Modeling Nonstationarity

BEE 6940 Lecture 10

March 27, 2023

Table of Contents

Review of Extreme Value Models

Two Common Approaches to Modeling Extremes

Block Maxima:

- Find maxima for independent blocks from time series;
- Can be inefficient use of data.

Peaks Over Thresholds:

- Set threshold and model level of exceedance *conditional on exceedance*;
- Choices of threshold and declustering length.

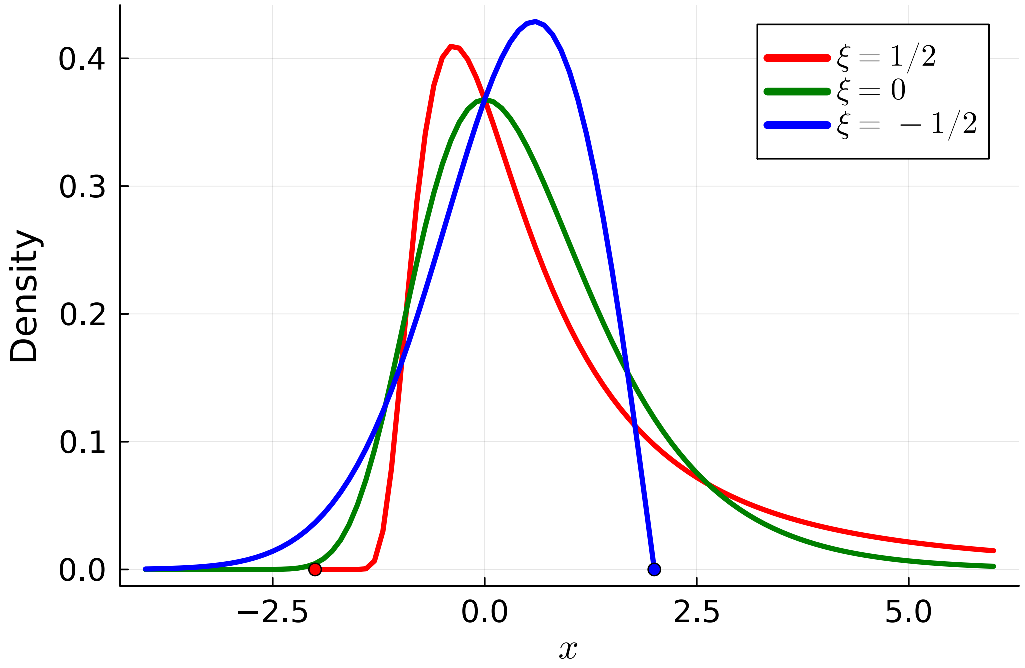
Block Maxima: Generalized Extreme Value Distributions

GEV distributions have three parameters:

- location μ;
- scale $\sigma > 0$;
- shape ξ.

