



**INRA**  
SCIENCE & IMPACT

**Mercredi 19 juin 2013**

**Evolution des rendements de culture à  
différentes échelles :**

**Estimation des tendances passées et futures  
en tenant compte des incertitudes**

**Séminaire final de  
restitution du projet CASDAR:**

« Associer un niveau d'erreur aux prédictions des  
modèles mathématiques pour l'agronomie et  
l'élevage »

*D. Makowski (INRA, UMR Agronomie)*

*L. Michel (INRA, UMR Agronomie)*



# Contexte

# Contexte

Etude de séries chronologiques de rendement  
de culture et de leur évolution

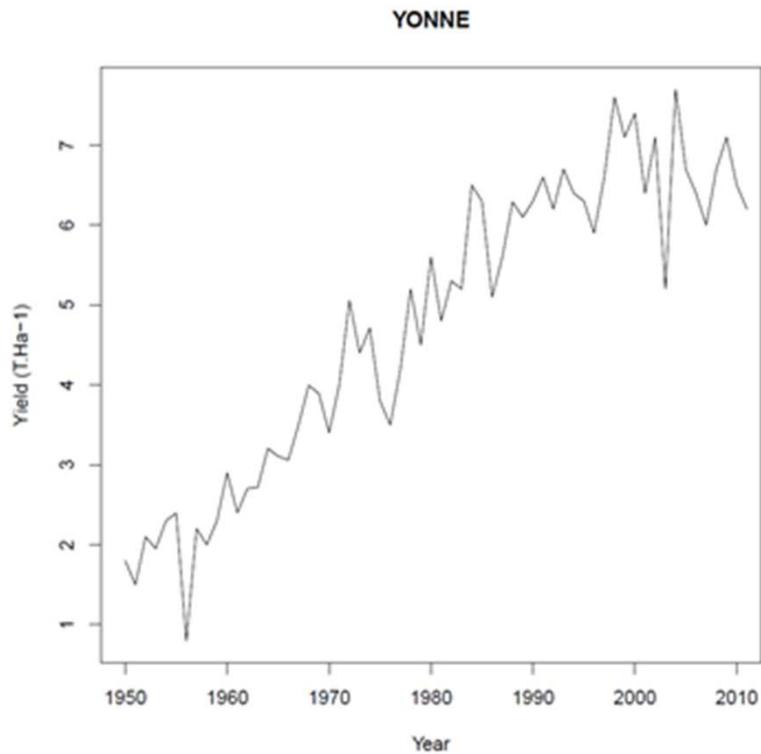
# Contexte: Quoi ?

## Série chronologique

Série chronologique = observations à des dates différentes.

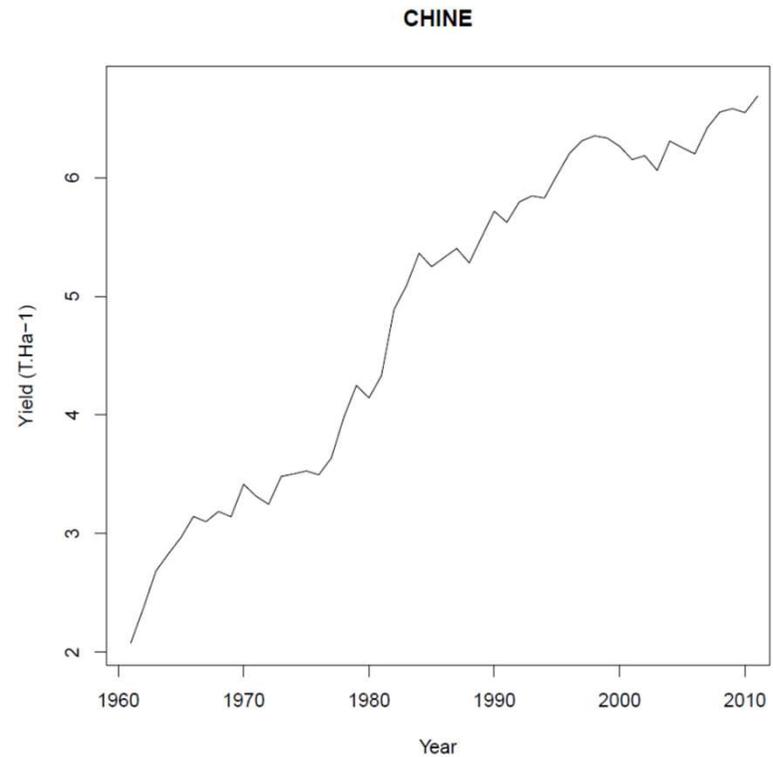
- Evolution au cours du temps d'un phénomène, dans le but de décrire, expliquer puis prédire ce phénomène dans le futur.

# Séries chronologiques de rendement



Evolution du rendement  
du blé tendre dans l'Yonne

Source : SSP - Agreste



Evolution du rendement  
du riz en Chine

Source : FAO - FAOSTAT

# Contexte: Comment?

## Modèles statistiques

- Modèle statistique permet d'estimer les tendances des séries chronologiques et prévoir les rendements futurs.

# Contexte

- **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.
  - Reflexion sur l'estimation de la demande mondiale future en céréales

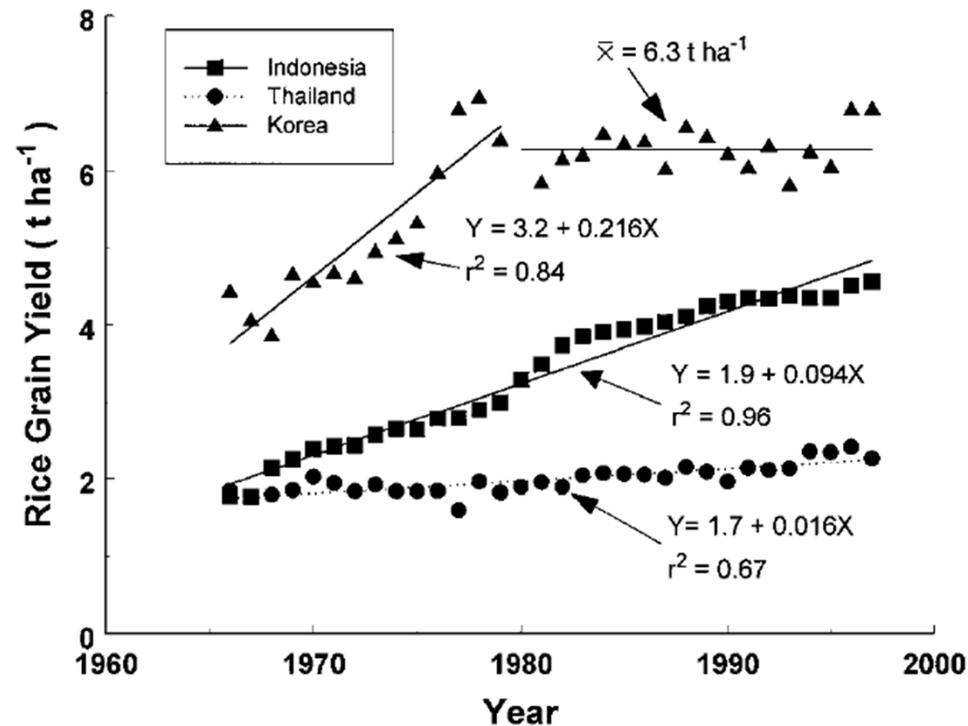
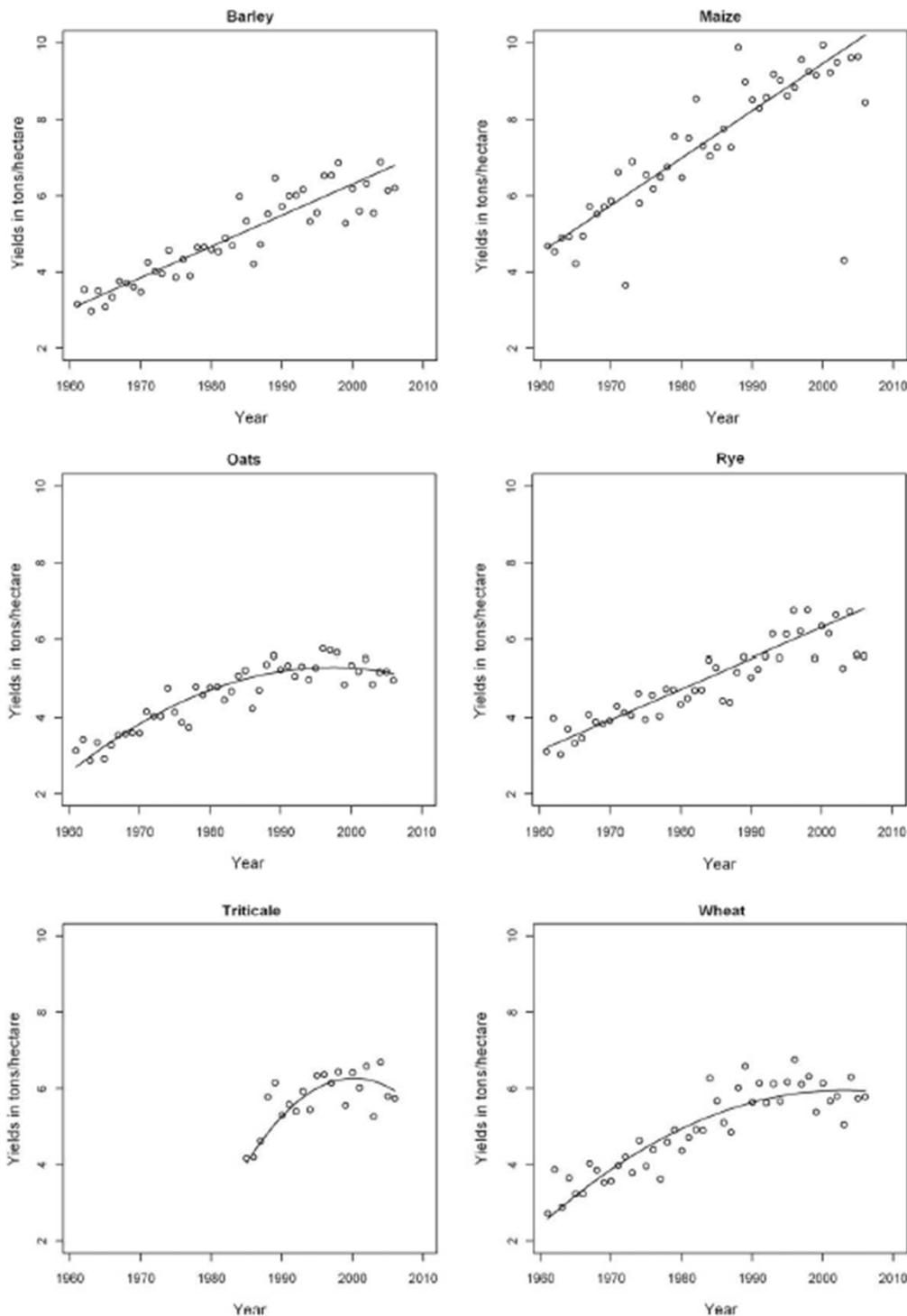


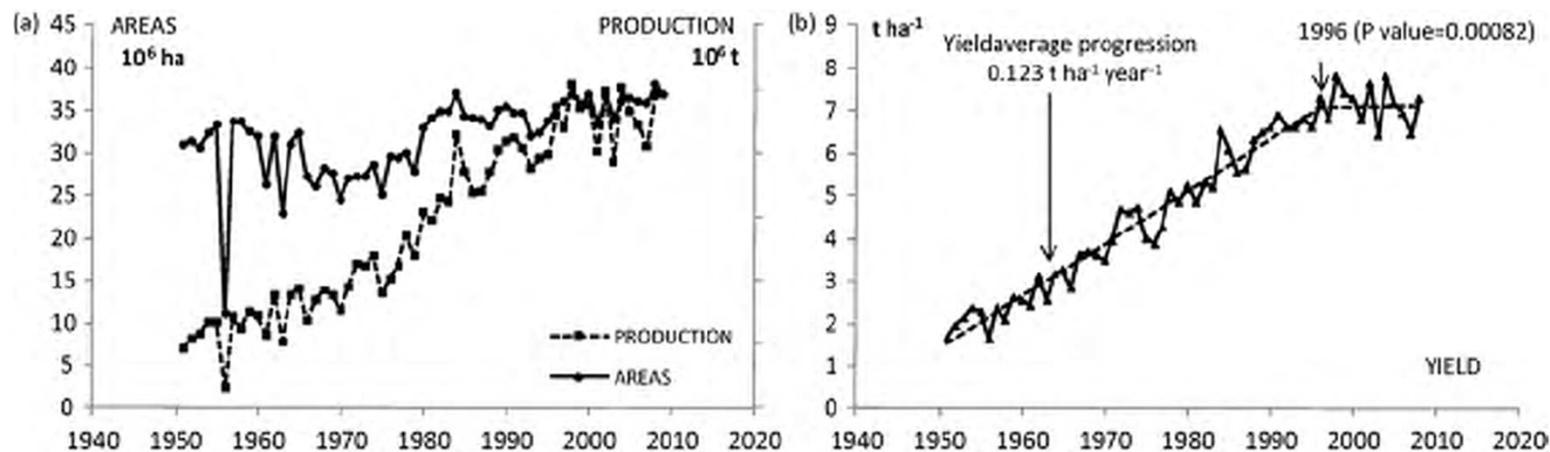
FIG. 1. National average rice yields from 1967 to 1997 in three Asian countries (<http://apps.fao.org>).



**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).

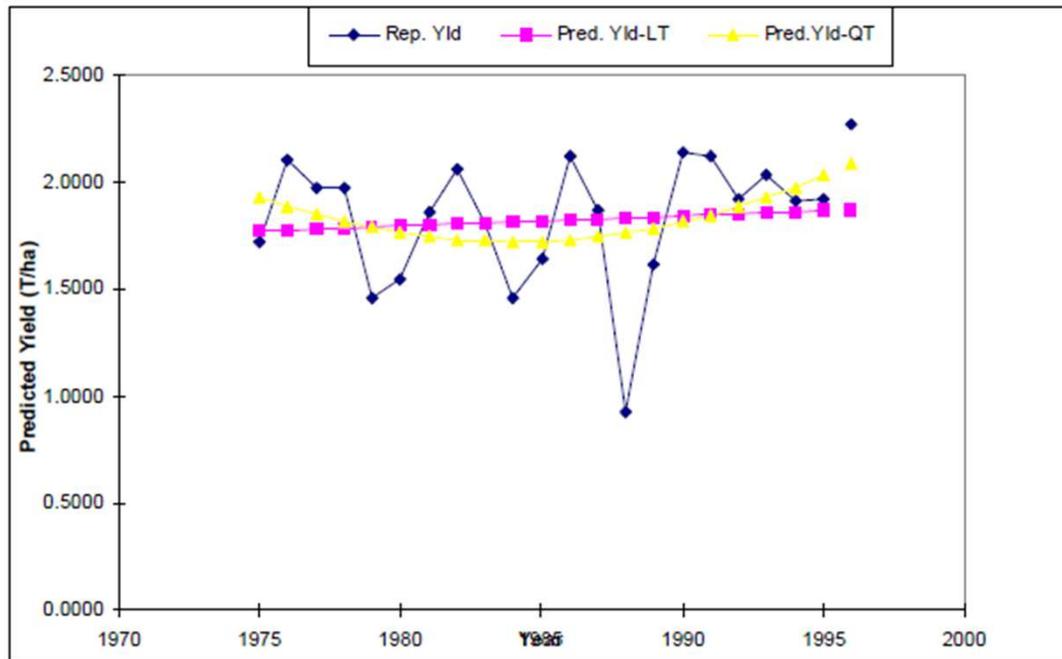
- **FINGER, R., 2010.**  
Evidence of slowing yield growth – The example of Swiss cereal yields. Food policy 35. p. 175-182.
- Etude de plusieurs cultures (orge, maïs, avoine, seigle, triticale, blé) en Suisse.
- Modèle linéaire et quadratique

- **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.**
  - Modèles régression linéaire, linéaire + plateau
  - Mise en évidence d'une stagnation à l'échelle de la France
  - Recherche de l'origine de cette stagnation (génétique, agronomique, climatique)



**Fig. 1.** Annual evolution of bread wheat areas and production in France (a) and the subsequent yield (b). Optimization of the rising-plateau statistical model gives 1996 as the year of inflexion. 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.



