



**HAL**  
open science

# Modélisation du champ MAGnétique principal par Inversion de données d'OBServatoires (MAGIOBS): étude de résolution spatiale

Erwan Thébault, Lydie-Sarah Gailler, Jean-François Oehler, Didier Rouxel

► **To cite this version:**

Erwan Thébault, Lydie-Sarah Gailler, Jean-François Oehler, Didier Rouxel. Modélisation du champ MAGnétique principal par Inversion de données d'OBServatoires (MAGIOBS): étude de résolution spatiale. Laboratoire Magmas et Volcans, Campus universitaire des Cézeaux, 6 avenue Blaise Pascal, 63170 Aubière. 2024. hal-04505145

**HAL Id: hal-04505145**

**<https://hal.science/hal-04505145>**

Submitted on 14 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

# Modélisation du champ MAGnétique principal par Inversion de données d'OBServatoires (MAGIOBS): étude de résolution spatiale

THEBAULT Erwan<sup>1\*</sup> | GAILLER Lydie<sup>1\*</sup> | OEHLER  
Jean-François<sup>2†</sup> | ROUXEL Didier<sup>2†</sup>

## REVISION

Rev	Status du document	Date	Doc propriété	Modifications
00	Version 1	12/01/2024	LMV	
01	Revision	04/02/2024	LMV	Contenu
02	Version 2	14/03/2024	LMV	Recommandations SHOM

Ce document est rédigé dans le cadre du projet MAGIOBS. Il ne relève d'aucune clause de confidentialité.

## Résumé

Les composantes vectorielles du champ magnétique terrestre sont mesurées en continu aux observatoires magnétiques terrestres. Certains de ces observatoires délivrent des mesures qui obéissent à des formats et des standards internationaux tels que définis par l'Association Internationale de Géomagnétisme et d'Aéronomie (IAGA/AIGA). La continuité de service et la stabilité de la ligne de base des séries temporelles sont deux critères majeurs de la labellisation INTERMAGNET. Cependant, ces standards se révèlent en pratique très contraignants, à la fois pour leurs raisons techniques (maintenance et opération des observatoires en territoires extrêmes et/ou isolés) que pour des raisons sociétales (expansion des zones urbaines et contaminations anthropiques). Les ouvertures/fermetures, déplacements, et changements de la ligne de base suite à des migrations techniques entraînent des discontinuités de mesures et des distributions spatiales et temporelles inhomogènes. Ce rapport présente des résultats de simulation synthétiques à partir de la distribution réelle des observatoires magnétiques terrestres enregistrés dans le réseau INTERMAGNET. Pour ce faire, nous avons pris les positions des observatoires compilés par le British Geological Survey (BGS, UK). Nous avons calculé un champ magnétique connu à ces positions et récupéré par des méthodes numériques le modèle correspondant que nous comparons avec le modèle initial. Nous estimons ainsi la résolution spatiale qui peut être atteinte dans la description du champ magnétique interne en harmoniques sphériques à partir des données d'observatoires seules et identifions des sources d'instabilités. En l'absence d'information a priori sur le comportement du champ global, la résolution spatiale maximum de la partie statique du champ global est estimée être entre  $n = 6 - 8$  et  $n = 8$  des harmoniques sphériques (pour une cible à  $n = 13$ ). Cela correspond à une longueur d'onde de résolution entre 5000 km et 6700km (contre environ 3200 km pour les modèles comme l'IGRF

<sup>1</sup>Laboratoire Magmas et Volcans, Université Clermont Auvergne, Blaise Pascal, 63170 Aubière, France

<sup>2</sup>SHOM (French Hydrographic and Oceanographic Service), 13, rue de Chatellier, CS 92803, 29228 Brest cedex 2, France

### Correspondence

Thebault E., Laboratoire Magmas et Volcans, France  
Email: erwan.thebault@uca.fr

### Funding information

Cette étude est partiellement financée par le CNES dans le cadre du projet "Exploitation scientifique de la mission Swarm" et par l'Agence de l'Innovation de Défense (AID) du Ministère des Armées dans le cadre du Projet de Technologies de Défense APOGé.

ou le WMM obtenus à partir des mesures spatiales). Ce travail préliminaire nous permet de poser des pistes pour améliorer la représentation spatiale du champ magnétique terrestre. Cette notice synthétique accompagne la restitution orale et illustrée qui a eu lieu le 20 novembre 2023 dans les locaux du SHOM et complétée du compte rendu de réunion rédigé par Lydie-Sarah GAILLER.

## RESUME

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Distribution géographique des observatoires et calcul des données vectorielles synthétiques</b>	<b>5</b>
2.1	Rappels sur la distribution des observatoires magnétiques terrestres . . . . .	5
2.2	Calcul des modèles synthétiques . . . . .	6
2.3	Résolution du problème inverse . . . . .	8
<b>3</b>	<b>Résultats</b>	<b>10</b>
3.1	Analyse préliminaire de la matrice d'inversion . . . . .	10
3.2	Effets de la variation temporelle et spatiale de la distribution . . . . .	10
3.3	Exploration statistique des sous-ensembles . . . . .	14
<b>4</b>	<b>Conclusions</b>	<b>16</b>

## 1 | INTRODUCTION

Le champ magnétique terrestre est une quantité vectorielle constituée de nombreuses sources d'origine interne (noyau, croûte, manteau) et externes (ionosphère, magnétosphère). La mesure en continu est réalisée soit depuis l'espace depuis environ deux décennies (Missions satellitaires OERSTED, CHAMP et Swarm) soit aux sols aux observatoires magnétiques. L'avantage de la technologie satellitaire est d'offrir une couverture globale avec une instrumentation homogène. Ses limites, néanmoins, sont liées au déplacement de la plateforme de mesure dans l'espace (une orbite toute les 90 min environ à 500 km d'altitude), à la durée de vie de la mission, et à sa couverture spatiale et temporelle: un satellite en orbite quasi-polaire revisite une même heure local et longitude tous les 3-4 mois. Ces limites entraînent une difficulté importante pour séparer les effets internes des effets externes du champ magnétique et un tel modèle est intrinsèquement obtenu à partir de données intégrées en temps sur plusieurs mois pour obtenir une description spatiale globale. Les observatoires magnétiques terrestres sont quant à eux fixes en espace (en principe) et délivrent de la mesure continue. Ils ont donc la capacité de détecter en temps réel des perturbations rapides dont l'origine est externe et de permettre en retour une meilleure identification des évolutions du champ interne. En revanche, situés sur des Terres émergées, la résolution spatiale des descriptions globales du champ que l'on peut obtenir à partir des observatoires magnétique souffre de la dichotomie océan/continent de la surface terrestre. Ils sont d'autre part posés sur la croûte terrestre qui contient une aimantation que nous pouvons supposer statique mais qui varie en espace. Ces effets de site ou « biais crustaux » sont parfois importants sur des structures très aimantées et créent des déformations artificielles dans les modèles qui tentent de décrire le champ global, typiquement pour les harmoniques sphériques  $n < 15$  (Thébault and Gailler, 2023). Dans un premier temps nous présentons la base de données des observatoires magnétiques et décrivons la construction du jeu de données synthétiques que nous avons calculé. Dans une seconde partie, nous écrivons le problème inverse afin de récupérer un modèle global à partir de ces mesures distribuées aux positions des observatoires. A partir de critères définis dans Thébault and Gailler (2023) dans la cadre du projet MAGIOBS nous comparons ensuite le modèle obtenu avec le modèle initial utilisé pour générer les données synthétiques. Cette simulation « end-to-end » nous permet de poser un intervalle de confiance pour la limite de résolution du champ global (interne) à partir des données d'observatoires uniquement. Enfin, nous explorons les différentes sources possibles de désaccords entre le modèle d'entrée et le modèle estimé afin de poser les éléments d'une amélioration par une exploration aléatoire de la base de données disponible.

