

Reducing the model dependence in the cosmic ray composition interpretation of X_{\max} distributions

Simon Blaess, Jose Bellido*, Bruce Dawson

The University of Adelaide, School of Physical Sciences, South Australia, 5005, Australia

E-mail: jbellido@physics.adelaide.edu.au

At higher energies the uncertainty in the estimated cosmic ray mass composition, extracted from the observed distributions of the depth of shower maximum X_{\max} , is dominated by uncertainties in the hadronic interaction models. Thus, the estimated composition depends strongly on the particular model used for its interpretation. To reduce this model dependency in the interpretation of the mass composition, we have developed a novel approach which involves fitting some of the model parameters to the data. Using this approach we extracted cosmic ray mass composition information from the published Pierre Auger X_{\max} distributions. We have obtained a consistent mass composition interpretation for EPOS-LHC and QGSJetII-04. The Sibyll2.3 model provides two possible composition interpretations, one of them in agreement with EPOS-LHC and QGSJetII-04. Our fit suggests a composition dominated by protons below $10^{18.5}$ eV, while between $10^{18.5}$ eV and $10^{19.4}$ eV it becomes dominated by helium, and in the two energy bins above $10^{19.4}$ eV a growing nitrogen component is hinted. The fit also suggests that the normalisation level for $\langle X_{\max} \rangle$ is consistent with the QGSJetII-04 model prediction, but the normalization level for $\sigma(X_{\max})$ is consistent with the EPOS-LHC prediction. These results remain sensitive to the other model parameters that we keep fixed, such as the $\langle X_{\max} \rangle$ elongation rate and the $\langle X_{\max} \rangle$ separation between p and Fe.

*35th International Cosmic Ray Conference -ICRC2017-
10-20 July, 2017
Bexco, Busan, Korea*

*Speaker.

1. Introduction

The origin of the highest energy cosmic rays is still one of nature's great mysteries. Cosmic rays have been observed with energies slightly above 10^{20} eV, with origins that are not understood. In order to understand their acceleration mechanisms (at their source) and propagation in the intergalactic and galactic magnetic fields, it is important to know the distribution of the cosmic ray mass composition observed at Earth. Cosmic rays are fully ionized nuclei, so their mass composition will also indicate their electric charge, which affects their acceleration process and their propagation through the magnetic fields. Several source models explain the observed high energy cosmic rays, predicting different mass composition distributions at Earth [1, 2, 3, 4]. For this reason, one of the hot topics in astrophysics is the identification of the cosmic ray mass composition at the highest energies, and we will address this point.

A common parameter used to extract mass composition information is X_{\max} , the atmospheric depth in g/cm^2 from the top of the atmosphere where the longitudinal development of an air shower reaches a maximum (i.e. maximum number of particles or maximum energy deposit). Different cosmic ray primaries propagate through the atmosphere differently, resulting in different observed distributions of X_{\max} [5, 6]. Due to statistical variability in the interactions between cosmic rays (of a specific primary mass) and the atmosphere, a cosmic ray's primary mass cannot be determined on an event by event basis by examining X_{\max} . Instead we study the X_{\max} distribution of cosmic rays of similar energy to infer the average mass composition. The mode, width and tail of the X_{\max} distribution provide information on the average mass composition distribution [7] and on the hadronic interaction properties [8].

In order to infer the cosmic ray mass composition from the observed X_{\max} distributions, we need to compare them with expectations according to hadronic models. However, the uncertainty in the hadronic interaction models introduces a large systematic uncertainty in the interpretation of the cosmic ray composition. We have developed a novel approach to infer the mass composition from X_{\max} distributions in such a way that the interpretation is less dependent on the hadronic models. In this paper we will apply this method to interpret the published Auger X_{\max} distributions in [9] in terms of the mass composition of cosmic rays.

In our approach we have developed a parameterisation which describes the expected X_{\max} distribution for cosmic rays of some energy and mass according to a particular hadronic interaction model. This parameterisation is then fitted to X_{\max} data to extract primary mass information from each energy bin. By allowing some of the coefficients of our X_{\max} parameterisation to be included in the fit, we reduce (but not eliminate!) the dependence on hadronic interaction models of the mass composition interpretation. Using mock data sets the performance has been evaluated.

