Analysis of How Well Regression Models Predict Radiation Dose from the Fukushima Daiichi Nuclear Accident

Stephen U. Egarievwe[1*], Jamie B. Coble[2], Laurence F. Miller[2]

[1] Nuclear Engineering and Radiological Science Center, Alabama A&M University, Normal, AL 35762 USA.
[2] Nuclear Engineering Department, University of Tennessee, Knoxville, TN 37996 USA.

* Corresponding author. Tel.: +1-2563728952; email: stephen.egarievwe@aamu.edu

Manuscript submitted January 10, 2016; accepted July 23, 2016.
doi: 10.17706/ijapm.2016.6.4.150-164

Abstract: The 2011 Fukushima Daiichi nuclear accident in Japan resulted in the release of radioactive materials into the atmosphere, the nearby sea, and the surrounding land. Based on the International Atomic Energy Agency (IAEA) Convention on Early Notification of a nuclear accident, several radiological data were collected on the accident. Among the radioactive materials monitored, are I-131 and Cs-137 which form the major contributions to the contamination of drinking water. The radiation dose in the atmosphere was also measured. This study focused on how well regression models predict radiation dose from the following predictor variables: I-131and Cs-137 concentrations in drinking water, radiation monitoring locations, and distance and direction of monitoring points from the accident location. The analysis covered 1) the correlations between the radiation dose and the predictor variables, and 2) how well simple regression methods could predict the radiation dose. The modeling techniques investigated include linear regression, principal component regression (PCR), partial least square regression (PLS), ridge regression, and locally weighted regression. The Venetian Blinds method was used to divide the data into training, test, and validation datasets. The concentrations ofI-131 and Cs-137 directly determine the output parameter dose, and thus have better correlations compared to the other predictor variables. The linear regression model with one variable (I-131 concentration in drinking water) was found to be the best with a root mean square error of 0.0133. For the other models, the root mean square errors are0.0148 for ridge regression cross validation,0.0198 for ridge regression L-curve, 0.0210 for PCR,0.0856 for PLS, 0.0892 for locally weighted linear regression, and 0.0993 for locally weighted kernel regression.

Key words: Nuclear accident, partial least square regression, principal component regression, radiation dose, radioactive materials, regression models, ridge regression.

1. Introduction

Radioactive materials released during nuclear accidents pose danger to the environment and people. The 2011 Fukushima Daiichi nuclear accident that happened in Japan resulted in the release of radioactive materials into the atmosphere, the nearby sea, and the surrounding land [1]. The major radioactive materials include I-131, Cs-137, Cs-134, Te-129m, Sr-90, and Pu isotopes [2]–[5]. While the surrounding areas, where people were thought to be at risk, were evacuated, other regions could still be at risk due to the dispersion of the radioactive materials released. Radioactive materials are dispersed through air, land and water. Based on the International Atomic Energy Agency (IAEA) Convention on Early Notification of a nuclear accident, several radiological data were collected, by the Japanese authorities, on the accident.

This paper analysis how well regression models predict radiation dose from the following predictor variables: I-131 and Cs-137 concentrations in drinking water, radiation monitoring locations, and distance and direction of monitoring points from the accident location.The goal of this study focused on investigating 1) the correlations between the radiation dose and the other parameters, and 2) how well simple regression methods could predict the radiation dose. The modeling techniques investigated include linear regression, principal component regression (PCR), partial least square regression (PLS), ridge regression, and locally weighted regression (kernel regression and local linear regression).

2. Methodology

2.1. Dataset and Variables

The data used in this study were obtained from the Fukushima Monitoring Database in the IAEA website [1]. The variables in the dataset are listed in Table 1. Each variable has 103 observations. The first variable is the prefecture where the data were recorded. The prefectures from which data were extracted are Chiba (Ichihara), Gunma (Maebashi), Ibaraki, Iwate (Morioka), Tochigi (Utsunomiya), Tokyo (Shinjyuku), Yamagata (Yamagata), and they were assigned identification numbers 1 – 7, respectively, in this investigation. The fifth variable is the direction of the radiation monitoring point from Fukushima (ground zero of the accident), and for this investigation they were assign numbers as follow: North (N) = 1, South West (SW) = 2, South West West (SWW) = 3, South South West (SSW) = 4, and North West (NW) = 5. Variables 2, 3, 4 and 6 are as stated in Table 1 along with their units of measurements. The predictor variables are location, I-131 concentration, Cs-137 concentration, distance from Fukushima, and direction from Fukushima. The output variable is radiation dose.

Table 1. Data and Variables

	Variable	Type
1.	Location – monitoring point (Prefecture)	Predictor
2.	I-131 concentration in drinking water (Bq/kg)	Predictor
3.	Cs-137 concentration in drinking water (Bq/kg)	Predictor
4.	Distance from Fukushima (km)	Predictor
5.	Direction from Fukushima	Predictor
6.	Radiation dose (μSv/h)	Output

2.2. Training, Test, and Validation Sets

The Venetian Blinds method was use to divide the data into training, test, and validation sets. The 103 observations were grouped into nine sets of twelve observations each except the last set which has seven. The allocations of the groups of observations to the training, test, and validation sets are shown in Table 2. The training set is assigned 53.4% of the data while the test and validation sets have 24.3% each.

Table 2. Training, Test, and Validation Datasets

Observations	Dataset
1 – 12	Training
13 – 24	Test
25 – 36	Training
37 – 48	Validation
49 – 60	Training
61 – 72	Test
73 – 84	Training
85 – 96	Validation
97 – 103	Training

2.3. Regression Models

While there are complex computer models for estimating radiation dose, such as RESRAD [6], this study focused on basic regression models and how well they predict radiation dose from concentrations of released radioactive materials. The modeling techniques used in this investigation include simple linear regression, principal component regression (PCR), partial least square regression (PLS), ridge regression, and locally weighted regression (LWR). These models have different advantages in terms of their abilities to model the set of data [7].

