Simulation-based procedure for predictor's selection - integrating machine-learning and model-based approach in survey sampling

Alicja Wolny-Dominiak, Tomasz Żądło

Department of Statistical and Mathematical Methods in Economics
Department of Statistics, Econometrics and Mathematics
University of Economics in Katowice

Acknowledgement

The work has been co-financed by
the Minister of Science under
the "Regional Initiative of Excellence" programme.

Minister of Science
Republic of Poland

Regionalna
Inicjatywa
Doskonałości

Motivating paper

Tzavidis et al. (2018) propose a three-stage framework for the production of small area official statistics:

- STAGE 1: specification
  - specify user needs
  - data availability and geographical coverage
- STAGE 2: analysis & adaptation
  - initial triplet of estimates
  - use of models for small area estimation
  - model building
- STAGE 3: evaluation
  - uncertainty assessment
  - method evaluation - design-based (or model-based) simulation study

Content

- Introduction
- Predicting a function of the response variable - the ex-ante approach
- The proposed WASP method
- Real-data application

Table: Terms in survey sampling vs machine learning

survey sampling	**machine learning**
auxiliary variable	predictor
class of predictors	predictive algorithm / prediction method
predictor	prediction strategy
multivariate predictor	joint prediction strategy

Introduction

Models:

- parametric,
- nonparametric.

Various predictive algorithms (predictors), e.g.:

- BLUPs and EBLUPs,
- BPs and EBPs,
- PLUG-IN.

Ex-ante approach:

- usage of various accuracy measures, e.g. prediction RMSE, QAPE,
- selection based on voting.

WASP:

Voting-based
ex-**A**nte method
for **S**electing
joint **P**rediction strategy

Introduction: proposal

Example of a voting matrix
for S voters choosing one of P candidates

	candidate 1	candidate 2	...	candidate P
voter 1	v_{11}	v_{12}	...	v_{1P}
voter 2	v_{21}	v_{22}	...	v_{2P}
...	...	...	...	
voter S	v_{S1}	v_{S2}	...	v_{SP}

Example of a matrix of values of accuracy measures
for P prediction strategies under S simulation scenarios

	strategy 1	strategy 2	...	strategy P
simulation scenario 1	r_{11}	r_{12}	...	r_{1P}
simulation scenario 2	r_{21}	r_{22}	...	r_{2P}
...	...	...	...	
simulation scenario S	r_{S1}	r_{S2}	...	r_{SP}

Introduction: proposal

A method with the following features:

- joint prediction of any vector of characteristics,
- using a set of any ex-ante prediction accuracy measures,
- utilizing any type of data (cross-sectional, longitudinal, time-series),
- considering various scenarios for out-of-sample data (also not observed in the sample) defined by any class of models (both parametric and nonparametric models can be used simultaneously).
- different proposals of selection criteria.

... to be used in any prediction problem (not only in survey sampling).

Predicting a function of the response variable: input data

Two sets:

- S of n sample observations
  - $\mathbf{X}_S = [x_{ij}]_{n \times q}$ - known, fixed
  - $\mathbf{y}_S = [y_1 \quad y_1 \quad \dots \quad y_n]^T$ - a realization of $\mathbf{Y}_S = [Y_1 \quad Y_1 \quad \dots \quad Y_n]^T$
- R of k out-of-sample observations
  - $\mathbf{X}_R = [x_{ij}]_{k \times q}$ - known or assumed, fixed
  - $\mathbf{Y}_R = [Y_1 \quad Y_1 \quad \dots \quad Y_k]^T$ - random with unknown realizations

Let $\mathbf{Y} = [\mathbf{Y}_S^T \quad \mathbf{Y}_R^T]^T$ and $\mathbf{X} = [\mathbf{X}_S^T \quad \mathbf{X}_R^T]^T$.

Predicting a function of the response variable: models and prediction strategies

The aim: prediction of a vector of characteristics:

$$\boldsymbol{\theta} = \left[\theta^{(1)}(\mathbf{Y}) \quad \theta^{(2)}(\mathbf{Y}) \quad \dots \quad \theta^{(C)}(\mathbf{Y})\right]^T, \tag{1}$$

where $c = 1, ..., C$.

Predicting a function of the response variable: models and prediction strategies

An example of a regression model:

$$\begin{cases} \mathbf{Y} = h(\mathbf{X}) + \xi \\ E(\xi) = \mathbf{0} \\ Var(\xi) = \mathbf{V} \end{cases} \tag{2}$$

where $h(.)$ is a fixed but unknown function of independent variables, ξ is a random term with $\mathbf{0}$ mean and unknown variance-covariance matrix $\mathbf{V}$.

Predicting a function of the response variable: models and prediction strategies

Special cases of (2) include e.g.:

- the multiple regression model (Baltagi, 2021),
- the linear mixed model (Jiang, 2007),
- machine learning models (Hastie et al., 2009).

Predicting a function of the response variable: models and prediction strategies

$\hat{H}^{(g)}(.)$ - an estimator of $h(.)$ based on gth model $M^{(g)}(\mathbf{Y}, \mathbf{X})$.

