Additional Material

5) Pg. 3827, last paragraph: why are the results in García and Giraldo (2011) "numerically quite different"? Please explain.

The following Figures represent the mean μ (Fig. 1), and the standard deviation σ (Fig. 2), of AMDR PDF, from the referenced work when R= R_B (García and Giraldo, 2011), and from the present paper (where $R=R_{B*}R_D$). Therefore, from these Figures, is possible to identify the spatial differences between the two works.
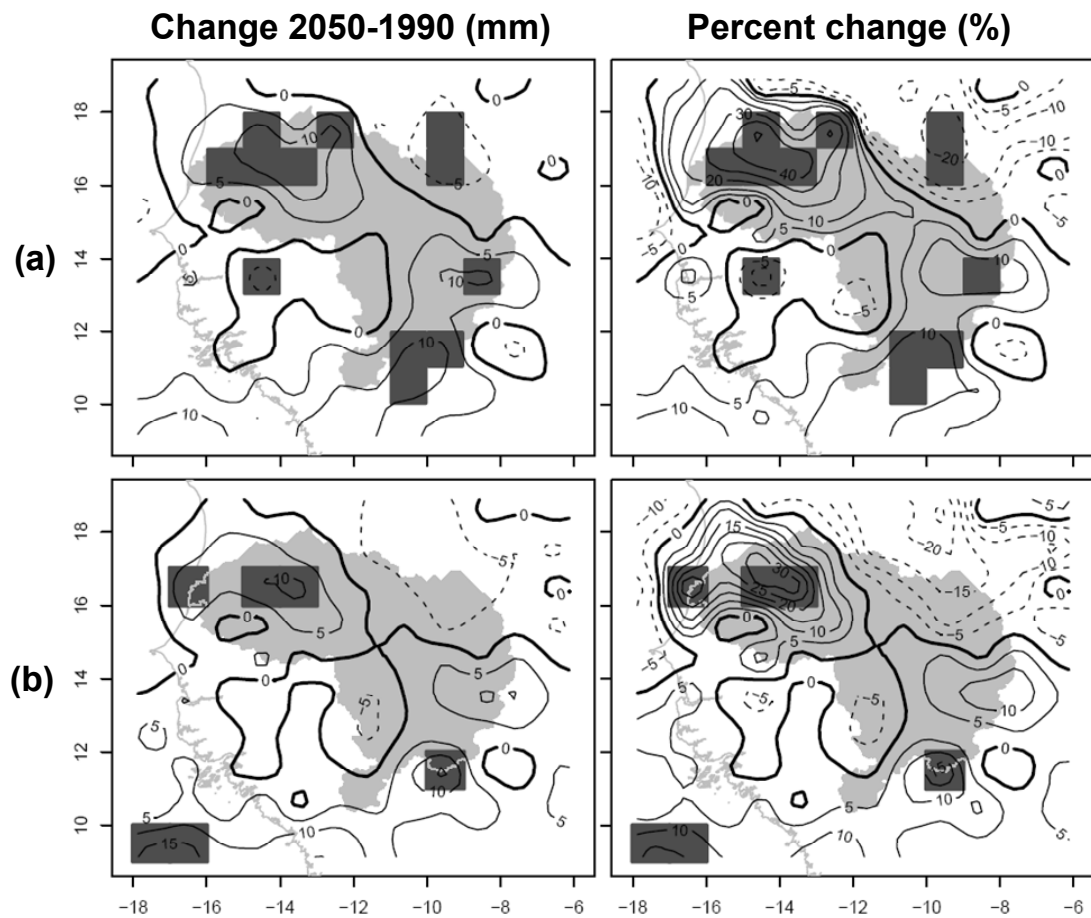
Fig. 1. Predicted changes of μ of AMDR PDF: (a) $R=R_B$ y (b) R) $R=R_{B*}R_D$. The dark-grey squares show the sites with statistically significant changes. The dashed lines represent negative changes.
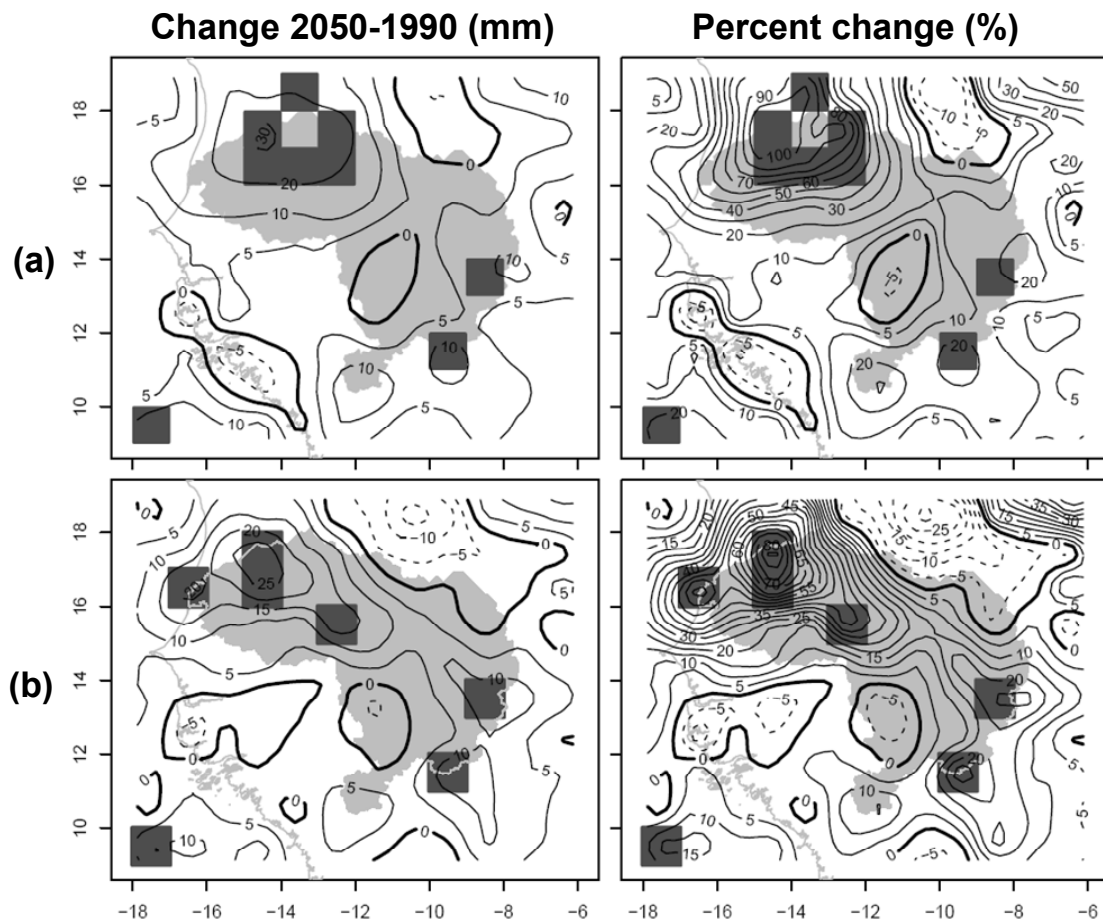
Fig. 2. The same as Fig. 1, but for σ of AMDR PDF.

