

Follow the Water: Finding Water, Snow and Clouds on Terrestrial Exoplanets with Photometry and Machine Learning

Dang Pham¹*, Lisa Kaltenegger²

¹David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada

²Department of Astronomy and Carl Sagan Institute, Cornell University, 302 Space Sciences Building, Ithaca, NY 14853, USA

Last updated YYYY MM DD; in original form YYYY MM DD

ABSTRACT

All life on Earth needs water. NASA’s quest to follow the water links water to the search for life in the cosmos. Telescopes like JWST and mission concepts like HabEx, LUVOIR and Origins are designed to characterise rocky exoplanets spectroscopically. However, spectroscopy remains time-intensive and therefore, initial characterisation is critical to prioritisation of targets.

Here, we study machine learning as a tool to assess water’s existence through broadband-filter reflected photometric flux on Earth-like exoplanets in three forms: seawater, water-clouds and snow; based on 53,130 spectra of cold, Earth-like planets with 6 major surfaces. XGBoost, a well-known machine learning algorithm, achieves over 90% balanced accuracy in detecting the existence of snow or clouds for $S/N \gtrsim 20$, and 70% for liquid seawater for $S/N \gtrsim 30$. Finally, we perform mock Bayesian analysis with Markov-chain Monte Carlo with five filters identified to derive exact surface compositions to test for retrieval feasibility.

The results show that the use of machine learning to identify water on the surface of exoplanets from broadband-filter photometry provides a promising initial characterisation tool of water in different forms. Planned small and large telescope missions could use this to aid their prioritisation of targets for time-intensive follow-up observations.

Key words: astrobiology – methods: statistical – techniques: photometric – planets and satellites: surface – terrestrial planets

1 INTRODUCTION

Thousands of exoplanets have been detected to date; including terrestrial planets. Dozens of these orbit in their host star’s habitable zone (HZ) (e.g. Kane et al. 2016; Johns et al. 2018), where liquid water could be available on their surface (Kasting et al. 1993). Missions, such as the Transiting Exoplanet Survey Satellite (Ricker et al. 2014), and ground-based searches continually expand our list of potentially habitable planets (e.g. Nutzman & Charbonneau 2008; Luque et al. 2019; Kaltenegger et al. 2021). Upcoming telescopes, such as the James Webb Space Telescope, will probe the atmosphere of Earth-sized exoplanets in the HZ (e.g. Zhang et al. 2018; Krissansen-Totton et al. 2018; de Wit et al. 2018; Wunderlich et al. 2019; Kaltenegger et al. 2020; Edwards et al. 2021).

Planetary formation models suggest that water-rich planets are common (Raymond et al. 2014), making planets with surface water interesting targets for upcoming telescope observations (e.g. Wunderlich et al. 2019; Smith et al. 2020a). However, being in the HZ does not imply habitability (Selsis et al. 2007; Kaltenegger 2017). Understanding the habitable characteristics of a planet requires a detailed characterisation of the exoplanet through time intense spectroscopic observations to analyze the planets’ atmosphere and surface composition. But spectroscopy of Earth-size planets in the HZ will be time-intensive, even with future telescope concepts recommended by the recent Decadal survey like HabEx or LUVOIR (e.g. Feng et al. 2018; The LUVOIR Team 2019; Gaudi et al. 2020). Thus, fast iden-

tification and prioritisation of targets is a critical aspect to enable a successful search. Several teams suggested photometry as a tool for such initial characterisation, testing it on Solar System objects as well as exoplanet models and explored optimal filters for such characterisation (e.g. Crow et al. 2011; Hegde et al. 2015; Krissansen-Totton et al. 2016; Madden & Kaltenegger 2018; Smith et al. 2020b).

Additionally, machine learning methods have shown potential for fast initial characterisation of extrasolar planets with photometry: For giant planets Batalha et al. (2018) explore how to characterise atmospheric properties, such as metallicity and cloud coverage through supervised machine learning for the 5 broadband filters of the Nancy Grace Roman Space Telescope. For Earth-like planets, Pham & Kaltenegger (2021) assess eight machine learning methods to characterise surface biota through photometric fluxes for the standard Johnson filters.

Broadband filter photometry could allow smaller telescopes like the Nancy Grace Roman Telescope (Spergel et al. 2015), Ariel (Tinetti et al. 2021) and Twinkle (Edwards et al. 2019) to help initially characterising nearby rocky exoplanets. It could also allow large missions concepts like HabEx (Gaudi et al. 2020) and LUVOIR (The LUVOIR Team 2019) rapid initial characterisation of exoplanets in the Habitable Zone to prioritise targets for time-intensive spectroscopic observations. This approach is complimentary to transit observations proposed for rocky exoplanet from JWST (Gardner et al. 2006), which won’t be able to probe the surface of a rocky exoplanet.

