PROJET FLASH
ANR-09-SYSC-004

ANNE JOHANNET

Gardon d'Alès 1907

FLOOD FORECASTING WITH MACHINE LEARNING, DATA ASSIMILATION AND SEMI-PHYSICAL MODELING

CONTEXTE SOCIÉTAL

Les crues rapides

- Montée de quelques dizaines de m^3/s à plus de 2000 m^3/s en quelques heures
- Épisode de 2002 dans le Gard (27 victimes, 1,2 milliards d'euros)
- Épisode de 2010 dans le Var (25 victimes, 1,5 milliards d'euros)

Pont SNCF de Ners : 1958 (gauche), 2002 (droite)

Les crues de nappes

- Crue de la Somme en 2001 (100 millions d'euros)

La réponse institutionnelle

- Restructuration en SPC (Services de Prévision des crues)
- Création du SCHAPI
- Mise en place du service vigicrues (http://www.vigicrues.gouv.fr/)

OBJECTIF SOCIÉTAL

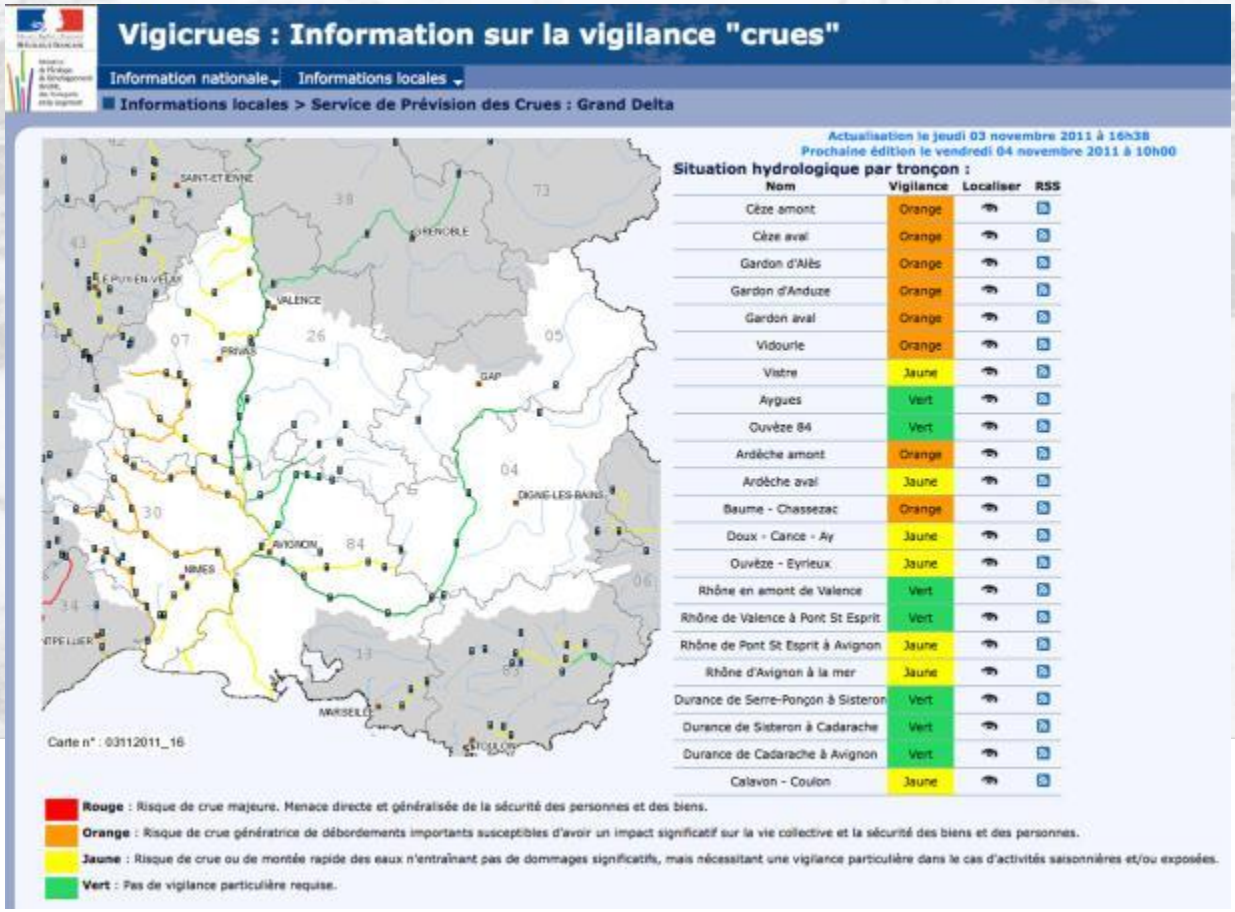
Prévision de l'aléa

- Définition de modèles de prévisions, temps réel, à très courte échéance (5h) pour les crues rapides.
- Définition de modèles de prévision à plus longue échéance (journée, semaine), pour les crues de nappe.

- Ces modèles ont pour finalité d'être intégrés à la plate-forme temps réel du SCHAPI afin de permettre aux prévisionnistes d'alimenter la carte vigicrues.

OBJECTIF SOCIÉTAL

- **Vigilance fiable accessible à la population, aux services de l'état et aux média**



CONTEXTE SCIENTIFIQUE (1/2)

Hydrologie : crues rapides

