Estimating parameters from multidimensional data

Coryn Bailer-Jones
Max Planck Institute for Astronomy, Heidelberg
http://www.mpia.de/~calj

ELSA School, Heidelberg
1 October 2009

Outline

- Estimation problems and methods for solving them
- The Gaia problem: estimating stellar parameters
- An estimation method based on forward modelling: ILIUM
- Application of ILIUM to BP/RP spectra
- A T_{eff} - A_V degeneracy
- Conclusions

Two distinct problems

1. Classification
   - estimation of best *discrete* class (generally, class probabilities)
   - e.g. star/galaxy/QSO discrimination using Gaia BP/RP and astrometry
2. Parametrization
   - estimation of best *continuous* parameters (generally, probability distribution over the parameters)
   - e.g. stellar astrophysical parameter (AP) estimation using Gaia BP/RP, RVS, parallax

↑

remember this!

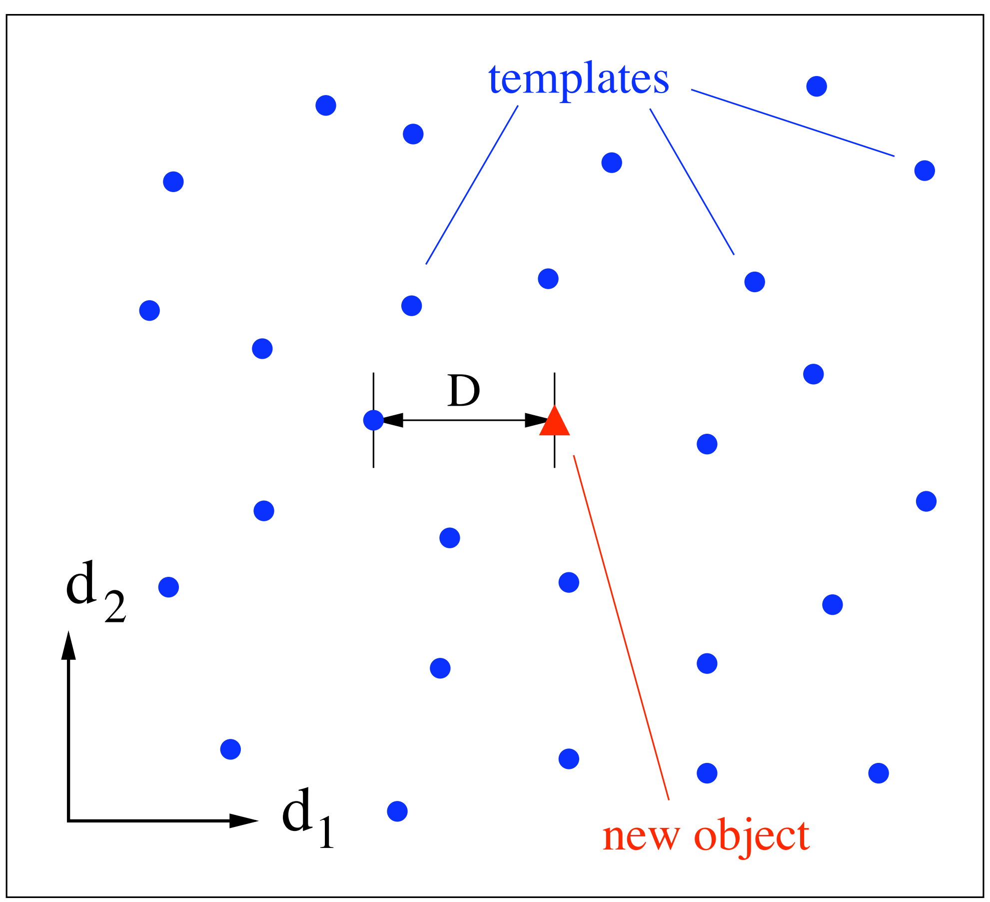
Methods of AP estimation (from spectra)

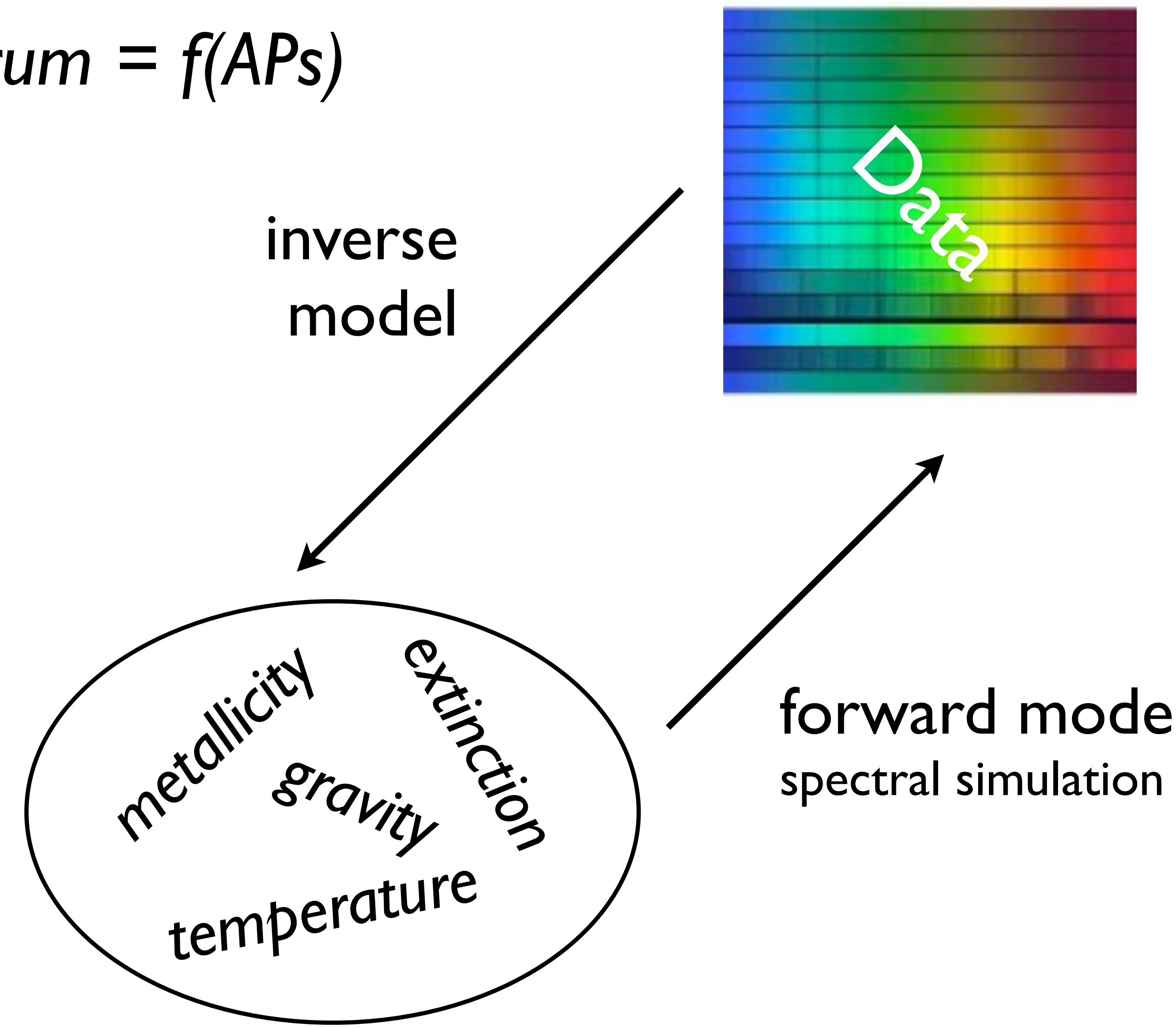
1. Physically-motivated metrics (equiv. widths, line indices etc.)
    - isolate specific features sensitive to phenomena of interest
    - generally restricted to narrow part of AP space
2. Global pattern recognition
    - attempt to make use of *all* information in spectrum
    - infer (learn) what is relevant for which APs

In all cases we must fit/calibrate/train the models on physical models to give physical parameters!

Methods of pattern recognition

1. direct, global inverse mapping: *APs = g(spectrum)*
2. template matching (k nearest neigbours)
3. forward modelling: *spectrum = f(APs)*





Methods of pattern recognition

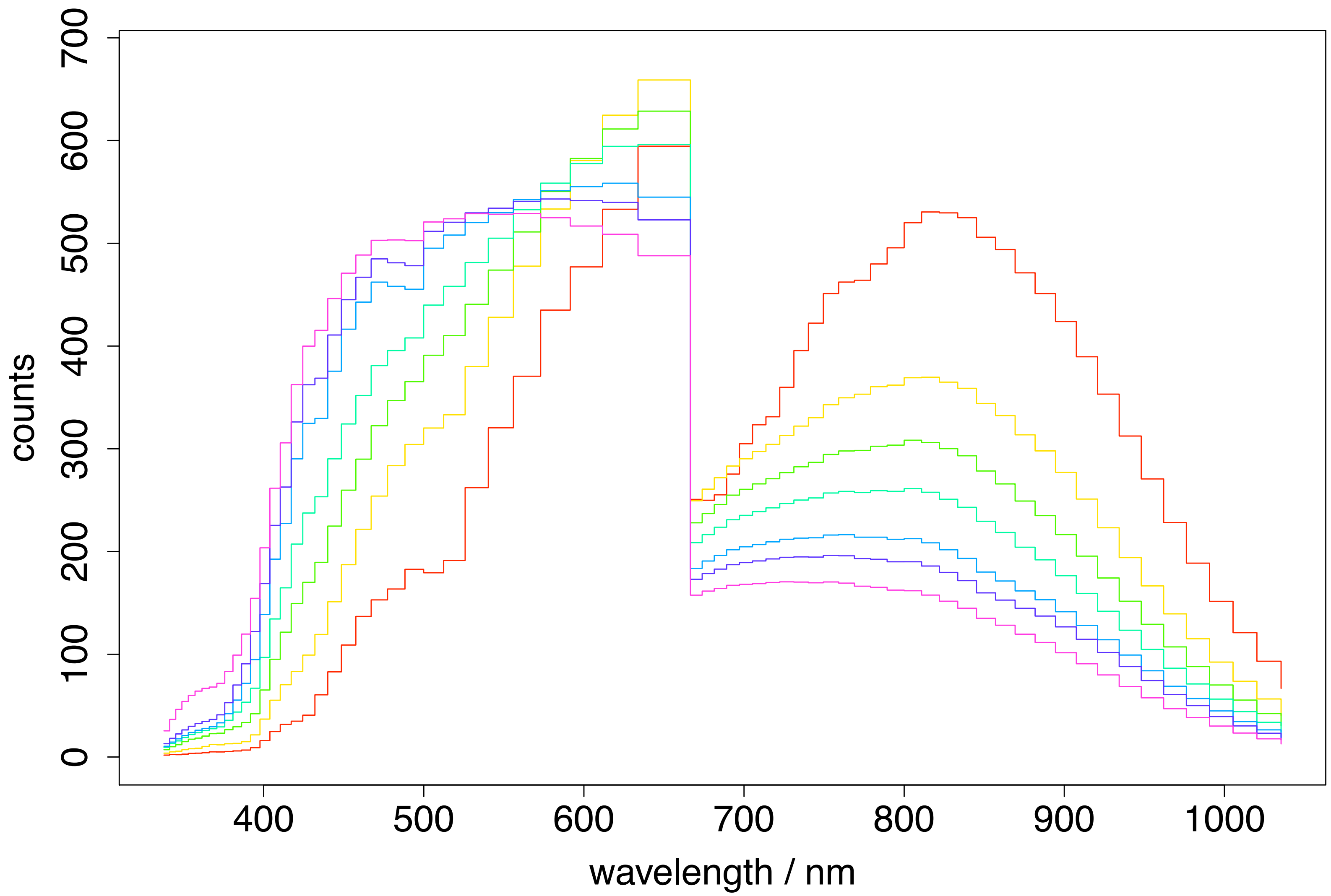
1. direct inverse mapping, *APs = g(spectrum)*
   - flexible; off-the-shelf methods (e.g. ANN, SVM); fast
   - complex; non-unique; must learn sensitivity of APs to data; no natural uncertainty estimates or goodness-of-fit (GoF)
2. template matching (k nearest neigbours)
   - direct; "ideal" method
   - need a large grid: curse of dimensionality; slow; insensitive to weak APs → need adaptive kernel size, appropriate distance metric
3. forward modelling, *spectrum = f(APs)*
   - model is unique; error/GoF estimates
   - difficult to simultaneously fit strong and weak APs; need scheme to invert to get APs

Gaia data

Coryn Bailer-Jones, MPIA Heidelberg

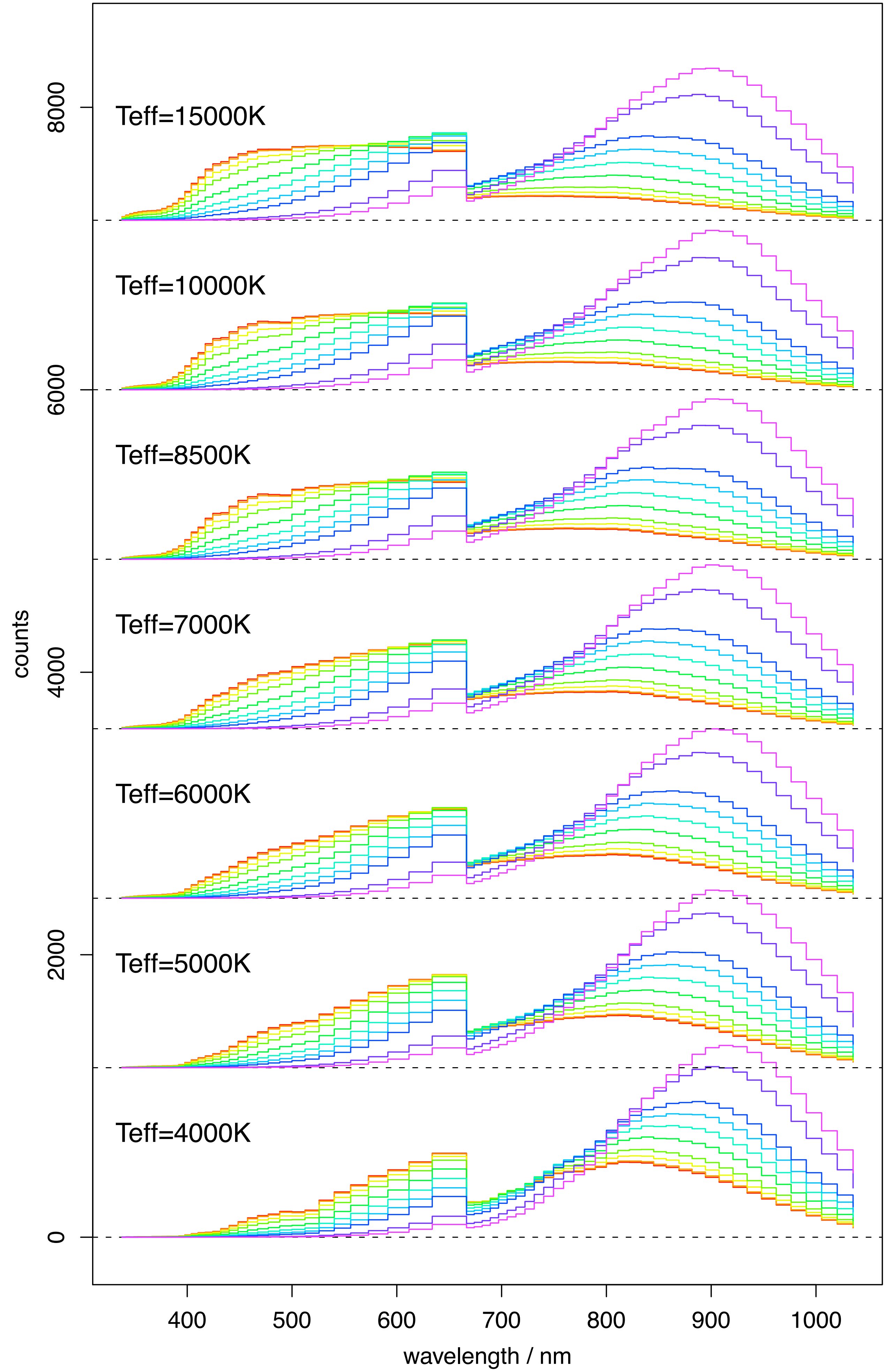
Gaia BP/RP spectra (GOG simulations)



A_V=0, [Fe/H]=0, [α/Fe]=0, logg=4

Teff / K = 4000, 5000, 6000, 7000, 8500, 10000, 15000

- Basel and Marcs stellar libraries
- 34 pixels in each of BP and RP retained
- Nominal sampling, LSF included, noise added, but no CTI etc.

