

## 気象データの分析に関する研究 Study on Analysis of Weather Data

生田目雅也<sup>†</sup>  
Masaya Namatame

小飼 敬<sup>†</sup>  
Kei Kogai

蓬莱尚幸<sup>‡</sup>  
Hisayuki Horai

### 1. はじめに

長い間、世界中で地球温暖化が問題となっている。地球温暖化の影響で極地方の氷が融け、巨大化した台風やハリケーン、熱波や洪水などが起こっている。このような、地球温暖化が進行していると感じられる現象による悪影響は、世界中の様々な世界中の様々な所で起きている。それらの悪影響は、様々な分野にわたって世界中の人々の暮らしに密接にかかわっている。

将来的な地球温暖化が地球環境に今より悪影響を与える恐れがあると考えられている。そのため、上記の現象と地球温暖化との関係性および、その対応策は多くの研究が行われている。しかし、地球温暖化が与える文化面、経済面、学力面などの社会現象への影響については、まだ研究の余地があると思われる[1]。

そこで、本研究は気象データを分析し、文化面などの今まで気象データとの関係性が明らかにされてこなかった社会現象と気象データの関係性を発見し、その関係性がもたらす社会的な利益に寄与することを目的とする。

### 2. 比較研究

本研究では、気象データとの今まで関係が明らかにされていなかった関係性について考えるため、多くの人が直感的に比較的わかりやすい相関係数や Accuracy, precision, recall, F-value を使用することとした。グラフにできるような値のあるデータは相関係数、値のないデータは Accuracy, precision, recall, F-value を使用し比較する。

比較した結果、関係性が明らかになっていない社会現象と気象データの相関係数や F-value が大きいなら関係性があり、この事実が社会の利益に寄与することになるだろう。また、比較結果が低い場合でもなぜ低いのかを考察することで社会貢献になるだろう。

例えば、マグロの月の価格などがある。マグロの旬は 12-2 月であるので、多くの都市では 12-2 月の価格が高い。しかし、12-2 月のマグロの旬の季節よりも 8,9 月のような夏の方が価格が高い都市がある年がある。このような場合、季節という要素ではなく気象と相関があるかもしれないと考えられる。

<sup>†</sup> 茨城工業高等専門学校

National Institute of Technology, Ibaraki College

<sup>‡</sup> 国立高等専門学校機構本部

National Institute of Technology, Headquarters

### 3. 気象データ

本研究は、広い種類と長い年月のデータを扱いやすいという理由から日本の気象データで行うことにした。気象庁のホームページにて取得した、AMeDAS、官署(気象台等)のデータを使用する。現在、気象に関するデータは気象庁等によって組織的に収集しデータベース化されている。その中の主なデータ収集方法を紹介する。

\$北海道						
月	気温	降雨量	日照率	風速	曇量	気圧
1	-4.9	138.0	37	1.8	7.0	1016.0
2	-0.7	75.0	33	2.0	8.0	1023.3
3	2.9	77.5	57	2.4	6.0	1014.5
4	7.9	86.5	50	2.9	6.5	1011.4
5	13.4	25.0	49	2.4	6.8	1011.2
6	17.7	37.5	36	2.4	7.7	1008.8
7	20.8	48.5	45	2.5	6.9	1009.0
8	23.1	137.5	39	2.2	7.5	1008.1
9	18.7	123.5	43	1.8	7.4	1015.0
10	13.0	87.0	53	1.8	5.1	1017.3
11	7.5	73.0	31	2.1	7.1	1016.8

図 1 北海道の月平均データ

#### 3.1. 気象データ収集方法

##### AMeDAS

アメダス(AMeDAS)とは「Automated Meteorological Data Acquisition System」の略で、「地域気象観測システム」という。雨、風、雪などの気象状況を時間的、地域的に細かく監視するために、降水量、風向・風速、気温、日照時間の観測を自動的におこない、気象災害の防止・軽減に重要な役割を果たしている。

##### 気象台

全国約 60 か所の気象台・測候所では、気圧、気温、湿度、風向、風速、降水量、積雪の深さ、降雪の深さ、日照時間、日射量、雲、視程、大気現象等の気象観測を行っている。雲、視程、大気現象等は観測者が目視によって観測しているが(一部の気象台を除く)、その他は地上気象観測装置によって自動的に観測を行っている。また、全国約 90 か所の特別地域気象観測所では、地上気象観測装置による自動観測のみを行っている。

### 3.2. 利用する気象データ

気象庁のホームページから取得したデータは、1872 年から 2019 年までの 147 年間あり、各年に気温、降水量、日射/日照、積雪/降雪、風速、湿度/気圧、雲量/天気 の 7 種類用意されている。さらに、7 種類のデータは月平均や月の最高最低などのデータに変換して都道府県ごとに、csv ファイルでダウンロードすることができる[2]。

気象庁では、主に気象データを扱うとき、複数年にわたる期間で行う統計は 30 年間のデータを使用している[2]。それに合わせて本研究でも 30 年分の気象データを使用することにした。

