

SW-SVR を用いた施設園芸環境向け知的制御システムの評価

兼田千雅^{†1} 井林宏文^{†2} 鈴木雄也^{†2}
黒田正博^{†3} 大石直記^{†4} 峰野博史^{†2}

農業分野では、センサネットワークを用いて農作物の栽培環境や生育状況を収集し、その時の状況に応じた制御を自動的に行うシステムの開発が進められている。しかし、既存システムでは、WSNを用いた場合に通信の信頼性が低下し、また農業従事者の経験や勘による制御を適切に再現しきれていないという課題があった。そこで本論文では、SW-SVRを用いた施設園芸環境向け知的制御システムを提案する。提案システムは、現在の環境に特化したモデル化を行うSW-SVRと、施設園芸環境での無線通信に適した無線方式を用いることで、農業従事者の経験や勘による制御を高精度かつ高速に再現する。現場実証実験の結果、無線通信部ではPERO.00%を実現し、データ予測部では窒素吸収量の予測誤差を既存SVRと比べ68%削減することを確認した。

Evaluation of agricultural environmental control system using SW-SVR

Yukimasa Kaneda^{†1} Hirohumi Ibayashi^{†2} Yuya Suzuki^{†2}
Masahiro Kuroda^{†3} Naoki Oishi^{†4} Hiroshi Mineno^{†2}

In the agriculture, systems to collect the environmental information using sensor networks and to control depending on the situation have been actively developed. However, existing systems lack the reliability of wireless sensor network, and it is difficult to reproduce the control based on the experience and the intuition of farmers. In this paper, we propose agricultural environmental control system using SW-SVR. This system can do farm work like farmers without relying on farmers. This system uses 429MHz band wireless communication that has high diffractive wave and SW-SVR that can predict future environment with high accuracy and low calculation cost. By applying 429MHz band wireless communication and SW-SVR, it also confirmed that this system build the highly reliable WSN and reduced nitrogen absorbed amount prediction error by 68% compared with SVR.

1. はじめに

センサ技術の進歩により様々な環境データを収集可能となり、また収集した環境データを利用して各種機器を制御するシステムも盛んに実用化されつつある。例えば農業分野では、センサネットワークを用いて農作物の栽培環境や生育状況を収集し、その時の状況に応じた制御を自動的に行うシステムの開発が進められている。しかし、センサネットワークを用いた農業環境向け機器制御システムの本格的な実用化のためには、解決すべき二つの課題が存在すると考える。

一つ目の課題は、無線センサネットワークを用いて数カ月の栽培期間中継続して稼働する機器制御システムの実現が困難なことである[1]。無線センサを用いた場合、有線センサと異なりセンサの設置が容易となるが、無線通信によってデータ通信の信頼性が低下するため、障害物が多数存在する施設園芸環境下で安定した無線通信を実現することは困難であった。

二つ目の課題は、農業従事者の経験や勘による制御を適切に再現しきれていないことである。従来の機器制御システムでは、得られたセンサデータとユーザ入力の設定値を比較し、比較結

果から制御必要性の有無を決定する方法が主であった[2][3]。しかし従来の制御方法では、植物の生育ステージや状態から適切に判断された制御を実現することは困難であると考える。例えば、トマト栽培の養分供給では、当日トマトが吸収する養分の量を判断し、過不足の無い量の養分をあらかじめ供給することで、高品質な作物の栽培だけでなく肥料の節約や病気の抑制を実現できるが、養分吸収量の判断は農業従事者の経験や勘が必要であり、機器制御システムでの判断の再現は困難であった。

本研究では、上記の課題を解決するため、施設園芸環境に適した無線方式を用いて予測に必要な環境データを確実に収集し、SW-SVR (Sliding Window-based Support Vector Regression)[4]を用いて農業従事者の経験や勘による制御の再現を可能とする施設園芸環境向け知的制御システムを提案する。SW-SVRは、現在の環境に特化した予測モデルの構築と、特性変化に応じたモデル再構築を自動的に繰り返す機械学習アルゴリズムの一つであり、詳細は3.4節で述べる。

以降、2章で関連研究について述べ、3章でSW-SVRを用いた施設園芸環境向け知的制御システムの提案を行う。4章でプロトタイプの実装について、5章で評価実験結果を述べる。6章で本論文をまとめる。

2. 関連研究

施設園芸環境の様々な環境データをセンシングし、制御機器を自動的に制御するシステムが普及しつつある。特に配線コストが不要となる無線センサネットワーク (WSN) を利用したセンシングの研究[5]-[7]が注目されている。ただし施設園芸環境

^{†1} 静岡大学情報学部

Faculty of Informatics Shizuoka University

^{†2} 静岡大学大学院情報学研究科

Graduation school of informatics Shizuoka University

^{†3} 情報通信研究機構

National Institute of Information and Communications Technology

^{†4} 静岡県農林技術研究所

Shizuoka Prefectural Research Institute of Agriculture and Forestry

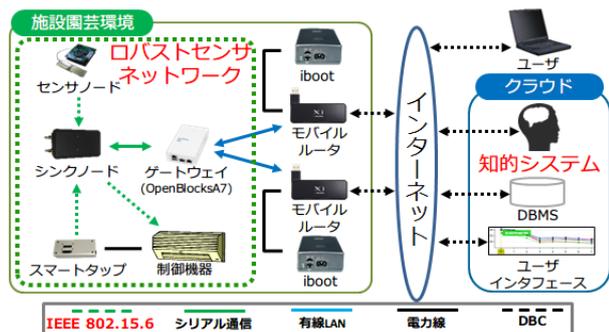


図1 システムアーキテクチャ

の WSN で広く使われている 2.4GHz 帯無線通信は直進性が強いので、金属パイプや植物等の障害物が多数存在する施設園芸環境での利用は適さない場合もある。2.4GHz 帯以外の周波数帯を用いた施設園芸環境向け WSN として、433MHz 帯を用いたじゃがいも畑でのセンシング[8]が挙げられる。433MHz 帯等の低周波数帯無線通信は、2.4GHz 帯無線通信と比べ波長が長く回折性に優れるため、障害物が多数存在する施設園芸環境下でも高信頼な WSN を構築可能であると考えられる。

