

令和6年度 成績概要書

課題コード（研究区分）：3104-215251（経常（一般）研究）

1. 研究課題名と成果の要約

- 1) 研究成果名：栽培管理履歴の有効活用に向けたデータ管理の現状整理と秋まき小麦の分析事例（研究課題名：生産現場のビッグデータを活用した栽培技術診断手法の開発）
- 2) キーワード：データ駆動型、大規模データ、栽培改善、秋まき小麦、地域
- 3) 成果の要約：関係機関に蓄積した各種データを総合的に分析するためには、各データ内の用語統一と、各データを紐付けるための共通項目を明確にする必要がある。生産者単位の集計データを分析した結果、秋まき小麦の栽培管理履歴から収量への影響が大きい栽培管理項目を地域ごとに抽出可能であり、本データの有効性が示された。

2. 研究機関名

- 1) 代表機関・部・グループ・役職・担当者名：中央農業試験場・農業環境部・環境保全グループ・主査・八木哲生
- 2) 共同研究機関（協力機関）：（道内の農業協同組合）

3. 研究期間：令和5～6年度（2023～2024年度）

4. 研究概要

1) 研究の背景

日々の営農管理を記録した栽培管理履歴は、各生産者が栽培改善を進める上で重要な情報となるが、地域に蓄積した多様なデータと連携することで、実態に基づくよりの確な改善案を導出できる可能性がある。

2) 研究の目的

生産現場に蓄積された大規模データの有効活用を図るため、秋まき小麦を題材に、栽培管理履歴等のデータ管理に関する現状を整理するとともに、各種データを紐付けた具体的な分析事例を提示する。

5. 研究内容

1) 生産現場におけるデータ管理に関する現状整理（R5～6年度）

- ・ねらい：関係機関から収集した各種データの特徴や紐付け工程で生じる問題点を整理する。
- ・試験項目等：〔対象農協〕畑作地域の4農協（生産者数約80～300および300以上）〔対象作物〕秋まき小麦〔対象期間〕2017～2021年〔データの種類〕生産物、栽培管理履歴、圃場、土壌分析、肥料銘柄（以上、農協所有）、気象、土壌型（いずれも農研機構：メッシュ農業気象データ、日本土壌インベントリ）〔方法〕(1) 栽培管理履歴を中心にデータの特徴の確認 (2) 現状データの課題整理 (3) データの品質や紐付け程度（結合率）の確認。

2) 栽培改善を目指した大規模データの分析事例（R5～6年度）

- ・ねらい：上記1)において生産者単位で紐付けたデータセットを供試し、具体的な分析事例を提示する。
- ・試験項目等：〔分析対象〕上記1農協の秋まき小麦（2011～2023年「きたほなみ」、延べ生産者数約7500）〔データ〕生産者単位で整理した製品収量、播種時期、各生育期節における追肥の有無および気象を適宜紐付けて使用〔検討内容〕(1) 実態把握：地区ごとに、栽培管理履歴データの作業日から時期別の追肥実績を確認（地区は、気温、降水量および日射量で区分（例：栽培期間の積算気温は、地域 B>C>A））(2) 栽培管理の重要度評価：上記の地区ごとに、製品収量に対する播種および追肥の影響度合を評価（3）気象による収量推定：当該地域全体における各栽培年の製品収量を推定

