

Solutions for Quality Control of multi-detector instruments and their application to CRIRES and VIMOS

Burkhard Wolff^a, Reinhard W. Hanuschik^a, Wolfgang Hummel^a, and Mark Neeser^a

^aESO, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany

ABSTRACT

Quality Control (QC) of calibration and science data is an integral part of the data flow process for the ESO Very Large Telescope (VLT) and has guaranteed continuous data quality since start of operations. For each VLT instrument, dedicated checks of pipeline products have been developed and numerical QC parameters to monitor instrumental behavior have been defined. The advent of the survey telescopes VISTA and VST with multi-detector instruments imposes the challenge to transform the established QC process from a detector-by-detector approach to operations that are able to handle high data rates and guarantee consistent data quality. In this paper, we present solutions for QC of multi-detector instruments and report on experience with these concepts for the operational instruments CRIRES and VIMOS. Since QC parameters scale with the number of detectors, we have introduced the concept of calculating averages (and standard deviations) of parameters across detectors. This approach is a powerful tool to evaluate trends that involve all detectors but is also able to detect outliers on single detectors. Furthermore, a scoring system has been developed which compares QC parameters for new products to those from already existing ones and gives an automated judgment about data quality. This is part of the general concept of information on demand: detailed investigations are only triggered on a selected number of products.

Keywords: quality control, multi-detector instruments, instrument performance, VLT, CRIRES, VIMOS

1. INTRODUCTION

Quality Control of VLT data involves several steps of different complexity. The main part requires pipeline processing of raw data: instrumental and atmospheric signatures are measured, analyzed, and removed from the final products. This part is executed at ESO headquarters in Garching by the Data Processing and Quality Control group (or QC group).

The analysis of pipeline products is based on two main components. Graphical *QC reports* are created for each set of pipeline-reduced files so that the product quality can be directly assessed. The second component consists of numerical representations of certain instrument or data reduction properties like, for example, the median bias level. These are called *QC parameters*. They can be regarded as an abstract of the information contained in the products and the graphical reports. Any change of the instrument behavior is reflected in the QC parameters. They provide an efficient tool for monitoring on short time scales (to detect ‘outliers’) and for following long-term trends. Based on QC parameters and reports, pipeline products are *certified* (or rejected) for further usage. Results of the monitoring and trending analysis are fed back to the observatory in Chile.

QC parameters and graphical reports are usually set up per instrument detector. While current operational instruments have at most four detectors, the new survey telescopes VISTA¹ and VST² will increase this number to 16 and 32, respectively. This adds a challenge in data volume on top of the complexity of already existing ≈ 1000 QC parameters that need to be followed and checked on a regular basis. In this paper, we present solutions for Quality Control on multi-detector instruments. The basic idea is information on demand with the main concepts of QC parameter *aggregates*³ and *scoring*.^{3,4} They are applied to the operational VLT instruments CRIRES and VIMOS.

CRIRES⁵ is a high resolution ($R \approx 80000$) echelle spectrograph operating at $\lambda \approx 1 \dots 5 \mu$. It has four $1k \times 1k$ detectors (read out to $1k \times 0.5k$) that are aligned in a row so that the complete array has effectively $4k \times 0.5k$

