

The *Kepler* Smear Campaign I: An Asteroseismic Catalogue of Bright Red Giants

Benjamin J. S. Pope,^{1,2,3*} Guy R. Davies,^{4,5} Keith Hawkins,^{6,7} Timothy R. White,^{5,8}
Daniel Huber,^{9,10,11} Ashley Chontos,⁹ Victor Silva Aguirre,⁵ Victoria Antoci,⁵ and friends

¹Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA

²NASA Sagan Fellow

³Oxford Astrophysics, Denys Wilkinson Building, University of Oxford, OX1 3RH, Oxford, UK

⁴School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK

⁵Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark

⁶Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA

⁷Department of Astronomy, Columbia University, 550 W 120th St, New York, NY 10027, USA

⁸Research School of Astronomy and Astrophysics, Mount Stromlo Observatory, The Australian National University, Canberra, ACT 2611, Australia

⁹Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

¹⁰Sydney Institute for Astronomy (SIfA), School of Physics, University of Sydney, NSW 2006, Australia

¹¹SETI Institute, 189 Bernardo Avenue, Mountain View, CA 94043, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Here we present the first data release of the *Kepler* Smear Campaign, using collateral ‘smear’ data obtained by *Kepler* to reconstruct light curves of 101 stars too bright to have been otherwise observed. We describe the pipeline developed to extract and calibrate these light curves, and show that we attain photometric precision comparable to stars ordinarily more observed in the nominal *Kepler* mission. In this Paper, we focus in particular on a subset of these consisting of 60 red giants for which we detect solar-like oscillations. Using high-resolution spectroscopy from the Tillinghast Reflector Échelle Spectrograph (TRES) together with asteroseismic modelling, we constrain the masses and evolutionary states of these benchmark red giants. All source code, light curves, TRES spectra, and asteroseismic and stellar parameters are publicly available as a *Kepler* legacy sample.

Key words: asteroseismology – techniques: photometric – stars: variable: general

1 INTRODUCTION

The *Kepler* Space Telescope, operated by NASA, was launched in 2009 to obtain photometry of hundreds of thousands of stars in a field in Cygnus-Lyra, in order to detect a statistically-useful sample of transiting exoplanets (Borucki et al. 2010). It achieved this primary goal, showing that exoplanets are common around Sun-like stars (Fressin et al. 2013; Petigura et al. 2013; Foreman-Mackey et al. 2014), though with the failure of two reaction wheels, the mission was cut short and there remain substantial uncertainties on these estimates. *Kepler* was revived as a two-wheeled mission, K2, with its third axis balanced against solar radiation pressure. K2 is therefore constrained to point in the ecliptic plane, which it surveys in a succession of ~ 80 day Campaigns. In this paper, we will deal exclusively with data from the nominal *Kepler* mission before this change.

Beyond searching for planets, *Kepler* has revolutionized the field of asteroseismology (Gilliland et al. 2010). It has yielded

the first detection of gravity-mode period spacings in a red giant (Beck et al. 2011), enabling probes of interior rotation of red giants (Beck et al. 2012) and distinguishing between hydrogen- and helium-burning cores (Bedding et al. 2011). It has also permitted the determination of ages and fundamental parameters of main-sequence stars (Silva Aguirre et al. 2013), including planet-hosting stars (Huber et al. 2013; Silva Aguirre et al. 2015; Van Eylen et al. 2018), revealing the most ancient known planetary system, dating back to the earliest stages of the galaxy (Campante et al. 2015). By comparing asteroseismic stellar ages to stellar rotation periods, Angus et al. (2015) have shown that gyrochronology models cannot fit the data with a single relation, leading van Saders et al. (2016) to suggest a qualitative change in dynamo mechanism as stars age through the main sequence.