In this article, we explore the capability of machine learning to detect water on the surface of a cool terrestrial planet in three forms: liquid seawater, water-clouds and snow. Note that we did not choose

* E-mail: dang.pham@astro.utoronto.ca (DP)

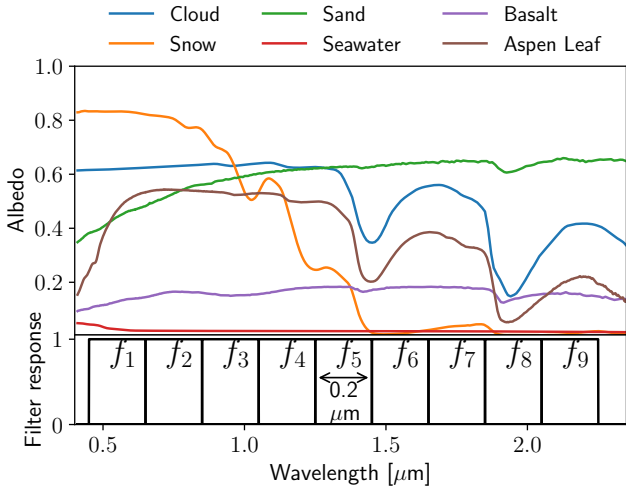


Figure 1. *Top:* Albedos of six reflecting components representative of modern Earth, with Aspen leaf as representative for Earth vegetation (USGS Spectra Library (Clark et al. 2003)). *Bottom:* Nine idealised, broadband filters (width = $0.2 \mu\text{m}$) between $0.45 - 2.25 \mu\text{m}$.

specific filters for any proposed telescope concept here, but we explore what the optimal choice would be to find water on an Earth-like exoplanet’s surface to inform open design choices.

We propose machine learning as an effective method to characterise terrestrial exoplanets’ reflected broadband photometry, along with other statistical methods (e.g. compared to MCMC as shown in this paper). Our machine learning method XGBoost performs on a model grid of 53,130 Earth-size planets with varying surface coverage of six major surfaces. Our aims are to i) assess machine learning’s efficacy in classifying the existence of water, clouds or snow, ii) identify optimal filters for characterisation for telescope designs and iii) perform mock Bayesian analysis on fluxes attained from these optimal filters.

First, we define an initial set of idealised, broadband filters. We then train a well-known, versatile machine learning model – XGBoost (Chen & Guestrin 2016) – with fluxes derived using the broadband filters, and assess its performance at various S/N. This approach delivers an estimated signal-to-noise ratio cutoff for performance maximisation. In addition, the trained XGBoost identifies a set of five¹ optimal filters through feature importance ranking. Using the five optimal filters, we perform Bayesian inference through Markov-chain Monte Carlo sampling. The analysis described in this paper can be extended to any detected exoplanet by training the algorithm on specific models pertaining to the specific planetary environment (see Truitt et al. 2020).

This article is structured as follows: in section 2, we introduce the planetary models; in section 3, we discuss the usage of XGBoost and its performance; in section 4, we perform mock Bayesian analysis on a case study of Earth and on one hundred random combinations and in section 5 we summarise our findings.

2 DATA GENERATION

We use a well-known 1D atmospheric model (EXO-Prime2, details in e.g. Kaltenegeger & Lin 2021) to model a slightly colder Earth with a surface temperature of 273K with modern Earth outgassing rates and a Solar spectrum with reduced luminosity of $0.875L_{\odot}$ corresponding to a planet further away from the Sun than Earth to generate a cool environment that would allow for snow to cover a large part of the planet. While more complex 3D models are used to model individual planets and explore the effects of surface topography, and rotation rate on specific planets, a computationally cheaper 1D model, which generates a disk integrated flux over an entire planet’s surface for broadband filters provides a suitable approximation for the parameter space exploration we undertake here.

We use the six major surface components of modern Earth (water, snow, basalt, vegetation, sand, clouds; following Kaltenegeger et al. 2007) varying in 5% steps for surface coverage (Pham & Kaltenegeger 2021) to generate the corresponding reflection spectra of 53,130 nominal Earth-size planets. These reflecting components are θ_i , where i ranges from 1-6 representing each component. We create spectra for each nominal planet at a resolution of 0.01 cm^{-1} using Exo-Prime2, a 1-D iterative climate-photochemistry code coupled to a line-by-line- radiative transfer code (details in Kaltenegeger & Lin 2021; Kaltenegeger et al. 2007; Kaltenegeger & Traub 2009). The model has been validated from the visible to infrared through comparison to Earth seen as an exoplanet by missions like the Mars Global Surveyor, EPOXI, multiple Earthshine observations and Shuttle data (Kaltenegeger et al. 2007; Kaltenegeger & Traub 2009; Rugheimer et al. 2013). It includes the most spectroscopically relevant molecules: C_2H_6 , CH_4 , CO , CO_2 , H_2CO , H_2O , H_2O_2 , H_2S , HNO_3 , HO_2 , N_2O , N_2O_5 , NO_2 , O_2 , O_3 , OCS , OH , SO_2 , using the HITRAN2016 line lists (Gordon et al. 2017) as well as Rayleigh scattering. We divide the planet atmosphere into 52 layers: for each atmospheric layer the code calculates line shapes and widths individually with Doppler- and pressure-broadening with several points per line width.

