



**HAL**  
open science

## Projet SNCF maths-entreprises AMIES.

Matthieu Dolbeault, Lucas Journal, Rémi Robin

► **To cite this version:**

Matthieu Dolbeault, Lucas Journal, Rémi Robin. Projet SNCF maths-entreprises AMIES.. [Rapport de recherche] Laboratoire Jacques-Louis Lions, Sorbonne université. 2021. hal-03345577

**HAL Id: hal-03345577**

**<https://hal.science/hal-03345577v1>**

Submitted on 15 Sep 2021

**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.



# Projet SNCF maths-entreprises AMIES, Premier Prix

Matthieu Dolbeault\*, Lucas Journal\*, Rémi Robin\*,  
Laboratoire Jacques-Louis Lions, Sorbonne Université

Septembre 2021

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Énoncé du problème . . . . .	2
1.2	Data sets . . . . .	2
1.3	Ordres de grandeur et sources d'erreurs . . . . .	3
<b>2</b>	<b>Un premier modèle spatial</b>	<b>5</b>
<b>3</b>	<b>Influence de la température et débruitage temporel</b>	<b>10</b>
<b>4</b>	<b>Un modèle d'inférence plus complexe</b>	<b>16</b>
4.1	Rail non parfait . . . . .	16
4.2	Modélisation statistique . . . . .	17
4.3	Estimation des erreurs . . . . .	19
4.4	Détermination des angles initiaux . . . . .	19
<b>5</b>	<b>Etude du cas de Farébersviller</b>	<b>22</b>
5.1	Introduction et qualité des données . . . . .	22
5.2	Reconstruction de l'affaissement . . . . .	24
<b>6</b>	<b>Conclusion et améliorations envisageables</b>	<b>30</b>

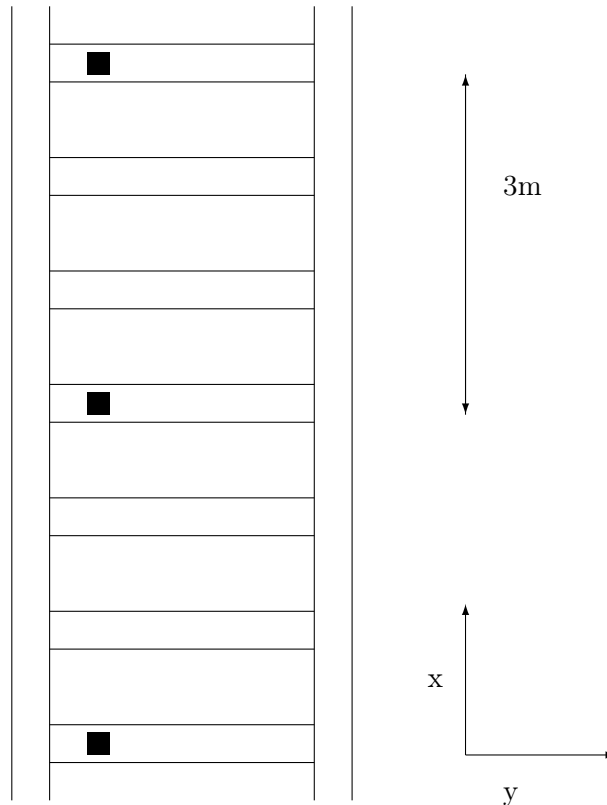
---

\*prenom.nom@sorbonne-universite.fr

# 1 Introduction

## 1.1 Énoncé du problème

Le but de ce projet est d'étudier les variations d'altitudes d'une voie suite à des travaux à proximité, à l'aide de capteurs inclinométriques.



Plus précisément, on installe à intervalles réguliers des capteurs sur les traverses d'un rail, espacés d'environ 3m (voir schéma 1.1). A intervalles de temps réguliers (de l'ordre de quelques minutes), les capteurs mesurent la variation de leur inclinaison (rotation autour de l'axe  $y$ , c'est-à-dire variation de pente dans la direction  $x$ ). Le but est ensuite d'obtenir la variation du profil longitudinal de la voie en fonction du temps.

Un travail similaire a déjà été effectué par la SNCF pour l'étude du dévers (la variation angulaire autour de l'axe  $x$ ). On se concentrera donc ici sur l'évolution de l'altitude (que l'on notera  $z$ ) en fonction de  $x$ .

## 1.2 Data sets

Nous disposons de trois jeux de données différents. Un premier jeu a été relevé lors des travaux près de la gare de Mittersheim. Le deuxième et le

troisième correspondent aux données de deux voies près de Farébersviller.

Ces données présentent l'évolution de la température et des mesures d'angles en fonction du temps. Elles contiennent également le profil initial de la voie, ainsi que des données de contrôle dans le cas de Farébersviller.

Data set	Mittersheim	Farébersviller (V1/V2)
nombre de capteurs (dont utilisables)	58 (54)	75 (75/74)
longueur de voie équipée (m)	167.5	222
date début	18/11/2019	03/03/2021
date des relevés	18/11/2019	03/03/2021 06/04/2021 10/05/2021 03/06/2021
date fin	15/12/2019	06/06/2021

Les données de Mittersheim nous ont été présentées comme plus fiables que celles de Farébersviller. En particulier les variations de température importantes lors des mesures à Farébersviller mettent en doute leur intégrité. Cependant les données de Mittersheim présentent l'inconvénient de ne pas présenter de relevés de contrôle. Nous avons donc décidé dans les trois premières parties d'illustrer nos méthodes sur les données de Mittersheim. La partie 5 appliquera notre modèle aux données de Farébersviller.

### 1.3 Ordres de grandeur et sources d'erreurs

L'ordre de grandeur de l'affaissement est le cm par mois de travaux (voir Figure 1). Afin de reconstruire l'affaissement, nous développerons plusieurs modèles ultérieurement. L'objet de cette partie est d'identifier les sources d'erreurs potentielles et d'essayer de les quantifier.

1. Les erreurs de relevé d'altitude. Elles apparaissent lors de l'initialisation du profil initial et lors des mesures de contrôle. Elles limitent intrinsèquement toutes les estimations que nous ferons. Nous les estimons à  $\pm 2\text{mm}$ .
2. Les erreurs de référentiel. Nous nous ramènerons systématiquement à un problème à 1 dimension et ne modéliseront pas les variations d'altitude due au devers (les capteurs ne sont pas au milieu de la voie). De plus, même si les voies étudiées sont rectilignes, on remarque par exemple une erreur sur l'alignement de l'axe x avec la voie dans les relevés initiaux de Mittersheim l'ordre de 0.7 degré. Cette erreur a pour seul impact de sous estimer légèrement les distances entre les capteurs. De même sur les donnée de Farébersviller nous n'avons pas la position précise des capteurs (nous interpolons donc tous les 3m). Encore une fois nous négligerons ces erreurs. Enfin, les capteurs ne

sont pas tous correctement alignés angulairement. L'erreur angulaire initiale selon la direction  $x$  (l'angle important) empêche l'utilisation directe des mesures ; seules les variations depuis les mesures initiales seront utilisées dans la suite. Par ailleurs, un écart d'un degré autour de l'axe  $z$  fausse les mesures d'environ 0.02% ce qui justifie de négliger ces erreurs.

3. Les erreurs dues aux variations des points de références. Des mesures angulaires seules ne permettent pas de récupérer une altitude. Il est nécessaire de disposer d'un point de mesure altimétrique, pour cela nous supposons plus loin disposer d'un point fixe (typiquement les extrémités de la zone de mesure seront supposé fixes). D'après les mesures sur Farébersviller, l'affaissement peut atteindre 0.5 cm sur les extrémités (Figure 1). Cependant cette erreur est systématique et n'affectera nos estimations "qu'à une constante près".
4. Les erreurs de mesures. Les capteurs possèdent une résolution de  $10^{-4}$  degrés, Les fabricants précisent cependant une répétabilité de  $2.5 \cdot 10^{-3}$  degrés. Or nous montrerons que les données présentent des erreurs de d'ordre de plusieurs  $10^{-2}$  degrés voir plus en cas de grandes variations de température. Ces erreurs ne sont par ailleurs pas indépendantes. Une erreur systématique de  $\Delta\theta = 0.01$  degré donne une erreur en altitude  $\Delta z$  au bout de  $l = 100$  m de l'ordre de  $\Delta z \approx \frac{\pi}{180} l \Delta\theta \approx 1.7$  cm ; c'est à dire du même ordre de grandeur que l'affaissement.

Dans ce rapport, nous nous focaliserons sur les bruits relatifs aux mesures afin de les quantifier précisément et de les éliminer au maximum.

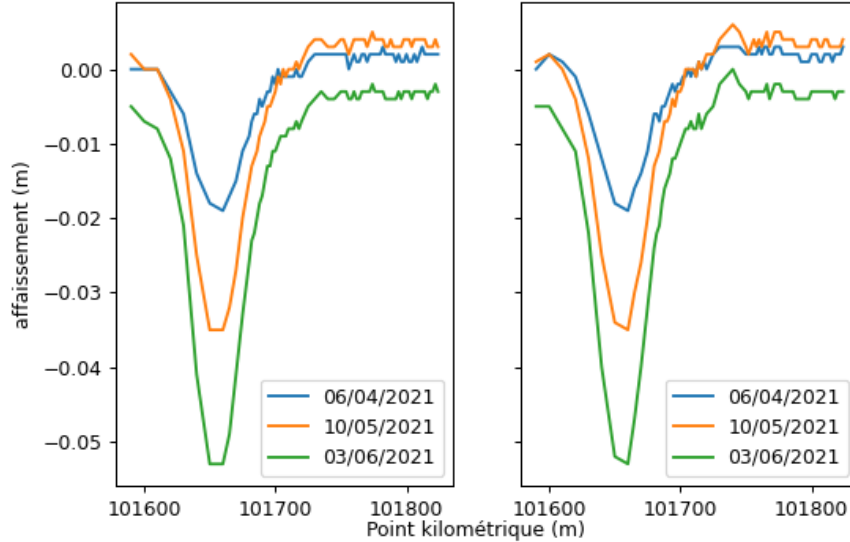


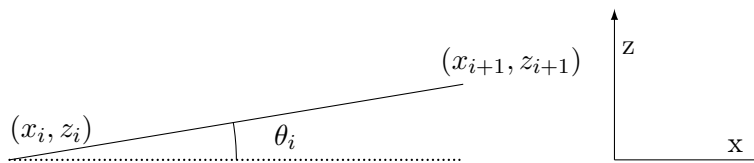
FIGURE 1: Affaissement relatif des deux voies de Farébersviller par rapport au premier relevé (le 03/03/2021)

## 2 Un premier modèle spatial

Un premier modèle simple est le suivant : Les capteurs sont aux extrémités de barres fixes d'environ 3m. Chaque mesure de capteur donne la variation de l'inclinaison de la barre le reliant au capteur suivant. A partir de la variation de l'angle et de l'altitude du capteur, on retrouve l'altitude de l'extrémité droite à partir de l'extrémité gauche.