Further author information: (Send correspondence to B.W.)
E-mail: bwolff@eso.org

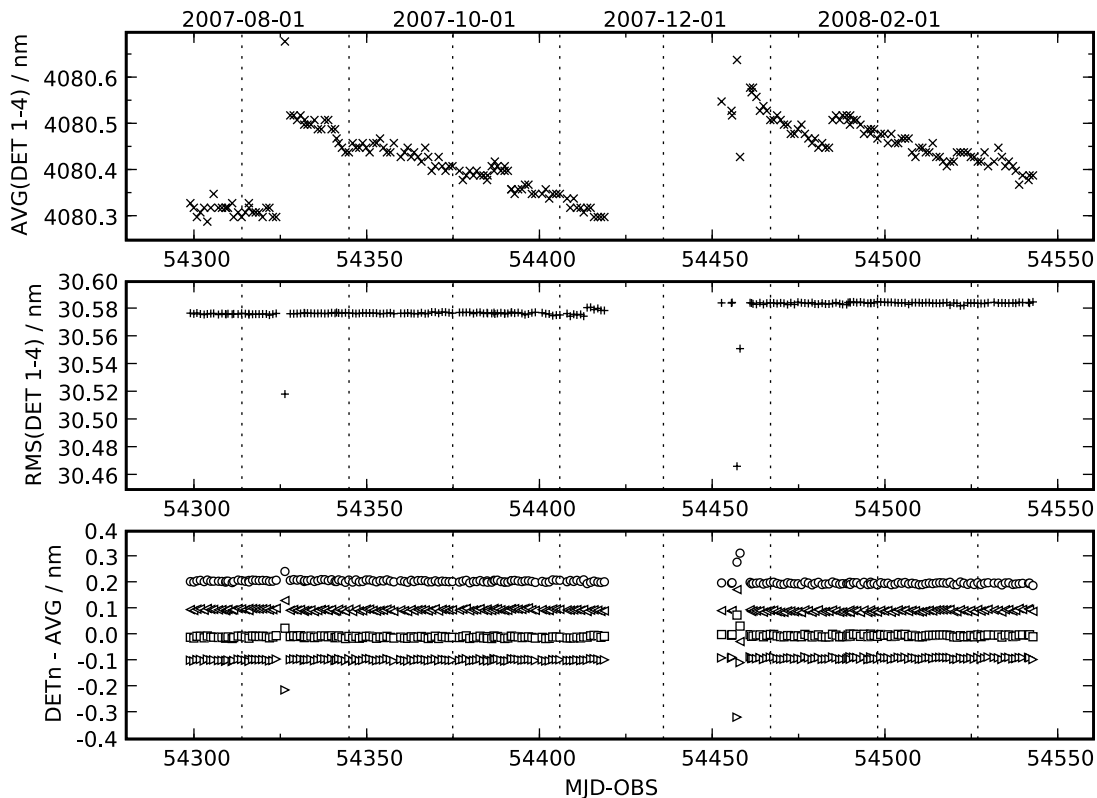


Figure 1. CRIFRES wavelength calibration for a setting with a reference wavelength of 4092.9 nm. Top: measured central wavelength on each detector, averaged over all four detectors. Middle: rms of the central wavelength. Bottom: differences of the individual values to the average (shifted for clarity).

pixels. One echelle order can be observed at a time; the order is dispersed across the four detectors. CRIFRES is operated on a regular basis since April 2007.

VIMOS⁶ is a multi-mode instrument for imaging, multi-object spectroscopy (MOS) using slit masks, and integral field spectroscopy (IFU) in the visual. It has four detectors that form a 2×2 mosaic. The detectors have two read-out modes: low gain with 2148×2440 pixels per detector and high gain with 2148×4096 pixels. VIMOS is operational since March 2003.

2. AGGREGATES OF QC PARAMETERS

QC parameters are calculated separately for each instrument detector because values on different detectors are in many cases uncorrelated. The increased number of parameters for multi-detector instruments makes the traditional approach of monitoring individual detectors unmanageable. Furthermore, it is not sufficient to detect single-detector outliers but it is also necessary to discover coherent changes that involve all detectors and to distinguish between both cases.

Aggregates of QC parameters provide a solution to this problem: we are using the *average* of the individual parameter values of all detectors and the *rms* (standard deviation) across the detectors. The total amount of information is, thereby, condensed into two values that give the necessary information about coherence. In the following, we illustrate typical examples of aggregate behavior and some pitfalls.

