

Directives sur le contrôle et l'assurance qualité des données d'observation en surface à des fins climatologiques

Édition 2021

TEMPS CLIMAT EAU



ORGANISATION
MÉTÉOROLOGIQUE
MONDIALE

OMM-N° 1269

Directives sur le contrôle et l'assurance qualité des données d'observation en surface à des fins climatologiques

Édition 2021



ORGANISATION
MÉTÉOROLOGIQUE
MONDIALE

OMM-N° 1269

NOTE DE L'ÉDITEUR

La base de données terminologique de l'OMM, METEOTERM, peut être consultée à l'adresse <https://public.wmo.int/fr/meteoterm>.

Il convient d'informer le lecteur que lorsqu'il copie un hyperlien en le sélectionnant dans le texte, des espaces peuvent apparaître après <http://>, <https://>, <ftp://>, <mailto:>, et après les barres obliques (/), les tirets (-), les points (.) et les séquences de caractères (lettres et chiffres). Il faut supprimer ces espaces de l'URL ainsi recopiée. L'URL correcte apparaît lorsque l'on place le curseur sur le lien. On peut aussi cliquer sur le lien et copier l'adresse qui s'affiche dans le ruban du navigateur.

OMM-N° 1269

© **Organisation météorologique mondiale, 2021**

L'OMM se réserve le droit de publication en version imprimée ou électronique ou sous toute autre forme et dans n'importe quelle langue. De courts extraits des publications de l'OMM peuvent être reproduits sans autorisation, pour autant que la source complète soit clairement indiquée. La correspondance relative au contenu rédactionnel et les demandes de publication, reproduction ou traduction partielle ou totale de la présente publication doivent être adressées au:

Président du Comité des publications
Organisation météorologique mondiale (OMM)
7 bis, avenue de la Paix
Case postale 2300
CH-1211 Genève 2, Suisse

Tél.: +41 (0) 22 730 84 03
Fax: +41 (0) 22 730 81 17
Courriel: publications@wmo.int

ISBN 978-92-63-21269-6

NOTE

Les appellations employées dans les publications de l'OMM et la présentation des données qui y figurent n'impliquent, de la part de l'Organisation météorologique mondiale, aucune prise de position quant au statut juridique des pays, territoires, villes ou zones, ou de leurs autorités, ni quant au tracé de leurs frontières ou limites.

La mention de certaines sociétés ou de certains produits ne signifie pas que l'OMM les cautionne ou les recommande de préférence à d'autres sociétés ou produits de nature similaire dont il n'est pas fait mention ou qui ne font l'objet d'aucune publicité.

REMERCIEMENTS

Ces directives ont été élaborées initialement par le personnel chargé de la gestion des données climatologiques au sein du Bureau météorologique australien (BOM) et intègrent bon nombre des méthodes utilisées par le logiciel du système de gestion de la qualité du BOM. Meaghan Flannery, John Flannery et Lynda Chambers sont tout particulièrement remerciés pour leurs efforts. Reinhard Spengler du Service météorologique allemand (*Deutscher Wetterdienst (DWD)*) et William Wright (ancien membre du BOM) ont révisé et dirigé la version initiale de cette publication, avec l'aide de David Sinclair (BOM), qui a également ajouté quelques tests, et Peer Hechler (OMM). Cette version a ensuite fait l'objet d'un processus d'examen collégial, au cours duquel les Services météorologiques et hydrologiques nationaux (SMHN) et les membres de l'Équipe d'experts de l'OMM sur l'élaboration et la gestion de données ont été invités à faire part de leurs commentaires. Ces commentaires ont grandement amélioré ce document et nous en sommes reconnaissants à leurs auteurs.

2. **PRINCIPES ET EXIGENCES GÉNÉRALES À RESPECTER POUR GARANTIR LA QUALITÉ DES DONNÉES DESTINÉES À DES APPLICATIONS CLIMATOLOGIQUES**

Les sections qui suivent répertorient les pratiques génériques considérées comme essentielles pour garantir la qualité des relevés climatologiques en surface des SMHN. Des indications plus détaillées figurent dans le [chapitre 3](#).

2.1 **Principes généraux**

- Il y a lieu de mettre en place un processus d'assurance qualité qui couvre le cycle de vie entier des données, depuis l'installation des équipements et l'observation jusqu'à l'archivage final.
- Toutes les méthodes et pratiques employées pour assurer la qualité et la traçabilité tout au long du cycle de vie des données doivent être étayées par une documentation complète; tous les documents pertinents doivent être tenus à jour et être accessibles.
- Il convient d'établir des accords avec les fournisseurs de données pour assurer l'exécution et la conservation des observations à visée climatologique, ainsi que le respect des dix principes relatifs à la surveillance du climat du Système mondial d'observation du climat (SMOC) (voir le [Guide des pratiques climatologiques](#) (OMM-N° 100)).
- Des tests et des procédures de contrôle qualité objectifs doivent être mis en œuvre, et les résultats des tests pour chaque élément de donnée doivent être clairement signalés.
- On conservera toujours une copie des premières données lues ou reçues, même si elles sont considérées comme douteuses¹.
- Chaque SMHN devrait désigner un responsable de l'assurance qualité des données, à qui il incombera de contrôler le processus complet d'assurance qualité tout au long du cycle de vie des données.
- Les processus d'assurance qualité comprennent un retour d'information sur les sources des erreurs. Il est recommandé que le responsable de l'assurance qualité et le personnel des SMHN chargé des systèmes d'observation s'échangent régulièrement des informations, afin que les sources d'erreurs systémiques puissent être repérées et les problèmes correspondants corrigés.

2.2 **Exigences sous-tendant l'adéquation des observations à des fins climatologiques**

- L'emplacement du site d'observation doit être représentatif de l'environnement défini.
- Chaque site d'observation et chaque variable doivent être classés conformément au [Guide des instruments et des méthodes d'observation](#) (OMM-N° 8), notamment pour ce qui concerne la classification du site et de la stabilité des performances.
- Les SMA doivent être équipées d'enregistreurs de données (enregistrement automatique des données), qui doivent être accessibles et pouvoir être utilisés pour restituer les informations au système de gestion des données climatologiques et permettre le recalcul automatique des paramètres dérivés, notamment la température du point de rosée, la pression moyenne au niveau de la mer, etc.

¹ Une première lecture peut n'être pas du tout représentative des conditions météorologiques mais avoir tout de même une valeur scientifique intrinsèque. Par exemple, un relevé de température élevé peut s'expliquer par la présence d'un feu de forêt à proximité d'une station. Dans pareil cas, la lecture initiale ne sera pas intégrée dans les relevés climatologiques officiels, mais sera conservée à des fins de recherche et autres. Voir également la note 4.

4 DIRECTIVES SUR LE CONTRÔLE ET L'ASSURANCE QUALITÉ DES DONNÉES D'OBSERVATION EN SURFACE À DES FINS CLIMATOLOGIQUES

- Les enregistreurs de données doivent satisfaire les exigences techniques fixées par le SMHN et doivent être testés et acceptés.
- Des tests de prédéploiement, d'étalonnage et d'acceptation doivent être effectués pour s'assurer que les capteurs des instruments ou le logiciel et les capteurs de la SMA répondent aux exigences précisées. Il s'agit notamment de s'assurer que les canaux de communication sont suffisamment robustes.
- Les manuels d'observation doivent faire partie intégrante de la documentation relative aux opérations des SMHN et doivent être facilement accessibles.
- Les observateurs doivent suivre régulièrement des formations, en particulier lorsque des modifications sont introduites dans les pratiques d'observation.
- Les SMHN doivent veiller à ce que les stations soient régulièrement inspectées; notamment, des visites sur site doivent être organisées à intervalles appropriés pour déterminer si des opérations de maintenance sont ou non nécessaires et si des capteurs doivent être remplacés, ainsi que pour maintenir des conditions d'observation adéquates sur le site. Des contrôles de réétalonnage des capteurs doivent être régulièrement effectués.
- En cas de changement de modèle de capteur (changement du type de capteur ou du fabricant par exemple) ou de changement de site d'observation, des mesures parallèles doivent être effectuées (avec l'ancien capteur et le nouveau capteur, ou sur l'ancien site et le nouveau site) sur une période d'au moins un an² (deux ans de préférence). Les données originales et parallèles, ainsi que les métadonnées d'observation associées, doivent être archivées et conservées pendant une durée illimitée.
- Les SMHN doivent définir des exigences minimales en matière de qualité à l'intention des fournisseurs de données externes, y compris des exigences concernant les métadonnées d'observation (la mise à disposition de matériels didactiques peut faciliter le respect de ces exigences).

2.3 Exigences en matière de surveillance technique

- **Surveillance des erreurs d'intégration des données:** Un système d'alerte est nécessaire pour garantir que les situations dans lesquelles des données peuvent être corrompues, écrasées ou supprimées par inadvertance soit ne se produisent pas, soit puissent être rapidement détectées et corrigées. Ce système doit être soumis à des tests d'acceptation avant d'être rendu opérationnel et les résultats des procédures de contrôle qualité doivent être analysés régulièrement. Par ailleurs, des processus automatisés devraient être en place pour garantir que les données attendues sont bien reçues, que l'horodatage est exact et que les données intégrées ne sont pas irréalistes.
- **Contrôle des flux de données:** Un système de contrôle en temps réel des équipements techniques (SMA, capteurs, matériel de transmission des données) doit être mis en place afin que les éventuels dysfonctionnements puissent être rapidement repérés et corrigés (on sait que le non-respect de cette procédure peut entraîner des pertes importantes de données climatologiques, qui sont susceptibles de n'être détectées que plusieurs mois plus tard).
- **Procédures de récupération des données:** Lorsque des alertes signalent que des données sont manquantes ou corrompues d'une quelconque manière, il doit exister des procédures normalisées d'exploitation pour récupérer la version originale des données en question; par exemple, les données de SMA doivent être mises en mémoire tampon ou pouvoir

² Le *Guide des pratiques climatologiques* (OMM-N° 100) recommande que «Dans la mesure où cela est réalisable, des observations simultanées devraient être effectuées – sur l'ancien et le nouvel emplacement ou avec les anciens et les nouveaux instruments – pendant une période de chevauchement d'au moins une année, et de préférence de deux ou plusieurs années, afin de déterminer les effets des modifications apportées aux instruments ou à l'emplacement des stations sur les données climatologiques.»

être récupérées à partir d'un enregistreur de données, et la période de conservation des données doit être suffisamment longue. Ces pertes de données sont souvent dues à des problèmes de communication plutôt qu'à une défaillance du capteur lui-même. Il faut agir promptement, car les enregistreurs de données des SMA ou les bases de données en amont du système de gestion des données climatologiques (où les données transmises à un SMHN sont souvent stockées dans un premier temps) ne conservent en général les données que quelques jours, avant que celles-ci ne soient écrasées. Par conséquent, si, faute d'un contrôle adéquat de ces procédures, les pertes de données ne sont découvertes qu'après un certain temps, la perte pour les relevés climatologiques peut être considérable.

- Tout problème lié aux capteurs et toute perte de données doivent être corrigés dès que possible et consignés, et l'occurrence et la correction du problème doivent être notifiées aux utilisateurs des données.
- Il incombe au fournisseur de données et/ou au service informatique d'identifier les processus qui entraînent la perte ou la corruption de données et d'y remédier, et des protocoles internes appropriés doivent être mis en place à cet effet.

2.4 Exigences sous-tendant l'adéquation des pratiques de gestion des données climatologiques

- Du point de vue de la traçabilité des données, il est important de pouvoir retracer la provenance des données, du capteur à l'archive. Pour ce faire, il est souhaitable de conserver un enregistrement des mesures initiales, ainsi que de toute modification apportée aux données.
- Les indicateurs de qualité sont une composante essentielle de la gestion des données et constituent une mesure de la confiance dans la véracité des données³ (la question des indicateurs de source et de qualité sera traitée plus en détail dans la [sous-section 3.4.5](#)).
- Les procédures de contrôle qualité doivent faire l'objet d'une documentation détaillée qui doit être accessible aux utilisateurs de données si nécessaire, de préférence en ligne.
- Toute modification des valeurs mesurées ou observées doit être signalée et consignée dans les métadonnées, la valeur originale et la valeur modifiée étant conservées l'une et l'autre.
- Le cycle d'assurance qualité des données comprend généralement plusieurs étapes (décrites dans le [chapitre 3](#)). Les résultats des tests de contrôle qualité doivent être consignés à chaque étape, notamment via l'attribution d'un indicateur de qualité. Ainsi, des indicateurs de qualité peuvent être attribués aux données au point d'observation, lors de l'intégration et après un contrôle qualité en différé au sein du système de gestion des données climatologiques.
- Les valeurs douteuses ou suspectes qui ont été modifiées ne doivent pas être supprimées, mais être stockées avec le résultat de l'évaluation de leur qualité⁴.
- Lorsqu'ils récupèrent des données mesurées ou observées, les utilisateurs de données doivent avoir accès à toutes les informations disponibles sur leur qualité (explications comprises). Il convient, le cas échéant, d'établir une distinction entre les utilisateurs

³ Au moment de la rédaction de ce document, il n'existe pas de directives normalisées de l'OMM concernant la définition des indicateurs de qualité des données; on s'attend néanmoins à ce que de telles directives soient élaborées par l'OMM dans un avenir proche.

⁴ Bien entendu, il incombe aux SMHN de veiller à ce que les relevés climatologiques présentés aux utilisateurs soient du meilleur niveau possible et capables de soutenir de manière fiable les services climatologiques. Les valeurs erronées ne doivent pas apparaître dans cette version publique des relevés climatologiques mais doivent pouvoir être consultées s'il y a lieu, à des fins d'enquête publique ou pour faciliter la détection et la résolution de problèmes de mesure ou de logiciel.

internes du SMHN, qui peuvent avoir besoin d'informations extrêmement détaillées, et les utilisateurs externes, qui ne seront probablement intéressés que par les données ayant passé les contrôles qualité avec succès.

- Les métadonnées d'observation soutiennent les données, fournissent des informations quant à leur provenance, peuvent faciliter les enquêtes sur les données suspectes et sont essentielles pour l'homogénéisation. Par conséquent, la conservation des métadonnées pendant une période au moins égale à la durée de vie des données doit être considérée comme une pratique obligatoire. Il est important que ces métadonnées soient facilement accessibles à l'opérateur de contrôle qualité, ainsi qu'aux utilisateurs des données.

Notes:

1. La publication intitulée *Climate Data Management System Specifications* (WMO-No. 1131) donne des informations détaillées sur les pratiques de gestion des données et de gestion de la qualité appropriées.
2. Le *Manuel sur le Cadre mondial pour la gestion de données climatologiques de qualité* (OMM-N° 1238), qui fait partie des textes réglementaires de l'OMM, définit des normes et des pratiques recommandées pour la gestion des données climatologiques.

2.5 **Considérations générales sur la conception d'un système de contrôle qualité**

2.5.1 **Considérations sur l'automatisation**

Les procédures de contrôle qualité peuvent être appliquées manuellement, automatiquement ou semi-automatiquement à chaque étape du cycle de vie des données. Bien que les détails de ces procédures varient selon les SMA, une procédure de contrôle qualité type peut s'articuler autour des séquences suivantes:

- Contrôle qualité au point d'observation. Par exemple, un observateur manuel sur le point d'enregistrer une observation par voie électronique pourra être invité à vérifier la valeur proposée si les algorithmes indiquent que cette valeur est improbable. Dans le cas d'observations provenant de SMA, divers contrôles automatisés peuvent être intégrés sur la base de la fréquence et des résultats détaillés de l'échantillonnage et des contrôles de portée;
- Contrôle qualité des bases de données en amont du système de gestion des données climatologiques, pour appuyer l'élaboration de produits en temps réel;
- Contrôle qualité au point d'intégration dans le système de gestion des données climatologiques, qui permet d'éliminer automatiquement les valeurs qui échouent à un test de portée étendu ou à d'autres vérifications (même si une copie de la valeur elle-même doit être conservée);
- Contrôle qualité post-intégration dans le système de gestion des données climatologiques (manuel, semi-automatique ou automatique).

L'exposé ci-dessous traite uniquement du contrôle qualité des données au sein d'un système de gestion des données climatologiques. On parle parfois de contrôle qualité *en différé*, du fait que le contrôle peut être effectué un certain temps après l'événement d'observation et l'intégration des données. Le contrôle qualité en différé peut faire intervenir des contrôles manuels, automatiques ou semi-automatiques, ou une combinaison de ces différents modes. Ce document ne formule pas de recommandations spécifiques concernant les contrôles optimaux, car ceux-ci dépendent de facteurs tels que la nature des données, le nombre de stations et les ressources disponibles pour le contrôle qualité (toutefois, les contrôles devraient avoir lieu le plus tôt possible afin que, dans l'éventualité où il faudrait vérifier une valeur auprès d'un observateur, celui-ci ait encore l'événement en mémoire).

