Estimating Atmospheric Turbulence (Cn^2) at Astronomical Sites Using Meteorological Data and Machine Learning

Edith T. Cuadros[1], Adriano P. Almeida[1], Deysi Cornejo[2], Franciele Carlesso[1], Fernando L. Guarnieri[1], Alisson Dal Lago[1], and Luis E. A. Vieira

[1]National Institute for Space Research, INPE, Brazil;
[2]Centro de Radio Astronomia e Astrofísica Mackenzie, CRAAM, Brazil

Corresponding author e-mail: edith.cuadros@inpe.br

ABSTRACT

Accurate estimation of the refractive index structure parameter (C_n^2) is fundamental for optimizing astronomical observations and adaptive optics performance. In this study, we explore the application of machine learning models to predict C_n^2 based on Differential Image Motion Monitor (DIMM) measurements using ERA5 reanalysis data provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). Three machine learning approaches were evaluated: a Multilayer Perceptron (MLP), a Long Short-Term Memory (LSTM) recurrent neural network, and the Extreme Gradient Boosting (XGBoost) algorithm. Data preprocessing was performed to temporally align the different time series and handle missing data. Model training and validation were conducted using data from Paranal Observatory, Chile. In the future, the trained models will be applied to estimate C_n^2 at the Laboratório Nacional de Astrofísica (LNA) Observatory located at Pico dos Dias, Brazil. The input-output relationship was modeled using time-synchronized datasets, where the features included ERA5-derived atmospheric parameters (temperature, pressure, relative humidity, and wind profiles), while the target variable was the DIMM-measured C_n^2. This study highlight the potential of combining reanalysis data with machine learning techniques for turbulence monitoring and prediction at astronomical sites, offering a cost-effective alternative to continuous in-situ measurements.

INTRODUCTION

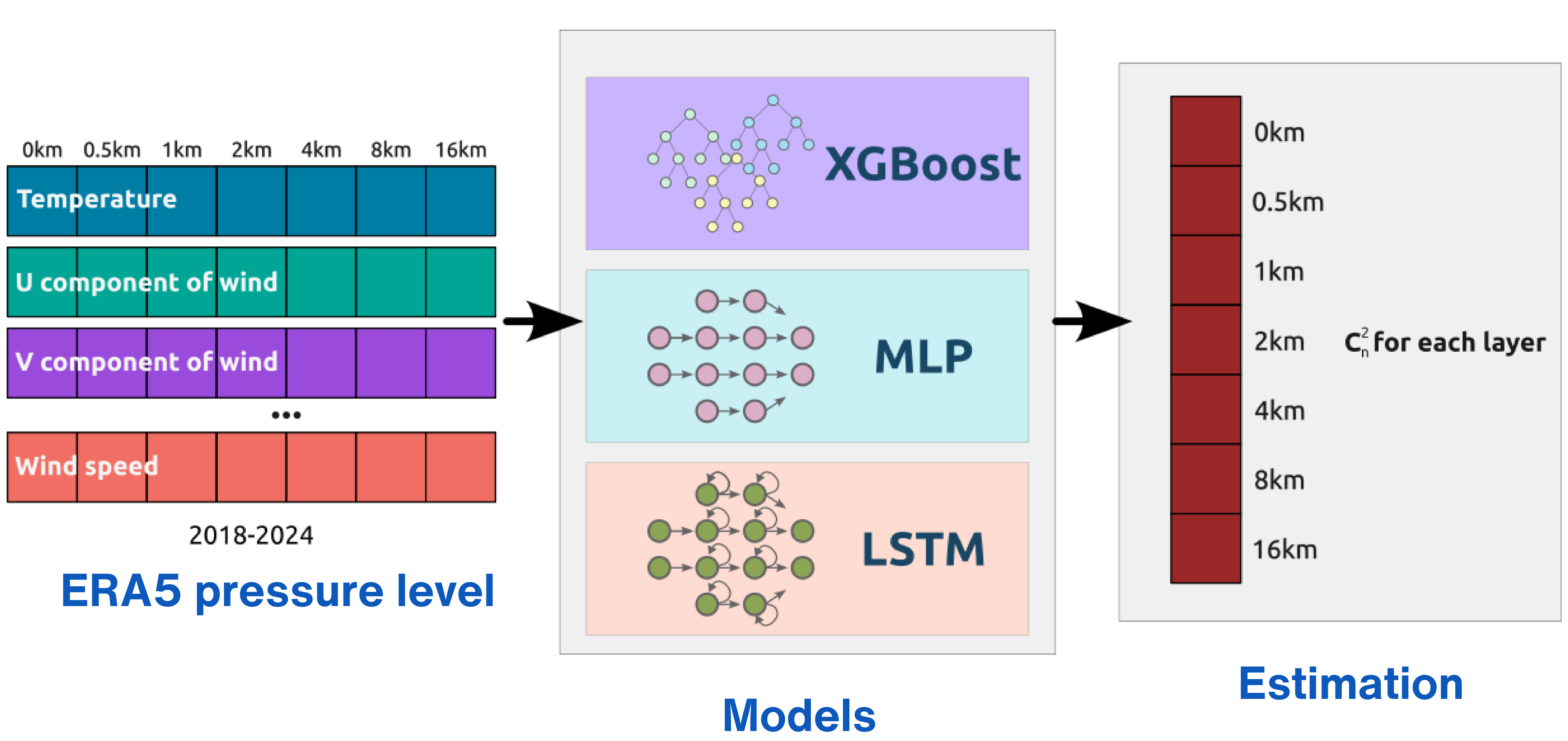
The quality of ground-based astronomical observations is affected by atmospheric turbulence, which induces spatial and temporal fluctuations in the Cn^2, degrading the optical resolution of telescopes. The intensity of this turbulence is commonly quantified by the refractive index structure parameter, Cn^2, which characterizes the vertical distribution of refractive fluctuations everywhere the atmospheric column (Tatarskii, 1971; Fried, 1966).

