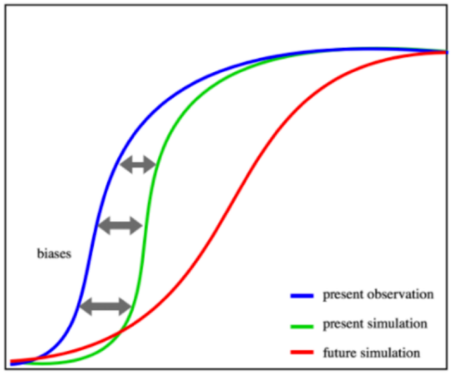
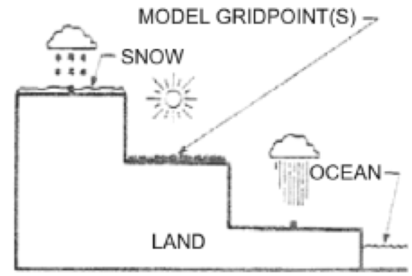
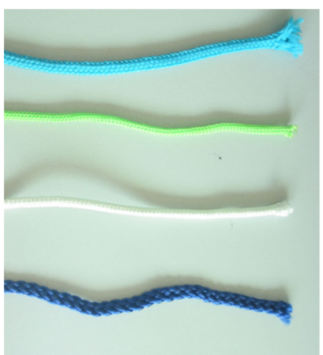




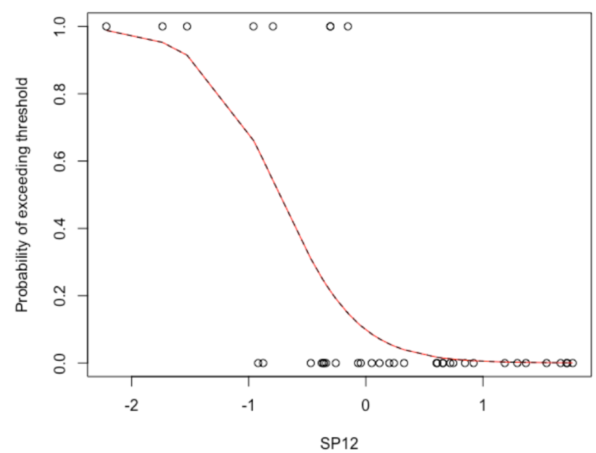
Statistical modeling frameworks for extremes

Erin Towler

C3WE, NCAR



July 26, 2019
Weather & Climate Extremes Tutorial
NCAR



This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Yesterday, Cindy introduced:

Dynamical Downscaling

Overview and Best Practices

Cindy Bruyère, C3WE/NCAR

And the associated considerations:

- Resolution
- Domain size
- Model Physics
- Input Data

Dynamical modeling frameworks, also called "dynamical downscaling" for regional climate or extremes have advantages and limitations.

Advantages

- Responses are physically consistent
- Produces finer resolution data

Limitations

- Strongly dependent on GCM forcing data
- Computationally intensive
- Some variables are not well modeled (insufficient resolution or inadequate process understanding)

(Fowler et al. 2007)

Some of the limitations can be addressed through statistical modeling frameworks, or “statistical downscaling” (SD)

Limitations

- Strongly dependent on GCM forcing data
- Computationally intensive
- Some variables are not well modeled (insufficient resolution or inadequate process understanding)

Some of the limitations can be addressed through statistical modeling frameworks, or "statistical downscaling" (SD)

SD can correct bias in forcing GCMs/RCMs →

Limitations

- Strongly dependent on GCM forcing data
- Computationally intensive
- Some variables are not well modeled (insufficient resolution or inadequate process understanding)

Some of the limitations can be addressed through statistical modeling frameworks, or “statistical downscaling” (SD)

Limitations

SD is computationally efficient, and impact studies often require climate information at **finer spatial scales** than provided by either GCMs or RCMs

- Strongly dependent on GCM forcing data
- Computationally intensive
- Some variables are not well modeled (insufficient resolution or inadequate process understanding)

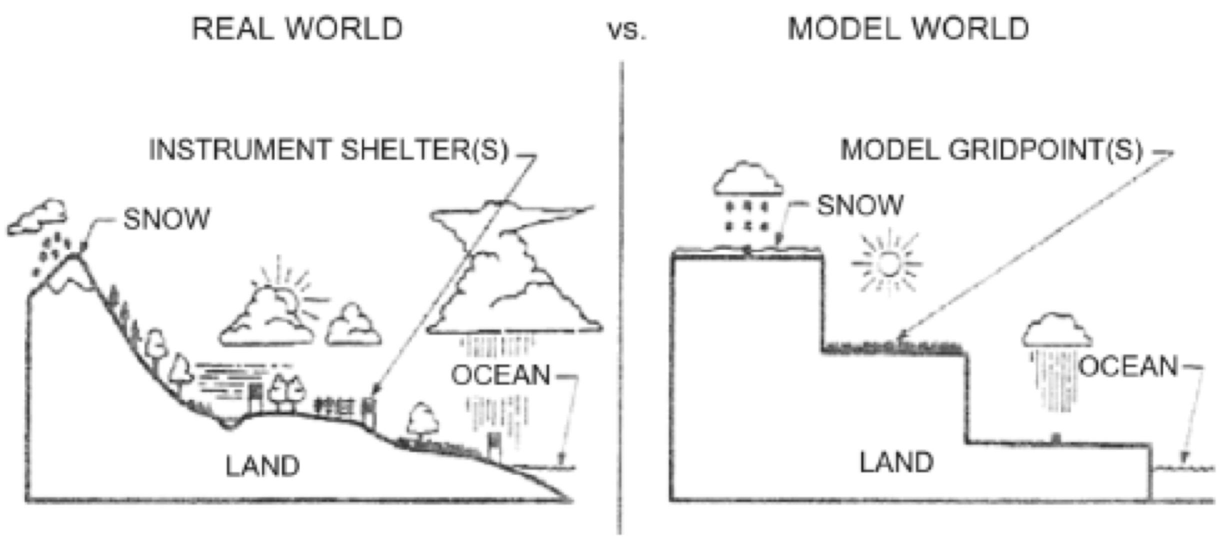
Some of the limitations can be addressed through statistical modeling frameworks, or "statistical downscaling" (SD)

Limitations

Can empirically derive variables not available or not well modeled by GCMs/RCMs.

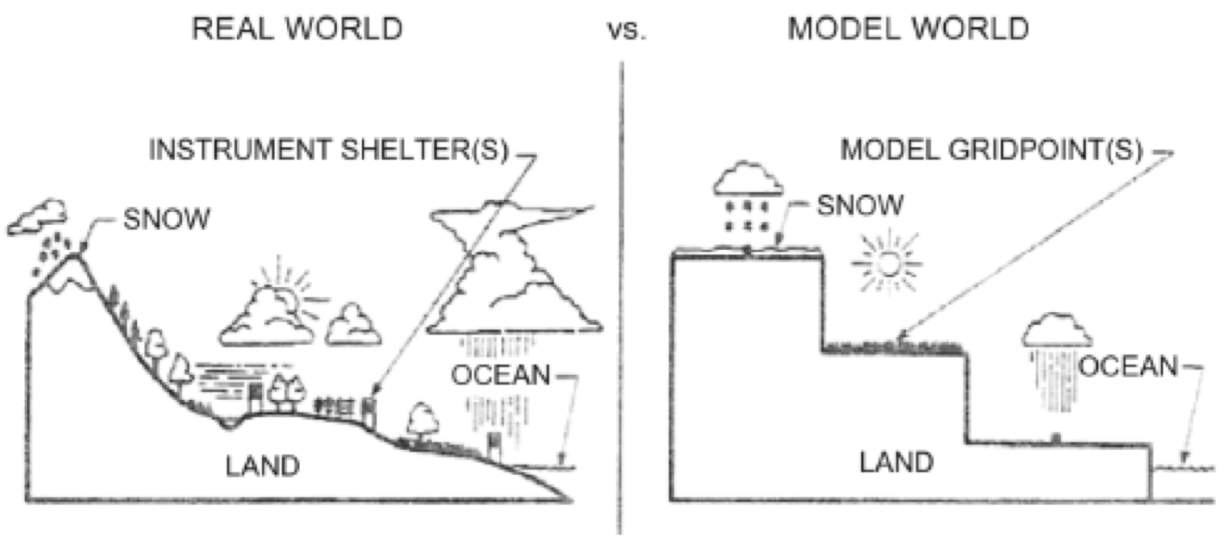
- Strongly dependent on GCM forcing data
- Computationally intensive
- Some variables are not well modeled (insufficient resolution or inadequate process understanding)

But, extremes were not the original focus of statistical downscaling…



(Wilks 2011 Fig 7.19)

But, extremes were not the original focus of statistical downscaling…



(Wilks 2011 Fig 7.19)

So it’s important to understand statistical downscaling techniques and how they handle extremes.

SD relates large-scale climate variables (predictors) to local or regional variables (predictants)

$$y = f(x)$$

predictant

statistical model

predictor

- Extremely flexible
- Makes SD difficult to "neatly" categorize or summarize
- SD requires some data analysis and statistical tools

3 SD classifications:

1. Perfect prognosis downscaling
2. Model output statistics (MOS)
3. Weather generators

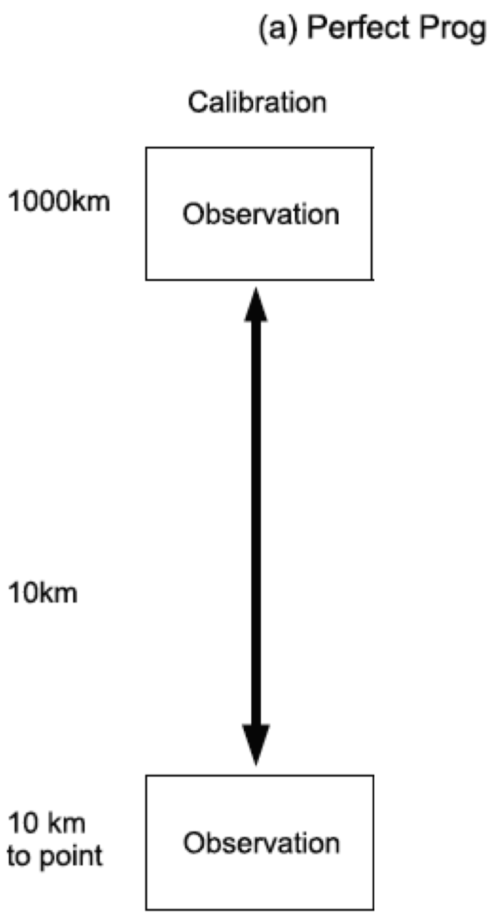
Differentiated based on nature of predictor (observation vs. model output)

Many incorporate MOS and PP aspects

- Note: Different fields use different terminology and there are often new hybrid approaches coming out, but we will provide the broad classifications here.