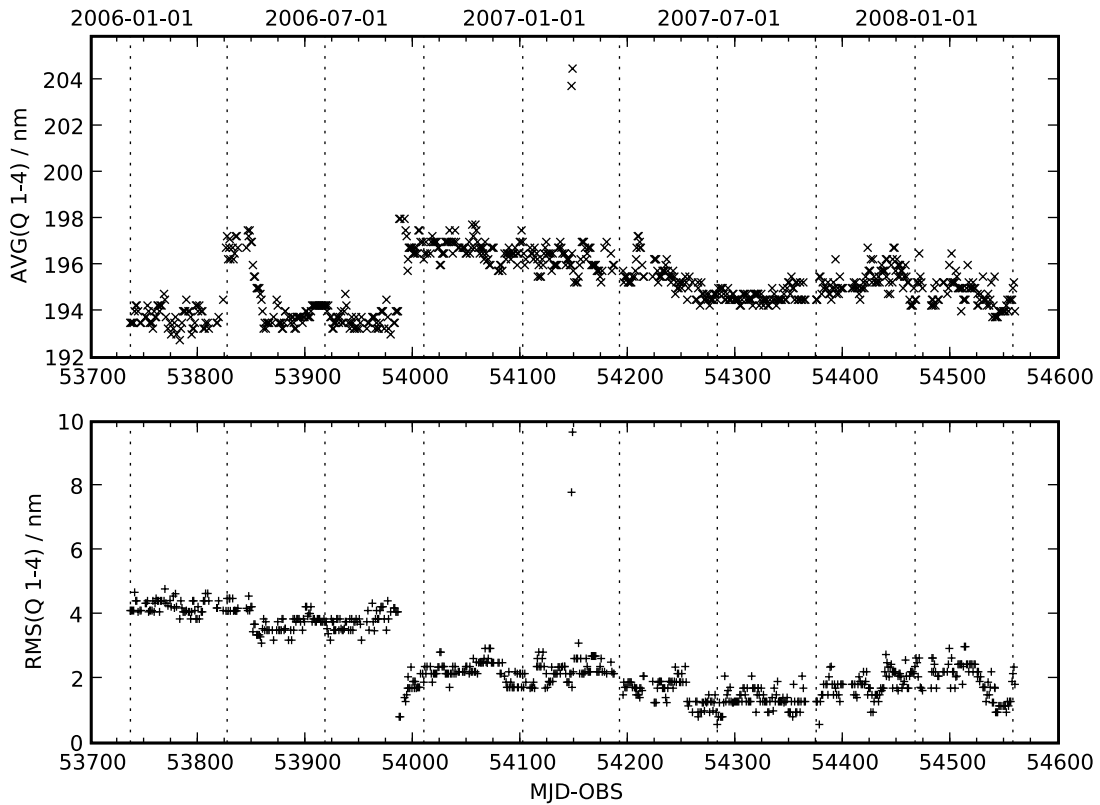


Figure 2. VIMOS bias values: averages of the four quadrants (top) and corresponding rms (bottom).

2.1 Coherent Changes

The first example shows results for wavelength calibration with CRIRES (see Fig. 1) from July 2007 to March 2008 which uses an N_2O gas absorption cell as light source. The grating was positioned so that the nominal (reference) wavelength on the center of detector 3 is 4092.9 nm. The reduction pipeline determines a wavelength solution from a cross-correlation with a reference line list separately for each detector. One important QC parameter is the actual wavelength λ_c of the central column on each detector as given by the solution.

The top panel in Fig. 1 presents averaged values of λ_c . An apparent feature is the continuous small movement of the central wavelength over time, interrupted only by occasional interventions on the instrument in mid August and December 2007 and in January 2008. The middle panel shows the rms of λ_c across the four detectors. With the exception of events in mid August and in mid December, the rms is extremely stable. The shift in central wavelength is, therefore, an example of a coherent trend that involves all detectors in a similar way.

The source of this behavior is a small drift of the cryostat temperature which causes small shifts of optical elements. Average and rms confirm the expectation that this would coherently change wavelength calibration.

