Package: yahtsee (via r-universe)

September 21, 2024

Title Yet Another Hierachical Time Series Extension and Expansion

Version 0.0.0.9000

Description An opinionated approach to building hierarchical time series models in R using INLA and inlabru.

License MIT + file LICENSE

URL <https://github.com/njtierney/yahtsee>

BugReports <https://github.com/njtierney/yahtsee/issues>

ByteCompile TRUE

Depends R (>= 4.0.0), tsibble

Imports cli (>= 2.5.0.9000), inlabru, magrittr, purrr, stringdist, rlang, stringr, stats, glue, tibble, methods, clock, tidyr, tidyselect, lifecycle, formula.tools, generics, vctrs

Suggests ggplot2, covr, dplyr, knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Additional_repositories <https://inla.r-inla-download.org/R/testing>

Config/testthat/edition 3

Encoding UTF-8

Language es

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Remotes r-lib/cli

Repository https://njtierney.r-universe.dev

RemoteUrl https://github.com/njtierney/yahtsee

RemoteRef HEAD

RemoteSha 94e4f8cb5e67aefdad9945e7b7ead42f0239c562

Contents

---

add_group_id	*Helper function to add group ID information*

Description

Some modelling software needs group data as an index. This means instead of `c("AU", "AU", "UK", "UK", "NZ")` you need `c(1, 1, 2, 2, 3)`, corresponding to each of the groups. This function makes it simpler to add that information to a dataframe.

Usage

```
add_group_id(.data, group)
```

Arguments

`.data`	data.frame
`group`	group to add the ID for

Value

`data.frame` with additional column named ".group_id", where "group" is named after the variable you would like to add the group ID for.

Examples

```
malaria_africa_ts %>% add_group_id(country)
```

---

as_hts_inla   *Create object of class "hts_inla"*

Description

Create object of class "hts_inla"

Usage

```
as_hts_inla(x, formula)
```

Arguments

x	model from inlabru
formula	formula passed to `fit_hts()`

Value

hts_inla model

Note

internal

Author(s)

Nicholas Tierney

---

augment.hts_inla   *augment: Add predictions and residuals to data*

Description

The `augment()` function returns the data used in the model

Usage

```
## S3 method for class 'hts_inla'
augment(x, newdata = NULL, exponentiate = FALSE, ...)
```

Arguments

x	object of class "hts_inla"
newdata	new data to pass to prediction? Default is NULL
exponentiate	default FALSE. Whether to exponentiate predictions
...	extra arguments to pass to augment(). Currently unused.

Value

dataframe with column names of original data, as well as extra columns, ".fitted" and ".resid".

Examples

```
hts_example_model
augment(hts_example_model)
```

---

available_likelihoods   *Helper function to show the available likelihoods that can be used*

Description

Helper function to show the available likelihoods that can be used

Usage

```
available_likelihoods()
```

Value

list of the available likelihoods

Examples

```
available_likelihoods()
# code taken from INLA::inla.models()$likelihood
```

---

build_bru_formula   *create formula to pass to inlabru::bru*

Description

create formula to pass to inlabru::bru

Usage

```
build_bru_formula(formula, .data, special_index)
```

Arguments

formula   formula with an hts() term

.data   tsibble data

special_index   special index term

Value

formula

Author(s)

Nicholas Tierney

Examples

```
bru_form <- build_bru_formula(
  formula = pr ~ avg_lower_age + hts(who_subregion, country),
  .data = malaria_africa_ts,
  special_index = month_num
)
bru_form
```

---

create_hts_data   *Create data to be used when fitting a model with fit_hts*

Description

Create data to be used when fitting a model with fit_hts

Usage

```
create_hts_data(.data, formula)
```

Arguments

.data   tsibble

formula   formula with hts() term

Value

data.frame

Examples

```
data_w_groups <- create_hts_data(
  .data = malaria_africa_ts,
  formula = pr ~ avg_lower_age + hts(who_subregion, country)
  )
data_w_groups
```

---

fitted.hts_inla   *Extract fitted values from* hts_inla *model object*

Description

This returns the fitted

Usage

```
## S3 method for class 'hts_inla'
fitted(object, exponentiate = FALSE, ...)
```

Arguments

object   hts_inla object

exponentiate   default FALSE

...   extra arguments. Currently not used.

