



New **E**nabling **V**isions and Tools for **E**nd-use**R**s and stakeholders thanks to a common **M**Odeling app**R**oach towards a Climat**E** neutral and resilient society

## D3.2 Production of high-resolution climate data for the case studies

November 2023



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101056858.

## Document history

<b>Project Acronym</b>	NEVERMORE
<b>Project ID</b>	101056858
<b>Project title</b>	New Enabling Visions and Tools for End-useRs and stakeholders thanks to a common MOdeling appRoach towards a ClimatE neutral and resilient society
<b>Project coordination</b>	Fondazione Bruno Kessler (Italy)
<b>Project duration</b>	1 <sup>st</sup> June 2022 – 31 <sup>st</sup> May 2026
<b>Deliverable Title</b>	D3.2 Production of high-resolution climate data for the case studies
<b>Type of Deliverable</b>	Dataset
<b>Dissemination level</b>	PU – Public
<b>Status</b>	Final
<b>Version</b>	1.0
<b>Work package</b>	WP3 – Climate science information
<b>Lead beneficiary</b>	CMCC
<b>Author(s)</b>	Giusy Fedele (CMCC), Stelios Karozis (NCSRd)
<b>Reviewer(s)</b>	Ivan Ramos (CARTIF), Thanasis Sftesos (NCSRd), Rita De Stefano (RINA-C), Alessia Torre (FBK)
<b>Due date of delivery</b>	30/11/2023
<b>Actual submission date</b>	29/11/2023

Date	Version	Contributors	Comments
19/09/2023	0.1	Stelios Karozis (NCSRd), Giusy Fedele (CMCC)	Table of Contents (ToC) defined and shared with the reviewers
15/11/2023	0.2	Stelios Karozis (NCSRd), Giusy Fedele (CMCC)	Document filled
27/11/2023	0.3	Stelios Karozis (NCSRd), Giusy Fedele (CMCC), Rita De Stefano (RINA)	Document updated
27/11/2023	0.4	Stelios Karozis (NCSRd), Giusy Fedele (CMCC), Rita De Stefano (RINA), Thanasis Sftesos (NCSRd)	Document reviewed by the selected reviewers
28/11/2023	0.5	Stelios Karozis (NCSRd), Giusy Fedele (CMCC)	Integration of comments and suggestions from reviewers, final draft shared for quality review to PC and TM
29/11/2023	1.0	Alessia Torre (FBK), Ivan Ramos Diez (CARTIF)	Final editing and submission



New Enabling Visions and Tools for End-useRs and stakeholders thanks to a common  
MOdeling appRoach towards a ClimatE neutral and resilient society

Copyright ©2022 NEVERMORE Consortium Partners. All rights reserved.

NEVERMORE is a Horizon Europe Project supported by the European Commission under contract No.101056858. For more information on the project, its partners, and contributors please see NEVERMORE website. You are permitted to copy and distribute verbatim copies of this document, containing this copyright notice, but modifying this document is not allowed. All contents are reserved by default and may not be disclosed to third parties without the written consent of the NEVERMORE partners, except as mandated by the European Commission contract, for reviewing and dissemination purposes. All trademarks and other rights on third party products mentioned in this document are acknowledged and owned by the respective holders. The information contained in this document represents the views of NEVERMORE members as of the date they are published. The NEVERMORE consortium does not guarantee that any information contained herein is error-free, or up to date, nor makes warranties, express, implied, or statutory, by publishing this document.

## Abbreviations and acronyms

---

Acronym	Description
C3S	Copernicus Climate Change Service
CDF	Cumulative Distribution Function
CMIP	Coupled Model Intercomparison Project Phase
CS	Case Study
ECS	Equilibrium Climate Sensitivity
ECV	Essential Climate Variable
EM	Ensemble-Mean
EQM	Empirical Quantile Mapping
GCM	Global Climate Model
RCP	Representative Concentration Pathway
SD	Statistical Downscaling
WP	Work Package

## Table of Contents

---

<b>DOCUMENT HISTORY</b> .....	<b>1</b>
<b>ABBREVIATIONS AND ACRONYMS</b> .....	<b>3</b>
<b>TABLE OF CONTENTS</b> .....	<b>4</b>
<b>LIST OF FIGURES</b> .....	<b>5</b>
<b>LIST OF TABLES</b> .....	<b>5</b>
<b>EXECUTIVE SUMMARY</b> .....	<b>6</b>
<b>1. INTRODUCTION</b> .....	<b>6</b>
1.1 Structure of the document .....	6
<b>2. STATISTICAL DOWNSCALING METHODOLOGY</b> .....	<b>6</b>
2.1 Bias-correction approach .....	7
2.1.1 Dataset Description .....	8
2.2 AI approach.....	10
2.2.1 Convolutional Operation and Convolutional Neural Networks .....	11
2.2.2 Autoencoder.....	12
2.2.3 Convolutional Autoencoders.....	12
2.2.4 Evidence Transfer .....	13
2.2.5 NEVEMORE AI climate data.....	14
<b>3. VISUALIZATION AT CASE STUDY LEVEL</b> .....	<b>14</b>
<b>4. CONCLUSIONS</b> .....	<b>17</b>
<b>5. REFERENCES</b> .....	<b>18</b>

## List of Figures

---

Figure 1. Workflow of statistical downscaling techniques.....	7
Figure 2. Statistical Downscaling Workflow.....	7
Figure 3. A representation of autoencoder consisting of encoder and decoder.....	12
Figure 4. Illustration of training steps of a Convolutional Autoencoder (CAE) with evidence transfer learning.....	13
Figure 5. Climatology of 2m temperature over the historical (1985-2014) period for ACCESS-CM2 before (tick black line) and after SD-EQM (shaded black line) against CERRA (in red) for all the case studies.....	14
Figure 6. Climatology of daily precipitation over the historical (1985-2014) period for ACCESS-CM2 before (tick black line) and after SD-EQM (shaded black line) against CERRA (in red) for all the case studies.....	15
Figure 7. The climatological signal of 2m temperature for ACCESS-CM2 before (tick line) and after SD-EQM (shaded line) are shown over the historical period (black) and two SSP245 and SSP585 scenarios (orange and red respectively) for all the case studies.....	15
Figure 8. The climatological signal of daily precipitation for ACCESS-CM2 before (tick line) and after SD-EQM (shaded line) are shown over the historical period (black) and two SSP245 and SSP585 scenarios (orange and red respectively) for all the case studies.....	16
Figure 9. Yearly mean surface temperature time series of CERRA (red line), ACCESS-CM2 before (blue line) and after SD-EQM (green line) for all the case studies.....	16
Figure 10. Yearly cumulated precipitation time series of CERRA (red line), ACCESS-CM2 before (blue line) and after SD-EQM (green line) for all the case studies.....	17

