Software requirements flow-down and preliminary software design for the G-CLEF spectrograph

Ian N. Evans, Jamie A. Budynkiewicz, Janet DePonte Evans, Joseph B. Miller, Cem Onyuksel, Charles Paxson, and David A. Plummer

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, MS-81, Cambridge, MA 02138, USA

ABSTRACT

The Giant Magellan Telescope (GMT)-Consortium Large Earth Finder (G-CLEF) is a fiber-fed, precision radial velocity (PRV) optical echelle spectrograph that will be the first light instrument on the GMT. The G-CLEF instrument device control subsystem (IDCS) provides software control of the instrument hardware, including the active feedback loops that are required to meet the G-CLEF PRV stability requirements. The IDCS is also tasked with providing operational support packages that include data reduction pipelines and proposal preparation tools. A formal, but ultimately pragmatic approach is being used to establish a complete and correct set of requirements for both the G-CLEF device control and operational support packages. The device control packages must integrate tightly with the state-machine driven software and controls reference architecture designed by the GMT Organization. A model-based systems engineering methodology is being used to develop a preliminary design that meets these requirements. Through this process we have identified some lessons that have general applicability to the development of software for ground-based instrumentation. For example, tasking an individual with overall responsibility for science/software/hardware integration is a key step to ensuring effective integration between these elements. An operational concept document that includes detailed routine and non-routine operational sequences should be prepared in parallel with the hardware design process to tie together these elements and identify any gaps. Appropriate time-phasing of the hardware and software design phases is important, but revisions to driving requirements that impact software requirements and preliminary design are inevitable. Such revisions must be carefully managed to ensure efficient use of resources.

Keywords: G-CLEF, GMT, instrument control software, data reduction pipelines, requirements analysis, software design process

1. INTRODUCTION

The Giant Magellan Telescope (GMT)-Consortium Large Earth Finder (G-CLEF)[1–3] is a fiber-fed optical echelle spectrograph that will be the first light instrument on the GMT. While G-CLEF is a general-purpose instrument, with multiple observing modes capable of addressing many fundamental questions in the areas of stellar astrophysics and cosmological studies, the precision radial velocity (PRV) mode ($R \sim 108,000$) that is designed for exoplanet science drives much of the instrument design. When operating in the PRV mode, G-CLEF will provide unprecedented radial velocity sensitivity, with the capability of making single PRV measurements with a precision of 40–50 $\mathrm{cm\,s^{-1}}$, and an ultimate PRV sensitivity requirement of 10 $\mathrm{cm\,s^{-1}}$ from multiple observations. Many of the complexities involved in computing precision radial velocities to this level of accuracy are described in the excellent review of Fischer et al.[4]

Such PRV sensitivity is achieved by maintaining the G-CLEF optical bench in an ultra-stable environment that is immune to mechanical perturbations, either intrinsically by design (*e.g.*, vibration isolation), or through active control by software (*e.g.*, thermal control, active leveling). These factors drive the flow-down to the software requirements and preliminary design of the instrument device control subsystem (IDCS), and significantly increase the complexity of the latter. In addition to the typical mechanical and detector subsystems, the IDCS must manage an active precision thermal control system that maintains the optical bench temperature to ±1 mK,

---

Send correspondence to ievans@cfa.harvard.edu

Software and Cyberinfrastructure for Astronomy III, edited by Gianluca Chiozzi,
Juan C. Guzman, Proc. of SPIE Vol. 9913, 99131Q · © 2016 SPIE
CCC code: 0277-786X/16/$18 · doi: 10.1117/12.2232163

and a precision active leveling system that ensures that the spectrograph assembly is maintained in a gravity invariant orientation with a precision of a few micro-radians. The IDCS must monitor telemetry from several hundred sensors and manage more than one hundred controllable devices. In addition to instrument control, achieving the specified radial velocity precision requires the use of data reduction algorithms that are carefully matched to the instrument operational concept in order to fully eliminate common sources of calibratable errors and minimize the impact of non-calibratable and systematic errors.

To ensure tight integration with the GMT software systems, the G-CLEF software uses a model-based systems engineering[5] approach and conforms to a reference architecture[6] developed by the GMT Organization (GMTO). While model-based systems engineering is commonly used for large scale industrial robotics systems, this approach has only relatively recently been utilized[7] in the development of ground-based astronomical instrumentation, where the focus is typically on the hardware. However, in the era of billion-dollar extremely large telescopes, and their concomitant operating costs, expectations for first light performance raise the bar for software development scope and schedule to near space-borne mission standards.

In this paper, we describe the approaches that we have used to identify and flow down the requirements for the G-CLEF IDCS and data reduction pipeline software, and discuss the preliminary designs for key components. Since this phase of the G-CLEF software effort is being conducted in parallel with both the internal G-CLEF hardware design work and also with the GMT's own software requirements and design effort, we discuss how we handle revisions to driving requirements that impact our software requirements and preliminary design process. Finally, we identify a number of lessons learned that have general applicability to the development of software for ground-based instrumentation under similar circumstances.

2. G-CLEF INSTRUMENT DESCRIPTION

The G-CLEF instrument includes three major functional subsystems that must be managed by the IDCS, as shown in Figure 1. These subsystems are the *G-CLEF front end assembly* (GCFEA), *G-CLEF calibration system* (GCCS), and the *G-CLEF spectrograph* (GCSP).

The GCFEA includes the optical interface to the telescope and the calibration system. The front end assembly is separated further into a subassembly (GCGIR) located on the rotating *GMT Gregorian Instrument Rotator* (GIR), and a subassembly (GCIP) located on the fixed *GMT Instrument Platform*. When the GIR is locked in the appropriate orientation, a deployable stage on the GCGIR can be extended to insert a steerable (tip-tilt) tertiary mirror into the telescope optical path to direct light through collimating optics and onto the fiber selection system on the GCIP. Since the front-end optical train is split between the GCGIR and the GCIP, the GCGIR also includes components necessary to ensure proper optical alignment. The GCIP includes several key mechanisms, including an *atmospheric dispersion corrector* (ADC) that corrects for differential refraction due to the atmosphere, a *fiber selector system* (FSS), and a *flexure control system* (FCS).

The G-CLEF instrument supports four observing modes with varying spectral resolution and throughput. These are an optically scrambled precision radial velocity (PRV) mode with spectral resolution $R \sim 108,000$; a non-scrambled PRV mode that has similar resolution and higher efficiency, but with degraded PRV accuracy due to the lack of optical scrambling; a medium resolution (MR) mode with $R \sim 35,000$ that provides a good balance between resolution and throughput; and a high throughput (HT) mode that has the highest overall efficiency but with lower spectral resolution ($R \sim 19,000$). For each observing mode there is an "object" fiber that is used for the science target (and that may consist of a set of pupil sliced apertures for some modes), and a pair of "sky" fibers (only one of which may be used at any time) that can be used to simultaneously observe the sky background. The FSS is used to insert the optical fibers for the selected instrument mode into the optical path to direct light to the spectrograph. The FSS also includes the shutters that control the duration of the individual exposures, as well as a pair of calibration masts that may be positioned to insert light from the GCCS into the object and/or either of the sky fibers for calibration purposes. A pupil alignment camera that is used to align the GCGIR and GCIP optical paths after the deployable stage inserts the tertiary mirror into the telescope beam is also mounted on the FSS.

The FCS incorporates a commercial CCD camera (and associated mechanisms) that is used to measure telescope pointing and focus offsets caused by differential flexure between the G-CLEF front end assembly and

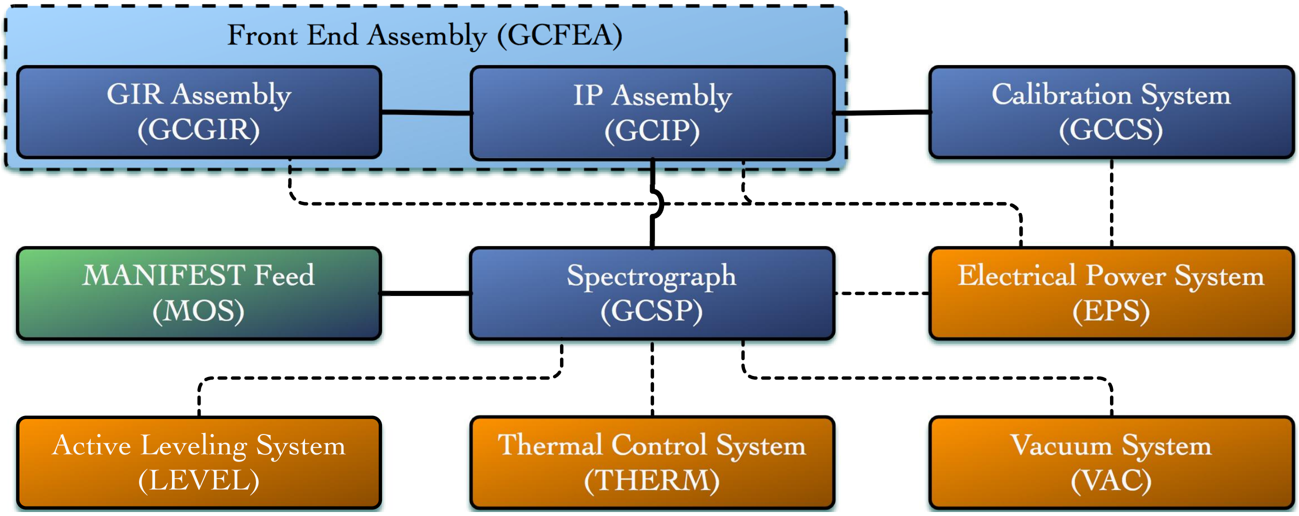


Figure 1. The G-CLEF instrument functional subsystems. The major functional subsystems are shown in blue. Supporting operational subsystems are shown in brown. The external MANIFEST feed is shown in green. Solid lines indicate optical connections, and dashed lines indicate electrical/control connections.

the telescope guider, so that these offsets can be corrected. The FCS camera is also used to acquire targets and position them correctly on the object fiber.

