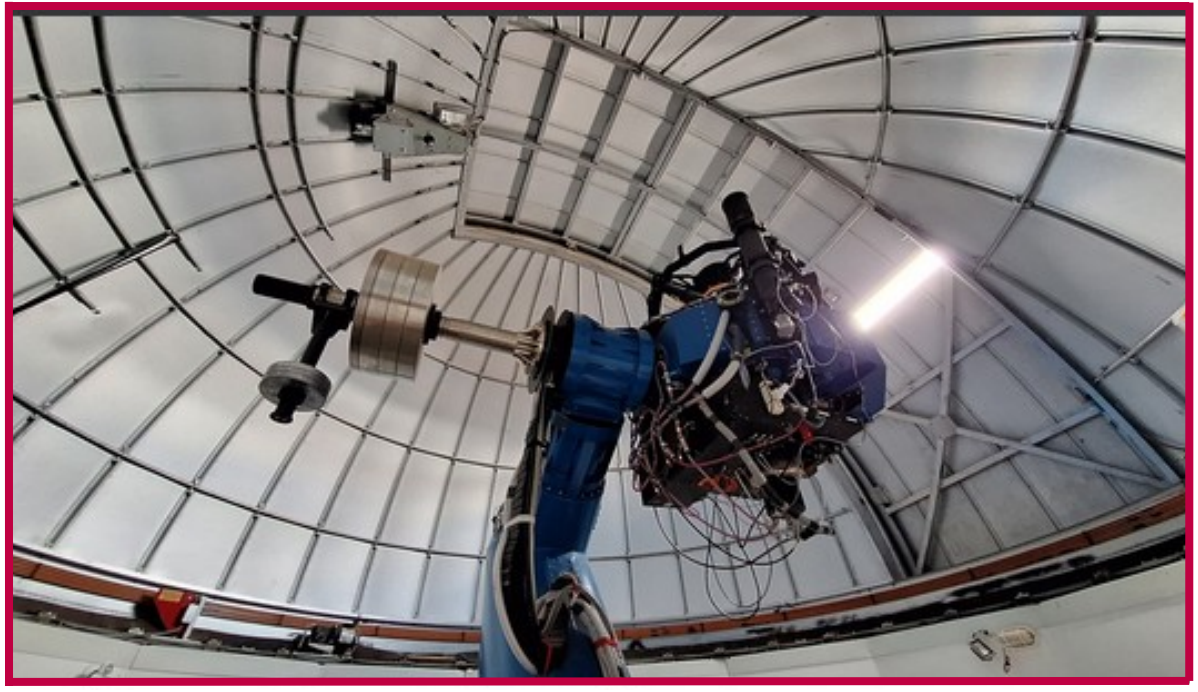
Unveiling the Hidden Population of Hα Excess Sources in the Southern Sky: A Synergy of S-PLUS Photometry and Machine Learning

Luis A. Gutiérrez-Soto
Instituto de Astrofísica de La Plata - CONICET-UNLP

XIV Edition of FRIENDS OF FRIENDS Meeting
7 to 11 April 2025

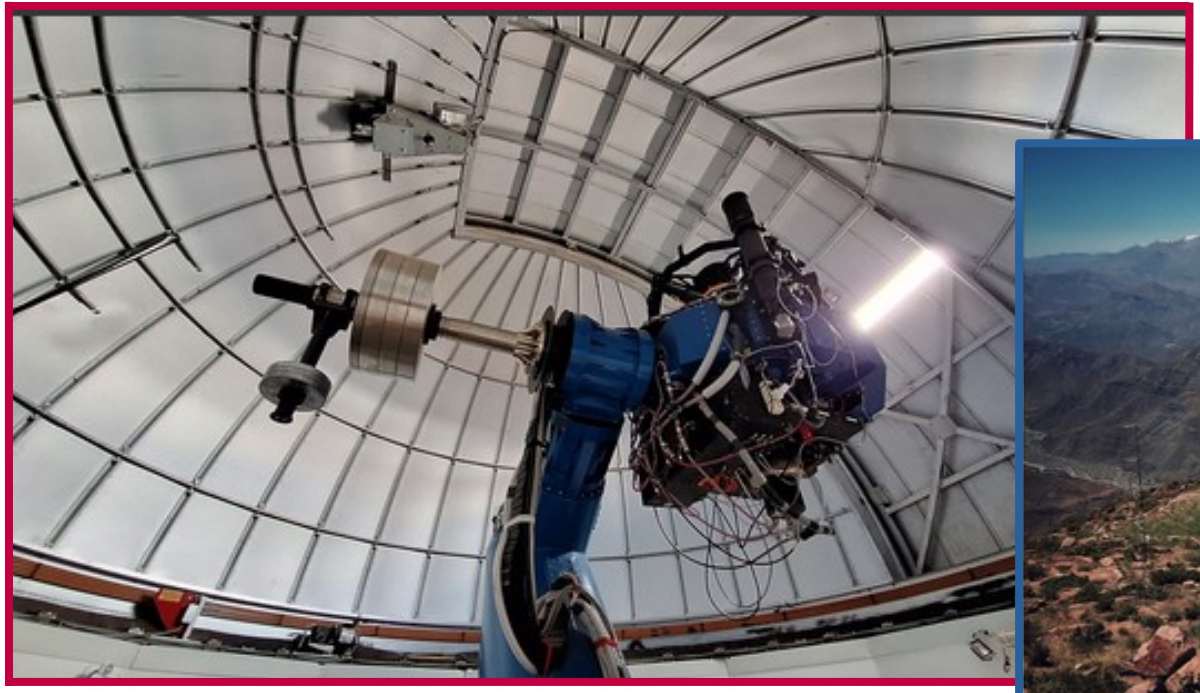
Context

The Southern Photometric Local Universe Survey (S-PLUS; Mendes de Oliveira et al. 2019) *is an international collaboration led by Brazil, Chile, and Spain, with contributions from Argentina, and other countries. It aims to map ~9300 square degrees of the Southern Sky using the* ***T80-South (T80-Sul)****, an 83-cm robotic telescope located at Cerro Tololo Inter-American Observatory (CTIO), Chile.*



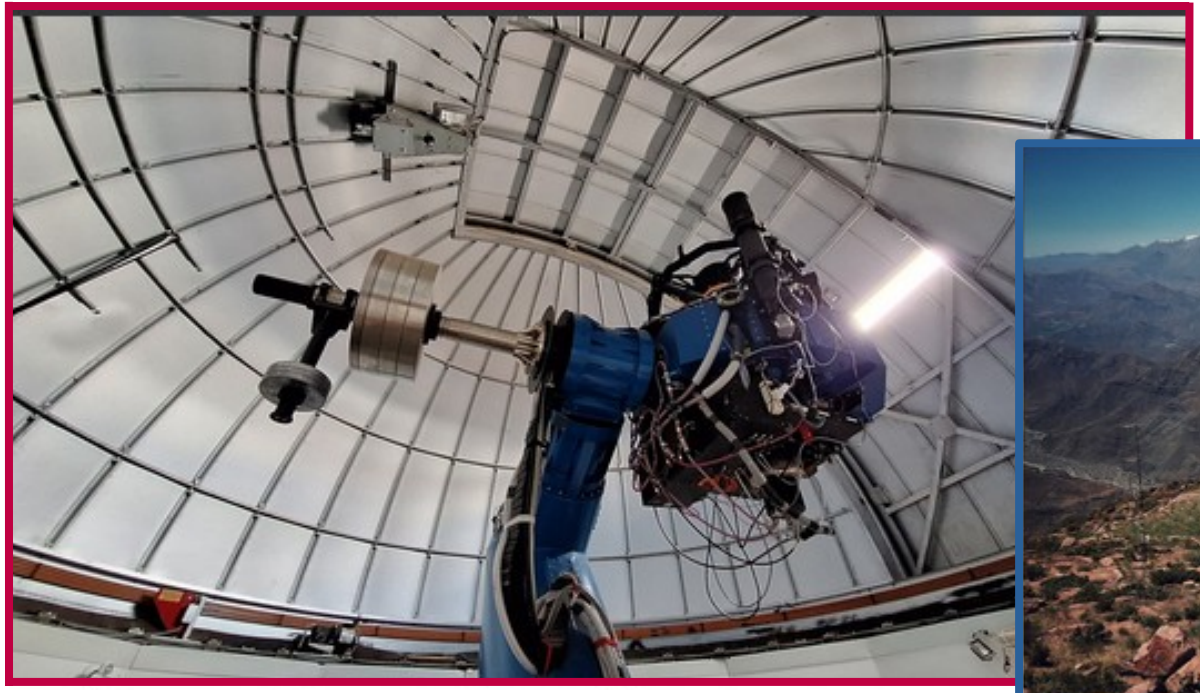
Context

The Southern Photometric Local Universe Survey (S-PLUS; Mendes de Oliveira et al. 2019) *is an international collaboration led by Brazil, Chile, and Spain, with contributions from Argentina, and other countries. It aims to map ~9300 square degrees of the Southern Sky using the* ***T80-South (T80-Sul)****, an 83-cm robotic telescope located at* ***Cerro Tololo Inter-American Observatory (CTIO), Chile.***



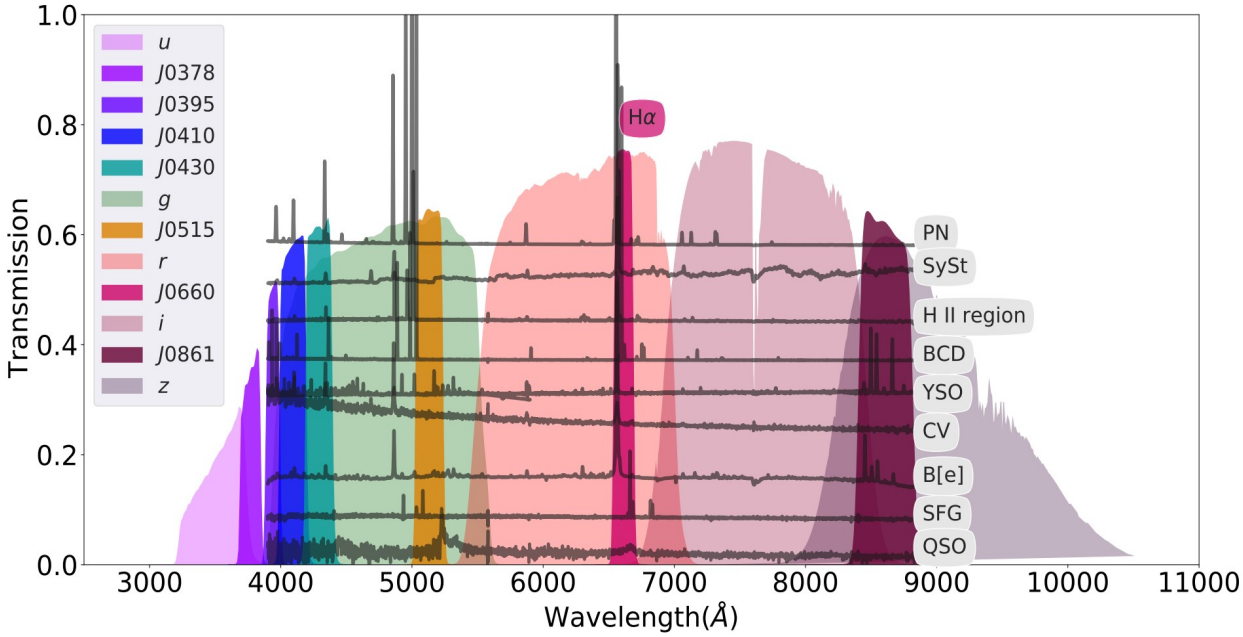
Context

The Southern Photometric Local Universe Survey (S-PLUS; Mendes de Oliveira et al. 2019) *is an international collaboration led by Brazil, Chile, and Spain, with contributions from Argentina, and other countries. It aims to map ~9300 square degrees of the Southern Sky using the* ***T80-South (T80-Sul)****, an 83-cm robotic telescope located at* ***Cerro Tololo Inter-American Observatory (CTIO), Chile.***



