MNRAS **000**, 1–20 (2020) Preprint 12 June 2020 Compiled using MNRAS LaTeX style file v3.0

Asteroseismic Ages of Subgiant Stars with Deep Learning

Marc Hon[1]⋆, Earl P. Bellinger[2,1,3], Saskia Hekker[3,2], Dennis Stello[1,2,4],
and James S. Kuszlewicz[3,2]

[1] *School of Physics, The University of New South Wales, Sydney NSW 2052, Australia*
[2] *Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*
[3] *Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany*
[4] *Sydney Institute for Astronomy (SIfA), School of Physics, University of Sydney, NSW 2006, Australia*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

With the observations of an unprecedented number of oscillating subgiant stars expected from NASA's *TESS* mission, the asteroseismic characterization of subgiant stars will be a vital task for stellar population studies and for testing our theories of stellar evolution. To determine the fundamental properties of a large sample of subgiant stars efficiently, we develop a deep learning method that learns from a grid of stellar models varied in eight physical parameters to estimate distributions of fundamental parameters including age, mass, and radius using classical and asteroseismic observations as input. We apply our method to four *Kepler* subgiant stars and compare our results with previously modelled estimates. Our results show good agreement with previous estimates for three of them (KIC 11026764, KIC 10920273, KIC 11395018). With the ability to explore a vast range of stellar parameters, we identify a candidate solution for the remaining star (KIC 10005473) that is 1 Gyr younger than its previously modelled estimate and reproduces the star's observed asteroseismic and spectroscopic measurements well. We further demonstrate our method's utility for ensemble asteroseismology by characterizing a sample of 30 *Kepler* subgiant stars, where we find a majority of our age, mass, and radius estimates agree within uncertainties from more computationally expensive grid-based modelling techniques.

Key words: asteroseismology – stars: oscillations – stars: evolution – methods: data analysis

1 INTRODUCTION

Asteroseismology of solar-like oscillations is a powerful approach to measure ages of individual field stars. By probing the stellar interior, asteroseismic measurements can reveal structural changes that are indicators of stellar evolution. This is especially the case for subgiant stars that have begun to show mixed modes in their oscillation spectra. These modes arise from the coupling of acoustic waves that propagate in the stellar envelope with gravity (g-) waves that propagate near the core (Osaki 1975), and result in perturbations to the near-uniform frequency spacing of acoustic (p-) modes (*avoided crossings*, Aizenman et al. 1977). As the interiors of subgiants evolve over relatively short timescales, the mixed mode behaviour of the star's oscillation spectra also changes rapidly (e.g., Christensen-Dalsgaard et al. 1995). Hence, detailed measurements of subgiant mixed modes not only provide valuable diagnostics of the stellar interior (e.g., Deheuvels & Michel 2011; Benomar et al. 2012; Benomar et al. 2014), but also yield precise stellar age estimates (e.g., Deheuvels et al. 2014; Metcalfe et al. 2014; Li et al. 2017).

Owing to high-quality photometric observations from the *Kepler* space mission (Borucki et al. 2010), precise oscillation frequencies have been measured for subgiant stars (e.g., Appourchaux et al. 2012). Such measurements have enabled the fundamental stellar parameters of subgiants to be determined using stellar modelling techniques (e.g., Metcalfe et al. 2010; Creevey et al. 2012; Doğan et al. 2013; Stokholm et al. 2019). Although only a small number of oscillating subgiants were observed by *Kepler*, this number is expected to be amplified by NASA's Transiting Exoplanet Survey Satellite (*TESS*), where at least a few hundred oscillating subgiants are expected to be observed for a year (Campante et al. 2016; Schofield et al. 2019). There will therefore be

---

⋆ E-mail: m.hon@unsw.edu.au

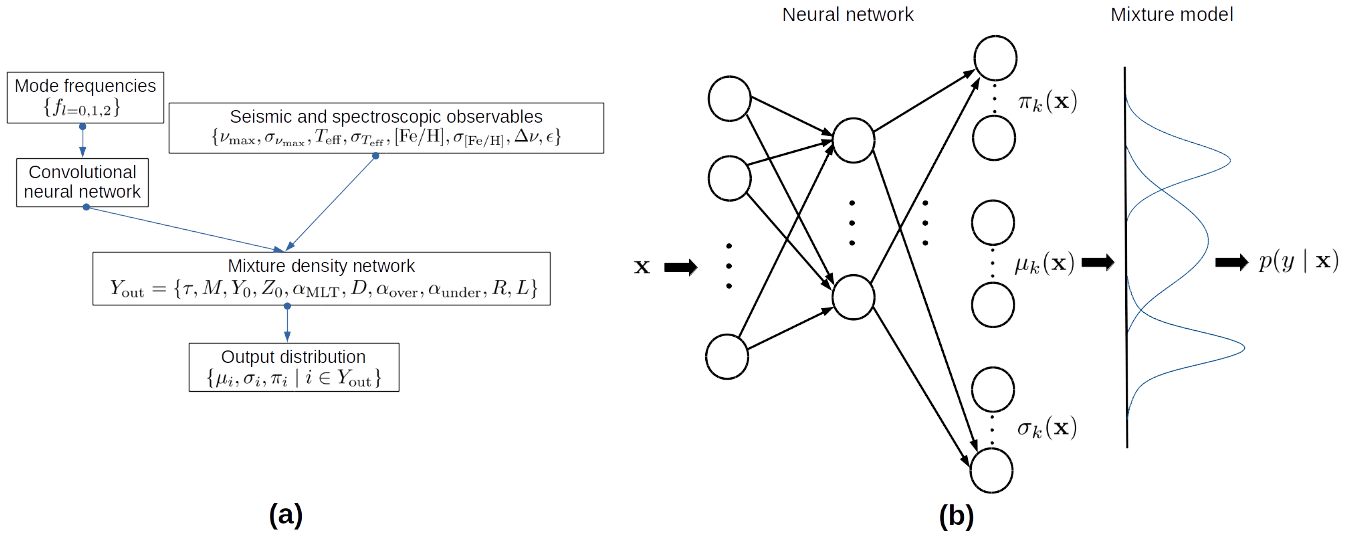

Figure 1. (a) General schematic of the deep neural network in this work. The network combines information from individual mode frequencies, global seismic parameters, and spectroscopic measurements to predict the parameters describing a 10-dimensional Gaussian mixture distribution of stellar model parameters, $Y_{\rm out}$. These parameters are the mean (μ), deviations (σ), and mixture coefficient (π) of each Gaussian in the mixture. (b) Schematic of a mixture density network. The network maps input $\mathbf{x}$ into conditional density $p(y \mid \mathbf{x})$ by predicting the shape parameters $\boldsymbol{\pi}(\mathbf{x}), \boldsymbol{\mu}(\mathbf{x}), \boldsymbol{\sigma}(\mathbf{x})$ for as many as k Gaussian functions, which are combined to form a mixture model.

further opportunities for studying subgiant stellar structure and evolution along the subgiant branch.

Stellar models are necessary for inferring stellar ages but the task of finding a model that best fits the observables from a star is computationally demanding. Such a task is a non-linear, high-dimensional optimization problem, where the complex relations governing stellar structure and evolution ($\mathcal{E}$) are sensitive to numerous input physical parameters that are being optimized ($\mathcal{P}$) such as the star's mass, initial composition, and mixing parameters. Traditional optimization methods find a best-matching set of parameters ($\mathcal{P}_*$) that best fits the observed properties of a star ($\mathcal{O}$) by solving the following:

$$\mathcal{P}_* = \arg\min_{\mathcal{P}} \left(\frac{\mathcal{E}(\mathcal{P}) - \mathcal{O})^2}{\sigma^2} \right), \tag{1}$$

where σ is the uncertainty from $\mathcal{O}$. However, as the dimensionality of $\mathcal{P}$ increases, the volume of the parameter space involved in the search increases exponentially. In an attempt to make stellar model searches tractable, traditional optimization methods typically deploy one or more of the following strategies: lowering the model grid density, grid interpolation (e.g., Rendle et al. 2019), or reducing the number of initial model parameters that are explored in the search. Lowering the grid density significantly reduces the number of models required to be generated, but comes at the cost of parameter coverage that may result in finding sub-optimal solutions. Grid interpolation methods mitigate the need for a very fine grid of models; however they still struggle with high computational complexity once additional dimensions are included in the search. A common alternative is to restrict the search to only a few free parameters and use approximations for other initial model parameters. These include the adoption of a solar-calibrated value for the mixing length parameter ($\alpha_{\rm MLT}$), or the use of the Galactic enrichment relation to estimate the initial helium abundance (Y_0) using the initial metal abundance (Z_0). These assumptions may lead to underestimated uncertainties and/or systematic errors when inferring stellar properties from models. An additional prohibiting factor in subgiant model searches is the time-consuming calculation of non-radial modes for evolved stars, which makes it expensive for search methods that require either a large grid of models or the on-the-fly calculation of stellar tracks (e.g., Metcalfe et al. 2009; Paxton et al. 2013).

Bellinger et al. (2016, hereafter BA16) showed that these problems can be mitigated for main-sequence stars by using machine learning to infer the parameters of stellar models from a given set of observables. Machine learning techniques, once trained, are able to statistically capture the complex relations connecting observations to stellar models at a fraction of the computational cost required for model grid searches. In other terms, machine learning algorithms can learn to approximate the inverse relation $\mathcal{E}^{-1}$ between model parameters and observed data. As a result, such algorithms output maximum likelihood estimates for $\mathcal{P}_*$ by computing $\mathcal{E}^{-1}(\mathcal{O})$. These algorithms have been shown by BA16 to be effective in the systematic age determination of all main-sequence stars within the high-quality *Kepler* LEGACY sample with an age precision closely comparable to those inferred from traditional grid-based optimization methods (Angelou et al. 2017; Bellinger et al. 2019a).

In this work, we seek to extend machine learning-based stellar model inference towards subgiant stars using deep learning. A major difference between our work with the BA16 study is the type of asteroseismic stellar age proxy used. The observed oscillation frequency ratios $r_{0,2}$, which are known to be sensitive towards core hydrogen abundance (Roxburgh & Vorontsov 2003), are typically used as a stellar age proxy for main-sequence stars (e.g., Christensen-Dalsgaard 1984; White et al. 2011; Bellinger & Christensen-Dalsgaard 2019). These ratios, however, are no longer effective age proxies for core hydrogen-depleted subgiant stars. Instead, observations of rapidly evolving mixed modes can be used to precisely constrain subgiant stellar ages. The mixed mode frequency pattern can be analytically described by fitting individual avoided crossings (e.g., Deheuvels & Michel 2009), however such an approach can be challenging to compute systematically across a large grid of models that contain both less-evolved and highly-evolved subgiant stars. Alternatively, the asymptotic relation of mixed modes (Shibahashi 1979) can be fit to the mixed mode pattern; however this approach works best for sufficiently evolved subgiants whose coupled g-modes are within the asymptotic regime. Another useful approach to parameterizing mixed modes is with an asteroseismic p–g diagram, which shows avoided crossing frequencies versus the p-mode large separation as a method to paramaterize subgiant evolution (Bedding 2014). While useful for a preliminary comparison with theoretical models, extracting precise age estimates with this method would still require detailed modelling of the avoided crossing frequencies. In our work, we introduce a novel machine learning-based method that learns frequency patterns from the èchelle diagram (Grec et al. 1983) and therefore does not require such patterns to be explicitly parameterized. As a result, our method can estimate the ages of oscillating stars from early post-core hydrogen exhaustion up to the base of the red-giant branch.

While machine learning has previously been applied for asteroseismic modelling (e.g., Verma et al. 2016; Bellinger et al. 2016; Hendriks & Aerts 2019), another novelty in our approach is the estimation of parameters in the form of distributions, rather than point estimates. Our method estimates a distribution across an 8D parameter space with relatively small computational cost. Besides five basic input model parameters, namely age (τ), mass (M), initial fractional helium abundance (Y_0), initial fractional metal abundance (Z_0) and mixing length parameter (α_{MLT}), we include additional physics in the form of convective core overshooting, envelope undershooting, and heavy element diffusion. These processes have their respective free parameters in the form of the overshooting parameter (α_{over}), undershooting parameter (α_{under}), and diffusion multiplication factor (D). By including these parameters, we not only explore a greater number of solutions within the grid, but we also open up the possibility of empirically estimating relations between model parameters such as the $M - \alpha_{\mathrm{over}}$ relation (Angelou et al. 2020) or the α_{MLT}-[Fe/H] relation (Viani et al. 2018).

Our work in this study constitutes a highly efficient method for subgiant star fundamental parameter estimation, which enables the characterization of subgiant ensembles and can support the detailed 'boutique' modelling of individual subgiant stars by providing informative estimates. First, we detail the construction of our deep learning algorithm and a novel sampling-based training procedure to increase the network's robustness towards measurement uncertainties and known systematics in stellar models. We then report the performance of our method on a hold-out set of subgiant stellar models and estimate the properties of real subgiant stars, which includes those modelled individually as well as those modelled as part of an ensemble.

Symbol	Name	Min	Max
$M/\mathrm{M}_\odot$	Mass	0.7	1.8
Y_0	Fractional helium abundance	0.22	0.34
α_{MLT}	Mixing length parameter	1	3
Z_0	Fractional metal abundance	0.0001	0.04
α_{over}	Overshooting parameter	0.0001	1
α_{under}	Undershooting parameter	0.0001	1
D	Diffusion multiplication factor	0.0001	3

Table 1. Ranges of initial parameters in the computed grid of stellar evolution models. The latter four parameters are varied logarithmically, and the latter three values are set to 0 if their value would otherwise be less than 0.001.

2 METHOD

We develop a deep neural network that predicts τ, M, Y_0, Z_0, α_{MLT}, D, α_{over}, and α_{under} of oscillating subgiant stars. We additionally estimate stellar radius (R) and luminosity (L), thus increasing the dimensionality of the network's output to ten. The network takes as input individual mode frequencies, the global seismic parameter ν_{max}, and spectroscopic observables (T_{eff}, [Fe/H]). We train the network with supervised learning on a grid of models that we describe in Section 2.1. In Section 2.2, we detail the deep neural network's structure and training procedure.

2.1 Models for Training

We use *Models for Experiments in Stellar Astrophysics* (MESA r12778, Paxton et al. 2011, 2013, 2015, 2018, 2019) to compute a dense grid of stellar models. The calculations begin at the pre-main-sequence evolutionary phase and span until the base of the red-giant branch. The set of input physics of the evolution is the same as described in BA16 and Bellinger et al. (2019a). The input parameters of each track ($M, Y_0, Z_0, \alpha_{\mathrm{MLT}}, \alpha_{\mathrm{under}}, \alpha_{\mathrm{over}}$, and D) are varied quasi-randomly (see Appendix B of BA16) in the ranges listed in Table 1.

