

Package ‘ArvindSt’

May 11, 2026

Title Five Novel Stochastic Regression Models with Arvind-Distributed Errors and Effects

Version 1.0.0

Description Implements the 'Arvind' distribution and five novel stochastic regression models that replace the traditional Gaussian error assumption with 'Arvind'-distributed errors. The 'Arvind' distribution is a flexible single-parameter continuous distribution on the positive real line characterised by a polynomial numerator with Gaussian-type decay. The package provides complete distribution functions (`darvind()`, `parvind()`, `qarvind()`, `rarvind()`), maximum likelihood estimation via `fit_arvind_mle()`, and five model-fitting routines: Random Walk on Coefficients via `fit_rw1()`, Time-Varying Coefficient Linear Model via `fit_tvlm()`, Simulation-Extrapolation via `fit_simex()`, Mixed-Effects Regression via `fit_mixed()`, and Regime-Switching Hidden Markov Model via `fit_hmm()`. Additionally, provides Monte Carlo forecasting with prediction intervals via `forecast_arvind()`, comprehensive goodness-of-fit diagnostics (21 metrics and 25 plots) via `diagnostics_arvind()` and `plot_arvind()`, k-fold and rolling-window cross-validation via `cv_arvind()`, and unified model comparison via `summary_arvind()`. For more details see Pandey, Singh, Tyagi, and Tyagi (2024) "Modelling climate, COVID-19, and reliability data: A new continuous lifetime model under different methods of estimation", *Statistics and Applications*, 22(2), <<https://ssca.org.in/journal.html>>.

License MIT + file LICENSE

Depends R (>= 4.0.0)

Imports stats, graphics, grDevices, utils, ggplot2, forecast, tvReg, lme4, depmixS4, reshape2, rlang

Suggests testthat (>= 3.0.0), knitr, rmarkdown

Config/testthat/edition 3

Encoding UTF-8

Language en-US

RoxygenNote 7.3.3

VignetteBuilder knitr

NeedsCompilation no

Author Shikhar Tyagi [aut, cre] (ORCID:
<<https://orcid.org/0000-0003-1606-0844>>),
Arvind Pandey [aut]

Maintainer Shikhar Tyagi <shikhar1093tyagi@gmail.com>

Repository CRAN

Date/Publication 2026-05-11 18:20:02 UTC

Contents

arvind_mean_fn	2
arvind_var_fn	3
cv_arvind	4
darvind	5
diagnostics_arvind	6
fit_arvind_mle	7
fit_hmm	8
fit_mixed	9
fit_rw1	10
fit_simex	11
fit_tv1m	12
forecast_arvind	13
parvind	14
plot_arvind	15
qarvind	16
rarvind	16
rarvind_centred	17
simulate_arvind_data	18
summary_arvind	18
Index	20

arvind_mean_fn	<i>Mean of the Arvind Distribution</i>
----------------	--

Description

Computes the theoretical mean of the Arvind distribution with parameter θ by numerical integration.

Usage

```
arvind_mean_fn(theta)
```

Arguments

theta positive numeric scalar; the distribution parameter.

Value

A numeric scalar giving the theoretical mean, or NA if integration fails.

Examples

```
arvind_mean_fn(1)
arvind_mean_fn(2)
```

arvind_var_fn *Variance of the Arvind Distribution*

Description

Computes the theoretical variance of the Arvind distribution with parameter theta by numerical integration.

Usage

```
arvind_var_fn(theta)
```

Arguments

theta positive numeric scalar; the distribution parameter.

Value

A numeric scalar giving the theoretical variance, or NA if integration fails.

Examples

```
arvind_var_fn(1)
arvind_var_fn(2)
```

`cv_arvind`*K-Fold and Rolling-Window Cross-Validation*

Description

Performs k-fold cross-validation and optionally rolling-window (expanding-window) cross-validation for an ArvindFit model.

Usage

```
cv_arvind(fit, k_folds = 5, rolling = TRUE, n0_frac = 0.5, seed = 42)
```

Arguments

<code>fit</code>	an object of class "ArvindFit".
<code>k_folds</code>	integer; number of cross-validation folds (default: 5).
<code>rolling</code>	logical; if TRUE (default), also performs rolling-window cross-validation.
<code>n0_frac</code>	numeric; fraction of data used as initial training window for rolling CV (default: 0.5).
<code>seed</code>	integer; random seed for reproducibility (default: 42).

Value

A list with components:

cv_rmse numeric vector of length `k_folds`; per-fold RMSE.

cv_mae numeric vector of length `k_folds`; per-fold MAE.

mean_cv_rmse numeric; average k-fold RMSE.

mean_cv_mae numeric; average k-fold MAE.

roll_rmse numeric; rolling-window RMSE (or NA).

