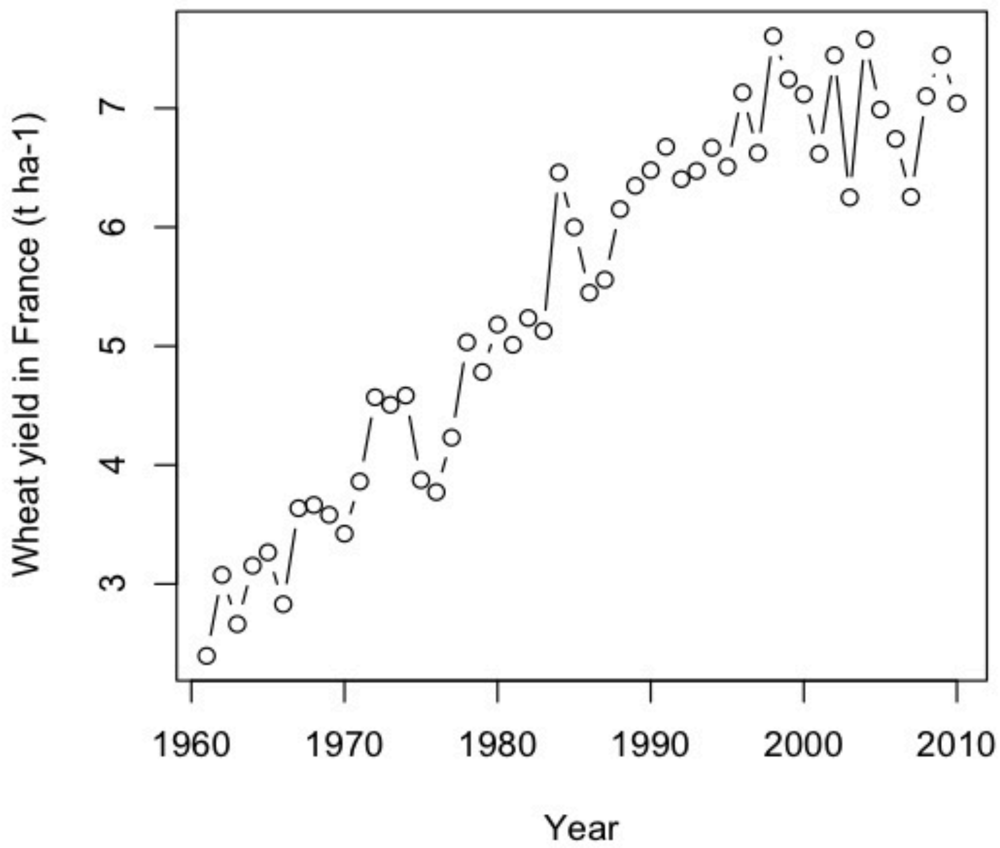
# Analysis of crop yield time series to estimate past and future trends

*Lucie Michel & David Makowski*
*(INRA,UMR 211)*

# Rational

# Le rendement plafonne-t-il ?




Données FAO-Stat

- **CASSMAN, K.G.,** 1999. Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. Proc. National Acad. Sci. (USA) 96: 5952-5959.

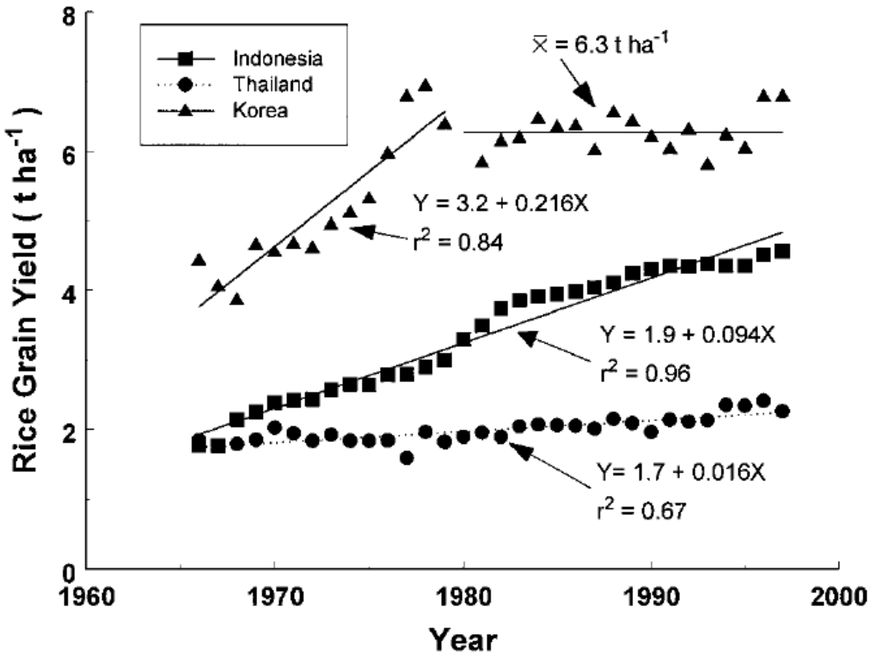



FIG. 1. National average rice yields from 1967 to 1997 in three Asian countries (http://apps.fao.org).