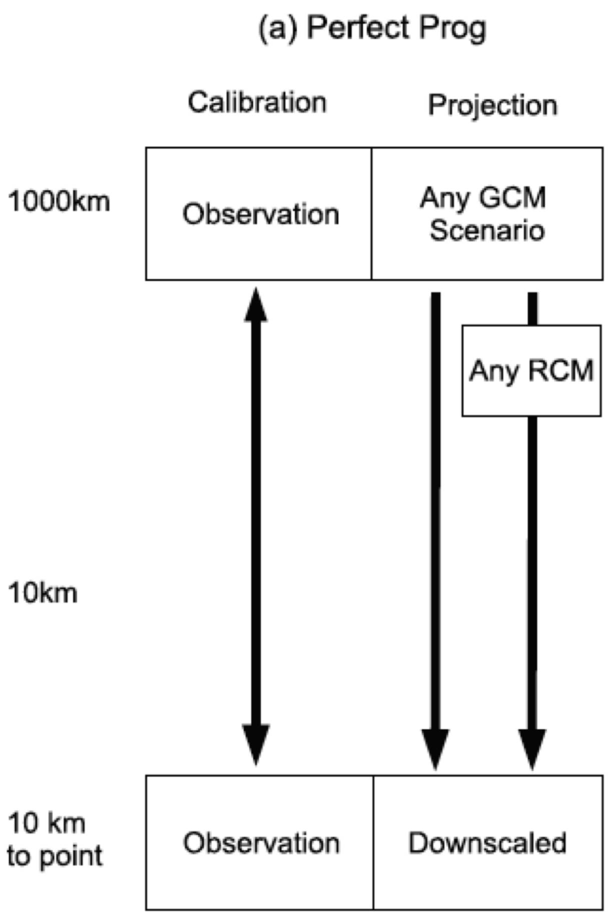
(Classifications based on Maraun et al 2010)

Perfect prognosis (PP) downscaling relates **observed** large-scale predictors to **observed** local-scale predictants



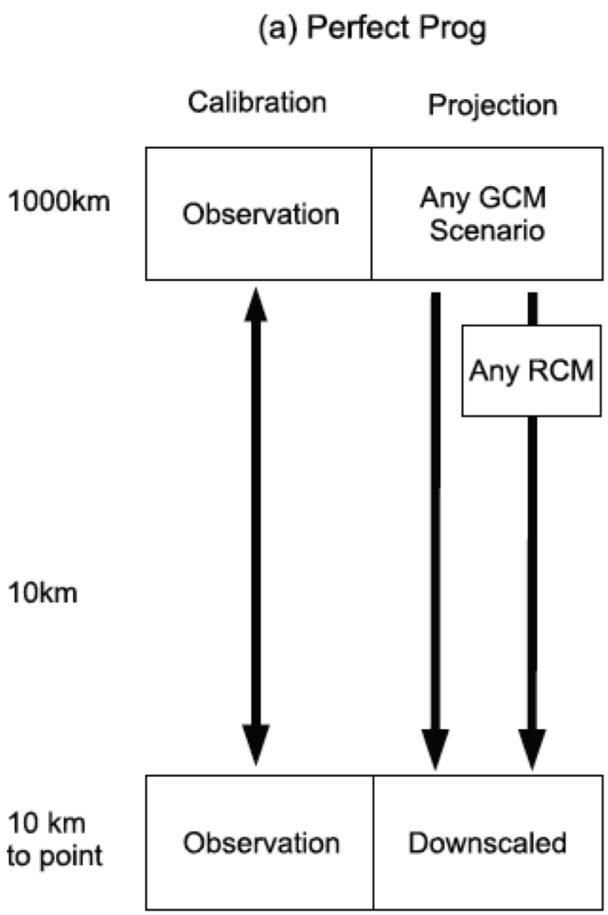
(Maraun et al 2010, Fig 2a)

Perfect prognosis (PP) downscaling relates **observed** large-scale predictors to **observed** local-scale predictants



(Maraun et al 2010, Fig 2a)

Perfect prognosis (PP) downscaling relates **observed** large-scale predictors to **observed** local-scale predictants



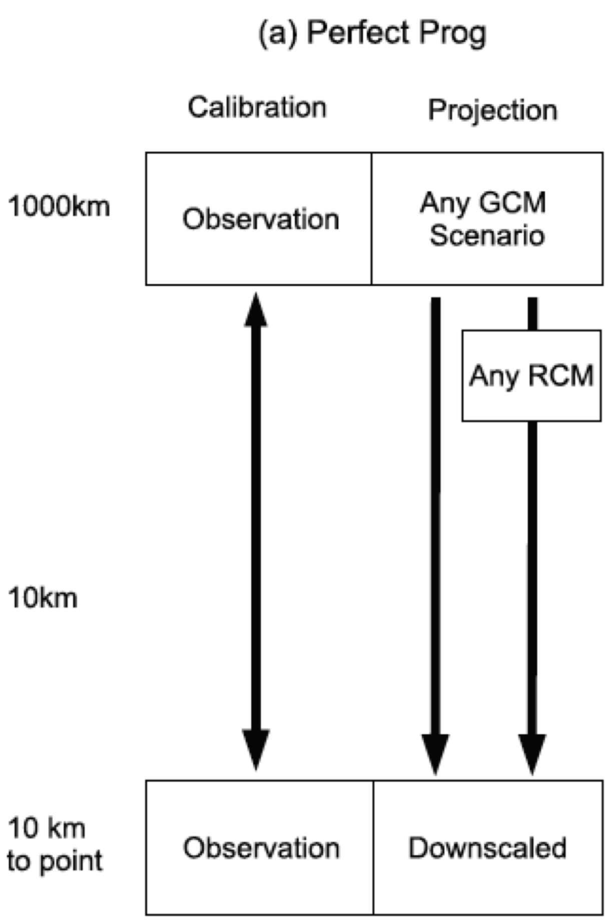
Key steps

1) Identify informative large-scale predictors (e.g., atmospheric circulation, humidity, etc.)

2) Develop statistical model (i.e., regression methods)

(Maraun et al 2010, Fig 2a)

Perfect prognosis (PP) downscaling relates **observed** large-scale predictors to **observed** local-scale predictants



Key steps

1) Identify informative large-scale predictors (e.g., atmospheric circulation, humidity, etc.)
2) Develop statistical model (i.e., regression methods)

Main assumptions

1) Large-scale predictor is well-simulated by climate model
2) Statistical relationship is constant in future (temporally stable)

(Maraun et al 2010, Fig 2a)

Statistical models commonly used for perfect prognosis (PP) downscaling:

1. Linear Regression

2. Generalized linear models (GLM)

3. Nonlinear regression

4. Analogue method

5. Extreme value statistics *(talk after break today)*

(See Maraun et al 2010 for additional models)

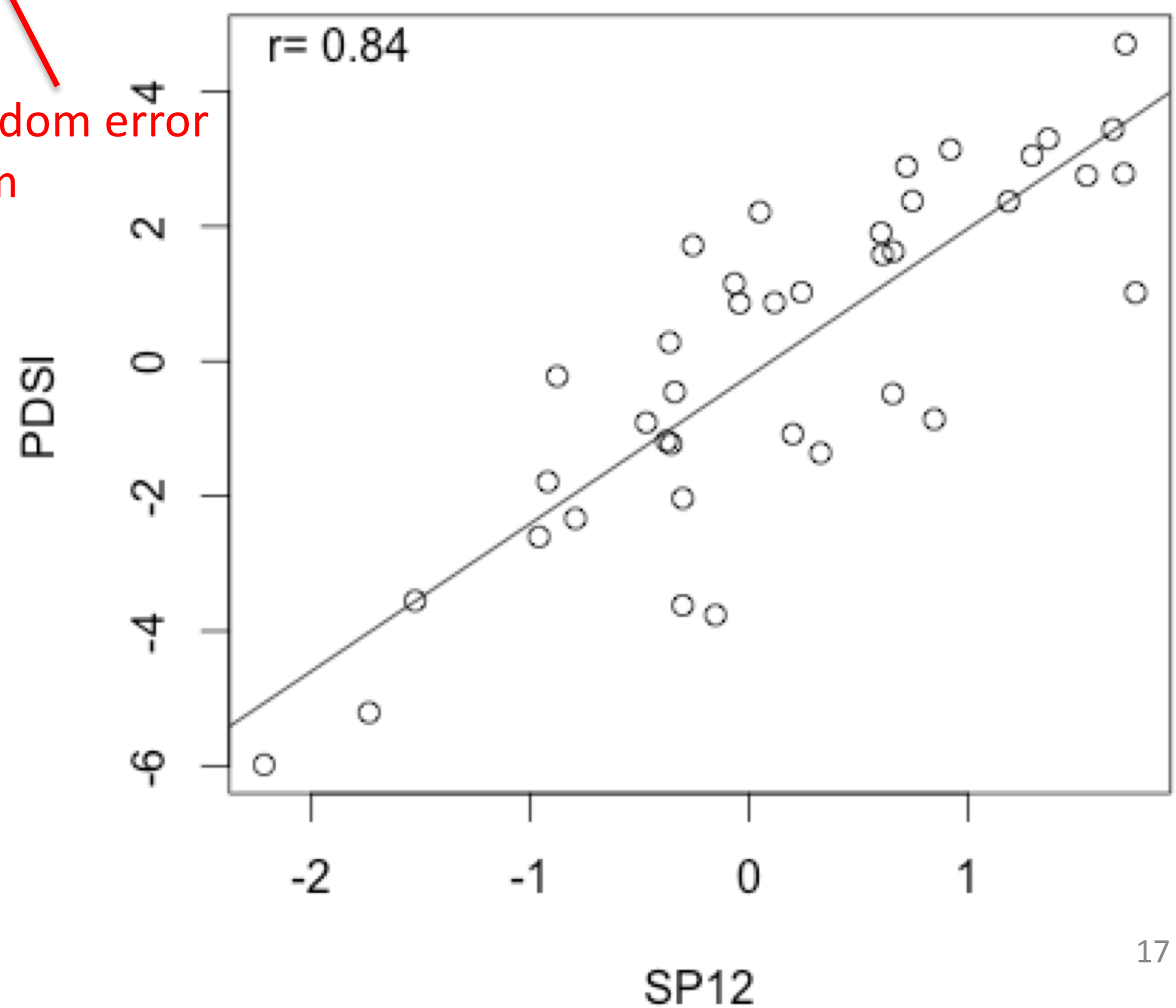
Linear regression is simple way to relate two variables

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Intercept

Slope

Random error term



Linear regression is simple way to relate two variables

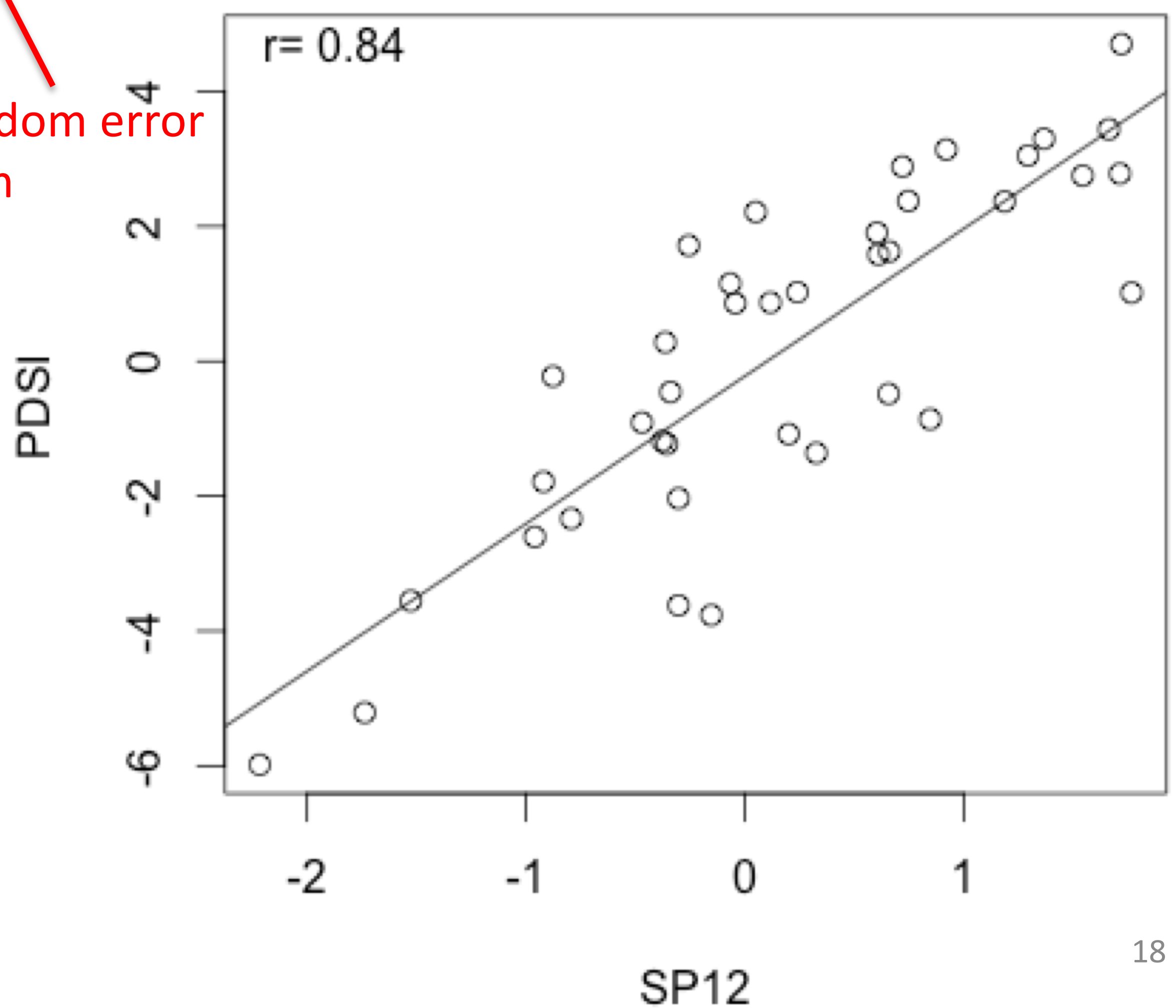
$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Intercept

Slope

Random error term

Good for estimating the expected value, or averages, but may not be appropriate for extremes.



