

Predicting the spatial distribution of seabed gravel content using random forest, spatial interpolation methods and their hybrid methods

J. Li^a

^a *Coastal, Marine and Climate Change, Environmental Geoscience Division, Geoscience Australia*
Email: jin.li@ga.gov.au

Abstract: Spatially continuous information of seabed sediments is often required for a variety of activities including seabed mapping and characterisation, prediction of marine biodiversity, and marine environmental planning and conservation. As seabed sediment data is often collected by point sampling, spatially continuous information must then be predicted from the point data. The accuracy of the predicted information is crucial to evidence-based decision making in marine environmental management and conservation. Improving predictive accuracy by identifying the most accurate methods is essential, but also challenging, since the accuracy is often data specific and affected by many factors.

Because of the high predictive accuracy of machine learning methods, especially Random Forest (RF), they were introduced into spatial statistics by combining them with existing spatial interpolation methods (SIMs), which resulted in new hybrid methods with improved accuracy. This development opened an alternative source of methods for spatial prediction. These hybrid methods, especially the hybrids of RF with inverse distance weighting (IDW) or ordinary kriging (OK) (i.e. RFOK or RFIDW), showed their high predictive capacity. However, their applications to spatial predictions of environmental variables are still uncommon. Model selection for RF and the hybrid methods is necessary and further test is required. Furthermore, model averaging has been argued to be able to improve predictive accuracy, but no consistent findings were observed in previous studies.

In this study, we aim to identify the most accurate methods for spatial prediction of seabed gravel content in the northwest Australian Exclusive Economic Zone. We experimentally examined: 1) whether input secondary variables affect the performance of RFOK and RFIDW; 2) whether the performances of RF, SIMs and their hybrid methods are data specific; and 3) whether model averaging improves predictive accuracy of these methods. For RF and the hybrid methods, up to 21 variables were used as predictors. The predictive accuracy was assessed in terms of relative mean absolute error and relative root mean squared error based on the average of 100 iterations of 10-fold cross-validation.

The findings of this study are:

- the predictive errors fluctuate with the input secondary variables;
- the existence of correlated variables can alter the results of model selection, leading to different models;
- the set of initial input variables affects the model selected;
- the most accurate model may be missed during the model selection;
- RF, RFOK and RFIDW proved to be the most accurate methods in this study, with RFOK preferred;
- these methods are not data specific, but their models are, so best model needs to be identified; and
- Model averaging is clearly data specific.

In conclusion, model selection is essential for RF and the hybrid methods. The best model needs to be identified for individual studies and application of model averaging should also be examined accordingly. RF and the hybrid methods have displayed substantial potential for predicting environmental properties and are recommended for further testing for spatial predictions in environmental sciences and other relevant disciplines. This study provides suggestions and guidelines for improving the spatial predictions of biophysical variables in both marine and terrestrial environments.

Keywords: *geostatistics, machine learning, spatial modelling, random forest, spatial interpolation*

1. INTRODUCTION

Environmental geoscience and geospatial information are often required for evidence-based decision making by governments, industry and the wider community on a range of economic, social and environmental issues relevant to managing Australia's natural resources. For example, seabed sediment data is often required for a variety of activities including seabed mapping and characterisation, prediction of marine biodiversity, and marine environmental planning and conservation (McArthur et al., 2010; Pitcher et al., 2008; Przeslawski et al., 2011). However, spatially continuous data are usually not available and information of many environmental variables including seabed sediments is usually collected by point sampling. Spatially continuous information must then be predicted from this often sparse, unevenly distributed point data. This is particularly true within the continental Australian Exclusive Economic Zone (AEEZ), due to the expense and practical limitations of acquiring samples.

Statistical and mathematical techniques for spatial prediction are essential tools for generating spatially continuous data from point data. The existing spatial interpolation methods (SIMs) can be classified into three groups (Li and Heap 2008, 2011): 1) deterministic or non-geostatistical methods (e.g. inverse distance weighting (IDW)); 2) stochastic or geostatistical methods (e.g. ordinary kriging (OK)); and 3) combined methods (e.g. regression kriging). These methods are, however, often data or even variable specific and their performance depends on many factors (Li and Heap 2011). No consistent findings have been obtained on how these factors affect the performance of spatial prediction methods, making it difficult to select an appropriate method for a given dataset (Li and Heap 2011). IDW is the most commonly applied interpolator because of its relative simplicity and availability. However, predictions using IDW are usually associated with large predictive errors (Li and Heap 2008). Therefore, it is often a challenge to select an appropriate spatial prediction method for a given study area.

