On the use of Deep Generative Models for "Perfect" Prognosis Climate Downscaling

Jose González-Abad
Santander Meteorology Group and
Advanced Computing and e-Science Group
Institute of Physics of Cantabria (CSIC-UC)
Santander, Spain
gonzabad@ifca.unican.es

Jorge Baño-Medina
Santander Meterology Group
Institute of Physics of Cantabria (CSIC-UC)
Santander, Spain
bmedina@ifca.unican.es

Ignacio Heredia Cachá
Advanced Computing and e-Science Group
Institute of Physics of Cantabria (CSIC-UC)
Santander, Spain
iheredia@ifca.unican.es

Abstract

Deep Learning has recently emerged as a "perfect" prognosis downscaling technique to compute high-resolution fields from large-scale coarse atmospheric data. Despite their promising results to reproduce the observed local variability, they are based on the estimation of independent distributions at each location, which leads to deficient spatial structures, especially when downscaling precipitation. This study proposes the use of generative models to improve the spatial consistency of the high-resolution fields, very demanded by some sectoral applications (e.g., hydrology) to tackle climate change.

1 Motivations for generative models in "perfect" prognosis downscaling

Global Climate Models (GCMs) are the main tools used nowadays to study the evolution of climate at different time-scales. They numerically solve a set of equations describing the dynamics of the climate system over a three-dimensional grid (latitude-longitude-height). In climate change modeling, these models are utilized to produce possible future pathways of the climate system based on different natural and anthropogenic forcings. However, due to computational limitations these models present a coarse spatial resolution —between 1° and 3°,— which leads to a misrepresentation of important phenomena occurring at finer scales. The generation of high-resolution climate projections is crucial for important socio-economic activities (e.g., the energy industry), and they are routinely used to elaborate mitigation and adaptation politics to climate change at a regional scale.

Statistical Downscaling (SD) is used to bridge the scale-gap between the coarse model outputs and the local-scale by learning empirical relationships between a set of large-scale variables (predictors) and the regional variable of interest (predictands) based on large simulated/observational historical data records [1]. In this study we focus on a specific type of SD, named the "Perfect" Prognosis (PP) approach. PP downscaling leans on observational datasets to learn empirical relationships linking the predictors and the predictands. For the former, reanalysis data —a global dataset which combines observations with short-range forecasts through data assimilation,— is typically used, whilst for the latter either high-resolution grids or station-scale records can be employed. Once the relationship is established in these "perfect" conditions, we feed the model/algorithm with the equivalent GCM predictor variables to obtain high-resolution climate projections. A wide variety of

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2021.

statistical techniques have been deployed to establish these links, such as (generalized) linear models [2], support vector machines [3], random forests [4], classical neural networks [5], and more recently deep learning (DL). In particular, DL has recently emerged as a promising PP technique, showing capabilities to reproduce the observed local climate [6, 7, 8], whilst showing plausible climate change projections of precipitation and temperature fields over Europe [9]. Nonetheless, currently the regression-based nature of most of the existing PP methods, leads to an underestimation of the extremes when the predictors lack from sufficient informative power —i.e., given a particular predictor configuration there are many possible predictand situations,— since they output the conditional mean [10]. To account for the uncertainty describing the possible extremes is crucial for some activities, and the community has driven its attention to probabilistic regression-based modeling. The probabilistic models used mostly estimate the parameters of selected probability distributions conditioned to the large-scale atmospheric situation. The choice of the distribution depends on the variable of interest to be modeled —for instance, the temperature follows a Gaussian distribution, whilst wind or precipitation fields present a heavy-tailed structure which better fits with Gamma, Poisson or log-normal density functions,— and the regression-based models are trained to optimize the negative log-likelihood of the selected distribution at each site [5, 7, 11, 12, 13]. To model the spatial dependencies among sites, ideally we would estimate multivariate distributions representing the whole predictand domain, instead of predicting independent probability functions at each predictand site. Nonetheless, this was in practice computationally intractable, and very few procedures aimed to downscale over low-dimensional predictand spaces have been successfully deployed [14, 15, 16].

