Article

Toward a Regional-Scale Seasonal Climate Prediction System over Central Italy Based on Dynamical Downscaling

Lorenzo Sangelantoni *, Rossella Ferretti and Gianluca Redaelli

CETEMPS—Department of Physical and Chemical Sciences of University of L’Aquila, 67100 L’Aquila, Italy; rossella.ferretti@aquila.infn.it (R.F.); gianluca.redaelli@aquila.infn.it (G.R.)

* Correspondence: lorenzo.sangelantoni@aquila.infn.it; Tel.: +39-3337351924

Received: 5 August 2019; Accepted: 3 October 2019; Published: 5 October 2019

Abstract: Anticipating seasonal climate anomalies is essential for defining short-term adaptation measures. To be actionable, many stakeholders require seasonal forecasts at the regional scale to be properly coupled to region-specific vulnerabilities. In this study, we present and preliminarily evaluate a regional-scale Seasonal Forecast System (SFS) over Central Italy. This system relies on a double dynamical downscaling performed through the Regional-scale Climate Model (RCM) RegCM4.1. A twelve-member ensemble of the NCEP-CFSv2 provides driving fields for the RegCM. In the first step, the RegCM dynamically downscales NCEP-CFSv2 predictions from a resolution of 100 to 60 km over Europe (RegCM-d1). This first downscaling drives a second downscaling over Central Italy at 12 km (RegCM-d2). To investigate the added value of the downscaled forecasts compared to the driving NCEP-CFSv2, we evaluate the driving CFS, and the two downscaled SFSs over the same (inner) domain. Evaluation involves winter temperatures and precipitations over a climatological period (1982–2003). Evaluation for mean bias, statistical distribution, inter-annual anomaly variability, and hit-rate of anomalous seasons are shown and discussed. Results highlight temperature physical values reproduction benefiting from the downscaling. Downscaled inter-annual variability and probabilistic metrics show improvement mainly at forecast lead-time 1. Downscaled precipitation shows an improved spatial distribution with an undegraded but not improved seasonal forecast quality.

Keywords: regional climate modeling; dynamical downscaling; ensemble seasonal forecasting

1. Introduction

In the last few decades, the availability of climate information on different temporal and spatial scales has increased tremendously. If on the one hand climate projections (i.e., over several decades temporal horizon) represent the basis to define the next decades climate impact adaptation measures, on the other hand, many systems (agriculture, energy production, hydrology, water reservoir, and tourism) management particularly relies on climate information on a seasonal time scale. The latter can represent an essential tool to reduce societal vulnerabilities to the inter-annual climate fluctuation through short-term (i.e., next season) climate impact mitigation measures.

Seasonal forecasting represents an intermediate field where aspects typical of the short-term (weather forecast) and long-term (climate projections) interact. This aspect makes seasonal forecasting facing issues traditionally found in weather forecasting (e.g., difficulties on initializing simulations) and in long-term climate projections (e.g., realistic representation of slow varying feedback processes).

Seasonal Forecast Systems (SFSs) provide climate predictions on a time scale ranging from a few weeks to several months [1]. SFSs generally produce multiple simulations (ensemble) defining
a probabilistic information about anomalous conditions regarding monthly/seasonal statistics (e.g., mean or extreme).

Despite the chaotic component of atmospheric dynamics, prediction attempts on a seasonal temporal scale can be performed by evaluating slowly varying climate system components. The predictability of the climate variables most commonly considered for human activities is connected to the forcing exerted by the oceanic circulation, including sea surface temperature (SST), significantly impacting global atmospheric circulation (see e.g., [2–5]). El Ninó–Southern Oscillation (ENSO) represents one of the most relevant sources of seasonal time scale predictability at tropical latitudes. Several studies demonstrate ENSO’s influence beyond tropical latitudes circulation, highlighting the ENSO prediction as fundamental also for extra-tropical areas seasonal forecast [6].

For the European winter season, factors such as stratospheric processes, soil moisture, sea-ice extension, and North Atlantic Oscillation (NAO) phase have been identified as sources of predictability [7–9]. Specifically, for the winter season evidence indicates a stratospheric variability that significantly impacts seasonal temperature anomalies [10]. In this context, it has been found that interactions between stratospheric (sudden stratospheric warming) and tropospheric (quasi-biennial oscillation) dynamics can have a substantial impact on temperature anomalies [11]. However, a poorer predictability characterizes the mid-latitudes compared to tropical latitudes. In fact, the latter are affected by a larger chaotic component which limits the skill/quality of the current seasonal forecast systems [9].

Coupled Atmosphere–Ocean General Circulation Models (AO-GCMs) are the established tool for reproducing atmospheric and ocean dynamics at global scale. The products of several operational global-scale SFSs are currently available, including Climate Forecast System from the National Centers for Environmental Prediction (NCEP-CFS, [12]), European Center for Medium-Range Weather Forecasts (ECMWF, [13]), UK Meteorological Office [14], Meteo-France [15], and the Euro-Mediterranean Centre on Climate Change (CMCC) Foundation [16].