6) Section 3.2: the authors use non-stationary statistical model to describe AMDR time series. However, have they first checked that there is statistical evidence against the use of a stationary model?

The GAMLSS procedure was applied using two-parameter distributions, in order to guarantee the parsimony of fitted models. Due to the huge volume of data, the fit procedure was automated. GAMLSS tools allow the validation of stationarity hypothesis of distribution parameters.

An example is presented for site 16 (Fig. 3(a)), for the GKSS/CLM model considering calls to GAMLSS function (R software). It is noted that, in addition, the summary of adjustment gives the p-values associated with the linear fit (constant and linear term).

The results in this particular case, for a significance of 0.05, are: - there is no trends of μ parameter (-0.01264, p-value=0.01963), while - the parameter σ presents a decreasing trend (-0.006798 , p-value=0.05695)

```
mod2 = gamlss(
        datos$value~cs( datos$date , df=dfmu[bestmodel] , c.spar=c(-1.5,2.0) ) ,
```

```
                        sigma.fo=~cs( datos$date , df=dfsigma[bestmodel] ) ,
                    data=datos , family=modelnames[bestmodel] , trace=F , control=control
               )
```

Where an adjustment with smoothing functions using family=WEI (considering the application of SBC), was invoked. The summary of the results for this line is the following:

```
********************************************************************
> summary(mod2)

        The following object(s) are masked _by_ .GlobalEnv :

         date value

********************************************************************
Family:  c("WEI", "Weibull")

Call:  gamlss(formula = datos$value ~ cs(datos$date, df = dfmu[bestmodel],
    c.spar = c(-1.5, 2)), sigma.formula = ~cs(datos$date, df = dfsigma[bestmodel]),
family = modelnames[bestmodel], data = datos, control = control,
    trace = F)

Fitting method: RS()

-------------------------------------------------------------------
Mu link function:  log
Mu Coefficients:
                                                Estimate  Std. Error  t value
Pr(>|t|)
(Intercept)                                     28.29777  10.629037    2.662
0.00940
cs(datos$date, df = dfmu[bestmodel], c.spar = c(-1.5, 2))  -0.01264    0.005306   -2.382
0.01963

-------------------------------------------------------------------
Sigma link function:  log
Sigma Coefficients:
                                  Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)                       13.562944   7.073711    1.917   0.05881
cs(datos$date, df = dfsigma[bestmodel])  -0.006798   0.003519   -1.932   0.05695

-------------------------------------------------------------------
No. of observations in the fit:  81
Degrees of Freedom for the fit:  4.000016
     Residual Deg. of Freedom:  76.99998
                    at cycle:  2

Global Deviance:     635.7243
           AIC:     643.7243
           SBC:     653.3022
```

```
*******************************************************************
```

If the previous results are compared with the results of the call for a lineal adjustment to parameters μ y σ, in the following form:

```
mod3 = gamlss(
            datos$value~datos$date ,
            sigma.fo=~datos$date ,
            data=datos , family=modelnames[bestmodel] , trace=F , control=control
          )
```

the result should be:

```
*******************************************************************
> summary(mod3)

        The following object(s) are masked _by_ .GlobalEnv :

         date value

*******************************************************************
Family:  c("WEI", "Weibull")

Call:   gamlss(formula = datos$value ~ datos$date, sigma.formula = ~datos$date,
family = modelnames[bestmodel], data = datos, control = control,
    trace = F)

Fitting method: RS()

-----------------------------------------------------------------
Mu link function:  log
Mu Coefficients:
            Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  28.29775  10.629039    2.662   0.00940
datos$date   -0.01264   0.005306   -2.382   0.01963

-----------------------------------------------------------------
Sigma link function:  log
Sigma Coefficients:
            Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  13.562949   7.073711    1.917   0.05881
datos$date   -0.006799   0.003519   -1.932   0.05695

-----------------------------------------------------------------
No. of observations in the fit:  81
Degrees of Freedom for the fit:  4
      Residual Deg. of Freedom:  77
                      at cycle:  2
```

```
Global Deviance:      635.7243
          AIC:        643.7243
          SBC:        653.3021
********************************************************************
```

In conclusion, the results (p-value) from lineal fitting and the corresponding from smoothing functions, are the same.

