www.iiste.org

# **Risk Assessment Model for Pluvial Flood Prediction Using Fuzzy-Based Classification Technique**

Oladapo Kayode Abiodun Department of Computer Science, School of Computing and Engineering Science, Babcock University, Ilisan-Remo, Ogun State, Nigeria oladapokayodeabiodun@gmail.com

Idowu, S.A. Department of Software Engineering, School of Computing and Engineering Science, Babcock University, Ilisan-Remo, Ogun State, Nigeria osisanwof@babcock.edu.ng Ayankoya, F.Y.

Department of Computer Science, School of Computing and Engineering Science, Babcock University, Ilisan-Remo, Ogun State, Nigeria idowus@babcock.edu.ng

Adekunle, Y.A

Department of Computer Science, School of Computing and Engineering Science, Babcock University, Ilisan-Remo, Ogun State, Nigeria adekunlea@babcock.edu.ng

#### Abstract

Both developed and developing countries are promoting risk management and refining the ability to alleviate the effects of disaster both man-made and natural, which have become a threat to human life and the world's economy. The variability in climate change, rapid urbanization and fast-growing socio-economic development has naturally increased the risk associated with flooding. A recent report showed that flood have affected more individuals than any other category of disaster in the 21st century with the highest percentage of 43% of all disaster events in 2019 and Africa been the second vulnerable continent after Asia. So, it is highly important to devise a scientific method for flood risk reduction since it cannot be eradicated. Machine learning can improve the risk management. The paper proposes a pluvial flood detection and prediction system based on machine learning techniques. The proposed model will employ a fuzzy rule-based classification approach for pluvial flood risk assessment.

Keywords: Machine Learning, Pluvial Flood, Risk, Fuzzy Rule-Based, Prediction DOI: 10.7176/CEIS/12-1-07 Publication date: January 31<sup>st</sup> 2021

# 1. Introduction

Technology and advancement in science drives transformation around the world. These advanced technologies can capture and analyse vast amount of data which has positively affected various industries and how the world do businesses. Machine learning is one of the technologies where machines reproduces human intelligence through learning, reasoning and self-correction which is actually poised to assist human decision-making. The intention of machine learning is to enable machine to learn by themselves using the provided data and make accurate prediction.

Flood management have embraced waves of automation over many years to improve the efficiency and effectiveness of reducing the risk associated with flooding. Handling flood risk with the intention of safety and comfort of the citizens as well as saving their environemnts is one of the major responsibilities of each country leadership especially in flood prone areas. At large, flood management has advanced from flood control tactics to flood risk administration. Governments, are therefore under burden to advance a consistent and precise maps of flood susceptible area, advance strategy for maintainable risk administration in flood with emphasis on preparation, avoidance and fortification though this is not a try to eradicate flood risk but its goal to alleviate it.

Flood are foreseen to transpire more sternly in urban areas across the globe (Nasiri, Mohd Yusof, & Mohammad Ali, 2016), findings have shown that there exists three key steps in pluvial flood risk management namely, flood planning mitigation, responsive measures and recovery (Tingsanchali, 2012; Nasiri, Mohd Yusof, & Mohammad Ali, 2016) which are subject generally to two approaches non-structural and structural ways. Structural approach are measures which involves physical construction such as flood defence, catchment and coastal defences (Geoffrey, 2010; Hussein, 2017) and are based on engineered solutions, such as river dike that modifies the river flow, dams or levees (Nasiri, Mohd Yusof, & Mohammad Ali, 2016). The basic idea comprise of storing, averting and confinement of floods to reduce risk. On the other hands, several methods were developed by various researches in non-structural measures which include enlightening, emergency services, reporting, building codes,

cautioning and forescasting, flood indemnity, assessing methods, health and social methods, land use development and public contribution (Nasiri, Mohd Yusof, & Mohammad Ali, 2016).

