

電力需給ひっ迫状況下におけるスーパーコンピュータシステム運用

山田 新^{1,a)} 宮寄 洋¹ 前田 光教¹ 佐藤 孝明¹ 山本 和男¹ 福沢 秋津¹ 中張 遼太郎¹
中島 研吾² 埴 敏博²

Operation of Supercomputer Systems under Power Supply and Demand Constraints

HAJIME YAMADA^{1,a)} HIROSHI MIYAZAKI¹ MITSUNORI MAEDA¹ TAKAAKI SATOH¹ KAZUO YAMAMOTO¹
AKITSU FUKUZAWA¹ RYOTARO NAKAHARI¹ KENGO NAKAJIMA² TOSHIHIRO HANAWA²

概要：本稿は、昨今の電力需給ひっ迫に対して、東京大学情報基盤センターで用意しているスーパーコンピュータシステムの消費電力削減策の内容と、2022年3月22日の「電力需給ひっ迫警報」及び、同年6月27日の「電力需給ひっ迫注意報」発令を受けて実際に消費電力削減を実施した際の成果について報告する。

1. はじめに

昨今、火力・原子力発電所設備の休廃止による電力供給力の低下と、異常気象による突発的な需要増加の発生により、電力需給のひっ迫が発生する状況が増加している。このような状況の中、政府・電力会社各社からは節電要請が広くなされている。

さて、東京大学情報基盤センター^{[1],[2]}(以下、当センター)でも節電要請に応えるべく節電への取り組みを実施しているところであるが、中でもスーパーコンピュータシステムは多くの電力を消費しており節電効果も大きくなる。本稿では、当センターが保有する「計算・データ・学習」融合スーパーコンピュータシステム(Wisteria/BDEC-01)^[3]、大規模超並列スーパーコンピュータシステム(Oakbridge-CX)^[4]の2システムでの消費電力削減の取り組みを紹介する。

2. 各システムの概要

最初に電力削減の対象となるシステムの概要を示す。

2.1 Wisteria/BDEC-01 スーパーコンピュータシステム

Wisteria/BDEC-01 は、2つのサブシステム Odyssey と Aquarius で構成(図 1)され、富士通株式会社により開発されたスーパーコンピュータである。総理論演算性能はそれぞれ 25.9 PFLOPS (Odyssey)、7.2 PFLOPS (Aquarius)、合計 33.1 PFLOPS である。Odyssey と Aquarius を利用するためのログインノード及び各ファイルシステムは、管理サーバ(ジョブ管理等)も含め、同じ環境が利用できるよう構成されている。

消費電力については、冷却設備を除いた Wisteria/BDEC-01 全体で約 2,400kW、Odyssey の計算ノードで約 2,000kW、Aquarius の計算ノードで約 250kW であり、計算ノードが消費電力の 9 割以上を占める。

2.1.1 シミュレーションノード群 (Odyssey)

Odyssey は、Armv8.2-A 命令セットアーキテクチャをスーパーコンピュータ向けに拡張した SVE(Scalable Vector Extension)を実装した CPU である A64FX を搭載した計算ノードで構成されたノード群である。各ノードは、48 個の演算コアと 2 個または 4 個のアシスタントコアを有しており、Odyssey 全体でのコア数は 368,640 コアとなる。ノード間ネットワークについては、バイセクションバンド幅が 13.0 TB/s の Tofu インターコネクト D で各ノードが結合されている。

2.1.2 データ・学習ノード群 (Aquarius)

Aquarius は CPU 2 基 (Intel Xeon Platinum 8360Y) と、

¹ 東京大学 情報システム部 情報基盤課
Information Systems Department, Information Technology Group,
The University of Tokyo.

² 東京大学 情報基盤センター
Information Technology Center, The University of Tokyo.

