Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

Article

Climatic extrapolations in hydrology: the expanded Bluecat methodology

Demetris Koutsoyiannis [1,*] and Alberto Montanari [2]

[1] Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Zographou, Greece
[2] Department DICAM, University of Bologna, Bologna, Italy; alberto.montanari@unibo.it
* Correspondence: dk@itia.ntua.gr

Abstract: Bluecat is a recently proposed methodology to upgrade a deterministic model (D-model) into stochastic (S-model), based on the hypothesis that the information contained in a time series of observations and the concurrent predictions by the D-model is sufficient to support this upgrade. Prominent characteristics of the methodology are its simplicity and transparency, which allow easy use in practical applications, without sophisticated computational means. Here we utilize the Bluecat methodology and expand it in order to be combined with climatic model outputs, which often require extrapolation out of the range of values covered by observations. We apply the expanded methodology to the precipitation and temperature processes in a large area, namely the entire territory of Italy. The results showcase the appropriateness of the method for hydroclimatic studies, as regards the assessment of the performance of the climatic projections, as well as their stochastic conversion with simultaneous bias correction and uncertainty quantification.

Keywords: Bluecat; climate models; stochastics; uncertainty.

1. Introduction

Hydrologists often need to consider in their studies climatic projections for the future. Such need comes from concerns about the future climate, the idea that "stationarity is dead" [1] (which however has been disputed, e.g. [2,3]), and the obligation to use climatic predictions for technical design of climate adaptation strategies. There is no doubt that the future climate would change (as ever) and therefore the assessment of the technical performances of climate models at local time and spatial scale is an important research topic [4-10].

For the above reasons, methods for extracting useful technical information from climate models are getting increasing attention. In this respect, Tyralis and Koutsoyiannis [11] have developed a Bayesian methodology for extracting such information and providing a stochastic framework of future climate based on the observations on the one hand and conditional on the climate model outputs on the other hand.

However, this methodology is rather complicated, which prohibits its use by practitioners. Here we provide a simple methodology that can be applied without sophisticated computational means. The methodology stems from hydrological modelling—in particular from deterministic hydrological models which are upgraded to stochastic (see section 2). Here we illustrate the methodology in hydroclimatic applications, in which the deterministic models are climate models (else known as general circulation models), while the stochastic model is built upon reanalysis data representing reality; both model outputs and reanalysis data are readily provided on the internet (section 3). Climate models provide outputs for atmospheric variables, such as temperature and precipitation, which in turn are used as inputs in hydrological modelling. Therefore, the usefulness of the study relies on (a) the assessment of how well climate model outputs represent reality, (b) the correction of the model outputs to approach reality, (c) the assessment of the involved uncertainty by means of comparing model outputs to reality (instead of the more common

© 2022 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

practice of comparing different models with each other), and (d) the potential of supporting the construction of a stochastic representation of climatic inputs for use in hydrologic modelling. This usefulness with respect to items (a)-(c) is illustrated by application of the methodology in a large area, namely the entire territory of Italy (section 4), while a study focused on item (d) is planned for future research.

2. The Bluecat and its expansion to deal with climatic projections

Hydrological models are transformations of inputs x_τ (e.g. rainfall) at discrete time τ to outputs Q_τ (e.g. river discharge) by means of a model, expressed as:

$$Q_\tau = G(\boldsymbol{x}_\tau) \tag{1}$$

where $\boldsymbol{x}_\tau$ is a vector containing a number of consecutive input variables, or even a matrix consisting of several input variables (such as rainfall, evapotranspiration, perhaps river discharge in an upstream basin, etc.). The transformation function is generally a complicated one, also involving additional state variables (e.g. soil moisture).

The climatic models do not differ from this logic, even though they are much more complicated. A model is never identical to reality and thus the true value of the output q_τ will be different from the model prediction Q_τ.

In their blueprint, Montanari and Koutsoyiannis [12] provided a framework to upgrade a deterministic model into stochastic, which provides the probability distribution of the output given the inputs and the deterministic model output, considering the uncertainty in model parameters and in input variables. This work has been discussed [13,14] and advanced in other studies [15-17].

In a recent follow-up paper, Koutsoyiannis and Montanari [18] proposed a simple, easy to use and transparent methodology to upgrade a deterministic model into stochastic based on the data only, which they named Bluecat. The basic hypothesis is that the information contained in the true outputs q_τ and the concurrent predictions by the deterministic model Q_τ is sufficient to support this upgrade. Simplicity is a primary objective of this methodology, which does not involve multiple simulations, likelihood estimation, Bayesian methods etc. Rather it uses a computational framework that can run even in a worksheet software. Here we utilize this latter methodology and expand it in order to be combined with climatic model outputs, which often require extrapolation out of the range of values covered by observations.

In upgrading a deterministic model into stochastic, all related quantities should be regarded as stochastic (random) variables and their sequences as stochastic processes. For notational clarity we underline stochastic variables, stochastic processes and stochastic functions. We use non-underlined symbols for regular variables and deterministic functions, as well as for realizations of stochastic variables and of stochastic processes, where the latter realizations are also known as time series.

We assume that the inputs $\underline{x}_\tau$, at discrete times τ, have a stationary probability density function $f_x(x)$ and distribution function $F_x(x)$. The output $\underline{q}_\tau$ depends on the inputs $\underline{x}_\tau$ and is given though some stochastic function (S-model) as:

$$\underline{q}_\tau = \underline{g}(\underline{\boldsymbol{x}}_\tau) \tag{2}$$

The stochastic process $\underline{q}_\tau$ is assumed to correspond to the real process, while the outcome of the deterministic model (D-model) of equation (1) is an estimate thereof. By writing the latter equation in stochastic terms, retaining however the function G ($\neq \underline{g}$) as a deterministic function, we obtain the estimator $\underline{Q}_\tau$ of the output $\underline{q}_\tau$ as:

$$\underline{Q}_\tau \coloneqq G(\underline{\boldsymbol{x}}_\tau) \tag{3}$$

To advance from the D-model, in its form (3), to the S-model in (2) we just need to specify the conditional distribution:

$$F_{q|Q}(q|Q) = P\left\{\underline{q} \le q | \underline{Q} = Q\right\} \tag{4}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

with q and Q assumed concurrent and referring to discrete time τ. In other words, here conditioning is meant in scalar setting. An extension where Q is a vector containing the current and earlier predictions by the D-model is possible but more laborious, thus not complying with our simplicity target.

It is relatively easy to infer from data the marginal distribution and density functions of the S-variable $\underline{q}$ and D-predicted variable $\underline{Q}$. Therefore, we may assume that $f_q(q)$ and $f_Q(Q)$ are known. Specifically, if we have time series of concurrent Q and q, each of size n, and if $Q_{(i:n)}$ is the ith smallest value in the time series of Q and $q_{(j:n)}$ is the jth smallest value in the time series of q, then we can use the approximations

$$F_Q\left(Q_{(i:n)}\right) \approx \frac{i-0.439}{n+0.122}, \qquad F_q\left(q_{(j:n)}\right) \approx \frac{j-0.439}{n+0.122} \tag{5}$$

which provide an unbiased estimate of the logarithm of the excess return period, or, equivalently, the quantity $-\ln(F/(1-F))$ (adapted from Koutsoyiannis [19], Table 5.5).

