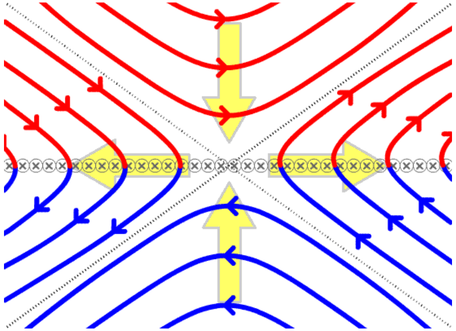
Automatic identification of magnetic reconnection events in 2D Hybrid Vlasov Maxwell simulations using Convolutional Neural Networks

A. Hu[1], M. Sisti[2,3], F. Finelli[2], F. Califano[2], J. Dargent[2], M.Faganello[3], E. Camporeale[1,4,5], J. Teunissen[1]

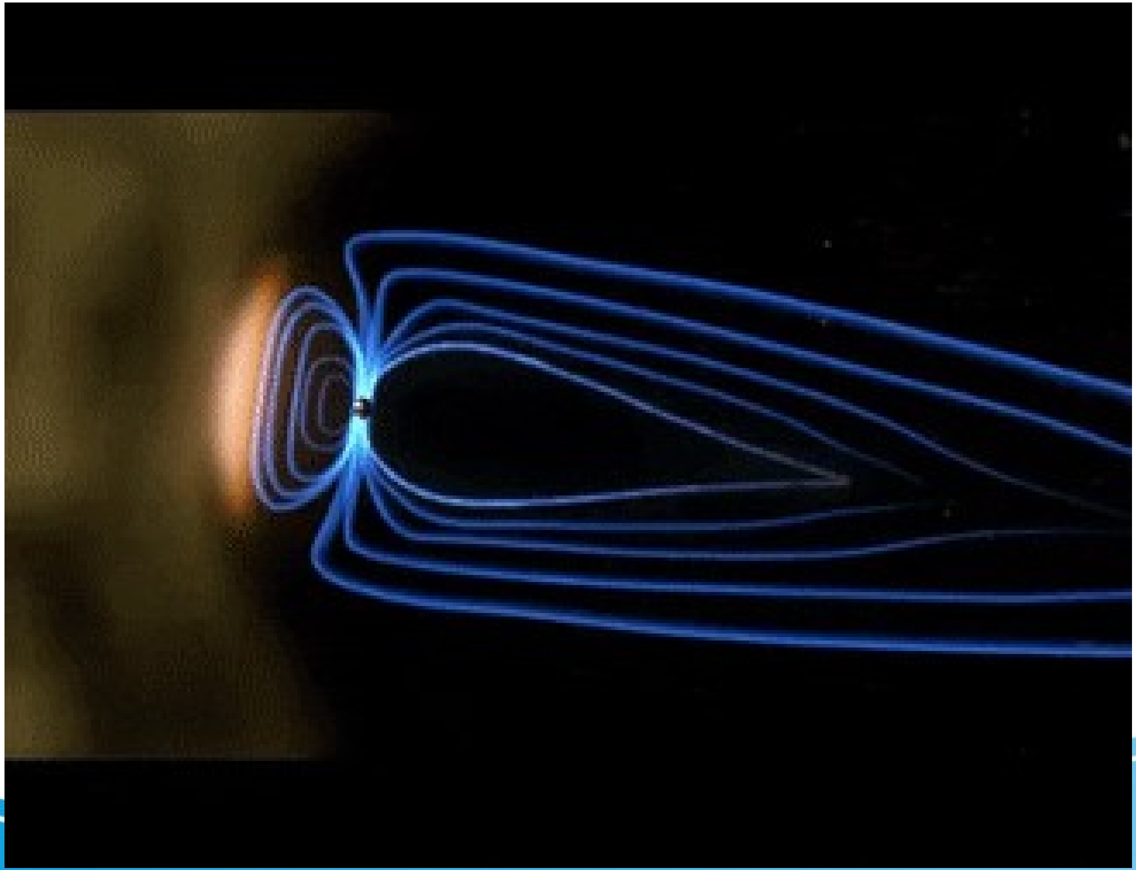
[1]Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
[2]Dipartimento di Fisica, Universita di Pisa, Italy, EU
[3]Aix-Marseille University, CNRS, PIIM UMR 7345, Marseille, France, EU
[4]CIRES, University of Colorado, Boulder, CO, USA
[5]NOAA Space Weather Prediction Center, Boulder, CO, USA



This research is developed in the framework of the European project AIDA (Artificial Intelligence Data Analysis)

