

# Améliorer les prévisions à court et moyen termes des modèles agronomiques en prenant mieux en compte l'incertitude des prévisions météorologiques

---

Ivana ALEKSOVSKA

Rencontres du réseau Mexico,  
Inria Centre Bordeaux Sud-Ouest

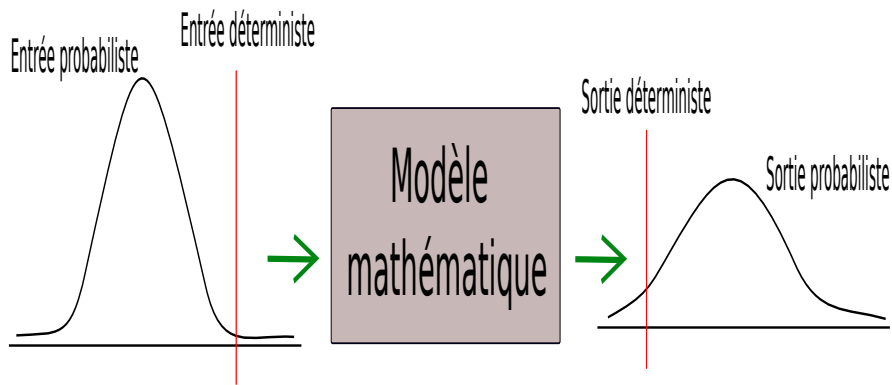
Encadrants: L. RAYNAUD (Météo-France), F. BRUN (ACTA), R. FAIVRE (INRA),  
M. RAYNAL (IFV), O. DEUDON (Arvalis), G. PIGEON (Météo-France)

(dans le cadre de la thèse d'Ivana ALEKSOVSKA)

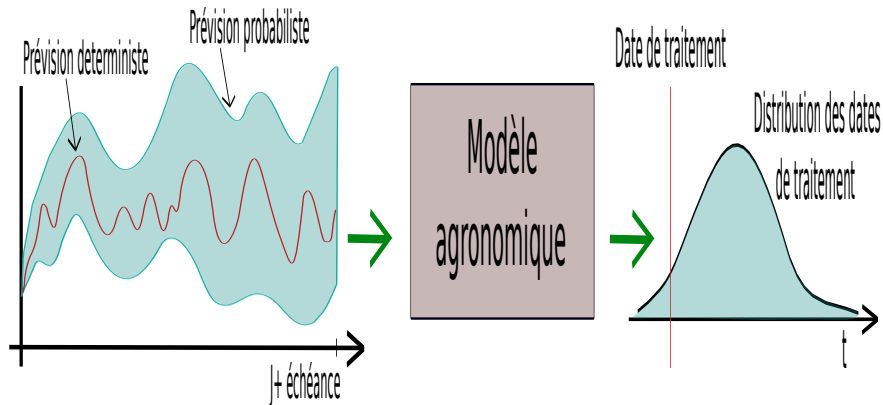
Novembre 12-13, 2018



## Schéma représentatif des entrées et sorties du modèle



## Schéma représentatif des entrées et sorties du modèle agronomique



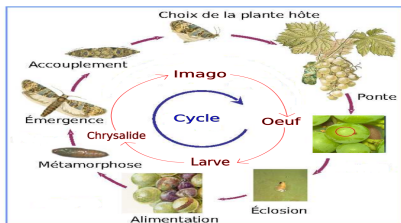
- 1 : Quelle est l'influence des paramètres du modèle EVA et leur incertitude associée sur les sorties ?
- 2 : Quelle est l'influence des prévisions météorologiques et leur incertitude associée sur les sorties ?
- 3 : Comment générer des prévisions météorologiques ?

## Cas d'étude

- ▶ Lutte contre un ravageur dont le développement dépend de la **température**: le ver de la grappe en vigne [3].  
→ modèle EVA
- ▶ Lutte contre une maladie dont la dynamique épidémique dépend des **précipitations** et de la **température**: la septoriose du blé [4].  
→ modèle Septo-LIS®
- ▶ Pilotage de l'irrigation, dont le bilan hydrique dépend de la **pluie** et de l'**évapotranspiration** (en vigne et/ou maïs) [5].  
→ modèle Wallis



# Cycles et stades de développement du ravageur



- ▶ Plusieurs cycles de reproduction (4 stades au cours d'un cycle: imagos, oeufs, larves (de L1 à L5), chrysalides) ;
- ▶ Nombre de cycles difficile à prévoir, très dépendant la température ;
- ▶ Traitement optimal doit être effectué lorsque les oeufs de la deuxième génération annuelle de l'insecte dépassent un seuil critique (2 et 15%) de l'effectif total de la population.