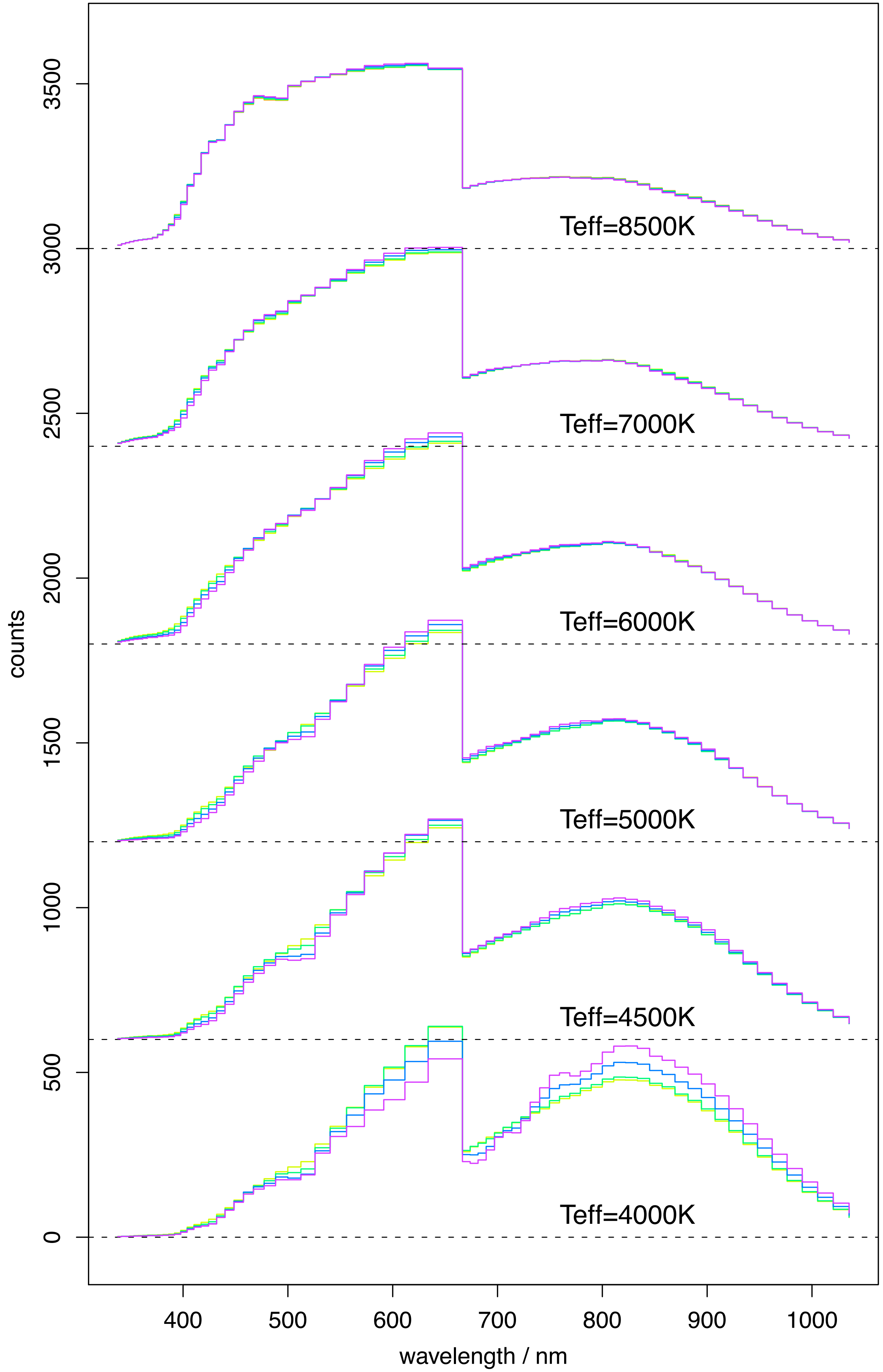


T_{eff} and A_V variation

A_V = 0, 0.1, 0.5, 1, 2, 3, 4, 5, 8, 10

(R_V=3.1)

T_{eff} and A_V: "strong" APs

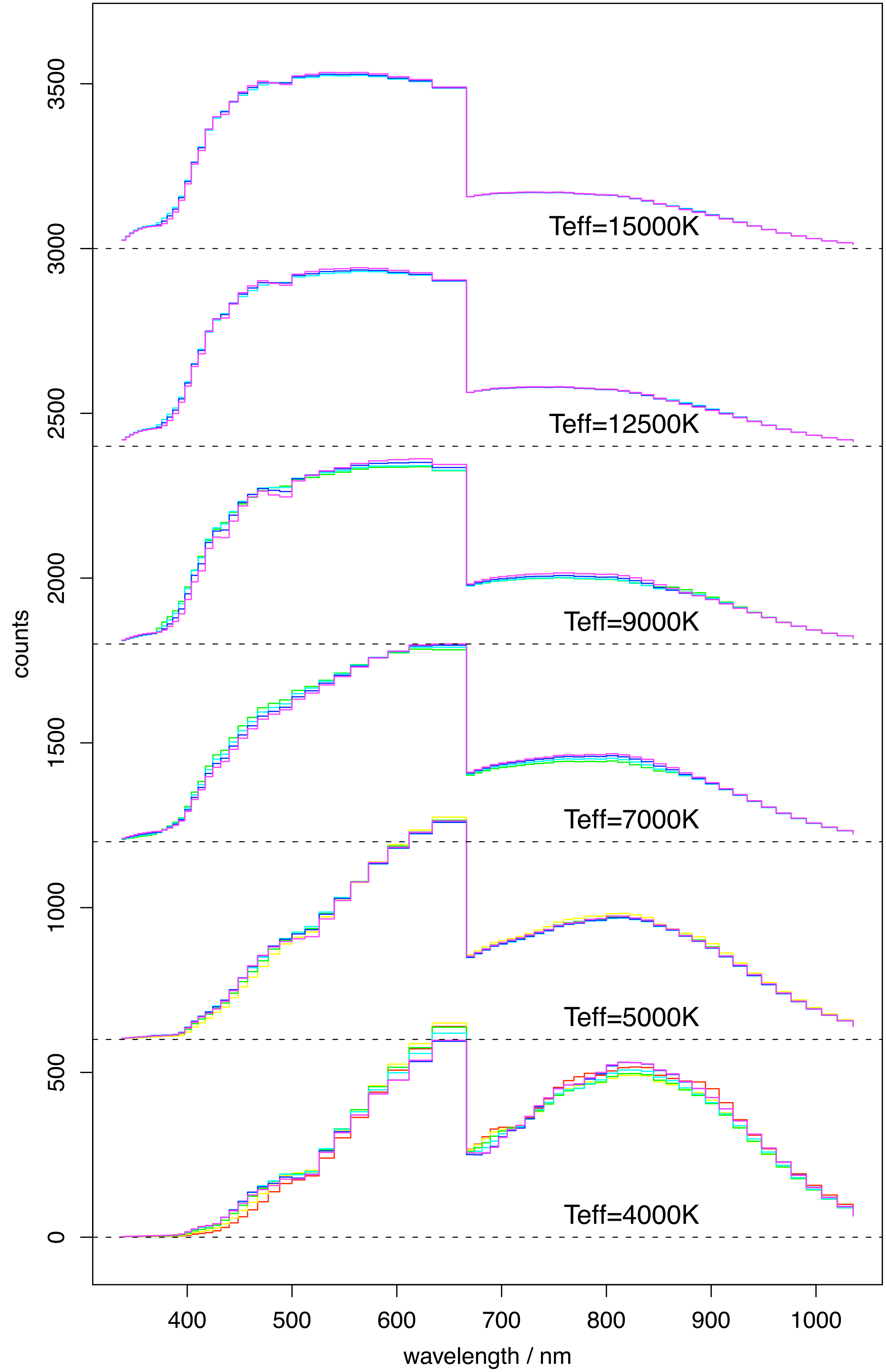


T_{eff} and [Fe/H] variation

[Fe/H] = -3, -2, -1, 0, +0.5

(A_V=0, [Fe/H]=0)

[Fe/H]: a “weak” AP



T_{eff} and logg variation

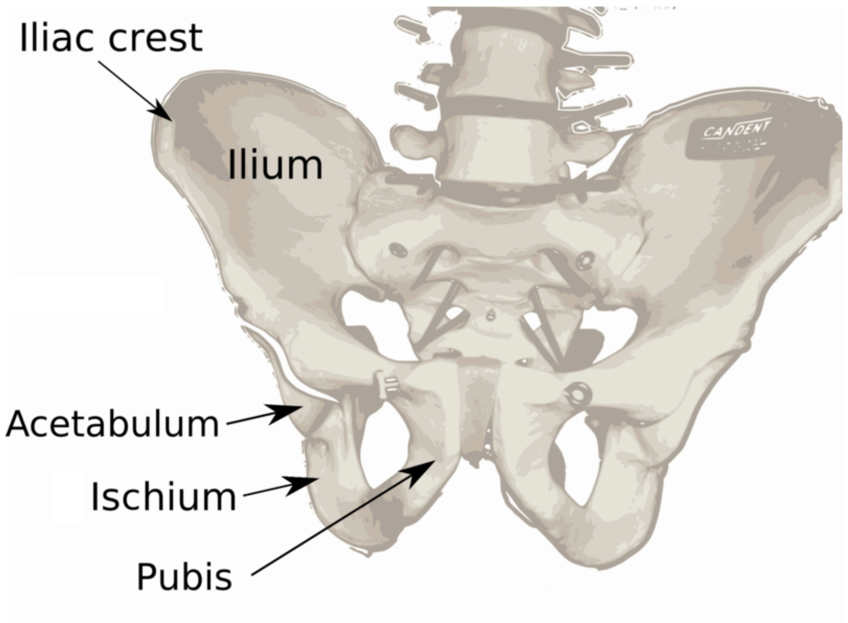
logg = -0.5, 0.5, 2, 3, 4, 5

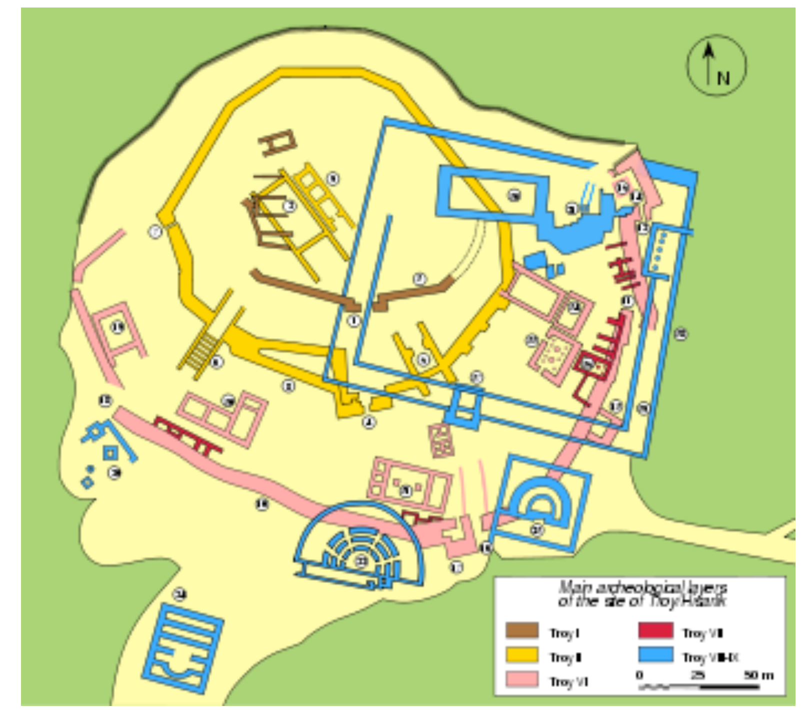
(A_V=0, [Fe/H]=0)

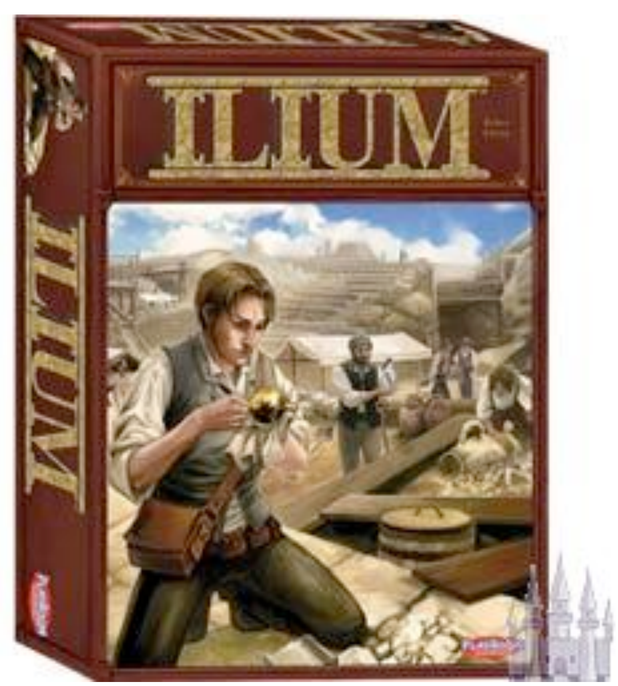
logg: a “weak” AP

AP estimation algorithm based on forward modelling:

ILIUM







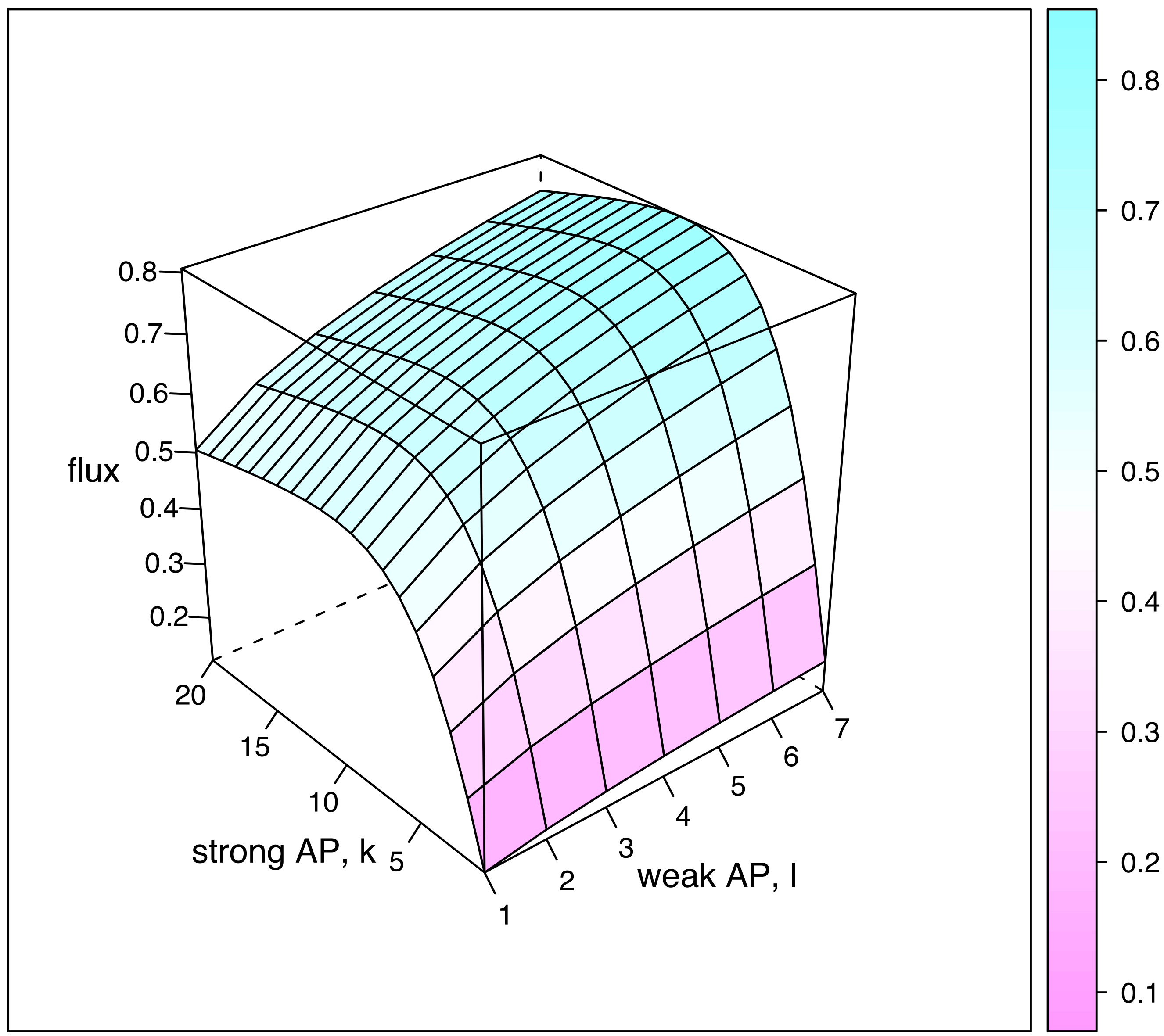
Further reading:
CBJ-042, -043, -046, 048
on Livelink

