Designing software for the science operations of Maunakea Spectroscopic Explorer

Szeto, Kei, Marshall, Jennifer, Flagey, Nicolas, Withington, Kanoa, Surace, Christian

Kei Szeto, Jennifer L. Marshall, Nicolas Flagey, Kanoa Withington, Christian Surace, "Designing software for the science operations of Maunakea Spectroscopic Explorer," Proc. SPIE 11452, Software and Cyberinfrastructure for Astronomy VI, 1145207 (13 December 2020); doi: 10.1117/12.2562846

Event: SPIE Astronomical Telescopes + Instrumentation, 2020, Online Only

Designing software for the science operations of Maunakea Spectroscopic Explorer

Kei Szeto[*ab], Jennifer L. Marshall[c], Nicolas Flagey[d], Kanoa Withington[a], Christian Surace[e]
[a]CFHT Corporation, 65-1238 Mamalahoa Hwy, Kamuela, Hawaii 96743, USA
[b]National Research Council Canada, Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC, Canada, V9E 2E7
[c]Mitchell Institute for Fundamental Physics and Astronomy and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843-4242 USA
[d]Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[e]Aix Marseille Univ, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France

ABSTRACT

The Maunakea Spectroscopic Explorer (MSE) will transform the Canada-France-Hawaii Telescope into an 11.25-m aperture telescope, dedicated to highly multiplexed, visible to near-IR spectroscopic studies with multiple spectral resolution modes. A metric of MSE's success is survey speed, i.e. how many scientifically useful spectra MSE will obtain in support of its surveys, which requires hardware and software to be designed and perform efficiently. In this paper, we describe the front-end software, which includes proposal review, a scheduler, an exposure time calculator, and a breaker to prepare and define the survey observations, and the back-end software, which includes data reduction and science pipelines, science archive, and science platform to deliver the data back to the science community. The interfaces, the flow of data, and the overarching object model will be explained. We also discuss the tools required to support the Design Reference Survey that describes and simulates the science operations of MSE.

Keywords: spectroscopy, survey, science operations, data product, data reduction, pipelines, scheduler

1. INTRODUCTION

Maunakea Spectroscopic Explorer (MSE) is the first of the future generation of massively multiplexed spectroscopic facilities. MSE is designed to enable transformative science, being completely dedicated to large-scale multi-object spectroscopic surveys, each studying thousands to millions of astrophysical objects. MSE uses an 11.25 m aperture telescope to feed 4,332 fibers over a wide 1.52 square degree field of view. It will have the capabilities to observe at a range of spectral resolutions, from R~3,000 to R~40,000, with all spectral resolutions available at all times and across the entire field. As a dedicated survey facility, the success of MSE will link to the quantity and quality of observations it obtains per night, per year, and ultimately over its lifetime.

In this paper, we propose an Operations Model and the supporting Program Execution System Architecture (PESA) concept to execute MSE survey programs in successive phases, from proposal submission to the distribution of the data. PESA is the end-to-end high-level software suite that directs remote science operations[†] at the MSE observatory beginning from proposal submission and ending at the distribution of science products. PESA also contains the processing and analytical tools along with the associated user interfaces for the MSE staff and scientific community. PESA will allow timely delivery of reduced and validated data products via a science portal that facilitates processing and publishing of their observational discoveries for the MSE users.

[*]Email: szeto@cfht.hawaii.edu; Telephone: 236-464-5433; Fax 808-885-7288
[†]Remote science operations imply no physical presence of telescope operators on-site at the summit of Maunakea and observations are conducted by remote observers from Waimea headquarters, which is two hours from the telescope by car.

Software and Cyberinfrastructure for Astronomy VI, edited by Juan C. Guzman, Jorge Ibsen, Proc. of SPIE Vol. 11452, 1145207 · © 2020 SPIE · CCC code: 0277-786X/20/$21 · doi: 10.1117/12.2562846

The PESA architecture follows the representative workflow of a ground-based observatory, Figure 1, where the workflow is divided into five sequential phases in order to produce the reduced and validated science data products envisioned for MSE user community:

1. Observing programs are selected via proposal announcement and selection process (Phase 1);
2. Target definitions of approved proposals are supplied along with the instrument configurations specified (Phase 2);
3. Observations are scheduled and executed (Phase 3);
4. Data are reduced with quality analysis performed (Phase 4); and
5. Data products are curated and distributed to the MSE community according to access rights (Phase 5).

Figure 1 Observing sequence to produce reduced and validated science data products

Figure 2 is a block diagram showing the ten PESA products and their workflow with respect to the observing sequence shown in Figure 1. The PESA products are separated into two operational domains, pre and post observation, separated by the product Observatory Execution System Architecture (OESA) that contains the different observatory control systems, hardware and software, and MSE's IT infrastructure. Organized by their functionalities, the PESA products are divided into three groups: Survey Preparation and Definition (SPD), Data Reduction Pipelines (DRP) and Science Archive and Platform (SAP). The products in the SPD group completes the pre-observing operations where the selected survey proposals are converted into observing commands that are understood by the OESA Observatory Control Sequencer. MSE will observe survey programs contemporaneously from an integrated target list, which contains all target definitions, managed by the Object Model (ObjMod). SPD encompasses the work of Phase 1 Proposal Selection and Phase 2 Targets Definition. The products in the DRP and SAP groups generate, validate and then deliver the final MSE science data products by processing the spectrographs' science detector readout collected from the Observations phase, Phase 3, via OESA. Individually, DPR contains the work of Phase 4 Reduction/Validation, and SAP contains the work of Phase 5 Distribution. The functionalities of each PESA product are discussed in the following sections.

