

Package ‘gdm’

January 9, 2016

Type Package

Title Functions for Generalized Dissimilarity Modeling

Version 1.1.7

Date 2015-01-08

Author Glenn Manion, Matthew Lisk, Simon Ferrier, Diego Nieto-Lugilde, Matthew C. Fitzpatrick

Maintainer Matthew C. Fitzpatrick <mfitzpatrick@al.umces.edu>

Description A toolkit with functions to fit, plot, and summarize Generalized Dissimilarity Models.

License GPL-2

Depends R (>= 2.15.2), Rcpp (>= 0.10.4), raster, reshape2, vegan, plyr

LinkingTo Rcpp

Imports

Suggests R.rsp

VignetteBuilder R.rsp

NeedsCompilation yes

Repository CRAN

Date/Publication 2016-01-09 16:14:18

R topics documented:

gdm-package	2
formatsitepair	3
gdm	7
gdm.transform	9
gdmDissim	11
gdmExpData	11
isplineExtract	12
plot.gdm	13
predict.gdm	15
stackedVars	16
summary.gdm	17

Index	19
--------------	-----------

Description

Generalized Dissimilarity Modeling is a statistical technique for modelling spatial variation in biodiversity between pairs of geographical locations. The **gdm** package currently provides basic functions to fit, summarize, and plot Generalized Dissimilarity Models and to make predictions (in both space and time) and map biological patterns by transforming environmental predictor variables. Future updates will incorporate support for genomic data.

Details

The functions in the **gdm** package provide the tools necessary for fitting GDMs, including functions to prepare biodiversity and environmental data. Major functionality includes:

- Formatting various types of biodiversity and environmental data to **gdm**'s site-pair format used in model fitting
- Fitting GDMs using geographic and environmental distances between sites
- Plotting fitted functions & extracting I-spline values
- Predicting pairwise dissimilarities between sites or times and transforming environmental predictors to biological importance and mapping these patterns.

To see the preferable citation of the package, type `citation("gdm")`.

I. Formatting input data

GDM fits biological distances to pairwise site geographical and environmental distances. Most users will need to first format their data to **gdm**'s site-pair table format:

formatsitepair	To convert biodiversity and environmental data to site-pair format
--------------------------------	--

II. Model fitting and summary

gdm	To fit a GDM model
summary	To summarize a GDM model

III. Model prediction and transformation of environmental data

predict	To predict biological dissimilarities between sites in space or between time periods
-------------------------	--

<code>gdm.transform</code>	To transform each environmental predictor to biological importance
----------------------------	--

IV. Plotting model output and fitted functions

<code>plot</code>	To plot model fit and I-splines
<code>isplineExtract</code>	To extract I-spline values to allow for custom plotting

Author(s)

The **gdm** development team is Glenn Manion, Matt Lisk, Simon Ferrier, Diego Nieto-Lugilde, and Matt Fitzpatrick. Where others have contributed to individual functions, credits are provided in function help pages.

The maintainers of the R version of **gdm** are Matt Fitzpatrick <mfitzpatrick@al.umces.edu> and Matt Lisk <mlisk@al.umces.edu>.

<code>formatsitepair</code>	<i>Converts Common Biological (Response) and Environmental (Predictor) Data Formats to a GDM Site-Pair Table</i>
-----------------------------	--

Description

This function takes input biological and environmental data in various formats and builds a site-pair formatted table required for fitting a Generalized Dissimilarity Model using the `gdm` function. NOTE: Sample site x-y coordinates MUST be present in either the biological or the environmental data.

The input biological data can be in one of the following four formats (note that "species" can optionally be replaced by any other classification of biological entities - e.g. "functional types", "haplotypes"):

1. site-species matrix
2. x, y, species list
3. site-site biological distance (dissimilarity) matrix
4. an existing site-pair table (see Details)

Environmental (predictor) data can be provided in three formats:

- a site-predictor matrix with a column for each predictor variable and a row for each site
- a raster stack, with one raster for each predictor variable
- one or more site-site distance predictor matrices using the "distPreds" argument (see below).