^{a)} yamada@cc.u-tokyo.ac.jp

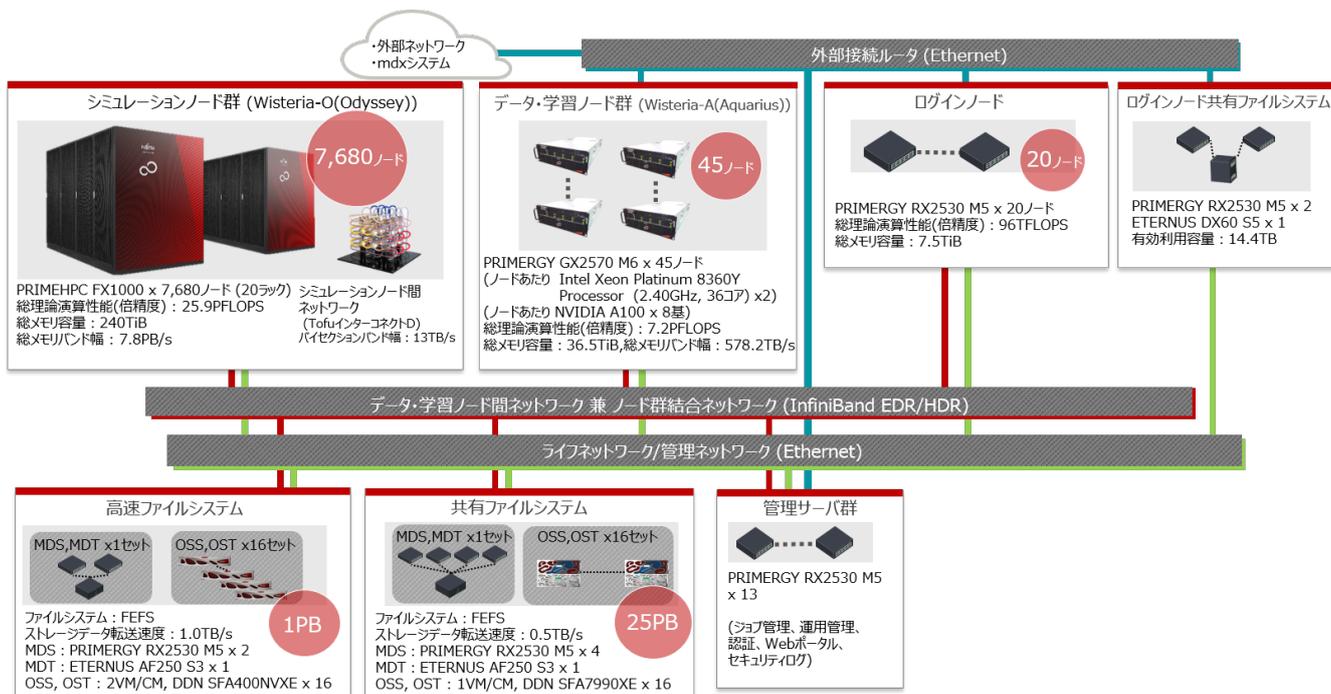


図1 Wisteria/BDEC-01の概要図

GPU (NVIDIA A100 Tensor コア) を 8 基搭載した計算ノードで構成されたノード群である。各ノードは、25 Gbps の Ethernet インタフェースを備えており、Aquarius は合計 800 Gbps のネットワーク転送速度で外部との通信が可能な構成となっている。当センターの既存システムである Oakbridge-CX と同様に、Aquarius の一部のノードは学術情報ネットワーク (SINET) 等の外部ネットワークを介して、様々な外部リソースに直接アクセス可能であり、観測データをリアルタイムに取り込んで解析、シミュレーションに利用することも可能である。

2.1.3 システムの特色

本システムは、「計算・データ・学習」融合スーパーコンピュータシステム (通称「BDEC(Big Data & Extreme Computing) システム」)^[5] 構築を目指した、「BDEC システム」の 1 号機として、設計、導入を実施し、運用を開始したシステムである。Odyssey と Aquarius を使用し、幅広いアプリケーションをカバーすることによって、最先端の科学技術計算を支える重要なインフラとなることだけではなく、東京大学各部局 (生産技術研究所、地震研究所、大気海洋研究所、物性研究所)、理化学研究所等との協力のもと、ものづくり、地球科学分野 (固体地球、大気・海洋)、物質科学などの分野における「シミュレーション (Simulation) + データ (Data) + 学習 (Learning) (S+D+L)」融合により、Society 5.0 を実現していく重要なプラットフォームとして活用がなされている。

2.2 Oakbridge-CX スーパーコンピュータシステム

Oakbridge-CX は Intel Xeon Platinum 8280(開発コード名:

CascadeLake) を搭載したスーパーコンピュータシステムである (図 2)。計算ノードは 1,368 台で、そのうち 128 台は SSD 搭載ノードである。計算ノード単体の理論演算性能は 4.8384TFLOPS、主記憶容量は DDR4 192GiB であり、全体では 6.61PFLOPS の性能、256.5TiB の主記憶容量を有している。消費電力は、冷却設備を除いた Oakbridge-CX 全体で約 950kW、計算ノードが約 850kW であり、消費電力の 9 割以上を計算ノードが占める。



