

## **Les Analogues, une approche statistique adaptée pour la prévision opérationnelle des crues et étendue à l’ensemble de la France**

*Analog sorting of meteorological patterns, a statistical approach adapted to operational flood forecast and extended to the whole of France*

**Auteurs :** MARTY Renaud<sup>1\*</sup>, GAUTHERON Alain<sup>2</sup>, EDOUARD Simon<sup>2</sup>, HORTON Pascal<sup>3</sup>, OBLED Charles<sup>4</sup>

\* *auteur correspondant*

<sup>1</sup> DREAL Centre-Val de Loire, 5 avenue Buffon, CS 96407, 45065 Orléans, France, renaud.marty@developpement-durable.gouv.fr

<sup>2</sup> DREAL Auvergne-Rhône-Alpes, 17 boulevard Joseph Vallier, 38030 Grenoble Cedex 2, France, alain.gautheron@developpement-durable.gouv.fr, simon.edouard@developpement-durable.gouv.fr

<sup>3</sup> Terranum, rue de l’Industrie 35b, 1030 Bussigny, Suisse, pascal.horton@terranum.ch

<sup>4</sup> Université Grenoble-Alpes – Institut des Géosciences de l’Environnement, CS 40700, 38058 Grenoble Cedex 9, charles.obled@univ-grenoble-alpes.fr

**Résumé :** La méthode des Analogues est une adaptation statistique des prévisions météorologiques à l’échelle synoptique. Son principe peut s’énoncer clairement : à situations météorologiques semblables, les effets seront comparables. Les Service de Prévision des Crues (SPC) Alpes-du-Nord et Loire-Allier-Cher-Indre disposent, au sein de leur système de prévision, d’outils opérationnels directement issus des travaux de recherche : l’arrivée de cette approche dans la sphère du prévisionniste Vigicrues est la conclusion d’une riche histoire débutée à la fin des années 1960, ponctuées de nombreuses améliorations grâce à l’arrivée des réanalyses météorologiques, la définition de critères de similarité et de performance pertinents dans le contexte de la prévision probabiliste. Le réseau Vigicrues mène un projet visant à étendre la méthode à la France métropolitaine et à mettre en place une suite logicielle unique. Cela nécessite notamment de réduire le nombre de modèles d’analogie par la répartition des mailles SAFRAN en 30 groupes, de déterminer de nouveaux paramétrages optimaux reposant sur une analogie saisonnière puis une analogie évaluée sur 4 champs géopotentiels, et d’adapter la suite logicielle AtmoSwing aux pratiques des SPC et aux contraintes de leurs systèmes opérationnels. Le dernier défi à accomplir réside dans la construction de scénarios hydrométéorologiques cohérents spatio-temporellement et adaptés à la modélisation hydrologique du réseau Vigicrues, rendue possible par l’arrivée de la prévision météorologique d’ensemble et de nouvelles méthodologies de post-traitement.

**Mots-clefs :** Adaptation statistique, prévision probabiliste de précipitations, prévision hydrologique opérationnelle

**Abstract :** The analog sorting approach is a statistical adaptation of meteorological forecasts at synoptic-scale. Its principle can be clearly stated : the effects of similar weather situations will be comparable. The Regional Flood Forecasting Service (SPC) Alpes-du-Nord and Loire-Allier-Cher-Indre have, within their forecasting system, operational tools resulting directly from research : the arrival of this approach in the sphere of Vigicrues is the conclusion of a rich history

started at the end of the 1960s, taking advantage of numerous improvements as the arrival of meteorological reanalyses, the definition of criteria of similarity and of relevant verification score in the context of probabilistic precipitation forecast. Vigicrues is leading a project aimed at extending the method to mainland France and setting up a unique software suite. This notably requires reducing the number of analogy models by classifying SAFRAN meshes into 30 groups, determining new optimal parameters based on a seasonal analogy and then on an analogy established on 4 geopotential fields, and adapting the AtmoSwing software suite to operational systems in order to facilitate the use in SPC. The last challenge to be accomplished is the construction of hydrometeorological scenarios that are spatially and temporally consistent and adapted to hydrological modeling. It is now possible thanks to the availability of ensemble weather forecasts and new post-processing technics.

**Keywords :** statistical postprocessing, probabilistic precipitation forecast, operational flood forecasting

## 1 ELABORATION DE LA METHODE ET SES PREMIERS USAGES

### 1.1 Des premières techniques de prévision de précipitations journalières aux systèmes opérationnels en France

La méthode des Analogues est une adaptation statistique des prévisions météorologiques à l'échelle synoptique. Son principe peut s'énoncer clairement : à situations météorologiques semblables, les effets seront comparables. Cette hypothèse correspond à la démarche naturelle du prévisionniste qui, à la vue d'une situation météorologique, se rappelle des situations passées similaires et construit sa prévision en s'aidant des souvenirs qu'il garde de ces situations. L'arrivée de cette approche dans la sphère du prévisionniste Vigicrues est la conclusion d'une riche histoire débutée à la fin des années 1960 où D. Duband s'intéresse à la prévision des précipitations journalières par la comparaison des champs géopotentiels 700 hPa sur 25 stations couvrant le proche Atlantique, l'Europe de l'Ouest et le Maghreb, complétés par 11 stations de pression au sol plus centrées sur la France, sur la période 1958-1969 (Duband, 1970, 1980).

