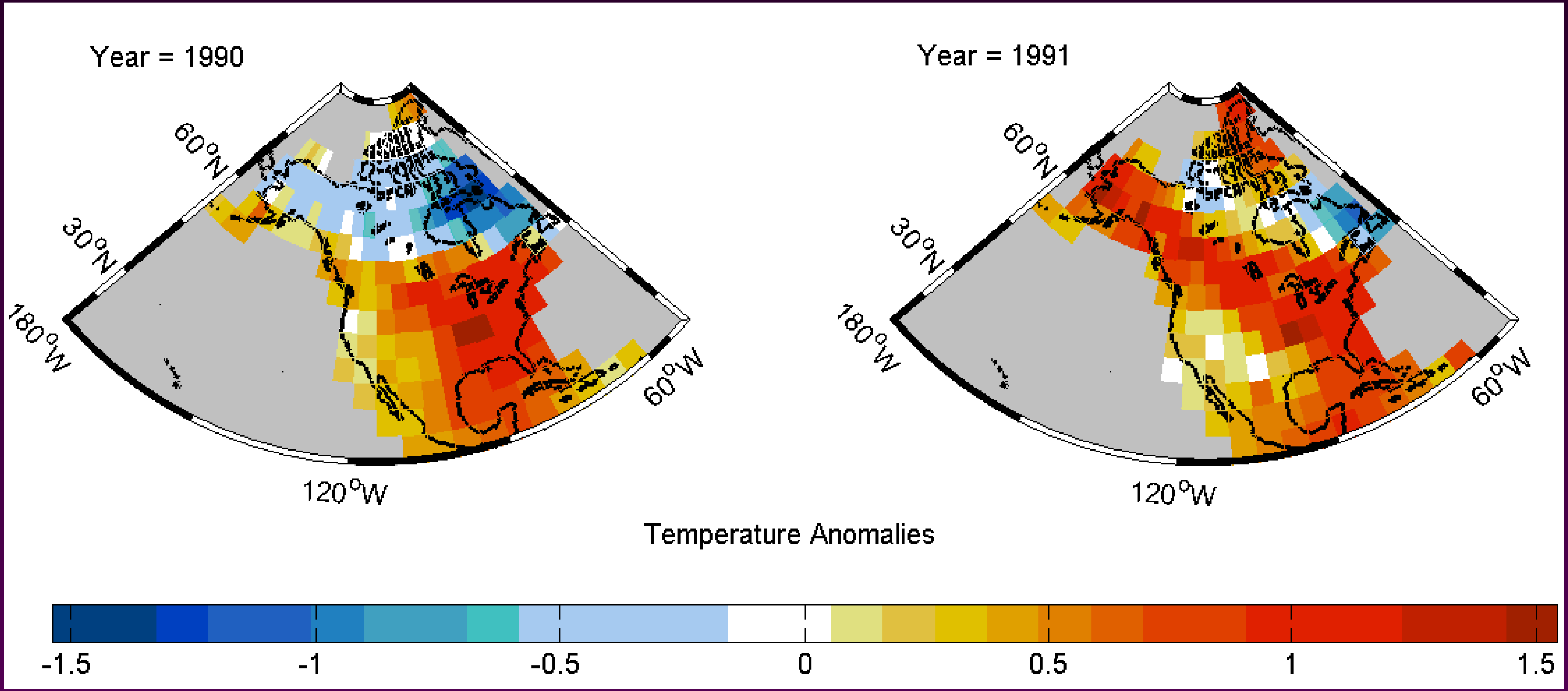
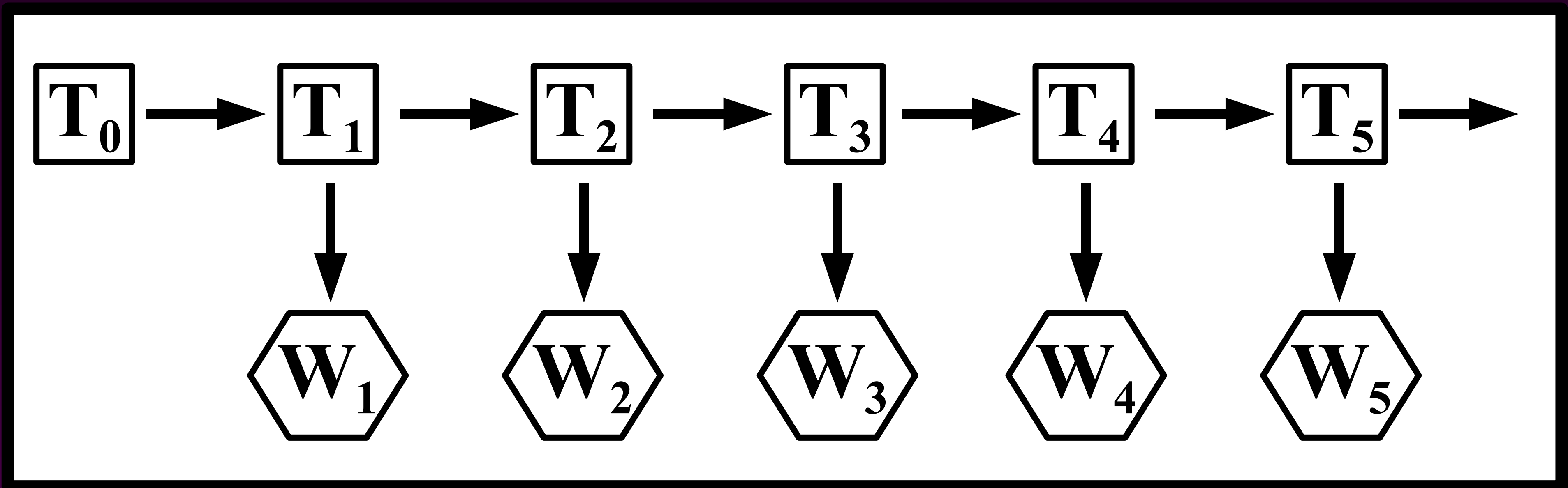
Temperature: a field with both spatial and temporal covariance