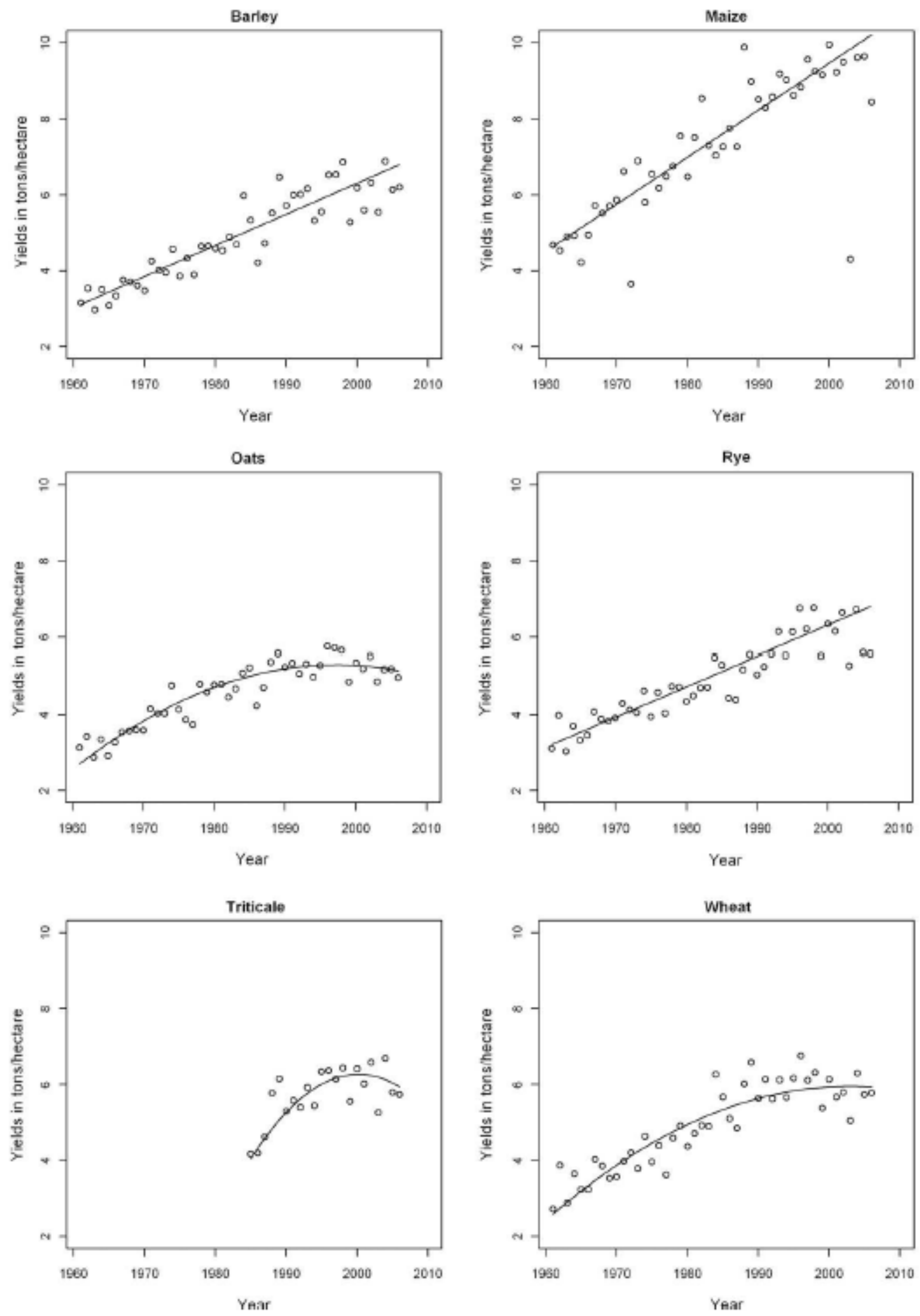


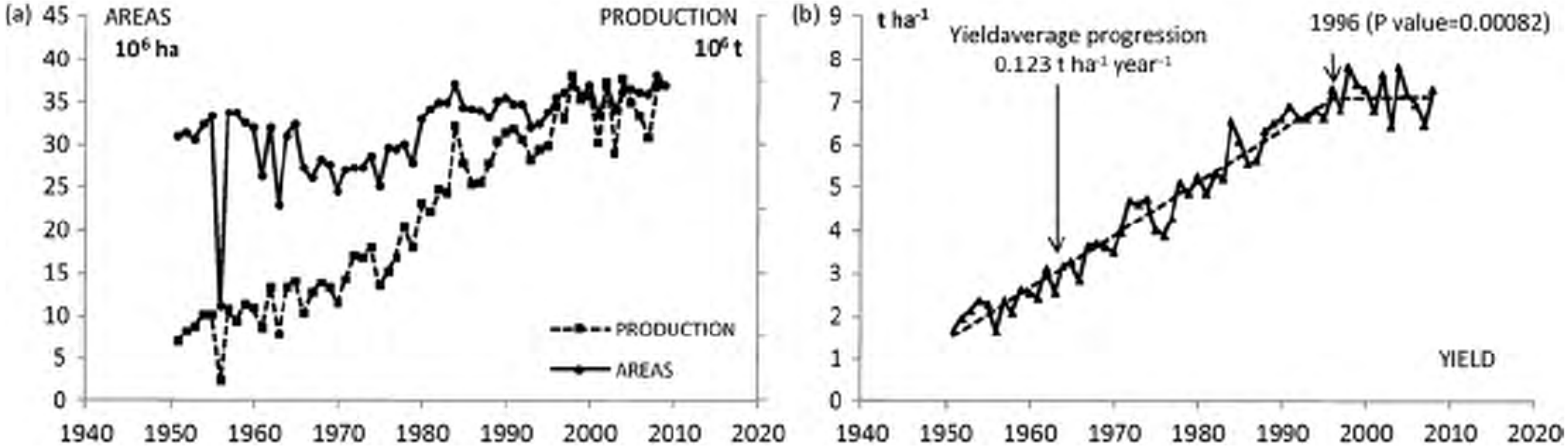

**Fig. 2.** Yields for barley, maize, oats, rye, triticale and wheat in Switzerland (1961–2006) and estimated trends. *Note:* Barley, maize and rye are fitted to a linear model (Eq. (1)). Oats, triticale and wheat are fitted to a quadratic model (Eq. (2)). Data for triticale is available only for 1985–2006. Source: FAO (2008).

- **FINGER, R.,** 2010. Evidence of slowing yield growth – The example of Swiss cereal yields. Food policy 35. p. 175-182.

Linear and quadratic models

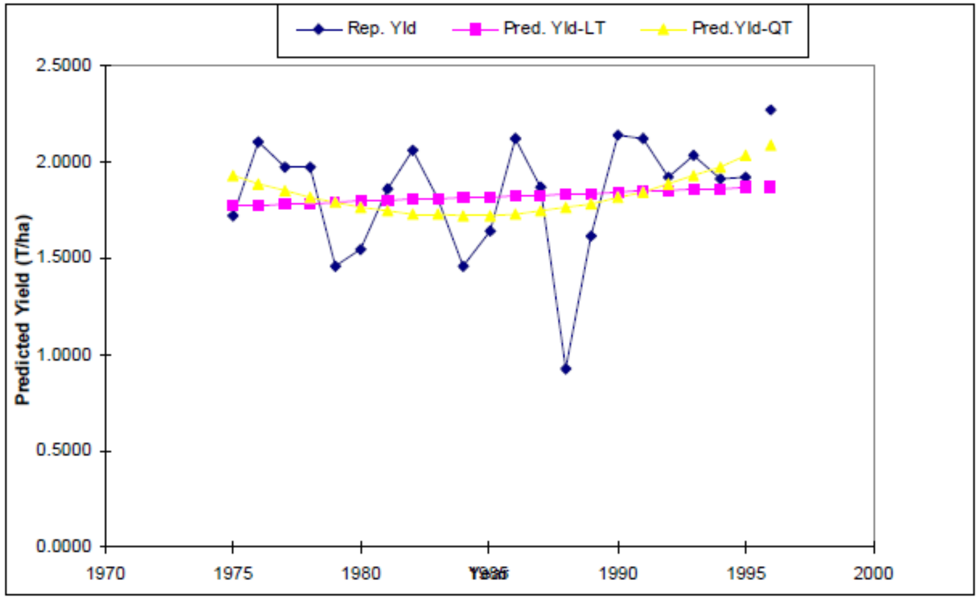
- **BRISSON, N., GATE, P., GOUACHE, D., CHARMET, G., OURY, F.-X., Huard F.,** 2010. Why are wheat yields stagnating in Europe ? A comprehensive data analysis in France. Field Crops Research 119. p. 201-212.

Linear-plus-plateau model




al evolution of bread wheat areas and production in France (a) and the subsequent yield (b). Optimization of the rising-plateau statistical mod
ıflexion. The *P* value results from the Fisher test comparing the rising-plateau model with the linear regression. *Data source*: AGRESTE.

- **KUMAR BOKEN, V.,** 2000. Forecasting Spring Wheat Yield Using Time Series Analysis: A Case Study for the Canadian Prairies. Agronomy Journal 92, n°6. p. 1047–1053.




Model comparison

spring wheat yield for Saskatchewan, during the model testing

| Year | Reported yield | Forecasted yield, ton/ha | | | | | |
|---|---|---|---|---|---|---|---|
| | | Nonstationary series | | | | Stationary series | |
| | | Trend analysis | | | | | |
| | | Linear trend | Quadratic trend | Double exponential smoothing | Double moving averaging | Simple exponential smoothing | Simple moving averaging |
| | t/ha | | | | | | |
| 1994 | 1.913 | 1.841 | 2.108 | 1.810 | 1.732 | 1.727 | 2.008 |
| 1995 | 1.921 | 1.782 | 2.105 | 1.844 | 1.837 | 1.727 | 1.626 |
| 1996 | 2.274 | 1.873 | 2.100 | 1.861 | 1.927 | 1.740 | 1.936 |
| Mean squared error | | 0.062 | 0.034 | 0.062 | 0.054 | 0.119 | 0.070 |

# Objective

Comparison of several methods for estimating yield trends

# Material & methods

# Data

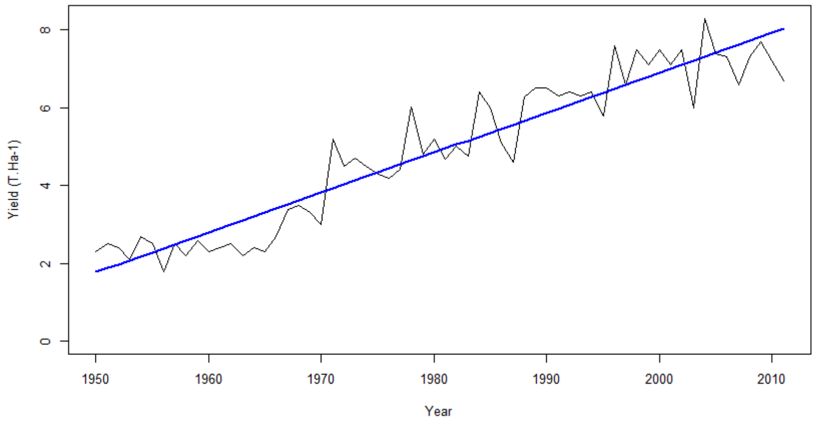
- Wheat yield time series
  - Départements (France):
    - SSP – Agreste (1950 - 2011)
      - ➢ France 1 (92 départements)
      - ➢ France 2 (59 départements)
  - World:
    - FAO – FAOSTAT (1961 – 2010)
      - ➢ World (120 pays)

# Statistical methods

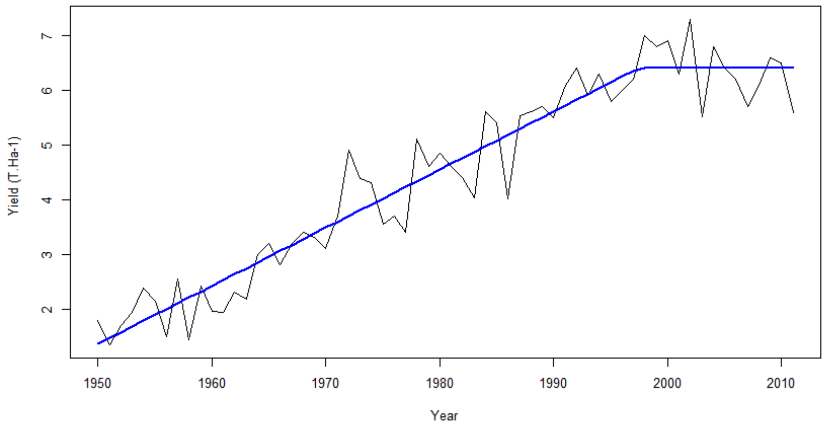
- Linear regression
- Linear + plateau
- Quadratic
- Cubic
- Exponential smoothing (Holt-Winters)
- Dynamic linear model

