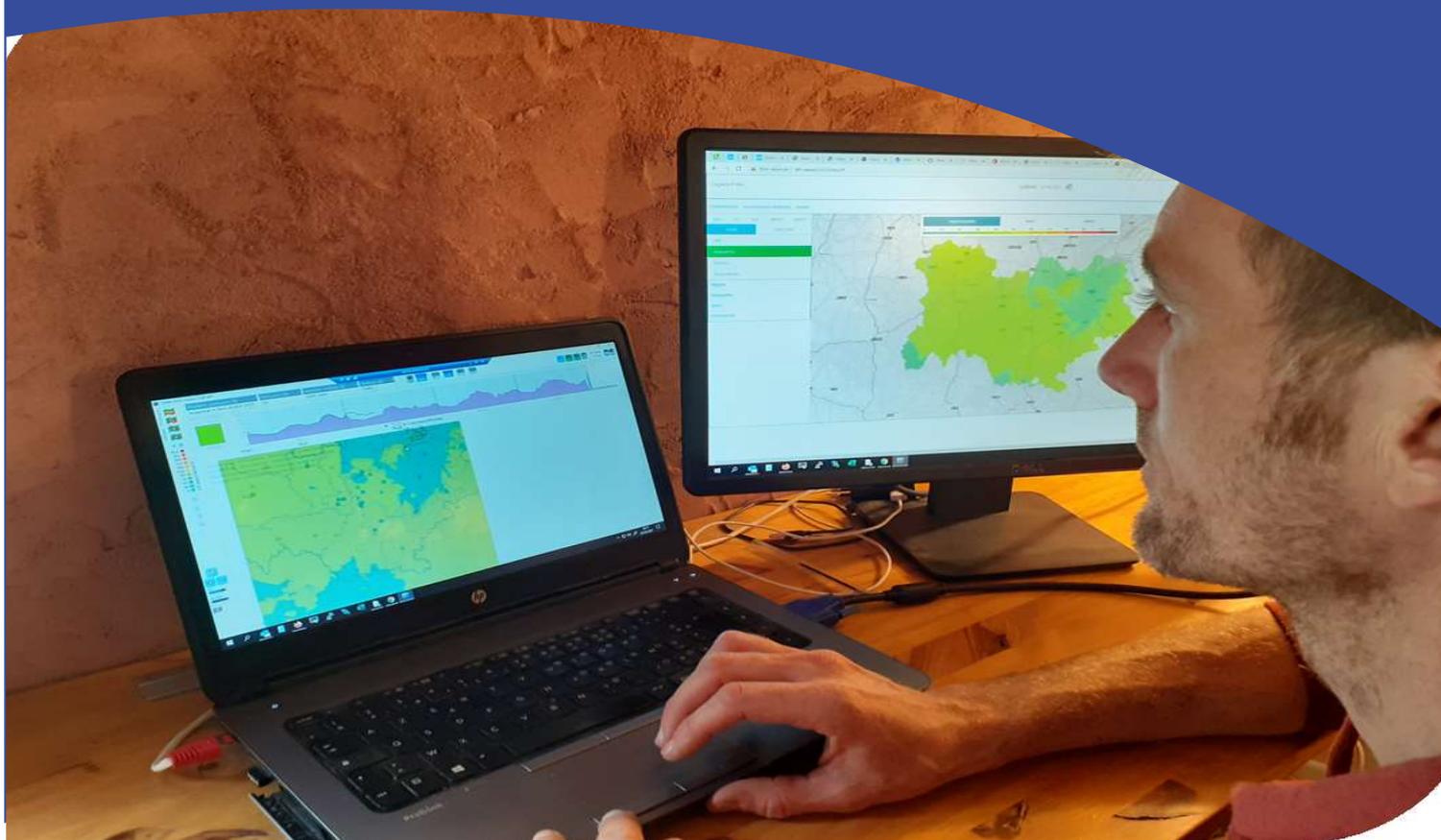


Evaluation des performances du système régional de prévision de la qualité de l'air

Bilan statistique global et focus sur les épisodes de pollution

Année 2021



Auteur : Unité Prévision et Vigilance

Diffusion : Juin 2022

Siège social :
3 allée des Sorbiers 69500 BRON
Tel. 09 72 26 48 90
contact@atmo-aura.fr

Conditions de diffusion

Atmo Auvergne-Rhône-Alpes est une association de type « loi 1901 » agréée par le Ministère de l'Ecologie, du Développement Durable et de l'Energie (décret 98-361 du 6 mai 1998) au même titre que l'ensemble des structures chargées de la surveillance de la qualité de l'air, formant le réseau national ATMO.

Ses missions s'exercent dans le cadre de la loi sur l'air du 30 décembre 1996. La structure agit dans l'esprit de la charte de l'environnement de 2004 adossée à la constitution de l'Etat français et de l'article L.220-1 du Code de l'environnement. Elle gère un observatoire environnemental relatif à l'air et à la pollution atmosphérique au sens de l'article L.220-2 du Code de l'Environnement.

Atmo Auvergne-Rhône-Alpes communique publiquement sur les informations issues de ses différents travaux et garantit la transparence de l'information sur le résultat de ses travaux.

A ce titre, les rapports d'études sont librement disponibles sur le site www.atmo-auvergnerhonealpes.fr.

Les données contenues dans ce document restent la propriété intellectuelle d'Atmo Auvergne-Rhône-Alpes.

Toute utilisation partielle ou totale de ce document (extrait de texte, graphiques, tableaux, ...) doit faire référence à l'observatoire dans les termes suivants : © Atmo Auvergne-Rhône-Alpes (2022) Evaluation des performances du système régional de prévision de la qualité de l'air.

Les données ne sont pas rediffusées en cas de modification ultérieure.

Par ailleurs, Atmo Auvergne-Rhône-Alpes n'est en aucune façon responsable des interprétations et travaux intellectuels, publications diverses résultant de ses travaux et pour lesquels aucun accord préalable n'aurait été donné.

En cas de remarques sur les informations ou leurs conditions d'utilisation, prenez contact avec Atmo Auvergne-Rhône-Alpes :

- par mail : contact@atmo-aura.fr
- par téléphone : 09 72 26 48 90

Le dispositif de gestion des épisodes de pollution atmosphérique est un dispositif d'urgence, qui a pour but de limiter l'exposition des populations à la pollution de l'air si le dépassement d'un seuil réglementaire de concentration est prévu. Il permet d'informer la population, de délivrer des recommandations sanitaires et comportementales et de mettre en place des actions de réduction des émissions sur les différentes sources concernées (trafic routier, industries, secteurs agricole et domestique...). Initialement, ces dispositifs étaient conçus pour contribuer à respecter les valeurs réglementaires annuelles. Mais les attentes de l'Etat, des collectivités et du grand public vis-à-vis des missions concernant la qualité de l'air ont rapidement évolué ces dernières années. En effet, la demande sociale s'accroît et les citoyens attendent une information de plus en plus individualisée, prévisionnelle, en lien avec leur mode de vie, leurs activités. Ainsi, dans les dernières décennies, plusieurs enjeux environnementaux, financiers et économiques, politiques et sanitaires sont associés explicitement à la prévision quotidienne de la qualité de l'air.

Il est par conséquent indispensable de disposer d'outils adaptés à une forte exigence d'expertise et à la diffusion d'une prévision quotidienne fiable et de bonne qualité. Cela nécessite une évaluation correcte et permanente des systèmes de prévision de qualité de l'air. Dans ce cadre, des nouveaux textes réglementaires sont apparus, notamment l'article 13 de l'arrêté du 19 avril 2017 abrogé par l'article 13 de l'arrêté du 16 avril 2021 qui insiste sur la nécessité d'étudier la performance du système de prévision quotidienne et d'harmoniser les critères d'évaluation des prévisions de qualité de l'air à l'échelle nationale. Dans ce contexte et à partir de 2017, chaque AASQA évalue son système de prévision en étudiant ses performances statistiques selon des critères préconisés par un référentiel technique national. Cette évaluation fait l'objet d'un rapport annuel adressé aux autorités au plus tard le 30 juin de l'année suivante. Comme il n'existe pas de méthode d'évaluation adaptée aux modèles de prévision de la qualité de l'air, un groupe de travail national piloté par l'INERIS a été mis en place pour en définir une.

Ce rapport étudie la performance de la chaîne de prévision en utilisant des indicateurs statistiques permettant de comparer les résultats de modélisation aux observations enregistrées par le réseau de mesures permanents d'Atmo Auvergne-Rhône-Alpes. L'évaluation des performances du système de prévision est réalisée avec les résultats de la modélisation avant et après l'expertise humaine de prévisionniste. Le calcul des scores est effectué selon les recommandations du référentiel technique national pour l'évaluation des performances des prévisions de la qualité de l'air.

Les résultats du système de prévision sont considérés comme assez satisfaisants voire bons pour les concentrations journalières des PM₁₀ et maximales pour l'O₃. Les valeurs des indicateurs sont dans la plage acceptable et même proches des valeurs « idéales » proposées par le référentiel technique national.

L'analyse effectuée montre que les résultats bruts (en sortie de chaîne de modélisation) ont une tendance à sous-estimer les épisodes de pollution. Ce comportement est plus contrasté (marquant) pour les particules fines. L'apport de l'adaptation statistique par krigeage est réel et permet d'améliorer nettement les performances statistiques des résultats. Les résultats avant et après ajustement par les prévisionnistes sont globalement similaires. Cependant, on constate que l'expertise humaine apporte une amélioration dans les statistiques de prévision des dépassements

de seuils pour le jour même J0. Elle permet par ailleurs de fiabiliser la chaîne tout au long du processus, par la détection d'anomalies.

En 2021, 25 jours ont connu une vigilance active en Auvergne Rhône-Alpes, cela en fait l'année avec le plus faible total de cette dernière décennie (53 jours en moyenne sur 10 ans, 34 jours en moyenne sur 5 ans). Le bassin lyonnais Nord-Isère et la vallée de l'Arve restent les territoires sur lesquels le plus grand nombre de vigilances atmosphériques ont été activées cette année (17 jours). Le taux de détection et le ratio de fausses alarmes ont été respectivement de 60% et 36% pour J0. A noter que 100% des journées ayant vu un dépassement des seuils réglementaires ont concerné les particules PM₁₀. Pour autant, même si les seuils réglementaires pour l'ozone n'ont pas été dépassés, des risques de dépassement de ces seuils ont été détectés pour le lendemain ou le surlendemain, entraînant la mise en place de vigilance sur les bassins d'air concernés pour les journées concernées.

En 2022, plusieurs travaux visant à améliorer la chaîne de prévision seront menés. Ils porteront notamment sur le modèle lui-même (nouvelle version de CHIMERE), les données de bord (conditions aux limites de notre région) et les données d'émissions.



Sommaire

1. Introduction.....	6
2. Présentation de la plateforme de prévision	7
3. Evaluation de la plateforme de prévision régionale de la qualité de l'air	8
3.1. Mesures utilisées pour la validation statistique	8
3.2. Prévision ozone (O ₃).....	9
3.2.1. Scores classiques (biais, RMSE et corrélation).....	9
3.2.1.1. Par rapport aux objectifs de qualité	9
3.2.1.2. Par typologie	13
3.2.1.3. Par échéance de prévision	14
3.2.1.4. Par gamme de concentration	14
3.2.1.5. Sur plusieurs années	16
3.2.2. Performances sur les seuils réglementaires	16
3.2.2.1. Scores aux stations.....	16
3.2.2.1.1. Séries temporelles	17
3.2.2.1.2. Eléments de la table de contingence	18
3.2.2.2. Scores en lien avec les critères réglementaires de déclenchement d'épisode de pollution	19
3.3. Prévision particules fines (PM ₁₀).....	20
3.3.1. Scores classiques (Biais, RMSE, Corrélation)	20
3.3.1.1. Par rapport aux objectifs de qualité	20
3.3.1.2. Par typologie	23
3.3.1.3. Par échéance de prévision	24
3.3.1.4. Par gamme de concentration	24
3.3.1.5. Sur plusieurs années	26
3.3.2. Performances sur les seuils réglementaires	26
3.3.2.1. Scores aux stations.....	26
3.3.2.1.1. Séries temporelles	26
3.3.2.1.2. Eléments de la table de contingence	27
3.3.2.2. Scores en lien avec les critères réglementaires de déclenchement d'un épisode de pollution	28
4. Conclusion	29

1. Introduction

Le système de prévision d'Atmo Auvergne-Rhône-Alpes a pour objectif de mettre à disposition quotidiennement des cartes de concentrations pour les polluants réglementés O₃, NO₂ et PM₁₀ pour le jour même, le lendemain et le surlendemain (J+2). L'ensemble est disponible en ligne sur le site web d'Atmo Auvergne-Rhône-Alpes (<https://www.atmo-auvergnerhonealpes.fr/>) de manière intégrée selon la méthode de calcul de l'indice Atmo, l'application AirToGo et différentes API.

La chaîne de prévision est composée de plusieurs maillons : météorologie, chimie/transport, statistique, expertise humaine... Sa fiabilité est fortement liée à la qualité des paramètres d'entrée, notamment les caractéristiques du terme source (émissions) qui sont encore difficilement quantifiables par des mesures directes. En effet, les débits d'émissions de polluants sont en général calculés par des méthodes génériques en utilisant des facteurs d'émissions qui ont une grande incertitude ou des cadastres d'émissions qui restent perfectibles. De plus, le système n'intègre pas d'une manière explicite la topographie dans le calcul des écoulements, les reliefs sont pris en compte uniquement pour établir les profils de vitesse de vent. Il en résulte certaines difficultés à reproduire les écoulements et les processus thermiques dans les vallées, dont la topographie est particulièrement marquée sur l'est de la région, avec les Alpes.

Atmo Auvergne-Rhône-Alpes évalue la performance de sa plateforme de prévision en se référant à une méthode établie dans le cadre d'un groupe de travail national¹ portant sur la prévision de la qualité de l'air. Cette méthode est décrite explicitement dans un référentiel technique national élaboré en réponse à l'article 13 de l'arrêté du 16 avril 2021 relatif au dispositif national de surveillance de la qualité de l'air ambiant.

¹Groupe de travail piloté par l'INERIS et constitué de représentants de toutes les associations agréées de surveillance de la qualité de l'air françaises, dont les travaux ont débuté en juin 2018.