As in Bellinger et al. (2019a), we define three evolutionary stages of interest: the main-sequence (MS), the MS turn-off (TO), and subgiant branch (SG). We define the beginning of the MS as when at least 99.99% of the stellar luminosity is generated by hydrogen fusion. We define the beginning of the TO (and end of MS) as the point when the central hydrogen abundance (X_c) drops below 10^{-1}. We define the beginning of the SG branch (and end of TO) as the point when X_c drops below 10^{-6}. Finally, we define the end of the SG branch as when the asymptotic period spacing drops below 150 seconds or when $\mathrm{d}\log L/\mathrm{d}\log T_{\mathrm{eff}} > -3$, whichever happens first. Alternatively, any phase can end if a maximum age of 15 Gyr is reached, after which no subsequent phases are computed.

From each of these phases, we retain 32 models which we select to be nearly equally spaced (see Appendix A of BA16) either in X_c (in the case of MS models) or in age (for TO and SG models). We use GYRE (Townsend & Teitler 2013; Townsend et al. 2018) to compute the radial (spherical degree $\ell = 0$) and non-radial ($1 \leq \ell \leq 3$) linear adiabatic mode frequencies and inertias of these models. In total, the grid contains 660,736 stellar models. For our training set in

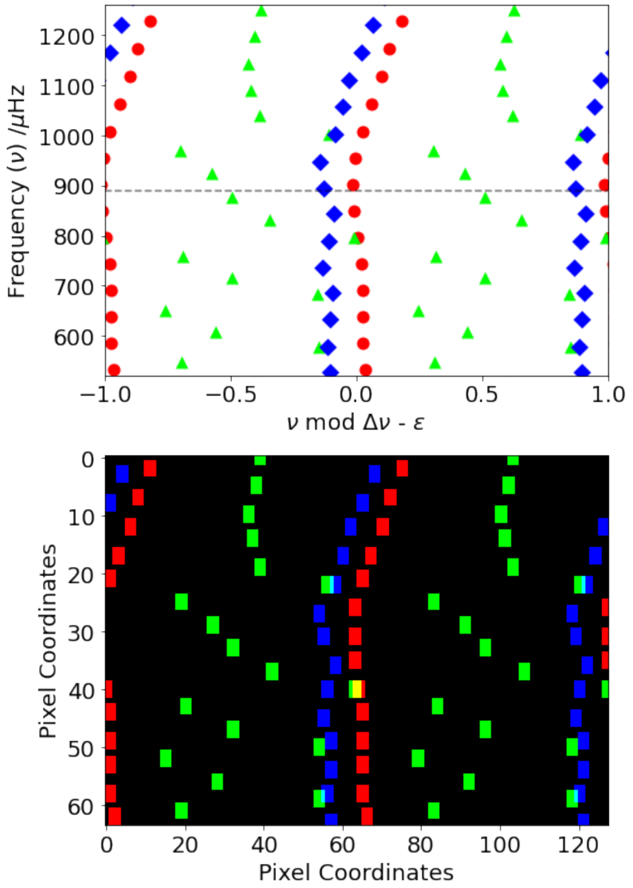

Figure 2. (Top) A repeated échelle diagram of a subgiant model's oscillation spectrum showing $l = 1$ avoided crossings. $l = 0$ modes are represented as red circles, $l = 1$ modes as green triangles, and $l = 2$ modes as blue diamonds. The diagram's vertical axis has a range of $\pm 7\Delta\nu$ around $\nu_{\rm max}$ (dashed line). Additionally, the oscillation modes are positioned such that the $l = 0$ ridge aligns with ϵ calculated from the 6 closest $l = 0$ modes to $\nu_{\rm max}$. (Bottom) The same échelle diagram binned into a 128x64 image as input for the convolutional neural network. Each $l = 0$ (red), $l = 1$ (blue), and $l = 2$ (green) mode occupies a 5x5 square within the image. The use of three separate colour channels allows overlapping modes in the image to still be visible to both viewer and network.

this study, we select models with ($X_c < 10^{-5}$), resulting in 271,631 models near the end of the TO phase up to the end of the SG branch.

2.2 Neural Network

The deep neural network, as visualized in Figure 1a, comprises two components: a convolutional neural network and a mixture density network. The detailed structure of the full network is described in Appendix A, and the code for performing estimates and training a network is made available at `https://github.com/mtyhon/deep-sub`. In the following, we describe the role of each network component.

2.2.1 Convolutional Neural Network: Analyzing Oscillation Modes

The role of the convolutional neural network in our method is to detect frequency patterns of oscillation modes by automatically learning pattern-matching filters from training data. Because we want to emphasize both the near-uniform regularity of p modes as well as the avoided crossing pattern, the oscillation modes are represented in a repeated échelle diagram that is provided as input to the network in the form of a 2D image. The advantages of such a representation are as follows:

- An échelle diagram distinctly shows the frequency pattern of oscillation modes without requiring the detailed parameterization of each mode frequency or avoided crossing.
- The network can easily adapt to missing oscillation modes that can occur for low S/N observations. Because most data-driven methods require their inputs to have a fixed size, we can only use a fixed number of modes per star/model if we use numerical frequency values as the input. This is circumvented by using an échelle diagram because the size of the 2D image of the diagram remains constant regardless of the number of modes present.
- Due to the binning of mode frequencies into a 128x64 image, the image representation is robust towards small perturbations (uncertainties) of mode frequencies. In the diagram shown in Figure 2, the position of a mode in the image will only shift vertically by one pixel if the mode is perturbed with a frequency magnitude of $7\Delta\nu/64$. Shifting a mode horizontally in the diagram would require a frequency perturbation of at least $2\Delta\nu/128$. Assuming a subgiant $\Delta\nu$ of ~ 50 μHz, a horizontal pixel shift would correspond to a frequency perturbation ~ 0.75 μHz, which is typically at the 3σ level for frequency measurements. Therefore, the binning of mode frequencies can prevent network overfitting due to noise.

To create the échelle diagram of a given model, we first estimate the frequency of maximum oscillation power, $\nu_{\rm max}$, using the following scaling relation (Brown et al. 1991; Kjeldsen & Bedding 1995):

$$\nu_{\rm max} = \frac{M/M_\odot}{(R/R_\odot)^2\sqrt{T_{\rm eff}/T_{\rm eff,\odot}}}\nu_{\rm max,\odot}, \tag{2}$$

with $\nu_{\rm max,\odot} = 3090$ μHz (Huber et al. 2011) and $T_{\rm eff,\odot} = 5772$ K (Prša et al. 2016). Using the 6 nearest $l = 0$ modes to $\nu_{\rm max}$, we calculate $\Delta\nu$ and ϵ using a weighted linear fit to the following equation:

$$\nu = \Delta\nu(n + \epsilon), \tag{3}$$

where n is the mode order and ν is the mode frequency. The fit is weighted by a Gaussian centered at $\nu_{\rm max}$ with a standard deviation of $0.1\nu_{\rm max}$.

Next, we construct a repeated échelle diagram with a range of $\pm\Delta\nu$ on the horizontal axis and $\pm 7\Delta\nu$ around $\nu_{\rm max}$ on the vertical axis. We additionally shift the abscissa of the échelle diagram by ϵ such that an $l = 0$ ridge is always positioned at the center of the diagram. Finally, we bin the diagram into a 2D array of size 128x64 as input to the network.

2.2.2 *Mixture Density Network*

After the pattern analysis of mode frequencies with the convolutional neural network, a mixture density network (MDN) combines the mode frequency information with other spectroscopic and global seismic parameters. A MDN, as introduced by Bishop (1994), models the output of a neural network as a Gaussian mixture model:

$$p(y \mid \mathbf{x}) = \sum_{k=1}^{K} \pi_k(\mathbf{x}) \frac{1}{(2\pi)^{N/2}\sigma_k(\mathbf{x})} \exp\left(-\frac{(y-\mu_k(\mathbf{x}))^2}{2\sigma_k(\mathbf{x})^2}\right), \tag{4}$$

where $\mathbf{x}$ is the network input, $p(y \mid \mathbf{x})$ is the conditional density of the output parameter vector $y = \{\tau, M, Y_0, Z_0, \alpha_{\text{MLT}}, D, \alpha_{\text{over}}, \alpha_{\text{under}}, R, L\}$, and $N = 10$ is the number of output parameters. Here, $p(y \mid \mathbf{x})$ is a sum of K Gaussian distributions that are parameterized by values $\mu_k(\mathbf{x})$ and $\sigma_k(\mathbf{x})$ and are each weighted by $\pi_k(\mathbf{x})$. For our study, we specify the network output to be described by as many as $K = 16$ distributions. The MDN output for each parameter is a vector w of length $3K$, comprising the following:

$$\begin{aligned} \mu_k(\mathbf{x}) &= w_k^{\mu}, \\ \sigma_k(\mathbf{x}) &= w_k^{\sigma}, \\ \pi_k(\mathbf{x}) &= \frac{\exp(w_k^{\pi})}{\sum_{k=1}^{K} \exp(w_k^{\pi})}, \end{aligned} \tag{5}$$

with $k \in [1, ..., K]$ and μ_k, σ_k, and π_k representing the respective mean, standard deviation, and mixing coefficient of the k−th mixture component with $\sum_{k=1}^{K} \pi_k(\mathbf{x}) = 1$. A schematic of the MDN is shown in Figure 1b. Because there are 10 parameters that are estimated by the MDN, both μ_k and σ_k in this work are 10-dimensional. Optimizing the MDN during training involves minimizing the negative log-likelihood E, given by the following:

$$E = \sum_{m=1}^{m_{\text{tot}}} -\ln p(y_m \mid \mathbf{x}_m), \tag{6}$$

where m_{tot} is the total number of models in the training set.

Fundamentally, we expect each output parameter in y to span a distribution within a grid of stellar models when given a set of subgiant star observables $\mathbf{x}$, which is why conditional density estimation with an MDN is useful. The MDN's output is effectively a region of parameter space that is expected to contain the global optimum, with uncertainties that can be estimated directly from the properties of the output parameter distribution. This is a highly efficient way of obtaining good initial guesses spanning a narrow region of parameter space for traditional grid optimization approaches. Additionally, output estimates in the form of distributions express more explicitly the presence of non-unique solutions within a grid of models, which are often the largest sources of uncertainty in subgiant star model fitting (e.g., Doğan et al. 2013). For instance, minor adjustments to the input physics of subgiant stellar models can cause them to share the same luminosity even though they have different masses, as discussed by Metcalfe et al. (2010, their Section 5.2).

2.3 Training the Network

The network is trained over 500 iterations, with early stopping if the network's performance on a hold-out validation set does not improve after more than 20 consecutive iterations. Network training only incurs a one-time cost of 2-3 hours using an NVIDIA Titan Xp GPU. Once trained, estimating the properties for a subgiant is extremely efficient, typically requiring less than one second per star.

During training, we perform bootstrapping of the input data, meaning that the values we pass to the network for each training iteration are randomly perturbed by noise or by artificially-included systematic offsets. The goal with bootstrapping is to train the network to recover the correct model values even when they have been perturbed by noise or systematic offsets. At the same time, it prevents the network from overfitting on the grid of models. The following sections describe each step in the bootstrapping procedure, with an outline in the form of pseudo-code presented in Appendix B.

2.3.1 *Surface Correction*

The improper modelling of the near-surface layers in 1D stellar models results in a systematic offset of model frequencies from the mode frequencies of real solar-like oscillators. This frequency offset is known as the *surface effect*, which varies proportionally with the inverse of mode inertia (Gough 1990). A correction term to the surface effect, $\delta\nu_{\text{surf}}$, was proposed by Ball & Gizon (2014) and is given by the following equation:

$$\delta\nu_{\text{surf}} = [c \cdot (\nu/\nu_{\text{ac}})^3 + a \cdot (\nu/\nu_{\text{ac}})^{-1}]/\mathcal{I}, \tag{7}$$

where ν is the mode frequency, ν_{ac} is the acoustic cut-off frequency, $\mathcal{I}$ is the normalized mode inertia, and both c and a are coefficients that are determined by matching the model frequencies to the observed frequencies.

When training the network, we randomly apply different levels of surface term corrections to the model frequencies. For implementation simplicity, we use only the cubic term in Equation 7 and determine for each model the range of parameter c required to obtain a $\delta\nu_{\text{surf}}$ between 0.22-0.38% of ν_{max} for the $l = 0$ mode closest to ν_{max}. This $\delta\nu_{\text{surf}}$ range is empirically estimated based on frequency offsets reported by Ball & Gizon (2017) for subgiant stars. Each stellar model in the training set thus has its own uniform range of values that c can take. In every training iteration, we randomly sample c for each model, calculate their corresponding $\delta\nu_{\text{surf}}$, and offset each model's oscillation frequencies to simulate the frequencies from a real star. Because c for each model is randomly sampled in every training iteration, different levels of surface term offsets are consistently simulated during training. By covering the range of variations expected for $\delta\nu_{\text{surf}}$, we aim to increase the network's robustness towards the surface effect.

2.3.2 *Frequency Perturbation*

Besides an artificial correction to the surface term, the input model frequencies are perturbed with random noise during training. The $l = 0$ modes of each stellar model are perturbed by Gaussian noise with a standard deviation of $\sigma_{l=0}$. The

Table 2. Summary of perturbation magnitudes of network inputs.

Input	Perturbation magnitude
$\nu_{\rm max}$	0.5 – 2.5%
$T_{\rm eff}$	50 – 150 K
[Fe/H]	0.05 – 0.15 dex
$\sigma_{l=0}$	0.1 – 1 μHz
$\sigma_{l=1}$	(0.5 – 1) $\sigma_{l=0}$
$\sigma_{l=2}$	(1 – 2) $\sigma_{l=0}$

value of $\sigma_{l=0}$ is uniformly sampled from a range of 0.1-1μHz in each training iteration. l = 1 and l = 2 modes for each stellar model are also perturbed with noise, but with $\sigma_{l=1}$ = (0.5 − 1) $\sigma_{l=0}$ and $\sigma_{l=2}$ = (1 − 2) $\sigma_{l=0}$, which are estimated from the relative uncertainties of mode frequencies of stars in the *Kepler* LEGACY sample (Lund et al. 2017). While this implementation may overestimate the uncertainties of mixed l = 1 modes (which have smaller linewidths than non-mixed l = 1 modes), we choose to be conservative with our uncertainties.

2.3.3 Simulating Missing Modes

For lower S/N observations of subgiant stars, it is common to have individual modes missing within oscillation spectra. To train our network to be robust towards this phenomenon, we randomly remove modes from the échelle diagram in each training iteration. The number of modes retained in the échelle diagram is dependent on l: we retain l = 0 modes within a 4 − 7 $\Delta\nu$ range from $\nu_{\rm max}$, while l = 1 modes are retained in a similar but independent manner from the l = 0 modes. Meanwhile, the $\Delta\nu$ range for retained l = 2 modes are constrained to be smaller or equal to the model's l = 0 range. In addition to varying the range of oscillation modes, we apply a 5% chance for each mode to be randomly removed from the set of model frequencies.

2.3.4 Noise in Spectroscopic and Global Seismic Parameters

We additionally perturb the $\nu_{\rm max}$, $T_{\rm eff}$, and [Fe/H] values of each model with random Gaussian noise with zero mean. In each training iteration, the magnitudes of $\sigma_{\nu_{\rm max}}$, $\sigma_{T_{\rm eff}}$, and $\sigma_{\rm [Fe/H]}$ are sampled uniformly from a range of values as in Table 2.