The accuracy of spatially continuous environmental information is critical to evidence-based decision making. Improving the accuracy by identifying the most accurate SIMs is essential, but also challenging since the accuracy is often data specific and affected by many factors. Machine learning methods, like random forest (RF) and support vector machine (SVM), have showed their predictive accuracy in data mining and other disciplines in terms of predictive errors (Cutler et al., 2007; Diaz-Uriarte and de Andres, 2006; Drake et al., 2006; Shan et al., 2006). Thus they were introduced to spatial statistics by combining them with commonly used SIMs to predict the spatial distribution of seabed sediments (Li, 2011; Li et al., 2010). These hybrid methods, especially the hybrids of RF with IDW or OK, showed their high predictive capacity (Li et al., 2011b; Li et al., 2011c; Li et al., 2012b).

RF can select the most important variable at each node split (Okun and Priisalu, 2007). RF can also deliver good predictive performance even when most predictive variables are noisy (Diaz-Uriarte and de Andres, 2006). Thus model selection is assumed to be unimportant for RF and hence for the hybrid methods. This assumption was preliminarily tested for spatial predictions of environmental variables in previous studies (Li et al., 2011a; Li et al., 2011b; Li et al., 2012a; Li et al., 2012b). These studies have suggested that model selection is necessary for RF.

Furthermore, since model averaging can often improve the predictive accuracy (Marmion et al., 2009), it was tested in the geostatistical context, but no consistent findings were observed in previous publications. Model averaging showed no significant improvement (Li et al., 2011a; Li et al., 2011b) for seabed mud predictions but displayed further improvement for seabed sand predictions (Li et al., 2012b). Therefore, further study is warranted.

In this study, we aim to identify the most accurate predictive model to predict the spatial distribution of seabed gravel content in the northwest region of the Australian continental margin. To achieve this, we tested: 1) whether the input secondary variables affect the performances of RF, RFOK and RFIDW; 2) whether the performances of RF, SIMs and their hybrid methods are data-specific; and 3) whether model averaging improves predictive accuracy of these methods in the study region.

2. METHODS

2.1. Study area and datasets

The study area is located in the northwest region of AEEZ (Figure 1). This region covers an area of about 227,000 km² and comprises two geomorphic provinces (i.e. shelf and slope) (Heap and Harris 2008). It is mostly located on the shelf, with water depth ranging from 0 to 378 m, and backed by a coastline of southwest-northeast orientation.

In total, 237 samples of seabed sediments were considered in this study following data quality control (Li et al., 2010; Li et al., 2012b) and further check to ensure location information suitable for 250 m resolution. All samples were collected from the shelf. Sample density is very low, with 1.04 samples per 1000 km² on average. The sediments are divided into three groups: mud, sand and gravel (Li et al., 2010). In this study, we only focused on seabed gravel content.

A range of predictors could be used as secondary information to improve the spatial prediction of marine environmental data. However, only six predictors that were justified and used in previous studies (Li et al., 2011b; Li et al., 2012b) were employed in this study based on their availability for the whole continental AEEZ at the resolution required. These predictors are: bathymetry (bathy), distance-to-coast (dist.coast), seabed slope (slope), seabed relief (relief), latitude (lat) and longitude (long). Of these predictors, bathymetry data was based on Whiteway (2009), and seabed slope and relief were derived from the bathymetry data. All datasets of these variables were generated in ArcGIS at a 250 m resolution using the methods detailed by Li et al. (2010; 2012b). The coordinates system for this study was based on WGS84 as explained in previous studies (Li et al., 2011b; Li et al., 2011c).