* E-mail: benjamin.pope@nyu.edu

A major outcome of the *Kepler* asteroseismology programme is a legacy sample of extremely well characterized stars which can serve as benchmarks for future work (Lund et al. 2016; Silva Aguirre et al. 2016). As well as asteroseismology, by also using optical interferometry, it has been possible to determine fundamental parameters of main-sequence and giant stars with unprecedented precision (Huber et al. 2012; White et al. 2013, 2015). Likewise by combining with spectroscopy, Hawkins et al. (2016c) have been able to produce a large sample of stars with precise elemental abundances by fitting spectroscopic data with $\log g$ and T_{eff} fixed to asteroseismically-determined values. It is necessary to calibrate such a study against benchmark stars with very precisely-determined parameters, which in practice means requires nearby bright stars that are amenable to very high signal-to-noise spectroscopy plus asteroseismology (Creevey et al. 2013), parallaxes (Hawkins et al. 2016a), and/or interferometry (Casagrande et al. 2014; Creevey et al. 2015). This is especially important in the context of the *Gaia* mission (Gaia Collaboration et al. 2016), which has recently put out its second data release of 1,692,919,135 sources, including 1,331,909,727 with parallaxes (Gaia Collaboration et al. 2018). These data will form the basis of many large surveys and it is vital that they are calibrated correctly. To this end, 34 FGK stars have been chosen as *Gaia*-ESO benchmark stars for which metallicities (Jofré et al. 2014), effective temperatures and surface gravities (Heiter et al. 2015), and relative abundances of α and iron-peak elements (Jofré et al. 2015) have been determined. This has been accompanied by the release of high resolution spectra (Blanco-Cuaresma et al. 2014) and formed the basis of extensions to lower metallicities (Hawkins et al. 2016b), stellar twin studies (Jofré 2016) and comparisons of stellar abundance determination pipelines (Jofré et al. 2017).

Brighter *Kepler* stars are therefore ideal benchmark targets, as photometry can be most easily complemented by *Hipparcos* parallaxes, interferometric diameters, and high resolution spectroscopy. Unfortunately, the *Kepler* field was deliberately placed to minimize overall the number of saturated stars, so that only a dozen stars brighter than 6th magnitude landed on silicon (Koch et al. 2010). This was because stars brighter than $Kp \sim 11$ saturate the CCD detector, spilling electrons up and down their column on the CCD and rendering these pixels otherwise unusable. Furthermore, due to the limited availability of bandwidth to download data from the satellite, only a fraction **What fraction?** of pixels on the *Kepler* detector are actually downloaded, these being allocated via a competitive proposal process. The result of these two target selection constraints is that photometry was obtained for only **a small number** of saturated stars in the *Kepler* field, while many bright targets were ignored.

Kolodziejczak & Caldwell (2011) noted that there is a way to obtain photometry of every target on-silicon in *Kepler* using a data channel normally used for calibration, even if active pixels were not allocated and downloaded. *Kepler* employs an inter-line transfer CCD as its detector, which successively shuffles each row of pixels down to the edges of the chip where they are ultimately read out. Because the *Kepler* camera lacks a shutter, the detector is exposed to light during the readout process, with the result that fluxes in each pixel are biased up by light collected from objects in the same column. This is a particularly serious issue for faint objects in the same detector column as brighter stars, and it is important to calibrate this at each readout stage. Six rows of blank ‘masked’ pixels are allocated in each column to measure the smear bias; furthermore, six ‘virtual’ rows are recorded at the end of the readout, with the result that twelve rows of pixels sample the smear bias in each column. Kolodziejczak & Caldwell (2011) realized that

these encode the light curves of bright targets in a 1D projection of the star field. The masked and virtual smear registers each receive $\sim 1/1034$ of the incident flux in each column; if this is dominated by the light from a single star, the flux combining both smear registers is equivalent to that of a star ~ 6.8 times fainter.

In Pope et al. (2016), we demonstrated a method for extracting precise light curves of bright stars in *Kepler* and K2, and presented light curves of a small number of variable stars as examples to illustrate this method. In this Paper we present light curves of all unobserved or significantly under-observed stars brighter than $V = 8$ in the *Kepler* field. This sample is biased towards red giants and hot stars, containing only a few FG dwarfs. We find no transiting planets, but detect **M** new eclipsing binaries, and solar-like oscillations in **N** red giants. We do not model hot stars or FG dwarfs in great detail, but provide some discussion and initial classification of interesting variability. For eclipsing binaries, we present the results of light-curve modelling to precisely determine their parameters. Finally, for the oscillating red giants, which constitute the bulk of the sample, we determine the asteroseismic parameters ν_{max} and $\Delta\nu$, and therefore stellar masses and $\log g$ measurements; and we obtain high-resolution spectroscopy with the Tillinghast Reflector Échelle Spectrograph (TRES), from whose spectra we derive stellar parameters and elemental abundances constrained by asteroseismic parameters. We discuss the potential for these as benchmark stars for other stellar surveys, in particular *Gaia*.

We have made all new data products and software discussed in this paper publicly available, and encourage interested readers to use these in their own research.

2 METHOD

In this Section we will discuss the methods used for characterizing our new benchmark stars. We have obtained smear light curves for our sample of red giant stars with the `keplersmear` pipeline as described in Section 2.1, performed asteroseismology on all of these to extract ν_{max} and therefore $\log g$ as described in Section 3, and combined these with TRES spectra to obtain chemical abundances as described in Section 3.1.

