MadMexDocumentation

Table of contents

ToC

Home

This is the user documentation for the current MAD-Mex Version 2.2.

MAD-Mex 2.2 is the successor of the CONABIO MAD-Mex 2.1 system. Version 2.2 features include: GUI, project configuration driven processing, elimination of SGE and Postgres dependencies. It is a consolidation of the MADMex system to support national land cover/land cover change products. Changes focus on simplification of system installation and maintenance, user interface and integrated product operations. Goal is the consolidation of the production process and its administration.

MADMex system is designed to process large amount of data in a cluster environment. It takes advantages of cloud provider services like s3 and ec2 from Amazon Web Services. There are no hard dependencies to these cloud services - a local on premise installation is possible via configuration.

All computing workers need access to centralized data structures. Required base data will be copied to each node in a temporary processing workspace. One computing node can execute multiple jobs in parallel if minimum resources (RAM, CPU, disk space) area available. Results are uploaded to centralized project storage.

MADMex components

- Processing (Landsat USGS product T1, Sentinel2 L1C)
  - Numerical data processing in Python
  - Land cover product adapter
  - Land cover change product adapter
  - Containerization with docker
- Job distribution
  - Tasking
  - Job load scaling
  - Job queues, prioritization and job management
- Logging
  - Production
  - Resource monitoring
- GUI
  - Project processing
  - Product catalog

Processing adapters

Adapters implement the workflows to process the different MADMex products and sub-products. They can be executed directly via CLI applications or send into a processing queue. Every. adapter needs its input parameter set to specify the corresponding workflow and its dependencies. There exist two CLI programs which execute adapter workflow. Additional shell scripts used by the Sun Grid Engine aren't necessary for the current version. More details about job execution can be seen in the corresponding notebooks.

Adapters call necessary functions located in MADMex base module structures. Adapters are located directly in the MADMEX repository and organised in packages: atomic, combination, external, project, workflow. Atomic functions contain only singular functions, like image masking. These functions are used in all higher level adapters. Combined adapters like change detection are also the base for the workflow adapters which implements the supported products for the different sensors. External adapters serve as wrappers for external functionality installed in included docker images.

In general the system is installed using various docker containers. MADMex is designed to execute additional software in these already created docker images. The system itself runs within a docker instance to ensure full compatibility on different platforms.

Job distribution

Celery is used to distribute & collect jobs in the cluster. It replaces former Sun Grid engine to simplify MADMex cluster installation and management. It consists of the following components:

- Celery (python module) https://github.com/celery/celery
- RabbitMQ (message broker) https://www.rabbitmq.com
- RedisDB (key/value database) https://redis.io

Celery serializes the execution of every decorated and configured python function and send the information to the message broker. The message broker handles job priorities, different execution queues and job persistency. A celery worker instance registers itself and is allowed to pull jobs from a given queue. After execution of the job at the worker instance the results and additional meta data is stored in a key/value database. The worker specifies the number of concurrent run instances and defines the using queues.

Because Celery is implemented in Python it is possible to send/monitor jobs directly in python via cli/web interface or as code / notebook. No additional scripts are necessary to wrap execution on any number of remote workers. Jobs can be generated automatically to process larges batches of data. These jobs are collected in the message broker component persistently. A dynamic number of worker can be executed to process these jobs. Queues redirect tasks to specific workers to allow job priorities.

The GUI takes advantage of the job execution/monitoring. The creation and handling of a number of jobs is more easily achieved with the current version of MADMex.

Logging

All MADMex components support a constant logging of its state and activity. Logger will be instantiated in a module, e.g. logger = AdapterLogger(sys.modules[*name*]) and send their log to a ELK stack (ElasticSearch, Logstash and Kibana). The web client is installed at https://kibana.cnf-da.tk/ and can be used to (I) analyse system resources and (II) logging activity of MADMex components.

Graphical user interface

A GUI is fully implemented to provide fast and easy access to MADMex functionality and supports comprehensive product operations. There is the web GUI component necessary and available in a local network and a browser client used by operators and supervisors. In the moment the GUI application is available here https://mm.cnf.gob.mx.

Project scope processing

The GUI application supports an end user to define the production of land cover and land cover change products by a project configuration file. This configuration file is a plain text JSON file describing all necessary parameters for each workflow - from data acquisition via ordering to mosaicing, post processing and meta data. generation.

Product catalog

MADMEX products are stored in a catalog system powered by ElasticSearch which already collected the processing log messages. The catalog consists of EO Data, MADMEX product components and the project data. Storage location of the data can be local partitions or cloud resources, like S3.

Actual MADMEX infrastructure

MADMEX installation of CONAFOR

The MADMEX installation for CONAFOR is located in Amazon Web Services (AWS) in the US-EAST-2 region.

S3 storage locations

All relevant data is located in two buckets in the region: eodata and mm-fao.

EC2 servers

Available services

AWS security guideline

AWS security guideline

General operations of MADMex production is performed in a cloud infrastructure like AWS. This ensure flexibility and scalability depending on the requirements of the production of a product. Like in all installation locations a set of security measurements should take place to secure the production process. The following guidelines are developed for AWS and use AWS specific services. But the concept should be easily adapted to other cloud vendors and can be applied to on-premise installations.

General recommendations

- limited access to specific users

  ```
  - Only admin should have access to AWS and servers
  ```

- monitor access

  ```
  - Configure CloudWatch to show raising / login to machines
  ```

- Monitor usage
- periodic rotate access keys and passwords
- Use IAM for services instead of individual API keys/secrets
- Create individual users/roles for single services instead of general admin accounts
- Read documentation to understand security issues
  - https://d1.awsstatic.com/whitepapers/Security/AWS_Security_Best_Practices.pdf
- Check location and access configuration of data resources
  - Limit public access to s3 buckets and objects and set permissions according to needs

Backup (for data security)

- Backup not used data from s3 to glacier
- Code repository serves as backup solution for code
- Backup components data if persistently needed
  - Log stack
  - RabbitMQ

Server level

- Install only necessary software and update software / kernel regularily
- Open only ports in security groups if necessary for the service
- passwordless access
  - Create individual ssh keys for maschines
- fail2ban service, change port number to obscure maschine login
- private VPC networks
  - Run services in private network, configure open ports directly to used IPs
- Create AMI as instance template to

Service level

- SSL secured connections with certificates
- Password secured api / web applications
- Secure external MADMex components (ElasticSearch, RabbitMQ, Redis)
  - Set passwords
    - Redirect traffic with HTTP servers (NGINX, Apache) using SSL and password authentification Recommended
    - Use private network to restrict access for necessary servers

Application level

- No development servers for python, instead uWSGI server and NGINX/APACHE to handle ssl and password access

Code level

- Passwords and credentials aren't stored in code or documentation, should be included as OS variables or environment files
- Logging won't show credentials or passwords

User groups

- *Admin*: administrators are allowed to access aws console and machines
- *User*: MADMex users are allowed to use services (e.g. GUI) and may start jobs, download data etc.

AWS configuration

There are a couple of security groups (SG) configured and used in a AWS installation. These groups are used for a MADMex worker and notebook services.

- *sg-fcf49f97* (celery): opens ports for celery services
- *sg-4f4ce027* (madmex-base): provides admin access to machines
- *sg-e9a39482* (madmex-logging): opens ports for log stack
- *sg-d1b750bb* (madmex-salt): manages ports for salt management

All ec2 instances use currently *madmex_conafor_180222* ssh keys for machine access but can be changed for All ec2 hosts currently run in *vpc-08107c61* as default VPC in AWS us-east-1 (Ohio) region.

New installation in AWS

New MADMex system installation

A Mad-Mex installation on a clean AWS EC2 instances is based on an Amazon Linux AMI 2018.03.0 (HVM), SSD Volume Type AMI to create a new AMI which can be used to launch quickly new processing workers. There are three installation types necessary:

- MADMex components
- MADMex worker installation
- MADMex notebook installation

Requisites:

- Hardware / instance:
    - 16GB min RAM
    - 50GB min harddisk
    - 4 CPUs
    - min 1GB network
- DNS services (with SSL certificates):
    - root (e.g. cnf-da.tk)
    - kibana.cnf-da.tk
    - cluster.cnf-da.tk
    - madmex-gui.cnf-da.tk
- Open network ports
    - see AWS security groups

Base OS installation

Following packages are necessary or provide base functionality working on the server:

```
mkdir /home/ec2-user/system
mkdir /LUSTRE/MADMEX/processing -p
sudo yum update -y
sudo yum install -y htop git fail2ban mosh tmux docker
sudo service fail2ban restart
sudo chkconfig fail2ban on
sudo yum erase ntp* -y
sudo yum install chrony -y
sudo service chronyd start
sudo chkconfig chronyd on

sudo usermod -aG docker $USER
sudo service docker restart
sudo chkconfig docker on
```

- Configure AWS cli with valid credentials: `aws configure`

Installation of metric collector (recommended)

```
cd /home/ec2-user/system
curl -L -O https://artifacts.elastic.co/downloads/beats/metricbeat/metricbeat-6.2.2-x86_64.rpm
sudo rpm -vi metricbeat-6.2.2-x86_64.rpm
```

- Edit `/etc/metricbeat/metricbeat.yml` or copy from s3 (s3://madmex-fao/madmex.v22/metricbeat.yml) to specify *kibana host* and *elasticsearch host list*

    ```
    sudo metricbeat modules enable system
    sudo metricbeat setup
    sudo service metricbeat start
    ```

MADMex components

MADMex components provide the base for a cluster based processing system on premise or in the cloud. The components include

- ELK stack (centralised logging)
    - Elasticsearch (document database)
    - Logstash (log collector)

    - Kibana (web application to browse Elasticsearch)

- RabbitMQ (message broker providing processing queues to celery)
- Redis (key/value database for celery results)
- Web services of these components can be redirected (reverse proxy configuration) via Nginx
- services need configuration files for each service at `/etc/nginx/conf.d`/

  ```
  sudo yum install -y nginx
  sudo service nginx restart
  sudo chkconfig nginx on
  ```

```
# MADMex system installation
mkdir /home/ec2-user/repo
cd /home/ec2-user/repo
git clone https://github.com/CONABIO/madmex-v2.git
git config --global credential.helper 'cache --timeout=2628000'
```

ELK stack installation

- adapt `/home/ec2-user/repo/madmex-v2/resources/docker-elk/docker-compose.yml` for ES2 data storage directory `` ` `` mkdir /home/ec2-user/system/elk/data -p sudo chown 1000 /home/ec2-user/system/elk/data

cd /home/ec2-user/system sudo curl -L https://github.com/docker/compose/releases/download/1.21.2/docker-compose-$(uname -s)-$(uname -m) -o /usr/local/bin/docker-compose sudo chmod +x /usr/local/bin/docker-compose cd /home/ec2-user/repo/madmex-v2/resources/docker-elk docker-compose up

```
* Kibana needs to build indices for elasticsearch
    * Management
    * Index Pattern
    * Create Index (specify index name)

*Test services:*
* Kibana with port IP:5601

---
### RabbitMQ
* RabbitMQ needs special security group to open ports
`docker run -d -p 5672:5672 -p 8080:15672 --hostname cnf-da.tk  -e RABBITMQ_DEFAULT_USER=admin -e RABBITMQ_DEFAULT_PASS=<password>  --name madmex_rabbitmq
 rabbitmq:3-management`

*Additional configuration*
* Create madmex user in RabbitMQ admin console: IP:8080
    * Admin
    * Create User madmex:password
    * Set all privileges to madmex user
* Access to RabbitMQ services only work with *external/public IP* !!!
---

### Redis DB
* Redis needs special security group to open ports
```

mkdir -p /home/ec2-user/system/redis sudo chown 999:999 /home/ec2-user/system/redis

docker run -d -p 6379:6379 --name madmex_redis -v /home/ec2-user/system/redis:/data redis redis-server --appendonly yes --requirepass

```
* No additional configuration is needed - maybe testing if service is reachable with `redis-cli`
---

### Flower (celery management GUI)
* flower needs special security group to open ports or Nginx configuration
* example Nginx config is at s3://madmex-fao/madmex.v22/config/nginx_flower.conf
* nginx needs virtual domain names (e.g. cluster.cnf-da.tk) registered and configured
* to secure connection in addition SSL certificates are necessary, too (e.g. letsencrypt.org)
```

docker run -p 8888:5555 --name madmex_flower -e TZ=UTC -v /etc/localtime:/etc/localtime:ro -e BACKEND_URL=XXX -e BROKER_API_URL=madmex:XXX@cnf-da.tk:8080 -e BROKER_URL=madmex:XXX@cnf-da.tk -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini -d -w /madmex/interfaces/worker/ -e WORKER_TZ=UTC -v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex flower

```
---

## MADMex worker installation
* Check variables and configuration
* Every new worker needs an unique name using `-n worker1@%h` or use starter script
```

aws s3 cp s3://madmex-fao/madmex.v22/opencv_contrib.tar.gz /home/ec2-user/repo/madmex-v2/resources/docker-containers/gdal-segment/ cd madmex-v2/ make worker # needs 2-3 min.

cd /home/ec2-user/repo/madmex-v2/resources/docker-containers && make # needs 15-20 min.

```
*Executing madmex worker*
```

docker run -e USGS_USER=XXX -e USGS_PASSWORD=XXX —ulimit nofile=98304:98304 -e USE_SERIAL_IMPLEMENTATION=1 -v /var/run/docker.sock:/var/run/docker.sock —hostname $(curl http://169.254.169.254/latest/meta-data/hostname) -e TZ=UTC -v /etc/localtime:/etc/localtime:ro -e BACKEND_URL=redis://:XXX@cnf-da.tk -e BROKER_API_URL=madmex:XXX@cnf-da.tk:8080 -e BROKER_URL=madmex:XXX@cnf-da.tk -e PUSHOVER_USERKEY=XXX -e PUSHOVER_APITOKEN=XXX -e PYTHONUNBUFFERED=1 -e DB_PORT=5432 -e AWS_ACCESS_KEY_ID=XXX -e AWS_SECRET_ACCESS_KEY=XXX -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini —rm -it -w /madmex/interfaces/worker/ -e WORKER_TZ=UTC -v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex celery -A worker worker —loglevel=info —concurrency 1 -n worker1@%h -Ofair

```
* or *adapt* and use start script at s3://madmex-fao/madmex.v22/scripts/start_worker.sh
```

```
-h  show this help text
-n  set number of worker (default is 1)
-m  Build worker instance
-u  Update repository from github
```

```
## MADMex notebook installation
* Jupyter needs special security group to open ports
* Jupiter needs valid SSL certificates to provide https (e.g. from [Let’s Encrypt - Free SSL/TLS Certificates](https://letsencrypt.org)
* (optional): create self-signed certificates using openssl  ([How To Create a Self-Signed SSL Certificate for Nginx in Ubuntu 16.04 | DigitalOcean](https://www.digitalocean.com/community/tutorials/how-to-create-a-self-signed-ssl-certificate-for-nginx-in-ubuntu-16-04)) and change references in following notebook docker start command
```

openssl req -x509 -nodes -days 365 -newkey rsa:2048 -keyout /home/ec2-user/system/cert/jupyter-selfsigned.key -out /home/ec2-user/system/cert/jupyter-selfsigned.crt

```
* Check configuration file (password, certificates, etc.) in `/home/ec2-user/repo/madmex-v2/jupyter_lab_config.py`
```

cd /home/ec2-user/repo/madmex-v2/ make # needs 10 min.

Starting notebook

docker run --rm -it -e BACKEND_URL=XXX -e BROKER_API_URL=XXX -e BROKER_URL=madmex: qKr49EXb86MB@cnf-da.tk -e AWS_ACCESS_KEY_ID=XXX -e AWS_SECRET_ACCESS_KEY=XXX -e CATALOG_DB=sqlite:////data/temp/catalog.sqlite3 -v /externo/notebooks/:/LUSTRE -v /externo/data/:/data --ulimit nofile=98304:98304 -v /etc/letsencrypt/live/cnf-da.tk/privkey.pem:/keys/mykey.key -v /etc/letsencrypt/live/cnf-da.tk/cert.pem:/keys/mycert.pem -p 7777:7777 madmex/lab jupyter notebook --ip='*' --port 7777 `

Starting workers on running AWS EC2 instances

Starting an AWS instance as MADMex worker

The fast launch of new MADMex workers needs the following prerequisites to ensure a working system:

- Existing worker AMI (Amazon Maschine Image) e.g. *ami-9c132cf9* (madmex_worker_020618)
- Existing security groups to open necessary ports
- Running celery components and logging stack (ELK) configured in MADMex configuration file (/madmex/resources/config/configuration.latest.local.ini)

Worker nodes can be added later during production following the section *Starting new ec2 instance from AWS Console*. To remove a running worker it can be shut down from AWS console (may result in current job losses).

AMI build steps

- Update start_worker.sh script in /home/ec2-user/repo/madmex.v22 directory.
- Build all docker images (MADMex and external dockers from resources/docker-containers)
- Check if everything works

Starting new ec2 instance from AWS Console

- login to console and switch to ec2 services
- launch ec2 instance
    - select private AMI (e.g. *madmex_worker_020618*)
    - select proper ec2 machine type (e.g. m5.2xlarge or m5.4xlarge) depending on job demands
    - define number of machines, request spot instance to reduce price
    - check storage demand
    - skip tags section or provide tag names to annotate machines
    - select security groups (from wiki page AWS security)
    - select latest ssh key and launch machines

The new MADMex worker should be shown in flower management console after a couple of minutes. In addition the server can be monitored in Kibana dashboard in the *[Metricbeat System] Overview* section or in the AWS console. A supervisord config exists in the current AMI and starts one worker using the /start_worker.sh/ script. If a worker is allowed to execute more jobs in parallel it is possible to login to the server manually and use the script. Moreover, workers and be started or stopped with the salt management tool.

AWS infrastructure configuration

Used AWS infrastructure

- Notebook server
- Processing nodes

Default settings:

- git settings `git config --global credential.helper 'cache --timeout=2628000'`
- aws configure
- server monitoring with Kibana) `service metricbeat restart`

Notebook server

Special configuration for notebook server:

- Notebook server services `sudo mount /dev/nvme1n1 /externo/` Several software components are installed and configured at the notebook server:
- rabbitMQ
- redis
- flower
- ELK stack

Additional service installation

- rabbitMQ `docker run -d -p 5672:5672 -p 8080:15672 --hostname cnf-da.tk -e RABBITMQ_DEFAULT_USER=admin -e RABBITMQ_DEFAULT_PASS=<password> --name madmex_rabbitmq rabbitmq:3-management`
- redis `docker run -d -p 6379:6379 --name madmex_redis -v /externo/system/redis:/data redis redis-server --appendonly yes --requirepass <password>` Testing redis: `redis-cli -h cnf-da.tk`
- https://cnf-da.tk:7777 for notebook
- https://kibana.cnf-da.tk for kibana/log access

Initialization: `docker run --rm -it -e BACKEND_URL=XXX -e BROKER_API_URL=XXX -e BROKER_URL=madmex:qKr49EXb86MB@cnf-da.tk -e AWS_ACCESS_KEY_ID=XXX -e AWS_SECRET_ACCESS_KEY=XXX -e CATALOG_DB=sqlite:////data/temp/catalog.sqlite3 -v /externo/notebooks/:/LUSTRE -v /externo/data/:/data --ulimit nofile=98304:98304 -v /etc/letsencrypt/live/cnf-da.tk/privkey.pem:/keys/mykey.key -v /etc/letsencrypt/live/cnf-da.tk/cert.pem:/keys/mycert.pem -p 7777:7777 madmex/lab jupyter notebook --ip='*' --port 7777`

`docker run --rm -it -e BACKEND_URL=XXX -e BROKER_API_URL=XXX -e BROKER_URL=XXX -e AWS_ACCESS_KEY_ID=X -e AWS_SECRET_ACCESS_KEY=XXX -e CATALOG_DB=sqlite:////data/temp/catalog.sqlite3 -v /var/run/docker.sock:/var/run/docker.sock --ulimit nofile=98304:98304 -v /externo/notebooks/:/LUSTRE -v /externo/data/:/data -v /etc/letsencrypt/live/cnf-da.tk/privkey.pem:/keys/mykey.key -v /etc/letsencrypt/live/cnf-da.tk/cert.pem:/keys/mycert.pem -p 7778:7778 madmex/lab jupyter lab --port 7778`

`docker run -e ES_JAVA_OPTS="-Xms1024m -Xmx2048m" --rm -it --name elasticsearch -v /externo/system/esdata:/usr/share/elasticsearch/data -p 9200:9200 -p 9300:9300 elasticsearch`

`docker run --rm -it --name logstash -v /externo/system/logstash/conf.d:/etc/logstash/conf.d:ro -v /externo/system/var/log:/host/var/log -p 5050:5050 --net host logstash logstash -f /etc/logstash/conf.d`

```
docker run -d --name kibana -p 5601:5601 -e ELASTICSEARCH_URL=http://localhost:9200 --net host kibana
```

```
docker run -e PYTHONUNBUFFERED=1 -e DB_PORT=5432 -e AWS_ACCESS_KEY_ID=XXX -e DB_USER=madmex_root
-e DB_URL=18.221.49.107 -e AWS_SECRET_ACCESS_KEY= XXX -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini \
-e DB_NAME=madmex_db2 -e DB_PASSWORD=123  --rm -it -v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex /madmex/interfaces/cli/madmex_processing.py -h
```

Execute flower manager

```
docker run -p 8888:5555  --name madmex_flower -e TZ=UTC  -v /etc/localtime:/etc/localtime:ro -e BACKEND_URL= XXX -e BROKER_API_URL=madmex:XXX@cnf-da.tk:80
80 -e BROKER_URL=madmex: XXX@cnf-da.tk  -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini -d -w /madmex/interfaces/worker/ -e WORKER_T
Z=UTC -v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex flower
```

Execute worker

```
docker run -e USGS_USER=XXX -e USGS_PASSWORD=XXX --ulimit nofile=98304:98304 -e USE_SERIAL_IMPLEMENTATION=1 -v /var/run/docker.sock:/var/run/docker.sock
--hostname $(curl http://169.254.169.254/latest/meta-data/hostname)  -e TZ=UTC  -v /etc/localtime:/etc/localtime:ro -e BACKEND_URL=redis://:XXX@cnf-da.tk
-e BROKER_API_URL=madmex:XXX@cnf-da.tk:8080 -e BROKER_URL=madmex:XXX@cnf-da.tk  -e PUSHOVER_USERKEY=XXX -e PUSHOVER_APITOKEN=XXX -e PYTHONUNBUFFERED=1 -e
DB_PORT=5432 -e AWS_ACCESS_KEY_ID=XXX -e AWS_SECRET_ACCESS_KEY=XXX -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini --rm -it -w /madm
ex/interfaces/worker/ -e WORKER_TZ=UTC -v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex  celery -A worker worker --loglevel=info  --concurrency 1 -n worker1@%h  -
Ofair
```

