

## MH2 確率勾配ブースティングを用いた耕作放棄の予測モデルの開発ならびに その発生要因の地域特性と経年変化の分析

Developing a Predicting Model of Cultivated Land Abandonment using Gradient Boosting Decision Tree and  
Analyzing Regional Characteristics and Temporal Change of Drivers

指導教員 町村尚准教授・地球循環共生工学領域

28H17040 左右田 篤 (Atsushi SOUDA)

**Abstract:** In Japan, farmland abandonment has been increasing in recent years, and causes upsetting of natural symbiotic systems in farmlands. Therefore, it is required to identify the driving factors of abandonment and forecast future abandonment in order to care the progress of abandonment in advance. In this study, I developed a process to build a diagnostic and predicting model of farmland abandonment with high accuracy by applying a machine learning algorithm: gradient boosting decision tree. Then, I identified the difference of important abandonment driving factors and characteristics by region and identified the temporal change of them. It was found that factors of abandonment changed depending on region and time. It was suggested that the most important factor was the share of non-farmers possessing farmland.

**Keywords:** cultivated land abandonment, machine learning, census of agriculture and forestry

### 1. 背景と目的

日本では耕作放棄が増加しており、耕作放棄の圧力が高い地域を特定するとともに、促進要因の緩和策を提案する意思決定支援システムの開発が求められる。これに対して先行研究では機械学習技術を用いて地域別の耕作放棄の要因分析と予測を同時に行えるモデルを開発されているが、より詳細な要因の特定の必要性が指摘されている<sup>1)</sup>。そこで本研究では、耕作放棄のモデルを開発し、構築したモデルによる地方別の耕作放棄要因の分析、5年の経過による耕作放棄要因の変化を分析する。

### 2. 耕作放棄予測モデルの構築プロセス

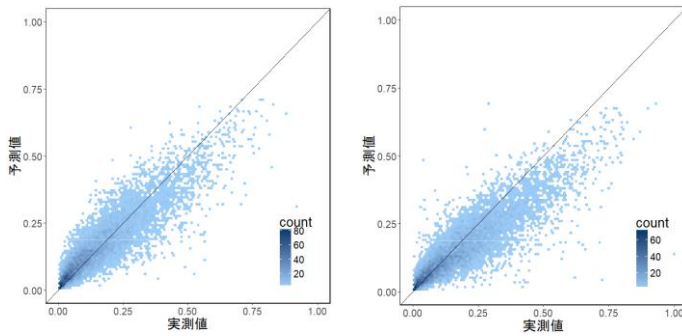
#### (1) 対象地の選定と分析データの収集

沖縄県を除く46都道府県を対象とした。旧市区町村を基礎として8,778のサンプルを選定した。モデルの説明変数と応答変数を抽出するため、農林水産省が公開している農林業センサス<sup>2)</sup>から経営体の営農情報、国土数値情報<sup>3)</sup>の5次メッシュスケールでの標高と傾斜、都市地域データから都市計画区域情報、人間活動量の空間指標である夜間光データ<sup>4)</sup>を使用した。

#### (2) 将来予測モデルの設計と学習

$$AFLR_i = \frac{AFL_i}{TFL_i + AFL_i} \quad (1) \quad AFLR_{i+5} = f_{i+5}(x_i) \quad (2)$$

式(1)の耕作放棄地面積率 $AFLR_i$ を応答変数とした。 $AFL_i$ 、 $TFL_i$ は*i*年の耕作放棄地面積 [ha]、経営耕地面積 [ha] を表す。*i*年の営農情報 $x_i$ から5年後の $AFLR_{i+5}$ を予測するためのモデル $f_{i+5}$ を式(2)で学習した。 $f_{i+5}$ に $x_{i+5}$ を入力して $AFLR_{i+10}$ の予測を行い、実測値と比較することでモデル $f_{i+5}$ の汎化性能を評価した。本研究では $x_{2005}$ から $AFLR_{2010}$ を予測する $f_{2010}$ モデル、 $x_{2010}$ から $AFLR_{2015}$ を予測する $f_{2015}$ モデルを構築した。また、地方別にデータセットを分割し、地方*j*特有の要因分析が可能な $f_{i+5,j}(x_{i,j})$ モデルを構築した。説明変数は労働力、経済的、経営体、農地流動性、農地集約度、作付け作物、地理的、都市計画要因に分類され、欠損値処理とVIF基準による前処理をし、計75変数を採択した。識別器*f*には回帰木を弱学習器とする確率勾配ブースティングを適用した。10×10入れ子交差検証により学習汎化性能を評価し、ベストモデルのハイパーパラメータの組み合わせを探索した。