$\hat{h}^{(g)}(.)$ - an estimate i.e. the realization of $\hat{H}^{(g)}(.)$.

If the gth model is either parametric or nonparametric, the estimates of $h(.)$ will be denoted by $\hat{h}^{(g)}_{PAR}(.)$ or $\hat{h}^{(g)}_{NPAR}(.)$, respectively.

Predicting a function of the response variable: models and prediction strategies

The PLUG-IN of C characteristics (1) under the gth model:

$$\hat{\boldsymbol{\theta}}^{(g)}_{PLUG-IN} = \left[\hat{\theta}^{(g,1)}_{PLUG-IN}(\mathbf{Y}_S) \quad \hat{\theta}^{(g,2)}_{PLUG-IN}(\mathbf{Y}_S) \quad \ldots \hat{\theta}^{(g,C)}_{PLUG-IN}(\mathbf{Y}_S)\right]^T, \quad (3)$$

where

$$\hat{\theta}^{(g,c)}_{PLUG-IN} = \hat{\theta}^{(g,c)}_{PLUG-IN}(\mathbf{Y}_S) = \theta^{(c)}\left(\begin{bmatrix} \mathbf{Y}_S \\ \hat{H}^{(g)}(\mathbf{X}_R) \end{bmatrix}\right), \quad (4)$$

$\hat{H}(.)$ is an estimator of $h(.)$ (see (2)).

Predicting a function of the response variable: two approaches to prediction accuracy

sample observations

known and fixed $\mathbf{X}_S$

random $\mathbf{Y}_S$

known realization of $\mathbf{Y}_S$: $\mathbf{y}_S$

training set (A)	**test set (B)**
$\mathbf{X}_{S(A)}$ $\mathbf{y}_{S(A)}$	$\mathbf{X}_{S(B)}$ $\mathbf{y}_{S(B)}$ predicted values: $\widehat{h}(\mathbf{X}_{S(B)})$

ex-post predicion accuracy

based on $\mathbf{y}_{S(B)} - \widehat{h}(\mathbf{X}_{S(B)})$

out-of-sample observations

known (or assumed) and fixed $\mathbf{X}_R$

random $\mathbf{Y}_R$

unknown realization of $\mathbf{Y}_R$: $\mathbf{y}_R$

predicted values: $\widehat{h}(\mathbf{X}_R)$
(realizations of $\widehat{H}(\mathbf{X}_R)$)

ex-ante prediction accuracy

based on random $\mathbf{Y}_R - \widehat{H}(\mathbf{X}_R)$

Predicting a function of the response variable: ex-ante prediction accuracy measures

The prediction error:

$$U = \hat{\theta}(\mathbf{Y}_S) - \theta(\mathbf{Y}) = \hat{\theta} - \theta$$

The prediction RMSE:

$$RMSE(\hat{\theta}) = \sqrt{E(\hat{\theta} - \theta)^2} = \sqrt{E(U^2)} \tag{5}$$

Predicting a function of the response variable: ex-ante prediction accuracy measures

The pth **Q**uantile of **A**bsolute **P**rediction **E**rror (Żądło, 2013; Wolny-Dominiak i Żądło, 2022):

$$QAPE_p(\hat{\theta}) = \inf\left\{x : P\left(\left|\hat{\theta} - \theta\right| \leq x\right) \geq p\right\} \tag{6}$$

This measure informs that at least $p100\%$ of observed absolute prediction errors are smaller than or equal to $QAPE_p(\hat{\theta})$, while at least $(1 - p)100\%$ of them are higher than or equal to $QAPE_p(\hat{\theta})$.

qape R package on CRAN: 13 000 downloads

Predicting a function of the response variable: Bootstrap under parametric and nonparametric models

Generated bootstrap realizations of prediction errors:

$$U_{gen} = \hat{\theta}(\mathbf{y}_{s\,gen}) - \theta(\mathbf{y}_{gen}), \tag{7}$$

where $\mathbf{y}_{s\,gen}$ and $\mathbf{y}_{gen}$, are generated sample and population vectors of the dependent variable, respectively.

Predicting a function of the response variable: Bootstrap under parametric and nonparametric models

Bootstrap under parametric model - the parametric bootstrap:

$$\mathbf{y}_{gen}^{(g,b)} = \hat{h}_{PAR}^{(g)}(\mathbf{X}) + \xi_{gen}^{(g,b)}, \quad (8)$$

where $g = 1, \ldots, G$; $b = 1, \ldots, B$; $\hat{h}_{PAR}^{(g)}(.)$ is a parametric estimate of $h(.)$ obtained based on the original sample under the assumption of the gth (here: parametric) model, $\xi_{gen}^{(g,b)}$ is a generated bth realisation of an error term from the estimated parametric distribution assumed for ξ under the gth model.