Some definitions

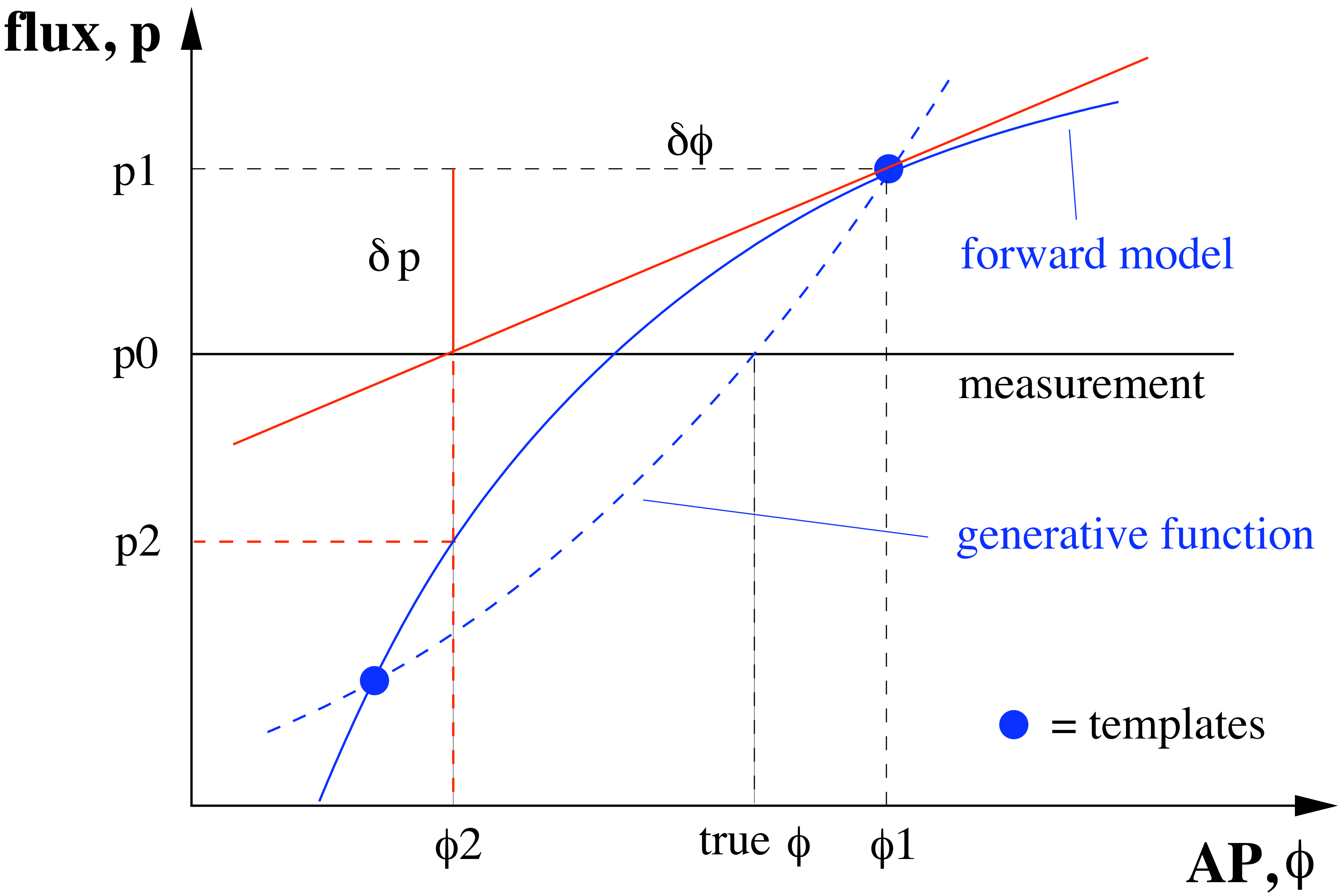
p_i	counts in i^{th} band (spectral pixel)	I bands
ϕ_j	j^{th} AP	J APs

Forward modelling

- Forward model for band i, $p_i = f_i(\boldsymbol{\phi})$ (multiple APs)
- sensitivities are its derivatives w.r.t ϕ_j, $s_{ij} = \partial p_i / \partial \phi_j$



AP update algorithm



1. find nearest neighbour: gives first AP estimate
2. calculate residual, δp
3. calculate sensitivity
4. make AP update
5. predict flux
6. iterate steps 2-5

$$\delta\phi(n) = \left(\frac{\partial\phi}{\partial p}\right)_{\phi(n)} \times \delta p(n)$$

$$\phi(n+1) = \phi(n) - \delta\phi(n)$$

Coryn Bailer-Jones, MPIA Heidelberg

Generalization to multiple bands and APs

I bands, J APs (I > J)
S is the I x J sensitivity matrix
with elements $s_{ij} = \partial p_i / \partial \phi_j$

$$\delta \boldsymbol{p} = \mathbf{S}\, \delta \boldsymbol{\phi}$$

$$\delta \boldsymbol{\phi} = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \delta \boldsymbol{p}$$

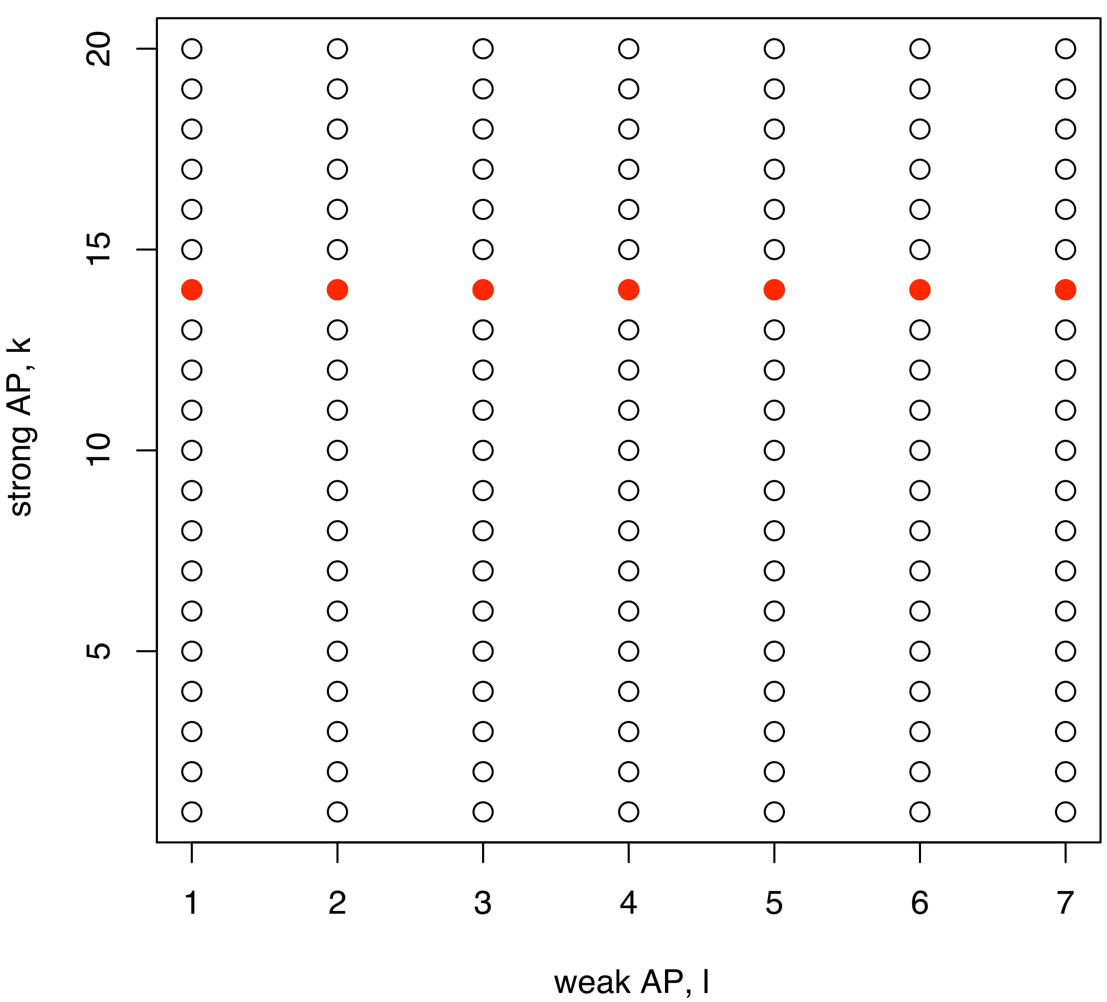
This can be written (for AP *j*) as:

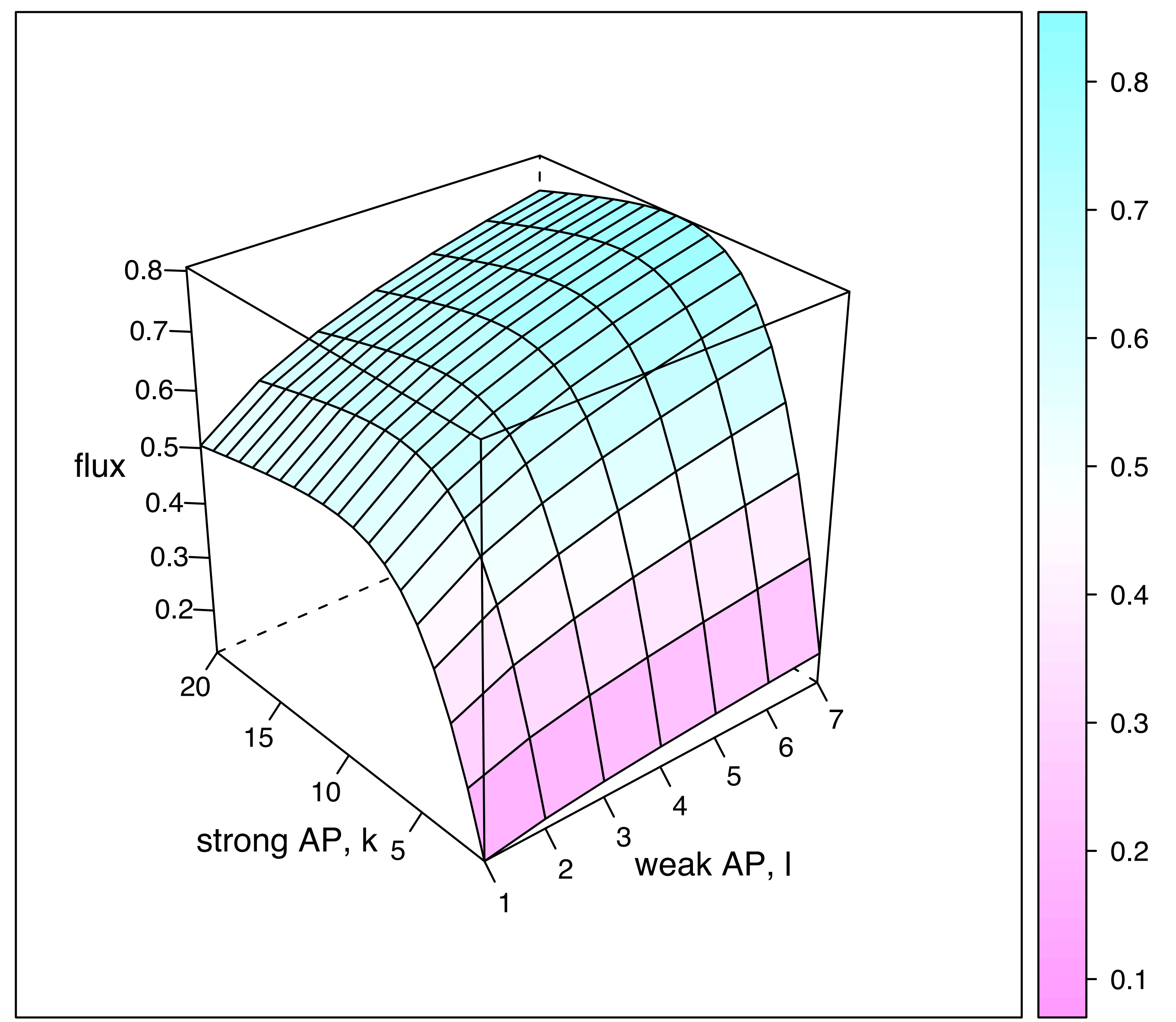
$$\phi_j(n+1) = \phi_j(n) - \underline{\boldsymbol{m}_j}\, \delta \boldsymbol{p}(n)$$

i.e. the update is a sensitivity-weighted sum over the bands

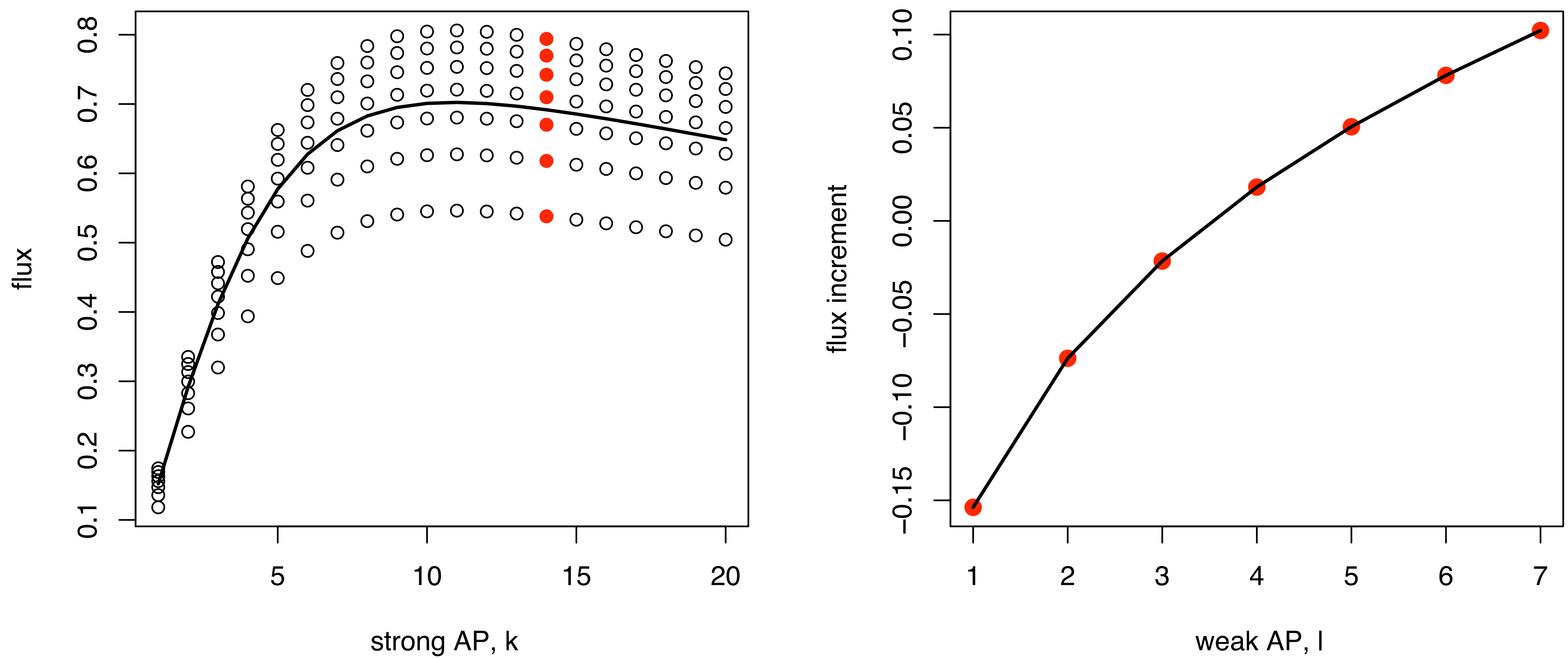
Form and fit of the forward models

Task: fit $p_i = f_i(\phi)$ for each band (separately) given a template grid of spectra with known APs