## 2 | DISTRIBUTION GÉOGRAPHIQUE DES OBSERVATOIRES ET CALCUL DES DONNÉES VECTORIELLES SYNTHÉTIQUES

### 2.1 | Rappels sur la distribution des observatoires magnétiques terrestres

Nous reprenons les positions géographiques des mesures d'observatoires définies dans le standard INTERMAGNET pour l'année 2020 (Figure 1). Les positions géographiques sont données dans le système WGS84 que nous convertissons en coordonnées géocentriques. Cette distribution spatiale n'est pas homogène avec une densité plus importante sur les continents, notamment en Europe, et avec une disparité importante entre les hémisphères Nord et Sud. De large zones comme le long de l'équateur (Afrique centrale) et en Sibérie ne sont pas couvertes par la mesure continue dans la base INTERMAGNET. La zone Antarctique quant à elle est couverte de manière quasi-circulaire en bordure du continent ce qui aura un effet bénéfique pour la caractérisation de certaines composantes du champ global sur tout le continent Antarctique. Notons à ce propos que le champ magnétique étant global et généré principalement à la frontière noyau-manteau, les mesures magnétiques sont susceptibles de contraindre une large zone. En d'autres

termes, une mesure magnétique ne contraint pas uniquement le champ magnétique à sa verticale mais est sensible à un rayon proportionnel à la profondeur de la source principal (soit un rayon d'action d'environ 3000 km). A partir de cet argument qui découle de la théorie du potentiel, nous entrevoyons déjà qu'une forte densité régionale de mesures n'est pas nécessairement plus avantageuse qu'une densité moindre mais mieux distribuée géographiquement. Pour les distances géographiques importantes supérieures à 3000 km (qui correspondent au degré 13 des harmoniques sphériques) entre deux séries de mesures aux observatoires nous pressentons déjà grâce à la Figure 1 la difficulté de pouvoir résoudre ces échelles à partir des seules mesures d'observatoires. A titre d'exemple, la distance entre les côtes américaines du Nord et de l'Europe étant d'environ 5 000 km, nous anticipons un déficit de résolution dans l'Atlantique Nord aussi bien qu'au dessus de l'Océan Pacifique.

La figure 2 montre le nombre d'observatoires disponibles chaque année entre 2003 et 2022. L'évolution du nombre d'observatoires sur cette période montrent plusieurs tendances. La première est qu'un effort international a été opéré jusqu'en 2015 pour une meilleure couverture, jusqu'à 152 observatoires disponibles en 2015, et que la maintenance n'a pas été assurée avec une chute rapide pour atteindre 93 observatoires en 2021 ; l'année 2022 n'étant pas représentative car les données définitives ne sont pas encore toutes disponibles. Cette évolution illustre la difficulté de maintenir des observatoires dans des contextes géopolitiques ou environnementaux instables. Elle illustre également une tendance actuelle pour trouver des solutions intermédiaires (type variomètre, observatoires « mobiles ») dont les standards ne respectent pas nécessairement les prescriptions INTERMAGNET. Si cette décroissance ne révèle pas nécessairement un abandon de la mesure terrestre, elle révèle néanmoins l'absence de cadre et de normes internationale pour accéder facilement à l'intégralité des mesures, quel que soit leur standard, dans une base de données collective et ouverte. En résumé, les études menées dans la suite de cette synthèse sont limitées par notre capacité à accéder à la diversité des données réellement disponibles ; diversité dont nous n'avons pas connaissance.

## 2.2 | Calcul des modèles synthétiques

Le champ magnétique terrestre est représenté en harmoniques sphériques (Thébault and Gailler, 2023). La méthode des harmoniques sphériques est la plus compacte possible pour la représentation globale, c'est-à-dire que le nombre de paramètres à évaluer afin de décrire globalement le champ est, parmi toutes les techniques existantes, le plus faible. Etant donné le nombre restreint d'observatoires cette solution est la seule voie possible.

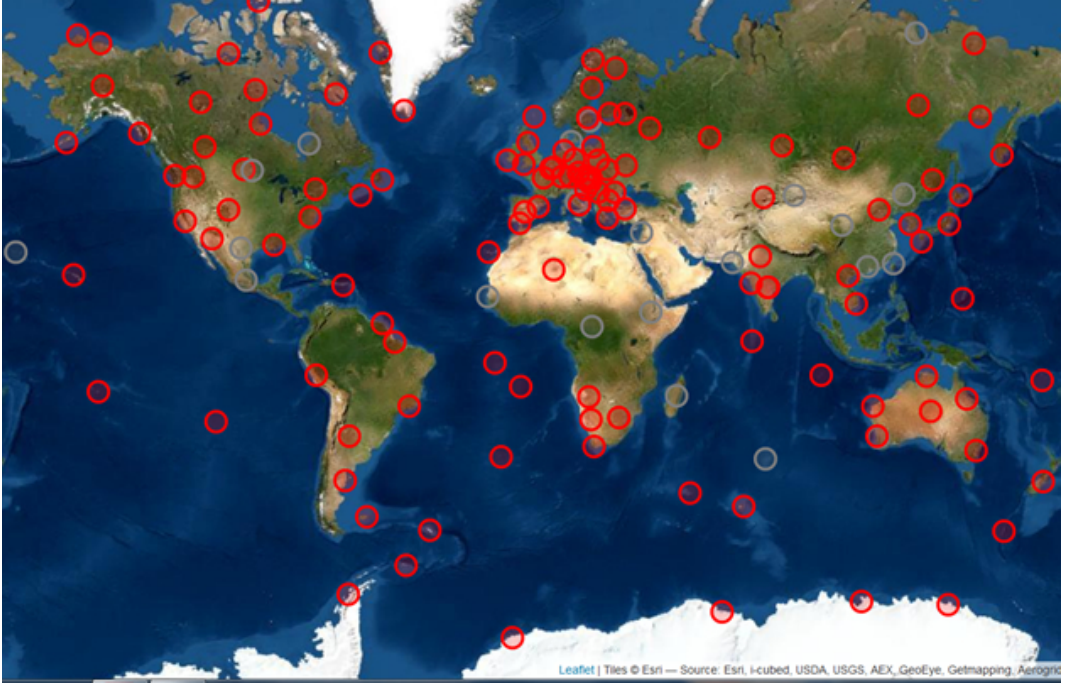
Rappelons que le potentiel magnétique est solution de l'équation de Laplace