Le contrôle qualité en différé se déroule généralement selon l'approche suivante: application de tests de contrôle qualité, attribution d'un indicateur de qualité, qui doit être consigné dans les métadonnées d'observation, et archivage des données et des métadonnées ainsi que des éléments détaillés relatifs aux corrections éventuelles, dans un journal d'audit. Comme son nom l'indique, le contrôle manuel de la qualité fait référence aux contrôles effectués sur les données par les opérateurs de contrôle qualité. Il s'agit d'un processus qui demande généralement beaucoup de travail et qui est subjectif et lourdement tributaire des connaissances et de l'expérience des opérateurs. Les techniques manuelles peuvent convenir lorsqu'il y a peu de stations et que l'accès aux logiciels de diagnostic de contrôle qualité est limité. Pour sa part, le contrôle automatique de la qualité consiste à parcourir les données et à attribuer les indicateurs de qualité de façon entièrement automatique, c'est-à-dire sans intervention manuelle. C'est le moyen le plus rapide de traiter les données entrantes. Cependant, les tests entièrement automatisés ont parfois tendance à signaler trop de valeurs potentiellement suspectes (faux positifs), ou au contraire trop peu, selon la façon dont les paramètres de filtrage sont définis, et peuvent signaler certaines valeurs extrêmes authentiques comme étant suspectes, ce qui n'est pas souhaitable dans la mesure où l'évolution des extrêmes suscite un intérêt scientifique croissant lié aux besoins de surveillance du climat.

Les techniques semi-automatiques représentent un compromis entre les techniques manuelles et les techniques entièrement automatiques. Dans ce cas, les données sont soumises à un contrôle automatisé au regard d'une série de critères de test, et les valeurs qui sont rejetées par un test (ou plusieurs tests) sont signalées comme suspectes et sont alors analysées manuellement par un opérateur formé au contrôle qualité. Pour déterminer si une valeur suspecte est erronée ou non, il est souhaitable que les opérateurs de contrôle qualité disposent d'un ensemble d'outils facilement accessibles, dont l'expérience a montré qu'ils sont utiles pour diagnostiquer les erreurs. Ces outils peuvent faciliter l'utilisation de données satellitaires et/ou radar complémentaires (pour évaluer les situations de pluie ou d'absence de pluie), ainsi que de données de radiosondage, et permettre d'accéder au message original (qui peut avoir été mal codé). Il existe d'autres moyens encore de déterminer si une valeur suspecte est erronée ou non: on peut par exemple reporter les valeurs sur une carte pour vérifier leur cohérence spatiale, tracer une série chronologique pour une station et la comparer avec les séries des stations voisines (dans le cas des précipitations, ce type de vérification est souvent utile pour repérer les cas où la valeur des précipitations a été enregistrée le mauvais jour ou représente un total cumulé), ou encore comparer les données aux résultats du modèle de prévision numérique du temps (PNT) pour la même période et le même emplacement.

Il est à noter qu'avec les données à fréquence élevée, telles que les données infrahoraires des SMA, les techniques automatiques sont probablement le seul mode de contrôle qualité réaliste, encore que, même dans ce cas, il soit souhaitable de procéder à un contrôle supplémentaire pour déterminer, dans la mesure du possible, si l'erreur était due à une défaillance du logiciel ou à un phénomène physique tel qu'un feu de forêt ou l'activation d'un système d'arrosage.

Il est probable qu'à l'avenir les techniques d'apprentissage automatique, y compris les approches par système expert, permettront d'améliorer le contrôle automatique de la qualité. Il est conseillé à cet égard aux SMHN de se tenir au courant des avancées des sciences des données et de leurs applications potentielles.

Une autre question qui se pose lors des tests de contrôle qualité est de savoir si ces tests doivent être effectués en parallèle ou successivement. Par exemple, il peut être utile, lorsqu'on effectue des comparaisons avec des stations voisines, d'éliminer d'abord du test les stations susceptibles d'avoir des valeurs erronées.

Enfin, il convient de préciser que, dans le cas de données transcrites dans un système de gestion des données climatologiques à partir de formulaires papier, les fautes de saisie peuvent constituer une source importante d'erreurs. Il est possible d'atténuer ce problème en insérant des contrôles à l'entrée qui préviennent le clavier qu'une valeur a peut-être été mal saisie. Si les ressources du SMHN le permettent, il est démontré que la double saisie est un moyen très efficace de limiter les erreurs de saisie. La saisie double ou multiple est particulièrement recommandée lorsque l'on numérise des données historiques dans le cadre d'une opération de sauvetage de données.

2.5.2 **Conception et mise en œuvre d'un système de contrôle qualité**

- Dans le cas de systèmes manuels ou semi-automatiques, il est fortement recommandé, lors de la conception du système, de la mise en place d'un contrôle qualité pour une nouvelle variable ou de la modification du système de contrôle qualité existant, d'associer dès le départ les opérateurs qui utiliseront le système aux phases de planification et de mise en œuvre. C'est un moyen de s'assurer que le système est en phase avec l'environnement opérationnel réel dans lequel les opérateurs seront amenés à travailler et de faire en sorte que les opérateurs s'approprient le système. En outre, les opérateurs de contrôle qualité peuvent formuler des recommandations sur les aides au diagnostic qui leur paraissent utiles. Pour concevoir ou mettre en œuvre un système informatisé, il est en outre recommandé d'adopter des méthodes «agiles» – les développeurs devraient ainsi travailler directement avec le personnel chargé du contrôle qualité pour échanger, par exemple, sur les processus décisionnels opérationnels et les lacunes du système.
 - Les statistiques sur les erreurs et les faux positifs (qui surviennent lorsque les routines de contrôle signalent une valeur comme étant suspecte, avant que d'autres vérifications ne concluent qu'elle est finalement valide) doivent être compilées et tenues à jour. Ces statistiques fournissent en effet des informations sur les erreurs systémiques devant être communiquées aux fournisseurs de données, ainsi que des informations qui peuvent aider à affiner le système de contrôle qualité et ses algorithmes.
 - Dans le prolongement du point précédent, il convient d'évaluer régulièrement l'efficacité du système de contrôle qualité afin de déterminer les techniques de test qui fonctionnent le mieux pour tels ou tels types de données et qui peuvent contribuer à affiner le système.
-

3. **ÉLÉMENTS DU PROCESSUS D'ASSURANCE QUALITÉ SUR LE CYCLE DE VIE DES DONNÉES**

3.1 **Vue d'ensemble**

Le processus d'assurance qualité sur le cycle de vie des données peut être décomposé en cinq grandes étapes décrites ci-dessous¹.

Étape 1: **Observation**

Cette étape a trait à des aspects de la qualité des données qui sont associés à la validité du processus d'observation, et notamment à la question de savoir si la station d'observation, qu'elle soit dotée de personnel ou automatique, a été mise en place et fonctionne correctement. Ces aspects doivent figurer dans les métadonnées d'observation, telles que celles relatives aux caractéristiques du site, les métadonnées des instruments et les métadonnées confirmant que les spécifications pour l'installation du site ont été respectées. Les métadonnées d'observation doivent également mettre en évidence le processus de capture des données, que l'observation soit automatique ou manuelle, ainsi que les processus d'entretien courant et de surveillance technique. Des contrôles logiciels peuvent également être mis en place pour déterminer si un relevé est raisonnable ou non, ou si un instrument a mal fonctionné ou non. Dans le cas d'un enregistrement manuel sur papier, le processus d'assurance qualité doit garantir que les formulaires papier sont transmis en temps utile au SMHN.

Étape 2: **Transmission et intégration des données**

Cette étape correspond à la surveillance de la qualité des données sur leur parcours entre le site d'observation et le système de gestion des données climatologiques. (Il peut aussi y avoir des bases de données en amont qui effectuent leurs propres contrôles qualité sur les données.) Les contrôles qualité des données devraient être effectués au point d'intégration électronique dans le système de gestion des données climatologiques (pour identifier les «valeurs impossibles»). Si les données sont saisies au clavier, des contrôles supplémentaires doivent être mis en place pour vérifier que les données ont été saisies correctement. Cette étape doit également permettre de détecter les données attendues, mais non reçues.

Étape 3: **Gestion de la base de données climatologiques (au sein du système de gestion des données climatologiques)**

Cette étape recouvre généralement le vaste domaine de la gestion (centrale) des données climatologiques², qui fait suite à la phase d'intégration. C'est là qu'intervient le contrôle qualité en différé, qui peut reposer sur des tests de contrôle qualité manuels, semi-automatiques ou automatiques. C'est cette étape qui est principalement abordée dans la suite de ce document.

Étape 4: **Archivage final**

Cette étape comprend les exigences relatives à l'archivage final dans le système de gestion des données climatologiques ainsi que les exigences relatives à l'archivage des documents papier dans un registre approprié.

¹ Dans ce document, le «cycle de vie des données» s'arrête à l'archivage des données, alors que dans l'usage courant, il s'étend jusqu'à la suppression des données.

² On trouvera beaucoup plus de précisions sur les exigences en matière de gestion des données dans la publication *Climate Data Management System Specifications* (WMO-No. 1131).

Étape 5: Récupération en cas de sinistre

Cette étape porte sur la protection contre la perte de données, qui peut comprendre des opérations telles que des sauvegardes électroniques et la numérisation de documents papier. Il est à noter que des procédures de récupération doivent également être mises en place pour les étapes intermédiaires.

3.2 Étape 1: Observation

Cette section traite des aspects de la qualité des données qui sont associés à l'observation des éléments météorologiques et au fait de s'assurer que les stations d'observation, qu'elles soient dotées en personnel ou automatiques, ont été mises en place et fonctionnent correctement.

3.2.1 Normes d'observation

La recherche d'un site et la décision d'installer ou non une station d'observation à un emplacement déterminé influent sur la qualité des données qui peut être obtenue compte tenu des besoins d'observation. À des fins climatologiques, le site doit satisfaire les besoins des utilisateurs de données sur la plus longue période possible³. Les conditions d'observation devraient donc rester largement inchangées sur une longue période. Le site doit également être représentatif des conditions météorologiques d'une zone plus vaste.

On trouve des normes et des orientations sur les observations météorologiques dans de nombreuses publications de l'OMM, notamment:

- *Manuel du Système mondial intégré des systèmes d'observation de l'OMM* (OMM-N° 1160)
- *Guide du Système mondial intégré des systèmes d'observation de l'OMM* (OMM-N° 1165)
- *Guide des instruments et des méthodes d'observation* (OMM-N° 8)
- *Guide des pratiques climatologiques* (OMM-N° 100)
- *Norme relative aux métadonnées du WIGOS* (OMM-N° 1192)
- *L'automatisation des réseaux d'observation météorologique et ses répercussions sur la surveillance à long terme du climat* (OMM-N° 1202)

Les éléments essentiels des normes et pratiques d'observation du point de vue du contrôle et de l'assurance qualité dans le contexte des applications et services climatologiques sont précisés ci-dessous.

- Procédures de prédéploiement des capteurs: Les capteurs doivent répondre à des normes spécifiques couvrant les aspects suivants:
 1. Spécifications d'essai du fabricant;
 2. Examen par le SMHN ou par un tiers externe de la capacité du capteur de respecter les normes applicables à la climatologie attendue (il est recommandé que les tests du capteur reproduisent la plage de valeurs attendue du paramètre);
 3. Normes de l'OMM pour les capteurs;
 4. Étalonnage.

³ Ces conditions sont nombreuses et variées, et les données destinées à la surveillance du changement climatique sont soumises à des exigences de qualité plus strictes que les données destinées à d'autres fins. Cependant, dans tous les cas, l'objectif est de s'assurer que les relevés climatologiques sont de la meilleure qualité possible.

- Procédures de prédéploiement des SMA: Les SMA doivent répondre à des normes spécifiques couvrant les aspects suivants:
 1. Documentation et spécifications d'encodage pour les algorithmes des SMA (si elles sont disponibles⁴), qui font partie intégrante des métadonnées d'observation que doit conserver le SMHN;
 2. Bancs d'essai et essais pratiques pour s'assurer que la plate-forme de la SMA est adaptée à sa finalité (c'est-à-dire qu'elle est robuste, fiable et utilisable) et pour garantir sa conformité aux normes du SMHN et de l'OMM;
 3. Capacité de l'enregistreur de données (qui dépend de la capacité de récupération des données du SMHN);
 4. Bonne qualité des communications, de l'alimentation électrique, etc., pour réduire les temps d'arrêt au minimum.

Tous les sites existants doivent être classés selon la classification des sites figurant dans le [Guide des instruments et des méthodes d'observation](#) (OMM-N° 8) pour chacun des paramètres observés sur le site; la classification doit être incluse dans les métadonnées des observations de la station.

3.2.2 **Métadonnées des observations**

Les métadonnées des observations doivent être consignées et stockées dans le système de métadonnées du SMHN, conformément aux pratiques normalisées de l'OMM et du SMHN ([Guide du Système mondial intégré des systèmes d'observation de l'OMM](#) (OMM-N° 1165)) et à la [Norme relative aux métadonnées du WIGOS](#) (OMM-N° 1192).

Pour les besoins du contrôle qualité des données climatologiques, il est recommandé que les métadonnées des observations comprennent un maximum des informations suivantes:

- Informations sur le site et l'exposition, y compris:
 - Latitude, longitude et altitude;
 - Relevé de la ligne d'horizon, si possible (y compris sa description);
 - Photographies du site prises du nord, du sud, de l'est et de l'ouest;
 - Description détaillée de l'environnement local, avec un relevé du site (diagramme polaire) faisant apparaître l'altitude des instruments et la distance entre les instruments et les arbres, les bâtiments, etc.
- Renseignements sur les instruments/capteurs, y compris:
 - Numéro de série;
 - Description détaillée des vérifications d'étalonnage;
 - Historique des déploiements/de l'aptitude au service;
 - Méthode d'établissement des rapports et moment de l'observation.

L'absence de métadonnées d'observation appropriées, correctes et à jour peut compromettre l'ensemble du processus de gestion de la qualité. Il est à noter que, pour les pratiques

⁴ Malheureusement, de nombreux fournisseurs de SMA n'autorisent pas l'accès aux algorithmes des capteurs.

d'évaluation de l'homogénéité, ces métadonnées doivent être régulièrement mises à jour pour faire apparaître tout changement susceptible d'affecter les mesures, par exemple la croissance des arbres ou les changements d'occupation des sols.

3.2.3 **Contrôle qualité au niveau de la station**

Cette section s'applique à la fois aux observations manuelles et aux observations électroniques générées automatiquement à partir de stations sans personnel.

Le fait d'effectuer la première étape du contrôle qualité à la station en temps réel présente un grand avantage. En effet, les données brutes de tous les capteurs sont disponibles sur place et, dans le cas de stations avec personnel, les observateurs peuvent procéder à des contrôles de cohérence par rapport à d'autres paramètres. Cela évite que des données manifestement erronées soient versées dans les archives de données et éventuellement incorporées dans des produits.

Observations manuelles

Tous les observateurs doivent recevoir des instructions normalisées concernant les observations météorologiques et s'y conformer. Ces instructions font partie intégrante de la documentation et de la conformité des procédures. Si des modifications sont apportées aux instructions, les observateurs doivent en être informés. Les observateurs doivent suivre des formations de mise à niveau à intervalles réguliers. Idéalement, des capteurs de rechange devraient être aisément accessibles en cas de panne d'un capteur, de manière à réduire au minimum les données manquantes.

Les relevés que les observateurs ont effectués manuellement doivent être envoyés à un bureau central (généralement le SMHN) immédiatement après la fin du mois. La station doit en conserver une copie.

Pour les observations qui sont lues entièrement ou en partie manuellement, il convient d'effectuer les vérifications suivantes (une liste récapitulative peut être utile pour s'assurer que ces vérifications sont systématiques):

- Contrôle de la cohérence par rapport à d'autres observations manuelles;
- Tests de domaine et de contraintes par rapport à des normes climatologiques raisonnables;
- Contrôle de la cohérence par rapport aux relevés de la station (c'est-à-dire les relevés historiques pour cette station, y compris les extrêmes passés);
- Calculs sur les observations (tels que le calcul de l'humidité relative);
- Pour les sites avec personnel, des contrôles indépendants, qui peuvent être effectués par la personne responsable ou un autre observateur.

On trouvera de plus amples précisions sur ces contrôles dans l'[annexe 1](#) et l'[annexe 2](#).

