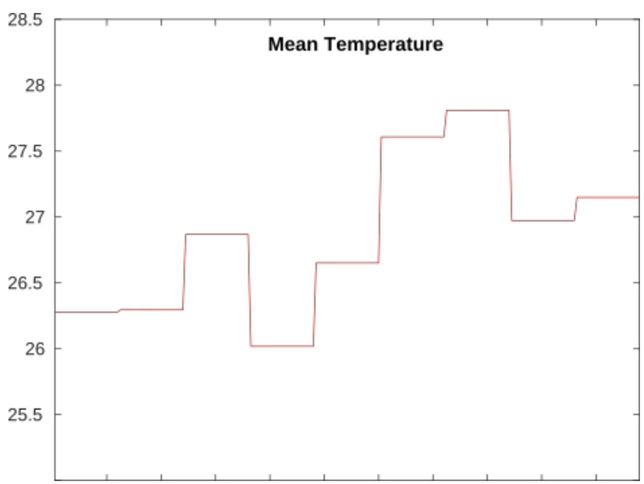
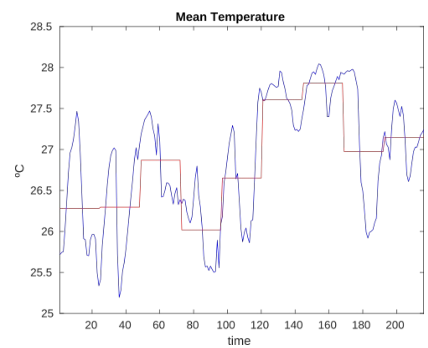
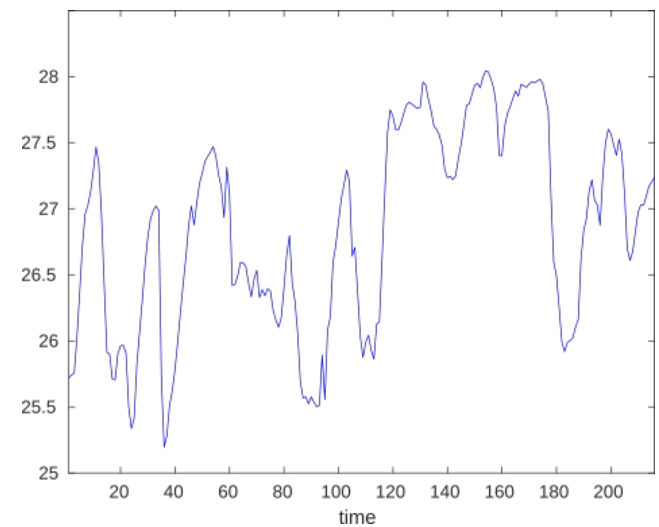
From daily to sub-daily climate projections: Increasing the temporal resolution by means of statistical downscaling





$$\mathbf{Y} = f\,(\mathbf{X};\theta)$$



Sixto Herrera
sixto.herrera@unican.es

Meteorology Group
Univ. de Cantabria – CSIC
MACC/IFCA

UNIVERSIDAD DE CANTABRIA
CSIC

Cross-sectoral ISIMIP and PROCLIAS online workshop 2021, January 11th (16:45-17:15 Berlin time)

Outline of the Talk

- ***Motivation***
  Impact models' requirements ↔ Statistical Downscaling
- ***Statistical downscaling***
  Transfer Functions
  Weather Typing - Analogues
  Weather Generators
  Analog-based methods
- ***Conclusions***

Motivation: Impact Models

climate projections
RCP scenarios from CMIP & CORDEX archives

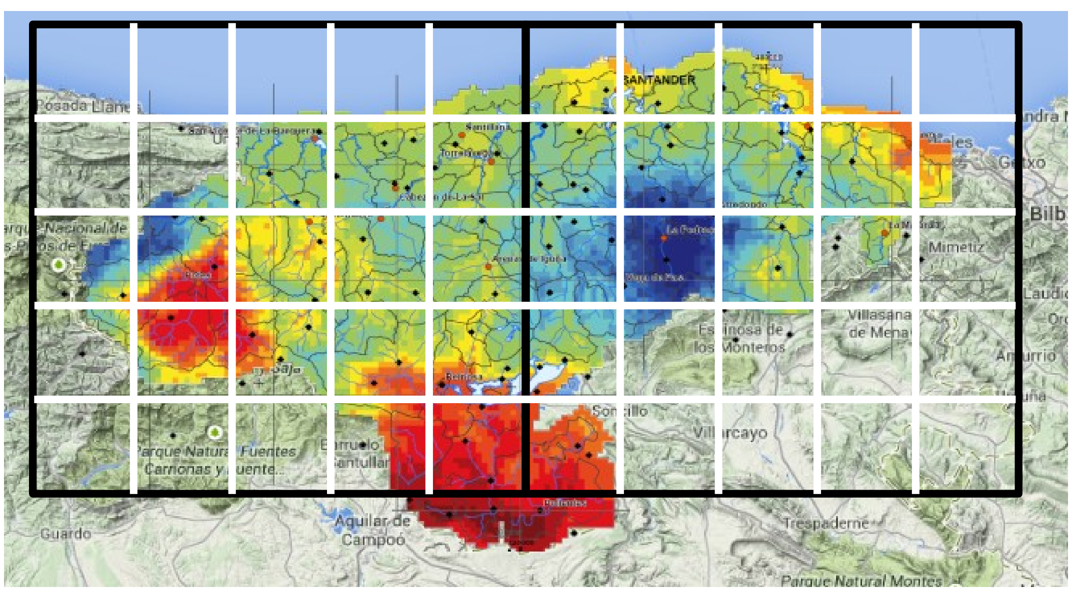
Socio-economic input
SSP scenarios

Impact models global & regional

agriculture	water
biomes	Forests
coastal infrastructure	health
fisheries	energy
agro-economics	permafrost

- Synthesis of impacts at different levels of global warming
- Quantification of uncertainties
- Model improvement
- Cross-sectoral interactions
- Cross-scale intercomparison
- Focus topics (e.g. extreme events, adaptation)

https://www.isimip.org/about/



Impact Models' needs:
- local or very high spatial resolution (~ 1km).

Spatial statistical downscaling
Challenges and prospects on the road to 1km ISIMIP3b daily climate data

Black cells ~ 100 km (GCMs, CMIP)

White cells ~ 10 km (RCMs, CORDEX)

Motivation: Impact Models

climate projections
RCP scenarios from CMIP & CORDEX archives

Socio-economic input
SSP scenarios

Impact models global & regional

agriculture	water
biomes	Forests
coastal infrastructure	health
fisheries	energy
agro-economics	permafrost

- Synthesis of impacts at different levels of global warming
- Quantification of uncertainties
- Model improvement
- Cross-sectoral interactions
- Cross-scale intercomparison
- Focus topics (e.g. extreme events, adaptation)

Floods
Urban drainage

Impact Models' needs:
- local or very high spatial resolution (~ 1km).
- Some extreme events have temporal resolutions lower than daily.

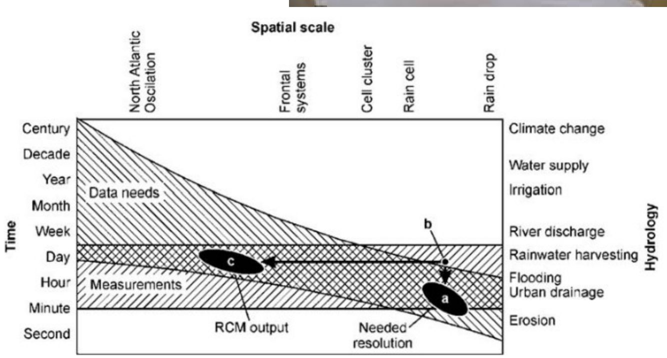


Figure 7.2 Statistical downscaling of RCM outputs down to the scale required for urban hydrological impact studies requires both temporal dowscaling (a – b) and spatial downscaling (b – c) (adapted from Arnbjerg-Nielsen, 2008).

