rebora, N., Ferraris, L., von Hardenberg, J., and Provenzale, A. (2006). RainFARM: Rainfall downscaling by a filtered autoregressive model. Journal of Hydrometeorology, 7(4), 724-738.
- Rolland, C. (2003). Spatial and seasonal variations of air temperature lapse rates in Alpine regions. Journal of Climate, 16(7), 1032-1046.
- Terzago, S., Palazzi, E., and von Hardenberg, J.: Stochastic downscaling of precipitation in complex orography: a simple method to reproduce a realistic fine-scale climatology, Nat. Hazards Earth Syst. Sci., 18, 2825–2840, <https://doi.org/10.5194/nhess-18-2825-2018>, 2018.
- Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, Int. J. Climatol., 31, 1530–1544, doi:10.1002/joc.2168, 2011.
- Themeßl, M. J., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, Climatic Change, 112, 449–468, doi:10.1007/s10584-011-0224-4, 2012.



## Appendix A – Code used for temperature downscaling

```
library(devtools)
library(zeallot)
library(fields)
library(raster)
library(ncdf4)
load_all("/work/users/giulio/git/cstools")

in_dir <- "/work/users/giulio/temperature/" #main directory
md <- c("0501","1101") #season - starting date
var <- "tas" #variable name
var_obs <- "t2m" #ERA5 variable name
model <- c("ecmwfS5","mfS6") #model name
clim_name <- c("era5_t2m_clim_summer_red.nc","era5_t2m_clim_winter_red.nc") #climatology
name
ye <- matrix(c(2016,2016,2018,2017),nrow = 2, ncol = 2) #end year

#directories for delta saved data
delta_dir <- paste0(in_dir,"downscaled/")
delta_name <- c("delta_oro","delta_oro0","delta_clim")

#interpolation method
method <- c("bil","nearn") #interpolation method name directories
met_n <- c("bilinear","nearest") #interpolation method name in function
ne <- 25 #number of ensemble members
df <- 4 #downscaling factor
coo_geo1 <- c(-18,41,19,53) #coarse domain ECMWF
coo_geo2 <- c(-19.5,42.5,17.5,54.5) #coarse domain Meteo France
lapse_rate <- c(5.5,0.0,6.5)

#orography
oro_in <- paste0(in_dir,"datasets/oro_red.nc")
oro_nc <- nc_open(oro_in)
oro_data <- ncvar_get(oro_nc, "topo")
oro_lon <- ncvar_get(oro_nc, "lon")
oro_lat <- ncvar_get(oro_nc, "lat")
oro <- list(data = oro_data, lat = oro_lat, lon = oro_lon)
attr(oro, 'class') <- 's2dv_cube'

for (i in 1:length(md)) {

  #ERA5 climatology
  clim_in <- paste0(in_dir,"datasets/",clim_name[i])
  clim_nc <- nc_open(clim_in)
  clim_data <- ncvar_get(clim_nc, "t2m")
  clim_lon <- ncvar_get(clim_nc, "lon")
  clim_lat <- ncvar_get(clim_nc, "lat")
  clim <- list(data = clim_data, lat = clim_lat, lon = clim_lon)
```



```

attr(clim, 'class') <- 's2dv_cube'

for (j in 1:length(model)) {

  #starting dates
  ys <- 1993
  ssd <- ys:ye[i,j]
  s <- length(ssd)
  sd <- array("",s)

  #starting dates selection
  for (k in 1:s) {
    sd[k] <- paste0(ssd[k],md[i])
  }

  #domain selection
  if (j == 1) {
    coo_geo <- coo_geo1
  } else {
    coo_geo <- coo_geo2
  }

  #seasonal forecast loading
  print(paste0("Load Seasonal Forecasts ",model[j]," season ",md[i]))
  tas_in <- paste0(in_dir, var, "/", model[j], "$VAR_NAME$$_START_DATE$.nc")
  tas_data <- CST_Load(var, exp= list(list(path = tas_in)), obs=NULL, sdates = sd, nmember =
ne, storefreq = "daily",
    sampleperiod = 1, latmin = coo_geo[3], latmax = coo_geo[4], lonmin = coo_geo[1],
lonmax = coo_geo[2],
    output = 'lonlat', nprocs = 6)
  dim(tas_data$data)

  #ERA5 Data loading
  obs_in <-
paste0(in_dir,"datasets/era5_obs/",model[j],"/era5_$VAR_NAME$$_START_DATE$.nc")
  obs_data <- CST_Load(var_obs, exp= list(list(path = obs_in)), obs=NULL, sdates = sd,
nmember = 1, storefreq = "daily",
    sampleperiod = 1, latmin = coo_geo[3], latmax = coo_geo[4], lonmin = coo_geo[1],
lonmax = coo_geo[2],
    output = 'lonlat', nprocs = 6)
  dim(obs_data$data)

  #Monthly splitting
  print("Monthly Splitting")
  tasm <- CST_SplitDim(tas_data, split_dim = c('ftime'))
  obsm <- CST_SplitDim(obs_data, split_dim = c('ftime'))
  dim(tasm$data)
  dim(obsm$data)

  #Quantile Mapping