Usage

```
formatsitepair(bioData, bioFormat, dist="bray", abundance=F, siteColumn=NULL, XColumn,
YColumn, sppColumn=NULL, abundColumn=NULL, sppFilter=0, predData, distPreds=NULL,
weightType="equal", custWeightVect=NULL, samples=NULL)
```

Arguments

bioData	The input biological (response) data table, in one of the four formats defined above (see Details).
bioFormat	An integer code specifying the format of bioData. Acceptable values are 1, 2, 3, or 4 (see Details).
dist	Default = "bray". A character code indicating the metric to quantify pairwise site distances / dissimilarities. Calls the vegdist function from the vegan package to calculate dissimilarity and therefore accepts any method available from vegdist .
abundance	Default = FALSE. Indicates whether the biological data are abundance data (TRUE) or presence-absence (0, 1) data (FALSE).
siteColumn	The name of the column in either the biological or environmental data table containing site codes/names. If a site column is provided in both the biological and environmental data, the column name must be identical.
XColumn	The name of the column containing x-coordinates of sample sites. X-coordinates can be provided in either the biological or environmental data tables, but MUST be in at least one of them. If an x-coordinate column is provided in both the biological and environmental data, the column name must be identical.
YColumn	The name of the column containing y-coordinates of sample sites. Y-coordinates can be provided in either the biological or environmental data tables, but MUST be in at least one of them. If a y-coordinate column is provided in both the biological and environmental data, the column name must be identical.
sppColumn	Only used if bioFormat = 2 (x, y, species list). The name of the column containing unique identifiers for each species.
abundColumn	If abundance = TRUE, this parameter identifies the column containing the measure of abundance at each site. Only used if bioFormat = 2 (i.e., x, y, species list), though in the case of abundance data, the format would be: x, y, species, abundance.
sppFilter	Default = 0. To account for limited sampling effort at some sites, sppFilter removes all sites at which the number of recorded species (i.e., species richness) is less than the specified value. For example, if sppFilter = 5, all sites with fewer than 5 recorded species will be removed.
predData	The environmental predictor data. Accepts either a site-predictor matrix or a raster stack.
distPreds	A list of distance matrices to be used as predictors, either in combination with, or as a substitute for, predData. For example, pairwise dissimilarity for one biological group (e.g., trees) can be used as a predictor for another group (e.g., ferns).

weightType	Default = "equal". Defines the weighting for sites. Can be either: (1) "equal" (weights for all sites set = 1), (2) "richness" (each site weighted according to number of species recorded), or (3) "custom" (user defined). If weightType="custom", the user must provide a vector of site weights equal to the number of rows in the full site-pair table (i.e., before species filtering or sub-sampling is taken into account).
custWeightVect	A numeric object of user-defined site weights. Required when weightType = "custom". Ignored otherwise.
samples	Default = NULL. The maximum number of site-pairs to be included in the output site-pair table. The function first constructs the full site-pair table. Then, if the value of the samples parameter is a less than or equal to the number of rows available in the site-pair table, a new table is created with randomly selected rows from the original table.

Details

bioData and bioFormat: The function accepts biological data in the following formats:

bioData = site-species matrix; bioFormat = 1: assumes that the response data are provided with a site ID column (specified by siteCol) and, optionally, two columns for the x & y coordinates of the sites. All remaining columns contain the biological data, with a column for each biological entity (most commonly species). In the case that a raster stack is provided for the environmental data (predData), x-y coordinates MUST be provided in bioData to allow extraction of the environmental data at site locations. The x-y coordinates will be intersected with the raster stack and, if the number of unique cells intersected by the points is less than the number of unique site IDs (i.e. multiple sites fall within a single cell), the function will use the raster cell as the site ID and aggregate sites accordingly. Therefore, model fitting will be sensitive to raster cell size. If the environmental data are in tabular format, they should have the same number of sites (i.e., same number of rows) as bioData. The x-y coordinate and site ID columns must have the same names in bioData and predData.

bioData = x, y, species list (optionally a fourth column with abundance can be provided); bioFormat = 2: assumes a table of 3 or 4 columns, the first two being the x & y coordinates of species records, the third (sppCol) being the name of the species observed at that location, and optionally a fourth column indicating a measure of abundance. If an abundance column is not provided, presence-only data are assumed. In the case that a raster stack is provided for the environmental data (predData), the x-y coordinates will be intersected with the raster stack and, if the number of unique cells intersected by the points is less than the number of unique site IDs (i.e. multiple sites fall within a single cell), the function will use the raster cell as the site ID and aggregate sites accordingly. Therefore, model fitting will be sensitive to raster cell size.

bioData = site-site distance (dissimilarity) matrix; bioFormat = 3: is used when a site-site distance (dissimilarity) matrix has already been created for the biological response (e.g., Fst for genetic data). Only the lower half (triangle) of the matrix is needed to create the site-pair output table, but this function automatically removes the upper half if present. This is the only case in which the environmental data CANNOT be provided as a raster stack.

bioData = site-pair table; bioFormat = 4: with an already created site-pair table, this option allows for one to add one or more distance matrices (see distPreds above) to the table and/or sub-sample the table (see sample above).

