A photograph of a large, smooth, grey boulder resting on a rocky ground in a forest setting. The background shows bare trees, suggesting a late autumn or winter scene. The boulder is the central focus, with other smaller rocks scattered around it.

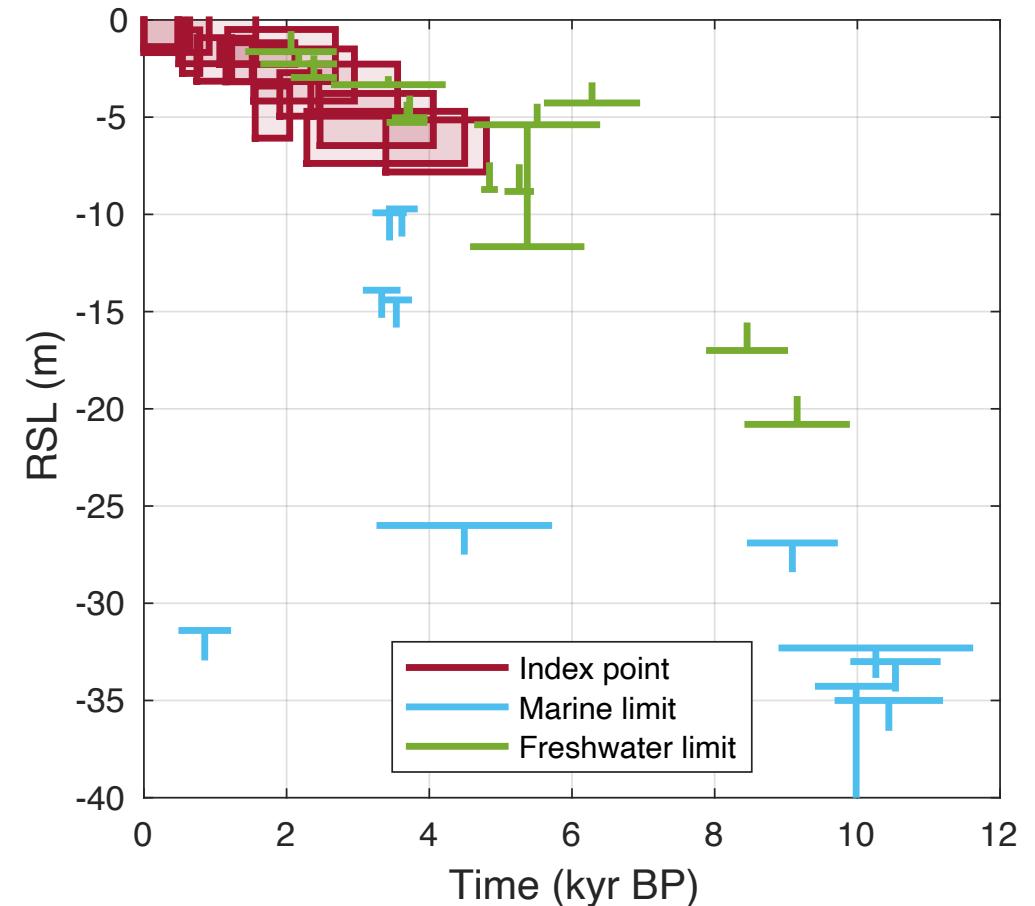
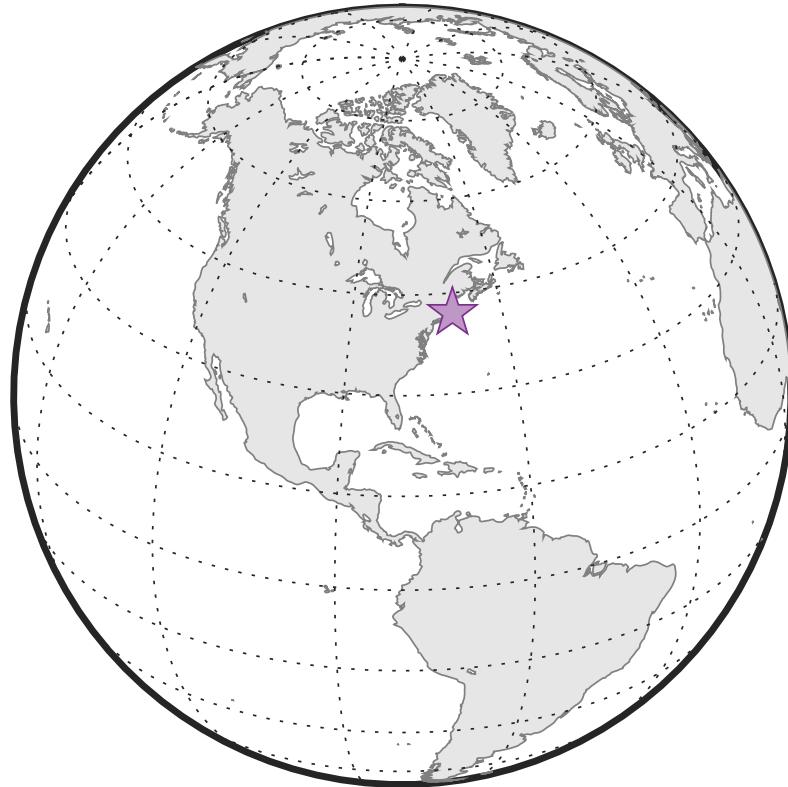
Doing Bayesian analysis for glacial isostatic adjustment

Chris Piecuch, Woods Hole Oceanographic Institution

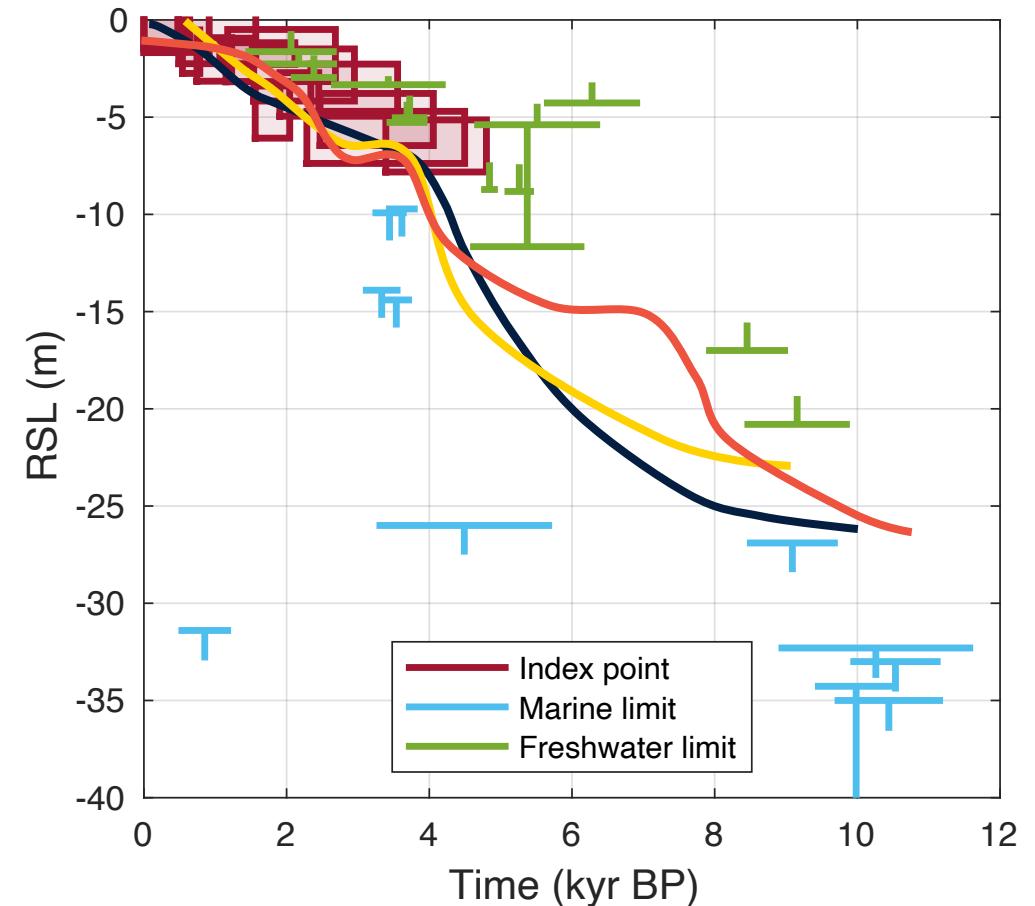
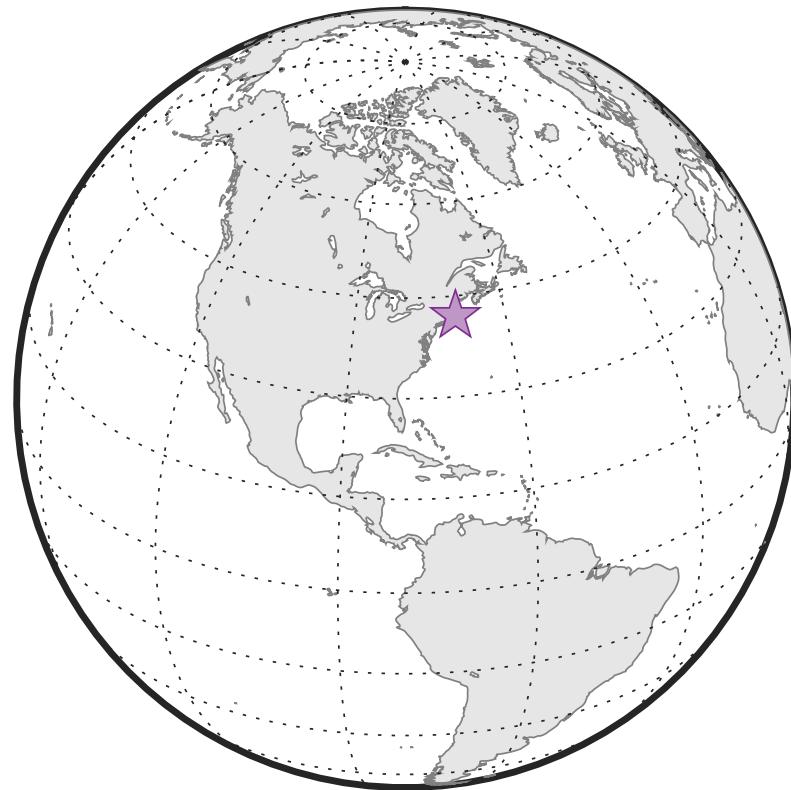
What do we have?

What do we want?

We have uncertain proxies; we want past rates of change



We have uncertain proxies; we want past rates of change



Known unknowns

random, parametric,
aleatoric, noise, modeled

Unknown unknowns

systematic, structural,
epistemic, bias, unmodeled

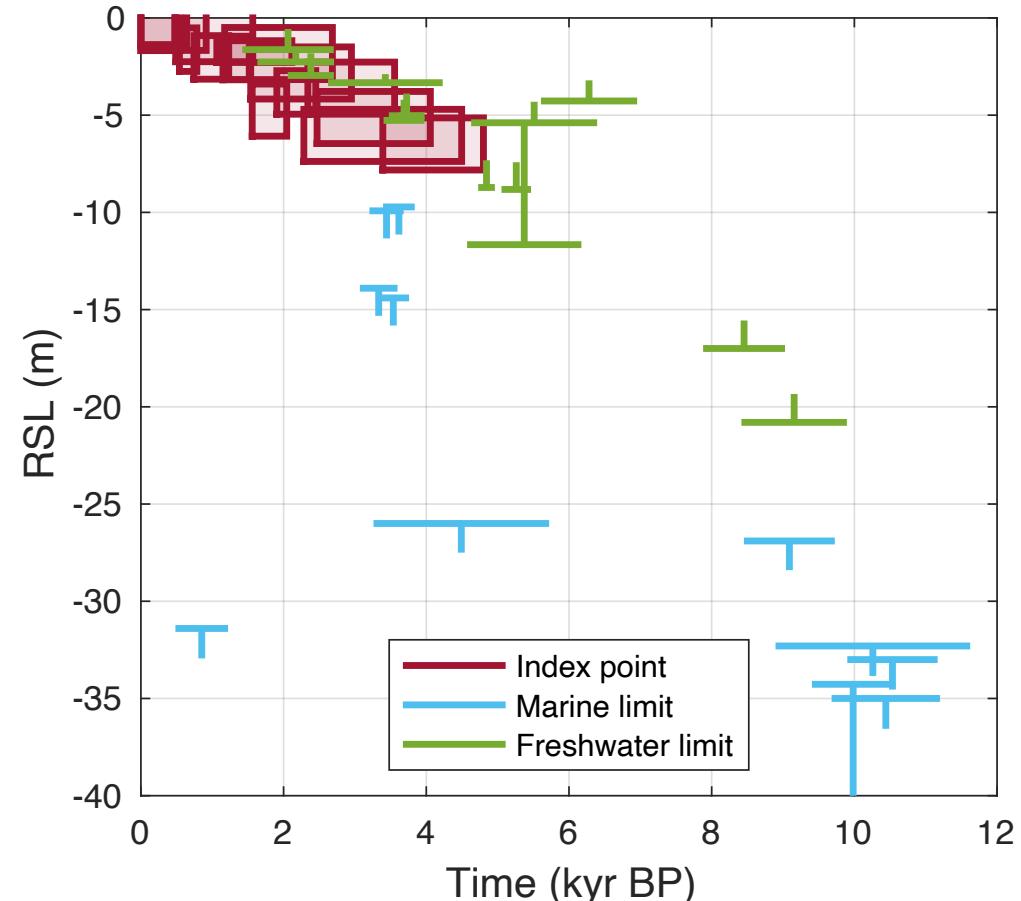


We have uncertain proxies; we want past rates of change

So how do we do this?