2.2 Incoherent Changes of Correlated Values

The jumps of the rms values in Fig. 1 are connected to interventions of the instrument calibration unit. They indicate that the instrument was in an undefined or unstable state at the time of the measurements since temperature drifts alone would result in coherent changes. The simultaneous changes of the averaged values do not provide this kind of information because the cryostat temperature could be different (as for the jump of

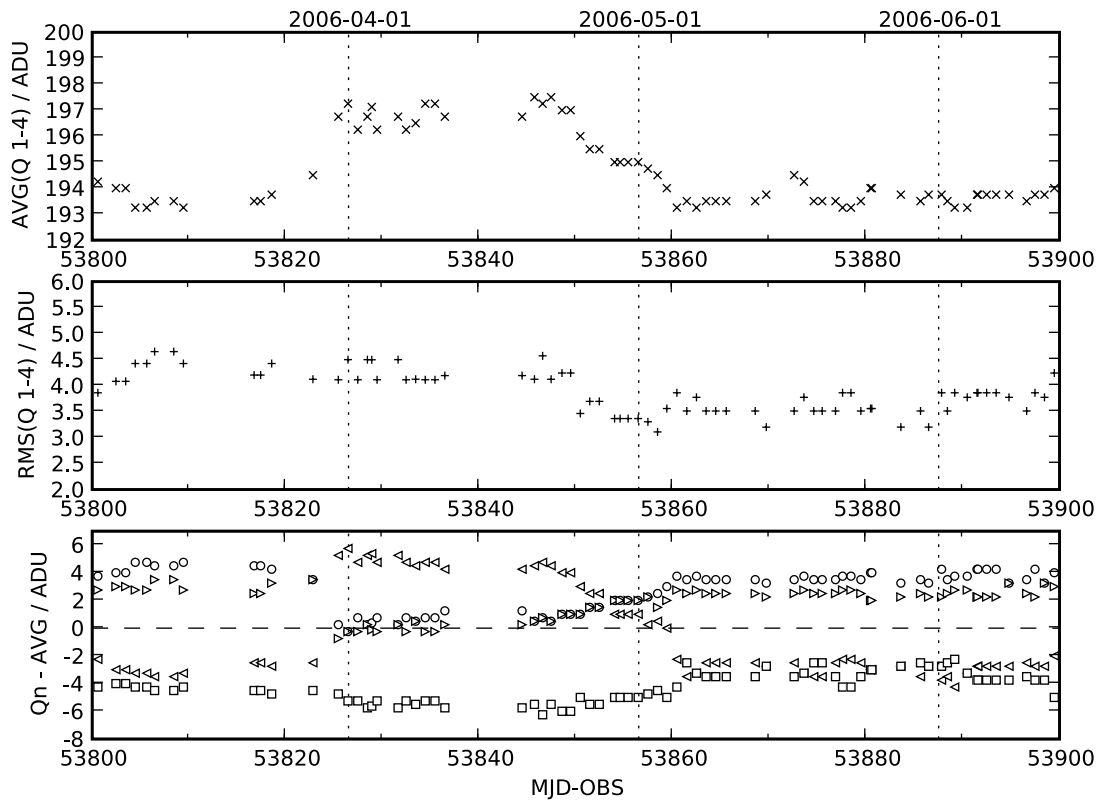


Figure 3. VIMOS median bias values from March to mid June 2006. Top: averaged values. Middle: rms. Bottom: differences of the individual quadrants to the average.

the average in January 2008) or the correlation between central wavelength and temperature could have been permanently changed. Only changes in rms indicate an unstable instrument state.

The bottom panel of Fig. 1 supports the above analysis. It shows values from all four detectors. For clarity, the average has been subtracted from the individual values and the resulting values have been shifted. As expected from the average and rms values, the individual values follow the global behavior (i.e. the difference to the average does not change) with the two mentioned exceptions where the relative values change. It is, therefore, sufficient to monitor only average and rms in order to follow the behavior of the wavelength calibration unit.

