



signals-and-markers-toolkit



Pregled

Klemen Lovenjak · GeoDev Meetup · 2025-04-10

AGRICULTURE · Chivilcoy, Argentina · February 9, 2022

Nadzor kmetijskih zemljišč

- Izplačilo subvencij
- Občutljiva tematika
- Debata o upravičenosti takšnega nadzora
- Debata o ustreznosti podatkov in algoritmov
- Kako poteka izračun?
- OpenSource?

Vi ste naredli sopotnika?

Zapret bi vas mogli takoj

[REDACTED] kdo si je to zmislo

Ce je glupo za popravit

Drugo ni, kot samo dodatno delo pa problemi

Neki 1984 bigbrother bulshit

Pol pa morem hodit na svetovalno pa pojasnjevat, da je na bregi, ko ze 5000 let samo trava rase, indeed samo trava

Area Monitoring @ Planet

- ~30 ljud
 - PM
 - DS
 - BE (Python, Java)
 - FE (React)
- Sentinel Hub -> Planet Insights Platform
- Vloga signals-and-markers-toolkit



Naloga

Monitoring kmetijskih zemljisc

Podlaga za odločanje o upravičenosti

Prilagodljiva logika: različni sistemi, različne zahteve

Število parcel: rang 10 mio

Več projektov, pogoste iteracije, striktni roki

Velikostni red (samo Parquet dataset s podatki):

- povprečna iteracija: 100 GB
- Izredni projekti: 1.5 TB

Vektorji, ne rastri!

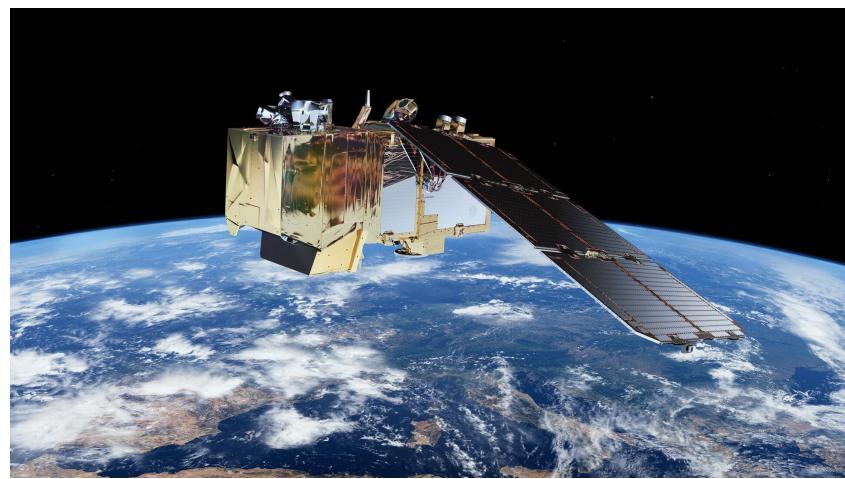
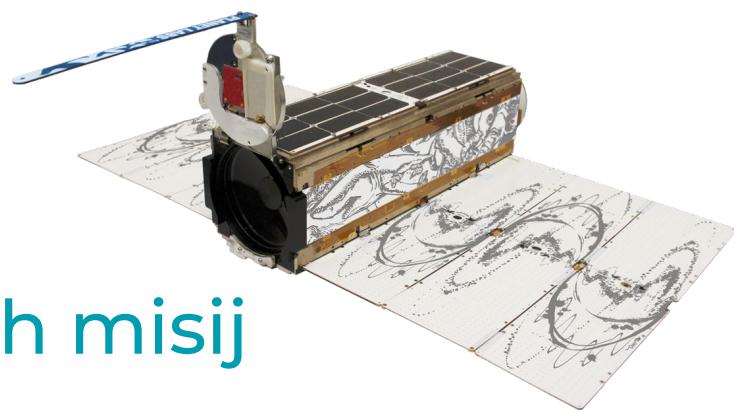
Zaenkrat ni potrebe po procesiranju v realnem času.

Splošno, zmogljivo in zanesljivo orodje*



Izhodišče

Produkti satelitskih misij



Produkt	Planet Fusion (PF)	Analysis-Ready PlanetScope (ARPS)	Sentinel-2 L2A (S2L2A)
Prostorska ločljivost	3,5 m	3 m (prevzorčeno)	10 m
Spektralni pasovi	4 pasovi	4 pasovi	13 pasov
Vir	Planet Labs	Planet Labs	ESA
Časovna ločljivost	Dan (??)	“Skoraj” dan	5 dni
Dostopnost	Plačljivo	Plačljivo	Brezplačno*





Orodja

Organska rast.

Python: Simpatičnost, odličen ekosistem

Amazon Web Services (AWS): Količina podatkov, kompleksnost obdelav

Dev Container: Večne težave z okolji, enostavno prehajanje med lokalnim in “oddaljenim” okoljem

Ray: Skaliranje (“lestvičenje!”)

Prefect: Orkestracija tokov, lep UI brez dodatnega truda

“Trial and error”

AWS

- S3 (Simple Storage Service):
 - Velike količine
 - Vsi podatki: tudi vmesni izračuni.
 - Rate limiting?
- EC2 (Elastic Compute Cloud):
 - Ni bistvene razlike ob delu
 - Potreba po različnih instancah
 - “In-house” CLI app za upravljanje
 - Dosegljivost instanc, SPOT



```
~ > amin

AMIN status:
General
Amin version: 25.3.2
Host OS: ubuntu
Git credentials manager: libsecret

AWS credentials status
Session status: Logged in

Amin worker state
Selected worker: local
Instance state: Not available
Worker status: worker connected

Worker stack
Local worker has no stack

Workspaces
Name Container status Workspace folder
eu-openem devcontainer destroyed ~/workspace/eu-openem-workspace
it-avepa devcontainer destroyed ~/workspace/it-avepa-workspace
markers devcontainer stopped ~/workspace/markers-workspace
nl-nea devcontainer stopped ~/workspace/nl-nea-workspace
si-aktrp devcontainer stopped ~/workspace/si-aktrp-workspace

~ > amin worker create
Worker name [worker_cli_code]: worker_cli_demo
Instance type [c5.4xlarge]:
Volume size [500]:
? Select one of the availability zones. (Use shortcuts or arrow keys)
  1) eu-central-1a
» 2) eu-central-1b
  3) eu-central-1c
```

Devcontainer

- VisualStudio Code
- Odprtokodna specifikacija
- Izolirano okolje, docker slika
- Odlična integracija z oddaljenimi okolji (vtičnik Remote - Containers)

Prefect

- Orkestracija flowov
- Self-hosted
- Zmogljiv UI (Pydantic)
- Uporabni vtičniki
- ...
- Križi in težave

```
@task(name="Report Images Used per BYOC SH Collection", task_run_name=get_report_task_run_name, **TASK_SETTINGS)
async def report_images_used_per_byoc_collection(
    sh_code: SentinelHubCollectionCodes,
    collection: CollectionToReport,
    report_base_path: str,
    sh_credentials: SentinelHubCredentials,
) -> None:
    """Generates a report on which COGs were used for a given SH collection."""

    images = get_cogs_used_in_byoc_collection(collection=collection, sh_credentials=sh_credentials)

    assume_role_arn = get_assume_role_arn(report_base_path)
    filesystem = get_filesystem(report_base_path, assume_role_arn=assume_role_arn)
    filesystem.write_text(f"{report_base_path}/{sh_code}.json", json.dumps(images, indent=4))

@flow(name="Report Images Used in Signal Packages", **FLOW_SETTINGS)      Matic Pecovnik, 2 months ago • Add retri
async def report_images_used_in_signal_packages(
    scope: str, signal_packages: list[signals.SignalPackageDto], report_base_path: str, geometry: Geometry
) -> None:
    """Iterates over the signal packages and generates a report on which COGs were used for each one."""

    sh_credentials = signals.get_sh_credentials(scope=scope)

    inputs = await prepare_input_data.submit(signal_packages=signal_packages)
    futures: list[PrefectFuture[None, Async]] = []
    for sh_code, collection in (await inputs.result()).items():
        if sh_code in get_args(PublicSentinelHubCollectionCodes):
            fut = await report_images_used_per_public_collection.submit(
                sh_code=sh_code,
                collection=collection,
                report_base_path=report_base_path,
                geometry=geometry,
                sh_credentials=sh_credentials,
            )
        else:
            fut = await report_images_used_per_byoc_collection.submit(
                sh_code=sh_code,
                collection=collection,
                report_base_path=report_base_path,
                sh_credentials=sh_credentials,
            )
        futures.append(fut)

    await asyncio.gather(*[fut.result() for fut in futures])
```

Deployments

16 Deployments



Q Deployment names All tags A to Z

Deployment name	Flow name	Schedule	Tags	Activity
External Job: NL-NEA Pixel Count Transformation from Exported Dataset • Created 2024/08/20 02:11:02 PM	External Job: NL-NEA Pixel Count Transformation from Exported Dataset			
External Job: Pixel-Level Signals Calculation • Created 2024/10/16 11:53:08 AM	External Job: Pixel-Level Signals Calculation			
External Job: Publish Datasets • Created 2025/02/03 12:11:40 PM	External Job: Publish Datasets			
External Job: Reference data transformation • Created 2024/08/20 02:11:02 PM	External Job: Reference data transformation			
General: 0 - Scope Creation and Configuration • Created 2024/08/20 12:14:50 PM	General: 0 - Scope Creation and Configuration		slack-notification	
General: 1 - Reference Data Ingest, Export and Transformation • Created 2024/08/20 12:14:50 PM	General: 1 - Reference Data Ingest, Export and Transformation		slack-notification	
General: 2 - Reference Data Ingest, Export, Transformation and Signal Download • Created 2025/02/12 11:09:04 AM	General: 2 - Reference Data Ingest, Export, Transformation and Signal Download		slack-notification	
General: 3 - Reference Data Ingest, Export, Transformation, Signal Download and Marker Computation • Created 2025/02/12 11:09:03 AM	General: 3 - Reference Data Ingest, Export, Transformation, Signal Download and Marker Computation		slack-notification	
Project: NL-NEA • Created 2024/08/20 12:14:50 PM	Project: NL-NEA		nl-nea-notification slack-notification	
Project: SI-AKTRP • Created 2024/10/25 06:09:03 AM	Project: SI-AKTRP		si-aktrp-notification slack-notification	
Test: End-to-End • Created 2025/03/31 02:41:08 PM	Test: End-to-End		slack-notification	
Utility: Class group mapping JSON coverage • Created 2025/02/03 03:44:17 PM	Utility: Class group mapping JSON coverage		slack-notification	
Utility: Class label mapping JSON coverage • Created 2025/02/03 03:44:16 PM	Utility: Class label mapping JSON coverage		slack-notification	
Utility: Mapped reference GeoPackage group count • Created 2024/08/20 02:48:09 PM	Utility: Mapped reference GeoPackage group count		slack-notification	
Utility: Mapping JSON validation • Created 2024/08/20 02:48:09 PM	Utility: Mapping JSON validation		slack-notification	
Utility: Reference GeoPackage validation • Created 2024/08/20 02:48:09 PM	Utility: Reference GeoPackage validation		slack-notification	

Flow Runs

/ General: 3 - Reference Data Ingest, Export, Transformation, Signal Download and Marker

Computation - Scope=ANCODEV_AI4SOILHEALTH_GAGEL_DEMO - Signals End

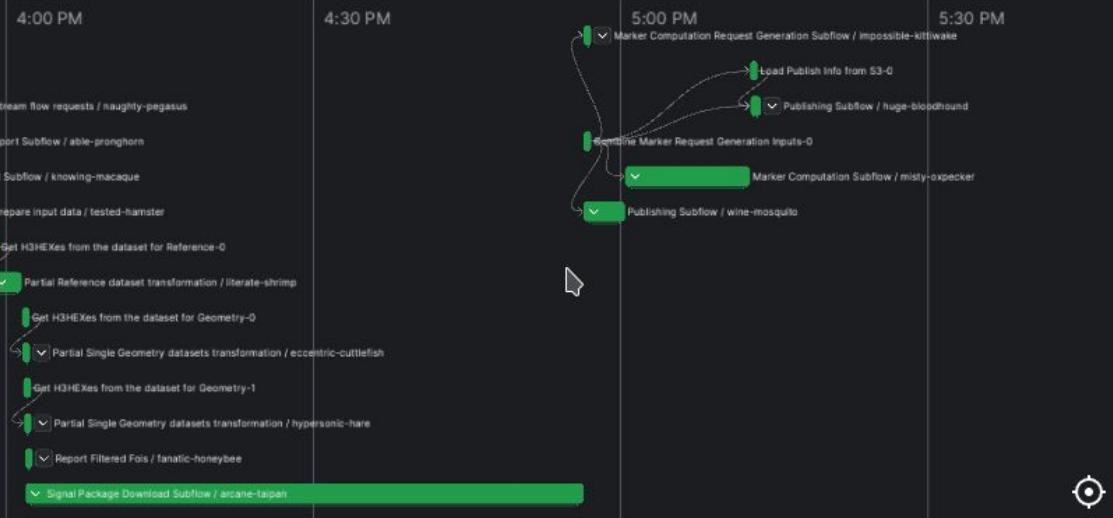
Date=2024-10-31 - Run Name=official-Gagel_farm

slack-notification

Completed 2025/04/02 03:49:31 PM 1h 24m 5 task runs

Flow General: 3 - Reference Data Ingest, Export, Transformation, Signal Download and Marker Computation

