Prediction of Regional Goods Demand Incorporating the Effect of Weather

Watanabe Takuya, Hiroaki Muroi, Motoki Naruke Edirium K.K.
Kyoto Yono Nowcast, Inc.
Gen Kobayashi University of Tsukuba
Masanori Yamasaki Acroquest Technology Co., Ltd.

IEEE Workshop on Big Data Analytics in Manufacturing and Supply Chains
2016 IEEE International Conference on Big Data

Outline

- Motivation
- Data
  - daily sales records
  - daily weather records
- Models
  - sales prediction models
- Methods
  - three methods of prediction
- Results
  - predictive performance
  - interaction between item and area
  - interaction between area and weather
- Discussion
- Summary

Motivation

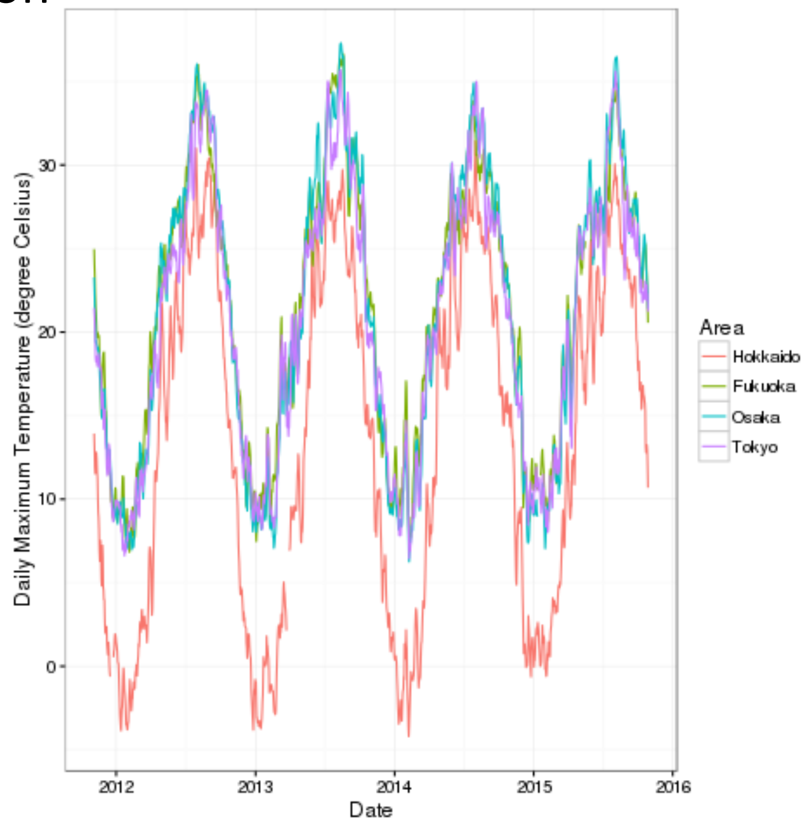
- Precise demand prediction of consumer goods is beneficial
  - supply chains can be optimised according to predicted demand
    - thus leading to higher profit
  - each item has different price elasticity
    - that is, effect of price change for sales is different for each item
  - goods demand is different in each area
  - effect of weather for goods demand would also be different in each area
- Big data can be utilised in two different ways
  - build complex models which need a vast amount of training data
    - image classification, speech recognition, machine translation etc.
  - sub-divide data according to some criteria for fine grained analysis
    - area, age, product model etc.
    - analyse each sub-divided data sets independently
    - being able to keep model's performance well because sub-divided data sets still have enough records

Data: Daily Sales Records

- Collected from Point of Sale (POS) systems
  - mainly of supermarkets in Japan
  - provided by Nikkei Digital Media, Inc.
  - 3 billion records, 1500 stores, 3 million items
  - from 1 Nov. 2011 to 31 Oct. 2015
- Consists of
  - store code
  - date
  - item code (Japan Article Number)
  - sales
  - quantity sold
    - target of the prediction
- Other tables
  - store code to area
  - item code to category

