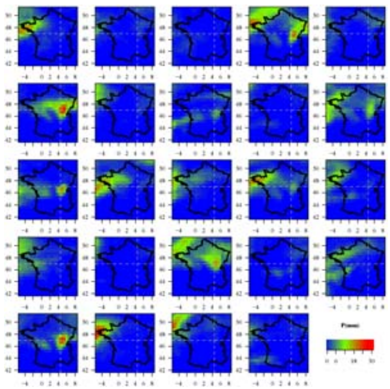
Impact of sample size on forecast verification scores and post processing parameters of hydrological ensemble predictions

Ioanna Zalachori, Annie Randrianasolo, Maria-Helena Ramos

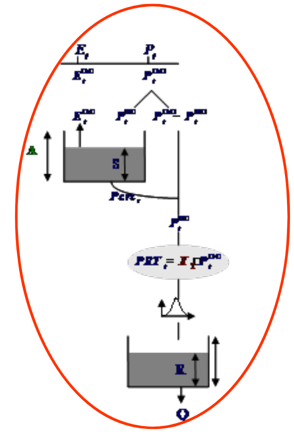
HEPEX workshop: UNESCO-IHE, The Netherlands
Delft 7-9 June 2011

Ensemble Prediction Chain

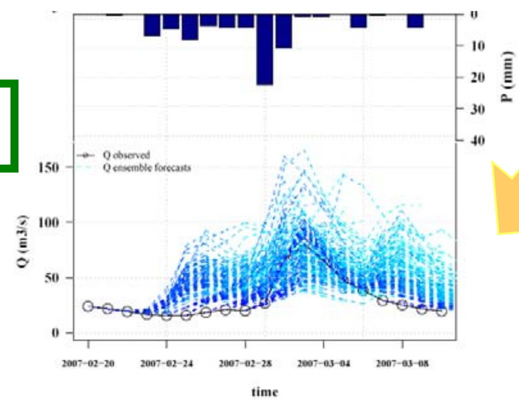


Meteorological Ensemble Forecasts

Hydrological Model



Hydrological Ensemble Forecasts



Evaluation

Statistical Processing

NO

GOOD?

YES

Hydrological Applications

Sample size limitations

«The availability of a large number of past forecast-observation pairs <…> is a major factor of the success of the calibration technique <…>. »

Hagedorn *et al.*, 2010

*** but, in practice, still « limited » (homogeneous) data series available :**

1) Case study for the floods in upper Mulde river basin (Dietrich *et al.*,2009)

« The time period covered by regular operation of the ensemble forecast systems is small (COSMO-LEPS operational from 2005, SRNWP-PEPS operational from 2004, COSMO-DE operational from 2007).»

2) Case study for the sub-catchment of the Alzette (Hostache *et al.*, 2011)

« Between October 9, 2006 and February 11, 2009, the half degree medium range rainfall forecasts provided by the Global Forecasting System(GFS) atmospheric model (Kalnay et al., 1971; Kanamitsu et al., 1991) <…> are provided by the National Oceanic and Atmospheric Administration (NOAA) through a web page every 6 h, as 3 h cumulated rainfall forecasts for the next 7 days.»

3), 4), etc….

Our Aim

- *Data series available are limited, ok, but are they sufficient?*
- *Is there a « minimum required length »* for our applications*?*

**Aims of this study:*

evaluate the impact of sample size on forecast verification scores and post processing parameters of hydrological ensemble predictions

