

Package ‘CompMix’

July 13, 2026

Type Package

Title A Comprehensive Toolkit for Environmental Mixtures Analysis

Version 1.1.0

Description Quantitative characterization of the health impacts associated with exposure to chemical mixtures has received considerable attention in current environmental and epidemiological studies. 'CompMix' package allows practitioners to estimate the health impacts from exposure to chemical mixtures data through various statistical approaches, including Lasso, Elastic net, Bayesian kernel machine regression (BKMR), hierNet, Quantile g-computation, Weighted quantile sum (WQS) and Random forest. Methods and recommendations are described in Hao et al. (2025) <doi:10.1289/EHP15305>.

License GPL-3

Imports Matrix, mvtnorm, hierNet, glmnet, SuperLearner, bkmr, qgcomp, gWQS, pROC, randomForest

Encoding UTF-8

Config/roxygen2/version 8.0.0

NeedsCompilation no

Author Wei Hao [aut, cre]

Maintainer Wei Hao <weihao@umich.edu>

Repository CRAN

Date/Publication 2026-07-13 07:10:20 UTC

Contents

Comp.Mix	2
lmi_simul_dat	4
simul_x_block	5
Index	7

Description

A comprehensive toolkit for environmental mixtures analysis

Usage

```
Comp.Mix(
  y,
  x,
  z = NULL,
  y.type,
  test.pct = 0.5,
  var.select = NULL,
  interaction = NULL,
  interaction.exp.cov = NULL,
  covariates.forcein = NULL,
  bkmr.pip = 0.5,
  bkmr.iter = 500,
  sl.validation = 0.4,
  sl.rep = 5,
  sl.max_iter = 1000,
  sl.tol = 1e-05,
  formula = NULL,
  expnms = NULL,
  seed = 1234,
  verbose = TRUE
)
```

Arguments

y	A vector of either continuous or binary values specifying the health outcome
x	A matrix of numeric values specifying the chemical mixtures
z	A matrix of numeric values specifying the covariates
y.type	A character value of either "continuous" or "binary"
test.pct	A numeric scalar between 0 and 1 specifying the proportion allocated as test samples
var.select	A logical value to indicate whether to perform variable selection
interaction	A logical value (TRUE/FALSE) specifying whether to include pairwise interaction terms between all the chemical mixtures x
interaction.exp.cov	A logical value (TRUE/FALSE) specifying whether to include pairwise interaction terms between all the chemical mixtures x and covariates z. If interaction.exp.cov=TRUE, interaction=TURE or interaction=FALSE will be ignored

<code>covariates.forcein</code>	A logical value (TRUE/FALSE) specifying whether to force in any covariates
<code>bkmr.pip</code>	A numeric scalar between 0 and 1 specifying the cutoff for the posterior inclusion probability in BKMR
<code>bkmr.iter</code>	A positive integer specifying the number of MCMC iterations for bkmr
<code>sl.validation</code>	A positive proportion specifying the fraction of data used for validation when estimating Super Learner weights
<code>sl.rep</code>	A positive integer specifying the number of replicates used when estimating Super Learner weights
<code>sl.max_iter</code>	A positive integer specifying the maximum number of iterations allowed when estimating Super Learner weights
<code>sl.tol</code>	A positive numeric value specifying the convergence tolerance when estimating Super Learner weight
<code>formula</code>	the formula for qgcomp and wqs
<code>expnms</code>	a vector of characters for names of exposure variables
<code>seed</code>	an integer value for seed
<code>verbose</code>	a logical value to show information

Value

A list object which may contain up to 8 cases

Case 1 variable selection on main effects for exposures and confounders

Each case may contain some of the following elements

betaest a numeric vector of coefficients for the exposures

z_betaest a numeric vector of coefficients for the covariates

sse A positive scalar to indicate sum of squares error

corr A numeric scalar between -1 and 1 to indicate correlation coefficient

Author(s)

Wei Hao <weihao@umich.edu>

Examples

```
dat <- lmi_simul_dat(n=1000,p=20,q=5,
  block_idx=c(1,1,2,2,3,1,1,1,1,1,2,2,2,2,3,3,3,3,3,3),
  within_rho=0.6,btw_rho=0.1,R2=0.2,
  effect_size=1,effect_size_i=1,
  cancel_effect = FALSE)
#Example 1: The users would like to perform variable selections
#on main effects of exposures and covariates, and outcome, exposures and
#covariates are entered. For any individual interactions that the users would
#like to include in the models, they can add those into the covariate z.
res_ex1 <- Comp.Mix(y.type="continuous",y=dat$y, x=dat$x, z=dat$z, test.pct=0.5,
```

```
var.select = TRUE, interaction = FALSE, interaction.exp.cov = FALSE,
covariates.forcein = FALSE,
bkmr.pip=0.5, seed=2023)
```

lmi_simul_dat

Simulate data from linear model with interactions

Description

Simulate data from linear model with interactions

Usage

```
lmi_simul_dat(
  n,
  p,
  q,
  block_idx = c(1, 1, 2, 2, 3, 1, 1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3),
  sigma2_x = 1,
  within_rho = 0.6,
  btw_rho = 0.2,
  R2 = 0.8,
  effect_size = 1,
  effect_size_i = 1,
  cancel_effect = TRUE
)
```

Arguments

n	a positive integer to indicate sample size
p	a positive integer to specify the number of exposures
q	a positive integer to specify the number of non-zero effects
block_idx	a vector of positive integers to indicate the block IDs. The length of the vector is p.
sigma2_x	a positive numeric scalar for variance of the covariates
within_rho	a numeric scalar between 0 and 1 for the within block correlation
btw_rho	a numeric scalar between 0 and 1 for the between block correlation
R2	a numeric scalar for R-squared
effect_size	a numeric scalar for effect size for main effect
effect_size_i	a numeric scalar for effect size for interaction effect
cancel_effect	a logic value to indicate whether there is effect cancelation

Value

a list object of the following

- x** covariate matrix of dimension n by p
- n** sample size
- p** number of covariates
- sigma2_x** variance
- within_rho** within block correlation
- btw_rho** between block correlation
- block_idx** block indices

Author(s)

Wei Hao <weihao@umich.edu>

simul_x_block	<i>Simulate covariate matrix with block structure</i>
---------------	---

Description

Simulate covariate matrix with block structure

Usage

```
simul_x_block(n, p, block_idx, sigma2_x = 1, within_rho = 0.6, btw_rho = 0.2)
```

Arguments

n	a positive integer to indicate sample size
p	a positive integer to specify the number of covariates
block_idx	a vector of positive integers to indicate the block IDs. The length of the vector is p.
sigma2_x	a positive numeric scalar for variance of the covariates
within_rho	a numeric scalar between 0 and 1 for the within block correlation
btw_rho	a numeric scalar between 0 and 1 for the between block correlation

Value

a list object of the following

- x** covariate matrix of dimension n by p
- n** sample size
- p** number of covariates
- sigma2_x** variance
- within_rho** within block correlation
- btw_rho** between block correlation
- block_idx** block indices

Author(s)

Wei Hao <weihao@umich.edu>

Examples

```
dat <- simul_x_block(n = 1000, p = 10, block_idx = rep(1:4,length=10))
```

Index

Comp.Mix, [2](#)

lmi_simul_dat, [4](#)

simul_x_block, [5](#)