

Rapport localisation – stratégie de sélection des ancres

25 mai 2020

I. Introduction

Il est apparu que le choix des ancres jouait un rôle très important dans la précision obtenue à l'issue de l'algorithme de localisation. Au-delà d'éventuels poids affectés à chacune des ancres, leur disposition est un critère important à prendre en compte avant de faire un choix.

À cause des contraintes de réalisation des expériences et d'espace ne permettant pas d'évaluer différentes stratégies, on décide de réaliser une simulation ce qui permet de réaliser de très nombreuses expériences en un temps restreint et donc de comparer la performance des différentes stratégies.

Une première série de simulations a été réalisée en créant un quadrillage d'ancres très réguliers avec un pas de 20 mètres. Cependant, dans une telle configuration, le meilleur choix était toujours de choisir les 4 ancres les plus proches, c'est à dire celles formant le carré 20 x 20 entourant le tag. Cette approche n'a pas permis de comparer objectivement les différentes stratégies de sélection.

Dans la suite, on décide de se baser sur des conditions plus proches de la réalité, où les ancres sont loin de suivre un quadrillage régulier et suivent plutôt les contraintes liées à l'environnement et aux murs.

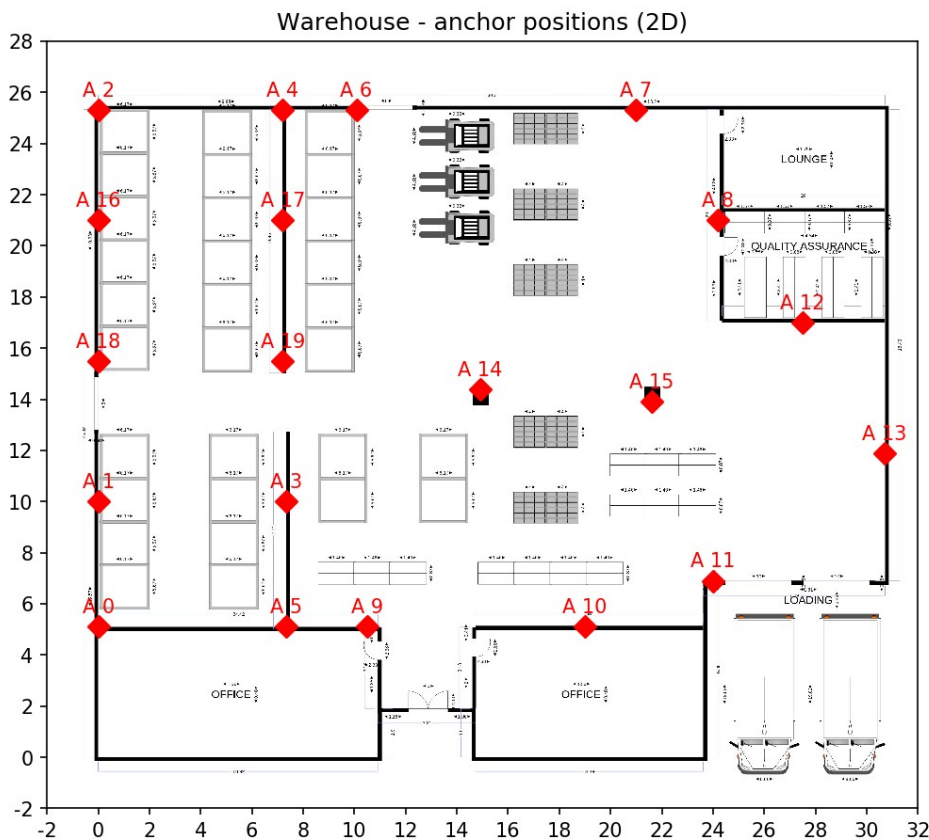


Figure 1. Carte de l'entrepôt et disposition des ancres utilisées dans les simulations.

On dispose donc 20 ancres numérotées de A0 à A19 sur la carte d'un entrepôt ci-dessus (basé sur un template de l'outil SmartDraw Cloud, distances en mètres). On essaie de les disposer assez

régulièrement sur les parois de l'entrepôt ou sur les deux piliers (A14 et A15) de sorte à couvrir aussi uniformément que possible l'ensemble de la surface. On pourrait discuter la densité de ces ancrs et leur disposition mais ce n'est pas l'objet premier de ce rapport.