- Méconnaissance du processus de génération des crues rapides
- Couplage hydrométéorologique entrainant une grande complexité
- Bases de données entachées de bruits de mesure et d'incertitudes considérables
- Grande incertitude sur la prévision du forçage atmosphérique

Hydrogéologie : crues de nappes

- Phénomène difficile à observer (souterrain) et dépendant de la structure géologique des terrains

CONTEXTE SCIENTIFIQUE (2/2)

Modélisation hydrologique :

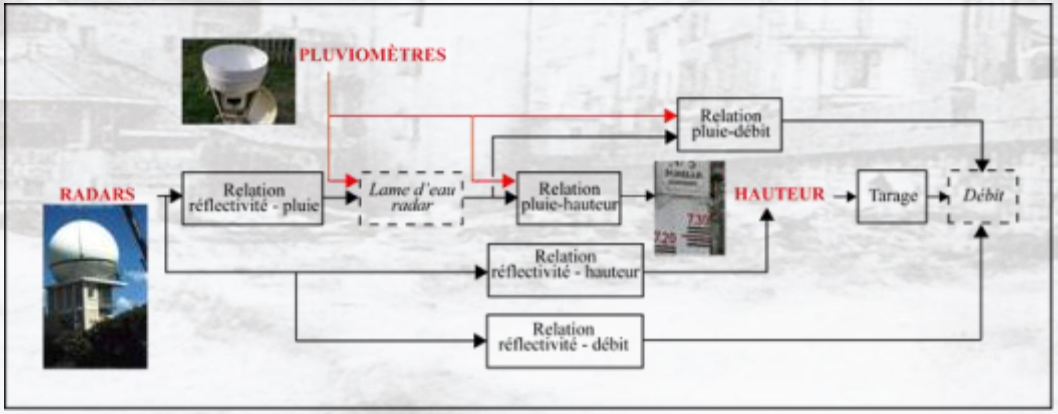
- Prise en compte du rôle du sol : humidité
- Initialisation du modèle et évolution au cours de l'événement

Apprentissage statistique

- Boîtes noires
  - Méfiance des hydrologues vis-à-vis des capacités de généralisation (événement extrême ?)
- Dilemme biais-variance/Méthodes de régularisation
- Modélisation semi-physique
- Réseaux de neurones et machines à vecteurs supports (SVM), dynamiques et non dynamiques

OBJECTIFS SCIENTIFIQUES (1/3)

- **Modélisation systémique de la chaine hydro-météorologique :**



- Diversité des sources de données (redondance-fiabilité)
- Bassin versant du Gardon à Anduze et jusqu'à Remoulins

OBJECTIFS SCIENTIFIQUES (2/3)

Changements d'échelles spatiales et temporelles

- Généralisation à la Cèze et à l'Ardèche
- Généralisation à la Somme

Crue de l'Ardèche au Pont d'Arc en 1890

OBJECTIFS SCIENTIFIQUES (3/3)

