

Modèles géostatistiques spatio-temporels pour la hauteur des vagues dans l'Atlantique Nord

Université de Bretagne Sud

Nicolas Raillard

`nicolas.raillard@gmail.com`

Valérie Monbet (UBS), Pierre Ailliot (UBO), Anne Cuzol (UBS)

Université de Rennes 1
2007–2008

Vendredi 03 Octobre 2008

Plan

- 1 Introduction
- 2 Données
- 3 Modèles de covariance
 - Présentation
 - Méthodologie
 - Résultats
- 4 Filtrage
 - Présentation
 - Résultats
- 5 Comparaison des méthodes
- 6 Conclusion - Perspectives

Plan

- 1 Introduction
- 2 Données
- 3 Modèles de covariance
 - Présentation
 - Méthodologie
 - Résultats
- 4 Filtrage
 - Présentation
 - Résultats
- 5 Comparaison des méthodes
- 6 Conclusion - Perspectives

Plan

- 1 Introduction
- 2 Données
- 3 Modèles de covariance
 - Présentation
 - Méthodologie
 - Résultats
- 4 Filtrage
 - Présentation
 - Résultats
- 5 Comparaison des méthodes
- 6 Conclusion - Perspectives

Plan

- 1 Introduction
- 2 Données
- 3 Modèles de covariance
 - Présentation
 - Méthodologie
 - Résultats
- 4 Filtrage
 - Présentation
 - Résultats
- 5 Comparaison des méthodes
- 6 Conclusion - Perspectives

Plan

- 1 Introduction
- 2 Données
- 3 Modèles de covariance
 - Présentation
 - Méthodologie
 - Résultats
- 4 Filtrage
 - Présentation
 - Résultats
- 5 Comparaison des méthodes
- 6 Conclusion - Perspectives

Plan

- 1 Introduction
- 2 Données
- 3 Modèles de covariance
 - Présentation
 - Méthodologie
 - Résultats
- 4 Filtrage
 - Présentation
 - Résultats
- 5 Comparaison des méthodes
- 6 Conclusion - Perspectives

Introduction

Données environnementales

- Réseaux de capteurs
- Evolution temporelle
- Modèles déterministes d'évolution

Introduction

Données environnementales

- Réseaux de capteurs
- Evolution temporelle
- Modèles déterministes d'évolution

Limitations

- Méthodes numériques lourdes
- Interpolation et extrapolation
- Coût de calculs

Introduction

Données environnementales

- Réseaux de capteurs
- Evolution temporelle
- Modèles déterministes d'évolution

Limitations

- Méthodes numériques lourdes
- Interpolation et extrapolation
- Coût de calculs

Cadre de l'étude

- Etat de mer
- Modèles de géostatistique spatio-temporel
- Modèles de filtrage

Données

ERA-40

- Hindcast
 - Grille régulière : 2.5°, toutes les 6 heures
 - Uniquement mois de décembre
 - H_s, Θ_m, T_m
 - $N_{an} = 21, N_t = 124, 5 \times 5$ points
 - Noté $Z(s, t, a)$, centré réduit

Données

ERA-40

- Hindcast
- Grille régulière : 2.5°, toutes les 6 heures
- Uniquement mois de décembre
- H_s, Θ_m, T_m
- $N_{an} = 21, N_t = 124, 5 \times 5$ points
- Noté $Z(s, t, a)$, centré réduit

Données

ERA-40

- Hindcast
- Grille régulière : 2.5°, toutes les 6 heures
- Uniquement mois de décembre
- H_s, Θ_m, T_m
- $N_{an} = 21, N_t = 124, 5 \times 5$ points
- Noté $Z(s, t, a)$, centré réduit

Données

ERA-40

- Hindcast
- Grille régulière : 2.5°, toutes les 6 heures
- Uniquement mois de décembre
- H_s, Θ_m, T_m
- $N_{an} = 21, N_t = 124, 5 \times 5$ points
- Noté $Z(s, t, a)$, centré réduit

Données

ERA-40

- Hindcast
- Grille régulière : 2.5°, toutes les 6 heures
- Uniquement mois de décembre
- H_s, Θ_m, T_m
- $N_{an} = 21, N_t = 124, 5 \times 5$ points
- Noté $Z(s, t, a)$, centré réduit

