CERFACS

CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

Statistical downscaling and bias correction

Christian Pagé
christian.page@cerfacs.fr

IS-ENES3 Central&Eastern Europe Autumn School
From Climate Projections to Climate impacts via Regional Downscaling
Prague, 28 November-2 December 2022

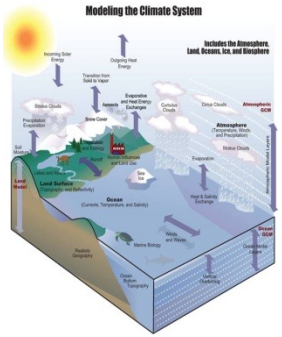
Structure

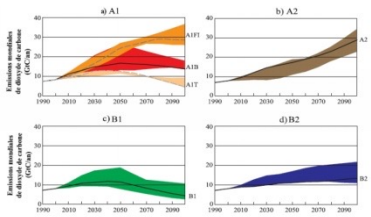
- General principles
- Methods
  - Simples
  - Dynamical: Bias correction
  - Statistical
    - General
    - Analogs
    - Weather Types
    - Validation
    - Uncertainties
  - Limitations
- Conclusions

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Climate Projections

Several climate models: *uncertainties*

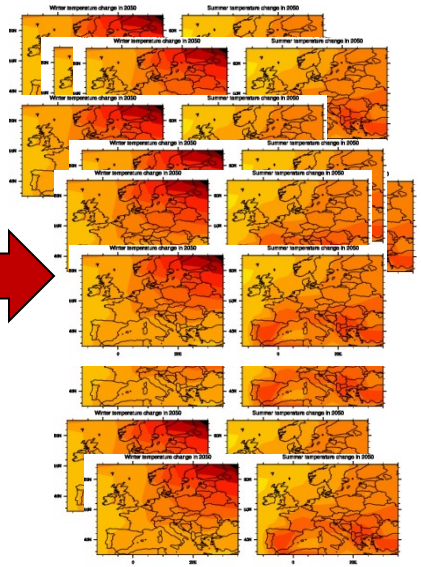




Greenhouse Gas Scenarios
More uncertainties

Only a few very large centres

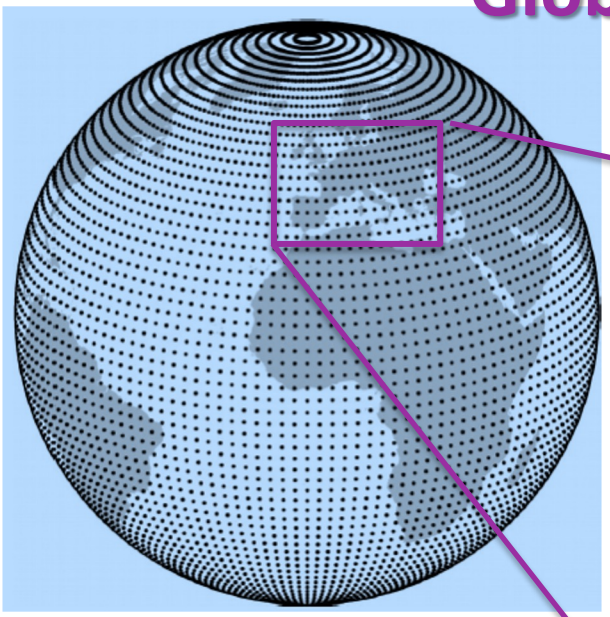
Very high computational costs



IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

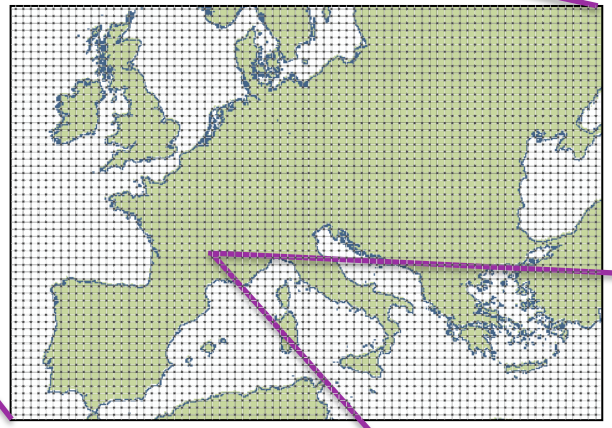
Downscaling: a required step

Global



Resolution:
≈ 100 km

Regional

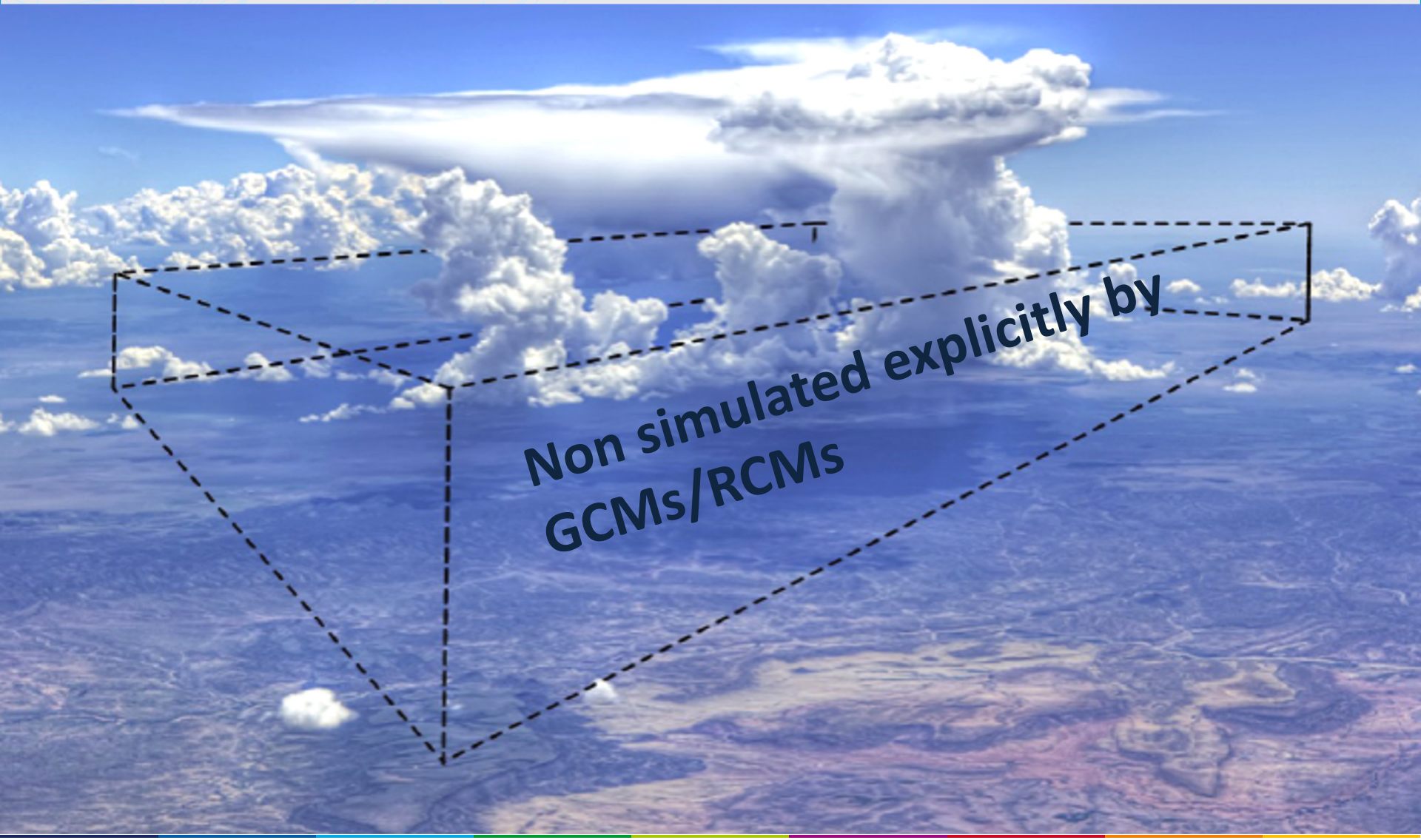


Resolution:
≈ 10 km

Local

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

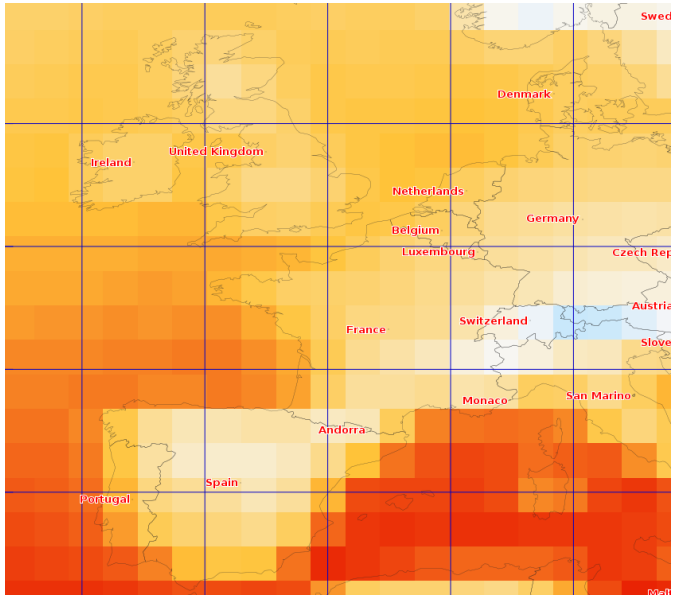
Downscaling and Bias Correction: a required step



IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Downscaling: a required step

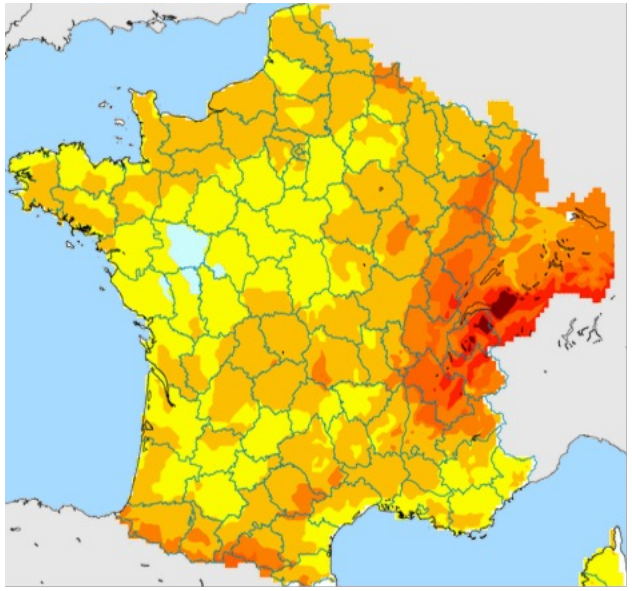
Typical spatial resolution of a global climate model CMIP6



≈125 km

Downscaling

Needed spatial resolution for impact modeling and studies

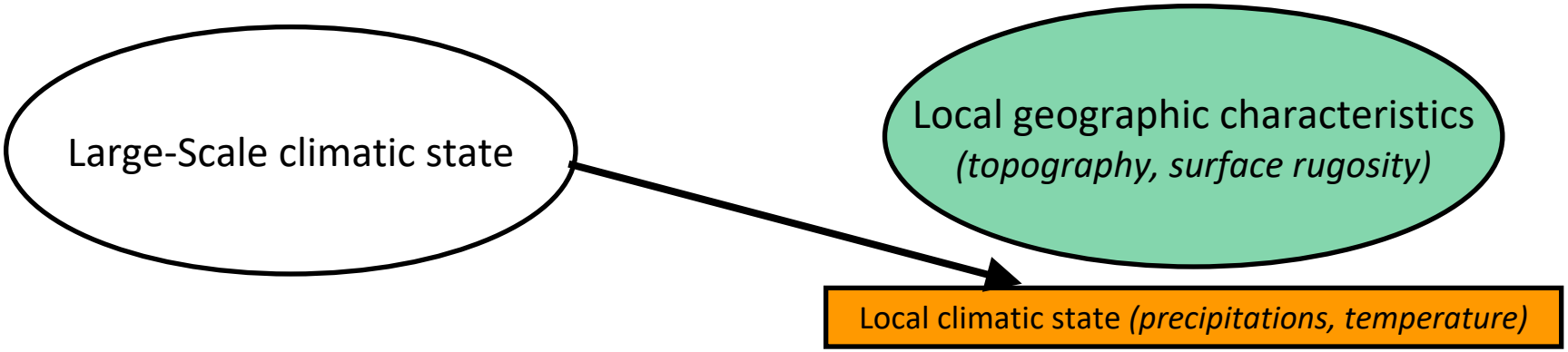


8 km

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Downscaling: a required step

Regionalisation / Downscaling



Statistical

Build a statistical model using relationships between local and large-scale climate variables

Dynamical

Resolve explicitly the physical equations of the regional climate system

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Simple Statistical Downscaling Method

Anomaly Method (Delta Method)

- Time Series of observations
- Deltas (anomalies) of climate models

Principles

- Reproduce a time series of observations in the future by repeating the time series
- Apply a tendency from a climate model simulation

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Simple Statistical Downscaling Method

Advantages

- Easy to interpret
- Reproduce the correct inter-daily and inter-annual variabilities
- No systematic bias

Limitations

- Does not reproduce the possible changes of future climate variabilities
- Decadal cycles are reproduced exactly like in the past
- Raises artificially the frequency of occurrence of extreme events that have a long return period

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

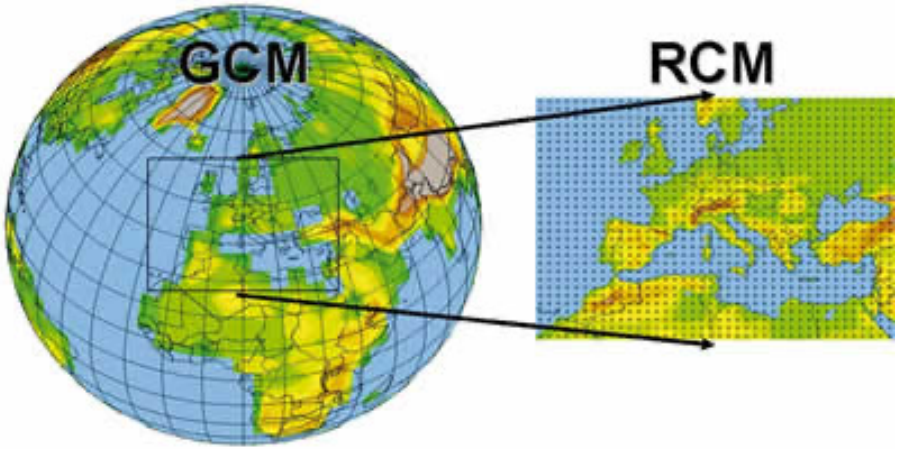
Dynamical Methods with Bias Correction

Based on

- Higher Spatial resolution numerical models (10 — 50 km)

How to deal with the computational costs ?

