Compression and Conditional Emulation of Climate Model Output

Joe Guinness
NC State University Department of Statistics

Dorit Hammerling
National Center for Atmospheric Research
Motivation

Computational power rightly receives considerable attention

Yellowstone Supercomputer, National Center for Atmospheric Research
Designed for climate and weather simulations

Planned Cheyenne Supercomputer has 2.5 times more computing power
This Project - Data Storage

NCAR Supercomputing Budgets:
- Yellowstone: 20% for storage hardware
- Cheyenne (under construction): 50% for hardware storage

CMIP5 Archive is 3.3 Petabytes of data
CMIP6 Archive > 10 Petabytes! (expected)

Climate models run with MANY different scenarios. Various
- Initial Conditions
- Model Settings (e.g. CO₂ emission projections)
- Model itself, combination of different versions of different components (oceans/atmosphere/land)

Data storage a limiting factor for climate science
This Project - Data Storage

NCAR Supercomputing Budgets:
- Yellowstone: 20% for storage hardware
- Cheyenne (under construction): 50% for hardware storage

CMIP5 Archive is 3.3 Petabytes of data
CMIP6 Archive > 10 Petabytes! (expected)
This Project - Data Storage

NCAR Supercomputing Budgets:
- Yellowstone: 20% for storage hardware
- Cheyenne (under construction): 50% for hardware storage

CMIP5 Archive is 3.3 Petabytes of data
CMIP6 Archive > 10 Petabytes! (expected)

Climate models run with MANY different scenarios. Various
- Initial Conditions
- Model Settings (e.g. CO$_2$ emission projections)
- Model itself, combination of different versions of different components (oceans/atmosphere/land)
This Project - Data Storage

NCAR Supercomputing Budgets:

- Yellowstone: 20% for storage hardware
- Cheyenne (under construction): 50% for hardware storage

CMIP5 Archive is 3.3 Petabytes of data
CMIP6 Archive > 10 Petabytes! (expected)

Climate models run with MANY different scenarios. Various

- Initial Conditions
- Model Settings (e.g. CO$_2$ emission projections)
- Model itself, combination of different versions of different components (oceans/atmosphere/land)

Data storage a limiting factor for climate science
Compression

Compression: representing a set of numbers by a smaller set of numbers
Compression

Compression: representing a set of numbers by a smaller set of numbers

2 Step Process:

- Compression: $X \rightarrow C$
- Decompression: $C \rightarrow \tilde{X}$
Compression

Compression: representing a set of numbers by a smaller set of numbers

2 Step Process:
- Compression: \( X \rightarrow C \)
- Decompression: \( C \rightarrow \tilde{X} \)

Types of Compression:
- Lossless: \( \tilde{X} = X \)
- Lossy: \( \tilde{X} \approx X \)
Compression

Compression: representing a set of numbers by a smaller set of numbers

2 Step Process:
  ▶ Compression: $X \rightarrow C$
  ▶ Decompression: $C \rightarrow \tilde{X}$

Types of Compression:
  ▶ Lossless: $\tilde{X} = X$
  ▶ Lossy: $\tilde{X} \approx X$

Lossless compression generally not efficient for scientific data
Compression

Compression: representing a set of numbers by a smaller set of numbers

2 Step Process:
- Compression: $X \rightarrow C$
- Decompression: $C \rightarrow \tilde{X}$

Types of Compression:
- Lossless: $\tilde{X} = X$
- Lossy: $\tilde{X} \approx X$

Lossless compression generally not efficient for scientific data

Lossy compression well studied for audio, video, and image data, but NOT for climate data
Study of Compression at NCAR

Several “black box” lossy compression algorithms have been applied to climate data

NCAR climate scientists study effect of compression on
- Distributions of variables
- Oversmoothing in space and time
- Statistical inferences about causal relationships among variables

Ongoing effort led by Allison Baker
Statistical Compression

We pursue lossy statistical compression algorithms

- store a set of summary statistics + conditional distribution of data
- decompress using conditional distributions (expectations, simulations)
Statistical Compression

