

Introduction au Connexionisme, ou *Machine Learning*

Anne Johannet, HSM, IMT Mines Alès

Contenus

- Inspiration biologique
- Mémoire distribuée : neurone « grand-mère »
- Éléments historiques
- Les études
- Quelques explications : la classification
- Quelques explications : l'identification de fonction
- Exemples de travaux en hydrologie

I. INSPIRATION BIOLOGIQUE : MODÉLISATION DU CERVEAU

- Le cerveau n'a pas été toujours identifié comme étant le siège central de notre intelligence
- Dès Platon, la nécessité d'avoir un cerveau non dégradé s'est imposée
- Au XXème siècle le cerveau a été clairement identifié comme étant le lieu de l'activité mentale : autant émotionnelle que cognitive.
- En parallèle à cette evolution :
 - *machine à calculer de Pascal*
 - *ordinateur de Von Neumann*
- Les deux champs d'études se sont fécondés l'un l'autre
 - *Neuromimétique*
 - *Approche machine*

Mémoire distribuée : le neurone « grand-mère »

Ainsi, un souvenir, ou un processus mental ne serait représenté que par l'état de l'ensemble des neurones [Changeux]. Et la mémoire serait implémentée de manière redondante dans les connexions entre ces neurones.

- A partir d'un certain âge, nous perdons environ plusieurs dizaines de milliers de neurones chaque jour :
 - *pas de différence brutale dans nos capacités mentales*
 - *neurone « grand-mère »*
- *⇒ le stockage de l'information dans le cerveau n'est pas localisé, mais distribué*

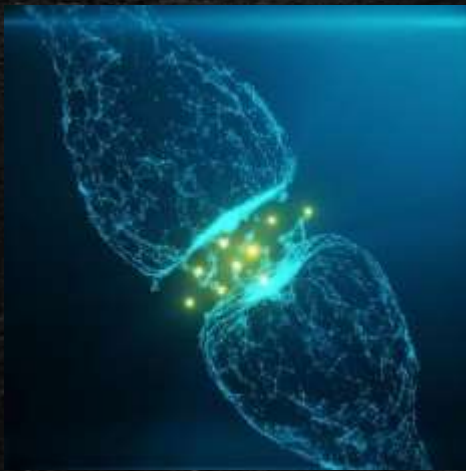
l'information dans le cerveau n'est pas codée comme dans une mémoire électronique (par adresse), mais est codée de manière distribuée et redondante : dans les connexions du schéma qui relie les neurones entre eux

[Changeux]. Jean-Pierre Changeux : « L'homme neuronal » : HACHETTE 1983

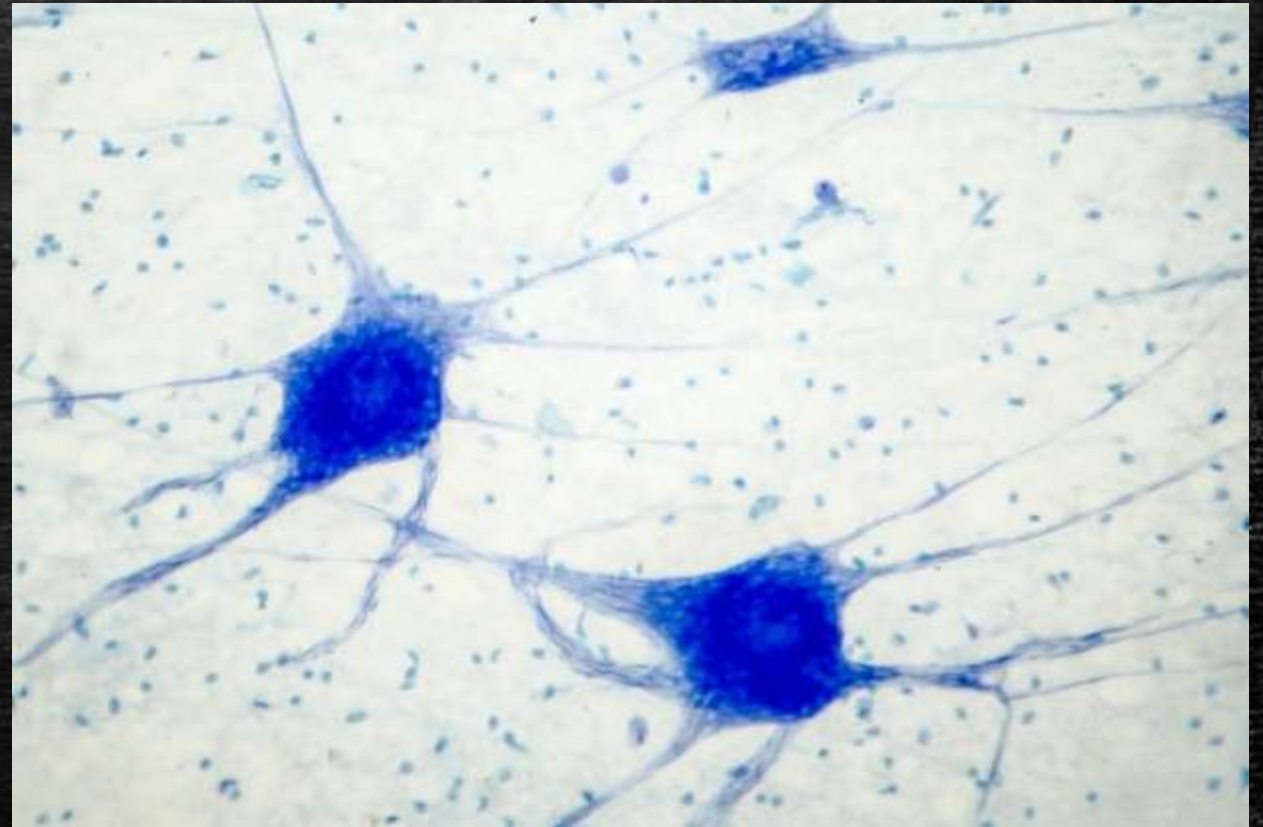
Inspiration biologique pour réaliser une machine

5

- Vue d'artiste de neurones biologique :
 - De très nombreux neurones
 - Un corps cellulaire,
 - Des connexions sortantes
 - Des connexions entrantes



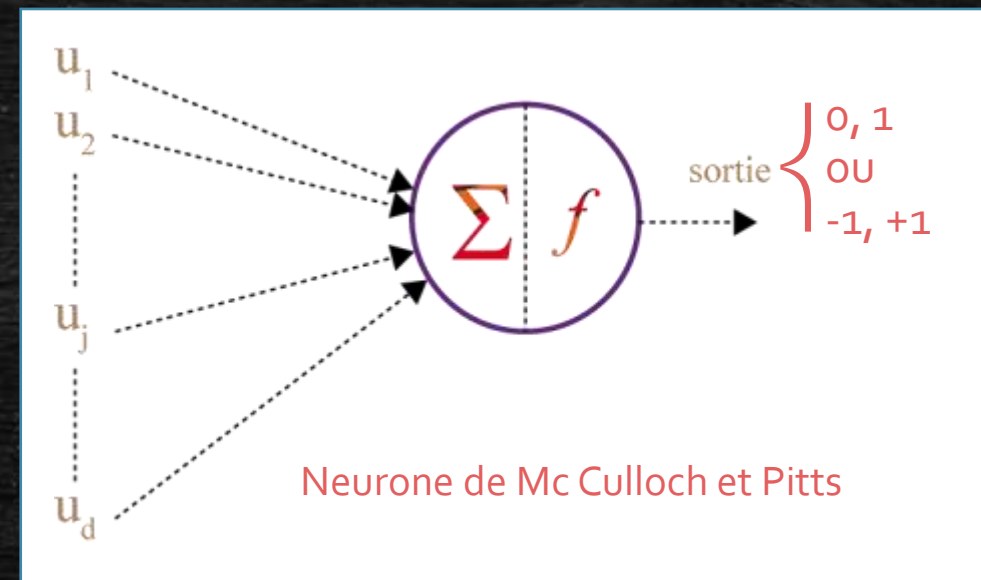
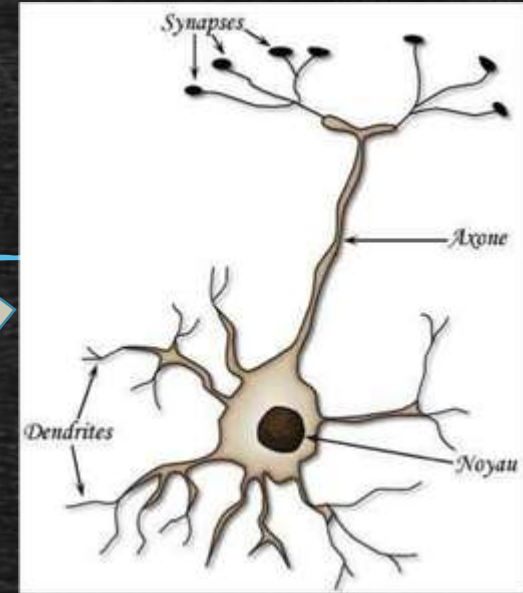
← Lieu des connexions : la synapse



Éléments historiques

Le premier modèle de neurone formel :

- proposé en 1943 par le mathématicien W. Pitts et le neurophysiologue Mc Culloch [[Mc Culloch](#)]
- Neurone à deux états : actif ou inactif (valeurs formelles : 0 ou 1)
- C'est ce modèle qui va initier les recherches en réseaux de neurones formels



