

環境データを高精度・高速に 予測可能な機械学習システム

研究者：静岡大学大学院 情報学研究科
准教授 峰野 博史

説明者：峰野 博史

新技術の特徴

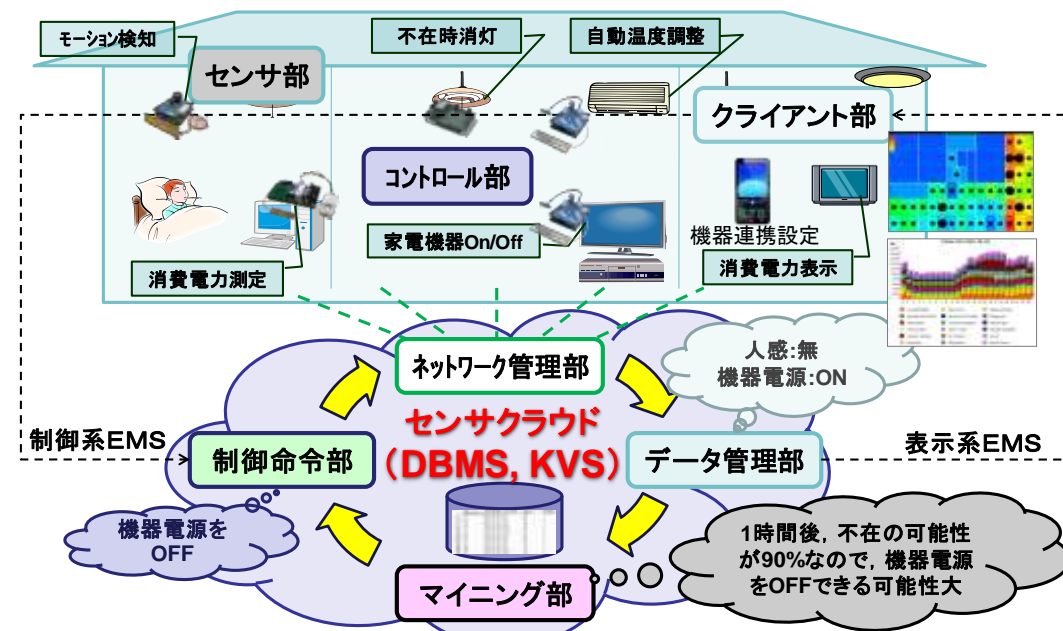
- 過去の環境データを機械学習させて将来の変化を予測するシステムにおいて、比較的近い未来の予測に特化した適切な学習データ量を選定する機能を組み合わせた新規の機械学習システム「**SW-SVR**」を開発し、高精度かつ高速な予測を可能とした

従来技術との比較

- 地域によって特性の異なる微気象(狭い領域の気象)を、精度よく予測するために、どのようなデータをどれだけ学習させモデルを構築するかには試行錯誤を必要とするが、本システムは予測に必要なデータ量を**自動調整**するだけでなく**高精度**かつ**高速な予測**を可能とする

背景

- ICTを駆使した各種制御システム
 - 可視化, 分析, 制御, 経営管理...
- ビッグデータの利活用
 - 環境データ, 作業記録, 品質...
- 将来予測に基づいた次世代制御支援
 - モデル予測制御(MPC)を用いた冷暖房空調設備(HVAC)の省電力化 [1]
 - 土壌水分量予測に基づく自動灌水制御の研究 [2]
- 予測制御システムに要求される機能
 - 高精度な予測
 - 適切な時間で終了できるデータ処理



文科省 地域イノベーションクラスター事業
「自立分散協調ユビキタスセンサネットワーク(H19-23)」

高精度かつ高速な予測を可能とするアルゴリズムが必要！

[1]Anil, Aswai, et.al, “Reducing Transient and Steady State Electricity Consumption in HVAC Using Learning-Based Model-Predictive Control”, Proc. of IEEE, pp.240-253, Vol.100, Jan.2012.

[2]Yuya Suzuki, et.al, “An SVM Based Irrigation Control System for Home Gardening”, IEEE GCCE2013, Oct.2013.

関連研究

Author	Outline	Prediction algorithm	Target
[1] Kadu et.al	アンサンブルモデルを用いたANNによる予測精度向上	Artificial Neural Network	気温
[2] Smith et.al	125万件のトレーニングデータを用いて、高精度予測可能なANNモデルを構築	Artificial Neural Network	気温
[3] Robert et.al	125万件のデータからトレーニングデータをランダムに抽出し、SVR学習に要する処理時間を短縮	Support Vector Regression	気温
[4] Sun et.al	トレーニングやモデル構築に分散処理を用いて高速化	Support Vector Regression	気温

