

サクラの開花日予測に関する考察

聖光学院中学校 矢向俊貴(中3) 田尻裕太(中3) 奥津新太郎(中3) 脇山大地(中3) 鈴木健太(中3) 李 博之(中3)

はじめに

サクラ以外の生物とサクラ開花日の相関関係を調べることによりサクラの開花を予測した。

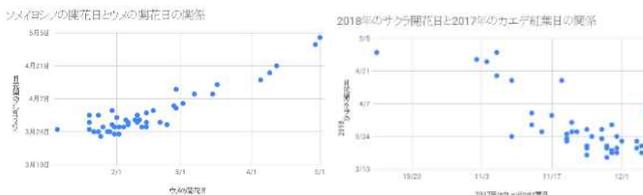
1. 研究目的・目標

同じ生物の中で関係があるのではないかという意見があり、サクラより先に見られる生物季節との相関を取り、関係式を作成し、それを補正する方法でサクラの開花日をなるべく早い段階で予測する。

2. 研究方法・結果・考察

① 他の生物による開花日予測

休眠打破日を正確に予測できる高い相関は今のところ見つかっていない。そこで、観測記録のある「ウメの開花日」「カエデの紅葉日」「ウグイスの初鳴き」とサクラの相関をそれぞれ取る。ウメの開花日とソメイヨシノの開花日の相関は左、「カエデの紅葉日」とサクラの開花日の相関は右のグラフのようになった。



各都市の20年分のデータからは、低緯度地域ほどサクラの開花とカエデの相関が高く、高緯度地域ではウメとサクラの相関が高いことが分かった。サクラとカエデは半年ほど離れているデータではあるがそれなりに高い相関がある。そこで、今回は高緯度地域と中緯度・低緯度地域に分けて、関係式を求めていく。

※ここで言う中緯度・低緯度地域と高緯度地域の定義は、高緯度地域=東北地方、中緯度・低緯度地域=東北地方以外とする。

(1) 高緯度地域の場合

高緯度地域ではウメを使う。ウメの開花日を x 軸、サクラの開花日を y 軸に取りると、グラフと式は下のようになった。

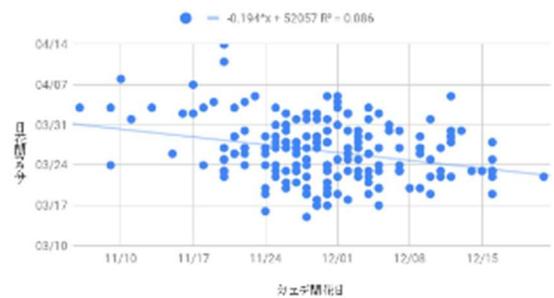


$$1.87x - 37828R^2 = 0.755$$

(2) 中緯度・低緯度地域の場合

中緯度・低緯度地域ではカエデを使い、同様にグラフと式を作成した。(高緯度地域より範囲が広いのでデータも多い)

2013~2018年シーズン カエデとサクラ



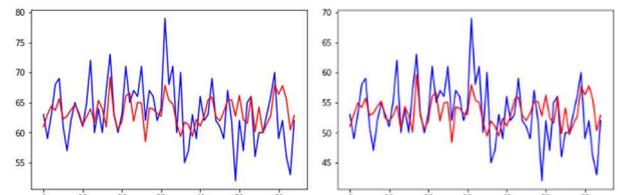
$$-0.194x + 52057R^2 = 0.086$$

高緯度地域のような良い結果は得られなかった。原因としては、データの数があまりにも多すぎたことが考えられる。これからは地域を精選し改善したい。

② 開花日予測

「梅の開花日」「ウグイスの初鳴き」「カエデの紅葉日」と「サクラの開花日」の間に相関関係が見られることが分かったため、Python を用いた機械学習によってサクラの開花日を直接的に予測することを試みた。

前者 3 つを説明変数(予測する元データ)、後者 1 つを目的変数(予測するデータ)とし、1953 年から 2018 年までの福島のデータとして予測をした。福島を選んだ理由としては、先ほどの考察でウメとサクラの相関は高緯度ほど高いことが分かっているためである。その結果が以下のグラフである。



横軸は年、縦軸はサクラ開花日を表し、青が実際のデータ、赤が予測したデータである。左側がサクラだけデータを整数化(その日付-2/5)した時のグラフ、右側が説明変数も整数化したグラフで、平均誤差は左側が 3.473...日、右側 3.456...日とほぼ差がない。本校の別の発表で述べた「休眠打破日の確定は日長と関係ない」ということの一理由のつともなりうる。他にも、整数化をすべてするか、サクラだけするかなど、説明変数の数などで 14 パターン機械学習を試みたが、いずれもこの 2 つのグラフほど成果は出なかった。しかし、いくら 1 か月前から開花日が予測できて、まだ誤差が平均 3.5 日と大きすぎるので、これからはこも修正していきたい。

3. まとめ

今回はサクラ開花日に関する考察を行い、高緯度地域と中緯度・低緯度地域で分けることによりサクラ開花日を予測することができること、また、「ウメの開花日」「ウグイスの初鳴き」「カエデの紅葉」の 3 つとサクラの開花日に高い相関があることが分かった。

参考文献：気象庁 HP