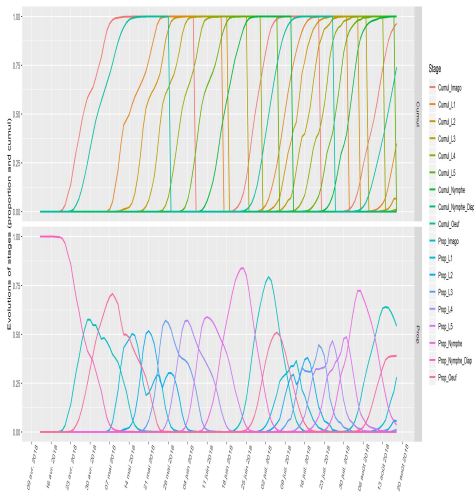
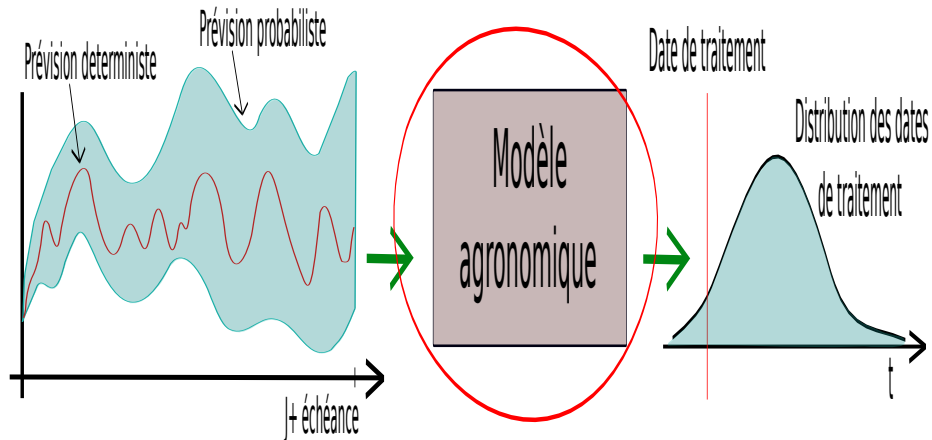


Figure: Sorties du modèle EVA à Pic Saint Loup.

## Quelle est l'influence des paramètres du modèle EVA et leur incertitude associée sur les sorties ?

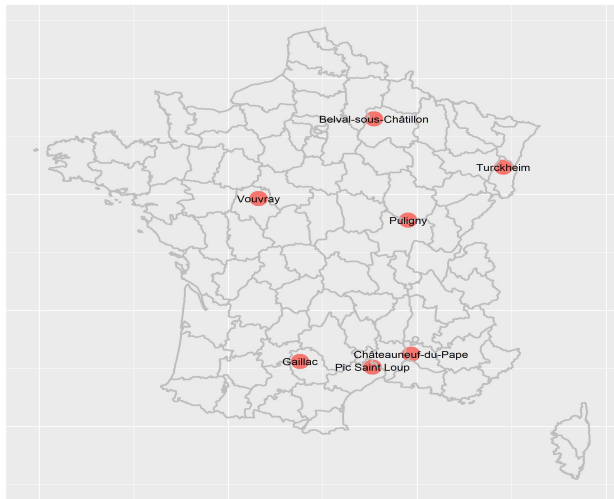


## Analyse de sensibilité des paramètres d'EVA

- ▶ Déterminer les paramètres à l'origine de la variabilité de la date de franchissement de seuil de 2%
- ▶ Prendre en compte la variabilité des condition météorologiques sur 5 ans et 7 lieux.

1. Méthode de criblage - Morris
2. Échantillonnage intensif - Sobol Jensen
3. Analyse temporelle - Influence des paramètres sur les sorties au cours du temps.


→ Travail stage : Olivier MAZE  
(Supervisé par R.FAIVRE et F.BRUN)





# Analyse par criblage - Morris

1. Sélection de 36 paramètres susceptibles d'être à l'origine de la variabilité du jour de dépassement du seuil de 2 %
2. Contraintes de temps de calcul liées à la lourdeur du modèle.
3. Mise en oeuvre d'une méthode de criblage (Morris)
4. Mise en place d'une étude plus poussée sur neuf paramètres :  
 ⇒ Analyse de sensibilité (méthode de Sobol)  
 ⇒ Analyse temporelle.

Classification des méthodes d'analyses de sensibilité 

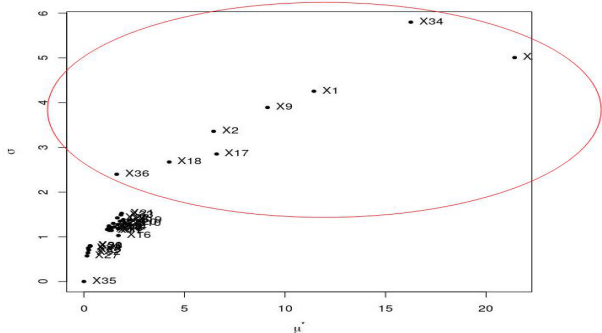
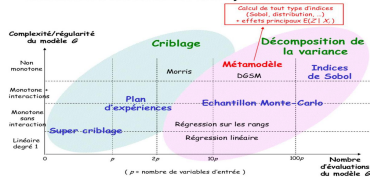
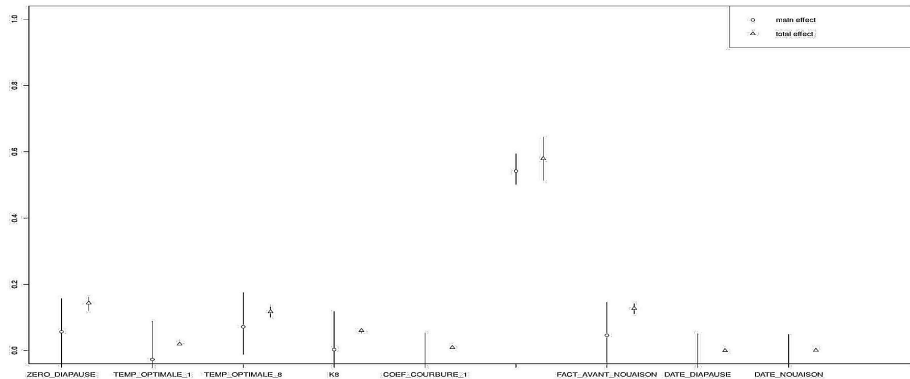