The other major surface components’ albedo are taken from the USGS Spectra Library (Clark et al. 2003), see Fig. 1 with Aspen leaf representative modern Earth vegetation (for the effects of different vegetation and biota and their identification through machine learning on Earth-like planets, see Pham & Kaltenegeger 2021) and one opaque water cloud layer at 6 km. The cloud reflectivity (following Madden & Kaltenegeger 2020) is based on the MODIS $20 \mu\text{m}$ cloud albedo model (King et al. 1997; Rossow & Schiffer 1999), which provides an average albedo for clouds of different droplet size and is representative of the 3 main water cloud layers on Earth at 1 km, 6 km and 12 km (see e.g. Kaltenegeger et al. 2007). We restrict our spectra wavelength to $0.41 - 2.35 \mu\text{m}$, because these are the available wavelengths of samples from the USGS Spectra Library. Note that we do not adjust the atmosphere profile for different combinations of reflective components, which introduces small changes in the planet’s atmosphere and influences their high resolution spectra (see Madden & Kaltenegeger 2020 for details). Here we are only interested in the broadband fluxes resulting from integrating the spectra over broadband filters, and require tens of thousands of spectra for machine learning and a similar number for model evaluations with Markov-chain Monte Carlo. Hence, keeping a constant atmosphere profile allows us to explore a wide parameter space and evaluate the model efficiently.

¹ We chose five filters based on previous instrument designs like the Nancy Grace Roman Space Telescope (Spergel et al. 2013). The methodology here can be easily extended to assess the trade-offs for more (or less) filters.