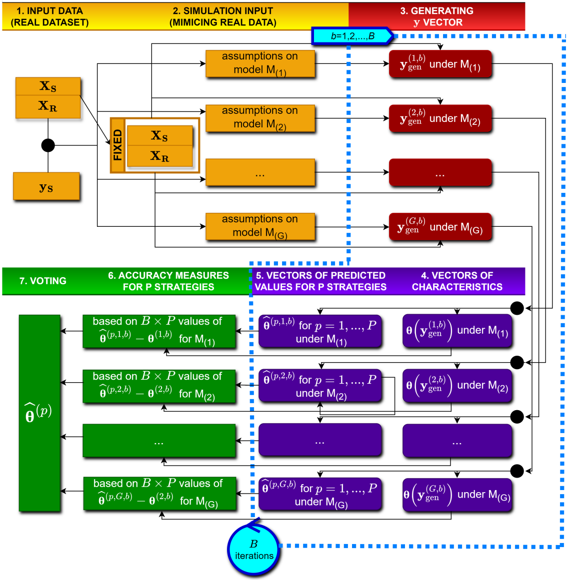
Predicting a function of the response variable: Bootstrap under parametric and nonparametric models

Bootstrap under nonparametric model (our proposal):

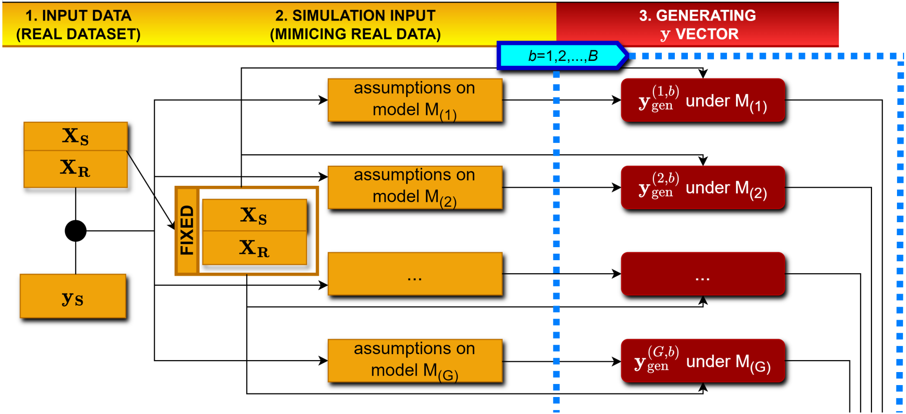
$$\mathbf{y}_{gen}^{(g,b)} = \hat{h}_{NPAR}^{(g)}(\mathbf{X}) + g\bar{e}n^{(b)}(\hat{f}(\mathbf{r}_S^{(g)}), (n+k)), \qquad (9)$$

where $g = 1, \ldots, G$; $b = 1, \ldots, B$; $\hat{h}_{NPAR}^{(g)}(.)$ is a nonparametric estimate of $h(.)$ obtained based on the original sample under the assumption of the gth (here: nonparametric) model, $\hat{f}(\mathbf{r}_S^{(g)})$ is a kernel density estimate of the distribution of residuals $\mathbf{r}_S^{(g)}$ computed under the gth model, and $g\bar{e}n^{(b)}(\hat{f}(\mathbf{r}_S^{(g)}), (n+k))$ is a $(n+k) \times 1$ zero centred vector of values generated based on kernel estimate $\hat{f}(\mathbf{r}_S^{(g)})$ in the bth iteration.

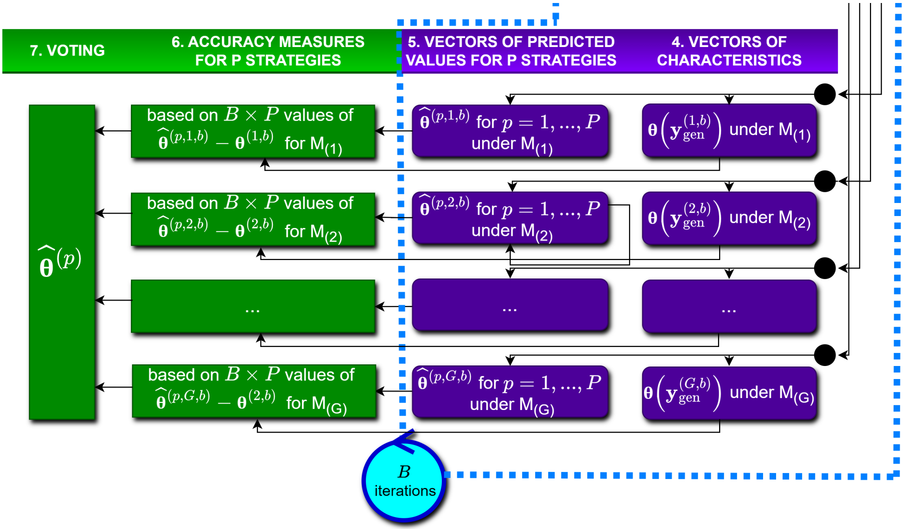
The proposed WASP method



The proposed WASP method



The proposed WASP method



Values of accuracy measures obtained in the simulation form a matrix of size $S \times P$, where:

- $S = G \times C \times M$ rows where G is the number of models used to generate data, C is the number of predicted characteristics, M is the number of considered accuracy measures ("voters"),
- P is the number of considered prediction strategies ("canditates").