一方、高品質な作物の栽培のためには、農業従事者の経験や勘で行われている制御を適切に再現し、植物に対して過不足の無い制御をリアルタイムに実施することが重要となる。そのため機器制御システムは、農業従事者が暗黙的に実施している将来予測を再現し、現在必要とされる制御を決定すべきであるが、農業従事者が予測に用いているデータは時間経過と共に特性が変化する時系列データであり、一般に予測が困難とされている。時系列データに対して予測を行う関連研究として、Adaboost[9]を応用した Boosting-SVM with Asymmetric Cost algorithm が挙げられる[10]。この研究では、学習データの重みを逐次的に変化させることで、高精度な予測モデルを構築する。実験の結果、多様なオープンデータで高精度な予測が可能であることが示された。またその他の関連研究として、125万件もの大規模な学習データを用いて高精度な ANN (Artificial Neural Network) 予測モデルを構築し評価を行った研究が挙げられる[11]。実験の結果、Mean absolute error (MAE) の評価値で1時間後の気温を 0.525度の誤差で予測することに成功した。しかし、これらの関連研究は学習データに特殊な前処理を施したり、学習データ量を増大させたりすることで予測精度を向上させることに着目しており、予測モデルの構築時間に関しては考慮されていない。予測を利用したアプリケーションを想定した場合、学習データの量に関わらず常に高精度な予測と高速なモデル構築を両立できることが望ましい。

3. SW-SVR を用いた施設園芸環境向け知的制御システム

3.1. システムアーキテクチャ

本研究では、施設園芸環境で農業従事者の経験や勘で行われている制御を適切に再現するため、多種多様なデータを確実に収集可能とする WSN と、時系列データ予測に特化した SW-SVR を用いた施設園芸環境向け知的制御システムを提案する。特に SW-SVR は、現在の環境に特化した予測モデルの構築と、特性変化に応じたモデル再構築を自動的に繰り返すことで高精度かつ高速な予測を維持し続ける。

SW-SVR を用いた施設園芸環境向け知的制御システムのアーキテクチャを図1に示す。提案システムは大別して、施設園

芸環境下で環境データのセンシングと制御信号の送信を行うロボストセンサネットワーク(図1左部)と、クラウド上で将来の環境の予測と制御信号の生成を行う知的システム(図1右部)から構成される。知的システムや DBMS は施設園芸環境内に配置することもできるが、クラウド上に配置することで高温多湿な現場環境へ設置する電子機器類の負荷を軽減し、またクラウド型サービスとして統合的な予測分析を可能とする。

提案システムの動作手順を示す。ロボストセンサネットワークで環境データを収集し、収集したデータをクラウド上 DBMS に蓄積する。蓄積されたデータを基に知的システムが SW-SVR で将来状況を予測し、予測結果に基づく制御信号を生成する。知的システムで生成された制御信号は、ロボストセンサネットワークを介して配送され、現場の制御機器を制御する。

3.2. 要求事項

提案システムは、システムの利用者が限定され、また使用するセンサノードや制御機器の増加など拡張性を必要とし、さらに栽培期間中の確実なシステム稼働を保証する必要があることから「社会的影響が限定されるシステム」[12]に該当すると考える。社会的影響が限定されるシステムでは、稼働率 99.99% 以上が必要とされることから、提案システムの目標稼働率も 99.99% 以上と設定する。特に稼働率 99.99% の実現のためには、無線通信部においてパケットエラーレート(PER)が 0.01% 以下である必要がある。

次に提案システムに要求される制御の妥当性とリアルタイム性について検討する。植物に対する制御の場合、満たすべき制御の妥当性とリアルタイム性は制御対象や実験環境で異なるため、これらを決定する要因である予測アルゴリズムの予測精度とモデル構築時間の具体的目標値は、システム利用シチュエーション毎に設定するものとする。

3.3. ロバストセンサネットワーク

3.3.1. 概要

ロボストセンサネットワークは、データの収集と制御信号の送信を無線通信で行う。無線通信を用いることで、センサノードや制御機器を任意の位置に設置することが容易となり、システム導入や運用の負担軽減に役立つ。

施設園芸環境には植物の葉や金属パイプ等の障害物が多数存在し、これら障害物は無線通信品質に悪影響を及ぼす。そこで提案システムでは、無線周波数帯を低くし波長を長くすることで障害物への回折性が高まりパスロスを抑えられることに着目し、429MHz 帯 IEEE802.15.6 方式[13]を採用する。

3.3.2. 429MHz 帯 IEEE802.15.6 方式

ISM 帯として多用途に用いられている 2.4GHz 帯無線通信と比べ、429MHz 帯無線通信は障害物を回折しやすいという特徴がある。波長が約 699mm である 2.4GHz 帯無線通信に比べ、波長が約 125mm である 429MHz 帯無線通信は、小さな障害物が存在しても迂回可能だけでなく、2.4GHz 無線通信と比べ同一出力で遠距離への伝送が可能である。無線周波数帯を低くすることで伝送ビットレートは低下し、また提案システムで採用した 429MHz 帯無線モジュールのペイロードサイズは 16 バイトと小さいが、提案システムで送受信されるセンサデータや制御命令の送受信に対しては遅延のない無線通信が可能である。

IEEE802.15.6 方式では、伝送スロット確保(TDMA)や再送処理によってパケット欠落を抑制し、また自動暗号キー生成および AES 暗号によって盗聴を防止する。本方式を施設園芸環境下の WSN に用いることで、センサノードが収集した環境データや改竄を防ぐべき制御信号を、無線通信で確実かつ高いセキュリティ性を保って送受信することが可能となる。



図 2 SW-SVR 概略図

3.4. 知的システム

3.4.1. 概要

知的システムは、ロバストセンサネットワークが収集した環境データを基に制御信号を生成し、生成した制御信号を制御機器に送信する。機械学習を用いて農業従事者の暗黙知を形式知化することで、機器制御システムで従来行われていなかった農業従事者の経験や勘による制御を自動的に実行可能とする。

予測アルゴリズムには SW-SVR を採用することで、農業従事者の予測を適切に再現する。SW-SVR は時系列データの特徴に着目することで、予測困難とされる時系列データに対して高精度かつ高速な予測を可能とするものであり、時系列データを扱う実環境向け機器制御システムへの適用に適している。