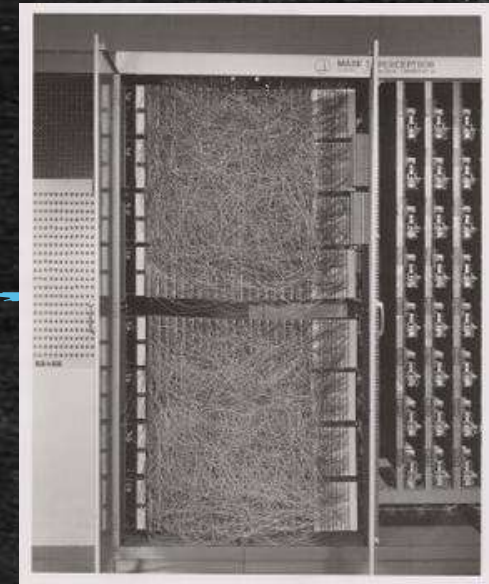
Éléments historiques et étapes clés

Difficultés d'études

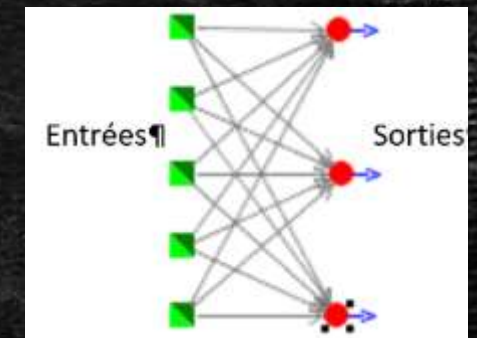
- La grande dimension du système
 - Comment représenter ces connexions ?
 - Comment les calculer ?
 - Quelle est la précision nécessaire ?

Années 50-60

- Rosenblatt (1958)
 - Modèle et machine de neurones binaires à seuil pour la classification et la reconnaissance de formes
- Minsky et Papers
 - Seules les séparations linéaires peuvent être réalisées



<https://en.wikipedia.org/wiki/Perceptron>



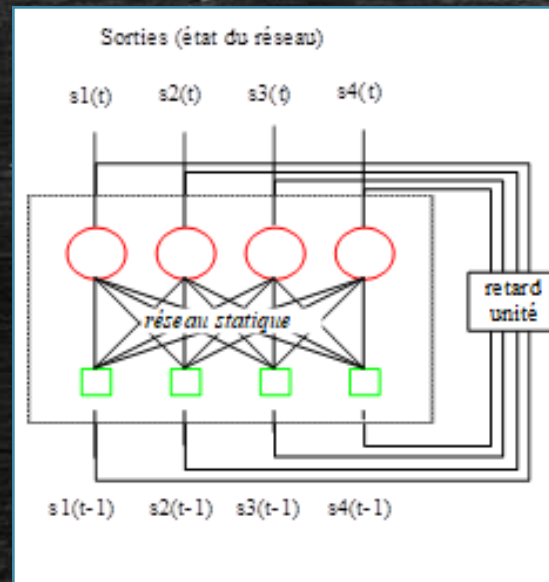
Un perceptron à 3 neurones et 5 entrées

Rosenblatt, Frank (1958), The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, Psychological Review, v65, No. 6, pp. 386–408. [doi:10.1037/h0042519](https://doi.org/10.1037/h0042519).
Minsky M. L. and Papert S. A. 1969. *Perceptrons*. Cambridge, MA: MIT Press.

Éléments historiques et étapes clés

Années 1980

- J. J. Hopfield (1982)
 - Physicien
 - Lien entre les “verres de spin” et un modèle de réseau de neurones
 - Utilise une règle d’apprentissage des coefficients synaptiques : la règle de Hebb
 - Définition d’une énergie
 - Les images à apprendre doivent minimiser cette énergie



N neurones, N^2 coefficients synaptiques.

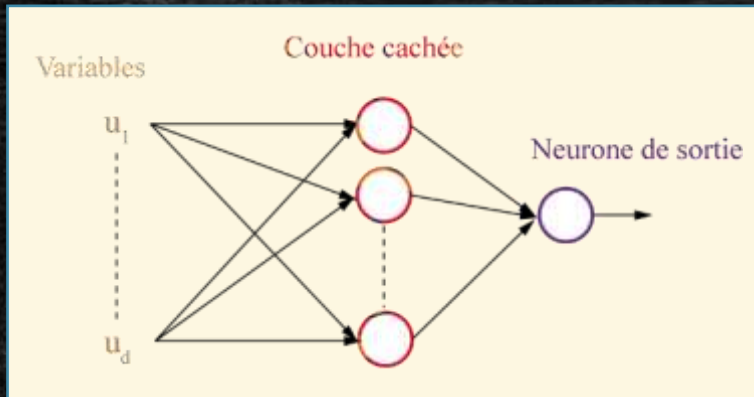
- T. Kohonen (1982)
 - Processus non supervisé
 - Cartes auto-organisatrices
 - SOM « Self Organizing memory »
 - ACP non linéaires
 - Application à la correction d’erreur sur des images

J. J. Hopfield, « Neural networks and physical systems with emergent collective computational abilities », *Proceedings of the National Academy of Sciences*, vol. 79, n° 8, 1^{er} avril 1982, p. 2554–2558

T. Kohonen, *Self-Organized Formation of Topologically Correct Feature Maps*, *Biological Cybernetics*, vol. 46, pp. 59–69, 1982.

Éléments historiques et étapes clés

La réalisation de réseaux de neurones formels complexes implique souvent l'utilisation de systèmes hiérarchisés ou structurés en couches.



Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (1986-10-09). "Learning representations by back-propagating errors". Nature. 323 (6088): 533–536. Bibcode:1986Natur.323..533R. doi:10.1038/323533a0. ISSN 1476-4687. S2CID 205001834.

K. Hornik, M. Stinchcombe, H. White. Multilayer Feedforward Networks are Universal Approximators. In Neural Networks, vol. 2, pp. 359–366 (1989).

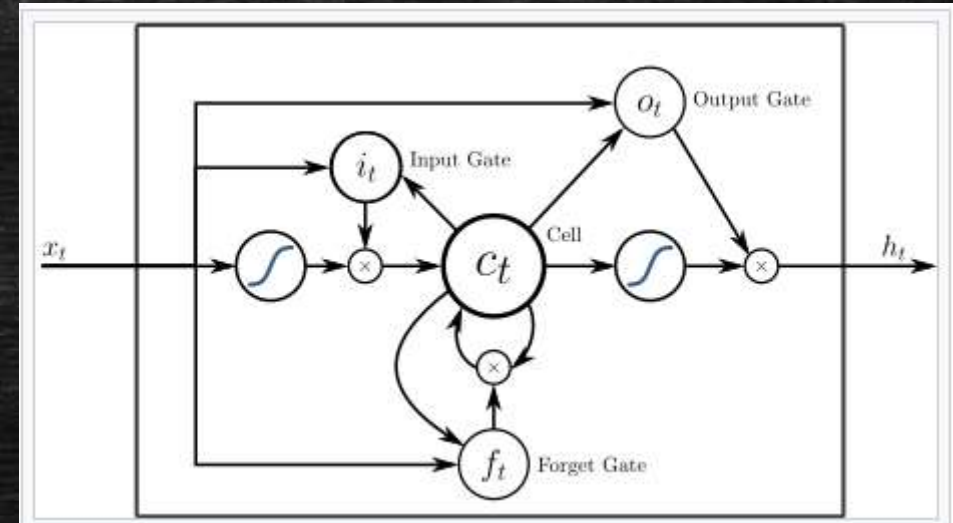
Werbos, P. (1990). "Backpropagation Through Time: What It Does and How to Do It". Proceedings of the IEEE. 78 (10): 1550–1560. doi:10.1109/5.58337

Werbos, Paul J. (1994). The Roots of Backpropagation : From Ordered Derivatives to Neural Networks and Political Forecasting. New York: John Wiley & Sons. ISBN 0-471-59897-6.

- Utilisation de réseaux multicouches (1986)
 - Algorithme de la rétropropagation (Rumelhart)
- 1989 : Démonstration de la propriété d'approximation universelle (Hornik)
- 1990 : rétropropagation sur les systèmes dynamiques (P. Werbos)
- 1994 : utilisation pour la prévision (P. Werbos)

Éléments historiques et étapes clés

- 1995-1997 : S. Hochreiter J. Schmidhuber proposent le LSTM (Long Short Term Memory)
 - Solution au problème du *gradient évanescant*
 - Peut prendre en compte plusieurs échelle de temps
 - Poids partagés
 - Conserve une information de contexte grâce à une rétroaction
 - Modèle récurrent
 - Particulièrement performant pour la traduction ou la correction des phrases

