ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Extrapolating Venusian Atmospheric Profiles using MAGMA Gaussian Processes

Simon Lejoly[1], Arianna Piccialli[2], Arnaud Mahieux[2,3], Ann Carine Vandaele[2] and Benoit Frénay[1]

1- University of Namur - NaDI Institute - Faculty of Computer Science
Rue Grandgagnage 21, 5000 Namur - Belgium

2- Royal Belgian Institute for Space Aeronomy - Planetary Atmosphere Department
Avenue Circulaire 3, 1180 Brussels - Belgium

3- The University of Texas at Austin - Department of
Aerospace Engineering and Engineering Mechanics
2617 Wichita St., Stop C0600, Austin, TX 78712 - USA

Abstract. In the field of spatial aeronomy, atmospheric profile datasets often contain partial data. Probabilistic models, particularly Gaussian processes (GPs), offer promising solutions for filling these data gaps. However, traditional GP algorithms encounter challenges when handling multiple sequences simultaneously, both in terms of performance and computational complexity. Recently, an algorithm named MAGMA was introduced to address these issues. This paper evaluates MAGMA's performance using the SOIR Venus atmosphere dataset, marking the first application of MAGMA to atmospheric profiles. Results indicate that MAGMA represents a significant advancement towards the efficient application of GPs for extrapolating atmospheric profiles.

1 Introduction

From May 2006 to November 2014, the Venus Express orbiter collected measurements of the Venusian atmosphere. One of the instruments on the satellite was SOIR (Solar Occultation in the InfraRed). By scanning the light rays coming from the Sun after they went through Venus' atmosphere, SOIR inferred various atmospheric properties such as the temperature and the abundance of a set of chemical species. Spanning multiple altitudes in the mesosphere and the thermosphere above the cloud layer, these observations constitute atmospheric profiles in the SOIR dataset [1].

Across all profiles, altitudes where measurements were taken range from 60km to 160km. However, due to inherent limitations in the measurement process, i.e., solar occultations, precise measurements could only be obtained within narrower altitude ranges. Most profiles in SOIR cover altitudes of merely 10 to 50 consecutive kilometres. The limited amount of usable data in each profile raises the question of what would have been observed at missing altitudes.

One could want to extrapolate the profiles to expand the available data, finding likely values for altitudes that could not be measured. A machine learning algorithm can use the whole dataset to discern a general shape of atmospheric profiles. This shape can then be tailored to individual profiles to infer values outside of their observed altitude range. However, such algorithms must be designed carefully, as their decision-making process and outputs should be understandable by experts to assert their plausibility. Probabilistic models, particularly Gaussian processes (GPs), are promising candidates for this task as they can offer both explainability and a measure of uncertainty in their predictions.

The MAGMA algorithm proposed by Leroy et al. [2] constitutes a recent advance in the field of GPs. It leverages data across multiple sequences to predict missing values more accurately, even far from known observations. This paper explores the benefits of using MAGMA rather than a traditional GP to extrapolate incomplete atmospheric profiles.

2 Related Works

GPs have applications across various domains, including Earth observation data analysis [3, 4], astronomy [5] and spatial aeronomy [6, 7, 8]. However, prior studies mostly focus on either time series modelling or parameter estimation for a single sequence of observations. In the setting of SOIR, the dataset contains multiple profiles that must be analysed together. This motivates the use of a specific extension of GPs called "multi-task GPs". In the context of SOIR, a task corresponds to an atmospheric profile. The literature in multi-task GPs is extensive [9, 10, 11]. However, most of these algorithms either require to specify the covariance between tasks explicitly or struggle to keep relatively tight confidence intervals and scale to larger datasets.

The MAGMA algorithm [2] is a variant of multi-task GPs tailored for datasets where tasks share a common set of inputs (e.g., altitude of observations). When an observation is missing within a specific task, the prediction relies not only on observations of this task close to the missing observation (similar to a conventional GP), but also on the mean value observed at that input across all tasks where it is present. This has the effect of narrowing confidence intervals when going further from observed data in a specific task. Moreover, MAGMA offers an implementation[1] with reasonable computational complexity, enabling applications to larger datasets. Despite its potential, MAGMA has not yet been applied to atmospheric profile analysis or extrapolation. The following sections explore the possible gains MAGMA can bring to the field of spatial aeronomy.