Examples

```
head(fitted(hts_example_model))
```

---

fit_hts   *Fit a hierarchical time series model*

Description

Fit a hierarchical time series model

Usage

```
fit_hts(
  formula,
  .data,
  special_index,
  family,
  verbose = TRUE,
  keep_data = TRUE,
  ...
)
```

Arguments

`formula`	formula must contain fixed effect terms and `hts()`
`.data`	a `tsibble`
`special_index`	The variable that represents time in your data
`family`	character.
`verbose`	default TRUE, whether to display message
`keep_data`	attach the data used to fit the model to the final model object?
`...`	list of options to pass to `inlabru::bru`, see `?inlabru::bru` to see these options.

Value

inlabru model

How to allow weighting of observations?

To weight your observations, say when using empirical logit, you must provide a vector of variances, for each observation, which represents the uncertainty of measurement. You can do this, provided the `family = "gaussian"`, add the argument, `scale = vec`, to `fit_hts`, where `vec` is a vector of variances for each observation.

Examples

```
## Not run:
m <- fit_hts(
#inputs are  the levels of hierarchy, in order of decreasing size
formula = pr ~ avg_lower_age + hts(who_subregion, country),
.data = malaria_africa_ts,
family = "gaussian",
special_index = month_num
)

## End(Not run)
```

---

formula_vars_in_data   *Check if formula variables are present in data*

Description

Check if formula variables are present in data

Usage

```
formula_vars_in_data(data, formula)
```

Arguments

data	a tsibble
formula	formula

Examples

```
formula_vars_in_data(
data = malaria_africa_ts,
formula = pr ~ dates
)
```

---

glance.hts_inla   *Glance: A single row summary of a model fit*

Description

Glance: A single row summary of a model fit

Usage

```
## S3 method for class 'hts_inla'
glance(x, ...)
```

Arguments

x	object of class "hts_inla"
...	extra arguments to pass to `glance()`. Currently unused

Value

dataframe with column names: deviance information criterion, wakaike information criteriot, log marginal-likelihood (integration), log marginal-likelihood (Gaussian).

Examples

```
hts_example_model
glance(hts_example_model)
```

---

hts_builder   *Build hts call*

Description

Build hts call

Usage

```
hts_builder(.data, new_index, ...)
```

Arguments

.data	tsibble
new_index	(temporary) new index we are using whilr working out model kinks
...	the levels of hierarchy, in order of decreasing size

Value

constructed expression of ar1 terms to pass to inla

Examples

```
# # hts_builder(malaria_africa_ts, who_region, who_subregion, country)
```

---

hts_example_model   *Example model for use in testing and examples*

Description

Example model for use in testing and examples

Usage

```
data(hts_example_model)
```

Format

An object of class hts_inla (inherits from bru, iinla, inla) of length 14.

---

install_yahtsee_deps   *Install yahtsee dependencies*

Description

Currently this just installs the latest version of INLA

Usage

```
install_yahtsee_deps()
```

Value

nothing

Examples

```
## Not run:
install_yahtsee_deps()

## End(Not run)
```

---

malaria_africa_ts *Prevalence Rate data of Malaria in Africa*

Description

Data containing information on malaria prevalence in Africa. Data was extracted using the `malariaAtlas` package. If using this data, please cite:

Usage

```
data(malaria_africa_ts)
```

Format

An object of class `tbl_ts` (inherits from `tbl_df`, `tbl`, `data.frame`) with 1046 rows and 15 columns.

Details

Pfeffer, D.A., Lucas, T.C., May, D., Harris, J., Rozier, J., Twohig, K.A., Dalrymple, U., Guerra, C.A., Moyes, C.L., Thorn, M., Nguyen, M., et al. 2018. malariaAtlas: an R interface to global malariometric data hosted by the Malaria Atlas Project. Malaria Journal, 17(1), p.352.