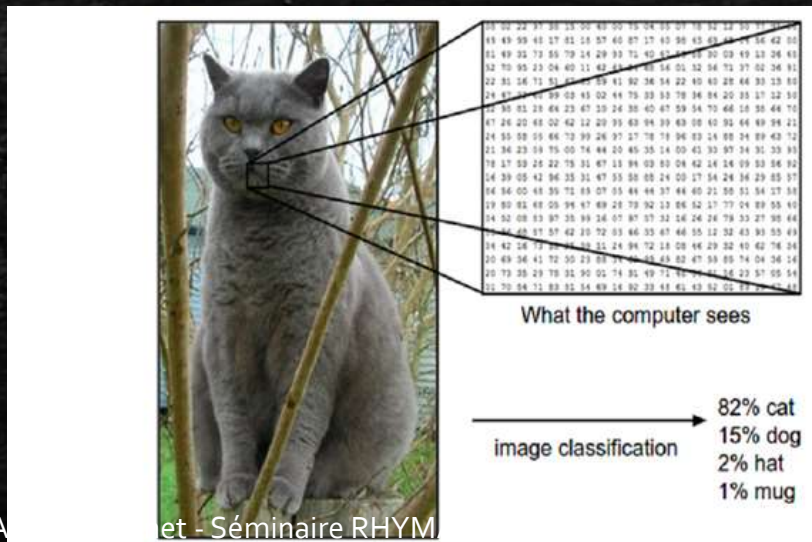


https://en.wikipedia.org/wiki/Long_short-term_memory

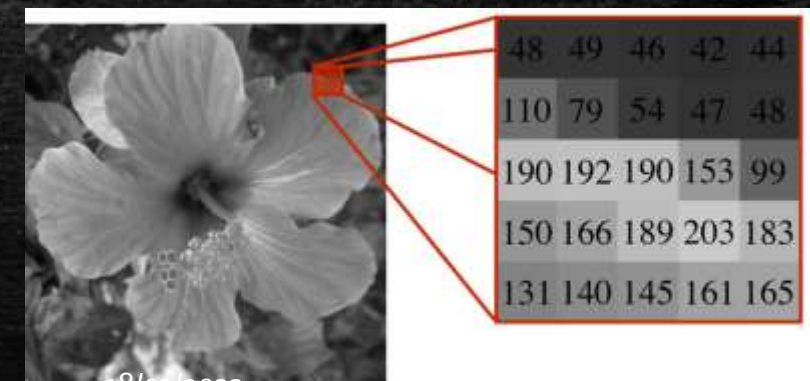
Ce n'est pas toujours le meilleur modèle : Jeannin, P.Y. et al, Karst modelling challenge 1: Results of hydrological modelling, Journal of Hydrology, Volume 600, 2021, 126508, <https://doi.org/10.1016/j.jhydrol.2021.126508>.

Éléments historiques et étapes clés

- 2012. Y. Le Cun utilise un réseau convolutionnel la reconnaissance d'image :
 - Guidage de robots mobiles
 - Gestion de l'invariance par translation
- 2015. Y. Le Cun Réseaux profonds convolutionnels pour prétraiter les signaux
 - Plusieurs couches enchainant les couches convolutionnelles et les filtrages/compression de l'information

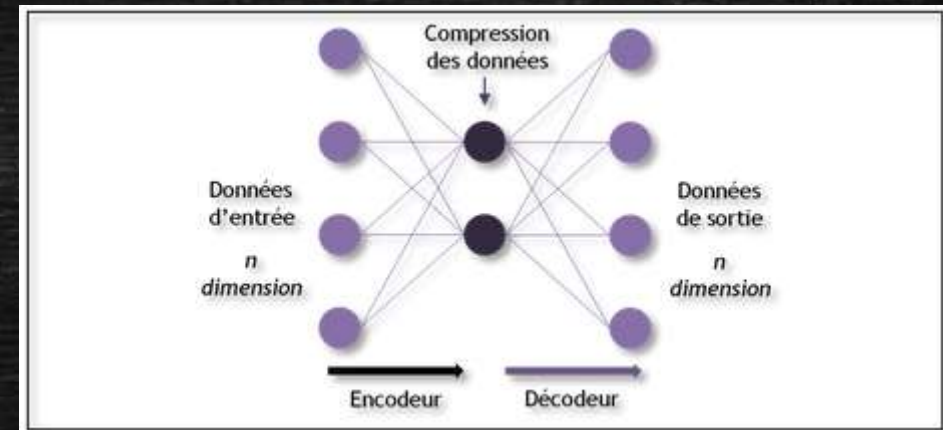


- Autoencodeurs – compression de l'information



Éléments historiques et étapes clés

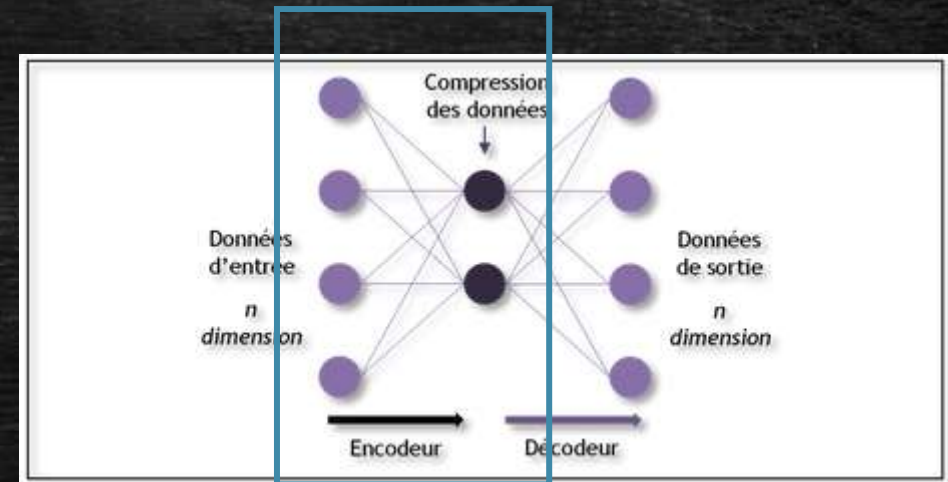
- Autoencodeurs (Hinton) :
 - Comment compresser l'information ?
- Comment apprendre les paramètres des couches profondes ?
 - On ajoute un autoencodeur dans les couches profondes
 - Compression de l'information
 - Incorporer la partie « compression » dans les couches profondes avec les paramètres déjà appris



<https://bigdatablog.skapane.com/les-auto-encodeurs/>

Éléments historiques et étapes clés

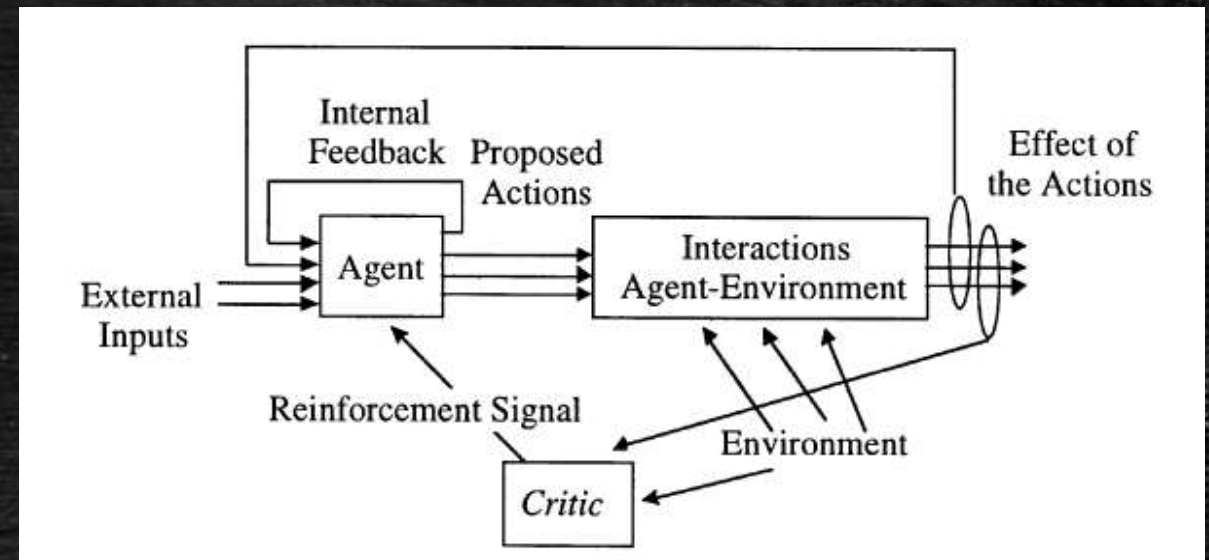
- Autoencodeurs : comment apprendre les paramètres des couches profondes ?
 - On ajoute un autoencodeur dans les couches profondes
 - Compression de l'information
 - Incorporer la partie « compression » dans les couches profondes avec les paramètres déjà appris



<https://bigdatablog.skapane.com/les-auto-encodeurs/>

Éléments historiques et étapes clés

- Années 80-90. Apprendre par renforcement
 - Fonction objectif = ce que l'on veut
 - L'agent produit des actions
 - L'environnement donne son évaluation
 - L'apprentissage se fait par "pénalité-récompense"
- Développements plus récents dans de nombreux domaines – Q-learning
- Liens avec la biologie
- Très utilisé pour les jeux



From: A. Johannek, I. Sarda. Goal-Directed Behaviours by Reinforcement Learning. In Neurocomputing, vol. 28, pp. 107-125 (1999-a)

Tom M. Mitchell, *Machine Learning*, McGraw-Hill International Editions, 1997, chap. 13 *Reinforcement Learning*, p. 367-390

R. S. Sutton & A. G. Barto. Reinforcement Learning, second edition: An Introduction (Adaptive Computation and Machine Learning series) second edition. MIT Press.

