

Rapport de stage

Réalisé à l'IPHC

**Sujet : Conception d'algorithmes de reconstruction dans
le cadre de l'expérience ANTARES.**

Maître de stage : M. Arnauld ALBERT
Tutrice de stage : Mme Adèle CARRADO

Période du 04 avril au 19 juin 2009

Remerciements

Je remercie le Professeur Daniel HUSS, directeur de l'IPHC, pour m'avoir permis de réaliser ce stage dans ses locaux.

Je tiens également à remercier mon maître de stage, M. Arnaud ALBERT, enseignant-chercheur au GRPHE, pour m'avoir proposé ce sujet et pour m'avoir guidé tout au long de ces trois mois.

Enfin, je remercie ma tutrice de stage, Mme Adèle CARRADO, pour ses conseils pour la rédaction de ce rapport et pour la préparation de la soutenance.

Sommaire

INTRODUCTION.....	3
I. Présentation : IPHC-ANTARES	4
I.1. IPHC et GRPHE	4
I.2. Le projet ANTARES	5
I.2.1. Les neutrinos	5
I.2.2. Principe de détection.....	8
II. Le télescope à neutrinos ANTARES	10
II.1. Le détecteur.....	10
II.2. Contraintes posées par l'environnement.....	13
III. Algorithmes de reconstruction de traces.....	15
III.1. Étude des algorithmes.....	15
III.2. Données simulées.....	18
III.2.1. Cas idéal	19
III.2.2. Incidence du bruit de fond.....	19
III.3. Données réelles	25
CONCLUSION.....	29
SOURCES	30

INTRODUCTION

Dans le but d'observer des régions et sources éloignées de l'Univers, un détecteur de neutrinos a été mis au point par la collaboration ANTARES (Astronomy Neutrino Telescope with Abyss environment RESearch).

La collaboration ANTARES regroupe des instituts de recherche provenant de toute l'Europe : l'IPHC à Strasbourg, le GRPHE à Colmar, ou encore l'INFN en Italie et le NIKHEF à Amsterdam par exemple.

Ce détecteur a pour but de repérer les neutrinos traversant la Terre. En effet, à ces distances, les particules usuelles ne peuvent plus être utilisées pour des études astronomiques: les photons sont absorbés par les vestiges du Big Bang, les protons sont déviés par les champs magnétiques et les neutrons ont une durée de vie trop courte pour se propager jusqu'à la Terre. Ainsi, avec sa faible masse, sa faible interaction et sa charge neutre, le neutrino constitue le candidat idéal pour cette étude.

La détection de ces neutrinos est rendue possible par la mise en place du détecteur ANTARES en mer Méditerranée à 2500 m de profondeur.

L'objectif de mon stage est concentré sur la mise au point d'algorithmes de reconstruction de trajectoire de neutrinos qui traversent notre planète. Cela est rendu possible grâce aux données issues du détecteur.

La première partie de ce rapport est une présentation de l'IPHC et du GRPHE, qui participent à ce projet, suivie d'explications sur les neutrinos et leur principe de détection.

La seconde partie détaille le détecteur ANTARES ainsi que les contraintes posées par son environnement.

La dernière partie traitera du travail effectué au cours de ce stage, les problèmes rencontrés, les solutions trouvées et les performances des algorithmes utilisés.

I. Présentation : IPHC-ANTARES

I.1. IPHC et GRPHE

L'Institut Pluridisciplinaire Hubert Curien, IPHC, de renommée internationale au cœur de l'Europe, est dirigé par le Professeur Daniel HUSS. L'IPHC est une unité mixte de recherche du Centre National de la Recherche Scientifique (CNRS) et de l'Université Louis Pasteur de Strasbourg (ULP). Sont ainsi associés trois laboratoires :

- *L'Institut de Recherches Subatomiques (IReS)*, laboratoire de l'IN2P3 (Institut National de Physique Nucléaire et de Physique des Particules) dont le principal domaine de recherche est la physique fondamentale de l'infiniment petit (noyau et particules).
- *Le Centre d'Ecologie, de Physiologie et d'Ethologie (CEPE)*, laboratoire propre du département de Sciences de la Vie du CNRS.
- *Le laboratoire des Sciences Analytiques et Interactions Ioniques Moléculaires et Biomoléculaires*, une unité mixte CNRS-ULP dont les principaux objectifs sont l'élaboration de nouvelles architectures permettant la reconnaissance ionique, le transport et le transfert.

Le Groupe de Recherche en Physique des Hautes Energies, GRPHE, se trouve à Mulhouse et participe au projet ANTARES. Ses domaines de recherche sont la physique des particules auprès des accélérateurs, concernant par exemple l'expérience CMS au CERN à Genève, et la physique des astroparticules avec ANTARES.

Le GRPHE a la charge du traitement des données qui gèrent le détecteur, et collabore aux programmes de contrôle d'ANTARES. L'institut mène également des études de physique dans le cadre de l'expérience, comme l'étude de sources potentielles de neutrinos.

C'est dans les bâtiments de l'IPHC de Strasbourg qu'au nom du GRPHE, j'ai réalisé mon stage, avec l'aide de M. Arnauld ALBERT.

I.2. Le projet ANTARES

I.2.1. Les neutrinos

Les différents types de neutrinos

Fin 1930, Wolfgang Pauli postule l'existence d'une nouvelle particule, du même type que l'électron, qui serait une réponse au problème posé par les désintégrations β : en effet, le spectre des électrons produits est continu alors que le principe de conservation de l'énergie stipule qu'il devrait être discret. Cette nouvelle particule devant être neutre, il l'appelle alors « neutron ». Toutefois, la découverte en 1932 du neutron actuel la renomme « neutrino ». Ses propriétés sont une masse très petite, une charge électrique nulle et une section efficace d'interaction extrêmement faible.

Cette hypothèse d'une nouvelle particule permettait alors de respecter le principe de conservation de l'énergie. C'est seulement dans les années 50 que Reines et Cowan mettent en évidence le neutrino électronique. En 1962, le neutrino muonique est découvert et en 2000 le neutrino tauique.

Dans le cadre d'ANTARES, ce sont surtout les neutrinos muoniques qui sont étudiés.

Astronomie à l'aide de neutrinos

En 1964, Bahcall et Davis suggèrent d'observer les neutrinos produits par le Soleil. Un détecteur constitué de tetracholoroéthylène est placé dans une mine d'or. Les neutrinos solaires transmutent le ^{37}Cl en ^{37}Ar radioactif. En comptant les noyaux d' ^{37}Ar formés, le nombre de neutrinos ayant interagi est obtenu.

La réussite de ce premier essai entraîne le développement de nombreuses autres expériences, utilisant des cibles de plus en plus grandes : IMB (USA), Kamiokande II (Japon), Baksan (Russie).

Sources potentielles de neutrinos

On peut distinguer deux sources de production de neutrinos : les sources galactiques et extragalactiques.

- **Sources galactiques :**

Haute atmosphère : ces neutrinos proviennent de l'interaction des rayons cosmiques primaires dans la haute atmosphère : on les nomme les neutrinos atmosphériques.

Les systèmes binaires : ce sont des objets compacts comme des trous noirs ou des étoiles à neutrons.

Les vestiges de Supernovae : les explosions de Supernovae entraînent l'éjection d'une quantité importante de matière qui va accélérer les particules jusqu'à de hautes énergies (PeV). L'interaction des protons avec cette matière produit alors des neutrinos et des photons. Cette production de neutrinos peut durer de un an à une dizaine d'années après l'explosion initiale. Les Supernovae sont des phénomènes très rares.



Fig. 1 – Image d'une supernovae du télescope Hubble.

Le Soleil : La température au centre du soleil est suffisamment élevée pour initier la combustion de l'hydrogène : $p + p \rightarrow d + e^+ + \nu_e$. Cette réaction initie le cycle de production de l'hélium ${}^4\text{He}$ dont sont issus la quasi-totalité des neutrinos solaires.

Les micro-quasars : Ce sont des sources capables d'accélérer des particules à des énergies extrêmement élevées qui pourraient contribuer à la production du rayonnement cosmique ainsi qu'à celle de neutrinos de hautes énergies.

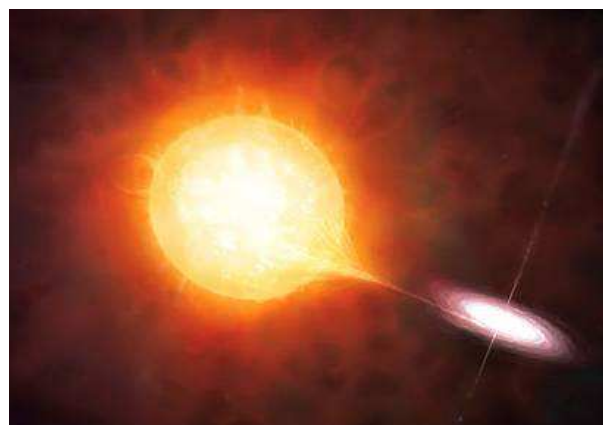


Fig. 2 – Le micro-quasar SS 433 de la constellation Aquila.

- **Sources extragalactiques :**

Les noyaux actifs de galaxies (AGN) : Ce sont les objets les plus lumineux dans l'Univers. Les phénomènes existants dans les AGN ne sont pas encore connus et dépendent des modèles utilisés.

Les sursauts gamma : Ce sont de courtes mais intenses émissions de photons détectés à des énergies de l'ordre d'une centaine de keV. Un des modèles prédit que ces sursauts seraient dus à des sources cosmologiques capables d'émettre des protons accélérés.