- Analyse rendement du blé de printemps dans Canadian Prairies entre 1975 et 1996
- Test de différents modèles via MSE

**Table 5. Evaluating the performance of various techniques to forecast spring wheat yield for Saskatchewan, during the model testing period.**

Year	Reported yield t/ha	Forecasted yield, ton/ha					
		Nonstationary series				Stationary series	
		Trend analysis		Double exponential smoothing	Double moving averaging	Simple exponential smoothing	Simple moving averaging
Linear trend	Quadratic trend						
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

# Contexte

Etude séries chronologiques de rendement de culture et de leur évolution

- Différents modèles statistiques utilisés
- Augmentation / Stagnation
- Origine de l'évolution

# Objectifs

1. Comparaison de plusieurs méthodes pour estimer les évolutions de rendements à différentes échelles.
2. Utilisation de la ou les meilleures méthodes pour analyser l'évolution des rendements de blé à l'échelle départementale (France) et mondiale.

# Matériels et méthodes

# Données utilisées

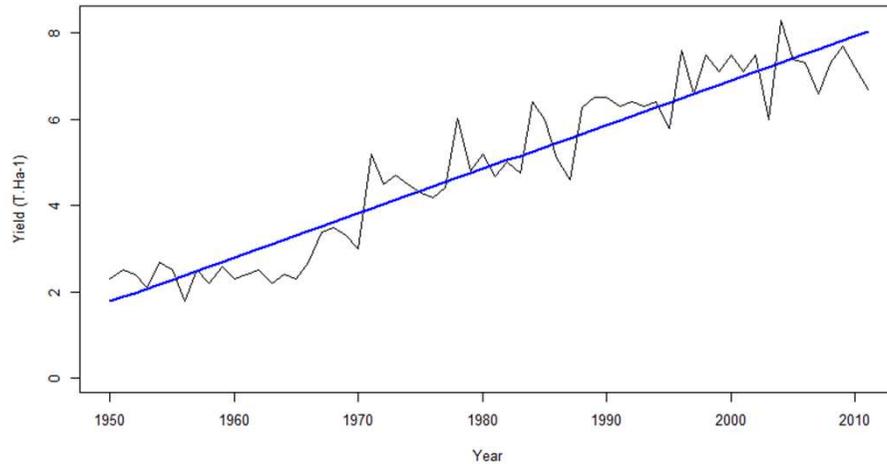
- Séries chronologiques de rendement de blé:
  - Mondiale:
    - FAO – FAOSTAT (1961 – 2010)
      - Dataset 1: World (120 pays)
  - Départementale et régionale (France):
    - SSP – Agreste (1950 - 2011)
      - Dataset 2: France (92 départements)
      - Dataset 3: France restricted (59 départements)

# Modèles utilisés

- Linéaire (L)
- Linéaire + plateau (LP)
- Quadratique (Q)
- Cubique (C)
- Lissage exponentiel (méthode Holt-Winters: HW)
- Linéaire dynamique (DLM)

# Linéaire / Linéaire + plateau

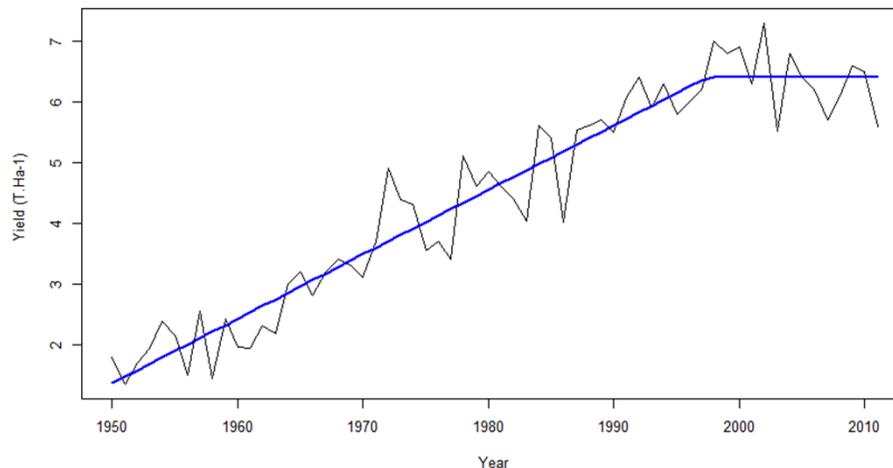
- Modèle linéaire (L)



$$Yield = a + b * Year + e$$

$$e \sim N(0, \sigma^2)$$

- Modèle Linéaire + plateau (LP)



Pour Year < Tmax:

$$Yield = Y_{max} + P * (Year - T_{max}) + e$$

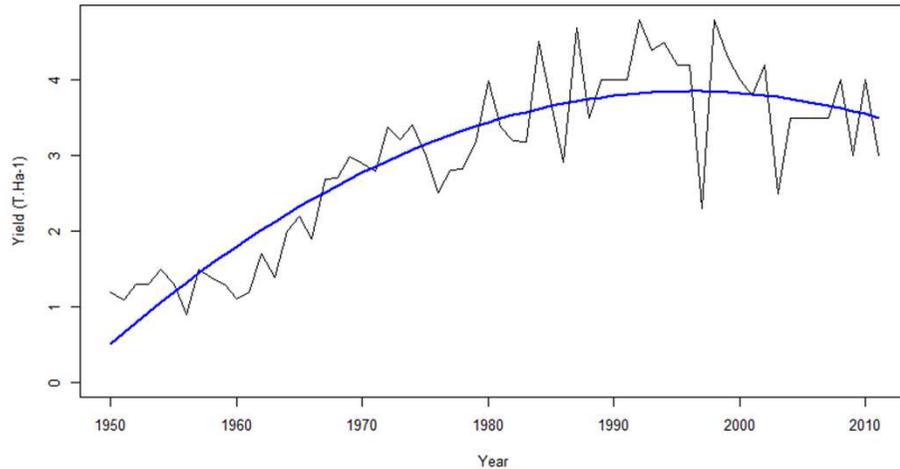
Pour Year > Tmax:

$$Yield = Y_{max} + e$$

$$e \sim N(0, \sigma^2)$$

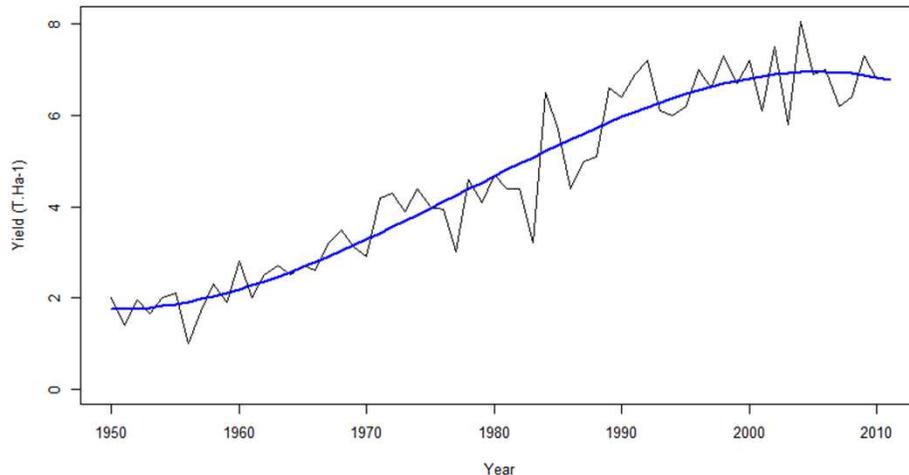
# Quadratique / Cubique

- Modèle quadratique (Q)



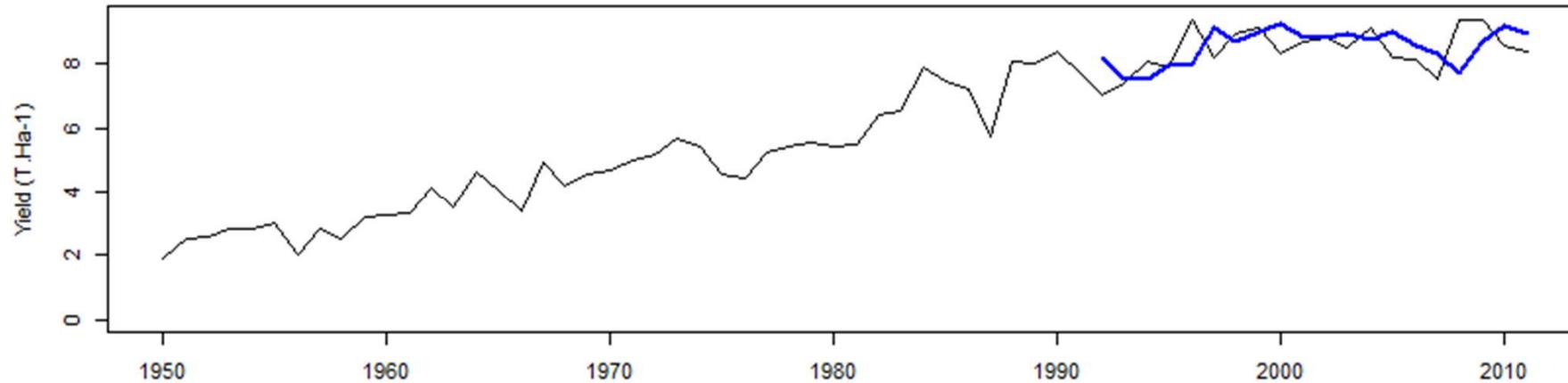
$$Yield = a + b * Year + c * Year^2 + e$$
$$e \sim N(0, \sigma^2)$$

- Modèle cubique (C)



$$Yield = a + b * Year + c * Year^2 + d * Year^3 + e$$
$$e \sim N(0, \sigma^2)$$

# Holt-Winters (HW)

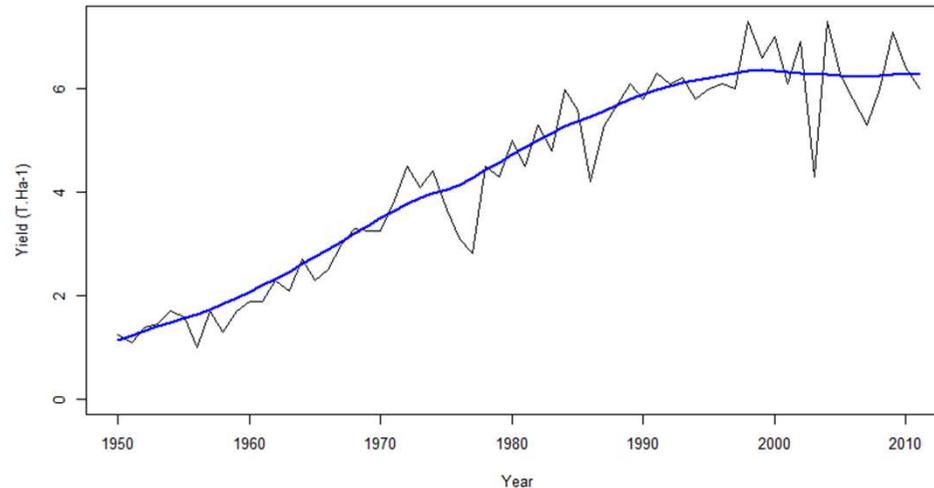


## Equation :

$$Y_{t+\Delta k} = m_t + b_t * \Delta t \quad \text{avec}$$

- $m_t = \lambda_0 * y_t + (1 - \lambda_0) * Y_t$
- $\Delta t = \text{écart de temps entre 2 rendements successifs}$
- $b_t = \lambda_1 * (m_t - m_{t-1}) + (1 - \lambda_1) * b_{t-1}$

# Dynamic Linear Model (DLM)



## Equation 1 (évolution de rendement):

$$Y_{t+\Delta t} = m_t + b_t * \Delta t + \varepsilon_t \quad \text{avec}$$

- $m_t = \text{intercept}$
- $b_t = \text{pente}$
- $\Delta t = \text{écart de temps entre 2 rendements successifs}$
- $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$

## Equation 2 (pente):

$$\theta_t = \theta_{t-1} + \eta_{t-1} \quad \text{avec } \theta_t = \begin{pmatrix} m_t \\ b_t \end{pmatrix}, \eta_{t-1} \sim N(0, \Sigma) \quad \text{et } \Sigma = \begin{pmatrix} \sigma_m^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix}$$

# Dynamic Linear Model (DLM)

Filtering

$$E(\theta_t | y_1, \dots, y_t) \quad V(\theta_t | y_1, \dots, y_t)$$

Smoothing

$$E(\theta_{t-\Delta t} | y_1, \dots, y_t) \quad V(\theta_{t-\Delta t} | y_1, \dots, y_t)$$

Predictions

$$E(\theta_{t+\Delta t} | y_1, \dots, y_t) \quad V(\theta_{t+\Delta t} | y_1, \dots, y_t)$$