Figure: Morris, seuil 2 %, Chateaufneuf du Pape (84) 2015

# Analyse de sensibilité du modèle - Sobol-Jensen

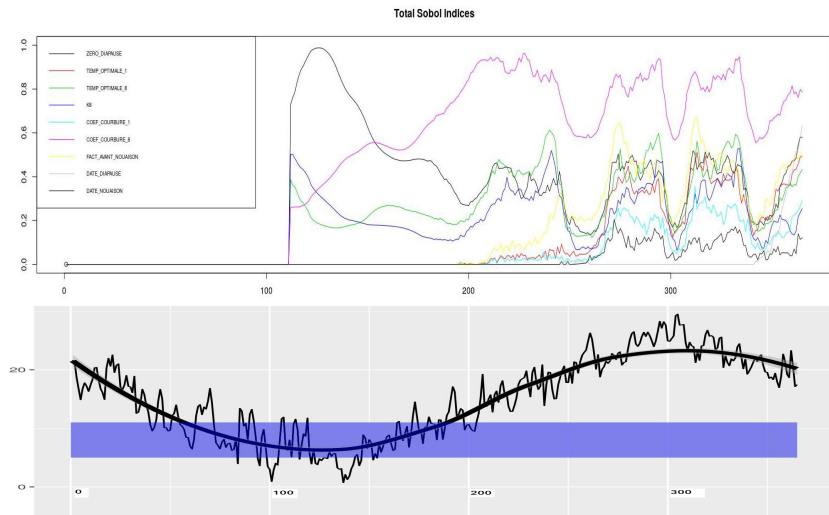


First order indices	original	bias	std error
ZERO_DIAPAUSE	0.055432420	-0.0012274518	0.05161602
TEMP_OPTIMALE_1	-0.026255357	0.0005243649	0.05066118
TEMP_OPTIMALE_8	0.070926030	-0.0006632766	0.04852297
K8	0.001437795	-0.0014638399	0.05461013
COEF_COURBURE_1	-0.054746490	-0.0011732303	0.05246288
COEF_COURBURE_8	0.543401335	0.0015386384	0.02414380

Total indices	original	bias	std error	min c.i.	max c.i.
ZERO_DIAPAUSE	1.430749e-01	-5.791263e-04	9.268860e-03	1.216029e-01	0.1615370404
TEMP_OPTIMALE_1	1.971334e-02	1.058623e-04	1.261151e-03	1.693415e-02	0.0219404071
TEMP_OPTIMALE_8	1.180694e-01	6.624970e-04	7.120549e-03	1.007683e-01	0.1317390301
K8	6.041678e-02	-1.246261e-05	3.901293e-03	5.230004e-02	0.0671206051
COEF_COURBURE_1	9.352344e-03	-6.832249e-06	6.872039e-04	7.780095e-03	0.0106715969
COEF_COURBURE_8	5.790793e-01	-7.081099e-04	3.555415e-02	5.142962e-01	0.6442333373
FACT_AVANT_NOUAISSON	1.275047e-01	1.563597e-05	7.620805e-03	1.109179e-01	0.141524221
DATE_DIAPAUSE	6.383852e-06	-3.860805e-08	7.136448e-06	-1.195674e-05	0.0000127677
DATE_NOUAISSON	8.171331e-04	-2.496249e-07	1.512445e-04	5.033137e-04	0.0010989117

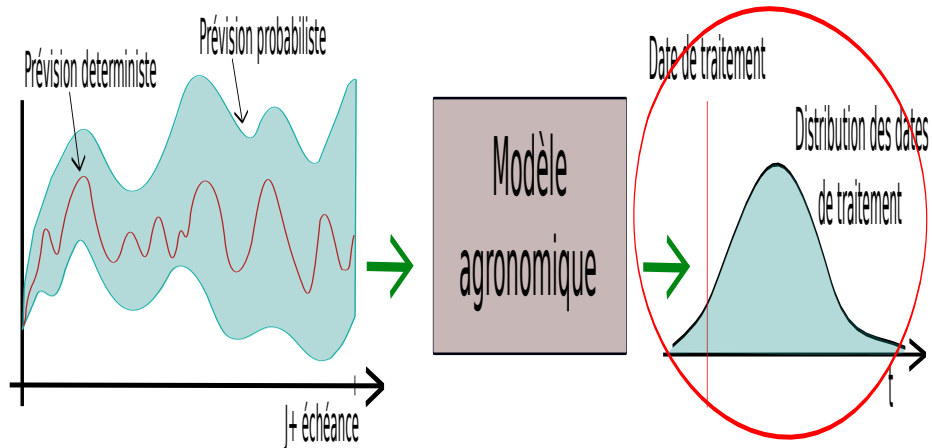
Figure: Indices de Sobol-Jansen pour Chateauneuf du Pape (84) en 2015.