3.4.2. SW-SVR

3.4.2.1. 概要

SW-SVR は、時系列データ予測に特化させた SVR(Support Vector Regression)の改良アルゴリズムである。SVR は SVM(Support Vector Machine)を回帰分析に応用したものであり、多様な分野で優れた認識性能を示すことが知られている学習器のひとつである。SVM は、カーネル関数で学習データを高次元空間に射影することで線形分離可能なデータに変換し、その後マージン最大化の基準で識別平面を定めることで、汎化性能の高い学習器を構築する。一方で SVR の計算量は学習データ量 n に対し $O(n^2)$ から $O(n^3)$ であるため[14]、膨大な情報量が必要となる自然環境の高精度なモデル化に SVR は適しておらず、リアルタイム性が必要とされる機器制御システムへの適用は困難だった。そこで SW-SVR では、全期間の特性変化を全て考慮したモデル化を行うのではなく、現在の環境に特化した予測モデルを常時維持し続ける。現時点の予測に必要な最低限の学習データから予測モデルを構築することで、学習データ量の削減と予測精度の向上が期待できる。

SW-SVR の概略図を図 2 に示す。SW-SVR は、学習データ抽出(図 2(1))と自動再構築(図 2(2))の二つの特徴を持つ。学習データ抽出では、時系列データが持つデータの規則性に着目することで、最新の観測データに関連した最低限のデータだけを抽出する。ただし抽出された学習データから構築される予測モデルは、その時点の環境に特化された予測モデルであるため、予測モデル構築時の環境から特性が変化するとつれて予測精度が悪化する。そのため自動再構築では、予測モデルの有効性を常時判定し、時間経過に伴う環境の特性変化に予測モデルが対応不能と判定した場合に、再びその時点の環境に特化した予測モデルを構築する。

SW-SVR のフローチャートを図 3 に示す。(i)データ入力部では、母集団として一定量の学習データを入力する。(ii)データ選択部では、入力された学習データから現時点の予測に必要な最低限の学習データだけを抽出する。ここで抽出した学習データを用いて、(iii)モデル構築部で現在の環境に特化した予測モデルを構築する。(iv)データ予測・予測モデル再構築判定部では、モデル構築部で構築した予測モデルを用いて予測を行い、同時に現在の予測モデルの予測結果と観測データを比較すること

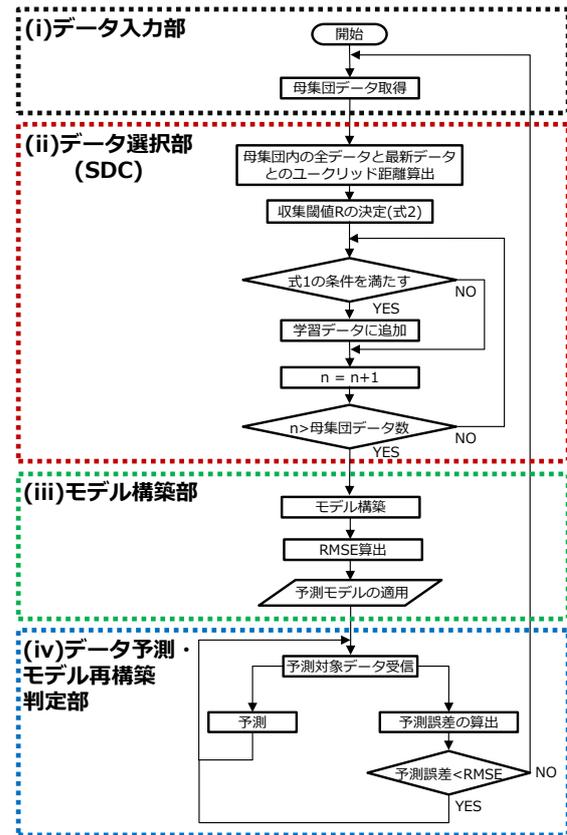


図 3 SW-SVR フローチャート

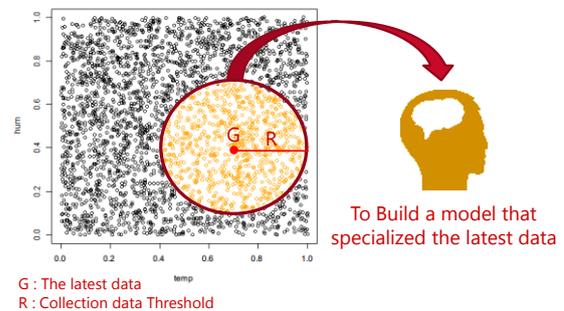


図 4 SDC 概略図

で予測精度を算出する。予測精度が閾値よりも悪化した時点で予測モデルの再構築を行う。

SW-SVR の計算量について述べる。一回のモデル構築の計算量は、データ選択部での学習データ抽出アルゴリズムの計算量と、モデル構築部での SVR の計算量から決定される。学習データ抽出アルゴリズムには後述する SDC を採用しており、SDC の計算量は $O(n)$ であるため[15]、SW-SVR の一回のモデル構築の計算量は SVR と同等の $O(n^2)$ から $O(n^3)$ となる。

3.4.2.2. 学習データ抽出

SW-SVR の(ii)データ選択部では、最新の観測データ G との類似度が収集閾値 R 以内の学習データのみをモデル構築に使用する学習データとして抽出する近傍データ収集法(Short-distance Data Collection : SDC)を用いる(図 4)[15]。SDC は、時系列データの遷移特性には規則性が存在し、ある環境とそれに類似した過去の環境は遷移特性も類似することに注目しており、最新の観測データと類似した学習データだけを用いること

