

Function	calvertModel()
File version	1.0
Last updated	6 December 2020
Associated source code file	www.github.com/bsh2/Fitness-Fatigue-Models/tree/docs-and-utilities/software/utilities/calvertModel.R
Dependencies (packages required)	caret, stats, GA, parallel, foreach, iterators, doParallel, doRNG, deSolve
Functionality	Fit the discrete two-component fitness-fatigue model with additional fitness delay (Calvert et al. 1976) using a standard optimiser (L-BFGS-B) or tuneable genetic algorithm with stochastic local search. Exponentially decaying initial component estimates can be included in the fitting process. That 'burn off' fairly quickly but can sometimes provide more accurate results.
Underlying model: Discrete difference series	$\hat{p}(t) = p^* + k_g \cdot \sum_{i=1}^{t-1} \omega(i) \cdot \left(e^{\frac{-(t-i)}{\tau_{g1}} \text{effect}} - e^{\frac{-(t-i)}{\tau_{g2}} \text{delay}} \right) - k_h \cdot \sum_{i=1}^{t-1} \omega(i) \cdot e^{\frac{-(t-i)}{\tau_h}}$ <p>Or with initial components q_g, q_h that are estimated within the function</p> $\hat{p}(t) = p^* + k_g \cdot \sum_{i=1}^{t-1} \omega(i) \cdot \left(e^{\frac{-(t-i)}{\tau_{g1}} \text{effect}} - e^{\frac{-(t-i)}{\tau_{g2}} \text{delay}} \right) - k_h \cdot \sum_{i=1}^{t-1} \omega(i) \cdot e^{\frac{-(t-i)}{\tau_h}} + q_g e^{-\frac{t}{\tau_g}} - q_h e^{-\frac{t}{\tau_h}}$

1. Usage

```
calvertModel(inputData, constraints, method = "bfgs", startingValues = NULL,
doTrace = TRUE, initialComponent = FALSE, parscale = NULL, initialWindow = NULL,
testHorizon= NULL, expandRate = NULL, doParallel = FALSE, maxIt = 10000, popSize =
120, gaSelection = "gareal_tourSelection", gaCrossover = "gareal_blxCrossover",
gaMutation = "gareal_rsMutation", gaElitism = 7.5)
```

2. Arguments

Argument	Required	Category	Default	Details
inputData	Yes	Core arg.	No default	Data frame. Contains time-series data in sequential order. Three columns in order from left to right: “days”, “performances”, “loads”. Data ordered in equal time steps, controlled by the “days” column (i.e. 1 day). Example dataset available in repository. 0 values used in loads column to represent no training on a given day. NA values in performance column to represent no observed measure on a given day.
constraints	Yes	Core arg.	No default	<p>Box constraints for the parameter values. Supplied as a data frame with two columns (“lower”, “upper”), each containing non-named numeric values in a specific order.</p> <pre>data.frame("lower" = c(), "upper" = c())</pre> <p>Where order of the values in the lower and upper bound vectors comprising the data frame are as follows:</p> <p>If <code>initialComponent = FALSE</code> (default) $c(p^*, k_g, \tau_{g_1}, \tau_{g_2}, k_h, \tau_h)$</p> <p>If <code>initialComponent = TRUE</code> $c(p^*, k_g, \tau_{g_1}, \tau_{g_2}, k_h, \tau_h, q_g, q_h)$</p>
method	NO	Core arg.	“bfgs”	Options are the limited memory modification of the first order (with second order approximation) BFGS method (method = “bfgs”) or a tuneable genetic algorithm with stochastic local search (method = “ga”, see further arguments below)

Argument	Required	Category	Default	Details
startingValues	Conditional Yes if method = "bfgs"	Core arg	NULL	Starting set for the BFGS algorithm. Not required if method = "ga". Supplied as un-named vector of values ordered as follows: If initialComponent = FALSE (default) $c(p^*, k_g, \tau_{g_1}, \tau_{g_2}, k_h, \tau_h)$ If initialComponent = TRUE $c(p^*, k_g, \tau_{g_1}, \tau_{g_2}, k_h, \tau_h, q_g, q_h)$
doTrace	No	Optimisation	TRUE	TRUE/FALSE argument. Controls whether optimisation output is printed to the console. For method = "bfgs", a value between 1-6 can be specified, with higher values producing more output.
initialComponent	No	Core arg	FALSE	TRUE/FALSE argument. Controls whether to include initial decaying component estimates for fitness and fatigue (q_g, q_h) respectively. If TRUE, startingValues and constraints arguments must be of appropriate length
parscale	No	Optimisation	NULL	A vector of scaling values for the parameters. Optimization is performed on par/parscale and these should be comparable in the sense that a unit change in any element produces about a unit change in the scaled value. If NULL, defaults to $c(1, 1, 1, 1, 1, 1, 1)$ (if initialComponent = FALSE) or $c(1, 1, 1, 1, 1, 1, 1, 1)$ (if initialComponent = TRUE)
initialWindow	No	Cross-validation	NULL	Initial model training window size for out-of-sample walk forward method (expanding window). Supplied as a percentage value (without the % sign). If NULL, default will be set to 60% of the data (i.e. initialWindow = 60)

Argument	Required	Category	Default	Details
testHorizon	No	Cross-validation	NULL	Walk forward testing size. Supplied as percentage value (without % sign). If NULL, default will be set to 20% (i.e. <code>testHorizon = 20</code>)
expandRate	No	Cross-validation	NULL	Walk forward increment rate. Supplied as percentage value (without % sign). If NULL, default will be set to 4% (i.e. <code>expandRate = 4</code>)
doParallel	No	Optimisation	FALSE	TRUE/FALSE argument. Controls whether the optimisation will be split across available logical nodes. <u>Only applicable if</u> <code>method = "ga"</code> .
maxit	No	Optimisation	1000	The maximum number of iterations to run before the L-BFGS-B search is halted
popSize	No	Optimisation	120	The population size <u>Only applicable if</u> <code>method = "ga"</code> .
gaSelection	No	Optimisation	"gareal_tourSelection"	The selection function (i.e. function that generates a new population of individuals from the current population probabilistically according to individual fitness). See GA package documentation for other options. <u>Only applicable if</u> <code>method = "ga"</code> .
gaCrossover	No	Optimisation	"gareal_blx Crossover"	The crossover function (i.e. a function which forms offsprings by combining part of the genetic information from their parents). See GA package documentation for other options. <u>Only applicable if</u> <code>method = "ga"</code> .
gaMutation	No	Optimisation	"gareal_rsMutation"	The mutation function (i.e. a function which randomly alters the values of some genes in a parent chromosome). See GA package documentation for other options. <u>Only applicable if</u> <code>method = "ga"</code> .
gaElitism	No	Optimisation	7.5	The number of best fitness individuals to survive at each generation. By default top 7.5% will survive at each iteration. <u>Only applicable if</u> <code>method = "ga"</code> .

3. Value (output)

Returns an object of class list, with the following slots

Slot	Name	Type	Relates to	Contains
1	mainSet	List	Primary fitted model	Further objects: Summary information and parameters (summary), fit statistics (stats), model prediction series (predictions), raw optimisation data (optim)
2	crossValidation	List	Cross validation slice data	Further objects: Summary information (summary), model fit metrics across the CV slices (metrics), parameter values across slices (parameters), prediction series across slices (predictions), raw slice data (raw)