It will be called the **accuracy measures matrix**, and it will be denoted by $\mathbf{A}$

The proposed WASP method: Algorithm 1

The first proposal: to mimic the first-past-the-post voting system (Felsenthal, 2012), where voters choose a single candidate, and the candidate with the highest number of votes wins the election.

The proposed WASP method: Algorithm 1

Algorithm 1 The proposed first-past-the-post voting algorithm of the prediction strategy selection

The input is an accuracy measure matrix $\mathbf{A}$.

1: **for** s in 1:S **do**

2: In the sth row of $\mathbf{A}$, assign rank 1 for the prediction strategy with the minimum value of mth accuracy measure and rank 0 for the rest of the strategies.

3: **end for**

4: Write the resulting ranks as a voting matrix $\mathbf{W_1}$ with S rows and P columns.

5: Compute column sums of the ranks for each out of P potential prediction strategies.

6: The prediction strategy with the highest sum is chosen.

The proposed WASP method: Algorithm 1

Table: Accuracy measures matrix **A** - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	30,409.52	92,010.24	30,525.97	67,948.94	75,471.81	91,967.10
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	33,892.15	82,222.11	33,973.64	79,599.19	78,954.10	82,234.57
$M_1, \theta_2, RMSE$	916.26	548.11	916.13	413.71	613.53	508.82
...	...	...	...	...	...	...

Table: First-past-the-post voting matrix $\mathbf{W_1}$ - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	1	0	0	0	0	0
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	1	0	0	0	0	0
$M_1, \theta_2, RMSE$	0	0	0	1	0	0
...	...	...	...	...	...	...
sum	**14**	5	7	3	1	6

The proposed WASP method: Algorithm 2

The second proposal: to mimic the Borda count system, which belongs to positional voting procedures (Felsenthal and Machover, 2012), where voters **order** candidates from the worst to the best, and in the original Borda count algorithm, the candidate with the highest sum/mean of points is elected.

In our proposal we choose candidate with the highest median because ranks are measured on the ordinal scale.

The usage of positional voting system (Felsenthal and Machover, 2012):

- in elections of the president of the Republic of Ireland,
- elections to the Australian House of Representatives,
- and some municipal elections in the United States.

The proposed WASP method: Algorithm 2

Algorithm 2 The proposed positional voting algorithm of the prediction strategy selection

The input is an accuracy measure matrix **A**.

for s in 1:S **do**

  In the sth row of **A**, rank prediction strategies according to the values of the mth accuracy measure from 1 (the maximum value in this row) to P (the minimum value in this row).

end for

Write resulting ranks as a voting matrix $\mathbf{W_1}$ with S rows and P columns.

Compute the median of ranks for each out of P prediction strategies.

The prediction strategy with the highest median rank is chosen.

Table: Accuracy measures matrix **A** - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	30,409.52	92,010.24	30,525.97	67,948.94	75,471.81	91,967.10
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	33,892.15	82,222.11	33,973.64	79,599.19	78,954.10	82,234.57
$M_1, \theta_2, RMSE$	916.26	548.11	916.13	413.71	613.53	508.82
...	...	...	...	...	...	...

Table: Positional voting matrix $\mathbf{W_2}$ - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	6	1	5	4	3	2
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	6	2	5	3	4	1
$M_1, \theta_2, RMSE$	1	4	2	6	3	5
...	...	...	...	...	...	...
median	**5**	4	4	3	3	3

The third proposal: inspired by

- evaluative voting, also called **utilitarian voting** (Baujard et al. (2014)
- and Majority Judgement (Balinski and Laraki, 2007).

In evaluative voting, each voter evaluates all candidates by assigning an **ordinal grade** that reflects their suitability (the better the candidate the higher the grade). Hence, the same grade can be given to several candidates. The candidate with the highest total/mean grade is the winner.

The procedure of the Majority Judgement is similar but the highest median grade is the rule to choose the candidate.

The proposed WASP method: Algorithm 3

In our approach the assessment is based on scaled accuracy measures (on a ratio scale) ...

... not based on the ranks (on the ordinal scale) as in the evaluating/utilitarian voting.

Scaled values of accuracy measures:

$$a' = 1 - a, \tag{10}$$

where a is a rescaled value (min-max normalization) of the considered accuracy measure.

Hence, $a' \in [0, 1]$ and the higher the value of (10), the better is accuracy.

The proposed WASP method: Algorithm 3

Algorithm 3 The proposed evaluative voting algorithm of the prediction strategy selection

The input is an accuracy measure matrix $\mathbf{A}$.

1: **for** s in $1{:}S$ **do**

2: In the sth row of $\mathbf{A}$, scale the values of the mth accuracy measure applying formula (10) to obtain P-values from interval $[0,1]$.

3: **end for**

4: Write resulting values as a matrix $\mathbf{W_3}$ with S rows and P columns (called the voting matrix).

5: Compute the median of scaled values for each out of P prediction strategies.

6: the prediction strategy with the highest median value is chosen.