2. Parameterisation

An X_{\max} distribution of some primary energy and mass can be modelled as the convolution of a Gaussian with an exponential [6]:

$$\frac{dN}{dX_{\max}} = \frac{1}{2\lambda} \exp\left(\frac{t_0 - t}{\lambda} + \frac{\sigma^2}{2\lambda^2}\right) \text{Erfc}\left(\frac{t_0 - t + \frac{\sigma^2}{\lambda}}{\sigma\sqrt{2}}\right) \quad (2.1)$$

where t_0 , σ and λ are shape parameters.

We fit Eq. (2.1) to simulated X_{\max} distributions of a particular primary energy and mass (either proton, helium, nitrogen or iron primaries) according to either the EPOS-LHC, QGSJetII-04 or Sibyll2.3 hadronic interaction model, obtaining the values of t_0 , σ and λ for that distribution. A parameterisation (shown in Eq. (2.2)) is then fitted to each set of shape parameters as a function of primary energy (for each mass type) for a particular model. Fig. 1 displays the results for the shape parameters t_0 , σ and λ for EPOS-LHC X_{\max} data.

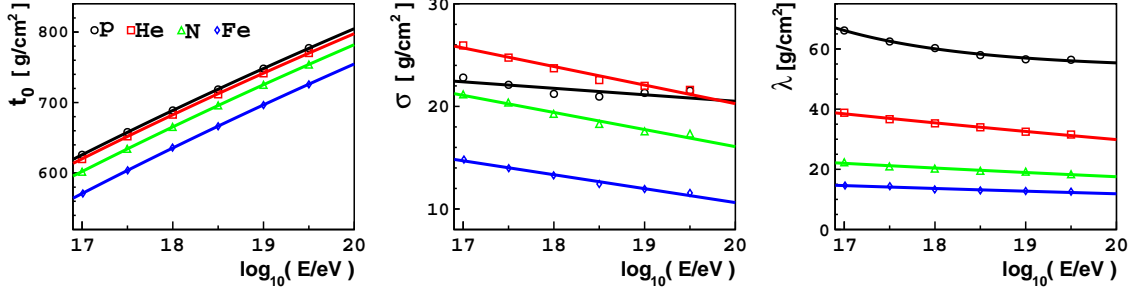


Figure 1: The t_0 , σ and λ values as a function of energy from fits of X_{\max} distributions generated with the EPOS-LHC hadronic model.

$$\begin{aligned}
 t_0(E) &= t_{0\text{norm}} + B \cdot \log_{10} \left(\frac{\log_{10} E}{\log_{10} E_0} \right), \\
 \sigma(E) &= \sigma_{\text{norm}} + C \cdot \log_{10} \left(\frac{E}{E_0} \right), \\
 \lambda(E) &= \lambda_{\text{norm}} - K + K \cdot \left(\frac{\log_{10} E}{\log_{10} E_0} \right)^{\frac{L}{\ln 10}},
 \end{aligned} \tag{2.2}$$

Combining Eq. (2.1) for each composition component (i.e. p, He, N, Fe) and accounting for its corresponding mass fraction, we produce the total X_{\max} distribution of cosmic ray events of a mixture of masses in a particular energy bin according to a hadronic interaction model:

$$\left. \frac{dN}{dX_{\max}}(E) \right|_{\text{total}} = N(E) \sum_{i=p, \text{He}, \text{N}, \text{Fe}} f_i(E) \left. \frac{dN}{dX_{\max}}(E) \right|_i \tag{2.3}$$

where $f_p(E)$, $f_{\text{He}}(E)$, $f_{\text{N}}(E)$ and $f_{\text{Fe}}(E)$ are the fractions of proton, helium, nitrogen and iron events respectively, and $N(E)$ is the total number of events.

3. Method

The parameters of our total X_{\max} parameterisation are fitted to a set of energy binned X_{\max} distributions by applying a Poisson log likelihood ratio minimisation. Each energy bin has a unique set of mass fractions fitted to that bin. However, the coefficients in Eq. (2.2) (for describing the shape parameters as a function of energy) are consistent across all energy bins. When fitting only the mass fraction parameters using our EPOS-LHC, QGSJetII-04 or Sibyll2.3 X_{\max} parameterisations, the mass composition result obtained reflects the hadronic physics assumptions of that model.

However, by also allowing some of the coefficients of our X_{\max} parameterisation to be fitted to data, the mass composition obtained has a reduced sensitivity to the hadronic interaction models.