# Analyse temporelle



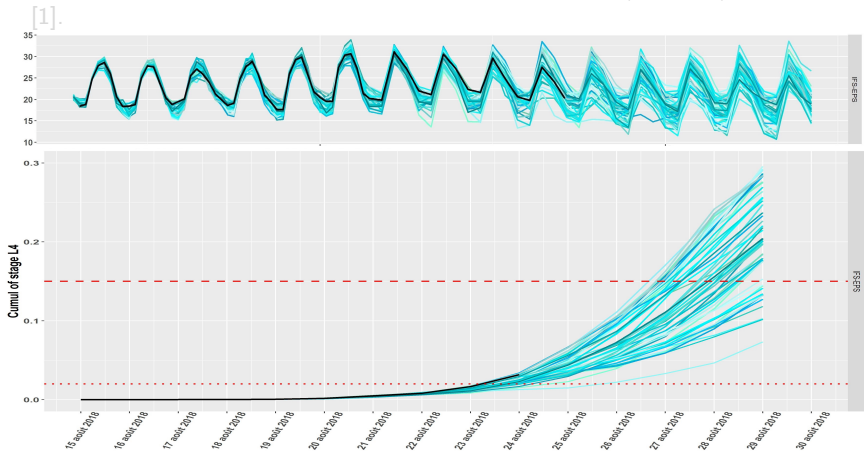
**Figure:** Haut : Evolution de la part relative de chaque paramètre sur la variation du cumul oeuf durant l'année (Chateauneuf du Pape (84) 2015). Bas : Températures moyennes journalières avec profil lissé et zone de franchissement du zéro de diapause (Chateauneuf du Pape (84) 2015)

## Quelle est l'influence des prévisions météorologiques et leur incertitude associée sur les sorties ?



## Sorties probabilistes du modèle EVA

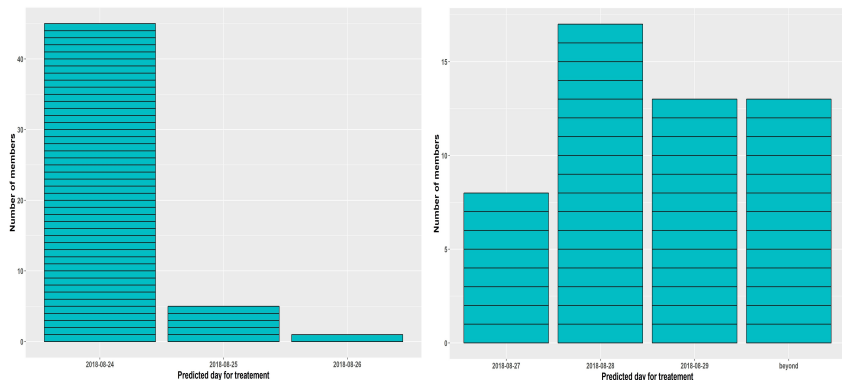
1. Exemple des entrées/sorties d'EVA en utilisant prévision probabiliste météorologique.
2. Distribution des dates de traitement en fonction des seuils/décision/prise de risque



**Figure:** Résultats des simulations d'ensemble pour le stade de larve L4 obtenus en utilisant des prévisions allant jusqu'à J+15 et franchissement du seuil de 2 % (ligne pointillée) et 15 % (ligne tirée).

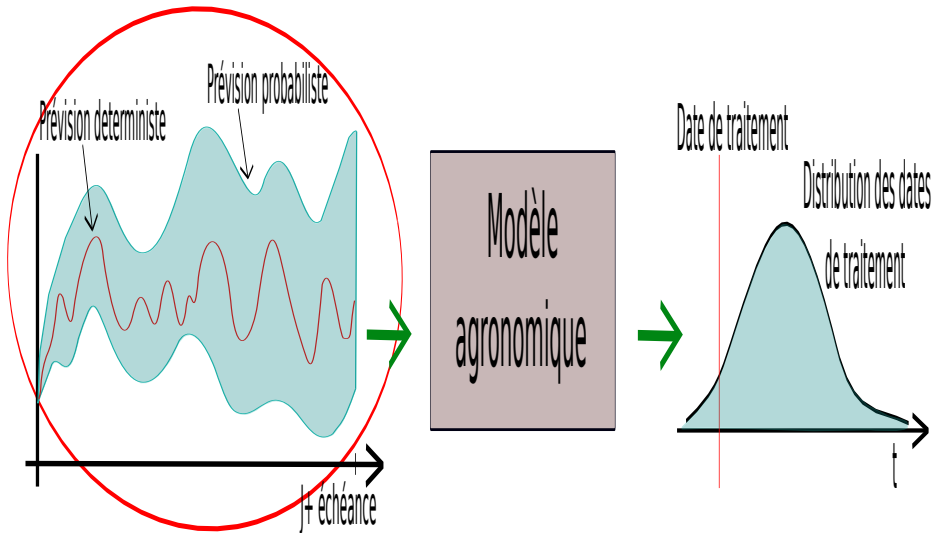
## Sorties probabilistes du modèle EVA

1. Exemple des entrées/sorties d'EVA en utilisant prévision probabiliste météorologique.
2. Distribution des dates de traitement en fonction des seuils/décision/prise de risque [1].