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Significance

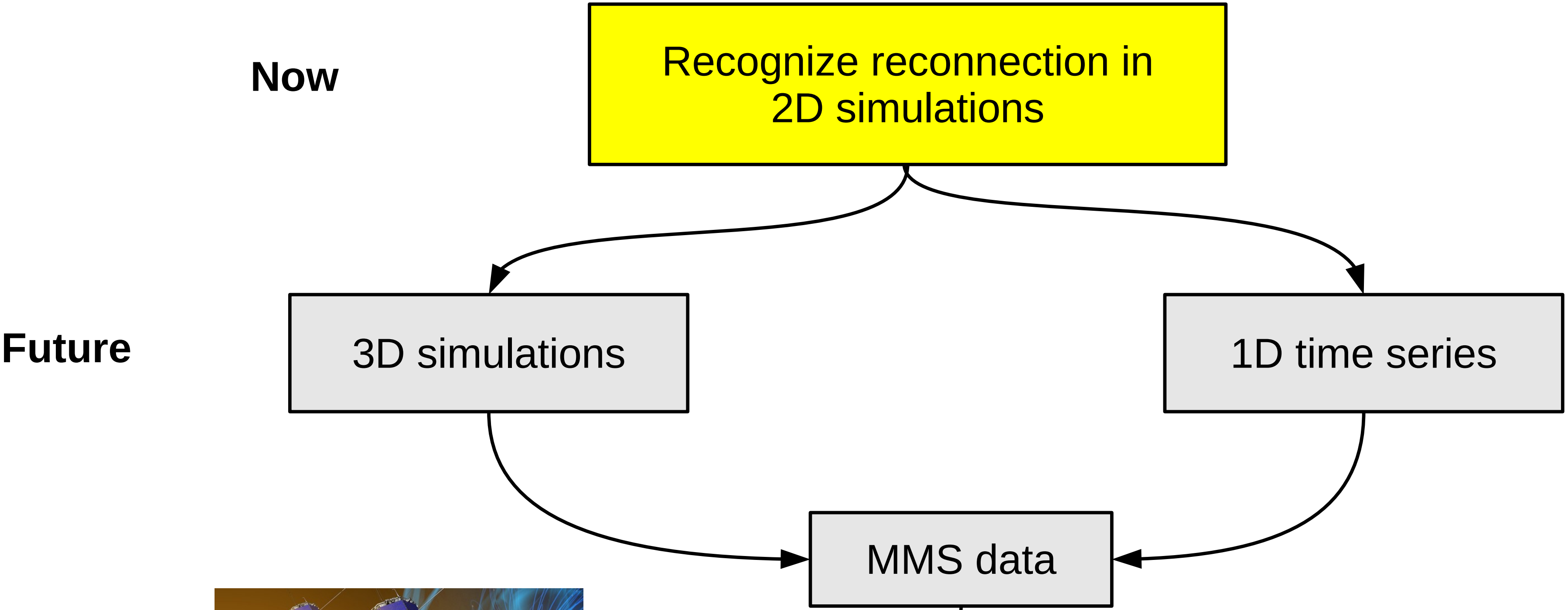


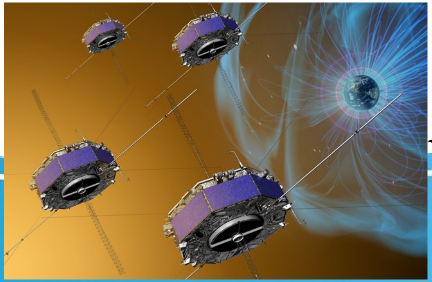
- From https://gfycat.com/thirdserpentinedrongo

- Magnetic reconnection is a fundamental process in space and laboratory plasmas in which magnetic energy is converted into kinetic energy, released in the form of accelerated particles, flows and heating.

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Goals

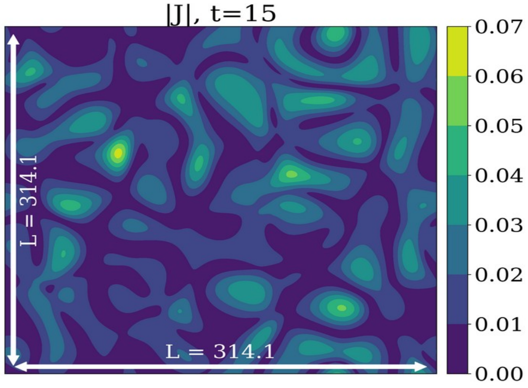




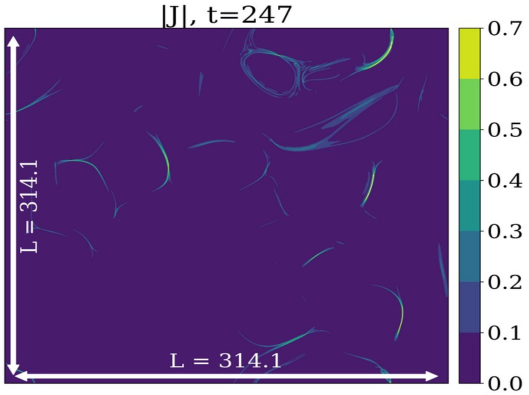
The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Current sheet evolution

