Efficient Imputation of Incomplete Petrophysical Dataset through Random & Robust Sequential Imputation Algorithms

Conference Paper · February 2016

DOI: 10.13140/RG.2.1.2126.5687

2 authors:

Watheq J. Al-Mudhafar
University of Texas at Austin
148 PUBLICATIONS **419** CITATIONS

SEE PROFILE

Ali Al-Mudhafar
University of Basrah
7 PUBLICATIONS **2** CITATIONS

SEE PROFILE

Efficient Imputation of Incomplete Petrophysical Dataset through Random & Robust Sequential Imputation Algorithms

W. J. M. Al-Mudhafar, Louisiana State University & A. Al-Mudhafar, University of Basrah

The incomplete dataset in petrophysical properties especially core permeability measurements is a crucial step in reservoir characterization. It affects the multivariate statistical inference of these measurements leading to non-accurate prediction and less accurate geospatial modeling.

Consequently, two imputation algorithms were presented in this paper to comparatively predict the missing values of horizontal and vertical core permeability for a well in sandstone reservoir in a southern Iraqi oil field. These two methods are Random Imputation of Missing Data and Robust Sequential Imputation Algorithms. These algorithms have been applied based on the deductive statistical inference to impute the incomplete data.

The two algorithms above were illustrated its robustness was depicted by visualizing the data distribution before and after the imputation process with respect to the histograms and the vertical data distribution given the well depth. The results have shown that the Random Imputation of Missing Data and Robust Sequential Imputation Algorithms are efficient tools for accurate imputation of incomplete petrophysical data.

Introduction

The missing data in petrophysical properties is a common problem that should be efficiently handled to obtain accurate inference. The most conventional way is to consider the correlation between the porosity and permeability given the well log records; however, it might result in inaccurate prediction (Al-Mudhafer, 2013). Therefore, it is important to look for modern algorithms to accurately impute the missing data in order to capture the real property distribution and to capture the reservoir heterogeneity in the spatial modelling.

In this paper, comparisons between Random Imputation of Missing Data and Robust Sequential Imputation Algorithms were adopted to efficiently impute the incomplete horizontal and vertical core permeability for a well in sandstone reservoir, southern Iraqi oil field. The histogram and vertical distribution of data before and after the imputation were done for comparison between the two algorithms.