2.1 Photometry

We selected as our sample all stars on-silicon in *Kepler* with $Kp < 8$ which were unobserved for more than 10 quarters **Tim: what was your cutoff in quarters for ‘underobserved’ stars?**, including those stars which were entirely unobserved. A number of these lay just at the edge of a detector, with the result that in some cadences the centroid of the star did not lie on the chip; light curves from these targets were found to be of extremely low quality and all of these objects were discarded. After applying these criteria we obtained a list of **101** targets. Aside from the restriction on stars falling on the edge of a chip, the addition of these objects to conventionally-observed stars makes the *Kepler* survey magnitude-complete down to $Kp = 8$.

In preparing light curves of the *Kepler* smear stars, we follow the methods described in Pope et al. (2016), with some improvements. We select using RA and Dec values from the *Kepler* Input Catalog (KIC) (Brown et al. 2011), and query MAST to find the corresponding mean pixel position for a given *Kepler* quarter. We measure the centroid of smear columns in the vicinity, and use these values to do raw aperture photometry. We find that the cosine-bell aperture used for raw photometry in Pope et al. (2016) can

in some light curves introduce position-dependent systematics and jumps. We instead in this work apply a super-Gaussian aperture,

$A \propto \exp \frac{-(x-x_0)^4}{w}$, where x_0 is the centroid and w a width in pixels. The very flat top of this function helps avoid significant variation with position, while still smoothly rolling off at the edges to avoid discontinuous artefacts. We calculate this on a grid of $10\times$ subsampled points in time so that the sharply varying edge changes column weights smoothly as a function of centroid. We extract photometry using apertures with a range of widths $w \in \{1.5, 2, 3, 4, 5\}$ pixels.

From this raw photometry we subtract a background light curve, which corrects for time-varying global systematics. Whereas in Pope et al. (2016) we then subtract a background estimate chosen manually, for this larger set of light curves, we now choose the lowest 25% of pixels by median flux as being unlikely to be contaminated by stars, and take our background level to be the median of this at each time sample. To denoise this, we fit a Gaussian Process with a 30-day timescale squared exponential kernel using GEORGE (Ambikasaran et al. 2014), and our final background light curve is taken to be the posterior mean of this GP.

The dominant source of residual systematic errors in nominal *Kepler* time series is a common-mode variation primarily due to thermal changes on board the spacecraft, an issue which is traditionally dealt with by identifying and fitting a linear combination of systematic modes (Twicken et al. 2010; Stumpe et al. 2012; Smith et al. 2012; Petigura & Marcy 2012). We adopt the same approach here, using the *Kepler* Pre-search Data Conditioning (PDC) Cotrending Basis Vectors (CBVs) available from MAST, finding least-squares fits of either the first 4 or 8 CBVs to each light curve. We note that this can subtract astrophysical signals on long timescales, such that we use and recommend 4 CBV light curves for stars with variability on timescales longer than ~ 5 days, but otherwise use the 8 CBV light curves. There is some room for improvement here by simultaneously modelling astrophysical and instrumental variations, but this is beyond the scope of this paper. In the following, we will use the light curves with the lowest 6.5 hr Combined Differential Photometric Precision (CDPP) (Christiansen et al. 2012) out of all apertures, as calculated with the $\kappa 2sc$ implementation (Aigrain et al. 2016). **This is primitive - how do we justify this? We only did it because we didn't have anything cleverer, but now we do. Should I redo this?.**

2.2 Asteroseismology

For all N red giants identified in this sample, we have attempted to extract the asteroseismic parameters ν_{\max} and $\langle \Delta\nu \rangle$. These constrain fundamental stellar parameters independently from spectroscopic or interferometric measurements:

$$\nu_{\max} \propto \frac{g}{g_{\odot}} \cdot \frac{T_{\text{eff}}}{T_{\text{eff}\odot}} \frac{1}{2} \quad (1)$$

and

$$\langle \Delta\nu \rangle \propto \sqrt{\langle \rho \rangle} = \sqrt{\frac{M}{M_{\odot}} \left(\frac{R}{R_{\odot}} \right)^{-3}} \quad (2)$$

We follow the method of Davies & Miglio (2016), obtaining a Lomb-Scargle periodogram of the smoothed time series according to the method of García et al. (2011). We then conduct a Markov

Chain Monte Carlo fit to this, applying the combined granulation and oscillation model of Kallinger et al. (2014), consisting of two Harvey profiles for the granulation (Harvey 1985), a Gaussian envelope for the stellar oscillations, and a white noise background for instrumental noise. We find that the marginal posterior distribution for the Gaussian envelope is well-approximated by a single Gaussian, and take its median and standard deviation to be our estimates for ν_{\max} and its uncertainty.