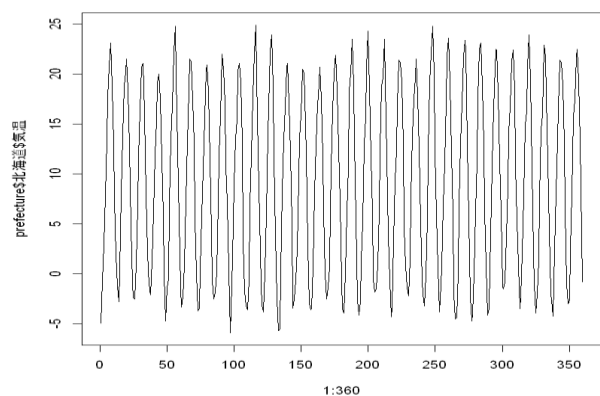


図 2 北海道の月平均気温グラフ

また、地理的に広い範囲のデータを対象とすることで様々な視点から変化を分析できると考えたため、各県庁所在地の気圧、降水量の合計、日照率、風速、現地気圧、雲量のデータを使用した。気象庁のデータは 7 種類あるにもかかわらず、この 6 種類を扱うことにした理由は、積雪/降雪のデータの抜けが多かったため分析対象としては適していないと判断したからである。風速もデータの抜けがあるが積雪/降雪ほど多くなかったため使用することにしたが、これから分析していく中で使用するかしなないかを決定していく。

取得したデータの例を図 1 に表形式で示す。図 1 は、気象庁のデータダウンロードオプションの一つである月平均を使用した北海道の気温である。比較対象データによってダウンロードオプションを変更できる。

図 2 は北海道の月平均気温のグラフである。このグラフは全都道府県、すべてのデータで書くことができる。

### 3. 比較対象データ

多くの日本の公開データは年平均や男女、年齢ごとに分かれているものが多く、時系列データが少ないため気象データとの比較には向いていない。そのため、比較対象には yahoo finance、e-stat のデータを使用することにした。しかし、現在データをとって分析ができていないため使用するデータは変更される可能性がある。

Yahoo finance のデータは yahoo が提供している api でダウンロードすることができる。ダウンロードできるデータの形式は json であり図 3 のような形でダウンロードできる。

Yahoo finance のデータは市場、株式、ニュース、会話の 4 種類ある。株式では各株の詳細情報、市場では収益情報などを取得することができる。

```
print(response.text)
{"marketSummaryResponse":{"result":[{"fullExchangeName":"CME","symbol":"ES=F","headSymbolAsString":"ES=F","gmtOffsetMilliseconds":-14400000,"language":"en","regularMarketTime":{"raw":1591788614,"fmt":"7:30AM EDT"},"regularMarketChangePercent":{"raw":-0.023397285,"fmt":"-0.02%"},"quoteType":"FUTURE","headSymbol":true,"contractSymbol":false,"tradeable":false,"regularMarketPreviousClose":{"raw":3205.5,"fmt":"3,205.50"},"exchangeTimezoneName":"America/New_York","regularMarketChange":{"raw":-0.75,"fmt":"-0.75"},"firstTradeDateMilliseconds":953528400000,"exchangeDataDelayedBy":10,"exchangeTimezoneShortName":"EDT","marketState":"REGULAR","regularMarketPrice":{"raw":3204.75,"fmt":"3,204.75"},"market":us24_market,"exchange":"CME","sourceInterval":10,"region":"US","shortName":"S&P Future","triggerable":false},{"fullExchangeName":"CBOT","symbol":"YM=F","headSymbolAsString":"YM=F","gmtOffsetMilliseconds":-14400000,"language":"en","regularMarketTime":{"raw":1591788612,"fmt":"7:30AM EDT"},"regularMarketChangePercent":{"raw":-0.15404929,"fmt":"-0.15%"},"quoteType":"FUTURE","headSymbol":true,"contractSymbol":false,"tradeable":false,"regularMarketPreviousClose":{"raw":27264.0,"fmt":"27,264.00"},"exchangeTimezoneName":"America/New_York","regularMarketChange":{"raw":-42.0,"fmt":"-42.00"},"firstTradeDateMilliseconds":1017982800000,"exchangeDat
```

図 3 json 形式の finance data

e-stat のデータは国土・気象、人工・世帯、労働・賃金、農林水産業、鉱工業、商業・サービス業、企業・家計・経済、住宅・土地・建設、エネルギー・水、運輸・観光情報通信・科学技術、教育・文化・スポーツ・生活、行財政、司法・安全・環境、社会保障・衛生、国際、その他の 17 種類ある。周囲が豊富で月別にデータが取得できるため気象と比較する対象データと最適なものがそろっていると考えられる。

### 4. 今後の課題

図 2 は図 1 をグラフにしたものである。このグラフではデータ数が多すぎるため、夏の気温が高い、冬の気温が低い年あるということくらいしかわからない。そのため、まずはこのグラフで相関をとり、基準を設定して期間を短くする必要があるだろう。また、図 2 のグラフはこのままでは見にくい。見やすくするために各対の値に点をつける、1 年ごとにグラフの色を変える、横軸を見て何月の点なのかわかるようにするなどして見やすくしていきたい。

図 3 のままでは区切りが見にくくデータが何を示しているかよくわからない。そのため、エンコーダ、デコーダし json 形式のデータをオブジェクト化する必要がある。

前述したデータより文化的な社会現象データを取得したい。そのために、twitter などを使用することを考えた。比較的毎日つぶやかれている単語を検索し、ツイート数の統計を取りことで文化的な社会現象データを取得できると考えている。

### 参考文献

- [1] 清水 浩、青木 陽二、森口 祐一 “地球温暖化対策技術とその評価”, 環境科学学会誌, Vol.3(1) 49-56(1990).
- [2] 気象庁  
<http://www.jma.go.jp/jma/index.html>
- [3] Yahoo finance:  
<https://finance.yahoo.co.jp/>
- [4] e-stat  
<https://www.e-stat.go.jp/>