=> data & methods

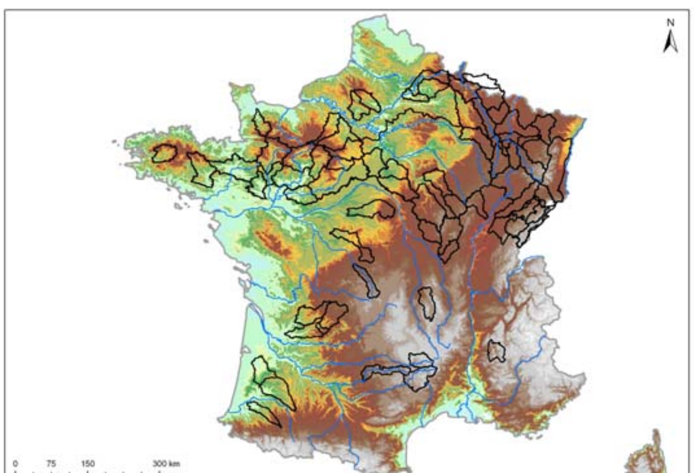
=> results

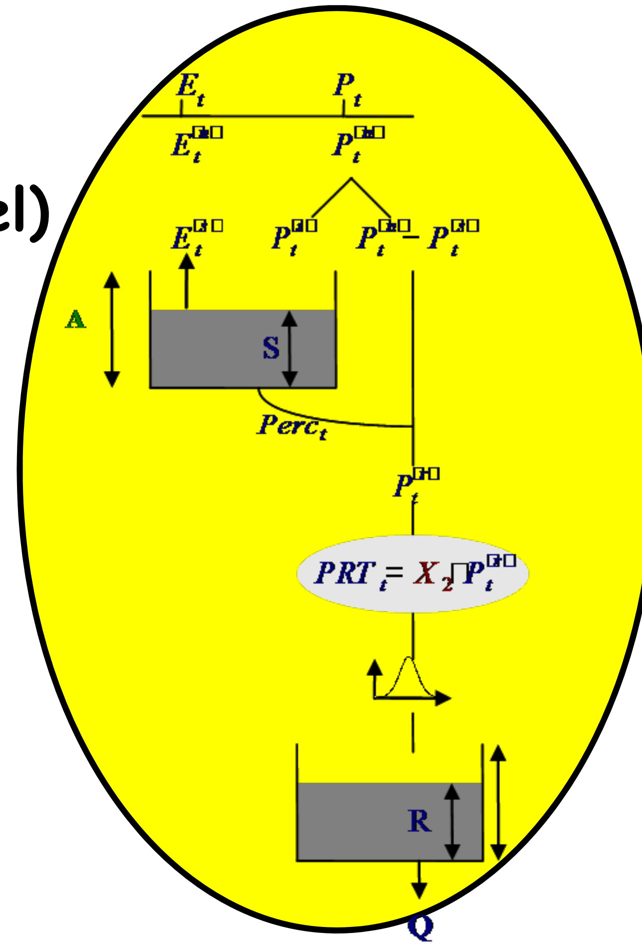
=> discussion

Set of data

- **74 Catchments in France**
- **A lumped hydrological model (GRP Model)**
- **Daily Meteorological Forecasts: ECMWF**

from 03/2005 to 10/2008 => 42 mois





Methodology

- Moving window of different sizes

Length of period (months)	1	3	6	12	18	24	30	36	40	42
Number of periods	42	40	37	31	25	19	13	7	3	1

Evaluation

Statistical Processing

Impact over typical scores (raw forecasts)

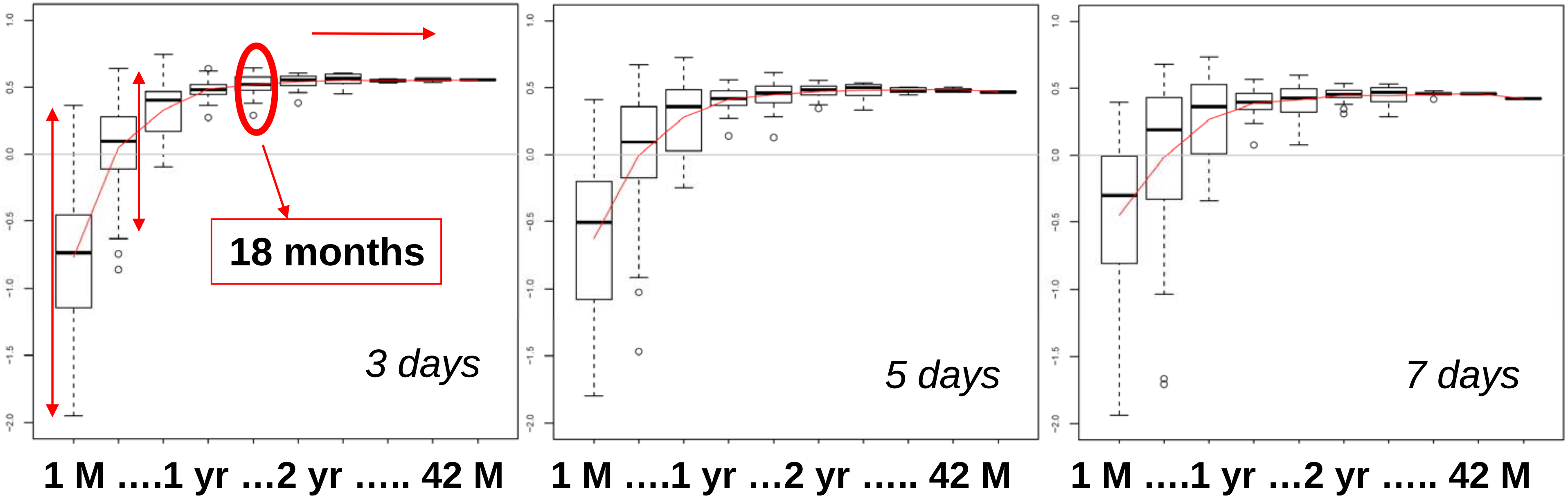
Impact over the parameters of the Best Member Method

Results : RPSS

$$RPSS = 1 - \frac{RPS}{RPS_{clim}}$$

* quantiles used to define the different bins calculated **for each period of verification**