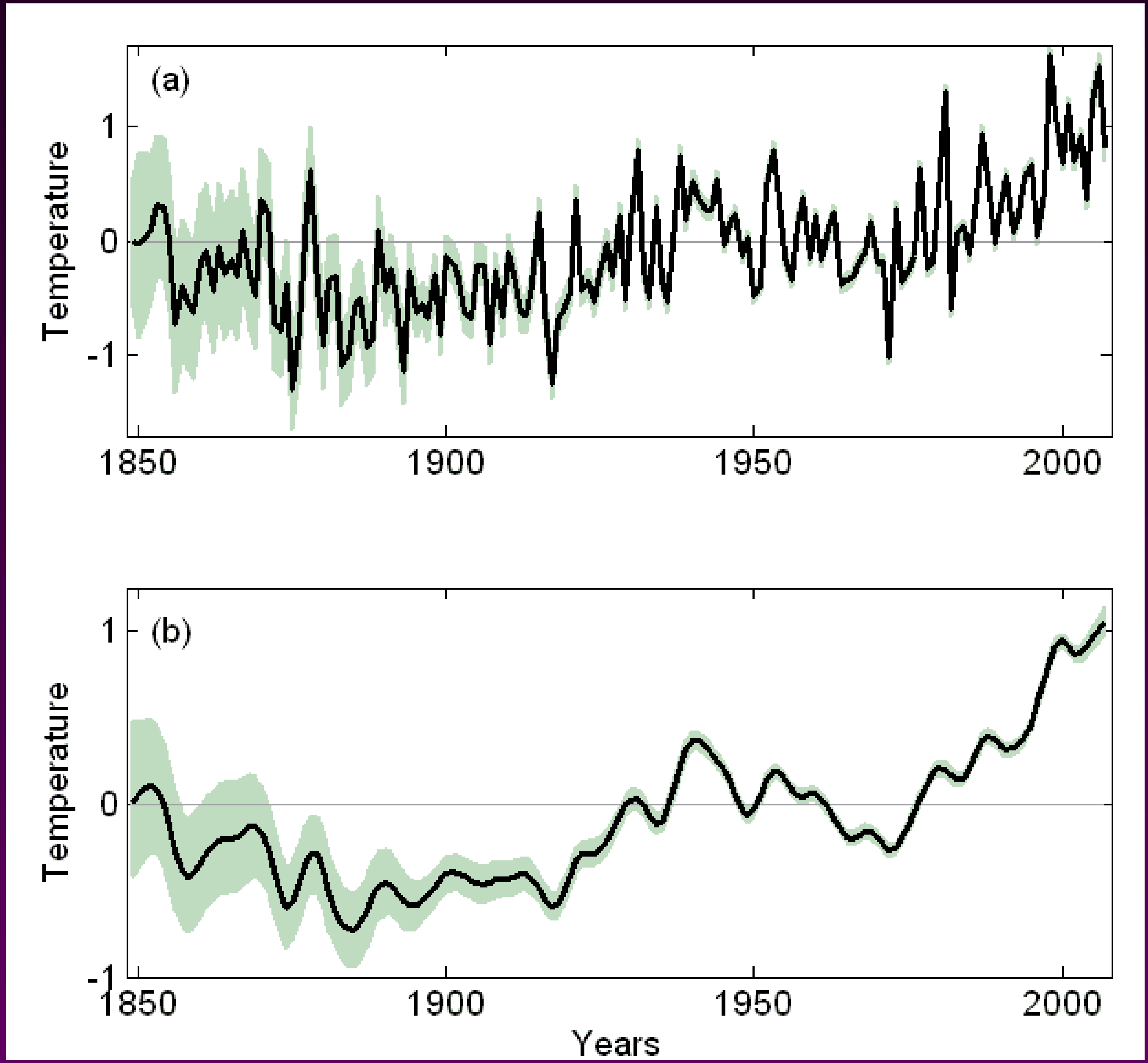
Different types of imperfect observations.