Then the conditional density sought should obey [18]:

$$\int_{-\infty}^{\infty} f_{q|Q}(q|Q)\mathrm{d}q = 1, \qquad \int_0^1 f_{q|Q}\left(q\middle|F_Q^{-1}(z)\right)\mathrm{d}z = f_q(q)\,, \qquad \int_0^1 F_{q|Q}\left(q\middle|F_Q^{-1}(z)\right)\mathrm{d}z = F_q(q) \tag{6}$$

However, determining $F_{q|Q}(q|Q)$ from the data alone without assuming a specific expression for the distribution is not that easy. As the variables of interest are usually of continuous type, we may expect that each value Q_τ in the available time series appears only once. Thus, we cannot form a sample for a particular value of Q. Koutsoyiannis and Montanari [18] proposed the following simple approximation of $F_{q|Q}(q|Q)$, by using a sample of Q-neighbours:

$$\begin{aligned} F_{q|Q}(q|Q) = P\left\{\underline{q} \le q \middle| \underline{Q} = Q\right\} &\approx P\left\{\underline{q} \le q \middle| Q - \Delta Q_1 \le \underline{Q} \le Q + \Delta Q_2\right\} = \\ &\approx P\left\{\underline{q} \le q \middle| F_Q(Q) - \Delta F_1 \le F_Q\left(\underline{Q}\right) \le F_Q(Q) + \Delta F_2\right\} \\ &=: F_{q|[Q]}(q|Q, \Delta F_1, \Delta F_2) \end{aligned} \tag{7}$$

where the increments ΔQ_i and ΔF_i can be chosen based on the requirement that the intervals below and above the values Q (or $F_Q(Q)$) contain appropriate numbers of data values, $m_1 := \Delta F_1 n$ and $m_2 := \Delta F_2 n$, respectively. The numbers m_1 and m_2 should not be too large, so that $F_Q(Q) \pm \Delta F_{1,2}$ be close to $F_Q(Q)$, nor too small, so that the probability $P\left\{\underline{q} \le q \middle| F_Q(Q) - m_1/n \le F_Q\left(\underline{Q}\right) \le F_Q(Q) + m_2/n\right\}$ can be estimated empirically, from a sample of size $m_1 + m_2 + 1$, as reliably as possible. Koutsoyiannis and Montanari [18] generally used $\Delta F_1 = \Delta F_2 = \Delta F$ and, hence, $m_1 = m_2 = m$, except for the m lowest and the m highest values of Q, for which they developed a special procedure, which however is not able to extrapolate beyond the lowest and highest Q values.

Here, we modify the procedure, with the aim to extrapolate beyond the range of Q values. First, we generalize the use of equal sizes of left and right subsamples, i.e. $m_1 = m_2 = m$, so that the subsample $[Q]$ always have size $2m + 1$. Similar to equation (5),

$$F_{q_{(i:2m+1)}|[Q]}(q|Q) = \frac{i-0.439}{2m+1+0.122} \tag{8}$$

Furthermore, we use a constant m in all cases. For example, setting $m = 50$, i.e. $2m + 1 = 101$, the lowest empirical probability we can estimate would be 0.55% and the highest one 99.45% for $i = 1$ and 101, respectively. However, these estimates are too uncertain; a more reasonable choice for $m = 50$ would be $i = 5$ and 97, resulting in probability estimates 4.5% and 95.5%, respectively.

Apparently, this procedure cannot work for $Q < Q_{(m+1:n)}$ and $Q > Q_{(n-m:n)}$. For these cases we choose a constant $c \neq 1$, so that $Q_{(m+1:n)} \le cQ \le Q_{(n-m:n)}$ and estimate the conditional distribution by enrolling the relationship

$$F_{q|Q}(q|Q) \approx F_{q|Q}(q + a(cQ)(1-1/c)|cQ) \tag{9}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

where a is a parameter representing a regression slope between Q and q. It can be shown that equation (9) is precisely valid if the process of interest is Gaussian. If not, it can be used as an approximation. The approximation can be improved by transforming the processes to Gaussian, e.g. using the transformation

$$q' = \lambda \ln\left(1 + \frac{q}{\lambda}\right) \tag{10}$$

where q and q' are original and transformed variables, respectively (and likewise for Q), and λ is a parameter. For $\lambda \to \infty$ and $\lambda \to 0$ equation (10) becomes equivalent to the identity and the logarithmic transform, respectively. Yet, in comparison to the standard logarithmic transform, equation (10) has the advantage that it maps zero to itself.

If Q_{L} and Q_{H} are the lowest and highest values for which estimates are sought, then we define two values c_{L} and c_{H} so that $c_{\mathrm{L}}Q_{\mathrm{L}} = Q_{(m+1:n)}, c_{\mathrm{H}}Q_{\mathrm{H}} = Q_{(n-m:n)}$. Likewise, we estimate two values (slopes) a_{L} and a_{H} by regression on the lowest and highest part (say quarter or third) of the entire sample of size n. Using the values $c_{\mathrm{L}}, c_{\mathrm{H}}, a_{\mathrm{L}}, a_{\mathrm{H}}$ along with equation (9) we are able to extrapolate below and above the values $Q_{(m+1:n)}$ and $Q_{(n-m:n)}$, respectively.

Bluecat has been tested in synthetic cases with a priori known conditional and marginal distributions [20] as well as in real-world hydrological modelling cases [18]. The results have been satisfactory as it always succeeds in its two targets, namely:

- to appropriately correct the D-model bias, which may differ for different ranges of Q, and
- to infer the model uncertainty in terms of confidence bands, whose width may also differ for different ranges of Q.

3 Data

To apply Bluecat in combination with climate models, we choose as a case study area the entire territory of Italy. The reason we chose a big area, rather than a more hydrologically relevant unit, such as a specific catchment, is the known fact [8,10] that the performance of climate models improves for increasing spatial scale. As case study variables of hydrological interest, we choose temperature and precipitation, both areally averaged across the entire Italian territory.