2. Présentation de la plateforme de prévision

La chaîne de prévision opérationnelle est un enchaînement de modèles imbriqués permettant d'obtenir des cartes de champs de concentrations à des échelles temporelles et spatiales définies sur un domaine d'étude régional. Ce dernier englobe l'ensemble de la région Auvergne-Rhône-Alpes, dont l'emprise spatiale est de 447x342km² et la résolution horizontale de 1x1 km². La Figure 1 illustre le synoptique général de la chaîne de modélisation d'Atmo AURA qui imbrique plusieurs étapes :

- Le calcul de l'évolution des champs météorologiques régionaux, qui utilise le modèle numérique WRF (National Center for Atmospheric Research). Les principales données d'entrée de ce modèle sont les prévisions à l'échelle globale (GFS), la topographie et l'occupation du sol. Le calcul des grandeurs météorologiques (vent, température, humidité...) par le modèle WRF à l'échelle régionale prend en compte d'une manière implicite les effets induits par le relief.
- La préparation du reste des données d'entrée nécessaires pour le modèle de dispersion utilisé par notre chaîne de modélisation. Ces données sont : les émissions de polluants, l'occupation des sols, les conditions aux limites et les conditions initiales, Les émissions, à savoir la répartition maillée des flux de polluants de tous les secteurs d'activité, sont calculées avec un module développé en interne. En dehors de la région Auvergne-Rhône-Alpes, les émissions anthropogéniques sont issues des données EMEP pour l'année de référence 2010.
- La dispersion et les transformations chimiques des polluants sont réalisées avec le modèle CHIMERE. Il s'agit d'un modèle eulérien 3D de dispersion atmosphérique à méso-échelle, développé par l'IPSL (Institut Pierre-Simon Laplace) en collaboration avec l'INERIS. CHIMERE calcule de manière déterministe les concentrations des espèces gazeuses et particulaires. Il permet de prévoir la pollution atmosphérique en prenant en compte l'évolution de nombreux processus physico-chimiques dans l'atmosphère.
- Le raffinement de la résolution spatiale via un module de descente d'échelle de 3km à 1km, fondé sur la méthode basique d'interpolation bilinéaire.
- Un calcul statistique de prévision des concentrations aux stations de mesures permanentes de la qualité de l'air. Cette étape consiste à prévoir les concentrations en chaque site de mesure par une méthode statistique (forêts aléatoires). Ces prévisions aux stations sont utilisées dans la prochaine étape pour corriger les cartes de la prévision brute (résultats non corrigés issus directement du modèle déterministe CHIMERE).
- Le couplage des concentrations calculées aux stations et de la carte de prévision brute par une correction géostatistique (dite le Krigeage).

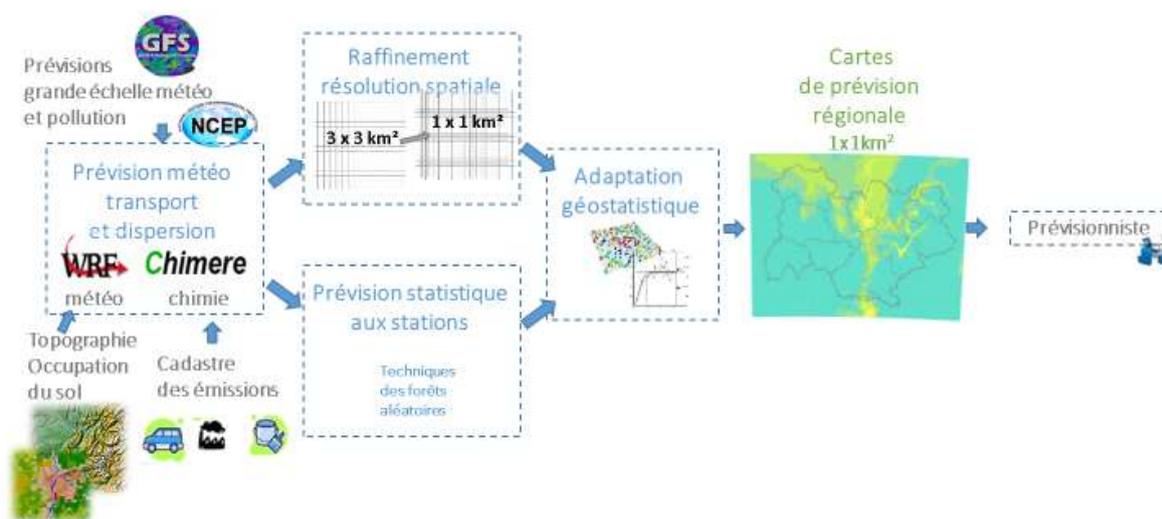


Figure 1 : Chaîne de prévision régionale d'Atmo Auvergne-Rhône-Alpes

Pour exécuter l'ensemble de ces processus, nous utilisons plusieurs scripts codés en Shell et en R pour manipuler et convertir les données d'entrées et post-traitées.

3. Evaluation de la plateforme de prévision régionale de la qualité de l'air

La performance du système de prévision est évaluée en comparant les concentrations prédites C_p de PM_{10} et d'ozone (O_3) aux concentrations mesurées (observées) C_m . Il convient de rappeler que les résultats de modélisation corrigés statistiquement peuvent être ajustés manuellement par les prévisionnistes. Ces ajustements sont réalisés dans un outil cartographique développé en interne par Atmo Auvergne-Rhône-Alpes, dénommé Carine. Ils sont réalisés sur la base de l'expertise du prévisionniste, en s'appuyant notamment sur la comparaison modèle/mesure des jours précédents, à la fois sur les modèles de prévision de météorologie et de qualité de l'air. La convergence ou la divergence de différents modèles sont également des paramètres pris en compte. Les ajustements peuvent dans certains cas entraîner des dépassements de seuils tels que définis dans les dispositifs de gestion des épisodes de pollution. Cette méthode permet d'assurer une cohérence entre les cartes de qualité de l'air diffusées au public et les dispositifs d'information ou d'alerte. Les déterminants pour chaque prévisionniste lors de la mise en œuvre d'un ajustement sont tracés dans les documents qualifiés. Un travail a été engagé en interne pour harmoniser cette pratique de correction en fixant les critères et les règles d'ajustement des cartes.

Dans ce rapport, les scores statistiques calculés sont présentés pour les deux approches, avec et sans expertise humaine sachant que l'ajustement des cartes n'est pas systématique. Ces scores sont calculés en utilisant le package Python « Evaltools² » codéveloppé par l'INERIS et Météo France.

La comparaison modèle-mesure, évaluée par différents indicateurs statistiques, est réalisée pour chaque mesure. Pour évaluer la qualité globale de la simulation, ces indicateurs statistiques sont moyennés sur l'ensemble des capteurs de la région ou par zone, selon les bassins d'air utilisés pour la gestion des épisodes de pollution (Annexe 1).

Les statistiques sont estimées à partir d'une donnée chaque jour pour tous les polluants : la moyenne journalière pour les particules PM_{10} , et la valeur maximale horaire du jour pour l'ozone (O_3).

La performance du système n'est pas évaluée pour le dioxyde de soufre (SO_2) et le dioxyde d'azote (NO_2), bien que faisant partie des polluants pris en compte pour les épisodes de pollution. En effet, ces deux polluants ne font pas l'objet d'exigence dans le référentiel technique national pour la prévision de la qualité de l'air.

Seules les stations de typologie urbaine, périurbaine et rurale sont utilisées pour calculer les paramètres statistiques.

3.1. Mesures utilisées pour la validation statistique

Les mesures utilisées pour évaluer les résultats de la chaîne de modélisation sont issues du réseau de stations fixes géré par Atmo Auvergne-Rhône-Alpes, représentant près de 90 points. Ces stations ont été classées par typologie (urbaine, rurale, périurbaine, industrielle, trafic). Les stations de typologie industrielle et trafic ont été écartées de cette analyse statistique car elles sont largement influencées par les émissions locales et de proximité, alors que le modèle est dit de « fond ». La répartition spatiale des stations et leurs mesures sont accessibles sur le site internet (<https://www.atmo-auvergnerhonealpes.fr/dataviz/mesures-aux-stations>) (Figure 2).

² <https://opensource.umr-cnrm.fr/projects/evaltools>

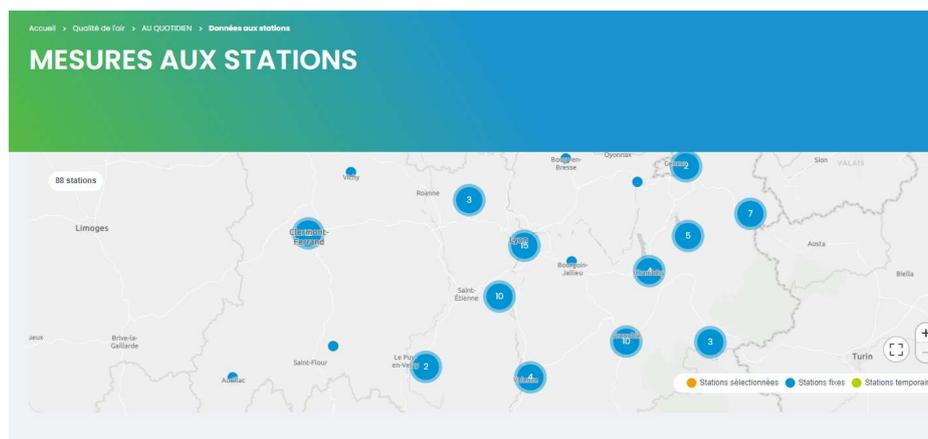


Figure 2 : Accès aux données de mesure sur le site web d'Atmo Auvergne-Rhône-Alpes

3.2. Prévision ozone (O₃)

L'approche utilisée pour évaluer la qualité des prévisions consiste à comparer les concentrations estimées à celles mesurées aux stations. La comparaison modèle-mesure, évaluée par différents indicateurs statistiques, est réalisée pour chaque mesure. Pour évaluer la qualité globale de la simulation, ces indicateurs statistiques sont moyennés sur l'ensemble des capteurs de la région ou par zone, selon les bassins d'air utilisés pour la gestion des épisodes de pollution (les zones AIP dans ce document).

Les statistiques sont estimées à partir de valeurs journalières pour tous les polluants : la moyenne journalière pour les particules PM₁₀, et la valeur maximale horaire du jour pour l'ozone (O₃).

Seules les stations de typologie urbaine, périurbaine et rurale sont utilisées pour calculer les paramètres statistiques. A noter que le calcul statistique a été effectué à la fois à partir des données expertisées (résultats ajustés manuellement), des données non expertisées par la prévisionniste (résultats krigés des modèles bruts) et les résultats bruts du modèle déterministe.

3.2.1. Scores classiques (biais, RMSE et corrélation)

Afin de quantifier les résultats des prévisions, il est nécessaire d'utiliser des indicateurs statistiques permettant d'étudier les écarts entre le modèle et les mesures. La bibliographie relate de nombreux travaux ayant abordé cette question récurrente de la quantification de la performance des prévisions et ces travaux ont conduit à retenir certains paramètres statistiques parmi une multitude de paramètres existants : le biais, l'erreur relative, l'erreur quadratique moyenne, la fraction dans un facteur de 2, le coefficient de performance, le coefficient de corrélation...

Sur la base de ces travaux bibliographiques, et conformément à « l'état de l'art », trois indicateurs statistiques complémentaires ont été retenus pour évaluer la performance des prévisions :

- **Le biais moyen (Biais)** : permet d'estimer l'erreur globale du modèle en quantifiant l'écart entre les moyennes des valeurs observées et modélisées. La valeur optimale de ce critère est 0.
- **La racine carrée de l'erreur quadratique des estimations (RMSE)** : peut être assimilée à une erreur locale. La valeur optimale pour ce paramètre est également 0.
- **Le coefficient de corrélation (R)** : paramètre statistique qui permet d'évaluer la corrélation entre les variations de deux variables. La valeur du coefficient de corrélation varie entre -1 et 1. La valeur optimale correspondant à des variables parfaitement corrélées est 1.

3.2.1.1. Par rapport aux objectifs de qualité

Puisque les modèles de dispersion atmosphérique ne peuvent pas prédire parfaitement la réalité (ou les concentrations mesurées par les capteurs), le groupe de travail national a proposé des fourchettes de valeurs associées aux indicateurs Biais, RMSE et la corrélation pour évaluer ces modèles : leurs résultats sont considérés comme acceptables lorsque les critères sont vérifiés (Tableau 1). A noter qu'il n'existe pas à notre connaissance de critères de qualité pour les indices R, Biais et RMSE.

Le Tableau 1 illustre les objectifs de qualité définis pour l’ozone. Ces valeurs doivent s’appliquer aux performances des prévisions en région et sur les zones AIP aux échéances J0 et J+1. Les prévisions de la qualité de l’air élaborées par Atmo Auvergne-Rhône-Alpes doivent donc satisfaire à ces critères.

Biais ($\mu\text{g.m}^{-3}$)	RMSE ($\mu\text{g.m}^{-3}$)	Corrélation
-10/+10	25	0,7

Tableau 1 : Objectifs de qualité définis pour l’O₃ s’appliquant aux performances des prévisions en région et sur les zones AIP aux échéances J0 et J+1.

Dans ce paragraphe, le focus est mis sur la performance de la chaîne de prévision par bassin d’air au sens des dispositifs de gestion des épisodes de pollution. Les résultats statistiques par zone d’activation de dispositif sont présentés sur les Figure 3a et Figure 3b qui illustrent les performances statistiques calculées pour les deux échéances J0 et J+1.

Ces figures montrent que les résultats de prévision non ajustés (Adapstat, en vert) sont globalement assez similaires à ceux expertisés par le prévisionniste (Outil Carine, en bleu). On remarque que l’intervention des prévisionnistes génère statistiquement une légère amélioration des prévisions d’ozone pour les différentes échéances. Le calcul des scores illustre des meilleures performances du modèle ajusté statistiquement pour les deux échéances de prévision. On observe également que les biais sont souvent plus élevés pour le modèle déterministe Chimère (en rouge). La Figure 3b montre que les deux modèles « Adapstat » et « Carine » ne respectent pas les objectifs de qualité pour le Biais dans l’Est-Drôme. En effet, le biais calculé pour ces deux modèles dépasse $10 \mu\text{g.m}^{-3}$.

Globalement, les résultats de la chaîne de prévision sont acceptables sur l’ensemble des zones AIP. Les Figure 3a et Figure 3b montrent que les objectifs de qualité sont largement respectés par le modèle ajusté à l’échéance J0. Ce comportement est moins confirmé pour l’échéance J+1 et plus particulièrement pour la zone Est-Drôme où le biais calculé dépasse légèrement l’objectif de qualité pour ce paramètre statistique. A noter également que les valeurs des coefficients de corrélations sont légèrement plus faibles pour J+1.