Impacts of Climate Change on Rainfall Extremes and Urban Drainage Systems P. Willems, J. Olsson, K. Arnbjerg-Nielsen, S. Beecham, A. Pathirana, I. Bülow Gregersen, H. Madsen and V.T.V. Nguyen

Motivation: Impact Models

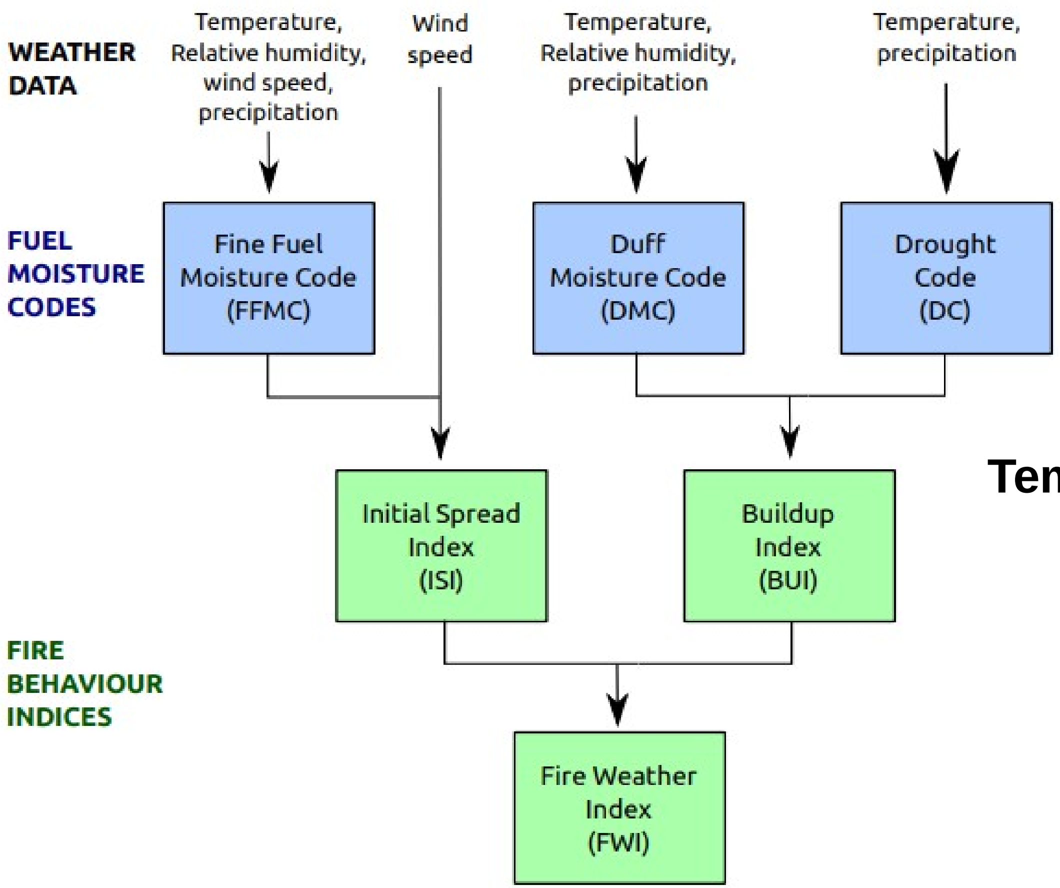


Figure 3.6: Block diagram of the CFFWIS (Adapted from van Wagner, 1987).

Fire

Temperature, humidity and wind speed at noon

Impact Models' needs:

- local or very high spatial resolution (~ 1km).
- Some extreme events have temporal resolutions lower than daily.
- Some communities need of subdaily data.

Motivation: Impact Models

WCRP CORDEX

ESGF Earth System Grid Federation

You are at the ***ESGF-DATA.DKRZ.DE*** *node*

Home | Technical Support

Project −
- [x] CORDEX (165003)

Product +

Domain +

Institute +

Driving Model +

Experiment +

Experiment Family +

Ensemble +

RCM Model +

Downscaling Realisation +

Time Frequency −
- [ ] 1hr (847)
- [ ] 3hr (8378)
- [ ] 6hr (10375)
- [ ] day (50004)
- [ ] fx (2628)
- [ ] mon (52854)
- [ ] sem (39917)

Variable +

Variable Long Name +

CF Standard Name +

Datanode +

Enter Text:

? Search Reset Display 10 results per page [More Search Options]

☐ Show All Replicas ☐ Show All Versions ☐ Search Local Node Only (Including All Replicas)

Search Constraints: ✖ CORDEX

Total Number of Results: 165003

-1- 2 3 4 5 6 Next >>

Please login to add search results to your Data Cart

Expert Users: you may display the search URL and return results as XML or return results as JSON

1. **cordex.output.AFR-44.DMI.ECMWF-ERAINT.evaluation.r1i1p1.HIRHAM5.v2.day.uas**
   Data Node: cordexesg.dmi.dk
   Version: 20140804
   Total Number of Files (for all variables): 5
   Full Dataset Services: [Show Metadata] [List Files] [THREDDS Catalog] [WGET Script]

2. **cordex.output.AFR-44.DMI.ECMWF-ERAINT.evaluatic**
   Data Node: cordexesg.dmi.dk
   Version: 20140804
   Total Number of Files (for all variables): 5
   Full Dataset Services: [Show Metadata] [List Files]

3. **cordex.output.AFR-44.DMI.ECMWF-ERAINT.evaluatic**
   Data Node: cordexesg.dmi.dk
   Version: 20140804
   Total Number of Files (for all variables): 5
   Full Dataset Services: [Show Metadata] [List Files]

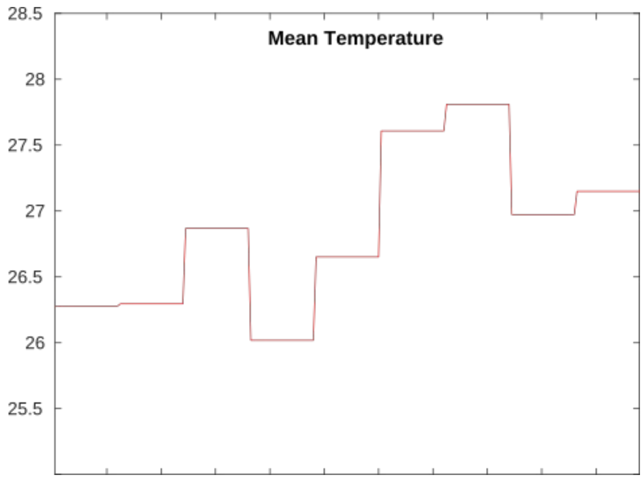
4. **cordex.output.AFR-44.DMI.ECMWF-ERAINT.evaluatic**
   Data Node: cordexesg.dmi.dk
   Version: 20140804
   Total Number of Files (for all variables): 5
   Full Dataset Services: [Show Metadata] [List Files]