Global-scale SFSs have to cope with a wide range of bias sources that could be discerned in model deficiencies resulting from the imperfect formulation of atmospheric dynamics, parameterizations, discretization, and quality of the initial condition fields. In the current operational global-scale SFSs, generally only the latter source of uncertainty is taken into account, by producing an ensemble of runs using slightly different initial conditions. To build an ensemble of seasonal forecasts, two main techniques are used. The “burst” mode, consisting of starting all the members on the same start date with slightly perturbed initial states to represent observation uncertainty (e.g., ECMWF), and the “lagged” mode, consisting of initializing members on different start dates (e.g., NCEP-CFS).

Whatever is the choice regarding the ensemble building technique, global scale SFSs still present too coarse spatial resolution for being climate informative and usable in the context of regional scale impact studies (e.g., to drive impact models as hydrological, energy, or agricultural production models). In this regard, to increase the spatial detail of seasonal climate information, two main approaches exist. Statistical approaches, relying on statistical models built on relationships observed on a large-scale circulation and local scale climate variables (perfect prognosis) or between simulations and observations over a common historical period (model output statistics [17]). On the other hand, a physical-based (dynamic) approach consists of nesting a high-resolution Regional Climate Model (RCM) within a driving Global-scale Climate Model (GCM) providing initial and boundary conditions.

In this study, we apply the latter dynamical downscaling technique, over a geographical domain centered over Central Italy. In recent years, many studies considered regional-scale dynamically downscaled SFSs, mainly focused on tropical Asian regions and Africa (see e.g., [18–21]). Based on the author’s knowledge, few studies involving regional-scale SFC over the European domain exist [9,22,23], and even less focused on the Mediterranean [24] basin and the Italian peninsula [25]. However, in recent years, an interesting multi-partner initiative aimed at investigating inter-annual variability mechanisms and at providing tailored climate services mainly on a seasonal time scale in the Mediterranean basin has been established (MEDSCOPE, https://www.medscope-project.eu/).
The Mediterranean basin is showing a sensitivity to climate change higher than the global average. For this reason, it has been identified as one of the climate change hot spots [26]. Furthermore, it is characterized by a large inter-annual variability as a consequence of the interaction between sub-tropical and mid-latitude climate dynamics. In this context, where most of the relevant economic sectors are deeply dependent on climate inter-annual variability (e.g., agricultural production and touristic activities) the availability of regional-scale seasonal climate information is essential in order to foster short-term climate impact reduction measures.

In this work, we present a novel regional-scale dynamically downscaled SFS, developed for operational purposes. It consists of a double dynamical downscaling performed over Central Italy by using a high-resolution RCM (RegCM4.1) driven by global-scale NCEP-CFSv2 fields. An evaluation of winter-season temperature and precipitation climatology, compared against reanalysis data sets, is performed and discussed. In this context, the potential added value of the downscaled RegCM forecast compared to the global-scale driving NCEP-CFS is investigated. The evaluation considers a climatological period of 22 years (1982–2003). Investigation focuses on climatological monthly mean values representation, statistical properties, anomalies inter-annual variability, and probabilistic hit-rate of anomalous conditions for temperature at 2 m and precipitation.

2. Modeling Chain and Evaluation Metrics

The modeling chain proposed in this study consists of a double dynamical downscaling of global-scale CFS simulations performed by a high-resolution RegCM.

The RegCM is the 4th generation of the regional climate model developed at the International Center of Theoretical Physics [27,28]. RegCM4.1 is a hydrostatic atmospheric model defined in the $\sigma$ vertical coordinate system. Large-scale precipitation is parameterized by the Subgrid Explicit Moisture Scheme (SUBEX; [29]). The Grell convection parameterization [30] uses the Arakawa–Schubert [31] closure schemes. Ocean fluxes computation are performed according to the [32] scheme. Interactions between atmosphere and biosphere are described according to the Biosphere–Atmosphere Transfer Scheme [33]. The convective precipitation component is parameterized by a cumulus-convection-scheme from [34].

The driving CSF fields consist of a subset of 22 winter seasons (1982–2003; November, December, January, and February months) retrospective forecasts (indicated as forecast throughout the paper), extracted from the NCEP-CFSv2 (Saha et al., 2014) data archive (https://nomads.ncdc.noaa.gov/modeldata). The CFS forecasts have initial conditions of 00:00, 06:00, 12:00, and 18:00 UTC every 5 days. To build the initial and boundary conditions for the RegCM, we consider three different start days (2, 7, and 12) of November for each year of the climatological 22-year period. Since every initialization day has in turn four initialization hours (00:00, 06:00, 1200, and 1800 UTC) we set up a lagged ensemble of twelve members. On each of the 22-year period, RegCM dynamically downscales each of the twelve members with a 3-month seasonal simulation from 1 December (corresponding to the first month with available simulations from all the twelve CFS members) to 28th February (29th). RegCM initial and boundary conditions consist of temperature, geopotential, specific humidity, and U-V wind components on 37 vertical pressure levels at 100 km horizontal resolution. More detailed information about the dynamical core and the physical parameterizations of the NCEP-CFSv2 can be found in [35].