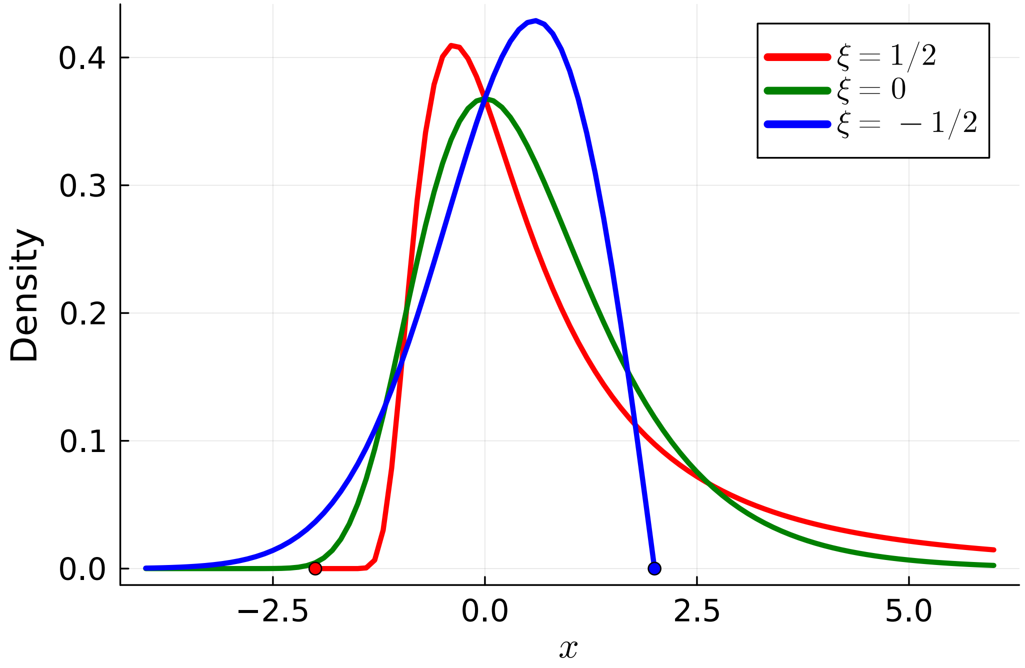
Generalized Extreme Value Distributions

The shape parameter ξ is particularly influential, as the GEV distribution can take on three shapes depending on its sign.



GEV Types

- $\xi > 0$: Frechet (*heavy-tailed*)
- $\xi = 0$: Gumbel (*light-tailed*)
- $\xi < 0$: Weibull (*bounded*)



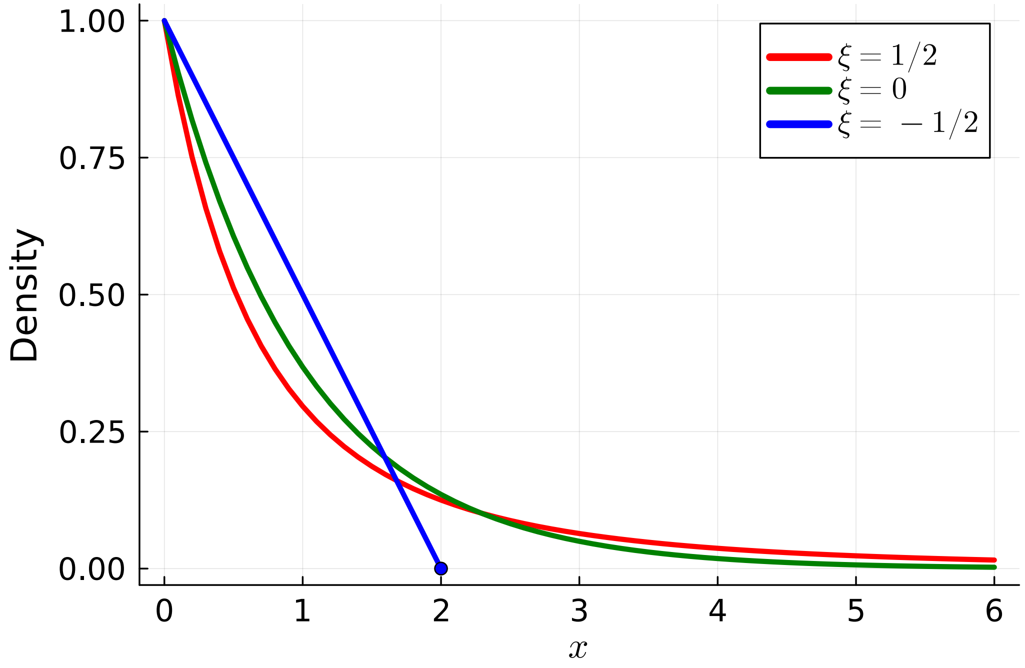
Peaks Over Thresholds: Generalized Pareto Distributions

Similarly to the GEV distribution, the GPD distribution has three parameters:

- location μ;
- scale $\sigma > 0$;
- shape ξ.