The proposed WASP method: Algorithm 3

Table: Accuracy measures matrix **A** - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	30,409.52	92,010.24	30,525.97	67,948.94	75,471.81	91,967.10
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	33,892.15	82,222.11	33,973.64	79,599.19	78,954.10	82,234.57
$M_1, \theta_2, RMSE$	916.26	548.11	916.13	413.71	613.53	508.82
...	...	...	...	...	...	...

Table: Scaled voting matrix $\mathbf{W_3}$ - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	1.000	0.000	0.998	0.391	0.268	0.001
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	1.000	0.000	0.998	0.055	0.068	0.000
$M_1, \theta_2, RMSE$	0.000	0.733	0.000	1.000	0.602	0.811
...	...	...	...	...	...	...
median	**0.999**	0.454	0.997	0.650	0.318	0.401

$\mathbf{W_3}$ will be used in the Algorithm 4 too ...

The drawback of Algorithm 3 - it only considers the median of the results in the selection process, not the whole distribution.

Hence, our fourth proposal ...

Algorithm 4 The proposed ECDF AUC-based voting algorithm of the prediction strategy selection

The input is an accuracy measure matrix $\mathbf{A}$.

1: **for** s in 1:S **do**

2: In the sth row of $\mathbf{A}$, scale the values of the mth accuracy measure applying formula (10) to obtain P-values from interval $[0, 1]$.

3: **end for**

4: Write resulting values as a matrix $\mathbf{W_3}$ with S rows and P columns (called the voting matrix).

5: Based on the values in the pth column, compute the ECDF for the scaled voting results obtained for the pth prediction strategy (where $p = 1, 1, \ldots, P$).

6: The prediction strategy with the smallest value of the area under the ECDF in interval $[0, 1]$ is chosen.

The proposed WASP method: Algorithm 4

Table: Scaled voting matrix $\mathbf{W_3}$ - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	1.000	0.000	0.998	0.391	0.268	0.001
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	1.000	0.000	0.998	0.055	0.068	0.000
$M_1, \theta_2, RMSE$	0.000	0.733	0.000	1.000	0.602	0.811
...	...	...	...	...	...	...
ECDF	ECDF 1	ECDF 2	ECDF 3	ECDF 4	ECDF 5	ECDF 6
AUC in $[0, 1]$	0.338	0.506	0.346	0.439	0.655	0.550

Automobile portfolio example

Input data - the portfolio from Polish insurance company for full 2007-2010 years (exposure $= 1$).
The longitudinal structure of data - each ith policy corresponds to the aggregate value of claims for single policy *Claim_Amount* and risk factors:

- *Gender* - category: 1 (Female), 0 (Male)
- *Kind_of_distr* - kind of district: urban, country, suburban
- *Kind_of_payment* - category: cash, transfer
- *Engine* - category: BEN, DIS
- *Age_group* - category: 1, 2, 3.

Two key characteristics describing the portfolio of policies: total value of claims $\theta_{(1)}$, the median of claims $\theta_{(2)}$.

Automobile portfolio example

The joint prediction of characteristics $\theta_{(1)}$ and $\theta_{(2)}$ with WASP - assumptions:

1. First assumption - how to obtain 'true' claims values
   - parametric M_{PAR}: GLM Gamma (GG), log-normal regression (LogN), GAM
   - non-parametric M_{NPAR}: decision tree (DT), SVM with linear kernel (SVML), SVM with polynomial kernel (SVMP)
2. Second assumption - which predictive models
3. Third assumption - which prediction algorithm for every predictive model

Table: Description of candidate prediction strategies

Strategy	Predictive model	Prediction algorithm
strategy 1	GG	PLUG-IN
strategy 2	LogN	PLUG-IN
strategy 3	GAM	PLUG-IN
strategy 4	DT	PLUG-IN
strategy 5	SVML	PLUG-IN
strategy 6	SVMP	PLUG-IN

4. Fourth assumption - the criteria for selecting a prediction strategy: $RMSE$, $QAPE_{0.5}$ and $QAPE_{0.95}$

Monte Carlo simulation to estimate ex-ante accuracy measures (5000 iterations) - single *b*th iteration stages:

- generating real values of the variable of *Claim_Amount* under 6 assumed 'true' models $M_{(1)}, ..., M_{(6)}$,
- calculating characteristics $\theta_{(1)}$ and $\theta_{(2)}$ for every generated "true" *Claim_Amount*,
- predicting characteristics $\theta_{(1)}$ and $\theta_{(2)}$ under 6 assumed models $M_{(1)}, ..., M_{(6)}$,
- calculating measures of ex ante prediction accuracy: $RMSE$, $QAPE_{0.5}$, $QAPE_{0.95}$.

In result, we obtain accuracy measures matrix **A** with 36 rows, as we involve 2 characteristics, 6 models and 3 accuracy measures.