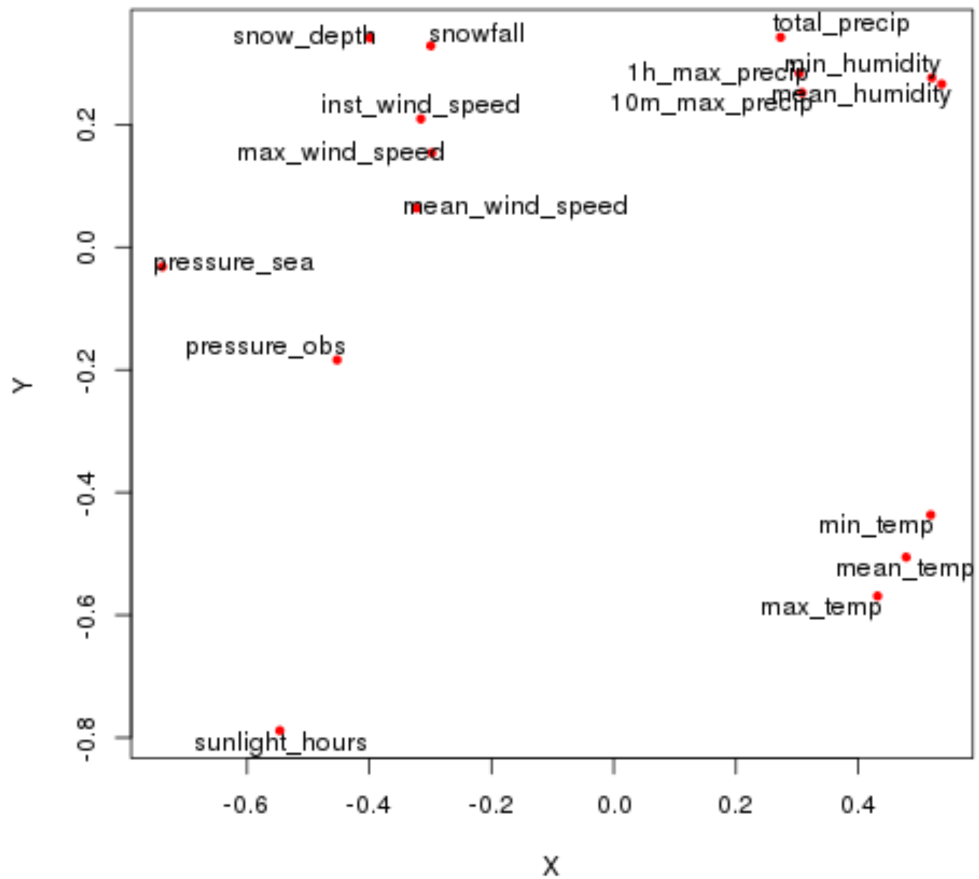
Data: Daily Weather Records

- Collected from the web site of Japan Meteorological Agency
  - http://www.jma.go.jp/jma/indexe.html
- Consists of 16 variables
  - atmospheric pressure at observatory/sea level
  - total/1h max/10m max precipitation
  - min/max/mean temperature etc.



Weather Variable Selection

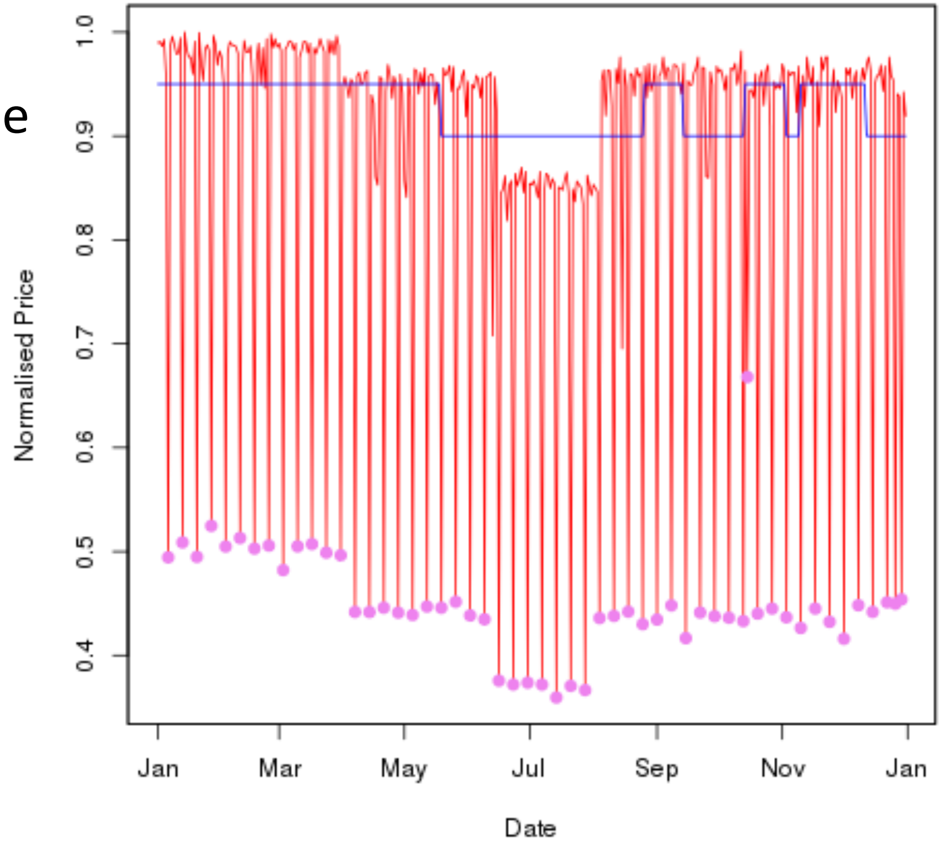
- Discard highly correlated variables
- Selected variables
  - average atmospheric pressure at observatory
  - total precipitation
  - maximum temperature
  - minimum humidity
  - maximum wind speed
  - sunlight hours
  - snow depth