## List of Tables

---

Table 1. List of CMIP6 models. The table reports the Equilibrium Climate Sensitivity (ECS) to assess the GCMs climate sensitivity to CO <sub>2</sub> .....	8
Table 2. The CMIP6 ECVs collected by Task 3.2.....	9
Table 3. Variables missing for the daily CMIP6 GCMs. The “N.A” stands for data Not Available.....	9

## Executive summary

---

Deliverable D3.2 “Production of high-resolution climate data for the case studies” provides high-resolution downscaled Essential Climate Variables (ECVs) at the Case Study (CS) level that could be used as input for climate change impact models. The document describes the downscaling procedures used to enhance the spatial resolution of basic-ECVs (i.e., total precipitation, maximum, minimum and mean temperature, and wind speed). Two different methodologies were adopted and described: The Statistical Downscale (SD) through Empirical Quantile Mapping and the SD through Artificial Intelligence (AI).

Throughout the document, results for each of the case studies are presented. These results have been used to characterize the climate in the case studies (Task 6.1) as well as to provide supporting data for the activities that are being developed in the municipalities of the Murcia region to meet their adaptation and mitigation objectives and in this way to be able to achieve their signature to the Covenant of Mayors for Climate and Sustainable Energy.

## 1. Introduction

---

This document aims to present the activities carried out by CMCC and NCSR D within the Task 3.2 “Downscaling of climate information” of the NEVERMORE project. The aim of Task 3.2 is to provide high-resolution downscaled GCMs at case study level as input for impact models.

The document describes the methodologies adopted to statistically downscale GCMs from a coarser spatial resolution to the finer one, presenting the data processing workflows, data distribution strategies and the technical validation to assess the reliability and consistency of the produced data. Results of the essential climate variables for each of the case studies are also included with the aim that the reader understands the data obtained by the downscaling models.

### 1.1 Structure of the document

Beyond the Introduction and this paragraph, the Document is structured as follows:

- Section 2: The Statistical Downscaled approaches are presented and discussed.
- Section 3: Some results are shown.
- Section 4: Main conclusions obtained from data downscaling.

## 2. Statistical downscaling methodology

---

Several methods can be adopted to statistically downscale climate information from a coarse resolution to a fine grid. The methodology used within the NEVERMORE project is described in this document, presenting first the approach used by CMCC through bias-correction techniques and then the Demokritos’s approach through AI/Data driven techniques (Figure 1). The Bias corrected dataset will be used as training dataset to validate the AI derived downscaled climate data via relevant statistical metrics (e.g. RMSE).

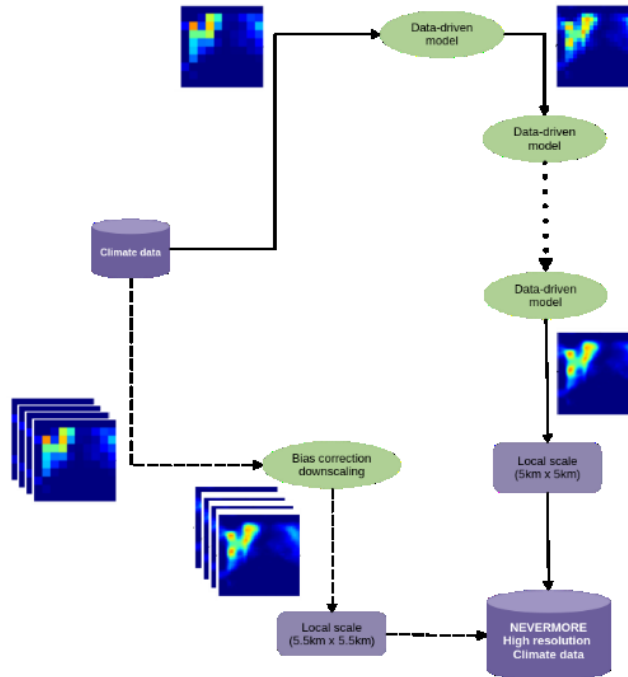


Figure 1. Workflow of statistical downscaling techniques.

## 2.1 Bias-correction approach

The workflow to statistically downscale climate information from a coarse to a fine grid consists of two main steps (see Figure 2). First, the climate information is interpolated from the “native” GCMs resolution of the coarse grid to the “reference” resolution of the finer grid. Once the climate information is remapped (or interpolated) over the finest grid, point-by-point statistical relationships are computed between raw “remapped” and “reference” time series during a shared training period. These relationships are expected to statistically translate the “interpolated” Cumulative Distribution Function (CDF) towards the “reference” one, removing (or reducing) the bias and properly preserving local scale physical processes and orographic features characterising the “reference” climate dataset. It is worth noting that the statistical relationships, known also as transfer functions, are here derived on monthly basis to preserve the seasonality of climate information. Finally, the same statistical relationships, retrieved over the training period between coarse and fine climate information, are applied to the climate information over the projection period.

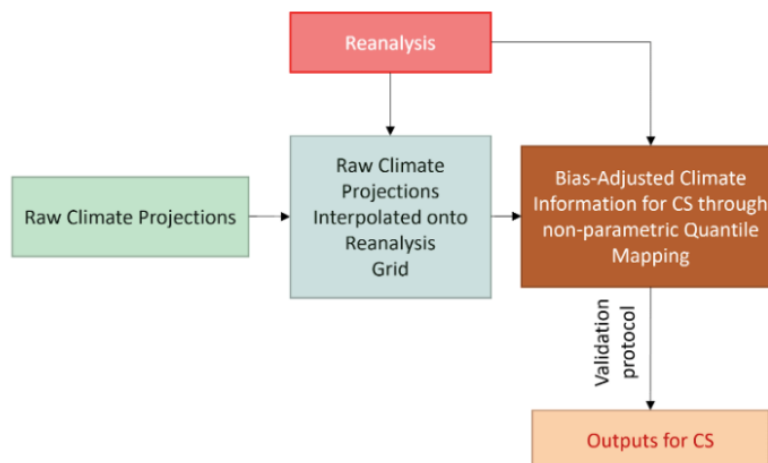


Figure 2. Statistical Downscaling Workflow.



