EGU21-8813
https://doi.org/10.5194/egusphere-egu21-8813
EGU General Assembly 2021

Sharing transferable methods in environmental data science: A Fuzzy changepoint approach to numerical model evaluation over Greenland.

Michael Hollaway[1], Peter Henrys[1], Rebecca Killick[2], Amber Leeson[3,4], and John Watkins[1]
[1]UK Centre for Ecology & Hydrology, Lancaster, United Kingdom
[2]Department of Maths and Statistics, Lancaster University, Lancaster, United Kingdom
[3]Lancaster Environment Centre, Lancaster University, Lancaster, United Kingdom
[4]Data Science Institute, Lancaster University, Lancaster, United Kingdom

Numerical models are essential tools for understanding the complex and dynamic nature of the natural environment and how it will respond to a changing climate. With ever increasing volumes of environmental data and increased availability of high powered computing, these models are becoming more complex and detailed in nature. Therefore the ability of these models to represent reality is critical in their use and future development. This has presented a number of challenges, including providing research platforms for collaborating scientists to explore big data, develop and share new methods, and communicate their results to stakeholders and decision makers. This work presents an example of a cloud-based research platform known as DataLabs and how it can be used to simplify access to advanced statistical methods (in this case changepoint analysis) for environmental science applications.

A combination of changepoint analysis and fuzzy logic is used to assess the ability of numerical models to capture local scale temporal events seen in observations. The fuzzy union based metric factors in uncertainty of the changepoint location to calculate individual similarity scores between the numerical model and reality for each changepoint in the observed record. The application of the method is demonstrated through a case study on a high resolution model dataset which was able to pick up observed changepoints in temperature records over Greenland to varying degrees of success. The case study is presented using the DataLabs framework, demonstrating how the method can be shared with other users of the platform and the results visualised and communicated to users of different areas of expertise.