Ex: Catchment Yonne à Gurgy

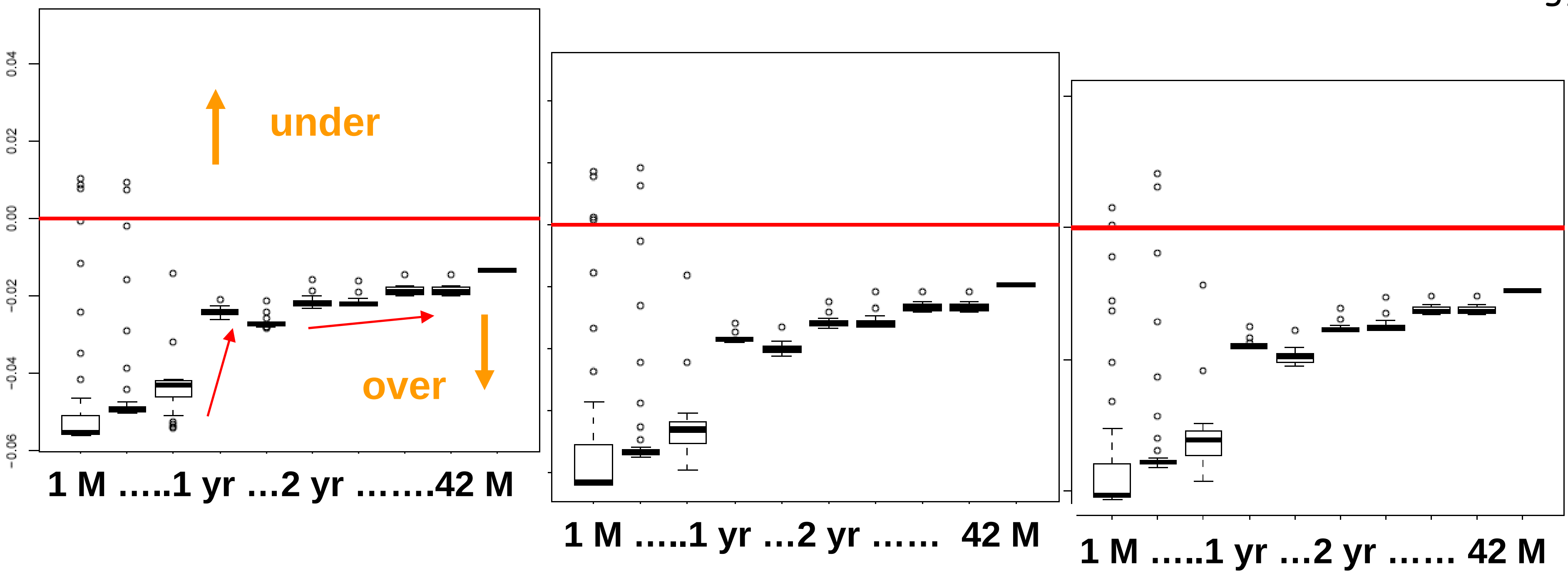


=> Reaches a sill after *12-18 months*
=> Similar behavior for all LT

Rank Divergence Score

$$DS(o_t \| f_t) = \sum_{i=1}^{n} [o_t] \log\left(\frac{[o_t]}{[f_t]}\right)$$

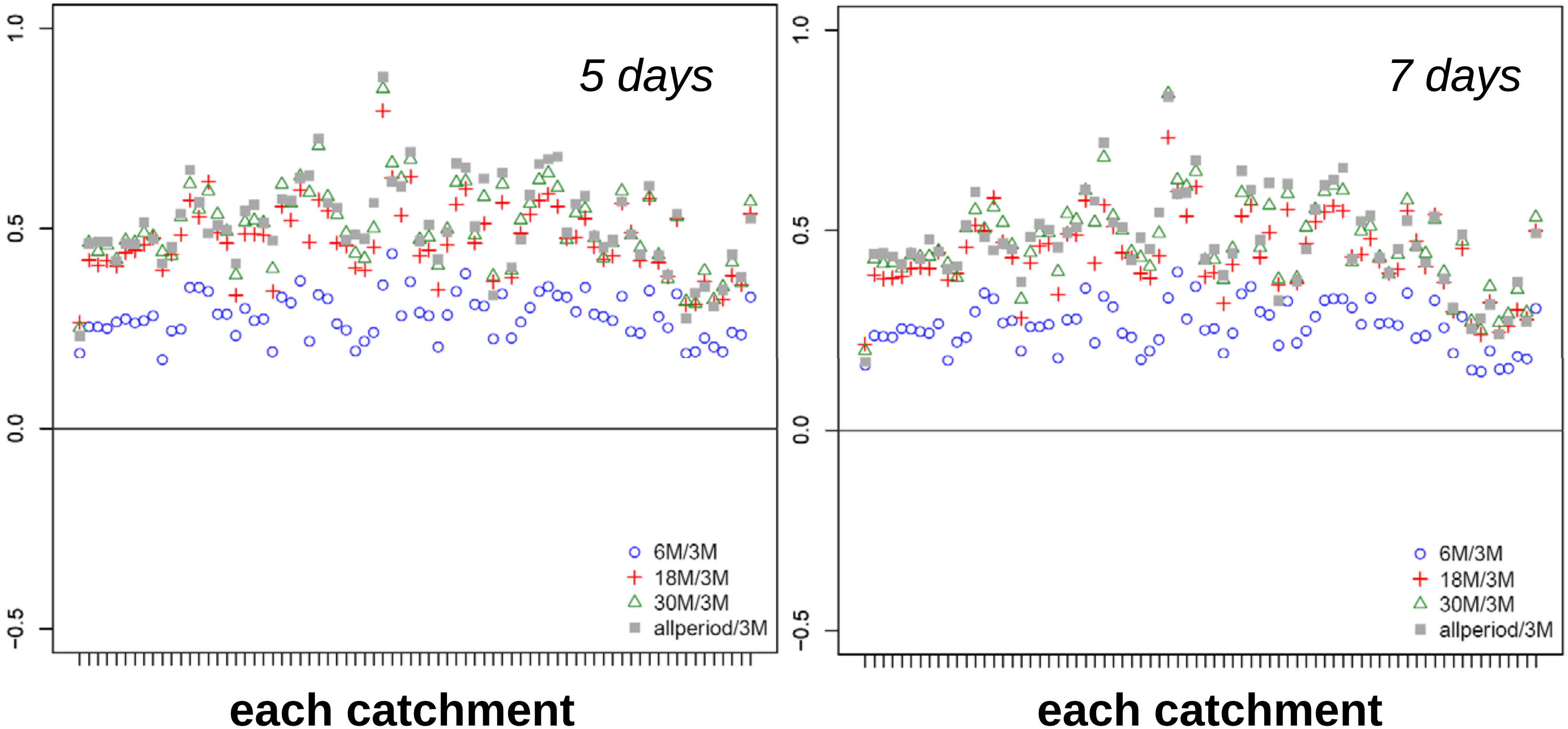
Ex: Catchment Yonne à Gurgy



=> Strong decreases for the shorter lengths
=> Improves after *12-18 months*
=> Does not reach a sill at perfect score (=0)
=> Similar behavior for all LT

Gain