Le nombre de paramètres à prendre en compte dans la simulation est gigantesque (à commencer par la disposition des ancrs). On décide donc fixer une partie :

1. La position des ancrs restera la même pour l'ensemble des simulations
2. L'erreur de *ranging* est modélisée par une variable aléatoire suivant une loi normale d'espérance μ et de variance σ^2 .
On choisit $\mu = 0.08 \text{ m} > 0$ pour tenir compte de l'effet *multipath* qui introduit un biais strictement positif sur les mesures. Dans le cas idéal, après calibration, on souhaiterait $\mu = 0$ mais cette hypothèse serait probablement trop optimiste.
On choisit $\sigma = 0.15 \text{ m}$ qui modélise l'erreur aléatoire sur les mesures de *ranging* (liée par exemple aux horloges).
3. On n'introduit pas de modélisation des phénomènes *line of sight* (LOS/NLOS). On a vu précédemment que la présence de gros obstacles (murs, etc.) peut fortement dégrader la qualité des mesures (erreur systématique et stabilité/variance).
4. La solution initiale de l'algorithme de localisation est choisie comme le barycentre des ancrs.

On se fixe deux objectifs :

1. Comparer différentes stratégies de sélection des ancrs
2. Évaluer l'intérêt de l'algorithme permettant d'améliorer la localisation sur l'axe vertical

Les différentes stratégies de sélection à évaluer sont :

1. Maximiser le déterminant de la matrice de covariance :
- des coordonnées XY des ancrs (*det_covariance_2D*)
- des coordonnées XYZ des ancrs (*det_covariance_3D*)
2. Maximiser la variance de la composante verticale des ancrs (*variance_z*)
3. Sélection aléatoire (*random*)
4. Sélection des ancrs les plus proches (*nearest*)

II. Description des algorithmes

On propose un algorithme permettant de réévaluer la composante verticale à l'issue de la minimisation de la fonction de coût. Voici la procédure :

1. On obtient une solution (x, y, z) à l'issue de l'algorithme de minimisation
2. On fixe les coordonnées x et y, et on calcule la coordonnée z comme la moyenne (pondérée) des z correspondant aux distances de *ranging* de chacune des N ancrs :

Pour $i = 1 \dots N$:

d_i = distance mesurée entre le tag et l'ancre i

On calcule la composante verticale égalisant la distance de *ranging* pour cette ancre :

$$\tilde{z}_i = z_i \pm \sqrt{d_i^2 - (x - x_i)^2 - (y - y_i)^2}$$

Si l'ancre est au-dessus du tag, il faut choisir la solution "-". Si elle est en dessous, il faut choisir la solution "+".

3. On calcule la coordonnée z finale comme la moyenne des \tilde{z}_i . On peut aussi utiliser une moyenne pondérée par les inverses des distances d_i comme indiqué ci-contre:

$$z = \sum \frac{\tilde{z}_i / d_i}{1 / d_i}$$

III. Résultats

1. Position initiale et correction sur la composante verticale

On va balayer un quadrillage en 3 dimensions sur le volume de l'entrepôt, dans la zone délimitée par les ancrs les plus extrêmes (A0 et A2 sur l'axe Y et A0 et A13 sur l'axe X). Pour chacune des positions, on réalise trois localisations.

La localisation s'effectue comme suit :

1. Faire un *pré-ranging* pour obtenir la liste de toutes les ancrs situées dans un rayon de 20 mètres.
2. Parmi ces ancrs, procéder à un tirage aléatoire de $N = 8$ ancrs selon une probabilité inversement proportionnelle à la distance.
- 3.

La première observation est que la correction de la composante Z permet une amélioration significative dans le sens où elle supprime complètement l'ambiguïté (qui apparaît fréquemment lorsque les ancrs sont placées à des altitudes similaires). Sans la correction, on observe deux pics dans les solutions situées de part et d'autre du plan moyen formé par les ancrs.