Quelques explications : la classification

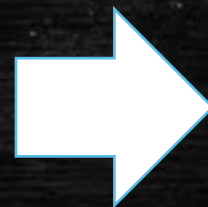
Exemple avec 2 entrées : aide à la conduite

- ▶ Entrées : vitesse, position
- ▶ On veut prendre une décision :
 - Je m'arrête au feu
 - Je continue à la même vitesse



On veut associer le code d'une classe à une (forme en) entrée

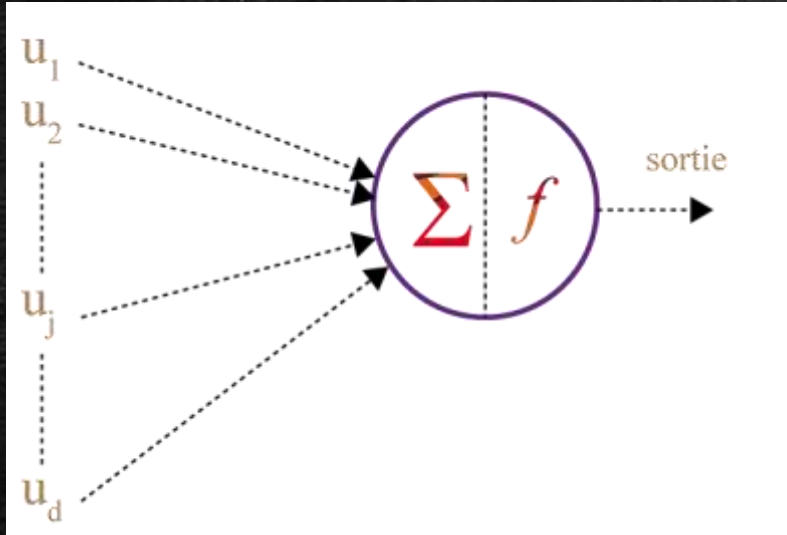
- ▶ On perd de l'information
- ▶ Le système prend une décision



Sortie : $\left\{ \begin{array}{l} +1 : \text{la voiture s'arrête} \\ -1 : \text{la voiture continue à} \\ \text{la même vitesse} \end{array} \right\}$

La classification

- La séparation linéaire



La fonction neurone

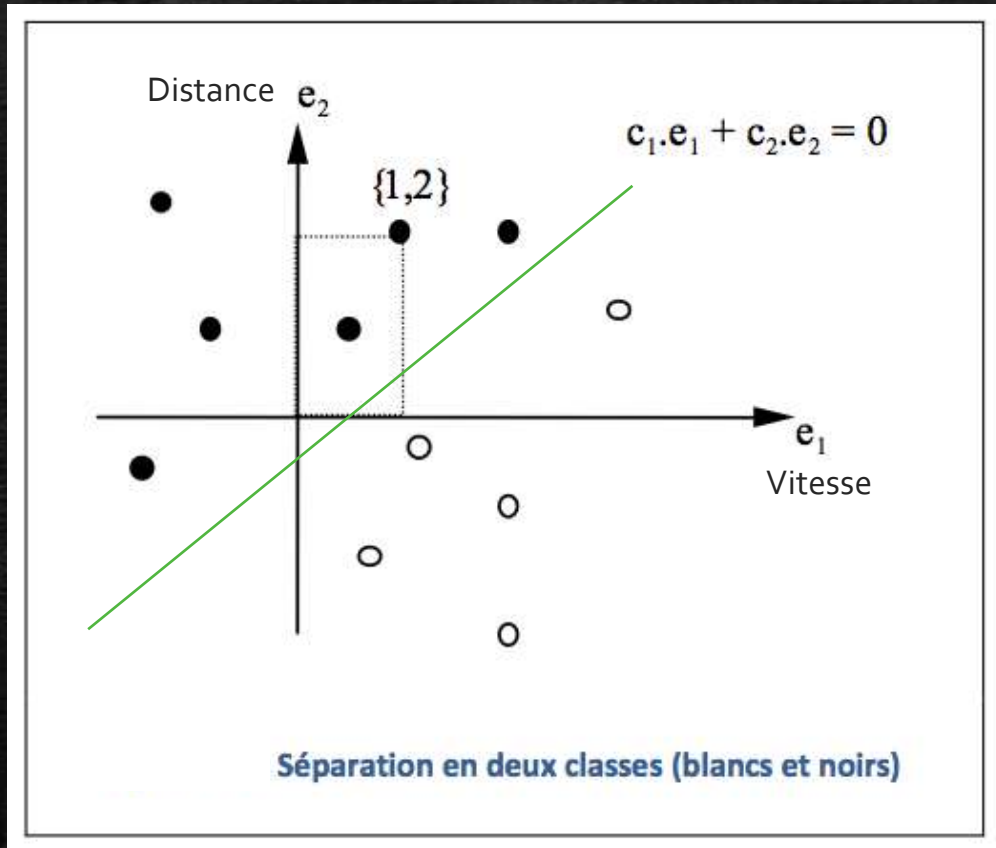
▪ **Modèle binaire à seuil :**

- Les synapses sont représentées par un coefficient
 - Elles connectent l'entrée au corps cellulaire
- Le potentiel synaptique est la somme pondérée des entrées :
 - Très simplifié !
- La sortie (potentiel d'action) est déclenchée par l'intensité du potentiel :

$$s=f(v) \text{ où } f(.) \text{ est un seuil}$$

La classification

- La séparation linéaire



Le calcul du neurone fait dans le plan des entrées

Que peut faire un neurone à 2 entrées tout seul ?

➤ Il calcule :

$$y = c_1 e_1 + e_2 c_2$$

➤ Il applique la fonction seuil

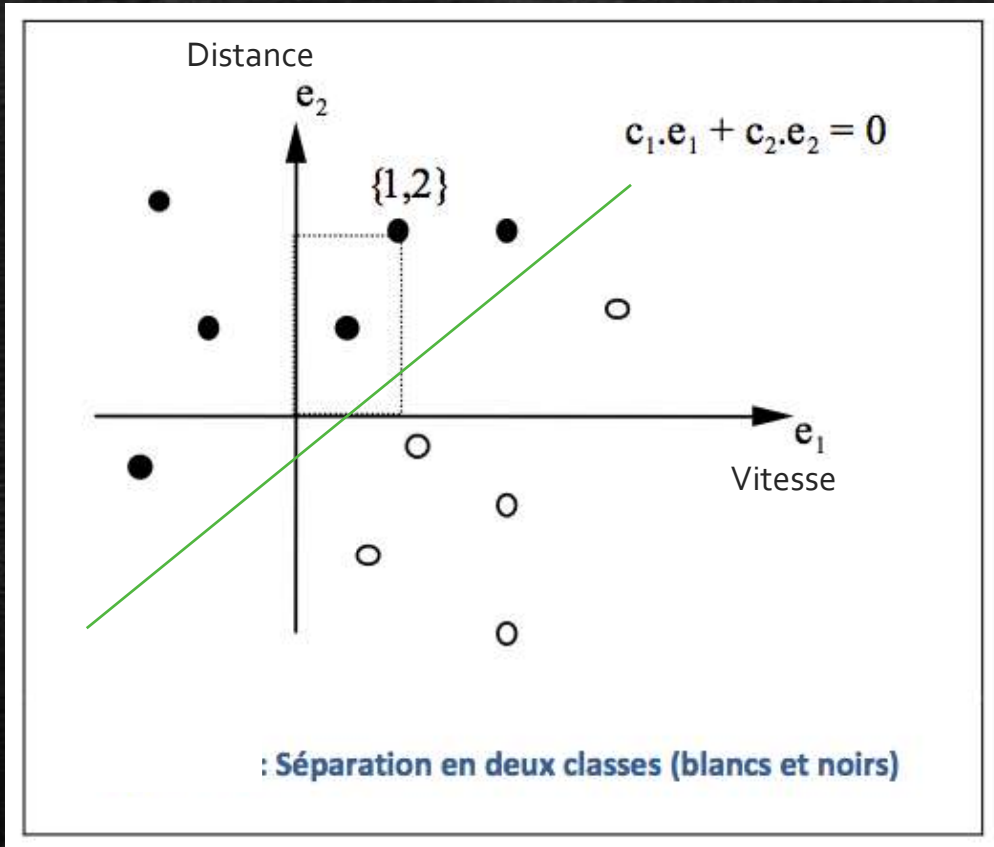
- L'équation du potentiel est celle d'une droite dans le plan :

$$e_2 = \frac{-c_1}{c_2} e_1$$

- Le neurone sépare le plan en 2 zones

La classification

- La séparation linéaire



Le calcul du neurone fait dans le plan des entrées

▪ Exemple d'application :

➤ Arrêt au feux orange ?

