# EVALUATION AND EXTENSION OF INDUCTIVE VENN-ABERS PREDICTIVE DISTRIBUTION

Ilia Nouretdinov, James Gammerman & Daljit Rehal

## ABSTRACT


- In this work we provide further development of Inductive Venn-Abers Predictive Distribution (IVAPD) for regression.
- The main contribution of this work is a new algorithm that allows combinations of underlying methods.
- We also review several evaluation metrics for the results.


## ALGORITHM

INPUT: proper training set
$T_P = \{(x_{-1}, y_{-1}), \ldots, (x_{-r}, y_{-r})\}$.
INPUT: cal. set $T_C = \{(x_1, y_1), \ldots, (x_h, y_h)\}$
INPUT: testing example $x_{h+1}$.
INPUT: underlying predictors $P^1$ and $P^2$
**for** $i := 1, \ldots, r$ **do**
  $s^1_{-i} := P^1(x_i, T \setminus \{(x_{-i}, y_{-i})\})$
  $s^2_{-i} := P^2(x_i, T \setminus \{(x_{-i}, y_{-i})\})$
**end for**
apply bivariate isotonic optimisation: find $(g_{-1}, \ldots, g_{-r})$ s.t. $\sum_{i=1}^{r}(g_{-i} - y_{-i})^2 \to \min$
$(s^1_{-i} \le s^1_{-j}) \& (s^2_{-i} \le s^2_{-j}) \Rightarrow (g_{-i} \le g_{-j})$
**for** $i := 1, \ldots, h+1$ **do**
  $s_i := P(x_i, T_P)$
  find $s_{-j}$ which is the closest to $s_i$
  $g_i := g_{-j}$
**end for**
let $A := \{i = 1, \ldots, h : g_i = g_{h+1}\}$
let $\hat{Y} := \{y_i : i \in A\}$
OUTPUT $(q = 0, 1)$:

$$\hat{P}_q\{y_{h+1} \le t\} := \frac{|\{\hat{y} \in \hat{Y} : \hat{y} \le t\}| + q}{|A| + 1}$$

## DATA SETS

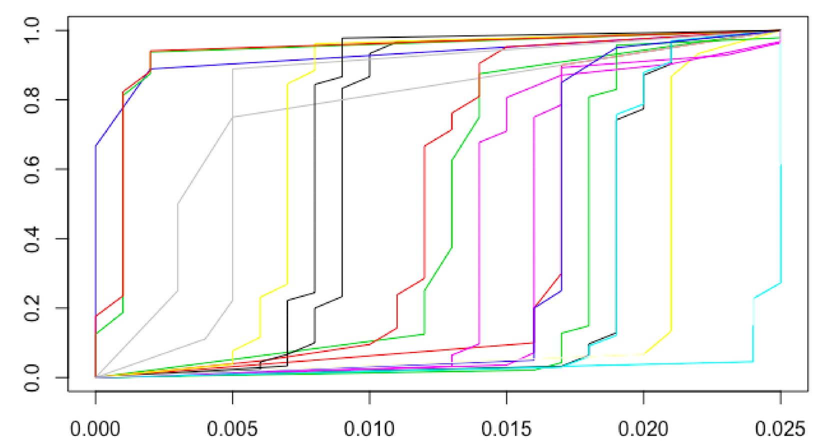
1) UCI public dataset on Propulsion Plants Maintenance (NPP), two labels:
**a) Compressor degradation coefficient;**
**b) Turbine degradation coefficient.**
2) UCI public dataset on household power consumption (ECP), one label:
**evening consumption (at 18:00).**

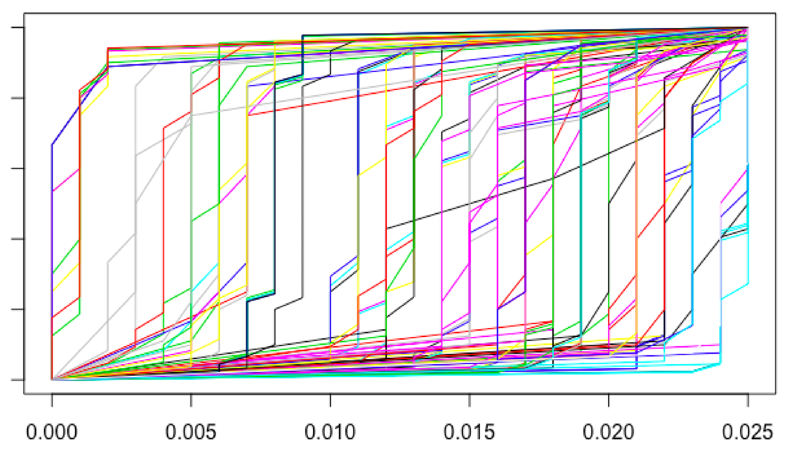
## EVALUATION METRICS

1. **(C)** Continuous Ranked Probability Score (requires the ground truth).
2. **(W)** Width of prediction interval for probability $(1 - \varepsilon)$.
3. **(V)** Variance of distribution.
4. **(P)** Average diff. between CDFs of $P_0$ and $P_1$.