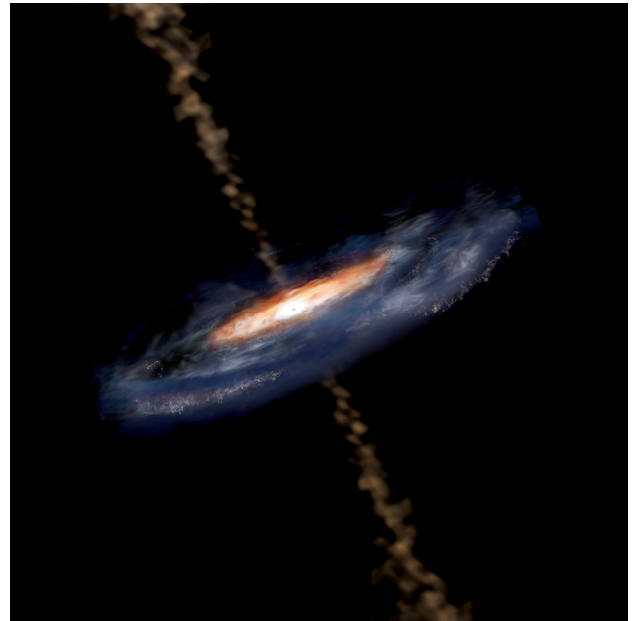


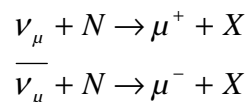
Fig. 3 – *Un AGN*

I.2.2. Principe de détection

L'interaction neutrino-nucléon.

Les neutrinos n'interagissent que par interaction faible. Cela leur permet de parcourir de grandes distances et de traverser la majorité des « obstacles ».

Pour ANTARES, l'interaction neutrino-nucléon est utilisée.



En traversant ainsi la Terre, un neutrino muonique peut interagir et former un muon. Ce muon aura une direction semblable à celle du neutrino. Les calculs cinématiques prédisent que l'angle entre le neutrino incident et le muon produit varie comme :

$$\sqrt{\langle \vartheta_{\nu\mu}^2 \rangle} \sim \sqrt{\frac{m_N}{E_{\nu}}}$$

Ainsi, par exemple, pour un neutrino de 300 GeV, cet angle vaut 1,16°.

Dans le cadre d'ANTARES, seuls les neutrinos muoniques sont exploitables. En effet, le muon engendré est une particule qui va peu interagir et qui peut parcourir une grande distance avant de disparaître. Pour un neutrino électronique, l'électron engendré réagit trop avec son environnement et il en résulte une « gerbe » électromagnétique dont il est difficile de tirer une trajectoire initiale. De même pour le tau qui interagira beaucoup lors de son parcours.

Radiation Cerenkov

Lorsqu'une particule chargée se déplace dans un milieu avec une vitesse supérieure à celle de la lumière dans ce milieu, il y a émission de photons dits photons de Cerenkov. Les muons considérés se trouvent au-delà du seuil d'émission de la lumière Cerenkov dans l'eau de mer, ils vont donc émettre de la lumière. Les photons sont émis principalement à de petites longueurs d'onde, typiquement entre 330 et 600 nm (couleurs allant du violet au jaune).

Les télescopes à neutrinos utilisent des photomultiplicateurs afin d'estimer le temps d'arrivée des photons permettant ainsi la reconstruction de la trajectoire du muon. La quantité de lumière permet de quantifier l'énergie des muons et de reconstituer la trace des neutrinos.

La figure 4 illustre le principe de détection de la radiation Cerenkov.

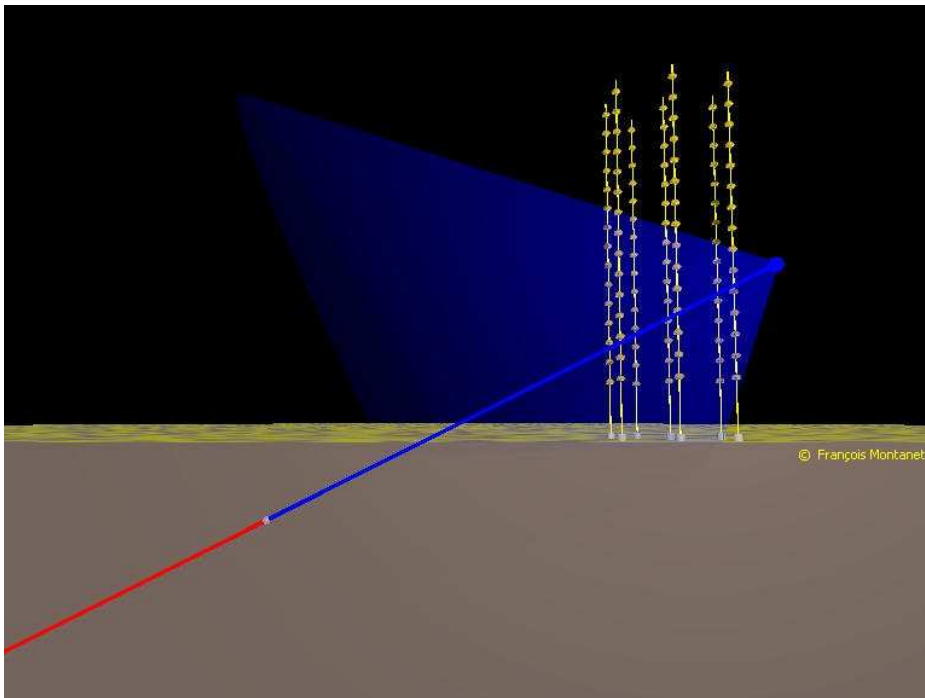


Fig. 4 - Illustration de l'effet Cerenkov. La trace en rouge est un neutrino muonique. La trace en bleu représente le muon engendré. Le cône bleu montre les photons émis par le muon dans l'eau.

II. Le télescope à neutrinos ANTARES

II.1. Le détecteur

ANTARES est situé au large de Toulon, à environ 40 km de la côte.

Il est constitué de 12 lignes verticales d'environ 450m de hauteur ancrées à 2500 m de profondeur. Ces lignes servent de support aux instruments de mesure. **La figure 5** représente une vue d'artiste du détecteur.

Positionnement absolu du détecteur

Ce positionnement nécessite la connaissance de trois coordonnées géographiques (latitude, longitude, altitude) et trois angles correspondant au référentiel terrestre. La position est obtenue grâce à une mesure GPS effectuée sur le bateau lors du déploiement d'une ligne, et la profondeur grâce à des capteurs de pression et à la longueur de câble déployée.

Ligne d'instrumentation

Chaque ligne comporte 75 photomultiplicateurs (PM). On parle aussi de modules optiques (OM) lorsque l'on considère le PM et son électronique associée. L'ensemble de trois modules optiques définit un étage de détection d'une ligne. Les étages sont séparés d'environ 14,5m.

L'annexe A est un rappel du principe de fonctionnement d'un photomultiplicateur et présente également ses principales caractéristiques du modèle utilisé dans ANTARES.

Un ensemble de balises acoustiques sont positionnées aux pieds des lignes et autour du détecteur. En communiquant entre elles, elles permettent ainsi de connaître les positions relatives des lignes.

Les modules de contrôle

Chaque ligne possède un SPM (Module de tension électrique), un SCM (Module de contrôle de la ligne).

Chaque étage possède un LCM (Module de contrôle local). Un secteur est composé de 4 LCM et d'un MLCM (Module de contrôle local maître) qui rassemble en plus les informations des modules. Il y a donc 5 MLCMs par ligne. Les informations de chaque étage sont regroupées au niveau du LCM, puis du MLCM, et envoyées au SCM. L'ensemble des données est finalement transporté sur terre via un câble électro-optique.

Les étages de détection

Outre les trois modules optiques, les étages contiennent des inclinomètres et des compas, éléments permettant de connaître la position et l'inclinaison exacte des modules optiques à tout moment. **La figure 6** représente une vue 3D d'un étage.