Données

ERA-40

- Hindcast
- Grille régulière : 2.5°, toutes les 6 heures
- Uniquement mois de décembre
- H_s, Θ_m, T_m
- $N_{an} = 21, N_t = 124, 5 \times 5$ points
- Noté $Z(s, t, a)$, centré réduit

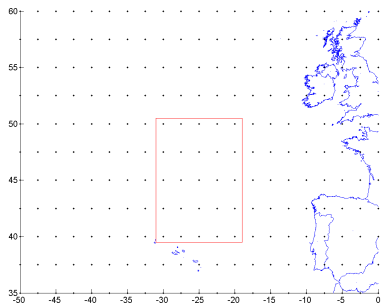
Données
ERA-40

Fig.: Zone retenue

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Fig.: Ordre des points de mesure

Données ERA-40

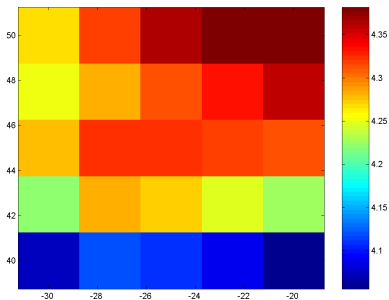


Fig.: Moyenne de la hauteur des vagues

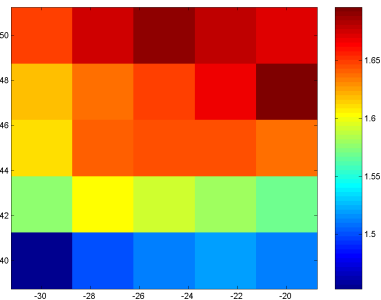


Fig.: Variance de la hauteur des vagues

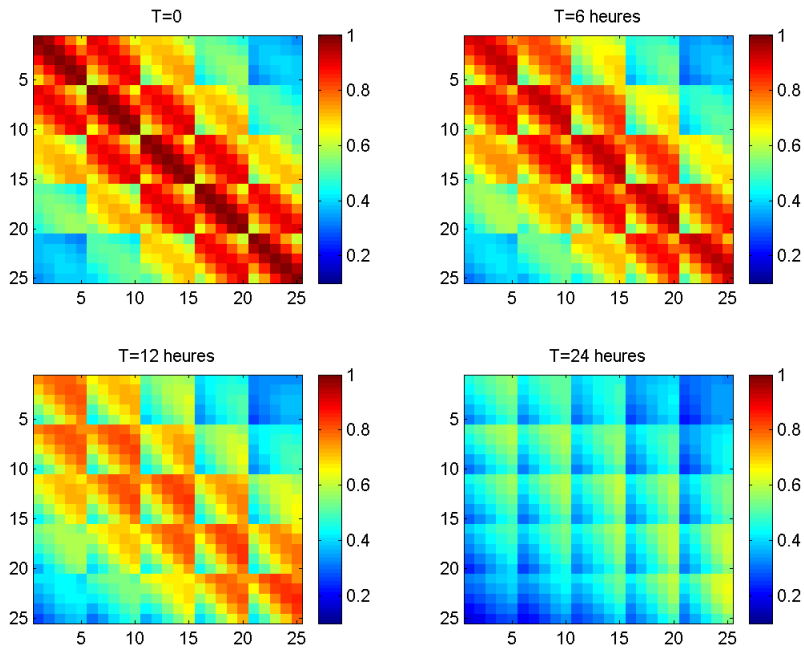


Fig : Covariance empirique

Modèles

Modèles Spatio-temporels

Modèles sur la covariance

- Décrit la dépendance spatio-temporelle du processus
- Fonction définie-positive
- Deux modèles :

$$C_1(s, s', t) = \rho \exp(-\alpha_1 d_x^2 - \alpha_2 d_y^2 - \alpha_3 |t - t'|) \quad \rho = \text{Cov}(Z(s, t), Z(s', t))$$

d_x et d_y sont les distances entre s et s'

$$C_2(s, s', t) = \rho \exp(-\alpha_1 d_x^2 - \alpha_2 d_y^2 - \alpha_3 |t - t'| - \alpha_4 |t - t'| d_x - \alpha_5 |t - t'| d_y) \quad \rho = \text{Cov}(Z(s, t), Z(s', t))$$

d_x et d_y sont les distances entre s et $s' - vt$

- Estimation par moindres carrés pondérés : $L_i(a, v, \sigma) =$

$$\sum_{(s, s') \in D^2} \sum_{t \in T} P(s, s', t) \left(C_i(s, s', t) - \widehat{C}_{\text{emp}} \right)^2$$