How do we get what we want
given what we have?

Simplest thing is to fit a line
through the best-estimate
ages and relative sea levels
from the index points using
ordinary least squares (OLS)



We have uncertain proxies; we want past rates of change

We specify the model

$$y_n = \alpha x_n + \beta + \varepsilon_n$$

where

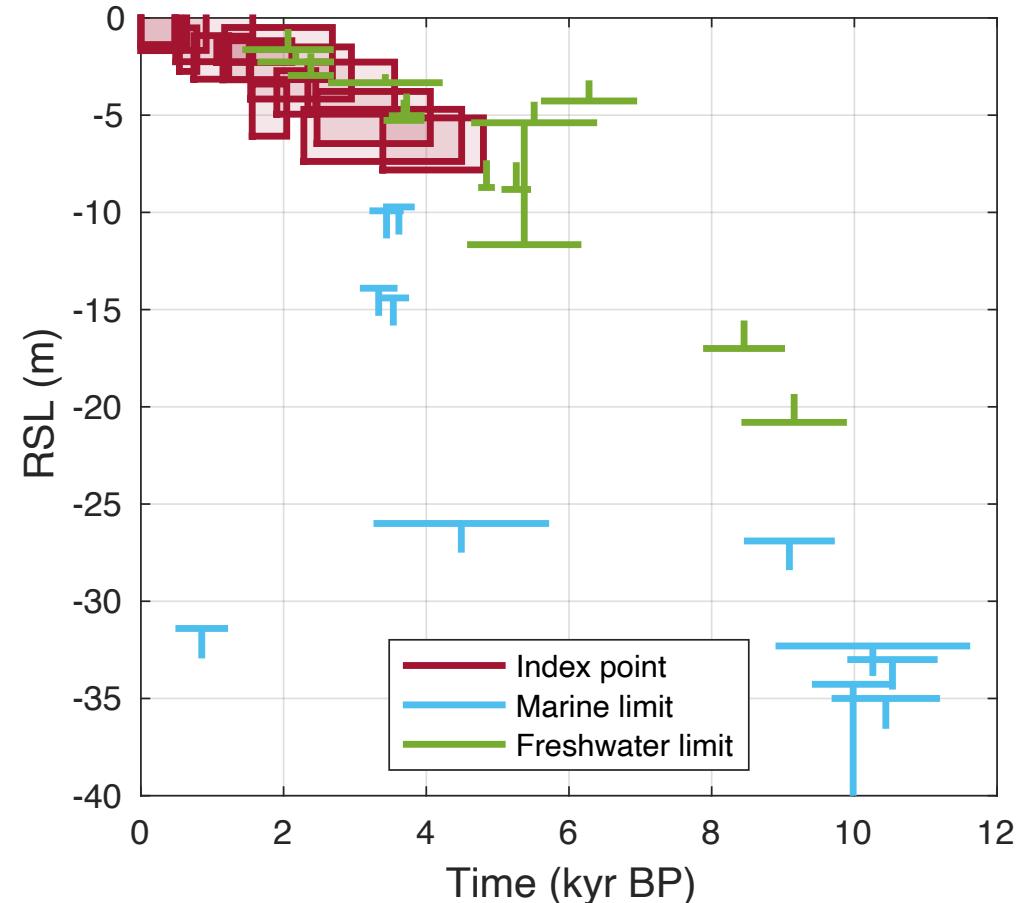
y_n is the nth value of sea level

x_n is the nth value of time

α is the slope

β is the intercept

ε_n is the nth residual (random
unstructured noise)



We have uncertain proxies; we want past rates of change

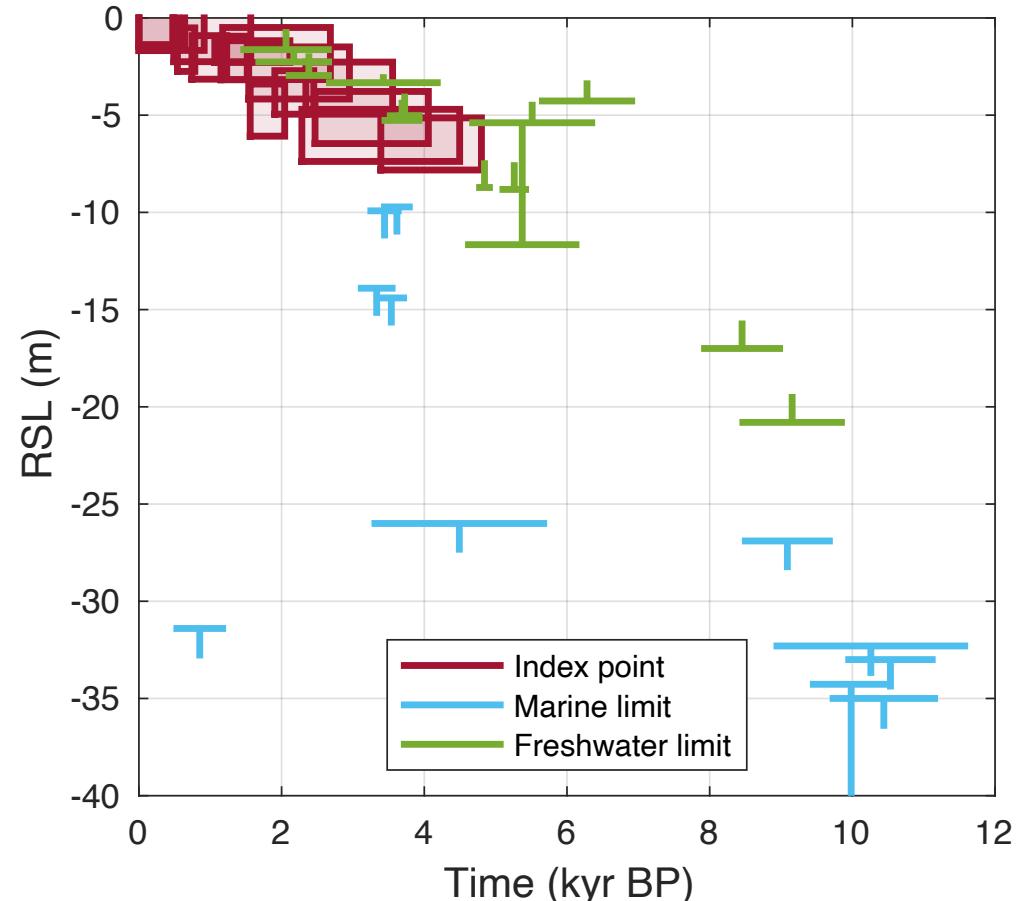
We solve it by minimizing the sum of the squared residuals

$$\sum_n (y_n - \alpha x_n - \beta)^2$$

OLS gives solutions $\hat{\alpha}$ for $\hat{\beta}$

$$\hat{\alpha} = \frac{\sum_n (x_n - \bar{x})(y_n - \bar{y})}{\sum_n (x_n - \bar{x})^2}$$

$$\hat{\beta} = \bar{y} - \hat{\alpha} \bar{x}$$



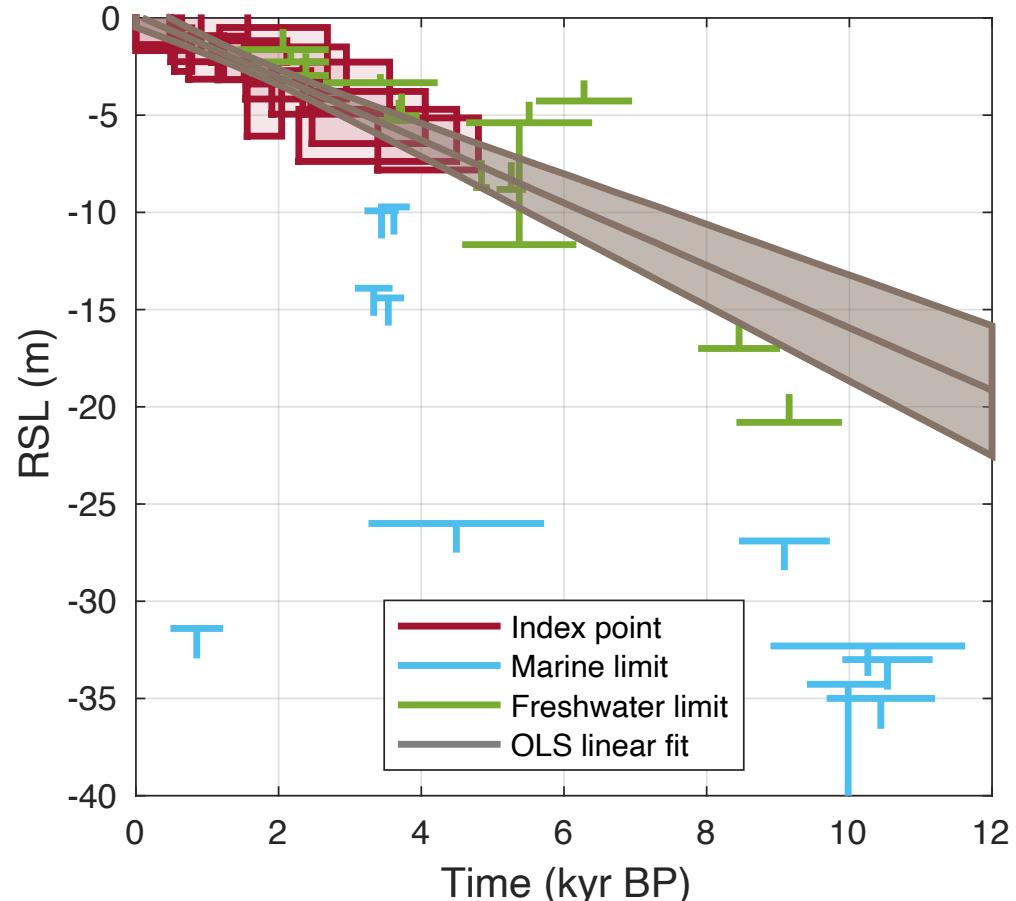
We have uncertain proxies; we want past rates of change

When we do this, we get the
value $\hat{\alpha} = 1.6 \pm 0.3 \text{ mm yr}^{-1}$

But what about the error bars
on the observations?

Do we have to prescribe a
particular functional form?

And how do we incorporate
limiting or bounding data?



Think probabilistically

Think hierarchically

Think conditionally

Observations are conditioned on **processes**

Processes are conditioned on **parameters**

Everything is uncertain!

Probability postulates

Say that S is the set (“sample space”) of all possible events and A is an event (“a subset of S ”)

Postulate #1

- The probability of A is a non-negative real number

$$P(A) \geq 0$$

Postulate #2

- Something must happen

$$P(S) = 1$$

Postulate #3

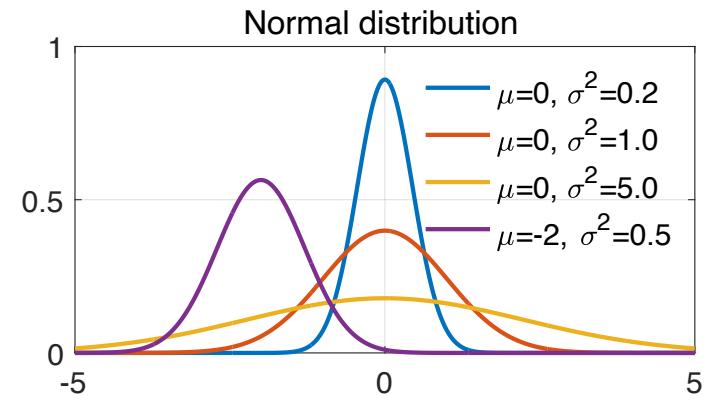
- If A, B, C, \dots are mutually exclusive events, then

$$P(A \cup B \cup C \cup \dots) = P(A) + P(B) + P(C) + \dots$$

Some standard distributions

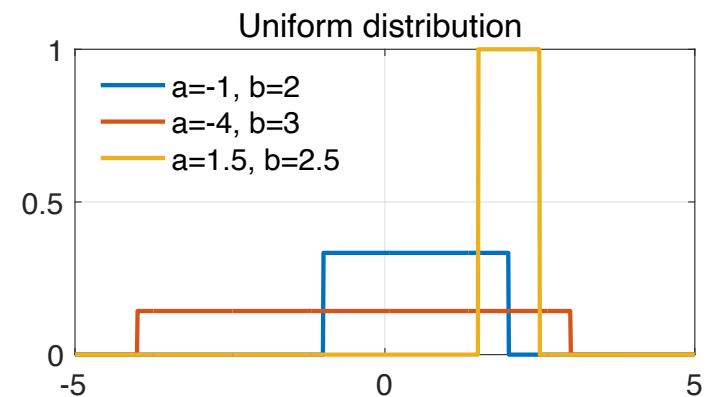
Normal

$$\bullet P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



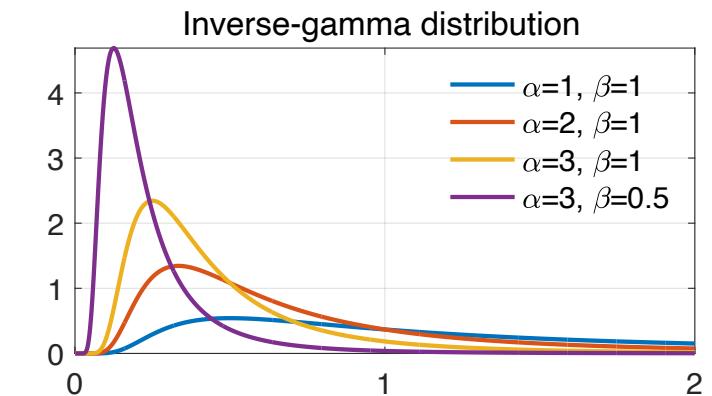
Uniform

$$\bullet P(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$



Inverse-gamma

$$\bullet P(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x}$$



Probability distributions

Marginal distribution

- $P(A)$ is the probability that A occurs

Joint distribution

- $P(A, B)$ is the probability that A and B both occur

Conditional distribution

- $P(A|B)$ is the probability that A occurs given that B is certain to occur (or already occurred)

