

# 高解像度気象データと病虫害発生情報に基づく てん菜収量の早期予測技術の開発

国立研究開発法人 農業・食品産業技術総合研究機構  
北海道農業研究センター 寒地畑作研究領域 環境病虫害グループ  
研究員 村上 貴一

### 【要約】

てん菜生産では収穫物の輸送・加工プロセスの効率化のため、高い空間解像度でかつ早期に収量を予測する技術が求められる。本研究では、高解像度な気象データに基づいて省力的かつ非属人的に根収量・根中糖分を予測するモデルを作成し、モデルの早期予測精度を検証した。また、モデル収量と統計収量の差を生産者圃場の巡回調査に基づく褐斑病発病株率を用いて補正する手法を検討した。

## はじめに

てん菜は北海道における輪作体系を構成する基幹作物である一方で、近年の砂糖需要の低下や労働負担の大きさから、作付面積・作付戸数が減少傾向にある<sup>1)</sup>。てん菜生産の省力化および高収益化は、てん菜のみならず大規模畑輪作体系全体の生産性の安定化に資する基盤的な課題である。てん菜製糖業では、圃場から製糖工場までの収穫物の輸送と製糖工場の操業の時期・期間および人員をより早期にかつ適正に手配することが重要となる。これらの計画のために、現場では圃場観察とサンプリングによる成分分析を判断材料として根収量および根中糖分が予測されている。しかし、年次変動する気象条件や局所的に発生する病虫害被害のもとで、十分な空間解像度<sup>(注)</sup>で根収量および根中糖分を正確に予測することは困難である。持続可能なてん菜生産のためには、早期に、正確に、かつ省力的に収量を予測することで輸送・加工プロセスを円滑に進める技術が強く求められている。本稿では、気象データを活用した省力的な作物収量予測モデルの開発例と病虫害発生に関する情報を予測に生かすための方法に関する

取り組みを紹介し、モデル開発に関するいくつかの考慮すべき点について議論する。

なお、本調査は独立行政法人農畜産業振興機構の令和3年度砂糖関係研究委託調査により実施したものである。

(注)空間を表現する最小格子単位の細かさ。空間分解能。

## 1. 方法

### (1) 先行研究による収量予測モデル

月別または上中下旬別のように、特定期間の気象データとてん菜根収量または根中糖分の関係に着目した相関解析研究は、これまでに多数なされてきた<sup>2-6)</sup>。例えば、柳沢(2018)は、作物統計として整理された十勝管内のてん菜根収量および根中糖分と帯広測候所における気象要素との関係を重回帰分析により評価した。解析の結果、根収量の説明変数として4月下旬から7月上旬の平均気温、6月から7月の日射量、4月中旬から7月中旬の降水量、9月から10月の最高気温の4変数を選択し、根中糖分の説明変数として7月上旬から10月上旬の最低気温、9月から10月の日射量の2変数を選択し

たモデルを提案した<sup>5)</sup>。このモデルのようにある地域を対象として、特定期間の気象を説明変数とするモデルは、栽培暦が異なる他地域へと転用した場合に、正確な推定ができないことが予想される。

限定された地域を対象とするのではなく、北海道におけるてん菜生産全体の気象応答特性を評価するためのモデルも提案されている。梶山（2011）は温暖化影響評価を目的として全道平均された旬別気象データとてん菜収量統計を用いて、根収量を4月中旬から6月下旬の積算最高気温から、根中糖分を7月上旬から10月上旬の積算最低気温から、それぞれ予測する単回帰式を提案し、相関係数は0.673および-0.934であったと報告した<sup>6)</sup>。