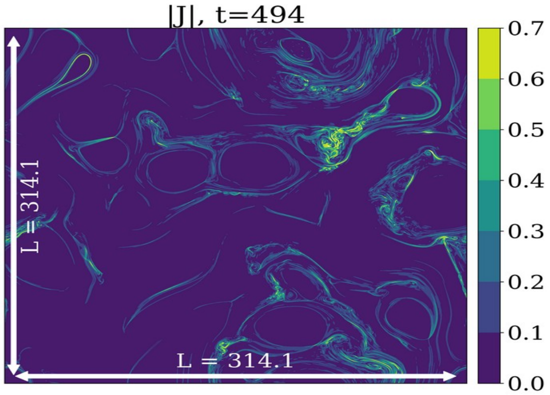
Initial phase



CS formation phase



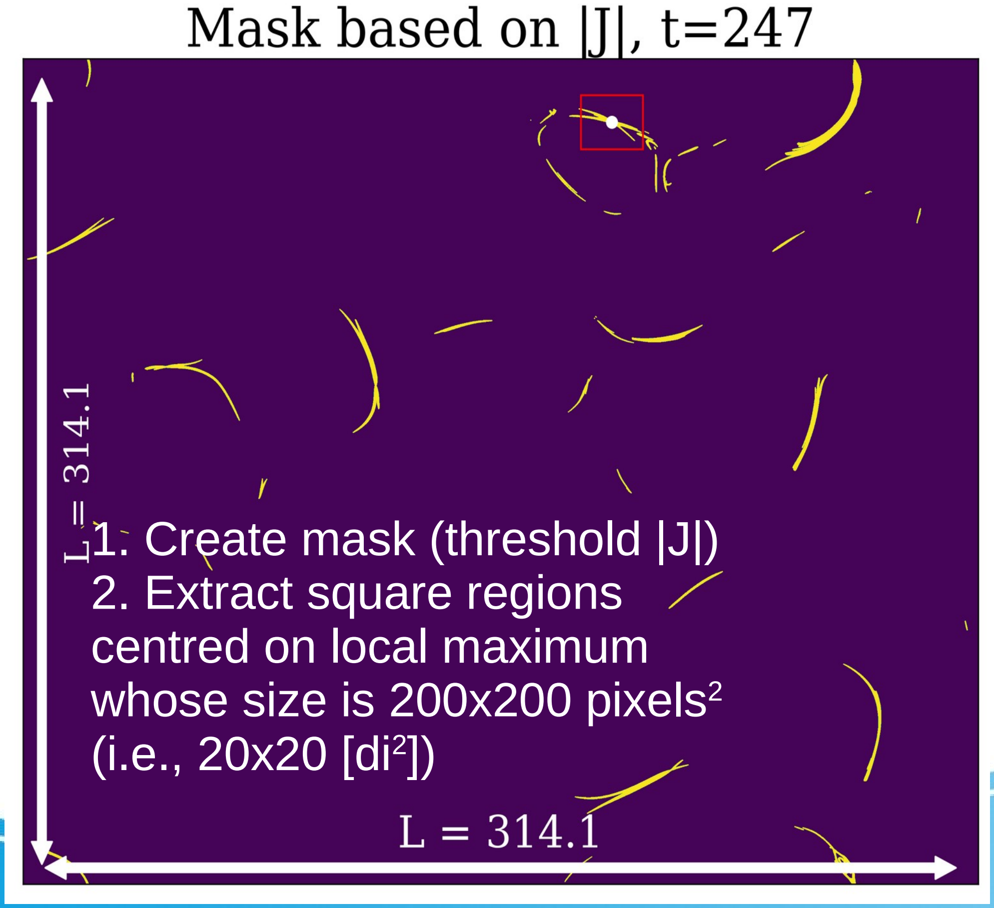
Fully turbulent phase



How the current sheets change with the time passing by in Sim 1.