Generalized Pareto Distributions Types

- $\xi > 0$: *heavy-tailed*
- $\xi = 0$: *light-tailed*
- $\xi < 0$: *bounded*



Poisson-GP Processes

GPD model exceedances over threshold.

Often pair with Poisson processes to model the number of exceedances in a unit period.

GEV vs. PP-GP

GEV Model: For each time period, what is the largest event?

PP-GP: For each time period, how many exceedances of threshold, and how large is each one?

Return Levels

m-period return level: How large is the expected event which occurs with this frequency?

Alternative explanation: Exceedance probability of $1 - 1/m$.

Nonstationarity

Climate Change and Nonstationarity

However, these models assume *no long-term trend* in the data, so no change in the distribution of annual extremes.

This situation is called **stationary**: the underlying probability distribution does not change over time.

Climate Change and Nonstationarity

But climate change risks are fundamentally about dynamic distributions!

- Storm tracks/intensities
- Frequencies of extremes (heat waves, droughts, atmospheric rivers, etc.)
- Correlations between extreme events

Climate Change and Nonstationarity

But climate change risks are fundamentally about dynamic distributions!

- Storm tracks/intensities
- Frequencies of extremes (heat waves, droughts, atmospheric rivers, etc.)
- Correlations between extreme events

This means that we need to consider **nonstationarity**: the statistical model has a dependence on time (explicitly or implicitly).

Testing for Nonstationarity

Commonly used: **Mann-Kendall Test**.

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}(y_i - y_j),$$

Null hypothesis (zero trend):

$$S \sim \mathrm{Normal}\left(0, \frac{2(2n+5)}{9n(n-1)}\right)$$

Aside: Null-Hypothesis Significance Tests

Mann-Kendall fits into the framework of **null-hypothesis significance tests (NHST)**.

This aligns with falsificationist scientific paradigm. The test is whether to reject a *null* hypothesis in favor of the existence of a relationship.

- *Null hypothesis*: Typically that the proposed relationship does not exist.
- *Alternative hypothesis*: The relationship does exist.

Aside: Null-Hypothesis Significance Tests

For example:

- Null: No effect in a regression model (coefficient is zero)
- Alternative: Effect is non-zero

Or:

- Null: No trend over time
- Alternative: Trend exists

Statistical Significance

The "significance" in NHST is based on the frequentist notion of sampling distributions.

Goal: try to identify whether the pattern in your data is strong enough that it likely did not emerge due to sampling chance.

This involves balancing *Type I* (false positive) and *Type II* (false negative) error rates.

Type I and Type II Errors

		Null Hypothesis Is	
		True	False
Decision About Null Hypothesis	Don't reject	True negative (probability $1 - \alpha$)	Type II error (false negative, probability β)
	Reject	Type I Error (false positive, probability α)	True positive (probability $1 - \beta$)

The **significance level** α is the probability of rejecting the null hypothesis **assuming that it is true** (Type I errror).

P-VALUES

The **p-value** captures the probability of observing results **at least as extreme as observed** under the null hypothesis.

Therefore, if a p-value is small (below α), it can mean:

1. The null hypothesis is not true for that data;
2. The null hypothesis *is* true and the data is an outlying sample.

Notice: the p-value is itself a random variable; it is contingent on the sample.

P-VALUES

What a p-value is **not**:

- Probability that the null hypothesis is true (this is meaningless in the frequentist paradigm);
- Probability that the effect was produced by chance alone (a p-value is conditional on the assumption that the null hypothesis is true)
- An indication of the effect size

These misunderstandings are behind the replication crisis...

Mann-Kendall Test

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(y_i - y_j),$$

Null hypothesis (zero trend):

$$S \sim \text{Normal}\left(0, \frac{2(2n+5)}{9n(n-1)}\right)$$

"Problems" with Mann-Kendall

However:

- Mann-Kendall only suggests the presence of a trend, not its magnitude (general problem with statistical significance tests: what is the effect size?);
- Doesn't work if the trend is oscillating.