NOTE: The function assumes that the x-y coordinates and the raster stack (if used) are in the same coordinate system. No checking is performed to confirm this is the case.

Value

A site-pair formatted table containing the response (biological distance or dissimilarity), predictors, and weights as required for fitting Generalized Dissimilarity Models.

Examples

```
##table data, species and environmental
load(system.file("../data/gdm.RData", package="gdm"))
sppData <- gdmExpData[c(1,2,13,14)]
envTab <- gdmExpData[c(2:ncol(gdmExpData))]

##environmental raster data
##commented out to reduce example run time
#rastFile <- system.file("../extdata/stackedVars.grd", package="gdm")
#envRast <- stack(rastFile)

#####table type 1
##site-species table without coordinates
testData1a <- reshape2::dcast(sppData, site~species)
##site-species table with coordinates
coords <- unique(sppData[2:ncol(sppData)])
testData1b <- merge(testData1a, coords, by="site")
##site-species, table-table
exFormat1a <- formatsitepair(testData1a, 1, siteColumn="site", XColumn="Long", YColumn="Lat",
predData=envTab)
##site-species, table-raster
##commented out to reduce example run time
#exFormat1b <- formatsitepair(testData1b, 1, siteColumn="site", XColumn="Long", YColumn="Lat",
# predData=envRast)

#####table type 2
##site xy spp list, table-table
exFormat2a <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
siteColumn="site", predData=envTab)
##site xy spp list, table-raster
##commented out to reduce example run time
#exFormat2b <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
# siteColumn="site", predData=envRast)

#####table type 3
##dissim matrix model
exFormat3 <- formatsitepair(gdmDissim, 3, XColumn="Long", YColumn="Lat", predData=envTab,
siteColumn="site")

#####table type 4
##adds a predictor matrix to an existing site-pair table, in this case, predData needs to be
##filled, but is not actually used
exFormat4 <- formatsitepair(exFormat2a, 4, predData=envTab, distPreds=list(as.matrix(gdmDissim)))
```

Description

For an overview of the functions in the gdm package have a look here: [gdm-package](#).

gdm is used to fit a generalized dissimilarity model to tabular site-pair data formatted as follows using the `formatsitepair` function: Response, Weights, X0, Y0, X1, Y1, Pred1SiteA, Pred2SiteA, ..., PredNSiteA, Pred1SiteB, Pred2SiteB, ..., PredNSiteB. The first column (Response) must be any ratio based dissimilarity (distance) measure between Site A and Site B. The second column defines any weighting to be applied during fitting of the model. If equal weighting is required, then all entries in this column should be set to 1.0 (default). The third and fourth columns, X0 and Y0, represent the spatial coordinates of the first site in the site pair (Site A). The fifth and sixth columns, X1 and Y1, represent the coordinates of the second site (Site B). Note that the first six columns are REQUIRED, even if you do not intend to use geographic distance as a predictor (in which case these columns can be loaded with dummy data if the actual coordinates are unknown). The next N*2 columns contain values for N predictors for Site A, followed by values for the same N predictors for Site B.

The following is an example of a GDM input table header:

```
Response, Weights, X0, Y0, X1, Y1, S1_Temp, S1_Rain, S1_Bedrock, S2_Temp, S2_Rain, S2_Bedrock
```