The spectrograph (GCSP) is mounted at the *Gravity Invariant Station* (GIS) on the *GMT Azimuth Disk*, and is fed by an optical fiber train from the FSS mounted on the GCIP. The spectrograph optics are mounted on a composite optical bench assembly that must be maintained at a constant temperature to achieve the required PRV precision. The spectrograph includes a pair of large format ($\sim 10\,\mathrm{K} \times 10\,\mathrm{K}$) CCD science cameras — one for each of the two (blue, red) spectrograph legs. The spectrograph optical bench is mounted on a vibration isolation system inside a vacuum chamber, and wrapped by a series of low-power heater panels. This entire assembly is in turn mounted inside a thermal enclosure that is maintained at a temperature that is biased $\sim 2.5\,^\circ\mathrm{C}$ lower than the target optical bench temperature, so that the bench temperature can be controlled solely by modulating the input power to the heater panels.

The calibration system (GCCS) provides calibration light sources that are fed by optical fibers to the calibration masts mounted on the FSS. The GCCS includes both continuum lamps that are used for flat-fielding the echelle spectral orders on the detectors (i.e., correcting for pixel-to-pixel quantum efficiency variations), and arc lamps that are used for wavelength calibration. Up to two lamps may be selected simultaneously, and the calibration light from the selected lamp passes through a neutral density filter and shutter mechanism prior to being injected into optical fibers that feed the calibration masts.

Finally, the G-CLEF spectrograph supports a direct multi-object spectroscopy feed that will be provided by the MANy Instrument FibEr SysTem (MANIFEST)[8,9] mounted at the GMT Gregorian focus. A complete definition of this mode of operation has yet to be determined. However, we anticipate that the GMT will be responsible for the software and controls interface to MANIFEST.

Several additional subsystems provide capabilities necessary to support operation of the G-CLEF instrument. These subsystems include the *thermal control system*, the *active leveling system*, the *vacuum control system*, and the *electrical power system*.

The thermal control system provides precision thermal control of both the spectrograph optical bench assembly and the CCD detectors mounted in the science cameras. Both subsystems must be maintained at constant temperatures (nominally $20.0\,^\circ\mathrm{C}$ for the optical bench, and -100 to $-120\,^\circ\mathrm{C}$ for the CCD detectors) with a precision of $\pm 1\,\mathrm{mK}$. The thermal control system also maintains the spectrograph thermal enclosure at the appropriate bias temperature relative to the optical bench assembly.

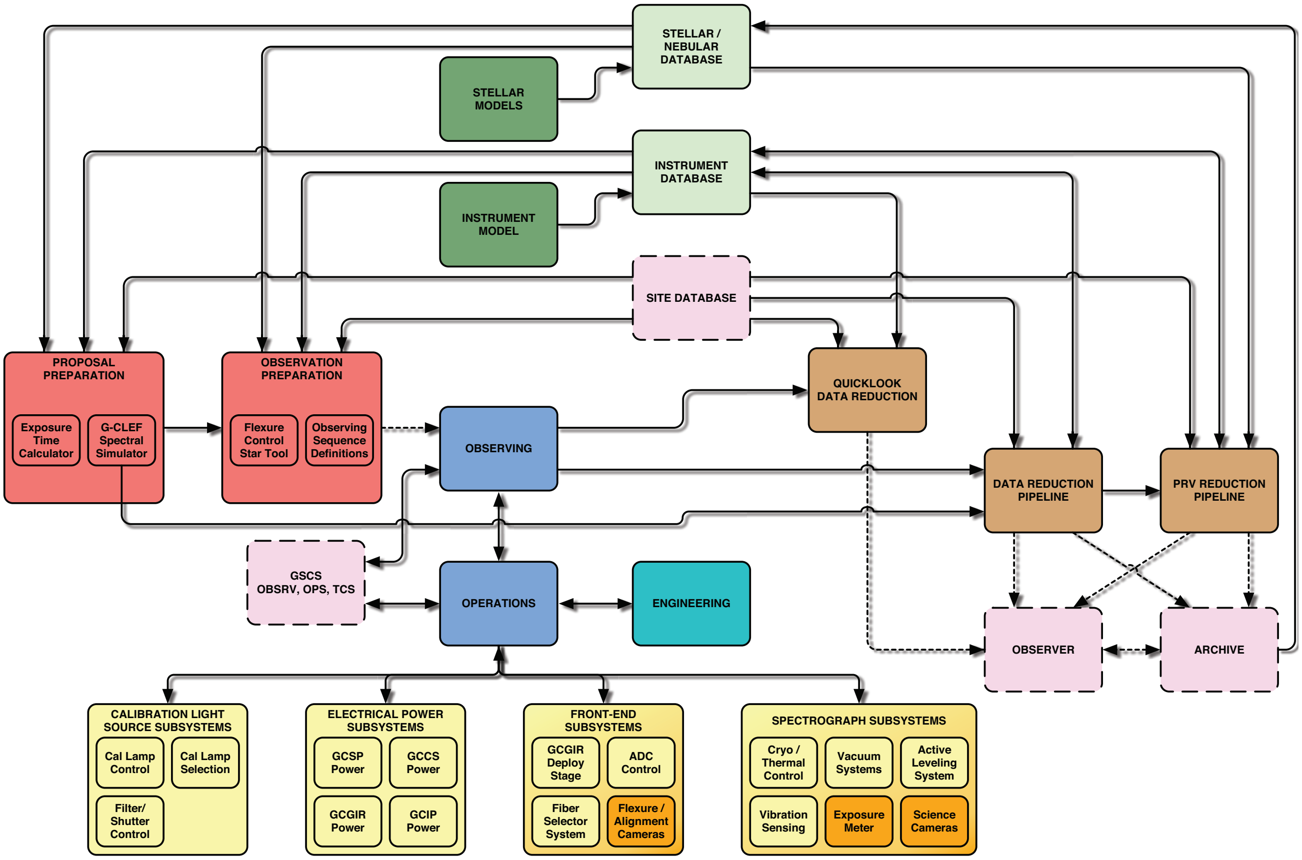


Figure 2. G-CLEF software and related components. See the text for a description of the individual components.

Although the spectrograph is located at the GMT gravity invariant station, the permissible misalignment between the GMT azimuth disk rotation axis and the true gravity vector, combined with localized distortions resulting from the finite stiffness of the azimuth disk, can result in variations in the true gravity vector seen by the optical bench assembly that exceed the maximum allowable to achieve the G-CLEF PRV precision requirement. The active leveling system provides precision leveling control of the spectrograph chamber by actively adjusting the lengths of the three vacuum chamber support legs to maintain reference inclinometers mounted on each leg oriented perpendicular to the instantaneous gravity vector.

The vacuum control system operates the vacuum pumps, valves, and gauges that maintain the spectrograph vacuum chamber, and each of the science camera assemblies, under vacuum. This system also provides startup and shutdown sequencing of the vacuum system.

Finally, the electrical power system controls and monitors electrical power to the front end, spectrograph, and calibration light source systems.

3. G-CLEF SOFTWARE SUITE

The GMTO has developed a reference architecture that requires tight integration between all components of each instrument's IDCS and the GMT software and controls system, and is developing software frameworks to facilitate this integration. The GMT Software and Controls Standards[10] (hereafter SWCS) identifies 13 "packages" that comprise a device control subsystem, each of which typically supports some aspect of instrument control or science operations. Packages that involve direct control of instrument hardware and detectors are termed *Device Control Packages* (DCP). The DCP run on a *Linux* platform.

