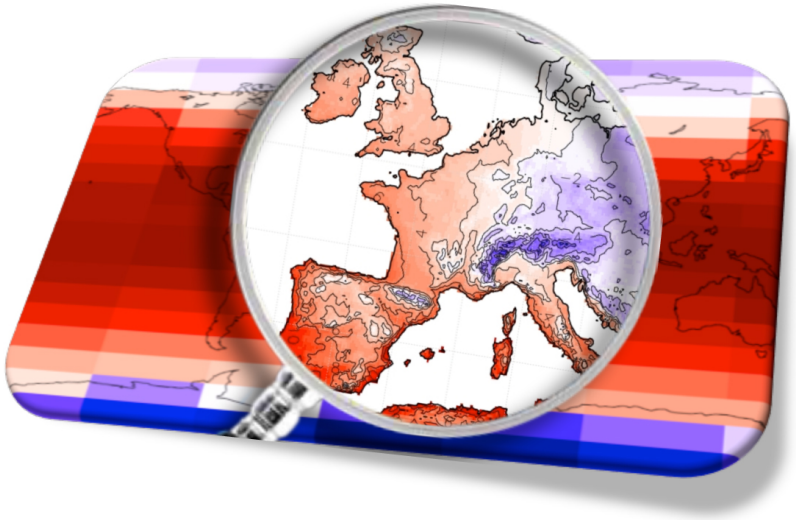
LSCE
LABORATOIRE DES SCIENCES DU CLIMAT & DE L'ENVIRONNEMENT

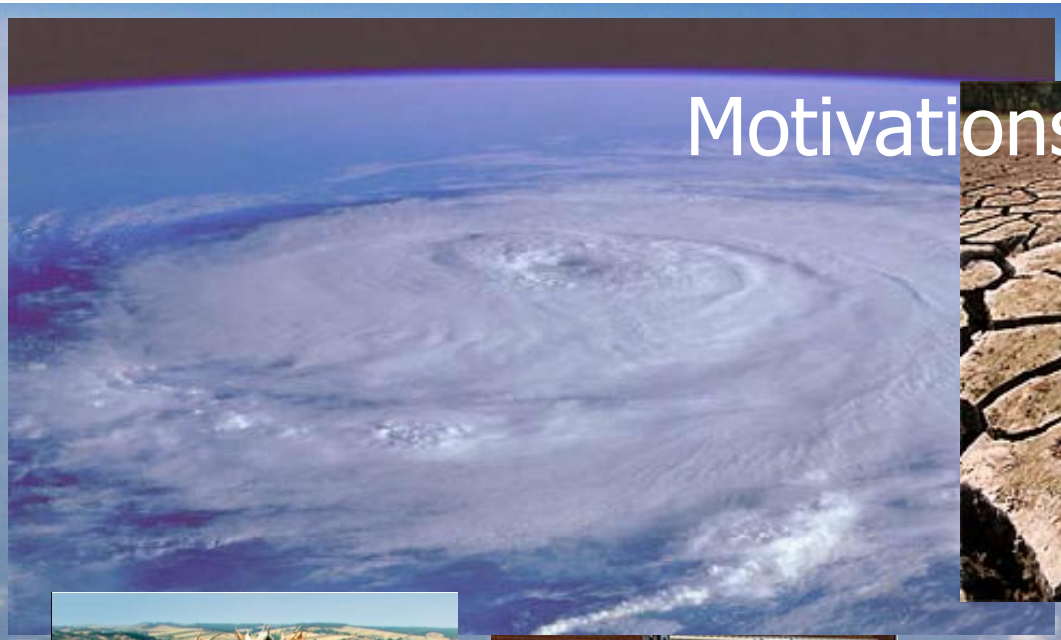
Statistical Downscaling: The current "Why?" and "How?" & Future steps…



Mathieu Vrac

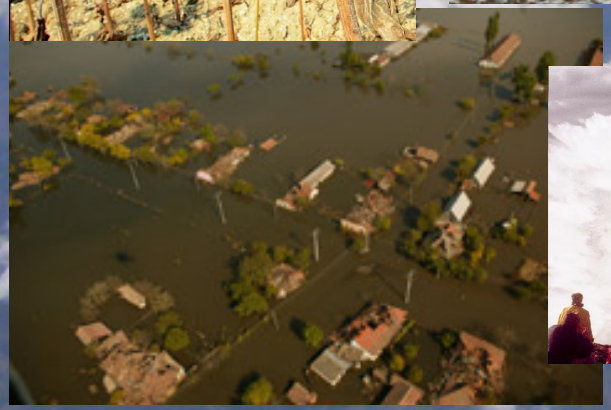
Statistical downscaling of climate scenarios for the impact communities: A CMIP5 perspective
Workshop IS-ENES, October 16-17, 2012, Paris

Motivations





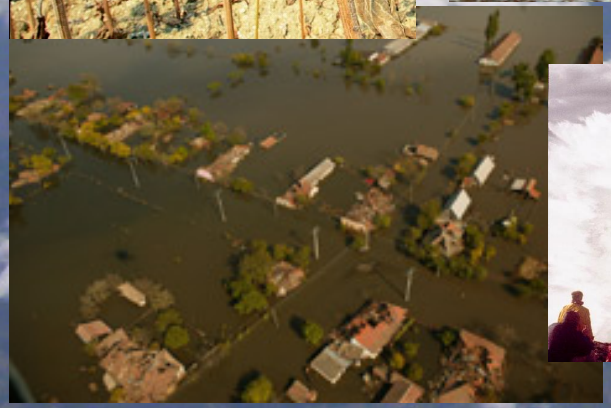



Motivations

- 30% of the world economic activities are affected by meteorological conditions (source: IPCC)





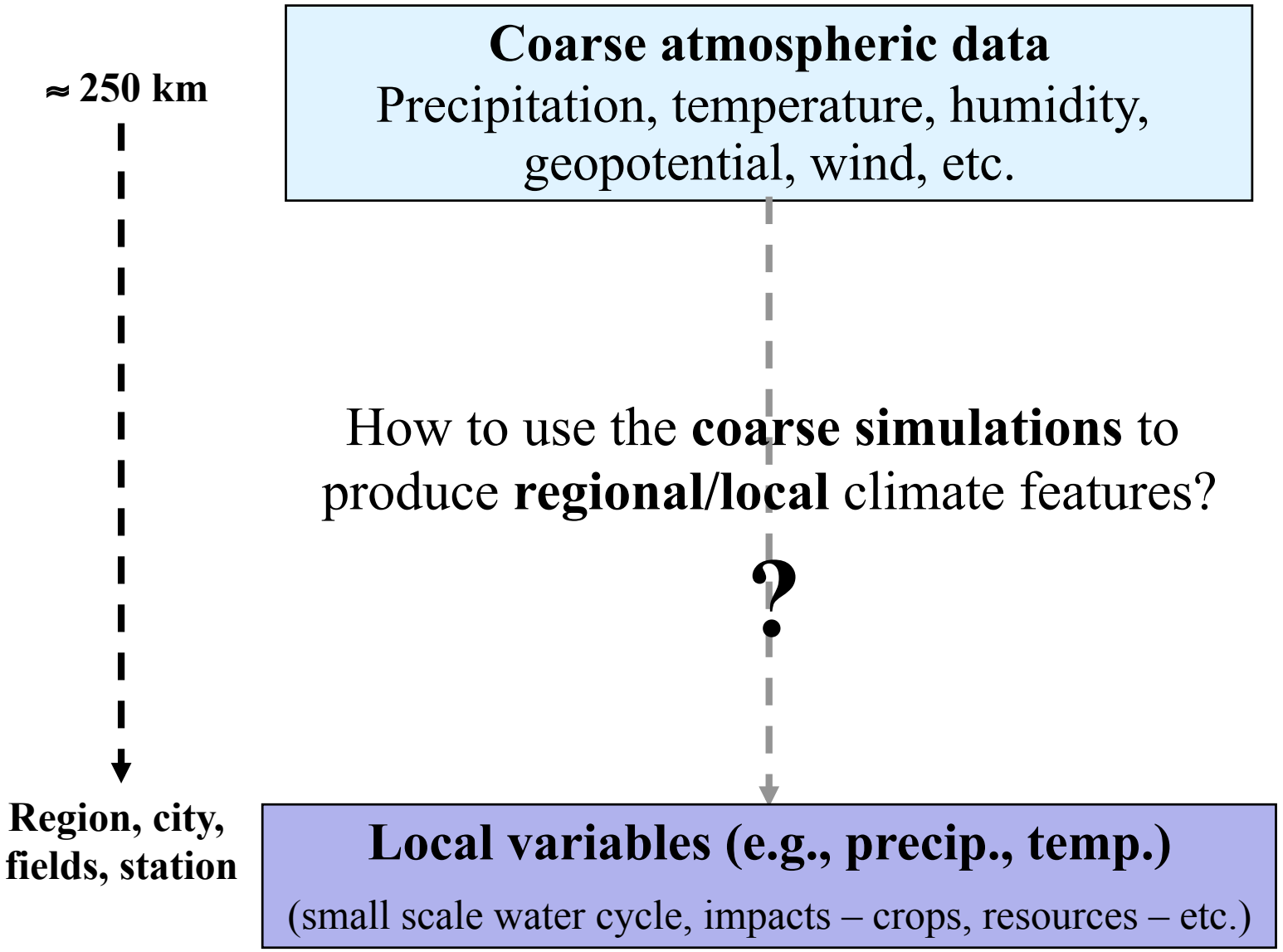
Motivations

- 30% of the world economic activities are affected by meteorological conditions (source: IPCC)
- **IPCC** scenarios of climate change have a **coarse spatial resolution !!**
  Not adapted to social and economic scales of impact studies
  - Social and economic impacts: water resources, hydrology, agriculture, air pollution, human health, etc.
  - How will climate change interact with existing environmental features at a regional/local scale ?