2.4. Simple Linear Regression

In simple linear regression model, using matrix notation [8], we have

$$y = Xw + \varepsilon \tag{1}$$

where y is a vector (nx1) of the samples of the response variables, X is a matrix (nxp) of predictor variables of n observations (the rows) and p variables (the columns), w is the weight matrix (px1) that linearly combines the predictors to form the response, and ε is a vector (nx1) of the predictor errors. The least square approach for solution solves for an optimum weight, w, with the assumptions that 1) the system is actually linear, 2) there are no errors in the measured values of the predictors, 3) all inputs are independent, 4) all inputs are available, and 5) the errors are homoscedastic, independent and normally distributed.

2.5. Principal Component Regression (PCR)

In PCR, we first transform the data, through principal component analysis (PCA), into a new coordinate system with orthogonal axes that form the principal components (PC) or loadings p, which are then used in the regression process. See Fig. 1. The first PC contains the maximum variance of the dataset. The second PC contains the second most variance of the dataset, and so on. The concept and computational technique for PCR are discussed in more details by Fekedulegn *et al.* [8].The PCA involves the following steps:

1) Collect and standardize the data.
2) Find the covariance matrix for the processed data.
3) Calculate the eigenvalues and eigenvectors of the covariance matrix. The eigenvectors are the PCs.

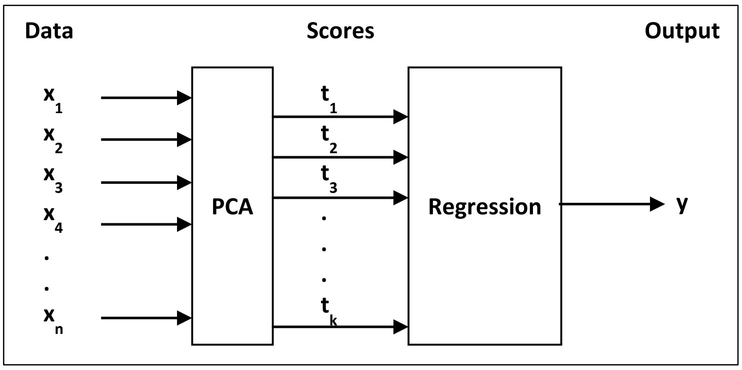


Fig. 1. Block diagram illustrating principal component regression. In PCA, the data is transformed into a new coordinate system with new components that are used in the regression.

The problems that are inherent in simple linear regression with collinear data are avoided in PCR. The PCR process uses a set of orthogonal, i.e. uncorrelated, inputs (the PC scores). Another advantage of the PCR is that the inputs in the model may be a subset of the PCs. Also, PCS could improve stability issues. When choosing the PCs to use in the regression, we take PCs that contain most of the information either by selecting those that explain up to a certain percentage (about 80% to 90%) of the data or by finding the PCs with eigenvalues that are greater than 1. We choose PCs that are well correlated with the response variable.

2.6. Partial Least Square (PLS) Regression

Partial least square regression is a factor based technique used to perform multi-linear regression. The inputs and outputs are transformed into uncorrelated factors called latent variable. As illustrated in Fig. 2, a linear mapping *b* is performed between the score vectors *t* and *u* on the latent variables *p* (of the inputs) and *q* (of the outputs).

The latent variables in the PLS algorithm give the maximal reduction in the covariance X^TY of the data. The PLS algorithm decomposes the input and output data into latent variables *p* and *q* respectively. Cassel *et al.* [9] discussed the PLS algorithm and the data generation in more details.

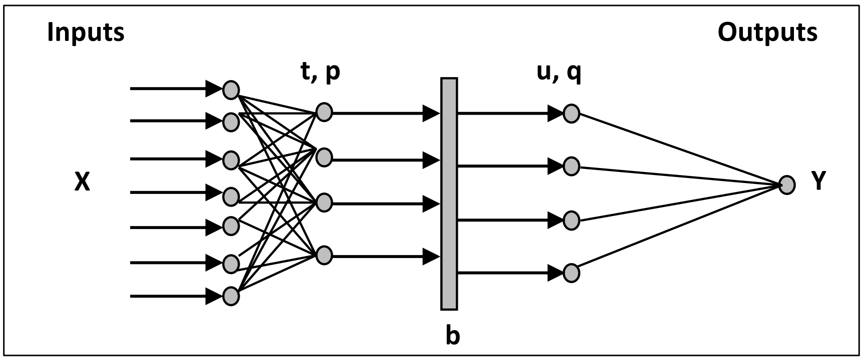


Fig. 2. Block diagram illustrating partial least square regression. A linear mapping *b*, is performed between the score vectors *t* and *u* on the latent variables *p* (of the inputs) and *q* (of the outputs).

Like in PCA, the coefficient vector to transform the measured variables to a latent variable is called the loading. The values of an observation in the latent variable space form the score vector. In PCR, the PCA focuses on the variance of the input data while PLS focuses on the correlation matrix between the inputs and the outputs. PLS transforms the inputs to explain both the variance in the input space and the covariance of the inputs with the output. PCA is an unsupervised method (it is only concerned with the inputs) while PLS is a supervised algorithm (its transformation is governed by the output).

2.7. Ridge Regression

Ridge regression is used in ill-conditioned problems to ensure that models give a better representation of the underlying process instead of noise in the training data. It balances fidelity to the data and fidelity for a priori knowledge that the process should be smooth. The cost function of the initial ill-conditioned problem can be represented by

$$Q = Y^T Y - 2b^T X^T Y + b^T X^T X b + \alpha^2 b^T b \tag{2}$$

where *X* is an *n*x*p* matrix of *p* predictors at each of *n* observations, *Y* is an *n*x1 vector of observed responses, *b* is vector of parameters (coefficients) and α is the ridge parameter. After simplification and differentiation we get the normal equations for regularized cost function:

$$(X^T X + \alpha^2 I)b = X^T Y \tag{3}$$

In (3), *I* is the identity matrix. Solving for *b* gives

$$b = (X^T X + \alpha^2 I)^{-1} X^T Y \tag{4}$$

This least square solution for the regularized cost function includes the cost for a small norm which makes the solution smooth. While there are several methods for the selection of the ridge regularization parameter, two methods are used here: the L-curve [10] and cross-validation [11], [12].

