

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 23 日現在

機関番号：53302

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23501084

研究課題名(和文) 気温データを用いた小中学生向け体験型データマイニング教育教材の開発と評価

研究課題名(英文) Developing and Evaluating a Hands-on Learning Tool Employing Air Temperature Data for Use by Primary and Junior High School Students in Introductory Data Mining Classes

研究代表者

今澤 明男 (Imazawa, Akio)

金沢工業高等専門学校・グローバル情報工学科・教授

研究者番号：20148141

交付決定額(研究期間全体)：(直接経費) 4,000,000円、(間接経費) 1,200,000円

研究成果の概要(和文)：本研究では、小中学生向けデータマイニング導入実技教材を開発し評価した。具体的にはまず、気温データのグラフを作成する専用のコンピュータソフトウェアを開発した。このソフトウェアには、47都道府県庁所在地を含む日本の60地点の気温データが収められている。ユーザが簡単な操作でデータを選択すると、ソフトウェアは折れ線グラフや散布図などを自動的に作成する。さらに、作成されたグラフ類を用いると、統計学やデータ解析の知識や技術を持たないユーザでも相関や傾向や循環といった統計的パターンを直感的に把握できることが確認された。

研究成果の概要(英文)：We have developed and evaluated a hands-on learning tool for use by primary and junior high school students in introductory data mining classes. This specialized computer software can make statistical graphs and diagrams based on air temperature information, and it contains air temperature data from 60 locations in Japan including all the prefectural capitals. It can automatically generate visualizations, such as line graphs and scatter diagrams, from data that the user has selected through a simple process. We have also confirmed that the user can use the generated graphs and diagrams to intuitively see statistical patterns such as correlations, cycles and trends even if they have no knowledge of statistics and data analysis.

研究分野：総合領域

科研費の分科・細目：科学教育

キーワード：データマイニング 気温 科学リテラシー 統計教育 科学教育 環境教育 体験型学習 小中学生

1. 研究開始当初の背景

(1) 問題の所在：グローバル化や技術革新、あるいは高齢化など、日本を取り巻く環境は急激に変化している。そのため、行政・産業・教育など様々な分野で、社会の変化を迅速にかつ的確に把握することが必要となってきた。一方、情報技術の飛躍的進歩により、大量のデータから意味のあるパターンを導出する「データマイニング技術」が可能となってきた。これを用いると、例えば企業の場合、消費者の新しいニーズや市場の未知の動向などを発見し、把握する能力が大きく向上すると期待されている。今後、技術と情報で変化を生き抜こうとする日本において、広く活用されるべき技術であり、そのためには専門家の間だけでなく、広く人々の間に普及させるべき技術である。しかしながら、初学者にとって扱いやすいデータをはじめとして、普及のための適切な教材が少ないのが現状である。

(2) 問題解決の糸口：本研究の研究者らは、平成 17 年から環境教育の研究を行い、その一環として小学生や中学生を対象に環境教育を実践してきた。その際、気温データの統計的分析が、小中学生に素直に、かつ興味を持って受け入れられることを実感した。さらに幸いなことに、日本では明治以来の各地の詳細な気温データが気象庁によりまとめられ、広く一般に公開されている。そこで、気温データを用いて小中学生に簡単なデータマイニング体験をさせることに思い至った。データマイニングの面白さを実感出来れば、将来データマイニングを本格的に学ぼうとしたり、社会に出てから積極的に活用しようとする期待できるからである。ただし、既存のデータマイニングソフトは高価で操作も難しく、一方、表計算ソフトを用いると操作が煩雑となり、教育の効率が低下する。小中学生でも理解できる気温データについてのマイニング機能がまとめられていて、直感的かつ簡単に操作できる専用ソフトが必要だと考えた。