# Dynamic Linear Model (DLM)

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

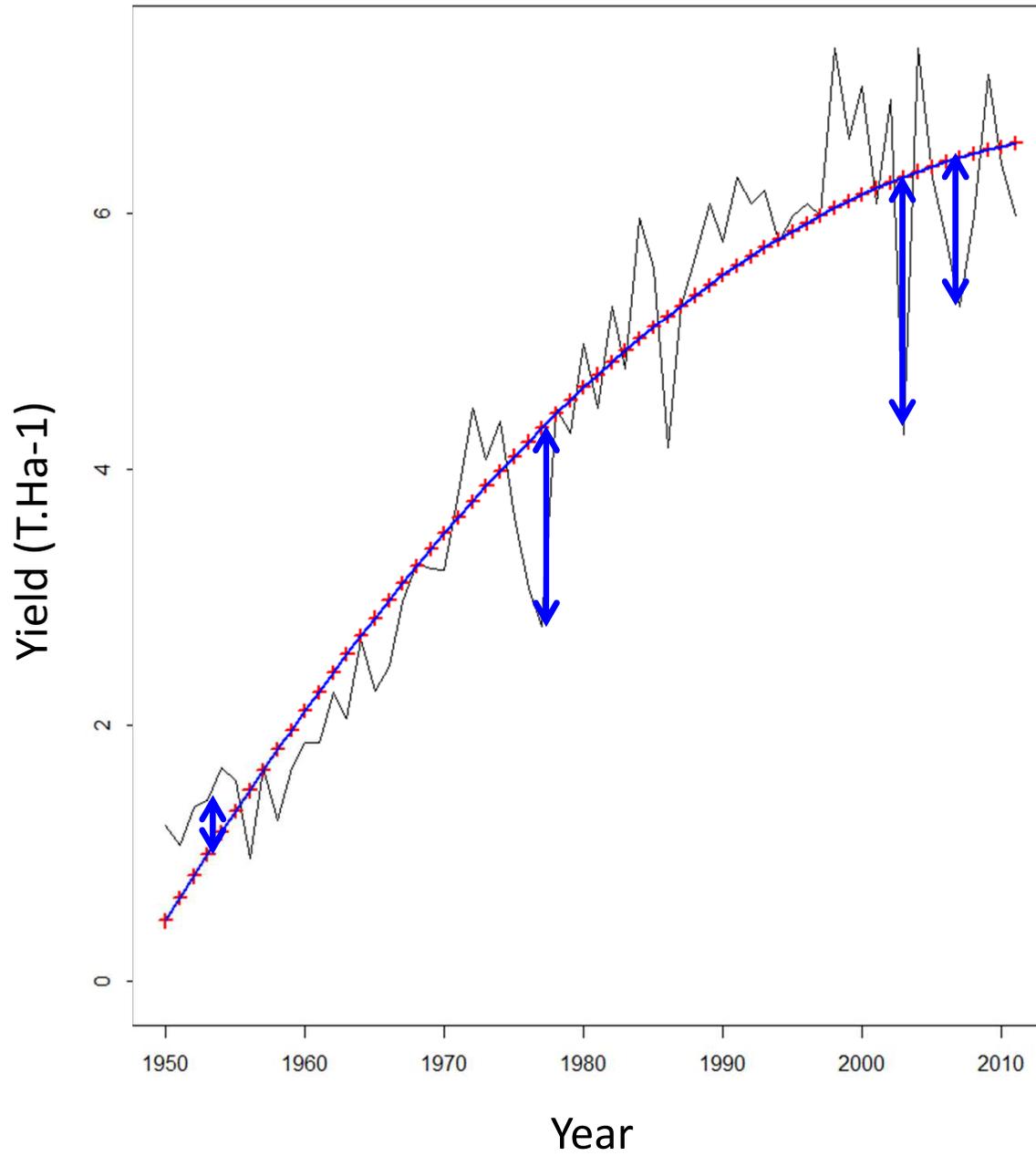
$$K = \text{Var}(\theta_t | y_1, \dots, y_{t-1}) F^T [F \text{Var}(\theta_t | y_1, \dots, y_{t-1}) F^T + V]^{-1}$$

# Choix du/des meilleur(s) modèle(s)

- Utilisation des critères RMSE et RMSEP
  - RMSE = Root Mean Square Error
  - RMSEP = Root Mean Square Error of Prediction
- Critères permettant d'évaluer les erreurs des estimations (RMSE) ou des prédictions (RMSEP) de rendement pour différents modèles.  
Quantification de l'écart entre le modèle et les mesures.
- $RMSE/P$  = racine carré de la moyenne arithmétique des carrés des écarts entre les estimations/prédictions et les observations
- Meilleur modèle = RMSE/RMSEP le plus petit !!!

# RMSE/RMSEP

# NIEVRE



# RMSE/RMSEP

- $RMSE = \sqrt{\frac{1}{T_1} \sum_{t=1}^{T_1} (y_t - Y_t)^2}$
- $RMSEP = \sqrt{\frac{1}{T_2} \sum_{t=1}^{T_2} (y_t - Y_{Pt})^2}$

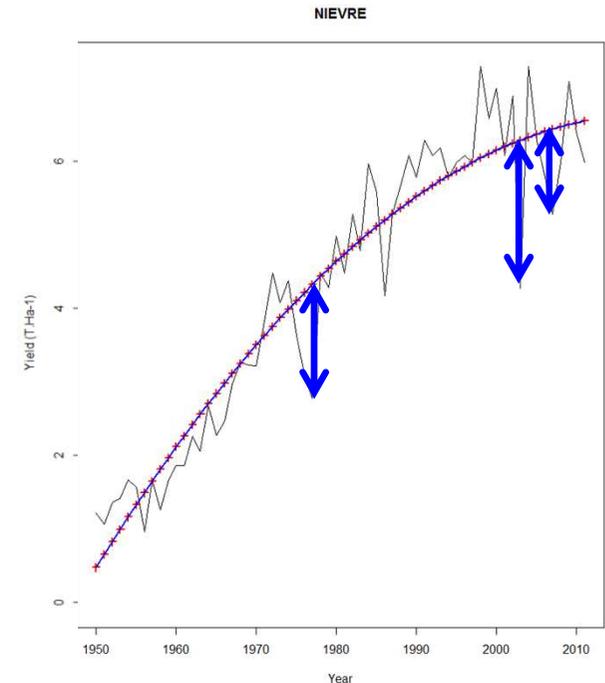
Avec:

- $y_t$ : observation
- $Y_t$ : estimation
- $Y_{Pt}$ : prédiction

$T_1 = 61$  ans pour les départements

$T_1 = 49$  ans pour les pays

$T_2 = 20$  ans (1992-2011)



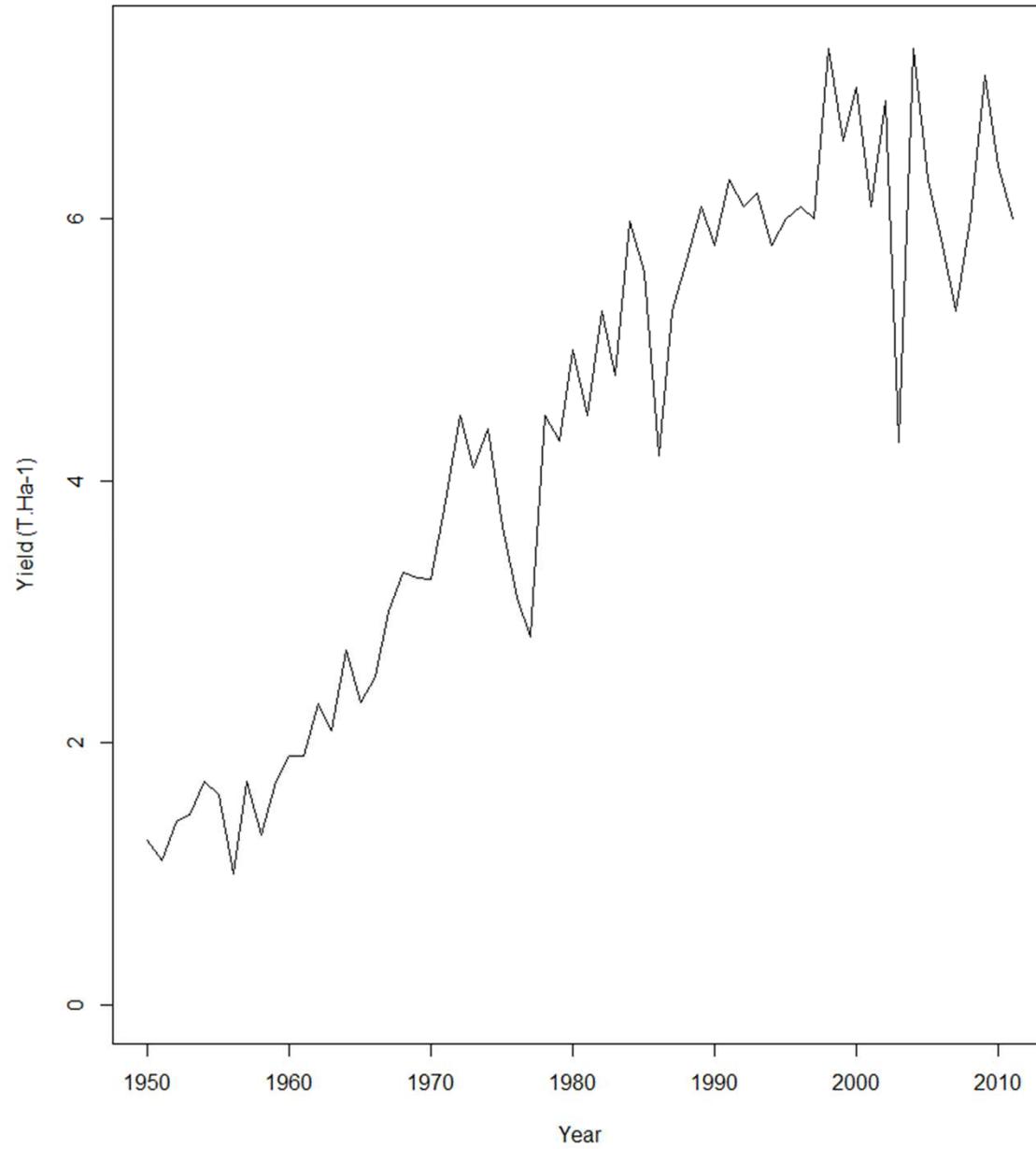
- Variables d'entrée pour les calculs:
  - Rendements observés
  - Rendements estimés (RMSE)
  - Rendements prédits par validation croisée (RMSEP)

# RMSE/RMSEP

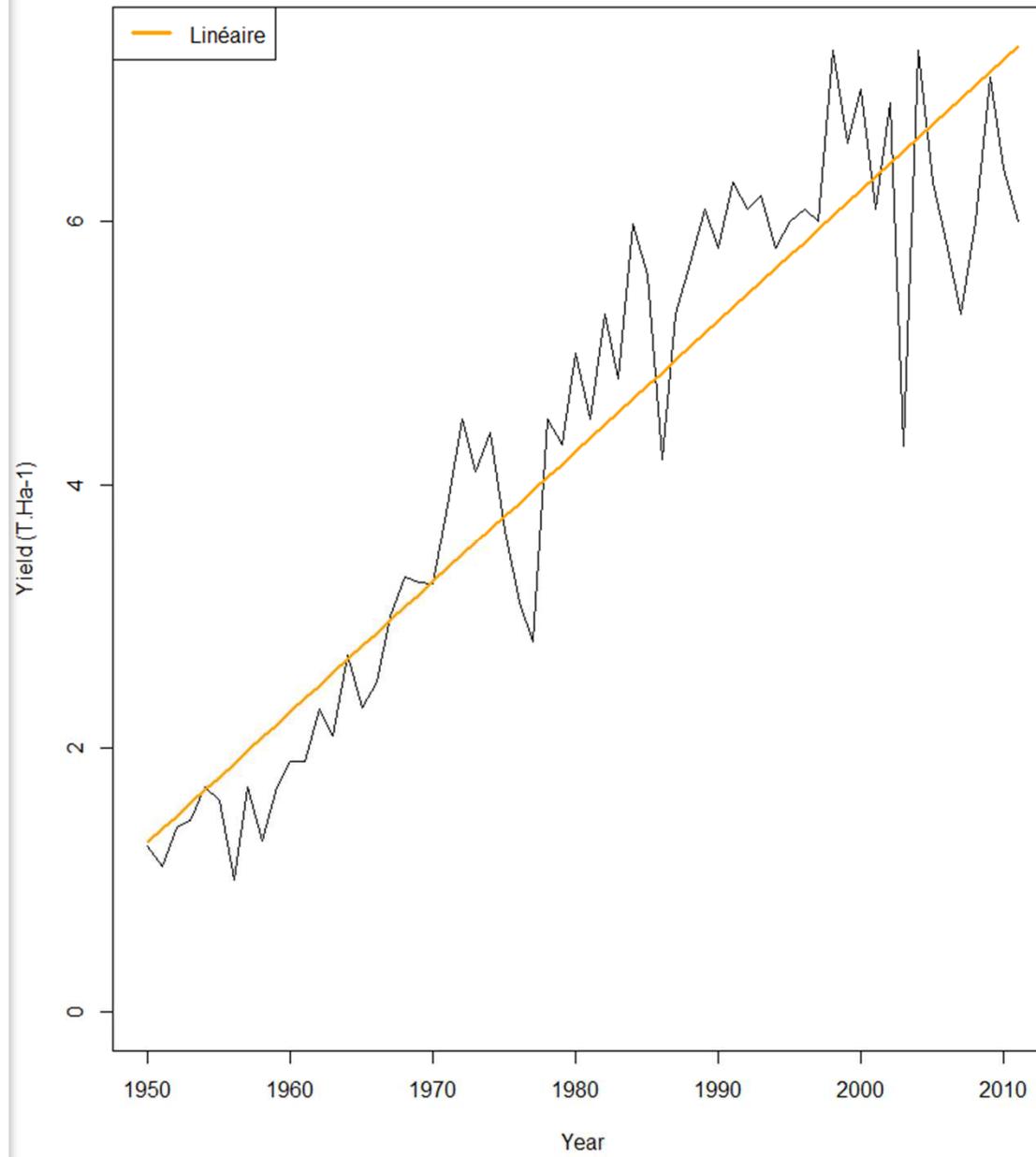
- RMSEP estimé par validation croisée
- Validation croisée: méthode utilisée pour sélectionner un modèle sur les données.
  - Découpage des données en deux sous-échantillons.
    - Premier échantillon utilisé pour estimer les paramètres du modèle
    - Second échantillon utilisé pour évaluer les prédictions en les comparant aux valeurs observées.
- Prédiction de 1992 -> 2011
  - RMSEP 1 an:
    - Estimation avec l'échantillon <1992 pour prédire à 1992
    - Estimation avec l'échantillon <1993 pour prédire à 1993
    - ...
    - Estimation avec l'échantillon <2011 pour prédire à 2011
  - RMSEP 10 ans:
    - Estimation avec l'échantillon <1992 pour prédire de 1992 à 2001
    - Estimation avec l'échantillon <2002 pour prédire de 2002 à 2011