Les premières réanalyses météorologiques mises à disposition au cours des années 1990 ont permis à la méthode des Analogues d'évoluer notablement : les données numériques fournies sous forme de grilles couvrent la terre entière à une résolution régulière sur plusieurs décennies. L'analogie de circulation atmosphérique, établie sur les champs géopotentiels, était initialement caractérisée par une analyse en composantes principales et une distance euclidienne. Elle est désormais réalisée directement sur les grilles de réanalyses par le critère « Teweles-Wobus Score » (Guilbaud, 1998). Les prévisions par analogie sont optimisées selon des critères probabilistes comme le RPS « Ranked Probability Score » qui compare la distribution de précipitations prévue par les analogues pour une date donnée avec l'observation à travers un découpage en classes. Guilbaud *et al.* (1998) appliquent ces nouveautés sur l'analogie des champs géopotentiels 700 et 1000 hPa : cette configuration est à l'origine de celle fonctionnant actuellement au sein d'EDF (voir la section 1.2). Bontron (2004) applique la version continue du RPS (« Continuous Rank Probability Skill Score ») et propose une optimisation plus poussée des champs géopotentiels, complétée par une seconde analogie sur les flux d'humidité. Il développe également une approche itérative d'optimisation des domaines d'analogie. Thévenot (2004) complète cette méthodologie en faisant varier le nombre de dates analogues à retenir en fonction de l'échéance afin de tenir compte de l'incertitude des prévisions météorologiques.

Le premier système opérationnel, RainFAST, est mis en place au LTHE (Marty *et al.* 2012): chaque jour, les prévisions synoptiques nécessaires pour définir les situations cible (attendues pour les jours J à J+6)

étaient empruntées au système américain GFS, cohérent avec les réanalyses NCEP. Pour chaque bassin d'intérêt, les résultats étaient résumés dans un graphique des quantiles (20, 60 et 90%) pour chacune des échéances J à J+6. Il s'agit du système utilisé opérationnellement par le SPC Loire-Allier-Cher-Indre (LACI) depuis 2010. Le système OPALE (Caillouet *et al.* 2022), développé et utilisé par la Compagnie Nationale du Rhône, dérive aussi de Bontron (2004), amélioré par Ben Daoud *et al.* (2016), en appliquant les techniques au pas de temps infra-journalier. OPALE est intégré dans le système opérationnel du SPC Alpes du Nord (AN) et maintenu par Kisters France.

Parallèlement, Horton *et al.* (2017) ont reconsidéré l'optimisation de la méthode grâce à une technique beaucoup plus combinatoire que l'optimisation séquentielle des prédicteurs utilisée jusqu'alors: l'optimisation par algorithmes génétiques. Elle permet de nombreux allers et retours entre variables, domaines spatiaux utilisés, etc. et améliore systématiquement les optimums obtenus. La performance de la méthode appliquée sur les précipitations de la Suisse ont conduit au développement de la suite logicielle AtmoSwing (Horton, 2019), décrite à la section 2.5.).

Il est à noter que l'approche par analogie présente des cas d'application variés sur différents territoires et grandeurs. Elle est également utilisée internationalement en vue de proposer de meilleures prévisions de précipitations à partir de modèle déterministe (Delle Monache *et al.* 2013), de prévision d'ensemble (Plenković *et al.* 2020), en tenant compte des erreurs des modèles numériques par le biais d'historique de prévision (Hamill *et al.* 2015, Jeworrek *et al.* 2023). D'autres grandeurs météorologiques sont ciblées, comment le vent (Alessandrini *et al.* 2019) ou la température de l'air (Hou *et al.* 2022).

## **1.2 Utilisation au sein de systèmes opérationnels de prévision hydrologique**

Les systèmes opérationnels de prévision du réseau Vigicrues sont principalement alimentés par la prévision expertisée de Météo-France. En complément de cette source d'information, les SPC Alpes du Nord (AN) et Loire-Allier-Cher-Indre (LACI) disposent de prévisions obtenues selon la méthode des Analogues à partir des configurations optimisées par Bontron (2004), respectivement par les logiciels OPALE et RainFAST. Sur le bassin de la Loire, ces prévisions de précipitations journalières sont alimentées par le modèle numérique américain (GFS, décrit à la section 2.2) et se limitent à la Loire en amont de Bas-en-Basset et à l'Allier en amont de Vieille-Brioude. Sur les Alpes du Nord, les prévisions sont produites sur l'ensemble du territoire pour une trentaine de sous-bassins avec une approche multi-modèle (GFS, ARPEGE, CEP). Les prévisions sont produites au pas de temps journalier sur la base d'une archive permettant de remonter jusqu'en 1958, mais aussi au pas de temps 6 heures mais cette fois-ci sur une archive ne démarrant qu'en 1992. Ces prévisions par analogie sont utilisées pour différents objectifs :

- Fournir un signal de pré-alerte pour activer la montée en charge de l'équipe des prévisionnistes ;
- Fournir des volumes de précipitations au pas de temps journalier ou infra-journalier (6h), répartis ensuite au pas de temps de la modélisation hydrologique (généralement horaire) par la prévision de Météo-France au pas de temps 3 heures ;
- Sélectionner des événements historiques présentant des similitudes météorologiques pour alimenter l'expertise du prévisionniste.

