Mitiga

Hitting the mark: Using AI in hail and windstorm identification

Dr. Alejandro Martí
Sam Macintyre

30.08.2021

Convective Storms

50%

of all-weather event insured losses globally (x2 from the 90’s)

Hail

$8-14B

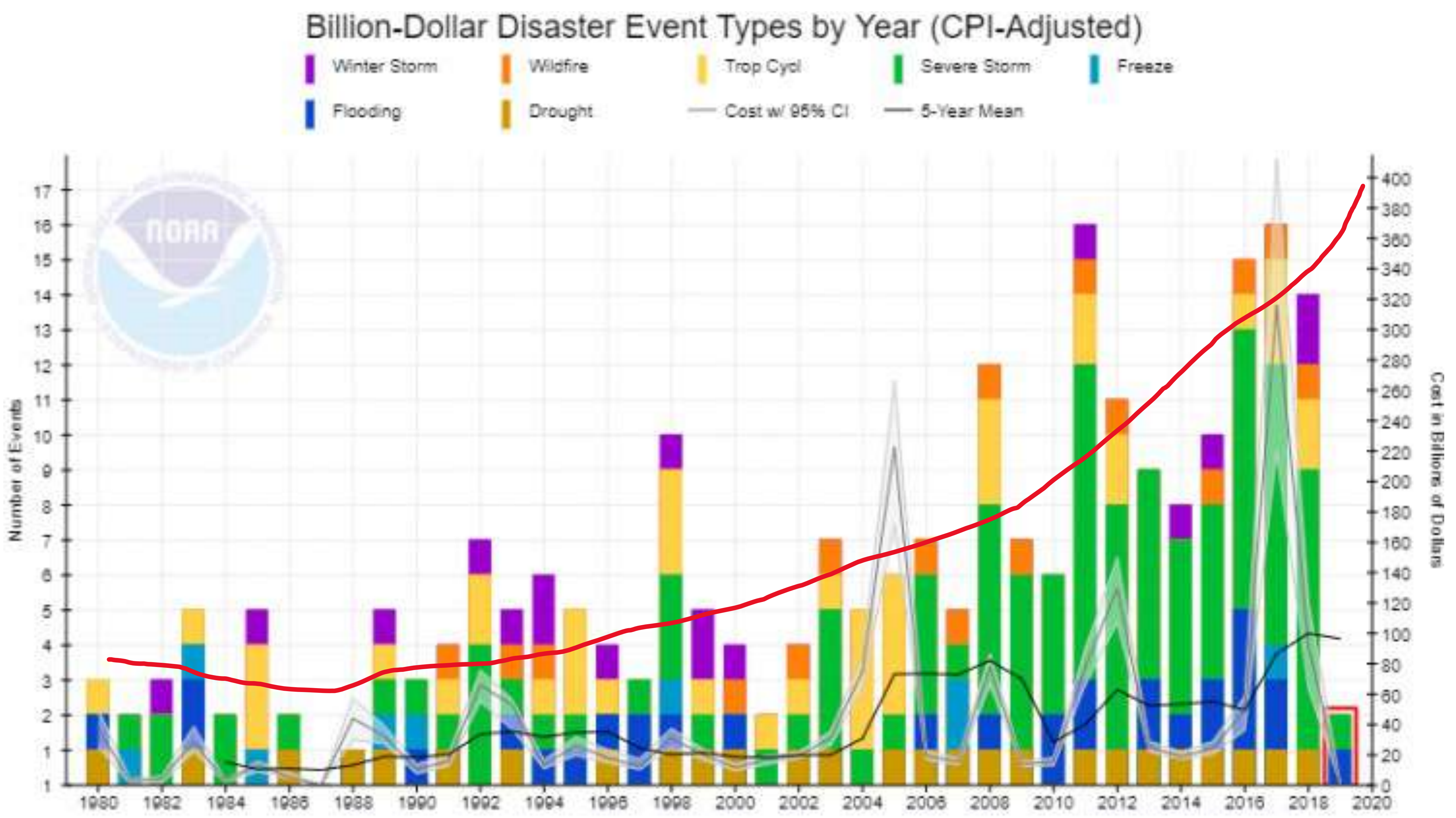
Average losses in the US alone
2020, Aon

Windstorms

$5B

Average losses in Europe alone

2015, JRC technical report



Adapted from NOAA, 2020

Need for a unified risk modeling

Need for a unified risk modeling

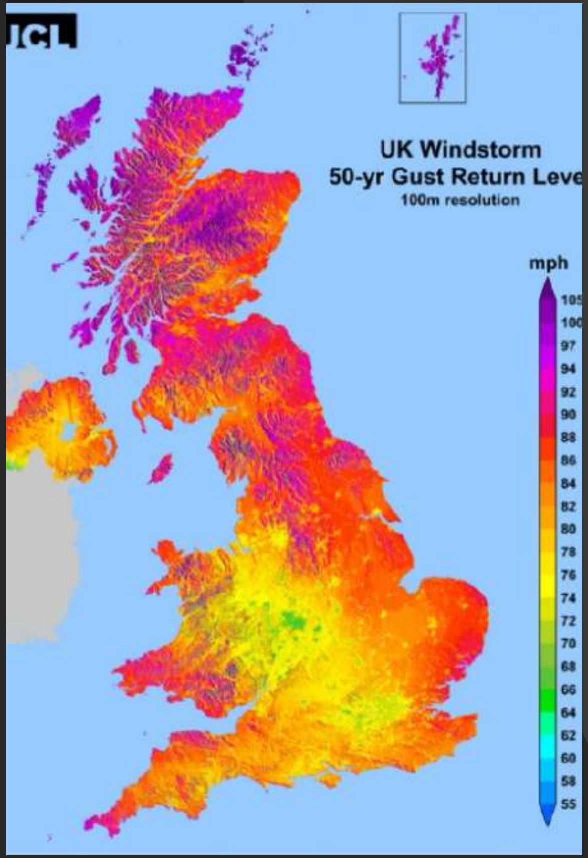
No standard guidelines for hail and windstorms hazard mapping

Current references dependent on well-established on-site observation SCS networks

Not the case in most countries (specially in LMIC)

Hazard mapping: rationale

Hazard maps display the location, frequency and severity of the respective hazard and are an essential part of hazard assessment and mitigation strategies.

