
A Novel Application of Conditional Normalizing Flows: Stellar Age Inference with Gyrochronology

Phil Van-Lane¹ Joshua S. Speagle (沈佳士)^{2 1 3 4} Stephanie Douglas⁵

Abstract

Stellar ages are critical building blocks of evolutionary models, but challenging to measure for low mass main sequence stars. An unexplored solution in this regime is the application of probabilistic machine learning methods to gyrochronology, a stellar dating technique that is uniquely well suited for these stars. While accurate *analytical* gyrochronological models have proven challenging to develop, here we apply conditional normalizing flows to photometric data from open star clusters, and demonstrate that a data-driven approach can constrain gyrochronological ages with a precision comparable to other standard techniques. We evaluate the flow results in the context of a Bayesian framework, and show that our inferred ages recover literature values well. This work demonstrates the potential of a probabilistic data-driven solution to widen the applicability of gyrochronological stellar dating.

1. Introduction

Stellar evolution and lifecycle models are critically important to astronomy on a wide range of scales, but require properly constrained age data. Since its conception (Demarque & Larson, 1964; Sandage & Eggen, 1969), isochrone fitting has been the most ubiquitous technique for dating star clusters. Isochrones are population models for stars of the same age but a variety of initial masses, so this technique performs poorly in the regime of low mass main sequence (MS) stars due to their degeneracy with age in the colour-

luminosity parameter space. This is particularly true in the case of non-coeval field stars with weakly constrained cluster membership, where the presence of multiple populations may hinder the identification of isochrones. Other stellar dating methods such as asteroseismology (Cunha et al., 2007; Kurtz, 2022), Lithium abundances (Jeffries, 2014; Beck et al., 2017; Deliyannis et al., 2019) and x-ray luminosity (Beck et al., 2017; Deliyannis et al., 2019) struggle in this regime as well. One technique that avoids such issues is gyrochronology, which relies on the concept of stellar spin-down to infer age using measurements of rotation period and MS location (ie. mass, proxied with an observable such as colour). This makes gyrochronology uniquely well-suited for dating low mass MS stars, especially in the current era of abundant stellar rotation data from missions such as *Kepler* and *Transiting Exoplanet Survey Satellite (TESS)*.

Beginning with several foundational papers by Barnes (2003; 2007; 2010), most gyrochronological studies have leveraged empirical models calibrated on stellar population ages that were benchmarked using other methodologies (Angus et al., 2019; Curtis et al., 2019; Boyle & Bouma, 2022; Kounkel et al., 2022). Open star clusters have often been used for validation in such analyses; these are gravitationally-bound groups of tens to hundreds of stars formed concurrently, and are useful gyrochronological calibrators due to their effectiveness with isochrone fitting. Other studies have relied on physical models of stellar spin-down rates to predict rotational evolution (van Saders et al., 2016; Gossage et al., 2021). The reliability challenges presented by these techniques (Angus et al., 2015; van Saders et al., 2016; Silva-Beyer et al., 2022), especially in stars older than ≈ 1 -2 Gyr, emphasize the difficulty in fitting empirical models to observations, and highlight the need for a data-driven approach. To the best of our knowledge, a gyrochronological model leveraging machine learning has not yet been implemented, and this work builds towards filling that gap by developing a probabilistic framework designed to infer stellar ages from real observational data.

¹Department of Astronomy & Astrophysics, University of Toronto, Toronto, ON, Canada ²Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada ³Data Sciences Institute, University of Toronto, Toronto, ON, Canada ⁴Dunlap Institute of Astronomy & Astrophysics, University of Toronto, Toronto, ON, Canada ⁵Department of Physics, Lafayette College, Easton, PA, United States. Correspondence to: Phil Van-Lane <phil.vanlane@mail.utoronto.ca>.

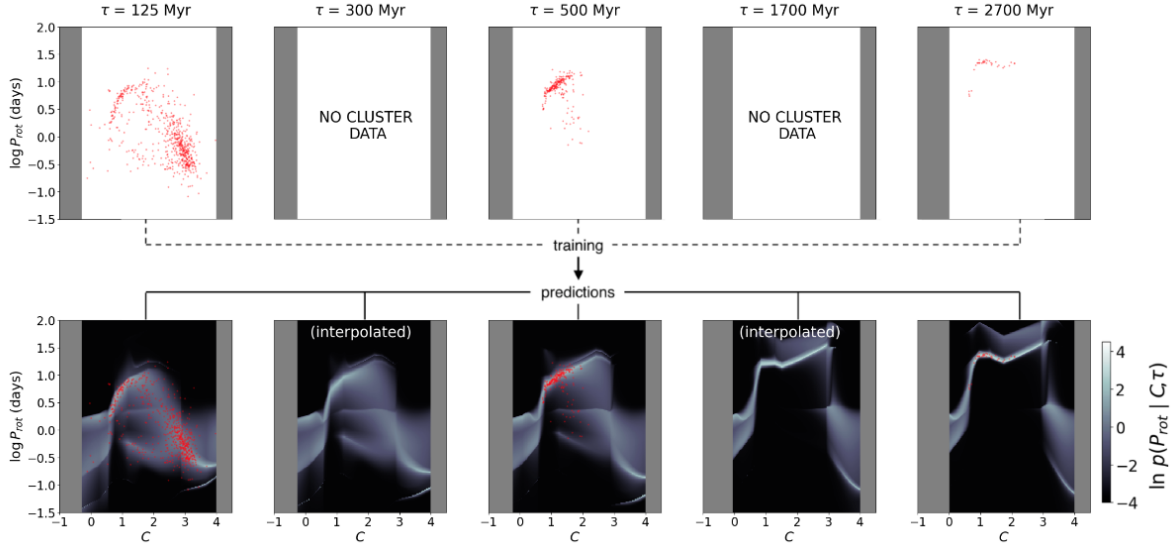


Figure 1. This figure shows subsets of our (i) observational data represented by solid red circles, and (ii) normalizing flow results represented by colour maps, both at 5 discrete sample ages. For ages aligned with our open clusters ($\tau = 125, 500,$ and $2,700$ Myr), we show results from our cross-validation exercise (see §3) in the bottom row. For ages that we do not have observational data for ($\tau = 300$ and $1,700$ Myr), the results are generated from a NF trained on all observations.