- **Développer la modélisation semi-physique pour estimer l'état d'humidité du sol**
- **Intégrer l'assimilation de données au Réseau de Neurones**
  - Mise à jour des paramètres (adaptativité)
  - Mise à jour des variables d'entrée (précipitation, erreur, débit précédent)
  - Assimiler des informations issues du modèle hydrologique
- **Résoudre les difficultés posées par l'apprentissage des SVM dynamiques**
  - Application à la relation pluie-débit à Anduze

CONSORTIUM

- **Partenaire 1 (coordinateur) : ARMINES/Ecole des Mines d'Alès**
  - (3 permanents, 2 doctorants, 2 stages de Master 2)
- **Partenaire 2 : ESPCI ParisTech**
  - (2 permanents, 1 chercheur non permanent,1 post-doc)
- **Partenaire 3 : EDYTEM**
  - (1 permanent, 1 post-doc)
- **Partenaire 4 : SCHAPI**
  - (3 permanents)

QUALIFICATIONS DES PARTENAIRES :
DOMAINES D'EXPERTISE

- **Partenaire 1 (coordinateur) : ARMINES/Ecole des Mines d'Alès**
  - Expérience dans l'application des réseaux de neurones à l'hydrologie
- **Partenaire 2 : ESPCI ParisTech**
  - Grande expérience dans les Réseaux de Neurones et SVM
- **Partenaire 3 : EDYTEM**
  - Grande expérience en hydrologie (crues rapides, TOPMODEL)
- **Partenaire 4 : SCHAPI**
  - Utilisateur final, responsable de l'alerte vigicrues

TACHES

- Tâche 0 : coordination (ARMINES/EMA)
- Tâche 1 : production et validation de la base de données (SCHAPI)
- Tâche 2 : modélisation de la chaine hydrométéorologique (ARMINES/EMA)
- Tâche 3 : modélisation semi-physique (EDYTEM)
- Tâche 4 : machines à vecteurs supports (ESPCI)
- Tâche 5 : assimilation de données (EDYTEM)
- Tâche 6 : généralisations spatiale et temporelle (ARMINES/EMA)
- Tâche 7 : spécifications du prototype logiciel (SCHAPI)

AVANCÉES SCIENTIFIQUES (1/7)

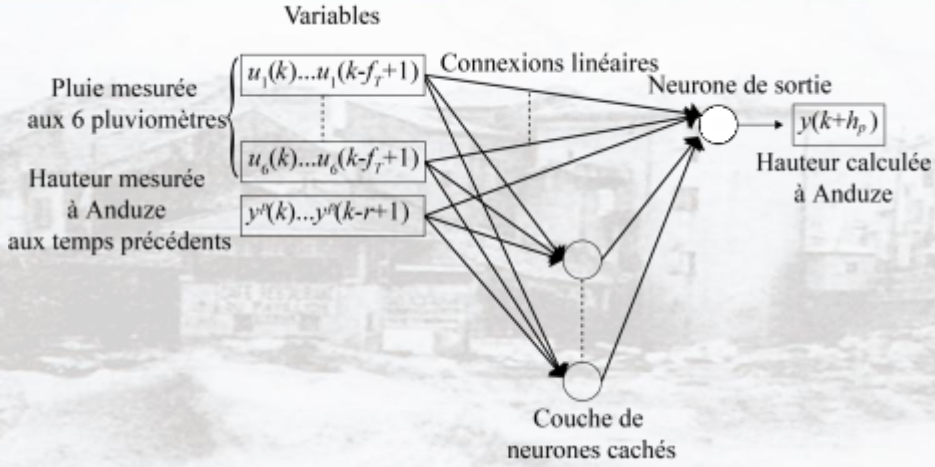
Relation pluie-débit du Gardon d'Anduze par réseaux de neurones

Thèse d'Audrey Bornancin Plantier, Co-direction entre EMA et ESPCI ParisTech

- Etude exhaustive du rôle des hyperparamètres sur la qualité de la solution en prévision
- Recherche d'un moyen de sélection du meilleur modèle pour 5 horizons de prévision (1/2h, 1h, 2h, 3h, 4h, 5h)
  - ➢ *Chaque événement est très spécifique, il ne semble pas y avoir de modèle de qualité simultanément pour tous les événements intenses*
- Justification de l'usage de l'assimilation de données : adaptation à l'événement en cours
- Du fait de l'utilisation en prévision, l'adaptativité n'apporte pas de bénéfice (la correction arrive trop tard)
- Meilleur gain apporté par la superposition de connexions directes à un modèle multicouches classique

