

Predicting Enzyme-Compound Interactions for Enzyme-Catalysed Reactions

Liam Brydon-Brown^{1,2*}, Gillian Dobbie¹,
Jörg Simon Wicker^{1,2}, Katerina Taškova¹

¹School of Computer Science, University of Auckland, Auckland, New Zealand.

²enviPath Limited, Auckland, New Zealand.

*Corresponding author(s). E-mail(s): lbry121@aucklanduni.ac.nz;

1 ECLIPSE Performance By Level One EC

Here we give the full tables of the ECLIPSE performance. EMap is shown in Table 1 and BBD is shown in Table 2.

2 Product Prediction By Level One EC

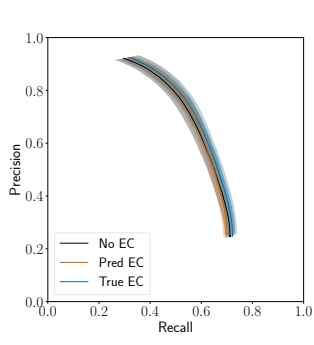
Here we give the precision-recall (PR) curves for product prediction grouped by level one EC across all datasets. Figure 1 shows the EMap plots. Figure 2 shows the BBD plots. Figure 3 shows the Soil plots.

EC	Metric	BECPred	F-ECLIPSE	H-ECLIPSE
1	F1	86.67% $\pm$ 0.91	90.25% $\pm$ 0.57	91.02% $\pm$ 0.50
	Precision	86.67% $\pm$ 0.91	90.56% $\pm$ 0.60	91.02% $\pm$ 0.50
	Recall	86.67% $\pm$ 0.91	89.97% $\pm$ 1.34	91.02% $\pm$ 0.50
2	F1	96.22% $\pm$ 0.35	97.22% $\pm$ 0.20	97.27% $\pm$ 0.22
	Precision	96.22% $\pm$ 0.35	97.22% $\pm$ 0.20	97.27% $\pm$ 0.22
	Recall	96.22% $\pm$ 0.35	97.22% $\pm$ 0.20	97.27% $\pm$ 0.22
3	F1	84.59% $\pm$ 0.67	88.61% $\pm$ 0.45	88.47% $\pm$ 0.44
	Precision	84.59% $\pm$ 0.67	88.61% $\pm$ 0.45	88.47% $\pm$ 0.44
	Recall	84.59% $\pm$ 0.67	88.61% $\pm$ 0.45	88.47% $\pm$ 0.44
4	F1	76.40% $\pm$ 1.48	80.51% $\pm$ 1.60	80.01% $\pm$ 1.36
	Precision	76.40% $\pm$ 1.48	80.51% $\pm$ 1.60	80.01% $\pm$ 1.36
	Recall	76.40% $\pm$ 1.48	80.51% $\pm$ 1.60	80.01% $\pm$ 1.36
5	F1	62.56% $\pm$ 4.02	64.70% $\pm$ 2.18	66.42% $\pm$ 3.14
	Precision	62.56% $\pm$ 4.02	64.70% $\pm$ 2.18	66.42% $\pm$ 3.14
	Recall	62.56% $\pm$ 4.02	64.70% $\pm$ 2.18	66.42% $\pm$ 3.14
6	F1	90.91% $\pm$ 1.19	91.70% $\pm$ 1.72	91.50% $\pm$ 1.36
	Precision	90.91% $\pm$ 1.19	91.70% $\pm$ 1.72	91.50% $\pm$ 1.36
	Recall	90.91% $\pm$ 1.19	91.70% $\pm$ 1.72	91.50% $\pm$ 1.36
7	F1	45.37% $\pm$ 8.85	50.13% $\pm$ 11.82	50.70% $\pm$ 12.31
	Precision	45.37% $\pm$ 8.85	50.13% $\pm$ 11.82	50.70% $\pm$ 12.31
	Recall	45.37% $\pm$ 8.85	50.13% $\pm$ 11.82	50.70% $\pm$ 12.31

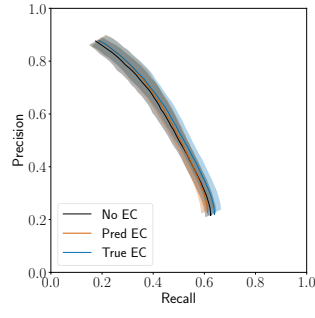
Table 1: Performance of BECPred, F-ECLIPSE and H-ECLIPSE on the ECMap dataset. The hierarchical F1, precision and recall are given, grouped by EC level one. The highest mean for each metric is bolded.

