A Bayesian method for inference of fundamental stellar parameters

November 7, 2013

Ian Czekala[1] and Sean M. Andrews[1]

iczekala@cfa.harvard.edu

1. Introduction

Other methods of stellar parameter inference: SPC(Buchhave et al. 2012), MOOG, SME (Valenti & Piskunov 1996).

2. The model

To summarize our model, we use Bayes rule to design a posterior probability function

$$\overbrace{p(\vec{\theta}|\vec{D})}^{\text{posterior}} \propto \underbrace{p(\vec{D}|\vec{\theta})}_{\text{likelihood}} \overbrace{p(\vec{\theta})}^{\text{prior}} \tag{1}$$

where the parameter vector is comprised of both stellar parameters and calibration parameters $\vec{\theta} = \{\vec{\theta}_\star, \vec{\theta}_\mathrm{N}\}$. Stellar parameters include $\vec{\theta}_\star = \{T_\mathrm{eff}, \log(g), v \sin i, v_z, A_v\}$ while the data vector $\vec{D} = f_\mathrm{D}(\lambda)$ represents the a high resolution spectrum. In our immediate case of our companion paper, this is a 51-order echelle spectrum from TRES at a resolution of $R = 48,000$ (6.8 km/s) and spanning the full optical range.

Although per-pixel photon counting errors are Poisson, we can safely assume that we have enough counts that we are in the Gaussian limit. Then our likelihood function becomes a pixel-by-pixel χ^2 comparison between the data spectrum f_D, and the model

[1]Harvard-Smithsonian Center for Astrophysics, 60 Garden Street MS 10, Cambridge, MA 02138

spectrum f_{M}, summed over the wavelength axis.[1]

$$p(\vec{D}|\vec{\theta}) = \mathcal{L} \propto \exp\left(-\frac{\chi^2}{2}\right) \tag{2}$$

$$\chi^2 = \sum_\lambda \left[\frac{f_{\mathrm{D}}(\lambda) - k(\lambda|\vec{\theta}_{\mathrm{N}}) f_{\mathrm{M}}(\lambda|\vec{\theta}_\star)}{\sigma(\lambda)}\right]^2 \tag{3}$$

In the future, we could modify our likelihood function to use a Student-t statistic instead of a Gaussian, such that it would be more robust in the presence of outlier pixels, however this does not seem necessary at the moment. To start, we will assume flat priors on $\vec{\theta}_\star$ and $\vec{\theta}_{\mathrm{N}}$ and then relax this assumption later. Then the logarithm of our posterior probability function is then

$$\ln\left[p(\vec{\theta}|\vec{D})\right] \propto -\frac{1}{2}\sum_\lambda \left[\frac{f_{\mathrm{D}}(\lambda) - k(\lambda|\vec{\theta}_{\mathrm{N}}) f_{\mathrm{M}}(\lambda|\vec{\theta}_\star)}{\sigma(\lambda)}\right]^2 \tag{4}$$

To generate a model spectrum given a set of parameters $f_{\mathrm{M}}(\vec{\theta})$, a model spectrum is created by linearly interpolating between grid points of a high-resolution PHOENIX model stellar spectra ($R = 500,000$), rotationally broadened, instrumentally broadened, Doppler shifted, extinction corrected, and downsampled to the exact pixels of the TRES spectrum. $\sigma(\lambda)$ is the RMS scatter of the continuum (measured in a line-free region) inversely scaled by the blaze function in each order, to account for Poisson photon counting errors. For some data reduction pipelines, a per-pixel "sigma spectrum" that accounts for spectral extraction errors due to night sky line contamination or low signal to noise is available, and should be used instead. The prefactor $k(\lambda|\vec{\theta}_{\mathrm{N}})$ is a systematic error term that aims to account for errors in the blaze-correction or flux-calibration. In our case, $\vec{\theta}_{\mathrm{N}} = \{c_0, c_1, \ldots, c_N\}$ are a set of 4 Chebyshev polynomial coefficients for each order. This amounts to a "re-fluxing" of the model following the techniques of Eisenstein et al. (2006). Additionally, we can add priors to limit the degree of re-fluxing (for example 20%) to be consistent with a systematic noise floor for flux-calibration determined independently through calibration tests on TRES using standard stars. We scale the model and introduce non-physical warps rather than attempt to re-flux the data since this operation would also need to re-scale the noise, and the χ^2 would be complicated function of k.