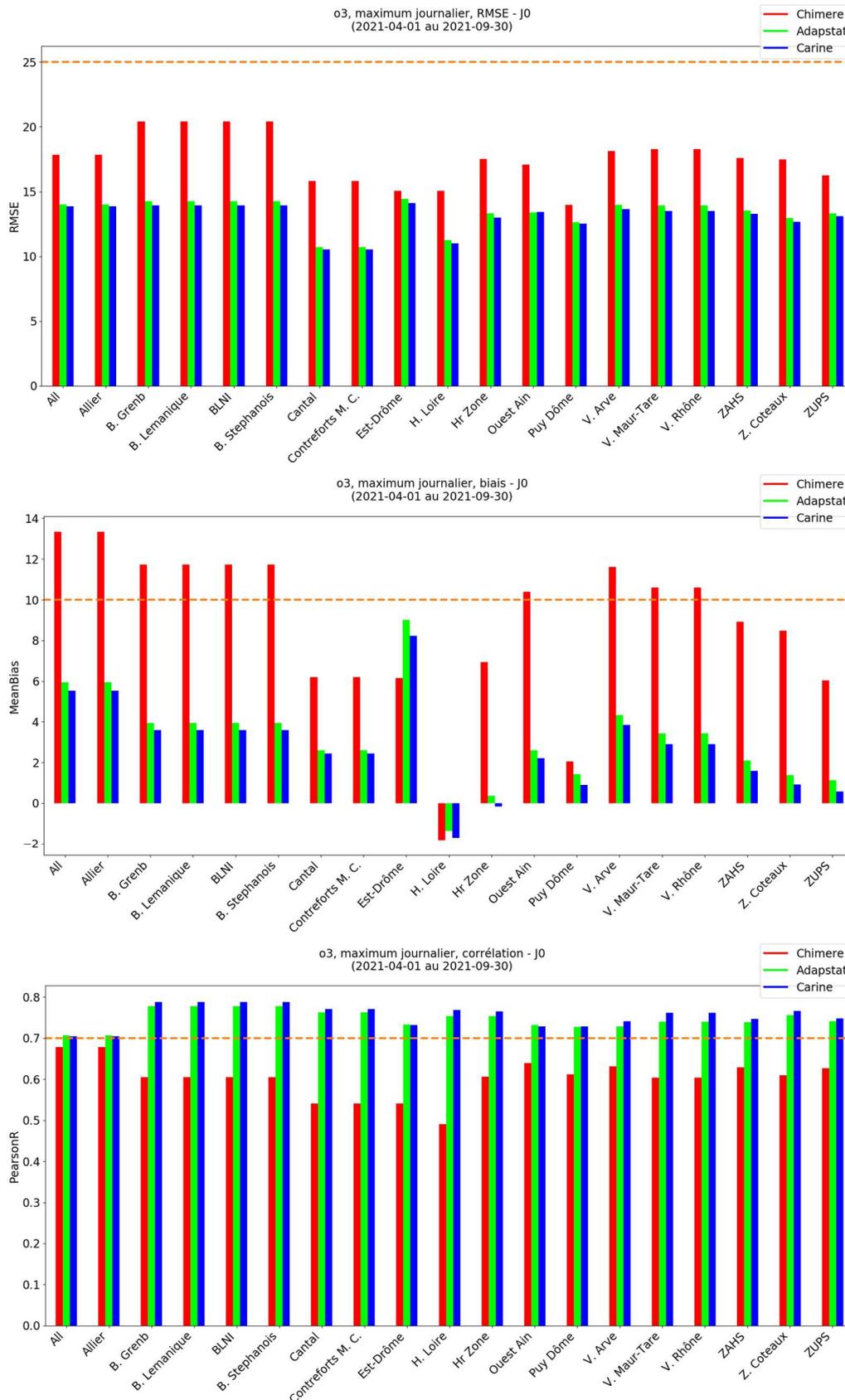


Figure 3a : Ozone - Scores de RMSE, biais et corrélation par zone AIP des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) d'avril à septembre 2021 sur l'échéance J0. La ligne en pointillé orange indique l'objectif de qualité.

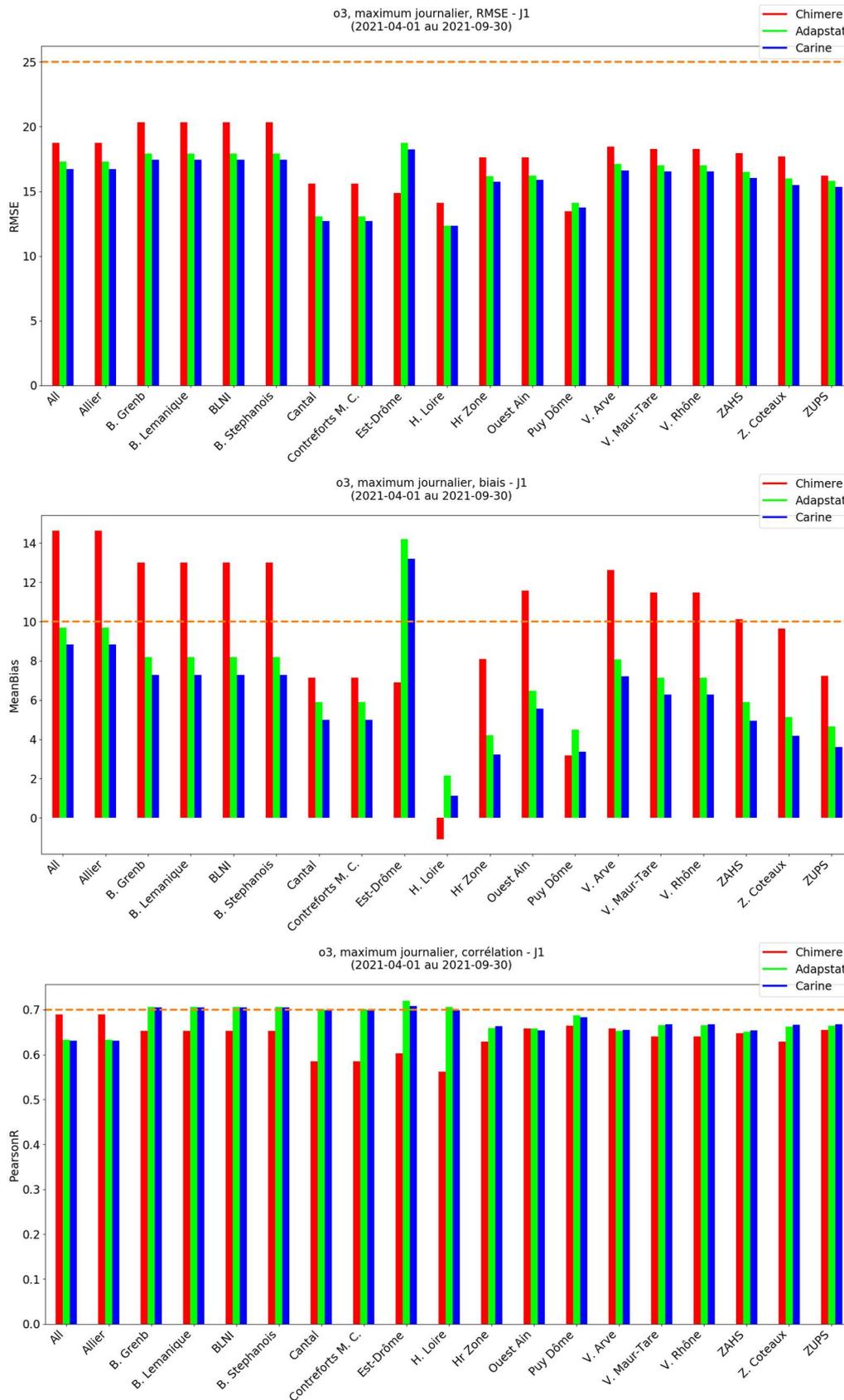


Figure 3b : Ozone - Scores de RMSE, biais et corrélation par zone AIP des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) d'avril à septembre 2021 sur l'échéance J+1. La ligne en pointillé orange indique l'objectif de qualité.

3.2.1.2. Par typologie

Les Figure 4a et Figure 4b illustrent les performances de la chaîne de prévision en fonction de la typologie des stations de mesures. Les meilleures performances statistiques sont globalement associées aux stations rurales. Cependant, les plus mauvais résultats sont souvent associés aux sites de typologie urbaines pour lesquels il y a une forte influence de sources locales et donc une très importante variabilité spatio-temporelle des émissions autour des sites de mesures. Les bonnes performances observées sur les sites ruraux peuvent s'expliquer par l'aptitude de la chaîne de modélisation à reproduire les faibles concentrations. Globalement, les indicateurs statistiques calculés dans les Figure 4a et Figure 4b indiquent un très bon accord entre les prévisions et les mesures. En effet, les indicateurs statistiques des prévisions ne dépassent pas la fourchette des valeurs considérées comme acceptables.

La comparaison des prévisions avant et après l'expertise humaine montrent encore une fois un résultat quasi-identique et très similaire pour l'O₃. Néanmoins, l'ajustement des cartes par les prévisionnistes peut **améliorer légèrement les prévisions de l'ozone quelle que soit la typologie des stations.**

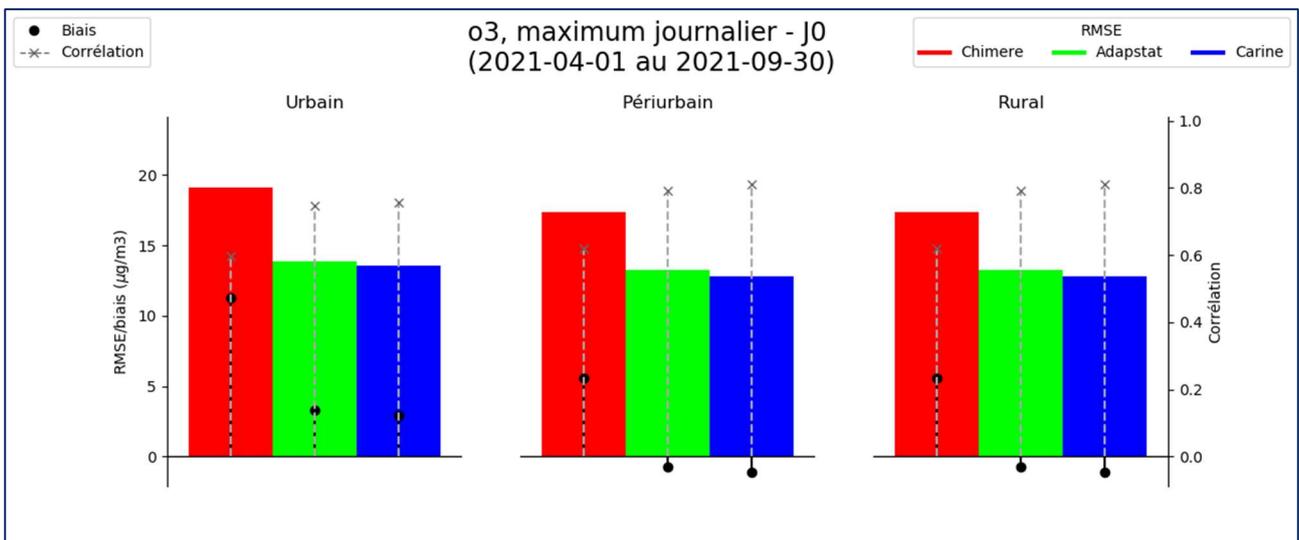


Figure 4a : Ozone - Synthèse des scores de RMSE, biais et corrélation en fonction de la typologie de station (urbaines, périurbaines, rurales) des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) d'avril à septembre 2021 sur l'échéance J0.

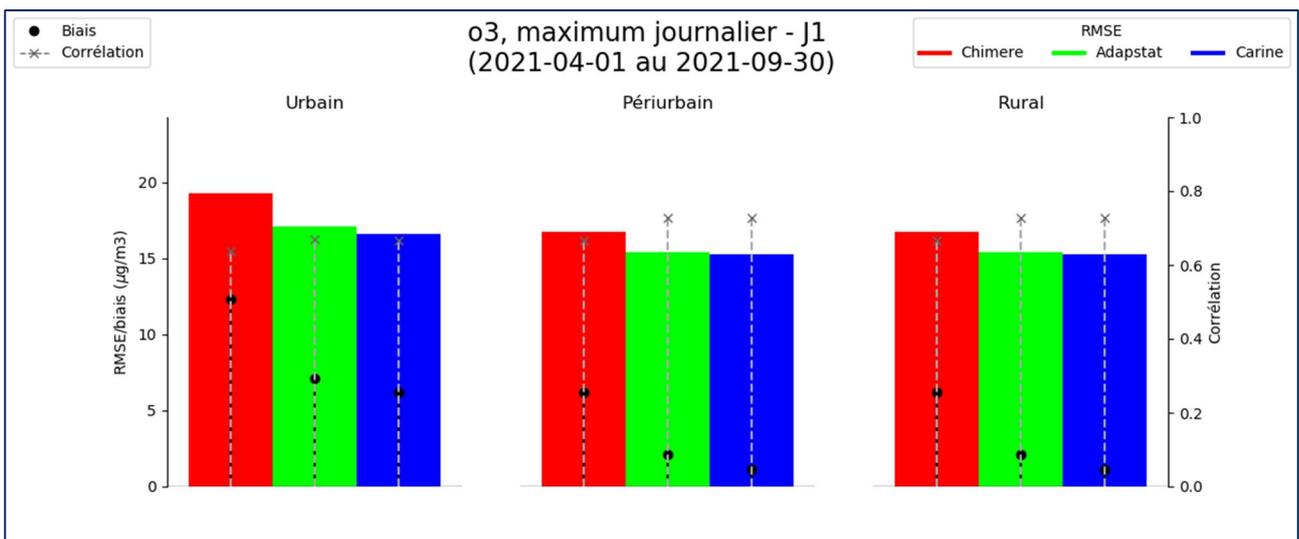


Figure 4b : Ozone - Synthèse des scores de RMSE, biais et corrélation en fonction de la typologie de station (urbaines, périurbaines, rurales) des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) d'avril à septembre 2021 sur l'échéance J+1.

3.2.1.3. Par échéance de prévision

La Figure 5 représente les performances statistiques calculées pour les différentes échéances de prévision (J0, J+1 et J+2) pour la période avril à septembre 2021. Ce calcul a été effectué à la fois pour les résultats ajustés par les prévisionnistes (en bleu), résultats adaptés (en vert) et les résultats bruts (en rouge). Cette figure montre que les résultats de prévision (adaptés et ajustés) sont plus fiables pour J0 (prévision pour le jour-même) que pour le lendemain (J+1) et le surlendemain (J+2). En effet, les biais et les RMSE sont plus faibles à J0 que les autres échéances. Les performances du modèle déterministe Chimère est quasi-identique pour toutes les échéances. Ce diagramme montre encore une fois que les résultats avec et sans ajustement sont globalement similaires. L'ajustement des cartes de prévision a permis d'améliorer légèrement les biais moyens à J+2 par rapport aux résultats adaptés statistiquement.