Zalachori *et al.* (2012) montrent que, sur un jeu de 11 bassins versants répartis sur le territoire national, l'utilisation des Analogues pour « habiller » les prévisions d'ensemble améliore significativement la performance de la prévision de précipitations. Cette approche permet en particulier de compenser la sous-dispersion des prévisions d'ensemble dans les courtes échéances et de profiter tout de même de la bonne capacité des EPS à distinguer l'occurrence de la pluie et non-pluie. Partant de ce constat, EDF-DTG a développé l'outil PREDICTOR au sein duquel la prévision de pluie est pré-estimée (avant expertise humaine) selon la méthode MIX qui s'appuie sur la bonne corrélation des prévisions d'ensemble avec la pluie observée, et corrige leur sous-dispersion en les modulant avec les prévisions

par analogie (Gailhard, 2014 ; Garçon, 2014). Chardon *et al.* (2012) présentent les dernières évolutions de cette méthode qui combine les Analogues et les prévisions d’ensemble. Il est important de noter que l’outil opérationnel PREDICTOR est utilisé quotidiennement par les prévisionnistes des Centres Hydro-météorologiques d’EDF-DTG (Grenoble, Toulouse) pour réaliser et expertiser les prévisions probabilistes de débits jusqu’à J+13 sur de nombreux bassins répartis sur la France, pas seulement en zone de montagne, (Bretagne, Massif Central, Alpes, Jura, Pyrénées...) et concernés par la production hydro-électrique. La Compagnie Nationale du Rhône (CNR) utilise aussi les prévisions par analogie pour définir la gestion de ses aménagements au fil de l’eau le long du Rhône et optimiser sa production hydro-électrique.

La méthode des analogues a également été implémentée en Suisse pour une prévision opérationnelle dans le cadre du projet MINERVE (García Hernández *et al.* 2009). Le but du projet était de mettre en place un système de prévisions hydrométéorologiques pour le bassin du Rhône à l’amont du lac Léman afin d’optimiser la gestion des aménagements hydroélectriques dans le but d’écarter les pics de crue et d’éviter des inondations dans la vallée. Une des plus grandes incertitudes dans la chaîne de prévision hydrologique provient des précipitations. En particulier, il est souvent incertain de prévoir si les plus forts cumuls, liés à des situations de remontée de masses d’air humide depuis le Sud, parviennent à franchir les Alpes. La prévision fournie par les analogues a donc pour objectif d’apporter une prévision complémentaire aux prévisions des modèles numériques dans un but d’expertise lors de la prise de décision en cas de crise. C’est précisément dans le cadre du projet MINERVE que le logiciel AtmoSwing a été développé.

## **2 VERS UN SYSTEME OPERATIONNEL UNIQUE ETENDU A LA FRANCE METROPOLITAINE**

### **2.1 Principes et objectifs liés à l’extension de la couverture de la méthode des Analogues**

L’usage de la méthode des Analogues est actuellement limité aux SPC AN et LACI, principalement pour des raisons historiques. Des premiers tests se sont déroulés en 2014 et d’autres SPC se sont déclarés intéressés. L’une des difficultés est la multiplicité des outils utilisés au sein du réseau Vigicrues. Il est donc nécessaire de converger vers un outil national unique. Le choix de la plateforme nationale s’est porté sur la suite logicielle AtmoSwing en raison de considérations sur les données (compatibilité avec les réanalyses ERA5 et les prévisions GFS), les configurations (fonctionnement aux pas de temps journalier et 6 heures, compatibilité avec les paramétrages actuels des SPC AN et LACI) et informatiques (open-source, multiplateforme, usage via une interface et en ligne de commande). Des développements sont néanmoins indispensables pour adapter au mieux l’alimentation d’AtmoSwing et ses exports au système d’information du réseau Vigicrues.

En parallèle, la méthode des Analogues est étendue à la France métropolitaine en spécifiant certains éléments en vue de faciliter son intégration dans les systèmes opérationnels existants au sein du réseau Vigicrues. Les archives pluviométriques sont extraites des réanalyses SAFRAN, au pas de temps journalier (0h à 0h TU) ou au pas de temps 6h. Les réanalyses météorologiques utilisées sont les réanalyses européennes les plus récentes (ERA5) avec une maille 0,25 degré et un pas de temps de 6h. Une approche multi-modèle est appliquée afin de tirer profit des différents modèles de prévision numérique à disposition (ARPEGE, IFS et GFS). Les analogies sont restreintes à 4 champs géopotentiels entre 300 et 1000 hPa. Le gain obtenu en calage par certaines configurations actuelles intégrant une analogie sur l’humidité et l’eau précipitable est réduit en opérationnel à cause d’une plus faible prévisibilité de ces champs thermodynamiques. Seulement 30 configurations d’analogie seront optimisées, impliquant le besoin de réaliser un zonage pluviométrique sur la France métropolitaine à partir des 9403 mailles de SAFRAN.

## 2.2 Réanalyses et modèles météorologiques

ERA5 (Hersbach *et al.* 2020) est un jeu de réanalyses généré par le Centre européen pour les prévisions météorologiques à moyen terme (CEPMMT) qui a remplacé ERA-Interim. Il bénéficie de multiples améliorations apportées au modèle IFS (Integrated Forecasting System) en termes de représentation des processus et d'assimilation des données. Il est caractérisé par des résolutions temporelles (horaires) et spatiales ( $0,25^\circ$ ) élevées. ERA5 assimile beaucoup plus de données qu'ERA-Interim, telles que des données provenant de radars au sol et de nouveaux capteurs satellitaires, et intègre une meilleure comparaison des résultats du modèle avec les observations. En outre, il bénéficie indirectement des améliorations apportées aux observations historiques des données conventionnelles et satellitaires.

Le Centre d'Etude de la Neige de Météo France (CEN) a développé, au cours des années 1990, une analyse météorologique à méso-échelle en vue de fournir des analyses objectives de paramètres atmosphériques requis par la modélisation de l'évolution du manteau neigeux sur les principaux massifs des Alpes, des Pyrénées et de la Corse (Durand *et al.* 1993). Plus récemment, les procédures développées ont été généralisées au territoire français afin d'alimenter une chaîne de modélisation du cycle de l'eau à l'échelle régionale (Rousset *et al.* 2004) et également de produire des réanalyses météorologiques (Vidal *et al.* 2009).