8) A much more extensive discussion of the statistical modeling should be provided. After reading the manuscript, I still have several questions: 1) how did the authors select their final distribution? 2) What distribution was generally chosen?

1) In order to execute the GAMLSS fit procedure, it should be taking into account that `data$value` is a vector with yearly AMDR values, and `data$date` is a vector of years. The most important part of the algorithm, which do the selection and optimization of statistical model, is detailed below. The algorithm begins by setting up `control` through the `gamlss.control` function. The `gamlss.control` function has the default setting in `gamlss` function, which control the parameters of the outer iterations algorithm (Stasinopoulos *et al.*, 2008).

After establishing `control`, `find.hyper` function finds the hyperparameters (degrees of freedom for smoothing terms) by minimizing the profile GAIC based on the global deviance. The procedure is executed for each statistical model (LN, GA, WEI and GU), beginning with hyperparameters values between [0.001,1] to encourage a quick convergence. The right side value of the hyperparameter interval can be increase, if it is necessary.

The statistical model is fitted using `gamlss` function and the optimal values of hyperparameters found with `find.hyper` function. The values Akaike information criterion (`mod2$aic`), the Schwartz Bayesian criterion (`mod2$sbc`) and the hyperparameters values (`mod2$mu.df-2` and `mod2$sigma.df-2`) are stored in several vectors. Finally, the best model is selected, which corresponds with the minimal value in the SBC vector. Again, the `gamlss` function is used to compute the GAMLSS fit with the best model and its optimal values of hyperparameters.


```
********************************************************************
modelnames = c( "LN" , "GA" , "WEI" , "GU" )

control = gamlss.control( c.crit=0.1 , n.cyc=20,
                  mu.step=0.1 , sigma.step=0.1 , nu.step=0.1 , tau.step=0.1 ,
                  gd.tol=5 , iter=0 , trace=F , autostep=TRUE , save=TRUE )

for ( m in 1:length(modelnames) ) {
```

```r
        mod1 = quote(
                gamlss(
                        datos$value~cs( datos$date , df=p[1] , c.spar=c(-1.5,2.0) )
                        sigma.fo=~cs( datos$date , df=p[2] , c.spar=c(-1.5,2.0) ) ,
                        data=datos , family=modelnames[m] , trace=FALSE , control=control
                        )
                )

        upperlimit=c(1,1)

        repeat {
                op1 = find.hyper( model=mod1 ,
                                parameters=c(0.001,0.001) ,
                                lower=c(0.001,0.001) ,
                                upper=c(upperlimit[1],upperlimit[2]) ,
                                steps=c(1e-06, 1e-06) , penalty=3.5 ,
                                trace=FALSE )

                                if(     (op1$par[1]+0.3)<upperlimit[1] &
                                        (op1$par[2]+0.3)<upperlimit[2] ||
                                        sum(upperlimit)>=8 ) {
                                                break
                                }
                                if((op1$par[1]+0.3)>=upperlimit[1]) {
                                        upperlimit[1]=upperlimit[1]+1
                                }
                                if((op1$par[2]+0.3)>=upperlimit[2]) {
                                        upperlimit[2]=upperlimit[2]+1
                                }

        }

        mod2 = gamlss(
                datos$value~cs( datos$date , df=op1$par[1] , c.spar=c(-1.5,2.0) ) ,
                sigma.fo=~cs( datos$date , df=op1$par[2] ) ,
                data=datos , family=modelnames[m] , trace=F , control=control
                )

        ##### Estimadores de desempeno del modelo
        modelaic[m] = mod2$aic
        modelsbc[m] = mod2$sbc
        dfmu[m] = round( mod2$mu.df-2 , 2 )
        dfsigma[m] = round( mod2$sigma.df-2 , 2 )

} ##### for ( m in 1:length(modelnames) ) #####


##### Selecciono el mejor modelo, de acuerdo con el criterio sbc ()
bestmodel = which( modelsbc==min(modelsbc) )

mod2 = gamlss(
```

```
            datos$value~cs( datos$date , df=dfmu[bestmodel] , c.spar=c(-1.5,2.0) ) ,
            sigma.fo=~cs( datos$date , df=dfsigma[bestmodel] ) ,
            data=datos , family=modelnames[bestmodel] , trace=F , control=control
        )
    ****************************************************************
```