(3) 気温データの教育素材としての特徴：日本の気温データは、データマイニング入門教育の素材として、以下の特徴を持つ。

①物理的特徴：気温は地点により、また季節や時刻あるいは天候により異なる。しかし一方で、緯度や標高などといった基本的決定要因との間に相関関係を持ち、さらに1日や1年といったサイクル(循環)や、温暖化などといったトレンド(傾向)も有している。また、間隔尺度であり平均値が意味をもつ。すなわち、気温データは解析対象として豊富な特徴を備えている。

②心理的特徴：気温を温度、すなわち数値で表すことに抵抗を感じる人はいない。また、

測定された数値は「暑い」「寒い」などと直感的に理解できる。さらに、緯度や標高あるいは季節や時刻などの基本的決定要因についても、日常の中で広く認識されている。すなわち、現象の数値化に抵抗がなく、測定値や解析結果を容易に理解し考察することができる。

2. 研究の目的

小中学生がデータマイニングを体験できる教材を開発し、評価する。具体的には、日本各地の気温データについての各種のグラフを、簡単な操作で短時間に作成することができる専用のコンピュータソフトウェアを開発する。さらに、そのソフトウェアを用いて作成されたグラフ類を用いると、統計学やデータ解析の知識や技術を持たないユーザでも、相関や傾向や循環といった統計的パターンを直感的に把握できるか否かを確認する。

3. 研究の方法

(1) 「気温データ・グラフ表示システム」の開発：日本各地の気温データについての各種のグラフを、簡単な操作で短時間に作成することができる専用のコンピュータソフトウェアを開発する。そのために、まずは多地点の長期の気温データを内蔵し、多種のグラフを作成することができるソフトウェアを作成する。それから、地点、データ、グラフの種類を絞り込むとともに必要な操作を少なくして単純化する。最終的には、必要最小限の操作で、ユーザが関心を持つグラフを作成できる使いやすいソフトウェアを作成する。なお、以下ではこのソフトウェアを「気温データ・グラフ表示システム」と呼ぶ。

(2) グラフを用いた統計的パターンの直感的把握の可否の確認：気温データ・グラフ表示システムで作成されたグラフから、相関(直線相関・曲線相関)、傾向、循環の各種の統計パターンが、視察によって直感的に把握できるか否かを、実際のグラフを用いて確認する。

4. 研究成果

(1) 気温データ・グラフ表示システムの概要：完成した気温データ・グラフ表示システムの概要は以下のとおりである。

① 使用環境：完成したソフトウェアは Windows XP、Windows Vista、Windows 7、Windows 8 の、どのオペレーティング・システムの上でも利用可能である。画面解像度は 1024×768 を想定している。この環境で Web サーバ上にプログラムを置き、ブラウザとして Chrome browser を用いて表示を行う。現在は同一のパーソナル・コンピュータ上でサーバとブラウザを組み合わせて使用しているが、単体のサーバと複数の端末とをネットワークで接続して利用することもできる。

② 対象とする観測地点：本ソフトウェアでは、気象庁のホームページ（<http://www.data.jma.go.jp/obd/stats/trn/>）に掲載されている気温データから、同庁の許可を得て一部を使用している。

対象観測地点は、47都道府県の都道府県庁所在地（ただし、埼玉県の県庁の所在地であるさいたま市、ならびに滋賀県の県庁の所在地である大津市には適当な観測地点がないので、それぞれ熊谷と彦根を代理として用いている）および気象庁が日本の平均気温算出に用いている17地点である。このうち4地点が両者に共通に含まれているので、合計60地点となる。都道府県庁所在地は一般に都市化が進み、気温の自然な状態を見るには不向きといえるが、地名から位置を容易に連想できる場所であり、都市化も現在の重要な気温決定要因と見なせるので、あえて用いている。

③ 対象とする気温データ：本ソフトウェアでは、気温データとしては以下の4種類を用いている。

- i) 時間ごとの気温（2010年2月1日1時～28日24時、5月1日1時～31日24時、8月1日1時～31日24時、11月1日1時～30日24時）
- ii) 日ごとの平均気温（2010年1月1日～12月31日）
- iii) 月ごとの平均気温（2001年1月～2010年12月）
- iv) 年ごとの平均気温（各地点観測開始翌年～2010年）

