

Complétion et désagrégation de données agricoles pour un calcul de surplus azoté

Cécile Poisvert¹ Yann Jullian²

Université de Tours

¹GéHCo - UFR Sciences et Techniques

²CaSciModOT - UFR Sciences et Techniques

26ième journée CaSciModOT, juin 2017



**AGENCE FRANÇAISE
POUR LA BIODIVERSITÉ**
MINISTÈRE DE L'ENVIRONNEMENT

Sujet de thèse (Cécile Poisvert)

*Evolution des surplus azotés (1960-2015) :
déploiement national, étude des temps de transfert et de
l'impact du changement des pratiques agricoles*

Encadrement : Florence Curie*, Nathalie Gassama*

* GéHCo, UFR Sciences et Techniques, Université de Tours

Plan

- 1 Contexte et objectifs
 - Contexte
 - Objectifs
- 2 Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3 Les données
 - Description générale
 - Les données départementales
 - Les données communales
- 4 Implémentation

Plan

- 1** Contexte et objectifs
 - Contexte
 - Objectifs
- 2** Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3** Les données
 - Description générale
 - Les données départementales
 - Les données communales
- 4** Implémentation

Plan

1 Contexte et objectifs

- Contexte
- Objectifs

2 Le modèle

- La balance azotée et le modèle CASSIS_N

3 Les données

- Description générale
- Les données départementales
- Les données communales

4 Implémentation

Dégradation de la qualité des eaux vis-à-vis des nitrates

- Augmentation du coût de traitement de l'eau potable (*MAAF et al., 2013*).

Dégradation de la qualité des eaux vis-à-vis des nitrates

- Augmentation du coût de traitement de l'eau potable (*MAAF et al., 2013*).
- Eutrophisation (prolifération d'algues lorsque le milieu reçoit trop de substances nutritives) des eaux de surface (*Minaudo et al., 2015, Passy et al., 2013*).

Dégradation de la qualité des eaux vis-à-vis des nitrates

- Augmentation du coût de traitement de l'eau potable (*MAAF et al., 2013*).
- Eutrophisation (prolifération d'algues lorsque le milieu reçoit trop de substances nutritives) des eaux de surface (*Minaudo et al., 2015, Passy et al., 2013*).

Dans les eaux de surface continentales (rivières, lacs) :

- mortalité piscicole,
- modification de la composition des communautés algales,
- modification de la toxicité algale (*Glibert, 2017*).

Dégradation de la qualité des eaux vis-à-vis des nitrates

- Augmentation du coût de traitement de l'eau potable (*MAAF et al., 2013*).
- Eutrophisation (prolifération d'algues lorsque le milieu reçoit trop de substances nutritives) des eaux de surface (*Minaudo et al., 2015, Passy et al., 2013*).

Dans les eaux de surface continentales (rivières, lacs) :

- mortalité piscicole,
- modification de la composition des communautés algales,
- modification de la toxicité algale (*Glibert, 2017*).

Au niveau des littoraux :

- marées vertes (*Chevassus-au-Luis et al., 2012*) :

Dégradation de la qualité des eaux vis-à-vis des nitrates

- Augmentation du coût de traitement de l'eau potable (*MAAF et al., 2013*).
- Eutrophisation (prolifération d'algues lorsque le milieu reçoit trop de substances nutritives) des eaux de surface (*Minaudo et al., 2015, Passy et al., 2013*).

Dans les eaux de surface continentales (rivières, lacs) :

- mortalité piscicole,
- modification de la composition des communautés algales,
- modification de la toxicité algale (*Glibert, 2017*).

Au niveau des littoraux :

- marées vertes (*Chevassus-au-Luis et al., 2012*) :
 - impact sur le tourisme,
 - dégagement de H_2S (gaz toxique),
 - impact économique (130 M euros entre 2017 et 2021, (*Le Parisien 08/16*)).

Source des composés azotés d'un hydrosystème

- Agrosystèmes (zones cultivées) = sources majeures des composés azotés dans l'hydrosystème (*Aquilina et al., 2012*; *Heathwaite, et al., 1996*; *Öborn et al., 2003*; *Oenema, et al., 2003*).

Source des composés azotés d'un hydrosystème

- Agrosystèmes (zones cultivées) = sources majeures des composés azotés dans l'hydrosystème (*Aquilina et al., 2012; Heathwaite, et al., 1996; Öborn et al., 2003; Oenema, et al., 2003*).
- Révolution verte : politique d'intensification de l'agriculture (hausse d'utilisation des engrais minéraux de synthèse).

Source des composés azotés d'un hydrosystème

- Agrosystèmes (zones cultivées) = sources majeures des composés azotés dans l'hydrosystème (*Aquilina et al., 2012; Heathwaite, et al., 1996; Öborn et al., 2003; Oenema, et al., 2003*).
- Révolution verte : politique d'intensification de l'agriculture (hausse d'utilisation des engrais minéraux de synthèse).
- Émission d'azote à partir des sols (dont sols agricoles) sous forme de nitrate.

Source des composés azotés d'un hydrosystème

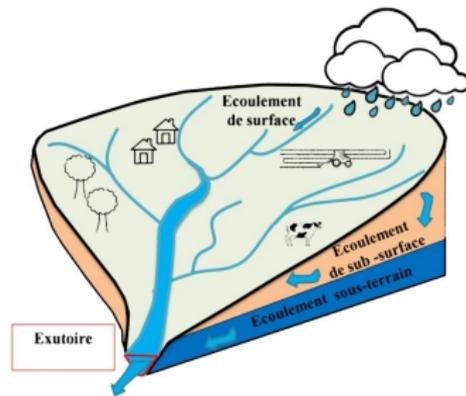
- Agrosystèmes (zones cultivées) = sources majeures des composés azotés dans l'hydrosystème (*Aquilina et al., 2012; Heathwaite, et al., 1996; Öborn et al., 2003; Oenema, et al., 2003*).
- Révolution verte : politique d'intensification de l'agriculture (hausse d'utilisation des engrais minéraux de synthèse).
- Émission d'azote à partir des sols (dont sols agricoles) sous forme de nitrate.
- Forte affinité des nitrates pour l'eau.

Source des composés azotés d'un hydrosystème

- Agrosystèmes (zones cultivées) = sources majeures des composés azotés dans l'hydrosystème (*Aquilina et al., 2012; Heathwaite, et al., 1996; Öborn et al., 2003; Oenema, et al., 2003*).
- Révolution verte : politique d'intensification de l'agriculture (hausse d'utilisation des engrais minéraux de synthèse).
- Émission d'azote à partir des sols (dont sols agricoles) sous forme de nitrate.
- Forte affinité des nitrates pour l'eau.
- Concept de bassin versant.

Source des composés azotés d'un hydrosystème

- Agrosystèmes (zones cultivées) = sources majeures des composés azotés dans l'hydrosystème (*Aquilina et al., 2012; Heathwaite, et al., 1996; Öborn et al., 2003; Oenema, et al., 2003*).
- Révolution verte : politique d'intensification de l'agriculture (hausse d'utilisation des engrais minéraux de synthèse).
- Émission d'azote à partir des sols (dont sols agricoles) sous forme de nitrate.
- Forte affinité des nitrates pour l'eau.
- Concept de bassin versant.



Problème : les relevés ne confirment pas l'effet de la législation

- Au niveau des agrosystèmes, directive nitrates (qui vise à réduire les apports d'azote) en 1991.

Problème : les relevés ne confirment pas l'effet de la législation

- Au niveau des agrosystèmes, directive nitrates (qui vise à réduire les apports d'azote) en 1991.
- Au niveau des hydrosystèmes (*eauxfrance, 2016*) :

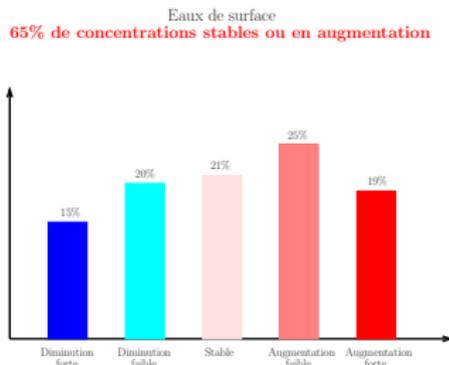


FIGURE – Évolution des concentrations moyennes en nitrates des stations en **eaux de surface** entre 1992 – 1993 et 2014 – 2015 (695 stations).

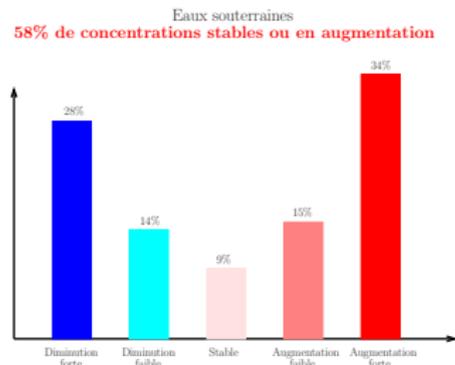


FIGURE – Évolution des concentrations moyennes en nitrates des stations en **eaux souterraines** entre 1992 – 1993 et 2014 – 2015 (582 stations).

Problème : les relevés ne confirment pas l'effet de la législation

- Au niveau des agrosystèmes, directive nitrates (qui vise à réduire les apports d'azote) en 1991.
- Au niveau des hydrosystèmes (*eauxfrance, 2016*) :

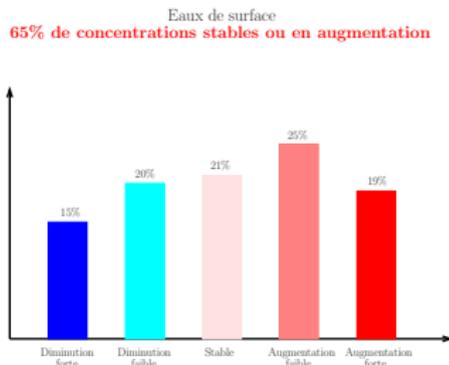


FIGURE – Évolution des concentrations moyennes en nitrates des stations en **eaux de surface** entre 1992 – 1993 et 2014 – 2015 (695 stations).