**Figure:** Gauche : Histogrammes des dates de déclenchement du traitement 2 %. Droite : Histogrammes des dates de déclenchement du traitement de 15 %.

## Comment générer des prévisions météorologiques ?



## Prévision d'ensemble

**Les conditions pour une prévision météorologique déterministe parfaite n'existent pas: de l'état initial aux prévisions à long terme, l'information météorologique est affectée par l'incertitude.**



La prévision probabiliste est un **ensemble de prévisions perturbées pour échantillonner la distribution de probabilité** (le modèle de prévision numérique de temps est exécuté plusieurs fois).



Il ne s'agit plus de prévoir l'état le plus probable de l'atmosphère mais la probabilité qu'un évènement spécifique se produise.



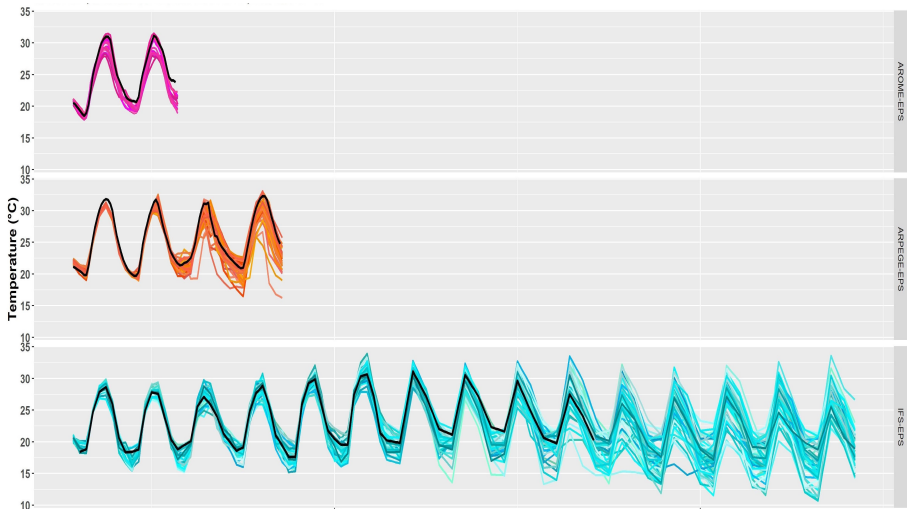
Approche probabiliste de la prévision pour estimer l'incertitude des prévisions (technique **prévisions d'ensemble** [2]).



cf. Présentation de Laure Raynaud : "Vers une prévision des incertitudes en météorologie" (Rencontres Mexico 2017)

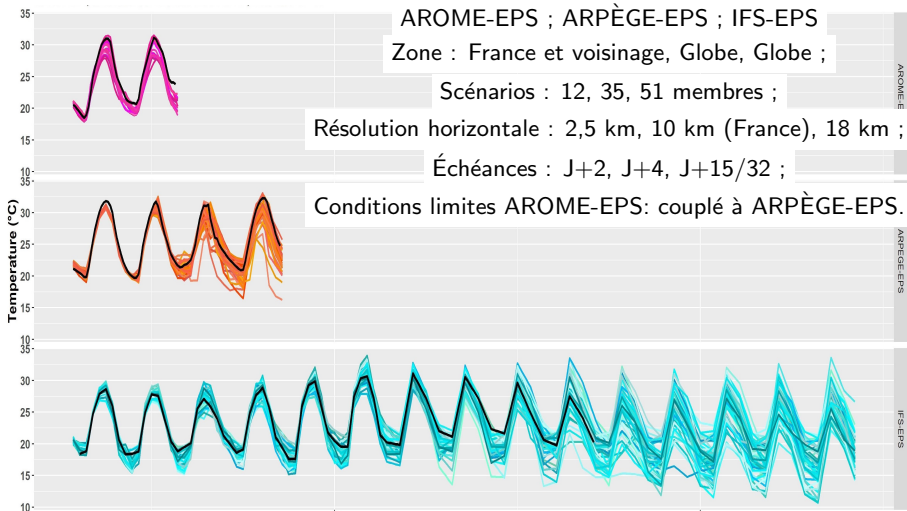


## Illustration des prévisions probabilistes



**Figure:** Prédiction probabiliste de la température AROME-EPs, ARPÈGE-EPs et IFS-EPs (avec leurs échéances respectives de 2, 4 et 15 jours) ainsi que la prédiction déterministe présentée en noir à Pic Saint Loup (34). Début des prévisions : 15/08/2018.

