Automated Data Exploration and Discovery of LOFAR Solar Observations

Kamen Kozarev (kkozarev@astro.bas.bg), Vergil Yotov, Maria Petkova, Vladimira Irincheva, Lyuben Branzalov, Oleg Stepanyuk – *Institute of Astronomy, Bulgarian Academy of Sciences*

Pietro Zucca, Mattia Mancini – *ASTRON, The Netherlands Institute of Radio Astronomy*

Abstract: The pan-European distributed LOFAR telescope is the biggest radio telescope in the world observing at the lowest accessible radio frequencies from the ground. LOFAR is very well suited for observing the solar corona and transient eruptive activity at frequencies below 250 MHz via interferometric imaging with unprecedented resolution in space, time, and frequency. Data from over 10 years of observations has been collected by the Solar Key Science Project (KSP) of LOFAR, which requires the development of automated pipelines for its discovery and analysis. In this work, we demonstrate our approach to automated data exploration and discovery of LOFAR Low-Band Antenna (LBA) solar observations between 20 and 90 MHz, applied to data catalogued by the Solar KSP / IDOLS team. We perform removal of low-quality observations, calibration and imaging, image stabilization, and segmentation of coronal features.

1. Introduction

The Low Frequency Array (**LOFAR**) is one of the largest radio telescopes in the world, a pan-European distributed telescope. LOFAR is designed to observe the universe at low radio frequencies, enabling scientists to study a variety of cosmic phenomena, including cosmic rays, pulsars, galaxies, and space weather. It is currently being updated to LOFAR2.0, doubling its capacity for continuous observations. **LOFAR-BG** is the Bulgarian station of the international LOFAR (Low-Frequency Array) network, planned for completion in late 2025. It will contribute to scientific research in fields such as radio astronomy, space weather monitoring, and solar studies. LOFAR-BG is part of Bulgaria's National Roadmap for Research Infrastructure, and once operational, it will help establish a more advanced radio astronomy community in the region. The objective of the **IDOLS** project is to **demonstrate space weather science and monitoring capabilities** using a single LOFAR station, with plans to gradually integrate upgraded dual-beam stations. The project aims to involve collaborators and International LOFAR partners in its development, including creating relevant science and monitoring tools. Additionally, IDOLS will be operational during the transition to LOFAR 2.0, ensuring continued contributions to space weather research and monitoring. **Here we outline our work towards automating data exploration and discovery in LOFAR spectral and interferometric solar observations, which will be applied in processing archival and future data products.**

2. Data Preparation

The tasks currently in progress are:

- Manual selection of good quality solar observational data to be used for algorithm development and tests
- Algorithm design for good/bad solar data recognition of multi-frequency LOFAR observations
- Algorithm design for quiet/active solar period recognition of multi-frequency LOFAR observations
- Image stabilization algorithm development applicable to multiple frequencies
- Removal of RFI from spectral data
- SDO/AIA EUV data selection corresponding to the radio imaging observations
- Stabilized data selection and labeling for quiet and active periods
- DL model design for automatic semantic segmentation of quiet time multi-frequency imaging LOFAR observations

