

QuickLook and QuickLook Framework Readiness Review

January 2019

Thank you for agreeing to review the QuickLook (QL) algorithms and the QuickLook Framework (QLF) web interface. Using the links below, please try using QL/F to evaluate the data quality of simulated exposures and provide feedback on its readiness for commissioning and survey operations. We will use this report to prioritize and plan the development work for 2019. This review will not be live presentations and demos, but rather a hands on "try it out and give comments" exercise. You are welcome to ask questions to the QL/F teams via desi-data@desi.lbl.gov.

Review Committee

- Connie Rockosi (Commissioning Lead)
- Kyle Dawon (Survey Validation co-Lead)
- Julien Guy (DESI Project Scientist)
- David Schlegel (DESI Project Scientist)

Others may optionally provide feedback, e.g. Klaus Honscheid (ICS, AIT&C Scientist), Paul Martini (DESI Instrument Scientist), Stephen Bailey (DESI Data Systems Manager), and Christophe Yeche (Survey Validation co-Lead).

QL/F Inputs to the Review Committee

Installed deployment of QL + QLF at both KPNO and NERSC

- KPNO: <http://desi-1.kpno.noao.edu:9000> (requires VPN)
- NERSC: <http://web.qlf.sandbox.stable.spin.nersc.org:60049/>
(no VPN or password required)

Caveat: when you view this, you will see that all exposures have a warning, which comes from a bug in the CCD amp bias offsets in the input simulations. These bias levels are jumping around from exposure to exposure and QL flags that as a warning. This bug in the input simulations is my (Stephen's) fault.

One short night (20191001) of data has already been processed with that version of QL/F. The raw QL file outputs are available at:

- KPNO: `/data/quicklook/exposures/redux/exposures/20191001/`

- NERSC /project/projectdirs/desi/datachallenge/quicklook/review-19.1/redux/ (transfer from KPNO still in-progress)

As a separate robustness validation test, we will also stage new data that QL/F has not yet seen to evaluate the outputs (Stephen will help with this).

Documentation Links

- QL Documentation: <https://desi.lbl.gov/trac/wiki/Pipeline/QuickLook>
- QLF Documentation: <https://github.com/desihub/qlf/wiki/QLF-Overview>
- QL Code: <https://github.com/desihub/desispec> (under py/desispec/quicklook/)
- QLF Code: <https://github.com/desihub/qlf>

Charge Questions to the Committee

Please use the provided version of QL/F to assess the (simulated) data quality and provide feedback on the following questions.

QuickLook (QL) Algorithms and Outputs

- Do the QL outputs provide sufficient information to assess the DESI data quality?
- Are the QL outputs documented clearly enough to be useful to end-users? This includes both what is measured and the data formats of files on disk.
- Are the QL algorithms correct?
- Is the QL code sufficiently flexible and written/documented clearly enough that new developers could reasonably contribute new metrics to be included in the outputs or modify existing metrics?

Quicklook Framework (QLF)

5. Does the QLF display all of the QL metrics?
6. Is the QLF interface clear and user friendly?
7. Is the QLF code sufficiently flexible and written/documented clearly enough that new developers could reasonably contribute new plots and refinements to existing plots?

While answering those questions, please assess QL/F at several levels:

- Using QLF to view QL outputs that were previously processed by the QL/F teams.
- Assessing QL spectra and metrics independently of the QLF interface.
- Looking "under the hood" at the feasibility of new developers adding/modifying QL metrics, and adding/modifying QLF plots, e.g. to implement new ideas that come up during commissioning.

Response from the Review Committee

Overall Reviewer Comments

Findings:

- The quicklook software provides a tool for on-site observers to assess data quality
- The QLF interface has excellent quality visuals to view the outputs of quicklook
- Together, these two tools could potentially provide the necessary framework to connect observers to the data quality in real time
- The main webpage for QLF is dominated by placeholders for future development rather than functional features
- The documentation and implementation of the quicklook software is advanced but not complete and appears to have some inaccuracies in its quality assurance diagnostics

Comments:

- The communications between the QLF team, QL team, and Data Systems team have not been terribly effective over the past two years; blame is to be fairly distributed, including amongst those on the review panel
- Many good tools exist and could potentially be preserved
- The review panel focused on the visual displays and not the code itself, thus limiting feedback to the code developers. Questions remain in the sustainability of the code and the accuracy of the algorithms due to lack of review.
- The steps to completion of the QLF interface would require significant coordination between the QLF team, QL team, and Data Systems team to ensure consensus on the final design.
- qproc now runs fast enough that there is no need for intermediate quality assessments during the data reduction. This will have implications for the QLF interface which shows progress of intermediate steps and logging.
- It is likely that new custom routines will be needed to assess the qproc outputs during commissioning and early survey operations. The development of these routines needs to occur quickly and accurately.
- Quicklook/QLF are roughly in the state that was expected in 2017. The QLF development is not progressing at a sufficient rate to complete all of the features that are presented in the main display. Progress has been limited by turnover of personnel on the QLF team. The structure of the code and database will place additional limitations on new features.