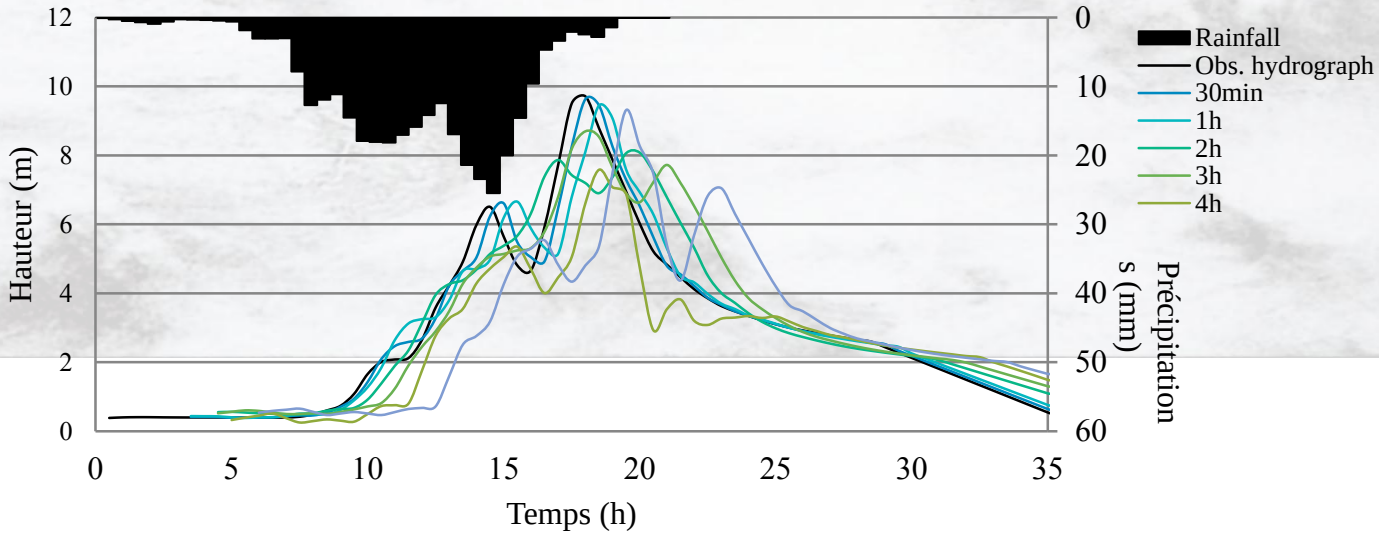
AVANCÉES SCIENTIFIQUES (2/7)

- **Thèse d'Audrey Bornancin Plantier (illustrations)**



Modèle incluant les contributions linéaires

Limnigrammes prévus pour l'événement de 2002 à Anduze



AVANCÉES SCIENTIFIQUES (3/7)

Changements d'échelles spatiales et temporelles

Thèse de Guillaume Artigue, Hydrologie, ED SIBAGHE, Montpellier 2

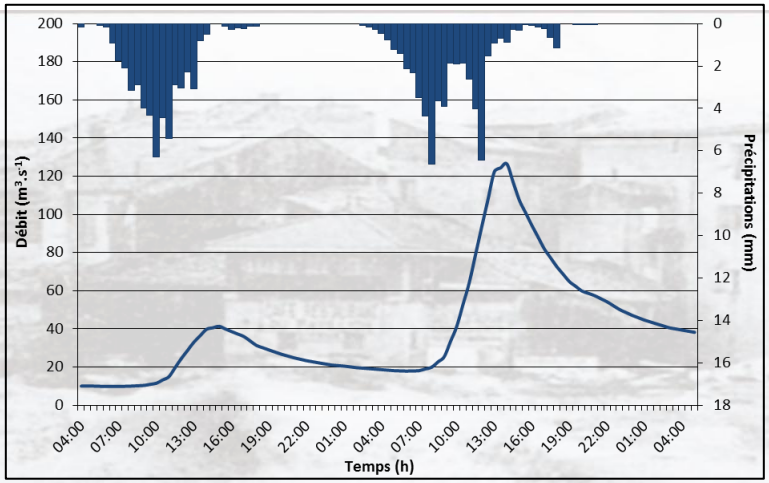
- Généralisations à la Cèze et à l'Ardèche : bassins versants non jaugés

