

# Package: MERO (via r-universe)

August 22, 2024

**Title** Performing Monte Carlo Expectation Maximization Random Forest  
Imputation for Biological Data

**Version** 0.1.2

**Author** Mohamed Soudy [aut, cre]

**Maintainer** Mohamed Soudy <MohmedSoudy2009@gmail.com>

**Description** Perform missing value imputation for biological data using  
the random forest algorithm, the imputation aim to keep the  
original mean and standard deviation consistent after  
imputation.

**License** GPL-3

**Encoding** UTF-8

**RoxygenNote** 7.1.2

**Imports** missForest, ggpubr, progress, doParallel, foreach

**NeedsCompilation** no

**Date/Publication** 2023-02-24 16:40:02 UTC

**Repository** <https://mohmedsoudy.r-universe.dev>

**RemoteUrl** <https://github.com/cran/MERO>

**RemoteRef** HEAD

**RemoteSha** 4988b8ecf7ac3ac4c2372081e1c33dfc85b68ee3

## Contents

EvalImp . . . . .	2
MERO . . . . .	2
PlotCorrelateMean . . . . .	3
RMSE . . . . .	3
<b>Index</b>	<b>5</b>

---

EvalImp	<i>Evaluate the imputed data sets and select the best data set</i>
---------	--

---

**Description**

The function is evaluate the imputed data sets based on the mean and standard deviation

**Usage**

```
EvalImp(Originaldata, ImputedSets ,Imputed.mean, Imputed.sd)
```

**Arguments**

Originaldata	data frame of original data containing the missing values
ImputedSets	list of imputed data frames
Imputed.mean	data frame of the means of the imputed data sets
Imputed.sd	data frame of the standard deviations of the imputed data sets

**Value**

The best data frame which mean and standard deviation are close to the original data

**Author(s)**

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

---

MERO	<i>Perform Monte Carlo Expectation Maximization Random Forest Imputation</i>
------	--

---

**Description**

The function is used to impute the missing data using Monte Carlo Expectation Maximization Random Forest Imputation

**Usage**

```
MERO(Data, ntree = 100, Nsets = 5)
```

**Arguments**

Data	a data matrix with missing values. The columns correspond to the variables and the rows to the observations.
ntree	number of trees to grow in each forest.
Nsets	number of simulations/ data sets to be generated.

**Value**

A list containing data sets and imputed means, and imputed standard deviation.

**Author(s)**

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

---

PlotCorrelateMean	<i>Plot the correlation in scatter plot between original mean and imputed mean</i>
-------------------	--

---

**Description**

The function is used to plot the correlation between the imputed mean and original mean

**Usage**

```
PlotCorrelateMean(OriginalMean, ImputedMean)
```

**Arguments**

OriginalMean	means of the original data
ImputedMean	means of the imputed data

**Value**

The scatter plot

---

RMSE	<i>Calculate Root Mean Square Error 'RMSE' between vectors</i>
------	--

---

**Description**

The function is used to calculate the root mean square error between two vectors

**Usage**

```
RMSE(Actual, Predicted)
```

**Arguments**

Actual	Vector of actual data
Predicted	vector of predicted data

**Value**

The root mean square error between the two input vectors

**Author(s)**

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

**Examples**

`RMSE(c(1,2,3), c(10,20,30))`

# Index

EvalImp, [2](#)

MERO, [2](#)

PlotCorrelateMean, [3](#)

RMSE, [3](#)