

# 桜の開花に関する統計的分析

2003MM081 丹羽章仁

指導教員: 田中豊

## 1 はじめに

桜は日本全国に分布し春の訪れを告げる花として日本人に古くから親しまれている。そのため気象庁が毎年各地の桜の開花日を予想し発表しており、とくに近年は桜に関するイベントも多く開花予想への注目が高まっている。本研究では地域特性をつかむとともに、気象データを使い気象と開花の関係を調べるとともにより精度の高い予測式作成を目指す。なお桜とはソメイヨシノを指し、開花とは花が数輪以上が咲いた状態のことを指す。

## 2 データについて

データは開花日、気象データそれぞれ各気象台で観測されたものを使用する。観測期間は1953年～2007年の55年、観測場所は鹿児島県:種子島～北海道:稚内までの全国78ヶ所のものを使用する。

## 3 開花の地域特性

### 3.1 散布図

全国78地点を西から順に番号をつけ過去10年の開花日の平均値をプロットした図である。縦軸に同年1月1日から開花日までの経過日数、横軸に観測地点を南から順に番号付けした値を使用した。3月中旬から5月上旬にかけて九州～北海道へ順に南から桜が開花していくのがわかる。

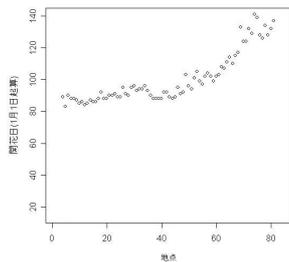


図 1: 散布図

### 3.2 クラスタ分析

各地点の過去10年間の開花日でクラスタ分析を行った。分析は標準化ユークリッド距離、最遠隣法で行った。結果、各群とも近い地点が集まっており、近郊の都市においては桜の開花傾向が似ていることがわかる。

**第1群:北海道～東北** 開花時期:4月下旬～5月上旬  
南は盛岡から北は稚内までの地域の都市。1年を通して日本の中では一番寒い地域の群となる。

**第2群:関東～九州** 開花時期:3月下旬  
今回の観測地点の半分の地点を占めている。本州・四国・九州のの大部分を占める地域の群となる。

**第3群:東北地方・日本海側** 開花時期:4月上旬～4月下旬  
南は島根県西郷から北は山形までの日本海側の都市が多く見られる。本州の都市の中でも比較的気温が低い地域の群となる。

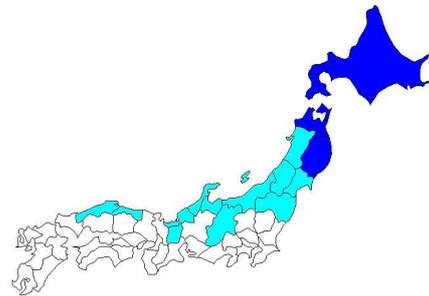


図 2: クラスタ分析の結果

### 3.3 主成分分析

10都市(札幌,仙台,新潟,東京,名古屋,大阪,広島,高松,福岡,鹿児島)の過去55年間(1953年-2007年)の開花日を使い主成分分析を行った。累積寄与率83.2%となった第2主成分までを考察する。

**第1主成分** 全て負の値を取ったことから、総合的な評価と判断できる。このことから開花の早遅の「開花時期」を示しているといえる。

**第2主成分** 新潟以北の3都市で正の値、東京以南の7都市で負の値となった。北の地域と南の地域で値が大きな差があることや北の都市ほど値が大きいことから「開花時期の地域差の大きさ」を示しているといえる。

また、主成分得点を時系列的にプロットすると図3,図4となる。第1主成分得点は「開花時期」を表している。次第に値が大きくなっていくことから時期にして1975年頃から開花の時期が早くなる傾向があることがわかる。

第2主成分得点は年によって変化はあるものの一定の傾向はみられる。つまり「開花時期の地域差の大きさ」という点ではこの55年間ほとんど変化がないといえる。

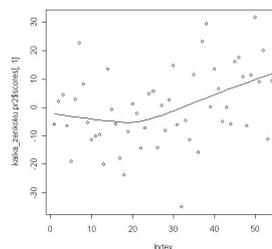


図 3: 第1主成分得点

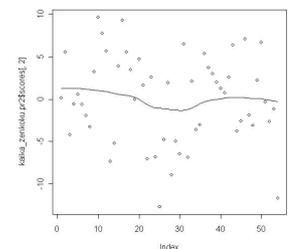


図 4: 第2主成分得点

## 4 開花と気象の関係

### 4.1 開花日の重回帰分析

気象条件が開花日にどのような影響をもたらすかを調べるため重回帰分析を行った。説明変数には開花前年の4月から開花の3月までの12ヶ月分の気象データ(降水量, 気圧, 気温, 湿度, 日照時間) 55年間分を用いた。これを3ヵ月ごとの4群に分け気象データから開花日の重回帰モデルを作成し, さらに変数選択を行う。そして変数を6ヵ月ごとの2群に, 最終的には1年1群になるまで繰り返すことにより当てはまりのよいモデルを求めた。

