

Le soleil comme superposition de soleils élémentaires

Pierre-Olivier AMBLARD¹, Saïd MOUSSAOUI², Thierry DUDOK DE WIT³, Jean LILENSTEN⁴, Mathieu KRETSZCHMAR⁵, Jean ABOUDARHAM⁶, Frédéric AUCHÈRE⁷

¹ GIPSA-lab, Dépt. Images et Signal (UMR CNRS 5083)
ENSIEG-BP 46 38402 Saint Martin d'Hères Cedex

²IRCCYN (UMR CNRS 5672)
ECNantes, 1 rue de la No, BP92101, 44321 Nantes cedex 3

³LPCE (UMR CNRS 5672, Université d'Orléans), 3A ave de la recherche scientifique, 45071 Orléans

⁴Physique Bat. D, BP53, 38041 Saint Martin d'Hères Cedex

⁵SIDC, Observatoire Royal de Belgique, Ringlaan 3, 1180 Brussel, Belgium

⁶LESIA, Observatoire de Paris, 5 Pl. Jules Janssen, 92195 Meudon

⁷IAS, Université Paris-Sud, 91045 Orsay
Bidou.Amblard@gipsa-lab.inpg.fr

Résumé – Le but de ce travail est la décomposition de l'activité solaire en des activités élémentaires interprétables physiquement. Nous utilisons les mesures du spectre solaire issues de l'instrument SEE embarqué par le satellite TIMED. Les mesures journalières du spectre VUV obtenues sur cinq années autorisent l'utilisation d'un algorithme MCMC de séparation bayésienne de sources positives. Nous montrons que le spectre solaire peut être décomposé en la somme pondérée de trois spectres élémentaires, deux étant clairement représentatifs de zones différentes du soleil, le dernier étant plus difficilement interprétable.

Abstract – The goal of the work is to decompose the solar activity into some elementary physically meaningful activities. We use the data collected by the instrument SEE on-board TIMED satellite. These data consist in a daily measure of the solar spectrum in the Vacuum Ultra-Violet band. Five years of data allows to apply an MCMC algorithm that implement the separation of positive sources using the Bayesian inference paradigm. We show that three elementary spectra are sufficient to represent the full spectrum. The interpretation of two of these spectra in terms zones of the sun seems possible, whereas the last source is more difficult to interpret.

1 Le spectre solaire

Le but du groupe de travail représenté par l'ensemble des auteurs de cette communication est l'étude des radiations solaires dans l'Extrême Ultra-Violet (EUV), les longueurs d'ondes concernées allant de 20 nm à 200 nm environ. L'étude du spectre solaire dans l'EUV est importante à double titre. Elle est fondamentale bien entendu pour la compréhension de la physique solaire, et importante également pour son rôle dans l'interaction Soleil-Terre. Ces radiations interagissent très fortement avec notre haute atmosphère et induisent des processus physiques jouant directement sur les activités humaines. Par exemple, le flux EUV influe sur la densité de la thermosphère dans laquelle se déplacent de nombreux satellites. Or la densité entre dans la détermination de la force de traînée, et le contrôle des trajectoires de nos satellites dépend ainsi du flux EUV solaire.

Les activités du groupe dans l'étude du spectre solaire ont pris différentes directions. Dans l'une, nous nous intéressons à la reconstruction du spectre à partir de quelques raies fondamentales. En effet, la physique montre que les émissions à une longueur d'onde sont liées aux émissions à d'autres longueurs d'onde. Nous avons caractérisé les liens statistiques existant entre ces différentes émissions sur des données issues de l'instrument SEE à bord du satellite TIMED, et montré que 6 raies fondamentales (parmi qua-

rante environ) permettent de reconstruire le spectre entier. L'intérêt de l'étude est également de montrer que le groupe de 6 raies doit être choisi en fonction des applications envisagées [1, 2]. Ces résultats sont très importants à des fins de mesures, puisqu'il suffit maintenant de mesurer les radiations uniquement dans quelques bandes bien spécifiées pour avoir accès à l'ensemble du spectre EUV. Cette remarque est fondamentale pour la conception d'instruments de mesure embarqués par des satellites. Une autre piste de recherche concerne l'étude des liens entre quelques grandeurs caractéristiques de l'activité solaire ("proxies") et l'irradiance totale reçue sur terre, à des fins de prévision.