既存研究：データの部分的除去，ビッグデータの活用，分散処理

➡ どのようなデータをどれだけ学習させモデルを構築するかには試行錯誤を必要とし，学習データの特徴を考慮した処理は想定していない

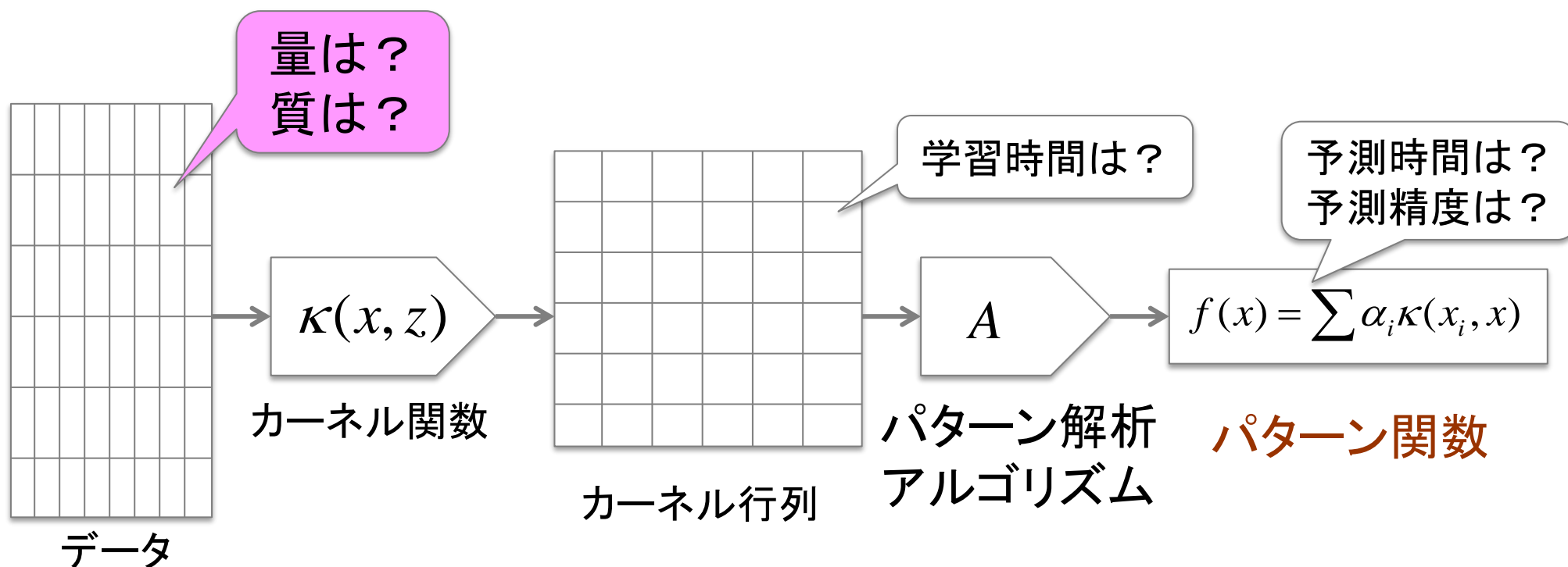
提案手法：データの特性を考慮して適切な学習量を自動調整

➡ 不要なデータを除去し，予測処理時間を短縮しつつ高精度化も両立

- [1]Kadu, et.al, "Analysis and Prediction of Temperature using Statistical Artificial Neural Network," *Intl. Journal of Computer Science*, 2012.
- [2]Smith, et.al, "Artificial neural networks for automated year-round air temperature prediction," *Computers and Electronics in Agriculture*, 2009.
- [3] Chevalier, et al. "Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks." *Neural Computing and Applications*, pp.151-159.,2011.
- [4] Sun,et.al. "Study on Parallel SVM Based on MapReduce." *Intl. Conf. on Parallel and Distributed Processing Techniques and Applications*. 2012.

カーネル法によるパターン解析の概要

- 与えられた有限のデータ標本(トレーニングデータ)において, パターンを線形の関係として発見可能な空間にデータを変換(カーネル関数)することで, 線形の関係としてパターンを発見(パターン解析アルゴリズム)する

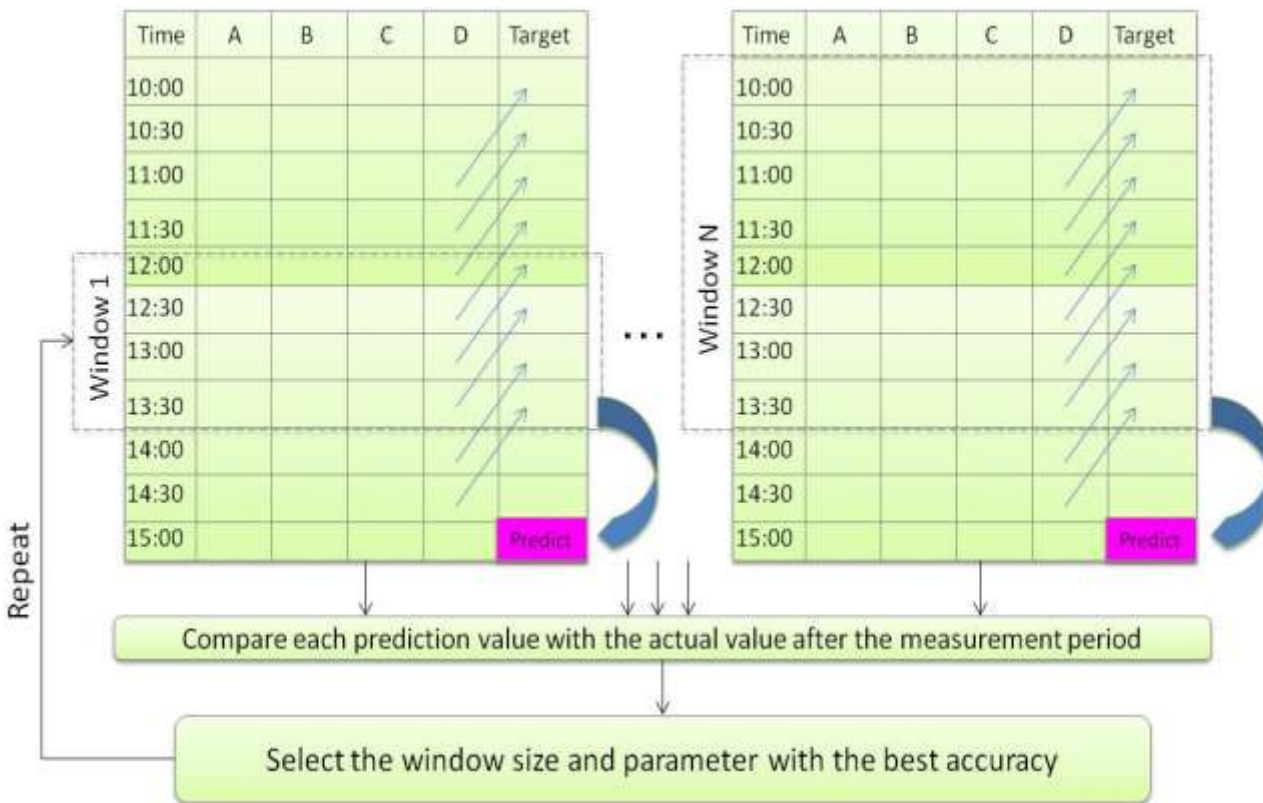


SW-SVRの概要

- SW-SVR: Sliding Window-based Support Vector Regression
 - SVR をベースとしたアルゴリズム
 - 高精度に予測可能な学習データ量を自動調整

ウィンドウサイズの分割例

Time	Season	Mo/Ev
JAN	Spring	Morning
FEB	Summer	Evening
MAR	Autumn	Morning
APR	Winter	Evening
MAY	Spring	Morning
JUN	Summer	Evening
JULY	Autumn	Morning



