ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Observations & Measurements In the Possible Futures

Or How to Solve the Irritating Time Problem for Forecast Data

Ilkka Rinne

Finnish Meteorological Institute (FMI)

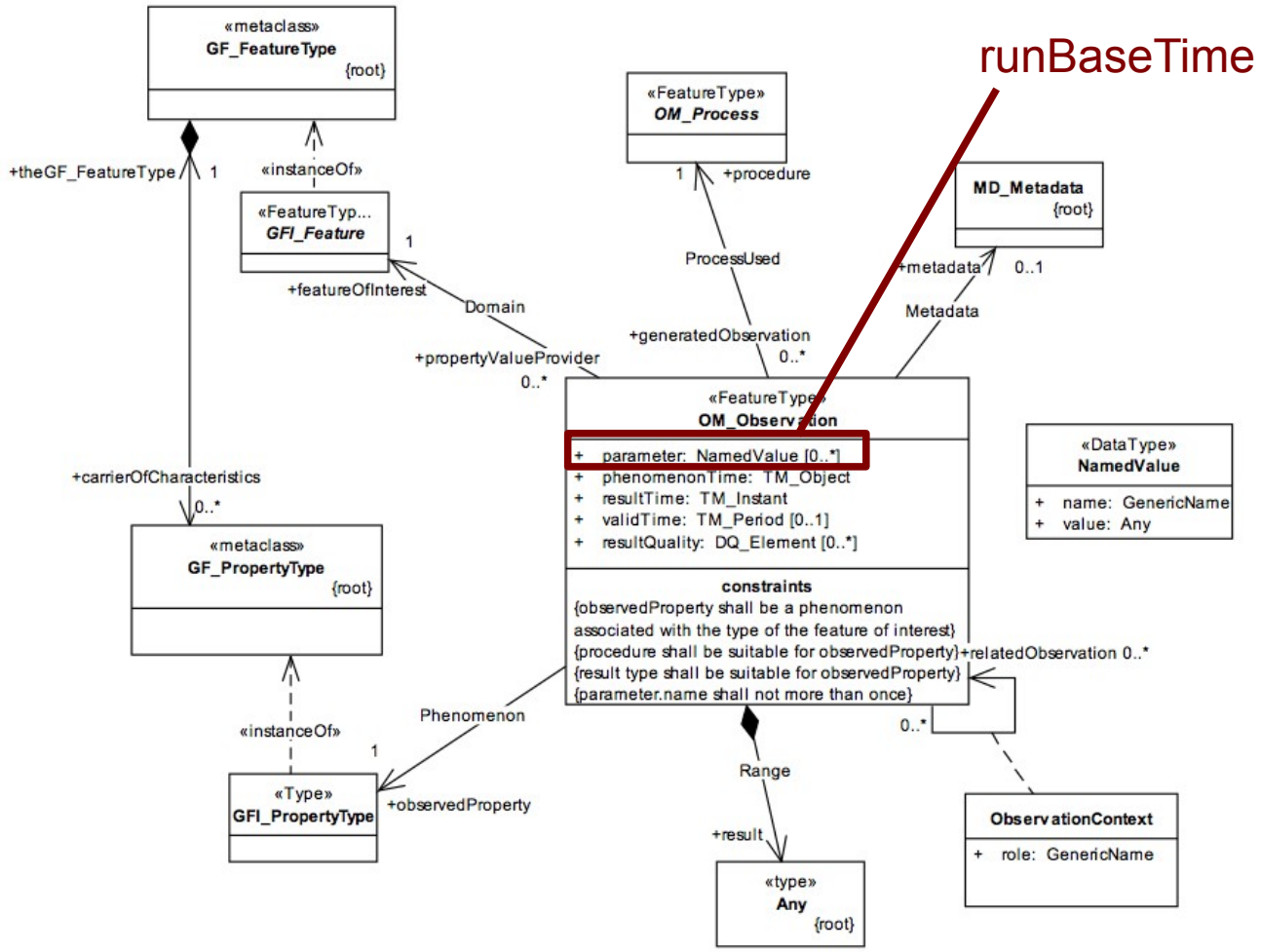
OGC TC/PC Meetings, MetOcean DWG

2nd March 2011, Bonn

O&M with Predicted Data – A Challenge

- **We have several relevant, but different data set types for overlapping phenomenonTimes**
  - Meteorological (station) observations, forecast model results from different models and different model runs, ensemble forecasts with several members,...
- **Would be nice if we could define a single conceptual model framework for all of them following the O&M model.**
  - Work-in-progress both in the MetOcean DWG Conceptual Modeling group and the INSPIRE TWG Atmospheric Conditions & Met. features (AC-MF).
  - Progress has somewhat been slow, because the issue is not a simple one solve elegantly.