2. Methods

2.1. Data

The observational data used in this work consist of rotation period (P_{rot}) and colour (standardized to the Gaia de-reddened photometric magnitudes $(G_{BP} - G_{RP})_0$ but denoted as C going forward for brevity) for 2,878 total stars from 8 different open clusters. Our data were curated from two different sources (Curtis et al., 2020; Godoy-Rivera et al., 2021), utilizing only the subset that each considered to be “probable” cluster members. De-reddening corrections have already been applied by Curtis (2020), but were calculated by us using the colour data presented by Godoy-Rivera (2021). These papers also provide age estimates of the open clusters, which we refer to throughout this work as the “literature” ages τ_{lit} , and use as the basis of evaluation for our results (see Appendix A for a detailed breakdown).

2.2. Rotation period inference with normalizing flows

The challenge motivating this work is predicting the evolution of a star’s rotation period from its age (τ) and colour. We describe this as the conditional probability: $p(P_{rot} | C, \tau)$, which fits into our Bayesian gyrochronological framework as formulated in eq. 1. This equation is expanded from Bayes’ theorem via a decomposition of our observational likelihood term $p(P_{rot}, C | \tau)$ into the product of $p(P_{rot} | C, \tau)$ and $p(C | \tau)$, and is expressed in terms of our posterior probability for the age of a star $p(\tau | P_{rot}, C)$:

$$p(\tau | P_{rot}, C) = \frac{p(P_{rot} | C, \tau) \cdot p(C | \tau) \cdot p(\tau)}{p(P_{rot}, C)} \quad (1)$$

To calculate $p(P_{rot} | C, \tau)$, we implement a normalizing flow (NF), taking advantage of this method’s unique ability to model conditional probability densities across a continuous parameter space. We train our NF to learn the conditional $p_f(P_{rot} | C, \tau)$ distribution across our observational data (p_f now representing the probability calculated by the flow). Our NF is built using the `pyro.ai`¹ framework (Bingham et al., 2018; Phan et al., 2019), a python-based package that extends the `PyTorch` deep learning library, and we implement the Adam optimizer for our training (Kingma & Ba, 2014). We use a series of 4 layers to train our flow: in order of implementation they are (i) linear spline (Durkan et al., 2019; Dolatabadi et al., 2020), (ii) matrix exponential (Miyato et al., 2018; Salimans & Kingma, 2016; Hoogeboom et al., 2020), (iii) householder (Tomczak & Welling, 2016), and (iv) affine autoregressive (Germain et al., 2015; Jimenez Rezende & Mohamed, 2015; Kingma et al., 2016) transformation layers. These were chosen based on empirical performance during our model evaluation stage. Each training run includes 150 000 steps with a learning rate of $5e-4$. Our loss function L minimizes the negative log probability of the conditional $p(P_{rot} | C, \tau)$ distribution. We further decompose this probability into two components: the p_f computed by our NF and a background probability p_b

¹<https://github.com/pyro-ppl/pyro>

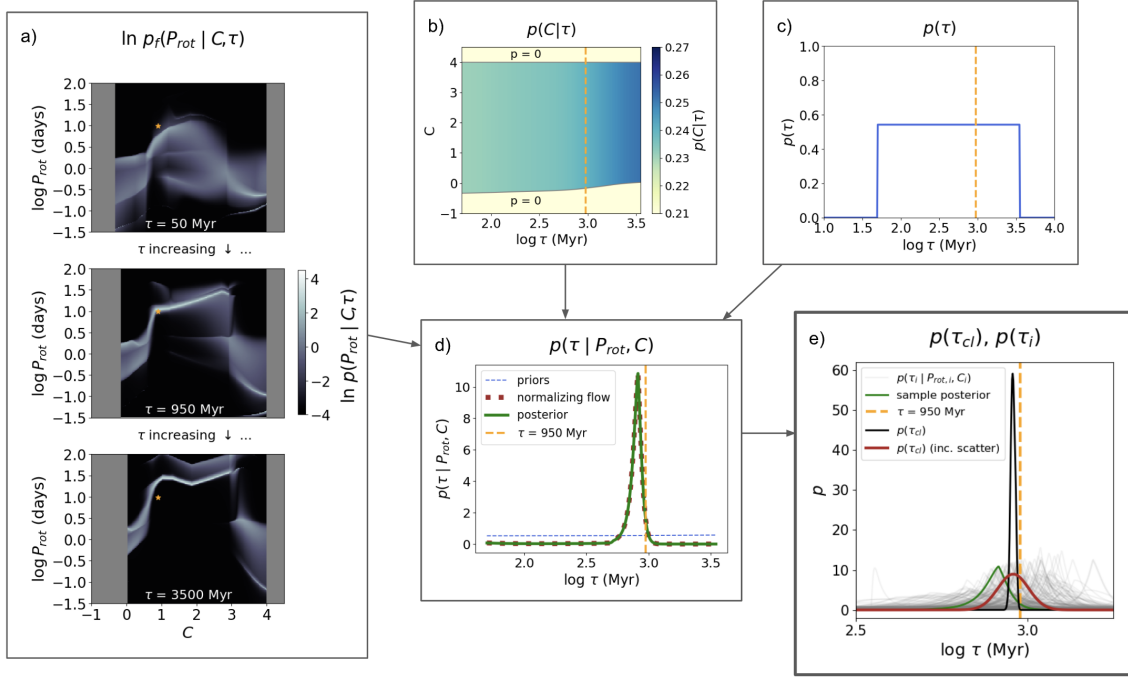


Figure 2. NF results, priors and posterior age distribution for a sample star from the NGC 6811 cluster ($P_{rot} = 9.75$ d, $C = 0.910$), and $p(\tau_{ci})$ for NGC 6811. Panel a) illustrates the NF-generated $p(P_{rot} | \tau, C)$; the star’s observed properties are indicated by the orange point. Panel b) shows the behaviour of the $p(C|\tau)$ across ages. Panel c) illustrates the uniform age prior. Panel d) demonstrates the calculation of the posterior age distribution for this sample star as described in eq. 1. Panels b), c), d), and e) all show the sample stellar age from literature as a dashed orange line. Panel e) shows $p(\tau_{ci})$ for NGC 6811 as calculated using eq. 6; the black line illustrates the distribution using only statistical uncertainties while the red line represents that posterior convolved with the intrinsic scatter that was generated with our cross-validation tests. Individual stellar posteriors are also shown in translucent grey; the green line corresponds to the sample star’s posterior.