- Variable grid projection
- Limited Area Model
- Limited number of simulations



Hidden Hypothesis

- Statistical relationships of physical parameterisations will be identical in the future climate

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Dynamical Methods

Still biases...?

- Spatial resolution: still not fine enough
- Vertical Interpolation
- Simplified representation of surface interactions
- Simplified Microphysics

→ Bias correction is needed

- In Weather Forecasting it is called Model Output Statistics (MOS): same exact idea

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

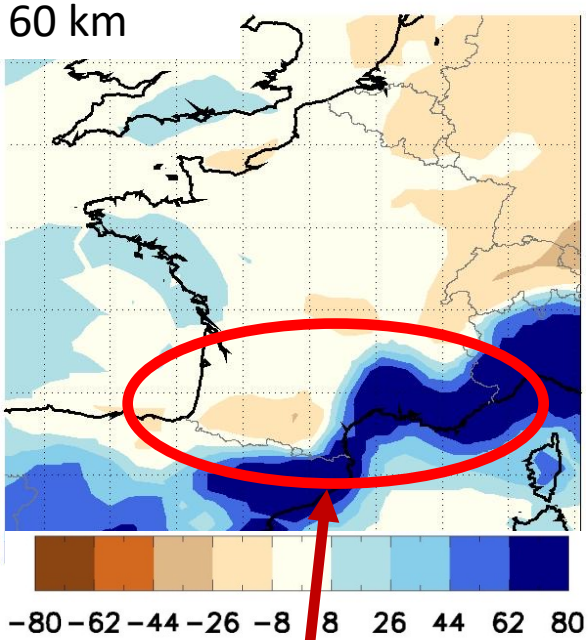
Bias Correction

High spatial resolution numerical model embedded in a global climate model

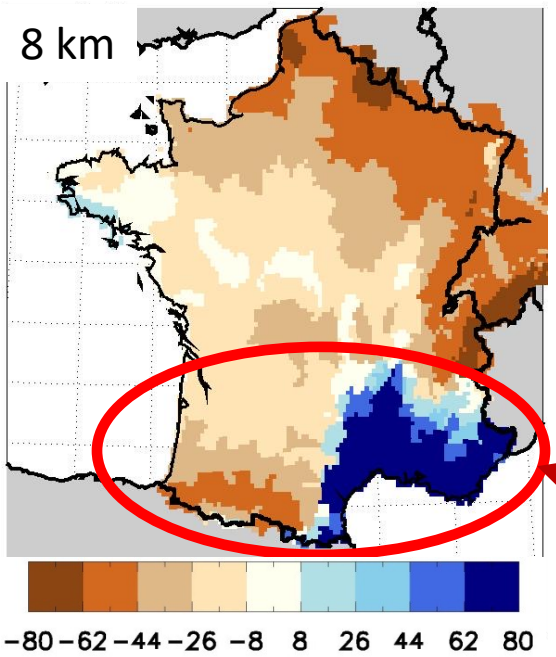
Precipitations Anomalies (%)

Regional Climate Model

60 km

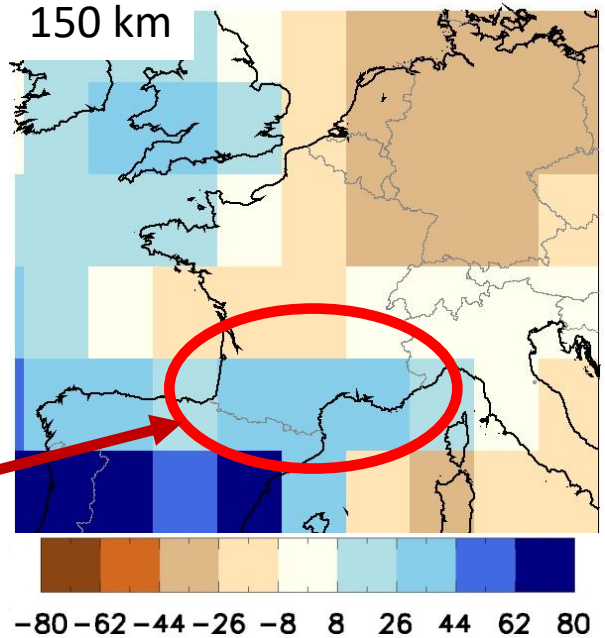


8 km



Observations

150 km



Global Climate Model

Biases partly corrected

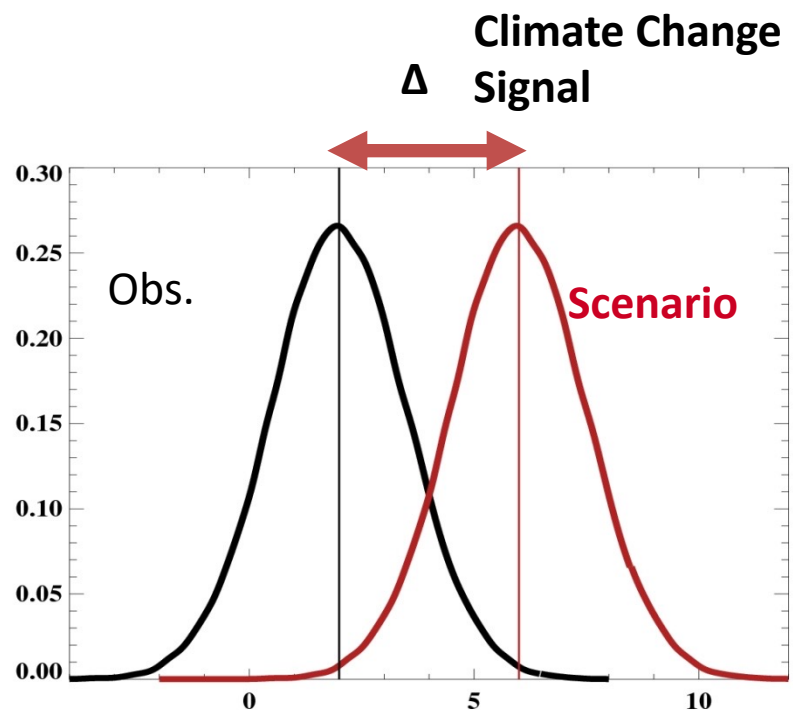
IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Bias Correction

High spatial resolution numerical model embedded in a global climate model

Bias Correction

=> Quantile-Quantile (Déqué, 2007)



PDF

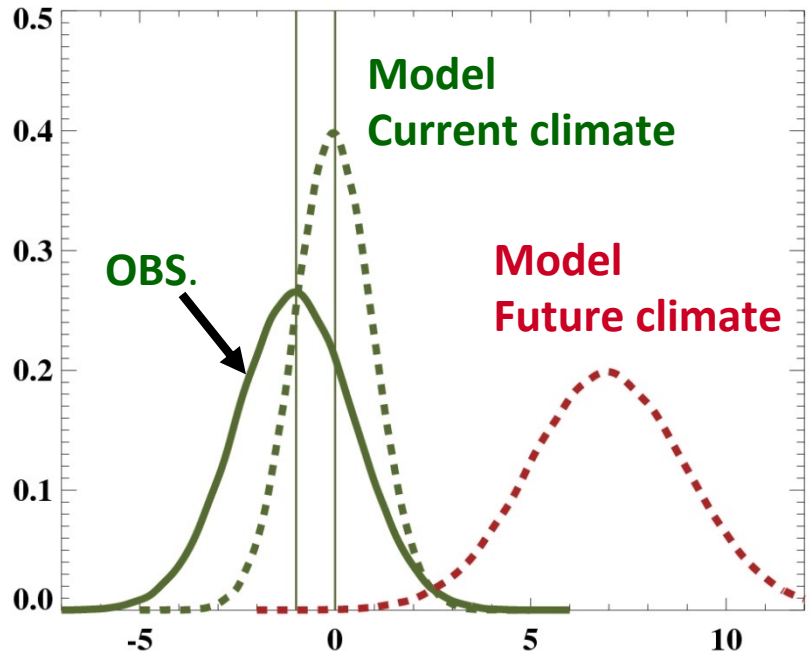
IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Bias Correction

High spatial resolution numerical model embedded in a global climate model

Bias Correction

=> Quantile-Quantile (Déqué, 2007)



PDF

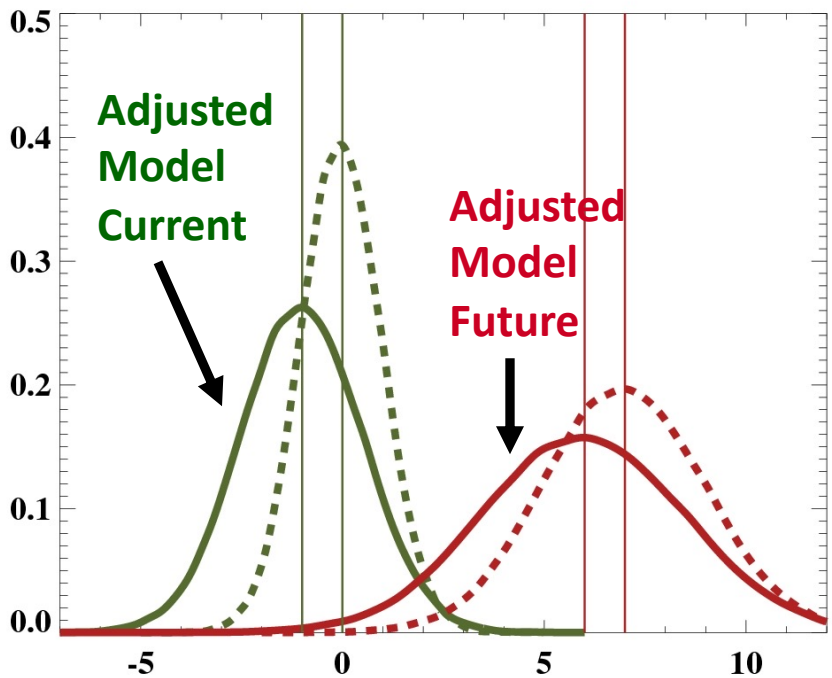
IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Bias Correction

High spatial resolution numerical model embedded in a global climate model

Bias Correction

=> Quantile-Quantile (Déqué, 2007)



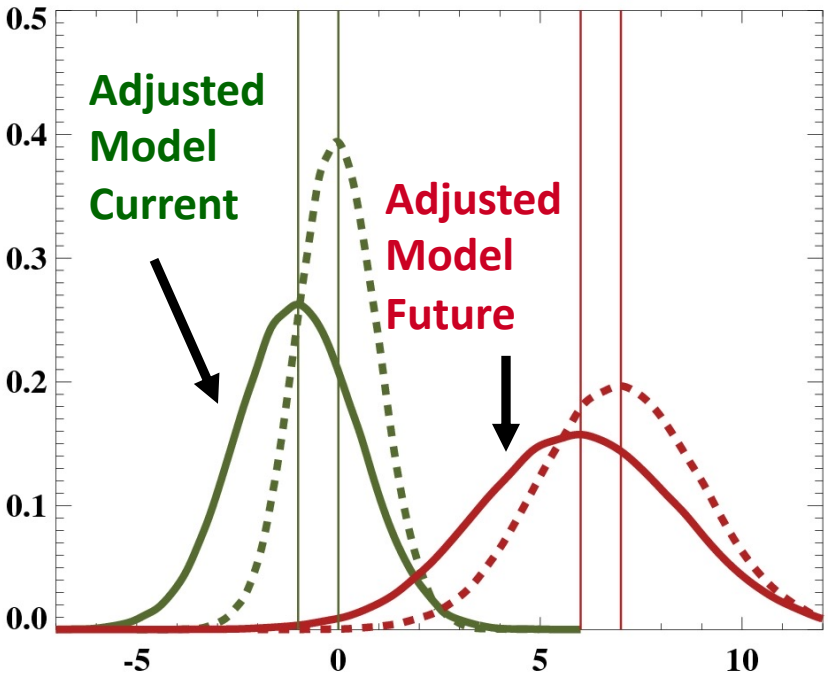
PDF

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Bias Correction

Bias correction: limitations?