RunBaseTime as property of OM_Observation



RunBaseTime as property of OM_Observation

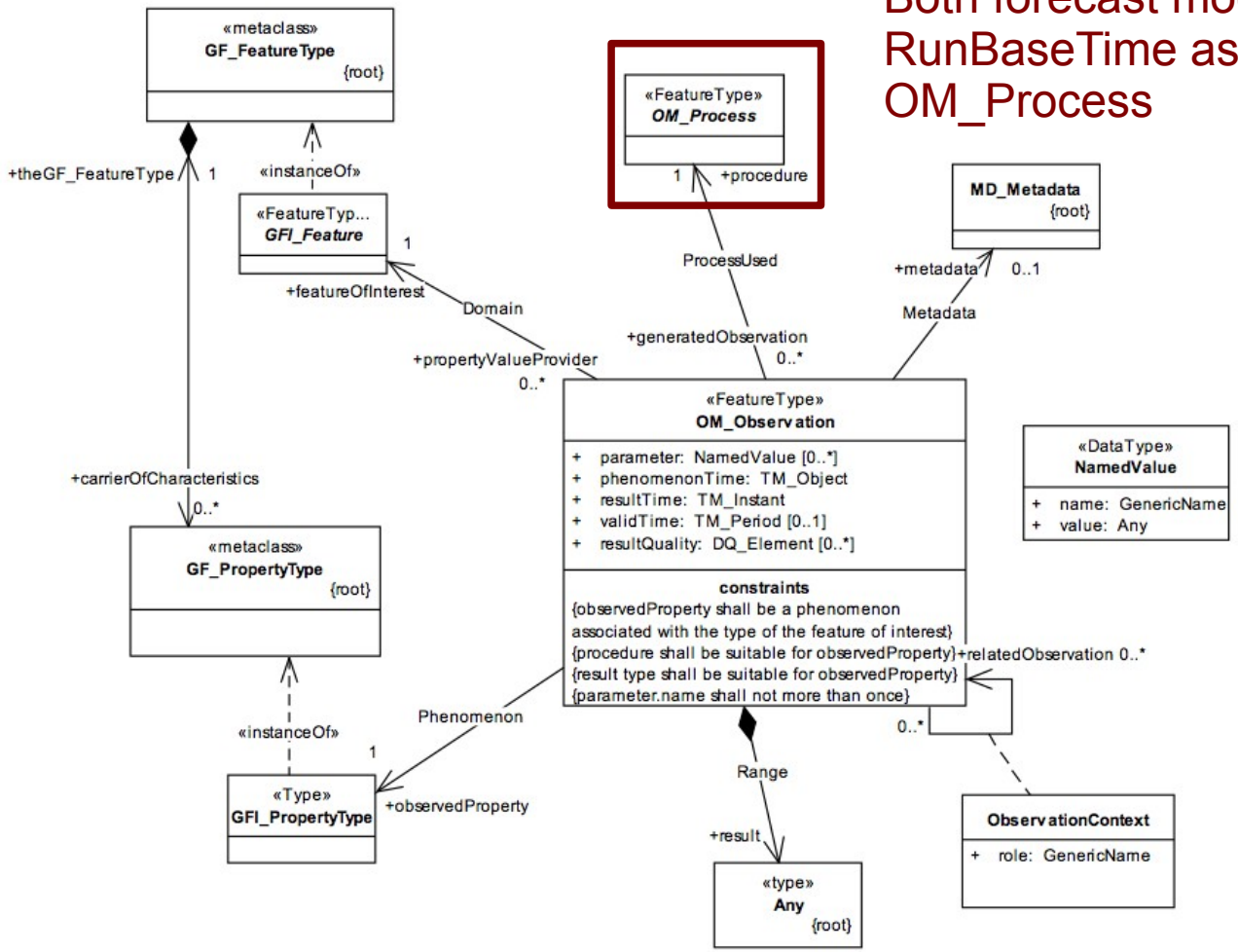
"If present, the attributes *parameter:NamedValue* shall describe an arbitrary **event-specific** parameter. This might be an environmental parameter, an instrument setting or input, or an event-specific sampling parameter that is not tightly bound to either the feature-of-interest (6.2.2.7) or to the observation procedure (6.2.2.10)"