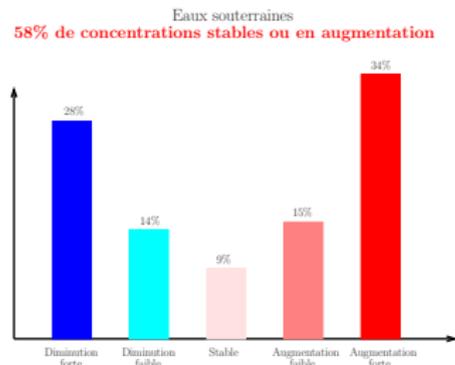


FIGURE – Évolution des concentrations moyennes en nitrates des stations en **eaux souterraines** entre 1992 – 1993 et 2014 – 2015 (582 stations).

- Temps de réponse de l'année à la décade (*Cherry et al., 2008* ; *Fovet et al., 2015* ; *Ma and Yamanaka 2016*).

Plan

- 1** Contexte et objectifs
 - Contexte
 - Objectifs
- 2** Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3** Les données
 - Description générale
 - Les données départementales
 - Les données communales
- 4** Implémentation

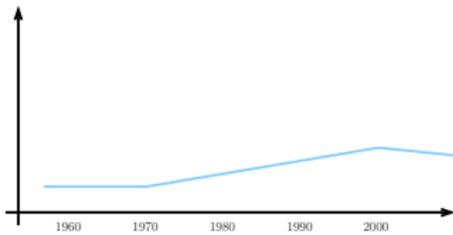
Objectifs

- Estimer la quantité d'azote potentiellement émise par les agrosystèmes.
- Estimer le temps de réponse des systèmes vis-à-vis des composés azotés.

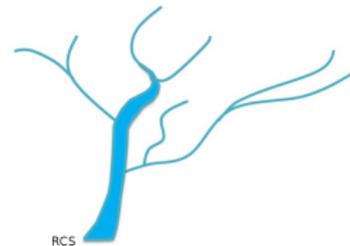
Estimation des temps de réponse des systèmes

- On dispose des **concentrations dans les rivières** au niveau des stations RCS (Réseau de Contrôle et de Surveillance).

(Unité arbitraire)



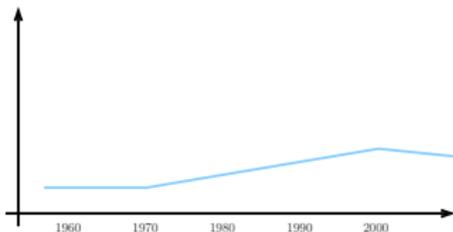
— Concentration nitrate en rivières (RCS)



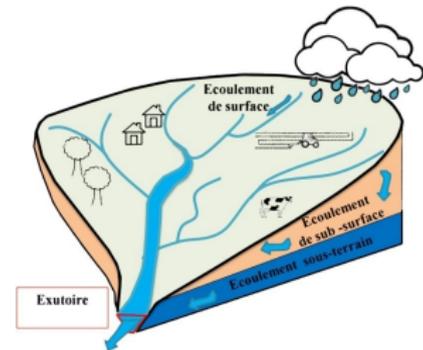
Estimation des temps de réponse des systèmes

- On dispose des **concentrations dans les rivières** au niveau des stations RCS (Réseau de Contrôle et de Surveillance).
- Pour chaque station RCS, on définit un bassin versant :
station de mesure = exutoire

(Unité arbitraire)



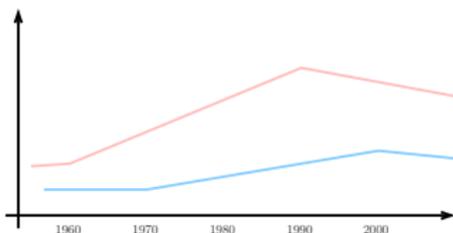
— Concentration nitrate en rivières (RCS)



Estimation des temps de réponse des systèmes

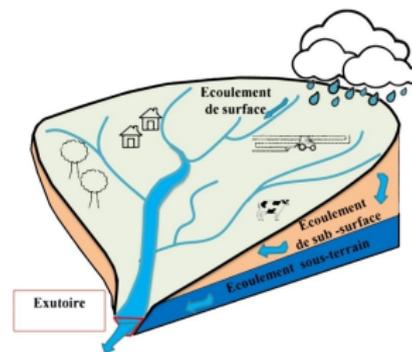
- On dispose des **concentrations dans les rivières** au niveau des stations RCS (Réseau de Contrôle et de Surveillance).
- Pour chaque station RCS, on définit un bassin versant :
station de mesure = exutoire
- On calcule le **surplus azoté** (azote potentiellement émis vers l'hydrosystème).

(Unité arbitraire)



— Surplus azoté

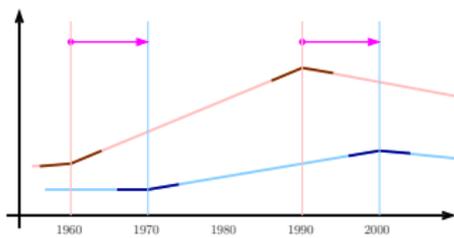
— Concentration nitrate en rivières (RCS)



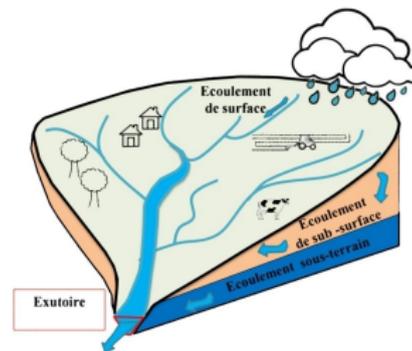
Estimation des temps de réponse des systèmes

- On dispose des **concentrations dans les rivières** au niveau des stations RCS (Réseau de Contrôle et de Surveillance).
- Pour chaque station RCS, on définit un bassin versant :
station de mesure = exutoire
- On calcule le **surplus azoté** (azote potentiellement émis vers l'hydrosystème).
- On évalue les temps de réponse en fonction des changements de tendances.

(Unité arbitraire)

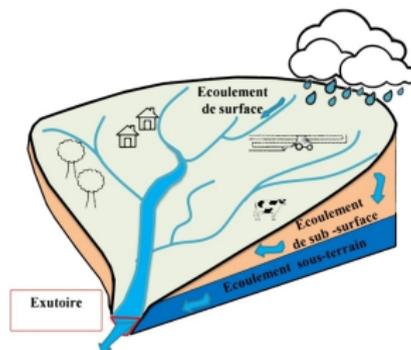


- Surplus azoté
- Changement de tendance
- Concentration nitrates en rivières (RCS)
- Changement de tendance
- Temps de réponse



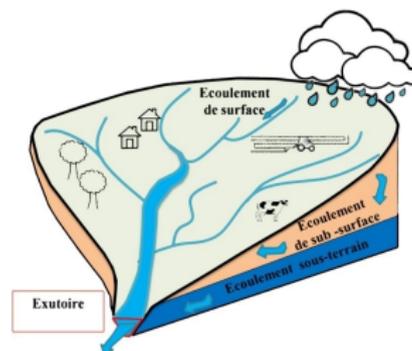
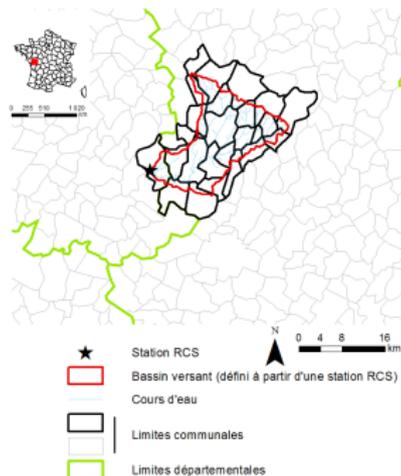
Échelle spatiale

- Pour obtenir des résultats sur des bassins versants de tailles *pertinentes*, on a combiné deux échelles d'étude :
 - échelle départementale,
 - échelle communale.

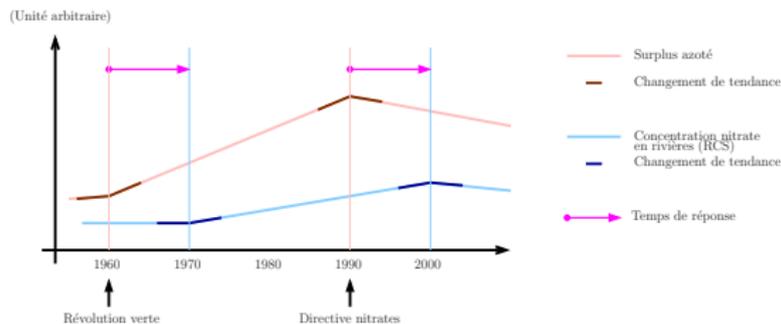


Échelle spatiale

- Pour obtenir des résultats sur des bassins versants de tailles *pertinentes*, on a combiné deux échelles d'étude :
 - échelle départementale,
 - échelle communale.