- Covariances between variables are not taken into account
- Extremes: extrapolations needed if all values did not happen in the past
- Statistical relationships of the biases between observations and model variables do not change in the future



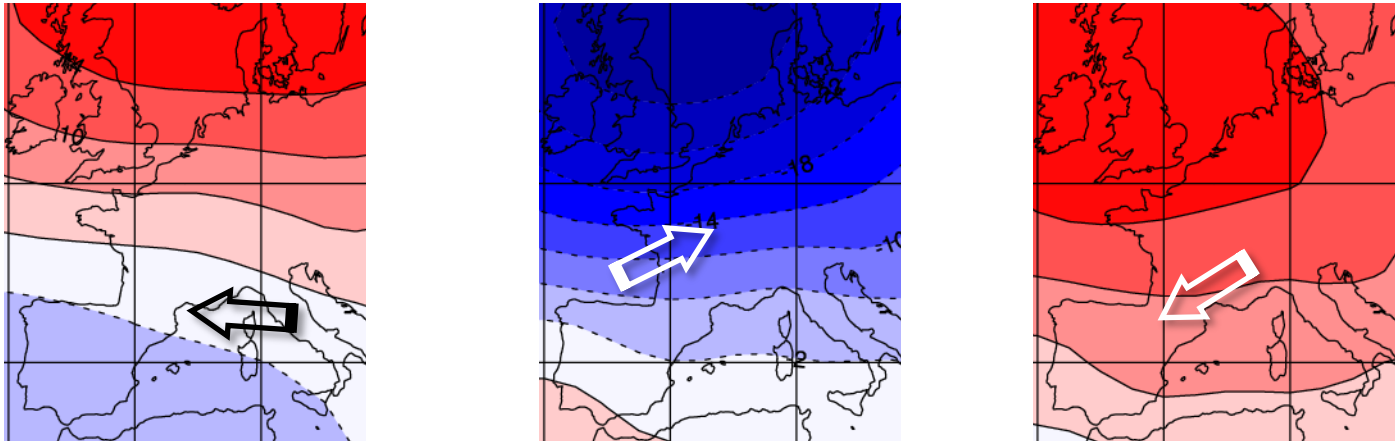
PDF

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

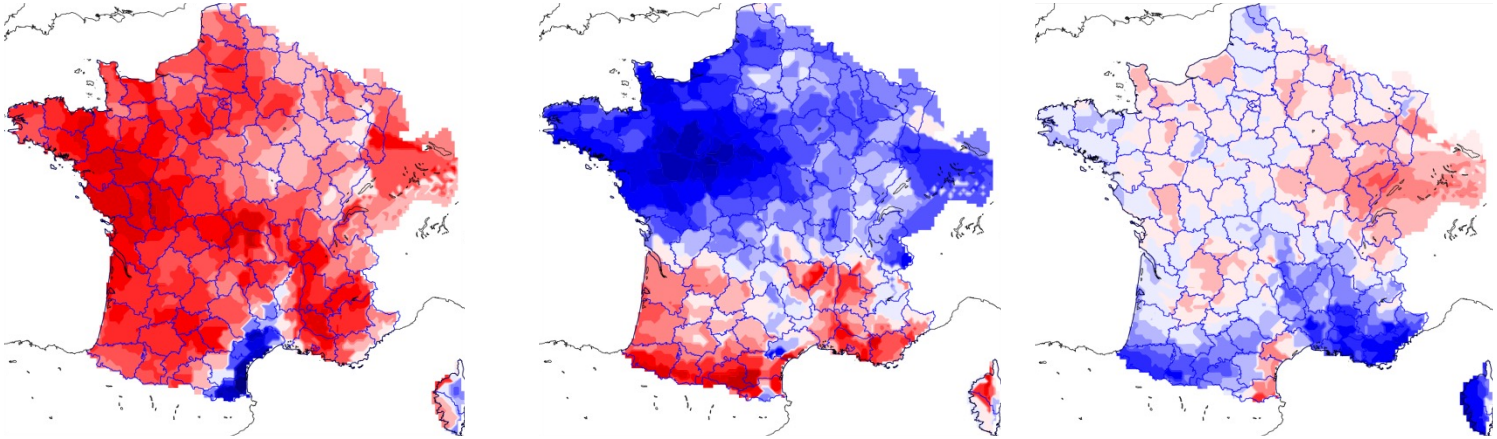
Statistical Methods

Examples of winter weather types and precipitation anomalies

Mean Sea-Level Pressure Anomalies



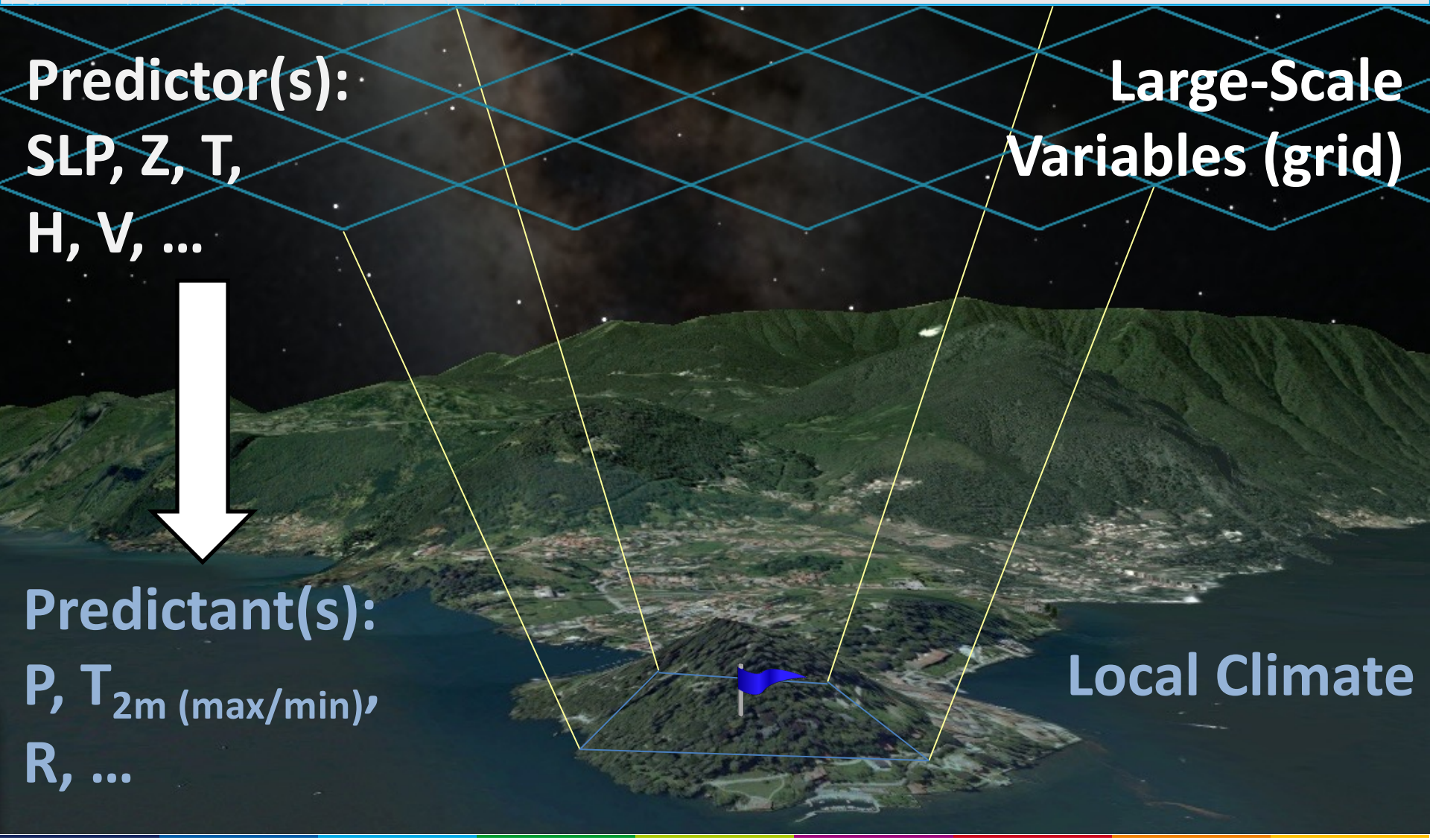
Physical Relationships



Precipitation Anomalies

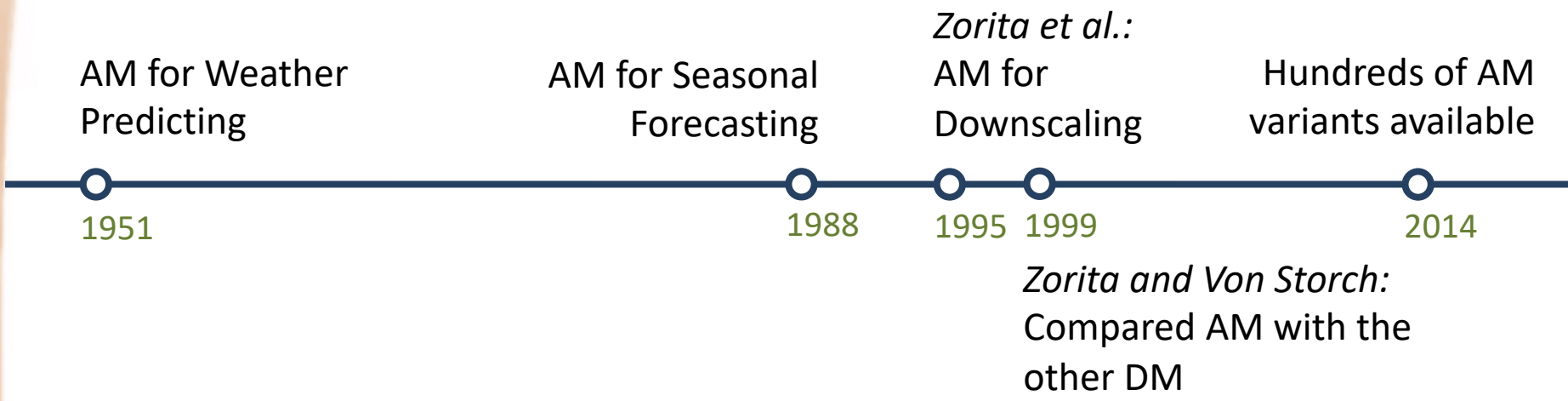
IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Statistical Methods



IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

A brief history of the Analogue Method (AM)