Specify parametric forms

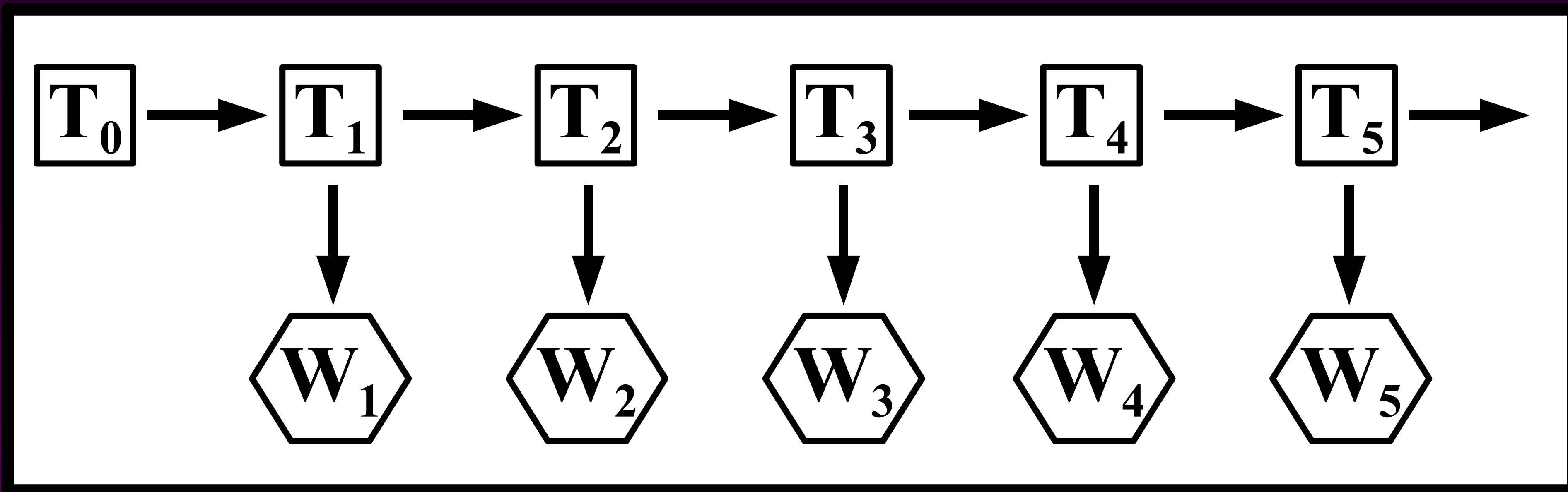


1. Structure of the spatial covariance.
2. An equation for the temperature evolution.
3. 'Observation equations' for instrumental and proxy data sets.

Block Average



BARSAT models the evolution of the field, and specifies 'observation equations'



The **T**s are field values and the **W**s observations.

The arrows denote conditional dependencies:

$$P\,(\,\mathbf{T}_2 \mid \cdot\,) = f\,(\mathbf{T}_1, \mathbf{T}_3, \mathbf{W}_2, \text{parameters})$$

BARSAT

A Bayesian Algorithm for Reconstructing Spatially Arrayed Temperatures

Martin Tingley and Peter Huybers

Outline

1. A few words about RegEM.

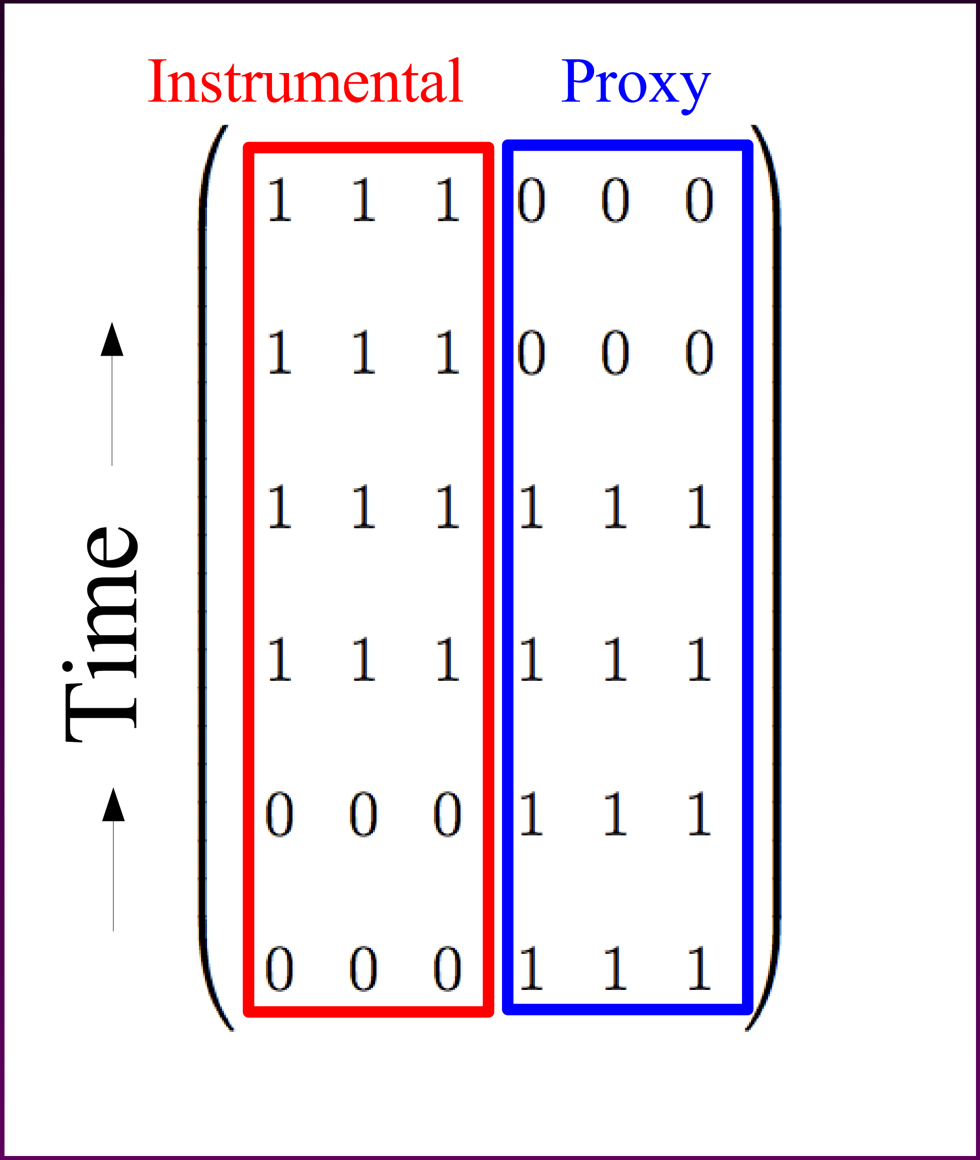
2. A few words about temperature fields.