The main difference between our EPOS-LHC, QGSJetII-04 and Sibyll2.3 X_{\max} parameterisations is the normalisation of t_0 and σ (i.e. $t_{0\text{norm}}$ and σ_{norm}), but the t_0 separation and σ ratios, among primaries, are very similar. Thus we fit $t_{0\text{norm}}$ and σ_{norm} for protons, determining the corresponding values for each other primary (i.e. He, N and Fe) by assuming that the hadronic models are correctly predicting the separation of t_0 between different primaries, and also the ratio of σ with respect to the σ for protons. Fitting $t_{0\text{norm}}$ and σ_{norm} in this way is sufficient to obtain fairly accurate mass fraction results when fitting different X_{\max} parameterisations (i.e. parameterisations based on different hadronic models) to data, provided there is a mixture of primary masses over the energy range, or there is a variation of the composition over the energy range. This condition is important, in addition to sufficient statistics, otherwise the degeneracy between the mass fractions and the two hadronic coefficients causes inconsistent mass composition results. Fitting additional hadronic coefficients is problematic as a greater spread of primaries and/or statistics is required to break the degeneracy between the fitted variables.

4. Performance

This section is about using MC simulations to evaluate the performance of the method described in the previous section. We have produced mock X_{\max} distributions using combinations of proton, He, N and Fe (for EPOS-LHC, QGSJetII-04 and Sibyll2.3 X_{\max} models) that resemble the published Auger X_{\max} distributions in [9]. When including $t_{0\text{norm}}$ and σ_{norm} in the fit of mock X_{\max} distributions, the mass fractions are reconstructed to within 15% (from the true values) in most energy bins. The higher energy bins are fitted less accurately due to the smaller statistics in these bins. The fits of $t_{0\text{norm}}$ and σ_{norm} are consistent with the true values within statistical uncertainties.

5. Results

We apply our EPOS-LHC, QGSJetII-04 and Sibyll2.3 X_{\max} parameterisation separately to X_{\max} data measured by the Pierre Auger Observatory fluorescence detector (FD) [9].

Fig. 2 displays the results from fitting the mass fractions and the coefficients $t_{0\text{norm}}$ and σ_{norm} of our EPOS-LHC, QGSJetII-04 and Sibyll2.3 X_{\max} distributions parameterisations. The top three panels display the fitted mass fractions for each model, and the bottom panel shows the p-values for these fits. The p-value is defined as the probability of obtaining a worse fit (larger likelihood ratio \mathcal{L}) than that obtained with the data. The resulting parameterisation and fractions from the fit of the X_{\max} distributions were used to generate sets of mock X_{\max} distributions to determine the p-values, and to calculate the mass composition statistical errors. Fitting $t_{0\text{norm}}$ and σ_{norm} improves the goodness of the fit of the X_{\max} distributions Fig. 2 (bottom panel). This is evident by comparing the QGSJetII-04 p-values for the $t_{0\text{norm}}$ and σ_{norm} fit to the QGSJetII-04 p-values for the fit of only the mass fractions.

We find that EPOS-LHC and QGSJetII-04 parameterisations of the X_{\max} distributions give a consistent mass composition result. The fit with the parameterisation based on Sibyll2.3 gives two minima (two solutions), where the second solution is consistent with EPOS-LHC and QGSJetII-04,

