Package ‘gimme’
August 19, 2021

Title Group Iterative Multiple Model Estimation
Version 0.7-6
Date 2021-08-20
Author Stephanie Lane [aut, trl],
       Kathleen Gates [aut, cre, ccp],
       Zachary Fisher [aut],
       Cara Arizmendi [aut],
       Peter Molenaar [aut, ccp],
       Michael Hallquist [ctb],
       Hallie Pike [ctb],
       Teague Henry [ctb],
       Kelly Duffy [ctb],
       Lan Luo [ctb],
       Adriene Beltz [csp],
       Aidan Wright [csp],
       Jonathan Park [ctb],
       Sebastian Castro Alvarez [ctb]
Maintainer KM Gates <gateskm@email.unc.edu>
Depends R (>= 3.5.0)
Imports lavaan (>= 0.6-9), igraph (>= 1.0-0), qgraph, data.tree,
       MIIVsem (>= 0.5.4), imputeTS (>= 3.0), nlptr, graphics, stats
Description Automated identification and estimation of group- and
       individual-level relations in time series data.
License GPL-2
LazyData true
URL https://github.com/GatesLab/gimme/
BugReports https://github.com/GatesLab/gimme/issues
ByteCompile true
RoxygenNote 7.1.1
NeedsCompilation no
Suggests knitr, rmarkdown
gimme-package

VignetteBuilder knitr
Encoding UTF-8
Repository CRAN
Date/Publication 2021-08-19 19:50:02 UTC

R topics documented:

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gimme-package</td>
<td>Group iterative multiple model estimation</td>
</tr>
</tbody>
</table>

Description

This package contains functions to identify group- and individual-level unified structural equation models.

Details

Researchers across varied domains gather multivariate data for each individual unit of study across multiple occasions of measurement. Generally referred to as time series (or in the social sciences, intensive longitudinal) data, examples include psychophysiological processes such as neuroimaging and heart rate variability, daily diary studies, and observational coding of social interactions among dyads. A primary goal for acquiring these data is to understand temporal processes. The gimme package contains several functions for use with these data. These functions include gimmeSEM, which provides both group- and individual-level results by looking across individuals for patterns of relations among variables. A function that provides group-level results, aggSEM, is included, as well as a function that provides individual-level results, indSEM. The major functions within the gimme package all require the user to specify the directory containing the data and a directory for output to be stored.
aggSEM

Author(s)
Stephanie Lane [aut, trl],
Kathleen Gates [aut, cre],
Zachary Fisher [aut],
Cara Arizmendi [aut],
Peter Molenaar [aut],
Michael Hallquist [ctb],
Hallie Pike [ctb],
Cara Arizmendi [ctb],
Teague Henry [ctb],
Kelly Duffy [ctb],
Lan Luo [ctb],
Adriene Beltz [csp]
Maintainer: KM Gates <gateskm@email.unc.edu>

Description
Concatenates all individual-level data files and fits a group model to the data.

Usage
aggSEM(data = "", out = "", sep = "", header = , ar = TRUE, plot = TRUE, paths = NULL, exogenous = NULL, conv_vars = NULL, conv_length = 16, conv_interval = 1, mult_vars = NULL, mean_center_mult = FALSE, standardize = FALSE, hybrid = FALSE)

Arguments

data The path to the directory where the data files are located, or the name of the list containing each individual’s time series. Each file or matrix must contain one matrix for each individual containing a T (time) by p (number of variables) matrix where the columns represent variables and the rows represent time. If in list form, each item in the list (i.e., matrix) must be named.
out  
The path to the directory where the results will be stored (optional). If specified, a copy of output files will be replaced in directory. If directory at specified path does not exist, it will be created.

sep  
The spacing of the data files when data are in a directory. "\" indicates space-delimited, "\t" indicates tab-delimited, "," indicates comma delimited. Only necessary to specify if reading data in from physical directory.

header  
Logical. Indicate TRUE for data files with a header. Only necessary to specify if reading data in from physical directory.

ar  
Logical. If TRUE, begins search for group model with autoregressive (AR) paths open. Defaults to TRUE.

plot  
Logical. If TRUE, figures depicting relations among variables of interest will automatically be created. For aggregate-level plot, red paths represent positive weights and blue paths represent negative weights. Dashed lines denote lagged relations (lag 1) and solid lines are contemporaneous (lag 0). Defaults to TRUE.

