Aide à la décision en phase de prévision de phénomènes inondants par un système à base de règles floues

Rule based system for decision aid in flood phenomenon prediction phase

IRIT-TCI

Petiot Guillaume Dalle Patrice

IRIT–TCI, Université Paul Sabatier, 118 route de Narbonne, 31062, TOULOUSE.

petiot@irit.fr, dalle@irit.fr

Résumé : Dans le domaine de la prévision des crues, de nombreuses connaissances sont utiles pour construire des indicateurs pertinents répondant aux besoins des previsionnistes en matiere d'aide à la decision. Les sources d'information prises en compte dans le calcul des indicateurs sont de natures differentes (numeriques, linguistiques, images,…). Certaines informations peuvent provenir de SIG (Système d'Information Geographique). On s'interesse au raisonnement dans les SIG. Pour cela, les connaissances expertes sont traduites sous la forme de regles floues. L'application des regles realise une combinaison des sources d'informations. Nous obtenons, en resultat une carte de potentialites par indicateur qui traduit son degre de satisfaction.

Mots clefs : Résolution de problèmes, ensembles flous, intelligence artificielle, traitement d'image, fusion de donnees.

Abstract : In the field of flood risk prediction, domain knowledge is required to build pertinent indicators that meet forecaster needs for decision aid. The information sources used to compute indicators have different types (numerical, linguistics, images,…), some of them may be present in GIS (Geographic Information System). We are focusing our interest to GIS reasoning, so we translate expert knowledge into fuzzy rules. The application of rules provides a fusion of information sources. We get as a result a potentiality map for each indicator that represents its satisfaction degree.

Keywords : Problem resolution, fuzzy sets, artificial intelligence, image processing, data fusion.

1. Introduction.

Cette recherche concerne les problèmes d'aide à la decision des experts previsionnistes des Services d'Annonce des Crues (SAC). Nous l'avons appliquee au cas des SAC geres par la DIREN Midi-Pyrenees. Un besoin majeur exprime par les previsionnistes est la combinaison, à partir de connaissances expertes formalisees, de nombreuses sources d'information (linguistiques, images, capteurs,…), dans le but de fournir des indicateurs pertinents pour les aider dans leur prise de decision. Actuellement, les outils d'aide à la decision utilises par les previsionnistes ne prennent en compte que des donnees capteurs et ne sont pas assez generiques. La presentation du resultat n'est pas toujours facile à interpreter notamment pour des previsionnistes manquant encore d'experience. Dans cet article nous proposons une etude d'indicateurs elabores à partir d'une methode de resolution de problemes. Les connaissances concernant les indicateurs mettent en relation des sources d'information de differents types dans un SIG, qui peuvent etre combinees en utilisant des regles floues [6], et un moteur

de raisonnement permettant de gérer l'incertain. En résultat nous fournissons, pour chaque indicateur, une carte de potentialités dans laquelle le niveau de gris du pixel est proportionnel au degré de satisfaction de l'indicateur. Les cartes de potentialités produites constituent une source d'information visuelle facile à interpréter et exploitable dans un Système d'Information et d'Aide à la Décision.

2. Les données.

Les données présentes au sein de la DIREN Midi-Pyrénées pour la surveillance du bassin de la Garonne sont des relèves capteurs issus de la télécollecte par réseau radio de 110 stations d'observation qui mesurent automatiquement les niveaux d'eau des rivières et, pour la plupart, les quantités de pluie tombées au sol. Les autres données proviennent de Météotel, de VALERHY[1] pour les pluies RADAR, du logiciel de modélisation SOPHIE2 pour les prévisions et des experts pour les connaissances qui ne sont pas toujours formalisées. Ces données peuvent se présenter sous la forme de bases de données, les hauteurs d'eau étant exprimées en mètres, les débits en litre par seconde, et les données pluviométriques en mm de pluie par heure.

Les prévisionnistes ont défini trois seuils aux stations de mesure de la hauteur d'eau, à partir des historiques de crues, des impacts sociaux économiques que peut avoir une crue, et de leurs connaissances personnelles du terrain. Ces seuils sont les seuils de vigilance, de pré-alerte et d'alerte. Il peut arriver que certains seuils soient manquant.

