

University of Groningen

## Machine learning as an ultra-fast alternative to Bayesian retrievals

Ardevol Martinez, Francisco; Min, Michiel; Kamp, Inga; Palmer, Paul I.

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2021

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Ardevol Martinez, F., Min, M., Kamp, I., & Palmer, P. I. (2021). *Machine learning as an ultra-fast alternative to Bayesian retrievals*. EPSC2021-71. Abstract from Europlanet Science Congress 2021. <https://ui.adsabs.harvard.edu/abs/2021EPSC...15..713A>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



## Machine learning as an ultra-fast alternative to Bayesian retrievals

Francisco Ardevol Martinez<sup>1</sup>, Michiel Min<sup>2</sup>, Inga Kamp<sup>1</sup>, and Paul I. Palmer<sup>3</sup>

<sup>1</sup>Kapteyn Astronomical Institute, University of Groningen, Groningen, The Netherlands

<sup>2</sup>SRON Netherlands Institute for Space Research, Utrecht, The Netherlands

<sup>3</sup>Centre for Exoplanet Science, University of Edinburgh, Edinburgh, UK

- <sup>1</sup>Kapteyn Astronomical Institute, University of Groningen, Groningen, The Netherlands
- <sup>2</sup>SRON Netherlands Institute for Space Research, Utrecht, The Netherlands
- <sup>3</sup>Centre for Exoplanet Science, University of Edinburgh, Edinburgh, UK

**Introduction:** Inferring physical and chemical properties of an exoplanet's atmosphere from its transmission spectrum is computationally expensive. A multitude of forward models, sampled from a high dimensional parameter space, need to be compared to the observation. The preferred sampling method is currently Nested Sampling [7], in particular, the MultiNest implementation [2, 3]. It typically requires tens to hundreds of thousands of forward models to converge. Therefore, simpler forward models are usually favoured over longer computation times.

A possible workaround is to use machine learning. A machine learning algorithm trained on a grid of forward models and parameter pairs can perform retrievals in seconds. This would make it possible to use complex models that take full advantage of future facilities e.g., JWST. Not only would retrievals of individual exoplanets become much faster, but it would also enable statistical studies of populations of exoplanets. It would also be a valuable tool for retrievability analyses, for example to assess the sensitivity of using different chemical networks.

The main obstacle to overcome is being able to predict accurate posterior distributions and error estimates on the retrieved parameters. These need to be as close as possible to their Bayesian counterparts.

**Methods:** Expanding on the 5-parameter grid in [5], we used ARCIS (ARtful modelling Code for exoplanet Science) [6] to generate a grid of 200,000 forward models described by the following parameters: isothermal temperature ( $T$ ), planetary radius ( $R_p$ ), planetary mass ( $M_p$ ), abundances of

water ( $\text{H}_2\text{O}$ ), ammonia ( $\text{NH}_3$ ) and hydrogen cyanide ( $\text{HCN}$ ), and cloud top pressure ( $P_{\text{cloud}}$ ). The models contain 13 wavelength bins, matching those of WASP-12b's observation with HST/WFC3 [4]. We added normally distributed random noise with  $\sigma=50$  ppm.

We trained a random forest following the details in [5] and a convolutional neural network (CNN). We divided the data into a training set of 190,000 spectra and a test set of 10,000. For the CNN we reserved 19,000 spectra (10%) from the training set for validation. These are needed to update the network weights at each training iteration.

The CNN was trained with the loss function introduced in [1] to output a probability distribution. To account for the observational noise, we combined the distributions predicted for multiple noisy copies of the spectrum.

To evaluate the performance of the machine learning algorithms, we retrieved all the spectra in the test set and plotted our predictions against the true values for the parameters. We repeated the experiment with only 1,000 spectra for Nested Sampling, reflecting the increased computational overhead of each of these retrievals. We then used a transmission spectrum of WASP-12b observed with HST/WFC3 [4] as a real-world test case.

**Results:** Although the random forest trains faster, the CNN provided better results. Figures 1 and 2 show the predicted versus the true parameters for the CNN and Nested Sampling bulk retrievals. Remarkably, we observe the same structures in both plots. This shows that the CNN is able to learn the relationship between spectral features and parameters. We also found that both the CNN and Nested Sampling provide correct error estimates, with  $\sim 60\%$  of predictions within  $1\sigma$  of the true value,  $\sim 98\%$  within  $2\sigma$ , and virtually all within  $3\sigma$ . This is in almost perfect agreement with expectation from statistical errors.