2.8. Locally Weighted Regression (LWR)

Locally weighted regression (LWR) is a non-parametric memory-based method of performing regression around a point using training data that are in the immediate region of that point [13]–[15]. The key features include 1) the storage of past data exemplars in memory for use in future queries, and 2) construction of a local model in the region of interest using appropriate method to determine the distance between the query and the training observation. The LWR algorithm has two steps for each new query:

Step 1: Locate training observations (exemplars) in the vicinity of the query.

Step 2: Perform a weighted regression with these nearby observations.

In step 1, the exemplars are weighted with respect to their proximity to the query point. This process involves 1) quantifying the distance between a training observation and the query, and 2) converting the distance to a similarity metric. Some of the methods used for this step are nearest neighbor, weighted averaging, kernel regression, and locally weighted regression. An example of nearest neighbor method is the use of Euclidian distance *d*:

$$d(x,q) = \sqrt{\sum_j (x_j - q_j)^2} \tag{5}$$

where *q* the query and *x* is the observation.

In the weighted average, each training data point is weighted by the inverse of the distance to the query point:

$$\hat{y}(q) = \frac{\sum\limits_{i=1:n} (y_i * w_i)}{\sum\limits_{i=1:n} w_i} \tag{6}$$

where

$$w_i = \frac{1}{d(x,q)} \tag{7}$$

In this analysis, we used Kernel regression. Kernel regression is the generalization of the weighted average method:

$$\hat{y}(q) = \frac{\sum\limits_{i=1:n} y_i w_i}{\sum\limits_{i=1:n} w_i} = \frac{\sum\limits_{i=1:n} y_i K(d(x_i,q))}{\sum\limits_{i=1:n} K(d(x_i,q))} \tag{8}$$

where the weight is given by a kernel operator *K* of the distance measure:

$$w_i = K(d) \tag{9}$$

Finding the kernel bandwidth involves the following steps:

Step1: Standardize the data.

Step 2: Select exemplars from the training data.

Step 3: Evaluate performance of a variety of bandwidths on the test data.

Standardizing the data makes all variables equally important in the distance measure. We can bound the expected range of distances, so we can bound our kernel bandwidth.

The weighted least squares regression equation can be solved for the optimal estimates of the regression coefficients. We also used the local linear regression in this analysis. Consider the following linear model:

$$y = X\beta + \varepsilon \tag{10}$$

where *y* is a vector (column matrix *n*x 1) of samples of the response variable, *X* is a matrix (*n*x*p*) of predictor variables (the columns are the variables and the rows are the observations), β is the vector of regression coefficients (*p*x1) that linearly combines the predictors to form the response, and ε is a vector (*n*x1) of the prediction errors. In LWR, we minimize a weighted sum square error (SSE) where the weighting factor considers the distance between the training exemplar and the query observation. The optimization function around the query is

$$Q(q) = \sum_{i}^{n} (y_i - x_i\beta)^2 K(d(x_i, q)) \tag{11}$$

We can solve for the optimal estimates of the regression coefficients:

$$\begin{aligned} \hat{\beta} &= (X^T W^T W X)^{-1} X^T W^T W X \\ &= (Z^T Z)^{-1} Z^T v \end{aligned} \tag{12}$$

where *Z* = *WX,* v = *WY*, and *W* is a diagonal matrix with the diagonal equal to the square roots of the kernel function:

$$w_{ii} = \sqrt{K(d(x_i, q))} \tag{13}$$

2.9. Qualitative Comparison of the Algorithms

In the ordinary least square (OLS) model, if the input matrix has near linear dependent columns (co-linear inputs), the least square solution is not unique and is unstable under small perturbations of the data. Every time we run the model we always have a different solution. The solutions may be close but they are never the same. These problems could lead to large regression coefficients that in theory may be giving the same result, but in practice adds to noise in the data. We however would like to minimize the effect of noise as much as possible; that is, we want to get small regression coefficients. One may suggest using principal component regression (PCR) for dealing with ill-conditioned regression problems, but it is a hard threshold method. The principal components (PCs) associated with the eigenvalues smaller than a predetermined certain tolerance are dropped by making them equal to zero. While PCR is appropriate for problems with a clear gap between two eigenvalues, ridge regression is more appropriate for problems without a clear gap. Unlike PCR, ridge regression dampens the minor components instead of completely removing them.

The major advantage of LWR is that it does not require a particular function to model the entire data set, unlike ordinary least square (OLS) that requires modeling the whole input space with a parametric model. Thus, LWR is more flexible and models are constructed as needed. Due to its features, LWR is more ideal for modeling complex systems that require more than one function to model. One disadvantage of LWR over OLS is that it makes less efficient use of the entire data input space. Also, unlike OLS, LWR does not produce any reusable regression function.

3. Results

3.1. Data Statistics

The data statistics is summarized in Table 3. The normal probability plots of the variables are shown in Fig. 3. The result to check for outliers is shown in Fig. 4. The significant outlier in the distance corresponds to Tokyo, which is a far distance from Fukushima compared to the other data collection points.