3. BARSAT: the main ideas.

4. A few equations.

5. An example and a short movie.

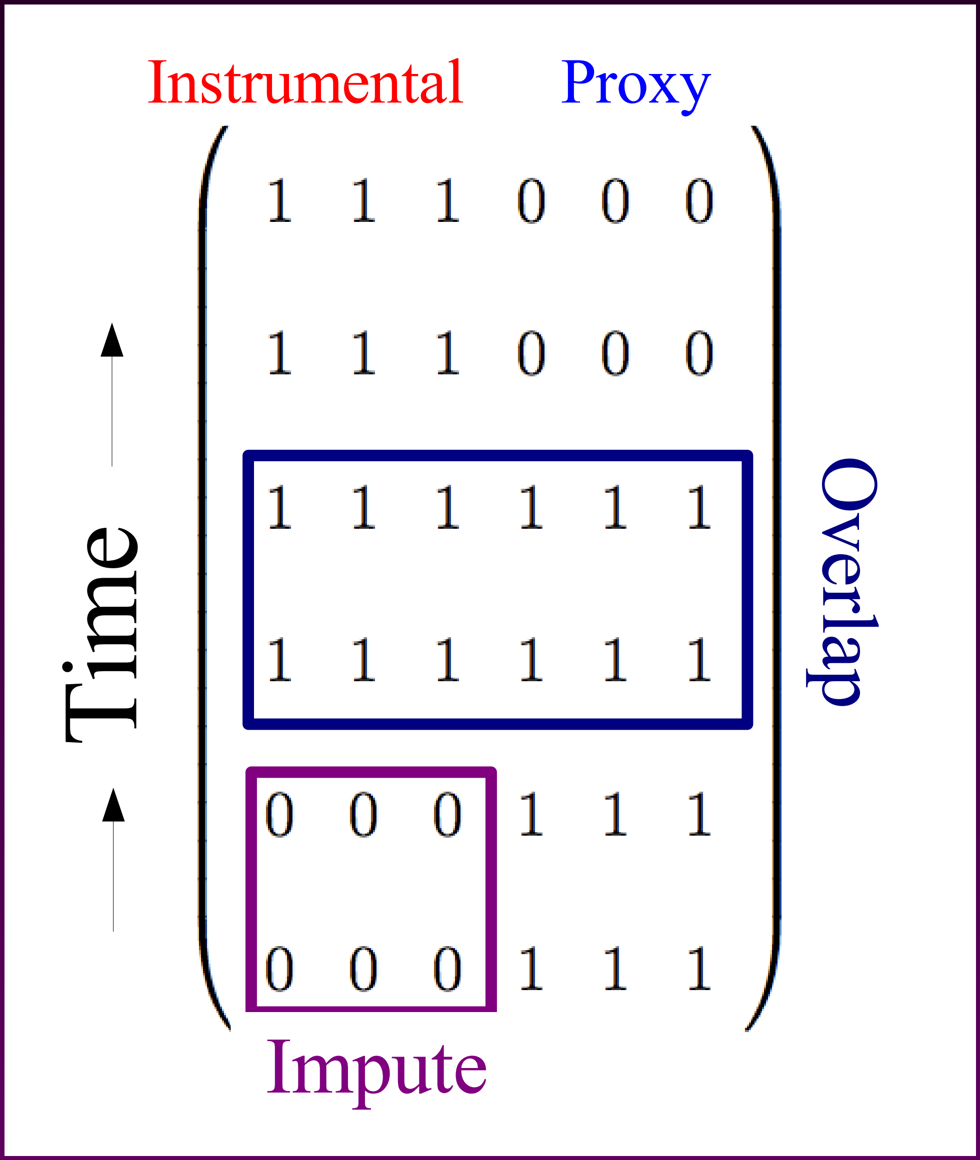
RegEM – A missing data approach



Two types of incomplete data time series.

1 = Observed
0 = Missing

RegEM – A missing data approach



- Assume that the complete data for each year is a draw from a MVN.

- Impute the missing instrumental values, using the information contained in the overlap.

Improving on RegEM

- An empirical estimate of the data covariance matrix.
  - The locations of the time series are not used.
  - Cannot impute where there are no data time series.

- Parametric form for the spatial covariance?

- What about temporal autocorrelation?

- What about the small, but non zero, uncertainty in the instrumental observations?

BARSAT Equations

Temperature Evolution:

$$\vec{T}_t - \mu\vec{1} = \alpha\left(\vec{T}_{t-1} - \mu\vec{1}\right) + \vec{\epsilon}_t$$

Spatial Covariance of Innovations:

$$\mathbf{\Sigma}_{ij} = \sigma^2 e^{-\phi|x_i - x_j|}$$

Instrumental Observation Equation:

$$\vec{W}_{I,t} = \vec{T}_{I,t} + \vec{e}_{I,t}$$

$$\vec{e}_{I,t}\left(t\right) \sim N\left(0, \tau_I^2 \mathbf{I}_{N_{I,t}}\right)$$

BARSAT Equations

Assume a statistically linear relationship between proxy values and the true temperature values:

$$\vec{T}_{P,t} = \frac{1}{\beta_1} \cdot \vec{W}_{P,t} + \beta_o \vec{1} + \vec{e}_{P,t}$$

Which leads to the proxy observation equation:

$$\vec{W}_{P,t} = \left(\vec{T}_{P,t} + \vec{e}_{P,t} - \beta_o \vec{1}\right) \beta_1$$

$$\vec{e}_{P,t} \sim N\left(0, \tau_P^2 \mathbf{I}_{N_{P,t}}\right)$$

"Discrete time, continuous state, Hidden Markov Model"

BARSAT Parameters

- $\vec{T}_t \rightarrow$ Vector of true temperatures at a number of spatial location.
- $\alpha \rightarrow$ AR(1) coefficient in the temperature evolution equation.
- $\sigma^2 \rightarrow$ Partial sill of spatial innovations.
- $\phi \rightarrow$ Inverse range of spatial innovations.
- $\tau_I^2 \rightarrow$ Error variance of instrumental observations.
- $\tau_P^2 \rightarrow$ Error variance of proxy observations.
- $\mu \rightarrow$ Constant in the equation for the mean of $\vec{T}$.
- $\beta_o \rightarrow$ Constant term in the proxy observation equation.
- $\beta_1 \rightarrow$ Scaling factor in the proxy observation equation.
- $\vec{\theta} \rightarrow$ Vector composed of the 8 scalar parameters.

