
Les données EROS-2, un cas d'étude pour l'archivage et la fouille de données

Jean-Baptiste Marquette*¹

¹Institut d'Astrophysique de Paris (IAP) – Université Pierre et Marie Curie (UPMC) - Paris VI :
UM7095, INSU, CNRS : UMR7095 – 98bis, bd Arago - 75014 Paris France, France

Résumé

La collaboration EROS-2 a observé entre juillet 1996 et mars 2003 (télescope dédié MARLY, ESO, Cerro La Silla) des champs stellaires encombrés (Centre galactiques, Bras spiraux, Nuages de Magellan) en vue de détecter la matière noire baryonique sous forme d'objets compacts (MACHOs) générant sur ces lignes de visée des événements de microlentilles gravitationnelles. Avec quelque 1.4 million d'images et de l'ordre de 100 millions de courbes de lumière en deux bandes photométriques larges non standard, cet ensemble unique de données est d'intérêt majeur pour l'étude de diverses populations d'étoiles variables, et au-delà. On montrera quelques exemples de valorisation astrophysique, et plus particulièrement en collaboration avec des projets en cours, comme le survey infrarouge VVV (ESO, télescope VISTA, Cerro Paranal), ou en utilisant des méthodes statistiques permettant de faire "parler" les données bien au-delà du thème original du projet.

La mise à disposition publique de ces données reste un objectif. Il nécessite de qualifier leur qualité, par des traitements automatiques et massifs avec le moins possible d'intervention humaine et visuelle. Ceci concerne particulièrement les images, dont il convient de vérifier la calibration astrométrique, le seeing, et tout autre paramètre à inclure dans leurs critères de qualité au sein d'une base de données.

Ceci est un préalable avant toute production systématique d'une nouvelle photométrie par différence d'images, une méthode qui a prouvé toute son efficacité pour mesurer toute variation d'éclat en champ encombré. C'est alors que pourront être appliquées les méthodes de fouilles d'archives pour la production d'une base de données de variabilité stellaire, avec classification automatique des types de courbes de lumière, mesure des périodes de pulsation, mesure des mouvement propres, etc. De tels exemples seront donnés, avant de conclure que les possibilités actuelles de traitement parallélisé et de stockage sont des facteurs essentiels à leurs production et pérennisation.

*Intervenant