that represents a uniform distribution across rotation period that we expect to see in non-cluster field stars:

$$p_b(P_{rot}) \sim \begin{cases} \frac{1}{P_{rot}^{max} - P_{rot}^{min}} & P_{rot}^{min} < P_{rot} < P_{rot}^{max} \\ 0 & otherwise \end{cases} \quad (2)$$

P_{rot}^{min} and P_{rot}^{max} are the limits of $\log P_{rot}$ (in days) that we evaluate our loss over; in this case -1.5 and 2.0 respectively. With this additional layer of nuance, and w_f applied as a weighting of our NF probability against our background probability, our loss function becomes (given data for each i^{th} star out of n total stars):

$$L = \sum_{i=1}^n [-\ln(w_f \cdot p_f(P_{rot,i} | C_i, \tau_i) + (1 - w_f) \cdot p_b(P_{rot,i}))] \quad (3)$$

The inclusion of p_b is important to minimize training artifacts when considering the inherent uncertainty in observational data, as there is always natural scatter and potential

for field star contamination. We apply $w_f = 0.90$ in this work, as this value is the criteria for “probable” cluster membership according to Godoy-Rivera et al. (2021). However, using a dynamic or fitted value for this weighting is a future optimization that we intend to apply.

3. Results

Figure 1 displays our NF results at a subset of discreet ages. The top row of subplots shows only observational data (where available), whereas the bottom row overlays the observations atop the NF-generated $p_f(P_{rot} | C, \tau)$ probability grid at each age value. The second and fourth panels in the bottom row are examples of NF results at ages for which we do not use any observational data in this work, and hence cannot validate directly. The behaviour of the flow at these ages is important when considering the challenge of heavily clustered data points in age space; only eight different age values are represented across all 2,878 data points, with significant distance in age space between many of the clusters. Despite these large gaps in training data, we see that the NF

interpolation does not behave erratically. Another important note is that the seemingly high probabilities in regions of sparse data at the extremes of the colour spectrum are due to the conditional nature of our probability distributions; since each subplot in the bottom row represents $p_f(P_{rot} | C, \tau)$ at a specific age slice, the probability distribution along any vertical line through a subplot will integrate to 1. Hence, these "wings" represent the $p_f(P_{rot})$ distribution that we would expect to see given stars at those ages and colours. This illustrates the importance of applying a *conditional* NF to mitigate selection effects, which otherwise could bias our NF in colour space.

To evaluate performance within the context of gyrochronological age estimation, we apply the NF results to eq. 1 to calculate posterior age distributions, and compare those to τ_{lit} from our data sources (Curtis et al., 2020; Godoy-Rivera et al., 2021). We perform this analysis using leave-one-out cross validation (LOOCV), training 8 different NFs and for each one excluding observational data from a single cluster (which is then subsequently used as the validation data set for the NF). Additionally, to apply our NF results to equation 1, we require (i) $p(\tau)$ and (ii) $p(C|\tau)$. These are calculated based on a uniform distribution of τ (in Myr) of:

$$\tau \sim U(50, 3500) \quad (4)$$

and a uniform distribution of C at any given τ of:

$$(C | \tau) \sim U(C_{iso}^{min}(\tau), C_{obs}^{max}) \quad (5)$$

We utilize isochrones generated from the `brutus`² package to constrain the lower limit on the colour values for a star at any given age: $C_{iso}^{min}(\tau)$. The upper limit C_{obs}^{max} is set to a constant of the largest (reddest) colour value present in the entire observational data set. These priors are relatively uninformative compared to stellar models available, which is an intentional choice that mitigates selection effects which may be generated by observational bias and inconsistent instrument behaviour. After computing all individual stellar age posteriors, we proceed to calculate overall cluster posteriors $p(\tau_{cl})$. A more thoughtful approach to generating these will be applied in future work, however in this initial phase we calculate $p(\tau_{cl})$ for each cluster as a naive product of the individual stellar posteriors for all n_{cl} stars in the cluster:

$$p(\tau_{cl}) = \prod_{i=1}^{n_{cl}} p(\tau_i | P_{rot,i}, C_i) \quad (6)$$

Once we have these cluster age posteriors (described with means τ_{inf} and statistical uncertainties $\sigma_{\tau_{inf}}$) inferred from

our NF results, we additionally apply Markov Chain Monte Carlo (MCMC) using the `emcee`³ package (Foreman-Mackey et al., 2013) to fit our data to:

$$\tau_{inf} \sim N(\theta \cdot \tau_{lit}, \sigma_{\tau_{inf}}^2 + s^2) \quad (7)$$

which includes an intrinsic scatter term s .

Figure 2 illustrates visually how eq. 1 is applied to generate posterior age distributions for single stars (**a** + **b** + **c** → **d**), how eq. 6 is applied to generate the cluster posteriors (**d** → **e**), and the effect of our fitted intrinsic scatter from eq. 7 on the shape of $p(\tau_{cl})$ (**e**). It is worth noting that the denominator $p(P_{rot}, C)$ of eq. 1 is constant across our model, so can be ignored given that we normalize the results. As expected, panel **d** demonstrates that since our $p(\tau)$ and $p(C | \tau)$ priors are nearly uniform (see eqs. 4 and 5), the posterior is NF-dominated. It is apparent in panel **e** that for this cluster, the posterior mean τ_{inf} is quite close to the nominal age of 950 Myr. It is also clear that including our intrinsic scatter term produces a much more reasonably constrained $p(\tau_{cl})$ than the over-precise distributions generated from only considering the statistical uncertainties (see Appendix B for individual cluster results).