Recently, deep generative models have been developed that seek to approximate high-dimensional distributions through DL topologies. Based on previous merits in other disciplines, such as image-super-resolution (see e.g., [17, 18]), some studies have searched for an analogy between this task and downscaling, deploying Generative Adversarial Networks (GAN, [19, 20]) to obtain stochastic samples of high-resolution precipitation and temperature fields conditioned to their counterpart low-resolution ones. Despite these first studies are far from the PP approach, —since they lean on surface variables in their predictor set, which are not well represented by GCMs (see [1, 21] for guidelines/details on PP),— they show the potential of generative models to attain impressive levels of spatial structure in their stochastic downscaled predictions. Following this idea, we state that these topologies may provide a tractable alternative to model multivariate conditional distributions over high-dimensional domains in a PP setting, providing stochastic and spatially consistent downscaled fields very demanded by some sectoral applications for climate impact studies. To prove the potential of this type of DL topologies for PP-based downscaling, we show in the next section a use-case where Conditional Variational Auto-Encoders (CVAE) are deployed to produce stochastic high-resolution precipitation fields over Europe.

2 A downscaling case study over Europe with CVAE

We develop a simple use-case [1] which seeks to illustrate the promising capabilities of CVAE topologies to generate spatially consistent stochastic downscaled fields, especially as compared to the recent state-of-the-art PP DL-based topologies, which are based on the estimation of conditional Bernoulli-Gamma distributions at each predictand site (we refer the reader to [7] for more details). To this aim, we deploy the CVAE in the same conditions than [7], which builds on the validation framework proposed in the COST action VALUE [22]. VALUE proposes the use of ERA-Interim [23] reanalysis variables as predictors —trimmed to an horizontal resolution of 2^o,— and the regular gridded 0.5^o E-OBS dataset [24] as predictand. For the predictor set we use five thermodynamical variables (geopotential height, zonal and meridional wind, temperature, and specific humidity) at four different vertical levels (1000, 850, 700 and 500 hPa), whilst as predictand we use the daily accumulated precipitation over Europe. The models are trained on the period 1979-2002 and tested on 2003-2008.

Figure 1 shows the scheme of the CVAE proposed. This models builds on three different neural networks —an embedding network, an encoder and a decoder,— to produce stochastic samples of precipitation by sampling from a latent distribution which represents the complex interactions between predictors and predictands. During training, the embedding network transforms the high-dimensional predictors X to a low-dimensional array $\mathbf{z}_x$. This array is then stacked with the high-resolution predictand fields Y to feed the encoder network. The encoder outputs the parameters of a Gaussian distribution (i.e., the mean μ and the standard deviation σ), which encodes the spatial dependencies

[1] The code of the use-case is available at https://github.com/jgonzalezab/CVAE-PP-Downscaling

