

論文要旨のテキストマイニング分析による 天文衛星の測定データを基にした研究動向の把握

- X線天文衛星「すざく」を題材にして -

中村 匡佑^{*1}, 水上 祐治^{*1}, 大畠 昭子^{*2}, 本多 啓介^{*3}, 中野 純司^{*3}

A Study on Grasp of Research Trend based on Measurement Data of Astronomical Satellite by Text-Mining Analysis of Abstract of Thesis: Using the Theses of X-ray Astronomical Satellite "SUZAKU"

NAKAMURA Kyosuke^{*1}, MIZUKAMI Yuji^{*1}, OHATA Akiko^{*2}, HONDA Kesuke^{*3}, NAKANO Junji^{*3}

ABSTRACT

Various observation data obtained from artificial satellites and explorers are expected to be widely used in all industry, government and academia. Meanwhile, since the space development project is large-scale activity, its contribution is required to be explained in all sessions including the National Assembly. In addition, it is said that research capability in Japan has decreased since the mid-1990s. Therefore, the improvement of research capability is urgently required and the necessity of objective explanation materials on academic contribution is increasing. This paper aims to derive a method to prepare objective information which is an auxiliary explanation of academic contribution to verify its effect. In the analysis of this paper, text mining analysis is applied to the abstract of the thesis to extract research trends per each year for 20 years and the effect of the method is verified. The academic papers related to the X-ray astronomical satellite "SUZAKU" that were collected by bibliographic databases, were used in this study.

Keywords: text mining, abstract, research trend, exploration satellite, institutional research

概要

宇宙開発プロジェクトは大規模な事業であるため、その貢献について、国会をはじめあらゆる場面で説明が求められる。特に、人工衛星・探査機から得られる様々な観測データは、産官学すべてにおいて広く活用が期待される。また、日本の研究力は1990年代中盤から低下しているとされ、研究力の向上が急務となっていることがあり、学術的貢献に関して客観的な説明資料の必要性が高くなっている。そこで本稿は、学術的貢献の補助的説明となる客観的情報を用意するための手法を導出すること、そして、その効果を検証すること目標とした。分析では、論文の要旨にテキストマイニング分析を施して年代別の研究トレンドを抽出、そして、その手法の効果を検証した。分析対象は、書誌データベースを用いて収集した

doi: 10.20637/JAXA-RR-19-006/0005

* 2019年12月17日受付 (Received December 17, 2019)

^{*1} 日本大学 生産工学部 (Nihon University, College of Industrial Technology)

^{*2} 宇宙航空研究開発機構 宇宙科学研究所 (Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency)

^{*3} 統計数理研究所 (The Institute of Statistical Mathematics)

X線天文衛星「すざく」関連の論文である。

1. はじめに

人工衛星・探査機から得られる様々な観測データは、広く活用が期待されている(2018)¹⁾。一方、水上ら(2019)²⁾は、宇宙開発プロジェクトは大規模な事業であるため、その貢献について、国会をはじめあらゆる場面で説明が求められているとしている。また、文部科学省(2017)³⁾は、1990年代中盤からの日本の国際的な研究力低下を指摘しており、研究力の向上が急務となっていることがあり、学術的貢献に関して客観的な補助的説明資料の必要性が高くなっている。そこで本稿は、学術的貢献の補助的説明となる客観的情報を用意するための手法を導出すること、そして、その効果を検証することを目標とした。

分析では、論文の要旨にテキストマイニング分析を施して年代別の研究トレンドを抽出、そして、その手法の効果を検証した。分析対象の論文は、X線天文衛星「すざく」関連として、書誌データベースを用いてそれら論文群を収集した。論文群を収集後、3段階で分析を行った。まず、テキストマイニングの単語抽出分析を行い、品詞別に「出現頻度の高い名詞と固有名詞の抽出」を行った。次に、テキストマイニングの共起ネットワーク分析を用いて、「すざく衛星の関連研究の研究対象と手法」を析出した。そして、テキストマイニングの対応分析を用いて「年代別の研究動向」を示した。なお、テキストマイニング分析では、KH Coder⁴⁾を用いた。

2. 関連研究レビューと本稿の目的

2.1. 研究開発成果オープン化モデルと本稿の目的

海老沢(2018)は、宇宙科学研究所 (ISAS: Institute of Space and Astronautical Science) のポリシーは、「公開データが広く利用されることが科学の進歩につながる」として、人工衛星・探査機から得られる観測データは、様々な分野での活用が期待されているとしている。図 1 に JAXA (Japan Aerospace Exploration Agency) の研究開発成果オープン化モデルを示す。なお、この図は、JAXA 第一宇宙技術部門 衛星利用推進 Web ページ⁵⁾をもとに筆者が作成したものである。

この研究開発成果オープン化モデルでは、まず、人工衛星・探査機から得た測定データは、生データとして蓄積される。しかし、この生データは、例えば機械的信号であること、ノイズが含まれていること、不完全なデータであること等があり、そのままでは広く活用することが難しい。そこで、この生データに対して、モデル化、複合・同化、アルゴリズム等の整理・解析の処理を施し、広く活用可能なプロダクト、データセット等の情報に加工して提供するものであり、これら情報は、データアーカイブとして JAXA 内で蓄積されている。データアーカイブとして JAXA 内で蓄積されている情報は、産学官 (産業界 (民間企業)、学术界 (教育・研究機関)、官公庁 (国・地方公共団体)) で活用され、企業戦略、派生研究、政策立案等に利用されている。また、産学官で派生した新たな知識は、JAXA へ戻り(掘り起し)、新たな活動の指針となっている。