3 RESULTS

3.1 Validation Set

To quantify how well the network can recover parameters from our grid of models, we measure its performance on a test set comprising 995 tracks from the grid that were not used for training. For each output parameter, we report the following metrics on the mean of the predicted distribution: the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the explained variance score. The explained variance score between a sample of predicted

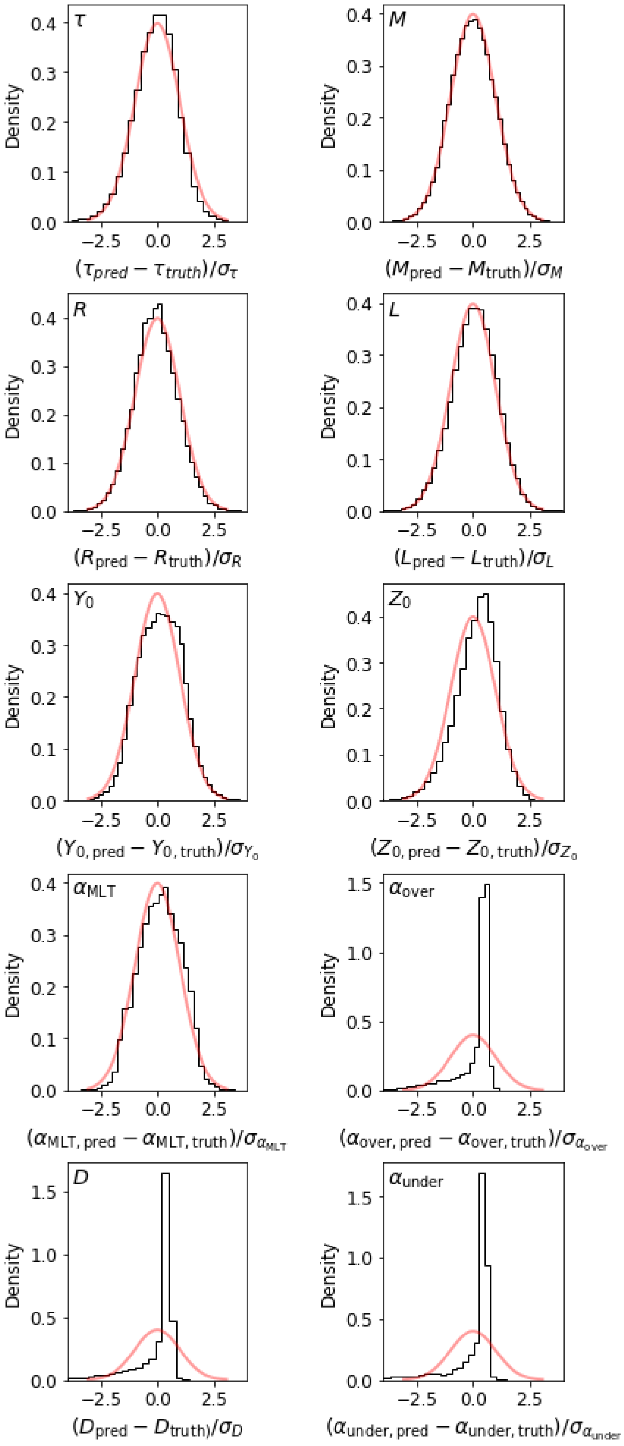

Figure 3. The z-score density for each predicted output parameter in the test set. The value of σ for each parameter is calculated as the square root of its estimated distribution's total variance. Because each distribution comprises a superposition of k Gaussian distributions with mean μ_k and deviation σ_k, the total variance is calculated by adding the expectation of σ_k^2 to Var(μ_k), i.e. the Law of Total Variance. The ideal distribution for each z-score is a normal distribution (red); this is when there are on average no systematic offsets in the output mean, and the predicted σ neither overestimates nor underestimates the true uncertainties.

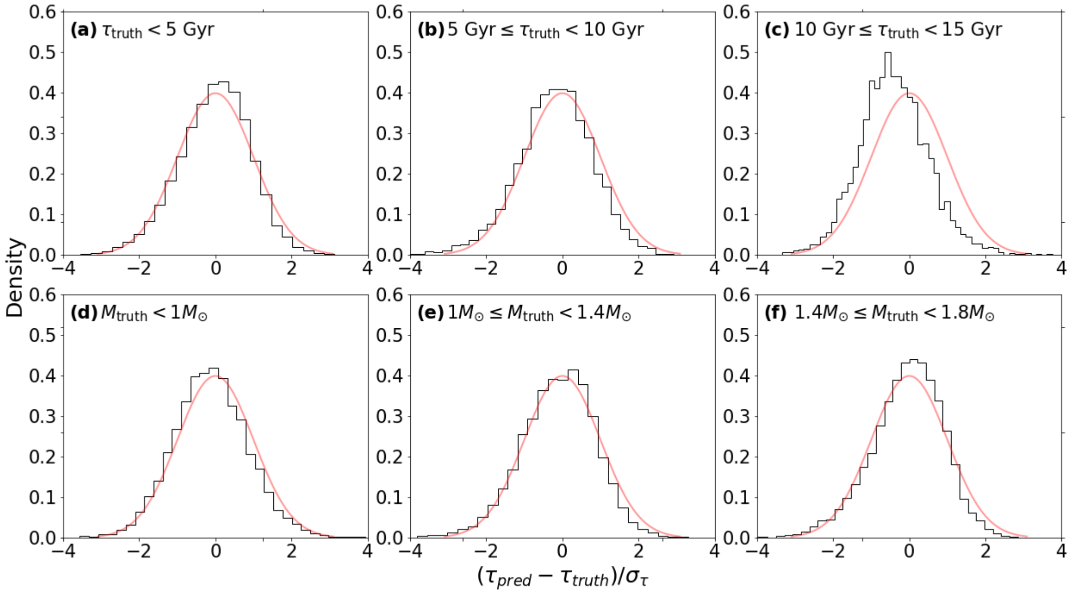

Figure 4. The z-score for stellar ages in the validation set, binned by age (top row) and mass (bottom row). Similar to Figure 3, a normal distribution (red) is plotted in each panel as reference to the ideal distribution for the z-score.

Table 3. Network performance on a hold-out validation set of stellar models. Metrics reported are the mean absolute percentage error (MAPE), mean absolute error (MAE), and the explained variance score V (Equation 8).

Output Parameter	MAPE	MAE	V
τ	7.87%	0.33 Gyr	0.97
M	2.41%	0.03 $M_\odot$	0.98
R	0.81%	0.02 $R_\odot$	0.99
L	4.34%	0.58 $L_\odot$	0.99
Y_0	5.94%	0.02	0.61
Z_0	17.15%	0.01	0.96
$\alpha_{\rm MLT}$	15.20%	0.26	0.55
D	143.35%	0.30	0.13
$\alpha_{\rm over}$	127.80%	0.09	0.10
$\alpha_{\rm under}$	150.14%	0.14	-0.10

quantities $\hat{\mathbf{y}}$ and their true values $\mathbf{y}$ is defined by the following:

$$V = 1 - \frac{\mathrm{Var}(\mathbf{y} - \hat{\mathbf{y}})}{\mathrm{Var}(\mathbf{y})}, \qquad (8)$$

with Var indicating the variance. This metric measures how well the network captures the variance of an output parameter in the test set, and ranges between negative infinity in the worst case scenario; and one for a perfect predictor. These metrics are tabulated in Table 3. From these results, we see that the network is the most accurate in its mass and radii predictions, with MAPE values of ~2% and ~1%, respectively. This is an expected result given that $\Delta\nu$, $\nu_{\rm max}$, and $T_{\rm eff}$ — all of which are parameters that can be used to infer mass and radii using the asteroseismic scaling relations (Brown et al. 1991; Kjeldsen & Bedding 1995) — are provided as inputs to the network. On average, stellar ages within the test set are well-constrained by the network, with an average error of ~8% and a V of 0.97. Y_0, and $\alpha_{\rm MLT}$, on the other hand, have moderate V scores of 0.61 and 0.55, respectively. Because the range of Z_0 varies logarithmically within the test set, the MAPE for this parameter is larger compared to other parameters. The variance for Z_0 between models, however, is explained well by the network, with a V of 0.96. We generally find that parameters D, $\alpha_{\rm over}$, and $\alpha_{\rm under}$ are poorly constrained by the data, similar to the results from BA16.

Besides performance metrics, we additionally evaluate the quality of our predicted uncertainties by visualizing each output parameter's z-score, defined as $(\mathbf{y} - \hat{\mathbf{y}})/\sigma_{\hat{\mathbf{y}}}$. Each parameter's z-score over the validation set is shown in Figure 3, where in each panel a comparison is made to a normal distribution (plotted in red). The skewness of the z-score relative to a normal distribution indicates an average systematic offset between predicted and true values in the test set. Furthermore, the increased or decreased sharpness of the z-score relative to a normal distribution indicates underestimated or overestimated uncertainties, respectively. Here, we find that the z-score for parameters M, R, L, τ, Y_0, and $\alpha_{\rm MLT}$ approximate a normal distribution well. The z-score for Z_0 is noticeably skewed from a normal distribution with a median value of 0.2. This indicates that our Z_0 estimates overestimate model values in the validation set with a typical offset of $+0.2\sigma_{Z_0}$. The poorly-constrained parameters D, $\alpha_{\rm over}$, and $\alpha_{\rm under}$ each have a z-score that sharply peaks at a small positive value. These indicate that the estimates for such parameters are highly uncertain.

To investigate our age estimates in greater detail, we show the age z-score across different ranges of model masses and ages in Figure 4. The z-score for stars with ages lower than 10 Gyr in Figures 4a-b generally follow a normal distribution with median values less than 0.1 from the origin. However, stars with 10 Gyr $\leq \tau \leq$ 15 Gyr in Figure 4c have a median z-score of -0.4, indicating that our age estimates for such stars tend to be underestimated by $-0.4\sigma_\tau$ on av-

Table 4. Summary results of the network on KIC 11026764 (Gemma). The quoted uncertainties in this work are based off the 16th and 84th percentile values from each estimated output distribution. The probability densities for each estimated parameter are shown in Appendix C.

Parameter	This work	Metcalfe et al. (2014)[1]
Age (Gyr)	$5.47^{+0.61}_{-0.63}$	5.00 ± 0.53
Mass ($M_\odot$)	$1.25^{+0.04}_{-0.05}$	1.27 ± 0.06
Radius ($R_\odot$)	2.093 ± 0.026	2.106 ± 0.025
Luminosity ($L_\odot$)	$4.09^{+0.27}_{-0.24}$	4.17 ± 0.27
Y_0	$0.252^{+0.028}_{-0.021}$	0.254 ± 0.016
Z_0	$0.020^{+0.004}_{-0.002}$	0.020 ± 0.003
α_{MLT}	$1.84^{+0.13}_{-0.11}$	2.10 ± 0.37
α_{over}	$0.011^{+0.246}_{-0.010}$	-
α_{under}	$0.013^{+0.327}_{-0.012}$	-
D	$0.029^{+0.841}_{-0.028}$	-

[1] Quoted literature uncertainties do not include systematic components of 12% for age, 3.7% for mass, and 1.3% for radius reported by authors.

erage. Because it is mainly low-mass subgiants that evolve to such old ages, we find that the age z-score of stars with $M < 1 M_\odot$ in Figure 4d are underestimated by $-0.15\sigma_\tau$ on average. The median z-score across other mass bins do not appear to have significant offsets.

3.2 Comparison with Boutique modelling

To further test our method, we apply it to four *Kepler* subgiant stars that have been modelled in detail with asteroseismology, which are KIC 11026764, KIC 10920273 , KIC 11395018, and KIC 10005473. The first three stars are colloquially known within the asteroseismic community as Gemma, Scully, and Boogie, respectively. We denote the final star by its bright star designation, HR 7322.

3.2.1 KIC 11026764 (Gemma)

KIC 11026764 is a bright (*Kp* ~9.3) G1V star that was monitored with a one minute cadence during the first 33.5 days of the *Kepler* telescope's science operations. It is therefore one of the earliest *Kepler* targets with an avoided crossing detected in its oscillation spectrum, which led to its subsequent characterization as an evolved subgiant star (Chaplin et al. 2010). Metcalfe et al. (2014) matched the oscillation frequencies and spectroscopic parameters of KIC 11026764 to a grid of stellar models using the Asteroseismic Modelling Portal (Woitaszek et al. 2009, AMP) to obtain $\tau = 5.00 \pm 0.53$ Gyr, $M = 1.27 \pm 0.06\ M_\odot$, and $R = 2.106 \pm 0.025\ R_\odot$. While this star had also been modelled by Metcalfe et al. (2010), we compare our results with the Metcalfe et al. (2014) estimates, which used a longer *Kepler* time series and updated spectroscopic parameters. In particular, they used oscillation mode frequencies provided by Appourchaux et al. (2012) and spectroscopic parameters $T_{\mathrm{eff}} = 5682 \pm 84$ K and [Fe/H] $= 0.05 \pm 0.09$ dex from Bruntt et al. (2012). We adopt these same measurements as input, and use a $\nu_{\max}$ value of $890 \pm 12\ \mu$Hz obtained from the APOKASC catalog of dwarf and subgiant stars (Serenelli et al. 2017).

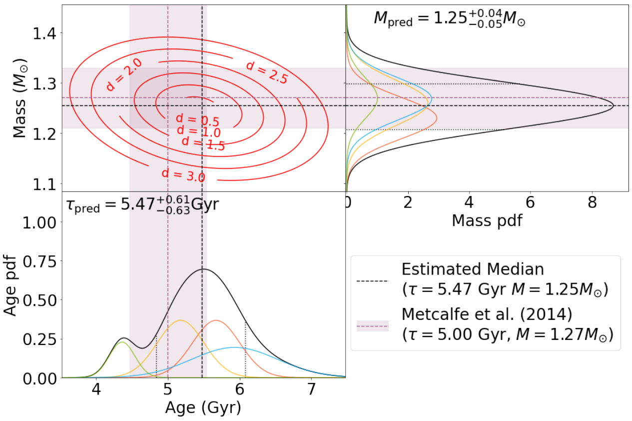

Figure 5. The estimated age and mass distribution for KIC 11026764 (Gemma). The probability density (black) in the bottom and right panels are the network's estimates for age and mass, respectively. The black dotted lines represent the 16th and 84th percentile values. Each probability density is a superposition of up to 16 Gaussians; here the 4 highest-weighted Gaussians are shown within each distribution. The dashed black lines correspond to the median of the age/mass distributions. Literature values of age and mass (including uncertainties) are shaded in purple. The center panel shows the joint age-mass distribution, where the red contours are lines of constant Mahalanobis distance.[2]

A full comparison of estimates is tabulated in Table 4, which shows a good agreement for τ, M, R, L, Y_0 and Z_0 with those from AMP (Metcalfe et al. 2014). We show our estimated age and mass distributions in Figure 5, where we find a distinct bi-modality in age, with peaks at $\tau \sim 4.4$ Gyr and $\tau \sim 5.4$ Gyr. This suggests that there are two distinct families of age solutions having model parameters within the range of our estimates. We also note that small values ($\lesssim 10^{-3}$) are generally favoured for $\alpha_{\mathrm{over}}, \alpha_{\mathrm{under}}$, and D by the network, which suggests that effects of convective over/undershooting and atomic diffusion processes are small for KIC 11026764. Equivalently, this implies that there exists a stellar model that matches the observables of KIC 11026764 well while having small values of α_{over}, α_{under}, and D.