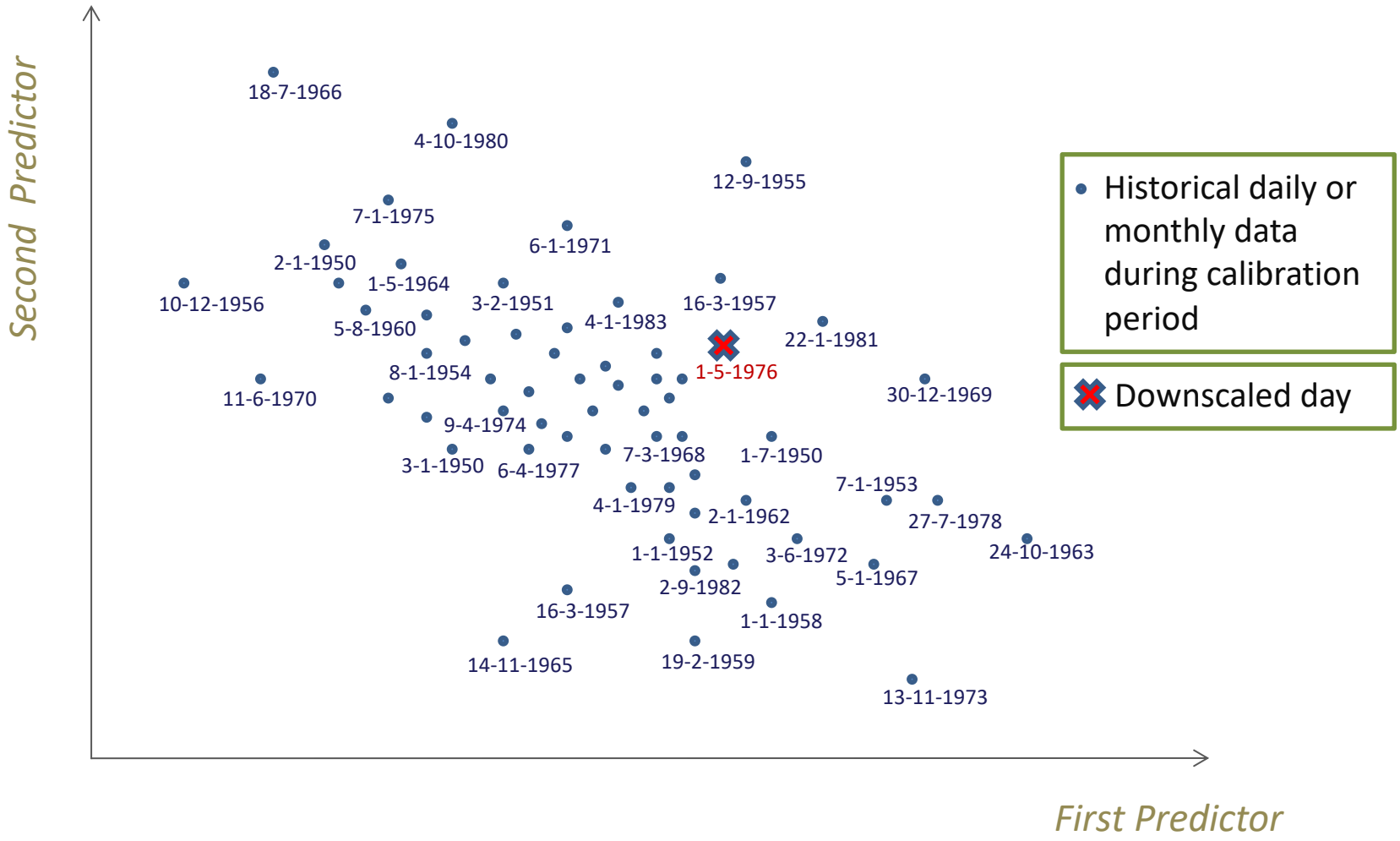


Slides Courtesy of Nicola Cortesi

Advantages of the AM

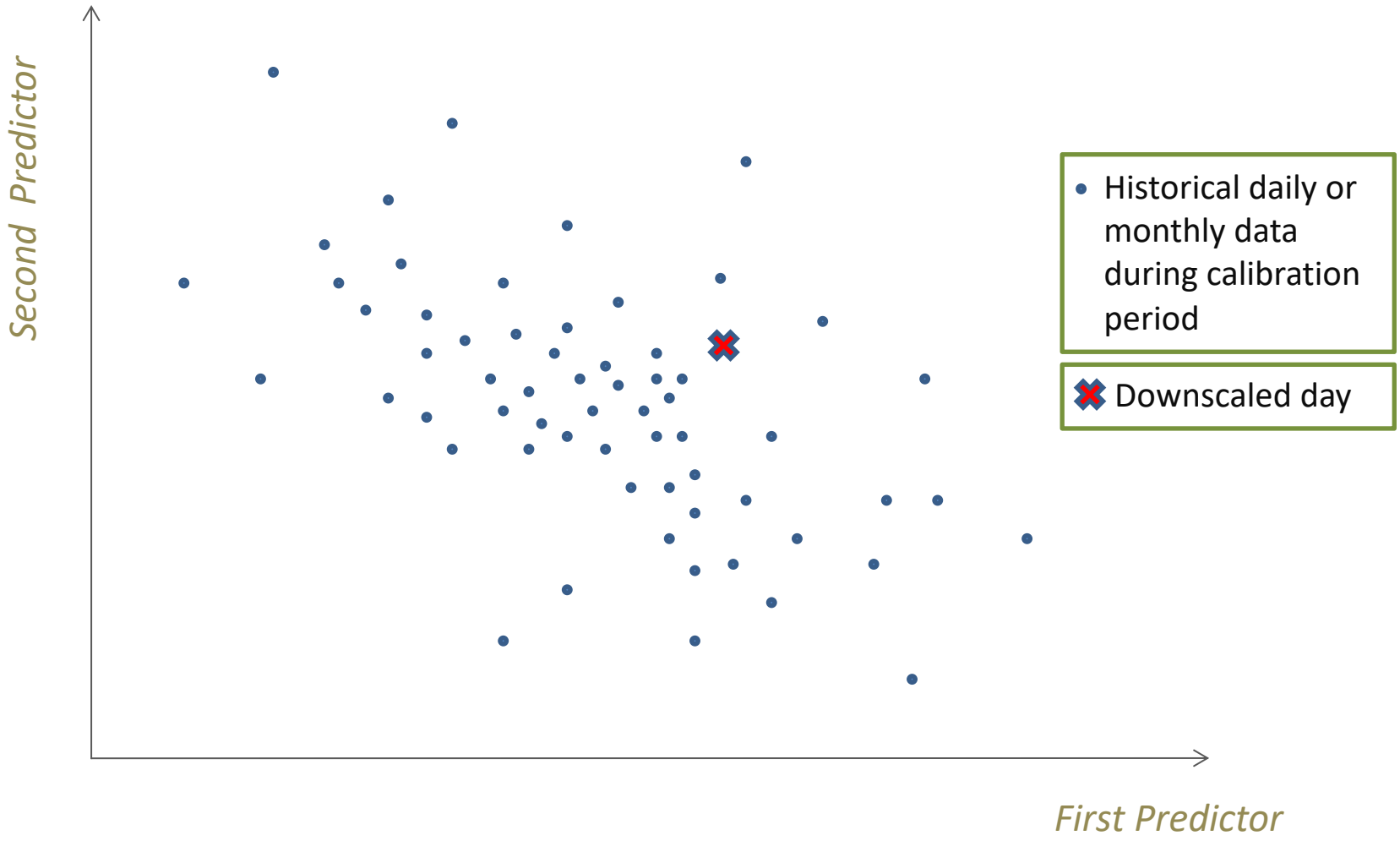
1. *Non-parametric*
2. *Non-linear*
3. *Do not underestimate the variance*
4. *Spatially Coherent*
5. *Easy to implement*
6. *Low computational cost*

How the AM works



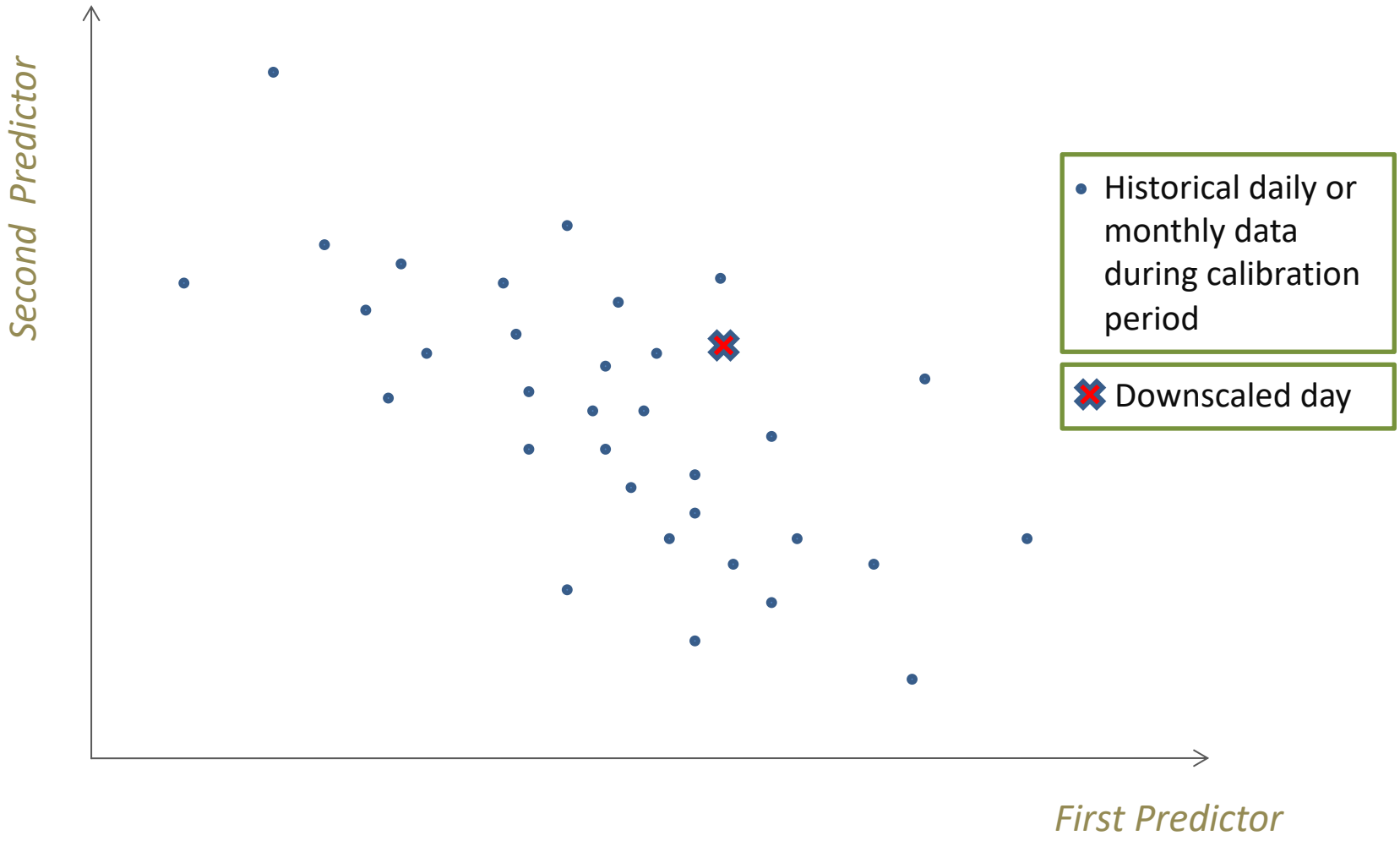
Slides Courtesy of Nicola Cortesi

Search Radius



Slides Courtesy of Nicola Cortesi

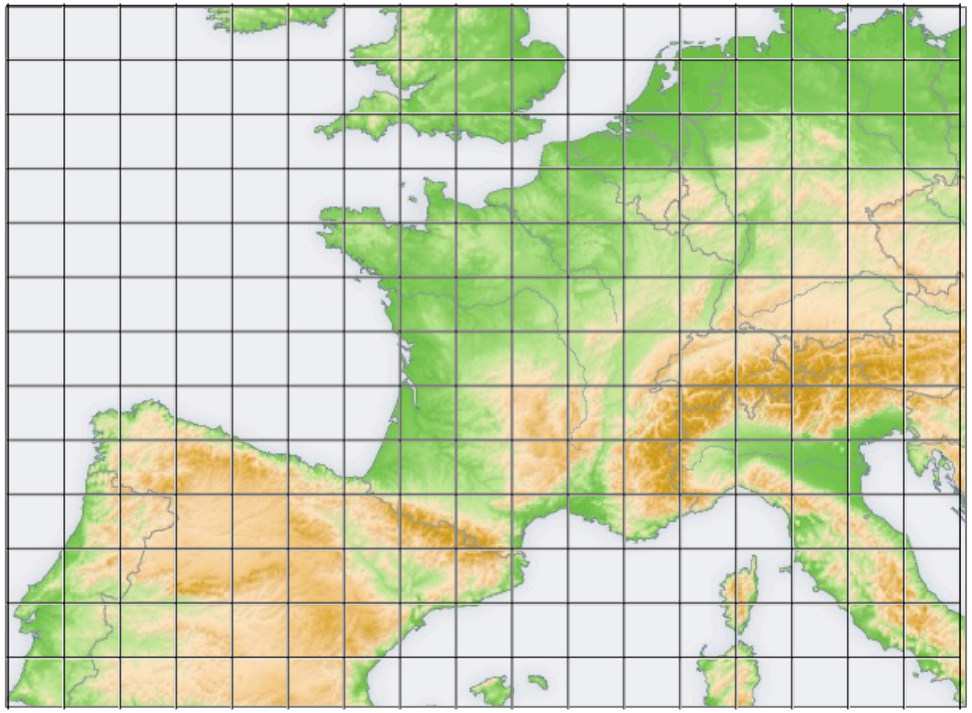
Search Radius



Slides Courtesy of Nicola Cortesi

How the AM works

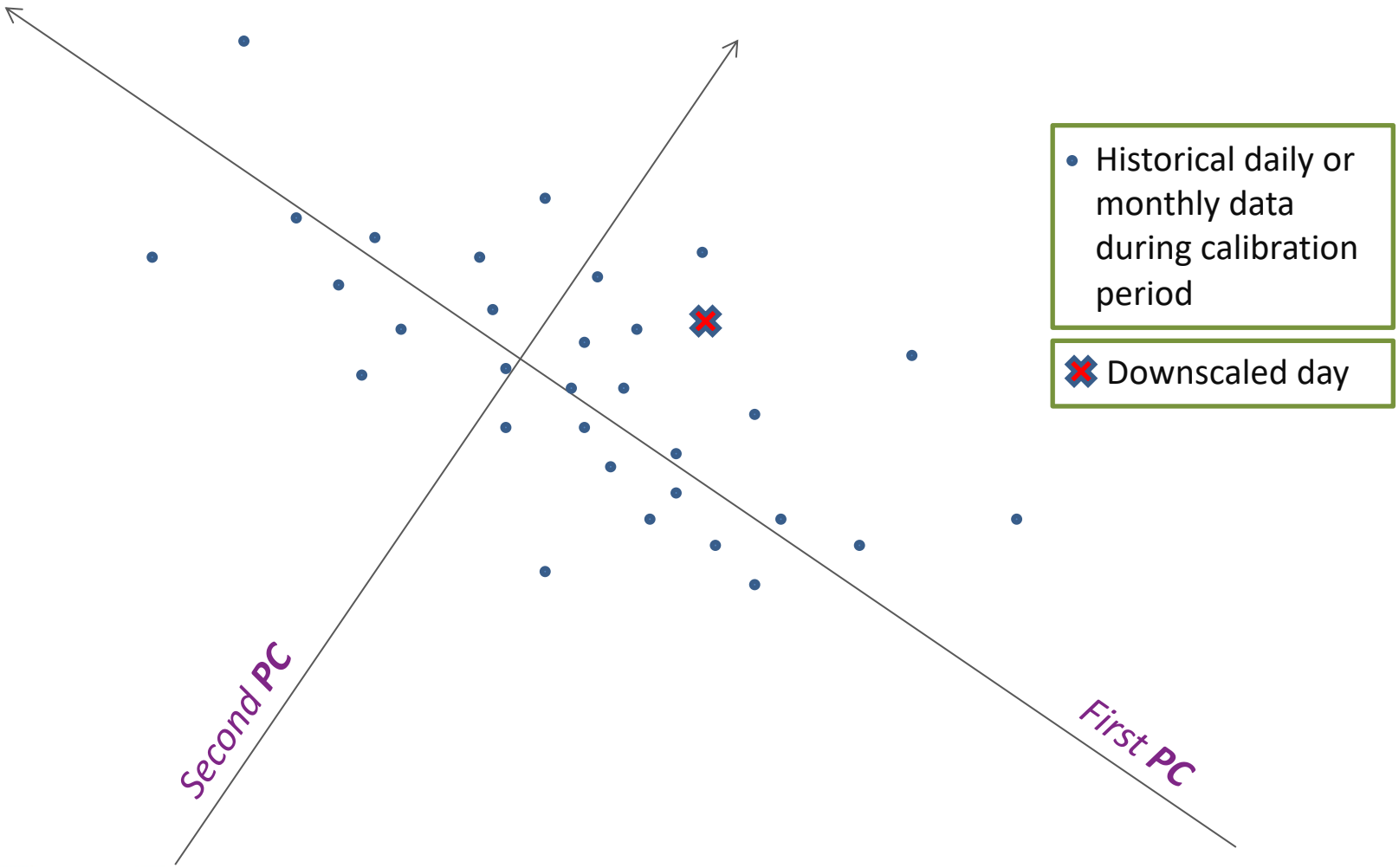
Typical Domain Size of a Predictor Field