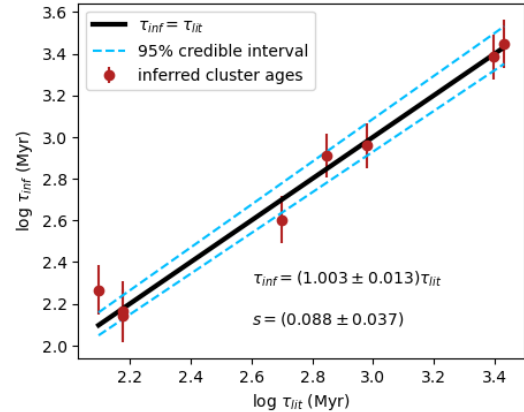


Figure 3. A summary of the cross-validation results by cluster, displaying the mean cluster posterior ages inferred by our framework (τ_{inf}) against the literature ages (τ_{lit}).

Figure 3 summarizes the accuracy of the inferred $p(\tau_{cl})$ distributions from our LOOCV exercise for each cluster relative to literature. The best fit value and 95% credible interval for θ from eq. 7 are also highlighted. The $\tau_{inf} : \tau_{lit}$ relationship is consistent with 1, indicating that the overall inference framework does not appear to suffer from systematic bias. One notable outlier here is our estimation for the Pleiades cluster (dated by Godoy-Rivera et al. (2021) as 125 Myr); this is unsurprising as (i) it is the youngest cluster

²<https://zenodo.org/record/3840241>

³<https://github.com/dfm/emcee>

we use so there is no lower bound to interpolate against in our LOOCV exercise, and (ii) the faster evolution of P_{rot} in younger stars requires more thorough calibration to enable precise inference for lower ages. More robust uncertainty handling will be addressed in future work, however it is encouraging that the recovered $p(\tau_{cl})$ distributions, particularly when intrinsic scatter is included, are reasonable.

4. Discussion

It is important to consider these results alongside the precision of other stellar age dating techniques, and within the context of historical performance of analytical and empirical gyrochronology models. The overall uncertainties on our inferred cluster ages are dominated by the intrinsic scatter, and range from 0.088 to 0.99 dex. These are consistent with isochrone fitting uncertainties, which are typically in the 10-20% range (Godoy-Rivera et al., 2021). The Gaia Collaboration et al. (2018) cite [-0.06,0.08] dex uncertainties for cluster ages greater than 100 Myr. The literature ages of NGC 6819 and Ruprecht 147 (Curtis et al., 2020) come with uncertainties of ≈ 0.03 dex based on a gyrochronological estimate of NGC 6819 calibrated to an isochrone model of Ruprecht 147. In a direct comparison of isochrone-only modeling to a combination of isochrone fitting and gyrochronology, Angus et al. (2019) reported uncertainties of 22% using only isochrones and 8% with both techniques. The results of these studies indicate that our precision is generally comparable to current stellar dating standards.

Building off of these results, and given that we still consider this work to be in the proof of concept stage, there are several features of our model that we intend to improve. First, further exploration and optimization of the NF parameters, using both empirical evaluation and more prescriptive approaches to tuning, are crucial to improving performance. Second, we need to account for observational uncertainties and selection biases; ensuring that we correctly propagate these is essential to realistically assessing the credibility of our model and systematic errors that may be introduced. In particular, a more robust method of calculating $p(\tau_{cl})$ is required to generate proper statistical uncertainties. Finally, a deeper investigation into a wider parameter space for our model could dramatically improve performance (eg. metallicity, a factor linked to stellar rotational evolution (See et al., 2023)).

There are further open questions that we will resolve through additional testing and validation. M50 and NGC 2516 (both 150 Myr) are the only coeval clusters in our observational data, so the effect of including more needs to be explored. This is particularly true in the case of young clusters, where the complexity of rotational evolution will require more extensive calibration. Furthermore, the LOOCV exercise demonstrated that our NF performed well within this POC,

however an entire cluster’s data was excluded from training during each iteration. Applying the same evaluation framework to subsets and mixtures of cluster data will provide a more nuanced assessment of how our model will behave once we incorporate new stellar observational data in which cluster membership is less clear. Finally, developing more thoughtfully motivated priors for our stellar age and conditional colour distributions will provide a more realistic indication of our best-case performance given the current model.

5. Conclusion

Our data-driven probabilistic model has successfully identified gyrochronological patterns in observational data to a precision comparable with empirical gyrochronological models and other stellar dating techniques. A particularly notable result is the successful interpolation of $p(P_{rot} | \tau, C)$, in the context where significant (age) gaps exist in the parameter space of training data. The efficacy of our framework will increase as more photometric data become available over time, but to properly leverage those observations it will be important to improve our model’s ability to handle uncertainties and observational bias. With the combination of these enhancements, additional training data, and an extensive validation pipeline, we are optimistic that this framework can become a valuable tool in the field of stellar dating.

References

- Angus, R., Aigrain, S., Foreman-Mackey, D., and McQuillan, A. Calibrating gyrochronology using Kepler asteroseismic targets. *MNRAS*, 450(2):1787–1798, June 2015. doi: 10.1093/mnras/stv423.
- Angus, R., Morton, T. D., Foreman-Mackey, D., van Saders, J., Curtis, J., Kane, S. R., Bedell, M., Kiman, R., Hogg, D. W., and Brewer, J. Toward Precise Stellar Ages: Combining Isochrone Fitting with Empirical Gyrochronology. *AJ*, 158(5):173, November 2019. doi: 10.3847/1538-3881/ab3c53.
- Barnes, S. A. On the Rotational Evolution of Solar- and Late-Type Stars, Its Magnetic Origins, and the Possibility of Stellar Gyrochronology. *ApJ*, 586(1):464–479, March 2003. doi: 10.1086/367639.
- Barnes, S. A. Ages for Illustrative Field Stars Using Gyrochronology: Viability, Limitations, and Errors. *ApJ*, 669(2):1167–1189, November 2007. doi: 10.1086/519295.
- Barnes, S. A. A Simple Nonlinear Model for the Rotation of Main-sequence Cool Stars. I. Introduction, Implications for Gyrochronology, and Color-Period Di-