Nous transformons certaines données des bases de données pour les représenter sous forme d'images. Les données des capteurs de hauteur d'eau ou de quantité de pluie tombée au sol, pour une même séquence d'enregistrement, dont les positions géoréférencées sont connues, sont replacées dans une image après application d'une fonction de transformation de coordonnées géoréférencées en coordonnées image. Le résultat obtenu est alors superposable au réseau hydrographique. De même, pour les données concernant les seuils, nous obtenons trois images une pour les seuils de vigilance, une pour les seuils de pré-alerte et une pour les seuils d'alerte. Ce choix a été fait pour faciliter un raisonnement sur des images qui permettent une meilleure appréhension globale des informations.

3. Représentation des connaissances à l'aide des ensembles flous.

3.1. Introduction.

Le cadre représentatif des ensembles flous permet de traduire les termes du langage utilisés par les experts. Les degrés d'appartenances aux ensembles flous peuvent alors être combinés par des opérateurs flous afin d'évaluer des règles. Enfin, les résultats obtenus en utilisant les ensembles flous sont plus proches de la réalité et plus faciles à interpréter. Les ensembles flous utilisés dans notre étude sont décrits ci-dessous.

3.2. Proximité de dépassement d'un seuil.

3.2.1. Construction des sous-ensembles flous.

La première étape a consisté à traduire les expressions: « proche de dépasser le seuil de vigilance », « proche de dépasser le seuil de préalerte », et « proche de dépasser le seuil d'alerte » en utilisant trois sous-ensembles flous notés respectivement E_v, E_{pa}, E_a. Ces sous-ensembles flous sont construits à partir de la donnée des seuils, d'une marge pour chaque seuil notée

[1] VALEHRY est un logiciel qui permet d'exploiter les images RADAR et de représenter les lames d'eau.

m_i avec $i \in \{v, pa, a\}$, de contraintes sur les marges, et d'un degré de contrainte $X \in [0,1]$ (Figure 1). On retrouve l'allure de ces sous-ensembles flous :

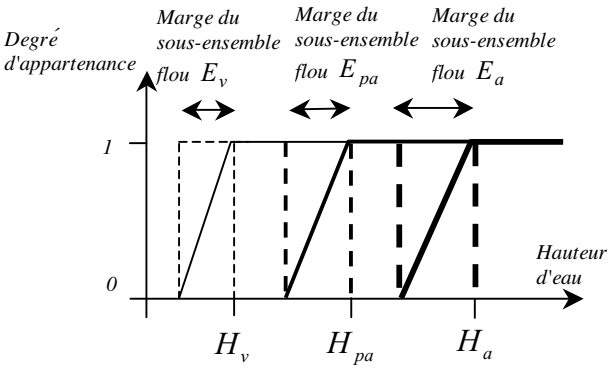


Figure 1 : proche de depasser un seuil.

Les marges des sous-ensembles flous sont calculees, si le seuil correspondant au sous-ensemble flou est connu, en utilisant une contrainte C_{maxi} differente pour chaque sous-ensemble flou. Elle symbolise la valeur maximale que peut prendre la marge. Si nous designons par H_v, H_{pa}, et H_a les valeurs respectives des seuils de vigilance de pre-alerte et d'alerte.

- Pour l'ensemble flou E_v : $C^{v}_{\text{maxi}} = H_v$
- Pour l'ensemble flou E_{pa} :

$$C^{\text{pa}}_{\text{maxi}} = \begin{cases} H_{pa} - H_v \text{ si } H_v \text{ connu,} \\ H_{pa} \text{ sinon.} \end{cases}$$

- Pour l'ensemble flou E_a :

$$C^{\text{a}}_{\text{maxi}} = \begin{cases} H_a - H_{pa} \text{ si } H_{pa} \text{ connu,} \\ H_a - H_v \text{ sinon si } H_v \text{ connu,} \\ H_a \text{ sinon.} \end{cases}$$