These three distributions are related according to

$$P(A|B) = P(A, B)/P(B)$$

Bayes theorem

Bayes theorem

$$P(A|B) = P(B|A)P(A)/P(B) \propto P(B|A)P(A)$$

↑ ↑ ↑ →
Posterior Likelihood Prior Normalizing constant

Example: what is the temperature θ outside?

- Prior (climatology) $p(\theta) = N(\mu, \sigma^2); \mu = 17^\circ\text{C}, \sigma = 2^\circ\text{C}$
- Likelihood (data) $p(T|\theta) = N(\theta, \delta^2); T = 20^\circ\text{C}, \delta = 1^\circ\text{C}$
- Posterior (temperature estimate θ given the data T)

$$p(\theta|T) = N\left(\frac{\frac{\mu}{\sigma^2} + \frac{T}{\delta^2}}{\frac{1}{\sigma^2} + \frac{1}{\delta^2}}, \frac{1}{\frac{1}{\sigma^2} + \frac{1}{\delta^2}}\right) = N(19.4^\circ\text{C}, 0.64^\circ\text{C}^2)$$

Back to the data

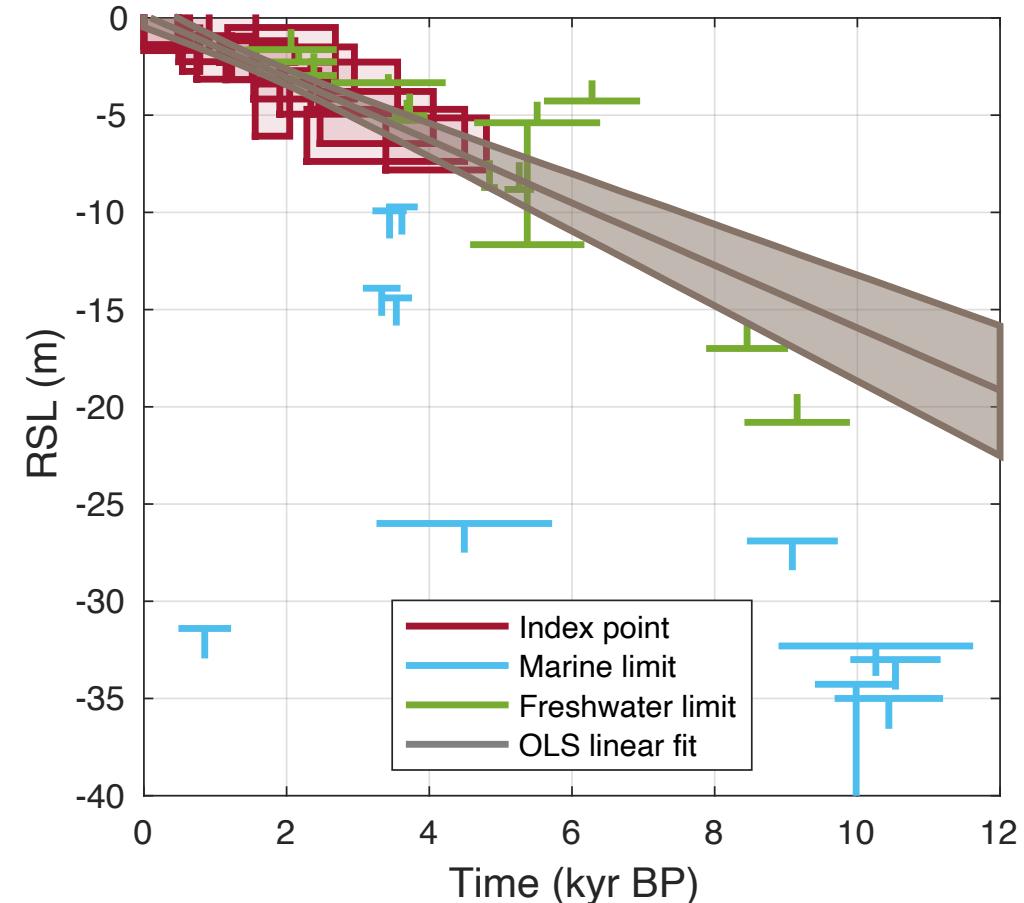
Let's think about what we did earlier

When we got the OLS slope, we did two things without realizing

1. We made a **modeling** choice
2. We made an **analysis** choice

What if we make different modeling and analysis choices?

Does that allow us to address the questions we had before?



A series of four Bayesian models

Start—
Ordinary least squares;
parametric;
perfect data;
without limiting dates



End—
Bayesian;
nonparametric;
uncertain data;
with limiting dates

**First model: linear/parametric process,
perfect data, without limiting data**

Model 1: linear/parametric process, perfect observations, without limiting data

Let's use the same model but
change the analysis approach

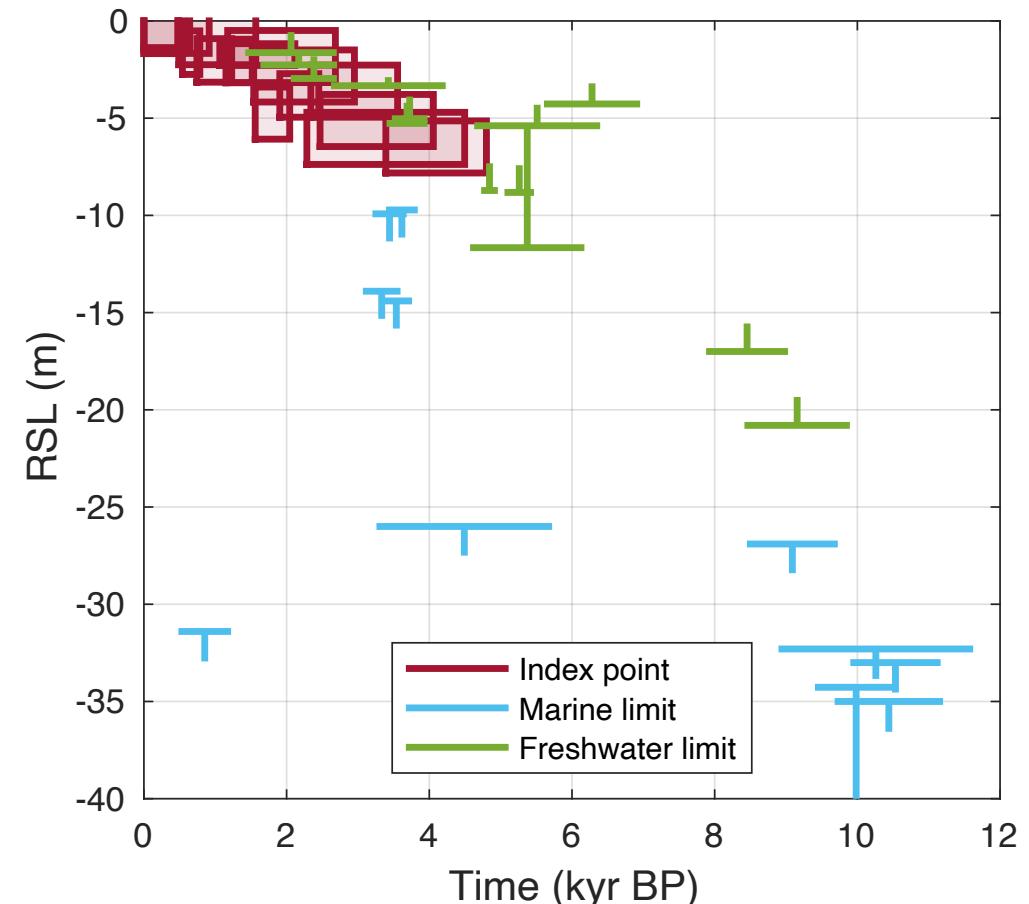
We want the **posterior** distribution

$$p(\alpha, \beta, \gamma^2 | x_1, \dots, y_1, \dots)$$

To get that, Bayes rule says that we
need **likelihood & prior** distributions

The likelihood is our linear model,
now written in a probabilistic form

$$p(y_n | x_n, \alpha, \beta, \gamma^2) = N(\alpha x_n + \beta, \gamma^2)$$



Model 1: linear/parametric process, perfect observations, without limiting data

The prior distributions express our belief about the parameters before we consider the data; let's choose

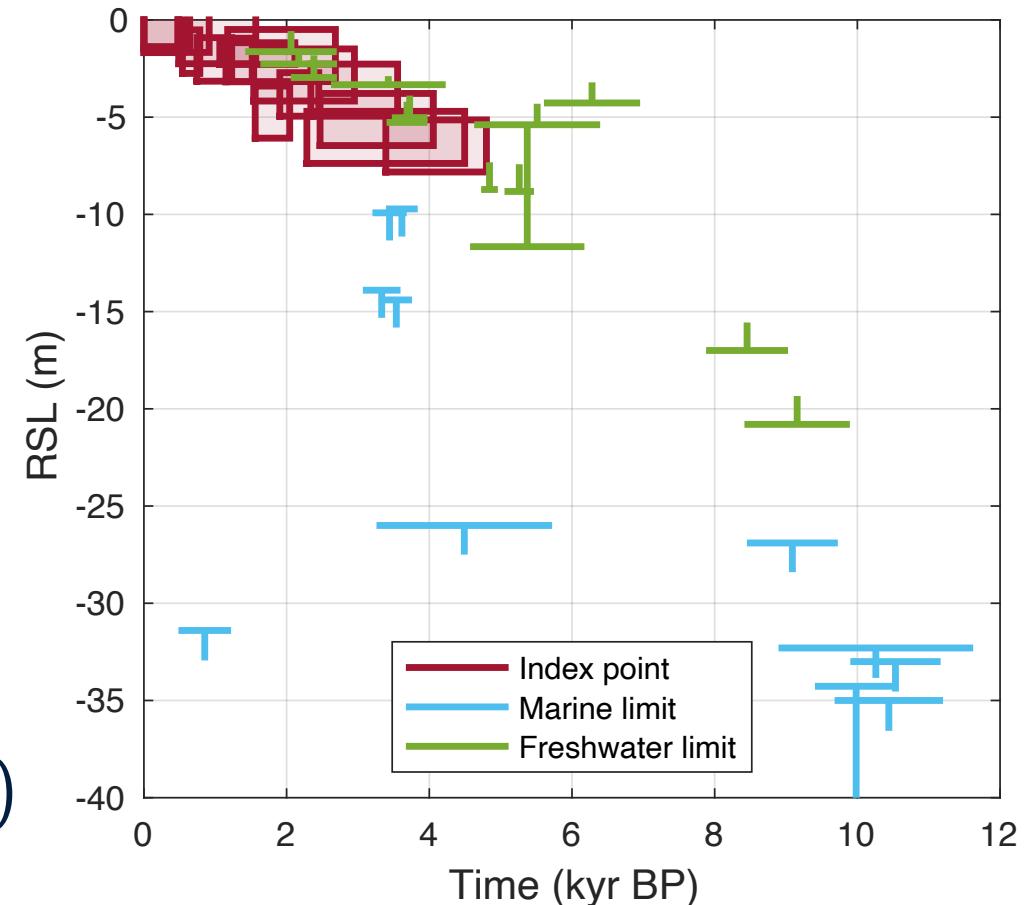
$$p(\alpha) = N(\tilde{\eta}_\alpha, \tilde{\zeta}_\alpha^2)$$

$$p(\beta) = N(\tilde{\eta}_\beta, \tilde{\zeta}_\beta^2)$$

$$p(\gamma^2) = G^{-1}(\tilde{\xi}_\gamma, \tilde{\chi}_\gamma)$$

Bringing the pieces together, we find

$$\begin{aligned} p(\alpha, \beta, \gamma^2 | x_1, \dots, y_1, \dots) \\ \propto p(\alpha)p(\beta)p(\gamma^2) \prod_n p(y_n | x_n, \alpha, \beta, \gamma^2) \end{aligned}$$