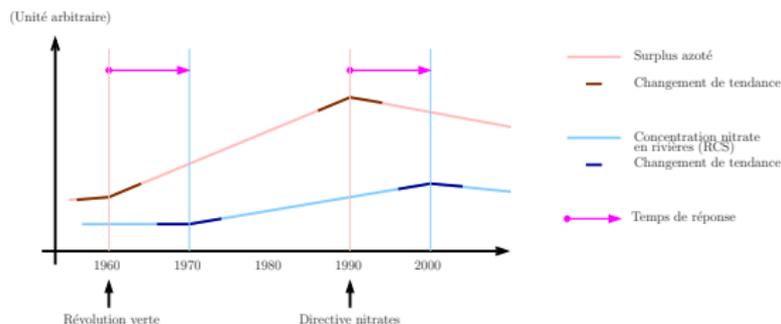


Échelle temporelle



Délimitation de la fenêtre d'étude. La chronique

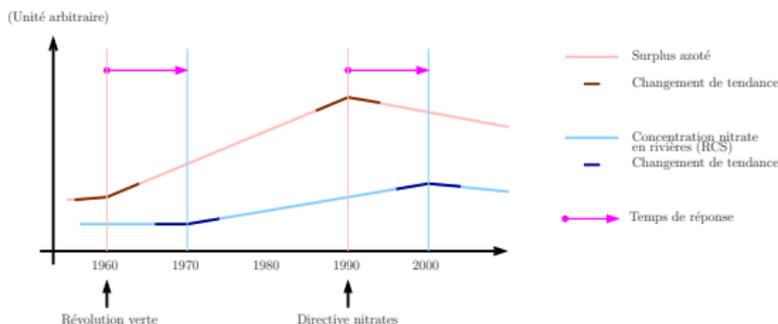
Échelle temporelle



Délimitation de la fenêtre d'étude. La chronique

- démarre avant 1960 pour mettre en évidence les effets de la révolution verte,

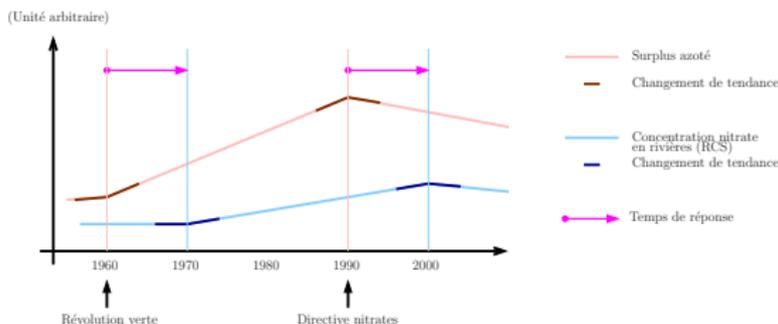
Échelle temporelle



Délimitation de la fenêtre d'étude. La chronique

- démarre avant 1960 pour mettre en évidence les effets de la révolution verte,
- va jusqu'à nos jours, incluant ainsi la directive nitrates,

Échelle temporelle



Délimitation de la fenêtre d'étude. La chronique

- démarre avant 1960 pour mettre en évidence les effets de la révolution verte,
- va jusqu'à nos jours, incluant ainsi la directive nitrates,
- doit contenir des données annuelles pour pouvoir estimer précisément les changements de tendance.

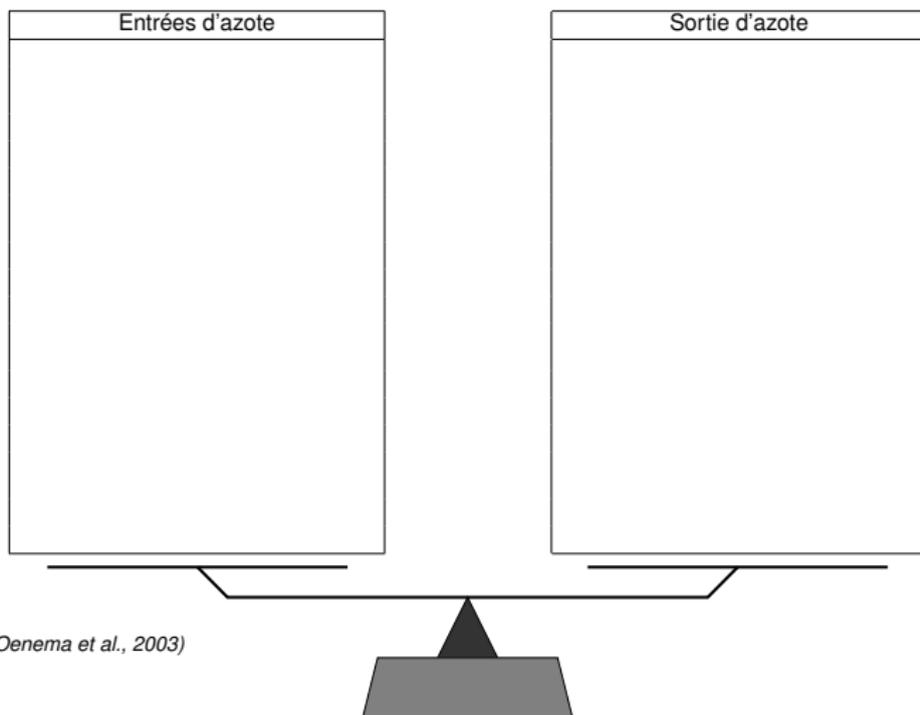
Plan

- 1 Contexte et objectifs
 - Contexte
 - Objectifs
- 2 **Le modèle**
 - **La balance azotée et le modèle CASSIS_N**
- 3 Les données
 - Description générale
 - Les données départementales
 - Les données communales
- 4 Implémentation

Plan

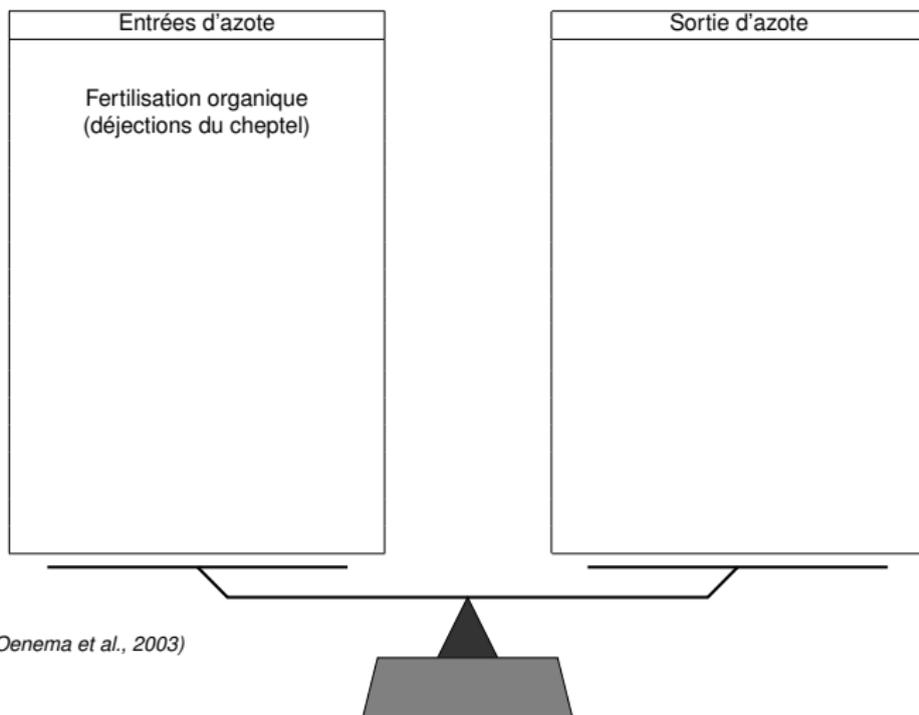
- 1 Contexte et objectifs
 - Contexte
 - Objectifs
- 2 Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3 Les données
 - Description générale
 - Les données départementales
 - Les données communales
- 4 Implémentation

Calcul de la balance azotée



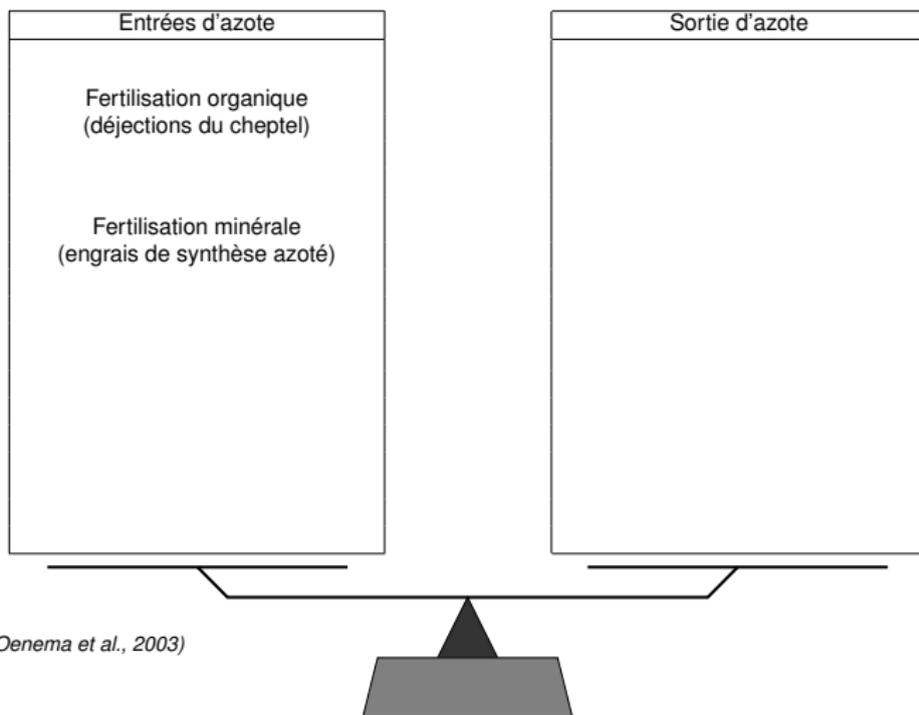
Balance azotée : (Oenema et al., 2003)

