

# BMA in FEWS

## BMA in FEWS.

The present version of BMA within FEWS uses an R Package for Probabilistic Forecasting using Ensembles and Bayesian Model Averaging - "Ensemble BMA Package". The package was developed by Chris Fraley, Adrian E. Raftery, J. McLean Sloughter and Tilmann Gneiting at the University of Washington. This package is distributed under [General Public License \(version >= 2\)](#) .

R Package can be downloaded from CRAN R project website or directly from [R package](#) . However it is recommended that you download the version of ensembleBMA from the link given below instead since the version 3.0-3 is tested and used within FEWS. The new version which is available from the above mentioned link may or may not work flawlessly within FEWS.

## Ensemble BMA Package Documentation.

For Ensemble BMA R package documentation, please refer to the online documentation on Cran R project website through this link [Ensemble BMA documentation](#)

## Versions

FEWS uses R version 2.7.0.

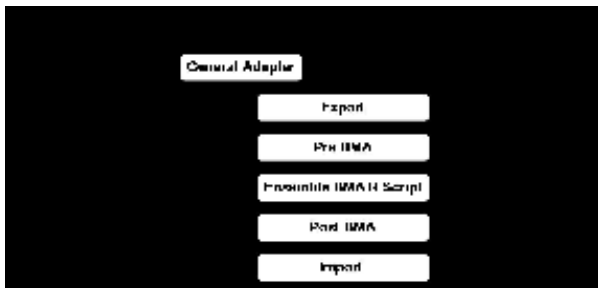
Package: ensembleBMA, Version: 3.0-3, Date: 2008-07-21

Supporting Package chron, Version: 2.3-24, Date: 2008-07-18

Please Note: The ensemble BMA is an older version that is no longer supported by the original developers

To run BMA in FEWS, firstly install the correct version of R on the computer where BMA Module is running .  
Copy the contents of Ensemble BMA [ensembleBMA.zip](#) and Chron [chron.zip](#) under library directory of R Package.  
Please Note: Use only the package versions as mentioned above for running BMA Module in Delft-FEWS.

## Systematic Diagram



## Preprocessor

BMA Module preprocessor prepares the input for ensemble BMA R Package. Ensemble BMA R Package uses input as CSV format. The General Adapter configuration of Preprocessor is shown as below.

```
<executeActivity>
  <command>
    <className>nl.wldelft.fews.adapter.bma.BmaPreAdapter</className>
  </command>
  <arguments>
    <argument>%ROOT_DIR%</argument> <!-- root directory -->
    <argument>piOutputTimeSeries/bmainputL0.csv</argument> <!-- outputfile -->
    <argument>%TIME0%</argument> <!-- Time0 -->
    <argument>0</argument> <!-- Start of Lead time period in days -->
    <argument>parameters.txt</argument> <!-- Parameter file - each column represents a row -->
    <argument>piOutputTimeSeries/forecast0.csv</argument> <!-- Number of (partly) complete
Forecasts used for calculating the training period -->
  </arguments>
  <timeOut>4000000</timeOut>
</executeActivity>
```

The above configuration has to be repeated for different lead time. Make sure that name of output file accordingly changed.

## BMA Module

BMA Module is a script written under R package which uses the ensembleBMA package written for R as briefly described above. The General Adapter configuration for running BMA Module is shown as below.

```
<executeActivity>
  <command>
    <executable>$R_EXE$</executable>
  </command>
  <arguments>
    <argument>--vanilla</argument>
    <argument>%ROOT_DIR%/config/BMA_FEWS_Script.R</argument>
    <argument>%ROOT_DIR%/piOutputTimeSeries/bmainputL0.csv</argument> <!-- inputfile -->
    <argument>%ROOT_DIR%/piOutputTimeSeries/bmaoutputL0.qan</argument> <!-- outputfile qauntile -->
    <argument>%ROOT_DIR%/piOutputTimeSeries/bmaoutputW0.wie</argument> <!-- outputfile weights-->
    <argument>%ROOT_DIR%/piOutputTimeSeries/bmaoutputB0.bia</argument> <!-- outputfile bias-->
    <argument>0</argument> <!-- input lead time in days (not used) -->
    <argument>%ROOT_DIR%/piOutputTimeSeries/forecast0.csv</argument> <!-- inputfile Forecast Length -->
  -->
  </arguments>
  <timeOut>1000000000</timeOut>
  <overrulingDiagnosticFile>%ROOT_DIR%/rscript.log</overrulingDiagnosticFile>
</executeActivity>
```

The above configuration has to be repeated for different lead time. Make sure that name of input and output files are accordingly changed.