We pursue lossy statistical compression algorithms

▶ store a set of *summary statistics + conditional distribution of data*
▶ decompress using conditional distributions (expectations, simulations)

Compression need not be lightning fast, but must be feasible

▶ NCAR has a supercomputer, after all
Statistical Compression

We pursue lossy statistical compression algorithms
▶ store a set of summary statistics + conditional distribution of data
▶ decompress using conditional distributions (expectations, simulations)

Compression need not be lightning fast, but must be feasible
▶ NCAR has a supercomputer, after all

Decompression should be fast
▶ Target user has a laptop
We pursue lossy statistical compression algorithms

- store a set of summary statistics + conditional distribution of data
- decompress using conditional distributions (expectations, simulations)

Compression need not be lightning fast, but must be feasible

- NCAR has a supercomputer, after all

Decompression should be fast

- Target user has a laptop

Objectives: goodness of fit of conditional model, small prediction errors, fast decompression
The Data

One year of one variable (surface temperature) is 19.97 million points!

\[ T = 365 \text{ days}, \quad n = 54,720 \text{ pixels} \]

[ show gif of temperature data ]
Will it float?

Why did we think this would work well? Data are strongly dependent!
Will it float?

Why did we think this would work well? Data are strongly dependent!
Will it float?

Why did we think this would work well? Data are strongly dependent!
Discrete Fourier Transform

Background: Discrete Fourier Transform of $Y(x, t)$

$$Y(\omega; x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y(x, t) \exp(-i\omega t)$$
Discrete Fourier Transform

Background: Discrete Fourier Transform of $Y(x, t)$

$$\mathcal{Y}(\omega; x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y(x, t) \exp(-i\omega t)$$

Simply a projection of the time series onto trigonometric functions
Discrete Fourier Transform

Background: Discrete Fourier Transform of $Y(x, t)$

$$
Y(\omega; x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y(x, t) \exp(-i\omega t)
$$

Simply a projection of the time series onto trigonometric functions

Stationary time series $\implies$ Approximately decorrelating transformation
Discrete Fourier Transform

Background: Discrete Fourier Transform of $Y(x, t)$

$$\mathcal{Y}(\omega; x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y(x, t) \exp(-i\omega t)$$

Simply a projection of the time series onto trigonometric functions

Stationary time series $\implies$ Approximately decorrelating transformation

$T^{-1/2}\mathcal{Y}(0; x)$ is real - mean over time

$T^{-1/2}\mathcal{Y}(2\pi/365; x)$ is complex - seasonal cycle
Discrete Fourier Transform

Background: Discrete Fourier Transform of $Y(x, t)$

$$\mathcal{Y}(\omega; x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y(x, t) \exp(-i\omega t)$$

Simply a projection of the time series onto trigonometric functions
Stationary time series $\implies$ Approximately decorrelating transformation

$T^{-1/2} \mathcal{Y}(0; x)$ is real - mean over time

$T^{-1/2} \mathcal{Y}(2\pi/365; x)$ is complex - seasonal cycle

Periodogram: $\hat{f}(\omega; x) = |\mathcal{Y}(\omega; x)|^2$

$\triangleright$ variation in $Y(x, t)$ explained by $\{e^{i\omega t}, e^{-i\omega t}\}$
Discrete Fourier Transform

Background: Discrete Fourier Transform of $Y(x, t)$

$$Y(\omega; x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y(x, t) \exp(-i\omega t)$$

Simply a projection of the time series onto trigonometric functions

Stationary time series $\implies$ Approximately decorrelating transformation

$T^{-1/2}Y(0; x)$ is real - mean over time

$T^{-1/2}Y(2\pi/365; x)$ is complex - seasonal cycle

Periodogram: $\hat{f}(\omega; x) = |Y(\omega; x)|^2$

$\triangleright$ variation in $Y(x, t)$ explained by $\{e^{i\omega t}, e^{-i\omega t}\}$

Fast with FFT - 0.5 seconds to transform all $n = 54,720$ time series
Proportion of variation explained by mean and seasonal cycle

Pixelwise mean: $Y(0; \mathbf{x})$, seasonal cycle: $Y(2\pi/365; \mathbf{x})$

Portion of variation left...