Calcul de la balance azotée



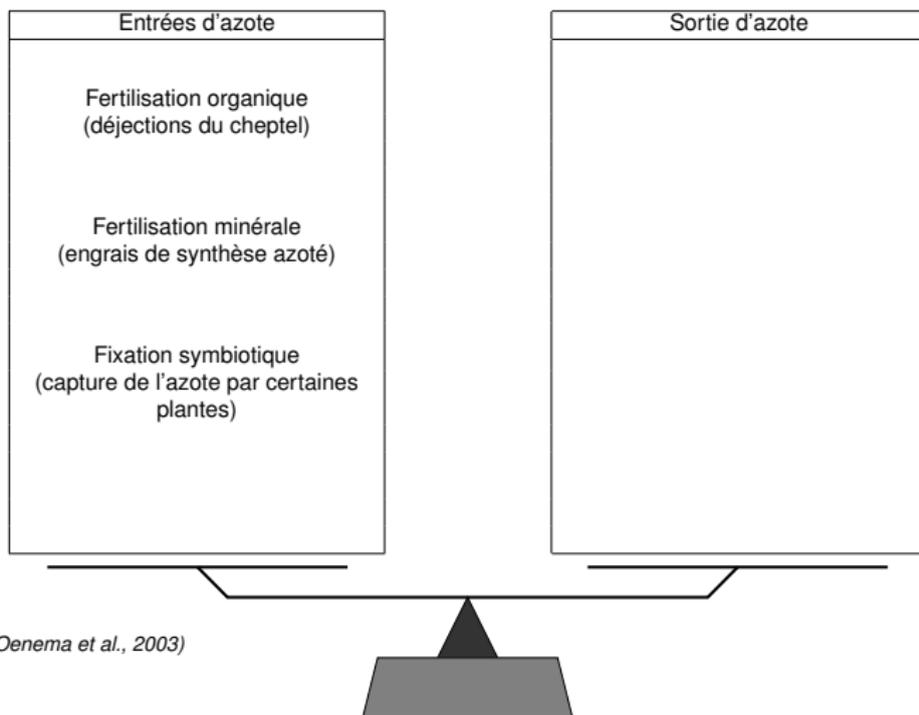
Balance azotée : (Oenema et al., 2003)

Calcul de la balance azotée

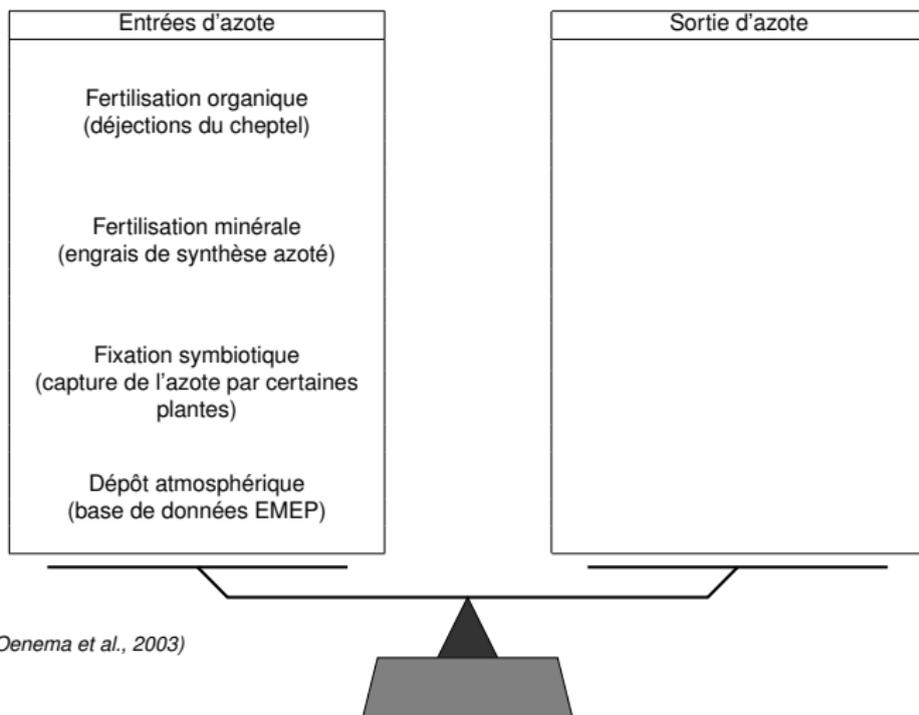


Balance azotée : (Oenema et al., 2003)

Calcul de la balance azotée

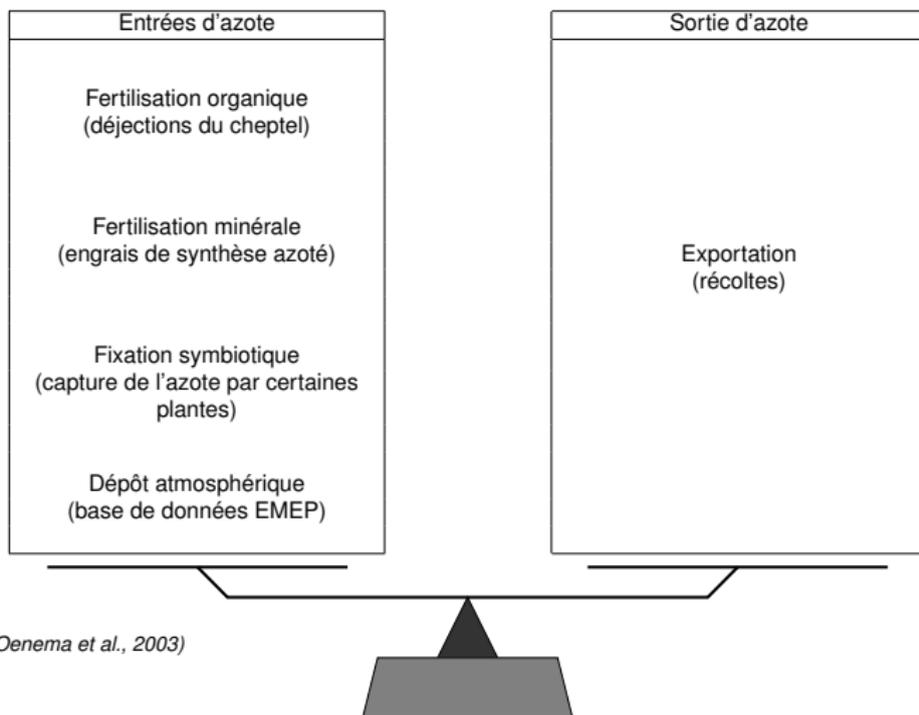


Calcul de la balance azotée

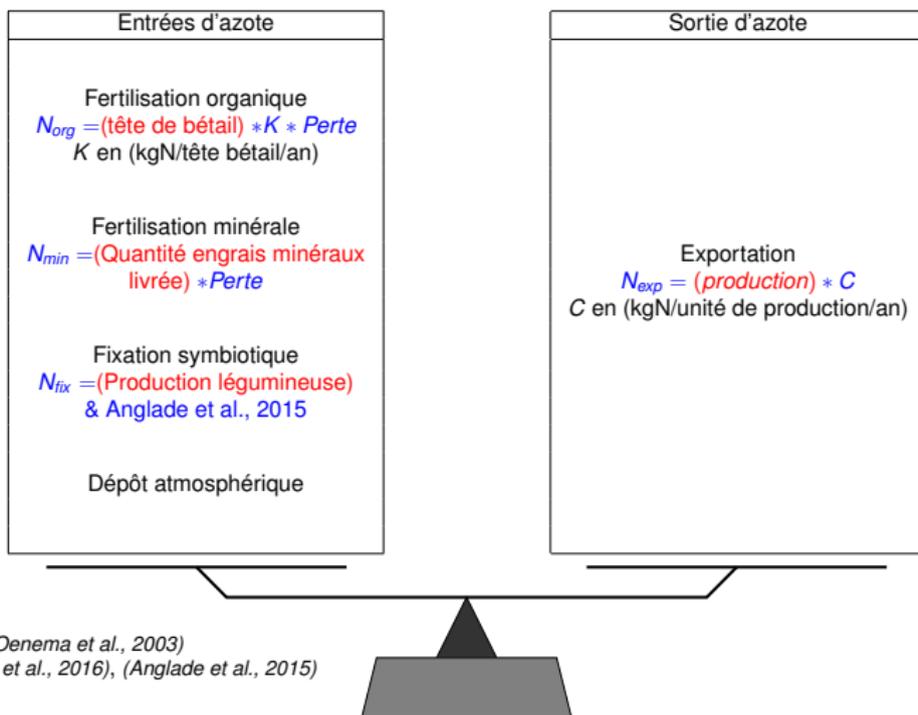


Balance azotée : (Oenema et al., 2003)

Calcul de la balance azotée



Calcul de la balance azotée



Enjeux pour les données agricoles

- Données annuelles.
- Données de production par type de culture, de bétail, données de livraison par type d'engrais.
- Données à fine échelle spatiale.
- Données sur 60 ans.

Plan

- 1 Contexte et objectifs
 - Contexte
 - Objectifs
- 2 Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3 Les données
 - Description générale
 - Les données départementales
 - Les données communales
- 4 Implémentation

Plan

- 1** Contexte et objectifs
 - Contexte
 - Objectifs
- 2** Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3** Les données
 - **Description générale**
 - Les données départementales
 - Les données communales
- 4** Implémentation

Deux simulations à échelles distinctes :

- départementale,
- communale.

Deux simulations à échelles distinctes :

- départementale,
- communale.

Trois jeux (incomplets) de données :

- régionales (SAA : Statistique Agricole Annuelle et UNIFA : Union des Industries de la Fertilisation),
- départementales (SAA et UNIFA),
- communales (RGA : Recensement Générale Agricole).