2) The Fig. 3 represents the spatial pattern of "best fit" distribution, with different colours. It could be concluded, the LN was generally chosen for models METO-HC/HAD and INM/RCA. However, there are some clusters, especially for KNMI/RACMO and SMHI/RCA.
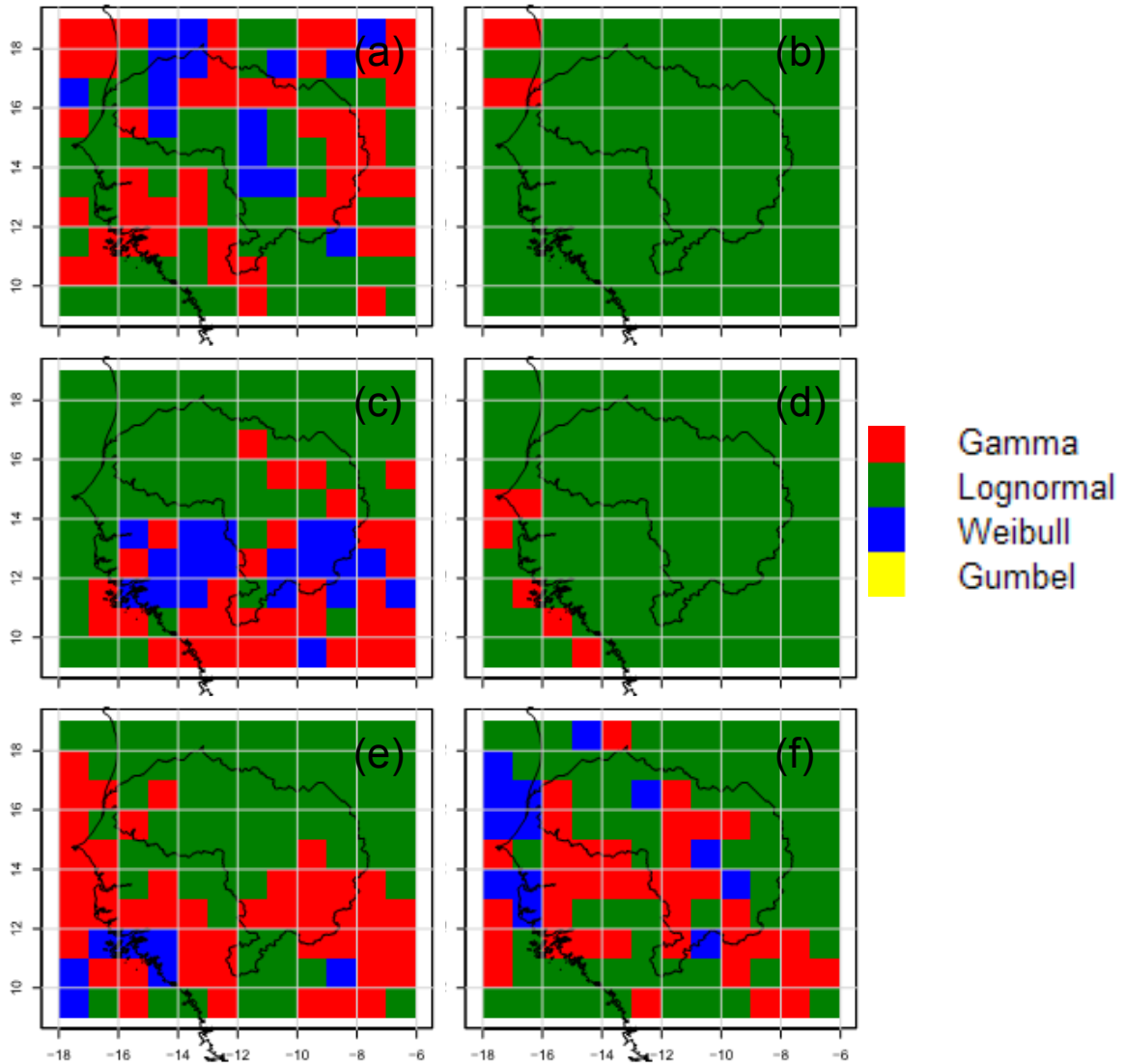
Fig. 3. "Best fit" maps to a particular distribution, according with the GAMLSS analysis. The RCMs considered are: (a) GKSS/CLM, (b) METO-HC/HAD, (c) KNMI-RACMO, (d) INM-RCA, (e) SMHI/RCA, and (f) MPI-M-REMO.

The following Table with the total number of sites considered for each distribution, is presented.

Table. Number of sites with "best fit" for each GAMLSS model and RCM considered.

| PDF	Gamma	Lognormal	Weibull	Gumbel
GKSS/CLM	48	57	15	0
METO-HC/HAD	4	116	0	0
KNMI/RACMO	33	68	19	0
INM/RCA	6	114	0	0
SMHI/RCA	45	68	7	0
MPI-M/REMO	37	70	13	0

9) Figure 3: based on this figure, I find it a bit hard to believe that the statistical models are really able to describe the data. I would be interested in seeing the corresponding residual plots used to assess the goodness of fit. I am not sure that the statement on pg. 3828, line 6 ("the good fit of the GAMLSS statistical model to simulated AMDR time") is really supported by the results presented.