Multi dimensional scaling plot of the weather variables

Mode Price and Sale Flag

- There would be an effect of "sale" itself
  - sales change which could not be explained by price change
- Calculate mode prices for each day
  - from the prices of the period of 28 days before and after each day
- Sale flag condition
  - prices below 1σ of the mode



Prediction Models

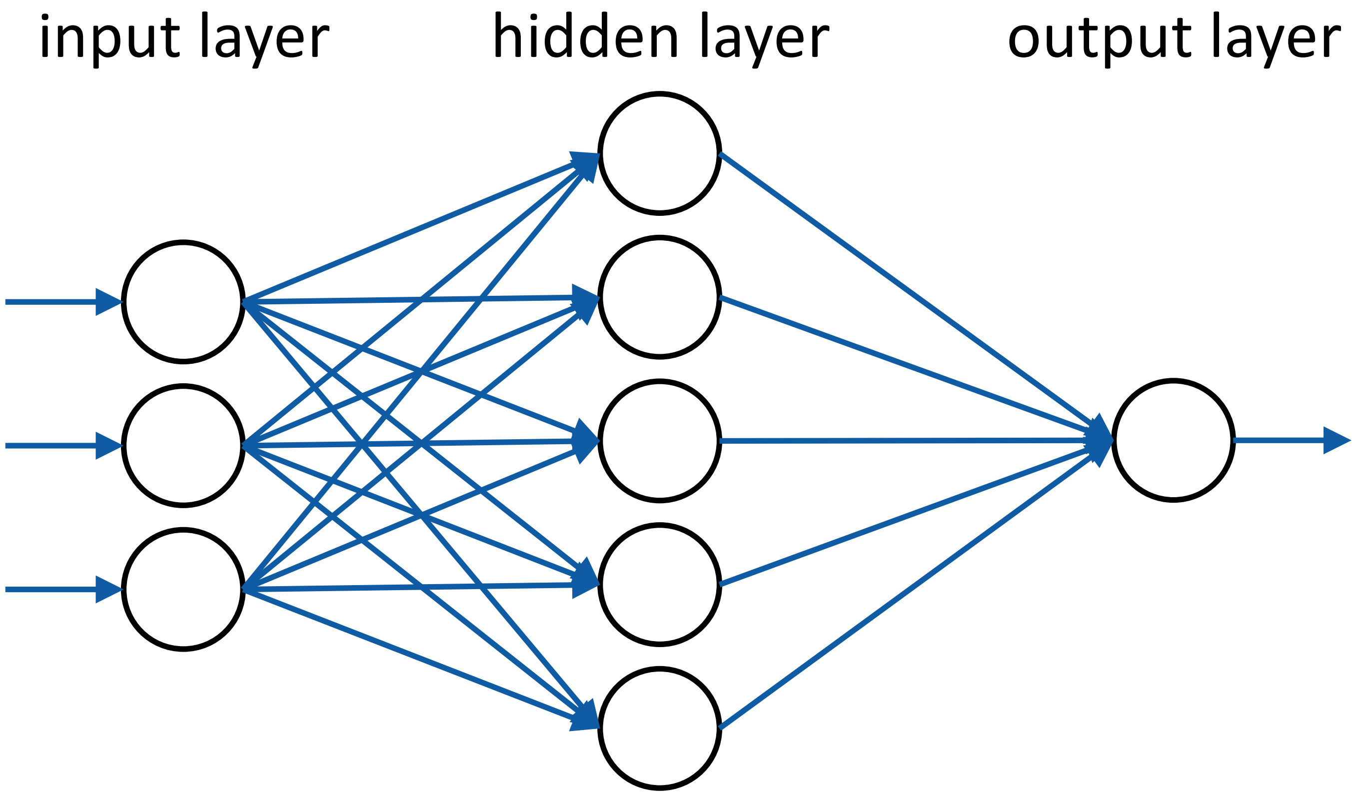
- Explanatory variables
  - quantity sold in the previous day
  - day of the week
  - season
  - weather related variables
  - price
  - sale flag
- Target variable: quantity sold
- Many fine grained models
  - dataset is separated by item and area
- Three prediction methods
  - linear regression
  - neural network
  - deep neural network