Several bias-correction techniques exist to remove the bias over the “interpolated” climate information, with different level of complexity (Cannon et al., 2015; Casanueva et al., 2020). Within this project, the EQM has been adopted, computing the transfer functions over the control period, mapping the quantiles from the “interpolated” empirical model CDF over the respective values in the reference distribution.

The statistical relationship providing the distribution functions of the modelled variables into the reference ones is:

$$x_r = h(x_m) = F_r^{-1}[F_m(x_m)] = x_m^{\blacksquare} \tag{1}$$

where  $x_r$ ,  $x_m$  and  $x_m^{\blacksquare}$  are the reference, model and model bias-adjusted values;  $h$  is the transformation function;  $F_r$  and  $F_m$  are the CDFs of  $x_r$  and  $x_m$ .

Operatively, EQM first estimates CDFs of reference and model daily data over a set of regularly spaced quantile levels, using linear interpolation to obtain values with respect to the quantile, and then a constant extrapolation (first and last corrections for values below and above the calibration range, respectively) for out-of-sample values is applied. Once the transfer function is computed, the empirical correction function  $h$  is applied to the projection climate data to obtain bias-adjusted model data over the scenario. Moreover, as above mentioned, the correction function  $h$  is computed on monthly basis to preserve the seasonality of the reference climate information. The collected data for SD-EQM are discussed in the section 2.1.1.

### 2.1.1 Dataset Description

In this section, the dataset used for the statistical downscale of climate information are presented, first describing the set of 8 GCMs used to retrieve climate information at coarse resolution (which can vary between ~100 and ~1000 km) and then presenting the CERRA climate re-analysis (Schimanke et al., 2021) used as reference dataset over the calibration period to statistically downscale down to ~5.5 km the selected ECVs.

#### 2.1.1.1 Global Climate Projection

The collected climate projection through C3S and the ESGF portal (<https://esgf-node.llnl.gov/search/cmip6/>) for eight CMIP6 climate models are listed below:

1. ACCESS-CM2 (Australia)
2. CNRM-ESM2-1 (France)
3. IPSL-CM6A-LR (France)
4. MIROC6 (Japan)
5. NorESM2-MM (Norway)
6. CESM2 (USA)
7. EC-Earth3-Veg-LR (Europe)
8. HadGEM3-GC31-LL (UK)

The Equilibrium Climate Sensitivity (ECS) of these models is shown in Table 1. The ECS indicates the long-term temperature rise (equilibrium global mean near-surface air temperature) that is expected to result from a doubling of the atmospheric CO<sub>2</sub> concentration.

**Table 1. List of CMIP6 models. The table reports the Equilibrium Climate Sensitivity (ECS) to assess the GCMs climate sensitivity to CO<sub>2</sub>.**

GCMs	ECS (°C)
ACCESS-CM2	4,7

GCMs	ECS (°C)
CESM2	5,2
CNRM-ESM2-1	4,8
EC-Earth3-Veg-LR	4,3
HadGEM3-GC31-LL	5,6
IPSL-CM6A-LR	4,6
MIROC6	2,6
NorESM2-MM	2,5

Both historical and projections for SSP2-4.5, and SSP5-8.5 pathways are collected for each GCMs above mentioned.

In this respect, SSP2-4.5 represents a future path in which social, economic, and technological trends do not shift markedly from historical patterns. While SSP5-8.5 describes a future with very high GHG emissions where the push for economic and social development is coupled with the exploitation of abundant fossil fuel resources and the adoption of resource- and energy-intensive lifestyles around the world. In this SSP, CO<sub>2</sub> emissions will triple by 2075.

For each CMIP6 model and SSP, the ECVs presented in Table 2 are collected. These include the near surface air temperature (mean, minimum, and maximum), the precipitation, the relative humidity and the near surface wind speed. The ECVs collected by Task 3.2 cover 150 years (1950-2100) and are available at daily frequency.

**Table 2. The CMIP6 ECVs collected by Task 3.2.**

Long Name	Name	Units	Temporal Resolution
Daily maximum near surface air temperature	tasmax	K	Daily
Near surface air temperature	tas	K	Daily
Daily minimum near surface air temperature	tasmin	K	Daily
Precipitation	pr	kg m <sup>-2</sup> s <sup>-1</sup>	Daily
Near surface wind speed	sfcWind	m s <sup>-1</sup>	Daily
Relative Humidity	relhum	%	Daily

It is important to mention that not all CMIP6 variables are available for all models and SSP. Table 3 shows what variables are missing for the daily data.

**Table 3. Variables missing for the daily CMIP6 GCMs. The “N.A.” stands for data Not Available.**

CMIP6 GCMs	EXPERIMENT	tasmax [daily]	tasmin [daily]	tas [daily]	pr [daily]	sfcWind [daily]	relhum [daily]
ACCESS-CM2	historical						
	SSP245						
	SSP585						
CESM2	historical	N.A.	N.A.				
	SSP245						
	SSP585						
CNRM-ESM2-1	historical						
	SSP245					N.A.	
	SSP585					N.A.	
IPSL-CM6A-LR	historical						N.A.
	SSP245						N.A.
	SSP585						N.A.

CMIP6 GCMs	EXPERIMENT	tasmax [daily]	tasmin [daily]	tas [daily]	pr [daily]	sfcWind [daily]	relhum [daily]
MIROC6	historical						
	SSP245						
	SSP585						
NorESM2-MM	historical						N.A.
	SSP245						N.A.
	SSP585						N.A.
EC-Earth3-Veg-LR	historical						
	SSP245					N.A.	
	SSP585					N.A.	
HadGEM_GC3_LL	historical						
	SSP245						
	SSP585						

The EQM-SD has been applied to downscale the GCM- CMIP6 climate projections (i.e., coarse-resolution model) to a target grid of 5.5 km x 5.5 km, assuming the CERRA climate reanalysis (Schimanke et al., 2021) as the fine-resolution model for training.

### 2.1.1.2 Climate Reanalysis Data as Observations for Training

A climate reanalysis gives a numerical description of the recent climate by combining models with observations – far more complete than any observational dataset can achieve. In detail, it estimates atmospheric parameters such as air temperature, pressure, relative humidity and wind at different altitudes, and surface parameters such as total precipitation, soil moisture content, ocean-wave height and sea-surface temperature. Its main advantages are that it provides:

- Regularly gridded data, even in places where there are no or few observations.
- A coherent, complete set of variables describing the atmospheric state.
- Reconstruction of the record of past weather since observations constrain it.

However, reanalysis systems differ in the observations assimilated, the climate model used, how the statistic errors are estimated, and how corrections are applied.