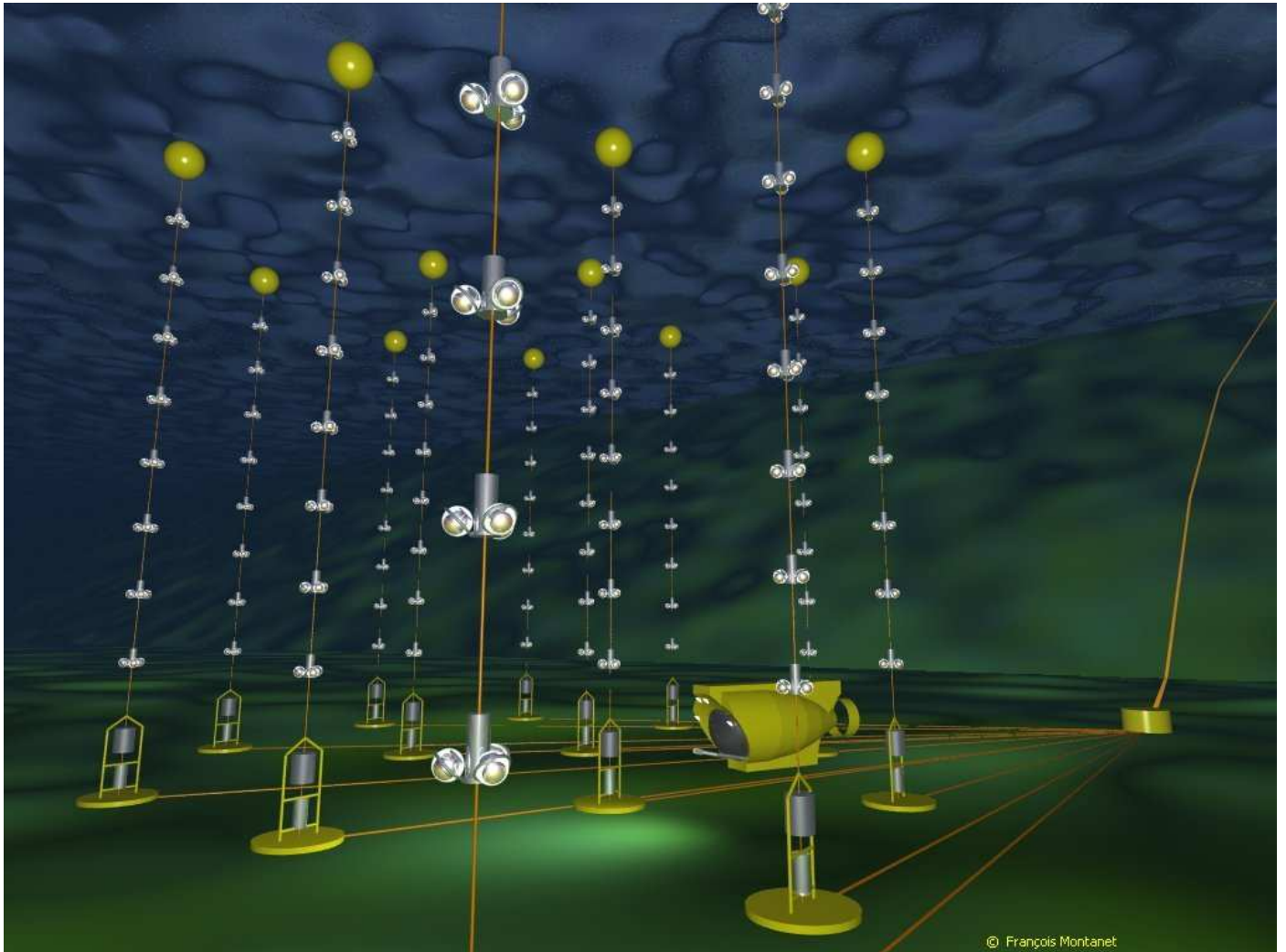


Fig. 5 – Représentation des lignes immergées du détecteur.

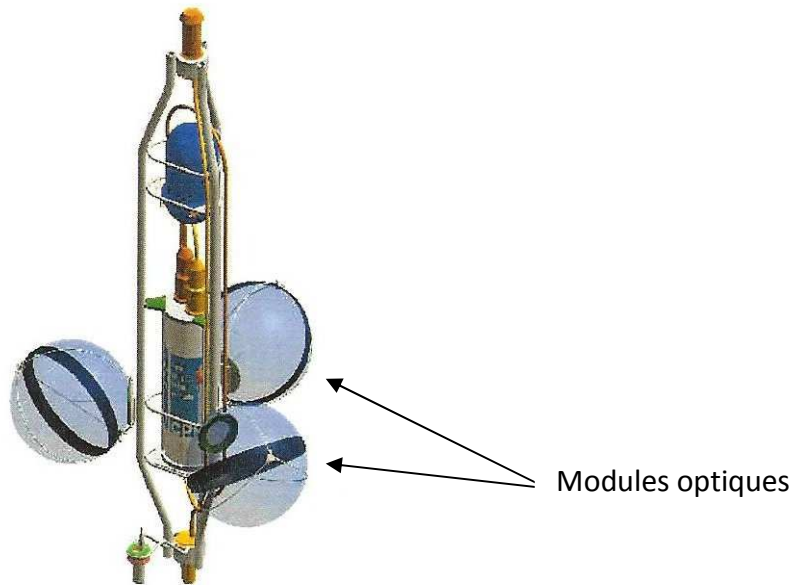


Fig. 6 – Illustration d'un étage du détecteur ANTARES. On peut y distinguer les trois modules optiques, le conteneur électronique (au centre, dans le cylindre)



Fig. 7 – Photos d'un module optique. A l'intérieur de la sphère se trouve le photomultiplicateur.



Les modules optiques (figure 7)

Un module optique se compose d'un photomultiplicateur, d'une sphère de verre transparent, d'une cage en mu-métal, d'une LED et de l'électronique du photomultiplicateur.

Les sphères sont résistantes à 700 bars. Elles ont un indice de réfraction (1.47) proche de celui de l'eau de mer (1.35) afin de minimiser les pertes par réflexion. Elles sont également imperméables à l'eau et à l'air, résistantes à la corrosion et sont chimiquement, électriquement et magnétiquement neutres.

La cage en mu-métal : le champ magnétique terrestre influence les trajectoires des électrons entre la photocathode et la première dynode. Une cage magnétique a donc été développée dans le but de réduire cet effet.

Le circuit intégré ARS (premier élément participant à la transmission des données)

Il faut numériser les signaux des photomultiplicateurs avant de les envoyer à terre à l'aide du câble électro-optique. Cette numérisation est effectuée par le circuit intégré ARS (Analog Ring Sampler). Deux ARS par photomultiplicateur sont utilisés afin d'avoir le moins de temps mort possible : lorsqu'un ARS s'occupera de digitaliser un signal, le second prendra le relais pour le signal suivant.

L'information, après avoir été digitalisée, est donc envoyée au SPM de la ligne via les MLCM. Le câble électro-optique envoie les données dans une salle informatique située à Seyne sur Mer. Là, elle est traitée automatiquement par des programmes qui vont créer des fichiers de données, avant de les envoyer à Lyon où elle sera centralisée et mise à la disposition de la collaboration via des serveurs.

II.2. Contraintes posées par l'environnement

La sélection du site nécessite de prendre en considération différents paramètres marins tels que la transparence de l'eau, le bruit de fond optique, le dépôt de salissures sur les modules optiques, les courants marins mais aussi des conditions météorologiques ou encore la profondeur. De nombreuses études de caractérisation du milieu ont été effectuées.

Dépôt de salissures sur les modules optiques

Une ligne test a été immergée deux fois sur un site situé à un mile nautique du site d'ANTARES.

Les résultats montrent que la perte de transmission due aux salissures est de l'ordre de 60% au sommet de la sphère après trois mois d'immersion et de 2% à l'équateur après huit mois d'immersion. Comme les modules optiques du détecteur sont orientés à 135° vers le bas, le dépôt de sédiment n'entraîne pas d'importante perte de transmission, même après plusieurs années d'immersion.

Transparence de l'eau

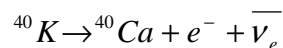
Cette donnée a un impact sur la résolution angulaire avec laquelle la direction du muon peut être reconstruite (plus la longueur de diffusion est grande, moins les photons vont être déviés). C'est un paramètre contrôlé régulièrement.

Bruits de fond optiques

Il s'agit de la bioluminescence et de la désintégration du potassium 40.

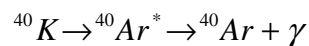
La bioluminescence est l'émission de lumière par les organismes vivants, pour de multiples raisons : éclairage, défense contre les prédateurs, etc. Les organismes marins émettent dans le bleu-vert, entre 450 et 520 nm environ, ce qui correspond à la longueur d'onde de l'émission Cerenkov.

Un OM enregistre aussi la lumière produite par désintégration du potassium 40 présent dans l'eau de mer. En effet, cet isotope radioactif du potassium produit des électrons qui entraînent une émission de lumière.



L'électron est émis avec une énergie cinétique moyenne de 1.12 MeV, ce qui est supérieur au seuil d'énergie d'émission Cerenkov dans l'eau dans la majeure partie des cas. Il produit des photons ayant une longueur d'onde entre 300 et 600 nm.

Il peut aussi se désintégrer par capture électronique :



Une capture électronique génère en moyenne 80 photons Cerenkov.

III. Algorithmes de reconstruction de traces

La première phase de mon stage a été la compréhension des méthodes utilisées par la collaboration ANTARES pour réaliser la reconstruction de trajectoire (dénommée également « trace »).

III.1. Étude des algorithmes

Toute la procédure de reconstruction est basée sur les données recueillies par les modules optiques. Lorsqu'un muon traverse les lignes, la lumière qu'il émet par effet Cerenkov est repérée par les photomultiplicateurs. Le muon se propage en ligne droite et émet des photons en perdant de l'énergie. Ces photons sont envoyés dans une certaine direction, toujours sous le même angle par rapport à la direction du muon, l'angle de Cerenkov. Dans l'eau de mer, il a été calculé que cet angle vaut environ 42° .

Il est donc possible, en rassemblant les données des modules touchés par cette lumière, de reconstruire la direction du muon.

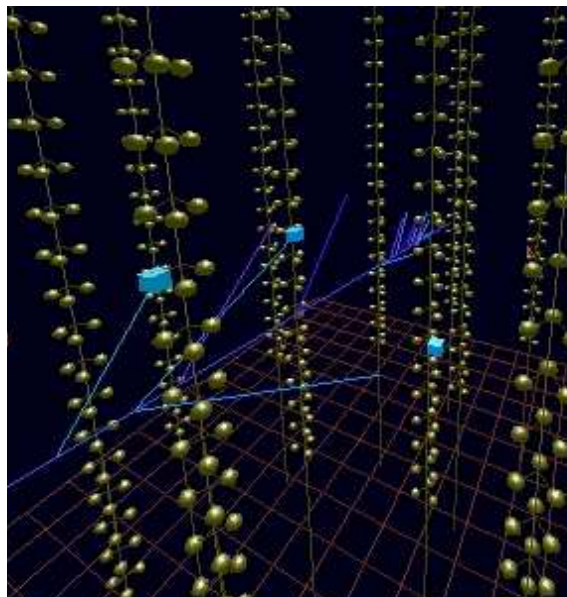


Fig. 9 – Schéma d'un muon passant au travers des lignes de détection. On distingue la trace du muon et les photons qui en sont issus, ainsi que les impacts sur les photomultiplicateurs.

Plusieurs algorithmes ont été mis au point par la collaboration ANTARES. Leur utilisation dépend de certains paramètres, comme l'énergie du muon incident.

Mon objectif a donc été de mettre au point des programmes de reconstruction des traces en me basant sur ces algorithmes. Pour ce faire, il a d'abord fallu utiliser des logiciels de simulations d'événements afin de juger de la fiabilité des méthodes utilisées et les adapter si nécessaire.