The first dynamical downscaling is performed at 60 km resolution over continental Europe (RegCM-d1 hereafter). The outputs of the first downscaling are used to build the initial and boundary conditions for the second dynamical downscaling (RegCM-d2 hereafter) with a domain centered over Central Italy at 12 km resolution (Figure 1).
We began by analyzing the capability of the three CFSs of reproducing physical values characterizing winter seasons of the climatological period. Operating over each grid node of the three systems, the data with different resolution from three CFSs have been interpolated on a common grid with the resolution of the reanalysis.

Figure 2 shows the mean bias affecting the ensemble mean of the three CFSs for the three winter months. The original NCEP-CFSv2 shows large positive biases (up to $5\,^\circ{C}$) in correspondence of the most complex-orography portion of the domain. At the same time, negative biases (down to $-5\,^\circ{C}$) are found over sea grid points. This bias pattern is similar for all the three lead time (December, January, and February) months. The first downscaling product, RegCM-d1, shows similar spatial distribution of the bias, even if the magnitude is reduced, especially over the sea grid points. A consistent reduction of biases can be seen in the RegCM-d2 product, with biases generally on the $-1: +1\,^\circ{C}$ range. Larger biases up to $+3\,^\circ{C}$ are found in the north-western and north-eastern border of the downscaled simulation, where numerical disturbance from the nesting can be expected. In general, as expected the higher resolution of the RegCM-d2, compared to the global-scale driven fields, provides an overall improved representation of the climatological temperature by removing orography-dependent bias resulted in the coarser resolution simulations.
CFSs have been interpolated on a common grid with the resolution of the reanalysis.

Figure 2. Climatological period (1982–2003 for December 1983–2004 for January and February) ensemble mean monthly mean temperature bias of the three forecasts products, compared to the reanalysis. Units are in °C.

Figure 2 shows the mean bias affecting the ensemble mean of the three CFSs for the three winter months. The original NCEP-CFSv2 shows large positive biases (up to 5 °C) in correspondence of the most complex-orography portion of the domain. At the same time, negative biases (down to -5 °C) are found over sea grid points. This bias pattern is similar for all the three lead time (December, January, and February) months. The first downsampling product, RegCM-d1, shows similar spatial distribution of the bias, even if the magnitude is reduced, especially over the sea grid points. A consistent reduction of biases can be seen in the RegCM-d2 product, with biases generally on the -1:+1 °C range. Larger biases up to +3 °C are found in the north-western and north-eastern border of the downscaled simulation, where numerical disturbance from the nesting can be expected. In general, as expected the higher resolution of the RegCM-d2, compared to the global-scale driven fields, provides an overall improved representation of the climatological temperature by removing orography-dependent bias resulted in the coarser resolution simulations.

Data plotted in Figure 3 concern the capability of SFS ensemble members and ensemble means to reproduce the climatological period statistical distribution (PDF). For this analysis, climatological period spatial means of the three SFSs are considered.

Compared to reanalysis, CFSv2 ensemble produces a flatter-shape distribution with a consistent underestimation of the central distribution values frequency. At the same time, a frequency overestimation characterizes distribution tails, especially over the cold left-tail values. These features generally characterize the three forecast lead times. The downscaled SFSs similarly improve the representation of the winter climatology distribution. Both the central part of the distribution and the distribution tail values are better represented in the RegCM-d1 and RegCM-d2 simulations. No consistent differences between the two downscaled products are found. For what concerns the three different lead times, it can be noticed that the larger improvement from the higher resolution simulations occurs in the first forecast month (December). In fact, though improved when compared to the CFSv2, the central portion of the PDF in January and February is underestimated also in the two downscaled SFSs.

In Figure 4 data are plotted over time in order to assess the capability of reproducing the winter monthly mean inter-annual variability for the three SFSs compared to the reanalysis (dotted lines). Solid horizontal lines represent the climatological-period mean values for the three CFSs. Ensemble members spread for the three SFSs are shown using shaded colors. Reanalysis and SFSs are spatially averaged in correspondence of the inner domain centered over Central Italy.
Figure 3. Climatological temperature statistical distribution, shown as probability density functions (PDFs) of the three seasonal forecast systems ensemble means (tick solid lines). Green, blue, and red lines represent the CFSv2, RegCM-d1, and RegCM-d2, respectively. Multi-colored thin lines refer to the individual ensemble member PDFs for three forecast systems. Black tick solid lines represent the reanalysis PDF.

