mvgam case study 3: distributed lag models

Nicholas Clark (n.clark@uq.edu.au (mailto:n.clark@uq.edu.au))

Here we will use the `mvgam` package, which fits dynamic GAMs using MCMC sampling via the `JAGS` software (Note that `JAGS` is required; installation links are found here (https://sourceforge.net/projects/mcmc-jags/files/)), to estimate paramaters of a Bayesian distributed lag model. These models are used to describe simultaneously non-linear and delayed functional relationships between a covariate and a response, and are sometimes referred to as exposure-lag-response models. If we assume $\tilde{\boldsymbol{y}}_t$ is the conditional expectation of a discrete response variable $\boldsymbol{y}$ at time $\boldsymbol{t}$, the linear predictor for a dynamic distributed lag GAM with one lagged covariate is written as:

$$log(\tilde{\boldsymbol{y}}_t) = \boldsymbol{B}_0 + \sum_{k=1}^{K} f(\boldsymbol{b}_{k,t}\boldsymbol{x}_{k,t}) + \boldsymbol{z}_t \, ,$$

where $\boldsymbol{B}_0$ is the unknown intercept, the $\boldsymbol{b}$'s are unknown spline coefficients estimating how the functional effect of covariate ($\boldsymbol{x}$) on $log(\tilde{\boldsymbol{y}}_t)$ changes over increasing lags (up to a maximum lag of ($\boldsymbol{K}$)) and $\boldsymbol{z}$ is a dynamic latent trend component.

To demonstrate how these models are estimated in `mvgam`, first we load the Portal rodents capture data, which are available from the `portalr` package

```
# devtools::install_github('nicholasjclark/mvgam')
library(mvgam)
library(dplyr)
portal_dat <- read.csv("https://raw.githubusercontent.com/nicholasjclark/mvgam/master/NEON_manuscript/Case studie
s/rodents_data.csv",
    as.is = T)
```

We'll keep data from the year 2004 onwards to make the model quicker to estimate for this simple example

```
portal_dat_all <- portal_dat %>% dplyr::filter(year >=
    2004) %>% dplyr::group_by(year, month) %>%
    dplyr::slice_head(n = 1)
```

Below is an exact reproduction of Simon Wood's lag matrix function (which he uses in his distributed lag example from his book Generalized Additive Models - An Introduction with R 2nd edition (https://www.taylorfrancis.com/books/mono/10.1201/9781315370279/generalized-additive-models-simon-wood)). Here we supply a vector and specify the maximum lag that we want, and it will return a matrix of dimension `length(x) * lag`. Note that `NAs` are used for the missing lag values at the beginning of the matrix. In essence, the matrix objects represent exposure histories, where each row represents the lagged values of the predictor that correspond to each observation in `y`

```
lagard <- function(x, n.lag = 6) {
    n <- length(x)
    X <- matrix(NA, n, n.lag)
    for (i in 1:n.lag) X[i:n, i] <- x[i:n -
        i + 1]
    X
}
```

Organise all data needed for modelling into a list. We will focus only on the species *Chaetodipus penicillatus* (labelled as `PP`), which shows reasonable seasonality in its captures over time

```
data_all <- list(lag = matrix(0:5, nrow(portal_dat_all),
    6, byrow = TRUE), y = portal_dat_all$PP,
    season = portal_dat_all$month, year = portal_dat_all$year,
    series = rep(as.factor("series1"), NROW(portal_dat_all)),
    time = 1:NROW(portal_dat_all))
data_all$precip <- lagard(portal_dat_all$precipitation)
data_all$mintemp <- lagard(portal_dat_all$mintemp)
```

The exposure history matrix elements of the data list look as follows:

```
head(data_all$lag, 5)
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]    0    1    2    3    4    5
## [2,]    0    1    2    3    4    5
## [3,]    0    1    2    3    4    5
## [4,]    0    1    2    3    4    5
## [5,]    0    1    2    3    4    5
```

```
head(data_all$precip, 5)
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,] 37.8   NA   NA   NA   NA   NA
## [2,]  8.7 37.8   NA   NA   NA   NA
## [3,] 43.5  8.7 37.8   NA   NA   NA
## [4,] 23.9 43.5  8.7 37.8   NA   NA
## [5,]  0.9 23.9 43.5  8.7 37.8   NA
```

```
head(data_all$mintemp, 5)
```

```
##        [,1]   [,2]   [,3]   [,4]  [,5] [,6]
## [1,] -9.710     NA     NA     NA    NA   NA
## [2,] -5.924 -9.710     NA     NA    NA   NA
## [3,] -0.220 -5.924 -9.710     NA    NA   NA
## [4,]  1.931 -0.220 -5.924 -9.710    NA   NA
## [5,]  6.568  1.931 -0.220 -5.924 -9.71   NA
```

All other elements of the data list are in the usual vector format

```
head(data_all$y, 5)
```

```
## [1]  0  1  2 NA 10
```

```
head(data_all$series, 5)
```

```
## [1] series1 series1 series1 series1 series1
## Levels: series1
```

```
head(data_all$year, 5)
```

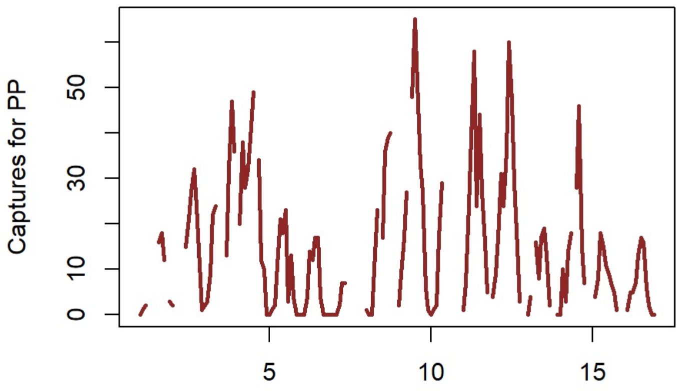
```
## [1] 2004 2004 2004 2004 2004
```

```
head(data_all$time, 5)
```

```
## [1] 1 2 3 4 5
```

View the raw series. There is a clear seasonal pattern to the data, and there are missing values scattered throughout

```
plot(ts(data_all$y, frequency = 12), ylab = "Captures for PP",
    xlab = "", lwd = 2.5, col = "#8F2727")
```



Create training and testing sets; start at observation 7 so that the `NA` values at the beginning of the covariate lag matrices are not included. Currently there is no option for on-the-fly imputation of missing covariate values in `mvgam` models, though this can easily be done in `JAGS` by specifying prior distributions over these missing entries

```
data_train <- list(lag = data_all$lag[7:174,
    ], y = data_all$y[7:174], series = data_all$series[7:174],
    season = data_all$season[7:174], year = data_all$year[7:174],
    time = 7:174, precip = data_all$precip[7:174,
        ], mintemp = data_all$mintemp[7:174,
        ])
data_test <- list(lag = data_all$lag[175:length(data_all$y),
    ], y = data_all$y[175:length(data_all$y)],
    series = data_all$series[175:length(data_all$y)],
    season = data_all$season[175:length(data_all$y)],
    year = data_all$year[175:length(data_all$y)],
    time = 175:length(data_all$y), precip = data_all$precip[175:length(data_all$y),
        ], mintemp = data_all$mintemp[175:length(data_all$y),
        ])
```