La valeur de la marge pour chaque sous-ensemble flou vaut donc :

$$\text{marge}_i = \begin{cases} X \times m_i \text{ si X} \times m_i < \text{C}^{\text{i}}_{\text{maxi}}, \\ C^{\text{i}}_{\text{maxi}} \text{ sinon.} \end{cases}$$

3.2.2. Filtrage.

Pour une hauteur d'eau, nous obtenons trois degres d'appartenance pour les trois sous-ensembles flous. On fait l'hypothèse qu'il existe une relation d'ordre entre les differents etats correspondant aux sous-ensembles flous decrits precedemment. Le sous-ensemble flou E_a sera considere comme le plus prioritaire, ensuite le sous-ensemble flou E_{pa}, le moins prioritaire etant le sous-ensemble flou E_v. Le filtrage consiste à choisir, pour une station, le degre d'appartenance non nul du sous-ensemble flou le plus prioritaire.

3.3. Les sous-ensembles flous « faible », « moyenne » et « forte ».

Nous utilisons egalement les ensembles flous pour modeliser certaines connaissances imprecises, concernant les pluies, les vitesses ou les accelerations. Les sous-ensembles flous modelises sont

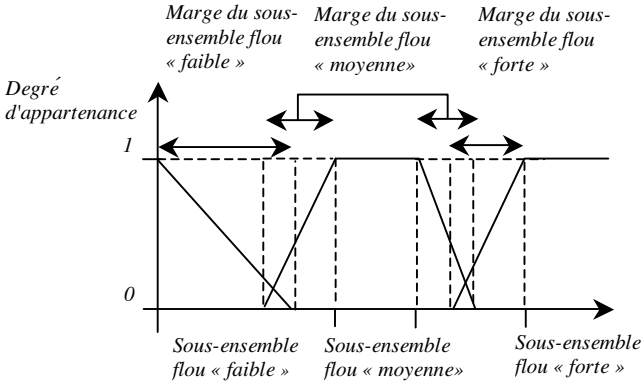


Figure 2 : sous-ensembles flous « faible», « moyenne », « forte».

« faible », « moyenne », et « forte ».

3.4. Sous-ensembles flou associés aux seuils.

3.4.1. Sous-ensemble flou « entre ».

Pour donner un sens à « entre les seuils *i* et *j* » ou i et j peuvent etre distinctement les seuils de vigilance, de pre-alerte et d'alerte, nous utilisons le sous-ensemble flou de la Figure 3.a calcule à partir des deux seuils *i* et *j* et la donnee d'une marge m.

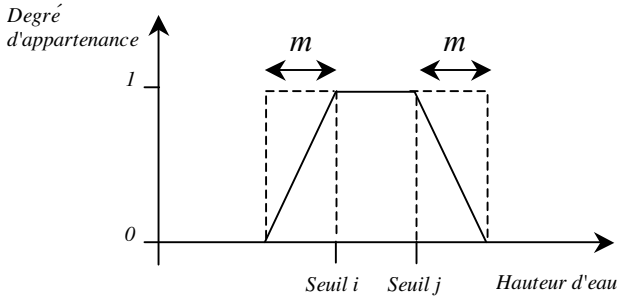


Figure 3.a : le sous-ensemble flou « entre ».

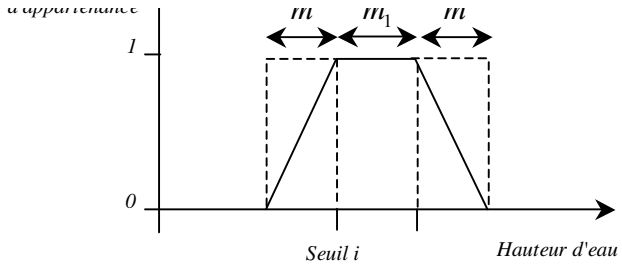


Figure 3.b : le sous-ensemble flou « dépassement ».