Form and fit of the forward models



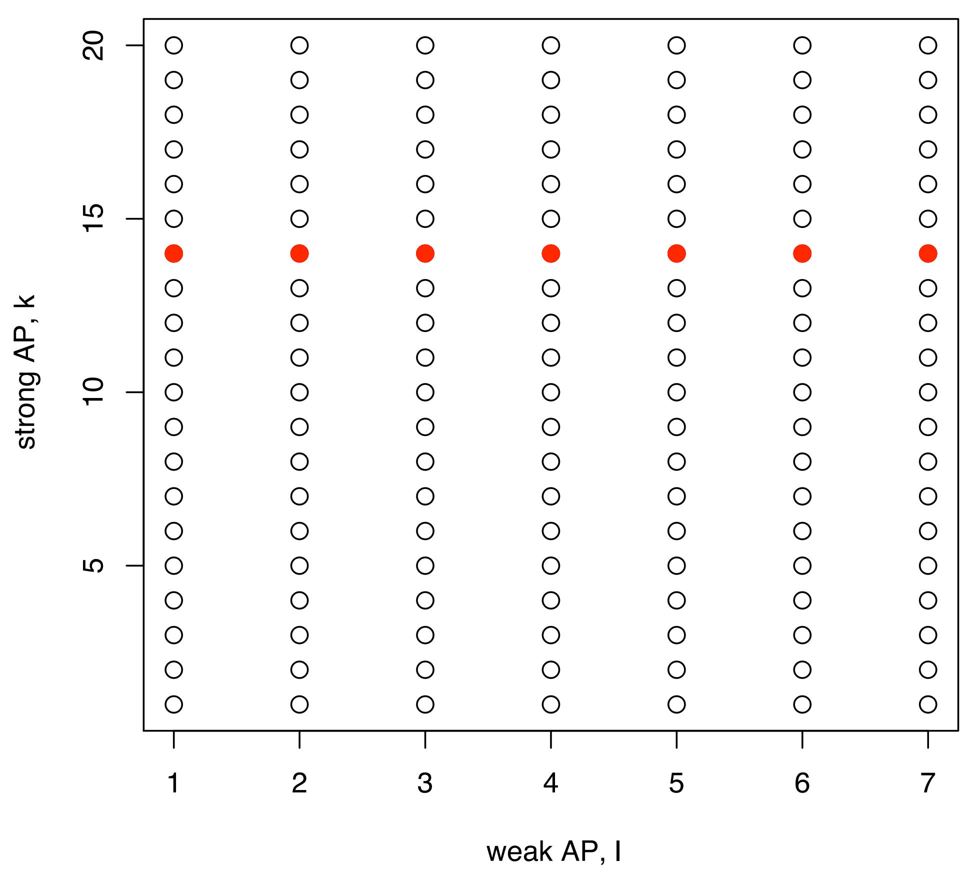
- Model strong and weak APs independently
- Fit strong APs by marginalizing over the weak, $f^{strong}(\phi^{strong})$
- At *each* strong AP, fit *residual* flux to the weak APs, $f_k^{weak}(\phi^{weak})$

Coryn Bailer-Jones, MPIA Heidelberg

Applying the forward model

Given the APs (ϕ^{strong}, ϕ^{weak})

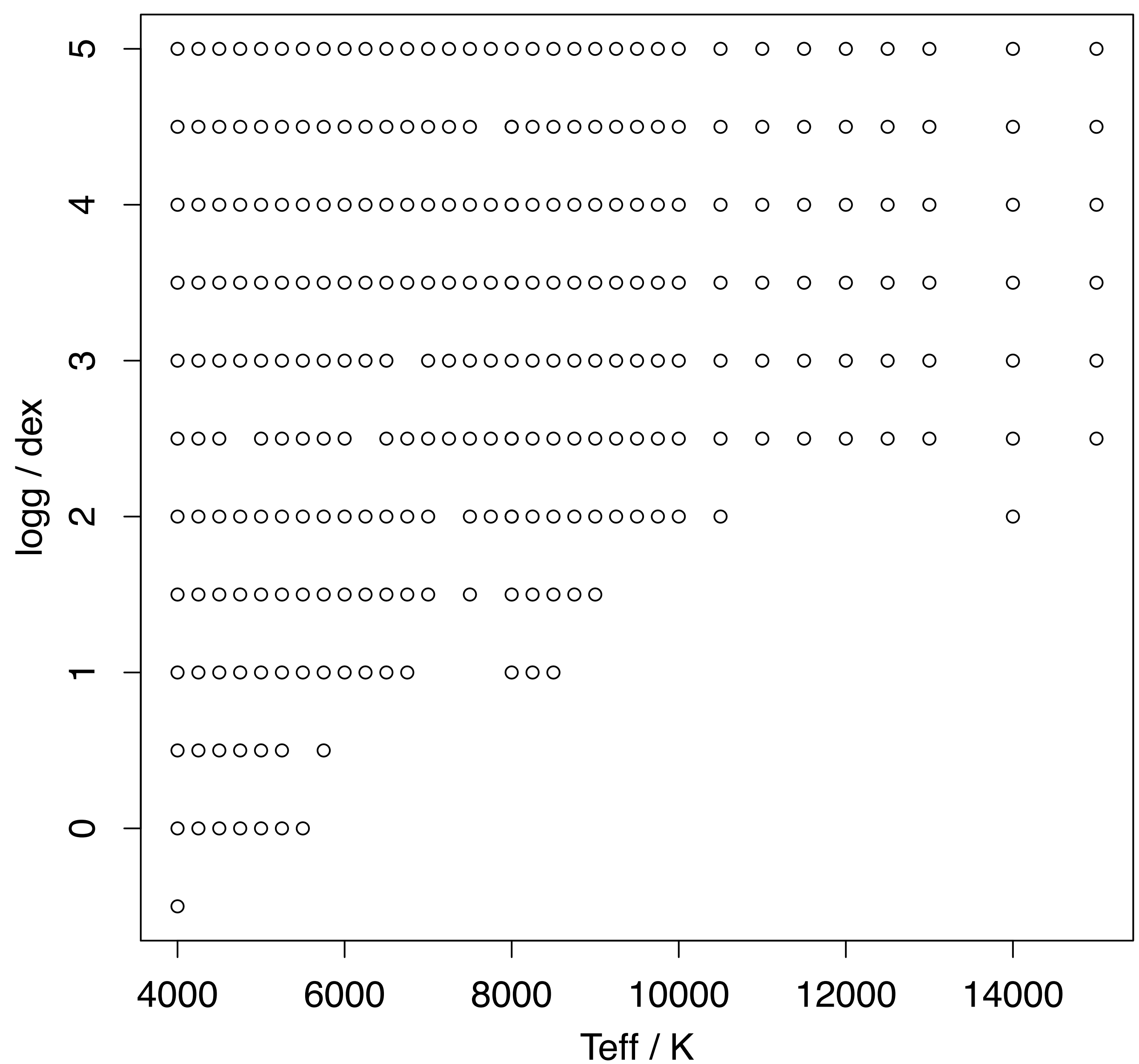
1. Evaluate $f^{strong}(\phi^{strong})$
2. Select weak component: find ϕ_k^{strong}, nearest neighbour in grid to ϕ^{strong}
3. Evaluate $f_k^{weak}(\phi^{weak})$
4. $f(\phi^{strong}, \phi^{weak}) = f^{strong}(\phi^{strong}) + f_k^{weak}(\phi^{weak})$



In current implementation, smoothing splines are used. Derivatives by first differences.

Applications

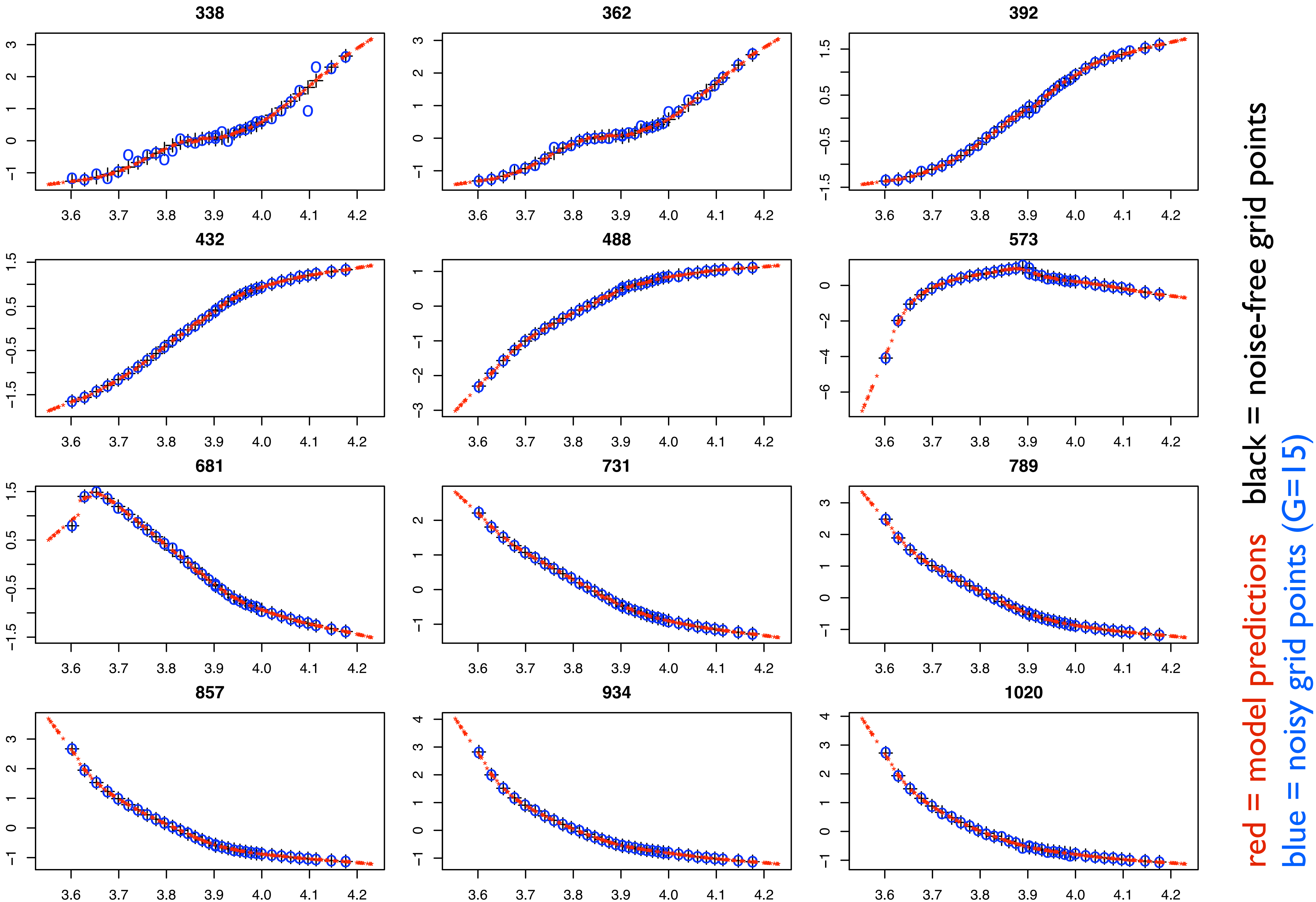
Application 1: T_{eff}/logg estimation



Template grid

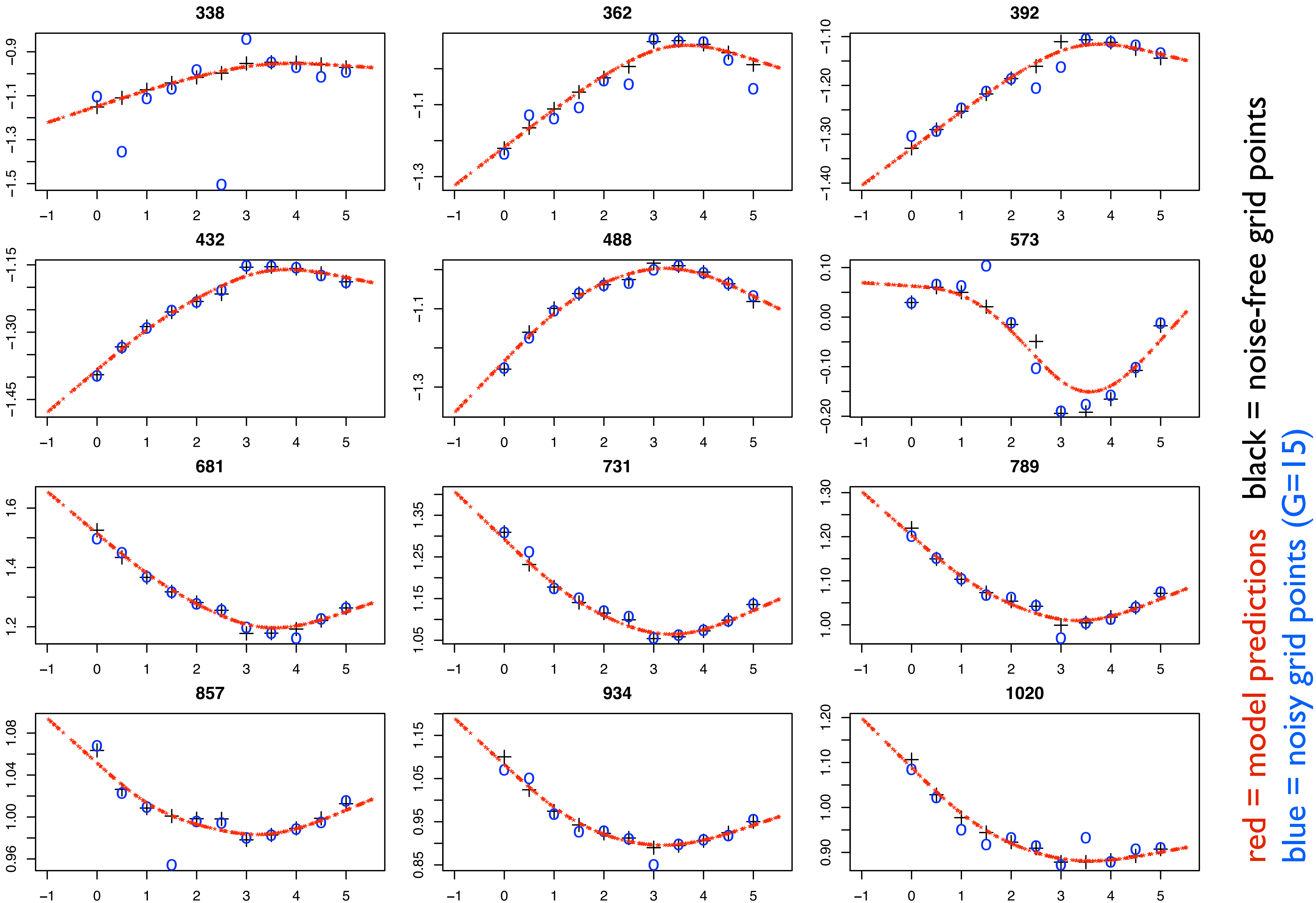
[Fe/H]=0
A_V=0

Forward model fit over T_{eff} (logg=4.0)