Plusieurs fichiers de données ont ainsi été générés, incluant les informations sur le muon simulé (point de départ, vrais paramètres, charge) ainsi que les différents « hits » dans les modules optiques. **On parle de « hit » lorsqu'un PM a détecté un photon.**

La reconstruction suivra ces deux étapes :

- le « Linear prefit », qui constitue une première estimation des paramètres de la trace, se base sur des relations linéaires.
- Le « M-Estimator », qui se base sur le résultat donné par le Linear Prefit, affine ces paramètres à l'aide d'estimateurs et de fonctions à minimiser.

Le Linear Prefit

Cette méthode va fournir une première estimation.

Il est supposé ici que les hits ont lieu en des points localisés sur la trace du muon. Dans ces conditions, on peut écrire :

$$y = H\theta$$

Avec : $y = [x_1, y_1, \dots, z_n]$ les positions des hits, $\theta = [p_x, d_x, p_y, d_y, p_z, d_z]$ vecteur contenant les paramètres de la trace et H une matrice contenant les temps des hits.

$$H = \begin{bmatrix} 1 & t_1 & & & & \\ & & 1 & t_1 & & \\ & & & & 1 & t_1 \\ 1 & t_2 & & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Si les erreurs sur les paramètres x, y et z des positions des hits sont approximativement égales, on peut écrire au final :

$$\theta = \frac{1}{\sum f_i \sum f_i c_i^2 t_i^2 - (\sum f_i c_i t)^2} \begin{bmatrix} \sum f_i c^2 t_i^2 \sum f_i x_i c - \sum f_i c t_i \sum f_i c t_i x_i \\ \sum f_i \sum f_i x_i c t_i - \sum f_i c t_i \sum f_i x_i \\ \sum f_i c^2 t_i^2 \sum f_i y_i c - \sum f_i c t_i \sum f_i c t_i y_i \\ \sum f_i \sum f_i y_i c t_i - \sum f_i c t_i \sum f_i y_i \\ \sum f_i c^2 t_i^2 \sum f_i z_i c - \sum f_i c t_i \sum f_i c t_i z_i \\ \sum f_i \sum f_i z_i c t_i - \sum f_i c t_i \sum f_i z_i \end{bmatrix}$$

Pour ANTARES, une erreur moyenne a été calculée pour les trois coordonnées x, y et z : $f \approx 10$ cm. Ce résultat fut obtenu en prenant en compte les erreurs de chaque instrument renseignant sur la position des OM (compas, inclinomètres, balises acoustiques, etc.). C'est donc cette dernière relation simplifiée qui va être utilisée.

Le M-Estimator

Les M-estimateurs constituent une large classe de statistiques obtenues par la minimisation d'une fonction dépendant des données et des paramètres du modèle.

La procédure de reconstruction doit estimer des paramètres de trace qui sont compatibles avec les hits observés. Ceci est réalisé en calculant le temps résiduel pour chaque hit, comme étant fonction des paramètres de la trace \vec{p} et \vec{d} :

$$r_i = t_i - t_i^{th}(\vec{p}, \vec{d})$$

Avec t_i le temps de l'i-ème hit. t_i^{th} se calcule de la façon suivante :

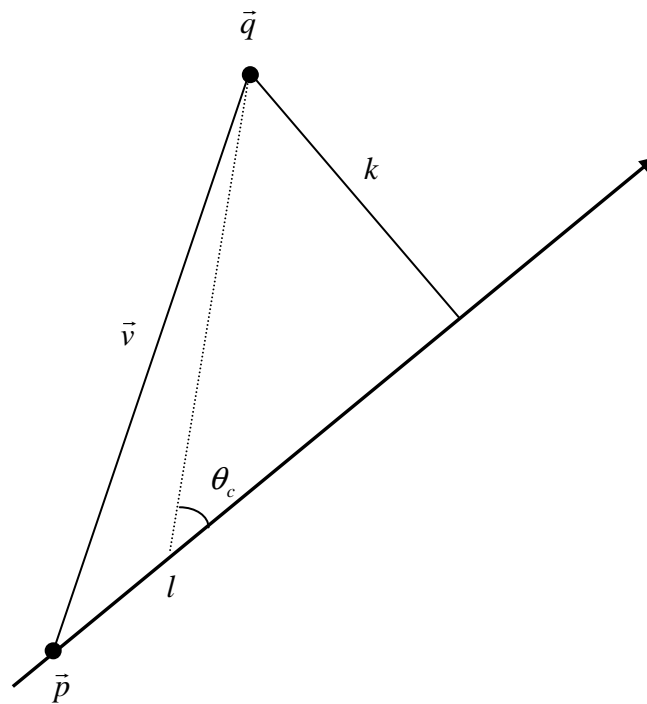


Fig. 10 – Schéma de la trace d'un muon et des variables utilisées pour le M-Estimator. \vec{p} est la position initiale du muon, \vec{q} la position de l'OM concerné. La ligne en pointillés représente l'émission d'un photon.

On calcule : $\vec{v} = \vec{q} - \vec{p}$

Les composants de \vec{v} sont : $l = \vec{v} \cdot \vec{d}$ et $k = \sqrt{v^2 - l^2}$

On calcule ensuite le temps théorique d'arrivée des photons par :

$$t^{th} = t_0 + \frac{1}{c} \left(l - \frac{k}{\tan \theta_c} \right) + \frac{1}{v_g} \left(\frac{k}{\sin \theta_c} \right)$$

c est la vitesse de la lumière dans le vide (0.3 m/ns), v_g la vitesse de la lumière dans l'eau (0.2175 m/ns), et θ_c l'angle de Cerenkov (ici, 42°).

Enfin, une nouvelle estimation des paramètres est alors donnée lorsque l'on a Q minimum :

$$Q = \sum_{i=1}^N g(r_i(\vec{p}, \vec{d}))$$

La fonction $g(r)$ est souvent $L(r)$ où L est la fonction de densité de probabilité de trouver un hit avec un temps résiduel r . On utilise aussi $g(r) = -r_i^2$ (méthode moindres carrés). Toutefois, un autre estimateur est utilisé et fonctionne correctement en association avec la méthode utilisé pour le Linear Prefit : le L1-L2.

$$g(r) = 2\sqrt{1 - \frac{a_i r_i^2}{2}} - 2$$

Avec a_i l'amplitude du ième hit.

La méthode est de calculer Q avec les paramètres du Linear Prefit, puis de faire varier ces paramètres dans un espace défini. Les paramètres pour lesquels Q est minimum sont sélectionnés.

L'étape suivante fut de réaliser un programme en langage C capable de lire les données et d'en tirer les paramètres des traces grâce aux calculs précédents. Un extrait du code C concernant le calcul du Linear Prefit est présenté dans *l'annexe B*.

Le programme fut tout d'abord testé sur des données simulées.

III.2. Données simulées

Le but de cette procédure est de vérifier que les algorithmes donnent des résultats corrects. Ainsi, des logiciels permettent de simuler le passage d'un muon à travers le détecteur. Les données que l'on obtient sont alors les hits engendrés par la particule et les paramètres utilisés par le logiciel pour créer le neutrino. On peut alors vérifier l'exactitude des résultats obtenus.

La figure 11 montre un exemple de fichier généré par simulation.

Nous avons tout d'abord testé le programme dans le cas idéal, puis nous avons étudié l'influence du bruit de fond sur les résultats.

start_event:	4	1																		
neutrino:	1	-163.444	217.099	-177.325	0.172463	-0.349385	0.920970	0.276731E+04	0.0	0.028394	0.556388	-1	14	2						
track_earthlepton:	1	-163.444	217.099	-177.325	0.177007	-0.347521	0.920813	0.122761E+04	0.0	0	6									
w2list:	0.1194E+36	0.7658E+07	0.1639E-38	0.8491E+14	0.9196E+00	0.2554E+04														
weights:	0.7310E+14	0.1007E+18	0.6001E+09																	
w3list:	0	0.5956E-08	0.6001E+09																	
track in:	1	-160.846	211.999	-163.812	0.177007	-0.347521	0.920814	0.122162E+04	0.000000E+00	5	-608.000									
hit:	1	751	0.1000E+01	1376.59	5	1	0	0.00												
hit:	2	456	0.1000E+01	1614.32	-5	1	0	0.00												
hit:	3	829	0.1000E+01	1422.42	-5	1	0	0.00												
hit:	4	752	0.1000E+01	1444.70	-3	1	0	0.00												
hit:	5	763	0.1000E+01	1243.25	-3	1	0	0.00												
hit:	6	763	0.1000E+01	1243.60	-3	1	0	0.00												
hit:	7	766	0.1000E+01	1210.94	3	1	0	0.00												
hit:	8	766	0.1000E+01	1212.86	-3	1	0	0.00												
hit:	9	767	0.1000E+01	1218.63	-3	1	0	0.00												
hit:	10	775	0.1000E+01	1184.13	3	1	0	0.00												
hit:	11	834	0.1000E+01	1407.41	-3	1	0	0.00												
hit:	12	841	0.1000E+01	1312.42	-3	1	0	0.00												
total_hits:	12																			
end_event:																				

Fig. 11 – Exemple de données simulées. En rouge, les paramètres du neutrino simulé (coordonnées initiales, direction, charge). En bleu, la trace calculée par le logiciel de simulation. En vert, les hits simulés ayant servi au calcul de la trace.

III.2.1. Cas idéal

Des premiers essais ont été réalisés dans le cas idéal : aucun bruit de fond optique, positions des OM fixes, etc.

A l'aide du programme de statistiques « ROOT », développé par le CERN, j'ai pu tracer des histogrammes des résultats.

La figure 12 montre la distribution des dx calculés en fonction des dx théoriques, dx étant la composante en x de la direction du muon. On constate que les calculs effectués par mon programme sont corrects : les valeurs calculées sont pour la majeure partie proches des valeurs théoriques. On obtient une erreur moyenne de 10 cm sur chaque paramètre. Les résultats pour dy et dz présentent les mêmes allures.