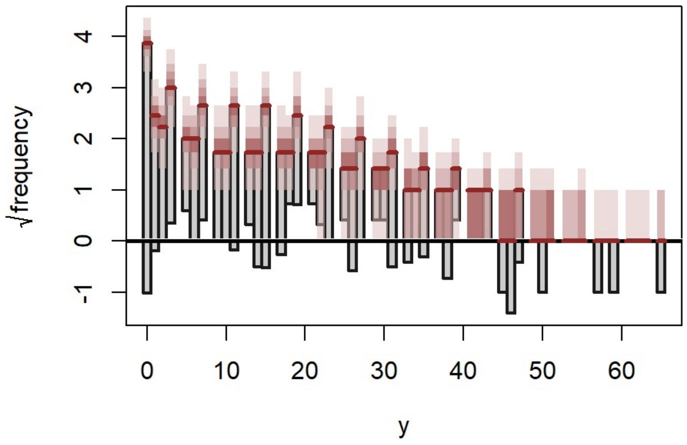
Now we can fit a Bayesian GAM with distributed lag terms for precipitation and minimum temperature. The distributed lags are set up as tensor product smooth functions (see `help(te)` for an explanation of tensor product smooth constructions in the `mgcv` package) between `lag` and each covariate. We will start simply by assuming our data follow a `Poisson` observation process

```
mod1 <- mvjagam(formula = y ~ te(mintemp,
    lag, k = c(8, 4)) + te(precip, lag, k = c(8,
    4)), data_train = data_train, data_test = data_test,
    family = "poisson", chains = 4, burnin = 25000,
    trend_model = "None")
```

```
## NOTE: Stopping adaptation
```

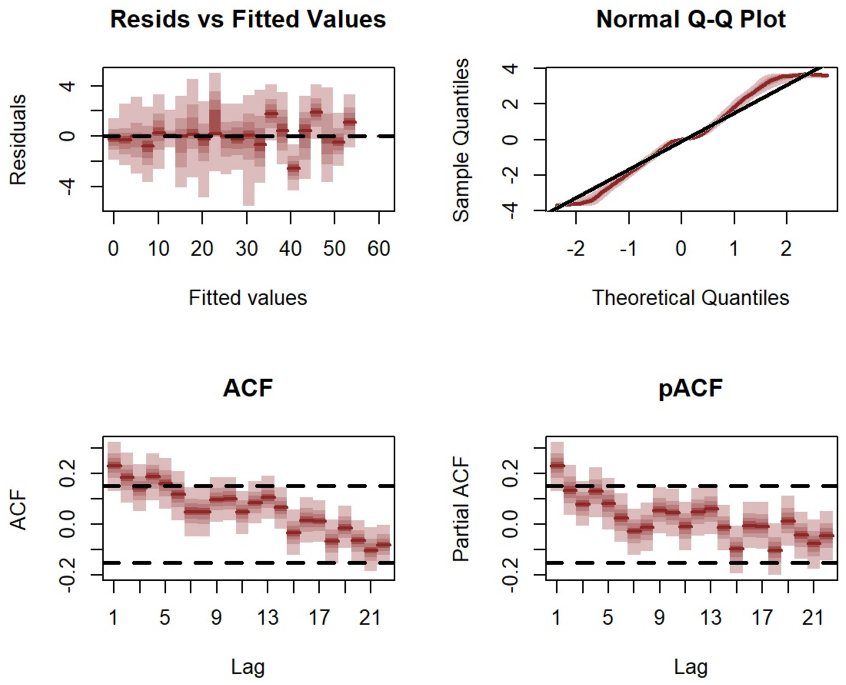
Posterior predictive rootograms are a useful way to explore whether a discrete model is able to capture relevant dispersion in the observed data. This plot compares the frequencies of observed vs predicted values for each bin, which can help to identify aspects of poor model fit. For example, if the gray bars (representing observed frequencies) tend to stretch below zero, this suggests the model's simulations predict the values in that particular bin less frequently than they are observed in the data. A well-fitting model that can generate realistic simulated data will provide a rootogram in which the lower boundaries of the grey bars are generally near zero

```
ppc(mod1, type = "rootogram")
```



The `Poisson` model is not doing a great job of capturing dispersion, underpredicting the zeros in the data and overpredicting some of the medium-range values (counts of `~5-30`). The residual `Q-Q` plot confirms that the `Poisson` is not an appropriate distribution for these data

```
plot(mod1, type = "residuals")
```



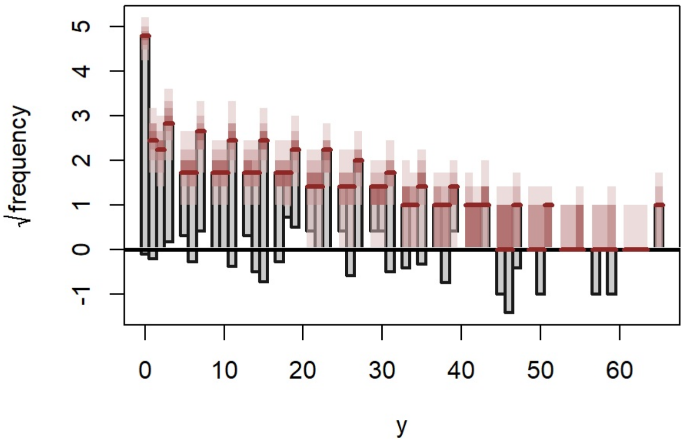
Given the overdispersion present in the data, we will now assume a `Geometric-Poisson` (https://en.wikipedia.org/wiki/Geometric_Poisson_distribution) observation model, which can be more flexible than the `Negative binomial` for modelling overdispersed count data (https://www.jstor.org/stable/2533492?origin=crossref&seq=1). In `mvgam` the `Geometric-Poisson` is estimated as a `Tweedie-Poisson` model with the power parameter `p` fixed at `1.5`

```
mod2 <- mvjagam(formula = y ~ te(mintemp,
    lag, k = c(8, 4)) + te(precip, lag, k = c(8,
    4)), data_train = data_train, data_test = data_test,
    family = "tw", chains = 4, burnin = 25000,
    trend_model = "None")
```

```
## NOTE: Stopping adaptation
```

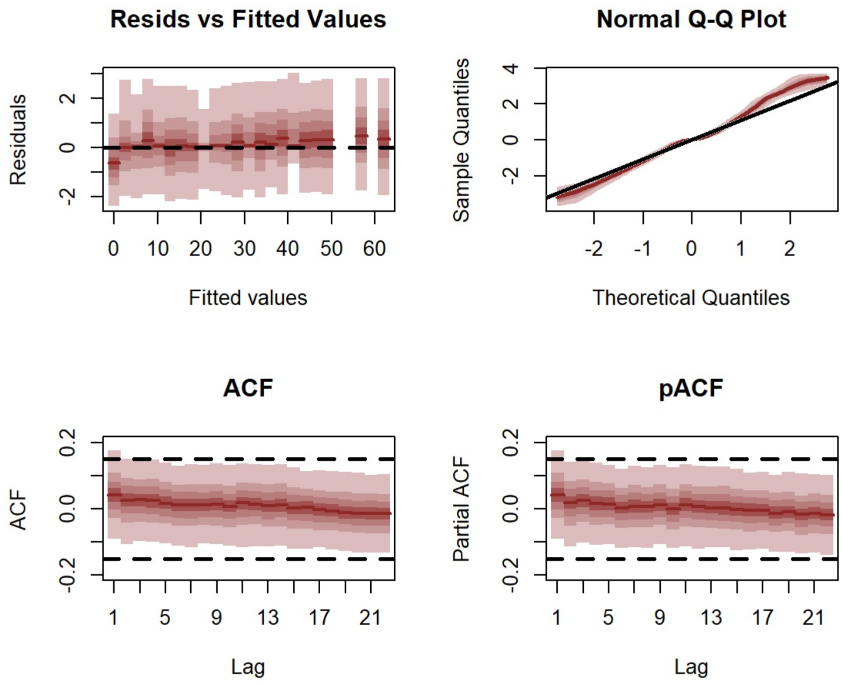
The rootogram for this model looks better, though of course there is still some overprediction of medium-range values

```
ppc(mod2, type = "rootogram")
```