2.3 Incoherent Changes of Non-Correlated Values

Other typical behavior of QC parameter aggregates can be found in VIMOS detector bias measurements. The top panel of Fig. 2 shows bias values for the low gain read-out mode from January 2006 to April 2008 averaged over four detectors (which are called quadrants in case of VIMOS). There are a few events when the average bias level suddenly changed: in April 2006 when the new average lasted for about a month, in September 2006 when the new value was stable afterward, and on two days in February 2007. The events in September 2006 and February 2007 were accompanied by changes in the rms across the quadrants as can be seen from the bottom panel of Fig. 2. Since each two quadrants share parts of the read-out electronics, changes in the bias frames often affect only two quadrants and average and rms vary simultaneously. However, it is also possible that changes in two quadrants cancel each other so that the average is not affected. An example has been presented in Ref. 3 for bias exposures with high gain. It is, therefore, not sufficient to monitor only the average.

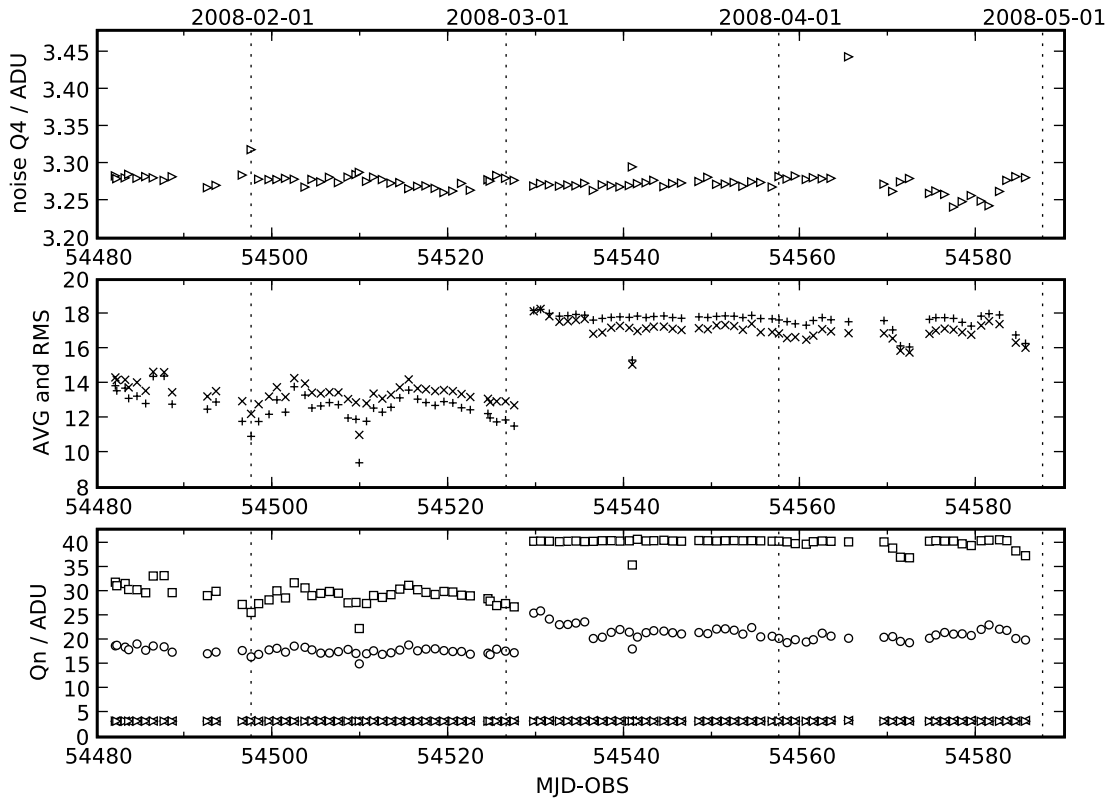


Figure 4. Total noise of VIMOS master bias frames. Top: quadrant 4 only. Middle: average (\times) and rms ($+$) across the four quadrants. Bottom: noise for individual quadrants.