L'objet de cette communication concerne la troisième direction suivie par ce groupe de recherches : la décomposition du spectre solaire en une superposition de spectres élémentaires. L'idée sous-jacente est de décomposer l'activité solaire en un nombre restreint d'activités solaires fondamentales, et d'examiner leurs variabilités temporelles.

2 Mesures SEE-TIMED

Le satellite TIMED (Thermosphere Ionosphere Mesosphere Energetics and Dynamics), lancé par la NASA en 2001, a pour objet l'étude de l'environnement planétaire à très haute altitude. Il est doté de plusieurs instruments de

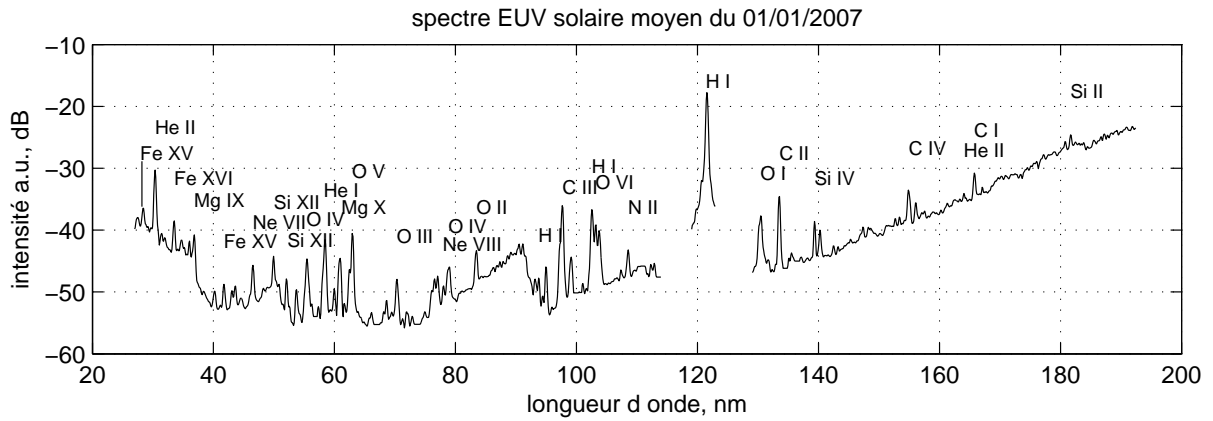


FIG. 1 – Spectre EUV solaire mesuré par l'instrument SEE-TIMED, le 1er janvier 2007. Quelques raies d'émissions sont labellées, le label étant à l'aplomb de la raie (voir la raie du Fer XV à environ 30nm).

mesure, dont l'instrument SEE (Solar Extreme ultraviolet Experiment) composé d'un spectromètre et de plusieurs photomètres. Le but de SEE est de mesurer l'énergie solaire de l'extrême ultra-violet déposée dans notre haute atmosphère (≈ 100 à 400 km). L'instrument observe le soleil 3 minutes par orbite lorsque le soleil est visible. Des traitements de données (moyennes journalières, corrections pour compenser des dégradations, éliminations des grandes éruptions...) permettent d'obtenir une mesure journalière du spectre solaire dans l'extrême ultra-violet, entre 27 nm et 192 nm, une gamme plus large que l'EUV et appelée VUV (pour Vacuum UV). Le spectre du 1er janvier 2007 est montré figure 1. Certaines raies ont pour caractéristique commune de se former dans ou près de la zone de transition, c'est-à-dire dans une région de l'atmosphère solaire où le flux de chaleur est non maxwellien. La température y passe en quelques centaines de kilomètres de $10\,000$ K en haut de la chromosphère à $1,5$ millions de Kelvin en bas de la couronne. Parmi ces raies, on distingue des raies caractéristiques d'émissions d'éléments comme l'hydrogène I (Lyman- α) à $121,6$ nm ou l'hélium II à $30,4$ nm. En général, les raies de cette région sont optiquement minces, c'est-à-dire que leur intensité est directement liée à la température du milieu et à la densité de l'élément chimique concerné, et qu'elles ne subissent pas d'absorption et de réémission le long de leur parcours optique vers l'observateur. Ces raies affichent donc un comportement directement lié aux conditions d'activité de l'atmosphère solaire. Certaines raies correspondent plutôt à des régions chaudes de la zone de transition, d'autres à des régions plus froides, car les transitions ne peuvent se faire que sous des conditions physiques bien définies. Une dernière catégorie de raies a cependant un comportement différent. Il s'agit des raies optiquement épaisses. Ces raies sont résonnantes avec le milieu, ou en recouvrement avec d'autres raies. Lors de leur parcours optique, elles peuvent être absorbées puis réémises, et leur forma spectrale peut être très fortement déformée. Pour comprendre leur comportement, il est nécessaire d'effectuer des calculs de transfert de rayonnement dans l'ensemble de la région de formation, et de prendre en compte les processus collisionnels du milieu. Ce qui explique que ces raies n'aient pas toujours un comportement aussi prévisible que les autres.