De même, les erreurs angulaires entre la trace théorique et la trace calculée restent pour la majeure partie inférieure à 20°, ce qui correspond aux résultats obtenus par la collaboration ANTARES.

III.2.2. Incidence du bruit de fond

On a ensuite généré à l'aide un fichier de données prenant en compte le bruit de fond (^{40}K , bioluminescence). Les premiers essais de calculs ont été faits sans aucune modification de l'algorithme.

La figure 13 illustre logiquement l'inadaptation de l'algorithme : on obtient des erreurs allant de la dizaine de cm au m. Les résultats sont corrects pour les petites valeurs des paramètres (aux alentours de 0), mais deviennent totalement faux pour des valeurs plus élevées. La présence de bruit de fond fausse donc les estimations des paramètres de la trace.

Plusieurs étapes de corrections des algorithmes ont alors été testées.

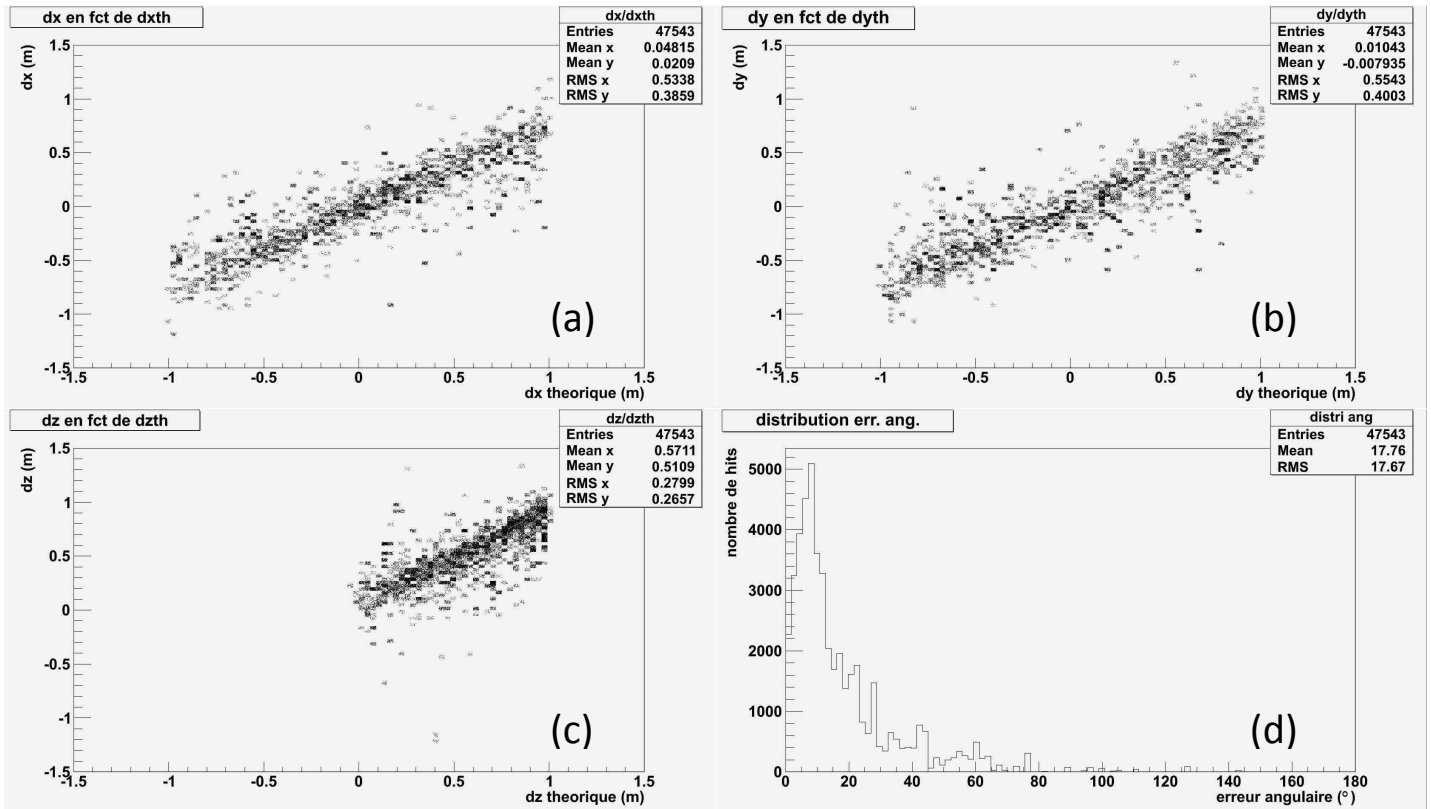


Fig. 12 – Graphiques obtenus avec les données simulées dans le cas idéal : a, b et c représentent les paramètres dx , dy et dz en fonction des paramètres théoriques donnés par le logiciel de simulation. En d est représenté l'écart angulaire entre la trace générée et la trace calculée.

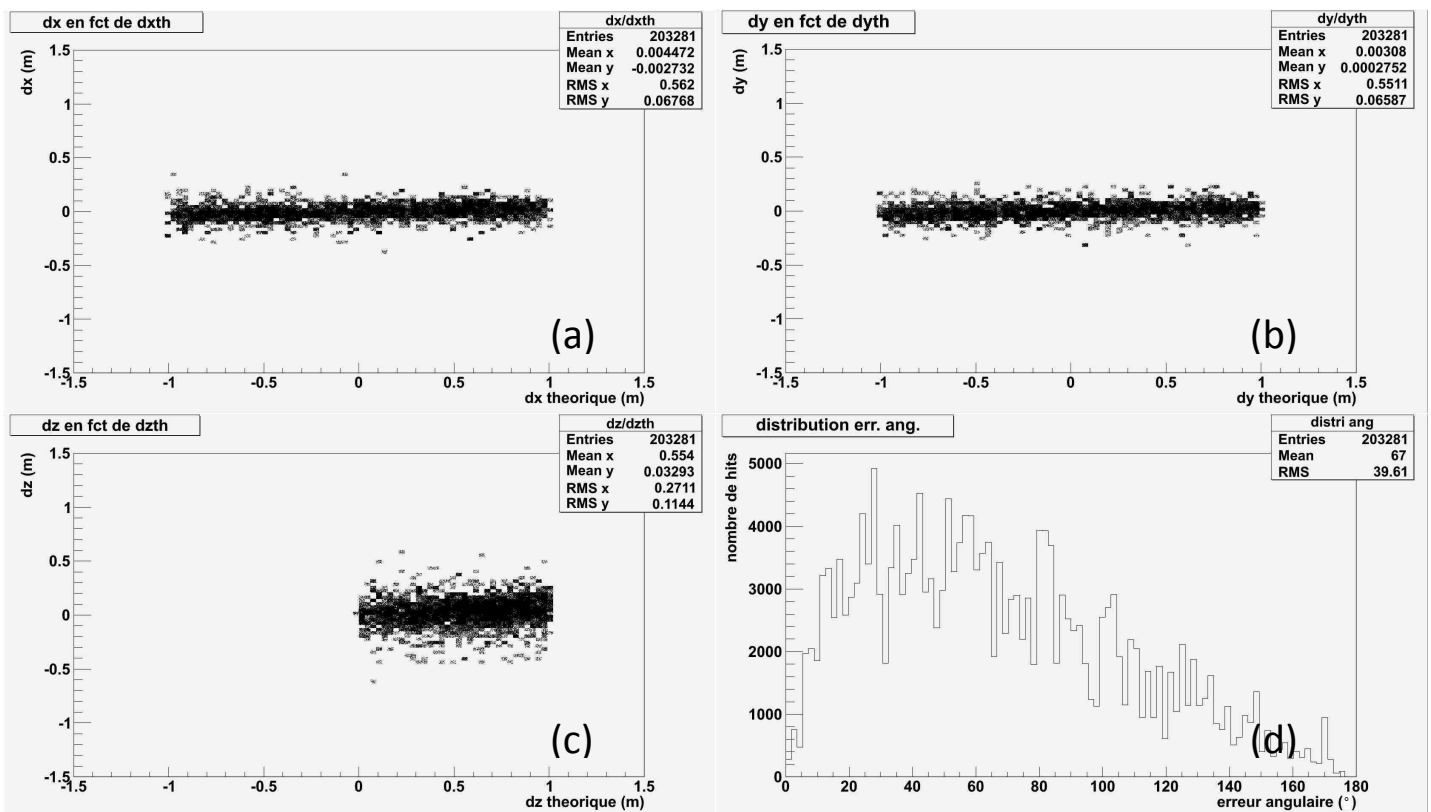


Fig. 13 – Graphiques obtenus avec des données simulées en incluant le bruit de fond.

Filtrage des hits sur PM par rapport à la distance à la trace

Le Linear Prefit nous donne une première estimation de la trace. Un des premiers critères de sélection de hits pour le M-Estimator est la distance du hit par rapport à la trace calculée.

La **figure 14** montre la répartition de k , distance par rapport à la trace, dans le cas idéal précédent, et pour le cas actuel. On remarque ainsi que dans le cas idéal, les valeurs de k sont pour la plupart situées autour de 30 cm, alors qu'ici, nous obtenons un pic autour de 30 cm et un large pic autour de 50 cm.

