

VNIVERSITAT
® VALÈNCIA



I Meeting Aragón – Comunidad Valenciana
PPCC-AstroFAE
Galáctica (Arcos de las Salinas) - 28 y 29 mayo 2025

Nueva infraestructura UPAD

Héctor Vázquez Ramió

1st Meeting Aragón – Comunidad Valenciana
Complementary Plans of Astrophysics and High Energy Physics
Galáctica, Arcos de las Salinas, Teruel
May 28 – 29, 2025



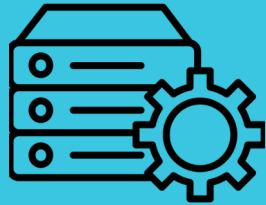
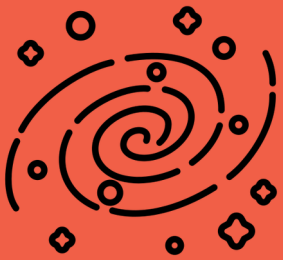
Departamento de Procesado y Archivo de Datos (DPAD) - The Team

Astrophysics

Pipelines & DA

Systems

Databases + Web



Head of the DPAD



Vázquez Ramió, Héctor

DPAD Team

SYSTEMS GROUP



López Sainz, Ángel



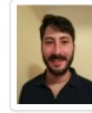
Castillo Ramírez, Juan



Akhlaghi, Mohammad



Hernán Caballero, Antonio



Moreno Signes, Alberto



Muniesa Gallardo, David José

ARCHIVE GROUP



Civera Lorenzo, Tamara



Hernández Fuertes, Javier



Vives Arias, Héctor

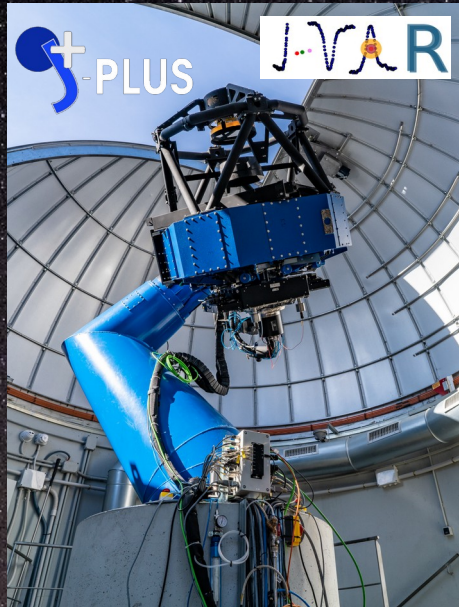


Lorenzetti, Giacomo



López San Juan, Carlos

OAJ – Infraestructura Científico-Técnica Singular (ICTS)



JAST80
Ritchey-Chrétien plus Field
Corrector
Diameter = 83 cm
FoV = 2 deg \varnothing
F/4,5
M2 hexapod

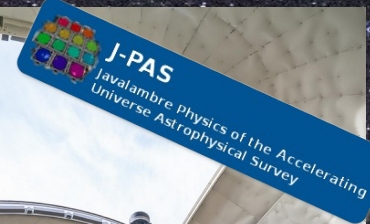
The *Observatorio Astrofísico de Javalambre (OAJ)* is conceived to carry out **large astronomical surveys**, starting with the **J-PAS** and **J-PLUS**

To this purpose, the OAJ has two unique telescopes and instrumentation of unusually **large FoV** and a **data center**.

The Unidad de Procesamiento y Archivo de Datos (UPAD) is an **essential infrastructure** of the **OAJ-ICTS**