Observations générées automatiquement

Pour les observations générées de façon entièrement automatique (via des SMA), les contrôles suivants doivent être effectués:

- Tests de contraintes imposés par l'intégration dans la base de données;
- Tests de contraintes pour déterminer si les observations sont raisonnables sur le plan climatologique;

- Contrôles de la cohérence par rapport aux autres paramètres de la station (cohérence interne).

On trouvera des exemples de ces contrôles dans l'[annexe 1](#) et l'[annexe 2](#).

Idéalement, les résultats des contrôles qualité automatiques devraient être stockés sous la forme d'indicateurs de qualité, transmis au bureau régional ou central avec la valeur de donnée correspondante.

Pour les observations électroniques générées automatiquement à partir de stations sans personnel, il est recommandé que l'enregistreur de données dispose d'une capacité suffisante pour pouvoir stocker les données pendant des interruptions éventuelles des communications et que, pour les sites cruciaux, l'on envisage une méthode de récupération des données robuste (par exemple, interrogation par satellite).

En cas de problèmes techniques dus à un capteur, à l'enregistreur de données ou à la SMA, les systèmes doivent émettre des «alarmes techniques» automatiques, que le SMHN pourra analyser et faire suivre de mesures appropriées.

3.2.4 **Maintenance**

Il convient d'effectuer des visites régulières dans toutes les stations pour vérifier les conditions du site et – plus important encore – réaliser des travaux de maintenance, qui comprennent notamment le remplacement des capteurs. Les résultats de ces visites doivent être consignés.

Lorsqu'un capteur est remplacé dans une station individuelle, un réétalonnage est nécessaire et ses résultats doivent être reportés dans les métadonnées des observations.

Quand on procède au remplacement des capteurs à l'échelle d'un réseau d'observation, des mesures parallèles doivent être effectuées avec les anciens et les nouveaux capteurs sur une période d'au moins un an, et de préférence de deux ans⁵ ou plus (voir le [Guide des pratiques climatologiques](#) (OMM-N° 100)) dans toutes les stations affectées, ou au moins dans un ensemble de stations de référence représentatives de l'éventail des conditions climatologiques du réseau, avant l'introduction des nouveaux capteurs dans l'ensemble du réseau.

Les résultats de ces mesures parallèles doivent être consignés dans les métadonnées des observations, et être ainsi accessibles aux utilisateurs des données qui souhaitent mener de plus amples analyses. Il est important d'enregistrer la date exacte du changement de capteur pour chaque site.

3.2.5 **Surveillance technique**

Cette section s'applique aux stations semi-automatiques et entièrement automatiques. Elle concerne les systèmes de surveillance technique semi-automatiques qui sont capables de détecter et, dans une certaine mesure, de corriger les dysfonctionnements des capteurs, des enregistreurs de données et des systèmes de communication.

Les erreurs au niveau des capteurs et des enregistreurs de données doivent être repérées au plus vite, grâce à la surveillance en temps réel des paramètres techniques. Cela réduit la probabilité que des données soient manquantes ou que des données défectueuses continuent d'être traitées sur une longue période au détriment de la qualité des séries chronologiques de données climatologiques et des produits climatologiques dérivés.

⁵ La raison principale pour laquelle il convient de recueillir des mesures parallèles sur une période de deux ans, plutôt que sur une seule année, est qu'il est utile de disposer de données supplémentaires au cas où la première année de la comparaison serait inhabituelle sur le plan climatique.

Il est fortement recommandé aux opérateurs de réseau d'observation de tenir compte des éléments suivants lors de l'élaboration des mécanismes de détection des systèmes de surveillance:

- Les capteurs intelligents sont capables de déterminer si l'alimentation électrique, le mécanisme d'étalonnage interne ou tout système interne, tel que le système de chauffage, fonctionnent correctement. Si les capteurs détectent un dysfonctionnement, les signaux correspondants sont transmis à l'enregistreur de données, et de l'enregistreur au logiciel de surveillance (centre de contrôle) pendant la récupération des données. Dans l'idéal, si l'alimentation électrique du capteur, le mécanisme d'étalonnage interne ou un système interne ne fonctionnent pas correctement, cette information devrait être rendue visible sous la forme d'un indicateur de qualité joint aux valeurs mesurées pendant le traitement du signal dans l'enregistreur de données, pour permettre de détecter facilement le dysfonctionnement.
- Les anomalies de fonctionnement peuvent toucher l'enregistreur de données, l'unité de traitement des données de la station, ou les deux. Le fait que les valeurs mesurées ne soient plus encodées ou que la récupération des données de la station ne soit plus possible donne une indication sur l'origine du problème.
- La surveillance technique de la communication des données permet d'obtenir rapidement des informations sur les éventuelles erreurs de processus. C'est à partir de ces informations que l'on pourra enquêter sur les omissions dans les entrées de données (c'est-à-dire les différences entre les données attendues et les données effectivement reçues), comprendre pourquoi certaines données sont manquantes et, sur cette base, engager des procédures de rectification. Il convient de noter que ces défaillances peuvent se produire en divers points de l'infrastructure de données du SMHN.

Les informations détaillées sur toutes les erreurs techniques doivent être enregistrées dans une base de données. Cette base de données devrait renfermer au moins les éléments suivants:

- Type d'erreur
 - Capteur;
 - Enregistreur de données/SMA;
 - Ligne de communication;
- Heure de début de l'erreur;
- Heure de fin de l'erreur;
- Mesures prises pour corriger l'erreur (Note: Il faut veiller à ce que ces mesures soient clairement décrites);
- Si l'erreur a entraîné ou non des omissions dans les données;
- Identité de l'opérateur qui a traité l'erreur.

3.3 **Étape 2: Transmission et intégration des données**

Cette section examine les problèmes de qualité des données associés au système de transmission des données entre le site d'observation et la base de données (centrale), ce qui comprend l'intégration des données.

3.3.1 **Saisie manuelle**

De nombreux SMHN reçoivent les relevés des stations sous la forme de documents papier, qui sont ensuite numérisés par le personnel du SMHN. Dans le cadre du processus d'assurance et de contrôle qualité, il est recommandé ce qui suit:

- Le relevé doit être enregistré sous forme d'image numérique et l'image doit être stockée⁶, dans le cadre de la stratégie de sauvetage de données du SMHN;
- Des saisies multiples doivent être effectuées pour se prémunir contre les erreurs de numérisation (la double saisie est recommandée);
- Un identifiant d'opérateur doit être joint au document numérisé et incorporé dans les métadonnées. Cet identifiant permettra d'évaluer la performance des opérateurs et donnera des indications sur les besoins de formation;
- La date et l'heure de la numérisation doivent être annexées au relevé numérisé. Cela facilite le suivi de tous les changements éventuels;
- Un contrôle qualité doit être pratiqué au stade de la numérisation pour réduire l'incidence des erreurs. En cas de numérisation dans un système de gestion des données climatologiques, les algorithmes de contrôle qualité au niveau de l'intégration des données devraient comprendre les éléments suivants:
 - Contrôles de domaine et de cohérence;
 - Contrôles pour vérifier que les données relatives à la station, à l'heure et à la date ne figurent pas déjà dans la base de données;
 - Contrôles des métadonnées, pour vérifier que le nom et l'emplacement de la station correspondent au numéro de station attribué;
 - Vérifications arithmétiques du calcul des paramètres (par exemple, calcul du point de rosée à partir des valeurs de température du thermomètre sec et du thermomètre mouillé);
 - Dans le cas de variables cumulatives telles que les précipitations et l'évaporation, addition arithmétique des valeurs quotidiennes pour vérifier la concordance avec le total mensuel enregistré.

3.3.2 **Intégration électronique des données**

Pour le contrôle qualité au niveau l'intégration des données, il est recommandé de limiter les tests à l'identification des données dont on peut montrer qu'elles sont manifestement erronées d'un point de vue statistique au moyen de contrôles de portée et de domaine. Si un contrôle qualité supplémentaire est effectué au moment de l'ingestion, il est recommandé que ces données douteuses soient signalées comme telles en vue de leur analyse par le système de contrôle qualité en différé du SMHN (voir la [section 3.4.1](#)).

⁶ En général, le document papier doit également être conservé, sauf si l'autorité nationale compétente en matière d'archives donne un avis contraire juridiquement contraignant.

Recommandations:

- Le processus d'intégration doit mettre en œuvre un décodage fiable des données et des métadonnées pour tous les paramètres;
- Le logiciel de contrôle qualité à l'intégration doit être indépendant du logiciel utilisé pour l'intégration même (afin de réduire les problèmes liés aux mises à niveau et aux mises à jour);
- Récupération en cas de sinistre: Il doit être possible de réintégrer les données récupérées si nécessaire;
- Le logiciel doit être facilement extensible, de manière à pouvoir traiter de nouveaux types, formats et flux de données;
- Il convient de tenir un journal d'erreurs, qui permettra de repérer les éventuels problèmes liés au système d'intégration des données.

3.3.3 **Détection des erreurs systémiques**

Même dans un système entièrement automatique, il doit exister des outils adéquats qui permettent de détecter et d'analyser les problèmes systémiques survenant sur le réseau. En outre, toute modification du système de contrôle qualité doit faire l'objet de tests d'acceptation ainsi que de tests dans l'environnement opérationnel, et ces tests et leurs résultats doivent être intégralement consignés. Des tests de bout en bout (du capteur à l'archive) sont également nécessaires.

Dans l'idéal, un dispositif approprié devrait être en place pour consigner les problèmes de données identifiés et analyser les schémas d'erreurs. Si un type d'erreur particulier apparaît de façon répétée, ce peut être le signe qu'il existe un problème systémique – au point d'observation, dans le système d'observation, dans le logiciel d'intégration, ou au niveau des opérateurs de saisie.

Le fait que la fréquence de données signalées comme «suspectes» pour un paramètre déterminé soit plus élevée qu'à l'accoutumée peut être le signe qu'il existe un problème systémique au niveau du réseau pour ce paramètre.

3.4 **Étape 3: Contrôle qualité dans la base de données climatologiques**

Cette section traite du contrôle qualité au sein du système de gestion des données climatologiques, en passant en revue les règles, les procédures et les tests.

Telles qu'elles sont définies actuellement, les procédures de filtrage appliquées aux données d'observation qui arrivent dans le système de gestion des données climatologiques visent généralement à détecter les erreurs les plus évidentes. Cependant, avant que les données ne soient archivées dans les relevés climatologiques officiels, il convient de les soumettre à un dernier contrôle qualité complet. Les processus de contrôle qualité diffèrent selon la source des données, le paramètre mesuré et la résolution temporelle. En fonction de la structure du SMHN, le contrôle qualité peut être exécuté par une entité centrale ou mené au niveau régional selon une procédure cohérente et bien définie, appliquée à l'échelle du SMHN. L'essentiel est que le contrôle final soit effectué selon une procédure uniforme, avec des outils et des approches homogènes, avant l'archivage final.

3.4.1 **Introduction**

Les tests de contrôle qualité recommandés ici sont conçus pour effectuer une analyse complète – mais non exhaustive – des données reçues dans la base de données climatologiques. Ils doivent être réalisés après l'intégration et, à ce titre, sont souvent désignés par le terme de contrôle qualité en différé⁷. Ces tests relèvent de cinq grandes catégories:

- **Tests de contraintes** (tests de portée fondés sur les capteurs, qui déterminent si les observations se situent en dehors des limites théoriques d'un capteur, et tests de domaine, qui déterminent si la valeur est possible du point de vue scientifique);
- **Tests de cohérence** (qui vérifient si les données provenant de deux sources ou plus sont cohérentes (par exemple: La durée d'insolation dépasse-t-elle le maximum possible pour cette période de l'année? La température du point de rosée est-elle supérieure à la température du thermomètre mouillé? Y a-t-il une incohérence entre les valeurs trihoraires et les températures maximales/minimales quotidiennes? La valeur des précipitations est-elle improbable compte tenu de l'humidité, des nuages et d'autres facteurs?));
- **Tests heuristiques** (qui s'appuient sur l'expérience et les connaissances en matière de processus d'observation et d'instrumentation pour détecter les valeurs incohérentes ou improbables; il s'agit par exemple de tests visant à déterminer si la mèche d'un thermomètre mouillé a séché);
- **Tests de transmission des données** (par exemple, tests consistant à comparer l'heure locale et l'heure de réception des données, tests pour déterminer si les intervalles entre les observations sont excessivement élevés, etc.);
- **Tests statistiques** (qui comparent les données reçues avec les données historiques ou les profils de variabilité spatiale afin de détecter les valeurs incohérentes ou improbables; par exemple, tests de courbe plate ou de pic; cette catégorie comprend également les tests de cohérence spatiale par régression linéaire ou fondés sur la cartographie et les tests de cohérence temporelle).

On trouvera dans l'[annexe 1](#) une vue d'ensemble de ces tests, ainsi que des indications quant à leur caractère obligatoire, recommandé ou facultatif. L'[annexe 2](#) les décrit de façon plus détaillée.

Des paramètres spécifiques de surveillance de la qualité (PSQ) sont définis pour déterminer si les valeurs testées ont «réussi» les tests décrits ci-dessus ou ont «échoué». Un PSQ est une valeur de référence numérique qui exerce une contrainte, souvent sous la forme d'une limite supérieure ou inférieure. Les PSQ peuvent être définis pour un élément météorologique, une station, un mois ou un type de test particulier et peuvent être générés statistiquement (par exemple, plus de x écarts types par rapport à la moyenne). Les valeurs qui sortent de la plage des PSQ pour un test particulier sont généralement considérées comme ayant échoué à ce test et peuvent être signalées comme suspectes, tout au moins jusqu'à ce que de nouvelles analyses les valident ou confirment qu'elles sont suspectes ou fausses.

La sensibilité d'un test dépend souvent du fait que le test soit ou non automatique ou semi-automatique. Les tests de contrôle qualité automatiques peuvent aboutir à des données moins fiables et sont susceptibles de surestimer le nombre d'erreurs ou de sous-estimer les extrêmes⁸. Néanmoins, ces tests sont adaptés aux données à fréquence élevée, en raison des variations limitées qu'elles enregistrent sur de courtes périodes et des volumes de données élevés que les tests doivent traiter.

⁷ Il convient de noter que certains des tests décrits ici peuvent également être effectués à d'autres étapes du cycle de vie des données.

⁸ On peut améliorer cette situation par différents moyens, par exemple en constituant un échantillon d'entraînement de grande taille, en procédant à des tests à variables multiples (voir ci-après), ou en utilisant des scénarios spécifiques. Les techniques d'apprentissage automatique peuvent être utiles à cet égard.

3.4.2 **Comment gérer les différences entre sites**

Les différences entre sites doivent être prises en considération lors des tests de contrôle qualité, tout particulièrement si le SMHN couvre un large éventail de zones climatiques et d'altitudes.

Par exemple, les plages de certaines variables diffèrent considérablement entre stations côtières et stations continentales, et cela peut avoir un impact sur le taux attendu de fausses erreurs si les mêmes paramètres de test sont utilisés. Pour les SMHN qui couvrent différentes zones climatiques, il existe deux approches principales pour traiter les différences entre sites et les variations des taux de fausses erreurs. La première consiste à calculer les PSQ et les constantes de normalisation séparément pour chaque zone. La deuxième est d'utiliser un seul système en suivant une procédure semi-automatique, c'est-à-dire en confiant à des opérateurs formés pour ce faire le soin d'identifier les fausses erreurs dues aux variations saisonnières et à la variabilité connue entre les sites.

Les différences d'altitude des sites peuvent également avoir un impact sur le taux attendu de valeurs suspectes si les mêmes paramètres de test sont utilisés. Les opérateurs de contrôle qualité doivent vérifier si les stations voisines sont situées à des altitudes sensiblement différentes ou dans des environnements sensiblement différents (par exemple si ce sont des stations à l'intérieur des terres alors que la station candidate est située sur la côte); si c'est le cas, l'utilisation de tests basés sur la station voisine la plus proche, par exemple, ne sera peut-être pas appropriée. On peut néanmoins remédier aux difficultés liées aux différences d'altitude des sites en comparant les anomalies des données vis-à-vis des normales climatiques des sites.

La durée des relevés climatologiques peut avoir un impact sur les PSQ et donc sur la fiabilité des tests. Par exemple, les PSQ générés par un site qui compte 30 ans de relevés seront beaucoup plus représentatifs que ceux générés par un site qui n'a que trois ans de relevés. Dans le contexte de l'exécution de statistiques spatiales, si les stations voisines sont des stations relativement récentes qui ne disposent que d'un court historique de relevés, cela aura une incidence sur les résultats du test. Les métadonnées des observations peuvent aider les opérateurs à déterminer la durée des relevés et l'efficacité attendue du test. Par exemple, le test de portée fondé sur le climat décrit dans l'[annexe 1](#) et l'[annexe 2](#) effectue des comparaisons par rapport aux extrêmes des relevés climatologiques locaux, de sorte qu'un site disposant d'un court historique de relevés sans large plage d'extrêmes devrait déclencher des suspicions de faux positifs plus fréquemment.