## （2）省力的な収量予測モデル開発

次項でも触れるが、前項のような汎用的モデルをより細かい空間スケールでの事例に適用した場合、その予測精度は全道平均で見られた精度よりも劣る。従って、モデルの予測精度を高めることを意図する場合、特定の地方、市町村、地区、あるいは個々の生産者にとって好適なモデルをオーダーメイド的に作る必要があるが、そのためには必要となる作業工数と要求される知識が空間解像度に応じて増加するというジレンマがある。先行研究のように変数選択を手動で行う場合、この欠点は特に顕著である。また、先行研究で利用されていた気象データは気象台や測候所あるいはアメダスでの観測によるものであり、必ずしも畑地での気象を反映しているものではない<sup>7)</sup>、作成されたモデルを気象観測が実施されていない地点での予測には利用できない、といった制約もある。これらの問題を解決するため、高解像度気象データセットと機械学習フレームワークを利用し、半自動化された属人性の低い作物収量予測モデルの開発方法を検討した。

一例として、振興局および市町村レベルでのてん菜根収量および根中糖分を予測する汎用的なモデル

を作成した。旬別の気象要素を主な説明変数とし、統計年鑑<sup>8)</sup>に整理された根収量と根中糖分を従属変数とする回帰モデリングを行い、両者の積として糖収量を推定した。振興局レベルでの解析では、文献<sup>6)</sup>の方法に準拠して品種遷移の影響が補正された根収量および根中糖分を従属変数とした。

旬別気象データの取得には農研機構メッシュ農業気象データシステム<sup>9)</sup>を利用した。同システムはおよそ1キロメートルの空間解像度で、国内地点における農業生産性に関する日別（一部要素は特別）の気象要素実況値および予報値を配信している。4月上旬から10月下旬までの主要な気象要素（平均・最高・最低気温、降水量、全日日射量）の旬別平均値をメッシュごとに計算し、振興局・市町村ごとに平成21年度時点で土地利用が畑として登録されている面積の割合が10%以上のメッシュを抽出して平均を取ることで旬別気象データセットを作成した。多数年のデータのうちある1年のデータをモデルの検証用として除き、残る年のデータを用いてモデルのパラメータを決定するleave-one-year-out交差検証により、モデルの予測精度を評価した。予測精度の評価指標として、二乗平均平方根誤差（RMSE）および相関係数を採用した。GPU計算を行わない一般的なパソコンを解析に使用し、R（ver. 3.6）<sup>10)</sup>とその機械学習フレームワークであるcaretパッケージ<sup>11)</sup>を利用した。回帰のアルゴリズムとして多変量解析に広く適用される偏最小二乗回帰（partial least squares regression; PLS）<sup>12)</sup>とサポートベクター回帰（support vector regression; SVR）<sup>13)</sup>を利用し、ハイパーパラメータの探索範囲などはデフォルトの設定を採用した。また、比較対照として梶山（2011）の単回帰モデルに基づいた根収量と根中糖分、糖収量を計算した。なお、本研究の調査方法は、JSPS科研費JP19H00963の助成を受けたコムギ収量予測に関する研究<sup>14)</sup>をモデルとしたものである。

## 2. 早期予測の精度

SVRモデルはすべての指標で、PLSモデルは振興局レベルでの糖収量予測に関する相関係数を除くすべての指標で、従来の単回帰モデルを上回った(表、図1)。また、前述の通り、対象とする空間スケールを高解像度化した場合に予測モデルのパフォーマンスが低下することが確認された。