で、最新の環境に特化した予測モデルを構築可能と考える。

SDC の抽出条件を式(1)に、収集閾値 R の導出式を式(2)に示す。ここで、最新の観測データ G の各成分を (g_1, g_2, \dots, g_D) 、 n 番目の学習データ P_n の各成分を $(p_{n1}, p_{n2}, \dots, p_{nD})$ 、説明変数の次元数を D 、母集団のデータ数を N とする。式(1)と式(2)では、類似度を表す距離の定義として広く用いられているユークリッド距離を採用する。ただし各次元のスケールの違いを除去するためデータの標準化を実施する。SDC は、式(1)の条件を満たす P_n を探索し、条件を満たした P_n のみを学習データとして採用する。収集閾値 R の導出式である式(2)は、母集団の全学習データと最新の観測データとの距離の標準偏差を求める式である。一般に自然環境の多くは正規分布に従うことが多いとされているため[16]、学習データが正規分布に従うと仮定した場合、最新の観測データとの距離の標準偏差から学習データ量を決定することで、母集団の全学習データから最新の観測データと類似しているデータが約 68%抽出されると考える。

$$R > \sqrt{\frac{D}{\sum_{i=1}^D (g_i - p_{ni})^2}} \quad (1)$$

$$R = \sqrt{\frac{N \sum_{n=1}^N \sum_{i=1}^D (g_i - p_{ni})^2}{N}} \quad (2)$$

3.4.2.3. 自動再構築

時系列データ予測の場合、一定時間経過後に得られる観測データから予測誤差を算出できるため、予測モデルの有効性の指標として予測誤差を用いることができる。そのため SW-SVR では、予測誤差が一定の閾値を上回ったとき、環境の特性変化に予測モデルが対応不能と判断し予測モデルの再構築を行う。

再構築の判定を行う閾値の決定方法について述べる。SW-SVR はモデル構築終了後に、モデル構築に用いた学習データに対して予測を行い、RMSE を算出する。この RMSE は学習データに対する予測誤差の標準的な大きさを示すものであり、学習データと最新の観測データの特性が類似しているならば、最新の観測データに対する予測誤差は RMSE を下回り、一方で学習データと最新の観測データの特性が大きく異なっているならば、最新の観測データに対する予測誤差は RMSE を上回ることが予想される。そのため、RMSE を再構築の判定を行う閾値に設定することで、最新の観測データと学習データの特性の差を識別可能となり、環境の特性変化に伴う自動的な再構築が実現される。

4. プロトタイプ実装

4.1. 散乱光センサを用いた窒素吸収量予測制御システム

SW-SVR を用いた施設園芸環境向け知的制御システムのプロトタイプとして、トマト水耕栽培において、空气中に散乱している光(散乱光)の強さをセンシングする散乱光センサを用いて、制御当日にトマトが吸収する窒素量を予測し、その予測結果に応じた窒素供給制御を行うシステムを開発した。毎日決められた時刻に行われる窒素供給制御に対し、どの程度の量の窒素を供給すべきかを知的システムで判断し、適切な量の窒素を供給し続けることで、高品質な果実の栽培を目指す。

本実験の栽培方式では、給液タンクに接続されているポンプが給液タンク内の溶液を汲み出し、植物が定植されている栽培



図 5 使用センサノード

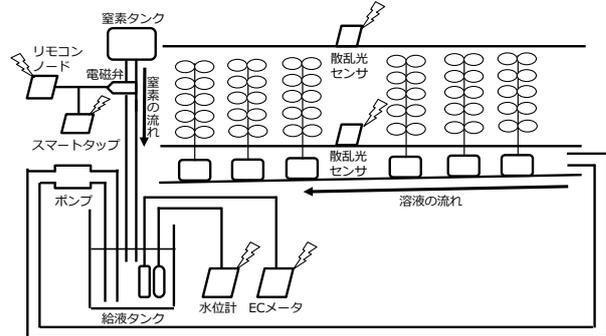


図 6 各種センサ配置図

ベッドへ溶液を流す。栽培ベッドを流下した溶液は再び給液タンク内に格納される。ポンプの稼働時間は 6 時から 18 時とし、稼働タイミングと稼働時間に関しては既存の機器制御システムが自動的に決定する。提案システムは、窒素タンクと給液タンク間に接続されている電磁弁を制御するものであり、予測結果に基づき供給すべき窒素量を決定し、その量を供給するために必要な時間だけ電磁弁を開放する。

プロトタイプ実装における目標予測精度について検討する。本実証実験では、農業従事者の暗黙知の形式化が可能であるかを検証するため、農業従事者が検討する窒素吸収量推定アルゴリズムの予測精度を目標予測精度に設定する。今回目標として使用する窒素吸収量推定アルゴリズムは、[17]で提案された回帰モデルから導出した回帰式とする。この回帰式では一株当たりの窒素吸収量を予測しており、予測精度は MAE で 0.66、RMSE で 0.90 であり、この値を目標予測精度に設定する。またプロトタイプ実装では予測時刻の観測データをセンシングしてから一分以内に予測結果が制御に反映されれば正常な稼働と定める。そのため目標モデル構築時間を一分に設定する。

4.2. ロバストセンサネットワークの実装

プロトタイプ実装で使用したセンサノードを図 5 に示す。各センサノードと 429MHz 帯 IEEE802.15.6 方式を搭載した無線モジュールを接続することで、各センサノードはセンサデータを 429MHz 帯 IEEE802.15.6 方式による無線通信でシンクノードへ送信する。図 5(a)は気温・相対湿度・光量を収集する散乱光センサである。光量センサは図 5(a)の筐体内に取り付けられており、直射日光が筐体内に差し込まないように開口部を北側に向けて設置することで、散乱光の強さをセンシングすることができる。図 5(b)は電気伝導率を収集する EC メータ(東亜ディーケーケー製、CM-31P)であり、図 5(c)は水位を収集する水位セ

ンサ(クリマテック株式会社製, CWG-H10-C20)である. EC メータと水位センサは給液タンク内窒素量を算出するために用いられる. 図 5(d)は制御機器の消費電力を計測するスマートタップであり, 制御確認に用いる. 図 5(e)は制御機器への制御信号を送信する赤外線リモコンノードである. 知的システムで生成された制御信号をシンクノードから受信し, 制御信号を基に適切な赤外線信号を制御機器に対して送信する.

現場環境における各センサノードの配置図を図 6 に示す. 散乱光センサは群落の上下に設置することで, 推定誤差 3%以内で LAI を推定可能であるため[18], 栽培ベッドの上部と下部に設置する. EC メータと水位計は給液タンク内に設置する. スマートタップとリモコンノードは窒素タンクの電磁弁に接続し, リモコンノードで電源を制御し, スマートタップで電源投入時の消費電力を計測する.

4.3. 知的システムの実装

4.3.1. 概要

知的システムのフローチャートを図 7 に示す. 知的システムは(ア)学習データ生成部, (イ)予測部, (ウ)制御部から構成され, 3.4 節で提案した SW-SVR は(イ)予測部で動作する. 本節ではプロトタイプ実装での学習データ生成部, 予測部, 制御部の詳細を述べる.