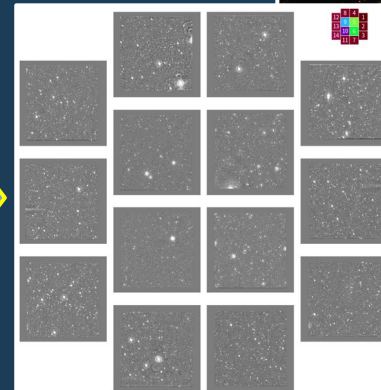
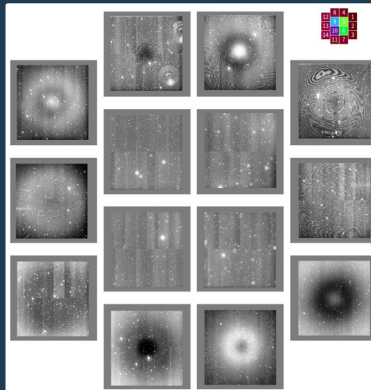
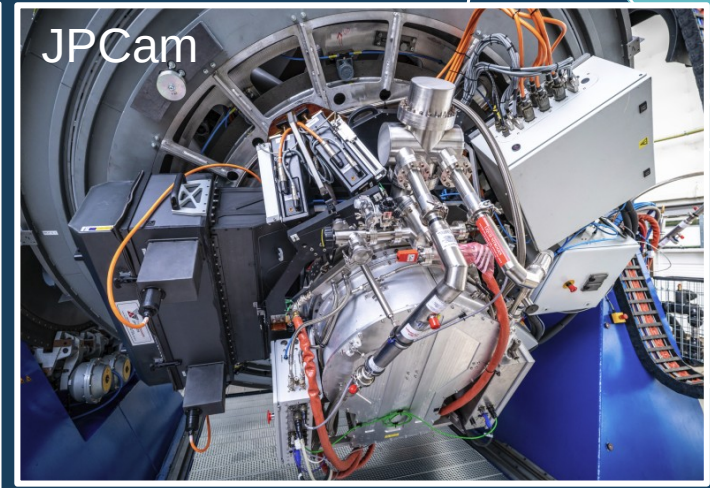
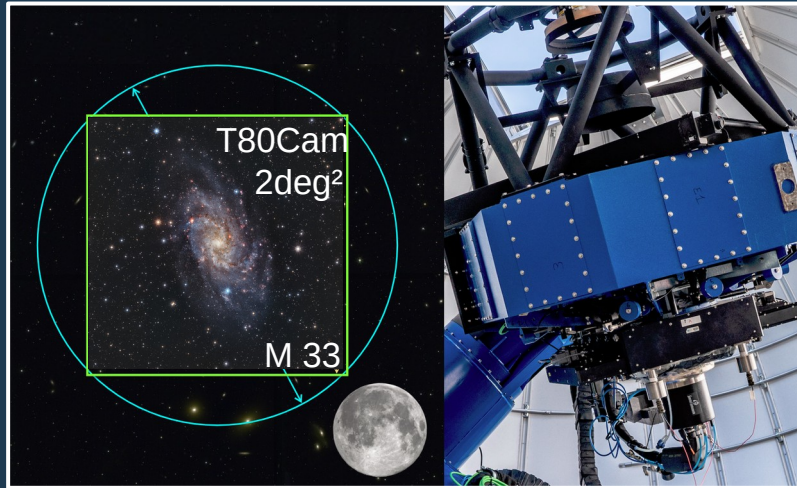


JST250
Ritchey-Chrétien plus Field Corrector
Diameter = 2,55 m
FoV = 3 deg \varnothing
F/3,5
M2 hexapod



DPAD duties

- Gathering data from the OAJ
- Storage and backup
- Data digestion, processing and analysis
- Data publication
- Also:
 - Data quality checks
 - Tools supporting Telescope Operation
 - HPC service



:~\$ jype



OAJ \Rightarrow UPAD - Data flow to the Data Center



100G

OAJ



OAJ/CPD
2x Service nodes
2x I/O nodes
2x Computing nodes
24 cores
192GB RAM

100TB storage

Data delivery to Teruel
Quick reduction of the images



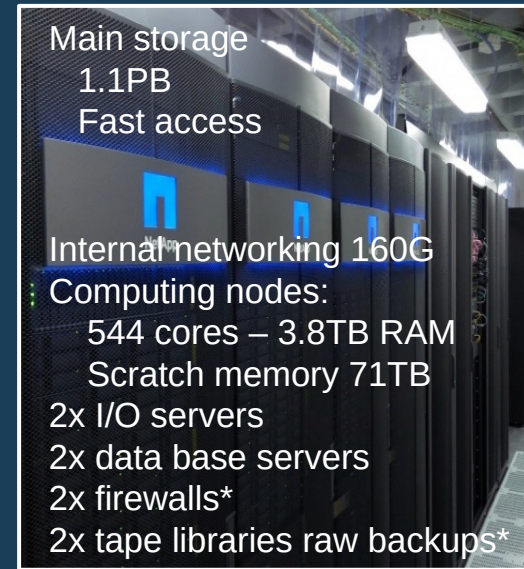
10G*

CEFCA's Headquarters

External Data Access Machine (EDAM-NG)