La densité du réseau d'observation a une incidence sur les tests de contrôle qualité qui utilisent des statistiques spatiales, par exemple, les tests de variabilité spatiale et les tests spatiaux fondés sur la cartographie (voir l'[annexe 1](#) et l'[annexe 2](#)). Il peut s'avérer nécessaire d'imposer des limites aux algorithmes spatiaux, afin que la distance entre la station cible et une station voisine ne puisse pas dépasser une distance «raisonnable», déterminée par un certain nombre de facteurs tels que la zone climatique, la topographie, la durée des relevés et le paramètre à tester. Pour les sites très reculés ou les réseaux très clairsemés, il peut être nécessaire de signaler le premier jour du mois comme valeur suspecte pour assurer la vérification manuelle des erreurs. D'autres méthodes et/ou paramètres peuvent être nécessaires pour valider une observation suspecte. Dans le cas de l'observation des précipitations, par exemple, on peut utiliser les images satellite, les images radar, la hauteur des cours d'eau ou l'écoulement fluvial. Une autre possibilité consiste à contacter directement l'observateur pour confirmer la quantité enregistrée, de préférence en temps réel ou quasi réel.

Les tests à variables multiples peuvent être utiles à l'opérateur de contrôle qualité. Par exemple, si le cumul des précipitations sur 24 heures pour un jour donné est noté comme étant élevé, on pourra comparer cette valeur aux observations horaires des précipitations, aux rapports sur le temps présent et le temps passé, aux types de nuages associés à de fortes précipitations, à l'humidité, aux données provenant d'autres instruments de mesure des précipitations (par exemple, les pluviomètres ou les diagrammes d'intensité des précipitations), aux images radar, aux images satellite, aux prévisions numériques du temps, aux analyses d'impact, aux bulletins des médias, etc. Cela suppose, bien entendu, que ces paramètres et images de substitution soient facilement accessibles. Il est à noter que ces approches à variables multiples peuvent aussi faciliter l'automatisation du processus de contrôle qualité.

3.4.3 **Comment traiter les données à fréquence élevée**

À l'heure actuelle, de nombreux SMHN exploitent des SMA qui produisent des données de résolution temporelle élevée (1 minute, 10 minutes), transmises de la station d'observation au siège à intervalles rapprochés. En plus des contrôles qualité automatiques effectués directement à la station, au niveau de l'unité centrale de traitement de la SMA, il est possible d'effectuer un contrôle qualité supplémentaire en temps quasi réel sur les données du système de gestion des données climatologiques. Ce contrôle supplémentaire permet d'éviter des erreurs grossières dans les valeurs quotidiennes ou mensuelles dérivées qui figurent dans les relevés climatologiques, dans la mesure où les données à résolution élevée à partir desquelles les valeurs quotidiennes ou mensuelles sont calculées ont déjà reçu un indicateur de qualité et/ou ont été corrigées.

Pour des raisons de traçabilité, les données originales à fréquence élevée doivent être stockées, ce qui nécessite des systèmes de stockage et des modèles de données complexes (ou tout au moins un journal d'audit contenant les modifications apportées). Le contrôle qualité des données à fréquence élevée est un processus qui prend énormément de temps. C'est pourquoi certains SMHN font le choix de cibler le contrôle qualité en temps quasi réel sur les valeurs horaires dérivées: les données à fréquence élevée ne sont examinées de plus près que si les valeurs horaires semblent douteuses. Il est toutefois souhaitable de procéder également à des tests de pic ou de courbe plate, car les pics, par exemple, peuvent ne pas être correctement pris en compte dans une valeur horaire.

Si cela est possible, il est recommandé que tous les paramètres des données à fréquence élevée fassent l'objet de contrôles qualité automatiques et que les tests de contrôle qualité comprennent des:

- Tests de cohérence;
- Tests de pic;
- Tests de changement rapide;
- Tests de courbe plate;
- Tests de domaine.

Comme on l'a indiqué dans la [section 2.5](#), les risques associés aux tests entièrement automatiques peuvent être atténués au moyen d'une analyse à variables multiples.

Contrôle qualité *a posteriori* dans le contexte de données à fréquence élevée

Dans certains cas, un changement significatif et inattendu de la valeur d'un paramètre de donnée à fréquence élevée, par exemple un pic dans les données, peut justifier de mener une enquête plus approfondie pour déterminer les circonstances plus générales de ce changement soudain. Un tel événement peut être dû par exemple à un phénomène météorologique violent authentique ou à un phénomène localisé tel qu'une microrafale, à un phénomène non météorologique tel qu'un feu de forêt à proximité, ou encore à un pic dû à une surtension interne.

Le contexte de l'événement doit être évalué à l'aide d'une analyse à variables multiples qui permettra de déterminer si l'événement est représentatif des conditions météorologiques au moment considéré ou s'il est dû à une autre cause.

3.4.4 **Contrôle qualité à valeur ajoutée**

Le contrôle qualité à valeur ajoutée peut prendre diverses formes, telles que l'estimation de valeurs de remplacement en cas de valeurs manquantes ou erronées⁹. Toutefois, l'examen de ces méthodes sort du cadre de ce document.

L'homogénéisation peut également être considérée comme une forme d'ajout de valeur. Elle implique de rectifier les relevés climatologiques lorsque cela est nécessaire, pour les débarrasser des effets de facteurs non climatiques (par exemple, un déplacement de station), de sorte que les données résultantes reflètent des variations non biaisées dues aux processus climatiques réels. Il est essentiel que le processus de contrôle et d'assurance qualité produise suffisamment de métadonnées fiables à l'appui de l'homogénéisation des données.

Les *Directives sur l'homogénéisation* (OMM-N° 1245) donnent des orientations sur le traitement des problèmes d'homogénéité. Le processus d'homogénéisation comprend généralement quatre étapes clés:

- Analyse des métadonnées des observations: Il s'agit, lors de cette étape, de rechercher les éventuels changements dans les mesures et d'examiner les procédures de contrôle qualité qui ont été effectuées.
- Création d'une série chronologique de référence: Pour créer cette série, on utilise généralement une moyenne pondérée des stations voisines (post-contrôle qualité), mais d'autres techniques, telles que l'analyse en composantes principales, peuvent également s'avérer utiles.
- Détection des points de rupture: Cette étape consiste à rechercher les défauts d'homogénéité en observant les divergences entre la série chronologique de référence et la série chronologique candidate.
- Ajustement des données: Au cours de cette étape, on décide quels points de rupture doivent être considérés comme des défauts d'homogénéité, en les comparant aux métadonnées disponibles et en se fiant aux avis d'experts. On corrige ensuite les discontinuités constatées dans les données pour que celles-ci correspondent aux conditions de la section homogène la plus récente.

Pour renforcer la traçabilité et la fiabilité des données, il est important de consigner chacun des ajustements effectués sur les données et de toujours conserver les données originales.

3.4.5 **Indicateurs de source et de qualité**

Même si, au moment de la publication de ce document, l'OMM n'avait pas encore édité d'orientations génériques en matière d'indicateurs de qualité, les SMHN devraient, pour des raisons de traçabilité, adopter une forme ou une autre de marquage par indicateurs de qualité de manière à pouvoir suivre la provenance et la source des données. Voici quelques suggestions à cet égard.

Les indicateurs de source indiquent le type de la source d'observation. Par exemple, ils établissent une distinction entre les observations manuelles et les observations automatiques, ainsi qu'entre les observations provenant du SMHN et celles issues de sources externes (qui peuvent être soumises à des pratiques d'observation et des exigences différentes en matière de degré d'exhaustivité des métadonnées). Les indicateurs de source peuvent également être utilisés pour comparer une même observation reçue par différents canaux: par exemple, des données sur les précipitations quotidiennes envoyées par voie électronique et sur un manuscrit papier. Ces valeurs ne concordent pas toujours, et dans ce cas, nécessitent une enquête plus approfondie. Les indicateurs de source permettent également de comparer les valeurs données

⁹ Ce type de substitution a toutes les chances d'être préférable à la simple mention «non disponible» dans le cas des précipitations, pour lesquelles une seule valeur manquante peut invalider le total d'un mois/an entier.

par des instruments situés au même endroit – par exemple, les valeurs des précipitations quotidiennes telles qu'elles ressortent des observations manuelles des précipitations et des relevés automatiques des pluviomètres et les valeurs provenant d'observations de tiers, telles que les jauges d'alerte de crue.

Les indicateurs de qualité livrent une indication qualitative du niveau de confiance dans les données. Ces indicateurs devraient être attribués à l'issue de tout contrôle qualité automatique effectué à la station (SMA, enregistreur de données, etc.), pendant le processus d'intégration/ de saisie (qui peut lui-même comporter une ou plusieurs étapes, comme indiqué ci-avant), ainsi que lors du contrôle en différé. Un indicateur de qualité type doit indiquer si une donnée a fait l'objet d'un contrôle qualité et, dans l'affirmative, si elle a été considérée comme valide, suspecte, douteuse ou erronée et si elle a été estimée ou modifiée.

Pour donner un exemple, l'utilisation d'un indicateur de qualité peut être justifiée lorsqu'une valeur maximale lue manuellement sur le thermomètre d'une station ne cadre pas avec les valeurs des stations voisines et que l'opérateur de contrôle qualité soupçonne une erreur de lecture. Dans ce cas, il peut être approprié d'indiquer que la donnée en question est une estimation brute ou, dans des circonstances plus extrêmes, une valeur suspecte ou incorrecte.

Les indicateurs de qualité peuvent également se référer spécifiquement au type de contrôle appliqué: contrôle à l'intégration, contrôle automatique (avec attribution automatique de l'indicateur de qualité), contrôle semi-automatique avec analyse à variables multiples par l'opérateur, etc.

Les valeurs d'indicateur attribuées, ainsi que les modifications qui leur sont éventuellement apportées, doivent être conservées dans un journal d'audit, et ces modifications doivent être accessibles à des fins d'analyse ou d'examen manuels.

Le modèle d'attribution des indicateurs ne doit pas être compliqué. Il est recommandé d'utiliser une série de sous-ensembles d'indicateurs correspondant aux différents aspects des métadonnées plutôt que d'établir une seule longue liste d'indicateurs englobant tous les scénarios de métadonnées possibles.

La quantité d'efforts consacrée à la planification du système d'indicateurs de qualité déterminera le degré de pérennité du système. Le système d'indicateurs doit être suffisamment souple pour pouvoir s'adapter aux changements futurs des technologies, des réseaux, etc. Un système renfermant plusieurs sous-ensembles d'indicateurs présente un certain nombre d'avantages, dont voici quelques exemples:

- Les sous-ensembles d'indicateurs peuvent être étendus si c'est nécessaire. Par exemple, un indicateur peut être ajouté à une série d'indicateurs existante pour signaler une nouvelle source de données, et une nouvelle série d'indicateurs peut être ajoutée pour décrire la qualité de l'exposition des instruments ou le contrôle qualité spécifique qu'a déjà réalisé un fournisseur de données externe.
- Des contrôles qualité rétrospectifs peuvent être effectués à mesure que le système de contrôle qualité se perfectionne.

Le *Guide des instruments et des méthodes d'observation* (OMM-N° 8) donne des exemples de types d'indicateurs de qualité. Il ne dresse pas une liste exhaustive de ces indicateurs, mais définit des orientations de base. Les indicateurs de qualité comprennent notamment le code correspondant:

- Au type de donnée;
- À la méthode d'acquisition;
- Au stade de validation.

Les indicateurs de qualité permettent d'établir des journaux d'audit et la provenance des données, c'est-à-dire de savoir si des modifications ont été apportées aux données,

par qui, comment et pourquoi. Ils facilitent également le suivi de la performance et de la responsabilisation des opérateurs de contrôle qualité. Par exemple, si un problème de formation a été constaté concernant un opérateur particulier, il est possible, grâce à l'utilisation des indicateurs de qualité, de retrouver toutes les corrections appliquées par cet opérateur qui sont en lien avec le problème de formation observé, de récupérer les données originales et de retraiter les données. Toute décision de marquer une donnée comme étant suspecte doit être étayée scientifiquement. Si l'opérateur n'est pas en mesure de prendre une décision raisonnable et étayée, il est recommandé de conserver les données telles quelles.

L'utilisation d'indicateurs de qualité est une composante obligatoire du contrôle qualité. Elle procure un certain niveau d'assurance qualité aux utilisateurs des données et les renseigne sur les sources des données et les résultats des contrôles qualité. Elle garantit que les données suspectes ou erronées sont non pas supprimées, mais signalées (comme indiqué précédemment, les données suspectes ou erronées doivent être conservées, mais ne doivent pas figurer dans les relevés climatologiques destinés à l'utilisateur). Enfin, elle offre aux utilisateurs des données la possibilité de choisir le niveau de qualité des données qui convient à leurs besoins.

Données provenant de sources autres que les SMHN

Les données qu'un SMHN a en sa possession peuvent avoir été collectées par des réseaux externes, qui en conservent la propriété. La qualité de ces données est amenée à varier en raison de différences dans les pratiques d'observation et/ou la quantité de métadonnées d'observation disponibles. Pour cette raison, les données peuvent être considérées comme appartenant à une classe de réseau particulière, à savoir un réseau exploité par le SMHN ou un réseau exploité par une tierce partie. La qualité des données (et des métadonnées) étant susceptible de varier en fonction de leur origine, il est recommandé d'utiliser un indicateur approprié (ou d'autres moyens de distinction) pour indiquer la source et décrire les traitements éventuellement appliqués (y compris, les contrôles qualité éventuellement effectués par le fournisseur de données). On trouvera des indications plus détaillées sur la gestion des données d'origine externe dans le [Manuel sur le Cadre mondial pour la gestion de données climatologiques de qualité](#) (OMM-N° 1238).

3.4.6 Perspectives futures pour le contrôle qualité

Les processus de contrôle qualité décrits ci-dessus sont principalement fondés sur des algorithmes qui utilisent les données à la surface du sol et sur la vérification manuelle de ces données par l'opérateur, qui s'appuie dans cette mission sur des données et produits satellite et radar et d'autres informations et outils.

À l'avenir, les processus de contrôle qualité pourraient intégrer ces catégories de données de télédétection ainsi que tout autre produit dérivé directement dans les algorithmes de vérification, dans le cadre d'une stratégie de renforcement de l'automatisation. À partir des données de télédétection, il pourrait devenir possible de vérifier intégralement et de façon automatique les paramètres suivants, en particulier pour les zones où les stations d'observation sont réparties de manière éparse et où les comparaisons spatiales avec d'autres stations de surface ne sont pas possibles:

- Rayonnement solaire;
- Nébulosité;
- Couverture de neige;
- Température (moyenne quotidienne);
- Précipitations (somme quotidienne).

Les produits de la PNT (par exemple, les champs de première approximation) peuvent livrer des informations utiles pour le contrôle en temps quasi réel de la qualité des données mesurées/observées, qui peuvent aider à détecter et corriger rapidement les données erronées. À partir des prévisions météorologiques propres à un site, il est possible de calculer l'écart historique (anomalies) entre les valeurs prévues et les valeurs observées. Si l'écart entre la prévision et l'observation est supérieur à un certain seuil, l'observation pourra être considérée comme suspecte. L'avantage de cette approche est que les écarts sont généralement faibles, de sorte que la valeur seuil à partir de laquelle une observation est considérée comme suspecte peut être définie de manière assez stricte. Revers de la médaille, si la prévision est erronée, une observation correcte peut être marquée comme valeur suspecte. De même, les technologies de réanalyse peuvent faciliter la détection d'anomalies dues à des données suspectes.

Il y aura également lieu de suivre les progrès réalisés dans le domaine des technologies de l'information et des communications (TIC), tels que l'apprentissage automatique et les approches par système expert, qui pourraient contribuer à accélérer l'automatisation de l'assurance et du contrôle qualité.

