



Evaluation des performances de prévision de la qualité de l'air sur la région Grand Est pour l'année 2022

Conformément au référentiel technique national

CONDITIONS DE DIFFUSION

Diffusion libre pour une réutilisation ultérieure des données dans les conditions ci-dessous :

- Les données produites par ATMO Grand Est sont accessibles à tous sous licence libre «**ODbL v1.0**».
- Sur demande, ATMO Grand Est met à disposition les caractéristiques des techniques de mesures et des méthodes d'exploitation des données mises en œuvre ainsi que les normes d'environnement en vigueur et les guides méthodologiques nationaux.
- ATMO Grand Est peut rediffuser ce document à d'autres destinataires.
- Rapport non rediffusé en cas de modification ultérieure des données.

PERSONNES EN CHARGE DU DOSSIER

Rédaction : *Nicolas Thorr, Florent VASBIEN, Ingénieurs d'étude*
Relecture : *Jérôme LE PAIH, Responsable Unité Modélisation*
Bérénice JENNESON, Responsable Unité Surveillance
Approbation : *Cyril PALLARES, Directeur opérationnel*

Référence du modèle de rapport : COM-FE-001_7

Référence du projet : MOD_PREVISION_RESOLUTION_KM

Référence du rapport : MOD-EN-084_1

Date de publication : juin 2023

ATMO Grand Est

Espace Européen de l'Entreprise – 5 rue de Madrid – 67300 Schiltigheim

Tél : 03 69 24 73 73

Mail : contact@atmo-grandest.eu

SOMMAIRE

1. SYNTHÈSE.....	3
2. PRÉSENTATION DE LA PLATEFORME DE PRÉVISION	4
3. ÉVALUATION DE LA PLATEFORME DE PRÉVISION.....	5
3.1. PRÉVISIONS OZONE.....	5
3.1.1. Scores classiques (Biais, RMSE et corrélation)	5
3.1.2. Performances sur les seuils réglementaires.....	13
3.2. PRÉVISIONS PM10	18
3.2.1. Scores classiques (Biais, RMSE et corrélation)	18
3.2.2. Performances sur les seuils réglementaires.....	24

1. SYNTHÈSE

Les résultats présentés dans ce document permettent d'établir un diagnostic des performances des prévisions brutes, des prévisions issues de l'adaptation statistique de 10h de la plate-forme PREVEST et des prévisions issues de l'outil PREDIQA (expertise prévisionniste) pour la région Grand Est en ozone et en particules PM10 sur l'année 2022.

Au regard des objectifs de qualité fixés sur les scores classiques (RMSE, corrélation et biais), l'ensemble des indicateurs sont respectés pour les prévisions issues de l'adaptation statistiques et pour celles expertisées, pour l'ozone comme pour les PM10. Tous les départements sont concernés, pour l'échéance J0 (jour-même) comme pour l'échéance J1 (prévision pour le lendemain).

L'outil EVALTOOLS dans sa version 1.0.7 a été utilisé pour cette évaluation (<https://opensource.umr-cnrm.fr/projects/evaltools>).

L'analyse par typologie n'a pas fait ressortir de profil (rural, périurbain, urbain) pour lequel la prévision brute, l'adaptation statistique ou l'expertise prévisionniste ressort particulièrement sur 2022. Les résultats sont sensiblement identiques d'un profil à l'autre, pour les PM10 comme pour l'ozone.

L'analyse par horizon montre une relative stabilité dans les scores de la prévision brute, que l'échéance soit proche (J0) ou lointaine (J2). Avec l'adaptation statistique de 10h, les scores sont nettement améliorés pour l'échéance du jour-même car cette prévision intègre des observations aux stations jusqu'à 10h. Entre les échéances J1 et J2, les scores se stabilisent, tout en bénéficiant là encore de la plus-value de la correction statistique par rapport au brut

L'année 2022 a été simulée avec des performances comparables à l'année 2021. On notera une dégradation du biais pour les PM10 et une légère dégradation de la RMSE pour l'ozone par le modèle brut. L'adaptation statistique de 10h a été profitable sur l'ensemble des indicateurs (biais, RMSE, corrélation) en ramenant les statistiques vers leur optimum. La nouvelle technique d'adaptation statistique en 2022 fondée sur la prévision d'ensemble a permis une amélioration sensible des scores comparativement aux autres années pour l'ozone en termes de corrélation et de RMSE. Les prévisions en sortie de PREDIQA sont du même ordre de performance que celles de l'adaptation statistique. Elles sont légèrement meilleures sur le biais mais dégradent un peu la RMSE pour l'ozone.

En termes de détection de la valeur seuil pour l'ozone ($180 \mu\text{g}/\text{m}^3$ en maximum journalier), l'année 2022 se distingue par de nombreuses détections manquées. Si PREDIQA a pu en détecter correctement une partie, cela se fait au prix de nombreuses fausses alertes. Concernant les PM10 ($50 \mu\text{g}/\text{m}^3$ en moyenne journalière), les dépassements sont repérés pour environ un quart par l'adaptation statistique à l'échéance J+0 et pour un tiers par PREDIQA. L'expertise prévisionniste (associée à l'agrégation communale) remplit son rôle correcteur en permettant une augmentation de la détection des épisodes à J+0 comme à J+1, et ce, sans occasionner de fausses détections supplémentaires par rapport aux autres modèles.

Pour les critères de dépassements AMU (critères surfaciques et de population confondus), le schéma est le même, à savoir, une meilleure détection des dépassements avec PREDIQA.

2. PRESENTATION DE LA PLATEFORME DE PREVISION

La plate-forme de prévision PREVEST est pilotée par ATMO Grand Est. Les prévisions de l'année 2022 sont issues d'un enchaînement de traitements faisant intervenir les éléments suivants :

- Modèles
 - Météo : WRF 4.0.3
 - Chimie : CHIMERE V2017r4 avec conditions aux limites PREVAIR par défaut
- Post-traitements

Les sorties de chimie brutes sont ensuite traitées par une adaptation statistique (AS) au niveau des stations de mesures. L'adaptation statistique est réalisée par un modèle de prévision d'ensemble (PE). Il s'agit d'une régression linéaire pénalisée Lasso à coefficients positifs, de plusieurs modèles (modèles Prev'Est, modèles d'AirParif, modèles d'ATMO Sud et modèles européens COPERNICUS). Les résultats de l'AS, ponctuels, sont ensuite restitués cartographiquement par un krigeage des innovations. Les corrections par l'AS sont réalisées 3 fois par jour (8h, 10h et 16h). Les sorties de 10h sont évaluées dans ce document car il s'agit des prévisions mises à la disposition des prévisionnistes pour leur prise de décision avant midi.

L'expertise prévisionniste a lieu entre 10h30 et 12h. Le prévisionniste dispose de la possibilité d'ajuster ces prévisions par un outil interne : PREDIQA (PREvision De l'Indice de Qualité de l'Air). Cet outil fonctionne à une résolution communale. Ainsi, si les données issues du modèle brut et AS sont sur une résolution kilométrique et extraites aux stations dans le cadre de cette évaluation, les données issues de l'outil PREDIQA sont quant à elle à résolution communale et correspondent à la maille maximale de la commune conformément aux règles d'agrégation définies pour la diffusion de l'indice de qualité de l'air. Les jeux de données entre brut/AS (maille) et PREDIQA (territoire) ne sont donc pas strictement comparables, mais l'information est malgré tout intégrée dans ce rapport puisqu'il s'agit de l'information finale diffusée, et couvrant les stations considérées, avec une légère surestimation potentielle au point des stations.

Enfin, notons également que les prévisions ajustées de PREDIQA les dimanches et jours fériés sont réalisées la veille par les prévisionnistes.

- Disponibilité

Les données brutes sont disponibles avant 6h. Les données corrigées sont généralement disponibles une demi-heure après le début de la correction (dès 8h30, 10h30, 16h30). Sur l'année 2022, le taux de mise à disposition des prévisions corrigées de 10h30 est de 99,1%.

Les données PREDIQA sont produites avant midi pour le bulletin quotidien (ponctuellement un peu plus tard en période d'épisode, pour prioriser le traitement du communiqué de déclenchement).

- Observations de fond utilisées dans l'évaluation :
 - Par typologie :

Typologie	Ozone	PM10
Rurale	7	4
Périurbaine	8	4
Urbaine	16	20

- Par département :

Département	Ozone	PM10
08-Ardennes	2	2
10-Aube	2	2
51-Marne	4	3
52-Haute-Marne	1	1
54-Meurthe-Et-Moselle	2	3
55-Meuse	2	2
57-Moselle	5	6
67-Bas-Rhin	6	2
68-Haut-Rhin	3	4
88-Vosges	3	2

Soit, pour la région Grand Est : 30 en ozone, 27 en PM10

3. EVALUATION DE LA PLATEFORME DE PREVISION

Les méthodes de calcul sont définies dans l'Annexe 1 : méthode de calcul des scores.

3.1. PREVISIONS OZONE

Les scores en O₃ sont établis (en maximum horaire journalier) sur une période d'avril à septembre inclus.

3.1.1. Scores classiques (Biais, RMSE et corrélation)

3.1.1.1. Par rapport aux objectifs de qualité

Polluant	Biais (µg/m ³)	RMSE (µg/m ³)	Corrélation
O ₃	-10/+10	25	0,7
PM10	-7,5/+7,5	15	0,6

Figure 1 : Objectifs de qualité définis pour O₃ et PM₁₀ pour le biais, la RMSE et la corrélation qui s'appliquent aux performances des prévisions en régions et sur les zones disposant d'un arrêté mesure d'urgence et aux échéances J0 et J1.