- **Comment apprendre sans base de données ?**

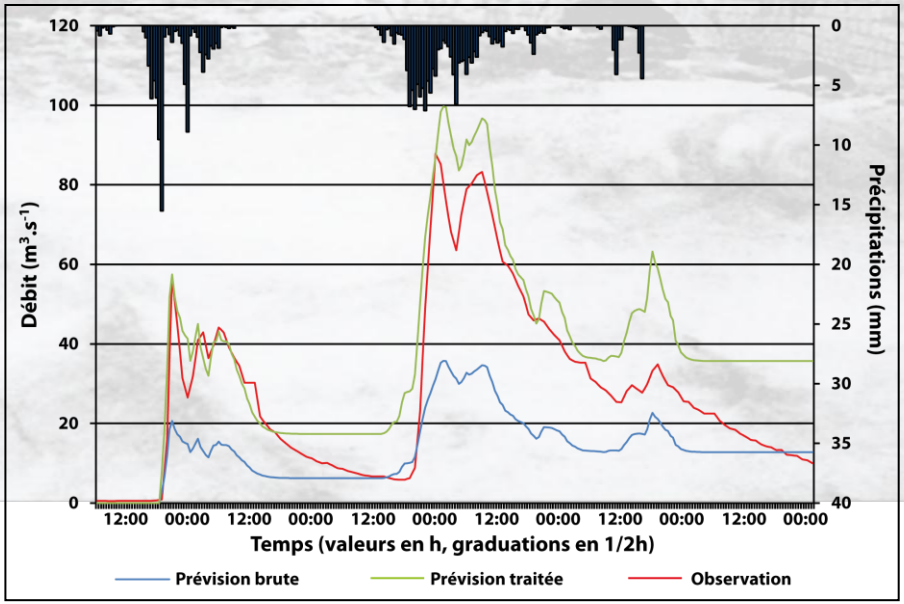
- Modèle dynamique établi pour un bassin pilote : Gardon de Mialet
- Généralisation à des bassins sans connaître les débits observés
    - **Modèle récurrent**
    - **Transformations pour s'adapter au bassin cible**
- A partir d'une superficie du bassin cible de 30km^2 les prévisions sont satisfaisantes jusqu'à un horizon de 2h (sans prévision de pluies)

AVANCÉES SCIENTIFIQUES (4/7)

- **Thèse de Guillaume Artigue (illustrations)**



Non-linéarité de la relation Pluie-Débit (Gardon de Mialet)



Prévision à 2h sur le Bassin de la Gagnière 4 fois plus petit que celui du Gardon de Mialet

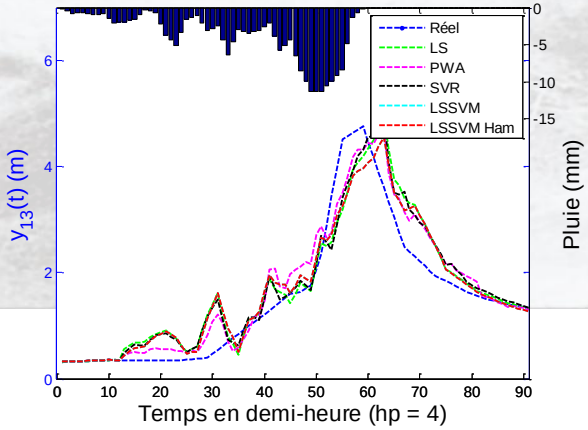
AVANCÉES SCIENTIFIQUES (5/7)