Generalized linear models (GLMs) are more flexible approach for modeling responses with different attributes (continuous, categorical, integer etc).

Linear Regression: $\hat{y} = \beta_0 + \beta_1 x$ (continuous)
Logistic: $logit(\hat{y}) = \beta_0 + \beta_1 x$ (categorical)
Poisson: $log(\hat{y}) = \beta_0 + \beta_1 x$ (Integer, count data)

(For details on GLM see McCullagh and Nelder 1989)

Categorical data can be modeled with a binomial distribution, or logistic regression

$$logit(\hat{y}) = \beta_0 + \beta_1 x$$

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

$$\frac{p}{1-p} = exp(\beta_0 + \beta_1 x)$$

$$p = \frac{exp(\beta_0 + \beta_1 x)}{1 + exp(\beta_0 + \beta_1 x)}$$

(Helsel and Hirsh 1995)

Categorical data can be modeled with a binomial distribution, or logistic regression

$$logit(\hat{y}) = \beta_0 + \beta_1 x$$

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

$$\frac{p}{1-p} = exp(\beta_0 + \beta_1 x)$$

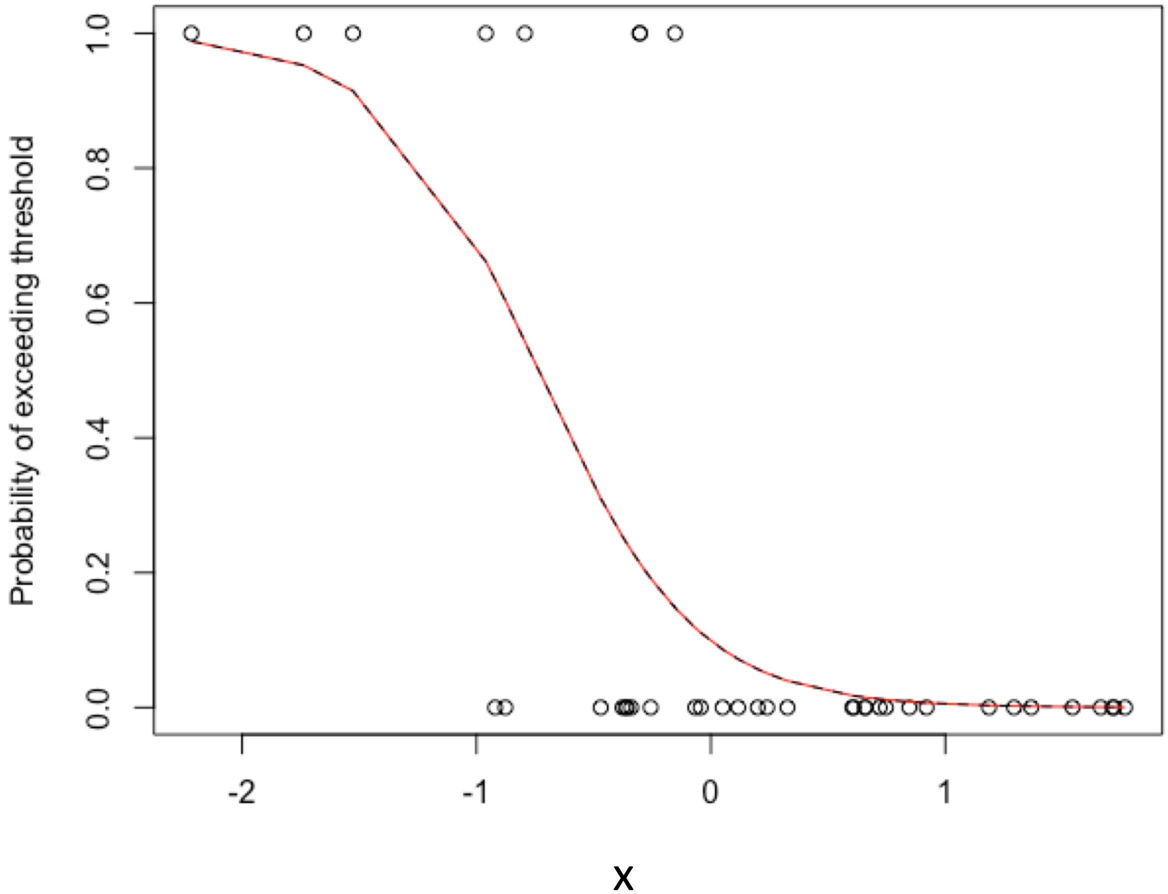
$$p = \frac{exp(\beta_0 + \beta_1 x)}{1 + exp(\beta_0 + \beta_1 x)}$$

(Helsel and Hirsh 1995)

Examples:
What's the probability of wet versus dry? *(# of wet days is an extreme characteristic)*

Or of being above or below an impact threshold? *(Threshold is subjective, but moves towards extremes)*

Use glm(y~x, family = “binomial”) to predict probability of exceedance



Integer, or count data can be modeled with a Poisson distribution

Use glm(y~x, family = “poisson”)

$$log(y) = \beta_0 + \beta_1 x_1$$

Integer, or count data can be modeled with a Poisson distribution

Use glm(y~x, family = "poisson")

$$log(y) = \beta_0 + \beta_1 x_1$$

Examples: How many hurricanes will make landfall in a year? *(These are extreme events!)*

How many days will we violate the minimum streamflow flow threshold? *(This is extreme for the fish!)*

PP statistical downscaling has some limitations:

- Requires long and reliable observed historical data for model development
- Dependent on choice of predictors
- Assumes the predictor-predictand relationship stays constant

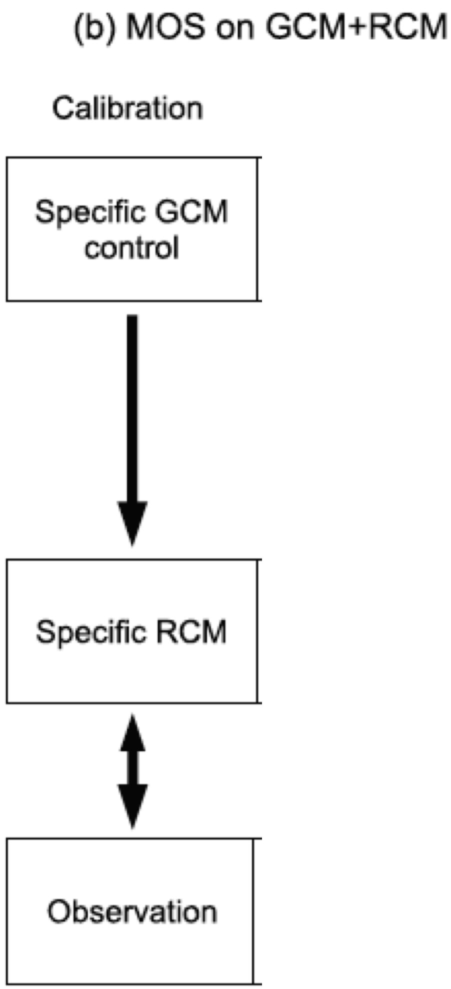
(Fowler et al. 2007, Table 1)

Summary of PP statistical downscaling for extremes:

- PP approach can be tailored to estimate characteristics of extremes, especially those that are not available or well simulated by GCMs/RCMs.
  - Flexible and portable to different users, regions, and subjective definitions of what is extreme.

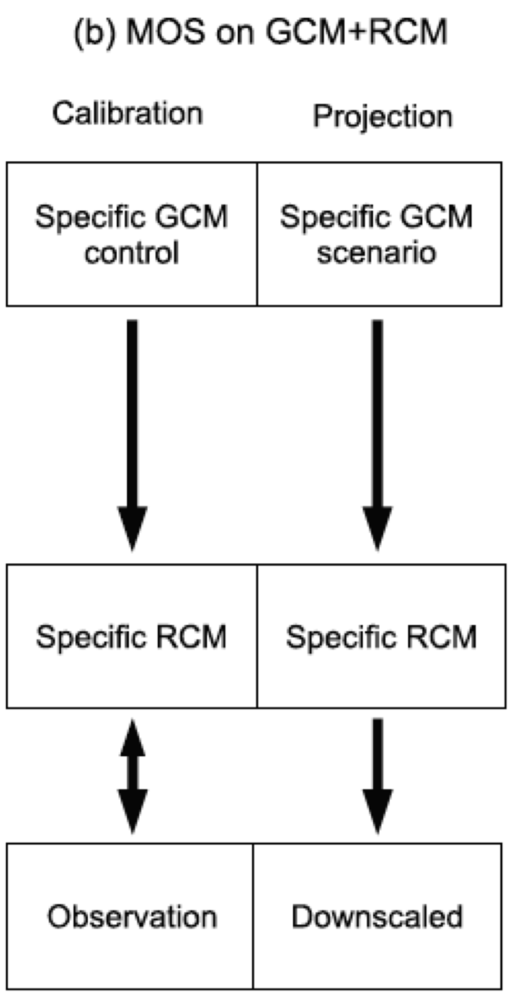
- Extreme value theory (talk after break) is a PP method for more rare extreme events.

Model output statistics (MOS) downscaling relates **modeled** large-scale predictors to **observed** local-scale predictants



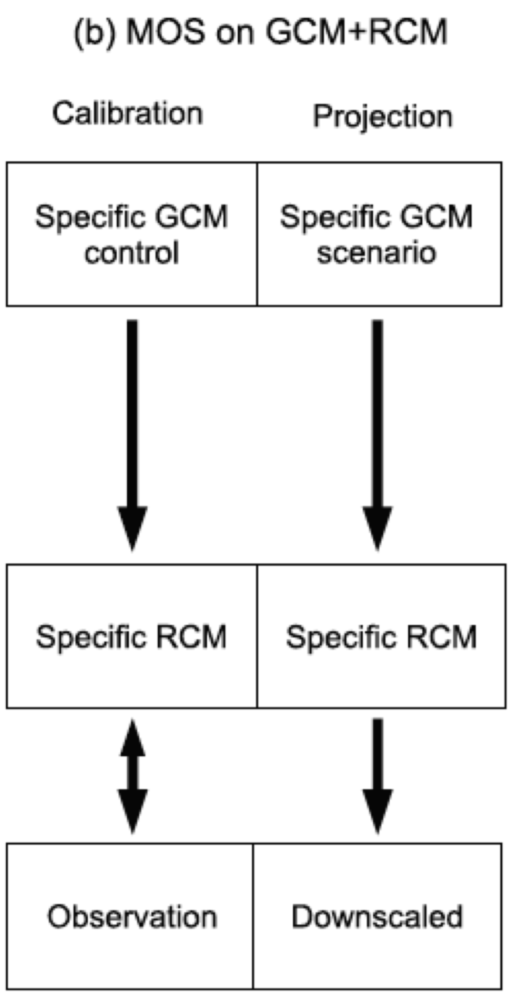
(Maraun et al 2010, Fig 2b)

Model output statistics (MOS) downscaling relates **modeled** large-scale predictors to **observed** local-scale predictants



(Maraun et al 2010, Fig 2b)

Model output statistics (MOS) downscaling relates **modeled** large-scale predictors to **observed** local-scale predictants

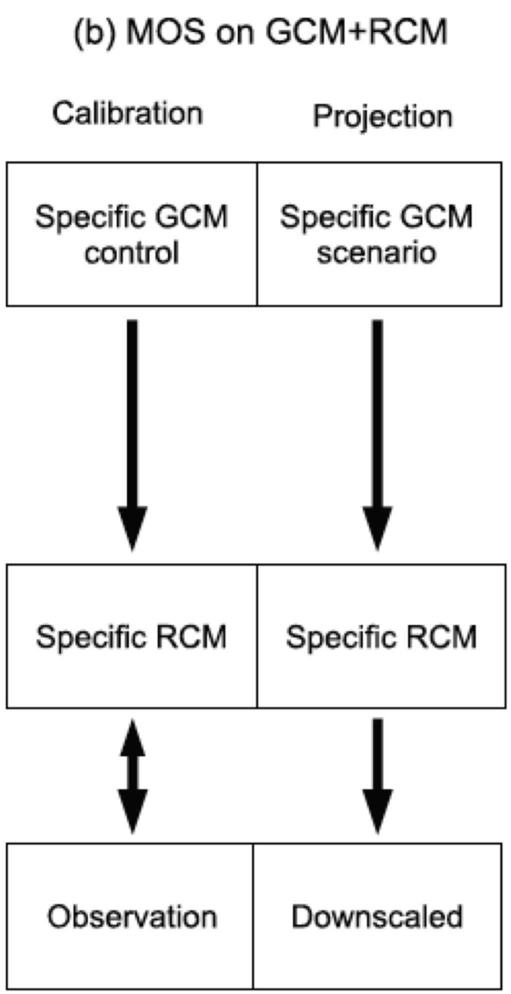


Key points

1) MOS can only be used to relate distributional characteristics (unless model is run by reanalysis or is for operational weather forecasting)