In recent years, methods of mitigating and preventing flood disasters have moved from defensing approach to management approach, this is based on the comprehensive risk assessment findings and cost with benefit analysis. A number of approaches as reported in literatures has been in use to detect and predict flood vulnerable zones such as multi-criteria method to define the operative deployment of geospatial systems for reducing flood threat, mathematical index for large-scale analysis of flood susceptibility, social research information integration into flood risk analyses, remote sensing, multi-criteria decision approach in evaluating flood susceptibility, multidimensional model for urban flood risk assessment, satellite imageries in identifying areas susceptible to flooding. Recent methods include machine learning means such as artificial neural network, linear regression, support vector machine, and many of this can predict areas susceptible to flooding and map flood susceptibly areas. Most of these methods unveil certain weaknesses. Indeed, previous methods of managing flood have generally demonstrated the continuing power of human expertise and the limits of machine, this underscore the need for this study to employ a fuzzy rule-based classification method as dimensionality reduction algorithm. This name focused on the development of a risk assessment model for prediction by using fuzzy.

This paper focused on the development of a risk assessment model for pluvial flood prediction by using fuzzy based classification technique alongside machine learning algorithm

# 2.0 Pluvial Flood Detection and Prediction Techniques

Pluvial flooding can broadly be seen as an end result of heavy rainfall as well as water overflow, which are caused by various conditioning variables (Falconer et al., 2009). There is no universal method to classify pluvial flood prediction model (method/techniques) models have been classified in several ways depending on the criteria of interest. Over the years, the research conducted in the field of pluvial flooding specific to techniques heavily concentrated on physical model followed by conceptual model and data driven model. Pluvial flooding frequently encompasses the use of computational methods and these methods are applied primarily to solve (i) physical-based flow governing equation (ii) Data-driven model and empirical stochastic model (Ghimire et al., 2013). Thorndahl et al., (2016) also shared parallel idea that, all together, the accessible susceptibility assessment approaches can be modified into physical based, conceptual model also known as expert-knowledge based model and data driven.

Nevertheless, with variation in techniques, it is hard to be confident that this is the best method in pluvial flooding prediction. Simões (2012) highlighted real time urban flood detection and prediction systems based on four different scenarios, (i) Real time flood forecast established only on rainfall statistics and empirical situations (ii) Real time flood forecast established on rainfall statistics and the pre-simulated situations (iii) Real time flood forecast established on real-time information integration (iv) Real time flood forecast with active response on the drainage structure operation.



Figure 1: Classification of pluvial flood prediction techniques. Source: Idowu et al., (2020)

# **3.0 Related Work**

Yahaya, Ahmad, & Abdalla (2010) proposed the integration of Multicriteria Evaluation and Multicriteria Decision Making Methods (MCDA) to analyze the flood vulnerable areas. The Boolean overlay and weighted linear combination were integrated with Geographical Information System (GIS), using pairwise comparison and ranking methods to analyse the weightiness of each factor. The technique does not consider topographic witness index that displays the quantity of water accumulation in several area as conditioning factor which is a great limitation.

Chen, Yeh, & Yu, (2011) developed a hierarchical structure by integrating geographic information system and analytic hierarchy process, to make available an ideal choices in flood risk breakdown for towns and semi-rural regions. Nonetheless, the method can be improved by using lengthy rainfall and flood data with the addition of more conditioning factors and risk categories.

Rahmati, Pourghasemi, & Zeinivand (2016) describes how to map flood vulnerability by using frequency ratio and weight-of-evidence models using area under the curves in measuring the performance of the models. It was concluded that the weight-of-evidence and frequency ratio can be applied as an instrument in flood vulnerability mapping based on the calculated area under the curves. The drawback is that the larger the size of the dataset the better the classifier, so the effectiveness cannot be ascertained due to the dataset involved.

Lee, Kim, Jung, Lee, & Lee (2017) used boosted-tree models and random forest in a geographic information system environment, after the selection of 12 conditioning variables that are connected to flooding. The study concluded that geology and the calculated digital elevation model (DEM) show up as the most significant predictors and the results were considered satisfactory when the model accuracies are over 75% for every result. Nevertheless, the accuracy of the data that were used was difficult to be established because it was built over an assessment inside the administrative-district structure and inappropriate data on location possibly will cause extensive problem in the study.