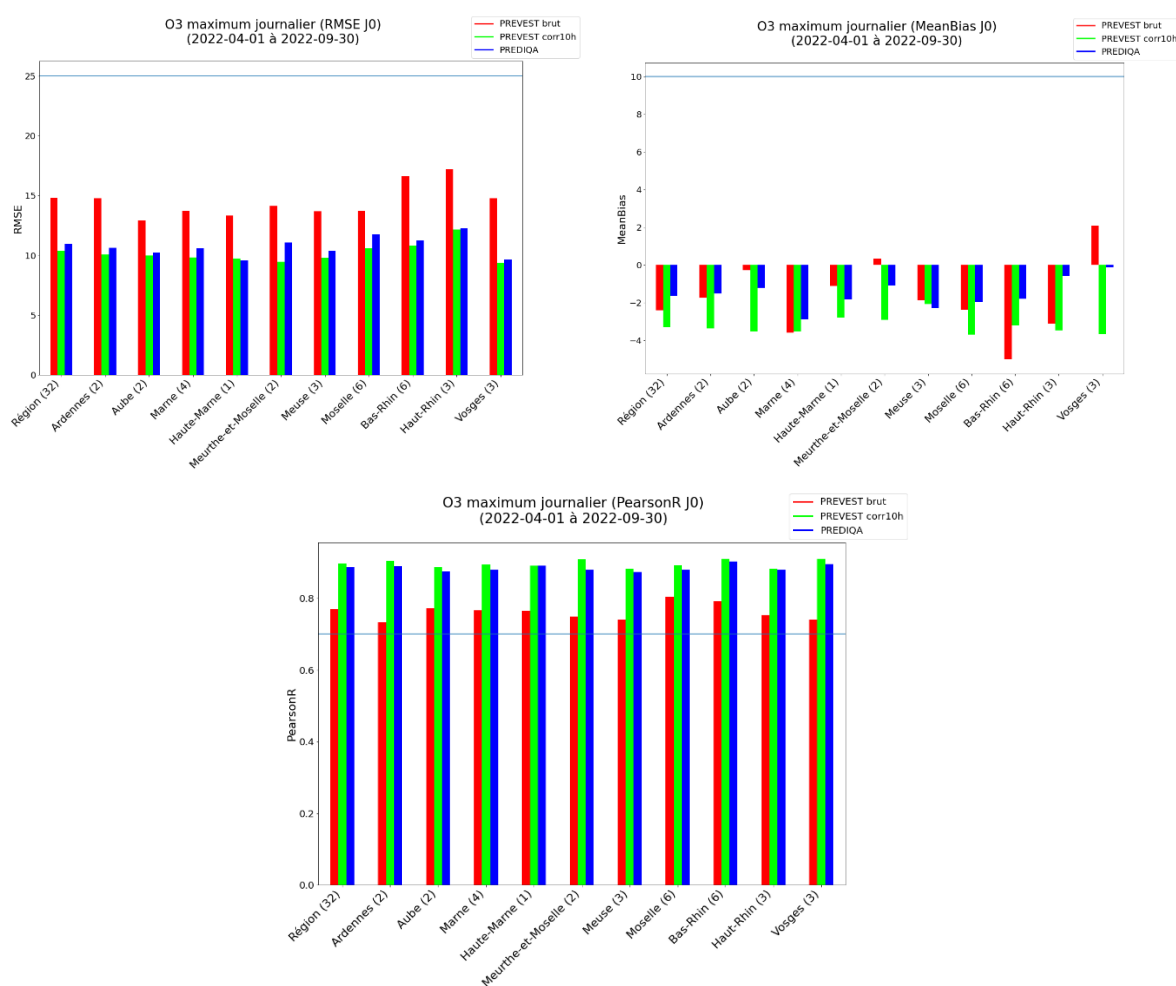


Figure 2 : Ozone - Scores de RMSE (gauche), biais (droite) et corrélation (bas) en $\mu\text{g}/\text{m}^3$ en Grand Est et par département des prévisions PREVEST brut (rouge), de l'adaptation statistique 10h (vert) et de l'expertise prévisionniste (bleu) d'avril à septembre 2022 sur l'échéance J0. Accolés entre parenthèses aux zones géographiques se trouvent le nombre de stations prises en compte pour le calcul. La ligne noire indique l'objectif de qualité.

Ces scores sont une moyenne des scores temporels de chaque station de fond. Les scores par département sont obtenus avec un filtre spatial au moment de calculer la moyenne.

Pour J0, à l'échelle de la région :

Le biais est de -2.4 pour le brut, -3.3 pour l'AS et -1.7 pour PREDIQA.

La RMSE est de 14.8 pour le brut, 10.4 pour l'AS et 11.0 pour PREDIQA.

La corrélation est de 0.77 pour le brut, 0.90 pour l'AS et 0.89 pour PREDIQA.

Tous les objectifs de qualité sont respectés pour les prévisions en ozone à J+0. On constate une amélioration importante des scores entre les prévisions brutes et les prévisions obtenues après l'adaptation statistique de 10h sur la RMSE et la corrélation pour tous les départements et la région.

On remarquera que sur ces mêmes statistiques (RMSE, corrélation), les scores entre les prévisions obtenues après l'adaptation statistique de 10h et celles issues de PREDIQA intégrant l'expertise prévisionniste sont très proches, avec un avantage pour l'AS.

Par rapport au brut, on constate une dégradation de l'adaptation statistique sur le biais. Ce dernier est davantage négatif que celui du modèle brut. PREDIQA permet toutefois de ramener le biais vers la valeur cible de 0.

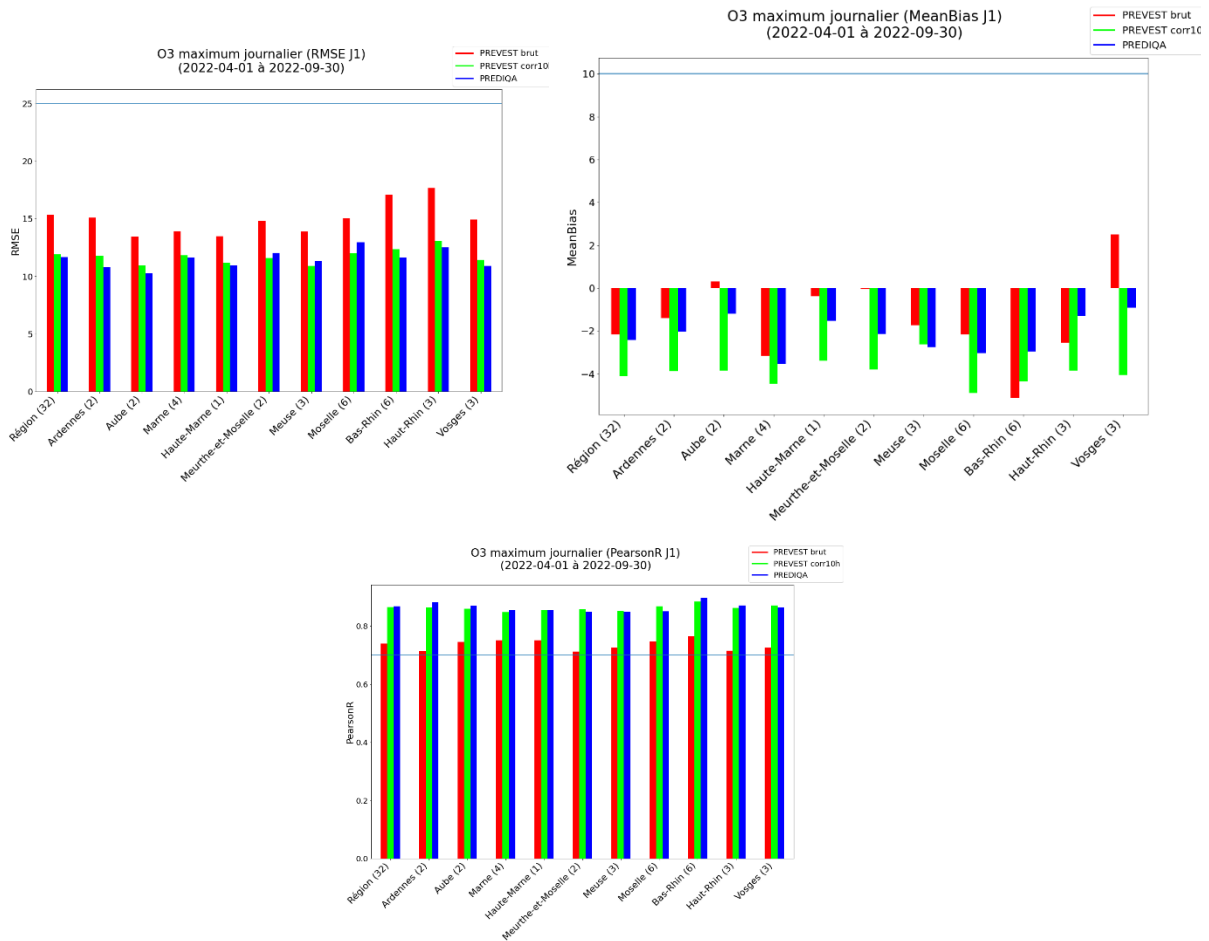


Figure 3 : Ozone - Scores de RMSE (gauche), biais (droite) et corrélation (bas) en $\mu\text{g}/\text{m}^3$ en Grand Est et par département des prévisions PREVEST brut (rouge), de l'adaptation statistique 10h (vert) et de l'expertise prévisionniste (bleu) d'avril à septembre 2022 sur l'échéance J+1. Accolés entre parenthèses aux zones géographiques se trouvent le nombre de stations prises en compte pour le calcul. La ligne noire indique l'objectif de qualité.

Pour J1, à l'échelle de la région :

Le biais est de -2.2 pour le brut, -4.1 pour l'AS et -2.4 pour PREDIQA.

La RMSE est de 15.3 pour le brut, 11.9 pour l'AS et 11.7 pour PREDIQA.

La corrélation est de 0.74 pour le brut, 0.86 pour l'AS et 0.87 pour PREDIQA.

Tous les objectifs de qualité sont respectés pour les prévisions en ozone à J+1.

Pour la RMSE et la corrélation, l’adaptation statistique de 10h permet d’améliorer sensiblement les scores par rapport aux prévisions brutes.

On remarquera que sur ces mêmes statistiques, les scores entre les prévisions obtenues après l’adaptation statistique de 10h et celles issues de PREDIQA intégrant l’expertise prévisionniste sont très proches.

Les biais constatés sur les prévisions J0 et J1 en ozone sont généralement négatifs, synonymes d’une faible sous-estimation moyenne des prévisions vis-à-vis des observations.

3.1.1.2. Par typologie

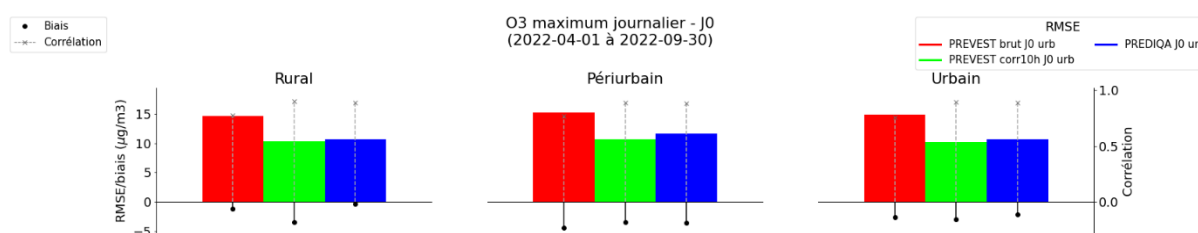


Figure 4 : Ozone - Synthèse des scores de RMSE, biais et corrélation en fonction de la typologie de station (urbaines, périurbaines, rurales) des prévisions PREVEST brut (rouge), de l’adaptation statistique 10h (vert) et de l’expertise prévisionniste (bleu) d’avril à septembre 2022 sur l’échéance J0

Ces éléments sont obtenus à partir du calcul des scores par station, puis ceux-ci sont moyennés avec un filtre sur les typologies de stations de fond (urbaines, péri-urbaines et rurales).

Les corrélations sont presque identiques entre l’AS et PREDIQA et ce pour toutes les typologies confondues. En revanche, c’est l’adaptation statistique qui s’en sort le mieux en termes de RMSE pour tous les profils. Notons que les biais sont négatifs et qu’on ne trouve pas de similitude entre les différentes typologies. Remarquons que les valeurs de ces biais moyens restent faibles puisqu’ils sont tous inférieurs à 5 µg/m³ en absolu.