$$\nabla^2 V(r, \theta, \varphi) = 0, \quad (1)$$

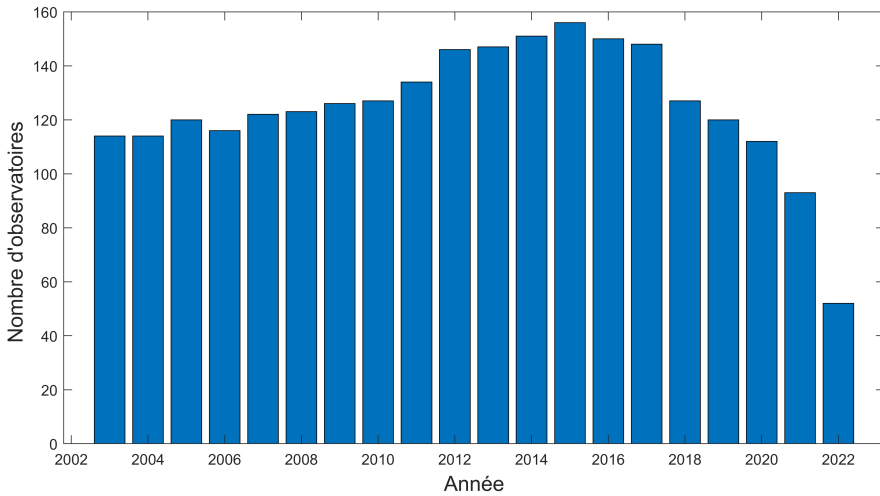
où le potentiel d'origine interne uniquement se développe en série infinie

$$V(r, t) = a \sum_{n,m} \beta_n^m(\theta, \varphi) \left[ \left( \frac{a}{r} \right)^{n+1} g_n^m(t) \right], \quad (2)$$

avec  $a$  le rayon de référence pour la Terre,  $r, \theta, \varphi$  les coordonnées géocentriques,  $g_n^m$  les coefficients de Gauss internes exprimés ici sous forme complexe en fonction du temps  $t$ . Les fonctions harmoniques sont écrites dans leur forme complexe  $\beta_n^m = e^{im\varphi} P_n^m(\cos\theta)$ , avec  $P_n^m(\cos\theta)$  les fonctions de Legendre au sens de Schmidt. Les indices  $n$  et  $m$  sont respectivement les degrés et les ordres entiers des harmoniques sphériques et  $i^2 = -1$  est le nombre imaginaire unitaire. Les trois composantes du champ magnétique selon les vecteurs unitaires  $\mathbf{u}_r, \mathbf{u}_\theta, \mathbf{u}_\varphi$ , sont obtenues



**FIGURE 1** Distribution géographique des observatoires magnétiques terrestres obéissant à la norme INTERMAGNET au 01/01/2022. Les cercles gris correspondent aux observatoires en maintenance en 2022.



**FIGURE 2** Nombre d'observatoires dont les données sont disponibles entre 2003 et 2022



en considérant le gradient (en coordonnées sphériques) du potentiel

$$\mathbf{B}(r, \theta, \varphi) = - \nabla V(r, \theta, \varphi); \quad (3)$$

le signe négatif étant posé par convention.

La construction du jeu synthétique des valeurs de  $\mathbf{B}$  aux positions des observatoires nécessite donc des modèles de champs magnétiques réalistes décrivant les sources internes. Pour représenter les contributions du champ principal nous utilisons le modèle CHAOS-7.17 (Finlay et al., 2020) dont la dernière version a été publiée Janvier 2024 (open source: <https://www.space.dtu.dk>). Le modèle CHAOS couvre plus de deux décennies et a été construit à partir des données des satellites Ørsted, CHAMP, SAC-C, Cryosat-2, Swarm et le réseau d'observatoires terrestres. Ce modèle se fonde sur une paramétrisation incluant les sources magnétosphériques, ionosphériques, celles induites dans le manteau et les sources crustales. La paramétrisation en temps est décrite par des splines d'ordre 6. Pour la construction du jeu de données synthétiques nous utilisons uniquement la partie décrivant le noyau pour les degrés harmoniques  $n=1$  à 20 ainsi que les variations temporelles du modèle; au-delà de 20 la partie statique interne contient les contributions de la partie crustale du champ magnétique et les contributions du champ externe ne sont pas calculées. Nous complétons ces valeurs du champ magnétique d'origine interne par des contributions crustales calculées aux observatoires à partir d'un modèle de champ lithosphérique (Thébault et al., 2021) mis à jour et calculé jusqu'au degré 1300 (échelles spatiales d'environ 30 km). Cette partie supplémentaire est destinée à simuler l'effet de site des observatoires magnétiques terrestres.

Nous avons ainsi un ensemble de coefficients de Gauss à partir du modèle CHAOS et les valeurs du champ vectoriel associé à la position de tous les observatoires INTERMAGNET.