Table 3. Analysis of the Data

	Variable	Maximum	Minimum	Mean	Median	Variance	Standard Deviation
1.	Location – monitoring point	7.0	1.0	3.8	5.0	3.6	1.9
2.	I-131 concentration in drinking water(Bq/kg)	110.0	0.2	9.4	4.6	260.8	16.1
3.	Cs-137 concentration in drinking water(Bq/kg)	18.0	0.1	1.9	0.6	7.9	2.8
4.	Distance from Fukushima (km)	250.0	110.0	187.2	206.0	1880.5	43.4
5.	Direction from Fukushima	5.0	1.0	2.8	2.0	00.8	0.9
6.	Radiation dose (μSv/h)	0.3	0.0	0.1	0.1	0.0	0.1

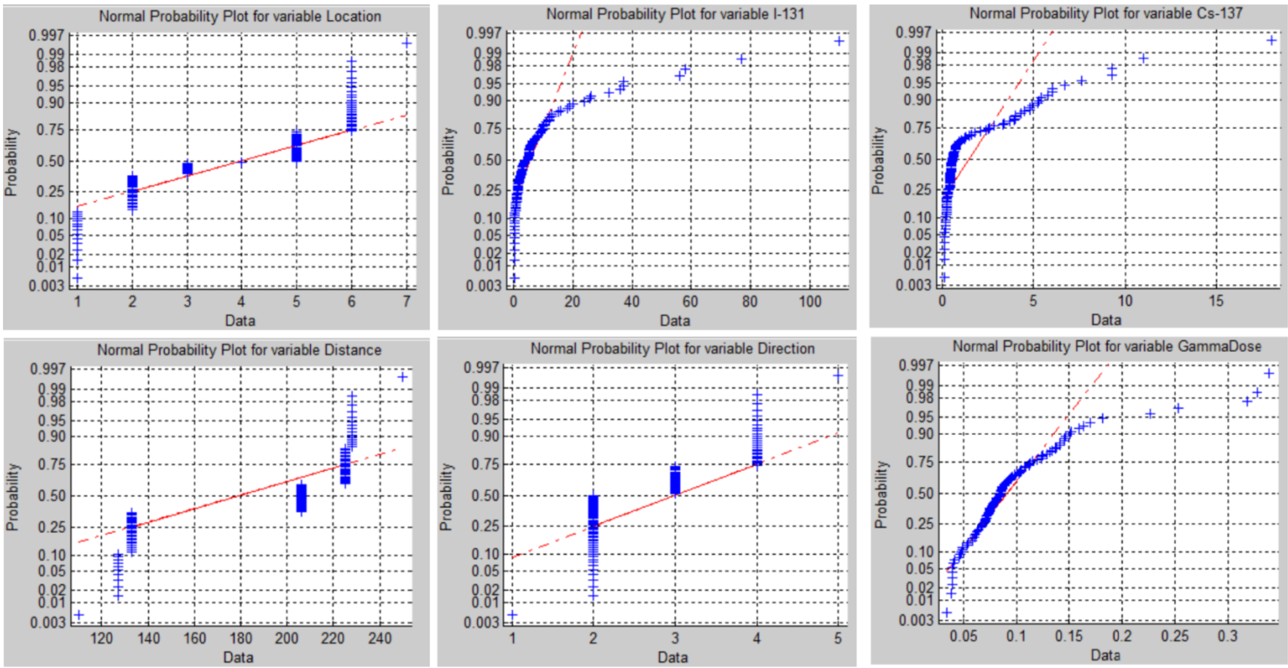


Fig. 3. The normal probability plots of the variables.

3.2. Simple Linear Regression

The training data were checked to see if it includes the maximum and minimum values in each of the predictor variables. Any missing maximum and minimum values were then included in the training dataset. This is to avoid extrapolating when data values are out of range. The correlation coefficients between the output variable and the predictor variables are shown in Table 4. The most correlated predictor variables, having correlation coefficients greater than |0.5|, are I-131 concentration in drinking water, Cs-137 concentration in drinking water, and the distance of monitoring point from Fukushima. The positive correlation values for I-131 and Cs-137 means an increase in radiation dose is associated with increase in their concentrations in drinking water. The negative correlation value for *distance*(see Table 4) means that as the distance of the monitoring point from Fukushima increase, the radiation dose decreases.

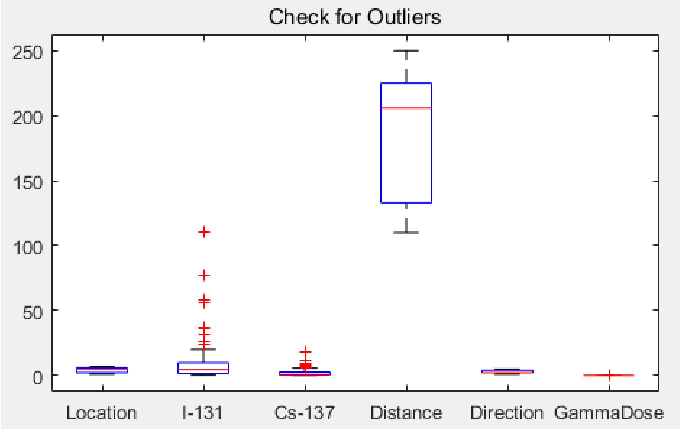


Fig. 4. Outliers. The significant outlier in the distance corresponds to Tokyo, which is a far distance from Fukushima compared to the other radiation data collection points.

Table 4. Data and Variables

	Variable	Correlation Coefficient
1.	Location – monitoring point (Prefecture)	0.4645
2.	I-131 concentration in drinking water (Bq/kg)	0.6962
3.	Cs-137 concentration in drinking water (Bq/kg)	0.6215
4.	Distance from Fukushima (km)	-0.5090
5.	Direction from Fukushima	0.1157
6.	Radiation dose (μSv/h)	1.0000

The root mean square error (RMSE) for each predictor variable in predicting the response variable (gamma dose) is shown in Table 5, along with those of using 1) all input variables, 2) a combination of input variables, and 3) the square and log of input variables. The regression with I-131 concentration (variable 2) as input gave the best result with a RMSE of 0.0133.

