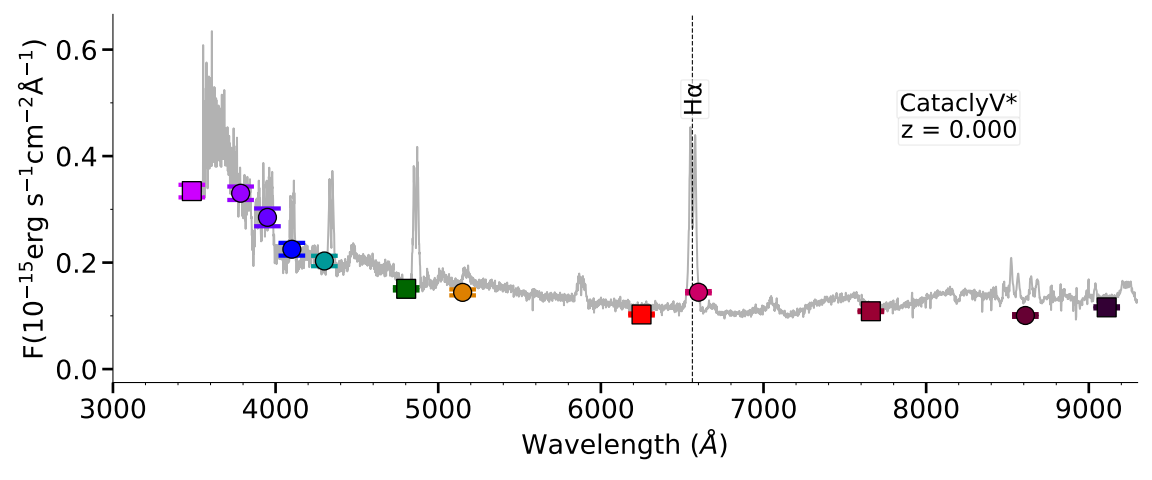
**Mapping H-alpha Excess Candidate Point Sources in the Southern Hemisphere Using S-PLUS Data**

The H-alpha emission line, the first in the Balmer series, is produced by electron transitions from the third to the second energy level in hydrogen. It is widely used in astronomy to identify various objects, including star-forming regions, planetary nebulae (PNe), young stellar objects (YSOs), active stars, and massive stars like Be, Wolf-Rayet (WR), and Luminous Blue Variables (LBVs). It is also observed in interacting binary systems (e.g., symbiotic stars, cataclysmic variables) and extragalactic sources such as QSOs and AGNs.

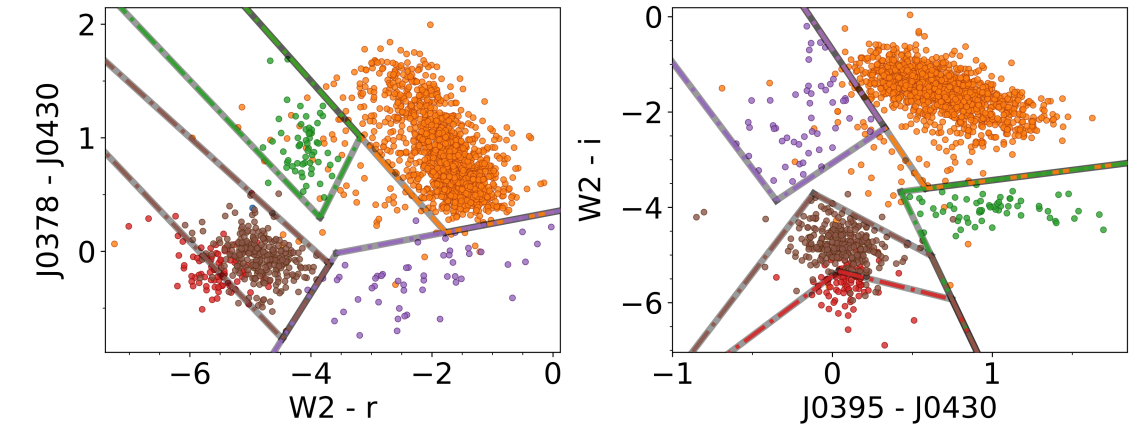


Figure 1. Top panel: SDSS spectrum of an H-alpha emitter (cataclysmic variable) with overlaid S-PLUS photometry. The colored symbols represent S-PLUS photometry in flux units for the following filters (from left to right): u, J0378, J0395, J0410, J0430, g, J0515, r, J0660, i, J0861, and z. Bottom panel: Two new color-color diagrams combining S-PLUS and WISE data, used to distinguish between different populations of H-alpha-excess sources. The colored lines indicate tentative color cuts for separating different classes of H-alpha excess sources.

Using data from the Southern Photometric Local Universe Survey (S-PLUS) Fourth Data Release (DR4), Gutiérrez-Soto et al. (2024, submitted A&A) applied automated color criteria to identify 6,956 H-alpha excess candidates. The (r - J0660) versus (r - i) diagram, with J0660 designed to detect H-alpha emission, was used for selection. Cross-matching with SIMBAD revealed a diverse set of sources, including emission-line stars, YSOs, PNe, binaries, CVs, variable stars, QSOs, AGNs, and galaxies. For extragalactic sources, the J0660 excess originates from redshifted lines rather than H-alpha. The top panel of Figure 1 shows an example of an object selected by our algorithm, clearly exhibiting H-alpha emission.

To separate galactic and extragalactic sources in the H-alpha list, Uniform Manifold Approximation and Projection (UMAP), a non-linear dimensionality reduction technique used in machine learning, was applied, followed by Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), a machine learning algorithm designed to identify clusters of varying densities. This approach not only enabled the separation of galactic and extragalactic sources, but also allowed for the differentiation of various classes of H-alpha excess sources, including cataclysmic variables (CVs), RR Lyrae stars, active galactic nuclei (AGNs), and others

We also constructed new color-color diagrams by combining S-PLUS and WISE data, providing a practical tool to refine the classification of H-alpha sources and distinguish between specific populations (see bottom panel of Figure 1). These diagrams improve our ability to classify a wide range of astrophysical objects, complementing the original (r - J0660) versus (r - i) diagram used to identify H-alpha emitters. Notably, these new diagrams offer a simpler alternative to complex machine learning techniques for classifying H-alpha sources. This work underscores the potential of multiwavelength surveys and machine learning techniques to study H-alpha-related phenomena and guide future spectroscopic follow-up.