EC	Metric	BECPred	F-ECLIPSE	H-ECLIPSE
1	F1	59.72% $\pm$ 5.25	63.74% $\pm$ 3.27	66.94% $\pm$ 3.96
	Precision	59.72% $\pm$ 5.25	64.53% $\pm$ 2.70	67.71% $\pm$ 4.52
	Recall	59.72% $\pm$ 5.25	63.03% $\pm$ 4.08	66.26% $\pm$ 3.97
2	F1	29.22% $\pm$ 19.37	40.08% $\pm$ 20.33	35.18% $\pm$ 21.73
	Precision	29.22% $\pm$ 19.37	40.86% $\pm$ 20.99	35.74% $\pm$ 22.11
	Recall	29.22% $\pm$ 19.37	39.48% $\pm$ 19.99	34.73% $\pm$ 21.52
3	F1	52.99% $\pm$ 4.29	57.02% $\pm$ 12.24	55.09% $\pm$ 12.04
	Precision	52.99% $\pm$ 4.29	57.03% $\pm$ 12.70	55.70% $\pm$ 11.80
	Recall	52.99% $\pm$ 4.29	57.10% $\pm$ 11.94	54.56% $\pm$ 12.31
4	F1	33.71% $\pm$ 11.66	47.30% $\pm$ 12.76	46.41% $\pm$ 14.24
	Precision	33.71% $\pm$ 11.66	47.30% $\pm$ 12.76	46.41% $\pm$ 14.24
	Recall	33.71% $\pm$ 11.66	47.30% $\pm$ 12.76	46.41% $\pm$ 14.24
5	F1	21.50% $\pm$ 7.43	48.83% $\pm$ 21.90	45.00% $\pm$ 19.89
	Precision	21.50% $\pm$ 7.43	48.83% $\pm$ 21.90	45.00% $\pm$ 19.89
	Recall	21.50% $\pm$ 7.43	48.83% $\pm$ 21.90	45.00% $\pm$ 19.89

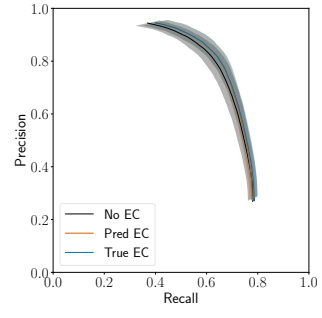
Table 2: Performance of BECPred, F-ECLIPSE and H-ECLIPSE on the BBD dataset. The hierarchical F1, precision and recall are given, grouped by EC level one. The configuration of Algorithm 1 removed EC classes six and seven as they were too small. The highest mean for each metric is bolded.