Alternative: Model Selection

We can also fit stationary and non-stationary models and see how they perform, and select accordingly.

Will discuss fitting today, selection after break.

Modeling Nonstationarity

Typically assume one (or more parameters) depend on another variable which can vary in time.

For example, could model block maxima as $\text{GEV}(\mu(t), \sigma, \xi)$, or frequency of occurrence as $\text{Poisson}(\lambda(t))$.

Often these are linear or generalized linear models:

$$\mu(t) = h(\sum_{i=0}^{n} \beta_i t^i).$$

Modeling Nonstationarity

- While any parameters can be treated as nonstationary, making models too complex can make them difficult to constrain given limited extremes data.

- Shape parameters are difficult to constrain normally, so are often best left constant.

Nonstationary Return Levels

Since we have a different model for each time t, we get different return levels for different times.

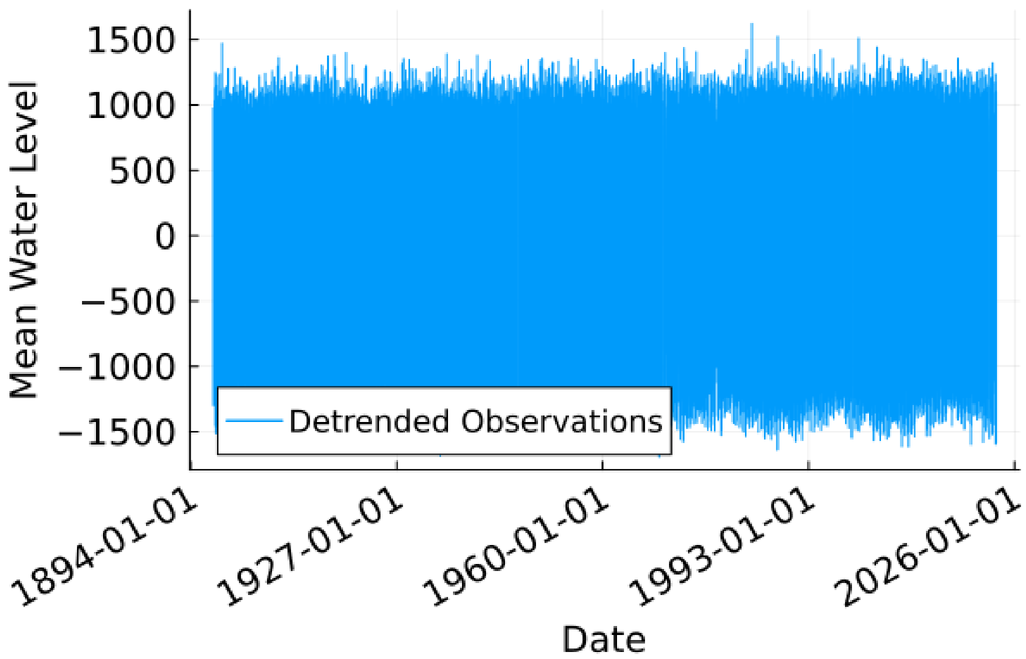
Contrast this with the stationary condition, in which we can just speak of "return levels".

Tide Gauge Example

Let's look at the San Francisco tide gauge data.