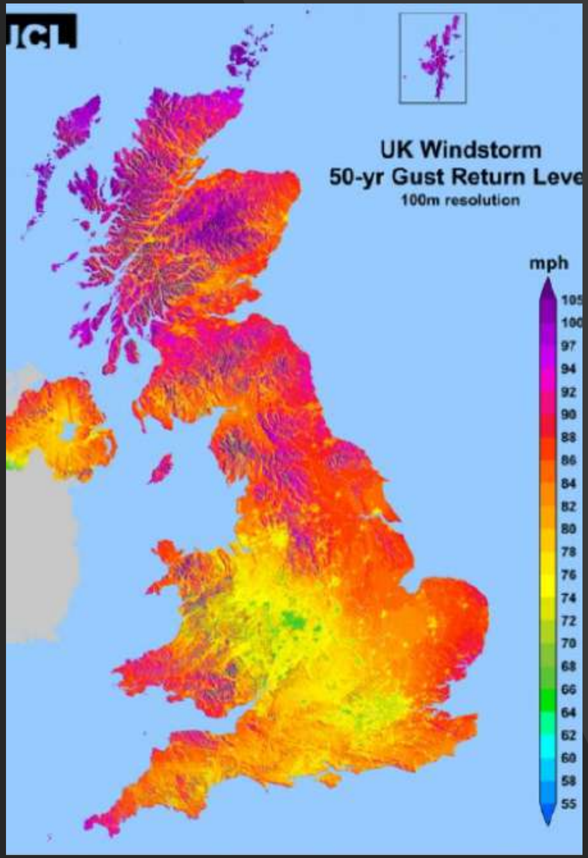


Windstorm Hazard Map – UK (UCL)

Hazard mapping: rationale

Accurate and reliable hazard mapping plays a key role in informing many operations and sectors:

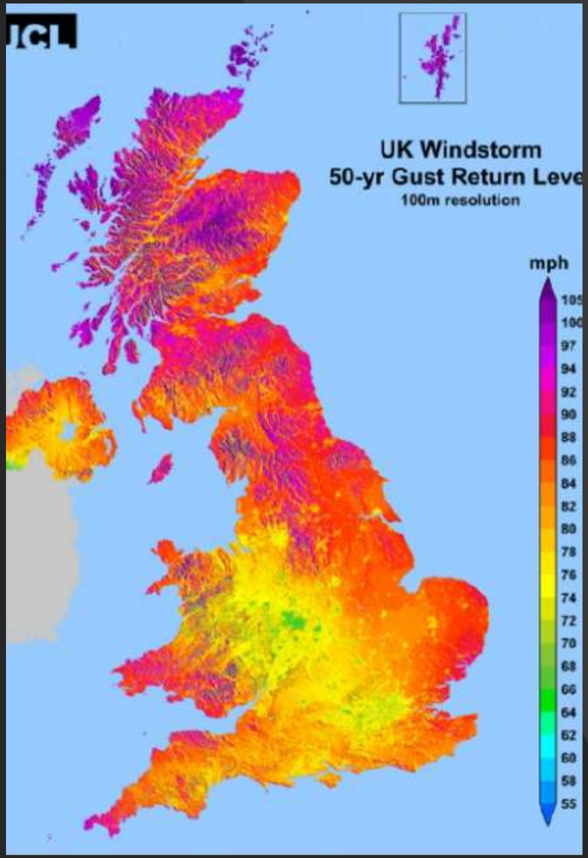
- **Operations**
  - Disaster management and planning
  - Infrastructure planning and resource allocation
- **Sectors**
  - Insurance (premiums)
  - Communities vulnerable to high impact hazards



Windstorm Hazard Map – UK (UCL)

Hazard mapping: challenges

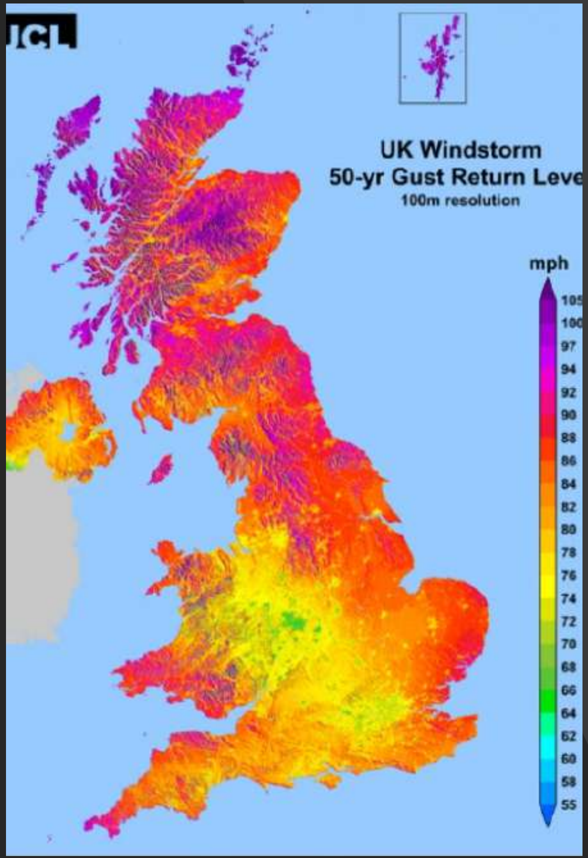
Hazard mapping requires estimates of event severity and frequency (recurrence) = initial **Severe Weather Event Dataset (iSWED)**