Comme mentionné dans l'introduction, nous n'avons accès qu'aux variations d'angles. Ce modèle ne modélise donc que les variations d'altitude et la donnée du profil d'altitude initial est nécessaire.

On suppose que le premier et le dernier capteur sont fixes. Ils sont suffisamment loin des travaux pour qu'on puisse supposer qu'ils sont immobiles.



On calcule alors l'altitude des capteurs de proche en proche à partir du premier capteur, considéré par hypothèse comme fixe.

On note  $(x_i, z_i)$  la position du capteur numéro  $i$ , pour  $1 \leq i \leq N$  (pour Mittersheim,  $N = 57$ ). On note  $m_i(t)$  la variation d'angle mesurée par le

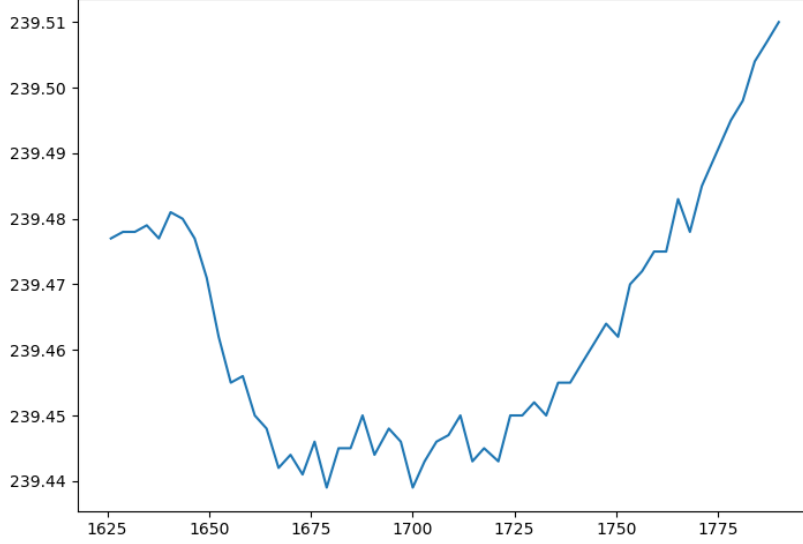


FIGURE 2: Profil initial de la voie de Mittersheim

capteur numéro  $i$  entre le temps initial et le temps  $t$ , et  $\theta_i(t)$  l'angle du capteur au temps  $t$ . On a la formule :

$$\theta_i = \arctan \frac{z_{i+1} - z_i}{x_{i+1} - x_i}.$$

Cette formule permet de calculer les angles initiaux des capteurs (voir aussi le profil initial figure 2). On calcule ensuite l'angle par la formule :

$$\theta_i(t) = \theta_i(0) + m_i(t),$$

et on obtient le profil de la voie par l'inverse de la formule précédente :

$$z_{i+1}(t) = z_i(t) + (x_{i+1}(t) - x_i(t)) \tan(\theta_i(t))$$

partant de  $z_1(t) = z_1(0)$ . On appelle ci-dessous ce modèle le modèle "forward". On peut observer que l'hypothèse que le dernier capteur est immobile n'est pas vérifiée dans ce modèle (voir figure 2).

Un modèle équivalent est le modèle dit "backward", où le dernier capteur est le point de départ du calcul récursif de l'altitude. Pour moyenner les erreurs des deux modèles, on considère le modèle "average"

$$z_i^a = \frac{z_i^f + z_i^b}{2}$$

qui correspond à une moyenne entre les résultats obtenue par les modèles "forward" et "backward", et le modèle "convex"

$$z_i^c = \frac{x_i - x_1}{x_N - x_1} z_i^b + \frac{x_N - x_i}{x_N - x_1} z_i^f$$

qui correspond à une combinaison convexe des deux. Ce dernier modèle a l'avantage de conserver l'immobilité des deux capteurs aux extrémités, et accorde en chaque point plus de confiance au modèle ("forward" ou "backward") qui part de l'extrémité la plus proche.

Toutes ces méthodes sont assez instables : les erreurs sur chaque mesure s'accumulent dans l'espace puisque l'on calcule les variations d'altitude de proche en proche. On remarque que deux estimations dans la même journée (avant les travaux) donnent des prédictions différentes, comme l'illustre la figure 4. Par ailleurs ces modèles ne permettent pas de donner des bornes d'erreurs sur le résultat (mis à part les variations d'altitude en  $x_N$  pour le modèle "forward", et en  $x_1$  pour le "backward", qui donnent une idée de l'erreur commise). Ces deux problématiques sont l'objet des deux parties suivantes.



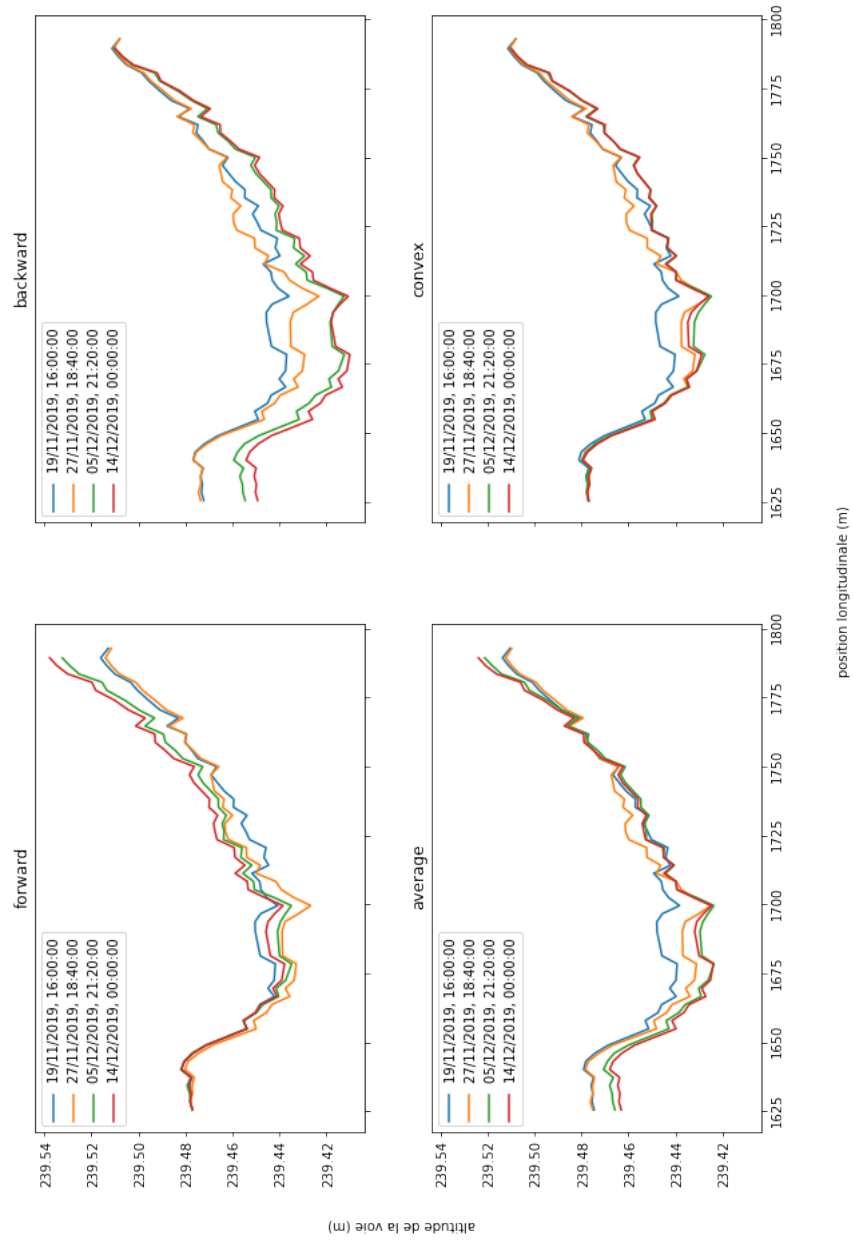


FIGURE 3: Evolution du profil de la voie selon le premier modèle.