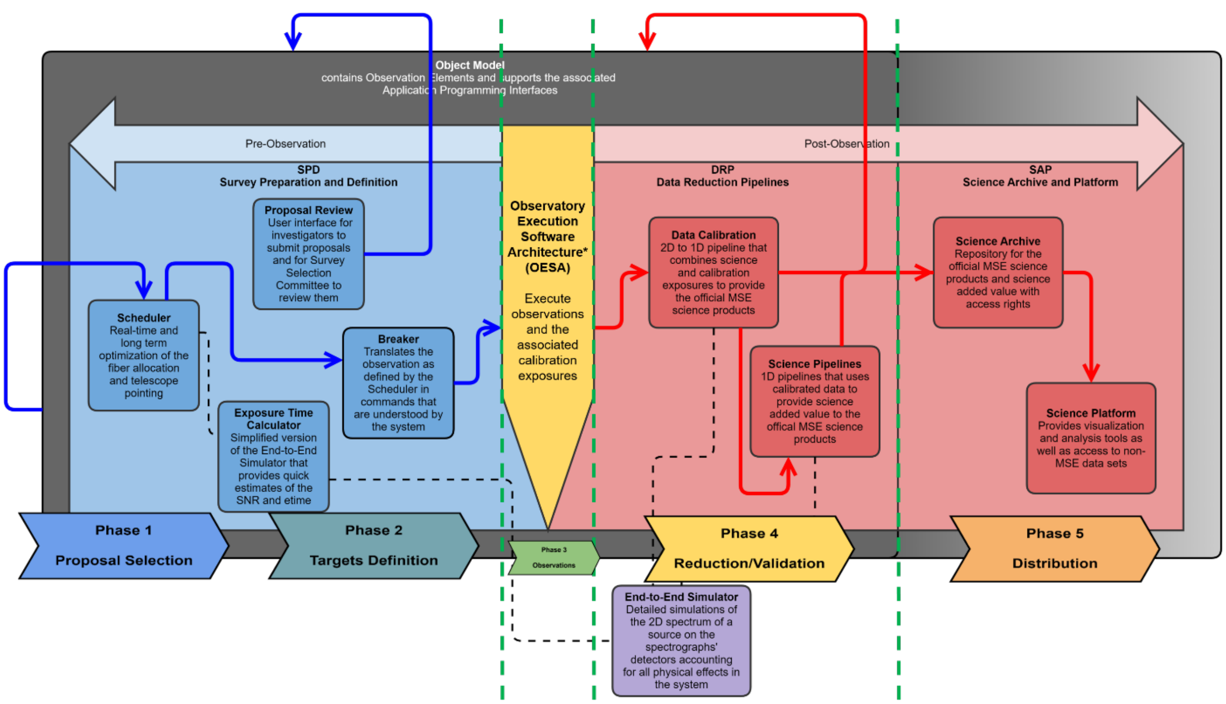

Figure 2 PESA block diagram illustrating the workflow with respect to the observing sequence, and showing the functions and product breakdown with data flow in the Pre-Observation domain (blue) and Post-Observation domain (red). Arrows indicate the flow of data and interfaces between the two domains facilitated by the Observatory Execution Software Architecture product (yellow), and linked by the Object Model.

PESA also includes the user interfaces and Application Programming Interfaces (APIs) to enable the MSE staff and scientific user community, survey teams and science teams, to perform science operations, such as design and submit surveys, schedule and verify observations, reduce and validate data, and finally distribute and access science products, etc. However, PESA does not include the computing equipment and network infrastructure, hardware, firmware and utilities, required to support the processing, reduction, archival and distribution of MSE's science products. The associated MSE product Observatory Execution System Architecture provides the computing equipment and network infrastructure at the summit and Waimea headquarters, and the Observatory Building Facilities provide the control room and related utilities. OESA is responsible for defining communication and control protocol required for PESA to execute daytime and nighttime science operations within the observatory system.

2. PESA PRODUCTS FOR PROPOSED OPERATIONS MODEL

Figure 3 shows the Product Breakdown Structure (PBS) with all six top-level MSE products, and the lower-level product branches within the Science Instrument Package, OESA and PESA expanded. The PESA products are conceptualized based on the operational workflow envisaged in the Operations Concept Document[1] according to MSE's Operations Model. Illustrated as blue and red arrows in Figure 2, they illustrate the information flow between pre-observations and post-observations according to the progression of the survey programs via the Object Model data structure.