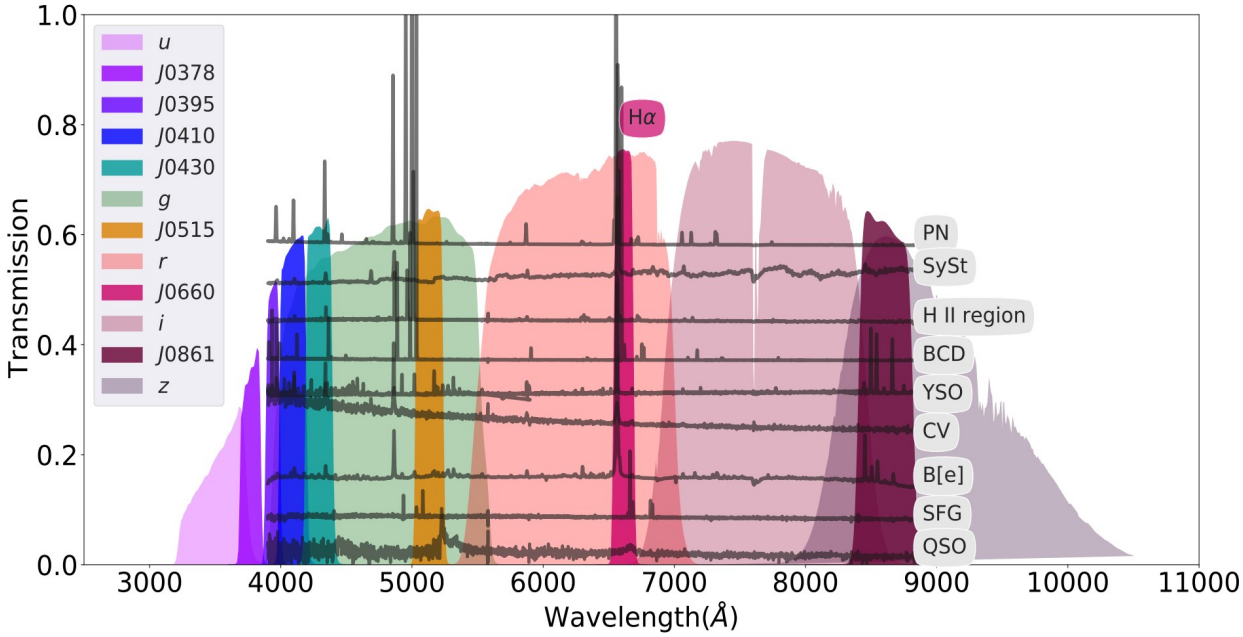
Context

*The T80-Sul is equipped with a 12-filter system (five broadband filters: u, g, r, i, z, and seven narrowband filters, including **J0660: Hα**), enabling precise photometric studies of stars, galaxies, and transient phenomena. S-PLUS has already released Data Release 4 (DR4), covering approximately 3000 square degrees of the southern sky, which include high-latitude regions and an area within the disk.*



Context

*The T80-Sul is equipped with a 12-filter system (five broadband filters: u, g, r, i, z, and seven **narrowband filters**, including **J0660: Hα**), enabling precise photometric studies of stars, galaxies, and transient phenomena. S-PLUS has already released Data Release 4 (DR4), covering approximately 3000 square degrees of the southern sky, which include high-latitude regions and an area within the disk.*



J0378: [O II]

J0395: Ca II H+K

J0410: Hδ

J0430: G band

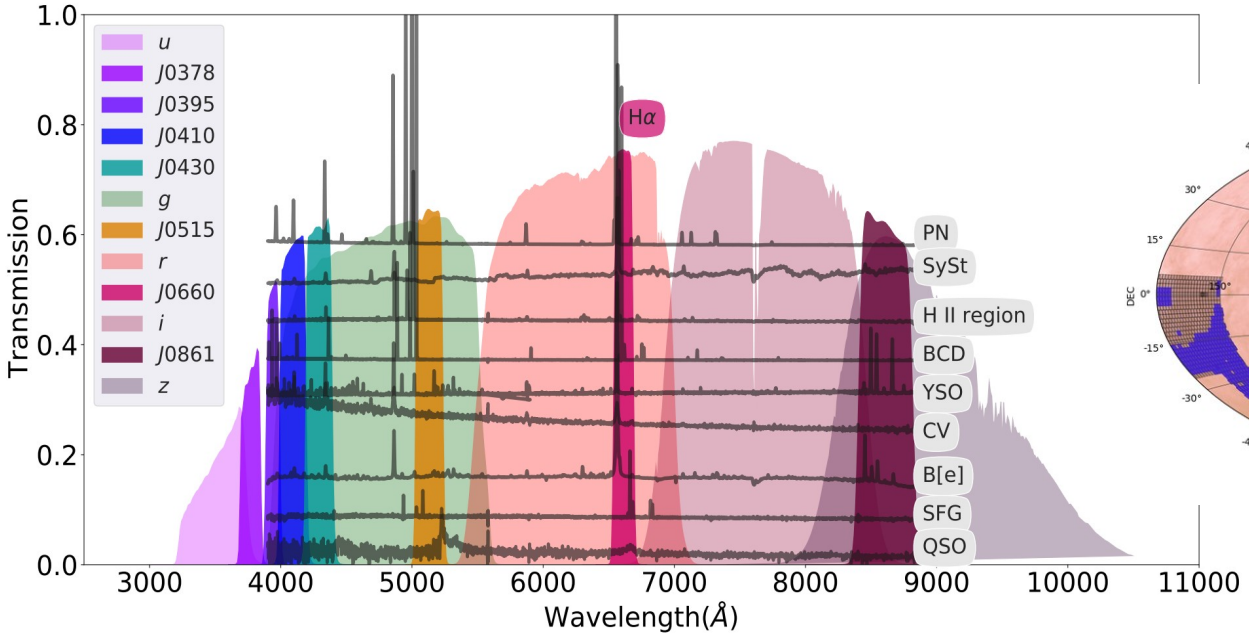
J0515: Mg b triplet

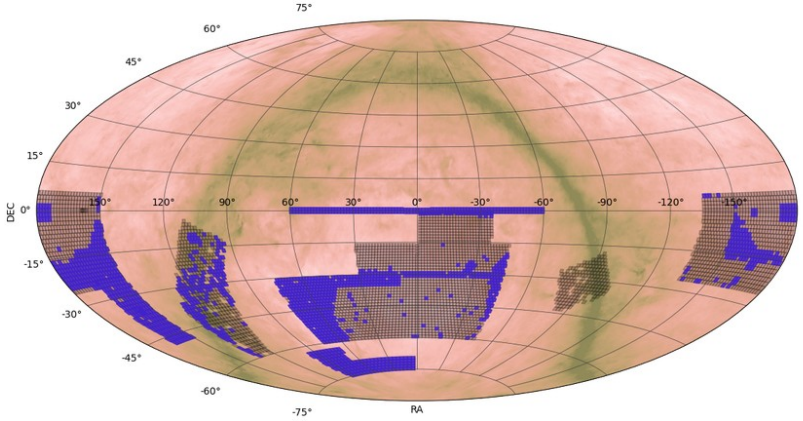
J0660: Hα + [N II]

J0865: Ca Triplet

Context

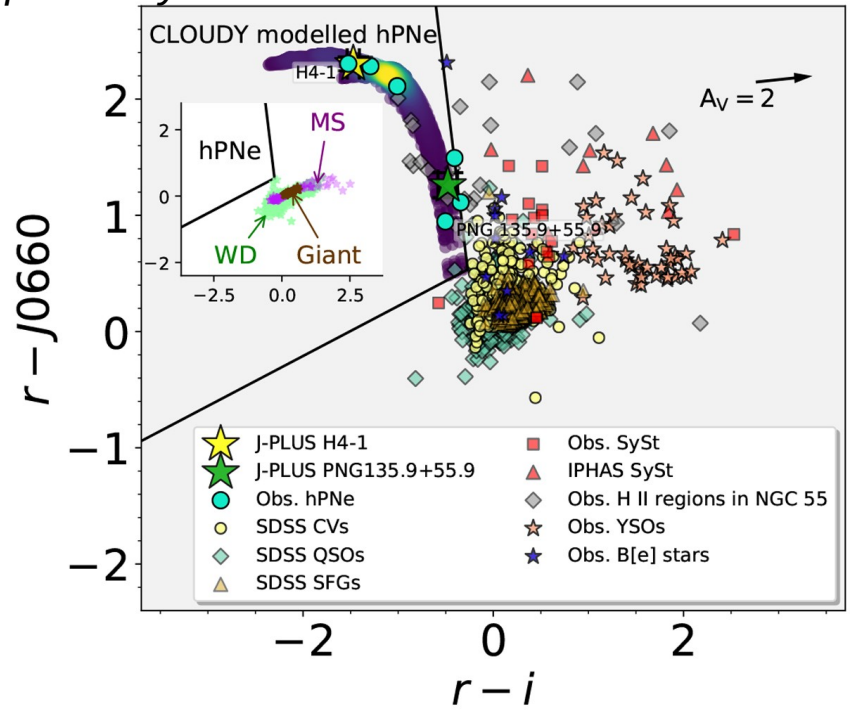
*The T80-Sul is equipped with a 12-filter system (five broadband filters: u, g, r, i, z, and seven narrowband filters, including **J0660: Hα**), enabling precise photometric studies of stars, galaxies, and transient phenomena. S-PLUS has **already released Data Release 4 (DR4), covering approximately 3000 square degrees** of the southern sky, which include high-latitude regions and an area within the disk.*





Photometric Tools for Identifying Compact Planetary Nebulae

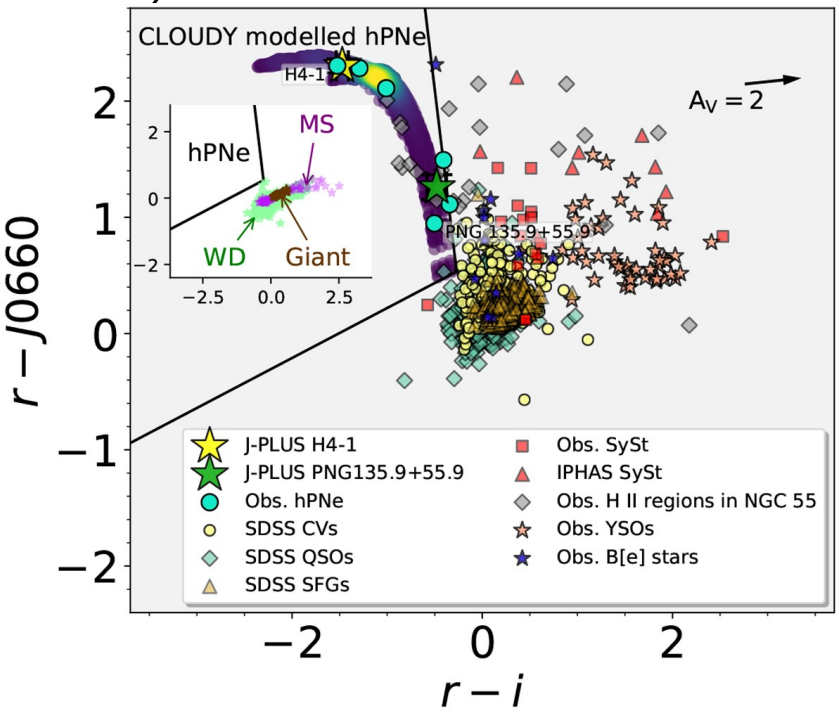
This study uses color–color diagrams, such as r – J0660 vs. r – i, alongside photometric spectra to effectively distinguish compact planetary nebulae from other emission-line objects. The approach, validated with J-PLUS and S-PLUS data, highlights its efficiency in searching for halo planetary nebulae.