Recommendations:

- While very good work has taken place with the QLF software, we recommend that the project pursue an alternate method to achieve visuals of the automated feedback.
- If the QLF team decides to continue development of the current code, we recommend a significant descope in the features based on time constraints, priority, options for independent tools, and limitations of the current framework.
- We recommend that another member of DESI contribute to the development of the qproc reductions and quality assurance measurements so that changes can be made quickly on the mountain during commissioning.
- We recommend that the data processing in quicklook (qproc) be fully separated from the QA evaluation.

While we do recommend an alternate software effort take place for both the visual interface and the quality assessment measurements, we appreciate the effort of both teams involved in the development of quicklook and the QA interface. If these two teams decide to continue the current effort, we provide below more detailed findings, comments, and recommendations for improvements to the QLF and quicklook quality assurance measurements. For each question posed to the review panel, we respond with a concise answer and follow with findings, comments, and recommendations.

Question: Do the QL outputs provide sufficient information to assess the DESI data quality?

Answer: Many important data quality assessments are included in the current version of quicklook, but a few are either incomplete or reported in a non-intuitive manner.

Findings:

- The reviewers independently worked through diagnostics of typical problems using the outputs of quicklook as a guide.
- An example of a problem is an out-of-focus spectrograph
- An example of a problem is misplaced fibers leading to reduced throughput
- An example of a problem is poorly behaving electronics
- An example of a problem is a stuck shutter
- Throughput information is encoded in DELTAMAG_TGT
- QA metrics for flat and arc frames are different from science frames
- The list of science QA metrics is close to adequate for now. More may be needed when real data arrives.

Comments:

- Some metrics are organized in a non-intuitive manner (e.g. XSIGMA and WSIGMA are a measure of the LSF and trace width but are presented under “check CCDs”).
- There is no clear evidence that the ETC estimates are given to quicklook

Recommendations:

- CHECKARC should provide a clear comparison of the measured wavelength solution to a reference wavelength solution
- CHECKFLAT should provide cross-section of fiber fluxes over a given wavelength interval as a function of fiberID. This will help to assess visually the light profile inside the spectrographs.
- CHECKFLAT should provide the ratio of the measured flux to a reference flux for some well-exposed region of the flat traces across each ccd
- The QA associated with the science exposures should provide a connection between the seeing, transparency, and throughput estimates of the ETC and the quicklook characterization of spectra.

Question: Are the QL outputs documented clearly enough to be useful to end-users? This includes both what is measured and the data formats of files on disk.

Answer: Documentation does cover much of the quicklook outputs but is at times incomplete or inaccurate. In addition, some reviewers had trouble finding the documentation in the maze of docDB.

Findings:

- The QL pipeline is described on the wiki page: <https://desi.lbl.gov/trac/wiki/Pipeline/QuickLook>
- The documentation of QL metrics for science exposures is found in docDB DESI-4701
- There is no obvious documentation of QL metrics associated with calibration exposures
- <https://github.com/desihub/qlf/wiki/QLF-Overview>, <https://desi.lbl.gov/trac/wiki/Pipeline/QuickLook>, DESI-3234 and DESI-1673 may not provide the correct information for understanding QL outputs
- The *structure* of the JSON hierarchy is documented in the "Example JSON output" section but the *meaning* of the fields is documented in a separate table.

Comments:

- The overview of the various QL metrics in DESI-4701 is quite clear for most science content.
- The short descriptions displayed in QLF are not sufficient explanations of the QA outputs
- NOISE AMP does not appear to be documented correctly and is possibly computed incorrectly. NOISE_AMP is described as the mean in the overscan but it should reflect the READNOISE. Documentation does not clearly reference readnoise as the diagnostic and would be easier to understand if the metric was named "READNOISE".
- LITEFRAC AMP is not documented correctly, appears to be the fraction of pixels above some threshold and not the number of pixels above threshold. Threshold needs to be clearly defined both in documentation and on the QLF display page. Consider renaming to COUNT PIXELS
- There are discrepancies between the names of the metrics in the quicklook documentation and the QLF descriptions. For example, NGOODFIB is also called COUNT SPECTRAL BINS on the QLF page.
- PEAKCOUNT/SKYPEAKS is not a clear metric because units are not described (photons per second?)
- Documentation of DELTAMAG_TGT is not clear. This should be computed using only std stars and it should be very clear about the bandpass that is being used, the conditions under which the reference calibration vector was calibrated, etc. This can be a very good metric to assess whether the throughput is inappropriate, but the current description fails to meet that standard. It also appears to be computed on all targets, but this may be an improper visual on QLF.