See Also

[diagnostics_arvind\(\)](#), [forecast_arvind\(\)](#)

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m1 <- fit_rw1(Y ~ X1 + X2 + X3, dat, seed = 42)
cv <- cv_arvind(m1, k_folds = 3, rolling = FALSE, seed = 42)
cv$mean_cv_rmse
```

darvind	<i>Arvind Distribution Density Function</i>
---------	---

Description

Computes the probability density function (PDF) of the Arvind distribution.

Usage

```
darvind(x, theta, log = FALSE)
```

Arguments

x	numeric vector of quantiles.
theta	positive numeric scalar; the distribution parameter.
log	logical; if TRUE, log-density is returned. Default FALSE.

Details

The Arvind distribution with parameter $\theta > 0$ has PDF

$$f(x; \theta) = \frac{\theta(1 + 2x + 2\theta x^2)}{(1 + \theta x)^2} \exp(-\theta x^2), \quad x > 0.$$

Value

A numeric vector of density values (or log-density values when `log = TRUE`).

Examples

```
# Evaluate the PDF at several points
darvind(c(0.5, 1, 2), theta = 1)

# Log-density
darvind(1, theta = 2, log = TRUE)

# Returns 0 for x <= 0
darvind(-1, theta = 1)
```

diagnostics_arvind *Goodness-of-Fit Diagnostics for Arvind Models*

Description

Computes 21 goodness-of-fit metrics for any fitted `ArvindFit` object, including MSE, RMSE, MAE, MAPE, R-squared, AIC, BIC, Kolmogorov-Smirnov test, Anderson-Darling statistic, and more.

Usage

```
diagnostics_arvind(fit)
```

Arguments

`fit` an object of class "ArvindFit" returned by any of the model-fitting functions.

Details

The following metrics are computed:

Model character; the model type.

MSE Mean Squared Error.

RMSE Root Mean Squared Error.

MAE Mean Absolute Error.

MAPE Mean Absolute Percentage Error.

R2 R-squared.

AdjR2 Adjusted R-squared.

AIC Akaike Information Criterion.

AICc Corrected AIC.

BIC Bayesian Information Criterion.

LogLik Log-likelihood at the MLE.

Bias Mean residual.

MASE Mean Absolute Scaled Error.

DW Durbin-Watson statistic.

LjungBox_stat Ljung-Box test statistic.

LjungBox_p Ljung-Box p-value.

Theta Estimated Arvind parameter.

KS_stat Kolmogorov-Smirnov test statistic.

KS_pvalue Kolmogorov-Smirnov p-value.

AD_stat Anderson-Darling test statistic.

CvM_stat Cramer-von Mises test statistic.

Value

A data frame with one row and 21 columns of diagnostics metrics. See Details for the full list.

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m1 <- fit_rw1(Y ~ X1 + X2 + X3, dat, seed = 42)
diagnostics_arvind(m1)
```

fit_arvind_mle

Maximum Likelihood Estimation for the Arvind Distribution

Description

Fits the Arvind distribution to a vector of positive observations by maximum likelihood. Optimisation is performed on the log-scale via the Brent method.

Usage

```
fit_arvind_mle(e_pos)
```

Arguments

`e_pos` numeric vector of strictly positive observations.

Value

A list with components:

theta numeric; the MLE of theta.

negloglik numeric; the minimised negative log-likelihood.

Examples

```
set.seed(42)
x <- rarvind(200, theta = 2)
fit_arvind_mle(x)
```

fit_hmm

*Fit Regime-Switching Regression (HMM)***Description**

Fits a hidden Markov model with state-dependent coefficients and Arvind-distributed errors. The EM algorithm with forward-backward recursions is used for parameter estimation, and the Viterbi algorithm decodes the most likely state sequence.

Usage

```
fit_hmm(formula, data, nstates = 2, seed = 42)
```

Arguments

formula	an object of class formula .
data	a data frame containing the variables in the formula.
nstates	integer; number of hidden states (default: 2).
seed	integer; random seed for reproducibility (default: 42).

Value

An object of class "ArvindFit", a list containing the same standard fields as [fit_rw1\(\)](#), plus:

hmm_fit the fitted depmixS4 object.
nstates integer; number of hidden states.
states integer vector; Viterbi-decoded state sequence.
trans_probs matrix; estimated transition probability matrix.
state_betas list of numeric vectors; state-specific coefficients.
state_sigmas numeric vector; state-specific standard deviations.