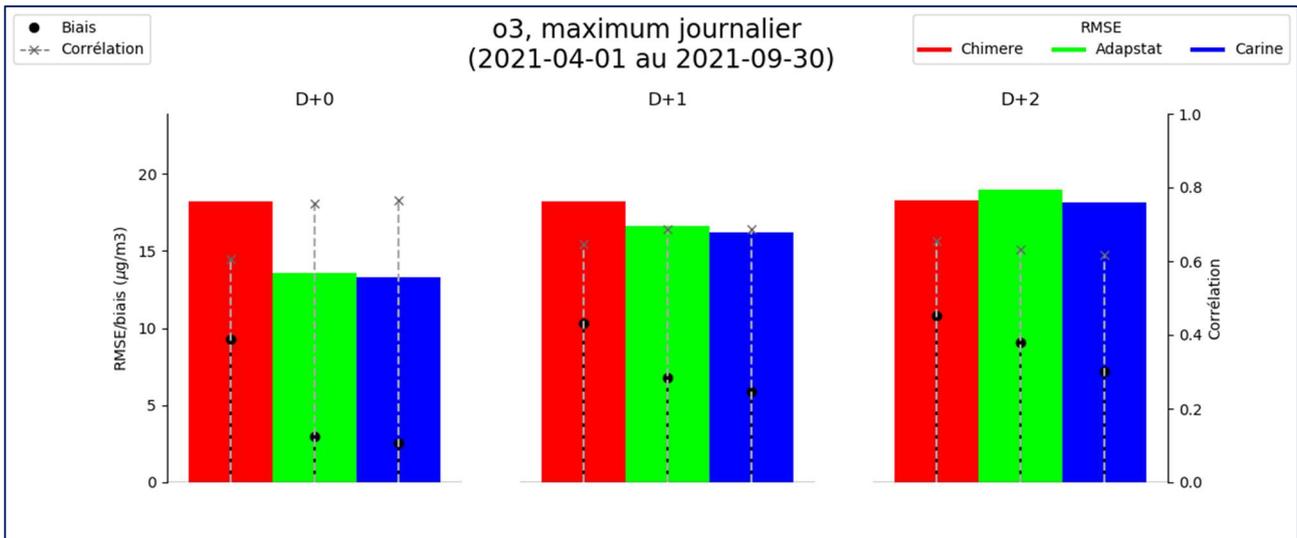


Figure 5 : Ozone - Synthèse des scores de RMSE, biais et corrélation en fonction du jour de prévision pour Carine (bleu), CHIMERE (rouge), et pour l'adaptation statistique (vert) d'avril à septembre 2021

3.2.1.4. Par gamme de concentration

Cette analyse consiste à grouper les données journalières selon leur classe de concentration (par rapport aux observations). Pour chaque groupe, on calcule les scores moyens sur les stations de fond. Les figures 6a et 6b illustrent l'évolution des scores (RMSE et biais moyen) en fonction des classes de concentrations d'O₃. Ces figures montrent que l'évolution des performances des modèles en fonction des gammes de concentrations est quasi-identique pour tous les modèles. En effet, ils sous-estiment les fortes concentrations d'O₃ et surestiment les faibles concentrations. Ce comportement est valable à la fois pour les prévisions de l'échéance J0 et celle de J+1. A noter que les trois modèles sont plus performants pour prévoir les concentrations qui varient entre 90 et 130 µg.m⁻³. Cette analyse a permis d'identifier l'efficacité et la performance des différents modèles dans cette même gamme de concentration (entre 90 et 130 µg.m⁻³) où les scores des modèles à J0 et J+1 sont meilleurs que ceux pour les autres gammes de concentrations.

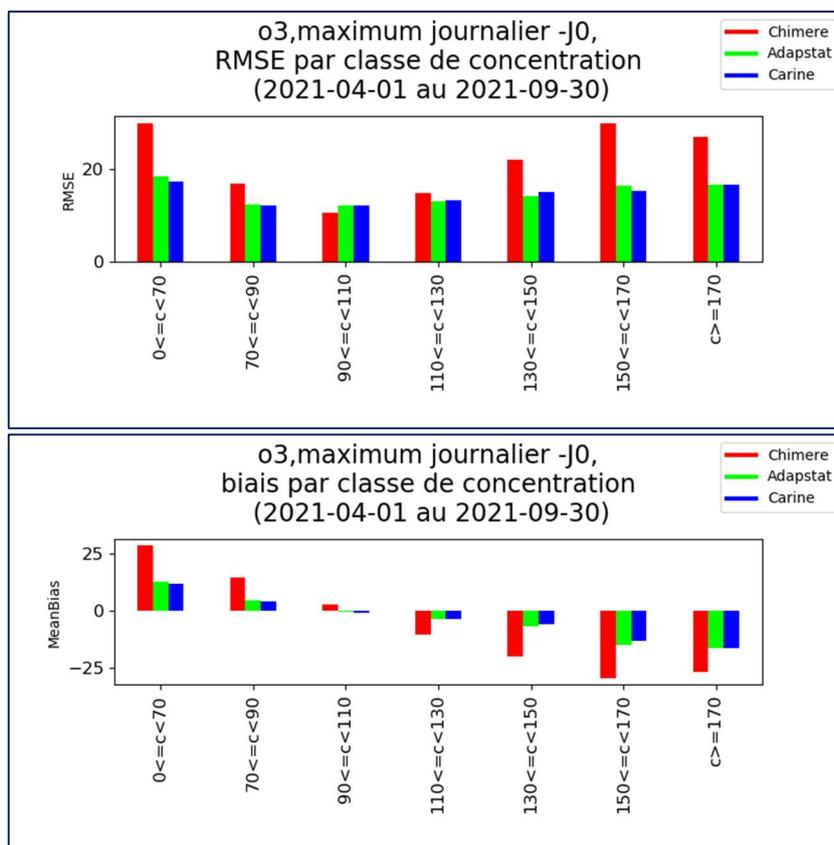


Figure 6a : Ozone - Scores de RMSE et biais par classe de concentrations des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) d'avril à septembre 2021 sur l'échéance J0.

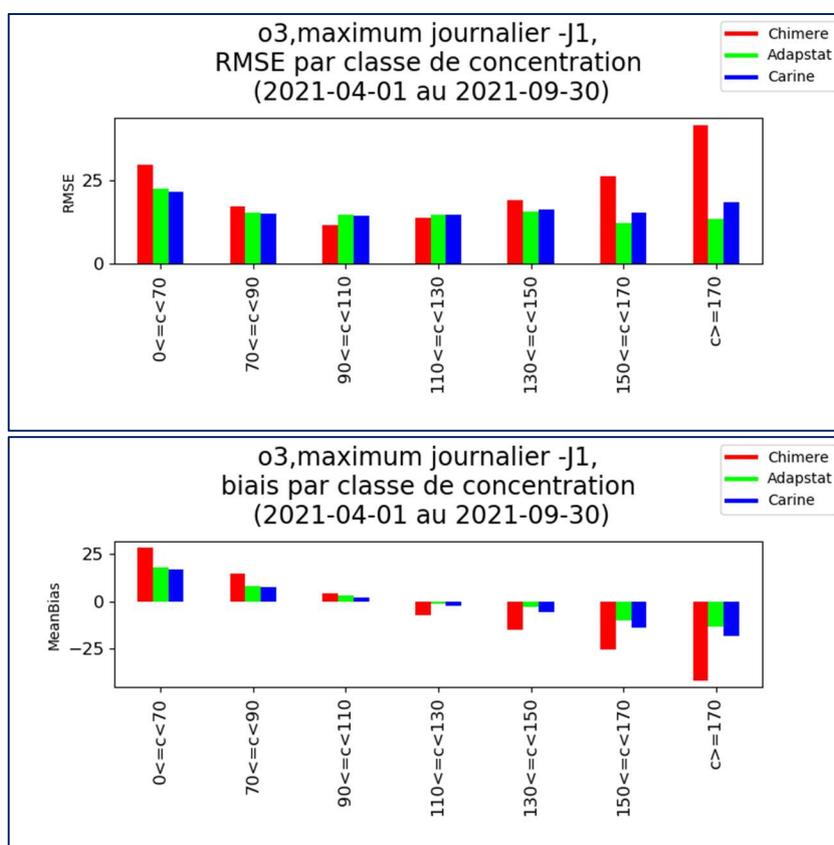


Figure 6b : Ozone - Scores de RMSE et biais par classe de concentration des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) d'avril à septembre 2021 sur l'échéance J+1.

3.2.1.5. Sur plusieurs années

L'objectif de cette partie est d'affiner les performances sur les années précédentes. La Figure 7 illustre l'évolution interannuelle des scores de performances pour les différents modèles. Cette figure montre que les résultats de prévision sont assez satisfaisants pour les modèles ajusté et adapté. En effet, les objectifs de qualité sont largement respectés pour ces deux modèles. Les indicateurs statistiques présentés dans ce graphique montrent que les deux prévisions avant et après l'expertise humaine ont globalement la même performance statistique. Les résultats ajustés par les prévisionnistes sont légèrement meilleurs que les résultats adaptés en 2018 mais ils sont moins bons en 2020. Comme l'année 2020, on constate que les biais calculés pour les modèles adaptés sont positifs. Le fait que le biais moyen global soit positif pour cette année signifie que la chaîne de prévision surestime la moyenne des concentrations pour l'O₃. La Figure 7 montre également que le modèle ajusté génère des biais plus faibles que les modèles brut et adapté (contrairement à l'année 2020). A noter que les performances du modèle ajusté sont meilleures en 2021 que les années précédentes.

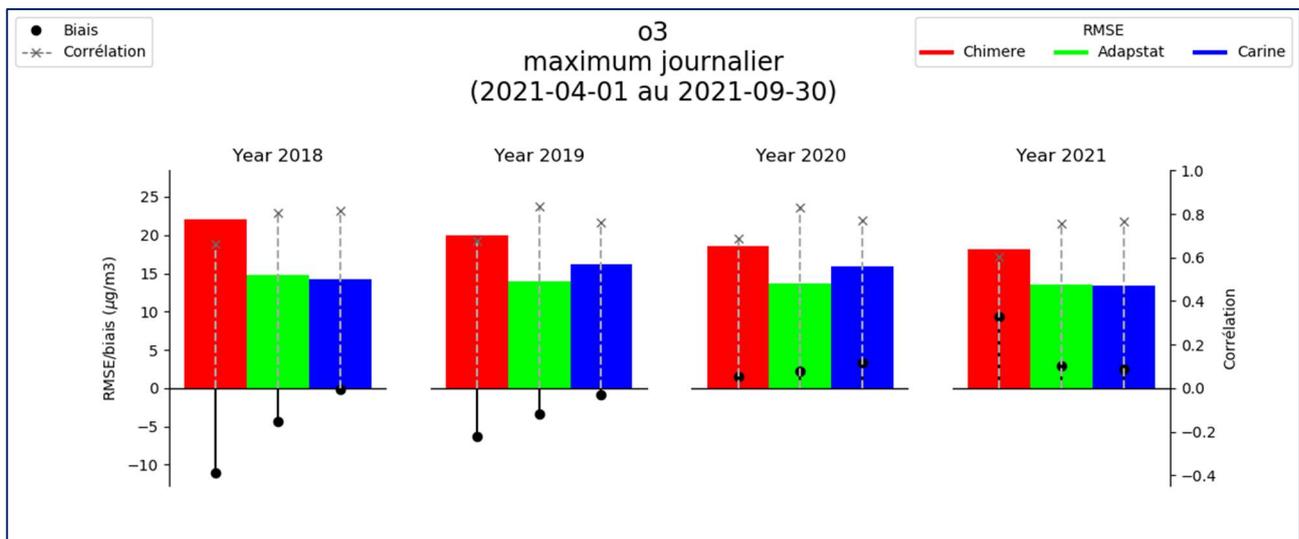


Figure 7 : Ozone - Synthèse des scores de RMSE (barres colorées), biais moyen (points noirs) et corrélation (pointillés gris) entre avril et septembre depuis 2018 pour CHIMERE, pour son adaptation statistique et pour le modèle Carine (modèle expertisé)

3.2.2. Performances sur les seuils réglementaires

Les scores présentés dans cette partie sont issus des méthodes d'évaluation de prévision par dichotomie. Ces méthodes sont fondées sur une table de contingence qui indique le nombre de « oui » et de « non » prévus et mesurés (Tableau 2). Les quatre combinaisons de prévisions et d'observations permettent de calculer des probabilités conjointes et conditionnelles. Le comptage des occurrences se fait par rapport à un seuil prédéfini. Il s'agit ici des seuils de dépassement des valeurs réglementaires (seuil d'information ou d'alerte). Les seuils utilisés comme critères de dépassement sont décrits ci-dessous :

Polluant	Seuil
O ₃ : maximum horaire	180 µg.m ⁻³
PM ₁₀ : moyenne journalière	50 µg.m ⁻³

Tableau 2 : Valeurs des seuils utilisées pour le calcul des critères statistiques

3.2.2.1. Scores aux stations

Dans cette partie est étudiée uniquement la capacité de la chaîne de prévision à reproduire les dépassements du seuil d'information aux stations pour l'O₃. Il convient de rappeler que ces dépassements aux stations n'expliquent pas systématiquement les vigilances de pollution activées sur la région en 2021. En effet, l'épisode de pollution se caractérise par le dépassement d'un seuil de recommandation ou d'alerte avec le respect d'autres critères définis dans l'arrêté zonal N°PREF-DIA-BCI-2018-05-22-01³ (par exemple, nombre d'habitants

³http://www.prefectures-regions.gouv.fr/auvergne-rhone-alpes/content/download/35211/238621/file/5-7-2018_recueil-84-2018-096-recueil-des-actes-administratifs-special.pdf

ou surface concernés). A noter que l'année 2021 ne compte aucun dépassement aux stations du seuil de 180 $\mu\text{g}/\text{m}^3$.

Le tableau de contingence (Tableau 3) compte les éléments suivants :

- Bonne détection (ou good detection) : événement prévu et observé
- Détection manquée (ou miss) : événement non prévu mais observé
- Fausse détection (ou false alarm) : événement prévu mais non observé
- Bonne non-détection (ou correct negative) : événement non prévu et non observé

		Observation		Total
		Oui	Non	
Prévision	Oui	Bonne détection (a)	Fausse détection (b)	Prévu (a+b)
	Non	Détection manquée (c)	Bonne non-détection corrects (d)	Non prévu (c+d)
Total		Observé (a+c)	Non observé (b+d)	

Tableau 3: Table de contingence

Cette table permet de visualiser quel type d'erreur a été commis. Une bonne prévision devrait logiquement ne contenir que des succès et des rejets corrects, autrement dit tous les dépassements et non dépassements sont validés par les mesures. Inversement, une mauvaise prévision correspond à des dépassements ou non dépassements non validés par les mesures. Le nombre total d'évènements prévus et observés est également étudié et est appelé distribution marginale. Une grande variété d'indicateurs statistiques sont calculés à partir des éléments de cette table de contingence.

3.2.2.1.1. Séries temporelles

La Figure 8 illustre le nombre de dépassements pour chaque jour de la période étudiée du seuil d'information et de recommandations pour le polluant O_3 . Elle représente à la fois le nombre de dépassements observés (en gris) et ceux prévus (en couleur) cumulés sur toutes les stations de fond. Ce graphe montre que les deux modèles adaptés (statistiquement et par le prévisionniste) ont affiché quelques dépassements des concentrations d'ozone sur la région. Cependant, aucun constat a été enregistré par les différentes stations de mesures. Le système de prévision a surestimé donc l'ampleur de l'épisode estival d'ozone de mi-août 2021 sur la région Auvergne-Rhône-Alpes (notamment à J+1).