※ 想定対象：規則変動（経験則）のありそうな量的データ
（例：気温，待ち時間，株価？）

提案手法

- SW-SVR: Sliding Window-based Support Vector Regression

- 経年変化等の特性変化に対応可能

1. 学習データの抽出: 現在の予測環境に適した学習データを抽出
2. 自動再構築: 予測精度の変化に伴って自動的にモデル再構築



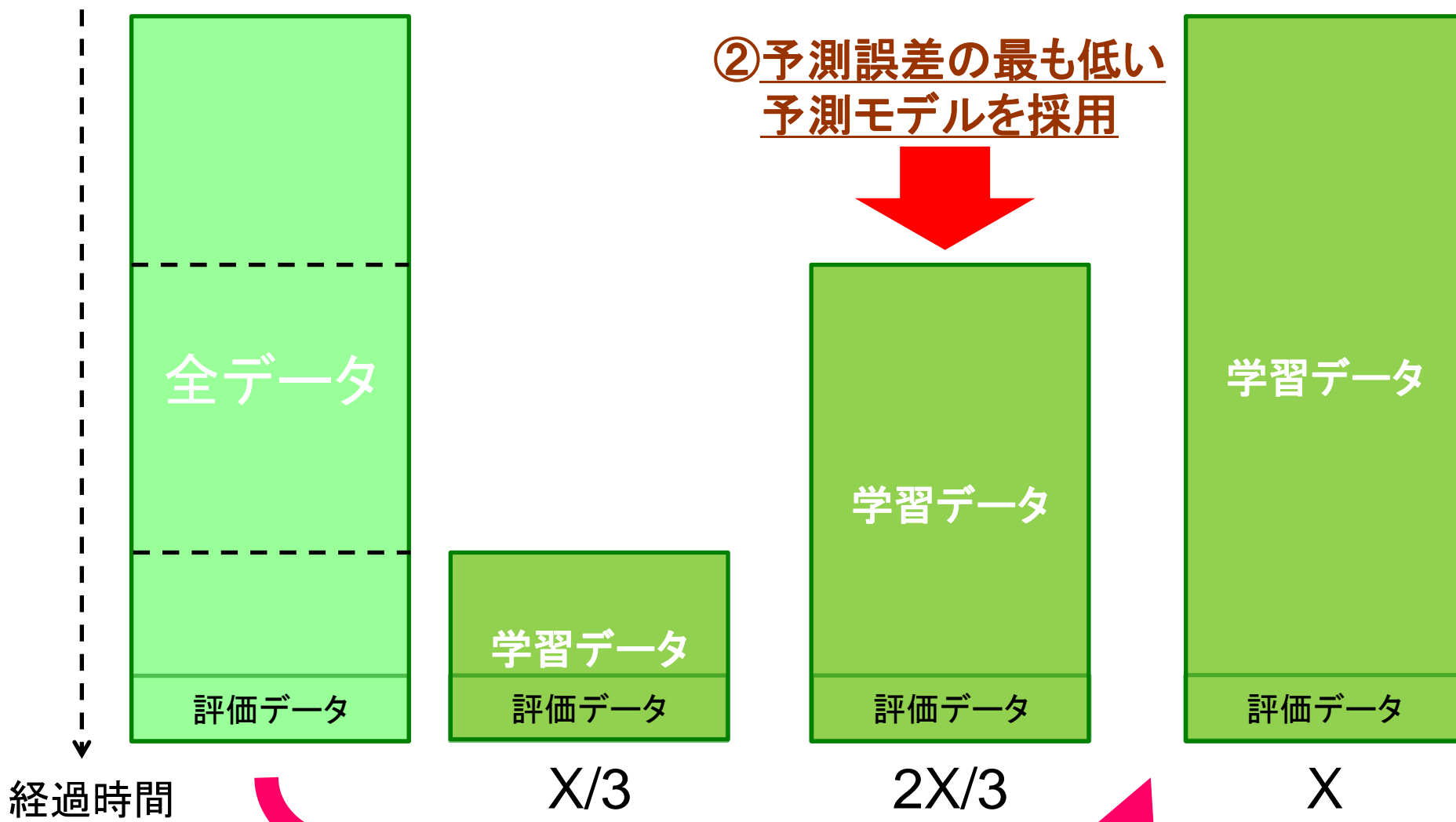
① 大規模データから複数のウィンドウサイズで予測モデルを構築

② 予測誤差の最も低い予測モデルを採用

③ 予測精度の低下(特性変化)を検知したらウィンドウサイズを調整して予測モデルを再構築

様々なウィンドウサイズでのモデル構築例

① 大規模データから複数のウィンドウサイズで予測モデルを構築



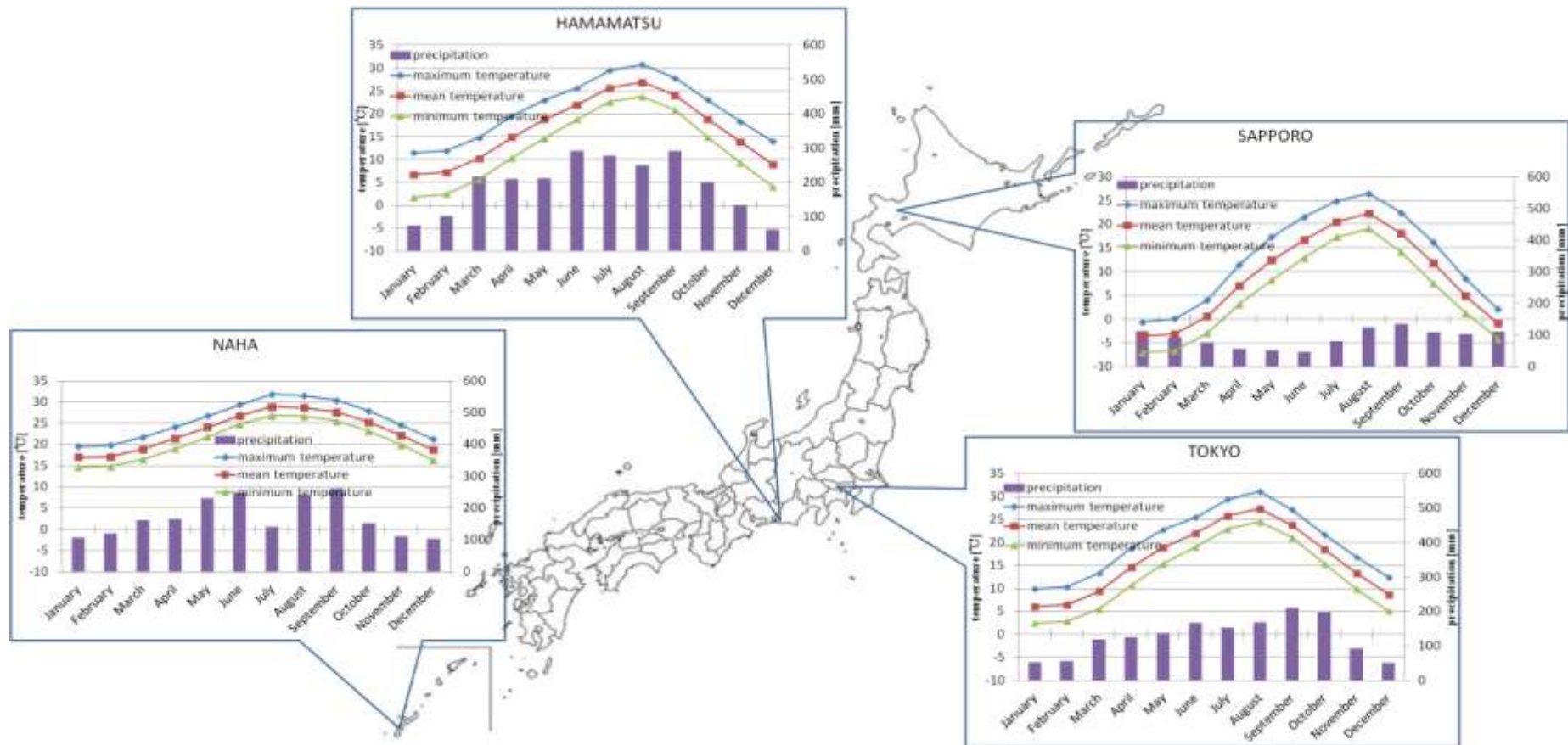
③ 予測精度の低下(特性変化)を検知したらウィンドウサイズを調整して予測モデルを再構築

基礎実験(1/2)

- 実験対象: 日本の気象データ@ AMeDAS

– 環境データ(時間経過による特性変化)

– 札幌, 東京, 浜松, 那覇の4地域を対象



基礎実験(2/2)