Table 5. RMSE for Each Variable in Predicting Gamma Dose

Variable	Correlation Coefficient
All impute variables	0.0201
Highly correlated variables (correlation ≥ 0.6)	0.0141
Well correlated variables (correlation ≥0.5)	0.0196
Variable 1 (Location – monitoring point)	0.0171
Variable 2 (I-131 concentration in drinking water)	0.0133
Variable 3 (Cs-137 concentration in drinking water)	0.0215
Variable 4 (Distance from Fukushima)	0.0238
Variable 5 (Direction from Fukushima)	0.0157
Quadratic of all variables	0.0240
Log of well correlated variables (correlation≥ 0.5)	0.0160

The next best linear regression, with RMSE of 0.0141, is that with the input of the highly correlated variables (correlation ≥ 0.6) which are I-131 and Cs-137 concentrations in drinking water. The actual and predicted values for the best model (simple regression with one variable, the I-131 concentration) are shown in Fig. 5.

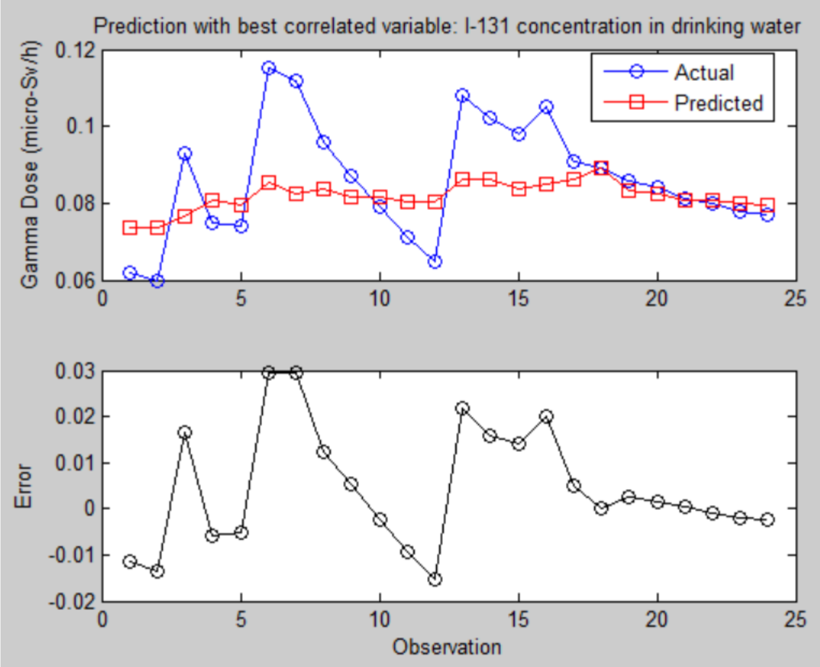


Fig. 5. The best model to predict Gamma dose, which has the lowest RMSE (0.0133), is the simple regression with one variable, the I-131 concentration in drinking water.

3.3. Principal Component Regression (PCR)

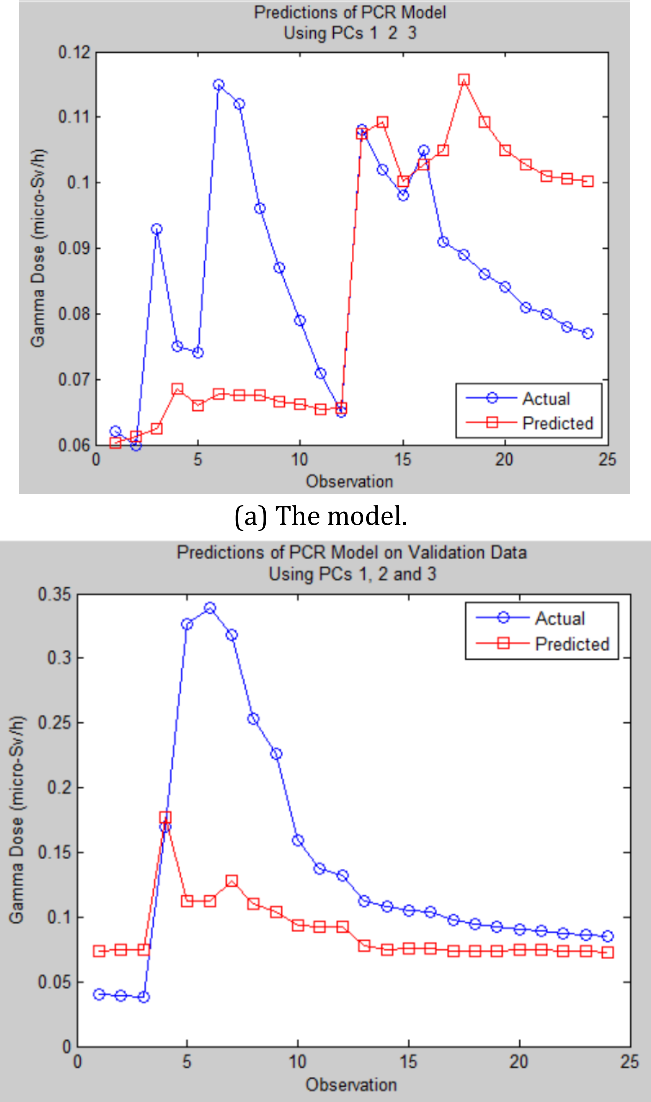
Table 6 shows the eigenvalue (latent) of each PC and the cumulative percentage of data (information) explained. The results of the PCR for several combinations of PCs are summarized in Table 7. This shows that PCR3, with PCs 1, 2 and 3 has the smallest RMSE (0.0210). The model and the validation result are shown in Fig. 6.

Table 6. Principal Components

	Eigenvalues (Latent)	Cumulative % Explained
PC_1	2.3379	46.7576
PC_2	1.4164	75.0863
PC_3	0.7166	89.4184
PC_4	0.3382	96.1833
PC_5	0.1908	100.0000

Table 7. RMSEs of PCR for Several Combinations of PCs

	PCs	RMSE
PCR1	1, 2, 3, 4	0.0213
PCR2	1, 2	0.0242
PCR3	1, 2, 3	0.0210
PCR4	1, 2	0.0242
PCR5	1, 2, 5	0.0226



(a) The model.

(b) The validation results.

Fig. 6. The PCR model with the smallest RMSE of 0.0210.