Gobierno de Aragón – Data Center



Main storage
1.1PB
Fast access

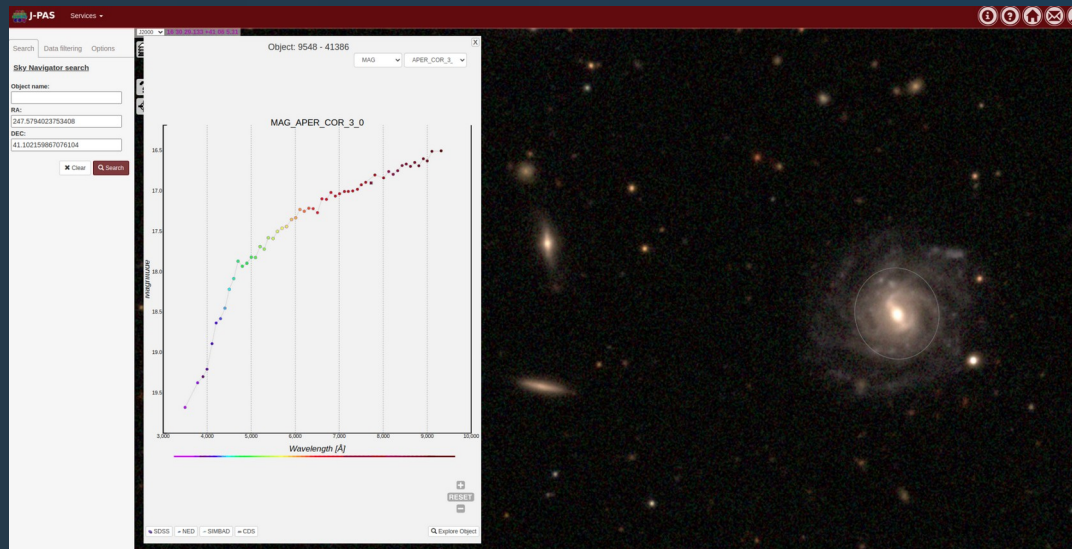
Internal networking 160G
Computing nodes:
544 cores – 3.8TB RAM
Scratch memory 71TB
2x I/O servers
2x data base servers
2x firewalls*
2x tape libraries raw backups*



Data archival
Robotic tape library
1600 tapes (4.1PB)
LTO6 tapes \Rightarrow LTO9

Data Publication

- CEFCa catalogs portal: <https://archive.cefca.es/catalogues/>
 - DRs of the main surveys: J-PAS, J-PLUS, J-VAR
 - Value Added Catalogs (VAC)
 - DRs of the legacy surveys with JAST80
- TAC Data portal: <https://tacdata.cefca.es/>
 - Delivery of Open Time programs' data