Deployment General: 3 - Reference Data Ingest, Export, Transformation, Signal Download and Marker Computation



Logs Task Runs Subflow Runs Results Artifacts Details Parameters

Level: all

Oldest to newest

Apr 2nd, 2025

INFO Runner 'runner-bd84cc0a-33be-44cd-816f-6bee7ba7f129' submitting flow run 'aaf799ca-070e-49ec-9501-9e1158ba4c18' 03:49:27 PM prefect.flow_runs.runner

INFO Opening process... 03:49:27 PM prefect.flow_runs.runner

INFO Completed submission of flow run 'aaf799ca-070e-49ec-9501-9e1158ba4c18' 03:49:27 PM prefect.flow_runs.runner

INFO Downloading flow code from storage at '.' 03:49:29 PM prefect.flow_runs

INFO Created subflow run 'naughty-pegasus' for flow 'Prepare downstream flow requests' 03:49:31 PM prefect.flow_runs

INFO Created subflow run 'able-pronghorn' for flow 'Refdata FOI import Subflow' 03:49:37 PM prefect.flow_runs

INFO Created subflow run 'knowing-macaque' for flow 'Dataset export Subflow' 03:49:41 PM prefect.flow_runs

INFO Created subflow run 'tested-hamster' for flow 'Prepare input data' 03:49:44 PM prefect.flow_runs

INFO Created task run 'Get H3HEXes from the dataset for Reference-0' for task 'Get H3HEXes from the dataset for Reference' 03:58:36 PM prefect.flow_runs

INFO Executing 'Get H3HEXes from the dataset for Reference-0' immediately... 03:58:36 PM prefect.flow_runs

INFO Finished in state Completed() 03:58:36 PM Get H3HEXes from the dataset for Reference-0 prefect.task_runs

INFO Created subflow run 'literate-shrimp' for flow 'Partial Reference dataset transformation' 03:58:37 PM prefect.flow_runs

INFO Created task run 'Get H3HEXes from the dataset for Geometry-0' for task 'Get H3HEXes from the dataset for Geometry' 04:01:38 PM prefect.flow_runs

INFO Executing 'Get H3HEXes from the dataset for Geometry-0' immediately... 04:01:38 PM prefect.flow_runs

Deployments

/ General: 3 - Reference Data Ingest, Export, Transformation, Signal Download and Marker Computation

/ Run

Flow General: 3 - Reference Data Ingest, Export, Transformation, Signal Download and Marker Computation

The enhanced deployment parameters UI improves support for lists, default values, validation, and more. If you're having trouble, please report an issue. To revert to the previous deployment parameters UI click here.

Run Name

upsilon5-morgulduin-adjunct

Parameters

params

Input parameter for a full SMT flow. It consists of reference data ingestion, export, transformation, signal download and marker computation.

Marker Computation Configuration

Marker computation parameters

GitLab Project Git Branch (Optional)

Name of the branch in gitlab_project repository used for marker calculation. Note: Should be left blank if the user is unfamiliar with the concept of providing custom computation logic.

GitLab Project (Optional)

Name of the associated GitLab project. Note: Should be left as 'markers' if the run requires no other GitLab project containing custom computation logic.

markers

EC2 Computation Instance Types (Optional)

List of AWS EC2 instance types that are used for calculation. The tooling chooses first available instance type from the list. Note: Should be left blank if user has no special preferences, the appropriate type is deduced in this case.

No items in this list

+ Add item

Marker Computation Run Config (Optional)

Run config for marker computation. Specifies the markers to be calculated and their configurations.

Run Name (Optional)

Unique ID for the specific marker computation run. If the Signal Download Configuration changes, e.g. user downloads additional signals by prolonging the TOI for a signal package, the run name can be reused. In case of markers recomputation using tweaked parameters or additional markers, a new run name must be chosen. Otherwise the flow will fail.

Marker Configurations (Optional)

Composition of all the marker configurations utilized in the run.

Type: Mowing (Optional)

Configurations for the Mowing marker.

Source: PF (Optional)

No items in this list

+ Add item

Source: ARPS (Optional)

No items in this list

+ Add item

Source: S2L2A (Optional)

Drop Ratio (Optional)

Relative measure in what the drop in the NDVI is required. In absolute terms it is calculated as $(\delta_{thr} * drop_r)$.

0.5

Grow Ratio (Optional)

Relative measure in what the rise after a drop in the NDVI is required. In absolute terms it is calculated as $(\delta_{thr} * grow_r)$. Condition $END_VALUE > EXTREMA_VALUE + \delta_{thr} * grow_r$ must be satisfied for a valid event.

0.8

Delta Threshold (Optional)

Absolute measure that defines the magnitude of rise and fall in NDVI required for an event to be valid. Condition $abs(START_VALUE - EXTREMA_VALUE) \geq \delta_{thr}$ must be satisfied for a valid event.

0.11



Ray

- Dosti lastne logike!
- Lokalno paralelizirana eksekucija po HEX
- LSTM Predict
- Clustri



Summary

Nodes	
- total:	1268
- running:	89
- successful:	147
- pending:	1032
- failed:	0
Time	
- elapsed:	0:36:30

Progress

REF_reference_crop_land	69% 450/650	--::--	RUNNING
PF_is_small_foi	19% 123/642	--::--	RUNNING
PF_AP_all_obs_counter	0% 0/4	--::--	RUNNING
PF_AP_valid_obs_counter	0% 0/4	--::--	RUNNING
PF_baresoil_input_timeseries	0% 0/33	--::--	RUNNING
PF_distance_wo_percentiles_main	0% 0/642	--::--	RUNNING
PF_distance_wo_percentiles_winter	0% 0/596	--::--	RUNNING
PF_event_fall_raw	0% 0/288	--::--	RUNNING

Pending

[Node]	[State]	[Inputs Ready]
S2L2A_similarity_meanhood_winter	INIT	4 / 4
S2L2A_std_ndvi	INIT	2 / 2
S2L2A_std_ndvi_2024_04_01_2024_08_01	INIT	2 / 2
S2L2A_std_ndvi_2024_08_01_2024_10_01	INIT	2 / 2
S2L2A_std_ndvi_2024_10_01_2025_02_01	INIT	2 / 2

Done

[Node]	[State]	[Exec. time]	[Failed]
PF_mean_ndvi_marker_report_full	complete	0:01:12	0
S2L2A_event_gh_rise_raw_main	complete	0:16:18	0
S2L2A_historic_mowing_2021_mowingproba	complete	0:16:18	0
S2L2A_event_gh_rise_raw_non-main	complete	0:15:26	0
PF_mean_ndvi_marker_full	complete	0:20:55	0

Logs: /home/ubuntu/workspace/si-aktzp-workspace/data/SI24/Iteration-i4/output/official/logs/2025-04-10-081417.log | Ray Dashboard: <http://127.0.0.1:8265>

A vertical stack of green rectangular boxes, each representing a task in the Ray workflow. The tasks are labeled with names such as "Compute Publish JSONs for multiple H3 hexagons-73", "Compute Publish JSONs for multiple H3 hexagons-53", and "Compute Publish JSONs for multiple H3 hexagons-46". The boxes are arranged in a descending staircase pattern from top to bottom, indicating the flow of data or dependencies between tasks.

Grafana

- Fragmentiranost komponent
- Perfect UI
- ... Količina



logs (grafanacloud-planet-logs)

Kick start your query Label browser Explain query

```
{namespace='smt', container="container-flow.develop"}
```

Options Type: Range Line limit: 1000 Direction: Backward

This query will process approximately 609.6 KIB.

+ Add query i Query inspector

> Logs volume

Logs

Time Unique labels Wrap lines Prettify JSON Deduplication None Exact Numbers Signature

Display results Newest first Oldest first

Common labels: container=container-flow.develop instance=ip-10-0-1-85 job=integrations/docker namespace=smt service=area-monitoring service_name=area-monitoring stream=stdout

Line limit: 1000 reached, received logs cover 87.58% (10h 30min 34sec) of your selected time range (12h) Total bytes processed: 1.31 MB

Download ▾

```
> 2025-04-10 11:37:01.421 {"severity": "INFO", "event": "No tasks to process. Going to sleep.", "flow_name": "Internal: Signal Scheduler", "deploy_type": "STAGE", "flow_run_id": "f7669b69-a32f-408e-a2e7-27d1c483e976", "task_name": null, "task_run_id": null}
> 2025-04-10 11:37:01.420 {"severity": "INFO", "metric_name": "Number CREATED Tasks", "metric_value": 0, "metric_unit": "1", "flow_name": "Internal: Signal Scheduler", "deploy_type": "STAGE", "flow_run_id": "f7669b69-a32f-408e-a2e7-27d1c483e976", "task_name": null, "task_run_id": null}
> 2025-04-10 11:37:01.407 {"severity": "INFO", "event": "Success. Got 0 tasks with status CREATED.", "flow_name": "Internal: Signal Scheduler", "deploy_type": "STAGE", "flow_run_id": "f7669b69-a32f-408e-a2e7-27d1c483e976", "task_name": "Get SH tasks", "task_run_id": "923e57aa-5540-469e-9b54-abde8c4bb8b0"}
> 2025-04-10 11:37:01.402 {"severity": "INFO", "event": "Getting tasks with status CREATED.", "flow_name": "Internal: Signal Scheduler", "deploy_type": "STAGE", "flow_run_id": "f7669b69-a32f-408e-a2e7-27d1c483e976", "task_name": "Get SH tasks", "task_run_id": "923e57aa-5540-469e-9b54-abde8c4bb8b0"}
> 2025-04-10 11:37:01.358 {"severity": "INFO", "event": "Getting Tasks with status CREATED.", "flow_name": "Internal: Signal Scheduler", "deploy_type": "STAGE", "flow_run_id": "f7669b69-a32f-408e-a2e7-27d1c483e976", "task_name": null, "task_run_id": null}
> 2025-04-10 11:37:01.357 {"severity": "INFO", "metric_name": "Number in-progress Tasks", "metric_value": 0, "metric_unit": "1", "flow_name": "Internal: Signal Scheduler", "deploy_type": "STAGE", "flow_run_id": "f7669b69-a32f-408e-a2e7-27d1c483e976", "task_name": null, "task_run_id": null}
> 2025-04-10 11:37:01.342 {"severity": "INFO", "event": "Success. Got 0 tasks with status STOPPED.", "flow_name": "Internal: Signal Scheduler", "deploy_type": "STAGE", "flow_run_id": "f7669b69-a32f-408e-a2e7-27d1c483e976", "task_name": "Get SH tasks", "task_run_id": "3bb168c4-24bf-4ea2-b0a8-edb19438dab7"}
```