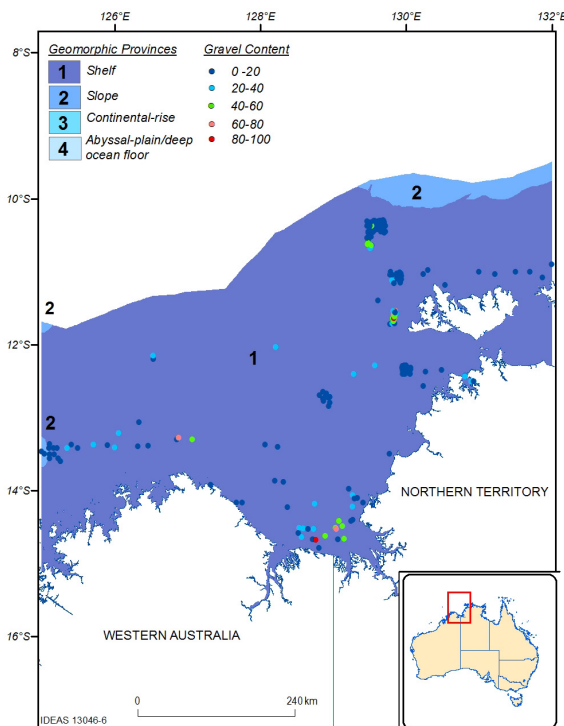


Figure 1. Spatial distribution of gravel content samples and their occurrence in the geomorphic provinces.

2.2. Methods for spatial prediction

RF, RFOK, RFIDW, and two most commonly compared methods, IDW and OK, were used in this study. For RFOK and RFIDW, six variables (bathy, dist.coast, slope, relief, lat, long) and their 15 derived variables (i.e. bathy², bathy³, dist.coast², dist.coast³, slope², slope³, relief², relief³, lat², long², lat*long, lat*long², long*lat², lat³ and long³) were used as predictors in the RF component based on previous studies (Li et al., 2011b; Li et al., 2011c). The residuals of RF were then interpolated using IDW with a searching window size of 5, and using OK with a Spherical model and a searching window size of 5 separately. For RF, the predictors used are identical to those used in the RF component in RFOK and RFIDW. For IDW, a distance power of 2 and a searching window size of 12 were used. For OK, log transformation was applied, and a Spherical variogram model and a searching window size of 12 were used. All these predictors and parameters were chosen based on our previous findings for predicting the seabed gravel content in AEEZ (Li et al., 2011d).

We further tested if model averaging could improve the prediction accuracy. In this study, we averaged the predictions of two methods (RFOK and RFIDW) or three methods (RFOK, RFIDW and RF) to produce the final predictions.

To compare the performance of these methods, a 10-fold cross-validation was employed. Randomness associated with the 10-fold cross-validation may lead to each method receiving different samples. To reduce such influence, we repeated the 10-fold cross-validation 100 times. Relative mean absolute error (RMAE) and relative root mean square error (RRMSE) (Li and Heap 2011) were used to assess the performance of the methods tested and to compare with findings in previous studies. The predictive errors were assessed based on the average of 100 iterations of 10-fold cross-validation.

The modelling was implemented in R 2.15.1 (R Development Core Team, 2012), using packages ‘raster’ for extracting data from different data layers, ‘gstat’ for geostatistical modelling and ‘randomForest’ for random forest modelling. Predictions were corrected by resetting the faulty estimates to the nearest bound of the data range (i.e. 0 or 100%) if applicable (Goovaerts, 1997).

3. RESULTS

3.1. Effects of input secondary variables on the performance of RF, RFOK and RFIDW

In total, 22 models were developed for RF, RFOK and RFIDW respectively (Table1). The predictive errors of all three methods fluctuated with the input secondary variables in terms of RRMSE (Figure 2). Since RMAE displayed a similar pattern to RRMSE in relation to the input secondary variables, it is not presented. RRMSE of models 6, 12, 20 and 21 was relatively lower than that of the rest models, while RRMSE of model 19 with only one predictor was prominently higher than that of all other models. Overall, model 21 for RF, RFOK and RFIDW was relatively more accurate than other models. One important phenomenon is that the most accurate model was missed during the initial model selection from model 1 to model 19, when the predictors in model 21 were gradually removed during model selection, leading to different models.