PCA

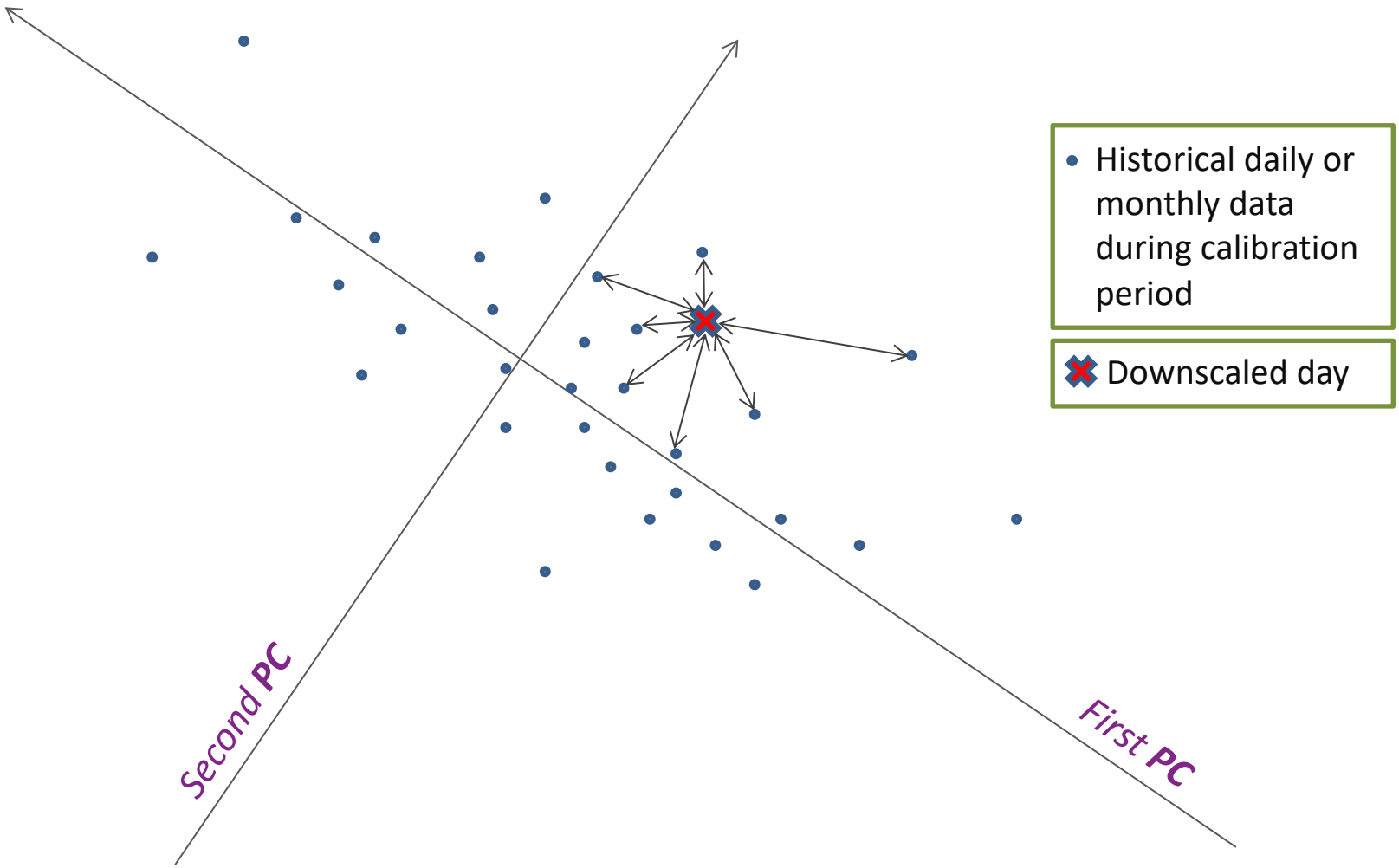
Slides Courtesy of Nicola Cortesi

How the AM works



Slides Courtesy of Nicola Cortesi

Similarity Measure



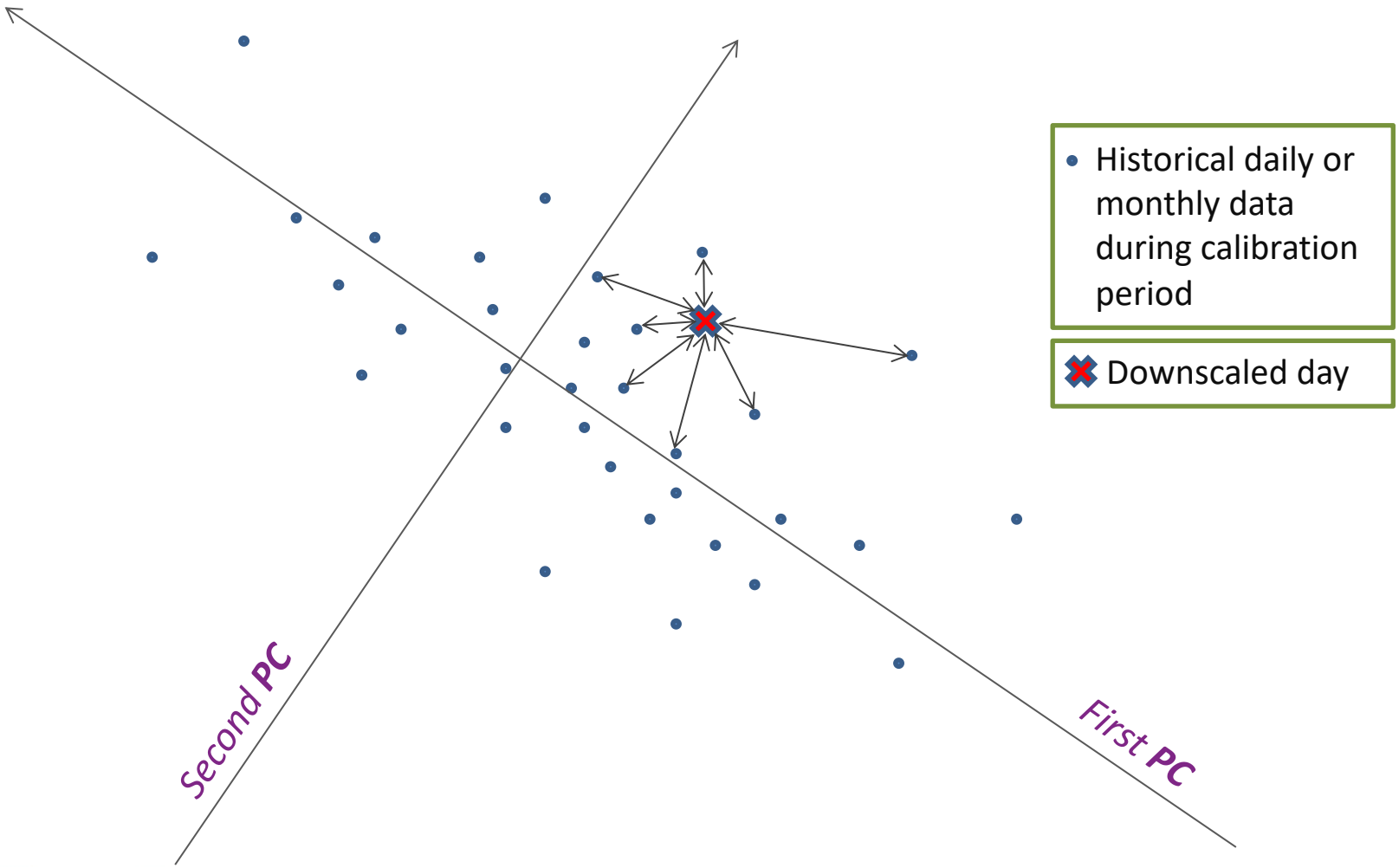
Slides Courtesy of Nicola Cortesi

Similarity Measure

- (Weighted) Euclidean distance: $$d(\vec{x},\vec{y}) = \sqrt{\sum_{i=1}^{n} w_i z_i^2}$$
- Sum of the absolute values of each component of ***z***: $$d(\vec{x},\vec{y}) = \sum_{i=1}^{n} |z_i|$$
- Cosine of the angle between ***x*** and ***y***: $$d(\vec{x},\vec{y}) = \cos\theta_{\vec{x},\vec{y}}$$
- Mahalanobis distance: $$d(\vec{x},\vec{y}) = -\sum_{i=1}^{n} \frac{x_i y_i}{\sqrt{\lambda_i}}$$
- Pattern correlation: $$d(\vec{x},\vec{y}) = r_{(P_1,P_2)}$$
- Teweles-Wobus score: $$d(\vec{x},\vec{y}) \approx \nabla P_1 - \nabla P_2$$
- Minimize the sequence of days instead: $$d(\vec{x},\vec{y}) = \sum_{t=0}^{7} w_t d(\vec{x}(t),\vec{y}(t))$$

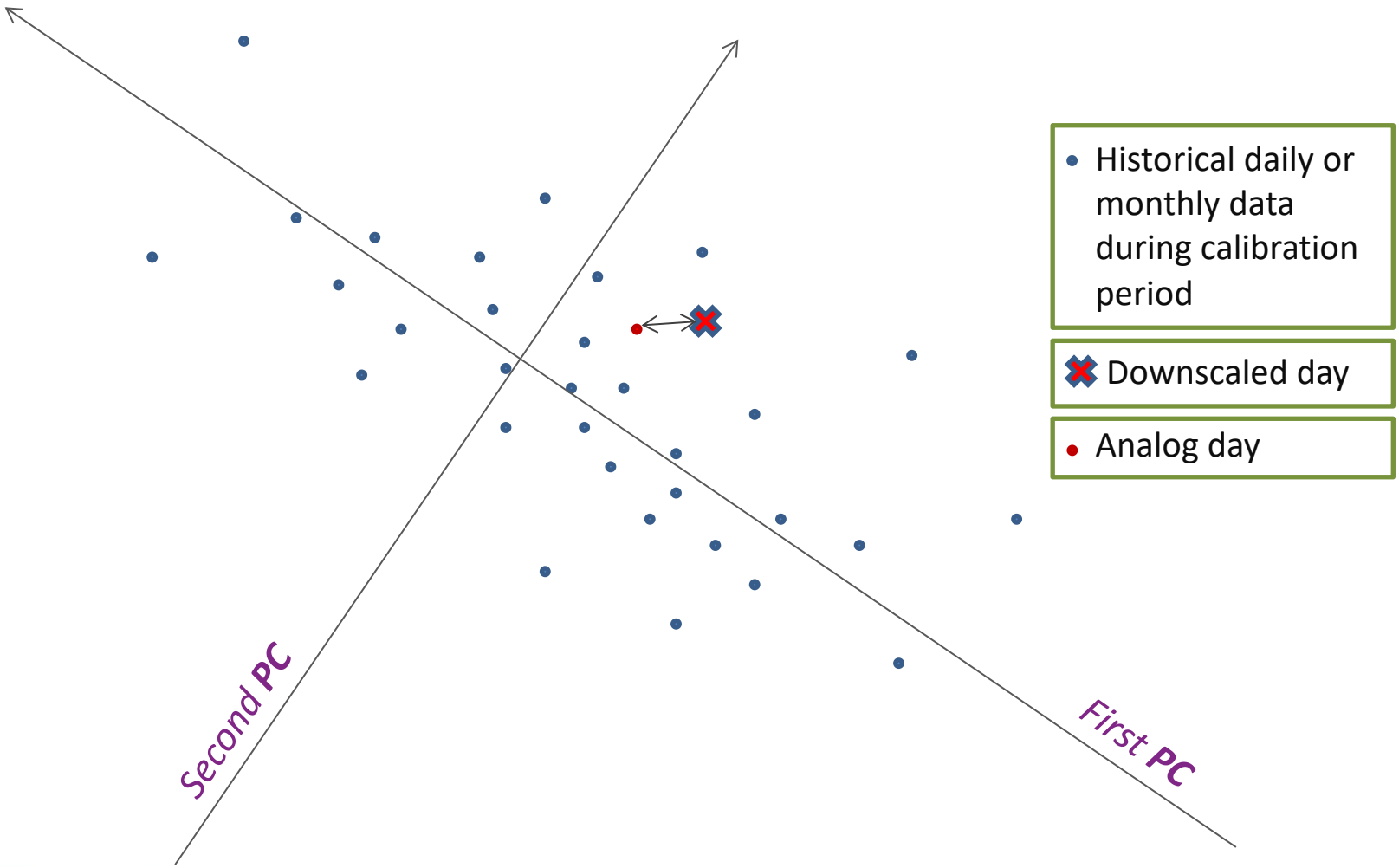
x and ***y*** are two vectors whose components are the values of the *n* Principal Components for the 1st and 2nd day, respectively, and: ***z*** = ***x*** - ***y***