Izhodišče

Referenčni podatki

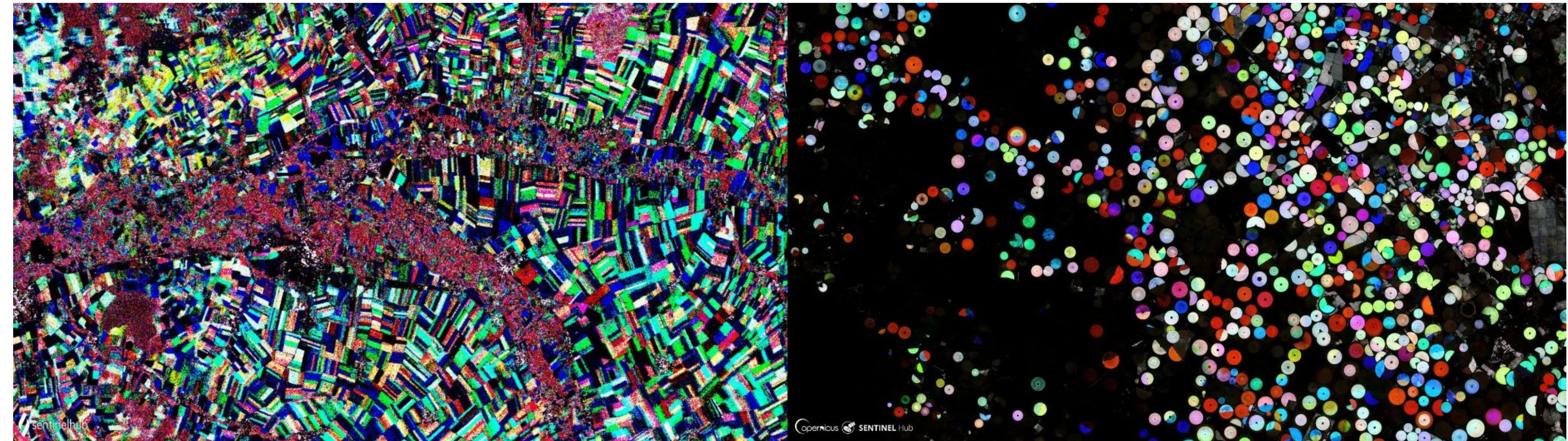
GEOPACKAGE

GPKG: Vsebina

Metrični koordinatni sistemi

Zahtevki





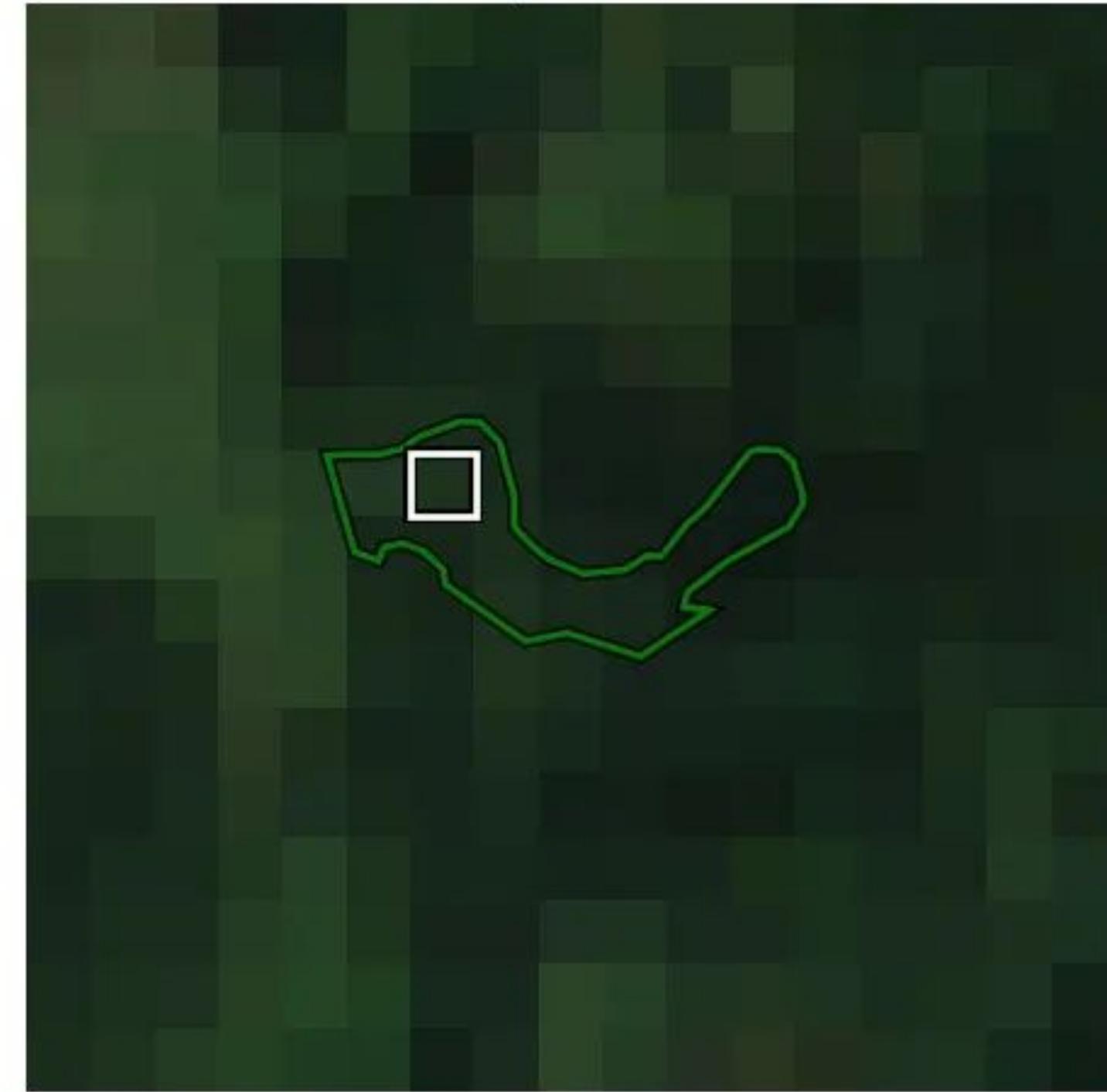
Struktura polj v Franciji (levo, vir: [Valtzen via Twitter](#)) in Savdski Arabiji (desno, vir: Harel Dan via [Twitter](#))

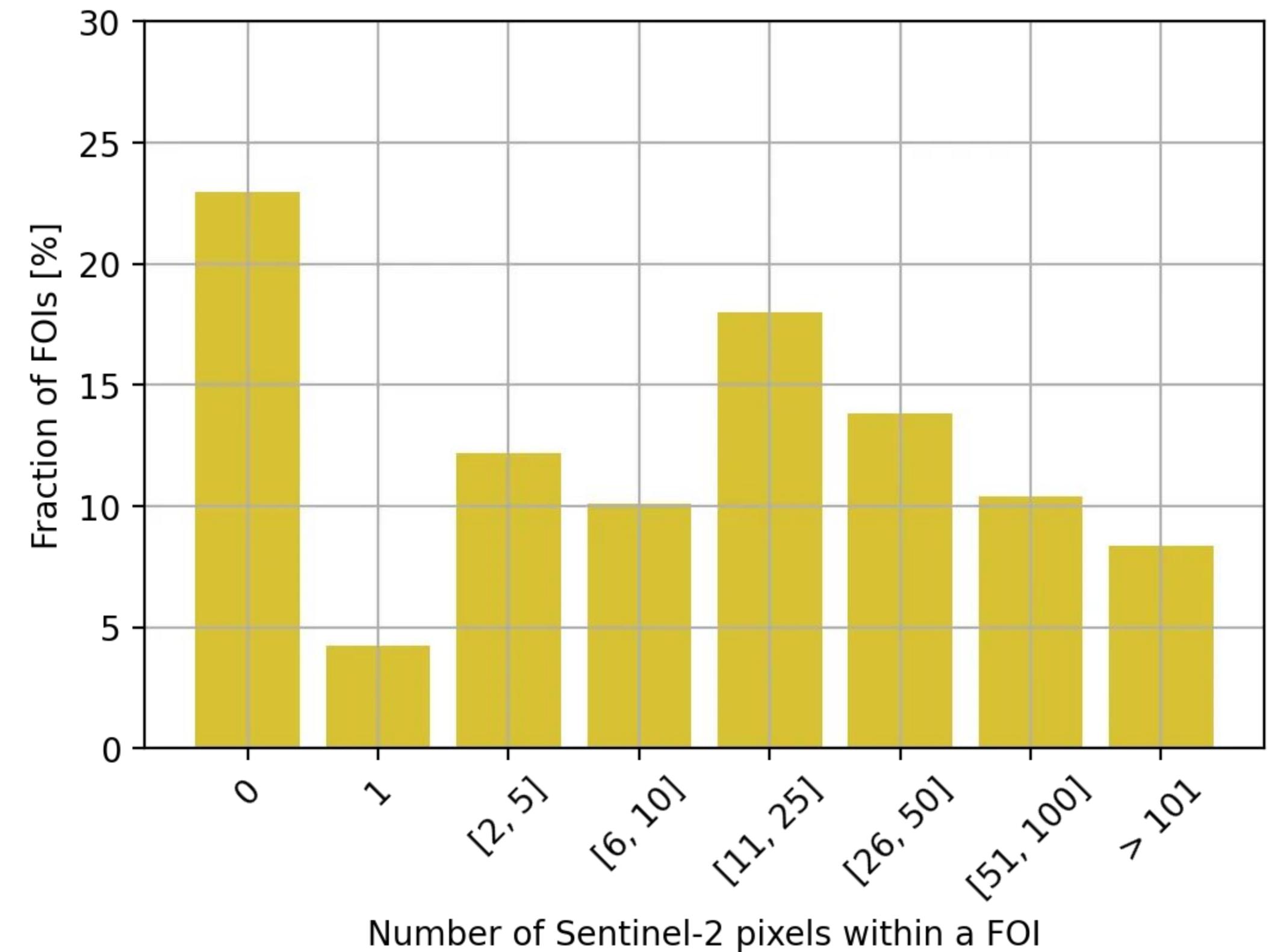


ID 5834833 | Date 2019-07-20



ID 4454355 | Date 2019-07-10







Prenos podatkov

Interakcija s SentinelHubom

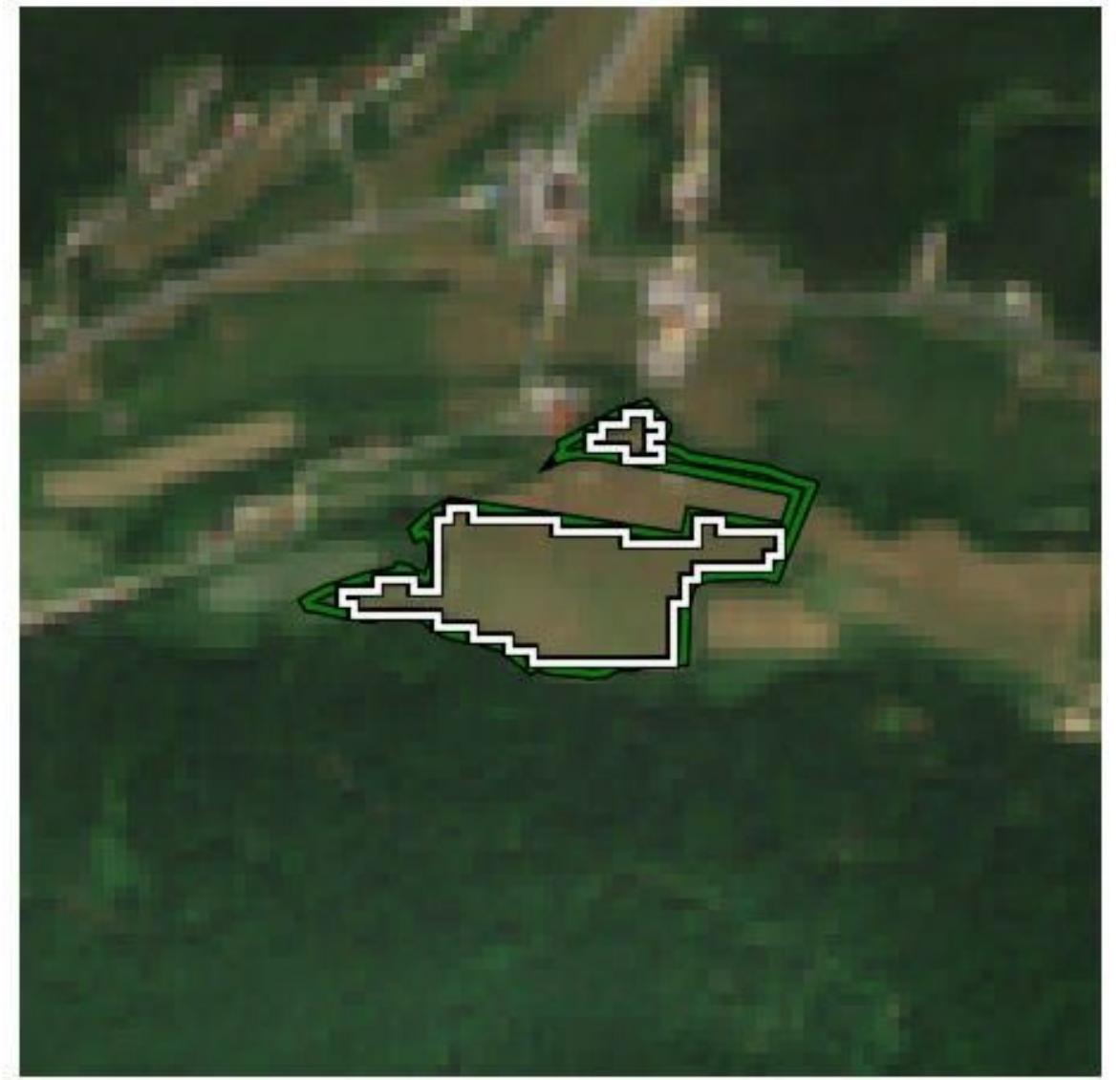
“Signali”

Agregirane vrednosti (srednja vrednost, std. odklon)*

Pikselacija

Ena časovna vrsta na FOI (Field of Interest - opazovano območje)

Poudarek na časovnih vrstah NDVI.



SentinelHub-Py <https://github.com/sentinel-hub/sentinelhub-py>

Batch Statistical API: <https://docs.sentinel-hub.com/api/latest/api/batch-statistical/>



Prenos podatkov

“Signal Blocks”

Spreminjanje reference med iteracijami.

Iterativni postopek:

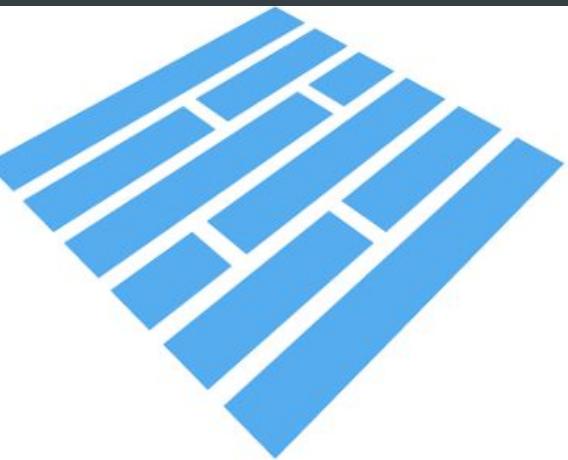
- Izhodišče: obstoječa podatkovna množica + ciljno stanje
- Izračun razlike po časovni komponenti in množico FOI.
- Prenos
- Združevanje

Parquet!

