

VAL-U-SUN

Valorization of 70 years Solar Observations from the Royal Observatory of Belgium

Contrat BR/165/A3/VAL-U-SUN

RESUME

Contexte

Le nombre de taches solaires est la plus longue expérience scientifique encore en cours et une référence cruciale pour étudier l'activité solaire, la météo spatiale ainsi que le changement climatique. L'Observatoire Royal de Belgique (ORB) joue un rôle central dans la poursuite de cette expérience, car il héberge le World Data Center Sunspot Index et Long-term Solar Observations (WDC-SILSO : <https://wwwbis.sidc.be/silso/>). Ce centre mondial de données a pour objectif de collecter des données solaires, ainsi que de produire et de diffuser le nombre international de taches solaires, qui est utilisé dans environ deux cents publications scientifiques chaque année.

Les taches solaires sont des taches sombres apparaissant en groupes sur la surface solaire. Elles sont la manifestation du magnétisme solaire et sont recensées depuis l'invention du télescope au début du XVIIe siècle. En 1843, Samuel Heinrich Schwabe (Schwabe, 1844) a découvert que les séries chronologiques des comptes de taches solaires présentaient une périodicité d'environ 11 ans. On la connaît aujourd'hui sous le nom de cycle solaire, et il affecte profondément l'ensemble du système solaire. Le champ magnétique solaire contenu dans les taches solaires est la force motrice de la variabilité solaire qui influence quotidiennement l'environnement spatial de la Terre. Au milieu du 19e siècle, le Dr Wolf de l'Observatoire de Zürich a créé un indicateur de l'activité solaire en additionnant le nombre total de taches solaires à 10 fois le nombre total de groupes de taches solaires. Avec ses 400 ans de données, ce « nombre de taches solaires » est la plus ancienne série d'observations de l'activité solaire et s'est avéré remarquablement utile, non seulement pour la physique solaire, mais dans tous les domaines scientifiques. La corrélation avec d'autres indices permet de reconstituer la variabilité solaire influençant le climat terrestre au cours des derniers siècles. Le nombre de taches solaires et le cycle solaire sont également une « pierre de Rosette » pour comprendre les étoiles variables de type solaire. La bibliothèque numérique « The NASA Astrophysics Data System » recueille plus de 70 articles arbitrés par an contenant la phrase exacte « sunspot number » dans le résumé ou dans le titre.

Au fil des siècles, l'ORB a joué un rôle essentiel dans l'étude du nombre de taches solaires. En 1939 et pendant la Seconde Guerre mondiale, l'ORB a lancé un programme d'observation solaire en collaboration avec les successeurs du Dr Wolf à l'Observatoire de Zürich qui est toujours en cours aujourd'hui. Les dessins solaires originaux, obtenus à l'instrument ROB Uccle Solar Equatorial Table (USET), ont été conservés et scannés et forment une série homogène de dessins ininterrompus depuis mars 1940. C'est l'une des plus longues collections d'observation stable du Soleil encore active. En 1981, l'Observatoire de Zürich cessa l'activité du World Data Center et en accord avec l'organisation astronomique internationale « COSPAR », le World Data Center fut transféré à l'ORB. A ce titre, le World Data Center « SILSO » est chargé de collecter les observations et de compter le nombre de taches solaires à partir d'un réseau international de stations d'observation et de produire à partir de ces données le nombre international de taches solaires.

Objectifs

Le but de ce projet est de valoriser deux collections de taches solaires conservées à l'ORB, qui, au début du projet VALUSUN n'étaient pas disponibles en ligne et n'étaient pas non plus au niveau des standards scientifiques contemporains.

La première collection est une série de plus de 20.000 dessins de taches solaires numérisés à USET depuis 1940 et la base de données correspondante qui n'a pas encore été contrôlée, c'est-à-dire qu'aucun contrôle de qualité complet n'a pu être effectué et que la base de données est encore incomplète en termes de paramètres extraits. La numérisation de ces dessins a commencé dans le cadre d'un projet européen mais le manque de financement ultérieur nous a empêché d'ajouter des métadonnées essentielles ainsi que de réaliser un contrôle qualité adéquat. Ainsi, il n'a pas été rendu accessible à l'ensemble de la communauté scientifique. Comme cet ensemble de données remonte à 1940, l'exploitation d'un ensemble de données aussi long et complet est extrêmement important pour les études de phénomènes à long terme tels que le climat terrestre, mais aussi afin d'évaluer la qualité des catalogues solaires parallèles.

