EVALUATION OF OVERTOPPING MODEL PERFORMANCE USING NOVEL EXPERIMENTAL DATA FROM INDUSTRIAL DESIGN PROJECTS

Steven Downie, Arup, steven.downie@arup.com
Jorge Vaz, Arup, jorge.vaz@arup.com
Eleni Anastasaki, Arup, eleni.anastasaki@arup.com

BACKGROUND
A critical requirement in successfully planning to mitigate wave overtopping is the ability to predict the frequency at which coastal defences will be overtopped.

Many empirical formulae have been developed to predict wave overtopping rates for specific structural typologies and hydrodynamic conditions. More recently, Machine Learning methods have been deployed in an effort to make models that generalize across a wide range of structures and environments. A critical enabling factor has been the compilation of systematically parameterized physical model data, culminating in the EurOtop extended database.

Practitioners now have the luxury of choosing between multiple, high-quality models. In addition, given the rapid advancement in usability of modern Machine Learning frameworks, training their own bespoke models is an increasingly realistic option. What is now needed is more information, showing how these models perform on unseen data, in a practical design context, in order to continue to refine guidance around their use.

RESEARCH OBJECTIVES
The objective of this study was to explore the behaviour of a range of overtopping models against physical model data not present in the EurOtop extended database. Models investigated include the most recent EurOtop ANN model (ANN2 – Zanuttigh et al. 2014), the Bayonet GPE model (Pullen et al. 2018) and a set of new models, developed by the authors and trained on the EurOtop database. The intent is not to rank or competitively assess model performance, rather it is to provide an understanding of variability between model predictions in real world scenarios and to consider how this can be managed by practitioners.

METHODS AND DATA
Experimental data (153 data points) have been collated from recent Arup projects, which involved physical modelling test series of a wide range of structural types:

- Complex geometry comprising a sloping armour with a double layer of Tetrapod armour and directional wave effects, including Typhoon conditions with substantial energy contribution from very long waves (3D)
- Seawall with alternative supportive structures against overtopping including set-back wall, rock berm, sloping rock armour (2D)
- Seawall at one and two levels, with and without a parapet wall and varying beach levels in front of the wall (2D)
- Sloping armour consisting of single layer Accropodes (2D/3D)
- Seawall covered with an armour made of Antifer units in double layer with a recurved parapet (2D).
- Seawall with a parapet crownwall; seawall with a set-back wall; seawall covered with a sloping rock armour (2D).

Test configurations have been parameterized using the EurOtop scheme and, while some of the physical models are relatively complex in comparison to those making up the EurOtop core training data, the resulting parameterizations are well represented within the parameter space of the EurOtop extended database.

Figure 1 shows 2D scatter plots produced using the t-SNE dimensionality reduction method, which provides a semi-quantitative, visual check on the proximity of data points in the original, high dimensional space. The top image shows the dimensionally reduced data for the EurOtop extended database, colored by structural type. The bottom figure shows a comparison including the Arup parameterizations, which indicates they are well encompassed by the EurOtop training data.

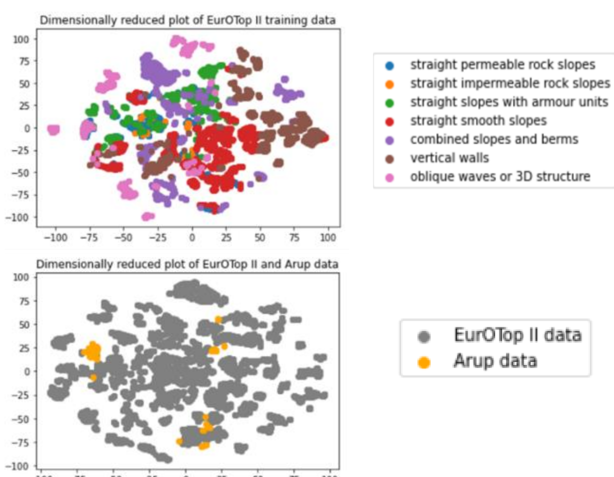

Figure 1 – Scatter plots of dimensionally reduced data points from the EurOTop extended database and the new experimental dataset presented in this study.

Alongside the ANN2 and Bayonet models, four new overtopping prediction models were trained by the authors including:

- A Gaussian Process regression model (GPR), similar in concept to the Bayonet model.

- Two Random Forrest regression models, with varying degrees of overfitting (RFR10 and RFR100).
- A Gradient Boosted regression model (GBR).

The models have been trained using the core data from the EurOTop extended database, excluding tests with overtopping rates measured at less than10^{-6} m^3/m.hr. A randomized subset, accounting for approximately 5% of the remaining data was held back to test model performance. Figure 2 shows predicted vs. measured values for both training and test data for all four models.