さらに、観測地点についてのデータとして、つぎを用いている。

- v) 緯度、vi) 経度、vii) 標高

(2) 気温データ・グラフ表示システムの操作手順：完成した気温データ・グラフ表示システムの操作手順は以下のとおりである。

① データの選択：本ソフトウェアを起動すると、「トップ画面」が現れる（図1参照）。この画面において、次に示す4つのデータ群から一つを選択する。以下では、月ごとの平均気温を選んだ場合について述べる。

- i) 時間ごとの気温
- ii) 日ごとの平均気温
- iii) 月ごとの平均気温
- iv) 年ごとの平均気温

② 表示方法の選択：データ群のひとつを選択すると「表示選択画面」が現れるので、データを表の形で数値で表示するかグラフで表示するかを選択する（図2参照）。

③ 対象地点の選択：「表示選択画面」でグラフ表示を選択したら、つぎに現れた画面で、対象とする地点をひとつ以上選ぶ。この際、一つひとつ選ぶこともできるが、メニューから以下のどれかをまとめて選ぶこともできる（図3、図4参照）。

- i) すべての地点
- ii) 都道府県庁所在地
- iii) 気象庁が日本の平均気温を算出する際に用いている17地点
- iv) 指定した範囲の緯度にある地点
- v) 指定した範囲の経度にある地点

ただし、緯度や経度は、利用者がプルダウンメニューから選択する。

④ 対象期間あるいは時点とグラフの選択：対象地点を選択したら、対象期間または時点ならびにグラフの種類を選択する。選択は以下の3つから行う。ただし、××で表した部分は、利用者がメニューあるいは表から選択する。例えば月ごとの平均気温をグラフ作成の対象とする場合は、以下の3つから選択する（図5参照）。

- i) ××年から××年を折れ線グラフで見る
- ii) ××年から××年を同じ月で平均値をとり、折れ線グラフで見る
- iii) 緯度と××年××月の平均気温とを比較した散布図を見る

⑤ グラフの表示：対象期間または時点ならびにグラフの種類を選択すると、新たなウィンドウが現れ、そこに当該のグラフが表示される。ここで、対象期間または時点を変えたり、さらには対象地点を変えたりして、グラフを幾つも連続して表示することができる。

⑥ 付加情報の表示：画面に表示されたグラフ上で、プロットされた点の上にポインタを置くと、付加情報が傍に表示される。折れ線グラフの場合は、i) 地点名、ii) 時点、iii) 気温が表示され、散布図の場合は、iv) 地点名、v) 緯度、vi) 気温が表示される。

⑦ グラフのサイズ調整と保存：表示されたグラフは、マウスで自由に縦横の長さを変えられることができる。また、クリック一つでPNGファイルとして保存することができる。

⑧ グラフの再作成と画面間の移動：対象期間あるいは時点、対象地点を変えて、グラフを幾つも連続して表示することができる。また、操作画面のどこからでもトップ画面に移動可能である。これらの機能により、試行錯誤的な解析を円滑に行えるようになっている。



図1. トップ画面

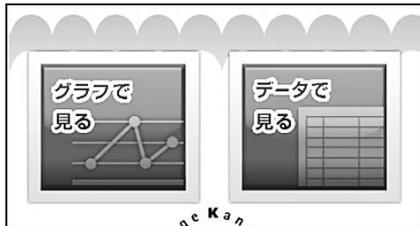


図2. 表示形式選択画面(部分)

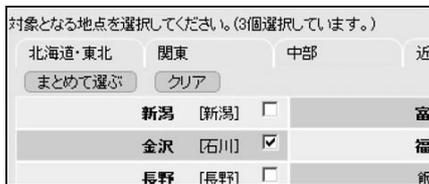


図3. 地点選択画面(部分1)