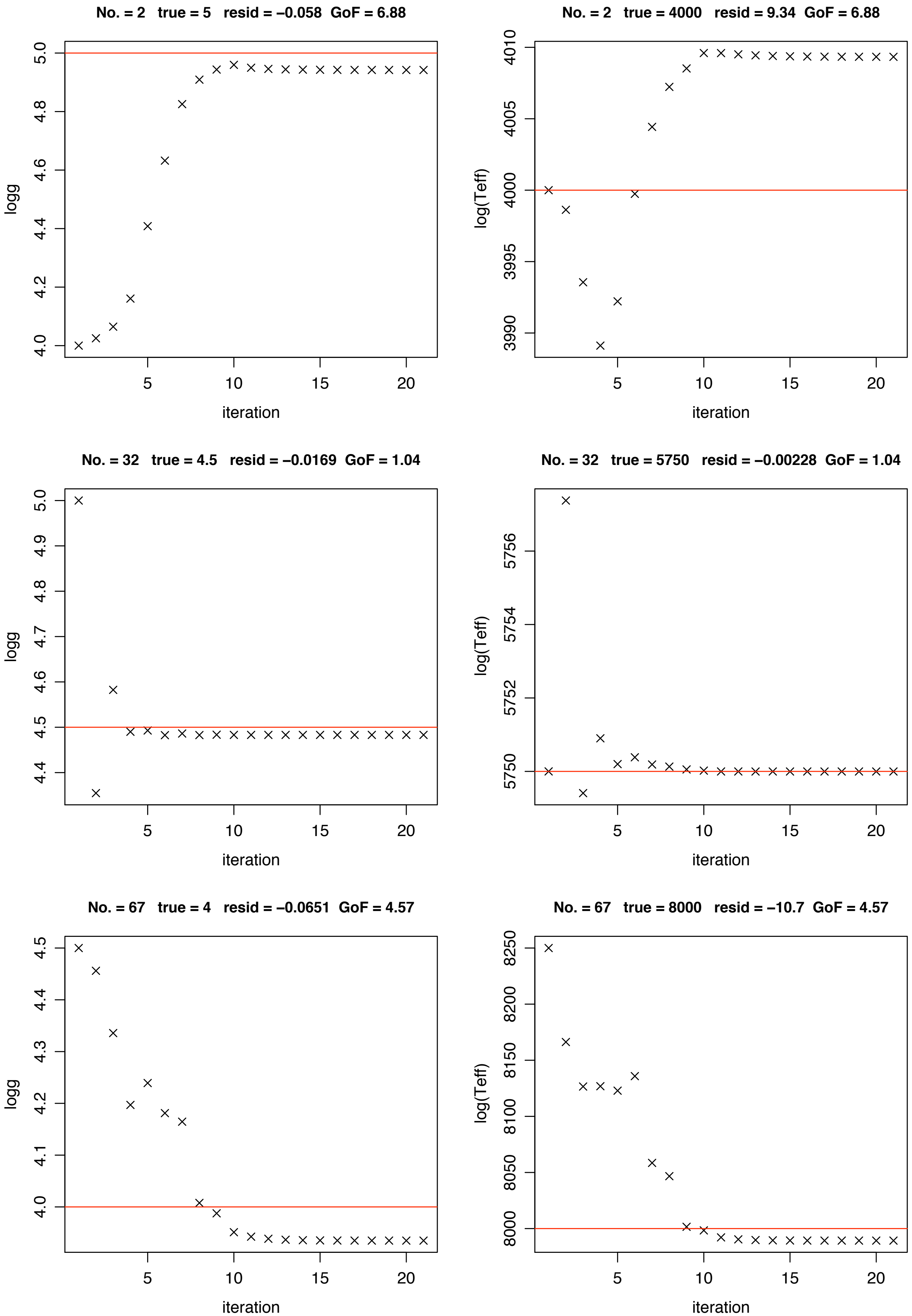


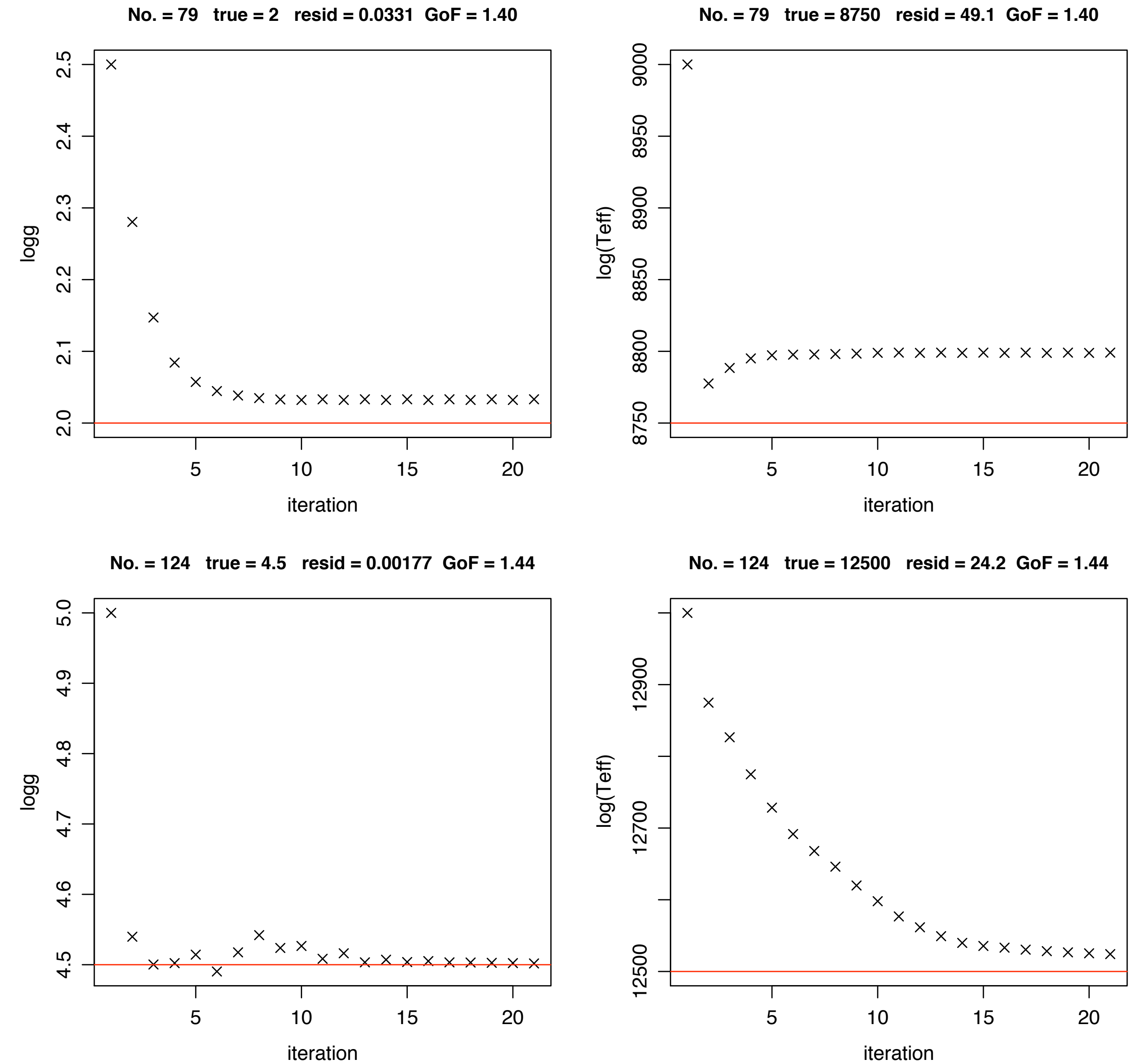
Coryn Bailer-Jones, MPIA Heidelberg

Forward model fit over logg (at 5000K)



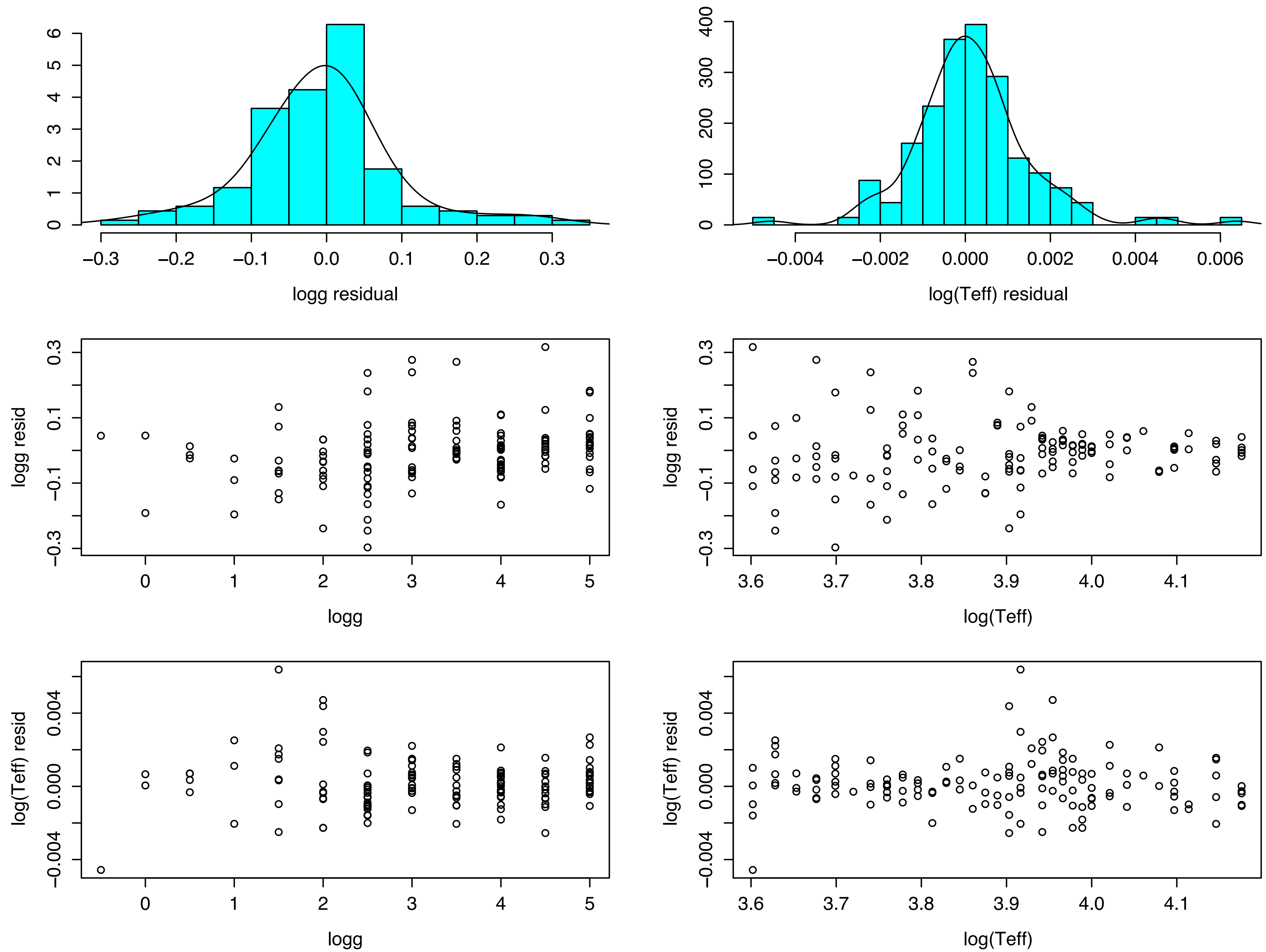
Coryn Bailer-Jones, MPIA Heidelberg





Apply ILIUIM to G=15 data

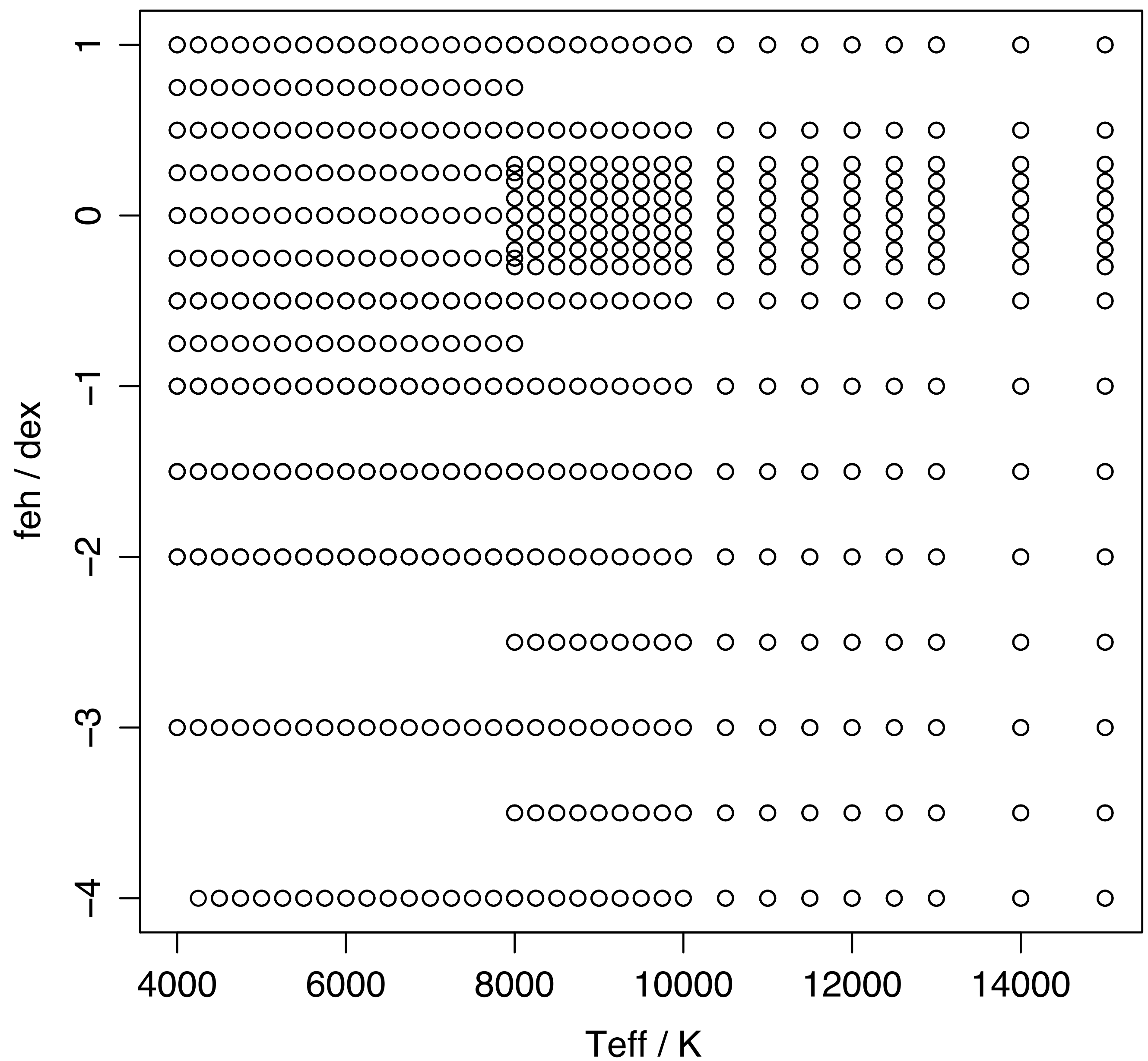
Evolution of APs during fitting for 5 stars (logg left, T_{eff} right)



	log(Teff)	logg	
Mean absolute errors:	log(Teff)=0.0010,	logg=0.065	(G=15)
	0.0057	0.35	(G=18.5)
	0.019	1.14	(G=20)

Coryn Bailer-Jones, MPIA Heidelberg

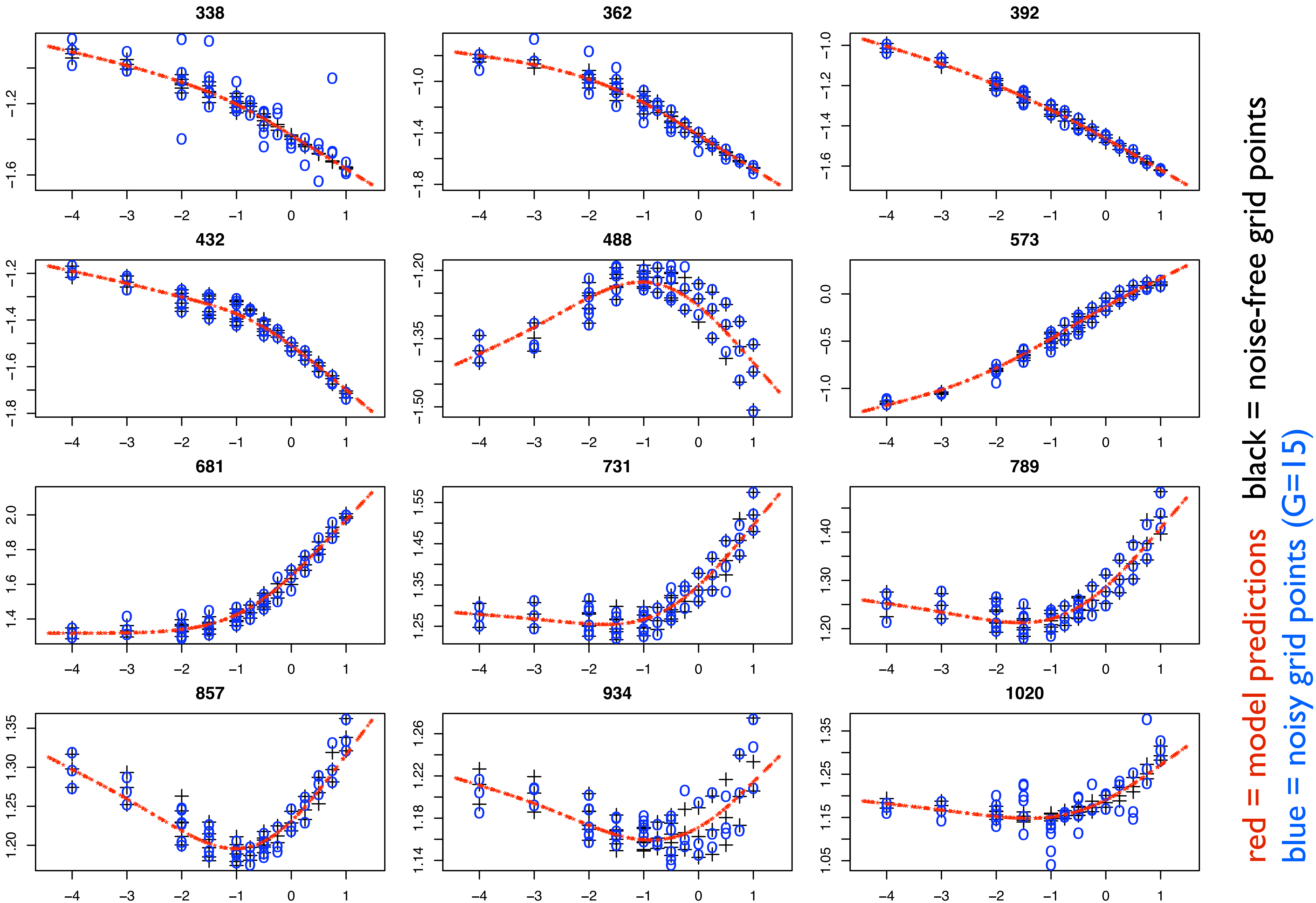
Application 2: T_{eff}/[Fe/H] estimation



