Gaussian Processes for Transmission Spectroscopy: Application to NICMOS and WFC3 observations of HD 189733

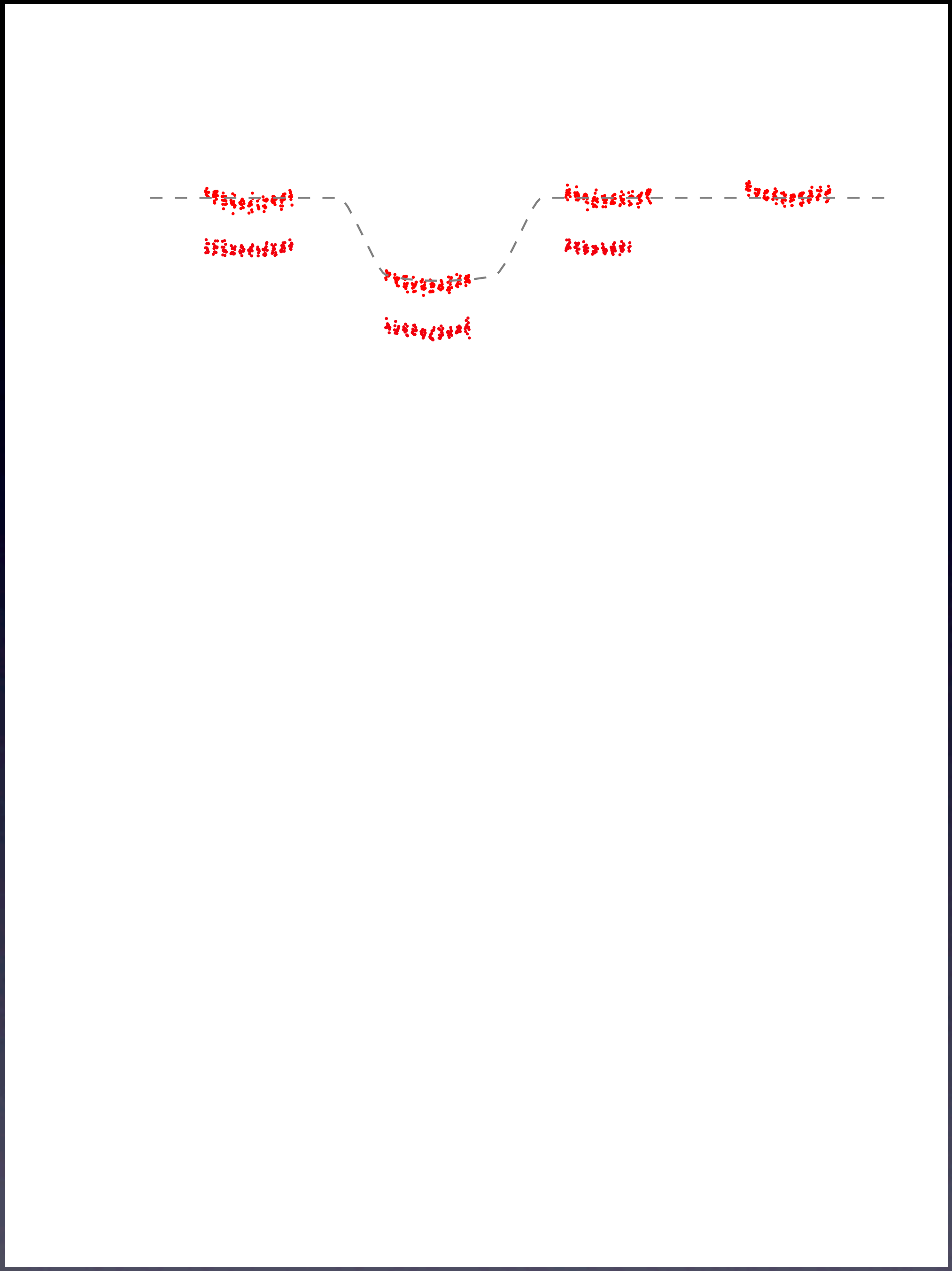
Neale Gibson
University of Oxford
Neale.Gibson@astro.ox.ac.uk

Collaborators:
Oxford Astrophysics: S. Aigrain, T. Evans
Oxford Information Engineering: S. Roberts, M. Osborne
Exeter: F. Pont, D. Sing

Transmission spectroscopy datasets

- Dataset consists of multiple light curves, in different wavelength channels.
- Systematics >> Transmission signal
- Choice of model is crucial to the spectrum
- Simultaneous *auxiliary measurements*, which describe the state of the instrument.
- 'Standard' method is to model the baseline flux as a linear function of auxiliary measurements:

$$f = \mathbf{X}\beta + \epsilon,$$



Why use a Gaussian process (GP) model?

- Using a non-parametric model makes fewer assumptions!
- GPs place a probability distribution over function space - this incorporates most commonly used functions.
- *Important, because often we have no physical justification to use a deterministic function of the systematics model.*

$$p(\boldsymbol{f}|\mathbf{X},\boldsymbol{\theta},\phi) = \mathcal{N}\left(T(\boldsymbol{t},\phi), \boldsymbol{\Sigma}(\mathbf{X},\boldsymbol{\theta})\right).$$

Why use a Gaussian process (GP) model?