Traditional instruments for measuring Cn^2, such as DIMM and the MASS-DIMM, provide reliable data but require specialized infrastructure, continuous operation, and high operational costs. These restrictions limit their deployment across multiple astronomical sites (Sarazin & Roddier, 1990; Tokovinin, 2002).

In this context, machine learning (ML) techniques are being explored as alternatives for estimating Cn^2 from meteorological parameters derived from reanalysis datasets ERA5(ECMWF) (Hersbach et al., 2020). ML models are particularly effective in capturing nonlinear patterns between atmospheric variables and turbulence behavior, offering potential for near real-time operational use at reduced costs (Masciadri et al., 2017; Osborn et al., 2018).

This study aims to analyze the performance of three ML approaches, Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM) networks, and the Extreme Gradient Boosting (XGBoost) algorithm, trained with data from the Paranal Observatory (Chile) to estimate Cn^2 based on ERA5 atmospheric profiles. The trained models were subsequently applied to the LNA, to assess their generalization capabilities and applicability to new astronomical sites.

METHODOLOGY



ERA5 + MLP + LSTM + XGBoost for Cn^2 Estimation

The schematic illustrates the methodology for forecasting the refractive index structure parameter Cn^2 using ERA5 meteorological time series. The data are processed and fed into three distinct models: MLP, LSTM, and XGBoost. Each model produces Cn^2 predictions at seven specific altitudes, ranging from 0 m to 16,000 m

Using ERA5 reanalysis data with machine learning models, we explored a cost-effective alternative for estimating turbulence profiles