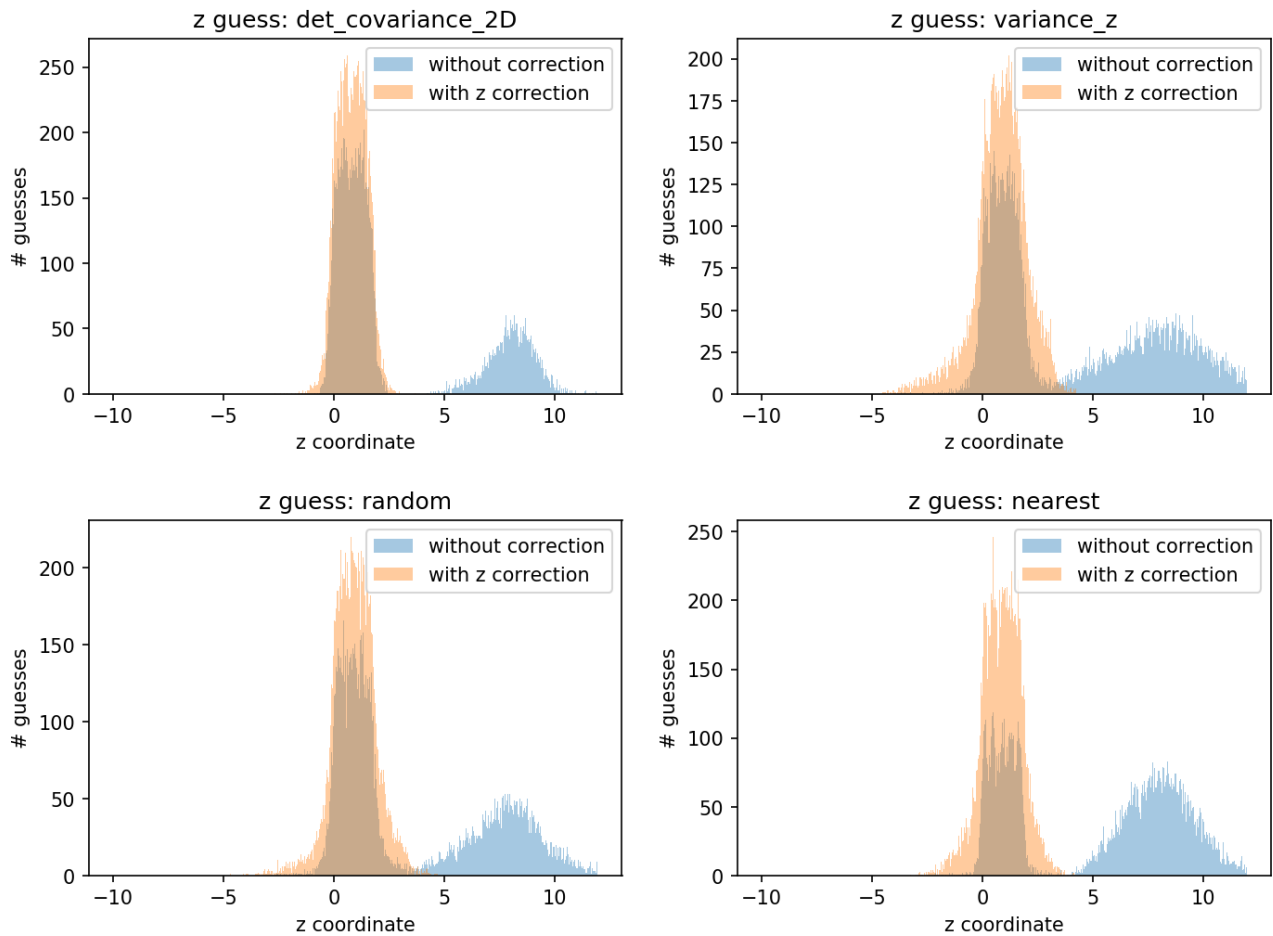


Figure 2. Histogramme des solutions sur l'axe z avec (en orange) et sans (en bleu) correction de la composante verticale selon l'algorithme décrit à la page précédente.

Cependant, cela n'est possible que dans le cas où l'on sait quel signe choisir dans l'équation. Si les ancrs sont toujours situées en hauteur par rapport au tag, alors on choisira toujours la solution avec le signe "-".

Cela implique de connaître *a priori* la position relative du tag et de l'ancre, ce qui n'est pas toujours possible, surtout dans un environnement où les ancrs ne seront pas toutes placées en hauteur par exemple. De plus, s'autoriser à mettre des ancrs à des hauteurs plus faibles permet d'augmenter la diversité des positions ce qui a généralement une influence positive sur la précision de la localisation.

Une contrainte plus facile à définir est une contrainte sur la zone de minimisation. On peut, selon le contexte, affirmer par exemple que le tag se trouvera toujours à une altitude comprise entre z_1 et z_2 m. Regardons l'effet d'une telle contrainte.

Dans la figure ci-dessous, tous paramètres égaux par ailleurs, on a reproduit la simulation en imposant $z < 4$ m pour la minimisation de la fonction de coût.