Geographic Information: Observations and Measurements OGC Abstract Specification Topic 20 (OGC 10-004r3 and ISO 19156), page 11

- **Standard mentions "samplingDepth" of water well quality observations as an example of using this parameter.**
- **Could be used together with OM_Process (the model metadata) to define both the model and the particular run.**
- **But wouldn't it be better to use this kind of parameter for secondary properties, like the "size of the used subsampling grid" for example?**

Forecast Model specified in OM_Process

Both forecast model and the RunBaseTime as properties of OM_Process



Forecast Model specified in OM_Process

"The purpose of an observation process is to generate an observation result. An instance of OM_Process is often an instrument or sensor, but may be a human observer, **a simulator**, or a process or algorithm applied to more primitive results used as inputs."

Geographic Information: Observations and Measurements OGC Abstract Specification Topic 20 (OGC 10-004r3 and ISO 19156), page 14

- **Seems like the closest thing to what we need in the O&M 2.0 standard's text: OM_Process may be "a simulator"**
- **But does the forecast model run really "generate an observation result" describing the properties of a future state of the atmosphere?**
  - It's a prediction that *might be* close or not so close to what can later be measured in the "real world"

A Possible Future: an observable state of a Prediction Universe

Making Observations in Possible Futures

Flickr CC licensed content by user “Sacred Destinations”

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Making Observations in Possible Futures

Flickr CC licensed content by user “Sacred Destinations”

Making Observations in Possible Futures

Flickr CC licensed content by user “Sacred Destinations”

Making Observations in Possible Futures

Flickr CC licensed content by user “Sacred Destinations”