4.3.2. 学習データ生成部

図 7(1)の学習不可期間について説明する. 時間経過に伴う給液タンク内窒素量の遷移の概観を図 8 に示す. 図 8(1)は植物の窒素吸収だけが行われている期間を示し, 図 8(2)は植物の窒素吸収と窒素滴下制御が行われている期間を示す. 知的システムでモデル化する対象は植物の窒素吸収モデルであり, 図 8(2)の期間は植物の窒素吸収以外の窒素量の変動要因が含まれている. そのため図 8(2)の期間は学習不可期間として学習データを生成せず, 図 8(1)の期間だけ学習データを生成する.

図 7(2)のフィルタリングについて説明する. 本実験で用いるセンサ類の内, 水位センサはポンプ稼働時の一時的な減少やセンサの誤差等が原因となり, 計測された値が不正値の場合がある. そのため水位センサのセンサデータに対しては, フィルタリングを適用する. 適用するフィルタリングは, ポンプ稼働時

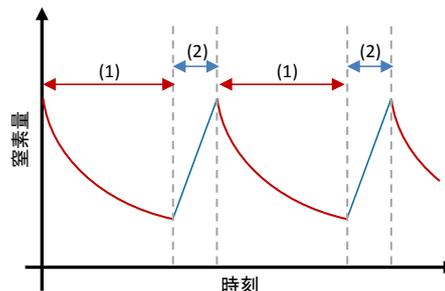


図 8 給液タンク内窒素量の遷移

表 1 学習データ詳細

項目	内容
説明変数	上部散乱光, 下部散乱光, 気温, 相対湿度, 最低相対光量
目的変数	6 時間後の給液タンク内窒素量

の一時的な水位減少を取り除くスパイク除去と, 水位センサ自身の誤差を取り除くローパスフィルタの二種類である.

図 7(3)の窒素量の算出について説明する. 給液タンク内窒素量は給液タンク内窒素濃度と給液タンク内水量の積をとることで算出する. 給液タンク内窒素濃度は電気伝導率と相関があるため, EC メータを用いて算出する. 給液タンク内水量は, 給液タンクが円筒型であるため水位センサを用いて算出する.

図 7(4)で作成される学習データの説明変数と目的変数を表 1 に示す. プロトタイプ実装では, 予測制御を行うために 6 時間後の給液タンク内窒素量を予測する. また説明変数には窒素吸収と関連があるとされる上部散乱光, 下部散乱光, 気温, 相対湿度, 最低相対光量の 5 項目を用いる. 最低相対光量とは一日毎に 10 時から 14 時までの上部散乱光と下部散乱光の比の平均を算出し, その最低値をとったものであり, 群落の繁茂度を示す指標となる. 上記 5 項目を用いることで蒸発散量を推定誤差 3%以内で推定可能となり[18], 蒸発散と窒素吸収は関係があるとされるため, 上記 5 項目を用いることで植物の窒素吸収予測モデルを構築可能と考える.

4.3.3. 予測部

予測部では図 2 に示した SW-SVR が動作する. ただし本実験での予測制御のため, 図 2 のデータ予測・モデル再構築判定部の仕様を一部変更した. 本来 SW-SVR は観測データが入力されるたびに予測と再構築判定を行うものであるが, プロトタイプ実装では制御時刻である朝 6 時のみ予測結果が必要とされるため, 予測は朝 6 時に入力される観測データに対してだけ行い, 再構築判定はその 6 時間後である 12 時に行う. また本実験では予測時刻である朝 6 時の予測を高精度に実施する必要があるため, 12 時に行われた再構築判定の結果再構築すべきと判断された場合, 翌日の朝 6 時の観測データに対して SDC を適用させ再構築を行う.

4.3.4. 制御部

制御部では予測結果を基に制御信号を生成する. 本実験では, 予測結果である制御時間帯の窒素吸収量に比例した量の窒素を給液タンク内に滴下する. 制御時間帯の窒素吸収量に一定の比例定数を掛け合わせる理由は, 制御時間帯の窒素吸収量と同量の窒素を供給しても, 必ずしも供給された窒素が全て吸収されるとは限らないからである. 窒素吸収量に対して何倍の窒素量を供給するのが適切かは, 本実験を通して農業従事者が明らかにすることとし, プロトタイプ実装での比例定数は農業従事者が値を変更させながら判断した.

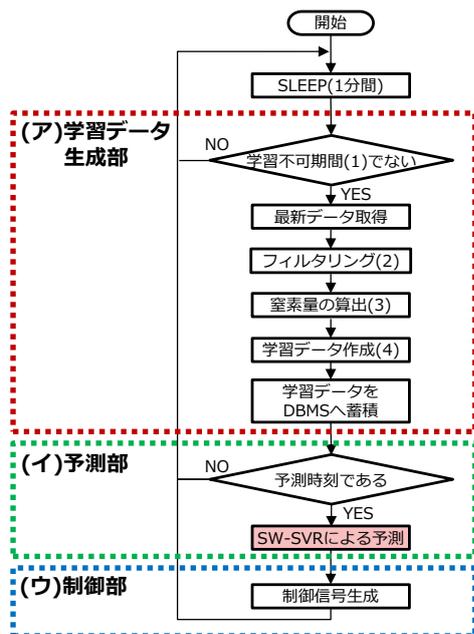


図 7 知的システムフローチャート

5. 評価実験

5.1. ロバストセンサネットワーク

ロバストセンサネットワークの評価実験では、施設園芸環境下に 2.4GHz 帯 IEEE802.15.4 方式の無線センサノードと 429MHz 帯 IEEE802.15.6 方式の無線センサノードを設置し、植物の繁茂度に対する各無線方式の PER の関係の評価した。本論文では知的システムに焦点を絞っているため、ロバストセンサネットワークの実験の詳細や考察は省略する。

