A Brief Introduction to PyCALI

Yan-Rong Li, liyanrong@ihep.ac.cn

December 4, 2021

1 Basic Equations

By taking 5100Å continuum flux densities and Hβ fluxes as examples, the intercalibration between datasets observed by different telescopes is written (Peterson et al. 1995)

$$
\begin{aligned}
F_\lambda(5100\ \text{Å}) &= \varphi \cdot F_\lambda(5100\ \text{Å})_{\rm obs} - G, \quad (1)\\
F(\text{H}\beta) &= \varphi \cdot F(\text{H}\beta)_{\rm obs}, \quad (2)
\end{aligned}
$$

where φ is a multiplicative factor and G is an additive factor.

To develop a Bayesian framework, we reform the above equations with vectors. First we derive the posterior probability for 5100Å continuum flux densities. The m measurements of a true signal $\boldsymbol{y}$ by k different telescopes can be written as (Li et al. 2014)

$$\boldsymbol{f}_{\rm obs} = \boldsymbol{\Phi}^{-1}(\boldsymbol{y}+\boldsymbol{G}) + \boldsymbol{bn} + \boldsymbol{\epsilon}, \quad (3)$$

$$\boldsymbol{f}_{\rm cali} = \boldsymbol{\Phi}(\boldsymbol{f}_{\rm obs} + \boldsymbol{bn} + \boldsymbol{\epsilon}) - \boldsymbol{G}, \quad (4)$$

where $\boldsymbol{n}$ are reported measurement noises, $\boldsymbol{\epsilon}$ are unknown systematic errors, $\boldsymbol{b}$ is an $m\times m$ diagnoal matrix for error scale, $\boldsymbol{\Phi}$ is an $m\times m$ diagonal matrix whose diagonal elements are formed out of k scale factors φ, $\boldsymbol{G}$ is a vector for k shift factors G, $\boldsymbol{L}$ is an $m\times k$ matrix, and

$$\boldsymbol{y} = \boldsymbol{s} + \boldsymbol{E}q, \quad (5)$$

where $\boldsymbol{E}$ is a vector with all unity elements, q is the mean of $\boldsymbol{y}$, and $\boldsymbol{s}$ is a stochastic process with a probability

$$P(\boldsymbol{s}) \propto \frac{1}{\sqrt{|\boldsymbol{S}|}} \exp\left[-\frac{1}{2}\boldsymbol{s}^T\boldsymbol{S}^{-1}\boldsymbol{s}\right]. \quad (6)$$

The probabilities of $\boldsymbol{n}$ and $\boldsymbol{\epsilon}$ are

$$P(\boldsymbol{n}) \propto \frac{1}{\sqrt{|\boldsymbol{N}|}} \exp\left[-\frac{1}{2}\boldsymbol{n}^T\boldsymbol{N}^{-1}\boldsymbol{n}\right], \quad (7)$$

and

$$P(\boldsymbol{\epsilon}) \propto \frac{1}{\sqrt{|\boldsymbol{N_\epsilon}|}} \exp\left[-\frac{1}{2}\boldsymbol{\epsilon}^T\boldsymbol{N}_{\boldsymbol{\epsilon}}^{-1}\boldsymbol{\epsilon}\right], \quad (8)$$

where $\boldsymbol{S}$, $\boldsymbol{N}$, $\boldsymbol{N_\epsilon}$ are covariance matrices of $\boldsymbol{s}$, $\boldsymbol{n}$, and $\boldsymbol{\epsilon}$, respectively. Generally, $\boldsymbol{N}$ and $\boldsymbol{N_\epsilon}$ are diagonal, and $\boldsymbol{S}$ can be described by a damped random walk process. Considering that $\boldsymbol{n}$ and $\boldsymbol{\epsilon}$ are independent, we have

$$P(\boldsymbol{bn}+\boldsymbol{\epsilon}) \propto \frac{1}{\sqrt{|\boldsymbol{b^T}\boldsymbol{N}\boldsymbol{b}+\boldsymbol{N_\epsilon}|}} \exp\left[-\frac{1}{2}(\boldsymbol{bn}+\boldsymbol{\epsilon})^T(\boldsymbol{b^T}\boldsymbol{N}\boldsymbol{b}+\boldsymbol{N_\epsilon})^{-1}(\boldsymbol{bn}+\boldsymbol{\epsilon})\right]. \quad (9)$$

