The Journal of Open Source Software

PyIRD: A Python-Based Data Reduction Pipeline for Subaru/IRD and REACH

Yui Kasagi [1], **Hajime Kawahara** [1,2], **Ziying Gu**[2], **Teruyuki Hirano** [3,4,5], **Takayuki Kotani** [3,4,5], **Masayuki Kuzuhara** [3,5], **and Kento Masuda** [6]

1 Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, 3-1-1 Yoshinodai, Chuo-ku, Sagamihara, Kanagawa, 252-5210, Japan **2** Department of Astronomy, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan **3** Astrobiology Center, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan **4** Astronomical Science Program, The Graduate University for Advanced Studies, SOKENDAI, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan **5** National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan **6** Department of Earth and Space Science, Osaka University, Toyonaka, Osaka 560-0043, Japan

DOI: 10.21105/joss.09126

Software

- Review
- Repository
- Archive

Editor: Arfon Smith

Reviewers:

- @crhea93
- @shihyuntang
- @profxj

Submitted: 07 February 2025
Published: 17 January 2026

Summary

PyIRD is a Python-based pipeline for reducing spectroscopic data obtained with IRD (InfraRed Doppler; Kotani et al. (2018)) and REACH (Rigorous Exoplanetary Atmosphere Characterization with High dispersion coronagraphy; Kotani et al. (2020)) on the Subaru Telescope. It is designed to process raw images into one-dimensional spectra in a semi-automatic manner. Unlike traditional methods, it does not rely on IRAF (Tody, 1986, 1993), a software used for astronomical data reduction. This approach simplifies the workflow while maintaining efficiency and accuracy. Additionally, the pipeline includes an updated method for removing readout noise patterns from raw images, enabling efficient extraction of spectra even for faint targets such as brown dwarfs.

Statement of need

The reduction of high-dispersion spectroscopic data has traditionally been performed using IRAF, one of the most widely used tools for astronomical data reduction and analysis. Although the National Optical Astronomy Observatory (NOAO) officially ceased development and maintenance of IRAF in 2013, community-based maintenance has continued. However, the official IRAF community distribution[1] and the Space Telescope Science Institute (STScI)[2] have both recommended that researchers transition away from IRAF due to its legacy architecture and lack of institutional support.

In recent years, several open-source, Python-based pipelines for the reduction of near-infrared echelle spectrographs have been developed. Some pipelines utilize PyRAF, a Python interface to IRAF, such as WARP for the WINERED spectrograph (Hamano et al., 2024), while others, including PLP for IGRINS (Sim et al., 2014) and PypeIt (Prochaska et al., 2020), do not rely on PyRAF or IRAF-based components. While these pipelines provide either general frameworks or instrument-specific solutions, PyIRD is designed to offer a simple pipeline optimized for IRD and REACH data reduction. Furthermore, recent advances combining adaptive optics with these instruments have enabled high-dispersion spectroscopic observations of faint companions

[1] IRAF Community Distribution website: https://iraf-community.github.io/

[2] STScI Newsletter (2018), “Removing the Institute’s Dependence on IRAF: You Can Do It Too”: https://www.stsci.edu/contents/newsletters/2018-volume-35-issue-03/removing-the-institutes-dependence-on-iraf-you-can-do-it-too

orbiting bright main-sequence stars. To support such observations, PyIRD implements improved detector noise reduction to extract high-quality spectra from these faint targets.

Together, these developments underscore the need for actively maintained, scalable, and flexible software for high-dispersion spectroscopic data reduction. PyIRD addresses this need by providing a modern, Python-based pipeline and has already been utilized in several studies (Kasagi et al., 2025; Kawahara et al., 2025; Kawashima et al., 2025; Tomoyoshi et al., 2024).

Key Features

PyIRD performs semi-automatic data reduction by following a general workflow for high-dispersion spectroscopic data, as illustrated in Figure 1. It primarily utilizes Astropy (Astropy Collaboration et al., 2022), NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), and pandas (The pandas development team, 2020).