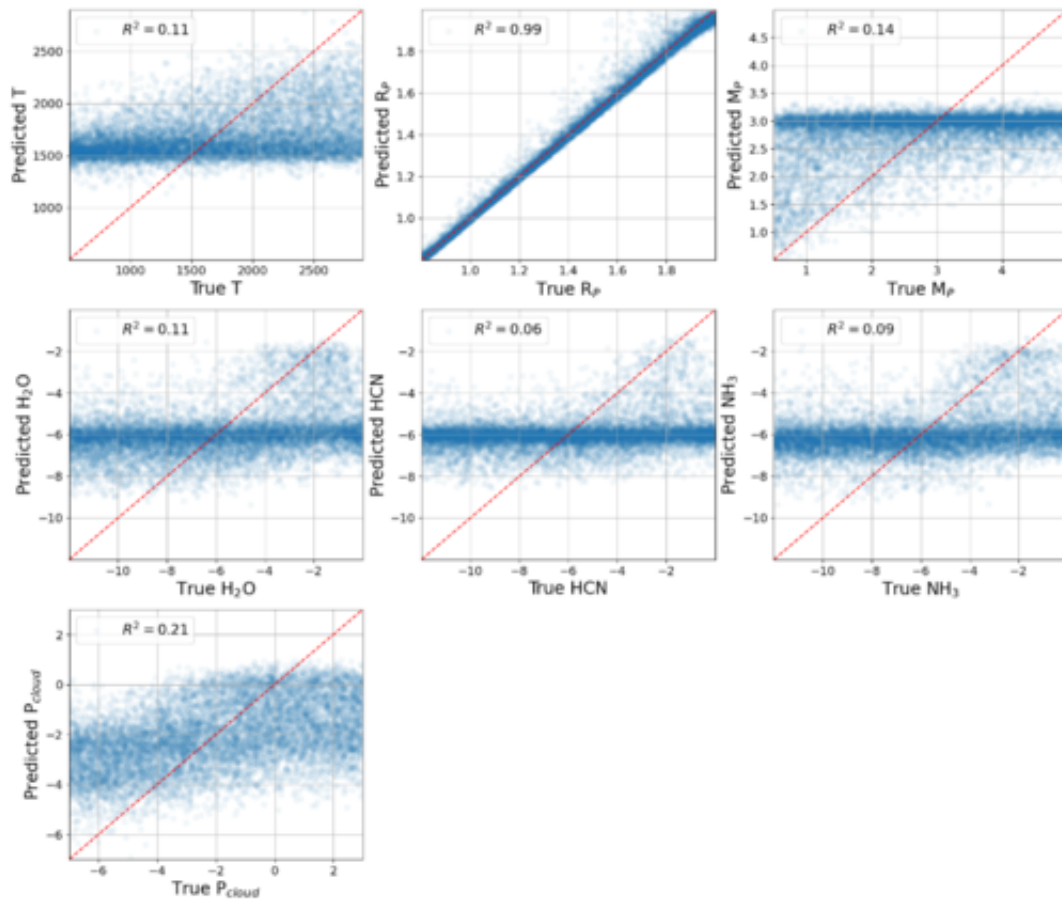


Figure 1: Predicted versus true values of the parameters for CNN retrievals of 10,000 synthetic spectra.

Figures 3 and 4 show the CNN and Nested Sampling retrieval of WASP-12b. Again, we see very similar results, although the CNN provides broader posterior distributions. Work is ongoing to try to fix this issue.