(Maraun et al 2010, Fig 2b)

Model output statistics (MOS) downscaling relates **modeled** large-scale predictors to **observed** local-scale predictants



Key points

1) MOS can only be used to relate distributional characteristics (unless model is run by reanalysis or is for operational weather forecasting)

Main assumptions

1) Large-scale predictor is well-simulated by climate model
2) Statistical relationship is constant in future (temporally stable) (Maraun et al 2010, Fig 2b)

Statistical methods commonly used for MOS downscaling:

1. Change factor/delta method

2. Bias correction/direct method

3. Distribution mapping/quantile Mapping

(See Maraun et al 2010 for details)

Change factor (CF) is simplest of MOS methods: Rescaling observations

Change factor (CF) is simplest of MOS methods: Rescaling observations

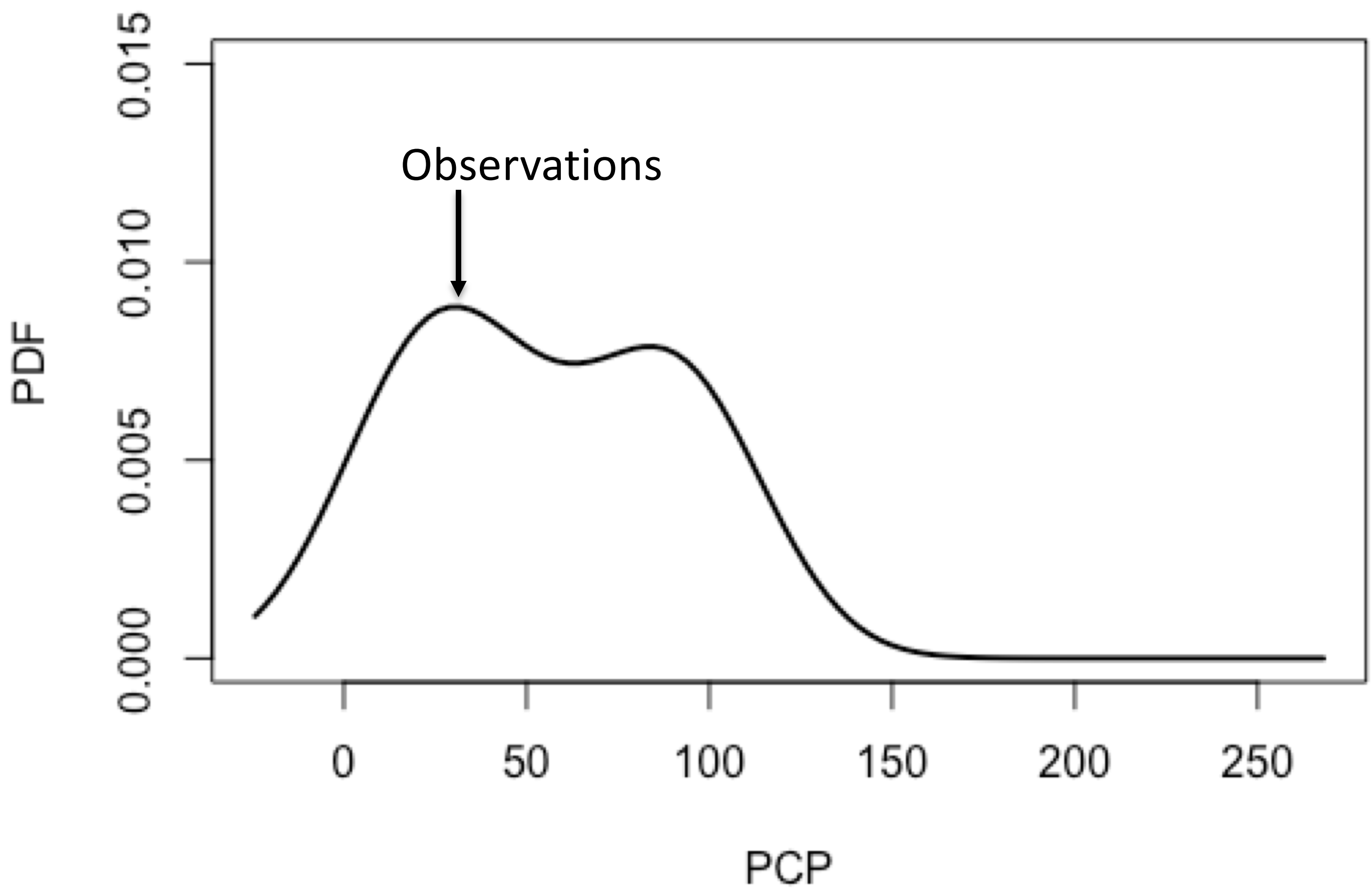
Temperature CF: Difference between mean control model (*Mc*) and mean future model (*Mf*) run added to observations (*O*):

TEMP (T): $CF_T = \overline{T_{Mf}} - \overline{T_{Mc}}$ $\quad T_f = T_O + CF_T$ (Also called "delta method")

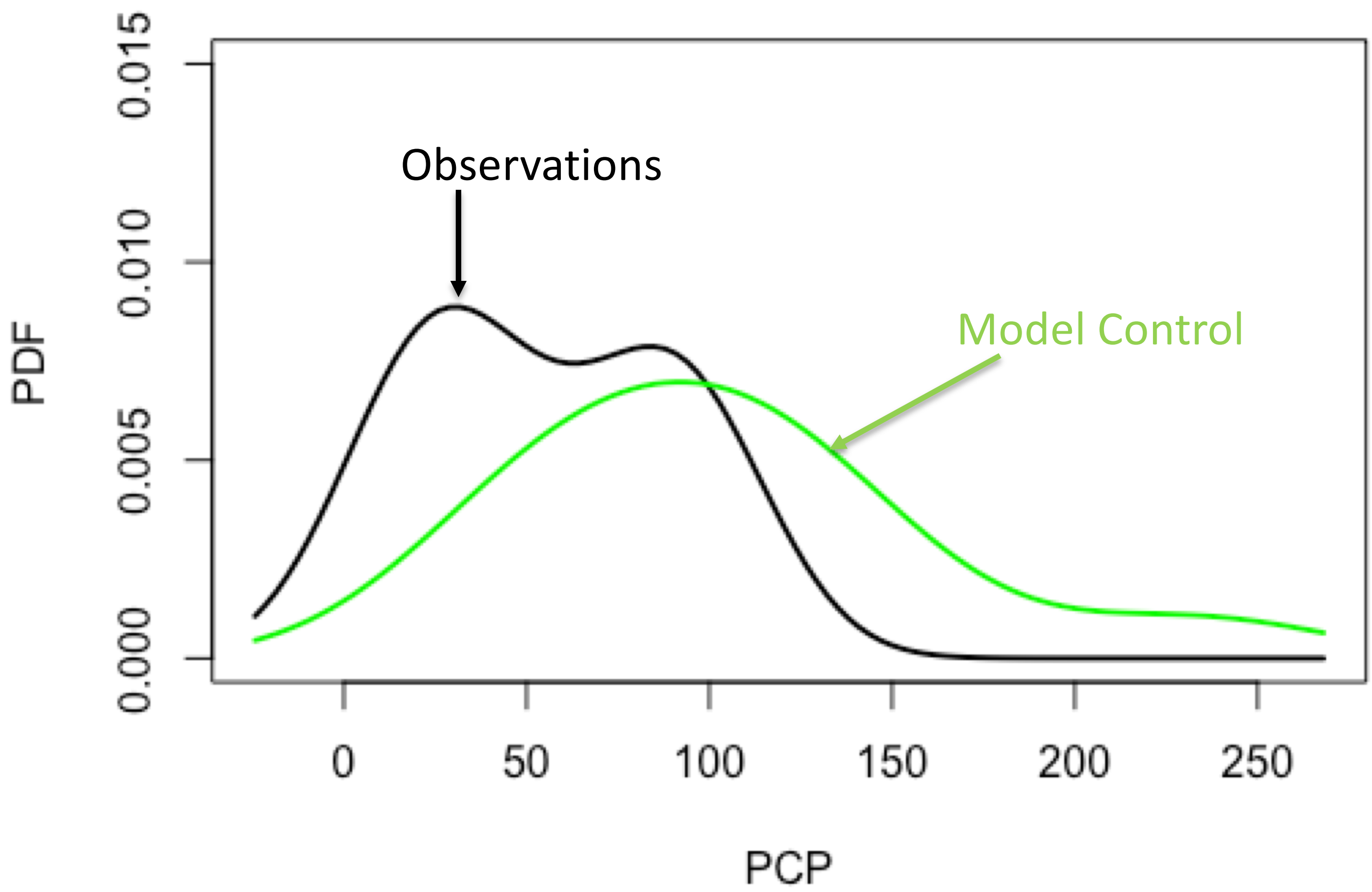
Precipitation CF: Ratio between future model (*Mf*) and control model (*Mc*) multiplied by observations (*O*):

PREC (P): $CF_P = \frac{\overline{P_{Mf}}}{\overline{P_{Mc}}}$ $\quad P_f = P_O * CF_P$ (Maraun et al 2010)