Or Priority worker

```
docker run --ulimit nofile=98304:98304 -v /var/run/docker.sock:/var/run/docker.sock  --hostname $(curl http://169.254.169.254/latest/meta-data/hostname)
-e TZ=UTC  -v /etc/localtime:/etc/localtime:ro -e BACKEND_URL=redis://: XXX@cnf-da.tk -e PUSHOVER_USERKEY= XXX -e PUSHOVER_APITOKEN= XXX -e BROKER_API_URL
=madmex: XXX@cnf-da.tk:8080 -e BROKER_URL=madmex: XXX@cnf-da.tk  -e PYTHONUNBUFFERED=1 -e DB_PORT=5432 -e AWS_ACCESS_KEY_ID= XXX -e AWS_SECRET_ACCESS_KEY=
 XXX -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini --rm -it -w /madmex/interfaces/worker/ -e WORKER_TZ=UTC -v /LUSTRE/MADMEX/:/LUS
TRE/MADMEX madmex  celery -A worker worker --loglevel=info --beat --concurrency 1 --beat -n prior1@%h -Q priority.high
```

Or

```
git pull && sudo make worker && sudo docker run --ulimit nofile=98304:98304 -v /var/run/docker.sock:/var/run/docker.sock  --hostname $(curl http://169.254
.169.254/latest/meta-data/hostname) -e TZ=UTC  -v /etc/localtime:/etc/localtime:ro -e PUSHOVER_USERKEY= XXX -e PUSHOVER_APITOKEN= XXX -e BACKEND_URL=redis
://: XXX@cnf-da.tk -e BROKER_API_URL=madmex: XXX@cnf-da.tk:8080 -e BROKER_URL=madmex: XXX@cnf-da.tk  -e PYTHONUNBUFFERED=1 -e AWS_ACCESS_KEY_ID= XXX -e DB
_USER=madmex_root -e AWS_SECRET_ACCESS_KEY= XXX -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini -e DB_NAME=madmex_db2 -e DB_PASSWORD
=123  --rm -it -w /madmex/interfaces/worker/ -e WORKER_TZ=UTC -v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex  python /madmex/interfaces/worker/starter.py start_
worker
```

send jobs

```
docker run -e BACKEND_URL=redis://: XXX@cnf-da.tk -e BROKER_API_URL=madmex: XXX@cnf-da.tk:8080 -e BROKER_URL=madmex: XXX@cnf-da.tk  -e PYTHONUNBUFFERED=1
-e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini --rm -it -w /madmex/interfaces/worker/ madmex  python /madmex/interfaces/worker/star
ter.py exec S2Download --kwarg entity_id S2A_OPER_PRD_MSIL1C_PDMC_20160130T060756_R012_V20160129T173241_20160129T173241 --queued True
```

execute job directly

```
docker run -v /var/run/docker.sock:/var/run/docker.sock --hostname worker1 -e TZ=UTC  -v /etc/localtime:/etc/localtime:ro -e BACKEND_URL=redis://: XXX@cnf
-da.tk -e BROKER_API_URL=madmex: XXX@cnf-da.tk:8080 -e BROKER_URL=madmex: XXX@cnf-da.tk  -e PYTHONUNBUFFERED=1 -e AWS_ACCESS_KEY_ID= XXX -e AWS_SECRET_ACC
ESS_KEY= XXX -e MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini --rm -it -w /madmex/interfaces/worker/ -e WORKER_TZ=UTC -v /LUSTRE/MADM
EX/:/LUSTRE/MADMEX madmex  python /madmex/interfaces/worker/starter.py exec S2GetPreprocess --kwarg entity_id S2A_MSIL1C_20170103T172712_N0204_R012_T13QFD
_20170103T173023 --queued True
```

execute job directly

```
docker run -v /var/run/docker.sock:/var/run/docker.sock --hostname worker1 -e TZ=UTC  -v /etc/localtime:/etc/localtime:ro -e BACKEND_URL=redis://:wP6z622J
VrmM@cnf-da.tk -e BROKER_API_URL=madmex: XXX@cnf-da.tk:8080 -e BROKER_URL=madmex: XXX@cnf-da.tk  -e PYTHONUNBUFFERED=1 -e AWS_ACCESS_KEY_ID= XXX -e DB_USE
R=madmex_root -e DB_URL=18.221.49.107 MRV_CONFIG=/madmex/resources/config/configuration.latest.local.ini --rm -it -w /madmex/interfaces/worker/ -e WORKER_
TZ=UTC -v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex  python /madmex/interfaces/worker/starter.py exec S2GetPreprocess --kwarg entity_id S2A_MSIL1C_20170103T17
2712_N0204_R012_T13QFD_20170103T173023 --kwarg tile 13QFD --queued False --waiting False --kwarg clean_up False
```

purge/clean queue

```
docker run  -e BACKEND_URL=redis://: XXX@cnf-da.tk -e BROKER_API_URL=madmex: XXX@cnf-da.tk:8080 -e BROKER_URL=madmex: XXX@cnf-da.tk -e MRV_CONFIG=/madmex/
resources/config/configuration.latest.local.ini  --rm -it -w /madmex/interfaces/worker/ madmex  celery -A worker purge
```

MADMEX environmental variables

MADMEX OS variables

The system requires a set of predefined variables before executing any MADMEX functionality. All variables should be set for a docker execution, within a shell script or directly after system start in e.g. `.bashrc` file.

Mandatory variables

- *MRV_CONFIG*: There are several configuration files included in the code repository in `resources/config/`. Developers are encouraged to create their own configuration file with their settings. The latest cloud configuration is `configuration.latest.aws-cnf.ini`. Without the variable set MADMEX won't run.
- *DOCKER_CONFIG*: used for intra docker communication, on Linux systems default is `/var/run/docker.sock`. Developer can redirect to their network docker services.
- *ES_USER*: ElasticSearch username, per default set to `madmex
- *ES_PASSWORD*: Password for the ElasticSearch server used by catalog functions.
- *ES_HOST*: ElasticSearch host name, per default `logging.mm.cnf.gob.mx`
- *CELERY_CONFIG*: Celery configuration already defined in MADMEX configuration
- *BACKEND_URL*: Url to Redis server used as output variable backend for celery processes. Layout is `redis://:[REDIS PASSWORD]@[REDIS SERVER]`
- *BROKER_API_URL*: Url to RabbitMQ server used to control/manage Celery. Layout is `[RABBIT USER]:[RABBIT PASSWORD]@[RABBIT MQ server]:8080`
- *BROKER_URL*: Url to RabbitMQ server used as processing queue server in Celery. Layout is `[RABBIT USER]:[RABBIT PASSWORD]@[RABBIT MQ server]`
- *MADMEX_LOG_FILE*: Location of external log file. In addition all logging will be send to Logstash - a centralised logging collector. Defined as well in Madmex configuration file
- *AWS_ACCESS_KEY_ID*: AWS credential key used for S3 access (needs full access)
- *AWS_SECRET_ACCESS_KEY*: AWS credential secret used for S3 access
- - USGS_USER*: User name of USGS account used to order Landsat data
- - USGS_PASSWORD*: Corresponding password of USGS account used to order Landsat data
- *GUI_USERS*: Yaml file already created with the user-manager CLI containing login credentials for the MADMEX GUI

Deprecated variables:

- *ANC_PATH*: auxiliary path containing DEM, etc. for leads processing
- *SunEarthDistance*: location of `SunEarthDistance.txt` used by Fmask docker
- *DB_USER*: postgres user
- *DB_PASSWORD*: postgres password
- *DB_URL*: psycopg2 compatible url
- *DB_PORT*: postgres port
- *DB_NAME*: postgres database name

SALT management tool

SALT management tool

Salt tool SaltStack is intelligent automation for a software-defined world helps to administrate computing nodes in a cluster.

Salt management

docker build -t madmex/salt_master . Start master

```
docker run --rm -it --name salt-master -v /externo/repos/madmex-v2/resources/management/salt/config/:/config -v /externo/repos/madmex-v2/resources/managem
ent/salt/data/:/data -p 4505:4505 -p 4506:4506 -p 8443:443 -p 8000:8000 madmex/salt_master
```

/Execute salt command/

```
docker exec   $(docker ps -q -f name=salt-master) salt-key -A -y
docker exec   $(docker ps -q -f name=salt-master) salt-key -d '*' -y
docker exec  $(docker ps -q -f name=salt-master) salt '*' test.ping
docker exec  $(docker ps -q -f name=salt-master) salt '*' cmd.run 'w'
```

/System management/

```
docker exec  $(docker ps -q -f name=salt-master)  salt '*' system.reboot
docker exec  $(docker ps -q -f name=salt-master)  salt '*' system.poweroff
```

Update process

- cd /home/ec2-user/repo/madmex-v2/ ; git pull ; make worker

Sentinel2 data acquisition

Sentinel2 data

Sentinel2 data is provided by ESA via different services. The ESA open data hub is the most important application. It requires a free user account and can be used to query for available data and data download. The EOSS catalog application simplifies the data search for geographic regions. IN collaboration with Amazon Web Services ESA mirrors complete Sentinel data in AWS S3. The latest data policy required an AWS account for s3 data download to pay for data transfer out of AWS.

All mentioned functionality is implemented as docker containers which needs to be build during MADMEX installation procedure.

Data ordering

No special data ordering is necessary to work with Sentinel-2 data. Because all available data is stored also in S3, demanded data can be downloaded using the tool sentinel hub. Based on the tool a docker is configured to provide a Madmex adapter `S2Download`. A valid AWS credential is necessary for the download. For the download the product name is necessary to find the data. The parameter can be identified with the above mentioned tools. In 2016 there was a naming/structure change of the S2 products. Before a product could consists of multiples tiles depending on the overpass. After the change one product consists only of one image tile. To speed up processing an optional `tile` parameter can specify the needed tiles of these older products. Otherwise the whole product is downloaded, rest. pre processed and imported into S3.

```
from adapters.external.sentinel_hub import S2Download

s2_download = S2Download()
s2_download.setInput("entity_id", "S2B_MSIL1C_20180920T171959_N0206_R012_T13QFB_20180920T221802")
s2_download.setInput("tile", "13QFB")
s2_download.setInput("compress", True)
s2_download.setInput("clean_up", False)
s2_download.setInput("output_folder", "/LUSTRE/MADMEX/processing/s2")

s2_download.run()
result = s2_download.output_variables["result"].result()
print result
```

The example will return a compressed ZIP of the data set, if compress is set to false the plain directory is returned. Data is available as level L1C. For the time series processing L2 surface reflectance data is needed.

Data preprocessing

After data download the preprocessing should be executed, including the following adapters: `Sen2CorProcessingSentinel`, `FmaskProcessingSentinel`, `RasterizeMask`. These adapters will convert L1C to L2, calculates additional cloud masks and applies the mask to the image data. All these adapters are executed with the `S2GetPreprocess` adapter.

```
from adapters.workflow.acquisition.s2_download_preprocessing import S2GetPreprocess

s2 = S2GetPreprocess()
s2.setInput("entity_id", "S2B_MSIL1C_20180920T171959_N0206_R012_T13QFB_20180920T221802")
s2.setInput("tile", "13QFB")
s2.setInput("clean_up", False)
s2.setInput("output_folder", "/LUSTRE/MADMEX/processing/s2")

s2.run()
result = s2.output_variables["result"].result()
print result
```

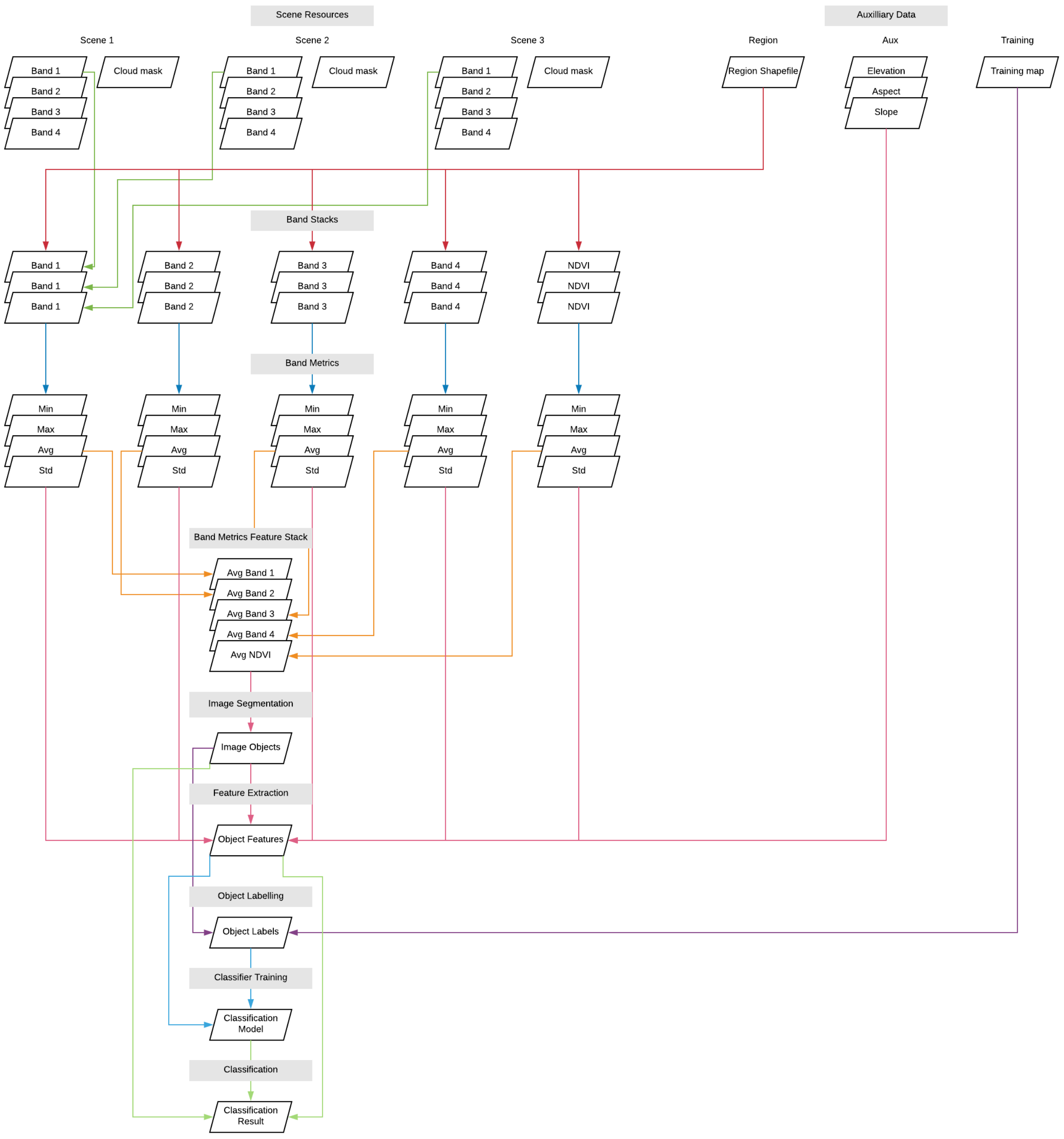
Ingested scenes are stored to s3://eodata as specified in the Mad-Mex configuration file after successful download and pre processing. The generated products/components are automatically synronized with the MADMEX catalog in these partitions.

Land cover concept

Method abstract

In a tile/granule based approach all defined Landsat/Sentinel2 imagery represented as surface reflectances and with calculated cloud mask are used for multi-temporal land cover classification. From each individual scene additional vegetation indices can be computed. All of those indices and the original spectral bands are combined to time-series as band stacks. From those selected descriptive statistics for each pixel are calculated as time-series metrics. In the processes of band stacking, pixels labeled by the cloud mask are omitted, thus producing cloud and cloud shadow free image metrics. Subsequently, image segmentation is done on a new spectral band stack extracted from a defined metric value (e.g. average or minimum) from all calculated metrics. Following that feature extraction is performed on all time-series metrics, and on optional auxiliary data like elevation, slope or aspect datasets. Segmented objects are then labeled against a training dataset by conserving only those objects, which are spatially intersected by only one class in the training dataset. In order to create a clean training dataset a class based outlier removal can optionally be performed. Based on the samples that passed the outlier removal a C5 decision tree classifier is trained and applied to all image objects. Optionally a 10-folded classification tree can be trained by enabling training boosting. Classified vector objects are then transformed to raster representation.

Data flows



Result structure

```
├── aux
│   ├── Entidades_2013.dbf
│   ├── Entidades_2013.prj
│   ├── Entidades_2013.shp
│   └── Entidades_2013.shx
├── bandmetrics
│   ├── landsat_stack_28046_NDVI_metrics.tif
│   ├── landsat_stack_28046_band2_metrics.tif
│   ├── landsat_stack_28046_band3_metrics.tif
│   ├── landsat_stack_28046_band4_metrics.tif
│   ├── landsat_stack_28046_band5_metrics.tif
│   ├── landsat_stack_28046_band6_metrics.tif
│   └── landsat_stack_28046_band7_metrics.tif
├── bandmetricsstacks
│   └── landsat_metricstack_28046_average.tif
├── bandstacks
│   ├── landsat_stack_28046_NDVI.tif
│   ├── landsat_stack_28046_band2.tif
│   ├── landsat_stack_28046_band3.tif
│   ├── landsat_stack_28046_band4.tif
│   ├── landsat_stack_28046_band5.tif
│   ├── landsat_stack_28046_band6.tif
│   └── landsat_stack_28046_band7.tif
├── classification
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp.c5
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp.cases
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp.data
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp.names
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp.result
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp.tmp
│   └── landcover_28046_classify_CSR25_JAL_2015_25m_warp.tree
├── features
│   ├── landsat_stack_28046_NDVI_metrics.tif.zstat
│   ├── landsat_stack_28046_band2_metrics.tif.zstat
│   ├── landsat_stack_28046_band3_metrics.tif.zstat
│   ├── landsat_stack_28046_band4_metrics.tif.zstat
│   ├── landsat_stack_28046_band5_metrics.tif.zstat
│   ├── landsat_stack_28046_band6_metrics.tif.zstat
│   └── landsat_stack_28046_band7_metrics.tif.zstat
├── result
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp_.dbf
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp_.prj
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp_.shp
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp_.shx
│   ├── landcover_28046_classify_CSR25_JAL_2015_25m_warp_.tif
│   └── landcover_28046_classify_CSR25_JAL_2015_25m_warp__confidence.tif
├── segmentation
│   ├── landsat_metricstack_28046_average.tif_40_02_08.dbf
│   ├── landsat_metricstack_28046_average.tif_40_02_08.prj
│   ├── landsat_metricstack_28046_average.tif_40_02_08.shp
│   ├── landsat_metricstack_28046_average.tif_40_02_08.shx
│   ├── landsat_metricstack_28046_average.tif_40_02_08.tif
│   └── landsat_metricstack_28046_average.tif_40_02_08_training_lcc.tif
└── training
    └── CSR25_JAL_2015_25m_warp.tif
```

Land cover production

Land Cover Production

The tile based multi-temporal land cover adapter

The tile based multi-temporal land cover process is implemented in the adapter `LandcoverMultitemporalWorkflow` which is defined by the following input parameters:

`working_dir` [optional]

This targets to a local directory where the processing results shall be stored. If this varaiable is not defined the workind directory will be created automatically.

`granule`

This variable defines the Landsat path/row (e.g. `018047`) or the Sentinel2 granule identifier (e.g. `13QED`).

`scene_folder_list`

This variable sets the locations of the `SceneResources` locations, either local or on AWS S3, as list or comma separated string. E.g. `['s3://madmex-fao/eodata/oli_tirs/29047/2016/2016-01-10/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-01-26/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-02-11/']` or `/Volumes/local_data/temp/madmex/28046/2017/2017-10-20,/Volumes/local_data/temp/madmex/28046/2017/2017-12-23,/Volumes/local_data/temp/madmex/28046/2017/2017-09-18,/Volumes/local_data/temp/madmex/28046/2017/2017-03-10,/Volumes/local_data/temp/madmex/28046/2017/2017-05-13,/Volumes/local_data/temp/madmex/28046/2017/2017-07-16,/Volumes/local_data/temp/madmex/28046/2017/2017-11-05,/Volumes/local_data/temp/madmex/28046/2017/2017-01-21,/Volumes/local_data/temp/madmex/28046/2017/2017-08-01,/Volumes/local_data/temp/madmex/28046/2017/2017-02-06`

`land_mask_shape`

This variable sets the location, either locally or on AWS S3, of a polygon Shapefile to be used for defining a validity mask for the classifications. This means, all image pixels outside the given polygons will be masked out in processing. This is especially helpful and important for masking out ocean and/or to run classifications for given regions or states only. Example: `s3://madmex-fao/mexico_admin/Entidades_2013.shp`

`aux_image_list`

This variable sets the locations, either locally or on AWS S3, as list or comma separated string, of additional images to be used in feature extraction and classification. This images must not share same projections as the satellite images but shall of course overlap the scenes/tile/granule geographic area. Example: `['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif]`

`training_url`

This variable sets the location, either locally or on AWS S3, of the training image to be used in classification. he training dataset shall overlap the scenes/tile/granule geographic area and must consist of one band with pixel values representing the respective class. No data must be valued 0. Example: `s3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif`

`mask_url` [optional]

This variable sets the location, either locally or on AWS S3, of a mask image to be used during classification. This variable is optional and might be applied if the user wants to exclude specific areas in the image region as for example a change mask, a water mask or a urban mask or a combination of many. The image must have one band and no data must be labelled 0. All image regions with mask pixel values different from zero will be used.

`bands` [optional]

This variable defines the spectral bands to include in land cover processing, as list or comma separated string. This parameter is optional, if it is not given all bands from the current sensor will be used. For the supported sensors those are:

- Landsat 5 and Landsat 7
  - `['band1','band2','band3','band4','band5','band7']`
- Landsat 8
  - `['band2','band3','band4','band5','band6','band7']`
- Sentinel 2
  - `['B02_10m','B03_10m','B04_10m','B08_10m', 'B04_20m', 'B05_20m','B06_20m', 'B07_20m', 'B8A_20m', 'B11_20m', 'B12_20m']`

`features` [optional]

This variable defines the additional spectral bands features to include in land cover processing, as list or comma separated string. This parameter is optional, if it is not given all features from the current sensor will be used. Note: This features are calculated from secific bands and these bands must be given in the `bands` variable. For the supported sensors those are:

- Landsat 5 and Landsat 7
  - `['NDVI','SR','EVI','ARVI']`
- Landsat 8
  - `['NDVI','SR','EVI','ARVI']`

- Sentinel 2
  - `['NDVI','NDVI20']` (either 10m or 20m resolution)

`metrics_features` [optional]

This variable defines the statistics to be calculated from the scene timeseries for any specified band and feature during claulcation of time series metrics, as list or comma separated string. This parameter is optional, if it is not given standard supported features will be used. The standard features are:

- `["minimum","maximum","range","average","stddev","quantiles25"]`

- Additional supported features `quantiles10`

`segmentation_metric`

This variable defines the metrics feature to be used for creating the band metrics feature stack to be used for image segmentation. This can be one of `"minimum","maximum","range","average","stddev"`. Note: The used feature must be given in the `metrics_features` variable.

`segmentation_method`

This variable defines the segmentation method to be used. This can bei either `bis` for the BerkeleyImageSegmentation or one of `SLIC, SLICO, LSC or SEEDS` for the OpenCSV segmentation. Note: `bis` requires license.