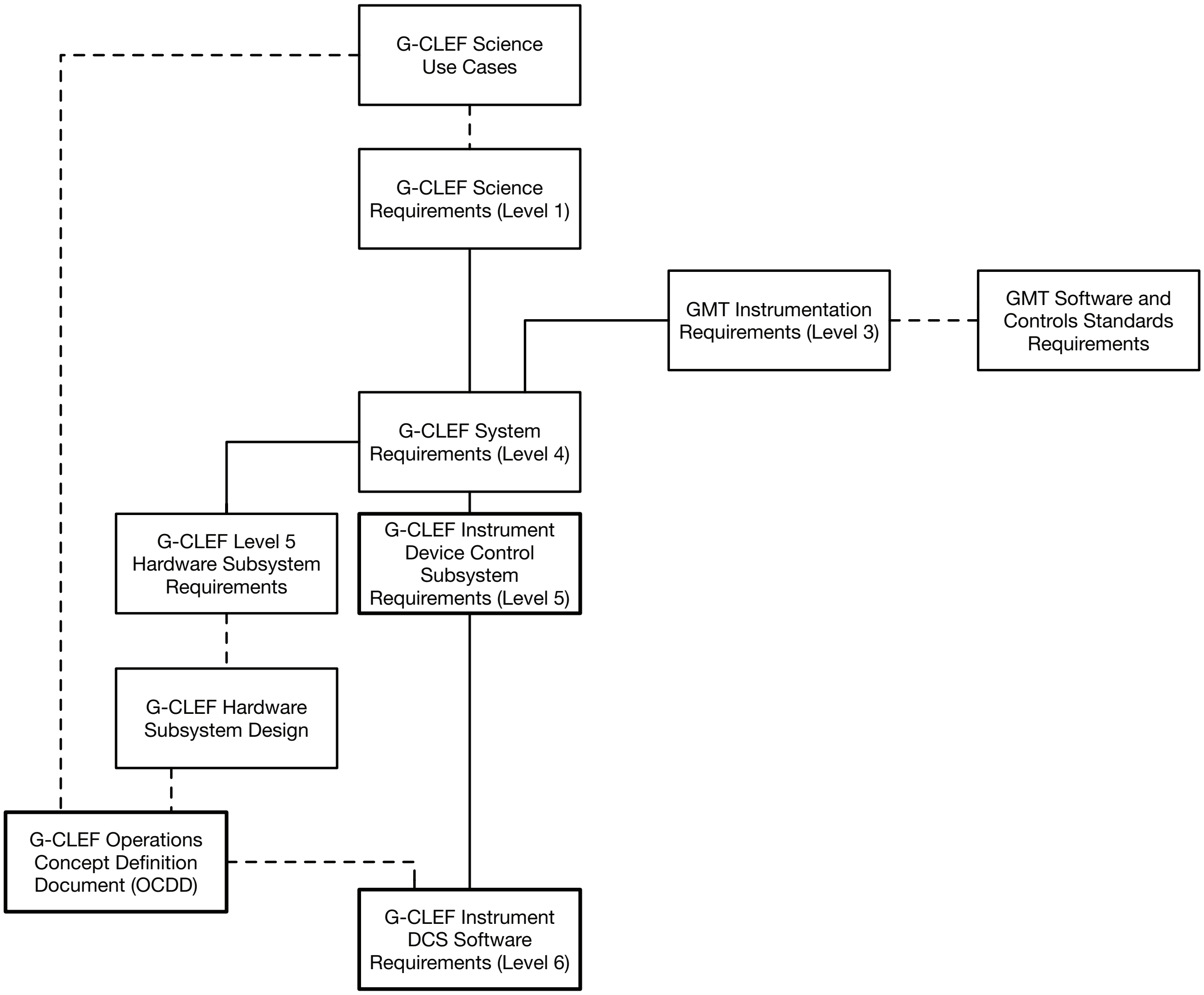
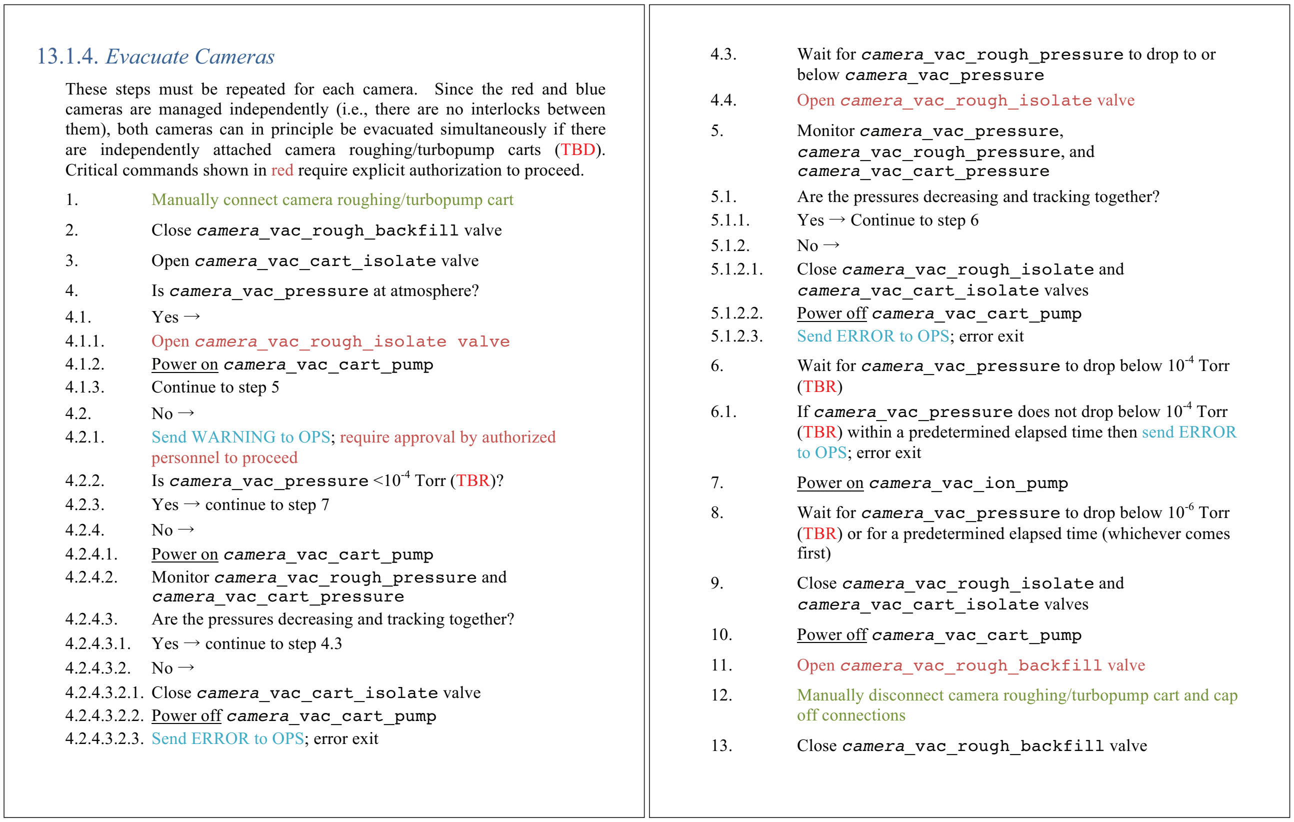


Figure 3. G-CLEF IDCS requirements flow-down. The documents described in this paper are highlighted with bold boxes. Solid lines indicate the formal requirements flow-down paths, while dashed lines highlight informational relationships.

In addition to the DCP, the G-CLEF software suite will include operation support packages, a data processing package, and a set of proposal and observation management tools. An example of an operation support package is a graphical user interfaces that enable the operator to interact with the instrument in a simple and controlled manner. The data processing package includes standard and PRV data reduction pipelines incorporate optimal algorithms designed to maximize the scientific utility of the data, even for observers who are unfamiliar with the details of PRV spectroscopy data reductions. The proposal and observation management tools allow prospective observers to estimate exposure times for observations and validate that the proposed observations will achieve their scientific goals.

A block diagram of the software and related components is shown in Figure 2. The yellow and orange blocks map to DCP (yellow for the *Control Package*; orange for the *Data Acquisition Package*), while the blue, turquoise, and brown blocks largely map to the operation support packages (blue for the *Sequencing*, *Visualization*, and *Safety Packages*; turquoise for the *Diagnosis* and *Calibration Packages*; brown for the *Data Processing Package*). Blocks shown in light green identify calibration data and/or model data recorded in persistent storage. They are loaded initially from the associated models shown in dark green. The pink blocks encased in long dashed lines represent functionality that we expect will be provided by GMTO.



13.1.4. *Evacuate Cameras*

These steps must be repeated for each camera. Since the red and blue cameras are managed independently (i.e., there are no interlocks between them), both cameras can in principle be evacuated simultaneously if there are independently attached camera roughing/turbopump carts (TBD). Critical commands shown in red require explicit authorization to proceed.

1. Manually connect camera roughing/turbopump cart
2. Close `camera_vac_rough_backfill` valve
3. Open `camera_vac_cart_isolate` valve
4. Is `camera_vac_pressure` at atmosphere?
4.1. Yes →
4.1.1. Open `camera_vac_rough_isolate valve`
4.1.2. Power on `camera_vac_cart_pump`
4.1.3. Continue to step 5
4.2. No →
4.2.1. Send WARNING to OPS; require approval by authorized personnel to proceed
4.2.2. Is `camera_vac_pressure` <10^{-4} Torr (TBR)?
4.2.3. Yes → continue to step 7
4.2.4. No →
4.2.4.1. Power on `camera_vac_cart_pump`
4.2.4.2. Monitor `camera_vac_rough_pressure` and `camera_vac_cart_pressure`
4.2.4.3. Are the pressures decreasing and tracking together?
4.2.4.3.1. Yes → continue to step 4.3
4.2.4.3.2. No →
4.2.4.3.2.1. Close `camera_vac_cart_isolate` valve
4.2.4.3.2.2. Power off `camera_vac_cart_pump`
4.2.4.3.2.3. Send ERROR to OPS; error exit
4.3. Wait for `camera_vac_rough_pressure` to drop to or below `camera_vac_pressure`
4.4. Open `camera_vac_rough_isolate` valve
5. Monitor `camera_vac_pressure`, `camera_vac_rough_pressure`, and `camera_vac_cart_pressure`
5.1. Are the pressures decreasing and tracking together?
5.1.1. Yes → Continue to step 6
5.1.2. No →
5.1.2.1. Close `camera_vac_rough_isolate` and `camera_vac_cart_isolate` valves
5.1.2.2. Power off `camera_vac_cart_pump`
5.1.2.3. Send ERROR to OPS; error exit
6. Wait for `camera_vac_pressure` to drop below 10^{-4} Torr (TBR)
6.1. If `camera_vac_pressure` does not drop below 10^{-4} Torr (TBR) within a predetermined elapsed time then send ERROR to OPS; error exit
7. Power on `camera_vac_ion_pump`
8. Wait for `camera_vac_pressure` to drop below 10^{-6} Torr (TBR) or for a predetermined elapsed time (whichever comes first)
9. Close `camera_vac_rough_isolate` and `camera_vac_cart_isolate` valves
10. Power off `camera_vac_cart_pump`
11. Open `camera_vac_rough_backfill` valve
12. Manually disconnect camera roughing/turbopump cart and cap off connections
13. Close `camera_vac_rough_backfill` valve

Figure 4. The G-CLEF OCDD *Evacuate Cameras* procedure detailed sequence.

3.1 G-CLEF Device Control Package

Development of G-CLEF DCP requirements follows a formal but pragmatic process. Emphasis is placed on establishing a complete and correct set of requirements rather than on the mechanism through which the requirements are gathered. The top-level requirements are determined based on a combination of historical experience with similar instrumentation, an analysis of the G-CLEF instrument engineering design, and requirements flowed-down from the G-CLEF System Requirements document.[11] The latter includes imposed requirements flowed-down from the GMTO. Software and controls subsystem-level requirements are recorded in the G-CLEF IDCS Requirements document,[12] while software-specific requirements are broken down in greater detail in the G-CLEF IDCS Software Requirements document.[13] To ensure compliance, all requirements at every level are associated with a verification method (analysis, demonstration, inspection, test) through a verification matrix.