Method: Linear Regression

- Simple linear regression
  - not much to say

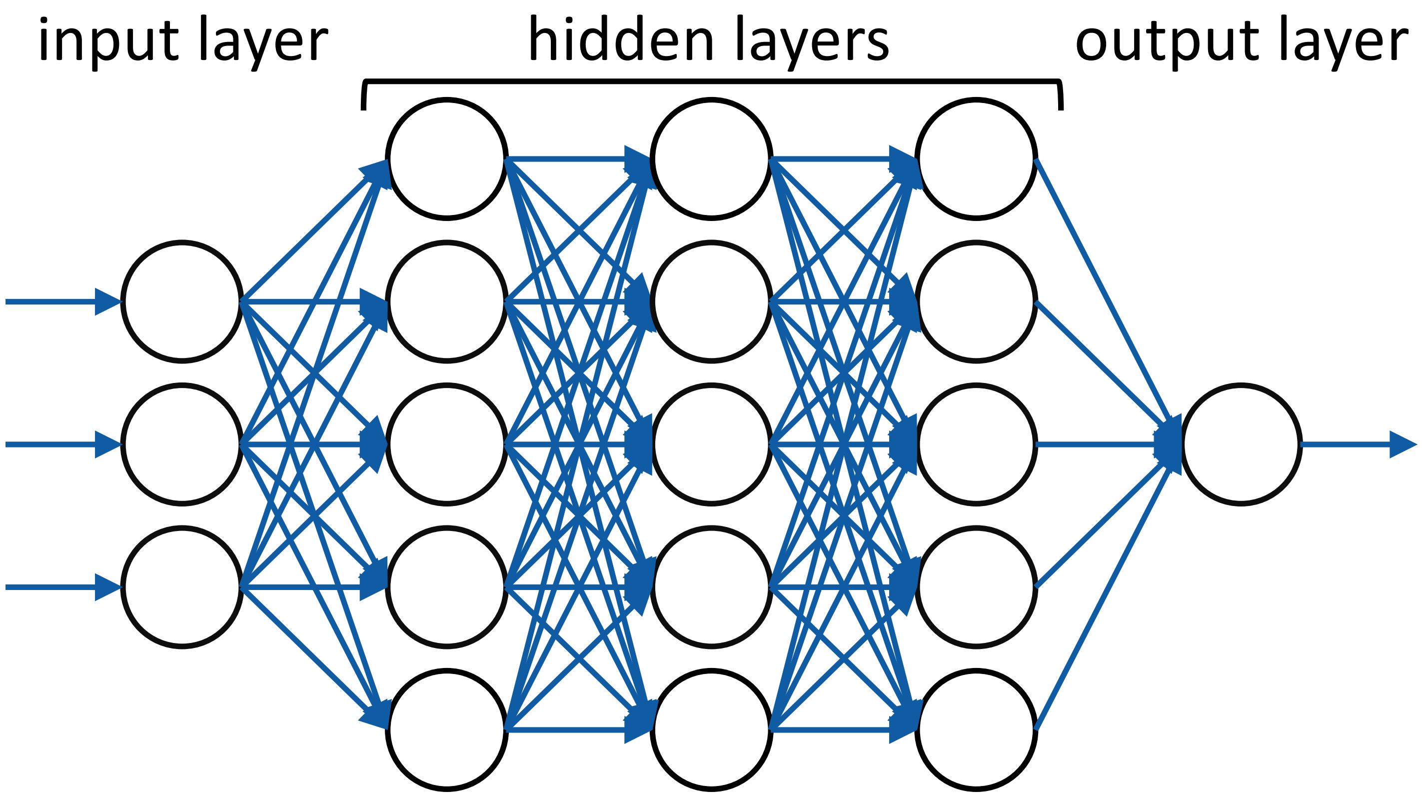
Method: Neural Network

- Basics
  - weight for each connection, activation function for each node
- Simple three layer network
  - only one hidden layer
  - single output node for regression