`segmentation_params`

This variable defines the parameters of the defined segmentation method, as list or comma separated string. For the supported segmentation methids these are:

- BIS
  - `[segmentation_threshold, segmentation_shape, segmentation_compactness]`

  - Example: `'40,0.2,0.8'` or `[40,0.2,0.8]`

  - whereby
    - segmentation_threshold [BIS threshold, the number of region merging iterations]

    - segmentation_shape [BIS shape parameters ranging from 0..1]

    - segmentation_compactness [BIS compactness parameters ranging from 0..1]

- SLIC, SLICO, LSC or SEEDS
  - `[iteration, region_size]`

  - Example: `'10,5'` or `[10,5]`

  - whereby
    - iteration [number of merge iterations, default 10]

    - region_size [approximate object size (diameter, edge length) in pixels]

`sampling_params` [optional]

This optional parameter sets up sampling option for the training dataset, as list or comma separated string. This will perform random stratified class weighted sampling of the training dataset to be used for classification. The parameters define percentage of samples per class, maximum naumber of samples per class, size of the sample chip in pixels, number of iterations e.g. `0.05,1000,30,2`. If iterations higher 1 is given, multiple classifications will be performed on multiple different sampled training datasets.

`training_percentage`

This variable defines in values of percentage [0.1 .. 99.9] how many samples shall be used in classificator training. Those are selected randomly to train the classifier, while the remaining samples will be used for classification evaluation. Example: `60` will use randomly selected 60% of all samples for training while the remaining 40% are used for classification validation.

`training_overlay_percentage`

This variable defines percentage of training object to segmentation object overlay in ranges between [0..1]. E.g. `0.7` means 70 percent of an image object must be covered by unique class values from the training dataset. If this is not the case, the image object will not be labelled with a class and will not be included as sample for classification training. This shall ensure the inclusion of the most clean image objects in classification training. Note: For training datsets that represent only small samples on the ground, which might be much smaller then the average object size from image segmentation this value must be very low.

`boosting`

This variable defines whether to use a boosted classifier or not. `0` means no boosting. `1` means boosting. If boosting is enabled by `1` 10 classifiers will be trained and their decision trees are combined to a final decision tree model.

`outlier` This variable defines whether to use sample outlier elimination or not. `0` means no elimination. `1` means do outlier elimination. If outlier elimination is enabled the class based iterative histogram trimming is applied in order to remove outlier samples for any class that do not fit in the class distribution. The outlier elimination is useful for very large training datasets, e.g. full coverage maps, to conserve only most clean training samples.

`outlier_init` [optional]

If outlier elimination is enabled this variable can be set to define the percentage of initial removal of unlikely samples per class. For example `.2` will remove the 20% most unlikely incoming samples for any class and will then continue the iterative histogram trimming and outlier removal for the remaining samples.

`aws_s3_bucket` [optional]

This optional variable defines the AWS S3 bucket where the final results shall be stored. For example `madmex-fao` indicates the madmex-fao bucket at `s3://madmex-fao/`.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`aws_s3_key` [optional]

This optional variable defines the AWS S3 key where the final results shall be stored. For example `projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047` together with bucket defined as `madmex-fao` indicates `s3://madmex-fao/projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047` as final storage destination.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`overwrite` [optional]

This optional boolean variable can be set to `True` or `False`. This is especially helpful during development and debugging. If overwriting is disabled the most time consuming calculations like band stacking or image segmentation will not be executed again, if those intermediate results are already available in the defined `working_dir`. Default value is `True`.

`clean_up` [optional]

This optional boolean variable can be set to `True` or `False` and defines whether to remove the `working_dir` after process has finished. Default value is `True`.

Python execution

```
from adapters.workflow.landcover.landcover_multitemporal import LandcoverMultitemporalWorkflow
landCoverProcess = LandcoverMultitemporalWorkflow()
landCoverProcess.setInput('working_dir', '/Users/tmp/')
landCoverProcess.setInput('scene_folder_list', ['s3://madmex-fao/eodata/oli_tirs/29047/2016/2016-01-10/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-01-26/', 's3://
madmex-fao/eodata/oli_tirs/29047/2016/2016-02-11/'])
landCoverProcess.setInput('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp')
landCoverProcess.setInput('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif')
landCoverProcess.setInput('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif])
landCoverProcess.setInput('granule', 029047)
landCoverProcess.setInput('bands', ['band2','band3','band4','band5'])
landCoverProcess.setInput('features', ['NDVI'])
landCoverProcess.setInput('metrics_features', 'minimum,maximum,average,quantiles25')
landCoverProcess.setInput('segmentation_metric', 'average')
landCoverProcess.setInput('segmentation_method', 'bis')
landCoverProcess.setInput('segmentation_params', '40,0.2,0.8')
landCoverProcess.setInput('training_overlay_percentage', 0.7)
landCoverProcess.setInput('training_percentage', 50)
landCoverProcess.setInput('boosting', 1)
landCoverProcess.setInput('sampling_params', [0.15,1000,30,3])
landCoverProcess.setInput('outlier', 1)
landCoverProcess.setInput('aws_s3_bucket', 'madmex-fao')
landCoverProcess.setInput('aws_s3_key', 'projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047')
landCoverProcess.setInput('clean_up',False)
landCoverProcess.run()
landover_results = landCoverProcess.output_variables["result"].getVariable()
```

Job execution

`kwargs = [('scene_folder_list', ['s3://madmex-fao/eodata/oli_tirs/29047/2016/2016-01-10/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-01-26/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-02-11/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-03-14/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-03-30/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-05-01/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-05-17/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-06-02/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-09-06/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-10-24/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-11-25/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-12-11/', 's3://madmex-fao/eodata/oli_tirs/29047/2016/2016-12-27/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'029047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]`

`print exec_adapter('LandcoverMultitemporalWorkflow', queued=True, waiting=False, kwarg=kwargs)`

Land cover RapidEye production

Land Cover RapidEye Production

The tile based land cover adapter for RapidEye scenes

The tile based RapidEye land cover process is implemented in the adapter `LandcoverRapidEyeSceneWorkflow` which is defined by the following input parameters:

`working_dir` [optional]

This targets to a local directory where the processing results shall be stored. If this varaiable is not defined the workind directory will be created automatically.

`granule`

This variable defines the RapidEye tile (e.g. `1348417`).

`scene_folder_list`

This variable sets the locations of the `SceneResources` locations, either local or on AWS S3, as list or comma separated string. E.g. `['s3://eodata/rapideye/1348417/2016/2016-02-17/', 's3://eodata/rapideye/1348417/2016/2016-10-11/']`

`land_mask_shape`

This variable sets the location, either locally or on AWS S3, of a polygon Shapefile to be used for defining a validity mask for the classifications. This means, all image pixels outside the given polygons will be masked out in processing. This is especially helpful and important for masking out ocean and/or to run classifications for given regions or states only. Example: `s3://madmex-fao/mexico_admin/Entidades_2013.shp`

`aux_image_list`

This variable sets the locations, either locally or on AWS S3, as list or comma separated string, of additional images to be used in feature extraction and classification. This images must not share same projections as the satellite images but shall of course overlap the scenes/tile/granule geographic area. Example: `['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif]`

`training_url`

This variable sets the location, either locally or on AWS S3, of the training image to be used in classification. he training dataset shall overlap the scenes/tile/granule geographic area and must consist of one band with pixel values representing the respective class. No data must be valued 0. Example: `s3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif`

`segmentation_method`

This variable defines the segmentation method to be used. This can bei either `bis` for the BerkeleyImageSegmentation or one of `SLIC, SLICO, LSC or SEEDS` for the OpenCSV segmentation. Note: `bis` requires license.

`segmentation_params`

This variable defines the parameters of the defined segmentation method, as list or comma separated string. For the supported segmentation methids these are:

- BIS
  - `[segmentation_threshold, segmentation_shape, segmentation_compactness]`
  - Example: `'40,0.2,0.8'` or `[40,0.2,0.8]`
  - whereby
    - segmentation_threshold [BIS threshold, the number of region merging iterations]
    - segmentation_shape [BIS shape parameters ranging from 0..1]
    - segmentation_compactness [BIS compactness parameters ranging from 0..1]
- SLIC, SLICO
  - `[iteration, region_size]`
  - Example: `'10,5'` or `[10,5]`
  - whereby
    - iteration [number of merge iterations, default 10]
    - region_size [approximate object size (diameter, edge length) in pixels]

`sampling_params` [optional]

This optional parameter sets up sampling option for the training dataset, as list or comma separated string. This will perform random stratified class weighted sampling of the training dataset to be used for classification. The parameters define percentage of samples per class, maximum naumber of samples per class, size of the sample chip in pixels, number of iterations e.g. `0.05,1000,30,2`. If iterations higher 1 is given, multiple classifications will be performed on multiple different sampled training datasets.

`training_percentage`

This variable defines in values of percentage [0.1 .. 99.9] how many samples shall be used in classificator training. Those are selected randomly to train the classifier, while the remaining samples will be used for classification evaluation. Example: `60` will use randomly selected 60% of all samples for training while the remaining 40% are used for classification validation.

`training_overlay_percentage`

This variable defines percentage of training object to segmentation object overlay in ranges between [0..1]. E.g. `0.7` means 70 percent of an image object must be covered by unique class values from the training dataset. If this is not the case, the image object will not be labelled with a class and will not be included as sample for classification training. This shall ensure the inclusion of the most clean image objects in classification training. Note: For training datsets that represent only small samples on the ground, which might be much smaller then the average object size from image segmentation this value must be very low.

`boosting`

This variable defines whether to use a boosted classifier or not. `0` means no boosting. `1` means boosting. If boosting is enabled by `1` 10 classifiers will be trained and their decision trees are combined to a final decision tree model.

`outlier` This variable defines whether to use sample outlier elimination or not. `0` means no elimination. `1` means do outlier elimination. If outlier elimination is enabled the class based iterative histogram trimming is applied in order to remove outlier samples for any class that do not fit in the class distribution. The outlier elimination is useful for very large training datasets, e.g. full coverage maps, to conserve only most clean training samples.

`outlier_init` [optional]

If outlier elimination is enabled this variable can be set to define the percentage of initial removal of unlikely samples per class. For example `.2` will remove the 20% most unlikely incoming samples for any class and will then continue the iterative histogram trimming and outlier removal for the remaining samples.

`cloud_dilation` [optional] An integer value can be provided to apply a dilation to the on-the-fly processed cloud and cloud shadow mask. This value must be in pixels.

`aws_s3_bucket` [optional]

This optional variable defines the AWS S3 bucket where the final results shall be stored. For example `madmex-fao` indicates the madmex-fao bucket at `s3://madmex-fao/`.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`aws_s3_key` [optional]

This optional variable defines the AWS S3 key where the final results shall be stored. For example `projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047` together with bucket defined as `madmex-fao` indicates `s3://madmex-fao/projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047` as final storage destination.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`overwrite` [optional]

This optional boolean variable can be set to `True` or `False`. This is especially helpful during development and debugging. If overwriting is disabled the most time consuming calculations like band stacking or image segmentation will not be executed again, if those intermediate results are already available in the defined `working_dir`. Default value is `True`.

`clean_up` [optional]

This optional boolean variable can be set to `True` or `False` and defines whether to remove the `working_dir` after process has finished. Default value is `True`.

Python execution

```
from adapters.workflow.landcover.landcover_rapideyescene import LandcoverRapidEyeSceneWorkflow
    landCoverProcess = LandcoverRapidEyeSceneWorkflow()
    landCoverProcess.setInput('scene_folder_list', ['s3://eodata/rapideye/1348417/2016/2016-02-17/', 's3://eodata/rapideye/1348417/2016/2016-10-11/'])
    landCoverProcess.setInput('land_mask_shape', 's3://mm-fao/aux/vectormask/mgm/areas_geoestadisticas_estatales.shp')
    landCoverProcess.setInput('training_url', 's3://mm-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif')
    landCoverProcess.setInput('granule', granule)
#     landCoverProcess.setInput('segmentation_method', 'bis')
#     landCoverProcess.setInput('segmentation_params', [50,0.2,0.8])
    landCoverProcess.setInput('segmentation_method', 'slic')
    landCoverProcess.setInput('segmentation_params', [70,100])
    landCoverProcess.setInput('training_overlay_percentage', 0.7)
    landCoverProcess.setInput('training_percentage', 50)
    landCoverProcess.setInput('cloud_dilation', 0)
    landCoverProcess.setInput('boosting', 1)
    landCoverProcess.setInput('sampling_params', [0.15,1000,30,3])
    landCoverProcess.setInput('outlier', 1)
    landCoverProcess.setInput('overwrite', False)
    landCoverProcess.setInput('aws_s3_bucket', 'mm-fao')
    landCoverProcess.setInput('aws_s3_key', 'test/landcover_rapideye/1348417')
    landCoverProcess.setInput('clean_up',False)
landCoverProcess.run()
landover_results = landCoverProcess.output_variables["result"].getVariable()
```

Job execution

`kwarg = [['scene_folder_list',['s3://eodata/rapideye/1348417/2016/2016-02-17/', 's3://eodata/rapideye/1348417/2016/2016-10-11/']], ['land_mask_shape','s3://mm-fao/aux/vectormask/mgm/areas_geoestadisticas_estatales.shp'], ['training_url','s3://mm-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'], ['granule',granule], ['segmentation_method','slic'], ['segmentation_params',[50,10]], ['training_overlay_percentage',0.7], ['training_percentage',50], ['boosting',1], ['cloud_dilation',10], ['sampling_params',[0.15,1000,30,3]], ['outlier',1], ['aws_s3_bucket','mm-fao'], ['aws_s3_key','test/landcover_rapideye/1348417'], ["clean_up", True], ["register_result", True],]`

`print exec_adapter('LandcoverRapidEyeSceneWorkflow', queued=True, waiting=False, kwarg=kwargs)`

Land cover mosaic production

Raster based land cover mosaicing has been implemented to reduce effects from tile based classifications in overlapping regions of neighbouring tiles.

Landcover mosaicing is implemented through the `LandcoverMosaicWorkflow` and can be executed on either Landsat, Sentinel or RapidEye land over classifications. It obtains the following input variables:

`lccfiles`

This variable sets the locations of the tile based landcover classification result image locations, either local or on AWS S3, as list or comma separated string. E.g. `['/Volumes/gdrive/work/conafor/landcover/2017/028045/result/landcover_028045_classify_mx_mapadereferencia2015_25m_warp_.tif', '/Volumes/gdrive/work/conafor/landcover/2017/028046/result/landcover_028046_classify_mx_mapadereferencia2015_25m_warp_.tif']`

`conffiles`

This variable sets the locations of the tile based landcover classification confidence image locations, either local or on AWS S3, as list or comma separated string. E.g. `['/Volumes/gdrive/work/conafor/landcover/2017/028045/result/landcover_028045_classify_mx_mapadereferencia2015_25m_warp__confidence.tif', '/Volumes/gdrive/work/conafor/landcover/2017/028046/result/landcover_028046_classify_mx_mapadereferencia2015_25m_warp__confidence.tif']`

`tileids`

This variable sets the corresponding tile identifiers, either local or on AWS S3, as list or comma separated string. Note: lccfiles, conffiles, and tileids must have same tile ordering. E.g. `[28045, 28046]`

`out_filename`

This variable defines the name of the final mosaic image. E.g. `'Landcover_mosaic.tif'`

`gridshapefile`

This variable defines the location of a Landsat WRS2 Tile grid Shapefile, either local or on AWS S3. E.g. `s3://mm-fao/grid/sentinel_footprints_mexico_states.shp` or `s3://mm-fao/grid/rapideye_footprints_mexico_states.shp` or `s3://mm-fao/grid/landsat_footprints_mexico_states.shp`

`gridcolumn`

This variable defines the name of the column with the path/row identifier for each tile. E.g. `'pathrow'`

`resolution` [optional]

Defines the output pixel size. Default is 30.

`proj4` [optional]

Defines the output projection as proj4 string. Default is '+proj=lcc +lat_1=17.5 +lat_2=29.5 +lat_0=12.0 +lon_0=-102 +x_0=2500000.0 +ellps=WGS84 +datum=WGS84 +units=m +no_defs'.

`masking` [optional]

Whether or not to apply the shapfile granule geometry as mask. Default is true.

`working_dir` [optional]

This targets to a local directory where the processing results shall be stored. If this varaiable is not defined the workind directory will be created automatically.

`aws_s3_bucket` [optional]

This optional variable defines the AWS S3 bucket where the final results shall be stored. For example `madmex-fao` indicates the madmex-fao bucket at `s3://madmex-fao/`.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`aws_s3_key` [optional]

This optional variable defines the AWS S3 key where the final results shall be stored. For example `projects/Cobertura_de_suelo/Jalisco_Landsat_2016/mosaic` together with bucket defined as `madmex-fao` indicates `s3://madmex-fao/projects/Cobertura_de_suelo/Jalisco_Landsat_2016/mosaic` as final storage destination.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`clean_up` [optional]

This optional boolean variable can be set to `True` or `False` and defines whether to remove the `working_dir` after process has finished. Default value is `True`.

Python execution:

```
a = LandcoverMosaicWorkflow()
a.setInput('lccfiles', ['/Volumes/gdrive/work/conafor/landcover/2017/028045/result/landcover_028045_classify_mx_mapadereferencia2015_25m_warp_.tif', '/Volumes/gdrive/work
/conafor/landcover/2017/028046/result/landcover_028046_classify_mx_mapadereferencia2015_25m_warp_.tif'])
a.setInput('conffiles', ['/Volumes/gdrive/work/conafor/landcover/2017/028045/result/landcover_028045_classify_mx_mapadereferencia2015_25m_warp__confidence.tif', '/Volumes
/gdrive/work/conafor/landcover/2017/028046/result/landcover_028046_classify_mx_mapadereferencia2015_25m_warp__confidence.tif'])
a.setInput('tileids', ['028045', '028046'])
a.setInput('out_filename', 'Landcover_mosaic.tif')
a.setInput('gridshapefile', 's3://madmex-fao/grid/landsat_footprints_mexico_states.shp')
a.setInput('gridcolumn', 'CODE')
a.setInput('aws_s3_bucket', 'madmex-fao')
a.setInput('aws_s3_key', 'projects/Cobertura_de_suelo/Jalisco_Landsat_2016/mosaic/')
a.setInput('clean_up', True)
a.run()
result = a.output_variables["result"].getVariable()
```

Job execution

`kwargs = [('lccfiles', ['/Volumes/gdrive/work/conafor/landcover/2017/028045/result/landcover_028045_classify_mx_mapadereferencia2015_25m_warp_.tif', '/Volumes/gdrive/work/conafor/landcover/2017/028046/result/landcover_028046_classify_mx_mapadereferencia2015_25m_warp_.tif']),('conffiles', ['/Volumes/gdrive/work/conafor/landcover/2017/028045/result/landcover_028045_classify_mx_mapadereferencia2015_25m_warp__confidence.tif', '/Volumes/gdrive/work/conafor/landcover/2017/028046/result/landcover_028046_classify_mx_mapadereferencia2015_25m_warp__confidence.tif']),a.setInput('tileids', [28045, 28046]),('out_filename', 'Landcover_mosaic.tif'),('gridshapefile', 's3://madmex-fao/grid/landsat_footprints_mexico.shp'),('gridcolumn', 'pathrow'), ('aws_s3_bucket', 'madmex-fao'),('aws_s3_key', 'projects/Cobertura_de_suelo/Jalisco_Landsat_2016/mosaic/'),('clean_up', True)]`

`print exec_adapter('LandcoverMosaicWorkflow', queued=True, waiting=False, kwarg=kwargs)`

Land cover production example

Landcover 2000 for the states of Campeche, Chiapas, Jalisco, Quintana Roo and Yucatan