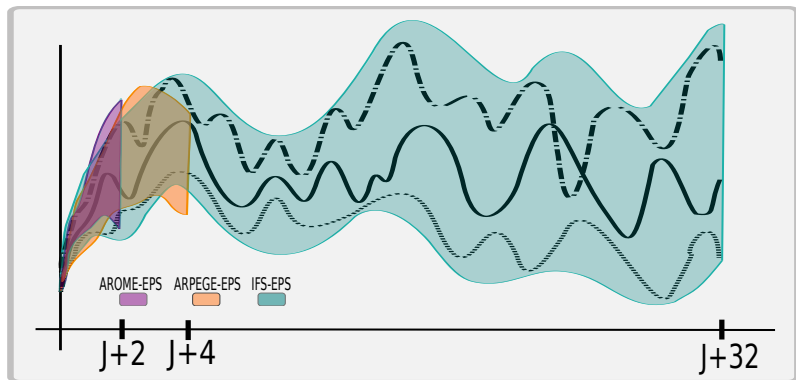
## Illustration des prévisions probabilistes



**Figure:** Prévision probabiliste de la température AROME-EPS, ARPÈGE-EPS et IFS-EPS (avec leurs échéances respectives de 2, 4 et 15 jours) ainsi que la prévision déterministe présentée en noir à Pic Saint Loup (34). Début des prévisions : 15/08/2018.

## Questions de recherche

Comment construire un ensemble de scénarii cohérents à partir des différents systèmes de prévision pour les différentes échéances ?

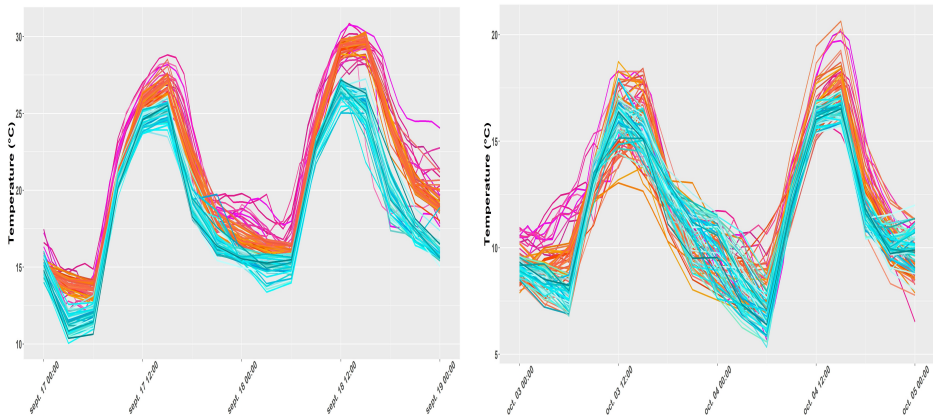


## Scénarii cohérents

Comment passer d'AROME à ARPEGE et puis à IFS ?

Stratégie : Considérer les distances entre les membres des différents systèmes.

## Exemples de prévisions

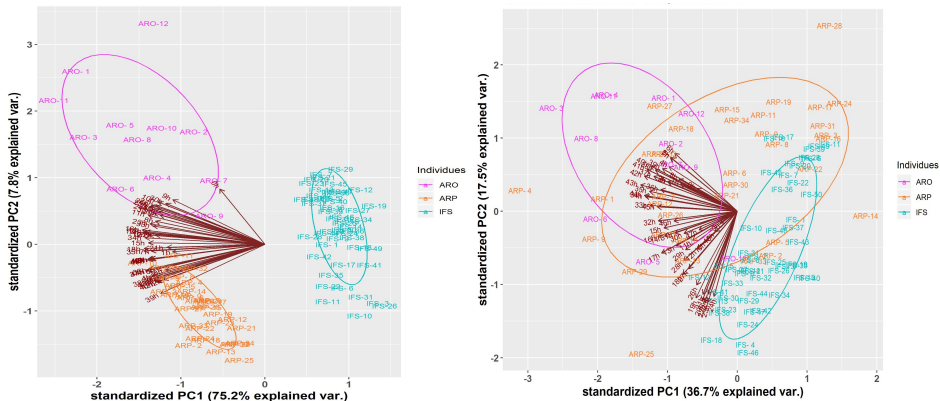


**Figure:** Gauche : 2018/09/17 (J+2), Droite :2018/10/02 (J+2) à Turckheim (68)

# Exploration statistique des prévisions météorologiques

- I : Méthodes factorielles : les méthodes dites factorielles de décomposition sur une base adaptée : les facteurs sur lesquels projeter les données pour des représentations graphiques en dimension réduite :
  - ▶ Analyse en Composantes Principales (ACP)
  - ▶ Positionnement Multidimensionnel (Multidimensional Scaling (MDS)).
  
- II : Classification non supervisée ou clustering : les méthodes ou algorithmes visent la recherche de classes, ou regroupements des individus, se ressemblant au mieux ou les plus proches au sens d'une mesure de distance :
  - ▶ Classification ascendante hiérarchique (CAH).
  - ▶ Choix de distance, choix de linkage ?