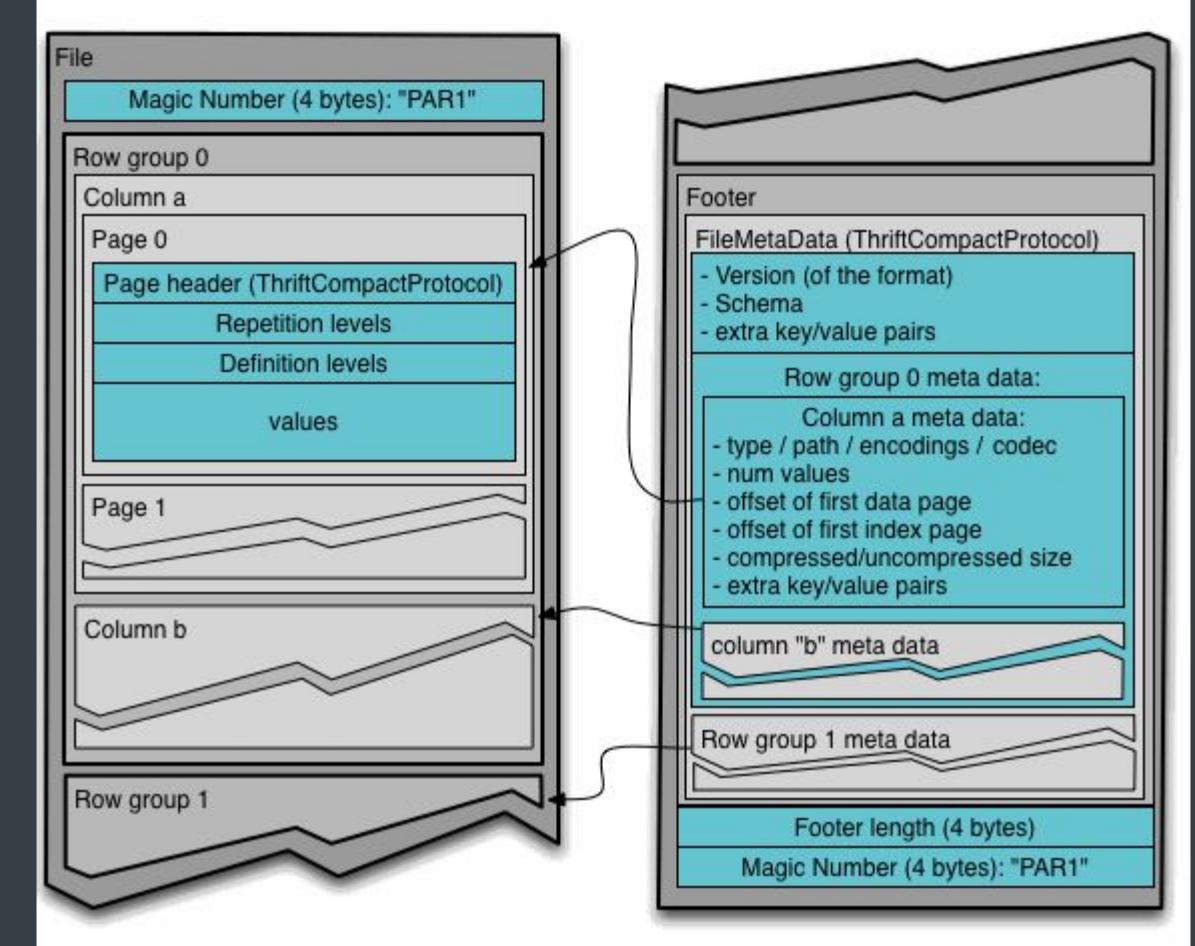


Parquet

- odprtakodni, “Column-oriented”, binarni format
- Učinkovito stiskanje*
- Učinkovito particioniranje: zasnovan za distribuirane sisteme*
- PyArrow
- Pandas (osnovno)

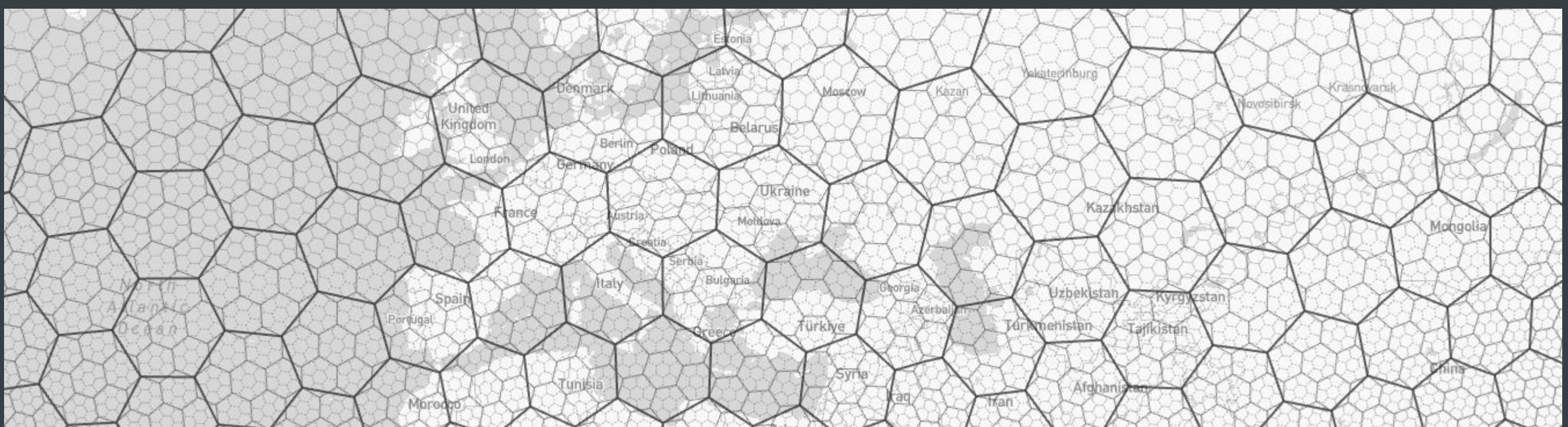
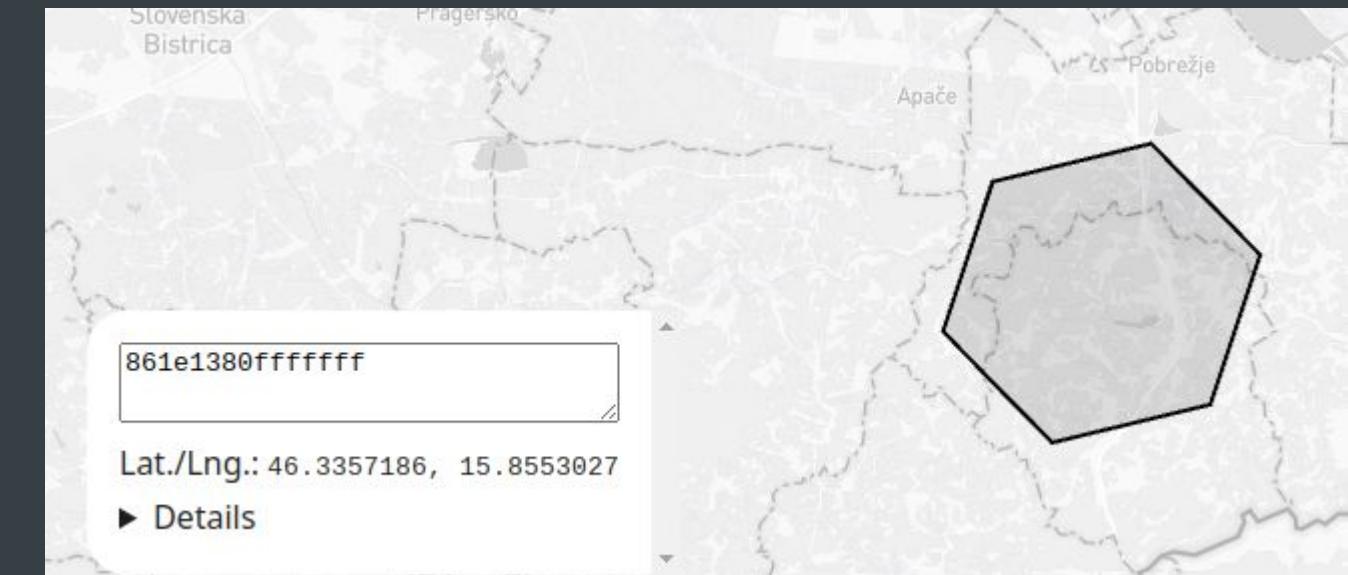


Parquet



H3

- Delitev površine Zemlje v šestkotne celice
 - Hierarhično, celica se deli na manjše
 - H3 indeksi: 64-bit številke
 - Algoritmi & optimizacija
 - Uber



Res	Total number of cells	Number of hexagons	Number of pentagons
0	122	110	12
1	842	830	12
2	5,882	5,870	12
3	41,162	41,150	12
4	288,122	288,110	12
5	2,016,842	2,016,830	12
6	14,117,882	14,117,870	12
7	98,825,162	98,825,150	12
8	691,776,122	691,776,110	12
9	4,842,432,842	4,842,432,830	12
10	33,897,029,882	33,897,029,870	12
11	237,279,209,162	237,279,209,150	12
12	1,660,954,464,122	1,660,954,464,110	12
13	11,626,681,248,842	11,626,681,248,830	12
14	81,386,768,741,882	81,386,768,741,870	12
15	569,707,381,193,162	569,707,381,193,150	12

Particioniranje Parquetov s H3

- Paralelizacija
- Omogoča operacije na sosedčini objektov
- PyArrow

<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e100dfffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e10217fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e1021ffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e1024ffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e10257fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e1025ffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e10287fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e1028ffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e10297fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e1029ffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102a7fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102affffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102c7fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102cffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102d7fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102dffffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102e7fffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102effffff/	Folder
<input type="checkbox"/>	<input type="checkbox"/> H3HEX=861e102f7fffff/	Folder



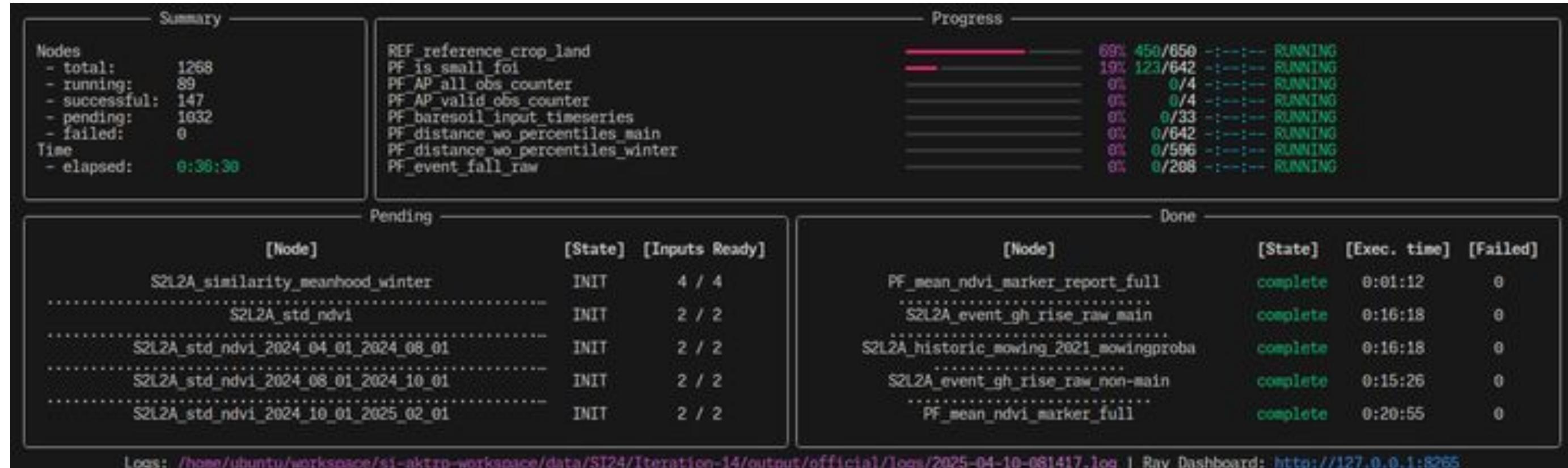
Markers

Logika

DAG: usmerjeni aciklični graf

Logika za custom node IO in njihovo povezovanje

Izvajanje “po meri”: Ray, CLI pregledna plošča



Pandas

- Python knjižnica
- Učinkovito manipuliranje velikih DataFramov





Markers

Node: definicija

“Glorified funkcije”

Logika je načeloma konstantna in se deli med različnimi projekti.

Domena DS in BE.