- Prédiction = Krigage : meilleure prédiction linéaire sans biais

Modèles

Modèles Spatio-temporels

Modèles sur la covariance

- Décrit la dépendance spatio-temporelle du processus
- Fonction définie-positive
- Deux modèles :

$$C_1(s, s', t) = \sigma e^{-\alpha_1 dx^2 - \alpha_2 dx - \alpha_3 dy^2 - \alpha_4 dy - \alpha_5 dx dy} e^{-\alpha_6 t^2 - \alpha_7 t}$$

dx et dy sont les distances entre s et s'

$$C_2(s, s', t) = \sigma e^{-\alpha_1 dx^2 - \alpha_2 dt^2 - \alpha_3 dy^2 - \alpha_4 dy - \alpha_5 dx' dy' - \alpha_6 t^2 - \alpha_7 t}$$

dx' et dy' sont les distances entre s et $s' - vt$

- Estimation par moindres carrés pondérés : $L_i(a, v, \sigma) = \sum_{(s,s') \in D^2} \sum_{t \in T} P(s, s', t) \left(C_i(s, s', t) - \widehat{C}_{emp} \right)^2$
- Prédiction = Krigage : meilleure prédiction linéaire sans biais

Modèles

Modèles Spatio-temporels

Modèles sur la covariance

- Décrit la dépendance spatio-temporelle du processus
- Fonction définie-positive
- Deux modèles :

- $C_1(s, s', t) = \sigma e^{-a_1 dx^2 - a_2 dx - a_3 dy^2 - a_4 dy - a_5 dx dy} e^{-a_6 t^2 - a_7 t}$
 dx et dy sont les distances entre s et s'

- $C_2(s, s', t) = \sigma e^{-a_1 dx'^2 - a_2 dx' - a_3 dy'^2 - a_4 dy' - a_5 dx' dy'} e^{-a_6 t^2 - a_7 t}$
 dx' et dy' sont les distances entre s et $s' - vt$

- Estimation par moindres carrés pondérés : $L_i(a, v, \sigma) =$

$$\sum_{(s, s') \in D^2} \sum_{t \in T} P(s, s', t) \left(C_i(s, s', t) - \widehat{C}_{\text{emp}} \right)^2$$

- Prédiction = Krigeage : meilleure prédiction linéaire sans biais

Modèles

Modèles Spatio-temporels

Modèles sur la covariance

- Décrit la dépendance spatio-temporelle du processus
- Fonction définie-positive
- Deux modèles :
 - $C_1(s, s', t) = \sigma e^{-a_1 dx^2 - a_2 dx - a_3 dy^2 - a_4 dy - a_5 dx dy} e^{-a_6 t^2 - a_7 t}$
 dx et dy sont les distances entre s et s'
 - $C_2(s, s', t) = \sigma e^{-a_1 dx'^2 - a_2 dx' - a_3 dy'^2 - a_4 dy' - a_5 dx' dy'} e^{-a_6 t^2 - a_7 t}$
 dx' et dy' sont les distances entre s et $s' - vt$
- Estimation par moindres carrés pondérés : $L_i(a, v, \sigma) = \sum_{(s, s') \in D^2} \sum_{t \in T} P(s, s', t) \left(C_i(s, s', t) - \widehat{C}_{\text{emp}} \right)^2$
- Prédiction = Krigeage : meilleure prédiction linéaire sans biais