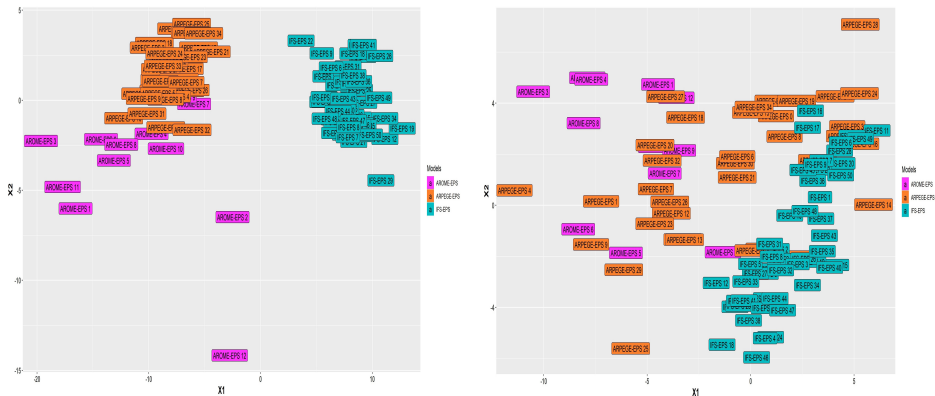
# Analyse en Composantes Principales (ACP)



**Figure:** Gauche : 2018/09/17 (J+2), Droite : 2018/10/02 (J+2) à Turkckheim (68)

Les trois systèmes sont distincts (gauche) et partiellement mélangés (droite) et les différences entre membres sont quelque soit l'heure (surtout à gauche).

## Positionnement Multidimensionnel (MDS)



**Figure:** Distance: Euclidien. Gauche : 2018/09/17 (J+2), Droite : 2018/10/02 (J+2) à Turckheim (68)

-Mêmes conclusions que pour l'ACP.

-Résultats analogues pour les méthodes de classifications : distances (Euclidean, Manhattan, Minkowski, DTW, Complexity Invariance Distance) et fonction de linkage (single, complete, average, centroid, ward).

# Perspectives

## 1 : Classification :

- ▶ Débiaisage ;
- ▶ Approche fonctionnelle ;
- ▶ Information de couplage (AROME-EPS et ARPÈGE-EPS).

## 2 : Jonction :

- ▶ AS sur des prévisions et les paramètres du modèle
- ▶ Tester différents méthodes de couplage (le plus proche, k-plus proche, à base de proba, pondération par performance de modèle)

## 3 : Validation :

- ▶ Validation des scénarios cohérents
- ▶ Généralisation à d'autres sites, périodes
- ▶ Extension à plus d'une variable météorologique pour d'autres cas d'étude.





I. Aleksovska, F. Brun, L. Raynaud, R. Faivre, M. Raynal, and O. Deudon.

Prendre en compte l'incertitude des prévisions météorologiques dans les oad utilisées pour gérer les maladies et ravageurs des cultures.

*Végéphyt - 12e conférence internationale sur les maladies des plantes - Tours (France) - 10, 11 et 12 décembre 2018.*



F. Bouttier and R. Buizza.

La prévision d'ensemble.

*La Météorologie*, (100):pp. 29–36, Fevrier, 2018.



F. Chavent.

Etude de la dynamique des populations d'eudémis de la vigne( lobesia botrana den. et s hi.). adaptation d'un modèle d'évolution aux conditions climatiques de provence, memoire de fin n d'étude, 1983.



D. Gouache, A. Bensadoun, F. Brun, C. Pagé, D. Makowski, and D. Wallach.

Modelling climate change impact on septoria tritici blotch (stb) in france: accounting for climate model and disease model uncertainty.

*Agricultural and forest meteorology*, 170:pp. 242–252, 2013.



S. Roux, F. Brun, and D. Wallach.

Combining input uncertainty and residual error in crop model predictions: A case study on vineyards.

*European journal of agronomy*, 52:191–197, 2014.

Merci de votre attention!

**ivana.aleksovska@meteo.fr**

## Definition

ACP est une procédure statistique qui utilise une transformation orthogonale en un ensemble d'observations de variables éventuellement corrélées dans un ensemble de valeurs de variables linéairement non corrélées appelées composantes principales.

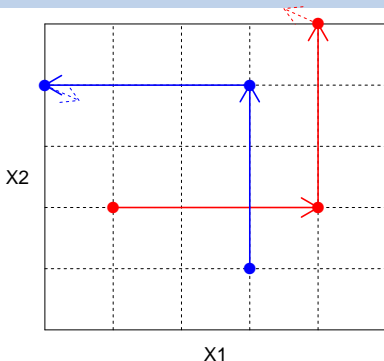
- ▶ Soit  $p$  variable statistique réelles  $X^j, j = 1, \dots, p$ , observées sur  $n$  individus,
- ▶ Soit  $\text{cov}(X^{j1}, X^{j2}), j1, j2 \in 1, \dots, p$  la matrice de variance-covariance des variables,

→ Les vecteurs propres de la matrice  $\text{cov}(X^{j1}, X^{j2})$  seront les composants principaux et les valeurs propres représenteront le pourcentage de la variance expliquée sur la composante principale.

