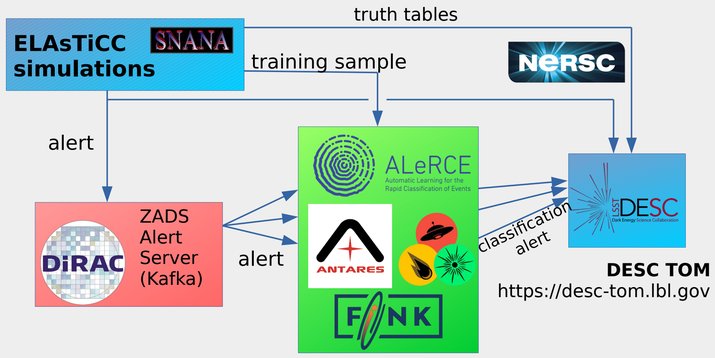
ELAsTiCC alert production and processing for LSST/DESC

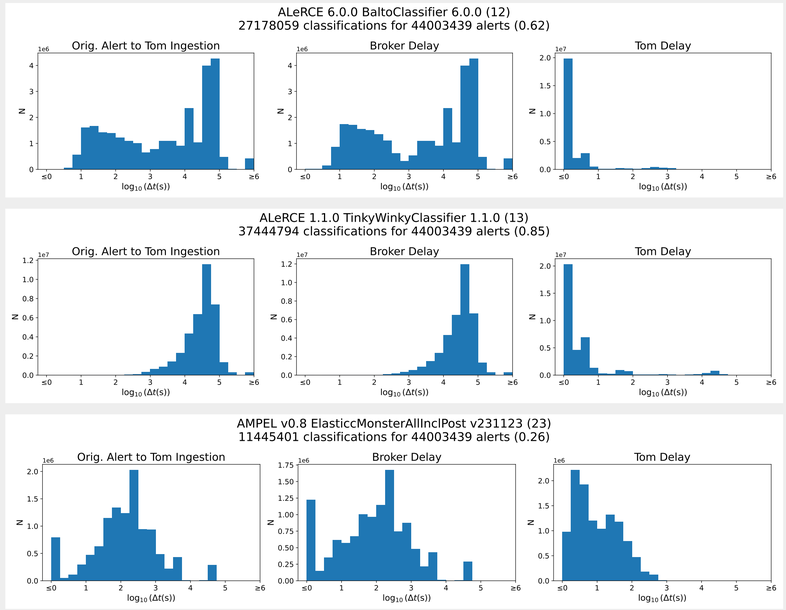
R.A.Knop (LBNL) and the LSST/DESC ELAsTiCC Team

ELAsTiCC Data Flow



For ELAsTiCC2 (Nov-Dec 2023), we simulated lightcurves for 4.1 million transient and variable objects, sampling those lightcurves using the LSST baseline 3.2 cadence over three years, yielding 62 million detections and 990 million forced photometry points. An alert is sent for each detection to the ZADS server. Brokers ingest those alerts, and through different mechanisms (kafka, Google pub-sub, and direct API posting) provide object classifications to DESC. DESC collects all of the data (both the original data and truth tables, and broker classifications) on the DESC TOM, a django-based server using the TOM Toolkit (https://lco.global/tomtoolkit/) with a PostgreSQL database backend. The TOM and associated processes run on the NERSC Spin kubernetes service.