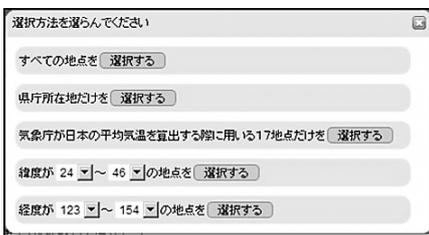


図4. 地点選択画面(部分2)

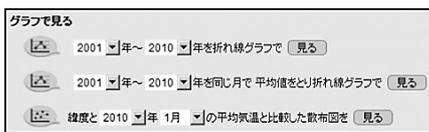


図5. 解析の種類・期間・時点の選択画面(部分)

(3) 抽出される基本的統計パターン：以下において、「気温データ・グラフ表示システム」に格納された60観測地点の4種の気温データからグラフを通じて視覚的・直感的に把握できる4つの基本的データパターン（直線相関・曲線相関・傾向・循環）を示す。

① 直線相関：データが格納されているすべての年において、観測地点の緯度と年平均気温との間に、グラフから視覚的に分かる直線相関が認められる。また、冬季において、各観測地点の緯度と月平均気温との間に、同様の直線相関が認められる。図6に例を示す。

② 曲線相関：夏季において、各観測地点の

緯度と月平均気温との間に、曲線相関が認められる。すなわち、高緯度の地域では緯度が上がると月平均気温は下がるが、低緯度のところでは緯度が上がっても月平均気温の大きな下降は見られない。図7に例を示す。

③ 傾向（トレンド）：すべての観測地点において、観測開始時点から2010年までの間に、明らかに気温の上昇傾向が見られる。特に近年、多くの地点で上昇傾向が強くなっていることもグラフから読み取れる。図8に例を示す。

④ 循環（サイクル）：各観測地点において、毎時の気温の年平均値は、24時間周期の滑らかな線を描く。図9に例を示す。また、多くの観測地点における毎時の気温の月平均値も、同様に24時間周期の滑らかな線を描く。また、毎月の平均気温は12カ月を周期とする滑らかな線を描く。

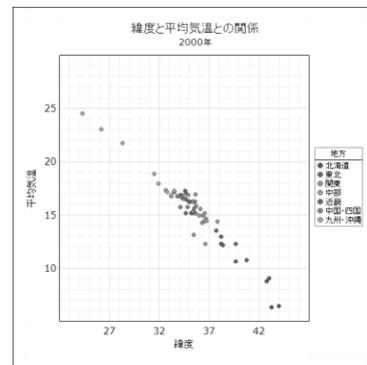


図6. 緯度と年平均気温との関係(2000年)

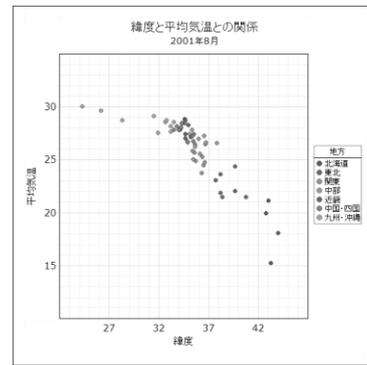


図7. 緯度と8月平均気温の関係(2001年8月)

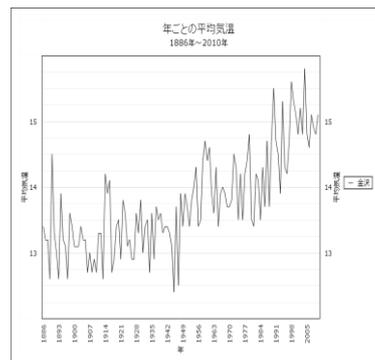


図8. 年平均気温の推移(金沢1886~2010年)

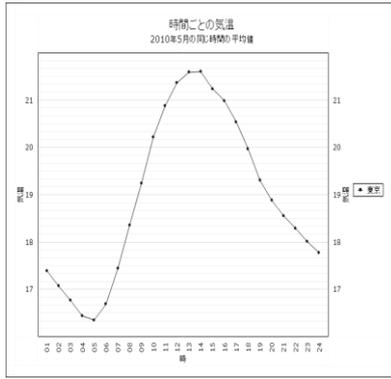


図9. 毎時の気温の月平均値の24時間の推移(東京・2010年5月)