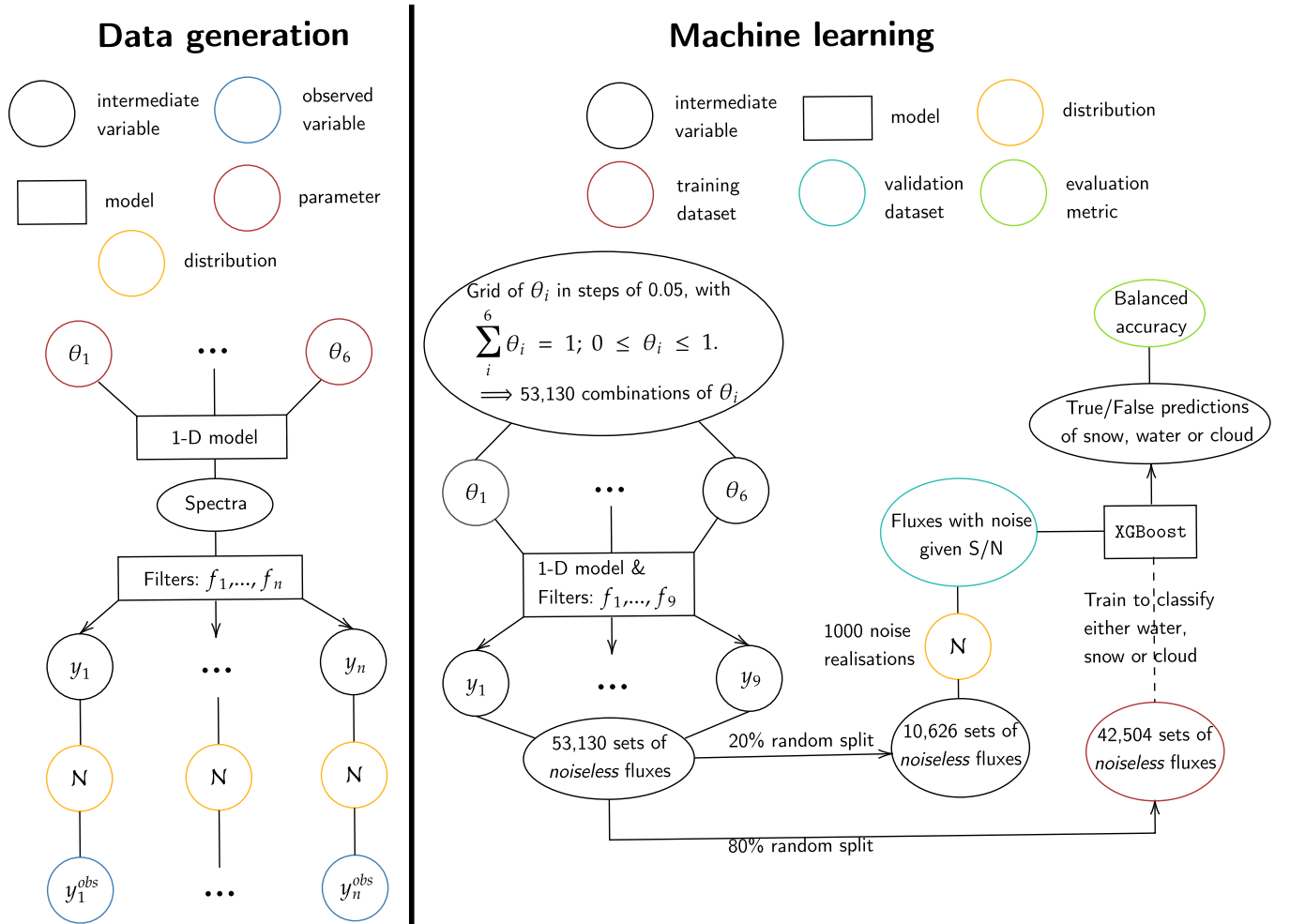


Figure 2. *Left:* Graphic representation of the data generation process. θ_i represents the composition of each reflective component (cloud, snow, etc.), f_i the filters' response, y_i the true integrated flux from each filter and y_i^{obs} the observed flux. Gaussian noise is added to the true flux (N) producing a signal-to-noise ratio (S/N). *Right:* Graphic of the machine learning pipeline – data generation, training and validation. θ_i represents the composition of a particular reflective component (cloud, snow, etc.), N is the Gaussian model (given some signal-to-noise ratio). To each set of noiseless flux in the 20% random split noise is added 1000 times (1000 noise realisations). The balanced accuracy is calculated from the prediction of *each* realisation. Hence, at each S/N , we have 1000 balanced accuracy scores. Yields in the shaded 95% confidence interval are shown in Fig. 3.

We initially created nine, idealised, broadband filters (f_i) between $0.45 - 2.25 \mu\text{m}$ with a width of $0.2 \mu\text{m}$ with an idealised filter response (Fig. 1) to identify which filters facilitate an effective characterisation of liquid water, snow and clouds on terrestrial exoplanets. We then integrate the model planet spectra with these filters to generate the corresponding true flux, y_i , for each filter. We add Gaussian noise to these true fluxes, producing the observed flux y_i^{obs} , given some S/N for our analysis (see Fig. 2 for a summary of the data generation and the machine learning pipeline). For machine learning validation (section 3), Gaussian noise with S/N in $[5, 100]$ are added. For the Bayesian data generation pipeline (section 4), Gaussian noise is added at $S/N = 10, 50, 100$.

3 MACHINE LEARNING WITH XGBOOST

We use XGBoost (Chen & Guestrin 2016), a versatile, gradient boosting algorithm (e.g. Tamayo et al. 2016), to perform binary classification of liquid water, snow and clouds on our model planets. We are

classifying if a certain component θ_i exists (True) or not (False). In our analysis, we train XGBoost to classify only either water, snow or cloud at a time. Through a grid search, we find the default hyperparameters for XGBoost work well for our problem².

The training and validation dataset are generated like in Pham & Kaltenecker 2021. First, we create a grid of combinations of reflective components θ_i , such that $\sum_{i=1}^6 \theta_i = 1; 0 \leq \theta_i \leq 1$ ³ where each component θ_i ranges from 0 to 1, in steps of 0.05, subject to the unity sum condition. Therefore, we create a total of 53,130 combinations. Then, we create spectra⁴ for each of these combination and nine filter fluxes per spectra. We randomly split this dataset of 53,130 set of nine fluxes into the training and validation datasets at a ratio of

² We set the `scale_pos_weight` parameter, the ratio of negative to positive samples in the training dataset, but the ratio is fixed for a given dataset.

³ This condition describes the $n - 1$ simplex ($n = 6$ here). In other words, we create a grid with resolution 0.05 on a 5-simplex.

⁴ The full spectra and filters dataset is publicly available at <https://doi.org/10.5281/zenodo.6234713>.

80-20. Thus, the training dataset contains 42,504 sets of fluxes (80%) and the validation dataset the remaining 10,626 (20%).

We keep the training dataset noiseless. Gaussian noise is added to the validation dataset for a given S/N. Each validation combination has 1000 noise realisations at a given S/N. This way, we can evaluate our trained model’s performance at various noise level. With this methodology, the model is exposed to two different distributions: noiseless for training data and noisy for validation data. Prior work by Batalha et al. (2018); Hayes et al. (2020) similarly trained on noiseless data. Our previous work (Pham & Kaltenegeger 2021) tested training on noisy and noiseless data and found the latter yielding better results. This approach is versatile because it explores the capability of the algorithms without specifying a noise profile to train on.