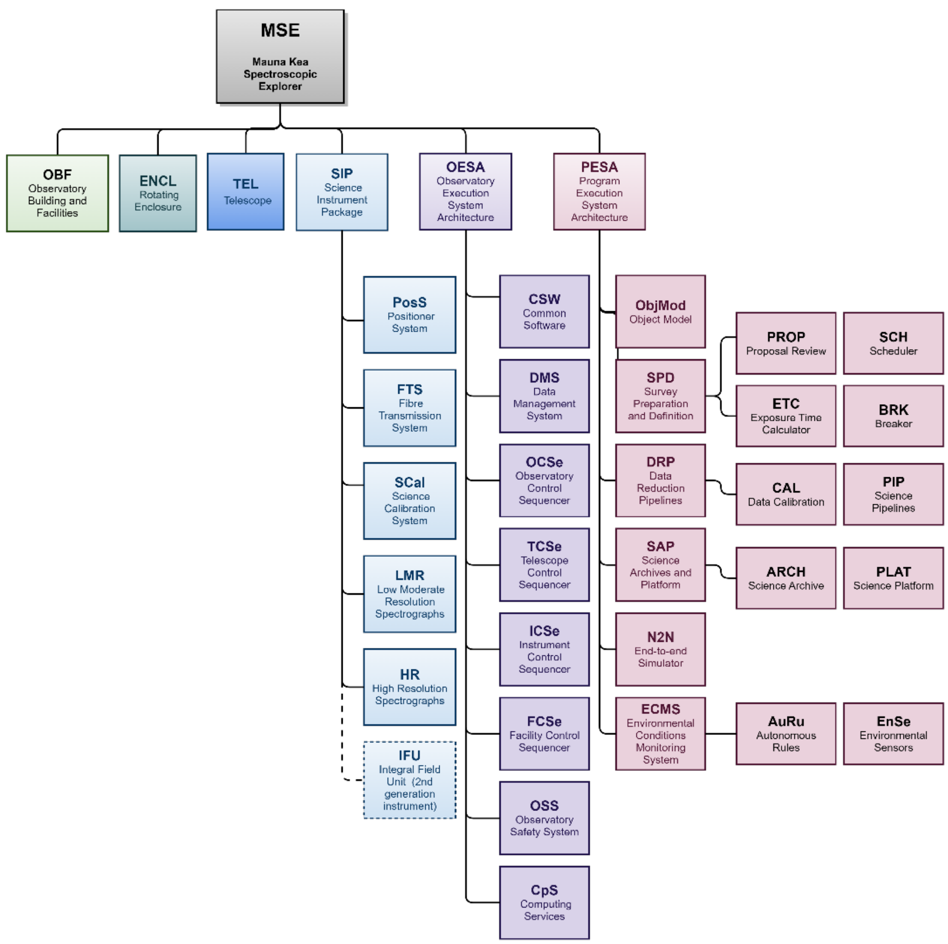

Figure 3 MSE PBS with top-level products and lower-level products under SIP, OESA and PESA

The MSE Operations Model, and the PESA products and their core functionalities are described in the following sections.

2.1 Operations Model

There are two categories of surveys:
Large Surveys: 80% of the available observing time annually is allocated to Large Surveys. These surveys represent collaborative scientific programs involving many scientists across multiple MSE member institutions. Large Surveys will be driven by primary science goals and require extensive observing time per year, over several years.
Small Surveys: 20% of the available observing time annually is allocated for Small Surveys. These surveys have more limited science goals, smaller scope and shorter duration than Large Surveys.

There are two suites of spectrographs:
Low/Moderate Resolution (LMR) Spectrographs: 75% of MSE's 4,332 fibers are allocated to a set of six LMR spectrographs, where each spectrograph has three visible arms, blue, green and red, and one NIR arm. In operations, each arm of each LMR spectrograph can observe in either low or moderation resolution independently while under a fixed fiber-to-target allocation, telescope pointing, and overall integration time. However, each arm of each spectrograph can expose and readout repeatability under the overall observing time.
High Resolution (HR) Spectrographs: 25% of MSE's 4,332 fibers are allocated to a set of two HR spectrographs, where each spectrograph has three visible arms, blue, green and red. In operations, each HR spectrograph can observe in different band-passes independently while under a fixed fiber-to-target allocation, telescope pointing, and overall integration time. However, each arm of each spectrograph can expose and readout repeatability under the overall observing time.

Given the survey categories and spectrograph suites, the Operations Model must be able to:

- Process and accept proposals at varying frequency of solicitations, where proposals are submitted by survey teams and reviewed by a Survey Selection Committee;
- Execute and manage combined Large and Small Surveys concurrently, while accommodating individual surveys with diverse scopes and durations, using different resolutions at different spectral coverages at the individual spectrograph level; and
- Provide reduced and validated science data products according the access rights within the MSE user community.

The ten products under PESA are designed to support MSE's Operations Model, they are:

- Object Model, PESA.ObjMod
- Survey Preparation and Definition, PESA.SPD
  - Proposal Review, PESA.SPD.PROP
  - Scheduler, PESA.SPD.SCH
  - Exposure Time Calculator, PESA.SPD.ETC
  - Breaker, PESA.SPD.BRK
- Data Reduction Pipelines , PESA.DRP
  - Data Calibration, PESA.DRP.CAL
  - Science Pipelines, PESA.DRP.PIP
- Science Archive and Platform, PESA.SAP
  - Science Archive, PESA.SAP.ARCH
  - Science Platform, PESA.SAP.PLAT
- End-to-end simulator PESA.N2N
- Environmental Conditions Monitoring System, PESA.ECMS
  - Environmental Sensors, PESA.ECMS.EnSe
  - Autonomous Rules, PESA.ECMS.AuRu

2.2 Glossary of Operations Model

Each survey program contains a set of targets information that is associated with its Observing Elements (ObSEles) in the Object Model. Each ObsEle defines the target and its requisite observing conditions, e.g., instrument configuration, time constraints, signal-to-noise ratio to reach. Once a survey program's observation has commenced, scientific data are extracted from the science detectors, and pixel values are processed by applying calibration and reduction techniques. Through the observing process, the status of each ObsEle is tracked and its science information is updated continuously until its observation objectives are met, e.g., signal-to-noise ratio. Then the data are released to the MSE community through the Science Archive and Platform.

There are four essential elements for the proposed PESA concept:

Observing Element (ObsEle): Each target is associated with an ObsEle that contains the required information about the proposed observations. Each ObsEle is defined by the following set of parameters:

- Unique set of coordinates, e.g., RAJ2000 and DECJ2000 with respect to a common reference frame such as GAIA;
- Ephemeris for proper motion objects[‡];
- Spectrograph setting for spectral resolution (low, moderate, or high resolution) and spectral coverage required[§];
- Non-standard calibrations required for that observation;
- Metric for validation, e.g., Signal to Noise Ratio (SNR) value and magnitude at a given wavelength;
- Status flag, with value such as Unobserved, Started or Completed;
- Survey program identification number;
- Priority, internal to the observing program, between 0 (lowest) and 9 (highest);
- Optionally, a set of Time constraints, e.g., start, finish, repeats, etc.;
- Optionally, a set of Weather constraints, e.g., sky background, seeing etc.; and
- Optionally, a set of Setting constraints, e.g., observe at the same time as another target, etc.