3 Atmospheric Profiles Data

The SOIR dataset contains 2616 profiles of CO_2 and temperature, each observed at the terminator of Venus at specific coordinates and time. A profile comprises multiple measurements with their corresponding 1-σ error estimation, obtained at varying altitudes. The altitude range for measurements varies across profiles.

As this study focuses on asserting the usability of one specific algorithm, only temperature observations, without their estimated error, are utilised. Profiles containing fewer than 10 observations are discarded, as they complicate the

[1] R package available at `https://github.com/ArthurLeroy/MagmaClustR`

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

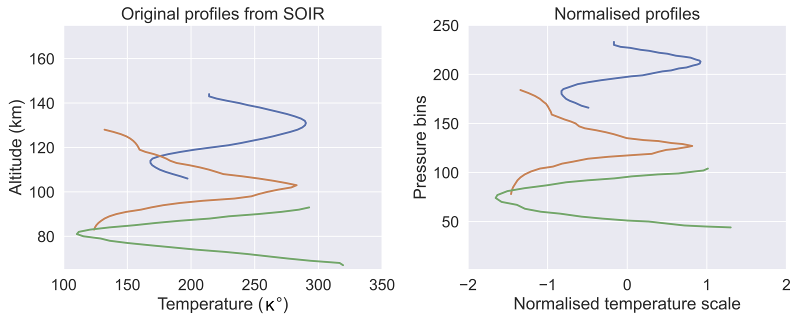

Fig. 1: Comparison of three profiles of the SOIR dataset, before and after being normalised and mapped to a discrete logarithmic pressure scale.

hyperparameter optimisation process of GPs, significantly hurting the performances of all models during experimentation.

As normalisation is known to enhance the performances of GPs, all temperature measurements in the dataset are standardised. In aeronomy, it is also common to use the logarithm of the pressure as a height indicator rather than the altitude. This raises a problem because pressure is measured on a continuous scale, but MAGMA requires discrete, common entries for each task in order to keep its computations feasible. To address this issue, we compute the logarithm of each pressure measurement and map it to its closest bin on a discrete scale. Preliminary experiments showed that a scale of 250 discrete bins gives good results. Figure 1 illustrates the difference between raw profiles from the dataset and their preprocessed versions.

4 Experiments

To assess the performance of MAGMA on the SOIR dataset, we design a simple, single-task GP with a zero mean prior and a standard Radial Basis Function (RBF) kernel to serve as a baseline. The kernel's hyperparameters, namely its length scale and variance, must be optimised in accordance with the working of MAGMA. MAGMA offers two ways of optimising hyperparameters: one where hyperparameters are shared across all tasks, and the other where each task has its own set of hyperparameter values. We call these setups "Common HP" and "Distinct HP", respectively. Both setups are applied during the experiments. To ensure a fair comparison, the baseline GP is trained under similar conditions. In the Common HP setup, the hyperparameters are optimised through stochastic descent using all profiles in the dataset. In Distinct HP, the GP hyperparameters are optimised individually for each profile.

We divide profiles from the SOIR dataset randomly into training and testing sets, with the test set comprising 10% of the profiles. Within each test profile, observations are further divided into given test observations and unseen test observations. During model evaluation, the given test observations serve

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

		Given test obs.		Unseen test obs.	
		MSE	CIC_{95}	MSE	CIC_{95}
Baseline GP	**Distinct HP**	0.0019	**96.59%**	0.3772	81.61%
	Common HP	0.0367	100%	0.4781	100%
MAGMA	**Distinct HP**	0.0016	98.62%	0.2804	78.73%
	Common HP	**0.0013**	98.47%	**0.2378**	**94.67%**

Table 1: Comparison of MAGMA and the baseline GP. The performances on unseen observations are the most relevant, as the model fits the given observations to make predictions.

as starting points for the model. The model can then make predictions for the whole scale of pressure bins, including unseen altitudes. In turn, the unseen test observations are compared to the model's predictions to evaluate its performances. Within each test profile, 33% of observations are designated as unseen test observations. Three variations of the test set are generated, depending on the location of unseen observations: at lower altitudes, at higher altitudes or in the middle of the profile.