However, the residual plot looks much better for this model

```
plot(mod2, type = "residuals")
```



The summary of the model provides useful information on convergence for unobserved parameters. Notice how strongly positive the overdispersion parameter is estimated to be, providing further evidence that this overdispersion is important to capture for these data

```
summary(mod2)
```

```
## GAM formula:
```

```
## y ~ te(mintemp, lag, k = c(8, 4)) + te(precip, lag, k = c(8,
##     4))
```

```
##
```

```
## Family:
```

```
## Tweedie
```

```
##
```

```
## Link function:
```

```
## log
```

```
##
```

```
## Trend model:
```

```
## None
```

```
##
```

```
## N series:
```

```
## 1
```

```
##
```

```
## N observations:
```

```
## 168
```

```
##
```

```
## Status:
```

```
## Fitted using runjags::run.jags()
```

```
##
```

```
## Dispersion parameter estimates:
```

```
##           2.5%      50%    97.5% Rhat n.eff
## twdis 1.189627 1.720465 2.641364 1.11   263
```

```
##
```

```
## Power parameter estimates:
```

```
##   2.5% 50% 97.5% Rhat n.eff
## p  1.5 1.5   1.5  NaN     0
```

```
##
```

```
## GAM smooth term approximate significances:
```

```
##                   edf Ref.df      F p-value    
## te(mintemp,lag) 18.80  30.00 14.004  <2e-16 ***
## te(precip,lag)  13.92  28.00  0.428   0.108    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## GAM coefficient (beta) estimates:
```

```
##                           2.5%          50%        97.5% Rhat n.eff
## (Intercept)         2.21515823  2.390286075  2.551999321 1.01   300
## te(mintemp,lag).1  -2.55907739 -1.070100877 -0.265471903 1.24    28
## te(mintemp,lag).2  -1.91051409 -0.815881257 -0.208315060 1.14    34
## te(mintemp,lag).3  -2.10041602 -0.577050299  0.532689169 1.20    28
## te(mintemp,lag).4  -0.86172508 -0.259425883  0.345407217 1.29    35
## te(mintemp,lag).5  -0.15092688  0.431543194  1.153983257 1.19    29
## te(mintemp,lag).6   0.24262519  0.785323465  1.473250332 1.08    31
## te(mintemp,lag).7   0.27845281  0.844835132  1.611084106 1.22    41
## te(mintemp,lag).8  -0.76208381 -0.358913164  0.200391341 1.30    44
## te(mintemp,lag).9  -0.34278182 -0.006525482  0.270464982 1.30    45
## te(mintemp,lag).10 -0.01599651  0.260952214  0.570505600 1.17    36
## te(mintemp,lag).11  0.01123885  0.468705397  1.098454005 1.21    37
## te(mintemp,lag).12 -0.45424500  0.059111958  0.657941263 1.30    34
## te(mintemp,lag).13 -0.32476123  0.160820366  0.726046671 1.36    39
## te(mintemp,lag).14 -0.02399687  0.321733551  0.877478699 1.14    29
## te(mintemp,lag).15 -0.05693148  0.504282048  1.226905646 1.29    30
## te(mintemp,lag).16 -0.28113840  0.175297664  0.757622682 1.29    28
## te(mintemp,lag).17 -0.13138760  0.251610173  0.744237594 1.14    35
## te(mintemp,lag).18 -0.09279971  0.262272907  0.678302489 1.11    39
## te(mintemp,lag).19 -0.25505360  0.342195715  1.041722083 1.29    35
## te(mintemp,lag).20 -0.31191807  0.272896962  0.806705910 1.59    40
## te(mintemp,lag).21 -0.09749038  0.323426708  0.724397348 1.20    57
## te(mintemp,lag).22 -0.09588235  0.269809557  0.867502230 1.20    56
## te(mintemp,lag).23 -0.26573476  0.221443687  1.031921269 1.40    48
## te(mintemp,lag).24  0.15037507  0.609289101  1.426606945 1.37    34
## te(mintemp,lag).25  0.08104738  0.489065815  1.040126577 1.19    38
## te(mintemp,lag).26 -0.17528312  0.234696247  0.769343915 1.12    29
## te(mintemp,lag).27 -0.67959516 -0.146385707  0.417578161 1.14    33
## te(mintemp,lag).28  0.01948300  0.622360588  1.419232875 1.41    38
## te(mintemp,lag).29 -0.05263451  0.301310046  0.688753976 1.18    51
## te(mintemp,lag).30 -0.44597385 -0.121728981  0.293326848 1.13    53
## te(mintemp,lag).31 -1.33711973 -0.651716552 -0.145049234 1.18    50
## te(precip,lag).1   -0.23289866  0.033500408  0.155467249 1.49    79
## te(precip,lag).2   -0.26251568 -0.093365646  0.060797156 1.15    73
## te(precip,lag).3   -0.51553181 -0.149424902  0.105963406 1.26    58
## te(precip,lag).4    0.05770746  0.216867858  0.527719830 1.40    74
## te(precip,lag).5   -0.23547368  0.010807956  0.096079471 1.61    79
## te(precip,lag).6   -0.31415047 -0.114165440  0.015600561 1.26    56
## te(precip,lag).7   -0.52385046 -0.169208271  0.040510351 1.31    51
## te(precip,lag).8   -0.05245172  0.127169908  0.429929717 1.24    59
## te(precip,lag).9   -0.24263515 -0.026001321  0.090484067 1.39    81
## te(precip,lag).10  -0.40022476 -0.129425082  0.007689021 1.25    62
## te(precip,lag).11  -0.62357858 -0.182426971  0.020444622 1.27    61
## te(precip,lag).12  -0.30006248 -0.036803363  0.334013232 1.10    41
## te(precip,lag).13  -0.24148192 -0.083192750  0.036858668 1.19    72
## te(precip,lag).14  -0.37636908 -0.126167037  0.043230070 1.14    52
## te(precip,lag).15  -0.67106299 -0.121345819  0.084804693 1.25    68
## te(precip,lag).16  -0.43050161 -0.115034956  0.228711116 1.09    44
## te(precip,lag).17  -0.35934866 -0.151355419 -0.008418841 1.26    79
## te(precip,lag).18  -0.45721121 -0.156650595  0.039536465 1.15    56
## te(precip,lag).19  -0.73317816 -0.170569354  0.110772365 1.23    52
## te(precip,lag).20  -0.13687621  0.081970236  0.400405197 1.24    62
## te(precip,lag).21  -0.14429101 -0.025125014  0.096402956 1.02    84
## te(precip,lag).22  -0.30001494 -0.138436036 -0.004212953 1.09    85
## te(precip,lag).23  -0.47123328 -0.196394875  0.029631232 1.21    71
## te(precip,lag).24  -0.21286577 -0.006454499  0.238309223 1.09    63
## te(precip,lag).25  -0.82698232 -0.089899828  0.516871802 1.05    34
## te(precip,lag).26  -0.50765928 -0.009553087  0.396038101 1.05    31
## te(precip,lag).27  -0.46712505  0.025662348  0.382817320 1.07    27
## te(precip,lag).28  -0.68079193 -0.066342797  0.565613223 1.19    27
```

```
##
```

```
## GAM smoothing parameter (rho) estimates:
```

```
##                        2.5%       50%    97.5% Rhat n.eff
## te(mintemp,lag)   1.0713950  2.735880 4.037811 1.17    52
## te(mintemp,lag)2  0.6116828  2.394352 4.041955 1.11    38
## te(mintemp,lag)3 -2.2348927 -0.192166 2.022092 1.06    67
## te(precip,lag)    2.4269898  3.735848 4.765889 1.01   184
## te(precip,lag)2  -0.2082284  2.329281 4.331516 1.53    65
## te(precip,lag)3  -0.3320276  2.151520 4.076637 1.08    67
```

```
##
```

```
## Latent trend drift (phi) and AR parameter estimates:
```

```
##     2.5% 50% 97.5% Rhat n.eff
## phi    0   0     0  NaN     0
## ar1    0   0     0  NaN     0
## ar2    0   0     0  NaN     0
## ar3    0   0     0  NaN     0
```

```
##
```