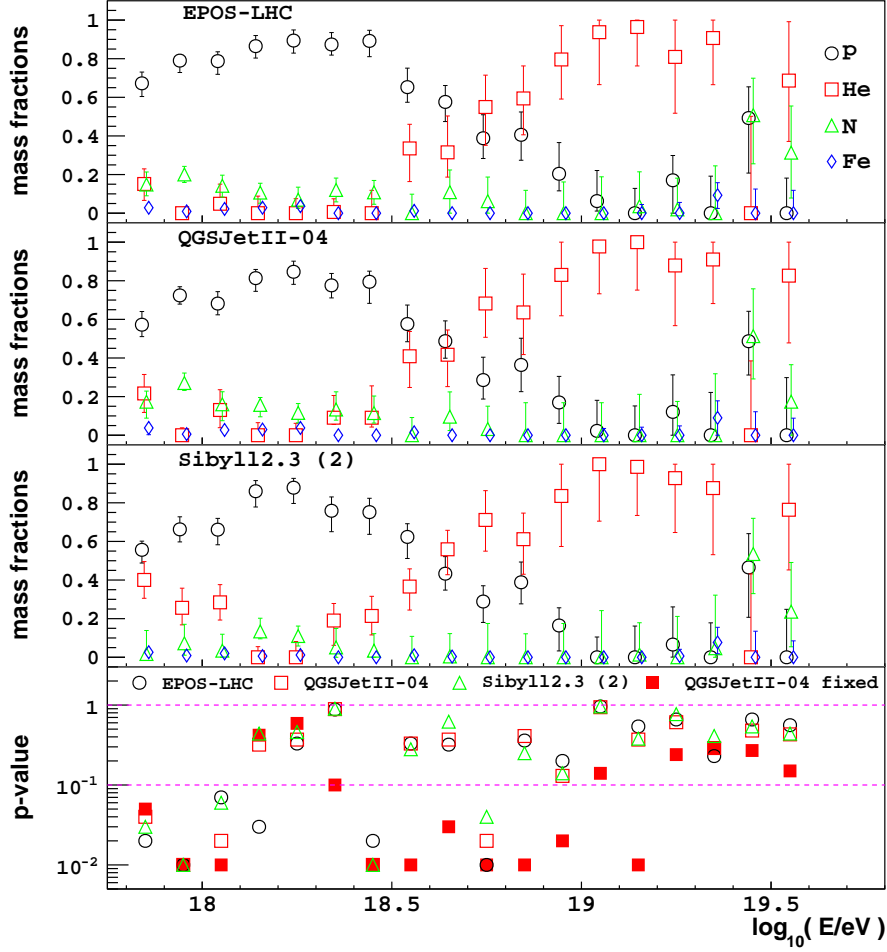


Figure 2: Fitting $t_{0,\text{norm}}$, σ_{norm} and the mass fractions of our parameterisations to FD X_{max} data measured by the Pierre Auger Observatory. The fitted mass fractions and p-values for each fitted model are shown. The red solid squares show the p-values (for QGSJetII-04 model) when fitting only the mass fraction ($t_{0,\text{norm}}$ and σ_{norm} fixed).

as shown in Fig. 4. The second minimum for Sibyll2.3 is not as deep, it can only show a contour which is 4σ away from the first minimum (Fig. 4). Fig. 2 shows the composition solutions (the Sibyll2.3 solution shown corresponds to the second minimum). Fig. 5 shows the corresponding moments of the $\ln A$ distribution, including the second Sibyll2.3 solution.

The common solutions suggest that proton is dominant around $10^{18.3}$ eV, with the mass composition then transitioning to a helium dominant composition around $10^{19.2}$ eV. A potentially growing nitrogen composition is hinted around $10^{19.5}$ eV, with more data required at the highest energies to confirm this. The transition towards heavier cosmic rays with increasing energies is consistent with the Peters cycle [10], where the maximum acceleration energy of a specie is proportional to its charge Z . A significant modification of the hadronic models would be required to accommodate a proton dominant composition at all energies above 10^{18} eV as suggested in [11].

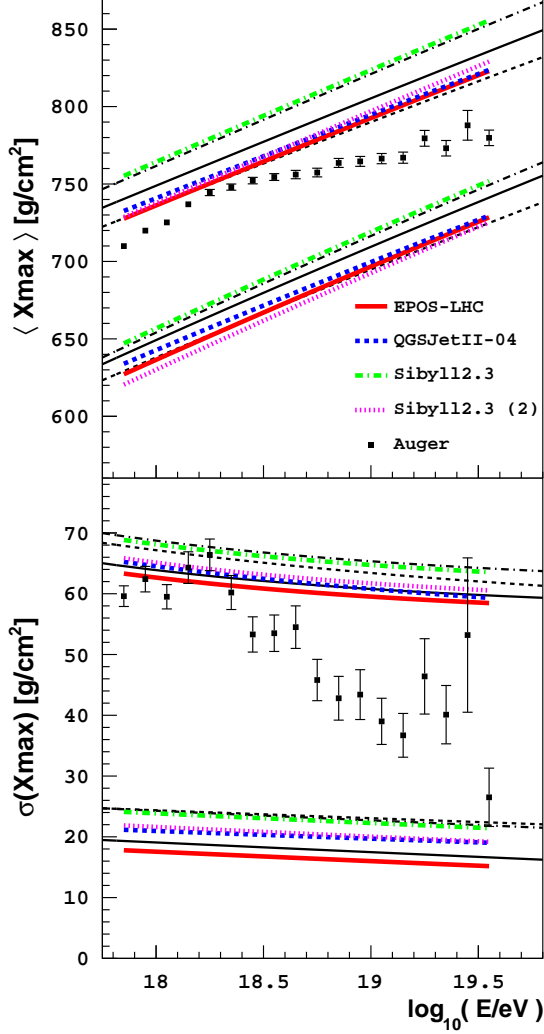


Figure 3: The red, blue and green lines show the new predictions for the $\langle X_{\max} \rangle$ and $\sigma(X_{\max})$ after fits of $t_{0\text{norm}}$, σ_{norm} and the mass fractions to FD X_{\max} distributions measured by the Pierre Auger Observatory.