Format: a 'tsibble' data frame with 1,117 observations on the following 15 variables.

who_region A character of WHO region, the largest region size.

who_subregion A character of WHO subregion, the second largest region size.

country A character of countries, the smallest region size.

country_id A three letter character code adhering to the standard, ISO 3166-1 alpha-3. For more detail see https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3.

continent_id A character of continent name

year integer of year data recorded

month integer of month data recorded

date date class of year+month recorded. Defaults to day = 1 as day was not provided.

lower_age lower age range

upper_age upper age range

examined integer, number of people examined

positive numeric, number of positive cases

pr positive rate - given from positive/examined

species Mosquito species

method method of measurement

predict.hts_inla   *Predict from hts_inla model*

Description

Predict from hts_inla model

Usage

```
## S3 method for class 'hts_inla'
predict(
  object,
  newdata = NULL,
  type = c("link", "response", "terms"),
  se.fit = FALSE,
  dispersion = NULL,
  terms = NULL,
  na.action = na.pass,
  ...
)
```

Arguments

object	a fitted object of class inheriting from `hts_model`
newdata	optionally, a data frame in which to look for variables with which to predict. If omitted, the fitted linear predictors are used.
type	the type of prediction required. The default is on the scale of the linear predictors; the alternative '"response"' is on the scale of the response variable. Thus for a default binomial model the default predictions are of log-odds (probabilities on logit scale) and 'type = "response"' gives the predicted probabilities. The '"terms"' option returns a matrix giving the fitted values of each term in the model formula on the linear predictor scale. The value of this argument can be abbreviated.
se.fit	logical switch indicating if standard errors are required.
dispersion	the dispersion of the GLM fit to be assumed in computing the standard errors. If omitted, that returned by 'summary' applied to the object is used.
terms	with 'type = "terms"' by default all terms are returned. A character vector specifies which terms are to be returned
na.action	function determining what should be done with missing values in 'newdata'. The default is to predict 'NA'.
...further	arguments passed to or from other methods.

Note

currently borrowing parameters + descriptions from predict.glm

Examples

```
## Not run:
predict(hts_example_model)

## End(Not run)
```

---

prediction_data   *Create prediction data*

Description

[Experimental]

Usage

```
prediction_data(model_data, key, index, ...)
```

Arguments

model_data	data used in the model
key	new key
index	new index
...	any other variables you want to create an example dataset of

Details

Create prediction data that has the same index and key as the model data for use in prediction.

Value

tsibble object with the same index and key as model_data.

Examples

```
date_range <- clock::date_build(2019, 2, 1:5)
date_range
countries <- c("Ethiopia", "Tanzania")
countries
prediction_data(
  model_data = malaria_africa_ts,
  key = countries,
  index = date_range
 )
```

---

tidy.hts_inla	*Tidy: summarise a model's statistical findings*

Description

The `tidy()` function provides a data frame of the mean, standard deviation, quantiles

Usage

```
## S3 method for class 'hts_inla'
tidy(x, ...)
```

Arguments

x   object of class "hts_inla"

...   extra arguments to pass to `tidy()`. Currently unused

Value

dataframe with column names: "variable", "mean", "sd", and quantiles for 0.025, 0.5, and 0.975, as well as mode and kld

Examples

```
hts_example_model
tidy(hts_example_model)
```

---

who_regions	*Who Regions*

Description

Data containing information on countries and WHO designated regions and ISO 3166-1 alpha-3 country code. From largest to smallest area it goes: "who_region", "who_subregion", then "country".

Usage

```
data(who_regions)
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 110 rows and 4 columns.

Details

Format: a data frame with 110 observations on the following 4 variables.

who_region A character of WHO region, the largest region size.

who_subregion A character of WHO subregion, the second largest region size.

country A character of countries, the smallest region size.

country_iso_code A three letter character code adhering to the standard, ISO 3166-1 alpha-3.
For more detail see https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3.

---

yahtsee_sitrep   *Checks if system requirements are present*

Description

Currently this just checks for INLA

Usage

```
yahtsee_sitrep()
```

Value

message on availability of system requirements

Examples

```
## Not run:
yahtsee_sitrep()

## End(Not run)
```

Index