(4) 基本的統計パターンから導出される高次のパターン：以上において、気温データから導き出される基本的パターンを紹介したが、これらのパターンを地域間あるいは時点間で比較することにより、パターン上のパターンとでも呼べる、より高次のパターンが導き出される。これを以下に示す。

① 経度と最低気温を示す時刻の関係：毎時の年平均気温を見ると、東日本では午前5時に最も気温が低い観測地点が多く、西日本に行くと6時が最も低い観測地点が多くなる。太陽が東から西へ移動することの一日の気温の変化に与える影響が見られる。ただし、最高気温に関しては、このような明確な傾向は見出されていない。

② 緯度と季節の変化による気温の変動幅の関係：月ごとの平均気温を見ると、夏季と冬季の気温差は、緯度が上がるほど大きく、緯度が下がるほど小さくなる傾向にある。例として図 10 に、札幌、横浜、那覇の比較を示す。

③ 緯度と近年の気温上昇の程度の関係：近年、たとえば 2001 年から 2010 年の 10 年間の毎年の平均気温を見ると、緯度が上がるほど大きく上昇しており、緯度が下がるほど上昇の程度が低い傾向にある。例として図 11 に、那覇、鹿児島、津、仙台、札幌の比較を示す。

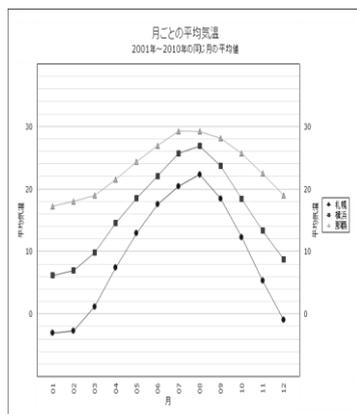


図10. 毎月の気温の平均値(2001年～2010年)
上：那覇，中：横浜，下：札幌

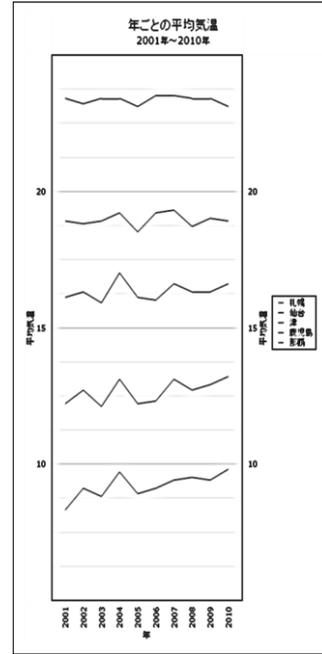


図11. 年平均気温の推移(2001年～2010年)
上から那覇, 鹿児島, 津, 仙台, 札幌

(5) まとめ

① 気温データ・グラフ表示システムの機能：完成した気温データ・グラフ表示システムでは、前述のように気温データが内在する基本的統計パターンを示すグラフを作成し保存することができる。データマイニング入門教材として最低必要な機能を備えていると考える。

② 気温データ・グラフ表示システムの操作性：気温データ・グラフ表示システムでは、多くとも数回のマウス操作でグラフを作成することができる。その際、表題に入力や表示範囲の設定などの煩雑な入力操作は不要である。また、地点や期間あるいは時点などを簡単に変更して複数のグラフを連続して作成することができる。これらは、短時間で多くの試行錯誤が可能になり、学習効率の向上ならびに学習内容の充実に寄与すると期待できる。

③ 基本的統計パターンの視認性：気温データ・グラフ表示システムで作成できるグラフから、前述のように直線相関、曲線相関、傾向、循環のパターンが認められる。しかも、いずれもデータ解析の初心者が見覚によって明確に認識できるものである。一方、そのようなパターンの存在は、一般の人々も日常の中で体験的に理解している。「何となく分かっていること」がデータを通じて明確になることは、「面白い」という感情を引き出すとともに、データ解析の有効性を認識するのに効果があると思われる。