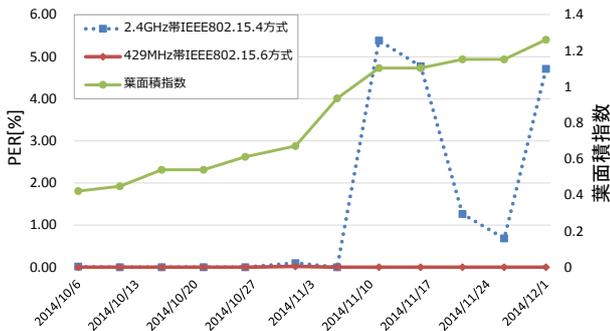


図9 植物の繁茂度と各無線方式の PER の関係

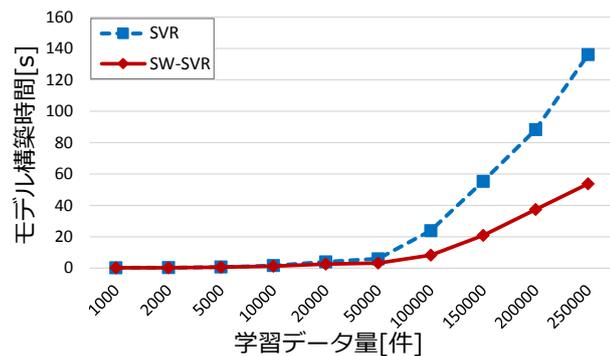


図10 データ量とモデル構築時間(基礎実験)

表2 学習期間と評価期間

学習期間	評価期間
期間：2014/10/31~2014/11/25	期間：2014/11/26~2014/12/7
日時：6時~18時	日時：6時
データ数：2000件	データ数：10件

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (3)$$

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N} \quad (4)$$

表3 SW-SVR 各種パラメータ設定値

項目	設定値
Method	Epsilon - SVR
カーネル関数	RBF
Cost Parameter (C)	1.0
Hyper parameter of RBF (γ)	0.2
Tube parameter	0.1

実験結果を図9に示す。葉面積指数は値が大きいほど植物の繁茂度が高いことを示す。実験結果から、2.4GHz 帯無線通信は植物の繁茂度と PER に関係が見られるが、一方で 429MHz 帯無線通信は植物の繁茂度に関わらず PER が増加していないことが確認できる。以上より 429MHz 帯 IEEE802.15.6 方式は施設園芸環境の WSN に適していると考えられる。

5.2. 知的システム

5.2.1. 実験内容

知的システムの評価実験では、既存手法 SVR と提案手法 SW-SVR の双方で制御時間帯の窒素吸収量予測モデルを構築し予測精度を評価する。モデル構築時間に関しては、AMeDAS の気象データ 25 万件を用いた基礎実験で、SW-SVR は一分以内に予測モデルを構築することが確認できているため(図10)[15]、本論文では現場環境での予測精度の評価を中心にまとめる。

5.2.2. 実験方法

学習期間と評価期間を表2に示す。ただしシステム更新等で学習期間と評価期間には一部データが欠損している期間が存在するため、学習期間のデータは 2000 件、評価期間のデータ数は 10 件となる。予測誤差の指標には RMSE(Root Mean Square Error: 式(3))と MAE(Mean Absolute Error: 式(4))を用いる。N は評価期間のデータ数、 y_i は i 番目の真値、 \hat{y}_i は i 番目の予測値を示す。

SVR の各種パラメータ設定値を表3に示す。各パラメータは R 言語内の SVM 用パッケージ e1071[19]で SVR を使用する際に必要となるパラメータであり、今回はパラメータチューニングを行わずデフォルト値を用いた。

5.2.3. 実験結果・考察

評価期間内の真値と予測値を図11に、評価期間全体での予測アルゴリズム毎の予測精度を表4に示す。ただし評価実験の予測対象は栽培ベッドに定植されている全ての植物の窒素吸収量の合計値であるため、表4の括弧内に一株当たりの窒素吸収量を予測したときの予測誤差を示す。図11から、SW-SVR は評価項目 10 項目の内 9 項目で SVR より低い予測誤差を示し、また表4から評価期間全体での予測誤差を MAE では 77%、RMSE では 68%削減することを確認した。さらに SW-SVR の予測誤差と SVR の予測誤差と目標予測誤差を比較したところ、SW-SVR を用いることで目標予測誤差よりも低い予測誤差で予測できることが確認できた。

SW-SVR が SVR と比べ予測誤差を削減した理由を考察する。

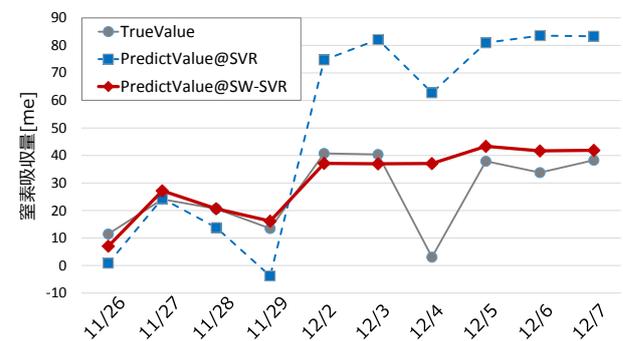


図11 真値と予測値

表4 予測アルゴリズム毎の予測精度

予測アルゴリズム	MAE	RMSE
SW-SVR	6.82 (0.49)	11.50 (0.82)
SVR	29.24 (2.09)	36.44 (2.60)
目標予測誤差	0.66	0.90

図 11 から SW-SVR は SVR と比べ、真値の変動に正確に対応していることが分かる。これは SW-SVR が毎日の環境に特化した予測モデルを繰り返し構築して予測しているのに対し、SVR は学習期間内の全データから構築した予測モデルで予測し続けているためと考える。特に SVR による予測では、2014/12/2 以降で予測誤差が増大していることから、2014/11/26 ~ 2014/11/29 の評価データと学習データには類似性が存在し、2014/12/2 ~ 2014/12/7 の評価データと学習データには類似度に差が生じていると考える。ここで評価データと学習データの類似度を示すため、学習期間のデータと評価期間のデータの特性を図 12 に示す。図 12 は学習期間のデータと評価期間のデータを、主成分分析で 2次元空間に変換したものである。第一主成分と第二主成分の累積寄与率は 93.138%であり、データ特性の概観を表していると言える。図 12 から、学習期間内の学習データは特性がばらついており、多様な環境から構成されていることが確認できる。また 2014/11/26 ~ 2014/11/29 の期間では評価データと学習データ間に類似性が見られるが、2014/12/02 ~ 2014/12/07 の期間では評価データと学習データ間に一定の違