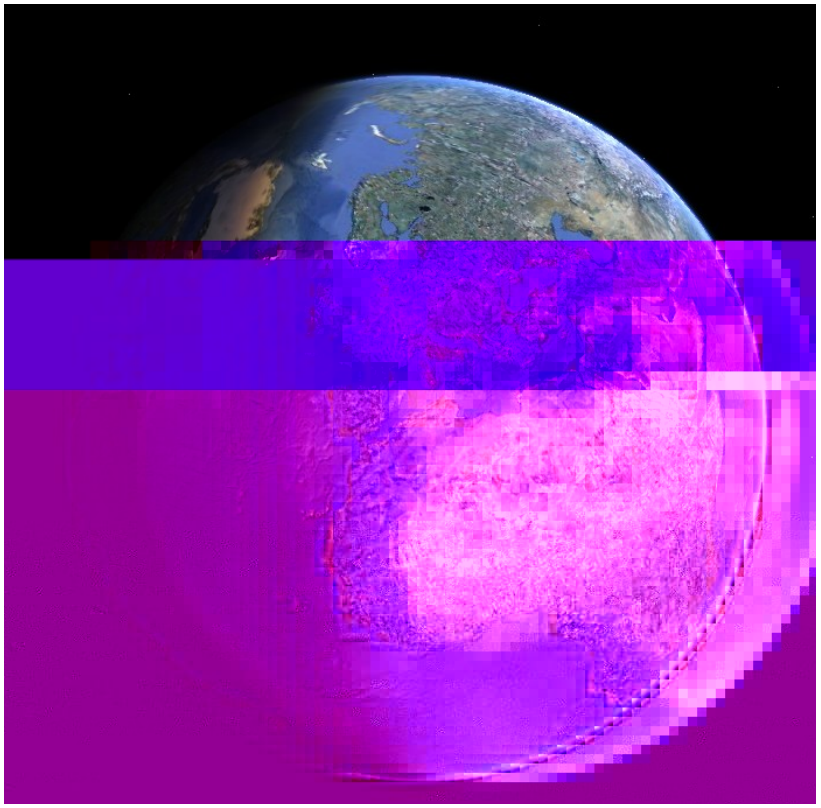
Making Observations in Possible Futures

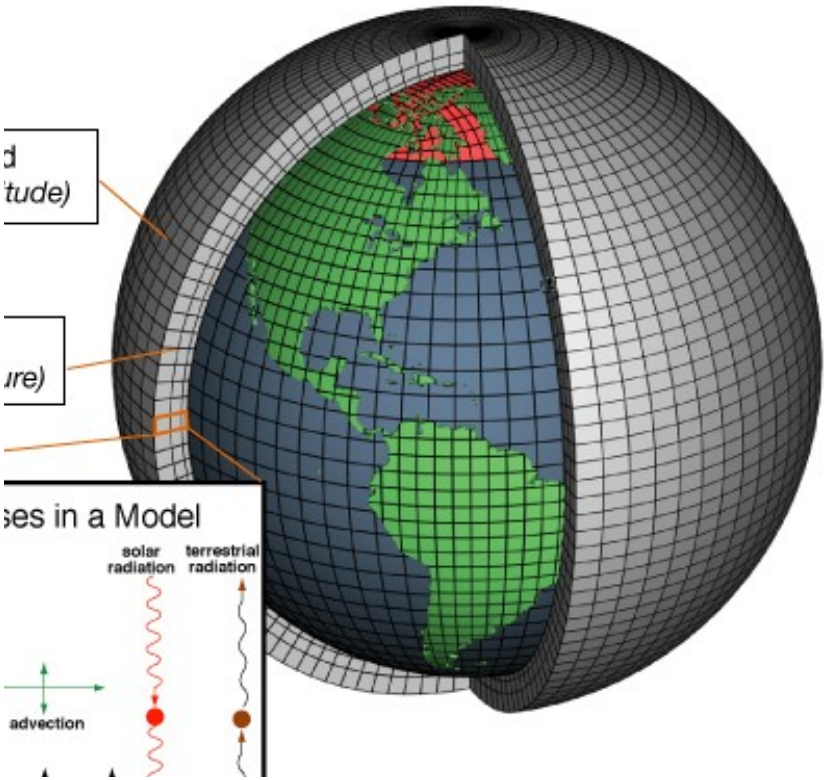
Flickr CC licensed content by user “Sacred Destinations”

Fantastic!

(but Observable)