The climate reanalysis used in this application is the Copernicus European Regional ReAnalysis (CERRA; Schimanke et al., 2021). This re-analysis provides spatially and temporally consistent sub-daily historical reconstructions of atmospheric and surface meteorological variables for Europe at 5.5 km x 5.5 km from 1984 to the present day, using ERA5 as lateral boundary conditions. It was produced using the HARMONIE-ALADIN limited-area numerical weather prediction, and data assimilation system called the CERRA system. Although the data assimilation system can resolve data holes, the much sparser observational networks in the past periods can impact the quality of analyses leading to less accurate estimates.

CERRA reanalysis over 1985-2014 has been assumed as the observational reference for EQM-SD training for the NEVERMORE Case Studies.


## 2.2 AI approach

In the current section a validated AI approach (Karozi et al., 2023) is customised to fit the needs of NEVERMORE project. The target is to learn to recreate the high-resolution dataset produced with the statistical downscaling method and provide a fast and reliable method to be applied in any region beyond the case studies. The latter will permit to utilize the model created in ICT Toolkit WP. In the following sections, the method is briefly described.

## 2.2.1 Convolutional Operation and Convolutional Neural Networks

Convolutional neural networks (CNNs; Karozis et al., 2015) are specialized neural networks designed for processing data with grid-like structures, such as images. These networks have shown remarkable success in real-world applications (Goodfellow and Bengio, 2017). In essence, CNNs process input data through a filtering mechanism. Multiple filters are applied to scan the image, one at a time, identifying different aspects of the input. Think of it like a small filter sliding across the image, searching for distinct features, such as dark edges. Whenever a match is found, it's recorded on an output image. The term "convolutional neural network" derives from the mathematical operation called convolution, which is a specialized linear operation that combines two functions to show how one modifies the other. The word "convolution" refers to both the result function and the process of calculating it. It's defined as the integral of the product of two functions, one of which is reversed and shifted. This integral is computed for all shift values, resulting in the convolution function.

$$(f * g)(t) = \int_{-\text{inf}}^{\text{inf}} (f(t - \tau)g(\tau)d\tau \quad (2)$$

The convolution formula can be described as a weighted average of function  $f(\tau)$  at time "t," with the weighting determined by function  shifted by an amount  $t$ . As  $t$  changes, the weighting function accentuates different parts of the input function (Smith, 1997).

CNNs rely on three critical concepts to enhance machine learning systems:

- **Sparse Interactions:** Achieved by using smaller kernels than the input, reducing the number of parameters needed, which saves memory and improves statistical efficiency. Fewer operations are also required for computing the output.
- **Parameter Sharing:** Involves using the same parameters for multiple functions in a model. Each kernel member is applied at every position in the input, except for boundary pixels in some cases. Parameter sharing ensures that we learn only one set of parameters rather than a separate set for each location (Liu et al., 2016).
- **Equivariant Representations:** This means that when the input changes, the output changes in the same manner. Note that convolution isn't naturally equivariant to every transformation, such as changes in scale or rotation. Other mechanisms are required for handling these transformations.

Furthermore, CNNs are distinguished from other neural networks by their exceptional performance in processing image, speech, and audio signal inputs. They consist of three primary types of layers:

- **Convolutional Layer:** The core building block of a CNN that connects hidden neurons to specific regions of gridded data in the input, where convolution operations take place.
- **Pooling Layer:** An additional subsampling layer used to reduce the number of features generated by convolutional layers. Pooling methods like max-pooling or average-pooling extract statistical features from small regions.
- **Fully-Connected (FC) Layer:** Connects each node in the output layer directly to a node in the previous layer. Its purpose is to transform the output of convolution and pooling into a vector of values, each representing the probability that a certain feature belongs to a label, which can be an input for the next stage.

The convolutional layer is the first layer in a CNN, and while it can be followed by additional convolutional or pooling layers, the fully-connected layer is typically the final layer. With each layer, the CNN becomes more sophisticated, identifying larger elements or shapes within the image. Initial

layers focus on basic features like colours and edges, while subsequent layers recognize more complex structures, ultimately identifying the intended object.

### 2.2.2 Autoencoder

An autoencoder is a type of neural network that aims to reproduce its input at its output. It consists of two main components: an encoder, which maps the input to a code, and a decoder, which maps the code back to a reconstruction of the original input. Autoencoders are typically constrained to approximate the input instead of copying it precisely. They retain only the most relevant aspects of the data (Diederik and Welling, 2013). Originally, autoencoders were used for dimensionality reduction and feature learning, but in recent years, they have found applications in learning generative models of data. Some of the most powerful artificial intelligence models in the 2010s employed sparse autoencoders integrated into deep neural networks (Domingos, 2015).

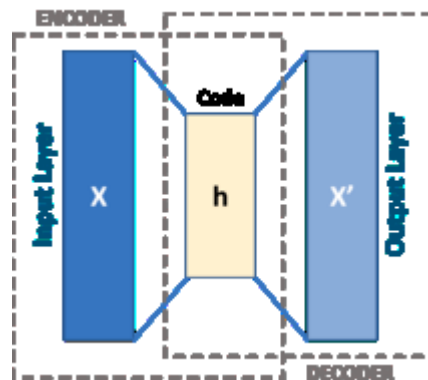


Figure 3. A representation of autoencoder consisting of encoder and decoder.

Depending on their design and objectives, autoencoders can be categorized into different types:

- **Undercomplete:** An autoencoder with a code dimension smaller than the input dimension, which forces it to capture the most important features of the training data.
- **Overcomplete:** In this case, the hidden code has a dimension greater than the input.
- **Regularized:** These autoencoders use a loss function that encourages properties beyond mere input replication, such as sparsity, smallness of derivatives, or robustness to noise or missing data.
- **Sparse:** A sparse autoencoder includes a sparsity penalty in its training criterion.
- **Denoising:** A denoising autoencoder receives a corrupted input and is trained to predict the original, uncorrupted data as output.