- 説明変数
 - 気温, 相対湿度, 気圧, 雨量, 平均風速, 最大風速, 日照時間
- 目的変数
 - 各地域の1時間後の気温を予測
- SW-SVRのウィンドウ数(並列処理数)
 - 8
- 評価指標
 - 予測誤差: RMSE (Root Mean Square Error)
- 学習期間・評価期間

$$RMSE = \sqrt{\frac{\sum_{i=1}^N \bar{e}_i^2}{N}}$$

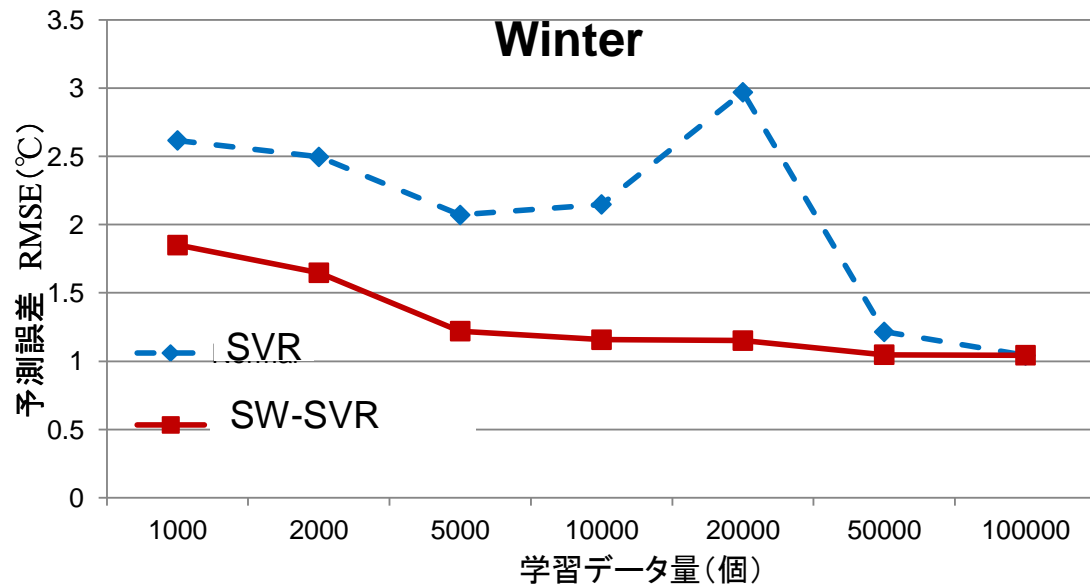
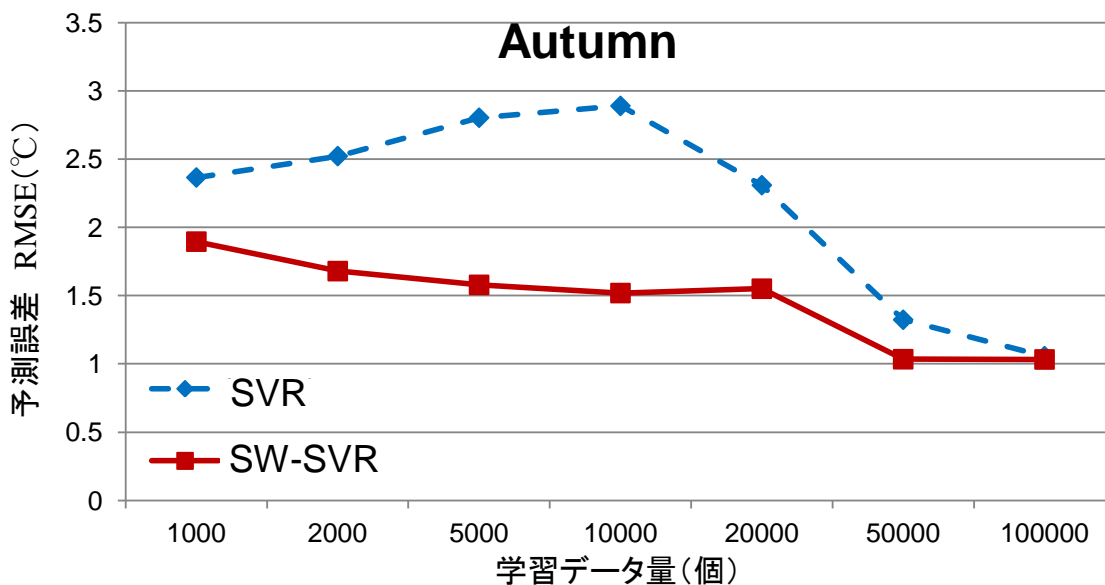
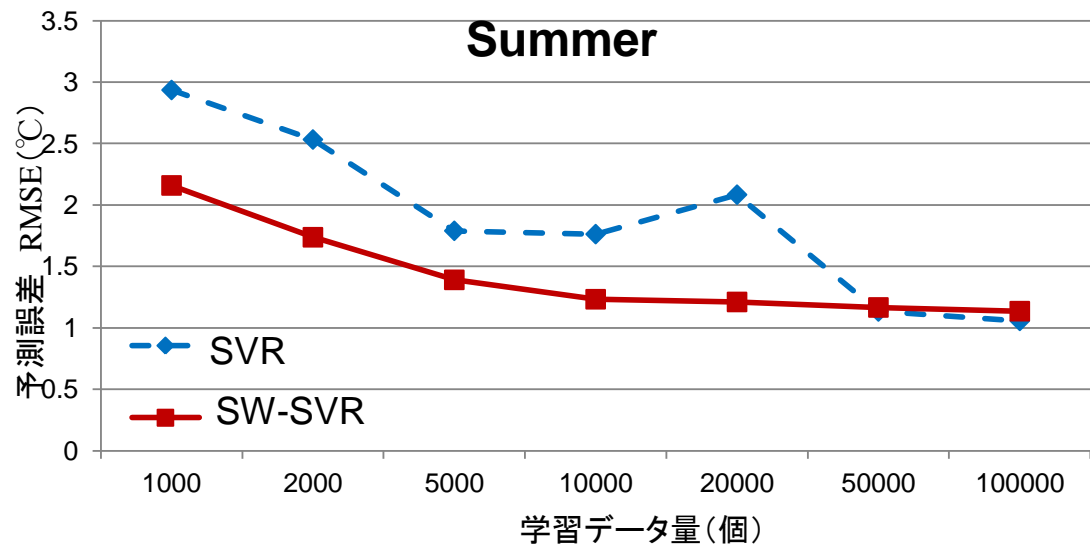
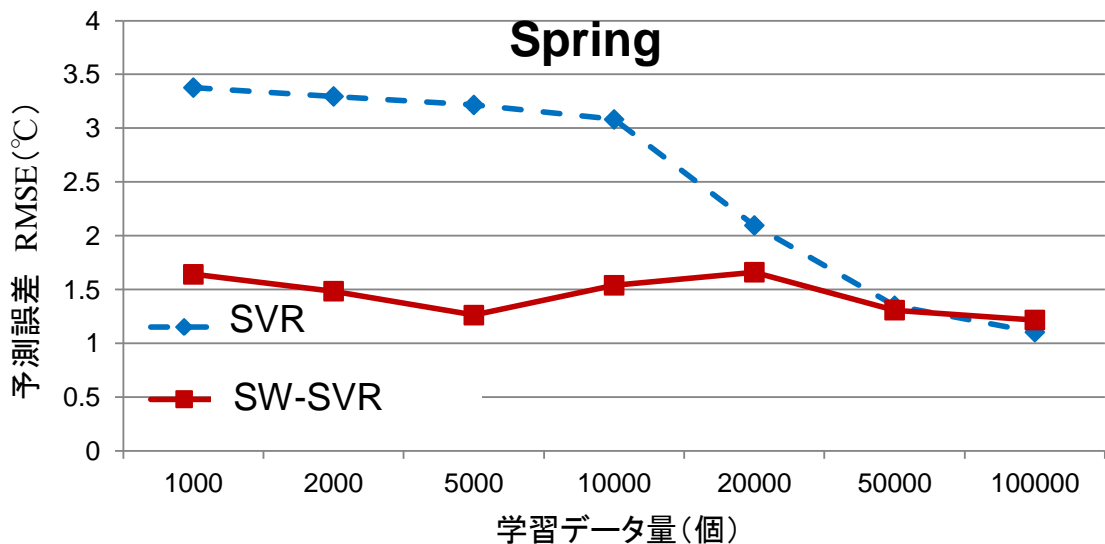
e: 誤差 N: 評価データ数