Making Observations in Possible Futures

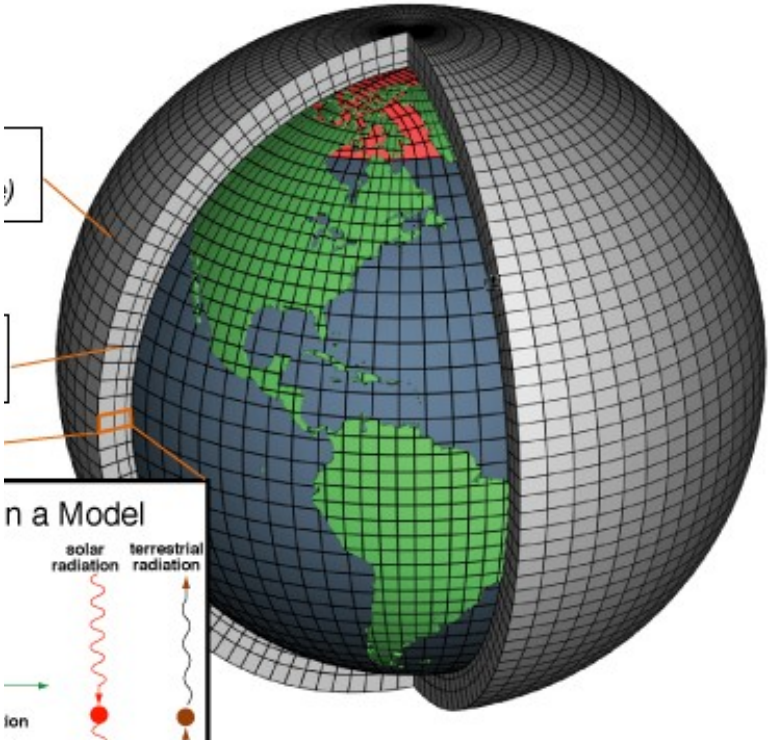




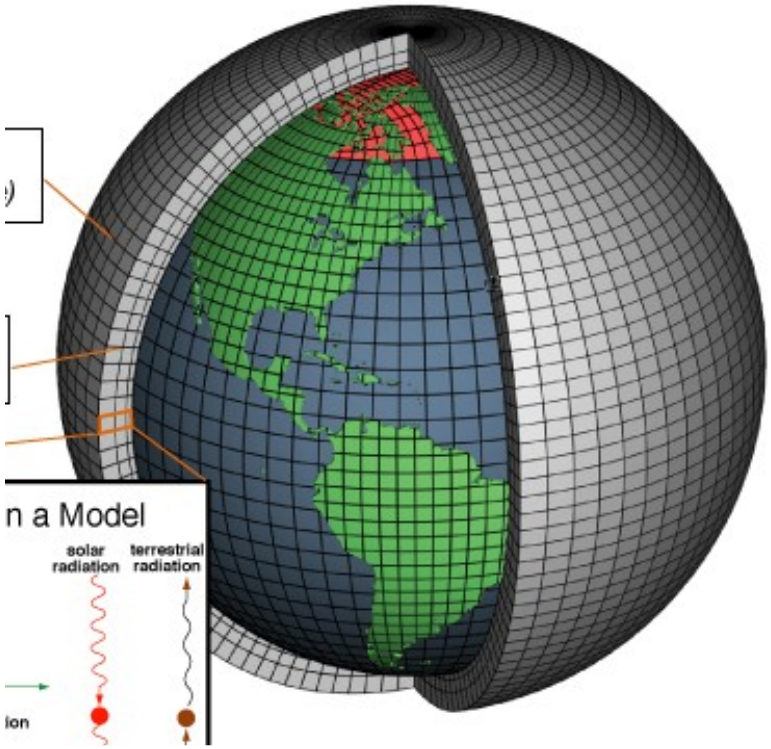
A forecast model run result is **not** a foreseen state of the real world. We are not observing the future (we cannot do that, unfortunately), **we are observing the results of a simulation**.

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Making Observations in Possible Futures



runBaseTime = 2011-02-25 12:00Z
phenomenonTime = 2011-02-26 00:00Z



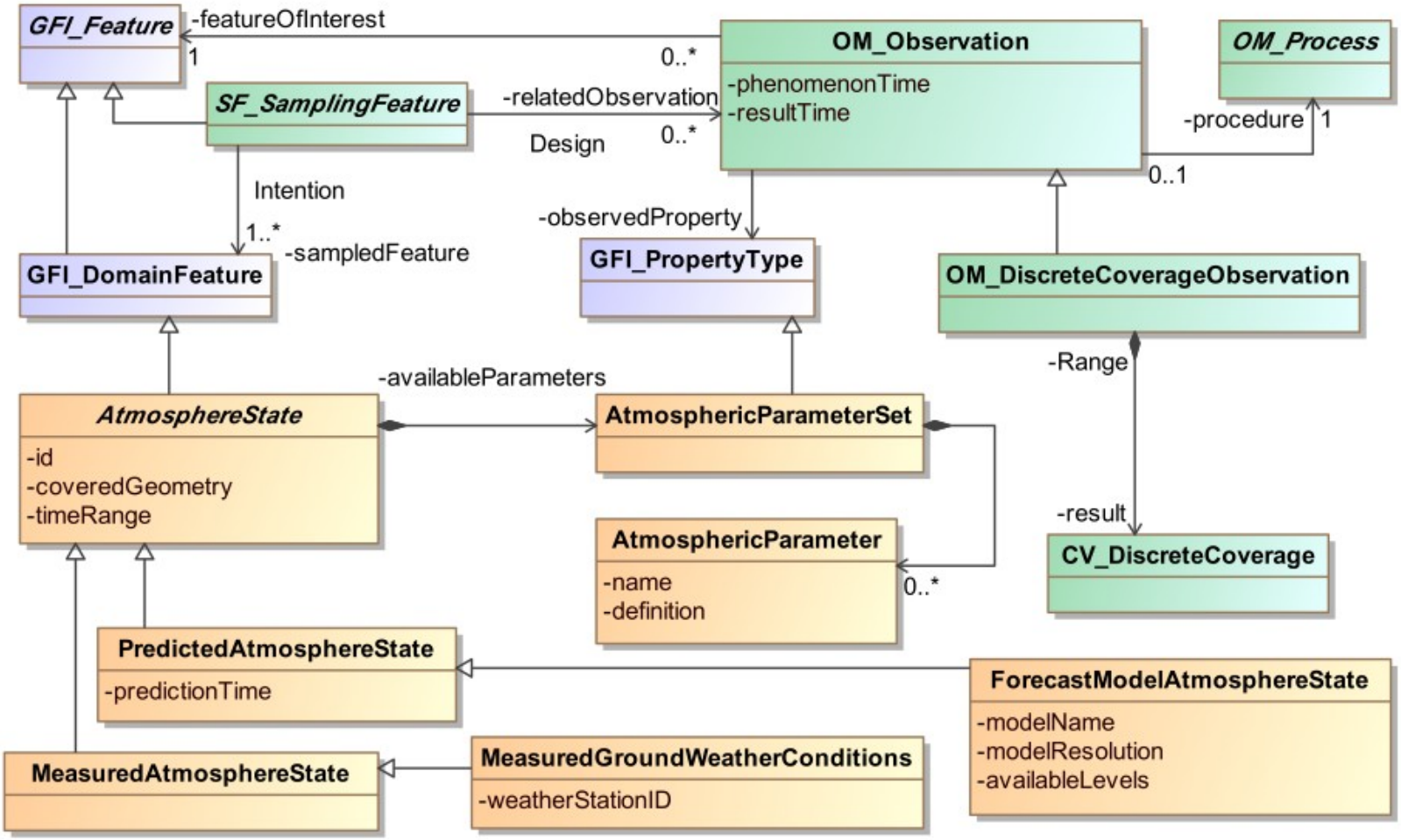
runBaseTime = 2011-02-25 18:00Z
phenomenonTime = 2011-02-26 00:00Z