「はじめに」で言及したように、てん菜の根収量および根中糖分の予測は輸送・加工プロセスの効率化を目的とするため、早期予測による経済的効果が

大きい。そこで次に、予測への気象予報の利用の有効性を検討した。

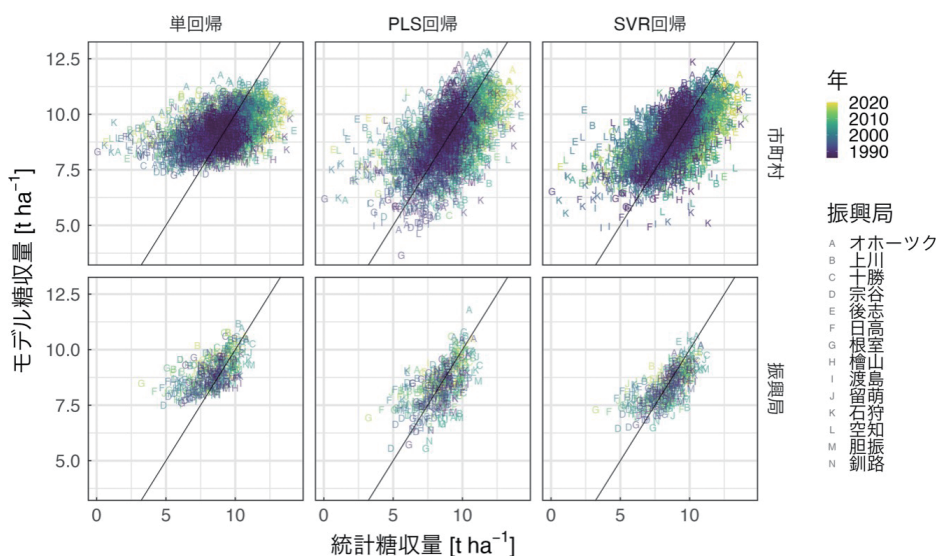
leave-one-year-out交差検証により作成されたモデルを検証データセットに適用する際に、検証データセットの旬別気象データを8月下旬までに限定し、9月上旬以降は予報値とする、というように、一部のデータを予報値に置換した場合の根収量および根中糖分を計算し、予測日に応じた予測精度を調査した。気象データの予報値もメッシュ農業気象データシステムを使用して取得されたものであり、日平均・最高・最低気温および降水量について

表 旬別気象データに基づく回帰モデルによるてん菜根収量、根中糖分および糖収量予測のパフォーマンス比較

対象スケール	モデル	相関係数			RMSE		
		根収量	根中糖分	糖収量	根収量	根中糖分	糖収量
					[t ha <sup>-1</sup> ]	[%]	[t ha <sup>-1</sup> ]
市町村	単回帰 <sup>5)</sup>	0.414	0.477	0.451	10.2	0.989	1.71
	PLS回帰	0.535	0.652	0.568	9.24	0.739	1.57
	SVR回帰	0.597	0.574	0.622	8.69	0.798	1.49
振興局	単回帰 <sup>5)</sup>	0.521	0.643	0.594	7.54	0.739	1.21
	PLS回帰	0.58	0.758	0.589	6.32	0.56	1.08
	SVR回帰	0.673	0.714	0.698	5.72	0.612	0.96

注：1986年から2020年のデータを解析に使用。

図1 比較した3種類の回帰モデルによるてん菜糖収量の予測値と統計値との関係



注1：1986年から2020年のデータを使用した市町村レベル(上)および振興局レベル(下)での解析の結果を示す。

注2：プロットはある年の1市町村あるいは1振興局のデータを表す。

は26日先まで、全天日射量については9日先までの予報値が、それ以降は平年値が与えられる。振興局レベルでの2011年から2018年の糖収量とその構成要素に関して、SVRを利用した早期予測モデルの予測精度を検証した。

根中糖分に関する予測では、予測日が遅くなるほど評価指標が向上する傾向にあった（図2）。特に8月から9月にかけて評価指標は大きく向上した。複数の先行文献で議論されているように、根中糖分は10月上旬といった比較的遅い時期まで気象の影響を受ける<sup>2, 3, 5, 6)</sup>ため、予報日を遅らせることがモデルに与える気象予報データを実際の値に近づけた結果として精度が向上したと考える。

一方で、根収量に関する予測では、予測日を遅らせることでRMSEはわずかに低下したものの、相関係数はむしろ低下する傾向にあった。根収量は比較的早期の4月から8月ごろの気象との相関が高い<sup>2-6)</sup>ことから、予測日を遅れさせることが予測精度の向上に寄与しなかった、と考えることができる。根収量と根中糖分の積として算出される糖収量の予測精度は、8月21日の予測までは向上する傾向が認められたが、それ以降大きく向上することはなかった。より正確な早期予測のためには、収量予測モデルの改善および気象予報の精度の改善のいずれもが求められる。