S stisnjениmi zobmi omogočamo tudi posebne “node” za posamezne projekte.

```
class Similarity(PublishableNode):
    """`Similarity` is a marker/Node whose goal is to map the percentiles (from [Percentiles](),
    from chi squared for various label hypothesis (from [SimilarityWoPercentiles]().`SimilarityWoPercentiles`).
    determine the classification label of the FOI and its percentile `classification_score`.

    An example of the output:

    +-----+-----+-----+-----+-----+
    | POLY_ID | MARKER_ID | score | classification | classification_score | mapped_a
    +-----+-----+-----+-----+-----+
    | 7244139001 | 7244139001 | 100 | 405-1100 | 35 | 005-1100
    | 7244139001 | 7244139001 | 100 | 405-1100 | 35 | 005-1100
    | 7244139001 | 7244139001 | 51 | 405-1100 | 35 | 005-1100
    | 7244139001 | 7244139001 | 100 | 405-1100 | 35 | 005-1100
    | 7244139001 | 7244139001 | 100 | 405-1100 | 35 | 005-1100

    `Similarity`'s output, more specifically the percentile scores of each hypothesis, can also
    be visualized as a histogram:
    ![[similarity]](class_doc_figs/similarity.png)
    """

Ziga Luksic, 4 weeks ago | 3 authors (Ziga Luksic and others)
class Inputs(PublishableNode.Inputs):
    foi_scope: FoiScope
    similarity_wo_percentiles = SimilarityWoPercentiles.Outputs.similarity_wo_percentiles
    percentiles: ParquetFile
    label_map = LabelMapNode.Outputs.label_map.subset(
        FOI_ID=FOI_ID,
        LABEL=ParquetColumn.define(name=None, dtype="string", default="NULL_LABEL"),
    )

Ziga Luksic, 4 weeks ago | 2 authors (Ziga Luksic and one other)
class Outputs(PublishableNode.Outputs):
    similarity = ParquetDataset.define(columns_class=PercentileSimilarityOutputCols)

Ziga Luksic, 3 weeks ago | 1 author (Ziga Luksic)
class PublishInfo(IntervalPublishInfo):
    hexes_kring: int

    inputs: Inputs
    outputs: Outputs

    def compute(self, hex_ids: HexIds, foi_ids: FoiIds):
        data = self.inputs.similarity_wo_percentiles.read(hex_ids=hex_ids, foi_ids=foi_ids)
        ...

```



Markers

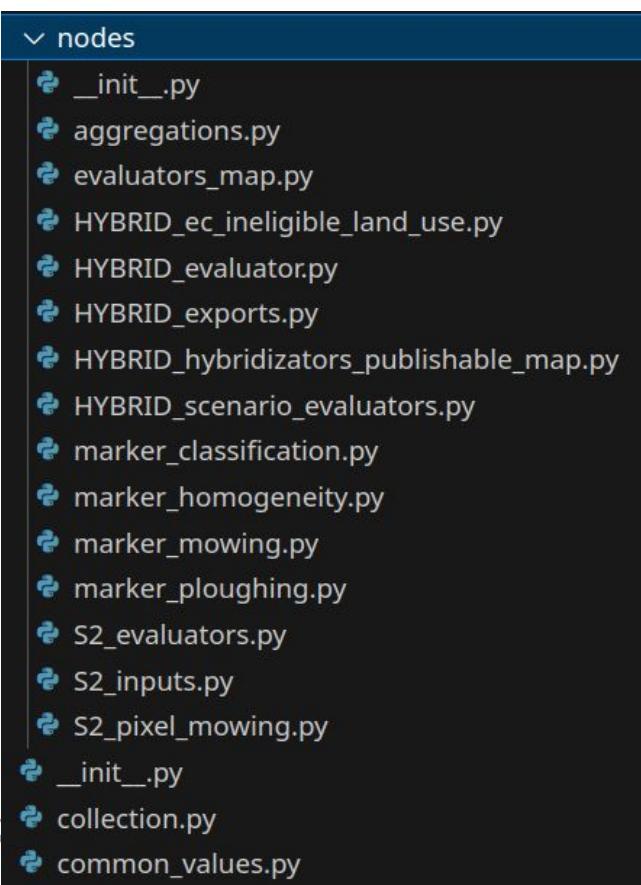
Node: uporaba

Instanciranje "nodov"

"Končni uporabnik" definira vhode, izhode in nastavitev

Nodi se medsebojno referencirajo!

Domena PM.



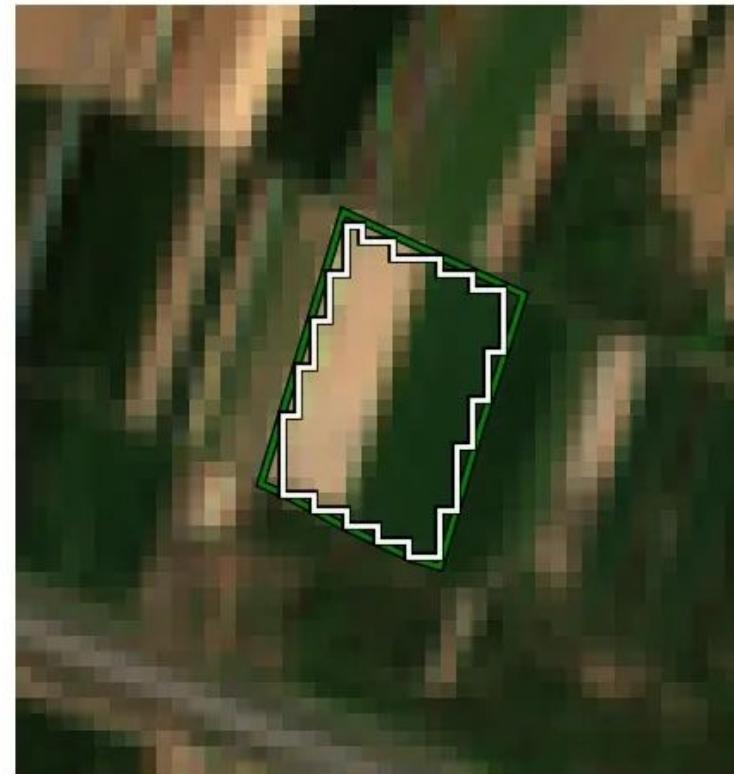
```
pixel_mowing_proba_node = PixelEventProba(
    name="S2L2A_pixel_mowingproba",
    config=PixelEventProba.Config(rgb_scaling_factor=2.7, temporal_features=["NDVI", "L", "a", "b"]),
    inputs=PixelEventProba.Inputs(
        {
            "foi_scope": foi_scope_node.outputs.foi_scope, # type: ignore[dict-item]
            "reference": reference_node.outputs.reference.model,
            "signals": pixel_signals_node.outputs.signals.model.copy(
                update={
                    "valid_query": f"{common.S2L2A_mowing_valid_query} and STATUS == 0",
                    "columns": {
                        "BLUE_BAND": _float_32_col("B02"),
                        "GREEN_BAND": _float_32_col("B03"),
                        "RED_BAND": _float_32_col("B04"),
                        "B01": _float_32_col("B01"),
                        "B05": _float_32_col("B05"),
                        "B06": _float_32_col("B06"),
                        "B07": _float_32_col("B07"),
                        "B08": _float_32_col("B08"),
                        "B8A": _float_32_col("B8A"),
                        "B09": _float_32_col("B09"),
                        "B11": _float_32_col("B11"),
                        "B12": _float_32_col("B12"),
                        "NDVI": _float_32_col("NDVI"),
                        "CLP": _float_32_col("CLP"),
                        "OUT_PROBA": _float_32_col("OUT_PROBA"),
                    },
                }
            ),
            "eventproba_model": event_proba_model_node.outputs.model, # type: ignore[dict-item]
        }
    ),
    outputs=PixelEventProba.Outputs(
        {
            "eventproba": DataItemModel(
                path=f"{local_path}/markers/pixel-mowingproba",
                remote=f"{remote_path}/markers/pixel-mowingproba/",
            )
        },
        origin_tag=ORIGIN_TAG,
        exec_config=ExecConfig(split_by="am.core.splitter:BalancingFoiNodeSplitter", chunk_min_size=500),
        variables=common.variables,
        storage=common.storage,
    )
)
collection.add_node(
    pixel_mowing_proba_node, [foi_scope_node, reference_node, pixel_signals_node, event_proba_model_node]
)
```