Motivations

- 30% of the world economic activities are affected by meteorological conditions (source: IPCC)
- **IPCC** scenarios of climate change have a **coarse spatial resolution !!** Not adapted to social and economic scales of impact studies
  - Social and economic impacts: water resources, hydrology, agriculture, air pollution, human health, etc.
  - How will climate change interact with existing environmental features at a regional/local scale ?
- **Downscaling**: To derive sub-grid scale (regional or local) weather or climate using General Circulation Models (GCMs) outputs or reanalysis data (e.g. NCEP)

How to downscale?: The basics



How to downscale?: The basics

≈ 250 km

Coarse atmospheric data
Precipitation, temperature, humidity, geopotential, wind, etc.

Dynamical downscaling (RCMs):

- GCMs to drive regional models (5-50km) determining atmosphere dynamics
- Requires a lot of computer time and resources => Limited applications

Region, city, fields, station

Local variables (e.g., precip., temp.)
(small scale water cycle, impacts – crops, resources – etc.)

How to downscale?: The basics

≈ 250 km

Coarse atmospheric data
Precipitation, temperature, humidity, geopotential, wind, etc.

Dynamical downscaling (RCMs):

- GCMs to drive regional models (5-50km) determining atmosphere dynamics
- Requires a lot of computer time and resources => Limited applications

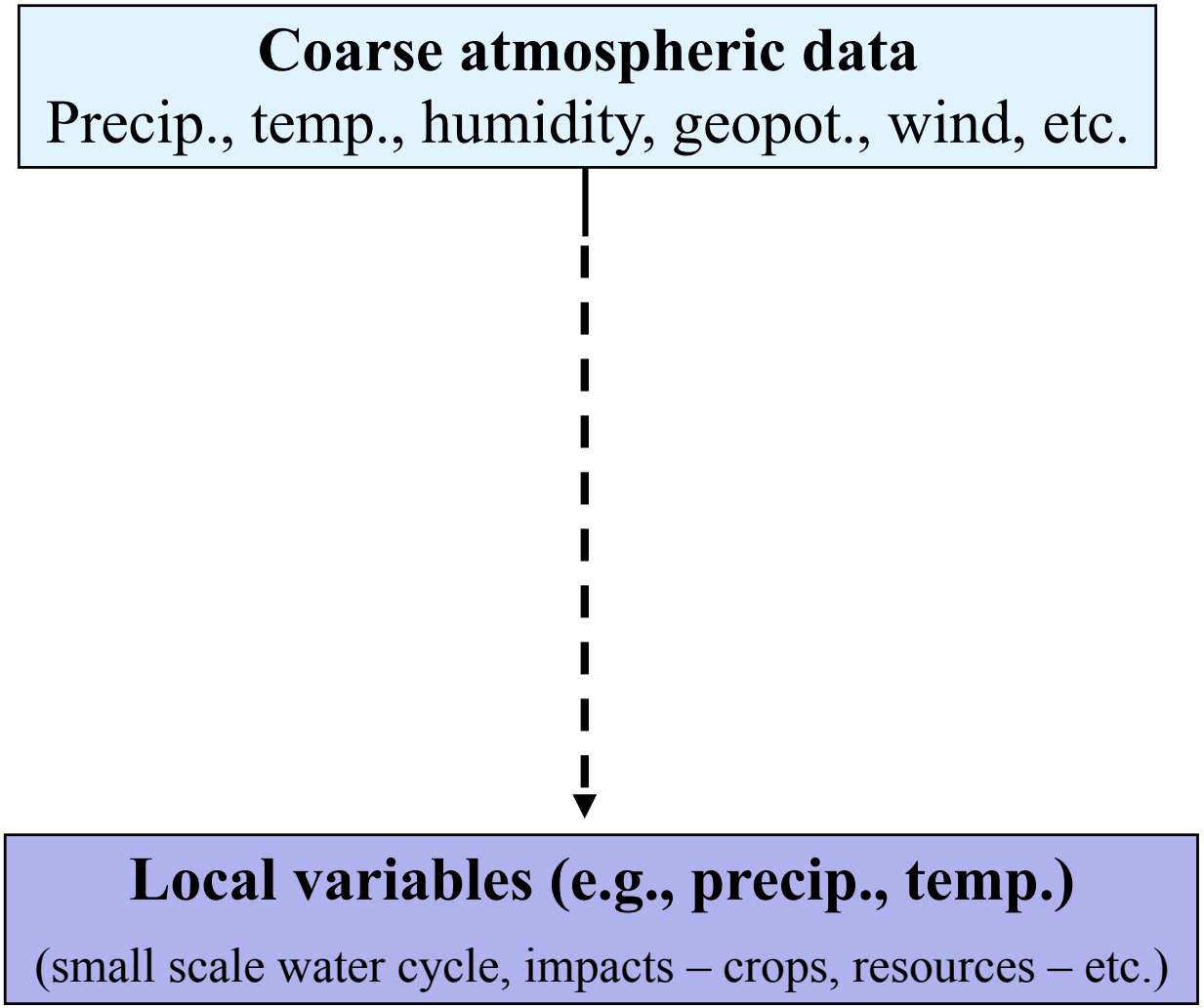
Statistical downscaling:

- Based on statistical relationships between large- and local-scale variables
- Low costs and rapid simulations applicable to any spatial resolution
- Uncertainties (results, propagation, etc)

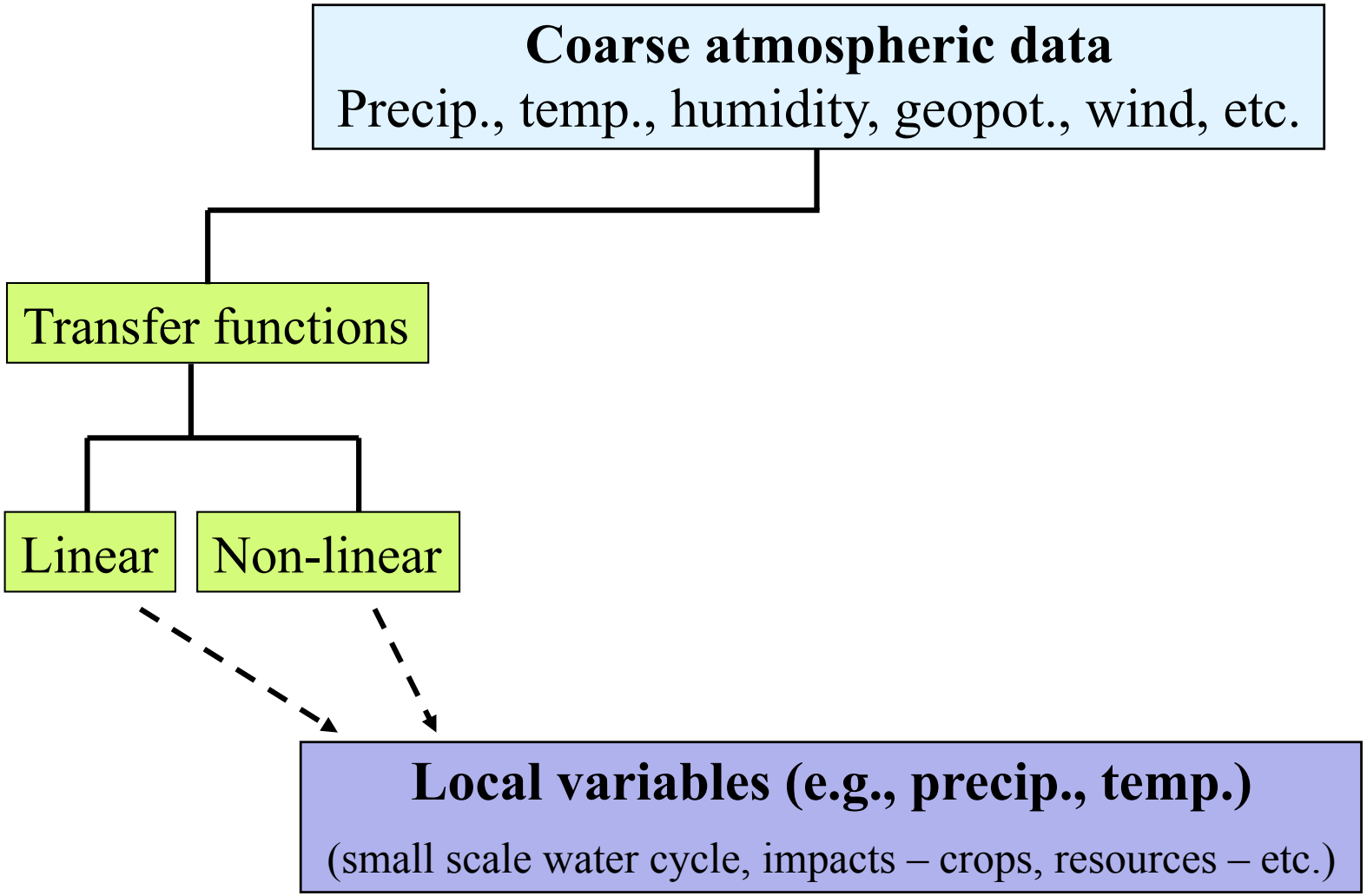
Region, city, fields, station

Local variables (e.g., precip., temp.)
(small scale water cycle, impacts – crops, resources – etc.)