# Ajustement des modèles

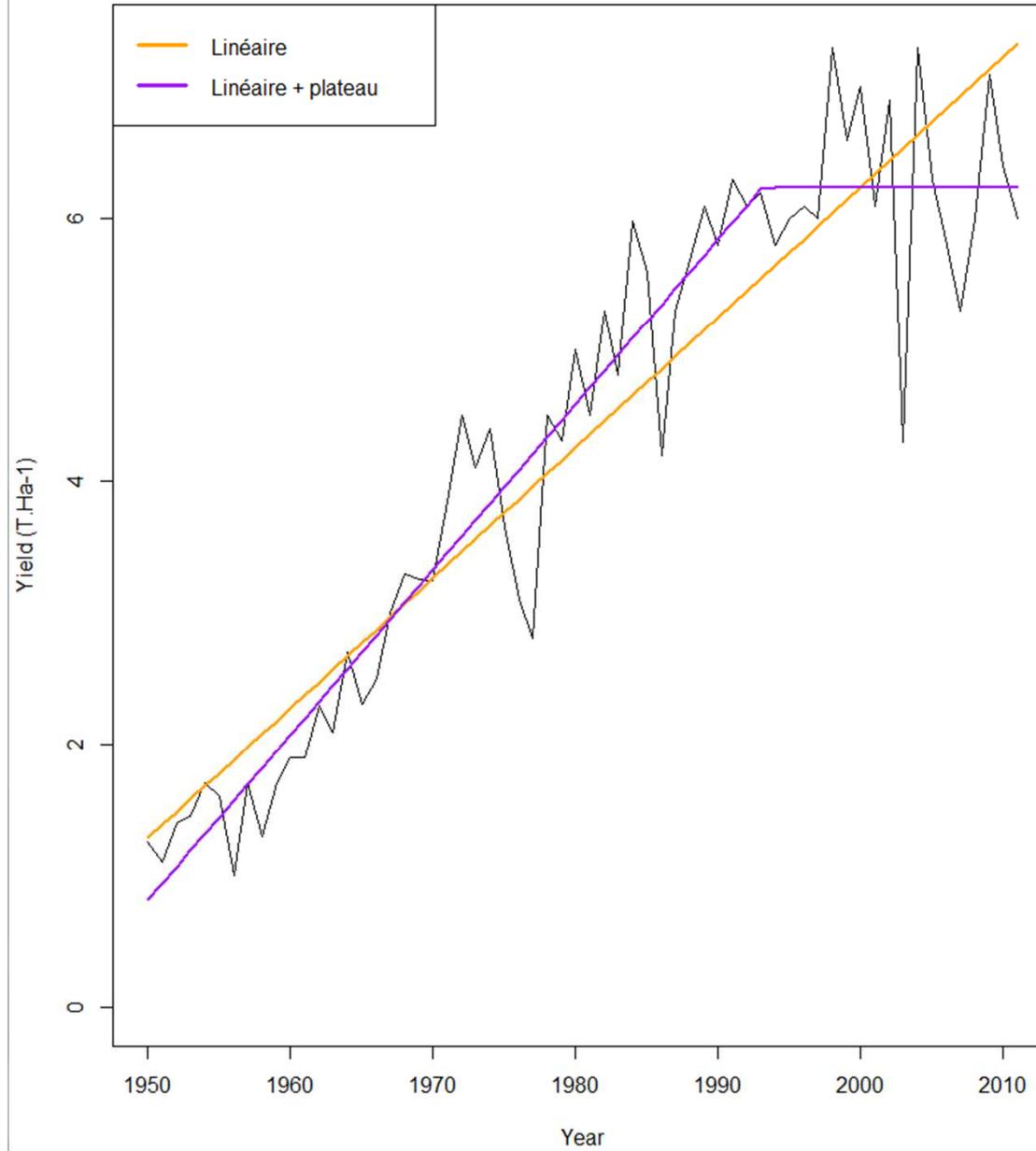
# NIEVRE



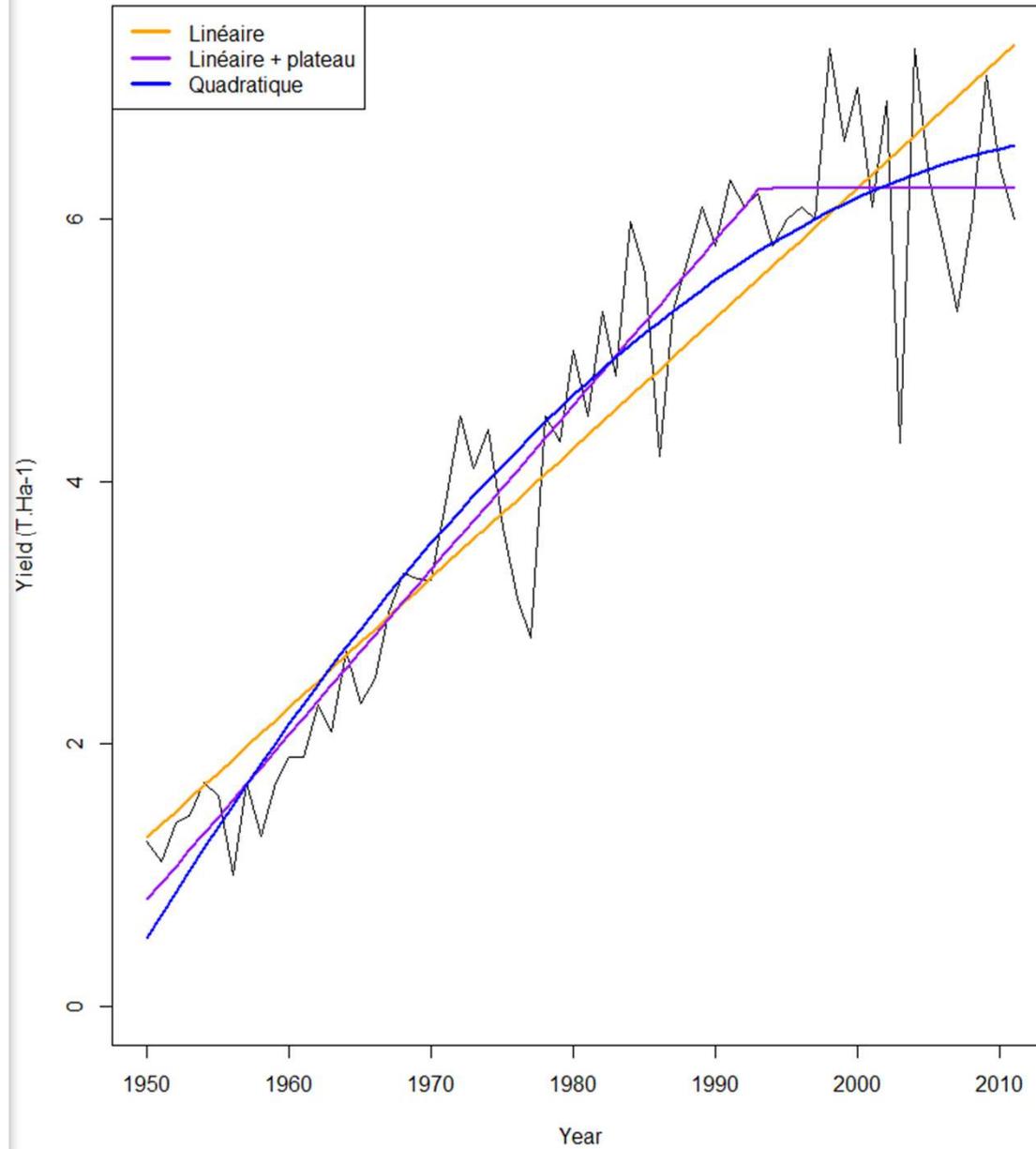
# NIEVRE



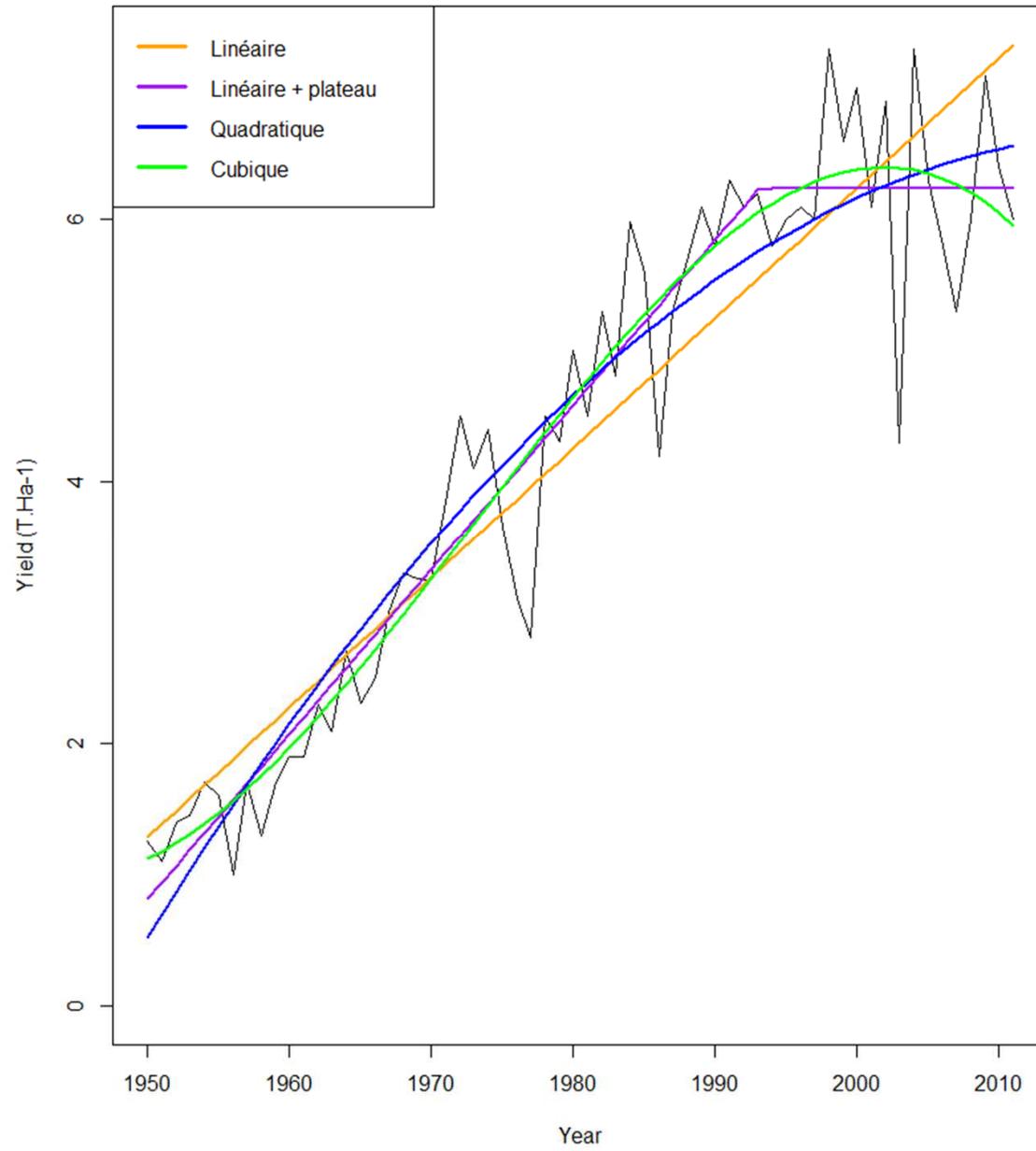
# NIEVRE



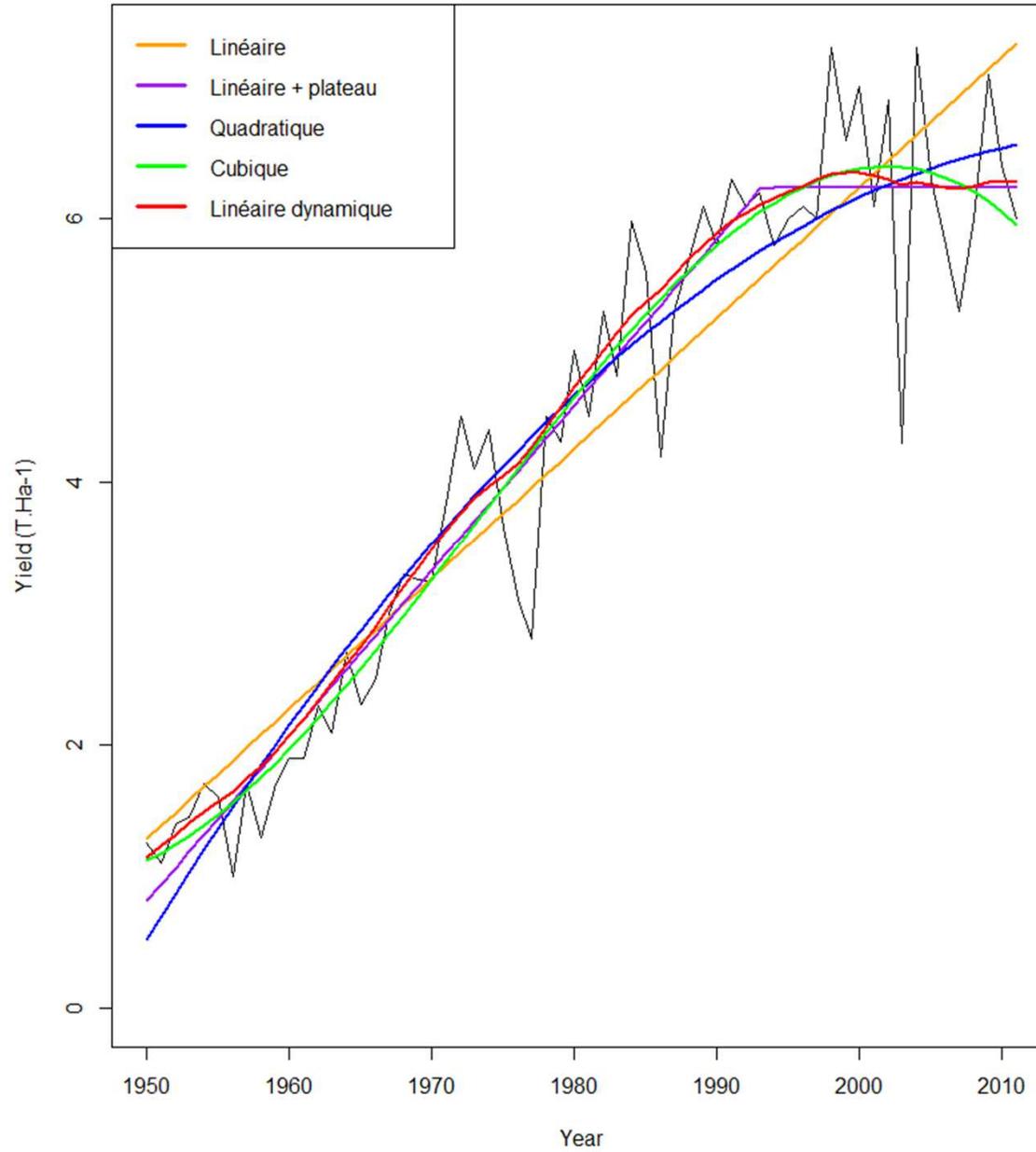
# NIEVRE



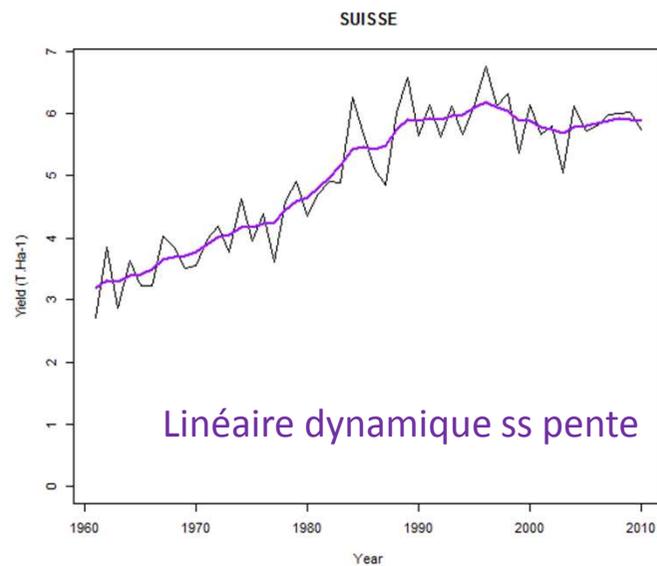
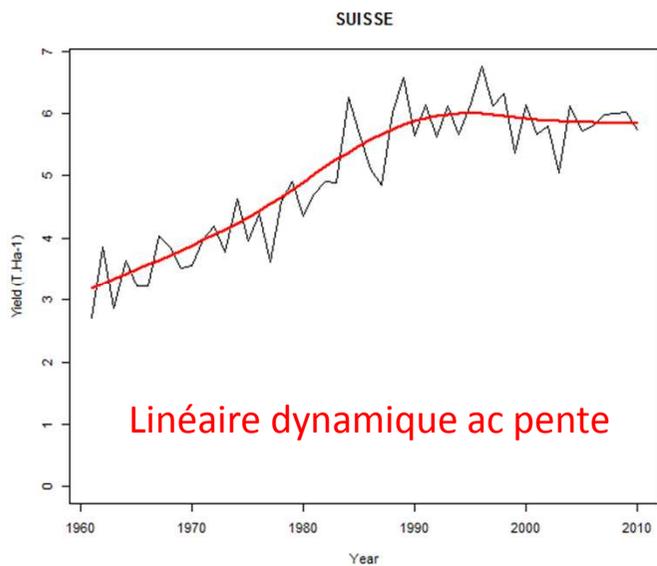
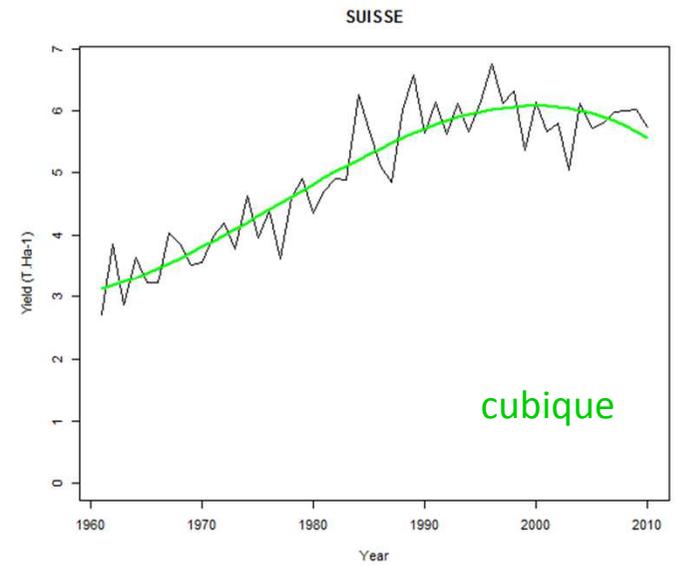
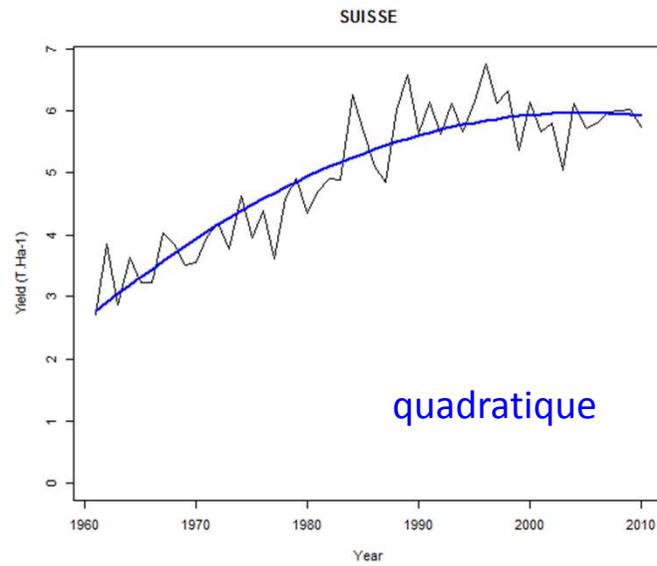
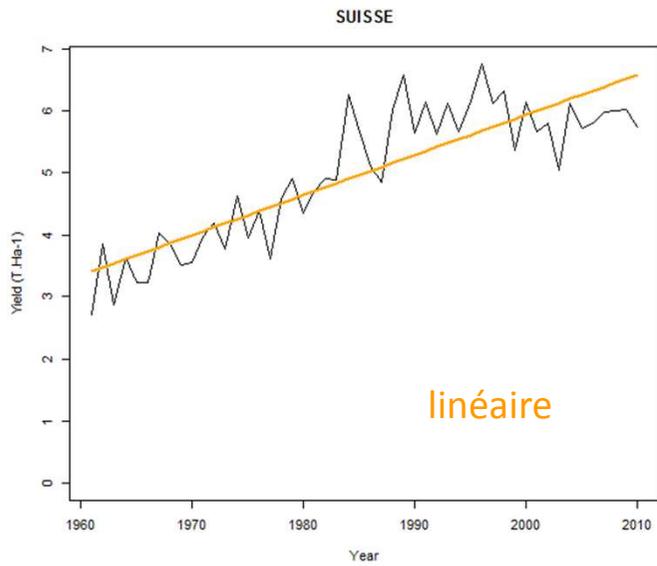
# NIEVRE



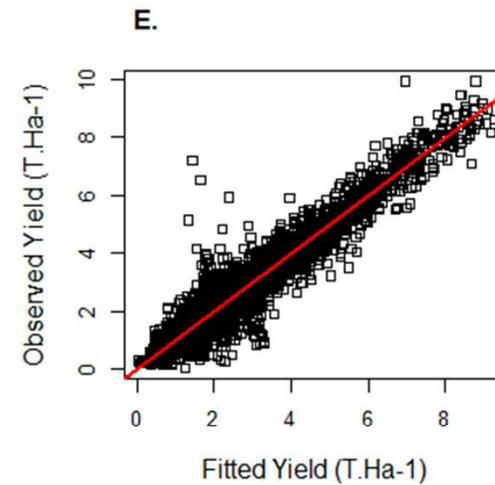
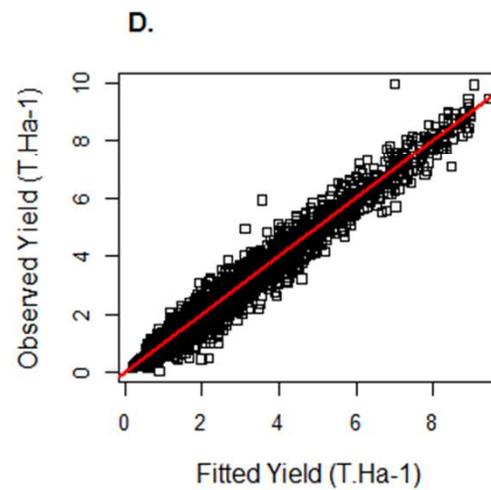