Following FAIR principles

Compliant with Virtual Observatory standards

UPAD

in the framework of the PPCC

UPAD in the PPCC

- LA 4. *Ciencia.*
- LA 8. *Computación, Big data e inteligencia artificial*

**UPAD is the Data Center
of J-PAS y J-PLUS**

Línea de Actuación	LA8 - Computación, <i>big data</i> e inteligencia artificial
Código de proyecto	LA8.A1
Título	Renovación y mejora del centro de datos UPAD
Líder del proyecto	Centro de estudios de Física del Cosmos de Aragón (CEFCA)
Entidades ejecutantes	Centro de estudios de Física del Cosmos de Aragón (CEFCA)
Descripción/Objetivos:	
<p>El objetivo del proyecto es mejorar la capacidad y prestaciones del centro de datos <i>Unidad de Procesado y Archivo de Datos</i> (UPAD) del CEFCA, que, con más de 5PBs de almacenamiento en disco y cintas, constituye en sí mismo una instalación esencial de la ICTS OAJ. Así, este proyecto potencia tanto las capacidades computacionales de la ICTS como los accesos al centro de datos por parte de la comunidad científica internacional. Se procederá por tanto a la renovación parcial del equipamiento de la UPAD, comenzando por aquellos equipamientos y sistemas que han llegado al final de su vida útil y han quedado obsoletos. En particular, esta primera fase de renovación de equipos incluye la actualización completa del equipo de acceso a datos de la UPAD desde el exterior, la máquina <i>External Data Access Machine</i> (EDAM), así como buena parte del almacenamiento de acceso rápido (almacenamiento en disco), donde se almacenan los datos utilizados de forma más frecuente. Además de proceder al acopio del hardware necesario mediante diversas licitaciones, se pretende contar con el apoyo de empresas tecnológicas, para el suministro de desarrollos necesarios para la nueva UPAD así como para el estudio de viabilidad de soluciones energéticas más adecuadas y ecológicas.</p> <p>La renovación y mejora de la UPAD es esencial para garantizar el almacenamiento, el análisis, la calibración científica y, en definitiva, la gestión y la publicación de los datos científicos de J-PAS, J-PLUS y de los demás cartografiados y proyectos del OAJ. Habida cuenta del gran volumen de datos astronómicos que producen cada noche los telescopios del OAJ, sin la UPAD ninguno de los proyectos anteriores sería posible.</p>	
Hitos:	
H8.1.1 – Suministro, instalación y puesta en funcionamiento de EDAM.	
H8.1.2 – Suministro, instalación y puesta en funcionamiento de parte de la renovación del almacenamiento de acceso rápido (almacenamiento en disco).	

LA8.2 Renewal of UPAD's main storage

Old storage: end of its life (disks failing)

Received in Jun 2024. In production since Dec 2024. Final position in Apr. 2025.

✓ 1st phase of the renewal complete:

- 1.1PB \Rightarrow 1.1PB @ 5000MB/s (avg r/w)
- 4 racks \Rightarrow 2 racks (reallocation HPC)

2nd phase: \sim x2 storage
(using same 2 racks)

Hitos:

H8.1.1 – Suministro, instalación y puesta en funcionamiento de EDAM.

H8.1.2 – Suministro, instalación y puesta en funcionamiento de parte de la renovación del almacenamiento de acceso rápido (almacenamiento en disco).



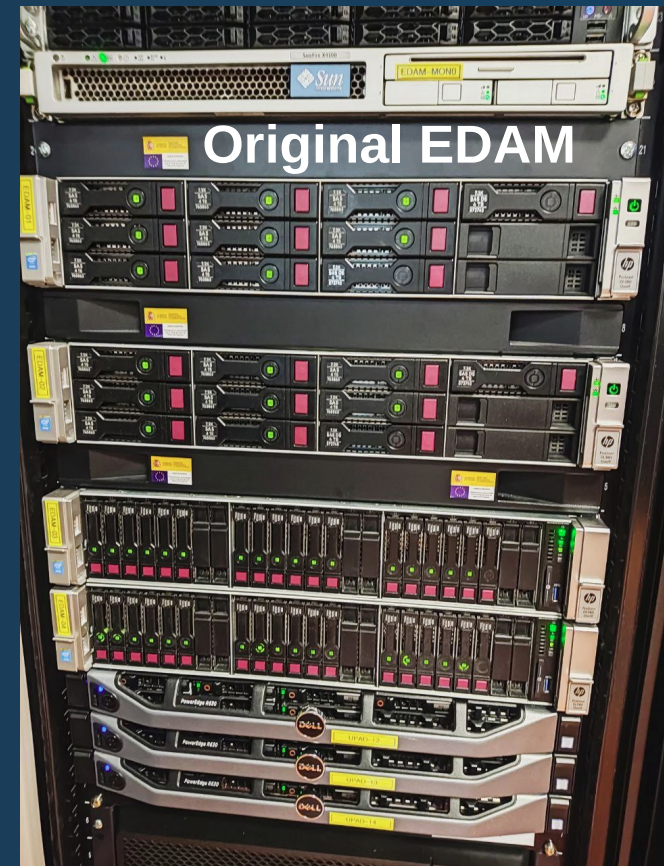
External Data Access Machine (EDAM)

OAJ Data is distributed to the world via EDAM:

- Data Releases (DRs) of the main surveys
- DRs of the Legacy Surveys
- Open time data
- Internal DRs

Previous system.