3.5 **Étape 4: Archivage final**

Pour les archives électroniques, au moment de l'archivage définitif, il est important de s'assurer de ce qui suit:

- Les données à archiver sont assorties d'un journal d'audit exhaustif des modifications, et/ou de versions distinctes des données permettant d'identifier les contrôles qualité et ajouts de valeur appliqués;
- Le message électronique original est conservé et peut être facilement récupéré avec la valeur d'observation et les métadonnées;
- La provenance, indication comprenant toutes les modifications apportées aux données, peut être extraite si nécessaire;
- Les données sont adaptées à leur finalité, c'est-à-dire qu'elles conviennent aux applications climatologiques;
- La qualité des données est connue, les tests de qualité effectués ont été dûment consignés et des indicateurs de qualité ont été attribués (pour renforcer la confiance du public dans le processus de contrôle qualité, il peut être utile de tenir une «fiche d'information» décrivant les tests effectués);
- Des procédures de gouvernance appropriées ont été mises en œuvre, de sorte que les données ne puissent pas être modifiées sans autorisation. Note: Des modifications peuvent être apportées après l'archivage, par exemple si un utilisateur repère une erreur manifeste dans les données, ou si la méthode de contrôle qualité est modifiée ultérieurement et que les données sont contrôlées rétroactivement à l'aide de nouvelles méthodes.

Une migration technologique régulière est nécessaire pour garantir que les archives finales restent accessibles pendant une durée illimitée.

Pour le stockage des documents papier, le système de gestion des documents doit disposer d'un environnement contrôlé qui:

- Utilise des matériaux de stockage sans acides;
- Soit protégé contre le feu et l'eau;
- Soit exempt de parasites;
- Contienne un répertoire complet des relevés détenus;

- Fasse l'objet de procédures de manipulation appropriées;
- Préserve tous les relevés;
- Contienne des dispositifs de sécurité adéquats.

Pour de plus amples détails sur la gestion des documents papier, se référer aux [Directives sur les bonnes pratiques en matière de sauvetage des données climatologiques](#) (OMM-N° 1182).

3.6 **Étape 5: Récupération en cas de sinistre**

L'étape de la récupération en cas de sinistre consiste à s'assurer que, en cas de sinistre entraînant la perte ou la corruption totales ou partielles des données, celles-ci pourront être récupérées. Pour les bases de données ou les données électroniques, il faut s'assurer que l'on dispose de plusieurs exemplaires des données et de sauvegardes hors site, et que des dispositions ont été prises pour pouvoir restaurer les archivages de façon efficace et complète à partir de ces sauvegardes; le [Manuel sur le Cadre mondial pour la gestion de données climatologiques de qualité](#) (OMM-N° 1238) donne de plus amples précisions sur les exigences de sécurité et de récupérabilité des données climatologiques des SMHN. Pour les documents papier, ces exigences comprennent un stockage et une sécurité adéquats (voir ci-dessus); il est fortement recommandé de scanner et numériser tous les documents papier.

4. **RÔLE DU RESPONSABLE DE L'ASSURANCE QUALITÉ**

Étant donné qu'il est de plus en plus important de disposer de données climatologiques fiables et dignes de confiance et que chaque SMHN a ses propres procédures et structures, la gestion du processus d'assurance qualité est très importante. Il est donc fortement recommandé que chaque SMHN désigne un responsable de l'assurance qualité des données, à qui il incombera de vérifier que toutes les composantes du processus d'assurance et de contrôle qualité des données climatologiques sont correctement appliquées et que tous les problèmes sont identifiés et traités de manière appropriée. Cette personne doit être au fait des aspects détaillés des processus tout au long du cycle de vie des données et avoir une vue d'ensemble de l'état du processus d'assurance qualité dans l'ensemble du SMHN, même si certaines opérations d'assurance et de contrôle qualité au sein du SMHN sont réparties entre différents secteurs organisationnels. Les responsables de l'assurance qualité devraient régulièrement échanger leurs données d'expérience au niveau international et se tenir au courant des nouvelles avancées scientifiques pertinentes. Dans l'idéal, les processus d'assurance et de contrôle qualité devraient viser à s'aligner sur une norme ISO reconnue, car cela permettrait de renforcer la confiance dans les processus et de s'assurer de la conformité à des normes de qualité reconnues.

Il est recommandé que le responsable de l'assurance qualité des données exerce les responsabilités recensées ci-dessous pour garantir la qualité de l'enregistrement, de la transmission et de l'archivage des données. Les tâches peuvent être effectuées par différentes sections du SMHN ou par des tiers externes, mais le responsable de l'assurance qualité doit s'assurer qu'elles sont réalisées de manière à ne pas compromettre l'intégrité des relevés climatologiques. Les métadonnées des observations et la documentation associée à ces procédures doivent être tenues à la disposition du responsable du contrôle qualité afin de garantir le caractère éclairé du processus de contrôle.

Responsabilités:

- Le responsable de l'assurance qualité des données doit s'assurer que les techniques d'instruments sont adaptées à leur finalité en:
 - Veillant à ce que les instruments soient correctement testés au moyen de tests de performance et d'acceptation dûment consignés, avant les essais sur le terrain;
 - Réalisant des essais parallèles sur le terrain. Pour les instruments qui remplacent une technologie plus ancienne, des essais parallèles sur le terrain doivent être menés pendant au moins un an (et si possible pendant deux ans ou plus) afin de quantifier correctement les écarts de performance entre les instruments. Idéalement, il faudrait que ces essais soient réalisés avant la mise en service opérationnelle des nouveaux instruments, pour s'assurer que les instruments fonctionnent de manière fiable et homogène dans le cadre des tolérances fixées;
 - Veillant à ce que les résultats des essais parallèles sur le terrain soient conservés et à ce qu'ils soient accessibles aux utilisateurs des données et aux clients sur demande;
 - S'assurant que l'instrument est installé de façon opérationnelle conformément aux spécifications de l'OMM ou du pays, et (s'il y a lieu) qu'il a été étalonné par rapport à une norme avant sa mise en service.
- Pour les instruments déjà en service, le responsable de l'assurance qualité des données doit:
 - S'assurer que les données produites par les instruments se situent dans les limites de tolérance grâce à un programme d'inspection régulier;
 - S'assurer que les observations sont effectuées à la fréquence requise;

- Consigner et surveiller régulièrement l'exposition des instruments sur le site, en notant tout changement significatif susceptible d'affecter la qualité des données (ces informations étant transmises au système de métadonnées des observations du SMHN);
- Vérifier que l'entretien de routine est planifié et a bien lieu.
- Pour les observations enregistrées manuellement, le responsable de l'assurance qualité des données doit:
 - Dans le cas où les observations sont consignées manuellement dans le journal de terrain de la station (registre d'observations), s'assurer que le personnel de la station vérifie l'exactitude de ces journaux¹. Pour les sites comptant plus d'un observateur, cette responsabilité incombe généralement au responsable de la station;
 - Enregistrer les métadonnées relatives à tout problème portant sur la qualité des instruments ou des données dans le journal de terrain ou dans un autre registre d'observations;
 - Dans la mesure du possible, veiller à ce que les observateurs et les réseaux d'observation utilisent des pratiques et des procédures normalisées (c'est-à-dire des pratiques et des méthodes d'observation homogènes) et à ce que les conseils concernant les changements requis dans les procédures d'observation soient largement diffusés et bien documentés;
 - Vérifier que des formations de mise à niveau sont organisées et que les nouveaux observateurs reçoivent une formation lorsqu'ils prennent leurs fonctions.
- En ce qui concerne l'intégration et l'archivage des données, le responsable de l'assurance qualité des données doit:
 - S'assurer que les données sont accompagnées de métadonnées adéquates et d'un journal d'audit;
 - S'assurer que les processus envoient des informations de retour au site et/ou au fournisseur des observations concernant les problèmes de qualité systémiques rencontrés;
 - Veiller à ce que des contrôles des flux de données soient en place, afin de remédier à des problèmes potentiels tels que: la non-réception des données attendues, la corruption des données, les problèmes d'horodatage, etc. (le responsable de l'assurance qualité des données doit être en communication permanente avec les responsables informatiques du SMHN pour s'en assurer);
 - S'assurer qu'il existe des procédures de récupération des données efficaces;
 - S'assurer que des procédures de sauvegarde et de récupération en cas de sinistre sont en place, qu'elles sont consignées et qu'elles sont parfaitement comprises (là encore, une liaison étroite avec le personnel informatique du SMHN est nécessaire);
 - Veiller à ce que les données reçues sur un support papier soient numérisées et, dans la mesure du possible, converties en images numériques, et à ce que les documents papier soient archivés.

¹ Ceci s'applique spécifiquement aux observations enregistrées sur des formulaires papier. Dans les stations où les données observées manuellement sont saisies dans un système électronique, une vérification doit également être effectuée sous une forme ou une autre.

- En ce qui concerne la qualité des données, le responsable de l'assurance qualité des données doit:
 - S'assurer que les opérateurs de contrôle qualité comprennent parfaitement les procédures de contrôle qualité du SMHN et sont en mesure de les appliquer (pour garantir le maintien de normes de contrôle qualité élevées, il est suggéré de procéder à des audits périodiques pour vérifier que les opérateurs appliquent des procédures normalisées et homogènes);
 - Veiller à l'exécution intégrale, continue et sans retard de toutes les procédures et de tous les tests de contrôle qualité;
 - Veiller à ce que les informations relatives à la qualité soient correctement enregistrées dans les métadonnées;
 - S'assurer que toutes les données sont assorties d'un indicateur de qualité, conformément aux procédures d'attribution d'indicateurs de qualité du SMHN;
 - Veiller à ce que toute la documentation relative aux procédures soit tenue à jour et soit accessible.
-

ANNEXE 1. TESTS D'ASSURANCE ET DE CONTRÔLE QUALITÉ: VUE D'ENSEMBLE

Cette annexe donne une vue d'ensemble non exhaustive des types de test qui peuvent être mis en œuvre au titre de l'assurance et du contrôle qualité des données climatologiques, accompagnée d'indications de leur importance relative suggérée. Ces tests sont décrits de façon plus détaillée dans l'[annexe 2](#).

Tableau 1. Sélection de tests d'assurance et de contrôle qualité et indication de leur importance suggérée

Légende: O = Obligatoire, R = Recommandé, F = Facultatif

Tests de contraintes: Tests visant à garantir que les observations sont techniquement et scientifiquement plausibles au regard des limites théoriques et climatologiques, des spécifications des capteurs ou des limites de la base de données en matière d'intégration des données.

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test de portée fondé sur les capteurs	Détecte les observations qui se situent en dehors des limites théoriques ou des spécifications du capteur		O
Test de portée au niveau de la base de données	Détecte des valeurs qui se situent en dehors de la plage des critères d'intégration du système de stockage	Ce test est effectué au moment de l'intégration des données dans le système de stockage	O
Test de domaine	Détermine si la valeur météorologique est scientifiquement possible, par exemple, $T > 70\text{ °C}$		O

Tests de cohérence: Tests utilisant des comparaisons avec d'autres paramètres pour faire en sorte que les combinaisons incohérentes, improbables ou impossibles soient rejetées ou signalées comme suspectes. Une enquête manuelle peut ensuite être menée pour déterminer la validité des valeurs suspectes.

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test infraquotidien	Ce test est effectué pour déterminer s'il existe des schémas cohérents entre les valeurs enregistrées lors des observations infraquotidiennes, par exemple les observations trihoraires, et les valeurs quotidiennes		O
Test de comparaison de la température minimale quotidienne et de la température minimale au sol quotidienne	Compare les températures minimales quotidiennes et les températures minimales au sol quotidiennes		F

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test de différence entre la pression moyenne horaire au niveau de la mer et la pression horaire au niveau de la station	Ce test sert à déterminer si la différence entre la pression moyenne au niveau de la mer et la pression au niveau de la station change de façon significative entre deux relevés consécutifs	Ce test peut être réalisé à partir d'observations horaires, trihoraires, etc.	F
Test de comparaison multisources des précipitations	Vérifie la cohérence entre les données provenant de différentes sources	Ces sources peuvent être des instruments différents, ou des canaux de communication différents à partir du même instrument	R
Test spatial de précipitations nulles	Détermine s'il y a des cas où des précipitations importantes sont enregistrées sur un site, mais pas sur les sites voisins, et vice versa. Le cas échéant, les opérateurs de contrôle qualité devront faire appel à leur expérience pour s'assurer que les stations voisines sont suffisamment représentatives du climat de la région.	En général, cela indique que les précipitations ont été enregistrées le mauvais jour ou que la valeur est un total cumulé	O
Test de voisinage insuffisant	Vérifie s'il existe un nombre suffisant de stations à une distance raisonnable de la station candidate pour effectuer des tests spatiaux	Ce test est généralement effectué le premier jour du mois et pour les données sur les précipitations	R
Test de la période des précipitations	Ce test recherche les surestimations et les sous-estimations potentielles des accumulations de précipitations, c'est-à-dire qu'il détermine s'il y a incompatibilité entre la valeur de la période du relevé et les dates réelles de précipitations nulles indiquées	Test conçu pour les relevés de précipitations	R
Test de concordance des trajectoires	Vérifie si deux ou plusieurs éléments, ou des occurrences du même élément dans deux stations voisines, augmentent et diminuent en tandem	Les deux éléments sont supposés augmenter et diminuer en tandem. Ce test est un moyen très efficace de déterminer si les valeurs doivent être rejetées ou signalées comme suspectes.	R
Test de cohérence de la température maximale (minimale) de l'air	Vérifie la cohérence entre la température maximale (minimale) quotidienne de l'air et les observations infraquotidiennes	Il s'agit d'une extension des tests infraquotidiens mentionnés ci-dessus	R
Test astronomique – temps d'insolation	Détermine la différence entre la durée d'insolation et la durée du jour calculée		O
Test de cohérence entre la température de l'air et la température du thermomètre mouillé	Compare la température de l'air à la température du thermomètre mouillé		O

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test de cohérence entre la température du thermomètre mouillé et la température du point de rosée	Détermine la différence entre la température du thermomètre mouillé et la température du point de rosée	Le test échoue si la différence est inférieure à zéro. Utilisé avec des observations manuelles.	F
Test de revérification température du thermomètre mouillé/ température du point de rosée	Le test vérifie la température du thermomètre mouillé par rapport à la température du point de rosée en recalculant la température du point de rosée à partir de la température du thermomètre mouillé de l'air	Utilisé avec des observations manuelles	F
Test de cohérence de la température du point de rosée et de la température de l'air	Détermine si la température du point de rosée est inférieure ou égale à la température de l'air		R
Test de cohérence de la visibilité	Vérifie la cohérence de la visibilité en croisant le code du temps présent et les indicateurs de phénomènes météorologiques (brouillard, tempête de sable, brume, tempête de poussière)	Visibilité horizontale. Utilisé avec des observations manuelles.	R
Test de cohérence de la nébulosité totale	Vérifie la cohérence entre la nébulosité totale et divers éléments	Utilisé avec des observations manuelles	R
Test des indicateurs de phénomènes météorologiques quotidiens	Vérifie la cohérence entre les indicateurs de phénomènes météorologiques quotidiens et les codes des tableaux infraquotidiens et divers autres éléments quotidiens		R
Test de cohérence du temps présent	Vérifie la cohérence entre les codes du temps présent figurant dans les tableaux infraquotidiens et divers autres éléments quotidiens		R
Test de la température du sol	Vérifie la cohérence des températures du sol à différentes profondeurs		R
Test de cohérence des messages	Dans les cas où plusieurs messages sont reçus pour une même observation d'un paramètre, le test vérifie si les messages contiennent tous la même valeur	S'applique à tous les paramètres	O
Test des données maximales/ minimales sur une minute	Compare la température maximale ou minimale indiquée avec la valeur correspondante sur 1 minute au moment du maximum ou du minimum de température, afin de s'assurer que les valeurs correspondent	Maximum quotidien, minimum quotidien	R

Tests heuristiques: Tests qui s'appuient sur l'expérience et la connaissance des processus, techniques et instruments d'observation pour détecter des valeurs incohérentes, improbables ou impossibles et les signaler comme suspects. Une enquête manuelle peut ensuite être menée pour déterminer la validité des valeurs suspects.

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test d'humidité relative (dessèchement de la mèche du thermomètre mouillé)	Test visant à déterminer si la mèche du thermomètre mouillé a séché		R

Tests de transmission des données: Tests permettant de s'assurer que les observations qui ne correspondent pas au programme d'observations prévu sont soit rejetées, soit signalées comme suspects.

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test des observations reçues du futur	Compare l'heure locale de l'observation et l'heure de réception de l'observation		R
Test de différences de périodes	Détermine si les périodes coïncident avec les relevés existants	S'applique aux observations quotidiennes	O
Test de longues périodes	Détermine si les périodes sont excessivement longues (plusieurs jours) ou inférieures à un jour	S'applique aux relevés quotidiens	R
Test de chevauchement avec les périodes futures	Test effectué pour identifier les cas où une observation est envoyée en retard, et après la réception d'une observation d'un jour ultérieur, afin de garantir la cohérence de la période d'observation déclarée		R
Test de réitération	Vérifie si la valeur, la période ou l'indicateur de qualité d'une observation ont changé depuis la dernière mise à jour	S'applique à tous les paramètres	R

Tests statistiques: Tests qui effectuent des analyses statistiques des données historiques pour détecter les valeurs incohérentes, improbables ou impossibles et les signaler comme suspects. Une enquête manuelle peut ensuite être menée pour déterminer la validité des valeurs suspects.