What are the implications of:

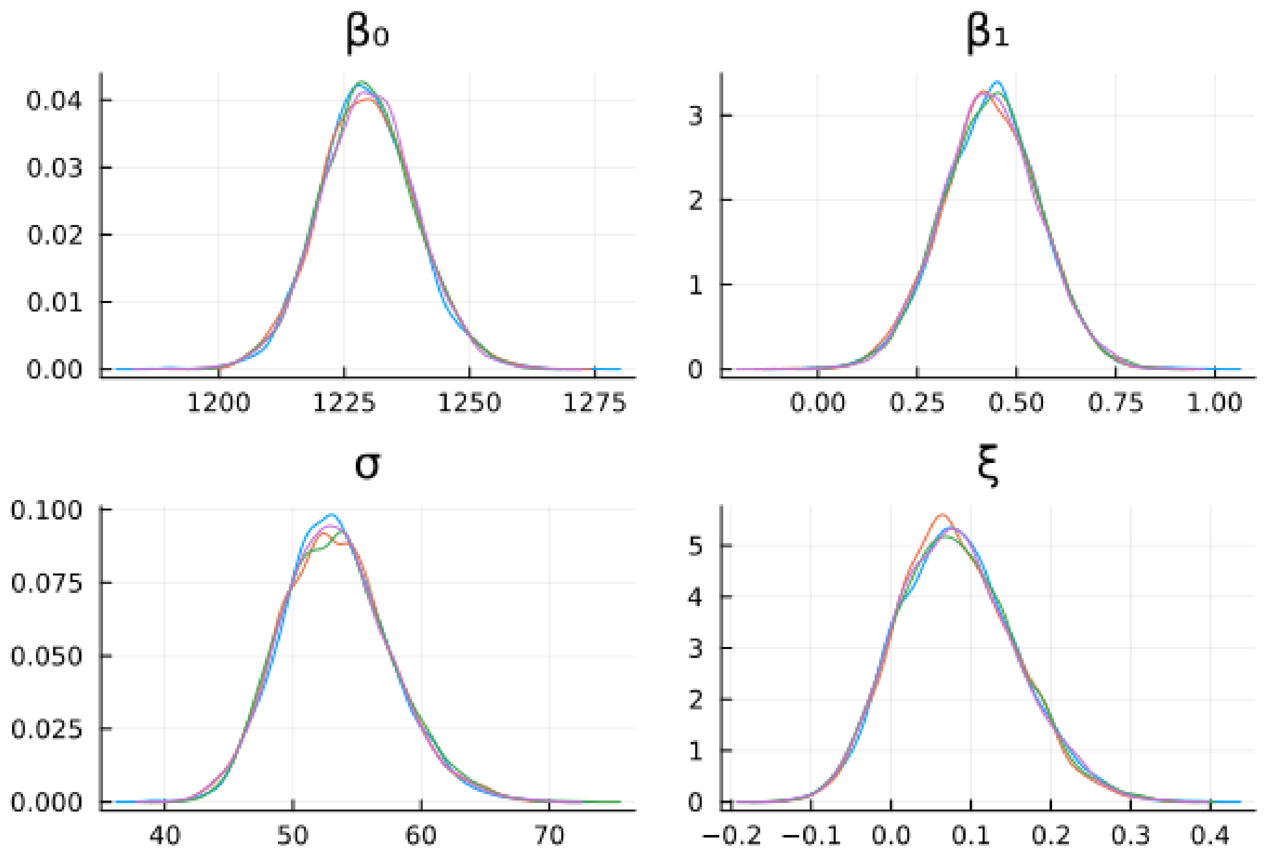
- Nonstationary GEV?
- Nonstationary Poisson rate?
- Nonstationary GPD?