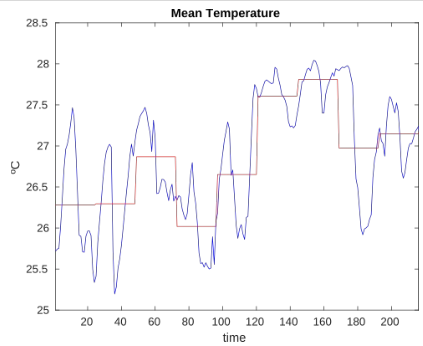
5. **cordex.output.AFR-44.DMI.ECMWF-ERAINT.evaluatic**
   Data Node: cordexesg.dmi.dk
   Version: 20140804
   Total Number of Files (for all variables): 5
   Full Dataset Services: [Show Metadata] [List Files]

CORDEX: requested variables

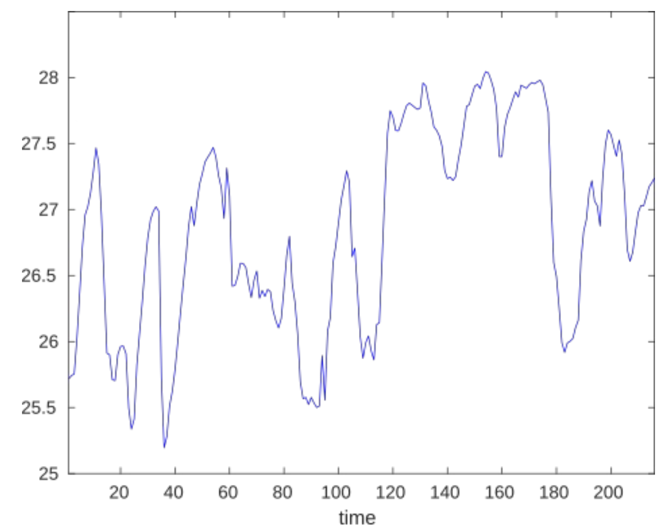
		Tier 2		Tier 1		Core			
output variable name	**units**	**frq [1/day]**	**ag**	**frq [1/day]**	**ag**	**frq [1/mon]**	**ag**	**frq [1/sem]**	**ag**
tas	K	8	i	1	8	1	m*8	1	s*8
tasmax	K			1		1	m	1	s
tasmin	K			1		1	m	1	s
pr	kg m-2 s-1	8	a	1		1		1	
ps	Pa	8	i	1	8				
psl	Pa	8	i	1	8	1	m*8	1	s*8
huss	1	8	i	1	8	1	m*8	1	s*8
hurs	%	8	i	1	8	1	m*8	1	s*8
sfcWind	m s-1	8	i	1	8	1	m*8	1	s*8
sfcWindmax	m s-1			1		1	m	1	s

Motivation: Downscaling





$$\mathbf{Y} = f\,(\mathbf{X};\theta)$$



CORDEX: requested variables

		Tier 2		Tier 1		Core			
output variable name	units	frq [1/day]	ag	frq [1/day]	ag	frq [1/mon]	ag	frq [1/sem]	ag
tas	K	8	i	1	8	1	m*8	1	s*8
tasmax	K			1		1	m	1	s
tasmin	K			1		1	m	1	s
pr	kg m-2 s-1	8	a	1		1		1	
ps	Pa	8	i	1	8				
psl	Pa	8	i	1	8	1	m*8	1	s*8
huss	1	8	i	1	8	1	m*8	1	s*8
hurs	%	8	i	1	8	1	m*8	1	s*8
sfcWind	m s-1	8	i	1	8	1	m*8	1	s*8
sfcWindmax	m s-1			1		1	m	1	s

Impact Models' needs:
- local or very high spatial resolution (~ 1km).
- Some extreme events have temporal resolutions lower than daily.
- Some communities need of subdaily data.

Temporal statistical downscaling:
Obtain sub-daily data from daily data

Statistical Downscaling

Statistical Downscaling

Transfer Functions | **Weather Generators** | **Analogues Weather Typing**

Links the local observed climate (**predictand** ***Y***) with the global simulations given by the GCMs (**predictors** ***X***), through some **function** ***f*** and/or **parameters** θ

Stochastic models able to generate artificial time series with similar statistical properties than the observed time series.

Similar atmospheric patterns lead to similar meteorological conditions (Lorenz, 1969). Based on the K-nearest neighbor algorithm.

Transfer Functions

Links the local observed climate (**predictand *Y***) with the global simulations given by the GCMs (**predictors *X***), through some **function *f*** and/or **parameters *θ***

A first approach is the so-called **delta method,** which assumes that the climate change signal at daily scale can be applied at hourly scale so it is *"added"* to the observations:

$$Y_{fut}^{h} = Obs^{h} + \left(Y_{fut}^{day} - Y_{ref}^{day}\right)$$

Several variations of this method have been proposed considering other moments of the statistical distribution (variance, quantiles, etc. See Willems and Vrac).

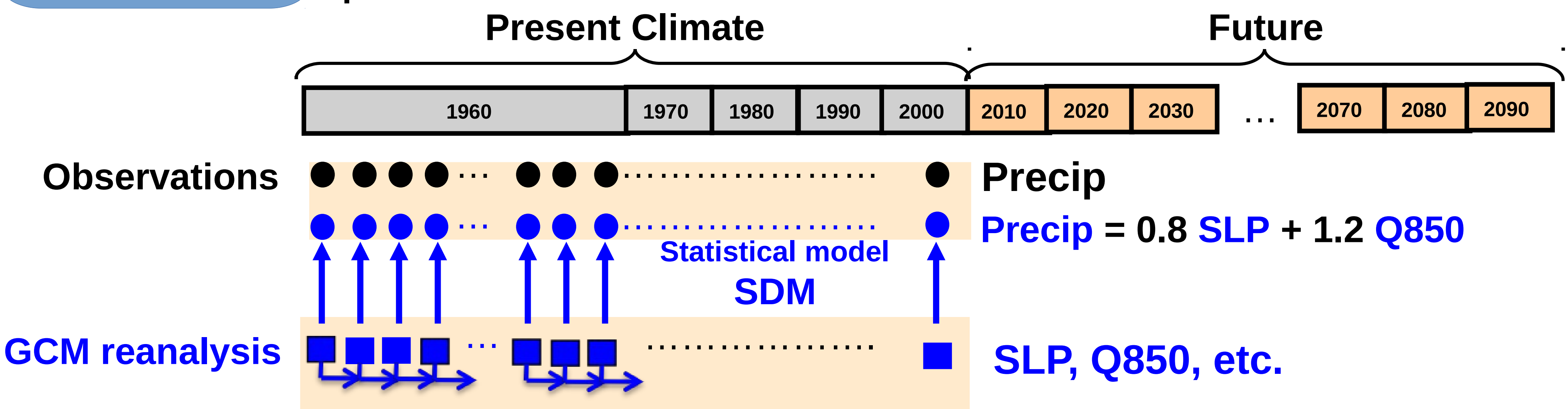
The main shortcoming is that the daily cycle can not be modified, the observed is preserved for the projections.

P. Willems, M. Vrac, Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change, Journal of Hydrology, Volume 402, Issues 3–4, 2011, Pages 193-205, https://doi.org/10.1016/j.jhydrol.2011.02.030.

