An automatic report for the dataset : monthly-critical-radio-frequenci

James Robert Lloyd
University of Cambridge

David Duvenaud
University of Cambridge

Roger Grosse
Massachussets Institute of Technology

Joshua B. Tenenbaum
Massachussets Institute of Technology

Zoubin Ghahramani
University of Cambridge

Abstract

This report was produced automatically by the Gaussian process structure search algorithm. See `http://arxiv.org/abs/1302.4922` for a preliminary paper and see `https://github.com/jamesrobertlloyd/gpss-research` for the latest source code.

1 Executive summary

The raw data and full model posterior with extrapolations are shown in figure 1.

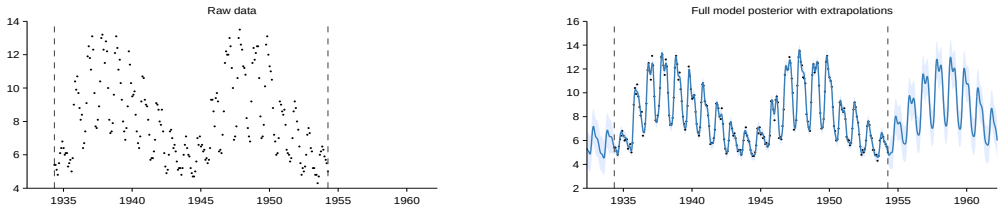

Figure 1: Raw data (left) and model posterior with extrapolation (right)

The structure search algorithm has identified six additive components in the data:

- A constant.
- An approximately sinusoidal function with a period of 9.9 years.
- An exactly periodic function with a period of 1.0 years.
- An approximate product of a periodic function and a sinusoid.
- A smooth function.
- A very approximately sinusoidal function with a period of 9.9 years.

#	R^2 (%)	ΔR^2 (%)	Residual R^2 (%)	Cross validated MAE	Reduction in MAE (%)
-	-	-	-	8.08	-
1	-0.0	-0.0	-0.0	2.06	74.5
2	53.7	53.7	53.7	1.40	32.0
3	87.5	33.8	72.9	0.80	42.9
4	92.7	5.3	41.9	0.64	19.3
5	99.0	6.3	86.4	0.64	0.7
6	99.1	0.1	11.6	0.64	-0.1

Table 1: Summary statistics for cumulative additive fits to the data. The residual coefficient of determination (R^2) values are computed using the residuals from the previous fit as the target values; this measures how much of the residual variance is explained by each new component. The mean absolute error (MAE) is calculated using 10 fold cross validation with a contiguous block design; this measures the ability of the model to interpolate and extrapolate over moderate distances. The model is fit using the full data so the MAE values cannot be used reliably as an estimate of out-of-sample predictive performance.

2 Detailed discussion of additive components

2.1 Component 1

This component is constant.

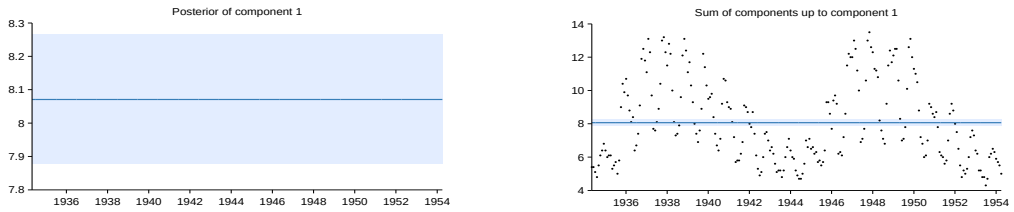

Figure 2: Posterior of component 1 (left) and the posterior of the cumulative sum of components with data (right)

2.2 Component 2

This component is approximately sinusoidal with a period of 9.9 years. Across periods the shape of the function varies very smoothly.