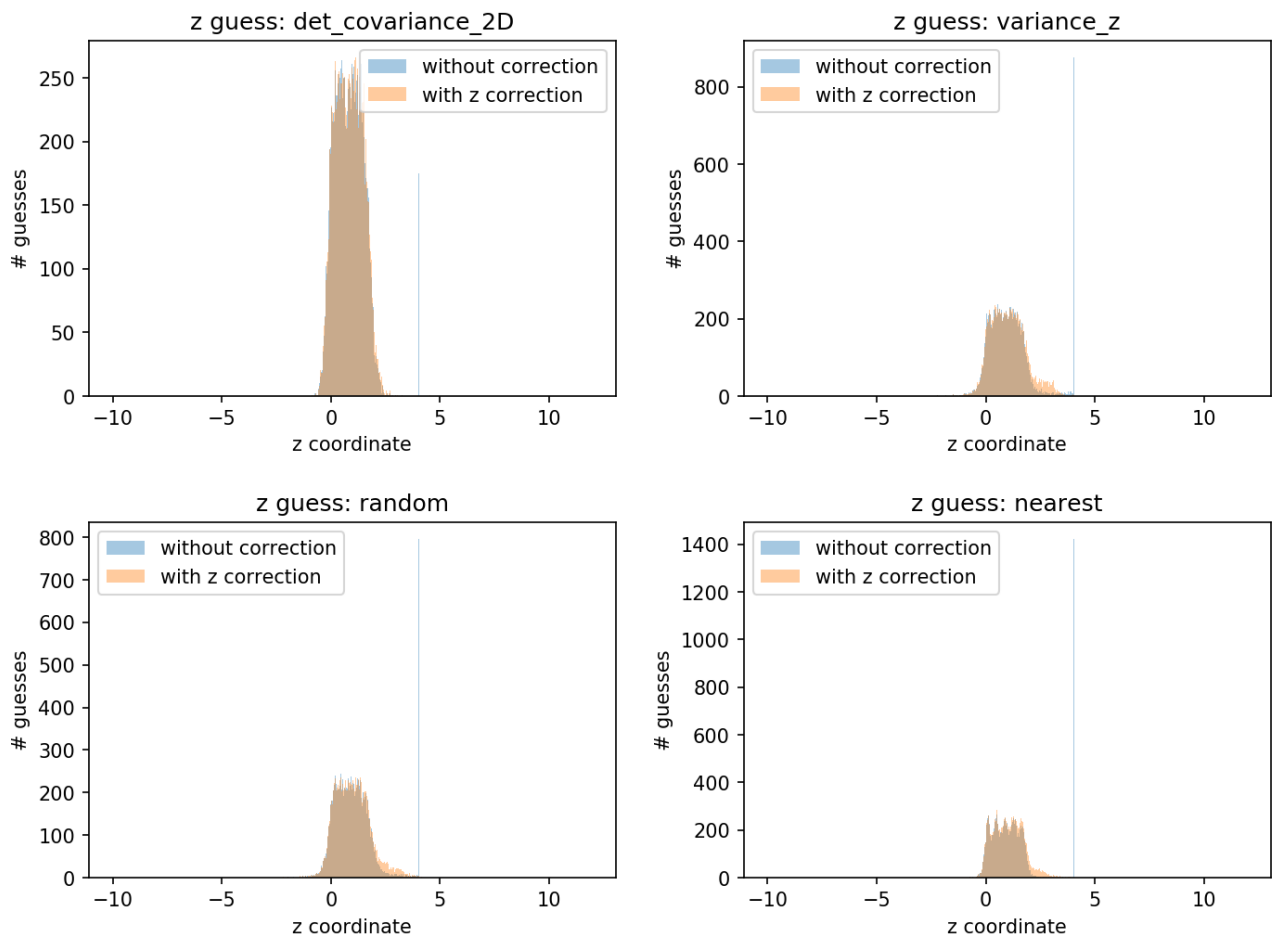


Figure 3. Histogramme des solutions sur z en imposant une contrainte ($z < 4$ m) sur l'axe vertical

On remarquera le pic très abrupt à $z = 4$ m. La fonction de coût n'étant pas convexe, l'algorithme de minimisation peut tomber dans un minimum local qui se trouve être sur la frontière. Il est alors nécessaire de bien choisir la solution initiale, choix qui peut dans certains cas s'avérer crucial pour la validité de la solution.

Pour comparaison, si au lieu de choisir le barycentre comme point de départ, on choisit un point proche de la vraie position du tag, on obtient les graphiques suivants :