- Using a non-parametric model makes fewer assumptions!
- GPs place a probability distribution over function space - this incorporates most commonly used functions.
- *Important, because often we have no physical justification to use a deterministic function of the systematics model.*

$$p(\boldsymbol{f}|\mathbf{X},\boldsymbol{\theta},\boldsymbol{\phi}) = \mathcal{N}\left(T(\boldsymbol{t},\boldsymbol{\phi}), \boldsymbol{\Sigma}(\mathbf{X},\boldsymbol{\theta})\right).$$

Transit light curve *mean function* with parameter vector ϕ

Why use a Gaussian process (GP) model?

- Using a non-parametric model makes fewer assumptions!
- GPs place a probability distribution over function space - this incorporates most commonly used functions.
- *Important, because often we have no physical justification to use a deterministic function of the systematics model.*

$$p(\boldsymbol{f}|\mathbf{X},\boldsymbol{\theta},\boldsymbol{\phi}) = \mathcal{N}\left(T(\boldsymbol{t},\boldsymbol{\phi}),\boldsymbol{\Sigma}(\mathbf{X},\boldsymbol{\theta})\right).$$

Transit light curve *mean function* with parameter vector ϕ

Covariance matrix: defined by kernel function which uses the inputs X, with parameter vector θ

Controls the systematics model and white noise

The Kernel function: Systematics + White Noise

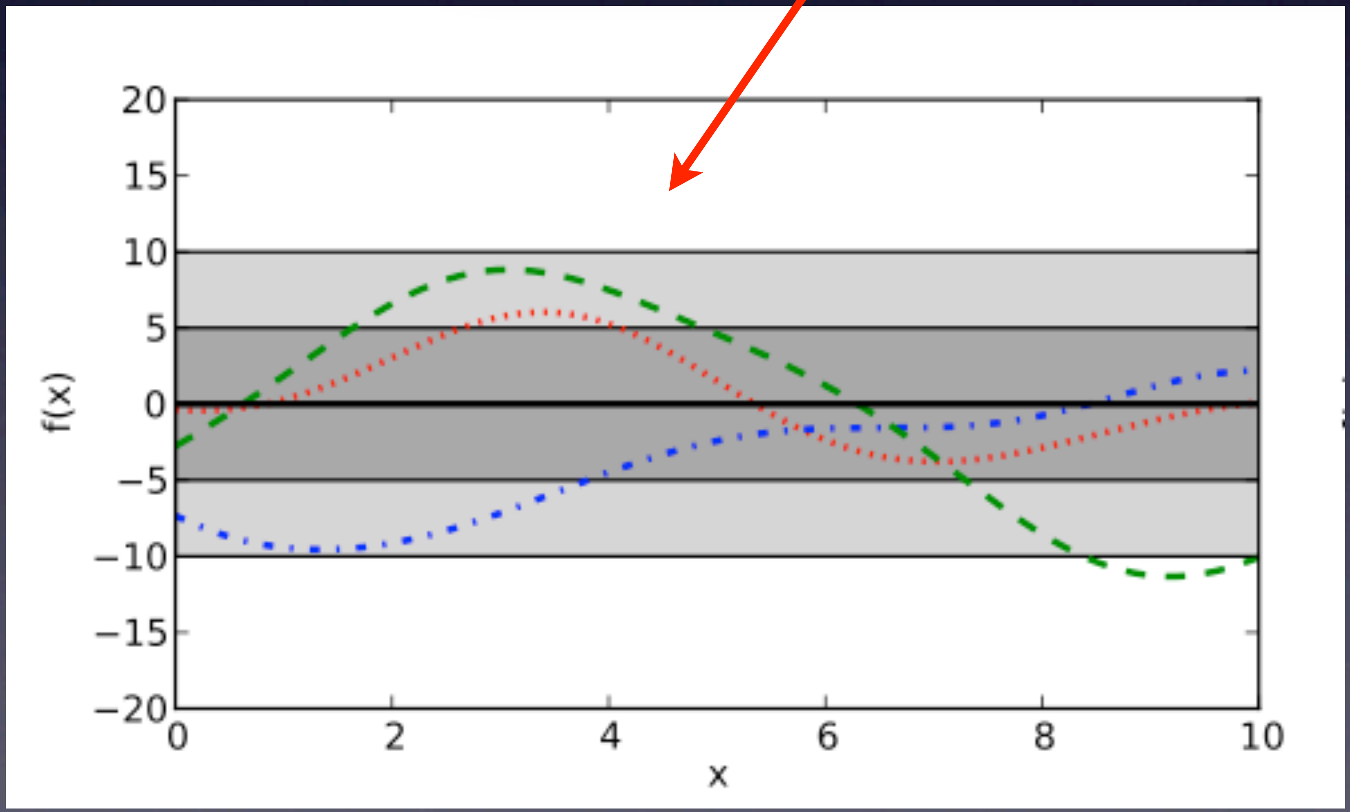
- Kernel function is a variation of the common squared exponential.

$$k(\boldsymbol{x}_n, \boldsymbol{x}_m) = \xi \exp\left(-\sum_{i=1}^{K} \eta_i (x_{n,i} - x_{m,i})^2\right) + \delta_{nm}\sigma^2.$$

The Kernel function: Systematics + White Noise

- Kernel function is a variation of the common squared exponential.

$$k(\boldsymbol{x}_n, \boldsymbol{x}_m) = \xi \exp\left(- \sum_{i=1}^{K} \eta_i (x_{n,i} - x_{m,i})^2 \right) + \delta_{nm}\sigma^2.$$