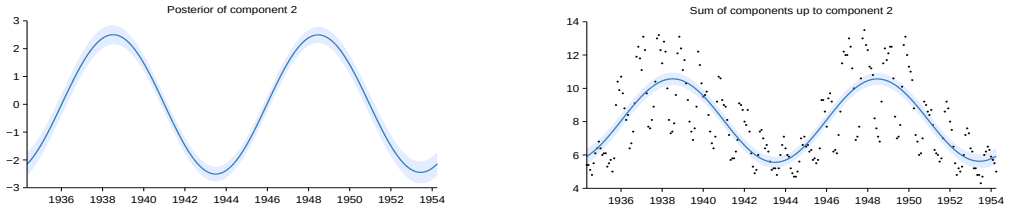

Figure 3: Posterior of component 2 (left) and the posterior of the cumulative sum of components with data (right)

2.3 Component 3

This component is exactly periodic with a period of 1.0 years. The shape of the function within each period has a typical lengthscale of 3.3 months.

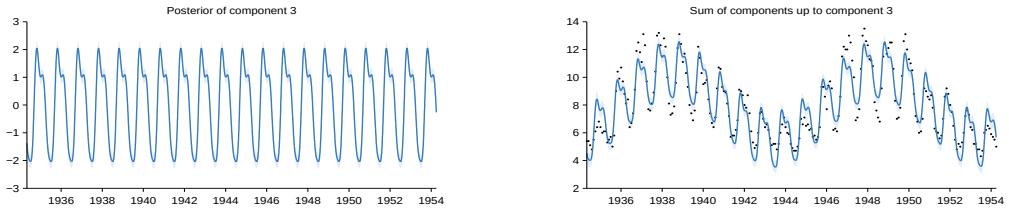

Figure 4: Posterior of component 3 (left) and the posterior of the cumulative sum of components with data (right)

2.4 Component 4

This component is a product of a periodic function and a sinusoid. Across periods the shape of the function varies smoothly with a typical lengthscale of 44.5 years. The periodic function has a period of 1.0 years. The shape of this function within each period has a typical lengthscale of 3.3 months. The sinusoid has a period of 9.9 years.

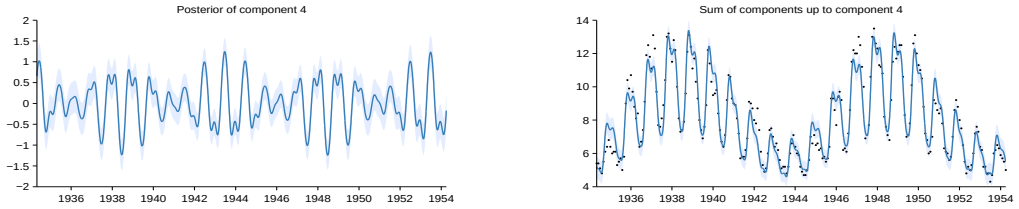

Figure 5: Posterior of component 4 (left) and the posterior of the cumulative sum of components with data (right)

2.5 Component 5

This component is a smooth function with a typical lengthscale of 2.1 months.

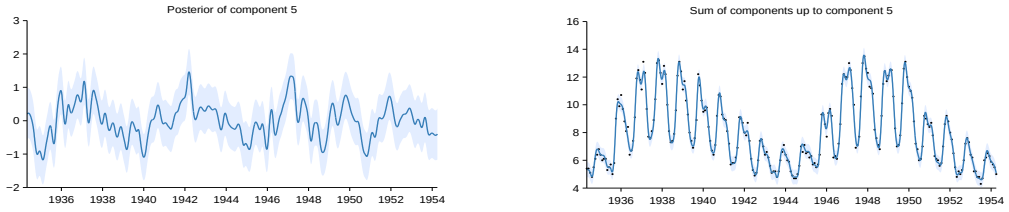

Figure 6: Posterior of component 5 (left) and the posterior of the cumulative sum of components with data (right)

2.6 Component 6

This component is very approximately sinusoidal with a period of 9.9 years. Across periods the shape of the function varies smoothly with a typical lengthscale of 2.1 months. Since this lengthscale is small relative to the period this component may more closely resemble a non-periodic smooth function.