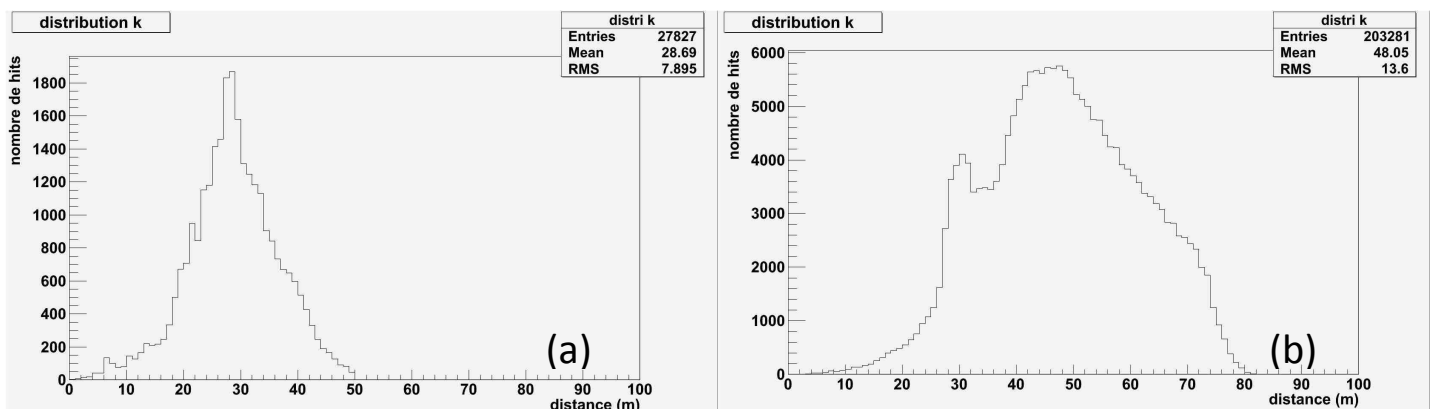


Fig. 14 – Distribution de la distance des hits à la trace calculée (k) ; *a* est le cas idéal, *b* le cas avec bruit de fond.

L'algorithme est alors testé en ne sélectionnant que les hits pour lesquels k est inférieur ou égal à 40 cm. Toutefois, les résultats obtenus présentent toujours les mêmes erreurs.

Filtrage par rapport à la charge

Les hits ayant moins de 0.5 p.e. (photons-électrons) de charge ne sont pas pris en compte, aucune amélioration n'est visible, mais cela élimine un autre champ d'investigation.

Prise en compte d'une seule ligne de mesure

On ne considère qu'une seule ligne de mesure, celle ayant le plus de hits, afin de réduire au maximum le bruit de fond.

La **figure 15** montre les résultats : cela m'a permis d'obtenir une meilleure précision dans les résultats, même s'il y a toujours présence de grandes erreurs.

Filtrage par rapport à l'angle d'incidence

La présence d'une « barre horizontale » sur les graphiques laisse penser que certains événements ne peuvent pas être bien reconstruits. Nous en avons déduit qu'il pourrait s'agir de muons ayant une trajectoire horizontale, c'est-à-dire perpendiculaire aux lignes du détecteur. Les algorithmes utilisés ici sont adaptés à des reconstructions de muons ascendants ou descendants.

Un nouveau filtre est donc ajouté pour éliminer un maximum de ces évènements. Après un premier Linear Prefit, on élimine les évènements dont l'angle entre la trace calculée et les lignes (c'est-à-dire l'axe Z du repère) est supérieur à 30°. **La figure 16** montre les résultats obtenus : on parvient ainsi à réduire ces barres horizontales mais on observe également une forte diminution du nombre d'évènements traités et il existe toujours des erreurs très aléatoires, notamment pour le paramètre Z.

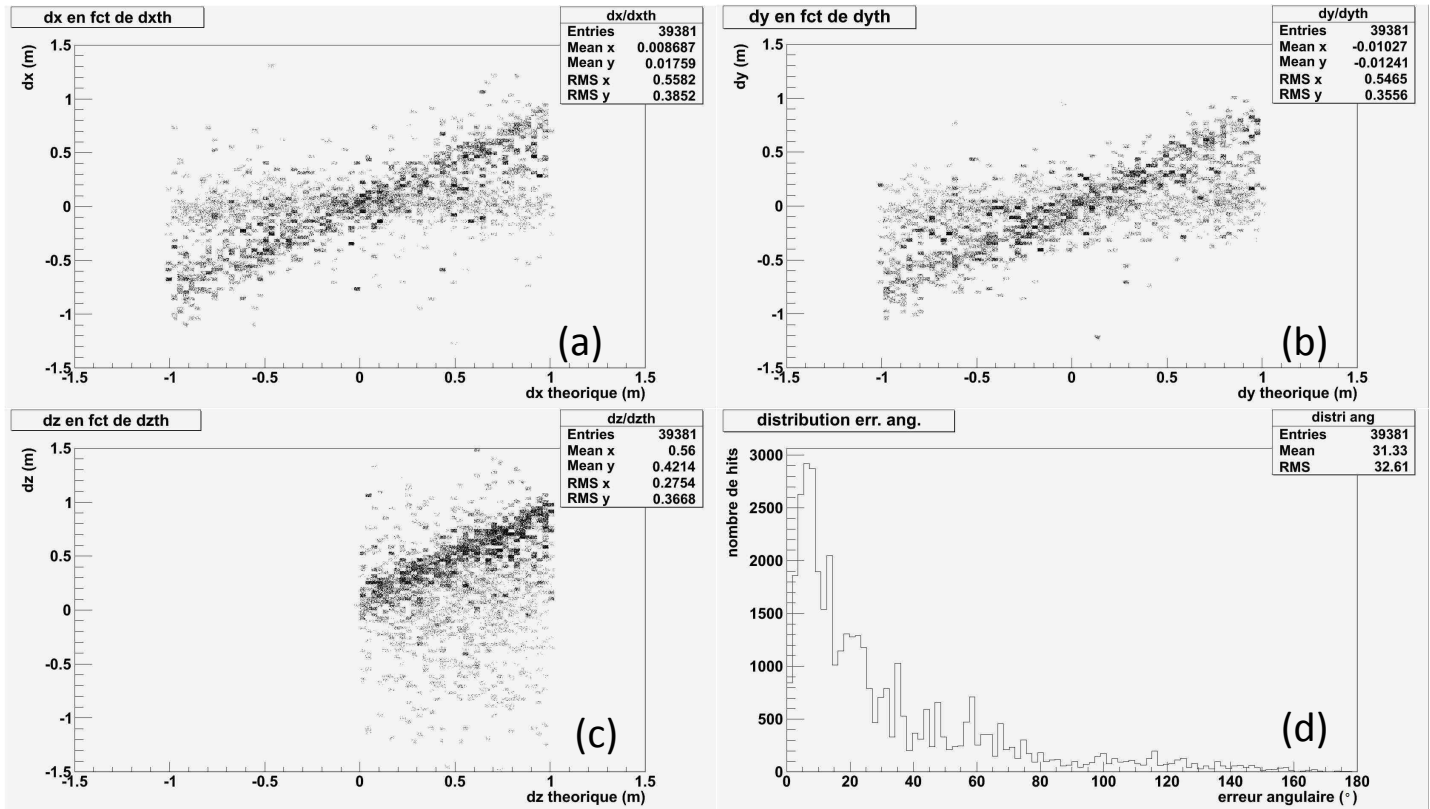


Fig. 15 – Graphiques obtenus avec les données simulées dans le cas réel, avec une seule ligne de mesure prise en compte.

Le Linear Prefit n'étant à ce stade pas correct, il fallait trouver un moyen efficace après tous ces tests pour ne sélectionner que les hits provenant de muons et non pas provenant du bruit de fond.

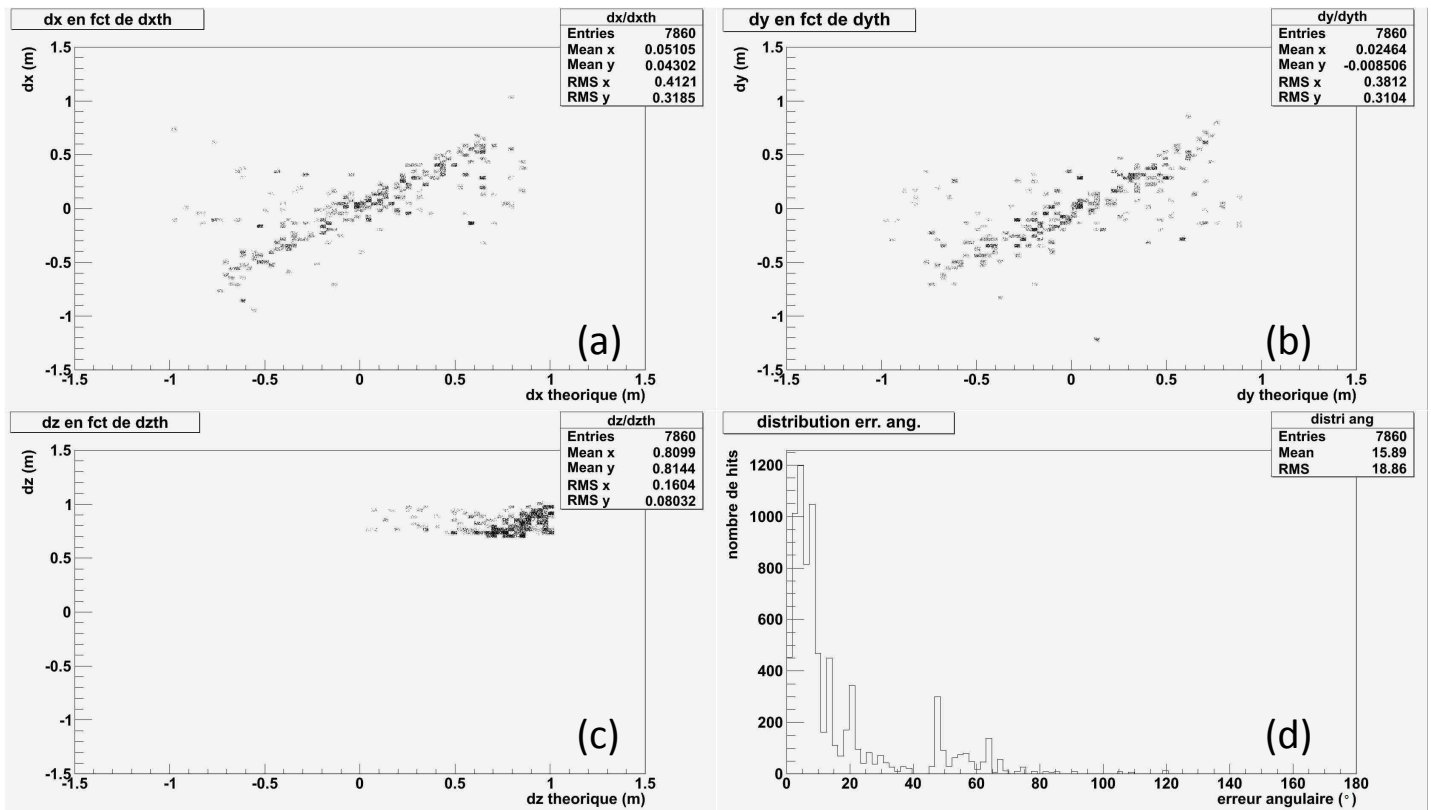


Fig. 16 – Graphiques obtenus avec les données simulées dans le cas réel, avec filtrage sur l'angle.