Modèles

Modèles Spatio-temporels

Modèles sur la covariance

- Décrit la dépendance spatio-temporelle du processus
- Fonction définie-positive
- Deux modèles :
 - $C_1(s, s', t) = \sigma e^{-a_1 dx^2 - a_2 dx - a_3 dy^2 - a_4 dy - a_5 dx dy} e^{-a_6 t^2 - a_7 t}$
 dx et dy sont les distances entre s et s'
 - $C_2(s, s', t) = \sigma e^{-a_1 dx'^2 - a_2 dx' - a_3 dy'^2 - a_4 dy' - a_5 dx' dy'} e^{-a_6 t^2 - a_7 t}$
 dx' et dy' sont les distances entre s et $s' - vt$
- Estimation par moindres carrés pondérés : $L_i(a, v, \sigma) = \sum_{(s, s') \in D^2} \sum_{t \in T} P(s, s', t) \left(C_i(s, s', t) - \widehat{C}_{\text{emp}} \right)^2$
- Prédiction = Krigeage : meilleure prédiction linéaire sans biais

Modèles

Modèles Spatio-temporels

Modèles sur la covariance

- Décrit la dépendance spatio-temporelle du processus
- Fonction définie-positive
- Deux modèles :
 - $C_1(s, s', t) = \sigma e^{-a_1 dx^2 - a_2 dx - a_3 dy^2 - a_4 dy - a_5 dx dy} e^{-a_6 t^2 - a_7 t}$
 dx et dy sont les distances entre s et s'
 - $C_2(s, s', t) = \sigma e^{-a_1 dx'^2 - a_2 dx' - a_3 dy'^2 - a_4 dy' - a_5 dx' dy'} e^{-a_6 t^2 - a_7 t}$
 dx' et dy' sont les distances entre s et $s' - vt$
- Estimation par moindres carrés pondérés : $L_i(a, v, \sigma) = \sum_{(s, s') \in D^2} \sum_{t \in T} P(s, s', t) \left(C_i(s, s', t) - \widehat{C}_{\text{emp}} \right)^2$
- Prédiction = Krigeage : meilleure prédiction linéaire sans biais

Modèles

Modèles Spatio-temporels

Modèles sur la covariance

- Décrit la dépendance spatio-temporelle du processus
- Fonction définie-positive
- Deux modèles :
 - $C_1(s, s', t) = \sigma e^{-a_1 dx^2 - a_2 dx - a_3 dy^2 - a_4 dy - a_5 dx dy} e^{-a_6 t^2 - a_7 t}$
 dx et dy sont les distances entre s et s'
 - $C_2(s, s', t) = \sigma e^{-a_1 dx'^2 - a_2 dx' - a_3 dy'^2 - a_4 dy' - a_5 dx' dy'} e^{-a_6 t^2 - a_7 t}$
 dx' et dy' sont les distances entre s et $s' - vt$
- Estimation par moindres carrés pondérés : $L_i(a, v, \sigma) = \sum_{(s, s') \in D^2} \sum_{t \in T} P(s, s', t) \left(C_i(s, s', t) - \widehat{C}_{\text{emp}} \right)^2$
- Prédiction = Krigeage : meilleure prédiction linéaire sans biais

Modèles de covariance

Méthodologie

- 2 modèles en concurrence
- prédicteurs : observations au pas de temps précédent
- prédiction effectuée du 2 au 31 décembre 2001

Modèles de covariance

Méthodologie

- 2 modèles en concurrence
- prédicteurs : observations au pas de temps précédent
- prédiction effectuée du 2 au 31 décembre 2001

Modèles de covariance

Méthodologie

- 2 modèles en concurrence
- prédicteurs : observations au pas de temps précédent
- prédiction effectuée du 2 au 31 décembre 2001

Modèles de covariance

Estimation des paramètres

Modèle sans vitesse

$$\begin{aligned}\sigma &= 0.94 & a_4 &= 2.3 * 10^{-4} \\ a_1 &= -3.0 * 10^{-7} & a_5 &= -2.1 * 10^{-7} \\ a_2 &= 1.5 * 10^{-4} & a_6 &= 2.2 * 10^{-4} \\ a_3 &= 3.8 * 10^{-7} & a_7 &= 0.01\end{aligned}$$

Modèle avec vitesse

$$\begin{aligned}\sigma &= 1.026 & a_5 &= -1.57 * 10^{-7} \\ a_1 &= -2.85 * 10^{-8} & a_6 &= 5.14 * 10^{-5} \\ a_2 &= 5.00 * 10^{-4} & a_7 &= 0.011 \\ a_3 &= 4.43 * 10^{-7} & v_1 &= 0.48 \\ a_4 &= 2.49 * 10^{-4} & v_2 &= -0.018\end{aligned}$$