Markers

Homogenost

ID 2302804 | Date 2019-04-21



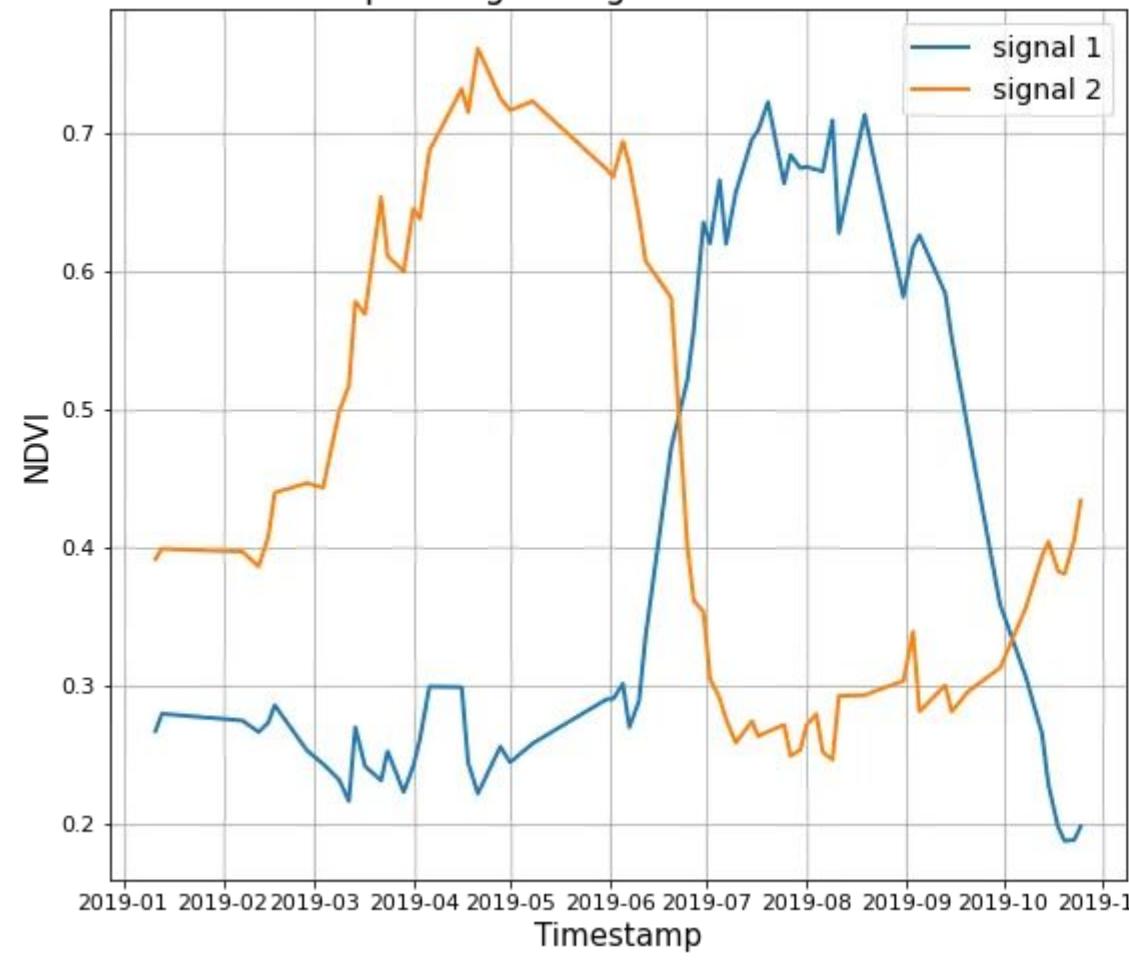
ID 2302804 | Date 2019-07-20



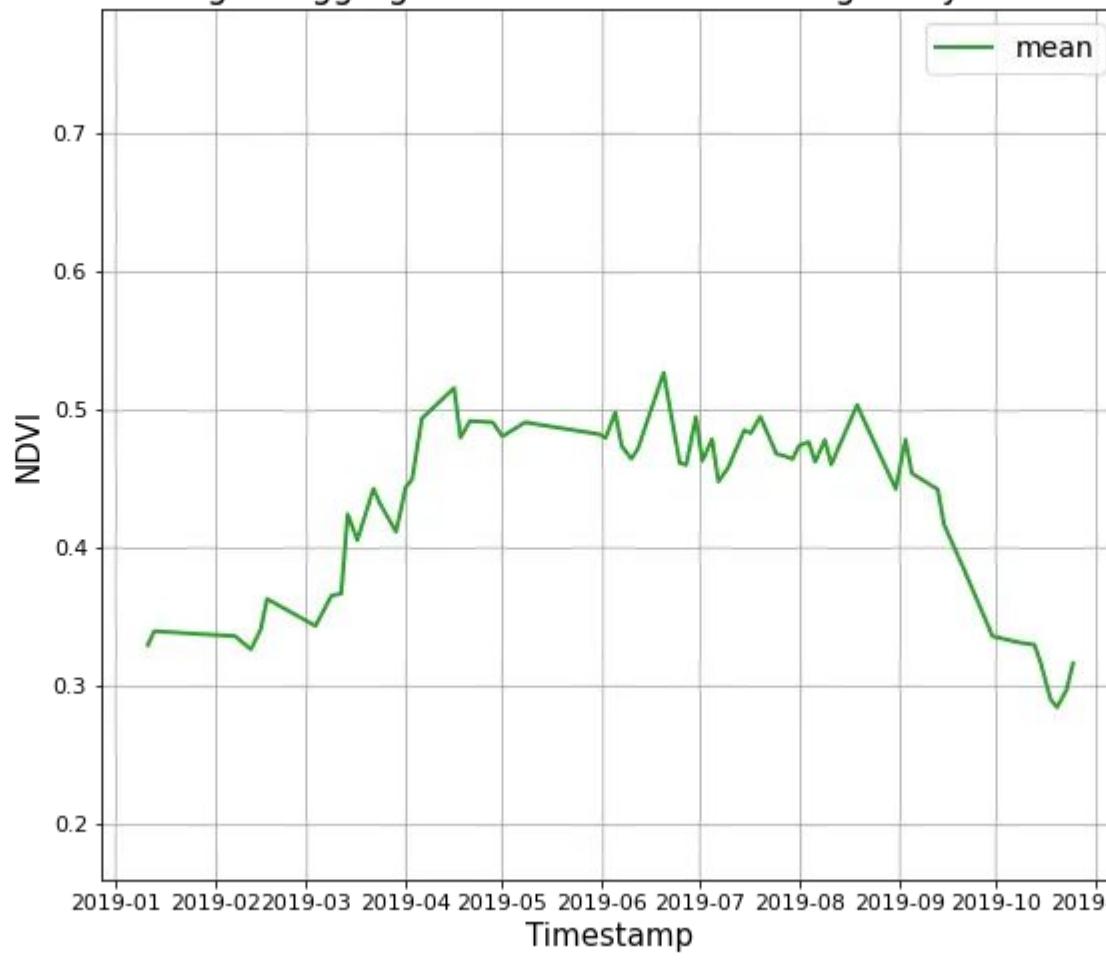
ID 2302804 | Date 2019-10-18



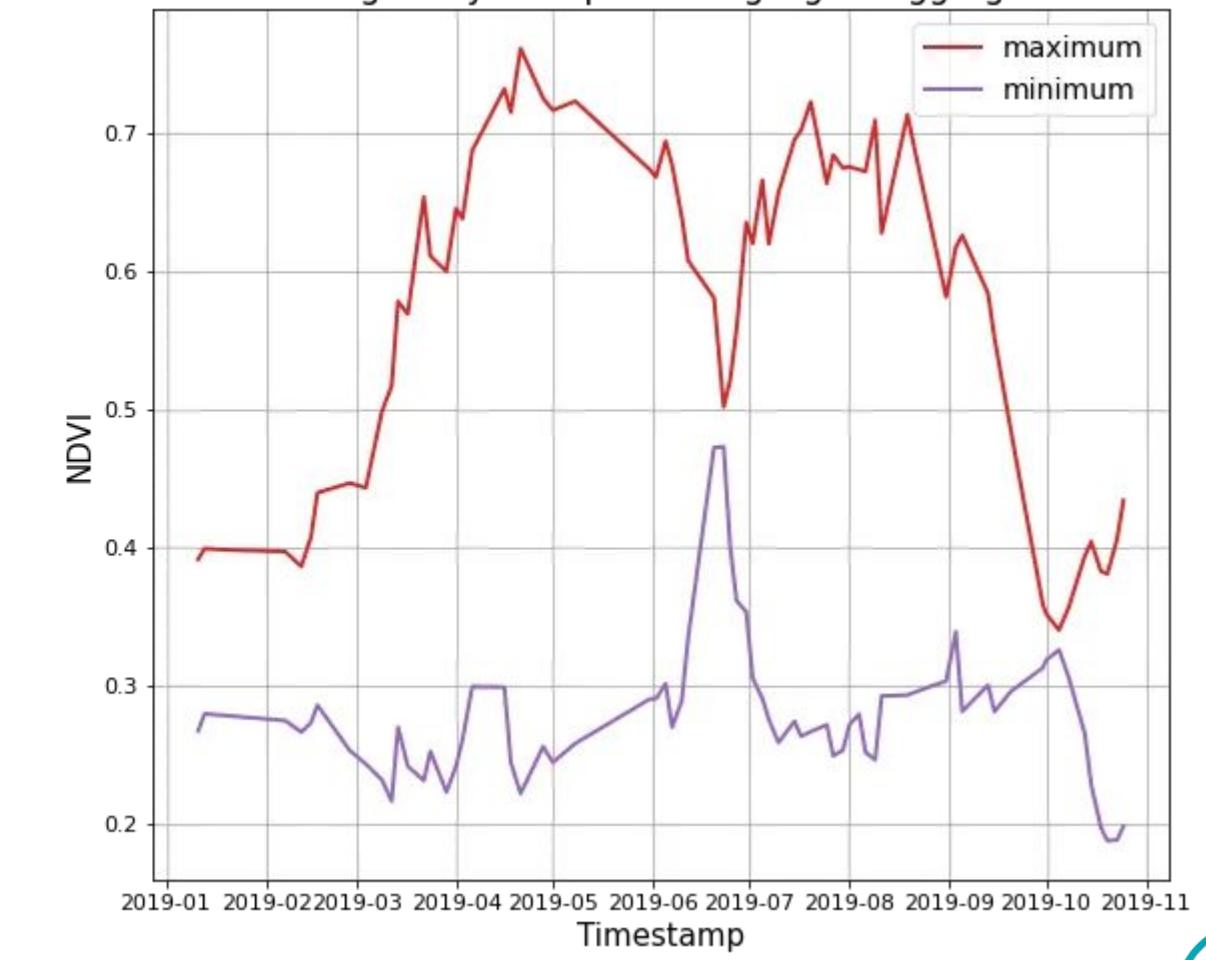
Mockup of original signals within some FOI



Signal aggregation with loss of inhomogeneity info.



Inhomogeneity info.-preserving signal aggregation





Markers

Podobnost

Primerjava zemljišč z ostalimi iste
kulture v okolici (HEX)

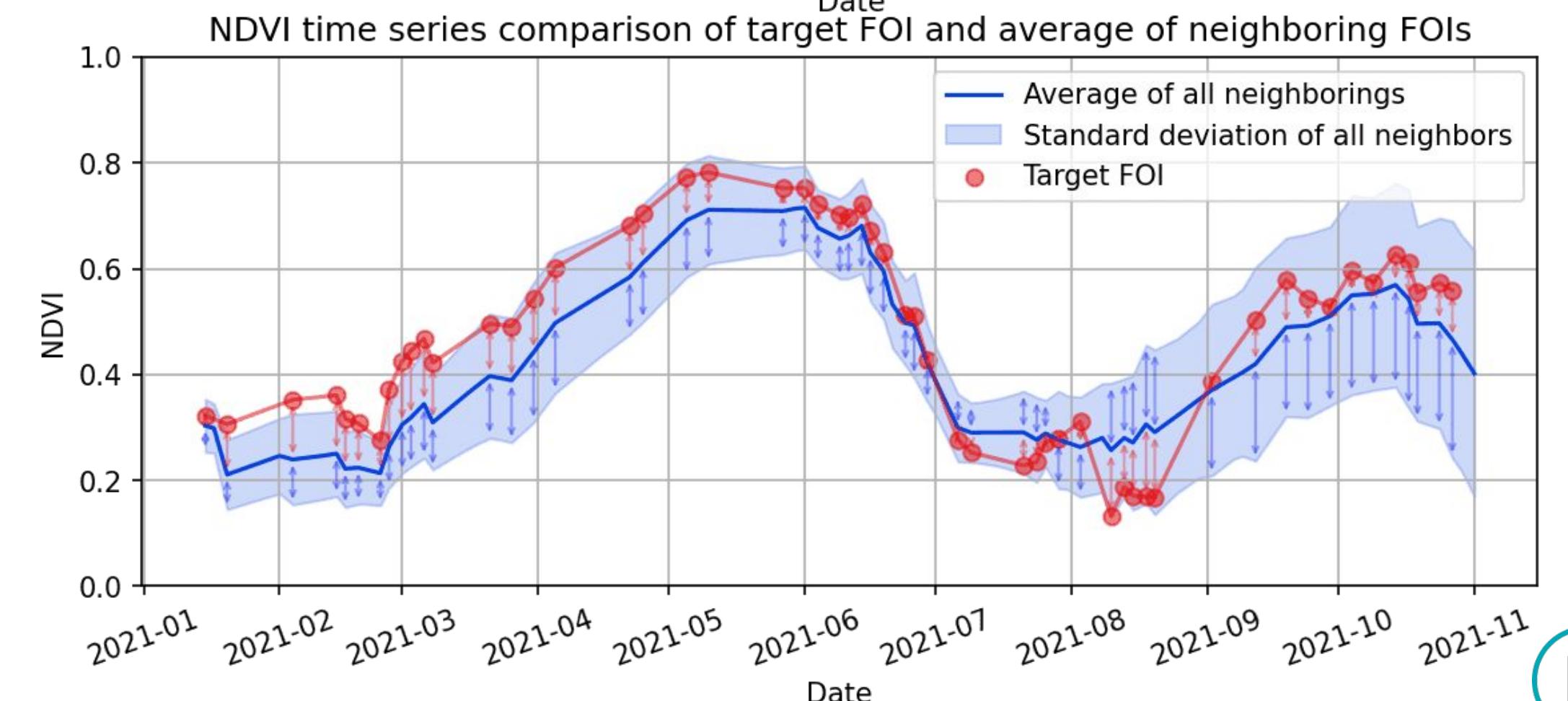
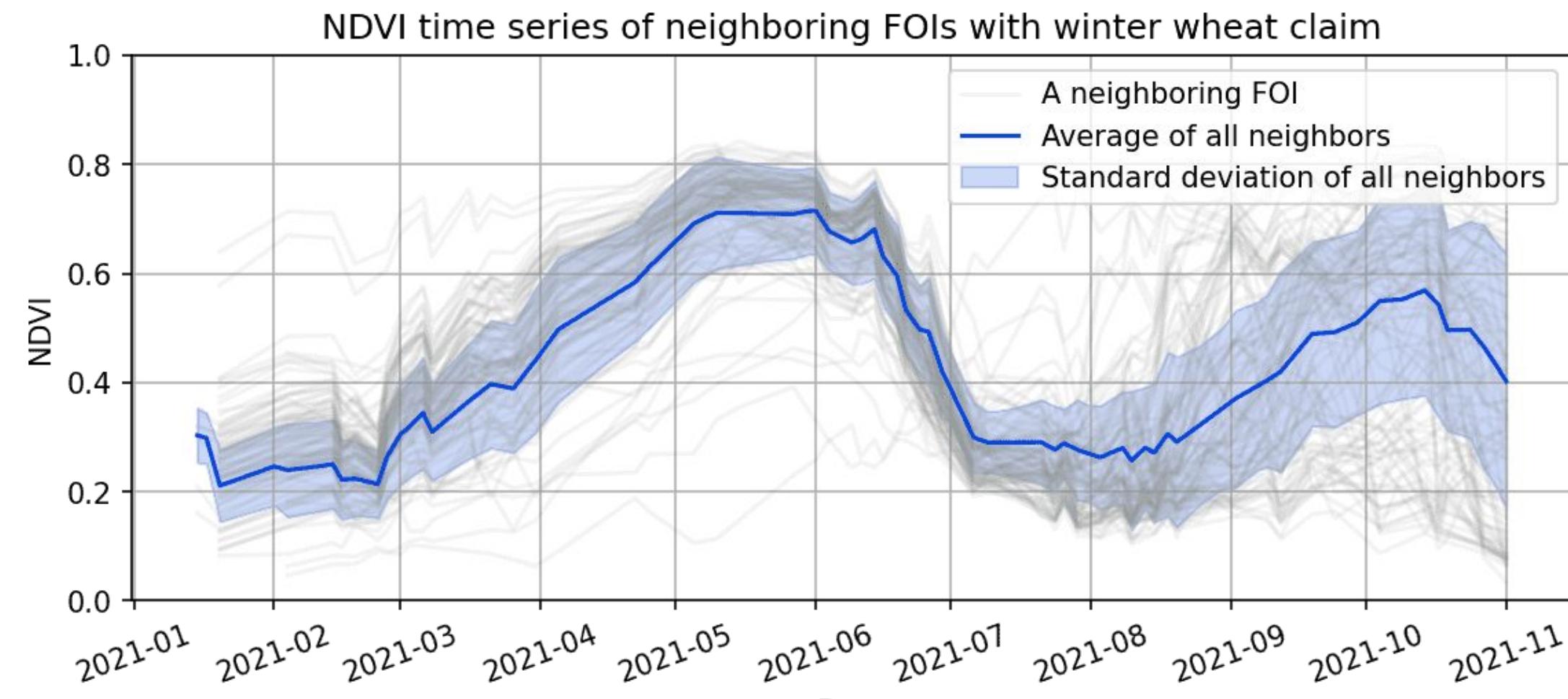




Markers

Podobnost

Primerjava zemljišč z ostalimi iste kulture v okolici (HEX)





