



SDI Review Form 1.6

PART 1:

Journal Name:	British Journal of Applied Science & Technology
Manuscript Number:	MS: 2012_BJAST_2698
Title of the Manuscript:	Data Analyzing by Attention to Weighted Multicollinearity in Logistic Regression Applicable in Industrial Data

General guideline for Peer Review process is available in this link:

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

- This form has total 9 parts. Kindly note that you should use all the parts of this review form.



SDI Review Form 1.6

PART 2: Review Comments

	Reviewer's comment	Author's comment <i>(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)</i>
<u>Compulsory</u> REVISION comments	1. Introduction is too long. The details of different methods on how to fit a logistic regression could be reduced and put in the section of method.	Some parts of the introduction are deleted. And some published papers are added. If I remove more, the concept of the introduction may be lost. Also some part of abstract are corrected.
	2. A table to compare the pros and cons of these different estimation methods would be very useful for the general readers.	It is added in the manuscript (highlighted with yellow color).
	3. The decimal points in almost every table were wrongly represented by commas, please correct them. The same applies to some equations.	They are corrected in the manuscript (highlighted with yellow color).
	4. Some discussion on how the parameter estimates from the different methods differ and why needs to be done as the manuscript mainly compares different methods.	Four methods are suggested for reducing the effect of multicollinearity between explanatory variables on logistic model. PC method produces PC components given by covariance structure between the explanatory variables, while PLS method produces the PLS components given by covariance structure between the explanatory and response variables. According to these methods, new components are orthogonal. After that they are used in model as a new explanatory variables. Also Ridge and Stein methods shrink both the estimated parameter vector and the associated standard errors by some different scaling techniques. These four methods are different because they



SDI Review Form 1.6

		have different procedures. They are the most practical and useful methods to solve multicollinearity problems. Data structure and assessment indicators (such as deviance and sum of coefficients variance) are very important to choose the best of these methods. For a few cases (data structures), may notice that none of these methods are totally satisfactory. Because of this reason all four methods are explained in this study and all of them are compared by paying attention to assessment indicators.
	5. The authors need to give a more conclusive recommendation on which model could be considered as the best model considering statistics and expert opinions.	It is added in the manuscript (highlighted with yellow color).
	6. Please note that most readers who are interested in reading the manuscripts may not be familiar with steel production.	The aim of this study is not to introduce the steel data set. We use it as a tool to investigate the effect of multicollinearity between explanatory variables on logistic regression model. Multicollinearity can be seen in different fields such as medicine, epidemiology and technology. We can solve this problem with mentioned methods in the manuscript.
<u>Minor</u> REVISION comments	Please cite more recently published papers.	They are added in the manuscript (highlighted with yellow color).
<u>Optional/General</u> comments		