Job generation

following listing provides input kwargs for job execution

```
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/18045/2000/2000-01-01/', 's3://madmex-fao/eodata/tm/18045/2000/2000-02-18/', 's3://madmex-fao/eodata/etm+/18045/2000/20
00-02-26/', 's3://madmex-fao/eodata/tm/18045/2000/2000-03-21/', 's3://madmex-fao/eodata/tm/18045/2000/2000-04-22/', 's3://madmex-fao/eodata/etm+/18045/2000/2000-04-30/',
's3://madmex-fao/eodata/tm/18045/2000/2000-05-24/', 's3://madmex-fao/eodata/tm/18045/2000/2000-07-11/', 's3://madmex-fao/eodata/etm+/18045/2000/2000-07-19/', 's3://madmex
-fao/eodata/tm/18045/2000/2000-07-27/', 's3://madmex-fao/eodata/tm/18045/2000/2000-08-12/', 's3://madmex-fao/eodata/tm/18045/2000/2000-08-28/', 's3://madmex-fao/eodata/et
m+/18045/2000/2000-09-05/', 's3://madmex-fao/eodata/tm/18045/2000/2000-09-29/', 's3://madmex-fao/eodata/tm/18045/2000/2000-10-15/', 's3://madmex-fao/eodata/etm+/18045/200
0/2000-12-10/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_
lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), (
'granule', u'018045'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_per
centage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite',
False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/018045'), ('features', ['NDVI']), ('metrics_features', 'minimum,ma
ximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/etm+/18046/2000/2000-01-25/', 's3://madmex-fao/eodata/tm/18046/2000/2000-02-18/', 's3://madmex-fao/eodata/etm+/18046/2000/
2000-03-29/', 's3://madmex-fao/eodata/tm/18046/2000/2000-04-22/', 's3://madmex-fao/eodata/etm+/18046/2000/2000-04-30/', 's3://madmex-fao/eodata/tm/18046/2000/2000-05-24/'
, 's3://madmex-fao/eodata/etm+/18046/2000/2000-07-19/', 's3://madmex-fao/eodata/tm/18046/2000/2000-07-27/', 's3://madmex-fao/eodata/etm+/18046/2000/2000-09-05/', 's3://ma
dmex-fao/eodata/tm/18046/2000/2000-09-29/', 's3://madmex-fao/eodata/etm+/18046/2000/2000-10-07/', 's3://madmex-fao/eodata/tm/18046/2000/2000-10-31/', 's3://madmex-fao/eod
ata/etm+/18046/2000/2000-12-10/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/
CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_l
cc_slope.tif']), ('granule', u'018046'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.
7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier',
1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/018046'), ('features', ['NDVI']), ('metrics_feat
ures', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/etm+/18047/2000/2000-01-25/', 's3://madmex-fao/eodata/tm/18047/2000/2000-03-21/', 's3://madmex-fao/eodata/etm+/18047/2000/
2000-03-29/', 's3://madmex-fao/eodata/tm/18047/2000/2000-04-22/', 's3://madmex-fao/eodata/etm+/18047/2000/2000-04-30/', 's3://madmex-fao/eodata/tm/18047/2000/2000-05-24/'
, 's3://madmex-fao/eodata/tm/18047/2000/2000-07-11/', 's3://madmex-fao/eodata/etm+/18047/2000/2000-07-19/', 's3://madmex-fao/eodata/tm/18047/2000/2000-07-27/', 's3://madm
ex-fao/eodata/etm+/18047/2000/2000-08-04/', 's3://madmex-fao/eodata/etm+/18047/2000/2000-09-05/', 's3://madmex-fao/eodata/etm+/18047/2000/2000-10-07/', 's3://madmex-fao/e
odata/tm/18047/2000/2000-10-31/', 's3://madmex-fao/eodata/etm+/18047/2000/2000-11-08/', 's3://madmex-fao/eodata/tm/18047/2000/2000-11-16/', 's3://madmex-fao/eodata/etm+/1
8047/2000/2000-12-10/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15
m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.t
if']), ('granule', u'018047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('trai
ning_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('ove
rwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/018047'), ('features', ['NDVI']), ('metrics_features', 'm
inimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/19045/2000/2000-01-08/', 's3://madmex-fao/eodata/tm/19045/2000/2000-01-24/', 's3://madmex-fao/eodata/tm/19045/2000/2000
-02-09/', 's3://madmex-fao/eodata/etm+/19045/2000/2000-03-20/', 's3://madmex-fao/eodata/tm/19045/2000/2000-03-28/', 's3://madmex-fao/eodata/etm+/19045/2000/2000-04-21/',
's3://madmex-fao/eodata/tm/19045/2000/2000-04-29/', 's3://madmex-fao/eodata/tm/19045/2000/2000-05-31/', 's3://madmex-fao/eodata/etm+/19045/2000/2000-06-24/', 's3://madmex
-fao/eodata/etm+/19045/2000/2000-07-10/', 's3://madmex-fao/eodata/tm/19045/2000/2000-07-18/', 's3://madmex-fao/eodata/tm/19045/2000/2000-08-03/', 's3://madmex-fao/eodata/
tm/19045/2000/2000-08-19/', 's3://madmex-fao/eodata/tm/19045/2000/2000-09-04/', 's3://madmex-fao/eodata/tm/19045/2000/2000-12-09/']), ('land_mask_shape', 's3://madmex-fao
/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/de
m/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'019045'), ('training_url', 's3://madmex-f
ao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segme
ntation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key
', 'projects/Cobertura_de_suelo/EAT_landcover_2000/019045'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/19046/2000/2000-01-08/', 's3://madmex-fao/eodata/tm/19046/2000/2000-01-24/', 's3://madmex-fao/eodata/tm/19046/2000/2000
-02-09/', 's3://madmex-fao/eodata/tm/19046/2000/2000-03-28/', 's3://madmex-fao/eodata/etm+/19046/2000/2000-04-21/', 's3://madmex-fao/eodata/tm/19046/2000/2000-12-09/']),
('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3:
//madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'01904
6'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('s
egmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_
bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/019046'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,qua
ntiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/19047/2000/2000-02-09/', 's3://madmex-fao/eodata/etm+/19047/2000/2000-02-17/', 's3://madmex-fao/eodata/tm/19047/2000/20
00-03-28/', 's3://madmex-fao/eodata/tm/19047/2000/2000-12-09/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex
-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcas
es_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'019047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('tr
aining_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average')
, ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/019047'), ('featu
res', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/19048/2000/2000-01-24/', 's3://madmex-fao/eodata/tm/19048/2000/2000-03-28/', 's3://madmex-fao/eodata/etm+/19048/2000/20
00-07-26/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_
dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('gra
nule', u'019048'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percent
age', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', Fal
se), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/019048'), ('features', ['NDVI']), ('metrics_features', 'minimum,maxim
um,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/etm+/20045/2000/2000-01-07/', 's3://madmex-fao/eodata/tm/20045/2000/2000-02-16/', 's3://madmex-fao/eodata/etm+/20045/2000/
2000-02-24/', 's3://madmex-fao/eodata/tm/20045/2000/2000-03-19/', 's3://madmex-fao/eodata/tm/20045/2000/2000-04-20/', 's3://madmex-fao/eodata/etm+/20045/2000/2000-04-28/'
, 's3://madmex-fao/eodata/tm/20045/2000/2000-06-23/', 's3://madmex-fao/eodata/etm+/20045/2000/2000-07-01/', 's3://madmex-fao/eodata/tm/20045/2000/2000-07-09/', 's3://madm
ex-fao/eodata/tm/20045/2000/2000-08-26/', 's3://madmex-fao/eodata/tm/20045/2000/2000-09-27/', 's3://madmex-fao/eodata/tm/20045/2000/2000-10-29/', 's3://madmex-fao/eodata/
etm+/20045/2000/2000-11-06/', 's3://madmex-fao/eodata/tm/20045/2000/2000-12-16/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_li
st', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-
fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'020045'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia20
15_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_m
```

```
15_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_m
etric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000
/020045'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/20046/2000/2000-02-16/', 's3://madmex-fao/eodata/tm/20046/2000/2000-03-19/', 's3://madmex-fao/eodata/tm/20046/2000/2000
-04-20/', 's3://madmex-fao/eodata/etm+/20046/2000/2000-04-28/', 's3://madmex-fao/eodata/etm+/20046/2000/2000-07-01/', 's3://madmex-fao/eodata/tm/20046/2000/2000-10-29/',
's3://madmex-fao/eodata/etm+/20046/2000/2000-11-06/', 's3://madmex-fao/eodata/tm/20046/2000/2000-12-16/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_20
13.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_as
pect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'020046'), ('training_url', 's3://madmex-fao/training/mapadereferenc
ia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2,
0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_s
uelo/EAT_landcover_2000/020046'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/20047/2000/2000-01-31/', 's3://madmex-fao/eodata/etm+/20047/2000/2000-03-27/', 's3://madmex-fao/eodata/tm/20047/2000/20
00-04-20/', 's3://madmex-fao/eodata/etm+/20047/2000/2000-04-28/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madm
ex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testc
ases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'020047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('
training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average
'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/020047'), ('fea
tures', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/20048/2000/2000-01-31/', 's3://madmex-fao/eodata/etm+/20048/2000/2000-03-27/', 's3://madmex-fao/eodata/tm/20048/2000/20
00-04-20/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_
dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('gra
nule', u'020048'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percent
age', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', Fal
se), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/020048'), ('features', ['NDVI']), ('metrics_features', 'minimum,maxim
um,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/etm+/20049/2000/2000-01-23/', 's3://madmex-fao/eodata/tm/20049/2000/2000-04-04/', 's3://madmex-fao/eodata/tm/20049/2000/20
00-04-20/', 's3://madmex-fao/eodata/tm/20049/2000/2000-06-23/', 's3://madmex-fao/eodata/tm/20049/2000/2000-09-11/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/En
tidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_in
egi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'020049'), ('training_url', 's3://madmex-fao/training/mapa
dereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params',
[50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobe
rtura_de_suelo/EAT_landcover_2000/020049'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/21045/2000/2000-01-22/', 's3://madmex-fao/eodata/etm+/21045/2000/2000-02-15/', 's3://madmex-fao/eodata/tm/21045/2000/20
00-02-23/', 's3://madmex-fao/eodata/tm/21045/2000/2000-03-10/', 's3://madmex-fao/eodata/tm/21045/2000/2000-03-26/', 's3://madmex-fao/eodata/etm+/21045/2000/2000-04-19/',
's3://madmex-fao/eodata/tm/21045/2000/2000-05-13/', 's3://madmex-fao/eodata/tm/21045/2000/2000-06-14/', 's3://madmex-fao/eodata/etm+/21045/2000/2000-06-22/', 's3://madmex
-fao/eodata/tm/21045/2000/2000-06-30/', 's3://madmex-fao/eodata/etm+/21045/2000/2000-07-08/', 's3://madmex-fao/eodata/tm/21045/2000/2000-07-16/', 's3://madmex-fao/eodata/
tm/21045/2000/2000-08-01/', 's3://madmex-fao/eodata/tm/21045/2000/2000-08-17/', 's3://madmex-fao/eodata/etm+/21045/2000/2000-09-10/', 's3://madmex-fao/eodata/etm+/21045/2
000/2000-10-28/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_ineg
i_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']),
('granule', u'021045'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_pe
rcentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite'
, False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/021045'), ('features', ['NDVI']), ('metrics_features', 'minimum,
maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/21046/2000/2000-01-22/', 's3://madmex-fao/eodata/etm+/21046/2000/2000-02-15/', 's3://madmex-fao/eodata/tm/21046/2000/20
00-02-23/', 's3://madmex-fao/eodata/tm/21046/2000/2000-03-10/', 's3://madmex-fao/eodata/tm/21046/2000/2000-03-26/', 's3://madmex-fao/eodata/etm+/21046/2000/2000-04-03/',
's3://madmex-fao/eodata/tm/21046/2000/2000-04-11/', 's3://madmex-fao/eodata/etm+/21046/2000/2000-04-19/', 's3://madmex-fao/eodata/tm/21046/2000/2000-04-27/', 's3://madmex
-fao/eodata/tm/21046/2000/2000-06-14/', 's3://madmex-fao/eodata/etm+/21046/2000/2000-06-22/', 's3://madmex-fao/eodata/tm/21046/2000/2000-06-30/', 's3://madmex-fao/eodata/
etm+/21046/2000/2000-07-08/', 's3://madmex-fao/eodata/tm/21046/2000/2000-07-16/', 's3://madmex-fao/eodata/tm/21046/2000/2000-08-01/', 's3://madmex-fao/eodata/tm/21046/200
0/2000-08-17/', 's3://madmex-fao/eodata/etm+/21046/2000/2000-09-10/', 's3://madmex-fao/eodata/etm+/21046/2000/2000-10-28/']), ('land_mask_shape', 's3://madmex-fao/mexico_
admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0
_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'021046'), ('training_url', 's3://madmex-fao/train
ing/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_
params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'proj
ects/Cobertura_de_suelo/EAT_landcover_2000/021046'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/etm+/21047/2000/2000-01-14/', 's3://madmex-fao/eodata/tm/21047/2000/2000-01-22/', 's3://madmex-fao/eodata/etm+/21047/2000/
2000-02-15/', 's3://madmex-fao/eodata/tm/21047/2000/2000-02-23/', 's3://madmex-fao/eodata/tm/21047/2000/2000-03-10/', 's3://madmex-fao/eodata/etm+/21047/2000/2000-04-03/'
, 's3://madmex-fao/eodata/etm+/21047/2000/2000-04-19/', 's3://madmex-fao/eodata/tm/21047/2000/2000-04-27/', 's3://madmex-fao/eodata/tm/21047/2000/2000-05-13/', 's3://madm
ex-fao/eodata/etm+/21047/2000/2000-06-22/', 's3://madmex-fao/eodata/tm/21047/2000/2000-06-30/', 's3://madmex-fao/eodata/tm/21047/2000/2000-07-16/', 's3://madmex-fao/eodat
a/etm+/21047/2000/2000-09-10/', 's3://madmex-fao/eodata/etm+/21047/2000/2000-10-28/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_imag
e_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://mad
mex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'021047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferenc
ia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentati
on_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_
2000/021047'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/21048/2000/2000-01-22/', 's3://madmex-fao/eodata/tm/21048/2000/2000-02-23/', 's3://madmex-fao/eodata/tm/21048/2000/2000
-03-10/', 's3://madmex-fao/eodata/etm+/21048/2000/2000-04-03/', 's3://madmex-fao/eodata/tm/21048/2000/2000-04-11/', 's3://madmex-fao/eodata/etm+/21048/2000/2000-04-19/',
's3://madmex-fao/eodata/tm/21048/2000/2000-05-13/', 's3://madmex-fao/eodata/etm+/21048/2000/2000-06-22/', 's3://madmex-fao/eodata/tm/21048/2000/2000-07-16/', 's3://madmex
-fao/eodata/etm+/21048/2000/2000-11-13/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_
aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m
_inegi_lcc_slope.tif']), ('granule', u'021048'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percent
age', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('o
utlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/021048'), ('features', ['NDVI']), ('met
rics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/21049/2000/2000-01-22/', 's3://madmex-fao/eodata/tm/21049/2000/2000-02-23/', 's3://madmex-fao/eodata/etm+/21049/2000/20
00-04-03/', 's3://madmex-fao/eodata/tm/21049/2000/2000-04-11/', 's3://madmex-fao/eodata/etm+/21049/2000/2000-04-19/', 's3://madmex-fao/eodata/etm+/21049/2000/2000-06-22/'
, 's3://madmex-fao/eodata/tm/21049/2000/2000-07-16/', 's3://madmex-fao/eodata/etm+/21049/2000/2000-11-13/', 's3://madmex-fao/eodata/tm/21049/2000/2000-12-07/']), ('land_m
ask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex
-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'021049'), ('t
raining_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentat
ion_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket',
'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/021049'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25'
)]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/21050/2000/2000-01-22/', 's3://madmex-fao/eodata/tm/21050/2000/2000-02-07/', 's3://madmex-fao/eodata/etm+/21050/2000/20
00-02-15/', 's3://madmex-fao/eodata/tm/21050/2000/2000-02-23/', 's3://madmex-fao/eodata/etm+/21050/2000/2000-04-03/', 's3://madmex-fao/eodata/tm/21050/2000/2000-04-11/',
's3://madmex-fao/eodata/etm+/21050/2000/2000-06-22/', 's3://madmex-fao/eodata/etm+/21050/2000/2000-10-28/', 's3://madmex-fao/eodata/tm/21050/2000/2000-11-21/', 's3://madm
ex-fao/eodata/tm/21050/2000/2000-12-07/', 's3://madmex-fao/eodata/etm+/21050/2000/2000-12-31/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'),
('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif',
's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'021050'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapad
ereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('s
egmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_l
andcover_2000/021050'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/22045/2000/2000-01-13/', 's3://madmex-fao/eodata/tm/22045/2000/2000-02-14/', 's3://madmex-fao/eodata/tm/22045/2000/2000
-03-01/', 's3://madmex-fao/eodata/tm/22045/2000/2000-03-17/', 's3://madmex-fao/eodata/tm/22045/2000/2000-04-18/', 's3://madmex-fao/eodata/tm/22045/2000/2000-05-04/', 's3:
//madmex-fao/eodata/etm+/22045/2000/2000-05-12/', 's3://madmex-fao/eodata/tm/22045/2000/2000-06-21/', 's3://madmex-fao/eodata/tm/22045/2000/2000-07-07/', 's3://madmex-fao
/eodata/etm+/22045/2000/2000-07-15/', 's3://madmex-fao/eodata/tm/22045/2000/2000-07-23/', 's3://madmex-fao/eodata/tm/22045/2000/2000-08-24/', 's3://madmex-fao/eodata/tm/2
2045/2000/2000-09-25/', 's3://madmex-fao/eodata/tm/22045/2000/2000-10-11/', 's3://madmex-fao/eodata/etm+/22045/2000/2000-10-19/', 's3://madmex-fao/eodata/tm/22045/2000/20
00-10-27/', 's3://madmex-fao/eodata/tm/22045/2000/2000-12-14/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex
-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcas
es_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'022045'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('tr
aining_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average')
, ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/022045'), ('featu
res', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/22047/2000/2000-01-13/', 's3://madmex-fao/eodata/etm+/22047/2000/2000-01-21/', 's3://madmex-fao/eodata/tm/22047/2000/20
00-02-14/', 's3://madmex-fao/eodata/etm+/22047/2000/2000-02-22/', 's3://madmex-fao/eodata/tm/22047/2000/2000-03-01/', 's3://madmex-fao/eodata/tm/22047/2000/2000-04-18/',
's3://madmex-fao/eodata/etm+/22047/2000/2000-05-12/', 's3://madmex-fao/eodata/tm/22047/2000/2000-07-07/', 's3://madmex-fao/eodata/etm+/22047/2000/2000-07-15/', 's3://madm
ex-fao/eodata/tm/22047/2000/2000-09-09/', 's3://madmex-fao/eodata/tm/22047/2000/2000-10-27/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('
aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', '
s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'022047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapade
referencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('se
gmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_la
ndcover_2000/022047'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/22048/2000/2000-02-14/', 's3://madmex-fao/eodata/tm/22048/2000/2000-03-01/', 's3://madmex-fao/eodata/etm+/22048/2000/20
00-03-25/', 's3://madmex-fao/eodata/etm+/22048/2000/2000-05-12/', 's3://madmex-fao/eodata/etm+/22048/2000/2000-07-15/']), ('land_mask_shape', 's3://madmex-fao/mexico_admi
n/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15
m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'022048'), ('training_url', 's3://madmex-fao/training/
mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_para
ms', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects
/Cobertura_de_suelo/EAT_landcover_2000/022048'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/22049/2000/2000-01-13/', 's3://madmex-fao/eodata/tm/22049/2000/2000-02-14/', 's3://madmex-fao/eodata/tm/22049/2000/2000
-03-01/', 's3://madmex-fao/eodata/etm+/22049/2000/2000-03-25/', 's3://madmex-fao/eodata/etm+/22049/2000/2000-07-15/', 's3://madmex-fao/eodata/etm+/22049/2000/2000-08-16/'
, 's3://madmex-fao/eodata/etm+/22049/2000/2000-10-19/', 's3://madmex-fao/eodata/etm+/22049/2000/2000-12-06/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidade
s_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lc
c_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'022049'), ('training_url', 's3://madmex-fao/training/mapaderefe
rencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50,
0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura
_de_suelo/EAT_landcover_2000/022049'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/28045/2000/2000-01-07/', 's3://madmex-fao/eodata/tm/28045/2000/2000-01-23/', 's3://madmex-fao/eodata/etm+/28045/2000/20
00-01-31/', 's3://madmex-fao/eodata/tm/28045/2000/2000-02-08/', 's3://madmex-fao/eodata/tm/28045/2000/2000-03-11/', 's3://madmex-fao/eodata/etm+/28045/2000/2000-03-19/',
's3://madmex-fao/eodata/tm/28045/2000/2000-04-12/', 's3://madmex-fao/eodata/tm/28045/2000/2000-04-28/', 's3://madmex-fao/eodata/etm+/28045/2000/2000-05-06/', 's3://madmex
-fao/eodata/tm/28045/2000/2000-05-14/', 's3://madmex-fao/eodata/tm/28045/2000/2000-07-17/', 's3://madmex-fao/eodata/tm/28045/2000/2000-08-02/', 's3://madmex-fao/eodata/tm
/28045/2000/2000-08-18/', 's3://madmex-fao/eodata/etm+/28045/2000/2000-08-26/', 's3://madmex-fao/eodata/etm+/28045/2000/2000-09-11/', 's3://madmex-fao/eodata/etm+/28045/2
000/2000-10-29/', 's3://madmex-fao/eodata/tm/28045/2000/2000-12-08/', 's3://madmex-fao/eodata/etm+/28045/2000/2000-12-16/', 's3://madmex-fao/eodata/tm/28045/2000/2000-12-
24/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.ti
f', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule',
u'028045'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 5
0), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('
aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/028045'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,aver
age,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/28046/2000/2000-01-07/', 's3://madmex-fao/eodata/tm/28046/2000/2000-01-23/', 's3://madmex-fao/eodata/etm+/28046/2000/20
00-01-31/', 's3://madmex-fao/eodata/tm/28046/2000/2000-02-08/', 's3://madmex-fao/eodata/tm/28046/2000/2000-03-11/', 's3://madmex-fao/eodata/etm+/28046/2000/2000-04-04/',
's3://madmex-fao/eodata/tm/28046/2000/2000-04-12/', 's3://madmex-fao/eodata/etm+/28046/2000/2000-04-20/', 's3://madmex-fao/eodata/tm/28046/2000/2000-04-28/', 's3://madmex
-fao/eodata/etm+/28046/2000/2000-05-06/', 's3://madmex-fao/eodata/etm+/28046/2000/2000-07-09/', 's3://madmex-fao/eodata/tm/28046/2000/2000-08-18/', 's3://madmex-fao/eodat
a/etm+/28046/2000/2000-09-11/', 's3://madmex-fao/eodata/etm+/28046/2000/2000-10-29/', 's3://madmex-fao/eodata/etm+/28046/2000/2000-11-30/', 's3://madmex-fao/eodata/tm/280
46/2000/2000-12-08/', 's3://madmex-fao/eodata/tm/28046/2000/2000-12-24/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s
3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madm
ex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'028046'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.t
if'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric',
'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/028046'
), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/28047/2000/2000-01-07/', 's3://madmex-fao/eodata/tm/28047/2000/2000-01-23/', 's3://madmex-fao/eodata/etm+/28047/2000/20
00-01-31/', 's3://madmex-fao/eodata/tm/28047/2000/2000-02-08/', 's3://madmex-fao/eodata/etm+/28047/2000/2000-04-04/', 's3://madmex-fao/eodata/tm/28047/2000/2000-04-12/',
's3://madmex-fao/eodata/tm/28047/2000/2000-04-28/', 's3://madmex-fao/eodata/etm+/28047/2000/2000-05-06/', 's3://madmex-fao/eodata/tm/28047/2000/2000-05-14/', 's3://madmex
-fao/eodata/tm/28047/2000/2000-08-18/', 's3://madmex-fao/eodata/etm+/28047/2000/2000-08-26/', 's3://madmex-fao/eodata/tm/28047/2000/2000-11-22/', 's3://madmex-fao/eodata/
etm+/28047/2000/2000-11-30/', 's3://madmex-fao/eodata/tm/28047/2000/2000-12-08/', 's3://madmex-fao/eodata/tm/28047/2000/2000-12-24/']), ('land_mask_shape', 's3://madmex-f
ao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/
dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'028047'), ('training_url', 's3://madmex
-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('seg
mentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_k
ey', 'projects/Cobertura_de_suelo/EAT_landcover_2000/028047'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/etm+/29044/2000/2000-01-22/', 's3://madmex-fao/eodata/tm/29044/2000/2000-01-30/', 's3://madmex-fao/eodata/etm+/29044/2000/
2000-02-23/', 's3://madmex-fao/eodata/etm+/29044/2000/2000-03-10/', 's3://madmex-fao/eodata/tm/29044/2000/2000-04-03/', 's3://madmex-fao/eodata/tm/29044/2000/2000-04-19/'
, 's3://madmex-fao/eodata/etm+/29044/2000/2000-06-30/', 's3://madmex-fao/eodata/tm/29044/2000/2000-07-08/', 's3://madmex-fao/eodata/tm/29044/2000/2000-07-24/', 's3://madm
ex-fao/eodata/tm/29044/2000/2000-08-25/', 's3://madmex-fao/eodata/tm/29044/2000/2000-09-10/', 's3://madmex-fao/eodata/etm+/29044/2000/2000-09-18/', 's3://madmex-fao/eodat
a/tm/29044/2000/2000-10-12/', 's3://madmex-fao/eodata/tm/29044/2000/2000-11-29/', 's3://madmex-fao/eodata/etm+/29044/2000/2000-12-07/', 's3://madmex-fao/eodata/tm/29044/2
000/2000-12-15/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_ineg
i_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']),
('granule', u'029044'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_pe
rcentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite'
, False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/029044'), ('features', ['NDVI']), ('metrics_features', 'minimum,
maximum,average,quantiles25')]
```

```
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/29045/2000/2000-01-14/', 's3://madmex-fao/eodata/etm+/29045/2000/2000-01-22/', 's3://madmex-fao/eodata/tm/29045/2000/20
00-01-30/', 's3://madmex-fao/eodata/etm+/29045/2000/2000-04-11/', 's3://madmex-fao/eodata/tm/29045/2000/2000-04-19/', 's3://madmex-fao/eodata/tm/29045/2000/2000-05-21/',
's3://madmex-fao/eodata/tm/29045/2000/2000-07-08/', 's3://madmex-fao/eodata/etm+/29045/2000/2000-07-16/', 's3://madmex-fao/eodata/tm/29045/2000/2000-07-24/', 's3://madmex
-fao/eodata/etm+/29045/2000/2000-08-17/', 's3://madmex-fao/eodata/tm/29045/2000/2000-08-25/', 's3://madmex-fao/eodata/tm/29045/2000/2000-09-10/', 's3://madmex-fao/eodata/
tm/29045/2000/2000-10-12/', 's3://madmex-fao/eodata/tm/29045/2000/2000-11-29/', 's3://madmex-fao/eodata/etm+/29045/2000/2000-12-07/', 's3://madmex-fao/eodata/tm/29045/200
0/2000-12-15/', 's3://madmex-fao/eodata/tm/29045/2000/2000-12-31/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://ma
dmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_tes
tcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'029045'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'),
('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'avera
ge'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/029045'), ('f
eatures', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/29046/2000/2000-01-14/', 's3://madmex-fao/eodata/etm+/29046/2000/2000-01-22/', 's3://madmex-fao/eodata/tm/29046/2000/20
00-01-30/', 's3://madmex-fao/eodata/tm/29046/2000/2000-02-15/', 's3://madmex-fao/eodata/etm+/29046/2000/2000-04-11/', 's3://madmex-fao/eodata/tm/29046/2000/2000-04-19/',
's3://madmex-fao/eodata/tm/29046/2000/2000-05-21/', 's3://madmex-fao/eodata/tm/29046/2000/2000-07-08/', 's3://madmex-fao/eodata/etm+/29046/2000/2000-07-16/', 's3://madmex
-fao/eodata/tm/29046/2000/2000-07-24/', 's3://madmex-fao/eodata/etm+/29046/2000/2000-08-17/', 's3://madmex-fao/eodata/tm/29046/2000/2000-08-25/', 's3://madmex-fao/eodata/
tm/29046/2000/2000-09-10/', 's3://madmex-fao/eodata/tm/29046/2000/2000-11-29/', 's3://madmex-fao/eodata/etm+/29046/2000/2000-12-07/', 's3://madmex-fao/eodata/tm/29046/200
0/2000-12-15/', 's3://madmex-fao/eodata/tm/29046/2000/2000-12-31/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://ma
dmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_tes
tcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'029046'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'),
('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'avera
ge'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/029046'), ('f
eatures', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/29047/2000/2000-01-14/', 's3://madmex-fao/eodata/etm+/29047/2000/2000-01-22/', 's3://madmex-fao/eodata/tm/29047/2000/20
00-01-30/', 's3://madmex-fao/eodata/tm/29047/2000/2000-02-15/', 's3://madmex-fao/eodata/etm+/29047/2000/2000-04-11/', 's3://madmex-fao/eodata/etm+/29047/2000/2000-05-13/'
, 's3://madmex-fao/eodata/etm+/29047/2000/2000-08-17/', 's3://madmex-fao/eodata/tm/29047/2000/2000-08-25/', 's3://madmex-fao/eodata/tm/29047/2000/2000-09-10/', 's3://madm
ex-fao/eodata/tm/29047/2000/2000-09-26/', 's3://madmex-fao/eodata/tm/29047/2000/2000-10-12/', 's3://madmex-fao/eodata/tm/29047/2000/2000-11-29/', 's3://madmex-fao/eodata/
etm+/29047/2000/2000-12-07/', 's3://madmex-fao/eodata/tm/29047/2000/2000-12-31/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_li
st', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-
fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'029047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia20
15_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_m
etric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000
/029047'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/30044/2000/2000-01-21/', 's3://madmex-fao/eodata/tm/30044/2000/2000-02-22/', 's3://madmex-fao/eodata/tm/30044/2000/2000
-03-09/', 's3://madmex-fao/eodata/etm+/30044/2000/2000-03-17/', 's3://madmex-fao/eodata/etm+/30044/2000/2000-04-18/', 's3://madmex-fao/eodata/tm/30044/2000/2000-05-12/',
's3://madmex-fao/eodata/tm/30044/2000/2000-06-29/', 's3://madmex-fao/eodata/etm+/30044/2000/2000-07-07/', 's3://madmex-fao/eodata/tm/30044/2000/2000-07-15/', 's3://madmex
-fao/eodata/tm/30044/2000/2000-10-03/', 's3://madmex-fao/eodata/tm/30044/2000/2000-10-19/', 's3://madmex-fao/eodata/etm+/30044/2000/2000-10-27/', 's3://madmex-fao/eodata/
tm/30044/2000/2000-12-06/', 's3://madmex-fao/eodata/etm+/30044/2000/2000-12-14/', 's3://madmex-fao/eodata/tm/30044/2000/2000-12-22/']), ('land_mask_shape', 's3://madmex-f
ao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/
dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'030044'), ('training_url', 's3://madmex
-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('seg
mentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_k
ey', 'projects/Cobertura_de_suelo/EAT_landcover_2000/030044'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/30045/2000/2000-01-21/', 's3://madmex-fao/eodata/etm+/30045/2000/2000-02-14/', 's3://madmex-fao/eodata/tm/30045/2000/20
00-02-22/', 's3://madmex-fao/eodata/tm/30045/2000/2000-03-09/', 's3://madmex-fao/eodata/etm+/30045/2000/2000-03-17/', 's3://madmex-fao/eodata/etm+/30045/2000/2000-04-18/'
, 's3://madmex-fao/eodata/tm/30045/2000/2000-04-26/', 's3://madmex-fao/eodata/tm/30045/2000/2000-05-12/', 's3://madmex-fao/eodata/etm+/30045/2000/2000-05-20/', 's3://madm
ex-fao/eodata/etm+/30045/2000/2000-07-23/', 's3://madmex-fao/eodata/tm/30045/2000/2000-10-19/', 's3://madmex-fao/eodata/etm+/30045/2000/2000-10-27/', 's3://madmex-fao/eod
ata/etm+/30045/2000/2000-12-14/', 's3://madmex-fao/eodata/tm/30045/2000/2000-12-22/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_imag
e_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://mad
mex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'030045'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferenc
ia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentati
on_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_
2000/030045'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/30046/2000/2000-01-21/', 's3://madmex-fao/eodata/tm/30046/2000/2000-02-06/', 's3://madmex-fao/eodata/tm/30046/2000/2000
-02-22/', 's3://madmex-fao/eodata/etm+/30046/2000/2000-03-17/', 's3://madmex-fao/eodata/etm+/30046/2000/2000-04-18/', 's3://madmex-fao/eodata/tm/30046/2000/2000-04-26/',
's3://madmex-fao/eodata/etm+/30046/2000/2000-05-20/', 's3://madmex-fao/eodata/tm/30046/2000/2000-07-15/', 's3://madmex-fao/eodata/etm+/30046/2000/2000-10-27/', 's3://madm
ex-fao/eodata/tm/30046/2000/2000-12-06/', 's3://madmex-fao/eodata/etm+/30046/2000/2000-12-14/', 's3://madmex-fao/eodata/tm/30046/2000/2000-12-22/']), ('land_mask_shape',
's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_
testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'030046'), ('training_url'
, 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method',
'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'
), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/030046'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/30047/2000/2000-01-21/', 's3://madmex-fao/eodata/tm/30047/2000/2000-02-06/', 's3://madmex-fao/eodata/etm+/30047/2000/20
00-02-14/', 's3://madmex-fao/eodata/tm/30047/2000/2000-02-22/', 's3://madmex-fao/eodata/etm+/30047/2000/2000-03-17/', 's3://madmex-fao/eodata/etm+/30047/2000/2000-04-18/'
, 's3://madmex-fao/eodata/etm+/30047/2000/2000-05-20/', 's3://madmex-fao/eodata/etm+/30047/2000/2000-08-08/', 's3://madmex-fao/eodata/etm+/30047/2000/2000-10-27/', 's3://
madmex-fao/eodata/tm/30047/2000/2000-12-06/', 's3://madmex-fao/eodata/etm+/30047/2000/2000-12-14/']), ('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp
'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.t
if', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'030047'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_
mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('segmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8])
, ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/
EAT_landcover_2000/030047'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,quantiles25')]
[('scene_folder_list', ['s3://madmex-fao/eodata/tm/31046/2000/2000-01-28/', 's3://madmex-fao/eodata/etm+/31046/2000/2000-03-24/', 's3://madmex-fao/eodata/tm/31046/2000/20
00-05-03/', 's3://madmex-fao/eodata/tm/31046/2000/2000-05-19/', 's3://madmex-fao/eodata/tm/31046/2000/2000-08-23/', 's3://madmex-fao/eodata/tm/31046/2000/2000-10-10/']),
('land_mask_shape', 's3://madmex-fao/mexico_admin/Entidades_2013.shp'), ('aux_image_list', ['s3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_dem.tif', 's3:
//madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_aspect.tif', 's3://madmex-fao/madmex_testcases_aux/dem/CEM3.0_R15m_inegi_lcc_slope.tif']), ('granule', u'03104
6'), ('training_url', 's3://madmex-fao/training/mapadereferencia/mx_mapadereferencia2015_25m.tif'), ('training_overlay_percentage', 0.7), ('training_percentage', 50), ('s
egmentation_method', 'bis'), ('segmentation_params', [50, 0.2, 0.8]), ('segmentation_metric', 'average'), ('boosting', 1), ('outlier', 1), ('overwrite', False), ('aws_s3_
bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cobertura_de_suelo/EAT_landcover_2000/031046'), ('features', ['NDVI']), ('metrics_features', 'minimum,maximum,average,qua
ntiles25')]
```