Lee, Kim, Jung, Lee, and Lee (2018) applied logistic regression and frequency ratio (FR) model for flood data, in deciding the relationship amid flooded region with the casual features also to derive flood susceptibility map based on the selected core variables. The results were demonstrated to be reliable with over 75% accuracy from both models. However, the drainage system was presumed to have no effect as an indicator but plays a vital role in flood occurrence, so if the sewer-related data can be secured it will enhance the result, which is a great limitation.

Luu, Von Meding, & Kanjanabootra (2018) presented a technique to analyse flood risk by means of flood mark data (flood depth and duration) with decision making judgement (analytic hierarchy process (AHP) method) which was merged into a single map by means of a weighted linear combination. The valuation was included into a geographic information system (GIS) structure to produce a flood hazard map. The 2013 flood occurrence in the city of Quang Nam, Vietman was used as the case study and the result match up with the hydraulic model MIKE11-GIS. However, a machine learning method could be used to progress the assessment.

Tang, Zhang, Yi, & Xiao (2018) created an integrated flood susceptibility assessment structure using probabilistic GIS-MCDA method to cover up for the conventional technique for the marking out on paper the flood susceptible zones. The method go well with the local ordered weighted averaging (OWA) and the probabilistic approaches through Monte Carlo simulation, considering the uncertainty related to criteria weights, risk attitude of the analyst and the spatial heterogeneity of preferences. Sensitivity analysis and uncertainty were implemented in examining the strength of the model which indicate that the ensemble method improves the outcome of in respect to the criteria weights and identify the most accountable weight for the inconsistency of the model outcome. However, the purpose of the evaluation benchmarks could be enhanced with approaches for instance machine learning techniques and factor analysis.

Samanta, Bhunia, Shit, & Pourghasemi (2018) proposed a GIS based-bivariate statistical analysis on eight flood conditioning factors. Each of the variables were categorised using the quantile technique and the relationship of each category was evaluated with frequency ratio probability model which was validated using the area under the curves (AUC). The study did not consider other conditioning factors like proximity to the river, water budget and therefore the optimal efficiency of the solution is a serious drawback.

Hong et al. (2018) implemented data mining technique namely random forest, logistic regression and support vector machine and a fuzzy weight-of-evidence for constructing flood susceptibility mapping and considering eleven flood related variable. The efficiency was appraised based on area under curve for measuring the realisation rates and prediction. The research maintained that, the proposed fuzzy weight-of-evidence make available a

substitute tool for flood vulnerability modelling. However, other machine learning techniques and performance measurement metrics can be applied to improve the model.

Zhao, Pang, Xu, Yue, & Tu (2018) mapped flood susceptibility in mountainous region of China based on historical flooding records using random forest model through factor contributing analysis to characterize the occurrence with conditioning variables. The result confirm that random forest model can identify the flood susceptibility with satisfactory accuracy and has an advantage over support vector machine and artificial neural network. However, the representation of flood characteristic of mountainous areas are still difficult which is a major drawback.

Seyed, Kornejady, Pourghasemi, & Keesstra (2018) proposed a premier flood susceptibility model merging adaptive neuro-fuzzy inference system with genetic algorithm, particle swarm and ant colony optimization then relating their accuracy to model flood susceptibility. Learning vector quantization was utilized to estimate the importance of the conditioning variables while the frequency ratio model was applied in assigning weights for each classification. The adaptive neuro-fuzzy inference system particle swarm optimization was reported to be a premier model. The result was attested with a chi-square which gave the same result. The major drawback with this solution is data paucity.

Ouma & Tateishi (2014) proposed an integrated multi-parametric approach. The approach created an easy to read pluvial flood vulnerability map, using analytical hierarchy process as multi-criteria analysis instrument in creating a three-level hierarchical structure that represent pluvial flood vulnerability and GIS used in the design of the criteria maps. The integration of AHP/GIS offers a way out to multi-parametric and multi-dimensional composite processes. It shows it's possible practicality for a broader model where multi-parametric geospatial analysis are engaged. Less expensive, and easy to use because the rating indices and arithmetical outcomes can be utilized as a mention point but there is difficulty in standardizing the dataset for flood risk evaluation.