# Linear / Linear + plateau

- linear

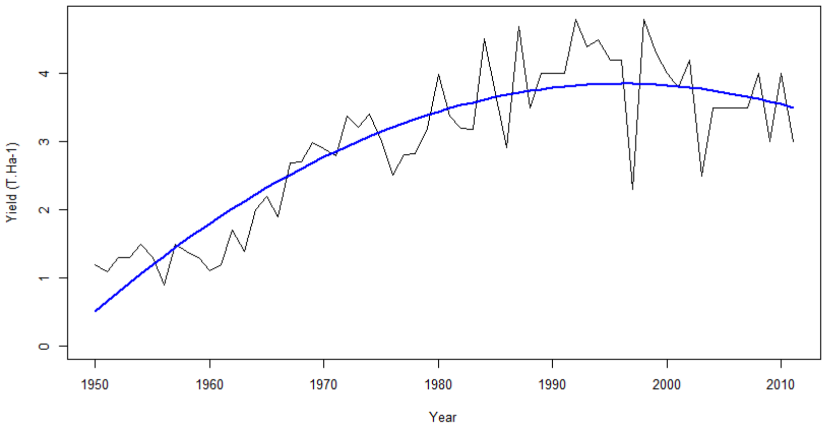



- linear + plateau

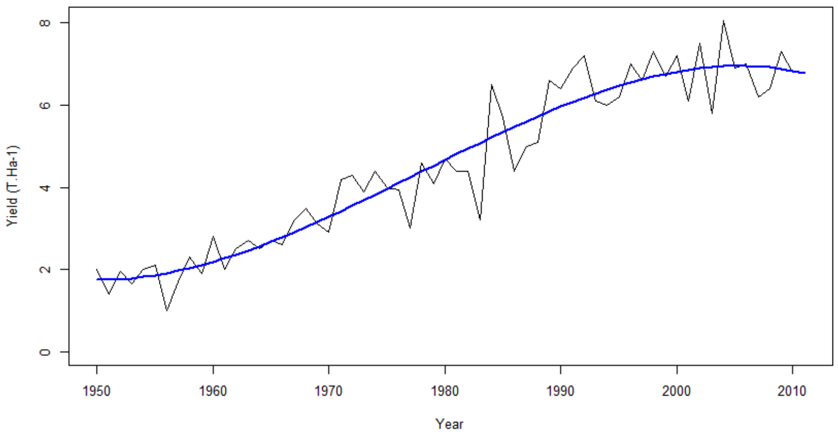


# Quadratic / Cubic

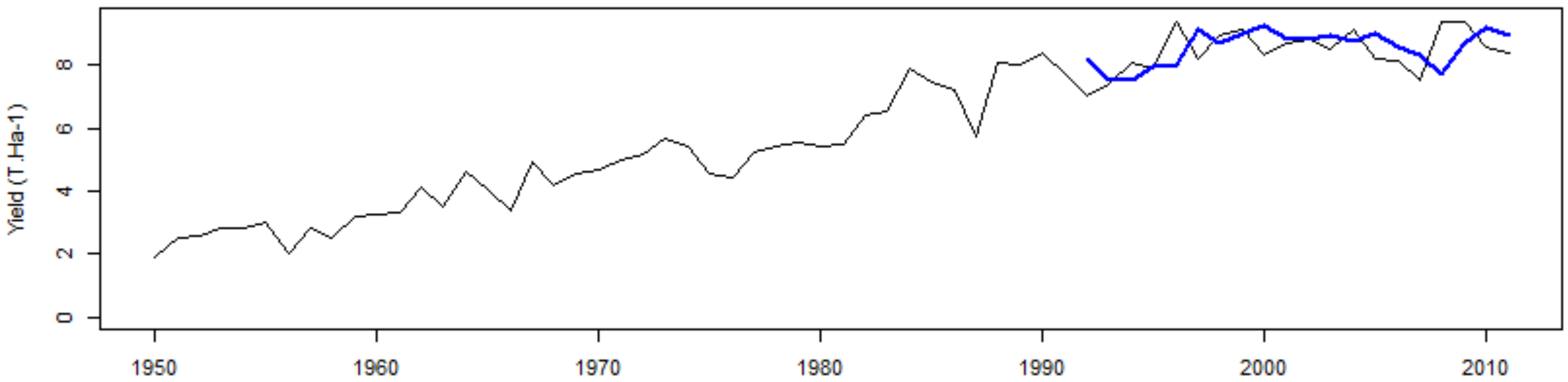
- quadratic




- cubic



# Holt-Winters

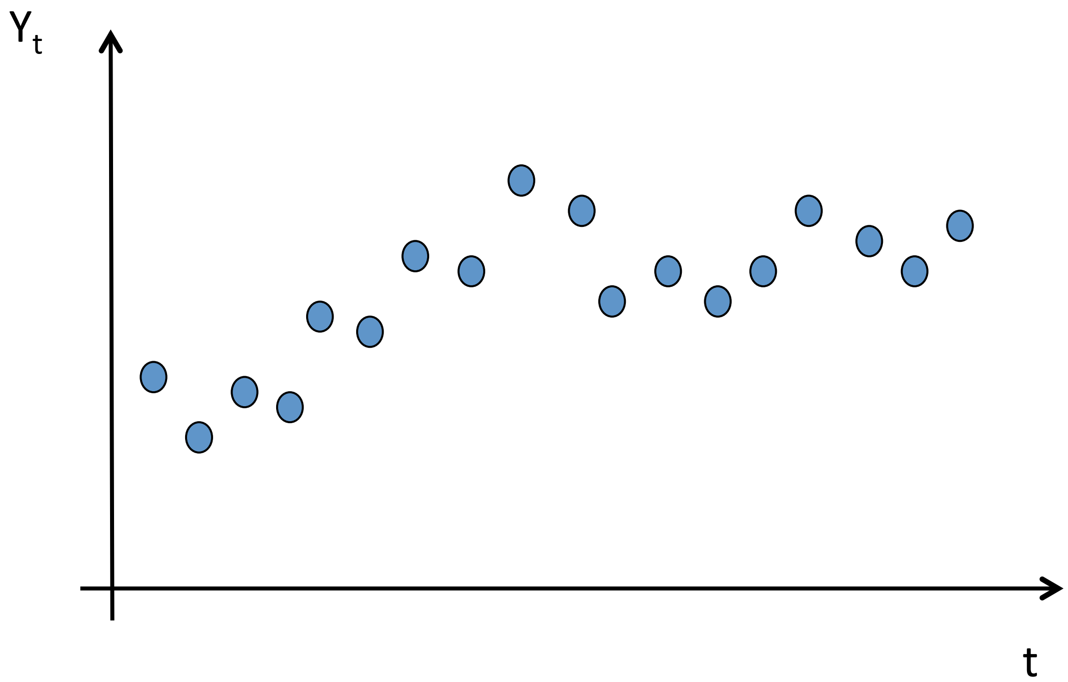



$$Y_{t+\Delta t} = m_t + b_t \times \Delta t$$

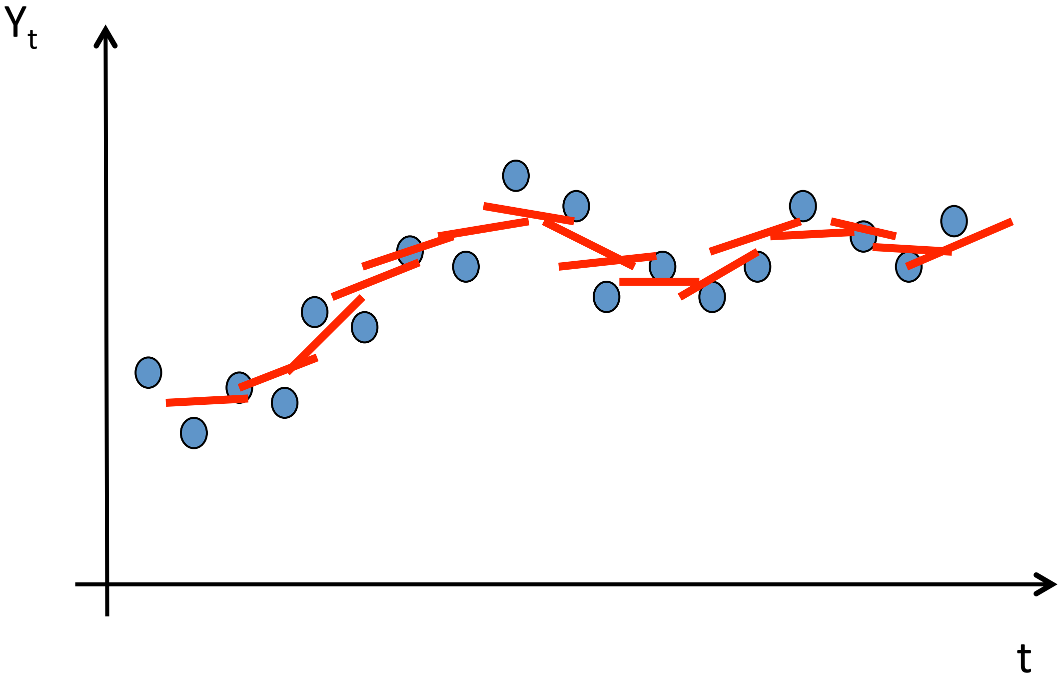
$$m_t = \lambda_0 y_t + \left(1 - \lambda_0\right) Y_t$$

$$b_t = \lambda_1 \left(m_t - m_{t-1}\right) + \left(1 - \lambda_1\right) b_{t-1}$$

# Modèle linéaire dynamique



# Modèle linéaire dynamique



# Filtrage et lissage

$\theta_t$ Vecteur des paramètres décrivant l'état du système

$\pi(\theta_t | Y_1,...,Y_t)$ Distribution de probabilité des paramètres à la date *t*, conditionnellement aux mesures obtenues jusqu'à cette date