As actual values of the processes of interest we regard those provided by the NCEP-NCAR Reanalysis 1 [21], jointly produced by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR). Its temporal coverage includes 4-times daily, daily and monthly values for 1948 to present at a horizontal resolution of 1.88° (~ 210 km at the equator). It uses a state-of-the-art analysis/forecast system to perform data assimilation using observations and a numerical weather prediction model. The data assimilation and the model used are identical to the global system implemented operationally at NCEP except in the horizontal resolution. A large subset of these data is available as daily and monthly averages (https://www.esrl.noaa.gov/psd/cgi-bin/data/testdap/timeseries.pl). Here the data have been retrieved from the climexp platform (http://climexp.knmi.nl/, section: monthly reanalysis fields), using as geographical mask that of Italy, readily provided by the platform. The time series are plotted in Figure 1 (temperature) and Figure 2 (precipitation) for their entire time period (1948-2021).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

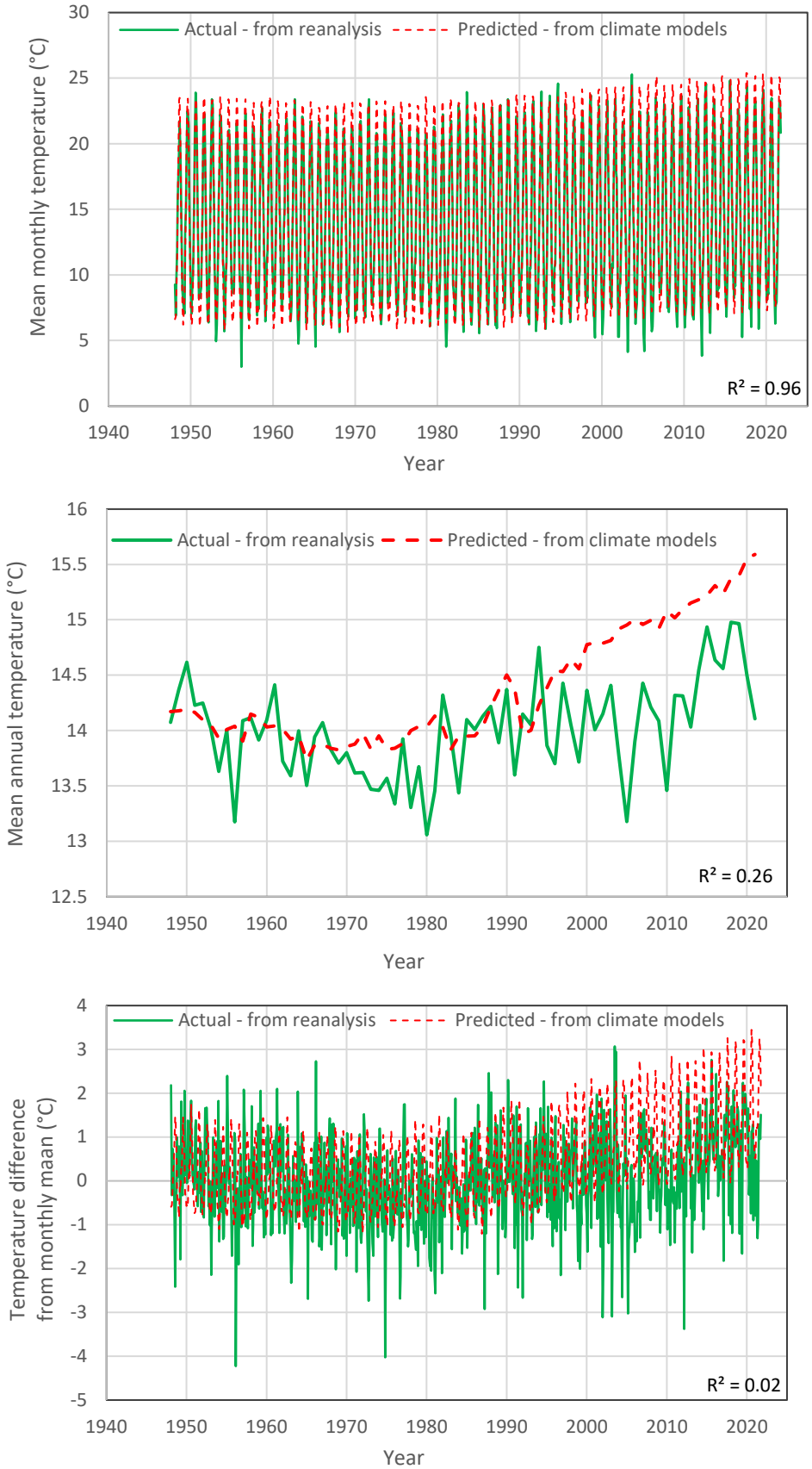

Figure 1. Comparison of actual and D-model temperature data (upper): original monthly data; (middle) after aggregation to annual scale; (lower) after subtraction of monthly means to reduce the effect of periodicity.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

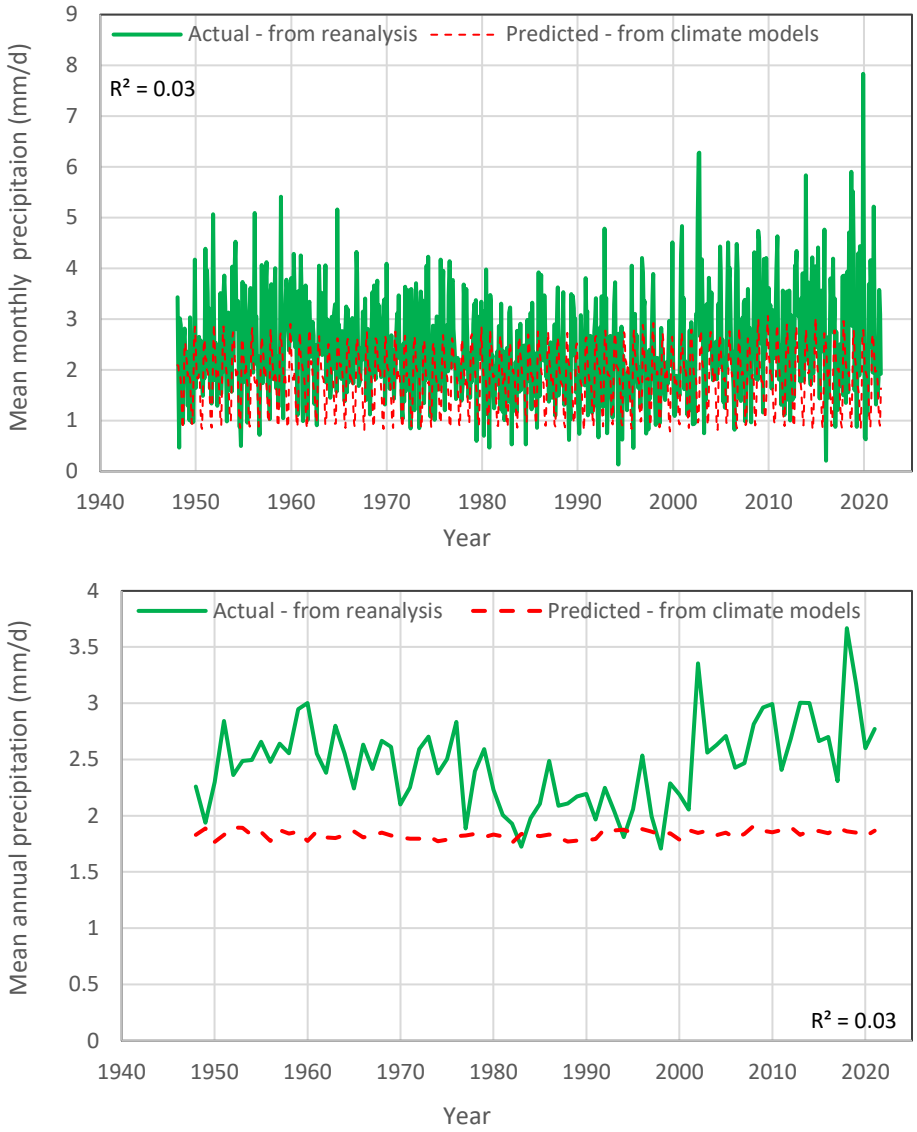

Figure 2. Comparison of actual and D-model precipitation data (upper): original monthly data; (lower) after aggregation to annual scale.