Le modèle GFS (Global Forecast System) est un modèle de prévision numérique du temps développé par le National Center for Environmental Prediction (NCEP) des États-Unis. Il offre des prévisions météorologiques mondiales à court et à moyen terme. Le GFS utilise une grille de résolution spatiale relativement grossière, généralement autour de  $0,25$  degré (environ 25 km). Il produit des prévisions horaires pour les 5 premiers jours, puis à pas de temps de 3h jusqu'à une échéance de 16 jours. Une nouvelle prévision est émise toutes les 6h.

IFS (Integrated Forecasting System) est un modèle de prévision météorologique développé par le Centre européen pour les prévisions météorologiques à moyen terme (CEPMMT). Il fournit des prévisions à l'échelle mondiale sur des échéances allant jusqu'à deux semaines. L'IFS utilise une grille de résolution spatiale modérée, généralement autour de  $0,125$  degré (environ 13 km) pour les prévisions à l'échelle globale. Il produit des prévisions horaires jusqu'à 10 jours à l'avance.

ARPEGE (Action de Recherche Petite Echelle Grande Echelle) est un modèle numérique de prévision météorologique développé par Météo France. Il offre des prévisions à différentes échelles, allant de l'échelle locale à l'échelle globale avec une grille de résolution spatiale relativement fine, généralement autour de  $0,1$  degré (environ 10 km) pour les prévisions à l'échelle globale. ARPEGE produit des prévisions horaires jusqu'à 3 à 5 jours à l'avance. Une nouvelle prévision est émise toutes les 6h avec des échéances maximales différentes.

## 2.3 Regroupement de séries pluviométriques

Artelys (2021) a réalisé le regroupement de séries pluviométriques issues des réanalyses SAFRAN afin de répartir les mailles au sein de 30 groupes (cluster) selon deux algorithmes. *K-means* est un algorithme de classification non supervisée reposant sur la distance euclidienne pour caractériser la similarité des séries. Bernus *et al.* (2007) l'ont appliqué pour obtenir un zonage pluviométrique en France métropolitaine. *funFEM* est une extension de *K-means* permettant la modélisation de clusters de formes et d'orientations variables est obtenue par Mixture de Gaussiennes. Celui-ci s'est révélé être peu adapté à des séries peu auto-corrélées. Différentes expériences ont été menées afin d'étudier l'impact du seuillage des précipitations et l'application d'une normalisation par la précipitation décennale sur la forme et la taille des clusters obtenus. Il ressort notamment que les secteurs soumis à des précipitations fortes et très variables dans l'espace (zones de montagne, pourtour méditerranéen) sont mieux pris en compte sans la normalisation. En revanche, celle-ci produit des clusters de très grande taille sur les secteurs de plaine. Le filtrage des jours de fortes précipitations n'influe pas (ou très peu) sur l'homogénéité de ces clusters. La construction du regroupement final repose sur l'application de l'algorithme *K-means*, avec un filtre

valant la moitié de la précipitation décennale (un jour est retenu si ce seuil est satisfait par au moins 5 pixels), et la normalisation des précipitations seulement sur la moitié sud-est de la France métropolitaine (limite : Bayonne-Strasbourg). Les 30 clusters de ce regroupement final sont illustrées à la Figure 1. Des analyses statistiques basiques sont réalisées sur les séries pluviométriques dans le but de caractériser certaines propriétés. Les lames d’eau moyennes les plus importantes se situent à l’Est et au Sud-Ouest du territoire, avec un maximum dans les Alpes du Nord. Pour ce qui est des lames d’eau maximales, la région Méditerranée Ouest présente la valeur la plus élevée (148 mm/j). Une grande variabilité spatiale intra-cluster lors des événements extrêmes existent, avec un fort écart entre les valeurs minimales et maximales des pixels par région. Spatialement, les précipitations les plus extrêmes sont observées dans les régions du Sud du pays, en concordance avec les connaissances climatologiques bien établies. Lorsque nous considérons les valeurs au pixel (et non régionales), la Corse détient la valeur de précipitations journalière la plus forte du jeu de données (512 mm/j). L’intérêt principal de ces analyses porte sur les valeurs de fortes précipitations et la détermination des précipitations d’un temps de retour de 10 ans (P10). En effet, la variable cible utilisée pour le calage des méthodes d’analogie est normalisée par la pluie décennale. Une loi de Gumbel est ajustée aux données afin de déterminer les valeurs des précipitations décennales, présentées dans la Figure 1