---

[1]We must necessarily mask certain wavelength regions of the T Tauri spectra that are contaminated (one might also say "made more interesting") by the astrophysical realities of accretion or stellar spots (such as the Balmer lines, which are in emission). We also accept the assumption that rotation or accretion onto the star does not fundamentally alter the structure of the model atmosphere that generates the model spectrum and that a spectral comparison of the "clean" regions of the spectrum is valid.

2.1. Nuisance parameters

In order to understand the effects of the $k(\lambda|\vec{\theta}_{\rm N})$ term, we now consider its effects on just one order. The first four Chebyshev polynomials are

$$\begin{aligned}
T_0(x) &= 1 \\
T_1(x) &= x \\
T_2(x) &= 2x^2 - 1 \\
T_3(x) &= 4x^3 - 3x
\end{aligned}$$

In our implementation, we map the full range of a specific order $\lambda \in [\lambda_{\rm min}, \lambda_{\rm max}]$ to $x \in [-1, 1]$ (or we can map pixel number). Then, for a given set of Chebyshev coefficients, $\vec{\theta}_{\rm N} = \vec{c} = \{c_0, c_1, c_2, \ldots, c_N\}$, we have

$$k(\lambda|\vec{\theta}_{\rm N}) = c_0 T_0(\lambda) + c_1 T_1(\lambda) + \ldots + c_N T_N(\lambda) \tag{5}$$

$$k(\lambda|\vec{\theta}_{\rm N}) = \sum_{i=0}^{N} c_i T_i(\lambda) \tag{6}$$

using this, we can rewrite Equation 4 as

$$\ln\big[p(\vec{\theta}|\vec{D})\big] \propto -\frac{1}{2}\sum_{\lambda}\left[\frac{f_{\rm D}(\lambda) - \big[\sum_{i=0}^{N} c_i T_i(\lambda)\big] f_{\rm M}(\lambda|\vec{\theta}_\star)}{\sigma(\lambda)}\right]^2 \tag{7}$$

To simplify the following discussion, we consider just one wavelength λ_i and then generalize this to a sum over λ later. If we expand out the square in Equation 7, then for a given λ_i we have

$$\ln\big[p(\vec{\theta}|\lambda_i)\big] \propto -\frac{1}{2\sigma^2}\big[f_{\rm D}^2 - 2 f_{\rm D} f_{\rm M} \sum_n c_n T_n + f_{\rm M}^2 \sum_n \sum_m c_n T_n c_m T_m\big] \tag{8}$$

where σ, $f_{\rm D}$, $f_{\rm M}$, and $T_{n,m}$ are all evaluated at λ_i

To simplify this math, we can rewrite the Chebyshev coefficients and polynomials as column vectors $\vec{c}$ and $\vec{T}(\lambda)$, respectively

$$\vec{c} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_N \end{bmatrix} \qquad \vec{T}(\lambda) = \begin{bmatrix} T_0(\lambda) \\ T_1(\lambda) \\ \vdots \\ T_N(\lambda) \end{bmatrix} \tag{9}$$

then

$$k(\lambda|\vec{\theta}_{\mathrm{N}}) = \sum_{i=0}^{N} c_i T_i = \vec{T}^{\mathsf{T}} \cdot \vec{c} \tag{10}$$

If we let

$$\boldsymbol{W}(\lambda) = \vec{T} \cdot \vec{T}^{\mathsf{T}} = \begin{bmatrix} T_0 T_0 & T_0 T_1 & \dots & T_0 T_N \\ T_1 T_0 & T_1 T_1 & \dots & T_1 T_N \\ \vdots & \vdots & \ddots & \vdots \\ T_N T_0 & T_N T_1 & \dots & T_N T_N \end{bmatrix} \tag{11}$$

then we have

$$\sum_n \sum_m c_n T_n c_m T_m = \vec{c}^{\mathsf{T}} \cdot \boldsymbol{W} \cdot \vec{c} \tag{12}$$

and we can rewrite Equation 8 as

$$\ln\left[p(\vec{\theta}|\lambda_i)\right] \propto -\frac{1}{2\sigma^2}\left[f_{\mathrm{D}}^2 - 2 f_{\mathrm{D}} f_{\mathrm{M}}\, \vec{T}^{\mathsf{T}} \vec{c} \, + f_{\mathrm{M}}^2\, \vec{c}^{\mathsf{T}} \boldsymbol{W} \vec{c}\right] \tag{13}$$