```

```

print("Performing the QuantileMapping")
tas.qm <- CST_QuantileMapping(tasm,obsm, method = "QUANT", wet.day=F,sample_dims =
c('member', 'sdate', 'ftime'))
dim(tas.qm$data)
# MergeDims
print("Restoring original dimensions")
tas.qmm <- CST_MergeDims(tas.qm, merge_dims=c("ftime", "monthly"),na.rm=TRUE)
# Adjust longitude to be monotonic [-180;180]
lon <- as.numeric(tas.qmm$lon)
tas.qmm$lon <- lon*(lon<=180)+(lon-360)*(lon>180)
lat=tas.qmm$lat
lon=tas.qmm$lon
l <- length(lon)
lf <- l*df
#directories for quantile mapping
qm_path <- paste0(in_dir,"tas_qm/",model[j],"/")
print("Save Bias-Adjusted")
CST_SaveExp(tas.qmm, qm_path)
rm(tas_data, tasm, tas.qm, tas.qmm,obs_data,obsm)

#Seasonal forecast loading (after quantile mapping)
print(paste0("Load Seasonal Forecasts ",model[j]," season ",md[i]," after quantile mapping"))
tas_in <- paste0(qm_path, "exp1/tas/$VAR_NAME$_$START_DATE$.nc")
tas_data <- CST_Load(var, exp= list(list(path = tas_in)), obs=NULL, sdates = sd, nmember =
ne, storefreq = "daily",
sampleperiod = 1, latmin = coo_geo[3], latmax = coo_geo[4], lonmin = coo_geo[1],
lonmax = coo_geo[2],
output = 'lonlat', nprocs = 6)
dim(tas_data$data)

for (l in 1:length(method)) {

  for (m in 1:length(delta_name)) {

    if (m == 3) {

      ref <- clim
      lapse_r <- TRUE

    } else {

      ref <- oro
      lapse_r <- FALSE

    }

    lr <- lapse_rate[m]

    #Delta computation
  }
}

```

```

delta <- CST_RFTemp(tas_data, ref, xlim = c(-11,40), ylim = c(20,52.5), lapse = lr,
lon_dim = "lon",
lat_dim = "lat", time_dim=c("member", "sdate", "ftime"), nolapse = lapse_r,
verbose = TRUE,
compute_delta = TRUE, method = met_n[1])
print(min(delta$data))

for (n in 1:s) {

  #seasonal forecast loading
  print(paste0("Load Seasonal Forecasts ",model[j]," season ",md[i]," method
",method[l]," delta ",delta_name[m]))
  tas_sd_data <- CST_Load(var, exp= list(list(path = tas_in)), obs=NULL, sdates =
sd[n], nmember = ne,
storefreq = "daily", sampleperiod = 1, latmin = coo_geo[3], latmax =
coo_geo[4],
lonmin = coo_geo[1], lonmax = coo_geo[2], output = 'lonlat', nprocs = 6)
dim(tas_sd_data$data)

  #temperature downscaling
  temp_ds <- CST_RFTemp(tas_sd_data, ref, xlim = c(-11,40), ylim = c(20,52.5), lapse
= lr, lon_dim = "lon",
lat_dim = "lat", time_dim=c("member", "sdate", "ftime"), nolapse = lapse_r,
verbose = TRUE,
delta = delta, method = met_n[1])

  save_name <- paste0(delta_dir,delta_name[m],"/",model[j],"/",method[l],"/")

  #saving data
  CST_SaveExp(temp_ds,save_name)

  rm(tas_sd_data,temp_ds)
}
rm(delta)
}
rm(tas_data)
}
rm(clim_nc,clim_data,clim_lon,clim_lat,clim)
}