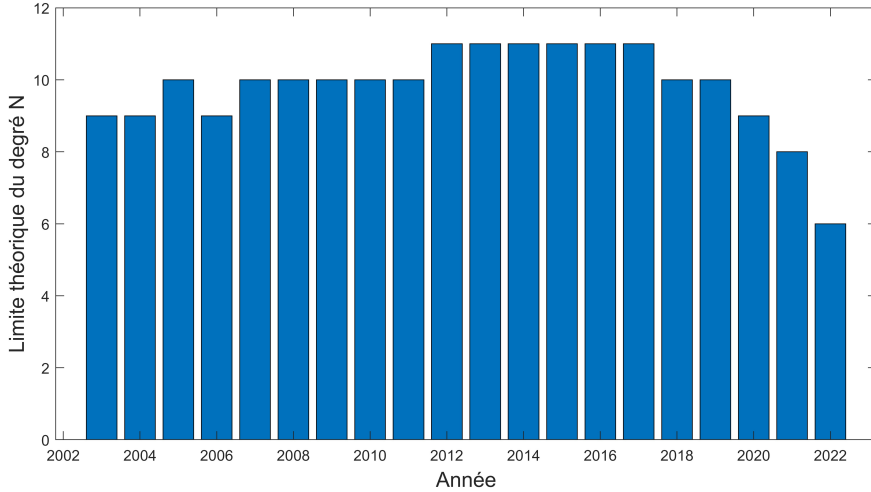
### 2.3 | Résolution du problème inverse

Pour estimer la résolution spatiale maximum représentable par les données d'observatoires magnétiques il convient à partir des données synthétiques du champ d'évaluer les coefficients obtenus et de confronter leurs valeurs aux coefficients initiaux de CHAOS. Le principe de modélisation revient à estimer les coefficients  $g_n^m$  des harmoniques sphériques à partir de la mesure discrète des composantes du champ magnétique  $\mathbf{B}(r, \theta, \varphi)$ .

Si  $\mathbf{A}$  est la matrice des harmoniques sphériques,  $\mathbf{b}$  le vecteur des valeurs vectorielles du champ et  $\mathbf{g}$  le vecteur des coefficients à estimer, alors le problème inverse se réduit à

$$\mathbf{g} = (\mathbf{A}'\mathbf{A})^{-1}(\mathbf{A}'\mathbf{b}). \quad (4)$$

Ce problème inverse est linéaire et ce formalisme n'est valable que pour un problème surdéterminé, c'est-à-dire lorsqu'il y a plus de données que de paramètres à estimer. La série infinie donnée par l'Eq. 2 est donc nécessairement tronquée en pratique à un degré maximum. Malgré la compacité des harmoniques sphériques, si  $N$  est le degré maximum de la série infinie (ou des harmoniques sphériques) alors le nombre de paramètres à estimer est  $N \times (N+2)$ . Les fonctions harmoniques sont orthogonales au sens du produit scalaire sur une sphère unitaire ce qui est une propriété essentielle pour la modélisation du champ à partir de mesures distribuées de manière discrète à la surface terrestre.



**FIGURE 3** Nombre d'observatoires par an et résolution  $N$  maximum atteignable dans le cas d'une distribution idéale.

Si les données sont suffisamment bien distribuées, cette propriété nous garantis que l'inversion de la matrice

$$(\mathbf{A}'\mathbf{A})^{-1}$$

est numériquement réalisable et que les coefficients obtenus sont indépendants les uns des autres ; c'est-à-dire que l'énergie est distribuée correctement selon les échelles spatiales ce qui permettra de calculer le champ en n'importe quel point en dehors des mesures initiales.

Le nombre disponible d'observatoires par année (figure 2) nous donne une limite sur le nombre maximum de coefficients, et donc le degré maximum des harmoniques sphériques, que nous pouvons estimer par une problème surdéterminé ( $N_d > N_g$ ). En imaginant que les observatoires sont distribués idéalement sur la sphère (ce qui reste à définir), un calcul simple nous permet d'estimer le degré maximum  $N$  compatible avec le nombre d'observatoires disponibles. Ce degré maximum permet de rester dans la validité du formalisme du problème inverse surdéterminé, c'est-à-dire pour  $N \times (N+2) < d$ . La figure 3 montre que la résolution numérique maximum possible avec cet ensemble d'observatoires est comprise entre  $N = 6$  et  $N = 11$  selon l'année considérée et que la cible de  $N = 13$  n'est pas atteignable quelle que soit l'année.

D'autre part, la figure 1 nous informe que ce nombre de données est mal distribué sur la sphère. Par conséquent, la matrice  $(\mathbf{A}'\mathbf{A})$  à inverser ne peut être orthogonale et des instabilités numériques sont à prévoir pour son inversion. Les degrés  $N$  de la figure 3 sont donc une estimation supérieure (optimiste) que cette complication supplémentaire ne permettra pas d'atteindre par une approche numérique simple sans ajouter de l'information destinée à régulariser cette matrice.