- e_1 est la vitesse du véhicule,
- e_2 est la position du véhicule

- Les exemples blancs correspondent aux cas où la voiture s'arrête
- Les exemples noirs au cas où la voiture continue de rouler

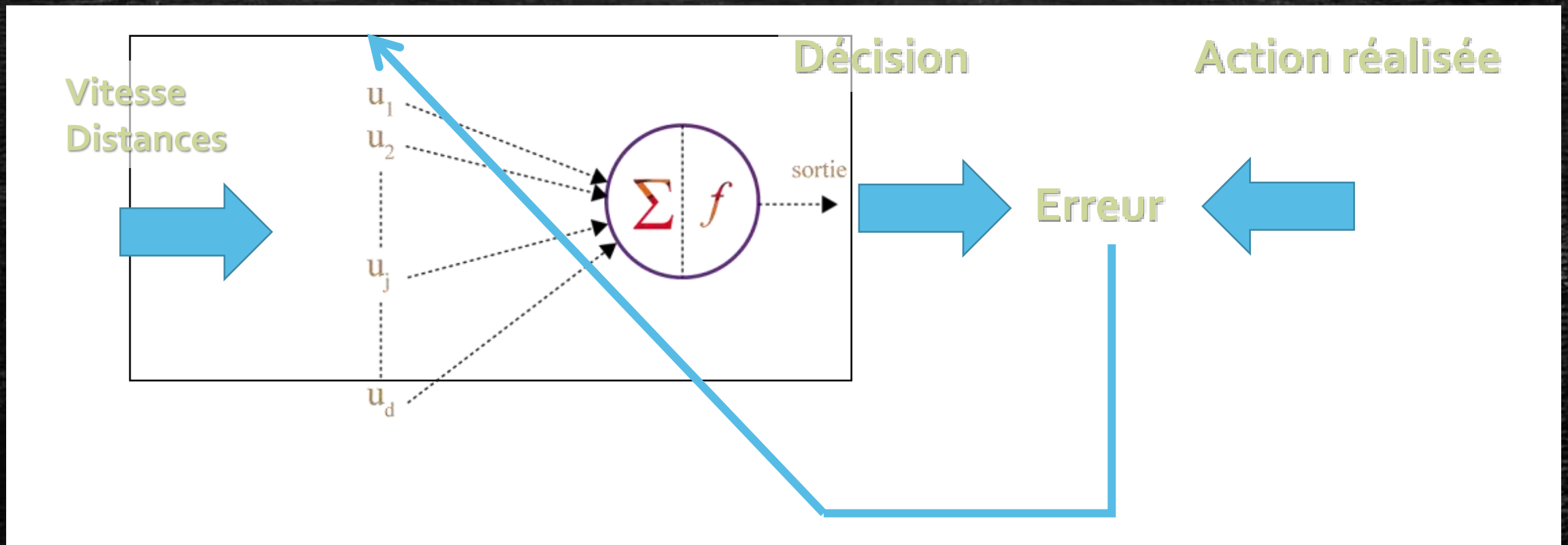
➤ Le neurone applique la fonction seuil :

- Il sépare le plan en 2 zones : il prend une **décision**

L'apprentissage

- Comment sont calculés les coefficients c_1 et c_2 ?

➤ Par apprentissage grâce à une base de données :



La classification

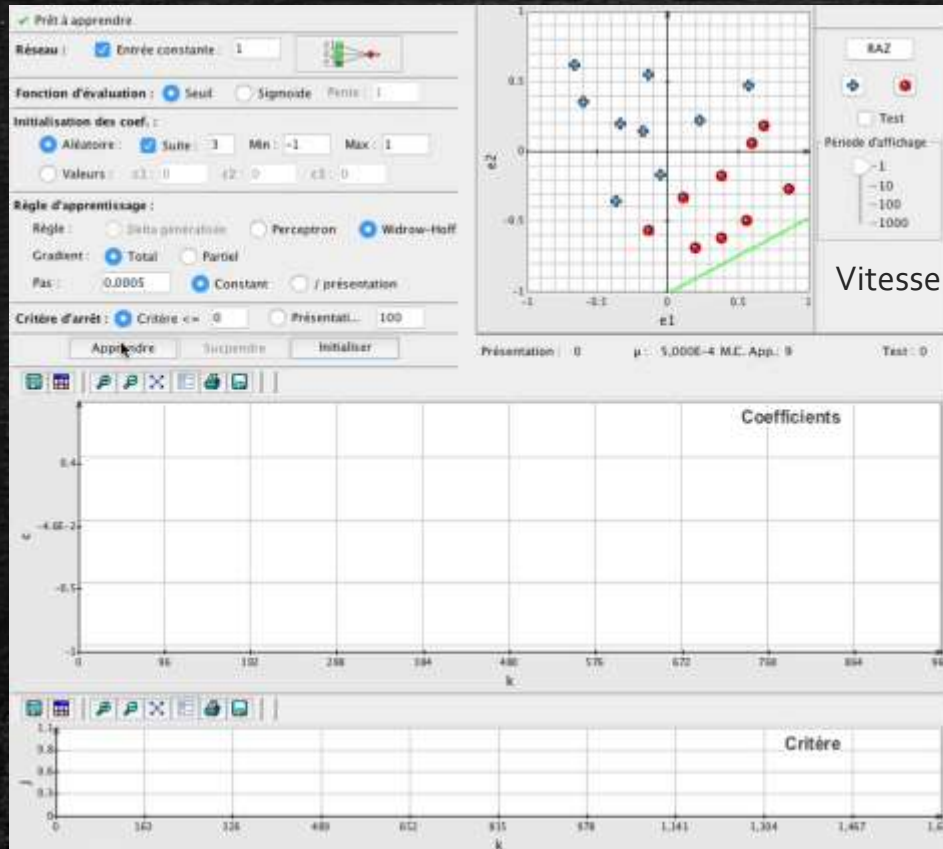
- La séparation linéaire

Exemple d'apprentissage :

➤ Arrêt au feu orange :

- e_1 est la vitesse du véhicule,
- e_2 est la position du véhicule

- Les exemples rouges correspondent aux cas où la voiture s'arrête
- Les exemples bleus au cas où la voiture continue de rouler
- Le neurone sépare le plan en 2 zones : il effectue une **décision**