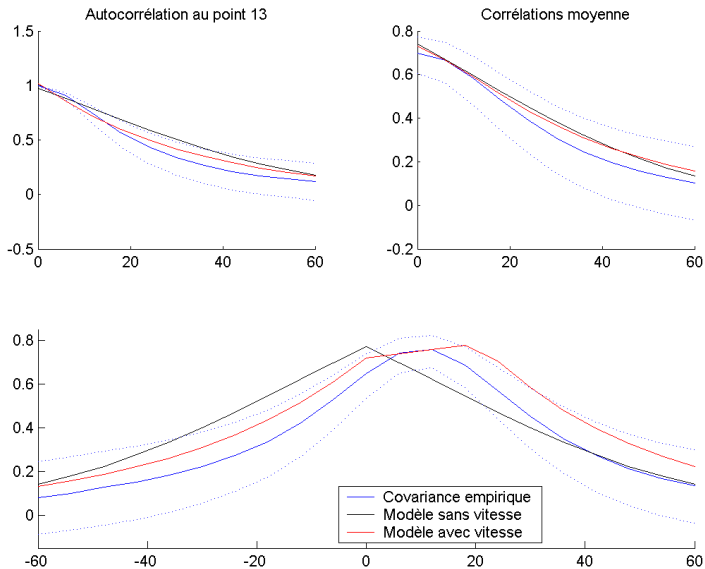


Fig.: Analyse de la modélisation

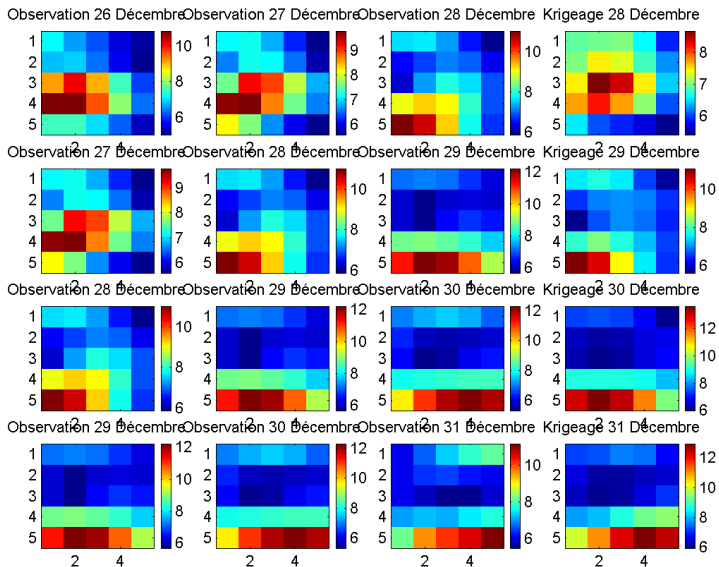


Fig.: Résultat du krigeage pour le modèle 1

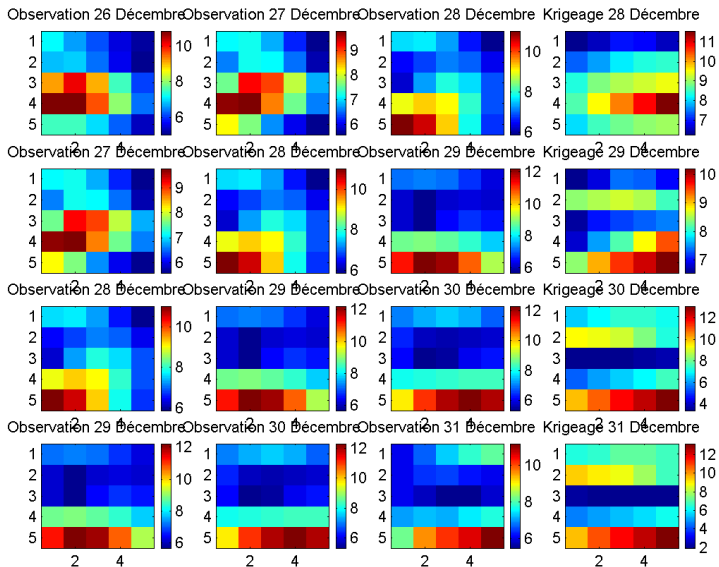


Fig.: Résultat du krigage pour le modèle 2

Filtrage

Présentation du modèle

Ici le champ observé est

$$Y(p, k) = (H_s(p, k)\cos(\Theta_m(p, k)), H_s(p, k)\sin(\Theta_m(p, k)))^T.$$

Modèles Particulaires

- Vitesse de déplacement : variable latente (cachée)
- Modèle dynamique pour l'état : $V(k) = AV(k-1) + E(k)$
- Equation d'observation :
 $Y(p, k) = Y(p - v(p, k)\Delta t, k-1) + w(p, k)$