The first two moments of the Auger X_{\max} distributions from [9] and their predictions (for proton and Fe) as a function of energy are shown in Fig. 3. It shows that the $t_{0\text{norm}}$ and σ_{norm} fits reduce the difference between predictions from the EPOS-LHC and QGSJetII-04 hadronic

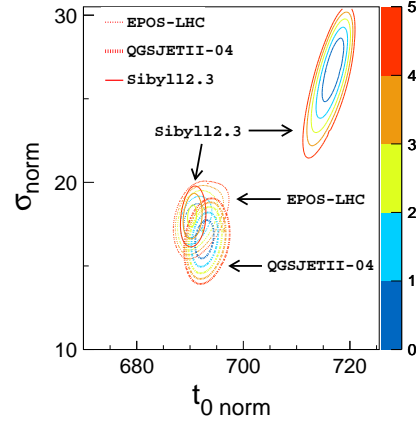


Figure 4: The $t_{0\text{norm}}$ and σ_{norm} parameter space scan over the Auger FD X_{\max} data. For specific values of $t_{0\text{norm}}$ and σ_{norm} (for each model parameterisation), the mass fractions are fitted to the data, and the first 5σ contours of the minimised Poisson log likelihood ratio are shown. Notice that Sibyll2.3 has a second minimum that overlaps with the EPOS-LHC 1σ contour.

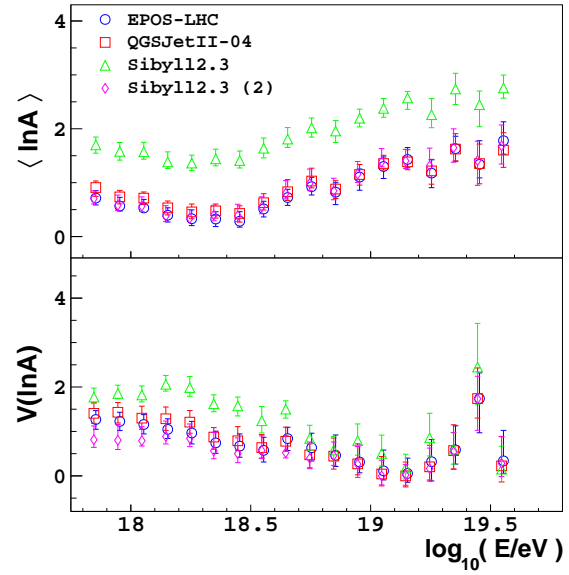


Figure 5: First two moments of the $\ln A$ distribution estimated with the fit fractions.

models. The EPOS-LHC X_{\max} distribution parameterisation fits (to the Auger data) gives a value for $t_{0,\text{norm}}$ which adjusts the EPOS-LHC $\langle X_{\max} \rangle$ values towards the QGSJetII-04 predictions, whilst the QGSJetII-04 parameterisation fit a value for σ_{norm} which adjusts the QGSJetII-04 $\sigma(X_{\max})$ predictions towards the EPOS-LHC model predictions. The values of the coefficients in Eq. (2.2) for proton, helium, nitrogen and iron primaries can be found in Table 1 for the EPOS-LHC, QGSJetII-04 and Sibyll2.3 models (assuming a normalisation energy of $E_0 = 10^{18.24}$ eV). The values fitted to the data for $t_{0,\text{norm}}$ and σ_{norm} are also shown in Table 1.

The statistical error in the estimated value of $\langle X_{\max} \rangle$ for protons or iron over the energy range is the same as the statistical error in the fitted value of $t_{0,\text{norm}}$, while for $\sigma(X_{\max})$ the statistical error is less than 0.5 g/cm^2 and 0.6 g/cm^2 respectively for both the EPOS-LHC and QGSJetII-04 fits. Our EPOS-LHC and QGSJetII-04 models have slightly different predictions for how the shape parameters change with mass and energy, but despite this there is reasonable agreement on the mass composition of the data when we allow their values for $t_{0,\text{norm}}$ and σ_{norm} to be fitted to the data.