As a deterministic model we use the mean of the output data of the Coupled Model Intercomparison Project (CMIP6), noting that nothing would change in the methodology if we chose any particular member of the ensemble instead of the mean. Among the scenarios provided, we use the Scenario Shared Socio-Economic Pathways 245 (SSP245) The model outputs also go back to the past, extending over the time period 1850-2100. They have again been accessed through the climexp platform (option: Monthly CMIP6 scenario runs; Case: CMIP6 mean).

A comparison of the model outputs for temperature is shown in Figure 1 for the common period of the datasets. In the upper panel, depicting the monthly data, it appears that the model agrees well with the actual temperature, with a very high coefficient of determination ($R^2 = 0.96$). This results from the fact that the model captures the annual periodicity, as materializing at the specific geographical location of Italy. If we average the process at the annual scale to eliminate the effect of periodicity (middle panel), then the agreement between the two datasets deteriorates ($R^2 = 0.26$). In particular, in the last years the models predicted too much temperature increase which does not correspond to reality. The lower panel of the Figure 1 also shows a comparison after reducing the effect of periodicity but at the monthly scale; this was done by subtracting the original values from the actual monthly means. Now the D-model data show little correspondence to reality ($R^2 =$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

0.03). Yet the Bluecat method can be applied without problems, as it has been tested even with models irrelevant to reality.

Likewise, Figure 2 shows similar comparisons for precipitation. Now the agreement of D-model with reality is poor both on the annual and the monthly scale ($R^2 = 0.03$). In particular, at the monthly scale the model exhibits too low variability and at the annual scale it shows a horizontal line of stagnancy which is different from what actually happens.

3. Results

For the temperature case, we apply the Bluecat to the monthly differences from means to avoid the construction of a more complex cyclostationary version of Bluecat, which would substantially reduce the available size of the dataset (e.g. in a monthly cyclostationary model, we would have 1/12 of the data for each month). Figure 3 shows the results of the application of Bluecat to temperature data. The graph depicts the true value and the S-prediction in terms of confidence limits vs. the D-prediction. The median S-prediction is also shown. The confidence limits are for confidence coefficient 90% ($F = 0.05$ and 0.95 for the low and high prediction, respectively). With reference to equation (8), to estimate the value corresponding to a certain F we solve the equation for i and round it to the nearest integer, and then calculate the q value as that corresponding to the ith smallest value in the sample of $2m + 1$ values. If the D-model were a good representation of reality the median line would be close to the equality line and the confidence band would be narrow. In contrast, a totally irrelevant representation, would result in a horizontal straight line with wide confidence band. Here the situation is in between these two extreme cases. Yet the performance of the D-model is not very good as the equality line is not enveloped by the confidence limit for high Q-predictions.

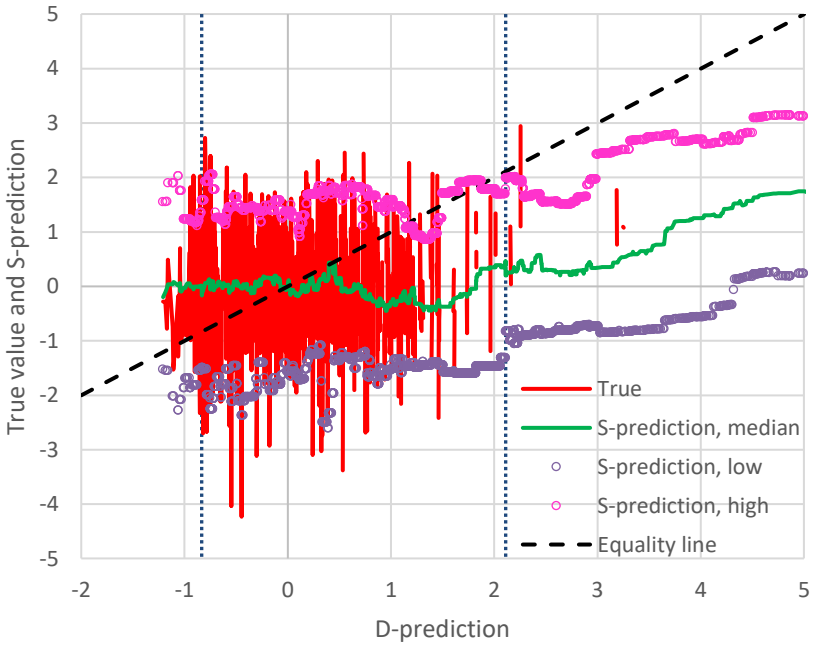

Figure 3. True values and S-model predictions vs. S-model predictions for temperature, along with confidence limits. The vertical dotted lines define the area out of which extrapolation is necessary.

Therefore, as shown in Figure 4, the evolution resulting from the S-model is quite different from that of the D-model. In the upper panel of the figure, it is seen that the increase rate of future temperature according to the S-model is smaller than that predicted by the D-model. The reason why this happens becomes obvious in the lower panel of the figure, which focuses on recent years. Specifically, in the past ten years, the D-model has already departed from reality and, therefore, the departure increases in the future.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