Le jeu de données à disposition comporte un spectre moyen journalier du $25/02/2002$ jusqu'au $12/05/2007$, soit 1903 jours. Les mesures sont effectuées sur 1546 longueurs d'ondes réparties uniformément sur la bande d'intérêt, avec toutefois deux coupures vers 115 nm et 125 nm (les données sont donc rangées dans une matrice $I(t, \lambda)$ de taille $(n_t, n_\lambda) = (1903, 1546)$). L'ensemble des données permet d'étudier la variabilité solaire sur 5 ans dans la bande concernée. L'évolution temporelle à différentes longueurs d'onde révèle de fortes similitudes; cette redondance suggère que la variabilité du spectre VUV ne comprendrait qu'un faible nombre de degrés de liberté. D'un point de vue physique, il serait très intéressant de savoir si ces quelques degrés de liberté correspondent à des spectres issus de régions différentes du Soleil (régions actives, calmes, etc.) ou bien permettent d'isoler des processus qui se déroulent à des températures différentes. Une étude récente [3] a en effet montré que la variabilité peut être entièrement reconstruite à partir de la mesure de 6 raies spectrales.

Le but de l'étude est de rechercher dans le spectre solaire des spectres caractéristiques, en nombre réduits, de sorte à décomposer l'activité solaire en caractéristiques élémentaires aisément interprétables physiquement. La première analyse pertinente afin d'extraire les caractéristiques fondamentales est d'effectuer une analyse en composantes principales.

2.1 Analyse en composantes principales

Une analyse en composantes principales et des techniques de clustering nous ont permis dans [1] de montrer que quelques raies suffisent à expliquer la majeure partie du spectre. Le traitement des données valide des idées issues de la physique. Cette analyse est très importante puisque les missions futures d'observation des émissions solaires pourront se limiter à la mesure de quelques longueurs d'onde particulières, à partir desquelles l'ensemble du spectre pourra être évalué. Ici, nous présentons rapidement l'analyse en composantes principales pour illustrer le nombre de caractéristiques pertinentes et mettre en évidence la variabilité temporelle de ces caractéristiques.

Nous montrons sur la figure 2 les évolutions temporelles des trois premières composantes principales. Si les compo-

santes 1 et 2 présentent des allures attendues (oscillations de 27 jours, croissance du cycle solaire, . . .) la composante 3 montre un comportement étrange, qui a pu être identifié comme un défaut de l'instrument de mesure. Le mode 3 exprime une modulation périodique du spectre, de période exactement deux mois. Cette modulation est donc une caractéristique de l'instrument, et a pu être associée aux variations de température que subit le satellite lorsque celui-ci se trouve dans l'ombre de la terre. Cette figure est placée ici pour illustrer la nécessité de corriger les données acquises. Ces corrections incluent l'élimination de l'effet des variations de température ou encore de quelques longueurs d'onde défectueuses.

