

令和 6 年 6 月 20 日現在

機関番号：12703

研究種目：挑戦的研究（萌芽）

研究期間：2020～2023

課題番号：20K21792

研究課題名（和文）線形計画法と深層学習による人工衛星データの復元と解析

研究課題名（英文）Recovery and Analysis of Satellite Data via Linear Programming and Deep Learning

研究代表者

土谷 隆（Tsuchiya, Takashi）

政策研究大学院大学・政策研究科・教授

研究者番号：00188575

交付決定額（研究期間全体）：（直接経費） 3,900,000円

研究成果の概要（和文）：人工衛星 Geotail のプラズマ粒子の観測データの深層学習による復元を試みた。プラズマ粒子は入射方向について緯度7区分、経度16区分、エネルギー32区分に仕分けされ、24時間中8時間は12秒ごとに $7 \times 16 \times 32 = 3584$ 個のビンの3次元ヒストグラムの情報が送られる。16時間は緯度方向の和をとった512個のビンの2次元ヒストグラムの情報が送られる。2次元ヒストグラムから3次元ヒストグラムの深層学習による復元を試み、U-netとRes-netを組合わせてVanilla GANsにE-tスペクトログラムのデータの識別器を追加したも
のによって、一定のレベルでの復元が可能であることを示した。

研究成果の学術的意義や社会的意義

人工衛星 Geotail のプラズマ粒子観測データについて、深層学習を用いて、2次元ヒストグラムデータから3次元ヒストグラムデータを復元することを試みた。2次元ヒストグラムは全体の2/3を占めており、本研究成果により、今後これらのデータから3次元ヒストグラムを復元して、興味深い現象とされる、磁気リコネクションが起きている様子などを知ることができると考えられる。

研究成果の概要（英文）：We developed deep learning neural network for recovery of plasma particle observation data from satellite Geotail. Observed plasma particles are categorized into 7 and 16 directions in latitude and longitude, respectively, and 32 levels in their energies. 3d histogram data of plasma particles in $7 \times 16 \times 32 = 3584$ bins based on aforementioned categorization is transmitted to a ground station every 12 seconds for 8 hours a day. For the rest 16 hours, 2d histogram data of 16×32 obtained by marginalizing latitude counts is sent due to transmission limit. Human observers analyzes the data by using a chart called E-t spectrogram computed from 3d histogram. We developed a system to recover 3d histogram from marginalized 2d histogram by deep learning neural networks. Utilizing U-net, Res-net, Vanilla GANs with some modification, we were able to obtain fairly good estimation of E-t spectrogram.

研究分野：統計数理・数理工学

キーワード：人工衛星データ 深層学習 プラズマ粒子 信号復元

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

1992年に打ち上げられた人工衛星 Geotail は、磁気圏尾部のプラズマ粒子を観測するための衛星である。衛星により観察されたプラズマ粒子は緯度方向7区分、経度方向16区分、そしてエネルギー32区分に分類され、24時間中8時間の間、12秒に一度3次元($7 \times 16 \times 32 = 3584$ 個)のビンから構成されたヒストグラムの情報が送られるものの、16時間は通信容量の関係で、緯度方向7方向について和をとった512個のビンからなるヒストグラムの情報が、3次元ヒストグラムの平均値と分散・共分散行列の9個の数値情報とともに送られている。

2. 研究の目的

深層学習と線形計画法を用いることで圧縮された2次元ヒストグラムデータから3次元ヒストグラムを復元することを目的とした。大量の学習データが存在するため、これを利用して、正解が分かっている3次元ヒストグラムデータが1日8時間分約10年間にわたり利用可能であるため、これを学習データに利用して、2次元ヒストグラムから3次元ヒストグラムを復元し、磁気リコネクションなどの惑星科学的に興味深い現象を観測する。

3. 研究の方法

研究当初は、以前の研究に基づいてヒストグラムの整数性に着目して、線形計画法を用いた復元を行う計画であったが、深層学習の急速な進展を踏まえて、深層学習による復元に焦点を絞った研究を行った。研究者は、3次元ヒストグラムのプラズマ分布データをE-tスペクトログラムというチャートで観察する。E-tスペクトログラムは粒子の入射方向によってNorthward, Equationward, Southward, (以上上下方向), Dawnward, Sunward, Duskward, Anti-Sunward (以上水平方向)の7つの帯状の色分けされたグラフが縦に7つ並べられたものである。後に説明に用いる図1は2枚のE-tスペクトログラムを並べたものである。各帯は以下のように作られている。帯はある特定の方向からの入射粒子数をエネルギーレベルごとに時系列で観測したものである。1時点は縦に並んだ小さな32個の正方形からなり(ほとんど直線のようにみえる)、その一つ一つが下から上に大きくなるエネルギー区分を表し、正方形の色で各レベルの粒子数を(近似的に)表現している。すなわち、1時点での観測値は色のついた7本の線分で表現され、時間の経過とともに、7つの彩色された帯状のグラフで入射方向ごとのプラズマ分布データが表現される。これは、3次元ヒストグラムの情報から容易に計算できる。

人間が観察するという立場からは、E-tスペクトログラムを復元できれば良いという考え方もあり、そこで、2次元ヒストグラムからE-tスペクトログラムを直接復元するというアプローチと3次元ヒストグラムを復元した上でE-tスペクトログラムを復元する、というアプローチが考えられるが、本研究では、3次元ヒストグラムを復元してE-tスペクトログラムを復元するというアプローチを採った。

プラズマの状態としては、プラズマの圧力/磁場の圧力として定義されるパラメータの値によって、ローブ(< 0.01)/プラズマ境界層($= 0.01 \sim 1$)/プラズマシート(> 1)の3つがあり、興味深いのはプラズマ境界層であるため、この復元を中心に研究を進めた。