Because matrix multiplication is associative, we can sum $\boldsymbol{W}(\lambda)$ and $\vec{T}^{\mathsf{T}}(\lambda)$ across all λ, and define

$$\boldsymbol{A} = \sum_{\lambda} \frac{f_{\mathrm{M}}^2(\lambda|\vec{\theta}_\star)}{\sigma^2(\lambda)} \boldsymbol{W}(\lambda) \tag{14}$$

$$\vec{B}^{\mathsf{T}}(\vec{\theta}_\star) = \sum_{\lambda} \frac{f_{\mathrm{D}}(\lambda) f_{\mathrm{M}}(\lambda|\vec{\theta}_\star)}{\sigma^2(\lambda)} \vec{T}^{\mathsf{T}}(\lambda) \tag{15}$$

$$g(\vec{\theta}_\star) = -\frac{1}{2} \sum_{\lambda} \frac{f_{\mathrm{D}}(\lambda)}{\sigma^2(\lambda)} \tag{16}$$

$$\tag{17}$$

Now we can rewrite Equation 7 as

$$\ln\bigl(p(\vec{\theta}_\star, \vec{\theta}_{\mathrm{N}}|\vec{D})\bigr) \propto -\frac{1}{2}\vec{c}^{\mathsf{T}} \boldsymbol{A} \vec{c} + \vec{B}^{\mathsf{T}}(\vec{\theta}_\star)\vec{c} + g(\vec{\theta}_\star) \tag{18}$$

where remember $\vec{\theta}_{\mathrm{N}} = \{c_0, c_1, \dots, c_N\}$. Now we have

$$\boxed{p(\vec{\theta}|\vec{D}) = p(\vec{\theta}_\star, \vec{\theta}_{\mathrm{N}}|\vec{D}) \propto \exp\left(-\frac{1}{2}\vec{c}^{\mathsf{T}} \boldsymbol{A} \vec{c} + \vec{B}^{\mathsf{T}}(\vec{\theta}_\star)\vec{c} + g(\vec{\theta}_\star)\right)} \tag{19}$$

This is one of the `lnprob` options to sample in, provided you are interested in the results of the Chebyshev coefficients. However, because this is a multi-dimensional Gaussian in the Chebyshev coefficients and the parameters we are most interested in are $\vec{\theta}_\star$, we can analytically marginalize out the nuisance parameters

$$p(\vec{\theta}_\star|\vec{D}) = \int p(\vec{\theta}_\star, \vec{\theta}_{\rm N}|\vec{D})\, {\rm d}\vec{\theta}_{\rm N} \tag{20}$$

The analytic multi-dimensional Gaussian integral with linear term (Stoof et al. 2009) yields

$$p(\vec{\theta}_\star|\vec{D}) = \int \exp\left(-\frac{1}{2}\vec{c}^{\mathsf{T}}\boldsymbol{A}\vec{c} + \vec{B}^{\mathsf{T}}(\vec{\theta}_\star)\vec{c} + g(\vec{\theta}_\star)\right)\, {\rm d}\vec{c} \tag{21}$$

$$p(\vec{\theta}_\star|\vec{D}) = \sqrt{\frac{(2\pi)^N}{\det|\boldsymbol{A}|}} \exp\left(\frac{1}{2}\vec{B}^{\mathsf{T}}\boldsymbol{A}^{-1}\vec{B} + g\right) \tag{22}$$

This result means that for any given set of parameters $\vec{\theta}_\star$, we only need to calculate $\vec{B}(\vec{\theta}_\star)$ and $g(\vec{\theta}_\star)$; $\boldsymbol{A}(\vec{D})$ can be computed once and stored for the remainder of the MCMC run. We have also reduced the dimensionality of our posterior space from 150+ parameters (~ 6 stellar parameters plus 3 Chebyshev coefficients for each of 51 echelle orders) to ~ 6, which will enable more rapid convergence of the Markov chain.

