Gaussian Processes for Systematic Mitigation

A Bayesian Method to Mitigate the Effects of Unmodelled Time-Varying Systematics for Global 21-cm Cosmology Experiments

Christian J. Kirkham[1,2], Dominic J. Anstey[1,2], Eloy de Lera Acedo[1,2]

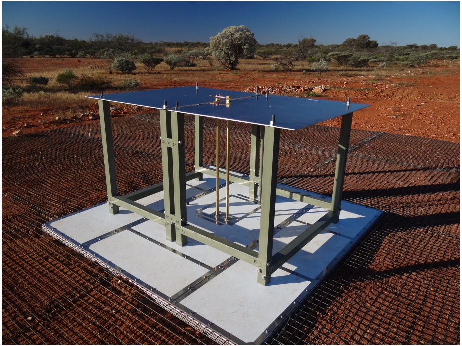
[1]Cavendish Astrophysics, University of Cambridge

[2]Kavli Institute for Cosmology, University of Cambridge

URSI AT-RASC, 23 May 2024

Table of Contents

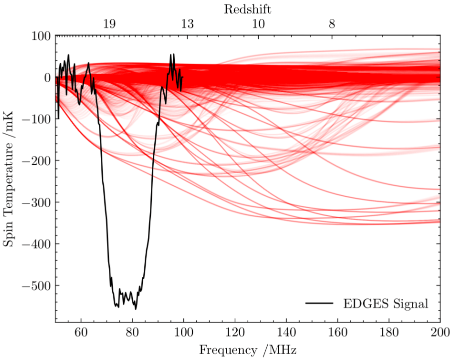
EDGES (Experiment to Detect the Global EoR Signature)



- EDGES is a low frequency radio experiment to detect the global sky-averaged 21-cm signal
- Collaboration between ASU and MIT
- Located in the Murchison Radio-astronomy Observatory in Western Australia

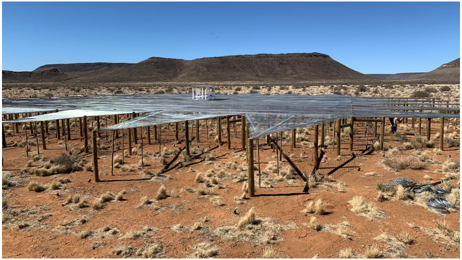
REACH

EDGES (Experiment to Detect the Global EoR Signature)



- Claimed detection of 21-cm signal at 78 MHz (Bowman+18)
- Unusually deep signal, requiring exotic physics
- Concerns that there is a residual systematic in the data (Hills+18, Sims and Pober 2020)

REACH (Radio Experiment for the Analysis of Cosmic Hydrogen)

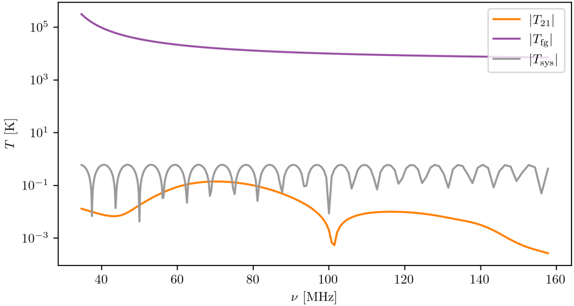


Credit: Saurabh Pegwal

- REACH is a global signal experiment designed to verify the EDGES detection
- Cambridge University, Stellenbosch University and others
- Located in the Karoo Desert on the SARAO facility in South Africa

REACH

Global Experiment Challenges



Credit: Harry Bevins

- **Galactic foregrounds**
  Smooth synchrotron emission
  $\sim 10^5$ K
- **Instrumental systematics**
  Generally sinusoidal
  $\sim 10^0$ K
- **21-cm Signal**
  Not spectrally smooth
  $\sim 10^{-2}$ K

Time Varying Systematics in the REACH System

- Some systematics are expected to be static, others vary with time
  - Reflections from the soil vary with rainfall (Bevins+21, 22)
  - Impedences in the system are temperature dependent
  - Improperly modelled beams can cause systematics which vary with the galactic foreground