Priors and Conditional Posteriors

	Prior Form	Conditional Posterior
α	Uniform	Truncated Normal
σ^2	Inverse-Gamma	Inverse-Gamma
ϕ	Log-Normal	Non Standard
τ_I^2	Inverse-Gamma	Inverse-Gamma
τ_P^2	Inverse-Gamma	Inverse-Gamma
μ	Normal	Normal
β_o	Normal	Normal
β_1	Normal	Non Standard
$\vec{T}_0$	MV Normal	MV Normal
$\vec{T}_{k=1\ldots\kappa}$	-	MV Normal

Our strategy is to use low information, but proper, priors, and show that the data is the major contributor to the posterior.

BARSAT Equations

Probability of the data given unknowns can be decomposed:

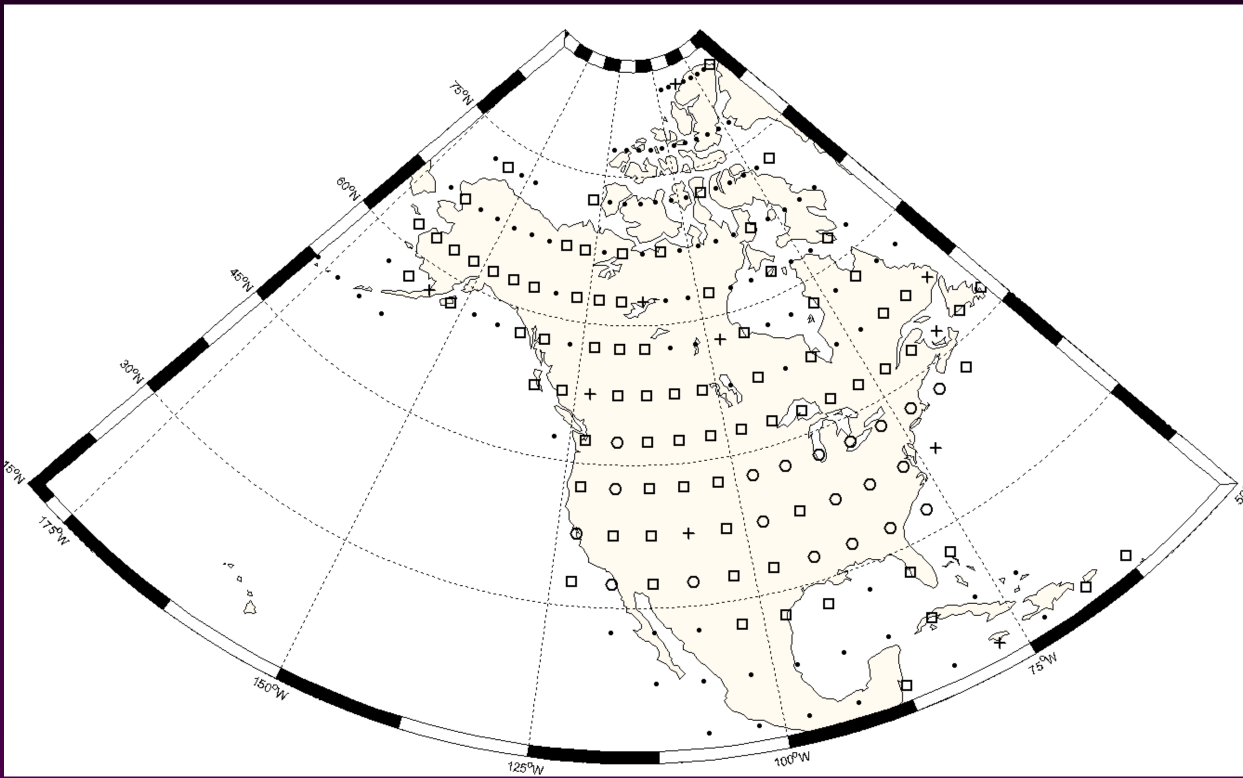
$$P\left(\vec{W}_1,\ldots,\vec{W}_\kappa|\vec{T}_1,\ldots,\vec{T}_\kappa,\vec{\theta}\right)=\prod_{k=1}^{\kappa}P\left(\vec{W}_k|\vec{T}_k,\vec{\theta}\right)$$

Solving for the posterior using Bayes' rule gives:

$$P\left(\vec{T}_1,\ldots,\vec{T}_\kappa,\vec{\theta}|\vec{W}_1,\ldots,\vec{W}_\kappa\right)\propto P\left(\vec{T}_0\right)\cdot P\left(\vec{\theta}\right)$$
$$\cdot\prod_{k=1}^{\kappa}P\left(\vec{W}_k|\vec{T}_k,\tau_I^2,\tau_P^2,\mu,\beta_o,\beta_1\right)P\left(\vec{T}_k|\vec{T}_{k-1},\sigma^2,\phi,\alpha\right)$$

A Gibbs sampler with two Metropolis-Hasting steps is used to draw from this monstrosity . . . .