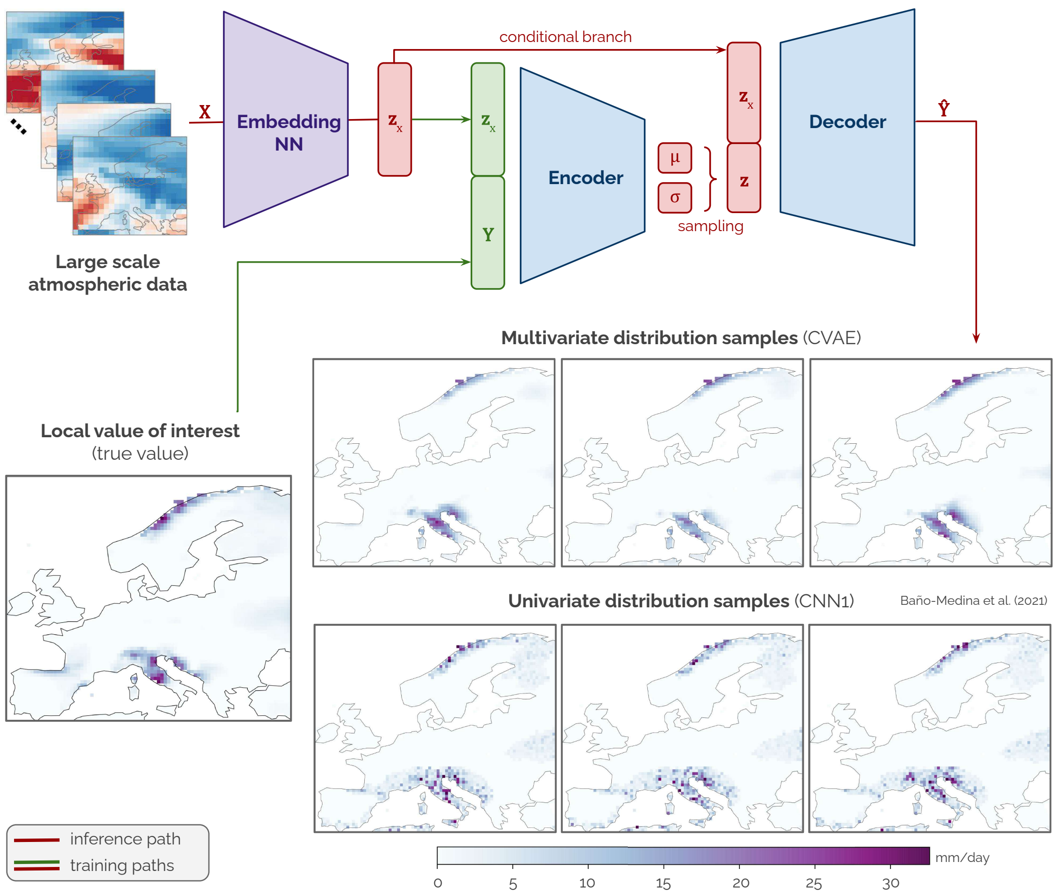

Figure 1: CVAE model architecture. Red lines represent the path followed by the model in the inference phase, during training both paths (green and red lines) are covered. At the bottom, a comparison between three different downscaled fields sampled from CVAE and CNN1 models, alongside the actual observation for 19/02/2004.

between both predictor and predictand fields. During both training and inference phases, stochastic realizations $\mathbf{z}$ sampled from this latent distribution are stacked with the low-dimensional predictor's embedding $\mathbf{z}_x$. This is used to feed the decoder network, which outputs the precipitation values $\widehat{Y}$ at each E-OBS predictand site considered. Therefore, different samples $\widehat{Y}$ conditioned on the same large-scale atmospheric situation X can be generated by sampling different vectors $\mathbf{z}$ from the latent distribution (see the three maps obtained for a particular day). We refer the reader to [25] for more details on CVAE.

For the sake of comparison, we select CNN1, which was one of the models that ranked first in [7], as an example of univariate model and compare its stochastic downscaled fields with those of CVAE. It can be seen how CNN1 fields present a spotty structure, characteristic of the sampling performed over the independent Bernoulli-Gamma distributions at each E-OBS site. In contrast, CVAE does not suffer from this problem improving the spatial consistency of the downscaled fields, as can be seen in the smoothness of the predictions.

3 Pathway of generative models to tackle climate change

Overall, we have showed the ability of CVAEs to produce spatially consistent stochastic fields in PP setups on a use-case over Europe. The generation of these high-resolution fields through generative models may foster the use of this type of downscaling into climate impact studies, since their products are very demanded by different sectors (e.g., agriculture, hydrology) to tackle climate change. In this line there are several challenges to address. For instance further research is needed in the evaluation of these models on aspects such as temporal consistency, and reproducibility of extremes. Also, in order to apply them to climate change projections, a study of its extrapolation capabilities is also required. The CVAE model developed here is a first approach, but further tuning this architecture may translate in improvements in the generated downscaled fields. For example, [26, 27] propose the use of normalizing flows to generate more complex latent distributions which could help capturing the complex non-linearities of the distribution of precipitation fields. Finally, the DL ecosystem offers a wide catalog of additional topologies which are of interest for PP downscaling (e.g., Conditional GANs [28]).

Acknowledgements. The authors acknowledge support from Universidad de Cantabria and Consejería de Universidades, Igualdad, Cultura y Deporte del Gobierno de Cantabria via the "instrumentación y ciencia de datos para sondear la naturaleza del universo" project. J. González-Abad would also like to acknowledge the support of the funding from the Spanish *Agencia Estatal de Investigación* through the *Unidad de Excelencia María de Maeztu* with reference MDM-2017-0765.

References

[1] Douglas Maraun and Martin Widmann. *Statistical downscaling and bias correction for climate research*. Cambridge University Press, 2018.

[2] José Manuel Gutiérrez, Douglas Maraun, Martin Widmann, Radan Huth, Elke Hertig, Rasmus Benestad, Ole Rössler, Joanna Wibig, Renate Wilcke, Sven Kotlarski, et al. An intercomparison of a large ensemble of statistical downscaling methods over europe: Results from the value perfect predictor cross-validation experiment. *International journal of climatology*, 39(9):3750–3785, 2019.

[3] Shien-Tsung Chen, Pao-Shan Yu, and Yi-Hsuan Tang. Statistical downscaling of daily precipitation using support vector machines and multivariate analysis. *Journal of hydrology*, 385(1-4):13–22, 2010.