To evaluate performance, we use balanced accuracy. The problem resulting from using accuracy instead of balanced accuracy for an imbalanced dataset is discussed in details in Pham & Kaltenegeger (2021). As a brief summary, a simple metric to evaluating machine learning efficiency is the accuracy (Eq. 1). However, this metric only works well if the dataset is balanced – when sample counts in all classes (True or False in this case) are equally distributed. We label any set of fluxes with (snow, cloud, or water) components at 0% as False, and True otherwise. Our dataset is imbalanced: for example, there are 42,504 combinations with snow (True prediction) yet only 10,626 combinations without snow (False prediction). Using the traditional accuracy usually leads to overly optimistic results (Brodersen et al. 2010).

For imbalanced datasets, the balanced accuracy should be used instead (Eq. 2), where sensitivity (True Positive Rate) gives the proportion of true positives and specificity (True Negative Rate) gives the proportion of true negatives. Balanced accuracy metric accounts for the true positive and negative rates, instead of averaging over the entire confusion matrix like the accuracy would. Note that another metric that can be used on an imbalanced dataset is a confusion matrix (see Pham & Kaltenegeger 2021). However, the balanced accuracy is a scalar value, which can be used when training the model and can easily measure performance quantitatively.

$$\text{accuracy} = \frac{\text{correct predictions}}{\text{total predictions}} \quad (1)$$

$$\text{balanced accuracy} = \frac{\text{sensitivity} + \text{specificity}}{2} \quad (2)$$

XGBoost’s (trained under nine filters) balanced accuracy over various S/N is shown in panel (a) of Fig. 3. Snow and clouds reach asymptotic to approximately unity, water to 0.7 balanced accuracy. Furthermore, the shaded area representing the 95% confidence interval acquired from 1000 noise realisations is small.

The difference in performance can be explained by the different reflectivity of snow, clouds and liquid water shown in Fig. 1. The reflectivity of water is low (< 0.1) and has a relatively featureless shape. In contrast, the reflectivity of both snow and clouds show specific features and reflect up to 80% of the incoming light for snow and up to 60% for clouds. Liquid water surface features are more difficult to identify than snow or clouds.

Note that Fig. 3a shows that for broadband filter photometry a S/N of 20 already provide high accuracy prediction for the existence of water on Earth-like planet models. Beyond S/N = 30, the balanced accuracy becomes asymptotic, meaning additional signal yields a relatively low increase in performance (in a statistical sense).

XGBoost results indicate that machine learning algorithms can be effective tools to classify the existence of snow, clouds and water on Earth-size exoplanets from photometric filter data even at low S/N.

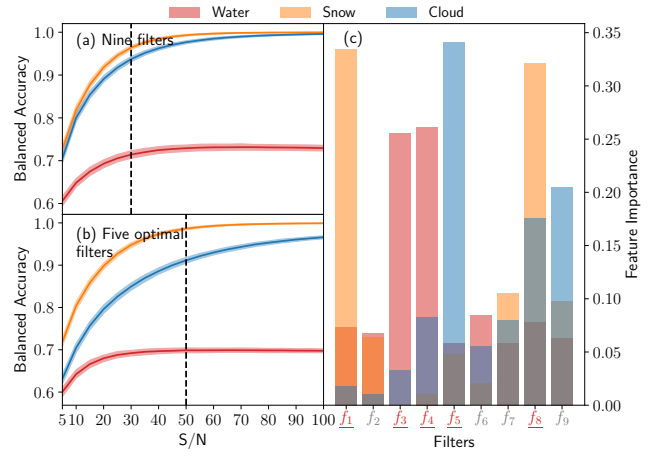


Figure 3. (a): Performance (measured with balanced accuracy) of XGBoost at classifying the existence of snow, water or clouds at different signal-to-noise ratio (S/N) with all nine idealised filters. The shaded area represents the 95% confidence interval from 1000 noise realisations. The algorithm approaches asymptotes at S/N \approx 30 (dashed line). (b): Same as (a), but for the five optimal filters. Here, the dashed line is chosen when both water and snow reach their asymptotes and cloud achieve a balanced accuracy $>$ 0.9. (c): Feature (filters) importance score ranking. A higher feature score implies greater importance in classification. We select the top five optimal filters (underlined in red).

With the trained model, we identify the set of five optimal filters for identifying water on the surface of an terrestrial exoplanet through photometry between 0.41 - 2.35 μm using XGBoost’s feature importance score. Here, ‘features’ are the filters for the surfaces we want to identify. Ranking the nine fluxes (as generated by the nine filters) by their importance score allows us to determine the optimal filters (Fig. 3c): f_1 (0.45 - 0.65), f_4 (1.05 - 1.25), f_5 (1.25 - 1.45), f_6 (1.45 - 1.65), and f_8 (1.85 - 2.05). Note that the choice of optimal filters are method (XGBoost) and model dependent (1D atmosphere EXO-Prime2, USGS albedo spectral library). In section 4, we study if the optimal filters for XGBoost yield effective performance with a Bayesian approach (MCMC).