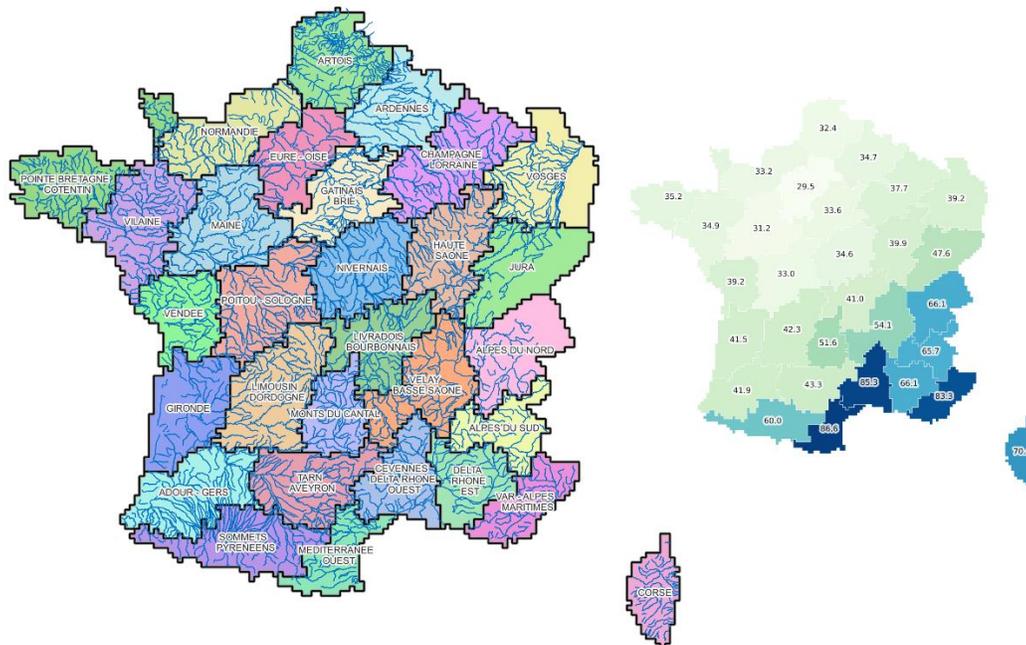


Figure 1 : Regroupement des séries pluviométriques SAFRAN en 30 clusters et carte des précipitations décennales

## 2.4 Optimisation de la méthode des Analogues sur la France

Les modèles d’analogie à optimiser sont constitués de deux niveaux d’analogie : (1) une présélection temporelle ayant pour but de sélectionner les situations candidates présentant des conditions saisonnières semblables à la situation, et consistant en une fenêtre temporelle de 120 jours centrée sur la date d’intérêt ; (2) une sélection des dates analogues sur la base de quatre champs géopotentiels. Ces champs sont comparés en termes de similitude de la circulation atmosphérique, par le critère d’analogie S1 (Teweless-Wobus). Les éléments à optimiser pour chaque champ géopotentiel sont : l’instant de validité (heure de la journée), le niveau de pression, et la fenêtre spatiale (propre à chaque niveau de pression). Le nombre de dates analogues à retenir au final est également à optimiser. Un modèle d’analogie doit

donc être optimisé pour chacune des 30 régions. L’optimisation globale et objective de ces caractéristiques est obtenue l’aide d’algorithmes génétiques (AGs) développés par Horton *et al.* (2017). Cette approche permet de se rapprocher d’un optimum global en ajustant tous les paramètres de la méthode de manière conjointe. Les niveaux de pression (options : 300, 400, 500, 600, 700, 850, 925, 1000 hPa), instants (0h à +30h), fenêtres spatiales (25°-65°N / 20°W-30°E), pondérations entre niveaux, et le nombre d’analogues (10 à 100) sont alors optimisés. Suivant les recommandations de Horton *et al.* (2017), plusieurs optimisations parallèles sont réalisées pour comparer les potentielles différentes solutions issues de l’optimisation, puisque les AGs permettent de s’approcher de l’optimum global, sans forcément garantir de l’atteindre.

Des analyses préliminaires ont montré qu’il n’est pas souhaitable d’optimiser la fenêtre temporelle pour la présélection ( $\pm 60$  jours) en raison de la grande variabilité résultant des optimisations et du faible gain en performance. De plus, l’utilisation des réanalyses à une résolution de 1° n’impacte pas significativement les résultats lorsque la variable considérée est le champ géopotentiel. Ceci s’explique par la forte autocorrélation spatiale des hauteurs du géopotentiel et a déjà été démontré par d’autres études (Horton & Brönnimann, 2019 ; Horton, 2021). Cette réduction de la résolution spatiale apporte par contre une réduction importante du temps de calculs par rapport à la résolution originale de ERA (0.25°).

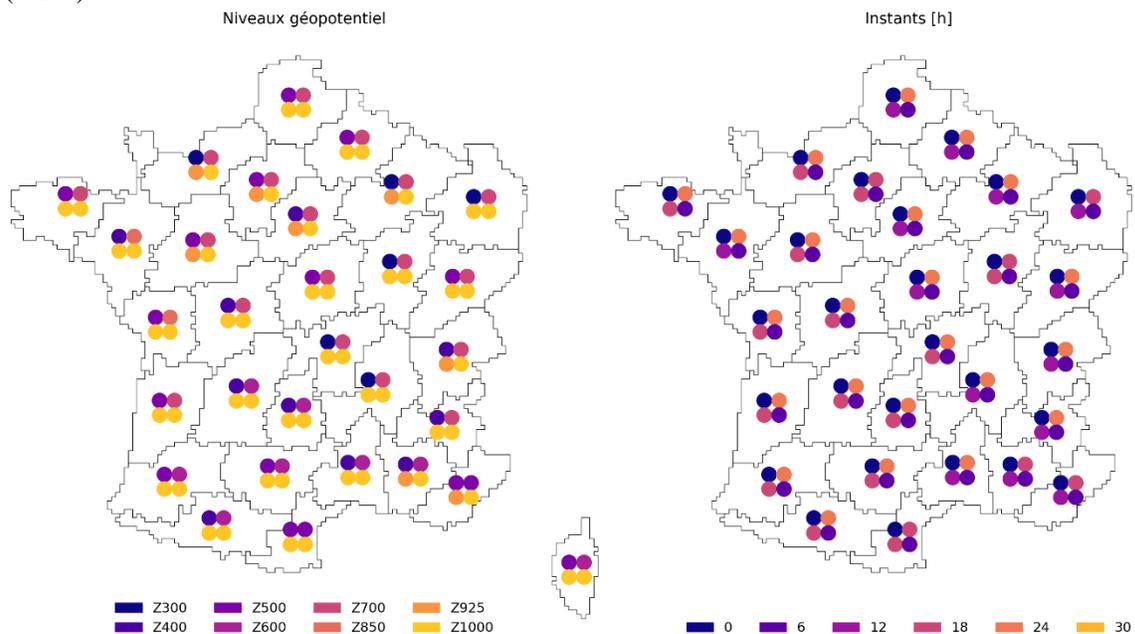


Figure 2: Niveaux de pression et instants des 4 analogies caractérisées sur les champs géopotentiels