For m measurements by k telescopes, the above variables take a form of

$$
\boldsymbol{\Phi}(m\times m) = \begin{bmatrix}
\varphi_1 & 0 & \cdots & 0 & \cdots & 0\\
0 & \ddots & 0 & 0 & \cdots & 0\\
0 & \cdots & \varphi_1 & 0 & \cdots & 0\\
0 & \cdots & 0 & \varphi_2 & 0 & 0\\
0 & \cdots & 0 & 0 & \ddots & 0\\
0 & \cdots & 0 & \cdots & 0 & \varphi_k
\end{bmatrix}, \quad
\boldsymbol{L}(m\times k) = \begin{bmatrix}
1 & 0 & \cdots & 0\\
\vdots & 0 & \cdots & 0\\
1 & 0 & \cdots & 0\\
0 & 1 & \cdots & 0\\
\vdots & \vdots & \vdots & \vdots\\
0 & 0 & 0 & 1
\end{bmatrix}, \quad
\boldsymbol{G} = \begin{bmatrix} G_1\\ G_2\\ \vdots\\ G_k \end{bmatrix},
$$

$$
\boldsymbol{b}(m\times m) = \begin{bmatrix}
b_1 & 0 & \cdots & 0 & \cdots & 0\\
0 & \ddots & 0 & 0 & \cdots & 0\\
0 & \cdots & b_1 & 0 & \cdots & 0\\
0 & \cdots & 0 & b_2 & 0 & 0\\
0 & \cdots & 0 & 0 & \ddots & 0\\
0 & \cdots & 0 & \cdots & 0 & b_k
\end{bmatrix}, \quad
\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_k \end{bmatrix}. \quad (10)
$$

The parameter set $\boldsymbol{\Theta}$ includes the parameters for variability model, multiplicative factors $\varphi_1, \ldots, \varphi_k$, the additive factors $G_1, \ldots, G_k$, the error scale factors $b_1, \ldots, b_k$, and the systematic error factors $\epsilon_1, \ldots, \epsilon_k$. `PyCALI` sets the multiplicative factor $\varphi_1 = 1$ and the additive factor $G_1 = 0$. The likelihood of measurements $\boldsymbol{f}$ with the parameter set $\boldsymbol{\Theta}$ is (e.g., Rybicki & Press 1992)

$$
\begin{aligned}
P(\boldsymbol{f}|\boldsymbol{\Theta}) &= \int P(\boldsymbol{s})P(q)P(\boldsymbol{n})P(\boldsymbol{\epsilon})\delta[\boldsymbol{f} - \boldsymbol{\Phi}^{-1}(\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G}) - (\boldsymbol{b}\boldsymbol{n} + \boldsymbol{\epsilon})]\mathrm{d}\boldsymbol{s}\mathrm{d}\boldsymbol{n}\mathrm{d}q \\
&= \int P(\boldsymbol{s})P(q)P(\boldsymbol{n} + \boldsymbol{\epsilon})\delta[\boldsymbol{f} - \boldsymbol{\Phi}^{-1}(\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G}) - (\boldsymbol{b}\boldsymbol{n} + \boldsymbol{\epsilon})]d\boldsymbol{s}\mathrm{d}\boldsymbol{n}\mathrm{d}q \\
&= \int P(\boldsymbol{s})P(q)P[(\boldsymbol{b}\boldsymbol{n} + \boldsymbol{\epsilon}) = \boldsymbol{\Phi}^{-1}(\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G})]\mathrm{d}\boldsymbol{s}\mathrm{d}q \\
&\propto \int \frac{P(q)}{\sqrt{|\boldsymbol{S}||\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}}|}} \exp\left\{-\frac{1}{2}\boldsymbol{s}^T\boldsymbol{S}^{-1}\boldsymbol{s}\right\} \\
&\quad \times \exp\left\{-\frac{1}{2}\left[\boldsymbol{f} - \boldsymbol{\Phi}^{-1}(\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G})\right]^T (\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})^{-1}\left[\boldsymbol{f} - \boldsymbol{\Phi}^{-1}(\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G})\right]\right\}\mathrm{d}\boldsymbol{s}\mathrm{d}q. \qquad (11)
\end{aligned}
$$

The argument in the above equation can be factorized into

$$
\begin{aligned}
& \boldsymbol{s}^T\boldsymbol{S}^{-1}\boldsymbol{s} + \left[\boldsymbol{f} - \boldsymbol{\Phi}^{-1}(\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G})\right]^T (\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})^{-1}\left[\boldsymbol{f} - \boldsymbol{\Phi}^{-1}(\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G})\right] \\
= \; & \boldsymbol{s}^T\boldsymbol{S}^{-1}\boldsymbol{s} + [\boldsymbol{\Phi}\boldsymbol{f} - (\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G})]^T\left[\boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}\right]^{-1}[\boldsymbol{\Phi}\boldsymbol{f} - (\boldsymbol{y} + \boldsymbol{L}\boldsymbol{G})] \\
= \; & (\boldsymbol{s} - \hat{\boldsymbol{s}})^T\left\{\boldsymbol{S}^{-1} + \left[\boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}\right]^{-1}\right\}(\boldsymbol{s} - \hat{\boldsymbol{s}}) + [\hat{\boldsymbol{y}} - \boldsymbol{E}q]^T\boldsymbol{C}^{-1}[\hat{\boldsymbol{y}} - \boldsymbol{E}q], \qquad (12)
\end{aligned}
$$