6. 研究成果

- 1) (1) 栽培管理履歴データは、圃場ごとに作業内容や使用資材に関する情報を追記する方式であり、そのデータ規模は年間で生産者50あたり1800～3000であった（データ略）。圃場データは、地理情報に加え、圃場番号、作付作物、面積等が属性情報として付帯する。土壌分析データは、受入日、生産者番号、圃場番号、分析値の項目がある。また、各データには、互いを紐付けるための共通項目（結合キー）がある。(2) 各種データを紐付けて総合的に分析するためには、各データ内において用語や表記を統一する「標準化」や、それをシステム全体で統一する「一貫性」を改善することが望ましいと考えられた（表1）。特に、生産者番号や圃場番号は結合キーの役割があるため、表記方法等のルールを明確化する必要がある。(3) 栽培管理履歴データには、必須作業の欠損や日付の不整合等が約4-16%確認された（データ略）。圃場単位での各データ間の紐付け割合（結合率：図1）は、栽培管理履歴と圃場、資材（②、③）で、49-91および31-75%、圃場と土壌分析（④）で27-45%、圃場と気象・土壌（⑤）で約100%であった。一方、生産者単位での各データ間（生産物、栽培管理履歴、圃場）の結合率は94-100%と高かった（データ略）。
- 2) (1) 地区Bの追肥実績は、低収群では生育期節（起生期、幼形期および止葉期）が進むにつれて低下する傾向であったが、高収群では低下しなかった（図2）。一方、地区A・Cの傾向は地区Bと異なった（データ略）。栽培管理履歴データを生産者単位で集計することにより、地域の特徴を把握できると考えられた。(2) 製品収量に対する栽培管理の影響は、地区Bでは追肥、地区A・Cでは播種時期が相対的に大きく、また、地区Aでは両者の交互作用も大きい傾向にあった（図3）。これらは各々10～39 kg/10a相当の効果が見込まれた。地域ごとにデータ分析に基づいた栽培改善指針を策定することが重要と考えられた。(3) 農協管内における生産年ごとの平均的な製品収量は、当該地域の気象データから推定できた（RMSE = 42 kg/10a, データ略）。今後は、栽培当年の気象と栽培管理履歴を関連付けた分析が期待される。

以上、生産物、栽培管理履歴、圃場、気象のデータを紐付けた総合的な分析は、地域の実態把握や、実態を反映した栽培改善に有効と考えられた。データ品質の向上により、更に多様で詳細な情報の導出が期待できる。

<具体的データ>

表1 データの標準化・一貫性の向上に向けた課題

ポイント	課題	対応(例)
日付	西暦と元号の混在	西暦に統一
文字の種類	全半角、大小文字、特殊記号、カタカナ・平仮名・漢字の混在	入力ルールを設定
データの種類	文字・数値の型、桁数の不統一	型や桁数の固定、枝番のルール設定
用語・固有名詞	同一の作業・資材で異なる表記・略称が混在	正式名称に統一、略称の設定
単位	面積・重量・濃度の単位の混在	単位を明確化
欠損	「作業なし」か「記録漏れ」かが不明	データが無い場合は「0」で明示
その他	資材等の基本情報が不足、年次間で様式が不統一、誤入力、結合キーが不明確	基本情報の整理、様式の変更を極力回避、選択式の導入、データテーブルごとに結合キーを明確化

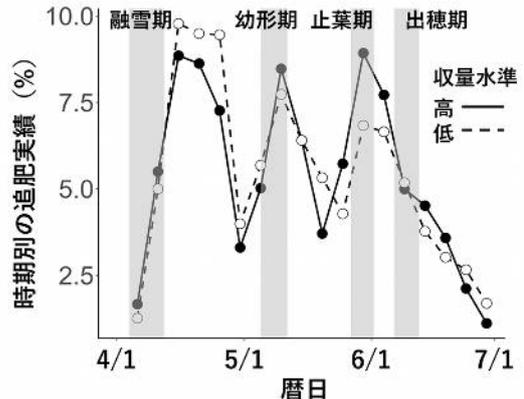


図2 地区Bにおける時期別の追肥実績

縦軸は、追肥の全実績数に対する時期別の実績割合。各プロットは、4/1-7/1を5日間ごとの時期別に集計した値(収量水準別の時期別の割合)。灰色の帯は、当該地域の気象データと発育モデル(H31指導参考事項)に基づき推定した各生育期節に達する平均的な範囲(2013-2023年の推定日の四分位範囲)。収量水準の区分は、各生産年において、製品収量が上位30%を「高」、下位30%を「低」として表示。