Six optimisations ont été réalisées par région avec des configurations des AGs différentes. De manière générale, le niveau 1000 hPa a été sélectionné de loin le plus fréquemment, et plus que une fois par région, suivi du 500 hPa. Les autres niveaux sont présents dans un même ordre de grandeur. Le choix des instants change avec le niveau de pression, avec une prédominance du début de journée (0-6 h) pour les niveaux supérieurs (Z300-Z500), de la fin de journée (18-24 h) pour les niveaux intermédiaires (Z600-Z850), et de la partie principale de la journée (6-18h) pour les niveaux inférieurs (Z925-Z1000). L’étendue des fenêtres spatiales varie principalement dans son orientation méridionale en fonction du niveau, avec une étendue maximale au Sud pour Z500. Les centres des fenêtres spatiales présentent un gradient Est-Ouest pour les longitudes et Nord-Sud pour les latitudes. Les tailles des fenêtres spatiales sont du même ordre de grandeur entre niveaux en ce qui concerne leur étendue zonale, mais présentent une plus forte réduction pour les niveaux inférieurs en termes d’étendue méridionale. Les pondérations du critère d’analogie sont notablement plus élevées pour la variable Z1000 que pour les niveaux

supérieurs. Le nombre d’analogues optimal en prévision parfaite se situe dans une fourchette relativement réduite (21-36), avec des tendances qui diffèrent par région. Les valeurs des niveaux de pression et des instants, issues des optimisations sont présentés sous forme de carte des régions afin de visualiser leur structure spatiale.

## 2.5 Implémentation de la suite logicielle Atmoswing

AtmoSwing (Horton, 2019) est un logiciel libre (licence CDDL-1.0) qui implémente différentes variantes de la méthode des analogues de manière flexible, de sorte qu’elles peuvent être facilement configurées. Il est écrit en C++, orienté objet et multiplateforme. AtmoSwing est constitué de quatre outils : l’Optimizer pour établir la relation entre le prédicand et les prédicteurs, le Downscaler pour appliquer la méthode aux études d’impact climatique, le Forecaster pour effectuer des prévisions opérationnelles et le Viewer pour afficher les résultats de la prévision dans un environnement SIG interactif.

Différents développements ont été réalisés pour l’utilisation d’AtmoSwing dans le cadre du réseau Vigicrues. Un nouveau module Python (nommé `atmoswing-vigicrues`) a été développé pour implémenter certaines fonctionnalités, alors que d’autres ont été implémentées directement dans AtmoSwing. Le choix du développement d’un module Python est motivé par la plus grande facilité de maintenance par des tiers, offerte par un langage de plus haut niveau que le C++, et par l’implémentation de traitements spécifiques au réseau Vigicrues. Le module Python a été implémenté de manière à pouvoir facilement configurer le flux des processus, par exemple pour des utilisations différentes entre SPCs. Les actions traitées par le module et leurs options sont définies dans un fichier de configuration (yaml). Le module permet d’effectuer différentes tâches :

- Téléchargement par SFTP et organisation locale des prévisions numériques issues de Météo-France (Arpege) et du CEPMMT (IFS) ;
- Exports dans d’autres formats compatibles avec les outils utilisés au sein du réseau Vigicrues ;
- Diffusion par SFTP des prévisions émises par AtmoSwing et des fichiers exportés. Cela faciliterait la robustesse du système en permettant la récupération de résultats voulus par SPC et obtenus par un autre SPC ;
- Les modifications apportées à AtmoSwing pour le réseau Vigicrues et disponibles pour l’ensemble des utilisateurs, concernent le Forecaster et le Viewer. Pour le Forecaster, le changement principal est l’ajout du support des nouvelles sorties de modèles numériques, en particulier ARPEGE. Les changements ont été plus importants pour le Viewer, avec en premier lieu un meilleur support de l’affichage des prévisions à pas de temps de 6h et l’affichage des champs météorologiques permettant de comparer les champs de la situation cible (prévision des modèles numériques) à ceux des situations analogues (issues de ERA5).

Les nouvelles configurations optimisées sur la France offrent de nouvelles perspectives dans les systèmes opérationnels des SPC Alpes du Nord et Loire-Allier-Cher-Indre.

L’approche multi-scénario météorologique employée au SPC LACI (Faucard *et al.* 2023) est ainsi enrichie par l’introduction des prévisions Analogues sur la totalité de son territoire et aussi par l’usage des modèles IFS et ARPEGE en complément du modèle GFS. Ainsi, le SPC LACI est capable de couvrir les bassins versants concernés par une modélisation pluie-débit. Les scénarios de cumuls journaliers sont ensuite interpolés au pas de temps horaire selon le scénario 3 heures de Météo-France puis par répartition horaire uniforme. Ces scénarios pluviométriques sont enfin utilisés comme entrées dans le système de prévision hydrologique du SPC. Le prévisionniste dispose alors d’une information plus exhaustive lors de son expertise. L’implémentation d’AtmoSwing dans la boîte à outils du prévisionniste lui permet de visualiser les résultats de la méthode des Analogues, de comparer les différentes configurations et d’accéder facilement aux dates analogues, sollicitant ainsi son expérience vis-à-vis des crues majeures passées.