Maps are on standard deviation scale
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients

![Real Part, frequency 20](image1)
![Imaginary Part, frequency 20](image2)
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Maps of Fourier coefficients

Even more promising - spatial correlation in Fourier coefficients
Summary Statistics

Our idea - store a subset of Fourier coefficients

Space-time Gaussian process model for the remaining coefficients
  - Can interpolate (with uncertainties) remaining coefficients
Summary Statistics

Our idea - store a subset of Fourier coefficients

![Real Part, frequency 20](image1.png) ![Locations of Saved Coefficients, frequency 20](image2.png)

Space-time Gaussian process model for the remaining coefficients
- Can interpolate (with uncertainties) remaining coefficients
Summary Statistics

Our idea - store a subset of Fourier coefficients

Space-time Gaussian process model for the remaining coefficients
  ▶ Can interpolate (with uncertainties) remaining coefficients
Summary Statistics

Our idea - store a subset of Fourier coefficients

Space-time Gaussian process model for the remaining coefficients
  - Can interpolate (with uncertainties) remaining coefficients
Major Modeling Considerations

1. Nonseparability in space and time
2. Heterogeneity of time series across space
Half-Spectral Space-Time Models (Process Representations)

Stationary space-time process on $\mathbb{R}^d \times \mathbb{Z}$

$$Y(\mathbf{x}, t) = \int_0^{2\pi} \int_{\mathbb{R}^d} g(\mathbf{\nu}, \omega) \exp(\mathbf{i}\mathbf{\nu}'\mathbf{x} + i\omega t) dB(\mathbf{\nu}, \omega) \leftarrow \text{orth. inc.}$$
Half-Spectral Space-Time Models (Process Representations)

Stationary space-time process on $\mathbb{R}^d \times \mathbb{Z}$

$$Y(x, t) = \int_0^{2\pi} \int_{\mathbb{R}^d} g(\nu, \omega) \exp(i\nu' x + i\omega t) dB(\nu, \omega) \iff \text{orth. inc.}$$

Integrating over $\nu$ (Stein, 2005)

$$Y(x, t) = \int_0^{2\pi} a(\omega) \exp(i\omega t) d\tilde{Z}(\omega; x) \iff \text{orth. inc. in } \omega$$
Half-Spectral Space-Time Models (Process Representations)

Stationary space-time process on $\mathbb{R}^d \times \mathbb{Z}$

$$Y(x, t) = \int_0^{2\pi} \int_{\mathbb{R}^d} g(\nu, \omega) \exp(i\nu'x + i\omega t) dB(\nu, \omega) \leftarrow \text{orth. inc.}$$

Integrating over $\nu$ (Stein, 2005)

$$Y(x, t) = \int_0^{2\pi} a(\omega) \exp(i\omega t) d\tilde{Z}(\omega; x) \leftarrow \text{orth. inc. in } \omega$$

Guinness and Stein (2013) considered nonstationary models over time

$$Y(x, t) = \int_0^{2\pi} a(\omega; t) \exp(i\omega t) d\tilde{Z}(\omega; x)$$
Half-Spectral Space-Time Models (Process Representations)

Stationary space-time process on \( \mathbb{R}^d \times \mathbb{Z} \)

\[
Y(x, t) = \int_0^{2\pi} \int_{\mathbb{R}^d} g(\nu, \omega) \exp(i\nu'x + i\omega t) dB(\nu, \omega) \leftarrow \text{orth. inc.}
\]

Integrating over \( \nu \) (Stein, 2005)

\[
Y(x, t) = \int_0^{2\pi} a(\omega) \exp(i\omega t) d\tilde{Z}(\omega; x) \leftarrow \text{orth. inc. in } \omega
\]

Guinness and Stein (2013) considered nonstationary models over time

\[
Y(x, t) = \int_0^{2\pi} a(\omega; t) \exp(i\omega t) d\tilde{Z}(\omega; x)
\]