3.4. Partial Least Square (PLS) Regression

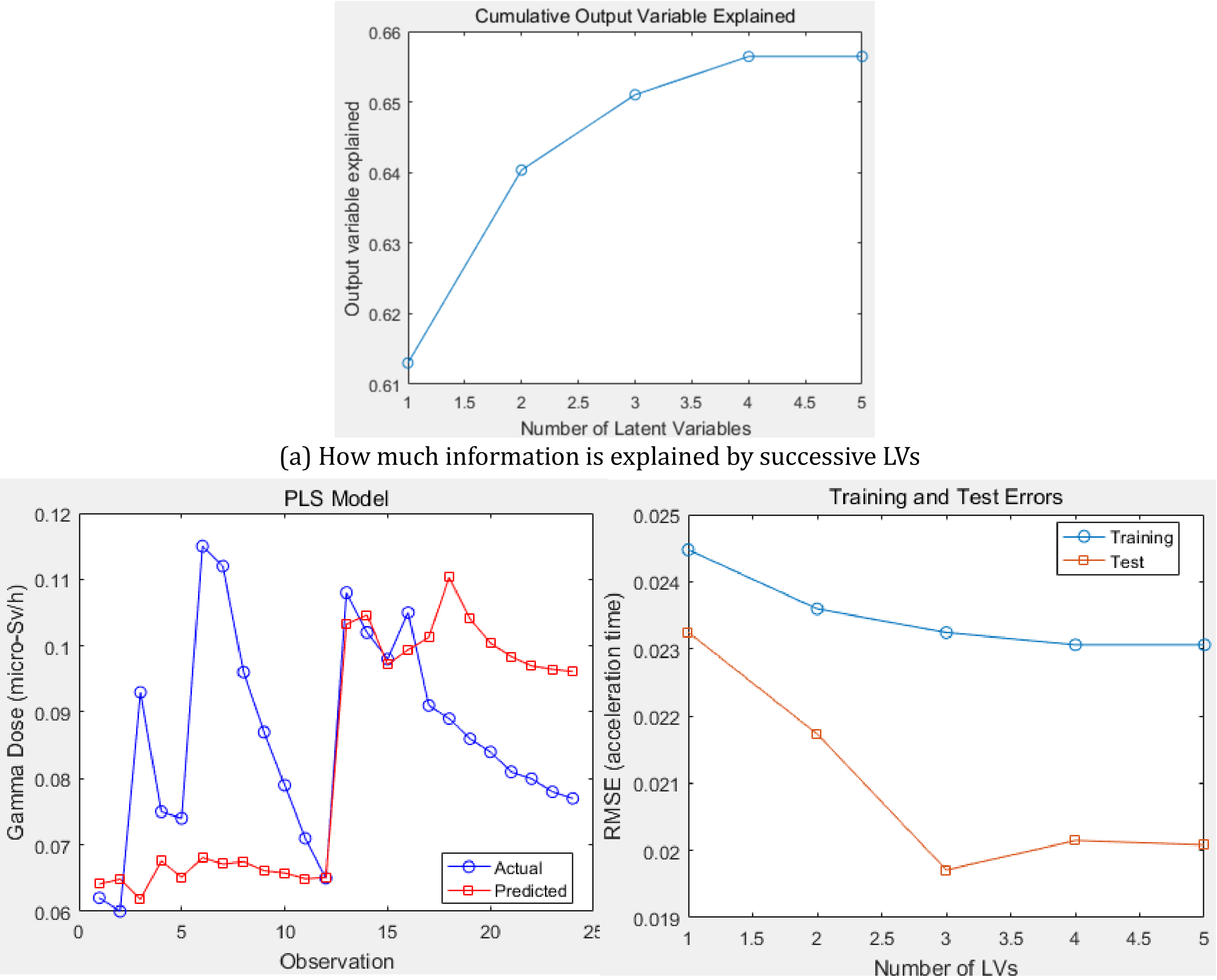
The cross validation method used in the PLS regression involved testing every possible latent variable. The variables were trained and the best latent variables were picked. The correlation coefficients of the PCs and LVs with the output are shown in Table 8.

Table 8. Correlation Coefficients of the PCs and LVs with the Output

PC/LV Number	PC Correlation Coefficient	LV Correlation Coefficient
1	0.7157	0.7830
2	0.3180	0.1653
3	0.1385	0.1033
4	0.0138	0.0736
5	-0.1540	0.0037
6	1.0000	1.0000

The cumulative output variable explained, the PLS regression, and the training and test errors of the cross validation LVs are shown in Fig. 7. As expected, the training error is always decreasing. We can see from Fig.

7 that the test error continue to decrease up to LV3 and then continue to increase after that. Thus, the number of useful latent variables is 3. The RMSE for the PLS regression is 0.0856.



(a) How much information is explained by successive LVs

(b) PLS regression.

(c) Training and test errors of the LVs.

Fig. 7. Results of the PLS model.

3.5. Ridge Regression

Prior to determining the regularization coefficient, the data were standardized. We test regularization parameters that cover the range of singular values. In this case, for convenience, we used ridge parameter α in the range of 0.01 – 100. For the L-Curve method, the plot of the RMSE vs norm of b is shown in Fig. 8. From the solution norm vs mean square error plot, we can determine that the best solution appear around $||b||$ = 0.55. The regularization coefficients obtained with the L-curve method are: optimum α = 2.9151 and condition number = 7.2173.The RMSE of the L-Curve method is 0.0198.

For the cross-validation method, the plots of root mean square error vs alpha for the training data and the test data are shown in Fig. 9a. Here, we have optimum α = 20.5651 (corresponding to the minimum root mean square error) and condition number = 1.2725. The RMSE of the CV method is 0.0145. The comparison of the L-Curve, CV and linear regression are shown in Fig. 9b. The L-Curve, CV and linear regression (using all variables) have RMSEs of 0.0198, 0.0145 and 0.0201 respectively. Thus, the CV method (having the smallest RMSE) performs best in ridge regression for this dataset.