REACH

Modelling Time Varying Systematics

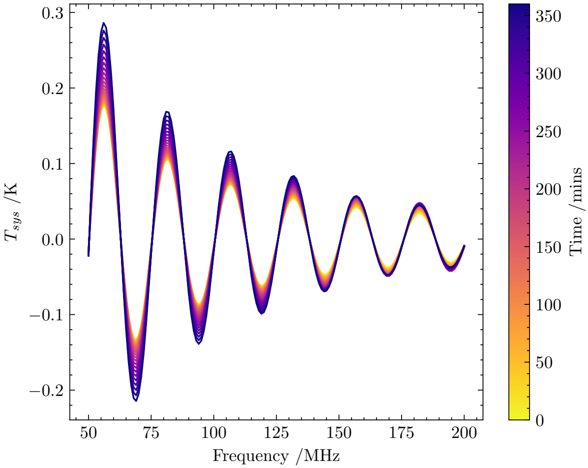
- Want to simulate a systematic to test robustness of data analysis techniques
- Model systematic as a sinusoid whose amplitude is modulated by the power of the galactic foreground

$$T_{\mathrm{sys}}(\nu) = A_{\mathrm{sys}} \left(\frac{\nu}{\nu_{0,\mathrm{sys}}} \right)^{-\alpha_{\mathrm{sys}}} \sin\left(\frac{2\pi\nu}{P_{\mathrm{sys}}} + \phi_{\mathrm{sys}} \right) \tag{1}$$

- Systematic amplitude is varied as

$$A_{\mathrm{sys}}(t_j) = A_{\mathrm{sys}}(t_0) \cdot \frac{T_{\mathrm{fg}}(\nu = \nu_0, t = t_j)}{T_{\mathrm{fg}}(\nu = \nu_0, t = t_0)} \tag{2}$$

Modelling Time Varying Systematics



Standard REACH Pipeline

- Fully Bayesian forward model of antenna beam, galactic foregrounds and 21-cm signal (Anstey+21, 22)
- Uses a Gaussian likelihood with uncorrelated noise, σ_0

$$\log \mathcal{L} = \sum_i -\frac{1}{2}\log(2\pi\sigma_0^2) - \frac{1}{2}\left(\frac{T_{\text{data}}(\nu_i) - (T_{\text{fg}}(\nu_i) + T_{21}(\nu_i) + T_{\text{CMB}})}{\sigma_0}\right)^2 \tag{3}$$

- Will be referred to as the "Standard Pipeline"

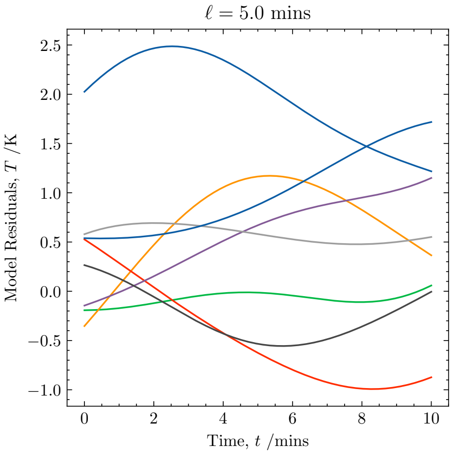
Gaussian Processes

- Define a GP as a collection of random variables which have consistent joint Gaussian distributions (Rasmussen 2004)
- Distribution is defined by the covariance matrix, or covariance function

$$\mathbf{C}_{ij} = K(t_i, t_j) \tag{4}$$

- K is called the "kernel" – form can be chosen freely

Squared Exponential Kernel

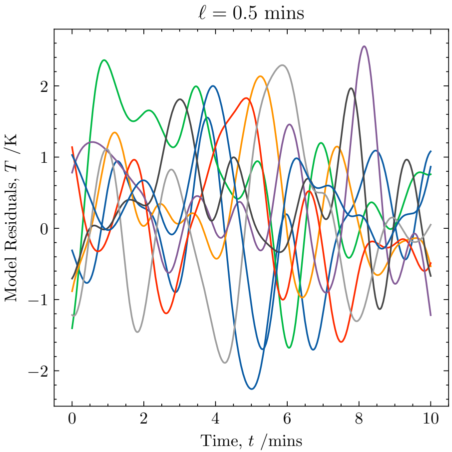


- In this work I use a squared exponential kernel with a white noise term

$$\mathbf{C}_{ij} = \sigma^2_{0,\mathrm{GP}} + \sigma^2_{\mathrm{SE}} \exp\left(-\frac{|t_i - t_j|^2}{2\ell^2}\right) \tag{5}$$