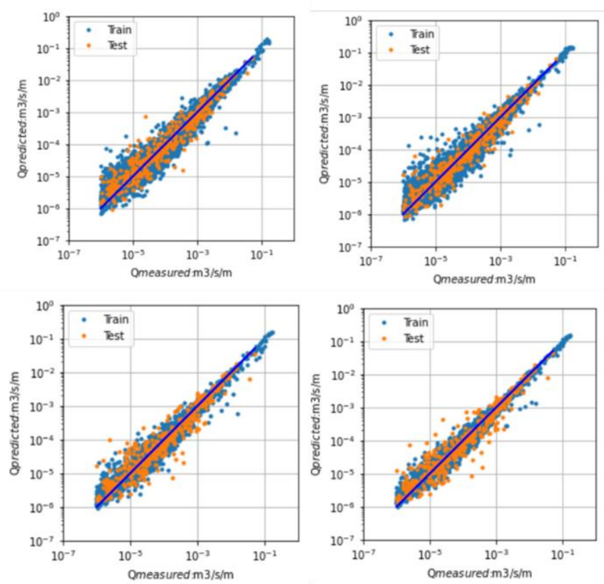

Figure 2 – Predicted vs. experimentally measured overtopping rates for the four new models trained from the EurOTop extended database. Training data shown as blue dots and test data shown as orange.

Root mean square error (RMSE) has been evaluated for the test dataset based on the logarithm of the predicted vs. measured overtopping rates. Values for the new models were found to be in the range 0.29 – 0.33, which compares well with values reported for other models (e.g. Zanuttigh et al. 2014). Further details can be seen in Table 1.

RESULTS AND CONCLUSIONS

Overtopping predictions have been generated for the Arup dataset using all four new models, along with the Bayonet and ANN2 models. RMSE values are shown in Table 1, alongside comparative values for the EurOTop extended dataset. In the latter case, values for the Bayonet and ANN2 models are those published by the model's developers.

As would be expected, error rates are higher for the Arup data, which contain designs which are not present in the training data and are generally more complex. There is no clear "winner" from the models included in the evaluation, with most of the differences likely to be explainable by natural variance in the input data (i.e. the order is likely to change if a different data set is used).

The distributions of predicted vs. measured values for the Arup dataset are shown via a series of scatter plots in Figure 3. Both Random Forrest regression models (RFR10 & RFR100) appear to systematically overpredict overtopping rates for experiments with low measured rates. Otherwise, there is little to distinguish the models based on their aggregate statistical performance.

Table 1 – RMSE values for all models, applied to both EurOTop extended database and new data set compiled from Arup projects.

Model	Data set	
	EurOTop	Arup
Bayonet	0.30	0.87
ANN2	0.28	0.86
RFR10	0.32	0.77
RFR100	0.33	0.73
GBR	0.32	0.75
GPR	0.29	0.92

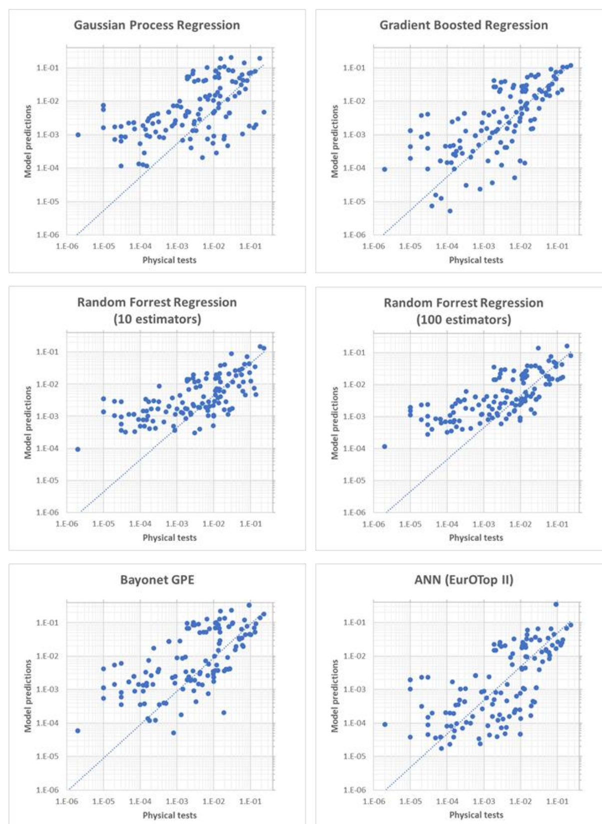

Figure 3 – Predicted vs. experimentally measured overtopping rates for all models using the Arup dataset.