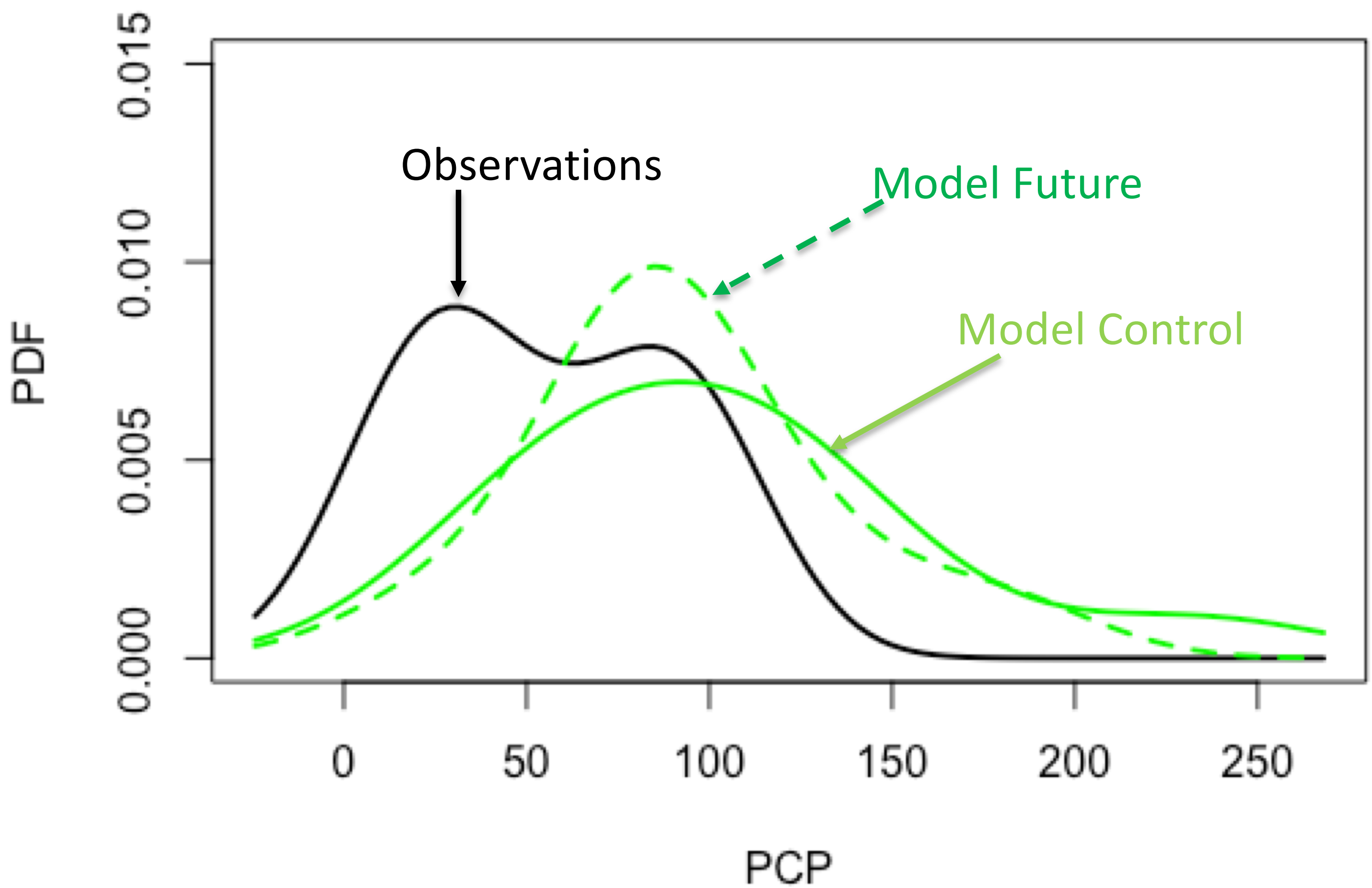
CF MOS example: Rescaling observations



CF MOS example: Rescaling observations

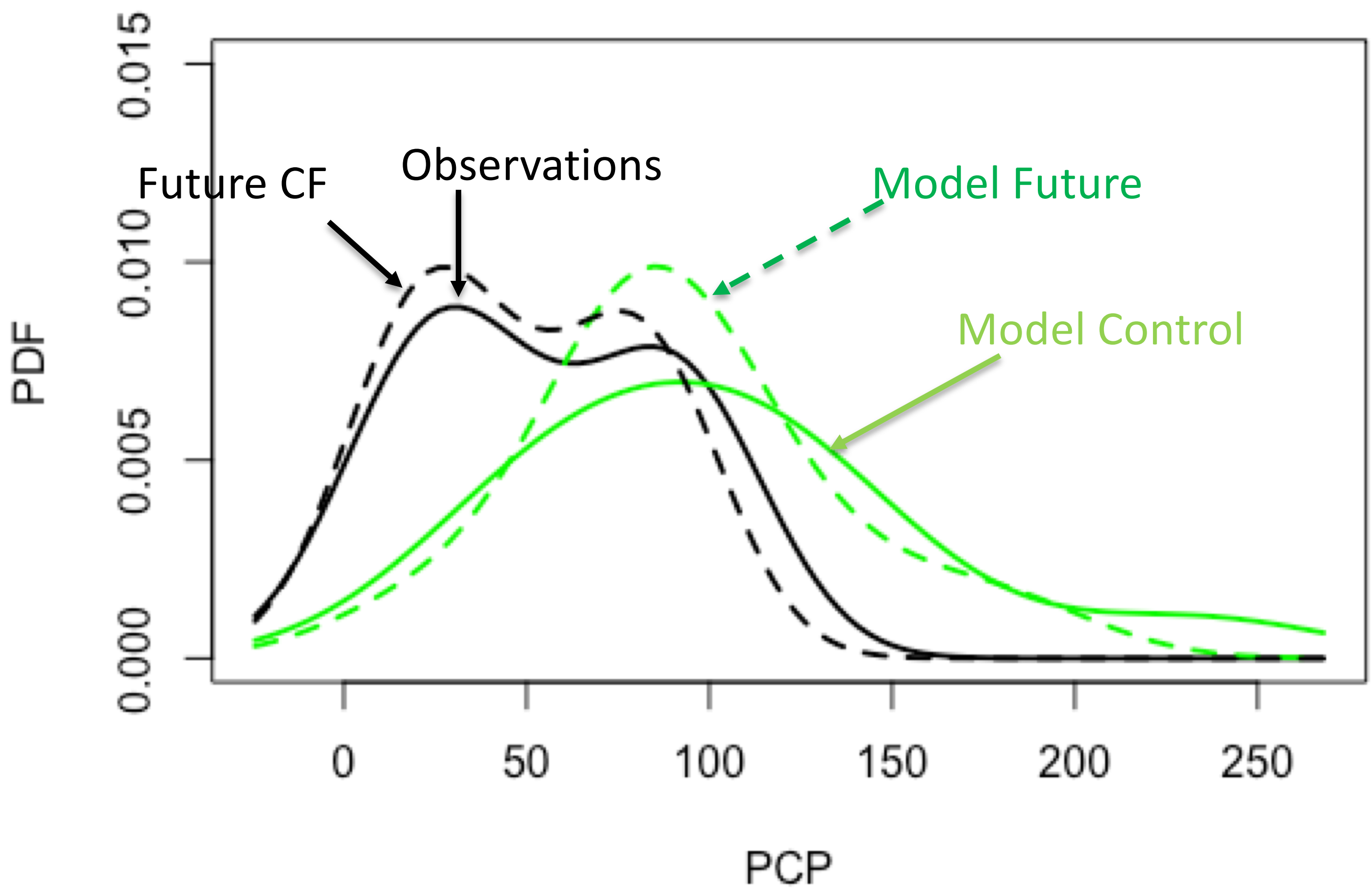


CF MOS example: Rescaling observations



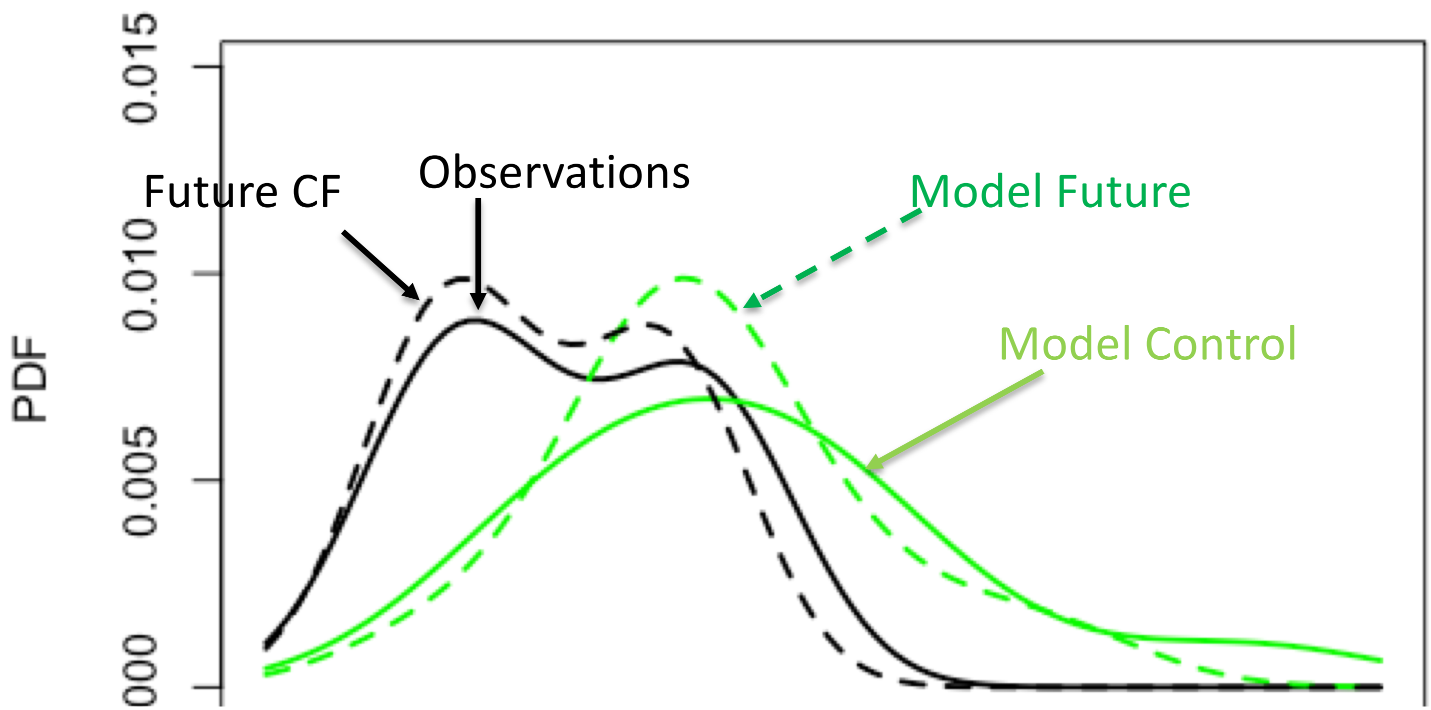
CF MOS example: Rescaling observations

* CF = .9, so multiply every PCP value by .9



CF MOS example: Rescaling observations

* CF = .9, so multiply every PCP value by .9



Limitations:

- Spatial pattern & temporal sequencing of present climate remains unchanged
- All parts of distribution (mean to extremes) are rescaled by single CF

MOS “bias correction” is another way to recalibrate observations to model output; rescaling model output

MOS “bias correction” is another way to recalibrate observations to model output; rescaling model output

Temperature Bias correction (BC): Difference between control model (*Mc*) and observations (*O*) added to future model (*Mf*)

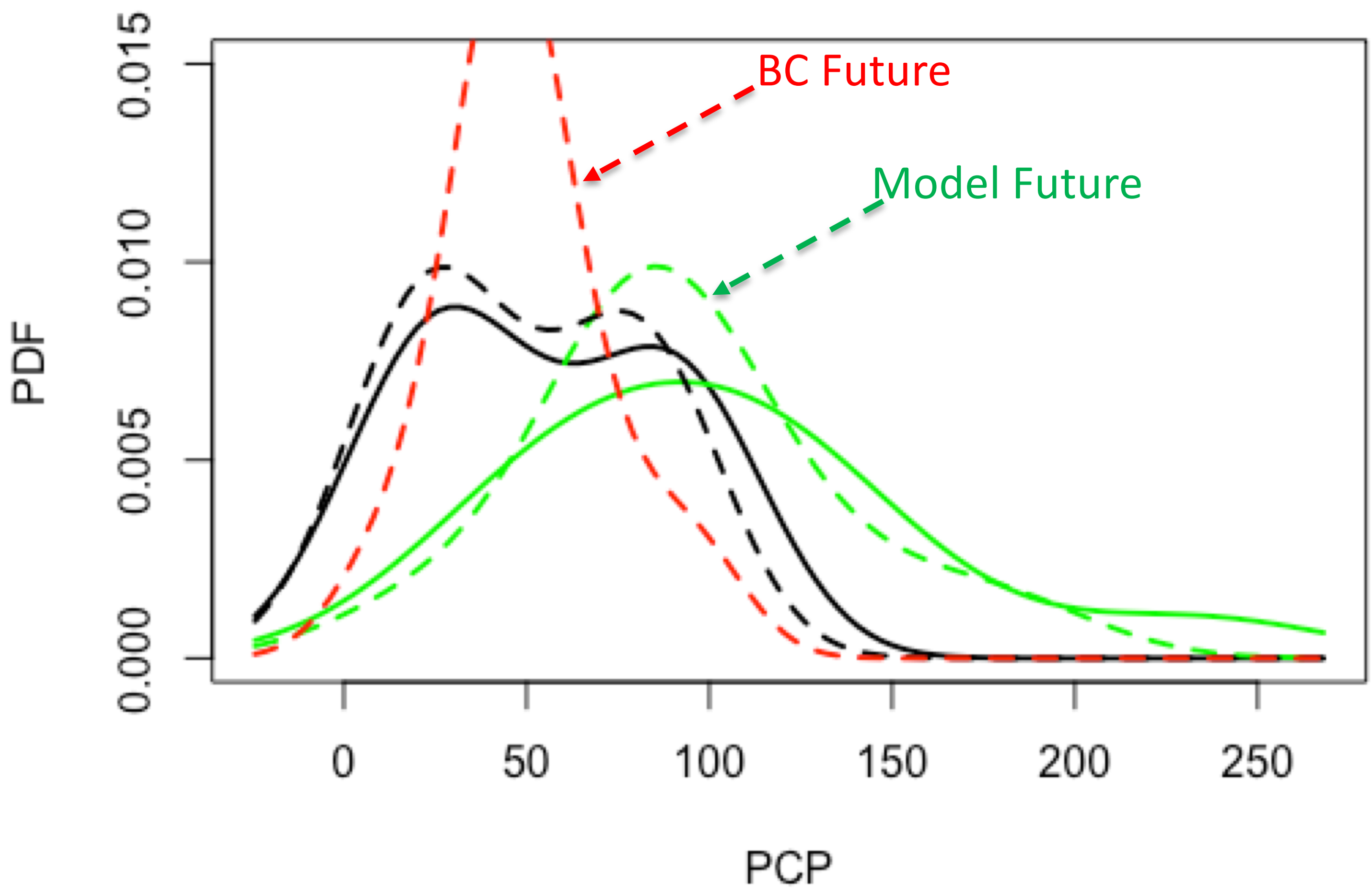
TEMP (T): $BC_T = \overline{T_O} - \overline{T_{Mc}}$ $\qquad T_f = T_{Mf} + BC_T$

