# Photometric Metallicities for Low-Mass Stars with Gaia and WISE

James. R. A. Davenport[1] and Trevor Z. Dorn-Wallenstein[1]

[1] *Department of Astronomy, University of Washington, Seattle, WA 98195, USA*

Corresponding author: James. R. A. Davenport
jrad@uw.edu

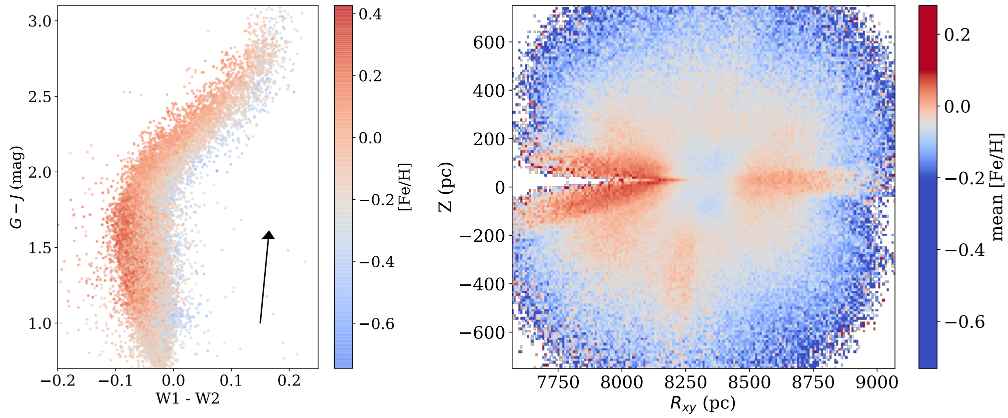


**Figure 1.** Left: Color–color training diagram for the 35,210 APOGEE stars colored by their measured metallicities. This sample is used to train our KNN algorithm `ingot`. A reddening vector from Sanders & Das (2018) for 1 mag of $V$-band extinction is shown for reference. Right: Average metallicity in our 3 million star ALLWISE-Gaia sample projected on cylindrical ($R_{xy}$, Z) coordinates. The metal-rich Galactic disk is clearly seen. The apparent break in the disk near the solar position is due to sample incompleteness of high proper-motion objects.

Multi-object spectroscopic surveys — e.g., the Apache Point Observatory Galactic Evolution Experiment (APOGEE, Majewski et al. 2017) — have increased the sample of stars with well-measured metallicities by orders of magnitude, but still don't reach brightness limits or sample sizes comparable with photometric surveys. While less precise than spectroscopic measurements, "photometric metallicities" from visible and infrared surveys can explore the large scale chemical structure and evolution of our Galaxy (e.g. Ivezić et al. 2008; Schmidt et al. 2016), though their use has so far been largely limited to legacy survey footprints.

Thanks to the Gaia mission (Gaia Collaboration et al. 2018), photometric metallicities are now possible for field stars across the entire sky. Here we present `ingot` (Davenport & Dorn-Wallenstein 2019), a $k$–nearest neighbors (KNN) tool to estimate stellar [Fe/H] using Gaia and WISE photometry for low-mass stars. We demonstrate this capability by estimating metallicities for three million cool stars, and advocate for more detailed explorations of this technique using WISE and Gaia data.

Following Figure 5 of Schmidt et al. (2016), we trained `ingot` on [Fe/H] as a function of ($W1-W2$, $G-J$), shown in Figure 1. We used 35,210 stars from the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASCAP, García Pérez et al. 2016), cross-matched to ALLWISE (Mainzer et al. 2014) and Gaia DR2 using the CDS X-Match service with a 1 arcsec search radius. The $J$-band magnitude comes from 2MASS (Skrutskie et al. 2006), and is pre-matched to WISE. $G-J$ (Gaia - 2MASS) spans a wide wavelength range, and is a good proxy for stellar effective temperature. As Schmidt et al. (2016) note, $W1 - W2$ shows a small amplitude gradient that correlates with [Fe/H] (∼1 dex in [Fe/H] over ∼0.1 mag in color).

To represent our data with a flexible model, we used `scikit-learn`'s $k$-nearest neighbors (KNN) regression, with the default neighbor distance of $k = 5$. This model can rapidly estimate [Fe/H] values for any new star given ($W1 - W2$, $G - J$), and be easily recomputed given additional axes including the absolute Gaia magnitude ($M_G$). Crude uncertainties can be computed by examining the standard deviation of the measured versus predicted [Fe/H] values, which found typical scatter of $\sigma \approx 0.11$ dex in our training sample. This does not account for photometric errors, nor correct for extinction in any of the bands.

We applied `ingot` to 3.8 million new low-mass stars selected from Gaia and WISE photometry for stars within the color–color box defined in Figure 1, reaching ∼600 pc over the entire sky. The mean [Fe/H] for these sources projected into the $(R_{xy}, Z)$ plane is shown in Figure 1. The metal-rich disk of the Milky Way, including the decreasing scale-height with increasing radius, is clearly recovered.

`ingot` is publicly available[1], including the data required to recreate our training sample. We consider `ingot` a demonstration of the potential for simple machine learning tools combined with multi-wavelength data from wide-field surveys to advance galactic chemical cartography. For example, the catalog of over 2 billion point sources from the unWISE project (Schlafly et al. 2019), matched to Gaia DR2, is an ideal dataset for such analysis.

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`https://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `https://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

This publication makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration.

*Software:* Python, IPython (Pérez & Granger 2007), NumPy (Oliphant 2007), Matplotlib (Hunter 2007), SciPy (Jones et al. 2001–), Pandas (McKinney 2010), Astropy (Astropy Collaboration et al. 2013), Scikit-Learn (Pedregosa et al. 2011)

## REFERENCES

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

[1] https://github.com/jradavenport/ingot

Davenport, J., & Dorn-Wallenstein, T. Z. 2019, jradavenport/ingot: Ingot: a demo, doi: 10.5281/zenodo.2600094. https://doi.org/10.5281/zenodo.2600094

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, ArXiv e-prints. https://arxiv.org/abs/1804.09365

García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, AJ, 151, 144, doi: 10.3847/0004-6256/151/6/144

Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90

Ivezić, Ž., Sesar, B., Jurić, M., et al. 2008, ApJ, 684, 287, doi: 10.1086/589678

Jones, E., Oliphant, T., Peterson, P., et al. 2001–, SciPy: Open source scientific tools for Python

Mainzer, A., Bauer, J., Cutri, R. M., et al. 2014, ApJ, 792, 30

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94, doi: 10.3847/1538-3881/aa784d

McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51 – 56

Oliphant, T. E. 2007, Computing in Science Engineering, 9, 10

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Pérez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21

Sanders, J. L., & Das, P. 2018, ArXiv e-prints. https://arxiv.org/abs/1806.02324

Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, ApJS, 240, 30, doi: 10.3847/1538-4365/aafbea

Schmidt, S. J., Wagoner, E. L., Johnson, J. A., et al. 2016, MNRAS, 460, 2611, doi: 10.1093/mnras/stw1139

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708