A Quick Example Using CRU data



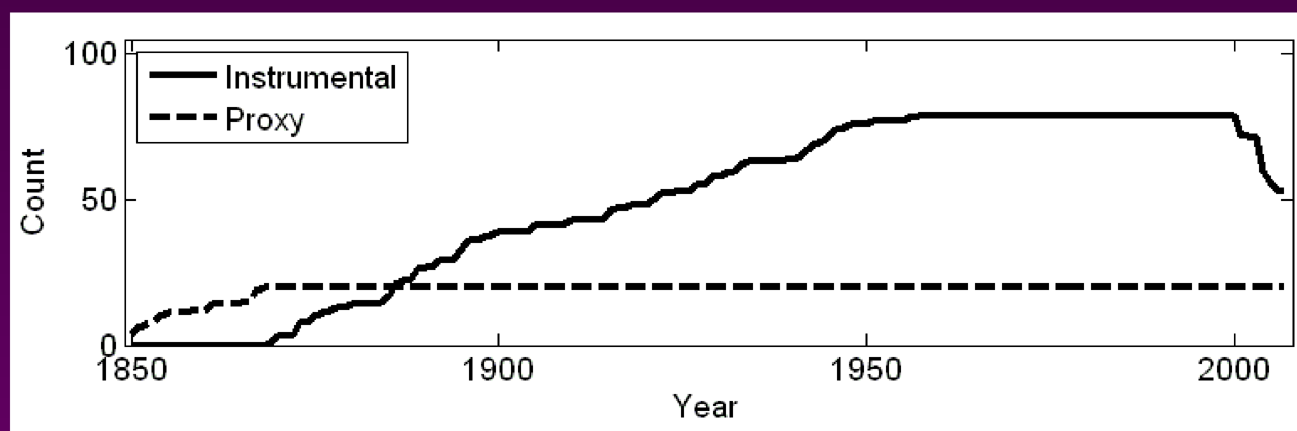
■ Instrumental

● Proxy

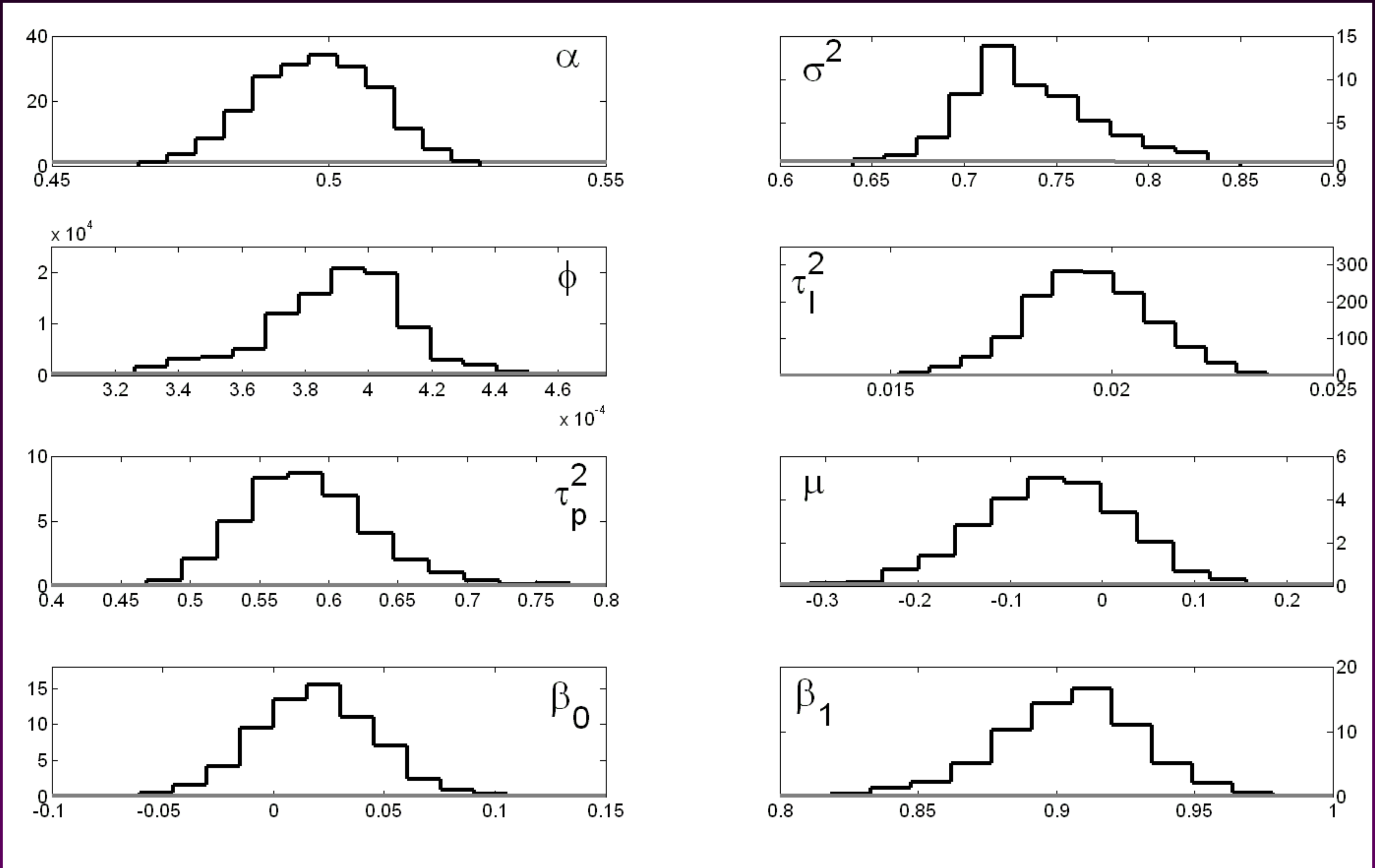
\+ Withheld

· No data

Proxies: constructed by adding WN with SNR of one to select CRU time series.