- Uncorrelated noise, $\sigma_{0,\mathrm{GP}}$, scale factor of the kernel, σ_{SE}, and the covariance length, ℓ
- Define smooth functions

Squared Exponential Kernel



- In this work I use a squared exponential kernel with a white noise term

$$\mathbf{C}_{ij} = \sigma^2_{0,\mathrm{GP}} + \sigma^2_{\mathrm{SE}} \exp\left(-\frac{|t_i - t_j|^2}{2\ell^2}\right) \quad (5)$$

- Uncorrelated noise, $\sigma_{0,\mathrm{GP}}$, scale factor of the kernel, σ_{SE}, and the covariance length, ℓ
- Define smooth functions

Gaussian Processes in the REACH Pipeline

- Hereafter called the “Gaussian Process pipeline” (Kirkham+24)

$$\mathcal{L}(\theta) = \frac{1}{\sqrt{(2\pi)^n|\mathbf{C}|}} \exp\left(-\frac{1}{2}(\mathbf{D} - \mathbf{M}(\theta))^T \boldsymbol{C}^{-1}(\mathbf{D} - \mathbf{M}(\theta))\right) \quad (6)$$

- **D** is the data vector, **M** is the model (foreground, beam model, signal model)
- Gaussian process covariance matrix, **C**, fits covariance of model residuals between time bins

Simulating Data

- We model the global signal as a Gaussian

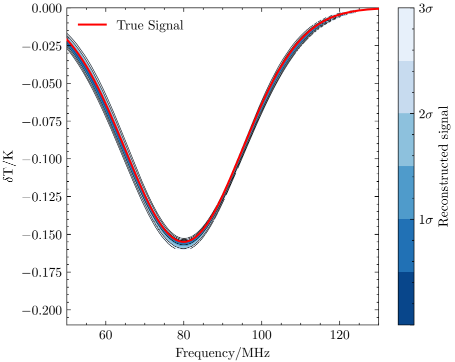
$$T_{sg}(\nu) = -A_{21} \exp\left(-\frac{(\nu - \nu_{0,21})^2}{2\sigma_{21}^2}\right) \tag{7}$$

- Simulated data produced with "true" signal parameters $A_{21} = 0.155$ K, $\sigma_{21} = 15$ MHz and $\nu_{0,21} = 85$ MHz
- 21-cm signal is fitted for using same Gaussian model, with three free parameters: A_{21}, σ_{21} and $\nu_{0,21}$

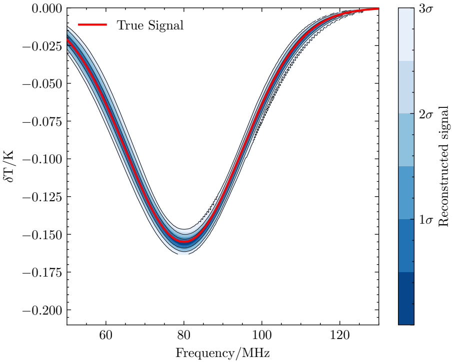
Varying Systematic parameters

- Will test both pipelines with a range of simulated systematics
- Data generated for 24 time bins of length 15 minutes – 6 hours of observation in total
- Posterior distribution was then sampled using PolyChord