## 3 | RÉSULTATS

### 3.1 | Analyse préliminaire de la matrice d'inversion

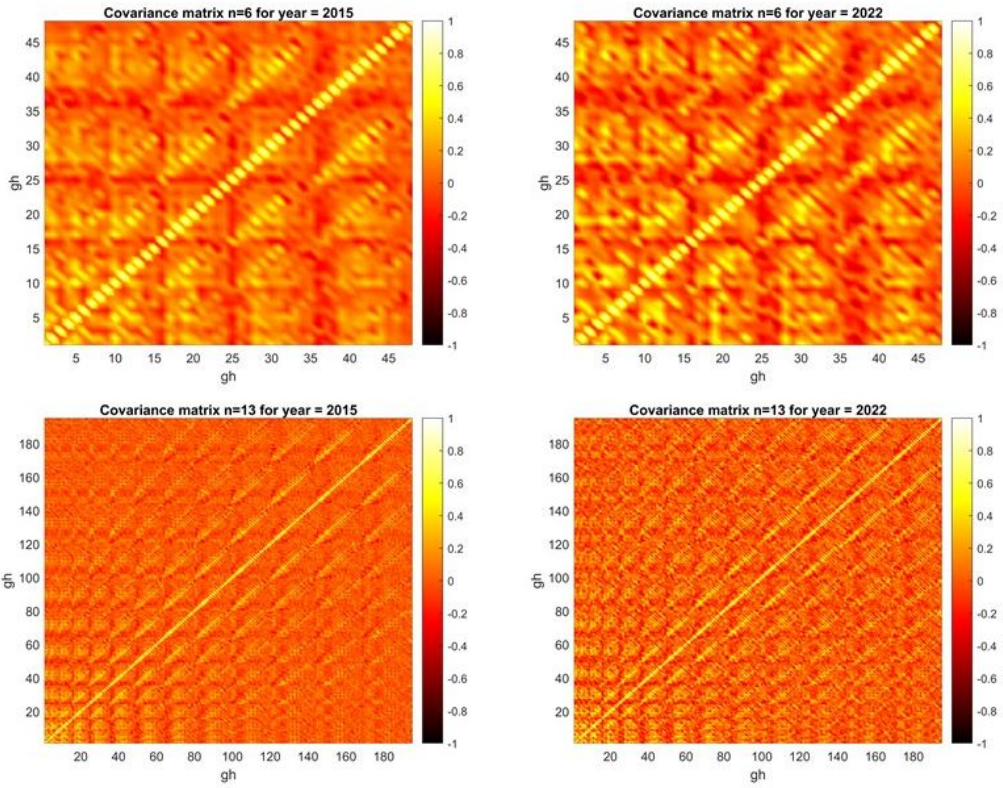
Nous analysons la matrice à inverser ( $A'A$ ) des fonctions harmoniques calculées aux positions des observatoires pour deux périodes offrant un nombre d'observatoires sensiblement différents. La figure 4 montre la matrice normalisée pour 2015, période qui comprend le plus de d'observatoires dans la base de données et pour 2022, période la moins favorable. Les développements sont réalisés jusqu'à  $n = 6$  et  $n = 13$ . Rappelons que pour  $n = 13$  aucune des matrices n'est invertible et cet exemple n'est choisi que pour la discussion. Idéalement, cette matrice devrait être diagonale en raison des propriétés mathématiques d'orthogonalité des harmoniques sphériques. Dans les deux cas, la matrice contient des éléments hors diagonale qui se présentent sous la forme de blocs dont l'intensité varie néanmoins entre 2015 et 2022 pour les deux expansions, à  $n=6$  et  $n=13$ . La matrice de covariance est à ce titre plus stable et propice à l'inversion en 2015 qu'en 2022. Néanmoins, il convient de préciser que les coefficients sont ordonnés de manière croissante et que la quadrant bas-gauche caractérise les très grandes échelles (bas degrés  $n$ ) et le quadrant haut-droite les plus petites échelles (hauts degrés  $n$ ) dans le modèle. Quelle que soit l'année, les covariances entre les coefficients de bas degrés et ordres (grandes échelles) sont importantes avec des valeurs atteignant  $\pm 0.6$  même en 2015 pour les degrés inférieurs ou égaux à  $n = 3$  ( $gh \leq 15$  sur la figure selon la règle que le nombre de coefficients pour les degrés est égal à  $n(n+2)$ ). Ces covariances sont à mettre en parallèle avec la distribution géographique des observatoires (Figure 1) et la dichotomie continent/océan. Sur Terre, la distribution longitudinale entre océan/continent, avec notamment l'océan Atlantique couvrant la zone du Nord au Sud sur des espacements de longitude allant de  $70^\circ$  -  $90^\circ$  environ (par exemple  $70^\circ$  entre Brest et New-York) entre deux continents, correspond déjà au degré des harmoniques sphériques  $n = 360/90 = 4$ . Les larges secteurs géographiques non couverts par la donnée dans les océans Pacifique et Antarctique contribuent également à cette limite de résolution qui est purement géométrique.

La présence d'éléments hors diagonale de valeur relativement significative indique que les coefficients lorsqu'ils seront estimés par inversion seront corrélés entre-deux. Ainsi, non seulement la solution est susceptible de varier significativement à chaque changement du degré de résolution par une réorganisation des relations linéaires entre les coefficients mais cette solution sera d'autre part potentiellement très sensible au niveau de bruit dans les données ou aux contaminations du champ qui n'est pas représentable par l'ensemble tronqué des paramètres du modèle. Ces covariances entraîneront une manque de stabilité et de confiance dans la prédiction du champ magnétique en dehors des points de mesures.

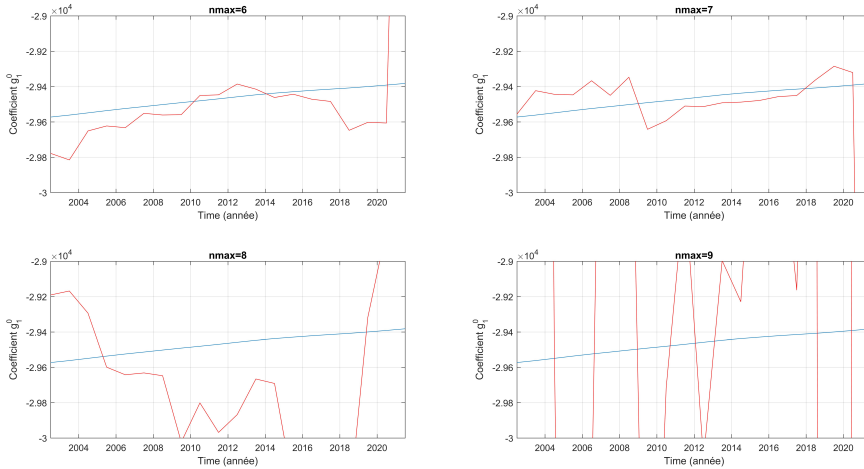
### 3.2 | Effets de la variation temporelle et spatiale de la distribution

La solution par problème inverse de l'équation Eq. 4 est obtenue pour chaque année entre 2003 et 2022 pour différents degrés maximums entre  $n = 6$  et  $n = 9$ . Chaque inversion consiste à estimer  $n(n + 2)$  degrés à partir des mesures vectorielles du champ synthétique aux observatoires. Nous utilisons plusieurs critères de comparaison entre les modèles estimés et le modèle théorique de CHAOS. Ces critères sont décrits en détails dans Thébault and Gailler (2023) (section 3.4 Eqs. 8-17). Le premier critère est la comparaison simple coefficient par coefficient. Pour illustrer la problématique, nous montrons en Figure 5 l'évolution du coefficient du dipôle magnétique axial théorique et sa reconstruction durant toute la période considérée.

Nous choisissons de discuter le coefficient axial car, décrivant la partie strictement axiale et dipolaire du champ magnétique, nous pourrions supposer que sa reconstruction ne pose pas de difficulté. Pourtant, la figure 5 montre une grande disparité entre les estimations selon le degré maximum de résolution  $n$ . Le résultat premier est que son



**FIGURE 4** Matrices d'inversion normalisée pour les périodes 2015 et 2022 et pour des degrés maximum de 6 et de 13, respectivement