Windstorm Hazard Map – UK (UCL)

Hazard mapping: challenges

- **Historical data availability**
  - Lack of monitoring systems and infrastructure
  - Not distributed covered (geographically)
- **Severity not recorded**
  - Event data may exist without intensity
- **Data consistency**
  - Data must be consistent in scales, format and have sufficient geographical distribution



Windstorm Hazard Map – UK (UCL)

Hazard mapping: challenges

Full iSWED (rare)

Partial iSWED (few locations)

Insufficient iSWED (most common)

The role of AI in hazard mapping: hail and windstorms

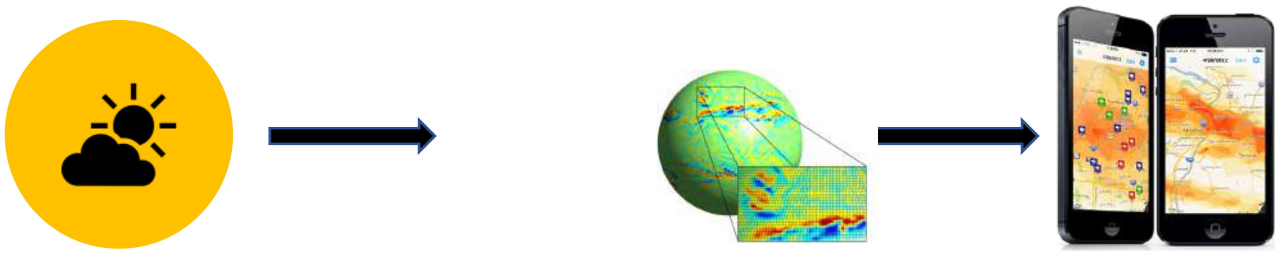

Full iSWED
Hazard Map

Full iSWED

Hazard Map

Full iSWED

Hazard Map



Partial iSWED

NWP + AI

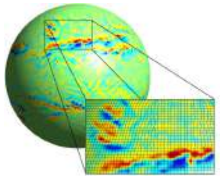
Hazard Map

- **NWP Simulations + Statistical theory** can be used to create event-frequency-intensity datasets that can be used for hazard mapping

Full iSWED

Hazard Map

Partial iSWED

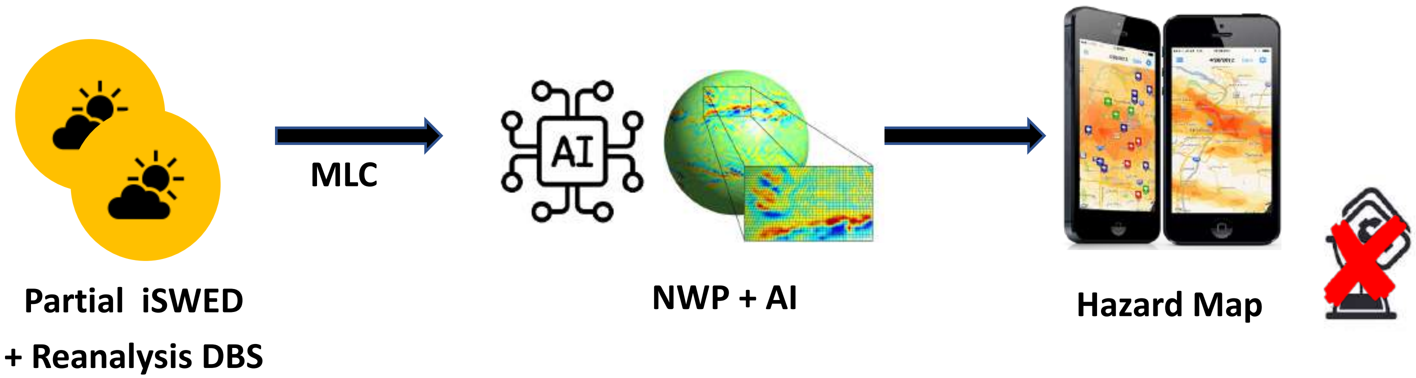


NWP + AI

Hazard Map

- **Costly** (HPC, data size, expertise)
- **Time consuming**

Inefficient as most "event-days" do not have meteorological conditions to produce desired hazard



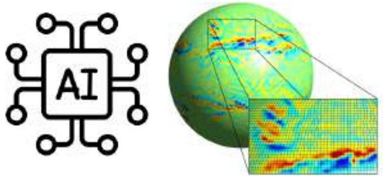
- **The Role of AI:** Machine Learning Classifiers (**MLC**) can identify the conditions that most likely lead to convective events, reducing the number of simulations and time required dramatically