Analog Day



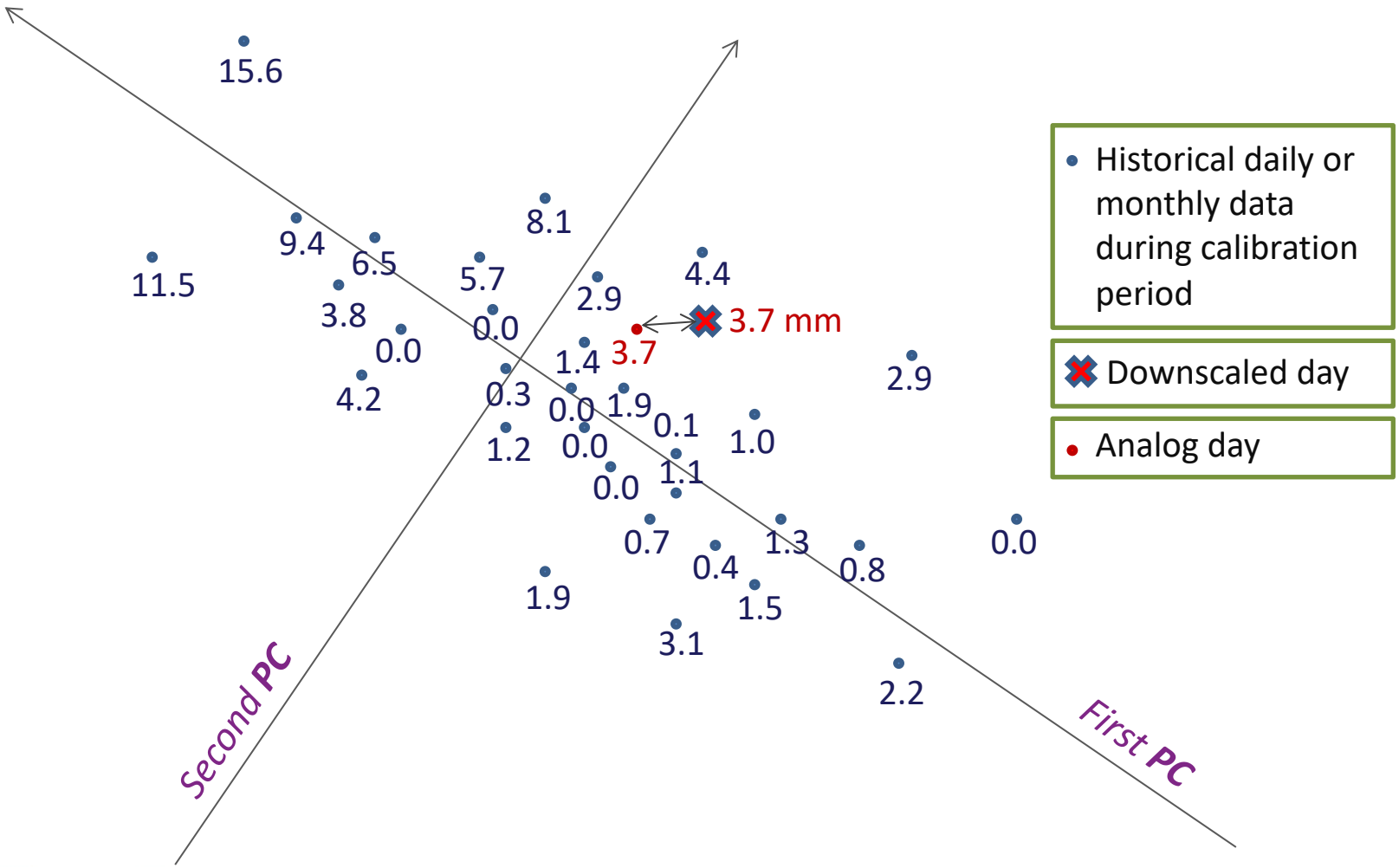
Slides Courtesy of Nicola Cortesi

Analog Day



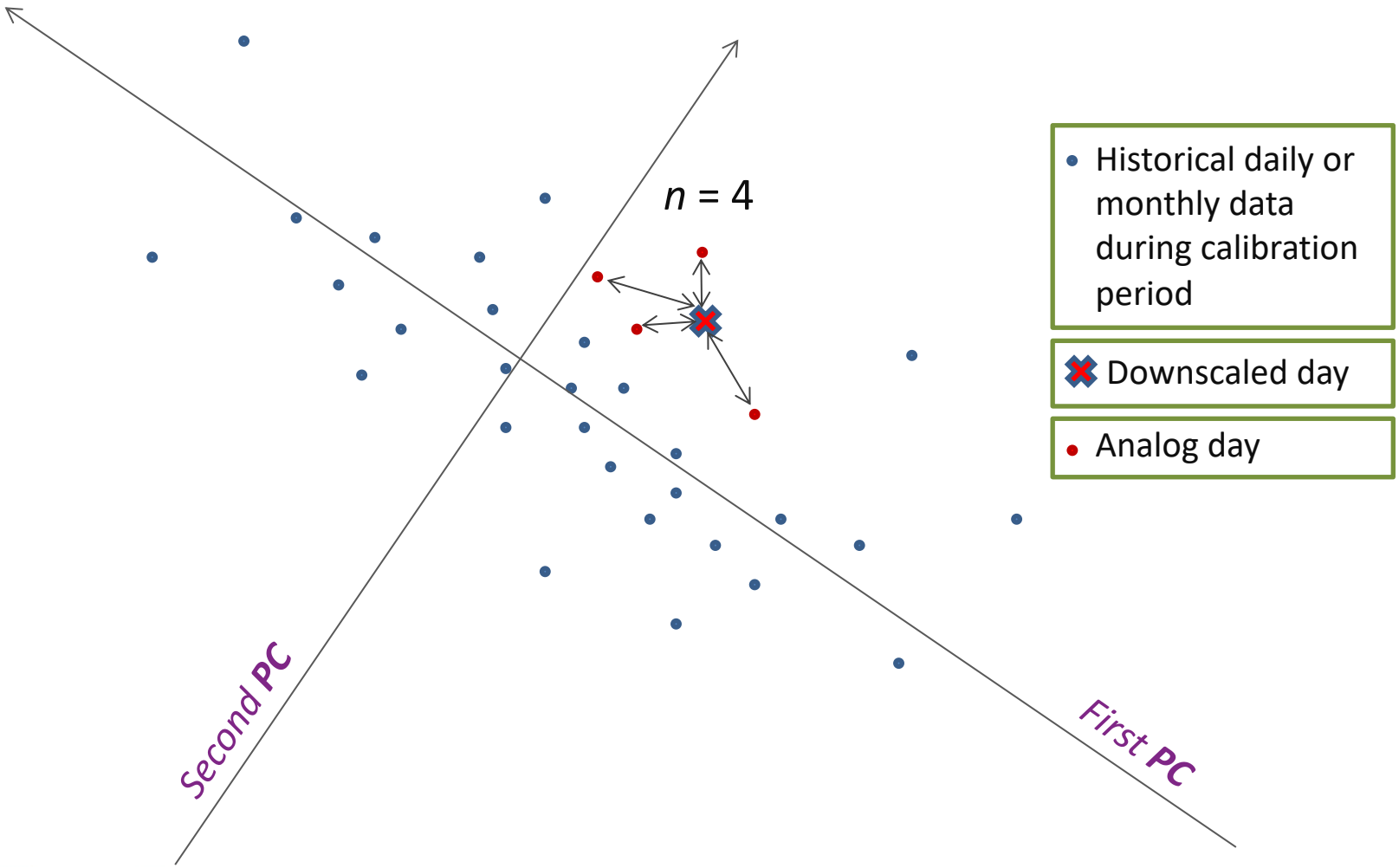
Slides Courtesy of Nicola Cortesi

Analog Day



Slides Courtesy of Nicola Cortesi

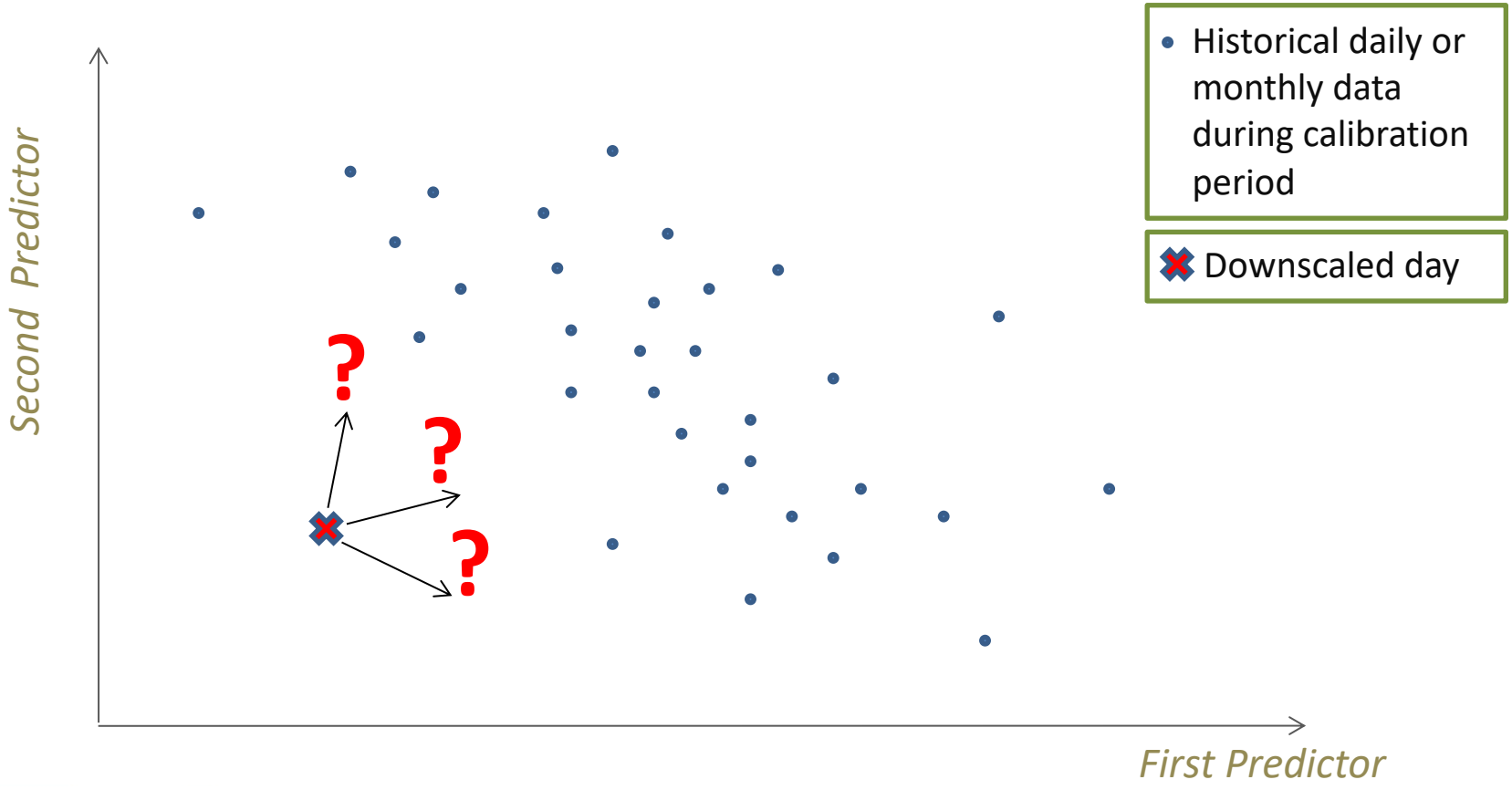
Nearest Neighbour Resampling



Slides Courtesy of Nicola Cortesi

Caveats of the AM

- It *cannot extrapolate values* outside the range of the observed data used in the calibration



Slides Courtesy of Nicola Cortesi

Caveats of the AM

- It needs *a large training sample* of observed data for calibration.

- Historical daily or monthly data during calibration period

Slides Courtesy of Nicola Cortesi

Caveats of the AM

- It doesn't preserve the *spatial autocorrelation structure* of the observed variable or its observed temporal *frequency distribution* or its *persistence* (for daily precipitation).

Recently, a number of AMs that successfully deal with these issues have been presented:

- Ribalaygua et al. (2013): a two step Analogue/Regression method
- Benestad (2009): a two step Analogue/Quantile mapping method using the Extended EOFs instead of the classic EOFs.
- Hwang and Graham (2013): a two step Bias-correction/Stochastic AM
- Matulla et al. (2007): an AM applied to sequences of 2-3 days

Suggested Readings

- *Introduction to the AM:*

Benestad et al. (2008) Empirical-Statistical Downscaling (Chapter 5). World Scientific Publishing Company.