En me basant sur les résultats de la thèse d'A. Heijboer, *Track Reconstruction and Point Source Searches with ANTARES* [3], nous avons décidé d'appliquer une méthode de présélection des hits de la manière suivante :

Présélection

On recherche tout d'abord le hit ayant la plus grande amplitude dans l'évènement. Il servira de référence pour la sélection des autres. Les hits conservés sont ceux pour lesquels on a :

$$|dt| \leq 100 + \frac{d}{v_g}$$

où dt est la différence de temps $t_{ref} - t_i$, v_g la vitesse de groupe de la lumière dans l'eau de mer, à $\lambda = 460nm$ (0,2175 m/ns) et d la distance entre les deux OM où ont lieu les 2 hits. De cette façon, une grande partie du bruit de fond est éliminée.

Linear Prefit

Une nouvelle sélection est faite : on ne prend en compte que les hits en coïncidence locale ou avec une amplitude inférieure ou égale à 3 p.e. Une coïncidence locale est définie comme la combinaison d'au moins 2 hits sur un même étage dans un espace de 25ns.

M-Estimator

Une nouvelle présélection est faite afin de rendre cet estimateur le plus précis possible. Pour être sélectionné, un hit doit avoir un temps résiduel compris entre -150 et 150 ns, ainsi qu'une distance à la trace précédemment calculée de 100 m au maximum. Les hits ayant une amplitude de 2,3 p.e. ou plus sont automatiquement pris en compte.

Les résultats de la figure 17 ainsi obtenus sont très proches des données du cas idéal. On remarque aussi que le M-Estimator affine correctement les paramètres calculés.

L'algorithme mis en place est donc correct et permettra la détermination de traces pour les vraies données.

La prochaine étape sera de traiter des données provenant d'ANTARES, où de nouveaux paramètres seront à prendre en compte (position des OM variant dans le temps à cause du courant marin par exemple

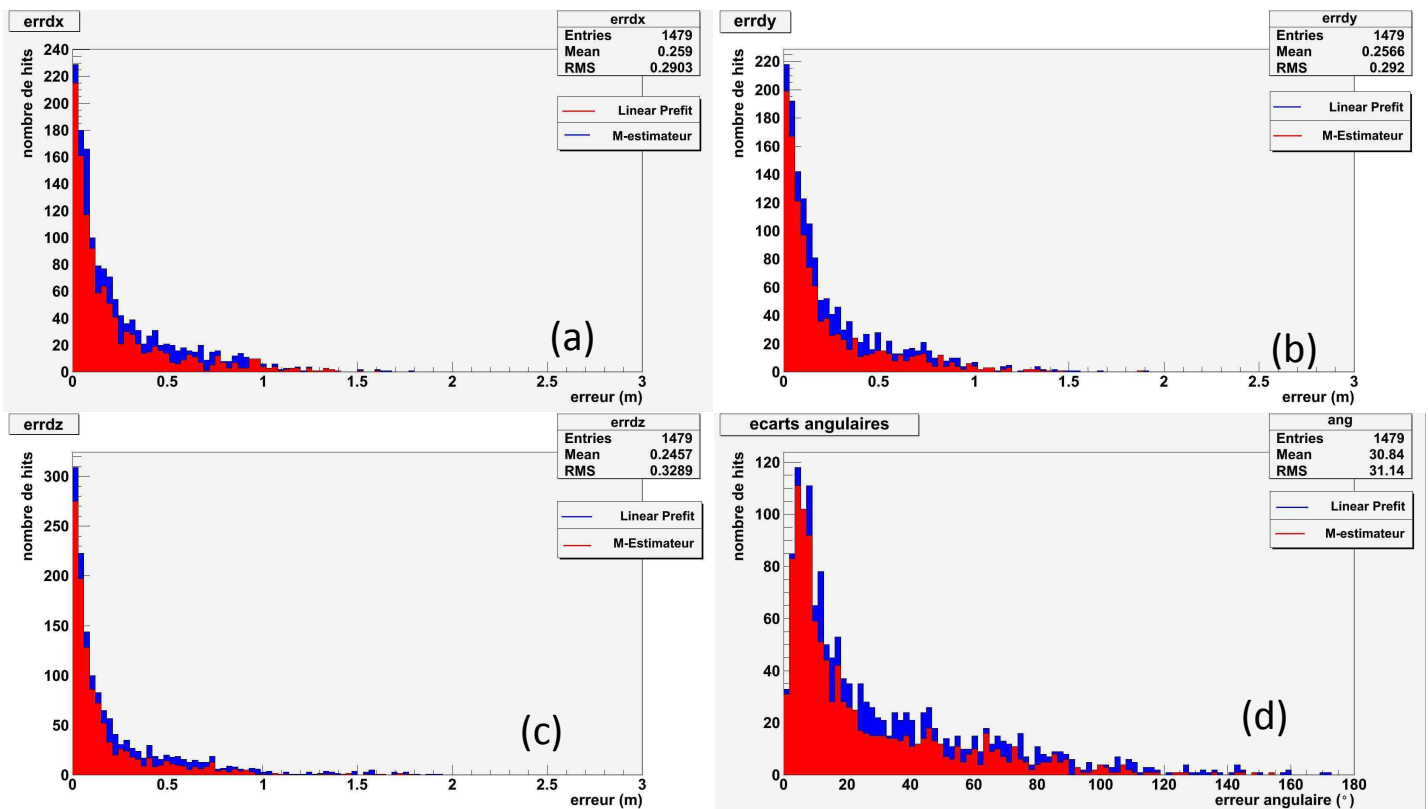


Fig. 17 – Graphiques finaux. En a, b et c, les erreurs sur dx, dy et dz. En d, l'écart angulaire entre la trace calculée et la trace théorique. La partie bleue représente les résultats obtenus après le Linear Prefit et la partie rouge après le M-Estimator.

III.3. Données réelles

Les données réelles se présentent différemment des données simulées. Comme le montre **la figure 19**, les adresses IP des OM sont indiquées et non pas leur numéro. Le « frame index » renseigne sur l'heure de l'évènement : un « frame index » correspond à un temps écoulé d'environ 104 ms. Ainsi, on peut aisément avoir l'heure de début de l'évènement.

A l'aide de cette donnée et d'un fichier renseignant les positions des OM (**voir figure 20**), les positions sont obtenues en fonction du temps par la méthode suivante :

Soit (x_0, y_0, z_0) les coordonnées absolues de l'étage concerné, (x', y', z') les coordonnées locales de l'OM (**voir figure 18**), et *Pitch, Roll, Heading* les trois angles de rotations donnés par le fichier renseignant sur les positions (**figure 19**).

On définit : $\alpha = -\arcsin\left(\frac{\sin(Roll)}{\cos(Pitch)}\right)$, $\beta = Pitch$ et $\gamma = Heading - \frac{\pi}{2}$.

Quatre étapes se succèdent alors pour obtenir les coordonnées finales absolues de l'OM.

- Rotation d'angle α autour de $x_3 = x'$

$$\begin{aligned}x_3 &= x' \\y_3 &= y' \cos \alpha + z' \sin \alpha \\z_3 &= -y' \sin \alpha + z' \cos \alpha\end{aligned}$$

- Rotation d'angle β autour de $y_2 = y_3$

$$\begin{aligned}x_2 &= x_3 \cos \beta - z_3 \sin \beta \\y_2 &= y_3 \\z_2 &= x_3 \sin \beta + z_3 \cos \beta\end{aligned}$$

- Rotation d'angle γ autour de $z_1 = z_2$

$$\begin{aligned}x_1 &= x_2 \cos \gamma + y_2 \sin \gamma \\y_1 &= -x_2 \sin \gamma + y_2 \cos \gamma \\z_1 &= z_2\end{aligned}$$

- Finalement, on rajoute ces coordonnées aux coordonnées absolues :

$$x = x_1 + x_0$$

$$y = y_1 + y_0$$

$$z = z_1 + z_0$$

Tous ces calculs ont donc été inclus dans le programme de reconstruction.

L'étude menée ici fut de comparer les résultats obtenus pour deux fichiers de données.

En effet, deux formats sont générés : un fichier avec des hits déjà sélectionnés par les programmes de récupération, et un autre fichier avec les données brutes, non-filtrées.

Un nouveau paramètre a dû être pris en compte dans les calculs. Il s'agit du temps T_0 de parcours de l'information via tous les câbles, jusqu'à la côte. Ce temps étant différent pour chaque étage de détection, il faut impérativement le prendre en compte afin d'avoir des données normalisées.