The requirements for the DCP that are recorded in the G-CLEF IDCS Software Requirements document either flow-down from the G-CLEF IDCS Requirements document, or are based on an analysis of the G-CLEF Operations Concept Definition Document (OCDD),[14] which is a project science document that describes in detail how the G-CLEF instrument will operate. The OCDD was developed from an analysis of the G-CLEF science use-cases that formed the input to the G-CLEF Science Requirements document,[15] and the procedures are extensively informed by the instrument hardware subsystem preliminary design. The G-CLEF IDCS requirements flow-down is presented graphically in Figure 3.

The OCDD identifies software controllable hardware components and sensors, and provides operational procedures for both routine and non-routine circumstances. The G-CLEF instrument includes approximately 180 software-controllable devices and another ∼270 software-readable sensors. Roughly 1/3 of the components are directly associated with the thermal control system.

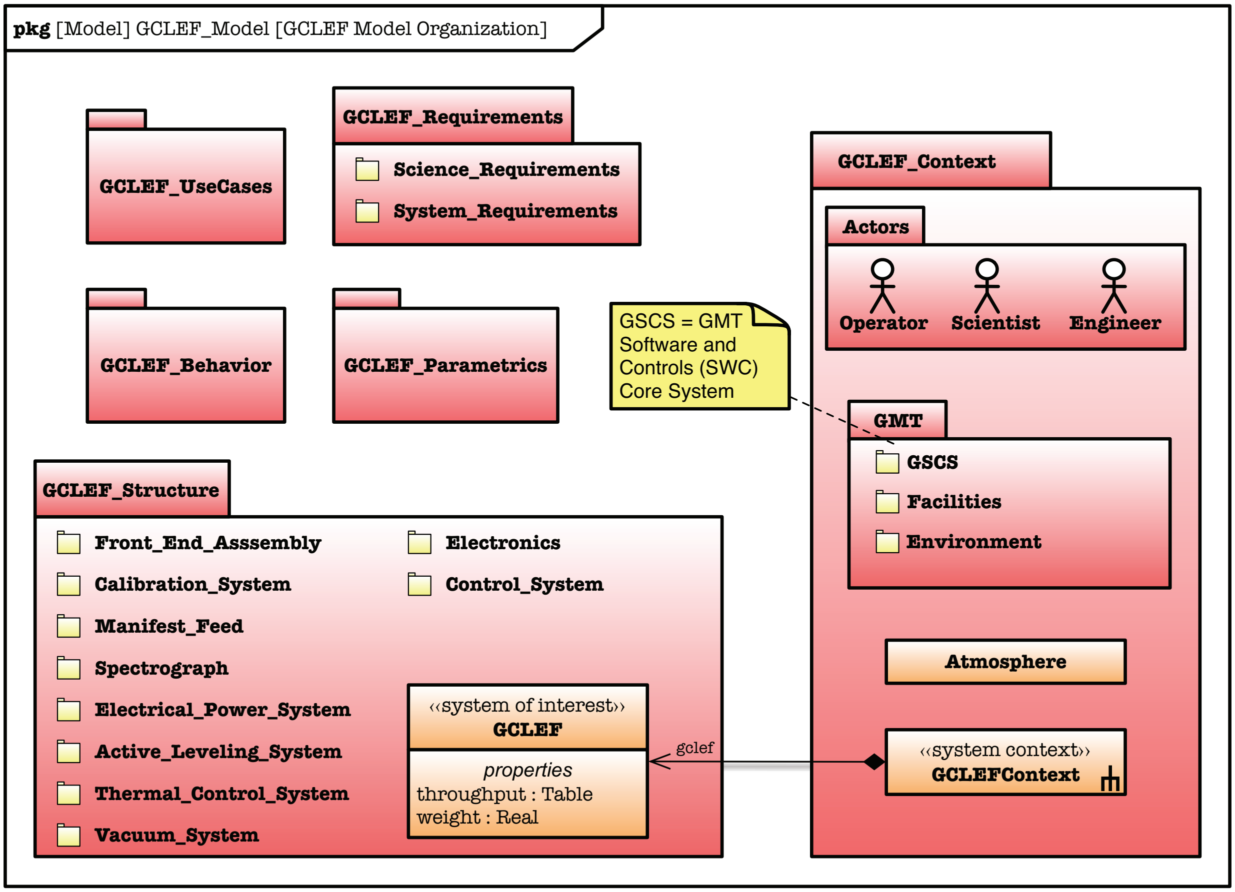


Figure 5. Preliminary organization of top-level G-CLEF-specific SysML models.

Routine operations include startup and shutdown procedures, general sequences for obtaining science observations and calibration exposures, ancillary procedures used when observing (such as target acquisition, flexure camera focus, and telescope focus), and procedures for the various control loops that must be executed whenever the instrument is in use. The latter include the flexure control, flexure camera rotator, ADC, focus control, and active leveling loops. In addition, daily operations procedures, and continuous system monitoring and verification are included in the set of routine operations.

Non-routine operations described in the OCDD include procedures for handling abnormal and emergency situations, including emergency shutdown and recovery procedures. Common maintenance procedures are also described, as are complex procedures used only during specific project phases (*e.g.*, assembly, integration, and test, or on-telescope commissioning).

The procedures included in the OCDD are provided textually so that the reader can understand the *intent* of the procedure, and also as detailed sequences so that the exact steps necessary to execute the procedure are understood. As an example of the latter, see the *Evacuate Cameras* procedure in Figure 4.

The procedures in the OCDD are analyzed together with the instrument engineering design documentation to derive a set of formal requirements that are captured in the G-CLEF Instrument Device Control Subsystem Software Requirements document. An example of the level of requirements identified using this process is the following:

> **GCSWCS Vacuum System Ion Pump HV** — The G-CLEF DCS Software shall provide the capability to set the voltage of the ion pump controller high voltage output.

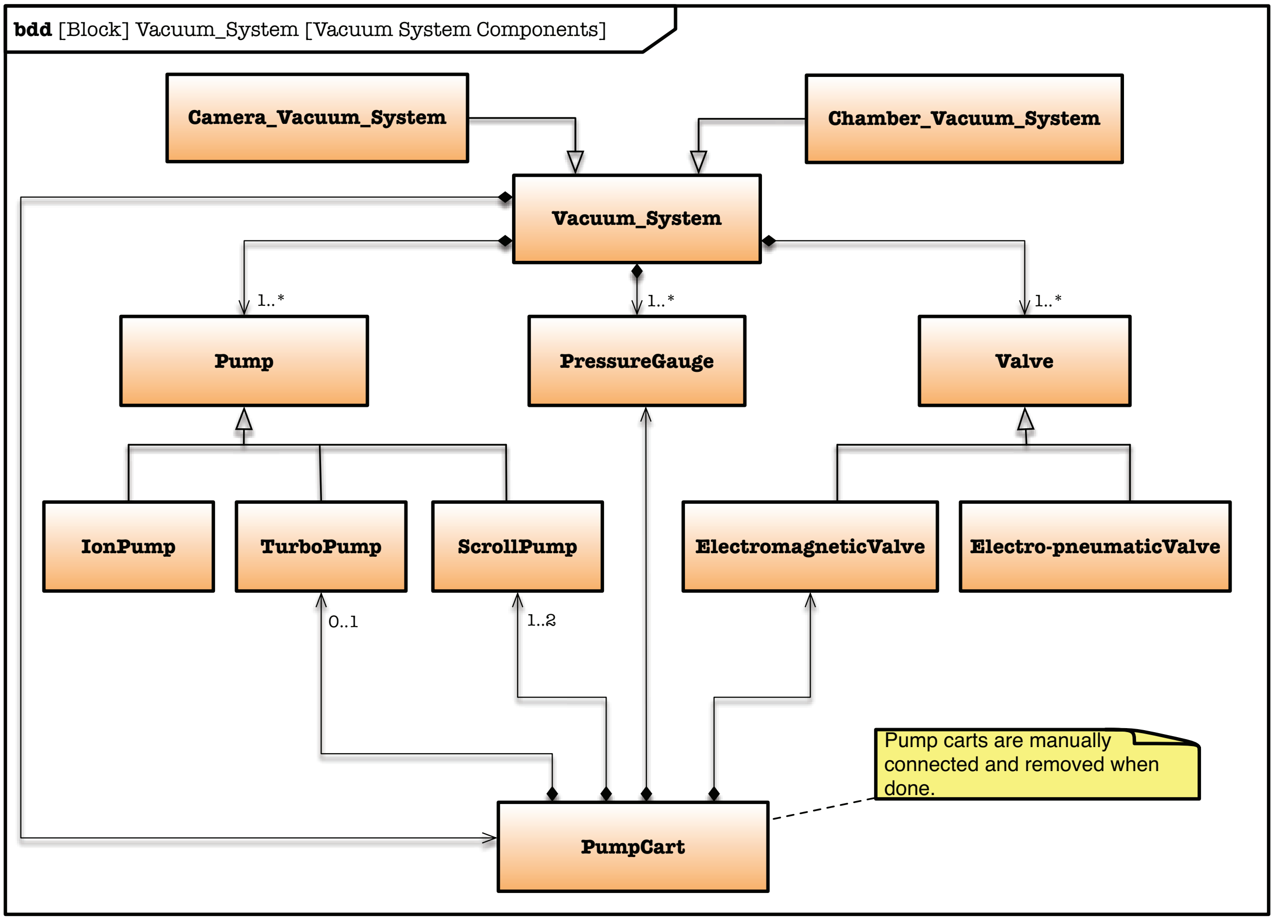


Figure 6. Preliminary G-CLEF vacuum system block definition diagram.