# Analyse de sensibilité globale

## Méthode de Morris

Une exploration astucieuse de l'espace



### Grands principes

- ▶ Discrétisation de l'espace
- ▶ Un déplacement sur un seul facteur à la fois (méthode OAT)
- ▶ Un point de départ aléatoire
- ▶ Une trajectoire correspond à K déplacements (K+1 points)

## Méthode de Morris

### Analyse

Mesures d'importance du facteur  $x_i$  évaluées sur  $R$  trajectoires  $T^r$  avec le segment  $[\mathcal{G}(T_{(i-1)}^r), \mathcal{G}(T_{(i)}^r)]$  correspondant au saut du facteur  $x_i$

- ▶ effet élémentaire :  $\Delta_i^r = \frac{\mathcal{G}(T_{(i-1)}^r) - \mathcal{G}(T_{(i)}^r)}{\delta}$ ,  $\delta$  la taille du saut
- ▶ effect moyen :  $\mu_i = \frac{1}{R} \sum_{r=1}^R \Delta_i^r$
- ▶ effect absolu moyen :  $\mu_i^* = \frac{1}{R} \sum_{r=1}^R |\Delta_i^r|$
- ▶ écart-type des effets élémentaires :  $\sigma_i = \sqrt{\frac{1}{R-1} \sum_{r=1}^R (\Delta_i^r - \mu_i)^2}$

# Théorème de la variance globale

## Autre formulation des indices

- ▶  $SI_i = \frac{Var[Esp(\mathcal{G}(\mathbf{x})|x_i)]}{Var[\mathcal{G}(\mathbf{x})]}$
- ▶  $TSI_i = \frac{Esp(Var[\mathcal{G}|\mathbf{x}^{(-i)}])}{Var[\mathcal{G}(\mathbf{x})]} = \frac{Esp(Var[\mathcal{G}|\mathbf{x}_{\sim i}])}{Var[\mathcal{G}(\mathbf{x})]}$

## Interprétation

- ▶  $SI_i$  : pour chaque valeur du facteur  $x_i$ , on fait bouger les autres facteurs et on calcule les moyennes en chaque valeur. Ensuite, on regarde si la moyenne (espérance) bouge.
- ▶  $TSI_i$  : pour tous les n-uplets de valeurs de tous les facteurs ( $\mathbf{x}^{(-i)}$ ) sauf  $x_i$ , on fait bouger le facteur  $x_i$  et on calcule la variabilité. On en fait ensuite la moyenne.

→ Principe des méthodes d'analyse de variance sur plan factoriel ou des méthodes dites de de Pick and Freeze ou de Sobol' par échantillonnage alternatif, méthodes nécessitant beaucoup d'évaluation du modèle.

# Méthode d'estimation des indices

Différentes manières de faire les calculs

$V_{X_i}(E_{X_{\sim i}}(Y X_i))$ pour calculer $S_i$	Calcul sensitivity <sup>1</sup>
$\frac{1}{N} \sum_{j=1}^N f(A)_j f(B_A^{(i)})_j - f_0^2$	Sobol' 1993 (sobol2002)
$\frac{1}{N} \sum_{j=1}^N f(B)_j (f(A_B^{(i)})_j - f(A)_j)$	sobol2007
$V(Y) - \frac{1}{2N} \sum_{j=1}^N (f(B_j) - f(A_B^{(i)})_j)^2$	soboljansen
$E_{X_{\sim i}}(V_{X_i}(Y X_{\sim i}))$ pour calculer $TS_i$	
$V(Y) - \frac{1}{N} \sum_{j=1}^N f(A)_j f(A_B^{(i)})_j + f_0^2$	Homma 1996 (sobol2002)
$\frac{1}{N} \sum_{j=1}^N f(A)_j (f(A)_j - f(A_B^{(i)})_j)$	sobol2007
$V(Y) - \frac{1}{2N} \sum_{j=1}^N (f(A_j) - f(A_B^{(i)})_j)^2$	soboljansen

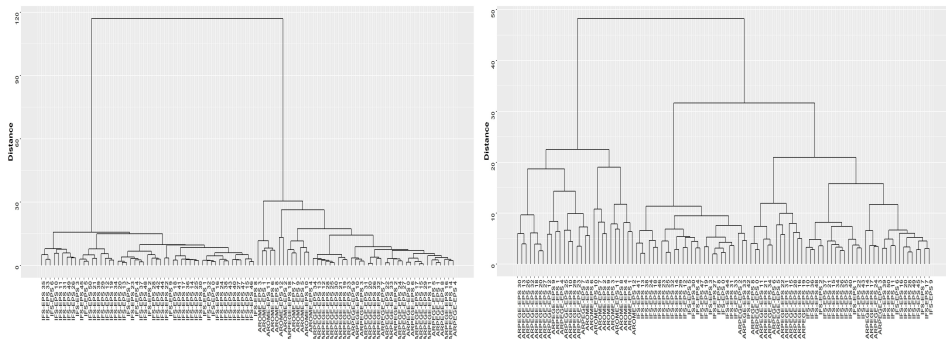
$A_B^{(i)}$  Matrice  $A$  pour laquelle la colonne  $i$  vient de  $B$  ( $C_i$ )

$B_A^{(i)}$  Matrice  $B$  pour laquelle la colonne  $i$  vient de  $A$

Table 2 de Saltelli A. *et al.*, 2010.

## CAH

CAH c'est une méthode de classification non supervisée (clustering) utilisée en analyse des données. Son but est de regrouper itérativement les individus dans différentes classes, en commençant par les deux plus proches et en construisant progressivement un dendrogramme. MDS nécessite une notion de distance et une saut (linkage).



**Figure:** Distance: Euclidien, Saut : Ward. Gauche : 2018/09/17 (J+2), Droite :2018/10/02 (J+2) à Turckheim (68)