paths  
lavaan-style syntax containing paths with which to begin model estimation (optional). That is, Y~X indicates that Y is regressed on X, or X predicts Y. If no header is used, then variables should be referred to with V followed (with no separation) by the column number. If a header is used, variables should be referred to using variable names. To reference lag variables, "lag" should be added to the end of the variable name with no separation. Defaults to NULL.

exogenous  
Vector of variable names to be treated as exogenous (optional). That is, exogenous variable X can predict Y but cannot be predicted by Y. If no header is used, then variables should be referred to with V followed (with no separation) by the column number. If a header is used, variables should be referred to using variable names. Defaults to NULL.

conv_vars  
Vector of variable names to be convolved via smoothed Finite Impulse Response (sFIR). Defaults to NULL.

conv_length  
Expected response length in seconds. For functional MRI BOLD, 16 seconds (default) is typical for the hemodynamic response function.

conv_interval  
Interval between data acquisition. Currently must be a constant. For fMRI studies, this is the repetition time. Defaults to 1.

mult_vars  
Vector of variable names to be multiplied to explore bilinear/modulatory effects (optional). All multiplied variables will be treated as exogenous (X can predict Y but cannot be predicted by Y). Within the vector, multiplication of two variables should be indicated with an asterik (e.g. V1*V2). If no header is used, variables should be referred to with V followed by the column number (with no separation). If a header is used, each variable should be referred to using variable names. If multiplication with the lag 1 of a variable is desired, the variable name should be followed by "lag" with no separation (e.g. V1*V2lag). Note that if multiplied variables are desired, at least one variable in the dataset must be specified as exogenous. Defaults to NULL.

mean_center_mult  
Logical. If TRUE, the variables indicated in mult_vars will be mean-centered before being multiplied together. Defaults to FALSE.

standardize  
Logical. If TRUE, all variables will be standardized to have a mean of zero and a standard deviation of one. Defaults to FALSE.
Description

This function estimates the basis vectors related to responses following a binary impulse and convolves that binary impulse vector.

Usage

```r
convolveFIR(ts_list = NULL, varLabels = NULL, conv_length = 16, conv_interval = 1)
```
Arguments

- **ts_list**: a list of dataframes.
- **varLabels**: a list of variable sets. Contains `varLabels$coln`, all column names, `varLabels$conv`, the names of variables to convolve, and `varLabels$exog`, a list of exogenous variables (if any).
- **conv_length**: Expected response length in seconds. For functional MRI BOLD, 16 seconds (default) is typical for the hemodynamic response function.
- **conv_interval**: Interval between data acquisition. Currently must be a constant. For fMRI studies, this is the repetition time. Defaults to 1.

---

**gimmeSEM**

*Group iterative multiple model estimation.*

---

**Description**

This function identifies structural equation models for each individual that consist of both group-level and individual-level paths.

**Usage**

```r
gimmeSEM(data = NULL,
          out = NULL,
          sep = NULL,
          header = NULL,
          ar = TRUE,
          plot = TRUE,
          subgroup = FALSE,
          sub_feature = "lag & contemp",
          sub_method = "Walktrap",
          sub_sim_thresh = "lowest",
          confirm_subgroup = NULL,
          paths = NULL,
          exogenous = NULL,
          conv_vars = NULL,
          conv_length = 16,
          conv_interval = 1,
          mult_vars = NULL,
          mean_center_mult = FALSE,
          standardize = FALSE,
          groupcutoff = .75,
          subcutoff = .5,
          diagnos = FALSE,
          ms_allow = FALSE,
          ms_tol = 1e-5,
          lv_model = NULL,
          )
```
lv_estimator = "miiv",
lv_scores = "regression",
lv_miiv_scaling = "first.indicator",
lv_final_estimator = "miiv",
lasso_model_crit = NULL,
hybrid = FALSE,
VAR = FALSE,
dir_prop_cutoff = 0)

Arguments

- **data**: The path to the directory where the data files are located, or the name of the list containing each individual’s time series. Each file or matrix must contain one matrix for each individual containing a T (time) by p (number of variables) matrix where the columns represent variables and the rows represent time. Individuals must have the same variables (p) but can have different lengths of observations (T).