Figure 4. CFSv2 (left column), RegCM-d1 (central column), and RegCM-d2 (right column) ensemble mean inter-annual variability (dotted lines) and climatological period mean (solid lines) compared to the reanalysis (black lines). Shades represent ensemble spread (between lowest and highest monthly values) for the three seasonal forecast products.
In term of climatological means, CFSv2 seems to reproduce extremely well the reanalysis data over the three lead times. However, it must be considered that it could represent an artifact originated by spatially averaging the negative and positive biases shown in Figure 2. For what concerns the inter-annual variability, the driving CFSv2 largely underestimates the reanalysis variability. The CFSv2 capability of reproducing the temporal sequence of temperature above or below the mean is fairly poor and also characterized by a relevant ensemble spread which can reach values up to 7–8 °C. Downscaled SFSSs improve the information about the inter-annual variability, especially on the second part of the climatological period time series and on the 1-month forecast lead time. The ensemble spread is also consistently reduced. Coherently, where CFSv2 misses to represent extreme monthly values (i.e., largely departing from the climatological mean), the downscaled simulations also fail to represent anomalous monthly values. However, no incoherent signal is introduced in the higher resolution simulations. The anti-phase resulting in the CFSv2 time series is partially adjusted in the downscaled simulations. A comparison between the two downscaled products shows that RegCM-d2 reduces the bias over the mean climatology respect to RegCM-d1, confirming results shown in Figure 2. However, a large inter-annual variability underestimation characterizing also the two downscaled SFSSs can be observed.

That is due to the weak variability signal coming from the large scale produced by the driving CFSv2, whose poor quality is only marginally amplified in the downscaled SFSSs. This feature is particularly evident in the 2- and 3-month forecast lead time where, in the downscaled SFSSs, the inter-annual variability is deeply underestimated by the downscaled SFSSs as well. This suggests that, in the presence of a weak large-scale forcing signal, the regional scale simulations tend to develop its own dynamic, which is mainly driven by orographic forcing, and consequently disclosing poor inter-annual variability.

In order to summarize the inter-annual variability produced in the three SFSSs compared to the reanalysis, in Figure 5 we show the mean absolute error (MAE) of the monthly anomalies during the climatological period.

For each year we calculate, grid-point by grid-point, the difference between the monthly mean anomaly produced by each ensemble member considering its own climatology. The individual ensemble member anomaly is computed with respect to the member’s own climatology. In this way, the ability of reproducing the seasonal mean departure rather than the climate variable physical values is considered.

An overall improvement of the inter-annual variability of the anomalies can be similarly found in the two downscaled SFSSs. In fact, a consistent reduction of the MAE can be observed for the two dynamically downsampling especially over the 1 and 3 month forecast lead time.

A summary of the deterministic evaluation metrics assessed so far is shown in Table 1, as an overview of the skill of the three SFSSs in reproducing the four statistical moments (mean, 5th, 50th, and
95th percentiles) of the reanalysis climatological period. Also, the inter-annual anomaly MAE of the three SFSs joined to the ensemble spread (i.e., standard deviation of the individual ensemble member’s MAE) is shown the rightmost column. Here we can see again how the second dynamical downscaling procedure contributes in consistently reducing, roughly halving, the driving CFSv2 biases over the four statistical moments. At the same time, when considering the inter-annual anomaly MAE, the two dynamical downscaling seem to provide a very similar impact, leading to the same MAE reduction (0.3 °C/year).

**Table 1.** Mean absolute error (MAE) for daily mean 2 m temperature of the climatological period in the, 5th, 50th, and 95th percentiles, and inter-annual anomaly variability for the three SFSs compared to the reanalysis. In the rightmost column, the standard deviation represents the ensemble spread from the ensemble mean anomaly MAE. The five metrics are averaged over the winter season. Units are °C. See text for details.

<table>
<thead>
<tr>
<th></th>
<th>Mean-MAE</th>
<th>5p-MAE</th>
<th>50p-MAE</th>
<th>95p-MAE</th>
<th>Inter-Annual Anomaly-MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFSv2</td>
<td>1.7</td>
<td>2.2</td>
<td>1.7</td>
<td>1.5</td>
<td>1.8 ± 0.20</td>
</tr>
<tr>
<td>RegCM-d1</td>
<td>1.4</td>
<td>1.7</td>
<td>1.4</td>
<td>1.4</td>
<td>1.5 ± 0.19</td>
</tr>
<tr>
<td>RegCM-d2</td>
<td>0.8</td>
<td>1.0</td>
<td>0.8</td>
<td>0.8</td>
<td>1.5 ± 0.20</td>
</tr>
</tbody>
</table>