用いたニューラルネットは、(i)3層ニューラルネットワーク、(ii)多層ニューラルネットワーク、(iii)U-net、(iv)GANsなどで、これらを適宜組み合わせ用いて実験を行い比較をした。学習のためのアルゴリズムとしては、Adamを用いた。

プラズマ境界層での磁気リコネクションを捉えた典型例として、1994年1月14日15時30分から16時10分をテストデータとして選び、学習データとしては、1993年10月から24ヶ月分のプラズマ境界層のデータ(265,168サンプル)を用いて学習した。

4. 研究成果

構築したニューラルネットの性能について、定量的な評価と定性的な評価を行った。定量的な評価としては、正答率として

$$\left(\left(\text{ビン毎の推定されたヒストグラムと実際のヒストグラムの小さい方の値} \right) / \left(\text{総粒子数} \right) \right)$$

を指標として用いた(この量は0と1の間で正解だと1となる)。

一番性能が良かったのは、U-netとResNetを組合わせてVanilla GANsによって学習し、E-tスペクトログラムのデータの識別器を追加したモデルによる学習であるため、このモデルの結果について述べる。このモデルの平均的な正答率は0.694、最高で0.894、最低で0.389であった。

図1にターゲットである、1994年1月14日15時30分から16時10分についての正しいE-tスペクトログラムと(左)と2次元ヒストグラムから復元したE-tスペクトログラム(右)を示す。正答率はプラズマ境界層に入ってから概ね80%前後であるが、それ以上の改良は中々困難であった。

より詳細な検討を行うと、EquationwardとAnti-Sunwardの低エネルギー粒子のランダムな

分布,そして Equationward と Sunward の高エネルギー粒子の分布は比較的良く再現されていると考えられる。磁気リコネクションが起こっていると考えられる,15時39分ごろについては,実際には Northward に見られる中エネルギー粒子の集中を再現するには至っておらず,このような点を改善することが今後の検討課題となる。

現時点での研究の限界としては,学習データが少ないことが挙げられる。より性能を向上するためには,特に,プラズマ境界層に入る前後の尖った分布に対応する学習データが必要となる。これを既存のデータから合成するなどの工夫が必要となる。また,現時点での結果を用いて現在あるデータについて解析を行い,興味深い現象が起こっている部分を抜き出して,惑星科学研究者に呈示して評価してもらうことも今後の課題となる。

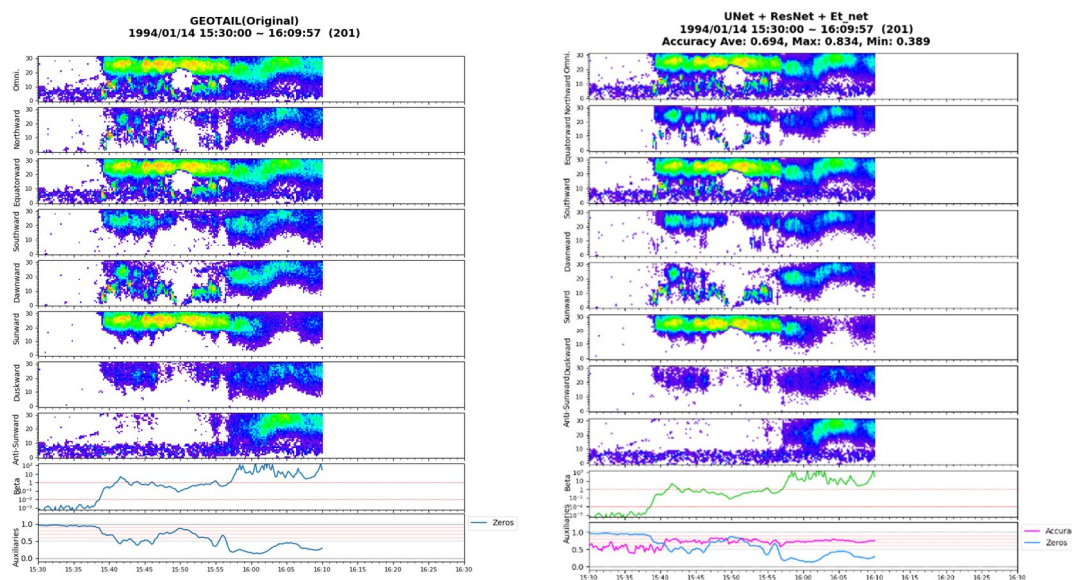


図1. 1994年1月14日15:30から16:30にかけてのE-tスペクトラム。
左図が実データ,右図が2次元ヒストグラムから復元したもの。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 柿原 聡, 土谷 隆, 上野 玄太	4. 巻 470
2. 論文標題 ニューラルネットワークを用いた人工衛星の観測データの復元	5. 発行年 2024年
3. 雑誌名 統計数理研究所共同研究レポート	6. 最初と最後の頁 161-173
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 柿原 聡, 土谷 隆, 上野 玄太
2. 発表標題 ニューラルネットワークを用いた人工衛星の観測データの復元
3. 学会等名 統計数理研究所共同研究集会 最適化：モデリングとアルゴリズム（於 統計数理研究所，2023年3月17日）
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	上野 玄太 (Ueno Genta) (40370093)	大学共同利用機関法人情報・システム研究機構（機構本部施設等）・データサイエンス共同利用基盤施設・教授 (82657)	
研究分担者	田中 未来 (Tanaka Mirai) (40737053)	統計数理研究所・数理・推論研究系・准教授 (62603)	
研究分担者	池上 敦子 (Ikegami Atsuko) (90146936)	成蹊大学・その他部局等・客員研究員 (32629)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------