

教師無し機械学習で日本一安全な市町村を突き止める

兵庫県立姫路西高校 藤原正道 播本涼太郎 横井友

1.研究動機、研究目標

自然災害が多発する日本において一番安全な場所があるとするならば、その地を様々な用途に利用できるのではないかと考えた。
本研究では、上記のために災害発生頻度を市町村別に導出し、可視化することで日本一安全な市町村を突き止めることを目標とする。
災害発生頻度は雨と地震に関するオープンデータから求める。

2.使用データ

①過去の地震動発生回数

使用データ：震源データベース(気象庁)
震度5弱以上の揺れを観測した回数
(期間：1919年1月1日～2020年12月31日)
データ収集：一つの地震につき、震度5弱以上を観測した地点のある市町村を記録していく。全557地震で実施。
一市町村につき震度5弱から震度7までの揺れを観測した回数を合計する。

②今後の地震動発生確率

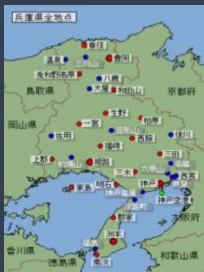
使用データ：確率的地震動予測地図(J-SHIS)
30年間で震度5弱以上となる確率
データ加工：5次メッシュデータをpythonを用いて市町村データに変換。
一市町村に該当するメッシュのデータの平均値をとった。

③一日降水量

④一時間降水量

使用データ：過去の気象データ(気象庁)
1日で降水量200mm以上を観測した回数
1時間で降水量50mm以上を観測した回数
(期間：1991年1月1日～2020年12月31日)
データ収集：赤色の観測点を持つ市町村のデータのみを収集。
赤色の観測点を持たない市町村は近くの赤観測点を持つ市町村のデータと同じ値を使用。
データ加工：目安となる降水量は気象庁のホームページを参考にした。

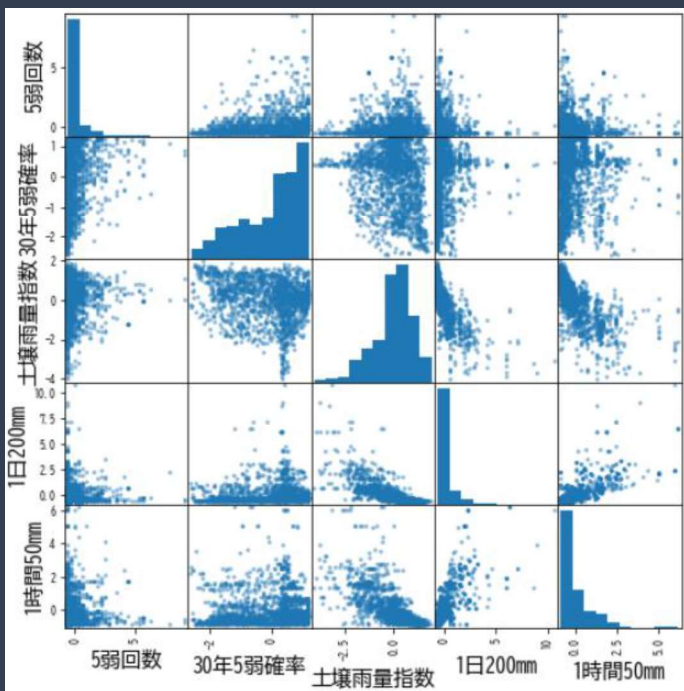
例：兵庫県



⑤土壌雨量指数

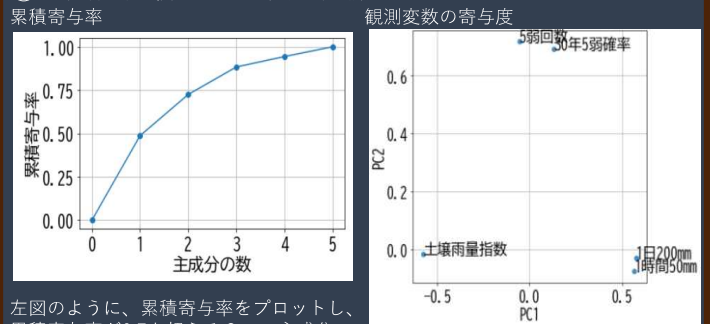
使用データ：雨に関する50年に一度の値一覧(気象庁)
50年に一度の土壌雨量指数
データ整理：全データを(-1)倍

使用データはすべて標準化する



3.分析

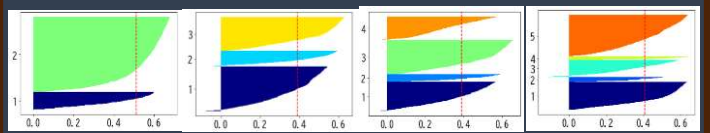
①主成分分析による次元削減



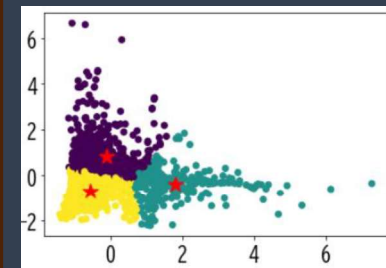
左図のように、累積寄与率をプロットし、累積寄与率が0.7を超える2つの主成分で表す。固有値は第一主成分が2.439、第二主成分が1.188である。
第一主成分と第二主成分における観測変数の寄与度をプロットし、右図で表した。第一主成分を豪雨災害の起きやすさ、第二主成分を地震災害の起きやすさと解釈。

②k-means法によるクラスタリング

シルエット分析によってクラスタ数を決定する。クラスタ数2～5までのシルエット図を作成した。



今回は合計クラスタ数の差が比較的小さいクラスタ数3でクラスタリングする。



左図において、紫色→地震災害に晒されやすい
緑色→豪雨災害に晒されやすい
黄色→比較的安全と解釈することとする。

③災害発生頻度を定義

本研究では、災害発生頻度について

$$\text{災害発生頻度} = \text{第一主成分} + \text{第二主成分}$$

と定義する。

4.結果の可視化と考察

クラスタリング結果と災害発生頻度をそれぞれ日本地図で可視化した。
主に太平洋側で災害が発生しやすいことがわかる。特に西日本では豪雨、東日本では地震の影響が顕著である。
また、災害発生頻度が一番低い市町村は北海道中頓別町であると結果が出た。
よって中頓別町を日本一安全な市町村であると考察する。それ以外にも下の地図における最も薄い黄色のエリアも日本では比較的安全であると考察する。

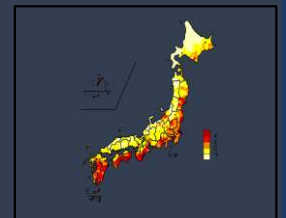
5.今後の展望

現時点では、導出された数値が有用かはまだ判断することができない。
今後は今回の結果の実証、また、その他の自然災害(降雪や火山噴火など)や人口密度なども考慮して、さらに多角的で、さらに意義のある研究をしていきたい。

↓クラスタリング結果



↓災害発生頻度



6.使用データURL

気象庁 震源データベース <https://www.data.jma.go.jp/svd/eqdb/data/shindo/index.html>

過去の気象データ検索 <https://www.data.jma.go.jp/obd/stats/etrn/index.php>

雨に関する各市町村の50年に一度の値 [1-50ame.pdf \(jma.go.jp\)](https://www.jma.go.jp/1-50ame.pdf)

J-SHIS 確率的地震動予測地図 <https://www.j-shis.bosai.go.jp/map/JSHIS2/download.html?lang=jp>