Results are stored to `s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2000`

Land cover change concept

Method abstract

MAD-Mex change detection module implements an image by image change detection, sourced by the IMAD MAF transformation and subsequent change detection by thresholding transformed MAF components. MAD-Mex, in addition, features the annual Tree Cover Density (TCD) product available for 1985 – 2015 provided by the University of Maryland for the country of Mexico. IMAD MAF is sourced by any satellite-, or synthetic image pair that share same conditions with respect to image bands and extent. It can be operated on selected scene pairs for different time periods or by calculated multi-temporal synthetic band stacks. In consequence it can be sourced by the tile based multi-temporal band metrics for any given band or vegetation index from Landsat or Sentinel time series, or by the synthetic band stacks calculated from these band metrics e.g. the average value for each pixel extracted from each band metric. The IMAD MAF transformations performs a image transformation towards maximizing the band correlation which in result provides transformed components that follow a gaussian distribution. Pixels values not within the normal distribution indicate changes. Those must be delineated applying thresholds to normally distributed MAF components. The result of this step are distinct image objects of potential changes. The second TCD based change detection is applied for imagery representing periods within the 1985 – 2015 time range. This method makes use of the annual available national tree cover density mappings and applies a given density threshold to delineate potential forest changes. The result of this step is a mask of potential changes with pixel values of:

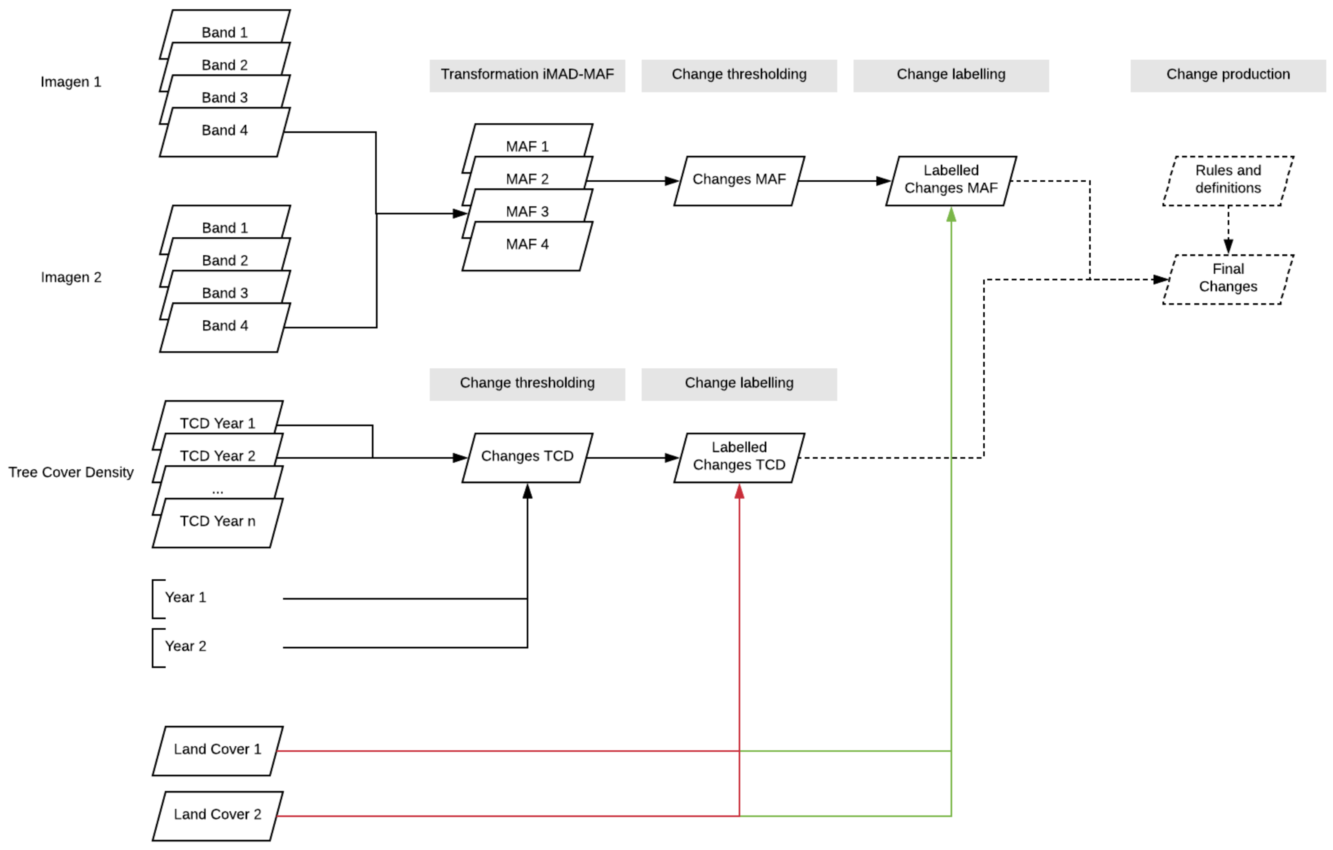
- 21: indicating deforestation (tree density loss towards < 10%)
- 22: indicating degradation (tree density loss)
- 23: indicating afforestation (tree density gain towards > 10%)
- 24: indicating regeneration (tree density gain)

Final change labelling of both derived change masks is performed using given land cover maps for same time period. Labelled changes are represented in 3 band raster images with:

- Band 1: change class derived from change detection (1..12 for IMAD MAF and 21,22,23,24 for TCD changes)
- Band 2: the land cover label of period/year 1 (before)
- Band 3: the land cover label of period/year 2 (after)

The so derived labelled changes must finally be converted to the final change map. This production must realize definitions on changes minimum mapping unit, change map coordinate reference system and must implement the developed class transition matrix which discriminates likely changes from impossible changes and with that performs a final filter to the change mask.

Data flows



Result structure

```
├── imadmaf
│   ├── i1_landsat_metricstack_020046_average.tif
│   ├── i2_landsat_metricstack_020046_average.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class.tif_pdf.png
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_mad.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_mad.tifcorrelations.csv
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_maf.tif
│   ├── m1_landsat_metricstack_020046_average_mask.tif
│   └── m2_landsat_metricstack_020046_average_mask.tif
├── imadmaf_label
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_labelfromto.tif
│   └── old_landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_labelfromto.tif
└── tcd_label
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    └── 90W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
```

Land cover change production

Change detection and labelling is implemented through the `ChangeDetectionFullWorkflow` adapter and receives the following input parameters:

`1st_image`

This variable sets the location, either locally or on AWS S3, of the first image to be used in change detection. No data must be valued 0. Example: `'s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2003/020046/bandmetricsstacks/landsat_metricstack_020046_average.tif'`

`2nd_image`

This variable sets the location, either locally or on AWS S3, of the second image to be used in change detection. No data must be valued 0. Example: `'s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2010/020046/bandmetricsstacks/landsat_metricstack_020046_average.tif'`

`1st_label`

This variable sets the location, either locally or on AWS S3, of the first land cover classification image to be used in change detection labelling. The land cover dataset must overlap the scenes/tile/granule geographic area and must consist of one band with pixel values representing the respective class. Example `'s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2003/mosaic/EAT_landcover_2003.tif'`

`2nd_label`

This variable sets the location, either locally or on AWS S3, of the second land cover classification image to be used in change detection labelling. The land cover dataset must overlap the scenes/tile/granule geographic area and must consist of one band with pixel values representing the respective class. Example `'s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2010/mosaic/EAT_landcover_2010.tif'`

`tcd_archives` [optional] This optional variable sets the locations, either locally or on AWS S3, as list or comma separated string, of UMN Tree Cover Density products available for 1985-2015 for the Mexican territory. This variable is optional. If given, changes for the region defined by the input images will also be calculates using tree cover density based change detection and labelling. All tree cover density datasets must overlap the scenes/tile/granule geographic area and must consist of one band with pixel values representing the respective class. Example: `['s3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/088W_19N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/089W_19N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/090W_19N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/088W_20N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/089W_20N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/090W_20N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/088W_21N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/089W_21N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/090W_21N.zip',]`

`tcd_interval`

This variable defines the time interval in years to be used in tree cover density change detection as list of numbers. This parameters are used to extract the corresponding bands from the UMN TCD dataset. Example: `[2003,2010]`

`tcd_threshold`

This variable defines the density threshold as absolute number of density percentage to be used in separating changes based on density changes. Example: `25`

`working_dir` [optional]

This targets to a local directory where the processing results shall be stored. If this varaiable is not defined the workind directory will be created automatically.

`aws_s3_bucket` [optional]

This optional variable defines the AWS S3 bucket where the final results shall be stored. For example `madmex-fao` indicates the madmex-fao bucket at `s3://madmex-fao/`.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`aws_s3_key` [optional]

This optional variable defines the AWS S3 key where the final results shall be stored. For example `projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047` together with bucket defined as `madmex-fao` indicates `s3://madmex-fao/projects/Cobertura_de_suelo/Jalisco_Landsat_2016/029047` as final storage destination.

Note: Only when both variables, `aws_s3_bucket` and `aws_s3_key`, are given, the results will be updated. Note: If AWS S3 upload is disabled the `clean_up` variable must be set to `False` in order to keep the results on local disk. If not, the workspace will be deleted automatically.

`overwrite` [optional]

This optional boolean variable can be set to `True` or `False`. This is especially helpful during development and debugging. If overwriting is disabled the most time consuming calculations like band stacking or image segmentation will not be executed again, if those intermediate results are already available in the defined `working_dir`. Default value is `True`.

`clean_up` [optional]

This optional boolean variable can be set to `True` or `False` and defines whether to remove the `working_dir` after process has finished. Default value is `True`.

Python execution:

```
a = ChangeDetectionFullWorkflow()
a.setInput('clean_up', False)
a.setInput('1st_image','s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2003/020046/bandmetricsstacks/landsat_metricstack_020046_average.tif')
a.setInput('2nd_image','s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2010/020046/bandmetricsstacks/landsat_metricstack_020046_average.tif')
a.setInput('1st_label','s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2003/020046/result/landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif')
a.setInput('2nd_label','s3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2010/020046/result/landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif')
a.setInput('tcd_archives',[
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/088W_19N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/089W_19N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/090W_19N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/088W_20N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/089W_20N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/090W_20N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/088W_21N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/089W_21N.zip',
        's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/090W_21N.zip',
    ])
a.setInput('tcd_interval',[2003,2010])
a.setInput('tcd_threshold',25)
a.setInput('working_dir','/Volumes/local_data/temp/madmex/processing/change')
a.setInput('overwrite',False)
a.setInput('aws_s3_bucket','madmex-fao')
a.setInput('aws_s3_key','projects/Cambio_de_cobertura_de_suelo/EAT_landcoverchange_2003-2010/020046/')
a.run()
```

Job execution:

```
k = [('1st_image', 's3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2000/018045/bandmetricsstacks/landsat_metricstack_018045_average.tif'), ('2nd_image', 's3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2003/018045/bandmetricsstacks/landsat_metricstack_018045_average.tif'), ('1st_label', 's3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2000/018045/result/landcover_018045_classify_mx_mapadereferencia2015_25m_warp_.tif'), ('2nd_label', 's3://madmex-fao/projects/Cobertura_de_suelo/EAT_landcover_2003/018045/result/landcover_018045_classify_mx_mapadereferencia2015_25m_warp_.tif'), ('tcd_archives', ['s3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/086W_20N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/085W_20N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/085W_21N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/087W_21N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/086W_21N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/085W_22N.zip', 's3://madmex-fao/madmex_testcases_aux/mexico_tcc_85_15/086W_22N.zip']), ('tcd_interval', [2000, 2003]), ('tcd_threshold', 25), ('aws_s3_bucket', 'madmex-fao'), ('aws_s3_key', 'projects/Cambio_de_cobertura_de_suelo/EAT_landcoverchange_2000-2003/018045/')]

exec_adapter('ChangeDetectionFullWorkflow', queued=True, waiting=False, kwarg=k)
```

S3 result structure in `s3://madmex-fao/projects/Cambio_de_cobertura_de_suelo/EAT_landcoverchange_2003-2010/019046`:

```
├── imadmaf
│   ├── i1_landsat_metricstack_020046_average.tif
│   ├── i2_landsat_metricstack_020046_average.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class.tif_pdf.png
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_mad.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_mad.tifcorrelations.csv
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_maf.tif
│   ├── m1_landsat_metricstack_020046_average_mask.tif
│   └── m2_landsat_metricstack_020046_average_mask.tif
├── imadmaf_label
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_labelfromto.tif
│   └── old_landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_labelfromto.tif
└── tcd_label
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    └── 90W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
```

Local result structure:

```
├── imadmaf
│   ├── i1_landsat_metricstack_020046_average.tif
│   ├── i2_landsat_metricstack_020046_average.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class.tif_pdf.png
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_mad.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_mad.tifcorrelations.csv
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_maf.tif
│   ├── m1_landsat_metricstack_020046_average_mask.tif
│   └── m2_landsat_metricstack_020046_average_mask.tif
├── imadmaf_label
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
│   ├── landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_labelfromto.tif
│   └── old_landsat_metricstack_020046_average_landsat_metricstack_020046_average_0_1_maf_class_labelfromto.tif
├── image_1
│   └── landsat_metricstack_020046_average.tif
├── image_2
│   └── landsat_metricstack_020046_average.tif
├── label_1
│   └── landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
├── label_2
│   └── landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
├── tcd
│   ├── 88W_19N_TS25.0_2003-2010_tcd_change.tif
│   ├── 88W_20N_TS25.0_2003-2010_tcd_change.tif
│   ├── 88W_21N_TS25.0_2003-2010_tcd_change.tif
│   ├── 89W_19N_TS25.0_2003-2010_tcd_change.tif
│   ├── 89W_20N_TS25.0_2003-2010_tcd_change.tif
│   ├── 89W_21N_TS25.0_2003-2010_tcd_change.tif
│   ├── 90W_19N_TS25.0_2003-2010_tcd_change.tif
│   ├── 90W_20N_TS25.0_2003-2010_tcd_change.tif
│   └── 90W_21N_TS25.0_2003-2010_tcd_change.tif
└── tcd_label
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 88W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 89W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_19N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_20N_TS25.0_2003-2010_tcd_change_labelfromto.tif
    ├── 90W_21N_TS25.0_2003-2010_tcd_change_after_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    ├── 90W_21N_TS25.0_2003-2010_tcd_change_before_landcover_020046_classify_mx_mapadereferencia2015_25m_warp_.tif
    └── 90W_21N_TS25.0_2003-2010_tcd_change_labelfromto.tif
```