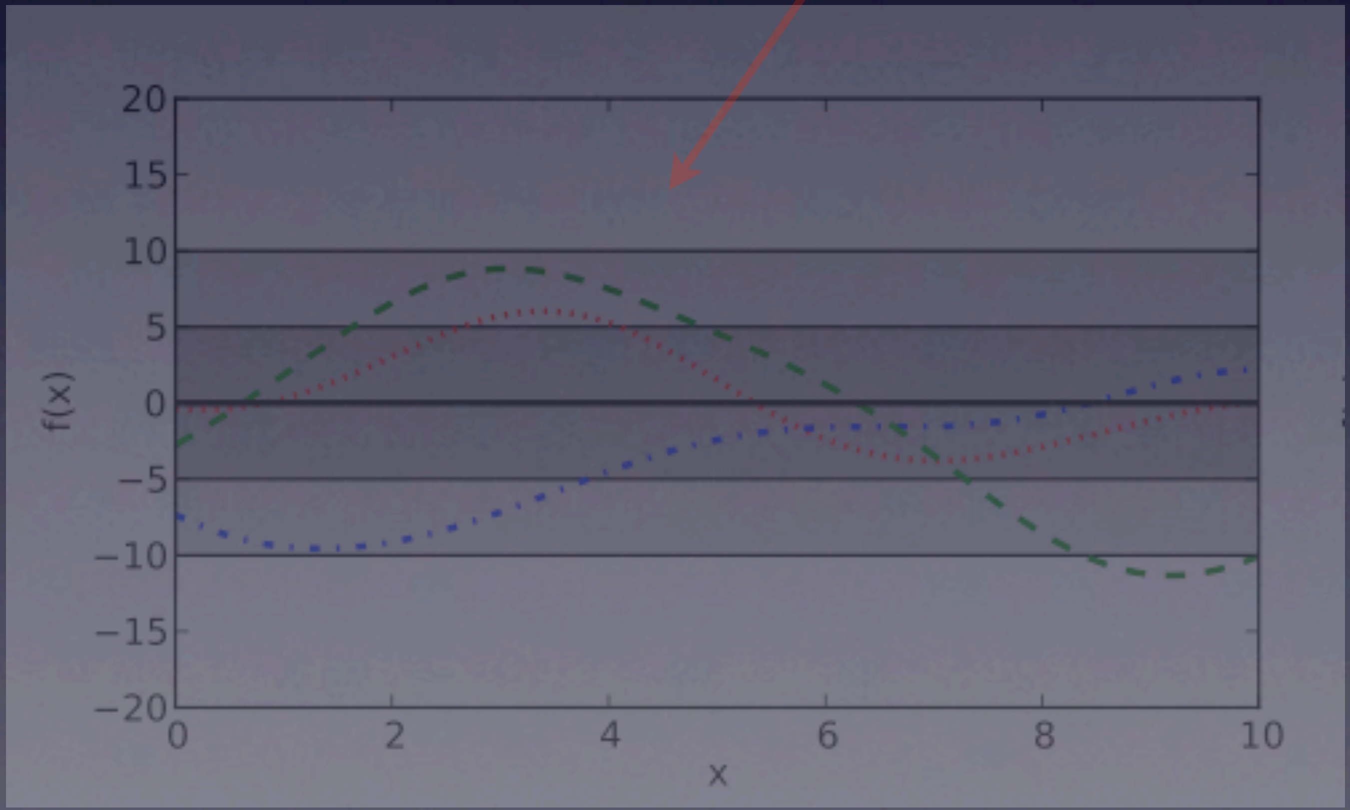
Probability distribution over smooth functions (but now multivariate!)

The Kernel function: Systematics + White Noise

- Kernel function is a variation of the common squared exponential.

$$k(\boldsymbol{x}_n, \boldsymbol{x}_m) = \xi \exp\left(-\sum_{i=1}^{K} \eta_i (x_{n,i} - x_{m,i})^2 \right) + \delta_{nm}\sigma^2.$$