Precipitation BC: Ratio between mean observations (*O*) and mean control model (*Mc*) multiplied by future model (*Mf*)

PREC (P): $BC_P = \frac{\overline{P_O}}{\overline{P_{Mc}}}$ $\qquad P_f = P_{Mf} * BC_P$ Direct method” in Maraun et al 2010)

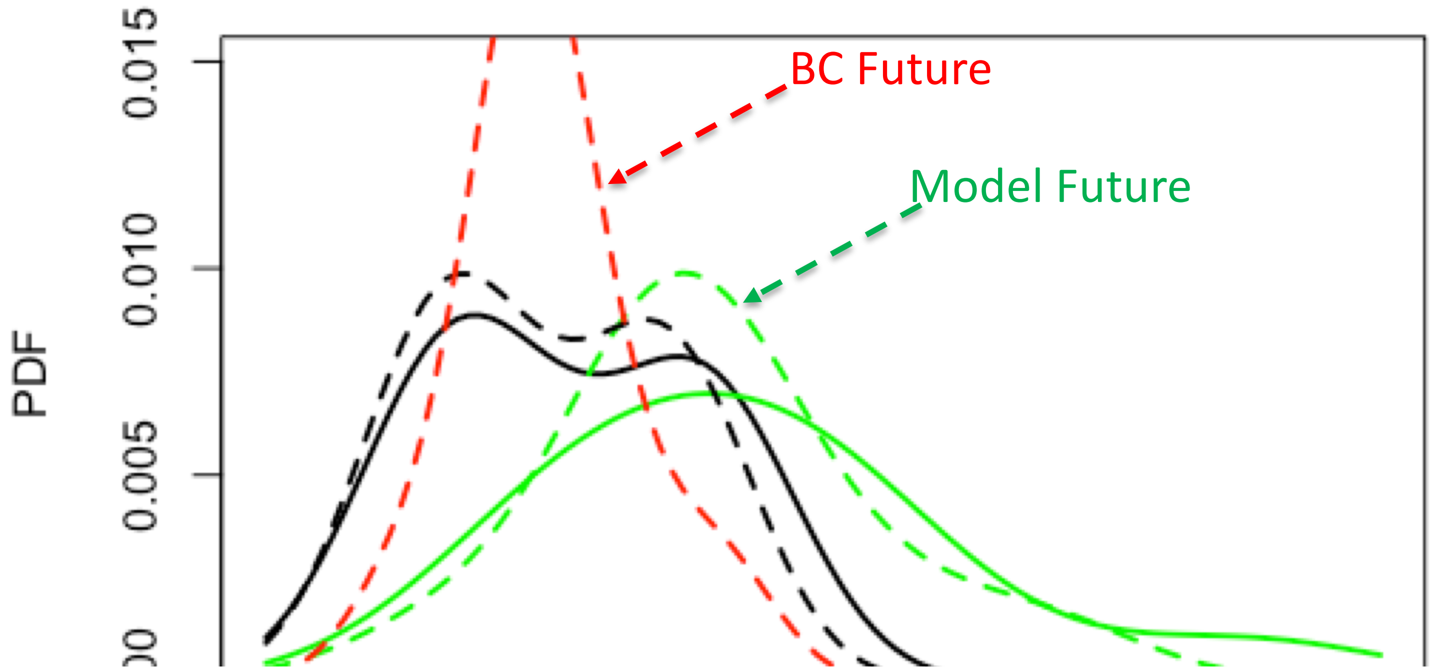
BC MOS example: Rescaling model output

* BC = .5, so multiply every future PCP value by .5



BC MOS example: Rescaling model output

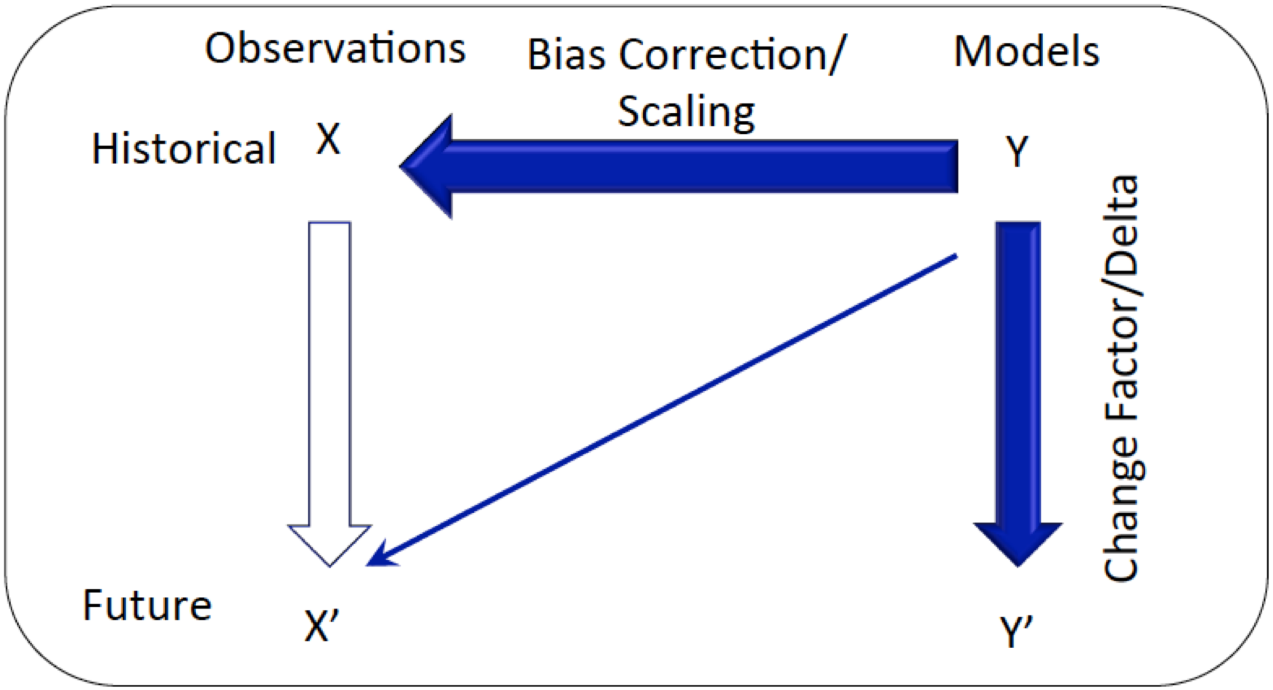
* BC = .5, so multiply every future PCP value by .5



- Advantage: variability in space and time is from climate model, not from present climate (Limitation in that future model variability can't be validated until future!)
- Limitation: All parts of model distribution are rescaled by single BC

MOS recalibration pathways don't yield same answer!

$$T_f \ from \ CF \neq T_f \ from \ BC!$$



BUT X' from X and Y (Bias Correction) ≠ X' from Y and Y' (Change Factor)!!

(Slide adapted from Mari Tye)

MOS "empirical CDF matching" (ECDF) is simple distribution mapping approach

Same as Temp BC & Precip BC, but on ***sorted* data at each point.**

Step 1. Sort T_O, T_{MC}, T_{MF}; P_O, P_{MC}, P_{MF}

Step 2. Calculate BC and apply to MF at each paired point

TEMP (T): $BC_T = \overline{T_O} - \overline{T_{Mc}}$ $\qquad T_f = T_{Mf} + BC_T$

PREC (P): $BC_P = \frac{\overline{P_O}}{\overline{P_{Mc}}}$ $\qquad P_f = P_{Mf} * BC_P$

MOS "empirical CDF matching" (ECDF) is simple distribution mapping approach

Same as Temp BC & Precip BC, but on ***sorted* data at each point.**

Step 1. Sort T_O, T_{MC}, T_{MF}; P_O, P_{MC}, P_{MF}

Step 2. Calculate BC and apply to MF at each paired point

TEMP (T): $BC_T = \overline{T_O} - \overline{T_{Mc}}$ $\quad T_f = T_{Mf} + BC_T$

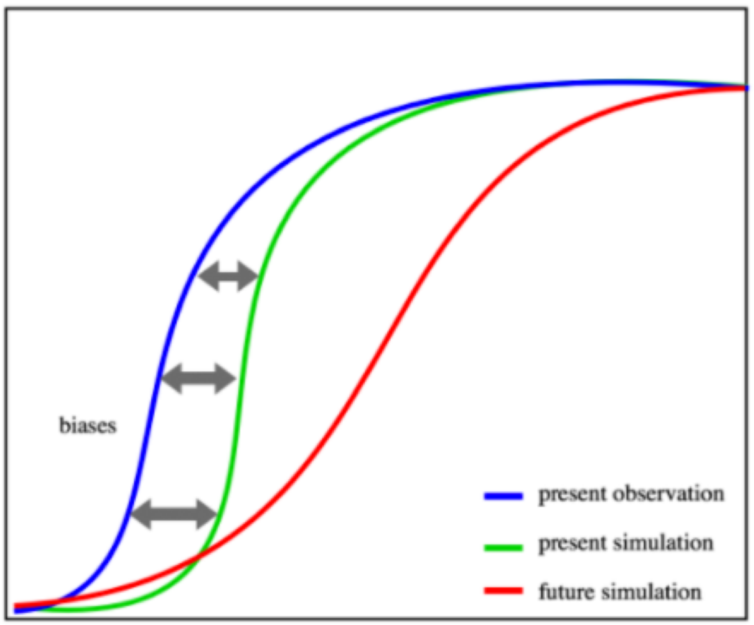
PREC (P): $BC_P = \frac{\overline{P_O}}{\overline{P_{Mc}}}$ $\quad P_f = P_{Mf} * BC_P$

Advantage: Correction done at each point (rather than single value)

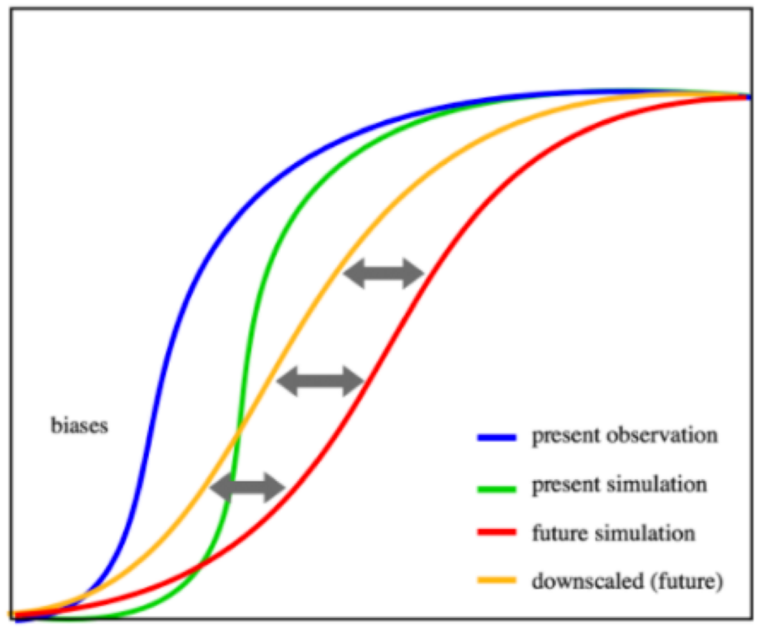
Limitation: Requires same number of points for obs, control, future

Distribution mapping at each quantile example:

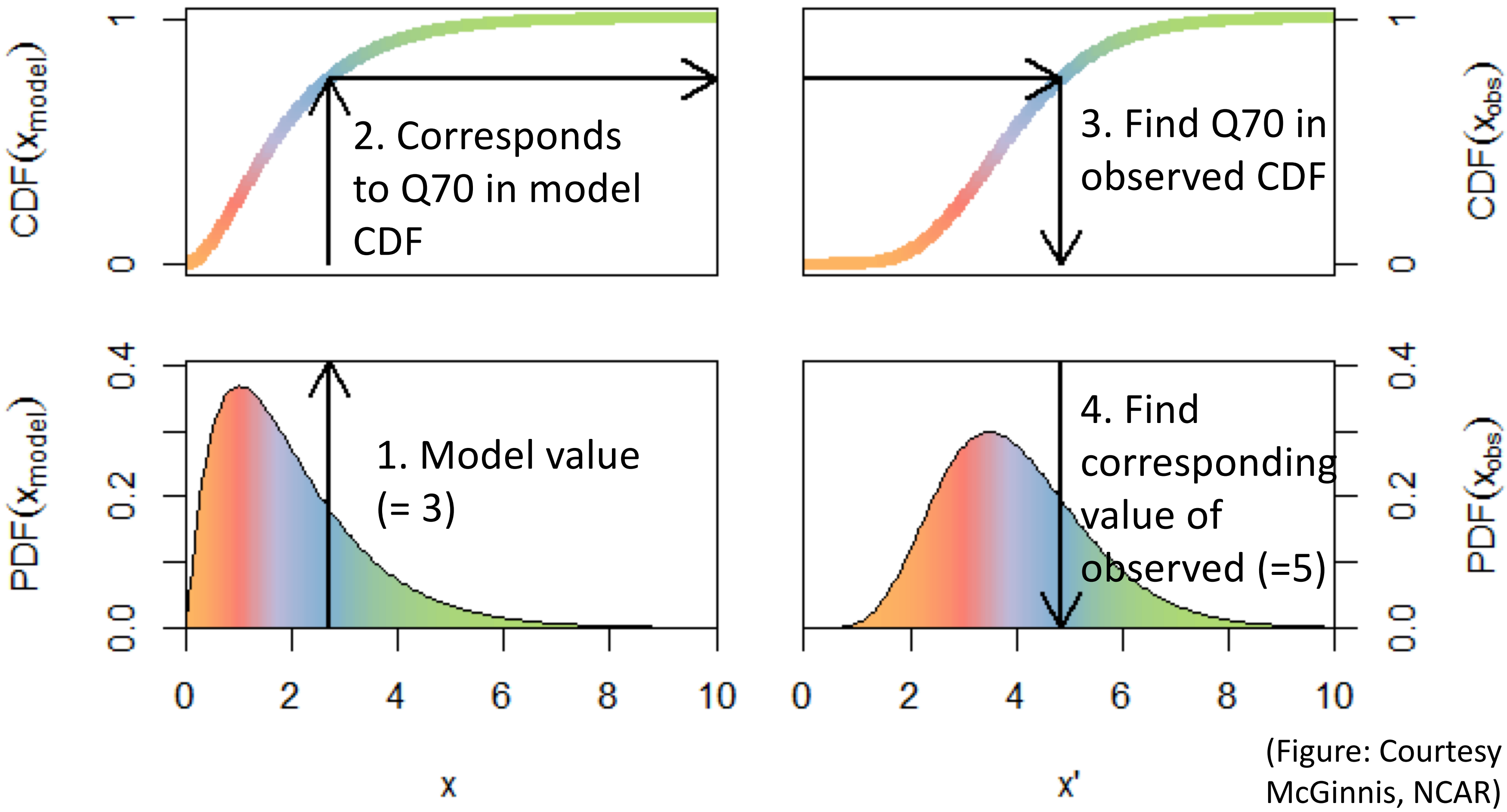
Biases are calculated for each quantile.



Calculated biases are added to future quantiles.



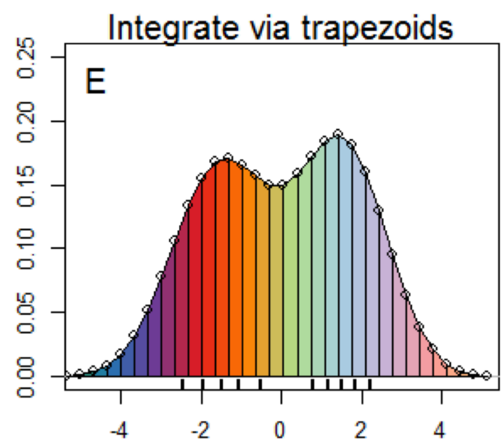
Distribution Mapping requires "transfer function" to go from value ➟ probability ➟ value



(Figure: Courtesy Seth McGinnis, NCAR)

Kernel Density Distribution Mapping is a nonparametric approach

- Transfer function equates CDFs by using trapezoid rule plus Kernel Density Estimation to get non-parametric estimate of CDFs
- Estimates CDF for unknown distributions



- Outperforms other distribution mapping approaches (McGinnis et al. 2017)

R package: https://github.com/sethmcg/climod

(Figure: Courtesy Seth McGinnis, NCAR)

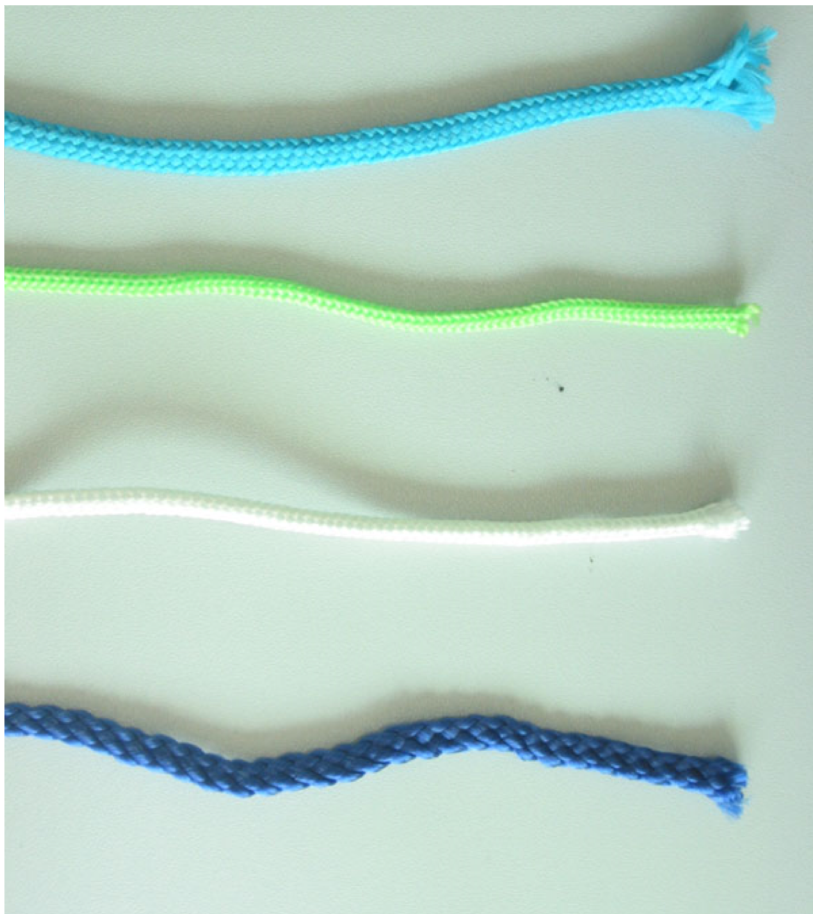
Summary of MOS statistical downscaling for extremes:

- MOS approaches are commonly used, but were not developed specifically for extremes.
    - Observation rescaling (CF) can change magnitude of distribution, but doesn’t change frequency or pattern.
    - Model rescaling (BC) captures distribution from models (but cannot be validated until future).
    - Assumes bias is constant in time.

Stochastic **weather generators** create synthetic sequences that preserve observed statistics

1960 2010

n.yr = 50 # 50 years



Observed Daily Weather Data

Sequence #1

Sequence #2

Sequence #N

- Generate many realizations (ensembles) of time series that can be used for impact models

2 Main Types of Weather Generators:

1. Based on parametric models like GLM (Richardson 1981)
2. Based on resampling (Rajagopalan & Lall 1999)

Weather generators usually have a precipitation generator at their core

Simple example:

1. Model occurrence with 1st order Markov chain:

WWDDDDDDWWWWWWWWDDDDDDWWDDDDDDDDWW

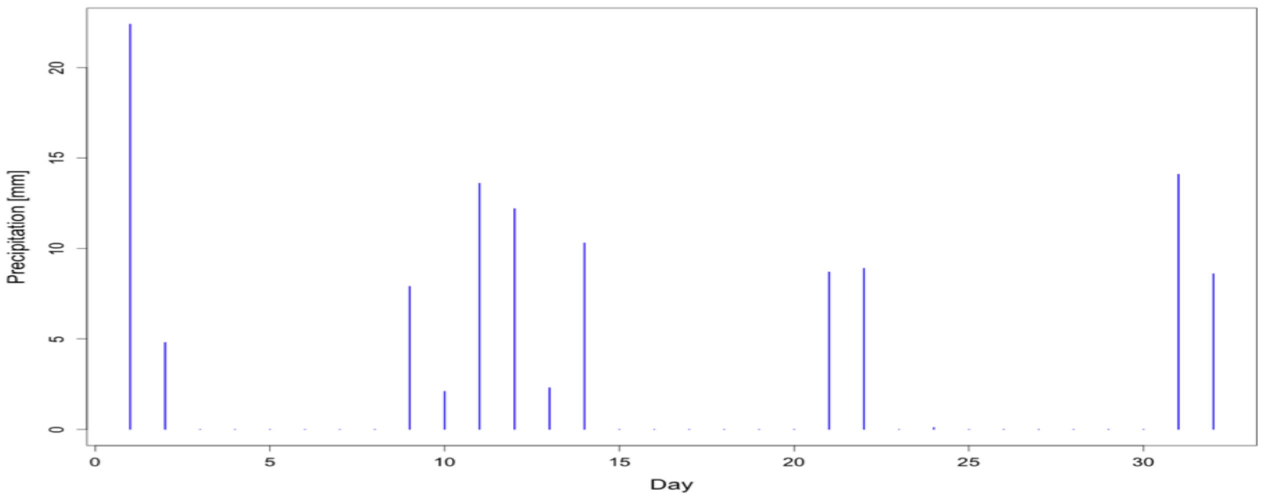
Weather generators usually have a precipitation generator at their core

Simple example:

1. Model occurrence with 1st order Markov chain:

WWDDDDDDWWWWWWDDDDDDWWDDDDDDDDWW

2. Model intensity with a distribution appropriate for precipitation (e.g., gamma)

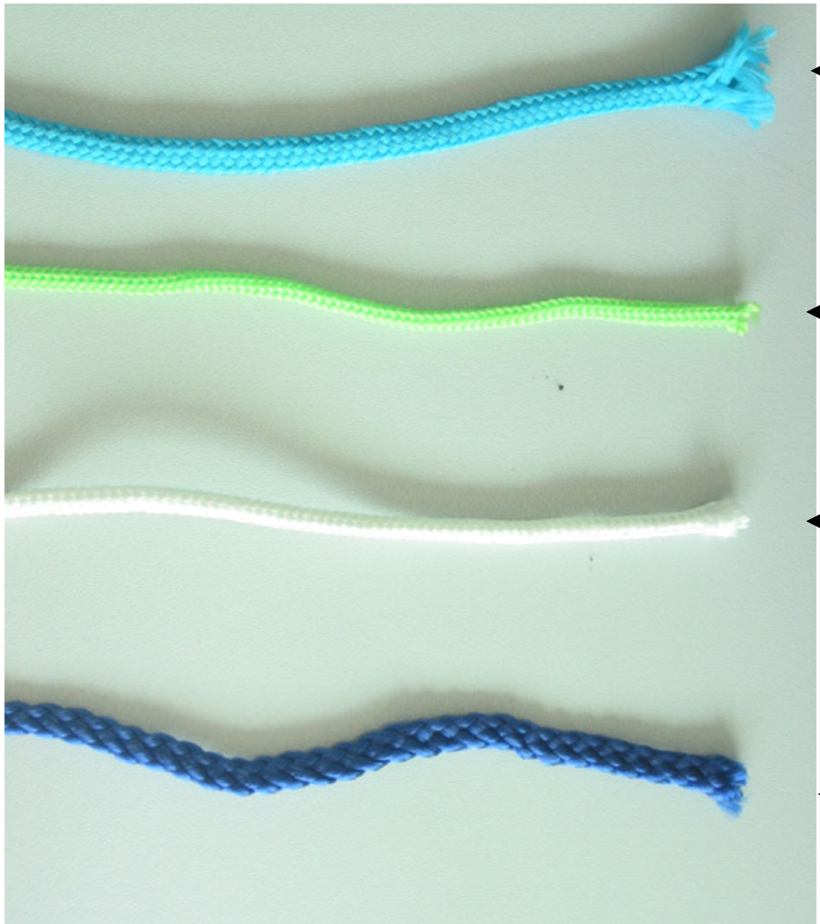


(See Maraun et al 2010 for details)

Weather generators can be used with MOS change factor time series

CF e.g., TMP + 3 deg,

PCP x 0.90



Change Factor (CF) Daily Weather

CF Sequence #1

CF Sequence #2

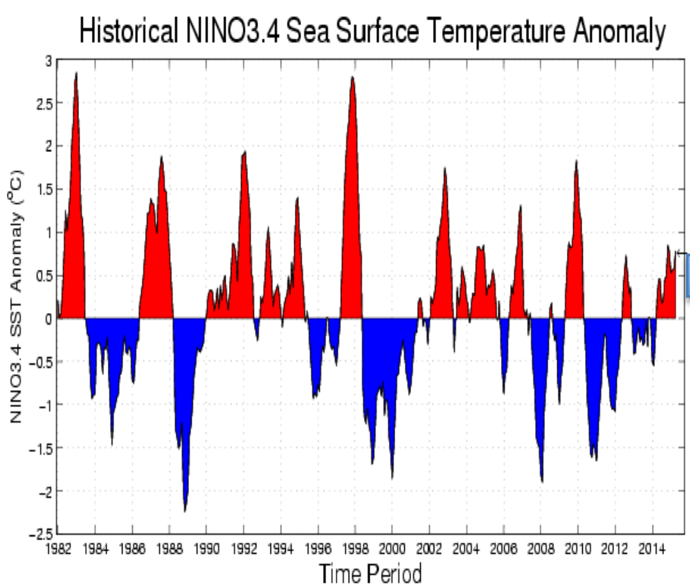
⋮

CF Sequence #N

- Generate many realizations (ensembles) of future (CF) time series that can be used for impact models, etc.