The April 2006 event in Fig. 2 is particularly interesting because it is an example where average and rms values alone do not provide enough information to understand the behavior of the instrument. Since the average increases while the rms remains constant, it seems to be plausible that all four quadrants are coherently affected. However, it is only quadrant 2 that causes the increase of the average whereas the other quadrants are almost unchanged. Figure 3 shows the average and rms values for this event in detail and has in its lower panel the differences of the individual values to the average. This shows that the increased average changes the contribution of the individual values to the rms so that the rms is virtually unchanged.

2.4 Impact on Definition of QC Parameters

For applying aggregates, it may be useful or even necessary to consider changes in the definition or calculation of QC parameters. The example of the total noise in VIMOS master bias frames may illustrate this (see Fig. 4). The total noise is measured by a standard deviation on the central parts of the master bias product. It is composed of read-out noise and contributions from dead columns and gradients across the detector. The top panel in Fig. 4 shows the total noise for quadrant 4 which has a well defined outlier on 9 April due to a small, unusual gradient in the bias frames for this detector. However, the average and rms values (middle panel) do not show any change because average and rms are dominated in this case by quadrants 1 and 3 (lower panel, denoted by squares and circles) whereas quadrants 2 and 4 are negligible. Only changes in quadrants 1 and 3 affect average and rms, as can be seen from the noise increase early March.

Aggregates reflect the behavior on all detectors only if the QC parameter values do not change significantly from detector to detector. Otherwise, the aggregates are dominated by some detectors only. In the presented

case, the differences in noise are due to some bad columns in quadrants 1 and 3. By introducing bad pixel maps for creating master bias frames, the differences could be reduced. This would not be necessary for QC on single detectors but enhances the usefulness of aggregates.

3. SCORING OF QC PARAMETERS

The definition of aggregates for QC parameters helps to handle the volume increase imposed by the increases of detectors of the survey instruments. However, the challenge to monitor a high number of different parameters and to immediately detect sudden changes remains. And in addition, not all changes of individual detectors can be correctly detected by aggregates, as examples in Sects. 2.3 and 2.4 show. We have, therefore, established an automated process to detect outliers: scoring of QC parameters. Ref. 4 gives a detailed description of concepts and applications. We focus here on the role of scoring in the context of outlier detection and product certification for multi-detector instruments.

Scoring is implemented on a subset, which is regarded as being critical for instrument health or product certification, of all QC parameters. Upper and lower limits are defined for each parameter. These thresholds can either be specified, based e.g. on experience (*static* limits), or can be calculated from statistics, i.e. average \pm standard deviation, of the recent parameter history (*dynamic* limits). Each parameter per pipeline product and per detector gets a score of 0 if the value is within the thresholds or a score of 1 if the value is outside. This score is used for outlier detection. The individual scores of all parameters and detectors are summed and give the total score for the product. In the context of certification, any product with a total score greater than 0 would have to be investigated in detail.

An example for outlier detection is the case of VIMOS bias noise presented in Fig. 4. Since the noise is very stable in quadrant 4, narrow static thresholds could be defined and the outlier in April was automatically detected. Scoring makes it, therefore, possible to concentrate on parameter aggregates for analyzing the global behavior without losing control on individual detectors.

In addition to individual detectors, scoring can also be defined on aggregates. This combination can provide better control on product quality than the usage of scoring on individual detectors alone. An application can be found for certification of CRIRES science products. The spectrograph disperses the light over the four detectors, with X being the dispersion direction. The Y position of the spectrum on each detector depends on the position of the object on the slit. While the absolute Y position cannot be predicted, the relative positions on the detectors should be almost the same.

The CRIRES science pipeline recipe is optimized for point-like continuum sources. It extracts the spectrum separately on each detector. The average Y position of the spectrum on each detector is one of the QC parameters. The rms of these four values should be near zero if the pipeline extraction was successful on all detectors. Therefore, scoring has been set up on the rms and is not used for the individual detectors. The rms score gives, together with additional parameter scores, a good indication on the pipeline extraction quality. A high rms is frequently encountered when the assumptions for the pipeline are violated: for weak or extended objects, double stars, or non-continuum sources. Therefore, it is not an indication of the instrument health but a valuable information for everyone who plans to use the pipeline products.