- agrams. *ApJ*, 722(1):222–234, October 2010. doi: 10.1088/0004-637X/722/1/222.
- Beck, P. G., do Nascimento, J. D., J., Duarte, T., Salabert, D., Tkachenko, A., Mathis, S., Mathur, S., García, R. A., Castro, M., Pallé, P. L., Egeland, R., Montes, D., Creevey, O., Andersen, M. F., Kamath, D., and van Winckel, H. Lithium abundance and rotation of seismic solar analogues. Solar and stellar connection from Kepler and Hermes observations. *A&A*, 602:A63, June 2017. doi: 10.1051/0004-6361/201629820.
- Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P., Horsfall, P., and Goodman, N. D. Pyro: Deep Universal Probabilistic Programming. *Journal of Machine Learning Research*, 2018.
- Boyle, A. W. and Bouma, L. G. Stellar Rotation and Structure of the α Persei Complex: When Does Gyrochronology Start to Work? *arXiv e-prints*, art. arXiv:2211.09822, November 2022. doi: 10.48550/arXiv.2211.09822.
- Cunha, M. S., Aerts, C., Christensen-Dalsgaard, J., Baglin, A., Bigot, L., Brown, T. M., Catala, C., Creevey, O. L., Domiciano de Souza, A., Eggenberger, P., Garcia, P. J. V., Grundahl, F., Kervella, P., Kurtz, D. W., Mathias, P., Miglio, A., Monteiro, M. J. P. F. G., Perin, G., Pijpers, F. P., Pourbaix, D., Quirrenbach, A., Rousselet-Perraut, K., Teixeira, T. C., Thévenin, F., and Thompson, M. J. Asteroseismology and interferometry. *A&A Rev.*, 14(3-4):217–360, November 2007. doi: 10.1007/s00159-007-0007-0.
- Curtis, J. L., Agüeros, M. A., Mamajek, E. E., Wright, J. T., and Cummings, J. D. TESS Reveals that the Nearby Pisces-Eridanus Stellar Stream is only 120 Myr Old. *AJ*, 158(2):77, August 2019. doi: 10.3847/1538-3881/ab2899.
- Curtis, J. L., Agüeros, M. A., Matt, S. P., Covey, K. R., Douglas, S. T., Angus, R., Saar, S. H., Cody, A. M., Vanderburg, A., Law, N. M., Kraus, A. L., Latham, D. W., Baranec, C., Riddle, R., Ziegler, C., Lund, M. N., Torres, G., Meibom, S., Aguirre, V. S., and Wright, J. T. When Do Stalled Stars Resume Spinning Down? Advancing Gyrochronology with Ruprecht 147. *ApJ*, 904(2):140, December 2020. doi: 10.3847/1538-4357/abbf58.
- Deliyannis, C. P., Anthony-Twarog, B. J., Lee-Brown, D. B., and Twarog, B. A. Li Evolution and the Open Cluster NGC 6819: A Correlation between Li Depletion and Spindown in Dwarfs More Massive Than the F-Dwarf Li-Dip. *AJ*, 158(4):163, October 2019. doi: 10.3847/1538-3881/ab3fad.
- Demarque, P. and Larson, R. The Age of Galactic Cluster NGC 188. *ApJ*, 140:544, August 1964. doi: 10.1086/147948.
- Dolatabadi, H. M., Erfani, S., and Leckie, C. Invertible Generative Modeling using Linear Rational Splines. *arXiv e-prints*, art. arXiv:2001.05168, January 2020. doi: 10.48550/arXiv.2001.05168.
- Durkan, C., Bekasov, A., Murray, I., and Papamakarrios, G. Neural Spline Flows. *arXiv e-prints*, art. arXiv:1906.04032, June 2019. doi: 10.48550/arXiv.1906.04032.
- Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. emcee: The MCMC Hammer. *PASP*, 125(925):306, March 2013. doi: 10.1086/670067.
- Gaia Collaboration, Babusiaux, C., van Leeuwen, F., Barstow, M. A., Jordi, C., Vallenari, A., Bossini, D., Bragaglia, A., Cantat-Gaudin, T., van Leeuwen, M., Brown, A. G. A., Prusti, T., de Bruijne, J. H. J., Bailer-Jones, C. A. L., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Panem, C., Pourbaix, D., Randich, S., Sartoretti, P., Siddiqui, H. I., Soubiran, C., Walton, N. A., Arenou, F., Bastian, U., Cropper, M., Drimmel, R., Katz, D., Lattanzi, M. G., Bakker, J., Cacciari, C., Castañeda, J., Chaoul, L., Cheek, N., De Angeli, F., Fabricius, C., Guerra, R., Holl, B., Masana, E., Messineo, R., Mowlavi, N., Nienartowicz, K., Panuzzo, P., Portell, J., Riello, M., Seabroke, G. M., Tanga, P., Thévenin, F., Gracia-Abril, G., Comoretto, G., Garcia-Reinaldos, M., Teyssier, D., Altmann, M., Andrae, R., Audard, M., Bellas-Velidis, I., Benson, K., Berthier, J., Blomme, R., Burgess, P., Busso, G., Carry, B., Cellino, A., Clementini, G., Clotet, M., Creevey, O., Davidson, M., De Ridder, J., Delchambre, L., Dell’Oro, A., Ducourant, C., Fernández-Hernández, J., Fouesneau, M., Frémat, Y., Galluccio, L., García-Torres, M., González-Núñez, J., González-Vidal, J. J., Gosset, E., Guy, L. P., Halbwachs, J. L., Hambly, N. C., Harrison, D. L., Hernández, J., Hestroffer, D., Hodgkin, S. T., Hutton, A., Jasniewicz, G., Jean-Antoine-Piccolo, A., Jordan, S., Korn, A. J., Krone-Martins, A., Lanzafame, A. C., Lebzelter, T., Löffler, W., Manteiga, M., Marrese, P. M., Martín-Fleitas, J. M., Moitinho, A., Mora, A., Muinonen, K., Osinde, J., Pancino, E., Pauwels, T., Petit, J. M., Recio-Blanco, A., Richards, P. J., Rimoldini, L., Robin, A. C., Sarro, L. M., Siopis, C., Smith, M., Sozzetti, A., Süveges, M., Torra, J., van Reeven, W., Abbas, U., Abreu Aramburu, A., Accart, S., Aerts, C., Altavilla, G., Álvarez, M. A., Alvarez, R., Alves, J., Anderson, R. I., Andrei, A. H., Anglada Varela, E., Antiche, E., Antoja, T., Arcay, B., Astraatmadja, T. L., Bach, N., Baker, S. G., Balaguer-Núñez, L., Balm, P., Barache, C., Barata, C.,