3.1.1.3. Par échéance de prévision

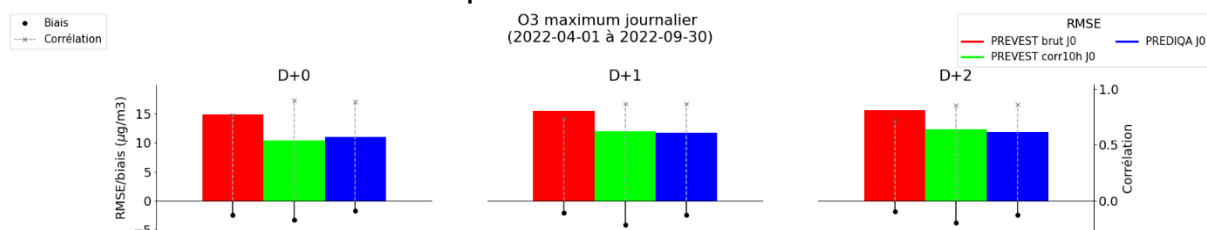
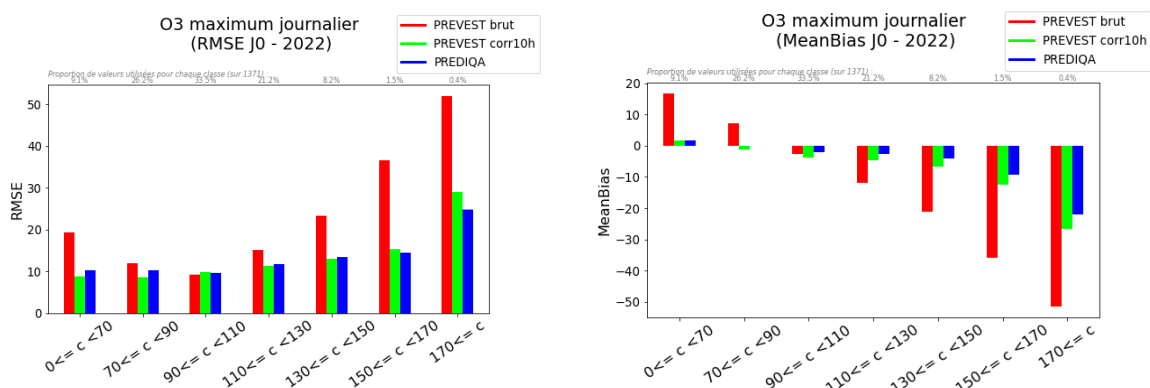


Figure 5 : Ozone - Synthèse des scores de RMSE, biais et corrélation en fonction du jour de prévision pour PREVEST brut (rouge), de l'adaptation statistique 10h (vert) et de l'expertise prévisionniste (bleu) d'avril à septembre 2022

Ces éléments sont obtenus à partir du calcul des scores par station, moyennés sur les différents horizons de prévisions (J0 à J2).

Les résultats obtenus en fonction de l'horizon de prévision sont sensiblement identiques entre les échéances J+1 et J+2. On constate, au niveau des scores de RMSE et de l'adaptation statistique de 10h et de PREDIQA, que l'échéance J0 est meilleure que les échéances J+1 et J+2. Ce résultat est logique, compte tenu que ces prévisions se réalisent en connaissance, à l'échéance J0, des données d'observation jusqu'à 10h (voire 11h pour l'expertise prévisionniste). Entre les échéances J+1 et J+2, les scores sont très proches. On notera que pour toutes les échéances, c'est l'AS qui s'écarte le plus en termes de biais bien qu'il soit faible.

3.1.1.4. Par gamme de concentration



Moins de 70	70 à 90	90 à 110	110 à 130	130 à 150	150 à 170	Plus de 170
9.1%	26.2%	33.5%	21.2%	8.2%	1.5%	0.4%

Figure 6 : Ozone - Scores de RMSE et biais par classe de concentration (en $\mu\text{g}/\text{m}^3$) des prévisions PREVEST brut (rouge), de l'adaptation statistique à 10h (vert) et de l'expertise prévisionniste (bleu) d'avril à septembre 2022 sur l'échéance J+0. Sous les figures, les taux d'observations pour chaque classe de concentrations.

Les données journalières sont groupées selon leur « classe de concentration » (par rapport aux observations). Pour chaque groupe, les scores sont ensuite calculés sur chaque station de fond puis moyennés.

L'adaptation statistique de 10h améliore l'ensemble des scores (RMSE ou biais en absolu) obtenus. La seule exception est la tranche de concentrations de $90 \mu\text{g}/\text{m}^3$ à $110 \mu\text{g}/\text{m}^3$ mais la différence est légère. Au-delà de la tranche de concentrations $110 \mu\text{g}/\text{m}^3$, les biais ressortent négatifs, sous-entendant que les différents modèles sous-estiment les max journaliers. On constate tout de même que l'AS et surtout l'expertise prévisionniste (associée à l'agrégation communale) se rapproche d'avantage des observations par rapport au modèle brut.

Notons que pour les classes de concentrations au-delà de $150 \mu\text{g}/\text{m}^3$, les données sont faiblement représentées. Même si la tendance à une sous-estimation croissante semble avoir lieu, les scores sont à interpréter avec prudence.

Remarque : Dans la partie 3.1.1.1., nous avons observé que le biais des prévisions issues de l'adaptation statistique semble moins bon que celui du modèle brut, en moyenne sur la région. Le graphique ci-dessous nous donne l'impression inverse. En regardant le détail par classe de concentrations, nous voyons en effet que l'adaptation statistique est meilleure que le modèle brut, en s'approchant davantage de la valeur cible de zéro. Le paradoxe se lève quand on constate que le brut surestime les faibles valeurs ($<90 \mu\text{g}/\text{m}^3$), qui sont nombreuses (35%), et que cela compense la forte sous-estimation des valeurs hautes, moins fréquentes. D'où en moyenne sur toute la gamme de concentration, un biais plus équilibré pour le modèle brut en ozone mais en réalité dégradé en

regardant le détail classe par classe. L'adaptation statistique ressort donc bien meilleure pour la statistique du biais.

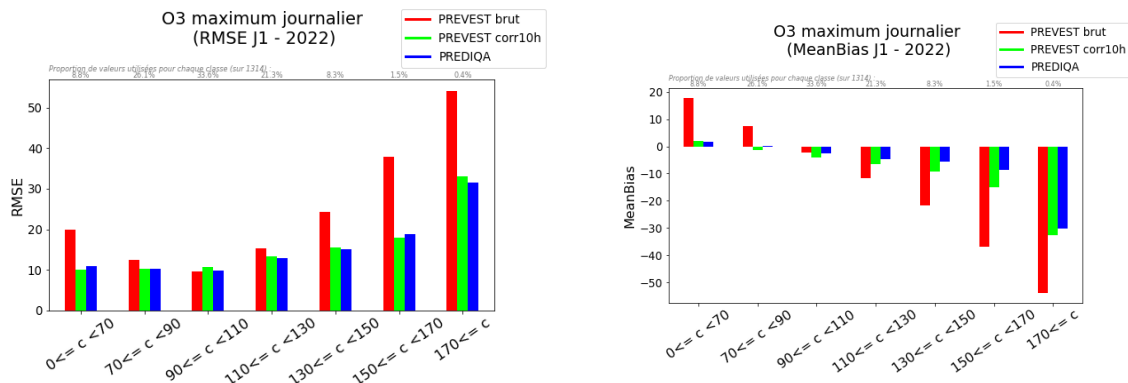


Figure 7 : Ozone - Scores de RMSE et biais par classe de concentration (en $\mu\text{g}/\text{m}^3$) des prévisions PREVEST brut (rouge), de l'adaptation statistique à 10h (vert) et de l'expertise prévisionniste (bleu) d'avril à septembre 2022 sur l'échéance J+1. Sous les figures, les taux d'observations pour chaque classe de concentrations.

Entre l'échéance J0 et J+1, les scores sont comparables pour le RMSE et le biais. On notera que la différence notable a lieu pour les classes de concentrations les plus hautes (au-delà de $130 \mu\text{g}/\text{m}^3$) où les scores se dégradent entre les deux échéances. L'AS et PREDIQA améliorent encore bien les prévisions brutes du maximum journalier.

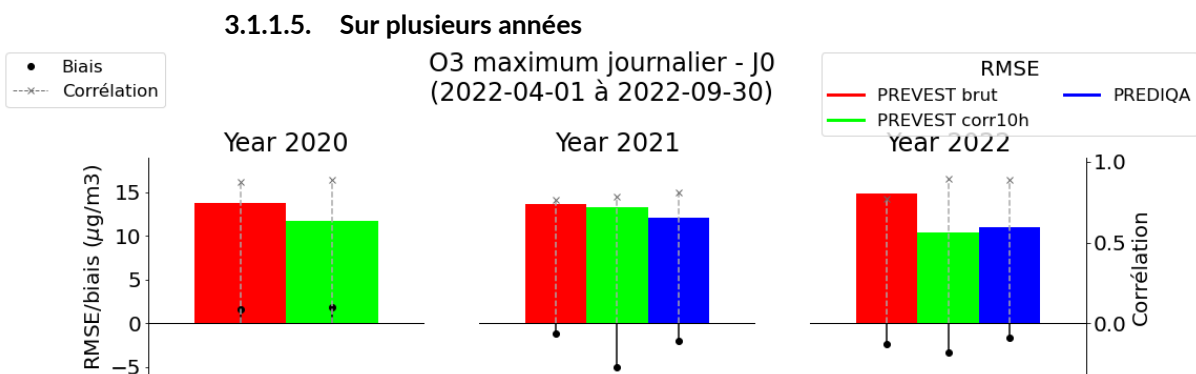


Figure 8 : Ozone - Synthèse des scores de RMSE (barres colorées), biais moyen (points noirs) et corrélation (pointillés gris) entre avril et septembre depuis 2019 pour PREVEST brut (rouge), pour son adaptation statistique de 10h (vert) et pour l'expertise prévisionniste (bleu)

Il s'agit de moyennes des scores temporels par station, pour chaque année présentée.

L'année 2022 a été simulée avec des performances plus faibles que les autres années. L'adaptation statistique de 10h en 2022 a permis toutefois de rattraper le tir en termes de RMSE et de corrélation par rapport au modèle brut. Cette marge de progression est la plus grande comparativement aux

autres années. PREDIQA et l'AS présente des performances comparables en corrélation en 2022. Si PREDIQA s'en sort mieux en termes de biais, l'AS ressort la plus performante pour la RMSE.

3.1.2. Performances sur les seuils réglementaires

3.1.2.1. Scores aux stations

3.1.2.1.1. Séries temporelles

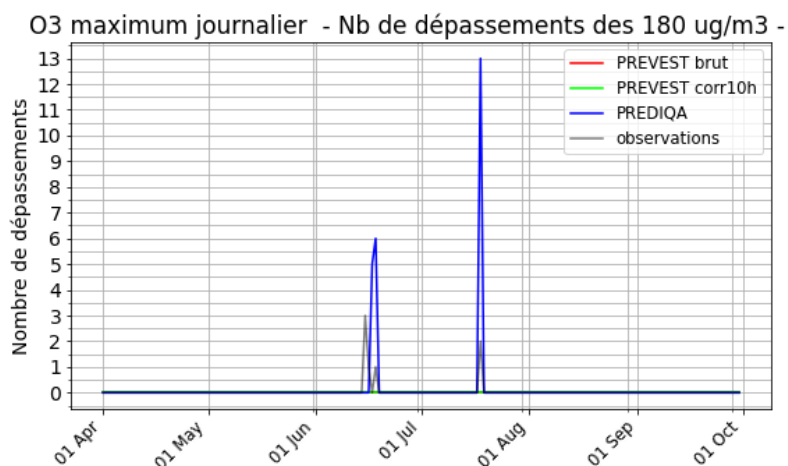


Figure 9 : Superposition des dépassements enregistrés et prévus à J0 par les différentes prévisions
Ces figures représentent le nombre de dépassements observés (en pointillés gris) et prévus (en couleur) cumulés sur toutes les stations de fond pour chaque journée de la période considérée.