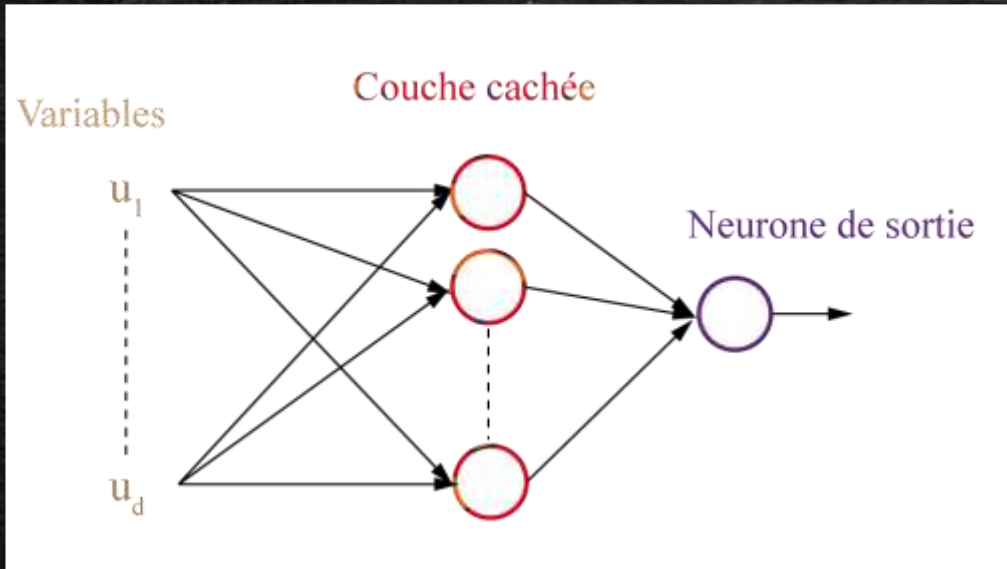


Le calcul du neurone fait dans le plan des entrées

La classification

- La séparation non linéaire

- On connecte entre eux plusieurs neurones : réseau

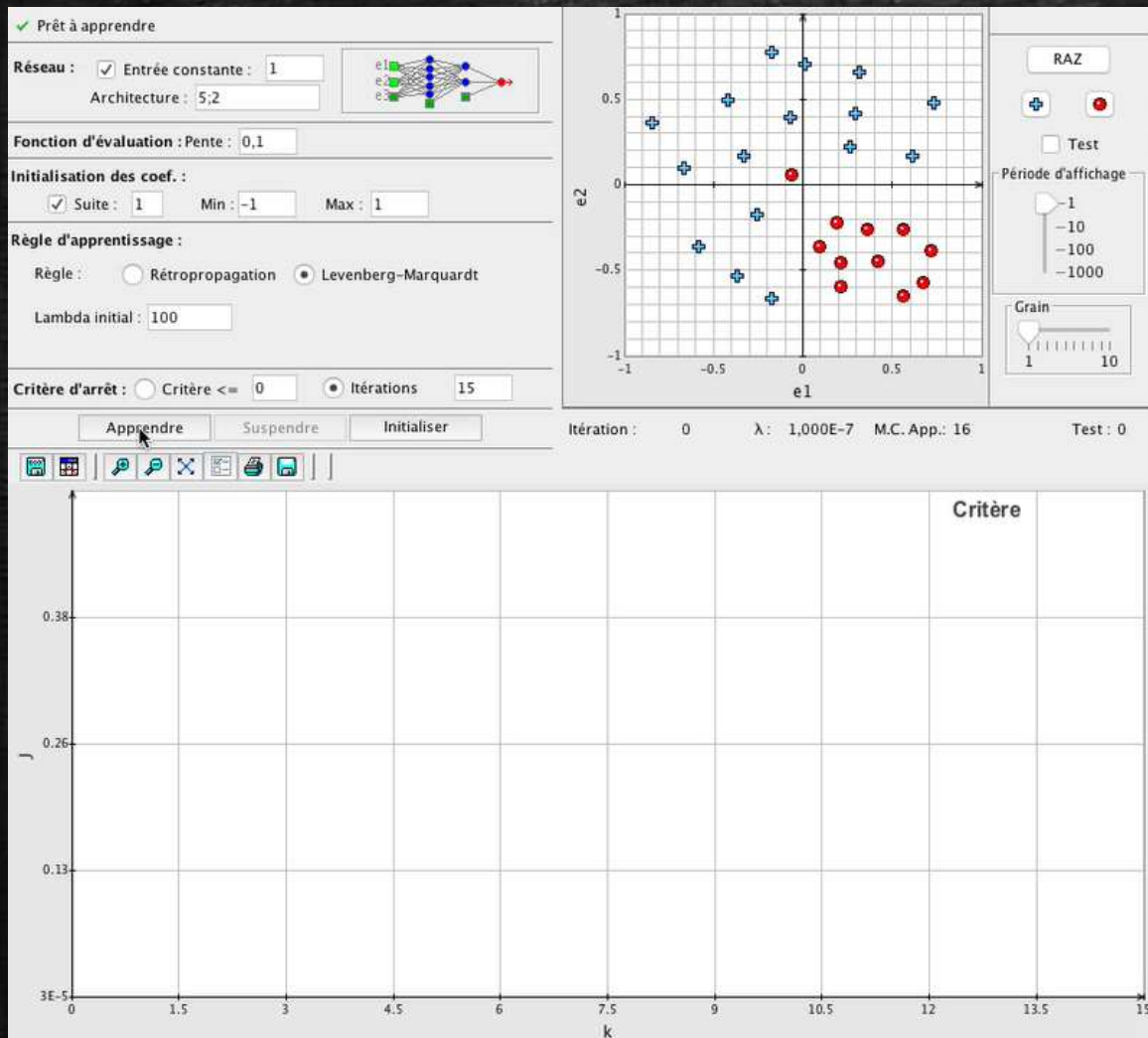


Perceptron à une couche cachée non linéaire

- Exemple du réseau appelé **Perceptron multicouches**
 - Connexion d'une couche à l'autre
 - Pas de connexion au sein d'une même couche
- Le neurone de sortie prend une décision plus complexe car le calcul de son potentiel est non linéaire

La classification

- La séparation non linéaire



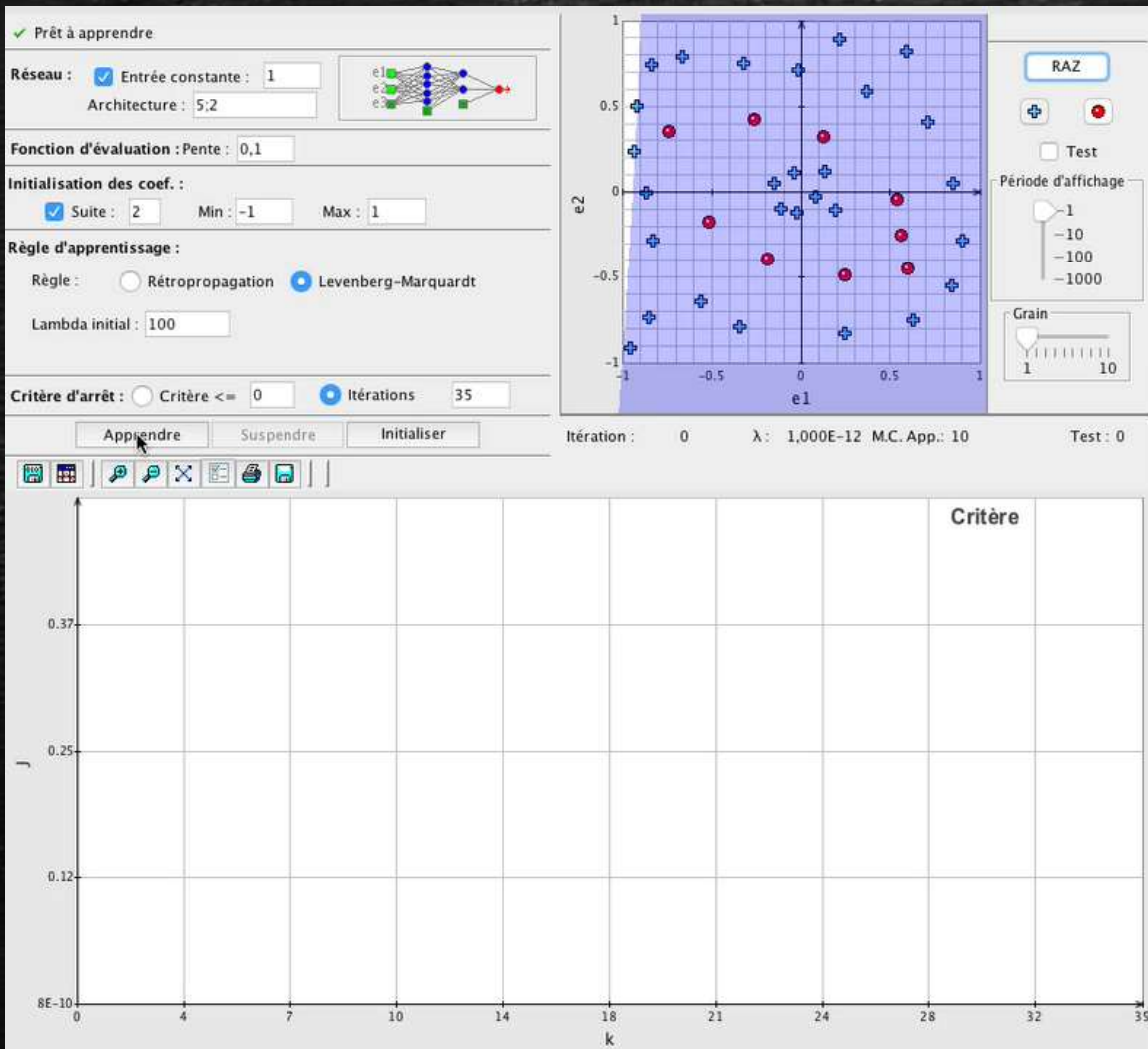
- Il faut déterminer :**

- Le nombre de couches
- Le nombre de neurones dans chacune des couches
- Les meilleurs variables d'entrée possibles
- Les paramètres propices à l'apprentissage

Le neurone de sortie prend une décision plus complexe que précédemment car le calcul de son potentiel est non linéaire

La classification

- La séparation non linéaire



- Il faut déterminer :**

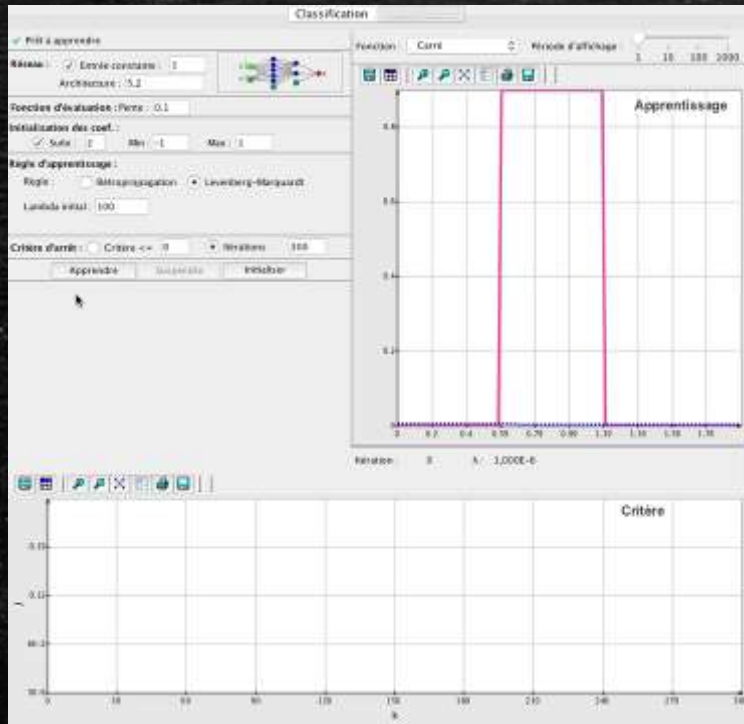
- Le nombre de couches
- Le nombre de neurones dans chacune des couches
- Les meilleures variables d'entrée possibles
- Les paramètres propices à l'apprentissage

Le neurone de sortie prend une décision plus complexe que précédemment car le calcul de son potentiel est non linéaire

Quelques explications : l'identification

▪ Il faut déterminer :

- Le nombre de couches
- Le nombre de neurones dans chacune des couches
- Les meilleures variables d'entrée possibles
- Les paramètres propices à l'apprentissage



Le neurone de sortie prend une décision plus complexe que précédemment car le calcul de son potentiel est non linéaire

Prévision des crues

25

- Thèse de Audrey-Bornancin-Plantier (co-Direction avec G. Dreyfus, ESPCI)
- Prévision de l'aléa
 - Définition de modèles de prévisions, temps réel, à très courte échéance (5h)