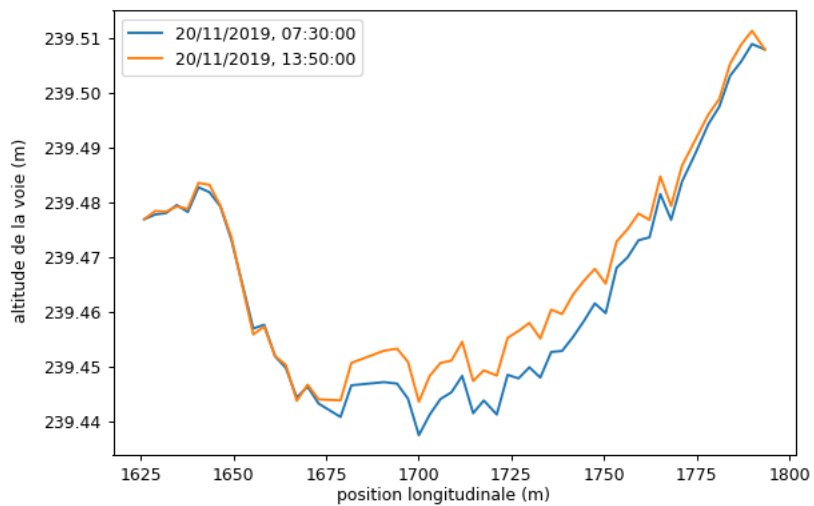


FIGURE 4: Utilisation de l'interpolation 'convexe' pour calculer deux profils datant du même jour, à des heures différentes.

### 3 Influence de la température et débruitage temporel

Le nombre élevé de mesures temporelles offre la possibilité d'améliorer nos estimations en agglomérant les mesures. Cependant les données capteurs ne sont pas seulement soumises à un bruit stationnaire. En effet, la figure 5 présente les données d'inclinaison de trois capteurs. Le capteur 1 et le 57 sont supposés hors de la zone d'affaissement, donc devraient être stables au cours du temps. A contrario le capteur 32 est dans la zone d'affaissement, on s'attend à ce qu'il indique des variations d'inclinaison, mais son signal semble quand même bruité.

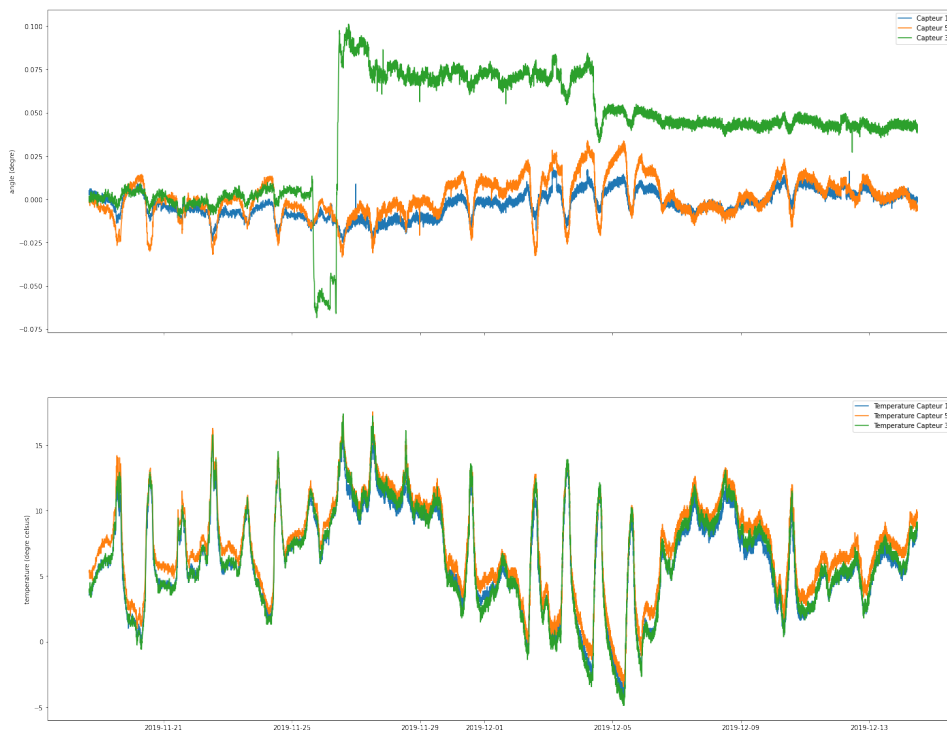


FIGURE 5: Valeurs d'inclinaison (en haut) et de température (en bas) mesurées par les capteurs 1, 32 et 57 en fonction du temps. Le capteur 32 enregistre des sauts d'inclinaison dus à l'affaissement de la voie, mais les trois capteurs ont des oscillations journalières, fortement corrélées avec la température.

On distingue clairement deux types de bruits :

- Un bruit haute fréquence d'amplitude de l'ordre de  $5 \cdot 10^{-3}$  degré, ce qui correspond à la variance des mesures indiquée par les spécifications des capteurs.
- Un bruit de fréquence beaucoup plus faible (période journalière), d'amplitude pouvant atteindre  $3 - 4 \cdot 10^{-2}$  degrés.

Le capteur 32 présente en plus des variations importantes (on observe 4 paliers) mais ces variations semblent être porteuse d'information et il nous faut les conserver. Comme le suggère la figure 5, il y a une corrélation entre ce bruit basse fréquence et la température.

Cette corrélation est illustrée par la figure 6. On y observe que la température et les mesures sont très corrélés (pas nécessairement positivement).

capteur	coefficient de corrélation	pente initiale (mm/m)
1	-0.884	0.34
19	0.925	2.05
32	0.002	-0.57
57	-0.942	-0.58

Par ailleurs la corrélation pour le capteur 32 ( qui se trouve dans la zone de travaux) semble exister. Seulement cette corrélation est locale en temps (en l'absence de mouvement d'affaissement, on a corrélation).

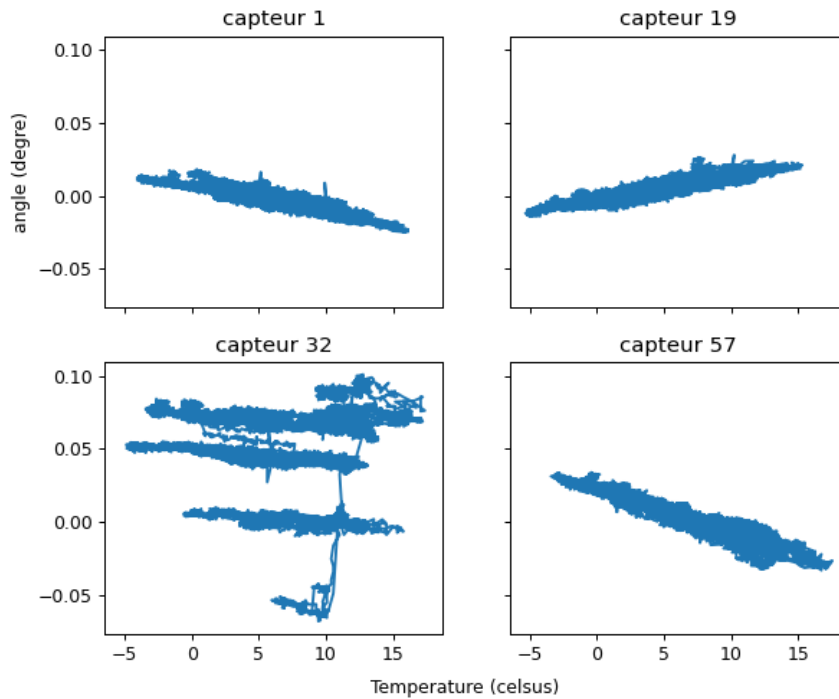


FIGURE 6: Corrélation mesures angulaires et de températures des capteurs 1,19 et 32 et 57

L'analyse précise de ce phénomène et de ses causes n'a pas été réalisée dans ce rapport, nous fournissons ici seulement quelques pistes de réflexions. Nous identifions les causes suivantes qu'il conviendrait de quantifier :

- Le capteur est sensible à la température et dispose d'un biais qui dérive (phénomène classique pour les capteurs inertiels de type MEMS).
- Le support du capteur a une inclinaison qui varie à cause de la température, par exemple à cause de la dilatation des pièces de fixation aux traverses.
- Le profil altimétrique de la voie varie en raison de la composition du ballast qui réagit de manière différenciée à la température ou de la dilatation du rail<sup>1</sup>.

Nous nous restreindrons donc à éliminer un peu naïvement ce biais en nous basant sur les observations suivantes :

- les mesures d'angle semblent être linéairement affectées par la température,
- le coefficient directeur est propre à chaque capteur, en particulier il ne semble pas être corrélé directement à la pente du rail (par exemple les capteurs 1 et 19 ont des pentes de mêmes signes, mais leurs coefficients directeurs sont de signes opposés),
- La corrélation doit être mesurée en dehors des périodes d'affaissement, mais le coefficient directeur est le même avant et après affaissement (les pentes de corrélation dans les 4 clusters de mesures du capteur 32 sont quasiment identiques).

Ceci nous a conduit à utiliser l'algorithme de débiaisage suivant pour chacun des capteurs :

1. Découper les données en tranches de 24h (ou d'une durée assez grande devant celle des changements de température mais telle que l'on dispose de suffisamment d'échantillons hors travaux)
2. sur chaque tranche, calculer le coefficient directeur de la corrélation entre angles et température.
3. prendre la médiane de ces coefficients directeurs, fixer une température cible et débiaisier linéairement toutes les données à l'aide de ce coefficient. Le choix de la médiane permet d'éliminer les valeurs aberrantes dues aux journées de travaux, et ainsi de corriger la variation thermique indépendamment de l'affaissement de la voie.

La figure 7 illustre le résultat sur les capteurs 1 et 32, avec en bleu les données brutes et en orange les données corrigées linéairement en température. On utilise une température de référence de  $5^{\circ}C$  pour les corrections, qui correspond à la température moyenne, mais ce choix n'affecte pas les résultats.

Une fois le biais de la température éliminé, on peut s'attaquer au bruit haute fréquence de nos mesures. Il est possible de le moyenner en temps,

---

1. En analysant les corrélations (surtout à Farébersviller) entre les capteurs on observe parfois des capteurs consécutifs anti-corrélés, comme si une déformation en forme de vague se formait lorsque la température augmente.