本稿の目的は、学术界 (教育・研究機関) への貢献度を客観的に測定する手法を導出すること、そして、その効果を検証することにある。

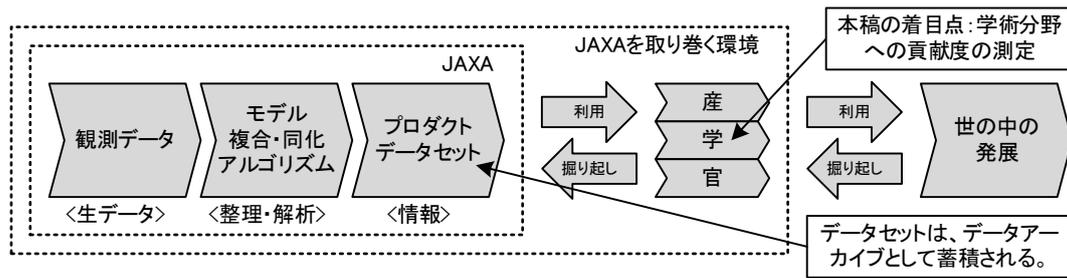


図 1 JAXA の研究開発成果オープン化モデル

JAXA 第一宇宙技術部門 衛星利用推進 Web ページをもとに筆者作成

2.2. 宇宙科学分野のデータアーカイブについて

NASA は 1990 年から X 線天文データアーカイブの開発・運用を開始⁶⁾、宇宙科学研究所/宇宙航空研究開発機構 (ISAS/JAXA) では 1997 年から宇宙科学データアーカイブの開発・運用を開始⁷⁾している。さらに、川上ら(2017)⁸⁾は、ISAS に蓄積されている歴史的に価値のある貴重な資料を含む技術資料の散逸を防止し、情報の共有化と有効活用を図ることを目的として、2007 年度より宇宙科学資料データベースシステム DARTS (Data ARchives and Transmission System)⁹⁾及びデータ登録を続けている。

宇宙科学分野では、これら観測データの活用促進を目的とした活動が盛んであり、DARTS に続く代表例として、2009 年度から 6 カ年計画でスタートした大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」(IUGONET: Inter-university Upper atmosphere Global Observation NETwork)¹⁰⁾が挙げられる。IUGONET は、複数の要因が複雑に絡み合う「超高層大気長期変動」のメカニズムを解明することを目的として、東北大学、名古屋大学、京都大学、九州大学、国立極地研究所の 5 機関が所有する多種多様な超高層大気データを統合して公開するプロジェクトである。林ら(2012)¹¹⁾によると、IUGONET プロジェクトでは、これまでに蓄積された多様な地上観測データに関するメタデータのデータベースを構築するとしており、このメタデータ・データベースによって、各機関・組織に分散して存在する観測データの横断的な検索・取得を可能にし、観測データの効率的な流通、さらには分野横断型の総合解析による新しい超高層大気研究を促進するとしている。また、田中ら(2012)¹²⁾は、IUGONET で開発された解析ソフトウェア UDAS (IUGONET Data Analysis Software)の概要を紹介しており、UDAS は、IDL (Interactive Data Language) で書かれた統合データ解析ツール TDAS (THEMIS Data Analysis Software suite) のプラグインソフトであり、TDAS の機能を利用することで、種類の異なる複数の時系列データの可視化・解析を容易にしている。さらに、小山ら(2012)¹³⁾は、SPASE コンソーシアムによって策定された SPASE データ・モデル/メタデータ・フォーマットをベースに、さらに拡張を施した IUGONET 共通メタデータ・フォーマットを策定した。そして、フリーのリポジトリ・ソフトウェアであり、デフォルトでは Dublin Core メタデータ・フォーマットのみを取り扱う DSpace に対し、IUGONET 共通メタデータ・フォーマットを取り扱えるようにカスタマイズを施している。また、堀ら(2012)¹⁴⁾は、IUGONET で策定・公開されたメタデータフォーマットと、そのフォーマットに沿って策定されたメタデータの登録・管理を行うシステムを構築したとしている。最後に小山ら(2014)¹⁵⁾は、IUGONET プロジェクトのメタデータ・データベースは広範な研究領域を対象としている為、専門分野外のユーザーにとって単語検索時における適切な検索語句の選択が容易でないことを指摘して、この問題を解決する連想検索の機能をシステムに付加している。

IUGONET 以外にも複数のプロジェクトがあり、北川ら(2014)¹⁶⁾は、地球惑星試料デポジトリ DREAM (Depository for references of earth and analytical materials)を整備、中川ら(2016)¹⁷⁾は、全天 X 線監視装置

表 1 研究評価の分類

系統	目的	分析手法
論文数調査	論文の生産性指数 研究活動の規模指数	単純集計
引用統計・分析	論文の消費指数 研究活動の品質指数	引用分析 共著分析 (謝辞分析) (共語分析) (共分類分析)

根岸・山崎(2001)と藤垣ら(2004)をもとに筆者作成

MAXI (Monitor of All-sky X-ray Image) の観測データを全世界の科学者が容易に利用できることを目的として、データ公開を進めており、2011年11月に開始したWebインタフェースを用いたオンデマンドデータの公開に続き、データを恒久的に後世に残すためのデータアーカイブの開発を進めている。さらに、高木ら(2016)¹⁸⁾は、科学衛星や探査機の運用に必要なとされる工学情報(テレメトリデータ、局設備からのデータなど)を一元的に収集管理し、また利用しやすい形で関係者に配布するための「科学衛星運用工学データベース」EDISON (Engineering Database for ISAS Spacecraft Operation Needs) の整備を進めている。そして、野口ら(2017)¹⁹⁾は、MRO (Mars Reconnaissance Orbiter) 搭載の MCS (Mars Climate Sounder) による観測データの利用を広めるための環境整備を進めている。