Comparison of Signal Posteriors - No Systematic

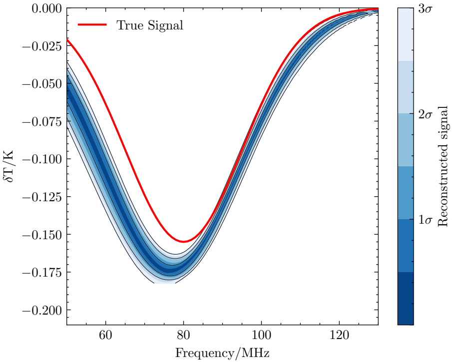


Standard pipeline

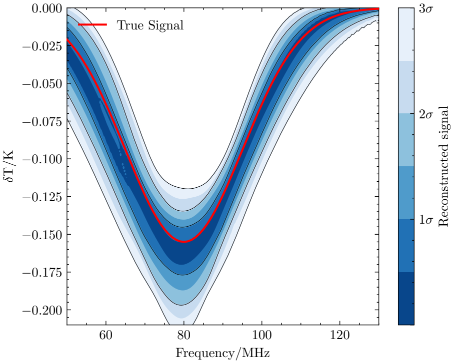


Gaussian Process pipeline

Comparison of Signal Posteriors - With Systematic



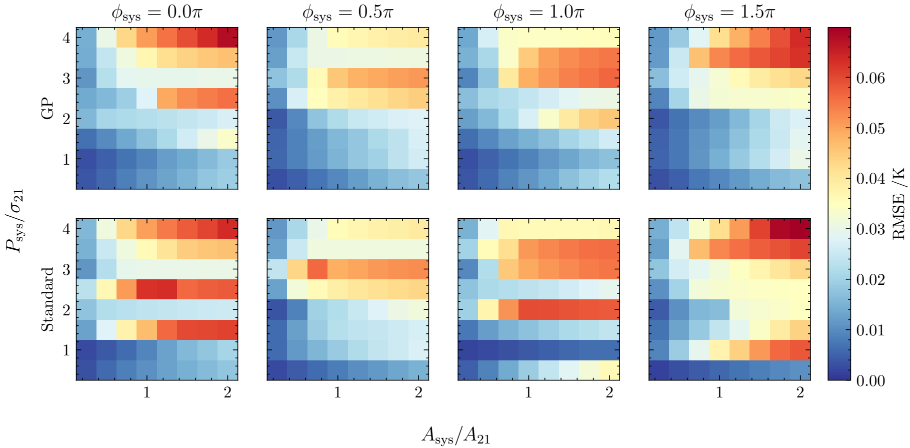
Standard pipeline



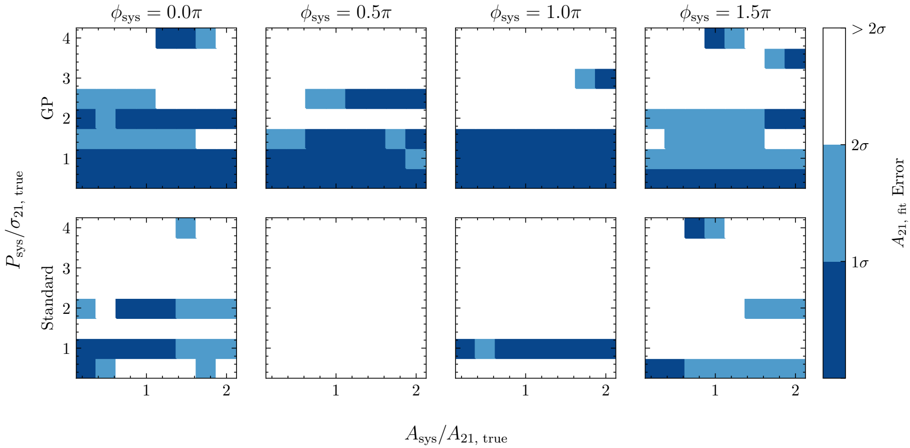
Gaussian Process pipeline

REACH

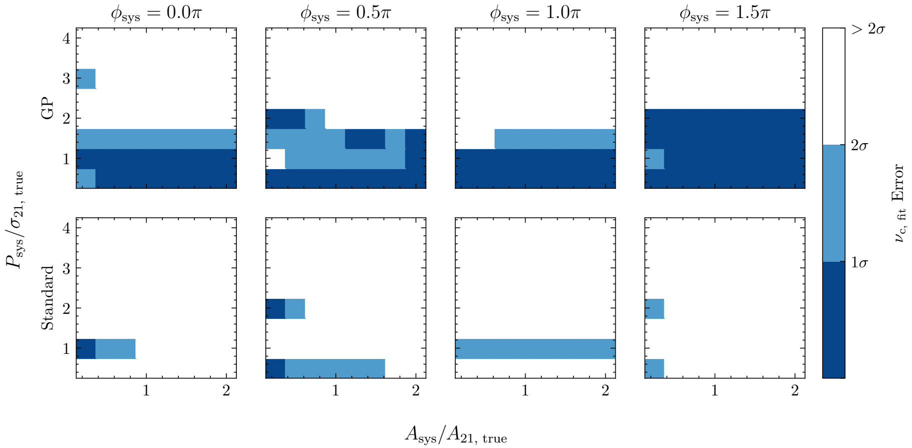
RMSE


$\phi_{sys} = 0.0\pi$
$\phi_{sys} = 0.5\pi$
$\phi_{sys} = 1.0\pi$
$\phi_{sys} = 1.5\pi$
GP
Standard
P_{sys}/σ_{21}
A_{sys}/A_{21}
RMSE /K
0.00
0.01
0.02
0.03
0.04
0.05
0.06