Full iSWED

Hazard Map

Partial iSWED
+ Reanalysis DBS

Class
imbalance

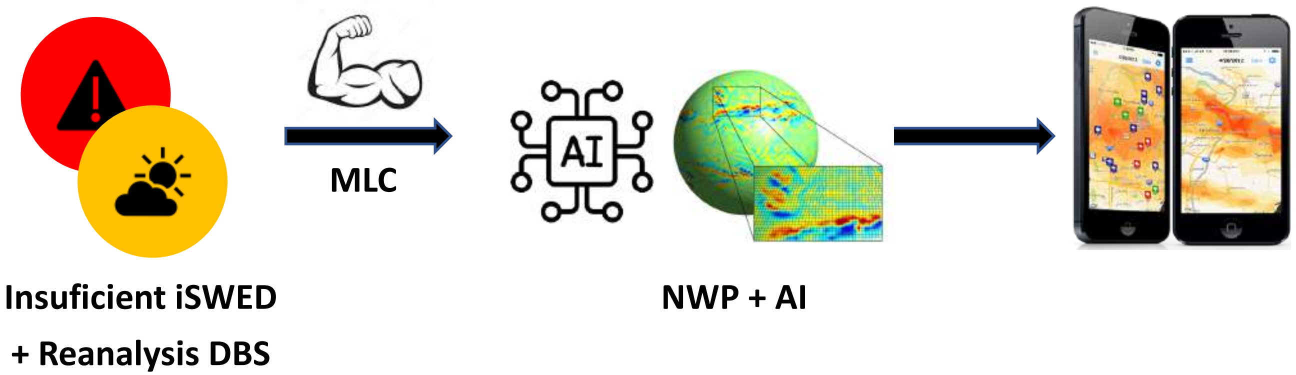


NWP + AI

Hazard Map

- **The Role of AI:** Machine Learning Classifiers (**MLC**) can identify the conditions that most likely lead to convective events, reducing the number of simulations and time required dramatically

- **The Role of AI:** Machine Learning Classifiers can identify the conditions that most likely lead to convective events, reducing the number of simulations and time required dramatically



- **Transfer learning** involves leveraging problem contexts which are well-defined and data rich
- It can be used in regions where data is too sparse/incomplete to train **Machine Learning models (MLC)**

Full iSWED

Hazard Map

Partial iSWED
+ Reanalysis DBS

Class
imbalance

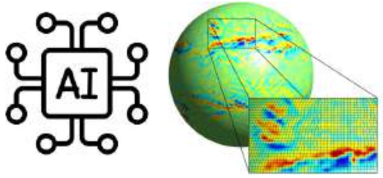
NWP + AI

Hazard Map

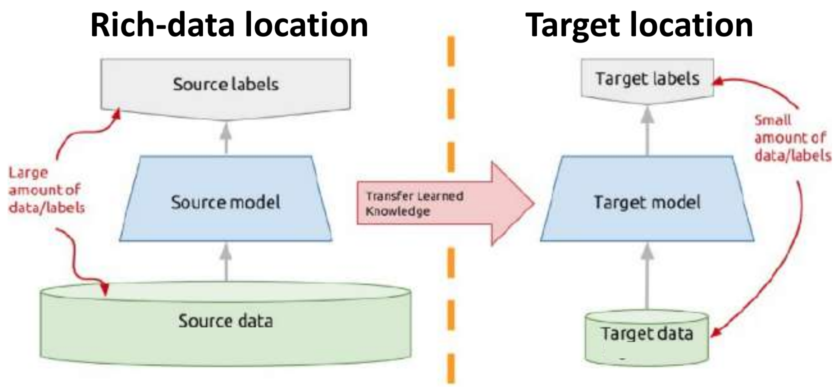
- **The Role of AI:** Machine Learning Classifiers can identify the conditions that most likely lead to convective events, reducing the number of simulations and time required dramatically

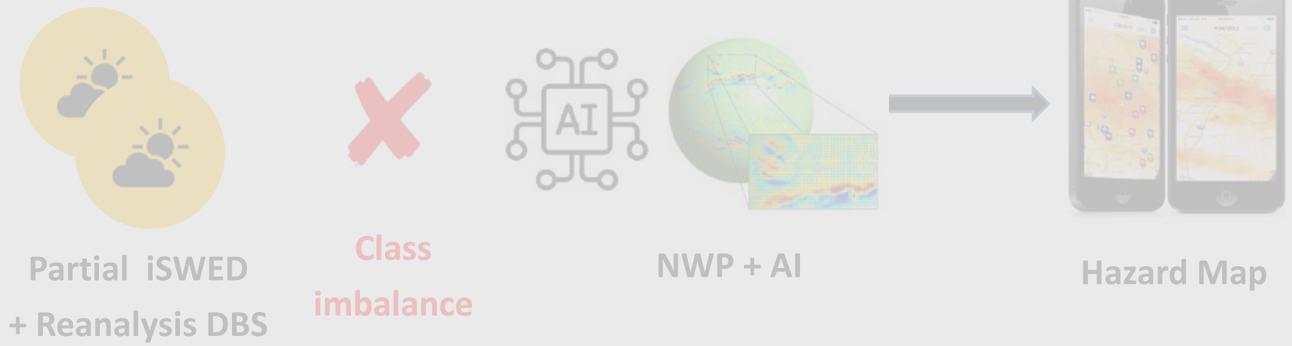
Insuficient iSWED
+ Reanalysis DBS

MLC

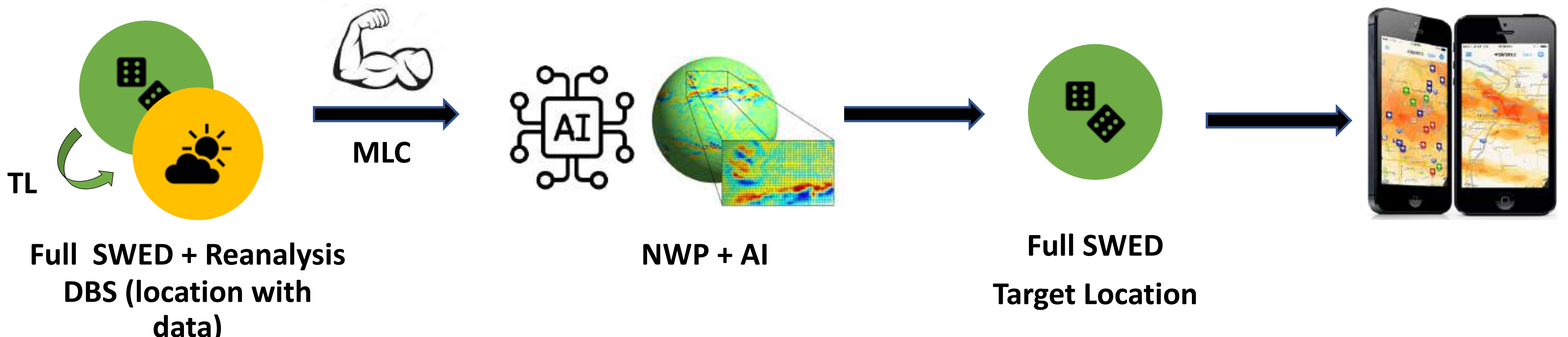


NWP + AI





- **The Role of AI:** Machine Learning Classifiers can identify the conditions that most likely lead to convective events, reducing the number of simulations and time required dramatically