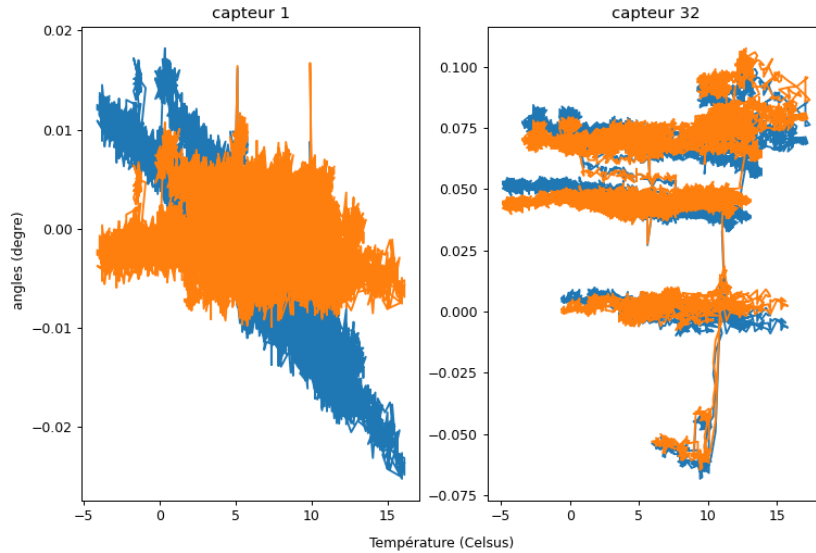


FIGURE 7: Corrélacion avec la température pour les capteurs 1 et 32.

mais nous avons préféré prendre la médiane des valeurs sur 24h pour éviter l'influence des données aberrantes.

En combinant la correction de température avec la sélection de la valeur médiane de chaque jour, on obtient le signal en bas à droite de la figure 8, qui semble beaucoup plus stable que le signal brut en haut à gauche. Dans la suite du rapport, on applique ces corrections à toutes les données ; il n'y a donc plus qu'une mesure d'inclinaison par jour et par capteur.

Le profil d'altitude est reconstruit dans la figure 9. En l'absence de données de contrôle, il est difficile de juger de la qualité du résultat. Nous nous contenterons donc de remarquer que :

- La première semaine est particulièrement stable (on observe une légère hausse d'altitude le 21/11/2019 vers le Point Kilométrique 101,740). C'est attendu car les travaux commencent une semaine après l'installation des capteurs.
- On observe un affaissement de près de 2cm le 26/11/2019 (un mardi) entre les PK 101,660 et 101,720 qui se stabilise.
- On observe une légère élévation de l'altitude autour du PK 101,725. Ce phénomène disparaît en fin de semaine.

Les résultats obtenus jusqu'à présent nous semblent très satisfaisants, cependant nous identifions deux faiblesses qui justifient la partie suivante. Premièrement notre méthode ne donne pas d'intervalles de confiance, et deuxièmement nous n'utilisons aucune corrélation spatiale pour notre modèle.

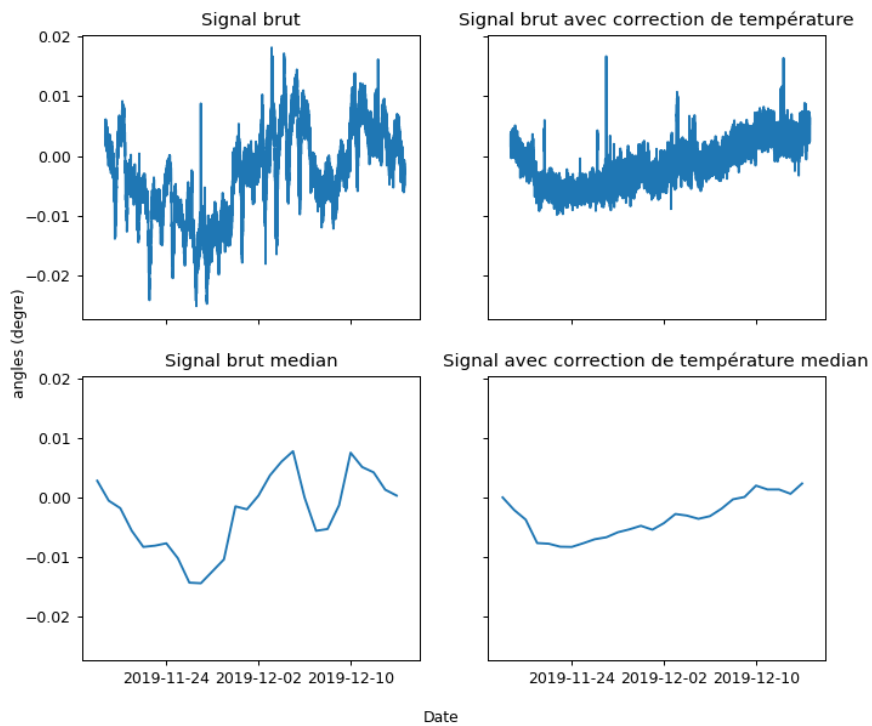


FIGURE 8: Signal corrigé pour le premier capteur. À partir du signal brut (en haut à gauche), on peut soit prendre directement les valeurs médianes journalières (en bas à gauche), soit corriger les biais dû à la température (en haut à droite) puis prendre les médianes (en bas à droite). On voit que cette dernière méthode donne les résultats les plus stables, ce qui correspond à l'hypothèse que l'inclinaison du capteur 1 est fixe car hors de la zone de travaux.

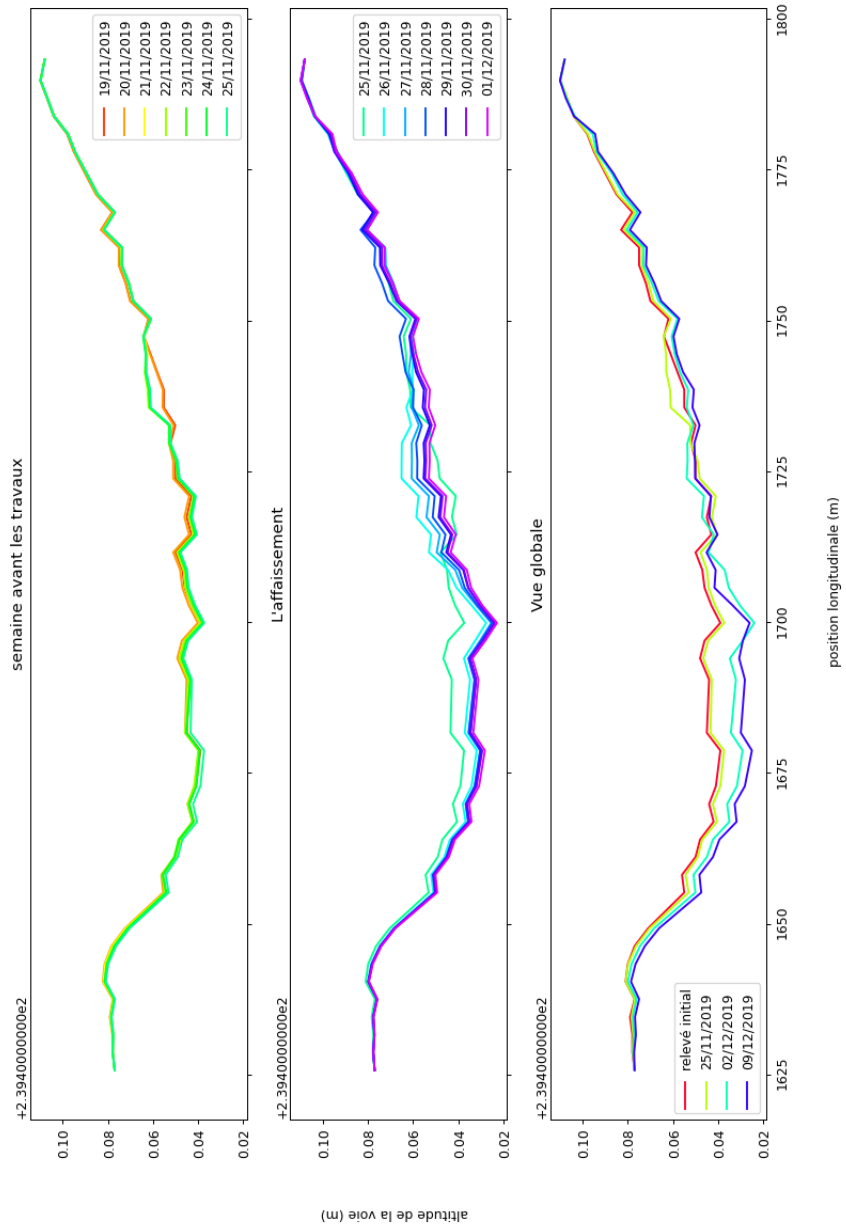


FIGURE 9: Profil d'altitude de Mittersheim reconstruit avec la méthode "convexe" avec données corrigées en température et lissées en temps



## 4 Un modèle d'inférence plus complexe

### 4.1 Rail non parfait

Malgré le lissage temporel précédent, on observe que les résultats ne sont toujours pas très réguliers en fonction de la coordonnée spatiale  $x$ , avec des flèches verticales allant jusqu'à 1cm entre un capteur et ses voisins, cf Figure 3. Puisque les capteurs sont placés tous les 3 mètres et que les rails font une douzaine de mètres de long, en supposant que les rails sont rigides, on s'attend à ce que les inclinaisons soient corrélées par groupes d'environ 4 capteurs consécutifs. Les positions des éclisses ne sont pas connues, donc on essaye de retrouver ces groupes uniquement à partir des données d'inclinaison.