宇宙科学分野では、IUGONETをはじめ、観測データの活用促進を目的とした複数のプロジェクトがあり、情報の公開と共に細分化が進んでしまった。そこで、根本ら(2018)²⁰⁾は、資料の種類や機関を横断した資料情報の整理を目的として、デジタルオブジェクト識別子 (DOI: Digital Object Identifier) を用いた「サブジェクトリポジトリ」の構築を推進している。

2.3. 書誌データベースを用いた研究力評価について

根岸・山崎(2001)²¹⁾は、書誌データを用いた研究評価の方法には、文献の発表数を対象に分析する論文数調査の系統と、論文への引用度数や論文刊の引用による結合関係を対象に分析する引用統計・分析の2系統があるとしている。研究評価の分類を表1に示す。根岸ら(2001)、藤垣ら(2004)²²⁾、Mizukami et al.²³⁾は、前者の「論文数調査」の目的は、論文の生産性指数、研究活動の規模指数を求めることであり、主に単純集計の手法が用いられる。一方、後者の「引用統計・分析」の目的は、論文の消費指数、研究活動の品質指数を求めることであり、引用分析、共著分析等の手法が用いられるとしている。また、少数ではあるが、謝辞分析、共語分析、共分類分析等の手法が用いられる。なお、謝辞分析は、論文中に記されている謝辞が分析対象、そして、共語分析は、論文中に記されている複数の語の間の関係が分析対象、最後に共分類分析は、分野分類の共出現現象が分析対象である。

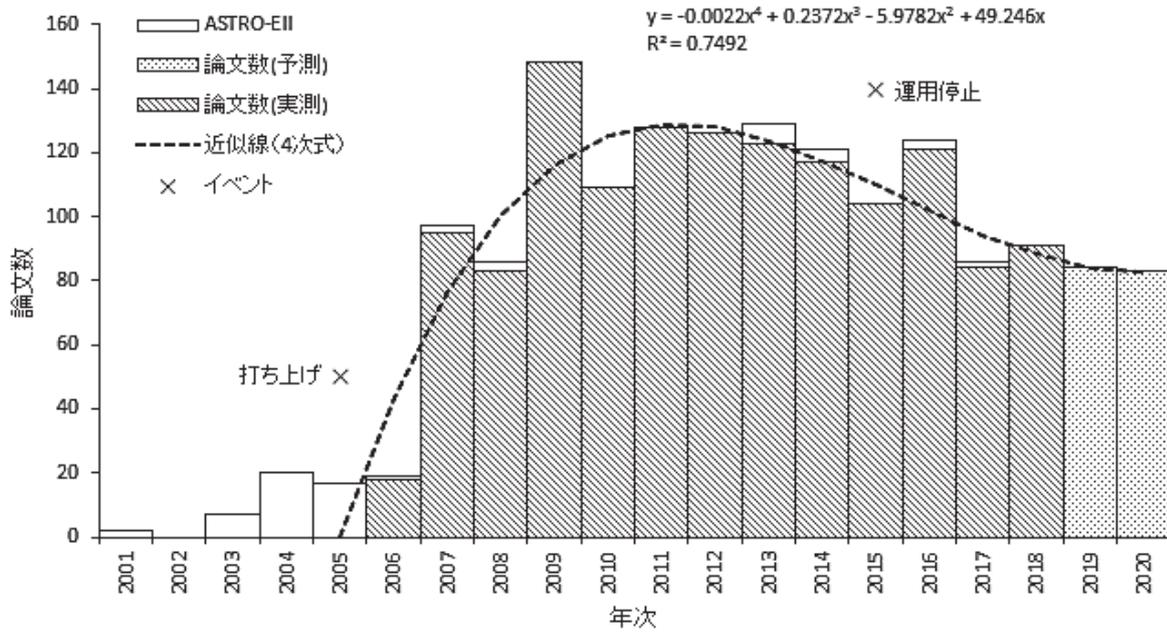


図 2 論文数の年次変化

3. 分析

3.1. 分析データ

分析対象の衛星は、X線天文衛星「すざく」である。この衛星は、遠距離にある天体のX線観測、宇宙の高温プラズマのX線分光観測等を目的としている衛星であり、2005年にM-Vロケット6号機にて内之浦宇宙空間観測所から打ち上げられ、10年間の運用後、2015年に観測を完了している。分析対象の衛星の選択においてこの衛星が選ばれた理由として、比較的最近の衛星であるので、現状の把握に適していると考えられること、10年間の運用期間があるので、時系列での分析に適していること、すでに運用が完了しているので、さらに測定データが増加することがなく、恒久的な分析が可能であることが挙げられる。

分析対象の論文は、Clarivate Analytics社の論文データベースWeb of Science – Core collection (以下、WOS)を用いて収集した。なお、WOSは、Impact Factorの算出に用いる書誌データベースであり、査読論文の網羅性が高いと考えられる。表2に論文の検索条件を示す。分析対象の論文は、トピックに”SUZAKU”が含まれる英語論文であり、2001年から2019年の18年間を対象とした。検索の結果、分析対象の論文は1419本であった。なお、検索条件のドキュメントタイプは、査読論文”Article”と”Review”のみとし、重複を避けるために発表論文等の査読なし論文”Proceedings”は対象外とした。