Static and dynamic thresholds require that the QC parameter is constant over time and shows only statistical fluctuations. They cannot be easily applied to a continuous trend like in the case of CRIRES wavelength calibration (see Fig. 1): narrow thresholds give false alarms and have to be adjusted regularly whereas wide thresholds would not detect outliers. There are two ways to resolve this situation: determining thresholds by a *function* and defining limits for *relative* changes.

A function can be used if the QC parameters depends on one or several variables. This is the case for the central wavelength λ_c determined by CRIRES wavelength calibration: Fig. 5 shows λ_c against the grating temperature for data from July to November 2007. There is a good correlation between both quantities with a gradient of $\Delta\lambda_c/\Delta T = 0.034 \text{ nm K}^{-1}$. This allows to predict λ_c and to define (static) thresholds around the predicted value. Any change in the relation $\lambda_c \sim T$ would immediately be detected.

The functional dependency does not detect sudden temperature changes. Those can only be monitored if relative changes from one measurement to another are investigated. This becomes important for calibrating science

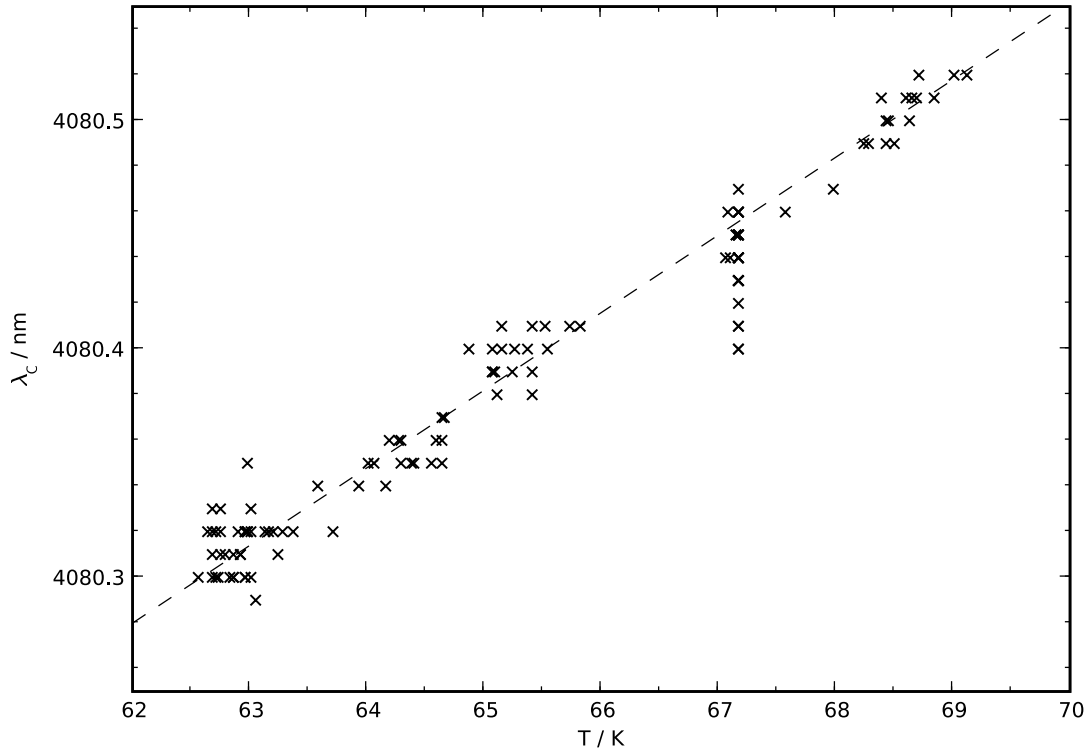


Figure 5. CRIFES: central wavelength λ_c vs. grating temperature. The dashed line has a slope of $\Delta\lambda/\Delta T = 0.034 \text{ nm K}^{-1}$. The single outlier with 4080.68 nm from 13 August (see Fig. 1) has been excluded. It would be at 75.19 K and does not fit the correlation. Several measurements with 67.17 K and different λ_c are caused by a malfunctioning of the temperature sensor.

data: the calibrations applied must have been measured under similar conditions as the science frames. It is, therefore, straightforward to implement a QC parameter that measures the difference of the grating temperature between the science observation and the associated wavelength calibration and to define the relative threshold within the scoring of the science data.