Figure 6. Cont.
Figure 6. Tercile plots for the CFSv2, RegCM-d1, and RegCM-d2 temperature winter months climatology (sub panels (a), (b), and (c) respectively). Shading of each square represents the portion of ensemble members predicting monthly values: below (≤1st tercile), normal (≥1st and ≤3rd tercile), and above (≥3rd tercile) the climatological mean of the relative ensemble member. Red dots show the corresponding reanalysis tercile for each year of the climatological period.
The effect produced on the inter-annual variability by the downscaling procedure is further investigated by means of probabilistic forecasts shown through tercile plots in Figure 6a–c for the CFSv2, RegCM-d1, and RegCM-d2 respectively. Here we can see the probabilistic skills of the CFSv2 and the two downscaled SFS simulations in reproducing the inter-annual sequence of anomalous conditions for the three forecast lead times. Each plot shows the year-by-year predicted probability (gray to black shade) of below-the average, normal, and above-the-average conditions represented by three terciles compared to the tercile produced by the reanalysis (red circles). The probability is here defined as the frequency (i.e., the portion) of the twelve ensemble members predicting one of the three terciles. More in detail, the first tercile bin includes monthly means ≤ 33rd percentile, second tercile monthly means ≥ 33rd and ≤ 66th percentile and third tercile monthly means ≥ 66th percentile. For each ensemble member, the three terciles are derived from its own specific climatological period statistical distribution. This analysis allows to investigate to what extent downscaling can modify the probabilistic inter-annual variability resulting from the global-scale CSFv2. While a general agreement between the global-scale and regional-scale can be seen, there are some cases where downscaled SFSs improve the probability of colder or warmer than average conditions mainly in December (e.g., 1998, 2000, and 2001) and very few times in February (e.g., 1993). In other cases, there is a reduction of probability of wrongly predicted anomalous conditions, again mainly in December.

At the same time, the downscaled SFSs are not able to significantly improve the probabilistic prediction, especially in the 2-month lead time (January). Overall, the downscaled SFSs seems to not show any substantial difference in the probabilistic forecast skill.

Finally, a second kind of probabilistic forecast evaluation metric is proposed through the reliability diagrams displayed in Figure 7a,b.

Reliability diagrams inform about the capability of a probabilistic forecast system to reproduce the observed frequencies of a particular event (i.e., occurrence of cold (warm) tercile temperature) conditioned to its forecast probability:

\[ p(y_i, o_j) = p(o_j | y_i) p(y_i) \]

where \( i \) is the forecast probability bins \((0, \ldots, 1)\) and \( j \), in the specific case of a binary event, can only take on a 1 or 0 value for occurrence (1) or not occurrence (0) of the considered event.

The first part of the factorization refers to the conditional distribution (i.e., frequency) of the observed event given the single probabilistic forecasts issued (i.e., given that the event was predicted at a certain probability, what has been observed?).

Let us suppose that the seasonal forecast probability for event \( e \) is equal to 0.9. In a reliable seasonal forecast system, \( e \) actually occurs in approximately 90% of the cases where \( e \) was predicted with a probability of 0.9. That is, a perfectly reliable forecast is when the predicted probability is equal to the conditional observed frequency, i.e., with the dots in the graph following the diagonal line. The second part of the factorization introduces the marginal distribution forecast, which represents the frequency of the forecast probability and provides information about the resolution (or sharpness) of the forecast.
Figure 7. Cont.
Reliability is derived as reference value the climatological period frequency of the considered event. In our case, considering the cold (lower) and warm (upper) terciles, the long-term climatological simulated and observed frequency of $e$ is one-third (0.3) indicated by the vertical and horizontal lines respectively. The dashed line identifies the delimitation between skillful and unskillful forecast. The area included between the dashed line and the vertical line represents the skillful forecast area. The latter is defined as the portion of the graph where the Brier skill score (BSS) is positive. BSS itself is defined as the mean squared error of the probability forecasts. Also, it is positive if the forecast is better of the reference forecast, which is given by the climatological period probability (0.3) (see e.g., [37,38]).

By comparing the reliability diagrams for the three forecasting systems, we can observe that for the cold (lower) tercile, the downscaling generally improves the climatological period reliability. More in detail, for the December month, according to the reference classification in [38], we can attribute a useful reliability category to both the downscaled systems. However, noteworthy improvement in the downscaled systems mainly characterizes the December month forecasts. In fact, in January, no substantial improvement can be observed compared to the CFSv2 which, in any case, provides a still marginally useful seasonal forecast. Reliability of the three systems sensibly decreases in the third lead time month forecast. In February, a negative relation between forecast probability and the conditioned observed frequency results shows that CFSv2 is not a useful forecast product. The downscaled systems only marginally improve the CFSv2’s poor results, in any case, without producing useful reliability.

Similar results have been obtained considering the reliability diagrams referred to in the warm (upper) temperature tercile. Also, in this case, we can observe a progressive degradation of the forecast reliability along the three lead times. It is also noteworthy how the downscaling is able to marginally improve the CFSv2 reliability but, again, mainly in the December month. Differently from the lower tercile results, downscaled systems are not able to improve the negative reliability characterizing the February month.

### 3.2. Precipitation Assessment

Before assessing the precipitation mean bias of the SFSs, it is worth considering the daily mean precipitation spatial pattern characterizing the climatological period for the reanalysis, CFSv2, RegCM-d1, and RegCM-d2 (Figure 8).
Figure 8. Spatial distribution of daily mean precipitation for the winter climatological period for: reanalysis, CFSv2, RegCM-d1, and RegCM-d2 (columns from left to right).