表 2 検索条件

WoS 検索条件 : (TS=(SUZAKU) OR TS=("ASTRO-E II") OR TS=("ASTRO-E 2") OR TS=("ASTROE II") OR TS=("ASTROE 2") OR TS=("ASTROEII") OR TS=("ASTROE2") OR TS=("ASTRO-EII") OR TS=("ASTRO-E2")) AND 言語: (English) AND ドキュメントタイプ: (Article OR Review)
索引=SCI-EXPANDED, SSCI, A&HCI, ESCI タイムスパン=全範囲

表 3 対応分析の設定

集計単位と抽出語の選択			対応分析の設定	
文章と見なす単位	H5		グラフの種類	抽出語 x 外部変数
条件	出現数	≧ 350	外部変数	論文誌名
	品詞	名詞のみ	差異が顕著な語を分析に使用	上位 100 語
上記条件での対象語数		150*1	バブルプロット(円の大きさ)	語の出現回数

*1: HKCoder の推奨値は 100-150 である。

図 2 にすぎく衛星の関連論文の年次変化を示す。なお、“ASTRO-EII”は、すぎく衛星を打ち上げる前のプロジェクト名である。また、JAXA の慣習により、衛星打ち上げ後はプロジェクト名が “SUZAKU” に変更になっている。本稿では、“ASTRO-EII”で検索された論文は、すぎく衛星の観測データを使用した論文ではないと判断して分析の対象外としている。実測の論文において、論文数は、打ち上げ翌年の 2006 年に 18 編の論文が発表されており、2009 年の 149 編が最も多くその後は減少傾向を示している。式 1 にすぎく衛星関連論文の論文数年次変化の近似線の式を示す。式 1 に当てはめると、2019 年、2020 年の論文数の予測は、84.078 編、82.770 編である。

3.2. 分析手法

論文の要旨にテキストマイニング分析を施して年代別の研究トレンドを抽出、そして、その手法の効果を検証する。論文の要旨は、論文の目的、手法、結論等の概要を簡潔にまとめたものであり、論文の特徴を示すものである。また、おおよそ 500 文字で構成されており、文字数の差が原因での分析結果の偏りが少ないと考えられる。分析は 3 段階で行った。まず、単語抽出分析を行い出現頻度の高い語句を抽出、次に、対応分析を施して年次別の研究の傾向を抽出した。最後に、すぎく衛星関連論文と新たに得られた他の衛星・宇宙望遠鏡の関連論文の関係を論文数比較する。

3.2.1. 単語抽出分析について

テキストマイニングの単語抽出分析を行い、品詞別に出現頻度の高い名詞と固有名詞の抽出を行う。分析ツールは、KH Coder Ver.3.Alpha.13k²⁴⁾を用いた。なお、テキストマイニング分析は、テキストデータに対して、自然言語解析の手法を使って、文章を単語（名詞、動詞、形容詞等）に分割し、それらの出現頻度

表 4 単語抽出分析の結果

#	名詞	出現数	#	名詞	出現数	#	固有名詞	出現数	#	固有名詞	出現数
1	x-ray	3350	11	Region	760	1	SUZAKU	1608	11	Compton	172
2	emission	2023	12	Galaxy	711	2	XMM-Newton	322	12	NuSTAR	162
3	observation	1535	13	Flux	660	3	AGN*	313	13	ICM*	155
4	spectrum	1505	14	Temperature	655	4	Galactic	285	14	Telescope	149
5	source	1197	15	Accretion	556	5	Chandra	276	15	Gamma	144
6	component	916	16	Gas	551	6	NGC*	271	16	kT	129
7	datum	894	17	Disk	530	7	Seyfert	243	17	Si	121
8	energy	888	18	Density	504	8	XIS*	240	18	Spectrometer	110
9	result	887	19	Absorption	496	9	SWIFT	211	19	MG	101
10	cluster	807	20	Time	489	10	SNR*	197	20	Imaging	100

* X線を用いた分析手法の短縮形表現である。

式 1 すぎく衛星関連論文の論文数年次変化の近似線の式 (2006-2018)