- 2 web servers + 2 DB servers:
 - Only 1 pair web+DB was running (+ a backup)
- Effective storage: 40TB
- Internal networking: 1GbE

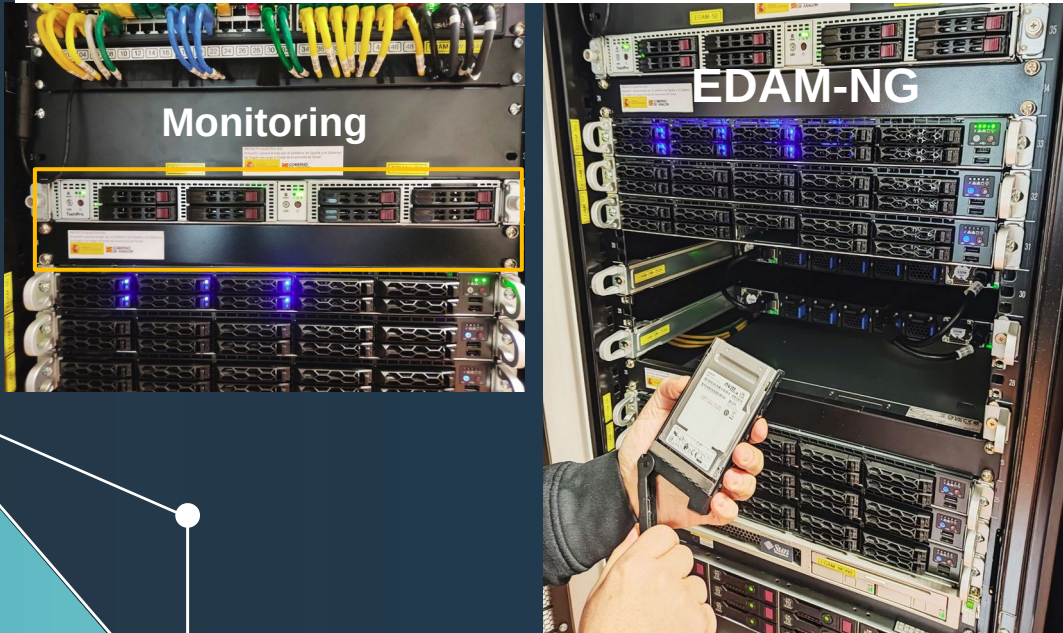


LA8.1 EDAM-NG (Next Generation)

Hitos:

H8.1.1 – Suministro, instalación y puesta en funcionamiento de EDAM.

H8.1.2 – Suministro, instalación y puesta en funcionamiento de parte de la renovación del almacenamiento de acceso rápido (almacenamiento en disco).



Installed in 2023. 2024: preparation of the system and the applications (into dockers) to be installed, including an upgrade of the CECA connectivity.

✓ In production in May 2025

- 6 general purpose + 2 monitoring servers
 - Effective storage (shared btw nodes): 100TB
 - NVMe solid state disks
 - Internal networking: 25GbE
- Several instances are running. **Highly scalable and transparent for the user.** Currently:
 - 3 instances of the Catalogs Web Portal
 - 3 instances of TAC Data (Open Time)
 - 2 instances of the DBs