Land cover change filter

Land Cover Change Filter

Conversion of detected changes in raster format to vector, filtering and attribute generation is performed by the `ChangeObjectsLabelFilter` adapter which receives the following input parameters:

- `change_label_images`
  - List of local or s3 imadmaf and tcd labelled change images
- `tcd_images` [optional]
  - List of local or s3 tcd images
- `maf_image`
  - local or s3 MAF image
- `metric1_image`
  - local or metric1 image
- `metric2_image` [optional]
  - local or metric2 image
- `filter_nochanges_ipcc` [optional]
  - Whether or not to filter out ipcc no-changes
- `gridshapefile`
  - local or s3 sensor grid shapefile
- `gridshapefile`
  - local or s3 sensor grid shapefile
- `gridcolumn`
  - shapefile column name with granule identifiers
- `tileid`
  - granule identifier
- `min_map_unit` [optional]
  - minimum mapping unit in sqm for change objects
- `working_dir` [optional]
  - Local working directory
- `aws_s3_bucket` [optional]
  - Target s3 bucket
- `aws_s3_key` [optional]
  - Target s3 key

The adapter takes some extra paramters defining basically external matrices xls files and their structure as provided by conafor:

```
    matrix_transition = MadMexInputVariable("matrix_transition")
    matrix_transition.documentation = "Transition matrix excel url"
    matrix_transition.type = VARTYPE_LITERAL_STRING
    matrix_transition.isMandatory = False
    matrix_transition.default_value = 's3://mm-fao/matriztransicion/MatrizVeriftransicion_MAD_MEX32 (001).xlsx'
    self.input_variables[matrix_transition.name] = matrix_transition

    matrix_dynamic = MadMexInputVariable("matrix_dynamic")
    matrix_dynamic.documentation = "Dynamic matrix excel url"
    matrix_dynamic.type = VARTYPE_LITERAL_STRING
    matrix_dynamic.isMandatory = False
    matrix_dynamic.default_value = 's3://mm-fao/matriztransicion/Dinamica_cambio_MADMEX_32cMR2015.xlsx'
    self.input_variables[matrix_dynamic.name] = matrix_dynamic

    dynamic_lct1_column = MadMexInputVariable("dynamic_lct1_column")
    dynamic_lct1_column.documentation = "dynamic_lct1_column"
    dynamic_lct1_column.type = VARTYPE_LITERAL_STRING
    dynamic_lct1_column.isMandatory = False
    dynamic_lct1_column.default_value = 'LC_T1'
    self.input_variables[dynamic_lct1_column.name] = dynamic_lct1_column

    dynamic_lct2_column = MadMexInputVariable("dynamic_lct2_column")
    dynamic_lct2_column.documentation = "dynamic_lct2_column"
    dynamic_lct2_column.type = VARTYPE_LITERAL_STRING
    dynamic_lct2_column.isMandatory = False
    dynamic_lct2_column.default_value = 'LC_T2'
    self.input_variables[dynamic_lct2_column.name] = dynamic_lct2_column

    dynamic_change_column = MadMexInputVariable("dynamic_change_column")
    dynamic_change_column.documentation = "dynamic_change_column"
    dynamic_change_column.type = VARTYPE_LITERAL_STRING
    dynamic_change_column.isMandatory = False
    dynamic_change_column.default_value = 'Cambio'
    self.input_variables[dynamic_change_column.name] = dynamic_change_column

    dynamic_description_column = MadMexInputVariable("dynamic_description_column")
    dynamic_description_column.documentation = "dynamic_description_column"
    dynamic_description_column.type = VARTYPE_LITERAL_STRING
    dynamic_description_column.isMandatory = False
    dynamic_description_column.default_value = 'descripcion T1'
    self.input_variables[dynamic_description_column.name] = dynamic_description_column

    transition_lct1_column = MadMexInputVariable("transition_lct1_column")
    transition_lct1_column.documentation = "transition_lct1_column"
    transition_lct1_column.type = VARTYPE_LITERAL_STRING
    transition_lct1_column.isMandatory = False
    transition_lct1_column.default_value = 'T1_ProMAD_MEX 34'
    self.input_variables[transition_lct1_column.name] = transition_lct1_column

    transition_lct2_column = MadMexInputVariable("transition_lct2_column")
    transition_lct2_column.documentation = "transition_lct2_column"
    transition_lct2_column.type = VARTYPE_LITERAL_STRING
    transition_lct2_column.isMandatory = False
    transition_lct2_column.default_value = 'T2_ProMAD_MEX 34'
    self.input_variables[transition_lct2_column.name] = transition_lct2_column

    transition_probability_column = MadMexInputVariable("transition_probability_column")
    transition_probability_column.documentation = "transition_probability_column"
    transition_probability_column.type = VARTYPE_LITERAL_STRING
    transition_probability_column.isMandatory = False
    transition_probability_column.default_value = 'Coef. Verf.porcentaje'
    self.input_variables[transition_probability_column.name] = transition_probability_column
```

Example adapter execution:

```
a = ChangeObjectsLabelFilter()
a.setInput('clean_up',False)
a.setInput('metric1_image',    "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/imadmaf/i1_landsat_metricstack_029045_average.tif")
a.setInput('metric2_image', "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/imadmaf/i2_landsat_metricstack_029045_average.tif")
a.setInput('tcd_images',["s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd/102W_20N_TS20.0_2000-2003_tcd_change.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd/102W_21N_TS20.0_2000-2003_tcd_change.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd/102W_22N_TS20.0_2000-2003_tcd_change.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd/103W_20N_TS20.0_2000-2003_tcd_change.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd/103W_21N_TS20.0_2000-2003_tcd_change.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd/103W_22N_TS20.0_2000-2003_tcd_change.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd/104W_21N_TS20.0_2000-2003_tcd_change.tif"])
a.setInput('change_label_images',["s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/imadmaf_label/landsat_metricstack_029045_average_landsat_metricstack_029045_average_0_1_maf_class_labelfromto.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd_label/102W_20N_TS20.0_2000-2003_tcd_change_labelfromto.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd_label/102W_21N_TS20.0_2000-2003_tcd_change_labelfromto.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd_label/102W_22N_TS20.0_2000-2003_tcd_change_labelfromto.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd_label/103W_20N_TS20.0_2000-2003_tcd_change_labelfromto.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd_label/103W_21N_TS20.0_2000-2003_tcd_change_labelfromto.tif", "s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/tcd_label/103W_22N_TS20.0_2000-2003_tcd_change_labelfromto.tif"])
a.setInput('filter_nochanges_ipcc',True)
a.setInput('merge',True)
a.setInput('maf_image',"s3://mm-fao/projects/Aguascalientes_Cambio_2000-2003/029045/imadmaf/landsat_metricstack_029045_average_landsat_metricstack_029045_average_maf.tif")
a.setInput('gridcolumn','CODE')
a.setInput('gridshapefile',"s3://mm-fao/grid/landsat_footprints_mexico_states.shp")
a.setInput('tileid','029045')
a.setInput('min_map_unit',10000)
a.setInput('working_dir','/Users/sgebhardt/tmp/label/')
a.run()
```

Land cover change mosaic

Land Cover Change Mosaic

Final vector mosaiced on beforehand created merged filtered changes Shapefiles is produced using the `ChangeDetectionMosaic` adapter which receives the following input paramters:

- `shapefilelist`
  - List of local or s3 merged filtered change shapefiles
- `outshapefilebasename`
  - Basename for resulting mosaic shapefiles (e.g. 'project1')
- `idcolumn` [optional]
  - Name of unique identifier. Defaults to 'oid'
- `resolution` [optional]
  - Pixelsize in intermediate rastered images. Defaults to '30'
- `splitsize` [optional]
  - Size in pixels of raster mosaic subsets to produce finale mosaic shapefiles. Defaults to '5000'
- `areascale` [optional]
  - Scale parameter for area attribute. Defaults to '0.0001' for sqm to ha conversion.
- `working_dir` [optional]
  - Local working directory
- `aws_s3_bucket` [optional]
  - Target s3 bucket
- `aws_s3_key` [optional]
  - Target s3 key

Example adapter execution:

```
a = ChangeDetectionMosaic()
a.setInput('clean_up',False)
a.setInput('shapefilelist',    [
    's3://mm-fao/projects/LCC_IRE_LANDSAT_2000-2010/020048/filter/changes_020048_label_merged_filter.shp',
    's3://mm-fao/projects/LCC_IRE_LANDSAT_2000-2010/020049/filter/changes_020049_label_merged_filter.shp',
])
a.setInput('outshapefilebasename', "mymosaic")
a.setInput('idcolumn', 'oid')
a.setInput('resolution',30)
a.setInput('splitsize',5000)
a.setInput('areascale',0.0001)
a.setInput('working_dir','/Users/sgebhardt/tmp/')
a.setInput('aws_s3_bucket','mm-fao')
a.setInput('aws_s3_key','test/changemosaic/')
a.run()
```

Image segmentation

Image Segmentation

This section describes the algorithms and usage implemented for performing image segmentation.

Berkeley Image Segmentation

In essence, threshold determines how large the segments get. Shaperate and compactness are weights that determine how the shapes look. Shaperate weighs the shape attributes versus the color attribute. Then within the shape calculation, the compactness rate weighs the compactness calculation over the smoothness calculation The main thing to control: A lower shaperate will let the segments go out further to follow similar colors. A higher shaperate will generally keep segments closer in, albeit less spectrally homogenious. For instance a threshold of 50, shaperate of 0.7, and compactness rate of 0.5 says "Give the regions 50 growth cycles while weighing the shape calculations over color homogeneity by 70/30 and (within the 70 percent bias) give equal weight to the compactness and smoothness calculations."

http://www.imageseg.com/faq

Execution from command line

The following workflows execute the Berkeley Image Segmentation.

```
BerkeleyImageSegmentationWorkflow
 --input_image  [path image to be segmented]
 --output_folder [path to output folder]
 --segmentation_threshold [BIS threshold, the number of region merging iterations]
 --segmentation_shape [BIS shape parameters ranging from 0..1]
 --segmentation_compactness [BIS compactness parameters ranging from 0..1]
 --baseiterations [Number of base iterations to be run parallel and or per tiling. Must be smaller then segmentation_threshold]
 --multiprocessing [Use multiprocessing boolean True/False]
 --tiling [Whether to use horizontal image tiling True/False. Will create horizontal image subsetting]
 --tilingrows [Number of pixels defining horizontal tile splitting size]
 --register [Whether to register the segmentation result or not 1/0]
```

```
docker-preamble docker-command BerkeleyImageSegmentationWorkflow
 --input_image  /madmex-data/products/metricslandsatworkflow/20046/indices/2010-01-01_2010-12-30/ls_stack_20046_2010-01-01_2010-12-30_NDVI_metrics.tif
 --output_folder /madmex-data/staging
 --segmentation_threshold 5
 --segmentation_shape 0.5
 --segmentation_compactness 0.5
 --multiprocessing True
 --tiling True
 --baseiterations 2
 --tilingrows 100
 --register 1
```

OpenCV Image Segmentation(s)

Implements various segmentation algorithms over raster images. It can be used to segment large aerial, satellite imagery of various formats supported by GDAL and various layouts like multispectral or hyperspectral. It uses OpenCV for it's core algorithms, and GDAL for undelying I/O. Implementation here follows several multithread and memory friendly optimizations, thus very large scenes are supported well. (https://github.com/cbalint13/gdal-segment)

```
GDALImageSegmentationWorkflow
 --input_image  [path image to be segmented]
 --output_folder [path to output folder]
 --algorithm [segmentation algoritm LSC, SLIC, SLICO or SEEDS]
 --iteration [number of merge iterations, default 10]
 --region_size [approximate object size (diameter, edge length) in pixels]
 -- register [Whether to register the segmentation result or not 1/0]
```

```
GDALImageSegmentationWorkflow
 --input_image  /madmex-data/products/metricslandsatworkflow/20046/indices/2010-01-01_2010-12-30/ls_stack_20046_2010-01-01_2010-12-30_NDVI_metrics.tif
 --output_folder /madmex-data/staging
 --algorithm SLIC
 --iteration 10
 --region_size 5
 --register 0
```

Geometric quality

Geometric quality represents the location error of the EOdata on pixel level. It can be measures with external GCP (ground control points) or relative to a reference data set. The geometric accuracy of pixel location is important for time series generation and therefore change detection.

Sensor product accuracies

Landsat

Landsat data is distributed by USGS in 2018 in different tiers Landsat collections:

- RT: for Real time data application
- T1: already image to image co registered and RMSE < 12m
- T2: RMSE > 12m due to effect like clouds insufficient GCPs, etc. Metadata contains product RMSE.

Madmex system needs USGS T1 and T2 data - older data models before collection 1 is not supported anymore.

Sentinel-2

ESA S2 MSI performance describes in their product guides radiometric and geometric accuracies of their products. With GCP support RMSE < 12.5m, without GCP is < 20m. There exist multi temporal registration products, as well.

RapidEye

Due to contract with data provider a geometric accuracy of 1 pixel (6m) was specified. Depending on the underlying DEM geolocation is in most areas given - RapidEye data is already coregistered with previous images.

Implementation

There are two adapters implemented using AROSICS - Automated and Robust Open-Source Image Co-Registration Software in the back. It is public available here Daniel Scheffler / arosics · GitLab. The code is installed in a docker image and can be used directly with *adapters.external.ortho.Coregistration*. A sensor wise co registration is implemented in *adapters.workflow.preprocess.sensor_orthocorrection.SensorCoregistration**. Adapter output is a directory containing the corrected band(s) and a PDF containing a quicklook with the applied GCPs and their RMSE.

Coregistration:

Input variables:

```
- [aws_s3_bucket]: (optional) AWS S3 bucket name
- [aws_s3_key]:  (optional) AWS S3 prefix key
- [clean_up]: (optional)  Remove working directory after adapter run
- [corrected_band]: url of to be corrected band
- [output_folder]: output_folder URL to mask image
- [reference_band]: url of reference band
- [register_result]: Update catalog
```

Output variables:

```
- [result]: result URL containing corrected band including qa metrics
```

SensorCoregistration

Input variables:

```
- [aws_s3_bucket]:  (optional) AWS S3 bucket name
- [aws_s3_key]:  (optional) AWS S3 prefix key
- [clean_up]:  (optional) Remove working directory after adapter run
- [correction_bundle_url]: Source local or remote archive of corrected image
- [output_folder]: Local processing directory
- [reference_bundle_url]: Source local or remote archive of reference image
- [register_result]: Update catalog
```

Output variables:

```
- [corrected_bundle]: Url of corrected image bundle
```

Radiometric quality

The radiometric quality affects the stability of the measurements taken by satellites over time / between sensors. Calibration is often in the responsibility of the satellite providers and done with LUT or scale/offset values.

Data coverage

Data coverage measures the 'quality' of existing time series by its numbers of observation per defined time period. It allows the comparison of different regions and the interpretation of mis classifications because of insufficient number of observations or different temporal distribution of data.

In MADMEX system there are two adapter implemented to calculate data coverage metrics:

```
* DataCoverage: uses an already calculated band image stack
* BandStackQualityControl:
```

DataCoverage - Calculating data coverage based on temporal stack: Input variables:

```
* [aws_s3_bucket]: AWS S3 bucket name
* [aws_s3_key]: AWS S3 prefix key
* [clean_up]: Remove working directory after adapter run
* [input_stack]: Source local or remote archive of reference image
* [nodata]: no data value to use
* [output_folder]: Local processing directory
* [percentage]: Calculate coverage in percentage
* [register_result]: Update catalog
```

Output variables:

```
* [data_coverage]: Url of corrected image bundle
```

BandStackQualityControl - Band Stack quality Input variables:

```
* [aws_s3_bucket]: AWS S3 bucket name
* [aws_s3_key]: AWS S3 prefix key
* [clean_up]: Remove working directory after adapter run
* [granule]: granule
* [land_mask_shape]: land_mask_shape.
* [overwrite]: overwrite
* [register_result]: Update catalog
* [scene_folder_list]: scene_folder_list
* [stack_metric_feature]: name of feature metric
* [working_dir]: working_dir
```

Output variables:

```
* [raster_absolut_sum]: Resulting raster file with absolut sum.
* [raster_percentage]: Resulting raster file with raster with the percentage .
* [synthetic_stack]: Resulting raster syntethic_stack.
```

Launch GUI

Launch GUI

The GUI applications runs inside the MADMEX docker as well. The following command executes the service on port 5000 (can be later redirected by docker/nginx). Passwords/credentials must be substituted before. The application creates a `gui.yaml` file in the starting directory. This file can be used as resource for the user-manager CLI to add new GUI users.

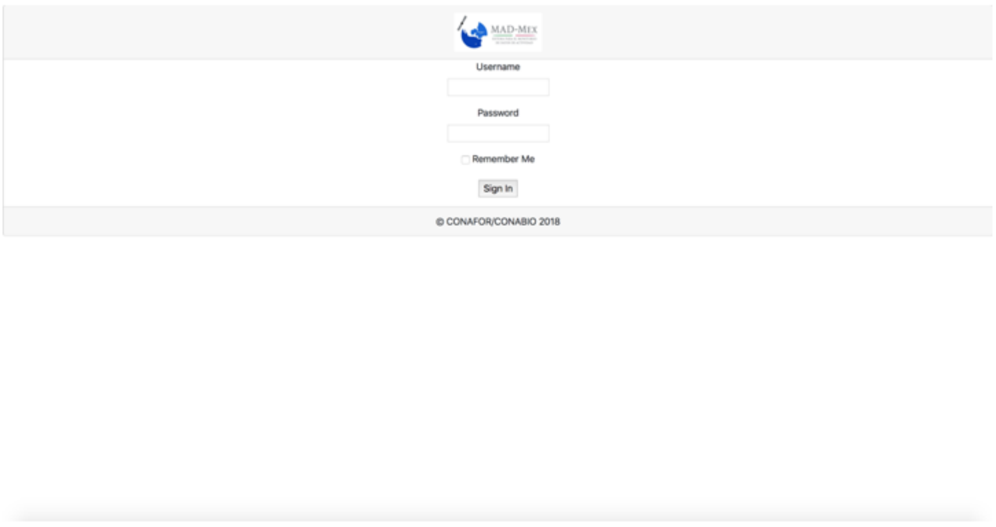
```
docker run -e FLASK_APP=serve.py -e FLASK_ENV=production -p 5000:5000 \
-v $(pwd)/gui.yaml:/tmp/gui.yaml -e GUI_USERS=/tmp/gui.yaml -e "ES_PASSWORD=xxx" \
-e USGS_USER=xxx -e USGS_PASSWORD=xxx --ulimit nofile=98304:98304 \
-v /var/run/docker.sock:/var/run/docker.sock  --hostname mm.cnf.gob.mx  -e TZ=UTC  \
-v /etc/localtime:/etc/localtime:ro -e BACKEND_URL=redis://:xxx@cluster.mm.cnf.gob.mx \
-e BROKER_API_URL=madmex:xxx@cluster.mm.cnf.gob.mx:8080 -e BROKER_URL=madmex:xxx@cluster.mm.cnf.gob.mx \
-e AWS_ACCESS_KEY_ID=xxx -e AWS_SECRET_ACCESS_KEY=xxx \
-e MRV_CONFIG=/madmex/resources/config/configuration.latest.aws-cnf.ini  \
--rm -it -w /madmex/interfaces/services/ -e WORKER_TZ=UTC \
-v /LUSTRE/MADMEX/:/LUSTRE/MADMEX madmex   gunicorn -b 0.0.0.0:5000 serve:app
```

Users guide

MAD-Mex Graphical User Interface

MAD-Mex principal processing workflows for land cover and change production can be executed and reviewed through the GUI. Supported sensors are Sentinel-2 and the Landsat sensors.

The GUI access is restricted and requires user authentication.



Username

Password

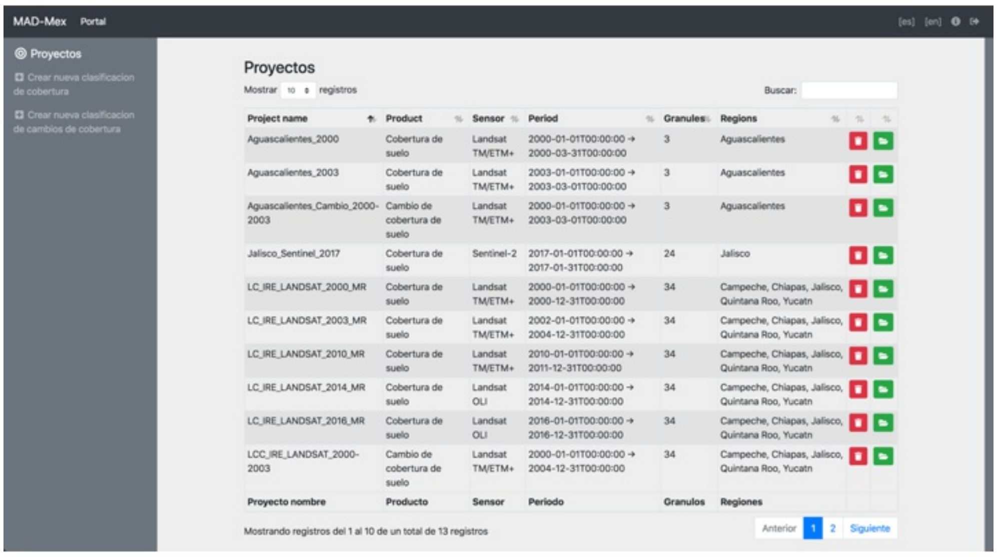
Remember Me

Sign In

© CONAFOR/CONABIO 2018

Portal

The project portal provides an overview of existing projects. Projects can either be deleted or opened for processing follwong the respective buttons. New land cover and landcover change projects can be created following the links on the left hand sidebar. The top toolbar allows switching GUI language to english or spanish, provides a link to page help and allows user logout.