As this is a timeseries and the residual plot hints at some autocorrelation remaining in the short-term lags, lets check if an `AR` latent trend process improves forecasts compared to the no-trend model. An important note here is the choice of prior for the overdispersion parameter `twdis`. This parameter and the latent trend variance can interact strongly, particularly when overdispersion is in the data high. This is because at high values of the dispersion parameter, there is less need for a latent trend to be able to capture any outliers and so the latent trend precision can go up toward infinity, approaching a space of very diffuse likelihood that forces the Gibbs samplers to take on frustratingly small step sizes. Likewise when there is not much need for overdispersion, the dispersion parameter can approach zero and move around in an equally uninformative parameter space. The latent trend operates on the log scale, so really we should not expect autocorrelated jumps in trappings of more than `6-8` from timepoint to timepoint (any larger and the trend will compete strongly with the overdispersion parameter, making it difficult for us to model the inherent overdispersion process and instead assuming it is all autocorrelation). A containment prior on the latent trend `sigma` will help achieve this

```
mod3 <- mvjagam(formula = y ~ te(mintemp,
    lag, k = c(8, 4)) + te(precip, lag, k = c(8,
    4)), data_train = data_train, data_test = data_test,
    family = "tw", sigma_prior = "dexp(2.5)T(0,2)",
    chains = 4, burnin = 25000, trend_model = "AR3")
```

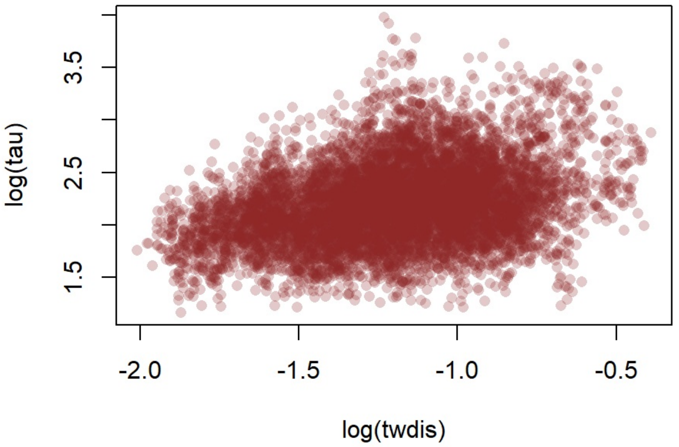
```
## NOTE: Stopping adaptation
```

```
summary(mod3)
```

```
## y ~ te(mintemp, lag, k = c(8, 4)) + te(precip, lag, k = c(8,
##     4))
## Tweedie
## log
## AR3
## 1
## 168
## Fitted using runjags::run.jags()
##            2.5%       50%     97.5% Rhat n.eff
## twdis 0.1654472 0.3059042 0.5180687 1.06    33
##   2.5% 50% 97.5% Rhat n.eff
## p  1.5 1.5   1.5  NaN     0
##                    edf Ref.df      F p-value
## te(mintemp,lag) 15.34  30.00 10.055  <2e-16 ***
## te(precip,lag)  14.69  28.00  0.099   0.904
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##                            2.5%          50%       97.5% Rhat n.eff
## (Intercept)         1.84756674  2.273995602  2.63894382 1.08    60
## te(mintemp,lag).1  -0.59876582 -0.121561567  0.24903938 1.26    59
## te(mintemp,lag).2  -0.85324575 -0.371446179  0.10581310 1.31    49
## te(mintemp,lag).3  -1.33318217 -0.766558937 -0.06441739 1.09    55
## te(mintemp,lag).4  -1.00011238 -0.640445163 -0.26598997 1.32    63
## te(mintemp,lag).5  -0.49641384 -0.175621069  0.15815146 1.26    42
## te(mintemp,lag).6  -0.07792297  0.194034219  0.50601653 1.18    52
## te(mintemp,lag).7   0.12739217  0.401021827  0.70752256 1.32    78
## te(mintemp,lag).8  -0.67293811 -0.405533355 -0.10095897 1.40    71
## te(mintemp,lag).9  -0.28043421 -0.113702587  0.10833580 1.35    50
## te(mintemp,lag).10 -0.01218030  0.134489550  0.31940159 1.15    61
## te(mintemp,lag).11  0.15237095  0.401433302  0.65071390 1.47    71
## te(mintemp,lag).12 -0.35806922 -0.087061959  0.22769078 1.38    61
## te(mintemp,lag).13 -0.25752400 -0.038567650  0.23402842 1.20    47
## te(mintemp,lag).14 -0.08552845  0.116141089  0.29459974 1.13    63
## te(mintemp,lag).15  0.11743614  0.352621766  0.62311785 1.54    70
## te(mintemp,lag).16 -0.24357205  0.025730110  0.29063726 1.15    53
## te(mintemp,lag).17 -0.20005199  0.019812870  0.24128071 1.09    48
## te(mintemp,lag).18 -0.13883100  0.090906936  0.26357522 1.06    68
## te(mintemp,lag).19 -0.03071544  0.251765845  0.51507825 1.51    65
## te(mintemp,lag).20 -0.20996192  0.066545701  0.36321145 1.19    63
## te(mintemp,lag).21 -0.19993205  0.046944928  0.23864361 1.18    52
## te(mintemp,lag).22 -0.18974116 -0.013189701  0.18139174 1.15    77
## te(mintemp,lag).23 -0.25568671  0.007575425  0.23148771 1.25    60
## te(mintemp,lag).24 -0.02140085  0.299969671  0.72870819 1.36    49
## te(mintemp,lag).25 -0.17963093  0.063963653  0.37563085 1.22    49
## te(mintemp,lag).26 -0.39565657 -0.203282739  0.03040414 1.13    66
## te(mintemp,lag).27 -0.72260956 -0.505011227 -0.15878010 1.23    87
## te(mintemp,lag).28 -0.04871209  0.354362465  0.79473135 1.37    54
## te(mintemp,lag).29 -0.18499865  0.083371236  0.35735846 1.10    50
## te(mintemp,lag).30 -0.54473963 -0.289672776 -0.04453311 1.22    56
## te(mintemp,lag).31 -1.03458027 -0.730799137 -0.29733653 1.12    83
## te(precip,lag).1   -0.14418297  0.017745020  0.14360519 1.11    53
## te(precip,lag).2   -0.20914458 -0.047956032  0.07660649 1.17    60
## te(precip,lag).3   -0.26900186 -0.068037597  0.17724196 1.44    57
## te(precip,lag).4   -0.03660461  0.117451215  0.27397660 1.19    41
## te(precip,lag).5   -0.10013953 -0.005371984  0.08477807 1.08    69
## te(precip,lag).6   -0.19900300 -0.047927530  0.04760728 1.31    66
## te(precip,lag).7   -0.26617928 -0.062983708  0.17700491 1.37    54
## te(precip,lag).8   -0.13166211  0.026619422  0.17743100 1.19    63
## te(precip,lag).9   -0.12650305 -0.032753788  0.08719536 1.06    74
## te(precip,lag).10  -0.22600443 -0.038300475  0.08266920 1.29    55
## te(precip,lag).11  -0.27460220 -0.067404908  0.18650439 1.20    51
## te(precip,lag).12  -0.23360524 -0.071377532  0.14420990 1.21    51
## te(precip,lag).13  -0.14566161 -0.032934062  0.08839007 1.04    66
## te(precip,lag).14  -0.17728246 -0.023979818  0.11469389 1.16    50
## te(precip,lag).15  -0.27378299 -0.060972989  0.15890749 1.14    48
## te(precip,lag).16  -0.35029766 -0.124887639  0.09247378 1.11    53
## te(precip,lag).17  -0.25341538 -0.091421137  0.08806972 1.10    52
## te(precip,lag).18  -0.27188373 -0.073572754  0.11773779 1.13    49
## te(precip,lag).19  -0.37198116 -0.091918407  0.15026854 1.18    52
## te(precip,lag).20  -0.12725284  0.058440869  0.23835292 1.25    67
## te(precip,lag).21  -0.16981487 -0.037220622  0.08789929 1.12    67
## te(precip,lag).22  -0.23286311 -0.088787763  0.02346266 1.05    67
## te(precip,lag).23  -0.26257358 -0.063739627  0.10230471 1.16    66
## te(precip,lag).24  -0.28753629 -0.074993197  0.09397680 1.18    57
## te(precip,lag).25  -0.78142761 -0.078617832  0.31628258 1.12    34
## te(precip,lag).26  -0.56445678 -0.029930702  0.21791400 1.16    36
## te(precip,lag).27  -0.41113176 -0.010832619  0.24307990 1.19    41
## te(precip,lag).28  -0.43622900 -0.015138945  0.32772613 1.17    53
##                         2.5%      50%    97.5% Rhat n.eff
## te(mintemp,lag)   2.22405908 3.382230 4.429000 1.09   108
## te(mintemp,lag)2  0.98420042 2.761682 4.187138 1.20   101
## te(mintemp,lag)3 -0.06265259 1.968263 3.970141 1.13    98
## te(precip,lag)    2.79809300 3.961591 4.895301 1.02   366
## te(precip,lag)2   0.97438562 2.527281 4.002889 1.11   112
## te(precip,lag)3   0.46786616 2.852954 4.309066 1.09   122
##           2.5%         50%     97.5% Rhat n.eff
## phi  0.0000000  0.00000000 0.0000000  NaN     0
## ar1  0.5001043  0.81330670 1.1417393 1.01   505
## ar2 -0.2565342  0.16049682 0.5310202 1.00   429
## ar3 -0.3708170 -0.07342271 0.2277921 1.00   504
```