While such requirements can be easily implemented in software and their performance verified, they may not include the entire set of details required to design and implement the DCP, because they do not readily capture many of the dynamic processes inherent in an operational hardware environment.

To ensure that the requirements are analyzed completely, the OCDD, software requirements developed to this stage, and instrument engineering design are used as a starting point to develop a state-based model of each G-CLEF subsystem in SysML using a formal systems modeling process. These models are refined through iteration with project science and engineering staff to validate the operations and design details. They incorporate further inputs such as the effects of safety system interlocks and over-rides, and detailed controls hardware design that may be developed subsequent to the hardware preliminary design review. Although this approach requires considerable effort, the result is a model of the instrument controls that captures the required software functionality with good fidelity.

The use of a state-based control system is mandated by the GMTO SWCS. A state-based control system is well-suited to a complex instrument such as G-CLEF because every subsystem, sub-subsystem, and so on, needs to know only its current state and how to transition from that state to a new desired state that is specified by an external goal mechanism. This approach immediately modularizes and parallelizes hardware control to the extent possible. Once the system modeling is completed, any new requirements identified as part of the modeling effort are incorporated into a revision of the G-CLEF Instrument Device Control Subsystem Software Requirements document so that the they can be formally tracked and ultimately verified. At the time of writing, the G-CLEF software team is just commencing the system modeling effort. A preliminary organization of the G-CLEF top level SysML models is shown in Figure 5.

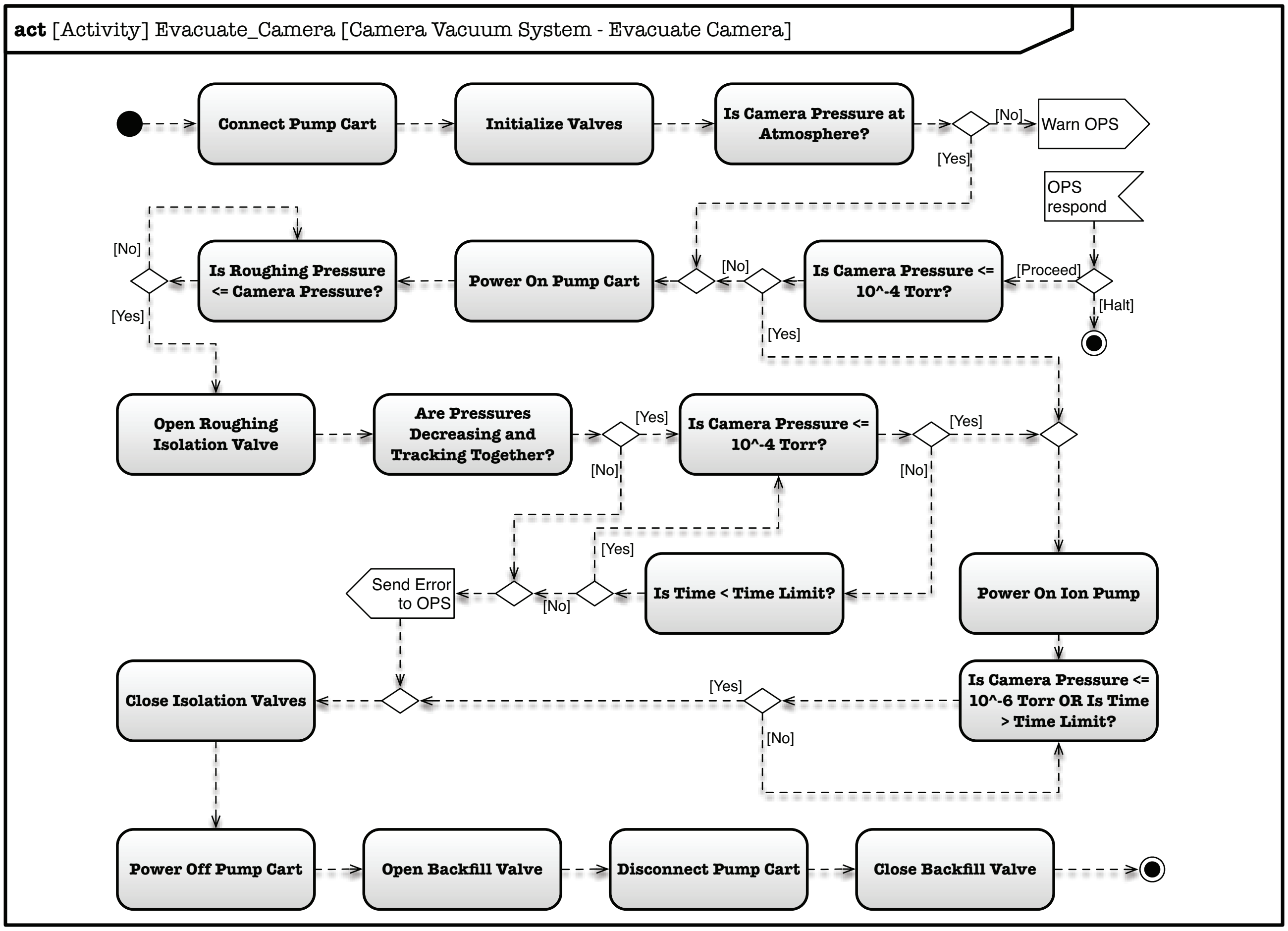


Figure 7. Example G-CLEF vacuum system *Evacuate Cameras* activity diagram. This activity corresponds to the OCDD *Evacuate Cameras* procedure detailed sequence shown in Figure 4

In order to validate G-CLEF use of the modeling approach, a model for the vacuum system is being developed initially. This is partly because the vacuum system engineering design is well advanced, and partly because the vacuum system includes numerous interlocks that will prevent unsafe conditions by inhibiting valves being opened or closed under inappropriate circumstances. Therefore the vacuum system is an excellent testbed for the modeling effort. A small subset of the SysML diagrams created as part of the vacuum system modeling effort are included below for illustration.

The preliminary vacuum system block definition is shown in Figure 6. Here the major controllable hardware component types are identified and their aggregation and generalization relationships shown. Many of these hardware blocks will be mapped directly to software blocks. The GMTO SWCS require the use of an EtherCAT[*] fieldbus for hardware control wherever possible, and will provide software frameworks for the *Linux*-based device control computer that will greatly simplify the software development effort when using EtherCAT-compliant devices. Since all of the vacuum system components either communicate directly using the EtherCAT fieldbus, or via a serial port that can be interfaced using an RS-232 EtherCAT I/O module, at this level, there is a relatively tight relationship between the defined software blocks and the corresponding hardware components.

Each of the procedures identified in the OCDD that manipulate any vacuum system components are next translated into activity diagrams. This step formalizes the procedures and helps to eliminate any ambiguities that may exist in the OCDD description. Interactions with external systems are also identified explicitly, and these

---

[*] https://www.ethercat.org/default.htm

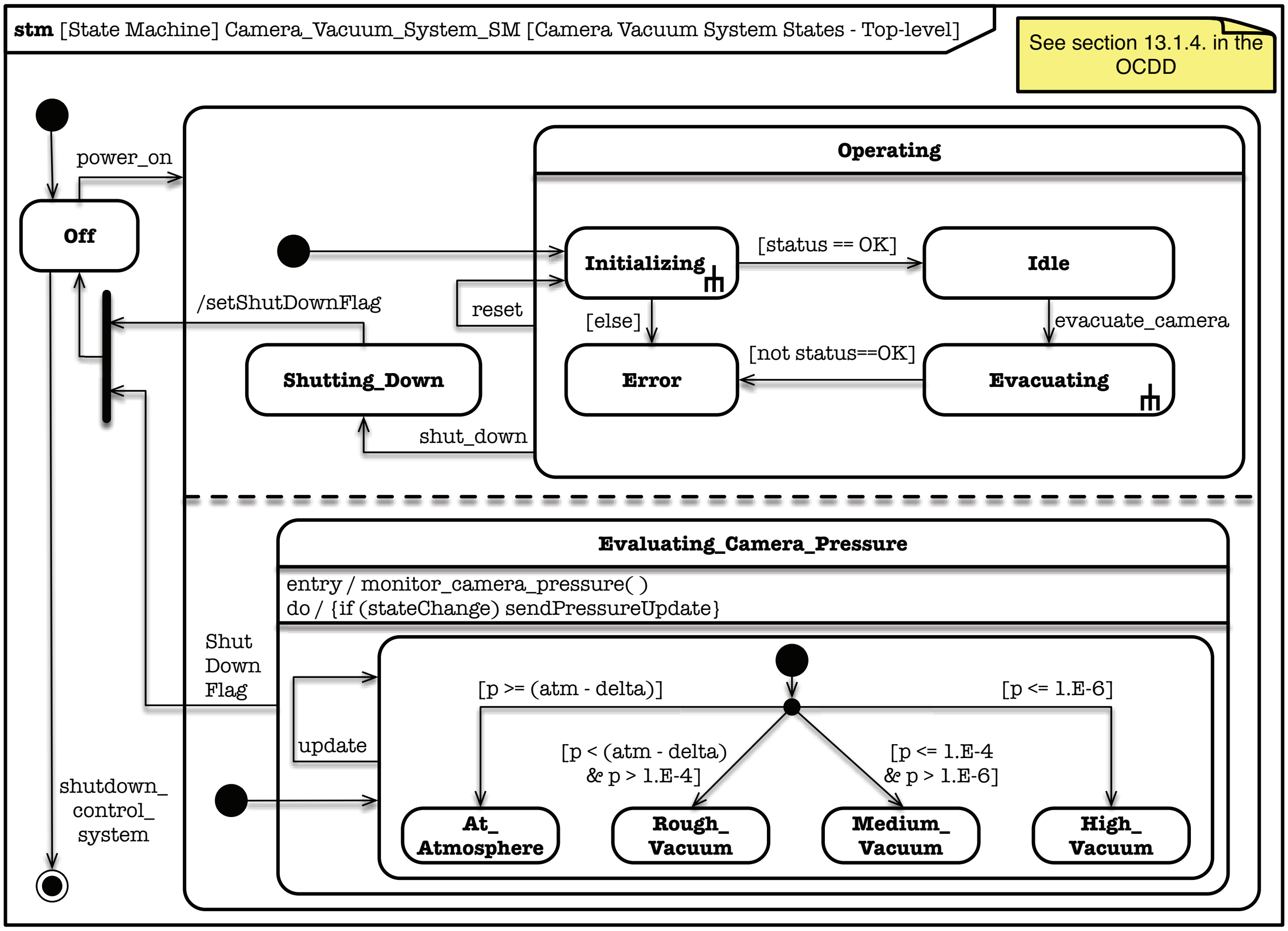


Figure 8. Preliminary G-CLEF camera vacuum system state machine diagram developed based on the *Evacuate Cameras* activity shown in Figure 7.

will form the basis for developing (internal or external) interface control documents between the subsystems. Figure 7 illustrates the preliminary activity diagram for the *Evacuate Cameras* procedure that is shown in Figure 4.