10) Figure 3: please add the corresponding time series from the data. It would be interesting to see whether the models can actually reproduce the patterns in the data.
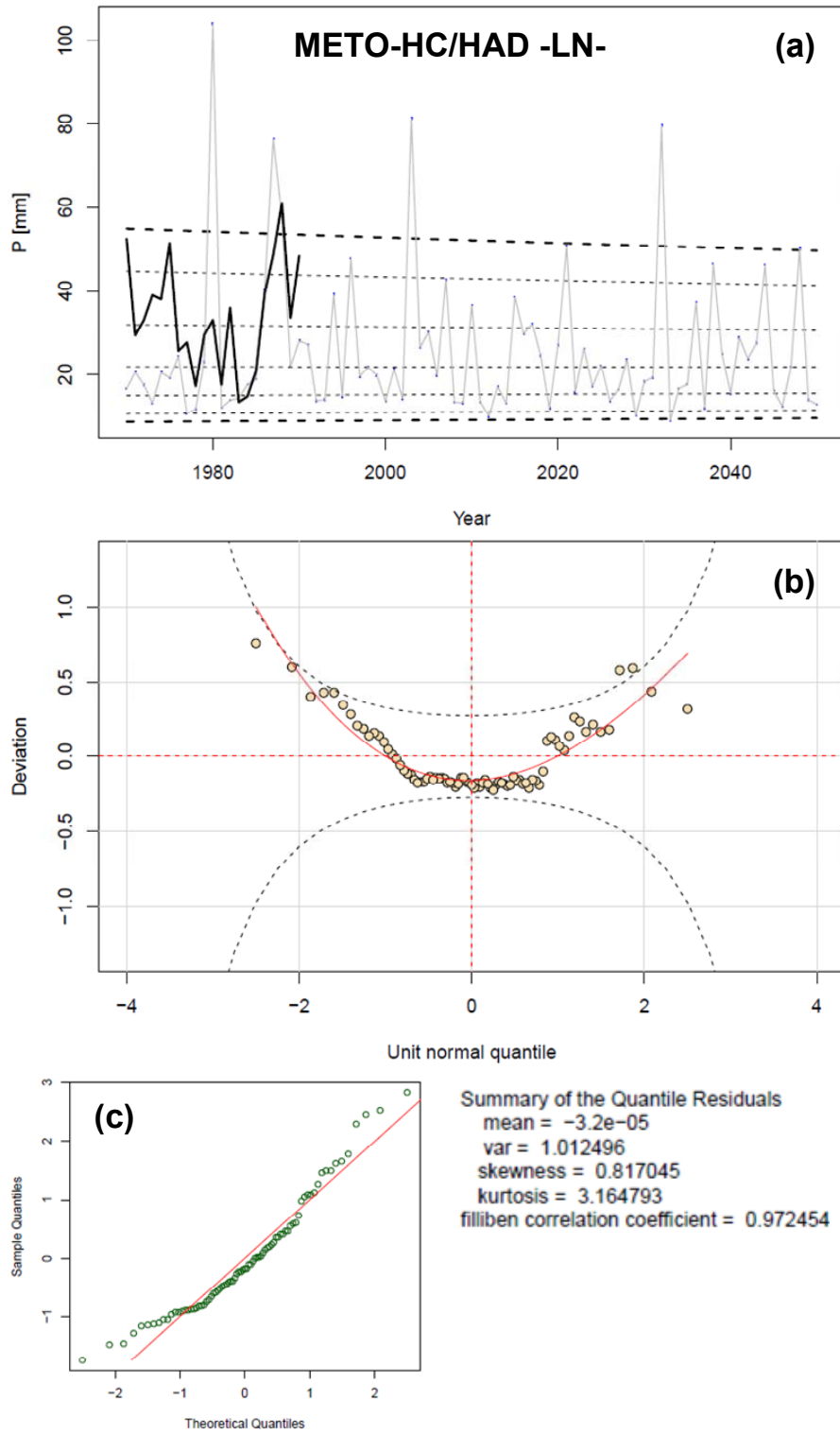
11) Pg. 3825, lines 20-24: the authors mention how they checked the goodness of the fit but don't state whether the residuals supported their choice of models.