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test de portée fondé sur le climat	Compare la valeur météorologique de l'observation avec les extrêmes climatologiques supérieur et inférieur	Des seuils peuvent être définis pour tenir compte des variations saisonnières des observations	O
Test de courbe plate	Vérifie la durée d'une série de valeurs météorologiques qui sont identiques, pour déterminer si le paramètre est resté inchangé pendant une durée improbable		R

<i>Nom du test</i>	<i>Description succincte</i>	<i>Notes</i>	<i>Importance suggérée</i>
Test de changement rapide	Vérifie que la différence entre l'observation précédente et l'observation présente ne dépasse pas un seuil raisonnable		R
Test de pic	Compare une observation météorologique déterminée avec les valeurs précédentes et suivantes	Semblable au test de changement rapide; cependant, le test de pic cherche à mettre en évidence une hausse puis une baisse (ou une baisse puis une hausse) improbable	R
Test de fréquence d'arrondis	Recherche les cas d'arrondis excessifs des valeurs	S'applique aux observations manuelles marquées par une tendance de l'opérateur à arrondir les valeurs au lieu de les interpoler	F
Test de variabilité spatiale	Climatologie quotidienne comparant les différences entre la station considérée et les stations voisines		F
Test spatial fondé sur la cartographie	Compare la valeur météorologique avec les valeurs des stations environnantes en utilisant une technique d'analyse cartographique telle que l'analyse de Barnes	Le test spatial fondé sur la cartographie est utilisé pour pondérer les valeurs des stations voisines de manière à estimer une valeur à l'emplacement de la station candidate. L'opérateur de contrôle qualité peut aussi évaluer visuellement la probabilité d'une valeur à un endroit suspect en la comparant aux valeurs des stations environnantes, telles qu'indiquées sur une carte contenant des informations topographiques.	R
Test spatial de régression linéaire	Compare la valeur météorologique avec les valeurs des stations environnantes en effectuant une régression linéaire		F
Test de variabilité spatiale par régression linéaire	Utilise la variabilité des stations voisines plutôt que les estimations des erreurs types pour calculer les limites de test	Variante du test spatial de régression linéaire et du test de variabilité spatiale	F
Test de régression linéaire sur plusieurs jours	Compare la climatologie des différences entre la station considérée et les stations voisines sur une période de plusieurs jours	Semblable au test spatial de régression linéaire, mais s'applique à plusieurs jours	R
Test maximum/minimum	Vérifie que la différence entre les températures maximales et minimales est réaliste (c'est-à-dire supérieure/égale à zéro et inférieure à une limite supérieure)	La limite supérieure est calculée à partir d'une climatologie utilisant au moins 5 ans de données (de préférence 30 ans)	R

Sources: «QMS Test Specification» (Bureau météorologique australien (document interne)), [Climate Data Management System Specifications](#) (WMO-No. 1131), [Guide des pratiques climatologiques](#) (OMM-N° 100)

ANNEXE 2. TESTS D'ASSURANCE ET DE CONTRÔLE QUALITÉ: PRÉSENTATION DÉTAILLÉE

1. TESTS DE CONTRAINTES

Tests visant à garantir que les observations sont techniquement et scientifiquement plausibles au regard des limites théoriques et climatologiques, des spécifications des capteurs ou des limites de la base de données en matière d'intégration des données.

1.1 Test de portée fondé sur les capteurs

Description succincte: Détecte les observations qui se situent en dehors des limites théoriques ou des spécifications du capteur.

Paramètres sur lesquels porte le test: Température, humidité, pression barométrique, vent.

Description détaillée: Pour les observations automatiques, le test de portée fondé sur les capteurs est généralement effectué à la station. Les limites sont généralement fixées par le fabricant, souvent avec l'aide du SMHN. Le capteur mesure à intervalles d'une seconde, et si la valeur dépasse la limite spécifiée, la valeur est exclue. Un nombre minimum d'échantillons valides doit être constitué avant qu'une valeur puisse être transmise – le nombre d'échantillons valides pouvant varier d'un fabricant à un autre. De ces échantillons valides, on extrait un maximum, un minimum et une moyenne. Des tests sont également effectués pour les observations manuelles; par exemple, lors de la réinitialisation de thermomètres à liquide sous verre, les valeurs de réinitialisation sont comparées à la température de l'air ambiant. Dans les stations d'observation dotées de stations manuelles et automatiques, on compare les lectures du thermomètre à liquide sous verre avec celles des capteurs de température, en appliquant des tolérances raisonnables.

Fréquence du test: Selon la fréquence d'observation.

1.2 Test de portée au niveau de la base de données

Description succincte: Détecte les observations qui se situent en dehors de la plage des critères d'acceptation pour l'intégration dans la base de données (contraintes techniques de la base de données).

Paramètres sur lesquels porte le test: Tous les paramètres stockés dans la base de données.

Description détaillée: Ce test peut être effectué en plusieurs points, par exemple au point de première entrée de la valeur observée dans le SMHN, puis au point d'intégration dans le système de gestion des données climatologiques. L'objectif est d'éviter les valeurs générées en conséquence d'un oubli, par exemple lors d'une correction ou de l'intégration de données dans une base de données. Exemples: température > 80 °C, pression négative.

Fréquence du test: Au moment où les données sont intégrées dans le système de gestion des données climatologiques et dans toute base de données située en amont.

1.3 **Test de domaine**

Description succincte: Détermine si la valeur météorologique se situe dans le domaine de la vraisemblance scientifique.

Paramètres sur lesquels porte le test: Température, humidité, pression barométrique, vent, etc.

Description détaillée: Ce test vérifie si la valeur observée se situe entre le 0,3^e centile et le 99,7^e centile de toutes les valeurs des 30 dernières années. Il s'applique aux observations manuelles et automatiques. Lorsque l'observation est une valeur cumulative, on la divise par la période pour la tester en termes de valeur par jour.

Fréquence du test: Quotidienne.

2. **TESTS DE COHÉRENCE**

Tests destinés à garantir que les valeurs incohérentes, improbables ou impossibles sont soit rejetées, soit signalées comme suspectes. Une enquête manuelle peut être menée pour déterminer la validité des valeurs suspectes.

2.1 **Test infraquotidien**

Description succincte: Ce test est effectué pour déterminer s'il existe des schémas cohérents entre les valeurs enregistrées lors des observations infraquotidiennes, par exemple les observations trihoraires, et les valeurs quotidiennes.

Paramètres sur lesquels porte le test: Températures maximales et minimales quotidiennes de l'air, température du thermomètre sec.

Description détaillée: Pour la température minimale quotidienne, la différence entre la valeur minimale enregistrée et la température sur trois heures la plus basse des 24 heures précédentes est calculée et testée.

Pour la température maximale quotidienne, la différence entre la valeur maximale enregistrée et la température sur trois heures la plus élevée des 24 heures suivantes est calculée et testée.

Dans les deux cas, une erreur de lecture de 0,5 °C est autorisée pour tenir compte des tolérances de l'instrument et d'observation dans les observations individuelles. Les différences significativement inférieures à zéro doivent être signalées, de même que les différences très importantes.

Fréquence du test: Quotidienne.

2.2 **Test de comparaison de la température minimale quotidienne et de la température minimale au sol quotidienne**

Description succincte: Compare les températures minimales quotidiennes et les températures minimales au sol quotidiennes.

Paramètres sur lesquels porte le test: Température minimale terrestre quotidienne, température minimale quotidienne de l'air.

Description détaillée: Ce test vérifie la cohérence entre la température minimale quotidienne, qui est prise à l'intérieur de l'abri (par exemple, un abri Stevenson), et la température minimale au sol, qui est prise au sol. Le minimum au sol est presque toujours inférieur au minimum sous abri. Le test permet généralement de détecter les problèmes portant sur la température minimale au sol, cette mesure pouvant être problématique si elle n'est pas réalisée dans les règles ou si des bulles se forment dans le thermomètre.

Ce test peut générer des faux positifs pour les sites qui affichent un niveau élevé de métaux lourds dans le sol (par exemple, de l'oxyde de fer).

Fréquence du test: Quotidienne.

Ordre d'exécution du test: Le test est lancé après la mesure de la température minimale de l'air.

2.3 **Test de différence entre la pression moyenne horaire au niveau de la mer et la pression horaire au niveau de la station**

Description succincte: Ce test sert à déterminer si la différence entre la pression moyenne au niveau de la mer et la pression au niveau de la station change de façon significative entre deux relevés consécutifs.

Paramètres sur lesquels porte le test: Pression moyenne horaire au niveau de la mer et pression horaire au niveau de la station.

Description détaillée: Ce test sert à déterminer si la différence entre la pression moyenne au niveau de la mer et la pression au niveau de la station change de façon significative entre deux relevés consécutifs. Il peut être réalisé à partir d'observations horaires, trihoraires, etc., et permet de détecter les erreurs de transposition et de calcul.

Note: Dans certaines stations qui utilisent la température présente dans l'algorithme pour déduire la pression moyenne au niveau de la mer de la pression au niveau de la station, la relation entre ces deux paramètres peut changer de façon substantielle au cours du cycle diurne de la température. Ce problème se pose principalement dans les stations de haute altitude, qui voient de fortes inversions thermiques se développer à basse altitude en hiver.

Fréquence du test: À chaque observation.

2.4 **Test de comparaison multisources des précipitations**

Description succincte: Vérifie la cohérence entre les données provenant de différentes sources, à savoir des instruments différents, ou les différents canaux de transmission ou d'intégration des données d'un même instrument.

Paramètres sur lesquels porte le test: Précipitations*.

Description détaillée: Ce test vérifie la cohérence entre les données provenant d'une source et les données provenant d'une autre source située au même endroit. Ces deux sources peuvent être des instruments différents, ou les différents canaux de transmission ou d'intégration des données d'un même instrument.

Si les deux sources sont les canaux de transmission des données du même instrument, les deux valeurs devraient être identiques. Lorsqu'il existe une différence entre les deux et qu'il n'est pas possible de prendre une décision objective quant à la valeur correcte, cela doit être signalé au moyen d'un indicateur approprié. Le même commentaire s'applique dans le cas de données provenant de différents instruments. Dans les deux cas, la réalisation d'autres tests et analyses (portant par exemple sur la répartition spatiale et temporelle) aidera à déterminer la valeur

correcte la plus probable. En marge de ce test, il y aura lieu d'effectuer d'autres analyses pour repérer les stations qui enregistrent des différences systématiques et déterminer les raisons de ces différences.

Fréquence du test: Quotidienne.

*Note: La même approche peut être utilisée pour d'autres paramètres mesurés à l'aide de plusieurs capteurs.

2.5 **Test spatial de précipitations nulles**

Description succincte: Détermine s'il y a des cas où des précipitations importantes sont enregistrées sur un site, mais pas sur les sites voisins, et vice versa. En général, cela indique que les précipitations ont été enregistrées le mauvais jour ou que la valeur représente un total cumulé sur plusieurs jours.

Paramètres sur lesquels porte le test: Précipitations quotidiennes.

Description détaillée: Ce test détecte les cas dans lesquels une station enregistre des précipitations alors que toutes les stations voisines affichent une valeur de zéro, et vice versa. Cela s'explique généralement par le fait que la valeur a été enregistrée le mauvais jour. Il peut également s'agir d'un total cumulé dans une ou plusieurs stations, par exemple un total sur trois jours relevé le lundi dans une station manuelle qui n'a pas de personnel pendant le week-end. L'anomalie n'est pas nécessairement détectée par les tests spatiaux, qui ne saisissent pas toujours de manière adéquate la variabilité spatiale associée aux averses. Il convient de noter que l'efficacité de ce test peut être amoindrie si la station «candidate» est soumise à un régime climatique différent de celui des stations considérées comme «voisines».

Ce test est supposé produire un faible taux de faux échecs et non pas un taux élevé de valeurs suspectes potentielles; par conséquent, de manière générale, la différence de valeur de précipitations entre la station candidate et les stations environnantes devrait être significative.

Note: En complément de ce test, il est possible de contrôler les valeurs nulles sur plusieurs jours pendant une période prolongée dans les situations où les stations voisines enregistrent régulièrement de petites quantités de précipitations qui ne sont pas assez importantes pour déclencher le test au cours d'une journée donnée. Cela peut se produire, par exemple, lorsqu'un pluviomètre se bloque. Dans ce cas de figure, bien que la valeur nulle individuelle ne s'écarte pas suffisamment des valeurs des stations voisines pour être considérée comme suspecte, les valeurs cumulées feront apparaître le problème.

Fréquence du test: Exécuté quotidiennement, mais analysé par blocs mensuels.

2.6 **Test de voisinage insuffisant**

Description succincte: Vérifie s'il existe un nombre suffisant de stations voisines pour effectuer des tests spatiaux. Ce test est généralement effectué le premier jour du mois et pour les données sur les précipitations.

Paramètres sur lesquels porte le test: Précipitations.

Description détaillée: Ce test cherche les cas où il n'y a pas suffisamment de stations voisines pour effectuer la série de tests spatiaux. Ce test est généralement effectué le premier jour du mois (les autres jours ne sont pas testés). Si le résultat du test est un «échec», l'opérateur de contrôle qualité vérifiera tous les jours du mois.

Fréquence du test: Uniquement le premier jour du mois.

Ordre d'exécution du test: À exécuter avant les autres tests spatiaux sur les précipitations.

2.7 Test de la période des précipitations

Description succincte: Test de cohérence qui recherche les surestimations et les sous-estimations potentielles des accumulations de précipitations.

Paramètres sur lesquels porte le test: Précipitations quotidiennes.

Description détaillée: Ce test détermine si la valeur de la période d'un relevé (le nombre de jours sur lequel le relevé des précipitations est cumulé) est supérieure ou inférieure au nombre réel de jours sur lequel les données relatives aux précipitations sont enregistrées comme valeur cumulée. Ce test est spécifiquement conçu pour les relevés de précipitations. Il comprend deux volets: surestimations et sous-estimations.

Exemple de surestimation: La valeur de la période d'un relevé est supérieure aux dates déclarées de précipitations nulles réelles. Par exemple, si la valeur de la période déclarée le 4 mai est de trois jours, mais qu'une valeur de précipitations non nulle a été enregistrée pour la dernière fois le 2 mai, l'écart de période réel (cumul de précipitations) pourrait n'être que de deux jours – les 3 et 4 mai. Dans cet exemple, on commencera par vérifier l'exactitude de la valeur des précipitations le 2 mai (il se peut que les précipitations aient été enregistrées le mauvais jour). S'il s'avère que les données sur les précipitations du 2 mai et du 4 mai sont correctes, il est probable qu'il y ait une erreur concernant la période. Pour déterminer si c'est effectivement le cas, l'opérateur de contrôle qualité utilise des données spatiales et d'autres outils tels que les radars pour confirmer les analyses. On évaluera également les relevés de la station pour s'assurer que les données ont été correctement numérisées.

Exemple de sous-estimation: La valeur de la période d'un relevé est inférieure aux dates déclarées de précipitations nulles réelles. Par exemple, si la période d'accumulation des précipitations déclarée le 6 mai est de deux jours, mais que le dernier relevé de précipitations non nulles remonte au 3 mai, l'écart de période réel est de trois jours.

Fréquence du test: Quotidienne.

2.8 Test de concordance des trajectoires

Description succincte: Vérifie si deux ou plusieurs éléments, ou des occurrences du même élément dans deux stations voisines, présentent le comportement synchrone attendu (c'est-à-dire augmentent et diminuent en tandem). Ce test est très efficace et présente un faible taux de fausses erreurs.

Description détaillée: Ce test examine la correspondance entre un élément et un ou plusieurs autres éléments, ou la correspondance dans le temps des observations du même élément entre un site et un site voisin. Il repose sur le concept très simple selon lequel deux ou plusieurs éléments étroitement liés sont supposés présenter des régimes de variation similaires. Pour évaluer les anomalies produites par ce test, l'opérateur de contrôle qualité doit prendre en considération les conditions météorologiques dominantes, y compris les effets locaux, et la densité spatiale du réseau.

Les combinaisons d'éléments auxquelles ce test peut être appliqué comprennent:

- Pression moyenne au niveau de la mer et pression au niveau de la station;
- Température minimale terrestre et température minimale de l'air;
- Température du sol à une certaine profondeur et température du sol à une autre profondeur;
- Températures horaires du thermomètre sec, du thermomètre mouillé et du point de rosée (en particulier pour les thermomètres à lecture manuelle);

- Évaporation quotidienne et précipitations quotidiennes;
- Valeurs du parcours du vent à moins de 3 mètres et à 10 mètres.