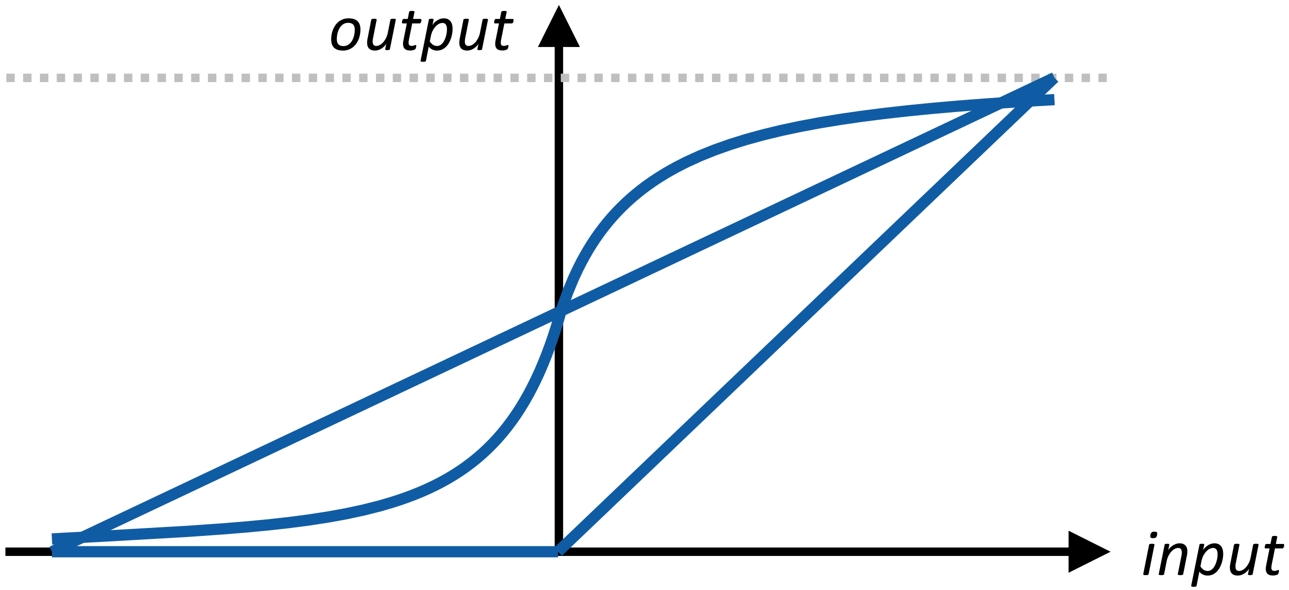
Method: Deep Neural Network

- Deeper structure
  - several number of hidden layers
    - actually not so deep, but adequate for this kind of simple regression problems?
  - single output node for regression (not changed)



Activation Functions

- Linear activation function (no activation function) for our NN
- Sigmoid function, traditionally used for classification tasks
- Rectified linear unit for our DNN
  - effective for faster convergence



System for Analysis

- Language
  - R and Python
- Libraries
  - *caret* for hyper parameter tuning
  - *nnet* for NN, *h2o* for DNN
- Database
  - PostgreSQL
    - single node database built on an Azure instance
    - approx. 1 TB data size including indices
    - too slow, even for relatively simple queries
  - Impala
    - rebuilt the database to seek higher performance
    - single node database built on an AWS EC2 instance
    - utilised I2 instance's SSD based instance storage for Avro format tables
    - significantly faster