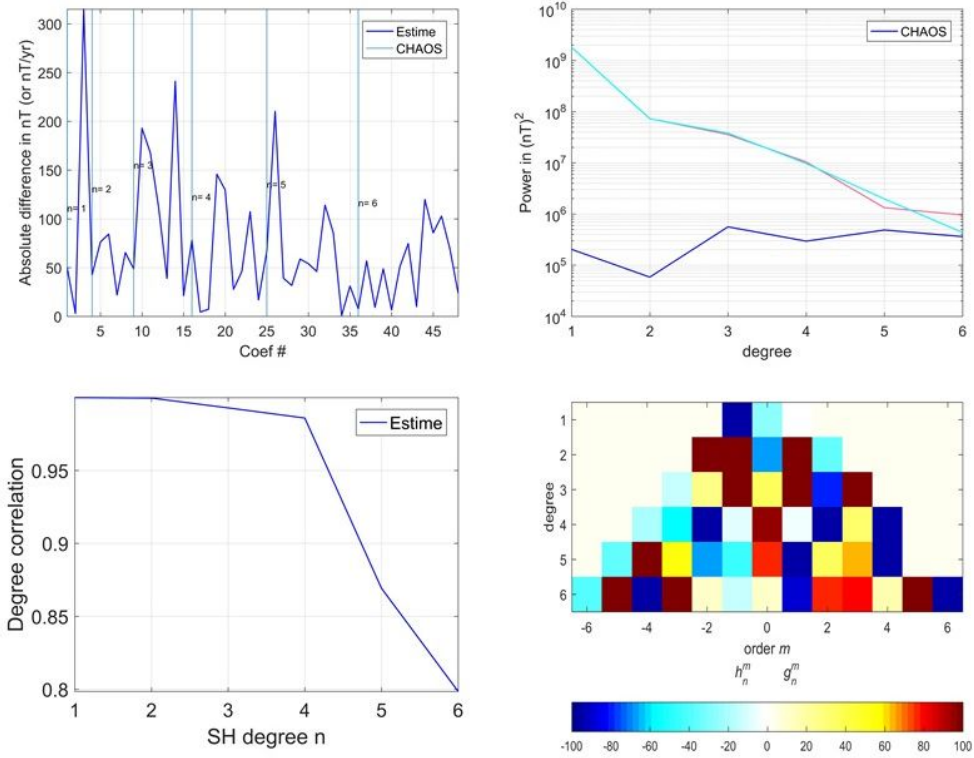


**FIGURE 5** Estimations du coefficient axial  $g_1^0$  du champ magnétique en fonction du temps à partir des séries temporelles des observatoires magnétiques (en rouge) et leurs comparaison avec la valeur réelle (en bleu).

estimation se dégrade très rapidement en fonction du degré maximum des harmoniques sphériques. Ce manque de stabilité était anticipé à l'examen de la matrice d'inversion dans laquelle chaque accroissement du degré de résolution impliquait effectivement une réorganisation des coefficients (et donc leur instabilité). Dans le cas le plus favorable ( $n = 6$ ), nous observons également un biais systématique dans l'estimation de ce coefficient qui est de plus couplé à une variation temporelle artificielle autour de la tendance réelle du coefficient du modèle CHAOS. La raison de ces biais provient de la variation du nombre d'observatoires avec le temps qui en créant une différence de couverture spatiale se traduit en variations temporelles. Les contributions du noyau non représentées (degrés  $6 < n < 20$ ) ainsi que celle du champ crustal (degré jusqu'à  $n = 1300$ ) entraînent cet effet de site qui se traduit par une « fuite spectrale » de ces composantes manquantes à plus haut degré vers les harmoniques du champ magnétique de bas degré.

Si l'on examine maintenant la partie statique des modèles en nous focalisant sur l'époque arbitraire de 2020.0 nous observons le détail des disparités. La figure 6-haut-gauche montre les différences absolues coefficients par coefficients. Les différences sont organisées par  $n$  et  $m$  croissants. Nous observons des pics de différences proches de chaque changement de degré  $n$  qui correspondent aux ordre  $m = 1$ . Géométriquement ces harmoniques correspondent aux fonctions  $\cos(\phi)$  et  $\sin(\phi)$  et cette analyse montre de larges secteurs longitudinaux (entre  $90^\circ$  et  $180^\circ$ ) très mal contraints par la disposition géographique des observatoires.

La figure 6-haut-droite montre le spectre magnétique de CHAOS (en bleu clair), le spectre du modèle reconstruit (en rouge) et le spectre de la différence entre les deux modèles. Nous remarquons que la différence entre les deux modèles suit au mieux un spectre blanc et que cette différence recoupe la valeur de l'énergie du spectre original. Le spectre reconstruit semble commencer à diverger à partir de  $n = 5$ . La figure 6-bas-droite montre la corrélation spectrale entre les deux modèles. Cette corrélation indique qu'en dépit d'une augmentation du spectre reconstruit à partir de  $n = 5$  et d'une différence importante à  $n = 6$  la corrélation spectrale est meilleure que 0.8. Ce résultat signifie que les coefficients du modèle reconstruit sont individuellement trop énergétiques mais que dans l'ensemble ils restent en accord acceptable avec le modèle initial. Ce constat permet d'envisager des techniques de régularisation destinée à réduire l'énergie artificielle des coefficients problématiques tout en gardant le contenu de l'information portée par



**FIGURE 6** Comparisons entre le modèle obtenu en 2020.0 pour  $n = 6$  et le modèle initial CHAOS. Les critères de comparaison sont décrits dans Thébault and Gailler (2023) les différences coefficient par coefficient (haut gauche), le spectre de puissance de CHAOS, le modèle reconstruit, et leur différence (haut droite), l'analyse de corrélation spectrale entre les deux modèles (bas gauche) et la matrice de sensibilité (bas droite).