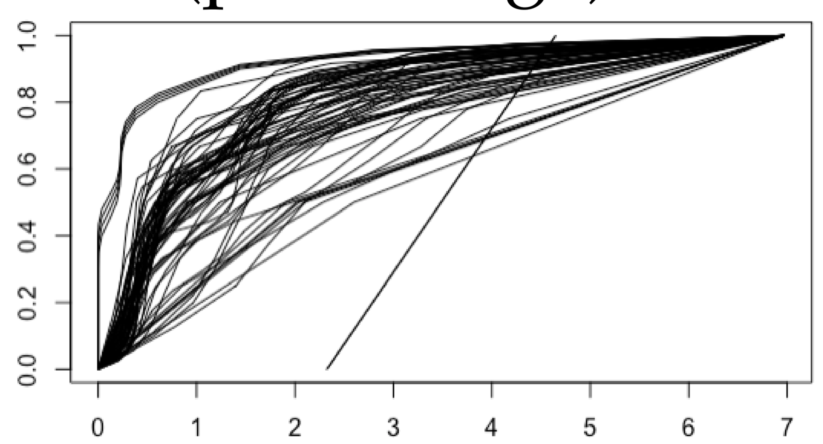
- We want to see which of W, V, and P best agrees with C.

## JOINT PLOTS OF PREDICTED CDFS





NPP: x-axis is 1- turbine degradation coefficient (percentage).



ECP: x-axis is energy consumption (kWh).

## $k$NN UNDERLYING METHOD

| Data | feat. | nei. | C | V | W $\varepsilon = 0.25$ | W $\varepsilon = 0.5$ | W $\varepsilon = 0.75$ | P |
|---|---|---|---|---|---|---|---|---|
| NPP1 | 5 | 5 | .00117 | .00225 | .00424 | .00213 | .000582 | 0.0493 |
| NPP1 | 5 | 20 | .00130 | .00237 | .00484 | .00263 | .000809 | 0.0379 |
| NPP1 | 5 | 100 | .00143 | .00250 | .00577 | .00338 | .00110 | 0.0189 |
| NPP1 | – | 5 | **.000964** | **.00207** | **.00339** | **.00160** | **.000369** | 0.0559 |
| NPP1 | – | 20 | .00124 | .00232 | .00454 | .00239 | .000744 | 0.0442 |
| NPP1 | – | 100 | .00143 | .00249 | .00579 | .00335 | .00107 | **0.0172** |
| NPP2 | 5 | 5 | .00183 | .00636 | .00773 | .00321 | .00116 | 0.108 |
| NPP2 | 5 | 20 | .00237 | **.00635** | .00964 | .00450 | .00166 | 0.0880 |
| NPP2 | 5 | 100 | .00379 | .00729 | 0.0153 | .00841 | .00329 | 0.0507 |
| NPP2 | – | 5 | **.00179** | .00642 | **.00757** | **.00283** | **.000954** | 0.112 |
| NPP2 | – | 20 | .00239 | .00648 | .00993 | .00449 | .00158 | 0.0925 |
| NPP2 | – | 100 | .00415 | .00747 | 0.170 | .00989 | .00399 | **0.0294** |
| ECP | 5 | 5 | 0.607 | 1.780 | 2.418 | 0.973 | 0.425 | 0.166 |
| ECP | 5 | 20 | 0.624 | 1.751 | 2.365 | 0.923 | 0.398 | 0.150 |
| ECP | 5 | 100 | 0.591 | **1.657** | **2.211** | **0.872** | **0.380** | **0.123** |
| ECP | 20 | 5 | 0.622 | 1.797 | 2.496 | 1.009 | 0.437 | 0.168 |
| ECP | 20 | 20 | 0.608 | 1.764 | 2.380 | 0.969 | 0.422 | 0.157 |
| ECP | 20 | 100 | 0.591 | 1.675 | 2.287 | 0.934 | 0.436 | 0.137 |
| ECP | – | 5 | 0.613 | 1.825 | 2.485 | 0.995 | 0.450 | 0.178 |
| ECP | – | 20 | 0.604 | 1.752 | 2.343 | 0.942 | 0.402 | 0.153 |
| ECP | – | 100 | **0.586** | 1.692 | 2.276 | 0.917 | 0.419 | 0.136 |

- W shows the best agreement with C.

## COMPARISON OF SINGLE $k$-VALUE vs TWO COMBINED

| Data | feat. | nei. | C | V | W $\varepsilon = 0.25$ | W $\varepsilon = 0.5$ | W $\varepsilon = 0.75$ | P |
|---|---|---|---|---|---|---|---|---|
| NPP2 | – | 5 | .00318 | .00793 | 0.0137 | .00679 | .00305 | 0.208 |
| NPP2 | – | 10 | .00398 | .00789 | 0.0150 | .00708 | **.00251** | 0.130 |
| NPP2 | – | 20 | .00458 | .00814 | 0.0169 | .00819 | .00267 | 0.0778 |
| NPP2 | – | 100 | .00442 | .00783 | 0.0170 | .00824 | .00301 | **0.0274** |
| NPP2 | – | 5,10 | .00306 | .00835 | 0.0146 | .00735 | .00364 | 0.247 |
| NPP2 | – | 5,20 | .00301 | .00831 | 0.0150 | .00728 | .00332 | 0.232 |
| NPP2 | – | 10,20 | .00401 | .00811 | 0.0155 | .00745 | .00271 | 0.151 |
| NPP2 | – | 5,100 | **.00298** | **.00744** | **0.0131** | **.00655** | .00263 | 0.144 |
| NPP2 | – | 10,100 | .00389 | .00773 | 0.0153 | 0.t00759 | .00279 | 0.0973 |
| NPP2 | – | 20,100 | .00448 | .00798 | 0.0170 | .00806 | .00273 | 0.0485 |

- Typically, a combination works better.

**Acknowledgements.** AstraZeneca, Mitie, RHUL, Centrica.

**Reference**

[1] Nouretdinov, I., Volkhonskiy, D., Lim, P., Toccaceli, P., Gammerman, A., 2018. Inductive Venn-Abers Predictive Distribution. Proceedings of Conformal Prediction with Applications (COPA 2018). Vol. 91, p. 1–22.