Heavy queries to the DB now take ~1/2 of the time

Currently, only ~10% of the total storage is used

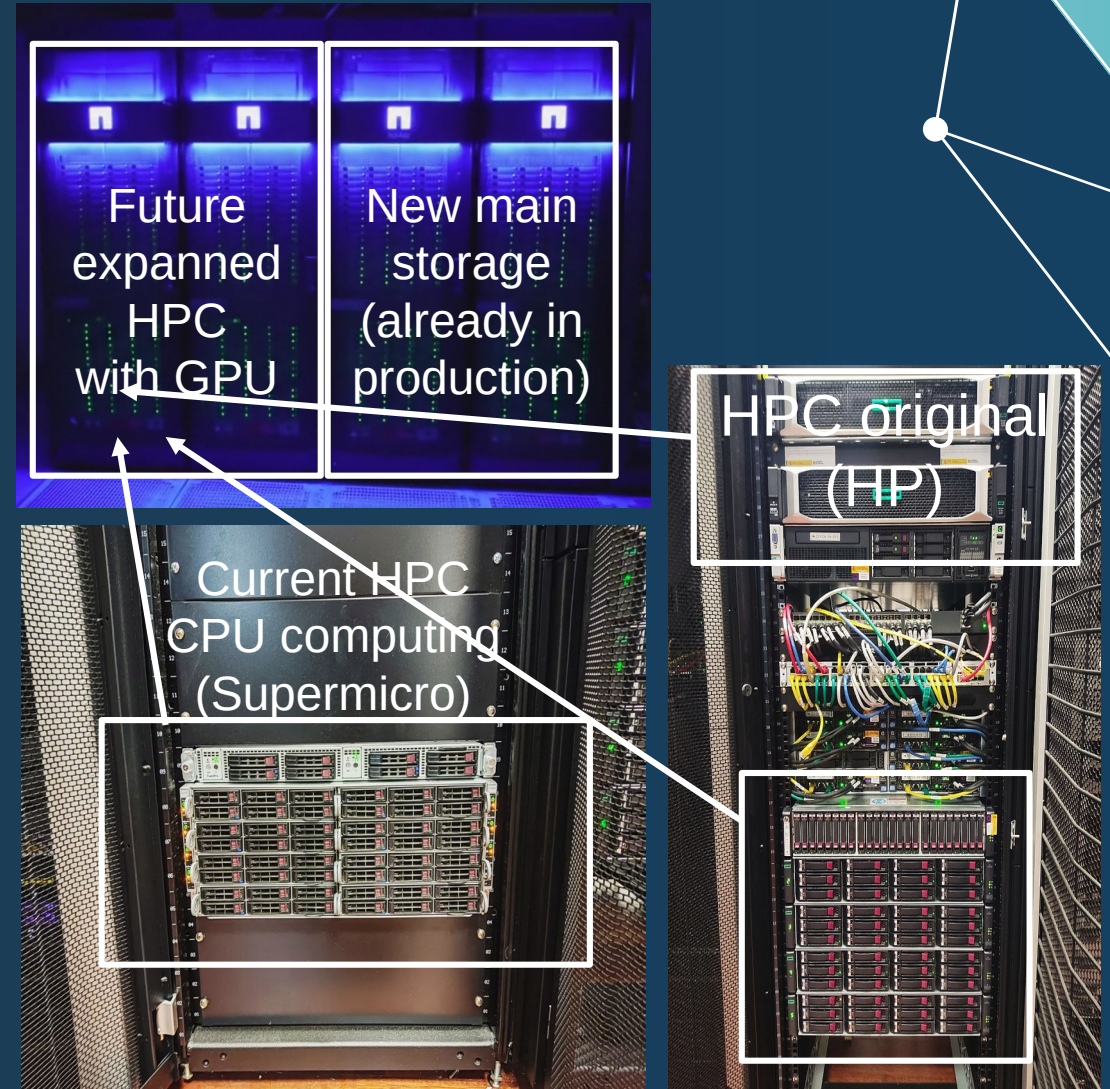
Expansion of the HPC

Recently added in the PPCC.

- Storage and most of the CPU computing from the current HPC
- Adding GPU. 5 servers with (per server):
 - At least 2GPUs
 - At least 800GB storage
 - At least 1.5TB NVMe scratch memory
 - At least 542GB RAM
 - At least 1 CPU (minimum 64 cores)
- Internal networking upgrade to 25GbE

System to be installed in the UPAD due to cooling constraints @ CEFCA's HQ: in the 2 freed racks

Tender open on May 14th 2025



PPCC Valencia actions

- Renewal 2x firewalls UPAD – CEFCA HQ: 1GbE⇒10GbE
- Renewal 2x LTO9 tape libraries for daily raw backups
- Power distribution units (PDU) and T sensors
- Establishment of a collaboration agreement UV-CEFCA

VNIVERSITAT
DE VALÈNCIA



GOBIERNO
DE ARAGON



MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



Plan de Recuperación,
Transformación
y Resiliencia



Financiado por
la Unión Europea
NextGenerationEU

Acuerdo de colaboración entre el Centro de Estudios de Física del Cosmos de Aragón y la Universitat de València en el marco del Plan Complementario de “Astrofísica y Física de Alta Energías” denominado “Tecnologías avanzadas para la exploración del universo y sus componentes” previsto en el Plan de Recuperación, Transformación y Resiliencia-Mecanismo de Recuperación y Resiliencia , financiado por la Unión Europea- NextGenerationEU

REUNIDOS



Computation,

big data and AI



LeMoNNADE

Learning Model With Neural Network Adaptive Distance Estimation

Photometric redshifts (J-PAS & J-PLUS) with **LeMoNNADE**: (Hernán-Caballero et al. 2025, in prep.)