季節	学習期間※1	評価期間※2
春	2011/1/1~2013/3/31	2013/4/1~2013/5/1
夏	2011/1/1~2013/6/30	2013/7/1~2013/8/1
秋	2011/1/1~2013/9/30	2013/10/1~2013/11/1
冬	2011/1/1~2013/12/31	2014/1/1~2014/2/1

※1: 1,000件~100,000件の範囲で学習量を変化

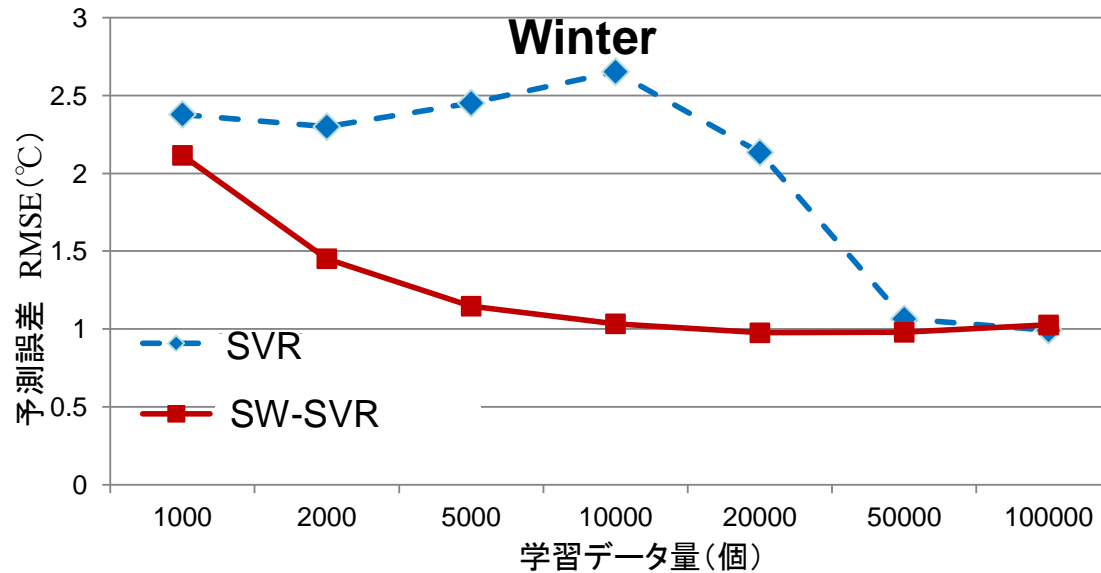
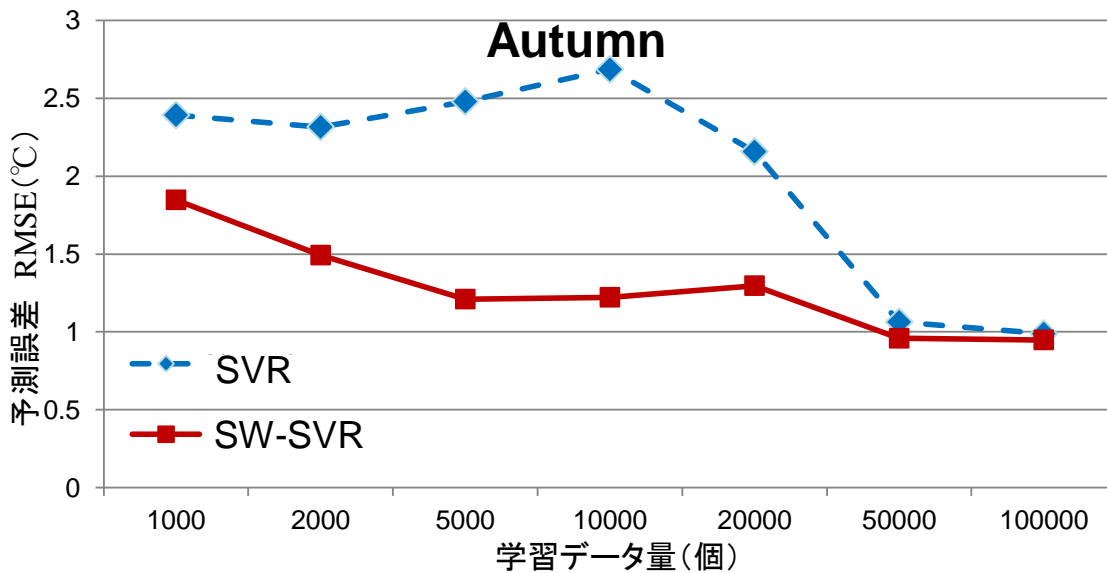
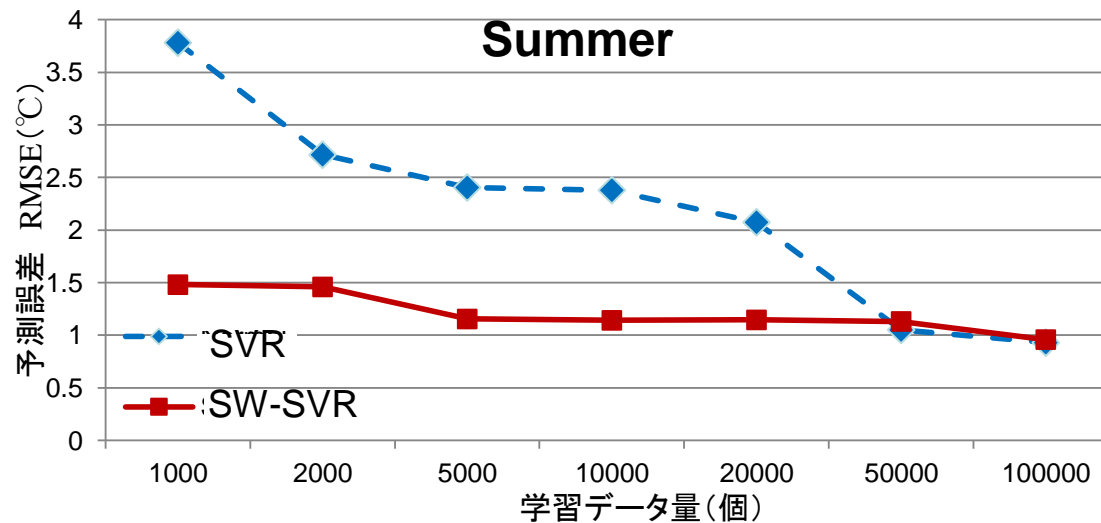
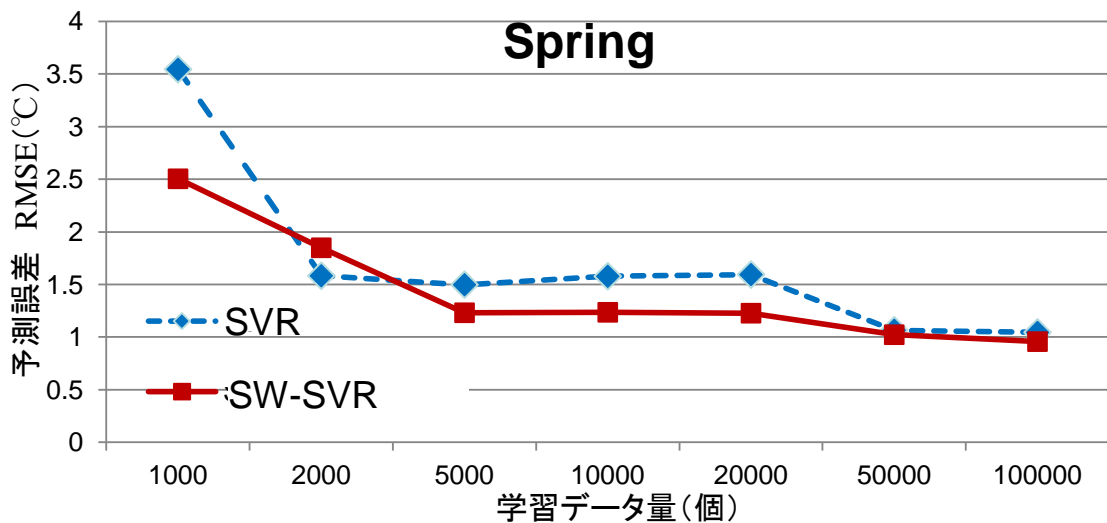
※2: 予測データ数は約4,500件