$$y = -0.0022x^4 + 0.2372x^3 - 5.9782x^2 + 49.246x \quad (R^2 = 0.7492)$$

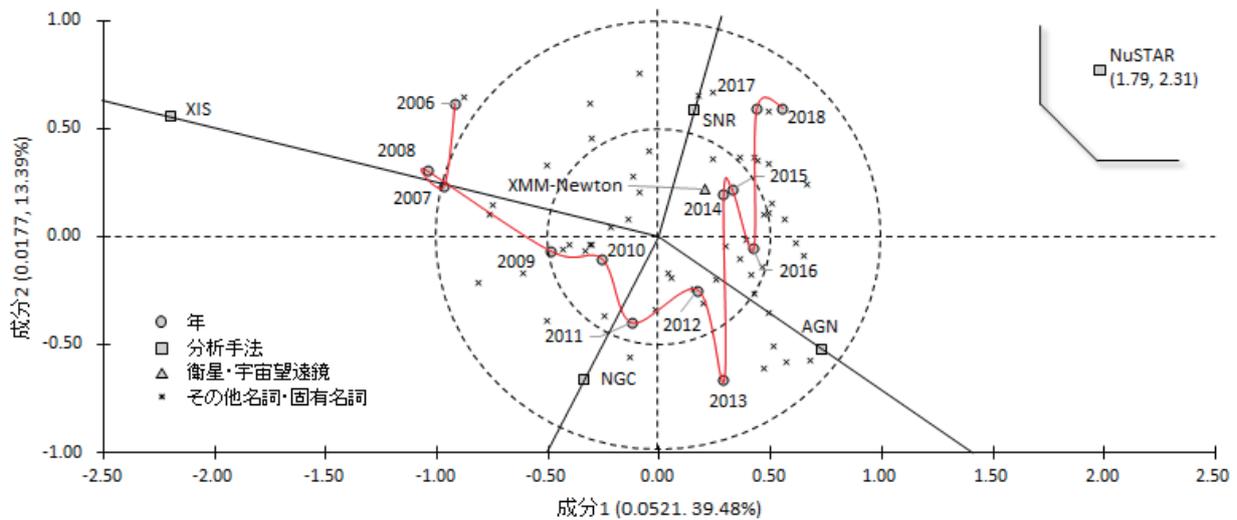


図 3 論文要旨情報の対応分析

や相関関係を分析することで有益な情報を抽出する手法である。那須川ら(2001)²⁵⁾は、テキスト・マイニングにおける分析対象は、基本的には名詞が中心であるとして、他の様々な技術を使い他の品詞を含める手法をまとめている。本稿では、分析対象を名詞と固有名詞と定めて分析した。また、1文字で抽出された語句は分析の対象外としている。

3.2.2. 対応分析について

すざく衛星関連論文の年代別の研究トレンドを示す。対応分析は、一般的にはコレスポンディング分析と言われ、クロス集計結果を2次元の散布図に展開する手法である。また、対応分析は、質的多変量解析の一種であり、数量化Ⅲ類と同等の分析が可能である。散布図は、掲載年に着目した場合、原点(0, 0)付近の年に含まれる語の分布に特徴がないこと、原点(0, 0)から遠いほどその分布に特徴があることを示す。表 3 にその対応分析の設定を示す。

3.2.3. すざく衛星と他の衛星・宇宙望遠鏡の論文数比較

すざく衛星関連論文と新たに得られた他の衛星・宇宙望遠鏡の関連論文の関係を年次比較する。対応分析で得られたすざく衛星に関連する他の衛星・宇宙望遠鏡に関して、論文数の比較を行いすざく衛星関連論文の論文数への影響を分析する。

3.3. 分析結果

3.3.1. 単語抽出分析の結果

表 4 に論文要旨情報に単語抽出分析を施した結果を示す。左側が出現頻度の高い名詞上位 20 位、右側が出現頻度の高い固有名詞上位 20 位である。左側の名詞においては、X線分析に関わる語句が多く示されている。左側の固有名詞では、衛星名の”SUZAKU”と”XMM-Newton”が上位 2 位、そして、”NuSTAR”が 11 位に入っている。なお、2 位の”XMM-Newton (X-ray Multi-Mirror Mission-Newton)”は、1999 年に欧州宇宙機関のギアナ宇宙センターからアリアン 5 ロケットで打ち上げられた X 線観測衛星である。そして、11 位の”NuSTAR (Nuclear Spectroscopic Telescope Array)”は、2016 年に NASA が打ち上げた高エネルギー X 線を観測するための X 線宇宙望遠鏡である。次の注目すべき点として、左側の固有名詞では、X線分析の分析手法である AGN が 3 位、NGC が 6 位、XIS が 8 位、SNR が 10 位、ICM が 13 位に入っている。