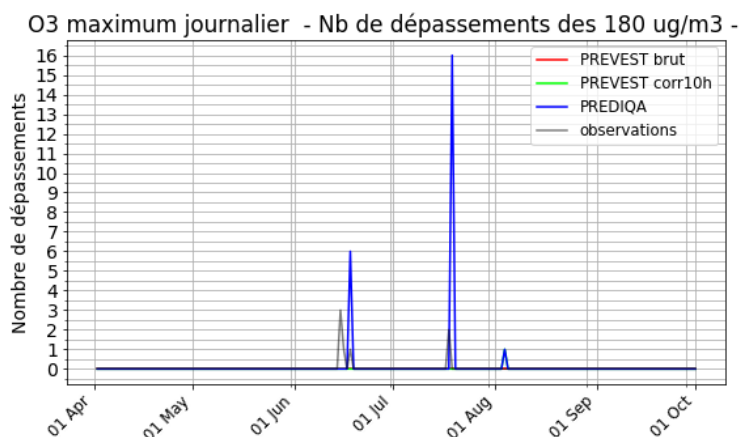


Figure 10 : Superposition des dépassements enregistrés et prévus à J1 par les différentes prévisions

L'année 2022 comporte sept dépassements des 180 $\mu\text{g}/\text{m}^3$ en ozone au niveau des observations (3 le 15 juin, 1 le 16 juin, 1 le 18 juin et 2 le 18 juillet). Le modèle brut et l'adaptation statistique n'ont pas réussi à détecter ces dépassements. A l'échéance J0, l'expertise PREDIQA a prédit des dépassements mais plus que nécessaires (24 au lieu de 7 réellement) amenant à de bonnes détections la journée du 18 juin et celle du 18 juillet. On notera toutefois des manqués comme la journée du 15 juin et du 16 juin ainsi que des fausses détections. En effet, l'expertise a prévu un dépassement le 17 juin, qui n'a pas été confirmé par les mesures (max à 175 $\mu\text{g}/\text{m}^3$).

A l'échéance J1, on retrouve le même schéma. On a de bonnes détections le 18 juin par PREDIQA (mais pas le 18 juillet). On notera que la correction de 10h et l'expertise ont vu un dépassement le 4 août, non vu par la mesure (maximum à 155 $\mu\text{g}/\text{m}^3$).

3.1.2.1.2. Eléments de la table de contingence

Les éléments pris en compte pour produire les figures de cette partie sont définis en

Annexe 2 - Eléments de table de contingence.

Les figures ci-dessous représentent les valeurs des tables de contingence (Bonne détection, Détection manquée, Fausse détection) cumulées sur les stations de fond pour la région à gauche et déclinées par département.

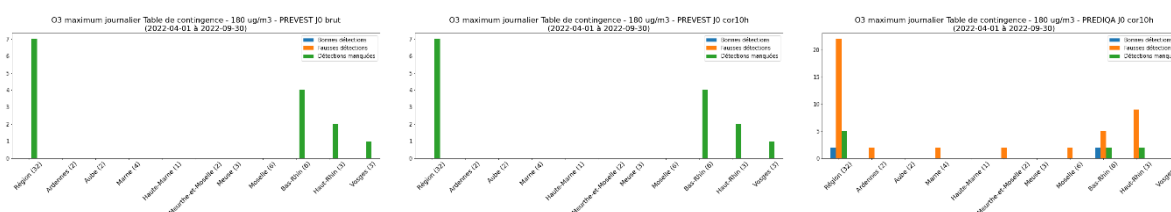


Figure 11 : Ozone - Tables de contingence sur la période d'avril à septembre 2022 à l'échéance J0 pour le seuil de 180 $\mu\text{g}/\text{m}^3$. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de PREVEST brut pour les départements du Grand Est. Milieu : équivalent pour l'adaptation statistique de 10h. Droite : expertise prévisionniste PREDIQA.

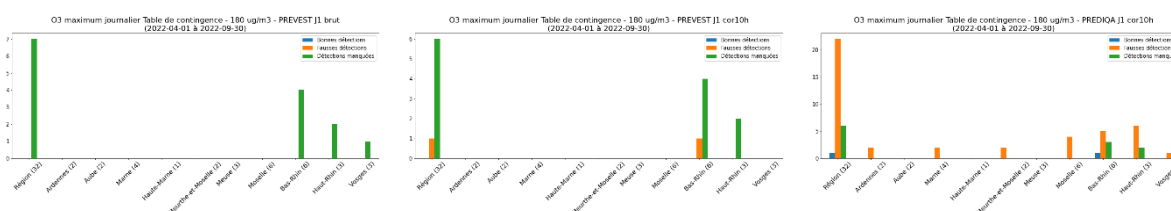


Figure 12 : Ozone - Tables de contingence sur la période d'avril à septembre 2022 à l'échéance J+1 pour le seuil de 180 $\mu\text{g}/\text{m}^3$. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de PREVEST brut pour les départements du Grand Est. Milieu : équivalent pour l'adaptation statistique de 10h. Droite : expertise prévisionniste PREDIQA.

Les figures de contingence ci-dessus font référence aux franchissements du seuil de 180 $\mu\text{g}/\text{m}^3$ observés en série temporelle (partie 3.1.2.1.1). On remarque que les franchissements concernent les départements du Bas-Rhin, du Haut-Rhin et des Vosges.

Seul PREDIQA a permis de détecter des dépassements (seulement sur le Bas-Rhin et que 2 sur 4 pour ce département). En revanche, la détection de dépassement par PREDIQA se fait au détriment de fausses détections, en nombre plus conséquent que le modèle brut et l'adaptation statistique.

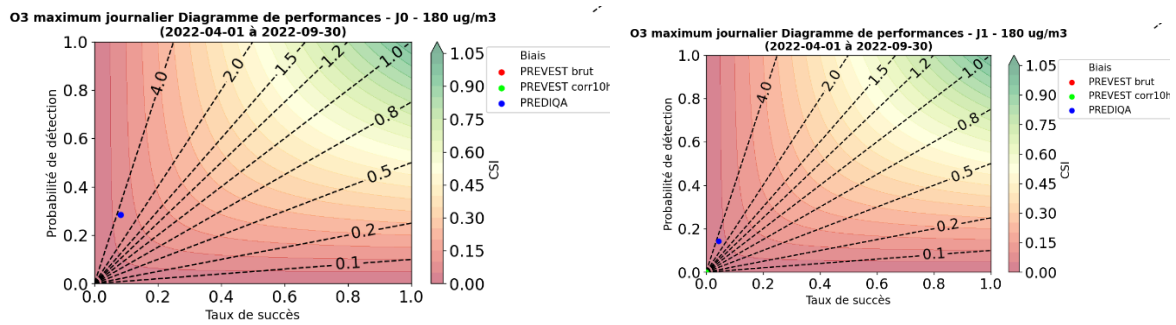


Figure 13 : Ozone - Diagrammes de performance pour PREVEST brut (rouge), l'adaptation statistique 10h (vert) et expertise prévisionniste (bleu) d'avril à septembre 2022 pour l'échéance J0 (gauche) et J1 (droite).

Les figures ci-dessus représentent 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 alertes par rapport aux alertes manquées
- Axe couleur : capacité du modèle à faire de bonnes prévisions parmi l'ensemble des dépassements (bons comme manqués) et des fausses détections. On ne considère pas les détections en dessous du seuil.

Pour plus de précisions, voir l'Annexe 2.

L'absence de point (ou sa présence en coordonnées (0 ;0)) symbolise l'absence de bonne détection par la prévision ou si la mesure n'a jamais présenté de dépassement du seuil.

3.1.2.2. Scores en lien avec les critères de l'Arrêté Mesure d'Urgence (AMU)

La référence pour calculer le dépassement des critères est le recalcul à la suite de la post-validation mensuelle (PMV) : les surfaces ou populations exposées à un dépassement recalculées en mode PVM avec intégration des mesures validées sont comparées à ce qui avait été prévu par le modèle en mode J ou J+1. Est considérée comme situation de dépassement toute journée dont la superficie départementale exposée dépasse 10 km² et dont la superficie régionale associée dépasse 100 km² d'exposition, ou lorsque la population exposée dépasse 10% (ou 50 000 habitants pour les départements des Ardennes, de l'Aube, de la Haute-Marne, de la Meuse et des Vosges) ; conformément à l'arrêté inter-préfectoral du 27 mai 2017 relatif à la gestion des épisodes de pollution atmosphérique et au déclenchement des procédures d'information-recommandation et d'alerte.

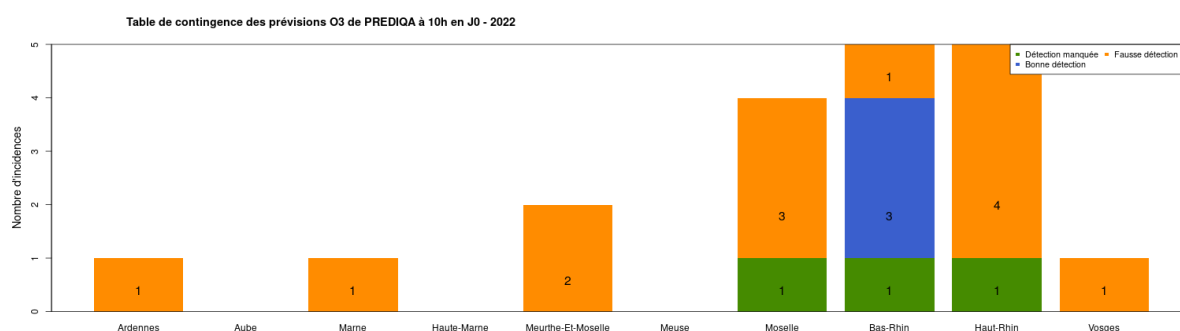
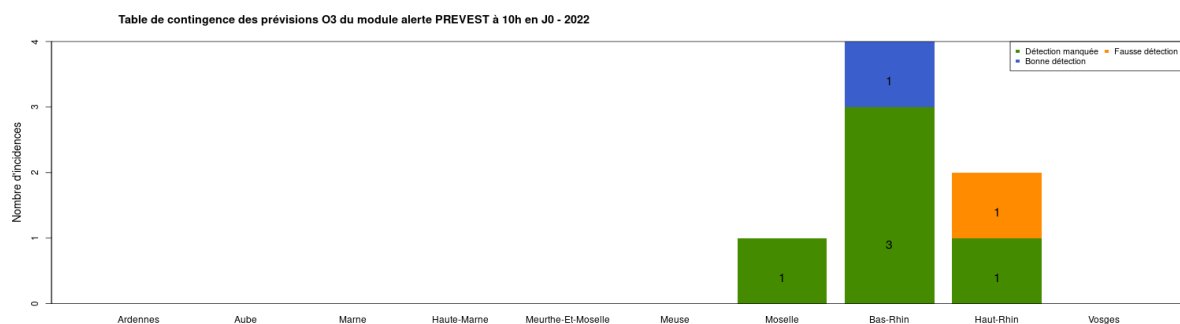


Figure 14 : Ozone - Tables de contingence des critères AMU (critères surfaciques et de population confondus) sur la période d'avril à septembre 2022 à l'échéance J0 pour l'adaptation statistique 10h (haut) et l'expertise prévisionniste (bas). Les bonnes détections sont en bleu, les fausses alertes en orange, et les alertes manquées en vert.