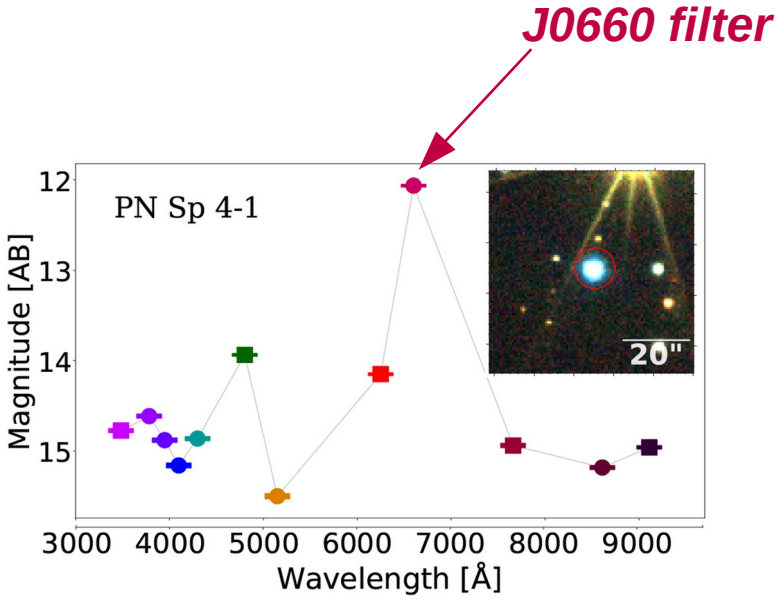


Gutiérrez-Soto, Gonçalves, Akras, Cortesi, Lopez-Sanjuan, et al. (2020, A&A, 633, A123)

Photometric Tools for Identifying Compact Planetary Nebulae

This study uses color–color diagrams, such as r – J0660 vs. r – i, alongside photometric spectra to effectively distinguish compact planetary nebulae from other emission-line objects. The approach, validated with J-PLUS and S-PLUS data, highlights its efficiency in searching for halo planetary nebulae.

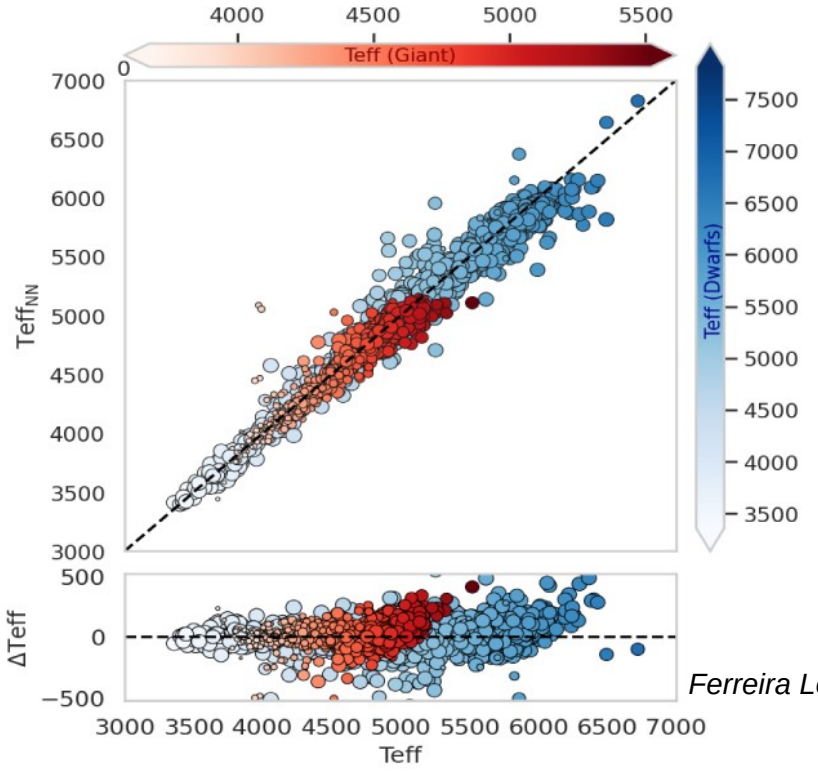




Gutiérrez-Soto, Gonçalves, Akras, Cortesi, Lopez-Sanjuan, et al. (2020, A&A, 633, A123)

Decoding the Milky Way's Stellar Populations

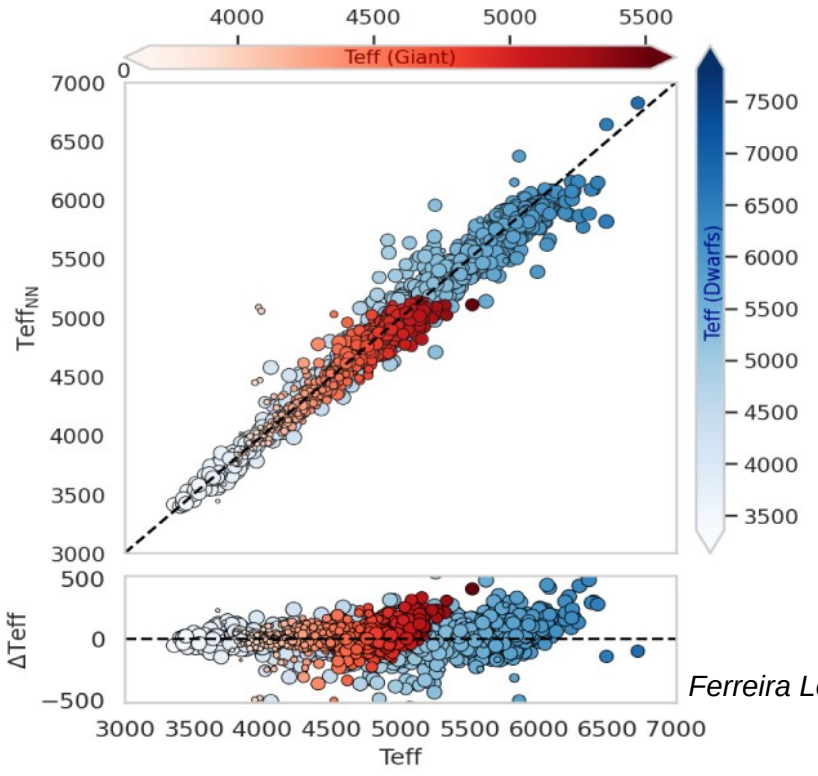
Optimized machine learning algorithms to derive stellar parameters (T_eff, log g, [Fe/H]) and chemical abundances for millions of stars.

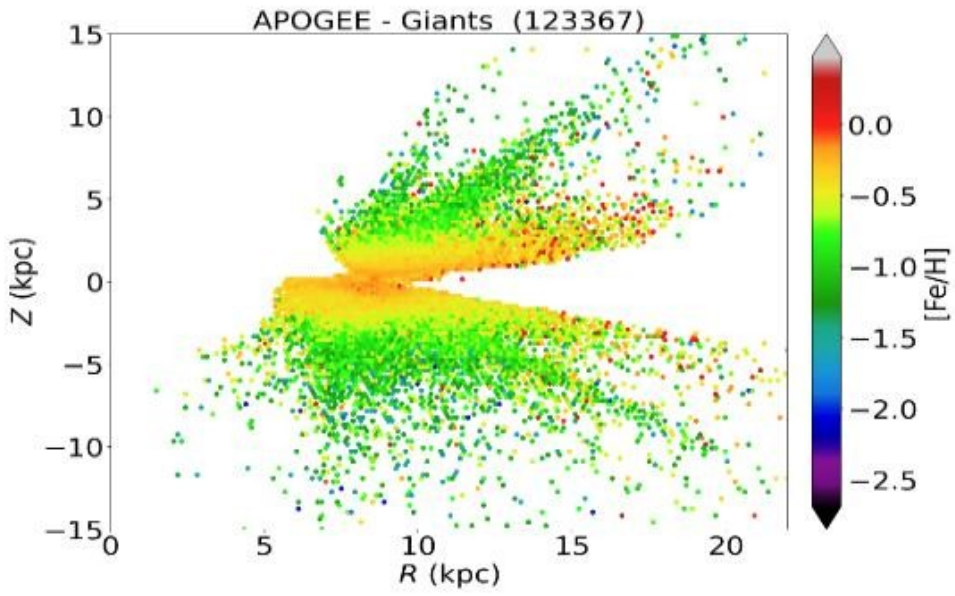


Ferreira Lopes, Gutiérrez-Soto, Ferreira Alberice, Monsalves, Hazarika, et al. (2025, A&A, 693, A306)

Decoding the Milky Way's Stellar Populations

Optimized machine learning algorithms to derive stellar parameters (T_eff, log g, [Fe/H]) and chemical abundances for millions of stars. 🌟 Large-scale metallicity map of the Southern Milky Way ([Fe/H] spanning −2.5<[Fe/H]<+0.5). 🌌 Revealed a radial metallicity gradient.