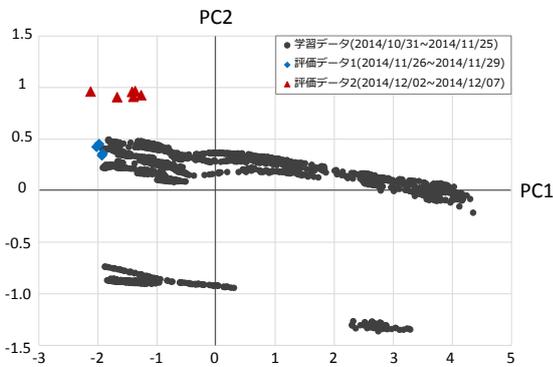


図 12 学習データと評価データの特性(SVR)

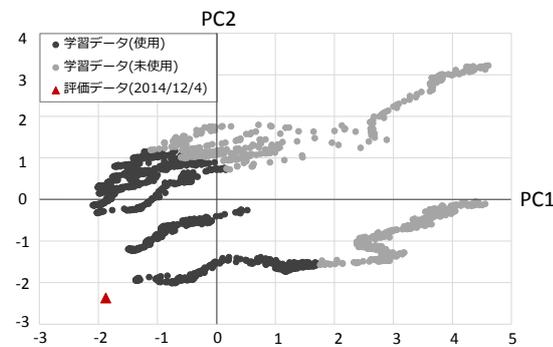


図 13 学習データと評価データの特性(2014/12/4)

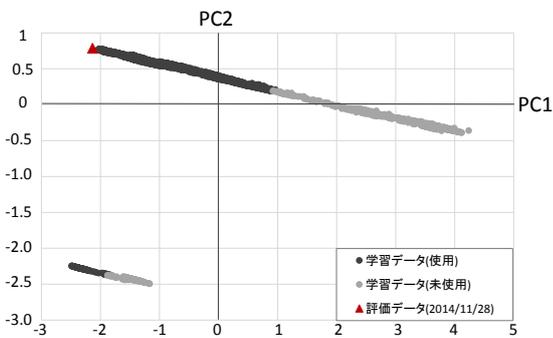


図 14 学習データと評価データの特性(2014/11/28)

いが見られることも確認できる。このことから、SVR は特性が類似していないものも含めた全ての学習データを学習しているため、様々な環境をモデル化することになり、その結果 SW-SVR よりも予測精度が悪化していると考えられる。また特に 2014/12/02 ~ 2014/12/07 の期間では、評価データと学習データ間の違いが大きいので、SVR は予測精度を大幅に悪化させていると考える。以上より、学習データと評価データの特性の差は、予測精度に影響を与えることを示しており、適切なタイミングで学習データと評価データの特性の差を小さくするように再構築を実施する SW-SVR は、時系列に特性変化が行われる農業データに対しても有効であると考えられる。

SW-SVR の予測に関して、予測誤差が高いモデルと予測誤差が低いモデルを比較し、SDC の効果を考察する。SW-SVR が評価期間で最も高い予測誤差を示したのは 2014/12/4 の予測であり、2014/12/4 の観測データと SW-SVR が用いた学習データの特性を図 13 に示す。一方で SW-SVR が評価期間で最も低い予測誤差を示したのは 2014/11/28 の予測であり、2014/11/28 の観測データと SW-SVR が用いた学習データの特性を図 14 に示す。いずれも図 12 と同様に主成分分析を用いてデータの概観を示しており、図 13 の累計寄与率は 91.102%、図 14 の累計寄与率は 94.032%となる。また図 13 と図 14 では SDC の抽出結果も合わせて示している。図 13 と図 14 では学習データの分布が大きく異なっているが、いずれの場合も SDC は全体の学習データの中から、最新の観測データとの類似度が高い学習データを抽出していることが確認できる。以上より、SDC は学習データの分布に関わらず入力された学習データの中から観測データと類似度が高いデータを収集することは可能であるが、SDC で抽出される学習データから予測モデルを構築しても予測精度が向上しない場合があることが確認できた。

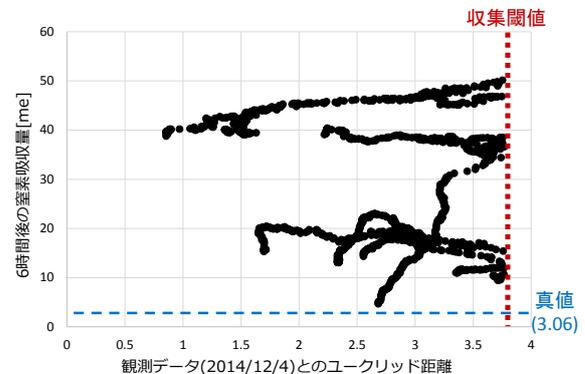


図 15 抽出された学習データの類似度と目的変数の関係

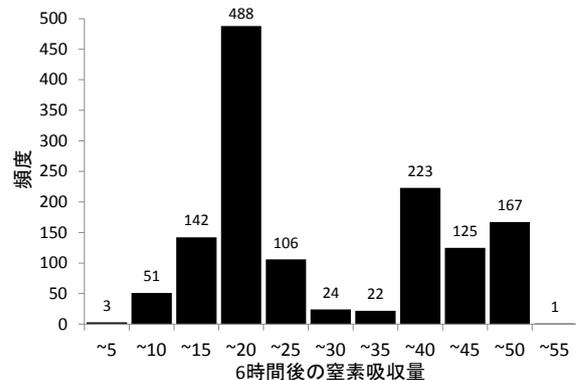


図 16 抽出された学習データの目的変数の分布

2014/12/4 の観測データに対して SDC で抽出された学習データを分析し、SW-SVR を用いても予測精度が悪化した理由を考察する。2014/12/4 の観測データに対して SDC が抽出した学習データの詳細を図 15 に示す。図 15 は各学習データの観測データとのユークリッド距離と目的変数(6 時間後の窒素吸収量)の関係を示したものである。図 15 からは、目的変数が真値付近の学習データの観測データとのユークリッド距離が、約 2.7 であることが確認できる。目的変数が真値付近の学習データは、全体の学習データの中で最も類似度が高い学習データの一つとして判断されるべきであるため、この結果から適切な類似性が判定されていないことが分かる。この原因は説明変数の情報量が少なかったためと考える。SDC は、ある環境とそれに類似した過去の環境は遷移特性も類似するという時系列データの特徴に着目し、観測データの説明変数を基に類似した過去の学習データを収集する。ただし適切な説明変数が与えられない場合、観測データとのユークリッド距離と類似度の関係性が失われ、現在の環境と類似度が低い環境の学習データが収集される可能性が生じる。その結果として現在の環境の遷移特性と異なる遷移特性を学習し、予測誤差増大の原因に繋がると考える。以上より SDC の有効性を最大限にするためには、ある時点での環境を高精度に再現できる情報量を持った適切な説明変数が必要になると考える。