Petrophysical Data Review

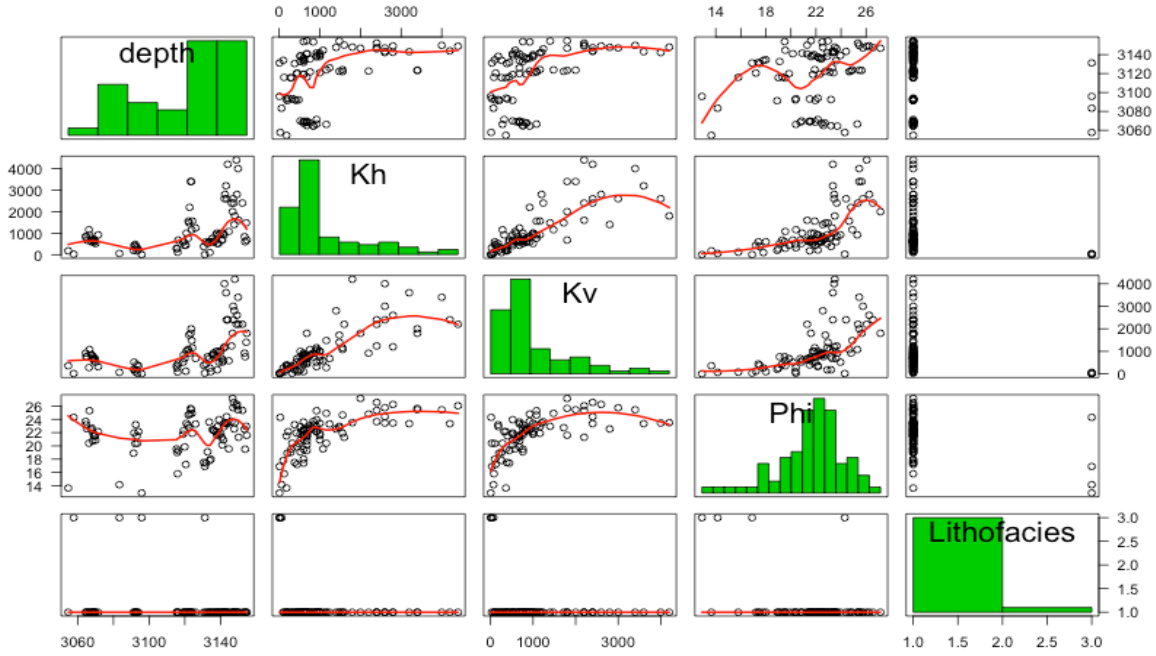

Figure 1: *Scatter Matrix Plot of Incomplete Petrophysical Data Parameters. It shows the distribution of each parameter in addition to the relationship between each two parameters.*

Random Imputation of Missing Data (RIMD)

It is the simplest procedure for data imputation for only one variable based on the observed data for this variable. The procedure is repeated for every variable in the dataset (Gelman and Hill, 2006). The imputed data that has been resulted from this algorithm has the most convenient outcomes compared to the other algorithms.

Robust Sequential Imputation Algorithm (impSeqRob)

Robust Sequential Imputation Algorithm was considered to impute the missing data. This algorithm starts from a complete subset of the dataset and estimates sequentially the missing values in an incomplete observation by minimizing the determinant of the covariance of the augmented data matrix. Then, the observation is added to the complete data matrix and the algorithm continues with the next observation with missing values (Verboven et.al, 2007).

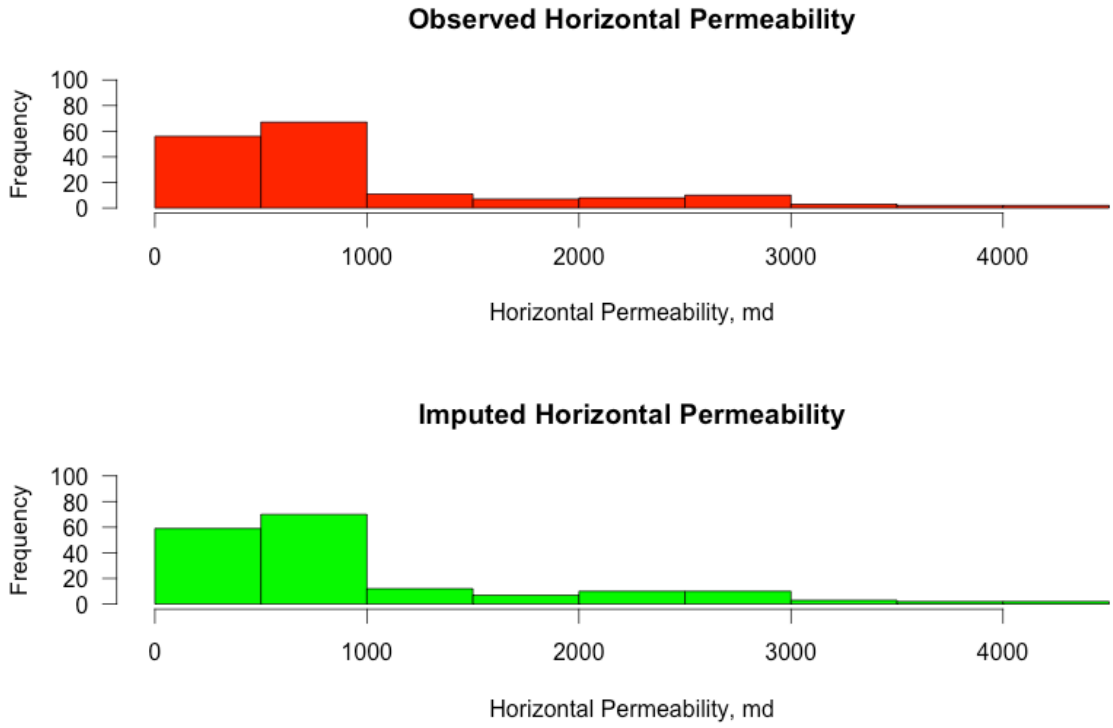

Figure 2: *Comparison of Histogram plots of observed (top) and imputed (bottom) Horizontal Permeability by Random Imputation of Missing Data (RIMD) Algorithm.*