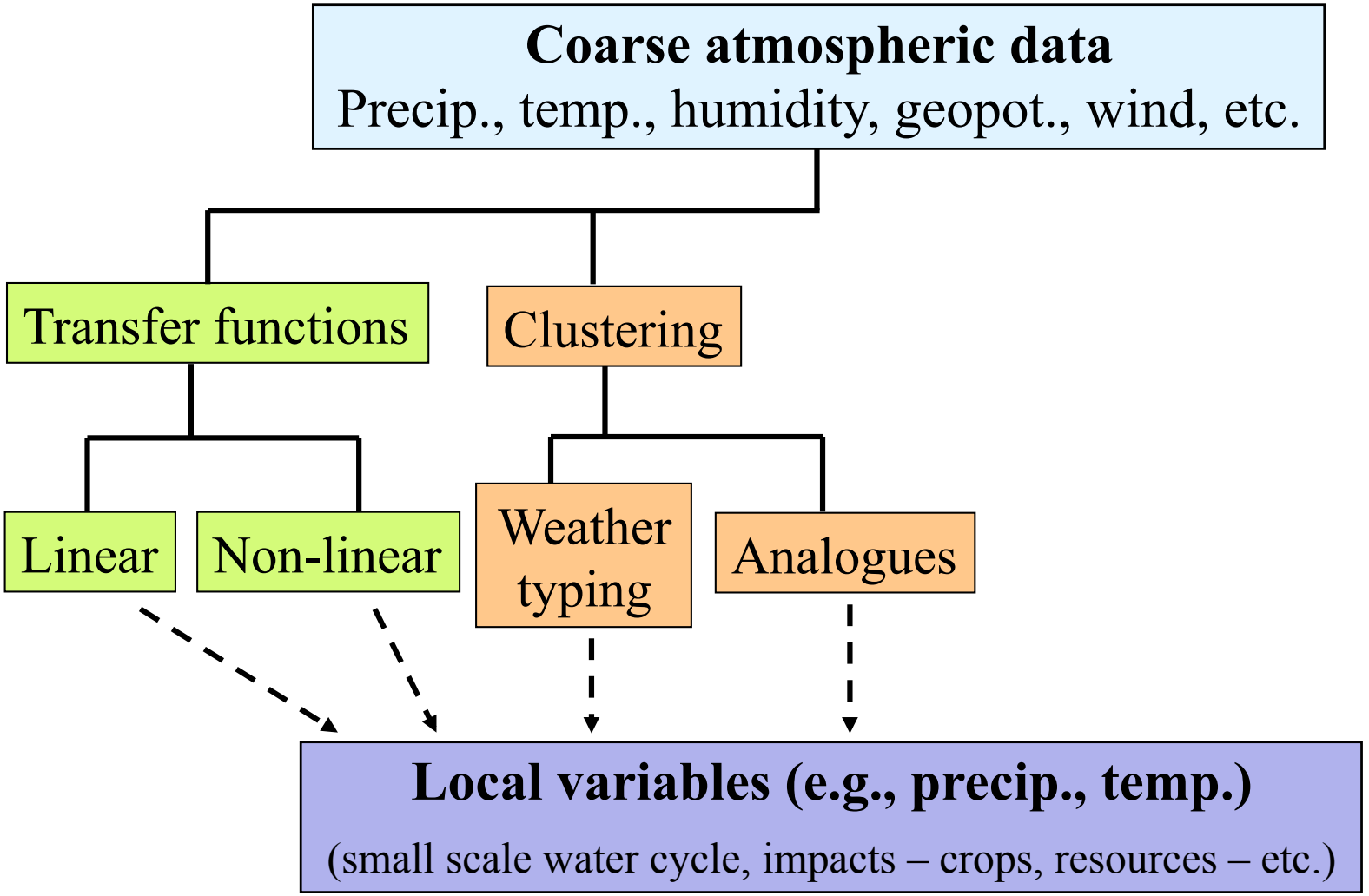
Main statistical approaches



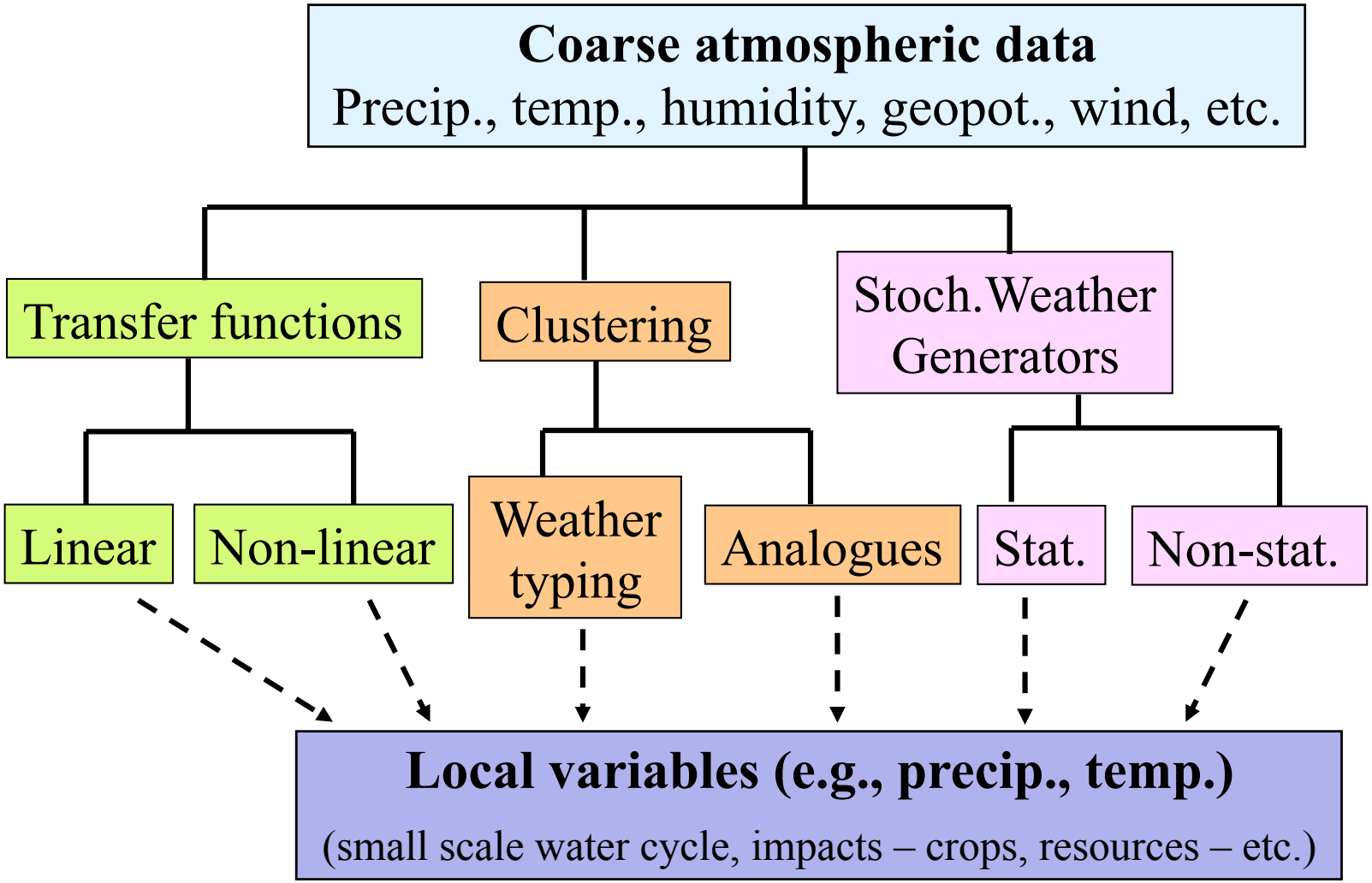
Main statistical approaches



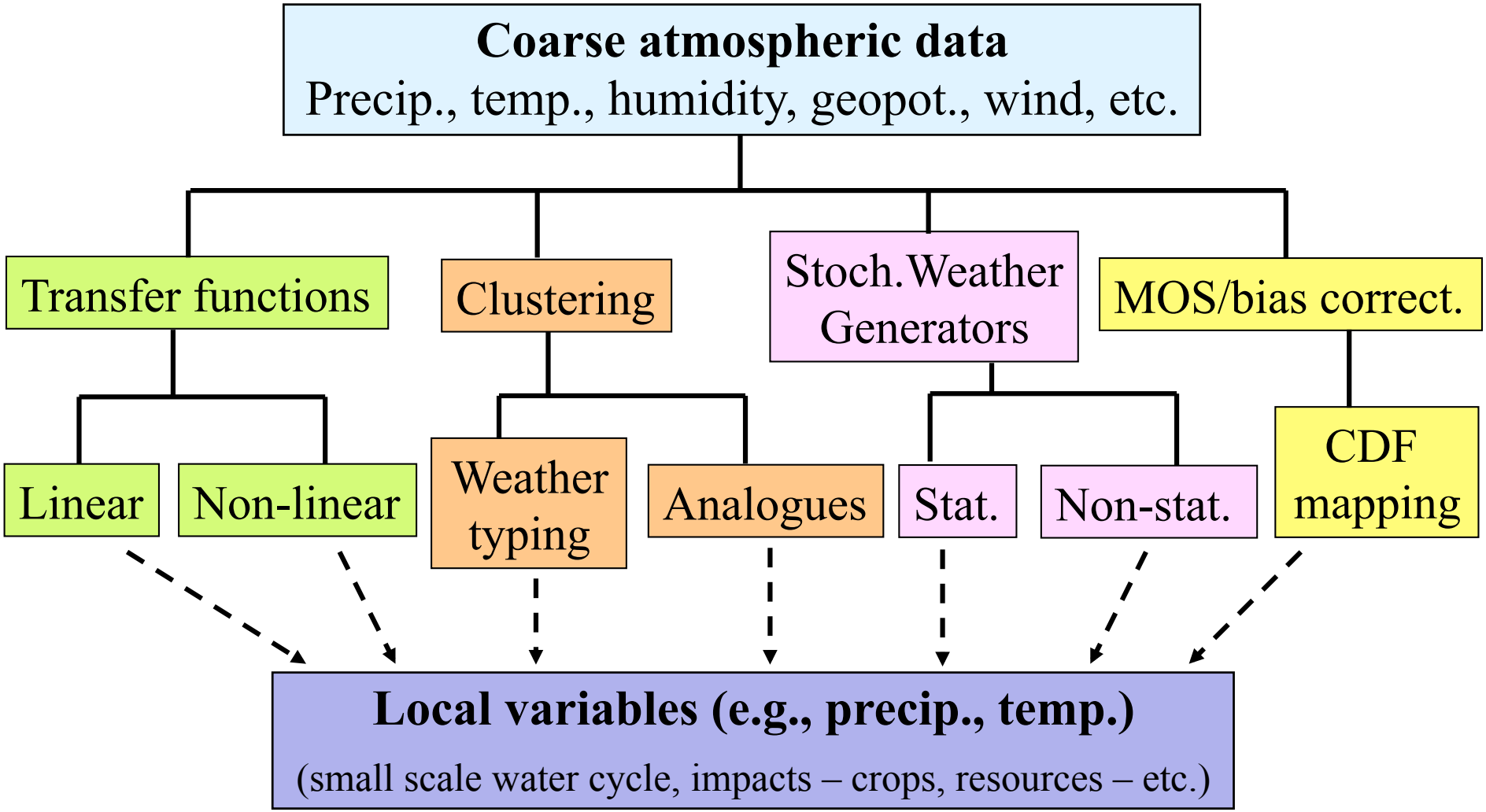
Main statistical approaches



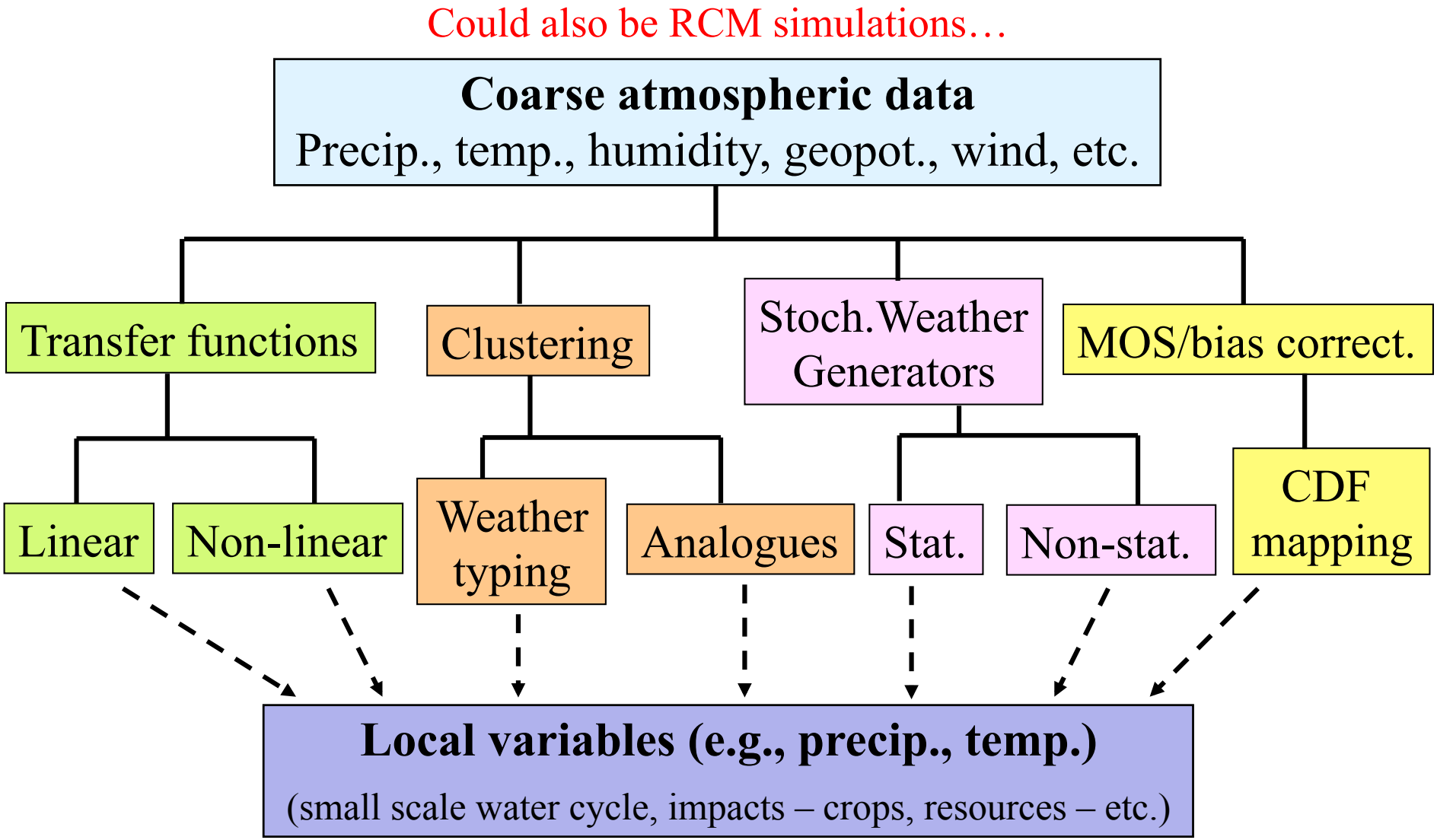
Main statistical approaches



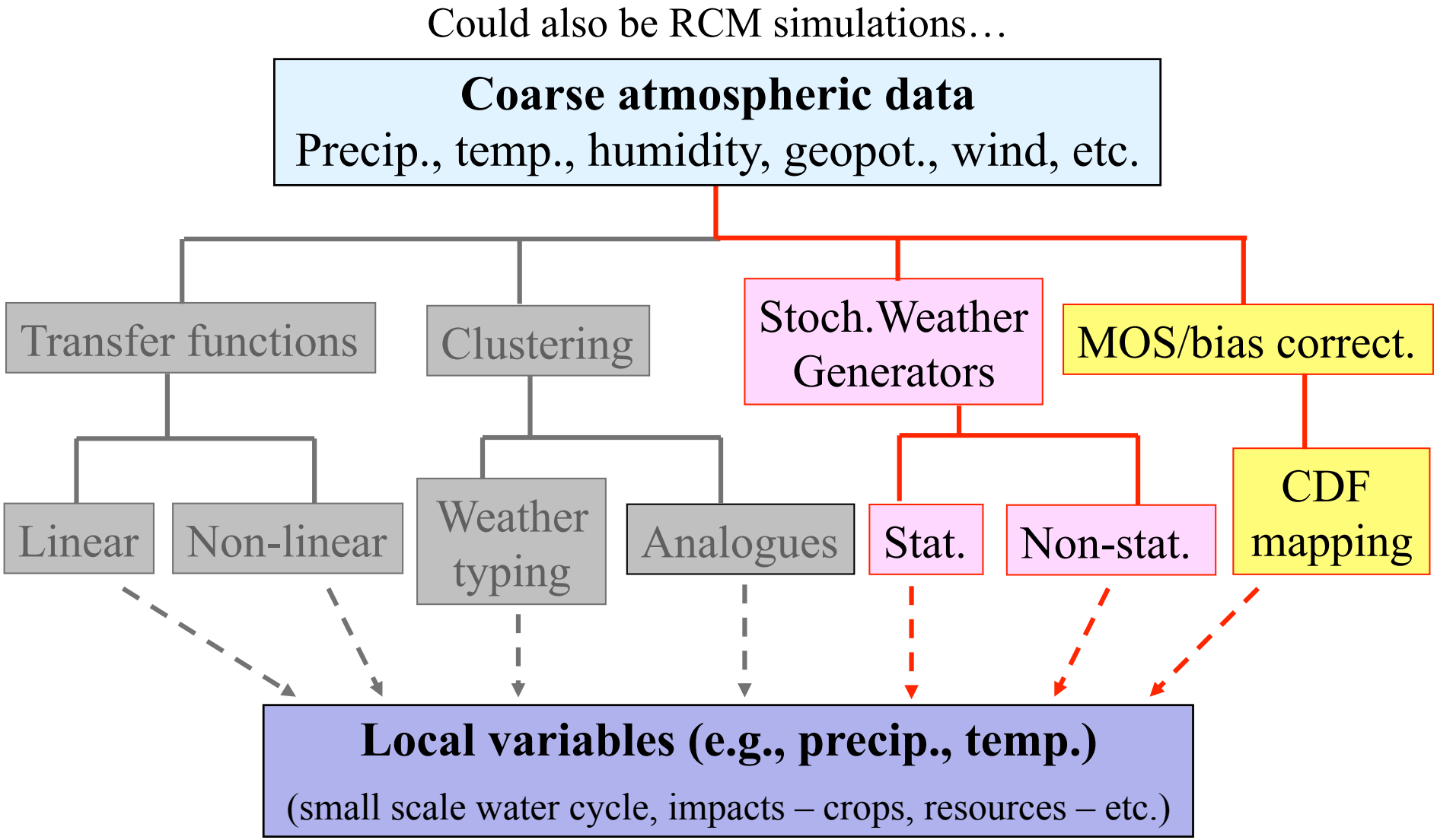
Main statistical approaches



Main statistical approaches



Main statistical approaches



Stochastic Weather Generators (WGs)

Principle: *A WG is a stochastic model simulating daily weather statistically similar to observations, based on parameters determined by historical records (Wilks and Wilby, 1999).*

Stochastic Weather Generators (WGs)

Principle: *A WG is a stochastic model simulating daily weather statistically similar to observations, based on parameters determined by historical records (Wilks and Wilby, 1999).*

Stochastic:

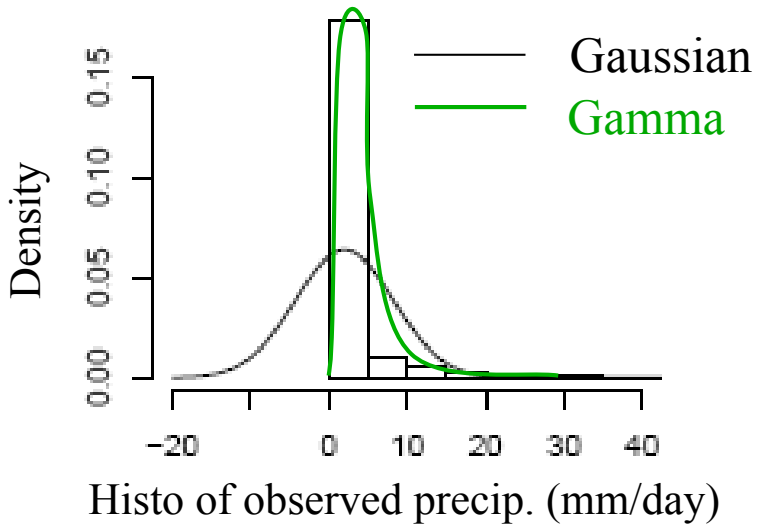
- The rainfall occurrence of *today* is conditional on the one of *yesterday* => Historical key-tool = **Markov Chains**

Stochastic Weather Generators (WGs)

Principle: *A WG is a stochastic model simulating daily weather statistically similar to observations, based on parameters determined by historical records (Wilks and Wilby, 1999).*

Stochastic:

- The rainfall occurrence of *today* is conditional on the one of *yesterday* => Historical key-tool = **Markov Chains**
- **Simulations** are performed according to **pdfs**

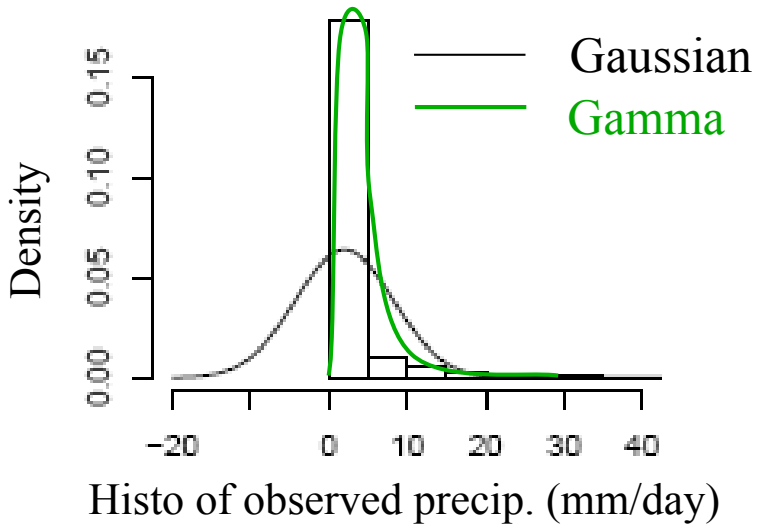


Stochastic Weather Generators (WGs)

Principle: *A WG is a stochastic model simulating daily weather statistically similar to observations, based on parameters determined by historical records (Wilks and Wilby, 1999).*

Stochastic:

- The rainfall occurrence of *today* is conditional on the one of *yesterday* => Historical key-tool = **Markov Chains**
- **Simulations** are performed according to **pdfs**



For rain intensity, most of the WGs **simulate** values in (0,+∞) according to a **Gamma distribution** (here in green)

"Non-homogeneous" WGs

- Recently: WGs for downscaling: **large-scale info is included**
  - Pryor et al. (2006) for wind: Weibull param. = GLM(GCM features)
  - Furrer & Katz (2007) for prec: Gamma param. = GLM(GCM data)

"Non-homogeneous" WGs

- Recently: WGs for downscaling: **large-scale info is included**
  - Pryor et al. (2006) for wind: Weibull param. = GLM(GCM features)
  - Furrer & Katz (2007) for prec: Gamma param. = GLM(GCM data)

⇒ Non-stationary model (Vrac et al., 2007; Carreau & Vrac, 2011):

$$Po_t = P(O_t \,|\, O_{t-1}, X_t)$$

Large-scale info
(e.g., simulations, WR, statistics)

"Non-homogeneous" WGs

- Recently: WGs for downscaling: **large-scale info is included**
  - Pryor et al. (2006) for wind: Weibull param. = GLM(GCM features)
  - Furrer & Katz (2007) for prec: Gamma param. = GLM(GCM data)

⇒ Non-stationary model (Vrac et al., 2007; Carreau & Vrac, 2011):

$$Po_t = P(O_t \,|\, O_{t-1}, X_t)$$

Large-scale info
(e.g., simulations, WR, statistics)

& PDF of intensity with parameters cond'l on (=function of) large-scale data X_t

$$f\left(. \,|\, \alpha(X_t)\right)$$

"Non-homogeneous" WGs

- Recently: WGs for downscaling: **large-scale info is included**
  - Pryor et al. (2006) for wind: Weibull param. = GLM(GCM features)
  - Furrer & Katz (2007) for prec: Gamma param. = GLM(GCM data)

⇒ Non-stationary model (Vrac et al., 2007; Carreau & Vrac, 2011):

$$Po_t = P(O_t \mid O_{t-1}, X_t)$$

Large-scale info
(e.g., simulations, WR, statistics)

& PDF of intensity with parameters cond'l on (=function of) large-scale data X_t

$$f\left(. \mid \alpha\left(X_t\right)\right)$$

Take-home story about Stochastic WGs:
Local-scale data are **simulated from conditional pdf**

⇒ If X evolves with time => $f(.|\alpha(X))$ evolves too
⇒ Uncertainty assessment (Semenov, 2007)

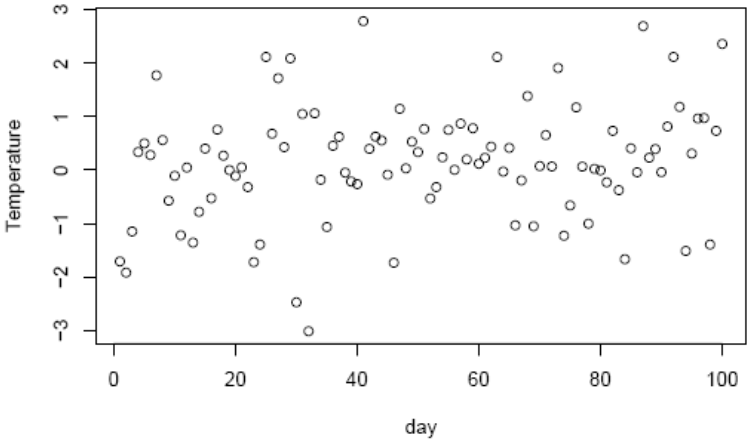
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

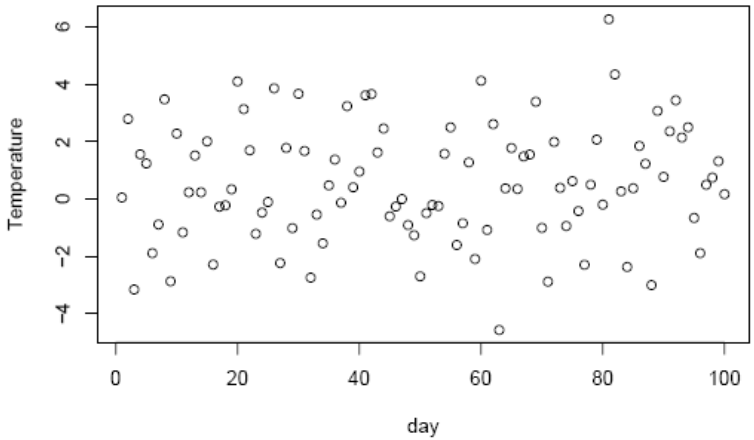
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