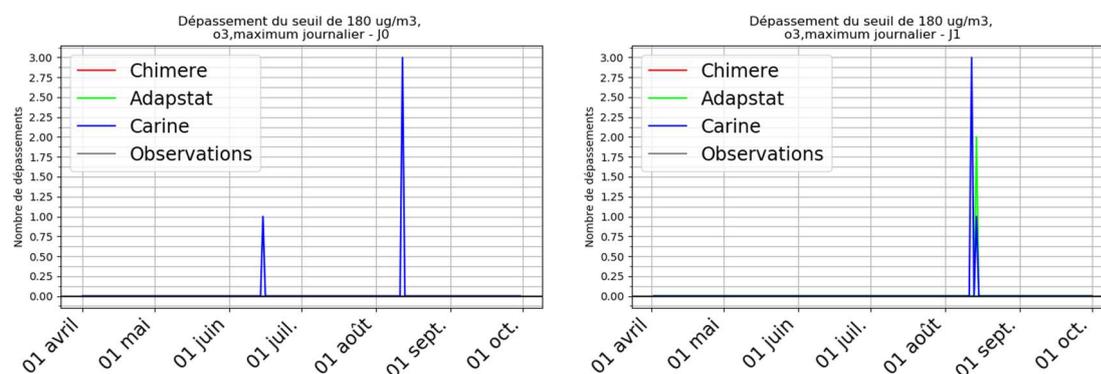


Figure 8 : Ozone - Superposition des dépassements enregistrés et prévus à J0 (à gauche) et à J+1 (à droite) à différentes étapes de la chaîne de prévision : modèle brut (Chimère en rouge), modèle adapté statistiquement (Adapstat - en vert) et le modèle ajusté par le prévisionniste (Carine - en bleu)

3.2.2.1.2. Éléments de la table de contingence

Les figures Figure 9a et Figure 9b illustrent les valeurs des tables de contingence (bonne détection, détection manquée et fausse détection) cumulées sur les stations de fond pour la région à gauche et déclinée selon les zones définies par l'arrêté interpréfectoral. Le graphique à gauche de chaque figure (9a et 9b) affiche la table de contingence pour le modèle de référence (ici Carine pour Atmo Auvergne-Rhône-Alpes).

Ces figures illustrent quelques prévisions non confirmées à J0 et J+1. Ces « fausses détections » reflètent la mise en vigilance par précaution lorsque les niveaux s'approchent des seuils d'activation, ceci afin de limiter le nombre d'événements non prévus. A noter que l'année 2021 ne compte aucun dépassement aux stations du seuil d'information et de recommandations de $180 \mu\text{g}/\text{m}^3$.

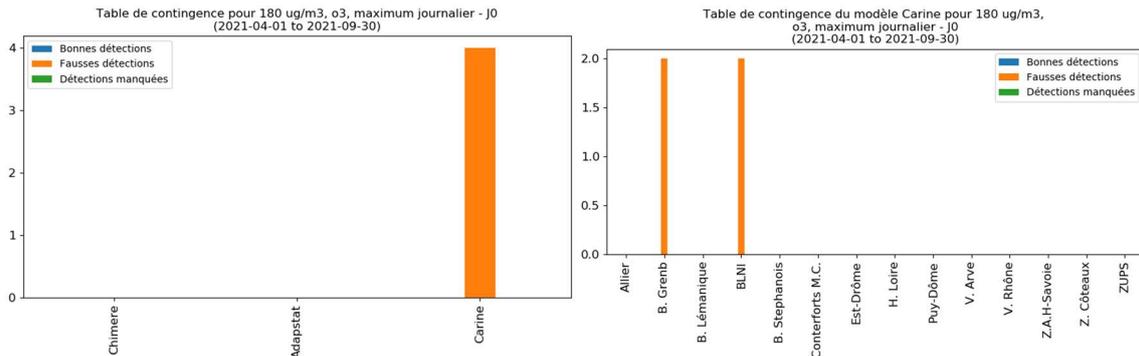


Figure 9a : Ozone - Tables de contingence sur la période d'avril à septembre 2021 à l'échéance J0 pour le seuil de $180 \mu\text{g}/\text{m}^3$. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de CHIMERE (groupe de gauche), ADAPSTAT (groupe du milieu), et de Carine (groupe de droite) pour la région Auvergne-Rhône-Alpes entière. Droite : Table de contingence de l'adaptation statistique déclinée par zones AIP.

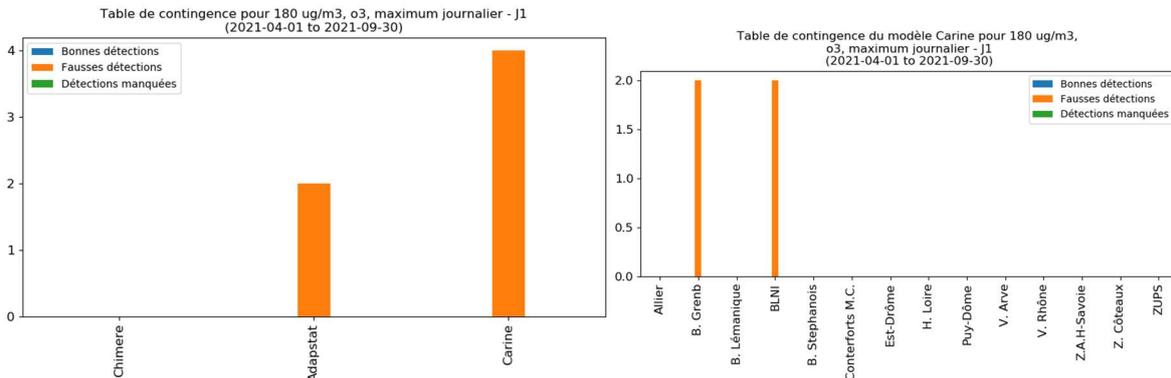


Figure 9b : Ozone - Tables de contingence sur la période d'avril à septembre 2021 à l'échéance J+1 pour le seuil de $180 \mu\text{g}/\text{m}^3$. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de CHIMERE (groupe de gauche), ADAPSTAT (groupe du milieu), et de Carine (groupe de droite) pour la région Auvergne-Rhône-Alpes entière. Droite : Table de contingence de l'adaptation statistique déclinée par zone AIP.

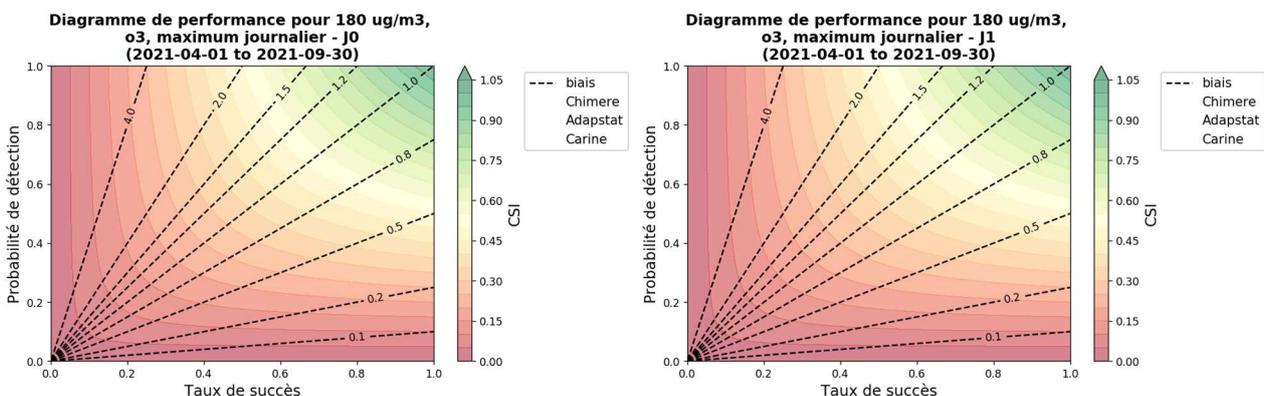


Figure 10 : Ozone - Diagramme de performance pour Chimère (rouge), Adapstat (vert) et Carine (bleu) d'avril à septembre 2021 pour l'échéance J0 (à gauche) et celle de J+1 (à droite)

La Figure 10 a été réalisée afin de représenter une vue complète des performances des prévisions par rapport à la détection des seuils réglementaires (l'objectif est de se situer au plus proche du coin supérieur droit) :

- Axe horizontal : capacité du modèle à faire de bonnes prévisions parmi les détections
- Axe vertical : capacité du modèle à prévoir des dépassements
- Axe transverse : proportion de fausses détections par rapport aux détections manquées
- Axe couleur : capacité du modèle à faire de bonnes prévisions sans évènements parasites (fausses détections, détections manquées)

Cette figure n'est pas exploitable car comme l'illustrent les figures Figure 9a et Figure 9b aucun dépassement du seuil de $180 \mu\text{g}/\text{m}^3$ aux stations n'a été observé pendant l'année 2021.

3.2.2.2. Scores en lien avec les critères réglementaires de déclenchement d'épisode de pollution

Le précédent paragraphe étudie la capacité de la chaîne de prévision à reproduire les dépassements du seuil d'information aux stations. Dans ce paragraphe, l'étude porte uniquement sur les journées d'épisodes de pollution, au sens de l'arrêté zonal N°PREF-DIA-BCI-2018-05-22-01. En effet, l'activation d'une vigilance de pollution n'est pas déterminée par des dépassements ponctuels des valeurs du seuil par une station de mesure mais elle se caractérise par d'autres critères définis dans l'arrêté cité précédemment. L'activation d'une vigilance de pollution dans un bassin d'air doit satisfaire d'au moins un de deux critères suivants :

- a. Critère de déclenchement associé à la superficie : au moins 25 km^2 du même bassin d'air sont concernés par un dépassement du seuil d'information et de recommandations ou/et du seuil d'alerte.
- b. Critère de déclenchement associé à la population :
 - o Dans les bassins d'air de plus de 500 000 habitants (dans les zones Puy-de-Dôme, bassin grenoblois, bassin lyonnais nord-Isère et vallée du Rhône) : lorsqu'au moins 10% de la population du bassin d'air est concernée par un dépassement de seuils.
 - o Dans les bassins d'air de moins de 500 000 habitants : lorsqu'au moins une population de 50 000 habitants au total est concernée par un dépassement de seuils.

Dans ce paragraphe, l'évaluation de performance du système de prévision consiste à calculer les occurrences de dépassements au sens des critères d'activation d'une vigilance de pollution. Chaque jour de l'année correspond exclusivement à un des évènements suivants :

1. Pas de dépassement prévu, pas de dépassement constaté
2. Dépassement prévu et constaté.
3. Dépassement prévu et non confirmé a posteriori
4. Dépassement avéré mais non prévu

L'évaluation des performances de prévision des épisodes de pollution consiste à calculer uniquement les occurrences des éléments type (2,3,4) pour chaque bassin d'air.

Ces résultats sont présentés dans les tables de contingence (Figure 11). Le comptage est effectué sans distinguer le type d'épisode (recommandation ou alerte). L'étude de performance est réalisée pour les prévisions pour le jour-même (J0) et les prévisions pour le lendemain (J+1). Il convient de rappeler que le calcul de pourcentage n'a pas de sens lorsque le nombre de dépassements est très faible. La Figure 11 illustre aucun dépassement d' O_3 constaté sur la région pour l'année 2021. Le nombre relativement élevé de dépassements « prévus non confirmés » traduit un "principe de précaution" : mise en vigilance lorsque les niveaux approchent les seuils d'activation afin de limiter le nombre d'évènements non prévus.

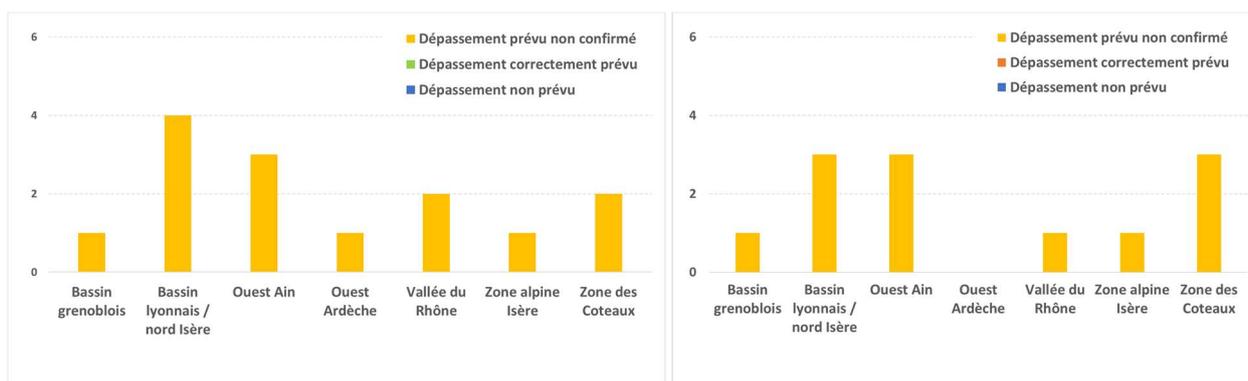


Figure 11 : Ozone - Table de contingence des critères AMU (critères surfaciques et de population confondus) sur la période d'avril à septembre 2021 à l'échéance J0 (à gauche) et à l'échéance J+1 (à droite).

3.3. Prédiction particules fines (PM₁₀)

3.3.1. Scores classiques (Biais, RMSE, Corrélation)

3.3.1.1. Par rapport aux objectifs de qualité

Le Tableau 4 illustre les objectifs de qualité définis pour les particules fines PM₁₀. Ces valeurs doivent s'appliquer aux performances des prévisions en région et sur les zones AIP aux échéances J0 et J+1. Les prévisions de la qualité élaborées par Atmo Auvergne-Rhône-Alpes doivent donc satisfaire à ces valeurs limites.

Biais (µg.m ⁻³)	RMSE (µg.m ⁻³)	Corrélation
-7,5/+7,5	15	0,6

Tableau 4 : Objectif de qualité définis pour les particules fines PM₁₀ s'appliquant aux performances des prévisions en région et sur les zones AIP aux échéances J0 et J+1.

Dans ce paragraphe, le focus est mis sur la performance de la chaîne de prévision par bassin d'air au sens des dispositifs de gestion des épisodes de pollution. Les résultats statistiques par zone d'activation de dispositif sont présentés sur les Figure 12a et Figure 12b qui illustrent les performances statistiques calculées pour les deux échéances J0 et J+1. Ces figures montrent que les résultats de prévision non ajustés sont globalement assez similaires à ceux expertisés par le prévisionniste. On remarque que l'intervention des prévisionnistes engendre une légère dégradation du biais mais l'ajustement des cartes a permis également d'améliorer les résultats de prévision des modèles sur plusieurs zones. On observe également que les biais sont toujours élevés pour le modèle déterministe Chimère (en rouge). Contrairement aux deux modèles ajustés, le biais moyen du modèle brut (Chimère) est négatif pour les particules PM₁₀. Cela signifie que la prévision brute sous-estime la moyenne des concentrations pour ce polluant pour toutes les échéances de prévision.