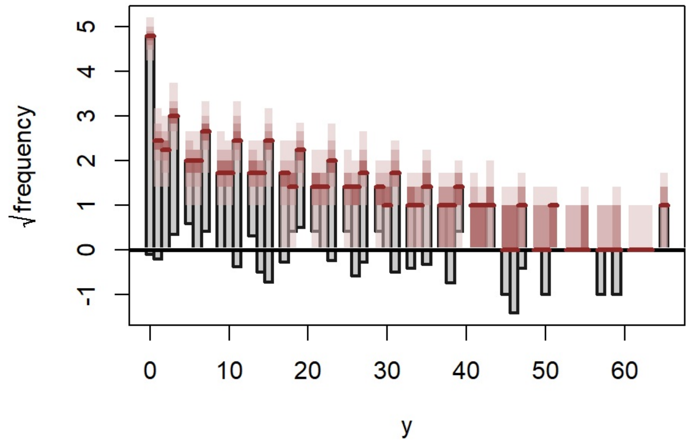
A pairs plot of the logged versions of the latent trend precision and the overdispersion parameter suggest there is no strange behaviour in the joint posterior

```
plot(log(MCMCvis::MCMCchains(mod3$jags_output,
    "twdis")), log(MCMCvis::MCMCchains(mod3$jags_output,
    "tau")), ylab = "log(tau)", xlab = "log(twdis)",
    pch = 16, col = "#8F272740")
```



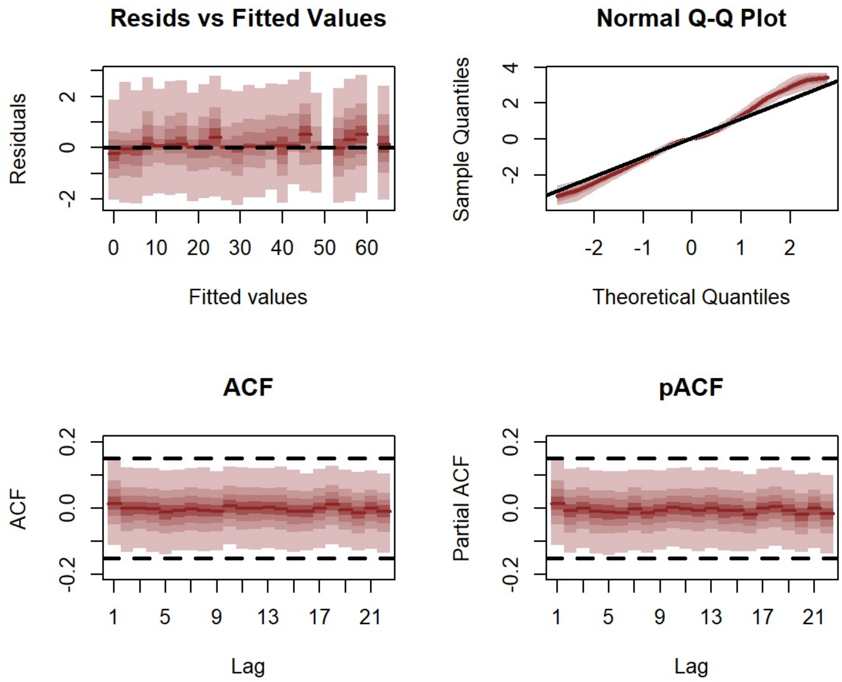
Our rootogram has not improved much with the addition of the latent trend

```
ppc(mod3, type = "rootogram")
```



But there is no more evidence of autocorrelation in the residuals

```
plot(mod3, type = "residuals")
```



The version with the latent trend does reduce in-sample DIC...

```
dic(mod2)
```

```
## Mean deviance:  667.6
## penalty 118.9
## Penalized deviance: 786.5
```