- **out**: The path to the directory where the results will be stored (optional). If specified, a copy of output files will be replaced in directory. If directory at specified path does not exist, it will be created.

- **sep**: The spacing of the data files. "" indicates space-delimited, "/t" indicates tab-delimited, "," indicates comma delimited. Only necessary to specify if reading data in from physical directory.

- **header**: Logical. Indicate TRUE for data files with a header. Only necessary to specify if reading data in from physical directory.

- **ar**: Logical. If TRUE, begins search for group model with autoregressive (AR) paths freed for estimation. If ms_allow=TRUE, it is recommended to set ar=FALSE. Multiple solutions are unlikely to be found when ar=TRUE. Defaults to TRUE.

- **plot**: Logical. If TRUE, graphs depicting relations among variables of interest will automatically be created. Solid lines represent contemporaneous relations (lag 0) and dashed lines reflect lagged relations (lag 1). For individual-level plots, red paths represent positive weights and blue paths represent negative weights. Width of paths corresponds to estimated path weight. For the group-level plot, black represents group-level paths, grey represents individual-level paths, and (if subgroup = TRUE) green represents subgroup-level paths. For the group-level plot, the width of the edge corresponds to the count. Defaults to TRUE.

- **subgroup**: Logical. If TRUE, subgroups are generated based on similarities in model features using the walktrap.community function from the igraph package. When ms_allow=TRUE, subgroup should be set to FALSE. Defaults to FALSE.

- **sub_feature**: Option to indicate feature(s) used to subgroup individuals. Defaults to "lag & contemp" for lagged and contemporaneous, which is the original method. Can use "lagged" or "contemp" to subgroup solely on features related to lagged and contemporaneous relations, respectively.

- **sub_method**: Community detection method used to cluster individuals into subgroups. Options align with those available in the igraph package: "Walktrap" (default), "Infomap", "Louvain", "Edge Betweenness", "Label Prop", "Fast Greedy", "Leading Eigen", and "Spinglass".
sub_sim_thresh  Threshold for inducing sparsity in similarity matrix. Options are: the percent of edges in the similarity matrix to set to zero (e.g., .25 would set the lower quartile), "lowest" (default) subtracts the minimum value from all values, and "search" searches across thresholds to arrive at one providing highest modularity.

confirm_subgroup  Dataframe. Option only available when subgroup = TRUE. Dataframe should contain two columns. The first column should specify file labels (the name of the data files without file extension), and the second should contain integer values (beginning at 1) specifying the subgroup membership for each individual. 

paths  lavaan-style syntax containing paths with which to begin model estimation (optional). That is, Y~X indicates that Y is regressed on X, or X predicts Y. If no header is used, then variables should be referred to with V followed (with no separation) by the column number. If a header is used, variables should be referred to using variable names. To reference lag variables, "lag" should be added to the end of the variable name with no separation. Defaults to NULL.

exogenous  Vector of variable names to be treated as exogenous (optional). That is, exogenous variable X can predict Y but cannot be predicted by Y. If no header is used, then variables should be referred to with V followed (with no separation) by the column number. If a header is used, variables should be referred to using variable names. The default for exogenous variables is that lagged effects of the exogenous variables are not included in the model search. If lagged paths are wanted, "&lag" should be added to the end of the variable name with no separation. Defaults to NULL.

conv_vars  Vector of variable names to be convolved via smoothed Finite Impulse Response (sFIR). Note, conv_vars are not not automatically considered exogenous variables. To treat conv_vars as exogenous use the exogenous argument. Variables listed in conv_vars must be binary variables. If there is missing data in the endogenous variables their values will be imputed for the convolution operation only. Defaults to NULL. If there are multiple variables listed in conv_vars they are not used in the convolution of additional conv_vars. You can’t do lagged variables.

conv_length  Expected response length in seconds. For functional MRI BOLD, 16 seconds (default) is typical for the hemodynamic response function.

conv_interval  Interval between data acquisition. Currently must be a constant. For fMRI studies, this is the repetition time. Defaults to 1.

