

論文の内容の要旨

論文題目 衛星データと深層学習を活用した農業保険開発の研究
(A study on development of agricultural insurance using satellite data
and deep learning)

氏 名 桑田 賢太郎

2050年の世界人口は、96億人に達するといわれており、2015年時点の約70億人から、およそ28%も増加すると予測されている。この増加の大部分には、アジアやサハラ以南アフリカの人口増加が寄与すると考えられており、食料需要を満たすためには現時点の穀物生産量の約1.6倍の増加が必要といわれている。穀物生産の安定化と増加が望まれている一方で、世界各地で気候災害による農業セクターへの甚大な被害が発生している。世界を代表する穀倉地帯として重要な役割を担うアメリカでは、2012年に歴史的な大干ばつが発生し、アメリカ本土のおよそ65%が被害を受けたとされている。400億ドルの損失が発生したとされ、その大部分が農業分野によるものであった。その他にも、大規模な台風や洪水などの気候災害が、アジアやアフリカなどでも頻発し、国家レベルの損失を与えている。このような天候リスクを軽減するため、農業保険が注目されている。

農業保険は、災害などで農作物が損害を受けた場合、その損失を補填し、農家の収入を保証するものである。世界の市場規模は、保険料ベースで約3兆円といわれており、大半をアメリカとカナダが占め、中国で急激に拡大している。アジアやアフリカ諸国でも導入が期待され、今後も成長市場であるといわれている。農業保険は、従来型とインデックス型に分類される。従来型は、調査員が被害を査定することで保険金を支払う保険であり、モラルハザードや逆選択、運用に要する高いコストが課題として考えられている。インデックス型は、従来型の課題を改善するものとして開発され、気象観測ステーションなどで観測される気象データをインデックスとし、ある条件を満たした際に保険金が自動的に支払われる保険である。しかし、インデックス型は、ベースリスクという支払われた保険金と実際の損失額が乖離するリスクが大きな課題とされ、途上国での普及が滞る原因とされている。本研究は、ベースリスクを改善するため、深層学習や衛星データを用いたトウモロコシ単収推定モデルを構築し、農業保険に適用することで、新しい単収推定保険を提案した。

本研究の対象は、アメリカのトウモロコシとした。アメリカは世界でも有数のトウモロコシ生産国であり、食料安全保障上、重要な作物であるといえる。2008年から2013年の1kmメッシュ気象データ Daymet と MODIS から推定した植生指数 EVI を入力データとし、アメリカ農務省が公開する郡レベルのトウモロコシ単収を推定するモデルを深層学習で構築した。単収推定モデルの精度を評価するために、cross validation の手法を適用した。結果、全米のトウモロコシ単収を単一の推定モデルで、 $R^2=0.780$ という高い精度で推定することが可能となった。既往研究は、州単位など非常に限定的な範囲を推定しているため、本研究の有用性は高いといえる。さらに、サポートベクターマシ

ンやオートエンコーダなど複数の機械学習アルゴリズムでも単収推定を行い、推定精度の比較を行うことで、深層学習の次元圧縮能力の有効性を確認した。

深層学習による単収推定モデルの性能を評価するため、「適用範囲の限定」、「汎化性の評価」、「衛星データの有効性」の3つの観点でそれぞれ実験を行った。適用範囲の限定では、推定対象範囲を全米からトウモロコシ主要産地である Heartland 地方に絞った。結果、 $R^2=0.89$ となり、推定精度の大幅な向上が確認された。トウモロコシは、アメリカの各地で栽培されているため、品種や栽培環境（施肥や灌漑など）が地域によって大きく異なる可能性がある。本研究の単収推定モデルでは、品種や栽培環境に関する特徴量が与えられていないため、対象地域を限定することで、推定精度が向上したと考えられる。汎化性の評価では、leave-one-out cross validation (LOOCV) による時間軸方向の汎化性を評価した。2008年から2013年までの6年間のデータのうち、各年を学習用データセットから取り除き、5年分のデータで学習させた単収推定モデルを、検証年のデータで評価した。結果、有意な推定精度は得られなかった。衛星データの有効性では、入力データを Daymet のみの場合と、MODIS EVI のみの場合で単収推定モデルを学習し、推定精度を確認した。結果、Daymet のみで単収推定した場合、Daymet と EVI で学習した場合より、推定精度が若干低下した。EVI のみでは、有意な推定精度が得られなかった。よって、深層学習で単収推定モデルを構築する場合、気象データや衛星の植生指数を組み合わせた入力データセットが有効であると考えられる。