a) 交差検証の散布図      b) 汎化誤差の散布図

図1  $f_{2010,全国}$ の精度評価

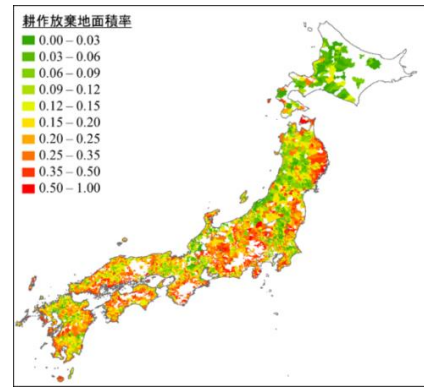


図2  $AFLR_{2020}$ の予測値 (旧市区町村スケール)

### 3. 結果

#### 3.1 予測汎化誤差の検証結果

図1 a), b)に、本研究で構築された $f_{2010,全国}$ の交差検証と予測汎化性能を評価した結果を示す。縦軸は予測結果、横軸に農林業センサスによる実測値を表した散布図を示している。交差検証の相関係数は0.89となり、学習の妥当性が示された。汎化誤差の相関係数は0.88となり、5年後の耕作放棄地面積率を高精度に予測できたといえる。図2に、 $f_{2015,全国}$ による2020年の耕作放棄予測マップを示す。前年度に高耕作放棄であった地域を中心に、耕作放棄が全国的に進展する結果となった。

#### 3.2 耕作放棄要因の地域差と経年変化の分析

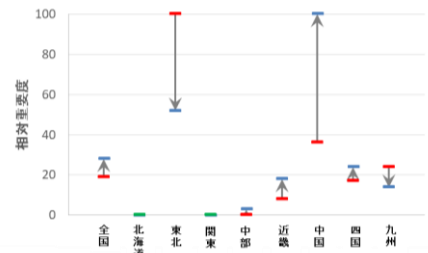
図3 a), b)に全国共通で重要であると特定された、非農家が貸付耕地を所有する割合、農業従事者の平均年齢の影響度の変化を示す。縦軸は最も重要度が高い変数を100とした相対重要度で、横軸は地方を表している。図3 a)では東北や中国、図3 b)では北海道や四国で耕作放棄要因の変化が大きくみられた。全国では地方の特色が平均化され変数重要度の変化は小さくなった。両年代において、土地持ち非農家の割合に関する変数が全国的に最も重要な変数であると特定された。夜間光は地理的要因のうち最も重要な変数と特定され、衛星データを活用した耕作放棄予測の有効性が示された。

### 4. 今後の課題

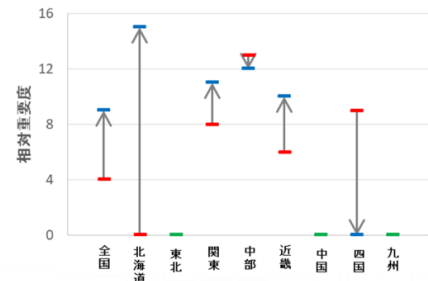
地域や時代の変化による耕作放棄要因の変化が明らかになったことから、マクロな社会経済などの動向のティッピングによる、システム変化に柔軟に対応できる将来予測モデルの構築が必要である。新たな変数候補として、中山間地域等直接支払制度による収益が少ない農家への交付金といった政策的な変数や、農地転用に関わる不動産価格などの社会的変数の導入を検討する。

### 参考文献

- 1) 松井孝典, 宇賀田徹, 町村尚:機械学習アルゴリズムによる耕作放棄の要因分析および予測モデルの開発, 土木学会論文集 G (環境), vol. 70, No. 6 (環境システム研究論文集第42巻), II\_131-139, 2014.
- 2) 農林水産省:農林業センサス, <<http://www.maff.go.jp/j/tokei/census/afc/>>, 2019年2月参照.
- 3) 国土交通省国土政策局国土情報課:国土数値情報ダウンロードサービス, <<http://nlftp.mlit.go.jp/ksj/>>, 2019年2月参照.
- 4) NOAA's National Centers for Environmental Information, <<https://www.ngdc.noaa.gov/ngdc.html>>, 2019年2月参照.



a) 非農家が貸付耕地を所有する割合



b) 農業従事者の平均年齢

図3 変数重要度の経年変化

— 2010 — 2015 — 変化無し