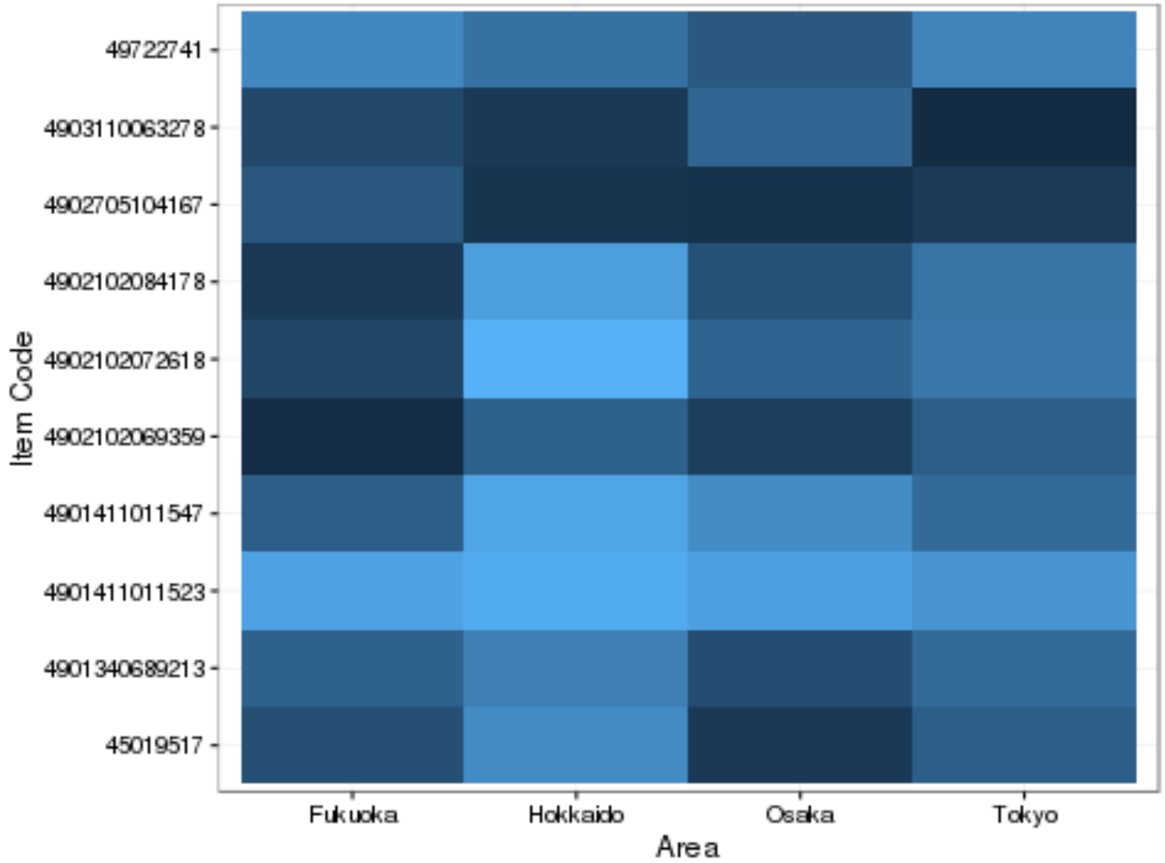
Results: Performance

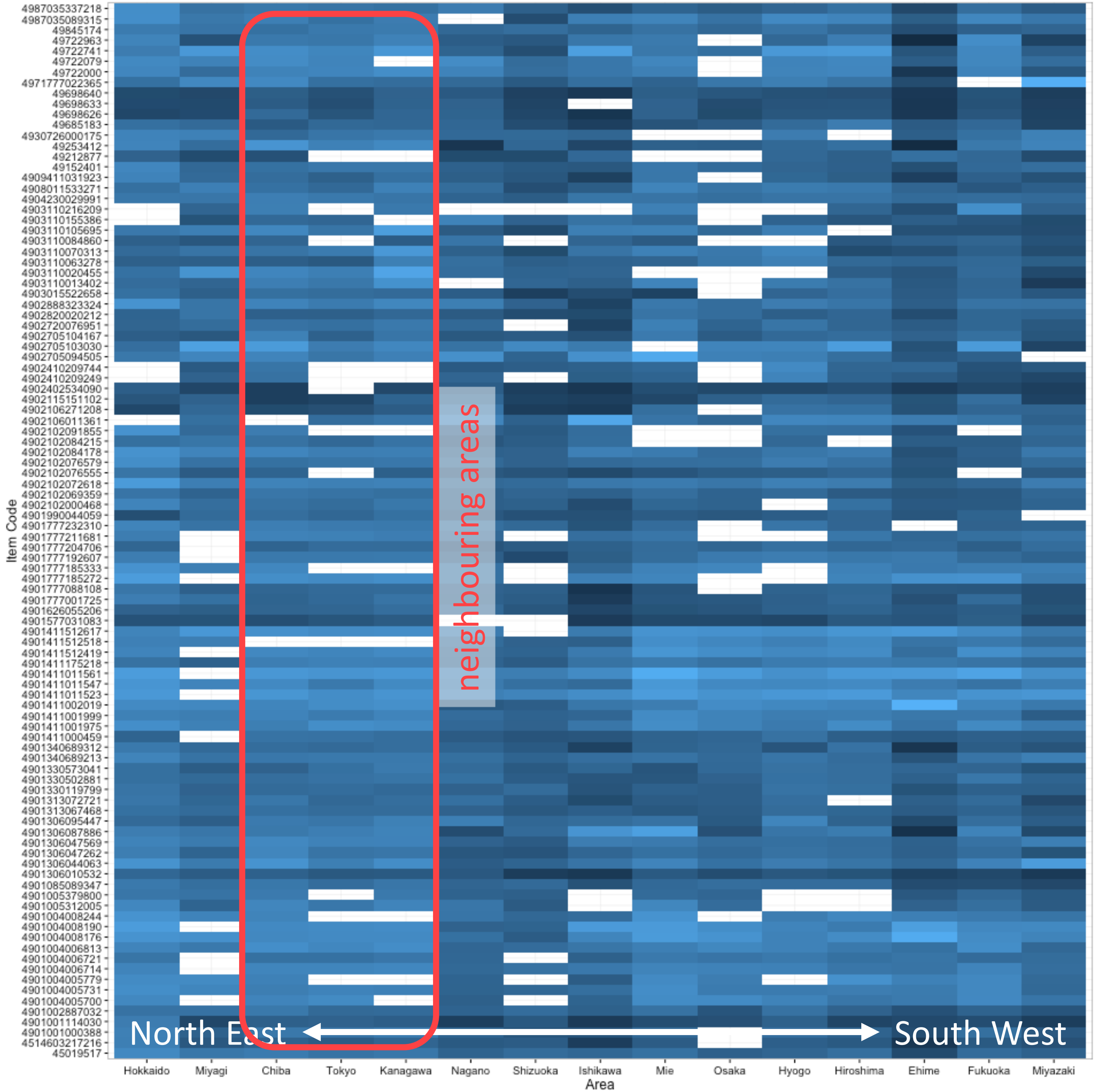
- Quantity sold is normalised to [0, 1] interval
  - predictive performance can simply be compared by root mean squared error (RMSE)
- Measured RMSE with 5-fold cross validation
- Neural network seems better than linear regression
- Deep neural network may be better than neural network

Area	**RMSE (Linear)**	**RMSE SD (Linear)**	**RMSE (Neural)**	**RMSE SD (Neural)**	**RMSE (Deep)**	**RMSE SD (Deep)**
Fukuoka	0.11320	0.00089	0.11137	0.00170	<u>0.11016</u>	0.00104
Hokkaido	0.09364	0.00356	0.09259	0.00327	<u>0.08527</u>	0.00101
Osaka	0.09448	0.00151	0.08918	0.00174	<u>0.08801</u>	0.00098
Tokyo	0.10308	0.00095	<u>0.09878</u>	0.00296	0.09934	0.00160