Detailed descriptions of the parameters are listed in **Appendix A**.

Observing Field (ObsFld): An ObsFld corresponds to a collection of Observing Elements, with each ObsEle allocated to a fiber, that can be observed together because they all fit within the science field of view under a single configuration of fiber-to-target allocation and telescope pointing. An ObsFld is selected and built from the content of Object Model. In addition, an ObsFld also contains fiber positions allocated for sky background, reference and calibration stars measurements.

Observing Sequence (ObsSeq): An ObsSeq describes the collection of the numbers and durations of all exposures for all arms of all spectrographs. An exposure is the time during which a detector collects photons, until readout. An ObsSeq is required since each arm of each spectrograph may have different exposure times. Since there are two suites of multiple spectrographs providing LR and/or MR, and HR resolutions, and each spectrograph is divided into separate wavelength arms, the ObsSeq enable each arm of each spectrograph to be controlled and exposed independently.

Observing Matrix (ObsMat): An ObsMat is a collection of one ObsFld and one ObsSeq, generated by the Scheduler, to form a unique block of observations. An ObsMat combines the target locations in the sky, i.e., ObsFld, with the exposure times for each arm of each spectrograph, i.e., ObsSeq.

2.3 Object Model Product

The Object Model product is the backbone connecting the Pre-Observation and the Post-Observations domains, and supports data exchanges during science operations. Conceptually, the ObjMod is an abstraction of the traditional data repositories utilized for tracking of target definitions (Phase 1 and 2), collecting observational information and status (Phase 3), and coordinating science data processing (Phase 4). ObjMod also links with the environmental and

[‡]Although MSE will not be able to track fast moving targets, slow moving targets should in principle be accessible
[§]Same coordinates targets with different spectral resolutions and coverages specifications required separate ObsEles

engineering data in the OESA Data Management System (DMS) and its system-level status sever in order for MSE's scientific and technical staff to assess, analyze and troubleshoot system performance during science operations.

Operationally, ObjMod contains information in the form of Observing Elements for every MSE target, and tracks the targets' observing status and their science data as they progress through the survey programs. ObjMod also defines the API protocol that enables interfacing with other PESA products. The MSE staff and science community will develop their own APIs in order to input, remove and modify ObsEles. A supplemental interface is also required from Object Model in order to synchronize with the OESA Data Management System, specifically communication with the observatory system-level status server.

Core Functionalities:
- Curate Observation Elements and supports the associated Application Programming Interfaces
- Accept ObsEle definitions from Proposal Review to compile the target list
- Provide Scheduler with targets and their status from target list
- Accept target information update from Data Calibration
- Links ObsEle science data with environmental and engineering data in the OESA DMS and its status sever

2.4 Proposal Review Product

Survey Preparation and Definition refers to the step and tools by which survey teams prepare proposals, a.k.a. survey programs, to observe large samples of targets, and with their scientific and technical justifications verified. Proposal Review (PROP) collects and processes survey programs via call for proposals that is coordinated by the observatory and selected by the Survey Selection Committee (SSC) at regular intervals. Depending on the type of surveys, Large or Small, the selection could happen every few months or every few years.

PROP automates and streamlines the proposals submission and review process between the survey teams and SSC. PROP curates proposals and verifies the sample target definitions and their science and technical justifications submitted in the Object Model. Once a proposal is approved, it facilitates Observing Elements upload and tags target information in the ObjMod for the Scheduler of those programs that have been approved for observations.

To facilitate the scheduling of observations and avoid duplications, we envision each survey team will define their sample of targets in a clear and concise way as ObsEles via the ObjMod process. Given the expected number of targets a survey program requests to observe, we do not envision that an exhaustive target set is required, initially. Instead, there will be provision for each submitted program to broadly define their target sample using a list of criteria, e.g., boundaries of a region in the sky, magnitude range, and object type. These criteria are later used to fill the target set once the survey program is approved,followed by automatic verification of each observing target.

Core Functionalities:
- Workspace where proposals are defined, received, evaluated, and selected
- Communication interface that processes proposals between the survey teams and SSC
- Provide proposal library to support the proposal assessment and selection process
- Enable survey teams to query their proposal status and receive notifications from the SSC
- Provide Object Model interface to facilitate target and observing conditions definition upload
- Check ObsEles for duplication and verify scientific and technical justifications

2.5 Scheduler Product

Efficient scheduling of the observations is critical for the success of MSE. The primary objective of the Scheduler (SCH) product is to optimize science observations by taking into account the overall priorities and requirements of the combined survey programs, while taking into account all the observational constraints, e.g., weather, current spectrographs configuration, and considering both short-term and long-term schedules. i.e., in the upcoming night and upcoming year. The creation of Observing Matrices, which corresponds to choosing ObsEles, defining an ObsFld and choosing instrument configuration and exposure times sequences to define an ObsSeq, with optimization is SCH's primary objective

Due to the complexity from the vast number of targets and the large number of fibers in the field of view, SCH autonomously generates ObjMats by analyzing information contained in the ObjMod, current and forecasted observing conditions, current and possible instrument configuration[**], remaining time available for science in the night, etc. During science operations, SCH defines the best next ObsMat and sequence of ObsMats continually.