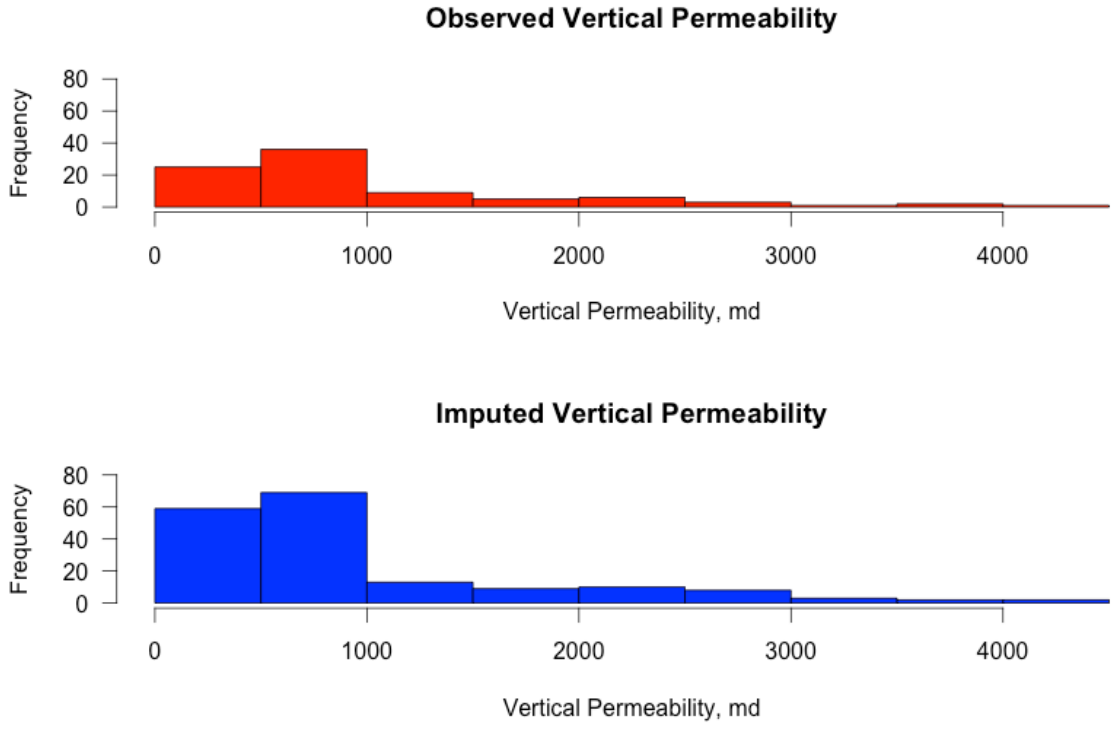

Figure 3: *Comparison of Histogram plots of observed (top) and imputed (bottom) Vertical Permeability by Random Imputation of Missing Data (RIMD) Algorithm.*