Plan

- 1** Contexte et objectifs
 - Contexte
 - Objectifs
- 2** Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3** Les données
 - Description générale
 - Les données départementales**
 - Les données communales
- 4** Implémentation

Caractéristiques des données départementales

- Disponibilité temporelle : annuelle.
- Types de données : production, surface, cheptel, engrais de synthèse.
- ~ 1.4 millions de données saisies depuis 2010 (intra-GÉHCo).

Défauts des données en vue de leur utilisation dans CASSIS_N.

- Hétérogénéité des types de données entre les années.
- Présence de données régionales sans détail départemental.
- Redécoupage des limites départementales de l'Île-de-France en 1968.
- Données manquantes entre 1917 et 2015 ($\sim 12\%$).

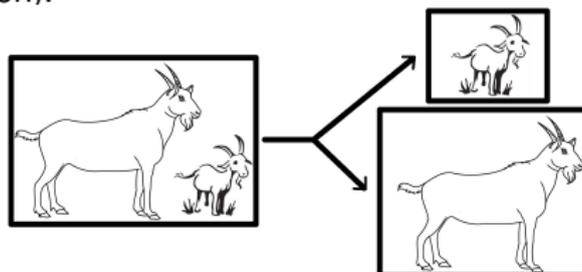


Hétérogénéité

Type de bétail	Coefficient d'excrétion	1968	1969	1970	1971	1972	1973	1974	1975		
Caprins (≤ 1 an)	5.7				400	500		445	600		
Caprins (> 1 an)	14.3					3000	3250	3600	3550		
Total caprins		2900	2900	3465							

TABLE – Répartition de caprins en Charente.

- Pour les années ≤ 1970 ,
 - on a seulement *Total caprins*,
 - on veut *Caprins (≤ 1 an)* et *Caprins (> 1 an)* (coefficients d'excrétion).



Hétérogénéité

Type de bétail	Coefficient d'excrétion	1968	1969	1970	1971	1972	1973	1974	1975	Somme partielle	Ratio
Caprins ($\leq 1 an$)	5.7				400	500		445	600	1545	0.132
Caprins ($> 1 an$)	14.3					3000	3250	3600	3550	10150	0.868
Total caprins		2900	2900	3465						11695	1

TABLE – Répartition de caprins en Charente.

- Pour les années ≤ 1970 ,
 - on a seulement *Total caprins*,
 - on veut *Caprins ($\leq 1 an$)* et *Caprins ($> 1 an$)* (coefficients d'excrétion).

- On calcule le ratio des types sur les années *pleines* (sur toute la chronique).

Hétérogénéité

Type de bétail	Coefficient d'excrétion	1968	1969	1970	1971	1972	1973	1974	1975	Somme partielle	Ratio
Caprins ($\leq 1 an$)	5.7				400	500		445	600	1545	0.132
Caprins ($> 1 an$)	14.3					3000	3250	3600	3550	10150	0.868
Total caprins		2900	2900	3465						11695	1

TABLE – Répartition de caprins en Charente.

- Pour les années ≤ 1970 ,
 - on a seulement *Total caprins*,
 - on veut *Caprins ($\leq 1 an$)* et *Caprins ($> 1 an$)* (coefficients d'excrétion).

- On calcule le ratio des types sur les années *pleines* (sur toute la chronique).
- On complète en multipliant *Total caprins* par les ratios obtenus.

Hétérogénéité

Type de bétail	Coefficient d'excrétion	1968	1969	1970	1971	1972	1973	1974	1975	Somme partielle	Ratio
Caprins ($\leq 1an$)	5.7				400	500		445	600	1545	0.132
Caprins ($> 1an$)	14.3		2527			3000	3250	3600	3550	10150	0.868
Total caprins		2900	2900	3465						11695	1

TABLE – Répartition de caprins en Charente.

- Pour les années ≤ 1970 ,
 - on a seulement *Total caprins*,
 - on veut *Caprins ($\leq 1an$)* et *Caprins ($> 1an$)* (coefficients d'excrétion).

- On calcule le ratio des types sur les années *pleines* (sur toute la chronique).

- On complète en multipliant *Total caprins* par les ratios obtenus.

Hétérogénéité

Type de bétail	Coefficient d'excrétion	1968	1969	1970	1971	1972	1973	1974	1975	Somme partielle	Ratio
Caprins ($\leq 1 an$)	5.7	383	383	458	400	500		445	600	1545	0.132
Caprins ($> 1 an$)	14.3	2527	2527	3007		3000	3250	3600	3550	10150	0.868
Total caprins		2900	2900	3465						11695	1

TABLE – Répartition de caprins en Charente.

- Pour les années ≤ 1970 ,
 - on a seulement *Total caprins*,
 - on veut *Caprins ($\leq 1 an$)* et *Caprins ($> 1 an$)* (coefficients d'excrétion).
- On calcule le ratio des types sur les années *pleines* (sur toute la chronique).
- On complète en multipliant *Total caprins* par les ratios obtenus.

Répartition des données régionales

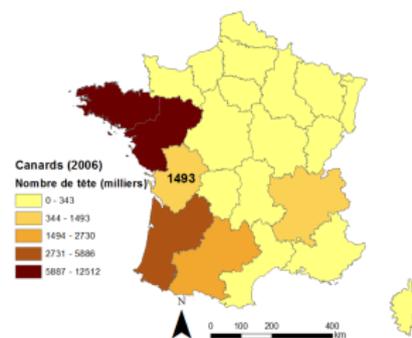
	2000	2001	2002	2003	2004		
Charente	57	59	60	59	60		
Charente-Maritime	72	72	72	72	70		
Deux Sèvres	1100	1200	1200	1280	1020		
Vienne	113	115	116	120	120		
Poitou-Charentes	908	1020	1162	1244	1222		

2005	2006	2007	2008	2009	2010
1342	1493	1561	1407	1512	1816

TABLE – Je collectionne des milliers de canards.

■ Années 2005 – 2010 :

- on a seulement le total *Poitou-Charentes*,
- on veut le détail par département.



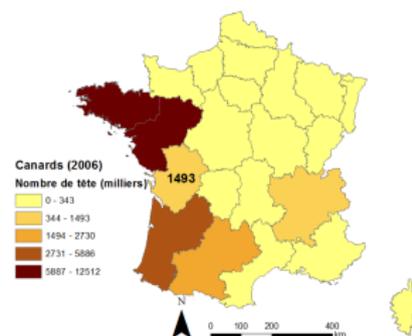
Répartition des données régionales

	2000	2001	2002	2003	2004	Somme	Ratio
Charente	57	59	60	59	60	295	0.042
Charente-Maritime	72	72	72	72	70	358	0.051
Deux Sèvres	1100	1200	1200	1280	1020	5800	0.824
Vienne	113	115	116	120	120	584	0.083
Poitou-Charentes	908	1020	1162	1244	1222	5556	1

2005	2006	2007	2008	2009	2010
1342	1493	1561	1407	1512	1816

TABLE – Je collectionne des milliers de canards.

- Années 2005 – 2010 :
 - on a seulement le total *Poitou-Charentes*,
 - on veut le détail par département.
- On calcule les ratios par département pour les années 2000 – 2004 (**années de référence**).



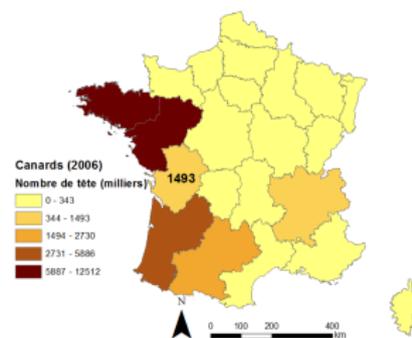
Répartition des données régionales

	2000	2001	2002	2003	2004	Somme	Ratio
Charente	57	59	60	59	60	295	0.042
Charente-Maritime	72	72	72	72	70	358	0.051
Deux Sèvres	1100	1200	1200	1280	1020	5800	0.824
Vienne	113	115	116	120	120	584	0.083
Poitou-Charentes	908	1020	1162	1244	1222	5556	1

2005	2006	2007	2008	2009	2010
1342	1493	1561	1407	1512	1816

TABLE – Je collectionne des milliers de canards.

- Années 2005 – 2010 :
 - on a seulement le total *Poitou-Charentes*,
 - on veut le détail par département.
- On calcule les ratios par département pour les années 2000 – 2004 (**années de référence**).
- On complète en multipliant *Poitou-Charentes* par les ratios obtenus.



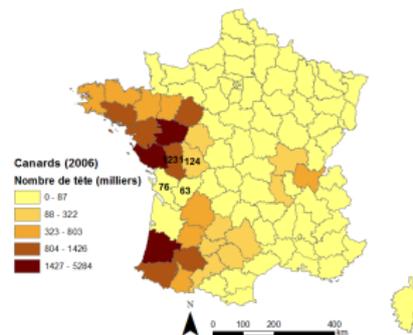
Répartition des données régionales

	2000	2001	2002	2003	2004	Somme	Ratio
Charente	57	59	60	59	60	295	0.042
Charente-Maritime	72	72	72	72	70	358	0.051
Deux Sèvres	1100	1200	1200	1280	1020	5800	0.824
Vienne	113	115	116	120	120	584	0.083
Poitou-Charentes	908	1020	1162	1244	1222	5556	1

2005	2006	2007	2008	2009	2010
	63				
1342	1493	1561	1407	1512	1816

TABLE – Je collectionne des milliers de canards.

- Années 2005 – 2010 :
 - on a seulement le total *Poitou-Charentes*,
 - on veut le détail par département.
- On calcule les ratios par département pour les années 2000 – 2004 (**années de référence**).
- On complète en multipliant *Poitou-Charentes* par les ratios obtenus.