Model temperature

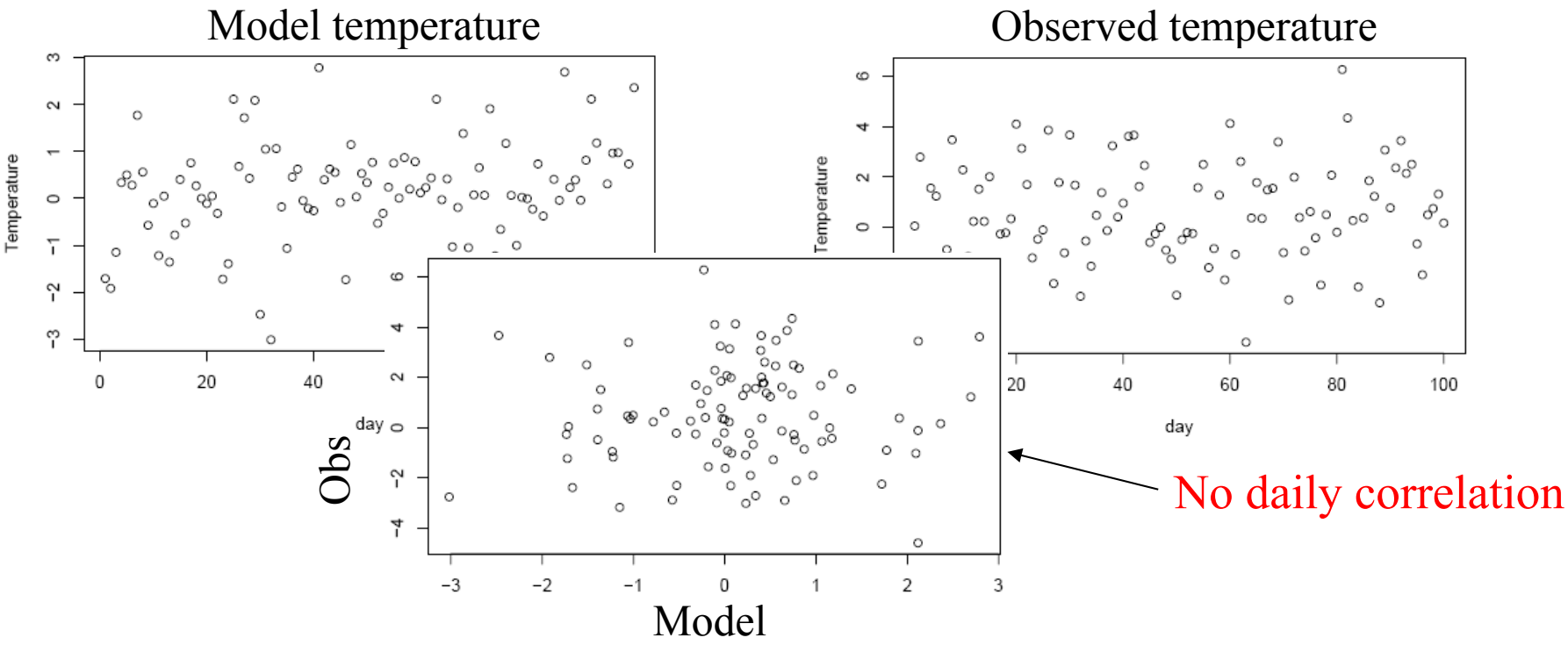


Observed temperature



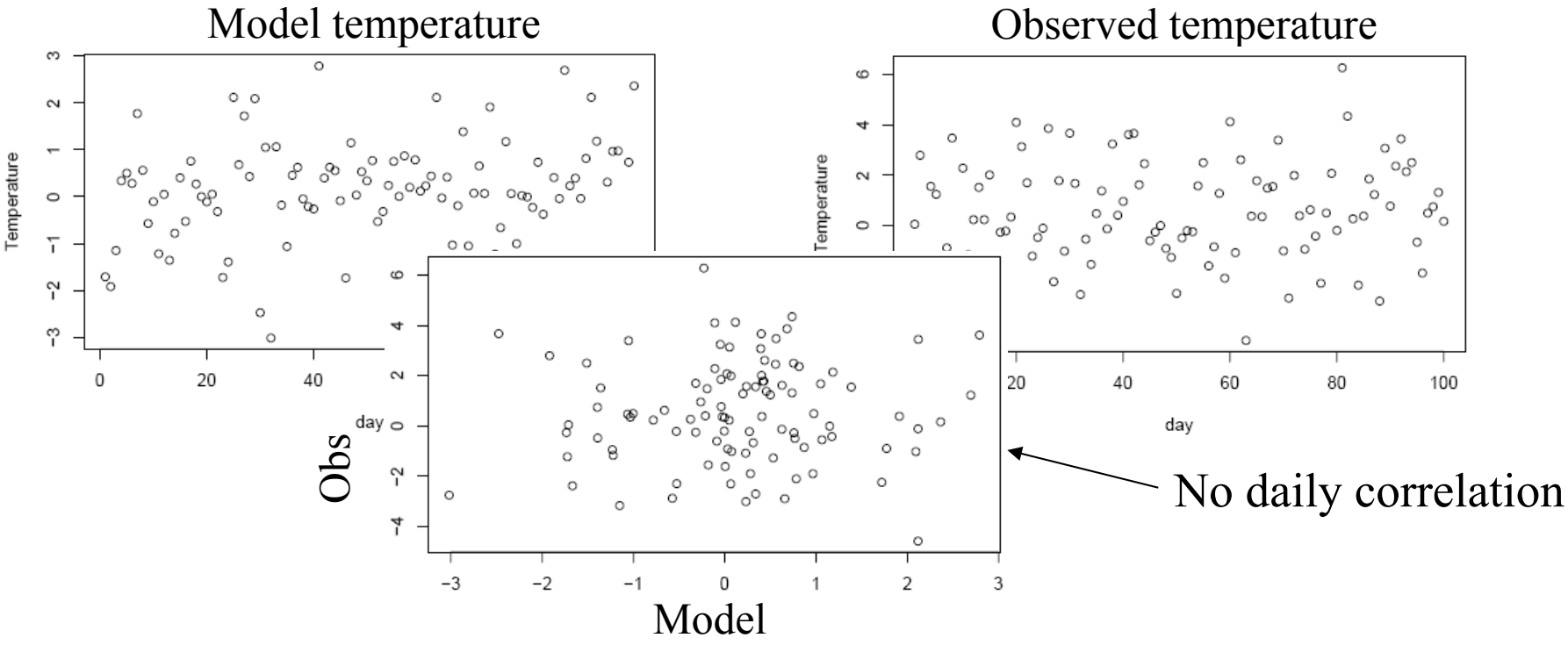
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

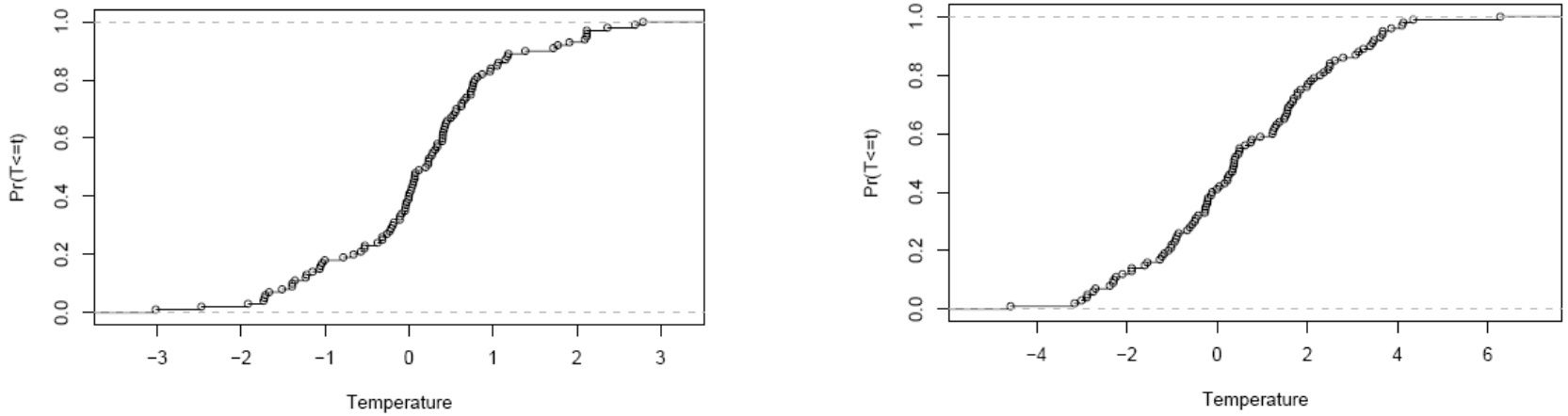


Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?



Main idea: Work on / relate (empirical) CDFs



Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

- Classical approach: Quantile-mapping
  - Station S; $X_S \sim F_S$; Gridcell G; $X_G \sim F_G$

$$F_S(x_S) = F_G(x_G) \quad \Leftrightarrow \quad x_S = F_S^{-1} F_G(x_G)$$

You **want** this — x_S

You **know** this — x_G

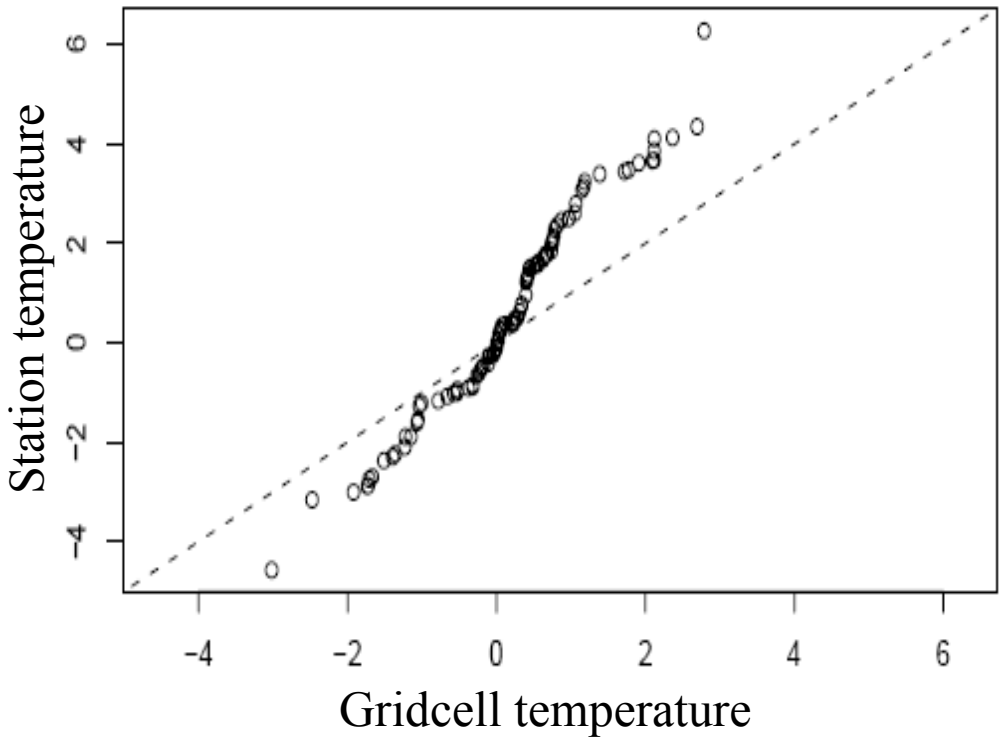
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

- Classical approach: Quantile-mapping
  - Station S; $X_S \sim F_S$; Gridcell G; $X_G \sim F_G$

$$F_S(x_S) = F_G(x_G) \quad \Leftrightarrow \quad x_S = F_S^{-1} F_G(x_G)$$

- Visual interpretation: Quantile-quantile plot (between F_S and F_G)