Template grid

logg=(4,4.5,5)
A_V=0

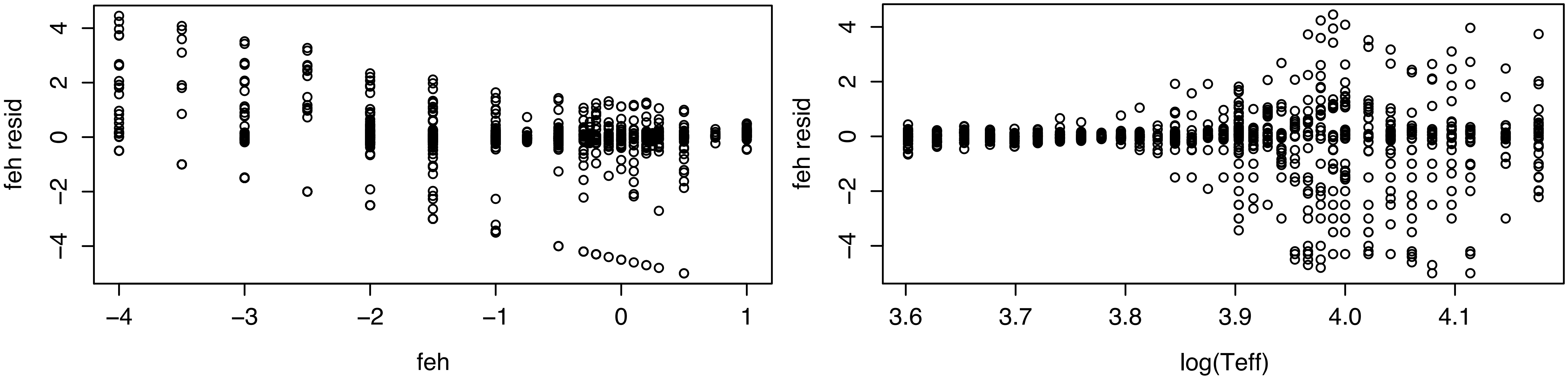
Forward model fit over [Fe/H] (at 5000K)



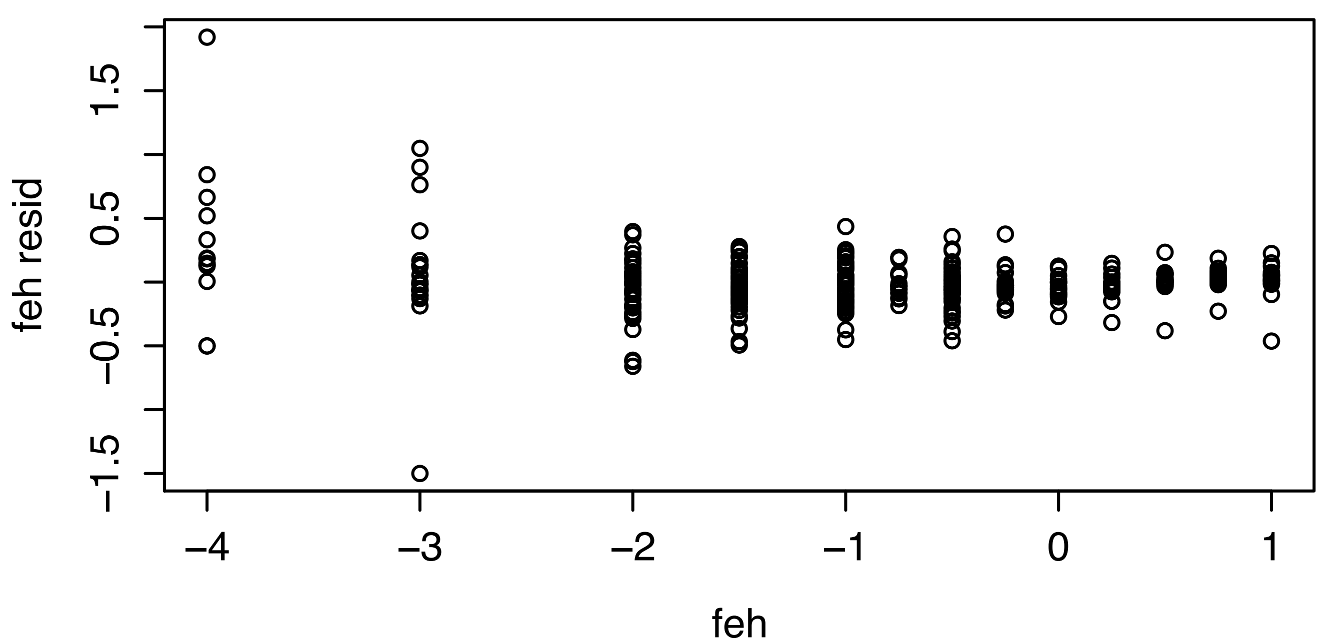
Coryn Bailer-Jones, MPIA Heidelberg

Performance on T_{eff}/[Fe/H] problem

Full T_{eff} range:



Cool stars:



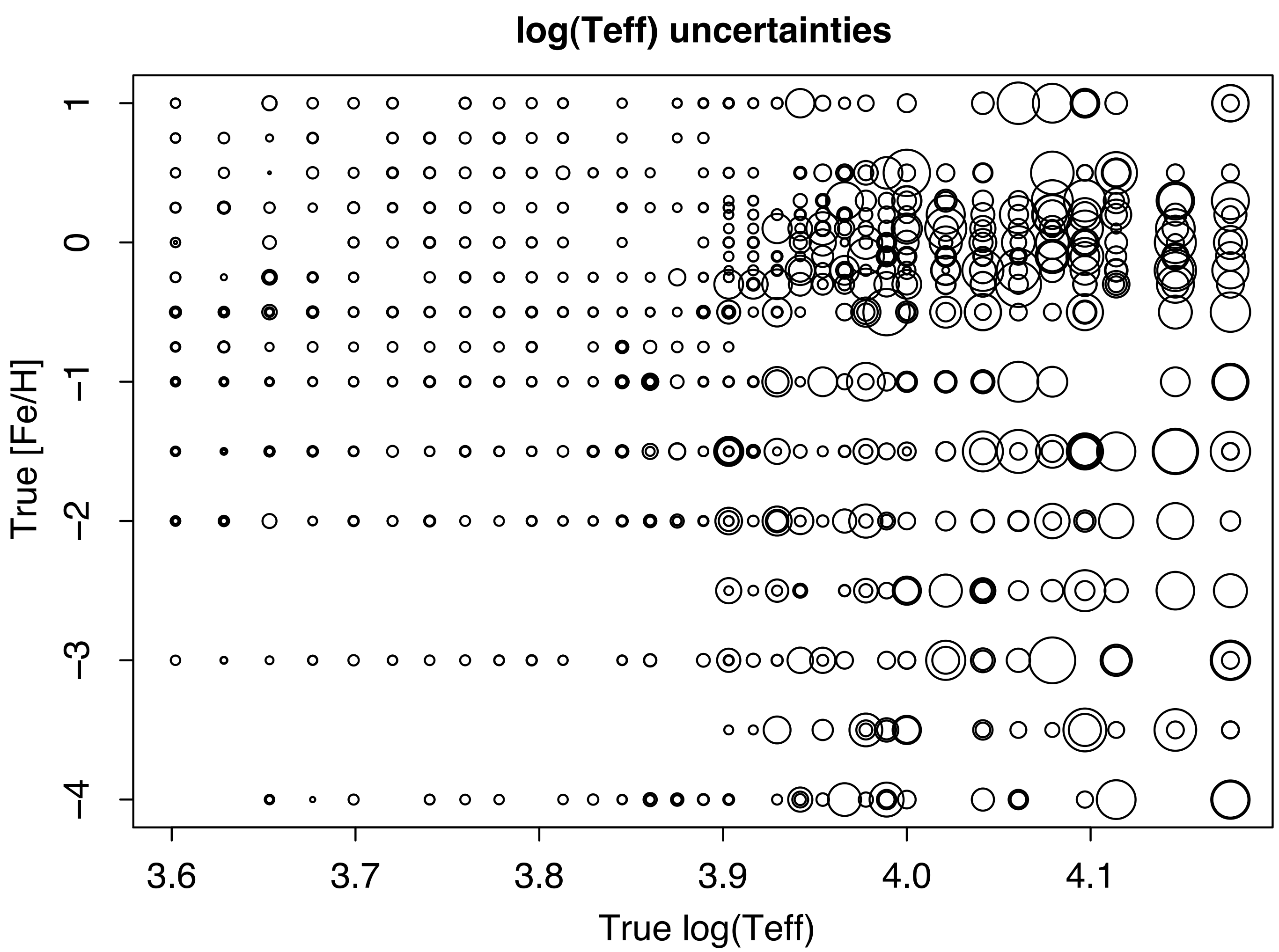
Mean abs. errors for Teff ≤ 7000K:

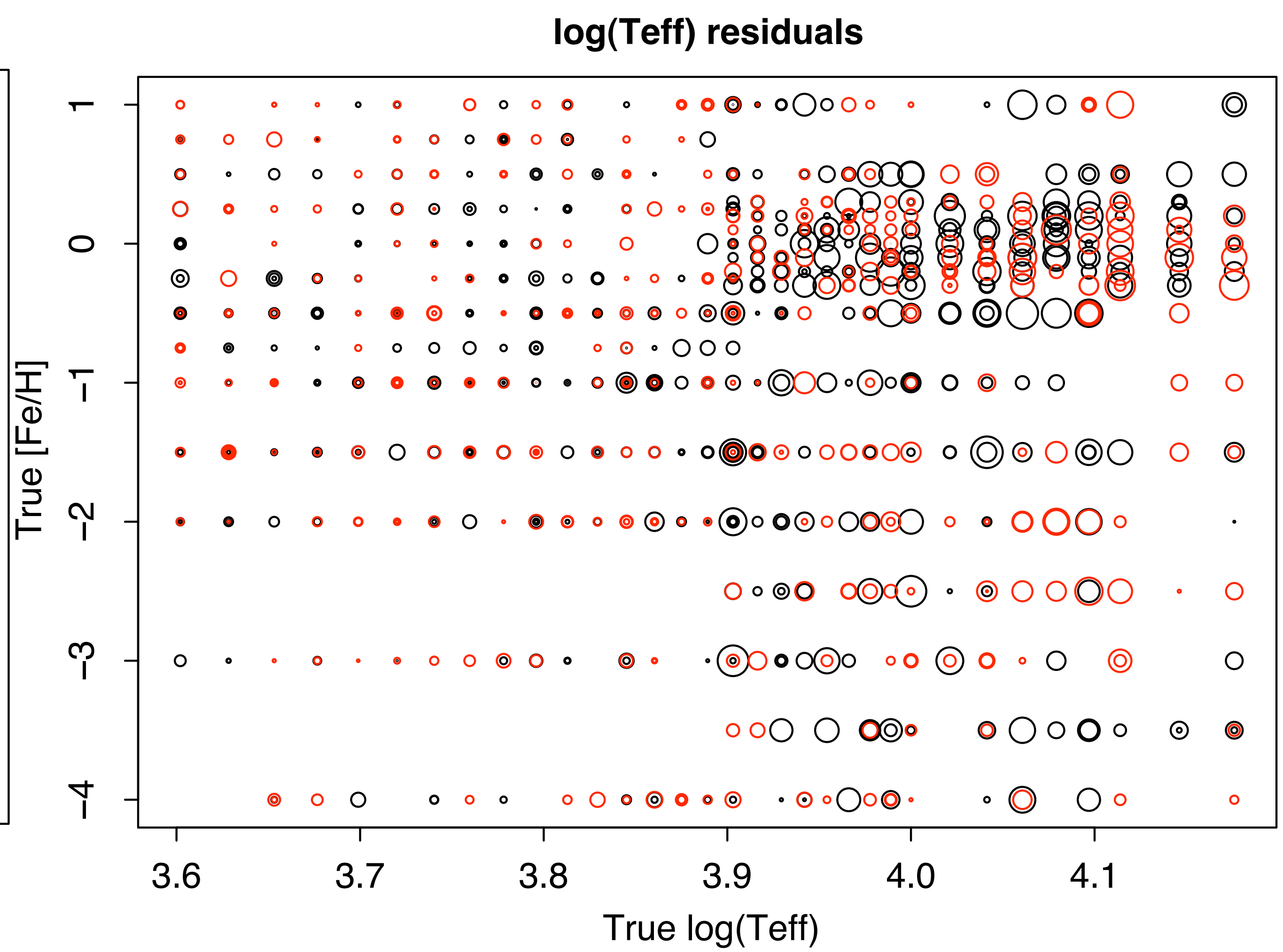
log(Teff)	[Fe/H]	
0.0017	0.14	(G=15)
0.0024	0.26	(G=18.5)
0.0070	0.82	(G=20)

ILIUM uncertainty estimation

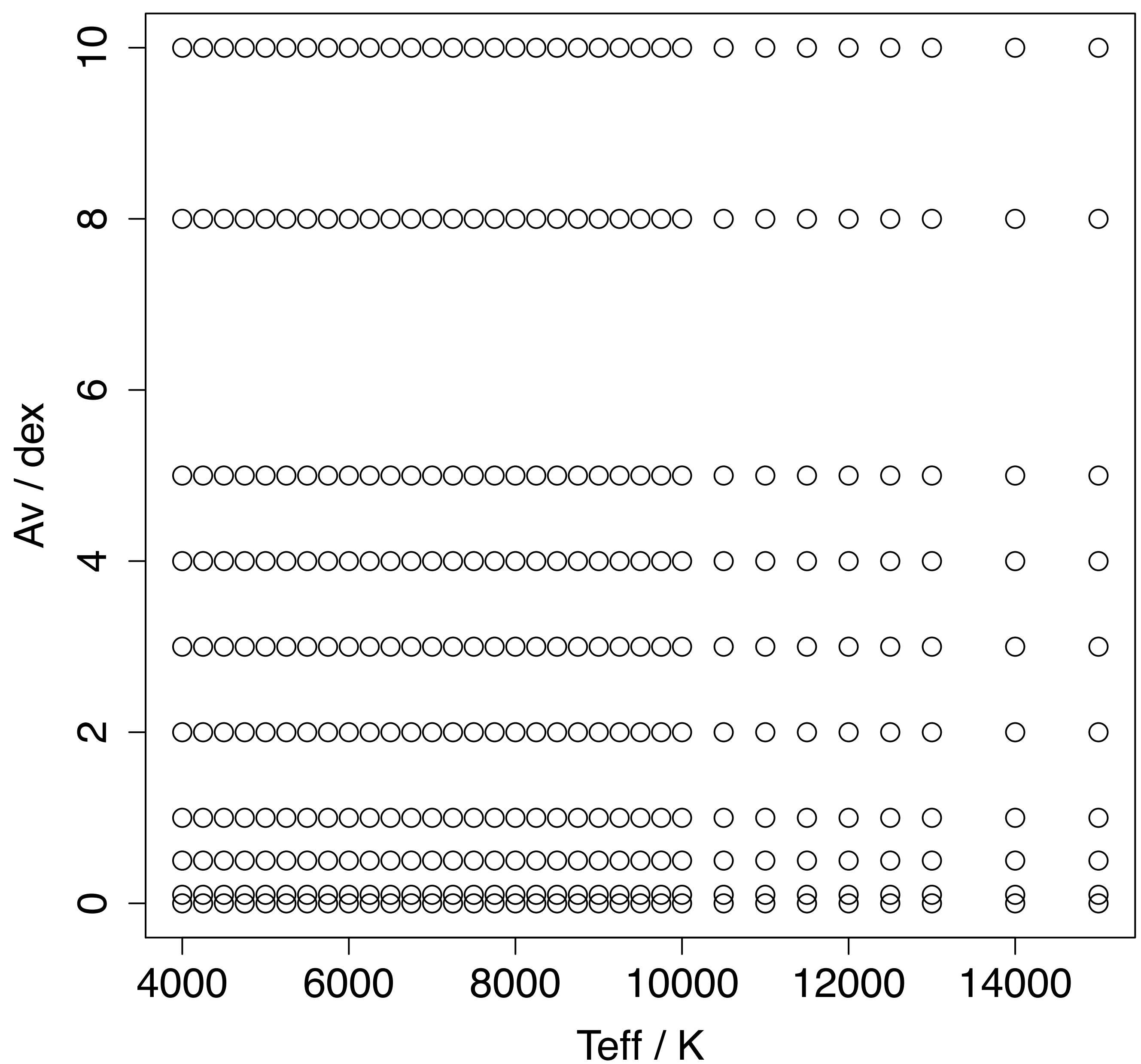
$$\mathbf{C}_\phi = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{C}_p\mathbf{S}(\mathbf{S}^T\mathbf{S})^{-1}$$