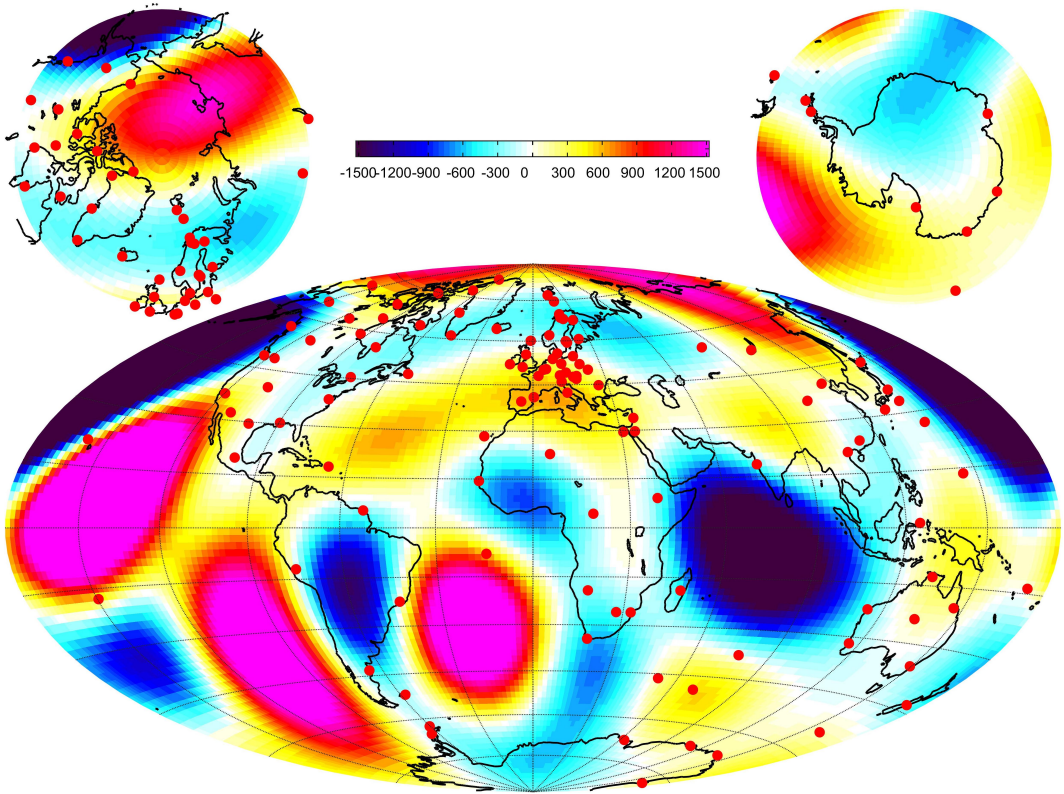
ces coefficients. Enfin, la figure 6-bas-gauche montre la matrice de sensibilité. Il s'agit d'une matrice qui identifie les différences normalisées entre les coefficients en pourcentages. Nous y distinguons clairement les coefficients les plus problématiques qui, terme à terme, sortent de la tendance. Parmi les coefficients reconstruits dont l'erreur atteint presque 100 pourcent nous reconnaissons les coefficients d'ordre  $m = 1$ . Certains coefficients d'ordre  $m = 1$  sont particulièrement bien contraints tels que le coefficient  $g_1^1$  et  $g_4^1$ . Nous reconnaissons aussi des différences importantes pour les coefficients d'ordre  $n = 2$ . Géométriquement, les harmoniques correspondantes décrivent la dichotomie Nord/Sud sur la sphère. Les erreurs croissantes sur ces coefficients sont donc dues à la disparité en couverture de données entre les hémisphères Nord et Sud. Tous ces indicateurs permettent de comprendre les différences entre le champ magnétique calculé à partir de CHAOS et celui calculé à partir du modèle reconstruit à la surface terrestre. La figure 7 montre la différence entre les composantes radiales du modèle CHAOS et le modèle reconstruit. Les différences spatiales sont très concentrées sur de larges zones non couvertes par les données, dans l'océan Atlantique, en Amérique du Sud, dans l'océan Indien etc. Notons à ce propos que la zone la plus dynamique du champ magnétique connue comme étant "l'anomalie de l'Atlantique Sud" est une zone particulièrement mal décrite par la distribution actuelle des données d'observatoires. Des projets sont en cours (station magnétique à Kourou, Guyane) pour aider à mieux contraindre cette structure magnétique. La zone Arctique est également problématique car il s'agit de la zone où se trouve le pôle nord magnétique dont la dynamique s'est accélérée ces 20 dernières années. Cette zone géographique est d'autre part cruciale pour la navigation aéronautique. Il est important de noter également que cette base INTERMAGNET étant limitée, un travail de moissonnage de données disponibles hors de ce réseau aux spécifications contraignantes serait très bénéfique. De manière toute à fait intéressante, le continent Antarctique présente une différence relativement faible par rapport à certaines autres régions du monde. La distribution circulaire des observatoires magnétiques terrestres en bordure du continent Antarctique semble favorable et bien contraindre le champ sur l'intégralité du continent Antarctique pour cette résolution maximum.

### 3.3 | Exploration statistique des sous-ensembles

À titre exploratoire nous avons réalisé des simulations numériques à partir de sous-ensembles pris au hasard dans la base de données complète des observatoires. L'idée sous-jacente est que l'orthogonalité des fonctions harmoniques sur la sphère est très sensible à la distribution discrète des données. Certaines configurations spatiales que nous ne connaissons pas encore sont donc nécessairement plus favorables que d'autres et en particulier plus favorables que l'ensemble complet. Par exemple, l'ensemble complet montre une concentration du nombre d'observatoires sur l'Europe (Figure 1) qui est susceptible de biaiser le modèle vers cette région au détriment des autres régions. Sur le plan numérique, un meilleur modèle est sans doute envisageable avec une distribution plus en accord avec les règles d'échantillonnage sur la sphère qui permettrait de mieux approcher l'orthogonalité numériques des fonctions harmoniques et donc de séparer correctement les coefficients. Ces règles d'échantillonnage strictes peuvent suivre une distribution de mesures équisurfaces sur la sphère, une distribution proche de l'échantillonnage de Fourier-Gauss etc. mais ne seront jamais atteintes dans la réalité. Néanmoins, des sous-ensembles d'observatoires sont peut-être plus proches de cette distribution mathématique formelle que ne l'est la distribution complète. Ces cas particuliers sont associés à une meilleure orthogonalité des fonctions harmoniques ce qui garantirait en retour une moindre contamination par l'information non modélisée (information portée par les harmoniques de degrés supérieures à  $n = 6$  dans le cas présent) et autoriserait une reconstruction à plus haut degré.

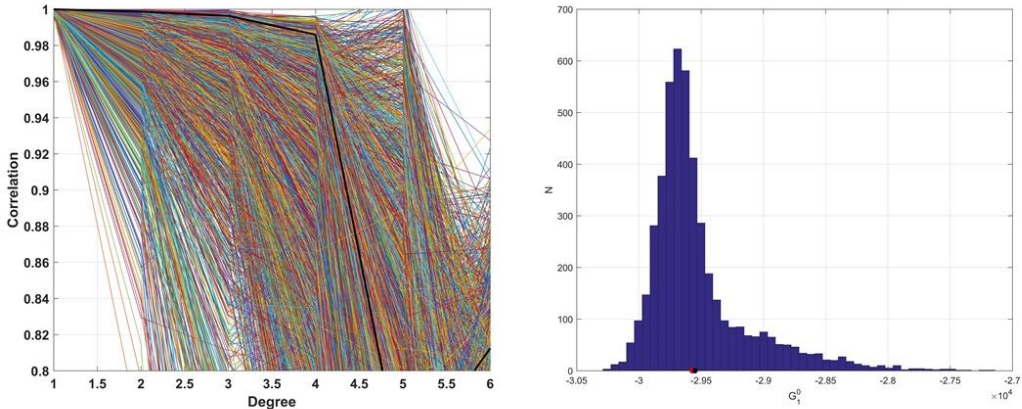
Sur le plan pratique nous appliquons une méthode de « bootstrap » qui consiste à explorer de manière aléatoire 100000 distributions d'observatoires différentes. Pour chaque tirage nous calculons un modèle jusqu'à  $n = 6$ . Chaque sous-ensemble aléatoire est constitué de 75 pourcent de la base complète. En prenant toujours l'année 2020





**FIGURE 7** Différence en nT entre la composante magnétique verticale du modèle initial et le modèle reconstruit à la surface terrestre.