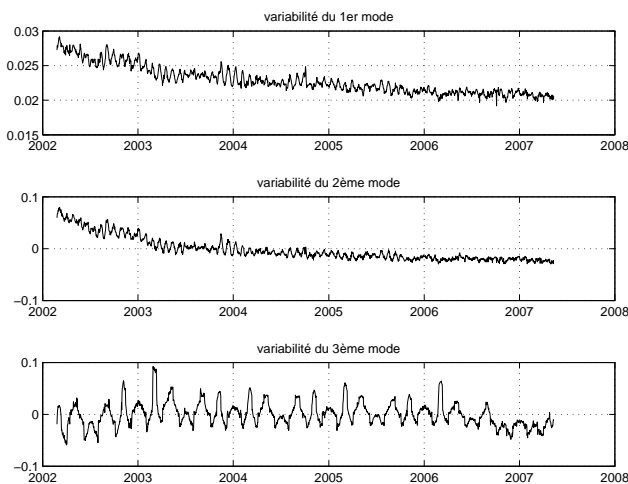


FIG. 2 – Variabilité des trois premiers modes orthogonaux calculés par SVD. Les deux premiers modes ont les variabilités attendues (décroissance de l'activité dans le cycle de 11 ans, période de 27 jours),. Le troisième mode révèle des problèmes de l'instrument du à des variations de températures.

Après corrections, une nouvelle analyse en composantes principales montre qu'un faible nombre de composantes permet d'expliquer la majeure partie du spectre. Toutefois, l'interprétation des termes principaux comme spectres laisse à désirer puisque rien n'assure dans la décomposition en valeurs singulières que les modes principaux restent positifs. Nous devons avoir recours à une autre décomposition assurant la positivité des modes principaux, et donc rendant possible leur interprétation en terme de spectre.

2.2 Séparation en composantes positives

La séparation en composantes positives consiste à chercher une factorisation de la matrice positive¹ en le produit de deux matrices positives, c'est-à-dire $\mathbf{I}(t, \lambda) = \mathbf{V}(t) \times \mathbf{S}(\lambda)$ où les matrices $\mathbf{V}(t)$ et $\mathbf{S}(\lambda)$ sont de tailles respectives (n_t, n_c) et (n_c, n_λ) , n_c représentant le nombre de composantes "principales" positives que l'on cherche. Ce problème de factorisation en matrices positives peut également être interprété comme un problème de séparation

¹Dans tout le papier, sauf mention explicitement contraire, le terme matrice positive concerne une matrice dont tous les termes sont positifs ou nuls, et n'a rien à voir avec la notion usuelle de définie-positivité.

de sources positives. Une telle séparation de sources est ici justifiée par le fait que le spectre mesuré résulte de la superposition linéaire et instantanée (non convolutive) de contributions provenant de différentes régions du Soleil. Cette justification, théoriquement correcte pour les raies optiquement minces, est seulement heuristique en ce qui concerne les raies optiquement épaisses.

La matrice $\mathbf{S}(\lambda)$ contient les spectres principaux alors que la matrice $\mathbf{V}(t)$ contient les variabilités temporelles de ces spectres. De nombreuses méthodes apportent des solutions pratiques au problèmes de factorisation en matrices positives.

Nous adoptons ici l'approche bayésienne développée dans [4], et dont nous rappelons les étapes fondamentales ici. On suppose que les mesures effectuées sont entachées d'erreurs et de bruits rangés dans une matrice $\mathbf{B}(t, \lambda)$, de sorte que le problème à résoudre est, étant donnée la matrice $\mathbf{I}(t, \lambda) = \mathbf{V}(t) \times \mathbf{S}(\lambda) + \mathbf{B}(t, \lambda)$, estimer les matrices \mathbf{V} et \mathbf{S} . L'inférence bayésienne consiste à utiliser un modèle probabiliste *a priori* $P(\mathbf{V}, \mathbf{S})$ sur les matrices à estimer, puis à combiner l'*a priori* avec l'équation de mesure (vraisemblance) à l'aide de la loi de Bayes pour obtenir la mesure *a posteriori*

$$P(\mathbf{V}, \mathbf{S} | \mathbf{I}) \propto P(\mathbf{I} | \mathbf{V}, \mathbf{S}) P(\mathbf{V}, \mathbf{S})$$

Cette mesure résume les connaissances apportées par l'observation sur les paramètres à estimer. A elle seule, la connaissance de la mesure *a posteriori* est donc une solution au problème d'inférence, solution toutefois difficilement exploitable. Il est en général préférable d'en déduire des estimateurs (bayésiens) comme sa moyenne, dont on montre qu'elle constitue l'estimateur de variance minimale, où la position de son maximum.