- FIDSNRTARGET is not well documented. It is not stated at magnitude the SN is assessed, whether photometry is corrected for MW dust, if photometry is based on the effective fiber magnitude, or what functional form is used to model the trend.
- The documentation mixes the order of the metadata fields (CAMERA, EXPID, PROGRAM, ...) with QA metrics (DIFF1SIG, DIFF2SIG, ...) with parameters about the thresholds on those metrics (NOISE_AMP_REF, PERCENTILES). Grouping metadata into its own section like METRICS and PARAMS would help.
- Some METRICS have corresponding PARAMS, while others do not, and some PARAMS (e.g. PERCENTILES) aren't clear about which metric they apply to (and aren't documented in the "Keyword Descriptions" table).

Recommendations:

- The documentation should include a description of the ARC, dark, bias, and FLAT quicklook outputs.
- The QLF team and QL team should iterate on the descriptions within QLF to ensure that the diagnostics are agreed upon and clear to the user
- Inaccuracies in science QA documentation should be corrected
- Details should be provided of how statistics are computed. For example, SKYCOUNT: state the wavelength range where mean (median?) is computed.
- Units should be provided throughout
- The clarity of the QL output data models should be improved to address the last two comments above, e.g. [ql-getrms-CAMERA-EXPID.json](#).

Question: Are the QL algorithms correct?

Answer: The panel did not test the QL outputs against other pipelines, nor did they fully assess the processing or QA analysis code. Some algorithms report results that are not physically motivated, but this may be confusion within the QL→QLF translation rather than improper algorithms.

Findings:

- The panel did not assess the QA algorithms
- The data processing was performed with the QPROC routines

Comments:

- XYSHIFTS is not showing or is set to zero in QLF. It is not clear if the fault lies in QL or QLF
- XWSIGMA reported by QLF: bottom panels report Standard Deviations on order of 1 pixel, which is unrealistic. It is not clear if the fault lies in QL or QLF
- DELTAMAG_TGT reports such limited results that it is difficult to assess whether the algorithm is correct.
- For Calculate SNR, it would be helpful to provide the full parameterization of the fit, fiducial values for the fit, and RMS about the fit.
- Due to time conflicts and panelist expertise, the panel considers a deep dive into the software to be beyond the scope of this review.

Recommendations:

- DESI should identify another individual(s) to independently assess the quicklook software. This individual should have expertise in processing spectroscopic data and not already have excessive responsibilities to the project. Introducing such an individual will ensure that the code is fully reviewed and will spread the knowledge base to other members of DESI.
- Tests on the quicklook reduced spectra should be done to make a pixel-level comparison to the same synthetic images processed by the full pipeline. These tests will help assess differences in cosmic ray rejection, signal-to-noise, sky artifacts, and other features in the data.

Question: Is the QL code sufficiently flexible and written/documented clearly enough that new developers could reasonably contribute new metrics to be included in the outputs or modify existing metrics?

Answer: Documentation provides the procedures to modify or add QA tests, but the panel did not attempt to do so. At face value, the code and documentation appear appropriate, but a deeper assessment would be required to answer this question fully.

Findings:

- The procedure to add QA tests or modify existing QA is described in DESI-4701
- Adding a QA requires modifying 4 files: qlconfig.py (lots of stuff in there), qa_quicklook.py (high level call), qalib.py (low level algorithms) and the qlconfig_xxxx.yaml files (what is run for this type of exposure).

Comments:

- The specific QAs in qlconfig.py make things complicated to follow.
- A different scheme could have been preferred using one python file per QA in desispec/py/desispec/qa/. The file could contain a class with everything related to this QA (low/high level, defaults), leaving no need to modify qlconfig.py at all.

Recommendations:

- DESI should identify another individual(s) to independently assess the quicklook software and documentation. This individual could stress-test the QL system by adding a new metric and also assist during commissioning and Survey Validation.
- Desi_quicklook should be updated so that it will derive the correct default configuration to use based upon the header keywords of the input exposure. Specifying a configuration file should only be required when experts want to test non-default parameters. In particular, QLF should not need to provide a config file to desi_quicklook.