**FIGURE 8** Exploration des statistiques possibles par la technique du bootstrap. Ensemble des corrélations harmoniques pour chaque ensemble (gauche) et histogramme des valeurs estimées pour le coefficient zonal axial  $g_1^0$  (droite). La valeur réelle du coefficient axial est indiquée en rouge.

comme référence arbitraire cela revient à prendre aléatoirement des sous-ensembles de 84 observatoires sur les 112 disponibles. D'après la Figure 3, ce nombre théorique permet d'aller jusqu'à une résolution maximum de  $n = 8$ . Pour chacune des 100000 configurations nous estimons le modèle et nous calculons sa corrélation spatiale avec le modèle réel CHAOS pour la même époque. La figure 8-gauche montre ces calculs d'ensembles aléatoires. Une première remarque évidente est que la plupart de ces sous-ensembles conduisent à des modèles fortement dégradés, ce qui est attendu dans ce genre d'exercice aveugle. Néanmoins, la comparaison avec la Figure 6 (qui montre que la corrélation au degré  $n = 6$  pour la base complète n'est pas meilleure que 0.8) nous permet de conclure effectivement que des sous-ensembles d'observatoires conduisent à des corrélations bien meilleures dont l'un en particulier atteint 0.92 à  $n = 6$ . Il s'agira par la suite d'identifier plus précisément ces distributions pour définir sa résolution maximum. La tendance à partir de  $n = 6$  laisse envisager une solution acceptable jusqu'à  $n = 7 - 8$ . Un autre résultat saillant de l'exercice est illustré par la figure 8-droite. Les valeurs des 100000 coefficients axiaux  $g_1^0$  que nous synthétisons par un histogramme suivent une distribution dont le mode est maintenant presque centré sur la valeur réelle du coefficient. Par conséquent, cette méthode d'ensemble permet bien de corriger en partie du biais (pour peu que nous considérons le mode de l'ensemble plutôt que sa moyenne) au moins pour ce coefficient.

## 4 | CONCLUSIONS

L'objectif de cette étude numérique était d'estimer la résolution maximum en harmoniques sphériques du champ magnétique terrestre d'origine interne qu'il est possible d'atteindre avec la distribution des observatoires magnétiques terrestres. Cette question initiale se heurte à un obstacle important concernant la définition même d'une base de données d'observatoires terrestres exhaustives pour cette étude. Les observatoires terrestres considérés ici sont ceux du réseau international INTERMAGNET (International Real-time Magnetic Observatory Network). Le programme INTERMAGNET a pour but d'établir un réseau mondial d'observatoires magnétiques numériques en coopération, en adoptant des spécifications standard modernes pour les équipements de mesure et d'enregistrement, afin de faciliter

les échanges de données et la production de produits géomagnétiques en temps quasi réel. Ces données sont collectées et diffusées sur un serveur dont l'un des hébergements est réalisé par le British Geological Survey (BGS). Ses principes, spécifications techniques et politique scientifique sont définis au niveau de l'Association Internationale de Géomagnétisme et d'Aéronomie depuis 1987 avec comme contrainte de base une qualification de la donnée et sa pérennisation dans le temps. Il ne peut donc s'agir d'une base de données exhaustive et exclusive. Il n'existe à notre connaissance pas de centre de données répertoriant les observatoires magnétiques terrestres en opération à une période donnée et délivrant des mesures standardisées. Les conclusions de ce rapport sont donc limitées à ce nombre restreint d'observatoires terrestres. Il existe vraisemblablement un nombre bien plus important de mesures disponibles qui ne sont pas référencées dans ce réseau faute d'atteindre ses spécifications techniques contraignantes. Or, partant du principe qu'une donnée de moindre qualité vaut mieux qu'aucune donnée, comme nous l'avons également montré dans les résultats précédents, une action envisageable future est de réaliser un travail de recensement et de veille des stations terrestres. Il s'agirait d'un travail conséquent de veille et de développement de moissonnage de la donnée pour s'adapter à leur format.

A partir du réseau INTERMAGNET nous avons montré que la résolution maximum atteignable pour le moment était l'harmonique au moins égale à  $n = 6$ . Néanmoins, nous avons également montré que la considération de sous-ensembles permettait d'améliorer cette résolution spatiale jusqu'à environ  $n = 8$ . Le travail à poursuivre est donc d'identifier ces sous-ensembles plus favorables. L'une des limites actuelles est liée à la fuite spectrale des contributions de plus haut degré dans ces bas degrés. Parmi ces contributions non modélisées nous trouvons toutes les harmoniques du champ supérieures à  $n = 6 - 8$  dont certaines proviennent du noyau et d'autres du champ crustal dont l'effet de site peut atteindre le micro-testla. Nous explorerons donc la valeur statistique de ces effets de site afin de prévoir leur prise en compte, soit de manière déterministe en corrigeant les données de cette prédiction, soit de manière statistique en les utilisant pour un arbitrage de la sélection des sous-ensembles ou en les incorporant comme des erreurs de modélisation. Enfin, une des pistes à explorer reste la régularisation du champ dont l'objectif est de réduire l'énergie des paramètres du modèle mal contraints par la distribution des données.

Les trois prochaines étapes sont donc de tester différentes régularisations, de réaliser une base de données concernant le biais crustal, pour enfin tenter de déterminer les sous-ensembles les plus favorables. En parallèle, il sera utile de réaliser dans la mesure du possible un inventaire des stations magnétiques de proche surface (observatoires, variomètre, etc.) qui permettent de compléter au mieux la base d'INTERMAGNET en regard de ses limites observées. L'absence de bases de données existantes concernant ces mesures « orphelines » ou hors spécifications INTERMAGNET est néanmoins compliqué et ne pourra pas faire l'objet d'une veille exhaustive.

## references

- Finlay, C. C., Kloss, C., Olsen, N., Hammer, M. D., Tøffner-Clausen, L., Grayver, A. and Kuvshinov, A. (2020) The chaos-7 geomagnetic field model and observed changes in the south atlantic anomaly. *Earth, Planets and Space*, **72**, 1–31.
- Thébault, E. and Gailler, L.-S. (2023) Modélisation du champ magnétique principal par inversion de données d'observatoires (magiobs): état de l'art. *Tech. rep.*, MAGIOBS, DOC001.
- Thébault, E., Hulot, G., Langlais, B. and Vigneron, P. (2021) A spherical harmonic model of earth's lithospheric magnetic field up to degree 1050. *Geophysical Research Letters*, **48**, e2021GL095147.