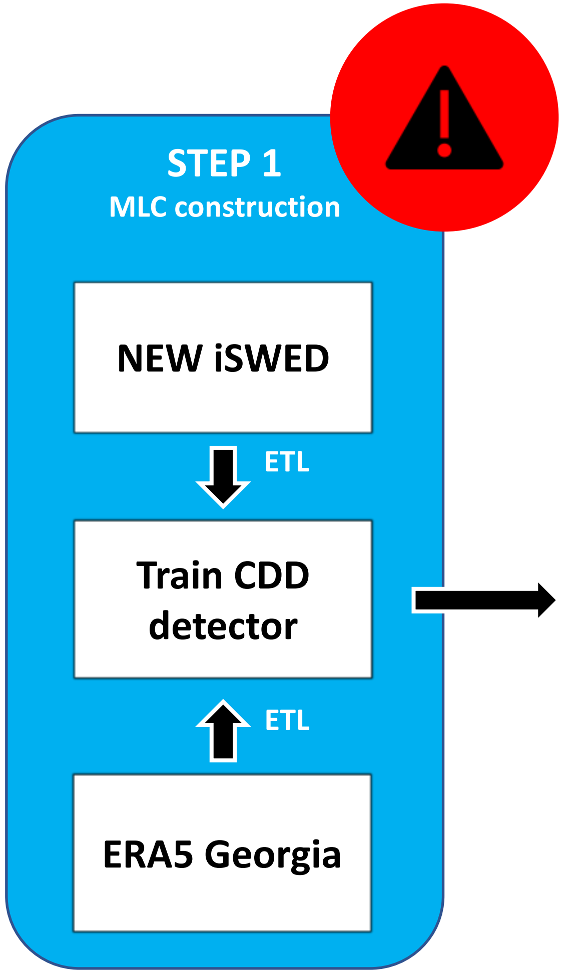


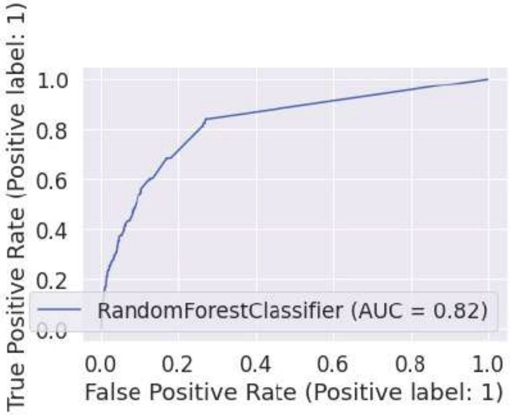
Unified methodology for hail and windstorm hazard mapping