Markers

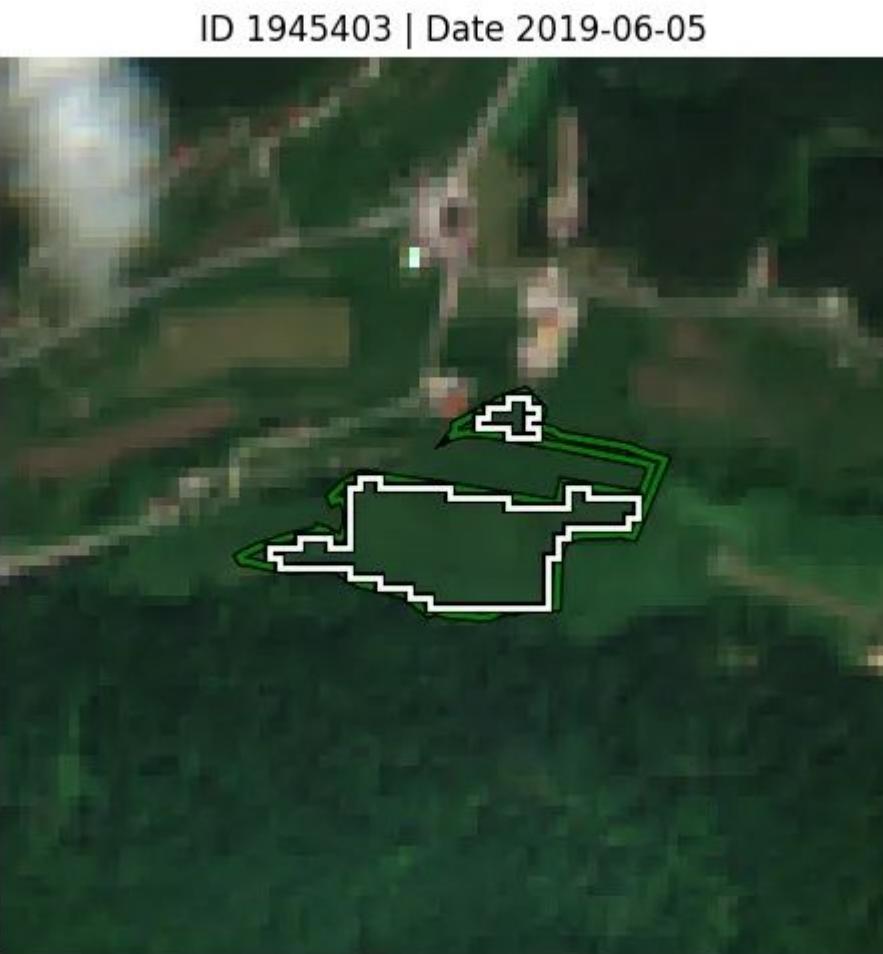
Gola zemlja





Markers

Košnja

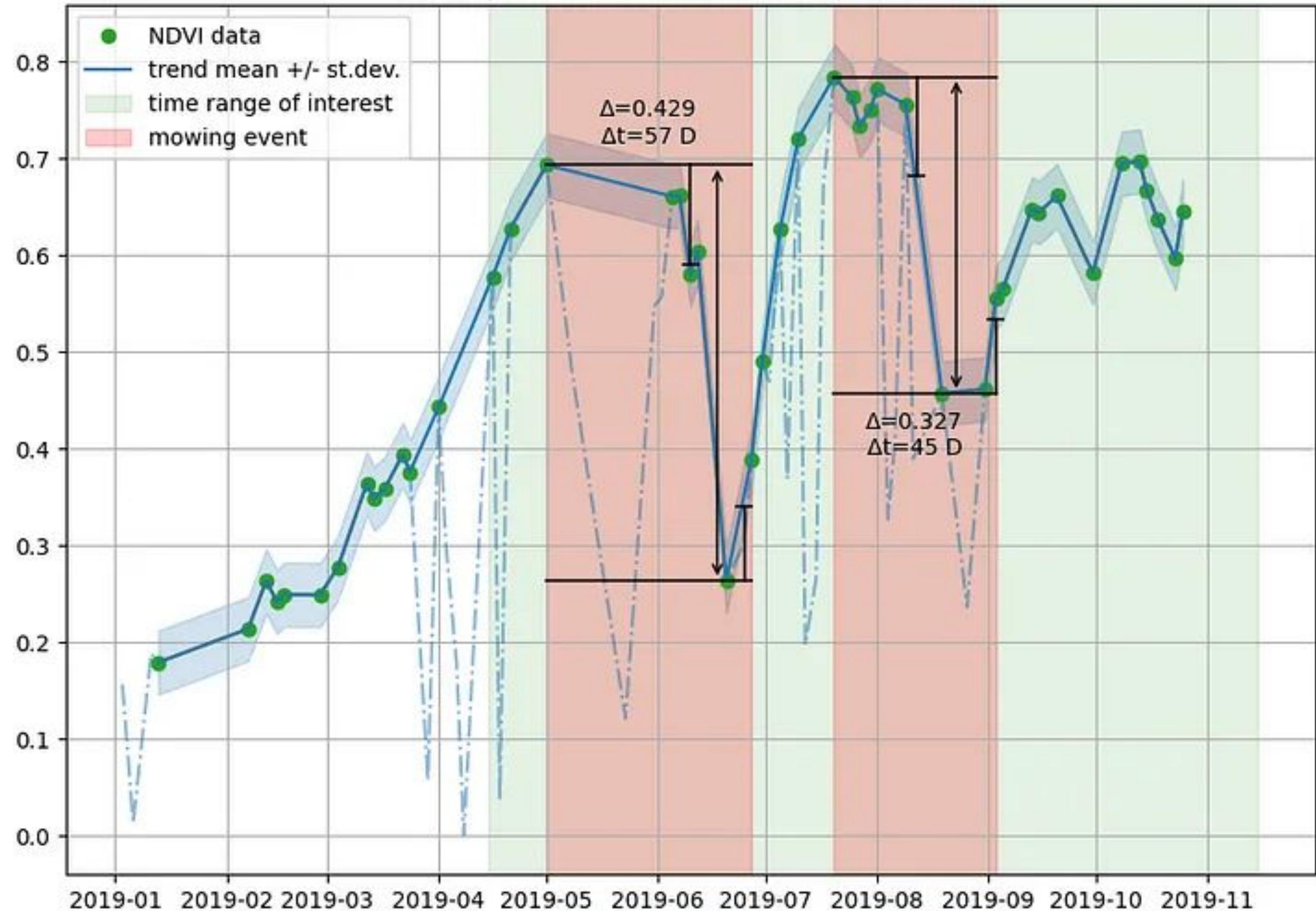




Mowing Marker V3 - ID 6046949 | Label 201

Markers Košnja

Identičen algoritem za
detekcijo oranja,
ozelenitev in žetev...





Markers

Klasifikacija

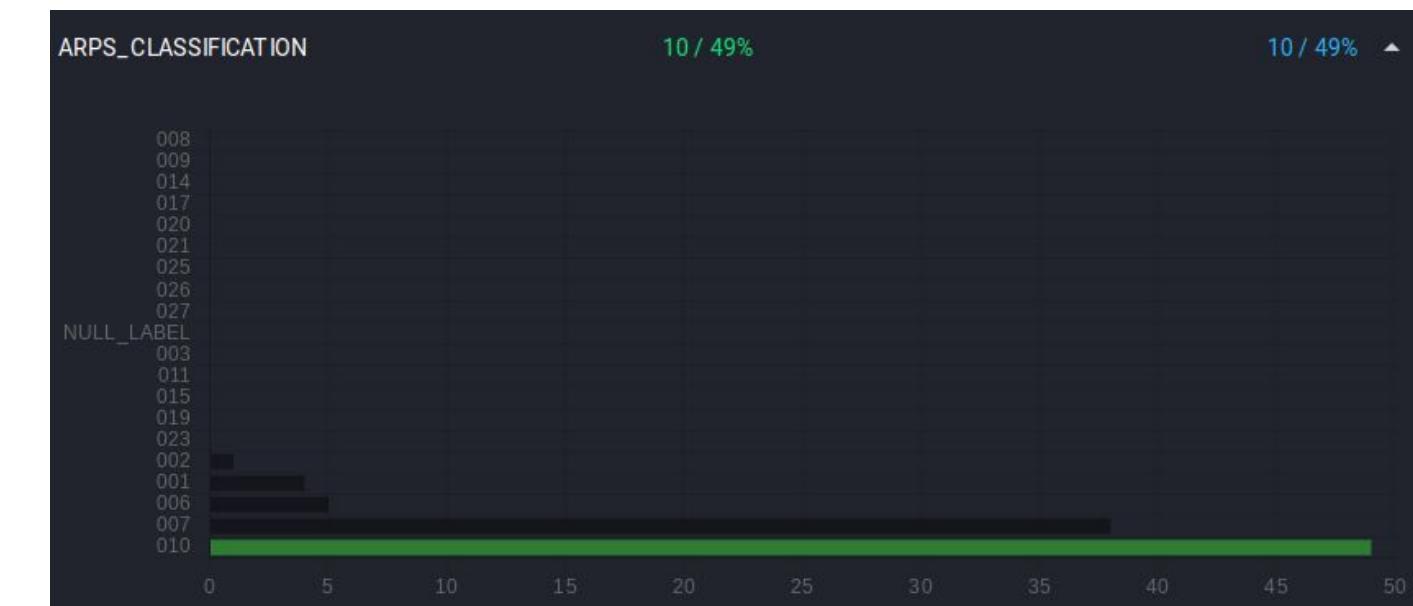
Problem: Kateri kategoriji pripadajo posevki na opazovani površini?

V našem primeru imamo vnaprej deklarirane razrede.

Učenje: Časovne vrste različnih kanalov in indeksov

Napovedovanje: Ocenjevanje verjetnosti, da površina pripada določenemu razredu

Morda bolj “outlier detection”.



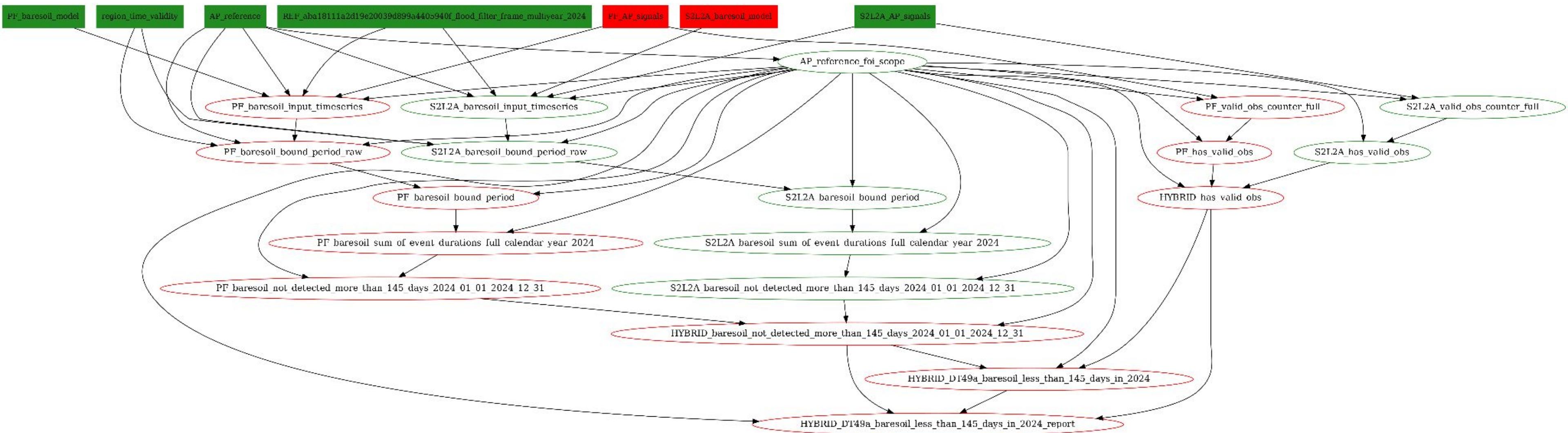
Long Short-Term Memory (LSTM)



Markers Collection

Zbirka nodov, ki gradijo usmerjen graf.

Rast kompleksnosti, od 10 do 1500 nodov na projekt.

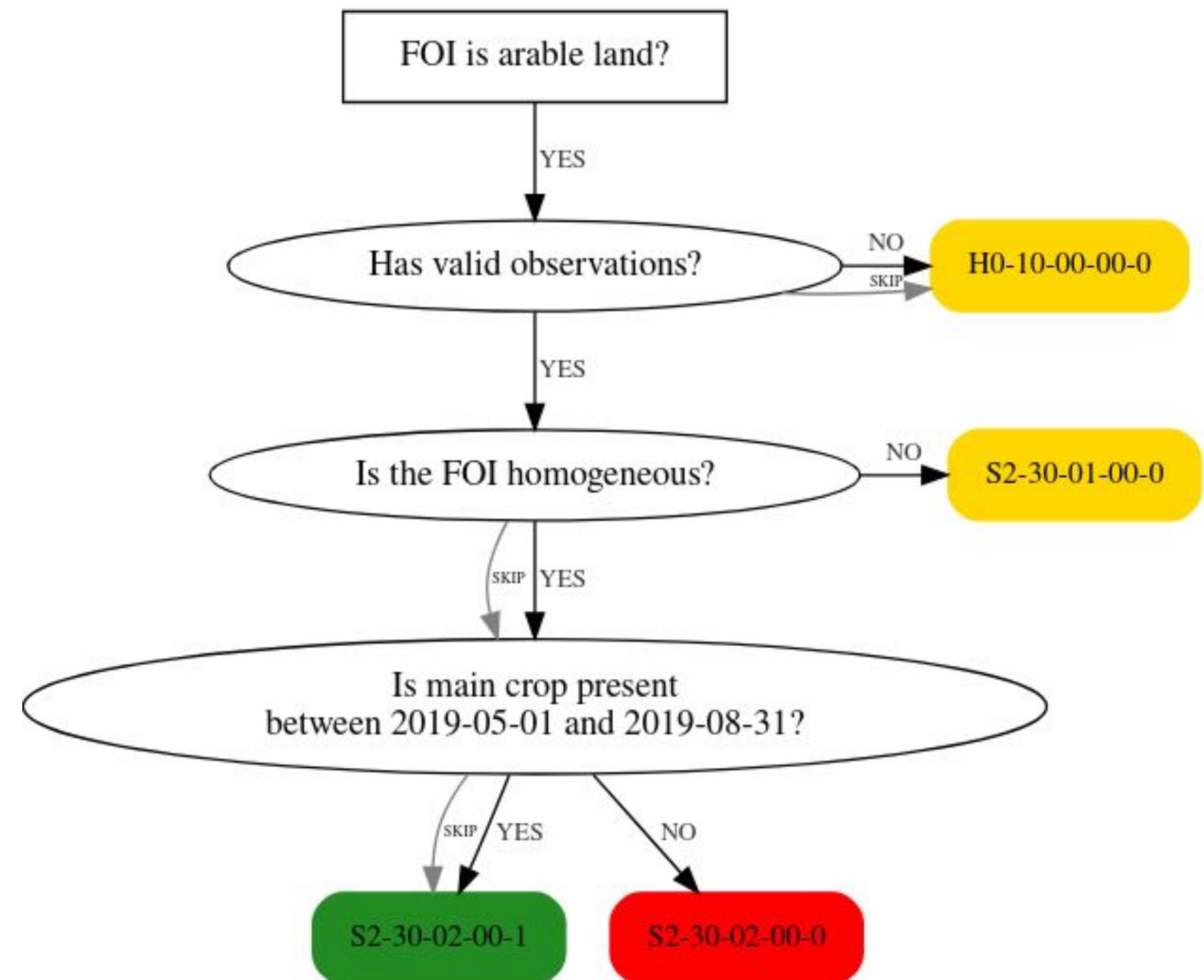




Markers

Odločitvena drevesa

Vsako zemljišče dobi dodeljeno odločitev glede na režim subvencioniranja.