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Data set creation

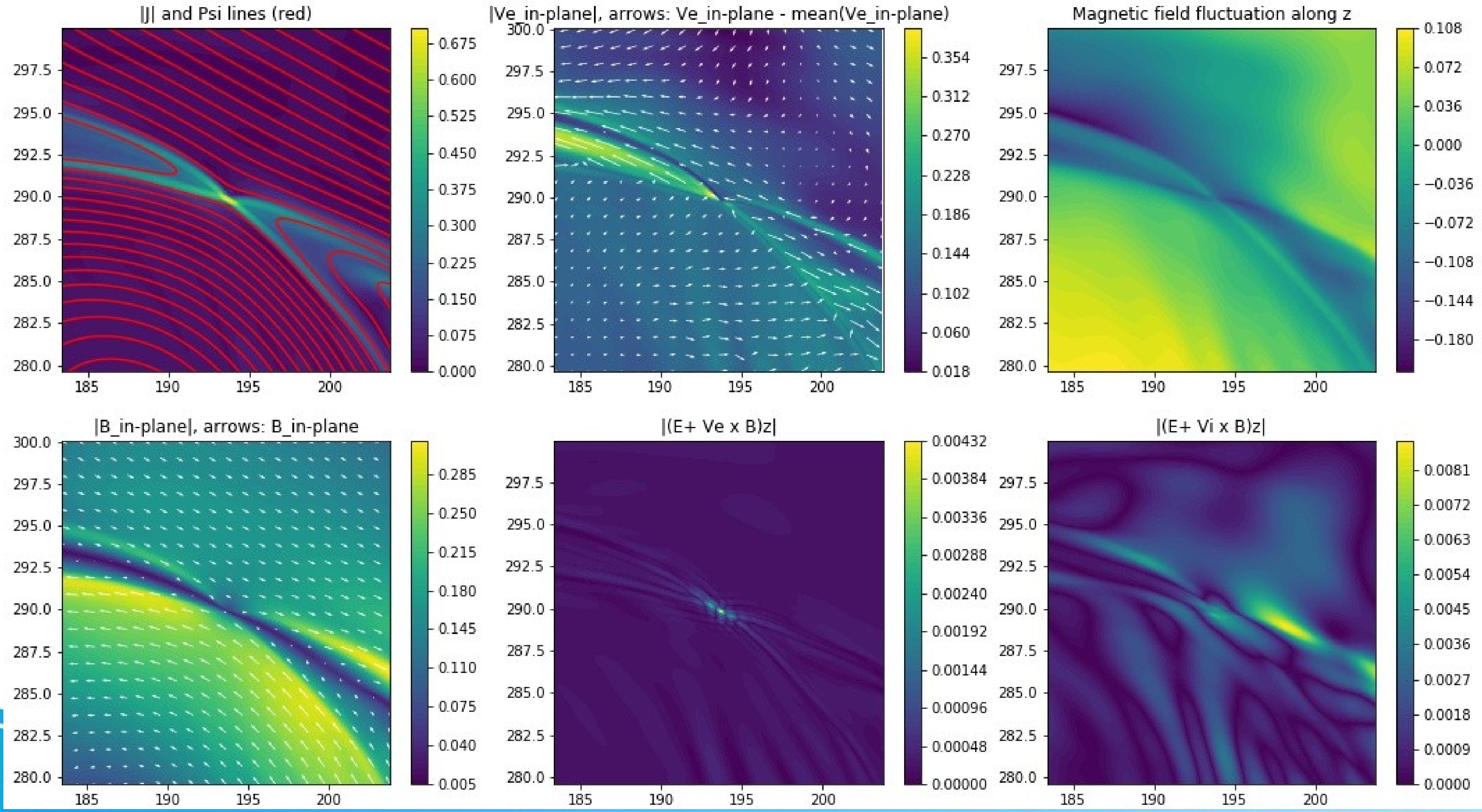


1. Create mask (threshold |J|)
2. Extract square regions centred on local maximum whose size is 200x200 pixels^2 (i.e., 20x20 [d_i^2])