### 2.2.3 Convolutional Autoencoders

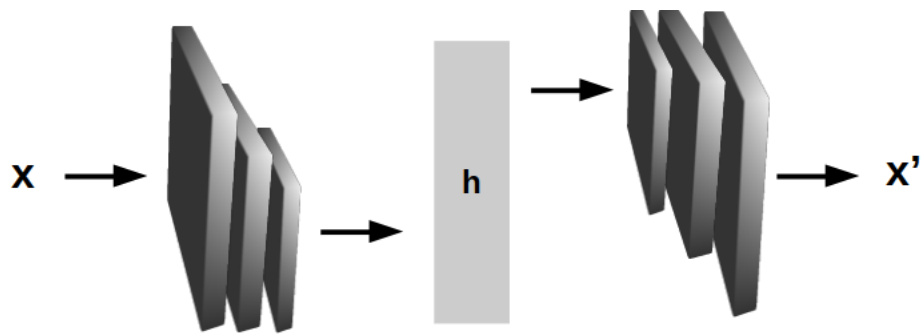
Convolutional Autoencoders (CAEs; Masci et al., 2011) are based on the standard autoencoder architecture but use convolutional layers instead of fully connected layers, with the exception of the output layer. This is a logical choice because convolutional layers offer advantages such as position invariance and reduced memory requirements through parameter sharing (Martinez-Murcia et al., 2020). While most pooling layers have been replaced by other techniques in modern CNNs, CAEs remain highly suitable for image processing due to their effective utilization of the inherent structure of convolutional neural networks (Sabour et al., 2017; Martinez-Murcia et al., 2020).

### 2.2.4 Evidence Transfer

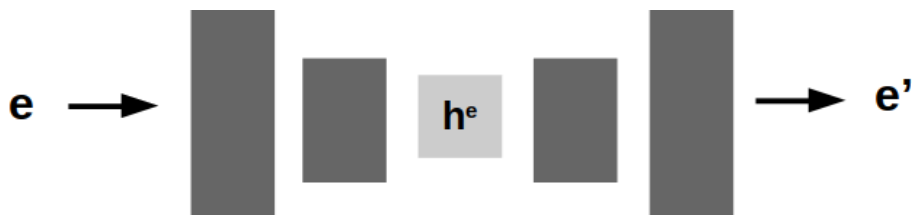
Evidence transfer (Davvetas et al., 2018) involves leveraging the outcomes of an auxiliary task to improve a primary task. It is a form of transfer learning where a pre-trained autoencoder's weights are adjusted to incorporate external evidence in the reconstruction task. In the case of CAE evidence transfer, the following steps are followed:

- **Initialization:** A baseline method is introduced by training the CAE as standard practice.
- **Evidence Preparation:** The evidence is prepared in either raw format or by training an additional autoencoder for each source of evidence to produce latent variables.
- **Input Evidence:** Additional layers are added to the output of the CAE, one for each source of evidence, to predict the raw or latent values obtained in step 2. It's important to note that there's no direct manipulation of the latent space; instead, a new training process adjusts the weights of the latent space (indirect manipulation). The weights decay in the case of low-quality evidence.

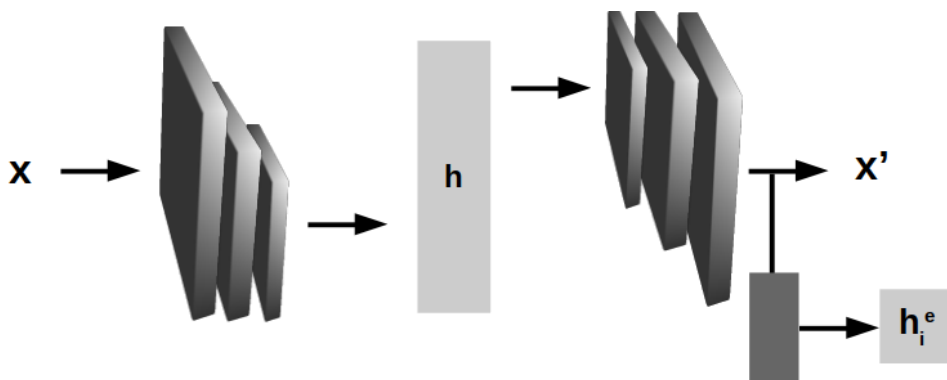
This workflow is depicted in Figure 4.



(a) Illustration of baseline CAE



(b) AE illustration for Evidence latent prediction



(c) Illustration of CAE with evidence transfer training

Figure 4. Illustration of training steps of a Convolutional Autoencoder (CAE) with evidence transfer learning.

### 2.2.5 NEVERMORE AI climate data

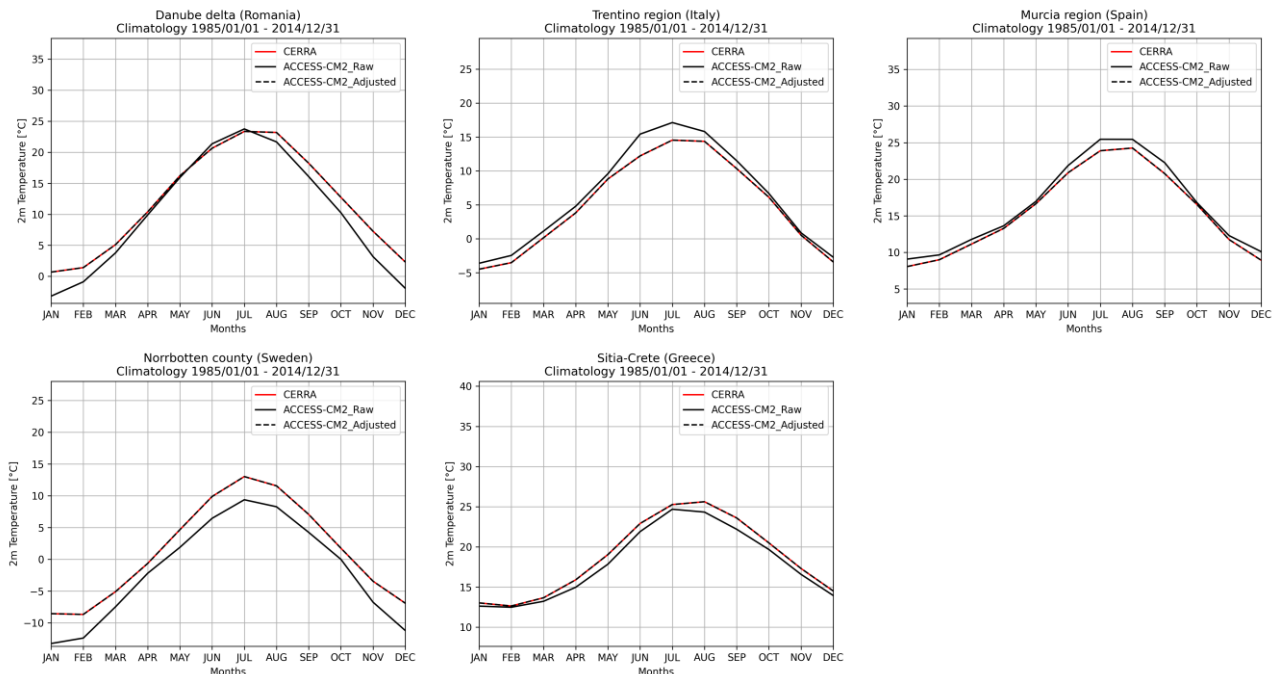
In the present task, a Convolutional Autoencoder (CAE) was trained using as training climate dataset mainly the bias corrected high-resolution dataset and a set of evidence relevant to the process. The latter consist of information that have an impact on the process of downscaling, such as the month or season that are more volatile in terms of events (e.g. Autumn and Spring) and information about the terrain which is becoming more detailed as the grid becomes denser. Subsequently, the evidence being currently tested are:

- Season
- Month
- Land use

The AI model is being trained for every relevant variable; temperature, precipitation, relative humidity and wind. These variables are planned to be used, with temperature being in the validation stage. The input data are based on climate models CMPI6 described in Section 2.1 (see Table 1) and the target resolution and dataset is the end production of bias correction gridded datasets, thus, providing a fast and reliable tool to downscale gridded datasets resulted from new socio-economic scenarios of WILIAM IAM model. The final trained models will be finalized and validated by the end of NEVERMORE project and will be used in integrating in WILIAM workflow as a fast and efficient post-process step to downscale the WILIAM information to local scale.

### 3. Visualization at case study level

In this section some results of the SD performed with EQM are shown for all the case studies for the temperature and precipitation variables related to a single GCM for simplicity (ACCESS-CM2).



**Figure 5. Climatology of 2m temperature over the historical (1985-2014) period for ACCESS-CM2 before (tick black line) and after SD-EQM (shaded black line) against CERRA (in red) for all the case studies.**

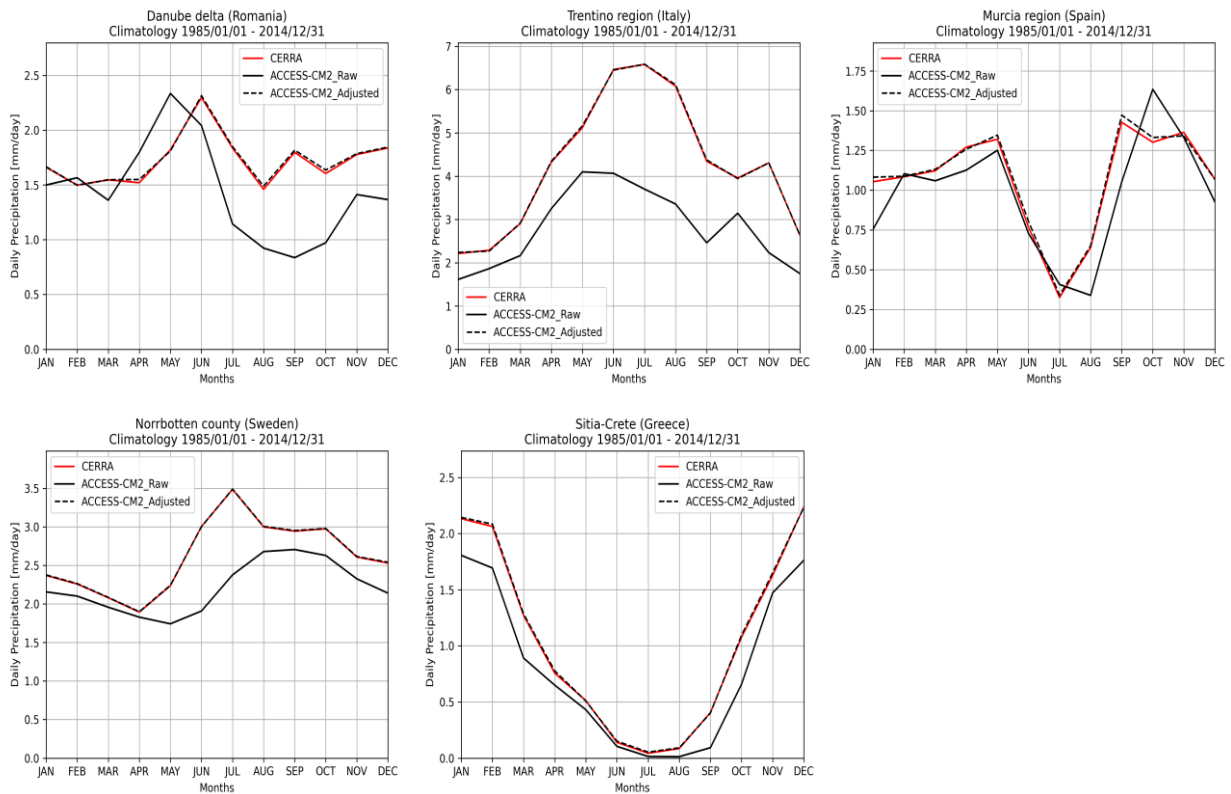


Figure 6. Climatology of daily precipitation over the historical (1985-2014) period for ACCESS-CM2 before (tick black line) and after SD-EQM (shaded black line) against CERRA (in red) for all the case studies.