2.2. Including Gaussian priors on nuisance parameters

In order to constrain the degree of "re-fluxing" to be consistent with an independent determination of the systematic error floor in flux-calibration (for example, 20%), we can add Gaussian priors that limit the coefficient of each Chebyshev polynomial to within a percentage of no correction. For example, we might wish to limit the multiplicative linear term c_1 to within an amplitude of ± 0.2.

This corresponds to a prior on the nuisance parameters of

$$p(\vec{\theta}_{\rm N}) \propto \exp\left(-\frac{1}{2}(\vec{c} - \vec{\mu})^{\mathsf{T}}\boldsymbol{D}(\vec{c} - \vec{\mu})\right) \tag{23}$$

where

$$\boldsymbol{D} = \begin{bmatrix} \sigma_0^{-2} & 0 & \dots & 0 \\ 0 & \sigma_1^{-2} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \sigma_N^{-2} \end{bmatrix} \tag{24}$$

σ_i represents the width of the Gaussian prior on the i-th Chebyshev coefficient. We can rewrite Equation 1 with priors as

$$p(\vec{\theta}_\star, \vec{\theta}_{\rm N}|\vec{D}) \propto p(\vec{D}|\vec{\theta}_\star, \vec{\theta}_{\rm N}) p(\vec{\theta}_{\rm N}) p(\vec{\theta}_\star) \tag{25}$$

$$p(\vec{\theta}_\star, \vec{\theta}_{\rm N}|\vec{D}) \propto \exp\left(-\frac{1}{2}\vec{c}^{\mathsf{T}}\boldsymbol{A}\vec{c} + \vec{B}^{\mathsf{T}}(\vec{\theta}_\star)\vec{c} + g(\vec{\theta}_\star)\right) \exp\left(-\frac{1}{2}(\vec{c} - \vec{\mu})^{\mathsf{T}}\boldsymbol{D}(\vec{c} - \vec{\mu})\right) p(\vec{\theta}_\star) \tag{26}$$

we can expand and rearrange the argument of Equation 23 to a similar form

$$-\frac{1}{2}(\vec{c} - \vec{\mu})^{\mathsf{T}}\boldsymbol{D}(\vec{c} - \vec{\mu}) = \frac{1}{2}\left(-\vec{c}^{\mathsf{T}}\boldsymbol{D}\vec{c} + \vec{c}^{\mathsf{T}}\boldsymbol{D}\vec{\mu} + \vec{\mu}^{\mathsf{T}}\boldsymbol{D}\vec{c} - \vec{\mu}^{\mathsf{T}}\boldsymbol{D}\vec{\mu}\right) \tag{27}$$

$$= -\frac{1}{2}\vec{c}^{\mathsf{T}}\boldsymbol{D}\vec{c} + (\boldsymbol{D}\vec{\mu})^{\mathsf{T}}\vec{c} - \frac{1}{2}\vec{\mu}^{\mathsf{T}}\boldsymbol{D}\vec{\mu} \tag{28}$$

then we can rewrite

$$\boldsymbol{A}'(\boldsymbol{D}) = \sum_\lambda \frac{f_{\rm M}^2(\lambda|\vec{\theta}_\star)}{\sigma^2(\lambda)}\boldsymbol{W}(\lambda) + \boldsymbol{D} \tag{29}$$

$$\vec{B}'^{\mathsf{T}}(\vec{\theta}_\star|\vec{\mu}, \boldsymbol{D}) = \sum_\lambda \frac{f_{\rm D}(\lambda) f_{\rm M}(\lambda|\vec{\theta}_\star)}{\sigma^2(\lambda)}\vec{T}^{\mathsf{T}}(\lambda) + (\boldsymbol{D}\vec{\mu})^{\mathsf{T}} \tag{30}$$

$$g'(\vec{\theta}_\star|\vec{\mu}, \boldsymbol{D}) = -\frac{1}{2}\sum_\lambda \frac{f_{\rm D}(\lambda)}{\sigma^2(\lambda)} - \frac{1}{2}\vec{\mu}^{\mathsf{T}}\boldsymbol{D}\vec{\mu} \tag{31}$$

$$\tag{32}$$

The result is the posterior probability function for our stellar parameters $\vec{\theta}$ that has already been marginalized over the nuisance parameters with Gaussian priors.

$$\boxed{p(\vec{\theta}_\star|\vec{D}) = \sqrt{\frac{(2\pi)^N}{\det|\boldsymbol{A}'|}} \exp\left(\frac{1}{2}\vec{B}'^{\mathsf{T}}\boldsymbol{A}'^{-1}\vec{B}' + g'\right) p(\vec{\theta}_\star)} \tag{33}$$

If we increase the permissible degree of re-fluxing by increasing σ_i, this will have the effect of smearing out or widening the posterior on a given $\vec{\theta}_\star$.