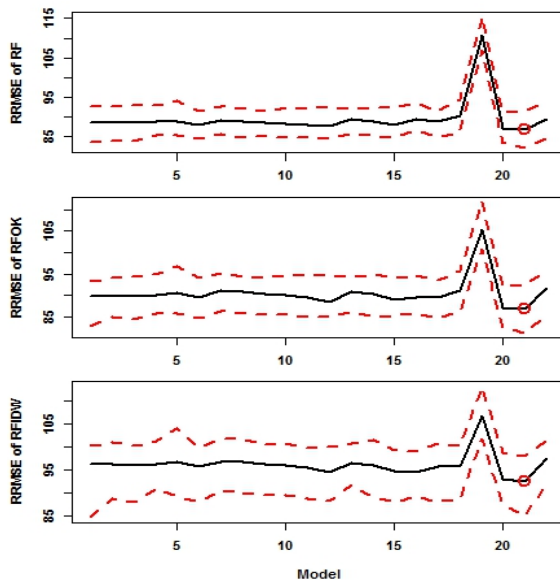


Figure 2. RRMSE (%) (mean: black line; minimum and maximum: dash lines) of 22 models of RF, RFOK and RFIDW for gravel content.

Table 1. Models developed for gravel content in the study region, with predictors contained in each model.

Modelling process	No of predictors
Model 1: All 21 predictors	21
Model 2: - sslope from model 1	20
Model 3: - cslope from model 2	19
Model 4: - clat from model 3	18
Model 5: - relief and crelief from model 4	16
Model 6: - latlon and slatlon from model 5	14
Model 7: - slope from model 6	13
Model 8: - cdist.coast from model 7	12
Model 9: - latslon from model 8	11
Model 10: - cbathy from model 9	10
Model 11: - slat from model 10	9
Model 12: - lat from model 11	8
Model 13: - srelief from model 12	7
Model 14: - sbathy from model 13	6
Model 15: - clon from model 14	5
Model 16: - slon from model 15	4
Model 17: - sdist.coast from model 16	3
Model 18: - bathy from model 17	2
Model 19: - lon from model 18	1
Model 20: lon, lat, bathy, dist, relief, slope	6
Model 21: lon, lat, bathy, dist, slope	5
Model 22: lon, lat, bathy, dist	4

Predictors: lon, lat, bathy, dist, relief, slope, sbathy, cbathy, sdist.coast, cdist.coast, srelief, crelief, cslope, slat, clat, slon, clon, latlon, latslon, sslope

3.2. Performance of RF, SIMs and their hybrid methods

The predictive error varied with the methods in terms of RMAE and RRMSE (Figure 3). RF, RFOK and RFIDW were the most accurate methods. They were significantly more accurate than the most commonly compared SIMs (i.e. IDW and OK), based on t-test for IDW in terms of RMAE, based on t-test for OK in terms of both RMAE and RRMSE (all with a p value < 0.0001). Of these three methods, RFIDW was significantly less accurate than RF and RFOK in terms of both RMAE and RRMSE based on paired t-test (with a p value < 0.0001). RF was significantly less accurate than RFOK in terms of RMAE based on paired t-test (with a p value < 0.0001), while there was no significant difference between RF and RFOK in terms of RRMSE based on paired t-test (with a p value = 0.2146). Overall, RFOK is preferred over RF and RFIDW.