mult_vars  Vector of variable names to be multiplied to explore bilinear/modulatory effects (optional). All multiplied variables will be treated as exogenous (X can predict Y but cannot be predicted by Y). Within the vector, multiplication of two variables should be indicated with an asterisk (e.g. V1*V2). If no header is used, variables should be referred to with V followed by the column number (with no separation). If a header is used, each variable should be referred to using variable names. If multiplication with the lag 1 of a variable is desired, the variable name should be followed by "lag" with no separation (e.g. V1*V2lag).
mean_center_mult Logical. If TRUE, the variables indicated in mult_vars will be mean-centered before being multiplied together. Defaults to FALSE.

standardize Logical. If TRUE, all variables will be standardized to have a mean of zero and a standard deviation of one. Defaults to FALSE.

groupcutoff Cutoff value for group-level paths. Defaults to .75, indicating that a path must be significant across 75% of individuals to be included as a group-level path.

subcutoff Cutoff value for subgroup-level paths. Defaults to .5, indicating that a path must be significant across at least 50% of the individuals in a subgroup to be considered a subgroup-level path.

diagnos Logical. If TRUE provides internal output for diagnostic purposes. Defaults to FALSE.

ms_allow Logical. If TRUE provides multiple solutions when more than one path has identical modification index values. When ms_allow=TRUE, it is recommended to set ar=FALSE. Multiple solutions are unlikely to be found when ar=TRUE. Additionally, subgroup should be set to FALSE. Output files for individuals with multiple solutions will represent the last solution found for the individual, not necessarily the best solution for the individual.

ms_tol Precision used when evaluating similarity of modification indices when ms_allow = TRUE. We recommend that ms_tol not be greater than the default, especially when standardize=TRUE. Defaults to 1e-5.

lv_model Invoke latent variable modeling by providing the measurement model syntax here. lavaan conventions are used for relating observed variables to factors. Defaults to NULL.

lv_estimator Estimator used for factor analysis. Options are "miiv" (default), "pml" (pseudo-ML) or "svd".

lv_scores Method used for estimating latent variable scores from parameters obtained from the factor analysis when lv_model is not NULL. Options are: "regression" (Default), "bartlett".

lv_miiv_scaling Type of scaling indicator to use when "miiv" selected for lv_estimator. Options are "first.indicator" (Default; the first observed variable in the measurement equation is used), "group" (best one for the group), or "individual" (each individual has the best one for them according to R2).

lv_final_estimator Estimator for final estimations. "miiv" (Default) or "pml" (pseudo-ML).

lasso_model_crit When not null, invokes multiLASSO approach for the GIMME model search procedure. Arguments indicate the model selection criterion to use for model selection: 'bic' (select on BIC), 'aic', 'aicc', 'hqc', 'cv' (cross-validation).

hybrid Logical. If TRUE, enables hybrid-V AR models where both directed contemporaneous paths and contemporaneous covariances among residuals are candidate relations in the search space. Defaults to FALSE.

VAR Logical. If true, V AR models where contemporaneous covariances among residuals are candidate relations in the search space. Defaults to FALSE.
dir_prop_cutoff

Option to require that the directionality of a relation has to be higher than the reverse direction for a prespecified proportion of individuals.

Details

In main output directory:

- **indivPathEstimates** Contains estimate, standard error, p-value, and z-value for each path for each individual. If subgroup = TRUE and subgroups are found, then a column is present containing the subgroup membership for each individual. Also contains the level at which each path was estimated: group, subgroup, or individual.

- **summaryFit** Contains model fit information for individual-level models. If subgroups are requested, this file also contains the subgroup membership for each individual.

- **summaryPathCountMatrix** Contains counts of total number of paths, both contemporaneous and lagged, estimated for the sample. The row variable is the outcome and the column variable is the predictor variable.

- **summaryPathCounts** Contains summary count information for paths identified at the group-, subgroup (if subgroup = TRUE), and individual-level.

- **summaryPathsPlot** Produced if plot = TRUE. Contains figure with group, subgroup (if subgroup = TRUE), and individual-level paths for the sample. Black paths are group-level, green paths are subgroup-level, and grey paths are individual-level, where the thickness of the line represents the count.

In subgroup output directory (if subgroup = TRUE):

- **subgroup\textsubscript{k}PathCounts** Contains counts of relations among lagged and contemporaneous variables for the \( k \)th subgroup.

- **subgroup\textsubscript{k}Plot** Contains plot of group, subgroup (if subgroup = TRUE), and individual-level paths for the \( k \)th subgroup. Black represents group-level paths, grey represents individual-level paths, and green represents subgroup-level paths.