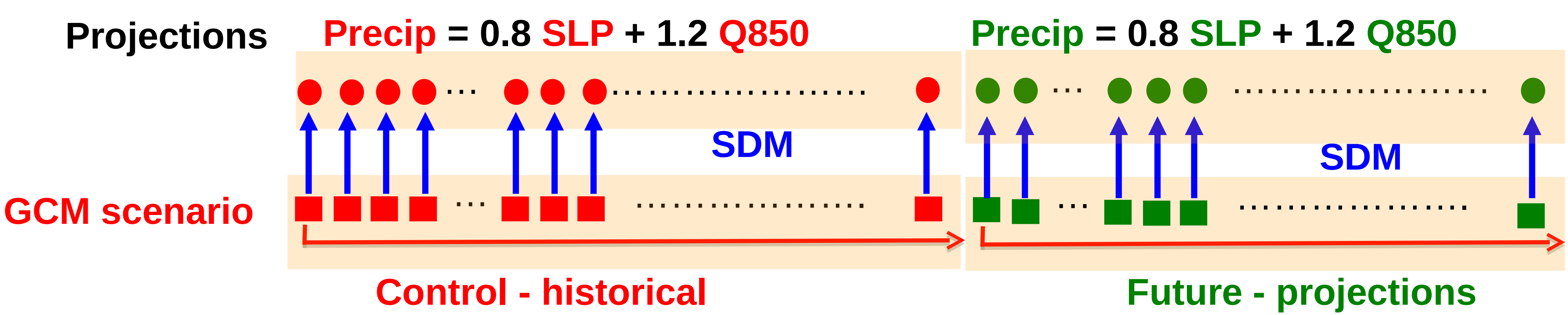
Statistical Downscaling

Transfer Functions

Links the local observed climate (**predictand *Y***) with the global simulations given by the GCMs (**predictors *X***), through some **function *f*** and/or **parameters θ**

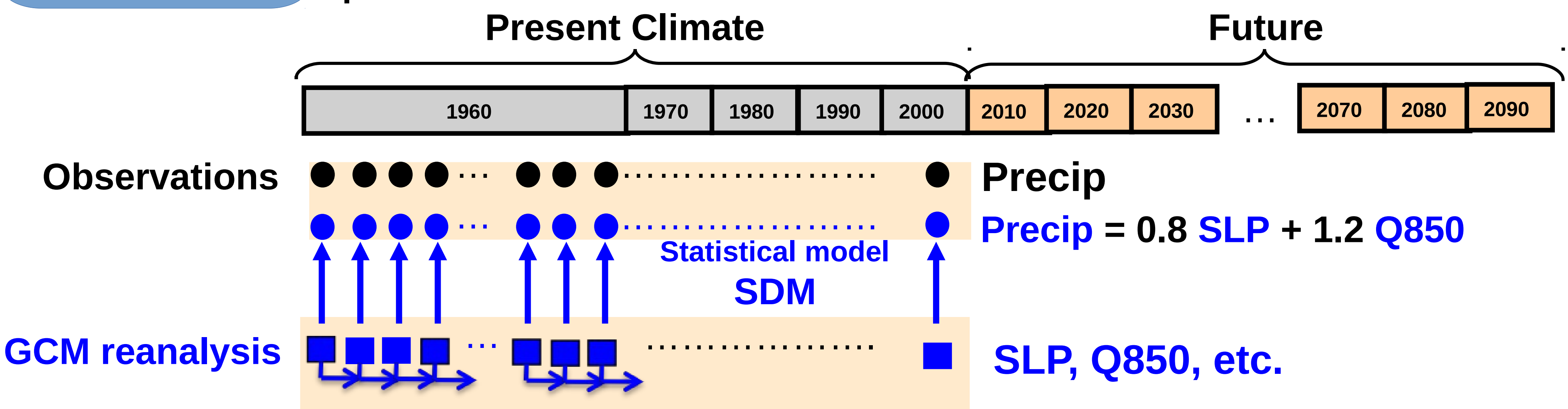


- **Assumptions: Reanalysis choice, choosing consistent predictors (e.g. sea level pressure and specific humidity) and stationarity/robustness.**



Transfer Functions

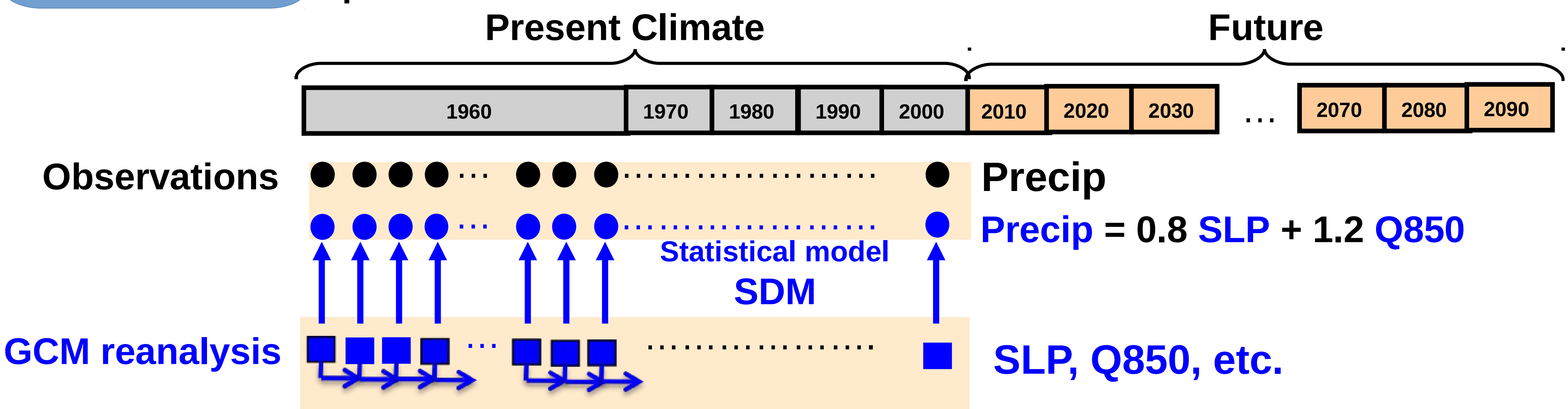
Links the local observed climate (**predictand *Y***) with the global simulations given by the GCMs (**predictors *X***), through some **function *f*** and/or **parameters *θ***



- **Assumptions:** **Reanalysis choice, choosing consistent predictors (e.g. sea level pressure and specific humidity) and stationarity/robustness.**
- **Assumptions:** **There must be temporal consistency between observations, reanalysis and projections.**

Transfer Functions

Links the local observed climate (**predictand *Y***) with the global simulations given by the GCMs (**predictors *X***), through some **function *f*** and/or **parameters *θ***



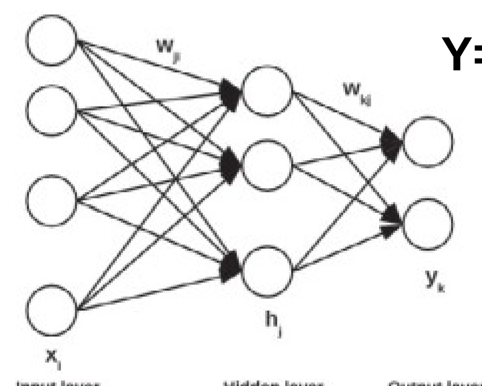
- **Adaptation: Consider a model for each time step based on the same predictors to maintain some coherence between the predicted values.**