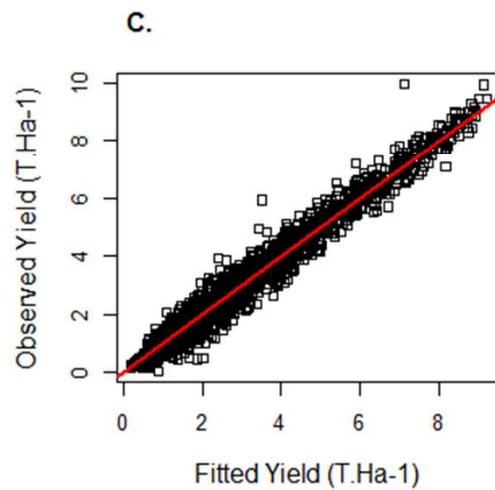
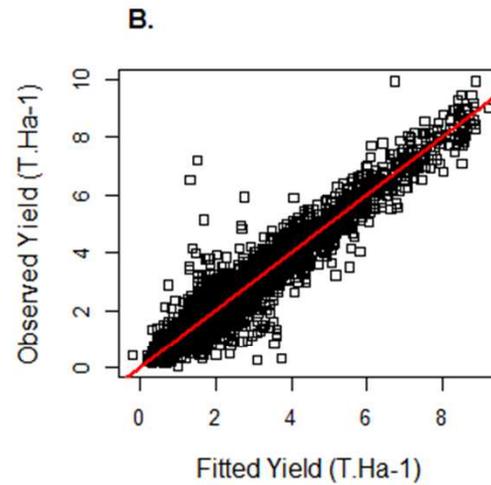
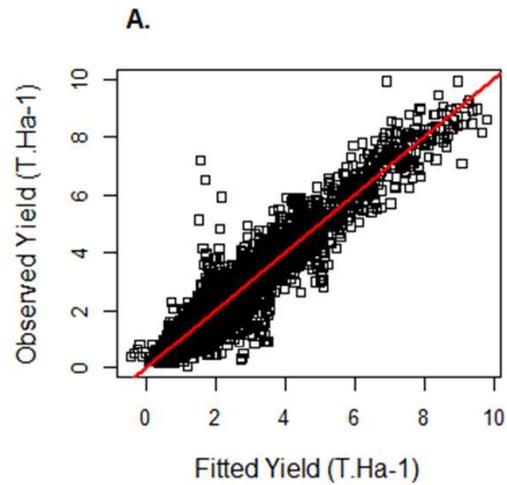
# NIEVRE



# Choix du modèle: RMSE



- A. Linéaire
- B. Cubique
- C. DLM0
- D. DLMS
- E. Quadratique

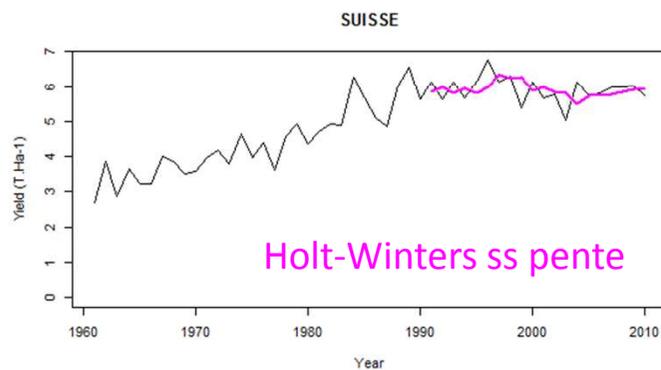
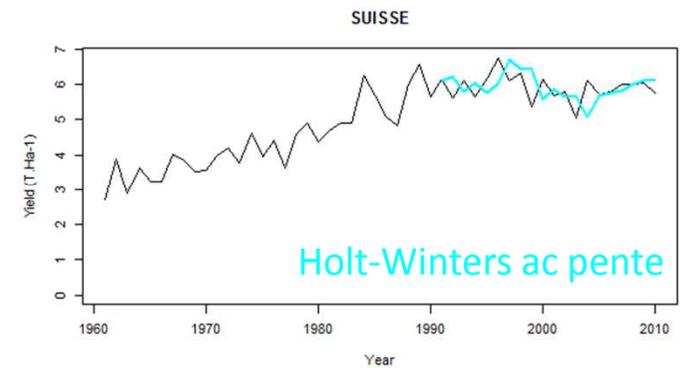
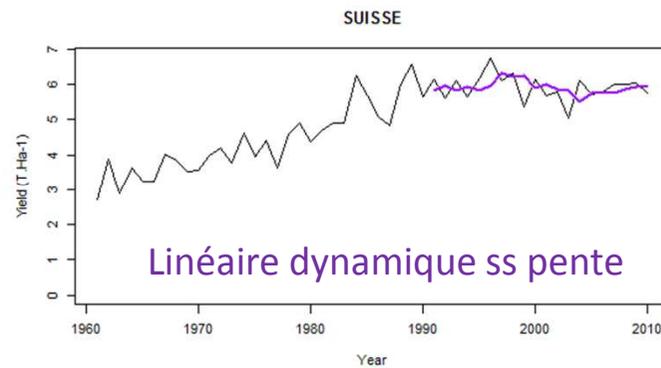
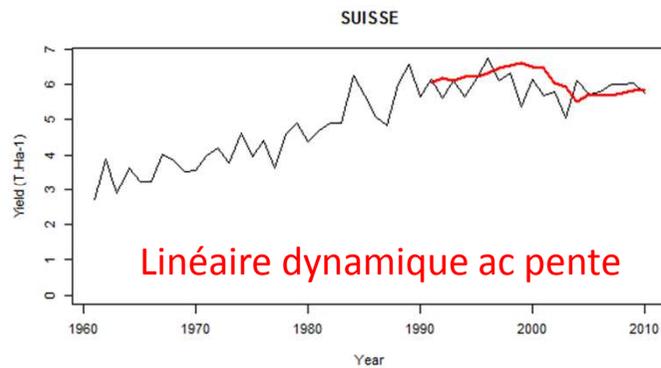
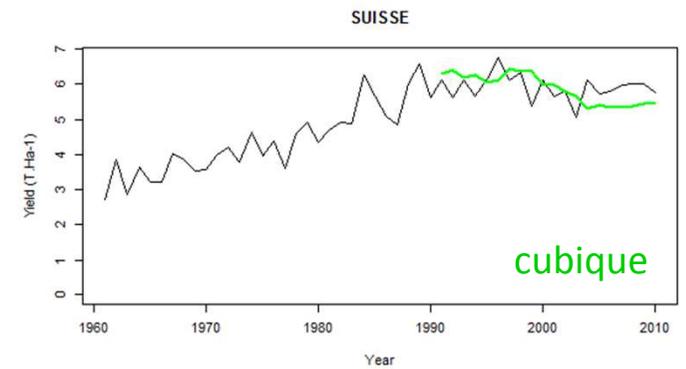
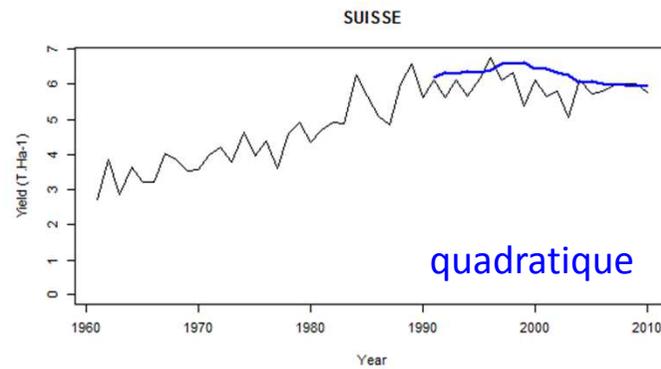
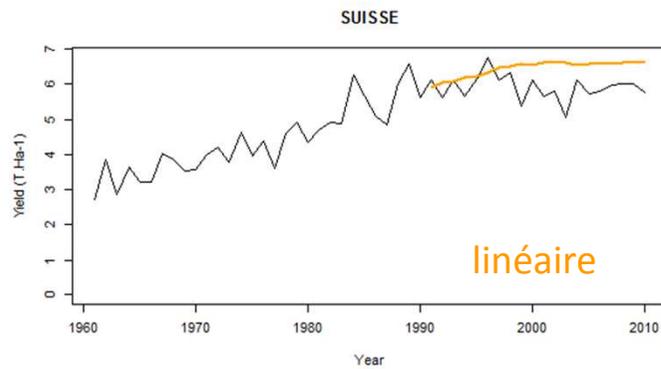