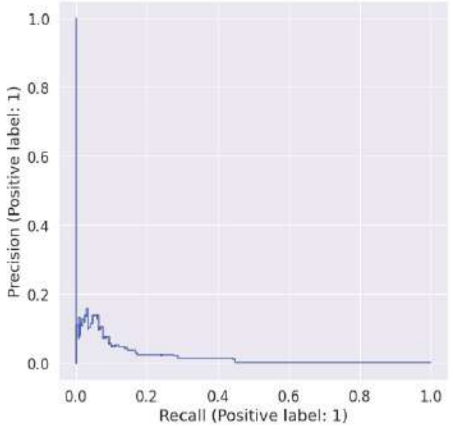
Georg-AI

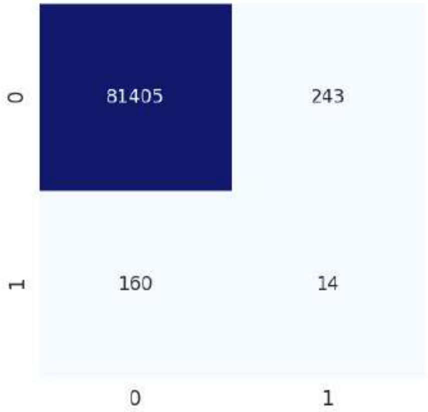
AI: Data Availability

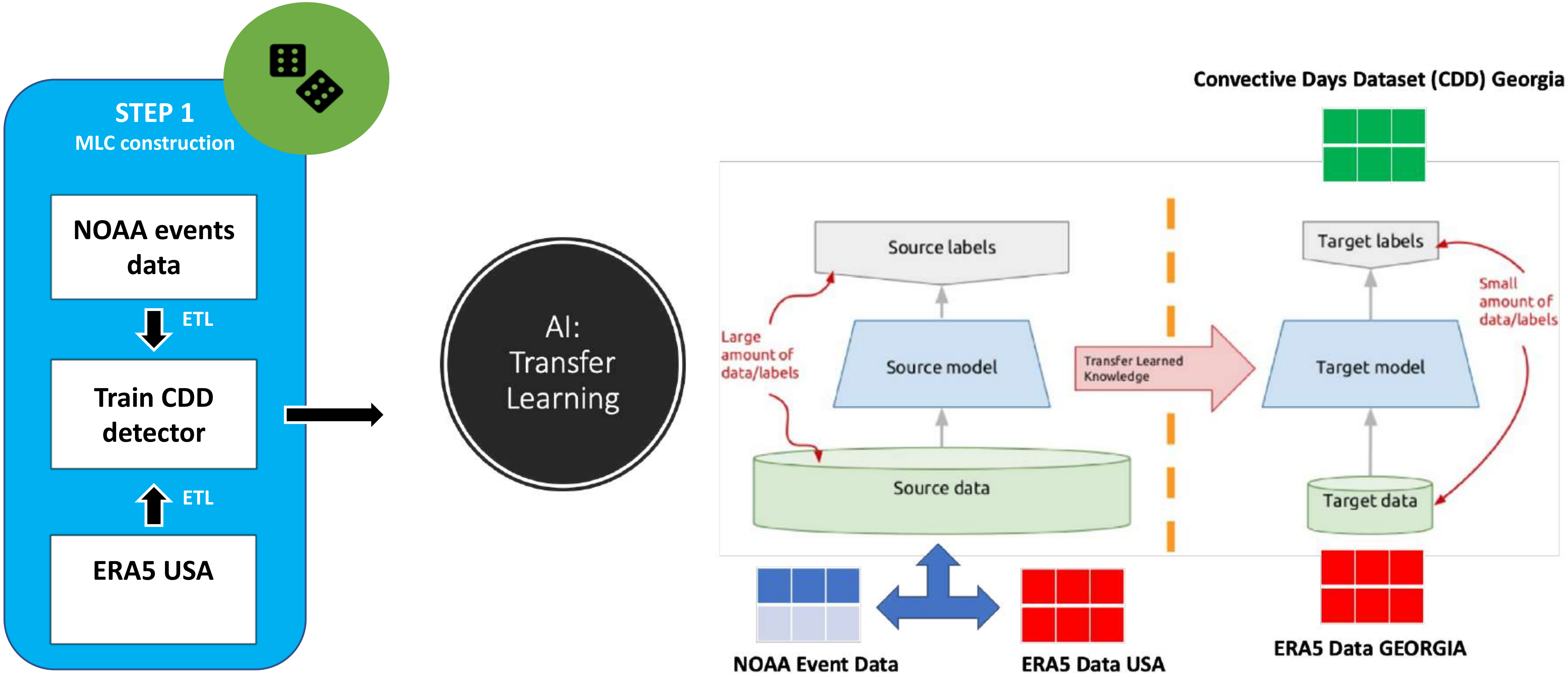
- Georgian National Environmental Agency's iSWED proved insufficient (~800 records)
  - Incomplete geographically and temporally
  - No or limited intensity information
  - Advanced monitoring network in its infancy