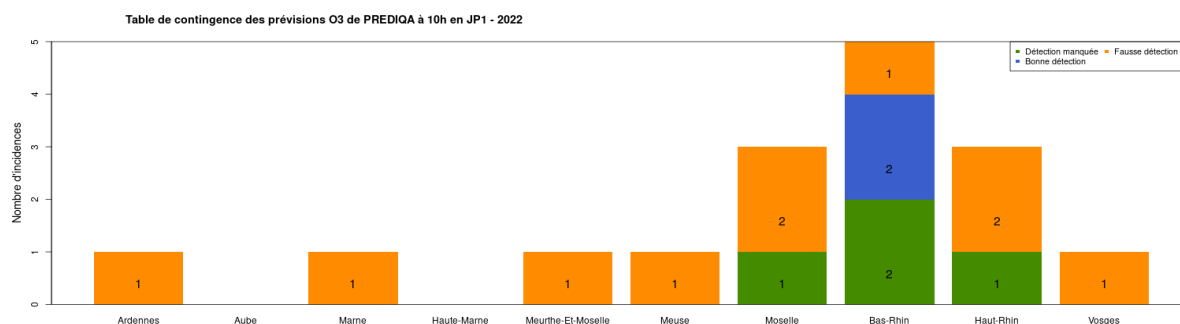
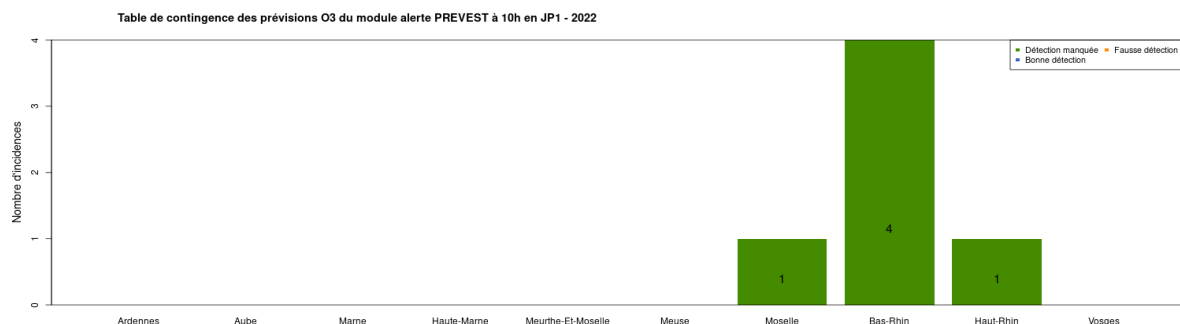


Figure 15: Ozone - Tables de contingence des critères AMU (critères surfaciques et de population confondus) sur la période d'avril à septembre 2022 à l'échéance J+1 pour l'adaptation statistique 10h (haut) et l'expertise prévisionniste (bas).

l'expertise prévisionniste (bas). Les bonnes détections sont en bleu, les fausses alertes en orange, et les alertes manquées en vert.

Les figures ci-dessus représentent les éléments de la table de contingence concernant les critères de l'arrêté mesure d'urgence en place dans les départements pour J0 et J+1.

Au regard des critères AMU (critères surfaciques et de population confondus), il y a eu des dépassements en Moselle, Bas-Rhin et Haut-Rhin en 2022 (respectivement 1, 4 et 1).

En comparant avec les parties 3.1.2.1.1 et 3.1.2.1.2, on remarquera que la Moselle a présenté un dépassement sur les critères de surface et de population alors qu'il n'y a pas eu de dépassement vu par la mesure. Cela est possible car une mesure est locale et les critères AMU concernent des surfaces et donc des résultats de modélisation.

On remarquera aussi qu'un dépassement de seuil a été constaté localement par la mesure dans les Vosges mais que cela n'a pas amené le département entier à être concerné par un dépassement sur les critères de surface et de population AMU.

Seul des dépassements sur le Bas-Rhin ont pu être détectés. PREDIQA s'en sort mieux que l'adaptation statistique pour les deux échéances.

3.2. PREVISIONS PM10

Les scores PM₁₀ sont établis (en moyenne journalière) à partir des données de l'ensemble de l'année.

3.2.1. Scores classiques (Biais, RMSE et corrélation)

3.2.1.1. Par rapport aux objectifs de qualité

Polluant	Biais ($\mu\text{g}/\text{m}^3$)	RMSE ($\mu\text{g}/\text{m}^3$)	Corrélation
O ₃	-10/+10	25	0,7
PM ₁₀	-7,5/+7,5	15	0,6

Figure 17 : Objectifs de qualité définis pour O₃ et PM₁₀ pour le biais, la RMSE et la corrélation qui s'appliquent aux performances des prévisions en régions et sur les zones disposant d'un arrêté mesure d'urgence et aux échéances J0 et J1.

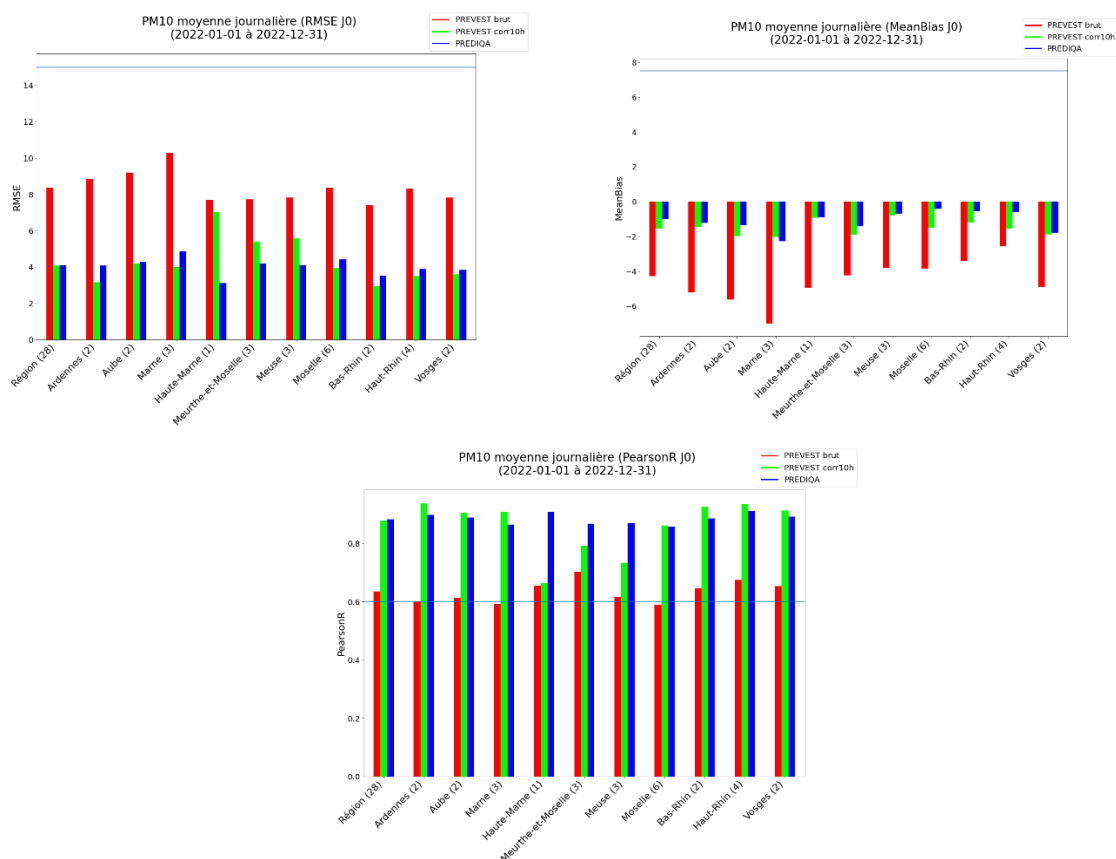


Figure 16 : PM₁₀ - Scores de RMSE (gauche), biais (droite) et corrélation (bas) en $\mu\text{g}/\text{m}^3$ en Grand Est et par département des prévisions PREVEST brut (rouge), de l'adaptation statistique 10h (vert) et de l'expertise prévisionniste (bleu) pour 2022 sur l'échéance J0. Accolés entre parenthèses aux zones géographiques se trouvent le nombre de stations prises en compte pour le calcul. La ligne noire indique l'objectif de qualité.

Ces scores sont une moyenne des scores temporels de chaque station de fond. Les scores par département sont obtenus avec un filtre spatial au moment de calculer la moyenne.

Pour J0, à l'échelle de la région :

Le biais est de -4.3 pour le brut, -1.57 pour l'AS et -1.0 pour PREDIQA.

La RMSE est de 8.4 pour le brut, 4.1 pour l'AS et 4.1 pour PREDIQA.

La corrélation est de 0.64 pour le brut, 0.88 pour l'AS et 0.88 pour PREDIQA.

Tous les objectifs de qualité sont respectés pour les prévisions en PM10 à J+0, à l'exception du modèle brut pour deux départements (la Marne et la Moselle) dont la corrélation est très légèrement en dessous de l'objectif de qualité fixé à 0.6.

On constate une amélioration importante des scores entre les prévisions brutes et les prévisions obtenues après l'adaptation statistique de 10h sur la RMSE, la corrélation et le biais pour tous les départements et la région. On remarquera que les scores entre l'adaptation statistique de 10h et ceux issus de PREDIQA sont très proches sur la région. Des disparités existent toutefois en regardant à l'échelle plus petite des départements. En ce qui concerne le biais, l'adaptation statistique et surtout PREDIQA (expertise prévisionniste couplée à l'agrégation à l'échelle communale) permettent de ramener le biais vers la valeur cible de 0.

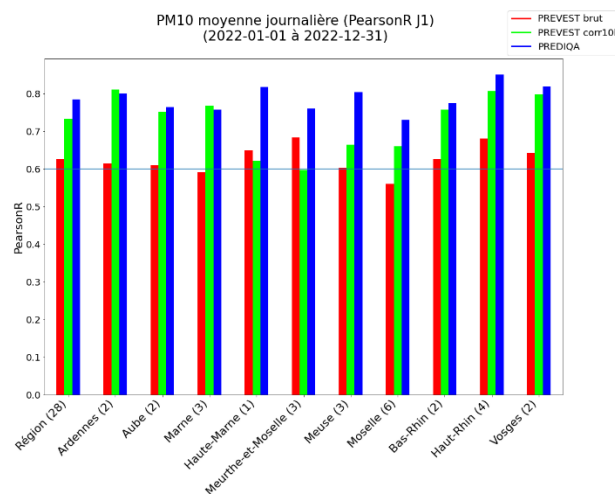
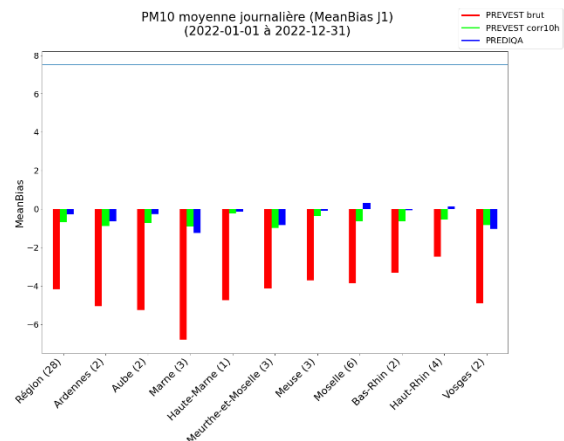
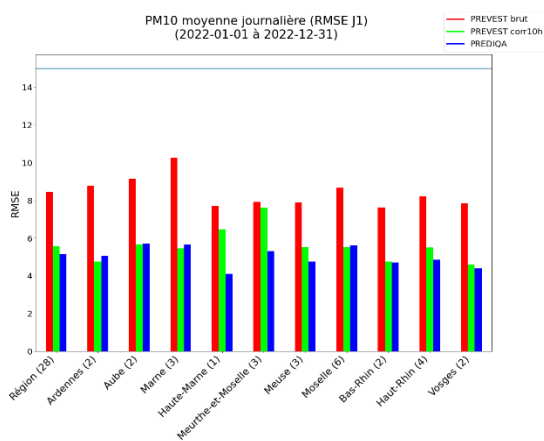


Figure 18 : PM10 - Scores de RMSE (gauche), biais (droite) et corrélation (bas) en $\mu\text{g}/\text{m}^3$ en Grand Est et par département des prévisions PREVEST brut (rouge), de l'adaptation statistique 10h (vert) et de l'expertise prévisionniste (bleu) pour 2022 sur l'échéance J+1. Accolés entre parenthèses aux zones géographiques se trouvent le nombre de stations prises en compte pour le calcul. La ligne noire indique l'objectif de qualité.