$\pi(\theta_t | Y_1,...,Y_t,...Y_N)$

Distribution de probabilité des paramètres à la date *t*, conditionnellement à toutes les mesures disponibles

**Sous certaines hypothèses, ces distributions sont gaussiennes et déterminées par leurs espérances et variances**

# Un modèle linéaire dynamique classique

**Equation des observations**

$$Y_t = \alpha_{0t} + \alpha_{1t} \times \Delta t + \varepsilon_t$$
$$\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$$

**Equation du système**

$$\theta_t = \theta_{t-1} + \eta_{t-1}$$

$$\theta_t = \begin{pmatrix} \alpha_{0t} \\ \alpha_{1t} \end{pmatrix} \qquad \eta_{t-1} \sim N\left(0, \Sigma\right) \qquad \Sigma = \begin{pmatrix} \sigma_{\alpha 0}^2 & 0 \\ 0 & \sigma_{\alpha 1}^2 \end{pmatrix}$$

# Filtrage et lissage

Le filtrage (données utilisées jusqu'à la date t) permet de calculer :

$$E\left(\alpha_{1t} \middle| Y_1,\ldots,Y_t\right) \qquad Var\left(\alpha_{1t} \middle| Y_1,\ldots,Y_t\right)$$

Le lissage (toutes les données sont utilisées) permet de calculer :

$$E\left(\alpha_{1t} \middle| Y_1,\ldots,Y_N\right) \qquad Var\left(\alpha_{1t} \middle| Y_1,\ldots,Y_N\right)$$

Prédiction à la date *N*+*k* :

$$\hat{Y}_{N+k} = E\left(\alpha_{0t} \middle| Y_1,\ldots,Y_N\right) + E\left(\alpha_{1t} \middle| Y_1,\ldots,Y_N\right) \times k$$

# Etape de correction du filtre de Kalman

$$E\left(\theta_t \left| y_1,\ldots,y_t \right.\right) = E\left(\theta_t \left| y_1,\ldots,y_{t-1} \right.\right) + K\left[ y_t - F\ E\left(\theta_t \left| y_1,\ldots,y_{t-1} \right.\right)\right]$$

$$K = Var\left(\theta_t \left| y_1,\ldots,y_{t-1} \right.\right) F^T \left[ F\ Var\left(\theta_t \left| y_1,\ldots,y_{t-1} \right.\right) F^T + V \right]^{-1}$$

# Estimation des paramètres inconnus

$$\sigma_{\varepsilon}^{2} \quad \sigma_{\alpha 1}^{2} \quad \sigma_{\alpha 0}^{2}$$

- Maximum de vraisemblance
- Méthode bayésienne

# Package R dlm (Petris et al., 2009): Estimation par max de vraisemblance

```
MyModel<-function(x) {
        return(dlmModPoly(2, dV=exp(x[1]), dW=c(exp(x[2]), exp(x[3]))))
        }
```

```
FittedModel<-MyModel(c(0,-5,-5))
```

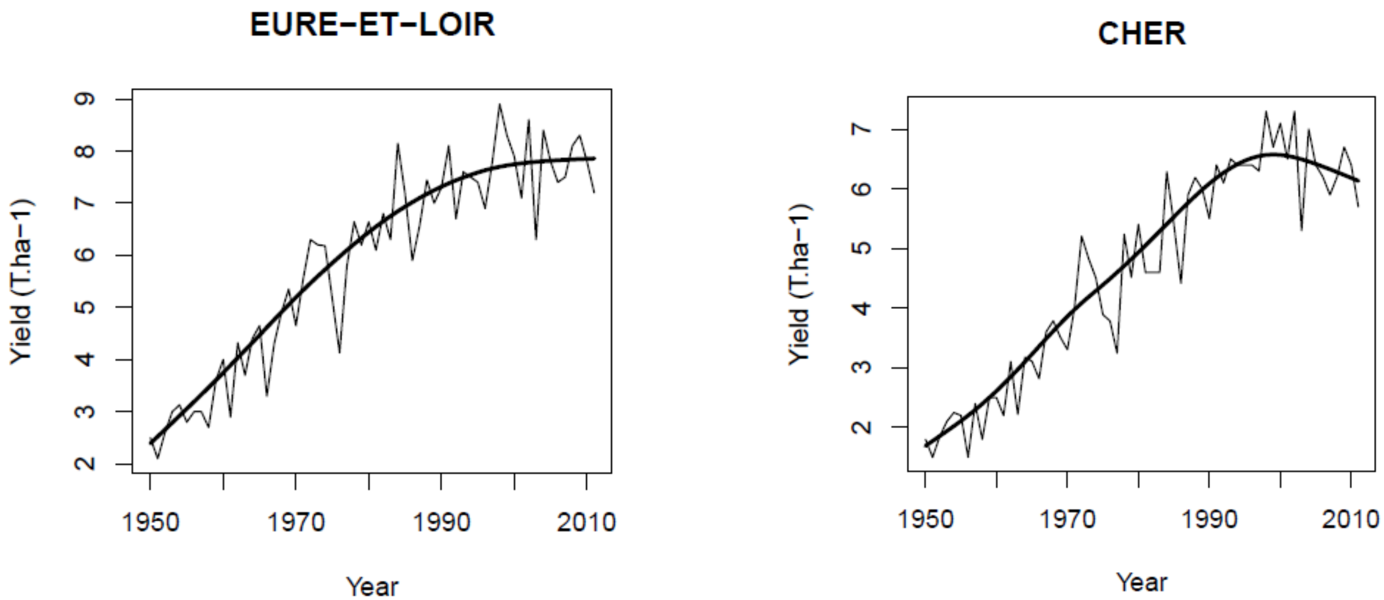
```
dlmMLE(Yield,parm=c(0,0,0), build=MyModel)
```

```
FittedModel<-MyModel(fitMyModel$par)

YieldFilter<-dlmFilter(Ym, FittedModel)

YieldSmooth<-dlmSmooth(Ym, FittedModel)
```

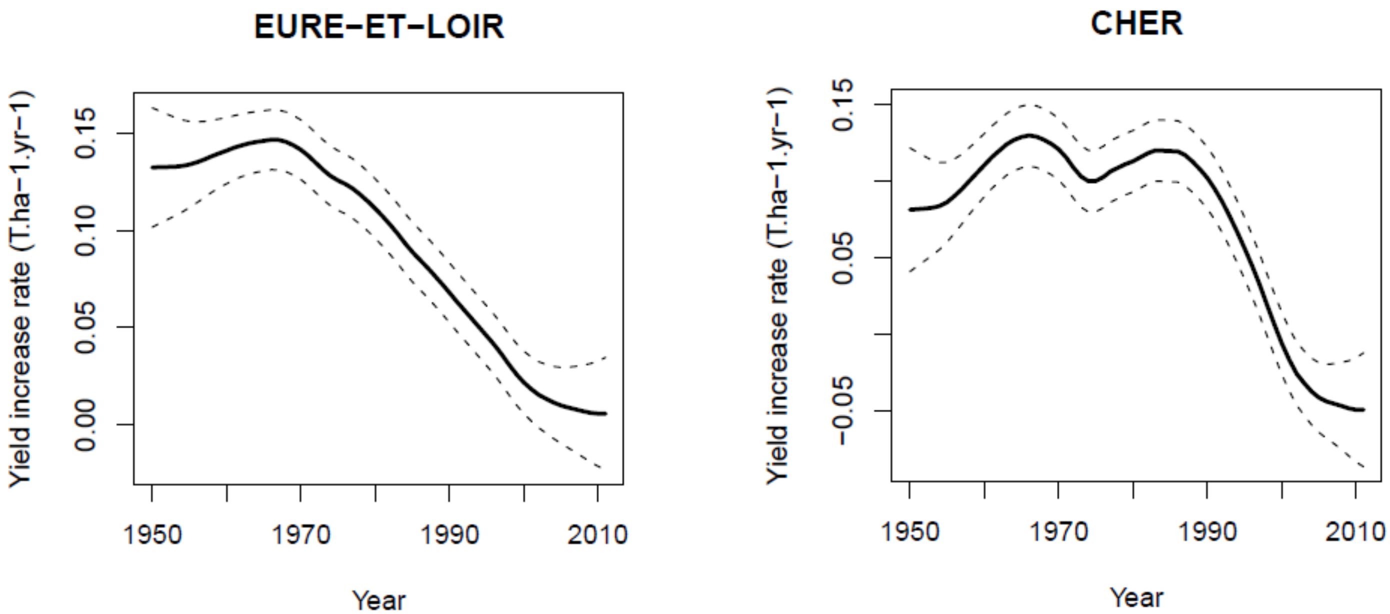
# Package R dlm (Petris et al., 2009): MCMC