MAD-Mex
Portal
[es]
[en]
Proyectos
Crear nueva clasificacion de cobertura
Crear nueva clasificacion de cambios de cobertura

Proyectos

Mostrar 10 registros

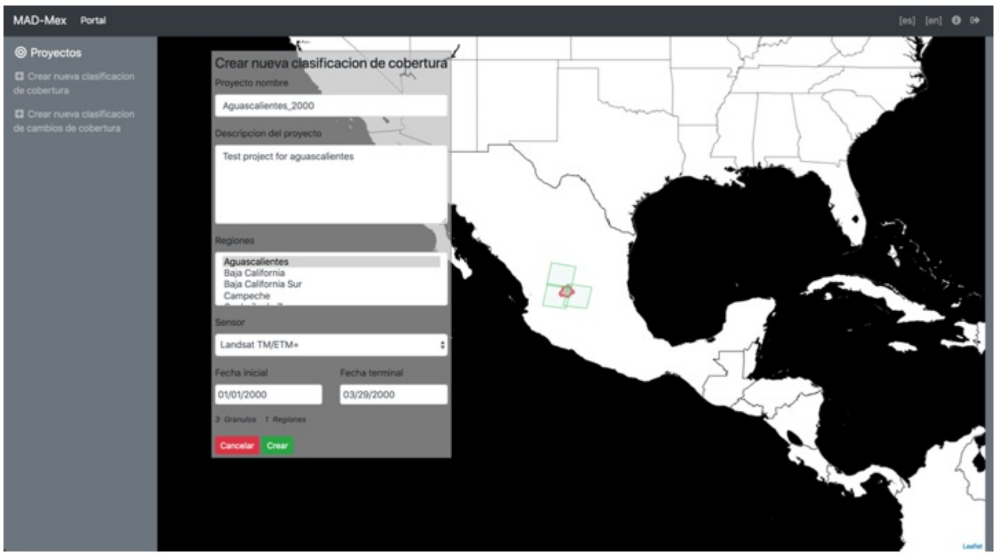
Buscar:

Project name	Product	Sensor	Period	Granules	Regions		
Aguascalientes_2000	Cobertura de suelo	Landsat TM/ETM+	2000-01-01T00:00:00 → 2000-03-31T00:00:00	3	Aguascalientes	🟥	🟩
Aguascalientes_2003	Cobertura de suelo	Landsat TM/ETM+	2003-01-01T00:00:00 → 2003-03-01T00:00:00	3	Aguascalientes	🟥	🟩
Aguascalientes_Cambio_2000-2003	Cambio de cobertura de suelo	Landsat TM/ETM+	2000-01-01T00:00:00 → 2003-03-01T00:00:00	3	Aguascalientes	🟥	🟩
Jalisco_Sentinel_2017	Cobertura de suelo	Sentinel-2	2017-01-01T00:00:00 → 2017-01-31T00:00:00	24	Jalisco	🟥	🟩
LC_IRE_LANDSAT_2000_MR	Cobertura de suelo	Landsat TM/ETM+	2000-01-01T00:00:00 → 2000-12-31T00:00:00	34	Campeche, Chiapas, Jalisco, Quintana Roo, Yucatn	🟥	🟩
LC_IRE_LANDSAT_2003_MR	Cobertura de suelo	Landsat TM/ETM+	2002-01-01T00:00:00 → 2004-12-31T00:00:00	34	Campeche, Chiapas, Jalisco, Quintana Roo, Yucatn	🟥	🟩
LC_IRE_LANDSAT_2010_MR	Cobertura de suelo	Landsat TM/ETM+	2010-01-01T00:00:00 → 2011-12-31T00:00:00	34	Campeche, Chiapas, Jalisco, Quintana Roo, Yucatn	🟥	🟩
LC_IRE_LANDSAT_2014_MR	Cobertura de suelo	Landsat OLI	2014-01-01T00:00:00 → 2014-12-31T00:00:00	34	Campeche, Chiapas, Jalisco, Quintana Roo, Yucatn	🟥	🟩
LC_IRE_LANDSAT_2016_MR	Cobertura de suelo	Landsat OLI	2016-01-01T00:00:00 → 2016-12-31T00:00:00	34	Campeche, Chiapas, Jalisco, Quintana Roo, Yucatn	🟥	🟩
LCC_IRE_LANDSAT_2000-2003	Cambio de cobertura de suelo	Landsat TM/ETM+	2000-01-01T00:00:00 → 2004-12-31T00:00:00	34	Campeche, Chiapas, Jalisco, Quintana Roo, Yucatn	🟥	🟩
Proyecto nombre	Producto	Sensor	Periodo	Granulos	Regiones		

Mostrando registros del 1 al 10 de un total de 13 registros

New land cover project

New land cover project definition requires the provision of a name, a description (optional), one or many regions (states), the sensor, and a range of dates. The map will automatically show selected regions and satellite image granules. Pushing the create button will store the project to AWS S3 and will register the project in the catalog.


MAD-Mex
Portal
[es] [en]
Proyectos
Crear nueva clasificacion de cobertura
Crear nueva clasificacion de cambios de cobertura
Leaflet

Crear nueva clasificacion de cobertura

Proyecto nombre: Aguascalientes_2000

Descripcion del proyecto: Test project for aguascalientes

Regiones: Aguascalientes, Baja California, Baja California Sur, Campeche

Sensor: Landsat TM/ETM+

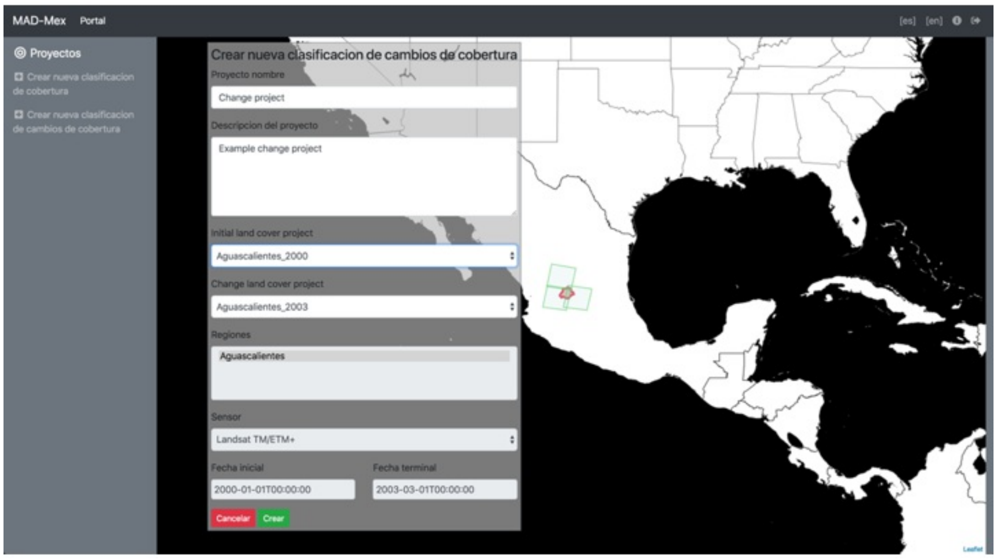
Fecha inicial: 01/01/2000 Fecha terminal: 03/29/2000

3 Granulos 1 Regiones

Cancelar Crear

New land cover change project

New land cover change project definition requires the provision of a name, a description (optional) as also the selection of two land cover projects. Selecting the first land cover project will filter the second land cover project options to those projects that are compatible (regions and sensor). Region, sensor and date parameters are extracted from the selected land cover projects. Pushing the create button will store the project to AWS S3 and will register the project in the catalog.


MAD-Mex
Portal
[es] [en]
Proyectos
Crear nueva clasificacion de cobertura
Crear nueva clasificacion de cambios de cobertura
Leaflet

Crear nueva clasificacion de cambios de cobertura

Proyecto nombre: Change project

Descripcion del proyecto: Example change project

Initial land cover project: Aguascalientes_2000

Change land cover project: Aguascalientes_2003

Regiones: Aguascalientes

Sensor: Landsat TM/ETM+

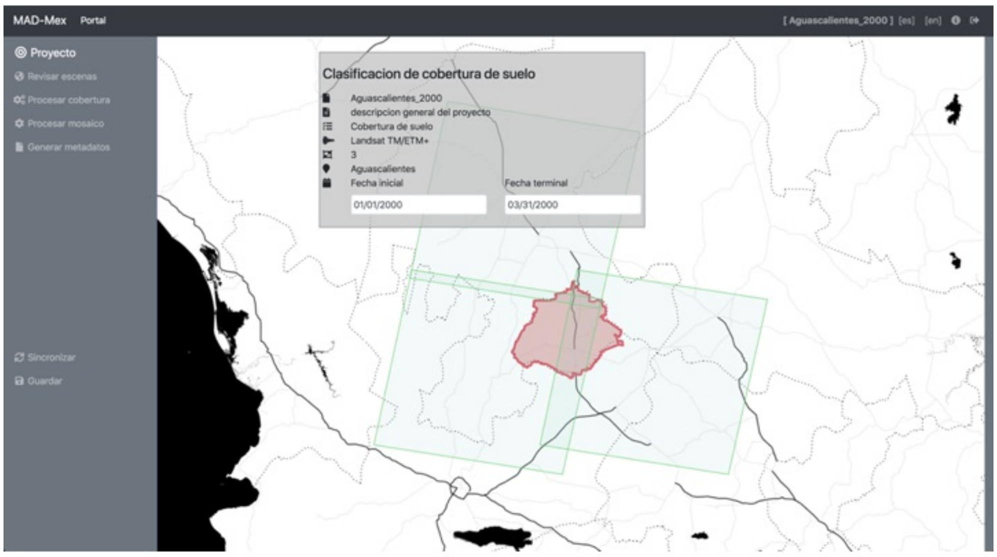
Fecha inicial: 2000-01-01T00:00:00 Fecha terminal: 2003-03-01T00:00:00

Cancelar Crear

Land cover processing

Project overview

Land cover processing page first provides an overview of the project definition. Here the user can modify the project dates. This is helpful when later in scene identification an inufficient amount of images is found for the given time interval. The left hand sidebar provides the links to the required processing steps and allows for permanant project synchrinisation and saving as also metadata generation and download.



Clasificacion de cobertura de suelo

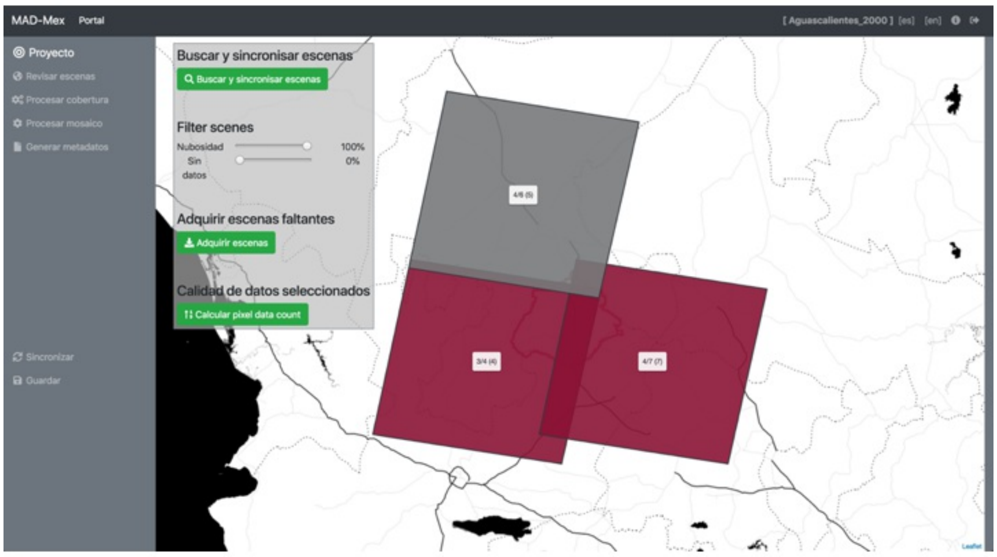
Aguascalientes_2000
descripcion general del proyecto
Cobertura de suelo
Landsat TM/ETM+
3
Aguascalientes
Fecha inicial 01/01/2000
Fecha terminal 03/31/2000

Scene identification, selection and acquisition

Scene selection section provides a map with the image granules drawn. Pushing the search button will trigger data search in the catalog. On termination of the search the map shows the number of selecetd scenes, the number of scenes found per granule and the number of scenes already acquired and available in aws s3. Global filters can be applied to select only those scenes fitting to the defined values for cloud and no data cover.

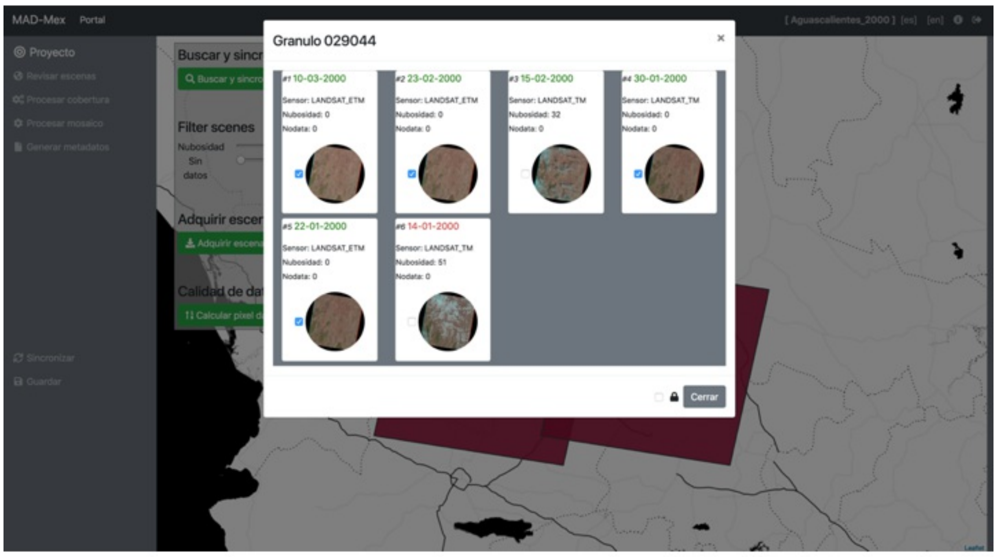
Pushing the acquire scenes button will send jobs to the defined processing queue. In case of Landsat first ESPA orders will be send and the download workflows are then started. These are runnung for 30 minutes and are polling ESPA order status. When ESPA order is finished, download is initiated. It is unlikely that all orders are finished within 30 minutes. The user should monitor the order status at the ESPA website or with the send ESPA emails. When all orders are completed the user can push the acquisition button again, no new orders will be created but the download workflow jobs are sent again to the queue. In the case of Sentinel-2 scene acquisition and pre-processing jobs are send to the queue.

For the scenes selected and available data count processes can be send to the processing queue by pushing the calculate pixel count button. For any given pixel in the granule based time-series, as defined by the scenes selected by the user, the number of pixels available, that are not affected by clouds, cloud shadows or no data, will be calculated and stored as image to the project aws s3 folder and will be registered to the catalog. Reviewing these images the user might go back to the scene selection and select mor images or even change the project period.


MAD-Mex
Portal
[Aguascalientes_2000] [es] [en]
Proyecto
Revisar escenas
Procesar cobertura
Procesar mosaico
Generar metadatos
Sincronizar
Guardar
Buscar y sincronisar escenas
Buscar y sincronisar escenas
Filter scenes
Adquirir escenas faltantes
Adquirir escenas
Calidad de datos seleccionados
Calcular pixel data count
4/6 (6)
3/4 (4)
4/7 (7)
Leaflet

Nubosidad 100%
Sin datos 0%

Clicking a granule on the map opens up a modal window with detailed information on images avaiable for this granule. Details provide information on acquisition date, sensor, cloud and no data cover, and quicklook. The acquisition date is coloured red if the scene is not yet available in aws s3 and green if it is available. Users select images that they want to include in processing by checking the respective checkbox. Finally, the granule can be locked, which prevents further global filters to be applied to this granule. Clicking on a preview image will open a modal window with enlarged preview image.



Granulo 029044

#1 10-03-2000
Sensor: LANDSAT_ETM
Nubosidad: 0
Nodata: 0

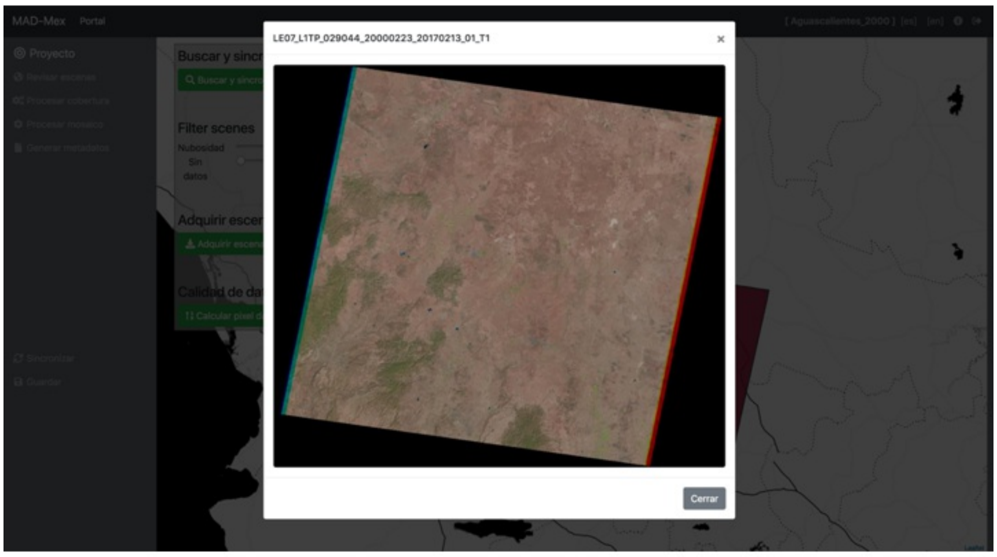
#2 23-02-2000
Sensor: LANDSAT_ETM
Nubosidad: 0
Nodata: 0

#3 15-02-2000
Sensor: LANDSAT_TM
Nubosidad: 32
Nodata: 0

#4 30-01-2000
Sensor: LANDSAT_TM
Nubosidad: 0
Nodata: 0

#5 22-01-2000
Sensor: LANDSAT_ETM
Nubosidad: 0
Nodata: 0

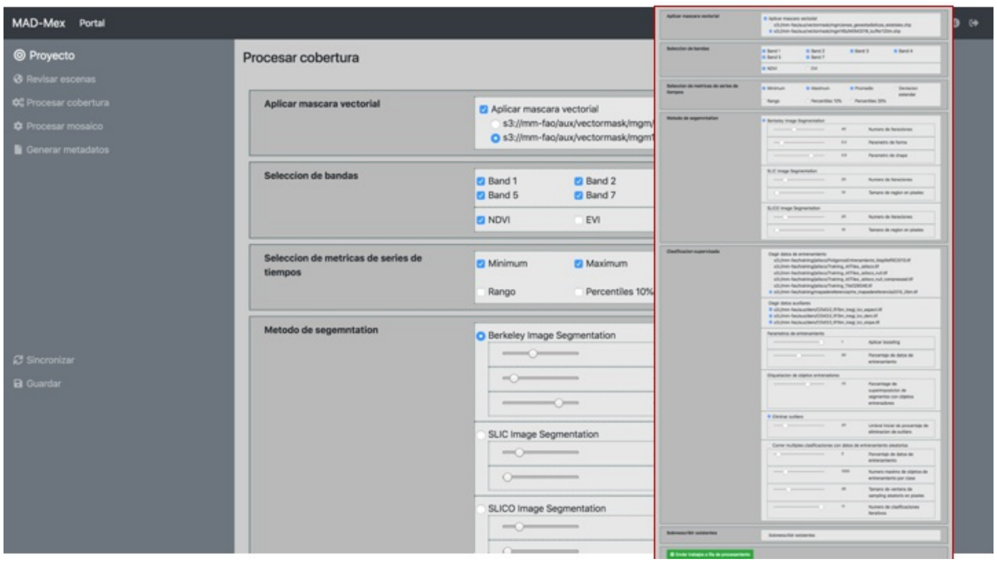
#6 14-01-2000
Sensor: LANDSAT_TM
Nubosidad: 51
Nodata: 0



LE07_L1TP_029044_20000223_20170213_01_T1

Land cover processing

Following the concept described in Land-Cover-Production-Concept and using the workflow adapter LandcoverMultitemporalWorkflow the user specifies the paramters required for granule/tile based land cover classification. The GUI helps the user by dedicated input elements and form validation to asure correct submission of job parameters to the processing queue. Jobs can only be submitted if all selected scenes are downloaded and available in the catalog. Classification results and intermediate results are stored to the project aws s3 folder and will be registered to the catalog.



Procesar cobertura

Aplicar mascara vectorial: ✓ Aplicar mascara vectorial; s3://mm-fao/aux/vectormask/mgm; ● s3://mm-fao/aux/vectormask/mgm

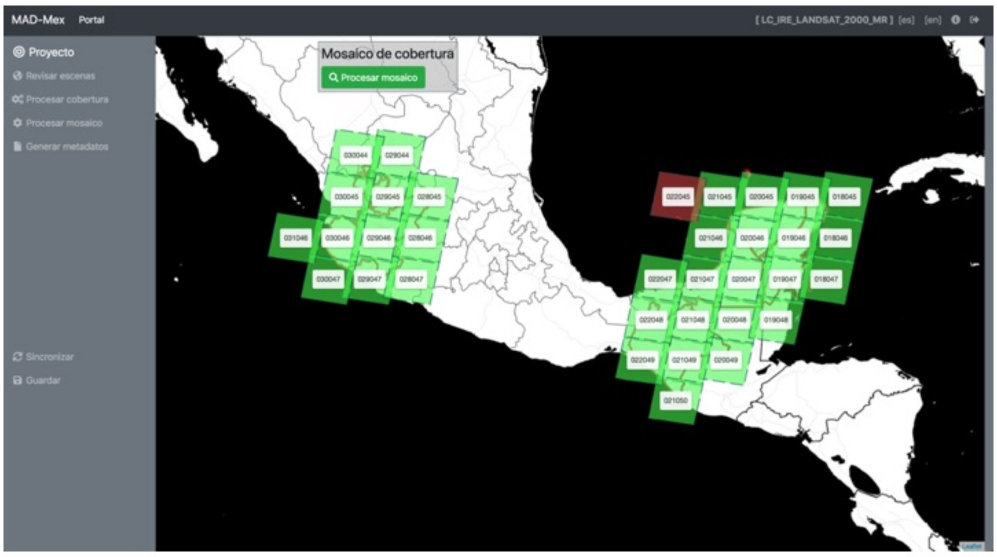
Seleccion de bandas: ✓ Band 1, ✓ Band 2, ✓ Band 5, ✓ Band 7, ✓ NDVI, EVI

Seleccion de metricas de series de tiempos: ✓ Minimum, ✓ Maximum, Rango, Percentiles 10%

Metodo de segemntation: ● Berkeley Image Segmentation; SLIC Image Segmentation; SLICO Image Segmentation

Land cover mosaic

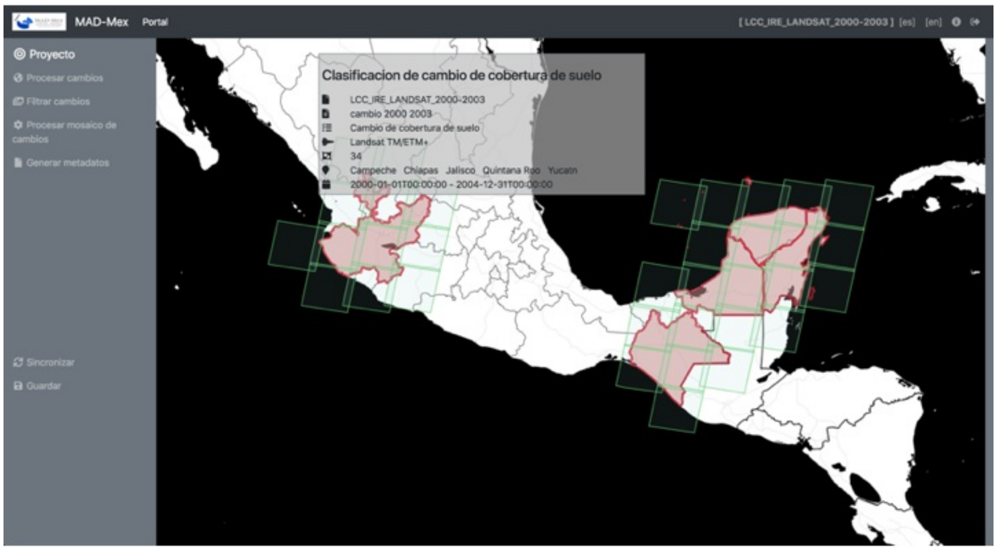
Final step in land cover production is the processing of the harmonized raster land cover mosaic. The map displays in green those granules which have benn succesfully been classified, in red are the granules which are still running or that failed. Pushin the process mosaic button will submit the mosaic procesing job to the processing queue and the results are stored to aws s3 project folder and are registered to the catalog.