Precip01h = **a01 SLP + b01 Q850**

……………………

Precip24h = **a24 SLP + b24 Q850**

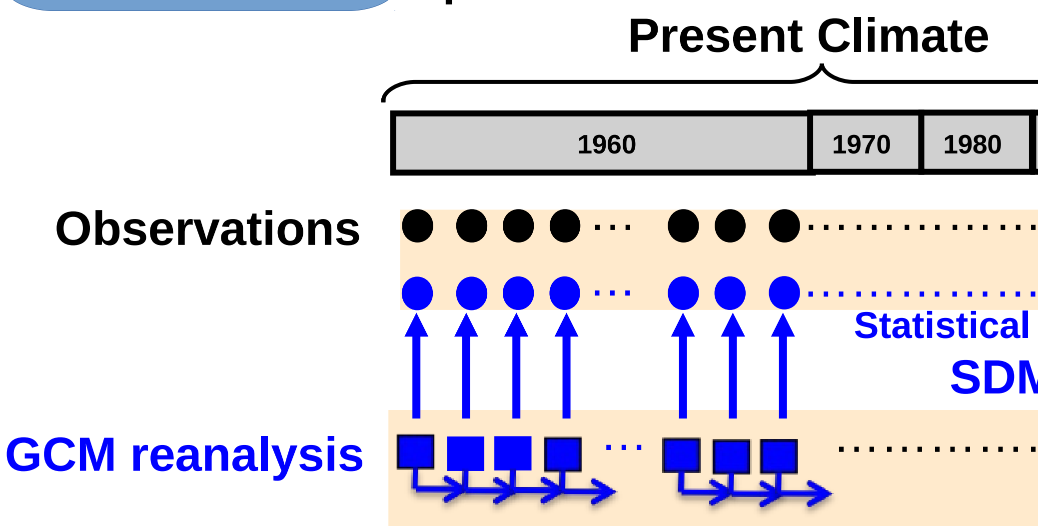
Y=[Y01,…,Y24]

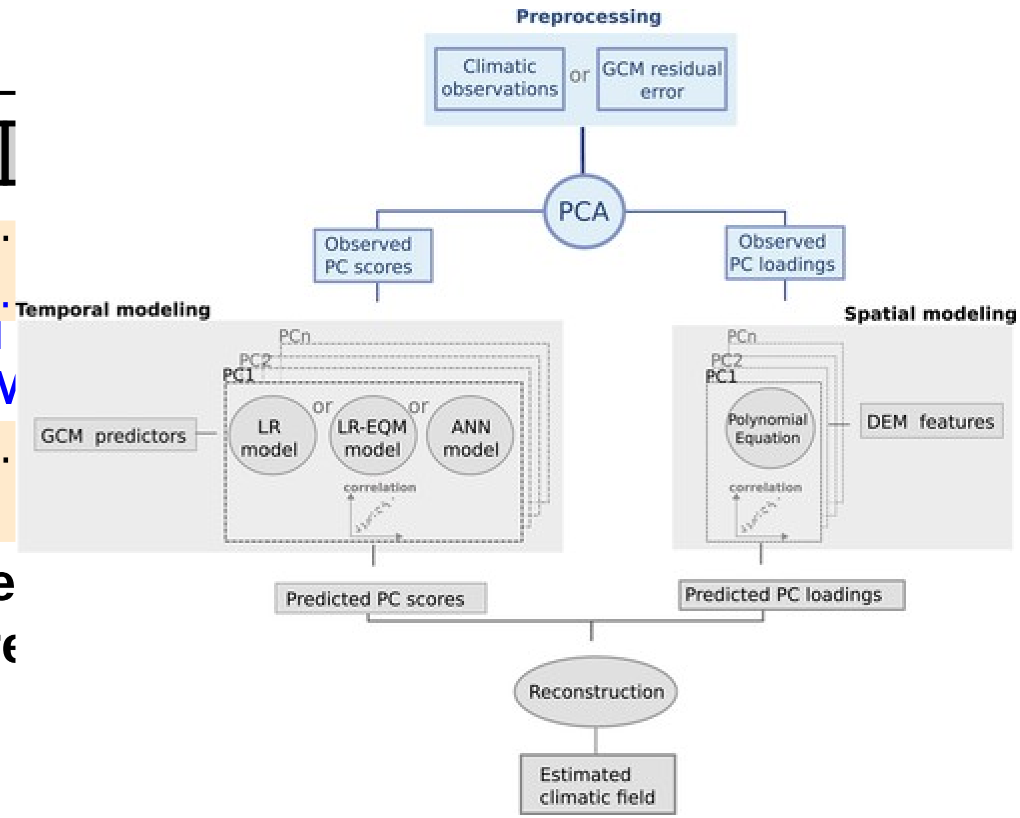


Kumar et al. 2012, Sub-daily Statistical Downscaling of Meteorological Variables Using Neural Networks, Procedia Computer Science, 9, pages: 887-896, https://doi.org/10.1016/j.procs.2012.04.095.

Transfer Functions

Links the local observed climate (**predictand *Y***) with the global simulations given by the GCMs (**predictors *X***), through some **function *f*** and/or **parameters θ**





- **Adaptation: Consider a model for e predictors to maintain some cohere**

  Precip01h = a01 SLP + b01 Q850

  ……………………

  Precip24h = a24 SLP + b24 Q850

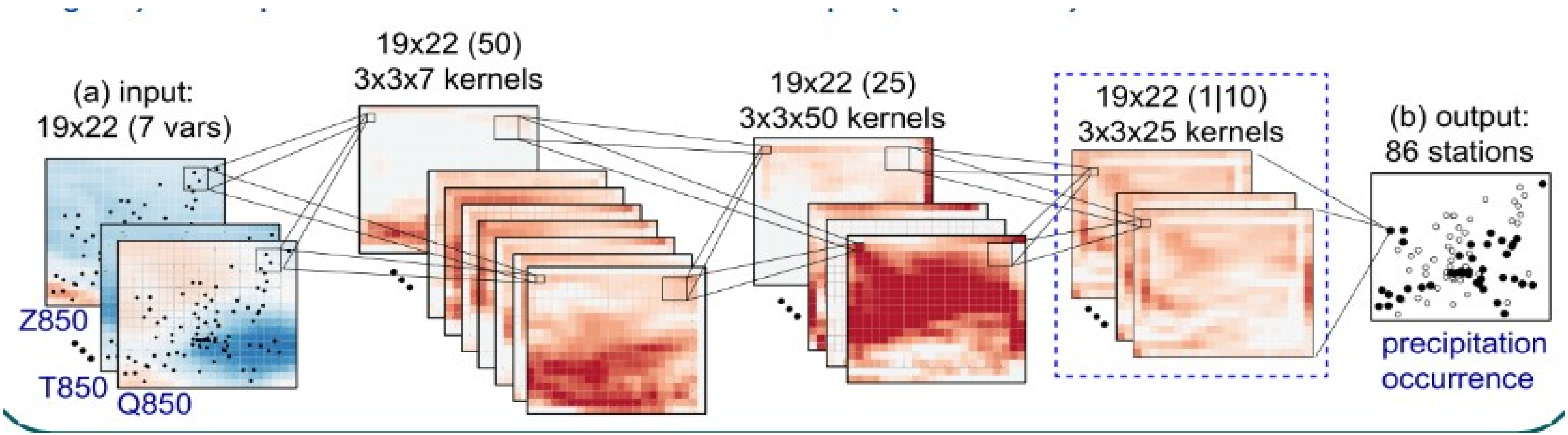
- **Adaptation: Relate daily and sub-daily predictors (e.g. Martin et al. 2021).**

Martin, TCM, Rocha, HR, Perez, GMP. Fine scale surface climate in complex terrain using machine learning. Int J Climatol. 2021; 41: 233– 250. https://doi-org.unican.idm.oclc.org/10.1002/joc.6617