Usage

```
gdm(data, geo=FALSE, splines=NULL, knots=NULL)
```

Arguments

data	A data frame containing the site pairs to be used to fit the GDM (typically obtained using the <code>formatsitepair</code> function). The observed response data must be located in the first column. The weights to be applied to each site pair must be located in the second column. If <code>geo</code> is TRUE, then the X0, Y0 and Y0, Y1 columns will be used to calculate the geographic distance between site pairs for inclusion as the geographic predictor term in the model. If <code>geo</code> is FALSE (default), then the X0, Y0, X1 and Y1 data columns must still be included, but are ignored in fitting the model. Columns containing the predictor data for Site A, and the predictor data for Site B, follow.
geo	Set to TRUE if geographic distance between sites is to be included as a model term. Set to FALSE if geographic distance is to be omitted from the model. Default is FALSE.
splines	An optional vector of the number of I-spline basis functions to be used for each predictor in fitting the model. If supplied, it must have the same length as the number of predictors (including geographic distance if <code>geo</code> is TRUE). If this vector is not provided (<code>splines=NULL</code>), then a default of 3 basis functions is used for all predictors.

knots An optional vector of knots in *units of the predictor variables* to be used in the fitting process. If knots are supplied and `splines=NULL`, then the knots argument must have the same length as the number of predictors * 3. If both knots and the number of splines are supplied, then the length of the knots argument must be the same as the sum of the values in the splines vector. Note that the default values for knots when the default three I-spline basis functions are 0 (minimum), 50 (median), and 100 (maximum) quantiles.

Value

`gdm` returns a `gdm` model object. The function `summary.gdm` can be used to obtain or print a synopsis of the results. A `gdm` model object is a list containing at least the following components:

<code>dataname</code>	The name of the table used as the data argument to the model.
<code>geo</code>	Whether geographic distance was used as a predictor in the model.
<code>gdmdeviance</code>	The deviance of the fitted GDM model.
<code>nulldeviance</code>	The deviance of the null model.
<code>explained</code>	The percentage of null deviance explained by the fitted GDM model.
<code>intercept</code>	The fitted value for the intercept term in the model.
<code>predictors</code>	A list of the names of the predictors that were used to fit the model.
<code>coefficients</code>	A list of the coefficients for each spline for each of the predictors considered in model fitting.
<code>knots</code>	A vector of the knots derived from the x data (or user defined), for each predictor.
<code>splines</code>	A vector of the number of I-spline basis functions used for each predictor.
<code>creationdate</code>	The date and time of model creation.
<code>observed</code>	The observed response for each site pair (from data column 1).
<code>predicted</code>	The predicted response for each site pair, from the fitted model (after applying the link function).
<code>ecological</code>	The linear predictor (ecological distance) for each site pair, from the fitted model (before applying the link function).

References

Ferrier S, Manion G, Elith J, Richardson, K (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity & Distributions* 13, 252-264.

See Also

`formatsitepair`, `summary.gdm`, `plot.gdm`, `predict.gdm`, `gdm.transform`

Examples

```
##fit table environmental data
##sets up site-pair table, environmental tabular data
load(system.file("../data/gdm.RData", package="gdm"))
sppData <- gdmExpData[c(1,2,13,14)]
envTab <- gdmExpData[c(2:ncol(gdmExpData))]
sitePairTab <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
siteColumn="site", predData=envTab)

##fit table GDM
gdmTabMod <- gdm(sitePairTab, geo=TRUE)
summary(gdmTabMod)

##fit raster environmental data
##sets up site-pair table
rastFile <- system.file("../extdata/stackedVars.grd", package="gdm")
envRast <- stack(rastFile)

##environmental raster data
sitePairRast <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
siteColumn="site", predData=envRast)
##sometimes raster data returns NA in the site-pair table, these rows will have to be removed
##before fitting gdm
sitePairRast <- na.omit(sitePairRast)

##fit raster GDM
gdmRastMod <- gdm(sitePairRast, geo=TRUE)
summary(gdmRastMod)
```

gdm.transform	<i>Transform Environmental Data Using a Generalized Dissimilarity Model</i>
---------------	---

Description

This function transforms geographic and environmental predictors using (1) the fitted functions from a model object returned from `gdm` and (2) a data frame or raster stack containing predictor data for a set of sites.

Usage

```
gdm.transform(model, data)
```

Arguments

model	A gdm model object resulting from a call to <code>gdm</code> .
data	Either (i) a data frame containing values for each predictor variable in the model, formatted as follows: X, Y, var1, var2, var3, ..., varN or (ii) a raster stack with one layer per predictor variable used in the model, excluding X and Y (rasters

for x- and y-coordinates are built automatically from the input rasters if the model was fit with `geo=T`). The order of the columns (data frame) or raster layers (raster stack) **MUST** be the same as the order of the predictors in the site-pair table used in model fitting. There is currently no checking to ensure that the order of the variables to be transformed are the same as those in the site-pair table used in model fitting. If geographic distance was not used as a predictor in model fitting, the x- and y-columns need to be removed from the data to be transformed. Output is provided in the same format as the input data.