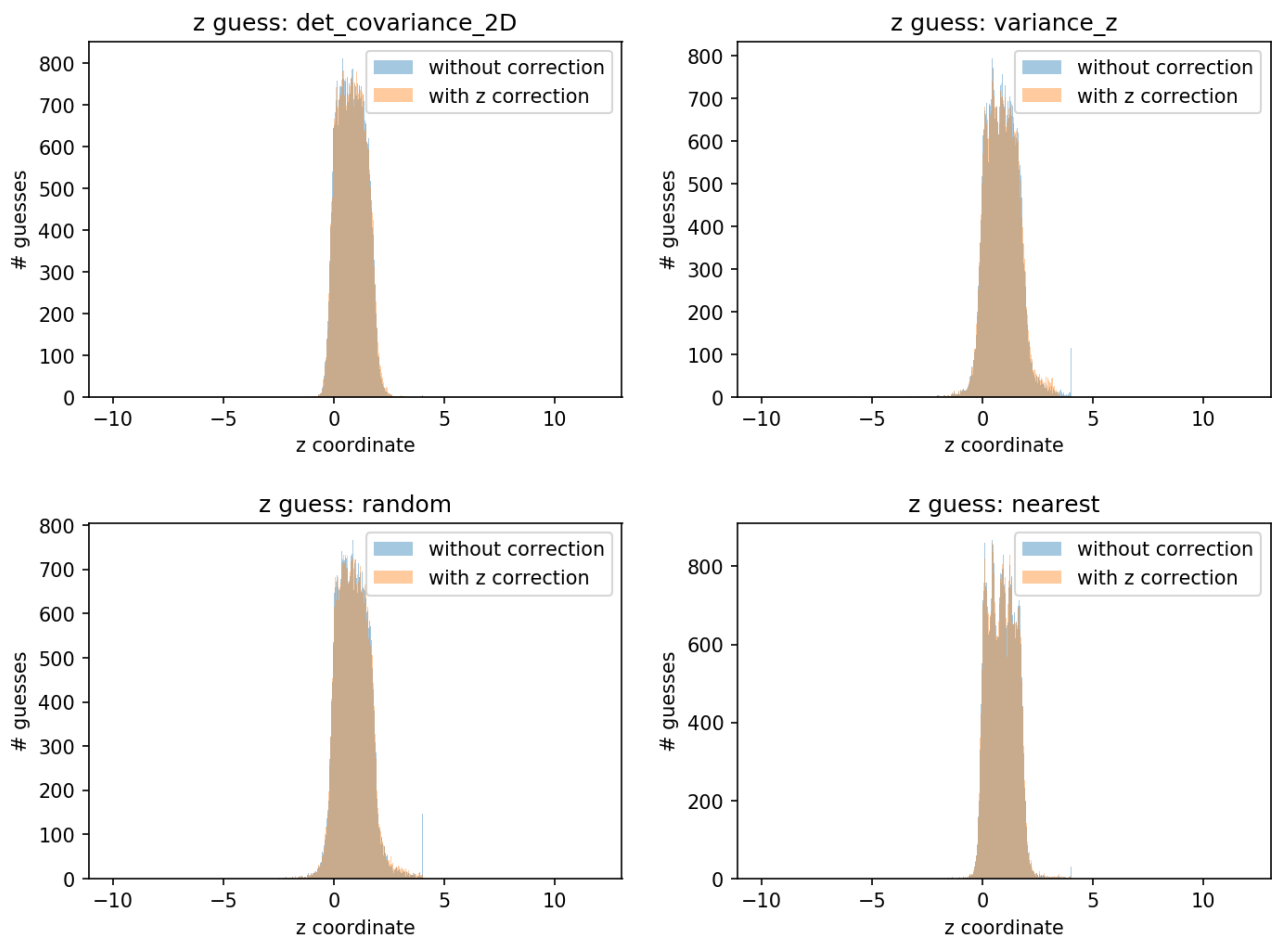


Figure 4. Histogramme des solutions sur z en imposant une contrainte ($z < 4$ m) sur l'axe vertical et en choisissant une position initiale proche (< 0.9 m de la vraie solution)

Ce qui améliore drastiquement les statistiques sur l'erreur de localisation (davantage sur l'erreur moyenne que sur l'erreur médiane, étant donné que la majorité des points sont quand même dans le bon demi-espace).

Pour finir, on retire la contrainte sur l'axe z , mais on choisit le point de départ déterminé par :

X, Y = moyenne des coordonnées X, Y des ancres

$Z = 0$

Comme on peut le voir sur la figure ci-dessous, on supprime complètement l'ambiguïté. Les résultats sont strictement équivalents avec et sans correction de la coordonnée sur l'axe vertical (comme le suggèrent les graphes ci-dessous, mais c'est également vérifié sur les statistiques de l'erreur : moyenne, médiane...).

Conclusion : le choix de la position initiale peut être crucial. Lorsque tous les appareils sont avec certitude au-dessus du plan $z = 0$, il semble judicieux de fixer $z = 0$ comme altitude initiale en entrée de l'algorithme de minimisation.

La correction sur z n'apporte alors pas d'amélioration.