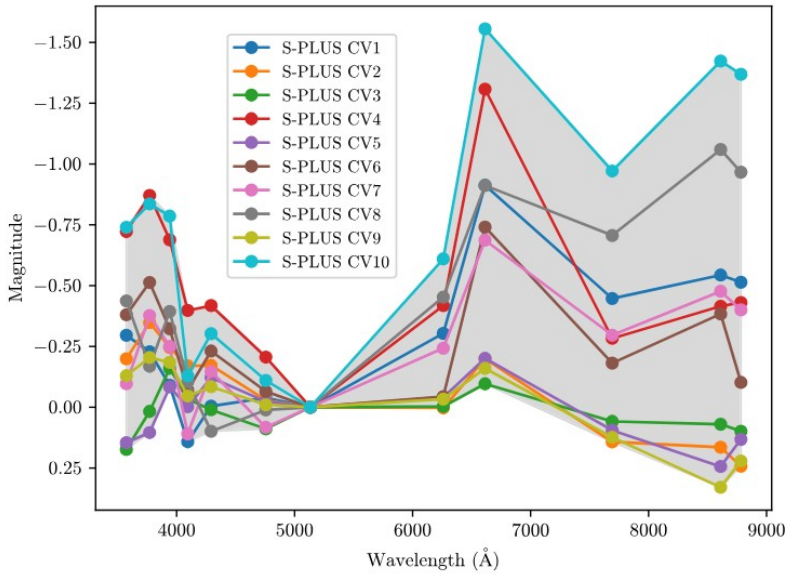




Ferreira Lopes, Gutiérrez-Soto, Ferreira Alberice, Monsalves, Hazarika, et al. (2025, A&A, 693, A306)

S-PLUS 12-Band Photometry Reveals a New Population

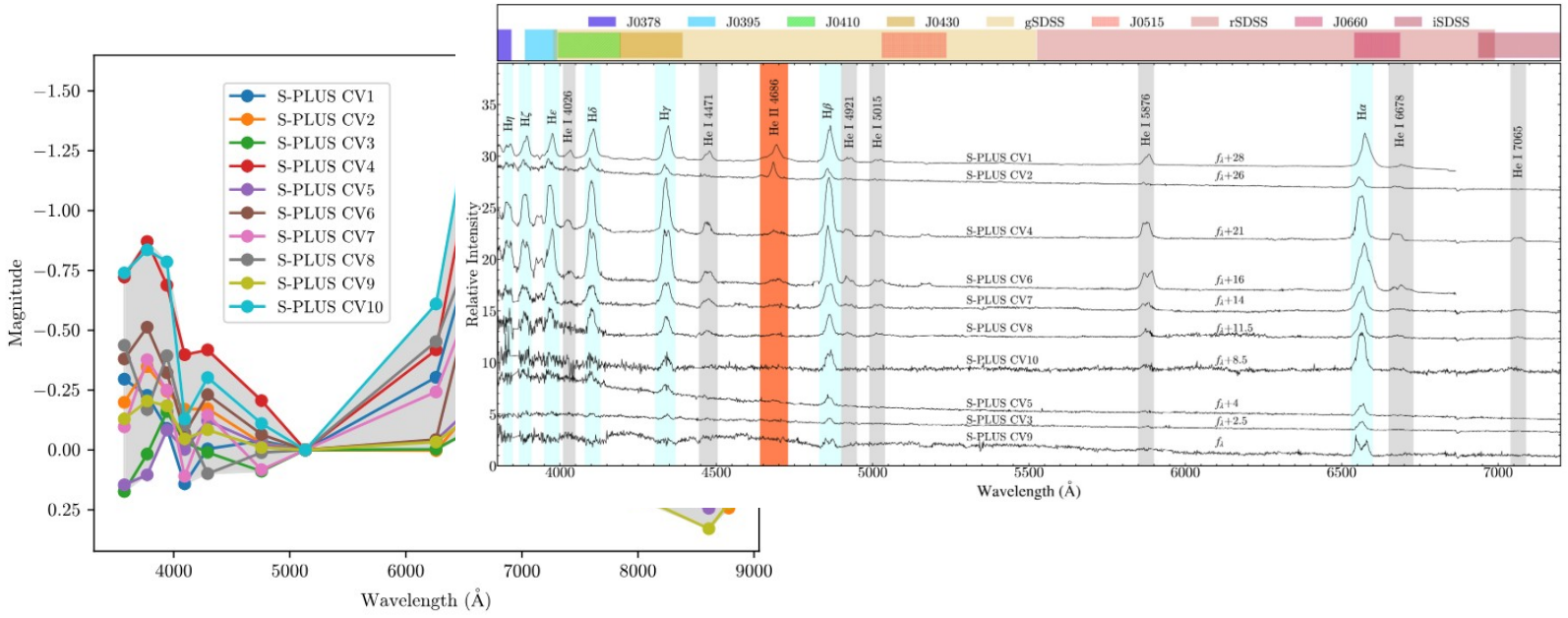
Using S-PLUS's 12-band photometry, we discovered 10 new cataclysmic variables, including rare WZ Sge-type dwarf novae and low-luminosity intermediate polars. Gemini spectra (A) and Swift X-rays (B) validate these systems, missed by traditional surveys.



Lopes de Oliveira, de Araujo, Krabe, Mendes de oliveira, Mukai, Gutiérrez-Soto, et al. (2025, ApJL, 981, L1)

S-PLUS 12-Band Photometry Reveals a New Population

Using S-PLUS's 12-band photometry, we discovered 10 new cataclysmic variables, including rare WZ Sge-type dwarf novae and low-luminosity intermediate polars. Gemini spectra (A) and Swift X-rays (B) validate these systems, missed by traditional surveys.



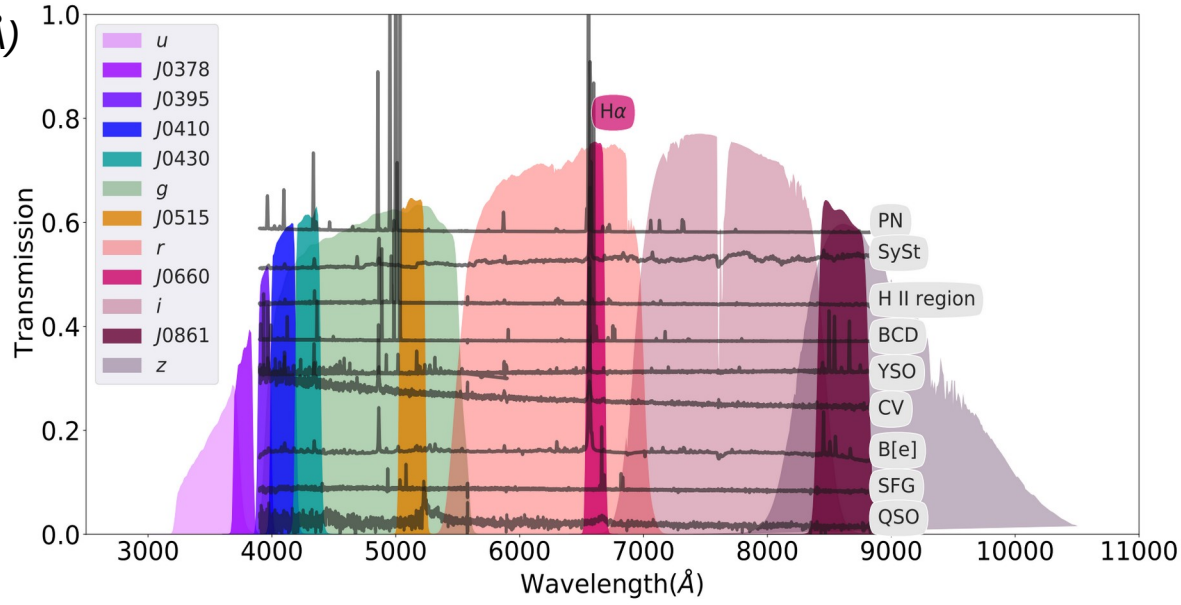
Lopes de Oliveira, de Araujo, Krabe, Mendes de oliveira, Mukai, Gutiérrez-Soto, et al. (2025, ApJL, 981, L1)

Selecting and Classifying Hα Excess Point Sources Using S-PLUS

S-PLUS: Hα sources

Hα Emission: A Universal Tracer of Cosmic Activity

- ***J0660 Filter: Capturing Hα***
  - *Hα at rest: λ = 6563 Å*
  - *Redshift range: z ≈ 0–0.15 (λ ≈ 6563–7500 Å)*
- ***Key Sources with Hα Emission:***
  - *Star-forming regions (H II regions, YSOs)*
  - *Stellar evolution: Planetary Nebulae (PNe), Be stars, Symbiotic Stars (SySt)*
  - *Accretion/Activity: Cataclysmic Variables (CVs), AGN*
  - *Transient phenomena: Supernovae (SN)*



Hα Emission: A Universal Tracer of Cosmic Activity

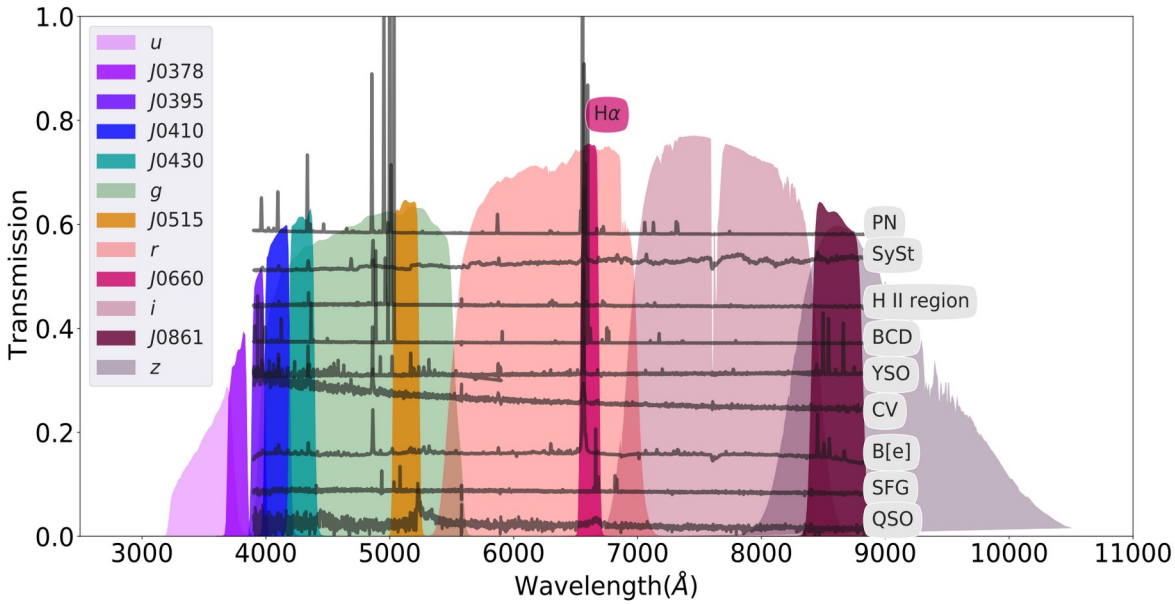
- **Origin of Hα Emission**
  - Photon emission at 6563 Å: Hydrogen's n=3 → n=2 transition.
  - Requires ionized gas (H^+) + recombination.
- **Key Drivers of Hα**
  - Birth:

    Massive stars ionize H II regions (Orion Nebula).
  - Death:

    PNe: Ionized shells around white dwarfs (Ring Nebula).

    SN: Shocks heat gas, emitting Hα.
  - Accretion:

    CVs, AGN, YSOs (e.g., jets, disks).
- **Why Hα Matters**
  - Star formation: $L_{H\alpha} \propto SFR$.
  - Stellar evolution: Maps PN shells and Be star disks.
  - Low-metallicity galaxies: Less dust extinction → stronger Hα.