While the activity diagrams illustrate the overall flow of tasks required to complete an operational procedure, they do not specify how the individual actions that make up those tasks are sequenced. In most cases, multiple actions can occur in parallel, which is desirable for reasons of efficiency. This information is elucidated using the sequence diagram.

The state machine diagram is amongst the most helpful in establishing the system design. The system states associated with each activity, and the transitions between them necessary to complete the activity, are identified. The steps required to transition the system from one state to another will be further broken down as sub-activities that will each have their own associated state machine diagrams, states, and state transitions, and so on down to an atomic level that requires only simple transitions between well-defined states.

Some iteration may be required when multiple activities manipulate the same subsystem. Transitions that are apparently atomic in the case of one activity may be broken down further in order to achieve another activity. An additional complication in the case of G-CLEF is that the engineering design in some areas makes extensive use of Fail Safe over EtherCAT (FSoE, also known as "black channel") safety systems that can autonomously initiate or inhibit state transitions to prevent unsafe conditions. FSoE is implemented using programmable logic controllers and safety I/O modules that implement communications time-out actions directly in hardware and that are independent of the *Linux* device control computer. A software state machine model that is appropriately homologous to the associated hardware states must consider all of these factors, and may require multiple iterations to develop. Fault analysis and maintenance activities often require the ability to manipulate individual hardware components directly, and some consideration for this is required when developing the state machine descriptions. A preliminary state machine diagram for the camera vacuum system developed based solely on the *Evacuate Cameras* activity is shown in Figure 8.

3.2 Data Reduction Pipelines

Reduction of G-CLEF data from raw data to scientifically useful calibrated data products is managed by a pair of data reduction pipelines. The main pipeline is responsible for performing reductions that apply to data obtained in all observing modes. These steps include standard CCD data reductions, tracing and extracting spectral orders from the blue- and red-camera 2-D echellograms, order-by-order blaze and flat-field corrections, and creating a 2-D wavelength calibration that is then applied to the spectral orders. This pipeline may also be run at the telescope with partial calibrations in a "quick-look" mode to validate the quality of the observations in quasi-real-time. A second reduction pipeline is applied in addition to PRV mode data to provide a radial velocity estimate with high precision by comparing the target spectra with a reference spectrum. The reference spectrum can be either model-based (such as a stellar atmosphere model of the same spectral type) or observation-based (for example, the target spectra observed at a different epoch).

The data reduction pipelines are based on a set of echelle and PRV data reduction algorithms developed by Buchhave.[16] An initial set of software requirements was developed by reverse-engineering prototype echelle pipelines developed in IDL[†] by Buchhave over several years, to ensure that the latest updates to the algorithms are captured.

Figure 9 is the UML activity diagram for the standard CCD reductions phase of the G-CLEF data reduction pipeline resulting from prototype pipeline analysis. The CCD reductions are largely standard, and start with bad pixel identification and masking, trimming overscan regions, combining and subtracting bias frames, and combining and subtracting dark frames, if appropriate. Since we expect that the data reduction pipelines will run autonomously (except for anomaly resolution), each step includes quality assurance assessment if appropriate. For example, each of the "combine" steps ensures that all of the frames being combined are of good quality (*e.g.*, they are not saturated) and are compatible (*e.g.*, they have statistically similar counts distributions).

Once the CCD reductions are complete, the individual spectral orders are identified and extracted, as shown in Figure 10. Prior to extracting the spectral orders, scattered light is removed by fitting a smooth function to the inter-order data. Cosmic-rays are identified at this point using a series of tests that match the local counts distribution to a scaled flat field profile, evaluating adjacent pixel counts differentials, and comparing the counts to previously-established cosmic ray flux thresholds. Next, one dimensional spectra for each order are extracted from the 2-D echellograms using an optimal extraction technique with the same order parameters as those determined from the order-by-order flat field reductions. After the individual orders are extracted, a post-reduction process uses the set of wavelength calibration lamp observations to calculate the wavelength solution. The exact details of the post-reduction process depends in detail on whether the exposure included a simultaneous sky or simultaneous wavelength calibration spectrum in addition to the target (object) spectrum.

After the wavelength solution is computed, the data from the individual spectral orders can be rebinned and merged to produce a single one dimensional spectrum of the source, which may be the most appropriate reduced data product for some observing modes and scientific inquiries. However, the un-rebinned individual spectral order data products (and their associated wavelength calibrations) are always preserved since these will have the best wavelength calibration accuracy.

PRV mode observations can be further processed through the PRV reduction pipeline. This pipeline takes the wavelength calibrated target spectra from the standard reduction pipeline and produces a high precision radial velocity estimate by comparing the target spectra with a reference *template* spectrum. The template spectrum may be a synthetic model stellar spectrum, a pre-existing spectrum of the science target, or a user-created template. Multiple pre-existing spectra may be combined to form a super-template to improve the

---

[†] `http://www.harrisgeospatial.com/ProductsandSolutions/GeospatialProducts/IDL.aspx`

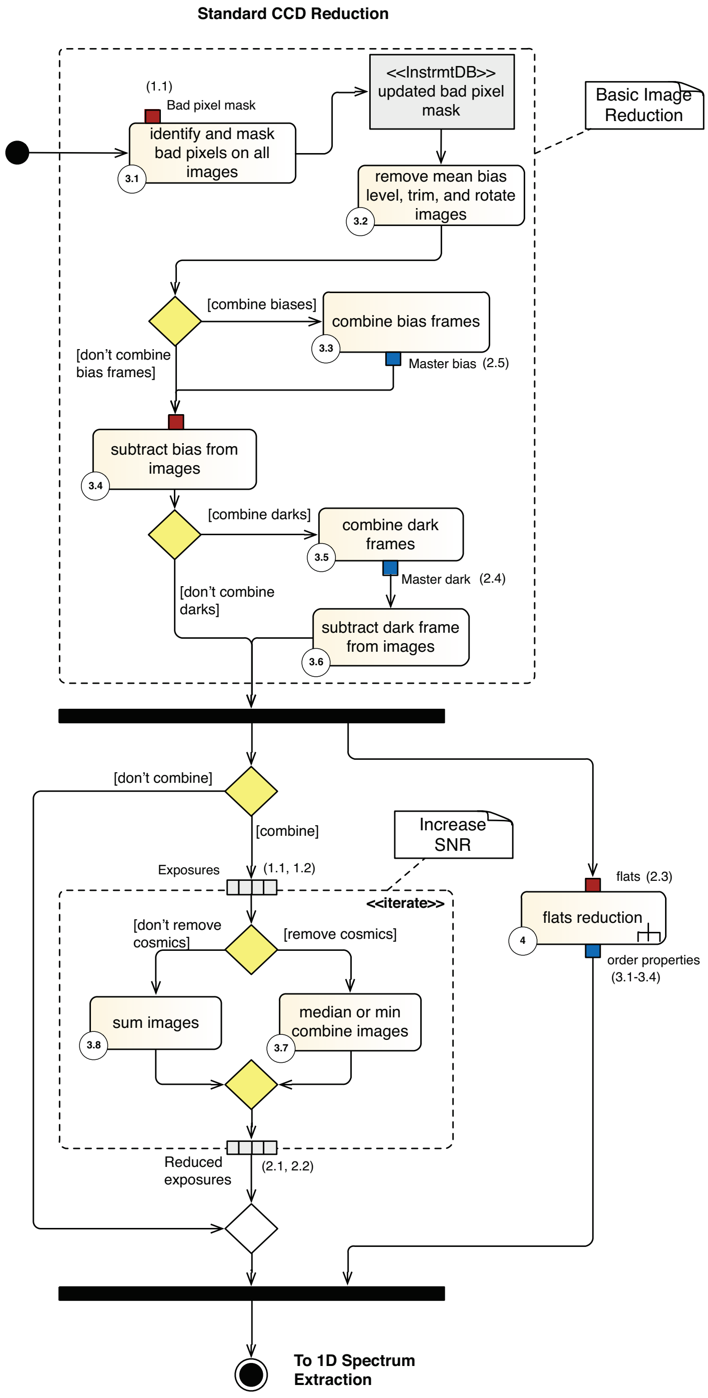


Figure 9. G-CLEF data reduction pipeline standard CCD reductions procedure.

signal-to-noise ratio of the reference spectrum. Combining multiple spectra will generally require that the wavelength calibration of each individual spectrum be corrected to solar system barycenter reference frame prior to combination, although alternative methods (*e.g.*, cross-correlation) may be used to ensure all of the spectra are wavelength-matched.