Inverse length scale parameter for each input vector



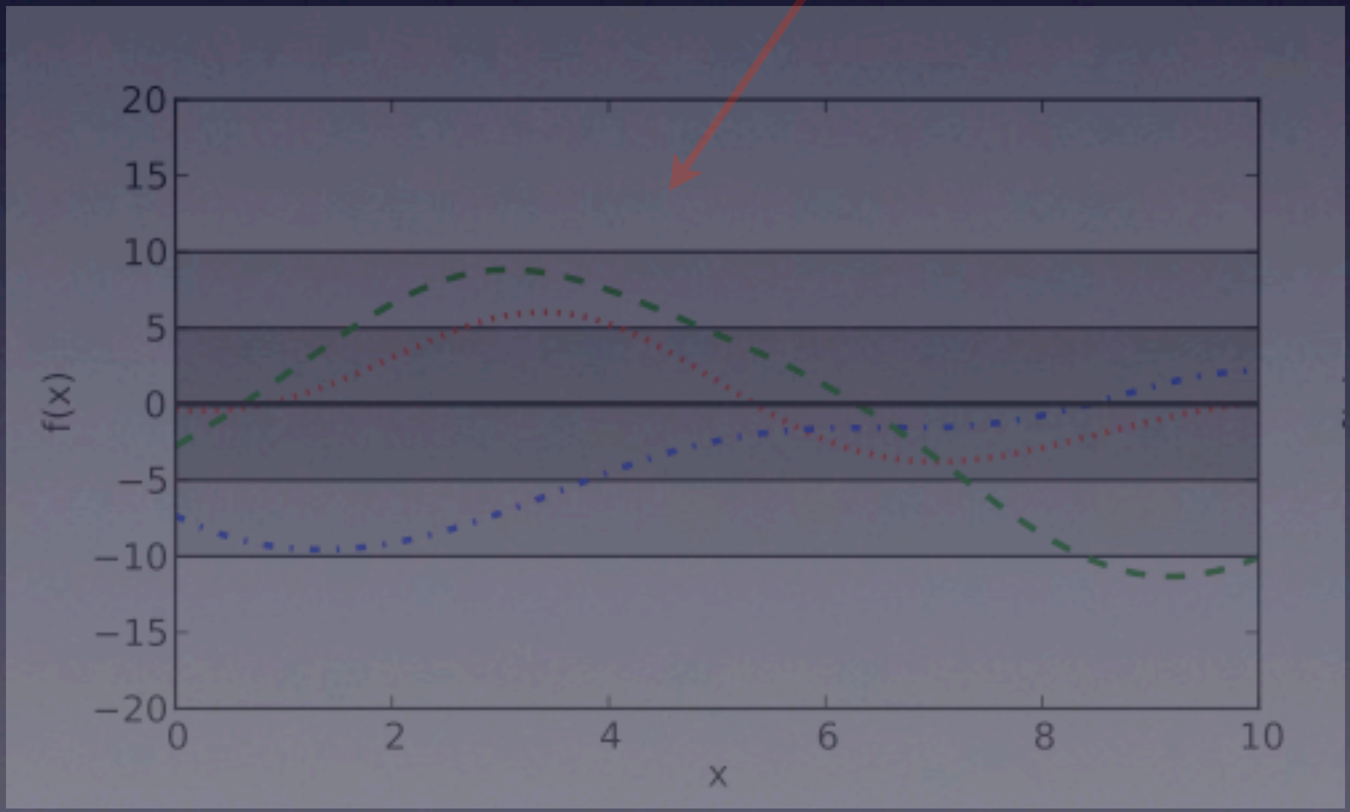
Probability distribution over smooth functions (but now multivariate!)

The Kernel function: Systematics + White Noise

- Kernel function is a variation of the common squared exponential.

$$k(\boldsymbol{x}_n, \boldsymbol{x}_m) = \xi \exp\left(-\sum_{i=1}^{K} \eta_i (x_{n,i} - x_{m,i})^2\right) + \delta_{nm}\sigma^2.$$

Inverse length scale parameter for each input vector



White noise added to the diagonal

Probability distribution over smooth functions (but now multivariate!)

Inferring the Gaussian process parameters

- In a GP, the (marginal) likelihood function is multivariate Gaussian;

$$p(\boldsymbol{f}|\mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\phi}) = \mathcal{N}\left(T(\boldsymbol{t}, \boldsymbol{\phi}), \boldsymbol{\Sigma}(\mathbf{X}, \boldsymbol{\theta})\right).$$

$$\mathcal{L}(\mathbf{r}|\mathbf{X}, \boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{r}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{r}\right),$$

(residuals | inputs, transit parameters, kernel parameters)

- Optimise with respect to the transit parameters and kernel (hyper-)parameters

or

- Explore the joint posterior distribution with MCMC (after multiplying by priors!)

Inferring the Gaussian process parameters

- In a GP, the (marginal) likelihood function is multivariate Gaussian;

$$p(\boldsymbol{f}|\mathbf{X},\boldsymbol{\theta},\boldsymbol{\phi}) = \mathcal{N}\left(T(\boldsymbol{t},\boldsymbol{\phi}),\boldsymbol{\Sigma}(\mathbf{X},\boldsymbol{\theta})\right).$$

$$\mathcal{L}(\mathbf{r}|\mathbf{X},\boldsymbol{\theta},\boldsymbol{\phi}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{r}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{r}\right),$$

(residuals | inputs, transit parameters, kernel parameters)