図2 Oakbridge-CXの概要図

2.2.1 システムの特色

本システムの SSD 搭載ノードは各計算ノードの SSD を個別に使用できる他、ジョブ割り当ての都度 BeeGFS

onDemand (BeeOND) により 200TB 以上の単一の並列共有ファイルシステム (/dist) を構成して利用可能であり、高速な I/O が必要な処理に最適である。また、128 ある SSD 搭載ノードの内 16 ノードは外部接続ノードとして構成しており、SINET を経由して、外部計算資源（サーバ、ストレージ、センサーネットワーク等）に直接接続が可能であるため、大量のデータをリアルタイムに処理する場合に威力を発揮する。実際の活用例として、地震観測装置やデータアーカイブと連携した利用が行われている。外部接続ノードの運用には、システムの構成やセキュリティ等、通常の計算ノードとは異なった配慮が必要であり、運用にあたって得られた知見は、先述の Wisteria/BDEC-01 導入・運用にも還元がなされている。

3. 電力需給ひっ迫警報等

3.1 電力の予備率

電力需要は常に一定ではなく季節や日内での変動があるが、電気を大量に貯蔵する事は困難であるため、想定される電力需要よりも多くの供給能力を常に確保し、猛暑による需要の急増や発電設備のトラブルによる供給能力の急減などの状況に備えている。この余分に確保された供給能力の事を供給余力といい、電力需要に対する電力供給余力の割合を電力の予備率という。電力の需要は 3% 程度のぶれがあることから、安定供給には予備率 3% が最低限必要^[7]とされており予備率がこれを下回った場合には、需給がひっ迫して不安定となり最悪の場合は停電を招く恐れがある。

なお、予備率には大手電力会社の管轄エリア毎に計算されるものと、電力を融通することのできる複数のエリアをひとまとめにして計算した広域の 2 種類があるが、電力需給ひっ迫時に使用されるのは主に後者である。

3.2 電力需給ひっ迫警報等の発令

電力需給ひっ迫を解消するための方法の 1 つとして電力需要を削減する事が有効だが、スーパーコンピュータシステムのように急に停止できない機器や設備も多く存在する。そのため、需要家が前もって準備できるように電力需給ひっ迫警報等が資源エネルギー庁によって用意され事前に発令される事となっている。発令される条件と発令される際の流れをまとめると図 3 の通りである。なお【】内に示したものは発令対象となる電力予備率の値である。

つまり、電力需給ひっ迫によって節電要請が出されるのは前日の 16 時頃であり、実際に停電が確定するのは当日となってからである。

4. 電力需給ひっ迫度に応じた対策

4.1 電力削減プラン

各システムを通常の保守時と同様の手順で停止した場合、作業に半日以上の時間を要するため、需給ひっ迫準備情報



図 3 電力需給ひっ迫警報等の発令流れ

の段階でシステムを停止し始める必要があり、ユーザーへの影響が過大となってしまう。そこで、電力需給ひっ迫度に応じた電力削減プランをプラン A ~ C まで策定し弾力的な電力削減が可能となるように工夫を凝らした。

なお、先述のように冷却設備を除いたシステムの消費電力は、9 割以上を計算ノードが占めていることから、以下のプランでも計算ノードを停止する事に重点を置いている。

4.2 プラン A

プラン A は、最初に実施が検討されるプランである。

Odyssey では新規ジョブを実行抑止し、計算ノードを 0, 960, 1920, 2880, 3840 のいずれかの数で停止する。960 ノード単位となっているのは、ノード間ネットワーク (Tofu インターコネクト D) の y 軸方向^[6]に停止するノードを選定するためである。

Oakbridge-CX では新規ジョブを実行抑止し、計算ノードを 0~700 ノードの範囲で停止する。この停止はジョブ実行中でないノードをリソースグループから順に切り離し、可能なノードから停止する方法を取る。Aquarius については特段の作業を実施しない。通常運用からプラン A の適用完了 (最大ノード数の場合) までに要する時間は約 1 時間である。

4.3 プラン B

プラン B は、プラン A よりも電力を削減するためのプランである。また後述するプラン C の実施を見据えたプランであり、本プランの状態を維持することで、通常運用状態からプラン C に直接移行するよりもプラン C 自体の作業時

間を短縮する事が出来る。本プランでは、Wisteria/BDEC-01(Odyssey、Aquarius の双方)と Oakbridge-CX においてすべてのジョブを打ち切り計算ノードを停止する。なお、通常運用からプラン B の適用完了までに要する時間は約 1 時間である。