Noymanee, Nikitin, & Kalyuzhnaya, (2017) designed a water level model. The use of machine learning techniques for open data using machine learning based model (Bayesian linear model) for predicting flood peak in the city. The design has the ability to handle complex task. It can be used as a temporary warning system but a large error occurred when escalating the period of time of ahead of 12 hours. The model was limited by time and resources and alteration of data set can create a radical difference in the model.

Semi-supervised Machine Learning Model was created by Nasiri, Mohd Yusof, & Mohammad Ali (2016) by using a weakly labelled support vector machine (WELLSVM) model as an enhanced support vector machine was technologically advanced for urban flood vulnerability. It can efficiently take care of weakly labelled data in large datasets and suitable for binary weakly labelled problems but the use of both uncertain labelled records and massive unlabelled data can advance the accuracy of the model.

Lee, Lee, Lee, & Jung (2018) conducted a data mining model using frequency ratio and logistic ratio models as data mining techniques with geographic information system (GIS) tools. This was used in generating susceptibility map to correlate amid flood data with related conditioning variables. This model was beneficial for clarifying the mechanism between flood incidents and related factors. There were difficulty in acquiring data set and the model validation was affected due to lack of data.

Seyoum, Willems, & Verbeiren (2019) propose a rainfall-runoff forecasting model. The model put forward an interactive and supportive system (data driven – a multilayer perception Artificial Neural Networks and Random Forest) at refining the monitoring and administration of urban pluvial flooding using historical rainfall and runoff data. It shows promise in the absence of hydraulic model. To improve the prediction of pluvial flows, test of different data transformation technique will be needed.

#### 4.0 Proposed Risk Assessment Model for Pluvial Flood Prediction

The purpose of pluvial flood detection and prediction is to alert the general public and concerned authorities of an impending pluvial flood as much in advance, and with as much reliability as possible. Flood prediction are essential utensils for municipal flood management, the groundwork of emergency response plans and for that reason will improve the security of lives with properties and the reduction in flood loss (Nkwunonwo, Whitworth, & Baily, 2015). Flood detection was suggested as part of the mechanisms of a flood warning structure by Sayers, (2015). The detection and prediction of pluvial flood depend importantly on the reliability of the available weather-related database. For building machine learning model, there are five basic flow which are data collection, pre-processing, model building, training and testing (Mosavi, Ozturk, & Chau, 2018). This can be classified into three fundamental levels. The pre-processing includes feature transformation and extraction, the processing includes model

generation, choosing of algorithm and parameters and post processing is where knowledge is represented (Richert & Coelho, 2013). The gaps identified in the existing models will aid as a guide to the improvement of a highly predictive conditioning variables subset for the proposed model. The proposed model will be aimed at evaluating the dataset using fuzzy rule based classification on the selected conditioning variables with the aim of classifying the pluvial flood susceptibilities and later be evaluated with machine learning algorithms. The model is proposed to have four components namely: data collection stage, data storage stage, model learning stage then mapping stage as shown in Figure 2. Data collection stage, which is the data collection phase where the conditioning variables are supplied into the model based on the categorization (ii) the data storage stage where the dataset are stored for processing (iii) model learning stage which comprises of the model building and decision module, this contained the code for the selected machine learning model to predict pluvial flood susceptibilities. (iv) mapping stage which displays the pluvial flood susceptibilities according to predicted multi-classification.

These are the basic steps for employing the proposed model:

- i. Selection of the identified category of the pluvial flood conditioning variables to generate the pluvial flood dataset.
- ii. Application of fuzzy logic on the generated pluvial flood dataset for classification.
- iii. The pluvial flood conditioning variables were stored as a geo-spatial database for processing.
- iv. Application of the selected machine learning algorithms on the fuzzified pluvial flood conditioning variables
- v. Performance evaluation of the machine learning algorithms using the selected metrics namely sensitivity, specificity, accuracy percentage and precision.
- vi. Testing of the selected model selected machine learning model to predict pluvial flood susceptibilities
- vii. Mapping of the pluvial flood susceptibilities according to predicted multi-classification to produce a pluvial flood susceptibility map with categorization.



Figure 2: Proposed Risk Assessment Model for Pluvial Flood Prediction

# 5.0 Preliminary Results and Future Research

The hydrological category of the conditioning variable categorization was identified as the most important of the three (3) categories related to pluvial flood. If those conditioning variables in the category do not exist, it has the greatest impact on the occurrence of pluvial flood and are highly dependent on the other categories. The next category is topographical followed by anthropological.