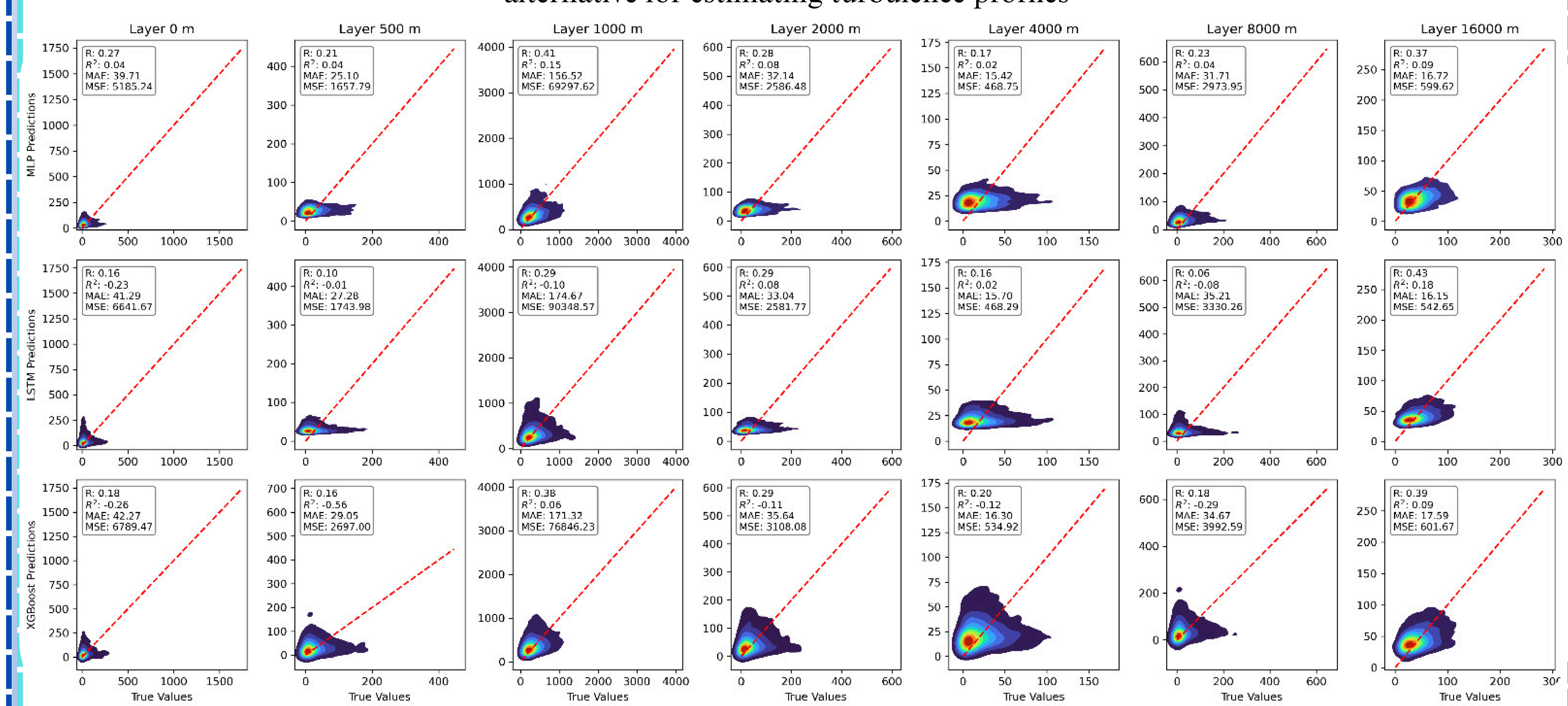


Figure1. Scatter plots of predicted vs. observed Cn^2 across atmospheric layers using MLP, LSTM, and XGBoost models. Metrics (R, R^2, MAE, MSE) show that XGBoost achieves the closest agreement with observations, particularly in the free atmosphere.