Taking into account each RCM, the time series together with the IRD data, worm-plot, qq-plot, and a summary of quantile residuals, are presented in the following figures (Fig. 4 to Fig. 9).
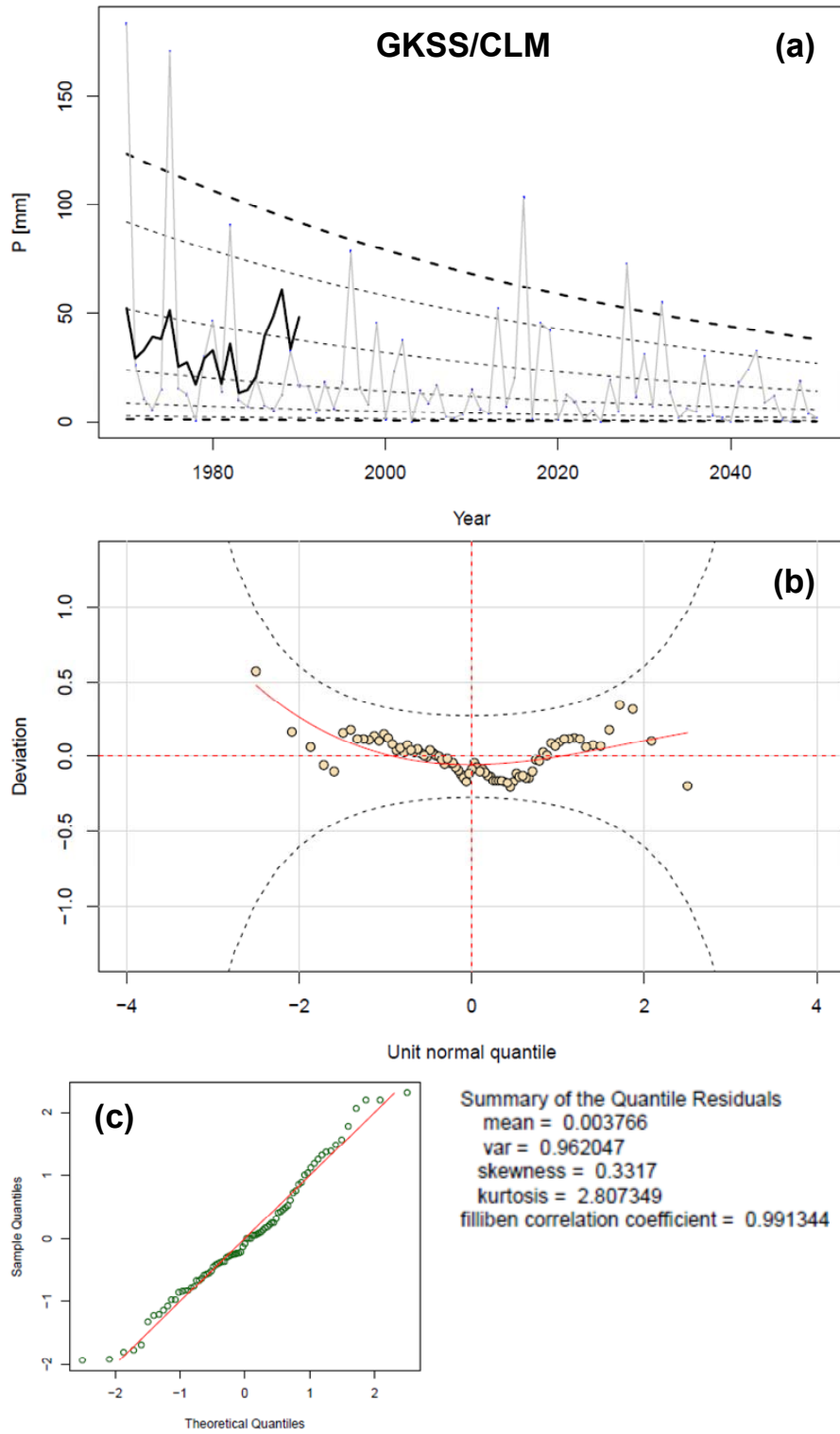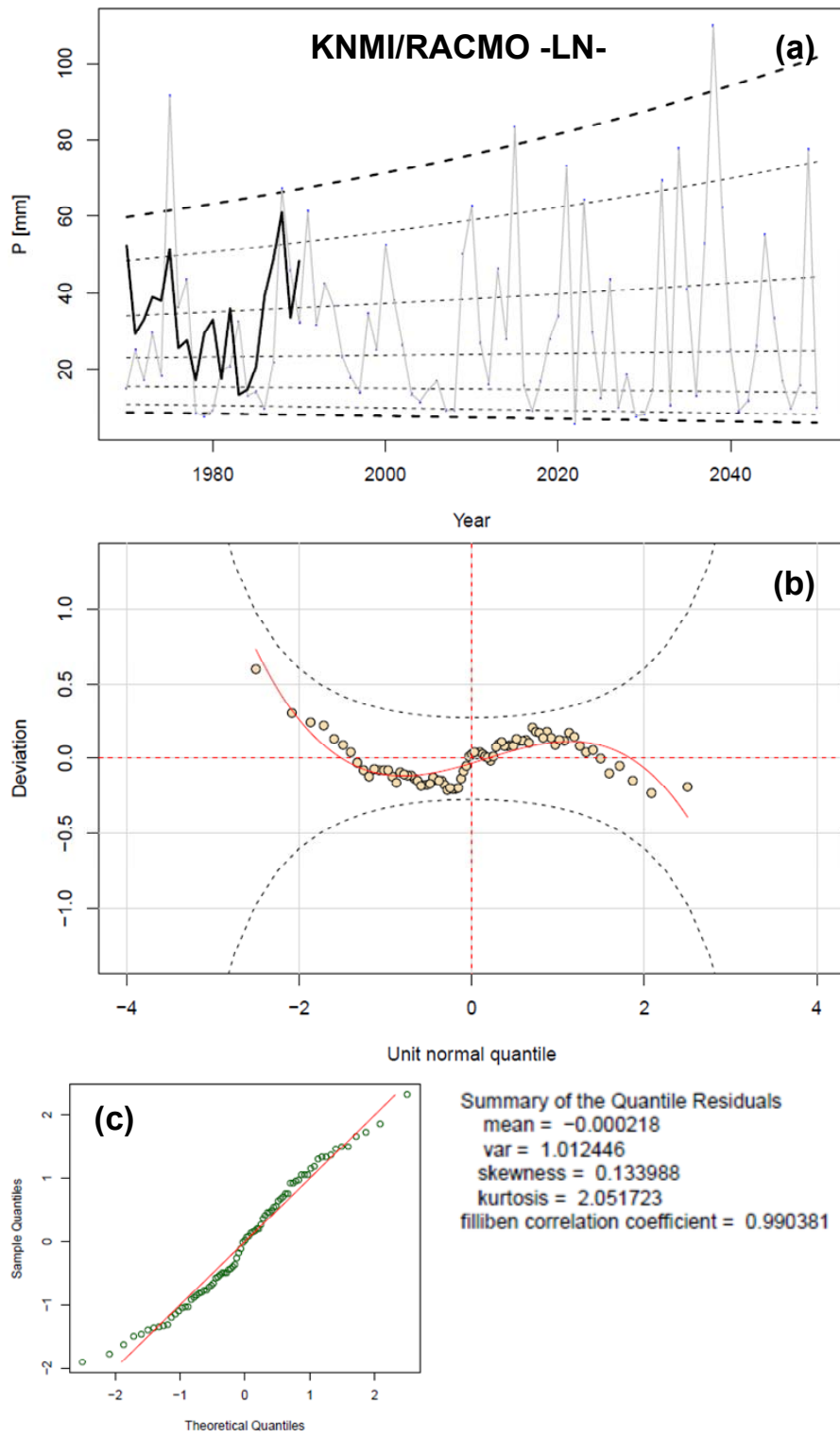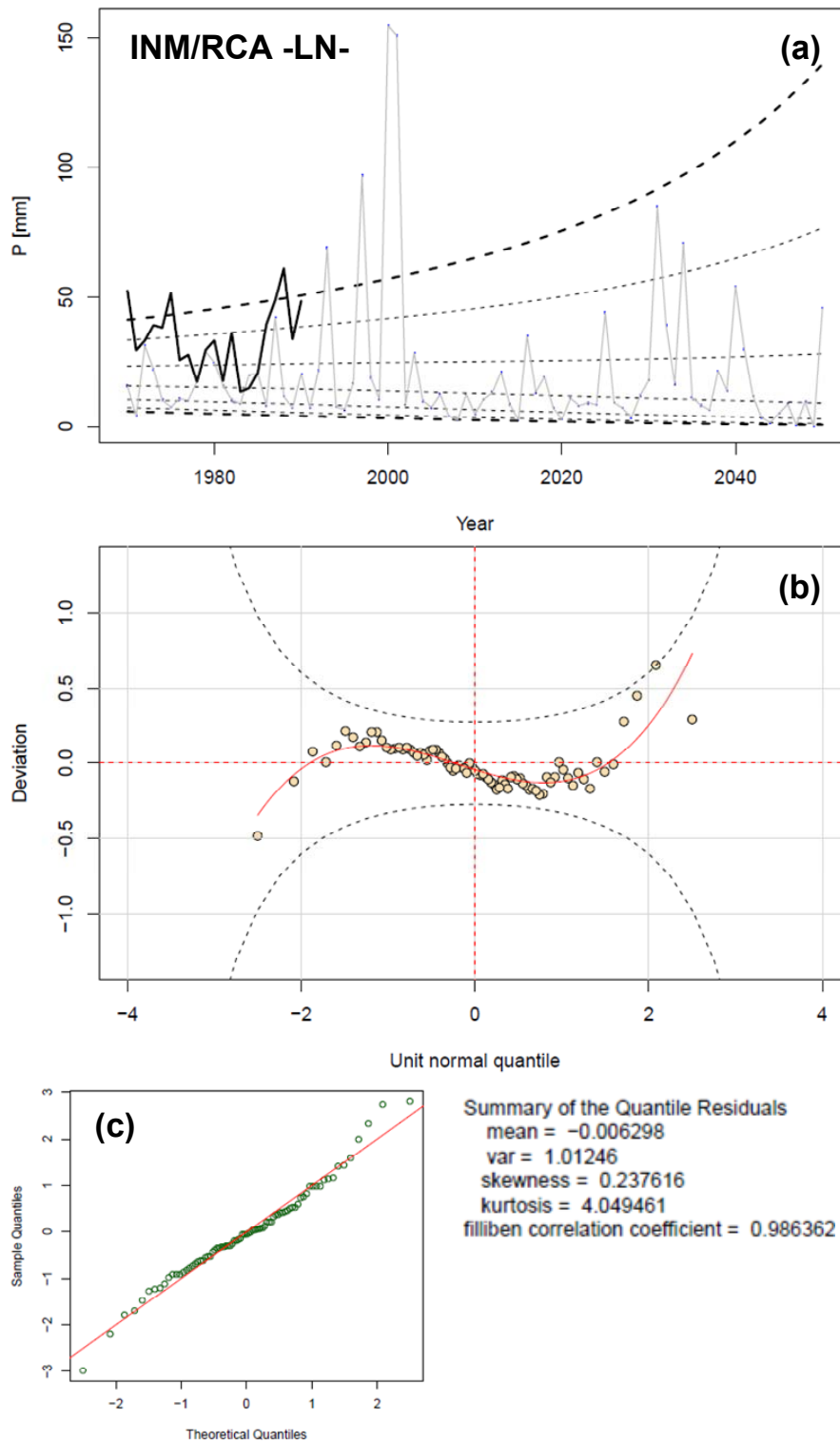
Fig. 4. GAMLSS analysis on site 16 for GKSS/CLM RCM. (a) Time