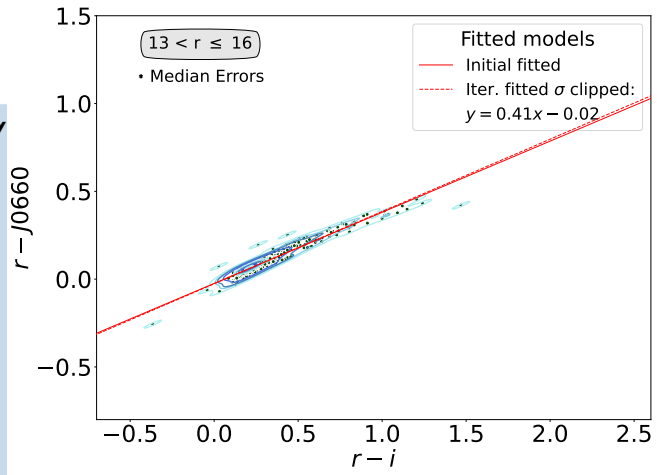


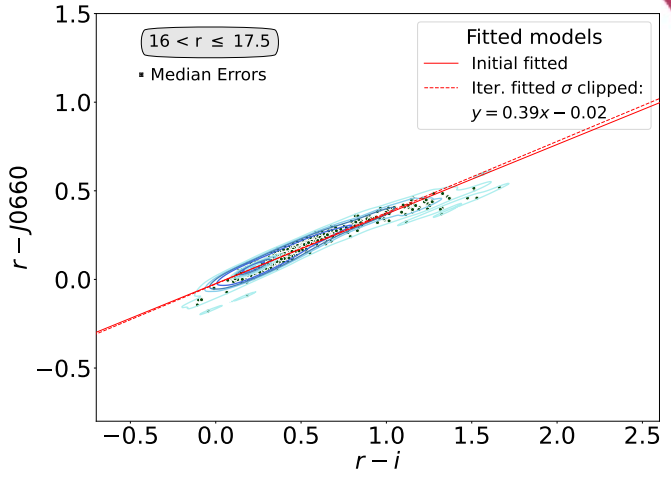
Selecting the Hα Excess Sources

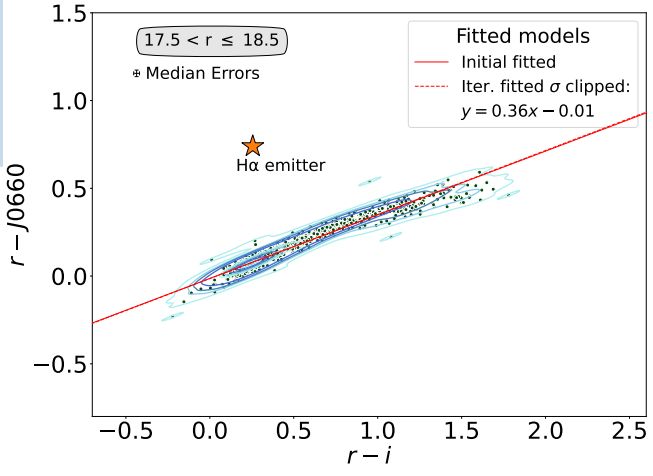
The identification of Hα excess sources was based on the method developed by Witham et al. (2008).

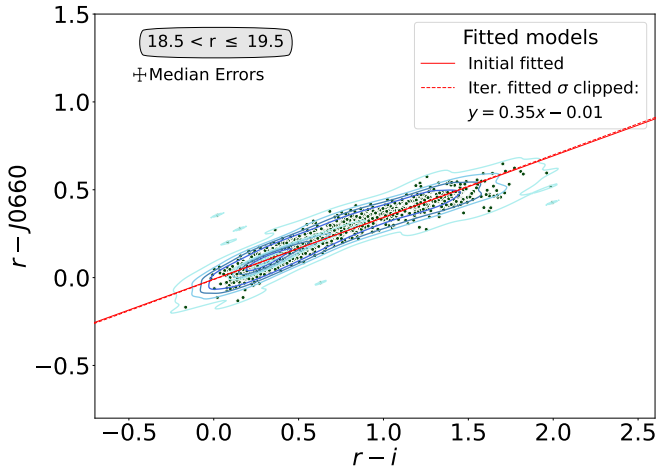
Iteratively fit the main stellar locus and select all Hα sources which lie more than 5σ away from the fitted locus .

Gutiérrez-Soto, Lopes de Oliveira, Akras, Gonçalves, Lomeli-Nunes, et al. (2025, A&A, 695, A104)

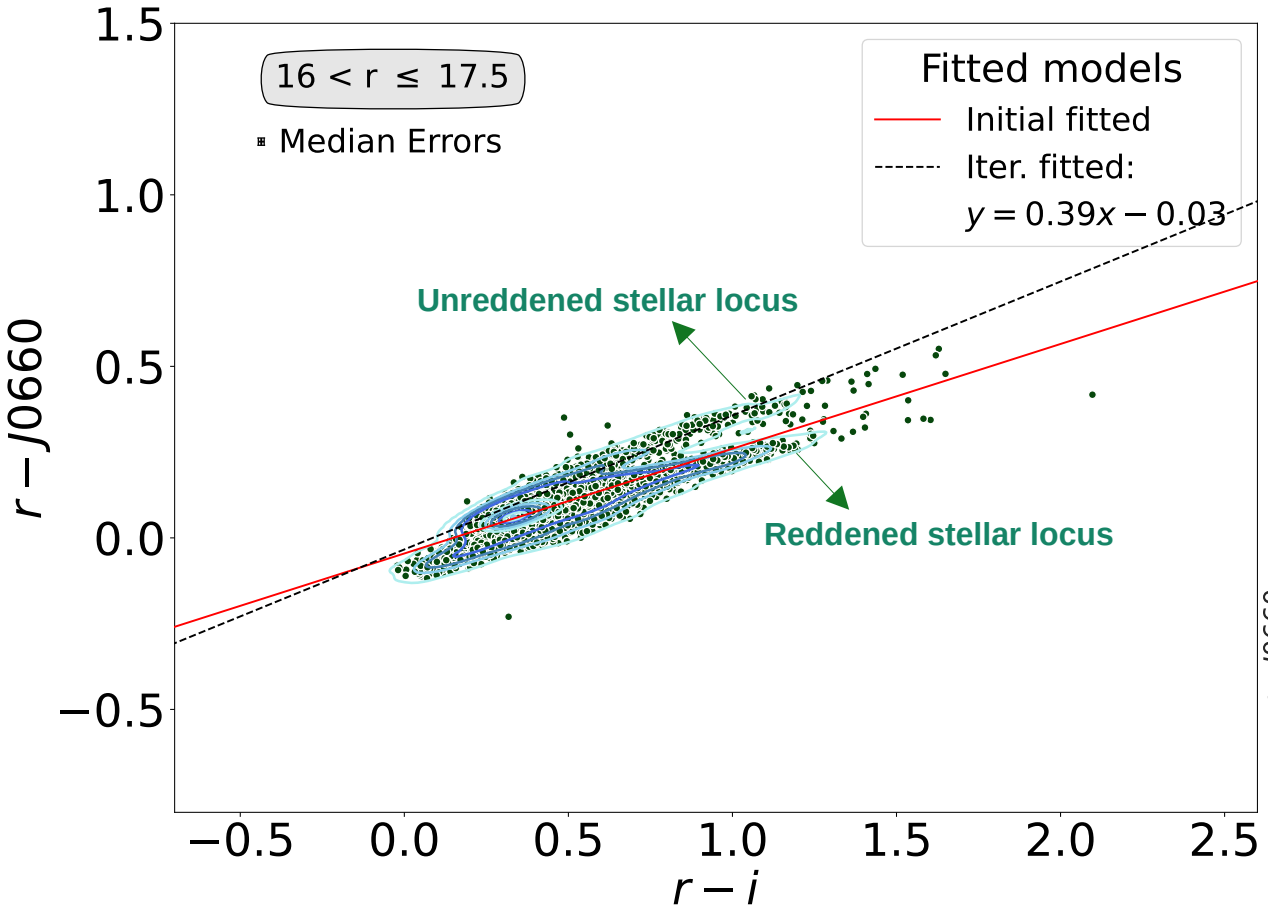








Selecting the Hα Excess Sources



Galactic Disk Challenges

Stellar Loci Duplication:

Disk fields show two stellar loci due to differential reddening and mixed populations (**main-sequence stars** and **giants**).

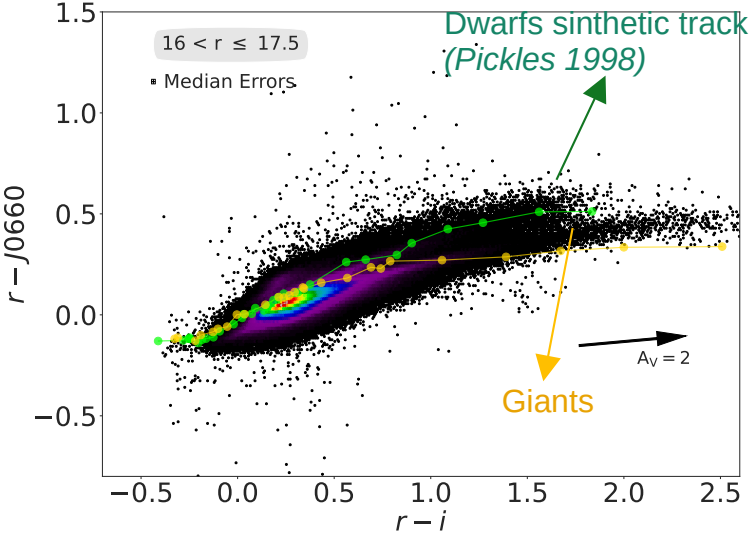
Fitting Issues:

Initial fit often aligns with the reddened locus, complicating H-alpha excess detection.

Solution:

Adjusted fit upward to align with the unreddened sequence.

Reverted to the original fit if adjustments worsened the accuracy.