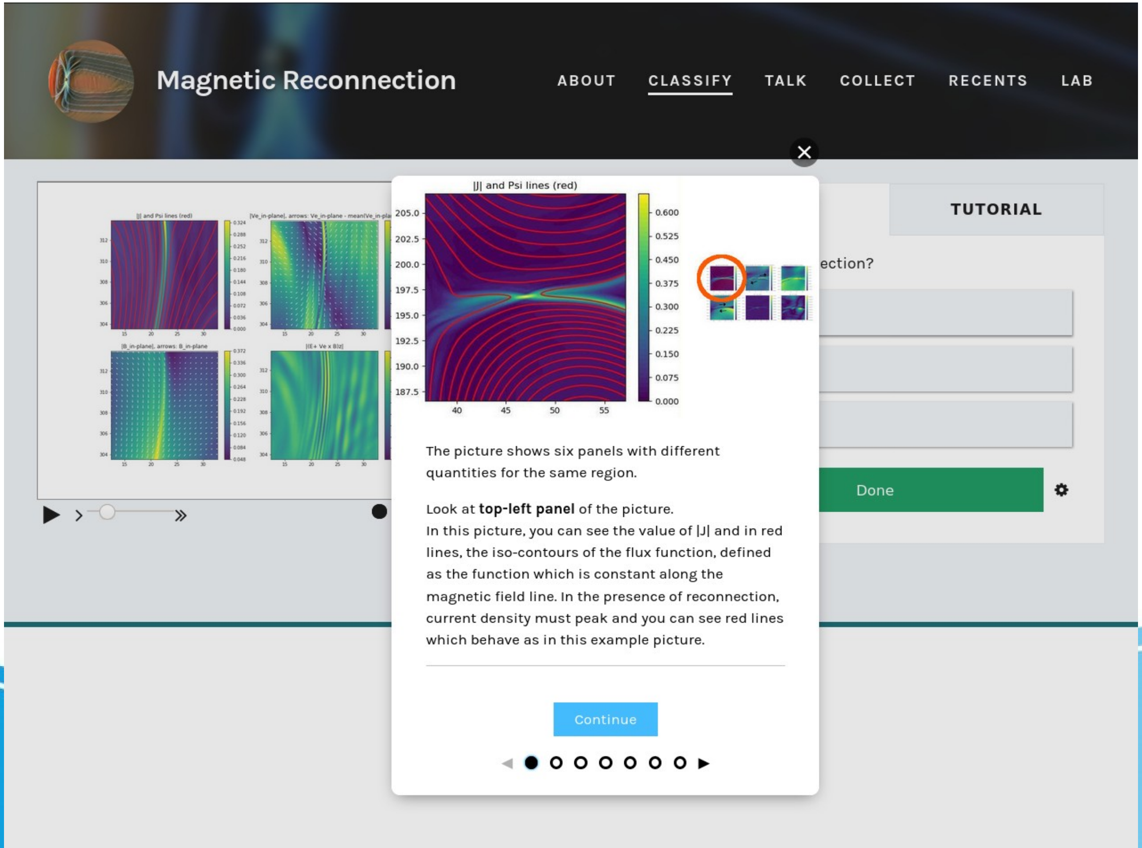
Variable name	Description
$\|\vec{J}\|$	L_2-norm of total current density $\vec{J}$
Ψ	flux function, $\vec{B} = \nabla\Psi \wedge \hat{z}$, $\vec{J} = -\nabla^2\Psi$
$V_{e,x}$	electron x-velocity
$V_{e,y}$	electron y-velocity
$V_{e,z}$	electron z-velocity
$V_{e,\text{plane}}$	$\sqrt{V_{e,x}^2 + V_{e,y}^2}$
B_z	z-component of magnetic field
B_{plane}	$\sqrt{B_x^2 + B_y^2}$
$E_{\text{dec},e}$	$(\vec{E} + V_e \times \vec{B})_z$ (decoupling electrons)
$E_{\text{dec},i}$	$(\vec{E} + V_i \times \vec{B})_z$ (decoupling ions)

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Variables in a selected region for human labelling



Human Labelling on-line Platform



As thousands of cases have to be classified, an automated workflow for labeling by human experts is important. For this, a project on Zooniverse has been created on **zooniverse.org/projects/taiyexingshang/magnetic-reconnection**

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

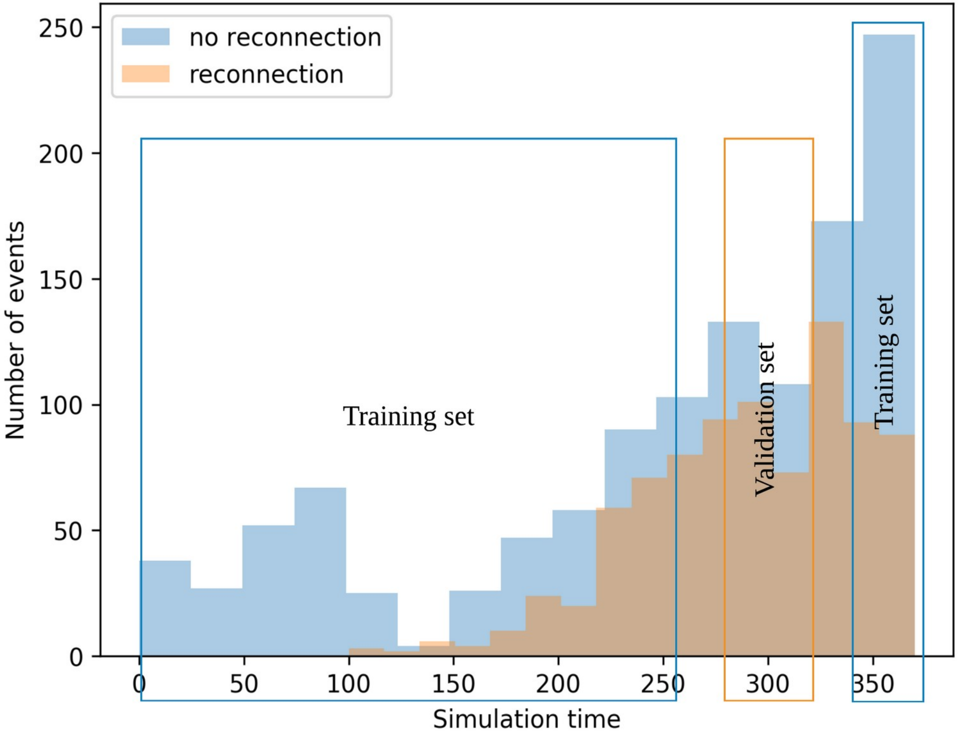
Simulations performed at CINECA on Marconi