- Optimise with respect to the transit parameters and kernel (hyper-)parameters

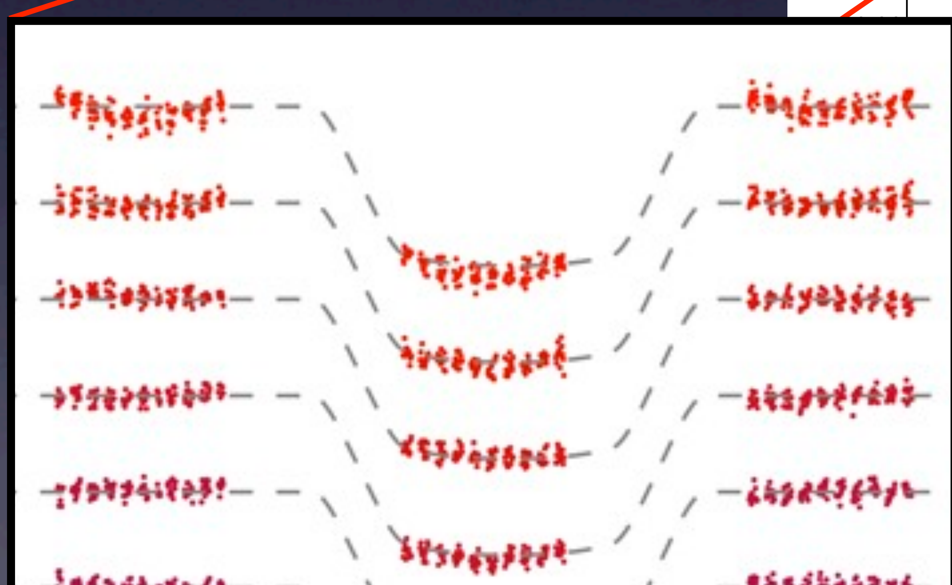
or

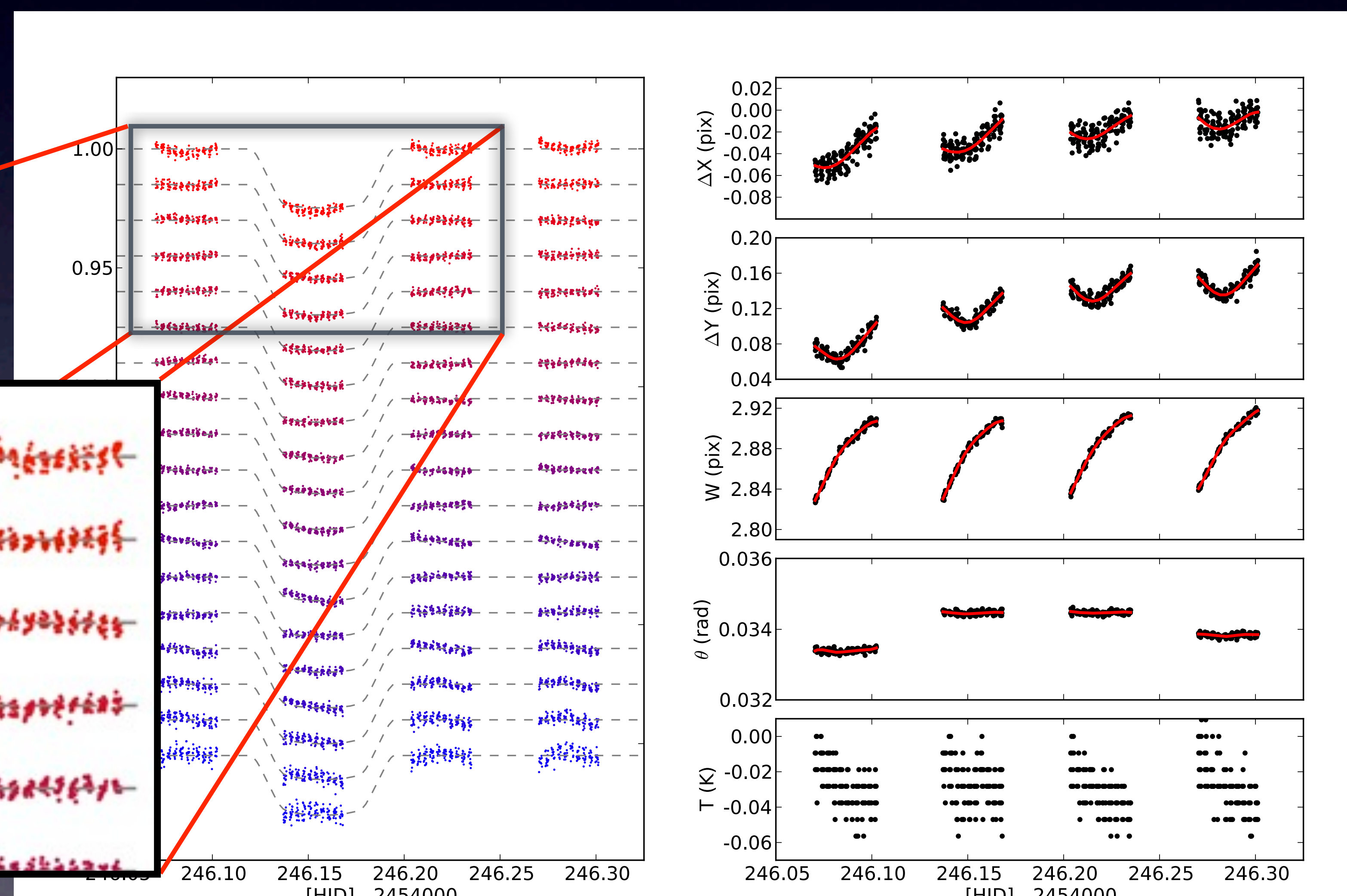
- Explore the joint posterior distribution with MCMC (after multiplying by priors!)

Inverse covariance matrix requires $O(n^3)$ calculations!!

NICMOS transmission spectroscopy of HD 189733

- First analysed by Swain, Vasisht & Tinetti (2008). Consists of 18 light curves from 1.4 to 2.5 μm (CH_4, H_20)
- Reanalysed by Gibson, Pont & Aigrain (2011).
- Auxiliary parameters: X,Y,Width,Temp,Phase,Angle
- Systematics ~10^-3 Spectra ~10^-4