# Choix du modèle: RMSE

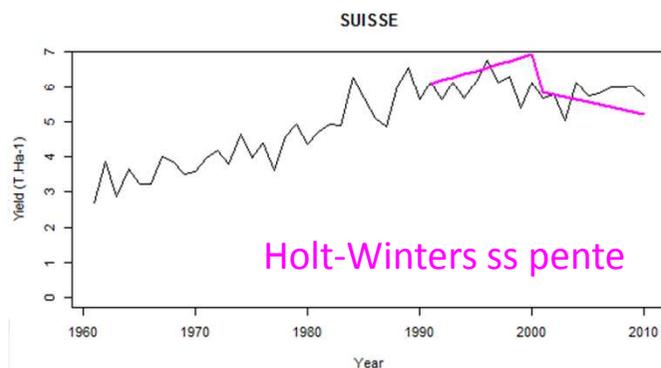
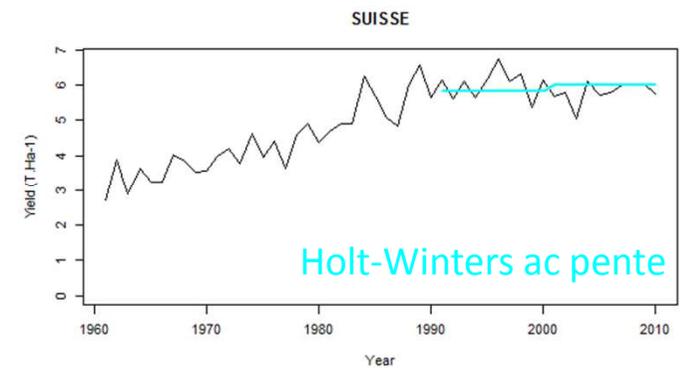
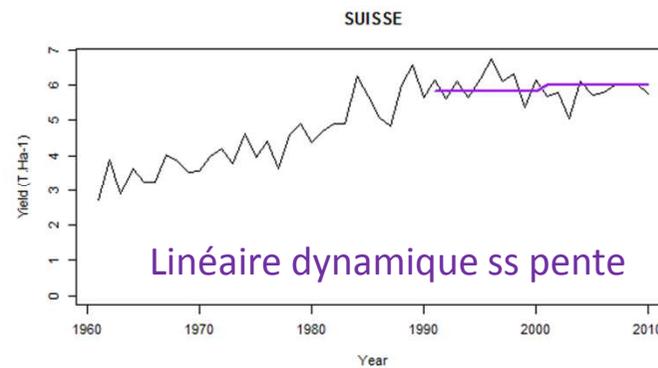
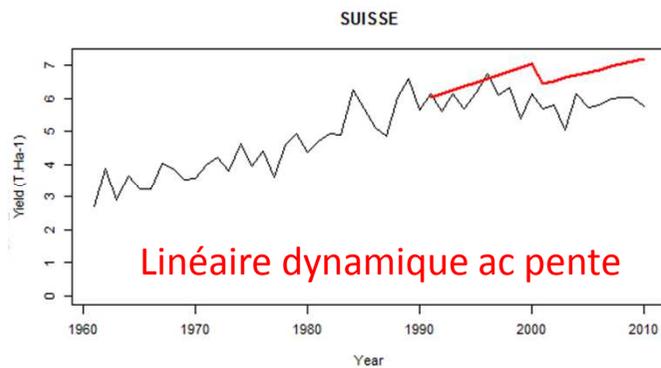
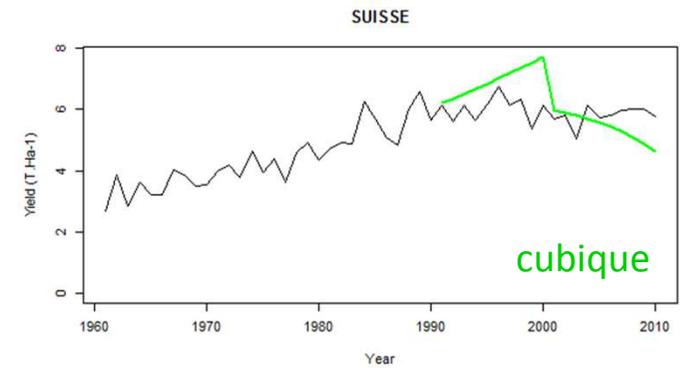
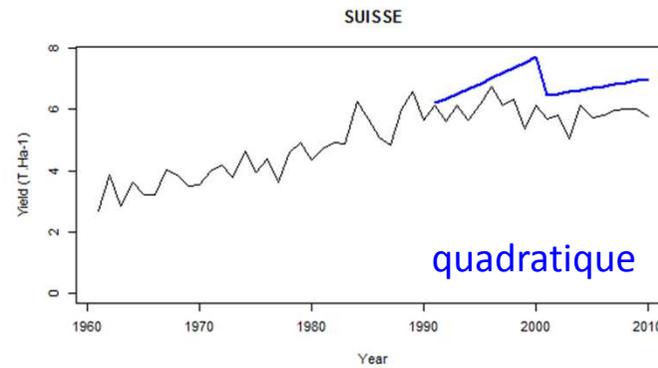
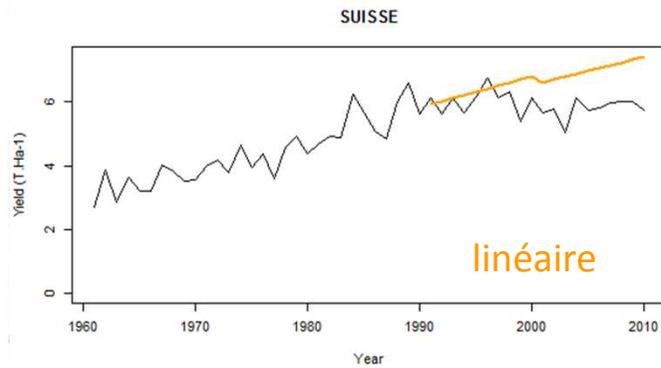
- RMSE

Scale	Unit	L	C	DLMs	DLM0	LP	
<i>France</i>	T.Ha-1	0.59	0.54	0.49	0.47	<b>0.38</b>	NA
	%	54.11	40.69	28.72	23.05	<b>0.00</b>	NA
<i>France (restricted)</i>	T.Ha-1	0.58	0.53	0.49	0.47	<b>0.38</b>	0.4966
	%	53.61	40.03	27.95	22.73	<b>0.00</b>	30.96
<i>World</i>	T.Ha-1	0.40	0.35	0.32	0.23	<b>0.20</b>	NA
	%	96.98	74.03	59.76	14.57	<b>0.00</b>	NA

# Choix du modèle: RMSEP 1an



# Choix du modèle: RMSEP 10ans



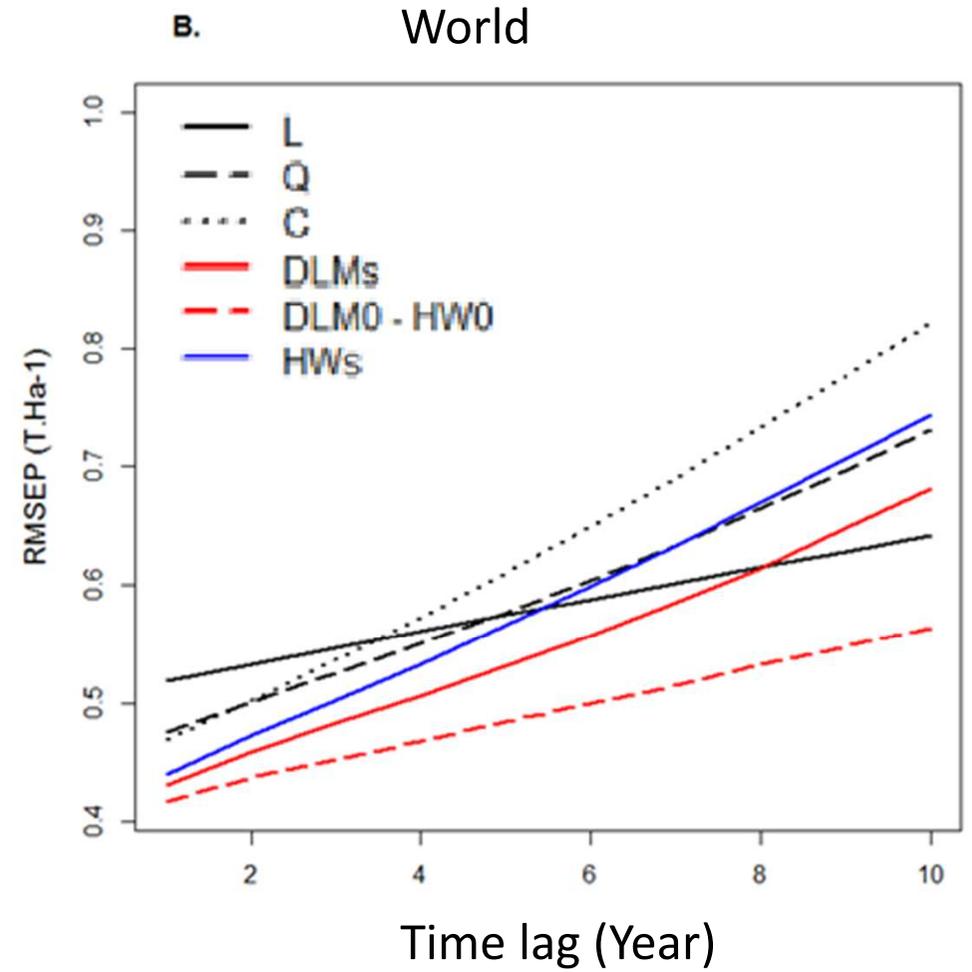
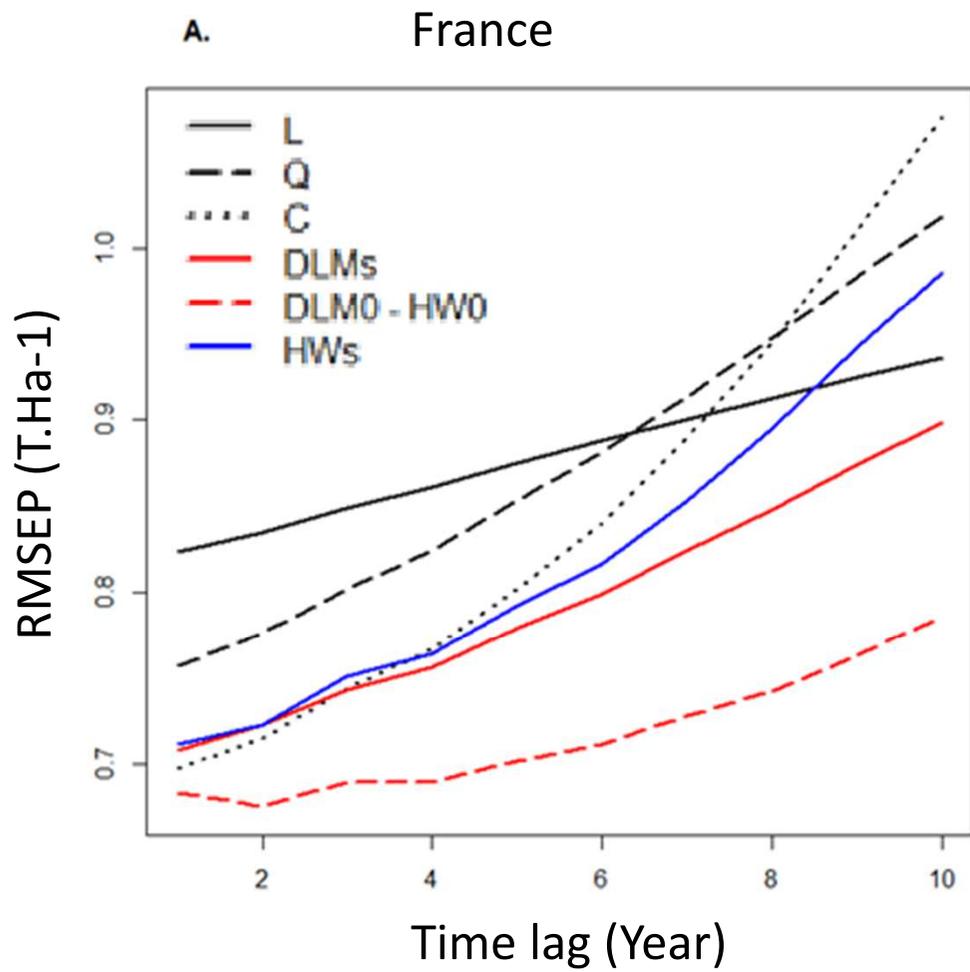
# Choix du/des meilleur(s) modèle(s)

- RMSEP1

Scale	Unit	L	Q	C	DLMs	DLM0	HWs	HW0
<i>France</i>	T.Ha-1	0.81	0.75	0.69	0.71	0.68	0.71	<b>0.68</b>
	%	18.73	10.49	1.64	3.42	0.09	3.98	<b>0.00</b>
<i>World</i>	T.Ha-1	0.52	0.48	0.48	0.43	0.42	0.44	<b>0.42</b>
	%	24.45	14.31	14.75	3.53	0.02	5.55	<b>0.00</b>

- RMSEP10

Scale	Unit	L	Q	C	DLMs	DLM0	HWs	HW0
<i>France</i>	T.Ha-1	0.94	1.07	1.06	0.93	<b>0.70</b>	0.97	0.70
	%	34.41	52.37	50.38	32.28	<b>0.00</b>	38.61	0.20
<i>World</i>	T.Ha-1	0.66	0.77	0.85	0.68	0.59	0.73	<b>0.59</b>
	%	11.61	30.66	43.74	15.02	0.07	23.69	<b>0.00</b>



# Conclusion: le meilleur modèle?

- Les meilleurs modèles dans l'absolue:
  - Linéaire dynamique sans pente (DLM0)
  - Holt-Winters sans pente (HW0)
- DLM vs HW:
  - Avantage DLM:
    - Distribution de probabilité
    - Modélisation de la variabilité interannuelle
    - Smoothing: estimation des tendances passées
- Le meilleur modèle permettant d'estimer des tendances:
  - Linéaire dynamique avec pente (DLMs)

# Conclusion

- Nombreux modèles possibles
  - Importance de comparer les erreurs
  - Contraste important entre modèles
- DLM: « outil puissant »
  - Estimation
  - Prédiction

# Fonctions R: outils graphique et cartographique

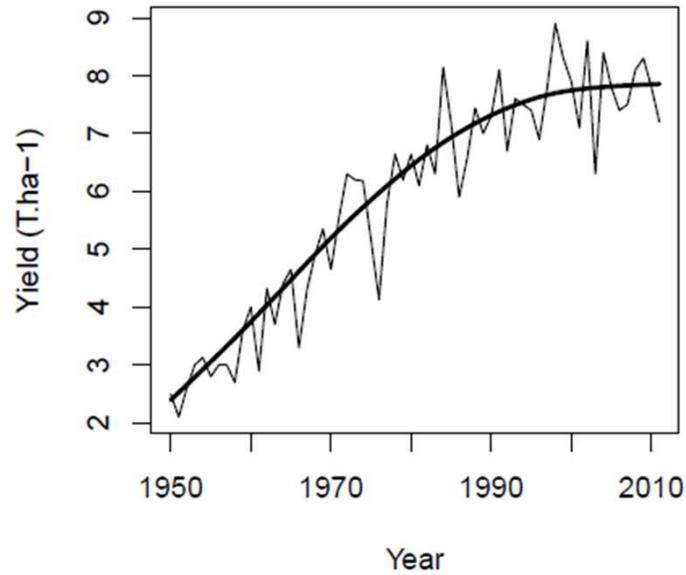
# Démarche

- Utilisation du modèle DLM0 et DLMs
- Création de graphiques:
  - Estimation et prédiction des rendements et des pentes
- Création de cartes annuelles:
  - Estimation et prédiction des rendements et des pentes du blé à l'échelle départementale et mondiale.

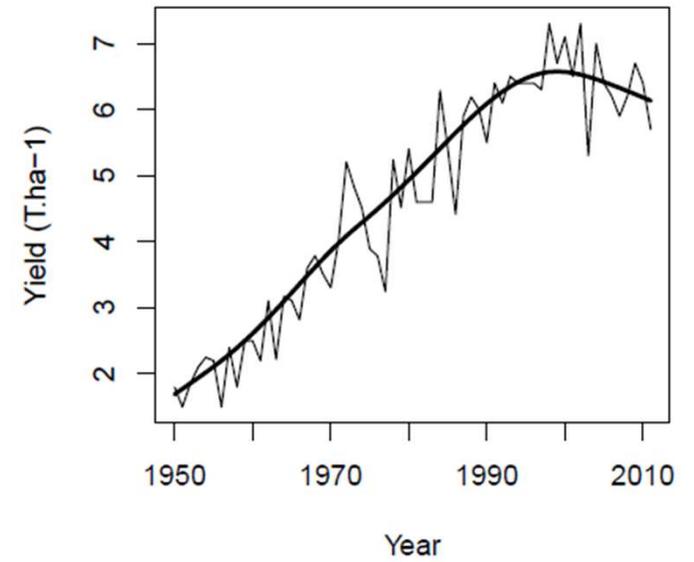
# Outil graphique

Rendement du blé.