3.4.2. Sous-ensemble flou « depassement ».

Dans le cas où l'on a besoin de savoir si une hauteur d'eau estimee depasse faiblement un seuil, on peut utiliser un sous-ensemble flou un peu similaire à celui decrit precedemment mais calcule à partir de la donnee d'un seuil *i* et de deux marges *m* et m_1 (Figure 3.b).

4. Exploitation des connaissances : description du système.

Notre objectif était d'améliorer l'approche basee sur un système expert existant classique utilise par la DIREN Midi-Pyrenees. Dans celui-ci, le resultat de l'application des regles donne pour chaque station un diagnostic, pouvant etre un etat de vigilance, de pre-alerte ou d'alerte, des consignes et des previsions. Aucune nuance n'est faite sur le degre de realisation d'une prevision ou d'un diagnostic. Dans notre approche, nous proposons un systeme à base de regles floues dans lequel on calcule des indicateurs sous formes d'images ou le niveau de gris des pixels est proportionnel au degre de realisation de l'indicateur. A chaque indicateur sont associees des consignes et des explications.

L'architecture fonctionnelle du système utilise est decrite figure 4.

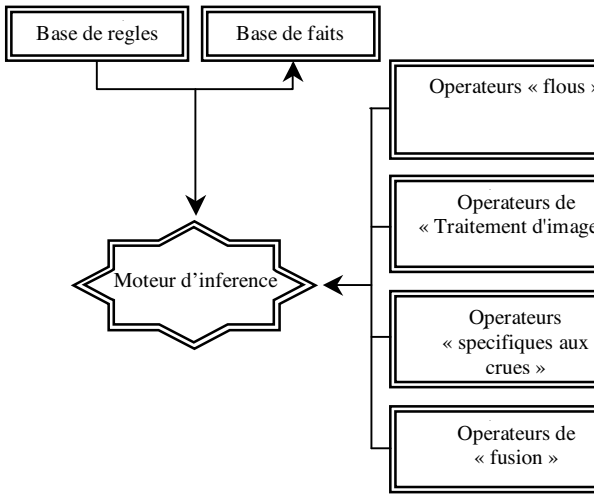


Figure 4 : architecture du système.

Nous avons defini une syntaxe permettant de traduire la connaissance des experts en regles et nous allons decrire le processus d'application des regles. Nous appelons « regle » la traduction d'une expertise, sous la forme « *si* condition *alors* conclusion ». Deux types de regles existent : des regles operant uniquement sur les donnees des stations, pour lesquelles le temps de calcul est assez faible et des regles combinant des donnees images, dont le temps de calcul est plus long. La condition d'une regle est decrite à partir de faits (« premisses ») et d'operateurs de combinaison flous. La conclusion d'une regle est un fait dont le nom peut deja figure dans la base de faits au moment de l'application de la regle. Il faudra alors mettre à jour ce fait. Dans la base d'operateurs « fusion », on retrouve donc des operateurs de fusion de donnees [10], comme par exemple l'operateur OWA[2], des operateurs utilises pour combiner les premisses [6] (actuellement la negation, la

[2] OWA [8,9] : « Ordered Weighted Averanging operator ». Creé par YAGER, il permet de combiner des donnees en utilisant une ponderation. La fusion des informations est obtenue par le produit scalaire du vecteur de donnees trie et du vecteur poids. On notera que l'operateur OWA permet de representer des connaissances relativement elaborees, mais que la definition des coefficients intervenant dans l'operateur n'est pas forcement simple à gerer.

conjonction floue, et la disjonction floue), et des operateurs qui permettent de combiner les conclusions.

Un « coefficient de confiance » $CF_r \in [0,1]$ est attribue à chaque regle en fonction de sa certitude. Ainsi, une regle toujours fausse a un CF_r egal à 0 et une regle garantissant avec certitude un resultat a un CF_r egal à 1.