See Also

[diagnostics_arvind\(\)](#), [forecast_arvind\(\)](#), [cv_arvind\(\)](#)

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m5 <- fit_hmm(Y ~ X1 + X2 + X3, dat, nstates = 2, seed = 42)
m5$states
m5$states
```

Description

Fits a mixed-effects regression model with Arvind-distributed random effects and observation-level errors. Estimation uses a two-stage approach: REML initialisation via **lme4**, followed by Arvind MLE on the residuals.

Usage

```
fit_mixed(formula, data, group_var = "Season", re_formula = NULL, seed = 42)
```

Arguments

<code>formula</code>	an object of class formula specifying the fixed-effects structure.
<code>data</code>	a data frame containing the variables in the formula and the grouping variable.
<code>group_var</code>	character string; the name of the grouping variable in data (default: "Season").
<code>re_formula</code>	optional random-effects formula (e.g., $(1 + X1 \text{group})$). If NULL (default), a random intercept model (1group_var) is used.
<code>seed</code>	integer; random seed for reproducibility (default: 42).

Value

An object of class "ArvindFit", a list containing the same standard fields as [fit_rw1\(\)](#), plus:

lme_model the fitted `lme4::lmer` object.

theta_re numeric; Arvind parameter estimated from random effects.

group_var character; the grouping variable name.

See Also

[diagnostics_arvind\(\)](#), [forecast_arvind\(\)](#), [cv_arvind\(\)](#)

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m4 <- fit_mixed(Y ~ X1 + X2 + X3, dat, group_var = "Group", seed = 42)
m4$theta
m4$theta
```

fit_rw1

*Fit Random Walk on Coefficients Model (RW1-approx)***Description**

Fits a stochastic regression model with time-varying coefficients evolving as a random walk with Arvind-distributed innovations. The observation errors also follow the Arvind distribution.

Usage

```
fit_rw1(formula, data, theta_innov = 2, rw_scale = 0.01, seed = 42)
```

Arguments

formula	an object of class <code>formula</code> specifying the model (e.g., $Y \sim X1 + X2$).
data	a data frame containing the variables in the formula.
theta_innov	positive numeric; the Arvind parameter for state innovations (default: 2.0).
rw_scale	numeric; proportion of OLS coefficients used as innovation scale (default: 0.01).
seed	integer; random seed for reproducibility (default: 42).

Value

An object of class "ArvindFit", a list containing:

- model_type** character; "RW1-approx".
- fitted** numeric vector; fitted values.
- residuals** numeric vector; raw residuals.
- theta** numeric; estimated Arvind parameter for residuals.
- sigma** numeric; residual scale.
- shift** numeric; shift applied to residuals.
- e_pos** numeric vector; positive standardised residuals.
- negloglik** numeric; negative log-likelihood.
- beta_t** matrix; time-varying coefficient paths.
- beta_final** numeric vector; final coefficient values.
- sigma_rw** numeric vector; random walk innovation scales.
- theta_innov** numeric; Arvind parameter used for innovations.
- n** integer; number of observations.
- p** integer; number of parameters.
- X** matrix; design matrix.
- Y** numeric vector; response variable.
- formula** the model formula.
- data** the input data frame.

See Also

[diagnostics_arvind\(\)](#), [forecast_arvind\(\)](#), [cv_arvind\(\)](#)

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m1 <- fit_rw1(Y ~ X1 + X2 + X3, dat, seed = 42)
m1$theta
```

fit_simex

Fit Simulation-Extrapolation (SIMEX) Model

Description

Fits a regression model correcting for measurement error attenuation using the SIMEX algorithm with Arvind-distributed measurement noise and residuals.

Usage

```
fit_simex(
  formula,
  data,
  me_vars = NULL,
  me_frac = 0.05,
  lambda_grid = c(0.5, 1, 1.5, 2),
  n_sim = 100,
  theta_me = 2,
  seed = 123
)
```

Arguments

formula	an object of class formula .
data	a data frame containing the variables in the formula.
me_vars	character vector of covariate names measured with error. If NULL (default), the first two term labels are used.
me_frac	numeric; fraction of marginal variance used as measurement error variance (default: 0.05).
lambda_grid	numeric vector; SIMEX lambda grid (default: <code>c(0.5, 1, 1.5, 2)</code>).
n_sim	integer; number of SIMEX simulation replicates (default: 100).
theta_me	positive numeric; Arvind parameter for measurement error (default: 2.0).
seed	integer; random seed for reproducibility (default: 123).

Value

An object of class "ArvindFit", a list containing the same standard fields as `fit_rw1()`, plus:

beta numeric vector; SIMEX-corrected coefficient estimates.

simex_coefs matrix; coefficient estimates at each lambda level.

lambda_grid numeric vector; the SIMEX lambda grid used.

me_vars character vector; covariate names with measurement error.

sigma2_me named numeric vector; measurement error variances.

See Also

`diagnostics_arvind()`, `forecast_arvind()`, `cv_arvind()`

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m3 <- fit_simex(Y ~ X1 + X2 + X3, dat,
               me_vars = c("X1", "X2"),
               n_sim = 20, seed = 123)

m3$beta
m3$beta
```

fit_tvlm

Fit Time-Varying Coefficient Linear Model (tvLM)