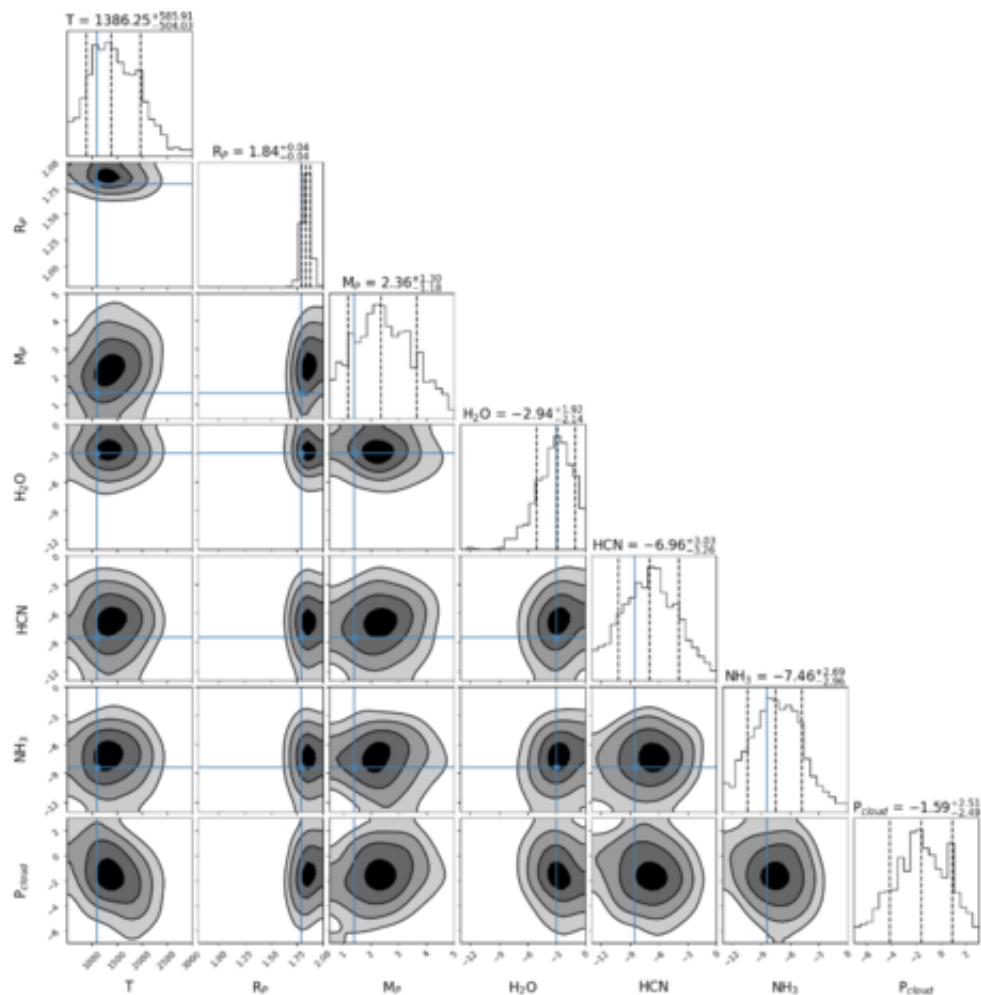


Figure 3: CNN retrieval of WASP-12b. Blue lines from Nested Sampling retrieval in [5].