上記で構築した単収推定モデルは、郡レベルのものとなっている。トウモロコシ主要産地の郡は、およそ35kmとなっており、郡内でもトウモロコシ生産が不均一であることから、ベースリスクが潜在することが懸念される。そこで、本研究では、郡レベルのトウモロコシ単収を1kmへ推定するダウンスケーリングを行った。深層学習で構築したトウモロコシ単収推定モデルに、1kmメッシュの入力データを入力し、郡レベルの単収から1kmメッシュの単収データにダウンスケーリングを行った。ダウンスケーリング単収は、収穫面積が小さい地域において過少推定の傾向があった。しかし、ダウンスケーリング単収と郡レベルの単収をそれぞれマッピングし、比較したところ、お互いの分布が非常に似ており、ダウンスケーリング単収はより詳細な単収分布を示していることがわかった。

ダウンスケーリングしたトウモロコシ単収を活用し、新しい農業保険である単収推定保険を提案した。保険設計には、アメリカ農地の大部分で適用されている収入保険を参考にした。単収推定保険が農家の収入の保証状況の評価するため、トウモロコシ生産による営業利益（売り上げからコストを差し引いた額）に保険金を合算することで、赤字であった営業利益が黒字に転換できたかを確認した。2012年は歴史的に深刻な干ばつが見舞われた年で、広い地域で大きな損害を被ったが、単収推定保険を適用することで、天候不順による損害をカバーできたことが確認できた。一方で、メッシュ気象データの最大気温を用いて、天候インデックス保険を設計し、単収推定保険と同様に営業利益の保証状況を確認した。天候インデックス保険でも、干ばつ年において多額の保険金支払いが発生し、農家の収入の保証が達成されたが、損失額に応じた金額の支払いではなかった。営業利益が黒字であった地域に対して、より多額の保険金支払いが発生したため、利益格差の拡大が生じたことから、ベースリスクが存在しているといえる。単収推定保険では、農家の営業利益に応じて補償が成されているため、損失が大きい地域に対して、多額の保険金支払いが実行された。よって、単収推定保険は、天候インデックス保険よりもベースリスクが低減された保険金支払いが可能であることが確認された。さらに、アメリカでは、その他の天候リスクや価格リスクによって損失を被るため、干ばつが発

生しなかった年においても営業利益の黒字を確保できない地域は存在した。単収推定保険では、天候インデックス保険が保険金支払いを行わなかった地域や年に対して保険金支払いを行い、全ての損失に対して補償が可能となった。

農業保険は、途上国への導入に向けて数多くのパイロットプロジェクトが実施されている。今後は、気候災害による減収の損失を補填するだけでなく、価格リスクなどあらゆるリスクを補填する農業保険商品のニーズが強まると予想されている。単収推定保険は、作物単収を直接的に推定することによって、収入保険など高度な保険金支払いの仕組みを設計することが可能となる。よって、天候災害や価格リスクなど、より多くのリスクへの対処に期待でき、幅広いニーズを満たすこととなる。また、衛星データや深層学習などのデータマイニング技術を適用することで、天候インデックス保険と変わらない低コストの保険運用が実現でき、モラルハザードや逆選択などのリスクを縮小することができる。今後は、打ち上げが計画されている将来の衛星の観測データを活用し、気象観測ステーションが十分に配備されていない地域において、本手法を導入する必要がある。また、単収推定の高精度化や汎化性向上を達成するため、長期間データの利用、さらに多くの特徴量入力やニューラルネットワーク構造の改善が課題となる。