$(E(k))$: bruit blanc $\mathcal{N}(0, \Sigma_{dyn})$

$(w(p, k\Delta t))$: bruit blanc $\mathcal{N}(0, \Sigma_{obs})$

Filtrage

Présentation du modèle

Ici le champ observé est

$$Y(p, k) = (H_s(p, k)\cos(\Theta_m(p, k)), H_s(p, k)\sin(\Theta_m(p, k)))^T.$$

Modèles Particulaires

- Vitesse de déplacement : variable latente (cachée)
- Modèle dynamique pour l'état : $V(k) = AV(k-1) + E(k)$
- Equation d'observation :
 $Y(p, k) = Y(p - v(p, k)\Delta t, k-1) + w(p, k)$

$(E(k))$: bruit blanc $\mathcal{N}(0, \Sigma_{dyn})$

$(w(p, k\Delta t))$: bruit blanc $\mathcal{N}(0, \Sigma_{obs})$

Filtrage

Présentation du modèle

Ici le champ observé est

$$Y(p, k) = (H_s(p, k)\cos(\Theta_m(p, k)), H_s(p, k)\sin(\Theta_m(p, k)))^T.$$

Modèles Particulaires

- Vitesse de déplacement : variable latente (cachée)
- Modèle dynamique pour l'état : $V(k) = AV(k-1) + E(k)$
- Equation d'observation :
 $Y(p, k) = Y(p - v(p, k)\Delta t, k-1) + w(p, k)$

$(E(k))$: bruit blanc $\mathcal{N}(0, \Sigma_{dyn})$

$(w(p, k\Delta t))$: bruit blanc $\mathcal{N}(0, \Sigma_{obs})$

Filtrage particulière

Présentation du modèle

- $A = \begin{pmatrix} A_N & 0 \\ 0 & A_N \end{pmatrix}$ $A_N(i, j) \propto \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_A)^2}\right)$ et $\lambda_A = 600 \text{ km}$
- $\Sigma_{dyn} = \begin{pmatrix} \Sigma_N^{dyn} & 0 \\ 0 & \Sigma_N^{dyn} \end{pmatrix}$ avec
 $\Sigma_N^{dyn}(i, j) = (\sigma_{dyn})^2 \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_{dyn})^2}\right)$, $\sigma_{dyn} = 2$, $\lambda_{dyn} = 600 \text{ km}$.
- $\Sigma_{obs} = \begin{pmatrix} \Sigma_N^{obs} & 0 \\ 0 & \Sigma_N^{obs} \end{pmatrix}$ avec
 $\Sigma_N^{obs}(i, j) = (\sigma_{obs})^2 \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_{obs})^2}\right)$, $\sigma_{obs} = 3$, $\lambda_{obs} = 220 \text{ km}$.

Filtrage particulaire

Présentation du modèle

- $A = \begin{pmatrix} A_N & 0 \\ 0 & A_N \end{pmatrix}$ $A_N(i, j) \propto \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_A)^2}\right)$ et $\lambda_A = 600 \text{ km}$
- $\Sigma_{dyn} = \begin{pmatrix} \Sigma_N^{dyn} & 0 \\ 0 & \Sigma_N^{dyn} \end{pmatrix}$ avec
 $\Sigma_N^{dyn}(i, j) = (\sigma_{dyn})^2 \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_{dyn})^2}\right)$, $\sigma_{dyn} = 2$, $\lambda_{dyn} = 600 \text{ km}$.
- $\Sigma_{obs} = \begin{pmatrix} \Sigma_N^{obs} & 0 \\ 0 & \Sigma_N^{obs} \end{pmatrix}$ avec
 $\Sigma_N^{obs}(i, j) = (\sigma_{obs})^2 \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_{obs})^2}\right)$, $\sigma_{obs} = 3$, $\lambda_{obs} = 220 \text{ km}$.

