

## 分野融合のためのアルゴリズム構築試験

### Algorithm construction test for interdisciplinary science

\*梅村 宜生<sup>1</sup>、阿部 修司<sup>2</sup>、田中 良昌<sup>3,4</sup>、中野 慎也<sup>5</sup>、新堀 淳樹<sup>1</sup>、齊藤 泰雄<sup>3</sup>、南山 泰之<sup>6</sup>

\*Norio Umemura<sup>1</sup>, Shuji Abe<sup>2</sup>, Yoshimasa Tanaka<sup>3,4</sup>, Shin ya Nakano<sup>5</sup>, Atsuki Shinbori<sup>1</sup>, Yasuo Saito<sup>3</sup>, Yasuyuki Minamiyama<sup>6</sup>

1. 名古屋大学宇宙地球環境研究所、2. 九州大学国際宇宙天気科学・教育センター、3. 国立極地研究所、4. 情報・システム研究機構データサイエンス共同利用基盤施設、5. 統計数理研究所、6. 国立情報学研究所

1. Institute for Space-Earth Environmental Research, Nagoya University, 2. International Center for Space Weather Science and Education, Kyushu University, 3. National Institute of Polar Research, 4. Joint Support-Center for Data Science Research, Research Organization of Information and Systems, 5. The Institute of Statistical Mathematics, 6. National Institute of Informatics

情報・システム研究機構は、分野融合型科学社会の実現に向けて、統計数理研究所、国立極地研究所、名古屋大学、九州大学の研究者らによる先行検証チームを構築し、各研究データが示す事象間の関係を統計的手法から導く手法の構築と、その計算結果の科学的妥当性の検証を行っている。先行チームは、国立極地研究所が所有する、気候・環境変動等の時系列データ、海水温測定等の同時期多拠点採取による空間分布データ、鉱物成分等の度数分布データをサンプルとし、以下の結果を得た。

時系列データ間には積率相関係数算出を使い、例えば、2014年12月22日の、南極宗谷海岸スカルブスネスにて観測された太陽放射照度変動と同宗谷海岸スカーレン大池にて観測された紫外線量変動には0.9871、同太陽放射照度とスカーレン大池にて観測された気温変動には0.4040、同太陽放射照度と南極昭和基地にて観測された地磁気変動には-0.0024の相関を得た。この例では異なる日時にも同程度の相関があることから、科学的には、太陽から放射照度として地球表面に流入した熱量は、紫外線量とは高い相互作用関係にあるものの、気温や地磁気変動との関係は低いことが推測される。また、度数分布データ間にはEarth Mover's Distance(EMD)を使い、例えば、1986年に南極のあすか基地とやまと山脈で採取した隕石の成分には2.23と類似関係にあるもの、同じあすか基地で採取した隕石であっても1028.05と類似しないものが見つかり、分析値の目視確認からもそのような分布と関係にあることを確認した。

また、隕石の例では、類似した試料の空間的特性、いわゆる群集合の空間的なずれは確認されなかったが、仮に空間的なずれが発見された場合、その特定地点では気温変動等に特異な事象は発生していないかの抽出も可能である。つまり、類似関係がブラインドされた状態からデータの比較、関連付け、詳細な区分化を実現したことは、関係が未知の状態にあってもその関係性を解き明かす単純な照合型システムとして機能するに留まらず、データ羅列や扱う分野が異なっても科学的な関係の究明に強く迫る分野融合型システムとしても機能することが言える。

実際にはデータの突合せはマトリックス状で行われ、関係性を網羅的に把握できる状態にある。計算結果はデータベースに格納されており、数十億件の結果に耐え得るための準備も進めている。また、メタデータを介して実際の研究データを取得する、研究機関および大学を横断する仕組みを採用している。2020年度は、データ点数や適用アルゴリズムを増やしたり、文献データや医薬分析データ等の度数分布データにも拡張したりするなどして、同分野や隣接分野から全くの異分野に至る知識の総合バーチャルネットワークを築き、公開し、共同利用機関法人による各研究機関および大学への研究基盤提供の使命を果たしながら、分野融合型科学社会の実現に寄与していく予定である。

キーワード：データサイエンス、分野融合、データ駆動型科学

Keywords: Data Science, Interdisciplinary Science, Data-Driven Science