Fitted Signal Amplitude error



Fitted Centre Frequency error



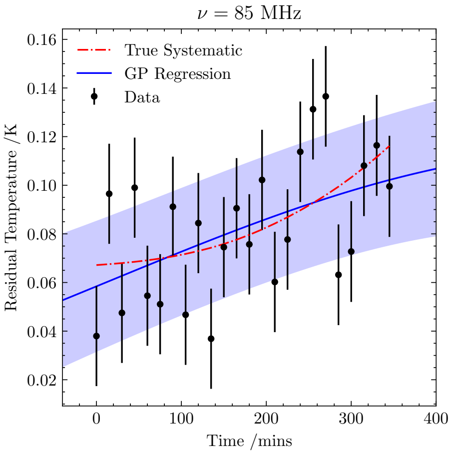
Gaussian Process Regression

- Can also use Gaussian Processes for regression
- Mean function of the Gaussian Process posterior given by

$$\mu(\mathbf{t}_{\text{pred}}) = K(\mathbf{t}_{\text{pred}}, \mathbf{t}_{\text{data}})K(\mathbf{t}_{\text{data}}, \mathbf{t}_{\text{data}})^{-1}\mathbf{T}_{\text{data}}, \tag{8}$$

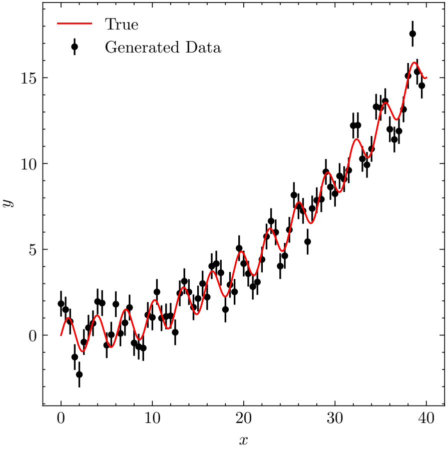
- Can be used to predict temperature, $\mathbf{T}_{\text{pred}}$, of model residuals at some time, $\mathbf{t}_{\text{pred}}$

Gaussian Process Regression



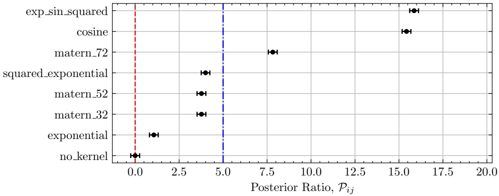
- Use weighted mean Gaussian Process hyperparameters to get a smooth fit to model residuals
- Could possibly be used to inform future time-varying systematic models

Automatic Kernel Selection



- Kernel choice is arbitrary so can use Bayes factor to inform us (Hee+15, Kroupa+24)
- Basic test with a quadratic curve with a sinusoidal residual
- Fit for quadratic but not sinusoid

Automatic Kernel Selection



- Uses PolyChord to sample over all kernels using a choice parameter, c
- Uses the posterior ratio of c as a proxy for Bayes Factor

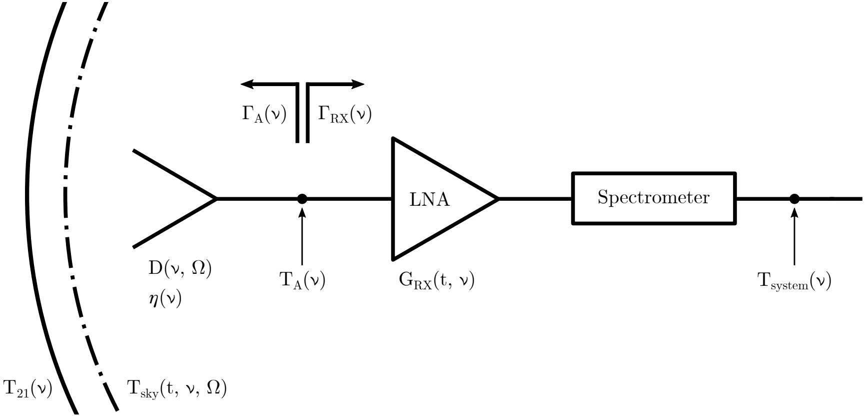
$$\mathcal{P}_{ij} = \log \frac{\Pr(c = j|\mathbf{D}, \mathbf{M})}{\Pr(c = i|\mathbf{D}, \mathbf{M})} \quad (9)$$