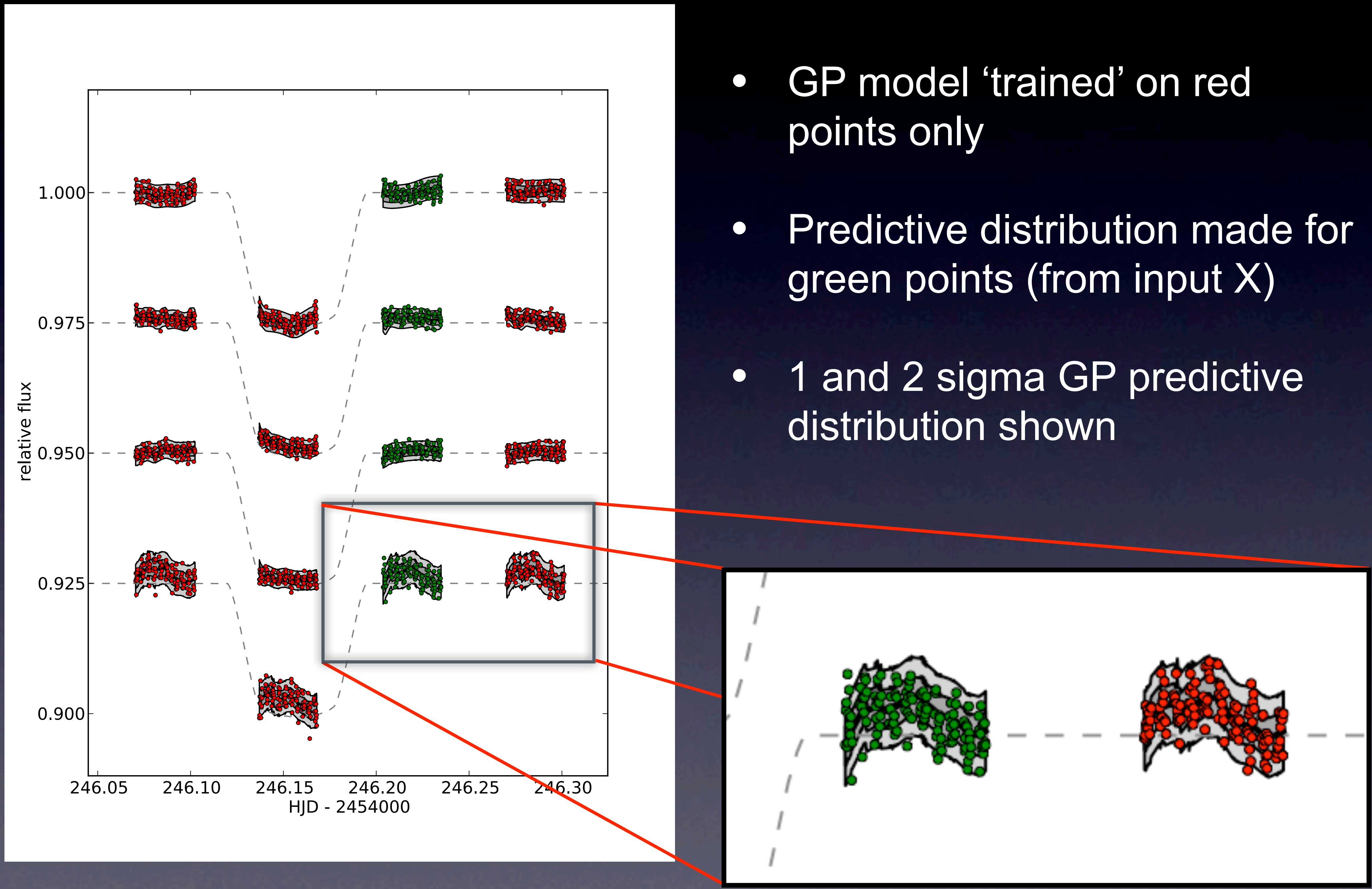




Gibson, Pont & Aigrain (2011)

NICMOS transmission spectroscopy of HD 189733

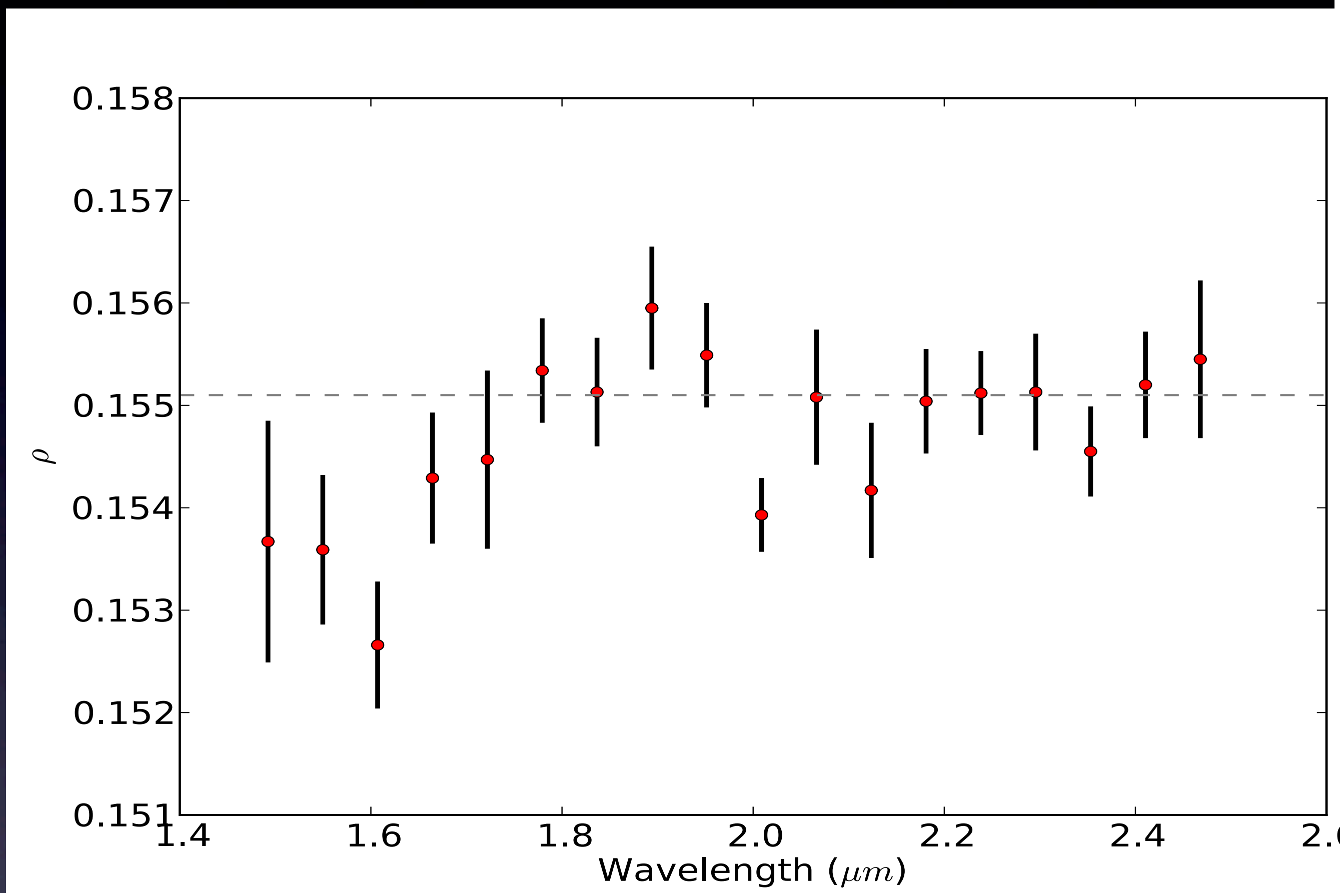
Gibson et al. (submitted)