series of AMDR from RCM (grey) and IRD data (black), together with the centiles curves (5, 10, 25, 50, 75, 90 and 95%) in dashed lines; (b) worm plot and (c) qq plot with residual statistics. The hypothesis of normality is rejected if the Filliben coefficient is less than 0.984.
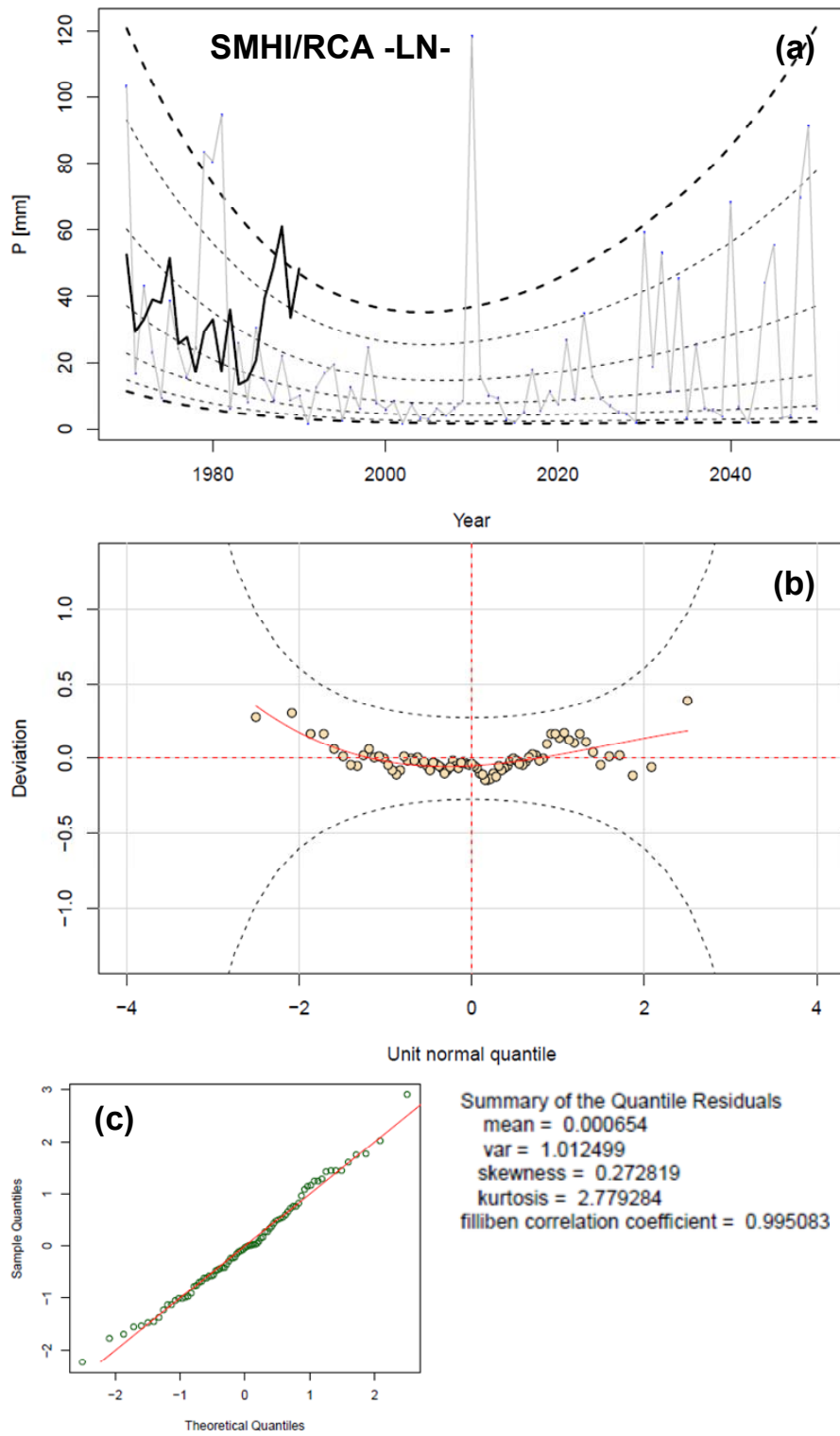
Fig. 5.   The same as Fig. 4, but METO-HC/HAD RCM.