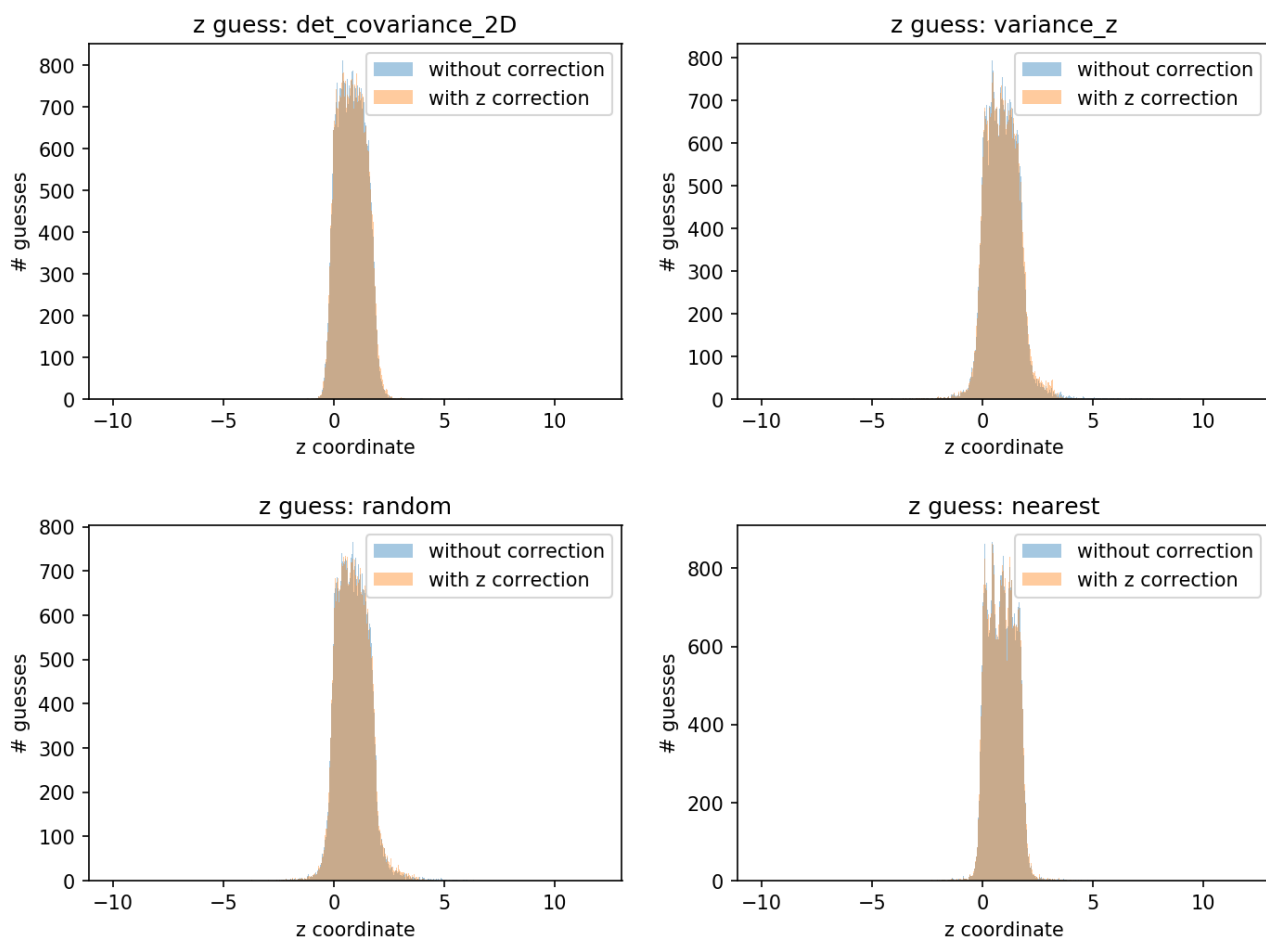


Figure 5. Histogramme des solutions sur z sans contrainte sur l'axe vertical mais en imposant $z = 0$ sur la position initiale de l'algorithme de minimisation.

2. Stratégie de sélection des ancres

On se place dans le dernier cas évoqué : pas de contrainte sur z , position initiale choisie comme le barycentre des coordonnées XY des ancres et $z = 0$.

On s'intéresse aux fonctions de répartition qui permettent de répondre à la question : « dans quel pourcentage des cas l'erreur est-elle inférieure à un certain seuil ? » (ou réciproquement).

On peut regarder à la fois l'erreur 2D et l'erreur 3D, pour chacune des différentes stratégies de sélection des ancres.

