CORPUS PUBLISHERS

Environmental Sciences and Ecology: Current Research (ESECR)

Volume 2 Issue 4, 2021

Article Information

Received date : May 06, 2021

Published date: June 02, 2021

*Corresponding author

Caio Atila Pereira Sena, Laboratory for Environmental Satellites Applications – Federal University of Rio de Janeiro, Brazil

Keywords

Meteorology; Topological data; Earth's system; Big data's; Velocity

Abbreviations

TDA: Topological Data Analysis;

MGSTD: Morse Graph Method for Stochastic Time-Series Data;

PH: Persistent Homology;

SVM: Support Vector Machine;

Tcs: Tropical Cyclones

Mini Review

Topological Data Analysis applications in Meteorology

Caio Atila Pereira Sena[1]* and Jose Ricardo de Almeida Franca[2]

[1]Laboratory for Environmental Satellites Applications, Federal University of Rio de Janeiro, Brazil

[2]Full Professor, Earth Sciences Institute of Federal University of Rio de Janeiro, Brazil

Abstract

Meteorology is currently built on a variety of measurements from several sources. The means to gain insight over the amount of available data are subject to intrinsic characteristics of the data. In this context, Topological Data Analysis (TDA) tools offer the possibility to overcome those limitations and harness a huge amount of data. Due to its roots in Group Theory, TDA can be intimidating and is yet underexplored in Meteorology. The current review aims to prospect some of TDA's capabilities in Atmospheric Sciences and provide insights on possible applications.

Introduction

Due to the intricacy of the Earth's system and the vast quantity of variables involved in weather and climate phenomena, meteorological research imposes quests that transcend traditional disciplinary boundaries [1, 2]. Climate and weather research mainly hinge on the availability of physical data. Instrumentation and numerical modelling provide a quantitative and objective mean of describing the current state of the atmosphere [3]. For that reason, climate science is one of the most data-rich domains in terms of volume, velocity, and variety (commonly referred to as Big Data's 3Vs) [4]. Whether produced by numerical simulations or gathered through remote sensing and in-situ measurements, the resulting data is often noisy and cumbersome, which restrains the analysis to techniques insensitive to those misfortunes. Additionally, the frequent and continuous development of new instruments leads to data heterogeneity in long-term studies [5], which may be unfavourable to some analysis tools. Earth sciences bring forth research questions related to the evolution of events in space and time. For this reason, meteorological data inevitably presents a spatiotemporal nature. This implies that measurements close in time and space tend to be highly correlated or similar, which means that the data tend to be smooth, blurring the feature boundaries. Therefore, some pattern mining methods that make implicit or explicit independence assumptions about the input data will most likely fail to track these features [6]. Some of the limitations described previously should be overwhelmed with the adoption of Topological Data Analysis (TDA) tools to explore the data. TDA offers a significant advantage over more conventional tools based on cluster analysis, as it focuses on global properties like the shape and connectivity of the data [7]. Primarily, Topological Data Analysis is an assemblage of statistical methods that and structure in data [8].