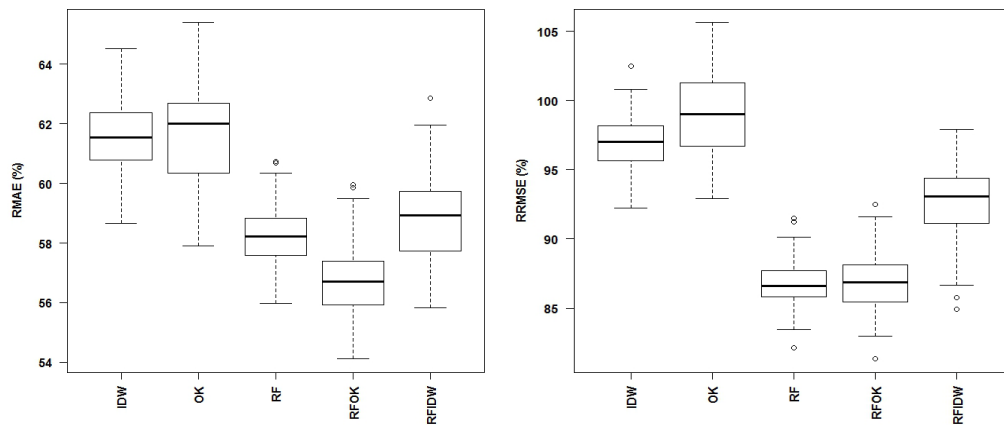


Figure 3. RMAE (%) and RRMSE (%) of IDW, OK, RF, RFOK and RFIDW for gravel content: summary statistics based on the 100 times 10-fold cross-validation.

3.3. Model averaging

Averaging of the predictions of the two or three most accurate methods produces two methods: RFOKRFDW and RFRFOKRFDW. The effects of this model averaging are summarised in Table 2. Both methods were significantly less accurate than RFOK in terms of both RMAE and RRMSE based on paired t-test (all with a p value < 0.0001). Thus RFOK is the most accurate method.

4. DISCUSSION AND CONCLUSIONS

4.1. Model selection for RF and its hybrid methods

Model selection is critical in model development. For RF, model selection is often assumed to be unimportant because: 1) RF is able to select the most important variable at each node split (Okun and Priisalu, 2007), thus is not sensitive to non-important variables; 2) RF can deliver good predictive performance even when most predictive variables are noisy (Diaz-Urriarte and de Andres, 2006); and 3) RF is relatively robust to outliers and noise and does not overfit (Breiman, 2001). This assumption proves to be invalid by preliminary studies being conducted in the application of RF to environmental sciences (Li et al., 2011a; Li et al., 2011b; Li et al., 2012a; Li et al., 2012b), and by recent studies in other disciplines (Genuer et al., 2010; Hapfelmeier and Ulm, 2013). In this study, four important findings were observed: 1) the predictive errors fluctuate with the input secondary variables; 2) the existence of correlated variables can alter the results of model selection, leading to different models; 3) the set of initial input variables affects the model selected; and 4) the most accurate model can be missed during the model selection and a less accurate model may be selected. Although the performance of RF is argued to depend only on the number of strong features, and not on how many noise variables if sample size is large (500 to 1000) (Biau, 2012), our findings have further demonstrated that model selection is essential for RF and its hybrid methods. This is perhaps because our sample is small, which is typical in environmental sciences. The difference between simulated data and environmental data may also play a significant role as discussed by Li (2013). In addition, although the model selected is relatively parsimonious for this study, the model selection adopted in this study often results in the most accurate model instead of the most parsimonious one. This is because the model selection is based on accuracy derived from cross-validation.

As observed by Li et al. (2013) and this study, randomness associated with RF often leads to a different order of variable importance between two different runs of RF model. Thus this results in different least important variable and then different model, which may cause difficulty in selecting the most accurate model and even lead to a suboptimal model being selected, as has been previously observed (Li et al., 2013). Apparently further research is required in this field, especially by considering some recently developed feature selection approaches (Genuer et al., 2010; Hapfelmeier and Ulm, 2013).

4.2. Are RF and the hybrid methods data specific?

RF, RFOK and RFIDW proved to be the most accurate methods in this study. This finding is consistent with previous studies (Li et al., 2011b; Li et al., 2011c; Li et al., 2012b) where RF and the hybrid methods (RFOK and RFIDW) showed their high predictive capacity. Of these three methods, RFOK is preferred over RF and RFIDW. These findings suggest that RF and the hybrid methods are not data specific. However, their models

Table 2. RMAE (%) and RRMSE (%) of RFOK and the averaged methods.