4. INFORMATION ON DEMAND

Aggregates and scoring are effective tools in the context of information on demand. Scoring can be set up hierarchically within the QC process (see Ref. 4 for further details). On the top level, there is one score summarizing all evaluations of individual products and QC parameters for an instrument. Only if this score is not zero, there is a need to investigate lower levels, eventually down to QC reports on individual products and detectors. Aggregates fit well into this approach since they reduce the volume of QC parameters; values for individual detectors may only be needed in case of unusual behavior of the aggregates.

Scoring and aggregates have been added to the well-established concepts of QC reports and parameters for the existing VLT instruments CRIFES and VIMOS. The operational experience shows that the new approaches are feasible, reliable, and can be applied to multi-detector instruments. It is used for efficiently setting up Quality Control on the survey telescopes VISTA and VST and for future VLT instruments.

REFERENCES

- [1] Emerson, J. and Sutherland, W., “Visible and Infrared Survey Telescope for Astronomy: Overview,” in [*Survey and Other Telescope Technologies and Discoveries*], Tyson, J. and Wolff, S., eds., *Proc. SPIE* **4836**, 35–42 (2002).
- [2] Mancini, D., Sedmak, G., Brescia, M., Cortecchia, F., Fierro, D., Garelli, V. F., Marra, G., Perrotta, F., Rovedi, F., and Schipani, P., “VST project: Technical overview,” in [*Telescope Structures, Enclosures, Controls, Assembly/Integration/Validation, and Commissioning*], Sebring, T. A. and Andersen, T., eds., *Proc. SPIE* **4004**, 79–90 (2000).
- [3] Wolff, B. and Hanuschik, R., “Quality Control for multi-detector instruments: a test bed with VIMOS,” in [*Observatory Operations: Strategies, Processes, and Systems*], Silva, D. and Doxsey, R., eds., *Proc. SPIE* **6270** (2006).
- [4] Hanuschik, R. W., Hummel, W., Neeser, M., and Wolff, B., “Scoring: a novel approach towards automated and reliable certification of pipeline products,” in [*Observatory Operations: Strategies, Processes, and Systems*], Brissenden, R. and Silva, D., eds., *Proc. SPIE* **7016**, these proceedings (2008).
- [5] Käuffl, H.-U., Ballester, P., Biereichel, P., Delabre, B., Donaldson, R., Dorn, R., Fedrigo, E., Finger, G., Fischer, G., Franza, F., Gojak, D., Huster, G., Jung, Y., Lizon, J.-L., Mehrgan, L., Meyer, M., Moorwood, A., Pirard, J.-F., Paufigue, J., Pozna, E., Siebenmorgen, R., Silber, A., Stegmeier, J., and Wegerer, S., “CRIRES: a High Resolution Infrared Spectrograph for ESO’s VLT,” in [*Ground-based Instrumentation for Astronomy*], Moorwood, A. and Masanori, I., eds., *Proc. SPIE* **5492**, 1218–1227 (2004).
- [6] LeFevre, O., Saisse, M., Mancini, D., Brau-Nogue, S., Caputi, O., Castinel, L., D’Odorico, S., Garilli, B., Kissler-Patig, M., Lucuix, C., Mancini, G., Pauget, A., Sciaretta, G., Scodreggio, M., Tresse, L., and Vettolani, G., “Commissioning and performances of the VLT-VIMOS instrument,” in [*Instrument Design and Performance of Optical/Infrared Ground-based Telescopes*], Iye, M. and Moorwood, A., eds., *Proc. SPIE* **4841**, 1670–1681 (2003).