To estimate $\Delta\nu$, we divide the power spectrum through by the granulation and noise models to obtain a signal-to-noise spectrum, and fit a sum of Lorentzians separated by mean large ($\Delta\nu$) and small ($\delta\nu$) separations to the part of this spectrum in the vicinity of ν_{\max} . In practice, for this dataset, $\delta\nu$ is poorly constrained, but mean $\langle \Delta\nu \rangle$ is typically well-constrained and its posterior marginal distribution is well-represented by a single Gaussian as with ν_{\max} .

We obtain good estimates of these asteroseismic parameters for 35 targets. In the remainder of cases, we find that the very-low-frequency ($\lesssim 2\mu\text{Hz}$) oscillations are affected by filter artefacts from detrending, and we are not able to obtain good estimates for these stars. **Should we try with these?**

3 ASTEROSEISMOLOGY

Guy - can you describe your pipeline here?

3.1 Spectroscopy

For the whole red giant sample, we have obtained high-resolution spectroscopy with TRES in order to constrain stellar parameters and elemental abundances. Operating with spectral resolving power $R = 44000$, we obtain signal-to-noise ratios of **something** per resolution element.

4 RESULTS

4.1 Red Giants

4.2 Other Stars

5 OPEN SCIENCE

We believe in open science, and have therefore made all substantive products of this research available to the interested reader. All code used to produce smear light curves is available under a GPL v3 license at github.com/benjaminpope/keplersmear. All smear light curves, both including the red giant sample studied in detail in Section 4.1, and other stars as discussed in Section 4.2, can be downloaded from the Mikulski Archive for Space Telescopes (MAST) as a High-Level Science Product. TRES spectra are available from **somewhere**, and all asteroseismic parameters and derived stellar parameters for the red giants in Section 4.1 are provided in an online-only table as Supplementary Material to this paper.

All smear light curves in this paper, as well as the L^AT_EX source code used to produce this document, can be found at github.com/benjaminpope/smearcampaign.

6 CONCLUSIONS

ACKNOWLEDGEMENTS

This work was performed in part under contract with the Jet Propulsion Laboratory (JPL) funded by NASA through the Sagan Fel-

lowship Program executed by the NASA Exoplanet Science Institute. B.P. also acknowledges support from Balliol College and the Clarendon Fund. D.H. acknowledges support by the Australian Research Council's Discovery Projects funding scheme (project number DE140101364) and support by the NASA Grant NNX14AB92G issued through the *Kepler* Participating Scientist Program.

BP acknowledges being on the traditional territory of the Lenape Nations and, today, we recognize that Manhattan continues to be the home to many Algonkian peoples. We thank the Lenape peoples for allowing us to carry out this work on the Lenape original homelands at New York University. BP and TW would like to acknowledge the Gadigal people of the Eora Nation and the Norongerragal and Gweagal peoples of the Tharawal Nation as the traditional owners of the land at the University of Sydney and the Sutherland Shire on which some of this work was carried out, and pay their respects to their knowledge, and their elders past, present and future.

This research made use of NASA's Astrophysics Data System; the SIMBAD database, operated at CDS, Strasbourg, France; the IPython package (Pérez & Granger 2007); SciPy (Jones et al. 2001); and Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013). Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX13AC07G and by other grants and contracts. We acknowledge the support of the Group of Eight universities and the German Academic Exchange Service through the Go8 Australia-Germany Joint Research Co-operation Scheme.