予測誤差(札幌)



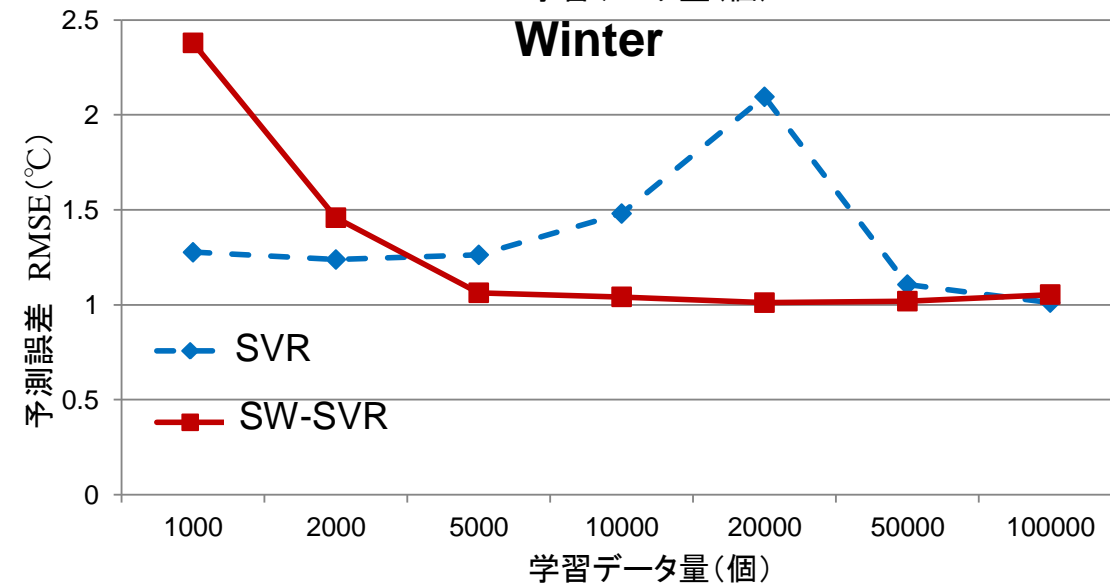
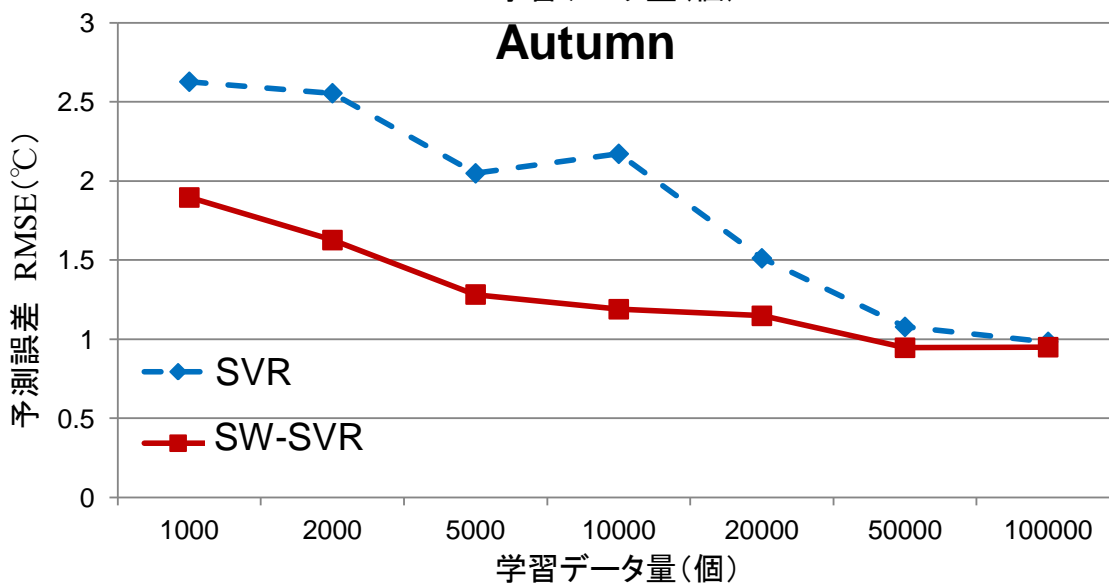
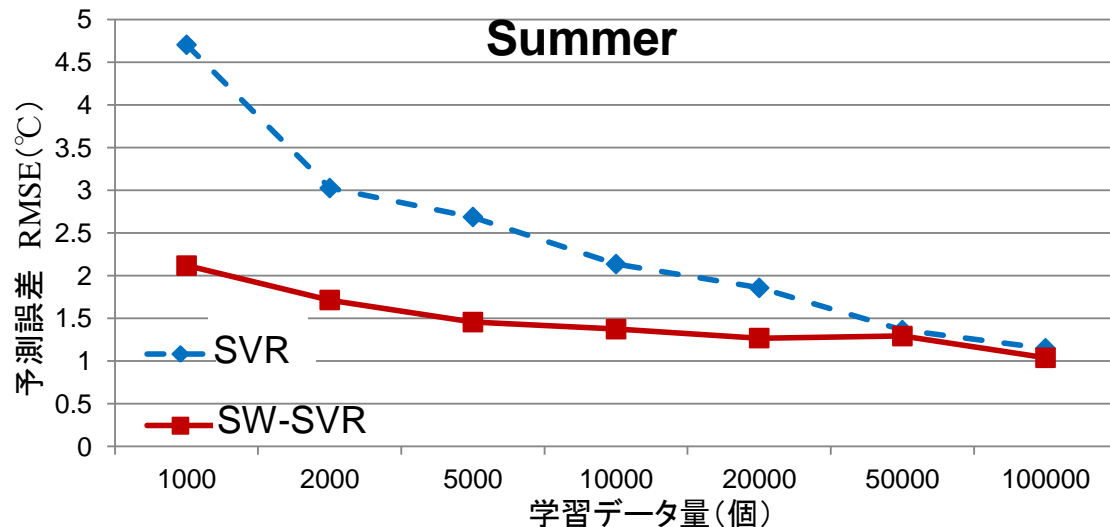
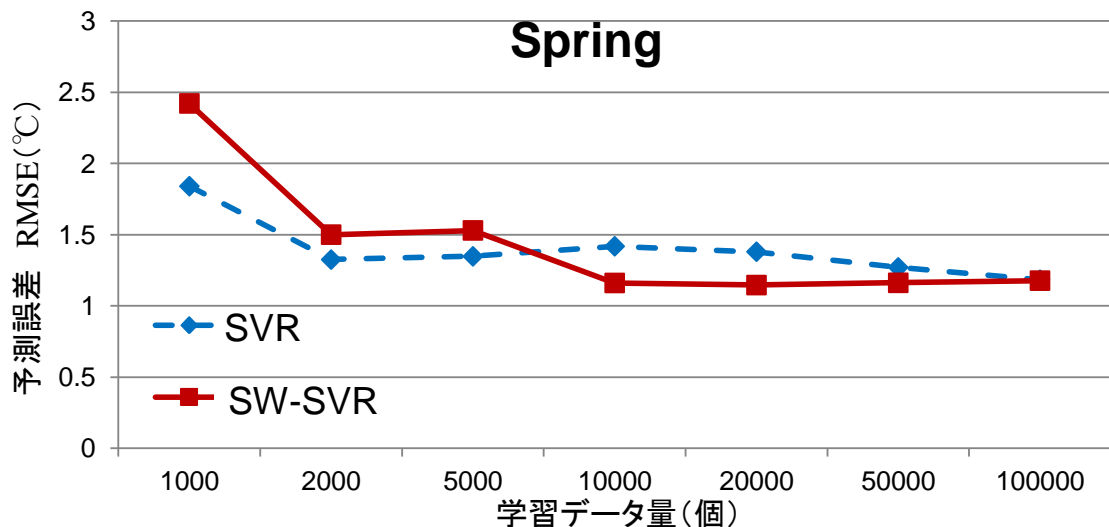
最大61%, 平均13% の予測誤差削減に成功