S-PLUS: Hα sources

Selecting the Hα Excess Sources

Identification Criteria

- ***Key Feature:*** *Sources significantly above the locus of main-sequence/giant stars in the J0660 filter (Hα line excess).*
- ***Surveys & Outliers:***
  - ***Main Survey (MS):*** *3,637 outliers identified.*
  - ***Galactic Disk Survey (GDS):*** *3,734 outliers identified.*

Object Types

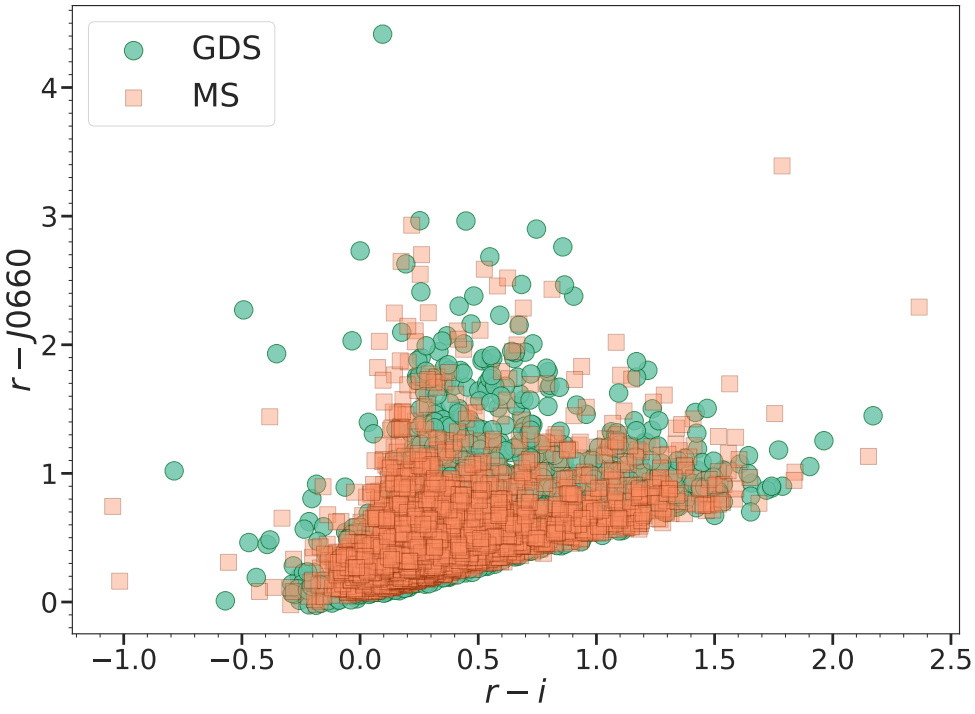
Main Survey (MS)

- ***Galactic Sources:***
  Cataclysmic Variables (CVs), binary systems, RR Lyrae stars, peculiar objects.
- ***Extragalactic Sources:***
  Active Galactic Nuclei (AGNs), Quasi-Stellar Objects (QSOs) with redshifted emission lines aligned to J0660.

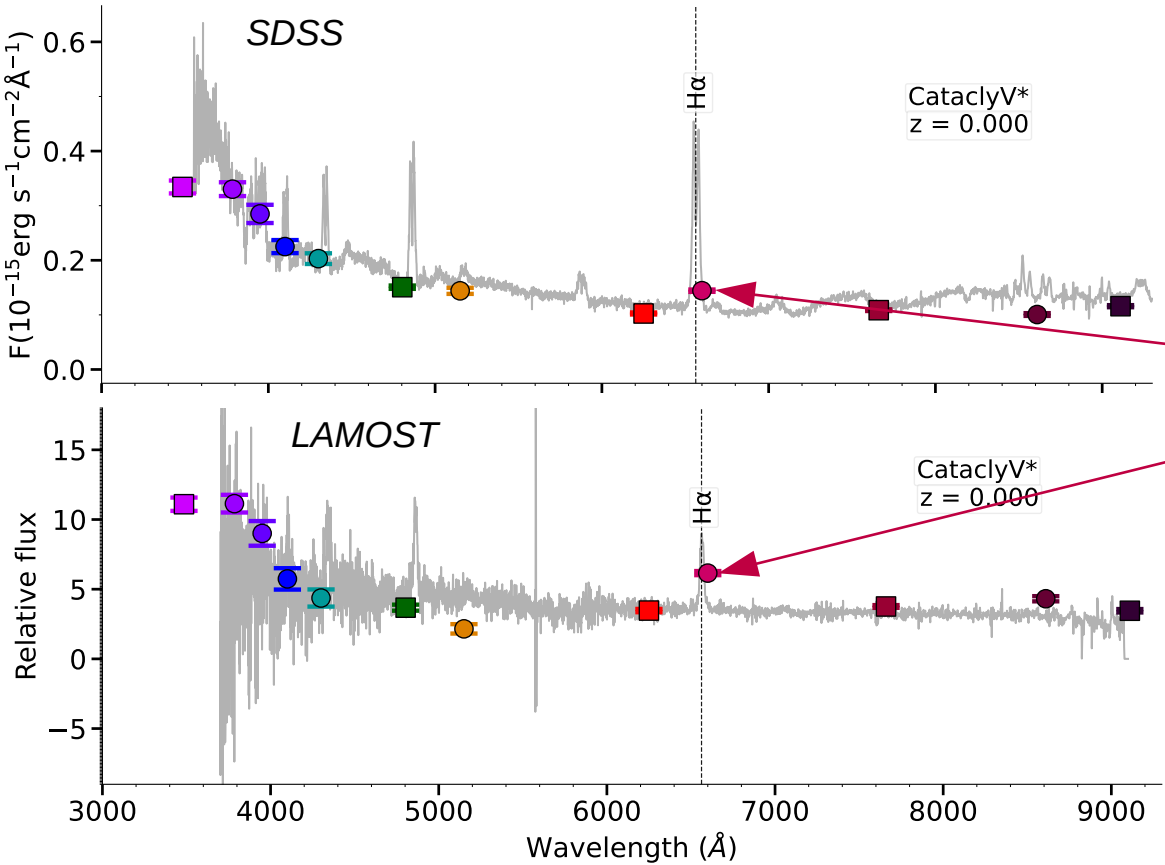
Galactic Disk Survey (GDS)

- ***Star Formation & Young Objects:***
  Emission-line stars (Em, Be*), Young Stellar Objects (YSOs: T Tauri, Herbig Ae/Be).*
- ***Nebular/Variable Sources:***
  Planetary nebulae, novae, reflection nebulae; cataclysmic variables, eclipsing binaries.

Note: The high prevalence of Em and YSOs in GDS highlights active star formation processes in the Galactic disk.*



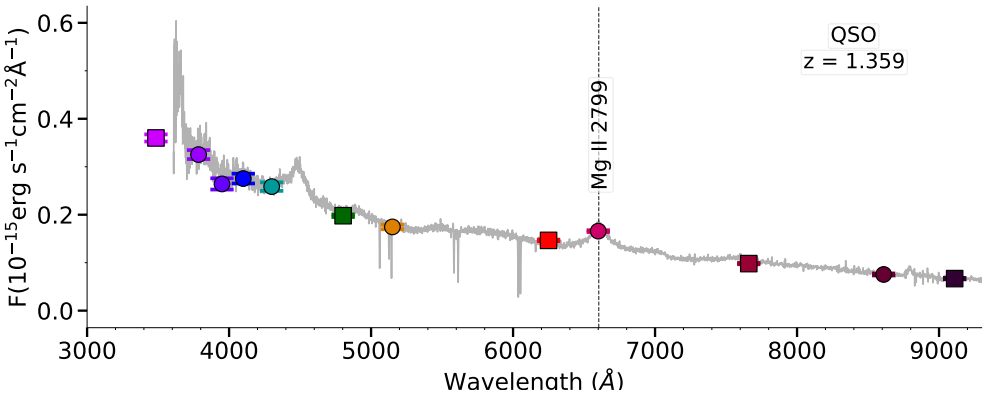
Example of two Hα Emitters

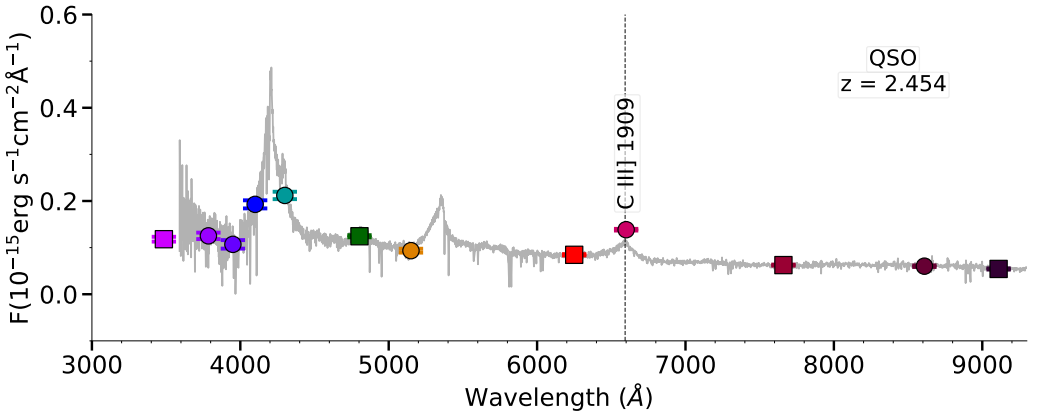


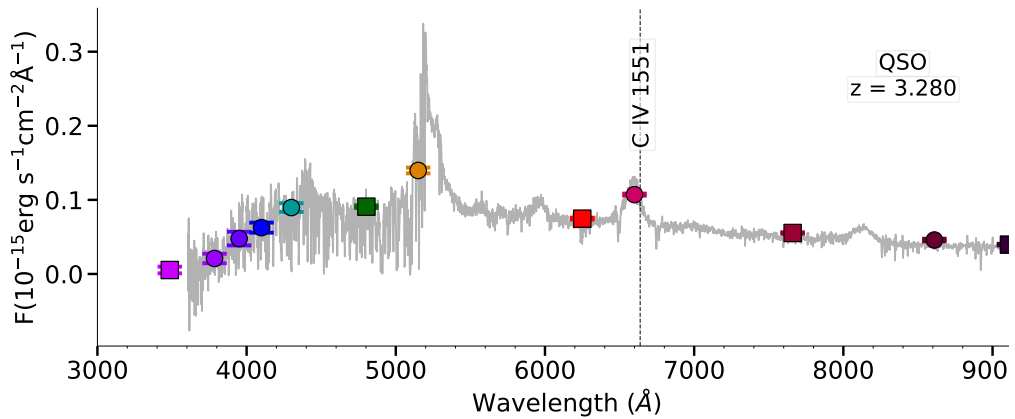
J0660 filter

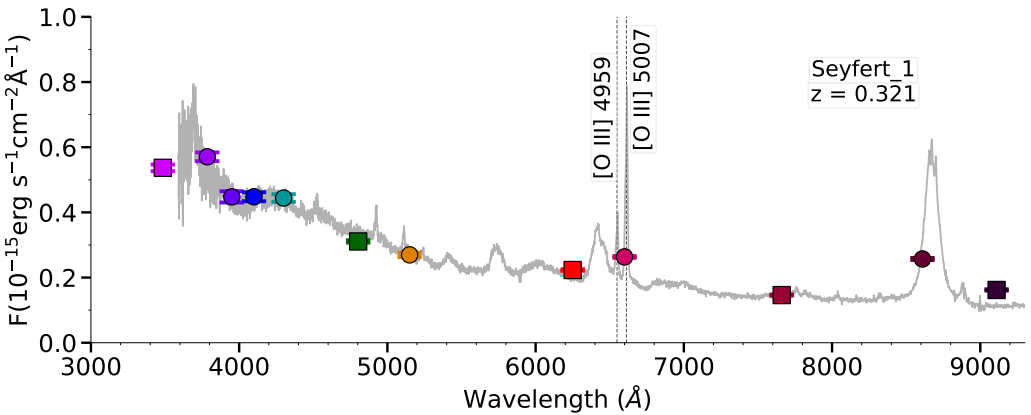
Gutiérrez-Soto, et al. (2025)