Value

`gdm.transform` returns either a data frame with the same number of rows as the input data frame or a raster stack, depending on the format of the input data. If the model uses geographic distance as a predictor the output object will contain columns or layers for the transformed X and Y values for each site. The transformed environmental data will be in the remaining columns or layers.

References

Ferrier S, Manion G, Elith J, Richardson, K (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity & Distributions* 13, 252-264.

Fitzpatrick MC, Keller SR (2015) Ecological genomics meets community-level modeling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* 18: 1-16

Examples

```
load(system.file("../data/gdm.RData", package="gdm"))

# grab the columns with xy, site ID, and species data
sppTab <- gdmExpData[, c("species", "site", "Lat", "Long")]

##fit gdm using rasters
rastFile <- system.file("../extdata/stackedVars.grd", package="gdm")
envRast <- stack(rastFile)
sitePairRast <- formatsitepair(sppTab, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
siteColumn="site", predData=envRast)
##remove NA values
sitePairRast <- na.omit(sitePairRast)

##fit raster GDM
gdmRastMod <- gdm(sitePairRast, geo=TRUE)

##raster input, raster output
transRasts <- gdm.transform(gdmRastMod, envRast)

# map biological patterns
rastDat <- sampleRandom(transRasts, 10000)
pcaSamp <- prcomp(rastDat)

# note the use of the 'index' argument
```

```
pcaRast <- predict(transRasts, pcaSamp, index=1:3)

# scale rasters
pcaRast[[1]] <- (pcaRast[[1]]-pcaRast[[1]]@data@min) /
  (pcaRast[[1]]@data@max-pcaRast[[1]]@data@min)*255
pcaRast[[2]] <- (pcaRast[[2]]-pcaRast[[2]]@data@min) /
  (pcaRast[[2]]@data@max-pcaRast[[2]]@data@min)*255
pcaRast[[3]] <- (pcaRast[[3]]-pcaRast[[3]]@data@min) /
  (pcaRast[[3]]@data@max-pcaRast[[3]]@data@min)*255

plotRGB(pcaRast, r=1, g=2, b=3)
```

gdmDissim

A sample biological dissimilarity matrix for use in GDM examples

Description

A full dissimilarity matrix.

Format

A data frame with 180 rows and 180 columns.

Source

Created using the example species data from southwest Australia.

gdmExpData

Species and environmental data used in GDM examples

Description

A dataset containing the species, site, and environmental data used to fit GDM in examples.

Format

A data frame with 42,142 rows and 14 variables

Details

- species: Species identifier.
- site: Site identification.
- awcA: plant-available water capacity of soil
- phTotal: soil pH
- sandA: percent sand content of soil
- shcA: saturated hydraulic conductivity of soil
- solumDepth: depth of solum layer (unknown units)
- bio5: Maximum Temperature of Warmest Month (C x 10)
- bio6: Minimum Temperature of Coldest Month (C x 10)
- bio15: Precipitation Seasonality (Coefficient of Variation)
- bio18: Precipitation of Warmest Quarter (mm)
- bio19: Precipitation of Coldest Quarter (mm)
- Lat: Latitude Coordinate of site
- Long: Longitude Coordinate of site

Source

Fitzpatrick MC, Sanders NJ, Normand S, Svenning J-C, Ferrier S, Gove AD, Dunn RR (2013). Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. *Proceedings of the Royal Society: Series B* 280, art. 1768

isplineExtract

Extract I-spline Values From a gdm Object.

Description

Extracts the I-spline values from a gdm object. There is one I-spline for each predictor that has at least one non-zero coefficient in the fitted model.

Usage

```
isplineExtract(model)
```

Arguments

model A gdm object from [gdm](#).

Value

A list with two items. The first item contains the x-values (actual values of the predictors) of the I-splines and the second item contains the y-values (partial ecological distances) of the fitted I-splines.

References

Ferrier S, Manion G, Elith J, Richardson, K (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity & Distributions* 13, 252-264.