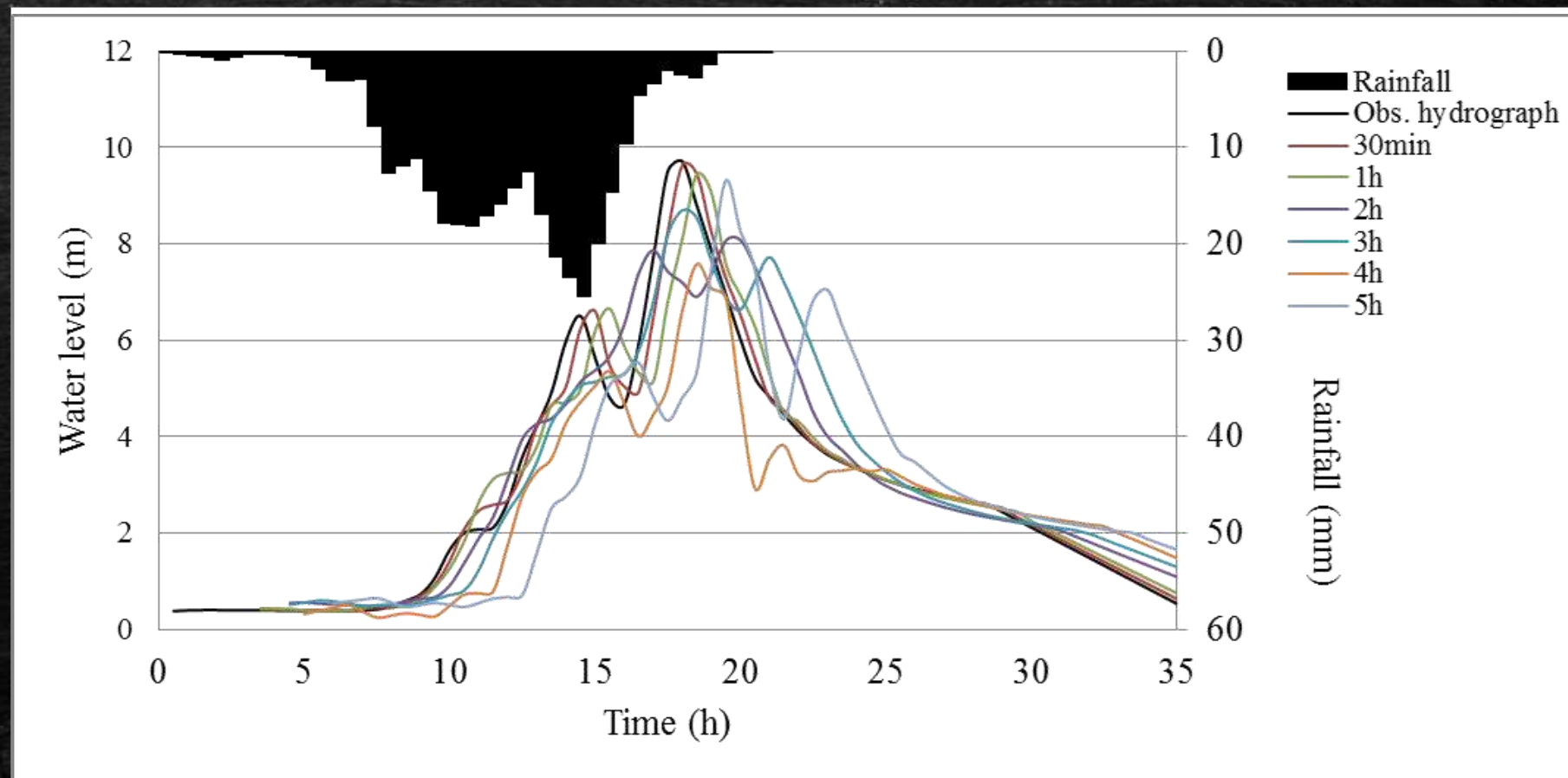
Prévision des crues

- Thèse de Audrey-Bornancin-Plantier (co-Direction avec G. Dreyfus, ESPCI)

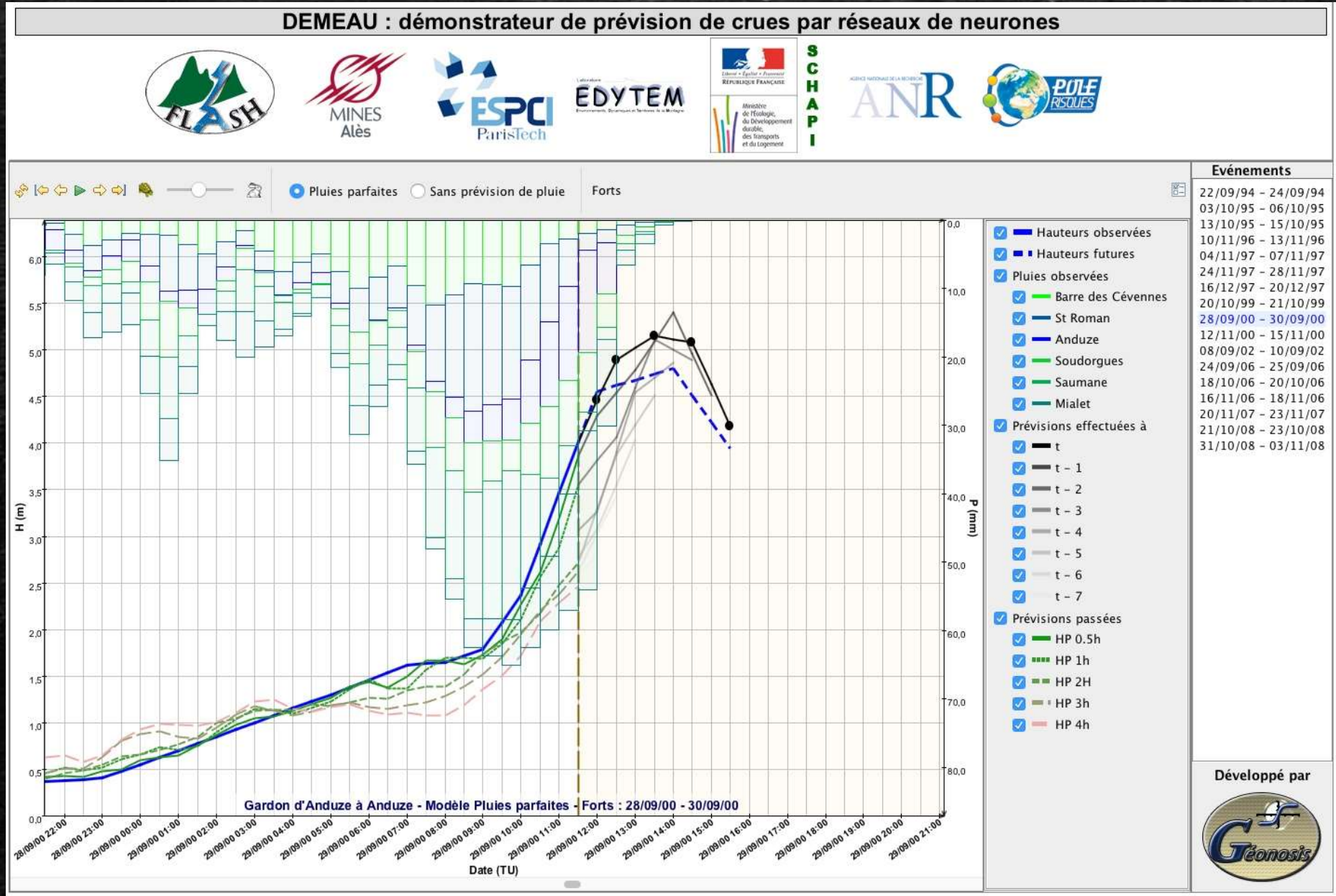


Prévision des hauteurs d'eau à Anduze, en l'absence de prévision de pluies, pour 6 horizons de prévision