4.4 プラン C

プラン C は、Wisteria/BDEC-01、Oakbridge-CX の全システムを停止して、計画停電に対応するためのプランである。通常運用やプラン A から本プランに直接移行した場合、Wisteria/BDEC-01 と Oakbridge-CX の実行ジョブはすべて打ち切りとなる。所要時間は、通常運用からの場合で約 3.5 時間を要する。この際の内訳は、Oakbridge-CX が約 2.5 時間、Wisteria/BDEC-01 が約 3.5 時間であり、並行して作業を実施するため Wisteria/BDEC-01 の作業が完了した段階で完了となる。また、プラン A(半数の計算ノード停止完了状態)からの場合は約 3.5 時間、プラン B(計算ノードすべて停止完了状態)からの場合で約 2.5 時間である。

4.5 停止時のプラン適用

電力削減による停止時については、次の順序でプランを適用する事とした。

- (1) 「前日 16 時目処」の需給ひっ迫注意報または需給ひっ迫警報発令に伴い「新規ジョブ実行抑止」を開始する。
- (2) ジョブ状況調査結果からプラン A の規模を検討し、プラン A を開始する。
- (3) 「当日」9:00 において、需給ひっ迫警報発令状態の場合、プラン B を開始する。
- (4) 「当日」において、計画停電の発表が行われた際に、プラン C を開始する。

復旧については、基本的に需給ひっ迫の解消から最短での作業完了を目指す。作業が深夜に及ぶ事を避けるために解消がおおむね 15 時より後となった場合にはジョブ実行抑止の解除のみを実施し、ノード起動は翌日とする。なお、停止時と復旧時の共通事項として、作業実施が確定した段階で速やかに Web ページとメールによるユーザアナウンスを実施する。

5. 2022 年 3 月 22 日の電力削減対応

5.1 対応の背景

2022 年 3 月 22 日は、悪天候（実際の天候^[8]は東京都心で昼間（06:00-18:00）が雨一時みぞれ、夜間が曇時々晴（18:00-翌日 06:00）、最高気温 9.9℃、最低気温 1.2℃）が予想されており、暖房による電力需要増加が見込まれていた。一方で、直前の 3 月 16 日に生じた福島県沖の地震の影響により、東北、東京エリアの火力発電所 6 基（計約 330 万 kW）が停止しており、電力供給力が低下している状況であったため「需給ひっ迫警報」が前日 3 月 21 日に発令

^[9]された。これを受けて、当センターでは、自主的に電力削減のためにプラン A を実施した。その際のタイムスケジュールは次のとおりである。

表 1 2022 年 3 月 22 日のタイムスケジュール

日時	システム	作業
03/22 13:40	Odyssey,Oakbridge-CX	新規ジョブ実行抑止
03/22 14:45	Oakbridge-CX	計算ノード部分停止
03/22 15:20	Odyssey	計算ノード部分停止
03/22 22:14	Odyssey,Oakbridge-CX	実行抑止を解除
03/23 13:00	Oakbridge-CX	計算ノード起動 (通常復帰)
03/23 14:15	Odyssey	計算ノード起動 (通常復帰)

5.2 Oakbridge-CX での電力削減効果

Oakbridge-CX では、ノード単位で停止を実施し、停止した計算ノード数は 600 ノードとなった。計算ノード停止日の前後 2 日を含めた電力の推移は図 4 に示す通りであるから、前日 (3/21) の電力値を「1 日あたりの値 (599 kW, 計算ノード+水冷)」と仮定した場合、3/22 から 3/23 にかけて 16.6% (199kW) の節電効果があったと考えられる。

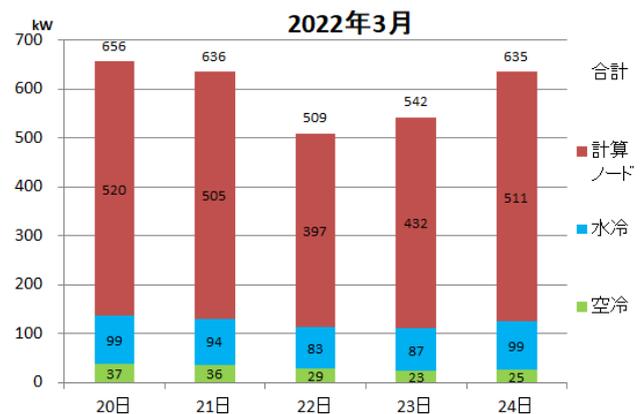


図 4 2022 年 3 月電力削減実施前後の電力推移 (Oakbridge-CX)