Our trained network is a deterministic function that provides estimates of stellar properties when given a set of input observables. While it is encouraging that our results in Table 4 agree well with a full boutique modelling analysis, it does not necessarily indicate that the estimated stellar properties are self-consistent. To test our estimates' self-consistency, we identify whether a stellar model using our estimates can reproduce the observables from KIC 11026764. First, generate a new model using the median estimated initial parameters $(Y_0, Z_0, \alpha_{\mathrm{MLT}}, \alpha_{\mathrm{over}}, \alpha_{\mathrm{under}}, D)$. We then compute the root mean density ($\rho^{1/2} = \sqrt{M/R^3}$) corresponding to the median mass and radius and explore stellar models

---

[2] The Mahalanobis distance, d, is a multi-dimensional generalization of the number of standard deviations that a point $\vec{x}$ is from the mean $\vec{\mu}$ of a distribution (Mahalanobis 1936). Mathematically, it is described as $d = (\vec{x} - \vec{\mu})^T S^{-1} (\vec{x} - \vec{\mu})$, where S is the covariance matrix of the distribution. In Figure 5, d is used to visualize the range of values that $\vec{x} = (\tau, M)$ can have when sampling from the joint age-mass distribution.

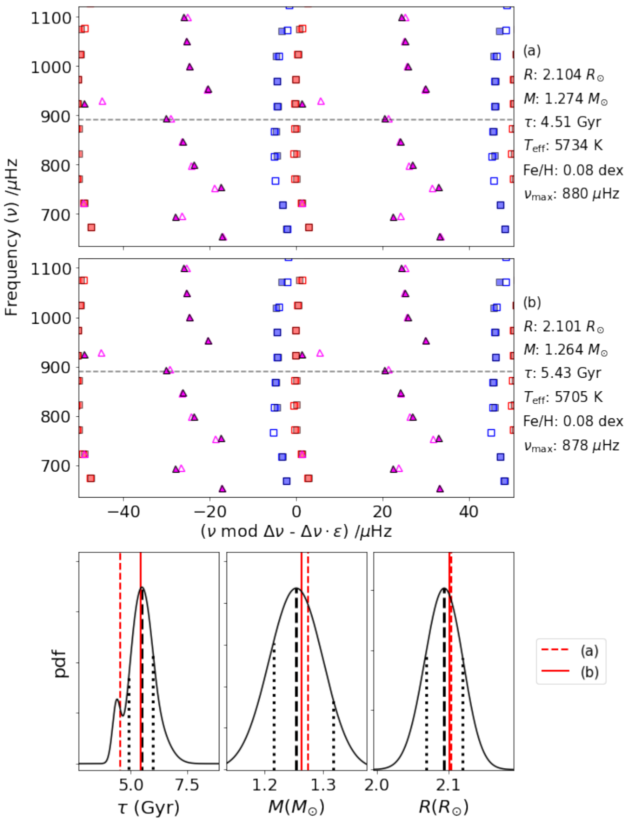

Figure 6. (Top) Échelle diagrams comparing model frequencies (open symbols) with observed frequencies (filled symbols) from KIC 11026764. Each panel corresponds to a different stellar model. Panel (a) corresponds to a model generated using the median estimated initial model parameters (Y_0, Z_0, $\alpha_{\rm MLT}$, $\alpha_{\rm over}$, $\alpha_{\rm under}$, D) from the network, with the model having similar root mean density ($\rho^{1/2} = \sqrt{M/R^3}$) to the corresponding median estimate. The model in panel (b) is another good-matching solution with similar initial parameters but $\alpha_{\rm MLT}$ smaller than the model in (a) (see text). (Bottom) A comparison of model τ, M, and R from both panels above with their corresponding estimated distributions.

Table 5. Summary results of the network on KIC 10920273 (Scully). The probability densities for each estimated parameter are shown in Appendix C.

Parameter	This work	Doğan et al. (2013)
Age (Gyr)	$6.27^{+1.89}_{-1.76}$	7.12 ± 0.47
Mass ($M_\odot$)	$1.09^{+0.05}_{-0.04}$	1.00 ± 0.04
Radius ($R_\odot$)	$1.837^{+0.025}_{-0.023}$	1.776 ± 0.021
Luminosity ($L_\odot$)	$3.51^{+0.24}_{-0.21}$	3.18 ± 0.13
Y_0	0.290 ± 0.028	0.294 ± 0.014
Z_0	$0.015^{+0.005}_{-0.002}$	0.011 ± 0.002
$\alpha_{\rm MLT}$	$1.94^{+0.14}_{-0.12}$	1.96 ± 0.09
$\alpha_{\rm over}$	$0.009^{+0.150}_{-0.008}$	-
$\alpha_{\rm under}$	$0.022^{+0.363}_{-0.021}$	-
D	$0.118^{+1.443}_{-0.117}$	-

with the same initial parameters and $\rho^{1/2}$ but with different M and R values. This exploration involves altering M and R simultaneously within $\pm 1\sigma$ from their median values in increments of 0.1σ. Each model generated has their mode frequencies corrected for the surface term offset using Equation 7. Next, the goodness of fit of each model's frequencies and spectroscopic properties ($\vec{x}_{\rm mod}$) with respect to the stellar observables ($\vec{x}_{\rm obs}$) is evaluated by computing $\chi^2 = (\vec{x}_{\rm obs} - \vec{x}_{\rm mod})^2/\sigma^2_{\rm obs}$, where $\sigma_{\rm obs}$ are observational uncertainties. The smallest-χ^2 model from this analysis reproduces avoided crossings at frequencies that are typically close to the observed avoided crossings of the star. To improve the match between model and observation further, we explore additional models within ± 0.005 $R_\odot$ from the smallest-χ^2 model and select the final best-matching model by χ^2 evaluation and by visual inspection of the match between model and observed frequencies. While this last step alters the final model's $\rho^{1/2}$ from the median value, it remains well within the estimated uncertainties for $\rho^{1/2}$. A comparison of model and observed frequencies for KIC 11026764 is shown in Figure 6a, demonstrating a good match across all modes. The model's age of τ = 4.51 Gyr is consistent with the peak observed at 4.4 Gyr in the age distribution, as shown in the bottom left panel of Figure 6.

Using similar initial parameters but with $\alpha_{\rm MLT}$ 0.25σ smaller than the median estimated value[3], we find another model showing a good match to the subgiant's observables as shown in Figure 6b. This particular model has an age of τ = 5.41 Gyr that corresponds to the main peak seen in the age distribution, and has a mass $0.01 M_\odot$ smaller than the model in Figure 6a. As discussed in Deheuvels & Michel (2011), a decrease in $\alpha_{\rm MLT}$ results in a lower mass required for a stellar model to achieve the same avoided crossing frequencies at the same $\Delta\nu$, which is consistent with our comparison of the two solutions shown for KIC 11026764.

3.2.2 *KIC 10920273 (Scully)*

KIC 10920273 is a relatively faint (Kp ~11.9) G-type subgiant that was monitored continuously throughout the *Kepler* mission. Following the asteroseismic analysis by Campante et al. (2011), KIC 10920273 was modelled in detail by Creevey et al. (2012) and Doğan et al. (2013). We use the results from the Doğan et al. (2013) study for comparison because they use individual mode frequencies to better constrain the star's characteristics. Their work combined results from five modelling teams and thus includes a variety of stellar evolutionary codes, fitting methods, and input physics used. They report $\tau = 7.12 \pm 0.47$ Gyr, $M = 1.00 \pm 0.04$ $M_\odot$, and $R = 1.776 \pm 0.021$ $R_\odot$ as their final consensus values.

For our analysis, we use the oscillation mode frequencies from the maximal frequency set extracted by Campante et al. (2011). We use spectroscopic parameters $T_{\rm eff}$ = 5790±74 K and [Fe/H]=−0.04±0.10 dex, which were adopted by Doğan et al. (2013) in their analyses. In addition, we use $\nu_{\rm max}$ = 990 ± 60 μHz as measured by Creevey et al. (2012, their Table 3). In Table 5, we show a summary of our estimates, where we find that our estimated age of

---

[3] A comparison of initial parameters for both models for Gemma in Figure 6 are shown in Appendix E

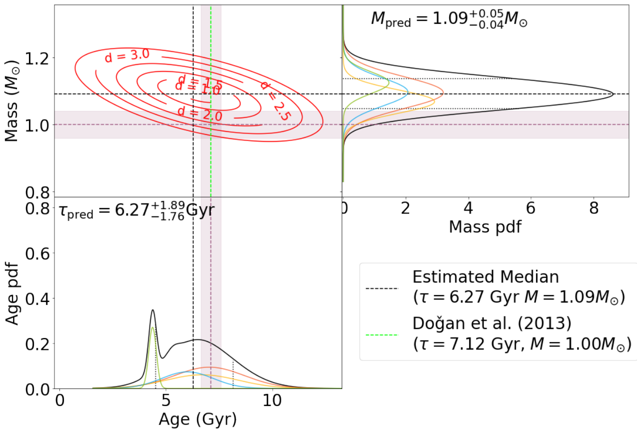

Figure 7. The estimated age and mass distribution for KIC 10920273 (Scully). The displayed features on the plot are the same as in Figure 5.

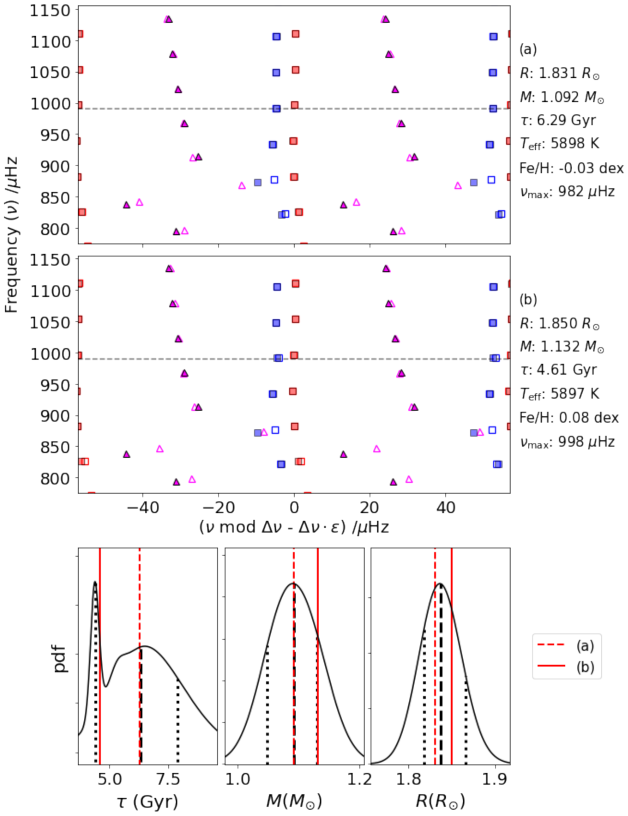

Figure 8. (a) Model generated using the median estimated initial parameters for KIC 10920273 (Scully). Model frequencies are represented by open symbols, and filled symbols represent observed frequencies. The stellar model in panel (b) has different model parameters than the model in (a), and has an age corresponding to the younger-age peak at $\tau = 4.50$ Gyr in the estimated age distribution. The bottom panel shows the τ, M, and R of models in (a) and (b) superimposed on their respective estimated distributions.

Table 6. Summary results of the network on KIC 11395018 (Boogie). The probability densities for each estimated parameter are shown in Appendix C.

Parameter	This work	Doğan et al. (2013)
Age (Gyr)	$4.55^{+0.57}_{-0.42}$	4.57 ± 0.23
Mass ($M_\odot$)	1.34 ± 0.06	1.27 ± 0.04
Radius ($R_\odot$)	$2.225^{+0.036}_{-0.034}$	2.184 ± 0.024
Luminosity ($L_\odot$)	$4.72^{+0.42}_{-0.36}$	4.54 ± 0.30
Y_0	$0.259^{+0.033}_{-0.026}$	0.276 ± 0.022
Z_0	$0.022^{+0.005}_{-0.002}$	0.023 ± 0.003
α_{MLT}	$1.94^{+0.14}_{-0.12}$	1.91 ± 0.09
α_{over}	$0.014^{+0.276}_{-0.013}$	-
α_{under}	$0.017^{+0.309}_{-0.016}$	-
D	$0.043^{+1.530}_{-0.042}$	-

$6.27^{+1.89}_{-1.76}$ Gyr is consistent with the Doğan et al. (2013) estimate of 7.12 ± 0.47 Gyr. Additionally, our estimates for M and R agree with the Doğan et al. (2013) estimates at the combined 2σ level.

In Figure 7, we show the network's estimated distributions for age and mass, where distinct bi-modality can once again be seen in the age distribution. The first peak is broad and is approximately centered at the median age estimate of τ = 6.27 Gyr, while the second peak has a much smaller dispersion and is centered at τ = 4.50 Gyr. Similar to the analysis on Gemma, we test our estimates' self-consistency by matching the observed properties of KIC 10920273 using models generated from our estimates. In Figure 8a, we show a model with a good match to the star's observables that uses our median estimates as initial parameters. This model's age of τ = 6.29 Gyr is consistent with the first, older-age peak seen in the age distribution.

We find a solution that matches the younger-age peak with a model that has α_{MLT} larger by 0.7σ, Y_0 larger by 0.2σ, and Z_0 larger by 1σ. An increase in Y_0 decreases the mass required for a stellar model to achieve the same avoided crossing frequencies at the same $\Delta\nu$ as the model in Figure 8a (Deheuvels & Michel 2011). However, increasing either α_{MLT} or Z_0 achieves the opposite effect (i.e. increases the model mass required). The net effect of these parameter differences results in the model in Figure 8b, which is $0.04 M_\odot$ more massive compared to the model in Figure 8a and has an age of τ = 4.51 Gyr that is consistent with the younger-age peak in the estimated age distribution[4]. Although this younger-age model has a larger [Fe/H] as a result of requiring a larger Z_0, it is still consistent with KIC 10920273's observed [Fe/H] value.