Model 1: linear/parametric process, perfect observations, without limiting data

We **cannot** sample this joint posterior,
but we **can** sample the full conditional
posteriors of each of the parameters

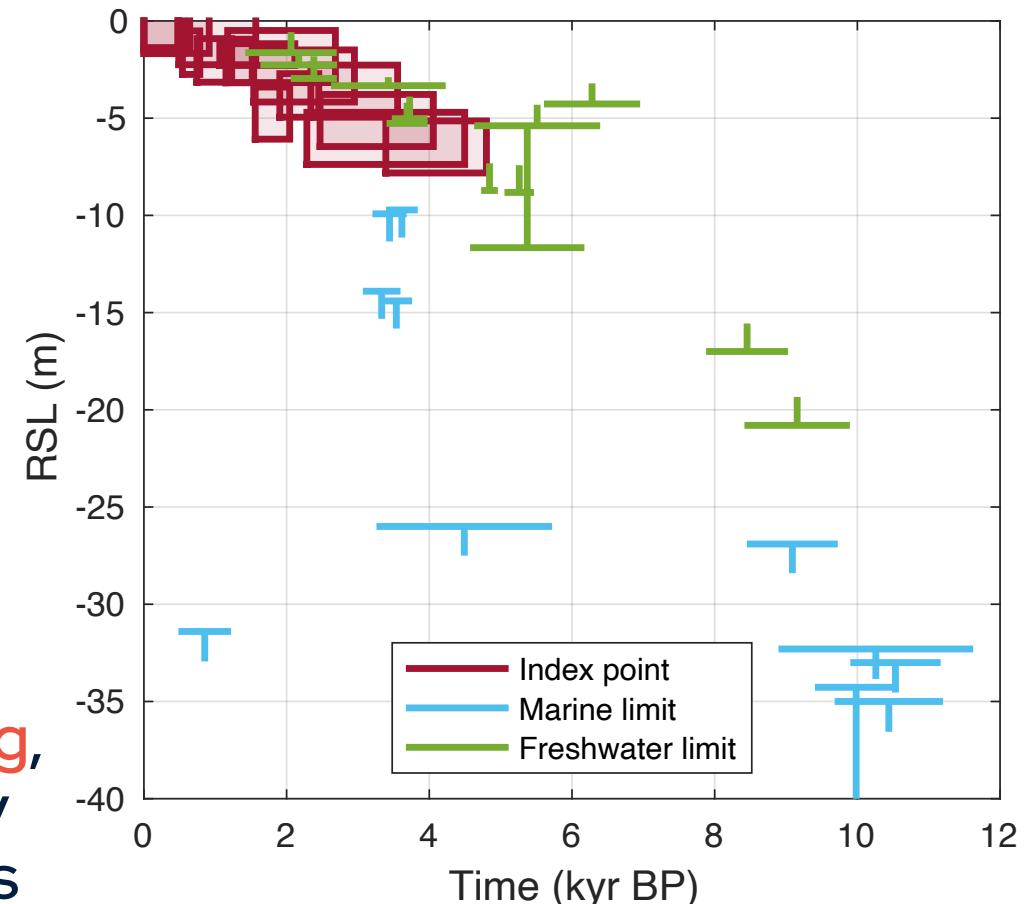
For example, for α , we have

$$p(\alpha | \cdot) = N(\psi_\alpha V_\alpha, \psi_\alpha)$$

$$V_\alpha = \tilde{\zeta}_\alpha^{-2} \tilde{\eta}_\alpha + \gamma^{-2} \sum_{k=1}^n x_k (y_k - \beta)$$

$$\psi_\alpha = 1 / (\tilde{\zeta}_\alpha^{-2} + \gamma^{-2} \sum_{k=1}^n x_k^2)$$

This forms the basis of **Gibbs sampling**,
which is a powerful way to numerically
sample from known distribution forms



Model 1: linear/parametric process, perfect observations, without limiting data

Gibbs sampler

First, make initial guesses for the parameters you seek $\alpha_{(0)}, \beta_{(0)}, \gamma^2_{(0)}$

Second, iteratively sample from the full conditional posterior distributions

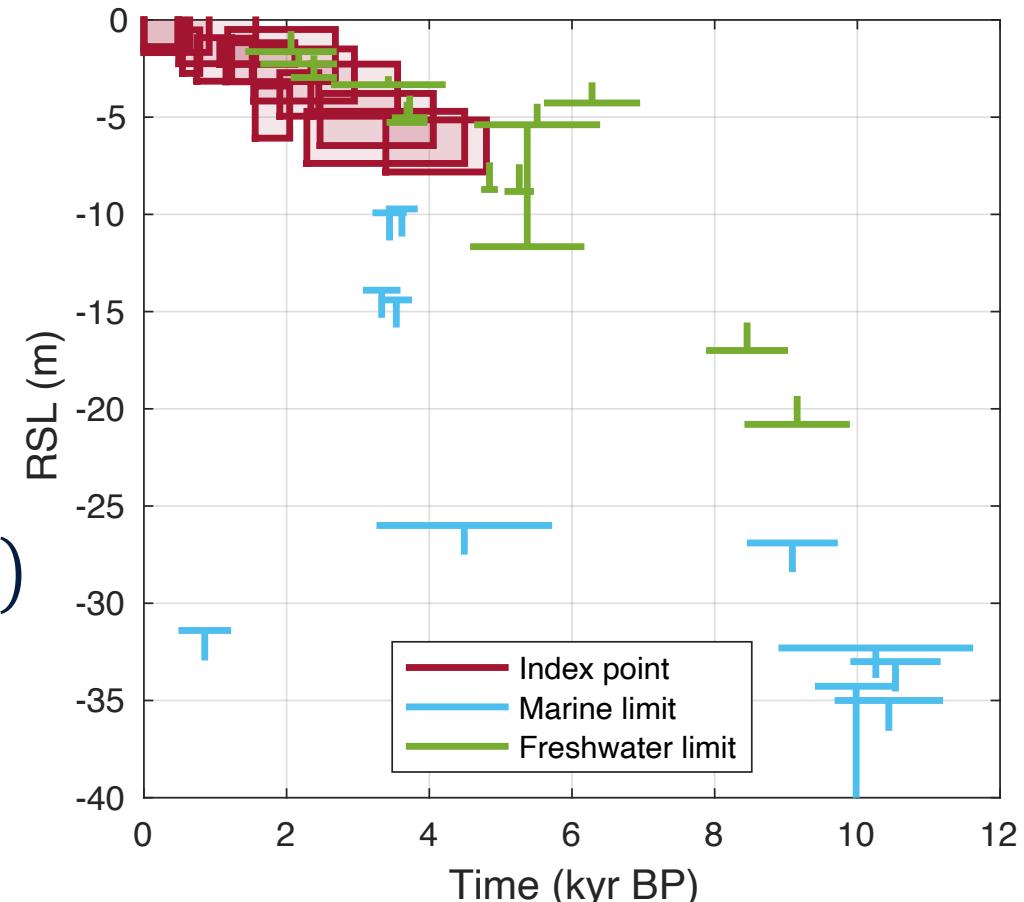
for $m = 1, \dots, M$

draw $\alpha_{(m)}$ from $p(\alpha | \beta_{(m-1)}, \gamma^2_{(m-1)}, \dots)$

draw $\beta_{(m)}$ from $p(\beta | \alpha_{(m)}, \gamma^2_{(m-1)}, \dots)$

draw $\gamma^2_{(m)}$ from $p(\gamma^2 | \alpha_{(m)}, \beta_{(m)}, \dots)$

end



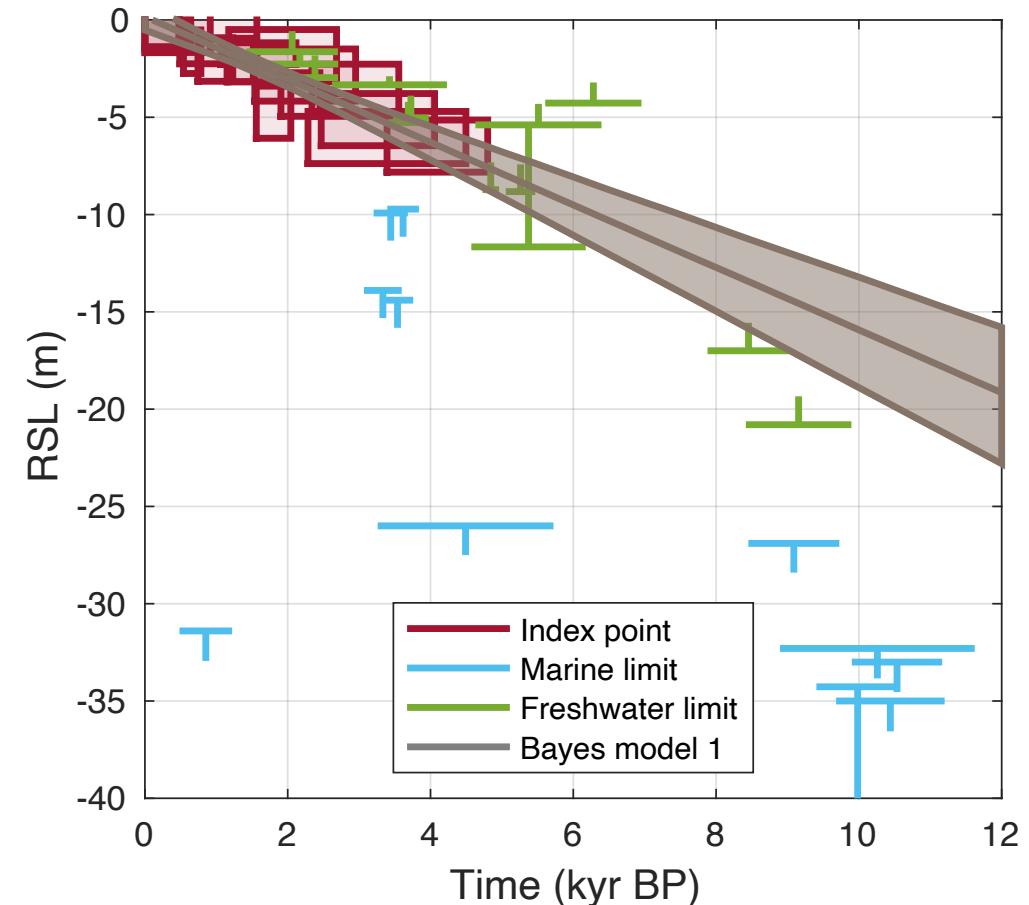
Model 1: linear/parametric process, perfect observations, without limiting data

We choose wide uninformative priors

When we run the Gibbs sampler, we
get $\alpha = 1.6 \pm 0.3 \text{ mm yr}^{-1}$ (95% CI)

Did we just do a lot of work to get the
same thing we had earlier with OLS?

Yes & no. You'll see the advantages
of the Bayesian approach shortly.



**Second model: linear/parametric process,
imperfect data, without limiting data**

Model 2: linear/parametric process, imperfect observations, without limiting data

How do we factor in the errors
on the proxy reconstructions?