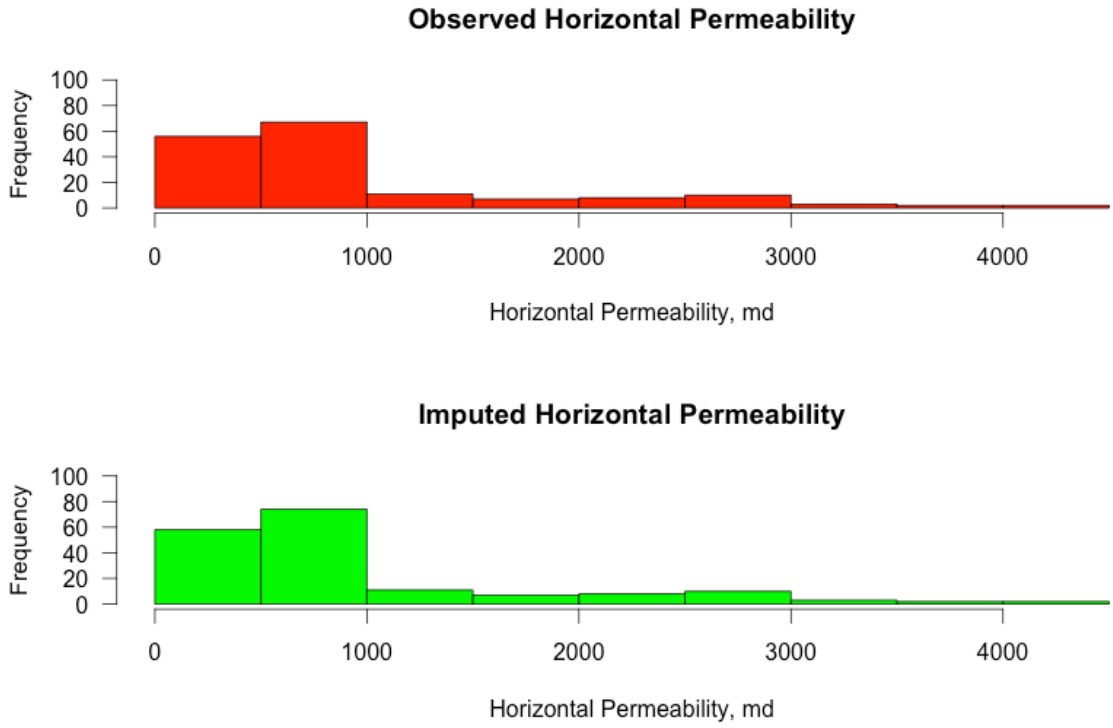

***Figure 4**: Comparison of Histogram plots of observed (top) and imputed (bottom) Horizontal Permeability by Robust Sequential Imputation (impSeqRob) Algorithm.*

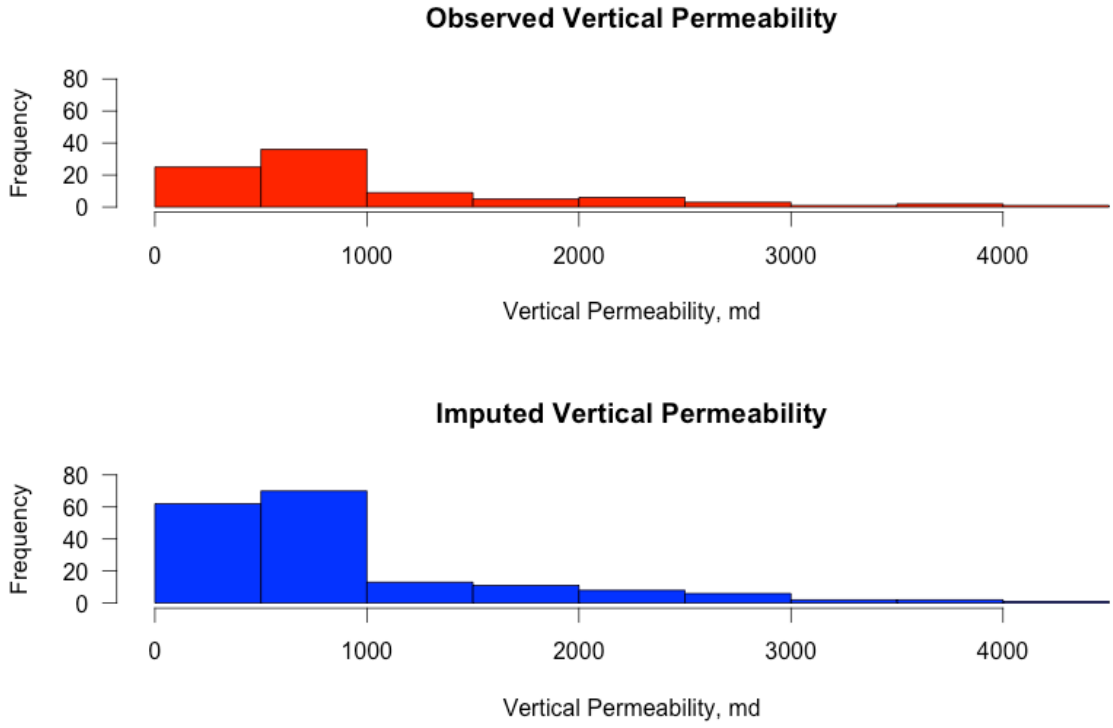

***Figure 5**: Comparison of Histogram plots of observed (top) and imputed (bottom) Vertical Permeability by Robust Sequential Imputation (impSeqRob) Algorithm.*