④ 高次のパターンの発見：基本的パターンの中に更に高次のパターンを見出す体験は、データ解析やデータマイニングの面白さや意義を理解するための教育に有効だと考えられる。その上、得られた知識から高次の新たな知識が得られる「知の連鎖」体験は、単にデータマイニングやデータ解析のみならず、科学そのものへの興味を駆り立てる切欠をつくることも期待できる。

⑤ 平均値の有効性とデータ蓄積の意義：前述した統計的パターンの多くは、月または年での平均値を用いたグラフから抽出されている。逆に平均値を用いない場合は、パターンを導き出すことが難しいことが多い。たとえば、24 時間における毎時の気温の推移は、月平均を用いてグラフを作ると 24 時間を周期とした滑らかな曲線が得られるが、1 日ごとの毎時の気温のグラフを見ると、毎日異なったパターンが示され、視覚から滑らかなパターンを想像することは容易ではない(図 12、図 13 参照)。このことを学習者が体験できると、平均値の有効性やデータを蓄積することの意義が認識できると期待される。

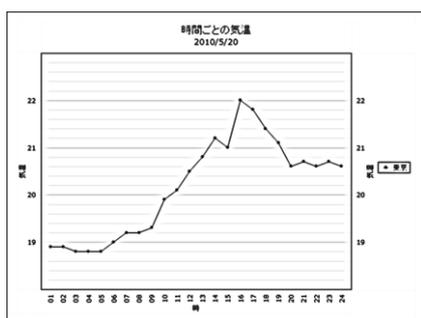


図12. 1時から24時までの毎時の気温の推移
(東京・2010年5月20日)



図13. 1時から24時までの毎時の気温の推移
(東京・2010年5月26日)

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 4 件)

- ① 今澤 明男、直江 伸至、「気温データ・グラフ表示システム」から得られるデータパターン、創造技術教育、査読有、14 巻 1 号、2014、ISSN 1884-1546

- ② 今澤 明男、直江 伸至、学習者のデータ科学への興味を喚起する気温データ解析ソフトウェアの紹介、創造技術教育、査読有、13 巻 1 号、2013、30-35、ISSN 1884-1546
- ③ 今澤 明男、直江 伸至、気温データを用いた環境教育と情報技術の活用、環境技術、査読有、Vol. 41、No. 11、2012、691-694、ISSN 0388-9459
- ④ 今澤 明男、直江 伸至、気温データを用いたデータマイニング入門教育の提案、創造技術教育、査読有、11 巻 1 号、2011、5-8、ISSN 1884-1546

[学会発表] (計 3 件)

- ① Akio Imazawa and Nobuyuki Naoe、Using Air Temperature Data of All the Prefectural Capitals in Japan as Teaching Material for Introductory Data-mining Classes、5th International Congress on Engineering Education (ICEED2013)、査読有、2013 年 12 月 4 日、Dorset Grant Subang Selangor (マレーシア)
- ② 今澤 明男、直江 伸至、計量感覚育成教材としての都道府県庁所在地気温データの検討、日本行動計量学会第 41 回大会、査読無、2013 年 9 月 4 日、東邦大学 (千葉県船橋市)
- ③ Akio Imazawa and Nobuyuki Naoe、Use of Air Temperature Data in Education to Develop Statistical Intuition、3rd International Congress on Engineering Education (ICEED2011)、査読有、2011 年 12 月 8 日、Sunway Lagoon Resort Malaysia (マレーシア)

[その他]

ホームページ

<http://trendy.co.jp/kion/>

6. 研究組織

(1) 研究代表者

今澤 明男 (IMAZAWA, AKIO)

金沢工業高等専門学校・グローバル情報工学科・教授

研究者番号： 20148141

(2) 研究分担者

直江 伸至 (NAOE, Nobuyuki)

金沢工業高等専門学校・電気電子工学科・教授

研究者番号： 00249781