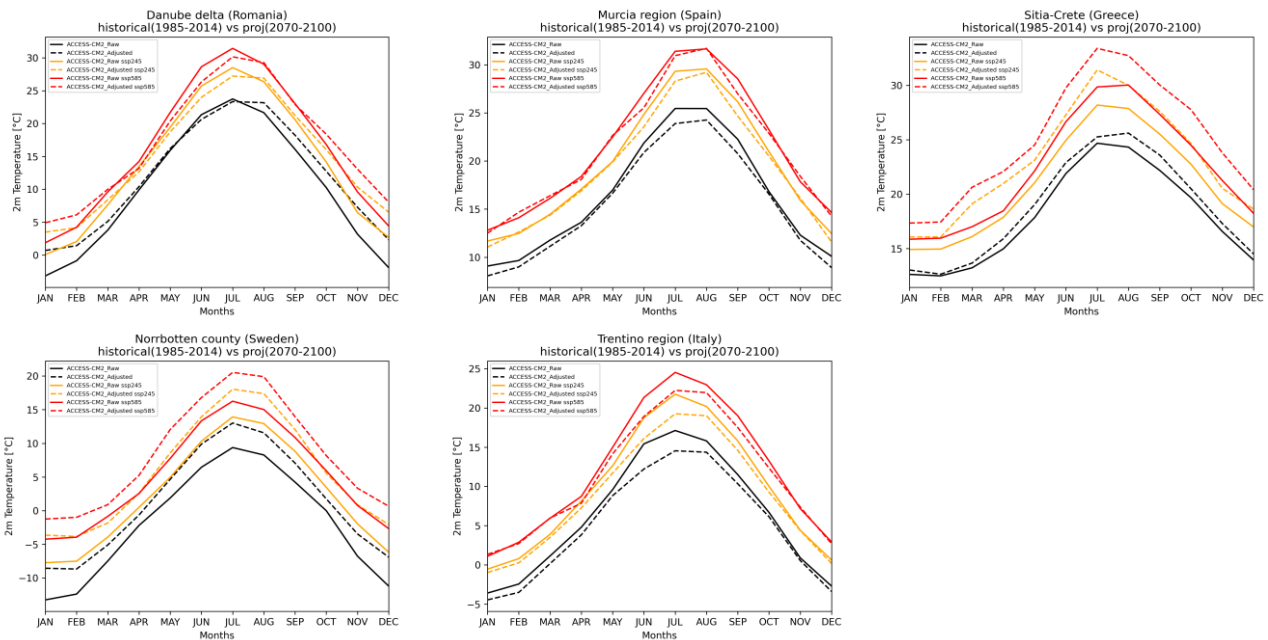


Figure 7. The climatological signal of 2m temperature for ACCESS-CM2 before (tick line) and after SD-EQM (shaded line) are shown over the historical period (black) and two SSP245 and SSP585 scenarios (orange and red respectively) for all the case studies.



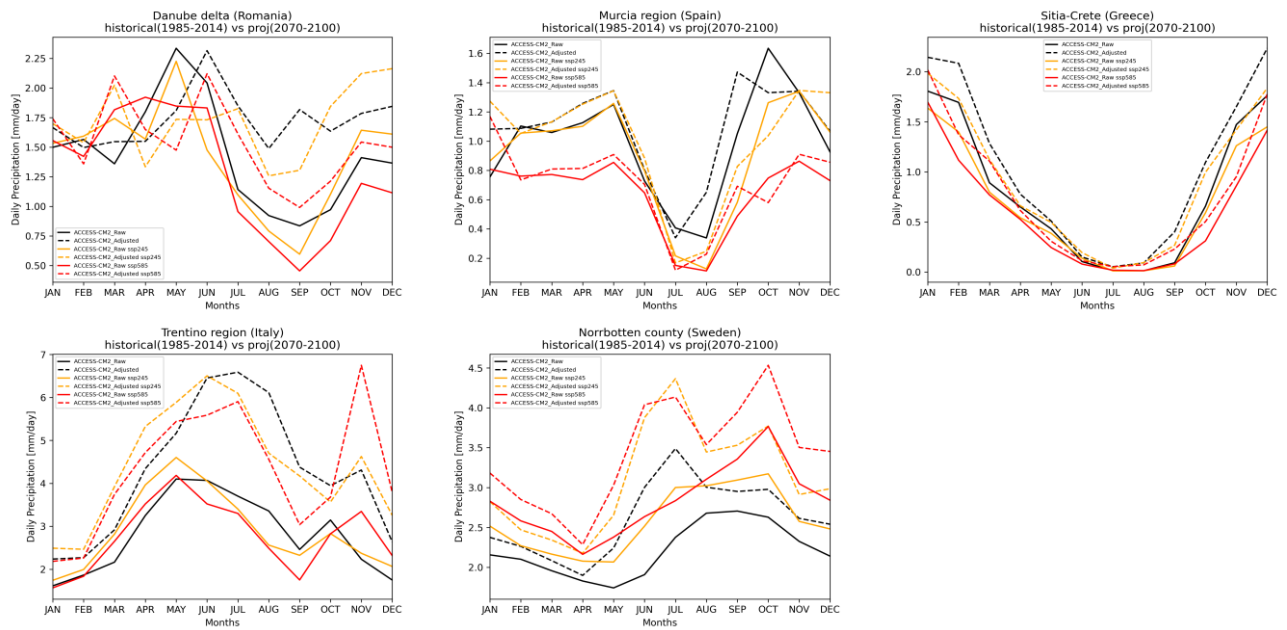


Figure 8. The climatological signal of daily precipitation for ACCESS-CM2 before (tick line) and after SD-EQM (shaded line) are shown over the historical period (black) and two SSP245 and SSP585 scenarios (orange and red respectively) for all the case studies.



Figure 9. Yearly mean surface temperature time series of CERRA (red line), ACCESS-CM2 before (blue line) and after SD-EQM (green line) for all the case studies.