$$Gain_{Xmois/3months} = 1 - \frac{RPS_{Xmonths}}{RPS_{3months}}$$



=> Significant gain when length increases up to *18 months*
=> Similar behavior for all catchments

Methodology

- Moving window of different sizes

Length of period (months)	1	3	6	12	18	24	30	36	40	42
Number of periods	42	40	37	31	25	19	13	7	3	1

Evaluation

Statistical Processing

Impact over typical scores (raw forecasts)

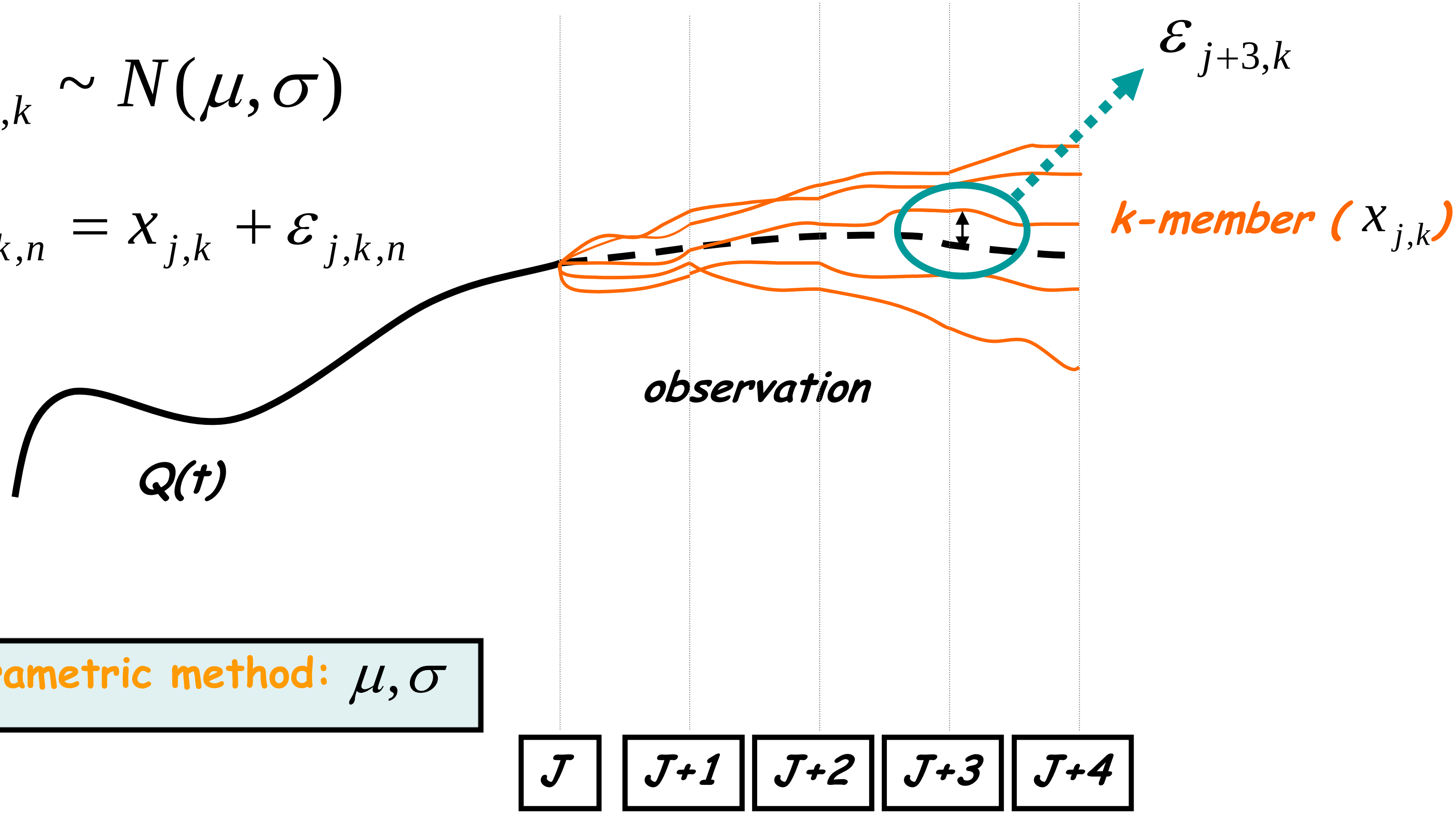
Impact over the parameters of the Best Member Method