The scheduling process is the most complex operation, and it runs concurrently with the Data Reduction Pipelines, where multiple threads and parallel access to the ObjMod are envisioned. SCH uses the ObjEle status information and the estimated time needed to complete observations predicted by the Exposure Time Calculator (ETC) under current observing conditions to maintain optimal observing schedule. The SCH also receives real-time feedback from the Data Reduction Pipelines, which provides ObsEle statuses post observationally, to adjust the exposure times dynamically. In addition, SCH provides a user interface for MSE scientific staff to modify its ObsMat, e.g., adjusting weighting factors in the optimization algorithm to repeat ObsEles that have already been observed or balancing the scientific priorities with LR, MR and HR spectroscopy due to ranking changes in the survey programs, etc.

Core Functionalities:

- Provide real-time and long-term optimization of the fiber allocation and telescope pointing
- Allocate targets to fibers[††], determine the instrument configuration and exposure sequence, and schedule the telescope pointing[‡‡];
- Queue ObsMat, i.e., send commands to the Breaker that interfaces with the OESA observatory control system to execute the planned ObsMat, including the associated nighttime and daytime calibration exposure frames;
- Monitor observing conditions via OESA Data Management System and utilize ETC to adjust exposure time dynamically
- Check target completion based on ObsEles signal-to-noise criteria in real-time while monitoring feedback from CAL to update the current ObsSeq, and the upcoming sequence of ObsMats to maintain an optimal forward-looking schedule;
- Receive and respond to ObsMat action status[§§] from the Breaker; and
- Provide user interface for MSE scientific staff, including remote observers, to query, monitor and manage the SCH actions

Additional SCH design considerations:

- Ability to modify scheduling optimization algorithm, e.g., providing user adjustable weighting factors
- Interface with the OESA DMS and its status server to access subsystems' telemetry
- Overall integration time to limit differential atmospheric refraction across the telescope focal surface and maintain the fiber injection efficiency[2]; and
- Accommodate Targets of Opportunity (ToO) observation for targets obtained from other observatories, e.g., LSST, by having the ability to incorporate targets that are defined on an as-needed basis nightly, where real-time modifications to ObsEles within ObjMod are possible as long as the affected ObsEles have not been queued for observation.

2.6 Exposure Time Calculator Product

The Exposure Time Calculator (ETC) is a SNR and exposure time-based algorithm that estimates system performance by including the target magnitude, sky characteristics, and subsystems performance with respect to throughput and noise contributions, etc. ETC is modular in design in order to facilitate revisions and future upgrade.

[**]Including considerations for spectrographs configuration, and remaining range of motion limits of all subsystems, e.g., remaining cable wrap travels of the Instrument Rotator and Telescope, in order to filter out ObsEles that cannot be observed and/or impose additional time constraints on ObsSeq

[††]Including for fibers for sky, reference and calibration stars measurements

[‡‡]Including considerations for available guide, reference and calibration stars, and voidance zones for bright objects in the field of view

[§§]Breaker action status for SCH's ObsMat during observations is **continue**, **abort,** or **repeat** an ObsMat, or **stop queuing** new ObsMat

Core Functionalities:

- Provide quick estimates of the SNR and exposure time
- Calculate exposure sequence in order to achieve target's SNR required;
- Accommodate future revisions as subsystems performances are refined through their design cycle; and
- Support web-based user access in addition to internal call by PESA's Scheduler and Data Calibration products.

2.7 Breaker Product

SCH queues the ObsMat for observation by sending the ObsFld and ObsEle information to the Breaker (BRK) product. BRK commands the OESA Observatory Control Sequencer[3] (OCSe) to execute the observation and monitor status messages for indications of success or failure of the commands. BRK serves as the intermediate communication link between PESA and OESA to accommodate their future upgradability, specifically hardware related changes.

Once the ObsMat is queued, it will be impossible to change its definition, except by aborting the observation.

Core Functionalities:

- Execute observations by converting SCH ObsMat into observing commands executed by the OESA OCSe
- Monitor system status via OESA and report ObsMat actions to SCH

2.8 Glossary of Science Data Products

MSE provides three levels of official science data products, but archives and distributes five levels of data products where teams outside of MSE scientific staff supply the additional two levels. The official data products produced by MSE are referred to as Level 0, 1 and 2, and those by external teams are referred to as Level 3 and 4.

Level 0 Science Data Products: They are the raw data, unprocessed readouts from the spectrograph science detectors, both science frames images and calibration frames.

Level 1 Science Data Products: They are the first version of the flat-fielded, and dark and bias corrected 2D images; and wavelength calibrated, flux calibrated, and sky subtracted 1D spectra. Level 1 data products are generated soon after the end of a science. Level 1 data products are intended for quick-look by MSE staff and the survey teams.

Level 2 Science Data Products: They are 2D images, 1D spectra and associated catalogs produced by the Data Calibration (CAL) product after applying all necessary calibrations, nighttime and daytime[***]. Level 2 data products utilize the best possible calibration exposures and recipe available. Level 2 data products are the main homogeneous, science ready data products provided by the observatory. A basic set of science parameters is included with the Level 2 data products with each ObsEle in the ObjMod.

Level 3 Science Data Products: They are science data products provided by survey teams with added scientific value. Each survey team is expected to have their own science specific reduction methods and unique science data requirements; as such their data products will be homogeneous within the survey. These specialized science data products will be different between surveys and from the Level 2 data releases. For example, the Science Pipelines (PIP) product processes the 1D spectra by adding specific attributes to the science data, such as stacking, line fluxes, time variations, and metallicities, etc. PIP may use different calibration methods from CAL by using different metadata set in of its data reduction. Consequently, PIP may report additional set of derived metadata to the Science Archive.

Level 4 Science Data Products: They are the same as the Level 3 data products, but their added values are provided by the broader science team within the international MSE community.