Land cover change processing

Project overview

Land cover change processing page first provides an overview of the project definition. The left hand sidebar provides the links to the required processing steps and allows for permanant project synchronisation and saving as also metadata generation and download.

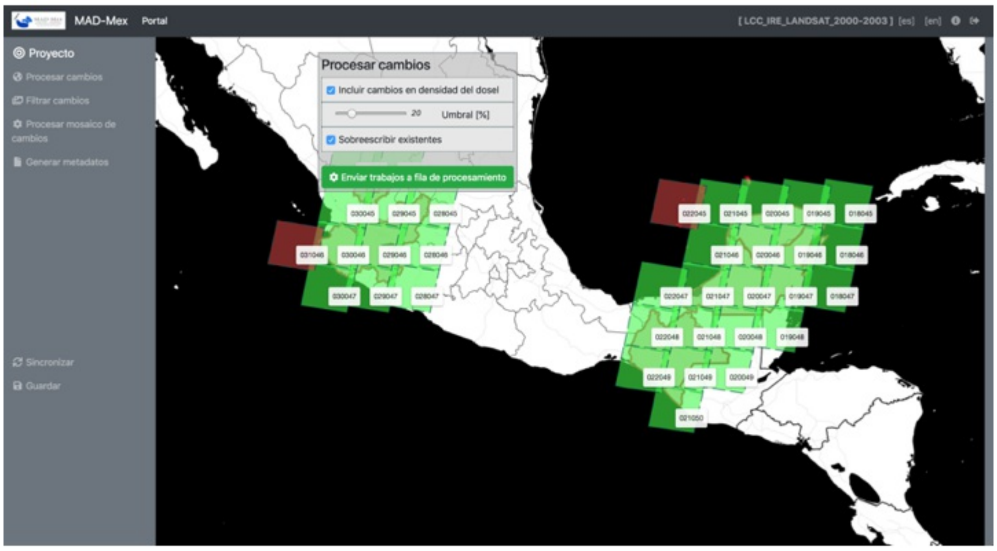


Clasificacion de cambio de cobertura de suelo

LCC_IRE_LANDSAT_2000-2003
cambio 2000 2003
Cambio de cobertura de suelo
Landsat TM/ETM+
34
Campeche Chiapas Jalisco Quintana Roo Yucatn
2000-01-01T00:00:00 - 2004-12-31T00:00:00

Change processing

The map provides an overview of granules with finished land cover classification in both base land cover projects in green. Red granules indicate incomplete land cover classifications. The actual land cover change processing can be triggered with optionally enabling additional change object identifaction using tree cover density products. Thereby the density threshold defining a change must be defined. Overwriting can be selected for full reprocessing. Pushing the process button will submit the change processing jobs to the processing queue and the results are stored to aws s3 project folder and are registered to the catalog.


MAD-Mex Portal
[LCC_IRE_LANDSAT_2000-2003] [es] [en]
Proyecto
Procesar cambios
Filtrar cambios
Procesar mosaico de cambios
Generar metadatos
Sincronizar
Guardar
Enviar trabajos a fila de procesamiento

Procesar cambios

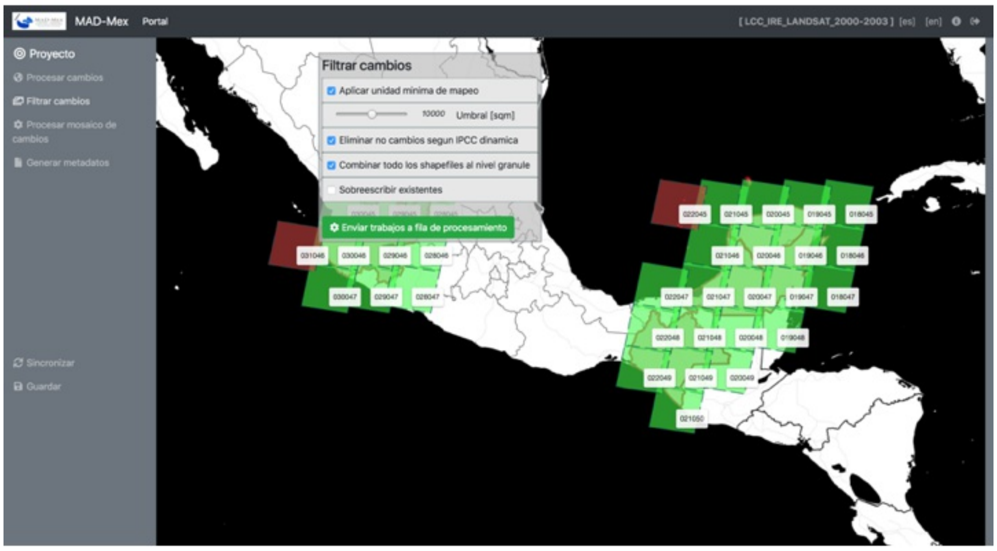
☑ Incluir cambios en densidad del dosel

20 Umbral [%]

☑ Sobreescribir existentes

Change objects extraction, filtering and labelling

Final processing step in change processing is the generation of labelled change vector shapefiles. The map provides an overview on finished and unfinished change detection jobs in the project. User can select an optional area threshold for change objects. The user can further define to filter out changes with land cover changes labelled as invalid according to the change dynamics definitions on IPCC level. When changes have also been caculated using the tree cover density product the user can force the production of a final merged shapefile with all change objects for a given granule in one shapefile. Pushing the process button will submit the change filtering processing jobs to the processing queue and the results are stored to aws s3 project folder and are registered to the catalog.


MAD-Mex Portal
[LCC_IRE_LANDSAT_2000-2003] [es] [en]
Proyecto
Procesar cambios
Filtrar cambios
Procesar mosaico de cambios
Generar metadatos
Sincronizar
Guardar
Enviar trabajos a fila de procesamiento

Filtrar cambios

☑ Aplicar unidad minima de mapeo

10000 Umbral [sqm]

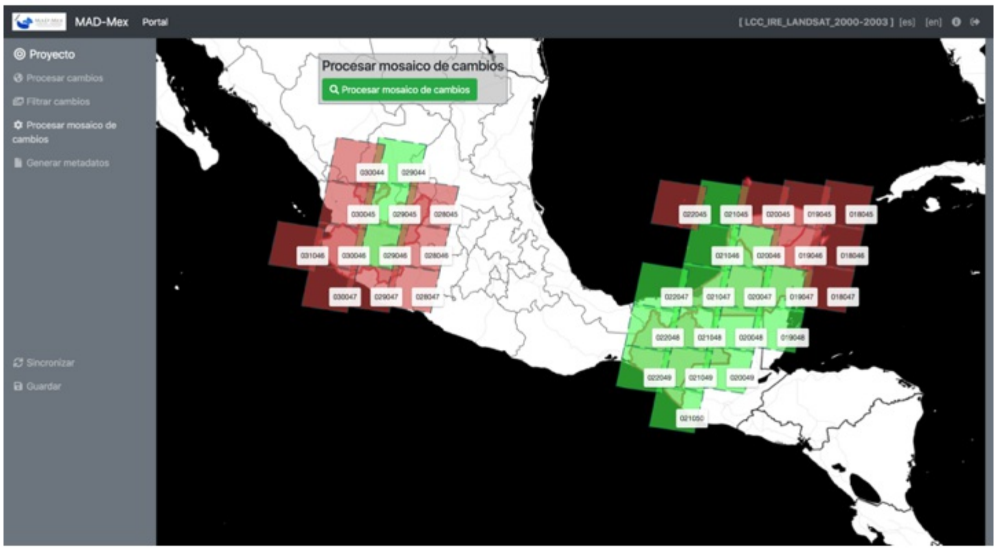
☑ Eliminar no cambios segun IPCC dinamica

☑ Combinar todo los shapefiles al nivel granule

☐ Sobreescribir existentes

Change objects mosaicing and topological cleaning

Final change vector mosaics are producing topological clean shapefiles over all merged filtered change shapefiles. Pushing the process button will submit the change filtering processing jobs to the processing queue and the results are stored to aws s3 project folder and are registered to the catalog.


MAD-Mex Portal
[LCC_IRE_LANDSAT_2000-2003] [es] [en]
Proyecto
Procesar cambios
Filtrar cambios
Procesar mosaico de cambios
Generar metadatos
Sincronizar
Guardar
Procesar mosaico de cambios
Procesar mosaico de cambios

GUI user management

GUI user management

GUI users must login into the application with their credentials. To manage users resource the CLI `user-manager.py` should be used. The application should be executed in a valid docker command including all MADMEX OS variables set. The system uses automatically a default `conafor` user for the GUI if no resource is set. To set a specific users resource a OS variable *GUI_USERS* must define a yam file already created by the user-manager CLI. This file should be located outside a docker instance and will be mounted in every GUI / user-manager docker container.

Basic functions of the CLI

```
Usage: user_manager.py [OPTIONS] COMMAND [ARGS]...

  MADMex processing system (c) 2018 CONAFOR, CONABIO, CATIE

  GUI user manager

Options:
  --version   Show the version and exit.
  -h, --help  Show this message and exit.

Commands:
  adduser     Add GUI user
  init        Initialize users resource
  listusers   List all users from users resource
  removeuser  Remove single user by its ID from users resource
  searchuser  Search particular user from users resource
```

Add users to resource

```
user-manager.py adduser {{YAML_FILE}} --username USER --email EMAIL   --password
```

Projects

Projects

Project definition

```
{
    "product": {
        "id": 1,
        "value": "lc",
        "label": "Cobertura de suelo"
    },
    "uuid": "f6732668-5e6d-44b4-ba0a-9a470750758c",
    "product_value": "lc",
    "scenes": null,
    "sensor_value": "ls57",
    "region": {
        "type": "FeatureCollection",
        "features": [{
            "geometry": {
                "type": "Polygon",
                "coordinates": [
                    [
                        [-102.287865181776, 22.416490039416786],
                        [-102.28537563213791, 22.417120584015603],
                        [-102.28183929975899, 22.413804174915516],
                        [-102.27398105615858, 22.41433115828133],
                        [-102.27091789390776, 22.411751778612796],
                        [-102.27305420539241, 22.40877260554491],
                        [-102.26739404606788, 22.405656682416646],
                        [-102.2643072103408, 22.399215317770572],
                        [-102.26095042746704, 22.398604052328835],
                        [-102.26081044946876, 22.395663808625795],
                        [-102.25563086463605, 22.391447303552486],
                        [-102.25524872423813, 22.38742865642485],
                        [-102.25364344028607, 22.38742135526703],
                        [-102.25362543974178, 22.38881789816681],
                        [-102.25029936337563, 22.387221706749436],
                        [-102.2491757011897, 22.385639349778298],
                        [-102.25056548786743, 22.38085392378658],
                        [-102.25494322487448, 22.378868375539717],
                        [-102.25320070663311, 22.374494503363238],
                        [-102.23557696697966, 22.37354387727419],
                        [-102.22813253189427, 22.374680900337516],
                        [-102.22517160185687, 22.374062665256638],
                        [-102.22333065481924, 22.371336400114565],
                        [-102.21198680863574, 22.372697320288648],
                        [-102.21197919494428, 22.36246518766482],
                        [-102.17737534226455, 22.362429735632432],
                        [-102.18072210203974, 22.350397937565006],
                        [-102.1498474709316, 22.34717505556252],
                        [-102.15323047689257, 22.29041021145548],
                        [-102.10658379972351, 22.285365454505495],
                        [-102.1078064990906, 22.282730467860798],
                        [-102.10524986517905, 22.280792829774697],
                        [-102.09092467978466, 22.29120718364915],
                        [-102.08957139400259, 22.28899886185833],
                        [-102.07562704201794, 22.306072278443693],
                        [-102.05673139361961, 22.299937805720013],
                                ...
                        [-102.29008501628718, 22.41994207621276],
                        [-102.287865181776, 22.416490039416786]
                    ]
                ]
            },
            "type": "Feature",
            "id": "0",
            "properties": {
                "CVE_ENT": "01",
                "NOM_ENT": "Aguascalientes",
                "OID": 1,
                "area": null
            }
        }]
    },
    "processing": {},
    "sensor": {
        "geojson": {
            "type": "FeatureCollection",
            "features": [{
                "geometry": {
                    "type": "Polygon",
                    "coordinates": [
                        [
                            [-101.7013623018939, 23.666852597424974],
                            [-102.03657953398758, 22.302284171189857],
                            [-102.055, 22.2273],
                            [-103.828, 22.485],
                            [-103.81061147139022, 22.5599435703236],
                            [-103.49342112043452, 23.927015960763583],
                            [-103.476, 24.0021],
                            [-101.683, 23.7416],
                            [-101.7013623018939, 23.666852597424974]
                        ]
                    ]
                },
                "type": "Feature",
                "id": "5"
                "id": "5",
                "properties": {
                    "pathrow": 29044,
                    "OID": 1,
                    "NOM_ENT": "Aguascalientes",
                    "CODE": "029044",
                    "CVE_ENT": "01"
                }
            }, {
                "geometry": {
                    "type": "Polygon",
                    "coordinates": [
                        [
                            [-102.03657953398758, 22.302284171189857],
                            [-102.036, 22.3022],
                            [-102.390496651119, 20.834719464179255],
                            [-102.402, 20.7871],
                            [-104.159, 21.0422],
                            [-104.14178818330669, 21.11726923958942],
                            [-103.811, 22.56],
                            [-103.81061147139022, 22.5599435703236],
                            [-102.03657953398758, 22.302284171189857]
                        ]
                    ]
                },
                "type": "Feature",
                "id": "10",
                "properties": {
                    "pathrow": 29045,
                    "OID": 1,
                    "NOM_ENT": "Aguascalientes",
                    "CODE": "029045",
                    "CVE_ENT": "01"
                }
            }, {
                "geometry": {
                    "type": "Polygon",
                    "coordinates": [
                        [
                            [-100.50909029074957, 22.227313116080136],
                            [-100.84443257241506, 20.839124342988903],
                            [-100.857, 20.7871],
                            [-102.614, 21.0422],
                            [-102.59678818330669, 21.11726923958942],
                            [-102.266, 22.56],
                            [-102.26561147139022, 22.5599435703236],
                            [-100.491, 22.3022],
                            [-100.50909029074957, 22.227313116080136]
                        ]
                    ]
                },
                "type": "Feature",
                "id": "183",
                "properties": {
                    "pathrow": 28045,
                    "OID": 1,
                    "NOM_ENT": "Aguascalientes",
                    "CODE": "028045",
                    "CVE_ENT": "01"
                }
            }]
        },
        "id": 1,
        "value": "ls57",
        "label": "Landsat TM/ETM+"
    },
    "period": [{
        "start": "01/01/2000",
        "end": "03/31/2000"
    }],
    "name": ["Aguascalientes_2000"]
}
```

Catalog introduction

MADMEX catalog system

All EO data, MADMEX products and MADMEX related data is organised in a catalog system to get fast access of available data. The base of the catalog system is an ElasticSearch (ES) database (Elasticsearch: RESTful, Distributed Search & Analytics | Elastic) which is a document based no SQL database which can handle and index structured data very well. Beside the catalog the database stores logging information of the processing, too.

Component installation

The ES server and its components are installed easily via docker images. The docker files are included in the Madmex repository. Server host name can be set with OS variable /ES_HOST/. Otherwise a default server is used (logging.mm.cnf.gob.mx). The ElasticSearch stack includes logstash for logging collection and kibana - a ES query client. Kibana is especially useful to query the processing logs and search for the MADMEX catalog or generate quick product inventories.

Search indices

There are several documents and indices configured to store the different data structures. Most documents have a file index for detailed (file based) queries and an aggregated index based on the EO product (e.g. masks, SR data, etc.). Indices organise a specific structure and provide a fast access based on a filter query.

EO data

EO datasets are EO imagery in different processing levels from different sensors. They can be managed in different locations and consists of multiple components (e.g. masks, surface reflectances, indices, etc.) The catalog incorporates EO data from providers (USGS for Landsat data and EOSS for Sentinel2 data) to support direct data querying of existing data and already downloaded/processed products in the local/cloud partitions. Provider based data should be updated regularly to include latest data takes - downloaded data locations must be updated after successful data ordering. Both data sets can be joined per sensor by its granule and acquisition date.

Used indices:

```
* madmex_eodata: Product level based existing EO data in the MADMEX system
* madmex_eodata_files: On file level based existing EO data in the MADMEX system
* madmex_eodata_usgs: Available Landsat data from USGS
* madmex_eodata_eoss: Available Sentinel2 data organised by EOSS catalog system ([EOSS Catalog](https://catalog.eoss.cloud))
```

File based items of downloaded resources include an id to reference their corresponding products (e.g. image files of a fmask product).

MADMEX product data

MADMEX products include all sub products of land cover and land cover change products. Similar to EO data and data file indices, products consists of grouped products (on component or sub product level) and product files. Products are organized with the following structure:

```
* product name
* Granule or general (e.g. config or mosaic)
* Component or sub product
```

Component folders contain the files for each individual component. Most adapters create results like files and upload these to a centralized partition. Results from different adapters should not be mixed to ensure correct extraction of attributes for the catalog system.

Used indices:

```
* madmex_product: on sub product (component) level aggregated information
* madmex_product_file: registered files for the scanned MADMEX products
```

Similar to the EO data concept a referencing id exists for product files to their corresponding component.

General data structures

File resources in general can be tagged and managed with the catalog system.

The catalog system is used in the MADMEX GUI environment, can be used with the kibana client and can be managed with a CLI application in MADMEX.

Updating indices

The ES database ensures a quick access to available files in a local/cloud partition. Depending on the number of files a direct access would be insufficiently slow. Hence, the catalog needs to be synced regularly with the file base to ensure an up to date state of the catalog. The update process can be triggered manually or in intervals using cron jobs. All partitions needs to be scanned.

Catalog use

Catalog use

The catalog must be synced with the storage partitions regularly. Otherwise the catalog will return non correct results after querying the system and shows an incorrect state in the web client. Management tasks can be done with the catalog cli application.

Catalog preparation

After the transfer of EO data or madmex product data to a specific local/cloud partition, files should be scanned and valid resources should be synced to the catalog.

Necessary syncs

```
* EO data data
* product data
* project data
```

A full sync includes a database clean up which removes older data. This ensures a consistent state of the data and a common time stamp of the data. The CLI must be started in the Madmex docker environment to include all necessary environment variables, python dependencies, etc. Valid index names are:

- madmex_eodata
- madmex_eodata_file
- madmex_product
- madmex_product_file
- madmex_project
- madmex_eodata_usgs
- madmex_eodata_eoss

The list should be actualized with the `indices` command of the catalog app. The `base_folder` option defines the location of the data.

```
python interfaces/cli/catalog.py sync --index madmex_project --base_folder s3://mm-fao/projects
```

After Madmex installation the catalog as one component of the system needs manual work, too. All indices which should be used/visualized in kibana needs to be created within kibana (Management -> Index patterns -> Create index) after index documents are uploaded to ElasticSearch.

Wiki ToC version 2.1

ToC

- Image Segmentation

https://github.com/CONABIO/madmex-v2/wiki/v22.Production

Production operations guide

Production operation consists of the following aspects

- Infrastructure administration
- Production monitoring
- Error analysis

Infrastructure administration

Every production needs a previous planning of necessary infrastructure components. Processing demands define the overall processing time. Number of computing nodes and executed MADMex worker instances reduce overall processing time to guarantee production on time. Every executed adapter workflow defines its requirements in term of minimum RAM, disk space and number of CPU cores. These requirements affect the selection of EC2 instance types and further configuration (e.g. partition size). Only the proper selection and configuration of the infrastructure ensures the successful production run. And on time processing of the selected project. Suggested configuration for major workflow adapters (for one worker instance):

- LandcoverMultitemporal: 32GB RAM, 100GB disk
- ChangeDetectionGeneral: 64GB RAM, 150GB disk

The current AMI supports an auto-start of one worker on a newly created and started ec2 instance. The auto-start process includes a code update from the MADMex github repository. After the pull of the new code new docker images are build and one instance of the worker is raised. The worker automatically connects to the default queue (`celery`). After successful start-up the worker is ready for the processing.

Certain errors and exceptions may crush the celery client and prevent the worker's temp workspace clean up process. In this case the worker must be cleaned manually or eliminated from the processing queue (ec2 instance termination). Errors can be detected using the monitoring components of the system (kibana or flower).

During the production process new worker can be raised to compensate failed instances. It must be validated whether the worker already had consumed the jobs or if celery still has the broken job in its queue. Missing or all jobs can be always resend to the queue by the operators or the administrator.

After the processing it must be validated if all results are finished for a production. In that case the infrastructure can be terminated to limit costs.

Production monitoring

There are two tools available to monitor the recent and historical processing flow.

Kibana

Kibana is a query client for elasticsearch. All adapters and many functions provide logging information which is collected by logstash and saves in elasticsearch. Basic functionality is described at https://www.elastic.co/guide/en/kibana/current/getting-started.html. Certain log messages contain additional information like used input parameters or detailed exception traceback. In addition to the logs the system collects automatically node metrics (CPU load, memory consumption, IO etc) for each processing node. There exist a system wide dashboard showing CPU load for all connected computing nodes and node specific metrics in detail.

Flower

Flower is a monitoring component of Celery. It allows to see the connected workers and their current work load. It shows the number of failed and successful run jobs of each worker. Moreover for each job the arguments and the resulting variables or exceptions can be queried.

In addition the final projects can be evaluated / checked for existence in s3 or the local storage system.

Error analysis

The previous described management tools allow to scan for errors during a production cycle. Errors occur for different reasons. These are the most common error types:

- Computing node related errors
  - Connection errors
  - Memory related errors
  - Disk space related errors
- MADMex code errors
  - Project / configuration related errors
  - Syntax errors