VOTING results

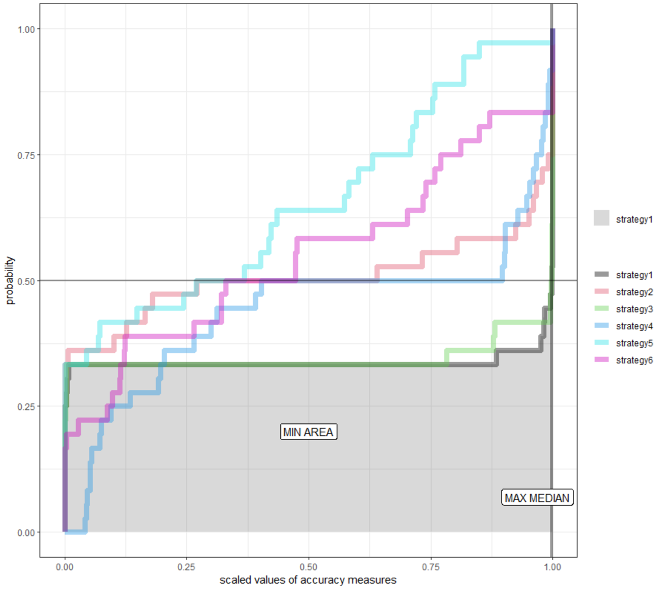
Four voting systems - the initiating step is to perform some transformation of the matrix **A** to the voting matrix **W** and then pointing the winner according to presented algorithms.

Values of selection criteria for four voting systems

Prediction strategy	FPTP	Positional	Evaluative	ECDF AUC
strategy 1	**14**	**5**	**0.999**	**0.338**
strategy 2	5	4	0.454	0.506
strategy 3	7	4	0.997	0.346
strategy 4	3	3	0.650	0.439
strategy 5	1	3	0.318	0.655
strategy 6	6	3	0.401	0.550

(a) sum of votes (the higher the better), (b) median of ranks (the higher the better), (c) median of scaled accuracy measures (the higher the better), (d) ECDF AUC (the smaller the better)

Automobile portfolio example: VOTING with ECDF AUC



Summing up, **strategy 1** is the winner in all voting systems. This means that the joint prediction of $\theta_{(1)}$ and $\theta_{(2)}$ should be performed using the **GLM-Gamma model**.

The prediction for 2011 amounts to $\hat{\theta}_{(1)} = 231,583.16$ and $\hat{\theta}_{(2)} = 3,141.69$.

Obtained assessments for the 'true' model GG

Accuracy Measure	$\theta_{(1)}$	$\theta_{(2)}$
RMSE	30,409.52	916.26
$QAPE_{0.5}$	20,337.75	829.49
$QAPE_{0.95}$	60,198.47	1,481.50

WASP METHOD

THANK YOU

References I

- Balinski, M. and Laraki, R. (2007). A theory of measuring, electing, and ranking. Proceedings of the National Academy of Sciences, 104(21):8720–8725. Publisher: Proceedings of the National Academy of Sciences.
- Baltagi, B. H. (2021). Econometrics. Classroom Companion: Economics. Springer International Publishing, Cham.
- Bauer, E. and Kohavi, R. (1999). An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants. Machine Learning, 36(1):105–139.
- Baujard, A., Igersheim, H., Lebon, I., Gavrel, F., and Laslier, J.-F. (2014). Who's favored by evaluative voting? An experiment conducted during the 2012 French presidential election. Electoral Studies, 34:131–145.

- Boubeta, M., Lombardía, M. J., and Morales, D. (2016). Empirical best prediction under area-level poisson mixed models. Test, 25(3):548–569.
- Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. D., editors (2016). Handbook of computational social choice. Cambridge University Press, New York, NY, USA.
- Burka, D., Puppe, C., Szepesváry, L., and Tasnádi, A. (2022). Voting: A machine learning approach. European Journal of Operational Research, 299(3):1003–1017.
- Butar, F. B. and Lahiri, P. (2003). On measures of uncertainty of empirical Bayes small-area estimators. Journal of Statistical Planning and Inference, 112(1-2):63–76.

References III

- Carpenter, J. R., Goldstein, H., and Rasbash, J. (2003). A novel bootstrap procedure for assessing the relationship between class size and achievement. Journal of the Royal Statistical Society: Series C (Applied Statistics), 52(4):431–443.
- Cassel, C.-M., Särndal, C.-E., and Wretman, J. H. (1977). Foundations of Inference in Survey Sampling. Wiley-Interscience, New York, 1st edition edition.
- Chambers, R. and Chandra, H. (2013). A Random Effect Block Bootstrap for Clustered Data. Journal of Computational and Graphical Statistics, 22(2):452–470.
- Chernick, M. R. and LaBudde, R. A. (2014). An Introduction to Bootstrap Methods with Applications to R. Wiley, 1st edition edition.

- Chwila, A. and Żądło, T. (2022). On properties of empirical best predictors. Communications in Statistics - Simulation and Computation, 51(1):220–253.
- Daneshvar, N. H.-N., Masoudi-Sobhanzadeh, Y., and Omidi, Y. (2023). A voting-based machine learning approach for classifying biological and clinical datasets. BMC Bioinformatics, 24(1):140.
- Dang, Q. and Yuan, J. (2023). A Kalman filter-based prediction strategy for multiobjective multitasking optimization. Expert Systems with Applications, 213:119025.
- De Jong, P. and Heller, G. Z. (2008). Generalized linear models for insurance data. Cambridge University Press.
- Denuit, M. and Lang, S. (2004). Non-life rate-making with bayesian gams. Insurance: Mathematics and Economics, 35(3):627–647.