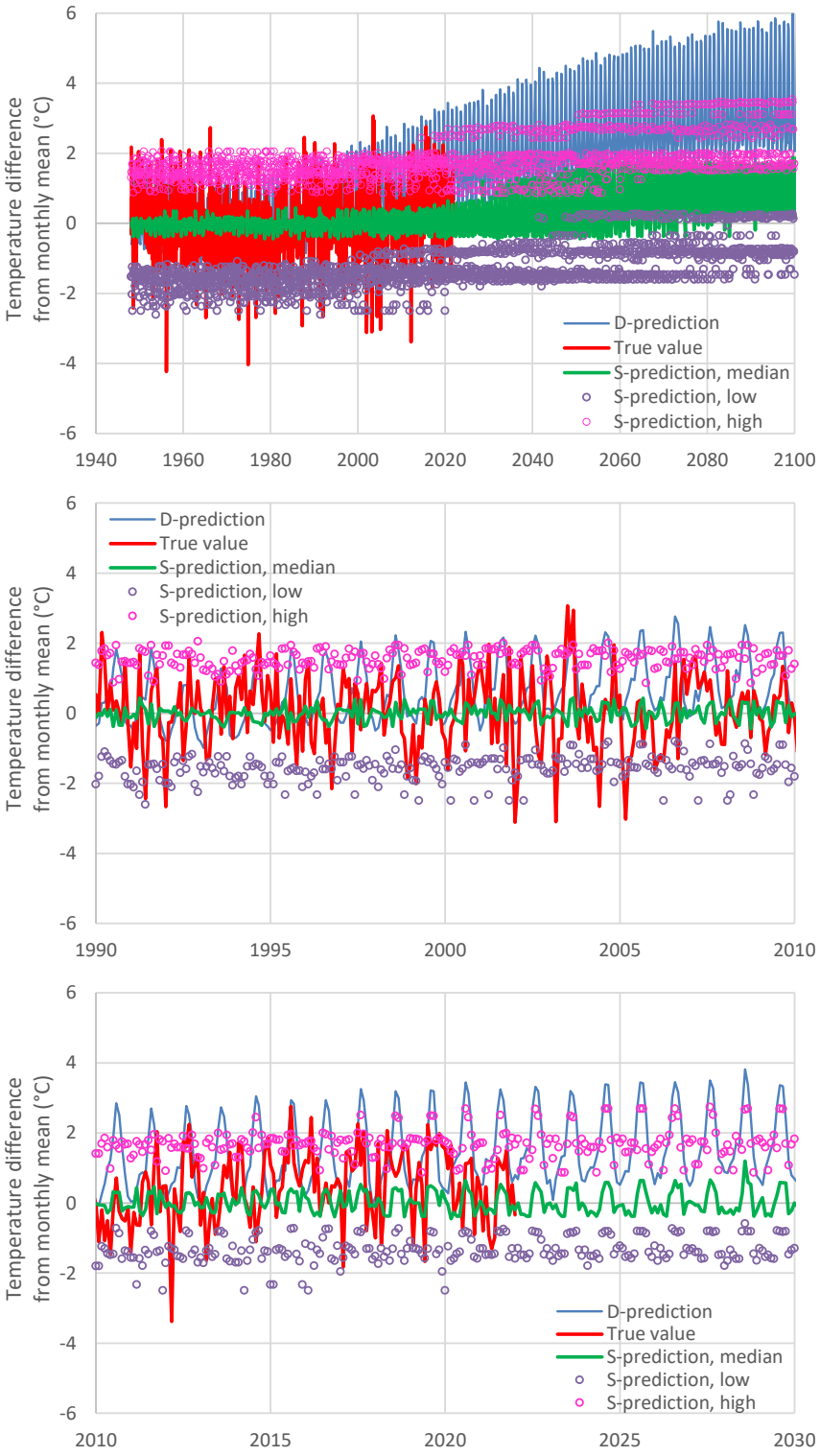

Figure 4. True values and predictions by D-model and S-model (median and 90% confidence limits) of temperature in Italy; (upper) entire period; (middle and lower) focus on 20-year periods.

Another notable feature in Figure 4 is the zig-zag shape of all predictions (D and S). This can easily be attributed to excessive intensity of periodicity in the D-model. A simple remedy of this problem is shown in Figure 5, where the monthly S-model values have been replaced by running averages with a 12-month time step.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

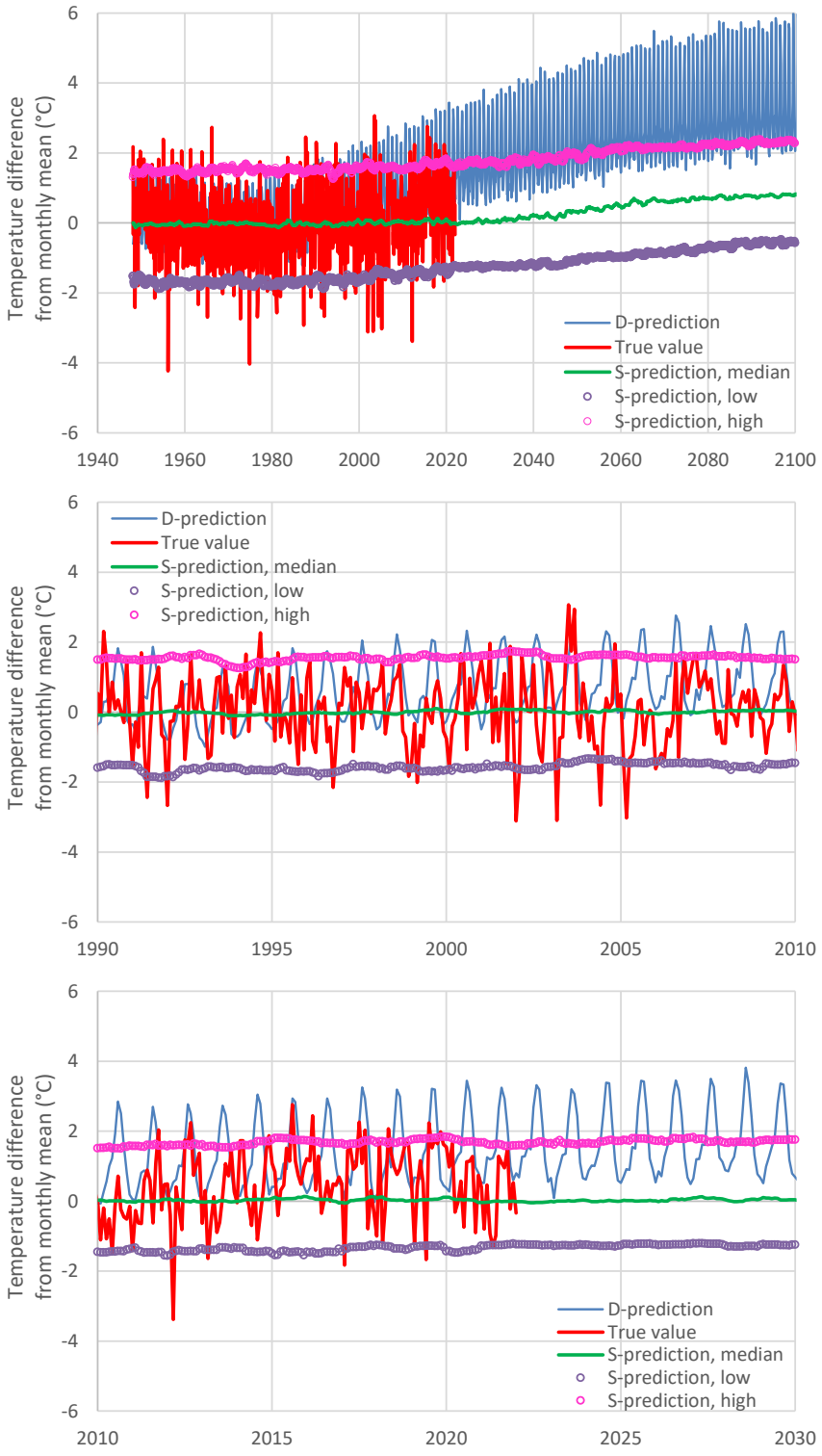

Figure 5. As Figure 4 but with S-model results (median and 90% confidence limits) replaced by running averages over 12 months.

One might potentially think to attribute the difference of the temperature increase rates of D-model and S-model to the fact that the latter is stationary, while the former is not. However, this is not the case. The actual reason of the difference is the departure of the model from reality in the latest period. In is easy to demonstrate that the Bluecat would behave well if the D-model had better performance, which means that its stationary formulation is not a problem. This is graphically depicted in Figure 6 (upper panel), where to make the D-model closer to reality, we replaced the D-model series with the weighted

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

sum of true and D-model values with weights 0.75 and 0.25, respectively. As seen in the figure, in this case the S-model remains very close to the D-model even in the distant future (2100). While the entire methodology is based on stationarity, it captures the trendy shape of the D-model, provided that the latter is close to reality.