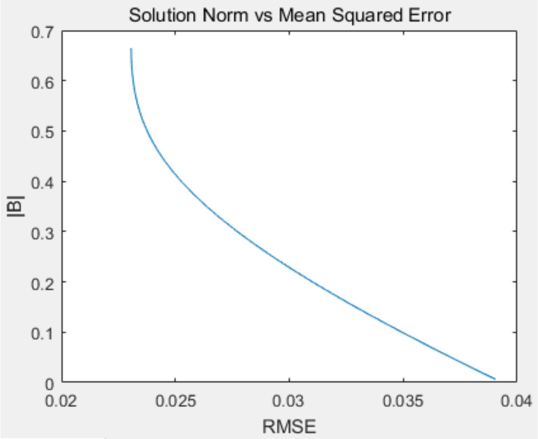
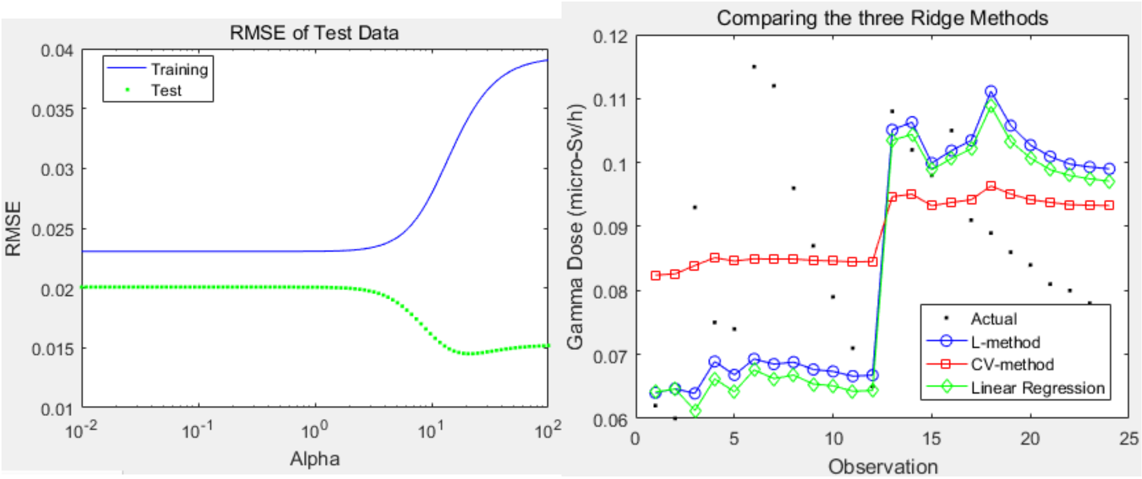


Fig. 8. L-curve method: The solution norm vs mean square error plot shows that the best solution appear around ||*b*|| = 0.55.



(a) Plots of root mean square error vs alpha. (b) Ridge regressions: L-Curve, CV and linear.

Fig. 9. Ridge regression results: L-Curve, CV and linear (using all variables) methods have RMSEs of 0.0198, 0.0145 and 0.0201 respectively.

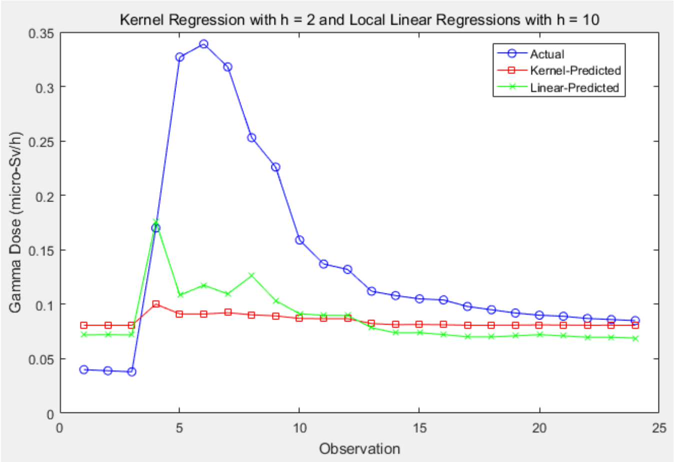


Fig. 10. Locally weighted regression using kernel regression and linear regression.

3.6. Locally Weighted Regression

The results of the locally weighted regression using kernel regression and linear regression are shown in

Fig. 10. The two locally weighted regression methods investigated are kernel and linear techniques. A visual observation of Fig. 10 shows that the shape of the predicted dose by the locally weighted linear regression resembles that of the actual gamma dose when compared to the prediction by locally weighted kernel regression. The RMSE for the local kernel regression is 0.0993 and that for the local linear regression is 0.0892. This implies that the local linear weighted method performs better than the kernel method for predicting radiation dose from the set of predictor variables.

Table 9. Comparison of the Regression Models

Model	RMSE
Linear Regression: One variable (I-131)	0.0133
Linear Regression: Two variables (I-131 and Cs-137)	0.0141
Linear Regression: All variables	0.0201
Principal Component Regression (PCR)	0.0210
Partial Least Square (PLS) Regression	0.0856
Ridge Regression: L-Curve	0.0198
Ridge Regression: Cross-validation	0.0145
Locally Weighted Regression: Kernel	0.0993
Locally Weighted Regression: Linear	0.0892

4. Conclusion

The performance summary of all the regression methods investigated are shown by the RMSEs listed in Table 9. For the set of data investigated in this study, the linear regression model with one variable (I-131 concentration in drinking water)was found to be the best with a root mean square error of 0.0133. Adding the Cs-137 concentration to the linear regression gave a RMSE of 0.0141. The Cs-137 concentration could have an added value to the linear regression model, and it increased the RMSE only by 0.0008. For the other models, we have from best to worst as follow: ridge regression, principal component regression, partial least square regression, and locally weighted regression. In ridge regression, the cross-validation method performed better that the L-curve. In locally weighted regression, the locally weighted linear method performed better than the kernel technique.