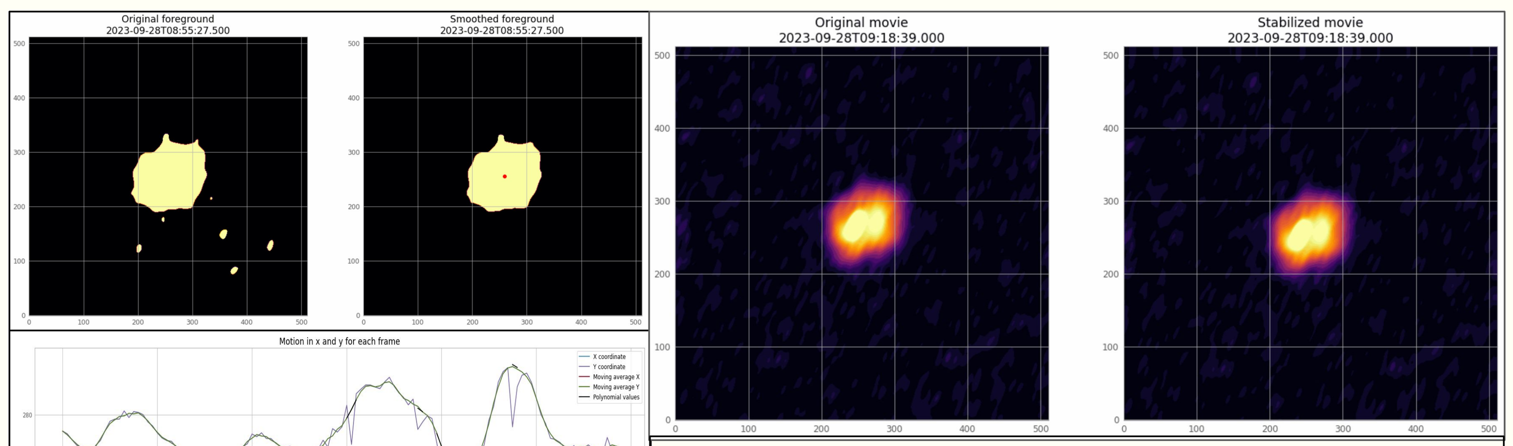


Fig. 1 - Steps in automated solar image stabilization. Images are smoothed and the motion of the center of mass is tracked then corrected for in the plane of sky. Result is in above panel.

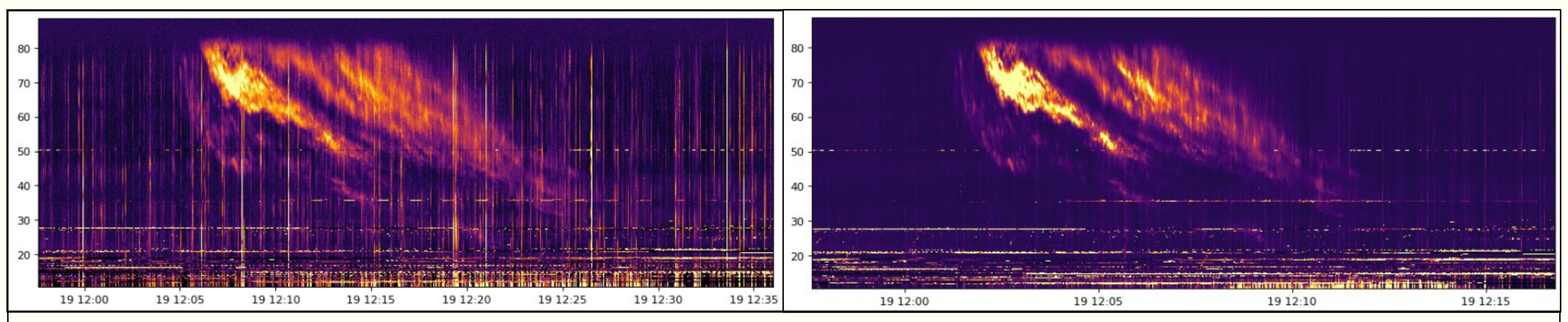
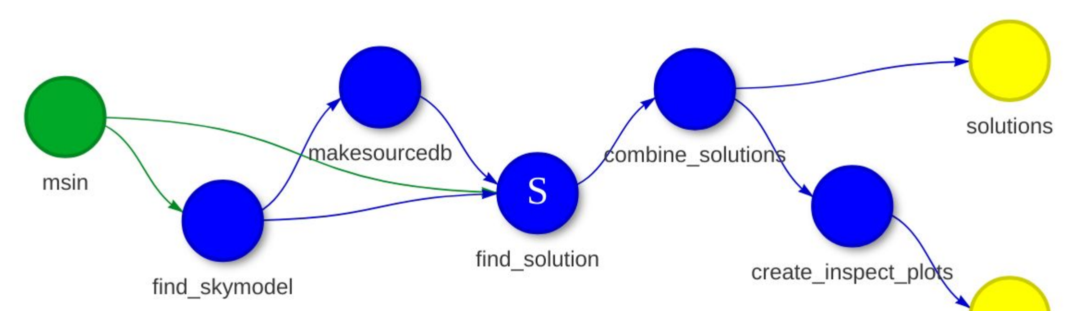


Fig. 2 – Removal of RFI in IDOLS dynamic spectra using an algorithmic approach and time averaging (Zhang et al. 2023).

3. Automated HPC processing of interferometric imaging data

- Application of modified LINC/IDOLS pipeline includes calibration and automated image generation from measurement set (MS) data.
- Currently application on Discoverer Petascale Supercomputer in Sofia.
- Combined with automated feature detection/extraction, will optimize processing and improve scientific product quality. Look-up images and video also produced.