The evaluation metrics used to assess the performance of each model are the Mean Squared Error (MSE) and the CI_{95} coverage (CIC_{95}), as in Leroy et al. [2]. The CIC_{95} is the ratio of observations actually located inside the estimated 95% confidence interval, which should ideally be close to 95%. While MSE assesses the precision of the model's mean predictions, CIC_{95} evaluates the plausibility of its predicted confidence intervals.

5 Results and Discussion

Results presented in Table 1 indicate that MAGMA significantly outperforms the baseline on unseen test observations. They also show that while the MSE of

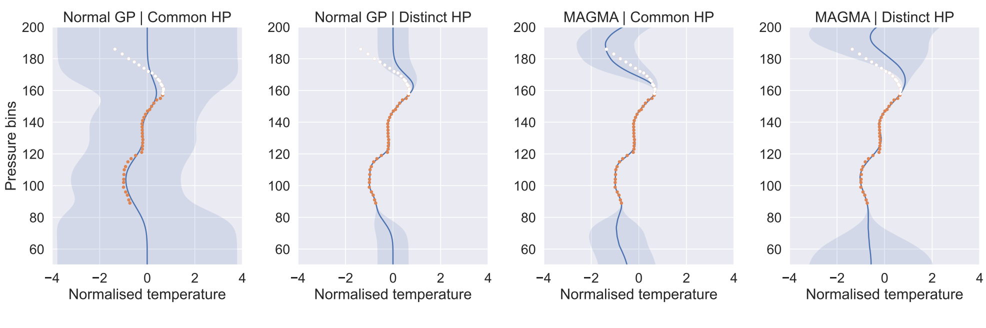

Fig. 2: Predictions from the baseline and MAGMA, both in Distinct and Common HP settings. The given and unseen test observations are represented in orange and white, respectively.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

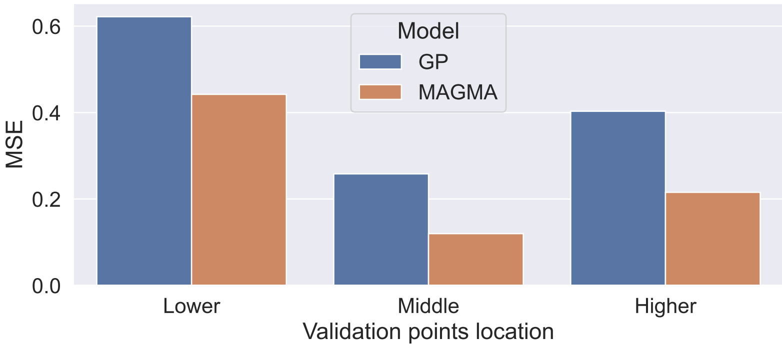

Fig. 3: Average performances of MAGMA and the baseline GP for the different locations of the unseen test observations.

the baseline is better when optimising specific hyperparameters for each profile, the performance of MAGMA is improved when using the same hyperparameters across the whole dataset. The influence of the hyperparameters setup can also be seen in Figure 2. It appears that the baseline trained in the Common HP setting overestimates its confidence intervals, explaining its 100% CIC_{95} value. As profiles vary in shape, using common hyperparameters tends to give the GP large variance, overly broadening the confidence intervals for each profile. Conversely, the Distinct HP setting makes both the GP and MAGMA overfit on the given test observations, deteriorating their predictions on unseen test observations. In the Common HP setting, MAGMA provides the most precise predictions while keeping 94.67% of unseen points in the predicted 95% confidence interval overall.