Globalement, les résultats de la chaîne de prévision sont acceptables sur l'ensemble des zones AIP. Les Figure 12a et Figure 12b montrent que les objectifs de qualité sont globalement respectés par le modèle ajusté pour les deux échéances J0 et J+1. Ces scores sont relativement faibles pour certaines zones comme le Cantal, le Haute-Loire.

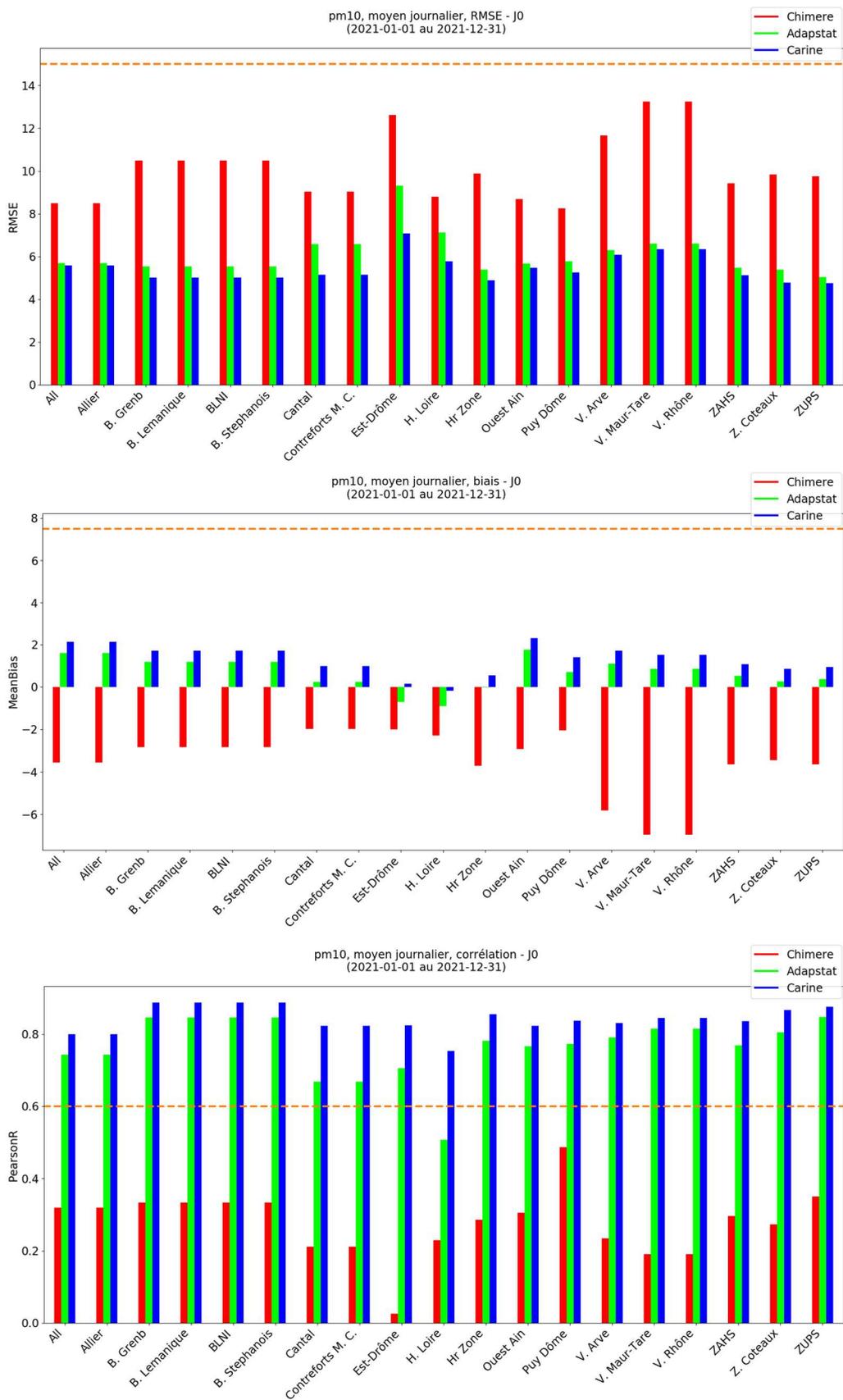


Figure 12a : PM₁₀ - Scores de RMSE, biais et corrélation par zone AIP des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) sur l'échéance J1. La ligne en pointillé orange indique l'objectif de qualité.

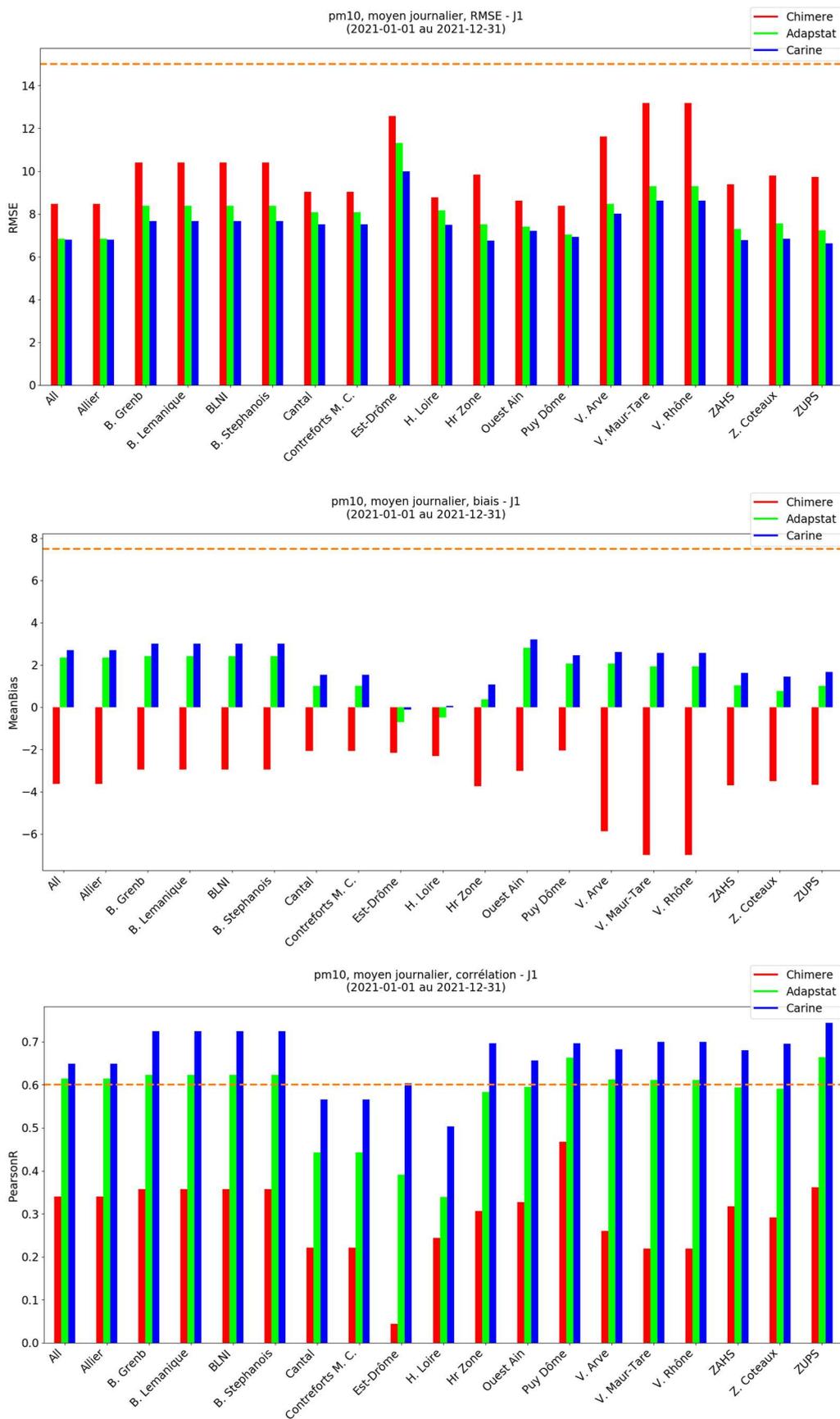


Figure 12b : PM₁₀ - Scores de RMSE, biais et corrélation par zone AIP des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) sur l'échéance J1. La ligne en pointillé orange indique l'objectif de qualité

3.3.1.2. Par typologie

Les Figures Figure 13a et Figure 13b illustrent les performances de la chaîne de prévision en fonction de la typologie des stations de mesures. Comme pour l’ozone, les meilleures performances statistiques sont globalement associées aux stations rurales. Cependant, les plus mauvais résultats sont souvent associés aux sites de typologie urbaines pour lesquels il y a une forte influence de sources locales et donc une très importante variabilité spatio-temporelle des émissions autour des sites de mesures. Les bonnes performances observées sur les sites ruraux peuvent s’expliquer également par l’aptitude de la chaîne de modélisation à reproduire les faibles concentrations. Globalement, les indicateurs statistiques calculés dans les figures Figure 13a et Figure 13b indiquent un très bon accord entre les prévisions et les mesures. En effet, les indicateurs statistiques des prévisions ne dépassent pas la fourchette des valeurs considérées comme acceptables pour toutes les typologies étudiées.

La comparaison des prévisions avant et après l’expertise humaine montrent encore une fois un résultat quasi-identique et très similaire pour les particules fines PM₁₀. Néanmoins, on constate que l’ajustement des cartes par les prévisionnistes a amélioré légèrement les prévisions à J0 et à J+1 quel que soit la typologie des stations.

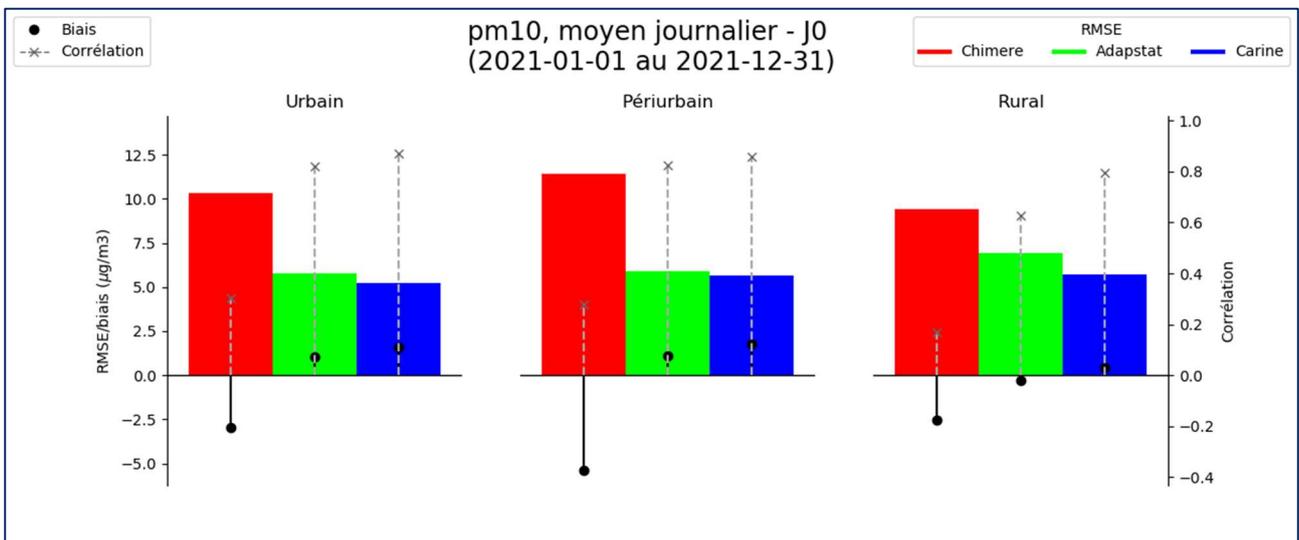


Figure 13a : PM₁₀ - Synthèse des scores de RMSE, biais et corrélation en fonction de la typologie de station (urbaines, périurbaines, rurales) des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) sur l'échéance J0

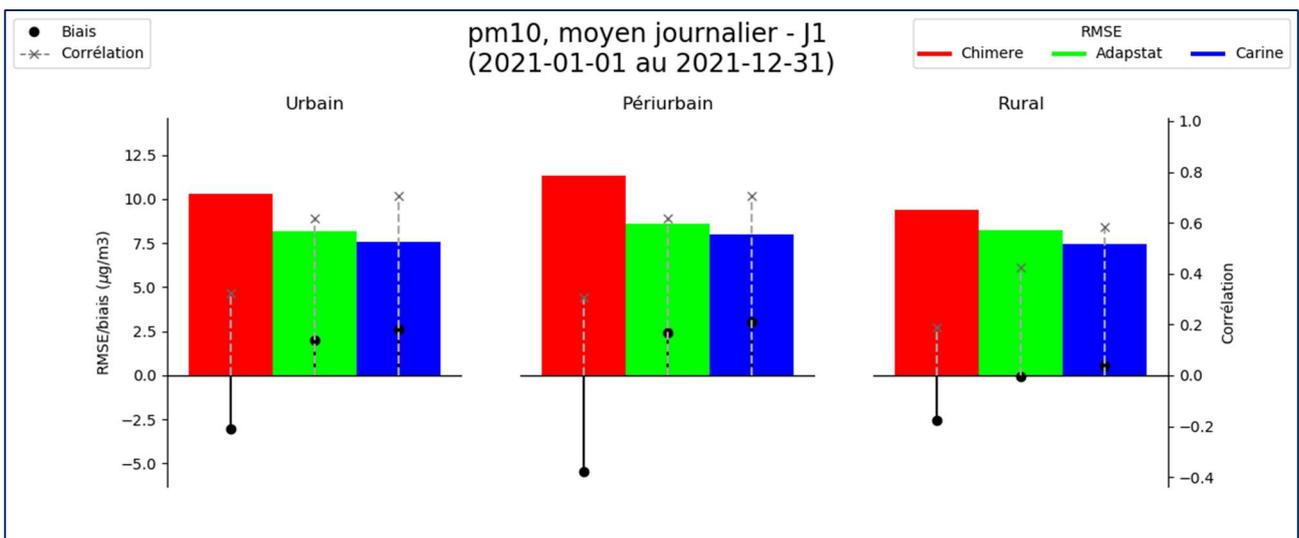


Figure 13b : PM₁₀ - Synthèse des scores de RMSE, biais et corrélation en fonction de la typologie de station (urbaines, périurbaines, rurales) des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) sur l'échéance J0

3.3.1.3. Par échéance de prévision

La Figure 14 représente les performances statistiques calculées pour les différentes échéances de prévision (J0, J+1 et J+2) pour l'année 2021. Ce calcul a été effectué à la fois pour les résultats ajustés par les prévisionnistes (en bleu), les résultats adaptés (en vert) et les résultats bruts (en rouge). Cette figure montre que les résultats de prévision (adaptés et ajustés) sont beaucoup plus fiables pour J0 (prévision pour le jour-même) que pour le lendemain (J+1) et le surlendemain (J+2). En effet, les corrélations, les biais et les RMSE sont plus faibles à J0 que les autres échéances. Les performances du modèle déterministe Chimère sont quasi-identiques pour toutes les échéances. Ce diagramme montre encore une fois que les résultats avec et sans ajustement sont globalement similaires. L'ajustement des cartes de prévision a permis d'améliorer légèrement les scores statistiques du RMSE et de la corrélation par rapport aux résultats adaptés statistiquement.