Pros and Cons

Transfer Functions

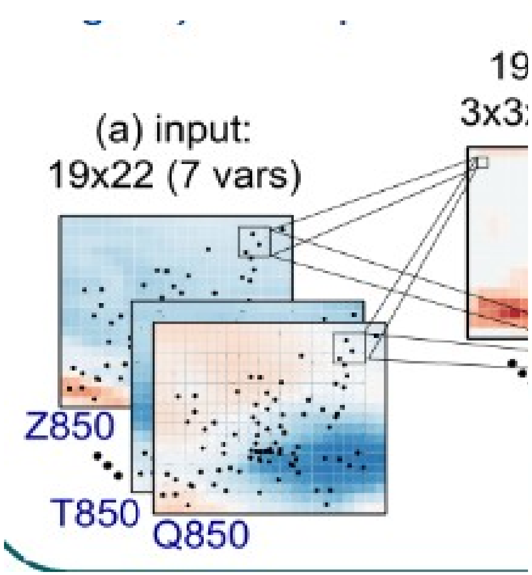
- **Transitions between the last hour of the previous day and the first one of the target day could not be preserved.**
- **Physical and spatial coherence between variables and location could not be preserved.**
- **These methods are commonly strongly dependent on the region, variable, season, etc. → New approaches, as Convolution Neural Networks, can partially solve this problem.**

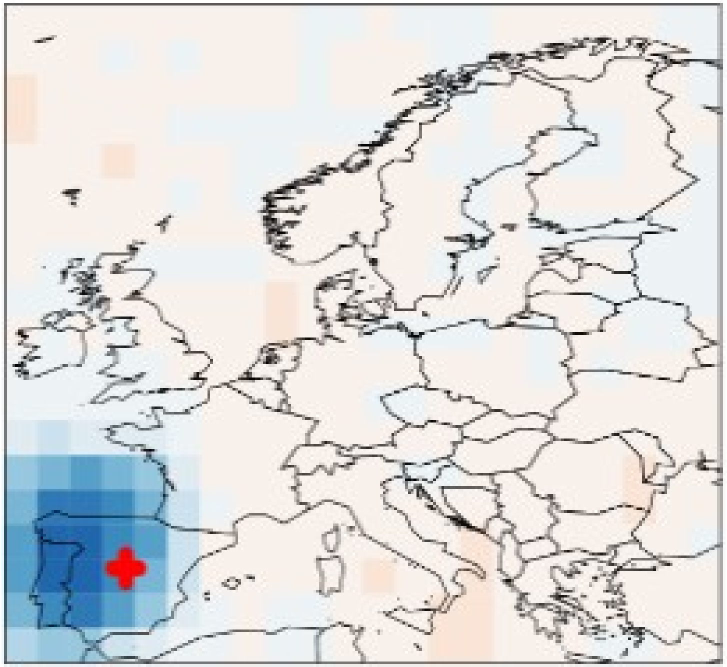


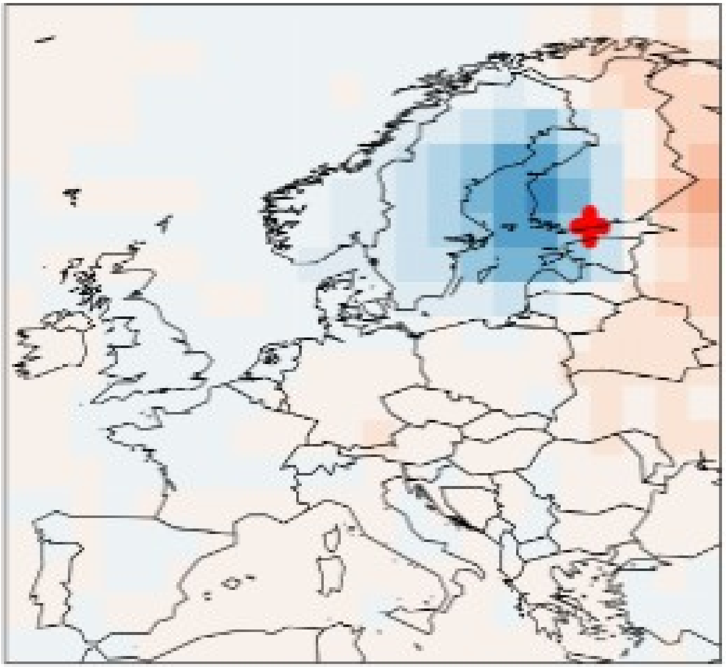
Pros and Cons

Transfer Functions

- **Transitions between the last hour of the previous day and the first one of the target day could not be preserved.**
- **Physical and spatial coherence between variables and location could not be preserved.**
- **These methods are commonly strongly dependent on the region, variable, season, etc. → New approaches, as Convolution Neural Networks, can partially solve this problem.**