Method	Mean RMAE	Mean RRMSE
RFOK	56.72	86.83
RFOKRFDW	57.65	89.19
RFRFOKRFDW	57.32	87.63

are data specific. This is because the predictive errors change with the input secondary variables. Therefore, we need to find an optimal set of predictors for the methods for individual primary variables, and best model needs to be selected according to individual situations. The high accuracy of RF and the hybrids has been further confirmed by their application to a terrestrial environmental variable (Sanabria et al., 2013a; Sanabria et al., 2013b). The high predictive performance should be attributed to their features as discussed in previous studies (Li et al., 2011a; Li et al., 2011b; Li et al., 2011c). These methods have displayed substantial potential in applications to both marine and terrestrial environmental properties. They are recommended for environmental sciences and other relevant disciplines.

4.3. Model averaging

Averaging the predictions of two or three most accurate methods (i.e. RFOKRFIDW and RFRFOKRFIDW) reduced the predictive accuracy. This is mainly due to the relatively poor performance of RFIDW. This negative effect of model averaging is not supported by previous studies where model averaging either showed no significant improvement (Li et al., 2011a; Li et al., 2011b) or displayed further improvement (Li et al., 2012b; Marmion et al., 2009). It is evident that no consistent findings were observed regarding model averaging that is clearly data specific. Therefore, application of model averaging should be examined for individual studies. In addition, averaging the predictions of RF with the predictions of RFOK or RFIDW was not considered in this study because RF is usually less accurate than RFOK and RFIDW according to previous research (Li et al., 2011a; Li et al., 2012b). However, the finding in this study suggests that such averaging should be considered and, thus, is recommended for future studies.

In summary, model selection is essential for RF and its hybrid methods. RF and the hybrid methods are not data specific, but their models are. RFOK is the most accurate method. Model averaging is also data specific. Hence the best model needs to be identified for individual studies and application of model averaging should also be tested for individual studies. RF and the hybrid methods are recommended for generating spatial predictions in environmental sciences and other relevant disciplines in the future.

ACKNOWLEDGMENTS

Augusto Sanabria, Fuqin Li and Riko Hashimoto provided valuable comments on an earlier draft of this manuscript. Xiaojing Li extracted sediment samples from MARS database. Scott Nichol and Tony Nicholas provided clarification about some recently collected samples. Zhi Huang provided bathymetry, distance to coast, slope and relief data. Chris Lawson produced a map. This paper is published with permission of the Chief Executive Officer, Geoscience Australia.

REFERENCES

- Breiman, L., 2001. Random forests. *Machine Learning* 45 5-32.
- Biau, G., 2012. Analysis of a random forest method. *Journal of Machine Learning Research* 13 1063-1095.
- Cutler, D.R., Edwards, T.C.J., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. *Ecography* 88(11) 2783-2792.
- Diaz-Uriarte, R., de Andres, S.A., 2006. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* 7(3) 1-13.
- Drake, J.M., Randin, C., Guisan, A., 2006. Modelling ecological niches with support vector machines. *Journal of Applied Ecology* 43 424-432.
- Genuer, R., Poggi, J.M., Tuleau-Malot, C., 2010. Variable selection using random forest. *Pattern Recognition Letters* 31 2225-2236.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Hapfelmeier, A., Ulm, K., 2013. A new variable selection approach using random forests. *Computational Statistics & Data Analysis* 60 50-69.
- Heap, A.D., Harris, P.T., 2008. Geomorphology of the Australian margin and adjacent seafloor. *Australian Journal of Earth Sciences* 55 555-585.
- Li, J., 2011. Novel spatial interpolation methods for environmental properties: using point samples of mud content as an example. *The Survey Statistician: The Newsletter of the International Association of Survey Statisticians* No. 63 15-16.
- Li, J., 2013. Predictive modelling using random forest and its hybrid methods with geostatistical techniques in marine environmental geosciences. The 11th Australasian Data Mining Conference (AusDM2013) proceedings: Canberra (submitted).