La seconde collection, la base de données SILSO, contient les nombres de taches et groupes de taches sur le Soleil observés par un réseau mondial depuis 1981 (plus de 530.000 mesures). Il est utilisé sur base mensuelle pour calculer le nombre international de taches solaires. Un seul contrôle est effectué sur chaque station chaque année, et sa base statistique remonte à la création du Nombre de Wolf dans les années 1850. Une supervision complète et cohérente de cette base de données est importante et comme SILSO est le World Data Center pour la détermination de cet indice, il est de notre responsabilité de faire entrer ce nombre international de taches solaires dans le 21^{ème} siècle en exploitant à son plein potentiel cette base de données existante. À cette fin, un effort international a débuté en 2011 et s'est concentré sur les données du passé, du XIX^e siècle à la fin du XX^e siècle. Malheureusement, l'étude réalisée sur cet ensemble de données spécifique de 1981 à aujourd'hui est restée limitée et les normes modernes ne peuvent être atteintes que par l'utilisation de techniques statistiques plus modernes.

L'objectif à atteindre sur nos deux collections est triple : (1) Amener nos bases de données dans le 21^e siècle en ajoutant des métadonnées essentielles, qu'il s'agisse de paramètres supplémentaires ou de techniques d'évaluation et de contrôle qualité. (2) Utiliser la valeur des paramètres solaires enregistrés dans ces bases de données pour répondre aux questions scientifiques d'aujourd'hui et (3) diffuser les collections et la valeur ajoutée à un ensemble prédéfini de publics.

Méthodologie et résultats

La collection de dessins USET a été enrichie avec les zones des groupes grâce au développement d'un logiciel que nous appelons DIGISUN 2.0 (pour VALUSUN, mais avec la main d'œuvre autofinancée de l'ORB). En plus d'être suffisamment convivial pour une exploitation de masse, il est également compatible avec de multiples plateformes. Il a été installé à la station de Locarno (la station pilote du réseau WDC-SILSO) et constitue ainsi une excellente base de comparaison pour les paramètres extraits de la base de données USET. Grâce à la nouvelle base de données de Locarno, nous avons pu réaliser un contrôle qualité en comparant les deux jeux de données sur 2019-2021. Les prochaines étapes consisteront à comparer les paramètres de taches solaires extraits par USET avec d'autres bases de données plus hétérogènes avec des paramètres de taches solaires ne se chevauchant que partiellement. Cette étude a déjà commencé et sera publiée au premier semestre 2022, sous la forme de deux articles : l'article I présentera DIGISUN2.0 et l'article II présentera le contrôle qualité approfondi de la base de données des paramètres des taches solaires USET.

Pour la deuxième collecte de données (c'est-à-dire les nombres de taches solaires de toutes les stations WDC-SILSO), nous avons développé un modèle d'erreur complet pour les nombres de taches solaires dans un cadre multiplicatif. Le modèle décompose les données en un signal physique, commun à différents observateurs, corrompu par trois types d'erreurs, à court terme, à long terme ainsi que pendant les minima solaires. Nous fournissons une analyse complète des

différents termes du modèle, y compris les ajustements paramétriques de leurs distributions. Ce modèle nous permet d'obtenir des estimateurs plus robustes du nombre de taches solaires et de fournir des erreurs pour ces données à chaque instant. Il met également en évidence les écarts à long terme qui se sont produits dans les séries de plusieurs stations d'observation. Bien que spécialement adapté au nombre de taches solaires, le modèle peut également servir de source d'inspiration pour traiter d'autres jeux de données ayant des propriétés similaires. Les résultats de cette partie de la recherche sont publiés dans Mathieu et al. (2019).