Le problème consiste à évaluer la mesure *a posteriori* puis à calculer les estimateurs qui en découlent. Ces calculs font en général intervenir l'évaluation d'intégrales n'ayant pas de solutions analytiques, et le recours aux techniques numériques est nécessaire. Parmi elles, les méthodes de simulation Monte-Carlo ont une place de choix, notamment grâce au développement des puissances de calculs numériques. La flexibilité des approches bayésiennes alliée aux approches numériques Monte-Carlo ont permis ces dernières années des développements considérables dans les problèmes d'inférence difficiles.

Dans le problème qui nous occupe, la vraisemblance $P(\mathbf{I} | \mathbf{V}, \mathbf{S})$ s'obtient facilement de l'équation de mesure $\mathbf{I}(t, \lambda) = \mathbf{V}(t) \times \mathbf{S}(\lambda) + \mathbf{B}(t, \lambda)$ en supposant les composantes de la matrice \mathbf{B} conjointement gaussiennes et indépendantes. La puissance de ces composantes est inconnue et constitue un paramètre à estimer également. Pour alléger la présentation, nous négligeons ici ce paramètre de puissance, sachant que les procédures pour son estimation sont développées dans [4]. Le choix du modèle *a priori* doit être guidé par notre connaissance physique. Il est dans un premier temps légitime de considérer les matrices \mathbf{V} et \mathbf{S} indépendamment l'une de l'autre, de sorte que $P(\mathbf{V}, \mathbf{S}) = P(\mathbf{V})P(\mathbf{S})$. En effet, \mathbf{V} dépend de l'évolution temporelle du champ magnétique solaire tandis que \mathbf{S} dépend des caractéristiques atomiques intrinsèques de la matière. Ensuite, comme nous cherchons des matrices posi-

tives, la mesure de probabilité associée à ces matrices rendra une mesure nulle pour les matrices avec au moins un coefficient négatif. Concrètement, le choix adopté consiste à considérer tous les éléments d'une matrice comme étant indépendants, et à leur conférer une loi Gamma comme densité de probabilité *a priori*. Un algorithme de simulation Monte-Carlo par chaîne de Markov est alors mis en œuvre. Il repose principalement sur un échantillonneur de Gibbs dont les principales étapes sont détaillées dans [4]. L'algorithme génère donc des échantillons selon une chaîne de Markov dont la densité stationnaire est précisément $P(\mathbf{V}, \mathbf{S} | \mathbf{I})$. Les échantillons sont alors utilisés (en oubliant une période de convergence) pour estimer les estimateurs empiriques de la moyenne *a posteriori* des spectres et des mélanges.

3 Des Soleils élémentaires ?

Les figures (3) et (4) montrent le résultat de l'application de la séparation de sources positives aux données SEE-TIMED corrigées. Pour l'analyse des résultats, nous avons superposé les éléments responsables des principales raies sur le tracé du spectre moyen, panneau haut. Cette procédure est itérée sur les trois sources extraites. Quelques

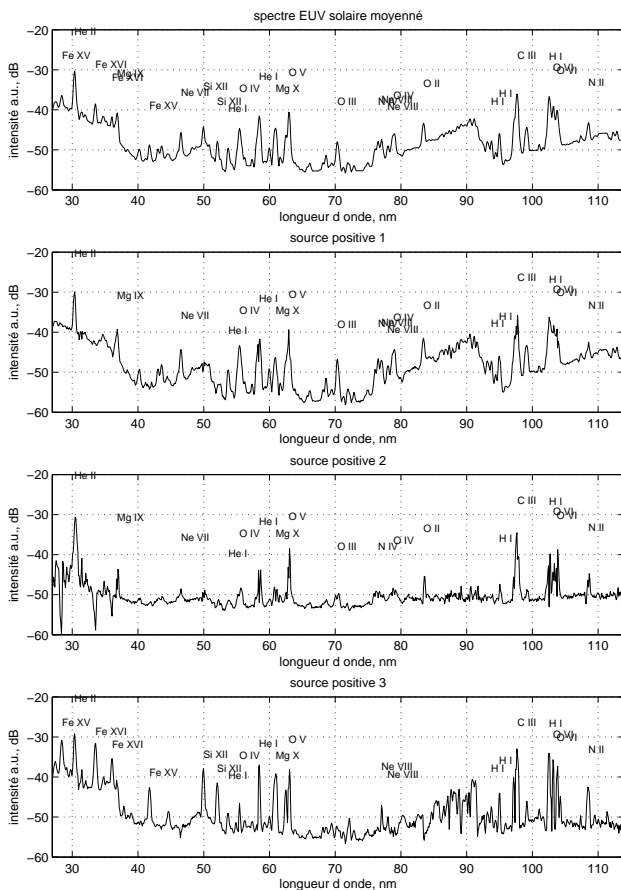


FIG. 3 – Spectre moyen (panneau supérieur) et trois spectres élémentaires obtenus par la séparation bayésienne. La bande est ici limitée à environ 28-113nm.