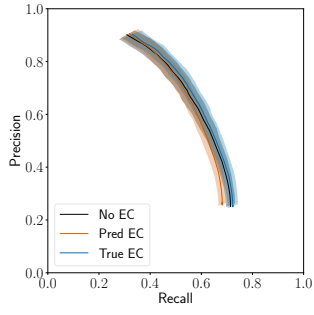
(a) ECMap PR curve overall



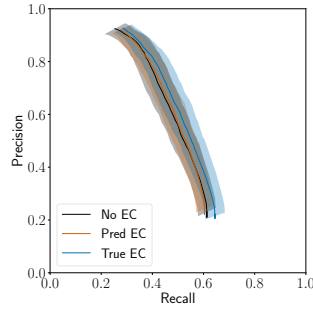
(b) ECMap PR curve on EC one



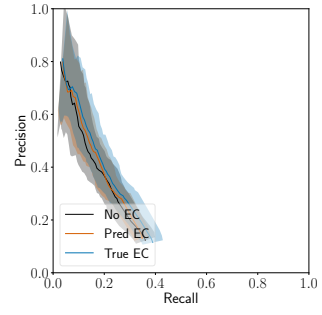
(c) ECMap PR curve on EC two



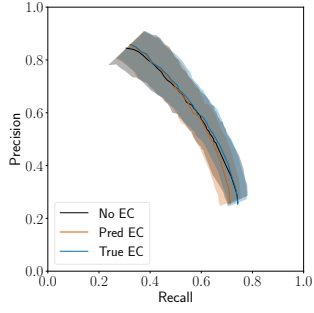
(d) ECMap PR curve on EC three



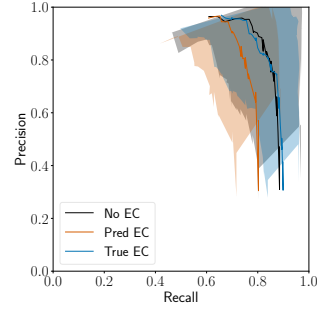
(e) ECMap PR curve on EC four



(f) ECMap PR curve on EC five



(g) ECMap PR curve on EC six



(h) ECMap PR curve on EC seven

Fig. 1: ECMap PR curves overall and by EC level one class

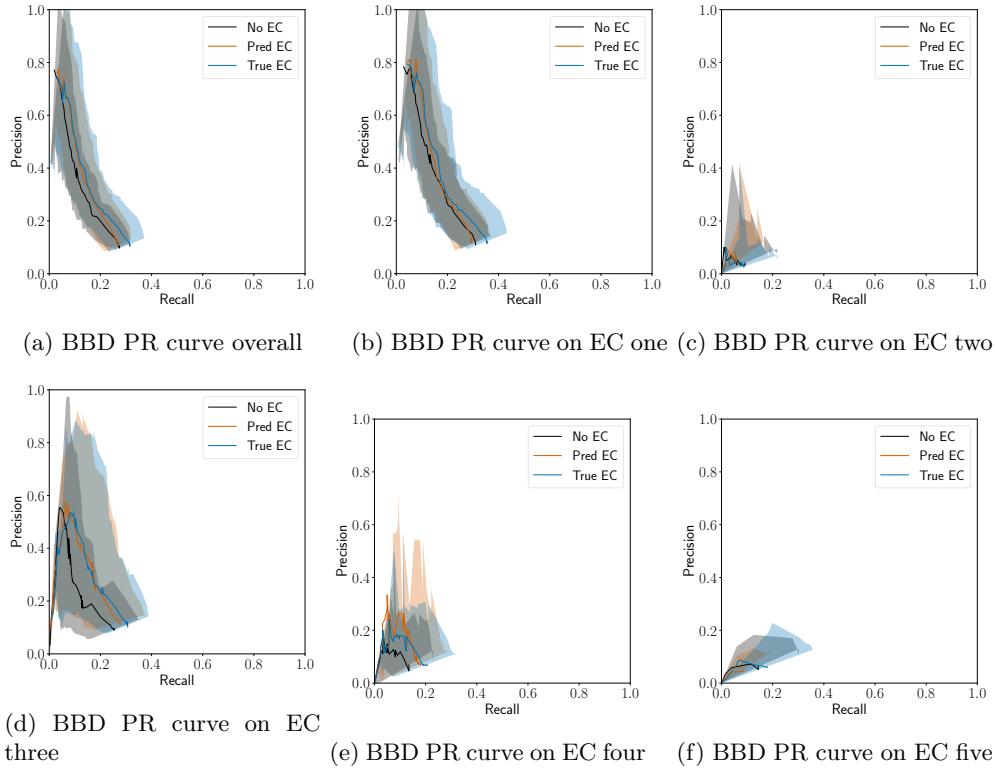
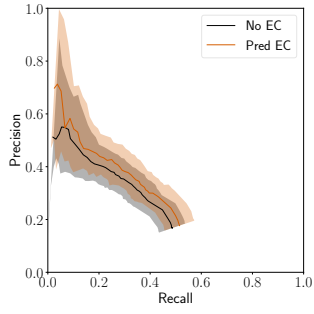
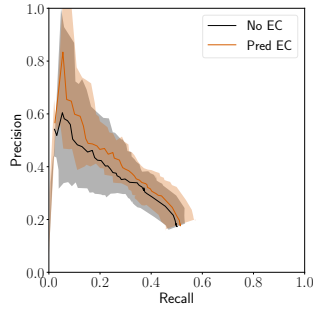


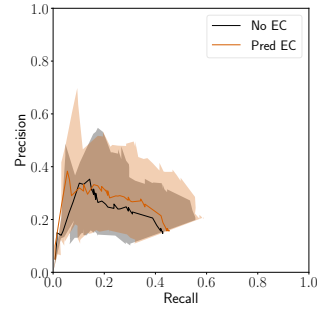
Fig. 2: BBD PR curves overall and by EC level one class



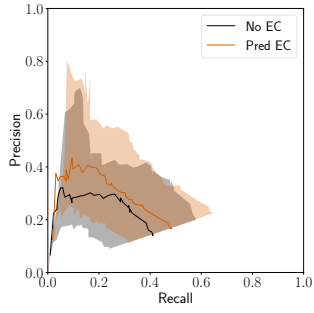
(a) Soil PR curve overall



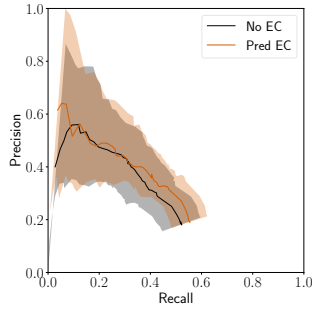
(b) Soil PR curve on EC one



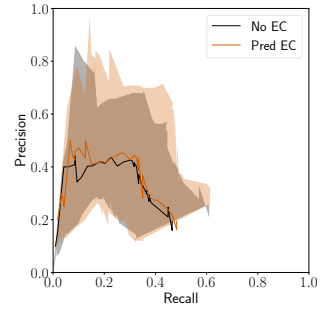
(c) Soil PR curve on EC two



(d) Soil PR curve on EC three



(e) Soil PR curve on EC four



(f) Soil PR curve on EC five

Fig. 3: Soil PR curves overall and by EC level one class