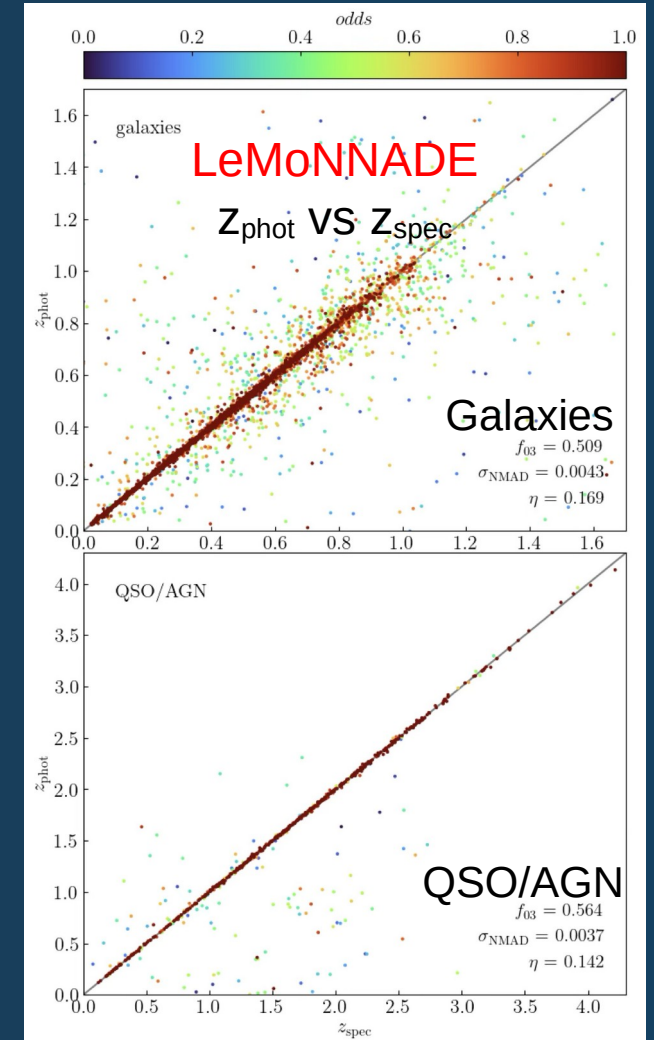
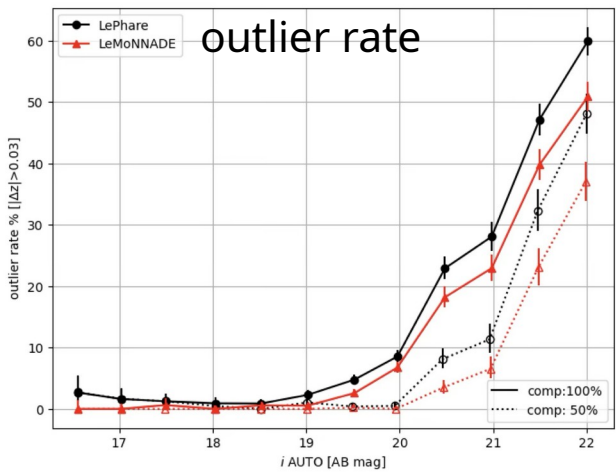
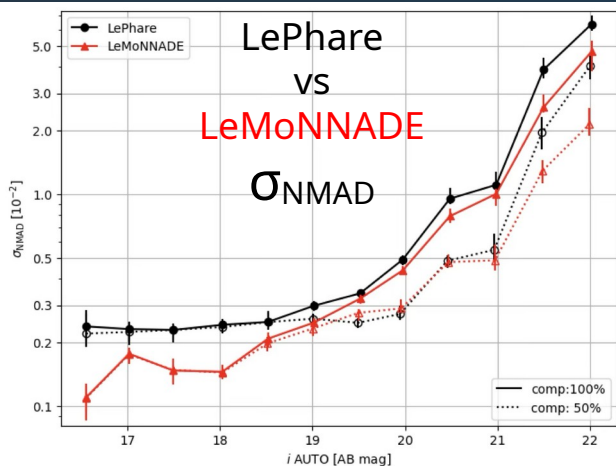
Technical aspects:

- Use of a dense network with a standard architecture.
- Pre-training with synthetic fluxes from DESI and adaptation with real data.
- Inputs: fluxes and errors. **Uncertainty** estimation (redshift PDF)

Performance:

- Significant **improvement** along all mag. range
- Trains a **unique model** for all objects: estimates redshift regardless object type \Rightarrow adds QSO/AGN estimates wrt LePhare (based on template fits of galaxies)

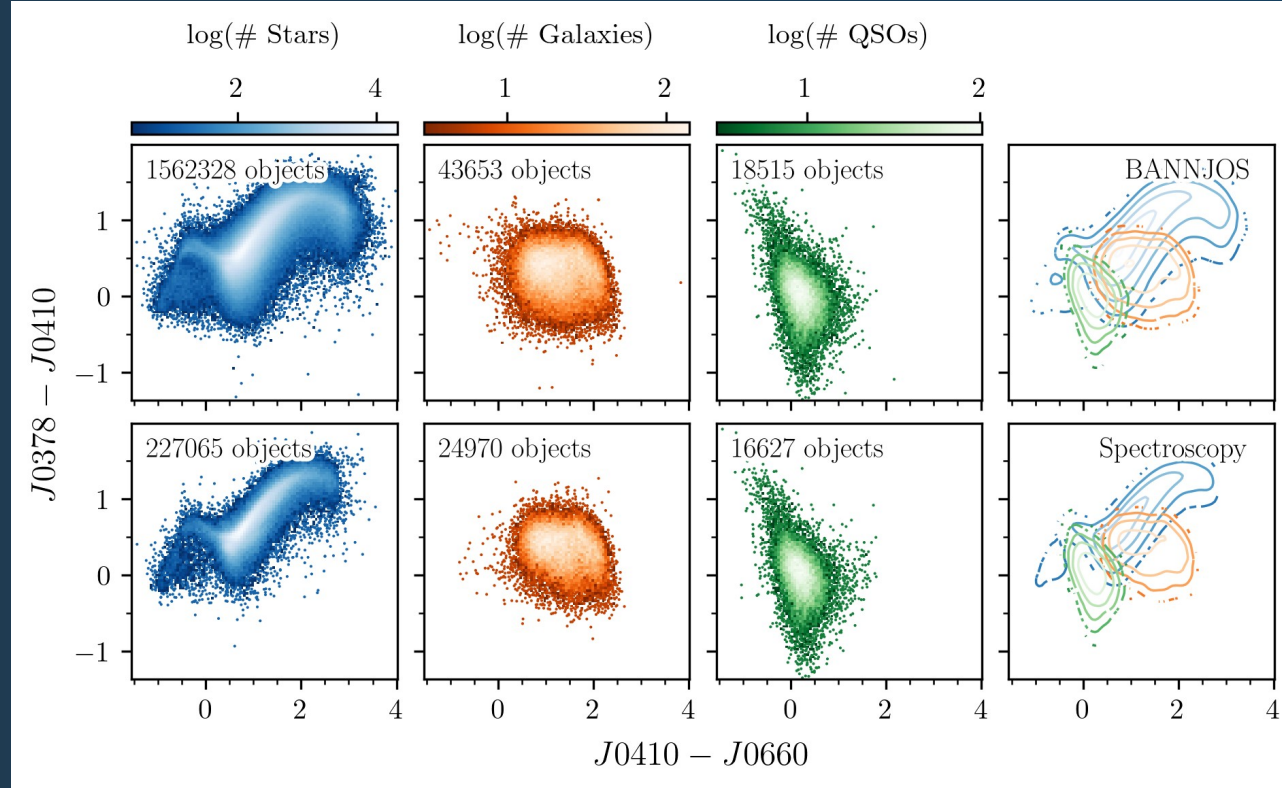
Model trained ©



BANNJOS: Star/Galaxy/QSO classification

Bayesian Artificial Neural Network for the Javalambre Observatory Surveys
(del Pino et al. 2024, A&A., 691, 221)

- ML pipeline that uses Bayesian NNs to provide full PDFs of the classification
- Trained on photometric, astrometric and morphological information from J-PLUS DR3, Gaia DR3, and CatWISE2020 over 1.2M objects with spectroscopy in SDSS DR18, LAMOST DR9, DESI EDR and Gaia DR3.



Applied to 47.4M objects in J-PLUS DR3
To be applied to J-PLUS DR4 too

- 95% accuracy for $r < 21.5$
- 90% accuracy for $r < 22.0$



Thank you for your attention!

Picture: Óscar Blanco Varela