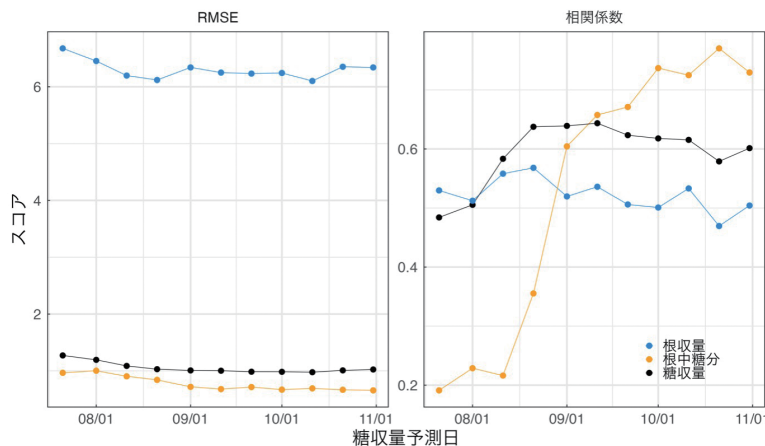
以上のように、機械学習フレームワークを利用した省力的かつ非属人的なモデル作成は、既存の経験的なモデル作成を置換するアプローチになり得るといえる。ただし、PLSモデル・SVRモデルによる予測値は従来のモデルと比較すれば統計値に近い値となったものの、予測値が統計値と大きく乖離<sup>かいり</sup>する年次・地域が多数あった。統計糖収量がモデル糖収量と比較して小さい年次に着目したところ、それらの年は褐斑病、根腐病および黒根病などの病害による被害が多発した年に該当した。このため、病害発生に関するデータを利用した病害による減収程度の補正に取り組んだ。

### 3. てん菜褐斑病による減収の評価と補正

#### (1) 減収の評価

前項で扱ったモデルは「てん菜の収量性は気象のみによって決定づけられる」というナイーブな設計のモデルであった。しかし、実際には病害虫による被害であったり、圃場の排水性の低さに起因する湿害であったり、多様な要因がてん菜の収量性を決定付ける。てん菜生産では、特に病害により顕著な減収が生じたと判断される事例が多数ある。試験研究機関の圃場で10月までの発病度に応じて減収率を

図2 てん菜糖収量および収量構成要素の早期予測のパフォーマンス



注：2011年から2018年の振興局レベルの統計値と気象実況・予報値を使用した解析の結果を示す。



調査した報告<sup>15</sup>など)が複数あるが、予測という観点での利用は難しく、実験レベルの制御がなされていない一般の生産者圃場での減収に应用可能かは明らかではない。

そこで、簡易に病害による減収程度を見積もるために、まずは前項の振興局別の解析で得られたモデル糖収量と統計糖収量から(モデル糖収量 - 統計糖収量) / モデル糖収量として糖収量のギャップを計算し、農作物有害動植物発生予察事業年報<sup>16</sup>)に記載された全道での褐斑病・根腐病の発生量(少、やや少、並、やや多、多の5段階)ごとに整理した(図3)。なお、農作物有害動植物発生予察事業年報では根腐病は「黒根病を含む」として集計されており、両者は区別されていないため本稿では単に根腐病と表記している。