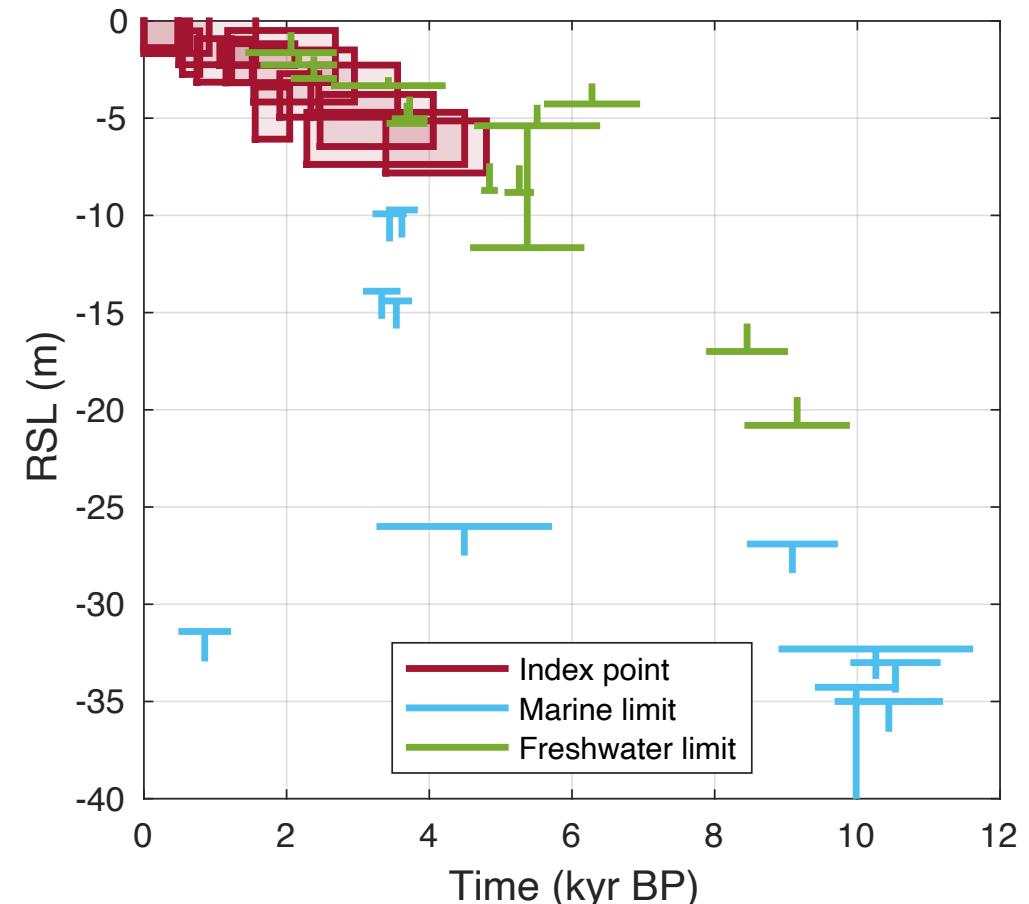
Earlier, we had a **process level**
and a **prior level**; now let's add
more equations in a **data level**

Say that index points are noisy
versions of the true processes

$$p(z_n) = N(y_n, \delta_n^2)$$

$$p(w_n) = N(x_n, \epsilon_n^2)$$

where z_n, w_n are sea level, age
data, and δ_n^2, ϵ_n^2 are the errors



Model 2: linear/parametric process, imperfect observations, without limiting data

All other equations are the same

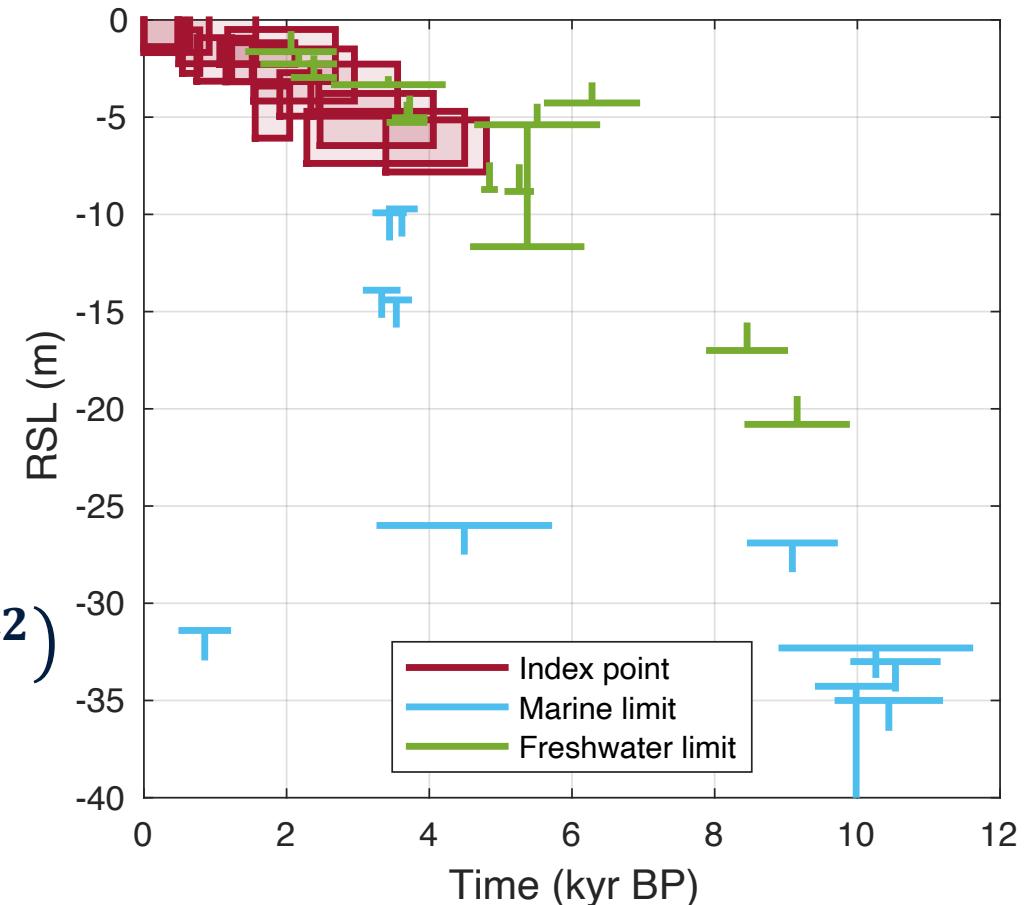
Now x_n and y_n are uncertain(!),
so our posterior distribution is

$$p(\alpha, \beta, \gamma^2, x_1, \dots, y_1, \dots | w_1, \dots, z_1, \dots)$$

Using Bayes, & our prior, data, &
process equations, this becomes

$$\propto p(\alpha)p(\beta)p(\gamma^2) \prod_n [p(y_n|x_n, \alpha, \beta, \gamma^2) \\ \times p(z_n|y_n) p(w_n|x_n)]$$

Draw solutions for $\alpha, \beta, \gamma^2, x, y$
again based on Gibbs sampling

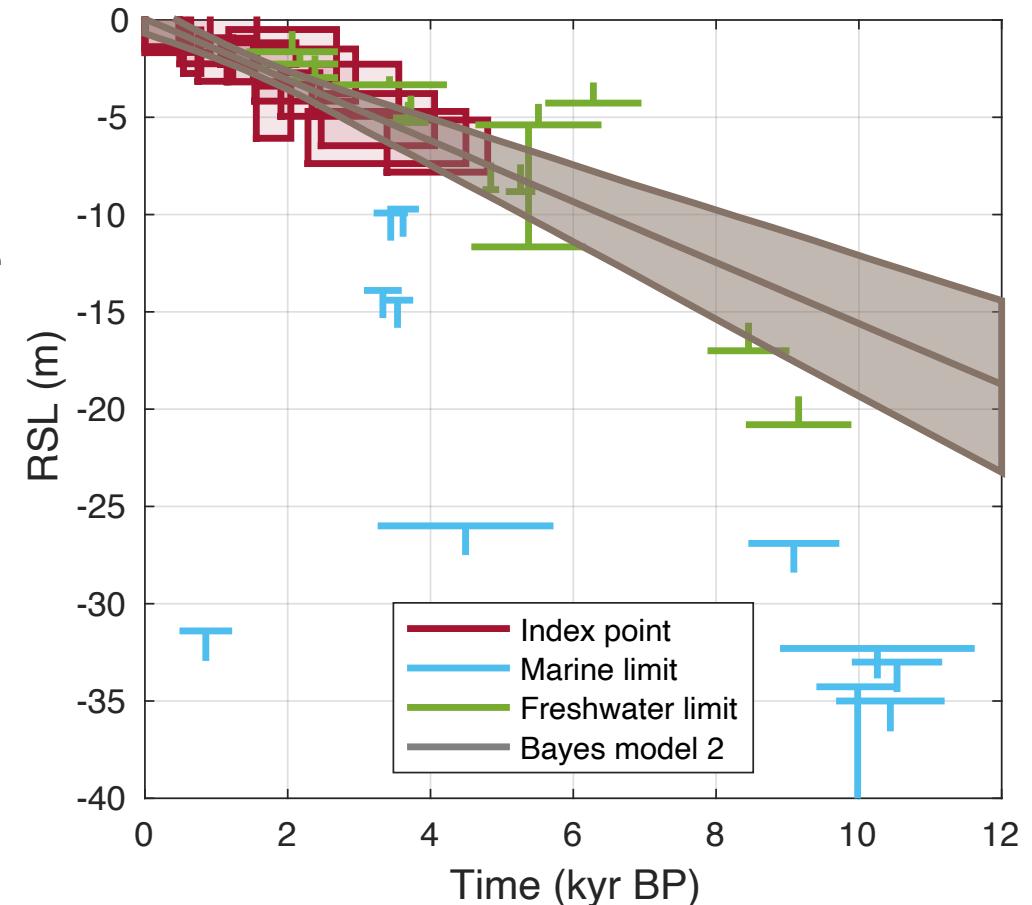


Model 2: linear/parametric process, imperfect observations, without limiting data

Now we determine a sea level rate of
 $\alpha = 1.6 \pm 0.4 \text{ mm yr}^{-1}$ (95% CI)

The error bars are now wider because
we include observational uncertainty
in the design of our Bayesian model

We arguably obtain a more realistic
linear trend in sea level, but is a linear
trend the most appropriate model?



**Third model: nonlinear/nonparametric process,
imperfect data, without limiting data**

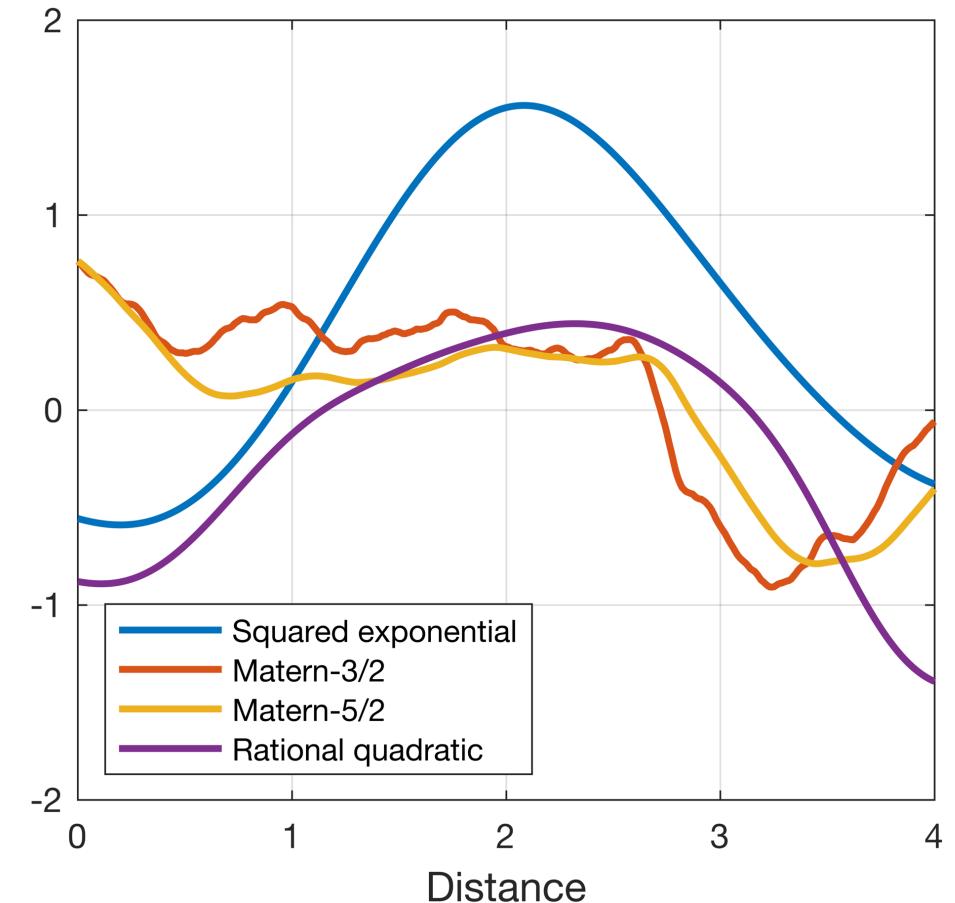
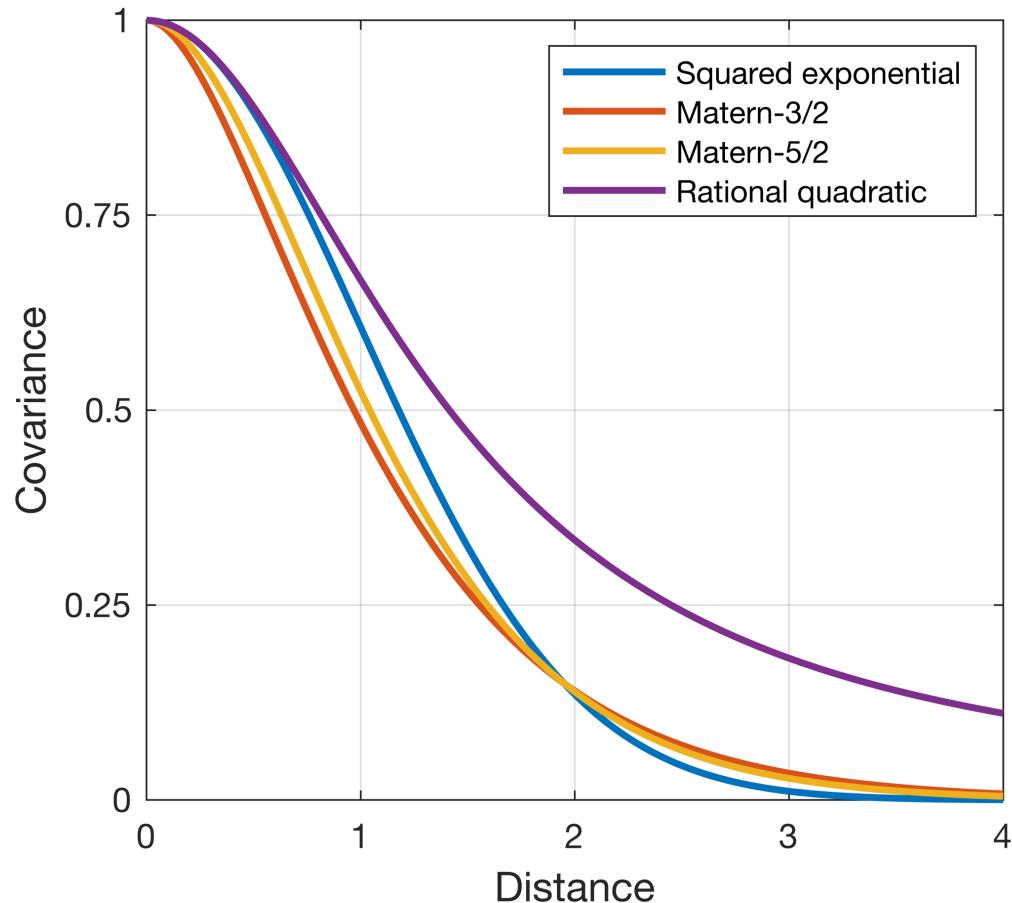
We must change how we view the process

We will still use a Gaussian distribution, but ...