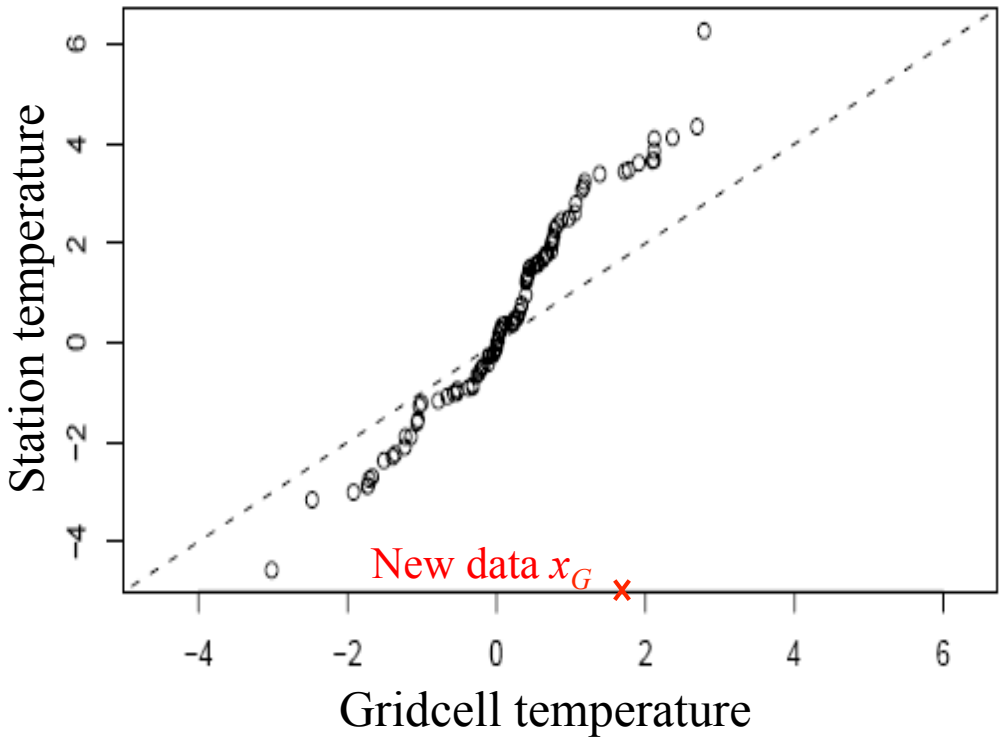
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

- Classical approach: Quantile-mapping
  - Station S; $X_S \sim F_S$; Gridcell G; $X_G \sim F_G$

$$F_S(x_S) = F_G(x_G) \quad \Leftrightarrow \quad x_S = F_S^{-1} F_G(x_G)$$

- Visual interpretation: Quantile-quantile plot (between F_S and F_G)



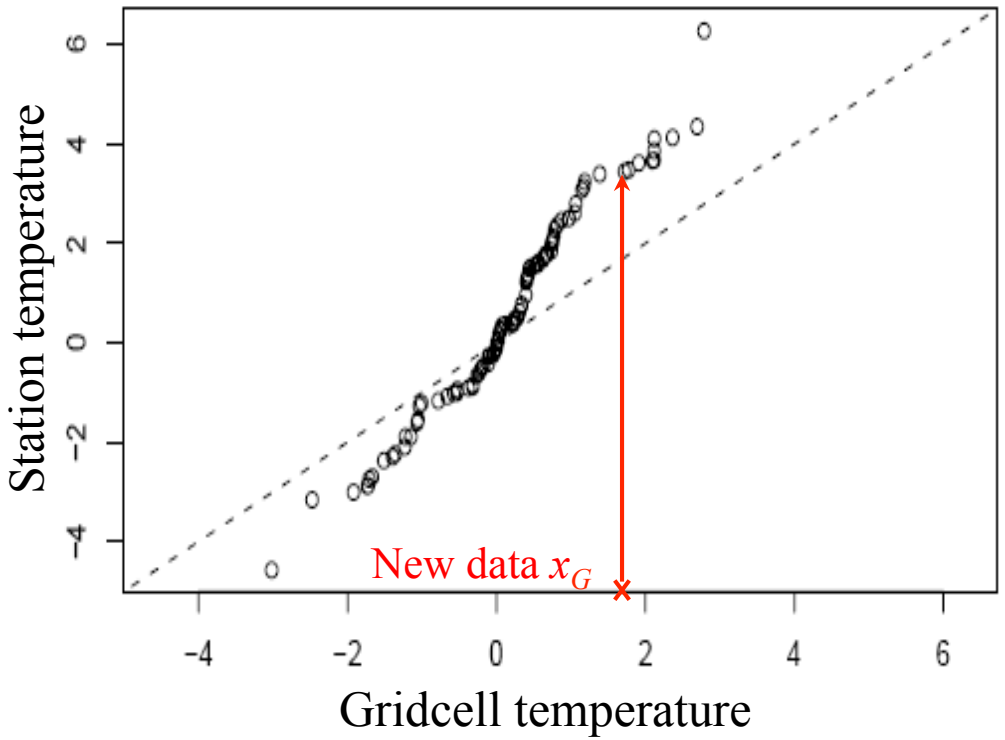
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

- Classical approach: Quantile-mapping
  - Station S; $X_S \sim F_S$; Gridcell G; $X_G \sim F_G$

$$F_S(x_S) = F_G(x_G) \quad \Leftrightarrow \quad x_S = F_S^{-1} F_G(x_G)$$

- Visual interpretation: Quantile-quantile plot (between F_S and F_G)



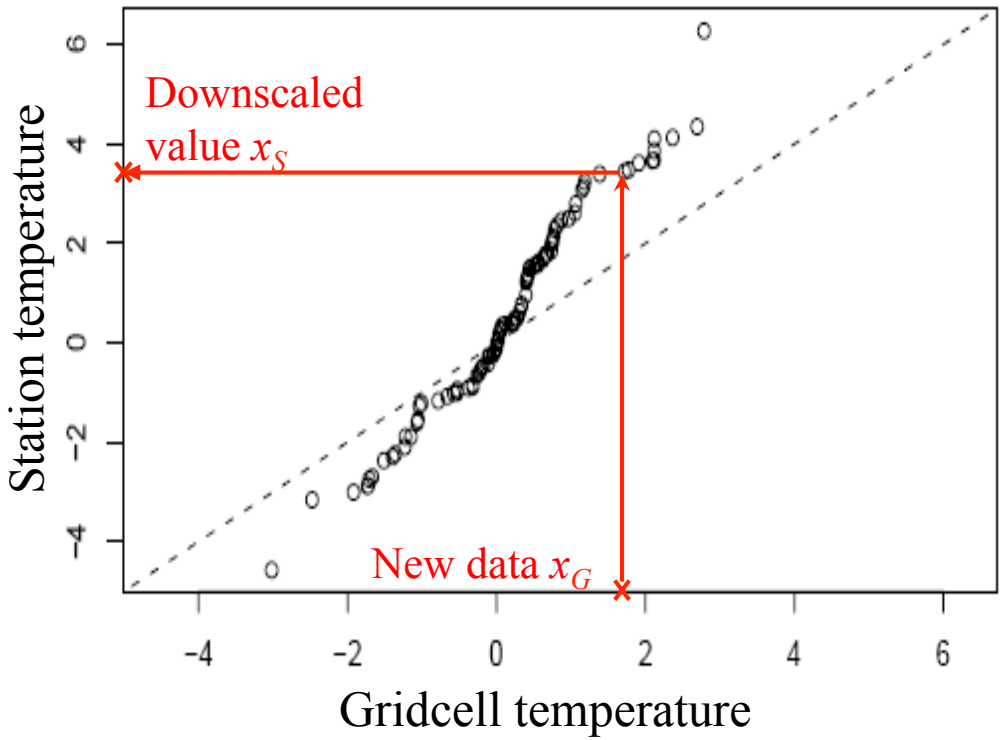
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

- Classical approach: Quantile-mapping
  - Station S; $X_S \sim F_S$; Gridcell G; $X_G \sim F_G$

$$F_S(x_S) = F_G(x_G) \quad \Leftrightarrow \quad x_S = F_S^{-1} F_G(x_G)$$

- Visual interpretation: Quantile-quantile plot (between F_S and F_G)



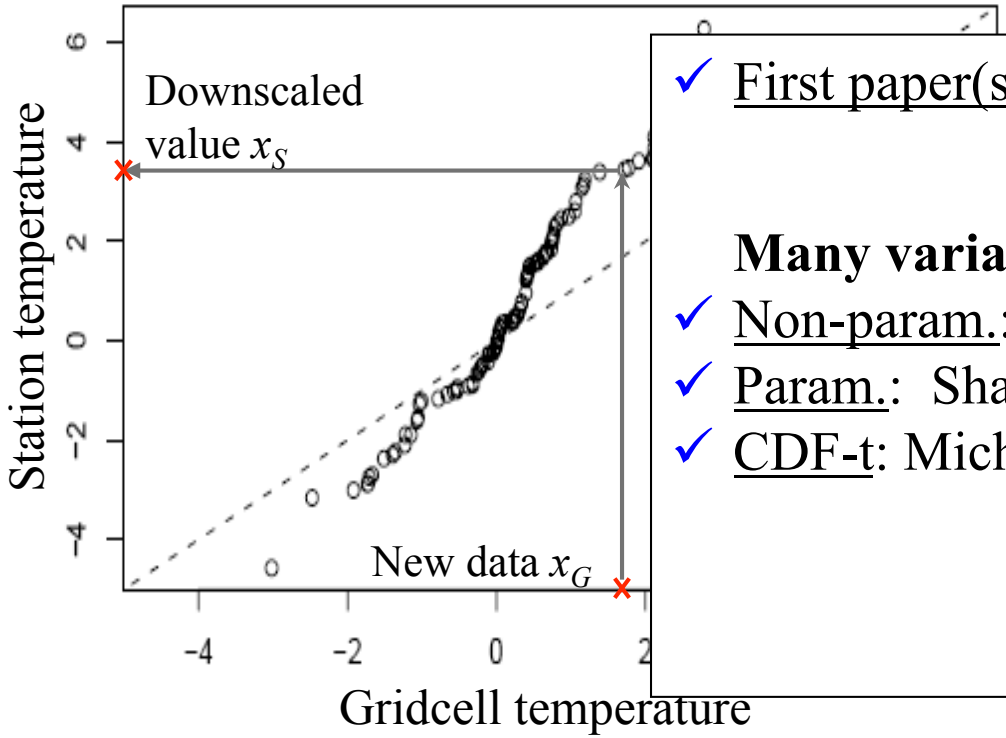
Model Output Statistics (MOS) / Bias correction

How to calibrate directly from GCM/RCM simulations?