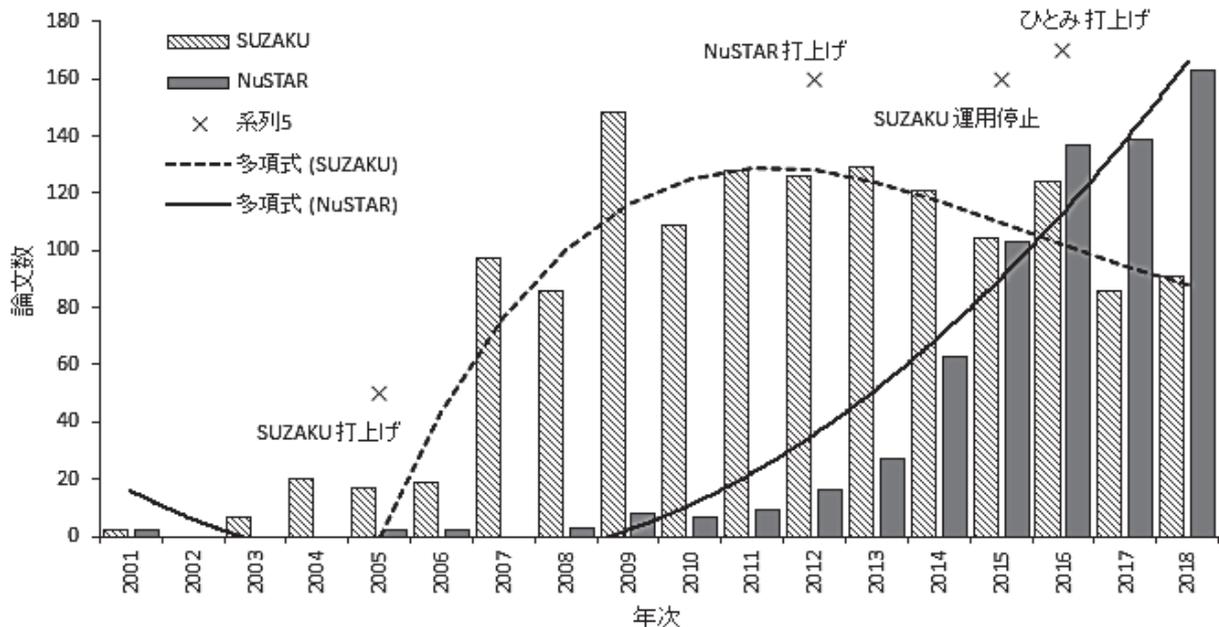


図 4 SUZAKU と NuSTAR の論文数の年次変化

3.3.2. 対応分析の結果

図 3 に論文要旨情報と論文発表年情報の対応分析の散布図を示す。この図は、第 1 成分と第 2 成分の散布図であり、全体の 52.87% (← 39.48% + 13.38%) を説明するものである。年次情報は、2006 年から 2018 年があり、分析手法として、XIS、NGC、AGN、SNR、関連の衛星・宇宙望遠鏡として、XMM-Newton, NuSTAR が示されている。また、中心からの距離を基にして、2006 から 2008 までの 3 年間は強い特性、2009, 2013, 2017, 2008 は特性、その他は、弱い特性が示されている。

3.3.3. すざく衛星と他の衛星・宇宙望遠鏡の論文数比較

図 3 の対応分析で強い特性を示した NuSTAR 宇宙望遠鏡を対象に分析を行う。図 4 にすざく衛星関連論文と NuSTAR 宇宙望遠鏡関連論文の論文数の年次別変化とその推移を示す。また、2016 年には、すざく衛星の後続衛星である ひとみ衛星の打ち上げの情報を示す。なお、ひとみ衛星は、2016 年 2 月 17 日に種子島宇宙センターから H-IIA ロケット 30 号機で打ち上げられたが、同年 3 月 26 日に通信が途絶して、4 月 28 日に運用を断念している。

すざく衛星関連論文の論文数 2009 年をピークとして、その後は減少傾向にある。一方、NuSTAR 宇宙望遠鏡関連論文は、打ち上げから 3 年の 2015 年には、すざく衛星関連論文の論文数に追いつき、その後も増加傾向がつづいている。式 2 に NuSTAR 宇宙望遠鏡関連論文の論文数年次変化の近似線の式を示す。

3.4. 考察

分析方法により年代別の特徴を 4 つのグループに分類することができる。まず、XIS (X-ray Imaging Spectrometer) グループがある。打ち上げ直後の 2006 年から 2010 年の 5 年間は、XIS 分析を中心とする

式 2 NuSTAR 宇宙望遠鏡関連論文の論文数年次変化の近似線の式 (2006-2018)

$$y = 1.1722x^2 - 13.441x + 28.181 \quad (R^2 = 0.9508)$$

グループを形成している。しかし、その特性は経年とともに減少している。松本ら(2006)²⁶⁾によるとすぎく衛星に掲載されている XIS は、撮像と X 線分光分析に同時対応した X 線 CCD カメラである。

次に、NGC (New General Catalogue) グループがある。2011 年は、NGC を中心とするグループを形成している。しかし、その特性は弱い。NGC は、星雲、星団や銀河など 7840 個の天体が載っている天体カタログである²⁷⁾

そして、AGN (Active Galactic Nuclei) グループがある。2012 年から 2013 年の 2 年間と 2016 年は、AGN を中心とするグループを形成している。しかし、2013 年以外はその特性は弱い。AGN (Active Galactic Nuclei) は、活動銀河核であり、太陽系程度の小さな領域から莫大なエネルギーを放出している天体を中心核に持つ銀河のことをいう。理化学研究所 (2014)²⁸⁾は、AGN はブラックホールにガスが激しく吸い込まれることで、強大なエネルギー量を放射していることが観測されている。

最後に SNR (Supernova Remnant) グループがある。2014 年から 2015 年の 2 年間、2017 年から 2018 年の 2 年間は、SNR を中心とするグループを形成している。しかし、2014 年から 2015 年の 2 年間はその特性が弱い。SNR は、超新星残骸であり、恒星が超新星爆発を起こし星間空間で飛び散った恒星物質が、巨大な高温プラズマ雲に形成される天体である²⁹⁾このように、すぎく関連論文の研究のトレンドは、13 年間の間に XIS、NGC、AGN、SNR にその特徴が移行した。

ここで図 3 の対応分析から NuSTAR という単語が現れた。NuSTAR とは、NASA から打ち上げられた X 線宇宙望遠鏡である。ブラックホールや超新星残骸などの高エネルギー X 線の観測が可能であり、すぎく衛星と同等の衛星であるといえる。NuSTAR は 2012 年に打ち上げられ、現在も運用中である³⁰⁾。同じ X 線天文衛星であることから、すぎく衛星と NuSTAR の比較研究など関連づけられた論文が書かれている可能性があると考えられる。

さらに、図 4 にすぎく衛星と NuSTAR の論文数の年次別変化とその推移を示した。すぎく衛星は運用が終了していることもあり、論文数は減少傾向にあるが、NuSTAR 論文は運用から 3 年ですぎく衛星の論文数に追いつき、その後も増加傾向にある。つまり SUZAKU 論文の減少とともに NuSTAR 論文が増加している。新たな性能を持った天文衛星が上げれば、その衛星の論文出版割合は増えて、それ以前の衛星の論文出版割合が減ることは自然な流れであると考えられる。