Within the Data Reduction Pipelines, Level 1 and 2 are supplied by Data Calibration, and Level 3 and 4 are supplied by Science Pipelines. As stated, Level 2 data products are the main homogeneous, science ready data products provided by MSE. The main addition to the Level 3 and 4 data products are enhanced scientific contents, which are

[***]Daytime calibration frames are generally obtained in the following day

survey and science programs dependent, and they are not linked in the ObjMod. However, all science data products are archived in an access controlled manner from the Science Archive and Science Platform to the MSE user community.

2.9 Data Calibration Product

Once the Level 0 data are acquired, they will be validated and reduced using the Data Reduction Pipelines product. DRP processes science exposures by applying calibration algorithms and generating calibrated science data products by way of the Data Calibration and Science Pipelines operations, respectively. In addition, CAL will provide real-time feedback to the Scheduler via the ObjMod to adjust the exposure time, as needed, and runs in parallel to support SCH operations.

The Level 0 raw data, and Level 1 and 2 (reduced and calibration) science data products provided by CAL are stored in the ObjMod and linked with the engineering data, such as system and subsystems parameters in the OESA DMS and environmental data extracted from the Environmental Conditions Monitoring System (ECMS) product. A user interface is included for the MSE scientific staff to perform manual quality analysis and troubleshooting, compare calibration methods, test pipeline improvements, and study performance trends, etc.

Each time a spectrum is reduced by CAL, it also goes through a Quality Assurance[†††] (QA) process where the appropriate metadata flags are assigned to each spectrum to indicate its quality and utility for further science analysis. While the QA is be based on the data themselves, the affiliated engineering and environmental data are helpful to provide context for troubleshooting. For instance, a significant discrepancy between the *expected* SNR and the *measured* SNR will likely be resolved with the benefit of additional engineering and environmental information.

CAL queries the Exposure Time Calculator while reducing observations to compare *expected* SNR and *measured* SNR. Therefore, CAL accesses seeing conditions, measurements of the sky background, extinction, and image quality during an ObsMat via ECMS's environmental sensors and imaging camera. This information is saved with the affiliated ObsEles in the ObjMod, while the Science Archive and Science Platform receive only a subset of pertinent environmental measurements.

Directed by SCH, nighttime calibration exposures are obtained before and after each science exposure, therefore the real-time feedback from CAL to SCH is based on the corresponding Level 1 data products. To minimize down time and loss of observing efficiency during the night, CAL is executed after each individual exposure readout during an ObsMat. Since each ObsMat has readouts in multiple spectrograph arms, CAL is executed multiple times and at the start and end of an ObsMat when the nighttime calibration exposures are taken.

After the nighttime exposures, CAL is executed during the day after the ObsMat's nighttime activities along with relevant daytime calibration exposures as directed by SCH and accord to the Science Calibration Plan[4] methodologies to generate the interim Level 2 data products. This leads to progressive revision of its Level 2 data products until all observations associated of an ObsEle is complete.

Core Functionalities:

- Deliver 2D to 1D pipeline that combines science and calibration exposures to provide the official MSE science products
- Apply nighttime and daytime calibration for each ObsMat exposure readout according to the Science Calibration Plan methodologies to determine Level 0, 1 &2 science data products, including measured SNR;
- Update ObsEle's science information and status in the ObjMod as its observation progresses, and upload Level 0, 1 & 2 science data products and basic set of science parameters to the Science Archive;
- Extract observing conditions and call ETC to obtain *expected* SNR to support CAL's QA process; and
- Provide interface for user performed QA check and quick-look.

[†††]Criteria of the QA assessment to be defined during PESA development

2.10 Science Pipelines Product

The Science Pipelines produces Level 3 & 4 science data products by utilizing additional calibration that the survey teams, and the border science team community deem necessary and relevant. The data reduction “recipes” for MSE are expected to evolve and improve over time as the throughput, stability, and repeatability of the observatory are better understood and characterized. The survey teams and science community will have opportunities to improve and test calibration combinations, and refine the reduction recipes. Over time, PIP improvement may be incorporated into CAL as a standardized procedure. Until then, PIP will not archive its results in the ObjMod but rather in the ARCH and PLAT with appropriate metadata identification.

Core Functionalities:

- Supply 1D pipelines that use non-standard calibration and reduction to provide science added value to the official MSE science products
- Apply survey team-provided calibration and reduction recipes to selected ObsEles to generate Level 3 science data products;
- Apply science team-provided calibration and reduction recipes to selected ObsEles to determine Level 4 science data products; and
- Upload Level 3 & 4 science data products to Science Archive for Science Platform access.

2.11 Science Archive Product

After the completed ObsEles are processed by CAL and PIP, the associated Level 0 to 4 science data products are transferred to the Science Archive and Platform (SAP) product for distribution. Generally, the Science Archive (ARCH) product is intended for direct access by MSE’s scientific staff and survey teams, and the Science Platform (PLAT) product is intended to facilitate access for the MSE science community. In addition, only selected environmental and engineering data relevant to science are archived and distributed by SAP.

Level 0 & 1 science data products are saved immediately to the ObjMod and ARCH after each readout and nighttime calibration applied by CAL. Level 2 science data products are saved immediately to the ObjMod and ARCH after daytime calibration applied by CAL. The ARCH science products are released to the survey teams after MSE’s scientific staff have completed their QA analysis.

Our goal is to have the Level 1 data products available to the survey team on the night they are obtained, typically within minutes after the last exposure has been readout, so that quick analysis can be performed via ObjMod access, or at the latest in the following day, after initial QA analysis, via ARCH. The Level 2 data products, because of the additional daytime calibrations, are typically ready for QA analysis on the same day of daytime calibration and released to the survey team within a week, via ARCH.

A user interface similar to those of other modern science archives will be defined and developed in the PESA next development phase.