Comparison at G=20, circle area ∝ residual or $\sqrt{c_{jj}}$





Application 3: T_{eff}-A_V/logg estimation

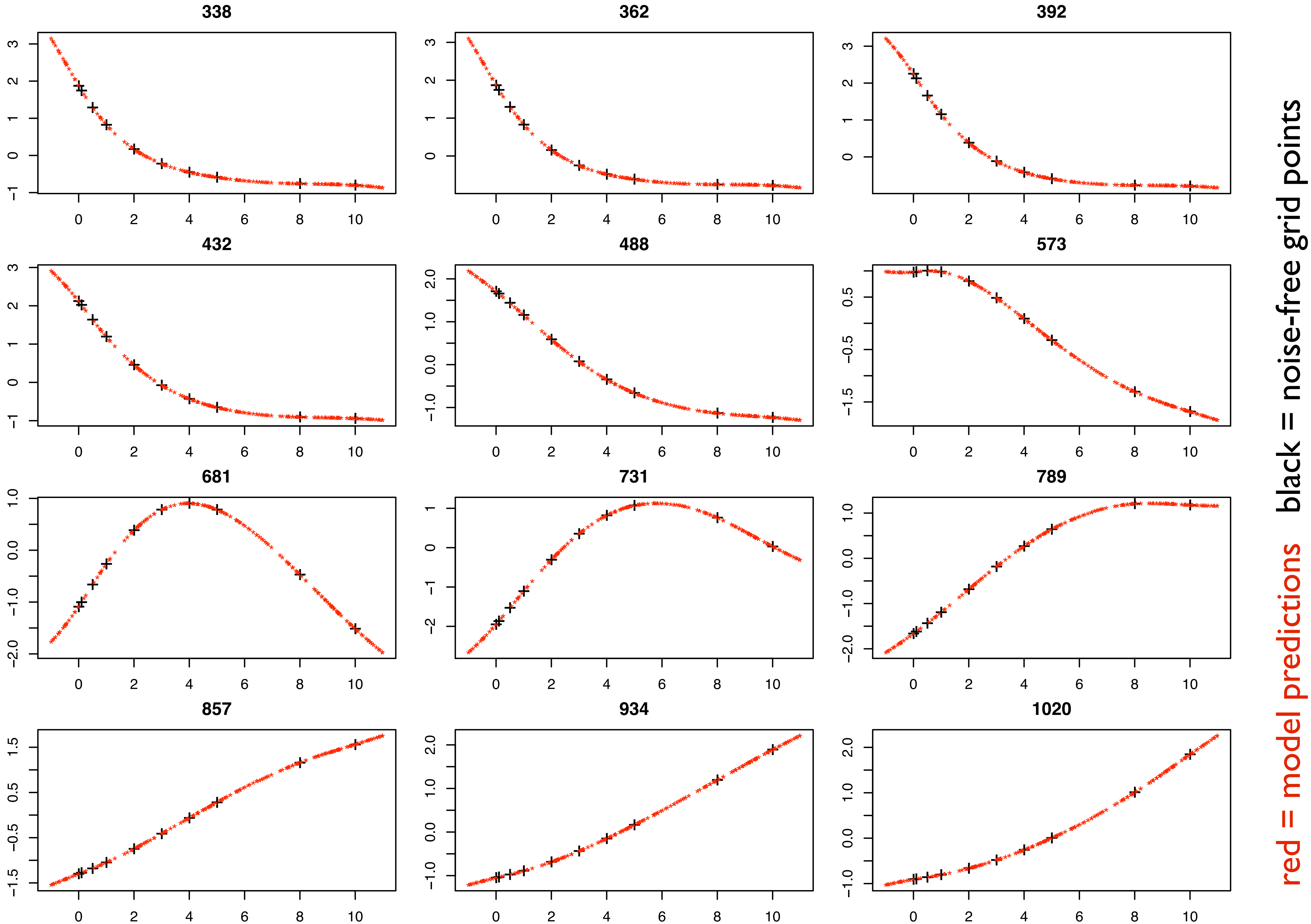


Template grid

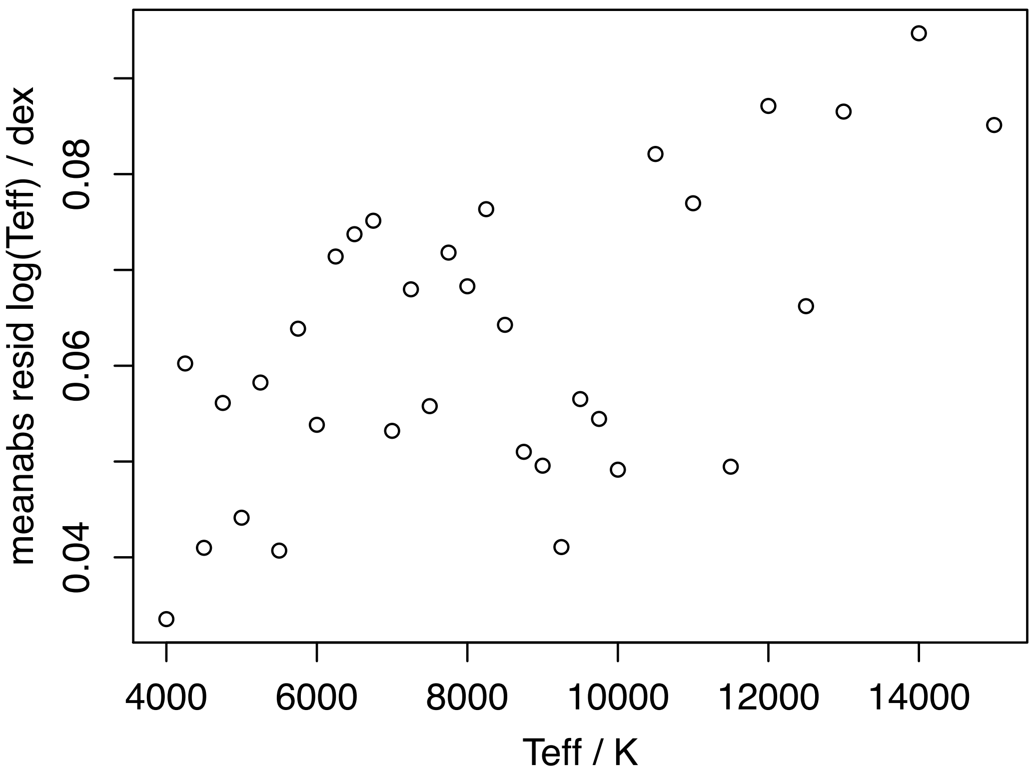
full range of logg

[Fe/H]=0

Forward model fit over A_V (at 10000K and logg=4.0)

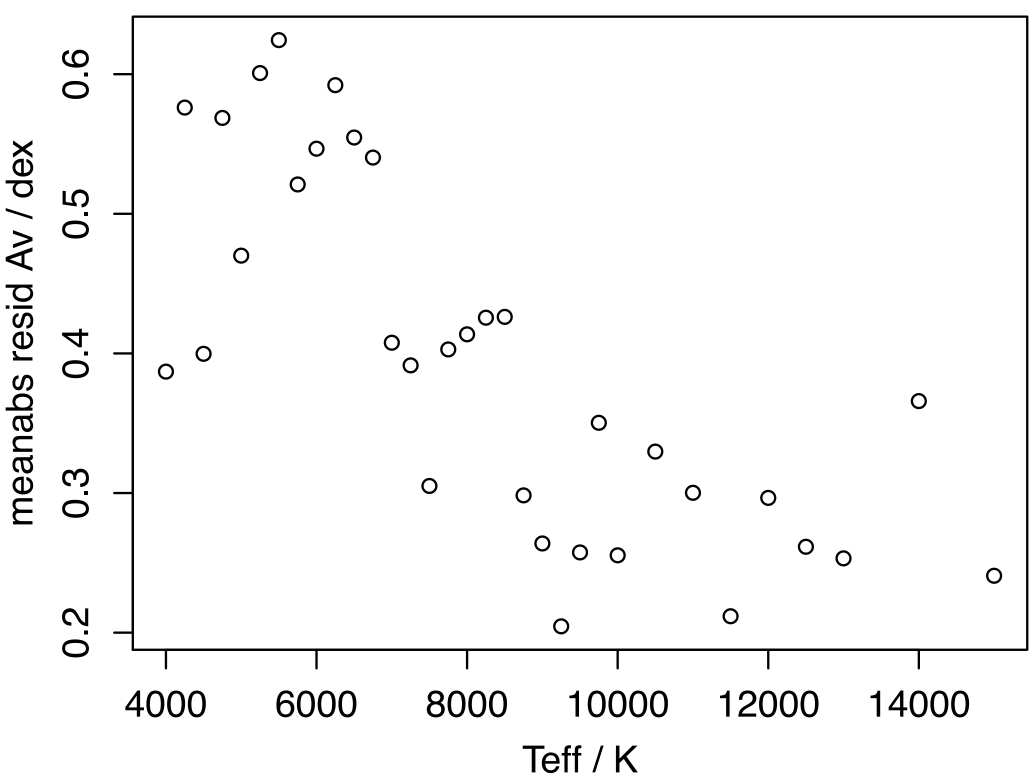


Performance on T_{eff}-Av/logg problem



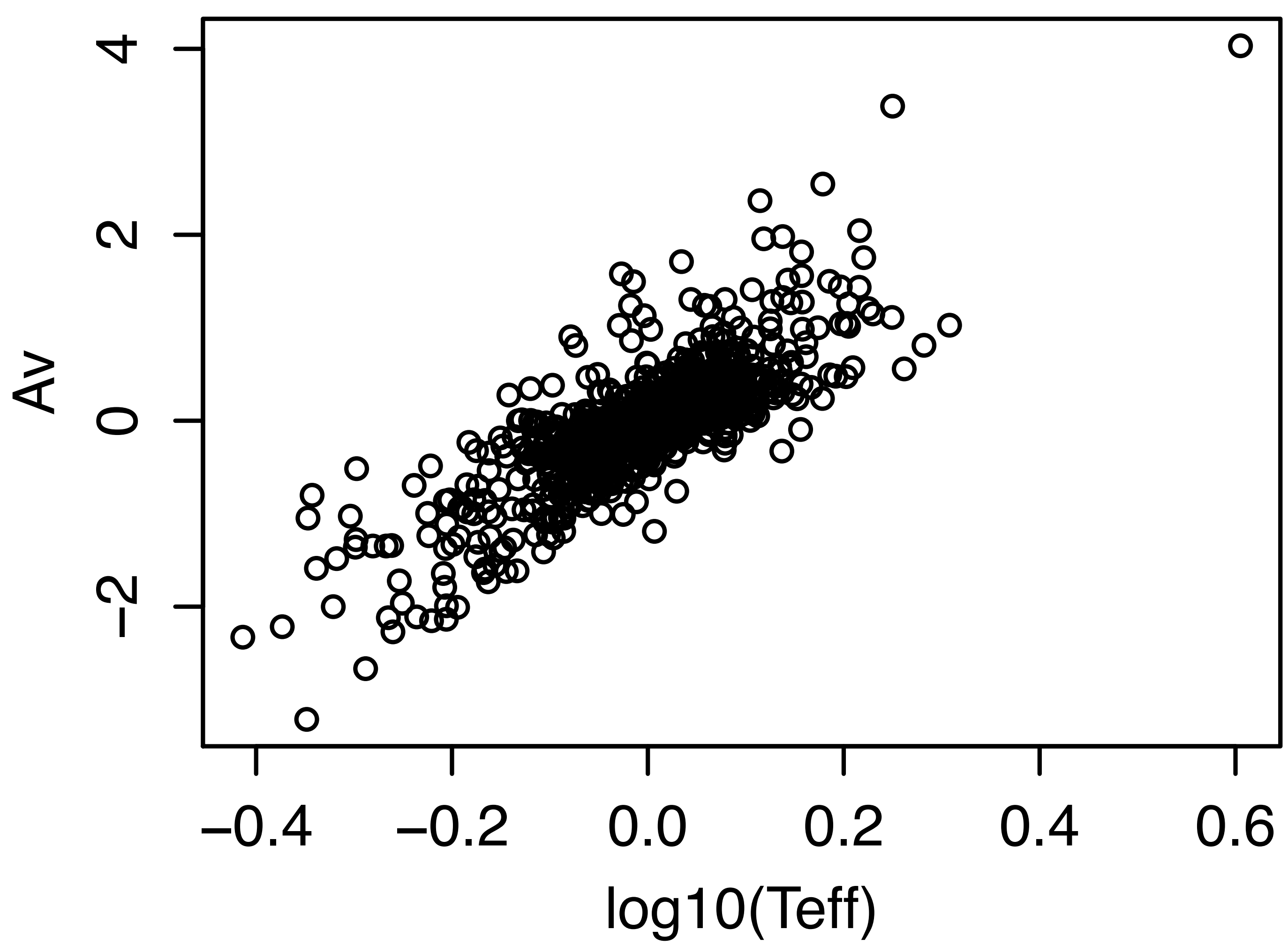
Mean absolute errors:

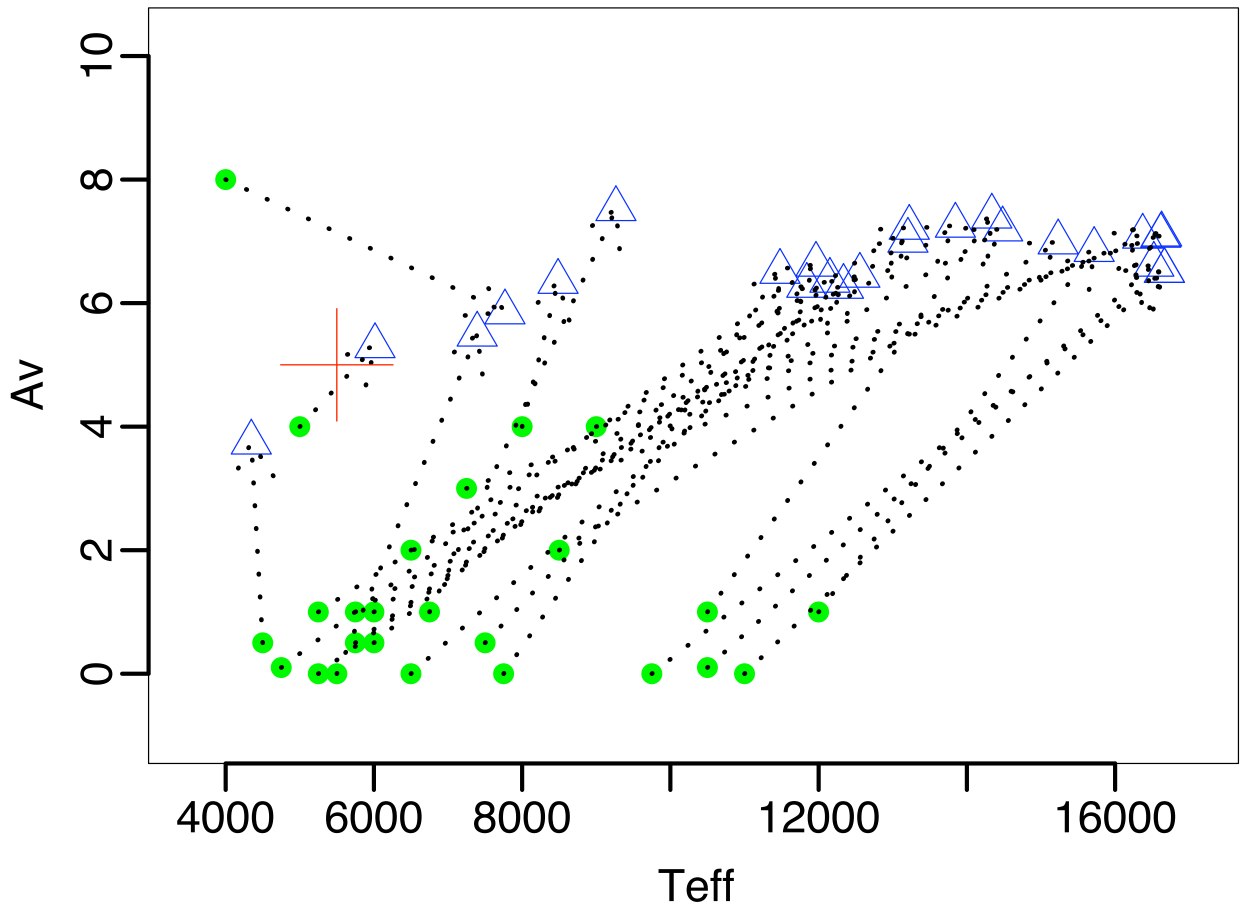
log(Teff)	Av	logg	
0.013	0.072	0.29	G=15
0.061	0.30	1.10	G=18.5