We **were** putting all structure in the **mean**;
now we'll put the structure in the **variance**

We **were** prescribing a **functional form**;
now we'll specify a **relation between points**

Covariance functions (kernels)



Model 3: nonlinear/nonparametric process, imperfect observations, without limiting data

We'll use the same data model, but write a new process model

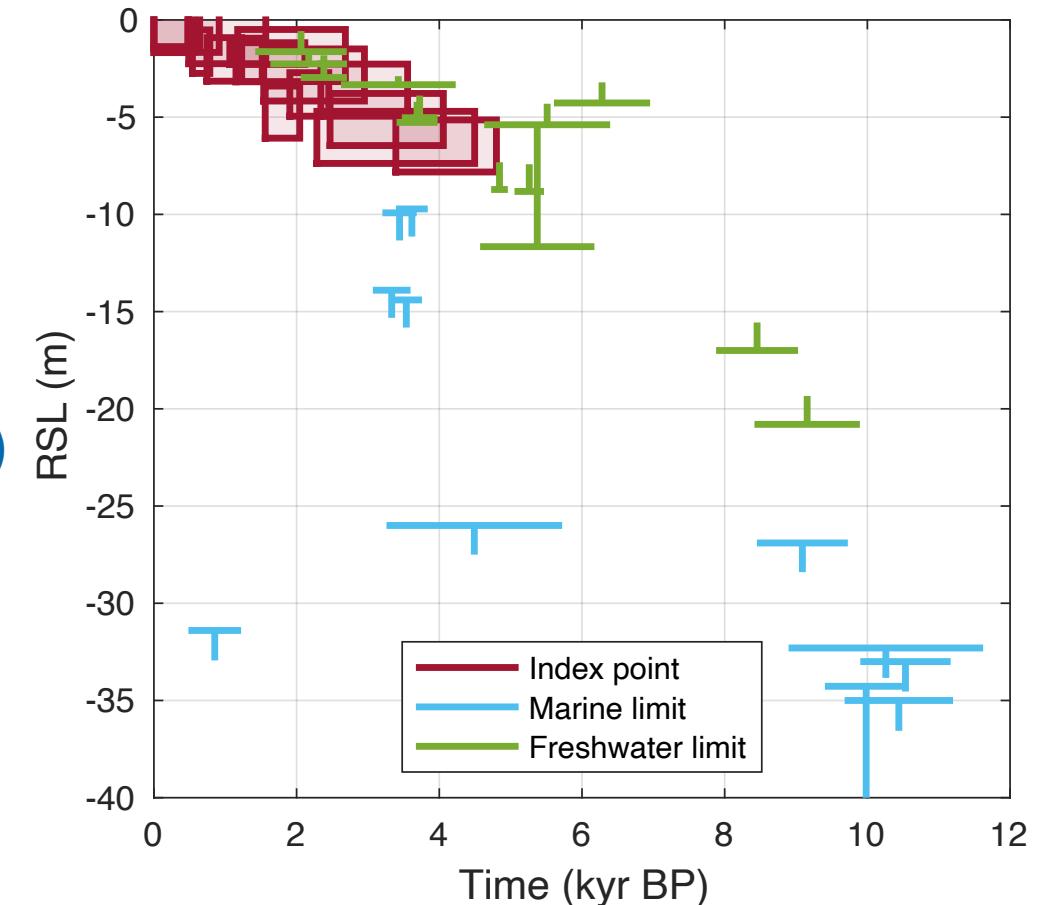
Let's represent sea level y as a random process with structure

$$y \sim N(0, L + G) \quad \text{local (uncorrelated)}$$

$$L_{jk} = \lambda^2 \delta(x_j, x_k) \quad \text{regional (correlated)}$$

$$G_{jk} = \rho^2 C(x_j, x_k, \varphi)$$

where δ is the Kronecker delta and C the covariance function (compare to Khan et al. 2015)



Model 3: nonlinear/nonparametric process, imperfect observations, without limiting data

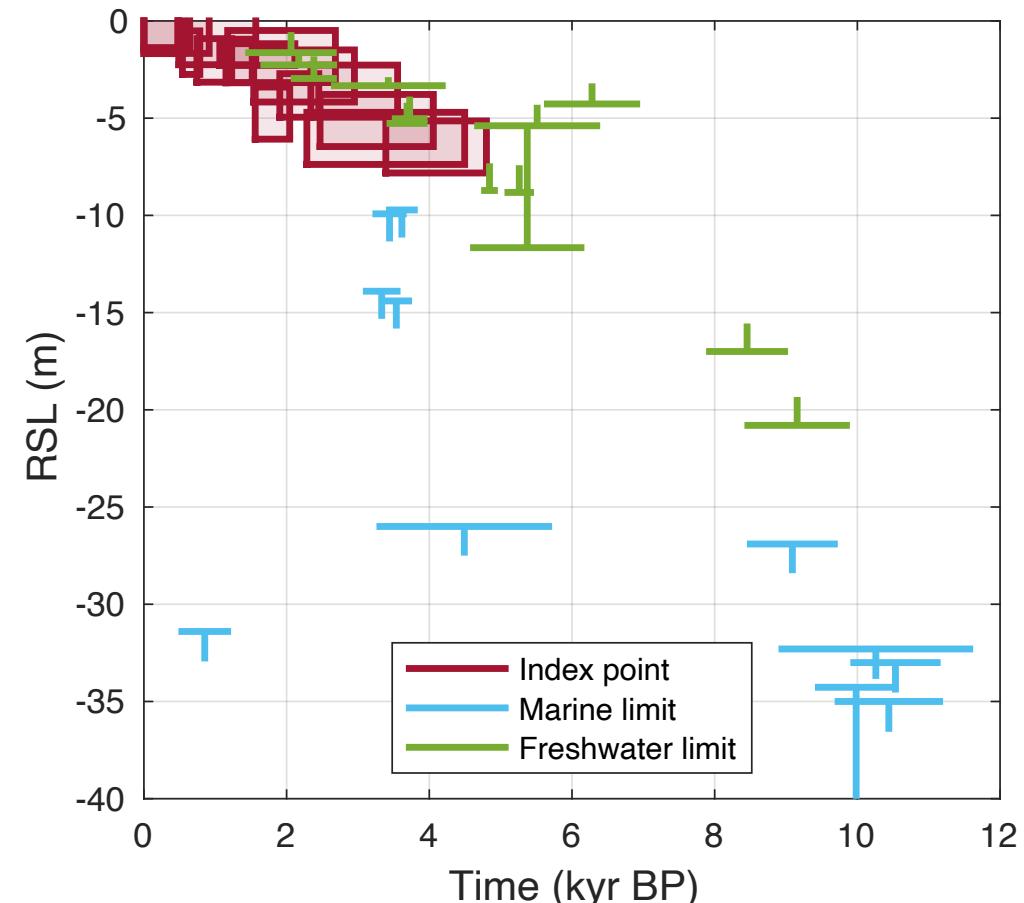
Place priors on $\lambda^2, \rho^2, \varphi$

The new posterior distribution is

$$\begin{aligned} p(\lambda^2, \rho^2, \varphi, x_1, \dots, y_1, \dots | w_1, \dots, z_1, \dots) \\ \propto p(\lambda^2)p(\rho^2)p(\varphi) \prod_n [p(z_n|y_n) \\ \times p(w_n|x_n)p(y_n|x, y_{m \neq n}, \lambda^2, \rho^2, \varphi)] \end{aligned}$$

Posterior is complicated, and full conditional posteriors don't all have standard distributions

We can't use Gibbs sampler, but we can use Metropolis sampler



Model 3: nonlinear/nonparametric process, imperfect observations, without limiting data

Metropolis sampler for parameter ϑ

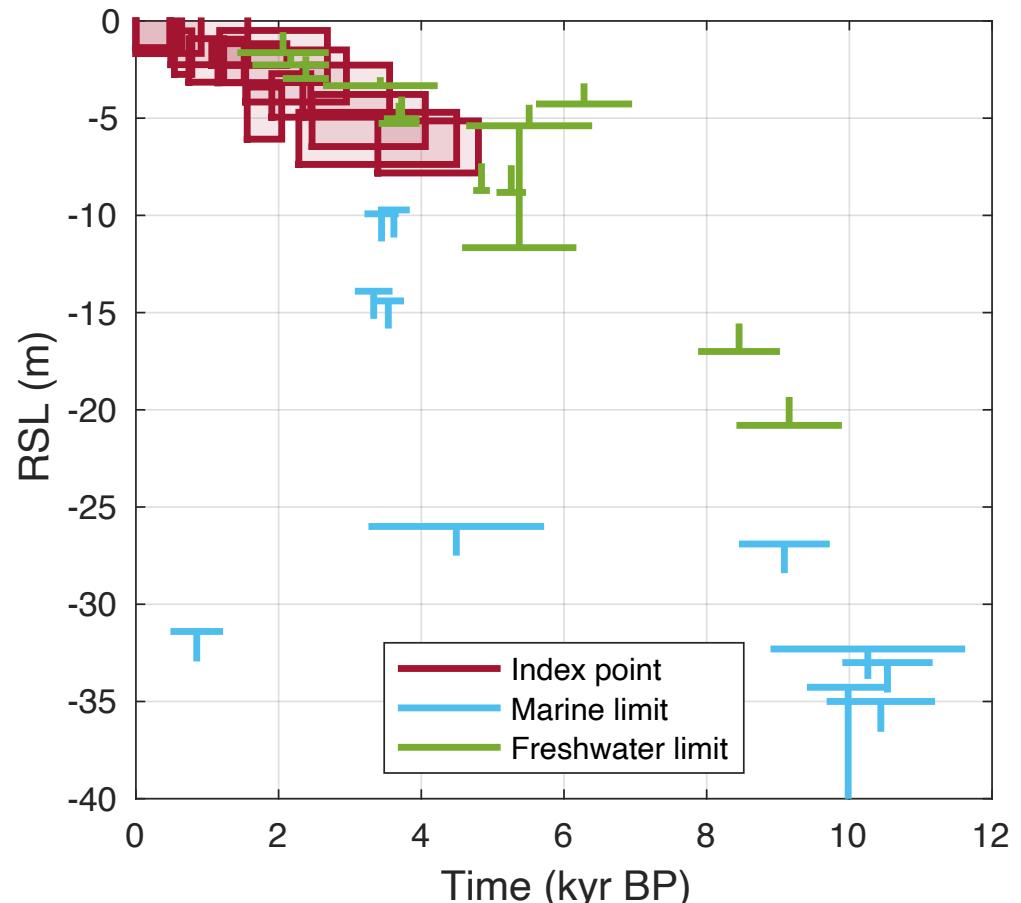
Goal is to sample a new estimate ϑ^{new} based on the current estimate ϑ^{now}