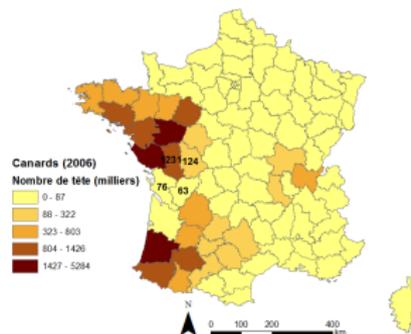
Répartition des données régionales

	2000	2001	2002	2003	2004	Somme	Ratio
Charente	57	59	60	59	60	295	0.042
Charente-Maritime	72	72	72	72	70	358	0.051
Deux Sèvres	1100	1200	1200	1280	1020	5800	0.824
Vienne	113	115	116	120	120	584	0.083
Poitou-Charentes	908	1020	1162	1244	1222	5556	1

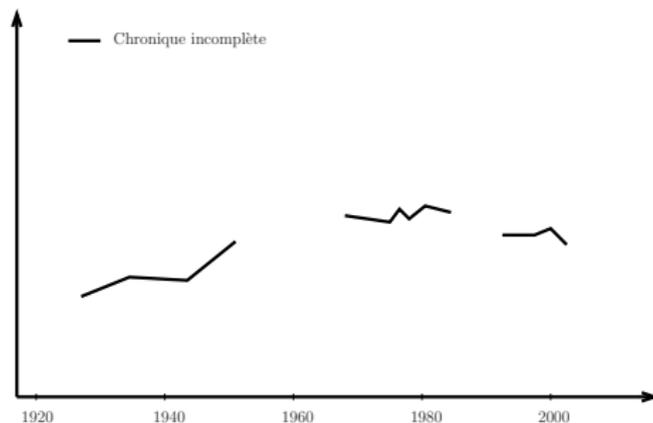
2005	2006	2007	2008	2009	2010
56	63	65	59	63	76
68	76	79	72	77	92
1106	1231	1287	1160	1246	1497
111	124	130	117	125	151
1342	1493	1561	1407	1512	1816

TABLE – Je collectionne des milliers de canards.

- Années 2005 – 2010 :
 - on a seulement le total *Poitou-Charentes*,
 - on veut le détail par département.
- On calcule les ratios par département pour les années 2000 – 2004 (**années de référence**).
- On complète en multipliant *Poitou-Charentes* par les ratios obtenus.

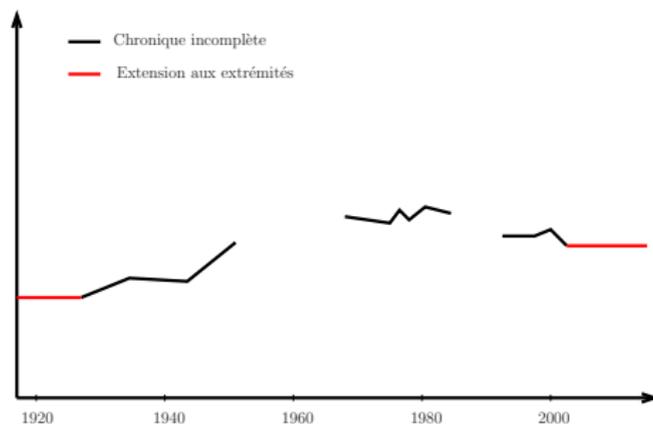


Complétion de chroniques



Deux types de complétion différents :

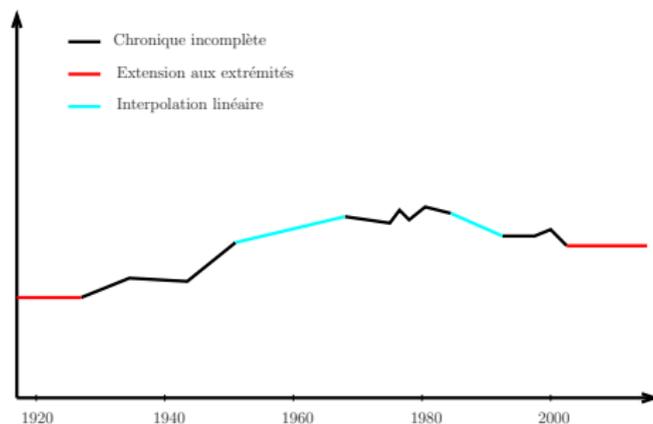
Complétion de chroniques



Deux types de complétion différents :

- On étend les premières et dernières valeurs trouvées jusqu'aux extrémités de la chronique.

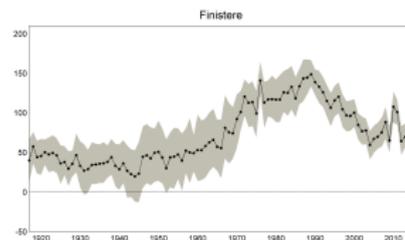
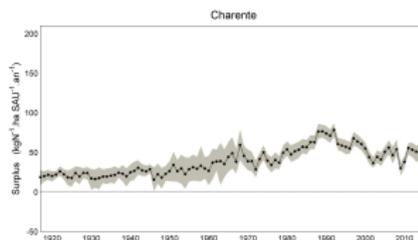
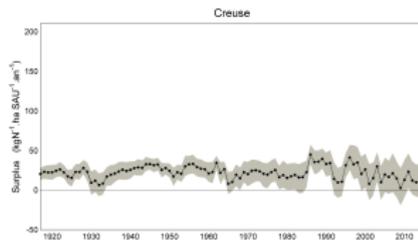
Complétion de chroniques



Deux types de complétion différents :

- On étend les premières et dernières valeurs trouvées jusqu'aux extrémités de la chronique.
- On fait des interpolations linéaires au milieu de la chronique.

Exemples de résultats



Plan

- 1** Contexte et objectifs
 - Contexte
 - Objectifs
- 2** Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3** Les données
 - Description générale
 - Les données départementales
 - **Les données communales**
- 4** Implémentation

Caractéristiques des données communales

- Disponibilité temporelle : ponctuelle (1955, 1970, 1979, 1988, 2000, 2010).
- Types de données : surface, cheptel. Pas de production, pas d'engrais.
- ~ 800 000 données saisies depuis 2015 (intra-GÉHCo).

Caractéristiques des données communales

- Disponibilité temporelle : ponctuelle (1955, 1970, 1979, 1988, 2000, 2010).
- Types de données : surface, cheptel. Pas de production, pas d'engrais.
- ~ 800 000 données saisies depuis 2015 (intra-GéHCo).

Défauts des données en vue de leur utilisation dans CASSIS_N.

- Hétérogénéité des types de données entre les années.
- Evolution fréquente des limites communales.
- Une partie des données est payante.
- Données manquantes entre 1955 et 2015 (~ 95%).

Hétérogénéité

Seulement sur les années du RGA.

Type de bétail	Coefficient d'excrétion	Brie	Cognac	Charente	
Porcs (20 à 50kg)	0.59			0	
Truies (> 50kg)	22.5			200	
Autres porcs (> 50kg)	4.2			7240	
Total porcins		526	12	7440	

TABLE – Répartition de porcins sur les communes de Charente en 1955.

- Pour les communes de Charente en 1955,
 - on a seulement *Total porcins*,
 - on veut *Porcs (20 à 50kg)*, *Truies (> 50kg)* et *Autres porcs (> 50kg)* (coefficients d'excrétion).

Hétérogénéité

Seulement sur les années du RGA.

Type de bétail	Coefficient d'excrétion	Brie	Cognac	Charente	Ratio
Porcs (20 à 50kg)	0.59			0	0
Truies (> 50kg)	22.5			200	0.027
Autres porcs (> 50kg)	4.2			7240	0.973
Total porcins		526	12	7440	1

TABLE – Répartition de porcins sur les communes de Charente en 1955.

- Pour les communes de Charente en 1955,
 - on a seulement *Total porcins*,
 - on veut *Porcs (20 à 50kg)*, *Truies (> 50kg)* et *Autres porcs (> 50kg)* (coefficients d'excrétion).
- On calcule le ratio des types sur le *département*.

Hétérogénéité

Seulement sur les années du RGA.

Type de bétail	Coefficient d'excrétion	Brie	Cognac	Charente	Ratio
Porcs (20 à 50kg)	0.59			0	0
Truies (> 50kg)	22.5			200	0.027
Autres porcs (> 50kg)	4.2			7240	0.973
Total porcins		526	12	7440	1

TABLE – Répartition de porcins sur les communes de Charente en 1955.

- Pour les communes de Charente en 1955,
 - on a seulement *Total porcins*,
 - on veut *Porcs (20 à 50kg)*, *Truies (> 50kg)* et *Autres porcs (> 50kg)* (coefficients d'excrétion).
- On calcule le ratio des types sur le *département*.
- On complète en multipliant *Total porcins* par les ratios obtenus.

Hétérogénéité

Seulement sur les années du RGA.

Type de bétail	Coefficient d'excrétion	Brie	Cognac	Charente	Ratio
Porcs (20 à 50kg)	0.59			0	0
Truies (> 50kg)	22.5			200	0.027
Autres porcs (> 50kg)	4.2	512		7240	0.973
Total porcins		526	12	7440	1

TABLE – Répartition de porcins sur les communes de Charente en 1955.

- Pour les communes de Charente en 1955,
 - on a seulement *Total porcins*,
 - on veut *Porcs (20 à 50kg)*, *Truies (> 50kg)* et *Autres porcs (> 50kg)* (coefficients d'excrétion).
- On calcule le ratio des types sur le *département*.
- On complète en multipliant *Total porcins* par les ratios obtenus.