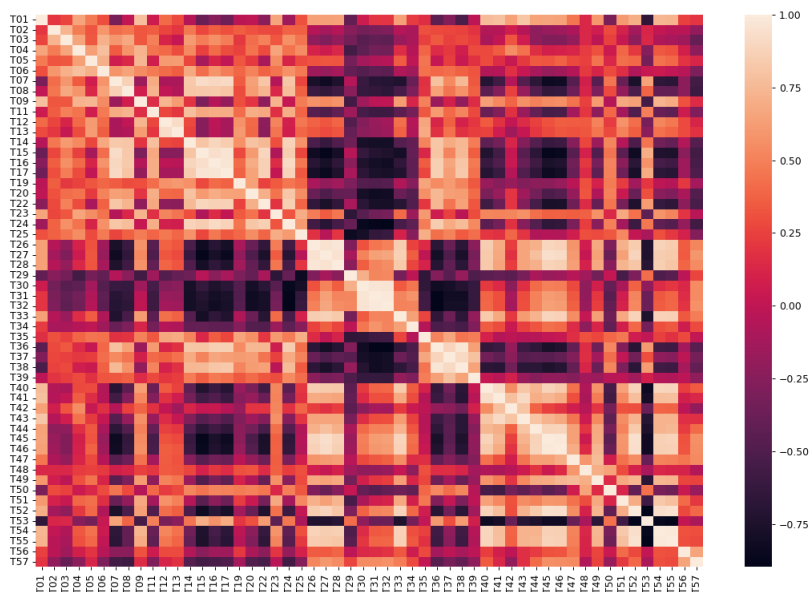


FIGURE 10: Corrélations entre les capteurs, après lissage temporel

Sur la figure 10, on affiche la matrice de corrélation entre les capteurs. Bien que certains groupes fortement corrélés se distinguent (T14 à T17, T26 à T28, T30 à T32), ce n'est pas le cas partout. Il y a même des capteurs corrélés négativement à leurs voisins, ce qui n'est pas compatible avec un modèle de rails rigides<sup>2</sup>.

Face à ces résultats peu concluants, nous proposons un modèle continu, où la voie a une rigidité uniforme sur la portion considérée.

2. Ce phénomène est très commun dans les données de Farébersviller, cela ne semble donc pas être simplement un capteur monté à l'envers

En notant  $f(x)$  l'altitude des rails après  $t$  relevés, la rigidité de la voie se traduit par l'absence de variations trop brutales de la pente  $f'$ , c'est-à-dire par une pénalisation de  $f''$ . Pour estimer la rigidité d'une portion de rail, nous allons utiliser les données de flèche verticale provenant des contrôles effectués à Farébersviller. On observe que la flèche verticale entre trois points espacés de 10m est de l'ordre de 10mm, ce qui donne des valeurs de  $f''$  de l'ordre de  $10\text{mm}/(10\text{m})^2 = 10^{-4}\text{m}^{-1}$ , soit  $\int_I (f'')^2$  de l'ordre de  $\frac{1}{r} := 10^{-8}\text{m}^{-1}$  sur tout intervalle  $I$  d'un mètre de long.

## 4.2 Modélisation statistique

Nous voulons rajouter de l'information à notre modèle en utilisant l'a priori de régularité justifié précédemment. Nous modélisons donc l'altitude par une fonction

$$f : x \in [x_1, x_N] \mapsto z = f(x) \in \mathbb{R}$$

sur laquelle nous imposons une loi a priori<sup>3</sup> de la forme

$$\pi(f) \propto e^{-\|f''\|^2/r}$$

où  $r$  est notre constante de rigidité et  $\|g\|$  est la norme  $L^2$  renormalisée :

$$\|g\|^2 = \frac{1}{x_N - x_1} \int_{x_1}^{x_N} g^2$$

On note

$$H^2 = \{f : [x_1, x_N] \mapsto \mathbb{R}, f \text{ est deux fois dérivable et } f, f', f'' \in L^2\}.$$

Il est intéressant de remarquer le lemme suivant

**Lemma 4.1.** *Si on impose  $f(x_i) = z_i$ ,  $f(x_{i+1}) = z_{i+1}$ ,  $f'(x_i) = \theta_i$  et  $f'(x_{i+1}) = \theta_{i+1}$ , le minimiseur  $f$  de  $\int_{x_0}^{x_1} (f'')^2$  dans  $H^2([x_1, x_{i+1}])$  est un polynôme de degré au plus 3.*

*Démonstration.* Par souci de complétude, nous justifions mathématiquement le caractère bien posé de notre problème :

L'espace affine  $\{f \in H^2([x_i, x_{i+1}]) \mid f(x_i) = z_i, f(x_{i+1}) = z_{i+1}, f'(x_i) = \theta_i, f'(x_{i+1}) = \theta_{i+1}\}$  est bien défini par injection continue de  $H^2$  dans  $C^1$ . On prend une suite minimisante et on peut en extraire une sous-suite faiblement convergente. La norme étant faiblement continue inférieurement et notre espace convexe, on obtient l'existence d'un minimiseur.

---

3. Ceci est uniquement formel puisque qu'il n'y a pas de mesure de Lebesgue sur  $H^2([x_i, x_f])$ .

Un minimiseur  $f$  satisfait alors une équation d'Euler-Lagrange : pour tout  $h$  tel que  $h(x_i) = h(x_{i+1}) = h'(x_i) = h'(x_{i+1}) = 0$ ,

$$\int_{x_i}^{x_{i+1}} f'' h'' = \int_{x_i}^{x_{i+1}} f^{(4)} h = 0,$$

d'où  $f^{(4)} = 0$ , i.e.  $f$  est un polynôme de degré 3 sur  $[x_i, x_{i+1}]$ .

Entre  $x_i$  et  $x_{i+1}$ , on sait que  $f$  est de la forme

$$f(x) = a + b(x - x_i) + c(x - x_i)^2 + d(x - x_i)^3,$$

et, en notant  $l = x_{i+1} - x_i$ , on obtient  $a = z_i$ ,  $b = \theta_i$ ,  $a + bl + cl^2 + dl^3 = z_{i+1}$  et  $b + 2cl + 3dl^2 = \theta_{i+1}$ . On peut écrire ensuite

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} l^2 & l^3 \\ 2l & 3l^2 \end{pmatrix}^{-1} \begin{pmatrix} z_{i+1} - z_i - \theta_i l \\ \theta_{i+1} - \theta_i \end{pmatrix} = \begin{pmatrix} 3l^{-2} & -l^{-3} \\ -2l^{-1} & l^{-2} \end{pmatrix} \begin{pmatrix} z_{i+1} - z_i - \theta_i l \\ \theta_{i+1} - \theta_i \end{pmatrix}.$$

soit

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 3l^{-2}(z_{i+1} - z_i) + l^{-3}(2\theta_i - \theta_{i+1}) \\ -2l^{-1}(z_{i+1} - z_i) + l^{-2}(\theta_i + \theta_{i+1}) \end{pmatrix}$$

Il est alors possible de calculer à partir de ces coefficients le coût de rigidité sur le segment  $[x_i, x_{i+1}]$  :

$$\int_{x_i}^{x_{i+1}} f''^2 = \int_0^l (2c + 6dx)^2 = 4lc^2 + 12l^2cd + 12l^3d^2.$$

□

Ceci nous amène à échanger notre variable continue  $f$  contre les variables  $z_i = f(x_i)$  et  $\theta_i = f'(x_i)$ . Soit  $X$  ce vecteur de dimension  $2N$ . Formellement

$$\pi(X) \propto e^{-X^T C X}$$

où  $C$  est la matrice obtenue par le lemme 4.1 en calculant le coût de chaque segment  $[x_i, x_{i+1}]$ . On appellera cette matrice la matrice de coût de rigidité. On se convainc facilement que  $C$  n'est pas de rang plein (nous avons fixé seulement  $2(N-1)$  degrés de libertés). On rajoute comme hypothèse que  $z_1$  et  $z_N$  suivent des lois gaussiennes centrées sur les valeurs initiales  $(z_1^0, z_N^0)$ . Nous avons choisi pour les applications numériques de mettre un écart type de  $\sigma_z = 1 \cdot 10^{-5}$  m sur ces deux variables aléatoires, ce qui revient implicitement à les fixer.

Nous avons alors une loi a priori bien définie :

$$\pi(X) \propto \exp\left(-\frac{1}{2}X^T C X - \frac{1}{2\sigma_z^2}((z_1 - z_1^0)^2 + (z_N - z_N^0)^2)\right)$$

Sachant  $X$ , nous modélisons alors les mesures d'angles  $M = (m_i)$  par des lois gaussiennes indépendantes de loi :

$$\pi(m_i|X) \sim \mathcal{N}(\theta_i, \sigma^2).$$

On prend  $\sigma = 5 \cdot 10^{-3}$  degré, estimé à l'aide des données sur les capteurs, comme dans la partie 3.

En appliquant la formule de Bayes, on obtient la loi a posteriori de  $X$  sachant  $M$  :

$$\pi(M|X) \propto \exp\left(-\frac{1}{2}X^T C X - \frac{1}{2\sigma^2} \sum_{i \in \{1, N\}} (z_i - z_i^0)^2 - \frac{1}{2\sigma} \sum_{i=1}^N (\theta_i - m_i)^2\right)$$

Donc  $X$  conditionnée à  $M$  est une gaussienne, de loi

$$\mathcal{N}(\Sigma_X P^T \Sigma_\theta^{-1} M, \Sigma_X)$$

avec  $P = (0_N, I_N)^T$  la matrice  $N \times 2N$  qui à  $X$  associe les angles  $(\theta_1, \dots, \theta_N)$ ,  $\Sigma_\theta = \sigma^2 I_N$  la covariance des mesures d'angle, et  $\Sigma_X = (P \Sigma_\theta^{-1} P^T + C)^{-1}$ . En particulier, nous disposons désormais de la matrice de covariance de nos estimateurs grâce à la matrice  $\Sigma_x$ .

Les résultats numériques sont présentés sur la figure 11.