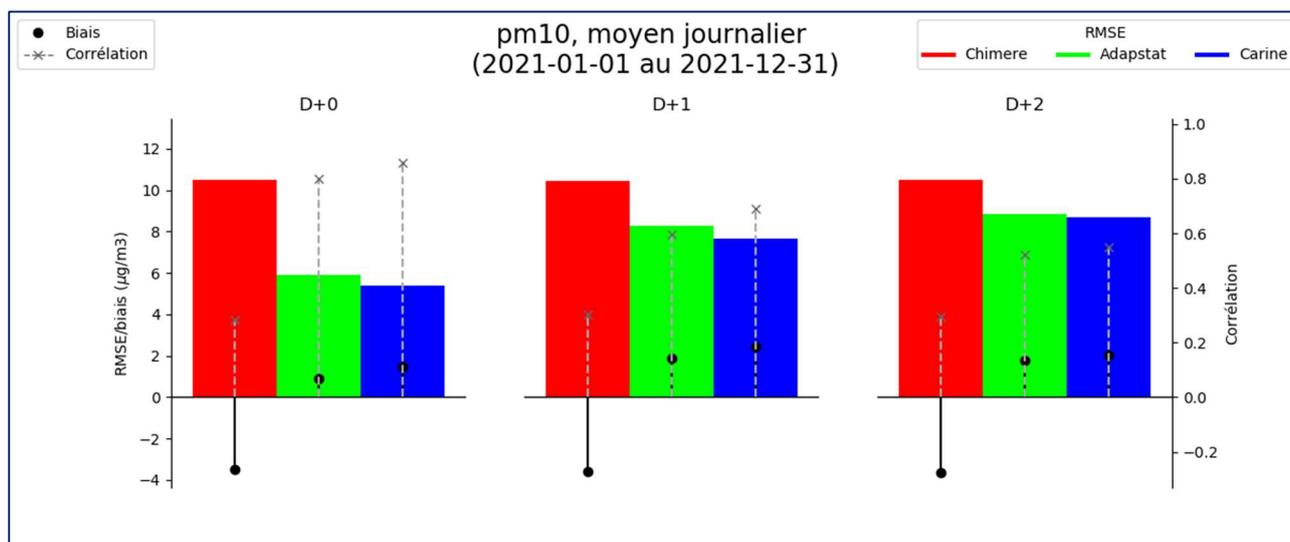


Figure 14 : PM_{10} - Synthèse des scores de RMSE, biais et corrélation en fonction du jour de prévision pour Carine (bleu), CHIMERE (rouge), et pour l'adaptation statistique (vert) en 2020.

3.3.1.4. Par gamme de concentration

Cette analyse consiste à grouper les données journalières selon leur classe de concentration (par rapport aux observations). Pour chaque groupe, on calcule les scores moyens sur les stations de fond. Les figures Figure 15a et Figure 15b illustrent l'évolution des scores (RMSE et biais moyen) en fonction des classes de concentrations des particules fines PM_{10} . Ces figures montrent que l'évolution des performances des modèles en fonction des gammes de concentrations est quasi-identique pour tous les modèles. En effet, ils sous-estiment les concentrations des particules. Ce comportement est valable à la fois pour les prévisions de l'échéance J0 et celles de J+1. En effet, on constate une dégradation monotone des scores par rapport aux concentrations. Les biais élevés observés pour les gammes des concentrations supérieures à $50 \mu g.m^{-3}$ traduisent une difficulté de la chaîne de prévision à bien reproduire les plus fortes concentrations des particules fines PM_{10} .

Ces figures mettent en évidence la capacité de l'adaptation statistique à améliorer les résultats bruts de prévision pour les particules fines PM_{10} . En effet, le modèle brut Chimère sous-estime largement les concentrations des particules fines PM_{10} . Les figures Figure 15a et Figure 15b montrent que les résultats de prévision adaptés statistiquement sont meilleurs à ceux expertisés par le prévisionniste.

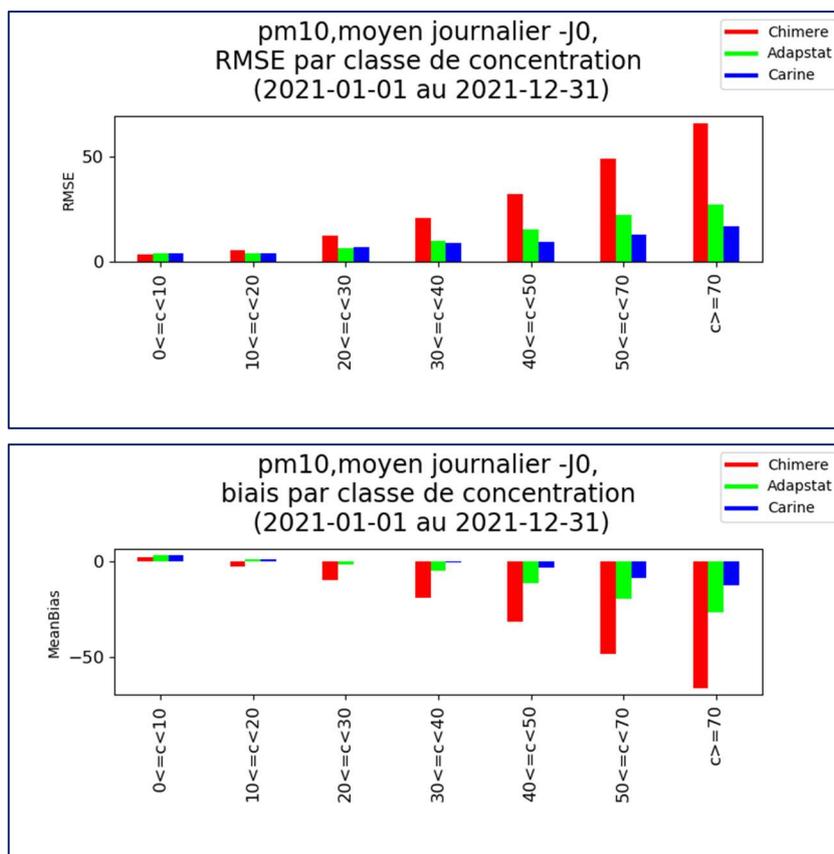


Figure 15a : PM_{10} - Scores de RMSE et biais par classe de concentration des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) sur l'échéance J0.

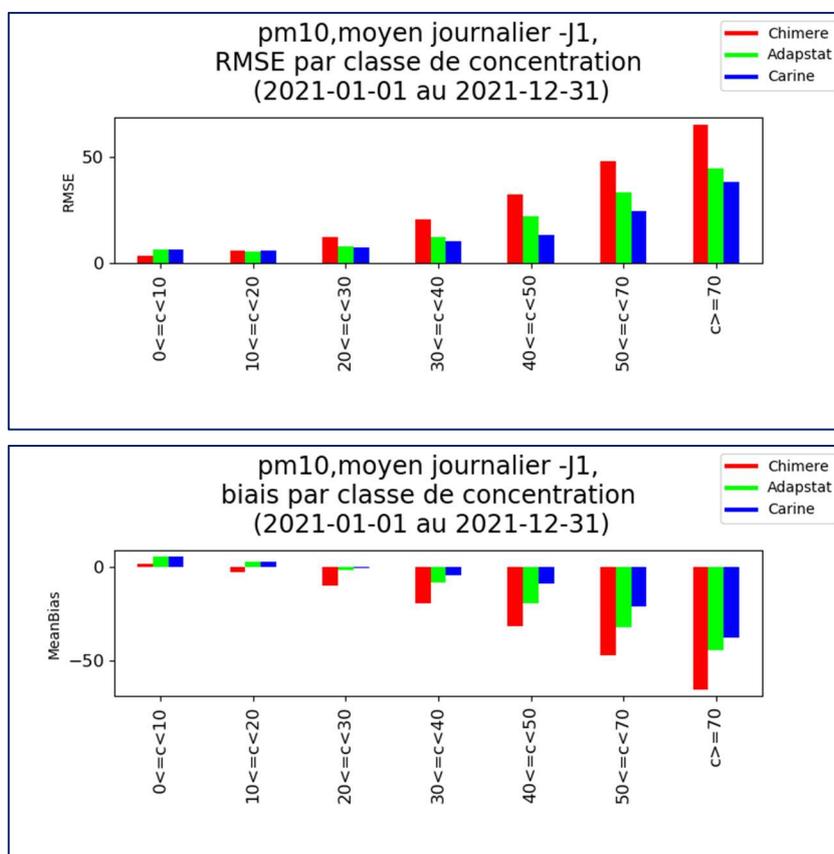


Figure 15b : PM_{10} - Scores de RMSE et biais par classe de concentration des prévisions Carine (bleu), CHIMERE (rouge), et de l'adaptation statistique (vert) sur l'échéance J+1.

3.3.1.5. Sur plusieurs années

L'objectif de cette partie est d'affiner les performances sur les années précédentes. La Figure 16 illustre l'évolution interannuelle des scores de performances pour les différents modèles. Cette figure montre que les résultats de prévision sont assez satisfaisants pour les modèles ajusté et adapté. En effet, les objectifs de qualité sont largement respectés pour ces deux modèles. Les indicateurs statistiques présentés dans ce graphique montrent que les deux prévisions avant et après l'expertise humaine ont globalement la même performance statistique. Les résultats adaptés statistiquement sont légèrement meilleurs que les résultats ajustés. Comme les années précédentes, les biais calculés pour les modèles adaptés sont positifs. Le fait que le biais moyen global soit légèrement positif pour cette année signifie que la chaîne de prévision surestime la moyenne des concentrations pour les particules fines. La Figure 16 montre que le modèle brut génère des biais négatifs beaucoup plus élevés que les modèles adaptés et ajustés.

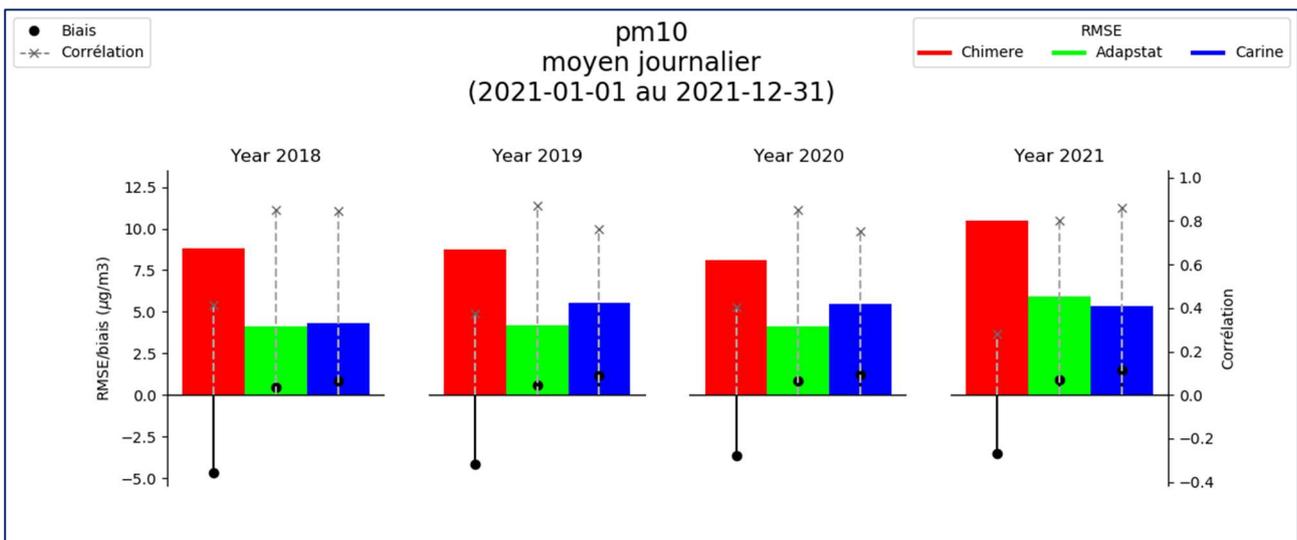


Figure 16 : PM10 - Synthèse des scores de RMSE (barres colorées), biais moyen (points noirs) et corrélation (pointillés gris) depuis 2018 pour CHIMERE, pour son adaptation statistique et pour le modèle Carine (modèle expertisé)

3.3.2. Performances sur les seuils réglementaires

Les scores présentés dans cette partie sont issus des méthodes d'évaluation de prévision par dichotomie. Ces méthodes sont fondées sur une table de contingence qui indique le nombre de « oui » et de « non » prévus et mesurés (Tableau 2) (cf. §3.2.2. - page16).

3.3.2.1. Scores aux stations

Dans cette partie est étudiée uniquement la capacité de la chaîne de prévision à reproduire les dépassements du seuil d'information aux stations pour les particules fines PM₁₀. Il convient de rappeler que ces dépassements aux stations n'expliquent pas systématiquement les vigilances de pollution activées sur la région en 2021. En effet, l'épisode de pollution se caractérise par le dépassement d'un seuil de recommandation ou d'alerte avec le respect d'autres critères définis dans l'arrêté zonal N°PREF-DIA-BCI-2018-05-22-01⁴ (cf. §3.2.2.1 – page 16).

3.3.2.1.1. Séries temporelles

La Figure 17 illustre le nombre de dépassements journaliers du seuil d'information et de recommandations pour le polluant PM₁₀. Elle représente à la fois le nombre de dépassements observés (en gris) et ceux prévus (en couleur) cumulés sur toutes les stations de fond pour chaque journée de la période considérée. Ce graphe montre que les deux modèles adaptés (statistiquement et par le prévisionniste) ont réussi à détecter et anticiper efficacement l'occurrence des épisodes de pollution. Cependant, quelques dépassements ponctuels

⁴http://www.prefectures-regions.gouv.fr/auvergne-rhone-alpes/content/download/35211/238621/file/5-7-2018_recueil-84-2018-096-recueil-des-actes-administratifs-special.pdf

(1 fois pour les particules fines) n'ont pas été prévus par les différents modèles de prévision. Il s'agit des épisodes de pollution mixte ou des épisodes de type Dust (poussières désertiques).