- Barbato, D., Barblan, F., Barklem, P. S., Barrado, D., Barros, M., Bartholomé Muñoz, L., Bassilana, J. L., Becciani, U., Bellazzini, M., Berihuete, A., Bertone, S., Bianchi, L., Bienaymé, O., Blanco-Cuaresma, S., Boch, T., Boeche, C., Bombrun, A., Borrachero, R., Bouquillon, S., Bourda, G., Bragaglia, A., Bramante, L., Breddels, M. A., Brouillet, N., Brüsemeister, T., Brugaletta, E., Bucciarelli, B., Burlacu, A., Busonero, D., Butkevich, A. G., Buzzì, R., Caffau, E., Cancelliere, R., Cannizzaro, G., Carballo, R., Carlucci, T., Carrasco, J. M., Casamiquela, L., Castellani, M., Castro-Ginard, A., Charlot, P., Chemin, L., Chivassa, A., Cocozza, G., Costigan, G., Cowell, S., Crifo, F., Crosta, M., Crowley, C., Cuypers, J., Dafonte, C., Damerdjì, Y., Dapergolas, A., David, P., David, M., de Laverny, P., De Luise, F., De March, R., de Martino, D., de Souza, R., de Torres, A., Debosscher, J., del Pozo, E., Delbo, M., Delgado, A., Delgado, H. E., Diakite, S., Diener, C., Distefano, E., Dolding, C., Drazinos, P., Durán, J., Edvardsson, B., Enke, H., Eriksson, K., Esquej, P., Eynard Bontemps, G., Fabre, C., Fabrizio, M., Faigler, S., Falcão, A. J., Farràs Casas, M., Federici, L., Fedorets, G., Fernique, P., Figueras, F., Filippi, F., Findeisen, K., Fonti, A., Fraile, E., Fraser, M., Frézouls, B., Gai, M., Galleti, S., Garabato, D., García-Sedano, F., Garofalo, A., Garralda, N., Gavel, A., Gavras, P., Gerssen, J., Geyer, R., Giacobbe, P., Gilmore, G., Girona, S., Giuffrida, G., Glass, F., Gomes, M., Granvik, M., Gueguen, A., Guerrier, A., Guiraud, J., Gutiérrez, R., Haignon, R., Hatzidimitriou, D., Hauser, M., Haywood, M., Heiter, U., Helmi, A., Heu, J., Hilger, T., Hobbs, D., Hofmann, W., Holland, G., Huckle, H. E., Hypki, A., Icardi, V., Janßen, K., Jevardat de Fombelle, G., Jonker, P. G., Juhász, Á. L., Julbe, F., Karamelas, A., Kewley, A., Klar, J., Kochoska, A., Kohley, R., Kolenberg, K., Kontizas, M., Kontizas, E., Kopusov, S. E., Kordopatis, G., Kostrzewa-Rutkowska, Z., Koubsky, P., Lambert, S., Lanza, A. F., Lasne, Y., Lavigne, J. B., Le Fustec, Y., Le Poncin-Lafitte, C., Lebreton, Y., Leccia, S., Leclerc, N., Lecoœur-Taïbi, I., Lenhardt, H., Leroux, F., Liao, S., Licata, E., Lindstrøm, H. E. P., Lister, T. A., Livanou, E., Lobel, A., López, M., Managau, S., Mann, R. G., Mantelet, G., Marchal, O., Marchant, J. M., Marconi, M., Marinoni, S., Marschalkó, G., Marshall, D. J., Martino, M., Marton, G., Mary, N., Massari, D., Matijevič, G., Mazeh, T., McMillan, P. J., Messina, S., Michalik, D., Millar, N. R., Molina, D., Molinaro, R., Molnár, L., Montegriffo, P., Mor, R., Morbidelli, R., Morel, T., Morris, D., Mulone, A. F., Muraveva, T., Musella, I., Nelemans, G., Nicastro, L., Noval, L., O'Mullane, W., Ordénovic, C., Ordóñez-Blanco, D., Osborne, P., Pagani, C., Pagano, I., Pailler, F., Palacin, H., Palaversa, L., Panahi, A., Pawlak, M., Piersimoni, A. M., Pineau, F. X., Plachy, E., Plum, G., Poggio, E., Poujoulet, E., Prša, A., Pulone, L., Racero, E., Ragaini, S., Rambaux, N., Ramos-Lerate, M., Regibo, S., Reylé, C., Riclet, F., Ripepi, V., Riva, A., Rivard, A., Rixon, G., Roegiers, T., Roelens, M., Romero-Gómez, M., Rowell, N., Royer, F., Ruiz-Dern, L., Sadowski, G., Sagristà Sellés, T., Sahlmann, J., Salgado, J., Salguero, E., Sanna, N., Santana-Ros, T., Sarasso, M., Savietto, H., Schultheis, M., Sciacca, E., Segol, M., Segovia, J. C., Ségransan, D., Shih, I. C., Siltala, L., Silva, A. F., Smart, R. L., Smith, K. W., Solano, E., Solitro, F., Sordo, R., Soria Nieto, S., Souchay, J., Spagna, A., Spoto, F., Stampa, U., Steele, I. A., Steidelmüller, H., Stephenson, C. A., Stoev, H., Suess, F. F., Surdej, J., Szabados, L., Szegedi-Elek, E., Tapiador, D., Taris, F., Tauran, G., Taylor, M. B., Teixeira, R., Terrett, D., Teysandier, P., Thuillot, W., Titarenko, A., Torra Clotet, F., Turon, C., Ulla, A., Utrilla, E., Uzzi, S., Vaillant, M., Valentini, G., Valette, V., van Elteren, A., Van Hemelryck, E., Vaschetto, M., Vecchiato, A., Veljanoski, J., Viala, Y., Vicente, D., Vogt, S., von Essen, C., Voss, H., Votruba, V., Voutsinas, S., Walmsley, G., Weiler, M., Wertz, O., Wevers, T., Wyrzykowski, Ł., Yoldas, A., Žerjal, M., Ziaepour, H., Zorec, J., Zschocke, S., Zucker, S., Zurbach, C., and Zwitter, T. Gaia Data Release 2. Observational Hertzsprung-Russell diagrams. *A&A*, 616: A10, August 2018. doi: 10.1051/0004-6361/201832843.
- Germain, M., Gregor, K., Murray, I., and Larochelle, H. MADE: Masked Autoencoder for Distribution Estimation. *arXiv e-prints*, art. arXiv:1502.03509, February 2015. doi: 10.48550/arXiv.1502.03509.
- Godoy-Rivera, D., Pinsonneault, M. H., and Rebull, L. M. Stellar Rotation in the Gaia Era: Revised Open Clusters' Sequences. *ApJS*, 257(2):46, December 2021. doi: 10.3847/1538-4365/ac2058.
- Gossage, S., Dotter, A., Garraffo, C., Drake, J. J., Douglas, S., and Conroy, C. MESA Models with Magnetic Braking. *ApJ*, 912(1):65, May 2021. doi: 10.3847/1538-4357/abebdf.
- Hoogeboom, E., Garcia Satorras, V., Tomczak, J. M., and Welling, M. The Convolution Exponential and Generalized Sylvester Flows. *arXiv e-prints*, art. arXiv:2006.01910, June 2020. doi: 10.48550/arXiv.2006.01910.
- Jeffries, R. D. Using rotation, magnetic activity and lithium to estimate the ages of low mass stars. In Lebreton, Y., Valls-Gabaud, D., and Charbonnel, C. (eds.), *EAS Publications Series*, volume 65 of *EAS Publications Series*, pp. 289–325, November 2014. doi: 10.1051/eas/1465008.
- Jimenez Rezende, D. and Mohamed, S. Variational Inference with Normalizing Flows. *arXiv e-prints*, art. arXiv:1505.05770, May 2015. doi: 10.48550/arXiv.1505.05770.