Decrivons plus en detail le processus d'inference de la base de regles [1,2,3,4]. Nous reutilisons des operateurs deja etudies dans [6]. Trois cas peuvent survenir :

Premier cas : la regle que l'on souhaite appliquer, de coefficient de confiance CF_r, a un membre droit $CF_c(i,j)$ qui ne se trouve ni dans la combinaison des premisses ni dans la base de faits. Si le calcul de la combinaison des premisses vaut $CF_p(i,j)$ pour un pixel (i,j), alors le nouveau fait $CF_c(i,j)$ associe vaut :

$$CF_c(i,j) = CF_r \times CF_p(i,j)$$

Deuxieme cas : la conclusion de la regle existe deja dans la base de faits. Il faut fusionner le fait existant x avec le nouveau fait y. Comme dans MYCIN [7], on se ramene à des valeurs de x et y comprises entre $[-1,1]$. Deux cas sont possibles : soit les informations sont « concordantes » (x et y sont de meme signe) et on utilise une des deux fonctions de fusion suivantes :

$$G_1(x,y) = x + y - xy \quad \text{si } x \geq 0 \text{ et } y \geq 0$$

$$G_2(x,y) = x + y + xy \quad \text{si } x < 0 \text{ et } y < 0$$

Soit elles sont « conflictuelles » (x et y sont de signe oppose) et on realise un compromis :

$$G_3(x,y) = \begin{cases} \dfrac{x+y}{1-\min(|x|,|y|)} & \text{si } xy \leq 0 \text{ et } |xy| \neq 1 \\ 1 & \text{si } xy \leq 0 \text{ et } |xy| \neq 1 \end{cases}$$

Troisieme cas : la regle est « auto-referencee » si son membre droit est present en tant que fait dans le membre gauche.

Si ($CF_c(i,j) \geq 0$ et $CF_r \times CF_p(i,j) \geq 0$) alors

$$NCF_c(i,j) = G_1(CF_c(i,j), CF_r \times CF_p(i,j))$$

Sinon si ($CF_c(i,j) < 0$ et $CF_r \times CF_p(i,j) < 0$) alors

$$NCF_c(i,j) = G_2(CF_c(i,j), CF_r \times CF_p(i,j))$$

Sinon si ($CF_c(i,j) \times CF_r \times CF_p(i,j) \leq 0$ et

$|CF_c(i,j) \times CF_r \times CF_p(i,j)| \neq 1$) alors

$$NCF_c(i,j) = G_3(CF_c(i,j), CF_r \times CF_p(i,j))$$

Sinon $NCF_c(i,j) = 1$

Dans ce cas, on applique l'algorithme suivant :

où $CF_c(i,j)$ designe le fait conclusion avant inference de la regle et $NCF_c(i,j)$ sa nouvelle valeur. L'interet de ce systeme est de permettre un effet de renforcement lorsque les resultats obtenus sont concordants et de realiser un compromis lorsqu'ils sont conflictuels. Les resultats se presentent sous la forme de carte de potentialites dans lesquelles la valeur d'un pixel est proportionnelle au degre de satisfaction du pixel au probleme pose.

5. Resultats.

5.1. Application à la crue du 29/01/01.

Nous avons realise une premiere application sur le bassin de la Garonne pour l'evenement du 29/01/2001 au 02/02/2001 ou le seuil d'alerte a ete depasse pour la station « le Fossat » et le seuil de pre-alerte pour la station « le Mas d'Azil ».

Nous illustrons l'application de regles definies uniquement pour les stations. Ci-dessous les hydrogrammes pour « le Fossat » (Figure 5) et « le Mas d'Azil » (Figure 6) :

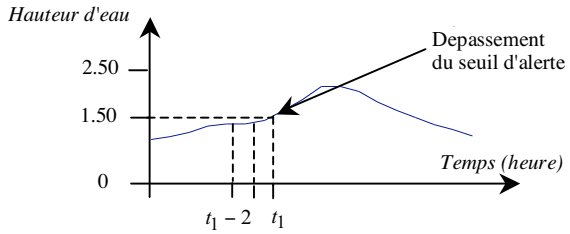


Figure 5 : hydrogramme « le Fossat ».