Results: Interactions between Item and Area

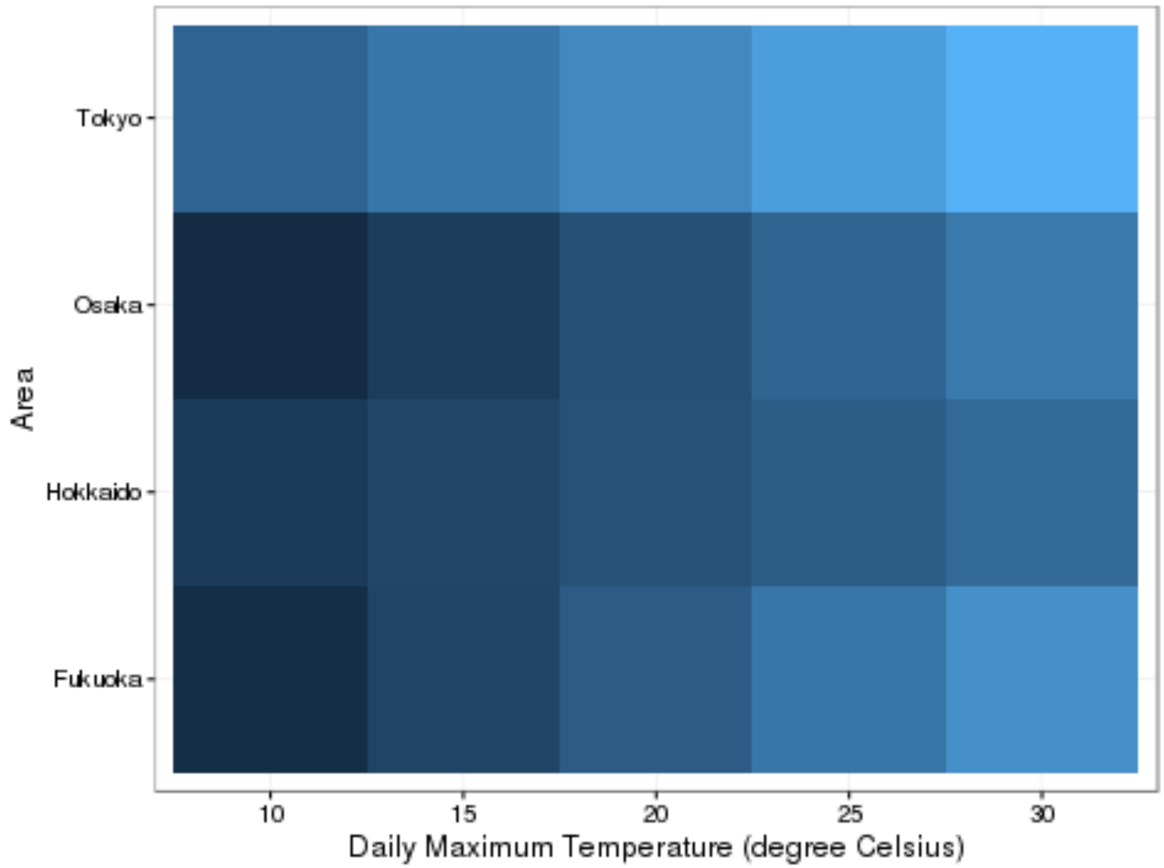
- Obtain demand curve for each item by changing the price
- Visualise demand by integrating the demand curve
  - lighter colour indicates more total demand
  - each item exhibits different pattern




Item Code
4987035337218
4987035089315
49845174
49722963
49722741
49722079
49722000
4971777022365
49698640
49698633
49698626
49685183
4930726000175
49253412
49212877
49152401
4909411031923
4908011533271
4904230029991
4903110216209
4903110155386
4903110105695
4903110084860
4903110070313
4903110063278
4903110020455
4903110013402
4903015522658
4902888323324
4902820020212
4902720076951
4902705104167
4902705103030
4902705094505
4902410209744
4902410209249
4902402534090
4902115151102
4902106271208
4902106011361
4902102091855
4902102084215
4902102084178
4902102076579
4902102076555
4902102072618
4902102069359
4902102000468
4901990044059
4901777232310
4901777211681
4901777204706
4901777192607
4901777185333
4901777185272
4901777088108
4901777001725
4901626055206
4901577031083
4901411512617
4901411512518
4901411512419
4901411175218
4901411011561
4901411011547
4901411011523
4901411002019
4901411001999
4901411001975
4901411000459
4901340689312
4901340689213
4901330573041
4901330502881
4901330119799
4901313072721
4901313067468
4901306095447
4901306087886
4901306047569
4901306047262
4901306044063
4901306010532
4901085089347
4901005379800
4901005312005
4901004008244
4901004008190
4901004008176
4901004006813
4901004006721
4901004006714
4901004005779
4901004005731
4901004005700
4901002887032
4901001114030
4901001000388
4514603217216
45019517
neighbouring areas
North East
South West
Hokkaido
Miyagi
Chiba
Tokyo
Kanagawa
Nagano
Shizuoka
Ishikawa
Mie
Osaka
Hyogo
Hiroshima
Ehime
Fukuoka
Miyazaki
Area

Results: Interactions between Area and Weather

- Visualise demand by a similar way as the item x area case
- Example: an ice cream product
  - higher temperature draws more demand
  - above effect is not apparent in Hokkaido (northern island)



Discussions

- Validity of the models and observations
  - a single big model with area variables could possibly perform better
  - no statistical check is done as to whether there is a real difference
- Effect of mid-term changes of economical circumstances
  - would be better to incorporate yearly trends into our models
- Effect of associations between items
  - would be better to incorporate category/manufacturer information into our models
- Better (meta-)models and/or prediction methods
  - there might be more suitable neural network models for this kind of regression problems
  - other prediction methods could be applied

Summary

- Analysed daily sales records (POS data) and weather records
  - to predict goods demand
- Compared performance of three prediction method
  - linear regression < neural network < deep neural network...?
- Visualised total demand by integrating the demand curve
  - interactions between item and area
  - interactions between area and weather