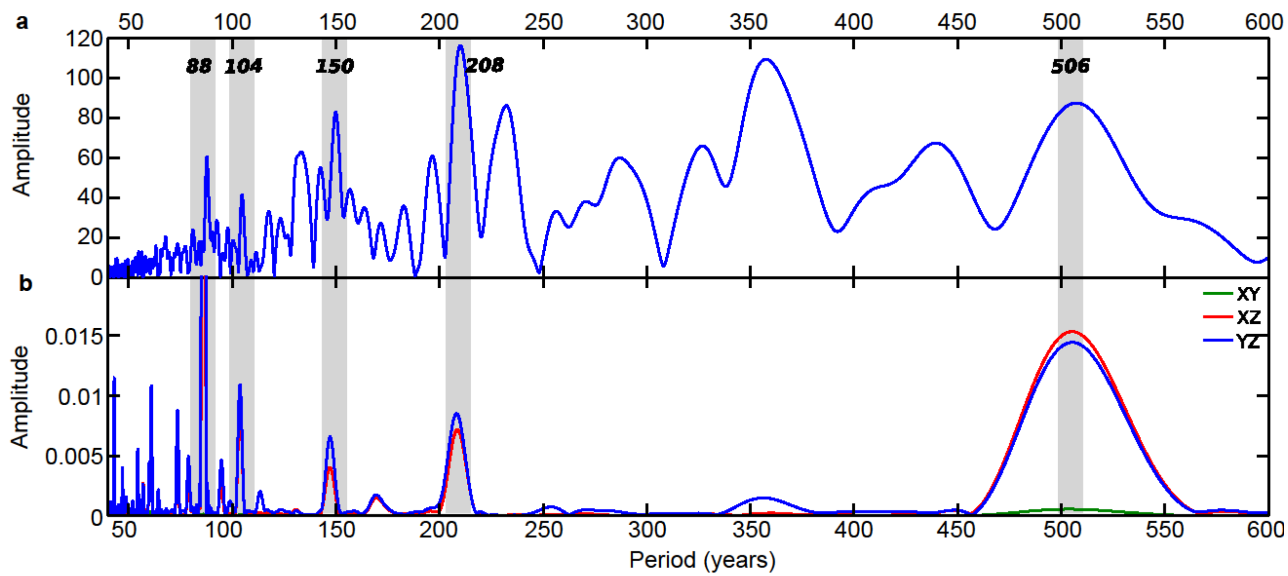
Bayesian inference for cycles in noisy time series

Thinkshop 2016, Dübendorf

Carlo Albert

Fourier Analysis



- Problem: Peaks have a **finite width**, even for pure harmonics, due to the finite observation window -> significance?
- This resolution has nothing to do with the **accuracy** lines can be determined with!
- How to measure accuracy -> **Bayesian statistics**!

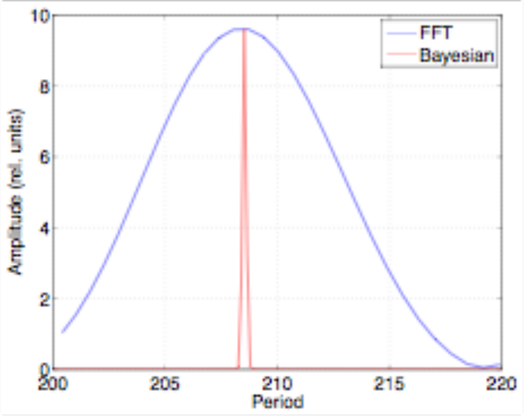
Bayesian Statistics

- Bayesian statistics conceptualises model-based **learning from data** and allows us to **quantify knowledge** of model parameters by means of **probability distributions**.
- We need **prior knowledge** in the form of
  - a **parameterised stochastic model,** for the data generating process, informed by our physical understanding of the system and/or analysis of the features of the data. Mathematically cast in the form of a **likelihood function** $L(\boldsymbol{y}|\boldsymbol{\theta})$.
  - a **probability distribution** $f_{\text{prior}}(\boldsymbol{\theta})$, expressing our prior knowledge about the parameters of the model.
- **Bayes' equation** tells us how to update knowledge about parameters when new data becomes available, which is believed to be a realisation of the model:

$$f_{post}(\boldsymbol{\theta}|\mathbf{y}) = \frac{L(\mathbf{y}|\boldsymbol{\theta})f_{prior}(\boldsymbol{\theta})}{\int L(\mathbf{y}|\boldsymbol{\theta}')f_{prior}(\boldsymbol{\theta}')d\boldsymbol{\theta}'}$$