Pour J1, à l'échelle de la région :

Le biais est de -4.2 pour le brut, -0.7 pour l'AS et -0.3 pour PREDIQA.

La RMSE est de 8.4 pour le brut, 5.6 pour l'AS et 5.1 pour PREDIQA.

La corrélation est de 0.63 pour le brut, 0.73 pour l'AS et 0.78 pour PREDIQA.

Comme en J0, à l'échéance J+1, tous les objectifs de qualité sont respectés pour les prévisions en PM10 (à l'exception du modèle brut pour la Marne et la Moselle dont la corrélation est inférieure à l'objectif de qualité fixé à 0.6).

En ce qui concerne la RMSE, elle est très stable entre les deux échéances pour le modèle brut. Elle baisse un peu pour l'AS de 10h et pour PREDIQA entre J0 et J+1.

En J+1, PREDIQA ressort globalement meilleure que l'AS de 10h pour la RMSE et la corrélation.

On notera que le biais, bien que faible, a diminué en J+1 par rapport à J0 pour l'adaptation de 10 h et pour PREDIQA.

3.2.1.2. Par typologie

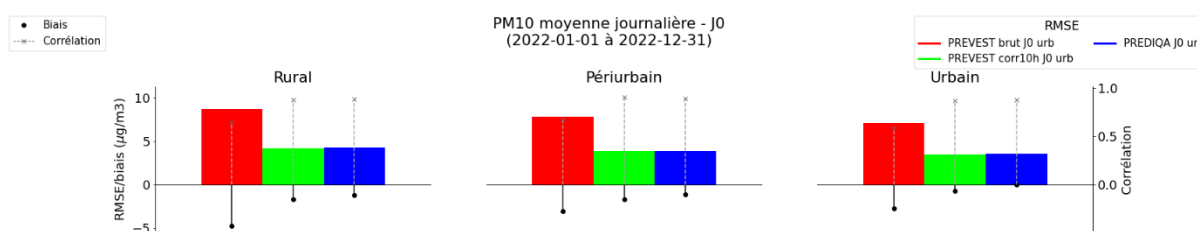


Figure 19 : PM10 - Synthèse des scores de RMSE, biais et corrélation en fonction de la typologie de station (urbaines, périurbaines, rurales) des prévisions PREVEST brut (rouge), de l'adaptation statistique 10h (vert) et de l'expertise prévisionniste (bleu) pour 2022 sur l'échéance J0

Ces éléments sont obtenus à partir du calcul des scores par station, moyennés avec un filtre sur les typologies de stations de fond (urbaines, péri-urbaines et rurales).

La RMSE et la corrélation sont presque identiques entre l'AS et PREDIQA et ce pour toutes les typologies confondues. Les biais sont négatifs. Ils se rapprochent davantage de la valeur cible pour les stations de typologie urbaine et pour PREDIQA.

3.2.1.3. Par échéance de prévision

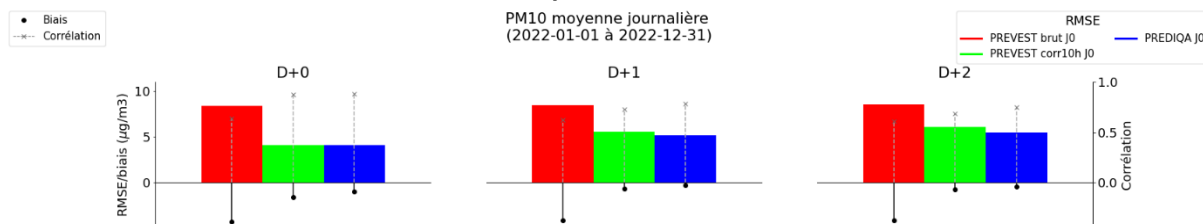
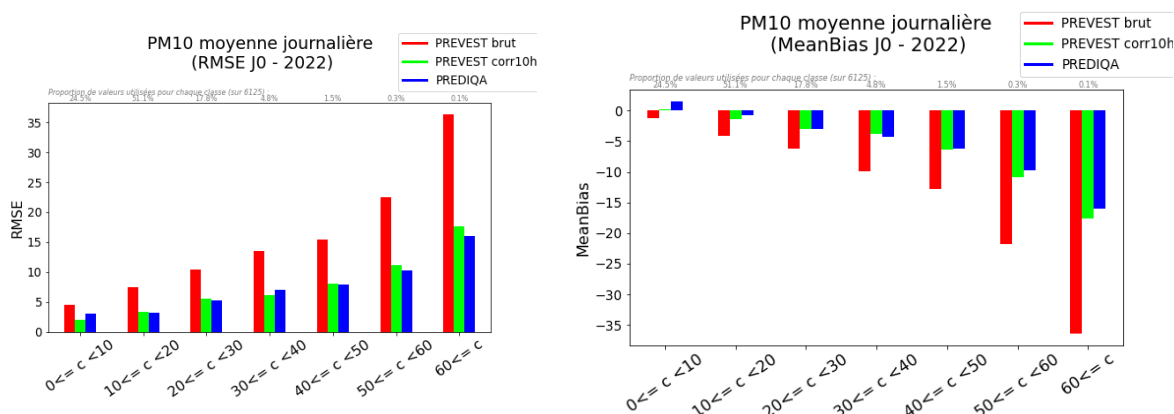


Figure 20 : PM10 - Synthèse des scores de RMSE, biais et corrélation en fonction du jour de prévision pour PREVEST brut (rouge), de l'adaptation statistique 10h (vert) et de l'expertise prévisionniste (bleu) pour 2022

Ces éléments sont obtenus à partir du calcul des scores par station, moyennés sur les différents horizons de prévisions (J0 à J2).

Les scores des prévisions brutes PM10 sont sensiblement proches quel que soit l'horizon de prévision. En revanche, la RMSE et la corrélation obtenues pour l'adaptation statistique se dégradent à mesure que l'échéance est lointaine. Ce résultat est logique, compte tenu que l'adaptation statistique de 10h apporte à l'échéance J0 des données d'observation jusqu'à 10h. L'expertise prévisionniste n'apporte pas de gain notable en J0 par rapport à l'adaptation statistique de 10h. Ses performances ressortent meilleures seulement en J+1 et J+2. On notera que le biais est légèrement plus faible à l'échéance J0 par rapport aux autres. Les résultats obtenus en fonction de l'horizon de prévision sont sensiblement identiques entre les échéances J+1 et J+2.

3.2.1.4. Par gamme de concentration



Moins de 10	10 à 20	20 à 30	30 à 40	40 à 50	50 à 60	Plus de 60
24.5%	51.1%	17.8%	4.8%	1.5%	0,3%	0,1%

Figure 21 : PM10 - Scores de RMSE et biais par classe de concentration (en $\mu\text{g}/\text{m}^3$) des prévisions PREVEST brut (rouge), de l'adaptation statistique à 10h (vert) et de l'expertise prévisionniste (bleu) pour 2022 sur l'échéance J+0. Sous les figures, les taux d'observations pour chaque classe de concentrations.

Les données journalières sont groupées selon leur « classe de concentration » (par rapport aux observations). Pour chaque groupe, les scores sont ensuite calculés sur chaque station de fond puis moyennés.

L'adaptation statistique de 10h améliore l'ensemble des scores (RMSE et biais). L'expertise prévisionniste n'apporte pas spécialement de gain supplémentaire sur les performances, sauf sur les gammes de concentrations élevées.

Les biais sont de plus en plus négatifs (sous-estimation du modèle) à mesure que les niveaux augmentent ; les valeurs les plus élevées et les moins fréquemment observées, sont logiquement celles qui sont le plus difficilement reproduites par les prévisions.

On constate tout de même que l'AS et surtout l'expertise prévisionniste (associée à l'agrégation communale) se rapprochent d'avantage des observations par rapport au modèle brut.

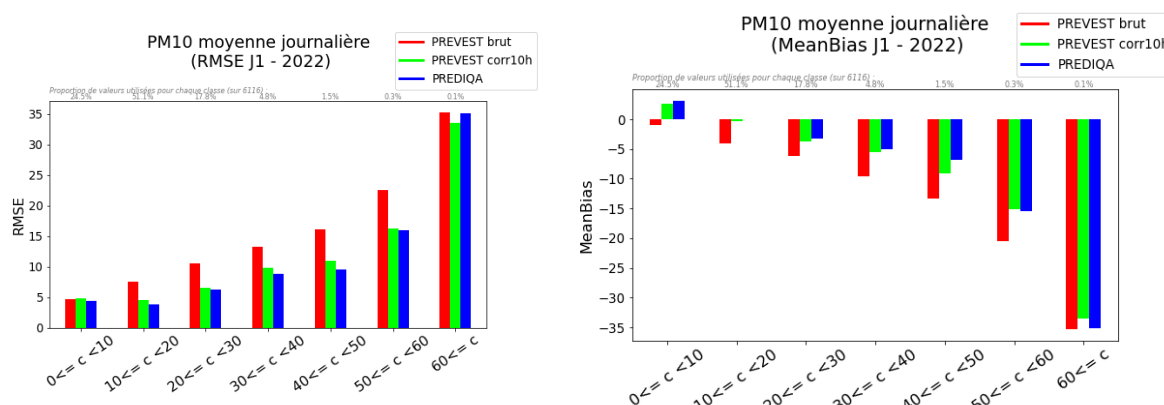


Figure 22 : PM10 - Scores de RMSE et biais par classe de concentration (en $\mu\text{g}/\text{m}^3$) des prévisions PREVEST brut (rouge), de l'adaptation statistique à 10h (vert) et de l'expertise prévisionniste (bleu) pour 2022 sur l'échéance J+1. Sous les figures, les taux d'observations pour chaque classe de concentrations.

L'adaptation statistique de 10h améliore l'ensemble des scores (RMSE et biais) pour les gammes de concentrations supérieures à $10 \mu\text{g}/\text{m}^3$. Pour les concentrations inférieures, le biais de l'AS est positif, indiquant une surestimation. En revanche, au-delà du seuil de $10 \mu\text{g}/\text{m}^3$, tous les modèles présentent une sous-estimation.

A l'échéance J+1, les indicateurs statistiques entre la prévision brute et l'adaptation statistique de 10h tendent à se rapprocher, synonymes d'un affaiblissement de l'efficacité de la correction à un horizon plus lointain que J+0.

Notons que pour les classes de concentrations au-delà de $50 \mu\text{g}/\text{m}^3$, les données sont faiblement représentées. Même si la tendance à une sous-estimation croissante semble avoir lieu, les scores sont à interpréter avec prudence.

3.2.1.5. Sur plusieurs années

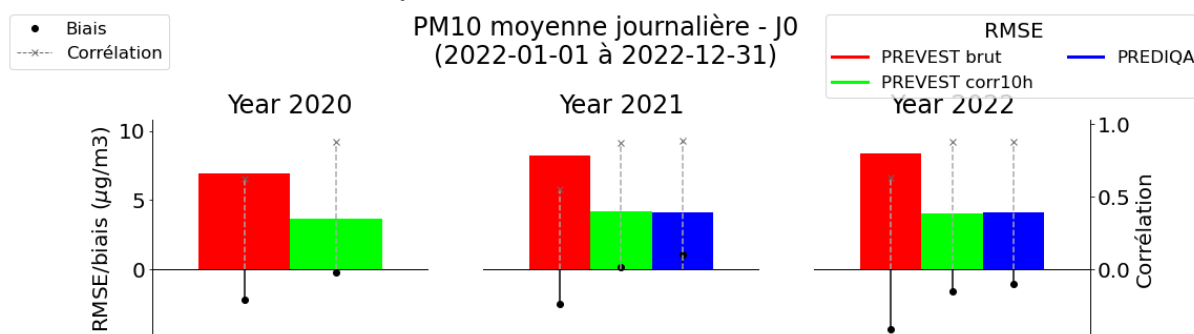


Figure 23 : PM10 - Synthèse des scores de RMSE (barres colorées), biais moyen (points noirs) et corrélation (pointillés gris) depuis 2020 pour PREVEST brut (rouge), pour son adaptation statistique de 10h (vert) et pour l'expertise prévisionniste (bleu)

Il s'agit de moyennes des scores temporels par station, pour chaque année présentée.