Remembering: Best Member Method

Roulston *et al.*, (2003)

$$\varepsilon_{j,k} \sim N(\mu, \sigma)$$

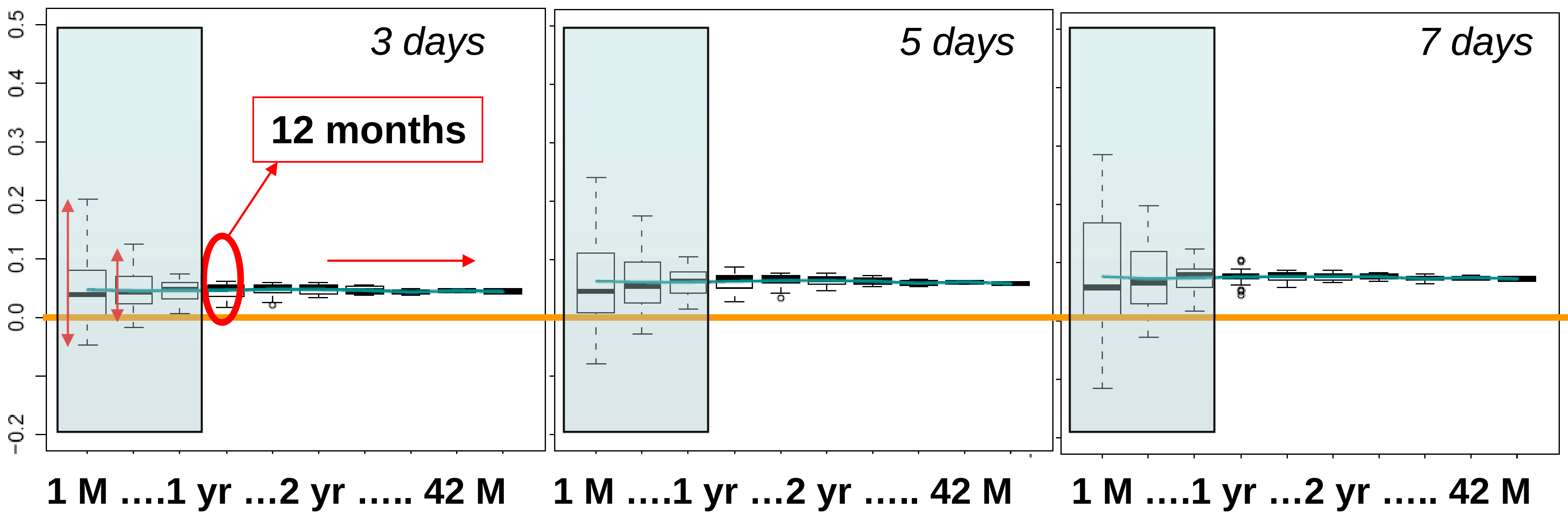
$$y_{j,k,n} = x_{j,k} + \varepsilon_{j,k,n}$$



Bi-parametric method: μ, σ

Parameter#1: μ (mean)