さらに 2014/12/4 の SW-SVR による予測精度が低い理由を考察するため、2014/12/4 の観測データに対して SDC で抽出された学習データの目的変数のヒストグラムを図 16 に示す。図 16 からは、目的変数が真値(3.06)付近の学習データの件数が絶対的にも相対的にも極めて少ないことが確認できる。そのため適切な説明変数を与え正しく類似性が判断されたとしても、SDC で全体の学習データの約 68%が収集されてしまうため、収集される学習データには類似度の低い学習データも多く存在する。この課題を解決するためには、類似度に合わせて動的に収集閾値を決定し、高い類似度の学習データだけを適切に収集する必要があると考える。また今回の評価実験では学習データ量が 2000 件と少なく、観測データに対して高い類似度の学習データの絶対量が少なくなる場合があったが、SW-SVR は SVR と比べ学習データ量を増やしてもモデル構築時間が増大し辛い特徴を持っているため、学習データ量を多くすることでより高精度な予測が期待できる。

6. おわりに

本研究では、SW-SVR を用いた施設園芸環境向け知的制御システムを提案し、プロトタイプシステムを実装した。高信頼な無線通信と高精度かつ高速な予測制御の実現のため、センサノード間無線通信には 429MHz 帯 IEEE802.15.6 方式を、予測アルゴリズムには提案手法である SW-SVR を採用した。提案システムの評価の結果、無線通信部では PER0.00%を実現し、データ予測部では窒素吸収量の予測誤差を従来手法と比べ MAE で 77%、RMSE で 68%削減した。特に SW-SVR を用いることで、学習データの蓄積数が少ない実験初期段階から高精度な予測制御が実現可能であることを示した。

今後の課題は、局所的な環境に対しても高精度な予測を実現するため、環境をより正確に表現する説明変数や、SDC の収集閾値を動的に決定する仕組みを検討する予定である。本実験終了後には、収穫される果実を評価して SW-SVR による予測制御の有用性について評価し、同時にインシヤルコストとランニングコストを算出してコスト面の評価も実施する予定である。

謝辞

本研究は、SCOPE 地域 ICT 振興型研究開発 (H25~26) と挑戦的萌芽研究(26660198)により実施したものである。

参考文献

- [1] 小林悠一, 他. “センサネットワークを用いた温湿度制御システムの研究”, 情処研報 UBI-37, pp.1-8, 2013.
- [2] 串岡 聡, 他. “無線センサネットワークを用いた施設園芸環境向け制御システムの開発と評価”, 情報処理学会第 74 回全国大会, 1, 7, 2012.
- [3] “FUJITSU Intelligent Society Solution 食・農クラウド Akisai(秋彩)” <http://jp.fujitsu.com/solutions/cloud/agri/>, (参照 2014/10/01)
- [4] Yuya Suzuki, et al., “Proposal to sliding window-based support vector regression”, Knowledge-Based and Intelligent Information & Engineering Systems, pp.1615-1624, 2014.
- [5] Alberto Camilli, et al., “From wireless sensors to field mapping: Anatomy of an application for precision agriculture”, Computers and Electronics in Agriculture, 58.1, pp.25-36, 2007.
- [6] Jing Huang, et al., “Development of a wireless soil sensor network”, ASABE annual international meeting, no.8, 2008.
- [7] George Vellidis, et al., “A real-time wireless smart sensor array for scheduling irrigation”, Computers and Electronics in Agriculture, 61.1, pp.44-50, 2008.
- [8] John Thelen, et al., “Radio wave propagation in potato fields”, Workshop on Wireless Network Measurements, 2005.
- [9] Yoav Freund, et al., “A decision-theoretic generalization of on-line learning and an application to boosting”, Journal of Computer and System Sciences, pp.119-139, 1997.
- [10] Benjamin Wang, et al., “Boosting support vector machines for imbalanced data sets”, Knowledge and Information Systems, 25.1, pp.1-20, 2010.
- [11] Brian Smith, et al., “Artificial neural networks for automated year-round air temperature prediction”, Computers and Electronics in Agriculture, 68.1, pp.52-61, 2009.
- [12] “IPA/SEC 重要インフラ情報システム信頼性研究会報告書”, <http://www.ipa.go.jp/files/000004741.pdf>, (参照 2014/10/01)
- [13] Masahiro Kuroda, et al., “Interoperable and Diligent Body Area Networks over IEEE802.15.6 for Real-time Monitoring”, IEEE Engineering in Medicine and Biology Society, pp.2224-2227, 2011.
- [14] 向山学, 他. “幾何学的学習アルゴリズムによる最大マージン識別法”, 信学論 J88-D-II(4), pp.687-694, 2005.
- [15] 鈴木雄也, 他. “近傍データ収集法を用いた SW-SVR の改良”, 情処研報 MBL, 1-8, 2014.
- [16] 阿部純義. “一般化された中心極限定理と Tsallis の非加法的統計力学 (第 8 回「非平衡系の統計物理 シンポジウム, 研究会報告)”, 物性研究, 75.5, 870-876, 2001.
- [17] Sago Yuki, et al., “Kinetics of Root Ion Absorption Affected by Environmental Factors and Transpiration: III. A Kinetic Model Integrated with Transpiration”, Environment control in biology, 49(1), pp.41-46, 2011.
- [18] 大石直記, “温室環境および植物生態情報取得のための散乱光マルチセンサユニットの開発”, 日本生物環境工学会, 2014.
- [19] David Meyer, “Support vector machines: The interface to libsvm in package e1071”, 2004.