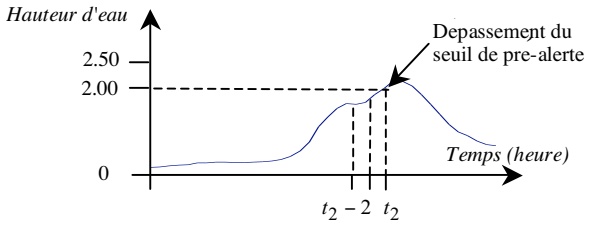


Figure 6 : hydrogramme « le Mas d'Azil ».

Les instants t_1 et t_2 sont distincts et correspondent au depassement du seuil d'alerte et de pre-alerte. La base de regles utilisee comprend une centaine de regles qui ont ete elaborees à partir des connaissances des experts de la DIREN Midi-pyrenees. Nous avons, concernant la station « le Fossat », obtenu le resultat de la Figure 7 et 8, qui correspond à une anticipation du depassement du seuil d'alerte. Les regles extraites de la base de regles qui nous ont permis d'obtenir cet indicateur sont les suivantes :

Si (*accélération_exceptionnelle* **ou** *accélération_très_forte*

ou *accélération_forte*) **et** (*proche_dépassement_alerte* **ou** *prévue_proche_dépassement_alerte*)

Alors *dépassement_probable_du_seuil_d_alerte* ; $CF_r = 1.0$

Si ***non*** *proche_dépassement_alerte* **et** *Entre*(*alerte*,*cote_prévue*,20,20)

Alors *dépassement_probable_du_seuil_d_alerte* ; $CF_r = 1.0$

Les autres regles ont donne des indicateurs dont le degre etait nul pour toute les stations. Les faits « acceleration exceptionnelle », « acceleration tres forte », et « acceleration forte » sont calcules à partir des images de hauteurs d'eau pour l'acceleration et des sous-ensembles flous « moyenne » et « forte ». Le fait « cote prevue » represente la prevision à 1h des hauteurs d'eau. Le fait « proche depassement alerte » et le fait « prevue proche depassement alerte » sont calcules à partir de l'ensemble flou « proche de depasser un seuil » pour la cote et la cote prevue. Enfin dans la derniere regle, la fonction « Entre » correspond à une situation ou la cote courante est eloignee du seuil et la cote prevue depasse de peu le seuil. Les deux parametres de la fonction sont les deux marges de l'ensemble flou decrit au §3.3. Le

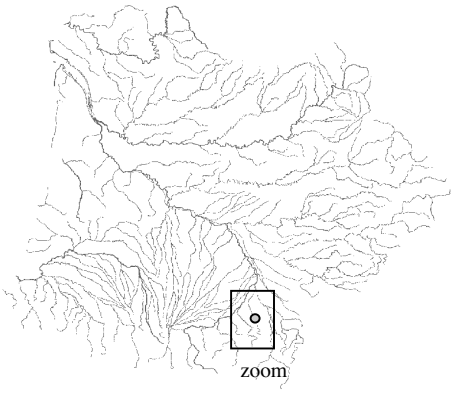


Figure 7 : indicateur dépassement du seuil d'alerte.

resultat de l'indicateur superpose sur l'image des rivieres est le suivant :

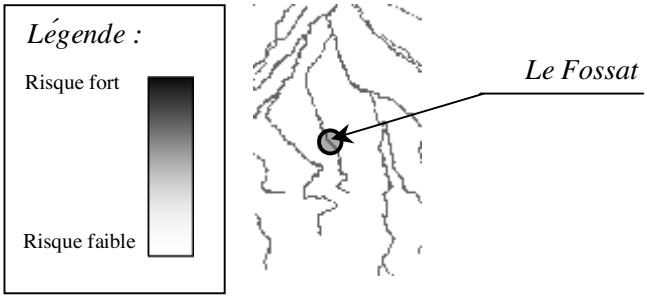


Figure 8 : station qui risque de dépasser le seuil d'alerte.