```

## Appendix B – Code used for precipitation downscaling

```
#!/usr/bin/env Rscript
# R script to downscale precipitation seasonal forecasts with the
RainFARM method (D'Onofrio et al., 2014; Terzago et al., 2018)
# using functions of the R CStools package
# 1. Load seasonal forecasts & ERA5 daily precip data, both at 1°
spatial resolution
# 2. Perform quantile mapping to correct ECDF on a monthly basis.
Reference ERA5 1° res
# 3. Get Monthly Slopes from ERA5 at 25 km
# 4. Estimate orographic weights (monthly basis) from ERA5 at 25
km
# 5. Downscale month by month from 1° to 0.25°
# Use:      Rscript
/work/users/silvia/medscope/downscaling_med/rainfarm-downscale.R
model      start_date_MMDD
# Example:  Rscript
/work/users/silvia/medscope/downscaling_med/rainfarm-downscale.R
ecmwfS5 0501
#
# Pre-requisite:
# Get ERA5 data:
./work/users/silvia/medscope/downscaling_med/era5_data/get_era5.sh
# Get model data:
./work/users/silvia/medscope/downscaling_med/seasfc_data/download_
med_mmday.sh

library(devtools)
library(zeallot)
library(fields)
library(raster)
load_all()

args = commandArgs(trailingOnly=TRUE)

# test if there is at least one argument: if not, return an error
if (length(args)<2) {
  stop("Two arguments must be supplied: model , startDate
MMDD", call.=FALSE)
} else {
  print(paste("Model: ", args[1]))
  print(paste("Starting date MMDD: ", args[2]))
}
# end test
#####
```

```
# get or set variables
model <- args[1]
mdd <- args[2]
wdir <- "/work/users/silvia/medscope/downscaling_med/"
outpath <- paste0(wdir, "output_qm/", model, "/")
outdir <- paste0(outpath, "experiment/prlr/")
downdir <-
"/work/users/silvia/medscope/downscaling_med/downscaled/"
system(paste("mkdir", outpath))

era5clim <-
"/work/users/silvia/medscope/CSTools/era5_tp_mon_latlon.nc"
grid <- "/work/users/silvia/medscope/CSTools/clim.grd"

infile <-
paste0("/work/users/silvia/medscope/downscaling_med/seasfc_data/",
model, "/$VAR_NAME$_$START_DATE$.nc")
reffile <-
'/work/users/silvia/medscope/downscaling_med/era5_data/era5_$VAR_N
AME$_$START_DATE$.nc'

# Spectral slopes (monthly scale) derived from ERA5 0.25°
resolution over the domain 2.5/18.25/37.5/53.25
# Slopes calculated from wavenumber >=5
slope.mon <-
c(2.45, 2.43, 2.43, 2.36, 2.22, 2.05, 1.98, 2.06, 2.37, 2.47, 2.52, 2.46)

if (mdd=="0501") {
  mon <- c(5:11)
  ltn <- c( 1, 32, 62, 93, 124, 154, 185)
  ltx <- c(31, 61, 92, 123, 153, 184, 214)
} else if (mdd=="1101") {
  mon <- c(11, 12, 1:5)
  ltn <- c( 1, 31, 62, 93, 121, 152, 182)
  ltx <- c(30, 61, 92, 120, 151, 181, 212)
}

var='prlr'
nens=25
lonn=-19
lonx=60.5
latn=0
latx=79.5
down_fact=4
nrealiz=5
y1=1993
y2=2018
```

```

yrs  <- y1:y2
nsd  <- length(yrs)
sd   <- vector()
for (i in 1:nsd) { sd[i] <- paste0(yrs[i], mddd) }
startDate <- sd

# end get or set variables
#####

# Load Seasonal Forecasts for the Bias Adjustment by Quantile
Mapping
print("Load Seasonal Forecasts for the Bias Adjustment by Quantile
Mapping")

exp <- list(name = 'experiment', path = infile)
obs <- list(name = 'observation', path = reffile)

exp %<-% CST_Load('prlr', exp=list(exp), obs=NULL, sdates
=startDates, nmember = nens, storefreq = "daily", sampleperiod =
1, latmin = latn, latmax = latx, lonmin = lonn, lonmax = lonx,
output = 'lonlat', nprocs=6)

obs %<-% CST_Load('prlr', exp=list(obs), obs=NULL, sdates
=startDates, nmember = 1, storefreq = "daily", sampleperiod = 1,
latmin = latn, latmax = latx, lonmin = lonn, lonmax = lonx, output
= 'lonlat', nprocs=6)

dim(exp$data)
dim(obs$data)

# SplitDim
print("Splitting ftime by month ...")
expm  <- CST_SplitDim(exp, split_dim = c('ftime'))
obsm  <- CST_SplitDim(obs, split_dim = c('ftime'))
obsm$data <- obsm$data[, , , , , 1:7]

dim(expm$data)
dim(obsm$data)

# CST_QuantileMapping by month:
# sample_dims: a character vector indicating the dimensions that
# can be used as sample for the same distribution
# Please note that wet.day=T would be optimal but not enough
# non-zero data in a month to work
print("Performing the QuantileMapping ...")
exp.qm <- CST_QuantileMapping(expm, obsm, method = "QUANT",
wet.day=F, sample_dims = c('member', 'sdate', 'ftime'))
dim(exp.qm$data)