References V

- Fabrizi, E., Ferrante, M. R., and Trivisano, C. (2016). Bayesian Beta Regression Models for the Estimation of Poverty and Inequality Parameters in Small Areas. In Pratesi, M., editor, Analysis of Poverty Data by Small Area Estimation, pages 299–314. John Wiley & Sons, Ltd.
- Fauzi, R. R. and Maesono, Y. (2023). Statistical Inference Based on Kernel Distribution Function Estimators. SpringerBriefs in Statistics. Springer Nature Singapore, Singapore.
- Felsenthal, D. S. (2012). Review of Paradoxes Afflicting Procedures for Electing a Single Candidate. In Felsenthal, D. S. and Machover, M., editors, Electoral Systems: Paradoxes, Assumptions, and Procedures, pages 19–91. Springer, Berlin, Heidelberg.

References VI

- Felsenthal, D. S. and Machover, M., editors (2012). Electoral Systems: Paradoxes, Assumptions, and Procedures. Studies in Choice and Welfare. Springer, Berlin, Heidelberg.
- González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D., and Santamaría, L. (2007). Estimation of the mean squared error of predictors of small area linear parameters under a logistic mixed model. Computational Statistics & Data Analysis, 51:2720–2733.
- González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D., and Santamaría, L. (2008). Bootstrap mean squared error of small-area eblup. Journal of Statistical Computation and Simulation, 78:443–462.
- Green-Armytage, J., Tideman, T. N., and Cosman, R. (2016). Statistical evaluation of voting rules. Social Choice and Welfare, 46(1):183–212.

References VII

- Greene, W. (2012). Econometric analysis. Pearson series in economics. Prentice Hall, Boston Munich, 7. ed edition.
- Gunawan, F. E., Budiman, A. S., Pardamean, B., Juana, E., Romeli, S., Cenggoro, T. W., Purwandari, K., Hidayat, A. A., Redi, A. A., and Asrol, M. (2023). Multivariate Time Series Deep Learning for Joint Prediction of Temperature and Relative Humidity in a Closed Space. Procedia Computer Science, 227:1046–1053.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning. Springer Series in Statistics. Springer, New York, NY.
- Henderson, C. R. (1950). Estimation of genetic parameters (Abstract). Annals of Mathematical Statistics, 21:309–310.
- Hobza, T. and Morales, D. (2016). Empirical Best Prediction Under Unit-Level Logit Mixed Models. Journal of Official Statistics, 32(3):661–692.

References VIII

- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R, volume 103 of Springer Texts in Statistics. Springer, New York, NY.
- Jiang, J. (2007). Linear and generalized linear mixed models and their applications. Springer series in statistics. Springer, New York.
- Kemm, F. (2023). Why Majority Judgement is not yet the solution for political elections, but can help finding it. arXiv:2302.10858.
- Levy, H. (1990). Stochastic Dominance. In Eatwell, J., Milgate, M., and Newman, P., editors, Utility and Probability, pages 251–254. Palgrave Macmillan UK, London.
- Matthews, S. and Hartman, B. (2022). Machine learning in ratemaking, an application in commercial auto insurance. Risks, 10(4):80.

References IX

- Molina, I. and Rao, J. N. K. (2010). Small area estimation of poverty indicators. Canadian Journal of Statistics, 38(3):369–385.
- Royall, R. M. (1976). The Linear Least-Squares Prediction Approach to Two-Stage Sampling. Journal of the American Statistical Association, 71(355):657–664.
- Thai, H.-T., Mentré, F., Holford, N. H., Veyrat-Follet, C., and Comets, E. (2013). A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed effects models. Pharmaceutical Statistics, 12(3):129–140.
- Valliant, R., Dorfman, A. H., and Royall, R. M. (2000). Finite Population Sampling and Inference: A Prediction Approach. Wiley-Interscience, New York.

References X

- van Setten, M., Veenstra, M., Nijholt, A., and van Dijk, B. (2004). Case-Based Reasoning as a Prediction Strategy for Hybrid Recommender Systems. In Favela, J., Menasalvas, E., and Chávez, E., editors, Advances in Web Intelligence, pages 13–22, Berlin, Heidelberg. Springer.
- Vovk, V. (2008). Leading strategies in competitive on-line prediction. Theoretical Computer Science, 405(3):285–296.
- Wolny-Dominiak, A. and Żądło, T. (2022). On bootstrap estimators of some prediction accuracy measures of loss reserves in a non-life insurance company. Communications in Statistics - Simulation and Computation, 51(8):4225–4240.
- Żądło, T. (2013). On parametric bootstrap and alternatives of MSE. In Vojáčková, H., editor, Proceedings of 31st International Conference Mathematical Methods in Economics 2013, pages 1081–1086. The College of Polytechnics Jihlava, Jihlava.