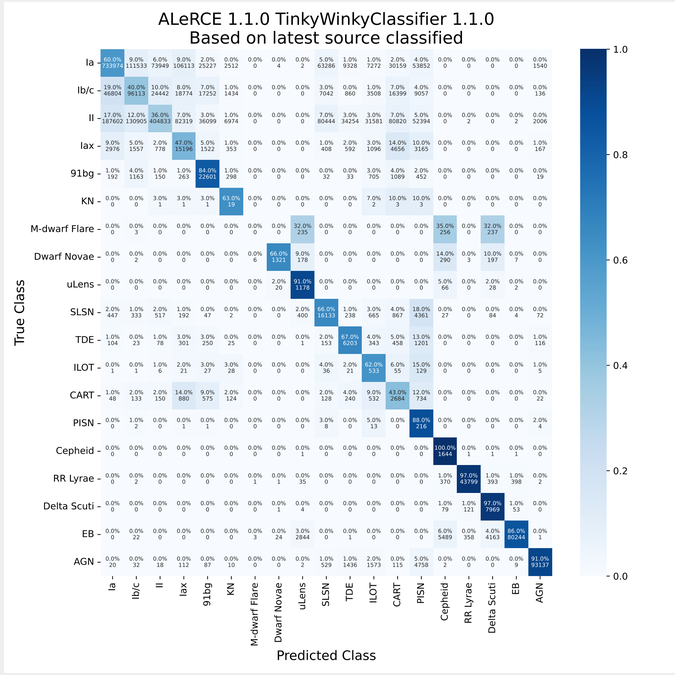
Receiving Broker Classifications



Left: Total time between when the original alert is sent out and when DESC has ingested classifications from a given broker's classifier. **Middle:** The time from the original alert to the broker publishing the classification. **Right**: The time between the broker publishing the classification and DESC ingesting it.

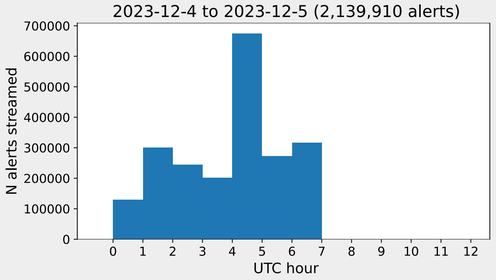
All brokers have demonstrated the ability to classify objects in less than one day (10^5 s), and often in less than 3 hours (10^4 s).

Classifier Metrics



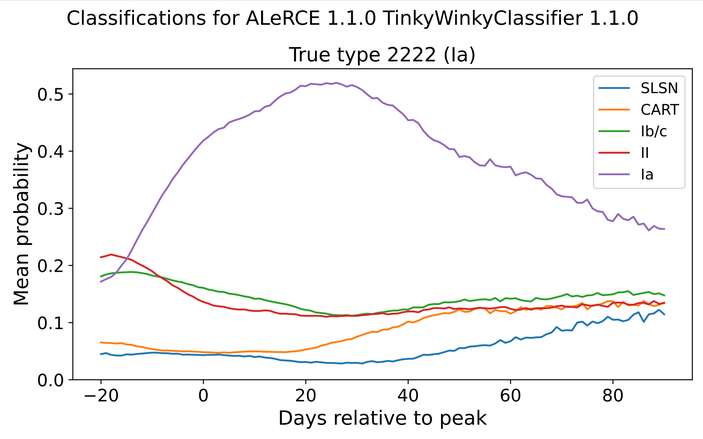
Confusion matrix from ELAsTiCC2 for one classifier from one broker, normalized by true type. In this case, the predicted type is the type that had the highest probability in a classification of the latest detection for a given object. This does not capture all of the data from that classifier, however, as classifications were sent for each detection of each object (see below).

Sending Alerts



Each day starting at 0h UTC, 36 simulated days of transient detection alerts are sent from a process running on NERSC Spin to the ZADS Kafka server at UW. Alerts go out at a rate of ~100-150 per second, which is within a factor of ~2 of what we expect for eventual LSST alerts.

***ELAsTiCC= "Extended LSST Astronomical Time Series Classification Challenge"**



Because each object is classified multiple times as additional detections become available, there is far more information than is captured in a confusion matrix. This plot shows the five highest-probability types reported by one classifier, averaged over all events whose true type was SNIa. The relative confidence the classifier has that these events are SNeIa changes as more information comes in; the highest confidence is 20-30 days after peak. Later in the lightcurve, when more data is available, the classifier still favors SNIa, but assigns less overall probability to that class.