Summary

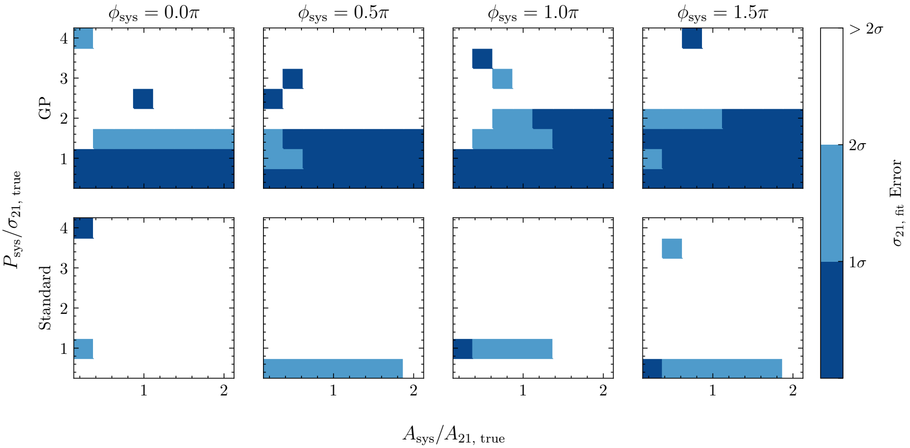
- Using Gaussian Processes to account for time correlated residuals improves fitting
- General method – no systematic model required
- Regression can inform future models of systematics
- Automatic Kernel Selection can use the data to select the most appropriate kernel

For more information

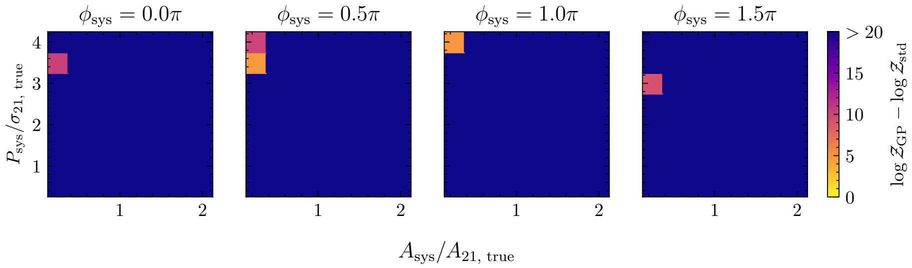
Antenna and Receiver System



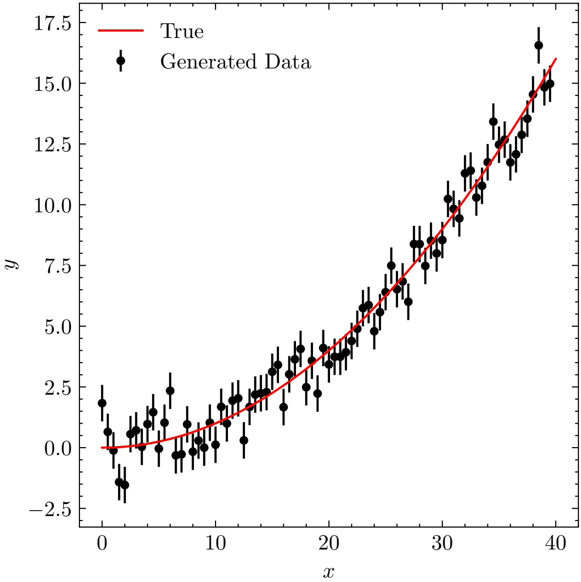
Signal Width



Bayes Factor



Automatic Kernel Selection - No Residual



Automatic Kernel Selection - No Residual