Having compared the general performance of the models across the unseen Arup dataset, it appears that roughly equivalent statistical error rates are obtained for both the Bayonet and ANN2 models, which we believe are the most commonly used amongst the coastal engineering community. In addition, models trained with a broad range of off the shelf Machine Learning approaches perform relatively well in comparison.

Reviewing the ability of models to reproduce qualitative behaviour across specific experimental test series, it is clear that caution is needed when considering absolute values against performance criteria but trends were generally reliably reproduced. Two illustrative examples are shown in Figures 4 & 5.

For Figure 4, the structure is a vertical seawall at two levels. Escalating storm return period and severity of climate change scenarios unsurprisingly result in increased overtopping rates. The absolute values are predicted to within a factor of 2-3 across a wide range of conditions, with a general bias towards underprediction. Based on a nominal performance criterion of e.g. 50 l/s/m, we note that reliance on data driven models alone would likely lead to a different conclusion as to the acceptability of the design.

For Figure 5, the structure is a revetment armoured with concrete units, with a crown wall at the crest and a parapet. Trends are reliably reproduced for all models, with absolute errors again within a factor of 2-3, which is sufficient to lead to alternative conclusions regarding acceptability.

Based on a comparison of model predictions, the following conclusions were reached:

- While model performance is usually presented in terms of statistical aggregates, such as RMS prediction error, this does not provide a clear indication if the models can reliably reproduce trends which can be used in design 3ptimization studies.
- Based on the test series used in this study, we have found that qualitative trends are generally well reproduced for a range of different models.
- The magnitude of absolute prediction errors can be unpredictable - models that perform best in some instances perform worst in others - and are sufficient to lead to different conclusions regarding acceptability of overtopping performance in comparison with physical tests.
- Comparing results across a suite of models can provide a useful measure of prediction uncertainty alongside the "native" uncertainty measures provided by individual models.
- Although numerical tools may be useful for initial assessment and optimization, physical modelling remains the most reliable way to estimate wave overtopping for design purposes.

REFERENCES

Van der Meer, Allsop, Bruce, De Rouck, Kortenhaus, Pullen, Schüttrumpf, Troch, Zanuttigh (2018): EurOtop 2018. Manual on wave overtopping of sea defences and related structures. www.overtopping-manual.com

Pullen, Liu, Otinar Morillas, Wyncoll, Malde, Gouldby (2018): A generic and practical wave overtopping model that includes uncertainty, Maritime Engineering, Vol. 171.

Zanuttigh, Formentin, van der Meer (2014): Advanced in modelling wave-structure interaction through artificial neural networks, Coastal Engineering, No. 34.

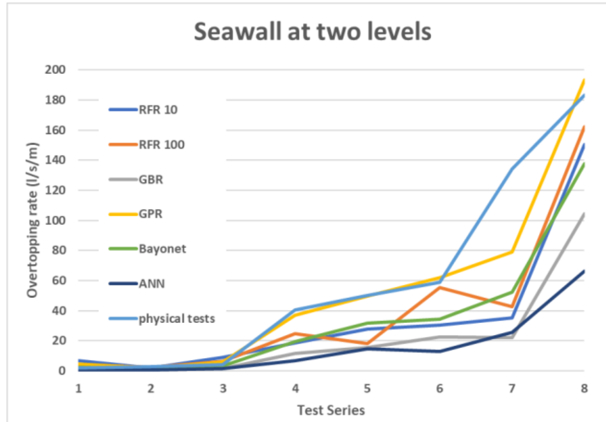

Figure 4 - Wave overtopping for an existing coastal structure, under a range of extreme wave conditions and future climate scenarios (test series of storm return period and severity of climate change scenarios in an ascending order)

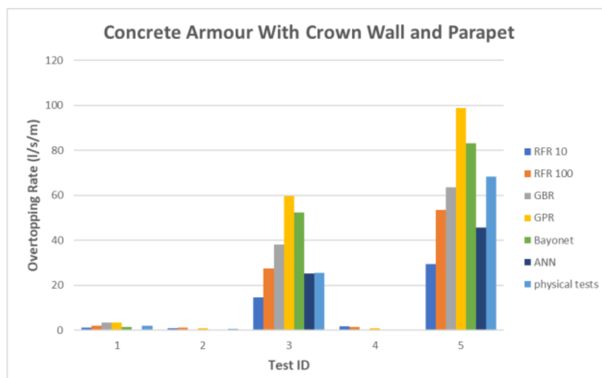

Figure 5 - Wave overtopping for a recently constructed breakwater under varying storm return periods in current (tests 1-3) and future climates (tests 4-5).