Fitzpatrick MC, Sanders NJ, Normand S, Svenning J-C, Ferrier S, Gove AD, Dunn RR (2013). Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. *Proceedings of the Royal Society: Series B* 280, art. 1768

Examples

```
##sets up site-pair table
load(system.file("../data/gdm.RData", package="gdm"))
sppData <- gdmExpData[, c(1,2,14,13)]
envTab <- gdmExpData[, c(2:ncol(gdmExpData))]
sitePairTab <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
siteColumn="site", predData=envTab)

##create GDM
gdmMod <- gdm(sitePairTab, geo=TRUE)

##extracts splines
exSplines <- isplineExtract(gdmMod)

##plot spline(s)
#spline for winter precip (bio19)
plot(exSplines[[1]][,"bio19"], exSplines[[2]][,"bio19"], type="l",
lwd=3, xlab="Winter precipitation (mm)", ylab="Partial Ecological Distance")
```

plot.gdm

Plot model fit and I-splines from a generalized dissimilarity model

Description

plot is used to plot the I-splines and fit of a generalized dissimilarity model created using the [gdm](#) function.

Usage

```
## S3 method for class 'gdm'
## S3 method for class 'gdm'
plot(x, plot.layout = c(2, 2), plot.color = "blue", plot.linewidth = 2,
include.rug = FALSE, rug.sitepair = NULL, ...)
```

Arguments

x A gdm model object returned from [gdm](#).

plot.layout	This specifies the row and column layout for the plots, including: (1) a single page plot of observed response data against the raw linear predictor (ecological distance) from the model, and (2) a single page plot of the observed response against the predicted response from the model, i.e. after applying the link function, $1.0 - \exp(-y)$, to the linear predictor, and (3) the I-splines fitted to the individual predictors. Default is 2 rows by 2 columns. To produce one predictor plot per page set plot.layout to c(1,1). The first two model plots are always produced on a single page each and therefore the layout parameter affects only the layout of the I-spline plots for those predictors that featured in the model fitting process (i.e., predictors with all-zero I-spline coefficients are not plotted).
plot.color	Color of the data points that are plotted for the overall plots.
plot.linewidth	The line width for the regression line over-plotted in the two overall plots to optimize the display of the line over the data points.
include.rug	Rather or not to include a rug plot of the actual site-pair values used to fit the gdm in the plots of the variable splines. When set to TRUE, the rug.sitepair argument should be filled. Default is FALSE.
rug.sitepair	A site-pair table used to add a rug plot of the tables values to the variable spline plots. This should be the same site-pair table used to fit the given gdm object, and there is no internal code to check that this is the case.
...	Ignored.

Value

plot returns NULL. Use [summary.gdm](#) to obtain a synopsis of the model object.

References

Ferrier S, Manion G, Elith J, Richardson, K (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity & Distributions* 13, 252-264.

See Also

[isplineExtract](#)

Examples

```
##sets up site-pair table
load(system.file("../data/gdm.RData", package="gdm"))
sppData <- gdmExpData[c(1,2,13,14)]
envTab <- gdmExpData[c(2:ncol(gdmExpData))]
sitePairTab <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
siteColumn="site", predData=envTab)

##create GDM
gdmMod <- gdm(sitePairTab, geo=TRUE)

##plot GDM
plot(gdmMod, plot.layout=c(3,3))
```

predict.gdm	<i>Predict Biological Dissimilarities Between Sites or Times Using a Generalized Dissimilarity Model</i>
-------------	--

Description

This function predicts biological distances between sites or times using a model object returned from `gdm`. Predictions between site pairs require a data frame containing the values of predictors for pairs of locations, formatted as follows: Distance, Weights, X1, Y1, X2, Y2, Var1Site1, Var2Site1, ..., VarNSite1, Var1Site2, Var2Site2, ..., VarNSite2. Predictions of biological change through time require two raster stacks or bricks for environmental conditions at two time periods, each with a layer for each environmental predictor in the fitted model.