On retrouve ci-dessous la station qui risque de depasser le seuil :

Les stations dont le degre est nul ne figurent pas dans le resultat. Seules les stations qui evoluent sont representees. Nous obtenons des resultats similaires pour l'anticipation du depassement du seuil de pre-alerte concernant le « Mas d'Azil ». La regle qui

nous a permis d'obtenir cet indicateur est la suivante :

Si *non* *proche_dépassement_préalerte* **et** *Entre*(*préalerte*,*cote_prévue*,20,20)

Alors *dépassement_probable_du_seuil_de_préalerte* ; $CF_r = 1.0$

5.2. Application à la crue du 04/07/01.

Nous nous intéressons cette fois aux règles combinant des données de type image et nous prenons comme exemple le dépassement du seuil de vigilance de la station Arreau-Aure. Nous sommes partis de la même base de règles que précédemment à laquelle nous avons rajouté l'indicateur VSP (Vitesse, Saturation des sols, Proche de dépasser un seuil). La saturation des sols est estimée à partir d'un cumul de pluies RADAR (Figure 9). Plus précisément, on calcule la saturation des sols par secteur du bassin Adour-Garonne (Figure 10) en prenant le cumul de pluie maximum par secteur (Figure 11). En effet une étude [5] a montré les bons résultats obtenus par cette approche. Cependant le même calcul est envisageable en fusionnant les données des capteurs pluviomètres afin d'en déduire la valeur pluviométrique la plus représentative du secteur. Pour cela, nous avons implémenté des opérateurs de fusion dans notre système. Des études sont en cours afin de confronter l'approche multicapteur et l'approche RADAR.

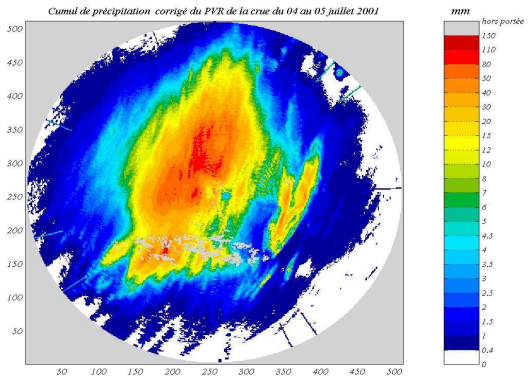


Figure 9 : cumul des pluies RADAR.

Les règles qui ont produit l'indicateur

Figure 11 : cumul maximum par secteur.

d'anticipation du dépassement du seuil de vigilance sont les suivantes :

Si *Forte*(*vitesse*,15,5) **et** Forte(*cumul*,100,10) **et** *proche_vigilance*

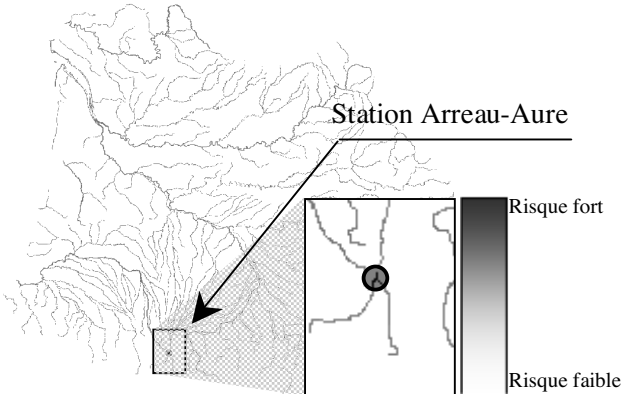
Alors *dépassement_probable_du_seuil_de_vigilance* ; $CF_r = 1.0$

Si non *proche_vigilance* **et**

(*Forte*(*Dépassement*(*vigilance*,*cote_prévue*),20,20) **ou** *Entre*(*vigilance*,*cote_prévue*,20,20))

Alors *dépassement_probable_du_seuil_de_vigilance* ; $CF_r = 1.0$

Dans ces règles, le fait « cumul » correspond à l'image du cumul Figure 11 dans laquelle le niveau de gris est proportionnel au cumul des pluies. La fonction dépassement calcule la différence entre la cote prévue et le seuil de vigilance, la valeur calculée vaut 0 si la cote prévue est inférieure au seuil de vigilance. Figure 12 on retrouve le résultat après application des règles :