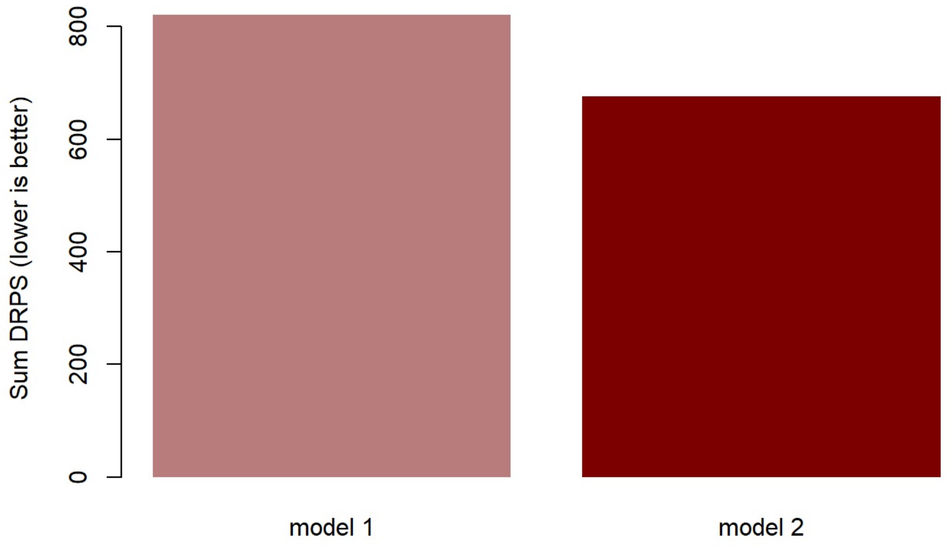
```
dic(mod3)
```

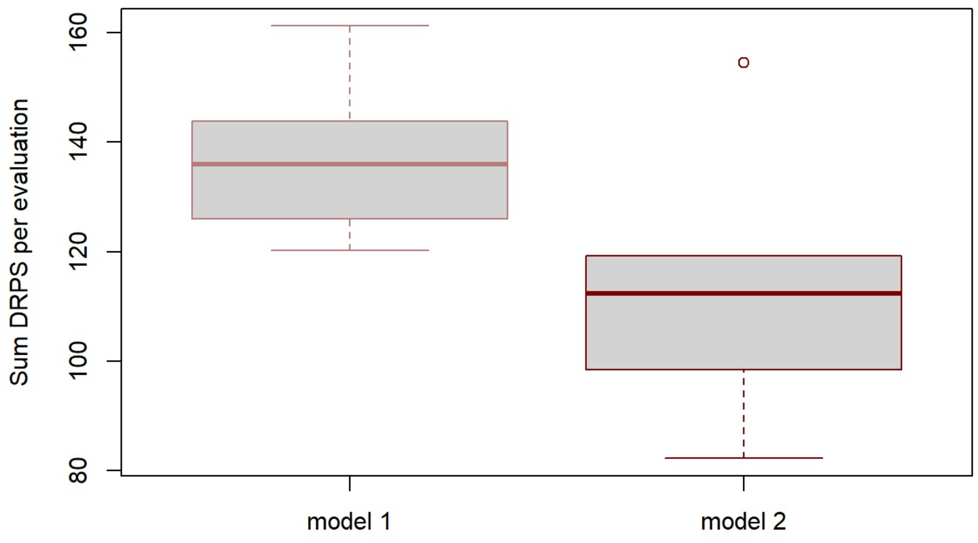
```
## Mean deviance:  657.3
## penalty 105.2
## Penalized deviance: 762.6
```

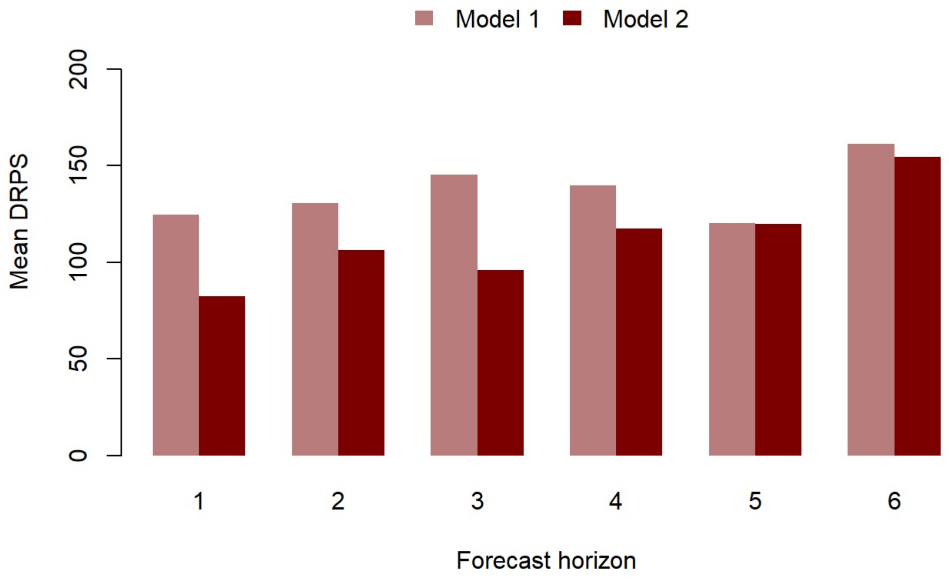
...and comparing models based on their forecasting abilities also suggests the latent trend model is superior. We will therefore stick with `mod3` , which is the distributed lag `Tweedie-Poisson` model with an `AR3` dynamic process

```
compare_mvgams(model1 = mod2, model2 = mod3,
    fc_horizon = 6, n_evaluations = 30, n_cores = 5)
```

```
## DRPS summaries per model (lower is better)
##              Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
## Model 1 120.16291 126.00038 135.0444 136.8593 143.8114 161.1846
## Model 2  82.27402  98.40519 111.9087 112.6984 119.2602 154.4841
##
## 90% interval coverages per model (closer to 0.9 is better)
## Model 1 0.9582604
## Model 2 0.9734119
```

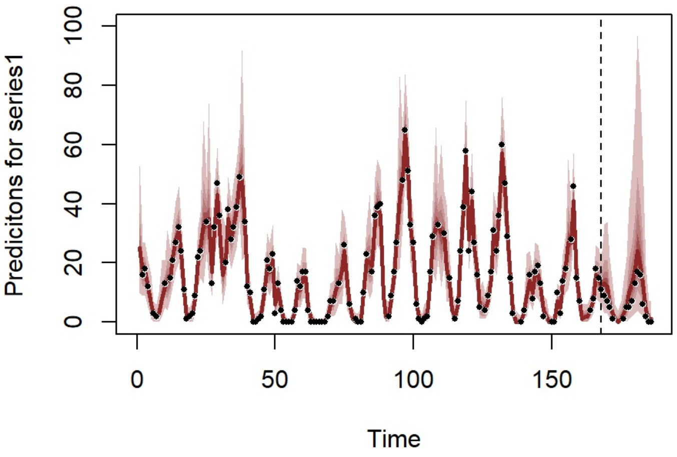

Sum DRPS (lower is better)
0
200
400
600
800
model 1
model 2


Sum DRPS per evaluation
80
100
120
140
160
model 1
model 2


Model 1
Model 2
Mean DRPS
0
50
100
150
200
1
2
3
4
5
6
Forecast horizon

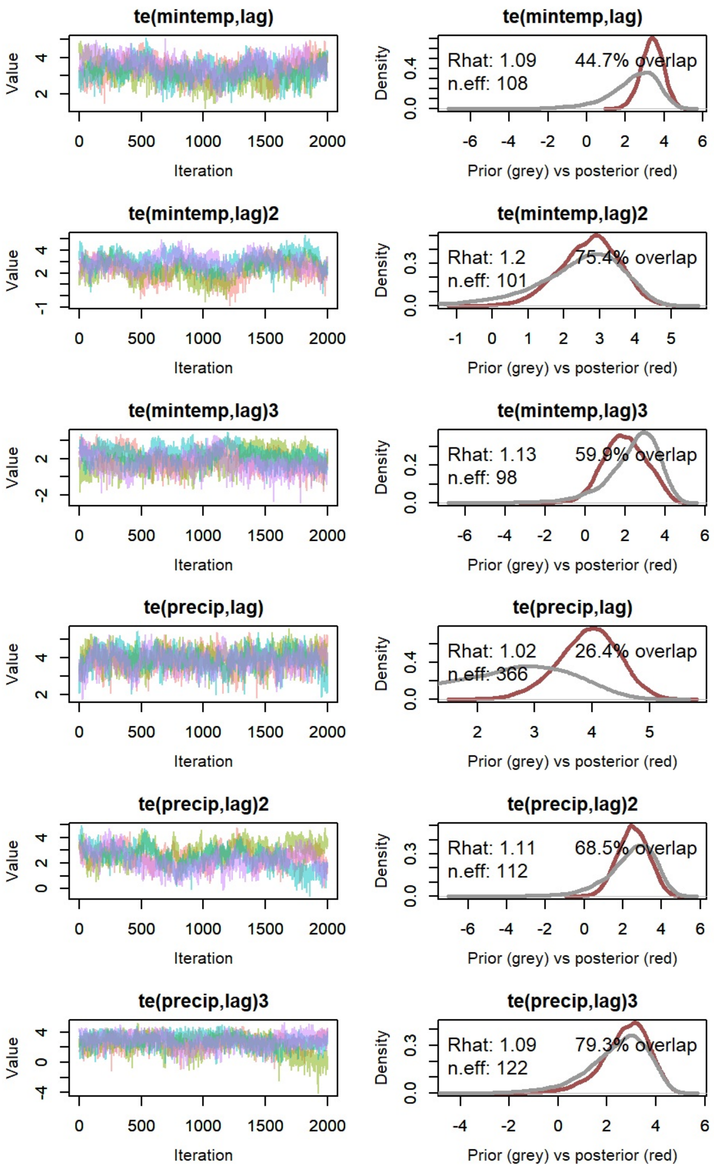
As with all other `mvgam` objects, we can create plots of the estimated forecast distribution

```
plot_mvgam_fc(mod3, series = 1, data_test = data_test,
    ylim = c(0, 100))
```