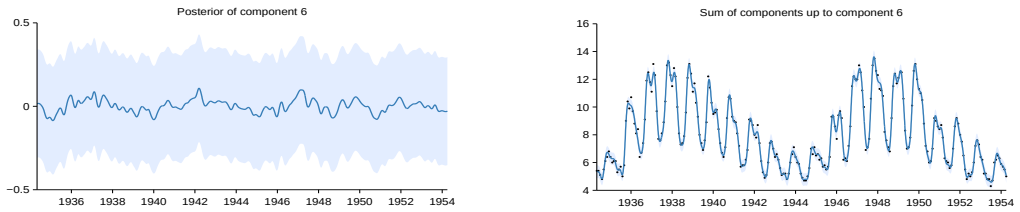

Figure 7: Posterior of component 6 (left) and the posterior of the cumulative sum of components with data (right)

3 Extrapolation

Summaries of the posterior distribution of the full model are shown in figure 8. The plot on the left displays the mean of the posterior together with pointwise variance. The plot on the right displays three random samples from the posterior.

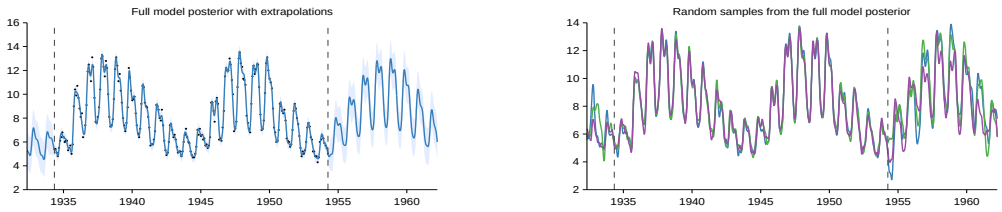

Figure 8: Full model posterior. Mean and pointwise variance (left) and three random samples (right)

3.1 Component 1 : A constant

Some discussion about extrapolation.

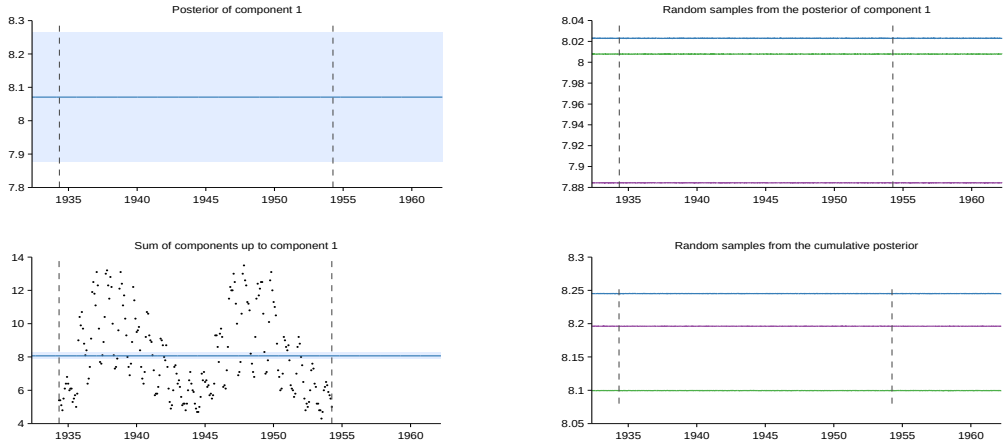

Figure 9: Posterior of component 1. Mean and pointwise variance (left) and three random samples from this distribution (right)

3.2 Component 2 : An approximately sinusoidal function with a period of 9.9 years

Some discussion about extrapolation.