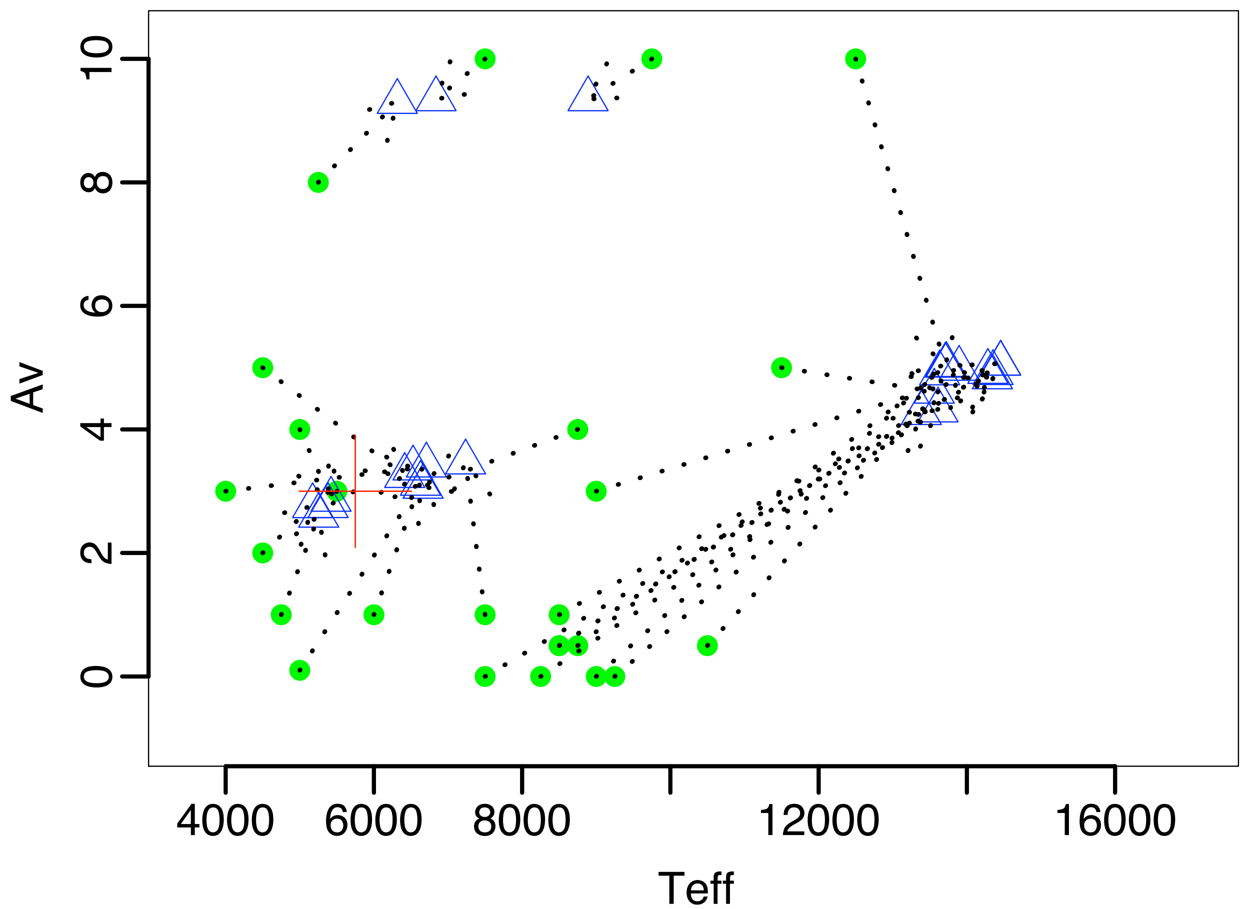


Plot:
Residuals as a function of T_{eff} at G=18.5

Correlated residuals



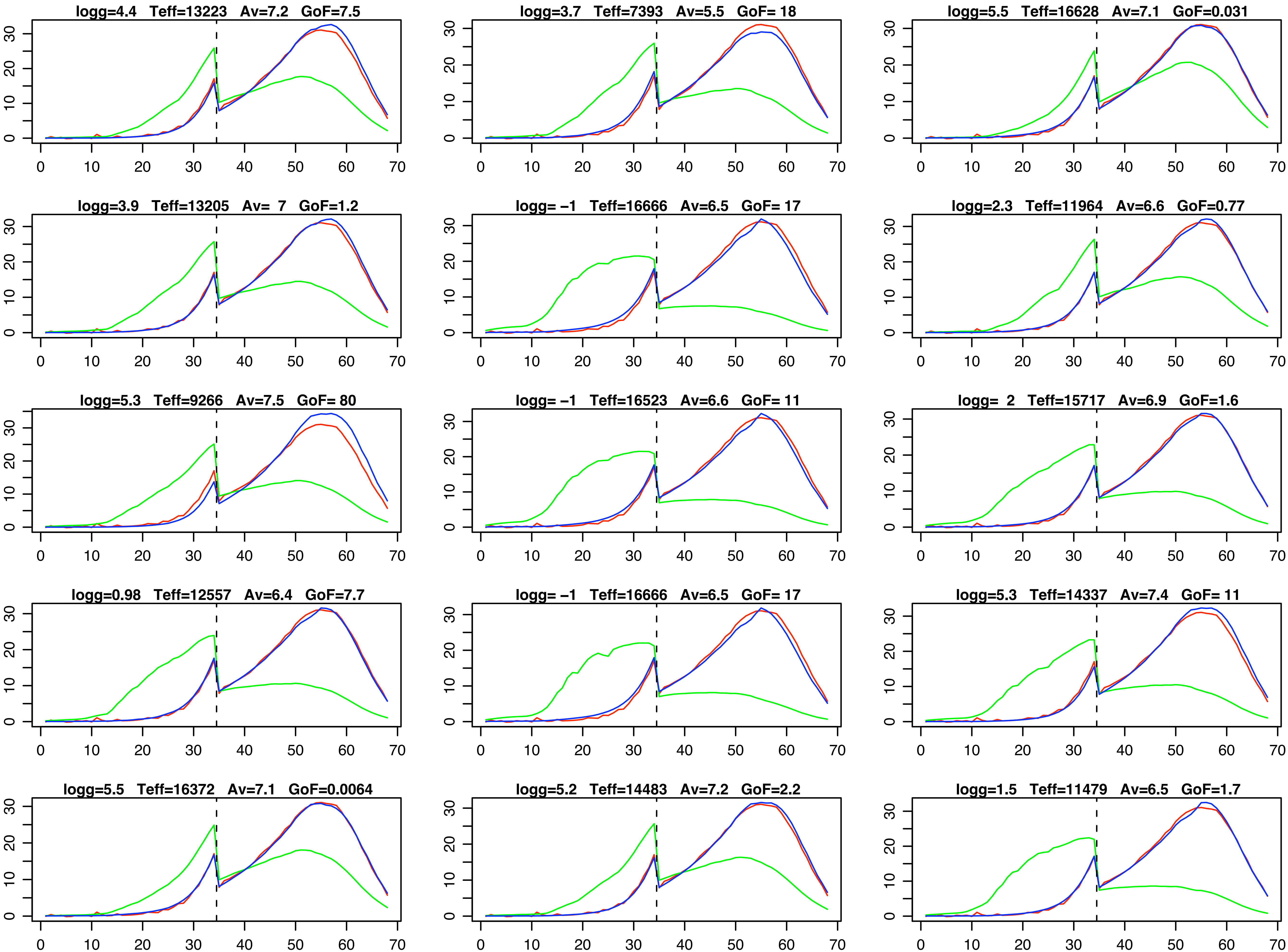




Multiple random ILIUM initializations for two stars

red cross: true APs
green circles: initial APs
blue triangles: ILIUM final APs

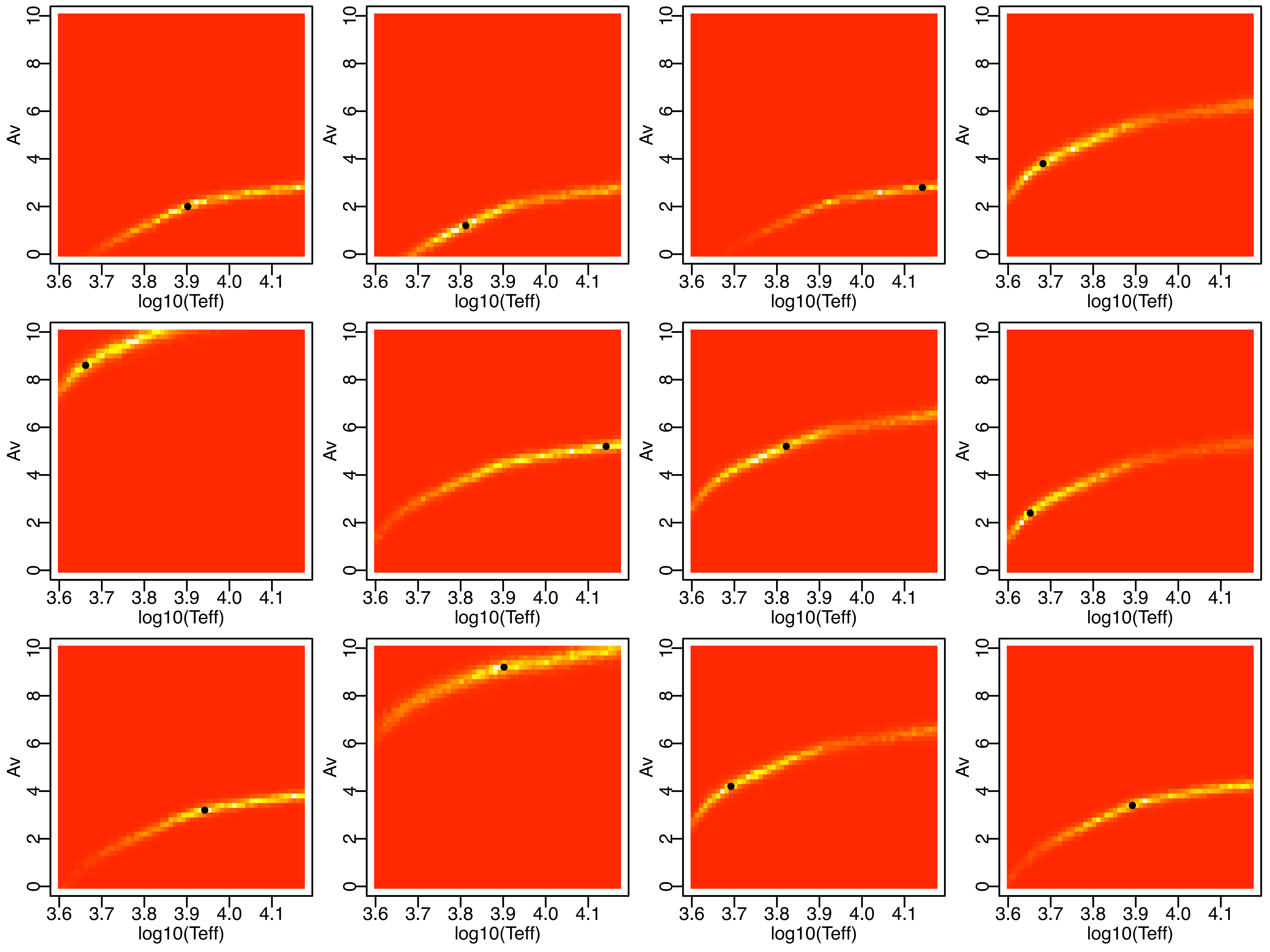
now look at corresponding (forward modelled) spectra (same colour scheme) for top plot


logg=4.4 Teff=13223 Av=7.2 GoF=7.5
logg=3.7 Teff=7393 Av=5.5 GoF= 18
logg=5.5 Teff=16628 Av=7.1 GoF=0.031
logg=3.9 Teff=13205 Av= 7 GoF=1.2
logg= −1 Teff=16666 Av=6.5 GoF= 17
logg=2.3 Teff=11964 Av=6.6 GoF=0.77
logg=5.3 Teff=9266 Av=7.5 GoF= 80
logg= −1 Teff=16523 Av=6.6 GoF= 11
logg= 2 Teff=15717 Av=6.9 GoF=1.6
logg=0.98 Teff=12557 Av=6.4 GoF=7.7
logg= −1 Teff=16666 Av=6.5 GoF= 17
logg=5.3 Teff=14337 Av=7.4 GoF= 11
logg=5.5 Teff=16372 Av=7.1 GoF=0.0064
logg=5.2 Teff=14483 Av=7.2 GoF=2.2
logg=1.5 Teff=11479 Av=6.5 GoF=1.7

T_{eff} - A_V degeneracy

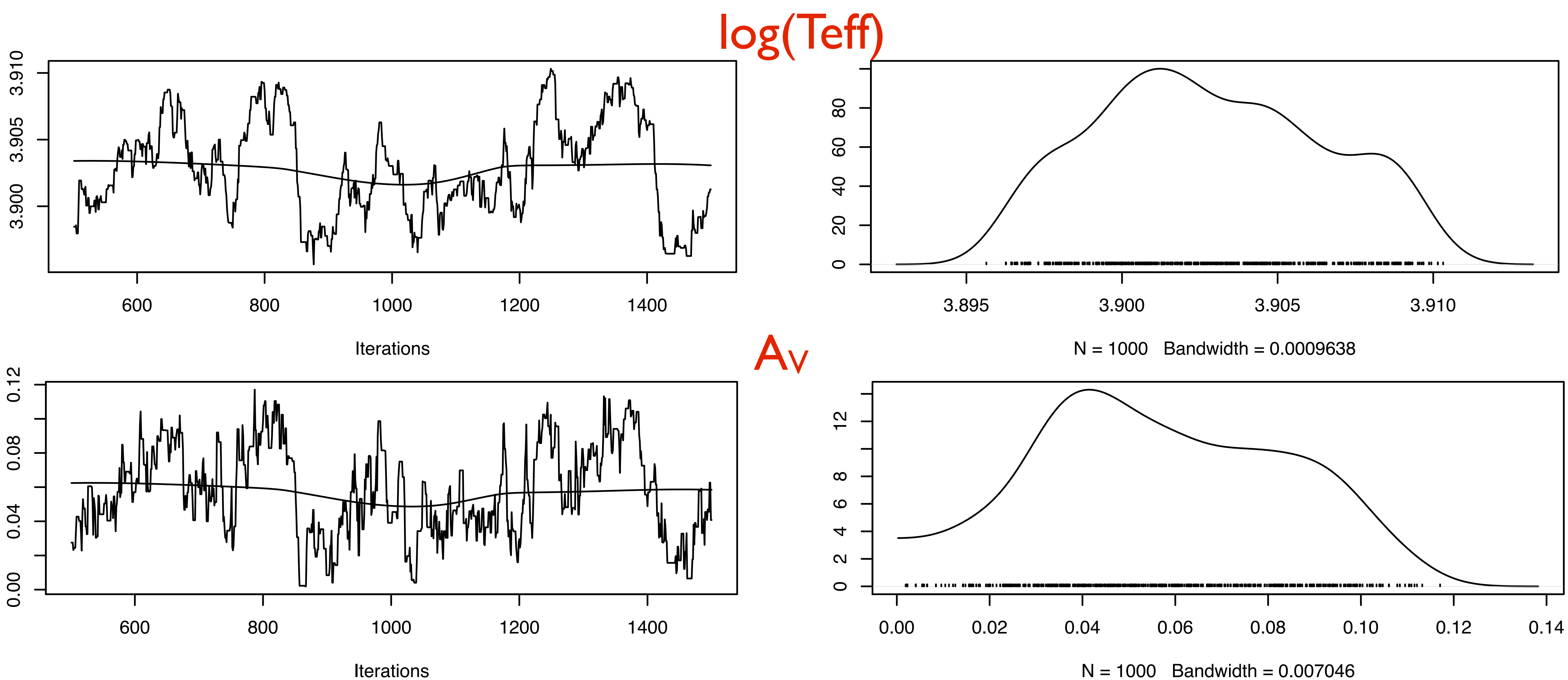
- Degeneracy: single spectrum corresponds to multiple AP solutions (spectra the same to within the noise)
- Can systematically map the degeneracy
  - next page: each panel shows inverse distance squared between some noise-free spectrum (black point) and every other (noisy) spectrum in T_{eff}/A_V grid, on a colour scale (white close, red far)

$$\begin{aligned} D^2 &= \delta\boldsymbol{p}^T \mathbf{C}_p^{-1} \delta\boldsymbol{p} \\ &= \sum_{i=1}^{i=I} \left(\frac{p_i - n_i}{\sigma_{p_i}} \right)^2 \end{aligned}$$


Av
log10(Teff)
0
2
4
6
8
10
3.6
3.7
3.8
3.9
4.0
4.1

Full probabilistic extension

- Given forward model, define a likelihood function
  - e.g. $P'(D^2) \propto e^{-D^2/2}$
- sample this (e.g. MCMC) to get a probability density function



Conclusions

- forward modelling brings significant advantages
  - avoid inverse problem
  - sensitivity weighting, uncertainty and GoF estimation, model strong and weak APs separately
- ILIUM: sensitivity-based, local, iterative method
- good T_{eff}, A_V performance to G=18.5; [Fe/H] and logg good to at least G=15
- but strong and ubiquitous T_{eff}-A_V degeneracy in BP/RP
  - no unique solution
  - can map in advance: ILIUM etc. to locate the degeneracy ridge
  - use additional information where possible to reduce it