1. Propose a new value $\vartheta^* = \vartheta^{now} + \xi$ with ξ from a symmetric distribution

2. Then compute the ratio

$$R_M = \frac{P(\vartheta^* | \cdot)}{P(\vartheta^{now} | \cdot)}$$

3. Accept the proposition and define $\vartheta^{new} = \vartheta^*$ with probability $\min(1, R_M)$



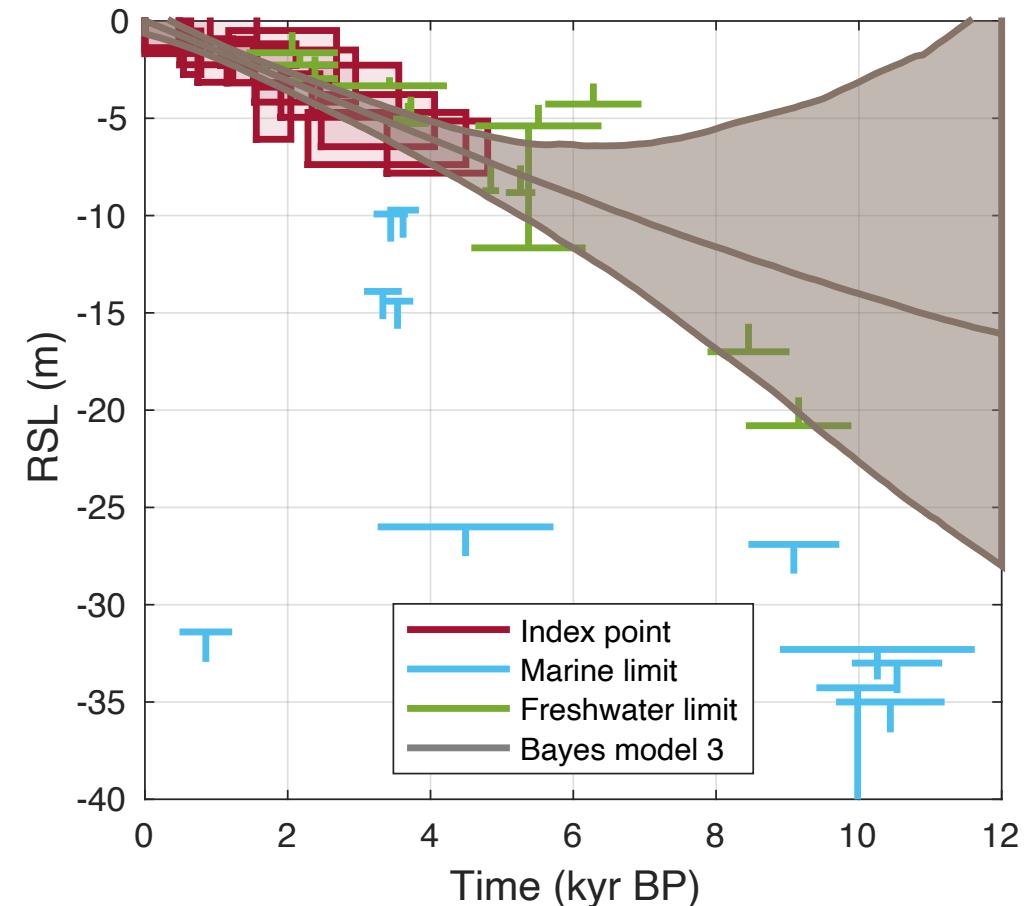
Model 3: nonlinear/nonparametric process, imperfect observations, without limiting data

Now we determine a variable rate of sea-level change over the Holocene

Rates are 1.5 ± 0.6 , 1.5 ± 0.7 , and $1.2 \pm 1.9 \text{ mm yr}^{-1}$ at 0, 4, and 8 kyr

Model more completely represents the structure of the index points, and the uncertainties characterizing them

But we're still omitting limiting dates!



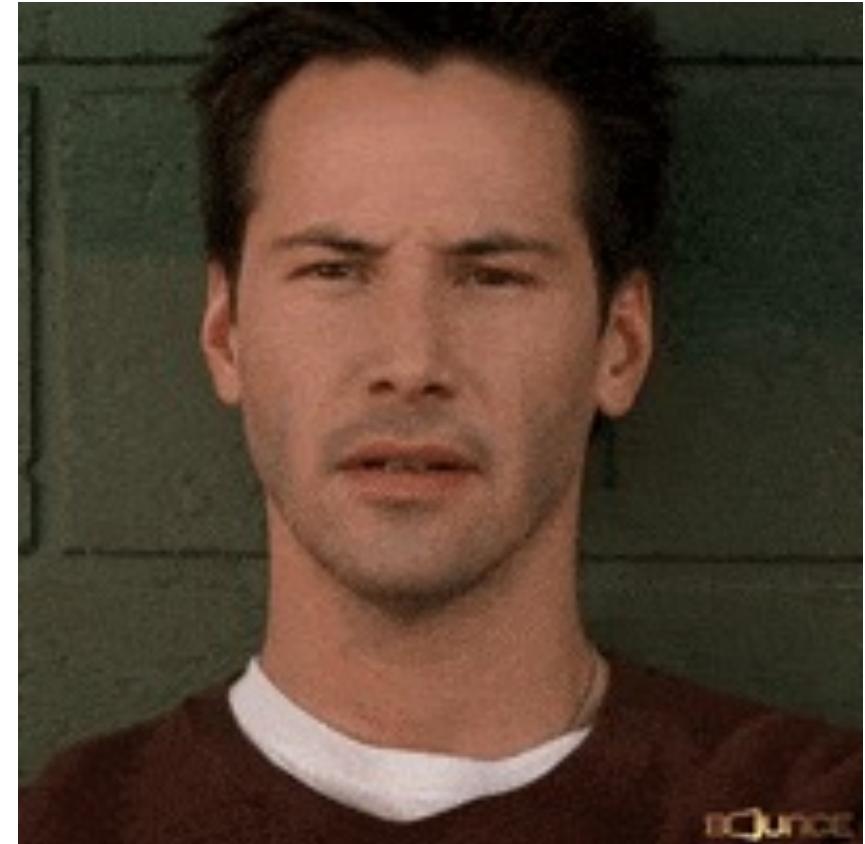
Model 3: nonlinear/nonparametric process, imperfect observations, without limiting data

Hold up. We didn't specify a function for sea level, so how did I just estimate sea-level values at regular time points?

To do this, I used posterior prediction: knowing sea levels y and times x , I can estimate sea levels y^* at other times x^*

$$p(y^*|x, y, x^*) = N \left(C(x^*, x) C(x, x)^{-1} y, \rho^2 C(x^*, x^*) - \rho^4 C(x^*, x) [\rho^2 C(x, x) + \lambda^2 I]^{-1} C(x, x^*) \right)$$

This follows from Gaussian identities and is at the core of Gaussian Process regression



**Fourth model: nonlinear/nonparametric
process, imperfect data, with limiting data**

Model 4: nonlinear/nonparametric process, imperfect observations, with limiting data

Last model! Almost there!

Adjust data equation and add process equations for marine limit m and freshwater limit f

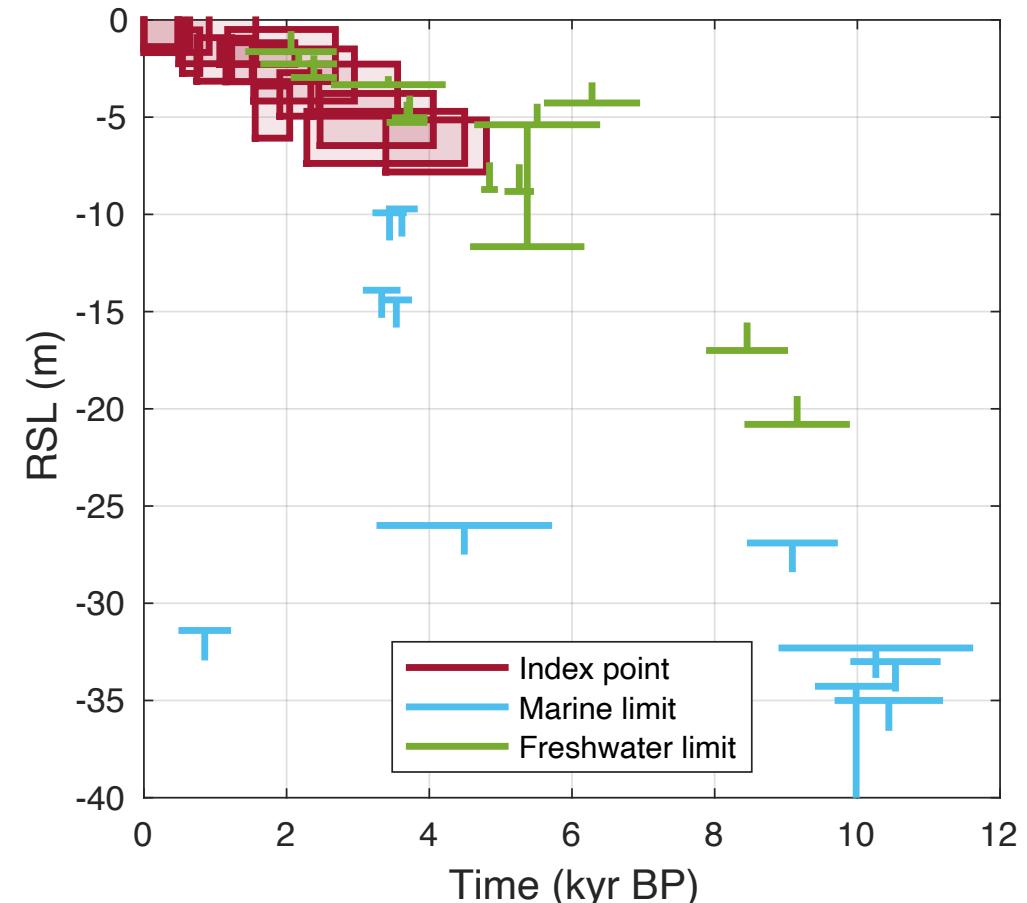
Data—

$$p(z_n) = \begin{cases} N(m_n, \delta_n^2) \\ N(y_n, \delta_n^2) \\ N(f_n, \delta_n^2) \end{cases}$$

Process—

$$p(m_n) = U(-\infty, y_n)$$

$$p(f_n) = U(y_n, +\infty)$$



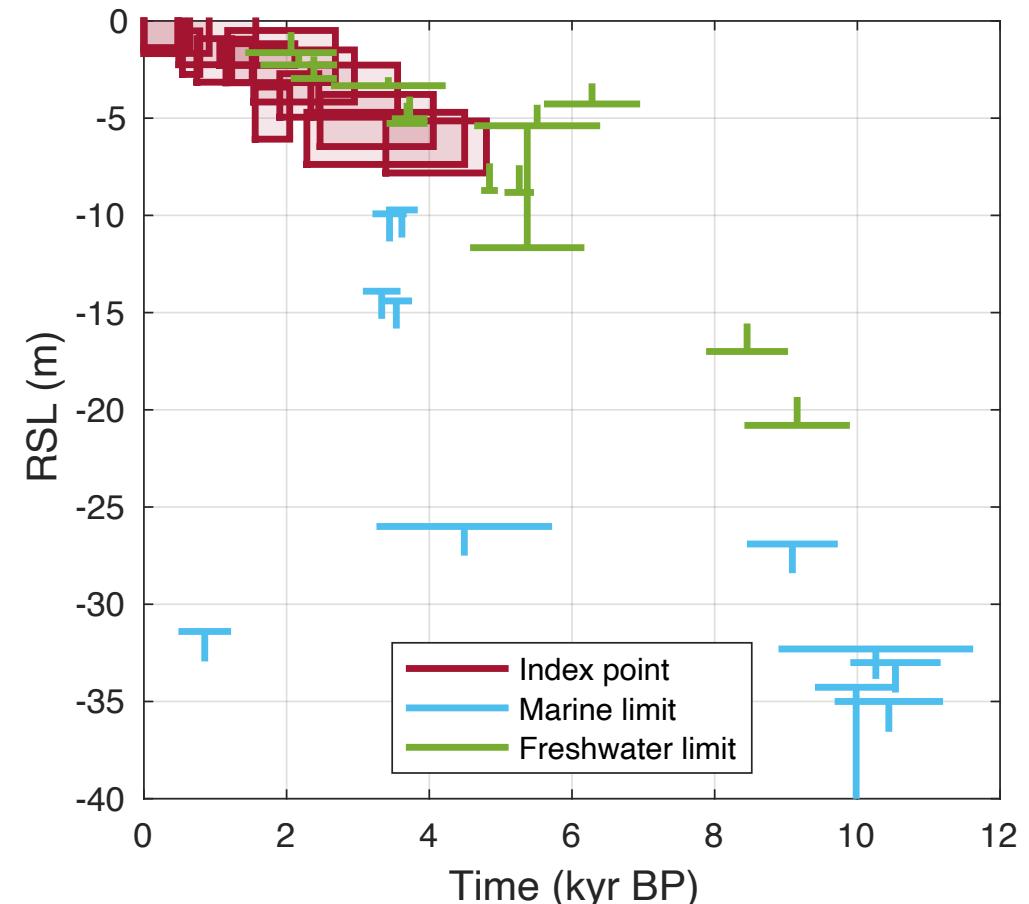
Model 4: nonlinear/nonparametric process, imperfect observations, with limiting data

Final posterior distribution is

$$p(\lambda^2, \rho^2, \varphi, x_1, \dots, y_1, \dots, m_1, \dots, f_1, \dots | w_1, \dots, z_1, \dots)$$

$$\propto p(\lambda^2)p(\rho^2)p(\varphi) \prod_n [p(w_n|x_n) \\ \times p(z_n|y_n, m_n, f_n) p(m_n|y_n) \\ \times p(f_n|y_n) p(y_n|x, y_{m \neq n}, \lambda^2, \rho^2, \varphi)]$$