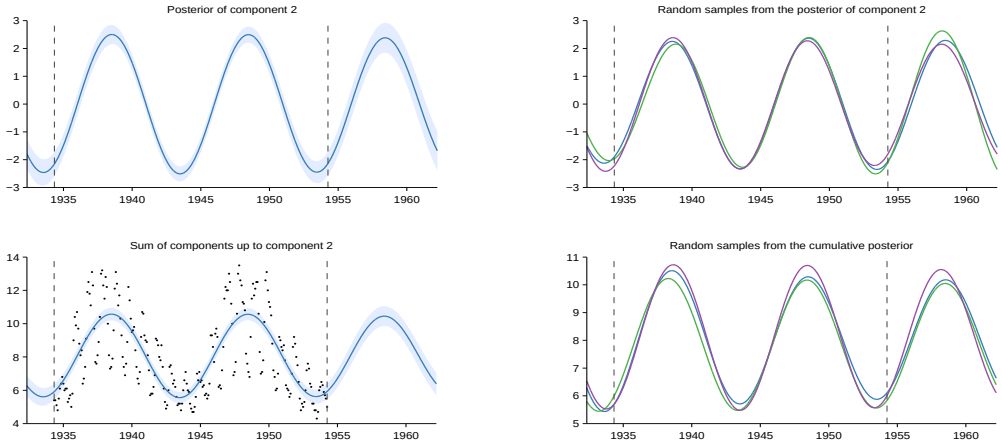

Figure 10: Posterior of component 2. Mean and pointwise variance (left) and three random samples from this distribution (right)

3.3 Component 3 : An exactly periodic function with a period of 1.0 years

Some discussion about extrapolation.

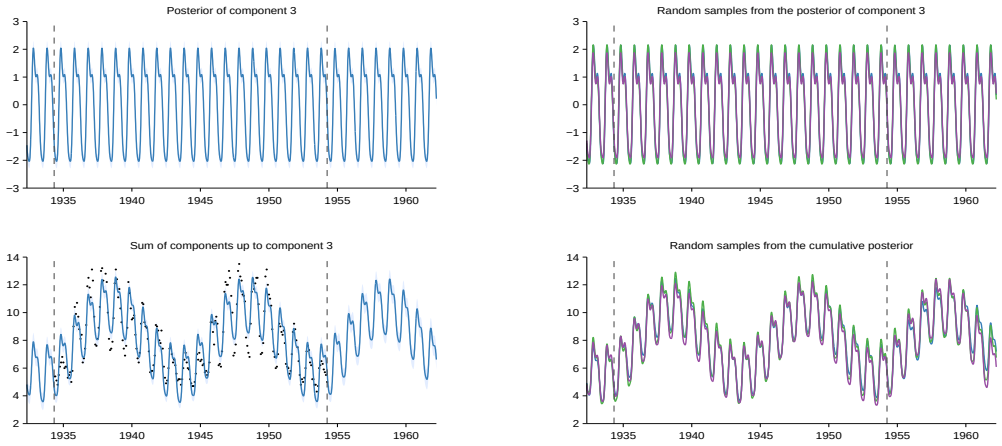

Figure 11: Posterior of component 3. Mean and pointwise variance (left) and three random samples from this distribution (right)

3.4 Component 4 : An approximate product of a periodic function and a sinusoid

Some discussion about extrapolation.