Core Functionalities:

- Serve as official repository of MSE science data products and added value science products
- Curate and provide secure storage for Level 0, 1 & 2 science data products;
- Maintain links between science data products and their “parent” ObsEles in the ObjMod, including the associated metadata;
- Provide user interfaces that facilitate the access of data products; and
- Manage access control according to users’ data rights.

2.12 Science Platform Product

The Science Platform (PLAT) product receives Level 3 & 4 science data products provided by the survey teams and science team, and facilitates access for Level 0, 1, 2, 3 & 4 science data products. In addition, PLAT includes visualization and analysis tools to the MSE community for timely processing and publishing of their MSE results.

A state-of-the-art user interface similar to those envisioned by the US NFS's NOIRLab will be defined and developed in the next PESA development phase. However, the interfaces will implement access controls based on authentication of authorized users and proprietary period. PLAT will automatically make the spectrum to the MSE community[‡‡‡] when the proprietary period has ended.

Core Functionalities:

- Curate and provide secure storage of Level 3 & 4 science data products;
- Maintain association between science data products and their "parent" targets in the survey programs
- Provide user interfaces that facilitate the access of data products
- Manage access control according to users' data rights and proprietary period; and
- Provide value added tools to facilitate processing and publishing of the science data products

2.13 End-to-End Simulator Product

In principle, the End-to-End Simulator (N2N) product models MSE's system performance and generates synthetic 2D spectra extracted from the LMR and HR spectrographs' detectors. N2N accounts for all physical effects affecting system performance at the subsystem level and contains modular representations of all subsystem parametrically. N2N interfaces directly with the Science Reduction Pipelines, and will be utilized extensively in the MSE development phase to validate design trades and calibration methodologies. During science operations, MSE staff will continue to rely on N2N to test and validate future software and hardware upgrades.

Basic Functionalities:

- Deliver detailed simulations of the 2D spectrum of science sources on spectrographs' detectors accounting for all physical effects in the system
- Interface with Science Reduction Pipelines to support Data Calibration development before real data are available
- Support ongoing MSE operations by allowing off-line testing and verification of software and hardware upgrades before actual on-site implementation

Detailed N2N functionalities will be defined during its development phase

2.14 Environmental Sensors Product

The Environmental Conditions Monitoring System (ECMS) product provides environmental information to enable safe operations in the scheduling, grading and monitoring of science observations. ECMS is the hardware and software required to monitor and respond to changing observing and operating conditions. The Environmental Sensors (EnSe) product is comprised of sensors for monitoring system-level conditions such as temperature probes, anemometers, humidity and precipitation sensors, seeing cameras, cloud/sky cameras, water column density monitor, dust particles, etc. Its intended usage is not specific to operations of individual MSE products, such as voltage, pressure, flow and temperature sensors on the telescope, enclosure, and instruments. In other words, EnSe contains sensors that provide environmental data relevant to the scientific performance and safe operations of the observatory, but it is not responsible for the sensors that monitor the health and proper functions of the machinery and equipment that are an integral part[§§§] of the Observatory Building and Facilities, Enclosure, Telescope and Science Instrument Package.

The sensors are located on a weather tower and distributed on the exterior of the Observatory Building and Facilities, on the interior and exterior of the Enclosure and on the Telescope. The environmental data are stored in the OESA Data Management System and its status server, and are accessible by the interfacing Scheduler and Data Calibration products.

[‡‡‡]Public access to MSE science data is under consideration by the Governing Board

[§§§]Such as the machinery and equipment supplied and delivered with the Observatory Building and Facilities, Enclosure, Telescope and Science Instrument Package as their deliverables

Core Functionalities:

- Provide environmental information to enable the Autonomous Rule to enforce safe science operations
- Interface with OESA DMS to upload environmental information
- Interface with SCH and CAL to enable scheduling, grading and assessment of observation progress

2.15 Autonomous Rules Product

In the context of remote observing, the Autonomous Rules (AuRu) product encompasses Business Rules[****] that discern environmental conditions, such as weather and system related, and respond autonomously to protect the observatory according to predefined Safety Rules. AuRu operates by using Business Rules to monitor a set of predetermined system and subsystems parameters and if they exceed their thresholds, in isolation or combinations, then apply the Safety Rules by placing the system into predefined safe modes, depending on violations. Examples of system parameters are wind speed, ambient temperature, relative humidity, precipitation, telescope hydraulic pressure, loss of coolant, drive over speed or over current, etc.

Core Functionalities:

- Ensure personnel, equipment and environment health and safety
- Interface with OESA DMS to extract environmental, system and subsystems information
- Notify responsible authorities when Safety Rules are activated
- Notify remote observer on-duty and responsible responders to resolve Business Rule infractions
- Log AuRu triggering and resolution history to OESA DMS

3. PESA PRODUCTS EXTERNAL AND INTERNAL INTERFACES

Table 1 illustrates the interfaces within and without the PESA products of the MSE Product Breakdown Structure presented in Figure 3. Among the PESA products, the internal interfaces are identified inside the dashed triangle and the external interfaces, with other products, are identified inside the dashed rectangle. Although Table 1 does not describe the nature of the interfaces, we intend to use Table 1 to strategize and prioritize PESA interface development in the next design phase. For instance, due to its many interface dependencies, the Object Model data structure should be defined early in the design phase and remain sufficiently stable throughout. However, the utilization of APIs to access ObjMod offers the needed interface flexibility for parallel design developments and alleviates some of our interface time phasing concerns. The external interface concern for ETC and N2N is that their development requires in-depth knowledge of many external products, which they simulate, in order to incorporate the appropriate design parameters and model their contributions to the system performance with fidelity.

We have developed a conceptual design level ETC, and further development is currently underway. However, the N2N development has not started due to the Design Reference Survey priorities.