References XI

- Żądło, T. (2020). On Accuracy Estimation Using Parametric Bootstrap in small Area Prediction Problems. Journal of Official Statistics, 36(2):435–458.

APPENDIX

Accuracy measures matrix **A**- the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	30,409.52	92,010.24	30,525.97	67,948.94	75,471.81	91,967.10
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	33,892.15	82,222.11	33,973.64	79,599.19	78,954.10	82,234.57
$M_1, \theta_2, RMSE$	916.26	548.11	916.13	413.71	613.53	508.82
...	...	...	...	...	...	...
$M_6, \theta_2, RMSE$	1,061.72	350.82	1,062.28	360.62	531.79	325.74
$M_1, \theta_1, QAPE_{0.5}$	20,337.75	86,634.24	20,583.34	60,212.39	62,486.63	87,039.66
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.5}$	23,137.12	74,980.98	23,372.24	72,309.64	67,400.49	74,869.13
$M_1, \theta_2, QAPE_{0.5}$	829.49	388.74	828.17	280.48	415.84	351.43
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.5}$	999.68	228.70	998.64	239.17	356.18	211.63
$M_1, \theta_1, QAPE_{0.95}$	60,198.47	135,643.60	60,587.21	112,127.20	132,336.46	133,635.47
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.95}$	66,258.19	129,184.87	66,403.74	126,818.35	136,122.08	127,664.58
$M_1, \theta_2 QAPE_{0.95}$	1,481.50	1,051.47	1,479.56	808.35	1,201.03	984.06
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.95}$	1,603.75	693.73	1,606.65	716.80	1,047.93	647.50

First-past-the-post voting matrix $\mathbf{W_1}$ - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	1	0	0	0	0	0
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	1	0	0	0	0	0
$M_1, \theta_2, RMSE$	0	0	0	1	0	0
...	...	...	...	...	...	...
$M_6, \theta_2, RMSE$	0	0	0	0	0	1
$M_1, \theta_1, QAPE_{0.5}$	1	0	0	0	0	0
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.5}$	1	0	0	0	0	0
$M_1, \theta_2, QAPE_{0.5}$	0	0	0	1	0	0
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.5}$	0	0	0	0	0	1
$M_1, \theta_1, QAPE_{0.95}$	1	0	0	0	0	0
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.95}$	0	0	1	0	0	0
$M_1, \theta_2 QAPE_{0.95}$	1	0	0	0	0	0
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.95}$	0	0	0	0	0	1
sum	14	5	7	3	1	6

Positional voting matrix $\mathbf{W_2}$ - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	6	1	5	4	3	2
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	6	2	5	3	4	1
$M_1, \theta_2, RMSE$	1	4	2	6	3	5
...	...	...	...	...	...	...
$M_6, \theta_2, RMSE$	2	5	1	4	3	6
$M_1, \theta_1, QAPE_{0.5}$	6	2	5	4	3	1
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.5}$	6	1	5	3	4	2
$M_1, \theta_2, QAPE_{0.5}$	1	4	2	6	3	5
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.5}$	1	5	2	4	3	6
$M_1, \theta_1, QAPE_{0.95}$	6	1	5	4	3	2
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.95}$	6	2	5	4	1	3
$M_1, \theta_2 QAPE_{0.95}$	1	4	2	6	3	5
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.95}$	2	5	1	4	3	6
median	5	4	4	3	3	3

Scaled voting matrix $\mathbf{W}_3$ - the short form

voter	strat 1	strat 2	strat 3	strat 4	strat 5	strat 6
$M_1, \theta_1, RMSE$	1.000	0.000	0.998	0.391	0.268	0.001
...	...	...	...	...	...	...
$M_6, \theta_1, RMSE$	1.000	0.000	0.998	0.055	0.068	0.000
$M_1, \theta_2, RMSE$	0.000	0.733	0.000	1.000	0.602	0.811
...	...	...	...	...	...	...
$M_6, \theta_2, RMSE$	0.001	0.966	0.000	0.953	0.720	1.000
$M_1, \theta_1, QAPE_{0.5}$	1.000	0.006	0.996	0.402	0.368	0.000
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.5}$	1.000	0.000	0.995	0.052	0.146	0.002
$M_1, \theta_2, QAPE_{0.5}$	0.000	0.803	0.002	1.000	0.753	0.871
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.5}$	0.000	0.978	0.001	0.965	0.817	1.000
$M_1, \theta_1, QAPE_{0.95}$	1.000	0.000	0.995	0.312	0.044	0.027
...	...	...	...	...	...	...
$M_6, \theta_1, QAPE_{0.95}$	1.000	0.099	0.998	0.133	0.000	0.121
$M_1, \theta_2 QAPE_{0.95}$	0.000	0.639	0.003	1.000	0.417	0.739
...	...	...	...	...	...	...
$M_6, \theta_2, QAPE_{0.95}$	0.003	0.952	0.000	0.928	0.583	1.000
median	0.999	0.454	0.997	0.650	0.318	0.401

Automobile portfolio example