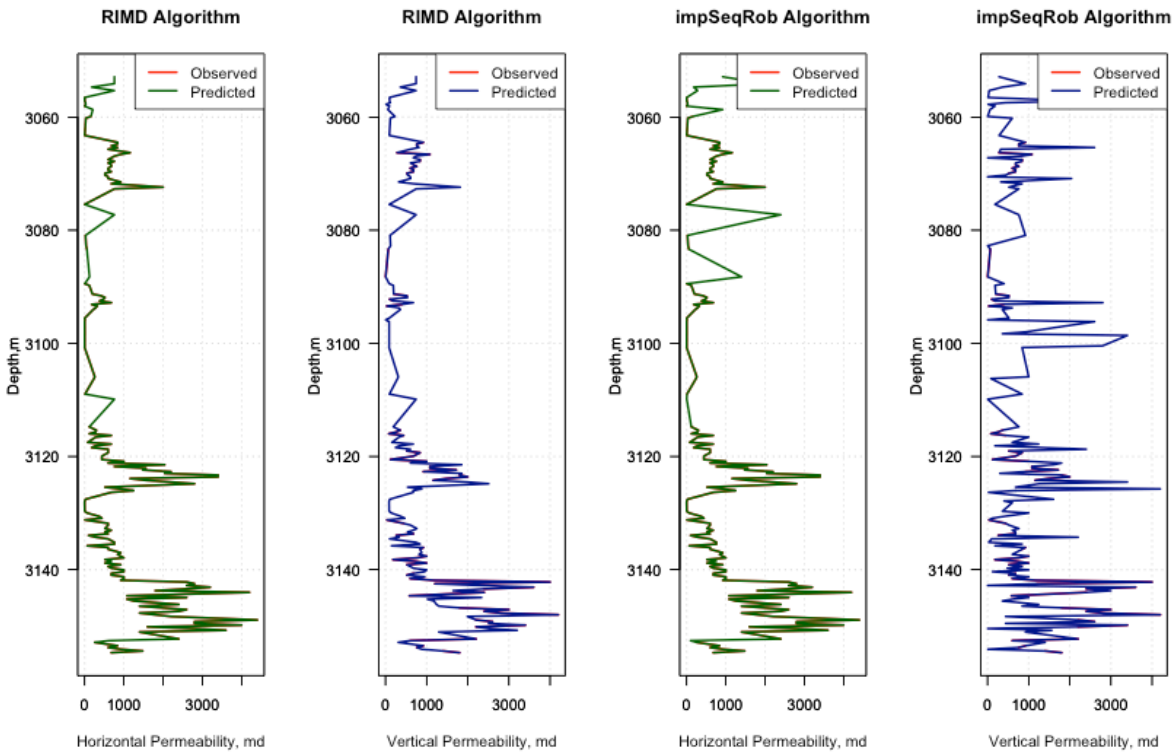

Figure 6: *Matching of horizontal & vertical permeability distributions between the observed and imputed data by Random Imputation of Missing Data (RIMD) Algorithm & Robust Sequential Imputation (impSeqRob) Algorithm. It can easily be noticed that there is excellent matching between the observed and predicted data after imputation through the two algorithms.*

Conclusions

From the diagnostic figures, it was easily noticeable that the Random Imputation of Missing Data and Robust Sequential Imputation Algorithms are very accurate tools to impute the incomplete horizontal and vertical core permeability. Additionally, the excellent matching between the vertical petrophysical data distributions by these two algorithms reflects their robustness to accurately imputation the missing data leading to efficient overall reservoir characterization.

References

1. Al-Mudhafer, W. (2013). Maximum Likelihood & Multiple Imputation of Incomplete Static and Dynamic Reservoir Data. *In Geoinformatics 2013,* Kiev, Ukraine.

2. Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. *Cambridge University Press.*

3. Verboven, S., Branden, K., Goos, P., (2007). Sequential imputation for missing values. Computational Biology and Chemistry 31, 320327.