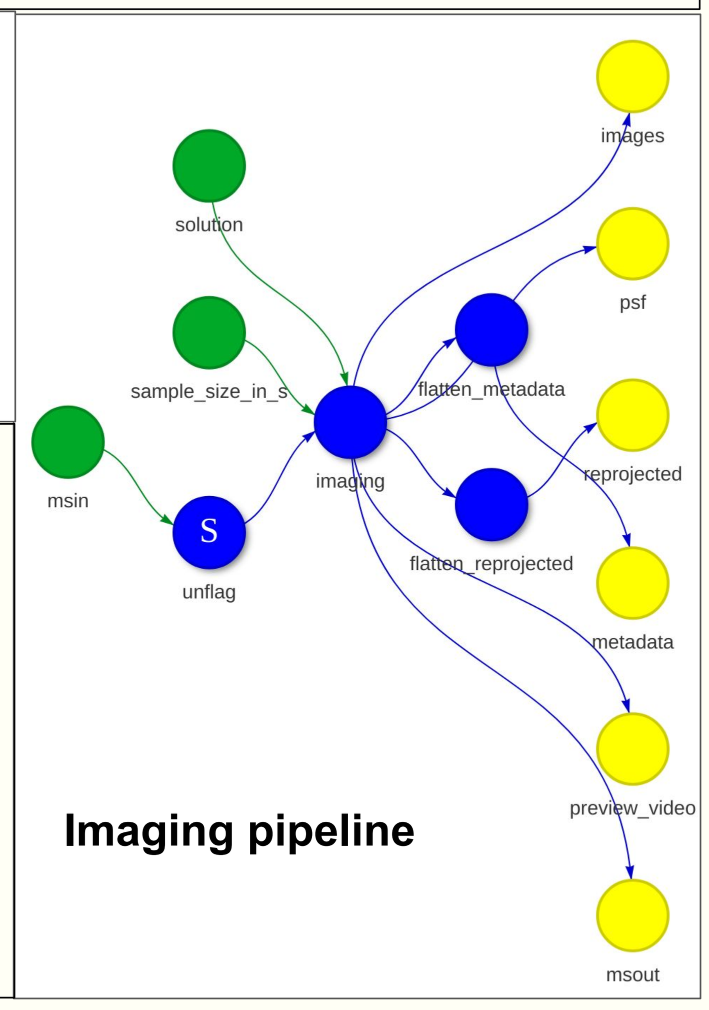


Fig. 3 – Calibration (top) and imaging (right) pipeline sections of the HPC workflow, being tested on Discoverer supercomputer. The measurement set (MS) data from the LOFAR long-term archive (LTA) is converted to a large number of flux-calibrated solar images.

4. Detection of time and frequency extent of emissions

- Algorithmic and ML approaches for detecting Type II and Type III burst regions in LOFAR dynamic spectra from IDOLS station
- Allows for developing smart imaging strategies for interferometric observations.

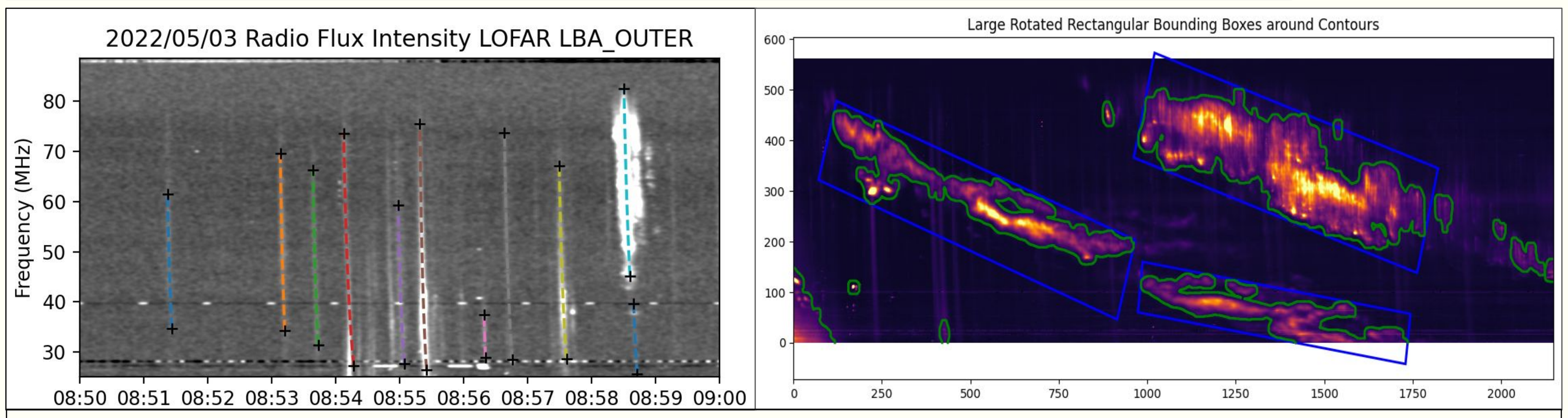


Fig. 4 – Automated detection of Type III (left panel) and Type II (right panel) radio bursts in dynamic spectra using computer vision approach allows for science, monitoring, DL dataset preparation, optimizing interferometric data processing and imaging.