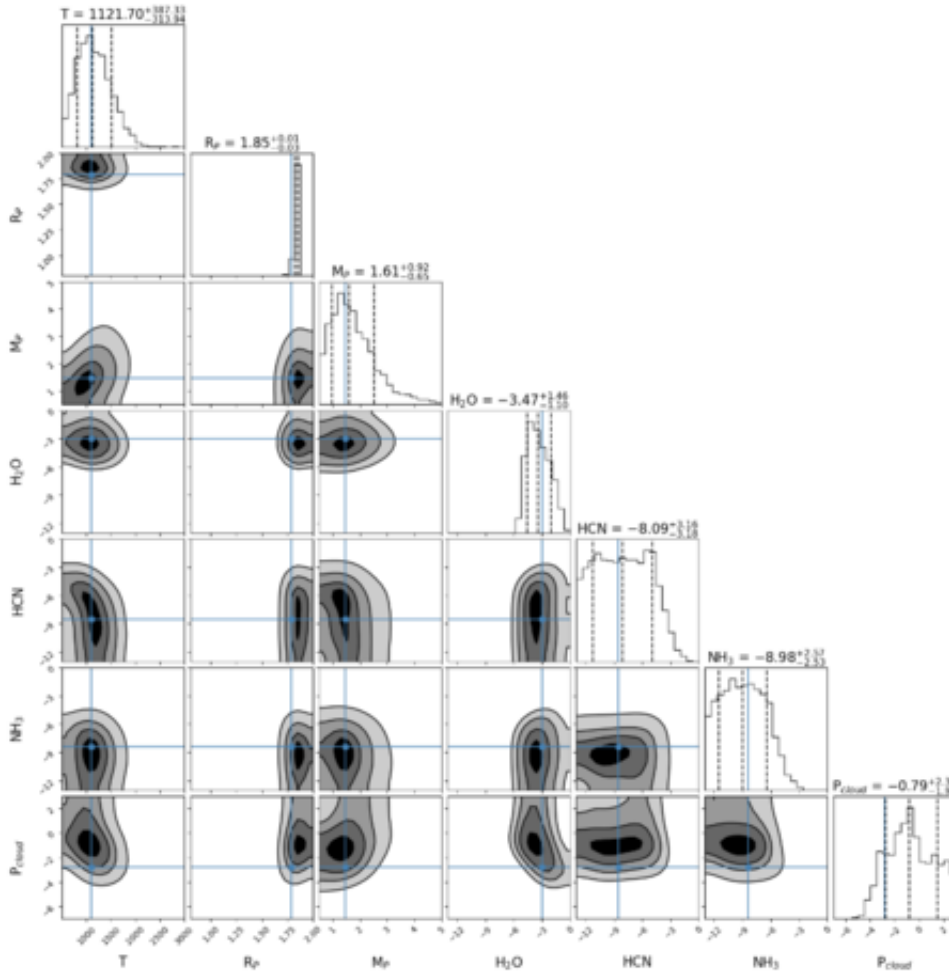


Figure 4: Nested Sampling retrieval of WASP-12b. Blue lines from Nested Sampling retrieval in [5].

We found that a training set of 180,000 spectra is unnecessarily large, and the same performance can be reached with only 20,000 spectra. This implies that the number of forward model computations needed to train a CNN is smaller than the number needed for a single Bayesian retrieval. If this holds for more complex forward models and higher quality spectra, it would make machine learning an extremely attractive alternative to Nested Sampling.

**Conclusion:** The existing literature on machine learning retrievals of exoplanet atmospheres only has comparisons between machine learning and Nested Sampling for a handful of test cases [1, 5, 8]. In this work we present a comparison of bulk retrievals done with both methods, showing that machine learning can indeed be a viable and fast alternative to Nested Sampling. We are currently working on extending these results to models with equilibrium chemistry and to JWST/NIRSpec simulated spectra.

**Acknowledgements:** This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 860470.

#### References:

[1] Cobb, A. D., Himes, M. D., Soboczenski, F., Zorzan, S., O’Beirne, M. D., Baydin, A. G., Gal, Y., Domagal-Goldman, S. D., Arney, G. N., Angerhausen, D. (2019, 5). An Ensemble of Bayesian

Neural Networks for Exoplanetary Atmospheric Retrieval.

[2] Feroz, F., Hobson, M. P. (2008). Monthly Notices of the Royal Astronomical Society 384 (2).

[3] Feroz, F., Hobson, M. P., Bridges, M. (2009). Monthly Notices of the Royal Astronomical Society 398(4).

[4] Kreidberg, L., Line, M. R., Bean, J. L., Stevenson, K. B., Desert, J.-M., Madhusudhan, N., Fortney, J. J., Barstow, J. K., Henry, G. W., Williamson, M. H., Showman, A. P. (2015). A DETECTION OF WATER IN THE TRANSMISSION SPECTRUM OF THE HOT JUPITER WASP-12b AND IMPLICATIONS FOR ITS ATMOSPHERIC COMPOSITION. Technical report.

[5] Marquez-Neila, P., Fisher, C., Sznitman, R., Heng, K. (2018). Supervised machine learning for analysing spectra of exoplanetary atmospheres.

[6] Min, M., Ormel, C. W., Chubb, K., Helling, C., Kawashima, Y. (2020). The ARCiS framework for exoplanet atmospheres. Astronomy & Astrophysics 642.

[7] Skilling, J. (2006). Nested sampling for general Bayesian computation. Bayesian Analysis 1 (4).

[8] Zingales, T., Waldmann, I. P. (2018). The Astronomical Journal 156 (6).