Suivi des corrélations spatiales entre sites voisins: L'opérateur de contrôle qualité peut évaluer visuellement les corrélations spatiales en comparant le tracé de la série chronologique d'un site avec celui d'un site voisin au moyen d'une interface utilisateur graphique au lieu d'un algorithme informatique. Cette approche peut être appliquée à de nombreux paramètres, notamment dans les réseaux ayant une bonne couverture spatiale. Les éléments suivants sont généralement évalués de cette manière:

- Température horaire du thermomètre sec;
- Température horaire du point de rosée;
- Pression moyenne au niveau de la mer et pression au niveau de la station;
- Vitesse du vent.

Fréquence du test: En fonction de la fréquence d'observation des deux paramètres.

2.9 **Test de cohérence de la température maximale (minimale) de l'air**

Description succincte: Vérifie la cohérence entre la température maximale (minimale) quotidienne de l'air et les valeurs enregistrées lors des observations infraquotidiennes. Il s'agit d'une application spécifique du test 2.1 (test infraquotidien).

Paramètres sur lesquels porte le test: Températures maximales et minimales de l'air.

Description détaillée: Vérifie la cohérence entre la température maximale quotidienne de l'air et la température horaire ou trihoraire de l'air, c'est-à-dire la température qui est enregistrée dans les relevés horaires ou trihoraires. Dans le cas d'observations trihoraires, aucune valeur trihoraire ne doit dépasser la température maximale quotidienne enregistrée. Une erreur de lecture de 0,5 °C est tolérée. Ce test peut être appliqué même si une seule observation quotidienne est disponible; cependant, son efficacité augmente avec la fréquence des observations infraquotidiennes.

Notes:

1. Ce test s'applique également aux températures minimales, comme indiqué dans le descriptif du test 2.1.
2. Il est possible d'utiliser une variante de ce test pour comparer la température maximale avec les températures à heures fixes, afin de s'assurer que la température maximale n'est pas supérieure aux températures à heures fixes de plus d'une quantité déterminée. Cette variante est particulièrement efficace pour les séries de données ayant une résolution temporelle relativement élevée.

Fréquence du test: Quotidienne.

2.10 **Test astronomique – temps d'insolation**

Description succincte: Détermine la différence entre la durée d'insolation et la durée du jour calculée en heures.

Paramètres sur lesquels porte le test: Durée d'insolation.

Description détaillée: Ce test détermine la différence entre la durée d'insolation et la durée du jour calculée. Le test échoue si la différence est supérieure à zéro.

Fréquence du test: Quotidienne.

2.11 **Test de cohérence entre la température de l'air et la température du thermomètre mouillé**

Description succincte: Compare la température de l'air à la température du thermomètre mouillé.

Paramètres sur lesquels porte le test: Températures du thermomètre sec et du thermomètre mouillé.

Description détaillée: Les températures du thermomètre sec et du thermomètre mouillé sont comparées. Si la température du thermomètre mouillé (avec une valeur de tolérance) est supérieure à la température de l'air, le test échoue.

Il est à noter que bien que la tolérance prévue dans ce test s'applique aux thermomètres à liquide sous verre à lecture manuelle, le même algorithme est généralement appliqué aux données issues de capteurs de température électroniques. Même si l'on s'accorde à reconnaître qu'une tolérance beaucoup plus réduite devrait s'appliquer aux capteurs de température électroniques, il est pratique, dans un réseau hybride, d'utiliser le même algorithme en raison de la difficulté d'établir des distinctions entre types d'instrument au stade du contrôle qualité.

Fréquence du test: Horaire, mais le test peut être appliqué à toute observation à heure fixe.

2.12 **Test de cohérence entre la température du thermomètre mouillé et la température du point de rosée**

Description succincte: Détermine la différence entre la température du thermomètre mouillé et la température du point de rosée.

Paramètres sur lesquels porte le test: Température du thermomètre mouillé et température du point de rosée.

Description détaillée: Détermine la différence entre la température du thermomètre mouillé et la température du point de rosée. Le test échoue si la température du thermomètre mouillé enregistrée est inférieure à la température du point de rosée. Ce test s'applique aux lectures de thermomètre manuelles.

Fréquence du test: Selon la fréquence d'observation.

2.13 **Test de revérification température du thermomètre mouillé/température du point de rosée**

Description succincte: La température du thermomètre mouillé est testée par rapport à la température du point de rosée.

Paramètres sur lesquels porte le test: Température du thermomètre mouillé et température du point de rosée.

Description détaillée: On teste la température du thermomètre mouillé par rapport à la température du point de rosée en recalculant la température du point de rosée à partir des températures du thermomètre sec et du thermomètre mouillé. Les valeurs observées et recalculées de la température du point de rosée doivent concorder à un degré près. Ce test est utilisé dans le cas d'observations manuelles, lorsque la température du point de rosée est calculée par l'opérateur. Il est plus sensible que le test 2.12 (test de cohérence entre la température du thermomètre mouillé et la température du point de rosée) et peut être exécuté après celui-ci pour repérer les températures de rosée mal calculées qui n'ont pas été rejetées par le test 2.12.

Lorsque la valeur de la température du point de rosée est rejetée par le test, l'opérateur de contrôle qualité doit recalculer le point de rosée en utilisant la méthode d'observation en vigueur dans le pays.

Fréquence du test: Selon la fréquence d'observation.

2.14 **Test de cohérence de la température du point de rosée et de la température de l'air**

Description succincte: Détermine si la température du point de rosée est inférieure ou égale à la température de l'air.

Paramètres sur lesquels porte le test: Température du point de rosée et température de l'air.

Description détaillée: La température de l'air doit toujours être supérieure ou égale à la température du point de rosée, les deux températures étant égales uniquement lorsque l'air est saturé. Les précipitations et le brouillard sont des situations dans lesquelles l'air peut devenir saturé. Il est important que l'opérateur de contrôle qualité détermine si la température du point de rosée résulte de la saturation de l'air et non d'une erreur de calcul du point de rosée de la part de l'observateur ou du dessèchement de la mèche du thermomètre humide.

Fréquence du test: Selon la fréquence d'observation.

2.15 **Test de cohérence de la visibilité**

Description succincte: Vérifie la cohérence entre la visibilité et le code du temps présent.

Paramètres sur lesquels porte le test: Temps présent et visibilité.

Description détaillée: Vérifie la cohérence de la visibilité en croisant le code du temps présent et les indicateurs de phénomènes météorologiques (brouillard, tempête de sable, brume, tempête de poussière). Ce test concerne la visibilité horizontale et est utilisé pour des observations manuelles.

Voici quelques exemples de combinaisons improbables/impossibles (à adapter aux caractéristiques du climat): visibilité supérieure à 1 000 m et brouillard, visibilité supérieure à 1 000 m ou inférieure à 200 m et tempête de sable ou de poussière légère ou modérée, visibilité supérieure à 10 km et brume, etc.

Fréquence du test: Selon la fréquence d'observation.

2.16 **Test de cohérence de la nébulosité totale**

Description succincte: Vérifie la cohérence entre la nébulosité totale et divers éléments. Utilisé avec des observations manuelles.

Paramètres sur lesquels porte le test: Nébulosité totale, temps présent.

Description détaillée: Vérification de la cohérence de la nébulosité totale par rapport à la hauteur des nuages et au temps présent.

Exemples de combinaisons improbables/impossibles: nébulosité totale inférieure à la nébulosité à l'étage inférieur, à l'étage moyen ou à l'étage élevé; la nébulosité totale est invisible, mais on observe des traînées de précipitations (virga); il n'y a aucun nuage, mais il y a des précipitations, etc.

Fréquence du test: Selon la fréquence d'observation.

2.17 **Test des indicateurs de phénomènes météorologiques quotidiens**

Description succincte: Vérifie la cohérence entre les indicateurs de phénomènes météorologiques quotidiens et les observations infraquotidiennes.

Paramètres sur lesquels porte le test: Phénomènes tels que la grêle, la neige, les orages, les tempêtes de poussière, les vents forts, les coups de vent, la gelée, la brume humide ou sèche, le brouillard, etc.

Description détaillée: Le test vérifie la cohérence entre les indicateurs de phénomènes météorologiques quotidiens (jours avec grêle, tonnerre, brouillard, etc.) et les codes figurant dans les tableaux infraquotidiens et divers autres éléments quotidiens. Il détecte les cas où un phénomène a été signalé dans un code synoptique de temps présent, mais pas dans les catégories de phénomènes de minuit à minuit¹, et vice versa. Il peut également détecter des incohérences entre les phénomènes météorologiques et d'autres variables, telles que la température, le point de rosée, etc.

Voici quelques exemples de combinaisons suspectes (qui pourront nécessiter une adaptation le cas échéant, en fonction des conditions climatiques): indicateur de neige et température minimale de l'air supérieure à 9 °C; pas d'indicateur d'orage alors qu'un orage a été observé dans la journée; pas d'indicateur de poussière/brume/brouillard, mais visibilité inférieure aux seuils correspondants pendant la journée; indicateur de gelée, mais température minimale de l'air supérieure à 4 °C, etc.

Fréquence du test: Quotidienne.

2.18 **Test de cohérence du temps présent**

Description succincte: Vérifie la cohérence entre les codes du temps présent figurant dans les tableaux infraquotidiens et divers autres éléments quotidiens.

Paramètres sur lesquels porte le test: Temps présent et température, etc.

Description détaillée: Le test signale les combinaisons de temps présent et de température improbables, telles que des précipitations verglaçantes ou du brouillard déposant du givre et des températures de l'air élevées. D'autres combinaisons peuvent être incluses à la discrétion du SMHN.

L'une des applications de ce test est la comparaison de la température de l'air et des observations de dépôts de givre et de précipitations verglaçantes.

Fréquence du test: Quotidienne et infraquotidienne.

2.19 **Test de la température du sol**

Description succincte: Vérifie la cohérence des températures du sol à différentes profondeurs.

Paramètres sur lesquels porte le test: Lectures de la température du sol à des profondeurs de 5 cm, 10 cm, 20 cm, 50 cm et 1 m.

¹ Ou une autre période de 24 heures, selon le SMHN.

Description détaillée: Le test du sol examine diverses combinaisons de températures du sol à différentes profondeurs (il peut s'avérer nécessaire d'adapter ces températures en fonction des caractéristiques du climat):

<i>Élément cible</i>	<i>Expression du test (suspect si vrai)</i>
Température du sol à une profondeur de 5 cm (temp SOL 5)	temp SOL 5 = temp SOL 10 ET temp SOL 5 = temp SOL 20 ET temp SOL 5 = temp SOL 50 ET temp SOL 5 = temp SOL 100 temp SOL 5 – temp SOL 10 > 16 OU temp SOL 10 – temp SOL 5 > 6
Température du sol à une profondeur de 10 cm (temp SOL 10)	temp SOL 10 – temp SOL 20 > 10 OU temp SOL 20 – temp SOL 10 > 6
Température du sol à une profondeur de 20 cm (temp SOL 20)	temp SOL 50 – temp SOL 20 > 6 OU temp SOL 20 – temp SOL 50 > 7
Température du sol à une profondeur de 50 cm (temp SOL 50)	temp SOL 50 – temp SOL 100 > 7 OU temp SOL 100 – temp SOL 50 > 4
Note: Les températures du sol sont exprimées en degrés Celsius.	

Fréquence du test: Selon la fréquence d'observation.

2.20 **Test de cohérence des messages**

Description succincte: Dans les cas où plusieurs messages sont reçus pour une même observation d'un paramètre, le test vérifie si les messages contiennent tous la même valeur.

Paramètres sur lesquels porte le test: Tous les paramètres stockés dans la base de données.

Description détaillée: Le test détecte les cas où plusieurs messages sont reçus pour une même observation. Il vérifie que les messages contiennent tous la même valeur pour un paramètre déterminé. Le test échoue si tous les messages n'affichent pas la même valeur.

Fréquence du test: En fonction des messages générés.

2.21 **Test des données maximales/minimales sur une minute**

Description succincte: Compare la température maximale ou minimale indiquée avec la valeur correspondante sur une minute au moment du maximum ou du minimum de température, afin de s'assurer que les valeurs correspondent.

Paramètres sur lesquels porte le test: Températures maximales et minimales de l'air.

Description détaillée: Le test compare la température maximale ou minimale indiquée à la valeur correspondante sur une minute au moment du maximum ou du minimum de température pour s'assurer que les valeurs correspondent. Ce test est conçu pour déterminer si l'horloge de la SMA connaît une dérive et si des données sur une minute sont manquantes. Si la température de l'air sur une minute ne correspond pas à la température maximale ou minimale quotidienne de l'air, ou si les données sur une minute sont manquantes au moment du maximum ou du minimum de température, le test échoue. Ce test comprend également un test de pic grossier qui compare la valeur sur une minute pour la minute qui précède et pour la minute qui suit la valeur quotidienne extrême, pour vérifier que la différence n'est pas supérieure à 0,5 °C pour la température minimale et à 1 °C pour la température maximale.

Fréquence du test: Quotidienne.

3. TESTS HEURISTIQUES

Tests (généralement multivariés) qui s'appuient sur l'expérience et la connaissance des processus, techniques et instruments d'observation pour détecter les valeurs incohérentes, improbables ou impossibles et les signaler comme suspects. Une enquête manuelle peut ensuite être menée pour déterminer la validité des valeurs suspects.

3.1 Test d'humidité relative (dessèchement de la mèche du thermomètre mouillé)

Description succincte: Test visant à déterminer si la mèche du thermomètre mouillé a séché ou a été souillée.

Paramètres sur lesquels porte le test: Températures du thermomètre sec et du thermomètre mouillé.

Description détaillée: Pour déterminer si la mèche d'un thermomètre mouillé a séché, le test examine les observations prises à 15 heures, heure à laquelle l'humidité relative est supérieure à 90 %. Pour réduire le taux de fausses erreurs, on élimine les cas où, en plus d'une humidité relative supérieure à 90 %, il y a eu des précipitations et/ou du brouillard, en se référant aux codes du temps passé et du temps présent.

Le test ne sera effectué que si:

- L'heure d'observation est 15 heures (ou l'heure équivalente à laquelle le thermomètre à maximum est lu dans l'après-midi);
- L'altitude est inférieure à 1 100 m (il n'est pas possible de distinguer les effets des nuages orographiques des effets du dessèchement de la mèche du thermomètre mouillé dans ce test).

Fréquence du test: Chaque jour à 15 heures (ou l'heure équivalente à laquelle le thermomètre à maximum est lu dans l'après-midi).

4. TESTS DE TRANSMISSION DES DONNÉES

Tests permettant de s'assurer que les observations qui ne correspondent pas au programme d'observations prévu sont soit rejetées, soit signalées comme suspects.

4.1 Test des observations reçues du futur

Description succincte: Généralement présenté sous la forme d'un rapport recensant les cas où l'observateur a envoyé l'observation en avance.

Paramètres sur lesquels porte le test: Compare l'heure locale de l'observation et l'heure de réception de l'observation.

Description détaillée: Le SMHN ou une entité régionale locale doit fixer une limite déterminant la «précocité» admissible de l'observation. Par exemple, une limite de 30 minutes signifie que les données d'observation transmises jusqu'à 30 minutes avant l'heure officielle de l'observation seront acceptées. Il est recommandé de tenir un journal identifiant les observations reçues d'une station systématiquement en avance, afin de mettre en évidence d'éventuels manquements dans les pratiques d'observation; ces journaux pourront être utilisés pour définir les priorités en matière de formation.

Fréquence du test: En fonction des observations générées.

4.2 **Test de différences de périodes**

Description succincte: Détermine si la période (cumul de jours) correspond aux relevés existants.

Paramètres sur lesquels porte le test: Paramètres quotidiens.

Description détaillée: Ce test s'applique aux observations quotidiennes et consiste simplement à vérifier si les périodes (cumuls de jours) correspondent aux relevés existants. Par exemple, un message synoptique est reçu pour le lundi matin pour une station qui n'a pas enregistré de données pendant le week-end. Il conviendra de vérifier les valeurs des températures maximales et minimales figurant dans la base de données pour ce lundi, car il peut s'agir de valeurs sur plusieurs jours même si elles sont archivées comme une seule valeur journalière. Voir également plus haut l'explication concernant le test 2.5 (test spatial de précipitations nulles).

Fréquence du test: Quotidienne.

4.3 **Test de longues périodes**

Description succincte: Détermine si les périodes (cumuls de jours) sont excessivement longues (plusieurs jours) pour des paramètres concernant lesquels un grand écart entre les lectures quotidiennes est susceptible d'affecter l'intégrité des données. Par exemple, si les précipitations ne sont pas relevées pendant cinq jours dans un climat chaud, le total des précipitations sera probablement réduit sous l'effet de l'évaporation.