### 4.3 Estimation des erreurs

Par ailleurs, ce modèle permet d'obtenir des intervalles de confiance. Dans notre modèle statistique, les altitudes et les angles sont représentés par des gaussiennes, dont on peut calculer la matrice de covariance. Cette matrice de covariance nous donne alors des intervalles de confiances pour l'altitude. Sur les figures sont représentés les intervalles de confiance à 90%, ce qui correspond à 1.6 fois l'écart-type autour de la moyenne. Pour ce niveau de confiance, on a une incertitude de l'ordre de  $10^{-2}m$ .

Cependant notre approche donne aussi accès aux corrélations de  $X$ , la matrice de covariance est donné par la figure 12. Nous pouvons donc nous intéresser à des quantités fonctions des altitudes et simuler (ou calculer) leur loi. Par exemple l'affaissement maximal défini comme l'écart maximal entre le profil initial et le profil au temps  $t$  (qui n'est pas gaussien) peut aisément se simuler par Monte-Carlo. Ainsi l'affaissement maximal le 12 décembre 2019 a une espérance de 2.3 cm pour un écart-type de 8mm.

### 4.4 Détermination des angles initiaux

Afin de mener à bien l'étude précédente, il faut aussi déterminer les angles à l'instant initial à partir de l'altitude. La méthode est similaire à ce qu'on a fait pour déterminer les différents niveaux d'altitude à partir des mesures d'angles.

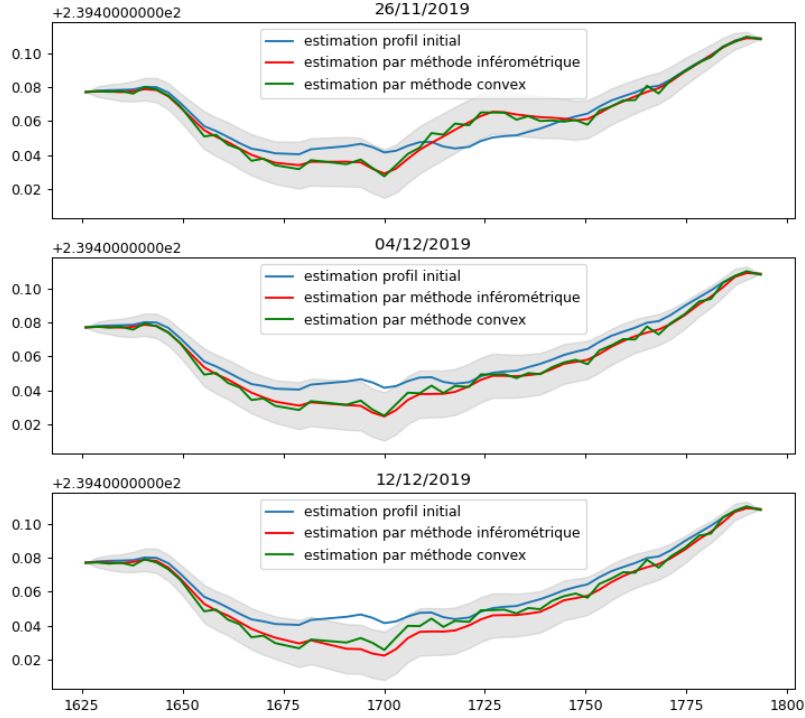


FIGURE 11: Résultat de notre méthode inférométrique pour 3 temps et comparaison avec le profil initial ainsi que la méthode développée dans la section précédente. En gris l'intervalle de confiance à 90%.

On dispose d'une mesure des altitudes ( $a_i$ ) des capteurs à l'instant initial, soumis à un bruit gaussien. On rappelle que  $C$  est la matrice de coût de rigidité, qui nous donne la norme  $L^2$  de la dérivé seconde d'un polynôme de degré au plus 4 en fonction de toute ses valeurs et dérivées au points ( $x_i$ ).

En utilisant un modèle statistique similaire à celui utilisé dans les angles dans la sous-section 4.2, on obtient une loi a posteriori sur le vecteur  $X = (z_1, \dots, z_N, \theta_1, \dots, \theta_N)$  des altitudes et des angles :

$$\pi(X|a_i) \propto \exp\left(-\frac{1}{2}X^t C X - \frac{1}{2\sigma_a^2} \sum_{i=1}^N (z_i - a_i)^2\right),$$

où  $\sigma_a$  est l'amplitude du bruit sur les mesures. On estime ce bruit de l'ordre de  $\sigma_a = 2 \cdot 10^{-3}m$ . En plus de toutes les valeurs d'angles, on obtient aussi le profil initial figure 13.

De même que pour les erreurs sur l'altitude de la partie précédente, on

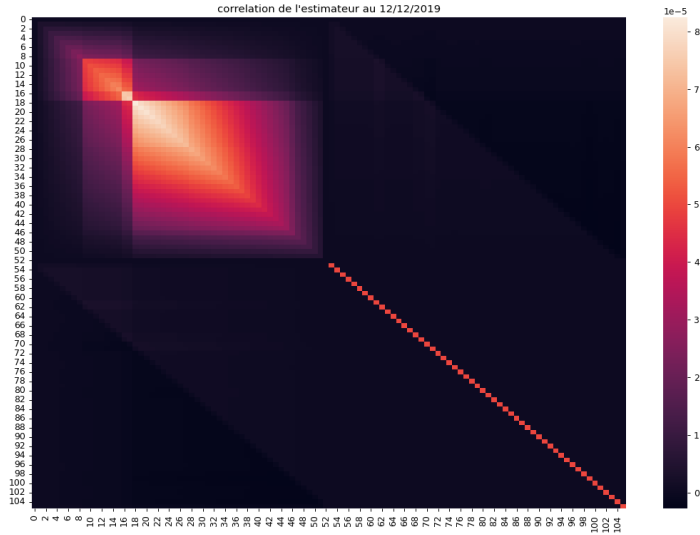


FIGURE 12: Matrice de corrélation. La matrice représente d'abord les altitudes puis les angles

obtient des intervalles de confiance pour la mesure de ces angles. L'intervalle de confiance à 90% correspond à une amplitude sur l'erreur de  $10^{-2}$  degré (en gris sur la figure 13).

Dans la suite, pour l'étude des données de Farébersviller, on utilisera toujours cette méthode d'analyse des angles initiaux, ainsi que la méthode présentée précédemment pour l'étude des profils successif de la voie.

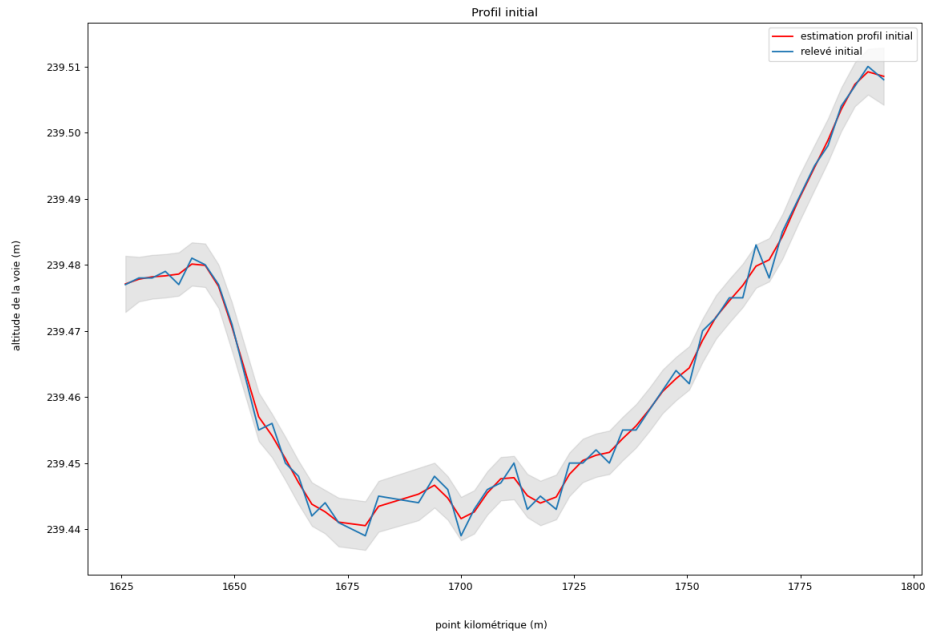


FIGURE 13: Profil initial de la voie à Mittersheim

## 5 Etude du cas de Farébersviller

### 5.1 Introduction et qualité des données

On étudie à présent le site de Farébersviller, où ont été installés des inclinomètres selon le même format, mais pour lequel on dispose de relevés d'altitude plus réguliers, effectués tous les mois. On applique les mêmes modèles que précédemment, en utilisant le relevé du 3 mars 2021 comme donnée initiale. Les relevés du 6 avril, du 10 mai et du 3 juin peuvent alors servir de données de contrôle, pour valider les résultats des modèles. Les mesures d'altitudes ne sont pas au niveau des capteurs (tous les 3m) mais tous les 10m, on interpole donc les altitudes.