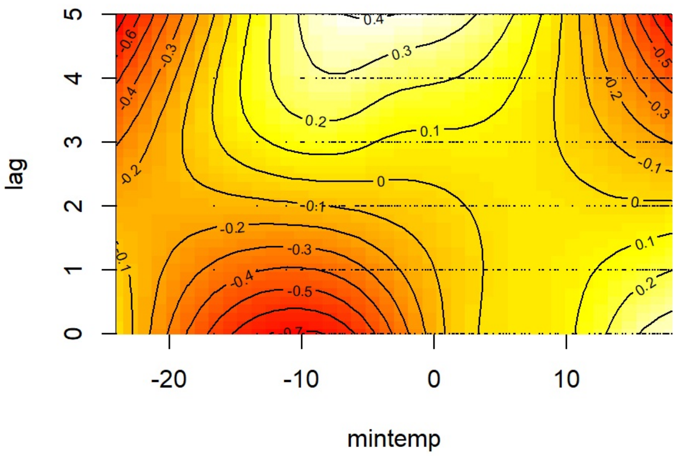
Traceplots of smooth penalties indicate good mixing and convergence of the four MCMC chains

```
plot_mvgam_trace(mod3, "rho")
```

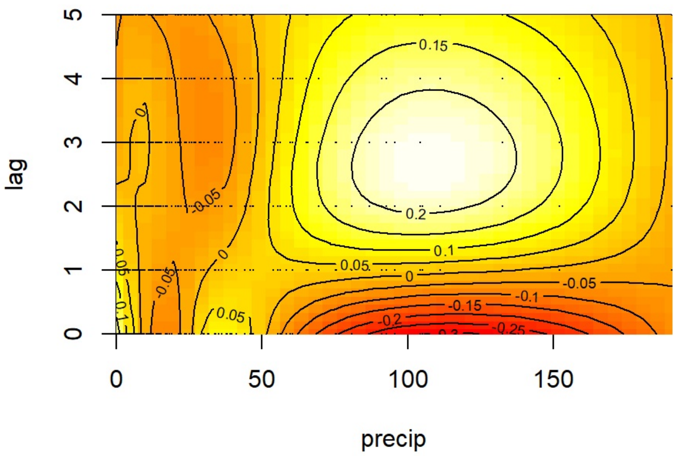


We can also create quick plots of the estimated smooth tensor product interactions for the distributed lag terms, which basically follow `mgcv` 's two-dimensional plotting utility but uses the `mvgam` 's estimated coefficients

```
plot_mvgam_smooth(mod3, series = 1, smooth = 1)
```



```
plot_mvgam_smooth(mod3, series = 1, smooth = 2)
```

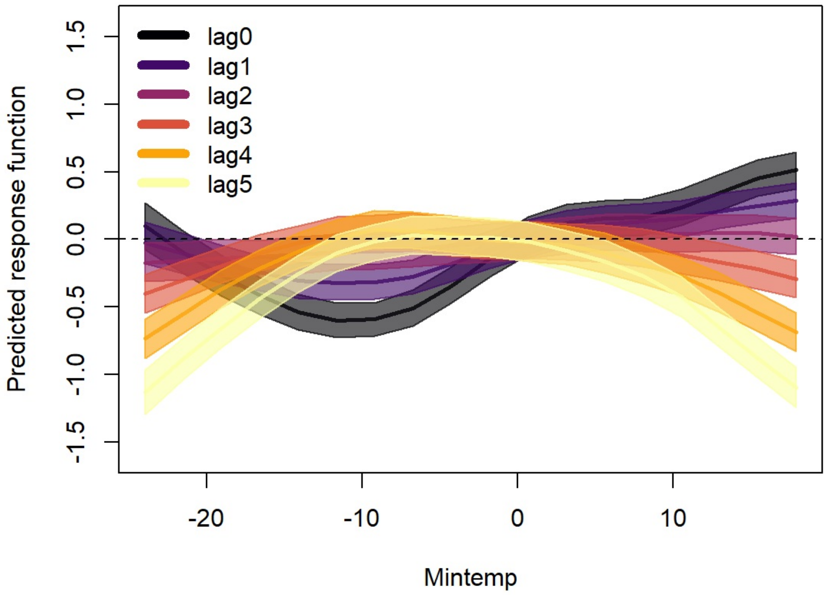


If you are like me then you'll find these plots rather difficult to interpret! The more intense yellow/white colours indicate higher predicted values, with the deeper red colours representing lower predicted values, but actually making sense of how the functional response is expected to change over different lags is not easy from these plots. HOwever, we can use the `predict_mvgam` function to generate much more interpretable plots. First we will focus on the effect of `mintemp` and generate a series of predictions to visualise how the estimated function changes over different lags. Set up prediction data by zeroing out all covariates apart from the covariate of interest

```
newdata <- data_test
newdata$year <- rep(0, length(newdata$year))
newdata$season <- rep(0, length(newdata$season))
newdata$precip <- matrix(0, ncol = ncol(newdata$precip),
    nrow = nrow(newdata$precip))
```

