Goldschmidt 2023 Abstract
https://doi.org/10.7185/gold2023.14355

Developing a Random Forest Classifier on Isotope Signatures of Methane from Different Sources

JIAWEN LI[1], JACOB BORTNIK[2], EDWARD D YOUNG[2] AND WIL LEAVITT[1]

[1]Dartmouth College

[2]University of California, Los Angeles

Presenting Author: jiawen.li.gr@dartmouth.edu

The discovery of methane on several extraterrestrial planetary bodies has led to the speculation that it could be a potential indicator of extraterrestrial life[1,2]. It is also a very important component of the carbon cycle on Earth and it has a large impact on the climate. Therefore, it is advantageous to trace methane of different origins for constraining global geochemical cycles, and for devising signatures of extraterrestrial life. In recent studies, the multiply substituted isotopologue ('clumped' isotope) signatures of methane ($D^{13}CH_3D$, $D^{12}CH_2D_2$) have been proposed as a tool for distinguishing methane of different origins[2-4]. Combined with bulk isotope signatures ($d^{13}C$, dD), they can provide useful information on the formation and destruction of methane. However, the boundaries that distinguish each type of methane are arbitrarily drawn in bulk and clumped isotope phase space, and the efficacy of the isotope clumping as opposed to bulk isotope ratios alone is not quantified. To address these problems, we are developing a random forest classifier to categorize methane into three sources (thermogenic, abiotic, and biogenic) using four isotope, or isotopologue, tracers ($d^{13}C$, dD, $D^{13}CH_3D$, $D^{12}CH_2D_2$) as the input features. The training and testing sets include published and unpublished methane isotopic data covering a range of experimental and environmental samples. The classifier is optimized by searching for the combination of tunable parameters that gives the best performance. The importance of each input feature during classification is quantified by the mean of the accumulation of impurity reduction[5]. Results suggest that the four input features are equally important in the classification. This study contributes to developing a more mathematical classification for methane gases, and the classifier will keep evolving as more data are available.

References

[1] M. A. Thompson, J. Krissansen-Totton, N. Wogan, M. Telus, J. J. Fortney, *Proceedings of the National Academy of Sciences*. **119**, e2117933119 (2022).

[2] E. D. Young *et al.*, *Geochimica et Cosmochimica Acta*. **203**, 235–264 (2017).

[3] D. A. Stolper *et al.*, *Geochimica et Cosmochimica Acta*. **161**, 219–247 (2015).

[4] D. T. Wang *et al.*, *Science*. **348**, 428–431 (2015).

[5] L. Breiman, *Machine Learning*. **45**, 5–32 (2001).