- *Robustness Analysis:*

Gutierrez et al. (2013) Reassessing Statistical Downscaling Techniques for Their Robust Application under Climate Change Conditions. *J. Climate,* 26, 171-188. DOI: 10.1175/JCLI-D-11.00687.1

- *Similarity Measures:*

Matulla et al. (2007) Influence of similarity measures on the performance of the analog method for downscaling daily precipitation. *Clim. Dyn.* DOI: 10.1007/s00382-007-0277-2

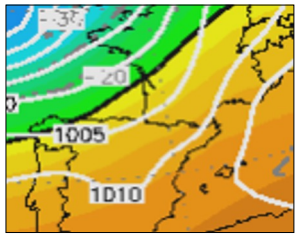
- *Two step Analogue/Regression method:*

Ribalaygua et al. (2013) Description and validation of a two step analogue/regression downscaling method. *Theor. Appl. Climatol,* 114:253-269. DOI 10.1007/s00704-013-0836-x

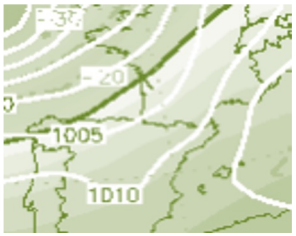
Weather Typing Method (WTM)

Composed by **2 steps**:

Reanalysis
SLP or Z Field



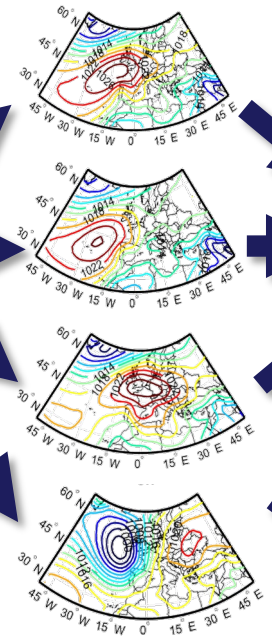
~ 1000 Km



GCM output
SLP or Z Field

Select Classification Scheme

4-43 Circulation or
Weather Types (WTs)



Select Downscaling Method

Observed Local Predictand 1950-2014

Relationships between WTs and the local predictand

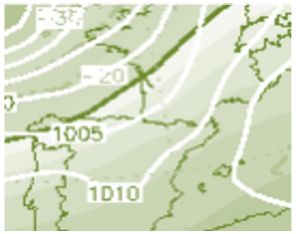
Downscaled Local Predictand

Slides Courtesy of Nicola Cortesi

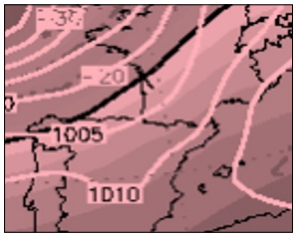
Weather Typing Method (WTM)

Composed by **2 steps**:

Reanalysis
SLP or Z Field



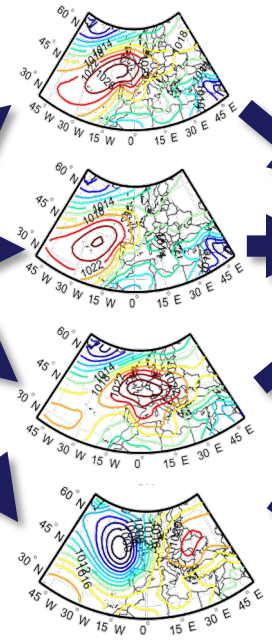
~ 1000 Km



GCM output
SLP or Z Field
2050-2099

Select Classification Scheme

4-43 Circulation or
Weather Types (WTs)



Select Downscaling Method

Observed
Local
Predictand
1950-2014

Relationships between WTs and the local predictand

Downscaled
Local
Predictand
2050-2099

Slides Courtesy of Nicola Cortesi

Advantages of the WTM

1. *Greater understanding* of the problems involved
2. Weather Types useful for downscaling *Extreme Indices*
3. *Spatially coherent*
4. *Low computational cost*
5. It can be *non-linear*
6. It can avoid *underestimation of the variance*

Statistical Downscaling Methods used with the WTM

- For Temperature: *Multiple Regression Models*
- For Daily Precipitation:
  - *Generalized Linear Models*
  - *Generalized Additive Models*
  - *Logistic (logit) Regression Model*
- *Weather Generators*
- *Analogue Method*
- *Canonical Correlation Analysis*

A brief history of the Weather Type Classifications

Hess and Brezowsky: Grosswetterlagen WTs

Spellman: Lamb classification for Spain

COST733 Catalogue: 72 WT classific. reported

1969 1972 *Lamb:* Lamb WTs

1977 *Jenkinson and Collison*: first automated Lamb classification.

1993 *Jones et al.:* assess the Lamb classification modified by Jenkinson and Collison

2000 *Trigo and DaCamara:* Lamb classification for Portugal

2007 *James:* Automated Grosswetterlagen classification

2010

Slides Courtesy of Nicola Cortesi

COST733 Catalogue

- 50 Classifications based on a *bottom-up approach*:

#	Abbreviation	Types	Parameters
PCA (PCA based methods)			
23	TPCA07	7	MSLP
24	TPCAC09	9	MSLP
25	TPCAC18	18	MSLP
26	TPCAC27	27	MSLP
27	TPCAV	6–12	MSLP
28	P27	27	Z500
29	P27C08	8	MSLP
30	P27C18	18	MSLP
31	P27C27	27	MSLP
32	PCAXTR	11–17	MSLP
33	PCAXTRC09	9–10	MSLP
34	PCAXTRC18	15–18	MSLP
LDR (methods based on leader algorithm)			
35	LUND	10	MSLP
36	LUNDC09	9	MSLP
37	LUNDC18	18	MSLP
38	LUNDC27	27	MSLP
39	ESLPC09	9	MSLP
40	ESLPC18	18	MSLP
41	ESLPC27	27	MSLP
42	EZ850C10	10	Z850
43	EZ850C20	20	Z850
44	EZ850C30	30	Z850
45	KHC09	9	MSLP
46	KHC18	18	MSLP
47	KHC27	27	MSLP

#	Abbreviation	Types	Parameters
OPT (optimization methods)			
48	CKMEANSC09	9	MSLP
49	CKMEANSC18	18	MSLP
50	CKMEANSC27	27	MSLP
51	PCACA	4–5	MSLP
52	PCACAC09	9	MSLP
53	PCACAC18	18	MSLP
54	PCACAC27	27	MSLP
55	PETISCO	25–38	MSLP, Z500
56	PETISCOC09	9	MSLP
57	PETISCOC18	18	MSLP
58	PETISCOC27	27	MSLP
59	PCAXTRKM	11–17	MSLP
60	PCAXTRKMC09	9–10	MSLP
61	PCAXTRKMC18	15–18	MSLP
62	SANDRA	18–23	MSLP
63	SANDRAC09	9	MSLP
64	SANDRAC18	18	MSLP
65	SANDRAC27	27	MSLP
66	SANDRAS	30	Z925, Z500
67	SANDRASC09	9	MSLP
68	SANDRASC18	18	MSLP
69	SANDRASC27	27	MSLP
70	NNW	9–30	Z500
71	NNWC09	9	MSLP
72	NNWC18	18	MSLP
73	NNWC27	27	MSLP

Table from *Philipp et al.* (2010)

COST733 Catalogue

- 22 Classifications based on a *top-down approach*:

#	Abbreviation	Types	Parameters
SUB (subjective methods)			
1	HBGWL	29	not specified
2	HBGWT	10	not specified
3	OGWL	29	MSLP, Z500
4	OGWLSLP	29	MSLP
5	PECZELY	13	not specified
6	PERRET	40	not specified
7	ZAMG	43	not specified
THR (threshold based methods)			
8	GWT	18	MSLP
9	GWTC10	10	MSLP
10	GWTC18	18	MSLP
11	GWTC26	26	MSLP
12	LITADVE	9	MSLP
13	LITTC	27	MSLP
14	LITTC18	18	MSLP
15	LWT2	26	MSLP
16	LWT2C10	10	MSLP
17	LWT2C18	18	MSLP
18	WLKC09	9	U/V700
19	WLKC18	18	U/V700, Z925
20	WLKC28	28	U/V700, Z925/500
21	WLKC733	40	U/V700, Z925/500, PW
22	SCHUEPP	40	MSLP, Z500, U/VSFC/500

Grosswetterlagen WTs

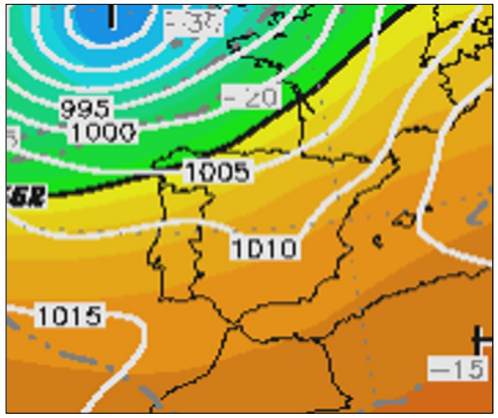
Modified Lamb WTs

Table from *Philipp et al.* (2010)

Slides Courtesy of Nicola Cortesi

Lamb classification modified by Jenkinson and Collison

Mean Daily SLP Field



Southerly Flow: $\mathbf{S} = 1.305\,[0.25\,(p_5 + 2p_9 + p_{13}) - 0.25\,(p_4 + 2p_8 + p_{12})]$

Westerly Flow: $\mathbf{W} = [0.5\,(p_{12} + p_{13}) - 0.5\,(p_4 + p_5)]$

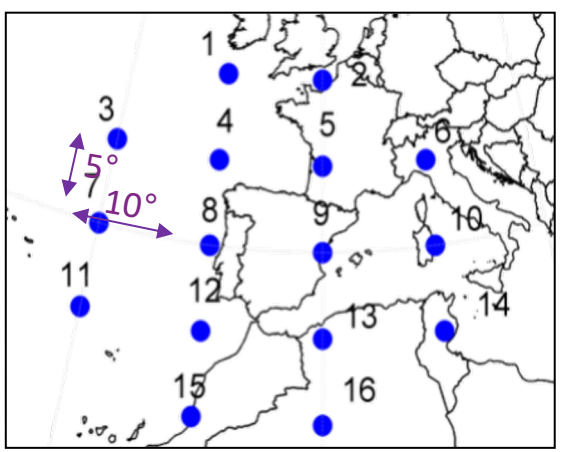
Total Flow: $\mathbf{F} = (\mathbf{S}^2 + \mathbf{W}^2)^{1/2}$ **Flow Direction:** $\mathbf{D} = \mathrm{arctg}\,(\mathbf{W}/\mathbf{S})$

Southerly Vorticity: $\mathbf{ZS} = 0.85\,[0.25\,(p_6 + 2p_{10} + p_{14}) - 0.25\,(p_5 + 2p_9 + p_{13}) - 0.25\,(p_4 + 2p_8 + p_{12}) + 0.25\,(p_3 + 2p_7 + p_{11})]$

Westerly Vorticity: $\mathbf{ZW} = 1.12\,[0.5\,(p_{15} + p_{16}) - 0.5\,(p_8 + p_9)] - 0.91\,[0.5\,(p_8 + p_9) - 0.5\,(p_1 + p_2)]$

Total Vorticity: $\mathbf{Z} = \mathbf{ZS} + \mathbf{ZW}$ (>0: Cyclonic ; <0: Anticyclonic)

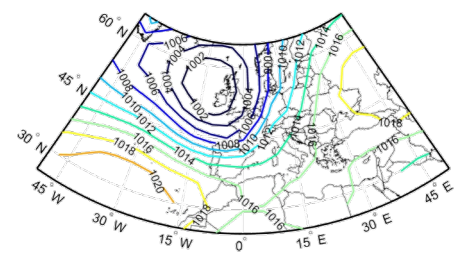
Selection of 16 grid points $p_1, \ldots, p_{16}$:



Extraction of 16 SLP time series

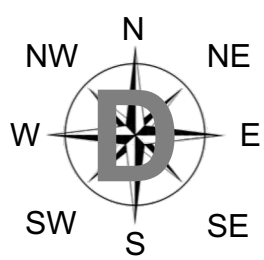
Calculation of Geostrophical Indexes

*Ex: **Westerly** WT:* $|\mathbf{Z}| < \mathbf{F}$, $248° < \mathbf{D} < 292°$



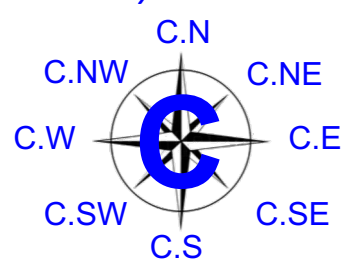
Selection of # WTs:

8 Directionals



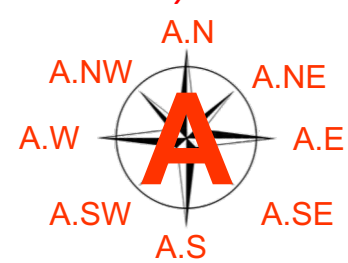
$|\mathbf{Z}| < \mathbf{F}$

9 Cyclonic



$|\mathbf{Z}| > \mathbf{F}, \mathbf{Z} > 0$

9 Anticyclonic

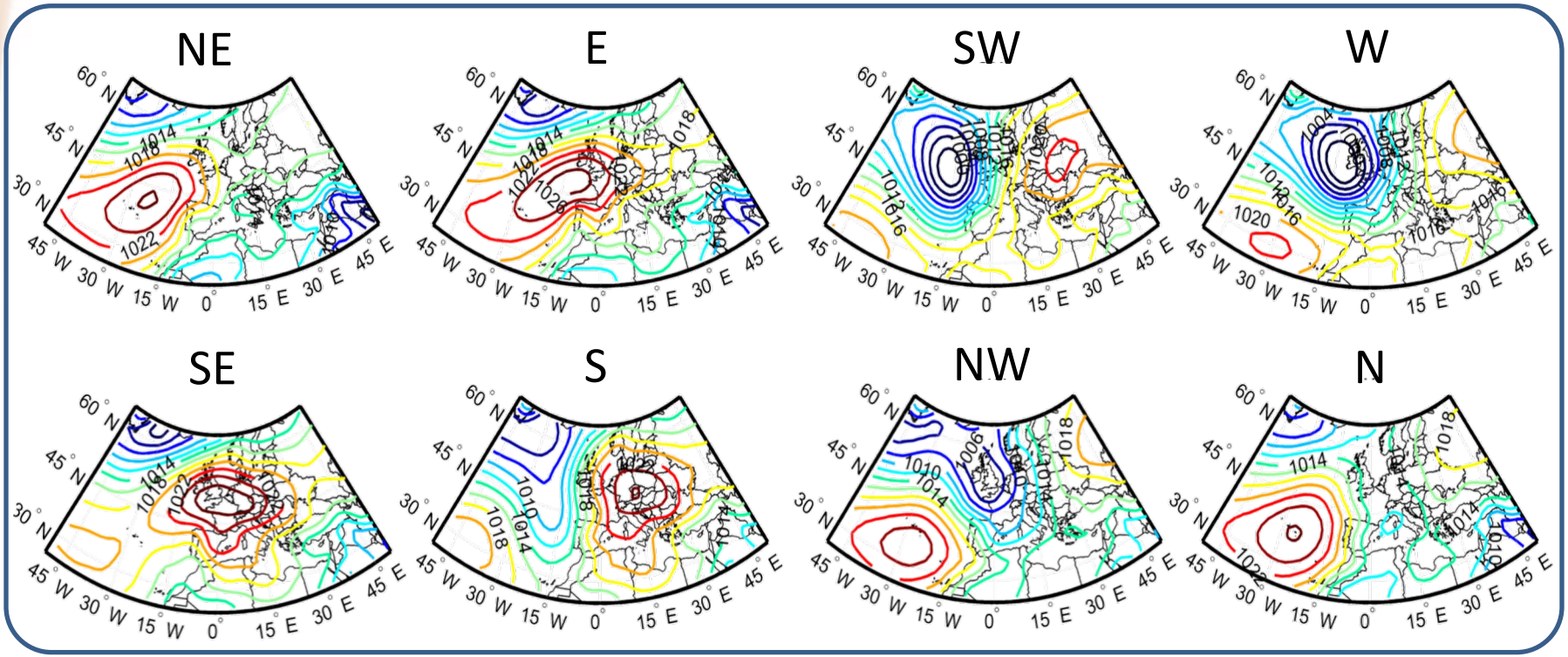


$|\mathbf{Z}| > \mathbf{F}, \mathbf{Z} < 0$

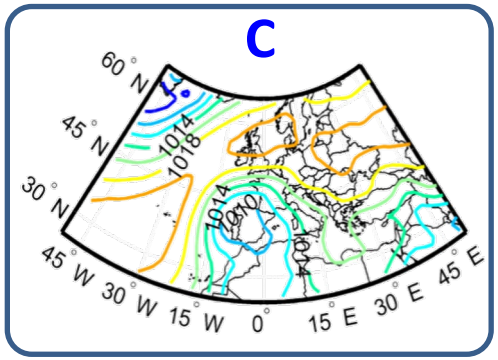
WTs Classification

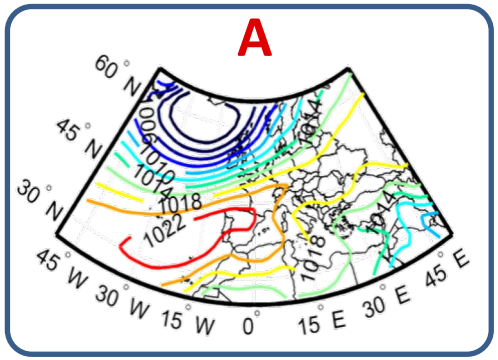
Slides Courtesy of Nicola Cortesi

Lamb classification modified by Jenkinson and Collison



Mean SLP fields





20th Century Reanalysis

Period: 1948-2003

Slides Courtesy of Nicola Cortesi

Caveats of the WTM

- Limit on the maximum number of Weather Types
  - Each WT must have a significant number of days of occurrence

- GCM pressure or geopotential predictors
  - They do not react sufficiently to climate change

- WTMs often underestimate the temporal climate variability, especially the linear models.

Caveats of both the AM and the WTM

- *Stationarity Hypothesis*
  - Relationships predictors/predictand are assumed to be constant in time

- *Convective structures* are still not well simulated
  - Limitations of GCMs: relies on parametrizations

- *Large-Scale Domain*
  - Must optimize the domain
  - Different predictors can also have different optimized domains

Suggested Readings for the WTM:

- *Weather Types Catalogue:*

Philipp, A., Bartholy, J., Beck, C., Erpicum, M., et al. (2010) *Cost733cat—a database of weather and circulation type classifications.* Phys. Chem. Earth 35, 360–373. Doi:10.1016/j.pce.2009.12.010

- *Lamb Classification:*

Jones, P. D., Harpham, C., and Briffa, K. R. (2012) *Lamb weather types derived from reanalysis products.* International Journal of Climatology. Doi: 10.1002/joc.3498

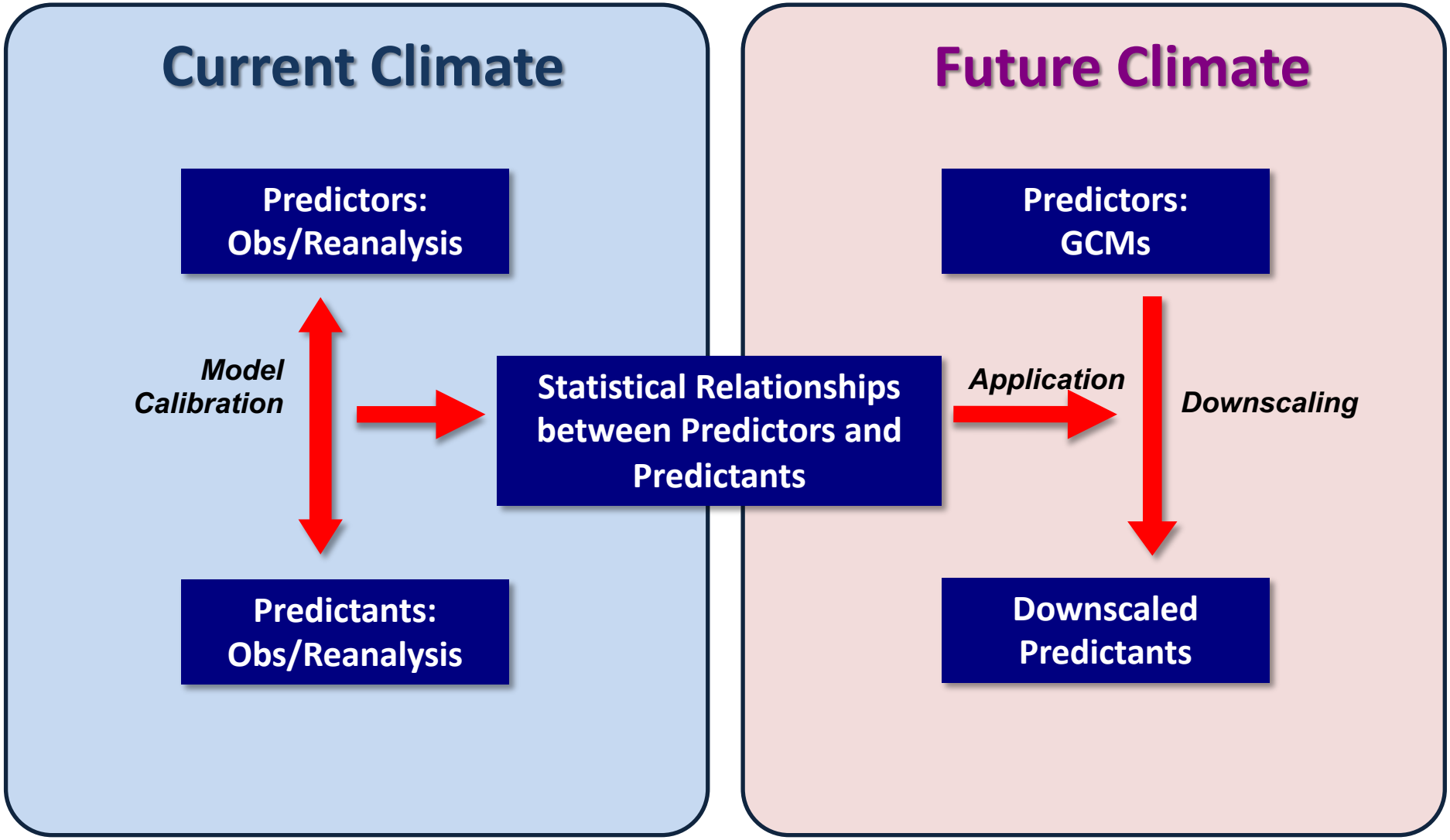
- *SANDRA Classification:*

Lutz, K., Jacobeit, J., Philipp, A., Seubert, S., Kunstmann, H. and Laux, P. (2012) *Comparison and evaluation of statistical downscaling techniques for station-based precipitation in the Middle East.* Int. J. Climatol., 32: 1579–1595. doi: 10.1002/joc.2381

- *A Classic:*

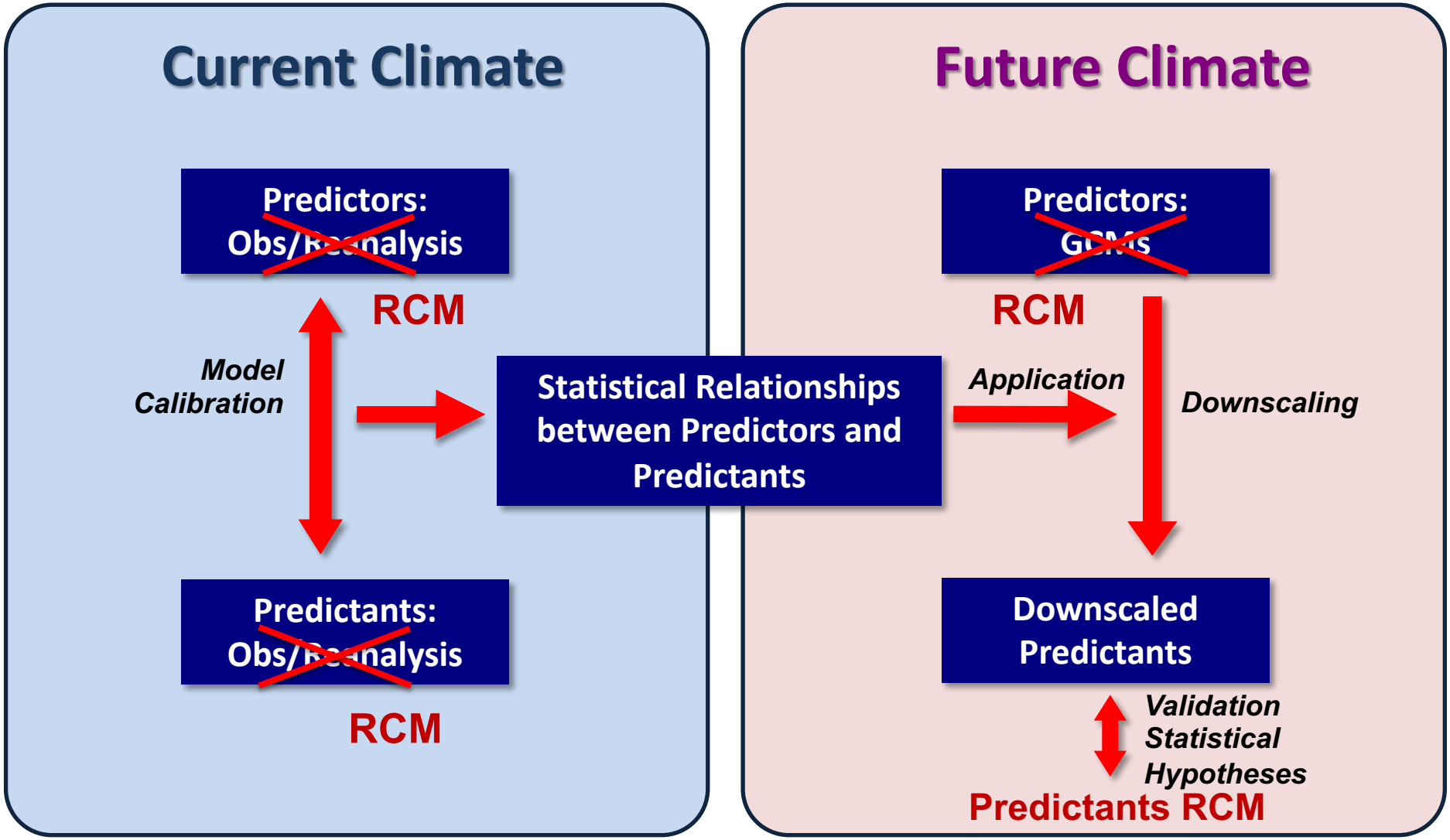
Goodess, C.M. and Palutikof, J.P. (1998) *Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling.* Int.J.Climatol. 18,1051-1083.

Statistical Downscaling: Recap



IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Perfect Model Approach : validate the transferability



IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Limitations and Advantages

Statistical Downscaling

Limits

- Uncertainties of the large-scale circulation
- Observations quality
- Limited by the number of observation locations
- Temperature tendencies are reproduced
- Most of methods are daily based
- New Extremes may not be reproduced
- Stationarity hypothesis

Advantages

- Covariance of variables are kept
- Low cost to evaluate uncertainties
- Can also produce hourly data
- High or very high spatial resolution

Dynamical Downscaling/Bias Correction

Limitations

- Bias correction is needed
  - Loss of covariances of variables (not always totally true)
  - Limited by the number of observation locations
  - Problems with some highly non-gaussian fields like precipitation
- Very high computational costs: limits uncertainties assessment and horizontal resolution

Advantages

- More appropriate for extremes (but bias correction is not)
- Historical Cases simulation: days, events
- Stationarity hypothesis is less strong

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Conclusions

- **Summary of Statistical Downscaling and Bias Correction**
  - Very fast to process, but sometimes complex to setup properly
    - Uncertainties Assessment
    - Multi-scenarios
  - Multiple configurations
  - Quality and performance dependent on the quality of the observations and reanalysis
  - Can only provide values at the observation locations
  - Need long time series of high-quality observations (25 years+)
  - Bias correction is always required for impact modeling and studies at local and regional scale
  - There exists a very large number of methods: not always easy to choose the best one(s)

Christian Pagé
christian.page@cerfacs.fr

Questions Time!

IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084