ROC movies, UROC curves, and CPA

E.-M. Walz[1,2] (eva-maria.walz@kit.edu) & T. Gneiting[2,1]

1) Institute of Stochastics, Karlsruhe Institute of Technology
2) Heidelberg Institute for Theoretical Studies

WAVES TO WEATHER

Motivation and Objective

Receiver operating characteristic (ROC) curves and associated area under the cure (AUC) measure constitude powerful tools for assessing predictive abilities of features, markers and tests in binary classification problems [2]. Despite its immense popularity, ROC analysis has been subject to a fundamental restriction, in that it applies to dichotomous outcomes only.

ROC movies, universal ROC (UROC) curves and an asymmetric coefficient of predictive ability (CPA) measure provide a natural generalization, that applies to any type of ordinal or real-valued outcome.

Receiver operating characteristic (ROC) movies and universal ROC (UROC) curves

Consider bivariate data of the form

$$(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R} \times \mathbb{R}, \tag{1}$$

where x_i is a real-valued point forecast, regression output, feature, marker, or covariate value, and y_i is a real-valued outcome, for $i = 1, \ldots, n$. Let

$$z_1 < \cdots < z_m$$

denote the $m \leq n$ distinct order statistics of $y_1, \ldots, y_n$ and

$$n_c = \sum_{i=1}^{n} \mathbb{1}\{y_i = z_c\}$$

for $c = 1, \ldots, m$. These groups of instances are referred to as classes.

We sucessively transform the original data into binary data of the form

$$(x_1, \mathbb{1}\{y_1 \geq z_{c+1}\}), \ldots, (x_n, \mathbb{1}\{y_n \geq z_{c+1}\}) \in \mathbb{R} \times \{0, 1\}, \tag{2}$$

where $c = 1, \ldots, m-1$. On these $m-1$ derived binary classification problems tools of classical ROC analysis apply.

Definition 1 For data of the form (1), the *ROC movie* is the sequence $(\text{ROC}_c)_{c=1,\ldots,m-1}$ of the ROC curves for the induced binary data in (2).

Definition 2 For data of the form (1), the *universal ROC* (UROC) *curve* is the curve associated with the function

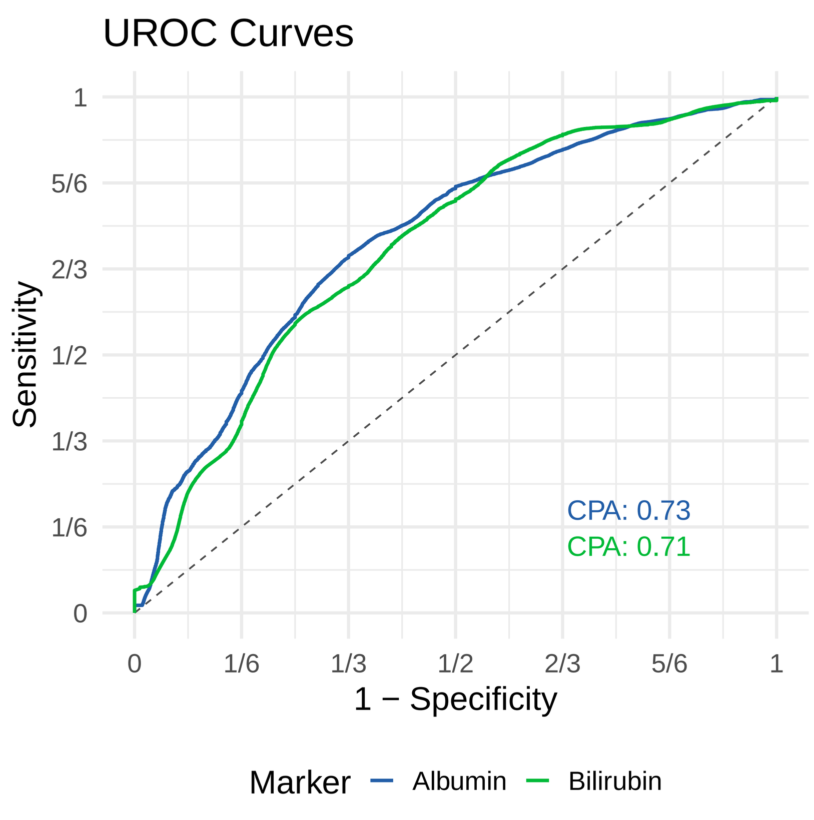
$$\sum_{c=1}^{m-1} w_c \, \text{ROC}_c \tag{3}$$