Two simulated worlds, two different sets of Observations for (partly) the same phenomenonTimes, two different sets of Observation results. **If the feature-of-interest is the same, this is a Data Quality problem** (at least the other is “wrong”). But which one?

Different features-of-interest for model runs?

- **Different values for observableProperties at the same phenomenonTimes are OK, because the Observation targets are in different "universes" defined by their domain features.**
- **Each model run *defines* it's simulation universe (a possible future)**
  - Limited by the model's spatial boundaries and by the forecast time range
  - Observable properties (atmospheric parameters) only have defined values at discrete vertical levels of provided by the model.
  - A set of OM_Observations for each time step
    - Possibly through via SamplingFeatures (like simulated soundings)

AtmosphereState as Met DomainFeature



Benefits

- **The semantic contradiction would be solved: Observations of model run results no longer claim to foresee the actual future state of the atmosphere**
- **Data Quality would be defined by how well these Observation results match the original result data of a model run (possibly affected by resampling, subsampling, re-projections etc.)**
  - Not affected by how close to the reality the values end up being (important, but can only be evaluated when the forecast time passes).
  - Ensemble forecast members' different Observation result values would not mean their Data Quality is not good.

Benefits (2)

- **OM_Procedure would be reserved only for information about the data extraction, post-processing, resampling, re-projection etc.**
  - Same as with “real world” measurements, SensorML directly usable both for simulated prediction data and “happened & measured” data.
- **Exactly the same data structures for observation events and the corresponding results could be used for all the data (forecast is not a special case)**
  - Numerical weather models, ground observation network, weather radar, satellite imaging,...
- **The different AtmosphereState instances would be directly usable as Data Set Catalog entries.**

Benefits (3)

- **Since the domain features of different model runs would be different, it would only be natural to serve Observations for each run from a different service end-point**
  - A WMS / WCS / WFS / SOS end-point for a particular model run would only serve (observation) results originating from that model run.
    - Using and defining Capabilities documents would be much easier, need for one TIME parameter only
    - This model already implemented by UCAR / THREDDS
  - Discovery of these service end-points by using a Data Set Catalog containing the “universe defining” AtmosphereState domain feature instances.

Data Comparison by Forecasters

- **Discovery of available data sets ("real world" measurements, forecast model runs) for particular forecasting time range and areas is done by spatio-temporal queries on the AtmosphericState instances:**
  - "Find me all the numerical weather model runs currently available covering this region and at least 2 days forward, and there must be at least temperature at 2m, ground pressure, and wind speed parameters available. Also the horizontal model resolution must be at less than 10 km."
  - "Found 6 runs of 2 different models, Hirlam & ECMWF deterministic. Here are the addresses for WMS and WCS services for each of them."

Data Discovery Across Application Domains

- **Each application domain has it's own DomainFeatures and related Observations**
  - A road intersection for transport administration, a runway for air traffic management, why not state of the atmosphere for meteorology?
  - Cross-domain data discovery is again done by spatio-temporal queries on domain features:
    - “Give me AtmosphereState instances that contain ground temperature values, and cover this RoadIntersection and the following 2 hours.”
    - “We have values from a coarse global forecast model, a fine-grained regional forecast model and recent observations road weather station network. Which ones would you like to retrieve?”

"A forecast model is an instrument for *measuring the future state* of the atmosphere"

- **Yes**, some models just yield results of better quality than others in some situations. There is no way to know for sure which is best in which situation before the forecast time passes, however.
  - The feature-of-interest should be same for all forecast model runs and real world measurements.
- **No**, a model run is just a simulation of an atmospheric model with a given set of preliminary conditions. The truth value evaluation (how close the predictions are to the actual "happened" state of the real world for certain atmospheric parameters) should be a separate concern from the data quality of a forecast data set.
  - It's natural to have a different feature-of-interest for each simulated world.