Depending on the spectral comparison technique used, additional steps may be required. For example, if a

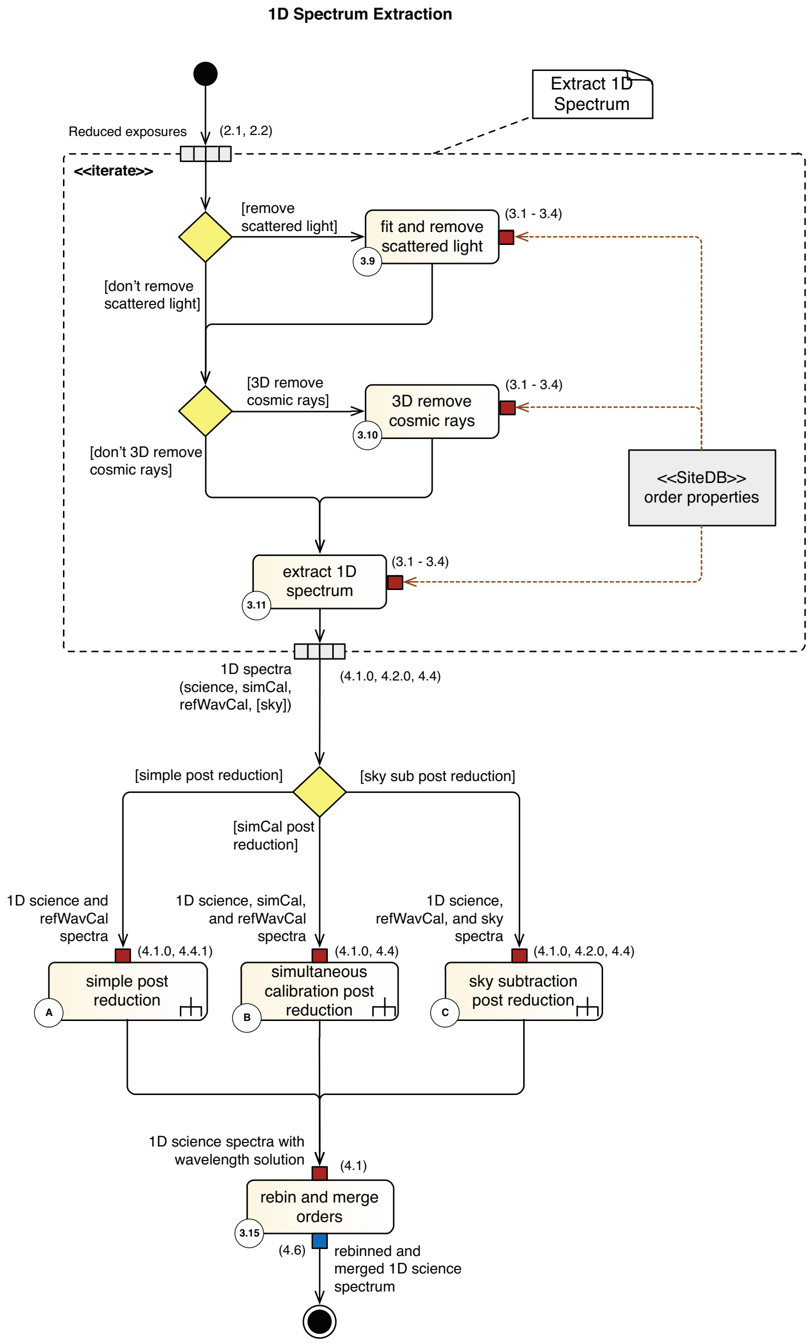


Figure 10. G-CLEF data reduction pipeline 1-D spectrum extraction procedure.

cross-correlation technique is used to compare the spectra, removing the normalized blaze function from the spectra, fitting and removing any continuum, and apodizing the ends of the spectra may improve the signal-to-noise of the cross-correlation peak significantly. Masking out regions of the spectrum that include undesirable features, for example telluric absorption lines from the Earth's atmosphere, will also improve the spectral comparison by removing contaminating signals. Similarly, some types of objects have strong spectral features only in a restricted waveband, and masking out the remainder of the spectrum will improve the comparison. If the object spectra have sufficient signal-to-noise, the radial velocity can be computed order by order. This has the advantage that the final mean radial velocity can be computed from the individual order radial velocities weighted by their individual errors. If this is not possible, the radial velocity can be computed for all of the orders simultaneously, which is equivalent to weighting by the orders' signal-to-noise values.

To ensure that the data reduction pipelines are optimized for G-CLEF data, G-CLEF project science conducted an extensive literature search to identify recently developed alternative algorithms that should be investigated further in some key areas. The first of these areas addresses improvements to optimal extraction algorithms that may be useful for order extraction. For example, Piskunov & Valenti[17] have employed clustering algorithms as an alternate means for tracing spectral orders and defining their profiles. Sharp & Birchall[18] have developed techniques for optimally extracting order profiles from tightly packed fibers that minimize crosstalk between the adjacent orders. This may be particularly relevant for observations using the G-CLEF PRV modes, since in these modes G-CLEF employs a pupil-sliced aperture consisting of 7 sub-apertures, resulting in spectral orders for the sub-apertures that are closely spaced together on the detector. We also plan to evaluate the "spectro-perfectionism" technique of Bolton & Schlegel,[19] which appears to have significant benefits as an extraction technique (although these may be more beneficial for lower signal-to-noise data). However, this technique may be too computationally intensive currently for the size of the echellograms produced by the G-CLEF detectors.

Alternative algorithms for precision radial velocity estimation will also be investigated as part of this evaluation. The cross-correlation technique to derive stellar radial velocities has been a standard in PRV studies for many years. However, forward-fitting (pixel-space) algorithms have been used in other wavebands for many years, and have more recently been used for lower-resolution optical spectroscopy studies for over two decades with good results. We plan to investigate whether the benefits of these techniques carry over to the high-resolution PRV regime in more detail by leveraging the knowledge and experience gathered developing the *Sherpa* multi-wavelength fitting engine.[20, 21]

The performance of alternate algorithms will be studied in the near future by using prototype implementations of the algorithms applied to a wide range of simulated G-CLEF data and comparing the results to the baseline prototype pipelines. The results of these studies will be used to revise the data reduction pipeline requirements and design as needed.

3.3 Proposal and Observation Preparation Tools

The OCDD identifies three G-CLEF-specific proposal and observation preparation tools. The first is an exposure time calculator (ETC) that can be used to compute the exposure time required to achieve a specified signal/noise (or vice-versa) from on a library of (simulated or observed) spectra, or using a spectrum supplied by the user. The second, an instrument spectral simulator that simulates the 2-D echellogram (*i.e.*, as detected by the science cameras) that can be processed using the G-CLEF data reduction system to create data products that may then be used to assess achievable radial velocity precision, equivalent width accuracy, and so on with a higher accuracy than is possible using the ETC. Finally, a flexure control star tool allows the observer to verify that acceptable offset guiding stars can be found in the field of view of the flexure control system, and also can be used to evaluate whether any stars present in the FCS field of view could rotate across the positions of the sky fibers during an exposure.

Similar to the data reduction pipelines, the top-level ETC software requirements were initially reverse-engineered from a science prototype. These top-level requirements were to develop more detailed derived requirements, and subsequently design and implement a production quality tool. Enhancement requests based on in-house usage experience of the latter by the G-CLEF project science team have been fed back to refine and develop additional requirements.

The requirements for the remaining proposal and observation preparation tools are currently being developed in collaboration with the G-CLEF project science team.

4. LESSONS LEARNED

Although the preliminary design of the G-CLEF software and controls system is still a work in progress, we have learned many valuable lessons that have general applicability to the wider instrumentation and software community.

The first lesson is to involve professional software and controls engineers early in the preliminary design of any instrumentation. This lesson is particularly relevant to instrumentation groups who have historically applied what may be best termed "minimal cost" development practices, where the focus is on the hardware design, and software is an afterthought and often the purview of graduate students or others who lack formal computer science or software engineering training. The software and controls systems being developed for state-of-the-art extremely large telescopes will certainly be using the most current computer science techniques and software practices, and instrumentation developed for these telescopes will have to conform to the observatory's requirements. Often, experienced software and controls engineers can work with the hardware engineering team to ensure that the control system design is not only compliant with imposed requirements but is also well-matched to the software architecture. Relatively minor design changes may dramatically simplify the software requirements and lead to substantial cost savings.

To ensure outstanding stewardship of investors' funds, the organizations developing and operating extremely large telescopes will, of necessity, require an approach to hardware and software development that has a degree of formality that approaches that of space-borne missions. This will certainly translate into a requirement for formal, traceable, and verifiable requirements (for both hardware and software), but the level of formality may well extend further, for example specifying requirements on systems modeling practices. For instrumentation groups lacking experience at this level, professional software and controls engineers should be at least somewhat familiar with these practices and can typically contribute to the overall group's capabilities in these areas.

A second lesson is to task an individual with overall responsibility for science/software/hardware integration. This individual is not the party responsible for the instrument design, or for the development of the software. Instead, this person is responsible for ensuring that the integrated instrument hardware and software will provide the science capabilities, and be responsive to the science goals, for which the instrument is being developed. This task optimally requires someone with practical experience in all of the areas, but good software and hardware focus is essential. Much of this individual's time will ultimately be occupied developing software requirements and systems modeling, but initially their focus should be on developing a robust OCDD.

The third lesson is to develop a detailed OCDD (or equivalent) that includes all of the routine and non-routine operations sequences, complete to the level of detail that is appropriate given the state of the design at any time. The G-CLEF OCDD has become a critical document that ties together the areas of science, hardware, and software, and the sequences help to flesh out any gaps in the design or understanding of how the system operates.

From the science point of view, the OCDD describes the many ways in which observations can be performed, including detailed science and calibration exposure sequences, and target acquisitions. Missing capabilities are immediately evident from a review of the OCDD sequences. Careful review of the G-CLEF OCDD observing sequences revealed enhanced functionality that could be provided with just minor changes to the sequences.