This Project - Nonstationary models over space

\[
Y(x, t) = \int_0^{2\pi} a(\omega; x) \exp(i\omega t) d\tilde{Z}(\omega; x)
\]
Semiparametric Time Series Model

Define $f(\omega; \mathbf{x}) = |a(\omega, \mathbf{x})|^2$ (spectral density)
Semiparametric Time Series Model

Define \( f(\omega; \mathbf{x}) = |a(\omega, \mathbf{x})|^2 \) (spectral density)

Semi-parametric model for \( f(\omega; \mathbf{x}) \):

\[
\log f(\omega; \mathbf{x}) = u_0(\omega) + \theta_1(\mathbf{x}) u_1(\omega) + \theta_2(\mathbf{x}) u_2(\omega) + \theta_3(\mathbf{x}) u_3(\omega)
\]

\( u_j(\omega) \) are data-driven bases for the log spectra
Semiparametric Time Series Model

Define $f(\omega; \mathbf{x}) = |a(\omega, \mathbf{x})|^2$ (spectral density)

Semi-parametric model for $f(\omega; \mathbf{x})$:

$$\log f(\omega; \mathbf{x}) = u_0(\omega) + \theta_1(\mathbf{x})u_1(\omega) + \theta_2(\mathbf{x})u_2(\omega) + \theta_3(\mathbf{x})u_3(\omega)$$

$u_j(\omega)$ are data-driven bases for the log spectra

Estimate $\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \theta_3(\mathbf{x})$ for every $\mathbf{x}$.

- Fast with Whittle likelihood
Semiparametric Time Series Model

Define $f(\omega; \mathbf{x}) = |a(\omega, \mathbf{x})|^2$ (spectral density)

Semi-parametric model for $f(\omega; \mathbf{x})$:

$$\log f(\omega; \mathbf{x}) = u_0(\omega) + \theta_1(\mathbf{x})u_1(\omega) + \theta_2(\mathbf{x})u_2(\omega) + \theta_3(\mathbf{x})u_3(\omega)$$

$u_j(\omega)$ are data-driven bases for the log spectra

Estimate $\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \theta_3(\mathbf{x})$ for every $\mathbf{x}$.

- Fast with Whittle likelihood

Time series model storage burden: $3n + 4T/2 = 164, 892 = (0.0083)nT$
Spatial Correlation in Fourier Coefficients

Define: \( Z(\omega; x) = Y(\omega; x)/\sqrt{f(\omega; x)} \)

Under half-spectral model:
\( Z(\omega; x) \) approximately uncorrelated across \( \omega \), variance 1.

Expect to see correlation across \( x \) in \( Z(\omega; x) \)

(a) Real part, frequency 10
(b) Imaginary part, frequency 10
(c) Real part, frequency 120
(d) Imaginary part, frequency 120
Specific Objectives

A given compression rate dictates we can store $N$ Fourier coefficients

- We consider 20:1, 10:1, and 5:1 compression
- Total Storage Burden: $N + 3n + 5T/2$
Specific Objectives

A given compression rate dictates we can store $N$ Fourier coefficients
- We consider 20:1, 10:1, and 5:1 compression
- Total Storage Burden: $N + 3n + 5T/2$

Select $N$ Fourier coefficients $\mathcal{Y}_1 = (\mathcal{Y}(\omega_1, x_1), \ldots, \mathcal{Y}(\omega_1, x_N))$
Specific Objectives

A given compression rate dictates we can store $N$ Fourier coefficients

- We consider 20:1, 10:1, and 5:1 compression
- Total Storage Burden: $N + 3n + 5T/2$

Select $N$ Fourier coefficients $\mathcal{Y}_1 = (\mathcal{Y}(\omega_1, \mathbf{x}_1), \ldots, \mathcal{Y}(\omega_N, \mathbf{x}_N))$

And model parameters $\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \theta_3(\mathbf{x}), \kappa(\omega)$
Specific Objectives

A given compression rate dictates we can store $N$ Fourier coefficients

- We consider 20:1, 10:1, and 5:1 compression
- Total Storage Burden: $N + 3n + 5T/2$