For the sake of comparison, all the three forecast products have been interpolated on a common resolution grid of 50 km, corresponding to the resolution of the reanalysis data set. From the reanalysis, December results on average the rainiest month (with values up to 4 mm/day). A relative drier area (1–2 mm/day) can be observed in the eastern side of Apennines ridge, which is known to exert a blocking effect to the humid flux of south-westerly air masses. The low resolution of the CFSv2 leads to a rough representation of the precipitation pattern. In this context, the poor representation of precipitation magnitude and spatial patterns drastically reduce the usability of current global-scale SFSs in any impact modeling research (e.g., hydrological modeling). The first RegCM downscaling introduces some spatial details, coherently to what shown by reanalysis. However, it could be also observed a consistent underestimation of spatial variability and also an underestimation of the precipitation on the western side of Apennines ridge. In this case also, the coarse resolution leads to a poor precipitation reproduction, being orography one of the most relevant precipitation forcing in this region. As expected, the orographical forcing on the precipitation pattern is consistently improved by the RegCM-d2. We can see a finer representation of the rainiest area located in the land points of the south-westerly portion of the domain.

In Figure 9, over the land points, the RegCM-d2 reduces the dry bias characterizing December and February months. Here, the biases are generally in the range of ±1 mm/day. However, the consistent dry bias (up to -3 mm/day) over sea grid points, found in the driving CFSv2, is preserved also in the downscaled SFSs, especially in the RegCM-d2 product. Furthermore, for the second downscaling, the dry bias over the borders of the domain corresponds to the numerical disturbance produced by the nesting also observed for the temperature case.
Figure 9. The same as in Figure 2 but for precipitation.

In Figure 9, over the land points, the RegCM-d2 reduces the dry bias characterizing December and February months. Here, the biases are generally in the range of ±1 mm/day. However, the consistent dry bias (up to -3 mm/day) over sea grid points, found in the driving CFSv2, is preserved also in the downscaled SFSs, especially in the RegCM-d2 product. Furthermore, for the second downscaling, the dry bias over the borders of the domain corresponds to the numerical disturbance produced by the nesting also observed for the temperature case.

Figure 10. The same as in Figure 3 but for the precipitation.

In Figure 10 the probability density functions (PDFs) of precipitation climatology for the ensemble means and each ensemble member for the three SFSs are compared to the reanalysis. Differently from what observed for the temperature, global CFSv2 well reproduces the reanalysis statistical properties and, at the same time, no clear improvements are shown by the downscaled SFSs. Looking at the ensemble mean, the RegCM-d2 is characterized by a slight overestimation of the low precipitation coupled to an underestimation of the high-to-heavy precipitation events frequency. This is particularly evident in December followed by February. In January, RegCM-d2 fairly outperforms the RegCM-d1.
In Figure 10 the probability density functions (PDFs) of precipitation climatology for the ensemble means and each ensemble member for the three SFSs are compared to the reanalysis. Differently from what observed for the temperature, global CFSv2 well reproduces the reanalysis statistical properties and, at the same time, no clear improvements are shown by the downscaled SFSs. Looking at the ensemble mean, the RegCM-d2 is characterized by a slight overestimation of the low precipitation coupled to an underestimation of the high-to-heavy precipitation events frequency. This is particularly evident in December followed by February. In January, RegCM-d2 fairly outperforms the RegCM-d1.

Figure 11 shows the spatial-averaged monthly mean time series for the climatological period. As for the temperature, the monthly mean inter-annual variability and climatological mean of the reanalysis are compared with the three SFSs.

CFSv2 ensemble mean heavily underestimates the climatological period mean (in December) and inter-annual variability over the three forecast lead times. A flatter inter-annual variability compared to the reanalysis is particularly evident for January and February. However, these two latter months shows an ensemble mean of the climatological period much closer to the reanalysis, even though characterized by a large ensemble spread. For what concerns the impact of the two downscaling, RegCM-d1 and d2 similarly fail to noticeably improve CFSv2 results. This is especially true for the inter-annual variability that is inherited by the driving CFSv2 and constantly underestimated on the three lead time forecast months. Downscaled simulations reduce the ensemble spread and RegCM-d2 ensemble mean climatology results very close to the reanalysis. As already partially observed for the temperature, and even more for the precipitation, the weak variability signal coming from the driving global model is preserved and not correctly amplified by the downscaling.
The lack of improvement in the forecast skill provided by the downscaled products is confirmed in Figure 12, that shows the ensemble anomaly MAE derived as described in the temperature results section.