3.2.3 *KIC 11395018 (Boogie)*

KIC 11395018 (Kp ~10.8) is another G-type subgiant star characterized by Doğan et al. (2013) based on the asteroseismic analysis by Mathur et al. (2011). Using the same methodology for characterizing KIC 10920273, they

---

[4] A comparison of initial parameters for both models for Scully in Figure 8 are shown in Appendix E

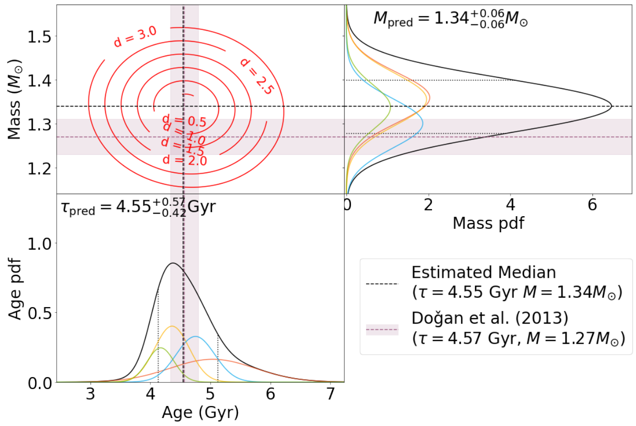

Figure 9. The estimated age and mass distribution for KIC 11395018 (Boogie). The displayed features on the plot are the same as in Figure 5.

report $\tau = 4.57 \pm 0.23$ Gyr, $M = 1.27 \pm 0.04$ $M_\odot$, and $R = 2.184 \pm 0.024$ $R_\odot$.

We use the maximal frequency set extracted by Mathur et al. (2011) as input to our network. We adopt the same spectroscopic parameters used by Doğan et al. (2013), which are $T_{\rm eff} = 5700 \pm 100$ K and [Fe/H] $= 0.13 \pm 0.1$ dex. We also use a $\nu_{\rm max}$ value of 847 ± 16 μHz, as reported by the APOKASC catalog (Serenelli et al. 2017). A summary of these estimates are tabulated in Table 6, where we note that our M and R estimates are within combined 2σ limits of the Doğan et al. (2013) solution. The estimated age distribution is shown in Figure 9, which shows a uni-modal solution. Using a mass 0.2σ above the median along with other initial model parameters at their respective median values, we find a model showing a good match to the observed properties of KIC 11395018 as shown in Figure 10. The age of the model of $\tau = 3.95$ Gyr is within uncertainties of our network's estimate of $\tau = 4.55^{+0.57}_{-0.42}$ Gyr.

3.2.4 *KIC 10005473 (HR 7322)*

KIC 10005473 is a bright ($Kp \sim 6$), F-type subgiant star that was subject to an in-depth analysis using interferometry, asteroseismology, and grid-based modelling by Stokholm et al. (2019). They find a best-fitting model based on a fixed $\alpha_{\rm MLT} = 1.6$ with no heavy element diffusion or convective overshooting included, which yielded the following values: $\tau = 4.27^{+0.05}_{-0.04}$ Gyr, $M = 1.20 \pm 0.01$ $M_\odot$, and $R = 1.954 \pm 0.006$ $R_\odot$.

As inputs to the network for our analysis of KIC 10005473, we used the oscillation mode frequencies extracted by Stokholm et al. (2019), as well as their measured parameters $\nu_{\rm max} = 960 \pm 15$ μHz, $T_{\rm eff} = 6313 \pm 50$ K, and [Fe/H] $= -0.23 \pm 0.06$ dex. Our estimates are tabulated in Table 7, showing a median age that is ~ 1 Gyr younger and a median mass that is ~ 0.1 $M_\odot$ more massive than the solution obtained by Stokholm et al. (2019). We find that our M and R estimates agree well with those obtained from asteroseismic scaling relations and interferometry as reported by Stokholm et al. (2019), but are discrepant by more than a combined 2σ limit with their modelled estimates.

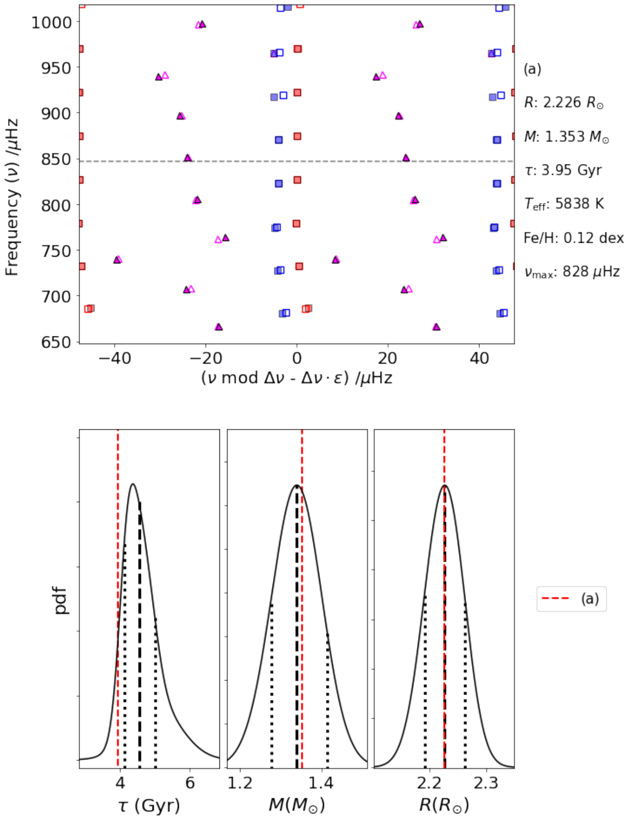

Figure 10. Model generated using the median estimated initial parameters for KIC 11395018 (Boogie), except for mass, which is 0.1σ above the median. Model frequencies are represented by open symbols, and filled symbols represent observed frequencies. (Bottom) τ, M, and R of the model superimposed on their respective estimated distributions.

In Figure 11, we show our estimated distributions for age and mass, which are both uni-modal. In Figure 12, we show a model generated using our median estimates that is a good match to the observables of KIC 10005473 with an age only 0.4σ from the distribution median and is near the distribution mode. In their study, Stokholm et al. (2019) found that their models can achieve a radius close to 2 $R_\odot$ only by allowing Y_0 to be less than the primordial helium abundance, $Y_p \simeq 0.2467$ (Planck Collaboration et al. 2016). Our results show that it is possible to obtain a good-matching model with $Y_0 > Y_p$.

3.3 Boutique modelling - Discussion

3.3.1 *Model Accuracy*

Our results show that the best match to the observed oscillation frequencies may not always correspond to the distribution median or mode for any particular parameter. This is because our method does not directly optimize the match between model and observed frequencies, as is done by conventional χ^2 optimization techniques. Because the network learns multiple realizations of input uncertainties and systematics for a given star during training, our estimates indicate credible intervals within which one realization (which is

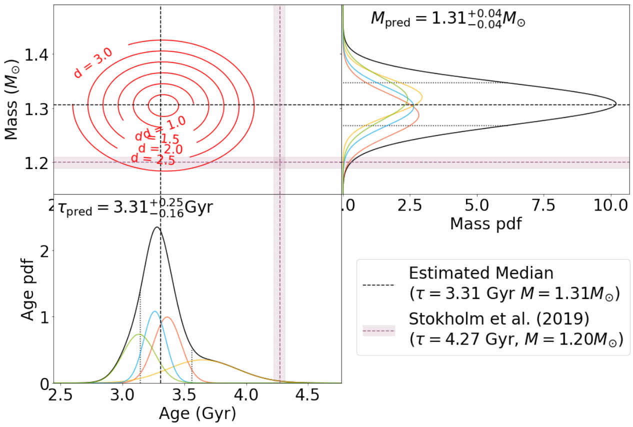

Figure 11. The estimated age and mass distribution for KIC 10005473 (HR 7322). The displayed features on the plot are the same as in Figure 5.

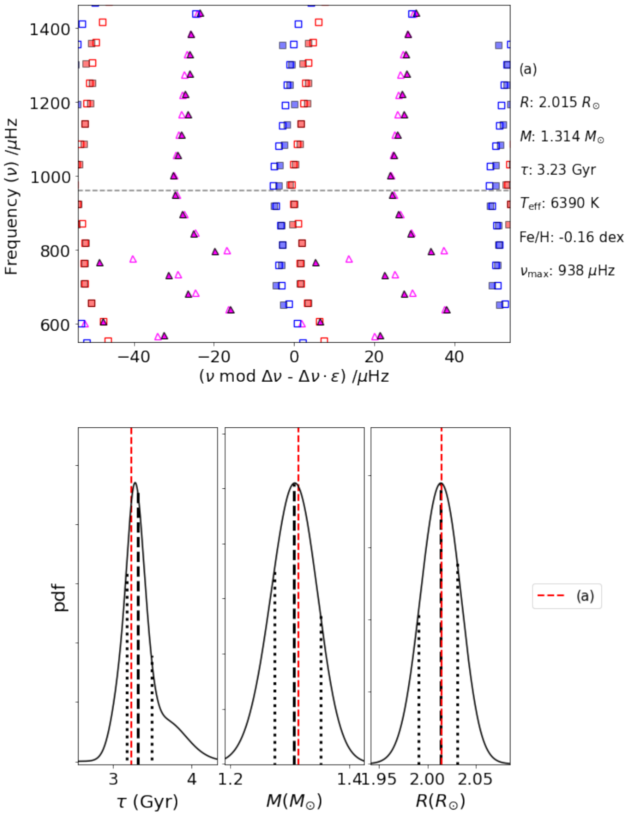

Figure 12. (Top) Model generated using the median estimated initial parameters for KIC 10005473 (HR 7322), with radius optimized to match the estimated median radius. Model frequencies are represented by open symbols, and filled symbols represent observed frequencies. (Bottom) Model τ, M and R superimposed on their respective estimated distributions.

Table 7. Summary results of the network on KIC 10005473 (HR 7322). The probability densities for each estimated parameter are shown in Appendix C.

Parameter	This work	Stokholm et al. (2019)
Age (Gyr)	$3.31^{+0.25}_{-0.16}$	$4.27^{+0.05}_{-0.04}$
Mass ($M_\odot$)	1.31 ± 0.04	1.20 ± 0.01 1.35 ± 0.07^a
Radius ($R_\odot$)	$2.013^{+0.021}_{-0.019}$	1.954 ± 0.006 2.04 ± 0.04^a 2.00 ± 0.03^b
Luminosity ($L_\odot$)	5.76 ± 0.24	5.37 ± 0.06
Y_0	$0.248^{+0.025}_{-0.019}$	0.261 ± 0.001^c
Z_0	$0.012^{+0.003}_{-0.002}$	0.010 ± 0.001^c
$\alpha_{\rm MLT}$	1.86 ± 0.10	1.60
$\alpha_{\rm over}$	$0.008^{+0.084}_{-0.007}$	-
$\alpha_{\rm under}$	$0.014^{+0.238}_{-0.013}$	-
D	$0.060^{+1.194}_{-0.059}$	-

[a] Value measured using asteroseismic scaling relations.
[b] Value measured by interferometry.
[c] Calculated using the Galactic chemical evolution law with $\Delta Y/\Delta Z = 1.4$.

the case when measuring the properties of a subgiant star) is likely to be found. Nonetheless, good matches to the observed subgiant properties are indeed found for each star analyzed in Section 3.2 with initial parameters within uncertainties from our median estimates. We stress, however, that this does not imply that these models presented in Figures 6, 8, 10, or 12 are the best solutions globally. Their purpose is to show that our estimates span a region of parameter space where a good-matching model can potentially be found.

3.3.2 *Interpreting Age Estimates*

In principle, stellar ages can be inferred very precisely using avoided crossing frequencies (e.g., Stokholm et al. 2019). The uncertainties from our estimates arise from variations of other stellar parameters like Y_0, Z_0, and $\alpha_{\rm MLT}$, which alter the mass (and hence age) required by a stellar model to match the observed avoided crossing frequencies and large frequency separation $\Delta\nu$ (Deheuvels & Michel 2011). Besides seismic requirements, having $T_{\rm eff}$ and [Fe/H] as inputs further constrains solutions from our estimates to span a localized region within the HR diagram. The estimated age distribution from the network from this study thus represents the range of stellar ages that a model can have within this localized region. In the case of Gemma and Scully, the presence of bi-modalities within their respective age distributions indicate that two distinct age ranges can be found, both with good-matching solutions as shown in Figures 6 and 8. The analysis by Deheuvels & Michel (2011) shows that when varying either $\alpha_{\rm MLT}, Y_0$, or Z_0, a corresponding linear change in model mass is required to reproduce the same avoided crossing frequencies at the same $\Delta\nu$. We infer that for stars like Gemma and Scully, such changes in mass

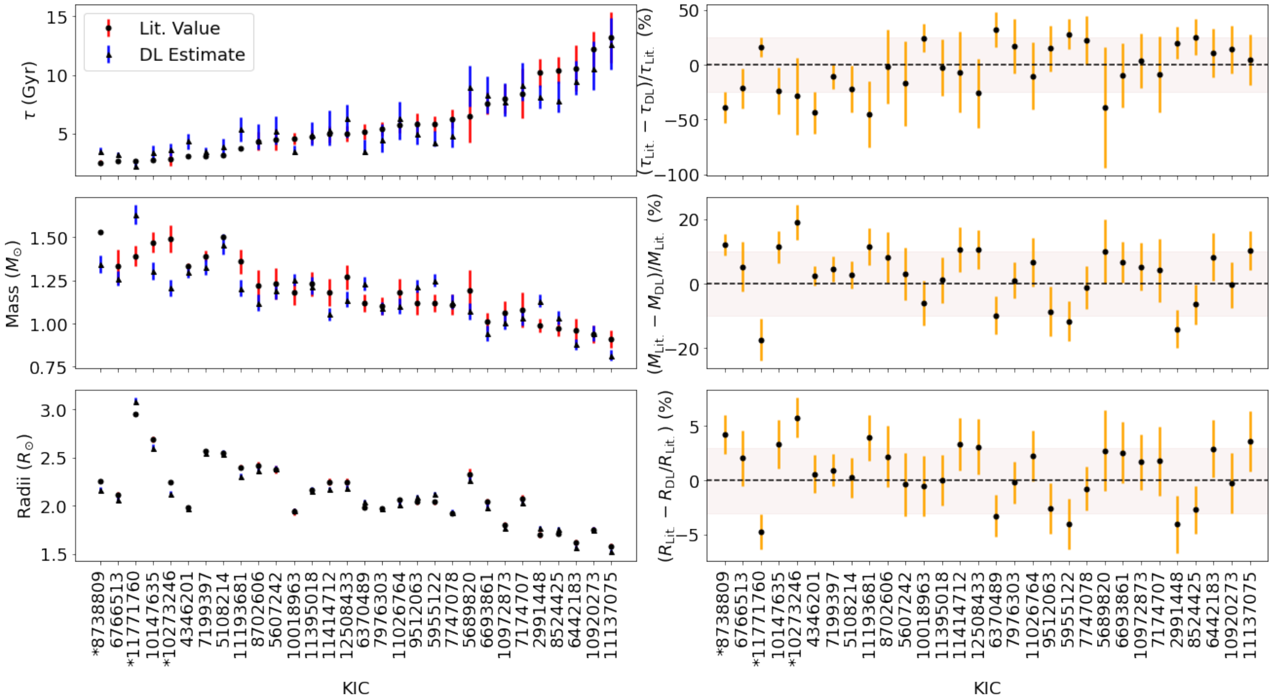

Figure 13. (Left panels) Estimates for age (top), mass (middle), and radius (bottom) values for a sample of 30 *Kepler* subgiant stars in this work (blue, subscript 'DL'), as compared to values inferred by Li et al. (2020a) using grid-based modelling (red, subscript 'Lit.') Stars in each panel are sorted by increasing $\tau_{\rm Lit.}$ from left to right. The errorbars for 'DL estimates' are the range of values between the 16th and 84th percentiles of the estimated distributions. (Right panels) Residuals of plots in the left. The errorbars of the residuals are the combined uncertainties from 'Lit.' and 'DL'. The shaded regions correspond to fractional difference intervals of 25% for age, 10% for mass, and 3% for radius. Stars with IDs beginning with an asterisk (*) have fractional differences larger than 2σ for both masses and radii.

can result in steep transitions in age, giving rise to the observed age bi-modalities. An investigation into how stellar ages vary at a fixed avoided crossing frequency and $\Delta\nu$ when varying stellar parameters can provide insight into such age transitions across different stars; however such a task is left for future research.