SVM

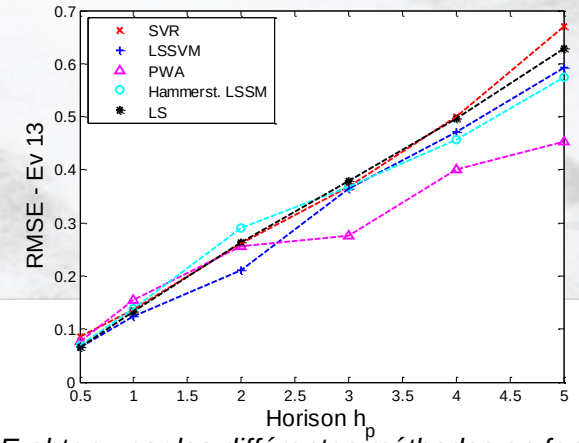
Post-doc en cours de Khaled Boukharouba, SIGMA, ESPCI Paris Tech

- Trois méthodes à noyaux (SVR, LSSVM, Hammerstein LSSVM) employées pour la modélisation de la relation : pluie – débit à Anduze
  - **Apprentissage dans un espace de plus grande dimension**
  - **Utilisation de fonction noyau : noyau gaussien**

- Performances très proches de celles des réseaux de neurones



Limnigrammes prévus par les différentes méthodes à noyaux et par le modèle linéaire (4h)



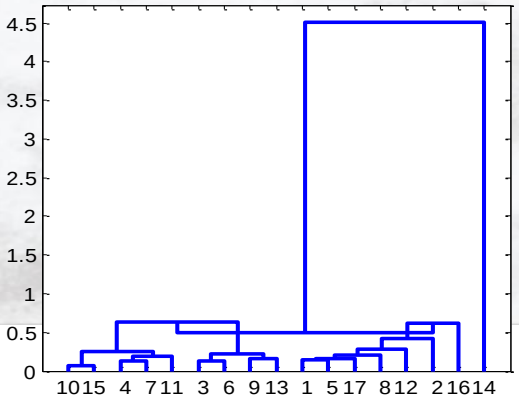
RMSE obtenu par les différentes méthodes en fonction de l'horizon de prédiction

AVANCÉES SCIENTIFIQUES (6/7)

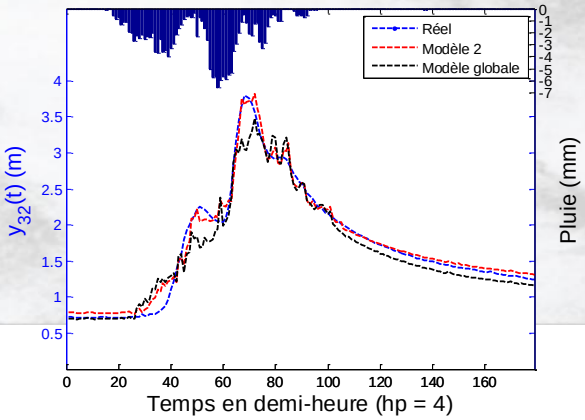
- **SVM**

Post-doc en cours de Khaled Boukharouba

- Construction de modèles spécifiques :
  - **Regroupement des événements en différents groupes ; chaque groupe comporte des événements similaires**
- Le meilleur modèle spécifique donne de meilleures performances que le modèle global (modèle issu de tous les événements)
- Sélection du meilleur modèle spécifique en cours d'utilisation (temps réel)?



Dendrogramme de la classification hiérarchique



Limnigrammes prévus par le meilleur modèle spécifique et par le modèle globale (4h)

AVANCÉES SCIENTIFIQUES (7/7)

TOPMODEL

- Assimilation de données par utilisation de techniques variationnelles : utilisation de modélisation adjointe comme correction de référence des prévisions du modèle hydrologique

- Test d'utilisation d'un RN comme apprentissage des biais systématiques du modèle hydrologique : détection et représentation des erreurs régulières pour une meilleure précision de la prévision.

- Avantage aux approches variationnelles mais piste d'amélioration du modèle neuronal en utilisant également l'information sur les variables d'état du modèle hydro à base physique : l'humidité du bassin
  - Le modèle neuronal comporte une entrée actualisée par un modèle hydrologique

PRODUCTION SCIENTIFIQUE

Thèses

- **Guillaume Artigue**, soutenue le 3 décembre 2012 :
  Prévision des Crues Éclair par Réseaux de Neurones : Généralisation aux Bassins non Jaugés. Thèse de l'Université Montpellier 2, spécialité Eaux Continentales et Sociétés. Ecole Doctorale : Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydrosciences et Environnement (SIBAGHE).
- **Audrey Bornancin-Plantier**, soutenue le 25 février 2013 :
  Conception de modèles de prévision des crues éclair par apprentissage artificiel. Thèse de l'Université Pierre et Marie Curie, spécialité Informatique. Ecole Doctorale Informatique, Télécommunications et Électronique (EDITE).

