- CHAIRS: Rochelle Schneider , Alan Geer, Alessandro Sebastianelli
- 16 in person, 74 virtual
- Organisations: Academia, Operational NWP/Forecasting, Industry, Space agencies

1. CURRENT ML APPLICATIONS IN THE THEMATIC AREA:
   a) Solar Energy/Radiation Forecasting
   b) Wildfires Detection
   c) Snow/Flooding Detection, Depth and Extent Estimation
   d) Precipitation Retrievals
   e) Environmental Health
   f) Pollution
   g) Oil spills
   h) Oceanography/Chlorophyll detection
   i) Crops classification
   j) Urbanization
   k) Geodesy
   l) Clouds detection/motion modelling

1. CURRENT ML TOOLS AND APPROACHES IN THE THEMATIC AREA:
   a) Segmentation
   b) Classification/Detection
   c) Forecasting
   d) Regression
   e) Domain Translation
   f) Super Resolution/Downscaling

2. LIMITATIONS, CHALLENGES AND OPPORTUNITIES:

a) Reproducibility → Scalable → Operational → Maintainable → Transferrable → Explainable
b) Continuous learning / retraining
c) Usability for society
d) Physics aware ML → prior physical knowledge + constraints
e) Interoperability and integration with existing non ML tools
f) Lack of training data
    a) citizen science + gamification (non-expert can label, e.g. zooniverse.org)
    b) Domain expert required for labelling
    c) Ground truth observation and generalization (e.g. from Europe to New Zealand)
g) supervised and semi-supervised
    a) Few shot learning (zebra)
    b) Meta-learning (learning how to learn)
    c) Non domain foundation models (e.g. cats and dogs VS crop types)

3. ADVANTAGES (DISADVANTAGES?) OF ML TECHNIQUES FROM TRADITIONAL STATISTICAL METHODS:
   a) Fill the gaps around existing classical models
   b) ML for diagnostics / evaluation / understanding outside of an existing framework
   c) Using ML to do things that existing systems can't do
      a) Replacing part of the process done by humans
      b) Complex non linear correlations
      c) Finding complex patterns in large datasets
   d) NLP for mining/indexing unstructured data
   e) Sector-specific transformative applications: forecasting future losses and gains in solar energy production → reducing CO_2 emissions
   f) Speed up in model performance and outcomes (training is slow, prediction is fast)

4. FUTURE DIRECTIONS

a) Federated learning
    a) data protection issues (e.g. health data)
b) AI on board satellites
    a) Events detection (smart satellites)
    b) Data compression / prioritization
    c) Cooperating / connected satellites
    d) Make sure we still keep the raw data for future learning
c) Smart sensors → connected environment → smart observing systems / grids
d) Emergent intelligence → aggregation of more limited systems / agents
e) Embodied intelligence → proactive human-like learning
f) Data fusion → connecting diverse sources of data

5. EXTRA DISCUSSION – CLOUD AI BASED PLATFORMS

a) More expensive than in house computing resources (re-train costs double)

b) On the other hand TPUs and GPUs speed a lot the processing

c) Multiple users interaction

d) Open online services (e.g. google earth engine) with private backends

e) Lock-in by the “open online services”

f) Rapid scale up

g) Lack of support

h) Data protection and security