The act of developing the OCDD sequences is helpful in exposing limitations in the current state of the hardware design, for example by identifying procedures that may require further automation to facilitate efficient operations, or by highlighting missing functionality. Detailed *non*-routine operations sequences, such as instrument alignment or subsystem maintenance procedures, may be most helpful in identifying missing functionality, since these procedures are generally not the focus when developing the initial design.

For the software and controls engineers, the OCDD provides a bridge between a functional representation of the interfaceable hardware and the operations sequences that must be managed by the software to complete each specific task. Although the OCDD does not specify every step required to manipulate the hardware (for example, stating "Read the `sp_level` inclinometer angles" rather than listing the steps required to execute the "read" action), the sequences do provide the information necessary to construct the requirements and system models that are the next steps towards developing the software design.

One final benefit of developing the OCDD sequences early is that they identify preliminary interfaces with other systems (*e.g.*, the telescope control system) and make explicit any assumptions about how these other systems interact. These assumptions can then be verified with the other systems. This aspect of the OCDD will be very beneficial when developing interface control documents between the instrument software and those systems.

A fourth lesson is that optimal time phasing of the software and controls requirements and preliminary design process may not be possible in practice, and planning for parallel or out-of-sequence development should be carefully considered. The primary inputs driving the G-CLEF software preliminary design are those that come from the instrument hardware design. Requirements imposed at a higher level can also drive the software design. The G-CLEF software requirements and preliminary design are highly dependent on the software and controls reference architecture, and software frameworks, that are being defined and developed by GMTO. While the software and controls design would ideally be developed only after all of the various inputs that drive the design stabilize, in practice the G-CLEF software requirements and preliminary design are being developed in parallel with the GMTO's own software and controls architecture efforts. This is partly because of the relative project schedule phasing between G-CLEF and the GMT, and partly because of the significant control system prototyping efforts required to validate the active thermal control loop. Therefore, revisions to the G-CLEF software requirements and preliminary design are inevitable. We have managed these revisions in two ways.

First, we identify snapshots in time of the driving inputs that we attempt to capture when revising the software requirements or preliminary software design. For example, we initially used the hardware design presented at the engineering preliminary design review (PDR) as a reference, and develop to that snapshot. As the instrument design is refined further, we identify key updates that warrant revisions to the software requirements or design, for example after a major revision to the vacuum system design or thermal control system design. In this way we are incrementally improving the fidelity of the software requirements and system model, while always having a baseline hardware design against which to develop the software design. As the various elements of the hardware design stabilize approaching the engineering critical design review (CDR), the software requirements and system model will converge to the final design also.

Revisions to the GMTO SWCS are handled similarly, by assuming a baseline and incrementally updating our requirements and design. This approach is somewhat less effective in cases where support software was required for laboratory pathfinder studies of various hardware options. In some circumstances, software tools have had to be developed that did not utilize the GMTO-developed frameworks, in order to maintain the schedule. Up to now, this has been a relatively infrequent occurrence and is an inevitable side-effect of parallel development.

Second, we attempt to time phase the various development efforts so that we work on the most stable tasks first. The G-CLEF data reduction pipeline requirements stabilized relatively early in the program because of our significant heritage in this area. As a consequence, we focused a significant fraction of our efforts on moving the data reduction pipeline preliminary design forward first. Although not all of the outstanding G-CLEF-specific pipeline questions could be addressed before completion of the final instrument optomechanical design, the vast majority of the pipeline preliminary design was able to be completed successfully. A relatively small number of open questions remain to be addressed, including the alternate algorithms discussed above.

By phasing the software development in this manner we are able to make efficient use of resources in areas where there are tasks that can be completed, while freeing up resources for latter use on tasks that are currently blocked. We expect to hold a CDR on the data reduction pipelines and associated tools on the same timeframe as the G-CLEF instrument CDR,[‡] and an instrument control software PDR shortly thereafter.

ACKNOWLEDGMENTS

The authors acknowledge the outstanding support of the G-CLEF science and hardware engineering teams, and the GMTO software and controls and systems engineering teams.

REFERENCES

[1] Szentgyorgi, A. et al., "The GMT-CfA, Carnegie, Catolica, Chicago Large Earth Finder (G-CLEF): a general purpose optical echelle spectrograph for the GMT with precision radial velocity capability," in [*Ground-based and Airborne Instrumentation for Astronomy IV*], McLean, I. S., Ramsay, S. K. R., and Takami, H., eds., *Proc. SPIE* **8446**, 84461H, SPIE, Bellingham, WA (2012).

[‡] Nominally May, 2017.

[2] Szentgyorgi, A. et al., "A preliminary design for the GMT-Consortium Large Earth Finder (G-CLEF)," in [*Ground-based and Airborne Instrumentation for Astronomy V*], Ramsay, S. K., McLean, I. S., and Takami, H., eds., *Proc. SPIE* **9147**, 914726, SPIE, Bellingham, WA (2014).

[3] Szentgyorgi, A. et al., "The GMT-Consortium Large Earth Finder (G-CLEF): An optical echelle spectrograph for the Giant Magellan Telescope (GMT)," in [*Ground-based and Airborne Instrumentation for Astronomy VI*], Evans, C. J.and Simard, L. and Takami, H., eds., *Proc. SPIE* **9908**, 9908–76, SPIE, Bellingham, WA (2016).

[4] Fischer, D. A. et al., "State of the Field: Extreme Precision Radial Velocities," *Pub. ASP* **128**, 066001 (2016).

[5] INCOSE, "Systems Engineering Vision 2020," Tech. Rep. INCOSE-TP-2004-004-02, Version 2.03, INCOSE, San Diego, CA (2007).

[6] Filgueira, J. M., Bec, M., Soto, J., Liu, N., and Peng, C. Y., "GMT software and controls overview," in [*Software and Cyberinfrastructure for Astronomy II*], Radziwill, N. M. and Chiozzi, G. R., eds., *Proc. SPIE* **8451**, 11, SPIE, Bellingham, WA (2012).

[7] Karban, R. et al., "Model Based Systems Engineering for Astronomical Projects," in [*Modeling, Systems Engineering, and Project Management for Astronomy VI*], Angeli, G. Z. and Dierickx, P., eds., *Proc. SPIE* **9150**, 91500L, SPIE, Bellingham, WA (2014).

[8] Saunders, W. et al., "MANIFEST: a multi-instrument fiber positions system for GMT," in [*Ground-based and Airborne Instrumentation for Astronomy IV*], McLean, I. S., Ramsay, S. K., and Takami, H., eds., *Proc. SPIE* **7735**, 773568, SPIE, Bellingham, WA (2010).

[9] Goodwin, M. et al., "MANIFEST instrument concept and related technologies," in [*Ground-based and Airborne Instrumentation for Astronomy IV*], McLean, I. S., Ramsay, S. K., and Takami, H., eds., *Proc. SPIE* **8446**, 84467I, SPIE, Bellingham, WA (2012).

[10] Filgueira, J. M., "GMT Software and Controls Standards," Tech. Rep. GMT-SWC-REF-00029, GMTO Corporation, Pasadena, CA (2015).

[11] Szentgyorgi, A. and Podgorski, W., "G-CLEF System Requirements," Tech. Rep. SAO-INST-REQ-00015, Smithsonian Astrophysical Observatory, Cambridge, MA (2014).

[12] Evans, I. N., "G-CLEF Instrument Device Control Subsystem Requirements," Tech. Rep. SAO-INST-REQ-00039, Smithsonian Astrophysical Observatory, Cambridge, MA (2014).

[13] Plummer, D. A., "G-CLEF Instrument Device Control Subsystem Sofwtare Requirements," Tech. Rep. SAO-INST-REQ-00094, Smithsonian Astrophysical Observatory, Cambridge, MA (2016).

[14] Evans, I. N., Evans, J. D., and Budynkiewicz, J. A., "G-CLEF Operations Concept Definition Document," Tech. Rep. SAO-INST-DOC-00065, Smithsonian Astrophysical Observatory, Cambridge, MA (2016).

[15] Frebel, A. and Szentgyorgyi, A., "G-CLEF Science Requirements Document," Tech. Rep. SAO-INST-REQ-00006, Rev 5, Smithsonian Astrophysical Observatory, Cambridge, MA (2015).

[16] Buchhave, L. C. A., *Detecting and Characterizing Transiting Extrasolar Planets*, PhD thesis, Niels Bohr Institute, University of Copenhagen (2010).

[17] Piskunov, N. E. and Valenti, J. A., "New Algorithms for Reducing Cross-Dispersed Echelle Spectra," *Astron. Astrophys.* **385**, 1095 (2002).

[18] Sharp, R. and Birchall, M. N., "Optimal Extraction of Fibre Optic Spectroscopy," *Pub. Astron. Soc. Aust.* **27**, 91 (2010).

[19] Bolton, A. S. and Schlegel, D. J., "Spectro-Perfectionism: An Algorithmic Framework for Photon Noise-Limited Extraction of Optical Fiber Spectroscopy," *Pub. ASP* **122**, 248 (2010).

[20] Doe, S. et al., "Sherpa: Goals and Design for Chandra and Beyond," in [*Astronomical Data Analysis Software and Systems XV*], Gabriel, C., Arviset, C., Ponz, D., and Solano, E., eds., *ASP Conf. Ser.* **351**, 77, ASP, San Francisco, CA (2006).

[21] Laurino, O., Siemiginowska, A., Evans, J., Aldcroft, T., Burke, D. J., McDowell, J., McLaughlin, W., and Nguyen, D., "Sherpa. On the Move to Open, Collaborative Development," in [*Astronomical Data Analysis Software and Systems XXIV*], Taylor, A. R. and Rosolowksy, E., eds., *ASP Conf. Ser.* **495**, 313, ASP, San Francisco, CA (2015).