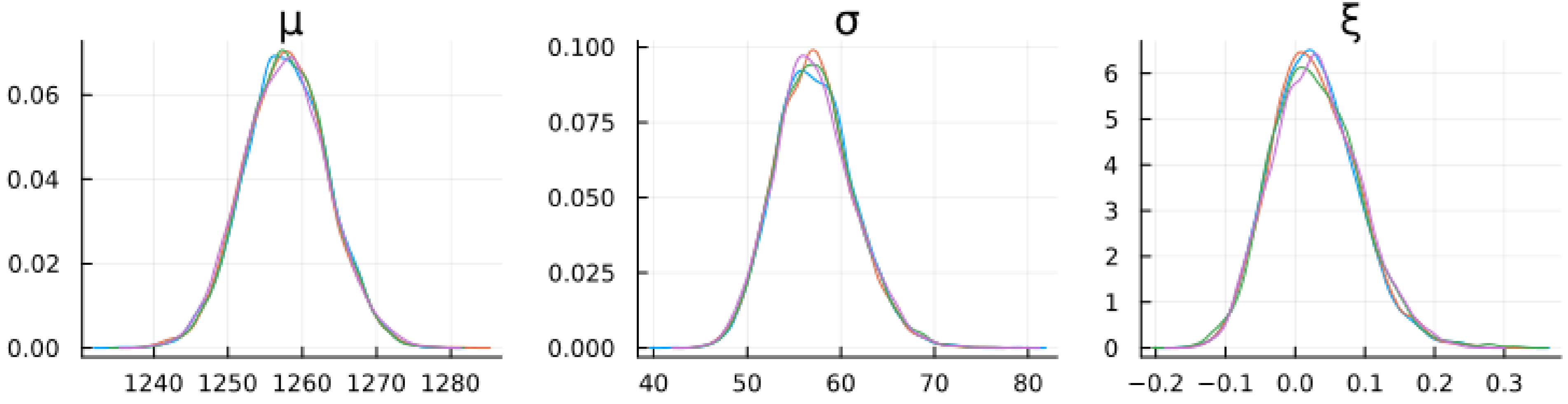
Nonstationary BLock Maxima Model

Let's fit a GEV with a linear trend in time: $\mu(t) = \beta_0 + \beta_1 t$, where t is in years.