Évènement de septembre 2002



■ Démonstrateur DEM'EAU

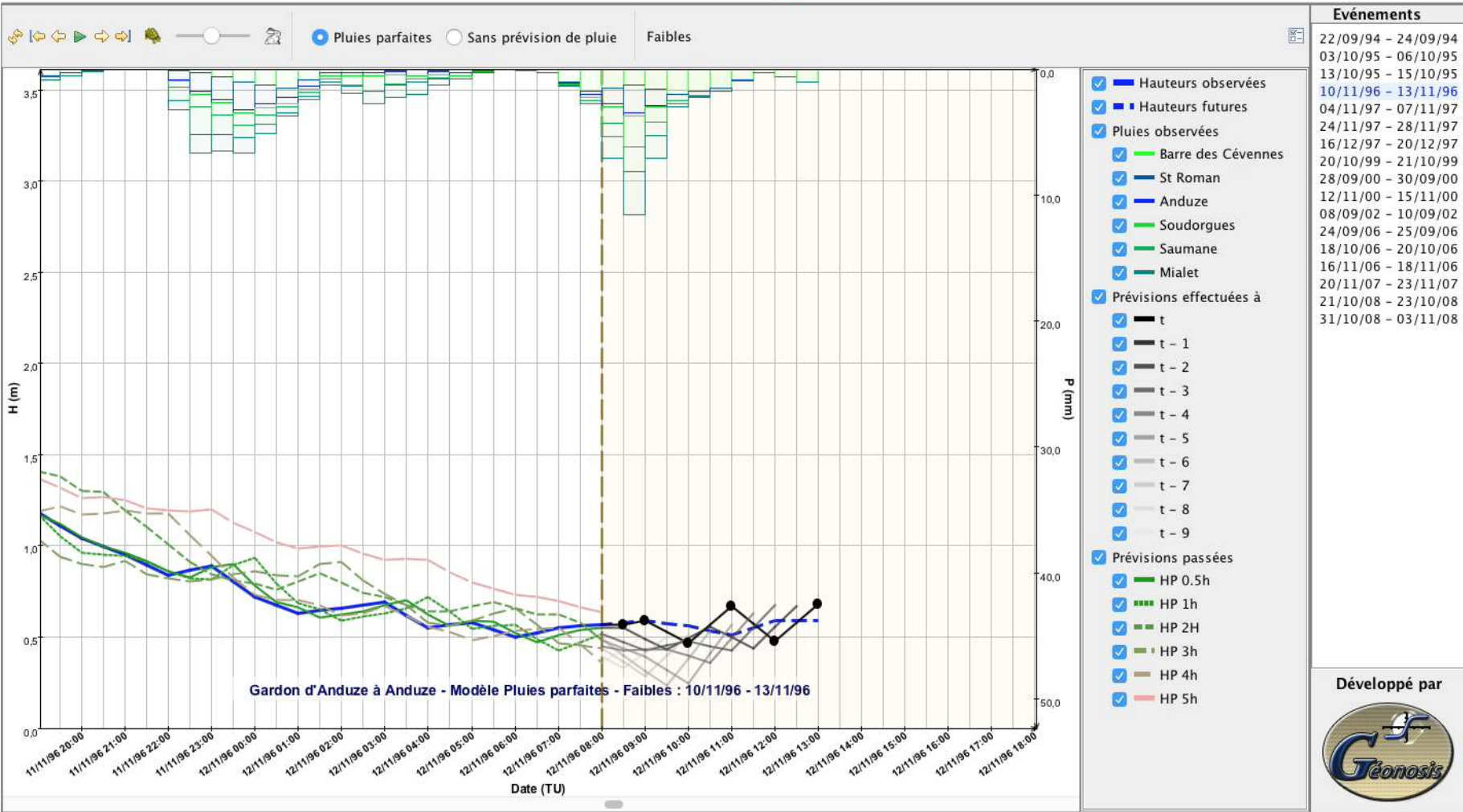


■ Démonstrateur DEM'EAU

DEMEAU : démonstrateur de prévision de crues par réseaux de neurones



S
C
H
A
P
I



Prévision des crues

- Thèse de Guillaume Artigue (co-Direction avec S. Piste et V. Borrell, UM)



Financement SCHAPI



Franchissement de l'Alzon, le 9 septembre 2002 à St-Jean-du-Pin

« Prévision des Crues Éclair par Réseaux de Neurones : Généralisation aux Bassins non Jaugés »



Prévision des crues

- Thèse de Guillaume Artigue (co-Direction avec S. Piste et V. Borrell, UM)



Financement SCHAPI



Franchissement de l'Alzon, le 9 septembre 2002 à St-Jean-du-Pin

- Objectif : réaliser des prévisions sur les petits bassins non équipés de station de mesures (les plus nombreux, les plus rapides)



Prévision des crues

- Thèse de Guillaume Artigue (co-Direction avec S. Piste et V. Borrell, UM)



Financement SCHAPI



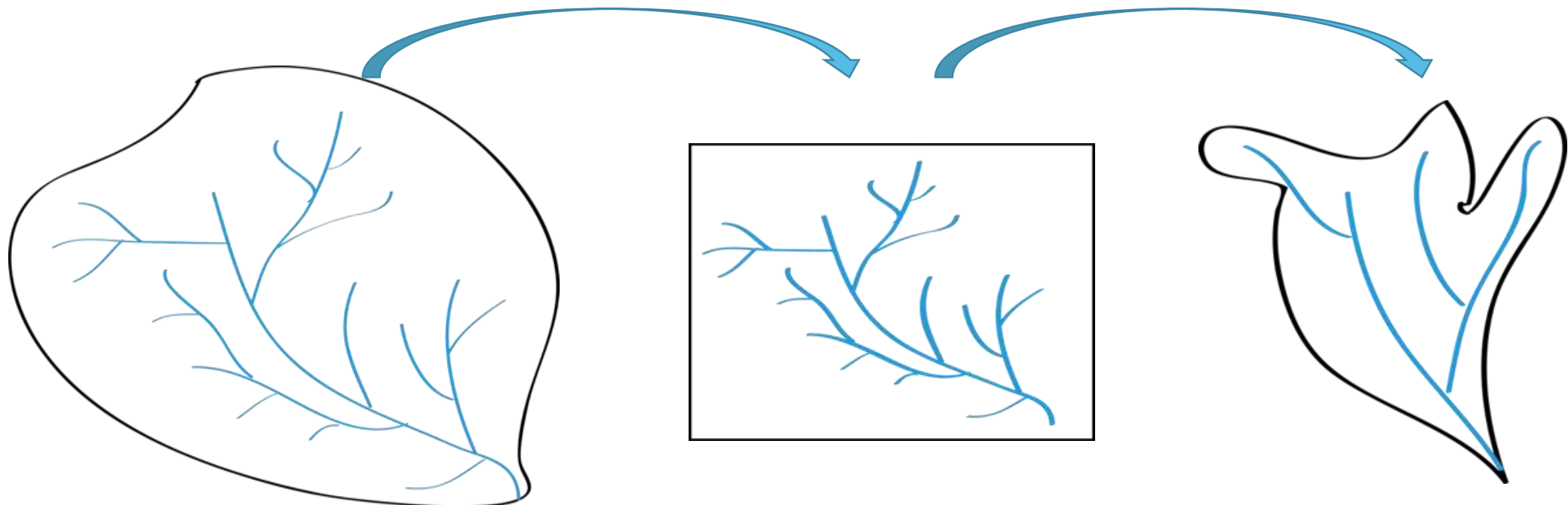
Franchissement de l'Alzon, le 9 septembre 2002 à St-Jean-du-Pin

■ Verrou scientifique : comment obtenir un modèle statistique sans base de données ?



Prévision des crues

- Thèse de Guillaume Artigue (co-Direction avec S. Piste et V. Borrell, UM)



Bassin versant pilote
Pour lequel on dispose d'une
base de donnée



Bassin versant normalisé
(facteurs géographiques
ou hydrologiques)



Bassin versant
cible

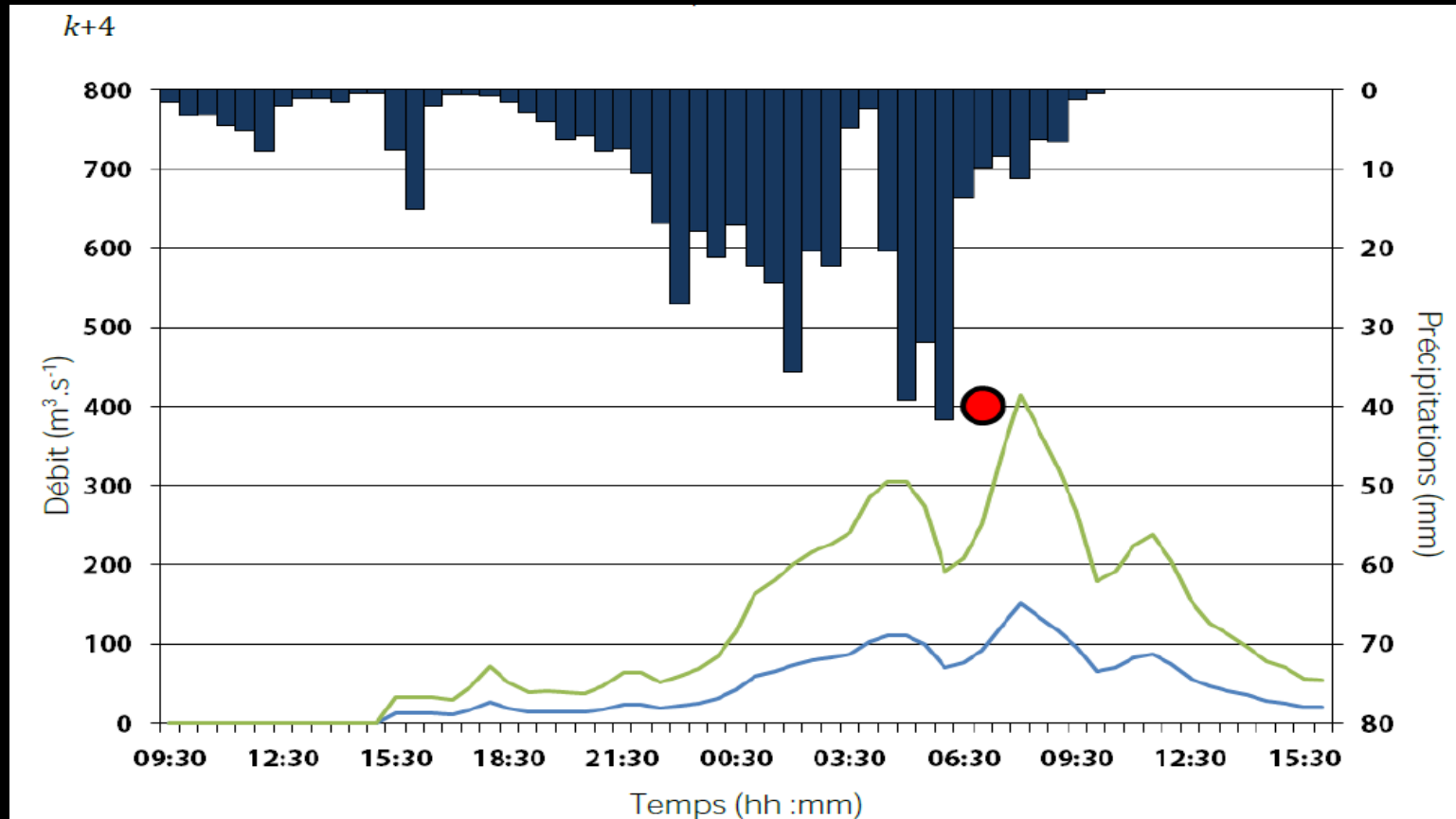


Prévision des crues

- Thèse de Guillaume Artigue (co-Direction avec S. Piste et V. Borrell, UM)



Prévision des débits de l'Avène, en l'absence de prévision de pluies, 2 h à l'avance (bleu : modèle normalisé ; vert : modèle dénormalisé)
Évènement de septembre 2002



Merci de votre attention !