Extragalactic Compact Sources









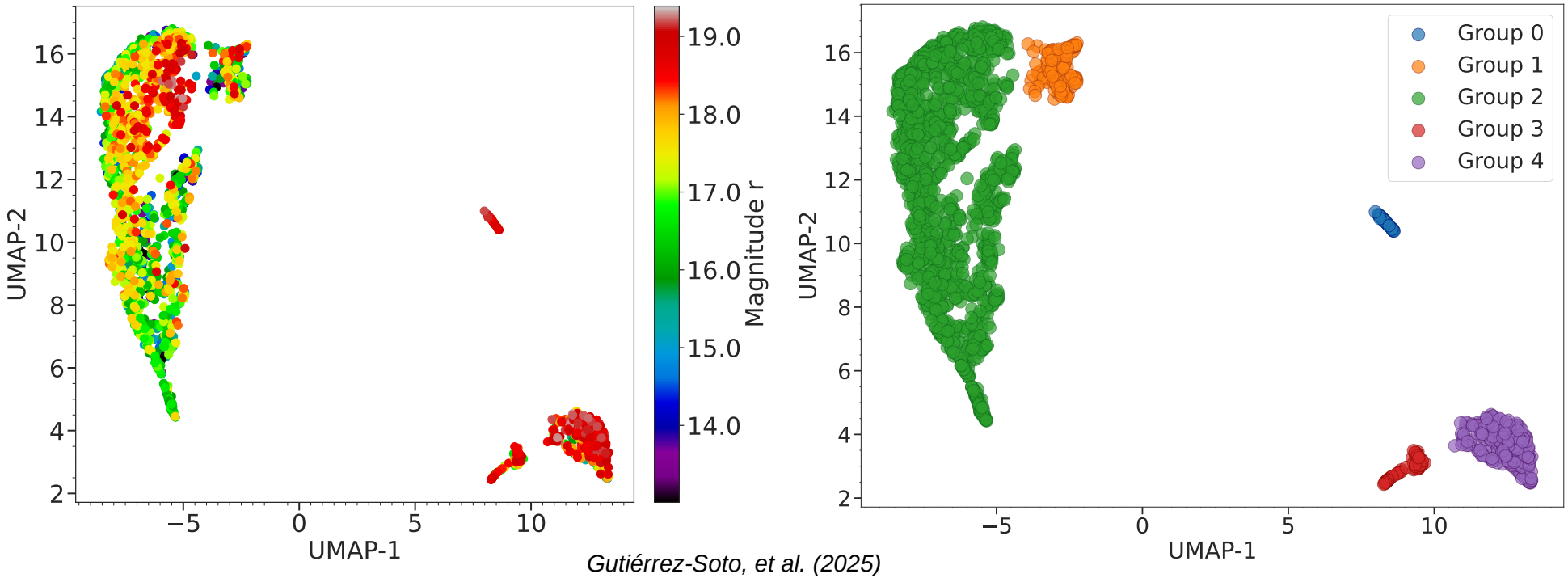
Our selection criteria effectively isolated compact sources, minimizing the inclusion of extended sources. In agreement with SIMBAD, only 23 AGN (3.1%) and 9 galaxies (1.2%) were identified in the main survey, while 143 QSOs (19.6%) were selected. The spectra of these QSOs and Seyfert galaxy show other emission lines that fall into the J0660 filter, resulting in a significant Hα excess.

Machine Learning Approaches

We applied ***UMAP*** for dimensionality reduction on S-PLUS data and used ***HDBSCAN*** for clustering Hα excess sources. Two experiments were conducted: one using only the ***66 S-PLUS*** colors and another incorporating WISE bands. The results revealed the formation of five distinct groups, effectively separating galactic and extragalactic sources.

Machine Learning Approaches

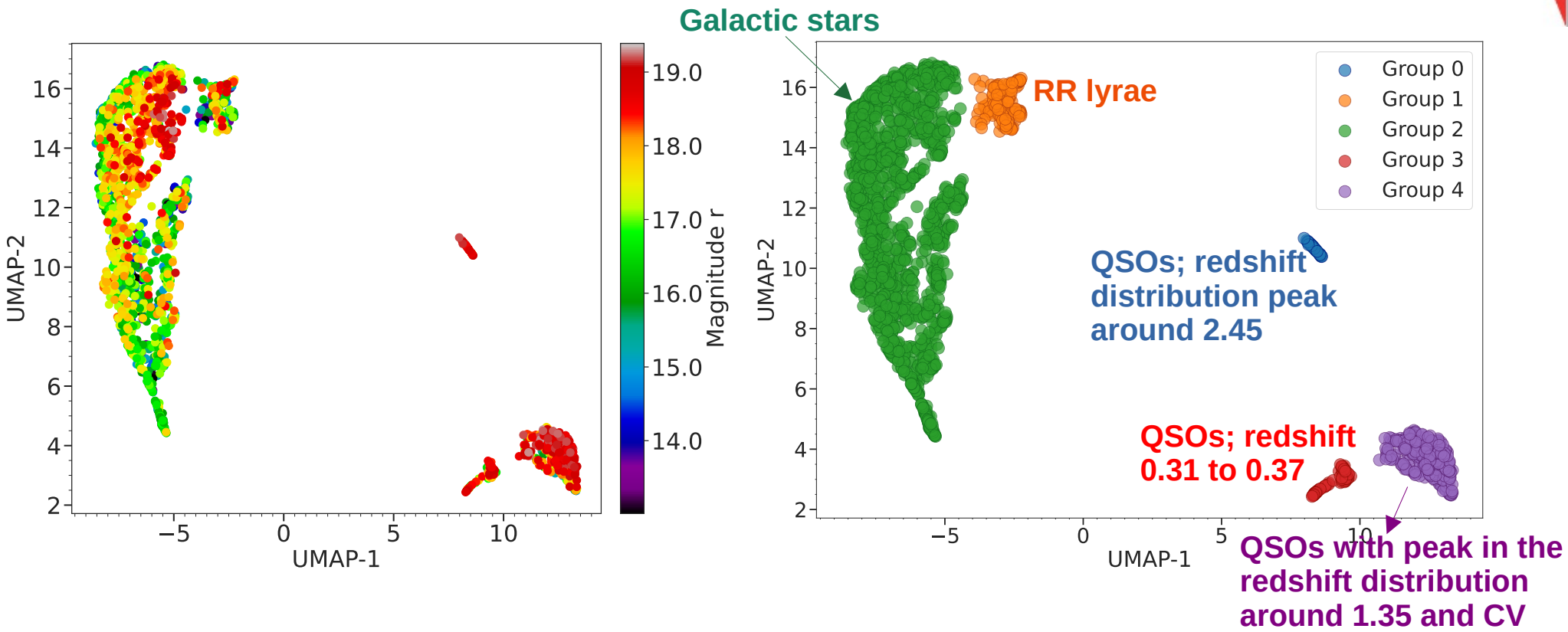
Only the ***66 S-PLUS*** colors



Gutiérrez-Soto, et al. (2025)

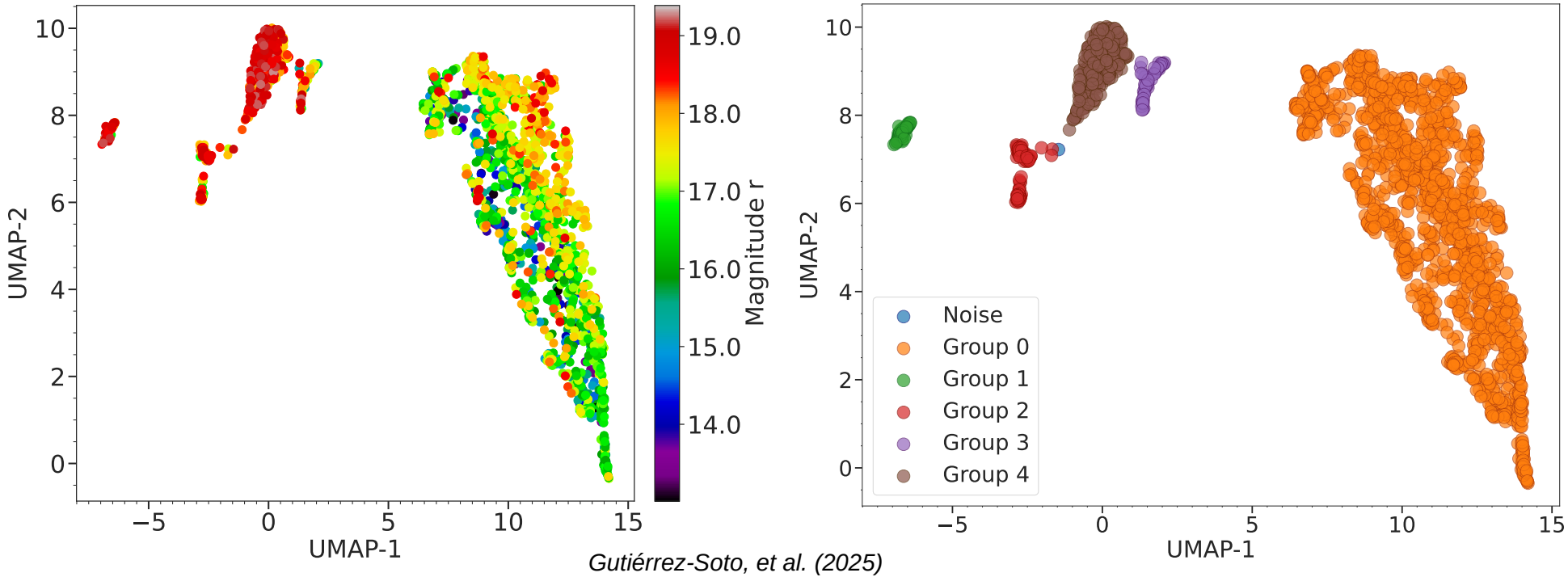
Machine Learning Approaches

Only the ***66 S-PLUS*** colors



Machine Learning Approaches

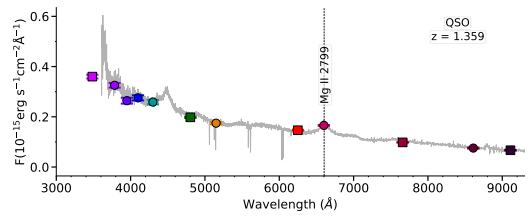
66 S-PLUS colors + WISE bands

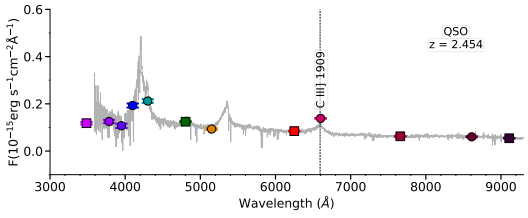


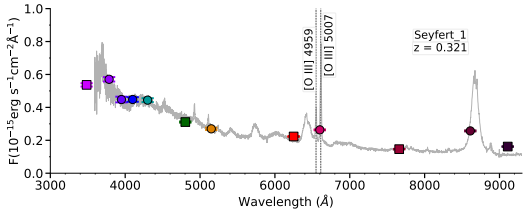
Gutiérrez-Soto, et al. (2025)