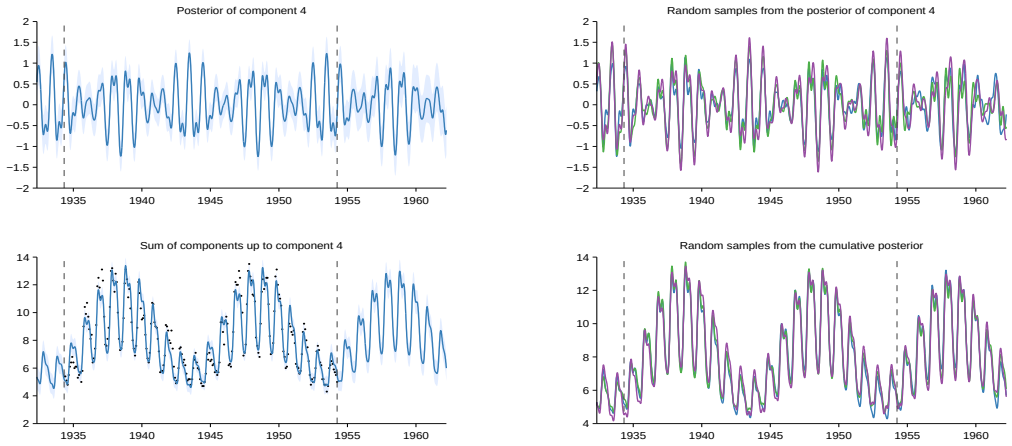

Figure 12: Posterior of component 4. Mean and pointwise variance (left) and three random samples from this distribution (right)

3.5 Component 5 : A smooth function

Some discussion about extrapolation.

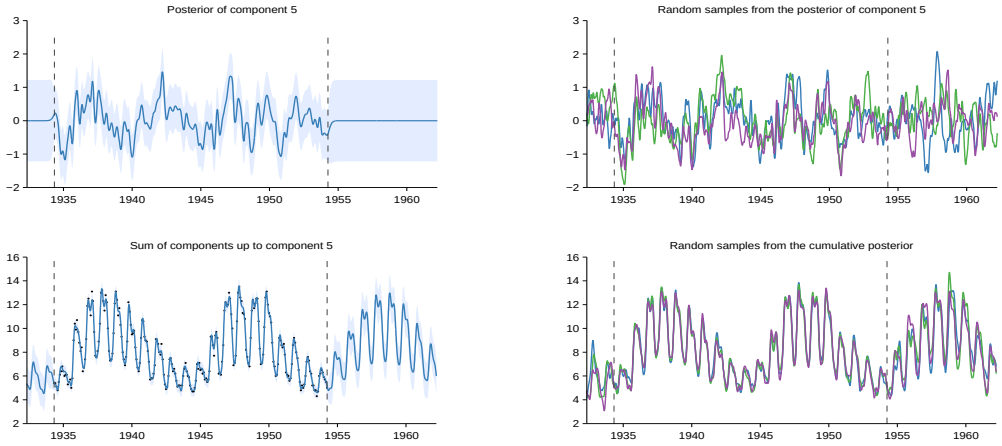

Figure 13: Posterior of component 5. Mean and pointwise variance (left) and three random samples from this distribution (right)

3.6 Component 6 : A very approximately sinusoidal function with a period of 9.9 years

Some discussion about extrapolation.

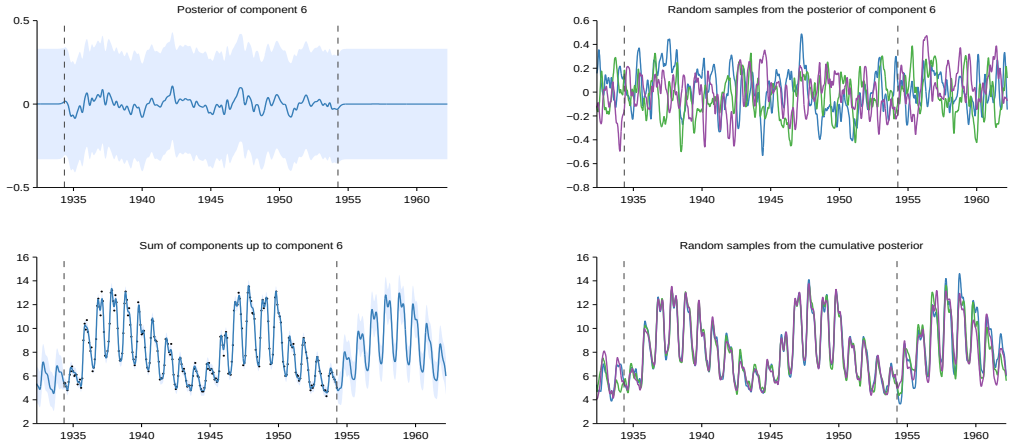

Figure 14: Posterior of component 6. Mean and pointwise variance (left) and three random samples from this distribution (right)