### 3 DISCUSSIONS ET CONCLUSIONS

La méthode des Analogues a été historiquement implémentée dans les zones de montagne. Les derniers développements étendent sa couverture à la France entière, dans une approche multi-modèle météorologique. Ceci s’accompagne de l’intégration de la suite logicielle AtmoSwing dans les systèmes opérationnels du réseau *Vigicrues*, améliorée à cette occasion. Les SPC AN et LACI disposent de nouveaux modèles d’analogie et de nouveaux scénarios météorologiques pour leur modélisation hydrologique. De nouvelles techniques d’apprentissage profond (*deep learning*) issues du domaine de l’intelligence artificielle sont appliquées dans le cadre de la présente étude. Les résultats, non détaillés ici, mettent en évidence l’intérêt d’optimiser par sous-échantillon, tout en maintenant l’évaluation sur l’échantillon complet. Le temps de calcul nécessaire à l’optimisation est nettement réduit et les performances sont équivalentes voire meilleures.

L’usage optimal de la méthode des Analogues nécessiterait plusieurs évolutions dont certaines sont déjà décrites dans la littérature : la prise en compte de l’incertitude de la modélisation météorologique par un nombre de dates analogues variables (Thévenot 2004), l’enrichissement des archives par une approche à pas de temps glissant, la prise en compte d’une saisonnalité par la température (air, mer), la complémentarité et le post-traitement de prévision d’ensemble afin d’obtenir des scénarios cohérents dans le temps et dans l’espace (Bellier *et al.* 2017). En effet, l’utilisation des Analogues permet d’obtenir facilement des informations sur la corrélation spatio-temporelle sans avoir à formuler d’hypothèses sur sa structure (Vannitsem *et al.*, 2021). Taillardat *et al.* 2016 proposent d’utiliser la méthode *Quantile Regression Forest* pour calibrer la prévision d’ensemble, méthode dont la philosophie est proche de celle de la méthode des Analogues dans la recherche des observations passées. Le post-traitement par analogie peut donc être vu comme une technique complémentaire à la construction d’ensemble. Une perspective intéressante serait de comparer la qualité des prévisions proposées par les Analogues sur la France avec les prévisions d’ensemble, à condition, évidemment, de disposer de l’historique de ces dernières.

La désagrégation temporelle des prévisions analogues (journalières ou au pas de temps 6h selon le cadre d’application) ne serait alors plus dépendante de l’unique scénario privilégié de Météo-France. Enfin, la mise en valeur des dates analogues dans les outils du prévisionniste, par le lien vers des données anciennes, vers des archives hydrologiques (dont l’Hydroportail), faciliterait son expertise réalisée en vue de diffuser des prévisions sur le site national *Vigicrues*.

### 4 REFERENCES

- Alessandrini, S., Sperati S., Delle Monache L. (2019) Improving the Analog Ensemble Wind Speed Forecasts for Rare Events. *Mon. Wea. Rev.*, 147, 2677–2692, <https://doi.org/10.1175/MWR-D-19-0006.1>
- Ben Daoud A., Sauquet E., Bontron G., Obled C., Lang M. (2016). Daily quantitative precipitation forecasts based on the analogue method: Improvements and application to a French large river basin. *Atmospheric Research*, 169 (A), pp.147-159. <https://doi.org/10.1016/j.atmosres.2015.09.015>
- Artelys (2021) Classification de séries temporelles de précipitations pour le zonage du territoire français métropolitain. Rapport d’étude Dreal Centre-Val de Loire, 63p.
- atmoswing-vigicrues : <https://github.com/atmoswing/atmoswing-vigicrues>

- Bellier J., Bontron G., Zin I. (2017). Using meteorological analogues for reordering postprocessed precipitation ensembles in hydrological forecasting. *Water Resources Research*, 53, 10, 085–10, 107. <https://doi.org/10.1002/2017WR021245>
- Bernus S., Pastuschak N., de Saint-Aubin C. (2007). Zonage pluviométrique de la France métropolitaine. Rapport de stage de l'ENM. Météo-France, 61p.
- Bontron G., 2004. *Prévision quantitative des précipitations: Adaptation probabiliste par recherche d'analogues. Utilisation des Réanalyses NCEP/NCAR et application aux précipitations du Sud-Est de la France.* Institut National Polytechnique de Grenoble.
- Caillouet L., Celie S., Vannier O., Bontron G., Legrand S. (2022). *Prévision hydrométéorologique opérationnelle du Rhône: sur le chemin d'une approche probabiliste "sans couture",* *La Houille Blanche*, 108 (1), <https://doi.org/10.1080/27678490.2022.2061312>
- Chardon J., T. Mathevet, M. Le Lay, J. Gailhard, Garçon R. (2012). Comparison of ensemble post-processing approaches, based on empirical and dynamical error modelisation of rainfall-runoff model forecasts, EGU General Assembly, European Geophysical Union.
- Delle Monache L., Eckel F. A., Rife D. L., Nagarajan B., Searight K. (2013) Probabilistic Weather Prediction with an Analog Ensemble. *Mon. Wea. Rev.*, 141, 3498–3516, <https://doi.org/10.1175/MWR-D-12-00281.1>
- Duband D. (1970). — *Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations.* Thèse de 3 e cycle à la Faculté des Sciences de Paris.
- Duband D. (1980). — *Dynamic selection of analogue flow patterns to produce quantitative precipitation forecasts.* WMO symposium on probabilistic and statistical methods in weather forecasting, Nice, France, September 1980, 487-492.
- Durand Y., Brun E., Merindol L., Guyomarc'h G., Lesaffre B., Martin E. (1993). A meteorological estimation of relevant parameters for snow models. *Annals of Glaciology* 18, pp. 65-71
- Gailhard J. (2014). *La prévision probabiliste des débits à EDF. Quelques rappels sur le pourquoi du comment...*, Rencontres Statistiques de Rochebrune, AgroParisTech/INRA.