Usage

```
## S3 method for class 'gbm'
## S3 method for class 'gdm'
predict(object, data, time=FALSE, predRasts=NULL, ...)
```

Arguments

object	A gdm model object resulting from a call to <code>gdm</code> .
data	<p>Either a data frame containing the values of predictors for pairs of sites, in the same format and structure as used to fit the model using <code>gdm</code> or a raster stack if a prediction of biological change through time is needed.</p> <p>For a data frame, the first two columns - Distance and Weights - are required by the function but are not used in the prediction and can therefore be filled with dummy data (e.g. all zeros). If <code>geo</code> is TRUE, then the X1, Y1 and X2, Y2 columns will be used for calculating the geographical distance between each site for inclusion of the geographic predictor term into the GDM model. If <code>geo</code> is FALSE, then the X1, Y1, X2 and Y2 data columns are ignored. However these columns are still REQUIRED and can be filled with dummy data (e.g. all zeroes). The remaining columns are for N predictors for Site1 and followed by N predictors for Site2. The order of the columns must match those in the site-pair table used to fit the model.</p> <p>A raster stack should be provided only when <code>time=T</code> and should contain one layer for each environmental predictor in the same order as the columns in the site-pair table used to fit the model.</p>
time	TRUE/FALSE: Is the model prediction for biological change through time?
predRasts	A raster stack characterizing environmental conditions for a different time in the past or future, with the same extent, resolution, and layer order as the data object. Required only if <code>time=T</code> .
...	Ignored.

Value

predict returns either a response vector with the same length as the number of rows in the input data frame or a raster depicting change through time across the study region.

See Also

[gdm.transform](#)

Examples

```
##sets up site-pair table
load(system.file("../data/gdm.RData", package="gdm"))
sppData <- gdmExpData[, c(1,2,14,13)]
envTab <- gdmExpData[, c(2:ncol(gdmExpData))]

# remove soils (no rasters for these)
envTab <- envTab[,-c(2:6)]
sitePairTab <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",
siteColumn="site", predData=envTab)

##create GDM
gdmMod <- gdm(sitePairTab, geo=TRUE)

##predict GDM
predDiss <- predict(gdmMod, sitePairTab)

##time example
rastFile <- system.file("../extdata/stackedVars.grd", package="gdm")
envRast <- stack(rastFile)

##make some fake climate change data
futRasts <- envRast
##reduce winter precipitation by 25%
futRasts[[3]] <- futRasts[[3]]*0.75

timePred <- predict(gdmMod, envRast, time=TRUE, predRasts=futRasts)
plot(timePred)
```

stackedVars

Worldclim climate variables used to fit GDM in examples

Description

A raster stack of five Worldclim gridded climate variables used as environmental data in the GDM examples. The data were originally downloaded from Worldclim using the `getData()` function in the raster package and clipped to the area encompassing the example data points.

Format

A raster stack of five layers.

Details

The specific variables included in this data are: bio5, bio6, bio15, bio18, and bio19. The clipped area is southwestern Australia.

Source

<http://www.worldclim.org/>

References

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.

summary.gdm

Summarize a Fitted Generalized Dissimilarity Model

Description

This function summarizes the model object returned from [gdm](#).

Usage

```
## S3 method for class 'gdm'  
summary(object, ...)
```

Arguments

object	A gdm model object resulting from a call to gdm .
...	Ignored.

Value

summary prints its output to the R Console window and returns no value.

See Also

[gdm](#)

Examples

```
##sets up site-pair table  
load(system.file("../data/gdm.RData", package="gdm"))  
sppData <- gdmExpData[, c(1,2,14,13)]  
envTab <- gdmExpData[, c(2:ncol(gdmExpData))]  
sitePairTab <- formatsitepair(sppData, 2, XColumn="Long", YColumn="Lat", sppColumn="species",  
siteColumn="site", predData=envTab)
```

```
##create GDM  
gdmMod <- gdm(sitePairTab, geo=TRUE)
```

```
##summary of GDM  
summary(gdmMod)
```

Index

*Topic **datasets**

- gdmDissim, 11
- gdmExpData, 11
- stackedVars, 16

*Topic **gdm**

- formatsitepair, 3
- gdm, 7
- gdm.transform, 9
- isplineExtract, 12
- plot.gdm, 13
- predict.gdm, 15
- summary.gdm, 17

formatsitepair, 2, 3, 7, 8

gdm, 2, 3, 7, 9, 12, 13, 15, 17
gdm-package, 2
gdm.transform, 3, 8, 9, 16
gdmDissim, 11
gdmExpData, 11

isplineExtract, 3, 12, 14

plot, 3
plot (plot.gdm), 13
plot.gdm, 8, 13
predict, 2
predict (predict.gdm), 15
predict.gdm, 8, 15

stackedVars, 16
summary, 2
summary (summary.gdm), 17
summary.gdm, 8, 14, 17

vegdist, 4