Question: Does the Quicklook Framework interface (QLF) display all of the QL metrics?

Answer: The QLF interface does seem to display all of the metrics in DESI-4701, although there are missing pieces that are not found in that document.

Findings:

- Based on the metrics listed in DESI-4701: all QA have pages, but there is nothing displayed for CHECKFLAT, CHECKARC. SKYRBAND=0
- From the main page, “Processing History”, “Observing History”, and “Afternoon Planning” all point to the same table of exposures taken over a finite window of time
- From the main page, the links to “Trend Analysis”, “Observing Conditions”, and “Survey Reports” appear to be placeholders for further development
- NOISE_OVERSCAN_AMP is not documented anywhere in the quicklook documentation and does not appear to be a necessary metric.

Comments:

- On the CHECK_CCDS page, it would be useful to display the CCD image with options for a histogram, binsize set to sample the noise.
- All items in CHECK_CCDS (BIAS_AMP, NOISE_AMP) could show all spectros, cameras and amps in single display + zoom per spectro
- Consider moving LITEFRAC_AMP and XWSIGMA to a new window (CHECK_SPECTROGRAPHS?)
- LITFRAC_AMP, XYSHIFT, XWSIGMA, NGOODFIB, SKYCONT have limited information in current version; consider consolidating into a single display or enhancing with additional useful information
- There are some inconsistent naming schemes between the metric and the title of each page (e.g. NGOODFIB versus COUNTSPECTRALBINS)
- SKYRBAND, CHECKFLAT, CHECKARC, and a few others appear to be empty
- The trend analysis functionality only offers readnoise and bias. There are surely other things that would be useful such as ETC parameters and spectral diagnostics

Recommendations:

- Given the rate of progress, DESI and the QLF development group should re-visit the goals of the full interface. As a Collaboration, we should consider descoping some of the original wishes for this interface such as the trend analysis, reporting of telemetry/instrument parameters, observing conditions, and afternoon planning displays
- The DESI Project should find other resources to provide visualizations for descoped features.
- The QLF group should focus first on finalizing the metrics that are in place and providing sufficient documentation for others to use the tools and develop tools in the future
- DESI Data systems members and the QLF team should agree how to best communicate in completing the displays for the current metrics so that the requests are matched to the existing software infrastructure

Question: Is the QLF interface clear and user friendly?

Answer: The interface was confusing in its first use but became a helpful and user-friendly tool once acquainted.

Findings:

- Clicking on “Observing History” brings up a table with all of the calibration and science exposures taken in a given time interval
- The “View”, “Com”, and “Logs” columns are not clear and appear to be placeholders
- One must click on the yellow cross under the “QA” column to access the wedge plots and warnings associated with that exposure
- The wedge plots provide a high level summary of all metrics and the corresponding status of the data
- Clicking on a wedge allows the user to drill down on a metric, but it does not seem possible to navigate back to the wedge plots from the new window

Comments:

- The interface is aesthetically pleasing.
- The initial navigation was not obvious: clicking on the yellow cross, the empty window that appears after clicking on the wedge, the identification of the chosen metric
- The SNR criteria are very compelling and should be optimized in the information they convey
- The QA metrics do not always need to be broken up by spectrograph (e.g. CHECK_CCDS, dome flats, NGOODFIBERS)
- In some places histograms may be more informative than per-fiber plots (e.g. sky brightness)
- There are differing opinions among the review panel on the manner in which the high-level and drill-down information should be displayed
- The wedge plots are great! Consistently providing information in the mouse-over such as fiber number and target magnitude would be useful
- DELTAMAG_TGT is defined as “List of the average fiber magnitude for each of N target types in this camera” but I only one number is listed and it is not labeled by target type. It is not clear what this test is supposed to be and the plot could be improved in the form of a histogram.
- NOISE AMP and Get RMS have different names. This should be changed to “READNOISE” as mentioned above in quicklook. The individual amplifiers are not identified in the table, although they are shown clearly in the visual of the CCD. The value for “Reference Exposure” in the table does not match the value in the visual. It is therefore unclear what reference exposure means. The color scheme is nice as are the tables that display alert thresholds. These should clearly state that the thresholds are relative to the reference value. Simply adding a title to the table that states

“DIFFERENCE FROM REFERENCE VALUES TO TRIGGER WARNINGS” would make this more clear.