Standardni markerji

Generalizacija

Standardne skupine nodov z določeno funkcionalnostjo.

Manj generalnosti, kot pričakovano.

Marker Configurations (Optional)
Composition of all the marker configurations utilized in the run.

Type: Mowing (Optional)
Configurations for the Mowing marker.

Source: PF (Optional)
No items in this list
+ Add item

Source: ARPS (Optional)
No items in this list
+ Add item

Source: S2L2A (Optional)

⋮
Drop Ratio (Optional)
Relative measure in what the drop in the NDVI is required. In absolute terms it is calculated as `(delta_thr * drop_r)`.
0.5

Grow Ratio (Optional)
Relative measure in what the rise after a drop in the NDVI is required. In absolute terms it is calculated as `(delta_thr * grow_r)`. Condition `END_VALUE > EXTREMA_VALUE + delta_thr * grow_r` must be satisfied for a valid event.
0.8

Delta Threshold (Optional)
Absolute measure that defines the magnitude of rise and fall in NDVI required for an event to be valid. Condition `abs(START_VALUE - EXTREMA_VALUE) >= delta_thr` must be satisfied for a valid event.
0.11

Slope Threshold (Optional)
Slope of NDVI where we stop extending the event. Must be strictly less than zero
-0.007

Valid Query (Optional)
Definition of the filter used to determine valid observations.
CLP <= 0.4 and OUT_PROBA <= 0.8

Summary Crop IDs (Optional)
An optional list of crop codes for subsetting a marker calculation report. If omitted, all the crops are regarded as relevant.
No items in this list
+ Add item

Identifier Suffix (Optional)
A suffix added to a marker to distinguish it from other markers that have the same type and source. It defaults to an empty string if not provided.

Subscription Name
Subscription code associated with the signal package user wants to run the marker on.

FOI Selector Query (Optional)
A reference dataset-based pandas query to define which FOIs to include in the computation of the corresponding standard marker.

Region Time Validity Filter (Optional)
Path to a Parquet dataset on SMT-accessible S3 bucket containing set of large polygons with associated temporal filters.

Postprocessing Queries (Optional)
An optional list of queries. It lets user control false-positive detections by querying them as in Pandas query.
No items in this list
+ Add item

Publish Mowing Marker (Optional)
Whether to publish the Mowing marker

Valley to Peak Upper Bound (Optional)
Upper bound for the jump/drop between the first valley and the second peak of a consequent rise and fall pair can have in order to be merged.
0.088

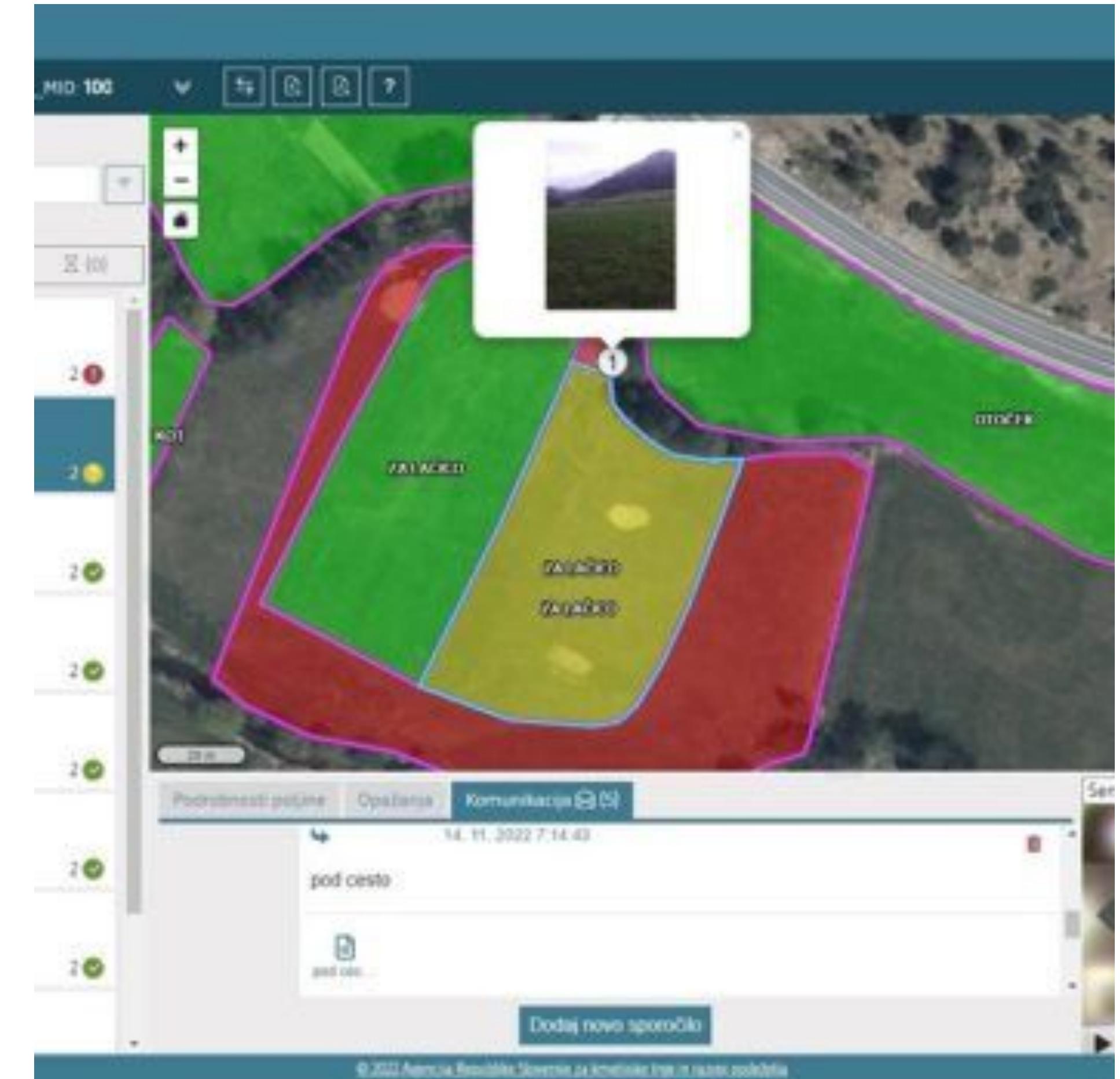


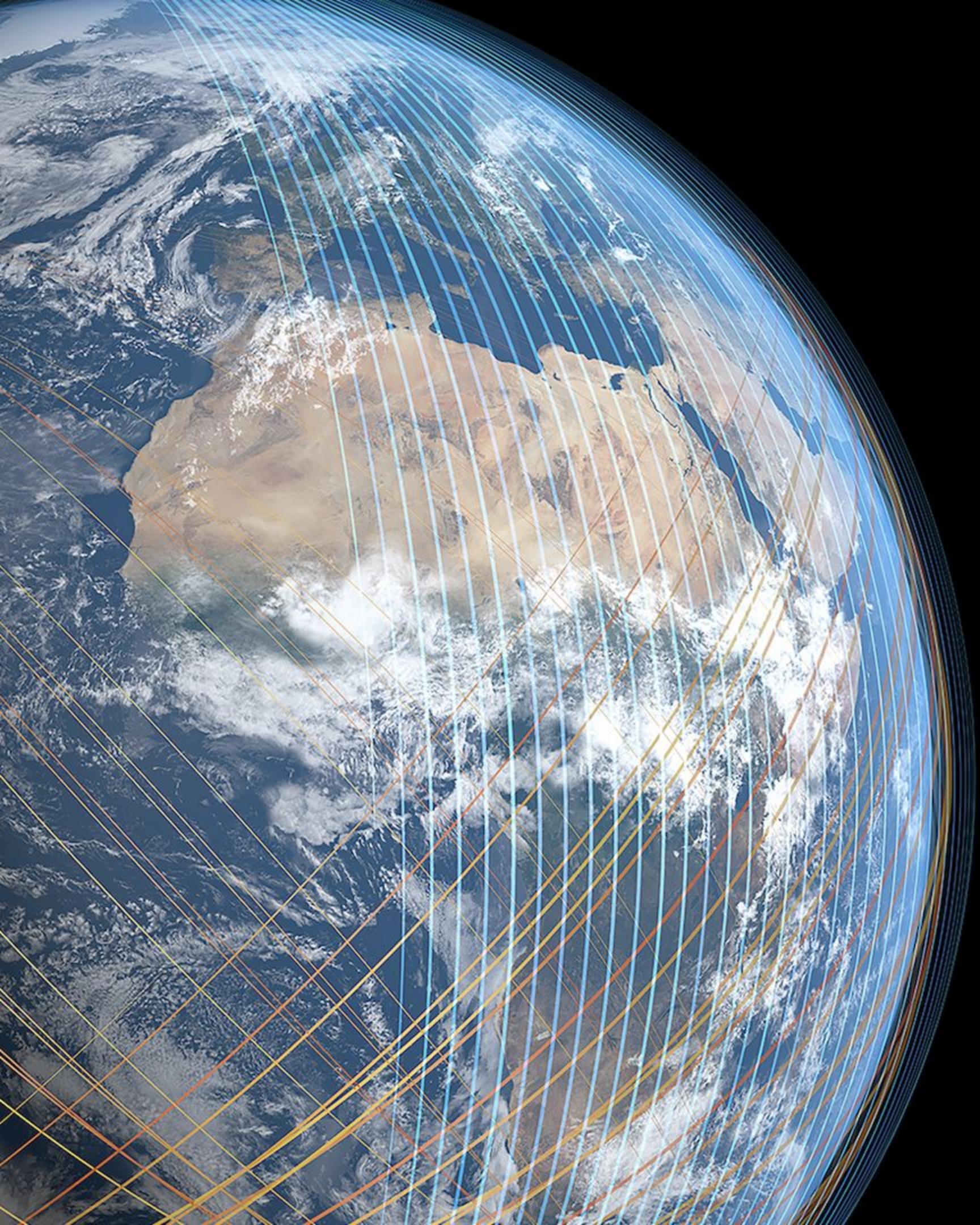
Rezultati

Objava v različnih servisih.



APLIKACIJA **SOPOTNIK**





Hvala za pozornost.



Klemen Lovenjak
Software Engineer
klemen.lovenjak@planet.com





EPP

"Wild Wild Space" je znanstvenofantastični film, ki raziskuje možnosti prihodnosti človeštva v vesolju. Glavna zgodba se osredotoča na raziskovanje oddaljenih planetov, kjer so naloge raziskovanja in odkrivanja novih svetov prepuščene naprednim vesoljskim agencijam in podjetjem, kot je Planet Labs.

Planet Labs, znan po svoji floti majhnih satelitov za spremljanje Zemlje, v filmu igra ključno vlogo pri zbiranju podatkov in obveščanju glavnih likov o stanju na drugih planetih. Ta tehnologija omogoča natančen vpogled v življenje na drugih svetovih, hkrati pa ponuja izzive, ki jih mora človeštvo premagati, da bi preživel v zvezništvu z naravnimi zakoni vesolja.

Film se dotika tudi vprašanja, kako tehnologija, kot je ta, ki jo ponuja Planet Labs, vpliva na naš življenje na vesolje in naše sposobnosti za preživetje v njem. Izpostavljeni so tudi etični in okoljski pomen tehnologije, ki pomembno vpliva na raziskovanje novih planetov in kako ravnamo z naravnimi viri na drugih svetovih.

Ozadje komercializacije
izstreljevanja raket v vesolje.

<https://www.imdb.com/title/tt3225885/>