Select $N$ Fourier coefficients $\mathbf{Y}_1 = (Y(\omega_1, x_1), \ldots, Y(\omega_N, x_N))$
And model parameters $\theta_1(x), \theta_2(x), \theta_3(x), \kappa(\omega)$

To maximize conditional likelihood of $\mathbf{Y}_2$ given $\mathbf{Y}_1$

$$p(\mathbf{Y}_1, \mathbf{Y}_2) = p(\mathbf{Y}_1)p(\mathbf{Y}_2 | \mathbf{Y}_1)$$
Specific Objectives

A given compression rate dictates we can store $N$ Fourier coefficients

- We consider 20:1, 10:1, and 5:1 compression
- Total Storage Burden: $N + 3n + 5T/2$

Select $N$ Fourier coefficients $\mathbf{y}_1 = (\mathcal{Y}(\omega_1, x_1), \ldots, \mathcal{Y}(\omega_N, x_N))$
And model parameters $\theta_1(x), \theta_2(x), \theta_3(x), \kappa(\omega)$

To maximize conditional likelihood of $\mathbf{y}_2$ given $\mathbf{y}_1$

$$p(\mathbf{y}_1, \mathbf{y}_2) = p(\mathbf{y}_1)p(\mathbf{y}_2|\mathbf{y}_1)$$

Why conditional likelihood?

- Model for $\mathbf{y}_1$ not relevant - we store them!
- Strictly proper scoring rule for prediction
- High conditional likelihood $\Leftrightarrow$ Sharp predictive distributions
Conditional Distributions

We store $\mathbf{y}_1$ and $p(\mathbf{y}_2|\mathbf{y}_1)$ (summary statistics + conditional distribution)
Conditional Distributions

We store $\mathbf{y}_1$ and $p(\mathbf{y}_2|\mathbf{y}_1)$ (summary statistics + conditional distribution)

Conditional model $p(\mathbf{y}_2|\mathbf{y}_1)$ used for decompression

- Best guess: $\hat{\mathbf{y}}_2 = E(\mathbf{y}_2|\mathbf{y}_1)$ (conditional expectation)
- Realistic value: $\tilde{\mathbf{y}}_2 \sim p(\mathbf{y}_2|\mathbf{y}_1)$ (conditional simulation)
Conditional Distributions

We store $\mathbf{Y}_1$ and $p(\mathbf{Y}_2|\mathbf{Y}_1)$ (summary statistics + conditional distribution)

Conditional model $p(\mathbf{Y}_2|\mathbf{Y}_1)$ used for decompression

- Best guess: $\hat{\mathbf{Y}}_2 = E(\mathbf{Y}_2|\mathbf{Y}_1)$ (conditional expectation)
- Realistic value: $\tilde{\mathbf{Y}}_2 \sim p(\mathbf{Y}_2|\mathbf{Y}_1)$ (conditional simulation)

Conditional expectations give best predictions, but tend to oversmooth
Conditional simulations add the “roughness” back in
Intractability of Conditional Likelihood Maximization

Global optimization is an intractable combinatorial optimization problem

- Find optimal $N$ (hundreds of thousands) out of 20M pixels
- Conditional likelihood changes for each choice
Intractability of Conditional Likelihood Maximization

Global optimization is an intractable combinatorial optimization problem

- Find optimal $N$ (hundreds of thousands) out of 20M pixels
- Conditional likelihood changes for each choice

Reducing computational burden

- Time series parameters estimated individually using Whittle likelihood
- Greedy search for selecting $\mathcal{Y}(\omega_1, x_1), \ldots, \mathcal{Y}(\omega_N, x_N)$
- Estimates of $\kappa(\omega)$ updated during greedy search
Intractability of Conditional Likelihood Maximization

Global optimization is an intractable combinatorial optimization problem

- Find optimal $N$ (hundreds of thousands) out of 20M pixels
- Conditional likelihood changes for each choice

Reducing computational burden

- Time series parameters estimated individually using Whittle likelihood
- Greedy search for selecting $\mathcal{Y}(\omega_1, x_1), \ldots, \mathcal{Y}(\omega_N, x_N)$
- Estimates of $\kappa(\omega)$ updated during greedy search