5.3 Wisteria/BDEC-01 Odyssey での電力削減効果

Odyssey では、16 ノード単位 (BoB) で計算ノードを停止し、停止した計算ノード数は 3,840 ノードとなった。Odyssey の 3 月の電力推移は図 5 に示す通りであるから、前日 (3/21) の電力値を「1 日あたりの値 (1066 kW, 計算ノード+水冷)」と仮定した場合、3/22 から 3/23 にかけて 23.2% (496kW) の節電効果があったと考えられる。

6. 2022 年 6 月 27 日の電力削減対応

6.1 対応の背景

2022 年 6 月 27 日は、季節外れの暑さ（実際の天候^[10]は東京都心で昼間（06:00-18:00）、夜間（18:00-翌日 06:00）ともに晴時々薄曇、最高気温 35.7℃、最低気温 24.1℃）

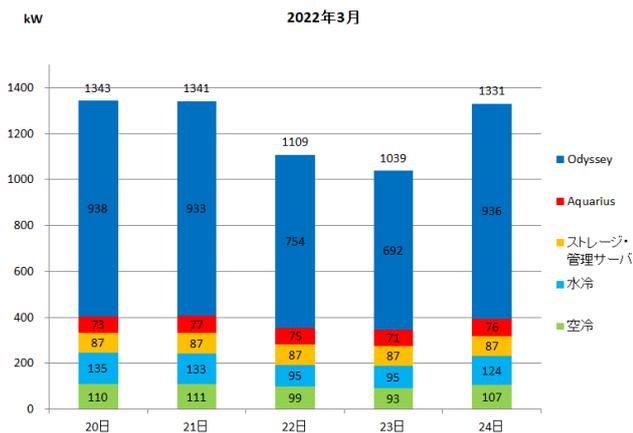


図5 2022年3月電力削減実施前後の電力推移 (Wisteria/BDEC-01)

が予想されており、冷房による電力需要増加が見込まれていたため「電力需給ひっ迫注意報」が前日6月26日に発令された。これを受けて、当センターでは、自主的に電力削減のためにプランAを実施した。その際のタイムスケジュールは次のとおりである

表2 2022年6月27日のタイムスケジュール

日時	システム	作業
06/27 10:00	Odyssey, Oakbridge-CX	新規ジョブ実行抑止
06/27 13:20	Oakbridge-CX	計算ノード部分停止
06/27 17:05	Odyssey	実行抑止を解除 (通常復帰)
06/27 17:10	Oakbridge-CX	実行抑止を解除
06/27 19:00	Oakbridge-CX	計算ノード起動 (通常復帰)

6.2 Oakbridge-CX での電力削減効果

Oakbridge-CX では、ノード単位で停止を実施し、停止した計算ノード数は700ノードとなった。計算ノード停止日の前後2日を含めた電力の推移は図4に示す通りであるから、前日(6/26)の電力値を「1日あたりの値(594kW, 計算ノード+水冷)」と仮定した場合、6/26は9.8%(57kW)の節電効果があったと考えられる。

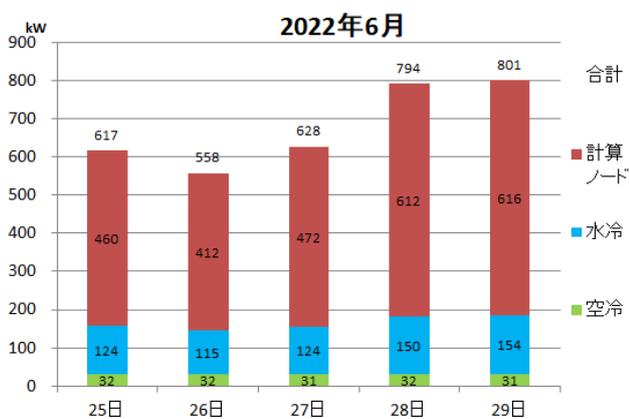


図6 2022年6月電力削減実施前後の電力推移 (Oakbridge-CX)

6.3 Wisteria/BDEC-01 Odyssey での電力削減効果

Odyssey では、計算ノードの停止は実施せずに、ジョブ抑止のみを実施した。計算ノード停止日の前後2日を含めた電力の推移は図4に示す通りだが、期間を通して消費電力はほぼ変化せず、節電の効果を確認する事はできなかった。