Li., Predicting spatial distribution of seabed gravel content using RF, SIMs and their hybrid methods

- Li, J., Heap, A., 2008. A Review of Spatial Interpolation Methods for Environmental Scientists. *Geoscience Australia, Record 2008/23*, 137pp.
- Li, J., Heap, A., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. *Ecological Informatics* 6 228-241.
- Li, J., Heap, A., Potter, A., Daniell, J.J., 2011a. Predicting Seabed Mud Content across the Australian Margin II: Performance of Machine Learning Methods and Their Combination with Ordinary Kriging and Inverse Distance Squared. *Geoscience Australia, Record 2011/07*, 69pp.
- Li, J., Heap, A.D., Potter, A., Daniell, J., 2011b. Application of machine learning methods to spatial interpolation of environmental variables. *Environmental Modelling & Software* 26 1647-1659.
- Li, J., Heap, A.D., Potter, A., Huang, Z., Daniell, J., 2011c. Can we improve the spatial predictions of seabed sediments? A case study of spatial interpolation of mud content across the southwest Australian margin. *Continental Shelf Research* 31 1365-1376.
- Li, J., Heap, A.D., Potter, A., Huang, Z., Daniell, J., 2011d. Seabed gravel content across the Australian continental EEZ 2011. *Geoscience Australia*.
- Li, J., Potter, A., Heap, A., 2012a. Irrelevant Inputs and Parameter Choices: Do They Matter to Random Forest for Predicting Marine Environmental Variables?, *Australian Statistical Conference 2012: Adelaide*.
- Li, J., Potter, A., Huang, Z., Daniell, J.J., Heap, A., 2010. Predicting Seabed Mud Content across the Australian Margin: Comparison of Statistical and Mathematical Techniques Using a Simulation Experiment. *Geoscience Australia, 2010/11*, 146pp.
- Li, J., Potter, A., Huang, Z., Heap, A., 2012b. Predicting Seabed Sand Content across the Australian Margin Using Machine Learning and Geostatistical Methods. *Geoscience Australia, Record 2012/48*, 115pp.
- Li, J., Siwabessy, J., Tran, M., Huang, Z., Heap, A., 2013. Predicting Seabed Hardness Using Random Forest in R. In: Zhao, Y., Cen, Y. (Eds.), *Data Mining Applications with R*. Elsevier (in press).
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W., 2009. Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* 15 59-69.
- McArthur, M.A., Brooke, B.P., Przeslawski, R., Ryan, D.A., Lucieer, V.L., Nichol, S., McCallum, A.W., Mellin, C., Cresswell, I.D., Radke, L.C., 2010. On the use of abiotic surrogates to describe marine benthic biodiversity. *Estuarine, Coastal and Shelf Science* 88 21-32.
- Okun, O., Priisalu, H., 2007. Random forest for gene expression based cancer classification: overlooked issues. In: Martí, J., Benedí, J.M., Mendonça, A.M., Serrat, J. (Eds.), *Pattern Recognition and Image Analysis: Third Iberian Conference, IbPRIA 2007 Lecture Notes in Computer Science: Girona, Spain*, pp. 4478: 4483-4490.
- Pitcher, C.R., Doherty, P.J., Anderson, T.J., 2008. Seabed environments, habitats and biological assemblages. In: Hutchings, P., Kingsford, M., Hoegh-Guldberg, O. (Eds.), *The Great Barrier Reef: biology, environment and management*. CSIRO Publishing: Collingwood, p. 377.
- Przeslawski, R., Daniell, J., Anderson, T., Vaughn Barrie, J., Heap, A., Hughes, M., Li, J., Potter, A., Radke, L., Siwabessy, J., Tran, M., Whiteway, T., Nichol, S., 2011. Seabed Habitats and Hazards of the Joseph Bonaparte Gulf and Timor Sea, Northern Australia. *Geoscience Australia, Record 2008/23*, 69pp.
- R Development Core Team, 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna.
- Sanabria, L.A., Cechet, R.P., Li, J., 2013a. Combining observational and modelling studies to develop a map of Australian fire weather potential, *The 20th International Congress on Modelling and Simulation (MODSIM2013): Adelaide*.
- Sanabria, L.A., Qin, X., Li, J., Cechet, R.P., Lucas, C., 2013b. Spatial interpolation of the FFDI fire weather Index: Observational study. *Environmental Modelling & Software* (submitted).
- Shan, Y., Paull, D., McKay, R.I., 2006. Machine learning of poorly predictable ecological data. *Ecological Modelling* 195 129-138.
- Whiteway, T., 2009. *Australian Bathymetry and Topography Grid, June 2009*. *Geoscience Australia*, p. 46pp.