KNMI/RACMO -LN- (a)

(b)

(c)

Summary of the Quantile Residuals
    mean = −0.000218
    var = 1.012446
    skewness = 0.133988
    kurtosis = 2.051723
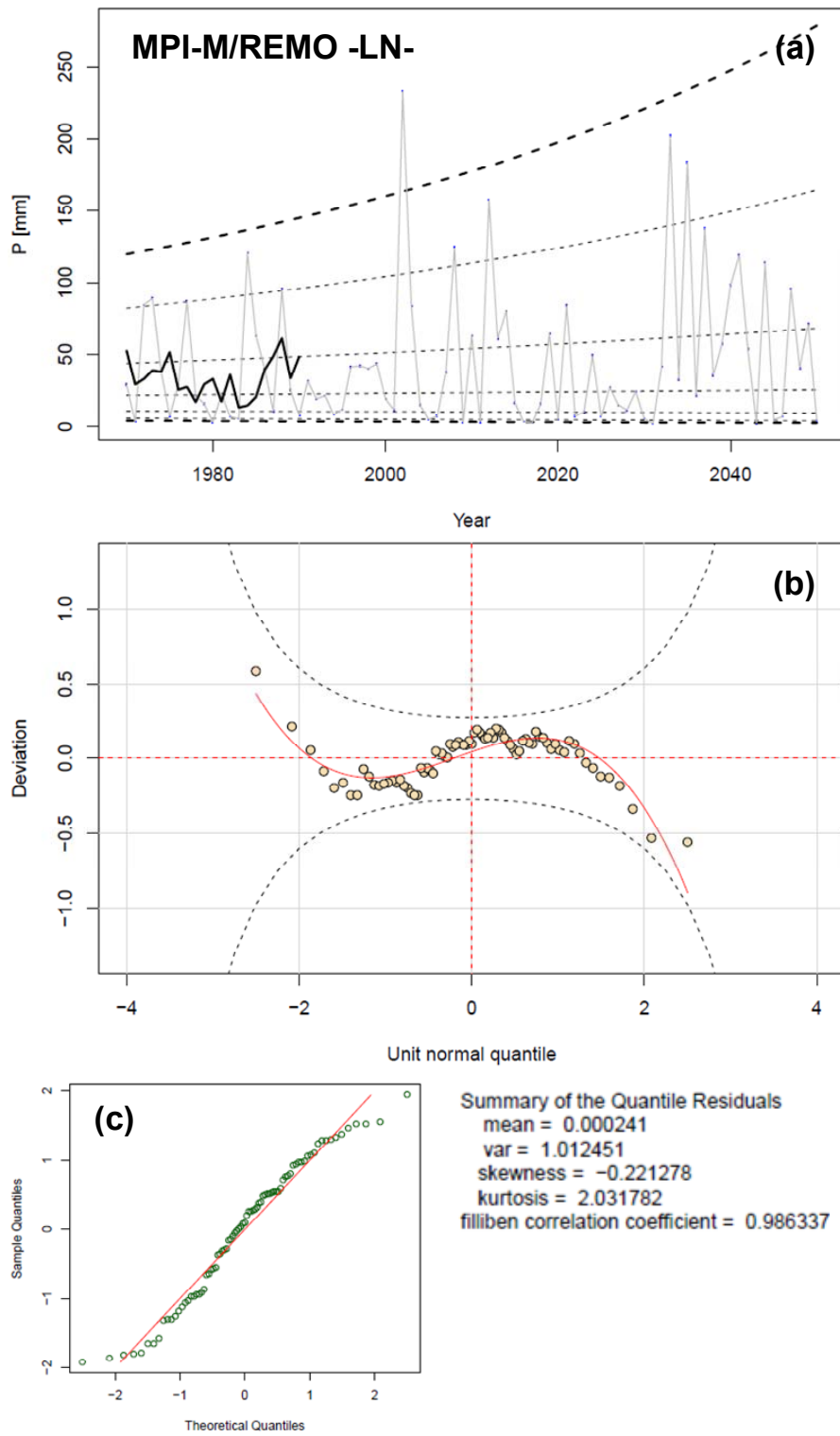   filliben correlation coefficient = 0.990381

Fig. 6. The same as Fig. 4, but KNMI/RACMO RCM.

Fig. 7. The same as Fig. 4, but INM/RCA RCM.

Fig. 8.   The same as Fig. 4, but SMHI/RCA RCM.

Fig. 9. The same as Fig. 4, but MPI-M/REMO.