予測誤差(東京)



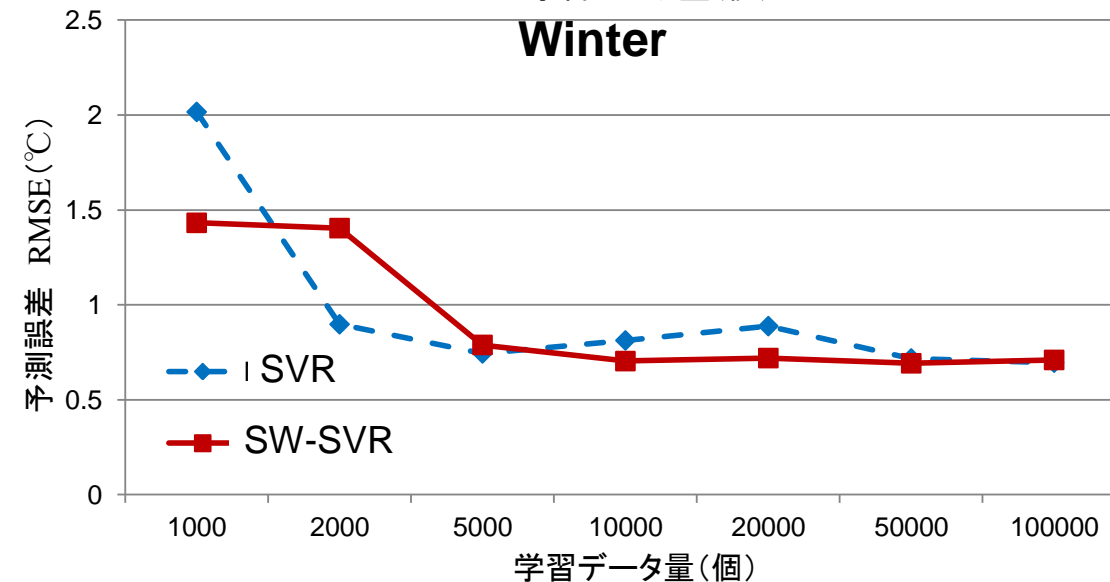
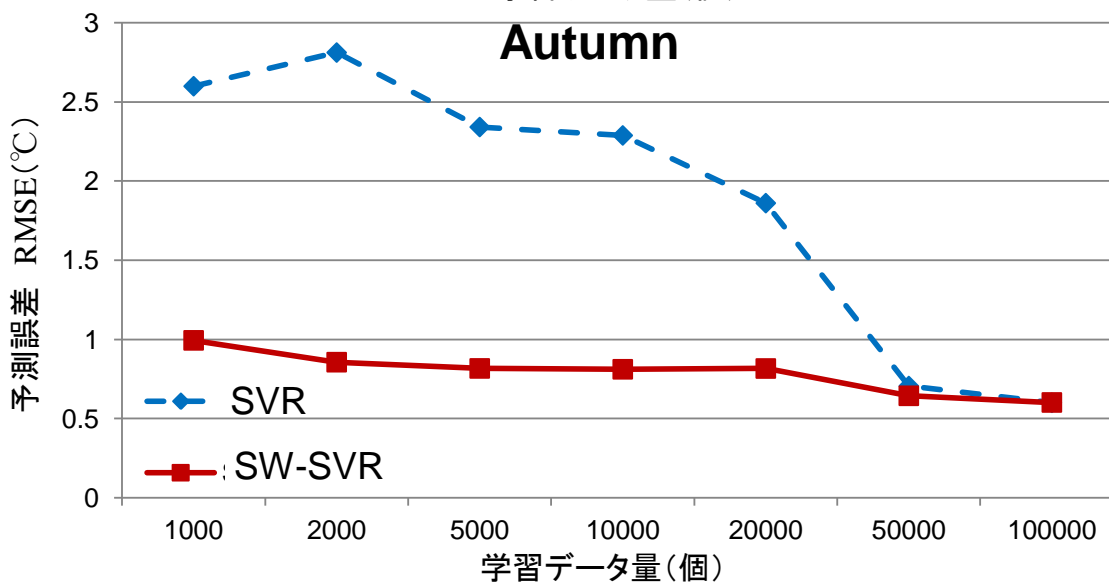
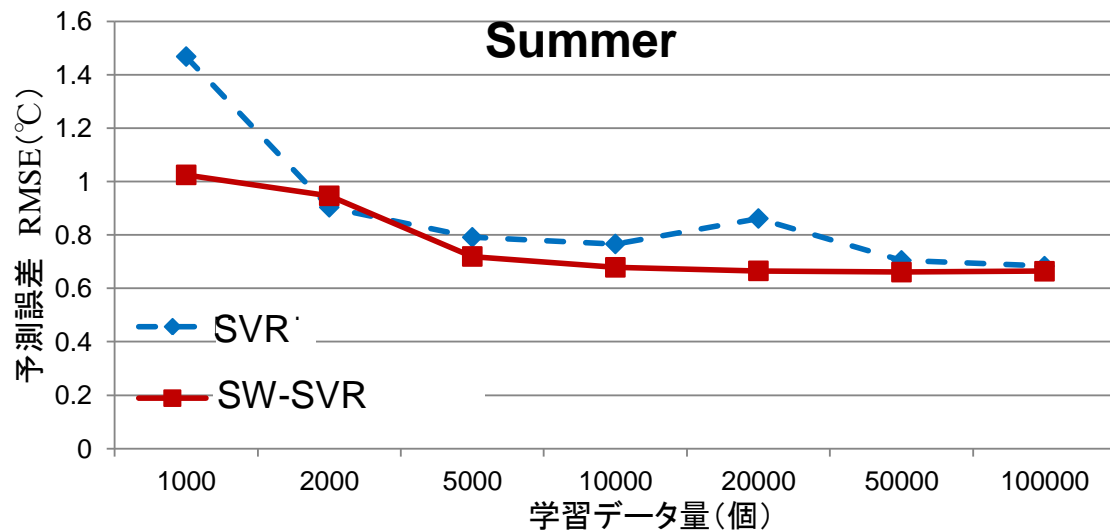
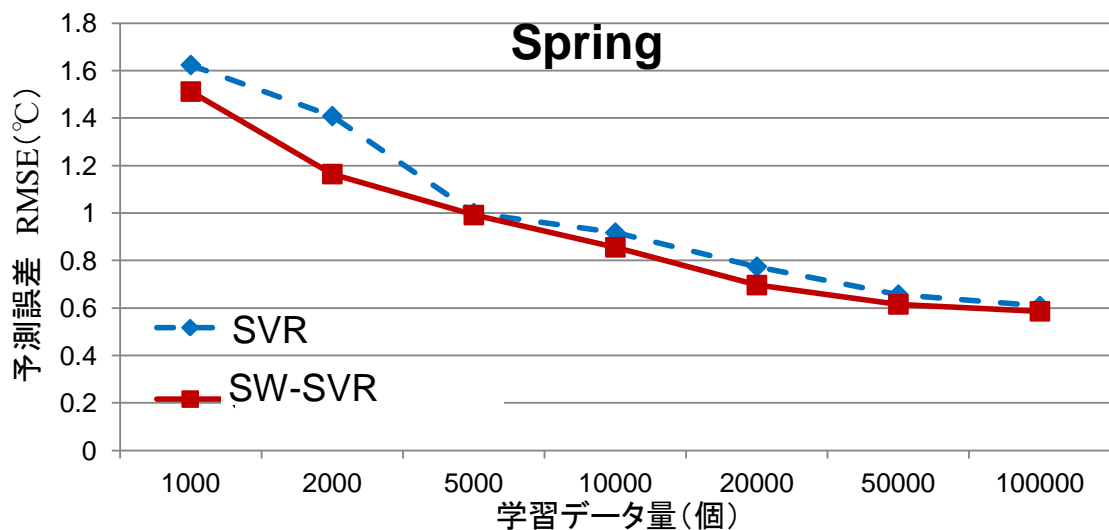
最大61%, 平均25% の予測誤差削減に成功

予測誤差(浜松)



最大70%, 平均33% の予測誤差削減に成功

予測誤差(那覇)



最大61%, 平均16% の予測誤差削減に成功

まとめ

背景

様々な予測制御システムにおいて、高精度かつ高速な予測を可能とするアルゴリズムが必要

提案手法

予測に必要なデータ量を、自動調整するだけでなく、高精度かつ高速に予測可能な、SW-SVRを研究開発

評価結果

AMeDASのオープンデータを用いて、気温の予測精度を比較した結果、SW-SVRはSVRに対し、同一学習データ量でも、最大70%、平均23%の予測誤差を削減可能なことを確認

今後の予定

- ① ウィンドウの分割手法について詳細検討
 - ・更なる予測精度向上を実現するアルゴリズムを研究開発中
- ② SW-SVRを用いた知的制御システムの効果を定量的に実証実験開始
 - ・システム全体の消費電力削減、収量増大、品質向上、など

これまでの外部資金獲得実績(代表のみ)

- ウィンドウサイズ自動調整型サポートベクターマシンによる高精度な微気象データ予測
 - H26~27 文科省 科研費 挑戦的萌芽研究
- モバイルデータ3Dオフローディングの研究
 - H26~29 文科省 科研費 基盤研究(B)