- BIAS AMP: same comment as above regarding the warning table. Also, description should state: “Value of bias determined from the mean overscan for each amplifier”. The caption below the table states “Photon counts”, is this really gain-corrected or is it in units of ADU?
- LITFRAC AMP: the caption below the visual is incorrect, looks like a cut-and-paste error from the BIAS AMP page. The metric of COUNT PIXELS does not match the title of this page “LITFRAC AMP”. The description of COUNT PIXELS should be more clear, if indeed we want to keep this metric. Consider clicking on the visual to retrieve a postage stamp of the raw CCD image.
- XWSIGMA: this is a nice metric. The visual of the wedge is nice, and the hover to retrieve information from each fiber is nice. The fact that the XSIGMA and WSIGMA values can be retrieved from each fiber by the hover is quite nice. The link of each fiber to the DECaLS viewer is quite nice, but there is a bug in the way that RA and DEC are propagated to the DECaLS webpage. The DECaLS viewer does not recognize the RA/DEC and retrieves a random galaxy on the sky instead. In the definition of XWSIGMA, it should be clearly stated what “X” and “W” are. This would not be obvious to a new user.
- XYSHIFTS does not look like it is being passed into QLF.
- COUNT SPECTRAL BINS/NGOODFIB have discrepant naming conventions. NGOODFIB is preferred. The wedge plot is nice, but clicking on a fiber does not retrieve the DECaLS viewer and instead highlights that fiber for no reason.
- PEAKCOUNT: bug when clicking on fiber in wedge plot. Units not clear, may be more informative if the user could retrieve a postage stamp of the relevant region of the image to assess deviant behavior.
- SkyCont: the wedge plots do correctly link to the DECaLS viewer. This is a nice trait and overall a better metric than the PEAKCOUNT metric. Units should be clear in table and in description. Need to explain in caption below the wedge plot that the fibers highlighted in green correspond to the sky targets that are used to compute the SKYCONT metric.
- DELTAMAG TGT/Integrate spectrum is not at all clear. This should also report the throughput (transparency and seeing) estimates for point sources. This should only report std stars.
- Where can we report the seeing, transparency, and sky background estimates from the ETC? Should that go on the SNR page? When the S/N is poorly behaved, we need to assess whether we are losing signal (does data produce the same throughput that is predicted by the ETC?) or whether we are getting excess noise. To do this comparison, we need to have direct predictions for signal and noise from the ETC and a means to compare these to the data.
- SKYRBAND could use a wedge plot for the R spectrograph. I do not mind having blank entries for the g and z spectrographs. This page should also display the values from the sky monitor

- Consider adding a high-level S/N score to the initial exposures table

Recommendations:

- All tables should give the correct units. For example, SkyPeaks/PEAKCOUNT is in photons per second (I hope) but it is not defined.
- There are detailed comments on each display in the QLF interface. Some of the ambiguity may arise from QL-->QLF communication, some may result from QLF--> Data systems communication, and some may result from an incomplete product. At highest priority, we should all coordinate to remove any ambiguity in the displays.
- Should find agreement on when a display would be more informative to show all spectrographs and all cameras
- Should find agreement on what will allow quickest assessment if exposure is ok without a single click
- The short text in the QLF display should explain much more clearly what the metrics are.
- Should agree on what development is needed for features that are currently maintained as placeholders

Question: Is the QLF code sufficiently flexible and written/documented clearly enough that new developers could reasonably contribute new plots and refinements to existing plots?

Answer: The review panel did not feel qualified to address the questions pertaining to the code. Documentation seems oriented to describing details for experts but is not sufficient for observers or non-LineA DESI developers.

Findings:

- The documentation for the QLF code is here: <https://github.com/desihub/qlf/wiki/QLF-Overview>
- The code for QLF is here: <https://github.com/desihub/desispec>

Comments:

- It is difficult to assess from the web interface what version of the dependencies are running (desispec, desimodel, calibration files, etc.)
- It is not clear how one would update QLF to use a new version of these dependencies or how to point QLF to a version of desispec outside of the docker container (e.g. a master version of desispec installed on desi-1)
- It is not clear how to restart the docker container in the middle of the night (e.g. after a power outage)
- It is not clear how to add a new plot or change an existing plot
- There does not exist a conceptual overview of how the pieces fit together (frontend, backend, external dependencies, docker)

Recommendations:

- Provide a tutorial with an example to modify or add a new a QA plot on a web page, change a web page design, or create new web page and run a test or provide a configuration interface that allows these changes without touching the code.
- Documentation should be improved for both users and developers