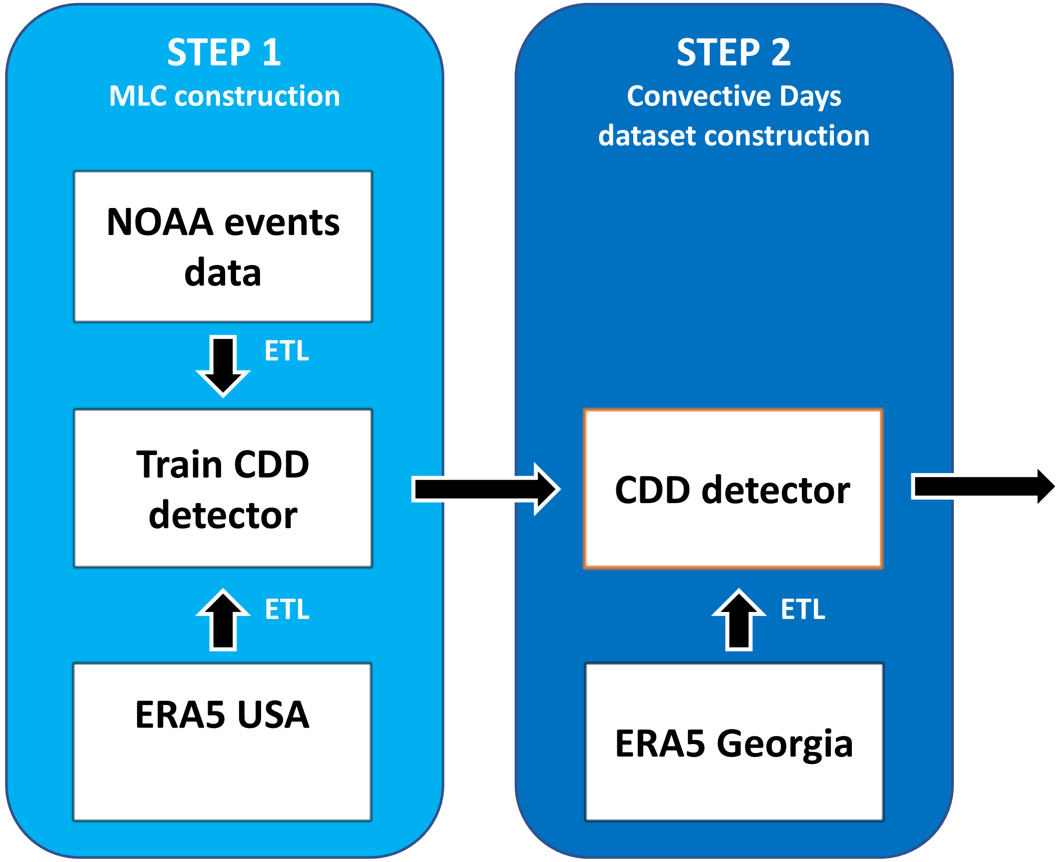

STEP 1
MLC construction
NEW iSWED
ETL
Train CDD detector
ETL
ERA5 Georgia

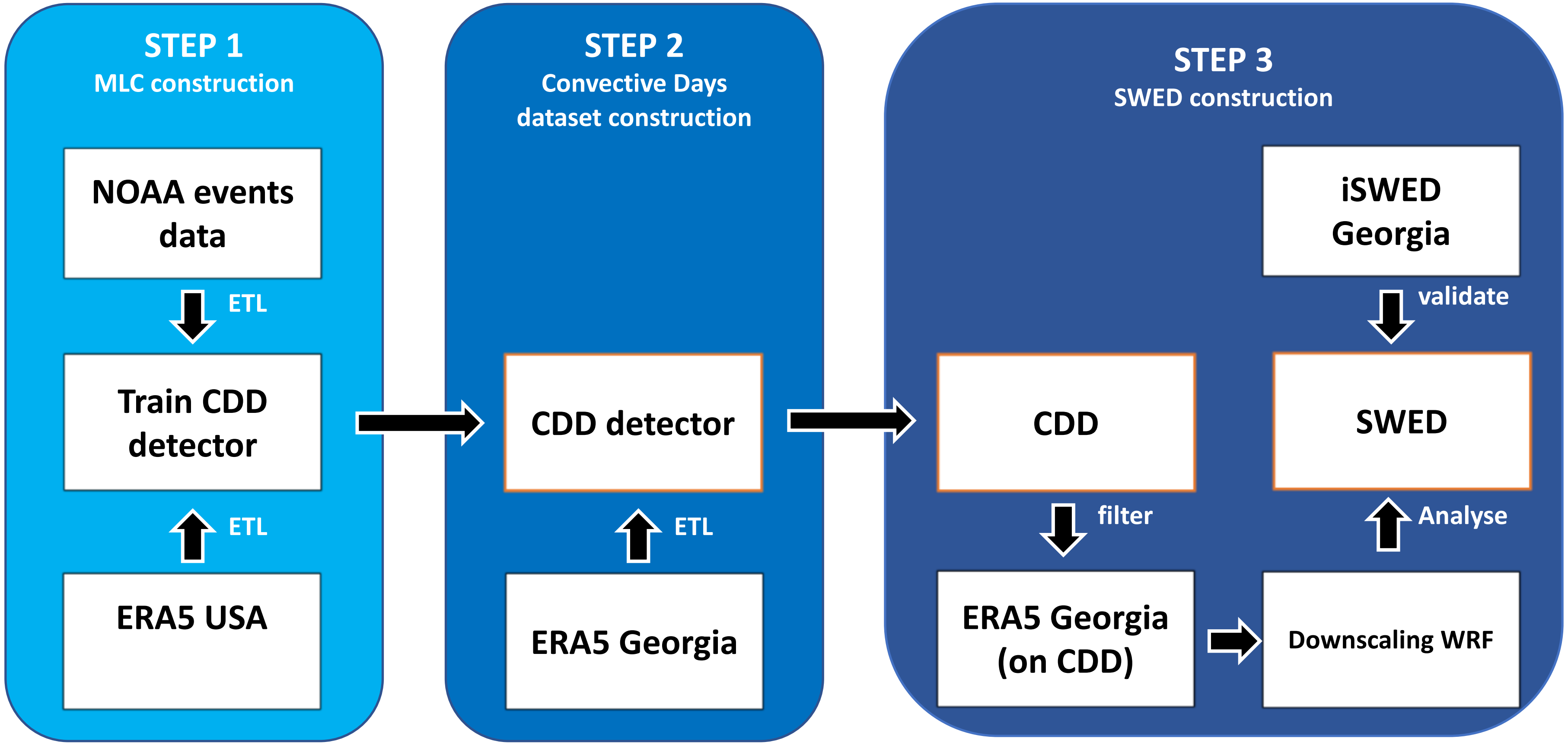

True Positive Rate (Positive label: 1)
1.0
0.8
0.6
0.4
0.2
0.0
RandomForestClassifier (AUC = 0.82)
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate (Positive label: 1)