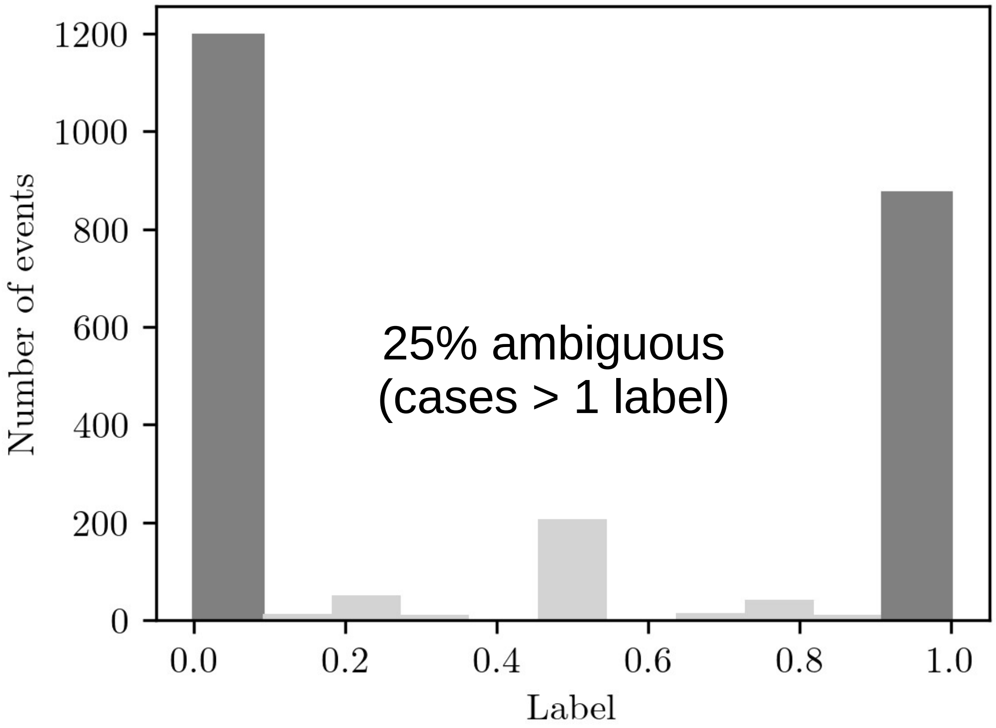
- 2D Hybrid Vlasov-Maxwell model
  - Ions: Vlasov (distribution function not yet used)
  - Electrons: fluid

Name	description	grid size	dl/d_i	N_{samples}	% reconnection	time range (Ω_{ci}^{-1})
Sim 1	all data	3072^2	0.1	2069	42 %	[0, 370]
	training set			1205	34.7 %	[0, 260], [340, 370]
	validation set			437	56 %	[280, 320]
Sim 2	test set	2048^2	0.15	124	56.5 %	[205, 233]

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Statistics Sim 1



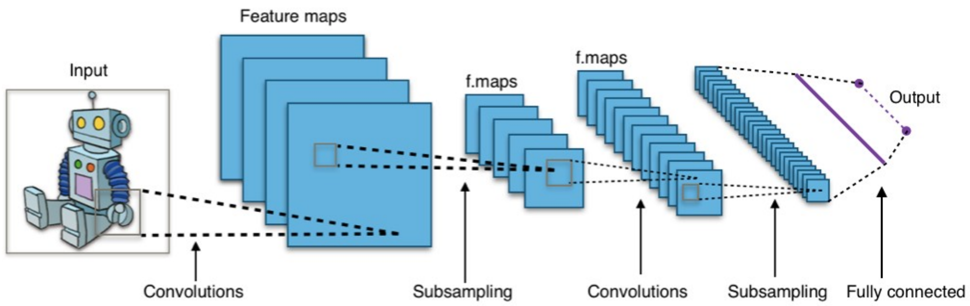


25% ambiguous
(cases > 1 label)