- Kingma, D. P. and Ba, J. Adam: A Method for Stochastic Optimization. *arXiv e-prints*, art. arXiv:1412.6980, December 2014. doi: 10.48550/arXiv.1412.6980.
- Kingma, D. P., Salimans, T., Jozefowicz, R., Chen, X., Sutskever, I., and Welling, M. Improving Variational Inference with Inverse Autoregressive Flow. *arXiv e-prints*, art. arXiv:1606.04934, June 2016. doi: 10.48550/arXiv.1606.04934.
- Kounkel, M., Stassun, K. G., Bouma, L. G., Covey, K., Hillenbrand, L. A., and Lee Curtis, J. Untangling the Galaxy. IV. Empirical Constraints on Angular Momentum Evolution and Gyrochronology for Young Stars in the Field. *AJ*, 164(4):137, October 2022. doi: 10.3847/1538-3881/ac866d.
- Kurtz, D. W. Asteroseismology Across the Hertzsprung-Russell Diagram. *ARA&A*, 60:31–71, August 2022. doi: 10.1146/annurev-astro-052920-094232.
- Miyato, T., Kataoka, T., Koyama, M., and Yoshida, Y. Spectral Normalization for Generative Adversarial Networks. *arXiv e-prints*, art. arXiv:1802.05957, February 2018. doi: 10.48550/arXiv.1802.05957.
- Phan, D., Pradhan, N., and Jankowiak, M. Composable effects for flexible and accelerated probabilistic programming in numpyro. *arXiv preprint arXiv:1912.11554*, 2019.
- Salimans, T. and Kingma, D. P. Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks. *arXiv e-prints*, art. arXiv:1602.07868, February 2016. doi: 10.48550/arXiv.1602.07868.
- Sandage, A. and Eggen, O. Isochrones, Ages, Curves of Evolutionary Deviation, and the Composite C-M Diagram for Old Galactic Clusters. *ApJ*, 158:685, November 1969. doi: 10.1086/150229.
- See, V., Roquette, J., Amard, L., and Matt, S. Further evidence of the link between activity and metallicity using the flaring properties of stars in the Kepler field. *arXiv e-prints*, art. arXiv:2307.01688, July 2023. doi: 10.48550/arXiv.2307.01688.
- Silva-Beyer, J., Godoy-Rivera, D., and Chanamé, J. The breakdown of current gyrochronology as evidenced by old coeval stars. *arXiv e-prints*, art. arXiv:2210.01137, October 2022.
- Tomczak, J. M. and Welling, M. Improving Variational Auto-Encoders using Householder Flow. *arXiv e-prints*, art. arXiv:1611.09630, November 2016. doi: 10.48550/arXiv.1611.09630.
- van Saders, J. L., Ceillier, T., Metcalfe, T. S., Silva Aguirre, V., Pinsonneault, M. H., García, R. A., Mathur, S., and Davies, G. R. Weakened magnetic braking as the origin of anomalously rapid rotation in old field stars. *Nature*, 529(7585):181–184, January 2016. doi: 10.1038/nature16168.