![Figure 12. The same as in Figure 5 but for the precipitation.](image)

As already done for the temperature, the deterministic evaluation metrics for the precipitation are summarized in Table 2, where the three statistical moments (mean, 50th, and 95th percentiles) and the inter-annual anomaly MAE are shown. Again, differently from the temperature, both downscaling procedures fail to provide improvements if compared to the driving CFSv2. Also, we can see a degradation of the 95th percentile produced by the second downscaling (as shown in Figure 10). However, the misrepresentation of the physical values associated to the different statistical moments does not lead to a degraded capability of predicting the inter-annual anomaly variability. In fact, the driving CFSv2 and the two downscaling show the same value of the inter-annual anomaly MAE (rightmost column of Table 2), while providing a reduction of the associated inter-model spread (i.e., standard deviation of the ensemble members from the mean annual anomaly).

**Table 2.** Mean absolute error (MAE) for daily accumulated precipitation climatological period, 50th and 95th percentiles, and inter-annual anomaly variability for the three SFSs compared to the reanalysis. In the rightmost column the standard represents the ensemble spread from the ensemble mean anomaly MAE. The four metrics are averaged over the winter season. Units are mm/day. See text for details.

<table>
<thead>
<tr>
<th></th>
<th>Mean-MAE</th>
<th>50p-MAE</th>
<th>95p-MAE</th>
<th>Inter-Annual MAE-MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFSv2</td>
<td>0.7</td>
<td>0.2</td>
<td>3.2</td>
<td>1.5 ± 0.16</td>
</tr>
<tr>
<td>RegCM-d1</td>
<td>0.6</td>
<td>0.2</td>
<td>3.4</td>
<td>1.5 ± 0.13</td>
</tr>
<tr>
<td>RegCM-d2</td>
<td>0.8</td>
<td>0.2</td>
<td>4.6</td>
<td>1.5 ± 0.12</td>
</tr>
</tbody>
</table>

Also, through the probabilistic forecast analysis shown in the tercile plots of Figure 13, a substantial low capability of reproducing the inter-annual variability of precipitation anomalies is confirmed. A further confirmation is obtained when considering reliability diagrams (not shown) where none of the three systems seem to improve the probabilistic forecast obtained considering the no-skill climatological reference.

Finally, especially for the precipitation, we have to remark the intrinsic limits regarding the quality of the precipitation reanalysis when used as a reference product. In fact, as shown in [36], the coarse resolution of the data set strongly limits the representation of regional-scale precipitation, in particular over sea areas.
Figure 13. Cont.
Figure 13. The same as in Figure 6 but for the precipitation. Tercile plots for the CFSv2, RegCM-d1, and RegCM-d2 temperature winter months climatology (sub panels (a), (b), and (c) respectively).

4. Summary and Conclusions

This work investigates the potentially added value of dynamical downscaling in the context of regional scale seasonal climate forecasts over Central Italy. At this aim, a regional climate model (RegCM4.1) has been used to dynamically downscale global scale seasonal forecast provided by the NCEP-CFSv2. A two-step downscaling has been performed on twelve-member winter season (December, January, and February) ensemble for a 22-year reference period (1982–2003). The first one downscales CFSv2 from 100 to 60 km resolution over European domain (RegCM-d1) and the resulting simulations drive the second downscaling with a 12 km resolution over Central Italy (RegCM-d2). Winter season months temperature at 2 m and precipitation forecasts from the CFSv2 and the two downscaled forecasts have been assessed and evaluated considering NCEP reanalysis CFSR as reference product.

Mean bias spatial distribution, probability density functions (PDFs) reproduction, inter-annual anomaly mean absolute error (MAE), and probabilistic metrics represented tercile plots and reliability diagrams have been considered as quality measures. The three seasonal forecast products have been evaluated over the inner common domain over Central Italy.

The CFSv2 ensemble is affected by relevant temperature mean biases over the climatological period ranging from about $-5\,^\circ C$ (over sea grid points) to about $+6\,^\circ C$ (over Apennines grid points). The first downscaling partially reduces CFSv2 biases. However, significant warm biases still characterize the most complex orography areas. The largest improvement in winter temperature climatology reproduction is obtained through the second downscaling which succeeds in consistently removing the biases characterizing the coarser scale SFSs. Residual biases are generally enclosed between $-1$ and $+1\,^\circ C$. 

\[ \text{(c)} \]
Differently, when we consider the spatial-averaged climatological period statistical distribution, the two downscaling perform very similarly in adjusting both the distribution central portion and tails toward the reanalysis distribution.

The two downscaled ensembles monthly mean time series similarly show a reduction of the inter-model spread from the climatological period ensemble means. The reanalysis inter-annual variability is still consistently underestimated and for the larger part inherited by the driving global-scale CFSv2. At the same time, again similarly in the two downscaled simulations, a consistent decrease of the cumulative inter-annual anomaly MAE results, especially for December where the CFSv2 anomaly MAE of ~2 °C is reduced to 1.5 °C in the RegCM-d1 and RegCM-d2 SFs.

Considering probabilistic metrics (i.e., tercile plots and reliability diagrams) the following aspects have to be underlined:

(i) The reliability of the driving CFSv2 decreases along the forecast lead time with very poor performance in correspondence of the third-month lead time (February);

(ii) The downscaling marginally improved the reliability of the CFSv2 forecasts. This has been especially observed for December, where the downscaled simulations provide useful predictions according to the reference reliability classification in [38];

(iii) The beneficial impact of the downscaling results mainly in the forecasts of the lower (cold) tercile;

(iv) Coherently to what was obtained for the inter-annual variability analysis, the second downscaling provides similar skill compared to the first downscaling.