**EURE-ET-LOIR**

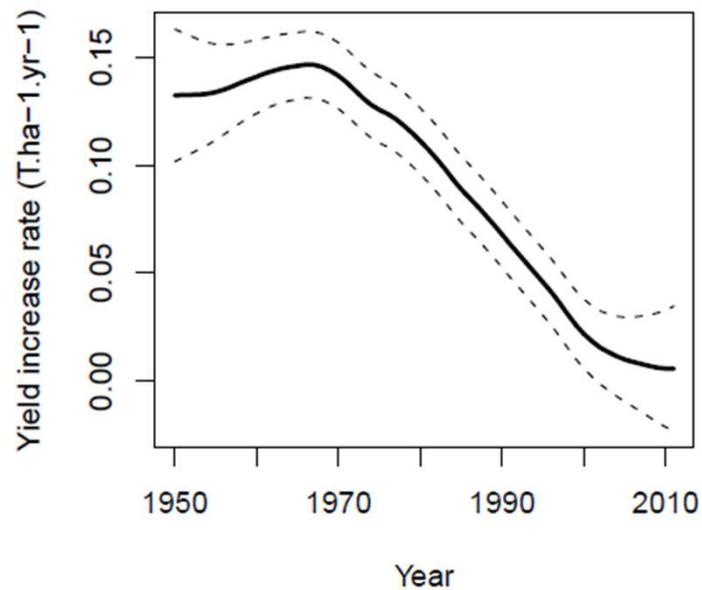


**CHER**

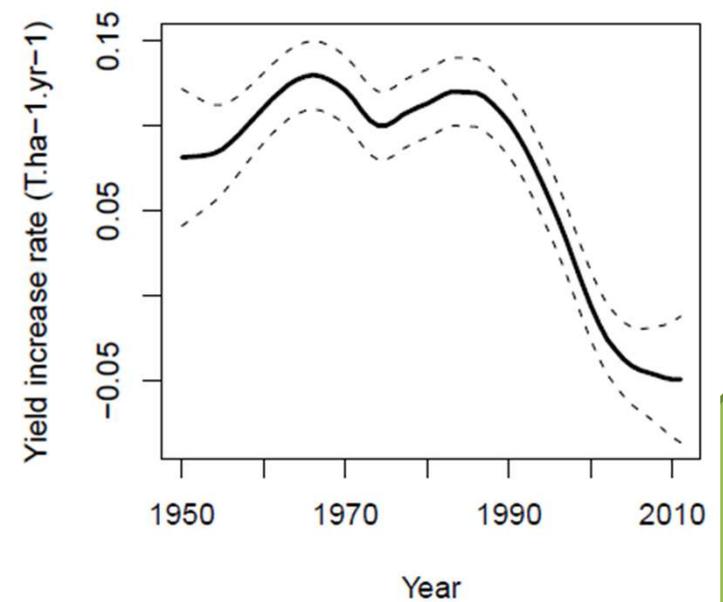


Accroissement de rendement du blé.

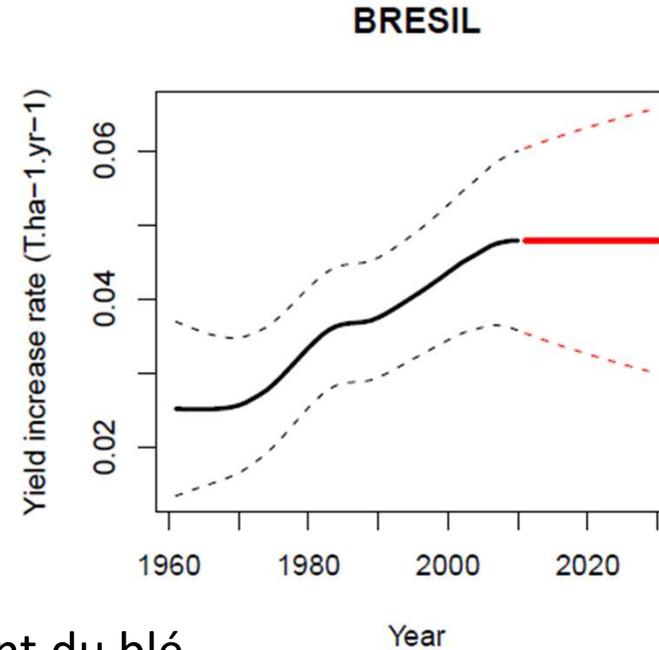
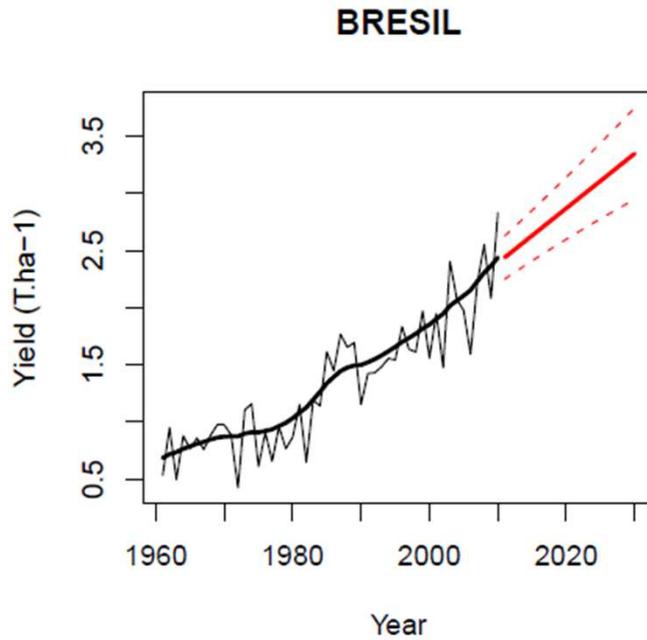
**EURE-ET-LOIR**



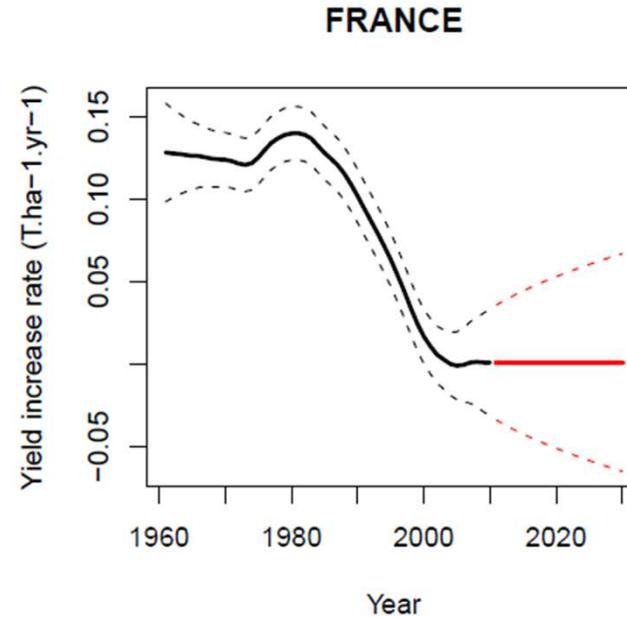
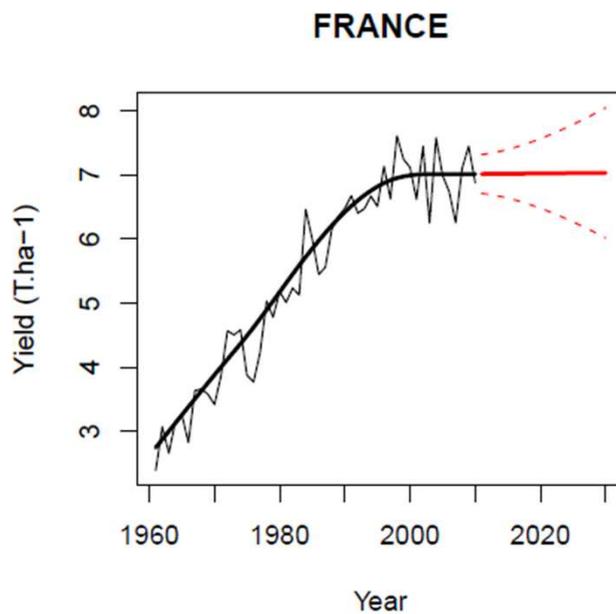
**CHER**

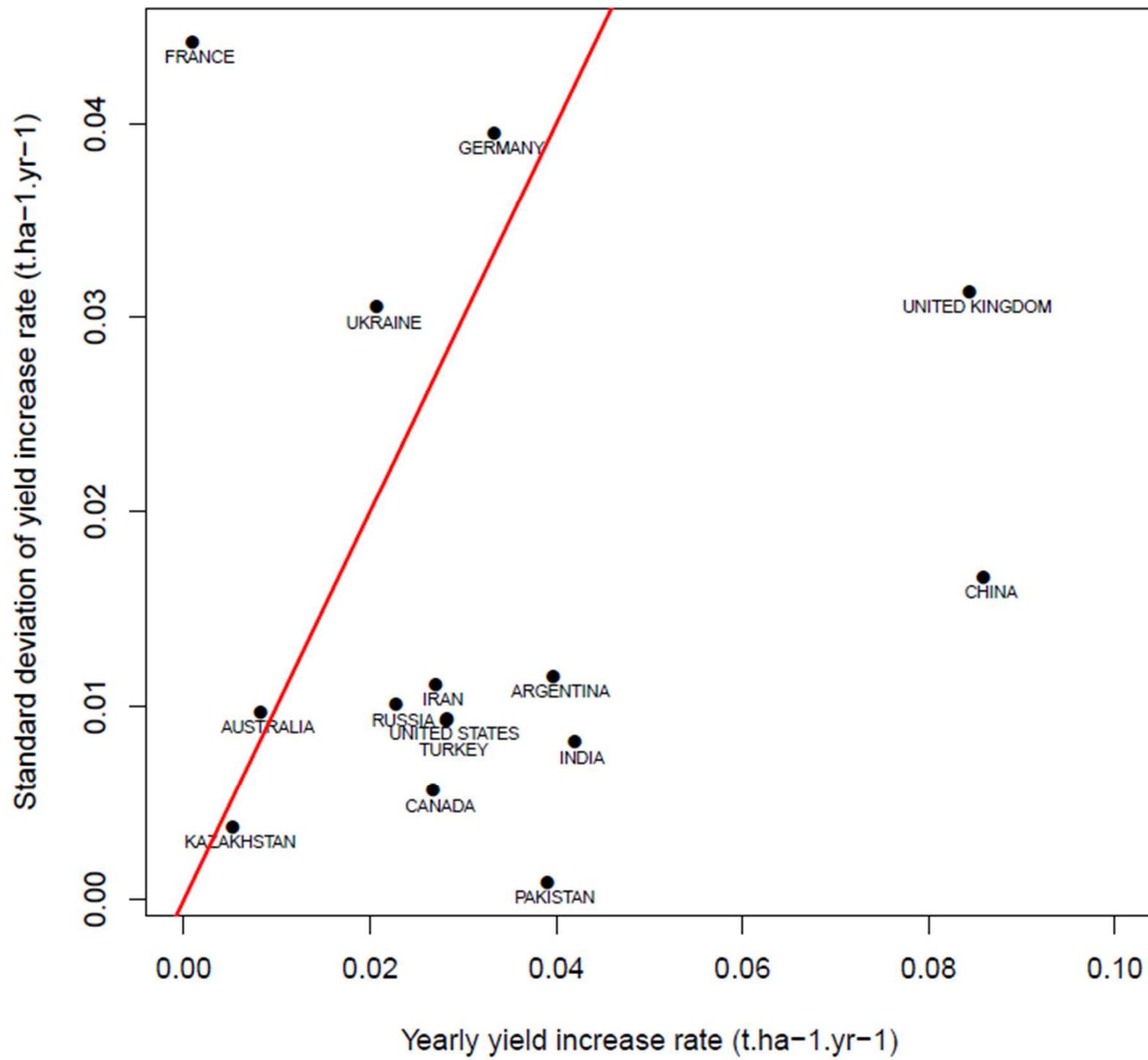


# Outil graphique

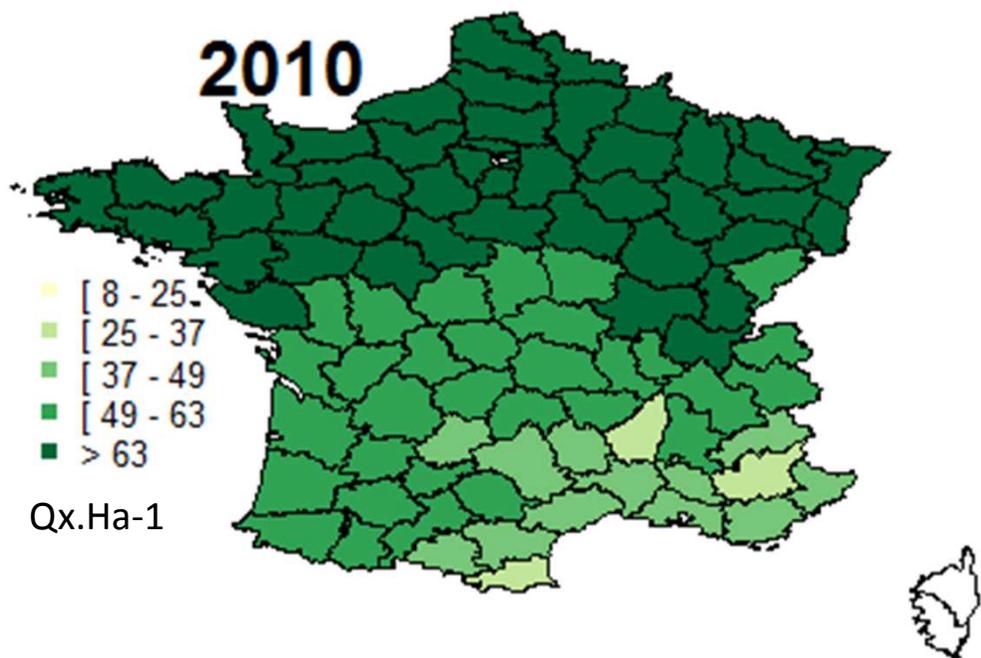
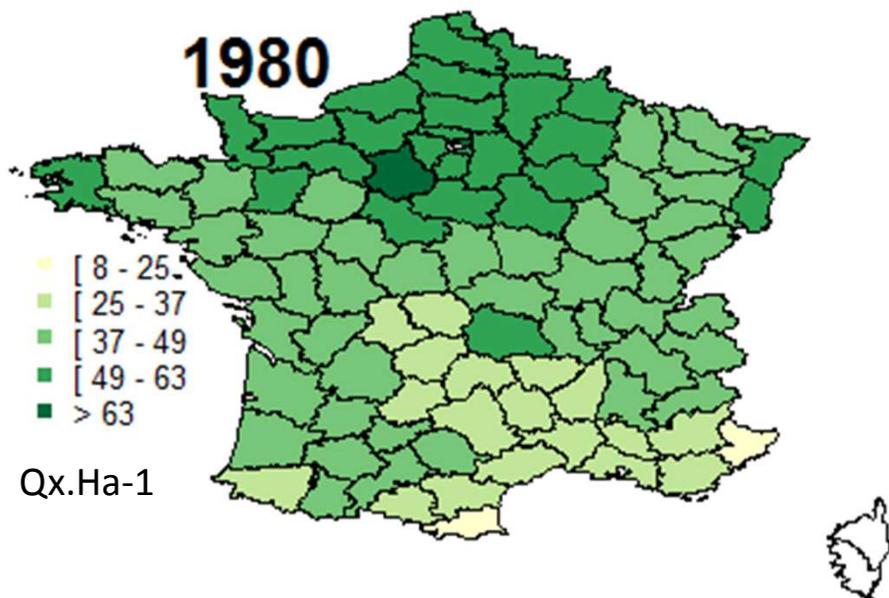