2.3. Sampling a un-flux calibrated model

The previously derived formalism works well in the case where the correction to c_0 is small in each order and can be approximated by a Gaussian. In reality, the correct formulation of $k(\lambda|\vec{\theta}_{\rm N})$ to use is

$$k(\lambda|\vec{\theta}_{\rm N}) = c_0 T_0(\lambda)\left[1 + c_1{}' T_1(\lambda) + \ldots + c_N{}' T_N(\lambda)\right] \tag{34}$$

When applied to scale the model flux $f_{\rm M}$, this becomes

$$k(\lambda|\vec{\theta}_{\rm N}) f_{\rm M} = c_0 f_{\rm M} + {c_1}' T_1(c_0 f_{\rm M}) + \ldots + {c_N}' T_N(c_0 f_{\rm M}) \tag{35}$$

Now, c_0 is allowed to scale by large amounts while the higher order terms (denoted by c_n) are perturbations on top of the *scaled* function. Previously, the higher order terms were only perturbations on the original, unscaled function, and if the c_0 term was large, then this gave an incorrect scaling. The correct priors to have on these small perturbations are Gaussian centered around 0

$$p({c_n}') = \frac{1}{\sqrt{2\pi}\sigma_{c_n}} \exp\left(-\frac{c_n'^2}{2\sigma^2_{{c_n}'}}\right) \tag{36}$$

Because c_0 is a scale factor, it should be equally probable to scale up by a factor of two as it is to scale down by a factor of two. This means we should be using a log-normal prior on c_0

$$p(c_0) = \frac{1}{\sqrt{2\pi}\sigma_{c_0} c_0} \exp\left(-\frac{(\ln c_0)^2}{2\sigma^2_{c_0}}\right) \tag{37}$$

instead of a Gaussian prior.

To gain insight, we can re-express this new version of k in the original form as

$$k(\lambda|\vec{\theta}_{\rm N}) = c_0 T_0 + c_1 T_1 + \ldots + c_N T_N \tag{38}$$

where $c_1 = c_0 {c_1}' T_0$ and $c_N = c_0 {c_N}' T_0$. Thus, if there needs to be a large perturbation in c_0, then there must also be a large perturbation in c_n. The prior on c_1 is now a product of a log-normal prior on c_0 and a Gaussian prior on ${c_1}'$, which is not analytically tractable. Therefore, in the case of large perturbations, it is better to use Equation 34 and use a log-normal prior on c_0. When we use the original framework, we are saying that perturbations in c_0 are small enough that we can approximate the log-normal prior on c_0 as Gaussian and that the priors on c_n can also be correctly approximated by a Gaussian. From tests, the original framework gives acceptable accuracy (using $\sigma_{c_n} = 0.05$) if the corrections in c_0 are $\sigma_{c_0} \leq 0.35$, or the one-sigma deviation of c_0 is between $0.7 \leq c_0 \leq 1.42$. Beyond this, one should switch to this framework, for example to correct for un-flux calibrated data.

From a flux-calibration experiment where we observed spectrophotometric standards BD+28 and BD+25 multiple times, we calibrated the dispersion in (Figure?) flux-calibration function, or IRAF's "sensitivity function." However, tests show that this effect is in fact small.

This invalidates the previous section on normalization. However, for corrections of less than XX%, this approximation is better than XX% accurate.

For the flux-calibration exercise, the overall spread in c_0 across all orders should not be more than $\sigma_{c_0} \leq 2$, or the one-sigma deviation of c_0 is between $0.125 \leq c_0 \leq 8.0$.