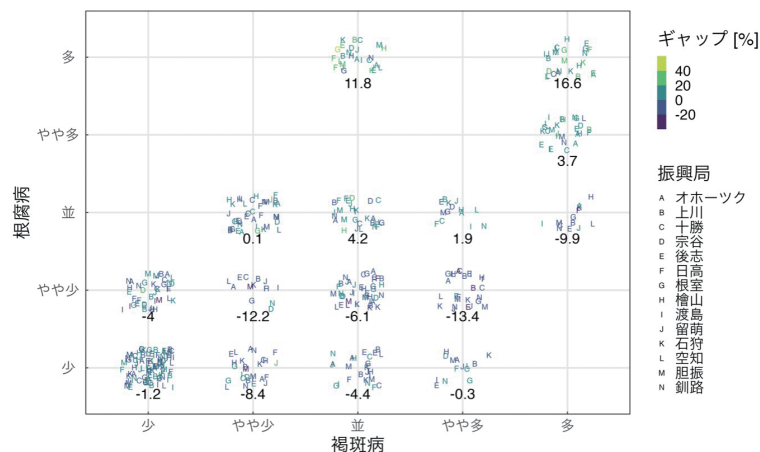
結果は、両病害が同年に多発する場合がある、特に根腐病は湿害と同時に発生する場合が多い、といった理由から病害ごとの寄与は明確ではないが、糖収量のギャップは病害発生量が多いほど大きい傾向にあり、褐斑病・根腐病の発生がともに「多」となった2000年と2010年には中央値で16.6%と推定された。すなわち、気象データのみに基づいたナイーブな収量予測では、病害の発生に起因する減収

程度が十分に考慮されないため、糖収量を10%以上過大に評価する可能性がある。そこで次に、北海道農業改良普及センターによる一般圃場巡回調査として記録された褐斑病発病株率データを利用し、糖収量を補正する手法を検討した。

## (2) 糖収量の補正

てん菜に関する一般圃場巡回調査では、生産者の圃場を巡回して調査集計した7月第3半旬、7月第6半旬、8月第3半旬、8月第6半旬における褐斑病発病株率が蓄積されている。この褐斑病発病株率の2001年から2020年のデータを市町村別に整理し、解析に供試した。前項の市町村レベルの収量予測SVRモデル作成のプロトコルを一部変更し、leave-one-year-out交差検証を行う際に、学習用データのうち褐斑病が多発した市町村・年次のデータを除くことにより、新たに作成した回帰モデルを用いて検証用データから予測収量を算出した。この値は「もし仮に褐斑病が多発していなかった場合に与えられた気象のもとで得られる仮想的な糖収量」を反映したものと考えられ得る。褐斑病が多発したとみなす基準を、巡回調査の最終時点(8月第6半旬)における発病株率が15%以上とした。褐斑病

図3 全道レベルで集計された褐斑病および根腐病の発生量がSVRモデル糖収量と統計糖収量のギャップに及ぼす影響



注1：ギャップを(モデル糖収量 - 統計糖収量) / モデル糖収量として計算した。  
 プロットはある年の1振興局のデータを表し、図内の数値はギャップの中央値を表す。  
 注2：1986年から2020年のデータを解析に使用。

による減収 = 仮想的な糖収量 - 統計糖収量、と定義し、4時点における発病株率および旬別気象と減収の間の相関を解析した。

具体的には得られた相関係数と、褐斑病は特に最低気温が高く多湿な気象条件のもとで発生しやすい、9月の発病度の進展が生産性に強く影響を及ぼす<sup>15</sup>など、という知見から、褐斑病による減収を8月中旬から9月上旬までの日最低気温・日降水量の平均値、8月第6半旬の発病株率により回帰する式を作成し、多発条件に該当する市町村・年次の糖収量を補正した(図4)。

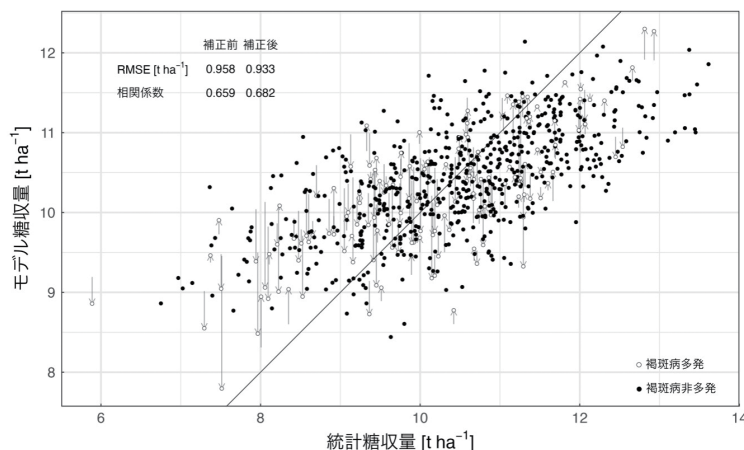
結果は補正後の糖収量は無補正の場合と比較して統計糖収量に近い値となった。同様のアプローチにより病害による減収を考慮することは、一定の収量予測精度の向上につながることを示唆された。ただし、この例ではデータの制約のため、モデルの作成と検証を同一のデータセットで実施しており、いわゆる過学習が起きていることに留意する必要がある。また、病虫害の発生は褐斑病に限らず、気象条件が同一とみなせる近隣圃場間あるいは圃場一筆内でも局所性を持って発生することから、気象データに基づく予測は不確実性が高いことが推察される。病虫害被害による減収リスクを議論する場合には確率的な視点が必要となる。