- García Hernández J., Horton P., Tobin C., Boillat J., 2009. MINERVE 2010: Prévision hydrométéorologique et gestion des crues sur le Rhône alpin. *Wasser Energie Luft – Eau Energie Air*, 4, 297–302.
- Garçon R. (2014). La prévision à EDF. Utilisation des Analogues, CIC Analogues, Schapi.
- Guilbaud S., Obled C. (1998). — L’approche par analogues en prévision météorologique. *La Météorologie*, 8, 24, 21-35. <https://doi.org/10.4267/2042/47185>
- Hamill T. M., Scheuerer M., Bates G. T. (2015) Analog Probabilistic Precipitation Forecasts Using GEFS Reforecasts and Climatology-Calibrated Precipitation Analyses. *Mon. Wea. Rev.*, 143, 3300–3309, <https://doi.org/10.1175/MWR-D-15-0004.1>
- Hersbach H., Bell B., Berrisford P., Hirahara S., Horányi A., Muñoz-Sabater J., Nicolas J., Peubey C., Radu R., ... Thépaut, J. N., 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Horton P., 2019. AtmoSwing: Analog Technique Model for Statistical Weather forecastING and downscalING (v2.1.0). *Geoscientific Model Development*, 12(7), 2915–2940. <https://doi.org/10.5194/gmd-12-2915-2019>
- Horton P., 2021. Analogue methods and ERA5: Benefits and pitfalls. *International Journal of Climatology*, September 2021, 4078–4096. <https://doi.org/10.1002/joc.7484>
- Horton P., Jaboyedoff M., et Obled C., 2017. Global Optimization of an Analog Method by Means of Genetic Algorithms. *Monthly Weather Review*, 145(4), 1275–1294. <https://doi.org/10.1175/MWR-D-16-0093.1>
- Horton P., et Brönnimann S., 2019. Impact of global atmospheric reanalyses on statistical precipitation downscaling. *Climate Dynamics*, 52(9–10), 5189–5211. <https://doi.org/10.1007/s00382-018-4442-6>
- Hou Z., Li J., Wang L., Zhang Y., Liu T. (2022) Improving the forecast accuracy of ECMWF 2-m air temperature using a historical dataset, *Atmospheric Research*, 273. <https://doi.org/10.1016/j.atmosres.2022.106177>
- Jeworrek, J., West G., Stull R. (2023) Improved Analog Ensemble Formulation for 3-Hourly Precipitation Forecasts. *Wea. Forecasting*, 38, 1323–1339, <https://doi.org/10.1175/WAF-D-23-0018.1>

- Marty R., Zin I., Obled C., Bontron G., Djerboua A. (2012) Toward Real-Time Daily PQPF by an Analog Sorting Approach: Application to Flash-Flood Catchments. *J. Appl. Meteor. Climatol.*, 51, 505–520, <https://doi.org/10.1175/JAMC-D-11-011.1>.
- Odak Plenković, I., Schicker I., Dabernig M., Horvath K., Keresturi E. (2020) Analog-based post-processing of the ALADIN-LAEF ensemble predictions in complex terrain. *Q. J. R. Meteorol. Soc.* 146, 1842–1860. <https://doi.org/10.1002/qj.3769>
- Radanovics S., Vidal J.-P., Sauquet E., Ben Daoud A., et Bontron G., 2013. Optimising predictor domains for spatially coherent precipitation downscaling. *Hydrology and Earth System Sciences*, 17(10), 4189–4208. <https://doi.org/10.5194/hess-17-4189-2013>
- Rousset F., Habets F., Gomez E., Le Moigne P., Morel S., J. Noilhan J., Ledoux E. (2004). Hydrometeorological modeling of the Seine basin using the SAFRAN-ISBA-MODCOU system. *Journal of Geophysical Research-Atmospheres* 109(D14), pp. 1-20.
- Taillardat M., Mestre O., Zamo M. et Naveau P. (2016). Calibrated Ensemble Forecasts Using Quantile Regression Forests and Ensemble Model Output Statistics. *Mon. Wea. Rev.*, 144, 2375–2393, <https://doi.org/10.1175/MWR-D-15-0260.1>
- Thévenot, N. (2004). *Prévision quantitative des précipitations par une méthode d’analogie: Utilisation de la prévision d’ensemble du CEPMMT*. LTHE, Institut National Polytechnique de Grenoble, Mémoire de master, 58 pp.
- Vannitsem, S., et co-auteurs (2021). Statistical Postprocessing for Weather Forecasts: Review, Challenges, and Avenues in a Big Data World. *Bull. Amer. Meteor. Soc.*, 102, E681–E699, <https://doi.org/10.1175/BAMS-D-19-0308.1>
- Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., Soubeyroux, J.-M. (2010) A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology* (30) , 1627-1644
- Zalachori I., M.-H. Ramos, R. Garçon, T. Mathevet et J. Gailhard (2012). Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies, *Advances in Science and Research*, 8, 135-141.