on the unit interval, with weights

$$w_c = \left(\sum_{i=1}^{c} n_i \sum_{i=c+1}^{m} n_i \right) \Bigg/ \left(\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (j-i) n_i n_j \right) \tag{4}$$

for $c = 1, \ldots, m-1$.

Example - UROC curve



- Survival data from Mayo Clinical trial on primary biliary cirrhosis [1]
- Two biochemical markers, serum albumin and serum bilirubin, as predictor for patient survival (in days)
- No need to artificially pick a threshold for survival time.
- Corresponding ROC movies in [3] or:

References

[1] Dickson, E. R., Grambsch, P. M., Fleming, T. R., Fischer, L. D., Langworthy, A. (1989). Prognosis in primary biliary cirrhosis: Model for decision making. *Hepatology* **10**, 1–7.

[2] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognit. Lett.* **27**, 861–874.

[3] Gneiting, T., Walz, E.-M. (2020). Receiver operating characteristic (ROC) movies, universal ROC (UROC) curves, and coefficient of predictive ability (CPA). http://arxiv.org/abs/1912.01956

[4] Rasp, S., Dueben, P.D., Scher, S., Weyn, J.A., Mouatadid, S. and Thuerey N. (2020). WeatherBench: A benchmark dataset for data-driven weather forecasting, https://arxiv.org/abs/2002.00469.

[5] Rosset, S., Perlich, C., Zadrozny, B. (2005). "Ranking-based evaluation of regression models", in *Proc. Fifth IEEE Int. Conf. Data Min. (ICDM'05)* (IEEE).

Coefficient of predictive ability (CPA)

Definition 3 For data of the form (1) and weights $w_1, \ldots, w_{m-1}$ as in (4), the *coefficient of predictive ability* (CPA) is defined as

$$\text{CPA} = \sum_{c=1}^{m-1} w_c \, \text{AUC}_c. \tag{5}$$

In words, CPA equals the area under the UROC curve.

Properties of CPA:

1. Rewriting data as

$$(x_{11}, z_1), \ldots, (x_{1n_1}, z_1), \ldots, (x_{m1}, z_m), \ldots, (x_{mn_m}, z_m) \in \mathbb{R} \times \mathbb{R}, \tag{6}$$

allows to reformulate CPA as a weighted probability of concordance, with weights that grow linearly in the class based distance between outcomes. Let

$$s(x, x') = \mathbb{1}\{x < x'\} + \frac{1}{2}\mathbb{1}\{x = x'\}, \tag{7}$$

where $x, x' \in \mathbb{R}$.

Theorem 1 For data of the form (6),

$$\text{CPA} = \frac{\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \sum_{k=1}^{n_i} \sum_{l=1}^{n_j} (j-i) \, s(x_{ik}, x_{jl})}{\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (j-i) \, n_i n_j}. \tag{8}$$

2. Representation of CPA in terms of covariances between classes and mid ranks of feature and outcome. Mid rank method handles ties by assigning the arithmetic average of the ranks involved.

