Paquete ‘DiMBio’ una guía rapida

Julio 2023

Versión 1.0.1

Fecha 2023-07-18

Título Discriminación de Modelos Biológicos

Autores

DiMBio es una herramienta estadística que permite la discriminación de los resultados de modelos de simulación de procesos biológicos.

EL paquete incorpora estadísticas univariante y multivariantes, las cuales no presentan supuestos basados en que corresponden a la comparación de dos bases de datos entre si para discriminar la similitud o no de las mismas

Los estadísticos univariantes incorporados son: Coeficiente de determinación (R2), coeficiente de correlación de Pearson (ρ), test de Student para datos pareados (t), índice de eficiencia de Nash y Sutcliffe (NS), porcentaje de la raíz del error cuadrático medio (%RMSE), raíz del error cuadrático de la predicción (MSEP), raíz del error cuadrático medio de la predicción (RMSEP), error de la media absoluta (MAE), porcentaje del error de la media absoluta (%MAE), coeficiente de desigualdad de Theil (U), determinación de los errores de la regresión lineal (MC, SC y RC), máxima verosimilitud (Lj), diferencia del coeficiente de información de Akaike (∆AIC), teorema de Bayes (PMi/Y ),

Los estadísticos multivariantes corresponden a distancias, la distancia de Mahalanobis (D2), distancia Euclidean (dE) and distancia de Manhattan (dL1),

Para su ejecución:

Instale el paquete y ejecute en R para lo cual se pone a disposición en el directorio el mismo (DIMBIO\_0.2.0.tar.gz)

Install.packages("DIMBIO”)

library(DIMBIO)

#Cargar datos desde el formato en que fueron almacenados

data <- read.table("C:/……………/Data.xlsx")

#Datos Reales. Creación de un data.frame con solo las columnas que corresponden a un grupo de datos en este caso los experimentales o los simulados de un primer modelo

Reales <- data.frame(data$variable\_1, data$variable\_2, …. data$variable\_n)

#Datos Simulados. Creación de un data.frame con solo las columnas que corresponden a un grupo de datos en este caso los simulados o los simulados de un segundo modelo

Simulados <- data.frame(data$variable\_1s, data$variable\_2s, …. data$variable\_ns)

#Transformación a log de los datos reales y simulados

logReales<-log10(reales)

logSimulados<-log10(simulados)

#Determinación de estadísticos multivariados

Manhattan(logReales, logSimulados)

Euclidiana(logReales, logSimulados)

Mahalanobis(logReales, logSimulados)

#Determinación de los estadísticos univariantes, se identifican las variables comparadas

nombres <-c(“variable\_1”, “variable\_2”, …. “variable\_n”)

Univariantes(reales, simulados, nombres)

Ver ejemplo anexo