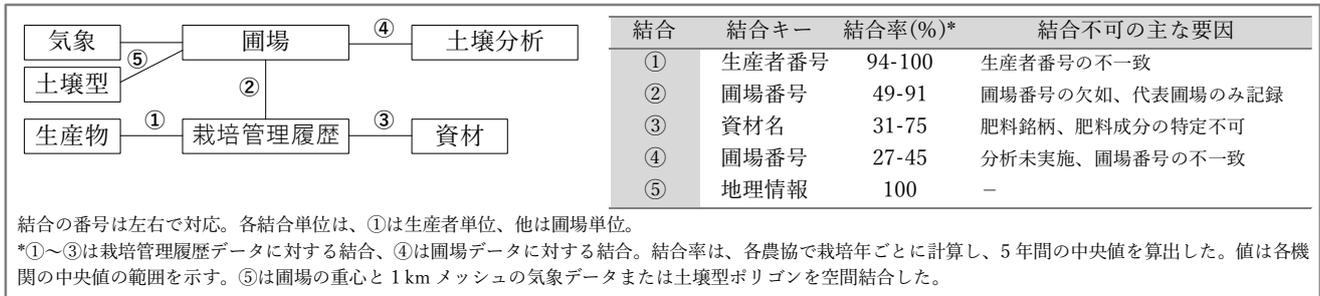


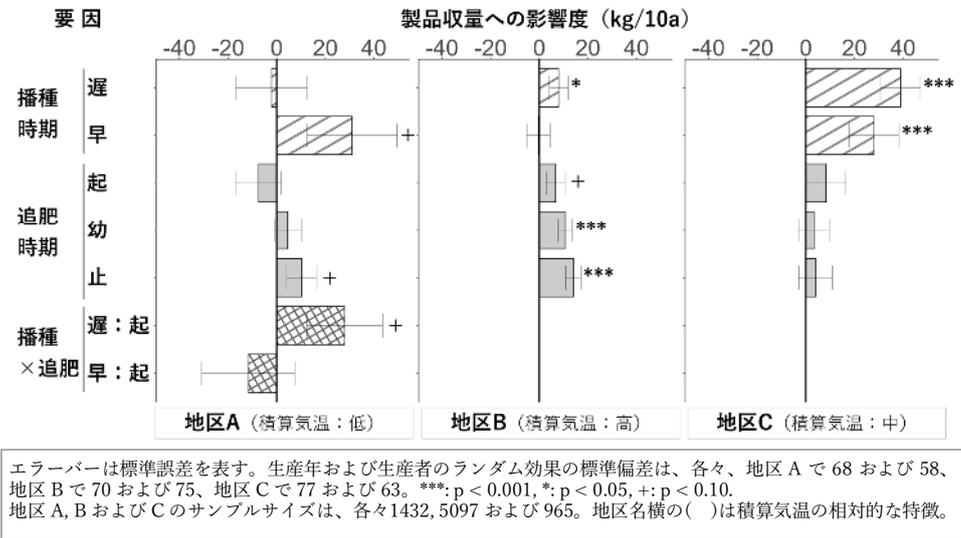
図1 各データの紐付け方法と結合状況

図3 播種・追肥の管理が製品収量に及ぼす影響

地域ごとに、製品収量を目的変数、播種時期(極遅:越冬前積算気温390°C未満、遅:同390~470°C、早:470°C以上)および追肥有無(あり、なし)を固定効果*、生産年および生産者をランダム効果**とした線形混合効果モデルにより、固定効果の影響を評価。

各固定効果の参照カテゴリは次の通り。播種時期は「極遅」、各生育期節の追肥は「なし」、播種と追肥の交互作用***は「播種「極遅」かつ追肥「なし」」。

* 全データに共通する平均的な傾向
** グループごとの偏りや変動を表す効果で、グループ内のバラツキから推定
*** 上記参照カテゴリに対して、「遅」かつ「起」の条件、または「早」かつ「起」の条件での効果を表す。



エラーバーは標準誤差を表す。生産年および生産者のランダム効果の標準偏差は、各々、地区Aで68および58、地区Bで70および75、地区Cで77および63。***: p < 0.001, *: p < 0.05, +: p < 0.10。地区A, BおよびCのサンプルサイズは、各々1432, 5097および965。地区名横の()は積算気温の相対的な特徴。

7. 成果の活用策

1) 成果の活用面と留意点

- (1) 農業生産現場においてデータを管理・運用する関係機関が、営農に関わるデータプラットフォームの整備等、データ駆動型農業の実現に向けた準備を進めるうえで参考にする。
- (2) 生産現場で収集された栽培管理履歴データを分析することで、地域全体の実態把握が可能となり、農協等が営農指導業務に活用することができる。

2) 残された問題とその対応

- (1) 本成績では、特定地域の播種および追肥に焦点を当ててデータの有効性を確認したため、令和7年度開始課題では、地域や栽培管理項目を拡大し、現場データの活用方法や解析方法を検討する予定である。

8. 研究成果の発表等 なし