Unfortunately, using this more correct formalism means that we can no longer analytically marginalize over c_0, we must sample in this parameter. We can, however, still analytically marginalize over the $c_n{}'$ as before.

$$\boldsymbol{A}_{ij} = \frac{c_0^2 f_{\rm M}^2}{\sigma^2} W_{ij} \tag{39}$$

$$B_i = \frac{-f_{\rm M}^2 c_0^2 + f_{\rm D} f_{\rm M} c_0}{\sigma^2} T_i \tag{40}$$

$$g = -\frac{1}{2\sigma^2}(f_{\rm M}^2 c_0^2 - 2 f_{\rm D} f_{\rm M} c_0 + f_{\rm D}^2) \tag{41}$$

and then

$$\boldsymbol{A}_{ij}^{-1} = \frac{\sigma^2}{f_{\rm M}^2 c_0^2} W_{ij}^{-1} \tag{42}$$

then, following the same integration as before, we have

$$p(\vec{\theta}_\star, c_0|\vec{D}) \propto \frac{1}{c_0^N} \sqrt{\frac{(2\pi)^N}{\det|A_{\rm old}|}} \exp\left(-\frac{f_{\rm M}^2 c_0^2}{\sigma^2} + \frac{f_{\rm M} f_{\rm D}}{\sigma^2}(T_i W_{ij}^{-1} T_j + 1) c_0 - \frac{f_{\rm D}^2}{2\sigma^2}(T_i W_{ij}^{-1} T_j + 1) \right) \tag{43}$$

and then we can put our log-normal prior on c_0. Where N is the number of higher order terms, in our case $N = 3$. This function is not able to be integrated analytically, and so two-step marginalization is not feasible.

The plot in plots/priortests shows that the log-normal prior is the correct choice. Since we care most about the inference on F, we want to know that the choice on prior of c_0 is not biasing the final outcome, while a Gaussian prior might if σ_{c_0} was large. How to quantify this error and create a dividing line? For 5% error? In peak? Do these plots need to be in the paper?

Compare distributions by doing a K-S test on the cumulative distributions?

2.4. Hierarchical modelling

Can have a c_1, etc, that are marginalized over but we sample c_0 in a hierarchical manner which informs the overall flux-level.

2.5. Sampling with emcee

Citation to DFM (Foreman-Mackey et al. 2012), GW (Goodman & Weare 2010), and CosmoHAMMER (Akeret et al. 2013).

3. Validation tests

Comparison to main sequence stars of HAT-P-9, and WASP-14, from the TRES archive (Torres et al. 2012).

Comparisons using HIRES data?

4. Processing TRES Data

We flux calibrate TRES data, but because the reduction pipeline does not produce a "sigma spectrum," we instead use the Poisson errors on each pixel. By taking the raw, still-blazed spectrum, in units of "counts," we can find the Signal-to-Noise ratio by,

$$S/N = \sqrt{\mathrm{cts}} \tag{44}$$

then the Noise-to-Signal ratio is the inverse of this

$$N/S = \frac{1}{\sqrt{\mathrm{cts}}} \tag{45}$$

then the errors on the flux-calibrated spectrum (in ergs/s/A/cm2) is simply the flux spectrum multiplied by N/S.

IC would like to graciously acknowledge Gregory Green, Doug Finkbeiner, and Daniel Eisenstein for many fruitful discussions about statistics and spectroscopy. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013).

REFERENCES

Akeret, J., Seehars, S., Amara, A., Refregier, A., & Csillaghy, A. 2013, CosmoHammer: Cosmological parameter estimation with the MCMC Hammer, 1212.1721, astrophysics Source Code Library

Astropy Collaboration et al. 2013, A&A, 558, A33

Buchhave, L. A. et al. 2012, Nature, 486, 375

Eisenstein, D. J. et al. 2006, ApJS, 167, 40, arXiv:astro-ph/0606700

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2012, ArXiv e-prints, 1202.3665

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Stoof, H. T., Gubbels, K. B., & Dickerscheid, D. B. 2009, Ultracold quantum fields (Springer)

Torres, G., Fischer, D. A., Sozzetti, A., Buchhave, L. A., Winn, J. N., Holman, M. J., & Carter, J. A. 2012, ApJ, 757, 161, 1208.1268

Valenti, J. A., & Piskunov, N. 1996, A&AS, 118, 595

This preprint was prepared with the AAS LaTeX macros v5.2.