Paramètres sur lesquels porte le test: Relevés quotidiens des précipitations et de l'évaporation.

Description détaillée: Le test détecte les périodes déraisonnablement longues qui sont susceptibles de compromettre la validité des lectures. Pour les précipitations et l'évaporation, un écart de période plus restreint peut être acceptable en fonction du climat et des exigences des utilisateurs du SMHN. Le SMHN peut s'appuyer sur ce test pour l'aider à fixer le nombre de jours au-delà duquel les totaux de précipitations et d'évaporation peuvent être considérés comme suspects.

Fréquence du test: Quotidienne.

4.4 **Test de chevauchement avec les périodes futures**

Description succincte: Test effectué pour identifier les cas où une observation est envoyée en retard, et après la réception d'une observation d'un jour ultérieur, afin de garantir la cohérence de la période d'observation déclarée.

Paramètres sur lesquels porte le test: Paramètres quotidiens.

Description détaillée: Le test est effectué pour identifier les cas où une observation est envoyée en retard, et après la réception d'une observation d'un jour ultérieur, afin de garantir la cohérence de la période d'observation déclarée. Si la période indiquée dans le message du jour suivant ne correspond pas à la période attendue, le test est un échec. Voici un exemple d'une situation qui entraînerait un échec au test:

- L'observation quotidienne du 8 n'est pas envoyée à temps.
- L'observation quotidienne du 9 correspond à une période de deux jours.
- Après la réception de l'observation quotidienne du 9, l'observation quotidienne du 8 est envoyée, ce qui entraîne une incohérence dans la période indiquée le 9.

Fréquence du test: Quotidienne.

4.5 **Test de réitération**

Description succincte: Vérifie si la valeur, la période ou l'indicateur de qualité d'une observation ont changé depuis la dernière mise à jour. S'il y a eu des changements depuis la dernière mise à jour, le test se soldera par un échec.

Paramètres sur lesquels porte le test: Tous les paramètres stockés dans la base de données.

Description détaillée: Le test vérifie si la valeur, la période ou l'indicateur de qualité d'une observation ont changé depuis la dernière mise à jour. S'il y a eu des changements depuis la dernière mise à jour, le test se soldera par un échec. Cependant, il n'échouera pas si la valeur, la période ou l'indicateur de qualité a été modifié par un opérateur de contrôle qualité qui figure sur la liste des utilisateurs autorisés. Par conséquent, les actions de contrôle qualité n'entraîneront pas un échec au test.

Fréquence du test: En fonction des observations générées.

5. **TESTS STATISTIQUES**

Tests qui effectuent des analyses statistiques des données historiques pour détecter les valeurs incohérentes, improbables ou impossibles et les signaler comme suspectes. Une enquête manuelle peut ensuite être menée pour déterminer la validité des valeurs suspectes.

5.1 **Test de portée fondé sur le climat**

Description succincte: Compare la valeur météorologique avec les valeurs climatologiques supérieure et inférieure. Différents seuils peuvent être appliqués pour tenir compte des variations saisonnières du climat local. Les valeurs qui se situent en dehors de ces limites seront signalées comme suspectes ou, à tout le moins, soumises à une enquête plus approfondie.

Paramètres sur lesquels porte le test: Généralement les précipitations quotidiennes. Cependant, ce test peut aussi être utilisé pour des paramètres infraquotidiens si un paramètre de surveillance de la qualité (PSQ) approprié a été déterminé.

Description détaillée: Compare la valeur météorologique avec les valeurs climatologiques supérieure et inférieure. Pour que ce test soit efficace, il est nécessaire de disposer d'une série chronologique de relevés suffisamment longue (un minimum de 30 ans est recommandé), afin de s'assurer que le PSQ représente toute la variabilité du climat. Par conséquent, ce test ne sera pas fiable pour les nouveaux sites et les sites dont les relevés chronologiques sont incomplets; pour ceux-ci, on peut utiliser le PSQ d'une station voisine ayant un climat similaire, mais une série de relevés plus longue.

Au lieu d'un PSQ fixe, on peut utiliser un PSQ fondé sur un test statistique (par exemple, plus de trois ou quatre écarts types par rapport à la moyenne). Cependant, la prudence doit être de rigueur lorsque l'on définit un PSQ de la sorte, car certaines variables météorologiques présentent des distributions de fréquence qui diffèrent sensiblement de la distribution normale (de Gauss).

Fréquence du test: Quotidienne et infraquotidienne.

5.2 **Test de courbe plate**

Description succincte: Recherche les cas où une série de valeurs météorologiques reste inchangée sur une période étonnamment longue, pour vérifier si le paramètre est effectivement resté inchangé dans la durée.

Paramètres sur lesquels porte le test: Paramètres quotidiens et infraquotidiens.

Description détaillée: Des courbes plates peuvent résulter d'un équipement défectueux ou d'une saisie manuelle incorrecte des données. Il est important de noter que ce test est d'une utilité limitée pour certains paramètres qui ont une faible variabilité, comme les températures du sol mesurées très en dessous de la surface.

Fréquence du test: Selon la fréquence d'observation.

5.3 **Test de changement rapide**

Description succincte: Examine l'importance de la différence entre la valeur actuellement observée et la valeur précédente.

Paramètres sur lesquels porte le test: Paramètres infraquotidiens dont la fréquence d'observation est de trois heures ou moins.

Description détaillée: Ce test détecte les problèmes de données dus au fait qu'une valeur a enregistré une hausse ou une baisse rapide en raison de facteurs non météorologiques. Comme ce test postule que le changement rapide n'est pas dû à la variabilité météorologique, il perd en sensibilité lorsque les intervalles d'observation sont plus longs.

La variabilité du paramètre peut ne pas être constante pour chaque station ou mois, de même qu'elle peut ne pas être constante pour chaque heure. Cette variation temporelle et spatiale peut justifier le calcul de PSQ discrets au niveau de la station, du mois et de l'heure.

Lorsque le délai entre la valeur observée et la valeur suivante ou précédente est supérieur à trois heures, ce test n'est généralement pas effectué.

Fréquence du test: Selon la fréquence d'observation.

5.4 **Test de pic**

Description succincte: Compare une observation météorologique déterminée avec les valeurs précédentes et suivantes.

Paramètres sur lesquels porte le test: Données infraquotidiennes. Ce test est particulièrement adapté aux données à fréquence élevée des SMA.

Description détaillée: Le test de pic est assez semblable au test de changement rapide; cependant, le test de pic recherche les sauts soudains suivis de chutes soudaines de la valeur d'une variable, et vice versa. Ces variations peuvent être dues, par exemple, à un pic de tension dans une SMA. Les limites PSQ appliquées au test de pic sont inférieures à celles appliquées au test de changement rapide. Lorsque le délai entre la valeur observée et la valeur suivante ou précédente est supérieur à trois heures, ce test n'est pas effectué. C'est sur les données à fréquence élevée et pour les périodes d'observation inférieures à 20 minutes que le test de pic donne les meilleurs résultats.

Fréquence du test: Selon la fréquence d'observation.

5.5 **Test de fréquence d'arrondis**

Description succincte: Recherche les cas d'arrondis excessifs des valeurs. Ce test s'applique aux observations manuelles, et plus précisément aux cas où un opérateur arrondit le relevé aux marques de graduation de l'instrument au lieu d'estimer la valeur entre les graduations de l'instrument.

Paramètres sur lesquels porte le test: Ce test est généralement utilisé dans les cas où les thermomètres et les pluviomètres sont lus manuellement.

Description détaillée: La fréquence d'apparition du dernier chiffre (par exemple, la première décimale pour des paramètres tels que la température) est examinée sur un certain nombre de jours et stockée en tant que PSQ. Si la fréquence d'apparition d'une des valeurs est «élevée», par exemple si le nombre de valeurs se terminant par un zéro est anormalement élevé, il y a échec au test. La fréquence est exprimée en termes de pourcentage des valeurs analysées. Il est à noter que les valeurs arrondies ne sont pas considérées comme suspectes; l'indicateur de qualité qui leur est attribué signale un niveau de confiance inférieur. Ce test est important pour identifier les questions à aborder en formation.

Le test de fréquence d'arrondis est effectué pour un mois entier et attribué à un élément et une date désignés.

Ce test n'est effectué que si au moins 75 % des valeurs existent (et sont non nulles) pour la période en question.

Fréquence du test: Mensuelle.

5.6 **Test de variabilité spatiale**

Description succincte: Les différences entre la station considérée et les stations voisines sont comparées.

Paramètres sur lesquels porte le test: Paramètres quotidiens et infraquotidiens.

Description détaillée: La différence entre une station et ses voisines est calculée et comparée à un seuil raisonnable, par exemple l'écart type de ces différences, déduit du relevé historique.

Fréquence du test: Selon la fréquence d'observation.

5.7 **Test spatial fondé sur la cartographie**

Description succincte: Compare la valeur de l'élément météorologique avec les valeurs de donnée environnantes. Une technique fondée sur la cartographie, du type de l'analyse de Barnes, est utilisée pour pondérer les stations proches².

Paramètres sur lesquels porte le test: Paramètres quotidiens et infraquotidiens.

Description détaillée: L'objectif de l'analyse cartographique est d'estimer une valeur pour une station cible à partir de valeurs réelles à des points situés dans la région générale de la station, en utilisant une fonction de pondération objective (les points proches devraient avoir davantage d'influence et devraient donc se voir attribuer un coefficient de pondération plus élevé).

² Il existe encore d'autres types d'analyse cartographique spatiale. Par exemple, certaines analyses incorporent l'altitude en tant que troisième dimension pour des variables sensibles à la topographie comme la température et les précipitations.

Le principe de base de ce dispositif est le suivant: il s'agit d'estimer une valeur de la variable considérée par interpolation cartographique à partir d'un certain nombre de points du réseau³ – les points qui correspondent aux emplacements des stations. À chaque emplacement de station correspondent une valeur réellement observée et une valeur interpolée fondée sur les données de toutes les autres observations proches (l'influence d'une observation sur un point étant déterminée à l'aide d'une fonction de pondération). Au point P analysé, la valeur observée est comparée à la valeur interpolée estimée par la cartographie.

Contraintes:

1. Ce test perd en sensibilité pour les stations de comparaison situées à plus d'une certaine distance de la station cible. Les SMHN doivent déterminer la portée maximale du test. En Australie, les zones où les données sont rares sont relativement plates et présentent un climat homogène; par conséquent, la portée (R) est fixée à 200 km, et les stations situées au-delà de cette distance sont exclues de l'analyse.
2. Les résultats de ce test doivent être interprétés à la lumière des caractéristiques locales du site. En particulier, certaines variables météorologiques peuvent présenter de grands gradients dans les zones à topographie complexe ou près des côtes. L'interprétation détaillée des résultats du test peut nécessiter d'avoir certaines connaissances sur les sites concernés. Par exemple, un site situé au sommet d'une colline ou sur une pente peut enregistrer des températures minimales systématiquement plus élevées que d'autres sites de la région situés en vallée ou en plaine.

Note: Pour déterminer si des valeurs sont raisonnables ou non, outre l'interpolation statistique, on peut comparer les valeurs de la station suspecte avec les valeurs des stations environnantes en procédant à une inspection visuelle. En cas d'inspection visuelle, les mêmes contraintes que celles énoncées ci-dessus s'appliquent.

Fréquence du test: Quotidienne et infraquotidienne, mais d'autres périodicités sont possibles, par exemple tous les mois.

5.8 Test spatial de régression linéaire

Description succincte: Compare la valeur météorologique avec les valeurs des stations environnantes en effectuant une régression linéaire.

Paramètres sur lesquels porte le test: Précipitations quotidiennes, mais le test peut être adapté à d'autres paramètres.

Description détaillée: L'analyse spatiale par régression linéaire⁴ produit une estimation d'un élément à une station et à un moment cibles déterminés sur la base du degré de corrélation observé au fil du temps entre cette station cible et les stations voisines. La corrélation est calculée par régression linéaire, et les stations qui présentent une meilleure corrélation avec la station cible se voient affecter une pondération statistique plus élevée que les stations qui présentent une corrélation moins marquée.

Fréquence du test: Quotidienne mais peut être adaptée à d'autres fréquences d'observation.

³ Dans cette description de l'algorithme, le réseau est simplement l'emplacement des stations, de sorte que l'interpolation est réalisée directement sur les emplacements des stations. Autre possibilité, si une analyse en points de grille du champ est disponible, l'interpolation sur les emplacements des stations peut être effectuée à partir de la grille. Cette solution pourrait être adaptée, par exemple, aux tests liés à la PNT.

⁴ Adapté d'une méthode décrite dans Hubbard, K. G., Goddard, S., Sorensen, W. D., Wells, N. et Osugi, T. T., Performance of Quality Assurance Procedures for an Applied Climate Information System. *Journal of Atmospheric and Oceanic Technology* **2005**, 22 (1), 105–112. <https://doi.org/10.1175/JTECH-1657.1>

5.9 **Test de variabilité spatiale par régression linéaire**

Description succincte: Utilise la variabilité des stations voisines plutôt que les estimations des erreurs types pour calculer les limites de test. Ce test est une variante du test spatial de régression linéaire.

Paramètres sur lesquels porte le test: Précipitations quotidiennes, mais le test peut être adapté à d'autres paramètres.

Description détaillée: Ce test utilise la variabilité des observations dans des stations voisines (au même moment d'observation), plutôt que les estimations de l'erreur type, pour calculer les limites de test. Il est conçu pour évaluer la fiabilité des stations voisines dans l'analyse des variations. Ce test est bien adapté aux précipitations lorsque des changements sont observés dans la variabilité d'un ensemble d'observations pour une région donnée.

Contraintes:

1. Ce test perd en sensibilité pour les stations de comparaison qui sont trop éloignées de la station cible ou dont l'altitude est trop différente de celle de la station cible. Les SMHN doivent déterminer la portée maximale du test. En Australie, la plage d'altitude à l'intérieur de laquelle un climat peut être considéré comme «homogène» est fixée à 300 m maximum.
2. Le type de précipitations doit également être pris en compte, comme les orages isolés; même les stations considérées comme proches les unes des autres peuvent alors présenter des différences de totaux quotidiens significatives.

Fréquence du test: Quotidienne mais peut être adaptée à d'autres fréquences d'observation.

5.10 **Test de régression linéaire sur plusieurs jours**

Description succincte: Compare la climatologie des différences entre la station considérée et les stations voisines sur une période de plusieurs jours. Ce test est semblable au test spatial de régression linéaire, mais s'applique à plusieurs jours. Il est indiqué pour détecter des stations pluviométriques potentiellement suspectes, dans lesquelles l'observateur omet régulièrement de lire le pluviomètre.

Paramètres sur lesquels porte le test: Précipitations quotidiennes, mais le test peut être adapté à d'autres paramètres.

Description détaillée: Bien que ce test soit semblable au test spatial de régression linéaire, les valeurs qu'il utilise doivent être définies différemment. Pour les cumuls, les précipitations sont la somme des relevés quotidiens des stations voisines, la température maximale est la plus élevée des relevés quotidiens des stations voisines et la température minimale est la plus basse des relevés quotidiens des stations voisines.

Fréquence du test: Quotidienne, mais peut être adaptée à d'autres fréquences d'observation.

5.11 **Test maximum/minimum**

Description succincte: Vérifie que la différence entre les températures maximales et minimales est réaliste (c'est-à-dire supérieure/égale à zéro et inférieure à une limite supérieure). Les limites sont calculées à partir d'une climatologie utilisant au moins cinq ans de données (de préférence 30 ans). Ce test est indiqué pour détecter les erreurs de lecture manuelle, qui ont tendance à être erronées de 5 ou 10 °C.

Paramètres sur lesquels porte le test: Températures maximales et minimales quotidiennes de l'air.

Description détaillée: La différence entre les températures maximales et minimales est calculée. Le test échoue si la différence est inférieure à zéro (le maximum étant dans ce cas inférieur au minimum) ou dépasse une limite supérieure.

Si la différence est inférieure à zéro, le test génère une erreur, mais si elle dépasse la limite supérieure, il génère une valeur suspecte.

Fréquence du test: Quotidienne.

Pour de plus amples informations, veuillez vous adresser à:

Organisation météorologique mondiale

7 bis, avenue de la Paix – Case postale 2300 – CH 1211 Genève 2 – Suisse

Bureau de la communication stratégique

Tél.: +41 (0) 22 730 87 40/83 14 – Fax: +41 (0) 22 730 80 27

Courriel: communications@wmo.int

public.wmo.int