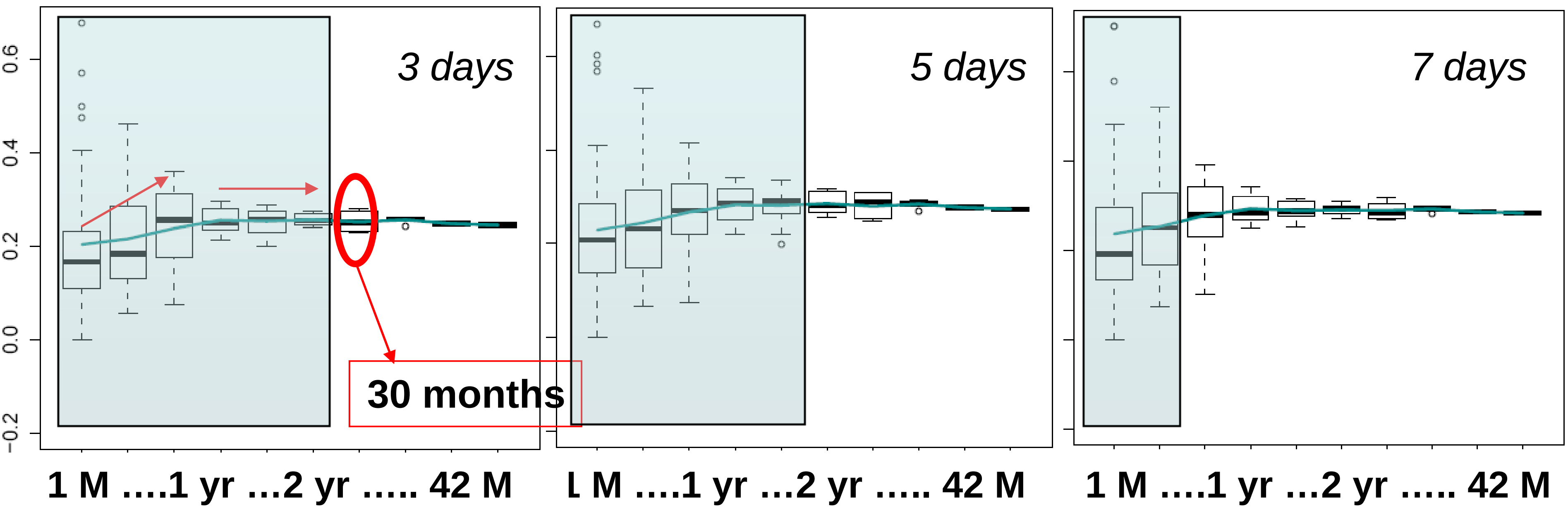
Ex: Catchment Yonne à Gurgy



=> From the *12-18 months** and on we obtain the final value of parameter#1

Parameter#2: σ (standard deviation)

Ex: Catchment Yonne à Gurgy



⇒From the *30 months** and on we obtain the final value of parameter#2

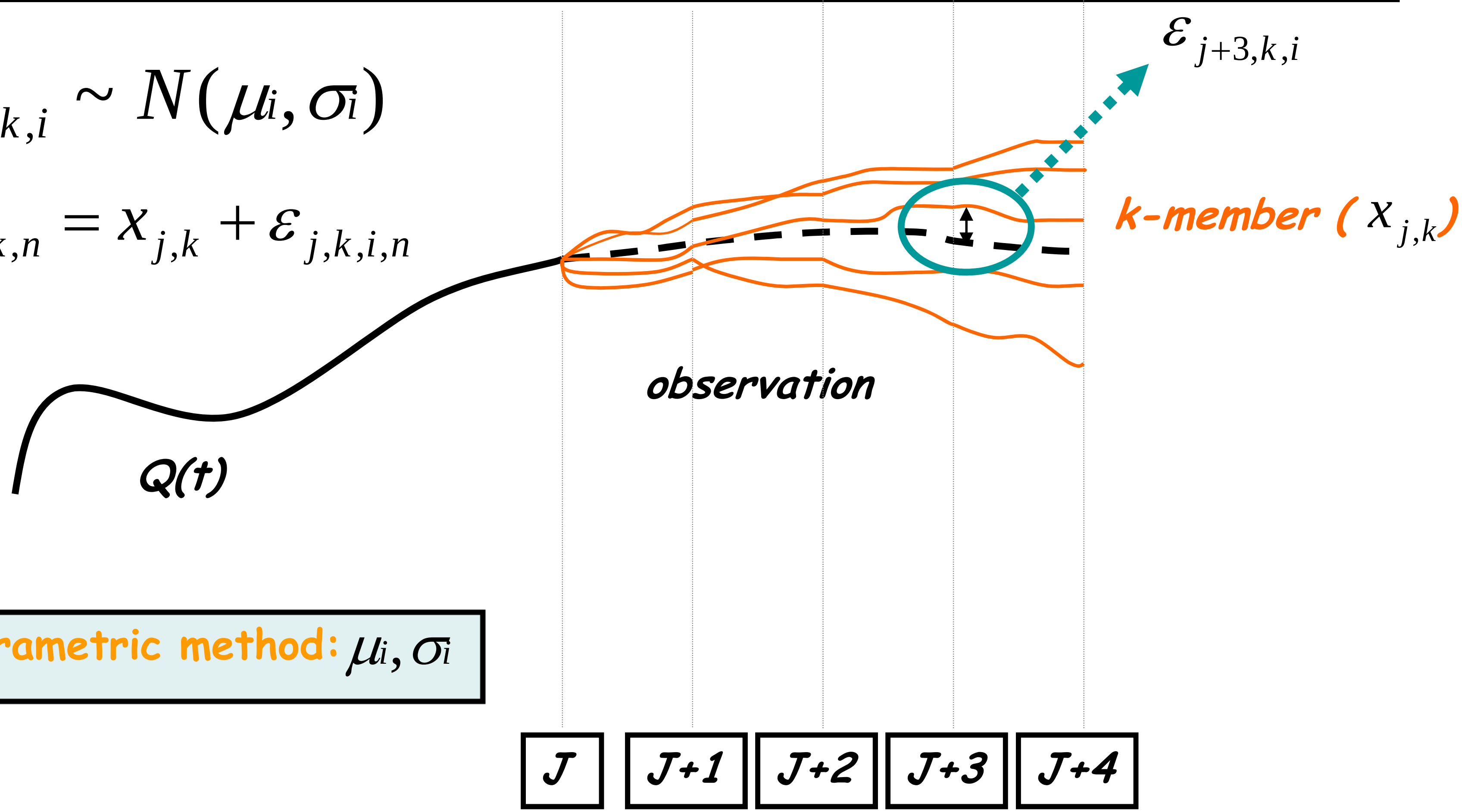
⇒For longer ldt we converse faster to the final value