Revues

- G. Artigue, A. Johannet, V. Borrell, and S. Pistre (2012). *Flash flood forecasting in poorly gauged basins using neural networks: case study of the Gardon de Mialet basin (southern France)*. Nat. Hazards Earth Syst. Sci., 12, 3307-3324, 2012 www.nat-hazards-earth-syst-sci.net/12/3307/2012/doi:10.5194/nhess-12-3307-2012.
- Maréchal D., Ayral P-A., Bailly J-S., Puech C., Sauvagnarges-Lesage S., (2012), *Sur l'origine morphologique des écoulements par l'analyse d'observations hydrologiques distribuées. Application à deux bassins versants cévenols (Gard, France)*. Géomorphologie : relief, processus, environnement, 1, 47-62.
- Borgonovo, E.; Castaings, W. & Tarantola, S. *Model Emulation and Moment Independent Sensitivity Analysis: An Application to Environmental Modelling*, Environmental Modelling and Software 38, 2012, pp. 13-26.

PRODUCTION SCIENTIFIQUE

Conférences

- Artigue, G., Johannet, A., Borrell, V., Pistre, S., 2011 : *Flash floods forecasting without rainfalls forecasts by recurrent neural networks. Case study on the Mialet basin (Southern France)*, Proceedings of the third world congress on nature and biologically inspired computing, NABIC 2011, Salamanca, Spain, pp. 303-310, doi : 10.1109/NaBIC.2011.6089612.
- Audrey Bornancin Plantier, Anne Johannet, Pierre Roussel Ragot, Gérard Dreyfus. *Flash Flood Forecasting using Neural Networks without Rainfall Forecasts : Model Selection and Generalization capability.* Communication Orale à l'assemblée Générale EGU (avril 2011)
- Castaings, W., E. Borgonovo et S. Tarantola (2010). *Sampling plans for the estimation of moment independent importance measures*. In Procedia – Social and Behavioral Sciences. T. 2. 6. Sixth International Conference on Sensitivity Analysis of Model Output, (Milan, IT), Vol 2, Issue 6, 2010, p 7619-7620 (chairman session 21-3A), p. 7629 –7630. doi : DOI:10.1016/j.sbspro.2010.05.151.
- Borgonovo, E., Castaings, W. et S. Tarantola (2010). *Theorems and Analytical Test Cases in Moment Independent Sensitivity Analysis*. In Procedia – Social and Behavioral Sciences. T. 2. 6. Sixth International Conference on Sensitivity Analysis of Model Output (Milan, IT), Vol 2, Issue 6, 2010, p 7619-7620, p. 7619 –7620. doi : DOI:10.1016/j.sbspro.2010.05.146.
- Le Dimet, F-X et Castaings, W. (2010). *Variational methods for sensitivity analysis and parameter estimation in hydrology*. In : Proceedings of the Asia Oceania Geosciences Society (AOGS) General Assembly (Hyderabad, India). Exposé invité.
- Castaings, W. (2010). *Sensitivity analysis as an ingredient of modelling environnemental systems : potential and challenges*. In Workshop on Modelling Soil & Water Conservation at different scales (Baeza, Spain). Séminaire invité.