5. Challenges:

- For the lower frequencies the image stabilization algorithm doesn't give great results.
- Ionospheric effects cannot be ignored for lower frequencies as the shape of the Sun becomes quite distorted. Improve the algorithms to determine the extent of ionospheric influence on the quality of the data.

6. Current Work

- **Detection and segmentation of solar radio features in radio images using EUV/white light observations**
- **Mapping of solar radio features to EUV and coronagraph images**
- **Translation of segmentation maps from high into low frequencies**
- **Automatic identification of the flaring onset in radio vs. EUV (both time series and imaging observations)**
- **Near-real time data processing and decimation of single station observations**

Morning Ionosphere | Midday Ionosphere

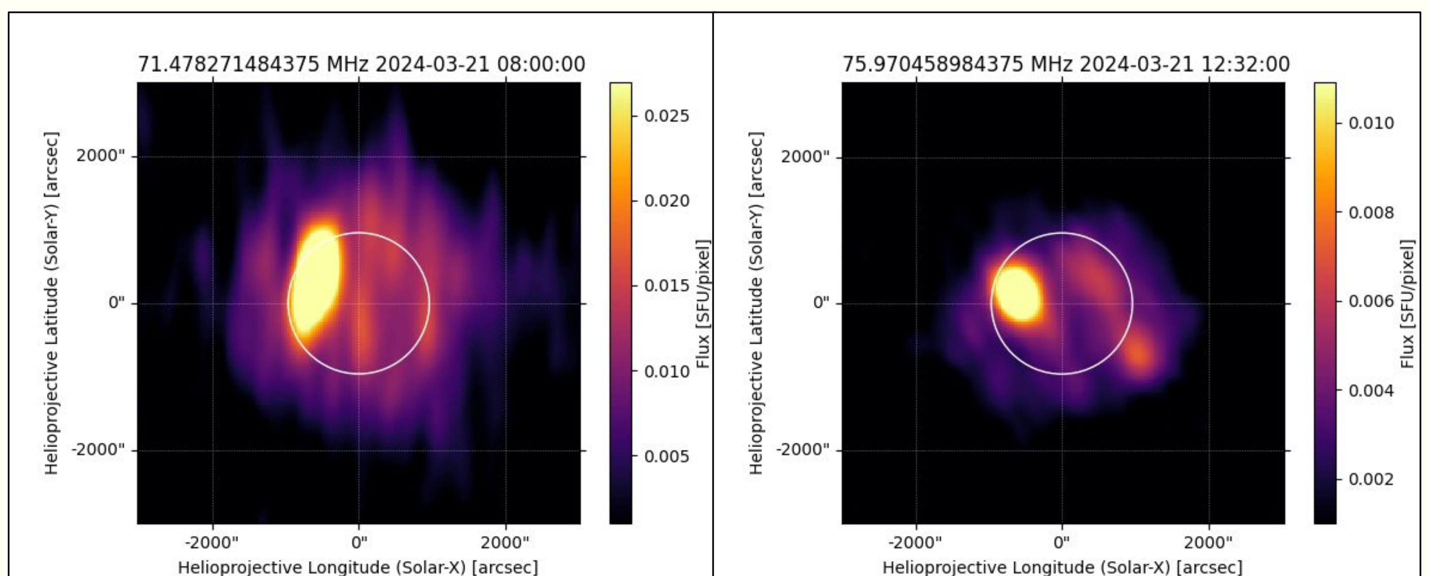


Fig. 5 – Solar interferometric snapshots from late morning and noon observations show the changing PSF and ionospheric conditions, which remain a challenge for continuous solar monitoring.

Acknowledgments: LOFAR ERIC defines and implements a common long-term strategy, joint fundraising, and consistent prioritisation of the development effort for LOFAR2.0, as well as optimising the availability of the collective partner and LOFAR ERIC owned facilities (including sensor, compute, and data storage resources) that form the LOFAR research infrastructure. LOFAR-BG is part of the National Roadmap for Scientific Infrastructure (2020-2027), coordinated by the Ministry of Education and Science of Bulgaria. The website of the Solar and Space Weather Key Science Project of LOFAR can be accessed at https://www.astron.nl/spaceweather/SolarKSP/

Zhang, P. et al. "RFI flagging in solar and space weather low frequency radio observations", MNRAS, 2023, 521, 630–637