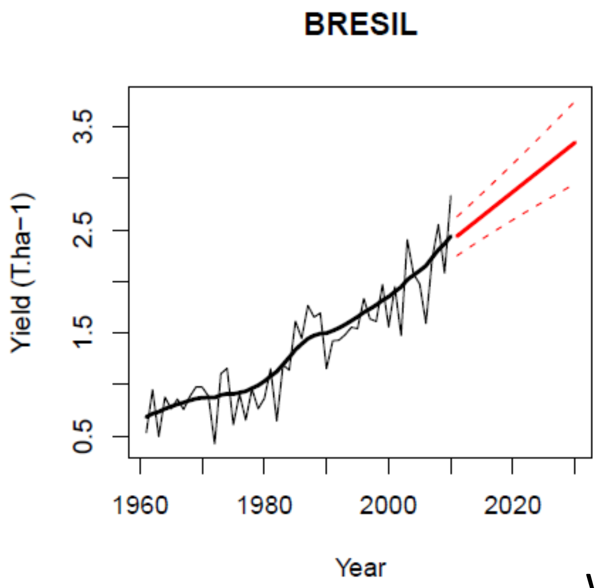
```
dlmGibbsDIG(     data,
                 mod=dlmModPoly(2),
                 a.y=1, b.y=1000,
                 a.theta=10, b.theta=10000
                 n.sample=10000,
                 thin=1
                         )
```