Theorem 2 Let the random vector (X, Y) be drawn from the empirical distribution of the data in (1) or (6). Then

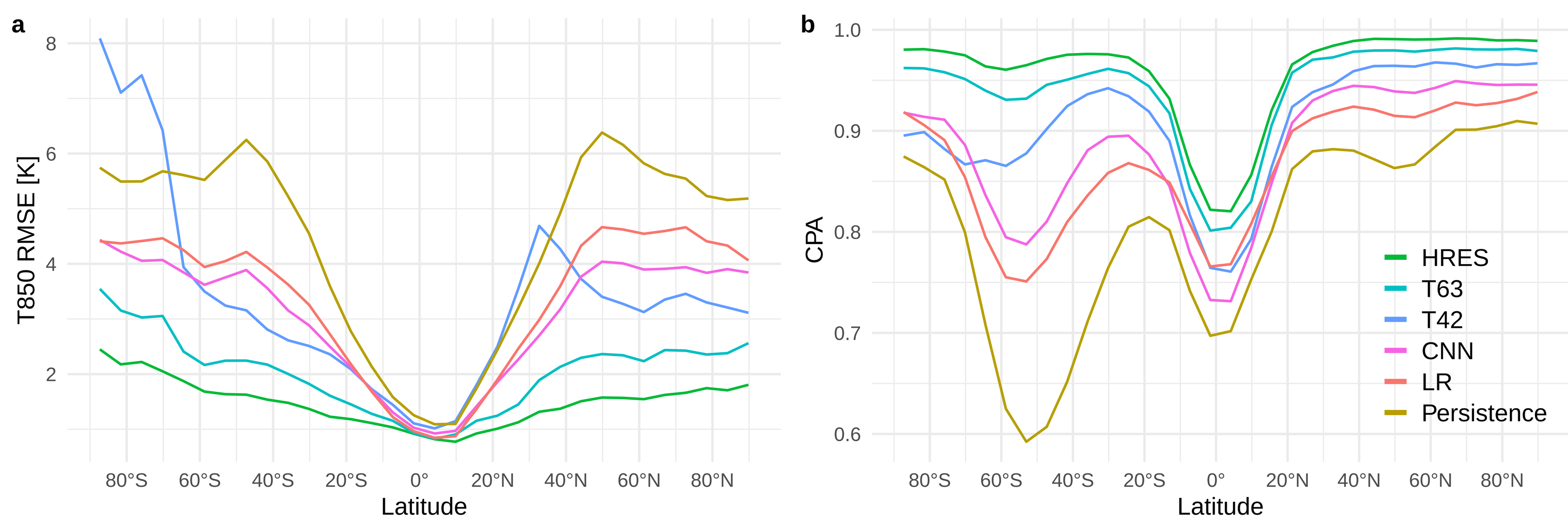
$$\text{CPA} = \frac{1}{2} \left(\frac{\text{cov}(\text{cl}(Y), \overline{\text{rk}}(X))}{\text{cov}(\text{cl}(Y), \overline{\text{rk}}(Y))} + 1 \right). \tag{9}$$

3. In case there are no ties in the data, CPA relates linearly to Spearman's rank correlation coefficient ρ_S [5].

Theorem 3 In the case of no ties,

$$\text{CPA} = \frac{1}{2}(\rho_S + 1). \tag{10}$$

Example – CPA



- WeatherBench [4] three days ahead forecast of 850 hPa temperature in 2017 and 2018 at different latitudes in terms of (a) RMSE and (b) CPA.
- RMSE and CPA reveal orthogonal facets of predictive performance.
- An example: Low CPA in the tropics indicates poor performance, well in line with meteorological expertise, in contrast to low RMSE suggesting superior performance in this region.

Conclusion

- For binary outcomes, ROC movies, UROC curves, and CPA reduce to ROC curves and AUC.
- They are straightforward to interpret in the sense *the larger the better*.
- CPA attains values between 0 and 1. For a perfect feature CPA = 1 and for a feature independent of the outcome, CPA = $\frac{1}{2}$.
- CPA can be represented in terms of covariances. If the outcomes are pairwise distinct, then CPA relates linearly to Spearman's correlation coefficient (3). The numerical value of CPA admits an interpretation as a weighted probability.
- ROC movies, UROC curves and CPA are purely rank based and therefore, invariant unter stricly increasing transformations

ROC movies, UROC curves and CPA nest and generalize ROC analysis, retain its attractive properties, and free researches from the need to artificially binarize real-valued outcomes.

contact: eva-maria.walz@kit.edu