Hétérogénéité

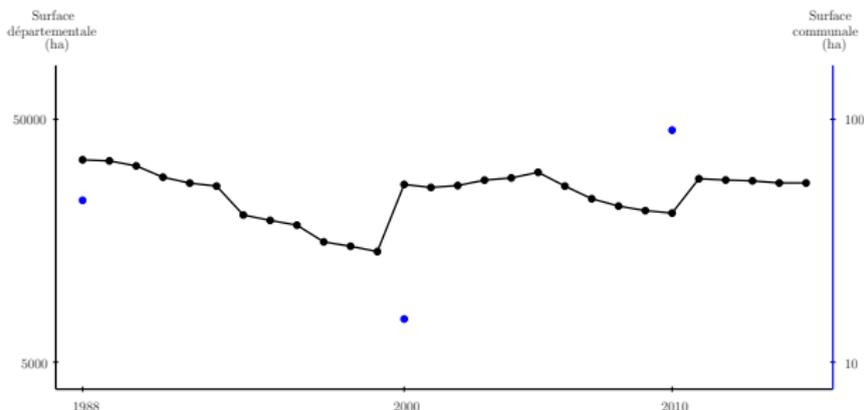
Seulement sur les années du RGA.

Type de bétail	Coefficient d'excrétion	Brie	Cognac	Charente	Ratio
Porcs (20 à 50kg)	0.59	0	0	0	0
Truies (> 50kg)	22.5	14	0	200	0.027
Autres porcs (> 50kg)	4.2	512	12	7240	0.973
Total porcins		526	12	7440	1

TABLE – Répartition de porcins sur les communes de Charente en 1955.

- Pour les communes de Charente en 1955,
 - on a seulement *Total porcins*,
 - on veut *Porcs (20 à 50kg)*, *Truies (> 50kg)* et *Autres porcs (> 50kg)* (coefficients d'excrétion).
- On calcule le ratio des types sur le *département*.
- On complète en multipliant *Total porcins* par les ratios obtenus.

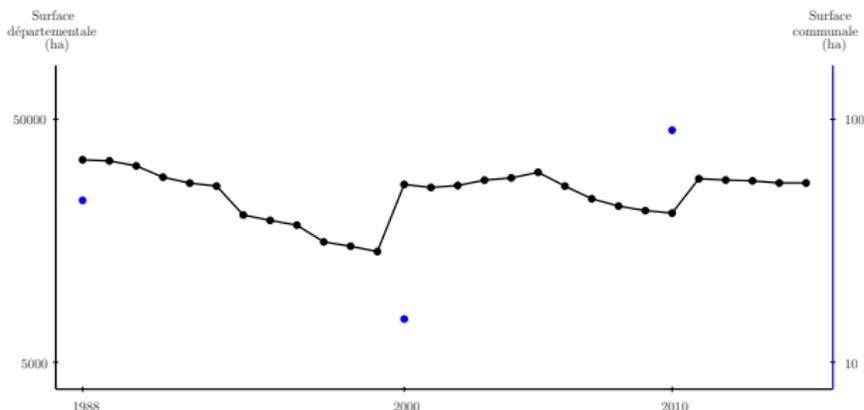
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26										96					
Ratio ($\times 10^{-3}$)																

- On dispose des données départementales sur toutes les années et des données communales sur les années du RGA.

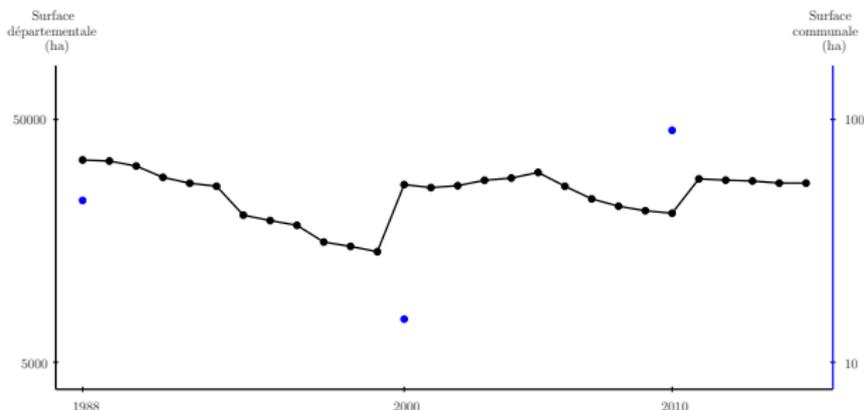
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26										96					
Ratio ($\times 10^{-3}$)	0.69										2.94					

- On calcule les ratios département / commune sur les années du RGA.

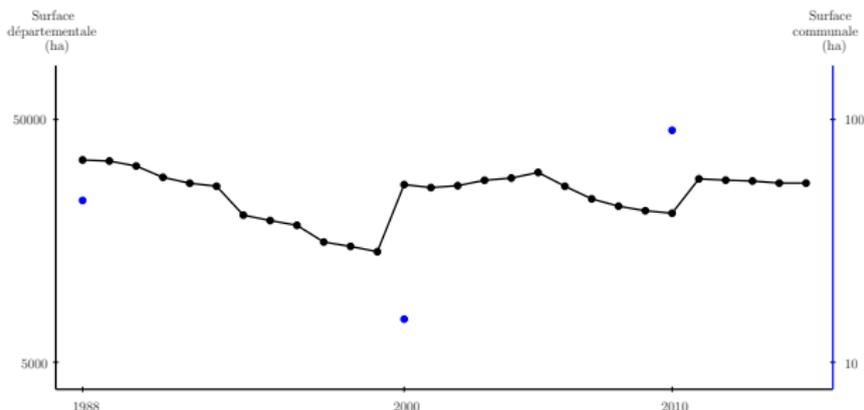
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26										96					
Ratio ($\times 10^{-3}$)	0.69										2.94					

- Extension : on applique le dernier ratio à toutes les années qui suivent.

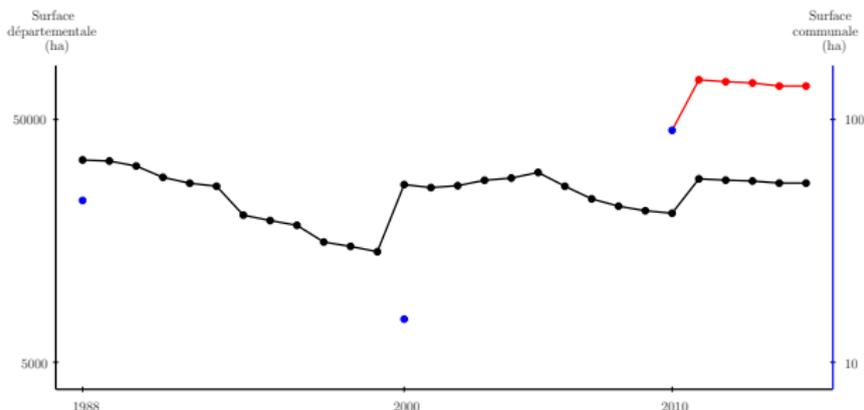
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26										96		113.99			
Ratio ($\times 10^{-3}$)	0.69										2.94					

- Extension : on applique le dernier ratio à toutes les années qui suivent.

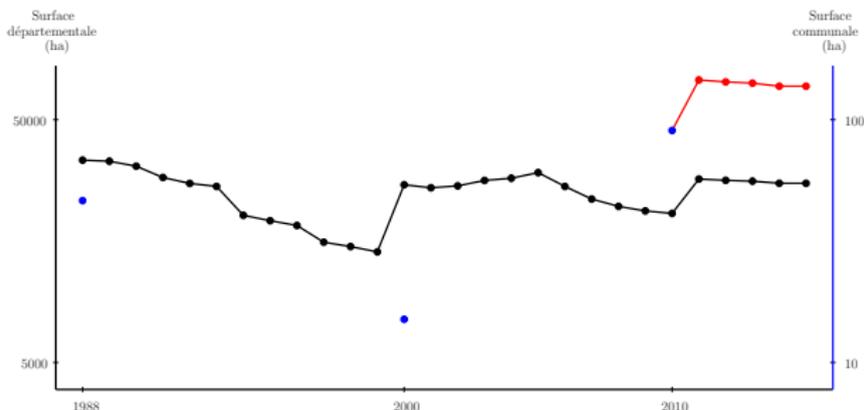
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26										96	114.71	113.99	113.54	112.39	112.39
Ratio ($\times 10^{-3}$)	0.69										2.94					

- Extension : on applique le dernier ratio à toutes les années qui suivent.

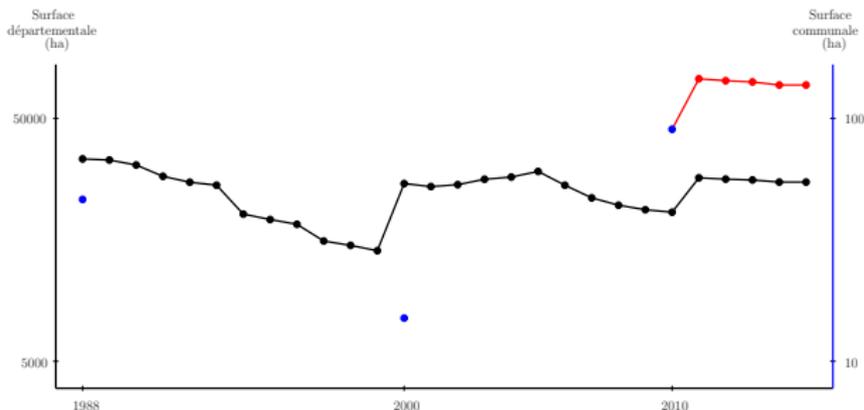
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26										96	114.71	113.99	113.54	112.39	112.39
Ratio ($\times 10^{-3}$)	0.69	0.91	1.14	1.36	1.59	1.81	2.04	2.26	2.49	2.72	2.94					

- Complétion : on interpole les **ratios**.