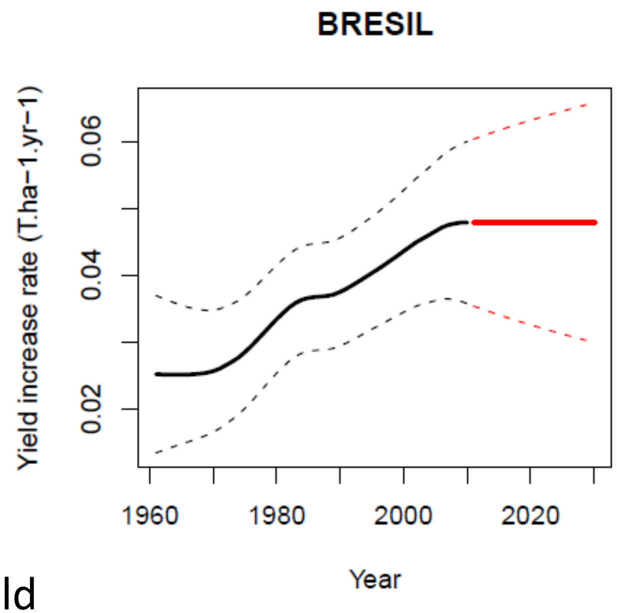
Wheat yield



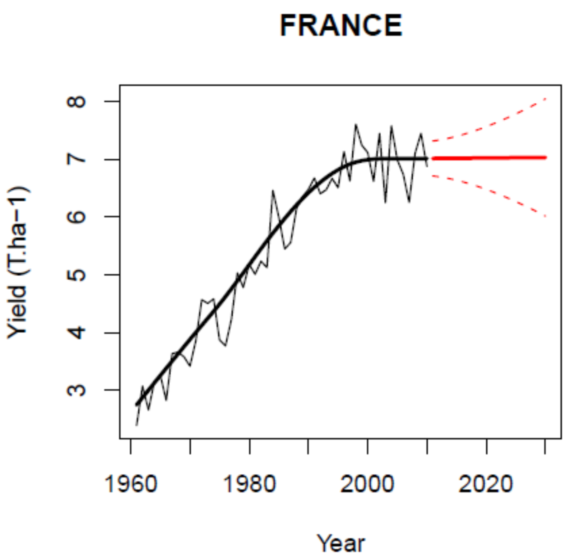

Wheat yield yearly increase

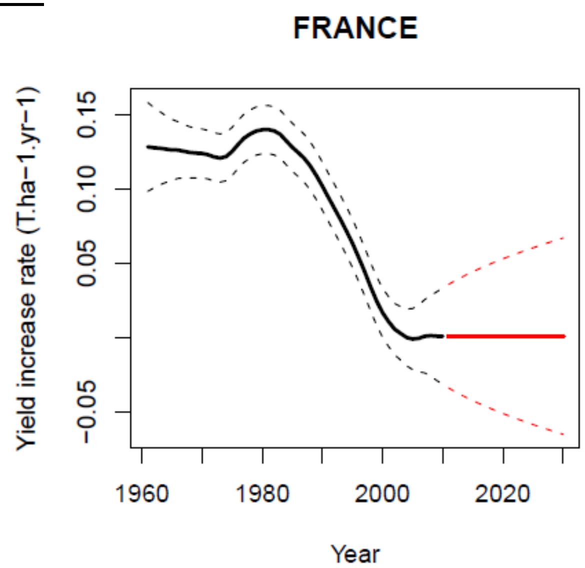


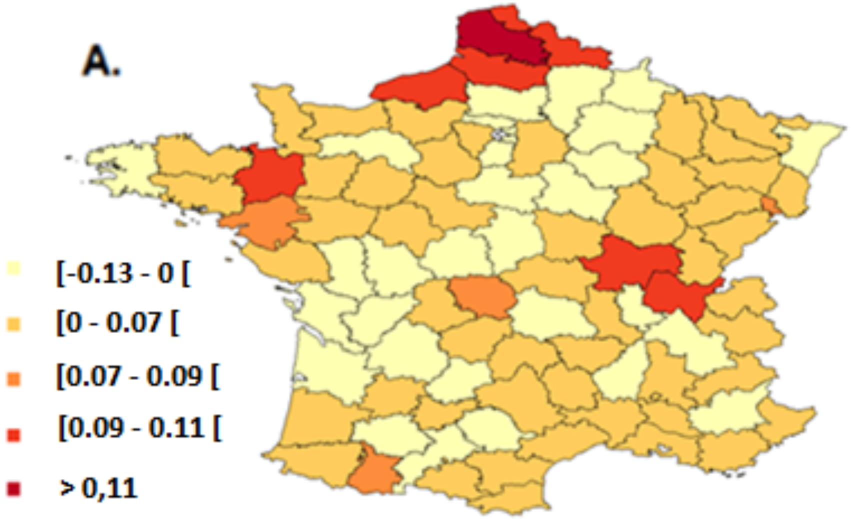

BRESIL
Yield (T.ha−1)
0.5
1.5
2.5
3.5
1960
1980
2000
2020
Year

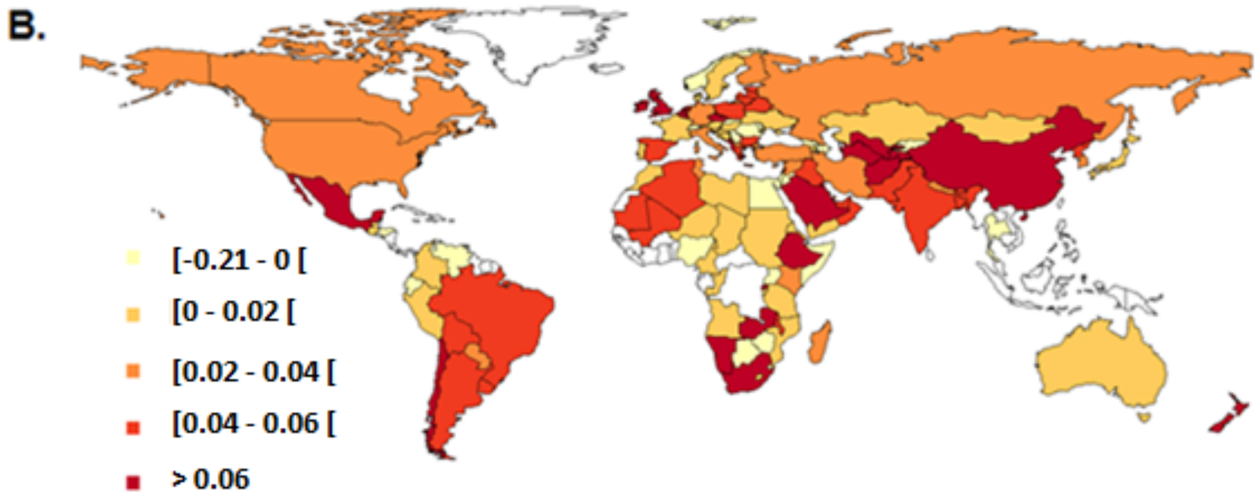

BRESIL
Yield increase rate (T.ha−1.yr−1)
0.02
0.04
0.06
1960
1980
2000
2020
Year

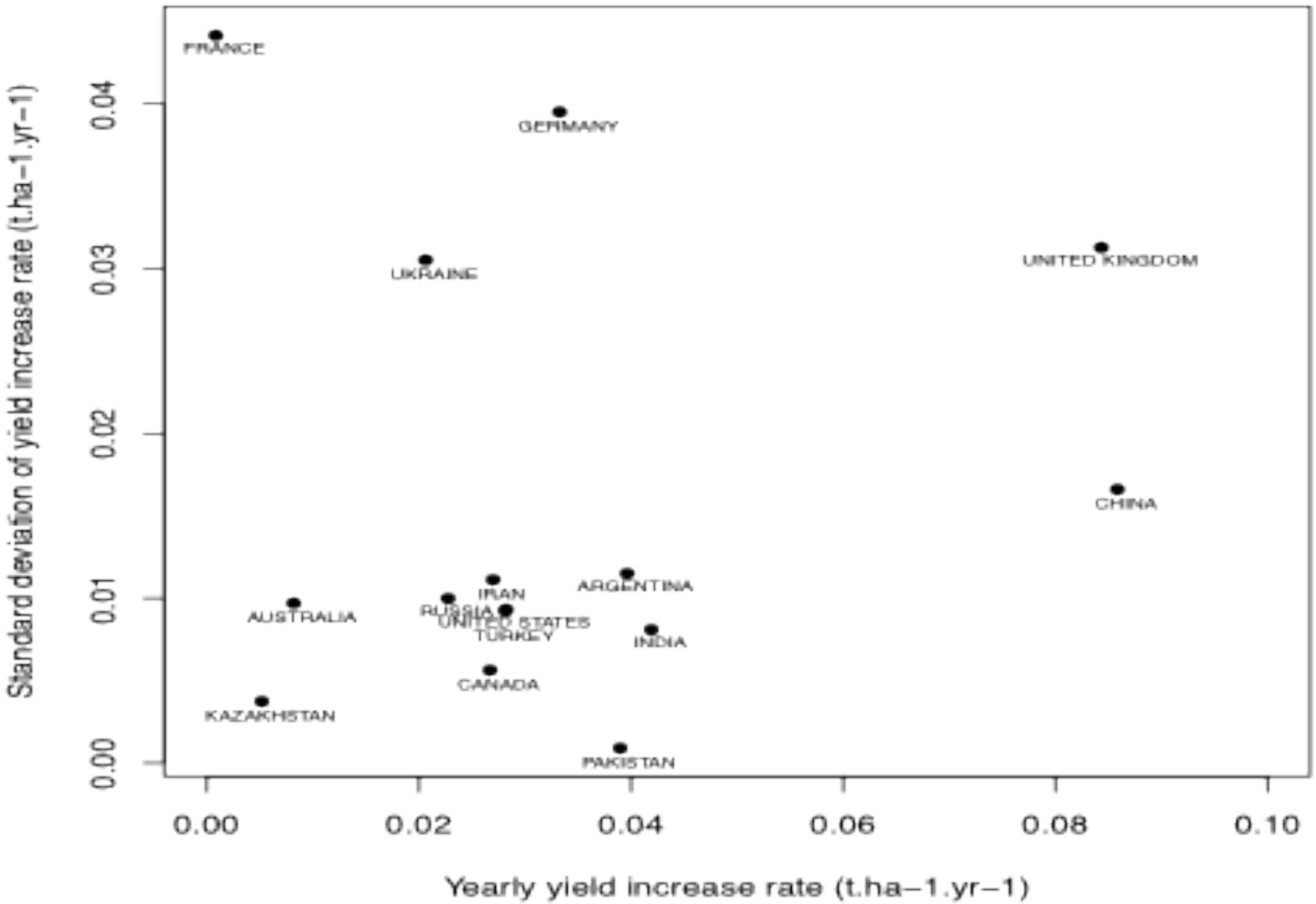
# Wheat yield


FRANCE
Yield (T.ha−1)
3
4
5
6
7
8
1960
1980
2000
2020
Year


FRANCE
Yield increase rate (T.ha−1.yr−1)
−0.05
0.05
0.10
0.15
1960
1980
2000
2020
Year


A.
[-0.13 - 0 [
[0 - 0.07 [
[0.07 - 0.09 [
[0.09 - 0.11 [
> 0,11


B.
[-0.21 - 0 [
[0 - 0.02 [
[0.02 - 0.04 [
[0.04 - 0.06 [
> 0.06



FRANCE
GERMANY
UKRAINE
UNITED KINGDOM
CHINA
IRAN
ARGENTINA
AUSTRALIA
RUSSIA
UNITED STATES
TURKEY
INDIA
CANADA
KAZAKHSTAN
PAKISTAN
0.00
0.01
0.02
0.03
0.04
Standard deviation of yield increase rate (t.ha−1.yr−1)
0.00
0.02
0.04
0.06
0.08
0.10
Yearly yield increase rate (t.ha−1.yr−1)

# RMSE/RMSEP

$$RMSE = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - Y_t\right)^2$$

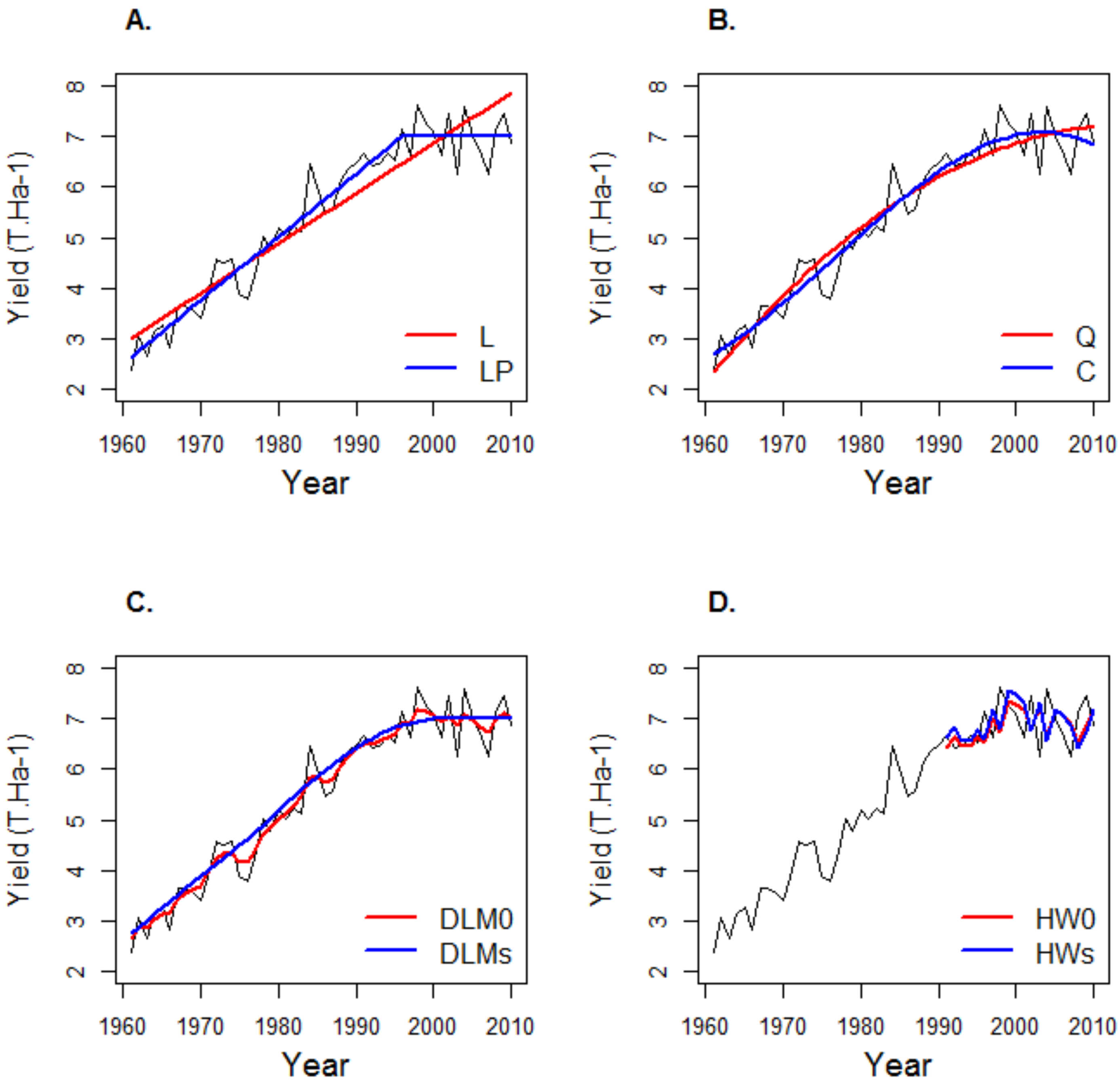
$$RMSEP = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - Y_{Pt}\right)^2$$