3.3.3 *Extending the Network's Capabilities*

Except for KIC 10005473 (HR 7322), the analysis for all stars in this Section are based on *Kepler* time series of observation length between 8-10 months. Thus, the network presented in this study can be readily applied to subgiant stars targeted by *TESS* within multiple Sectors, primarily those within the Continuous Viewing Zone. The application of our method towards subgiants observed only for a month by *TESS* is also possible, however the sparsity of detected oscillation modes due to low S/N becomes a limiting factor for this version of the network. The minimum capability of the version presented here is a frequency range of $\pm 3\Delta\nu$ around $\nu_{\rm max}$, which may not be sufficiently small for certain 1-month observations. Instead of training our network to generalize to both cases where the number of mode frequencies are sparse or plentiful, we propose in future work to train a network that focuses exclusively on observations where oscillation modes are sparse.

3.4 Ensemble Prediction

We now apply our method on a sample of 30 oscillating *Kepler* subgiant stars that were seismically analyzed by Li et al. (2020b). Using these extracted oscillation frequencies, Li et al. (2020a, hereafter T20) used a grid of stellar models to estimate ages for each subgiant in the sample. Because they find that changes to Y_0 and $\alpha_{\rm MLT}$ do not strongly influence the ages of subgiant stars, they construct a grid of models varied only in M and [Fe/H]. Consequently, they adopt a solar-calibrated $\alpha_{\rm MLT}$ of 1.9 and estimated Y_0 using the Galactic chemical evolution law. Their formulation neglects heavy element diffusion and includes an exponential overshooting scheme at the boundaries of convective cores and hydrogen-burning shells with a fixed overshooting parameter.

A comparison of our age, mass, and radius estimates[5] with the grid-based modelling approach on this ensemble is shown in Figure 13, where we find most of our estimates to be within a combined 2σ limit with the T20 estimates. Specifically, our median age estimates are typically below a 25% fractional difference to the T20 estimates. We find that our age estimates tend to be above $\tau_{\rm Lit}$ for stars with $\tau_{\rm Lit} < 5$ Gyr (equivalently, $M_{\rm Lit} > 1.4 M_\odot$), despite not seeing such tendencies when testing our network's performance in

---

[5] Our estimates of all predicted parameters for this sample are tabulated in Appendix F.

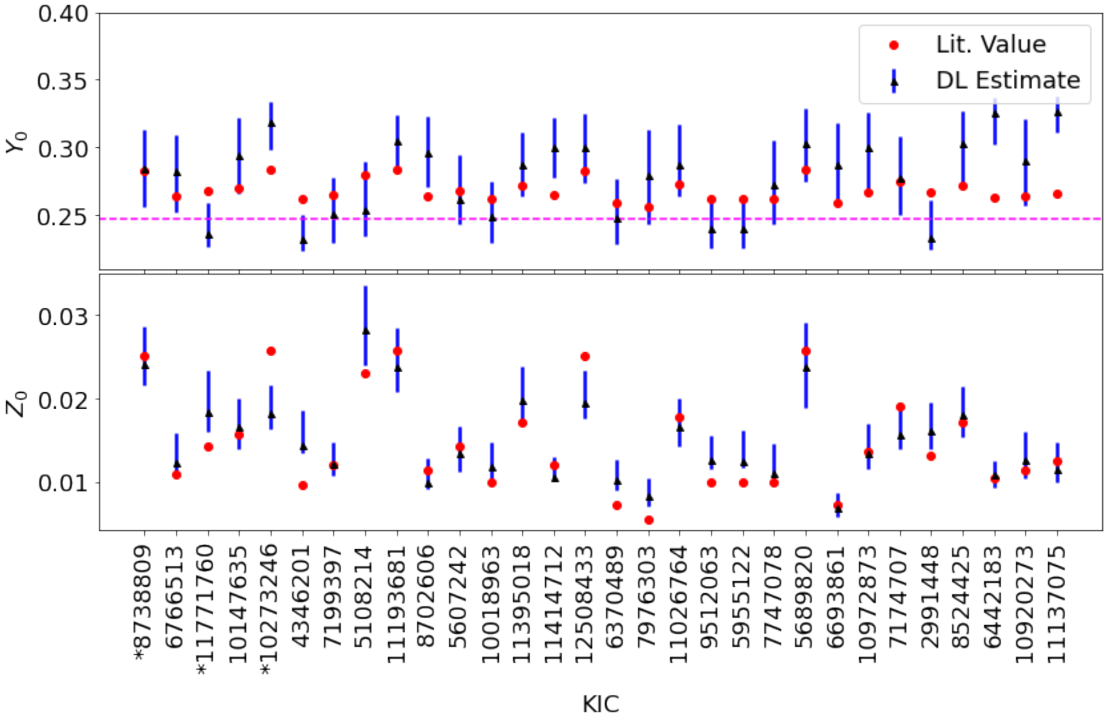

Figure 14. Comparison of initial helium abundance, Y_0, and the initial metal abundance, Z_0, for a sample of 30 *Kepler* subgiant stars. The pink line in the top panel corresponds to the primordial helium abundance, $Y_p = 0.2467$ (Planck Collaboration et al. 2016). Errorbars are 16th and 84th percentile values. Here, stars with IDs marked with an asterisk have both mass, or radii residuals larger than their respective 2σ values in Figure 13.

Section 3.1. Interestingly, we do not appear to systematically underestimate the ages of stars with $\tau > 10$ Gyr, which was another trend seen when validating our results against models.

For M and R, no significant trend can be seen in their residuals. Here, it is encouraging that M and R co-vary with one another (i.e. as M_{DL} increases, so does R_{DL}), because this indicates that the network has learned to relate these two parameters with the root mean density (or $\Delta\nu$) of the star. Our median estimates are typically below fractional differences of 10% for M and 3% for R. Stars with fractional differences in both M and R larger than 2σ are marked with asterisks in Figure 13 and are identified as KIC 8738809, KIC 10273246, and KIC 11771760. The disagreement between estimates for KIC 8738809 and KIC 11771760 is potentially due to the insufficient grid sampling from the model analyses by T20, which affected stars with $M_{\mathrm{Lit.}} > 1.3\ M_{\odot}$.

In Figure 14, we compare our Y_0 and Z_0 estimates with the values used by T20. We do not find a strong correlation between differences of Y_0 and Z_0 estimates with differences of τ estimates. This result supports T20's statement that subgiant ages are relatively insensitive to initial helium abundances, but stands in contrast to another result by T20, which states that τ is sensitive to [Fe/H]. These outcomes warrant further investigation into the systematic variation of input physics parameters such as the diffusion multiplier (D) to determine how the fundamental parameters of subgiant models change accordingly.

Additionally, we find 5 stars with an estimated median Y_0 marginally below the primordial helium abundance, $Y_p = 0.2467$ (Planck Collaboration et al. 2016). The occurrence of sub-primordial Y_0 solutions is a poorly-understood problem in fitting models of solar-like oscillators, and has been attributed to unknown systematic errors (e.g., Mathur et al. 2012), or the inadequacy of the input model physics used (Bonaca et al. 2012). As a result, work-around methods to this problem involve artificially penalizing sub-primordial Y_0 solutions during a grid search (e.g., Metcalfe et al. 2014) or more commonly, the use of the Galactic chemical evolution law, which effectively removes Y_0 as a free parameter. The prevalence of sub-primordial Y_0 values in our estimates may suggest that this issue cannot be solved by only having more free parameters with our current prescription of input physics in 1D stellar models. An inverse analysis, such as that which has been done for the Sun (e.g., Basu 2016) and main-sequence stars (Bellinger et al. 2017, 2019b) would be useful to identify missing physics from the evolutionary simulations. Nonetheless, we find that the 5 stars with estimated Y_0 marginally below Y_p do not correlate to large differences between our estimates and literature values for either age, mass, or radius. In subsequent work, we will further analyze correlations between other parameter estimates of the subgiant ensemble.

Because our estimates on an ensemble show good levels of agreement while only requiring very little computational time, our method is expected to appeal towards the inference of fundamental stellar properties over a large sample of subgiant stars. Such an inference task will be particularly valuable for characterizing stellar populations from *TESS* as well as those from the PLATO mission (Rauer et al. 2014) in the coming years. Following subgiant stars, we envision in future research that fundamental parameter inference may also be attempted with deep learning algorithms for evolved red giant stars showing solar-like oscillations.

4 CONCLUSIONS

We have developed a deep learning algorithm that estimates the fundamental parameters of oscillating subgiant stars. By training a neural network on a grid of stellar models, our method takes as input the observed oscillation frequencies as well as spectroscopic and asteroseismic parameters, and subsequently outputs a 10D distribution comprising estimates of age, mass, radius, luminosity, the mixing length parameter, overshooting and undershooting coefficients, and the diffusion multiplier. Besides a large degree of freedom in exploring various combinations of model physics for subgiant stars, additional novelties in our approach include the use of échelle diagrams to represent oscillation frequency patterns and the use of a mixture density network to estimate parameter distributions instead of point estimates.

We applied our method to three oscillating subgiant stars previously modelled based on *Kepler* observations of 8-10 months: KIC 11026764 (nicknamed Gemma), KIC 10920273 (nicknamed Scully), and KIC 11395018 (nicknamed Boogie). Our estimates on KIC 11026764 showed good agreement with previously modelled estimates for all parameters. Additionally, our age, mass and radius estimates for KIC 10920273 and KIC 11395018 also agree with their previously modelled estimates comprising a consensus solution from multiple stellar evolutionary codes. We showed that stellar models generated using our estimates result in good matches to the observed frequency and spectroscopic measurements for these *Kepler* subgiants. Additionally, we applied our method to the asteroseismic benchmark subgiant star HR 7322 and estimated a solution 1 Gyr younger than previously modelled estimates, but with a mass and radius that agree well with estimates from asteroseismic scaling relations and interferometry.

Finally, we estimated the fundamental parameters of a sample of 30 *Kepler* subgiant stars and find good agreement with solutions obtained by traditional grid-based modelling using different prescriptions of input model physics. In particular, a majority of our estimates have fractional differences of below 25% for age, below 10% for mass, and below 3% for radius, with only three stars with mass and radius discrepant above the 2σ level. The method presented in this study brings utility to the detailed modelling of individual subgiant stars in the form of initial estimates, and can reliably determine the fundamental parameters of a large sample of subgiant stars extremely efficiently, which will be a valuable task for stellar population studies with the *TESS* mission.

ACKNOWLEDGEMENTS

Funding for this Discovery mission is provided by NASA's Science Mission Directorate. We thank the entire *Kepler* team without whom this investigation would not be possible. D.S. is the recipient of an Australian Research Council Future Fellowship (project number FT1400147). We acknowledge funding and support from the Stellar Astrophysics Center (SAC), which initiated this project at the Max Planck Institute for Solar System Research in Göttingen, Germany, in cooperation with the Stellar Ages and Galactic Evolution (SAGE) group. Funding for the Stellar Astrophysics Centre is provided by The Danish National Research Foundation (Grant agreement no.: DNRF106). The research leading to the presented results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 338251 (StellarAges). We thank Tanda Li and Yaguang Li, for interesting and fruitful discussions. We also thank Jørgen Christensen-Dalsgaard for useful comments on the manuscript. Finally, we gratefully acknowledge the support of NVIDIA Corporation for the donation of the Titan Xp GPU used for developing the neural networks in this research.

REFERENCES

Aizenman M., Smeyers P., Weigert A., 1977, A&A, 58, 41
Angelou G. C., Bellinger E. P., Hekker S., Basu S., 2017, ApJ, 839, 116
Angelou G. C., Bellinger E. P., Hekker S., Mints A., Elsworth Y., Basu S., Weiss A., 2020, MNRAS, 493, 4987
Appourchaux T., et al., 2012, A&A, 543, A54
Ball W. H., Gizon L., 2014, A&A, 568, A123
Ball W. H., Gizon L., 2017, A&A, 600, A128
Basu S., 2016, Living Reviews in Solar Physics, 13, 2
Bedding T. R., 2014, Solar-like oscillations: An observational perspective. p. 60
Bellinger E. P., Christensen-Dalsgaard J., 2019, ApJ, 887, L1
Bellinger E. P., Angelou G. C., Hekker S., Basu S., Ball W. H., Guggenberger E., 2016, ApJ, 830, 31
Bellinger E. P., Basu S., Hekker S., Ball W. H., 2017, ApJ, 851, 80
Bellinger E. P., Hekker S., Angelou G. C., Stokholm A., Basu S., 2019a, A&A, 622, A130
Bellinger E. P., Basu S., Hekker S., Christensen-Dalsgaard J., 2019b, ApJ, 885, 143
Benomar O., Bedding T. R., Stello D., Deheuvels S., White T. R., Christensen-Dalsgaard J., 2012, The Astrophysical Journal, 745, L33
Benomar O., et al., 2014, ApJ, 781, L29
Bishop C. M., 1994, Technical report, Mixture density networks. Aston University
Bonaca A., et al., 2012, ApJ, 755, L12
Borucki W. J., et al., 2010, Science, 327, 977
Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, ApJ, 368, 599
Bruntt H., et al., 2012, MNRAS, 423, 122
Campante T. L., et al., 2011, A&A, 534, A6
Campante T. L., et al., 2016, ApJ, 830, 138
Chaplin W. J., et al., 2010, ApJ, 713, L169
Christensen-Dalsgaard J., 1984, in Mangeney A., Praderie F., eds, Space Research in Stellar Activity and Variability. p. 11
Christensen-Dalsgaard J., Bedding T. R., Kjeldsen H., 1995, ApJ, 443, L29
Creevey O. L., et al., 2012, A&A, 537, A111
Deheuvels S., Michel E., 2009, Astrophysics and Space Science, 328, 259
Deheuvels S., Michel E., 2011, A&A, 535, A91
Deheuvels S., et al., 2014, A&A, 564, A27

Doğan G., et al., 2013, ApJ, 763, 49

Gough D. O., 1990, Comments on Helioseismic Inference. p. 283, doi:10.1007/3-540-53091-6

Grec G., Fossat E., Pomerantz M. A., 1983, Sol. Phys., 82, 55

Hendriks L., Aerts C., 2019, PASP, 131, 108001

Huber D., et al., 2011, ApJ, 743, 143

Kjeldsen H., Bedding T. R., 1995, A&A, 293, 87

Li Y.-G., Du M.-H., Xie B.-H., Tian Z.-J., Bi S.-L., Li T.-D., Wu Y.-Q., Liu K., 2017, Research in Astronomy and Astrophysics, 17, 044

Li T., Bedding T. R., Christensen-Dalsgaard J., Stello D., Yaguange L., Keen M. A., 2020a, MNRAS

Li Y., Bedding T. R., Li T., Bi S., Stello D., Zhou Y., White T. R., 2020b, MNRAS,

Lund M. N., et al., 2017, ApJ, 835, 172

Mahalanobis P. C., 1936.