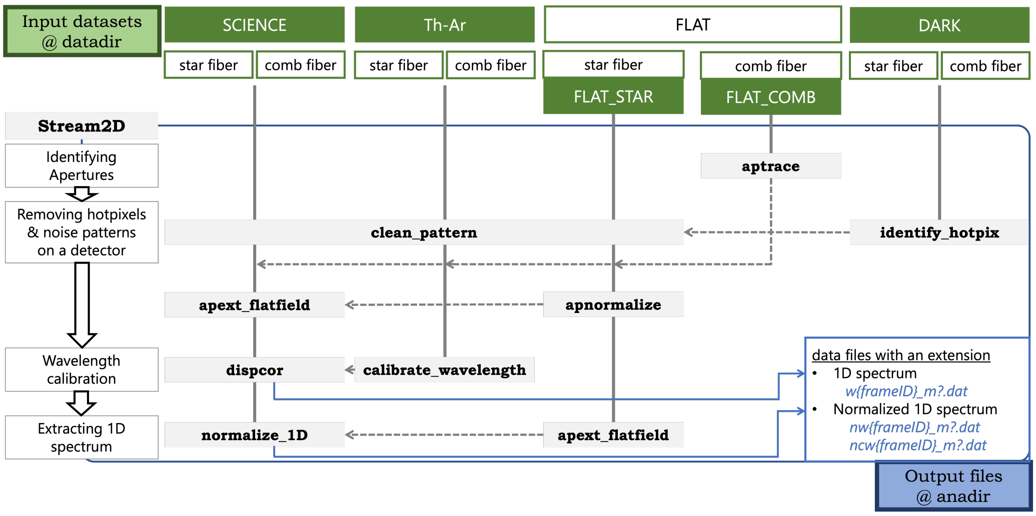

Figure 1: Flowchart of the reduction process for IRD and REACH data. The reduction process follows from top to bottom of this figure. Text in the grey boxes represents the instance names of each reduction step used in PyIRD.

To simplify the handling of a large number of input FITS-format files, PyIRD introduces a Python class called FitsSet. Once initialized with parameters such as the file IDs and the directory containing those files, FitsSet automatically organizes and manages the input files and their metadata. It also allows users to apply reduction functions collectively to a specified list of FITS IDs, enabling efficient and consistent data processing through the Stream2D class.

Since all functions in PyIRD are written in Python rather than IRAF's Subset Preprocessor Language (SPP), the package is easy to develop and maintain. This also significantly reduces the time required for the reduction process: users only need to execute a single Python script without complex IRAF configuration. For example, reducing data with PyIRD typically takes several tens of minutes to produce one-dimensional spectra from raw data obtained during a single observing night, compared to approximately half a day with traditional IRAF methods.

Moreover, PyIRD achieves a higher level of readout noise pattern removal in the final reduced data. This feature is particularly important for processing data from faint objects such as brown dwarfs, where the astronomical signal is often comparable in strength to systematic noise. The dominant noise source is the readout pattern from the H2RG detector used in IRD. To address this, PyIRD models the noise by calculating a median profile for each readout channel and applying a 2D Gaussian Process using gpkron (Kawahara, 2022). This innovative method effectively mitigates the readout pattern, as shown in Figure 2, and improves data quality for faint targets.

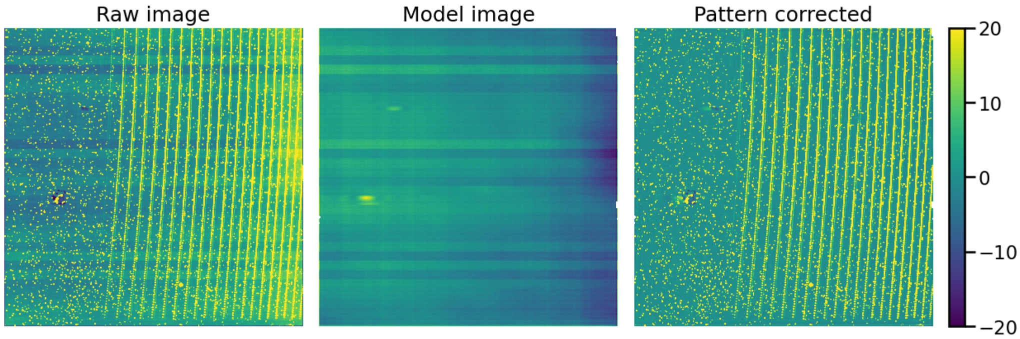

Figure 2: (Left) Raw image; (Middle) Readout pattern model created by `PyIRD`; (Right) Pattern-corrected image

Acknowledgements

Y.K. acknowledges support from JST SPRING, Grant Number JPMJSP2104 and JSPS KAKENHI grant No. 24K22912. Z.G. acknowledges support from Forefront Physics and Mathematics Program to Drive Transformation (FoPM), a World-leading Innovative Graduate Study (WINGS) Program, the University of Tokyo.

References

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., Earl, N., Starkman, N., Bradley, L., Shupe, D. L., Patil, A. A., Corrales, L., Brasseur, C. E., N"othe, M., Donath, A., Tollerud, E., Morris, B. M., Ginsburg, A., Vaher, E., Weaver, B. A., Tocknell, J., Jamieson, W., … Astropy Project Contributors. (2022). The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. *The Astrophysical Journal*, *935*(2), 167. https://doi.org/10.3847/1538-4357/ac7c74

Hamano, S., Ikeda, Y., Otsubo, S., Katoh, H., Fukue, K., Matsunaga, N., Taniguchi, D., Kawakita, H., Takenaka, K., Kondo, S., & Sameshima, H. (2024). WARP: The Data Reduction Pipeline for the WINERED Spectrograph. *Publications of the Astronomical Society of the Pacific*, *136*(1), 014504. https://doi.org/10.1088/1538-3873/ad1b38