## おわりに

本研究では、気象データに基づく収量予測モデルを半自動的・非属人的に作成することを目標の一つとし、機械学習フレームワークを利用することで、従来の方法と比較して省力的かつ高精度な気象に基づく収量予測モデルを構築する例を示した。また、普及センターによる一般圃場の巡回調査により得られた褐斑病の発生状況を説明変数の一つとしてモデルを作成する方法を提示した。気候変動に伴う気温上昇により植物病害の発生地域が高緯度へ拡大し、日本を含む地域では病害が多様化する可能性が指摘されており<sup>17</sup>、病虫害リスクの軽減は今後より重要となることが予想される。適切な防除と被害の把握のためにも、気象データに基づいた病虫害発生の予察が有効であると考えられる。

特に予測に主眼を置いたモデルに関する研究では、その精度が興味を中心となり、相関係数やRMSEのみに議論が集中するケースが多い。しかし、これらの評価指標は空間スケールに応じて変化するため、対象を明確に定めずこれらの値だけで議論することは有意義ではないだろう。また、本稿の褐斑病による減収の予測を含め、モデルの作成に利用したデータセットとモデル検証データセットを同一

図4 褐斑病による減収率を補正した市町村レベルのモデル糖収量と統計糖収量



注1: 褐斑病多発の基準を満たした場合に矢印で表される減収率補正をかけた値を表す。  
注2: 2001年から2020年のデータを解析に使用。

としたが、予測モデルの作成という観点では評価方法が妥当でない例が散見される。加えて、変数間の相関が強い気象データを説明変数とした重回帰モデルでは、推定結果が不安定化する多重共線性の問題があることが繰り返し指摘されている<sup>18</sup>など。以上のように、予測のための研究では落とし穴が多数あるため、データのオープン化と第三者による検証を活発化することが必要であると考えられる。

本研究を実施するにあたり北海道病害虫防除所の橋邦宏氏には、データの収集と整理にご対応いただいた。この場を借りて御礼申し上げます。今後データ活用型の解析では、特に公的機関により体系的に実施され、均一性が担保された調査データの重要性がより高まる。一方で過去の調査データはすでに一部が保存されていない、フォーマットが統一されていないなど、データの価値を十分に発揮できていないものが多数ある。このようなレガシーデータの効果的な利活用は今後の農業研究の課題となることが予

想される。

モデルの精度は対象とする空間スケールに依存する。対象となる空間スケールが大きい場合、気象データを利用することで比較的精度の高い推定が可能となる一方で、高解像度なモデルで正確な推定を行うためには、低解像度データでは平均化されて顕在化していなかった土壌、施肥管理、<sup>はしゅ</sup>播種・移植・収穫日といった変動要因が植物生育に及ぼす影響を捉える必要がある。本稿では考慮できていない多数の要因の中でも、<sup>ちよくはん</sup>直播栽培と移植栽培の収量性差、品種ごとの病害抵抗性差および気象応答特性差は、収量予測を行う上で特に重要な要因だろう。要因の数に応じて必要とされるモデルの複雑性は指数関数的に増加し、モデル開発に必要なデータセットも相応に肥大化することに留意すべきである。レガシーデータの利用と並行して、人工衛星および無人航空機（ドローン）を活用したリモートセンシングデータによる植生指数の活用<sup>19</sup>なども検討が必要である。