Discretization by classes of streamflow

5 Classes : quantiles 20%, 40%, 60%, 80%, 100% of Q(t)

$$\varepsilon_{j,k,i} \sim N(\mu_i, \sigma_i)$$

$$y_{j,k,n} = x_{j,k} + \varepsilon_{j,k,i,n}$$

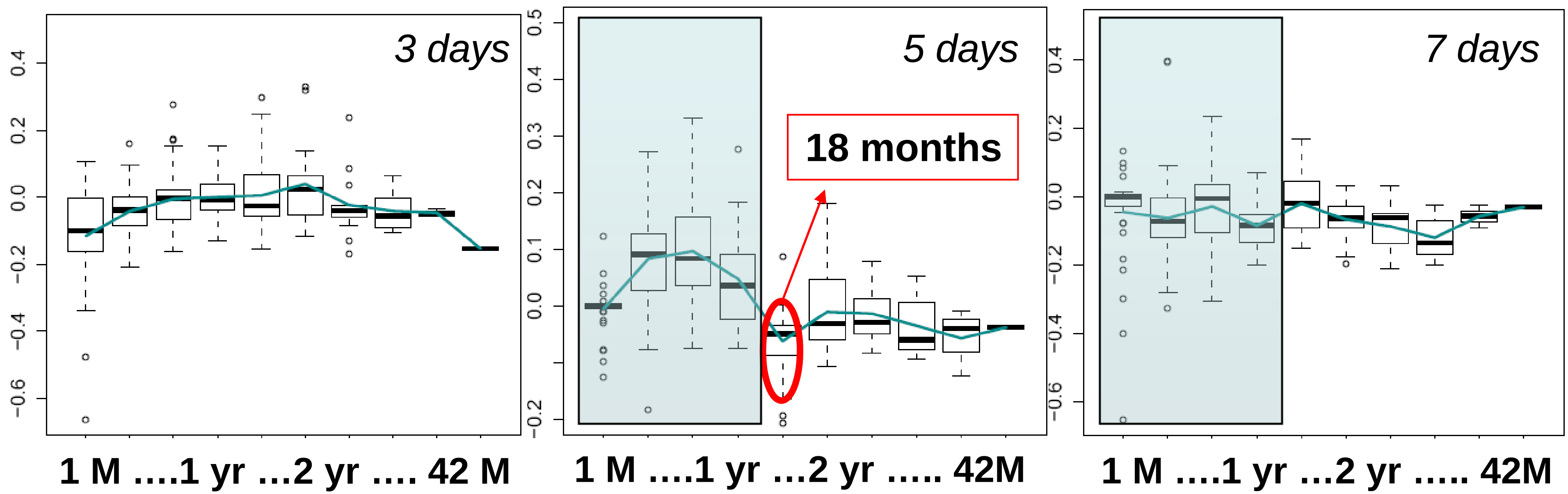


10-parametric method: μ_i, σ_i

Parameter#1: μ (mean)