Please note: \$R\_EXE\$ is attribute which is defined in global.properties file as "R\_EXE=C:/Program Files/R/R-2.7.0/bin/Rscript.exe"

## BMA Module R Script

A typical BMA Module R Script can be downloaded from [here](#).

The contents of BMA Module R script is briefly described as below.

```
--- Read Arguments
--- Check if files exists
--- Read Forecast Length file
--- Load Ensemble R
--- Read input data

-- Assign labels (hard coded - similar to parameter file) (R-Code - Make sure to update this line for your
model)
labels <-c("SBK_MaxLob_DWD_GME_Q.fs", "SBK_MaxLob_DWD_LM_Q.fs", ..... )

--- Perform ensembleBMA analysis (R-Code)
enRData<-ensembleData(forecasts=rdata[,labels], dates=rdata$TIME, observations=rdata$OBS)
trainingrule=list(length=forecastlen,lag=1)
rDataBMA <- ensembleBMA(enRData,model="normal",trainingRule=trainingrule, control = controlBMAnormal
(maxIter=20))

--- output Quantile to File
--- output Wiegths to File
--- output Bias to File
```

Please make sure that the line "labels <-c("SBK\_MaxLob\_DWD\_GME\_Q.fs", "SBK\_MaxLob\_DWD\_LM\_Q.fs", .....)" is changed according to the number of models used.

Output of each BMA Module run are 3 files, with extension ...

\*.wei -> weights

\*.bia -> bias

\*.qan -> quantiles - value for the next first forecast - (used only for checking)

### Format of Weights file (\*.wei)

forecast-date, weight for model one, weight for model 2 , .... and so on .... , sigma

### Format of Bias file (\*.bia)

B value for model 1, B value for model 2, ... and so on ...

A value for model 1, A value for model 2, ... and so on ...

## Postprocessor

BMA Module postprocessor reads the output of R model prepares the data which is to be later imported into FEWS database. The General Adapter configuration of Postprocessor is shown as below.

```
<executeActivity>
  <command>
    <className>nl.wldelft.fews.adapter.bma.BmaPostAdapter</className>
  </command>
  <arguments>
    <argument>%ROOT_DIR%</argument> <!-- root directory -->
    <argument>piOutputTimeSeries</argument> <!-- outputDirectory -->
    <argument>%TIME0%</argument> <!-- Time0 -->
    <argument>3</argument> <!-- max lead time in days -->
    <argument>parameters.txt</argument> <!-- Parameter file - each column represents a row -->
  </arguments>
  <timeOut>4000000</timeOut>
  <overrulingDiagnosticFile>%ROOT_DIR%/piDiagnostic.xml</overrulingDiagnosticFile>
</executeActivity>
```

The postprocessor uses the output of BMA Module run (i.e. quantiles, weights and bias) and the input to generate new forecasted timeseries + quantiles (10 , 25, 75 and 90) timeseries.

## Generating Forecast Timeseries

The forecasted timeseries are generated using the weights, sigma and bias correction.

```
FOR EACH models_i  (skip missing forecasts)

  BMA += weight_i * (bias_a_i * forecast_i + bias_b_i)
  sumweights += weight_i

END
BMA = BMA / sumweights

Quantile_10 = BMA - 0.842 * sigma
Quantile_25 = BMA - 0.675 * sigma
Quantile_50 = BMA
Quantile_75 = BMA + 0.675 * sigma
Quantile_90 = BMA + 0.842 * sigma
```