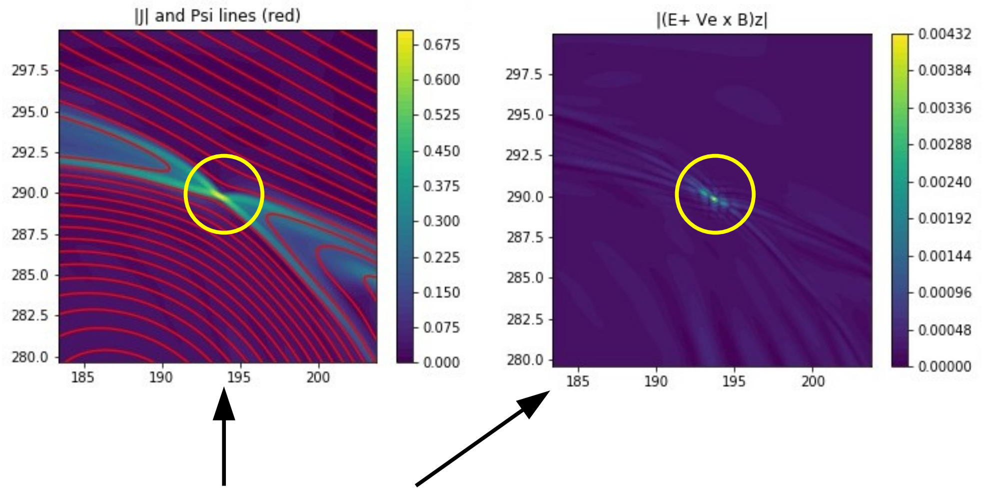
train/test split in time

Machine learning models

Convolutional Neural Network (CNN)



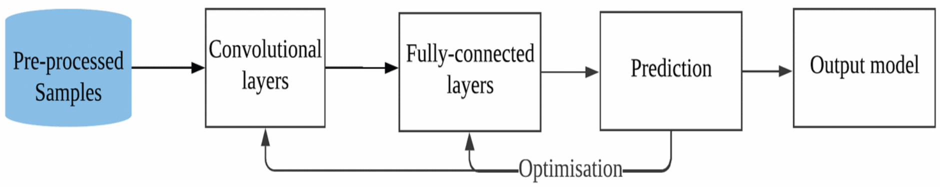
Heuristic model



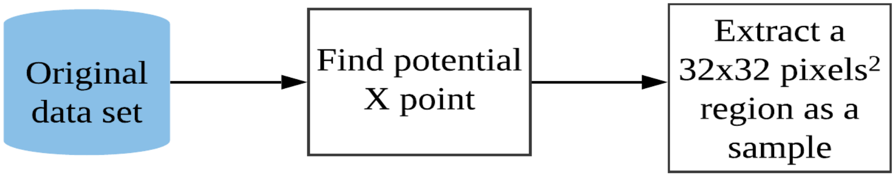
1. Detect peaks
2. Extract min/max/mean around peaks if they are close together
3. Use decision tree

Modelling methods

CNN-X Architecture



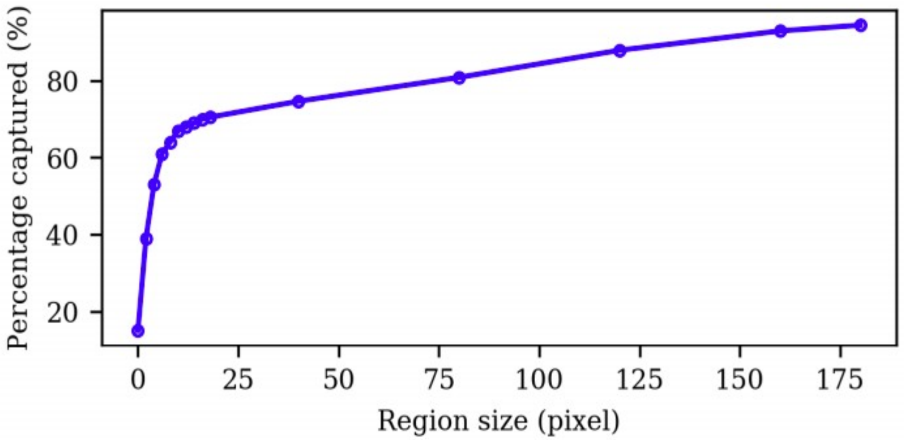
Feature engineering (X=32)



Details can be found in **Fig. 6 & 7** of the paper "**Hu, A., Sisti, M., Finelli, F., Califano, F., Dargent, J., Faganello, M., ... & Teunissen, J. (2020). Identifying Magnetic Reconnection in 2D Hybrid Vlasov Maxwell Simulations with Convolutional Neural Networks. The Astrophysical Journal, 900(1), 86.**" (https://arxiv.org/pdf/2008.09463.pdf)

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

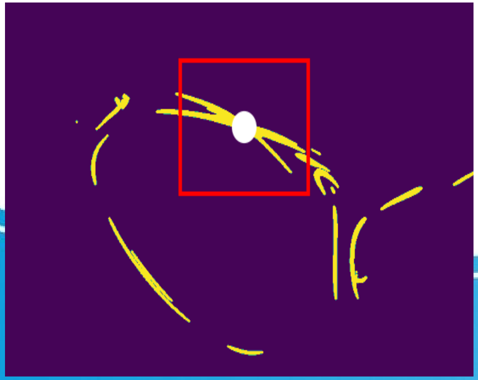
Analysis



Percentage of reconnection sites captured versus window size for the image cropping approach

Window size	TSS	MCC	TP	FP	TN	FN
16	0.29	0.32	85	**12**	**176**	158
32	**0.56**	**0.55**	**170**	28	161	**70**
64	0.42	0.41	138	29	159	103
128	0.43	0.44	133	23	166	108
200	0.39	0.44	154	48	141	87

Accuracy of the CNN-X models with different window sizes X.