Note: if a subgroup of size \( n = 1 \) is discovered, subgroup-level output is not produced.

In individual output directory (where \( id \) represents the original file name for each individual):

- **idBetas** Contains individual-level estimates of each path for each individual.

- **idStdErrors** Contains individual-level standard errors for each path for each individual.

- **idPlot** Contains individual-level plots. Red paths represent positive weights and blue paths represent negative weights.

**Author(s)**

Stephanie Lane and Zachary Fisher
**HRFsims**

**Hemodynamic Response Function (HRF) GIMME example.**

**Description**

This object contains a list of simulated time series data for twenty-five individuals. Each data set has 500 time points and five variables. The fifth variable represents an onset vector for stimulation.

**Usage**

```
HRFsims
```

**Format**

A list of data frames with 25 individuals, who each have 500 observations on 5 variables.
Description

This function identifies structural equation models for each individual. It does not utilize any shared information from the sample.

Usage

```r
indSEM(data = NULL,
        out = NULL,
        sep = NULL,
        header = NULL,
        ar = TRUE,
        plot = TRUE,
        paths = NULL,
        exogenous = NULL,
        conv_vars = NULL,
        conv_length = 16,
        conv_interval = 1,
        mult_vars = NULL,
        mean_center_mult = FALSE,
        standardize = FALSE,
        hybrid = FALSE)
```

Arguments

data: The path to the directory where the data files are located, or the name of the list containing each individual’s time series. Each file or matrix must contain one matrix for each individual containing a $T$ (time) by $p$ (number of variables) matrix where the columns represent variables and the rows represent time.

out: The path to the directory where the results will be stored (optional). If specified, a copy of output files will be replaced in directory. If directory at specified path does not exist, it will be created.

sep: The spacing of the data files. "" indicates space-delimited, "/t" indicates tab-delimited, ",," indicates comma delimited. Only necessary to specify if reading data in from physical directory.

header: Logical. Indicate TRUE for data files with a header. Only necessary to specify if reading data in from physical directory.

ar: Logical. If TRUE, begins search for individual models with autoregressive (AR) paths open. Defaults to TRUE.

plot: Logical. If TRUE, graphs depicting relations among variables of interest will automatically be created. Defaults to TRUE. For individual-level plots, red paths represent positive weights and blue paths represent negative weights.
paths  lavaan-style syntax containing paths with which to begin model estimation. That is, Y~X indicates that Y is regressed on X, or X predicts Y. If no header is used, then variables should be referred to with V followed (with no separation) by the column number. If a header is used, variables should be referred to using variable names. To reference lag variables, "lag" should be added to the end of the variable name with no separation. Defaults to NULL.

exogenous  Vector of variable names to be treated as exogenous. That is, exogenous variable X can predict Y but cannot be predicted by Y. If no header is used, then variables should be referred to with V followed (with no separation) by the column number. If a header is used, variables should be referred to using variable names. Defaults to NULL.

conv_vars  Vector of variable names to be convolved via smoothed Finite Impulse Response (sFIR). Defaults to NULL.

conv_length  Expected response length in seconds. For functional MRI BOLD, 16 seconds (default) is typical for the hemodynamic response function.

conv_interval  Interval between data acquisition. Currently must be a constant. For fMRI studies, this is the repetition time. Defaults to 1.

mult_vars  Vector of variable names to be multiplied to explore bilinear/modulatory effects (optional). All multiplied variables will be treated as exogenous (X can predict Y but cannot be predicted by Y). Within the vector, multiplication of two variables should be indicated with an asterisk (e.g. V1*V2). If no header is used, variables should be referred to with V followed by the column number (with no separation). If a header is used, each variable should be referred to using variable names. If multiplication with the lag 1 of a variable is desired, the variable name should be followed by "lag" with no separation (e.g. V1*V2lag). Note that if multiplied variables are desired, at least one variable in the dataset must be specified as exogenous. Defaults to NULL.

mean_center_mult  Logical. If TRUE, the variables indicated in mult_vars will be mean-centered before being multiplied together. Defaults to FALSE.

standardize  Logical. If TRUE, all variables will be standardized to have a mean of zero and a standard deviation of one. Defaults to FALSE.

hybrid  Logical. If TRUE, enables hybrid-V AR models where both directed contemporaneous paths and contemporaneous covariances among residuals are candidate relations in the search space. Defaults to FALSE.