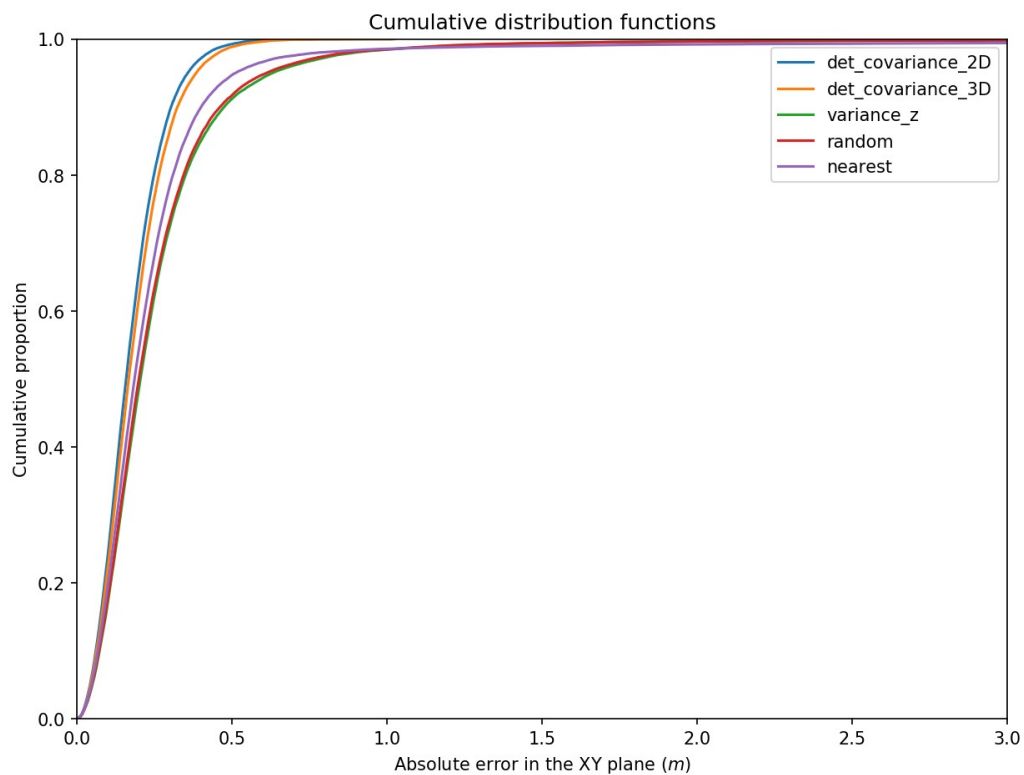


Figure 6. Fonctions de distribution cumulative pour les différentes stratégies de sélection des ancrs (erreur dans le plan XY).

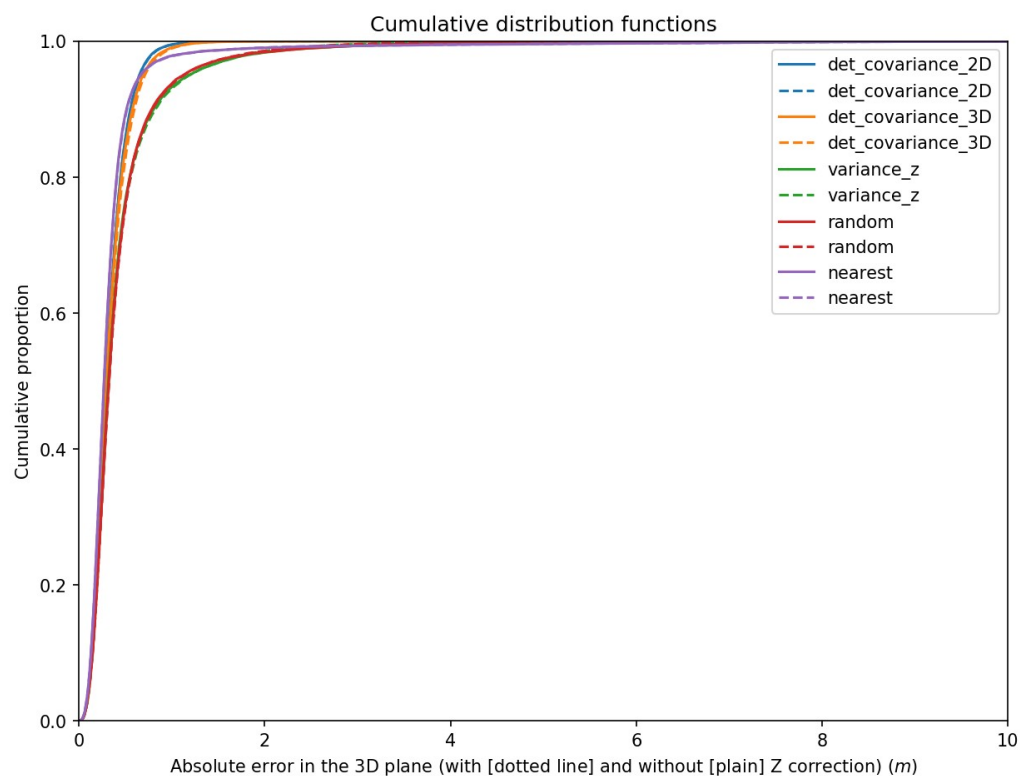


Figure 7. Fonctions de distribution cumulative pour les différentes stratégies de sélection des ancrs (erreur dans le plan 3D), avec et sans correction de l'altitude

Voici enfin les statistiques sur l'erreur pour les différentes stratégies :