Comparing the importance scores (Fig. 3c) with the wavelength dependent albedo of each component (Fig. 1) shows peaks of reflectivity of the surfaces in the corresponding filters. The results are based on the built-in XGBoost method, and are consistent with the expectations from surface albedos in Fig 1. For example, three of the best ranked filters for cloud detection are f_5 , f_8 and f_9 , which correspond to cloud reflection peaks between 1.6 - 1.9 μm and 2.1 - 2.3 μm . Similar correspondence between reflection features and optimal filters exist for snow and water.

Figure 3b shows the balanced accuracy of XGBoost using only the five optimal filters. Water and clouds reach a lower overall balanced accuracy asymptote, snow reaches its earlier asymptote value at higher S/N. These effects are expected because the five filter provide less data compared to the nine filters used in the top panel and our earlier analysis. In Fig. 3b, we identify S/N \approx 50 (dashed line) as the cutoff for additional signal rewards. However, note that our results show that even a S/N of \approx 20 yields a strong performance in accurately predicting the existence of water on the surface of an Earth-like exoplanet.

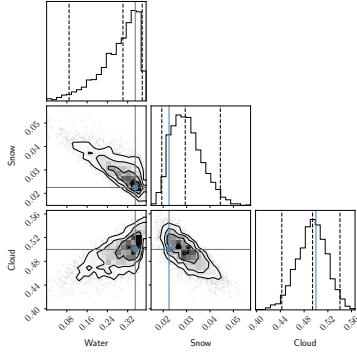


Figure 4. Corner plot showing the 1-D and 2-D marginalised distributions of water, snow and clouds for one random realisation with modern Earth surface composition at $S/N = 50$. The dashed lines represent the median and the 95% credible interval, the solid blue lines mark the true value.

4 BAYESIAN ANALYSIS

We assess the surface composition of our model planets regarding liquid water, snow or clouds with Bayesian inferencing using Markov-chain Monte Carlo (MCMC) for data from the set of five optimal filters identified earlier. That is, inferring θ_i given $\{y_1^{\text{obs}}, y_3^{\text{obs}}, y_4^{\text{obs}}, y_5^{\text{obs}}, y_7^{\text{obs}}\}$, assuming the model planet spectra and Gaussian noise.

Our posterior distribution function is

$$\mathcal{P}(\theta|y^{\text{obs}}) \propto \mathcal{U}_{\theta}^*(0, 1) \times \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2} \frac{(y_i^{\text{obs}} - y_i^{\theta})^2}{\sigma_i^2}\right) \quad (3)$$

where $\theta = \{\theta_1, \dots, \theta_6\}$ are the components' composition, y_i^{obs} are the (selected optimal) observed fluxes, y_i^{θ} are the modelled planetary fluxes in each filter θ , and σ_i are the measurement error. The product term is the likelihood function, assuming a Gaussian noise model. Each component $\theta_i \in [0, 1]$, and the sum of all components must be exactly unity. Since we don't know the distribution of exoplanets' surface compositions, we assume that any combination of θ_i summing to unity is equally likely. Hence, we propose a uniform 5-simplex prior, defined as:

$$\mathcal{U}_{\theta}^*(0, 1) \propto \begin{cases} 1 & \sum_{i=1}^6 \theta_i = 1; 0 \leq \theta_i \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

We sample this posterior distribution function using `emcee` (Foreman-Mackey et al. 2013).

We first perform a case study of *one* realisation for a model planet with modern Earth's surface composition at $S/N = 50$ for the five identified filters. This corresponds to a sample case of a planet-by-planet analysis that would be done for a specific exoplanet (here we choose the Earth model case). Modern Earth surface seen from space consists on average of about 50% clouds. The 50% surface is composed of about 70% seawater, 5.4% basalt, 4.5% snow, 2.1% sand, and 18% vegetation (Kaltenegger et al. 2007). These factors represent our θ_i 's for this case, which sum to unity.

Figure 4 shows the 1-D and 2-D marginalised distributions of water, snow and clouds for this one particular realisation at $S/N = 50$: It shows the median and 95% credible interval estimates from MCMC and the true composition. The MCMC's 95% credible interval covers the true value, and the median estimate is close as well. The results

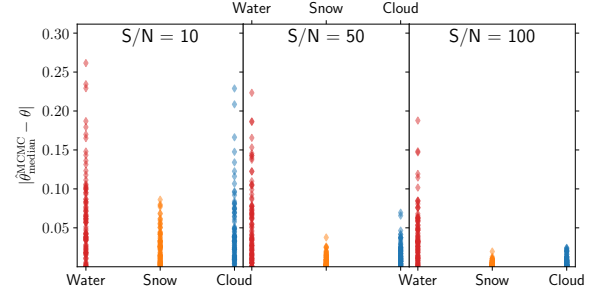


Figure 5. Residual between MCMC median estimates and the true values at $S/N = 10, 50, 100$, for one hundred random compositions. $\hat{\theta}_{\text{median}}^{\text{MCMC}}$ is the median estimate from MCMC and θ is the true value. The greatest residual for all three S/N is $\max |\hat{\theta}_{\text{median}}^{\text{MCMC}} - \theta| = 0.26$.

demonstrate that it is feasible to infer the surface composition regarding water as liquid, snow and clouds on a terrestrial exoplanet based on the proposed five optimal filters, within a 95% credible interval for this particular combination.