6. Conclusions

We have presented a method to parameterise X_{\max} distributions according to different hadronic interaction models. We have shown that these parameterisations can be used to interpret the X_{\max} distribution in terms of the cosmic ray mass composition. We have evaluated the possibility of including in the mass composition fits the fit of some of the coefficients of our X_{\max} distribution parameterisations. We concluded that it is possible to include in the fit the coefficients $t_{0,\text{norm}}$ and σ_{norm} , reducing in this way the model dependency in the mass composition interpretation. These results remain sensitive to the other model parameters that we keep fixed, such as the $\langle X_{\max} \rangle$ elongation rate and the $\langle X_{\max} \rangle$ separation between p and Fe. It is important to mention that the systematics in the measured X_{\max} values are absorbed completely by the fits of $t_{0,\text{norm}}$. Thus, the composition fractions are not affected by systematics in X_{\max} .

References

- [1] A. Aab *et al.*, ”*JCAP* **1704 no. 04** (038), 2017.
- [2] D. Hooper and A. M. Taylor, *Astropart. Phys.* **33** (151), 2010.
- [3] N. Globus, D. Allard, and E. Parizot, *Phys. Rev.* **D92, no. 2** (021302), 2015.
- [4] M. Unger, G. R. Farrar, and L. A. Anchordoqui, *Phys. Rev.* **D92, no. 12** (123001), 2015.
- [5] T. K. Gaisser and A. M. Hillas, *International Cosmic Ray Conference* **8** (353), 1977.
- [6] C. J. Todero Peixoto, V. de Souza, and J. A. Bellido, *Astropart. Phys.*, **47** (18), 2013.
- [7] K.-H. Kampert and M. Unger, *Astroparticle Physics* **35** (660), 2012.
- [8] Abreu, and others, *Physical Review Letters*, **109** (062002), 2012.
- [9] A. Aab *et al.*, *Phys. Rev.* **D90, no. 12** (122005), 2014.
- [10] B. Peters, *Il Nuovo Cimento (1955-1965)*, **22, no. 4** (800), 1961.
- [11] V. S. Berezhinsky and S. I. Grigor’eva, *Astron. Astrophys.* **199**, (1), 1988.

Table 1: Coefficients for Eq. (2.2) for EPOS-LHC, QGSJetII-04 and Sibyll2.3 X_{\max} distribution predictions, assuming a normalisation energy of $E_0 = 10^{18.24}$ eV. Also in the table, we show the fitted $t_{0\text{norm}}$ and σ_{norm} for each of the three models (the Sibyll2.3 values correspond to the second minimum).

EPOS-LHC	Proton	Helium	Nitrogen	Iron
$t_{0\text{norm}}$	703	697	680	650
B	2533.29	2515.46	2548.31	2603.31
σ_{norm}	21.61	23.46	19.01	13.01
C	-0.63	-1.81	-1.67	-1.36
λ_{norm}	59.12	34.74	20.06	13.41
K	5.80	-1913.99	-1828.11	-1406.72
L	-25.93	0.063	0.035	0.027
fitted $t_{0\text{norm}}$	691 (stat.) $^{+4}_{-2}$	684	667	638
fitted σ_{norm}	18 (stat.) $^{+1}_{-1}$	19	16	11
QGSJetII-04	Proton	Helium	Nitrogen	Iron
$t_{0\text{norm}}$	688	679	660	635
B	2444.88	2410.38	2422.37	2460.32
σ_{norm}	24.82	26.83	23.07	16.54
C	-1.32	-1.24	-0.99	-0.91
λ_{norm}	61.29	37.5	25.84	17.46
K	9.35	19.32	-1818.36	-986.08
L	-17.63	-6.08	0.041	0.040
fitted $t_{0\text{norm}}$	693 (stat.) $^{+3}_{-3}$	684	665	639
fitted σ_{norm}	17 (stat.) $^{+1}_{-2}$	18	15	11
Sibyll2.3	Proton	Helium	Nitrogen	Iron
$t_{0\text{norm}}$	715	701	678	650
B	2666.31	2705.43	2695.22	2714.41
σ_{norm}	28.30	24.28	19.61	14.24
C	-1.08	-0.82	-1.20	-0.77
λ_{norm}	61.52	40.31	29.48	19.20
K	5.81	23.70	-1362.17	-1349.93
L	-27.47	-6.84	0.083	0.044
fitted $t_{0\text{norm}}$	690 (stat.) $^{+4}_{-2}$	677	654	626
fitted σ_{norm}	18 (stat.) $^{+2}_{-3}$	15	13	9