High-level overview of Greedy Search: given a selection of coefficients,

- Add coefficients that are not well predicted by conditional model
- Update estimates of spatial coherence parameters
- Iterate
Fitted Time Series Model Parameters

\[ f(\omega; x) = u_0(\omega) + \theta_1(x)u_1(\omega) + \theta_2(x)u_2(\omega) + \theta_3(x)u_3(\omega) \]
Fitted Time Series Model Parameters

\[ f(\omega; \mathbf{x}) = u_0(\omega) + \theta_1(\mathbf{x})u_1(\omega) + \theta_2(\mathbf{x})u_2(\omega) + \theta_3(\mathbf{x})u_3(\omega) \]

Fitted Semiparametric Time Series Parameters, j = 2
Fitted Time Series Model Parameters

\[ f(\omega; x) = u_0(\omega) + \theta_1(x)u_1(\omega) + \theta_2(x)u_2(\omega) + \theta_3(x)u_3(\omega) \]
Selected Coefficient Locations
Selected Coefficient Locations
Selected Coefficient Locations

Frequency 2

20:1

10:1

5:1
Selected Coefficient Locations
Selected Coefficient Locations
Selected Coefficient Locations
Selected Coefficient Locations
Selected Coefficient Locations
Selected Coefficient Locations
Selected Coefficient Locations

![Frequency 60](image)

- 20:1
- 10:1
- 5:1
Selected Coefficient Locations
Selected Coefficient Locations
Selected Coefficient Locations
Fitted Spatial Coherence Parameters

Estimated Spatial Coherence Parameters (inverse range)
Proportion of variation explained by mean and seasonal cycle

Recall from earlier: Portion of variation left...

Maps are on standard deviation scale
Root Mean Squared Prediction Error of Conditional Expectations

Average RMSPE (Celsius)

<table>
<thead>
<tr>
<th></th>
<th>Land</th>
<th>Ocean</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>20:1</td>
<td>0.7461</td>
<td>0.3844</td>
<td>0.5333</td>
</tr>
<tr>
<td>10:1</td>
<td>0.4414</td>
<td>0.2396</td>
<td>0.3214</td>
</tr>
<tr>
<td>5:1</td>
<td>0.2083</td>
<td>0.1345</td>
<td>0.1631</td>
</tr>
</tbody>
</table>
Contrast Variances

Other Statistical Properties: Contrast Variances

- Assess “smoothness” of decompressed datasets

Computed for original data and conditional simulations

\[
\frac{1}{T} \sum_{T=1}^{T} \left( Y(x, t) - Y(x + \delta_{\text{lat}}, t) \right)^2, \\
\frac{1}{T} \sum_{T=1}^{T} \left( Y(x, t) - Y(x + \delta_{\text{lon}}, t) \right)^2, \\
\frac{1}{T-1} \sum_{T=1}^{T-1} \left( Y(x, t) - Y(x, t+1) \right)^2,
\]
Contrast Variance for Conditional Simulations

North-South, Original Data

East-West, Original Data

Temporal, Original Data

5:1 Decompressed Data

5:1 Decompressed Data

5:1 Decompressed Data
Contrast Variance for Conditional Simulations
Contrast Variance for Conditional Simulations

North-South, Original Data

East-West, Original Data

Temporal, Original Data

20:1 Decompressed Data

20:1 Decompressed Data

20:1 Decompressed Data
Contrast Variance for Conditional Simulations
Original and Decompressed Time Series
Original and Decompressed Time Series

10:1 Decompressed Data

Temperature (Celsius)
Conditional Simulations

[ show gif of conditional simulations ]
**Discussion**

Compression is important because it allows us to save more model runs under more scenarios.

General strategy for compression via storing **summary statistics + conditional distributions**
- Allows for conditional simulations (avoiding oversmoothing)
- Flexible: Can choose which and how many statistics to store

**Nonstationary in space** half-spectral Gaussian process models
- Needed this model for these data
- Other datasets will require different statistical models

Decompression for 20M observations in **under 4 minutes**