Details
In main output directory:

- **indivPathEstimates** Contains estimate, standard error, p-value, and z-value for each path for each individual
- **summaryFit** Contains model fit information for individual-level models.
- **summaryPathCountMatrix** Contains counts of total number of paths, both contemporaneous and lagged, estimated for the sample. The row variable is the outcome and the column variable is the predictor variable.
• **summaryPathCounts** Contains summary count information for paths identified at the individual-level.

• **summaryPathsPlot** Contains counts of total number of paths, both contemporaneous and lagged, estimated for the sample. The row variable is the outcome and the column variable is the predictor variable.

In individual output directory (where *id* represents the original file name for each individual):

• **idBetas** Contains individual-level estimates of each path for each individual.

• **idStdErrors** Contains individual-level standard errors for each path for each individual.

• **idPlot** Contains individual-level plots. Red paths represent positive weights and blue paths represent negative weights.

**Author(s)**

Stephanie Lane

**Examples**

```r
## Not run:
fit <- indSEM(data = "C:/data100",
              out = "C:/data100_indSEM_out",
              sep = ",",
              header = FALSE)
print(fit, file = "group1.1", estimates = TRUE)
plot(fit, file = "group1.1")
## End(Not run)
```

---

**ms.fit**  
*Fitted gimme object with multiple solutions*

**Description**

This object contains a fitted gimme object where multiple solutions gimme was used. The simulated data had 25 individuals, each with 100 time points.

**Usage**

`ms.fit`

**Format**

A fitted gimme object, where multiple solutions gimme was used.
**simData**

| simData | Large example, heterogeneous data, group, subgroup, and individual level effects. |

**Description**

This object contains a list of simulated time series data for twenty-five individuals with 200 time points and 10 variables, or regions of interest.

**Usage**

simData

**Format**

A list of data frames with 25 individuals, who each have 200 observations on 10 variables.

---

**simDataLV**

| simDataLV | Latent variable example, heterogeneous data, group, subgroup level effects. |

**Description**

This object contains a list of simulated time series data for twenty individuals with 500 time points and 9 variables, or regions of interest.

**Usage**

simDataLV

**Format**

A list of data frames with 20 individuals, who each have 500 observations on 9 variables.
solution.tree  

Solution trees for multiple solutions gimme.

Description

This function allows for the exploration of divergences in multiple solutions gimme for both the group and individuals.

Usage

```r
solution.tree(x, 
  level = c("group", "individual"), 
  cols = NULL, 
  ids = "all", 
  plot.tree = FALSE)
```

Arguments

- **x**: A fitted gimme object.
- **level**: A character vector indicating what levels of the solution tree you would like returned. Options are "group", "individual", or c("group", "individual"). Defaults to c("group", "individual").
- **cols**: A character vector indicating additional information to include in tree plot. Options include "stage", "pruned", "rmsea", "nfi", "cfi", "srmr", "grp_sol", "bic", "aic", "modularity." Defaults to NULL.
- **ids**: A character vector indicating the names of subjects to print. Defaults to "all."
- **plot.tree**: Logical. If TRUE, plot of tree is produced. Defaults to FALSE.

Details

solution.tree

ts  

Small example, heterogeneous data, group and individual level effects

Description

This object contains a list of simulated time series data for five individuals with 50 time points and 3 variables, or regions of interest.

Usage

```r
ts
```

Format

A list of data frames with 5 individuals, who each have 50 observations on 3 variables.
Index

* datasets
  HRFsim, 11
  ms.fit, 14
  simData, 15
  simDataLV, 15
  ts, 16
* gimmeSEM
  gimmeSEM, 6
* gimme
  gimme-package, 2
* indSEM
  indSEM, 12
* setupConvolve
  convolve, 5
  aggSEM, 2, 3
  convolve, 5
  convolveFIR (convolve), 5
  gimme (gimmeSEM), 6
  gimme-package, 2
  gimmeSEM, 2, 6
  HRFsim, 11
  indSEM, 2, 12
  ms.fit, 14
  setupConvolve (convolve), 5
  simData, 15
  simDataLV, 15
  solution.tree, 16
  ts, 16