Overall, the findings in Oladapo, Idowu, Adekunle, & Ayankoya (2020) are a step towards providing a framework or guidelines in the selection of pluvial flood conditioning variables. This finding is a step toward providing a clear defined category to effectively inform the preparation for the development of the pluvial susceptibility map using machine learning techniques.

These preliminary findings have provided insights into the basic requirements for the proposed risk assessment model for pluvial flood prediction, which is the most important remaining future work of this research. The proposed system is yet to be implemented hence future work will include full system implementation and improvement of system accuracy through the deployment of machine learning algorithms. Furthermore statistical metrics such as sensitivity, specificity and precision shall be used in evaluating the predictive capability of the proposed model as well as producing a pluvial flood susceptibility map with categorization

#### **6.0** Conclusion

Risk assessment is the step by step procedure of appraising the impending damage that may be involved in a projected activity. Therefore, preparing ahead of potential pluvial flood incident and risk is important. The risk assessment solution that will allow the identification of flood susceptibilities in any developing country cannot be overstressed. This paper discusses the method of detecting and predicting pluvial flood susceptibilities using machine learning techniques and as well proposes a model for predicting pluvial flood using machine learning algorithms. The proposed model will use a fuzzy rule-based classification technique as dimensionality reduction algorithm for feature selection and evaluate the machine learning algorithm performance on each set of conditioning variables based on the conditioning variable predictive power that is whether it is high or low.

#### References

- Chen, Y. R., Yeh, C. H., & Yu, B. (2011). Integrated application of the analytic hierarchy process and the geographic information system for flood risk assessment and flood plain management in Taiwan. *Natural Hazards*, 59(3), 1261–1276. https://doi.org/10.1007/s11069-011-9831-7
- Choubin, B., Moradi, E., Golshan, M., Adamowski, J., Sajedi-Hosseini, F., & Mosavi, A. (2019). An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Science of the Total Environment*, 651, 2087–2096. https://doi.org/10.1016/j.scitotenv.2018.10.064
- Falconer, R. H., Cobby, D., Smyth, P., Astle, G., Dent, J., & Golding, B. (2009). Pluvial flooding: New approaches in flood warning, mapping and risk management. *Journal of Flood Risk Management*, 2(3), 198–208. https://doi.org/10.1111/j.1753-318X.2009.01034.x
- Geoffrey, J. (2010). *Climate change and future flooding in the UK* (University of East Anglia). Retrieved from https://core.ac.uk/download/files/124/2771227.pdf
- Ghimire, B., Chen, A. S., Guidolin, M., Keedwell, E. C., Djordjević, S., & Savić, D. A. (2013). Formulation of a fast 2D urban pluvial flood model using a cellular automata approach. *Journal of Hydroinformatics*, 15(3), 676–686. https://doi.org/10.2166/hydro.2012.245
- Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A. X., & Chen, W. (2018). Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. *Science of the Total Environment*, 625(January), 575–588. https://doi.org/10.1016/j.scitotenv.2017.12.256
- Idowu, S. ., Adekunle, Y. A., Ayankoya, F. Y., & Oladapo, K. . (2020). Pluvial Flood Detection and Prediction Methods : A Review. International Journal of Computer Trends and Technology (IJCTT), 11(3), 1–9.
- Lee, S., Kim, J. C., Jung, H. S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomatics, Natural Hazards* and Risk, 8(2), 1185–1203. https://doi.org/10.1080/19475705.2017.1308971
- Lee, S., Lee, S., Lee, M. J., & Jung, H. S. (2018). Spatial assessment of urban flood susceptibility using data mining and geographic information system (GIS) tools. *Sustainability (Switzerland)*, 10(3). https://doi.org/10.3390/su10030648
- Luu, C., Von Meding, J., & Kanjanabootra, S. (2018). Assessing flood hazard using flood marks and analytic hierarchy process approach: a case study for the 2013 flood event in Quang Nam, Vietnam. *Natural Hazards*, 90(3), 1031–1050. https://doi.org/10.1007/s11069-017-3083-0
- Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. Water (Switzerland), 10(11). https://doi.org/10.3390/w10111536