The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

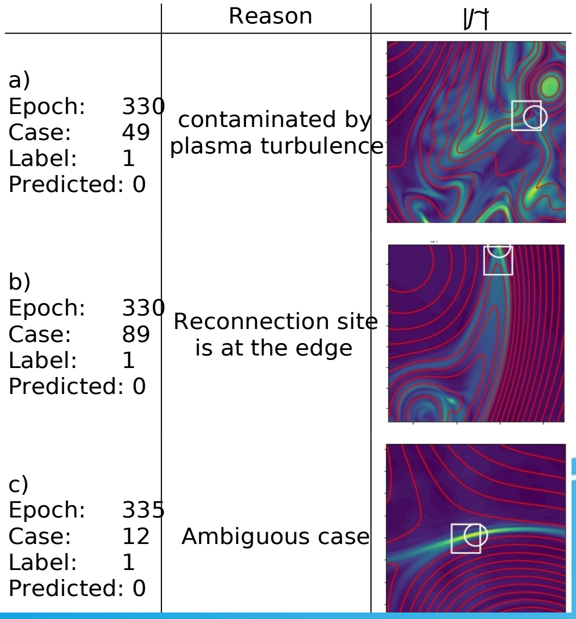
Results

Model	TSS	MCC	TP	FP	TN	FN
CNN-32	0.50	0.51	48	10	43	22
Decision tree	0.28	0.30	55	27	26	15

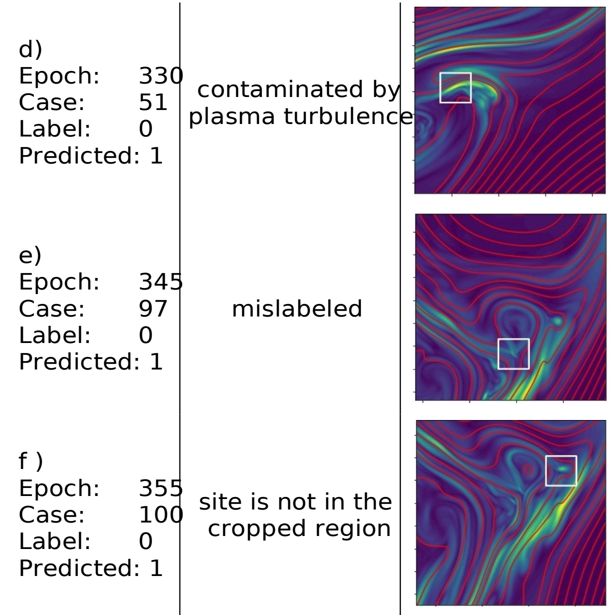
Accuracy of the machine learning models evaluated on test data set.

1 variable	MCC	2 variables	MCC	3 variables	MCC
$\|\mathbf{J}\|$	0.44	$\|\mathbf{J}\|$, B_{plane}	0.51	$\|\mathbf{J}\|$, $V_{e,z}$, B_{plane}	0.56
$V_{e,z}$	0.39	$\|\mathbf{J}\|$, $V_{e,z}$	0.49	$\|\mathbf{J}\|$, B_{plane}, $E_{\text{dec},e}$	0.55
$V_{e,plane}$	0.13	$\|\mathbf{J}\|$, $E_{\text{dec},e}$	0.44	$\|\mathbf{J}\|$, $V_{e,z}$, Ψ	0.50

MCC scores of CNN-32 models that only take the listed variables as input.



Examples of false negatives (rows a-c) and false positives (rows d-f).



Conclusion

1. The image cropping method can improve the accuracy of the CNN models by increasing the signal-to-noise ratio.

2. The developed CNN-32 model is generic and can be applied to other simulations. Furthermore, in some cases, the CNN-32 model was able to find reconnection sites that were initially missed by a human expert.

3. Three variables were found to be the most important reconnection markers: the current density |J|, the out-of-plane electron velocity $V_{e,z}$ and the in-plane magnetic field B_{plane}.

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Outlook

This study is a first step in adopting machine learning for the automatic identification of magnetic reconnection. We think that with more labeled data from different types of simulations the model's accuracy would improve.

Meanwhile, we are working on identifying reconnection events from 1D simulations based on this labelled data set in order to further implement this model into real satellite measurements. This is because 1D simulations are more similar to satellite trajectories.

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.