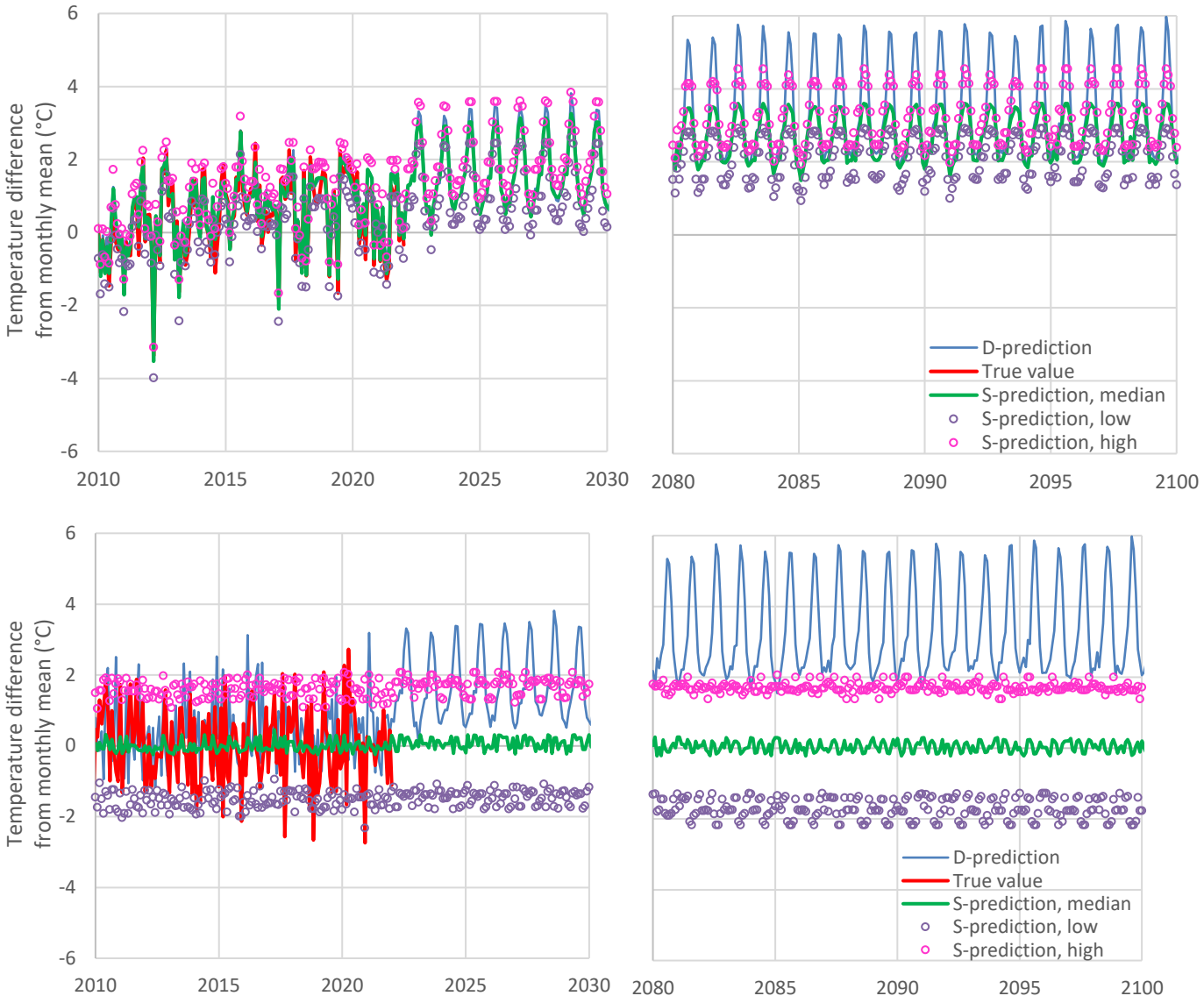

Figure 6. Modified Figure 4 for two hypothetical cases, in which the D-model values are reconstructed so as to be (upper) very close to reality, by replacing the D-model series with the weighted sum of true and D-model values with weights 0.75 and 0.25, respectively, and (lower) completely irrelevant to reality, by randomly rearranging the time order the true values.

For completeness of the investigation, the lower panel of Figure 6 depicts a case where the D-model if totally irrelevant to reality. To implement this hypothetical case, we reordered the D-model series at random, so as to become uncorrelated to the true values. As expected, in this case the S-model in effect disregards the D-model, producing horizontal lines of the future evolution, as expected for a stationary model.

For the precipitation case, in which the seasonal variation in Italy is not prominent, we apply the Bluecat to the original monthly values. A second difference from the temperature case is that the rainfall distribution is far from normal and therefore we have used transformation in equation (10) to normalize it. The resulting value of parameter λ is 2 mm and the transformation indeed yielded a perfect fit to the Gaussian distribution. Figure 7 shows the results of the application of Bluecat to transformed precipitation data.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

The graph depicts the true value and the S-prediction in terms of confidence limits vs. the D-prediction. The median S-prediction is also shown while the confidence limits are again for confidence coefficient 90% ($F = 0.05$ and 0.95 for the low and high prediction, respectively). It is manifest in the figure that for low precipitation values the D-model is poor—the median S-model prediction is flat—but this is improved for the highest predicted precipitation values, where the median line is close to the equality line. As already discussed (Figure 2), the D-models underpredicts the actual variability of precipitation. Clearly, as shown in Figure 7, this is corrected by Bluecat (compare the ranges of plotted values on the horizontal and vertical axes).

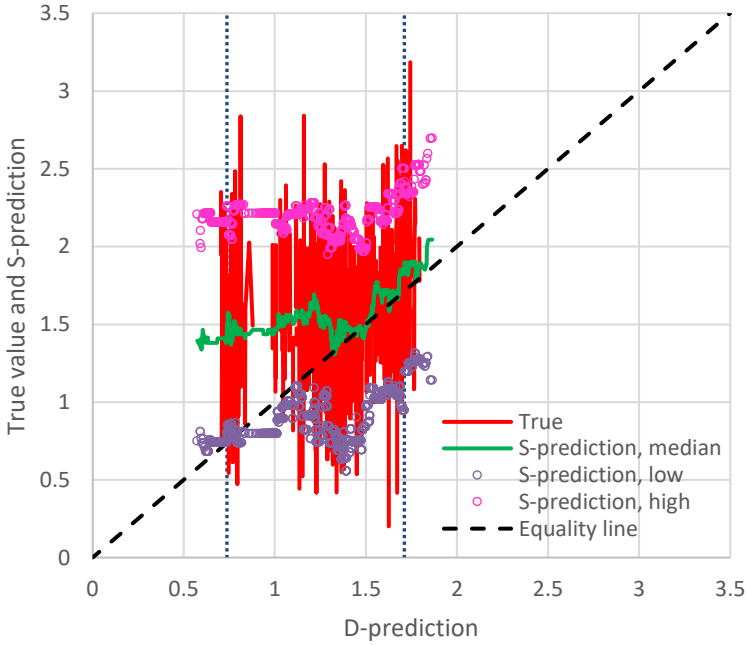

Figure 7. True values and S-model predictions vs. S-model predictions for precipitation, transformed for normalization, along with confidence limits. The vertical dotted lines define the area out of which extrapolation is necessary.