Nasiri, H., Mohd Yusof, M. J., & Mohammad Ali, T. A. (2016). An overview to flood vulnerability assessment

methods. Sustainable Water Resources Management, 2(3), 331–336. https://doi.org/10.1007/s40899-016-0051-x

- Nkwunonwo, U. C., Whitworth, M., & Baily, B. (2015). Flooding and Flood Risk Reduction in Nigeria: Cardinal Gaps. Journal of Geography and Natural Disasters, 5(1), 1–12. https://doi.org/10.4172/2167-0587.1000136
- Noymanee, J., Nikitin, N. O., & Kalyuzhnaya, A. V. (2017). Urban Pluvial Flood Forecasting using Open Data with Machine Learning Techniques in Pattani Basin. *Procedia Computer Science*, 119(2017), 288–297. https://doi.org/10.1016/j.procs.2017.11.187
- Oladapo, K. A., Idowu, S. A., Adekunle, Y. ., & Ayankoya, F. . (2020). Categorization of Conditioning Variables for Pluvial Flood Risk Assessment. *International Journal of Scientific & Engineering Research*, 11(8), 355–368.
- Ouma, Y. O., & Tateishi, R. (2014). Urban flood vulnerability and risk mapping using integrated multiparametric AHP and GIS: Methodological overview and case study assessment. *Water (Switzerland)*, 6(6), 1515–1545. https://doi.org/10.3390/w6061515
- Rahmati, O., Pourghasemi, H. R., & Zeinivand, H. (2016). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto International*, 31(1), 42–70. https://doi.org/10.1080/10106049.2015.1041559
- Richert, W., & Coelho, L. P. (2013). Building Machine Learning Systems with Python (First). Birmingham: Packt Publishing.
- Samanta, R. K., Bhunia, G. S., Shit, P. K., & Pourghasemi, H. R. (2018). Flood susceptibility mapping using geospatial frequency ratio technique: a case study of Subarnarekha River Basin, India. *Modeling Earth Systems and Environment*, 4(1), 395–408. https://doi.org/10.1007/s40808-018-0427-z
- Sayers, W. K. (2015). Artificial intelligence techniques for flood risk management in urban environments (University of Exeter). https://doi.org/10.1016/j.proeng.2014.02.165
- Seyed, V. R. T., Kornejady, A., Pourghasemi, H. R., & Keesstra, S. (2018). Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Science of the Total Environment*, 615, 438–451. https://doi.org/10.1016/j.scitotenv.2017.09.262
- Seyoum, S., Willems, P., & Verbeiren, B. (2019). Data-Driven Models for Pluvial Flood Forecasting in Brussels. *NOVATECH*, 1–4. Stormwater and Urban Planning.
- Simões, N. E. da C. (2012). Urban Pluvial Flood Forecasting. Imperial College London.
- Tang, Z., Zhang, H., Yi, S., & Xiao, Y. (2018). Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis. *Journal of Hydrology*, 558, 144–158. https://doi.org/10.1016/j.jhydrol.2018.01.033
- Thorndahl, S., Nielsen, J. E., & Jensen, D. G. (2016). Urban pluvial flood prediction: A case study evaluating radar rainfall nowcasts and numerical weather prediction models as model inputs. *Water Science and Technology*, 74(11), 2599–2610. https://doi.org/10.2166/wst.2016.474
- Tingsanchali, T. (2012). Urban flood disaster management. *Procedia Engineering*, *32*, 25–37. https://doi.org/10.1016/j.proeng.2012.01.1233
- Yahaya, S., Ahmad, N., & Abdalla, R. F. (2010). Multicriteria analysis for flood vulnerable areas in hadejiajama'are river Basin, Nigeria. *European Journal of Scientific Research*, 42(1), 71–83. Retrieved from http://www.eurojournals.com/ejsr.htm
- Zhao, G., Pang, B., Xu, Z., Yue, J., & Tu, T. (2018). Mapping flood susceptibility in mountainous areas on a national scale in China. Science of the Total Environment, 615, 1133–1142. https://doi.org/10.1016/j.scitotenv.2017.10.037