Filtrage particulaire

Présentation du modèle

- $A = \begin{pmatrix} A_N & 0 \\ 0 & A_N \end{pmatrix}$ $A_N(i, j) \propto \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_A)^2}\right)$ et $\lambda_A = 600 \text{ km}$
- $\Sigma_{dyn} = \begin{pmatrix} \Sigma_N^{dyn} & 0 \\ 0 & \Sigma_N^{dyn} \end{pmatrix}$ avec
 $\Sigma_N^{dyn}((i, j)) = (\sigma_{dyn})^2 \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_{dyn})^2}\right)$, $\sigma_{dyn} = 2$, $\lambda_{dyn} = 600 \text{ km}$.
- $\Sigma_{obs} = \begin{pmatrix} \Sigma_N^{obs} & 0 \\ 0 & \Sigma_N^{obs} \end{pmatrix}$ avec
 $\Sigma_N^{obs}((i, j)) = (\sigma_{obs})^2 \exp\left(-\frac{\|r_i - r_j\|^2}{(\lambda_{obs})^2}\right)$, $\sigma_{obs} = 3$, $\lambda_{obs} = 220 \text{ km}$.

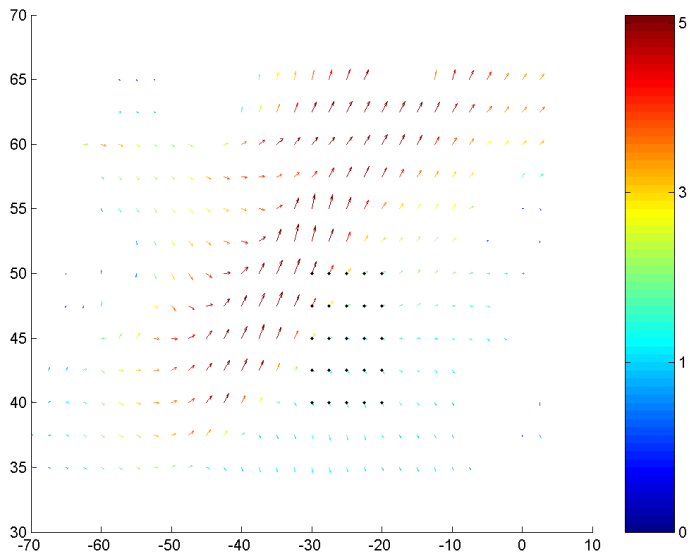


Fig.: Champs observé à $t-1$

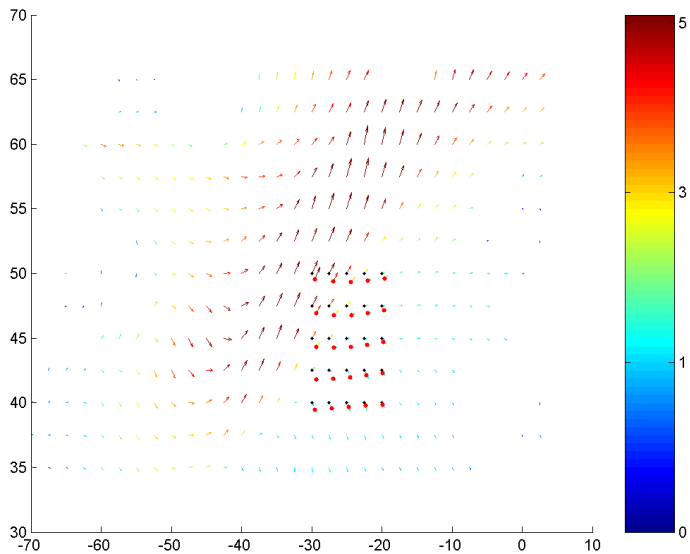


Fig.: Champs observé à t

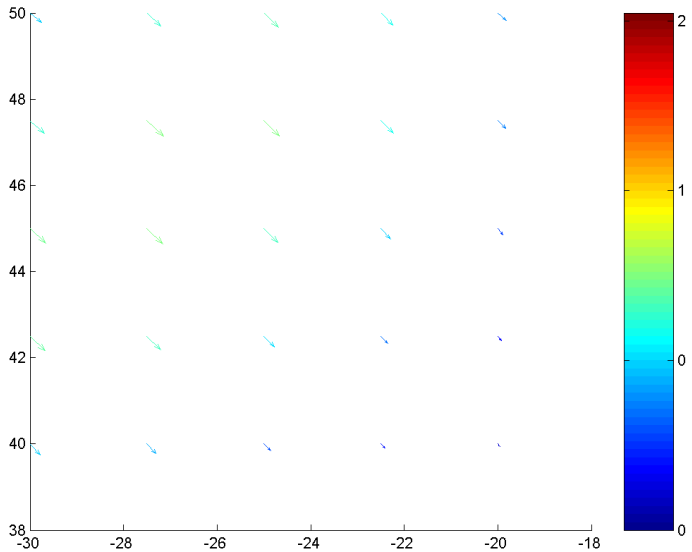