L'année 2022 a été simulée avec des performances comparables à l'année 2021 en RMSE et corrélation mais avec une dégradation du biais. L'adaptation statistique de 10h est profitable sur l'ensemble des indicateurs (biais, RMSE, corrélation) en ramenant les statistiques vers leur optimum. L'expertise prévisionniste présente des performances tout à fait comparables à l'adaptation statistique, avec un léger avantage sur la statistique du biais. On remarque une certaine stabilité des scores entre les deux dernières années.

3.2.2. Performances sur les seuils réglementaires

3.2.2.1. Scores aux stations

3.2.2.1.1. Séries temporelles

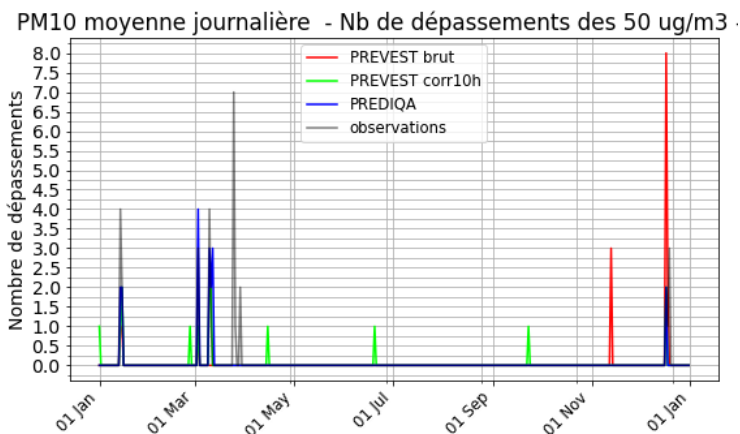


Figure 24 : Superposition des dépassements enregistrés et prévus à J0 par les différentes prévisions

Cette figure représente le nombre de dépassements observés (en pointillés gris) et prévus (en couleur) cumulés sur toutes les stations de fond pour chaque journée de la période considérée.

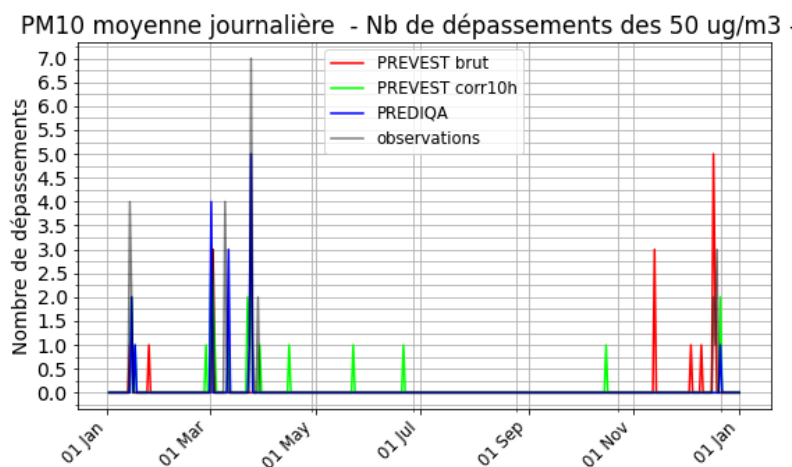


Figure 25 : superposition des dépassements enregistrés et prévus à J1 par les différentes prévisions

L'année 2022 comporte 30 dépassements du seuil de $50 \mu\text{g}/\text{m}^3$ en PM10 par la mesure. Ces dépassements concernent 11 jours de l'année. Les jours où il y en a eu le plus sont le 25 mars (7 dépassements), le 14 janvier (4 dépassements), le 10 mars (4 dépassements), le 3 mars (3 dépassements) et le 19 décembre (3 dépassements). L'adaptation statistique a permis de détecter une partie des dépassements du début d'année et de fin d'année 2022. Toutefois, elle a aussi conduit à de faux dépassements en milieu d'année. Le modèle brut a fait de nombreux faux dépassements en fin d'année. Notons qu'il a quand même fait de bonnes prévisions en J0, par exemple le 17 décembre, en détectant les deux dépassements (mais en en prévoyant huit cette même journée, conduisant donc

aussi à six faux dépassements). Sur les 11 jours où il y a eu des dépassements, il y a eu 8 jours où au moins un des modèles (brut, AS de 10h ou PREDIQA) a détecté un dépassement en J0 et en J+1. PREDIQA se révèle performant en détectant des dépassements tout en limitant les faux. Le résumé détaillé des statistiques se retrouvent dans la partie suivante 3.2.2.1.2.

3.2.2.1.2. Eléments de la table de contingence

Les éléments pris en compte pour produire les figures de cette partie sont définis en

Annexe 2 - Eléments de table de contingence

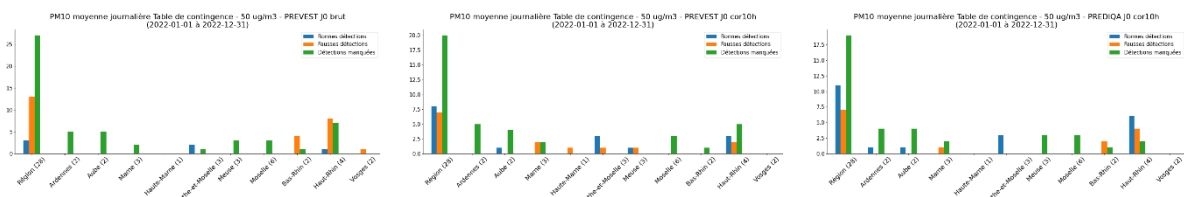


Figure 26 : PM10 - Tables de contingence sur la période d'avril à septembre 2022 à l'échéance J0 pour le seuil de 50 µg/m³. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de PREVEST brut pour les départements du Grand Est. Milieu : équivalent pour l'adaptation statistique de 10h. Droite : expertise prévisionniste PREDIQA.

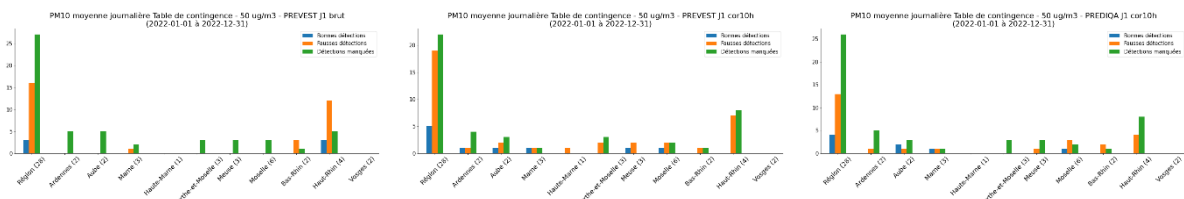


Figure 27 : PM10 - Tables de contingence sur la période d'avril à septembre 2022 à l'échéance J+1 pour le seuil de 50 µg/m³. Les bonnes détections sont en bleu, les fausses détections en orange, et les détections manquées en vert. Gauche : Tables de contingence de PREVEST brut pour les départements du Grand Est. Milieu : équivalent pour l'adaptation statistique de 10h. Droite : expertise prévisionniste PREDIQA.

Les tables de contingence ci-dessus révèlent que l'issue principale, sur les journées avec une moyenne journalière en dépassement des 50 µg/m³, est la détection manquée (vert) en sortie de PREDIQA pour l'échéance J+0 (11 bonnes détections et 19 manquées). Pour les échéances J0 et J+1, le nombre de bonnes détections est plus élevé avec PREDIQA par rapport à l'adaptation statistique. Cette dernière dépasse le modèle brut pour ces deux échéances, amenant un gain de performance avec les différentes corrections apportées.

L'expertise prévisionniste associée à l'agrégation communale remplit son rôle correcteur en permettant une augmentation de la part en bonne détection (bleu) pour les prévisions J+0 et J+1, et ce, sans occasionner de fausses détections supplémentaires par rapport aux autres modèles.

Notons qu'entre les deux échéances J0 et J+1, pour tous les modèles, le nombre de bonnes détections diminue tandis que les fausses détections augmentent. Cela traduit une dégradation des performances avec l'horizon de prévision.

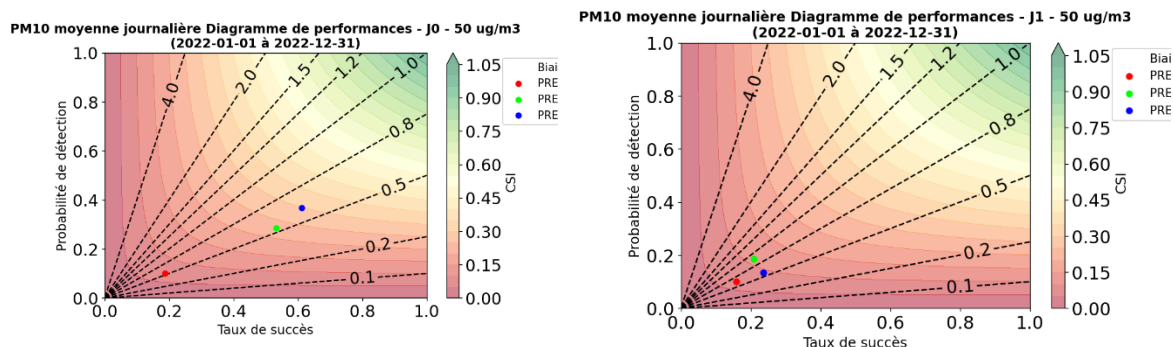


Figure 28 : PM10 - Diagrammes de performance pour PREVEST brut (rouge), l'adaptation statistique 10h (vert) et l'expertise prévisionniste (bleu) pour 2022 pour l'échéance J0 (gauche) et J1 (droite).

Les figures ci-dessus représentent 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 alertes par rapport aux alertes manquées
- Axe couleur : capacité du modèle à faire de bonnes prévisions parmi l'ensemble des dépassements (bons comme manqués) et des fausses détections. On ne considère pas les détections en dessous du seuil.

Pour plus de précisions, voir l'Annexe 2.

L'absence de point (ou sa présence en coordonnées (0 ;0)) symbolise l'absence de bonne détection par la prévision ou si la mesure n'a jamais présenté de dépassement du seuil.

3.2.2.2. Scores en lien avec les critères AMU

La référence pour calculer le dépassement des critères est le recalcul à la suite de la post-validation mensuelle (PMV) : les surfaces ou populations exposées à un dépassement recalculées en mode PVM avec intégration des mesures validées sont comparées à ce qui avait été prévu par le modèle en mode J ou J+1. Est considérée comme situation de dépassement toute journée dont la superficie départementale exposée dépasse 10 km² et dont la superficie régionale associée dépasse 100 km² d'exposition, ou lorsque la population exposée dépasse 10% (ou 50 000 habitants pour les départements des Ardennes, de l'Aube, de la Haute-Marne, de la Meuse et des Vosges) ; conformément à l'arrêté inter-préfectoral du 27 mai 2017 relatif à la gestion des épisodes de pollution atmosphérique et au déclenchement des procédures d'information-recommandation et d'alerte.