	回帰係数	p値
(Intercept)	-767.8	0.0011
前年5月:気圧	0.819	0.0070
前年5月:日照時間	0.015	0.0465
前年11月:降水量	0.012	0.0412
12月:気圧	0.313	0.0606
1月:日照時間	-0.034	0.0065
2月:降水量	-0.027	0.0065
2月:気圧	-0.253	0.0508
2月:日照時間	-0.035	0.0082
2月:気温	-0.974	0.0007
3月:湿度	0.176	0.0040
3月:気温	-3.251	3.42e-16
調整済決定係数:0.9039		

表 1: 開花日の重回帰分析結果

### 4.2 開花の予想-気象庁予想方式の改良-

現在気象庁では3月から4月にかけて週一回各地の予想開花日を発表している。

$$19.6 \leq \sum_D \exp \frac{9.5 \times 10^3 \times ((t + 273.15) - 288.15)}{((t + 273.15) \times 288.15)}$$

上記式( $t$ : 予想気温)より温度変換日数を求め, 起算日を2月15日とし前年秋からの気温経過と気温予報を当てはめた値の積算値が19.6を超えた日を開花予想日として発表している。図5, 図6は全国78地点の2002年, 2007年の実測値を使用し, 求めた予測日と実際の開花日をプロットした図である。

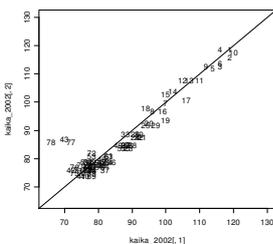


図 5: 2002年予想と実際

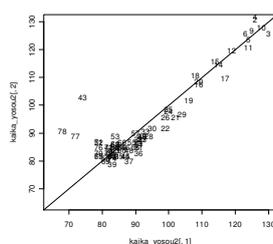


図 6: 2007年予想と実際

2007年のデータを使用し予測式と実際の開花日の残差を目的変数に, 説明変数に気象データ(気圧, 降水量, 最高気温, 最低気温, 気温差, 湿度, 日照時間)を使い重回帰分析を行うことにより残差が減らせないか試みた。データは予想式で2月15日から実際に開花した日までの実測値を使用した。結果, 自由度調整済決定係数は0.132と低いがp値から気圧, 最低気温, 湿度が関係があると言える。

	回帰係数	標準誤差	p値	
(定数項)	7.865	5.532	0.159	
気圧	0.013	0.007	0.044	*
降水量	0.002	0.006	0.714	
最高気温	0.108	0.275	0.695	
最低気温	-0.792	0.241	0.001	**
気温差				
湿度	-0.251	0.092	-0.008	**
日照時間	-0.008	0.005	0.126	
自由度調整済決定係数: 0.1322				

表 2: 気象庁予想残差の重回帰分析

求めた回帰式を使用することにより2007年予想で平均±2.9日あった残差が±2.2日に修正できた。2007年のデータで導出した重回帰式に2002年のデータを使用したところ, こちらでも±2.41日あった残差が±2.08日とわずかながらも改良できた。予想式から開花日を導出する際は, 気温と同時に気圧, 最低気温, 湿度のデータの傾向を見ることにより精度の高い予想が可能と考えられる。

## 5 おわりに

桜の開花には春先の気象条件が大いに関係あるということが確認でき, 気象データにより開花の予想も十分可能であることがわかった。中でも2-3月の気温の高さが開花と密接な関係があり, 他の気象条件に比べ圧倒的に影響が大きいと言える。また, これと関連してか, 全国的に開花が全国的に早まっているのが確認できた。現在地球は温暖化に向かっている。このままでは将来桜は冬に開花する花になってしまうかもしれない。

## 参考文献

- [1] 青木繁伸: Rによる統計処理:  
<http://aoki2.si.gunma-u.ac.jp/R/>.
- [2] 気象庁 Japan Meteorological Agency:  
<http://www.jma.go.jp/>.
- [3] 北山貴之: 花粉飛散数に関する統計的分析, 南山大学数理情報学部数理科学科卒業論文(2007).
- [4] 田中豊・脇本和昌: 多変量統計解析法, 現代数学社(1983).