5 Posters

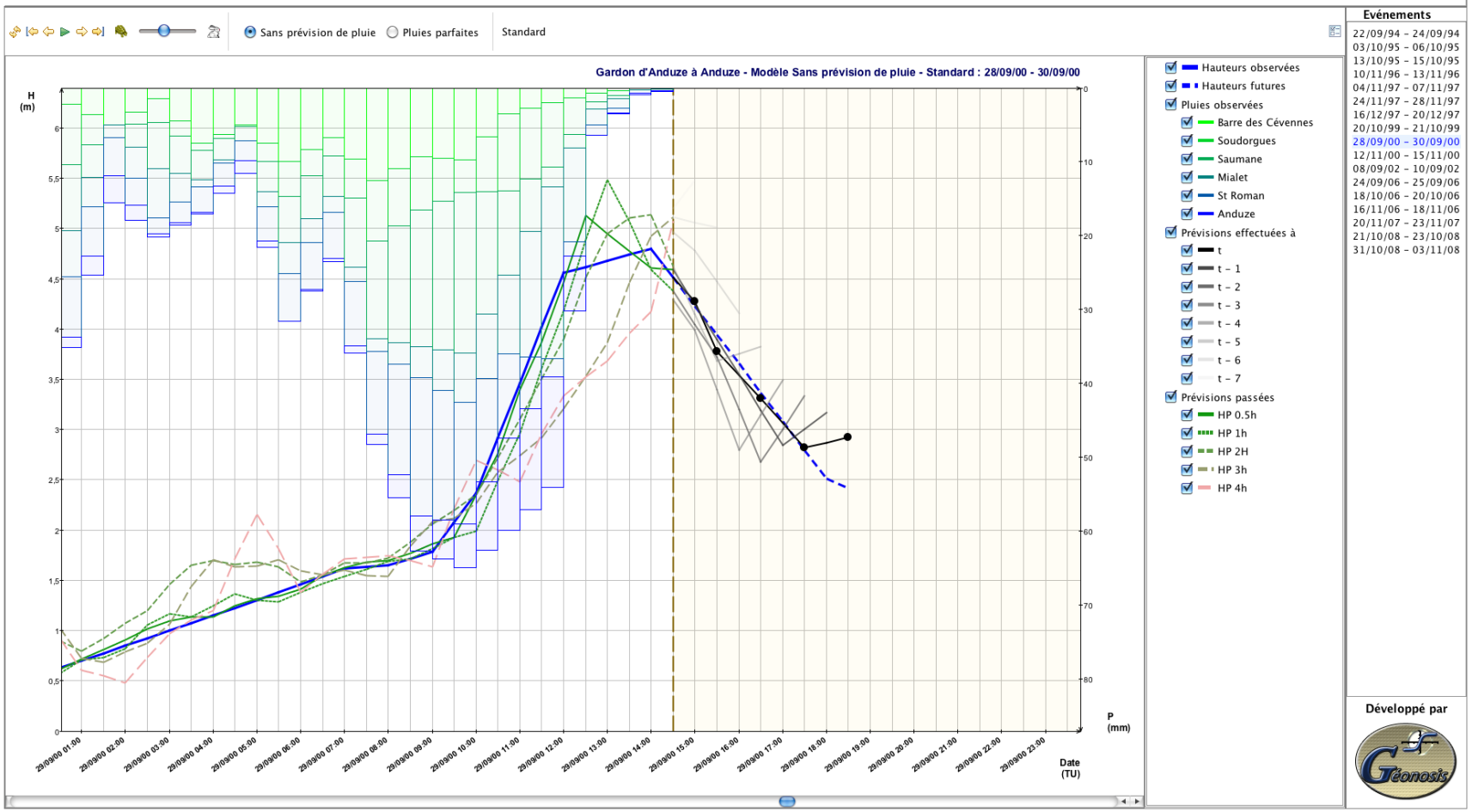
PRODUIT

- **Pour l'utilisateur final : le SCHAPI**
  - A partir des travaux d'Audrey Bornancin–Plantier (**partenaires 1 et 2**), **livraison d'un prototype de prévision** de la hauteur d'eau à Anduze jusqu'à un horizon de 5h.
  - Prototype en phase d'évaluation sur la chaine temps réel
  - Dépasse l'objectif initial de la seule définition d'un cahier des charges
- **Démonstrateur logiciel de prévision de crues DEMEAU accessible depuis le site internet du projet (**http://blog.espci.fr/flash/**)**

DÉMONSTRATEUR

- Développé par la Société Géonosis

Démonstrateur de prévision de crues par réseaux de neurones



DÉVELOPPEMENT ÉCONOMIQUE

- Aide à la création de la Société Géonosis (http://www.geonosis.fr), crée en juin 2011 et spécialisée dans la modélisation environnementale (développements scientifiques et gestion des informations environnementales, spatialisées).

Merci de votre attention