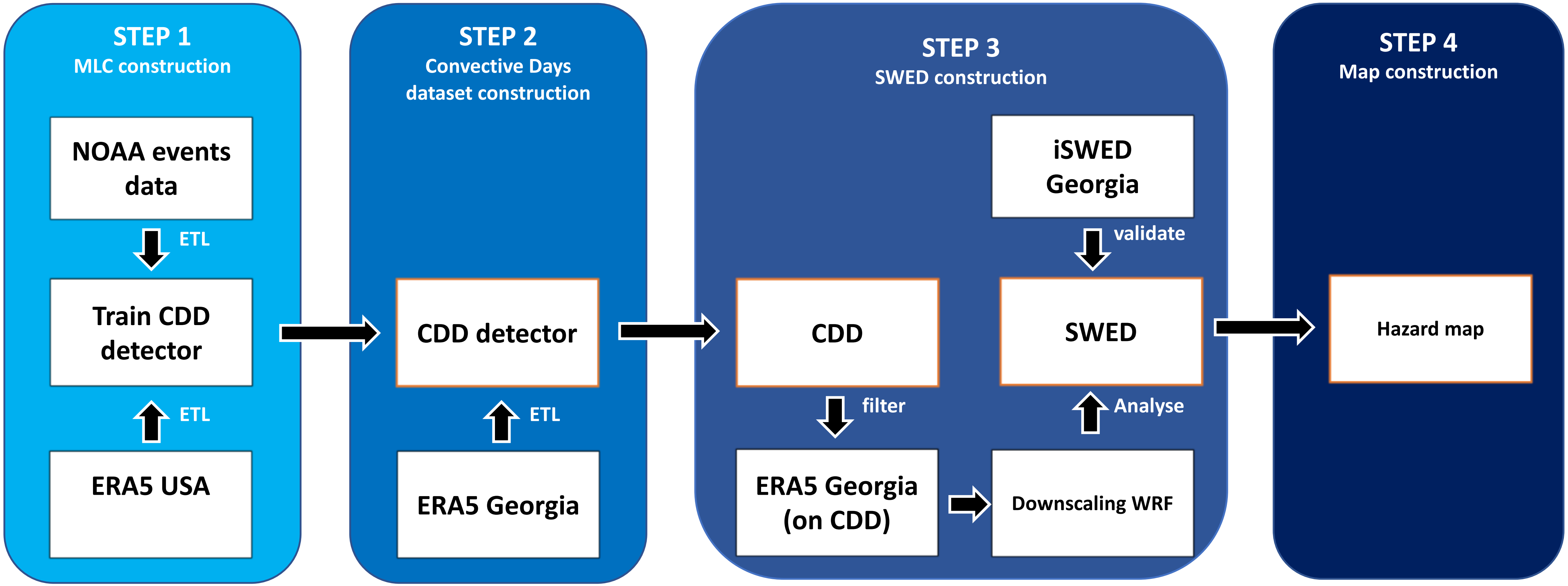

Precision (Positive label: 1)
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Recall (Positive label: 1)


0
1
81405
243
160
14
0
1


STEP 1
MLC construction
NOAA events data
ETL
Train CDD detector
ETL
ERA5 USA
AI: Transfer Learning
Convective Days Dataset (CDD) Georgia
Source labels
Large amount of data/labels
Source model
Source data
Transfer Learned Knowledge
Target labels
Small amount of data/labels
Target model
Target data
NOAA Event Data
ERA5 Data USA
ERA5 Data GEORGIA

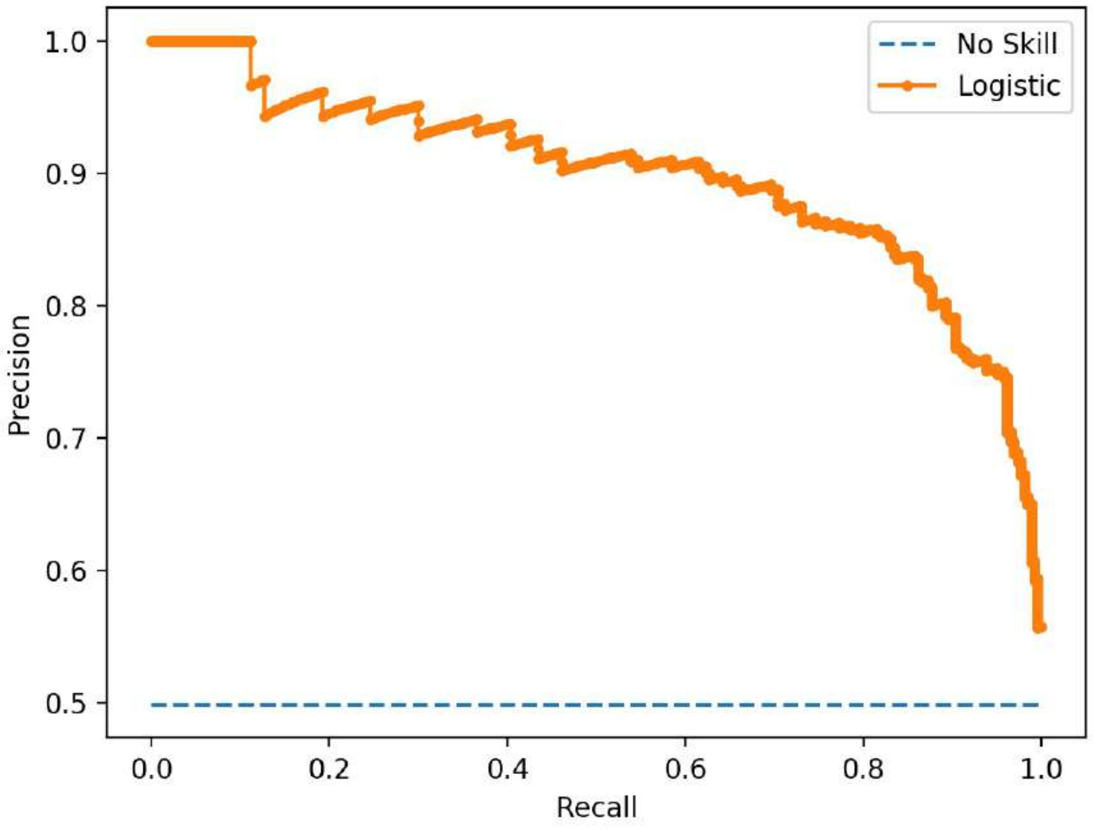

STEP 1
MLC construction
NOAA events data
ETL
Train CDD detector
ETL
ERA5 USA
STEP 2
Convective Days dataset construction
CDD detector
ETL
ERA5 Georgia


STEP 1
MLC construction
NOAA events data
ETL
Train CDD detector
ETL
ERA5 USA
STEP 2
Convective Days dataset construction
CDD detector
ETL
ERA5 Georgia
STEP 3
SWED construction
iSWED Georgia
validate
CDD
SWED
filter
Analyse
ERA5 Georgia (on CDD)
Downscaling WRF


STEP 1
MLC construction
NOAA events data
ETL
Train CDD detector
ETL
ERA5 USA
STEP 2
Convective Days dataset construction
CDD detector
ETL
ERA5 Georgia
STEP 3
SWED construction
iSWED Georgia
validate
CDD
SWED
filter
Analyse
ERA5 Georgia (on CDD)
Downscaling WRF
STEP 4
Map construction
Hazard map

AI: Expected result

- Perfrormant Machine Learning Classification with algorithm with a good precision-recall trade-off
- High recall particularly important (capturing as many of the true convective events as possible)



Mitiga

Thank you

Dr. Alejandro Martí Alejandro.marti@mitigasolutions.com