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., R'ıo, J. F. del, Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Kasagi, Y., Kawashima, Y., Kawahara, H., Kotani, T., Masuda, K., Ahn, K., Guyon, O., Hirano, T., Jovanovic, N., Kuzuhara, M., Lozi, J., Tamura, M., Uyama, T., Vievard, S., & Yoneta, K. (2025). Unveiling the Atmosphere of HR 7672 B from the Near-infrared High-resolution Spectrum Using REACH/Subaru. *The Astronomical Journal*, *170*(4), 211. https://doi.org/10.3847/1538-3881/adf730

Kawahara, H. (2022). Gpkron. In *GitHub repository*. GitHub. https://github.com/HajimeKawahara/gpkron

Kawahara, H., Kawashima, Y., Tada, S., Ishikawa, H. T., Hosokawa, K., Kasagi, Y., Kotani, T., Masuda, K., Nugroho, S. K., Tamura, M., Yama, H., Kitzmann, D., Minesi, N., & Morris, B.

M. (2025). Differentiable Modeling of Planet and Substellar Atmosphere: High-resolution Emission, Transmission, and Reflection Spectroscopy with ExoJAX2. *The Astrophysical Journal*, *985*(2), 263. https://doi.org/10.3847/1538-4357/adcba2

Kawashima, Y., Kawahara, H., Kasagi, Y., Ishikawa, H. T., Masuda, K., Kotani, T., Kudo, T., Hirano, T., Kuzuhara, M., Nugroho, S. K., Livingston, J., Harakawa, H., Nishikawa, J., Omiya, M., Takarada, T., Tamura, M., & Ueda, A. (2025). Atmospheric Retrieval of Subaru/IRD High-resolution Spectrum of the Archetype T-type Brown Dwarf Gl 229 B. *The Astrophysical Journal*, *988*(1), 53. https://doi.org/10.3847/1538-4357/adddbd

Kotani, T., Kawahara, H., Ishizuka, M., Jovanovic, N., Vievard, S., Lozi, J., Sahoo, A., Guyon, O., Yoneta, K., & Tamura, M. (2020). Extremely high-contrast, high spectral resolution spectrometer REACH for the Subaru Telescope. In L. Schreiber, D. Schmidt, & E. Vernet (Eds.), *Adaptive optics systems VII* (Vol. 11448, p. 1144878). https://doi.org/10.1117/12.2561755

Kotani, T., Tamura, M., Nishikawa, J., Ueda, A., Kuzuhara, M., Omiya, M., Hashimoto, J., Ishizuka, M., Hirano, T., Suto, H., Kurokawa, T., Kokubo, T., Mori, T., Tanaka, Y., Kashiwagi, K., Konishi, M., Kudo, T., Sato, B., Jacobson, S., … Genda, H. (2018). The infrared Doppler (IRD) instrument for the Subaru telescope: instrument description and commissioning results. In C. J. Evans, L. Simard, & H. Takami (Eds.), *Ground-based and airborne instrumentation for astronomy VII* (Vol. 10702, p. 1070211). https://doi.org/10.1117/12.2311836

Prochaska, J. X., Hennawi, J. F., Westfall, K. B., Cooke, R. J., Wang, F., Hsyu, T., Davies, F. B., Farina, E. P., & Pelliccia, D. (2020). PypeIt: The python spectroscopic data reduction pipeline. *Journal of Open Source Software*, *5*(56), 2308. https://doi.org/10.21105/joss.02308

Sim, C. K., Le, H. A. N., Pak, S., Lee, H.-I., Kang, W., Chun, M.-Y., Jeong, U., Yuk, I.-S., Kim, K.-M., Park, C., Pavel, M. D., & Jaffe, D. T. (2014). Comprehensive data reduction package for the Immersion GRating INfrared Spectrograph: IGRINS. *Advances in Space Research*, *53*(11), 1647–1656. https://doi.org/10.1016/j.asr.2014.02.024

The pandas development team. (2020). *Pandas-dev/pandas: pandas* (latest). Zenodo. https://doi.org/10.5281/zenodo.3509134

Tody, D. (1993). IRAF in the Nineties. In R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (Eds.), *Astronomical data analysis software and systems II* (Vol. 52, p. 173).

Tody, D. (1986). The IRAF Data Reduction and Analysis System. In D. L. Crawford (Ed.), *Instrumentation in astronomy VI* (Vol. 627, p. 733). https://doi.org/10.1117/12.968154

Tomoyoshi, M., Masuda, K., Hirano, T., Kasagi, Y., Kawahara, H., Kotani, T., Kudo, T., Tamura, M., & Vievard, S. (2024). Weighing Single-lined Spectroscopic Binaries Using Tidal Effects on Radial Velocities: The Case of V723 Monocerotis. *The Astrophysical Journal*, *977*(2), 151. https://doi.org/10.3847/1538-4357/ad8cdd

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2