Fig.: Vitesses estimées

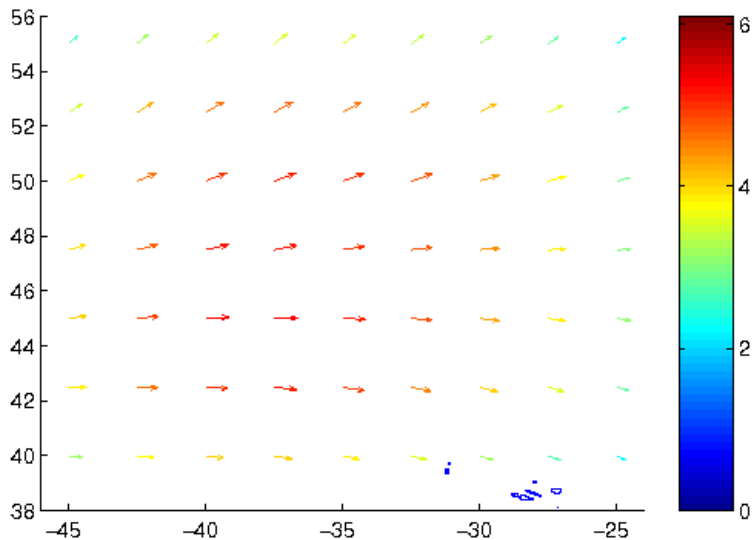


Fig.: Champs moyen des vitesses (km/h)

Comparaison des méthodes

RMSE pour le modèle sans vitesse

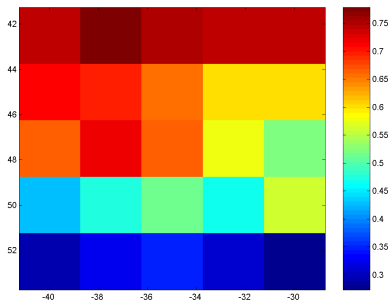


Fig.: Erreur de prédiction : cas du krigeage pour le modèle sans vitesse

Comparaison des méthodes

RMSE pour le modèle à vitesse constante

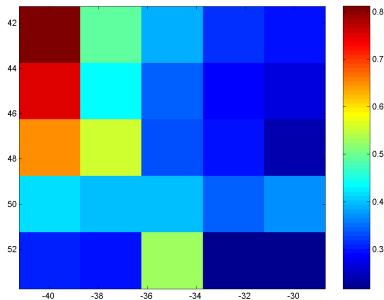


Fig.: Erreur de prédiction : cas du krigeage pour le modèle avec vitesse constante

Comparaison des méthodes

RMSE pour le modèle à vitesse variable

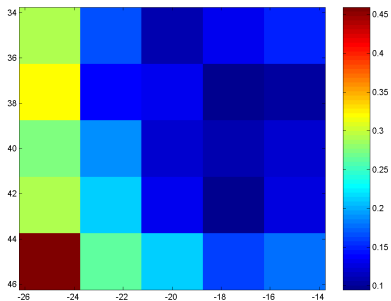


Fig.: Erreur de prédiction : cas du filtre particulaire

Conclusion - Perspectives

- Modèle : connaissance sur la structure des données
- Estimer finement les vitesses améliore beaucoup la prediction
- Questions :
 - Significativité des coefficients
 - Estimation des paramètres pour le modèle à espace d'état
 - Application à des données satellites