- 高度農業ICTを実現する高信頼双方向多点無線センサ/アクチュエータネットワークの研究開発
 - H25~26 総務省 SCOPE地域ICT振興型研究開発
- 多様な大規模センシングデータ間の時系列相関分析技術の研究
 - H22~23 文科省 科研費 挑戦的萌芽研究
- 高精密農業を可能とするマルチベンダセンサグリッドの実証的研究
 - H21~23 文科省 科研費 若手研究(A)
- 自律分散協調ユビキタスセンサネットワークの実現
 - H19~23 地域イノベーションクラスター事業(旧知的クラスター創成事業(第II期))(H21~代表)
- 大規模無線センサネットワークにおける複数基地局を用いたソフトウェア配送手法の研究
 - H20 浜松科学技術研究振興会 村田基金研究助成
- ハイブリッド無線通信応用ユビキタスネットワークングミドルウェアの開発
 - H16~18 文科省 科研費 若手研究(B)

※本技術に大きく関係するもの

(KDDI研究所, NTTドコモ研究所, 富士通研究所, 三菱電機研究所, RSO, NTT研究所, NICT等)

本技術に関する知的財産権、問合せ先

- ・ 発明の名称
： 予測システム、予測方法、および予測プログラム
- ・ 出願番号：特願2014-047431(3月11日)
- ・ 発明者：峰野博史、鈴木雄也
- ・ 出願人：国立大学法人 静岡大学

◎共同研究および関連する特許については、
静岡大学イノベーション社会連携推進機構にお問い合わせください。

コーディネーター：由比藤

TEL :053-478-1702

Email : ip-office@cjr.shizuoka.ac.jp



国立大学法人
静岡大学