## Rendement du blé.

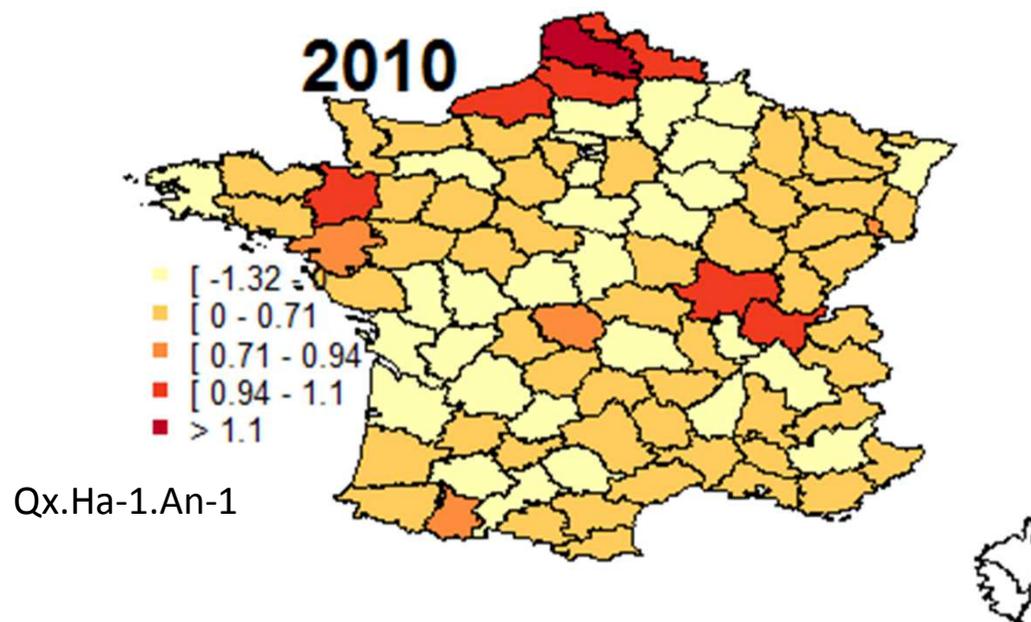
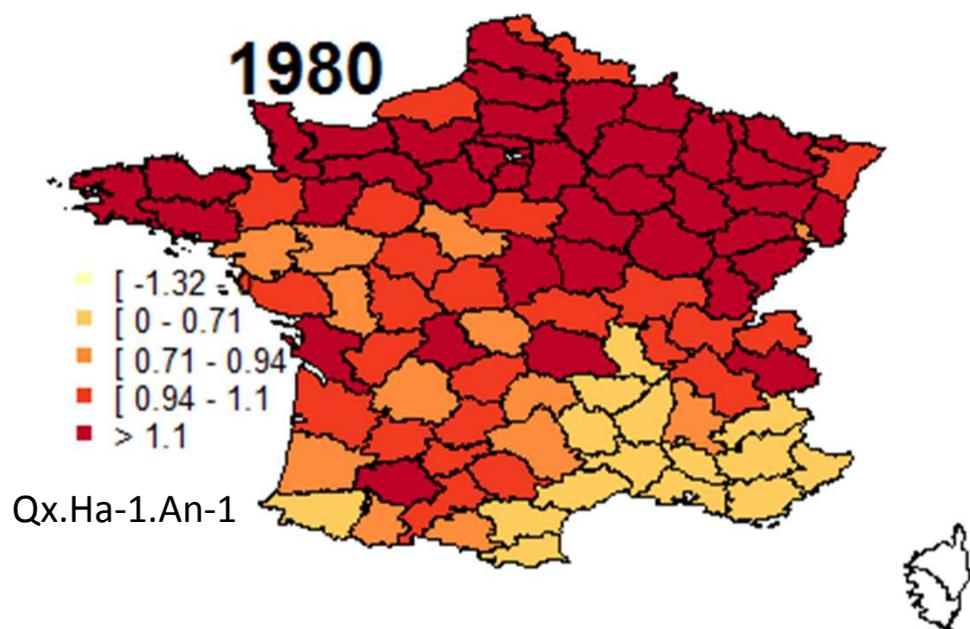




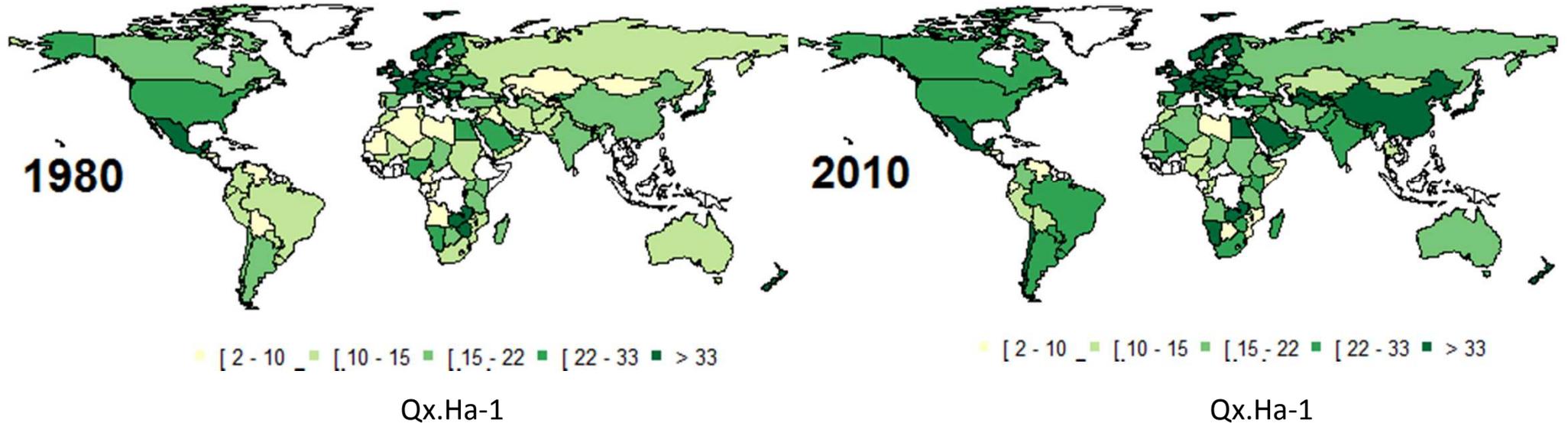
# Outil cartographique



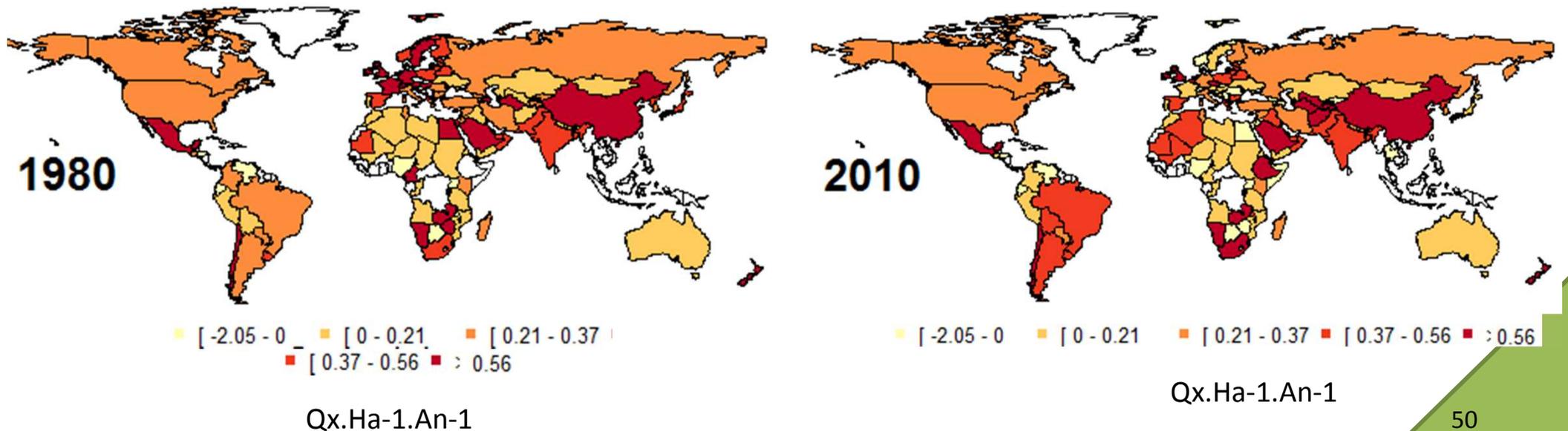
Evolution cartographique des rendements (ci-dessus) et des pentes (ci-dessous) du blé.

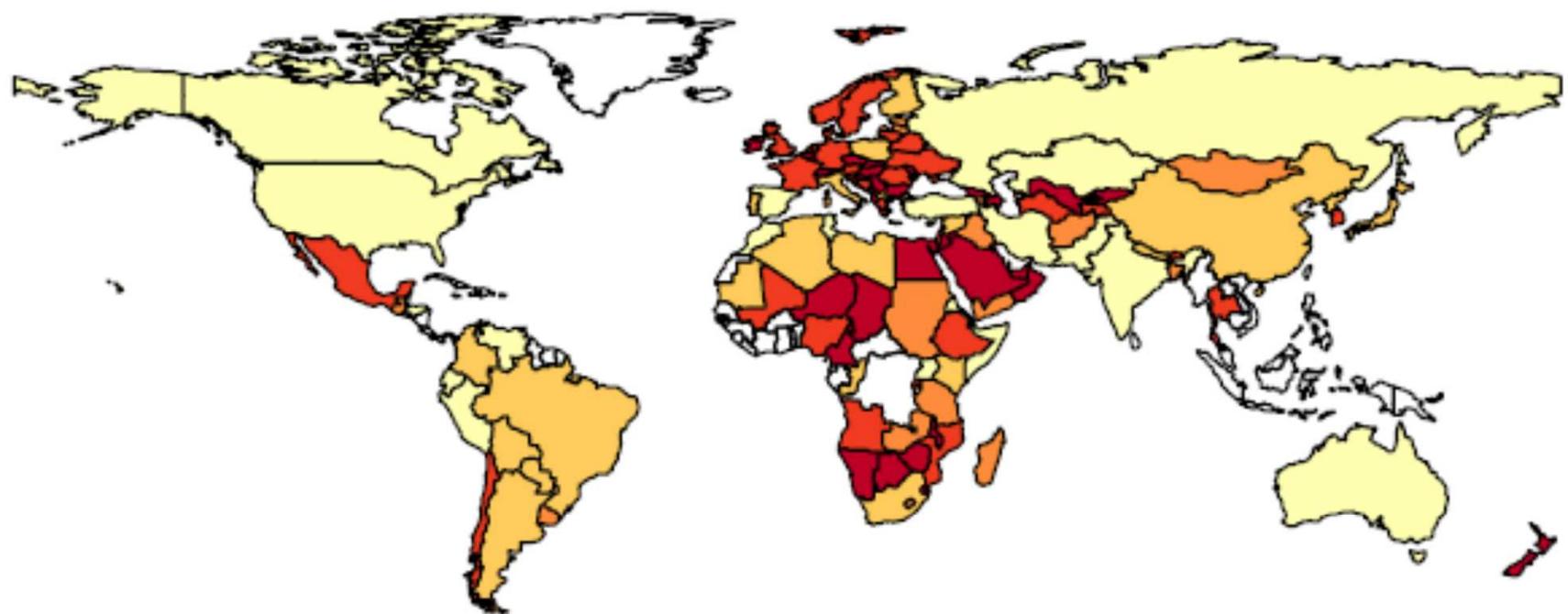


# Outil cartographique



Evolution cartographique des rendements (ci-dessus) et des pentes (ci-dessous) du blé.





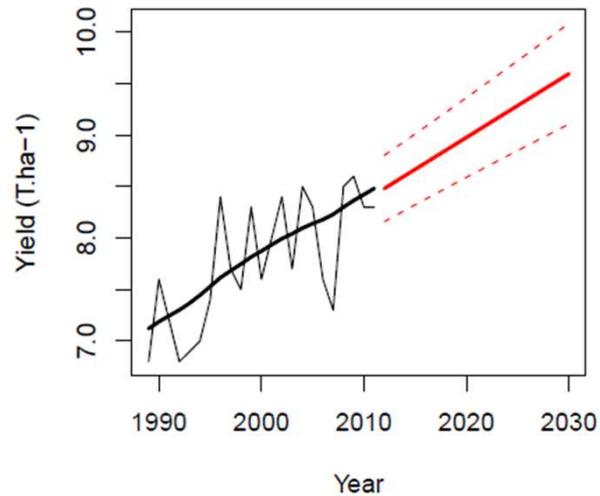
■ [ 0 – 0.01   ■ [ 0.01 – 0.02   ■ [ 0.02 – 0.03   ■ [ 0.03 – 0.05   ■ > 0.05  
t.ha<sup>-1</sup>.yr<sup>-1</sup>

# Outil générique

- Cet outil peut traiter plusieurs échelles et cultures différentes.

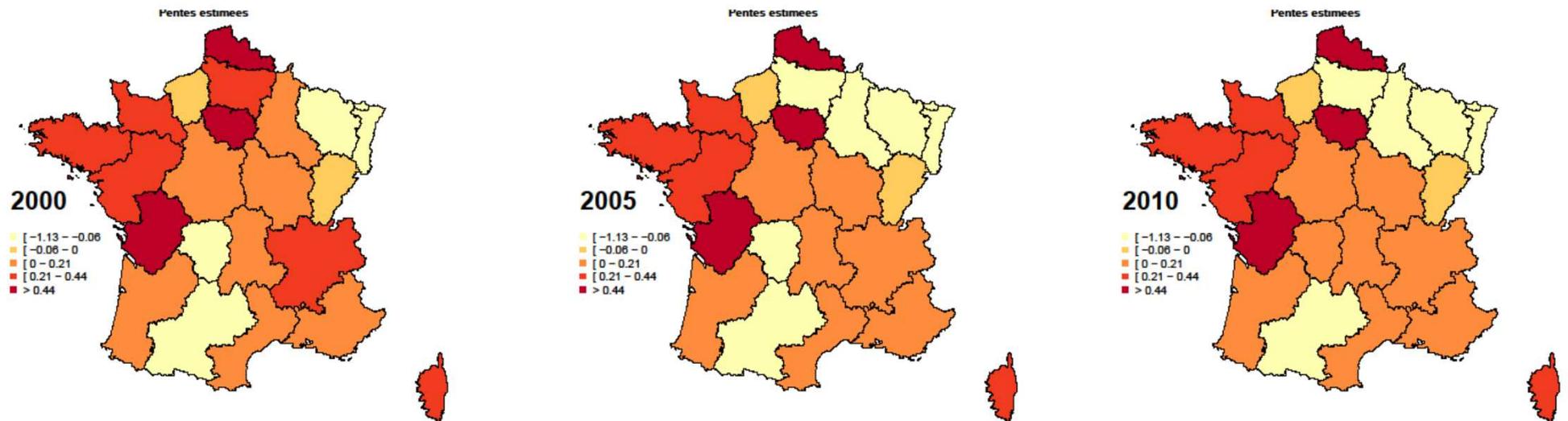
# Outil générique

NORD-PAS-DE-CALAIS



Evolution des rendements  
d'Orge en Nord-Pas-De-Calais

Evolution cartographique des pentes des rendements d'Avoine en Région



Qx.Ha-1.An-1

# Perspective

- Développer un package R pour automatiser les calculs

Merci de votre attention.