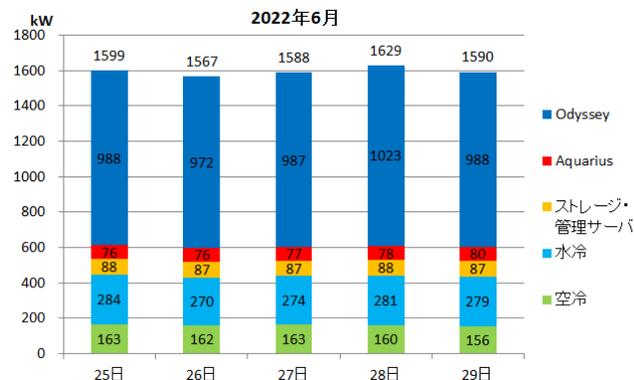


図7 2022年6月電力削減実施前後の電力推移 (Wisteria/BDEC-01)

7. まとめ

実際の対応を通して、策定した電力削減プランによって、節電効果を発揮しつつ利用者への影響も減少させるという目標を達成可能である事が確認できた。しかしながら、ジョブ停止のみで得られる消費電力削減はさほど効果が大きくないことが判明したのでプラン実行時のノード停止数については見直しが必要である。他の課題としては、ジョブが実行できないなどの問い合わせがユーザからあったことや、ノード停止状態からの復元時にノードやネットワーク障害が発生したことなどが挙げられる。ユーザへの告知方法や復旧時の手順見直しなど、よりよい電力削減運用ができるように改善を続けていきたい。

謝辞 本論文を作成するにあたって、様々なご教示をいただいた情報基盤課、情報基盤センターの関係者の皆様に深く感謝致します。また、システム運用に携わっていただいている富士通株式会社の方々、加えて、各システムにおいて電力削減にご協力頂いた多くの利用者様に深く感謝致します。この場を借りてお礼を申し上げます。

参考文献

- [1] 東京大学情報基盤センター: <https://www.itc.u-tokyo.ac.jp/>.
- [2] 東京大学情報基盤センタースーパーコンピューティング部門: <https://www.cc.u-tokyo.ac.jp/>.
- [3] Wisteria/BDEC-01 スーパーコンピュータシステム: <https://www.cc.u-tokyo.ac.jp/supercomputer/wisteria/service/>.
- [4] Oakbridge-CX スーパーコンピュータシステム: <https://www.cc.u-tokyo.ac.jp/supercomputer/obcx/service/>.
- [5] Iwashita, T, Nakajima, K., Shimokawabe, T., Nagao,

H., ita, *h3-Open-BDEC: Innovative Software Platform for Scientific Computing in the Exascale Era by Integrations of (Simulation + Data + Learning)*, Project Poster, ISC-HPC 2020

- [6] 富士通株式会社. "ジョブ運用ソフトウェアエンドユーザ向けガイド". FUJITSU 日本ポータル. 2022. <https://software.fujitsu.com/jp/manual/manualfiles/m220007/j2ul2452/02z007/j2ul-2452-02z0.pdf>, (参照 2022-10-12).
- [7] 資源エネルギー庁. "電力需給状況". 資源エネルギー庁ホームページ. 2022. https://www.enecho.meti.go.jp/category/electricity_and_gas/electricity_measures/dr/jokyo.html, (参照 2022-10-08).
- [8] 気象庁. "東京 2022年3月(日ごとの値)主な要素". 気象庁ホームページ. 2022. https://www.data.jma.go.jp/obd/stats/etrn/view/daily_s1.php?prec_no=44&block_no=47662&year=2022&month=3&day=&view=p1, (参照 2022-10-17).
- [9] 経済産業省. "3月22日は電力需給が厳しくなる見込みのため東京電力管内で節電のご協力をお願いします【需給ひっ迫警報】". 経済産業省ホームページ. 2022. <https://www.meti.go.jp/press/2021/03/20220321001/20220321001.html>, (参照 2022-10-11).
- [10] 気象庁. "東京 2022年6月(日ごとの値)主な要素". 気象庁ホームページ. 2022. https://www.data.jma.go.jp/obd/stats/etrn/view/daily_s1.php?prec_no=44&block_no=47662&year=2022&month=6&day=&view=p1, (参照 2022-10-17).
- [11] 経済産業省. "6月27日は東京エリアで電力需給が厳しくなる見込みのため節電のご協力をお願いします【需給ひっ迫注意報】". 経済産業省ホームページ. 2022. <https://www.meti.go.jp/press/2022/06/20220626001/20220626001.html>, (参照 2022-10-11).