Priors and Posteriors

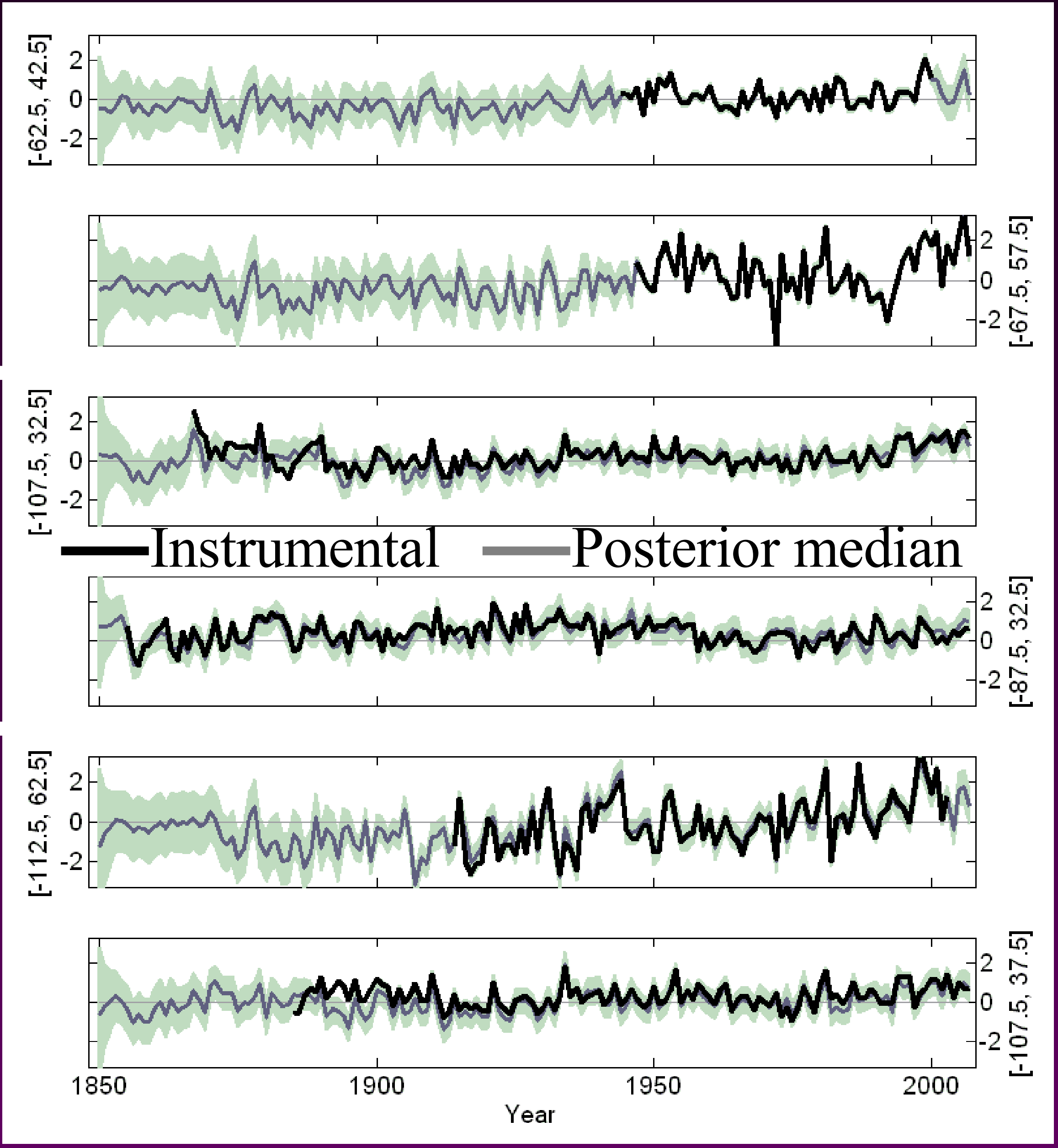


Time series at several locations

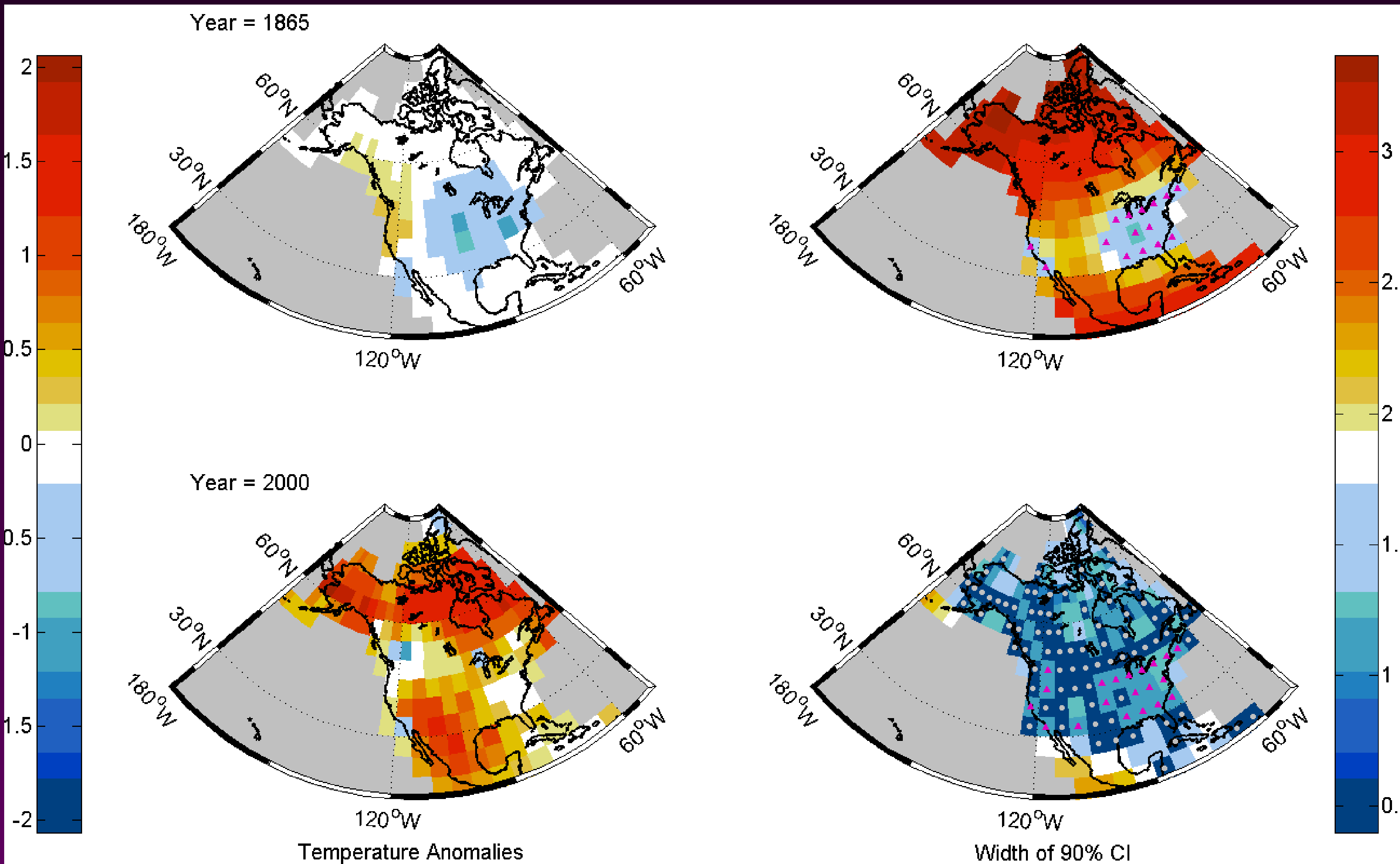
Instrumental data

Proxy data

No data



Temperature field for two years