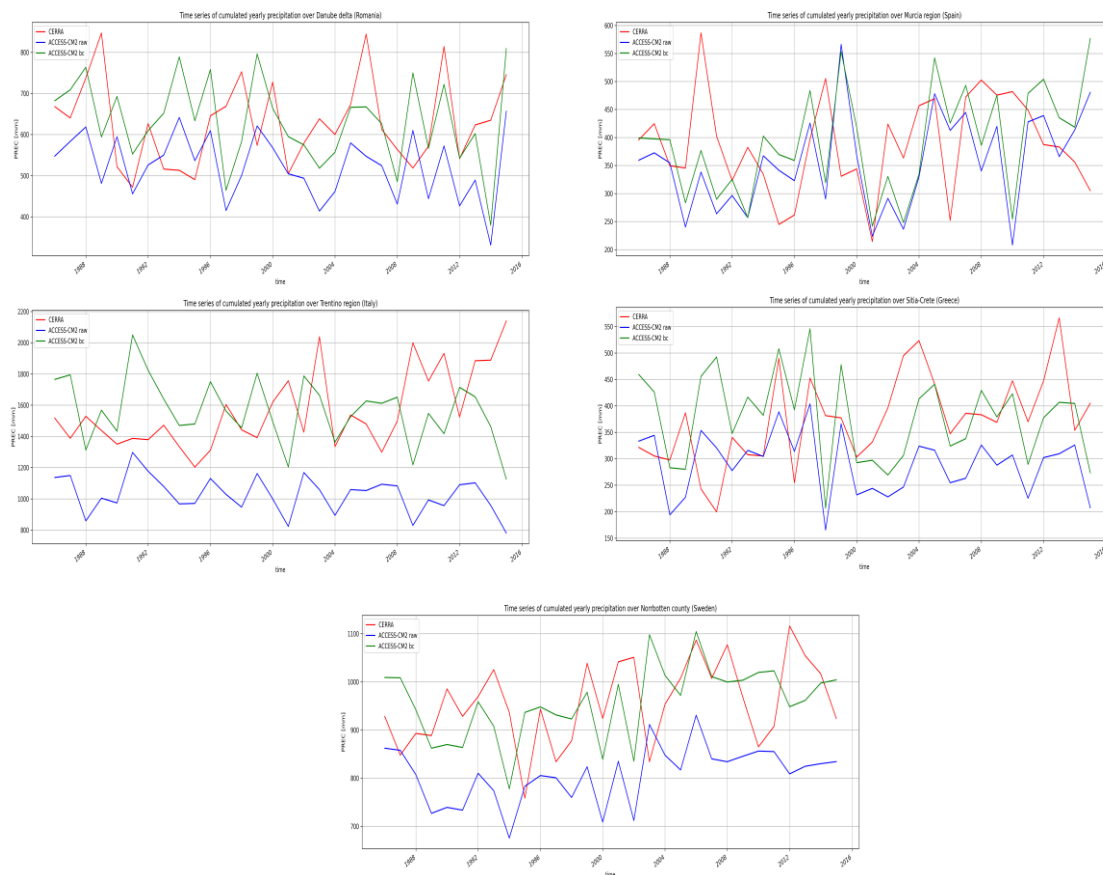


Figure 10. Yearly cumulated precipitation time series of CERRA (red line), ACCESS-CM2 before (blue line) and after SD-EQM (green line) for all the case studies.

## 4. Conclusions

The statistically downscaled variables obtained through EQM techniques have been evaluated against the reference CERRA re-analysis and the respective variables over their native grid. The improvements related to the performed downscaling are assessed investigating both the climatology and the signal at monthly basis. Figures shown in the previous section highlight a strong improvement in the ability of capturing the climatology and the signal of the reference data over the calibration period for the considered ECVs for each case study and GCM, moreover the methodology adopted allows to maintain the signal in the projection period. This evidence is important because it demonstrates the ability of the EQM SD technique to properly correct the intrinsic bias which belongs by construction to the climate atmospheric projections, preserving the observed climatology (Figures 5-6) without perturbing the climate signal of the climate simulation (Figures 7-8). Moreover, the yearly timeseries of the mean temperature and cumulated precipitation show a clear general improvement of the model skills in performing the reference variables (Figures 9-10).

## 5. References

---

- Cannon, A.J., Sobie, S.R., Murdock, T.Q. (2015). Bias correction of GCM precipitation by Quantile mapping: how well do methods preserve changes in quantiles and extremes? *J. Climate*, 28, 6938– 6959, <https://doi.org/10.1175/JCLI-D-14-00754.1>.
- Casanueva, A.; Herrera, S.; Iturbide, M.; Lange, S.; Jury, M.; Dosio, A.; Maraun, D.; Gutiérrez, J.M. (2020). Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. *Atmos. Sci. Lett.*, 21, e978, <https://doi.org/10.1002/asl.978>.
- Davvetas, A., Klampanos, I. A., & Karkaletsis, V. (2018). Evidence Transfer for Improving Clustering Tasks Using External Categorical Evidence. <https://doi.org/10.1109/IJCNN.2019.8852384>
- Diederik P Kingma; Welling, Max (2013). "Auto-Encoding Variational Bayes". arXiv:1312.6114
- Domingos, Pedro (2015). "4". *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books. "Deeper into the Brain" subsection. ISBN 978-046506192-1.
- Ian Goodfellow, Yoshua Bengio, A. C. (2017). *The Deep Learning Book*. MIT Press, 521(7553), 785. <https://doi.org/10.1016/B978-0-12-391420-0.09987-X>
- Karozis, S., Klampanos, I. A., Sfetsos, A., & Vlachogiannis, D. (2023). A deep learning approach for spatial error correction of numerical seasonal weather prediction simulation data. *Big Earth Data*, 1–20. <https://doi.org/10.1080/20964471.2023.21728>
- LeCun, Y. (1989). Generalization and network design strategies. Technical Report CRG-TR-89-4, University of Toronto. 330 , 352
- Liu, S., McGree, J., Ge, Z., & Xie, Y. (2016). Computer vision in big data applications. In *Computational and Statistical Methods for Analysing Big Data with Applications* (pp. 57–85). Elsevier. <https://doi.org/10.1016/B978-0-12-803732-4.00004-0>
- Martinez-Murcia F. J., A. Ortiz, J. -M. Gorriz, J. Ramirez and D. Castillo-Barnes, "Studying the Manifold Structure of Alzheimer's Disease: A Deep Learning Approach Using Convolutional Autoencoders," in *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 1, pp. 17-26, Jan. 2020, doi: 10.1109/JBHI.2019.2914970.
- Martinez-Murcia, Francisco J.; Ortiz, Andres; Gorriz, Juan M.; Ramirez, Javier; Castillo-Barnes, Diego (2020). "Studying the Manifold Structure of Alzheimer's Disease: A Deep Learning Approach Using Convolutional Autoencoders". *IEEE Journal of Biomedical and Health Informatics*. 24 (1): 17–26. doi:10.1109/JBHI.2019.2914970
- Masci, Jonathan, et al. "Stacked convolutional auto-encoders for hierarchical feature extraction." *International Conference on Artificial Neural Networks*. Springer Berlin Heidelberg, 2011.
- Sabour S., N. Frosst and G. E. Hinton, "Dynamic routing between capsules", *Proc. Adv. Neural Inf. Process. Syst.*, pp. 3857-3867, 2017.
- Schimanke S., Ridal M., Le Moigne P., Berggren L., Undén P., Randriamampianina R., Andrea U., Bazile E., Bertelsen A., Brousseau P., Dahlgren P., Edvinsson L., El Said A., Glinton M., Hopsch S., Isaksson L., Mladek R., Olsson E., Verrelle A., Wang Z.Q., (2021): CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: [10.24381/cds.622a565a](https://doi.org/10.24381/cds.622a565a) (Accessed on 20-10-2023).
- Smith, Stephen W (1997). "13.Convolution". *The Scientist and Engineer's Guide to Digital Signal Processing* (1 ed.). California Technical Publishing. ISBN 0-9660176-3-3. Retrieved 22 April 2016.



## Contacts

Fondazione Bruno Kessler  
E-mail: [NEVERMORE-info@fbk.eu](mailto:NEVERMORE-info@fbk.eu)  
Phone: +39 0461 314444  
Fax: +39 0461 314444

via Sommarive, 18,  
cp: 38123 Povo TN, Italia



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101056858.