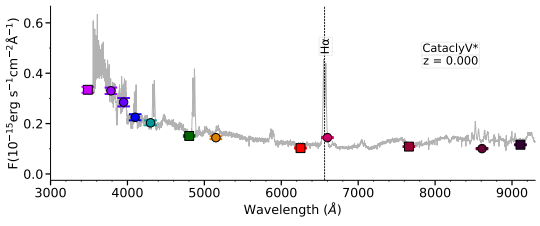
Machine Learning Approaches

66 S-PLUS colors + WISE bands

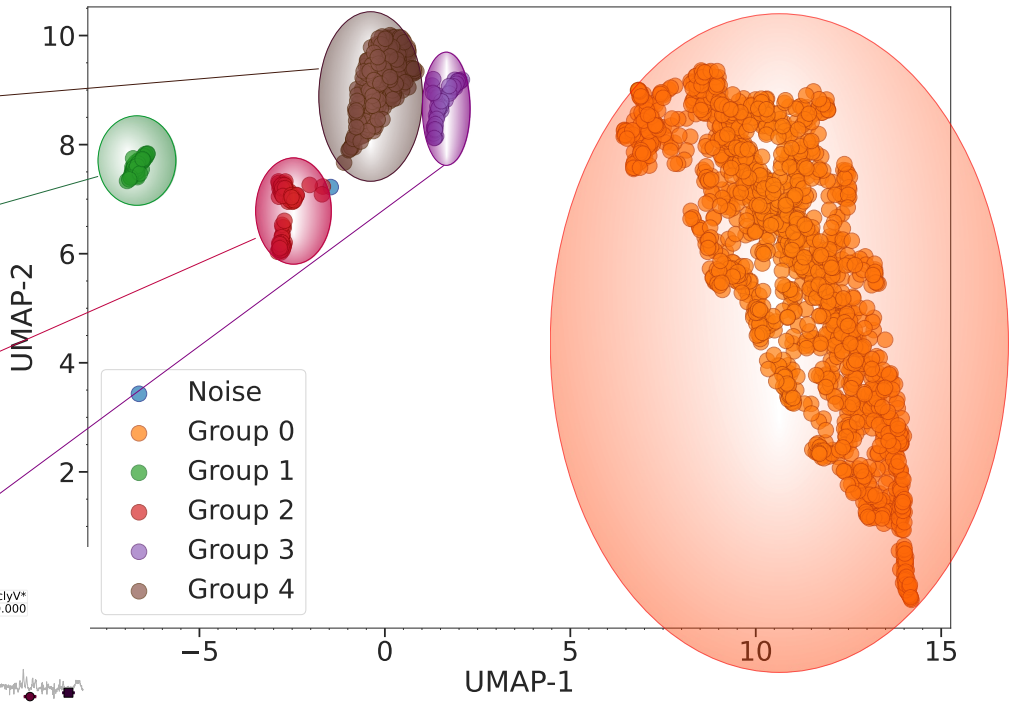








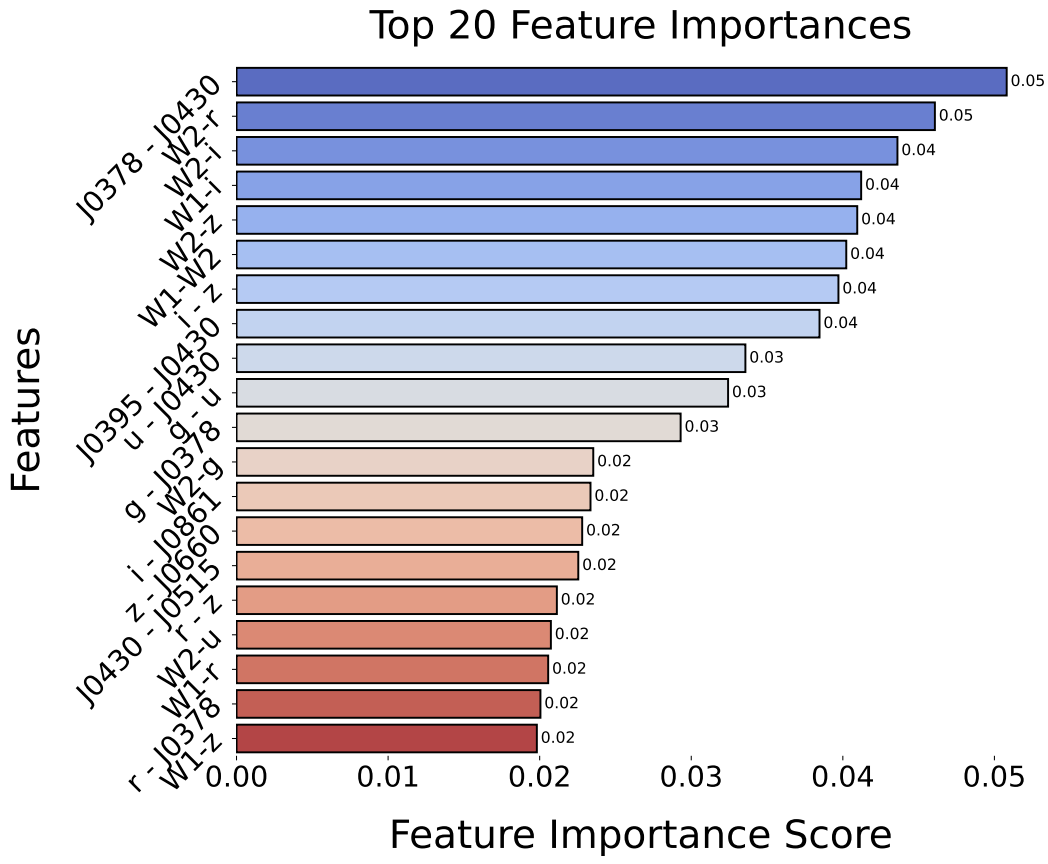
Galactic stars



Extracting Main Features: Color Analysis

Analysis: Key colors from S-PLUS and WISE filters identified by **Random Forest** as most influential in classifying Hα excess objects.

Results: **Top 20 features** contribute significantly to distinguishing between different object classes identified by UMAP + HDBSCAN.



Gutiérrez-Soto, et al. (2025)

Extracting Main Features: Color Analysis

Group 0:

Primarily Galactic stars.

Group 1:

Dominated by QSOs, with at a redshift of approximately 2.45.

Group 2:

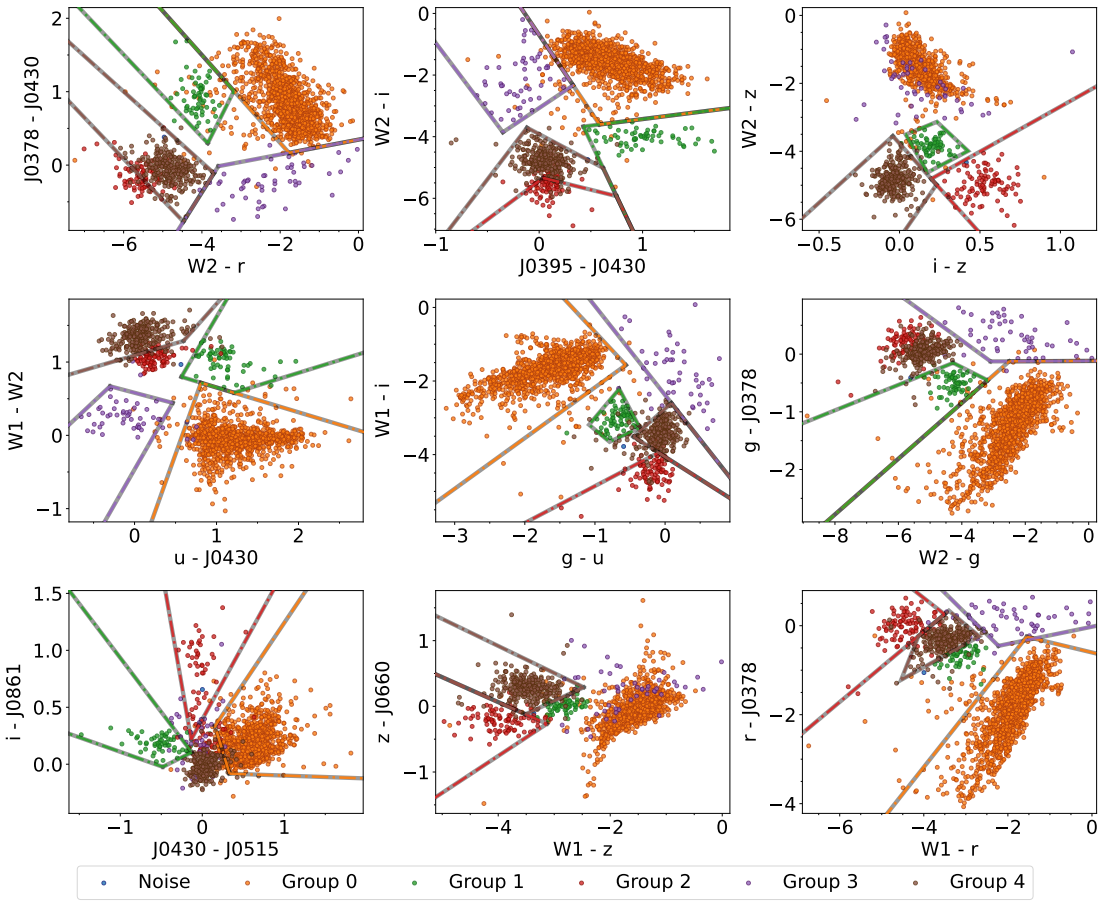
Consists mostly of active galactic nuclei—QSOs, Seyfert 1 galaxies, and AGN —with at redshifts around 0.31–0.37.

Group 3:

Characterized by a predominance of cataclysmic variables.

Group 4:

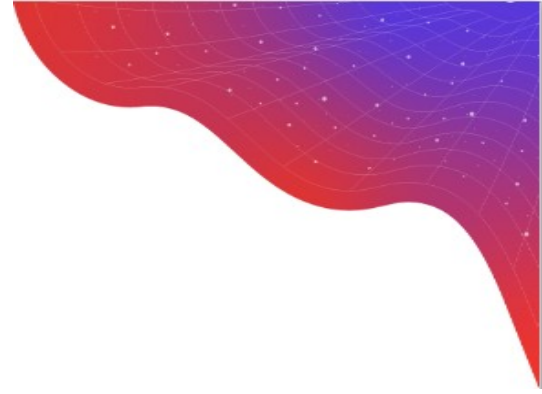
A diverse mix featuring mainly QSOs and AGNs peak around a redshift of 1.35.



Gutiérrez-Soto, et al. (2025)

Conclusions

- ***Identification of H-alpha Excess Candidates:***
- ✔ 7,371 candidates identified using the S-PLUS J0660 filter.
- ✔ Distribution: 3,637 in the high-latitude main survey and 3,734 in the Galactic disk.
- ***Classification of Objects:***
- ✔ Cross-referencing with SIMBAD: EM stars, YSOs, Be stars, CVs, PNe, QSOs, non-local galaxies, and RR Lyrae stars.
- ***Machine Learning Techniques:***
- ✔ UMAP and HDBSCAN: Effective differentiation between Galactic and extragalactic objects
- ✔ Challenges in distinguishing cataclysmic variables from QSOs or AGN with $z \approx 1.35$.
- **Integration of WISE Data:**
- ✔ Improved separation of extragalactic and Galactic sources.
- **Random Forest Model:**
- ✔ WISE data crucial for identifying significant features.
- ✔ Improved color-color diagrams and understanding of H-alpha related phenomena.



THANK!!

gsoto.angel@gmail.com