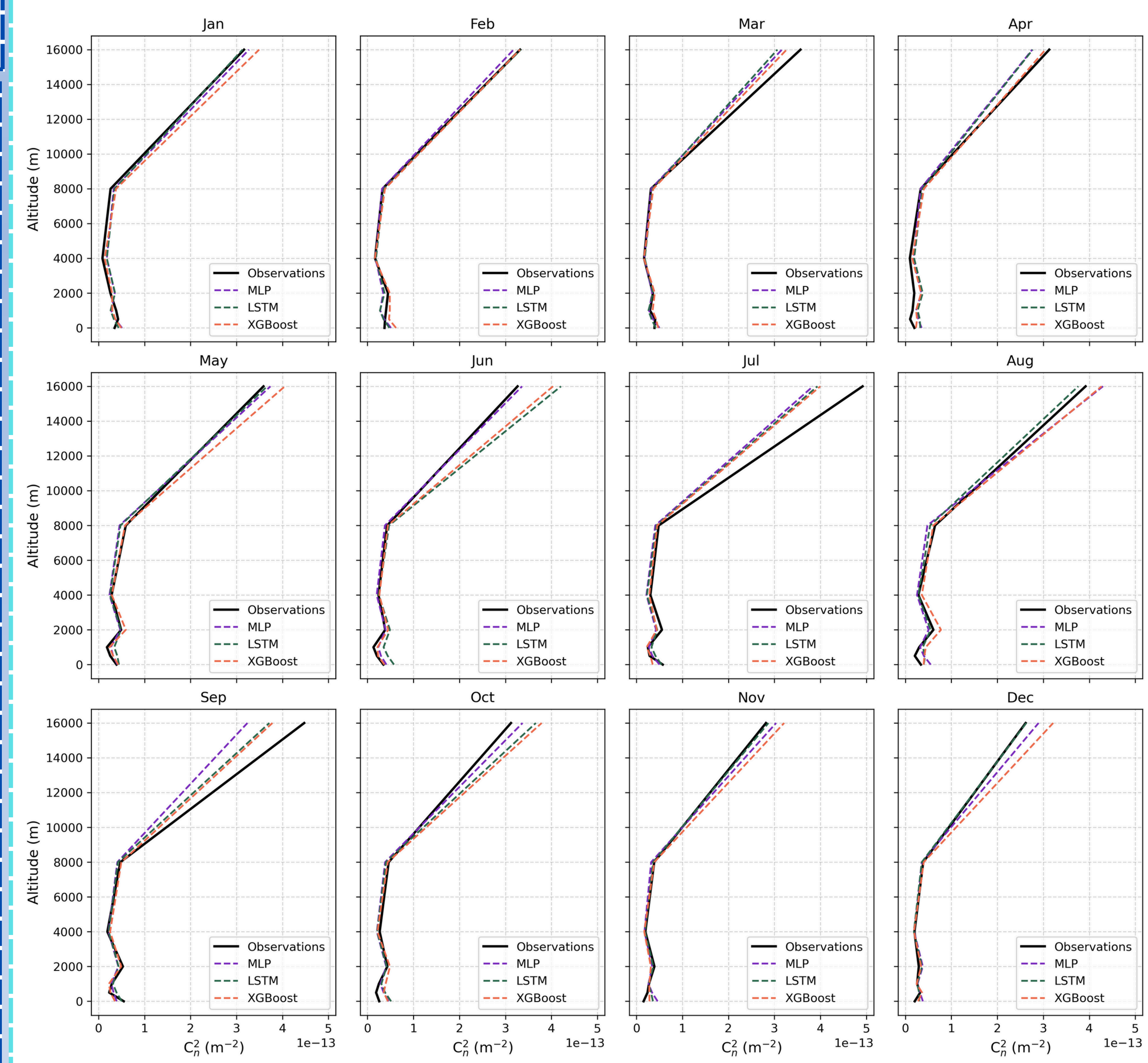


Figure 2. Monthly vertical profiles of Cn^2 at OPD, comparing observations (MASS-DIMM) with predictions from MLP, LSTM, and XGBoost. The models capture the seasonal variability, with XGBoost showing the closest agreement to observations across most months.

PRELIMINARY RESULTS

ERA5 reanalysis combined with machine learning provides an able and cost-effective means to estimate Cn^2 at astronomical sites. Among the tested models, XGBoost demonstrated the best balance of accuracy, robustness, and operational applicability. This methodology complements optical instruments, improves turbulence monitoring, and supports adaptive optics planning.

REFERENCES

Tatarskii, V. I. 1971, The effects of the turbulent atmosphere on wave propagation, *Journal: Jerusalem: Israel Program for Scientific Translations*, 1971

Fried, D. L. (1966). Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *Journal of the Optical Society of America*, **56**(10), 1372–137

Sarazin, M., & Roddier, F. (1990). The ESO differential image motion monitor. *Astronomy and Astrophysics*, **227**, 294–300

Tokovinin, A. (2002). From differential image motion to seeing. *Publications of the Astronomical Society of the Pacific*, **114**(800), 1156–1166

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Simmons, A. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**(730), 1999–2049.

Masciadri, E., Lascaux, F., & Turchi, A. (2017). Optical turbulence forecast: ready for an operational application. *Monthly Notices of the Royal Astronomical Society*, **466**(4), 520–531.

Osborn, J., Wilson, R. W., Sarazin, M., Butterley, T., & Shepherd, H. (2018). Atmospheric turbulence profiling with Stereo-SCIDAR for VLT and ELT site characterization. *Monthly Notices of the Royal Astronomical Society*, **478**(1), 825–834.

ACKNOWLEDGMENTS

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
GOVERNO FEDERAL BRASIL UNIÃO E RECONSTRUÇÃO
CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico
AEB AGÊNCIA ESPACIAL BRASILEIRA
LNA LABORATÓRIO NACIONAL DE ASTROFÍSICA UNIDADE DE PESQUISA DO MCTI
European Southern Observatory www.eso.org