# RMSE/RMSEP

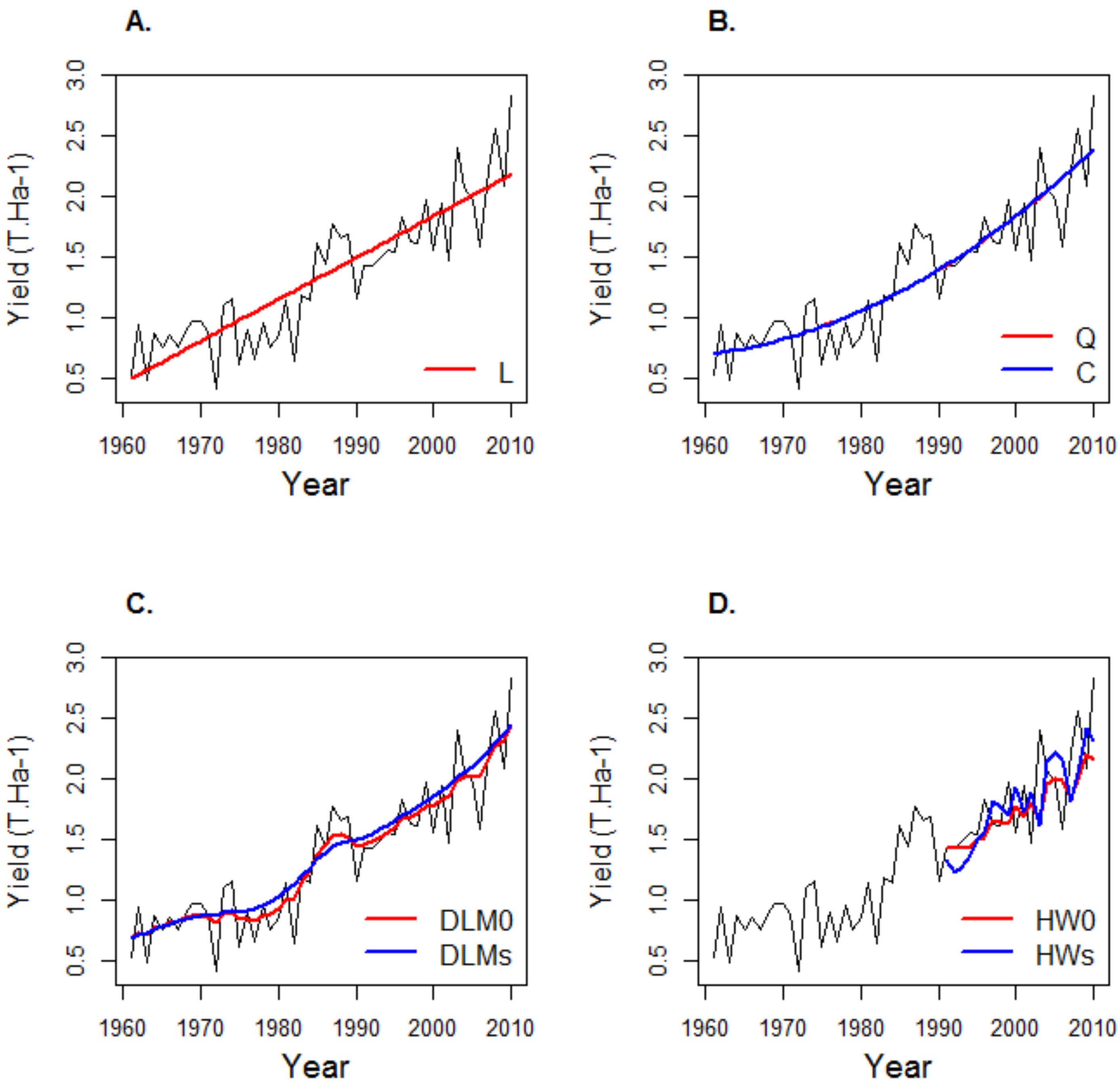
- RMSEP estimated by cross-validation
- Dataset splitted into two parts.
    - ✓ One part used for model fitting (parameter estimation)
    - ✓ Second part used for prediction

- Predictions from 1992 to 2011
    - RMSEP one-year ahead
        - Estimation with data <1992 for predicting 1992
        - Estimation with data <1993 for predicting 1993
        - …
        - Estimation with data <2011 for predicting 2011

    - RMSEP 10-year ahead:
        - Estimation with data <1992 for predicting 1992 to 2001
        - Estimation with data <2002 for predicting 2002 to 2011



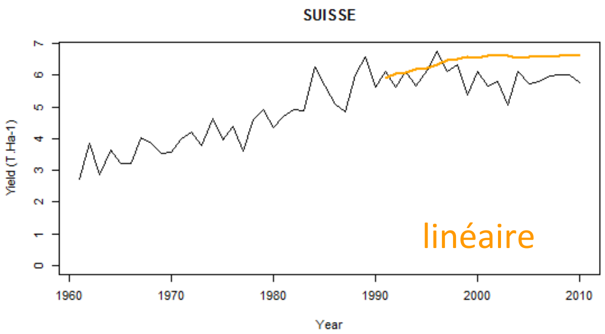
A.
B.
C.
D.
Yield (T.Ha-1)
Year
1960
1970
1980
1990
2000
2010
2
3
4
5
6
7
8
L
LP
Q
C
DLM0
DLMs
HW0
HWs

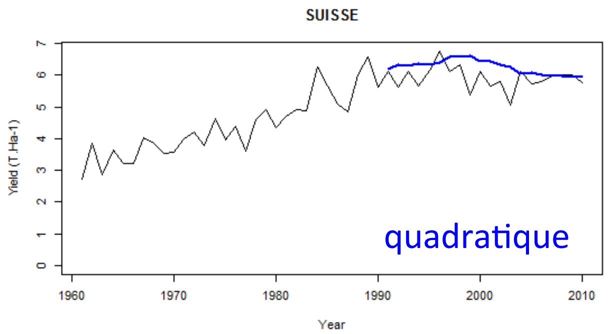

A.
B.
C.
D.
Yield (T.Ha-1)
0.5
1.0
1.5
2.0
2.5
3.0
1960
1970
1980
1990
2000
2010
Year
L
Q
C
DLM0
DLMs
HW0
HWs

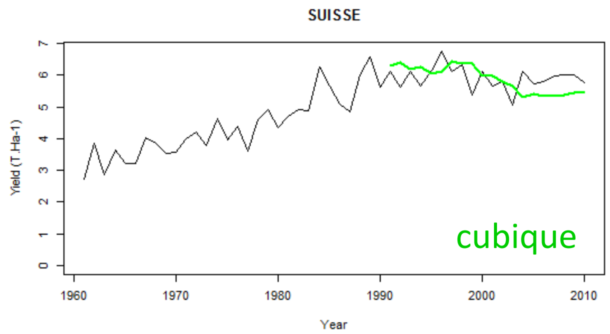
- RMSE

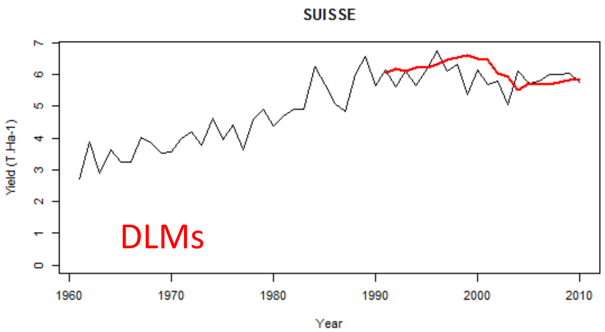
| Scale | Unit | Linear | Quadratic | Cubic | DLMs | DLM0 | Linear plateau |
|---|---|---|---|---|---|---|---|
| ***France 1*** | **T.Ha-1** | 0.59 | 0.54 | 0.49 | 0.47 | **0.38** | NA |
| | *%* | *54.11* | *40.69* | *28.72* | *23.05* | ***0.00*** | *NA* |
| ***France 2*** | **T.Ha-1** | 0.58 | 0.53 | 0.49 | 0.47 | **0.38** | 0.4966 |
| | *%* | *53.61* | *40.03* | *27.95* | *22.73* | ***0.00*** | *30.96* |
| ***World*** | **T.Ha-1** | 0.40 | 0.35 | 0.32 | 0.23 | **0.20** | NA |
| | *%* | *96.98* | *74.03* | *59.76* | *14.57* | ***0.00*** | *NA* |