REFERENCES

- Aigrain S., Parviainen H., Pope B. J. S., 2016, *MNRAS*, **459**, 2408
- Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2014
- Angus R., Aigrain S., Foreman-Mackey D., McQuillan A., 2015, *MNRAS*, **450**, 1787
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Beck P. G., et al., 2011, *Science*, **332**, 205
- Beck P. G., et al., 2012, *Nature*, **481**, 55
- Bedding T. R., et al., 2011, *Nature*, **471**, 608
- Blanco-Cuadras S., Soubiran C., Jofré P., Heiter U., 2014, *A&A*, **566**, A98
- Borucki W. J., et al., 2010, *Science*, **327**, 977
- Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, *AJ*, **142**, 112
- Campante T. L., et al., 2015, *ApJ*, **799**, 170
- Casagrande L., et al., 2014, *MNRAS*, **439**, 2060
- Christiansen J. L., et al., 2012, *PASP*, **124**, 1279
- Creevey O. L., et al., 2013, *MNRAS*, **431**, 2419
- Creevey O. L., et al., 2015, *A&A*, **575**, A26
- Davies G. R., Miglio A., 2016, *Astronomische Nachrichten*, **337**, 774
- Foreman-Mackey D., Hogg D. W., Morton T. D., 2014, *ApJ*, **795**, 64
- Fressin F., et al., 2013, *ApJ*, **766**, 81
- Gaia Collaboration et al., 2016, *A&A*, **595**, A1
- Gaia Collaboration Brown A. G. A., Vallenari A., Prusti T., de Bruijne J. H. J., Babusiaux C., Bailer-Jones C. A. L., 2018, preprint, ([arXiv:1804.09365](https://arxiv.org/abs/1804.09365))
- García R. A., et al., 2011, *MNRAS*, **414**, L6
- Gilliland R. L., et al., 2010, *PASP*, **122**, 131
- Harvey J., 1985, in Rolfe E., Battrick B., eds, ESA Special Publication Vol. 235, Future Missions in Solar, Heliospheric & Space Plasma Physics.
- Hawkins K., et al., 2016a, *A&A*, **592**, A70
- Hawkins K., Masseron T., Jofré P., Gilmore G., Elsworth Y., Hekker S., 2016c, *A&A*, **594**, A43
- Heiter U., Jofré P., Gustafsson B., Korn A. J., Soubiran C., Thévenin F., 2015, *A&A*, **582**, A49
- Huber D., et al., 2012, *ApJ*, **760**, 32
- Huber D., et al., 2013, *ApJ*, **767**, 127
- Jofré P., 2016, *Astronomische Nachrichten*, **337**, 859
- Jofré P., et al., 2014, *A&A*, **564**, A133
- Jofré P., et al., 2015, *A&A*, **582**, A81
- Jofré P., et al., 2017, *A&A*, **601**, A38
- Jones E., Oliphant T., Peterson P., Others 2001, SciPy: Open source scientific tools for Python, <http://www.scipy.org/>
- Kallinger T., et al., 2014, *A&A*, **570**, A41
- Koch D. G., et al., 2010, *ApJ*, **713**, L79
- Kolodziejczak J., Caldwell D., 2011, Technical Report 20120003045, Science from Kepler Collateral Data: 150 ksec/year from 13 Million Stars?, <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120003045.pdf>. NASA Marshall Space Flight Centre, <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120003045.pdf>
- Lund M. N., et al., 2016, preprint, ([arXiv:1612.00436](https://arxiv.org/abs/1612.00436))
- Pérez F., Granger B. E., 2007, *Computing in Science and Engineering*, **9**, 21
- Petigura E. A., Marcy G. W., 2012, *PASP*, **124**, 1073
- Petigura E. A., Howard A. W., Marcy G. W., 2013, *Proceedings of the National Academy of Science*, **110**, 19273
- Pope B. J. S., et al., 2016, *MNRAS*, **455**, L36
- Silva Aguirre V., et al., 2013, *ApJ*, **769**, 141
- Silva Aguirre V., et al., 2015, *MNRAS*, **452**, 2127
- Silva Aguirre V., et al., 2016, preprint, ([arXiv:1611.08776](https://arxiv.org/abs/1611.08776))
- Smith J. C., et al., 2012, *PASP*, **124**, 1000
- Stumpe M. C., et al., 2012, *PASP*, **124**, 985
- Twicken J. D., Chandrasekaran H., Jenkins J. M., Gunter J. P., Girouard F., Klaus T. C., 2010, in Software and Cyberinfrastructure for Astronomy. p. 77401U, doi:10.1117/12.856798
- Van Eylen V., Agentoft C., Lundkvist M. S., Kjeldsen H., Owen J. E., Fulton B. J., Petigura E., Snellen I., 2018, *MNRAS*, **479**, 4786
- White T. R., et al., 2013, *MNRAS*, **433**, 1262
- White T. R., et al., 2015, in European Physical Journal Web of Conferences. p. 06068, doi:10.1051/epjconf/201510106068
- van Saders J. L., Ceillier T., Metcalfe T. S., Silva Aguirre V., Pinsonneault M. H., García R. A., Mathur S., Davies G. R., 2016, *Nature*, **529**, 181

This paper has been typeset from a \LaTeX file prepared by the author.