The evolution resulting from the S-model is shown in Figure 4, after back-transforming to the natural rainfall vales. The substantial effect of the S-model in this case is that it widened the range of variability and corrected the bias. Generally, the D-model did not provide any substantial information and, even if we disregarded it, the results would not differ much. It must be further noted that, as the required range for extrapolation is not wide (see the vertical dotted lines in Figure 7), even without using the normalizing transformation, the results are practically the same (not included in figures as the lines would be indistinguishable from those already shown).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

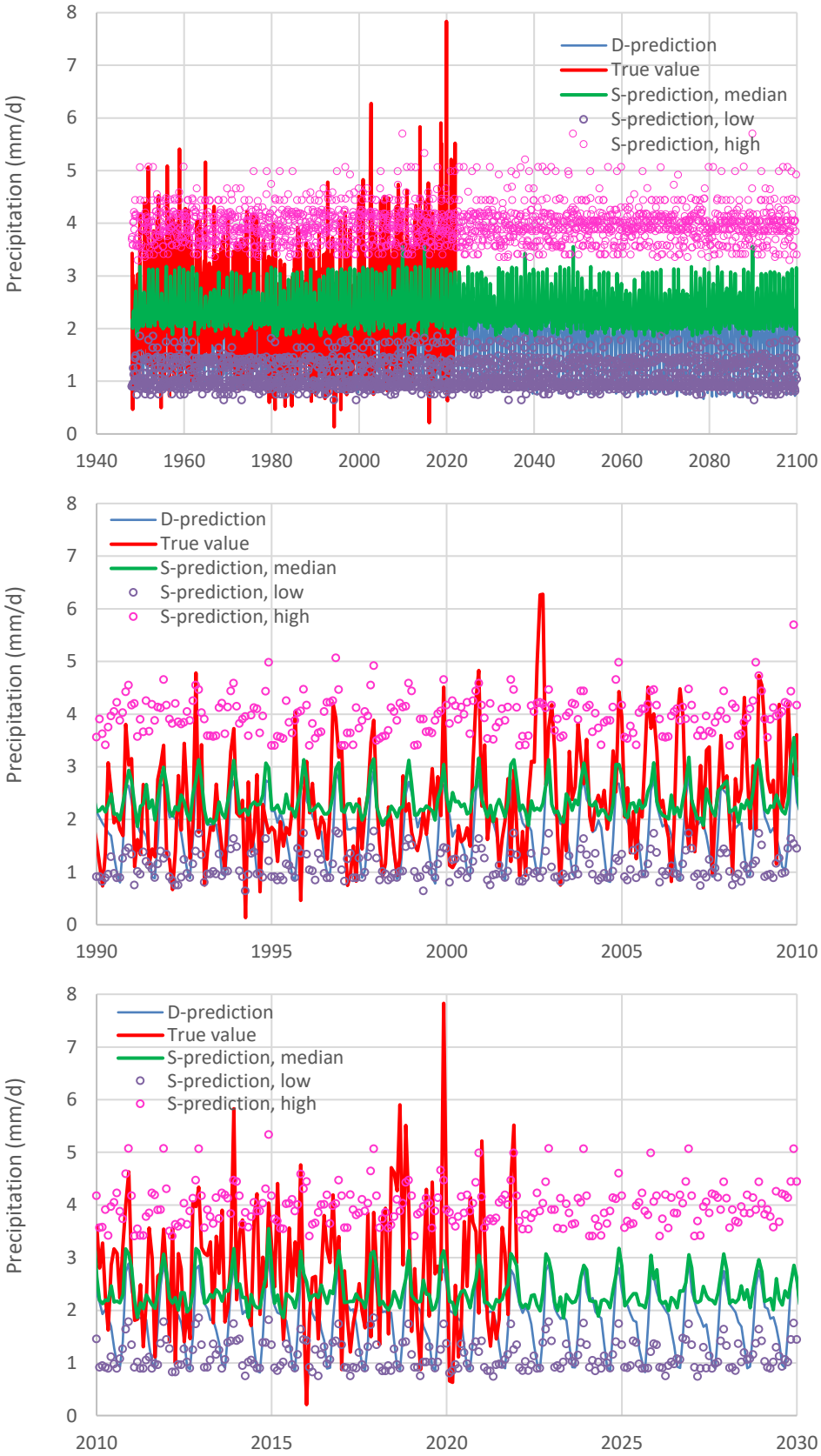

Figure 8. True values and predictions by D-model and S-model (median and 90% confidence limits) of precipitation in Italy; (upper) entire period; (middle and lower) focus on 20-year periods. (Nb. a normalizing transformation by equation (10) with $\lambda = 2$ mm/d is used, while the plotted values are back transformed.)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 April 2022 doi:10.20944/preprints202204.0260.v1

4. Discussion and conclusions

The Bluecat methodology, which was initially proposed as a hydrological modelling tool, appears to provide scientific means for incorporating climatic predictions within hydrological modelling. Its theory is simple, transparent and easily understandable. Its application is also very simple as it does not require any sophisticated computational means. These characteristics make it an ideal tool for enhancing collaboration of scientists with practitioners and decision makers to make a successful contribution to society.

Traditionally, the task of combining climatic models within hydrology is termed "downscaling", as hydrological models require finer scales than those resolved in climatic models. An additional (and most important) job of downscaling techniques is to adjust climate model outputs (such procedures are sometimes termed "bias correction" of statistics) making them consistent with local measurements. In this, downscaling techniques resemble the proposed method. However, our method has substantial differences from downscaling techniques, including the following:

- The scales of application of the Bluecat's S-model are not necessarily different from that of the climatic D-model. They could be precisely the same, as in the present application to Italy. In this respect, the Bluecat makes the D-model consistent with reality irrespective of spatial scale. Apparently, the case where the scales are different (e.g. smaller area of the S-model with respect to that corresponding to the D-model) is also served by the proposed methodology, without any change with respect to what is described above.
- The Bluecat methodology considers both the time sequence and the amplitude of the D-model and actual series. It does not make a lumped fitting on the entire set of past data to find unique parametric relationships to be applied to adaptation of the future values (as usually done in downscaling techniques). Nor does it regard the future values as correct that only need downscaling. Rather, it treats past and future values produced by the D-model in the same manner—as representing a model and not the truth.
- In addition to modifying a D-model series, correcting it for bias, the Bluecat advances it to a stochastic representation, thus characterizing the uncertainty which is illustrated here in terms of confidence bands.

In fact, the stochastic representation is much more than drawing confidence bands. Its real power emerges in stochastic simulation, by which a system can be tested in a Monte Carlo approach so that the uncertainty of the process of actual interest (e.g. river stage or discharge, reservoir outflow, inundated area) is eventually evaluated. This has not been included in the scope of this paper but is currently under research, whose results will be reported in the future.