Nous avons également développé une procédure de surveillance non paramétrique basée sur des cartes de contrôle et SVM pour détecter efficacement les déviations du nombre de taches solaires au fil du temps. Le schéma est conçu pour fonctionner avec des processus non normalement distribués et auto-corrélés avec des valeurs manquantes. Il fonctionne à différentes échelles et ne nécessite aucune hypothèse paramétrique sur les données pour fonctionner correctement. Cette méthode nous permet d'identifier automatiquement de nombreux écarts dans le nombre de taches solaires, pour la plupart invisibles dans les analyses précédentes et nous aide à trouver les causes profondes de certains changements importants. À la suite de ce projet, ce schéma de contrôle sera mis en œuvre pour surveiller toutes les stations d'observation impliquées dans le processus de comptage, afin d'éviter l'accumulation future d'écarts importants au fil du temps, tels que ceux observés précédemment. Les résultats ont été soumis au Journal of Quality Technology (JQT : <https://asq.org/quality-resources/pub/journal-of-quality-technology>)

L'ensemble des logiciels développés dans le cadre de VALUSUN par la doctorants S. Mathieu, est disponible sur (<https://github.com/sophiano/SunSpot>). L'adaptation au pipeline de réduction WDC-SILSO sera réalisée avant mi-2022 pour une exploitation plus efficace des données. Toutes les données de la base de données WDC-SILSO ont été vérifiées pour les erreurs et les incohérences, et les métadonnées ont été ajoutées et vérifiées chaque fois que possible. En conséquence, les données de USET et du WDC-SILSO ont désormais été soumises à un contrôle qualité et contextualisées, une étape importante pour rendre les données de SILSO-USET FAIR.

Conclusions et Recommendations

L'équipe USET-WDC-SILSO et les stagiaires embauchés grâce au projet ont mis à jour les bases de données tout au long du projet. Cela comprenait le contrôle de la qualité des données, la collecte et l'amélioration des métadonnées, l'ajout d'un paramètre essentiel à la liste existante : la superficie des groupes enregistrée sur les dessins des taches solaires. À cette fin, le développement d'un nouveau logiciel d'analyse des dessins était nécessaire car le précédent s'est avéré inefficace pour déterminer les zones de taches solaires sur les dessins. La base de données USET est désormais accessible sur www.sidc.be/valusun/usetdb/.

Grâce à nos partenaires de l'UCL et à leur expertise, la base de données SILSO (de 1981 à aujourd'hui) a été largement analysée et utilisée, afin d'amener la production du Nombre international de taches solaires dans le 21ème siècle. Une analyse complète des propriétés statistiques du jeu de données a été réalisée et les résultats sont publiés dans Mathieu et al. (2019). Ensuite, une méthode pour surveiller (détecter toute anomalie dans) les données a été développée et soumise au Journal of Quality Technology (JQT : <https://asq.org/quality-resources/pub/journal-of-quality-technology>), et la doctorante Sophie Mathieu a développé un ensemble complet de programmes en Python afin que le World Data Center SILSO puisse désormais appliquer la méthode de détection en temps réel sur les données (<https://github.com/sophiano/SunSpot>). La base de données WDC-SILSO est désormais disponible sur <http://www.sidc.be/valusun/silsodb/>.

Parallèlement à ce travail, un contrôle qualité complet des données d'entrée dans les deux bases de données a été réalisé par l'équipe WDC-SILSO, les métadonnées correspondantes ont été mises à jour et les données ont été rendues conformes au RGPD (https://ec.europa.eu/info/law/law-topic/data-protection_en).

En conclusion, nous tenons à souligner que la préservation, l'amélioration et la mise à disposition des jeux de données du WDC-SILSO et de la station USET sont de la plus haute importance, surtout

à l'heure où l'on se rend compte que préserver la Terre, c'est comprendre l'évolution de son climat. Et ce n'est pas seulement dans le passé, mais aussi dans l'avenir, car les deux ensembles de données continuent de croître avec les observations en cours.

Ces jeux de données contiennent toutes les clés pour lier les observations passées aux observations futures du Soleil (USET observe via des dessins et des CCD), et également pour lier le nombre international de taches passé à son futur. Bien que la station d'observation qui permet à ce « lien » de survivre dans le temps présente des coûts minimes par rapport aux missions spatiales, les observations au sol du Soleil sont rarement au centre des appels scientifiques belges ou européens. Le nombre international de taches solaires est essentiel à la compréhension de l'évolution du rayonnement solaire sur l'atmosphère terrestre, et donc à la compréhension du réchauffement climatique et de ses processus. Cela signifie que la maintenance des collections et des observations associées doit être une priorité lorsque les décisions sont prises aux plus haut niveau. Et comme un fonctionnement à moindre coût est toujours intéressant, l'option au sol doit être largement envisagée, notamment pour des structures comme USET et SILSO qui existent déjà.

Mots clés

Nombre de taches solaires, taches solaires, statistiques, barres d'erreur, surveillance