4. まとめ

人工衛星・探査機から得られる様々な観測データは、産官学すべてにおいて広く活用が期待されている。一方、宇宙開発プロジェクトは大規模な事業であるため、その貢献について、国会をはじめあらゆる場面で説明が求められている。また、日本の研究力は 1990 年代中盤から低下しているとされ、研究力の向上が急務となっていることがあり、学術的貢献に関して客観的な説明資料の必要性が高くなっている。そこで本稿は、学術的貢献の補助的説明となる客観的情報を用意するための手法を導出すること、そして、その効果を検証すること目標とした。分析では、論文の要旨にテキストマイニング分析を施して年代別の研究トレンドを抽出、そして、その手法の効果を検証した。

人工衛星・探査機の測定データの活用度を論文数とその特徴で表すことを目指し、X 線天文衛星「すぎく」プロジェクトの情報を用いて、テキストマイニング分析の適用の試行を行った。分析の結果、すぎく

衛星関連論文の研究のトレンドは、13年間の間にXIS(X線CCDカメラ)、NGC(天体カタログ関連の分析)、AGN(活動銀河核関連の分析)、SNR(超新星残骸関連の分析)にその特徴が移行していることを掴んだ。この結果から、すざく衛星は、研究のニーズに合わせた柔軟な運用が可能であったと考えられる。

次に、論文数に関する分析では、すざく衛星関連論文の論文数2009年をピークとして、その後は減少傾向にあることを掴んだ。一方、NuSTAR宇宙望遠鏡関連論文は、打ち上げから3年の2015年には、すざく衛星関連論文の論文数に追いつき、その後も増加傾向がつづいている。

すざく衛星にはないNuSTARの特長の一つに「硬X線による高精度撮像観測」がある。2016年に打ち上げられたひとみ衛星もまた「硬X線による高精度撮像観測」が可能であったが、運用期間が短く、すでに運用が断念されている。このため、高精度撮像観測の需要がNuSTARに集中しており、NuSTARの論文数が増加していると考えられる。このように、これら現在の構図は、ひとみ衛星の運用断念が影響していると考えられる。なお、ひとみ衛星は優れた分光観測性能を備えており、短期間の運用であったにも関わらず、ひとみ衛星プロジェクト関連の論文は、2015年から2019年の5年間に247編あり、その性能と注目度の高さをうかがい知ることができる。

今後の方向性として、論文要旨のテキストデータのみではなく、観測データの活用度確かめる方法を検討する。テキストマイニング分析では、分析時間が短い特徴がある。手法の精度向上とともに、特徴を生かした月報用簡易分析など、その活用方法を模索することがある。

謝辞

本研究はJSPS科研費JP17K04710の助成を受けたものです。本研究は統計数理研究所共同研究プログラム(30-共研-4212 研究者の異分野融合度と多様度の客観的な評価指標研究の深化)の助成を受け、統計数理研究所の資産を活用したものです。

参考文献

- [1] 海老沢研、「[宇宙科学研究所(ISAS)のデータポリシー]策定の経緯について」、第7回 国際的動向を踏まえたオープンサイエンスの推進に関する検討会、資料2、2018
- [2] 水上祐治、中村匡佑、家富紗妃、大島昭子、中野 純司、「テキストマイニング分析による研究動向把握の簡易的手法 -衛星プロジェクトの関連論文の要旨分析-」、平成30年度「宇宙科学情報解析シンポジウム」、2019
- [3] 文部科学省(2017)、科学研究のベンチマーキング 2017 <<http://www.nistep.go.jp/wp/wp-content/uploads/NISTEP-RM262-FullJ.pdf>> (最終確認: 2019年3月3日)
- [4] KHCoder Web ページ, <<http://khcoder.net/>> (最終確認: 2019年3月3日)
- [5] JAXA 第一宇宙技術部門 衛星利用推進 Web ページ <<http://www.sapc.jaxa.jp/use/rd/>> (最終確認: 2019年3月3日)
- [6] NASA's HEASARC-A Brief History of the High Energy Astrophysics Science Archive Research Center <<https://heasarc.gsfc.nasa.gov/docs/heasarc/history.html>> (最終確認: 2019年11月23日)
- [7] ISAS PLAIN センターニュース 第39号 1997年1月20日発行