Weather generators can be conditioned on covariates, e.g., large-scale circulation patterns (PP downscaling)

E.g., Can use generalized linear models (GLM) to incorporate covariates such as ENSO or seasonal cycle (sin/cosine)



P(Rain) = f(ENSO, seasonal cycle…)
(the occurrence of rain depends on ENSO and/or other covariates)

I(Rain) = f(ENSO, seasonal cycle…)
(the intensity of rain depends on ENSO and/or other covariates)

(Katz and Furrer 2007)

Summary of weather generators for extremes:

- Weather generators were developed to provide additional realizations (ensembles) of observed time series.
  - Long synthetic series can be used to get at probabilities of extreme characteristics.
  - Precipitation generator is flexible (gamma is typical, but extreme distribution could be used – talk after break)

See Wilks 2010 for overview.

Two commonly applied statistical downscaling techniques:

1. Bias correction with spatial disaggregation (BCSD; Wood et al. 2004)

1. "Constructed analog"–based techniques

$$y = f(x)$$

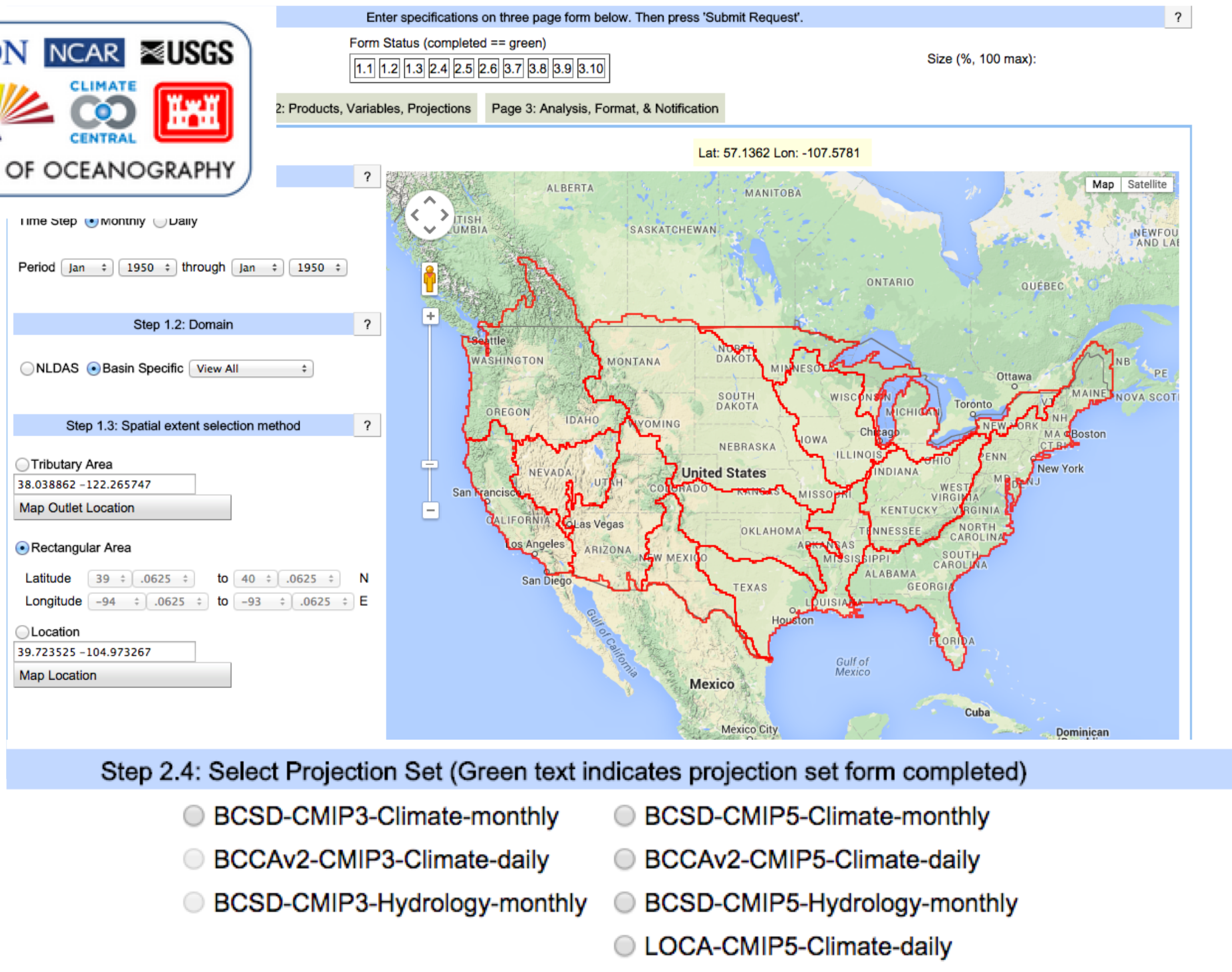
Predictant (Finer scale of climate variable z)

statistical model

Predictor (Coarse climate variable z)

Reclamation provides CMIP3 & CMIP5 Climate & Hydrology Projections for US using several downscaling methods

RECLAMATION NCAR USGS
Santa Clara University
CLIMATE CENTRAL
SCRIPPS INSTITUTION OF OCEANOGRAPHY



http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/

Intercomparison of statistical downscaling methods can reveal deficiencies

Table 2. Summary Statistics for Each Downscaling Method[a]

	Mean Annual Precipitation (mm/yr)	Interannual Variation (mm/yr)	50 yr Return Interval (mm/d)	Wet Day Fraction (0, 1 mm Threshold)	Wet Spell (Days)	Dry Spell (Days)
BCSDm	805	132	145	0.43, 0.26	2.4	7.7
BCSDd	850	139	109	0.88, 0.36	4.5	5.8
AR	817	161	149	0.34, 0.24	2.1	8.1
BCCA	579	101	85	0.79, 0.27	2.0	7.4
Observed	776	142	140	0.39, 0.24	2.1	7.6

- 4 methods in current climate
- Some methods have problems with wet days, wet/dry spells, and extreme events
- Most methods have problems with spatial scaling and interannual variability

(Guttman et al. 2014 WRR DOI: 10.1002/2014WR015559)

BCSD has been widely applied, but has limitations

Method has two steps (Wood et al. 2004):

1. Apply quantile mapping to P&T (MOS)
2. Add a spatial variability component.

BCSD has been widely applied, but has limitations

Method has two steps (Wood et al. 2004):

1. Apply quantile mapping to P&T (MOS)
2. Add a spatial variability component.

Limitations:

- Developed as a monthly technique, so limits changes in shape of daily distribution and does not consider daily GMC weather sequences
- Captures extremes and wet day fractions well, but is limited to rescaling current weather patterns. (Gutmann et al. 2014)

Constructed analog methods identify the *N* best matching analog days that reproduce a particular pattern

Disadvantage: Uses averaging, which tends to produce too much drizzle

Localized constructed analogs (LOCA) technique downscales point-by-point, and avoids the averaging issues of the other CA methods.

Advantages: Preserves GCM daily sequences, better at extremes, avoids drizzle.
Disadvantage: More computationally intensive than other CAs.

http://loca.ucsd.edu/

Conclusions

- Statistical and dynamical downscaling both have advantages and limitations

	Statistical downscaling	Dynamical downscaling
Advantages	• Comparatively cheap and computationally efficient • Can provide point-scale climatic variables from GCM-scale output • Can be used to derive variables not available from RCMs • Easily transferable to other regions • Based on standard and accepted statistical procedures • Able to directly incorporate observations into method	• Produces responses based on physically consistent processes • Produces finer resolution information from GCM-scale output that can resolve atmospheric processes on a smaller scale
Disadvantages	• Require long and reliable observed historical data series for calibration • Dependent upon choice of predictors • Non-stationarity in the predictor-predictand relationship • Climate system feedbacks not included • Dependent on GCM boundary forcing; affected by biases in underlying GCM • Domain size, climatic region and season affects downscaling skill	• Computationally intensive • Limited number of scenario ensembles available • Strongly dependent on GCM boundary forcing

(Fowler et al. 2007, Table 1)

Conclusions

- Statistical and dynamical downscaling both have advantages and limitations
- Statistical downscaling is complementary to dynamical downscaling, and even address some of the limitations.

Conclusions

- Statistical and dynamical downscaling both have advantages and limitations
- Statistical downscaling is complementary to dynamical downscaling, and even address some of the limitations.
- But, extremes were not the original focus of statistical downscaling, and there are many approaches, so it's important to understand the strengths/weaknesses of them with respect to extremes.

Questions?

References

Akaike, H. (1974), New look at statistical model identification, IEEE Trans. Autom. Control, 19, 716–723.

Bowman, A.W. and Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations.* Oxford University Press, Oxford.

Fowler H, Blenkinsop S, & Tebaldi C (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. International Journal of Climatology, 27, 1547–1578. doi:10.1002/joc

Gutmann, E., Pruitt, T., Clark, M., Brekke, L., Arnold, J., Raff, D., & Rasmussen, R. (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. Water Resources Research, 7167–7186. doi:10.1002/2014WR015559.Received

Katz RW, Furrer EM (2007). Generalized Linear Modeling Approach To Stochastic Weather Generators. Climate Research, 34, 129–144. doi:10.3354/cr034129

Helsel, D. R., and R. M. Hirsch (1995), Studies in Environmental Science, in Statistical Methods in Water Resources, vol. 49, 529 pp., Elsevier, Amsterdam.

Holland GJ, Done JM, Douglas R, Ge M. Global Tropical Cyclone Damage Potential. Natural Hazards Review (in review)

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. Reviews of Geophysics, 48(3), 1–34. doi:10.1029/2009RG000314

Maraun, D., Widmann, M., Gutiérrez, J. M., Kotlarski, S., Chandler, R. E., Hertig, E., … Wilcke, R. A. I. (2015). VALUE : A framework to validate downscaling approaches for climate change studies. Earth's Future, 3, 1–14. doi:10.1002/2014EF000259

McCullagh, P., and J. A. Nelder (1989), Generalized Linear Models, Chapman and Hall, London.

References (cont).

McGinnis S, Nychka D, Mearns LO (2017). A New Distribution Mapping Technique for Climate Model Bias Correction

Miller K and Yates D (2006). Climate change and water resources: A primer for municipal water providers. AwwaRF, Denver.

Montgomery DC, Runger GC (2011). Applied Statistics and Probability for Engineers. Fifth Edition. John Wiley and Sons, Inc., Hoboken, NJ.

Pierce DW, Cayan DR, Thrasher BL (2014). Statistical Downscaling Using Localized Constructed Analogs (LOCA). Journal of Hydrometeorology, 15(6), 2558–2585. http://doi.org/10.1175/JHM-D-14-0082.1

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

Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour Res 17: 182–190

Towler E, PaiMazumder D, Holland G (2016) A framework for investigating large-scale patterns as an alternative to precipitation for downscaling to local drought. Clim Dyn doi:10.1007/s00382-016-3116-5

Wilks DS (2010). Use of stochastic weathergenerators for precipitationdownscaling, WIREs Climate Change, DOI: 10.1002/wcc.85

Wilks DS (2011). Statistical methods in the atmospheric sciences 3rd ed., Waltham, MA: Academic Press Inc.

Wood AW et al (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62:189-216