

Raw categorical outcomes

Confusion matrix

		Predicted condition	
		Positive (PP)	Negative (PN)
		Positive (P)	Negative (N)
Observed condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

(a)

Deterministic model evaluation

Categorical metrics

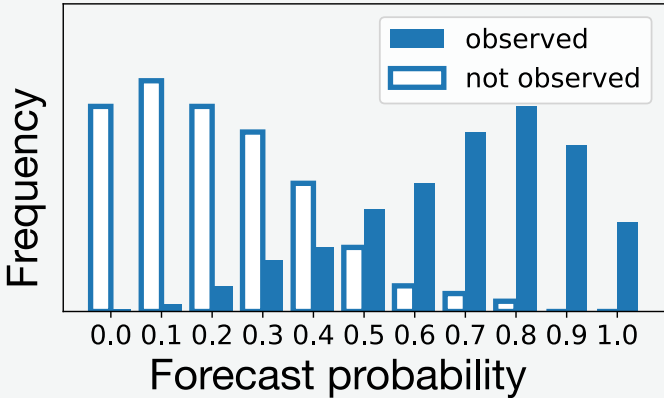
Metric	Definition	Range	Perfect
Frequency Bias Index (FBI)	$\frac{PP}{P} = \frac{TP + FP}{TP + FN}$	[0,inf]	1
True Positive Rate / Prob. of Detection (POD)	$\frac{TP}{P} = \frac{TP}{TP + FN}$	[0,1]	1
False Positive Rate (FPR)	$\frac{FP}{N} = \frac{FP}{FP + TN}$	[0,1]	0
Precision (Prec)	$\frac{TP}{PP} = \frac{TP}{TP + FP}$	[0,1]	1
Matthews Correlation Coefficient (MCC)	$\frac{TP \cdot TN - FP \cdot FN}{\sqrt{PP \cdot P \cdot N \cdot PN}}$	[-1,1]	1

(b)

Probabilistic model evaluation

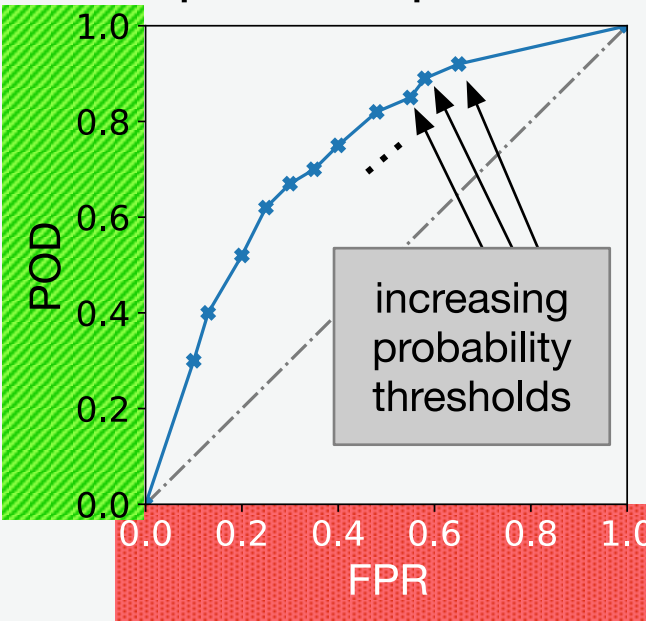
(i) Distributions of forecast probabilities

Discrimination diagram



(ii) Threshold-dependent performance

ROC curve



(c)