Stratégie		det_cov_2D	det_cov_3D	variance_z	random	nearest
Erreur 3D	MSE* (m ²)	0.1397	0.152	0.4813	0.4384	0.4483
	Erreur médiane (m ²)	0.0870	0.0893	0.1031	0.1055	0.0751
	Max erreur ²	22.79	16.70	247	117.74	217.62
	Min erreur ²	5.88e-05	4.78e-5	9.07e-5	6.83e-5	9.21e-5
Erreur 3D + corr. z	MSE* (m ²)	0.1397	0.1580	0.4472	0.3988	0.4438
	Erreur médiane (m ²)	0.0875	0.0931	0.1068	0.1070	0.0757
	Max erreur ²	12.45	5.87	247.67	117.72	217.67
	Min erreur ²	7.57e-05	4.99e-5	9.20e-5	1.5580e-5	4.03e-5
Erreur plan XY	MSE* (m ²)	0.040	0.045	0.2041	0.1525	0.3518
	Erreur médiane (m ²)	0.0252	0.0279	0.0431	0.0407	0.0355
	Max erreur ²	1.03	1.05	239.40	115.86	216.35
	Min erreur ²	1.01e-6	1.35e-7	2.95e-7	9.34e-7	1.65e-6

MSE = Mean Squared Error = Erreur Quadratique Moyenne

On observe que les meilleurs résultats en termes d'erreur moyenne sont dominés par la stratégie de la matrice de covariance. Cela s'explique par le fait que **la pire des situations correspond à des ancres alignées** (par exemple, sélectionner 4 ancres sur un même mur). Or, ce sont les deux seules stratégies qui vont éliminer totalement ces configurations (le déterminant étant nul pour 4 ancres situées sur un même mur, la configuration ne serait choisie que s'il n'y en avait aucune autre possible).

En revanche, l'erreur médiane est un peu meilleure pour la stratégie des ancres les plus proches (parmi les 8 pré-sélectionnées). L'erreur moyenne des autres stratégies (variance_z, random et nearest) est beaucoup plus élevée car elle ne permettent pas d'éviter les configurations dégénérées. Dans le cas général, elles peuvent cependant être meilleures (cf. *nearest*).

La remarque précédente explique aussi les valeurs énormes pour le maximum de l'erreur quadratique pour certaines stratégies. Un regard sur les fonctions de distribution cumulatives permet cependant de relativiser quant à l'importance de ce phénomène.

IV. Conclusion et limitations

Conclusions

La procédure de sélection par le critère de maximisation du déterminant de la matrice de covariance donne les meilleurs résultats, ce qui s'explique par le fait qu'il assure de ne jamais sélectionner une configuration extrême qui produira une localisation aberrante.

La minimisation sous contrainte (borner les valeurs permises de l'altitude) permet de réduire les erreurs liées à des ambiguïtés sur l'altitude, mais n'est pas tout à fait satisfaisante car certaines solutions restent « bloquées » sur la frontière de la contrainte.

En revanche, un choix judicieux des conditions initiales, notamment sur l'altitude, permet de s'affranchir de ce problème d'ambiguïté et d'erreurs dramatiques sur la localisation. C'est donc sûrement cette option qu'il faut privilégier.

À l'avenir, surtout si l'on découvre que l'erreur d'autant plus élevée que la distance augmente, il pourra être utile d'envisager un autre critère mêlant les atouts du critère sur le déterminant de la matrice de covariance et des ancrages les plus proches, en s'assurant simplement d'éliminer les configurations dégénérées, sans pour autant privilégier des ancrages très éloignées par rapport à des ancrages plus proches. En effet, voici la distance tag-ancrage moyenne et médiane mesurée pour chacune des stratégies :

	Distance médiane (m)	Distance moyenne (m)
det_covariance_3D	13.00	13.01
det_covariance_2D	12.15	12.14
variance_z	10.63	10.65
random	11.02	11.03
nearest	7.13	7.36

Le classement est sans surprise : il faudra donc bien penser évaluer l'importance du problème susmentionné dans le futur.

Limitations

Le modèle de bruit utilisé ici pour les erreurs de *ranging* est très simpliste (bruit gaussien). Cependant, on ne dispose pas d'assez de mesures réelles pour envisager un autre modèle qui serait peut-être plus proche de la réalité.

Il est par exemple probable que la distance tag-ancrage ait une influence sur l'erreur de *ranging* (hors considérations de LOS). Ici, l'erreur relative est d'autant plus faible que la distance est grande, comme l'erreur est indépendante de la distance.

On n'a pas tenu compte des erreurs dues à l'obstruction du canal (NLOS).

Toutes les simulations ont été faites à position d'ancres fixées. Il est probable que d'autres configurations soient plus favorables et améliorent la précision générale de la localisation.