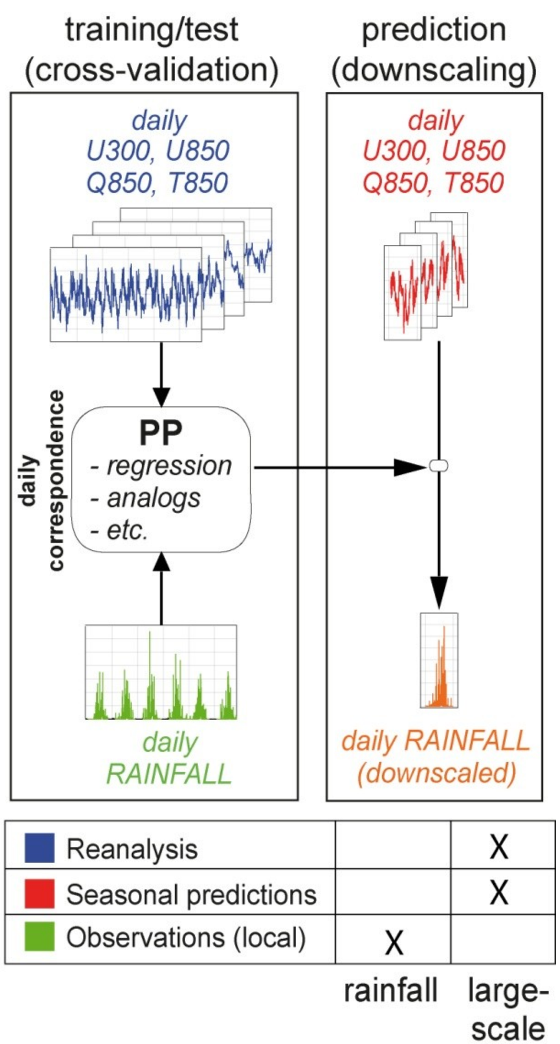


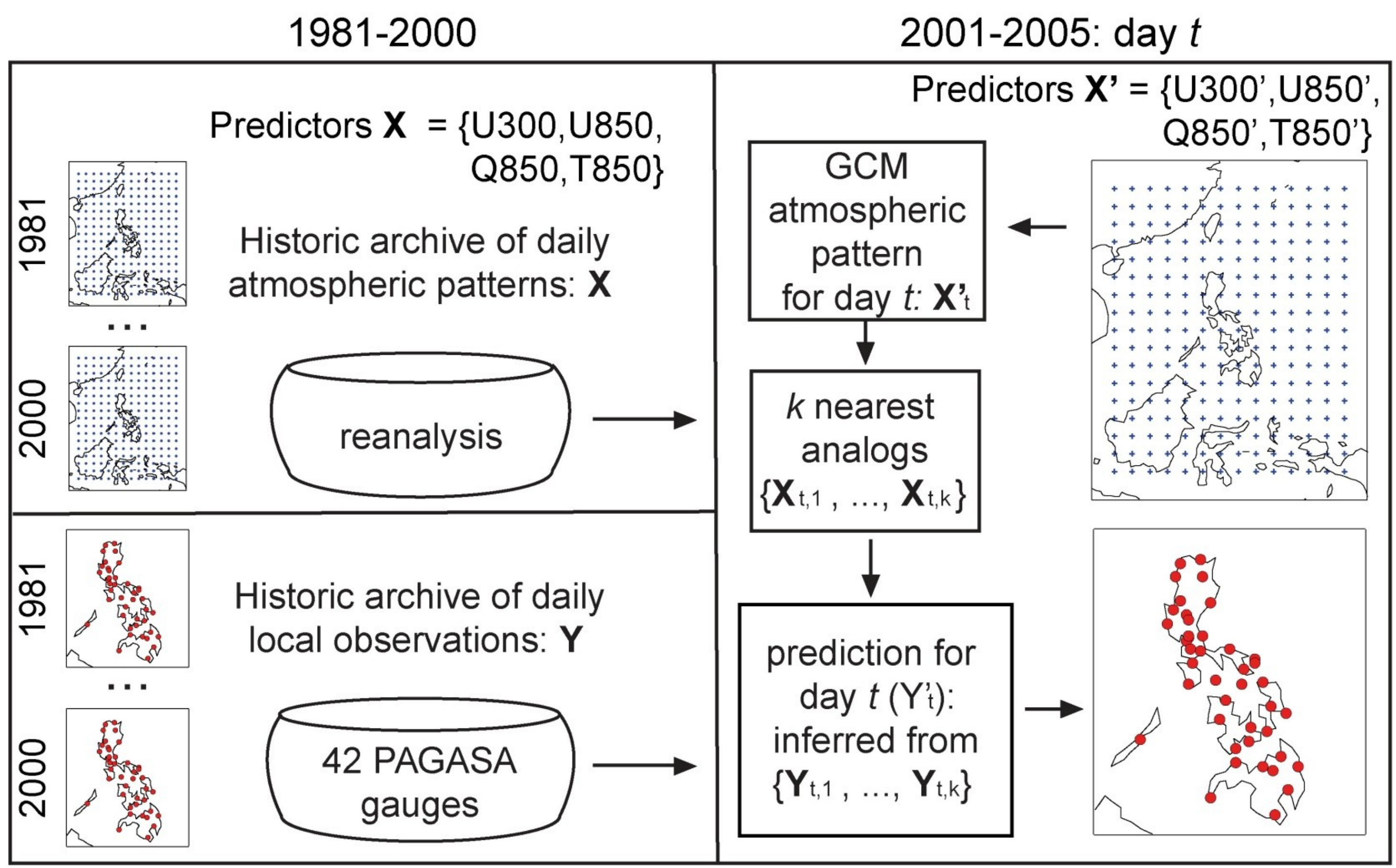


Statistical Downscaling

Analogues Weather Typing

Similar atmospheric patterns lead to similar meteorological conditions (Lorenz, 1969). Based on the K-nearest neighbor algorithm.





Analogues Weather Typing

Similar atmospheric patterns lead to similar meteorological conditions (Lorenz, 1969). Based on the K-nearest neighbor algorithm.

- **Deterministic Adaptation: Consider the sub-daily data observed for the analogue day.**
- **Stochastic Adaptation: Randomly chosen a day from the k-nearest and consider the sub-daily data observed for this day.**
- **Adaptation: Weather generator based on k-nearest neighbors.**

P. Willems, M. Vrac, Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change, Journal of Hydrology, 402 (3-4), 2011, Pages 193-205, https://doi.org/10.1016/j.jhydrol.2011.02.030.

Pros and Cons

Transfer Functions

- Transitions between the last hour of the previous day and the first one of the target day could not be preserved.
- Physical and spatial coherence between variables and location could not be preserved.
- These methods are commonly strongly dependent on the region, variable, season, etc.

Analogues Weather Typing

- The optimum number of nearest neighbors should be chosen.
- A purely analogue method can not obtain non-observed values or sub-daily patterns.
- Physical and spatial coherence between variables and location is preserved if the predictors and the analogue day are the same.

Weather Generators

Stochastic models able to generate artificial time series with similar statistical properties than the observed time series.

Weather Generators have been commonly used to generate artificial time series, mainly for precipitation based on:

- **A 1-lag Markov Model to simulate the occurrence:**

$$p_{01}(k) = P[X_t(k) = 1 | X_{t-1}(k) = 0]$$
$$p_{11}(k) = P[X_t(k) = 1 | X_{t-1}(k) = 1].$$

- **A theoretical statistical distribution for the precipitation amount:**

$$f(x) = \frac{(x/\beta)^{\alpha-1} e^{-x/\beta}}{\beta\Gamma(\alpha)}; \quad x, \alpha, \beta > 0.$$

- **All the parameters are adjusted using observations.**

http://www.ipcc-data.org/guidelines/pages/weather_generators.html

De Vera, A., & Terra, R. (2018). A Stochastic Precipitation Generator Conditioned by a Climate Index, Journal of Applied Meteorology and Climatology, 57(11), 2585-2603. Retrieved Jan 9, 2021, from https://journals.ametsoc.org/view/journals/apme/57/11/jamc-d-17-0307.1.xml

Weather Generators

Stochastic models able to generate artificial time series with similar statistical properties than the observed time series.

The most commonly approach is based on the work of Richarson (1981) and they are referred to as Richardson-type weather generators:

- **A multi-variate weather generator: precipitation, temperatures and solar radiation.**
- **The seasonal cycle is explicitly considered/adjusted.**
- **Lag-0 and lag-1 cross-correlations between variables are preserved by means of a first-order linear autoregessive model for the residuals.**
- **Temperature and solar radiation are conditioned to the precipitation occurrence.**

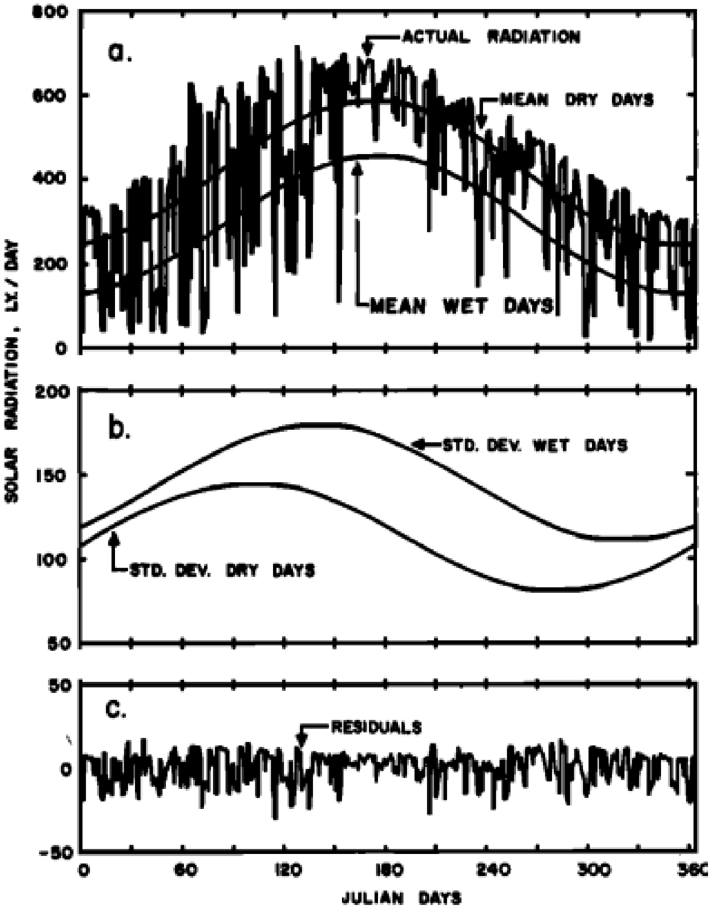


Fig. 1. Technique for reducing a daily solar radiation series to a series of residual elements, conditioned on the wet or dry status of the day.