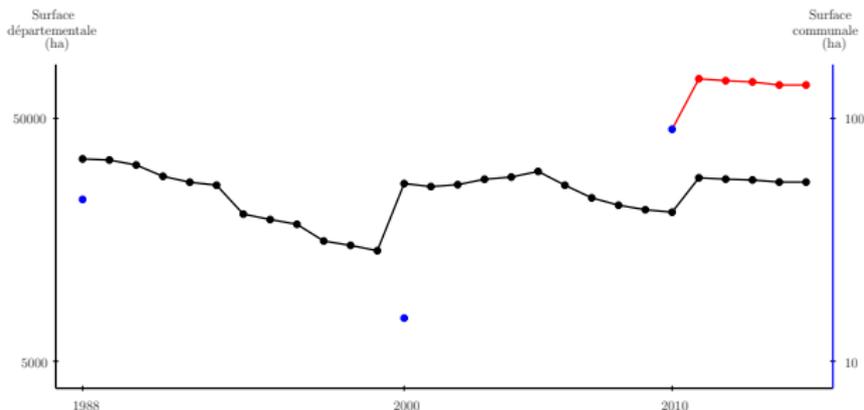
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26										96	114.71	113.99	113.54	112.39	112.39
Ratio ($\times 10^{-3}$)	0.69	0.91	1.14	1.36	1.59	1.81	2.04	2.26	2.49	2.72	2.94					

- Complétion : on multiplie chaque nouveau ratio par la valeur départementale.

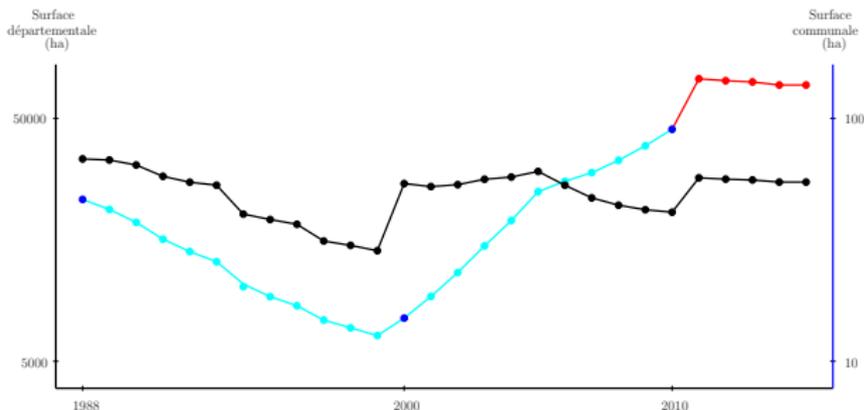
Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26					72.87					96	114.71	113.99	113.54	112.39	112.39
Ratio ($\times 10^{-3}$)	0.69	0.91	1.14	1.36	1.59	1.81	2.04	2.26	2.49	2.72	2.94					

- Complétion : on multiplie chaque nouveau ratio par la valeur départementale.

Complétion des données de cheptel et surface



Année	2000										2010					
Département	37930	37373	37730	38731	39145	40185	37631	35288	33931	33089	32637	38997	38754	38600	38210	38210
Commune	26	34.05	42.89	52.76	62.16	72.87	76.73	79.92	84.50	89.86	96	114.71	113.99	113.54	112.39	112.39
Ratio ($\times 10^{-3}$)	0.69	0.91	1.14	1.36	1.59	1.81	2.04	2.26	2.49	2.72	2.94					

- Complétion : on multiplie chaque nouveau ratio par la valeur départementale.

Calculs de la production et des engrais de synthèse

- Les productions départementales sont réparties sur les communes au prorata des surfaces des cultures considérées.

Calculs de la production et des engrais de synthèse

- Les productions départementales sont réparties sur les communes au prorata des surfaces des cultures considérées.
- Les engrais départementaux sont répartis sur les communes au prorata des SAU (toute surface destinée à l'agriculture).

Plan

- 1 Contexte et objectifs
 - Contexte
 - Objectifs
- 2 Le modèle
 - La balance azotée et le modèle CASSIS_N
- 3 Les données
 - Description générale
 - Les données départementales
 - Les données communales
- 4 Implémentation

Programme initial

Écrit en MATLAB (Cécile Poisvert, Florence Curie, . . .) pour les *départements seulement*.

- Première phase : une seule simulation.

Programme initial

Écrit en MATLAB (Cécile Poisvert, Florence Curie, . . .) pour les *départements seulement*.

- Première phase : une seule simulation.
- Deuxième phase : 200 simulations pour pouvoir calculer les imprécisions sur les résultats.

Programme initial

Écrit en MATLAB (Cécile Poisvert, Florence Curie, . . .) pour les *départements seulement*.

- Première phase : une seule simulation.
- Deuxième phase : 200 simulations pour pouvoir calculer les imprécisions sur les résultats.

Problèmes :

- interruptions fréquentes (nécessité de le relancer régulièrement),
- lent,

~ 1 mois pour arriver au bout.

Programme initial

Écrit en MATLAB (Cécile Poisvert, Florence Curie, . . .) pour les *départements seulement*.

- Première phase : une seule simulation.
- Deuxième phase : 200 simulations pour pouvoir calculer les imprécisions sur les résultats.

Problèmes :

- interruptions fréquentes (nécessité de le relancer régulièrement),
- lent,

~ 1 mois pour arriver au bout.

- Très peu de vérifications des données d'entrées.

Réécriture

Entièrement réécrit en C++. Pour les départements :

Réécriture

Entièrement réécrit en C++. Pour les départements :

- 7 minutes sous linux,
- 1h13 sous windows avec cygwin (mini-environnement linux sous windows).

Cause : accès disque (lecture/écriture) très lents.

Réécriture

Entièrement réécrit en C++. Pour les départements :

- 7 minutes sous linux,
- 1h13 sous windows avec cygwin (mini-environnement linux sous windows).
Cause : accès disque (lecture/écriture) très lents.
- Lenteur d'écriture sous cygwin non réglée.
- Pas de parallélisation.

Réécriture

Entièrement réécrit en C++. Pour les départements :

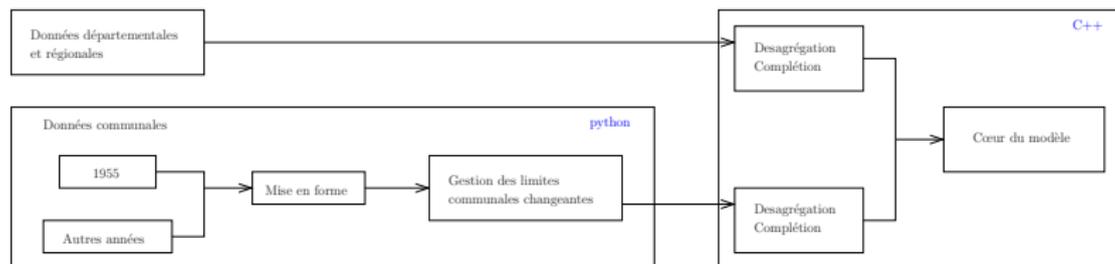
- 7 minutes sous linux,
- 1h13 sous windows avec cygwin (mini-environnement linux sous windows).
Cause : accès disque (lecture/écriture) très lents.
- Lenteur d'écriture sous cygwin non réglée.
- Pas de parallélisation.

Remarque : temps de calculs obtenus sur mon portable (linux) et sur le serveur emprisonnant les données communales (windows).

Organigramme



Organigramme



Adaptation du code pour les communes

~ 36000 communes contre ~ 100 départements.

Adaptation du code pour les communes

~ 36000 communes contre ~ 100 départements.

Adaptation naïve :

- le chargement de toutes les données communales provoque un dépassement de mémoire sur nos machines (8 Go).

Adaptation du code pour les communes

~ 36000 communes contre ~ 100 départements.

Adaptation naïve :

- le chargement de toutes les données communales provoque un dépassement de mémoire sur nos machines (8 Go).

⇒ Réorganisation du code pour traiter les communes d'un seul département à la fois. Le nouveau code tourne à nouveau sur nos machines.

Adaptation du code pour les communes

~ 36000 communes contre ~ 100 départements.

Adaptation naïve :

- le chargement de toutes les données communales provoque un dépassement de mémoire sur nos machines (8 Go).

⇒ Réorganisation du code pour traiter les communes d'un seul département à la fois. Le nouveau code tourne à nouveau sur nos machines.

Temps de calcul raisonnable sans optimisation additionnelle :

- 38m sous linux,
- 2h31 sous windows avec cygwin.

(Calculs communaux restreints par rapport aux départements).

Vérification de code

Test unitaire

Vérification de code

Test unitaire

Pour une fonction F donnée du code, on écrit une autre fonction, la *fonction test* qui

- appelle F dans les cas *généraux* et *critiques*,

Cas critiques : (ici) données manquantes ou nulles.

Vérification de code

Test unitaire

Pour une fonction F donnée du code, on écrit une autre fonction, la *fonction test* qui

- appelle F dans les cas *généraux* et *critiques*,
- compare la sortie F avec des résultats renseignés manuellement.

Cas critiques : (ici) données manquantes ou nulles.

Vérification de code

Test unitaire

Pour une fonction F donnée du code, on écrit une autre fonction, la *fonction test* qui

- appelle F dans les cas *généraux* et *critiques*,
- compare la sortie F avec des résultats renseignés manuellement.

Cas critiques : (ici) données manquantes ou nulles.

- Outil : CPPUNIT.

Vérification de code

Test unitaire

Pour une fonction F donnée du code, on écrit une autre fonction, la *fonction test* qui

- appelle F dans les cas *généraux* et *critiques*,
- compare la sortie F avec des résultats renseignés manuellement.

Cas critiques : (ici) données manquantes ou nulles.

- Outil : CPPUNIT.
- 45 fonctions testées sur 96 fonctions au total.

cecile.poisvert@univ-tours.fr
yann.jullian@univ-tours.fr