Some features of the Bluecat framework have not been included in this study in order to keep it as simple as possible. These features include a more sophisticated calibration approach, relying on a split-sample technique with different calibration and validation periods, which was not necessary here. Also, they include the concept of knowable moments (K-moments [19]) for a more robust estimation of empirical distribution quantiles, while it is also useful for the construction of a simulation framework.

Author Contributions: Conceptualization, DK and AM; methodology, DK and AM; software, DK and AM; validation, DK and AM; formal analysis, DK and AM; investigation, DK and AM; resources, DK and AM; data curation, DK and AM; writing—original draft preparation, DK and AM; writing—review and editing, DK and AM; visualization, DK and AM; supervision, DK and AM; project administration, DK and AM; funding acquisition, n/a. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding but was conducted for scientific curiosity.

Data Availability Statement: All data used in the study are freely available and, as described in the article, they were retrieved from the climexp platform. The NCEP–NCAR reanalysis data are described in https://www.esrl.noaa.gov/psd/cgi-bin/data/testdap/timeseries.pl. The Coupled Model Intercomparison Project (CMIP6) model data are produced by CSIRO is licensed under a Creative

Commons Attribution-ShareAlike 4.0 International License (https://creativecommons.org/licenses/).

Acknowledgments: We thank the Editor-in-Chief Ezio Todini and the Editor Tammo Steenhuis for the honourable invitation of this article. We are grateful to the organizations and colleagues who made available on the internet the data sets we have used and the platforms to process them.

Conflicts of Interest: We declare no conflict of interest. We believe that anonymous reviewing may hide possible conflicts of interest and hence we favour eponymous reviewing.

References

1. Milly, P.C.D.; Betancourt, J.; Falkenmark, M.; Hirsch, R.M.; Kundzewicz, Z.W.; Lettenmaier, D.P.; Stouffer, R.J. Stationarity is dead: whither water management? *Science* **2008**, 319, 573-574.
2. Montanari, A.; Koutsoyiannis, D. Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research* **2014**, 50 (12), 9748–9756, doi: 10.1002/2014WR016092.
3. Koutsoyiannis, D.; Montanari, A. Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal* **2015**, 60 (7-8), 1174–1183, doi: 10.1080/02626667.2014.959959.
4. Padrón, R.S.; Gudmundsson, L.; Seneviratne, S.I. Observational constraints reduce likelihood of extreme changes in multidecadal land water availability. *Geophysical Research Letters* **2019**, 46, 736– 744, doi: 10.1029/2018GL080521.
5. Moon, H.; Gudmundsson, L.; Seneviratne, S I. Drought persistence errors in global climate models. *Journal of Geophysical Research: Atmospheres* **2018**, 123, 3483– 3496, doi: 10.1002/2017JD027577.
6. Koutsoyiannis, D.; Efstratiadis, A.; Georgakakos, K. Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-based approaches. *Journal of Hydrometeorology* **2007**, 8 (3), 261–281, doi: 10.1175/JHM576.1.
7. Koutsoyiannis, D.; Efstratiadis, A.; Mamassis, N.; Christofides, A. On the credibility of climate predictions. *Hydrological Sciences Journal* **2008**, 53 (4), 671–684, doi: 10.1623/hysj.53.4.671.
8. Anagnostopoulos, G.G.; Koutsoyiannis, D.; Christofides, A.; Efstratiadis, A.; Mamassis, N. A comparison of local and aggregated climate model outputs with observed data. *Hydrological Sciences Journal* **2010**, 55 (7), 1094–1110, doi: 10.1080/02626667.2010.513518.
9. Kundzewicz, Z.W.; Stakhiv, E.Z. Are climate models "ready for prime time" in water resources management applications, or is more research needed? Editorial. *Hydrological Sciences Journal* **2010**, 55 (7), 1085–1089.
10. Koutsoyiannis, D.; Christofides, A.; Efstratiadis, A.; Anagnostopoulos, G.G.; Mamassis, N. Scientific dialogue on climate: is it giving black eyes or opening closed eyes? Reply to "A black eye for the Hydrological Sciences Journal" by D. Huard. *Hydrological Sciences Journal* **2011**, 56 (7), 1334–1339, doi :10.1080/02626667.2011.610759.
11. Tyralis, H.; Koutsoyiannis, D. On the prediction of persistent processes using the output of deterministic models. *Hydrological Sciences Journal* **2017**, 62 (13), 2083–2102, doi: 10.1080/02626667.2017.1361535.
12. Montanari, A.; Koutsoyiannis, D. A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research* **2012**, 48, W09555, doi: 10.1029/2011WR011412.
13. Nearing, G. Comment on "A blueprint for process-based modeling of uncertain hydrological systems" by Alberto Montanari and Demetris Koutsoyiannis. *Water Resources Research* **2014**, 50, doi: 10.1002/2013WR014812.
14. Montanari, A.; Koutsoyiannis, D. Reply to comment by G. Nearing on ''A blueprint for process-based modeling of uncertain hydrological systems''. *Water Resources Research* **2014**, 50 (7), 6264–6268, doi: 10.1002/2013WR014987.
15. Sikorska, A.; Montanari, A.; Koutsoyiannis, D.,. Estimating the uncertainty of hydrological predictions through data-driven resampling techniques, *Journal of Hydrologic Engineering (ASCE)* **2015**, 20 (1), doi: 10.1061/(ASCE)HE.1943-5584.0000926.
16. Papacharalampous, G.; Koutsoyiannis, D.; Montanari, A. Quantification of predictive uncertainty in hydrological modelling by harnessing the wisdom of the crowd: Methodology development and investigation using toy models. *Advances in Water Resources* **2020**, doi: 10.1016/j.advwatres.2019.103471.
17. Papacharalampous, G.; Tyralis, C.; Koutsoyiannis, D.; Montanari, A. Quantification of predictive uncertainty in hydrological modelling by harnessing the wisdom of the crowd: A large-sample experiment at monthly timescale. *Advances in Water Resources* **2020**, doi: 10.1016/j.advwatres.2019.103470.
18. Koutsoyiannis, D.; Montanari, A. Bluecat: A local uncertainty estimator for deterministic simulations and predictions, *Water Resources Research* **2022**, 58 (1), e2021WR031215, doi: 10.1029/2021WR031215.
19. D. Koutsoyiannis, Stochastics of Hydroclimatic Extremes - A Cool Look at Risk, ISBN: 978-618-85370-0-2, 333 pages, Kallipos Open Academic Editions, Athens, 2021.
20. Koutsoyiannis, D.; Montanari, A. A brisk local uncertainty estimator for hydrologic simulations and predictions (blue cat). EGU General Assembly Conference Abstracts (p 10125). doi: 10.5194/egusphere-egu2020-10125m 2020, http://www.itia.ntua.gr/2044 (accessed 2022-03-13).
21. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; and Zhu, Y. The NCEP/NCAR 40-year reanalysis project. *B. Am. Meteorol. Soc.* **1996**, 77, 437–472.