EPSC Abstracts
Vol. 18, EPSC-DPS2025-316, 2025, updated on 17 May 2026
https://doi.org/10.5194/epsc-dps2025-316
EPSC-DPS Joint Meeting 2025


# Uncertainty Quantification of Machine-Learning-Based Atmospheric Retrievals of Exoplanets fromTransmission Spectra

**Gaurav Shukla**[1], Katia Matcheva[2], Konstantin Matchev[2], and Sergei Gleyzer[2]
[1]National Institute of Science Education and Research, Homi Bhabha National Institute, India (gaurav.shukla@niser.ac.in)
[2]The University of Alabama, Tuscaloosa, USA

Transmission spectroscopy is a common tool for the characterization of transiting exoplanets and their atmospheres. Machine learning (ML) techniques are being increasingly applied to the inverse problem of determining the exoplanet parameters from observed transit spectra. A necessary ingredient in such studies is a database of synthetic spectra representing a wide class of exoplanets of interest. We focus on the issue of data uncertainty quantification and explore the performance improvement of ML retrievals with increasing the size of the training data. For this purpose, we create the largest ever publicly available database of 10 million transit spectra, for varying stellar and planet parameters and different atmospheric chemical compositions. We use the created database to train several popular ML architectures and compare the resulting performance. We provide a useful recipe for choosing the minimum size of the training database, for a given level of instrumental noise in the observations.