* Q < 20%

Ex: Catchment Yonne à Gurgy

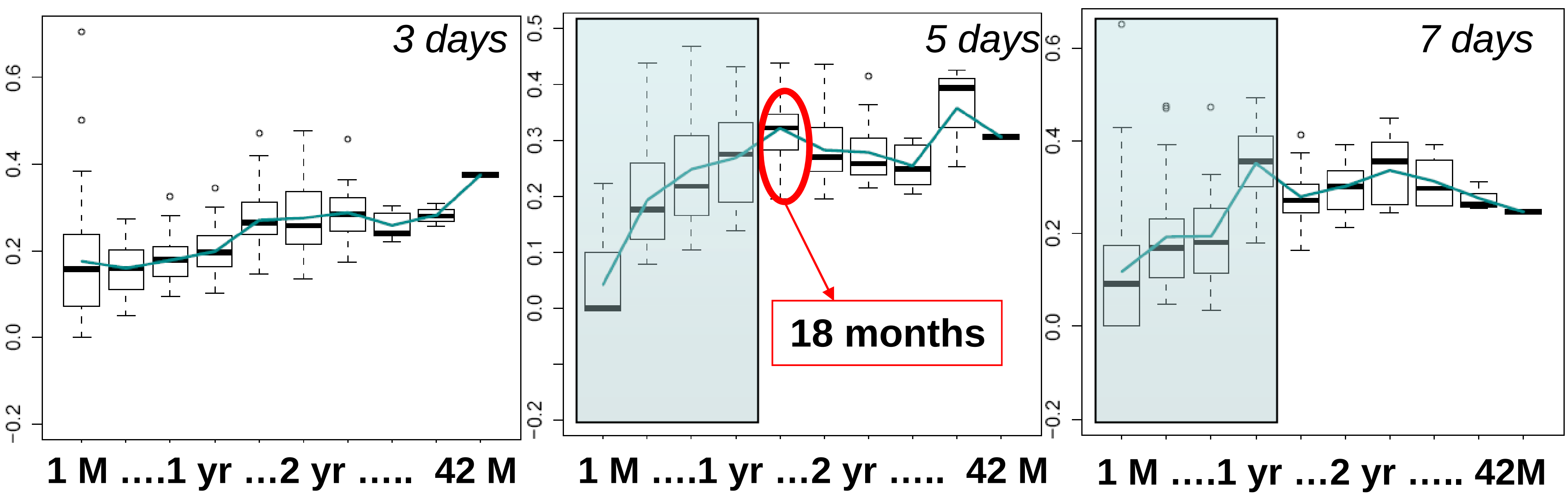


⇒From the *18-24 months** and on we obtain the final value of parameter#1

⇒For the first ldt, larger uncertainty in parameter#1 values

Parameter#2: σ (standard deviation)

Ex: Catchment Yonne à Gurgy



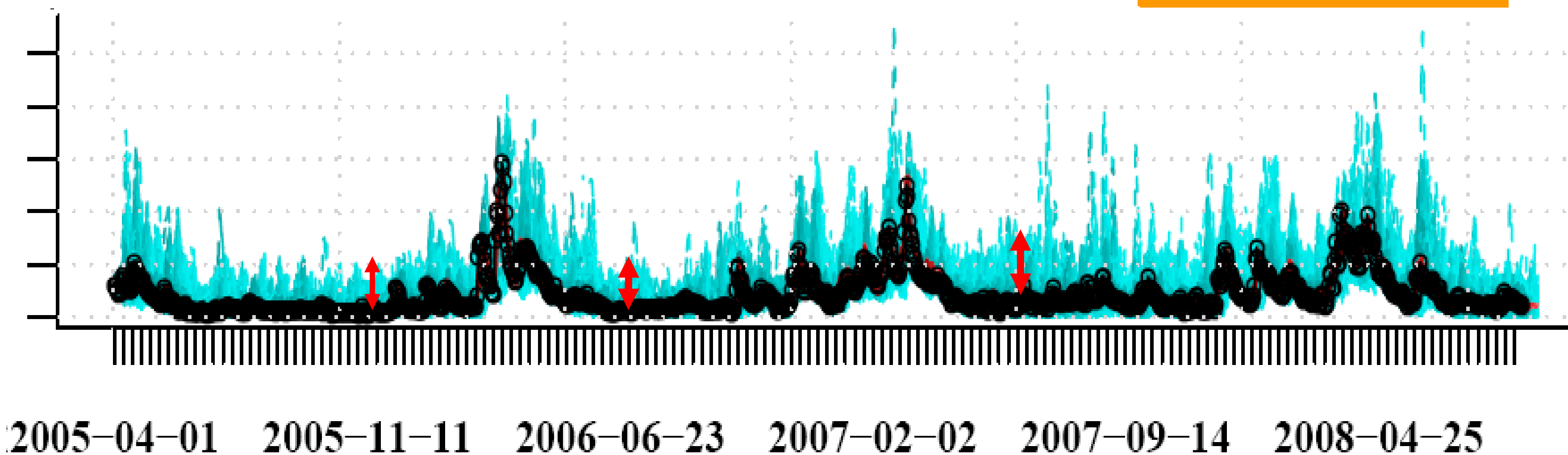
⇒From the *18 months** and on we obtain the final value of parameter#2

⇒For the first ldt, larger uncertainty in parameter#2 values

And what do we gain by applying BMM?

Ex: Catchment Yonne à Gurgy

BMM discretized



⇒BMM (2003) introduces some extra uncertainty to the raw forecasts

⇒By discretising we further increase the uncertainy selectively where the existing uncertainty isn't sufficient

Discussion

- Evaluation : a sample of pairs observations – forecasts period with a length greater than of 12-18 months is sufficient

- Calibration:

  *BMM (2003): a length of over 30 months is required for the estimation of both parameters

  *BMM (discretized): even though 18 months for the majority of classes & ldt is sufficient, we will need the longer sample possible for the first ldts.

- This behavior is consistent for all catchments

...To be continued

Thank you!