Patterns in Meteorology

In the atmosphere, clouds present itself usually isolated or bundled together. The patterns described by their organisation are often meaningful, such that some atmospheric phenomena are distinguishable, when viewed from above, based solely on the clouds features (i.e. temperature, type, content) and formation [9]. For example,in hurricanes the clouds form a unique circular structure around the eye, similar to a doughnut surface, making it very distinct from other types of storm. When it comes to surface studies, it is the geometry, organisation and cover of the surface structures that dictates the interaction with the atmosphere locally. For instance, cities with closely built high constructions and surface mostly covered with impervious materials of high heat capacity tend to experience higher temperature than the surrounding rural hinterlands [10]. The constraints described above indicate that traditional data mining tools involving linear separation may not withstand the limitations imposed by the aspects of physical measurements as that linear decision boundaries are often not sufficient to classify the data patterns with high accuracy [11]. Opportunely, the approximate shape of cloud as well as surface cover and texture can be meaningful features to describe the state of the atmosphere and local meteorological patterns from the TDA perspective. Recently, meteorological time-series data was examined from the viewpoint of dynamical systems with the aid of Morse Graph method for Stochastic Time-Series Data (MGSTD) [12]. The outcomes of the study illustrate an extension of such dynamical time-series analysis, which can capture unstable dynamics and attractors (i.e. characteristic behaviour in dynamic systems). Another TDA tool, named Persistent Homology (PH) was performed over five dimensional dataset of buoy measurements containing around 180000 data points [13]. The filtration process made it viable to detect anomalous changes in the measurements of wind components through the chain graphs, that are related to periods of El Niño occurrence. PH was also applied together with machine learning to identify atmospheric rivers in the west coast of the United States [14]. The data used for that was the integrated water vapour from climate reanalysis. The main challenge in this case was to construct the input to a Support Vector Machine (SVM) that would perform the class cations in a threshold-free form. PH was used then to mark the exact moment of birth of the complete structure of an atmospheric river, and the descriptor of these states were then fed to the SVM. Exploiting remotely sensed data, PH was performed to track H1 topological invariants in brightness temperature field over Tropical Cyclones (TCs) [15]. Those features were related to the TC diurnal cycle. Through that, the authors were able to identify the behaviour and quantify diurnal pulses in Hurricane Felix in an automated threshold free approach.

Summary

Atmospheric research is currently a data rich domain. The ability to harness the large amount of available data is toughly affected by intrinsic aspects of the measurements: such as high dimensionality, presence of noise, strongly correlated variables, missing values and the great heterogeneity of origins. Still, some relevant meteorological characteristics are preserved in hidden data features of shape and connectivity. Therefore, TDA shows great potential for meteorological research for being resilient tom, the listed adverse properties of the data and for enabling a high dimensional analysis. Then, TDA tools may provide new insights and bring forth high dimensional relations among meteorological variables that were not yet perceived.

References

1. ZP Majdik, CA Platt, M Meister (2011) Calculating the weather: Deductive reasoning and disciplinary telos in cleveland abbe's rhetorical transformation of meteorology. Quarterly Journal of Speech 97(1): 74-99.
2. F Williamson (2021) Building a long-time series for weather and extreme weather in the Straits Settlements: a multi-disciplinary approach to the archives of societies. Climate of the Past 17(2): 791-803.
3. J Coleman, K Law (2015) Meteorology.
4. JT Overpeck, GA Meehl S Bony, DR Easterling (2011) Climate data challenges in the 21st century. science 331: 6018: 700-702.
5. NB Chang, K. Bai (2018) Multisensor Data Fusion and Machine Learning for Environmental Remote Sensing. CRC Press.
6. JH Faghmous and V Kumar (2014) A Big Data Guideto Understanding Climate Change: The Case for Theory-Guided Data Science. Big Data 2(3): 155-163.
7. J Murugan, D Robertson (2019) An introduction to topological data analysis for physicists: From lgm to frbs. arXiv preprint arXiv: 1904-11044.
8. L Wasserman (2018) Topological data analysis. Annual Review of Statistics and Its Application 5: 501-532.
9. M Bader, G Forbes, J Grant, R Lilley, A Waters, et al. (1982) Images in weather forecasting, a practical guide for interpreting satellite and radar images. Atmospheric Research 43(2): 203-203.
10. TR Oke (1982) The energetic basis of the urban heat island (Symons Memorial Lecture, 20 May1980). Quarterly Journal. Royal Meteorological Society 108(455): 1-24.
11. G Mountrakis, J Im, C Ogole (2011) Support vectormachines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66(3): 247-259.
12. H Morita, M Inatsu, H Kokubu (2019) Topological computation analysis of meteorological time-series data. SIAM Journal on Applied Dynamical Systems 18(2): 1200-1222.
13. B Rieck, H. Leitte (2014) Structural analysis of multivariate point clouds using simplicial chains. in Computer Graphics Forum 33: 28-37.
14. G Muszynski, K Kashinath, V Kurlin, M Wehner (2019) Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets. Geoscientic Model Development 12(2): 613-628.
15. S Tymochko, E Munch, J Dunion, K Corbosiero, R Torn (2020) Using persistent homology to quantify a diurnal cycle in hurricanes. Pattern Recognition Letters.