```



```
# MergeDims
print("Restoring original dimensions ...")
exp.qmm <- CST_MergeDims(exp.qm, merge_dims=c("ftime", "monthly"),
na.rm=TRUE)

# Adjust longitude to be monotonic [-180;180]
lon <- as.numeric(exp.qmm$lon)
exp.qmm$lon <- lon*(lon<=180)+(lon-360)*(lon>180)
lat=exp.qmm$lat
lon=exp.qmm$lon
l <- length(lon)
lf <- l*down_fact

print (paste("Save Bias-Adjusted data to", outdir))
CST_SaveExp(exp.qmm, outpath)

rm(exp, expm, exp.qm, exp.qmm)

#####
# Downscaling by month (slope by month)
print("Subsetting data by month ... ")

infile <- paste0(outdir, "$VAR_NAME$_$START_DATE$.nc")
exp <- list(name = 'experiment', path = infile)

for (i in 1:length(mon)){

  # Check month
  mm <- sprintf("%02d", mon[i])
  print(mm)
  system(paste0("mkdir ", downdir, "mon_", mm))

  # Use slopes from ERA5
  slo <- slope.mon[mon[i]]

  # Get weights from ERA5
  print("Get weights from ERA5")
  climmon <- paste0("clim_" ,mm, ".nc")
  system(paste("rm", climmon))
  cmd0 <- paste0("cdo remapnn,", grid, " -selmon,", mm, "
", era5clim, " ", climmon)
  system(cmd0)
  w_clim_tmp<-CST_RFWeights(climmon, lat=lat, lon=lon,
nf=down_fact)
  # Get rid of NaN in the weights
  r <- raster(nrow=lf, ncol=lf)
  r[] <- w_clim_tmp
```

```

base <- raster(nrow=lf, ncol=lf)

if ( length(table(is.na(as.matrix(r))))>1 ) {
  while (table(is.na(as.matrix(r)))[2]>0) {
    r <- resample(r, base, method='bilinear')
    if (length(table(is.na(as.matrix(r)))) ==1
  ) break
  }
}

w_clim <- as.matrix(r)

# Get seasonal forecast for month mm, loop on start-date
print("Get seasonal forecast for month mm, loop on
start-date")
for (j in 1:length(startDates)){

  if ((mmdd=="1101") && (startDates==1995 |
startDates==1999 | startDates==2003 | startDates==2007
|startDates==2011 | startDates==2015)){
    print ("Anno Bisestoxxx")
    ltn      <- c( 1,31,62, 93,122,153,183)
    ltx      <- c(30,61,92,121,152,182,213)
  }
  print(startDates[j])
  ls %<-% CST_Load('prlr', exp=list(exp), obs=NULL,
sdates =startDates[j], nmember = nens, leadtimemin = ltn[i],
leadtimemax=ltx[i],
                                storefreq = "daily", sampleperiod = 1,
latmin =latn, latmax = latx, lonmin = lonn, lonmax = lonx,
                                output = 'lonlat',nprocs=6)
  lon      <- ls$lon
  ls$lon <- lon*(lon<=180)+(lon-360)*(lon>180)

  # Orographic downscaling
  print("Orographic downscaling")
  for (k in 1:nrealiz){

    k2 <- sprintf("%03d", k)
    fs <- CST_RainFARM(ls,nf=down_fact,
weights=w_clim, slope=slo, kmin=1, nens=1, verbose=TRUE,
                                time_dim=c("member","sdate",
"ftime"), nprocs=6, drop_realization=T)

    # Save output

```

```

        print(paste0("Saving output to: ",
downmdir, "mon_", mm, "experiment/prlr/"))
        CST_SaveExp(fs,  paste0(downmdir, "mon_",
mm))

        filetmp=paste0(downmdir, "mon_", mm,
"/experiment/prlr/prlr_", startDates[j], ".nc")
        fileout=paste0(downmdir, "mon_", mm,
"/experiment/prlr/prlr_", startDates[j], "_rainfarm_", k2 ,".nc")
        system(paste("mv", filetmp, fileout))
        rm(fs)
    }
    rm (ls)
}

# Mergetime
print (paste0("Merge monthly downscaled data by starting date,
then write output to: ", downmdir, model, "/"))
system(paste0("mkdir ", downmdir, model, "/"))
for (j in 1:length(startDates)){
    for (k in 1:nrealiz){
        k2 <- sprintf("%03d", k)
        nfiles      <- paste0(downmdir,
"mon_*/experiment/prlr/prlr_", startDates[j], "_rainfarm_", k2
, ".nc")

        downfile    <- paste0(downmdir, model, "/prlr_",
startDates[j], "_rainfarm_", k2 , ".nc")
        system(paste("cdo mergetime", nfiles, downfile))
    }
}

# End downscaling

```