[4] Bo Pang, Jiajia Yue, Gang Zhao, and Zongxue Xu. Statistical downscaling of temperature with the random forest model. *Advances in Meteorology*, 2017, 2017.

[5] Peter M Williams. Modelling seasonality and trends in daily rainfall data. In *Advances in neural information processing systems*, pages 985–991, 1998.

[6] Baoxiang Pan, Kuolin Hsu, Amir AghaKouchak, and Soroosh Sorooshian. Improving precipitation estimation using convolutional neural network. *Water Resources Research*, 55(3):2301–2321, 2019.

[7] Jorge Baño-Medina, Rodrigo Manzanas, and José Manuel Gutiérrez. Configuration and intercomparison of deep learning neural models for statistical downscaling. *Geoscientific Model Development*, 13(4):2109–2124, 2020.

[8] Lei Sun and Yufeng Lan. Statistical downscaling of daily temperature and precipitation over china using deep learning neural models: Localization and comparison with other methods. *International Journal of Climatology*, 41(2):1128–1147, 2021.

[9] Jorge Baño-Medina, Rodrigo Manzanas, and José Manuel Gutiérrez. On the suitability of deep convolutional neural networks for continental-wide downscaling of climate change projections. *Climate Dynamics*, pages 1–11, 2021.

[10] SC Pryor and JT Schoof. Differential credibility assessment for statistical downscaling. *Journal of Applied Meteorology and Climatology*, 59(8):1333–1349, 2020.

[11] Julie Carreau and Mathieu Vrac. Stochastic downscaling of precipitation with neural network conditional mixture models. *Water Resources Research*, 47(10), 2011.

[12] Anna Vaughan, Nicholas D Lane, and Michael Herzog. Multivariate climate downscaling with latent neural processes. 2021.

[13] Anna Vaughan, Will Tebbutt, J Scott Hosking, and Richard E Turner. Convolutional conditional neural processes for local climate downscaling. *Geoscientific Model Development Discussions*, pages 1–25, 2021.

[14] Alex J Cannon. Probabilistic multisite precipitation downscaling by an expanded bernoulli–gamma density network. *Journal of Hydrometeorology*, 9(6):1284–1300, 2008.

[15] Mohamed Ali Ben Alaya, Fateh Chebana, and Taha BMJ Ouarda. Probabilistic gaussian copula regression model for multisite and multivariable downscaling. *Journal of Climate*, 27(9):3331–3347, 2014.

[16] MA Ben Alaya, Taha BMJ Ouarda, and Fateh Chebana. Non-gaussian spatiotemporal simulation of multisite daily precipitation: downscaling framework. *Climate dynamics*, 50(1):1–15, 2018.

[17] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015.

[18] Zhen Li, Jinglei Yang, Zheng Liu, Xiaomin Yang, Gwanggil Jeon, and Wei Wu. Feedback network for image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3867–3876, 2019.

[19] Jussi Leinonen, Daniele Nerini, and Alexis Berne. Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network. *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[20] Bastien François, Soulivanh Thao, and Mathieu Vrac. Adjusting spatial dependence of climate model outputs with cycle-consistent adversarial networks. *Climate Dynamics*, pages 1–31, 2021.

[21] Douglas Maraun, Martin Widmann, and José M Gutiérrez. Statistical downscaling skill under present climate conditions: A synthesis of the value perfect predictor experiment. *International Journal of Climatology*, 39(9):3692–3703, 2019.

[22] Douglas Maraun, Martin Widmann, José M Gutiérrez, Sven Kotlarski, Richard E Chandler, Elke Hertig, Joanna Wibig, Radan Huth, and Renate AI Wilcke. Value: A framework to validate downscaling approaches for climate change studies. *Earth's Future*, 3(1):1–14, 2015.

[23] Dick P Dee, S M Uppala, AJ Simmons, Paul Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, d P Bauer, et al. The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656):553–597, 2011.

[24] Richard C. Cornes, Gerard van der Schrier, Else J. M. van den Besselaar, and Philip D. Jones. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres*, 123(17):9391–9409, 2018.

[25] Kihyuk Sohn, Honglak Lee, and Xinchen Yan. Learning structured output representation using deep conditional generative models. *Advances in neural information processing systems*, 28:3483–3491, 2015.

[26] Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pages 1530–1538. PMLR, 2015.

[27] Durk P Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, and Max Welling. Improved variational inference with inverse autoregressive flow. *Advances in neural information processing systems*, 29:4743–4751, 2016.

[28] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.