



SDI Review Form 1.6

Journal Name:	British Journal of Mathematics & Computer Science
Manuscript Number:	2013_BJMCS_6050
Title of the Manuscript:	Modeling Crop Residue Yield Potential using Geographically Weighted Regression
Type of the Article	Research Paper

General guideline for Peer Review process:

This journal's peer review policy states that **NO** manuscript should be rejected only on the basis of '**lack of Novelty**', provided the manuscript is scientifically robust and technically sound.

To know the complete guideline for Peer Review process, reviewers are requested to visit this link:

(<http://www.sciencedomain.org/page.php?id=sdi-general-editorial-policy#Peer-Review-Guideline>)



SDI Review Form 1.6

PART 1: Review Comments

	Reviewer's comment	Author's comment (if agreed with reviewer, correct the manuscript and highlight that part in the manuscript. It is mandatory that authors should write his/her feedback here)
Compulsory REVISION comments	This paper presents an analysis of crop residue yield as a function of temperature and precipitation. Two regression methods are compared; OLS regression and GWR. The paper does not present anything new in terms of regression methodology, only an application of existing methodology. Moreover the comparison with OLS regression is not a fair one in my opinion. OLS is a linear method, and clearly will never perform better than a nonlinear method on nonlinear data. I therefore think it would be much more interesting to see a comparison of GWR with other nonlinear methods, especially other locally linear methods like e.g. Hierarchical Cluster-based Partial Least Squares Regression. I think the paper should also describe other locally linear methods as well as other nonlinear regression methods, and explain why GWR was chosen to be used instead of alternative methods. Moreover, an explanation of the metrics used to assess predictive ability should be given, as well as a description of why these were chosen. Here cross-validation was used, but at least for the large dataset there should be more than enough observations to use the more robust validation method test set validation, where an independent set of observations is used only for validation.	We have demonstrated that GWR performs better than OLS model and more importantly that spatial heterogeneity of GWR parameters depends on the sample size of the data (we demonstrated, the individual parameters are not spatially heterogeneous in small size data set, and are spatially heterogeneous for large size data set). So, the new thing here is discussion of relation between sample size and spatial heterogeneity. As far as the comparison of OLS with GWR is concerned, we followed a common approach (Brown et al, 2012; Tulbure et al 2011, for example). GWR can be extended to a generalized linear model (if the response is a count or binary). But, our response is continuous, and hence, our GWR is a linear model, not a nonlinear model. Because of this, we did not use any nonlinear model. Moreover, GWR is the only method that addresses spatial dependency of the data and provides local estimates. (Even CAR and SAR give only a global parameter estimates.) Our data is spatial data, and spatial dependency must be addressed, which is only possible by the use of GWR. This is the reason why GWR was chosen over other many more



SDI Review Form 1.6

	The paper does not provide the resulting regression equations giving the crop residue yield as a function of temperature and precipitation. I would like to see this. It is not clear whether polynomial terms is taken into account in the regression here.	alternative methods. Cross validation is the most widely used method for estimating kernel bandwidth parameter. Corrected AIC_c approach could have been used as well, but the results from GWR are not much sensitive to the choice of these methods. In GWR, there is not a single resulting regression equation giving the response as a function of the predictors. Note that there are 743 locations, and in each location, one model is fitted. There are potentially large number of parameter estimates in GWR output (See, for example, the last paragraph on page 10 of GEOGRAPHICALLY WEIGHTED REGRESSION, WHITE PAPER, by MARTIN CHARLTON, A STEWART FOTHERINGHAM, National Centre for Geocomputation, National University of Ireland Maynooth). Following the tradition in the literature, we didn't write these 743 fitted models as we do in OLS model, where there is only one resulting equation. (The 743 models could be written, but it would be tedious and time consuming to do all that. It is not mentioned in any books or papers yet.) We have not taken polynomial (quadratic, or interaction) terms into our account in the regression here.
--	--	---



SDI Review Form 1.6

<u>Minor</u> REVISION comments	It should be explained whether any preprocessing of the data was done.	We mentioned in the manuscript that imputation technique was used to handle the missing data. No further pre-processing was done.
<u>Optional/General</u> comments		