Nonstationary Block Maxima Model Fit



Stationary Block Maxima Model Fit



Choice of Models

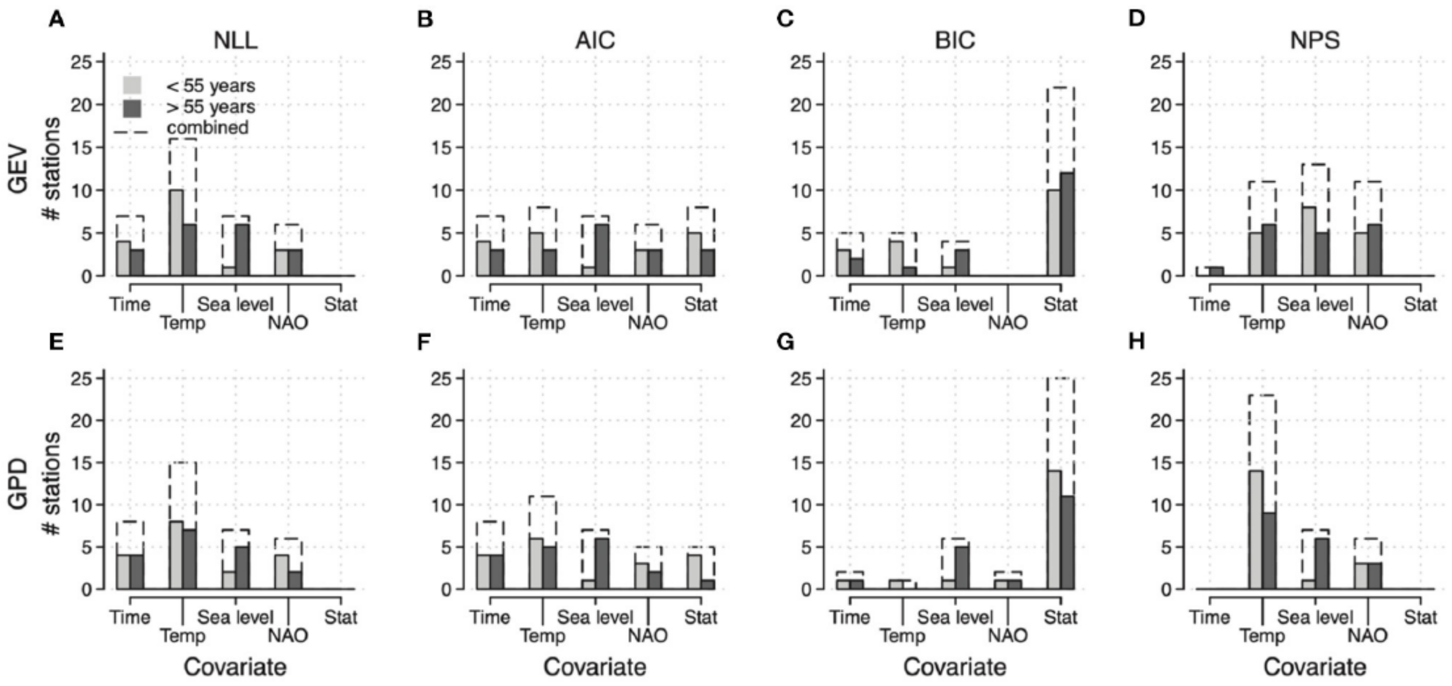
Possible Covariates

The candidate set of covariates is going to depend on the application.

For example, for storm surge, changes could be related to:

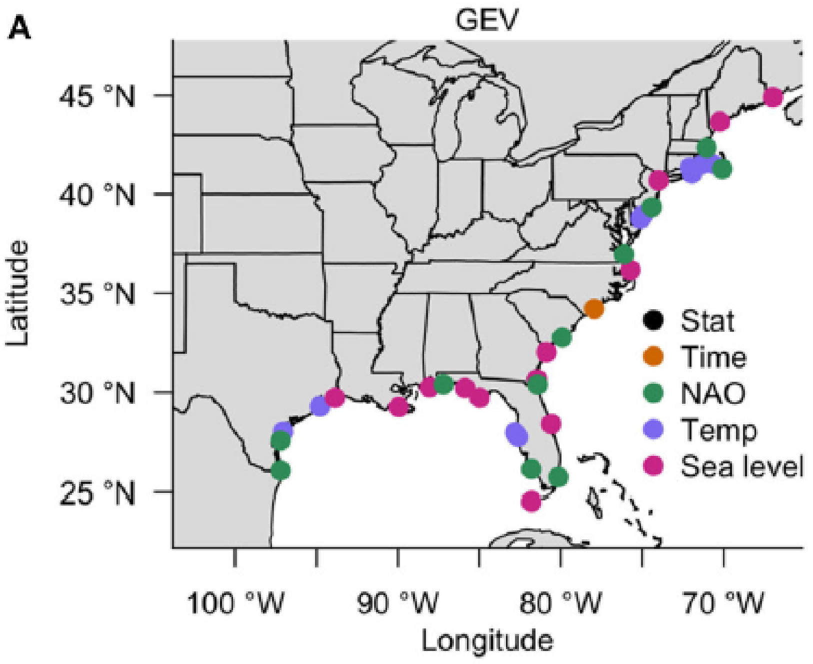
- sea-surface temperatures
- climate indices (North Atlantic Oscillation, Southern Oscillation)
- local mean sea level
- global mean temperature (as a broad proxy)
- time (general trend)

Space of Possible Models Is Difficult to Constrain



Wong et al (2022)

Space of Possible Models Is Difficult to Constrain



Wong et al (2022)

Key Takeaways

Key Takeaways

- **Nonstationarity**: Dynamic changes in the probability distribution
- Can be particularly hard to model/constrain with extremes due to limited data.
- Wise to avoid changing shape parameters.
- Nonstationary models can have very different return levels, so there are real implications for risk management.
- One possible path: adaptive decisions based on learning.

Upcoming Schedule

Upcoming Schedule

Wednesday: Discussion of Read & Vogel (2015).

Monday after break: Model selection