Figure 5 shows the results when we extend the analysis to one hundred random combinations, each with added noise at S/N of 10, 50 and 100. It shows the residual between MCMC median estimates, $\hat{\theta}_{\text{median}}^{\text{MCMC}}$, and the true values, θ . Note that the residual for snow and cloud reduces significantly as S/N increases. This corresponds to the balanced accuracy performance from XGBoost with increasing S/N (Fig. 3) due to the reflection features of the three surfaces (see Fig. 1). Snow and clouds reflect more incoming star light and show features over the wavelength range covered. In contrast the albedo of liquid seawater is comparably smooth, making liquid water more difficult to identify than snow or clouds. Finally, most residuals are accurate to within 5% for snow and clouds at $S/N \geq 50$ – these result highlighting the potential of using photometric filters to assess the existence of water on a terrestrial exoplanet.

5 CONCLUSION

We explored the feasibility of detecting water on the surface of terrestrial exoplanets in its various forms (snow, clouds and liquid water) through broadband photometry using machine learning and Markov-chain Monte Carlo (MCMC).

First, we trained a well known, versatile machine learning algorithm, XGBoost, to perform binary classification on the existence of snow, clouds and water on an exoplanet's surface using broadband photometric flux. The performance show promise to use machine learning on photometric data to identify water on the surface of terrestrial exoplanets (Fig. 3a,b): the algorithm achieves a high ($> 90\%$) balanced accuracy for snow and clouds for $S/N \geq 20$) and up to 70% balanced accuracy for liquid water.

Second, we identified five optimal filters to identify snow, clouds and liquid water on a terrestrial planet's surface between 0.45 - 2.25 μm based on XGBoost's feature importance ranking. These optimal filters (Fig. 3c), could be implemented in telescopes designs that search for water on exoplanets.

Third, we tested these optimal filters by performing Bayesian inferences using MCMC on i) an Earth case study, and ii) one hundred random realisations of our planetary models. The results show promise, with most predictions at $S/N \geq 50$ within 5% of the true value for finding snow and clouds. Detecting liquid water is more challenging, but most predictions are within 20% of the true values.

In a typical workflow, we expect to use machine learning for fast prioritisation and use MCMC for further constraints. The advantage of machine learning is its fast run time (in our runs, about 1 minute for 53,000 sets of fluxes), comparing to MCMC sampling's (about 15 minutes per set of fluxes). We find that machine learning can be an additional fast and accurate retrieval pathway. Note that MCMC would only require ~ 2500 hours for 10,000 samples.

The results showed that broadband filter photometry combined with machine learning provides a promising tool for initial characterisation of water in different forms on the surface of terrestrial exoplanets and their prioritisation for time-intense follow-up observations for a smaller observing time cost.

Broadband filter photometry could allow large missions like HabEx and LUVOIR to initially characterise targets for follow-up and smaller telescopes like the Nancy Grace Roman telescope to help with initial characterisation of nearby rocky worlds.

ACKNOWLEDGEMENTS

We thank the anonymous reviewer for their insightful suggestions that have improved the clarity and strengthened our paper. We would like to thank David Ruppert, Joshua S. Speagle, Samantha Berek and Zifan Lin for helpful discussions and comments.

DATA AVAILABILITY

The data underlying this article are available in <https://doi.org/10.5281/zenodo.6234713>.