Stochastic simulation of daily precipitation, temperature, and solar radiation	Richardson, C.W.	1981	Water Resources Research 17(1), pp. 182-190

Weather Generators

Stochastic models able to generate artificial time series with similar statistical properties than the observed time series.

The most commonly approach is based on the work of Richarson (1981) and they are referred to as Richardson-type weather generators:

- **A multi-variate weather generator: precipitation, temperatures and solar radiation.**
- **The seasonal cycle is explicitly considered/adjusted.**
- **Lag-0 and lag-1 cross-correlations between variables are preserved by means of a first-order linear autoregessive model for the residuals.**
- **Temperature and solar radiation are conditioned to the precipitation occurrence. ← Some authors don't consider this dependence.**
- **Has been generalized to a multi-site approach and/or other variables (Peleg et al. 2017, Legasa et al. 2020).**

Legasa, M. N., & Gutiérrez, J. M. (2020). Multisite weather generators using Bayesian networks: An illustrative case study for precipitation occurrence. Water Resources Research, 56, e2019WR026416. https://doi.org/10.1029/2019WR026416

Peleg, N., S. Fatichi, A. Paschalis, P. Molnar, and P. Burlando (2017), An advanced stochastic weather generator for simulating 2-D high-resolution climate variables, J. Adv. Model. Earth Syst., 9, 1595–1627, doi:10.1002/2016MS000854.

Pros and Cons

Transfer Functions

- Transitions between the last hour of the previous day and the first one of the target day could not be preserved.
- Physical and spatial coherence between variables and location could not be preserved.
- These methods are commonly strongly dependent on the region, variable, season, etc.

Analogues Weather Typing

- The optimum number of nearest neighbors should be chosen.
- A purely analogue method can not obtain non-observed values or sub-daily patterns.
- Physical and spatial coherence between variables and location is preserved if the predictors and the analogue day are the same.

Weather Generators

- Parametric approaches should define the theoretical distributions of the different variables considered. In addition, the relations between variables is also established by the auto-regressive model considered.
- The statistical parameters adjusted by the weather generator are predefined by the user leading to very local and overfitted models.
- The transitions between days could be adjusted.

Analogue-based downscaling

Most of the temporal statistical downscaling methods are based on function/algorithm/weather generator applied to a previous set of k analogue days.

In summary, this is the basic algorithm:

- **K-nearest neighbors are obtained with daily data:** $[Y_1^d, \dots, Y_k^d]$
- **For each neighbor, the distance to the target day (*d*) is estimated:**

$$D_1^d = \sqrt{\sum (Y_1^d - d)^2}$$

- **This distance is used to obtain the hourly values of the target day (*d*):**

$$[d_{01}, \dots, d_{24}]$$

- **The hourly values are rescaled to preserve the daily value.**

De Vera, A., & Terra, R. (2018). A Stochastic Precipitation Generator Conditioned by a Climate Index, Journal of Applied Meteorology and Climatology, 57(11), 2585-2603. Retrieved Jan 9, 2021, from https://journals.ametsoc.org/view/journals/apme/57/11/jamc-d-17-0307.1.xml

Analogue-based downscaling

Most of the temporal statistical downscaling methods are based on function/algorithm/weather generator applied to a previous set of k analogue days.

In summary, this is the basic algorithm:

- **K-nearest neighbors are obtained with daily data:** $[Y_1^d, \ldots, Y_k^d]$
- **For each neighbor, the distance to the target day (*d*) is estimated:**

$$D_1^d = \sqrt{\sum (Y_1^d - d)^2}$$

- **This distance is used to obtain the hourly values of the target day (*d*):**

$$[d_{01}, \ldots, d_{24}]$$

- **The hourly values are rescaled to preserve the daily value.**

The differences between algorithms are on this step.

Rajagopalan, B., and Lall, U. (1999), A k-nearest-neighbor simulator for daily precipitation and other weather variables, Water Resour. Res., 35(10), 3089– 3101, doi:10.1029/1999WR900028.

Analogue-based downscaling

Most of the temporal statistical downscaling methods are based on function/algorithm/weather generator applied to a previous set of k analogue days.

In summary, this is the basic algorithm:

- **K-nearest neighbors are obtained with daily data:** $[Y_1^d, \dots, Y_k^d]$
- **For each neighbor, the distance to the target day (*d*) is estimated:**

$$D_1^d = \sqrt{\sum \left(Y_1^d - d\right)^2}$$

- **This distance is used to obtain the hourly values of the target day (*d*):**

$$[d_{01}, \dots, d_{24}]$$

  - **Some authors consider a random resampling.**
  - **Genetic algorithms (reproduction, crossover and mutation).**
  - **...**

Rajagopalan, B., and Lall, U. (1999), A k-nearest-neighbor simulator for daily precipitation and other weather variables, Water Resour. Res., 35(10), 3089– 3101, doi:10.1029/1999WR900028.

Buishand, T. A., and Brandsma, T. (2001), Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest□neighbor resampling, Water Resour. Res., 37(11), 2761– 2776, doi:10.1029/2001WR000291.

Taesam Lee, Changsam Jeong, Nonparametric statistical temporal downscaling of daily precipitation to hourly precipitation and implications for climate change scenarios, Journal of Hydrology, 510, 2014, Pages 182-196, https://doi.org/10.1016/j.jhydrol.2013.12.027.

Pros and Cons

Analogue-based downscaling

- **Are the most commonly used methods.**
- **Physical and spatial coherence between variables and locations could be preserved, or partially preserved, by means of the analogue method.**
- **The transitions between days could be adjusted.**
- **Non-observed values or sub-daily patterns could be obtained.**

Pros and Cons

Analogue-based downscaling

- **Are the most commonly used methods.**
- **Physical and spatial coherence between variables and locations could be preserved, or partially preserved, by means of the analogue method.**
- **The transitions between days could be adjusted.**
- **Non-observed values or sub-daily patterns could be obtained.**

Thank you !!!!

Conclusions

Transfer Functions

- Transitions between the last hour of the previous day and the first one of the target day could not be preserved.
- Physical and spatial coherence between variables and location could not be preserved.
- These methods are commonly strongly dependent on the region, variable, season, etc.

Analogues / Weather Typing

- The optimum number of nearest neighbors should be chosen.
- A purely analogue method can not obtain non-observed values or sub-daily patterns.
- Physical and spatial coherence between variables and location is preserved if the predictors and the analogue day are the same.

Weather Generators

- Parametric approaches should define the theoretical distributions of the different variables considered. In addition, the relations between variables is also established by the auto-regressive model considered.
- The statistical parameters adjusted by the weather generator are predefined by the user leading to very local and overfitted models.
- The transitions between days could be adjusted.

Analog-based algorithms

- Are the most commonly used methods.
- Physical and spatial coherence between variables and locations could be preserved, or partially preserved, by means of the analogue method.
- The transitions between days could be adjusted.
- Non-observed values or sub-daily patterns could be obtained.