- Classical approach: Quantile-mapping
  - Station S; $X_S \sim F_S$; Gridcell G; $X_G \sim F_G$

$$F_S(x_S) = F_G(x_G) \quad \Leftrightarrow \quad x_S = F_S^{-1} F_G(x_G)$$

- Visual interpretation: Quantile-quantile plot (between F_S and F_G)



- ✓ First paper(s): Panofsky and Brier (1958); Haddad and Rosenfeld (1997)

Many variants:

- ✓ Non-param.: Wood et al. (2004); Déqué (2007)
- ✓ Param.: Shabalova et al. (2003); Piani et al. (2010)
- ✓ CDF-t: Michelangeli et al. (2009); Oettli et al. (2011); Kallache et al. (2011); Vigaud et al. (2012); Vrac et al. (2012); Vautard et al. (2012), etc.

A few (last?) remarks on statistical downscaling

- **Many** (and many) **applications** of downscaling
  - Past-Present-Future in climate and impacts related studies

    ⇒ **Many** Statistical downscaling **approaches** (still increasing)

A few (last?) remarks on statistical downscaling

- **Many** (and many) **applications** of downscaling
  - Past-Present-Future in climate and impacts related studies

    ⇒ **Many** Statistical downscaling **approaches** (still increasing)
  - My favorite ones:
    - ✓ *Stochastic WGs*: cond'l event-wise variability/uncertainty
    - ✓ *MOS / Bias correction*: DS of CDFs from CDFs

A few (last?) remarks on statistical downscaling

- **Many** (and many) **applications** of downscaling
  - Past-Present-Future in climate and impacts related studies

    ⇒ **Many** Statistical downscaling **approaches** (still increasing)
  - My favorite ones:
    - ✓ *Stochastic WGs*: cond'l event-wise variability/uncertainty
    - ✓ *MOS / Bias correction*: DS of CDFs from CDFs
- **RCMs vs. SDMs**: Not a conflict => complementary approaches

A few (last?) remarks on statistical downscaling

- **Many** (and many) **applications** of downscaling
  - Past-Present-Future in climate and impacts related studies

    ⇒ **Many** Statistical downscaling **approaches** (still increasing)
  - My favorite ones:
    - ✓ *Stochastic WGs*: cond'l event-wise variability/uncertainty
    - ✓ *MOS / Bias correction*: DS of CDFs from CDFs
- **RCMs vs. SDMs**: Not a conflict => complementary approaches
- DS of **extreme** values **vs.** DS of ***indices*** of extremes

A few (last?) remarks on statistical downscaling

- **Many** (and many) **applications** of downscaling
  - Past-Present-Future in climate and impacts related studies

    ⇒ **Many** Statistical downscaling **approaches** (still increasing)
  - My favorite ones:
    - ✓ *Stochastic WGs*: cond'l event-wise variability/uncertainty
    - ✓ *MOS / Bias correction*: DS of CDFs from CDFs
- **RCMs vs. SDMs**: Not a conflict => complementary approaches
- DS of **extreme** values **vs.** DS of ***indices*** of extremes
- There is not one good SDM for all variables and regions

  ⇒ Different skills according to regions/variables/applications, etc.

And now, what?... (A "StaRMIP" perspective)

And now, what?... (A "StaRMIP" perspective)

- **Intercomparisons of SDMs**
  - Only few SDMs are inter-compared:

  ⇒ Need more: between them & with RCMs

  ✓ Not only "transfer functions" approaches to be tested

  (Statistical component of Med-CORDEX)

And now, what?... (A "StaRMIP" perspective)

- **Intercomparisons of SDMs**
  - Only few SDMs are inter-compared:

  ⇒ Need more: between them & with RCMs

  ✓ Not only "transfer functions" approaches to be tested

  (Statistical component of Med-CORDEX)

- Various **evaluations**

  ⇒ Statistical **indicators of quality** of high-resolution simulations

  ✓ Spatial & temporal dependences, extremes, inter-var. (COST Value)

And now, what?... (A "StaRMIP" perspective)

- **Intercomparisons of SDMs**
  - Only few SDMs are inter-compared:

  ⇒ Need more: between them & with RCMs

  ✓ Not only "transfer functions" approaches to be tested

  (Statistical component of Med-CORDEX)

- Various **evaluations**

  ⇒ Statistical **indicators of quality** of high-resolution simulations

  ✓ Spatial & temporal dependences, extremes, inter-var. (COST Value)

  ⇒ **Uncertainty** modelling (TWP3 L-IPSL)

  ✓ Ensembles of SDMs

  ✓ "*Model-merging*": (non-) linear, Bayesian, etc., combinations

And now, what?... (A "StaRMIP" perspective)

- **Intercomparisons of SDMs**
  - Only few SDMs are inter-compared:

  ⇒ Need more: between them & with RCMs

  ✓ Not only "transfer functions" approaches to be tested

  (Statistical component of Med-CORDEX)

- Various **evaluations**

  ⇒ Statistical **indicators of quality** of high-resolution simulations

  ✓ Spatial & temporal de… inter-var. (COST Value)

  ⇒ **Uncerta…** …PSL)

  ✓ Ensem…les of SDMs

  ✓ "*Model-merging*": (non-) linear, Bayesian, etc., combinations

Including RCMs!!

And now, what?... (A "StaRMIP" perspective)

- SDMs are often univariate (although covariates):
  - ⇒ Needs for **inter-sites** models (stations or grid-cells - PLEIADES)
    - ? *Latent* (i.e. cond'l ind.)? Or *Complete* dependence structure?
  - ⇒ Needs for **inter-variables** models (b/ climate variables – Ex. Earth?)
    - ? MOS ? SWGs?
  - ⇒ Needs for **spatial models**: SD even at locations where no data
    - ? Continuous spatial processes? Inter/Extra-polation of parameters?
  - ➢ Especially for **dependence of extremes** (ANR McSIM)

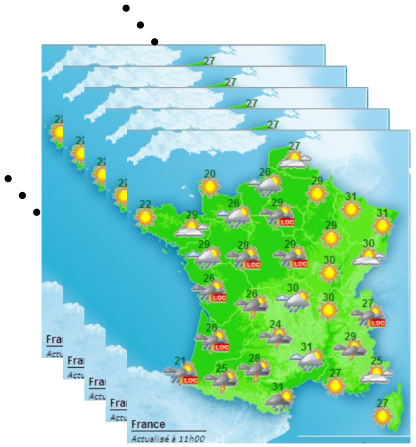
And now, what?... (A "StaRMIP" perspective)

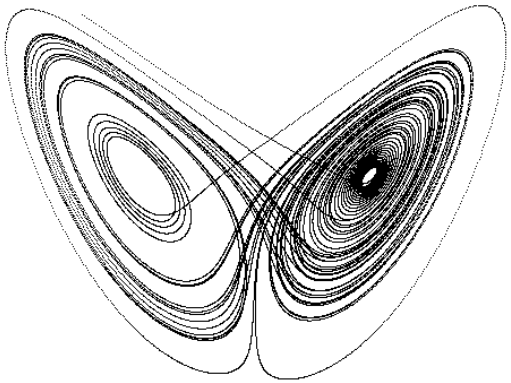
- SDMs are often univariate (although covariates):
  - ⇒ Needs for **inter-sites** models (stations or grid-cells - PLEIADES)
    - ? *Latent* (i.e. cond'l ind.)? Or *Complete* dependence structure?
  - ⇒ Needs for **inter-variables** models (b/ climate variables – Ex. Earth?)
    - ? MOS ? SWGs?
  - ⇒ Needs for **spatial models**: SD even at locations where no data
    - ? Continuous spatial processes? Inter/Extra-polation of parameters?
  - ➢ Especially for **dependence of extremes** (ANR McSIM)

- Dissemination / Valorization: **Simulations & R packages**
  - ✓ A real strength for applications and proper use of developments

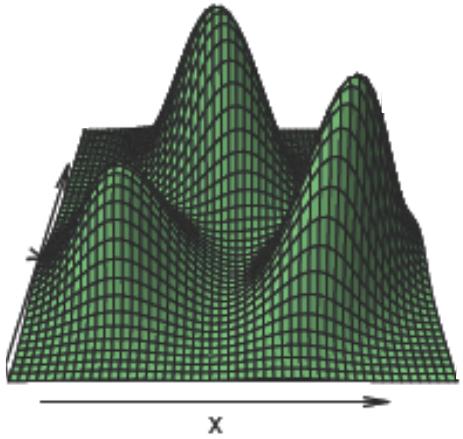
Meteorology ≠ Climate

- Time: ~1 week vs. 100 years





- Dynamics: 1 trajectory vs. the "attractor"



- Statistics:
  1 realization vs. its **random variable**

Main thread of this talk:

What we need is the **pdf or CDF** describing the climate variables

The *Med-CORDEX* initiative

- CORDEX: International program (WCRP):
  - ✓ Provide **regional** (i.e., downscaled) **climate data** for **recent past** and **21st century projections** on 12 major regions of the world
  - ✓ Provide relevant **climate information** and provide **expertise** to groups leading **impacts and adaptation studies**

= CMIP5 for the regional models

- Med-CORDEX

  = CORDEX + HYMEX + MEDCLIVAR

  - ✓ **Dynamical** (DD) and **statistical** (SD) downscaling on Mediterranean domain
  - ✓ Focus on **extreme events**
  - ✓ **SD** of Med-CORDEX **= StaRMIP**