REFERENCES

- Batalha N. E., Smith A. J. R. W., Lewis N. K., Marley M. S., Fortney J. J., Macintosh B., 2018, *The Astronomical Journal*, 156, 158
- Brodersen K. H., Ong C. S., Stephan K. E., Buhmann J. M., 2010, in 2010 20th International Conference on Pattern Recognition. IEEE, doi:10.1109/icpr.2010.764, <https://doi.org/10.1109/icpr.2010.764>
- Chen T., Guestrin C., 2016, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16. ACM, New York, NY, USA, pp 785–794, doi:10.1145/2939672.2939785, <http://doi.acm.org/10.1145/2939672.2939785>
- Clark R. N., Swayze G. A., Wise R. K., Livo E., Hoefen T. M., Kokaly R. F., Sutley S. J., 2003, USGS Digital Spectral Library splib05a, doi:10.3133/ofr03395, <https://doi.org/10.3133/ofr03395>
- Crow C. A., et al., 2011, *The Astrophysical Journal*, 729, 130
- Edwards B., et al., 2019, *Experimental Astronomy*, 47, 29
- Edwards B., et al., 2021, *AJ*, 161, 44
- Feng Y. K., Robinson T. D., Fortney J. J., Lupu R. E., Marley M. S., Lewis N. K., Macintosh B., Line M. R., 2018, *The Astronomical Journal*, 155, 200
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Gardner J. P., et al., 2006, *Space Sci. Rev.*, 123, 485
- Gaudi B. S., et al., 2020, arXiv e-prints, p. arXiv:2001.06683
- Gordon I. E., et al., 2017, *J. Quant. Spectrosc. Radiative Transfer*, 203, 3
- Hayes J. J. C., et al., 2020, *MNRAS*, 494, 4492
- Hegde S., Paulino-Lima I. G., Kent R., Kaltenecker L., Rothschild L., 2015, *Proceedings of the National Academy of Sciences*, 112, 3886
- Johns D., Marti C., Huff M., McCann J., Wittenmyer R. A., Horner J., Wright D. J., 2018, *The Astrophysical Journal Supplement Series*, 239, 14
- Kaltenecker L., 2017, *Annual Review of Astronomy and Astrophysics*, 55, 433
- Kaltenecker L., Lin Z., 2021, *ApJ*, 909, L2
- Kaltenecker L., Traub W. A., 2009, *The Astrophysical Journal*, 698, 519
- Kaltenecker L., Traub W. A., Jucks K. W., 2007, *The Astrophysical Journal*, 658, 598
- Kaltenecker L., MacDonald R. J., Kozakis T., Lewis N. K., Mamajek E. E., McDowell J. C., Vanderburg A., 2020, *ApJ*, 901, L1
- Kaltenecker L., Pepper J., Christodoulou P. M., Stassun K., Quinn S., Burke C., 2021, *AJ*, 161, 233
- Kane S. R., et al., 2016, *The Astrophysical Journal*, 830, 1
- Kasting J. F., Whitmire D. P., Reynolds R. T., 1993, *Icarus*, 101, 108
- King M., Tsay S.-C., Platnick S., Wang M., Liou K., 1997, MODIS Algorithm Theoretical Basis Document, No. ATBD-MOD-0
- Krissansen-Totton J., Schwietzman E. W., Charnay B., Arney G., Robinson T. D., Meadows V., Catling D. C., 2016, *The Astrophysical Journal*, 817, 31
- Krissansen-Totton J., Garland R., Irwin P., Catling D. C., 2018, *AJ*, 156, 114
- Luque R., et al., 2019, *Astronomy & Astrophysics*, 628, A39
- Madden J., Kaltenecker L., 2018, *Astrobiology*, 18, 1559
- Madden J., Kaltenecker L., 2020, *The Astrophysical Journal*, 898, L42
- Nutzman P., Charbonneau D., 2008, *Publications of the Astronomical Society of the Pacific*, 120, 317
- Pham D., Kaltenecker L., 2021, *MNRAS*, 504, 6106
- Raymond S. N., Kokubo E., Morbidelli A., Morishima R., Walsh K. J., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 595 (arXiv:1312.1689), doi:10.2458/azu_uapress_9780816531240-ch026
- Ricker G. R., et al., 2014, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 014003
- Rossow W. B., Schiffer R. A., 1999, *Bull. Amer. Meteorol. Soc.*, 80, 2261
- Rugheimer S., Kaltenecker L., Zsom A., Segura A., Sasselov D., 2013, *Astrobiology*, 13, 251
- Selsis F., Kasting J. F., Levrard B., Paillet J., Ribas I., Delfosse X., 2007, *Astronomy & Astrophysics*, 476, 1373
- Smith A. J. R. W., Feng Y. K., Fortney J. J., Robinson T. D., Marley M. S., Lupu R. E., Lewis N. K., 2020a, *AJ*, 159, 36
- Smith A. J. R. W., Mandell A. M., Villanueva G. L., Dan Moore M., 2020b, *The Astronomical Journal*, 160, 204
- Spergel D., et al., 2013, arXiv e-prints, p. arXiv:1305.5425
- Spergel D., et al., 2015, arXiv e-prints, p. arXiv:1503.03757
- Tamayo D., et al., 2016, *ApJ*, 832, L22
- The LUVOIR Team 2019, arXiv e-prints, p. arXiv:1912.06219
- Tinetti G., et al., 2021, arXiv e-prints, p. arXiv:2104.04824
- Truitt A. R., Young P. A., Walker S. I., Spacek A., 2020, *AJ*, 159, 55
- Wunderlich F., et al., 2019, *A&A*, 624, A49
- Zhang Z., Zhou Y., Rackham B. V., Apai D., 2018, *AJ*, 156, 178
- de Wit J., et al., 2018, *Nature Astronomy*, 2, 214

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.