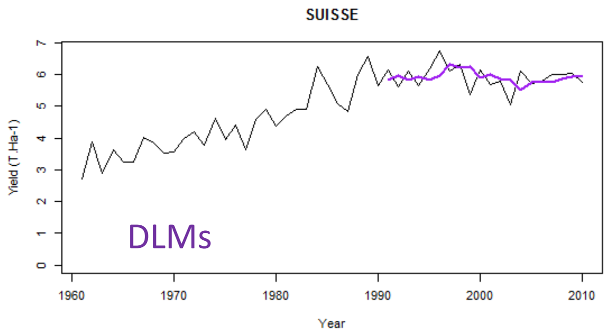
# RMSEP one-year ahead

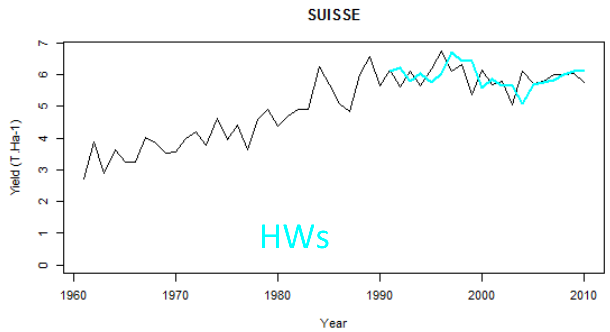


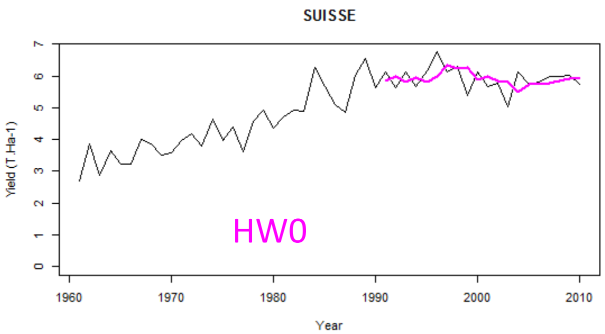



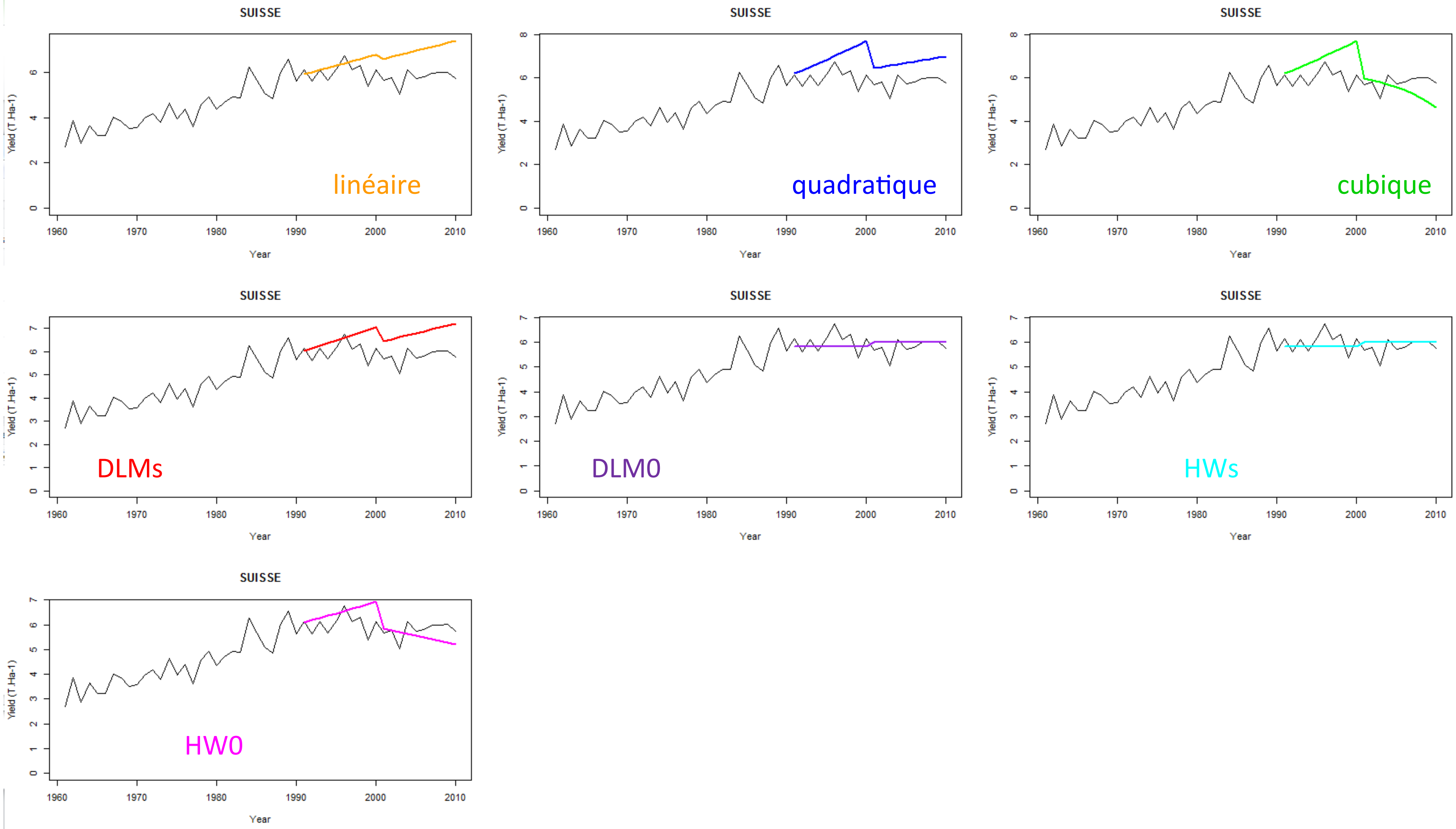



# RMSEP 10-year ahead

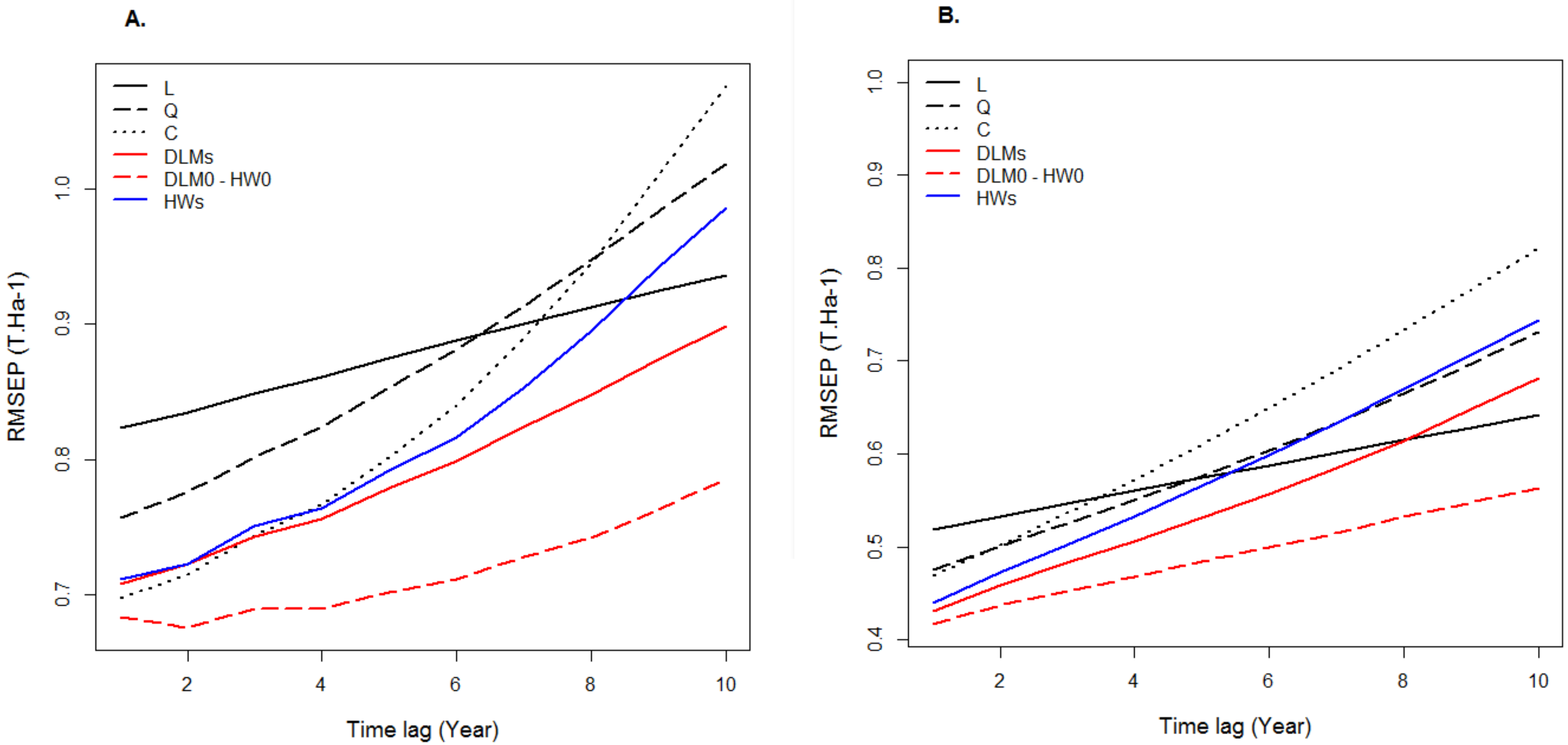


- RMSEP1

| Scale | Unit | Linear | Quadratic | Cubic | DLMs | DLM0 | HWs | HW 0 |
|---|---|---|---|---|---|---|---|---|
| *France 1* | T.Ha-1 | 0.81 | 0.75 | 0.69 | 0.71 | 0.68 | 0.71 | **0.68** |
| | % | *18.73* | *10.49* | *1.64* | *3.42* | *0.09* | *3.98* | ***0.00*** |
| *World* | T.Ha-1 | 0.52 | 0.48 | 0.48 | 0.43 | 0.42 | 0.44 | **0.42** |
| | % | *24.45* | *14.31* | *14.75* | *3.53* | *0.02* | *5.55* | ***0.00*** |

- RMSEP10

| Scale | Unit | Linear | Quadratic | Cubic | DLMs | DLM0 | HWs | HW 0 |
|---|---|---|---|---|---|---|---|---|
| *France 1* | T.Ha-1 | 0.94 | 1.07 | 1.06 | 0.93 | **0.70** | 0.97 | 0.70 |
| | % | *34.41* | *52.37* | *50.38* | *32.28* | ***0.00*** | *38.61* | *0.20* |
| *World* | T.Ha-1 | 0.66 | 0.77 | 0.85 | 0.68 | 0.59 | 0.73 | **0.59** |
| | % | *11.61* | *30.66* | *43.74* | *15.02* | *0.07* | *23.69* | ***0.00*** |



A.
L
Q
C
DLMs
DLM0 - HW0
HWs
RMSEP (T.Ha-1)
0.7
0.8
0.9
1.0
2
4
6
8
10
Time lag (Year)
B.
L
Q
C
DLMs
DLM0 - HW0
HWs
RMSEP (T.Ha-1)
0.4
0.5
0.6
0.7
0.8
0.9
1.0
2
4
6
8
10
Time lag (Year)

# What is the best approach?

- The winners are
  - Holt-Winters (HW0)
  - DLM (DLM0)
- DLM vs HW:
  - DLM has several advantages:
    - Give probability distributions instead of point value
    - Smoothing: estimation of past trends
- Among the methods including a trend, DLMs seems to be the best

Petris G., Petrone S., Campagnoli P. 2009.
Dynamic linear model with R. Springer