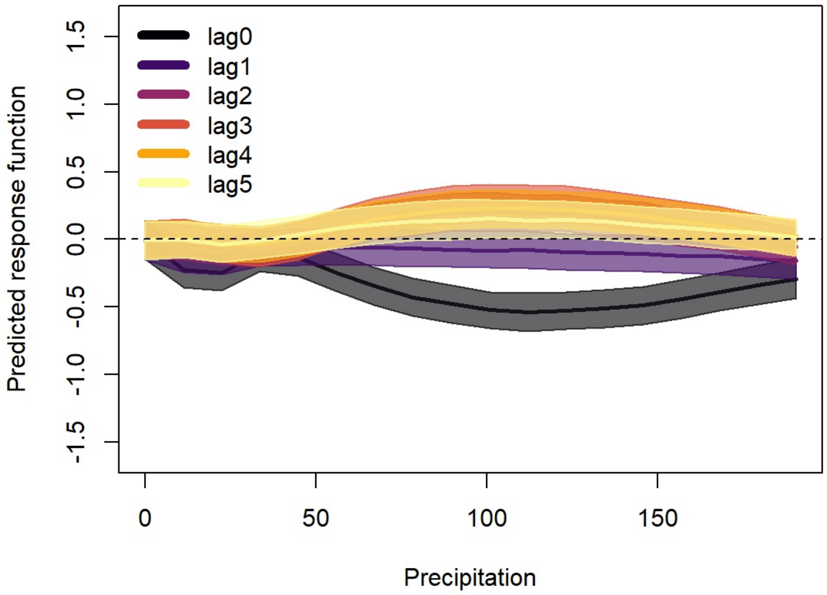
Set up `viridis` plot colours and initiate the plot window to be centred around zero. We will then keep all `mintemp` values at zero apart from the particular lag being predicted so that we can visualise how the predicted function changes over lags of `mintemp`. Predictions are generated on the link scale in this case, though you could also use the response scale. Note that we need to first generate predictions with all covariates (including the `mintemp` covariate) zeroed out to find the 'baseline' prediction so that we can shift by this baseline for generating a zero-centred plot. That way our resulting plot will roughly follow the traditional `mgcv` partial effect plots

```
cols <- viridis::inferno(6)
plot(1, type = "n", xlab = "Mintemp", ylab = "Predicted response function",
    xlim = c(min(data_train$mintemp), max(data_train$mintemp)),
    ylim = c(-1.6, 1.6))

# Calculate predictions for when mintemp
# is all zeros to find the baseline value
# for centring the plot
newdata$mintemp <- matrix(0, ncol = ncol(newdata$mintemp),
    nrow = nrow(newdata$mintemp))
preds <- predict(mod3, series = 1, newdata = newdata,
    type = "link")
offset <- mean(preds)

for (i in 1:6) {
    # Set up prediction matrix for mintemp
    # with lag i as the prediction sequence;
    # use a sequence of mintemp values across
    # the full range of observed values in
    # the training data
    newdata$mintemp <- matrix(0, ncol = ncol(newdata$precip),
        nrow = nrow(newdata$precip))
    newdata$mintemp[, i] <- seq(min(data_train$mintemp),
        max(data_train$mintemp), length.out = length(newdata$year))

    # Predict on the link scale and shift by
    # the offset so that values are roughly
    # centred at zero
    preds <- predict(mod3, series = 1, newdata = newdata,
        type = "link") - offset

    # Calculate empirical prediction
    # quantiles
    probs = c(0.05, 0.2, 0.3, 0.4, 0.5, 0.6,
        0.7, 0.8, 0.95)
    cred <- sapply(1:NCOL(preds), function(n) quantile(preds[,
        n], probs = probs))

    # Plot expected function posterior
    # intervals (40-60%) and medians in
    # varying colours per lag
    pred_upper <- cred[4, ]
    pred_lower <- cred[6, ]
    pred_vals <- seq(min(data_train$mintemp),
        max(data_train$mintemp), length.out = length(newdata$year))
    polygon(c(pred_vals, rev(pred_vals)),
        c(pred_upper, rev(pred_lower)), col = scales::alpha(cols[i],
            0.6), border = scales::alpha(cols[i],
            0.7))
    lines(pred_vals, cred[5, ], col = scales::alpha(cols[i],
        0.8), lwd = 2.5)
}
abline(h = 0, lty = "dashed")
legend("topleft", legend = paste0("lag",
    seq(0, 5)), bg = "white", bty = "n",
    col = cols, lty = 1, lwd = 6)
```



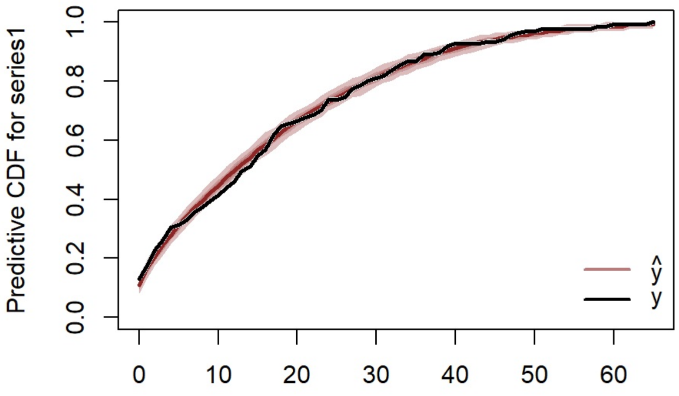
This plot demonstrates how the effect of `mintemp` is expected to change over different exposure lags, with the 3 - 5 month lags showing more of a cyclic seasonal pattern (catches expected to increase in the summer and autumn, roughly 3 - 5 months following cold minimum winter temperatures) while the recent lags (lags 0 and 1) demonstrate a more linear response function (catches broadly increasing as minimum temperature increases). This is hopefully a useful example for developing a better understanding of how a distributed lag model is attempting to recreate the data generating process. And here is the same plot for precipitation, which demonstrates how a u-shaped functional relationship diminishes toward a flat function at lags 2 - 5 (though this effect is clearly less important in the model than the mintemp * lag effect above)

```
newdata <- data_test
newdata$year <- rep(0, length(newdata$year))
newdata$season <- rep(0, length(newdata$season))
newdata$mintemp <- matrix(0, ncol = ncol(newdata$mintemp),
    nrow = nrow(newdata$mintemp))
newdata$precip <- matrix(0, ncol = ncol(newdata$precip),
    nrow = nrow(newdata$precip))
preds <- predict(mod3, series = 1, newdata = newdata,
    type = "link")
offset <- mean(preds)
plot(1, type = "n", xlab = "Precipitation",
    ylab = "Predicted response function",
    xlim = c(min(data_train$precip), max(data_train$precip)),
    ylim = c(-1.6, 1.6))

for (i in 1:6) {
    newdata$precip <- matrix(0, ncol = ncol(newdata$precip),
        nrow = nrow(newdata$precip))
    newdata$precip[, i] <- seq(min(data_train$precip),
        max(data_train$precip), length.out = length(newdata$year))
    preds <- predict(mod3, series = 1, newdata = newdata,
        type = "link") - offset
    probs = c(0.05, 0.2, 0.3, 0.4, 0.5, 0.6,
        0.7, 0.8, 0.95)
    cred <- sapply(1:NCOL(preds), function(n) quantile(preds[,
        n], probs = probs))
    pred_upper <- cred[4, ]
    pred_lower <- cred[6, ]
    pred_vals <- seq(min(data_train$precip),
        max(data_train$precip), length.out = length(newdata$year))
    polygon(c(pred_vals, rev(pred_vals)),
        c(pred_upper, rev(pred_lower)), col = scales::alpha(cols[i],
            0.6), border = scales::alpha(cols[i],
            0.7))
    lines(pred_vals, cred[5, ], col = scales::alpha(cols[i],
        0.8), lwd = 2.5)
}
abline(h = 0, lty = "dashed")
legend("topleft", legend = paste0("lag",
    seq(0, 5)), bg = "white", bty = "n",
    col = cols, lty = 1, lwd = 6)
```



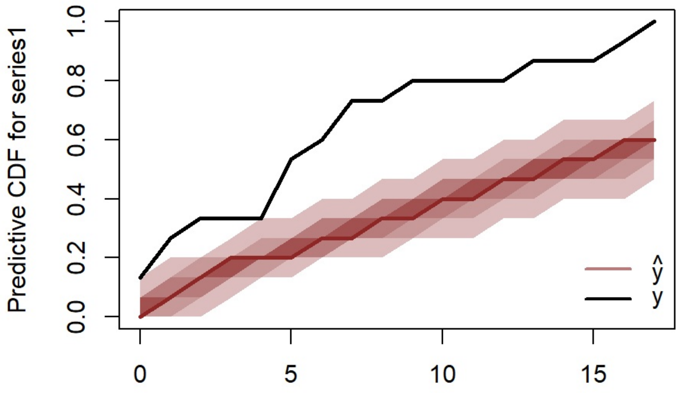
All of the usual functions in `mvgam` can also be used for list data objects, so long as they contain the necessary fields `series` , `season` and `year` . For example, posterior retrodictive checks for the in-sample training period:

```
ppc(mod3, series = 1, type = "cdf")
```



and predictive checks for the out of sample forecast period (which demonstrates how the model tends to overpredict for the forecast period in this particular example):

```
ppc(mod3, data_test = data_test, series = 1,
    type = "cdf")
```



Logical next steps for interrogating this model would be to trial different trend types (i.e. random walk), replace the distributed lag function for `precip` with a standard smooth function (that does not include lag interactions, as clearly the model above indicates that these are not supported) and inspect whether different covariates (such as `ndvi` or `maxtemp`) might play a role in modulating catches of `PP`. Finally, once we are satisfied that we have a well-performing model that we can understand and interrogate, we could expand up to a multivariate model by including other species as response variables. This would allow us to capture any possible unobserved dependencies in the catches of multiple co-occurring species in a single unified modelling framework