- <<http://www.isas.jaxa.jp/docs/PLAINnews/PLAINno39/PLAINno39.html#mark2>> (最終確認: 2019 年 11 月 23 日)
- [8] 川上修司, 本田秀之, 小野縁, 長木明成, 斎藤宏, 平井寿美子, 周東三和子, 竹島敏明, 「宇宙科学資料データベースの開発」, 宇宙航空研究開発機構研究開発報告 第 6 号, JAXA-RR-16-007, pp. 89-98, 2017
- [9] JAXA 宇宙科学資料データベース DARTS Web ページ <<https://www.darts.isas.jaxa.jp/index.html.ja>> (最終確認: 2019 年 3 月 3 日)
- [10] 大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」Web ページ <<http://www.iugonet.org/index.jsp>> (最終確認: 2019 年 3 月 3 日)
- [11] 林寛生, 小山幸伸, 堀智昭, 田中良昌, 新堀淳樹, 鍵谷将人, 阿部修司, 河野貴久, 吉田大紀, 上野悟, 金田直樹, 米田瑞生, 田所裕康, 元場哲郎, 「大学間連携プロジェクト『超高層大気長期変動の全球地上ネットワーク観測・研究』」, 宇宙航空研究開発機構研究開発報告 第 1 号, JAXA-RR-11-007, pp. 113-120, 2012
- [12] 田中良昌, 新堀淳樹, 鍵谷将人, 堀智昭, 阿部修司, 小山幸伸, 林寛生, 吉田大紀, 河野貴久, 上野悟, 金田直樹, 米田瑞生, 田所裕康, 元場哲郎, 三好由純, 関華奈子, 宮下幸長, 瀬川朋紀, 小川泰信, 「IUGONET 解析ソフトウェアの開発」, 宇宙航空研究開発機構研究開発報告 第 1 号, JAXA-RR-11-007, pp. 91-98, 2012
- [13] 小山幸伸, 河野貴久, 堀智昭, 阿部修, 吉田大紀, 林寛生, 田中良昌, 新堀淳樹, 上野悟, 金田直樹, 米田瑞生, 元場哲郎, 鍵谷将人, 田所裕康, 「超高層物理学分野の為のメタデータ・データベースの開発」, 宇宙航空研究開発機構研究開発報告 第 1 号, JAXA-RR-11-007, pp. 99-104, 2012
- [14] 堀智昭, 鍵谷将人, 田中良昌, 林寛生, 上野悟, 吉田大紀, 阿部修司, 小山幸伸, 河野貴久, 金田直樹, 新堀淳樹, 田所裕康, 米田瑞生, 「IUGONET 共通メタデータフォーマットの策定とメタデータ登録管理システムの開発」, 宇宙航空研究開発機構研究開発報告 第 1 号, JAXA-RR-11-007, pp. 105-111, 2012
- [15] 小山幸伸, 阿部修司, 八木学, 梅村宜生, 堀智昭, 新堀淳樹, 佐藤由佳, 家森俊彦, 田中良昌, 橋口典子, 上野悟, 谷田貝亜紀代, 「超高層物理学分野のメタデータ・データベースへの連想検索の適用」, 宇宙航空研究開発機構研究開発報告 第 3 号, JAXA-RR-13-010, pp. 89-98, 2014
- [16] 北川宙, 谷内勇介, 国広卓也, 中村栄三, 「地球惑星試料デポジットリ DREAM の構想と現状」, 宇宙航空研究開発機構研究開発報告 第 3 号, JAXA-RR-13-010, pp. 71-78, 2014
- [17] 中川 友進, 海老沢 研, 中平 聡志, 上野 史郎, 富田 洋, 木村 公, 杉崎 睦, 芹野 素子, 三原 建弘, 根来 均, 「全天 X 線監視装置『MAXI』のデータアーカイブの開発」, 宇宙航空研究開発機構研究開発報告 第 5 号, JAXA-RR-15-006, pp. 71-80, 2016
- [18] 高木 亮治, 北條勝己, 「第二世代 EDISON (衛星運用工学データベースシステム) の開発」, 宇宙航空研究開発機構研究開発報告 第 5 号, JAXA-RR-15-006, pp. 23-28, 2016
- [19] 野口克行, 林寛生, 「火星大気データを地球大気解析ツールで解析できるようにするためのデータ整備 —MRO/MCS データの netCDF 化と GrADS による可視化—」, 宇宙航空研究開発機構研究開発報告 第 6 号, JAXA-RR-16-007, pp. 109-116, 2017
- [20] 根本しおみ, 高田良宏, 堀井洋, 堀井美里, 飯野孝浩, 林正治, 「DOI を用いた天文学研究資料の

横断的な整理・公開に向けた取り組み」, 宇宙航空研究開発機構研究開発報告 第7号, JAXA-RR-17-009, pp. 9-18, 2018

- [21] 根岸正光, 山崎茂明, 「研究評価—研究者・研究機関・大学におけるガイドライン」, 丸善, 2001
- [22] 藤垣裕子, 平川秀幸, 富澤宏之, 調麻佐志, 林隆之, 牧野淳一郎, 「研究評価・科学論のための科学計量学入門」, 丸善, 2004
- [23] Yuji Mizukami, Yosuke Mizutani, Keisuke Honda, Shigenori Suzuki, Junji Nakano, An International Research Comparative Study of the Degree of Cooperation between disciplines within mathematics and mathematical sciences: proposal and application of new indices for identifying the specialized field of researchers, Springer, Behaviormetrika, Vol.1, 19 pages, On-line, 2017
- [24] 樋口耕一, 「社会調査のための計量テキスト分析—内容分析の継承と発展を目指して—」, ナカニシヤ出版, 2014
- [25] 那須川哲哉, 河野浩之, 有村博紀, 「テキストマイニング基盤技術」, 人工知能学会, 人工知能学会誌 Vol. 16、No.2、pp.201-211、2001
- [26] 松本浩典, 片山晴善, 山口弘悦, 田和憲明, 「X線天文衛星「すぎく」搭載X線 CCD (XIS) のバックグラウンド」, 2006, <http://www-cr.scphys.kyoto-u.ac.jp/member/matumoto/presentation/2006/XIS_bgd_20060328_HM.pdf> (最終確認 2019年3月3日)
- [27] 天文学辞典 NGC カタログ web ページ <<http://astro-dic.jp/ngc-catalogue/>> (最終確認 2019年3月3日)
- [28] 理化学研究所, 「巨大ブラックホールが支配する「AGN エンジン」の解明へ」, 2014, <http://www.riken.jp/pr/press/2014/20140918_1/> (最終確認: 2019年3月3日)
- [29] JAXA 超新星残骸の X線精密分光観測 <<http://www.isas.jaxa.jp/j/forefront/2014/katsuda/>> (最終確認: 2019年3月3日)
- [30] X線衛星「NuSTAR」 Web ページ <<http://www.astroarts.co.jp/news/2012/06/14nustar/index-j.shtml>> (最終確認: 2019年3月3日)