APPENDIX

A. Cluster data details

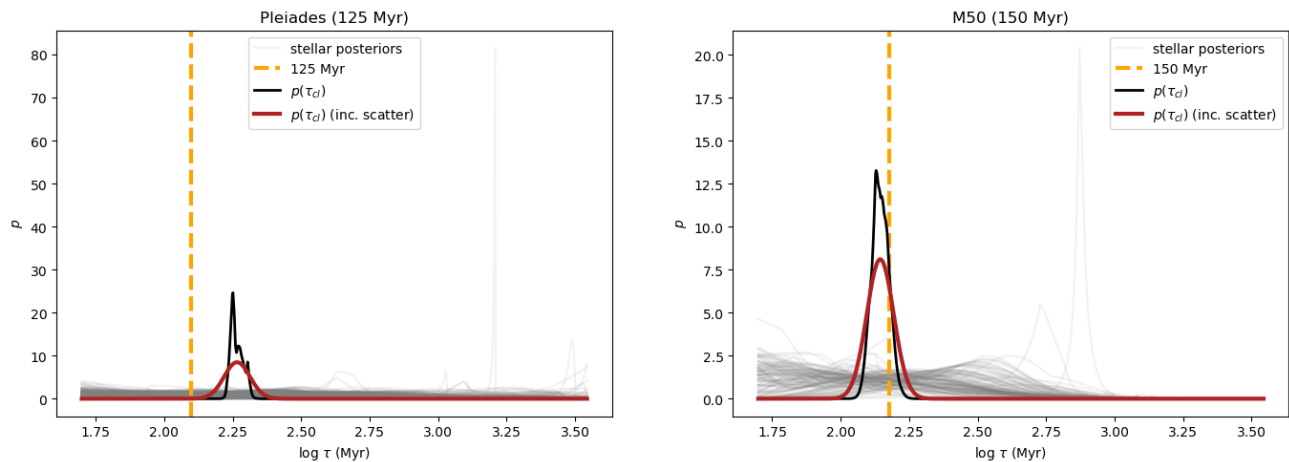
Table 1 provides the breakdown of observational data from open clusters used in this work.

Table 1. Open cluster data used. The age of each cluster is the literature age as cited per its respective source: [1] (Godoy-Rivera et al., 2021) and [2] (Curtis et al., 2020).

Cluster	Source	# Stars	Age (Myr)
Pleiades	[1]	960	125
M50	[1]	131	150
NGC 2516	[1]	316	150
M37	[1]	213	500
Praesepe	[1]	1032	700
NGC 6811	[1]	161	950
NGC 6819	[2]	30	2500
Ruprecht 147	[2]	35	2700

B. All LOOCV results

Here we present the naive $p(\tau_{cl})$ distributions from our LOOCV exercise for each cluster, both with and without intrinsic scatter, analogous to panel e) in figure 2.



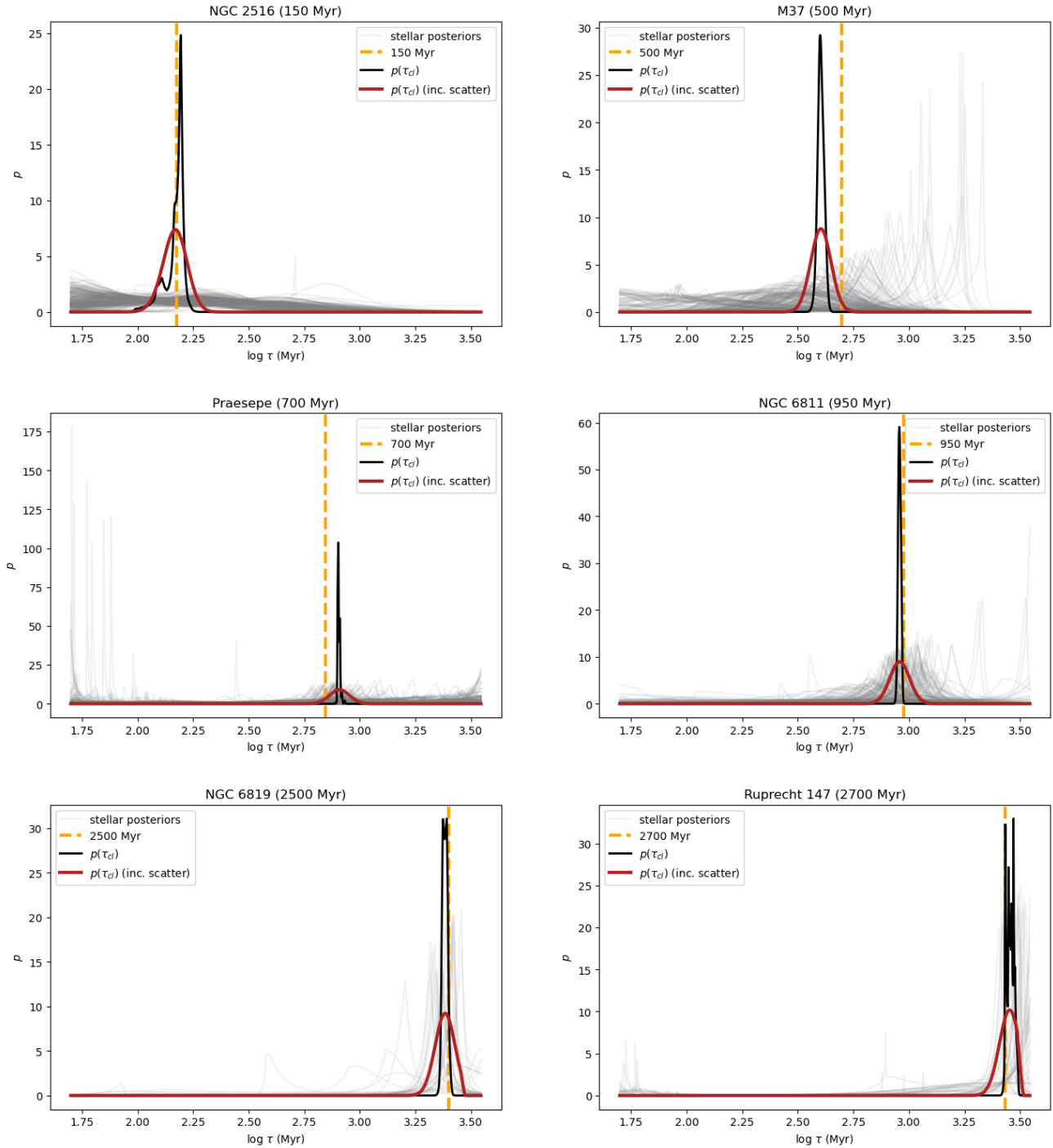


Figure 4. Plots of the naive $p(\tau_{cl})$ distributions, both with (red) and without (black) the fitted intrinsic scatter for each cluster. Also shown are all of the individual stellar posteriors (grey) and the literature cluster age (dashed orange line). Results are all from the NFs trained during our LOOCV exercise.