We use Metropolis algorithm and posterior prediction to generate solutions at regular points in time

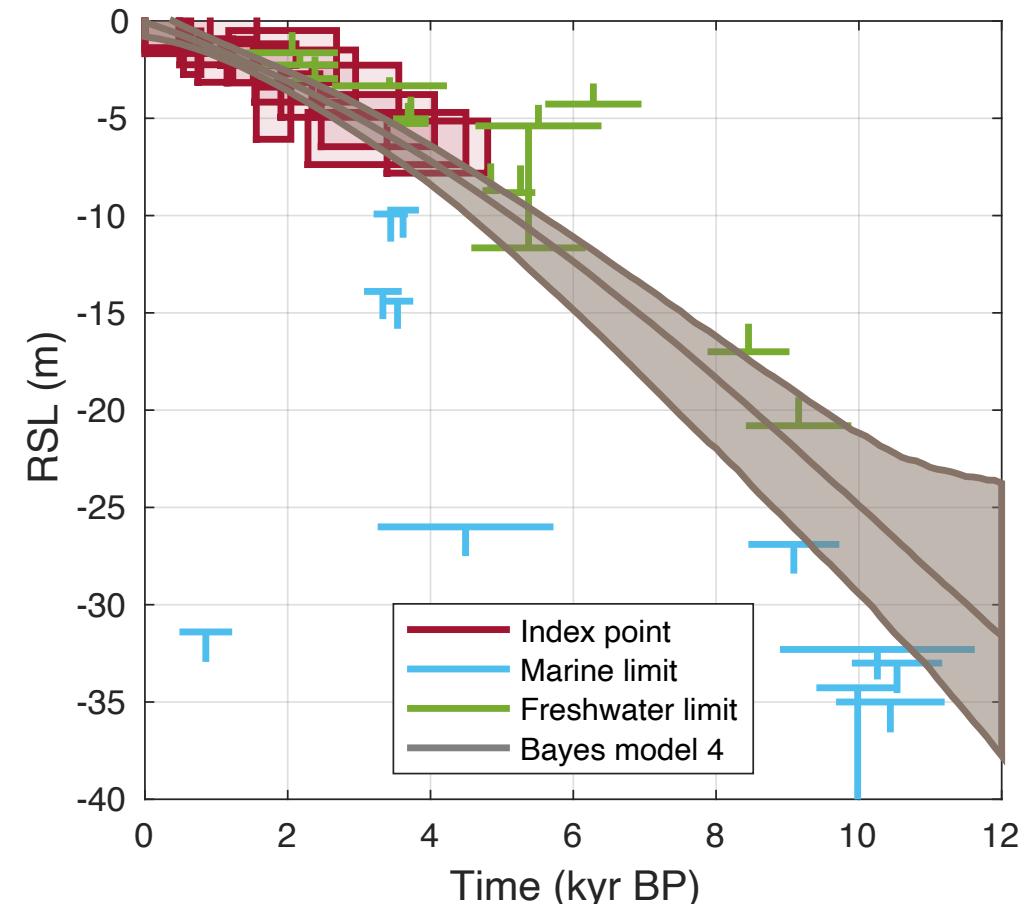


Model 4: nonlinear/nonparametric process, imperfect observations, with limiting data

Now we determine variable rates of sea-level change over the Holocene that are consistent with all data types

Model shows higher rates earlier in time, which is more consistent with our physical intuition

Rates are 1.3 ± 0.8 , 2.4 ± 0.6 , and $3.1 \pm 0.9 \text{ mm yr}^{-1}$ at 0, 4, and 8 kyr

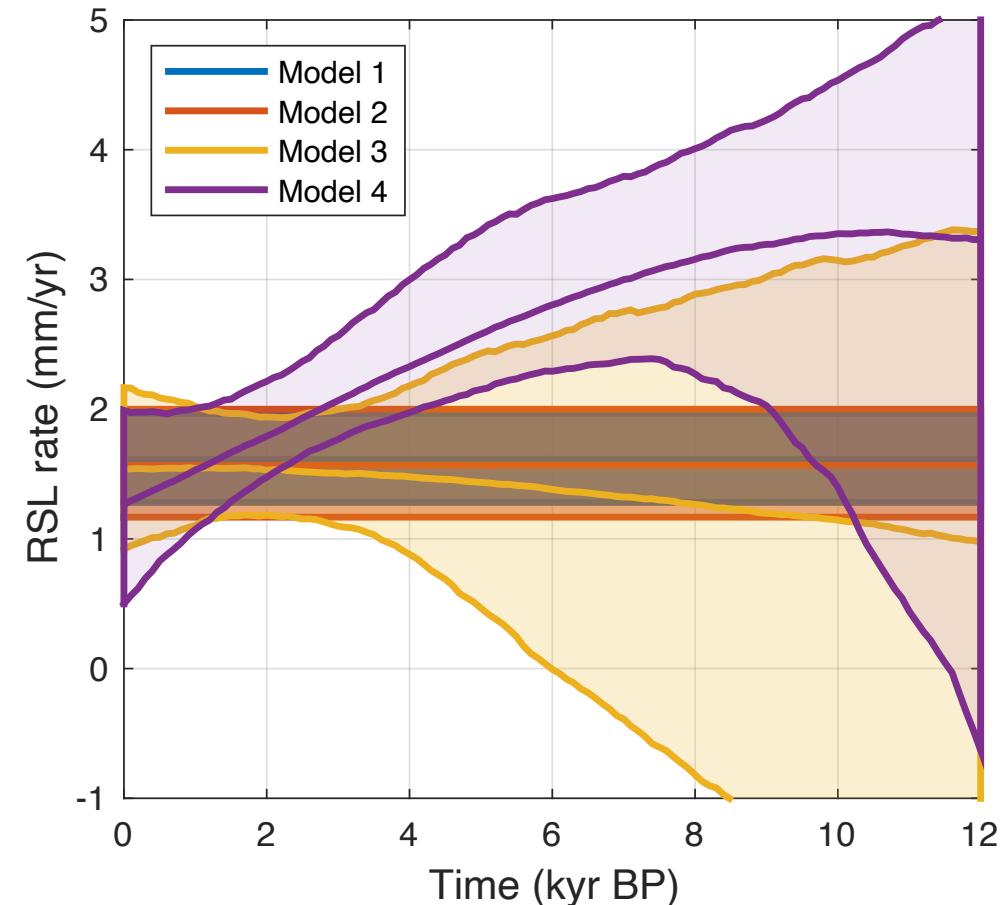


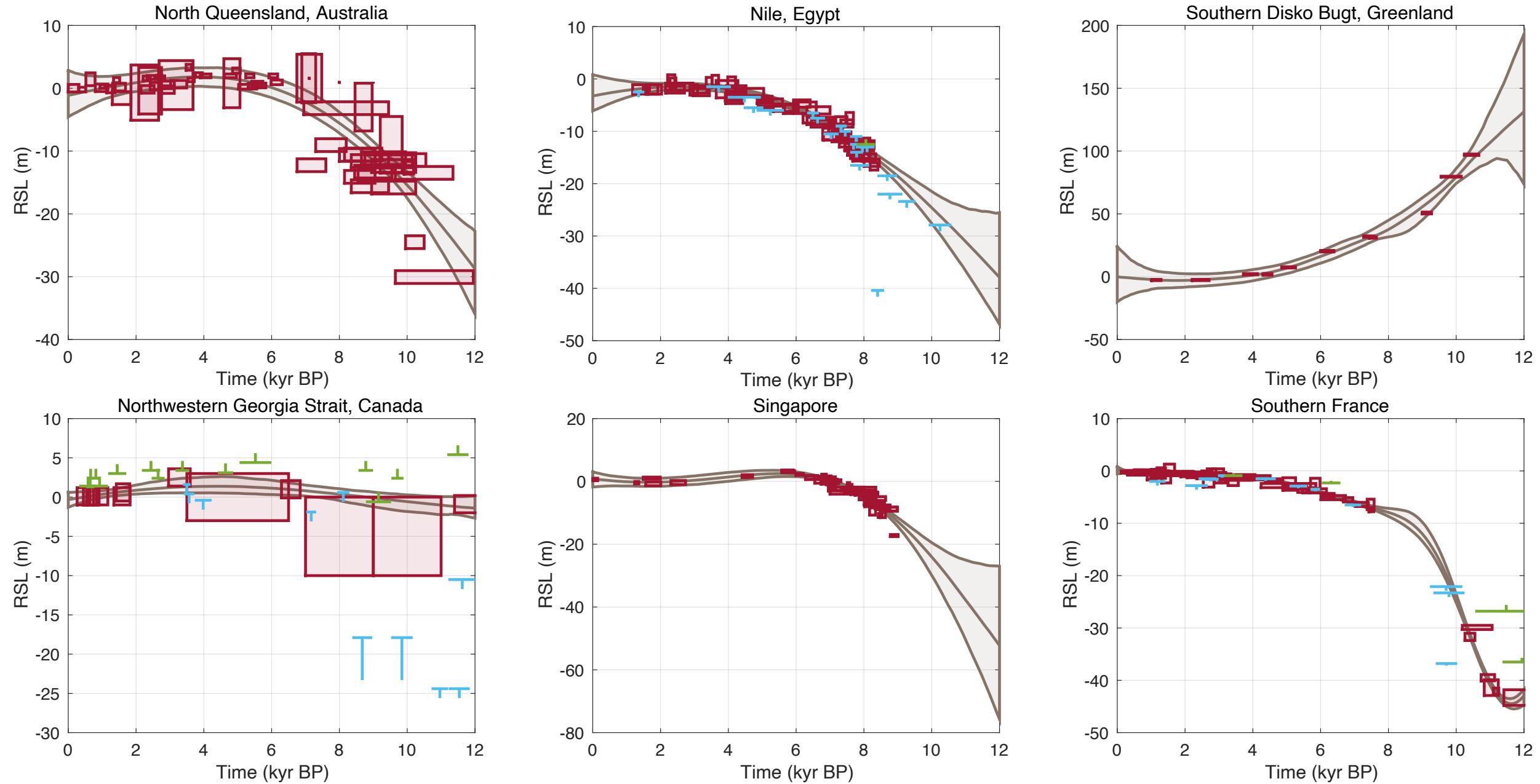
Model 4: nonlinear/nonparametric process, imperfect observations, with limiting data

Now we determine variable rates of sea-level change over the Holocene that are consistent with all data types

Model shows higher rates earlier in time, which is more consistent with our physical intuition

Rates are 1.3 ± 0.8 , 2.4 ± 0.6 , and $3.1 \pm 0.9 \text{ mm yr}^{-1}$ at 0, 4, and 8 kyr





Summary

Bayesian modeling involves thinking about problems probabilistically, hierarchically, and conditionally

This approach offers a flexible framework for analyzing sea-level reconstructions and quantifying uncertainty

You don't need to be a statistician to understand these methods and use them informedly in your own research

Summary

codes/slides available at

github.com/christopherpiecuch/bayesGavle

get/stay in touch

cpiecuch@whoi.edu

twitter.com/chrisspiecuch

Some helpful (and free!) textbook on Bayesian methods and spatiotemporal statistics

Cressie and Wikle (2019)

<https://spacetimewithr.org/>

Gelman et al. (2013)

<http://www.stat.columbia.edu/~gelman/book/BDA3.pdf>

Rasmussen and Williams (2006)

<https://gaussianprocess.org/gpml/chapters/RW.pdf>

Examples of Bayesian methods in sea level and glacial isostatic adjustment literatures

Ashe et al. (2019) <https://www.sciencedirect.com/science/article/pii/S0277379118302130>

Ashe et al. (2022) <https://ascmo.copernicus.org/articles/8/1/2022>

Cahill et al. (2015) <https://www.jstor.org/stable/24522592>

Cahill et al. (2016) <https://cp.copernicus.org/articles/12/525/2016>

Creel et al. (2022) <https://doi.org/10.1016/j.quascirev.2022.107422>

Hay et al. (2015) <https://www.nature.com/articles/nature14093>

Hay et al. (2013) <https://www.pnas.org/doi/10.1073/pnas.1117683109>

Khan et al. (2015) <https://doi.org/10.1007/s40641-015-0029-z>

Khan et al. (2019) <https://doi.org/10.1016/j.quascirev.2019.07.016>

Kopp et al. (2016) <https://www.pnas.org/doi/10.1073/pnas.1517056113>

Kopp (2013) <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/grl.50781>

Kopp et al. (2009) <https://www.nature.com/articles/nature08686>

Piecuch et al. (2018) <https://www.nature.com/articles/s41586-018-0787-6>

Open-source code (from other folks)

Erica Ashe <https://github.com/ericaashe>

Niamh Cahill <https://github.com/ncahill89>

Bob Kopp <https://github.com/bobkopp>

Maeve Upton <https://maeveupton.github.io/reslr/>

A photograph of a large, light-colored boulder resting on a rocky ground. The background consists of many bare, leafless trees, suggesting a late autumn or winter setting. The lighting is natural, with sunlight filtering through the branches.

Tack! Fråga?