Mathur S., et al., 2011, ApJ, 733, 95

Mathur S., et al., 2012, ApJ, 749, 152

Metcalfe T. S., Creevey O. L., Christensen-Dalsgaard J., 2009, The Astrophysical Journal, 699, 373

Metcalfe T. S., et al., 2010, ApJ, 723, 1583

Metcalfe T. S., et al., 2014, ApJS, 214, 27

Osaki Y., 1975, PASJ, 27, 237

Paszke A., et al., 2019, in Wallach H., Larochelle H., Beygelzimer A., d'Alché-Buc F., Fox E., Garnett R., eds, , Advances in Neural Information Processing Systems 32. Curran Associates, Inc., pp 8024–8035

Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, ApJS, 192, 3

Paxton B., et al., 2013, ApJS, 208, 4

Paxton B., et al., 2015, ApJS, 220, 15

Paxton B., et al., 2018, ApJS, 234, 34

Paxton B., et al., 2019, ApJS, 243, 10

Planck Collaboration et al., 2016, A&A, 594, A13

Prša A., et al., 2016, AJ, 152, 41

Rauer H., et al., 2014, Experimental Astronomy, 38, 249

Rendle B. M., et al., 2019, MNRAS, 484, 771

Roxburgh I. W., Vorontsov S. V., 2003, A&A, 411, 215

Schofield M., et al., 2019, ApJS, 241, 12

Serenelli A., et al., 2017, ApJS, 233, 23

Shibahashi H., 1979, PASJ, 31, 87

Stokholm A., Nissen P. E., Silva Aguirre V., White T. R., Lund M. N., Mosumgaard J. R., Huber D., Jessen-Hansen J., 2019, MNRAS, 489, 928

Townsend R. H. D., Teitler S. A., 2013, MNRAS, 435, 3406

Townsend R. H. D., Goldstein J., Zweibel E. G., 2018, MNRAS, 475, 879

Verma K., Hanasoge S., Bhattacharya J., Antia H. M., Krishnamurthi G., 2016, MNRAS, 461, 4206

Viani L. S., Basu S., Ong J. M. J., Bonaca A., Chaplin W. J., 2018, ApJ, 858, 28

White T. R., Bedding T. R., Stello D., Christensen-Dalsgaard J., Huber D., Kjeldsen H., 2011, ApJ, 743, 161

Woitaszek M., Metcalfe T., Shorrock I., 2009, in Proceedings of the 5th Grid Computing Environments Workshop. p. 1, doi:10.1145/1658260.1658262

APPENDIX A: NETWORK ARCHITECTURE

We detail the structure of the network in Table A1. The network is developed using the Pytorch version 1.1.0 deep learning library (Paszke et al. 2019).

Table A1. Structure of the neural network for stellar model inference.

Component	Layer	Weight Shape	Output Shape
Convolutional Network	conv1[a]	(8,5)	(128,128,8)
	pool1	-	(64,64,8)
	conv2	(16,3)	(64,64,16)
	pool2	-	(32,32,16)
	conv3	(32,3)	(32,32,32)
	pool3	-	(16,16,16)
	flatten	-	4096
	concatenate[b]	-	4096×9
	dense1	(36864, 512)	512
	dense2	(512, 512)	512
Mixture Density Network	μ-dense1	(512, 256)	256
	μ-dense (output)	(256, 10×10)	512
	σ-dense1	(512, 256)	256
	σ-dense (output)	(256, 10×10)	512
	π-dense1	(512, 256)	256
	π-dense (output)	(256, 10)	512

[a] For convolutional layers, weight shapes are in format (number of filters, receptive field size), while output shapes are in format (height, width, number of filters).

[b] Each input observable (except the échelle diagram) is multiplied with a copy of the flatten layer output and concatenated with the same layer's output.

APPENDIX B: BOOTSTRAPPING PROCEDURE

A summary of data generation for each training iteration is described by the following pseudo-code, with the notation (') implying perturbed quantities:

```
for each training iteration do
    for each stellar model do
        Obtain mode frequencies, $\nu$
        Sample artificial surface term $c$ and calculate $\delta\nu_{\mathrm{surf}}$
        $\nu' \longleftarrow \nu - \delta\nu_{\mathrm{surf}}$
        Sample $\sigma_{l=0,1}$
        $\nu' \longleftarrow \nu' + \sigma_{l=0,1}$
        Calculate missing mode factor $P$
        $\nu' \longleftarrow P \times \nu'$
        Calculate $\Delta\nu$, $\epsilon$
        Sample $\sigma_{\nu_{\mathrm{max}}}, \sigma_{T_{\mathrm{eff}}}, \sigma_{[\mathrm{Fe/H}]}$
        $\nu'_{\mathrm{max}} \longleftarrow \nu_{\mathrm{max}} + \sigma_{\nu_{\mathrm{max}}}$
        $T'_{\mathrm{eff}}, [\mathrm{Fe/H}]' \longleftarrow T_{\mathrm{eff}} + \sigma_{T_{\mathrm{eff}}}, [\mathrm{Fe/H}] + \sigma_{[\mathrm{Fe/H}]}$
        Construct échelle diagram, $\mathcal{E}$
        return $\mathcal{E}, \nu'_{\mathrm{max}}, T'_{\mathrm{eff}}, [\mathrm{Fe/H}]', \sigma_{l=0,1}, \sigma_{\nu_{\mathrm{max}}}, \sigma_{T_{\mathrm{eff}}}, \sigma_{[\mathrm{Fe/H}]}$
    end for
    Calculate network negative log-likelihood, $E$
    Update network weights
end for
```

Table D1. Input parameters for the good-matching models for Gemma. The age of each model is included for reference.

Age	M	Y_0	Z_0	$\alpha_{\rm MLT}$	$\alpha_{\rm over}$	$\alpha_{\rm under}$	D
4.51 Gyr	1.274 $M_\odot$	0.252	0.020	1.84	0.011	0.013	0.029
5.43 Gyr	1.264 $M_\odot$	0.253	0.020	1.81	0.011	0.010	0.033

Table E1. Input parameters for the good-matching models for Scully. The age of each model is included for reference.

Age	M	Y_0	Z_0	$\alpha_{\rm MLT}$	$\alpha_{\rm over}$	$\alpha_{\rm under}$	D
4.61 Gyr	1.132 $M_\odot$	0.296	0.019	2.03	0.017	0.009	0.025
6.29 Gyr	1.092 $M_\odot$	0.290	0.015	1.94	0.009	0.022	0.118

APPENDIX C: ESTIMATED 10D DISTRIBUTION FOR BOUTIQUE-MODELED SUBGIANT STARS

In Figures C1, C2, C3, and C4, we show the full estimated probability densities for KIC 11026764, KIC 10920273, KIC 11395018, and KIC 10005473, respectively. The output parameters of the mixture density network in this study that reproduce the plots in each Figure are available at `https://github.com/mtyhon/deep-sub`.

APPENDIX D: INITIAL PARAMETERS FOR GEMMA

Table E1 lists the input parameters used to generate the two models for Gemma in Figure 6.

APPENDIX E: INITIAL PARAMETERS FOR SCULLY

Table E1 lists the input parameters used to generate the two models for Scully in Figure 8.

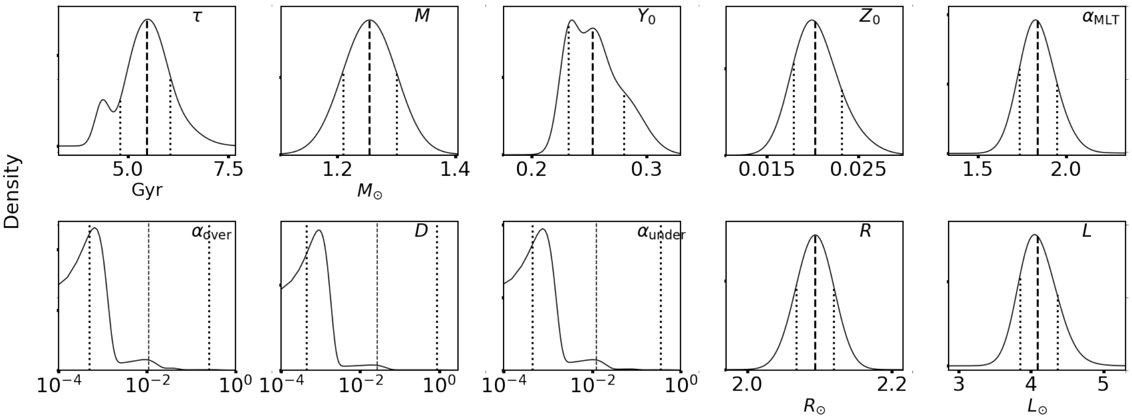

Figure C1. Probability density of each estimated parameter for KIC 11026764 (Gemma). The black dashed lines indicate the median estimated values, with dotted black lines representing the 16th and 84th percentile values.

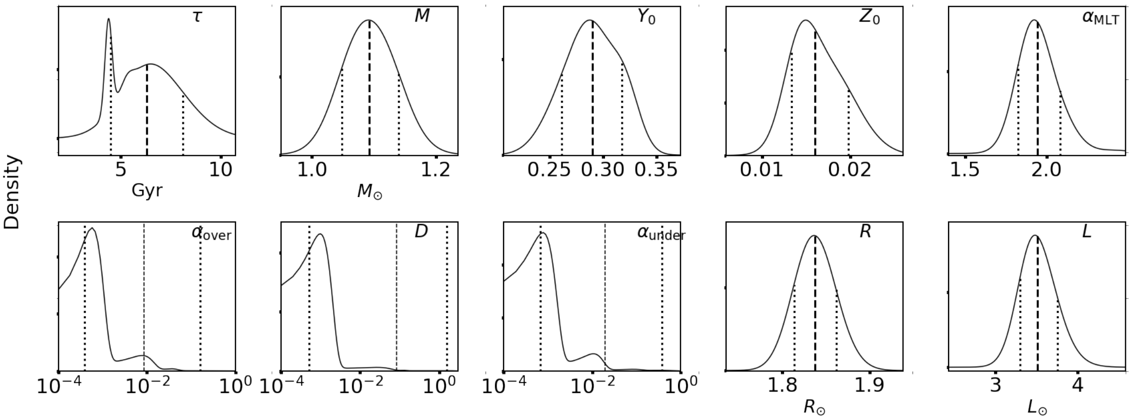

Figure C2. Probability density of each estimated parameter for KIC 10920273 (Scully). The black dashed lines indicate the median estimated values, with dotted black lines representing the 16th and 84th percentile values.

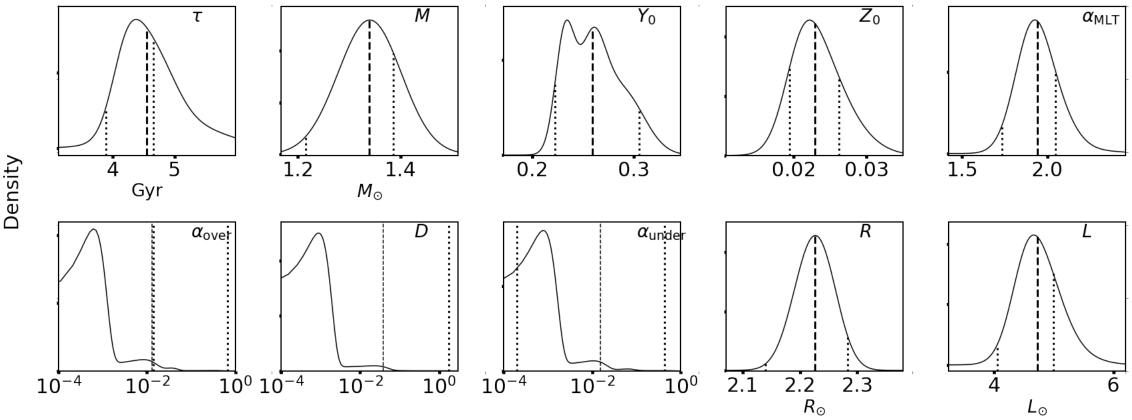

Figure C3. Probability density of each estimated parameter for KIC 11395018 (Boogie). The black dashed lines indicate the median estimated values, with dotted black lines representing the 16th and 84th percentile values.

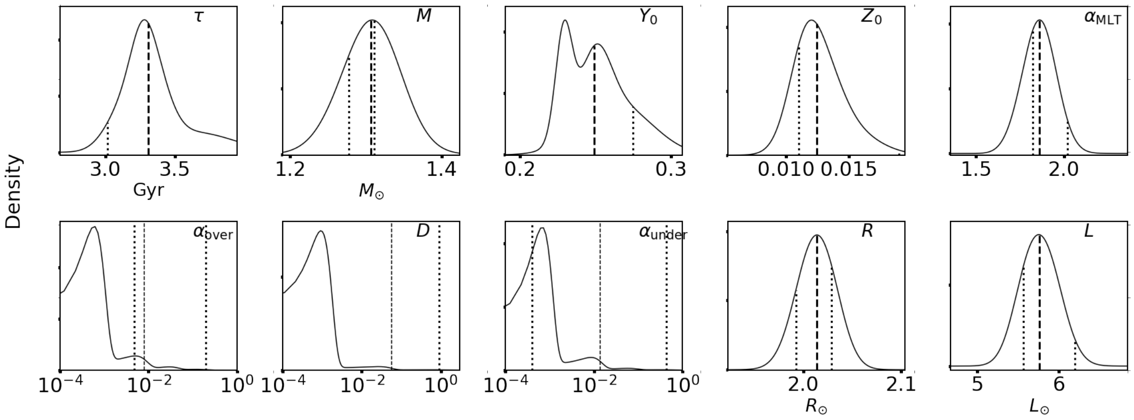

Figure C4. Probability density of each estimated parameter for KIC 10005473. The black dashed lines indicate the median estimated values, with dotted black lines representing the 16th and 84th percentile values.

Table F1. Estimates for the ensemble of 30 oscillating *Kepler* subgiant stars in this study.