résultats remarquables sont à noter. Le premier concerne les raies du fer et du silicium qui se retrouvent uniquement dans la source 3. Ces éléments qui interviennent

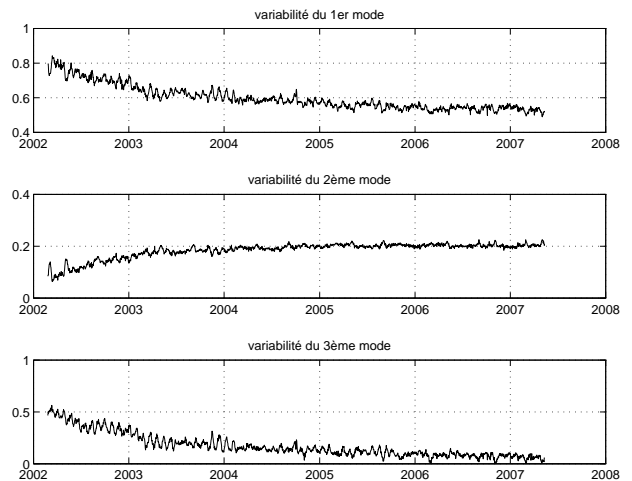


FIG. 4 – Variabilités des sources positives élémentaires.

avec un degré d'ionisation élevé définissent les raies dites "chaudes". La source positive 2 a pour composantes les plus marquées des raies à degré moyen d'ionisation (comme O V à environ 63 nm). Enfin, dans la source 1, les raies à degré faible et moyen d'ionisation prédominent.

La source 3 est dans notre interprétation essentiellement relative aux sources chaudes, alors que la source 1 correspond à des sources plus froides. La source 2 semble être un compromis entre les deux, mélangeant des éléments froids et tièdes. La variabilité temporelle, figure (4), indique que la proportion du spectre composé de raies froides est plus importante que celle du spectre composée de raies chaudes. On note également une tendance inverse pour la source 2. Cette troisième source ressemble à une sorte de mélange qu'il faut comprendre. Cependant, cette source ne peut se réduire en une combinaison linéaire des sources 1 et 2 car son comportement temporel est différent. S'agit-il de raies qui ont un comportement atypique (optiquement épaisses, par exemple), ou dont la région de formation peut être très variable (froide et dense ainsi que chaude et ténue) ?

La discussion reste ouverte et peut permettre aux physiciens solaires d'adopter une approche nouvelle de la compréhension de l'activité solaire.

Références

- [1] T. DUDOK DE WIT, J. LILENSTEN, L. ABOUDARAM, P. O. AMBLARD, and M. KRETZSCHMAR. Retrieving the solar euv spectrum from a reduced set of spectral lines. *Annales Geophysicae*, 23 :3055–3069, 2005.
- [2] J. LILENSTEN, T. DUDOK DE WIT, P. O. AMBLARD, M. KRETZSCHMAR, L. ABOUDARAM, and F. AUCHERE. How to choose an observed set of solar lines for aeronomy driven applications. *Annales Geophysicae*, to appear, 2007.
- [3] J. LILENSTEN, T. DUDOK DE WIT, M. KRETZSCHMAR, P. O. AMBLARD, S. MOUSSAOUI, L. ABOUDARAM, and F. AUCHERE. Retrieving the euv solar spectrum from a selected set of lines for space weather purposes : a review of theories, models and experiments. *Annales Geophysicae*, to appear, 2007.
- [4] S. MOUSSAOUI, D. BRIE, A. MOHAMMAD-DJAFARI, and C. CARTERET. Separation of non-negative mixture of non-negative sources using a Bayesian approach and mcmc sampling. *IEEE Transactions on Signal Processing*, 54(11) :4133–4145, Nov 2006.