Our analysis has shown the limitations of these regression techniques in predicting the radiation dose from I-131and Cs-137 concentrations in drinking water, radiation monitoring locations, and distance and direction of monitoring points from the accident location. The use of a more complex model, such as neural network, could give a better prediction.

Acknowledgment

This work was supported in part by the U.S. Nuclear Regulatory Commission under Grant No. NRC-27-10-514 and the U.S. Department of Homeland Security Domestic Nuclear Detection Office under competitively awarded contract/IAA award number 2012-DN-077-ARI065-04.

References

[1] Fukushima Monitoring Database, IAEA. Retrieved from the website: https://iec.iaea.org/fmd/default.aspx

[2] Normile, D. (2011, May). Fukushima revives the low-dose debate. *Science, 332(6032)*, 908–910.

[3] Hoeve, J. H., & Jacobson, M. Z. (2012). Worldwide health effects of the Fukushima Daiichi nuclear accident. *Energy & Environmental Science, 5(9)*, 8743–8757.

[4] Wolf, A. (2012, July). Response to an unexpected mortality increase in the United States follows arrival of the radioactive plume from Fukushima: Is there a correlation? *International Journal of Health Services, 42(3)*, 549–551.

[5] Mangano, J. J., & Sherman, J. D. (2012). An unexpected mortality increase in the United States follows arrival of the radioactive plume from Fukushima: Is there a correlation? *International Journal of Health Services, 42(1)*, 47–64.

[6] Yu, C., Zielen, A. J., Cheng, J. -J., LePoire, D. J., Gnanapragasam, E., Kamboj, S., Arnish, J., Wallo III, A., Williams, W. A., & Peterson, H. (2001, July). *User's Manual for RESRAD Version 6.* Environmental Assessment Division, Argonne National Laboratory, Argonne, IL, USA.

[7] Egarievwe, S. U., Coble, J. B., & Miller, L. F. (2015, April). Analytics of radioactive materials released in the Fukushima Daiichi nuclear accident. Advancements in Nuclear Instrumentation Measurement Methods and their Applications Conference. Lisbon, Portugal, 20-24 April 2015.

[8] Fekedulegn, D. B., Colbert, J. J., Hicks, Jr., R. R., & Schuckers, M. E. (2002, September). Coping with multicollinearity: An example on application of Principal Components Regression in Dendroecology. *Research Paper NE-721*. United States Department of Agriculture (USDA) Forest Service.

[9] Cassel, C., Hackl, P., & Westlund, A. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics*, *26(4),* 435–446.

[10] Hansen, P. C. (2000). The L-curve and its use in the numerical treatment of inverse problems. *Computational Inverse Problems in Electrocardiology, Advances in Computational Bioengineering, 4,* 119–142.

[11] Golub, G. H., Heath, M., & Wahba, G. (1979, May). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics, 21(2)*, 215–223.

[12] Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistical Surveys, 4,* 40–79.

[13] Altman, N. S. (1992, August). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician, 46(3),* 175–185.

[14] Ruppert, D., & Wand, M. P. (1994, September). Multivariate locally weighted least squares regression. *The Annals of Statistics, 22(3),* 1346–1370.

[15] Cleveland, W. S., & Devlin, S. J. (1988, September). Locally-weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association, 83(403)*, 596–610.

Stephen U. Egarievwe is an associate professor in the Department of Electrical Engineering and Computer Science at Alabama A&M University, Normal, Alabama, USA. He is also the director of the Nuclear Engineering and Radiological Science Center at Alabama A&M University. He is guest scientist at Brookhaven National Laboratory since 2008 till date. He obtained his Ph.D. degree in applied physics from Alabama A&M University. His current research is in the development of semiconductor materials for applications in the detection of radiological and nuclear threats. He also does research in the areas of computational science, nanotechnology, and cyber security. He is a Stanford Linear Accelerator Center (SLAC) Scholar (2002 and 2004), and a 2012 Scholar of the Accreditation Board for Engineering and Technology (ABET) Institute for the Development of Excellence in Assessment Leadership (IDEAL). He is currently the secretary of the Interdisciplinary Consortium for Research and Educational Access in Science and Engineering (INCREASE), a consortium of universities with headquarters at Hampton University.

Jamie Baalis Cobleis an assistant professor in the Nuclear Engineering Department at the University of Tennessee, Knoxville, Tennessee, USA. She obtained her Ph.D. degree in nuclear engineering from the University of Tennessee. Dr. Coble's expertise is primarily in statistical data analysis, empirical modeling, and advanced pattern recognition for equipment condition assessment, process and system monitoring, anomaly detection and diagnosis, and failure prognosis. Dr. Coble is currently pursuing research in prognostics and health management for active components and systems. Her research interests expand on past work in monitoring and prognostics to incorporate remaining useful life estimates into risk assessment, operations and maintenance planning, and optimal control algorithms. Prior to joining the UT faculty, she worked in the applied physics group at Pacific Northwest National Laboratory. Her work there focused primarily on data analysis and feature extraction for detecting anomalies and degradation in large passive components, advanced active components, and nuclear fuel reprocessing systems.

Laurence F. Miller is professor emeritus in nuclear engineering at the University of Tennessee, Knoxville, Tennessee, USA. He worked for Westinghouse for six years, and received his Ph.D. from Texas A&M University (1976). He served on the faculty at the University of Tennessee (UTNE) from September of 1976 through June of 2015, and he is continuing with some teaching, research and service commitments. He taught reactor theory and nuclear engineering laboratory for about 25 years, developed the Radiation Protection program that begun in 1988 in collaboration with area professionals, and has chaired the committees of 83 M.S. and 20 Ph.D. students. During the past 10 years he published 20 journal papers and 31 conference summaries. His research focused on the nuclear fuel cycle and on radiation detection and measurement. He served as interim radiation safety officer for UT (April-June 2012), chaired the UT radiation safety committee for 18 years, was the primary author of two ABET self-studies for UTNE, and has served on a number of other departmental, college, and university committees.