KIC	τ (Gyr)	M ($M_\odot$)	R ($R_\odot$)	L ($L_\odot$)	Y_0	Z_0	$\alpha_{\rm MLT}$	$\alpha_{\rm over}$	$\alpha_{\rm under}$	D
2991448	$8.15^{+0.94}_{-1.04}$	$1.13^{+0.03}_{-0.03}$	$1.769^{+0.017}_{-0.018}$	$2.84^{+0.19}_{-0.18}$	$0.233^{+0.026}_{-0.009}$	$0.016^{+0.003}_{-0.003}$	$1.77^{+0.12}_{-0.11}$	$0.007^{+0.178}_{-0.007}$	$0.013^{+0.355}_{-0.013}$	$0.031^{+1.314}_{-0.031}$
4346201	$4.36^{+0.53}_{-0.72}$	$1.30^{+0.03}_{-0.03}$	$1.969^{+0.017}_{-0.018}$	$4.72^{+0.28}_{-0.28}$	$0.232^{+0.016}_{-0.009}$	$0.014^{+0.003}_{-0.002}$	$1.70^{+0.10}_{-0.10}$	$0.010^{+0.051}_{-0.010}$	$0.025^{+0.219}_{-0.025}$	$0.676^{+1.142}_{-0.675}$
5108214	$3.86^{+0.67}_{-0.56}$	$1.46^{+0.06}_{-0.06}$	$2.544^{+0.034}_{-0.038}$	$6.61^{+0.43}_{-0.41}$	$0.254^{+0.033}_{-0.020}$	$0.028^{+0.005}_{-0.005}$	$1.79^{+0.11}_{-0.10}$	$0.010^{+0.236}_{-0.010}$	$0.094^{+0.378}_{-0.093}$	$0.329^{+0.869}_{-0.328}$
5607242	$5.20^{+1.27}_{-1.10}$	$1.19^{+0.05}_{-0.05}$	$2.389^{+0.029}_{-0.034}$	$4.69^{+0.32}_{-0.29}$	$0.262^{+0.030}_{-0.020}$	$0.014^{+0.003}_{-0.003}$	$1.78^{+0.12}_{-0.10}$	$0.010^{+0.430}_{-0.010}$	$0.033^{+0.392}_{-0.032}$	$0.161^{+1.078}_{-0.160}$
5689820	$8.98^{+1.80}_{-1.71}$	$1.07^{+0.05}_{-0.05}$	$2.266^{+0.038}_{-0.036}$	$3.50^{+0.23}_{-0.21}$	$0.303^{+0.024}_{-0.028}$	$0.024^{+0.005}_{-0.005}$	$2.10^{+0.19}_{-0.15}$	$0.015^{+0.584}_{-0.015}$	$0.033^{+0.442}_{-0.032}$	$0.021^{+0.624}_{-0.021}$
5955122	$4.20^{+0.98}_{-0.28}$	$1.25^{+0.04}_{-0.04}$	$2.122^{+0.020}_{-0.022}$	$4.86^{+0.30}_{-0.29}$	$0.240^{+0.021}_{-0.015}$	$0.013^{+0.003}_{-0.001}$	$1.85^{+0.11}_{-0.10}$	$0.022^{+0.050}_{-0.022}$	$0.041^{+0.131}_{-0.040}$	$1.385^{+0.973}_{-1.384}$
6370489	$3.48^{+0.95}_{-0.15}$	$1.23^{+0.04}_{-0.04}$	$2.045^{+0.020}_{-0.021}$	$5.66^{+0.30}_{-0.32}$	$0.248^{+0.026}_{-0.020}$	$0.010^{+0.002}_{-0.002}$	$1.95^{+0.11}_{-0.11}$	$0.024^{+0.050}_{-0.024}$	$0.046^{+0.124}_{-0.045}$	$1.532^{+0.784}_{-1.529}$
6442183	$9.48^{+1.61}_{-1.22}$	$0.88^{+0.03}_{-0.03}$	$1.573^{+0.017}_{-0.017}$	$2.55^{+0.15}_{-0.15}$	$0.326^{+0.010}_{-0.024}$	$0.011^{+0.001}_{-0.002}$	$1.75^{+0.14}_{-0.11}$	$0.014^{+0.339}_{-0.013}$	$0.039^{+0.334}_{-0.039}$	$0.056^{+0.662}_{-0.054}$
6693861	$8.29^{+1.51}_{-1.58}$	$0.94^{+0.04}_{-0.04}$	$1.988^{+0.029}_{-0.028}$	$3.52^{+0.24}_{-0.23}$	$0.287^{+0.029}_{-0.026}$	$0.007^{+0.001}_{-0.002}$	$1.70^{+0.13}_{-0.12}$	$0.017^{+0.400}_{-0.017}$	$0.015^{+0.252}_{-0.015}$	$0.042^{+1.341}_{-0.041}$
6766513	$3.19^{+0.22}_{-0.23}$	$1.26^{+0.04}_{-0.04}$	$2.067^{+0.022}_{-0.022}$	$6.02^{+0.32}_{-0.32}$	$0.282^{+0.026}_{-0.030}$	$0.012^{+0.003}_{-0.002}$	$1.72^{+0.10}_{-0.10}$	$0.011^{+0.084}_{-0.011}$	$0.018^{+0.172}_{-0.018}$	$0.482^{+1.186}_{-0.481}$
7174707	$9.14^{+1.69}_{-1.38}$	$1.04^{+0.04}_{-0.04}$	$2.033^{+0.029}_{-0.028}$	$2.98^{+0.16}_{-0.15}$	$0.278^{+0.030}_{-0.028}$	$0.016^{+0.003}_{-0.002}$	$1.80^{+0.12}_{-0.12}$	$0.016^{+0.409}_{-0.016}$	$0.016^{+0.330}_{-0.015}$	$0.030^{+0.954}_{-0.029}$
7199397	$3.43^{+0.33}_{-0.24}$	$1.33^{+0.04}_{-0.05}$	$2.546^{+0.028}_{-0.031}$	$6.94^{+0.44}_{-0.41}$	$0.251^{+0.026}_{-0.022}$	$0.012^{+0.002}_{-0.002}$	$1.72^{+0.10}_{-0.09}$	$0.019^{+0.143}_{-0.019}$	$0.037^{+0.236}_{-0.036}$	$0.612^{+1.413}_{-0.611}$
7747078	$4.83^{+1.48}_{-0.98}$	$1.12^{+0.04}_{-0.04}$	$1.935^{+0.021}_{-0.023}$	$4.27^{+0.26}_{-0.25}$	$0.273^{+0.031}_{-0.030}$	$0.011^{+0.003}_{-0.002}$	$1.82^{+0.11}_{-0.10}$	$0.020^{+0.204}_{-0.020}$	$0.032^{+0.241}_{-0.031}$	$0.955^{+1.069}_{-0.953}$
7976303	$4.46^{+1.39}_{-1.05}$	$1.09^{+0.03}_{-0.04}$	$1.974^{+0.021}_{-0.021}$	$4.83^{+0.26}_{-0.27}$	$0.279^{+0.032}_{-0.036}$	$0.008^{+0.002}_{-0.002}$	$1.75^{+0.10}_{-0.10}$	$0.031^{+0.295}_{-0.030}$	$0.018^{+0.275}_{-0.018}$	$1.125^{+0.868}_{-1.086}$
8524425	$7.80^{+1.41}_{-1.04}$	$1.03^{+0.04}_{-0.04}$	$1.756^{+0.022}_{-0.021}$	$2.85^{+0.19}_{-0.19}$	$0.304^{+0.023}_{-0.033}$	$0.018^{+0.003}_{-0.003}$	$1.83^{+0.14}_{-0.11}$	$0.011^{+0.436}_{-0.011}$	$0.018^{+0.373}_{-0.018}$	$0.043^{+0.939}_{-0.042}$
8702606	$4.40^{+1.38}_{-0.69}$	$1.12^{+0.05}_{-0.05}$	$2.367^{+0.032}_{-0.034}$	$4.84^{+0.39}_{-0.34}$	$0.296^{+0.025}_{-0.026}$	$0.010^{+0.002}_{-0.001}$	$1.81^{+0.15}_{-0.12}$	$0.015^{+0.486}_{-0.015}$	$0.014^{+0.280}_{-0.013}$	$0.081^{+1.485}_{-0.081}$
8738809	$3.50^{+0.27}_{-0.24}$	$1.34^{+0.05}_{-0.05}$	$2.165^{+0.028}_{-0.026}$	$5.41^{+0.32}_{-0.29}$	$0.284^{+0.027}_{-0.029}$	$0.024^{+0.004}_{-0.003}$	$1.62^{+0.08}_{-0.08}$	$0.005^{+0.102}_{-0.005}$	$0.012^{+0.272}_{-0.012}$	$0.021^{+0.444}_{-0.021}$
9512063	$4.96^{+1.13}_{-0.95}$	$1.22^{+0.04}_{-0.04}$	$2.093^{+0.021}_{-0.023}$	$4.52^{+0.29}_{-0.28}$	$0.240^{+0.021}_{-0.015}$	$0.013^{+0.002}_{-0.002}$	$1.77^{+0.11}_{-0.10}$	$0.019^{+0.060}_{-0.019}$	$0.037^{+0.170}_{-0.036}$	$1.272^{+1.065}_{-1.270}$
10018963	$3.45^{+0.47}_{-0.14}$	$1.25^{+0.03}_{-0.03}$	$1.950^{+0.018}_{-0.017}$	$5.09^{+0.27}_{-0.27}$	$0.250^{+0.023}_{-0.020}$	$0.012^{+0.002}_{-0.002}$	$1.68^{+0.09}_{-0.10}$	$0.009^{+0.049}_{-0.009}$	$0.044^{+0.191}_{-0.043}$	$0.675^{+1.117}_{-0.674}$
10147635	$3.41^{+0.58}_{-0.46}$	$1.30^{+0.05}_{-0.05}$	$2.601^{+0.034}_{-0.033}$	$7.37^{+0.46}_{-0.43}$	$0.294^{+0.027}_{-0.029}$	$0.017^{+0.003}_{-0.003}$	$1.73^{+0.12}_{-0.11}$	$0.014^{+0.392}_{-0.014}$	$0.025^{+0.361}_{-0.025}$	$0.046^{+0.855}_{-0.045}$
10273246	$3.65^{+0.52}_{-0.46}$	$1.21^{+0.05}_{-0.05}$	$2.120^{+0.027}_{-0.028}$	$5.98^{+0.56}_{-0.47}$	$0.319^{+0.013}_{-0.021}$	$0.018^{+0.003}_{-0.003}$	$1.97^{+0.17}_{-0.15}$	$0.018^{+0.223}_{-0.017}$	$0.022^{+0.192}_{-0.022}$	$0.045^{+0.891}_{-0.045}$
10920273	$10.52^{+2.25}_{-1.78}$	$0.94^{+0.04}_{-0.04}$	$1.754^{+0.027}_{-0.026}$	$2.46^{+0.18}_{-0.17}$	$0.291^{+0.029}_{-0.035}$	$0.013^{+0.003}_{-0.003}$	$1.63^{+0.13}_{-0.11}$	$0.011^{+0.364}_{-0.011}$	$0.017^{+0.343}_{-0.017}$	$0.039^{+1.441}_{-0.038}$
10972873	$7.67^{+1.45}_{-1.20}$	$1.01^{+0.04}_{-0.04}$	$1.769^{+0.023}_{-0.023}$	$3.10^{+0.20}_{-0.19}$	$0.300^{+0.024}_{-0.031}$	$0.013^{+0.003}_{-0.002}$	$1.86^{+0.14}_{-0.11}$	$0.011^{+0.379}_{-0.011}$	$0.022^{+0.364}_{-0.021}$	$0.049^{+1.053}_{-0.048}$
11026764	$6.28^{+1.28}_{-1.78}$	$1.10^{+0.05}_{-0.05}$	$2.013^{+0.026}_{-0.027}$	$3.70^{+0.25}_{-0.23}$	$0.288^{+0.027}_{-0.025}$	$0.017^{+0.003}_{-0.003}$	$1.80^{+0.12}_{-0.10}$	$0.009^{+0.301}_{-0.009}$	$0.016^{+0.372}_{-0.015}$	$0.154^{+0.991}_{-0.153}$
11137075	$12.59^{+2.21}_{-2.10}$	$0.82^{+0.03}_{-0.03}$	$1.523^{+0.018}_{-0.020}$	$2.08^{+0.14}_{-0.14}$	$0.327^{+0.010}_{-0.017}$	$0.012^{+0.003}_{-0.002}$	$1.83^{+0.14}_{-0.13}$	$0.012^{+0.021}_{-0.011}$	$0.083^{+0.297}_{-0.081}$	$0.059^{+1.031}_{-0.059}$
11193681	$5.40^{+0.97}_{-1.00}$	$1.20^{+0.04}_{-0.05}$	$2.306^{+0.030}_{-0.031}$	$4.76^{+0.30}_{-0.28}$	$0.305^{+0.017}_{-0.021}$	$0.024^{+0.004}_{-0.004}$	$1.86^{+0.11}_{-0.09}$	$0.009^{+0.067}_{-0.009}$	$0.194^{+0.338}_{-0.192}$	$0.130^{+0.495}_{-0.129}$
11395018	$4.86^{+1.09}_{-0.82}$	$1.22^{+0.05}_{-0.05}$	$2.159^{+0.030}_{-0.029}$	$4.62^{+0.43}_{-0.38}$	$0.288^{+0.023}_{-0.024}$	$0.020^{+0.003}_{-0.003}$	$1.98^{+0.14}_{-0.12}$	$0.009^{+0.216}_{-0.009}$	$0.036^{+0.415}_{-0.035}$	$0.119^{+1.038}_{-0.118}$
11414712	$5.27^{+1.61}_{-1.28}$	$1.06^{+0.03}_{-0.04}$	$2.175^{+0.023}_{-0.025}$	$4.29^{+0.28}_{-0.26}$	$0.300^{+0.021}_{-0.022}$	$0.011^{+0.002}_{-0.001}$	$1.62^{+0.10}_{-0.08}$	$0.026^{+0.379}_{-0.025}$	$0.017^{+0.153}_{-0.017}$	$0.246^{+1.516}_{-0.246}$
11771760	$2.22^{+0.22}_{-0.17}$	$1.63^{+0.05}_{-0.06}$	$3.090^{+0.032}_{-0.036}$	$9.83^{+0.63}_{-0.60}$	$0.236^{+0.021}_{-0.011}$	$0.018^{+0.004}_{-0.003}$	$1.77^{+0.12}_{-0.11}$	$0.012^{+0.176}_{-0.012}$	$0.015^{+0.301}_{-0.015}$	$0.040^{+1.084}_{-0.039}$
12508433	$6.27^{+1.17}_{-1.58}$	$1.14^{+0.05}_{-0.04}$	$2.180^{+0.029}_{-0.028}$	$3.58^{+0.22}_{-0.19}$	$0.300^{+0.024}_{-0.026}$	$0.020^{+0.003}_{-0.003}$	$1.83^{+0.12}_{-0.10}$	$0.012^{+0.355}_{-0.012}$	$0.015^{+0.375}_{-0.014}$	$0.032^{+0.662}_{-0.032}$

APPENDIX F: ESTIMATED PARAMETERS FOR ENSEMBLE OF 30 *KEPLER* SUBGIANT STARS

In Table F1, we tabulate our full estimates on the sample of 30 *Kepler* subgiant stars that were modelled by Li et al. (2020a).

This paper has been typeset from a TEX/LATEX file prepared by the author.