N° OM	x' (m)	y' (m)	z' (m)
0	0.578	0	-0.142
1	$-\frac{0.578}{2}$	-0.578	-0.142
2	$-\frac{0.578}{2}$	0.578	-0.142

Fig. 18 – Valeurs de x' , y' et z' , coordonnées d'un OM par rapport à son étage.

```

RUN:                28034
event:              5525
event type:         80000021
frame time 1:       0
frame time 2:       14680069
frame target:       6
frame index:        7
Minimal time:       2.72698e+08
Maximal time:       2.72698e+08
Maximal amplitude:  5.26317
Total amplitude:    19.1267
Number of hits:     9

triggered hits:
0 1031 2 64 10907823 86 63
1 1031 4 64 10907824 143 55
2 1151 2 64 10907828 220 100
3 1151 5 193 10907828 39 71
4 1149 1 193 10907829 108 56
5 1154 3 193 10907837 123 67
6 1154 4 64 10907837 92 71
7 1134 0 64 10907804 180 48
8 1134 4 64 10907803 142 70
    
```

Fig. 19 – Exemple de données récupérées d’ANTARES. En rouge, le frame index informe sur l’heure de l’évènement. En bleu, les informations sur les hits : l’adresse IP de l’OM touché, le temps en ns, la charge.

```

"APP_START", "APP_STOP", "LCM_ID", "POSITION", "X", "Y", "Z", "HEADING", "PITCH", "ROLL"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1624", "101", "8221.6", "2479.7", "-2386.06568", "3.57288", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1631", "102", "8221.6", "2479.7", "-2371.53663", "2.78716", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1622", "103", "8221.6", "2479.7", "-2357.00357", "3.59217", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1641", "104", "8221.6", "2479.7", "-2342.46652", "1.84659", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1625", "105", "8221.6", "2479.7", "-2327.93846", "3.28034", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1627", "106", "8221.6", "2479.7", "-2313.41041", "2.15115", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1645", "107", "8221.6", "2479.7", "-2298.88535", "6.88525", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1626", "108", "8221.6", "2479.7", "-2284.3563", "4.70708", "0", "0"
"01/06/2007 01:00:00.000000000", "01/06/2007 01:12:00.000000000", "1643", "109", "8221.6", "2479.7", "-2269.83924", "5.74915", "0", "0"
    
```

Fig. 20 – Exemple du fichier renseignant sur les positions des OM en fonction du temps. De gauche à droite, la date et l’heure de début, la date et l’heure de fin, l’adresse de l’OM, les positions absolues X, Y et Z, et les trois angles utilisés dans les calculs des coordonnées finales.

Résultats et conclusion

La figure 21 montre quelques résultats obtenus : il s’agit des écarts entre les paramètres calculés dans les deux cas.

Les écarts présentent un pic autour de 0. Dans le cas de dx , on observe également un pic autour d’environ 1,9 m. Les valeurs s’étalent sans grande incidence jusqu’à 1 m environ pour la plus grande partie d’entre eux dans le cas de dy .

L’influence de $T0$ est bien prouvée par la figure 22. Celle-ci fait un comparatif des écarts observés entre les deux fichiers, sans et avec $T0$. Lorsque $T0$ est inclus dans les calculs, les écarts sont fortement diminués et se rapprochent de 0.

L'objectif du stage a donc été rempli : les programmes que j'ai réalisé donnent des résultats corrects, que ce soit sur des données simulées ou des données réelles.

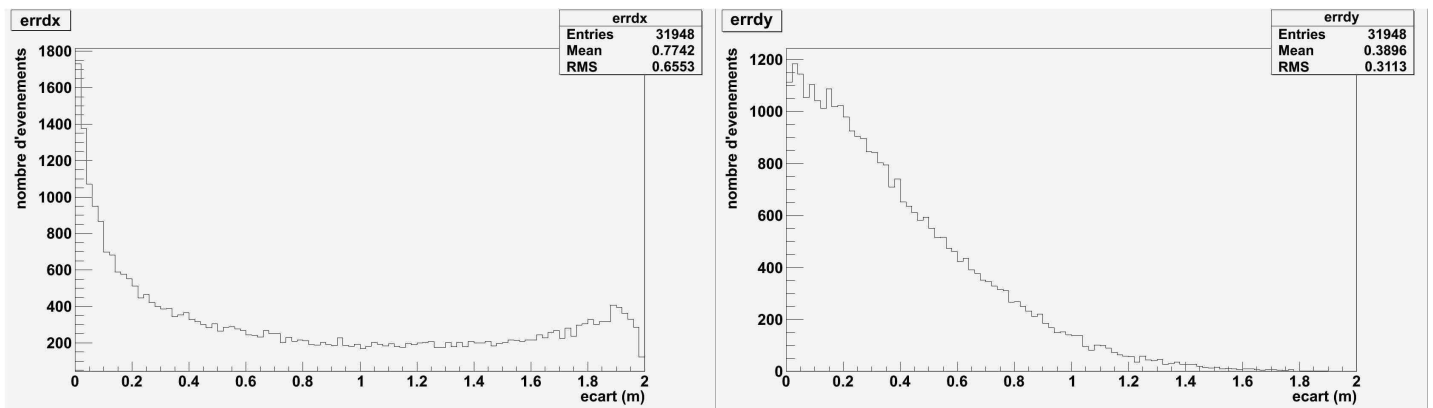


Fig. 21 – Ecart entre les deux fichiers de données, pour dx et dy.

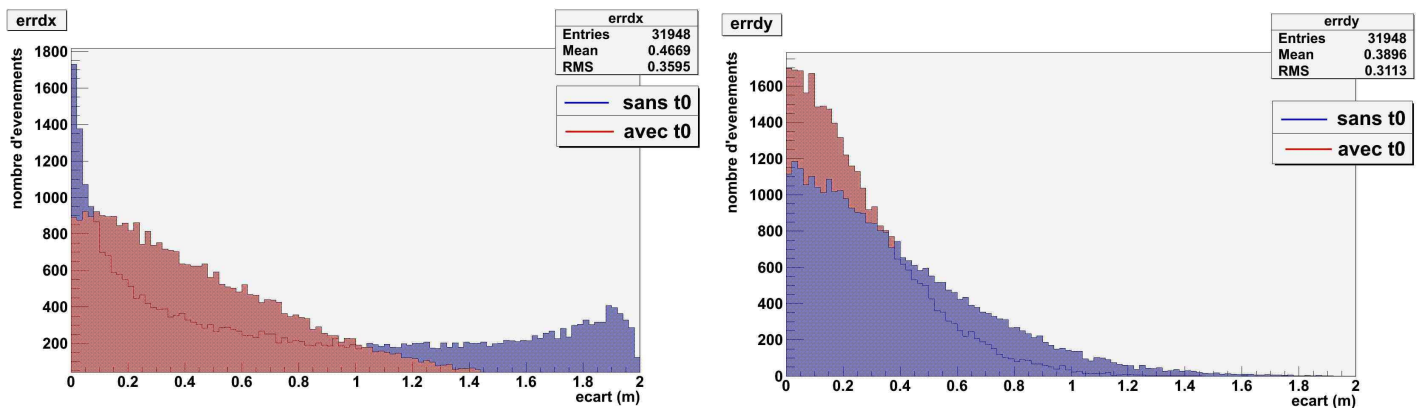


Fig. 22 – Comparaison des écarts obtenus pour dx et dy. En bleu, les écarts sans T_0 . En rouge, T_0 est pris en compte.

CONCLUSION

Ce stage m'aura permis de découvrir une expérience de grande envergure, ANTARES, ayant pour but l'exploration et la découverte de phénomènes provenant de régions éloignées de l'Univers.

J'ai pu découvrir une façon de travailler très différente de mes précédents stages en entreprises. En effet, il s'agit ici d'une expérience européenne, où plusieurs instituts se sont associés afin de mener à terme ce projet nécessitant plusieurs années de développement et de théories. Il m'a fallu comprendre les différentes thèses et documents écrits par la collaboration ANTARES afin de pouvoir remplir correctement ma mission.

Par ailleurs, mes connaissances acquises au cours de mon DUT m'ont été très utiles.

J'ai pu utiliser mes notions de mathématiques et de physique, apprises jusqu'à aujourd'hui, en dehors du cadre scolaire, appliquées sur un sujet concret, réel. J'ai ainsi pu voir une utilisation concrète de ce qui nous est enseigné dans ces matières.

J'ai aussi principalement développé mes connaissances en programmation C. De même, j'ai découvert le langage C++ et un programme de statistiques, « ROOT », notamment utilisé par le CERN. Ces nouveautés ont nécessité des périodes d'apprentissages plus ou moins compliquées mais profitables. Grâce à cela, j'ai pu étudier les influences de paramètres à l'aide d'outils de statistiques et corriger convenablement mes programmes.

Les documentations et les thèses consultées étaient le plus souvent en anglais, en raison du caractère européen d'ANTARES. J'ai ainsi pu approfondir mon vocabulaire technique dans cette langue.

SOURCES

- [1] D. Stubert, thèse de doctorat, *Contribution à l'expérience ANTARES : évaluation de site et étude de la sensibilité à la détection de neutralinos* (2002)
- [2] S. Ferry, thèse de doctorat, *Étude de la production de neutrinos associés aux Sursauts Gamma dans le modèle du Boulet de Canon. Possibilité d'observation de ces neutrinos par le détecteur ANTARES, et étude du bruit de fond optique enregistré par le prototype d'un secteur de ligne* (2004)
- [3] A. Heijboer, thèse de doctorat, *Track Reconstruction and Point Source Searches with ANTARES* (2004)
- [4] A. Heijboer, Antares-Soft, *An algorithm for track reconstruction in ANTARES* (2002)
- [5] ANTARES internal notes, *The data acquisition system for the ANTARES neutrino telescope* (2006)
- [6] Y. Becherini, *Offline reconstruction and analysis software v1r6* (2007)
- [7] S. Navas et L. Thompson, *KM3, User Guide and Reference Manual*(1999)
- [8] F. Cassol Brunner, *Gendet 1.2 : Cards and Tags* (2000)
- [9] M. de Jong, *Antares data formats* (2004)
- [10] Site web d'ANTARES : <http://antares.in2p3.fr/>