Observational errors and innovations

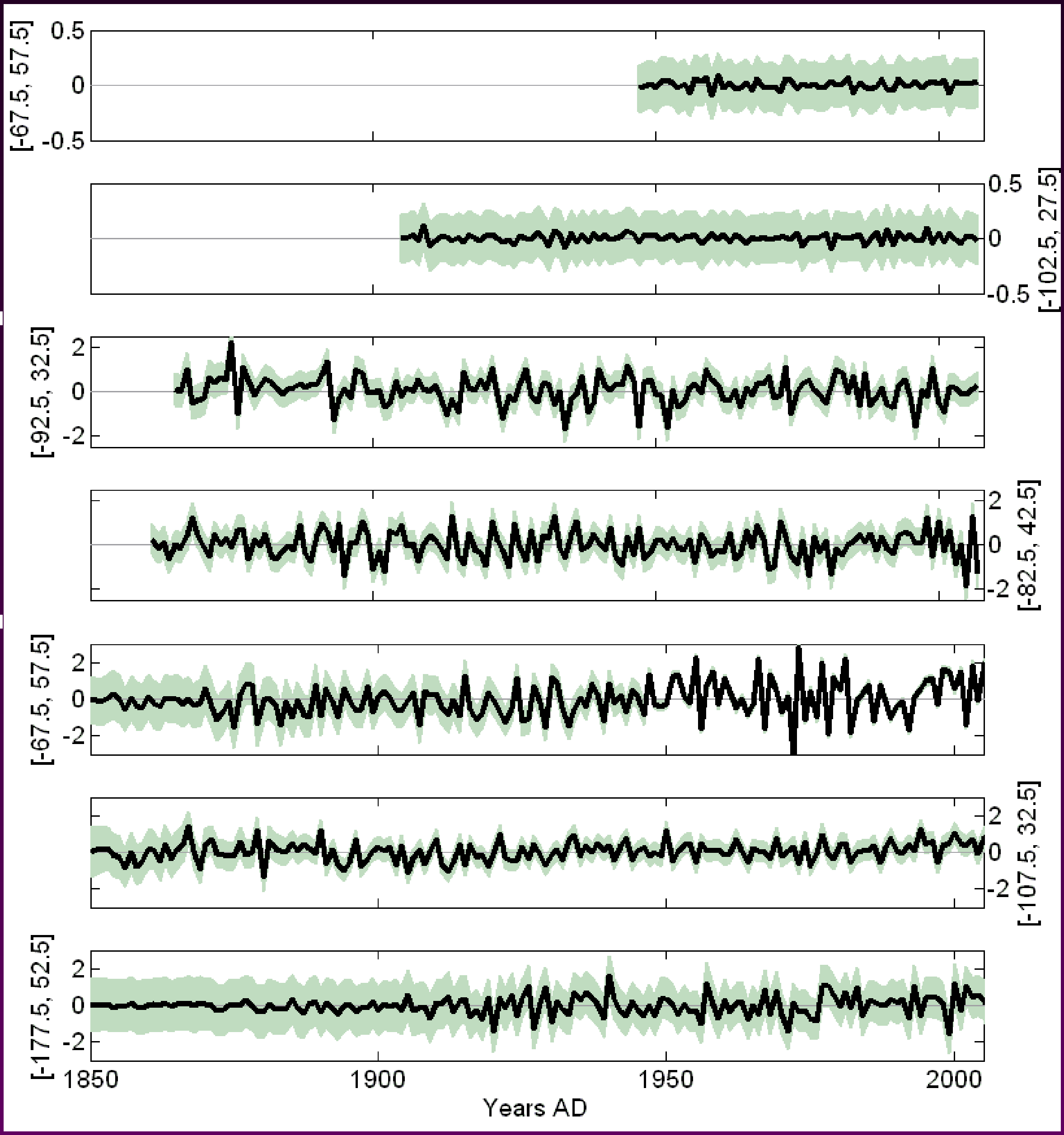
Instrumental observational errors

Proxy observational errors

Innovations:
Proxy data

Instrumental data

No data



{ *fin* }