#### 【引用文献】

- 1) 北海道農業協同組合中央会 (2022) 「てん菜・てん菜糖をめぐる情勢」 (<https://hokkaido-nosan.or.jp/manager/wp-content/uploads/kousyuukai-tensai-2.pdf>)
- 2) 富山信夫 (1980) 「収量と糖分の予測法に関する考察」 『てん菜研究会報』 21, pp.189-195. 甘味資源振興会
- 3) 松崎康範、吉田俊幸、三分一敏 (1983) 「気象要因とテンサイの生育および収量」 『てん菜研究会報』 25, pp.63-68. 甘味資源振興会
- 4) 中里秀昭、今木一喜 (1998) 「テンサイ生育期における気象条件と収量との関係について」 『てん菜研究会報』 39, pp.101-107. 甘味資源振興会
- 5) 梶山努 (2011) 「てんさい」 『北海道立総合研究機構農業試験場資料』 39, pp.32-39. 地方独立行政法人北海道立総合研究機構農業研究本部中央農業試験場
- 6) 柳沢朗 (2018) 「十勝地方における畑作物の生産性向上と収量変動要因について」 『砂糖類・でん粉情報』 (2018年4月号), pp. 46-54. 独立行政法人農畜産業振興機構
- 7) Sameshima R, Hirota T, Hamasaki T, Suzuki S (2007) 「Temperature trends at the national agricultural research center for Hokkaido region in the 40 years from 1966 to 2005」 『J. Agric. Meteorol.』 63, pp.95-102.
- 8) 北海道てん菜協会 (1987-2021) 『てん菜糖業年鑑』 北海道てん菜協会、北海道農産協会
- 9) 大野宏之、佐々木華織、大原源二、中園江 (2016) 「実況値と数値予報、平年値を組み合わせたメッシュ気温・降水量データの作成」 『生物と気象』 16, pp.71-79. 日本農業気象学会
- 10) R Core Team (2019) 「R: A language and environment for statistical computing. R Foundation for Statistical Computing」 Vienna, Austria. (<https://www.R-project.org/>)
- 11) Max Kuhn (2021) 「caret: classification and regression training」 (<https://CRAN.R-project.org/package=caret>.)
- 12) Liland KH, Mevik BH, Wehrens R (2021). 「pls: partial least squares and principal component regression」 (<https://CRAN.R-project.org/package=pls>.)
- 13) Karatzoglou A, Smola A, Hornik K, Zeileis A (2004) 「kernlab - an S4 package for kernel methods in R」 『J. Stat. Softw.』 11, pp.1-20.
- 14) Murakami K, Shimoda S, Kominami Y, Nemoto M, Inoue S (2021) 「Prediction of municipality-level winter wheat yield based on meteorological data using machine learning in Hokkaido, Japan」 『PLOS ONE』 16, e0258677.
- 15) 池谷聡、池谷美奈子 (2015) 「テンサイ褐斑病の被害解析の試み」 『てん菜研究会報』 56, pp. 45-47. 甘味資源振興会
- 16) 北海道農政部、北海道病害虫防除所、北海道立総合研究機構農業研究本部 (1987-2021) 『農作物有害動物発生予察事業年報』 北海道病害虫防除所・地方独立行政法人北海道総合研究機構中央農業試験場
- 17) Chaloner TM, Gurr SJ, Bebber DP (2021) 「Plant pathogen infection risk tracks global crop yields under climate change」 『Nat. Clim. Change』 11, pp.710-715.
- 18) Katz RW (1977) 「Assessing the impact of climatic change on food production」 『Clim. Change』 1, pp.85-96.
- 19) 本郷千春、丹羽勝久 (2013) 「農業空間情報を用いたてん菜の収量予測」 『精密工学会誌』 79, pp.991-994. 公益社団法人精密工学会