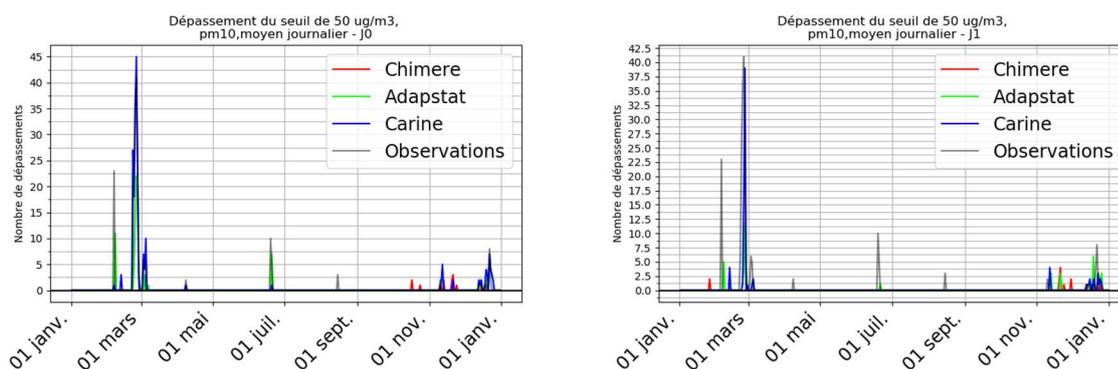


Figure 17 : PM₁₀ - Superposition des dépassements enregistrés et prévus à J0 (à gauche) et à J+1 (à droite) à différentes étapes de la chaîne de prévision : modèle brut (Chimère en rouge), modèle adapté statistiquement (Adapstat - en vert) et le modèle ajusté par le prévisionniste (Carine – en bleu)

3.3.2.1.2. Éléments de la table de contingence

Les figures Figure 18a et Figure 18b illustrent les valeurs des tables de contingence (Bonne détection, Détection manquée et Fausse détection) cumulées sur les stations de fond pour la région à gauche et décliné par zone AIP disposant d'un arrêté mesure d'urgence. Le graphique à gauche de chaque figure (Figure 18a et Figure 18b) affiche la table de contingence pour le modèle de référence (ici Carine pour Atmo Auvergne-Rhône-Alpes).

L'intercomparaison entre les résultats adaptés statistiquement et ajustés par le prévisionniste a montré que l'expertise humaine des cartes améliore nettement le taux de détection des seuils d'information à J0 et améliore légèrement ce taux de détection pour J+1. La Figure 18a illustre une nette amélioration des performances du modèle ajusté par rapport au modèle adapté statistiquement. Le nombre relativement élevé de dépassements « prévus non confirmés » sur certaines zones reflète la mise en vigilance par précaution lorsque les niveaux s'approchent des seuils d'activation, ceci afin de limiter le nombre d'évènements non prévus.

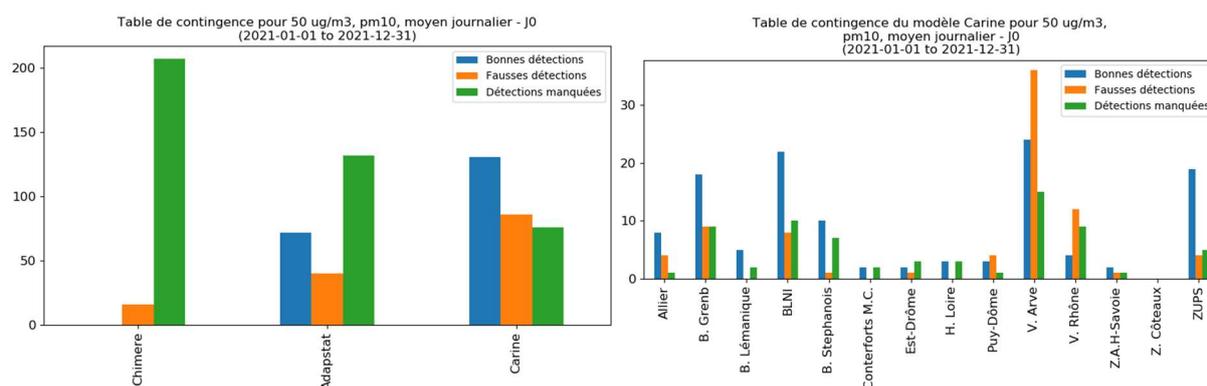


Figure 18a : PM₁₀ - Tables de contingence pour l'année 2021 à l'échéance J0 pour le seuil de 50 µg/m³. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de CHIMERE (groupe de gauche), ADAPSTAT (groupe du milieu), et de Carine (groupe de droite) pour la région Auvergne-Rhône-Alpes entière. Droite : Table de contingence de l'adaptation statistique déclinée par zone AIP.

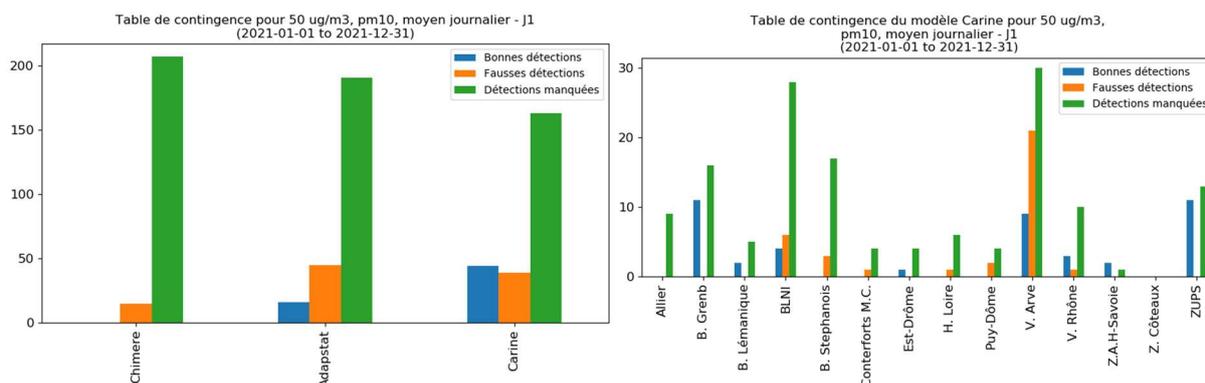


Figure 18b : PM₁₀ - Tables de contingence calculée sur l'année 2021 à l'échéance J+1 pour le seuil de 50 µg/m³. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de CHIMERE (groupe de gauche), ADAPSTAT (groupe du milieu), et de Carine (groupe de droite) pour la région Auvergne-Rhône-Alpes entière. Droite : Table de contingence de l'adaptation statistique déclinée par zone AIP.

La Figure 19 représente une vue complète des performances des prévisions par rapport à la détection des seuils réglementaires (l'objectif est de se situer au plus proche du coin supérieur droit) :

- Axe horizontal : capacité du modèle à faire de bonnes prévisions parmi les détections
- Axe vertical : capacité du modèle à prévoir des dépassements
- Axe transverse : proportion de fausses détections par rapport aux détections manquées
- Axe couleur : capacité du modèle à faire de bonnes prévisions sans événements parasites (fausses détections, détections manquées).

Cette figure confirme notre analyse pour les prévisions des particules fines PM₁₀. Elle illustre des meilleures performances du modèle ajusté pour J+1 et une meilleure efficacité du modèle adapté sur J0. On constate que l'ajustement par le prévisionniste a permis d'assurer une probabilité de détection bien supérieure à 60% à J0 et un taux de succès supérieur à 50% pour les deux échéances (J0 et J+1).

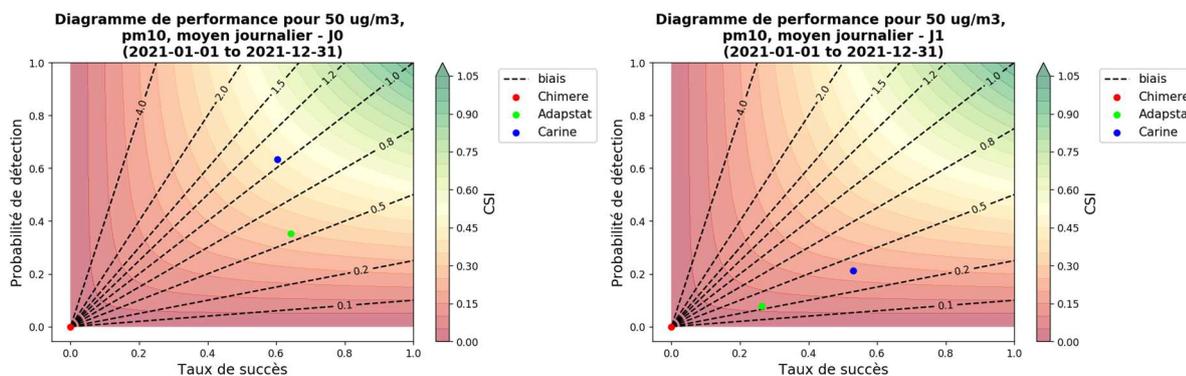


Figure 19 : PM₁₀ - Diagramme de performance pour Chimère (rouge), Adapstat (vert) et Carine (bleu) en 2021 pour l'échéance J0 (à gauche) et celle de J+1 (à droite)

3.3.2.2. Scores en lien avec les critères réglementaires de déclenchement d'un épisode de pollution

Le précédent paragraphe étudie la capacité de la chaîne de prévision à reproduire les dépassements du seuil d'information aux stations. Dans ce paragraphe, l'étude porte uniquement sur les journées d'épisodes de pollution, au sens de l'arrêté zonal N°PREF-DIA-BCI-2018-05-22-01. En effet, l'activation d'une vigilance de pollution n'est pas déterminée par des dépassements ponctuels des valeurs du seuil dans une station de mesure mais elle se caractérise par d'autres critères définis dans l'arrêté cité précédemment. (cf. §3.2.2.2 - page 19)

Dans ce paragraphe, l'évaluation de performance de notre système de prévision consiste à calculer les occurrences de dépassements au sens des critères d'activation d'une vigilance de pollution. Chaque jour de l'année correspond exclusivement à un des événements suivants :

1. Pas de dépassement prévu, pas de dépassement constaté
2. Dépassement prévu et constaté.
3. Dépassement prévu et non confirmé a posteriori
4. Dépassement avéré mais non prévu.

L'évaluation des performances de prévision des épisodes de pollution consiste à calculer uniquement les occurrences des éléments type (2,3,4) pour chaque bassin d'air.

Ces résultats sont présentés dans les tables de contingence (Figure 20). Le comptage est effectué sans distinguer le type d'épisode (recommandation ou alerte). L'étude de performance est réalisée pour les prévisions du jour-même (J0) et les prévisions de lendemain (J+1). Il convient de rappeler que le calcul de pourcentage n'a pas de sens lorsque le nombre de dépassements est très faible. La Figure 20 illustre un nombre de dépassements beaucoup plus élevés pour les particules par rapport à celui de l'ozone (Figure 11). Le nombre relativement élevé de dépassements « prévus non confirmés » reflète la mise en vigilance par précaution lorsque les niveaux s'approchent des seuils d'activation, ceci afin de limiter le nombre d'événements non prévus. Ainsi, les prévisions « non confirmées » se situent systématiquement en début et fin d'épisodes (difficulté à cerner les bornes des épisodes). Cependant, les niveaux atteints lors de ces journées étaient souvent proches du seuil d'activation.

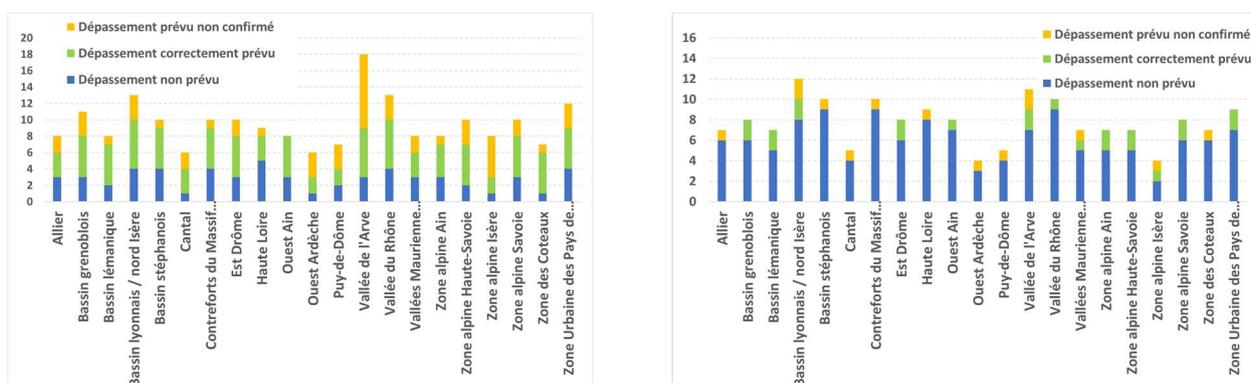


Figure 20 : PM10 - Table de contingence des critères AMU (critères surfaciques et de population confondus) sur toute l'année 2020 à l'échéance J0 (à gauche) et à l'échéance J+1 (à droite).

4. Conclusion

Dans ce rapport, a été étudiée la performance des différentes étapes du système de prévision de la qualité de l'air d'Atmo Auvergne-Rhône-Alpes : résultats fournis par le modèle brut, résultats adaptés statistiquement et résultats expertisés par les prévisionnistes de la qualité de l'air.

Selon les valeurs de références proposées par le référentiel technique national pour l'évaluation des performances des prévisions de la qualité de l'air, les résultats du système de prévision sont considérés comme satisfaisants pour les concentrations journalières des PM₁₀ et maximales pour l'O₃. Les valeurs des indicateurs sont dans la plage acceptable et même proches des valeurs « idéales ». Le calcul de ces indicateurs statistiques a été fondé sur les résultats prenant en compte tous les sites de mesures.

L'analyse effectuée a montré que les résultats bruts ont une tendance à sous-estimer les épisodes de pollution. Ce comportement est plus affirmé pour les particules fines.

Les performances statistiques des résultats ajustés et non ajustés par les prévisionnistes sont globalement similaires. Cependant, on constate qu'en 2021, l'expertise humaine a apporté une amélioration dans les statistiques de prévision des dépassements de seuils (épisodes de pollution) pour J et J+1.

Enfin, en prenant en compte les critères d'activation d'une vigilance de pollution (épisode de pollution), on constate que les performances du système de prévision dépendent du bassin d'air et du nombre d'épisodes.

En 2021, la région Auvergne Rhône-Alpes a vu 25 jours avec une vigilance activée, cela en fait l'année avec le plus faible total de cette dernière décennie (53 jours en moyenne sur 10 ans, 34 jours en moyenne sur 5 ans). Le bassin lyonnais Nord-Isère et la vallée de l'Arve restent les territoires sur lesquels le plus grand nombre de vigilances atmosphériques ont été activées cette année (17 jours). Le taux de détection et le ratio de fausses alarmes ont été respectivement de 60% et 36% pour J0. A noter que 100% des journées ayant vu un de vigilance ont concerné un risque de dépassement des seuils réglementaires ont concerné les de particules PM10. Pour autant, même si les seuils réglementaires pour l'ozone n'ont pas été dépassés, des risques de dépassement de ces seuils ont été détectés pour le lendemain ou le surlendemain, entraînant la mise en place de vigilance sur les bassins d'air concernés pour les journées concernées.

En 2022, plusieurs travaux visant à améliorer la chaîne de prévision seront menés. Ils porteront notamment sur le modèle lui-même (nouvelle version de CHIMERE), les données de bord (conditions aux limites de notre région) et les données d'émissions.

Annexe 1

La région est découpée en 21 zones, les critères sont vérifiés quotidiennement dans chaque zone.