The Bretthorst Method

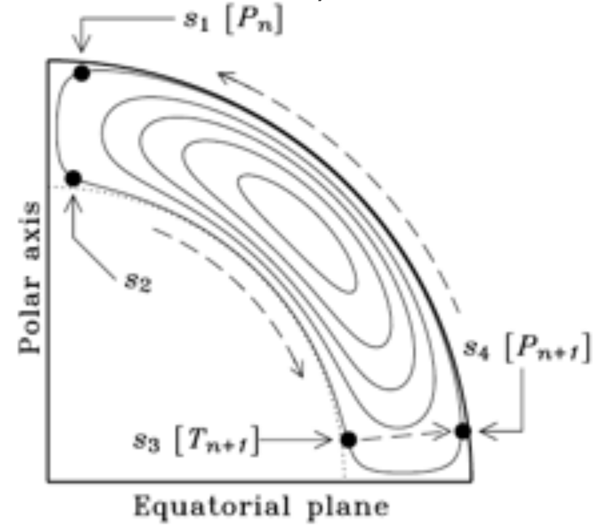
- **Data generating model**: Single harmonic(s) plus red noise (Palonen 2004),
  $\boldsymbol{\theta}$ = period, phase and amplitude of harmonic(s) and amplitude of red background noise.
- Allows us to determine lines (periods of harmonics) with **phenomenal accuracy**!



$\Delta^{14}C$	88.1 ± 0.2	104.3 ± 0.2	148.8 ± 0.4	208.7 ± 1.1	348.4 ± 2.2	503.7 ± 5.3
Planet. Torque	86.5 ± 0.2	103.9 ± 0.3	148.1 ± 0.5	208.5 ± 1.6	350.0 ± 2.5	507.0 ± 4.9
Diffe-rence	1.6	0.4	0.7	0.2	1.6	3.3

Outlook: More realistic models

- 1D iterative map model for solar dynamo (Durney 2000, Charbonneau 2001), derived from Babcock Leighton mechanism:



$$P_{n+1} = \alpha_n f(P_n) P_n + \epsilon_n$$

- P_n normalised amplitude of cycle n
- α_n multiplicative noise parameter
- ε_n additive noise parameter
- f nonlinear map modelling poloidal field production from the decay of active regions as a function of the deep-seated toroidal magnetic component

- Promising features:
  - period doubling (higher harmonics?)
  - intermittency (Maunder minima?)

Outlook: Hamiltonian Monte Carlo

- Bayesian inference with stochastic models requires calculating a (hugh-) dimensional integral over all possible model realisations that are compatible with the data (and the prior):

$$f_{post}(\boldsymbol{\theta}|\mathbf{y}) \propto \int \exp[-\mathcal{A}_{model}(\boldsymbol{\alpha},\boldsymbol{\epsilon}) - \mathcal{A}_{data}(\boldsymbol{\alpha},\boldsymbol{\epsilon};\mathbf{y}) - \mathcal{A}_{prior}(\boldsymbol{\theta})] d\boldsymbol{\alpha} d\boldsymbol{\epsilon}$$

- Re-interpret this problem as a **statistical mechanics** problem (Albert, Ulzega, Stoop, PRE **93**, 2016):

  Pair each degree of freedom, ***α***, ***ε***, and ***θ***, with a conjugate momentum, add a Gaussian kinetic term to the action above, and generate parameter samples from the posterior via simulating the resulting Hamiltonian system.

$$f_{post}(\boldsymbol{\theta}|\mathbf{y}) \propto \int \exp[-\mathcal{H}_{\mathbf{y}}(\boldsymbol{\alpha},\boldsymbol{\epsilon},\boldsymbol{\theta};\mathbf{p}_\alpha,\mathbf{p}_\epsilon,\mathbf{p}_\theta)] d\boldsymbol{\alpha} d\boldsymbol{\epsilon} d\mathbf{p}_\alpha d\mathbf{p}_\epsilon d\mathbf{p}_\theta$$