where

$$\hat{\boldsymbol{s}} = \boldsymbol{S}\left[\boldsymbol{S} + \boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}\right]^{-1}(\hat{\boldsymbol{y}} - \boldsymbol{E}q), \qquad (13)$$

$$\hat{\boldsymbol{y}} = \boldsymbol{\Phi}\boldsymbol{f} - \boldsymbol{L}\boldsymbol{G}, \qquad (14)$$

and

$$\boldsymbol{C} = \boldsymbol{S} + \left[\boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}\right]. \qquad (15)$$

Note that

$$\left\{\boldsymbol{S}^{-1} + \left[\boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}\right]^{-1}\right\}^{-1} = \boldsymbol{S}^{-1}\boldsymbol{C}[\boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}]^{-1}. \qquad (16)$$

By eliminating the intergration over $\boldsymbol{s}$, we can now rewrite the likelihood into

$$
\begin{aligned}
P(\boldsymbol{f}|\boldsymbol{\Theta}) &\propto \int \frac{\sqrt{|\boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}|}}{\sqrt{|\boldsymbol{C}||\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}}|}} \exp\left\{-\frac{1}{2}[\hat{\boldsymbol{y}} - \boldsymbol{E}q]^T\boldsymbol{C}^{-1}[\hat{\boldsymbol{y}} - \boldsymbol{E}q]\right\}P(q)dq \\
&\propto \frac{\sqrt{|\boldsymbol{\Phi}^T(\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}})\boldsymbol{\Phi}|}}{\sqrt{|\boldsymbol{C}||\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{N}\boldsymbol{b} + \boldsymbol{N}_{\boldsymbol{\epsilon}}||\boldsymbol{E}^T\boldsymbol{C}^{-1}\boldsymbol{E}|}} \exp\left\{-\frac{1}{2}[\hat{\boldsymbol{y}} - \boldsymbol{E}\hat{q}]^T\boldsymbol{C}^{-1}[\hat{\boldsymbol{y}} - \boldsymbol{E}\hat{q}]\right\}, \qquad (17)
\end{aligned}
$$

where

$$\hat{q} = (\boldsymbol{E}^T\boldsymbol{C}^{-1}\boldsymbol{E})^{-1}\boldsymbol{E}^T\boldsymbol{C}^{-1}\hat{\boldsymbol{y}}. \qquad (18)$$

For Hβ fluxes, we can similarly derive the corresponding posterior probability.

2 Implementation

In real implementation, the first dataset is set to be the reference with $\varphi = 1$ and $G = 0$. **The input light curve of each dataset is first normalized** before being passed to intercalibration. The continuum fluxes are normalized as

$$f'_{i,j} = \frac{f_{i,j}}{C_i}, \qquad (19)$$

where $f_{i,j}$ is the j-th point of the i-th continuum dataset and C_i is the mean, calculated as

$$C_i = \frac{1}{N_i}\sum_j f_{i,j}, \qquad (20)$$

where N_j is the number of data points in the i-th continuum dataset. The emssion line fluxes are normalized as

$$f'_{i,j} = \frac{f_{i,j}}{L_i} \times \frac{L_i}{L_0}\frac{C_0}{C_i} = \frac{f_{i,j}}{L_0\frac{C_i}{C_0}} = \frac{f_{i,j}}{L'_i}, \tag{21}$$

where L_i is the mean of the i-th line dataset and

$$L'_i = L_0\frac{C_i}{C_0}. \tag{22}$$

This normalization is to enforce that the fluxes are scale with a same factor as those of continuum. The obtained posterior samples of parameters refer to normalized light curves. That is to say, one needs to mannually do some convertion to obtain the real parameter values. For scale and shift parameters,

$$\varphi_i \to \frac{C_0}{C_i}\varphi_i, \qquad G_i \to C_0 G_i, \tag{23}$$

For systematic error factor and error scale factors of continuum datasets,

$$\epsilon_i \to C_i\epsilon_i, \qquad b_i \to b_i. \tag{24}$$

For systematic error factor and error scale factors of line datasets,

$$\epsilon_i \to L'_i\epsilon_i = L_0\frac{C_i}{C_0}\epsilon_i, \qquad b_i \to b_i. \tag{25}$$

References

Li, Y.-R., Wang, J.-M., Hu, C., Du, P., & Bai, J.-M. 2014, ApJL, 786, 6

Peterson, B. M., Pogge, R. W., Wanders, I., Smith, S. M., & Romanishin, W. 1995, PASP, 107, 579

Rybicki, G. B., & Press, W. H. 1992, ApJ, 398, 169