- GP model 'trained' on red points only
- Predictive distribution made for green points (from input X)
- 1 and 2 sigma GP predictive distribution shown

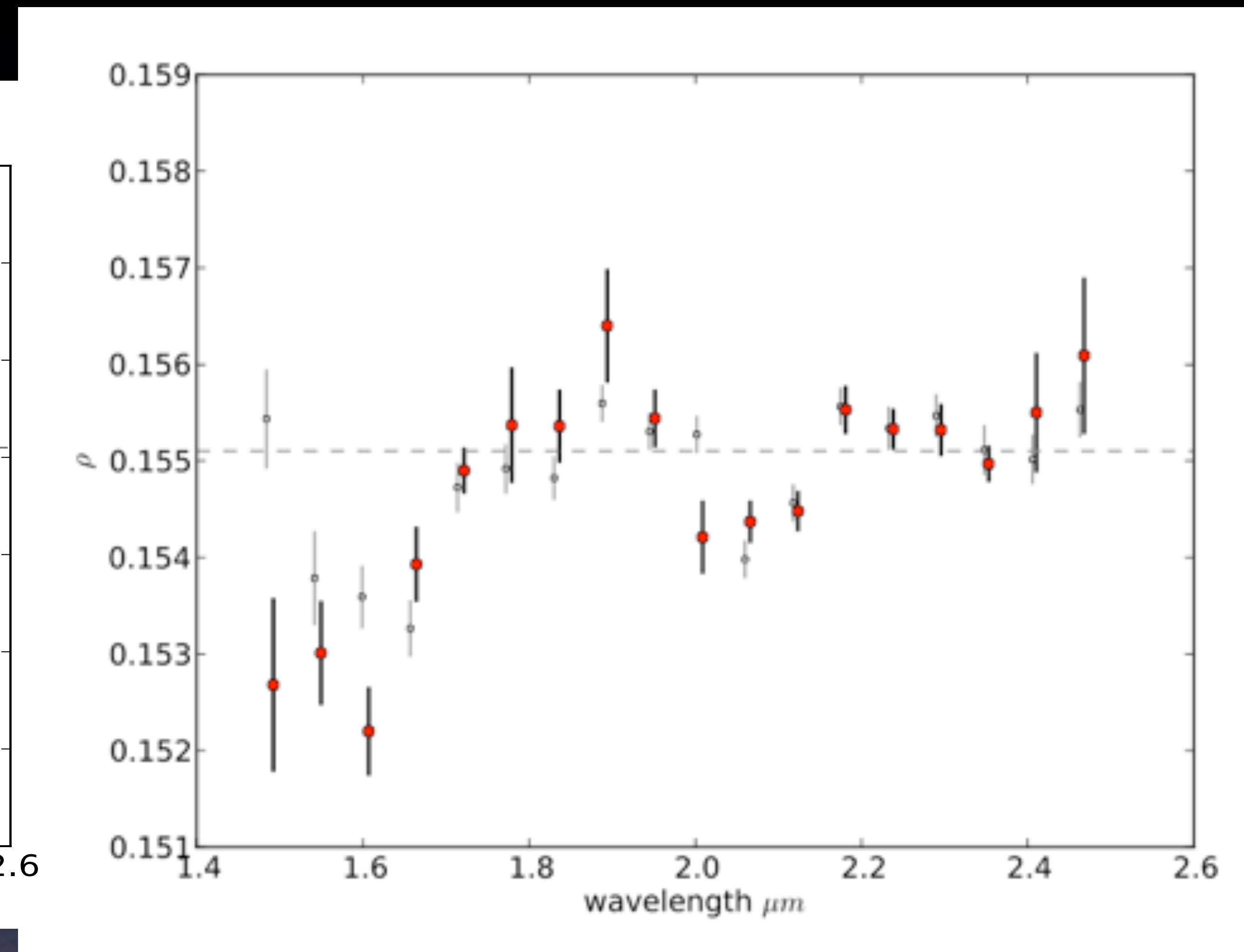
NICMOS transmission spectroscopy of HD 189733

Gibson et al. (submitted)



Gaussian process model

Gibson, Pont & Aigrain (2011)



Linear basis functions

- This shows the linear basis function model is not sufficient to explain the NICMOS systematics.