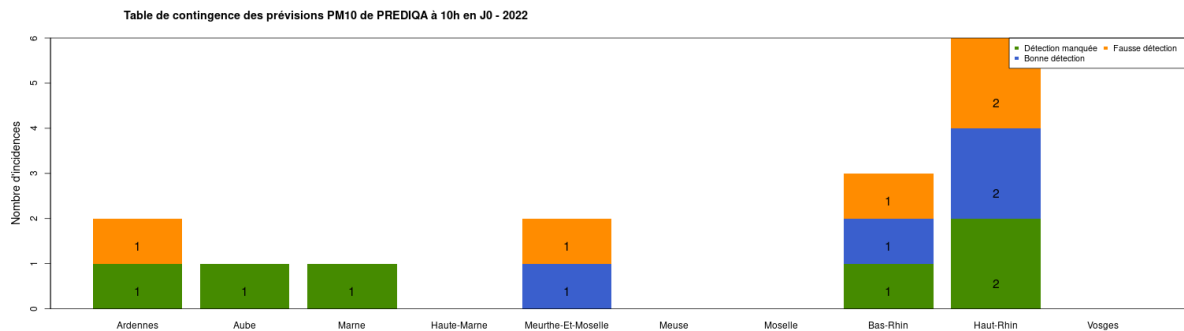
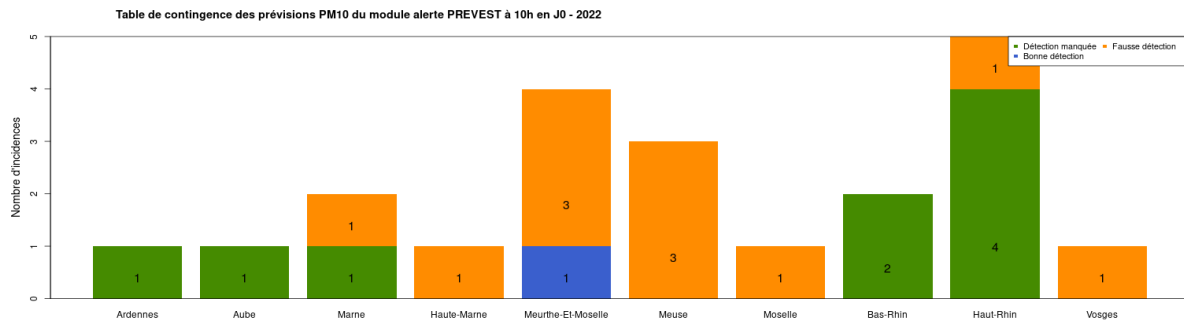


Figure 29 : PM10 - Tables de contingence des critères AMU (critères surfaciques et de population confondus) sur 2022 à l'échéance J0 pour l'adaptation statistique 10h (haut) et l'expertise prévisionniste (bas). Les bonnes détections sont en bleu, les fausses alertes en orange, et les alertes manquées en vert.

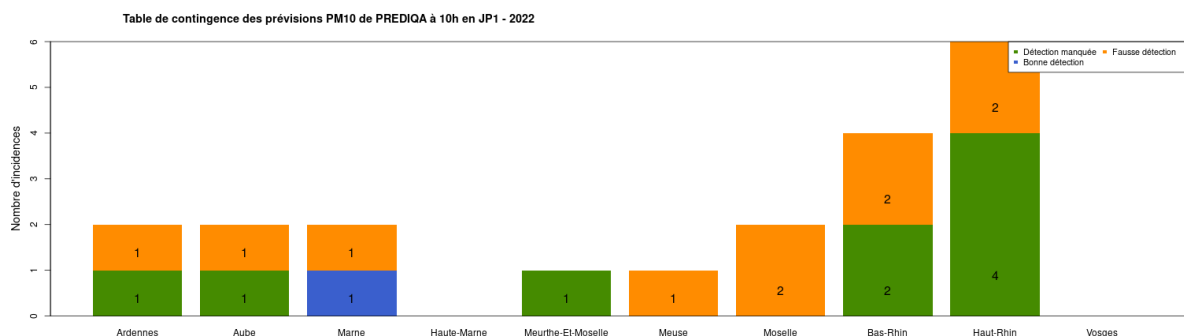
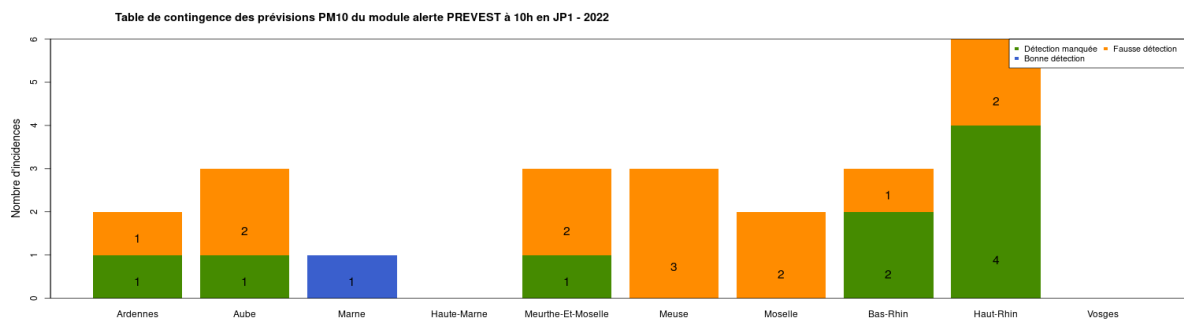


Figure 30 : PM10 - Tables de contingence des critères AMU (critères surfaciques et de population confondus) sur 2022 à l'échéance J+1 pour l'adaptation statistique 10h (haut) et l'expertise prévisionniste (bas). Les bonnes détections sont en bleu, les fausses alertes en orange, et les alertes manquées en vert.

Les figures ci-dessus représentent les éléments de la table de contingence concernant les critères de l'arrêté mesure d'urgence en place dans les départements pour J0 et J+1.

Au regard des critères AMU (critères surfaciques et de population confondus), il y a eu des dépassements dans tous les départements à l'exception de la Haute-Marne, la Meuse, la Moselle et les Vosges. C'est dans le Haut-Rhin qu'il y a eu le plus de dépassements, en s'élevant au nombre de 4. Sur l'ensemble de la région, on en dénombre 10.

Sur l'année 2022, ce sont majoritairement les fausses détections (en orange) qui ressortent le plus. Cela est plus marqué pour l'adaptation statistique de 10h que pour l'expertise prévisionniste PREDIQA.

En comparant les parties précédentes, on remarquera que la Moselle a présenté un dépassement vu par la mesure mais pas selon les critères de surface et de population. Cela est possible car une mesure est locale tandis que les critères AMU concernent des dépassements surfaciques.

A l'échéance J0, alors que le nombre de fausses alertes aux stations (partie 3.2.2.1.2) est le même pour l'expertise prévisionniste que pour l'adaptation statistique (7), ils ont généré moins de fausses alertes, au regard des critères AMU, pour PREDIQA par rapport à l'adaptation statistique.

A l'échéance J+1, seul le dépassement sur la Marne a pu être détectés par l'adaptation statistique et PREDIQA sur les dix de la région.

PREDIQA s'en sort mieux que l'adaptation statistique pour les deux échéances en prévoyant davantage les dépassements selon les critères AMU tout en limitant globalement les fausses détections.

ANNEXE 1 - METHODE DE CALCUL DES SCORES

Taux de disponibilité des données fixé à 75%. Cela signifie qu'il faut que sur la période d'évaluation, 75% des couples journaliers (ou horaires) mesures-concentrations prévues soient disponibles pour qu'une station soit prise en compte dans le calcul des scores.

L'ensemble des scores est calculé aux stations puis fait l'objet d'une moyenne selon les critères de la représentation (par zones géographiques, par typologies, par année, ...)

- **Biais** : Estimateur des écarts entre prévisions et observations :

$$\frac{1}{N} \cdot \sum_{i=1}^N (P_i - O_i)$$

N est le nombre de stations de mesure, P_i la valeur de prévision à la station i et O_i la valeur observée à la station.

Si le biais est établi pour le maximum journalier, alors le maximum journalier des 2 valeurs doit être pris, s'il s'agit d'un biais pour la moyenne journalière, alors la moyenne journalière des 2 valeurs doit être prise. Valeurs exprimées en $\mu\text{g}/\text{m}^3$

- **RMSE** : Root Mean Square Error (RMSE) donne une information sur la qualité du modèle à reproduire les niveaux de concentrations observées. Elle doit être la plus faible possible.

$$\sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (P_i - O_i)^2}$$

N est le nombre de stations, P_i correspond à la prévision à la station i et O_i l'observation à la station i . Valeurs exprimées en $\mu\text{g}/\text{m}^3$.

- **Corrélation** : Corrélation r informe sur la cohérence des variabilités temporelle ou spatiale des concentrations observées et prévues. Plus celle-ci est proche de 1 plus la prévision est en mesure de reproduire les bonnes variations des concentrations observées.

$$r = \frac{\text{cov}(P, O)}{\sqrt{\text{var}(P)} \cdot \sqrt{\text{var}(O)}} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

N est le nombre de stations, P_i correspond à la prévision à la station i et O_i l'observation à la station i . Valeurs sans dimension. (cov correspond à la covariance et var à la variance).

ANNEXE 2 - ELEMENTS DE TABLE DE CONTINGENCE

En préambule, un dépassement est comptabilisé lorsqu'une valeur observée ou prévue excède strictement le seuil de concentration considéré.

La table de contingence des prévisions d'un modèle se présente comme suit :

	<i>Obs > seuil</i>	<i>Obs < seuil</i>	<i>Total</i>
<i>Prévi > seuil</i>	Bonne détection (a)	Fausse détection (b)	a+b
<i>Prévi < seuil</i>	Détection manquée (c)	Bonne non-détection (d)	c+d
<i>Total</i>	a+c	b+d	a+b+c+d

Bonne détection (a) : bonne prévision de dépassement du seuil

Bonne non-détection (d): bonne prévision de non-dépassement du seuil

Détection manquée (c) : dépassement de seuil observé en station mais non prévu par le modèle

Fausse détection (b) : dépassement de seuil prévu par le modèle mais non observé en station

A partir de ces paramètres, il est possible de calculer plusieurs indicateurs permettant de mieux formaliser les performances d'un modèle vis-à-vis de la détection des dépassements.

POD = « Probability of Detection », ou probabilité de détection

C'est le taux de bonnes prévisions de dépassements de seuil par rapport au nombre total de dépassements observés.

$$POD = \frac{a}{a + c}$$

FAR = « False Alarm Ratio », ou taux de détection manquée

C'est le taux de mauvaises prévisions de dépassements de seuil par rapport au nombre total de dépassements prévus par le modèle.

$$FAR = \frac{b}{a + b}$$

SR = « Success Ratio », ou taux de succès

C'est le taux de bonnes prévisions de dépassements de seuil par rapport au nombre total de dépassements prévus par le modèle.

$$SR = \frac{a}{a + b} = 1 - FAR$$

FB = « Frequency Bias », ou biais de prévision

C'est le biais de prévision des dépassements de seuils. S'il est inférieur à 1, la prévision a tendance à faire plus d'alertes manquées que de détections manquées.

$$FB = \frac{a + b}{a + c}$$

CSI = « Critical Success Index » ou « Threat Score »

C'est le taux de bonnes prévisions de dépassements de seuil par rapport au nombre total d'évènements prédits et manqués.

$$CSI = \frac{a}{a + b + c}$$



AtMO

GRAND EST

Metz - Nancy - Reims - Strasbourg

Air • Climat • Energie • Santé

Espace Européen de l'Entreprise - 5 rue de Madrid - 67300 Schiltigheim

Tél : 03 69 24 73 73 - contact@atmo-grandest.eu

Siret 822 734 307 000 17 - APE 7120 B

Association agréée de surveillance de la qualité de l'air