L'analyse de la corrélation entre la température et les mesures (figure 14) montre une dispersion beaucoup plus importante. Alors que l'on suppose les extrémités de la voie stable, on observe ici des variations d'angles de l'ordre de 0.1 degré. Ceci est un ordre de grandeur de plus que ce que l'on observait dans les Figures 6 et 7 pour Mittersheim. Sans surprise les coefficients de corrélation sont beaucoup plus proches de 0.

capteur	coefficient de corrélation
V1I01	-0.491
V1I75	-0.046
V2I01	0.561
V2I75	-0.625

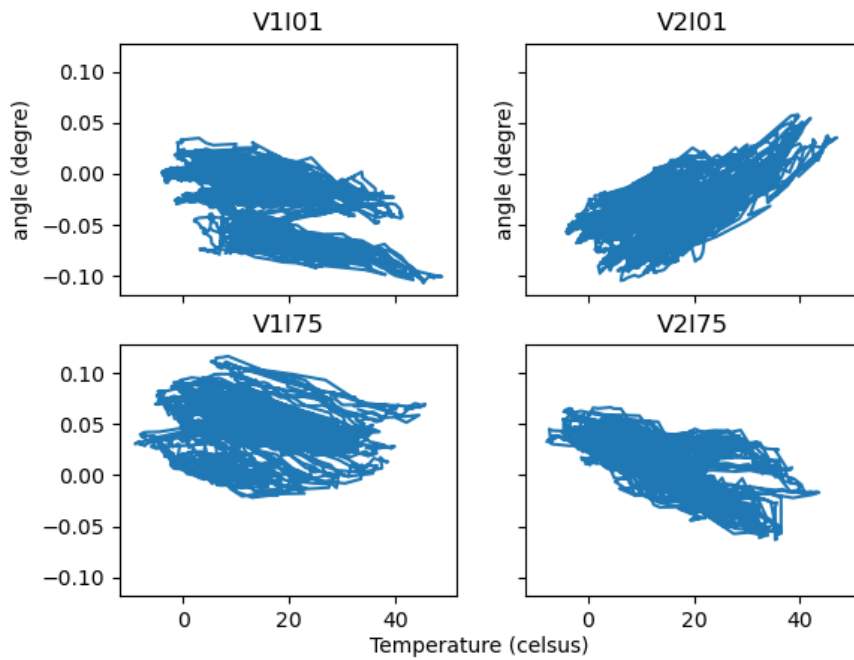


FIGURE 14: Corrélation des mesures angulaires et des températures pour le premier et le dernier capteur de chaque voie

Ainsi même après notre algorithme développé dans la partie 3, nous observons des variations d'amplitude de 0.1 degré par exemple sur le dernier capteur de la voie 2 (cf Figure 15).

La corrélation entre température et mesures est même clairement mise en défaut par des phénomènes d'hystérésis. Sur la figure 16, on distingue nettement des cycles dans le sens horaire, à l'échelle de la journée.

Par ailleurs, nous disposons de données de contrôles sur ce segment, nous pouvons donc estimer les variations d'angles entre les segments lors des relevés et les comparer avec les variations d'angles mesurées par les capteurs. Le résultat est présenté dans la figure 17. Il est à noter que l'on dispose d'un sous-échantillon de mesures pour reconstruire les angles (un relevé d'altitude tous les 10 m pour un capteur d'inclinaison tous les 3m) et qu'il n'est donc pas surprenant de ne pas déduire correctement les variations d'inclinaisons



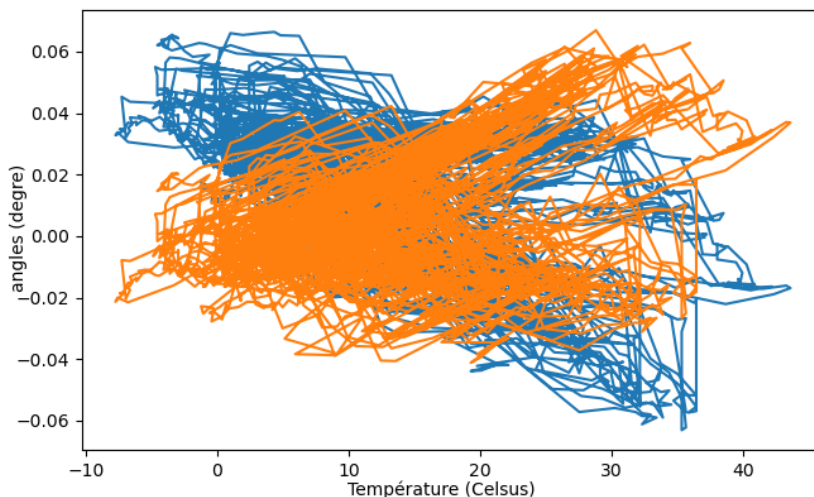


FIGURE 15: Les mesures du dernier capteur de la voie 2 (V2I75) brut (bleu) ou après débiaisage en température (orange).

de chaque capteur mais seulement les variations moyennes. Le point problématique concerne la voie 2 entre les capteurs 1 et 20 où il y a une dérive importante de l'estimation d'angle par les capteurs.

## 5.2 Reconstruction de l'affaissement

Nous présentons dans la figure 18 les résultats obtenus par notre méthode en initialisant avec les altitudes du relevé du 3 mars pour la voie 1. Les résultats sont très satisfaisant pour les 2 premières dates. Cependant l'estimation de l'affaissement est sous estimée pour le 03 juin. L'affaissement de la voie au niveau du premier capteur (PK 101,601) n'est cependant pas en notre faveur. Le profil est tout de même assez bien reconstruit, y compris dans ce cas, et il est simplement légèrement décalé vers le haut.

La figure 19 présente les données pour la voie 2. Malheureusement les résultats ne sont pas satisfaisants mais cela était attendu à cause des erreurs de mesures observées dans la figure 17. On y avait observé un écart type de  $0.1^\circ$  entre les angles estimés par les capteurs et ceux provenant des données altimétriques entre le premier et le dernier relevé. En prenant en compte cet écart-type, on retrouve bien les données contrôles dans l'intervalle de confiance et on peut tout de même affirmer qu'il y a bien eu un affaissement.

On peut aussi appliquer une correction affine en espace aux données capteurs pour forcer le premier capteur à toujours indiquer un angle stable. Les résultats obtenus pour la voie 2 sont alors nettement plus satisfaisants (Figure 20). Cette correction a pour objectif de corriger les biais des capteurs

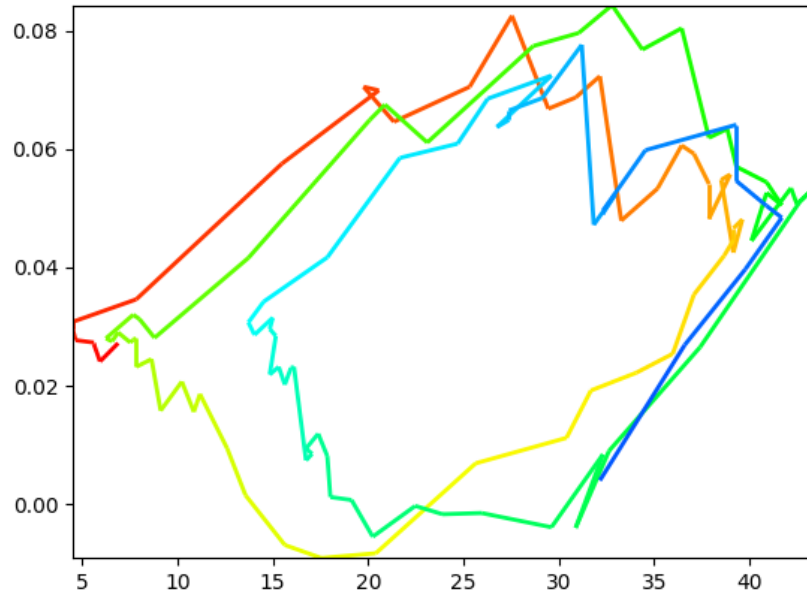


FIGURE 16: Corrélation entre la température (en abscisse) et les angles mesurés (en ordonnée) entre le 6 et le 8 mai 2021 par le capteur 38 de la voie 2. L'évolution des couleurs est fonction du temps, le rouge correspond au 6 mai au matin et le bleu au 8 mai au soir.

qui semble dériver au cours du temps. Cette méthode détériore cependant les résultats de la voie 1 et ne semble donc pas très robuste.

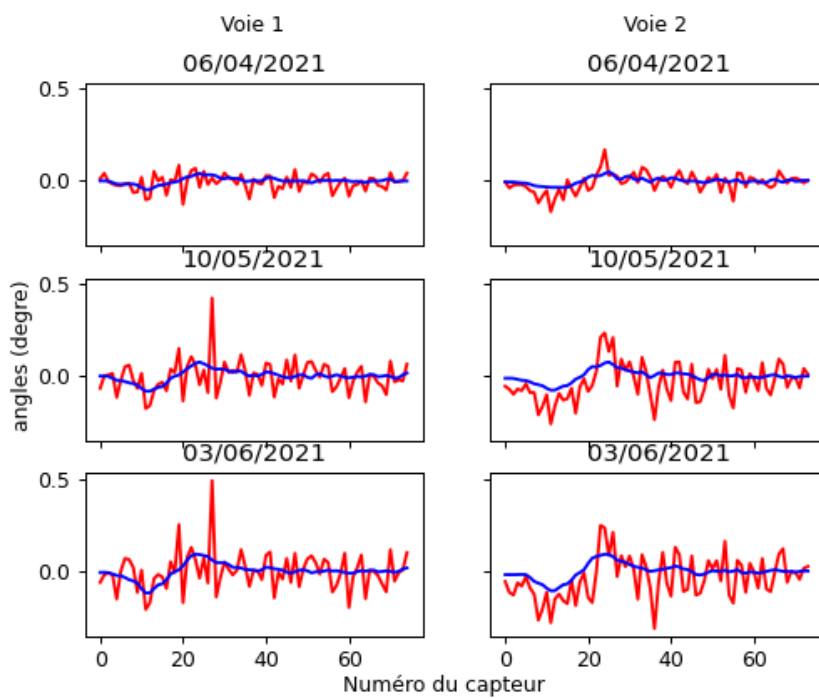


FIGURE 17: comparaison des variations d'angles entre les contrôles par rapport aux angles initiaux. En rouge les variations d'angles mesurés par les capteurs inclinométriques après correction en température (mais le résultat est très similaire sans cette correction) et en bleu les variations d'angles déduit des variations de profil altimétrique.