Gaussian processes for transmission spectroscopy

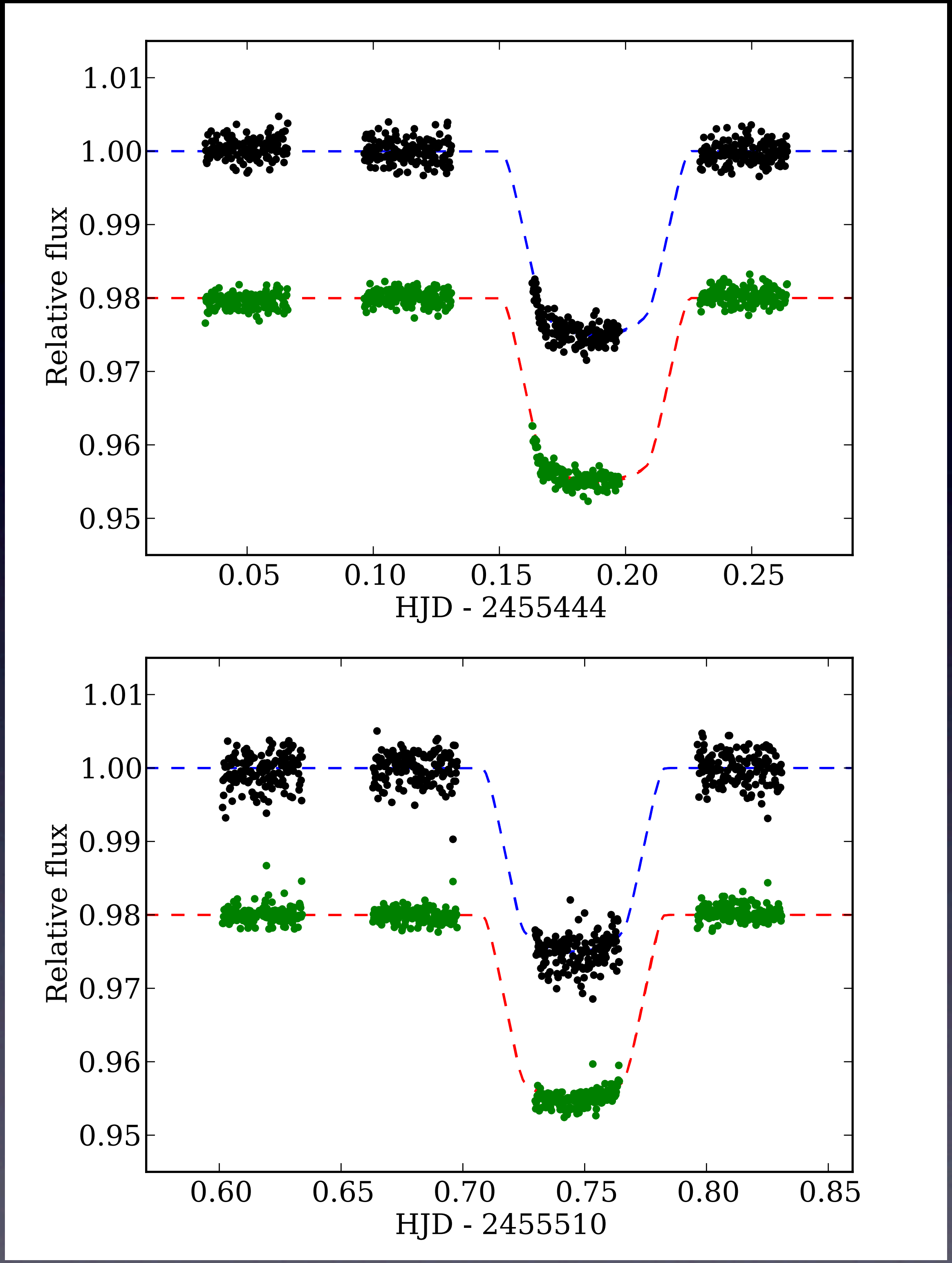
- Paper will appear on astro-ph tonight!

A Gaussian process framework for modelling instrumental systematics: application to transmission spectroscopy

N. P. Gibson[1*], S. Aigrain[1], S. Roberts[2], T. M. Evans[1], M. Osborne[2] and F. Pont[3]
[1] *Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[2] *Robotics Research Group, Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK*
[3] *School of Physics, University of Exeter, Exeter, EX4 4QL, UK*

WFC3 transmission spectroscopy of HD 189733

- G141 grism - Idea was to bridge the gap between ACS and NICMOS data
- 2 transits observed
- Light curves at ~1.1 and 1.6 μm
- GP type-II max likelihood only
- Analysis on-going...



Gibson et al. (in prep)

WFC3 transmission spectroscopy of HD 189733