Description

Fits a time-varying coefficient linear model using kernel-weighted least squares (via the **tvReg** package) with Arvind-distributed residuals.

Usage

```
fit_tvlm(formula, data, bw = NULL, seed = 42)
```

Arguments

formula	an object of class <code>formula</code> .
data	a data frame containing the variables in the formula.
bw	numeric or NULL; the bandwidth for kernel smoothing. If NULL (default), bandwidth is selected automatically via leave-one-out cross-validation.
seed	integer; random seed for reproducibility (default: 42).

Value

An object of class "ArvindFit", a list containing the same standard fields as `fit_rw1()`, plus:

tv_coefs matrix; time-varying coefficient estimates.

tv_fit the fitted tvReg: : tvLM object.

See Also

`diagnostics_arvind()`, `forecast_arvind()`, `cv_arvind()`

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m2 <- fit_tvlm(Y ~ X1 + X2 + X3, dat, bw = 0.5, seed = 42)
m2$theta
```

forecast_arvind

Monte Carlo Forecasting for Arvind Models

Description

Generates Monte Carlo forecasts with 80 percent and 95 percent prediction intervals for any fitted ArvindFit model. Covariates are forecast using SARIMA models (via the **forecast** package) if not supplied.

Usage

```
forecast_arvind(
  fit,
  newdata_sims = NULL,
  h = 120,
  nsim = 5000,
  covariate_models = NULL,
  seed = 123
)
```

Arguments

fit an object of class "ArvindFit".

newdata_sims optional named list of pre-computed covariate simulation matrices, each of dimension $h \times nsim$.

h integer; forecast horizon in time steps (default: 120).

nsim integer; number of Monte Carlo replicates (default: 5000).

covariate_models optional list of fitted SARIMA models for covariates (auto-fitted if NULL).

seed integer; random seed for reproducibility (default: 123).

Value

A list with components:

sims matrix (h x nsim); full simulation matrix.

mean numeric vector length h; mean forecast.

median numeric vector length h; median forecast.

lo80 numeric vector; lower 80 percent prediction interval.

hi80 numeric vector; upper 80 percent prediction interval.

lo95 numeric vector; lower 95 percent prediction interval.

hi95 numeric vector; upper 95 percent prediction interval.

See Also

[fit_rwl\(\)](#), [diagnostics_arvind\(\)](#), [cv_arvind\(\)](#)

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m1 <- fit_rwl(Y ~ X1 + X2 + X3, dat, seed = 42)
fc <- forecast_arvind(m1, h = 12, nsim = 100, seed = 42)
head(fc$mean)
```

 parvind

Arvind Distribution Function (CDF)

Description

Computes the cumulative distribution function (CDF) of the Arvind distribution.

Usage

```
parvind(q, theta, lower.tail = TRUE)
```

Arguments

q numeric vector of quantiles.

theta positive numeric scalar; the distribution parameter.

lower.tail logical; if TRUE (default), probabilities are $P(X \leq q)$; otherwise $P(X > q)$.

Details

The CDF is given by

$$F(x; \theta) = 1 - \frac{1}{1 + \theta x} \exp(-\theta x^2), \quad x > 0.$$

Value

A numeric vector of probabilities.

Examples

```
parvind(1, theta = 1)
parvind(c(0.5, 1, 2), theta = 2)
parvind(1, theta = 1, lower.tail = FALSE)
```

plot_arvind

Diagnostic Plots for Arvind Models

Description

Generates up to 25 diagnostic plots for a fitted `ArvindFit` object, including observed vs fitted, residual histogram with Arvind density overlay, Q-Q plot, ACF, ECDF comparison, and more.

Usage

```
plot_arvind(fit, output_dir = tempdir(), prefix = NULL)
```

Arguments

<code>fit</code>	an object of class "ArvindFit".
<code>output_dir</code>	character; directory where plots are saved. Defaults to a temporary directory.
<code>prefix</code>	character or NULL; prefix for plot filenames. If NULL, derived from the model type.