The results presented in Figure 3 also indicate that observations located lower in the atmosphere are more challenging to predict for both models, while observations removed in the middle of a profile are the easiest to estimate. One possible reason is that the pressure scale places observations at low altitudes further from one another. Given that the performances of both models deteriorate as they move further from known observations, this larger span of values is harder to predict accurately.

6 Conclusion

GPs are promising models for atmospheric profile extrapolation. The first step towards their practical application is to explore how they can leverage data from multiple profiles to improve their prediction accuracy. This paper examined the possibility of applying the MAGMA algorithm for this task. We compare MAGMA to a single GP baseline trained across profiles. Ultimately, we demonstrate that MAGMA outperforms the baseline, offering more precise and credible predictions by leveraging information from the whole dataset for each prediction.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

These findings represent preliminary results, making a first step towards a concrete application of GPs to spatial aeronomy. Future enhancements for our models include incorporating additional features as covariates, adopting more advanced preprocessing techniques, considering error measurements in a heteroskedastic model and refining the kernel to make it more specific to the task at hand. Another potential next step would be to use MagmaClustR [12], a variant of the MAGMA algorithm that may reduce prediction uncertainty by performing clustering on profiles. These enhancements could establish GPs as highly effective models for spatial atmosphere analysis.

Acknowledgements

This project has received funding from the Belgian Science Policy Office (Belspo) in the frame of the Brain.be program under grant agreement B2/233/P2/VAMOS.

References

[1] L. Trompet, Y. Geunes, T. Ooms, et al. Description, accessibility and usage of SOIR/Venus Express atmospheric profiles of Venus distributed in VESPA (Virtual European Solar and Planetary Access). *Planetary and Space Science*, 150:60–64, 2018.

[2] Arthur Leroy, Pierre Latouche, Benjamin Guedj, and Servane Gey. MAGMA: inference and prediction using multi-task Gaussian processes with common mean. *Machine Learning*, 111(5):1821–1849, 2022.

[3] Gustau Camps-Valls, Valero Laparra, Fernando Mateo-Jiménez, et al. A Survey on Gaussian Processes for Earth Observation Data Analysis. *IEEE Geoscience and Remote Sensing Magazine*, 2015.

[4] Gustau Camps-Valls, Dino Sejdinovic, Jakob Runge, et al. A perspective on Gaussian processes for Earth observation. *National Science Review*, 6(4):616–618, 2019.

[5] Suzanne Aigrain and Daniel Foreman-Mackey. Gaussian Process Regression for Astronomical Time Series. *Annual Review of Astronomy and Astrophysics*, pages 329–371, 2023.

[6] Elyar Sedaghati, Henri M. J. Boffin, Ryan J. MacDonald, et al. Detection of titanium oxide in the atmosphere of a hot Jupiter. *Nature*, 549(7671):238–241, 2017.

[7] N. P. Gibson, S. Aigrain, S. Roberts, et al. A Gaussian process framework for modelling instrumental systematics: application to transmission spectroscopy. *Monthly Notices of the Royal Astronomical Society*, 419(3):2683–2694, 2012.

[8] Emily K Pass, Nicolas B Cowan, et al. Estimating dayside effective temperatures of hot Jupiters and associated uncertainties through Gaussian process regression. *Monthly Notices of the Royal Astronomical Society*, 489(1):941–950, 2019.

[9] Edwin V Bonilla, Kian Chai, and Christopher Williams. Multi-task Gaussian Process Prediction. In *Advances in Neural Information Processing Systems*, volume 20, 2007.

[10] Edwin V. Bonilla, Felix V. Agakov, and Christopher K. I. Williams. Kernel Multi-task Learning using Task-specific Features. In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*, pages 43–50. PMLR, 2007.

[11] Carlos Ruiz, Carlos M. Alaíz, and José R. Dorronsoro. A survey on kernel-based multi-task learning. *Neurocomputing*, 577:127255, 2024.

[12] Arthur Leroy, Pierre Latouche, Benjamin Guedj, and Servane Gey. Cluster-Specific Predictions with Multi-Task Gaussian Processes. *Journal of Machine Learning Research*, 24(5):1–49, 2023.