If we analyze the downscaling of precipitation, the increase of resolution does not generally improve the forecast quality in terms of both deterministic and probabilistic metrics. Only a slightly improved representation of the orography-driven precipitation spatial pattern can be observed in the second downscaling (Figure 8). A summary of the deterministic metrics shows that RegCM-d1 provides similar MAEs when compared to the CFSv2 while severe precipitation is substantially underestimated by the higher resolution RegCM-d2. However, it is noteworthy that anomaly inter-annual variability MAE is not degraded in both of the dynamical downscaling. Also, when considering probabilistic metrics (i.e., tercile plots), the downscaling does not provide any substantial improvement when compared to the global-scale CFSv2.

In agreement with recent studies [23,39,40], the precipitation results seem to confirm the evidence that an increased resolution, if on the one hand, at least in principle, ensures an improved physical values reproduction (e.g., smaller mean bias), on the other hand does not necessarily improve the forecast skill intended as a better representation of the inter-annual variability.

These findings underline that the more computationally expensive second downscaling entails, on the one hand, an improved representation of climatological period mean temperature physical values, but, on the other hand, fails to improve inter-annual variability compared to the first, computationally cheaper, downscaling for both temperature and precipitation. From an operational perspective, this evidence could suggest to aim at configurations where only one downscaling, at an intermediate resolution (e.g., 25 km), would be performed.

However, sensitivity tests (e.g., considering different cumulus convection schemes) are required to properly define the potential added value of a second downscaling, and will be performed in future studies.

In the context of the lack of noticeable added value provided by the precipitation downscaled simulations, we also acknowledge another relevant aspect represented by the nested runs spin-up time not considered in the present study. This latter represents a portion of simulation devoted to reach equilibrium between e.g., humidity, soil moisture, and heat exchanges, that can play a crucial role for the seasonal predictions in terms of temperature and especially precipitation patterns. However, mainly for the soil, this may take up to a few months and further increase the computing sources demand.

Further research efforts should also be devoted to test the seasonal predictability over the same domain, while deriving initial and boundary conditions from a different global-scale seasonal forecast
system (i.e., ECMWF, [13]). Also, while using the same global-scale driving global model, the impact of dynamical downscaling component can be tested through considering different high-resolution numerical models (i.e., AR-WRF, [41]).

Besides the dynamical approach also the capability of empirical-statistical downscaling techniques could also be explored. In this context recent studies (e.g., [40]) compared dynamical and statistical approaches. The latter, that can be discerned according to the model output statistics and perfect prognosis principles, have been shown to successfully reduce model biases during the calibration (past) period. However, it is worth underlining that the reduction of biases produced by statistical downscaling is a direct result of the calibration phase based on observations, where statistical downscaling is expected to reduce biases by construction. In fact, they have been found to provide irrelevant quality/skill improvement in terms of essential aspects of the inter-annual variability reproduction [39].

As already mentioned, in operational systems, the choice between the statistical and dynamical approaches to refine regional-scale climate information is a relevant strategic aspect. If, on the one hand, physical-based dynamical approaches generally, at least in principle, offer more consistent basis, they require a considerable investment in terms of computational time and facilities. On the other hand, statistical approaches are computationally cheap and can be performed without the need of particularly advanced computational infrastructures. However, it should be underlined that statistical and dynamical methods should not be seen as alternative, but rather an integrated approach should be preferred.

Regardless the downscaling approach considered (dynamical or statistical), it should be stressed that a consistent part of the improvements achievable in the seasonal forecasts at the regional scale will be in any case a function of improvements involving the driving large-scale. In this regard, in recent years the global-scale forecasting systems are showing an encouraging improvement to the representation and prediction of large-scale teleconnections. However, it is still not translating into an improved reliability of inter-seasonal anomalies variability over the mid-latitudes [42].


**Acknowledgments:** The Authors would like to thank the National Centers for Environmental Prediction for making publicly available the reanalysis data sets (NCEP/CFSR Data Archives at NCAR; https://rda.ucar.edu/#!lfd?nb=y&projv=NCEP%20Climate%20Forecast%20System%20Reanalysis) and for making publicly available the forecasts of the NCEP Climate Forecast System Version (https://nomads.ncdc.noaa.gov/modeldata/) which have been used in this study as initial and boundary conditions for the dynamical downscaling. Authors would also like to thank the International Centre for Theoretical Physics (ICTP) for making publicly available climate model RegCM code which has been compiled and run on the internal computing resources at CETEMPS (University of L’Aquila). The Authors thank the Santander Meteorology Group for making publicly available the MATLAB MeteoLab tool code. The authors would also like to thank the anonymous reviewers for their helpful and constructive comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).