Value

The fit object is returned invisibly.

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m1 <- fit_rw1(Y ~ X1 + X2 + X3, dat, seed = 42)
plot_arvind(m1, output_dir = tempdir())
```

qarvind	<i>Arvind Distribution Quantile Function</i>
---------	--

Description

Computes quantiles of the Arvind distribution by numerical inversion of the CDF using [uniroot](#).

Usage

```
qarvind(p, theta)
```

Arguments

p	numeric vector of probabilities ($0 \leq p \leq 1$).
theta	positive numeric scalar; the distribution parameter.

Value

A numeric vector of quantiles.

Examples

```
qarvind(0.5, theta = 1)
qarvind(c(0.25, 0.5, 0.75), theta = 2)
```

rarvind	<i>Random Generation from the Arvind Distribution</i>
---------	---

Description

Generates random variates from the Arvind distribution using a rejection sampling algorithm with a half-normal proposal distribution.

Usage

```
rarvind(n, theta)
```

Arguments

n	positive integer; number of random variates to generate.
theta	positive numeric scalar; the distribution parameter.

Value

A numeric vector of length n containing positive random variates.

Examples

```
set.seed(42)
x <- rarvind(100, theta = 1)
summary(x)
```

rarvind_centred

Centred Random Generation from the Arvind Distribution

Description

Generates centred Arvind variates with approximately zero mean, suitable for use as error terms and innovation terms in stochastic regression models.

Usage

```
rarvind_centred(n, theta)
```

Arguments

n positive integer; number of random variates to generate.
theta positive numeric scalar; the distribution parameter.

Details

The centred variate is computed as $\tilde{\varepsilon} = \varepsilon - \mu_A(\theta)$, where $\varepsilon \sim \text{Arvind}(\theta)$ and $\mu_A(\theta)$ is the mean of the Arvind distribution.

Value

A numeric vector of length n with approximately zero mean.

Examples

```
set.seed(42)
eps <- rarvind_centred(1000, theta = 2)
mean(eps) # approximately 0
```

simulate_arvind_data *Generate Simulated Data for Examples*

Description

Creates a small simulated dataset that mimics the structure needed for demonstrating the **ArvindSt** model-fitting functions. Useful for examples and testing.

Usage

```
simulate_arvind_data(n = 60, seed = 42)
```

Arguments

n integer; number of observations to generate (default: 60).
seed integer; random seed for reproducibility (default: 42).

Value

A data frame with columns:

Y numeric; simulated response variable.

X1 numeric; first covariate.

X2 numeric; second covariate.

X3 numeric; third covariate.

Group factor; grouping variable with 4 levels.

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
head(dat)
```

summary_arvind *Summary and Comparison of Multiple Arvind Models*

Description

Accepts multiple ArvindFit objects, computes diagnostics for each, produces a unified comparison table, and prints the best model by RMSE, R-squared, and AIC.

Usage

```
summary_arvind(..., comparison_plots = TRUE, output_dir = tempdir())
```

Arguments

`...` one or more objects of class "ArvindFit".
`comparison_plots` logical; if TRUE (default), generate comparison plots.
`output_dir` character; directory to save comparison plots. Defaults to a temporary directory.

Value

A data frame of diagnostic metrics (one row per model) is returned invisibly.

See Also

[diagnostics_arvind\(\)](#), [plot_arvind\(\)](#)

Examples

```
dat <- simulate_arvind_data(n = 50, seed = 1)
m1 <- fit_rw1(Y ~ X1 + X2 + X3, dat, seed = 42)
summary_arvind(m1)
```

Index

arvind_mean_fn, [2](#)
arvind_var_fn, [3](#)

cv_arvind, [4](#)
cv_arvind(), [8](#), [9](#), [11–14](#)

darvind, [5](#)
diagnostics_arvind, [6](#)
diagnostics_arvind(), [4](#), [8](#), [9](#), [11–14](#), [19](#)

fit_arvind_mle, [7](#)
fit_hmm, [8](#)
fit_mixed, [9](#)
fit_rw1, [10](#)
fit_rw1(), [8](#), [9](#), [12–14](#)
fit_simex, [11](#)
fit_tv1m, [12](#)
forecast_arvind, [13](#)
forecast_arvind(), [4](#), [8](#), [9](#), [11–13](#)
formula, [8–12](#)

parvind, [14](#)
plot_arvind, [15](#)
plot_arvind(), [19](#)

qarvind, [16](#)

rarvind, [16](#)
rarvind_centred, [17](#)

simulate_arvind_data, [18](#)
summary_arvind, [18](#)

uniroot, [16](#)