6. Conclusion.

Les objectifs fixés sont atteints : nous avons traduit les connaissances expertes sous forme de regles floues afin d'anticiper des depassements de seuil en combinant des sources d'informations de natures differentes. Ces techniques utilisant l'approche flou et la fusion nous permettent de definir relativement simplement de nouveaux indicateurs (exemple VSP §5.2). Il apparaît cependant que si trop d'indicateurs sont produits par le systeme on risque de creer un effet de surcharge pouvant gener le previsionniste. Des considerations ergonomiques doivent donc etre prises en compte pour realiser l'interface. A chaque indicateur est associe une couleur. L'intensite des pixels dans l'image est proportionnelle au degre de realisation de l'indicateur. Pour tous les indicateurs nous obtenons une carte de potentialites à laquelle sont associees des consignes destinees aux previsionnistes et des explications. Le temps de calcul des indicateurs est de l'ordre de 15 secondes pour des images (768,1024). Cependant la frequence de rafraîchissement des donnees est de 15 minutes pour les images RADAR et d'une heure pour les releves des stations, le temps de calcul est donc negligeable. Ces techniques vont nous permettre d'etudier des indicateurs plus elabores permettant de prendre en compte des informations spatiales, des hauteurs d'eau spatialisees issues de modeles hydrauliques, des informations extraites d'un MNT (Modele Numerique de Terrain),... D'autre part, nous etudions la possibilite de definir des indicateurs plus globaux tel que le contexte d'une situation. Ces nouveaux indicateurs seront etudies en utilisant le systeme à base de regles existant et d'autres approches dont les reseaux bayesiens.

Remerciements :

Cette étude a bénéficié d'un soutien financier de la DIREN Midi-Pyrenees qui a egalement fourni l'expertise des prévisionnistes. Les auteurs tiennent à remercier tout particulierement P. Cornuau, N. Watrin, P. Fontana et J. J. Vidal pour leurs remarques pertinentes et leurs aides tout au long de ce travail.

7. Références.

[1] I. Bloch, *Information combination operation for data fusion : a comparative review with classification*, IEEE trans. on systems, Man and cybernetics, 26, 1996, 52-67.

[2] J. Desachy. *Interprétation automatique d'images satellites : le système ICARE*. These d'etat, Universite Paul Sabatier, Toulouse, France, 1991.

[3] J. Desachy. *Numeric and symbolic fuzzy information fusion applied to satellite imagery classification*. Conference invitee, CIMA'99, International Congress on Computational Intelligence : Methods and Applications, Rochester, USA, juin 1999.

[4] J. Desachy, V. Bessettes, P. Dherete. *Remote sensed images processing and geographic features*. Chapitre d'un livre « Information Processing for Remote Sensing », edite par C. H. Chen, World Scientific Publishing Co, 2000.

[5] A. Mimouni, H. Sauvageot, P. Cornuau, J.J. Vidal, P. Fourmigue, J. Lavabre. Des lames d'eau radar dans la modelisation pluie debit pour la prevision des crues : L'exemple du bassin versant du Mas d'Azil. Atelier de modelisation de l'athmosphere, CNRM, Meteo-France, Toulouse, 2001.

[6] D. Saint-Joan. *Données géographiques et raisonnement : le système GEODES*. These d'etat, Universite Paul Sabatier, Toulouse, France, 1995.

[7] E. H. Shortliffe, B. Buchanan. *A model of inexact reasoning in medecine*. Mathematical Biosciences, vol. 23, pages 351-379, 1975.

[8] R. R. Yager. *On ordered weighted averaging agregation operators in multicriteria decision making*. Didier Dubois, Henri Prade et Ronald R. Yager. dans Fuzzy sets for intelligent systems, pages 80-87, 1993.

[9] R. R. Yager, J. Kacprzyk. *The Ordered Weighted Averaging Operation : Theory, Methodology and Applications*. Kluwer, Norwell, MA, USA, 1997.

[10] D. Dubois, H. Prade, R. R. Yager. *Merging fuzzy information*. Fuzzy sets in approximate and information systems, the handbook of fuzzy sets series, Kluwer Academic Publ.,Derdrec ht, 1999.