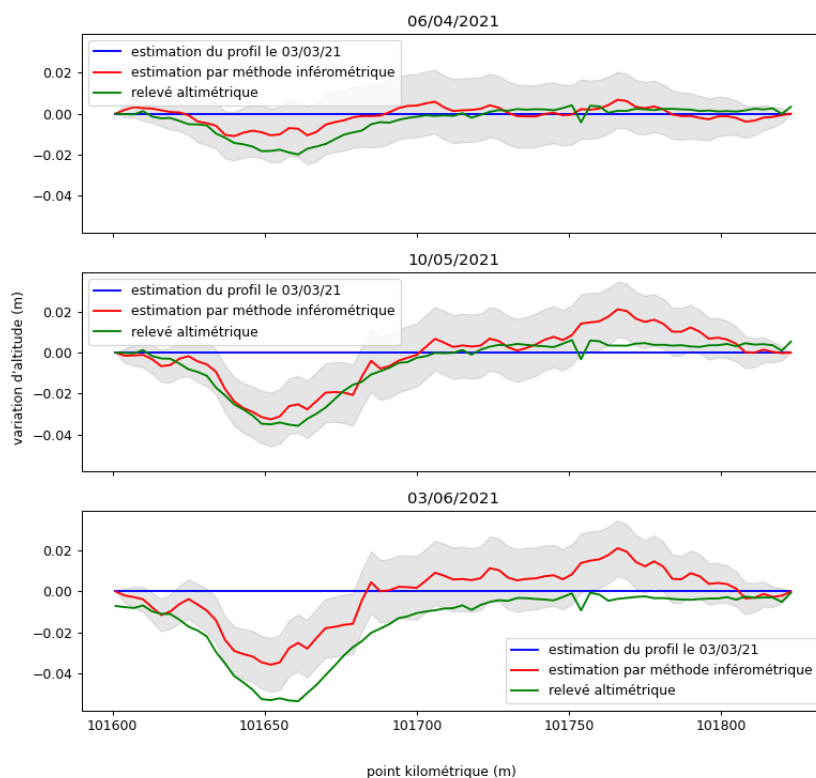


FIGURE 18: Estimation des variations d'altitudes à partir des données inclinométriques des capteurs et du relevé du 3 mars, et comparaison avec les données contrôles pour la Voie 1 de Farébersviller. En gris le seuil de confiance de 90% en supposant un écart type sur les mesures inclinométrique de  $0.005^\circ$ .

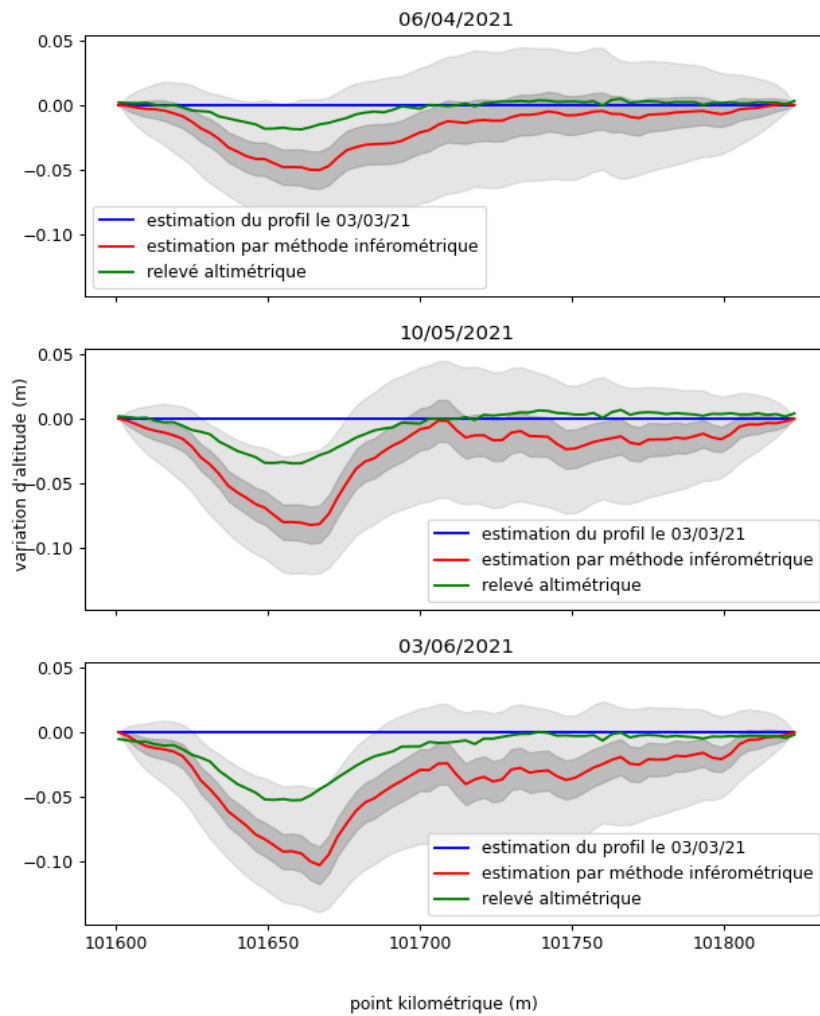


FIGURE 19: Estimation des variations d'altitudes à partir des données inclinométriques des capteurs et du relevé du 3 mars, et comparaison avec les données contrôles pour la Voie 2 de Farébersviller. En gris clair le seuil de confiance de 90% en supposant un écart type sur les mesures inclinométrique de  $0.005^\circ$ . En gris foncé l'intervalle de confiance à 90% en supposant un écart type de  $0.1$  degré.

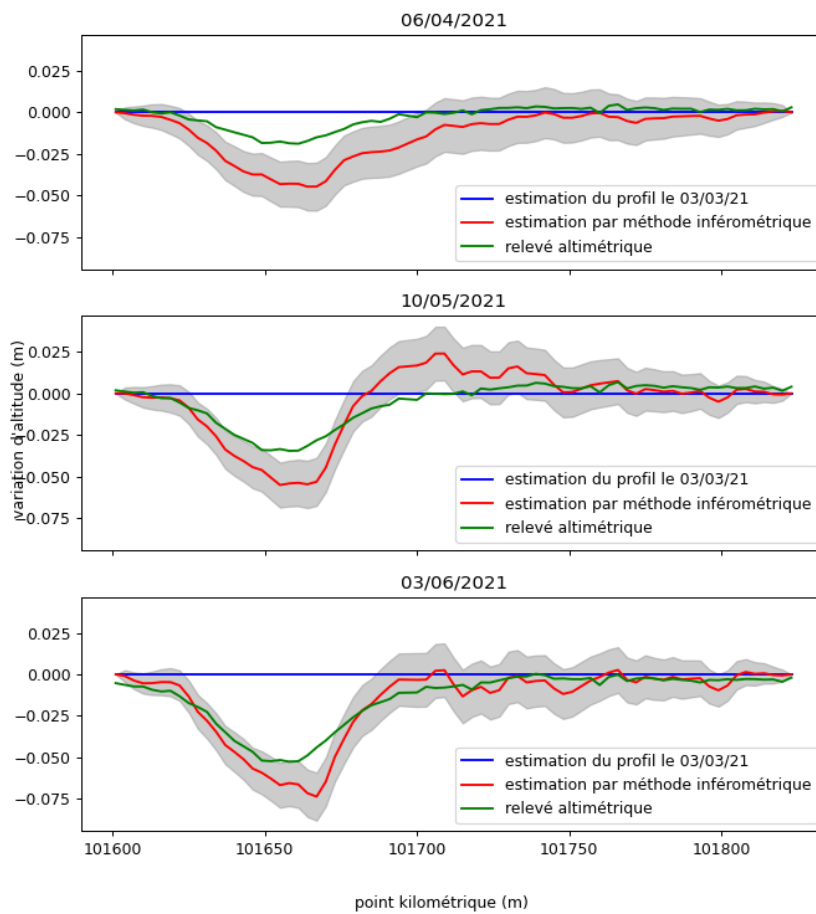


FIGURE 20: Même données que la figure 19 (donc voie 2) avec incertitudes fondées sur des erreurs d'écart-type 0.005 degré et une correction affine des angles pour forcer le premier à être nul.

## 6 Conclusion et améliorations envisageables

Nous avons développé une méthode de reconstruction de profil altimétrique fondée sur un modèle continu avec un a priori de rigidité. Ce modèle nous permet d'estimer les angles initiaux à partir d'un premier relevé altimétrique, et ensuite d'estimer le profil à l'aide des variations angulaires données par les mesures des capteurs inclinométriques. Ce modèle versatile présente l'avantage de fournir des bornes d'incertitudes réalistes et de ne pas nécessiter l'inférence de paramètres cachés.

La difficulté est alors de filtrer et corriger autant que possible les mesures capteurs. Un biais évident est la température, elle influence profondément les mesures. Sur Mittersheim, les variations de températures restaient assez faibles (entre -5 et 15 degrés Celsius) et une corrélation linéaire importante se dégageait. Cette corrélation linéaire couplée avec une régularisation temporelle a permis de faire des estimations qui semblent être très satisfaisantes. Les phénomènes thermiques sont cependant beaucoup plus importants à Farébersviller (entre -10 et 40 degrés Celsius), la corrélation est alors très complexe et présente même de l'hystérésis. En l'absence de modèle thermique satisfaisant, nous observons des erreurs importantes pour la voie 2 entre nos résultats et les relevés de contrôles.

Le développement d'un modèle plus robuste capable d'estimer sur plusieurs mois avec une large amplitude thermique les variations altimétriques nécessite donc une modélisation plus précise de l'influence de la température.

Une autre piste d'amélioration est la détection d'outliers (les données aberrantes). Notre lissage temporelle élimine déjà les points de mesures aberrants. Cependant si un capteur développe systématiquement des données non cohérentes (capteur monté à l'envers, détérioration importante, décrochage du capteur...), une intervention humaine est nécessaire. Notre modèle devrait permettre facilement de détecter ces éventuels capteurs donnant des mesures "peu probables" et faciliter leur détection.