[****]Business rules are lists of logical statements that enable or disable the observatory from operations based on the environmental criteria and conditions for decision making.

Table 1 N2N interface chart of the products listed in the MSE Product Breakdown Structure where presence of interface is indicated by the highlighted 'X' mark.

System	Product																												
Observatory Building and Facilities	OBF																												
Enclosure	ENCL	X																											
Telescope Structure	STR	X	X																										
Fibre Positioner System	PosS	X	X	X																									
Fibre Transmission System	FTS		X	X	X																								
Science Calibration System	SCal	X	X	X	X	X																							
Low/Moderate Resolution Spectrograph	LMR	X	X	X		X	X																						
High Resolution Spectrograph	HR	X	X			X	X																						
Common Software	CSW	X	X	X	X	X	X	X	X																				
Data Management System	DMS	X	X	X	X		X	X	X	X																			
Observatory Control Sequencer	OCSe						X			X	X																		
Telescope Control Sequencer	TCSe			X						X	X	X																	
Instrument Control Sequencer	ICSe				X			X		X	X	X																	
Facility Control Sequencer	FCSe		X	X						X	X	X																	
Observatory Safety System	OSS	X	X	X	X	X	X	X	X	X	X	X	X	X	X														
Computer Services	CpS	X								X	X	X	X	X	X														
Object Model	ObjMod									X	X																		
Proposal Review	PROP																	X											
Scheduler	SCH										X							X											
Exposure Time Calculator	ETC		X	X	X	X	X	X	X		X								X	X									
Breaker	BRK											X								X									
Data Calibration	CAL						X	X			X							X		X	X								
Science Pipelines	PIP						X	X			X							X					X						
Sciecne Archive	ARCH										X							X					X						
Science Platform	PLAT																	X						X	X				
End-to-End Simulator	N2N	X	X	X	X	X	X	X	X	X	X	X	X	X	X					X	X	X							
Autonomous Rules	AuRu									X	X	X				X						X							
Environmental Sensors	EnSe	X	X	X			X				X						X											X	
	Product	OBF	ENCL	STR	PosS	FTS	SCal	LMR	HR	CSW	DMS	OCSe	TCSe	ICSe	FCSe	OSS	CpS	ObjMod	PROP	SCH	ETC	BRK	CAL	PIP	ARCH	PLAT	N2N	AuRu	EnSe

4. PESA PRIORITIES FOR DESIGN REFERENCE SURVEY

The Design Reference Survey[5] (DRS) is a key recommendation of the 2018 system conceptual design review panel. The review panel stated for a survey driven observatory like MSE, DRS is an enormously useful tool for evaluating at every phase of development whether or not the science goals can be met by the planned observatory.

DRS is a mock observing plan for the first few years of science operations to determine whether the as-designed observatory can achieve MSE's science objectives. In other words, DRS simulates execution of a consolidated survey plan that contains diverse science cases[††††] while taking into account both external constraints, e.g., weather, lunar cycle, sky availability as a function of time of the year, as well as the technical constraints, e.g., observatory, instrument, and calibration using design information collected at each design phase. By examining step-by-step the execution of the envisaged survey plan, DRS informs functionally and operationally the adequacy of MSE's Product Breakdown Structure, and specifically the PESA software products to support the intended observations.

For the first phase of our DRS, our priority is to understand the survey speed in terms of fiber-to-target allocation and telescope pointing to achieve the targets' SNR required given the target list representing the selected science cases. MSE's survey speed is directly related to ability of the Scheduler to optimize Observing Matrices that maximize number of science targets observed and SNR, and minimize telescope pointing and instrument configuration changes. Since the current ETC is adequate for SNR predictions, our PESA work plan is to develop a preliminary version of SCH that provides optimal fiber-to-target allocations associated with discrete telescope pointings and then an optimal sequence of telescope pointings under the overall constraints. Utilizing SCH and ETC, we anticipate being able to estimate the survey speed and gain insight into the pertinent SCH design parameters required to improve survey efficiency.

[††††]Current observing plan combines four diverse science cases of Milky Way halo star metallicities, "Cosmic noon" survey, AGN reverberation mapping and cosmology.

5. CONCLUSION

Given a spectroscopic survey facility's success is linked to the quantity and quality of spectra it obtains per night, per year, and ultimately over its lifetime, we have proposed a high-level end-to-end software suite, PESA, to ensure that MSE's data products will enable and ensure the success of the project. We have prioritized the PESA development with respect to the Design Reference Survey in order to determine and optimize MSE's survey speed and efficiency.

The proposed PESA design is consistent with MSE's science operations. Figure 4 summaries how the operations workflow is realized by the PESA products. The five color compartments show the processes associated with each phase of the workflow shown in Figure 1, and the operator's actions to enable the processes, user interfaces, repositories and software tools provided by the participating PESA products to execute the workflow.

Phase 1 Proposal Selection: Survey teams submit their proposals via PROP that includes a website interface, a proposal library and an API to access and define their target lists in the Object Model. The Survey Selection Committee evaluates and selects proposals via PROP. PROP facilitates communication, enquiries and notifications, tags and forwards selected proposals to SCH for processing. ETC and SCH provide utilities to justify and validate the feasibility of the proposals' target lists.

Phase 2 Target Definition: Selected survey teams upload their target definitions and observing conditions into the ObjMod using either website interface or via direct API access support by PROP. ETC and SCH provide utilities to validate the feasibility of the proposed target lists.

Phase 3 Observation: Remote observers schedule and execute the observations using the Scheduler's queue. During observation, targets' science information is updated in the ObjMod and engineering and environmental conditions information is logged in the OESA Data Management System repository. SCH provides utilities to monitor progress and update Observing Matrix dynamically, and BRK interfaces with OESA Observatory Control Sequencer to execute ObsMat.

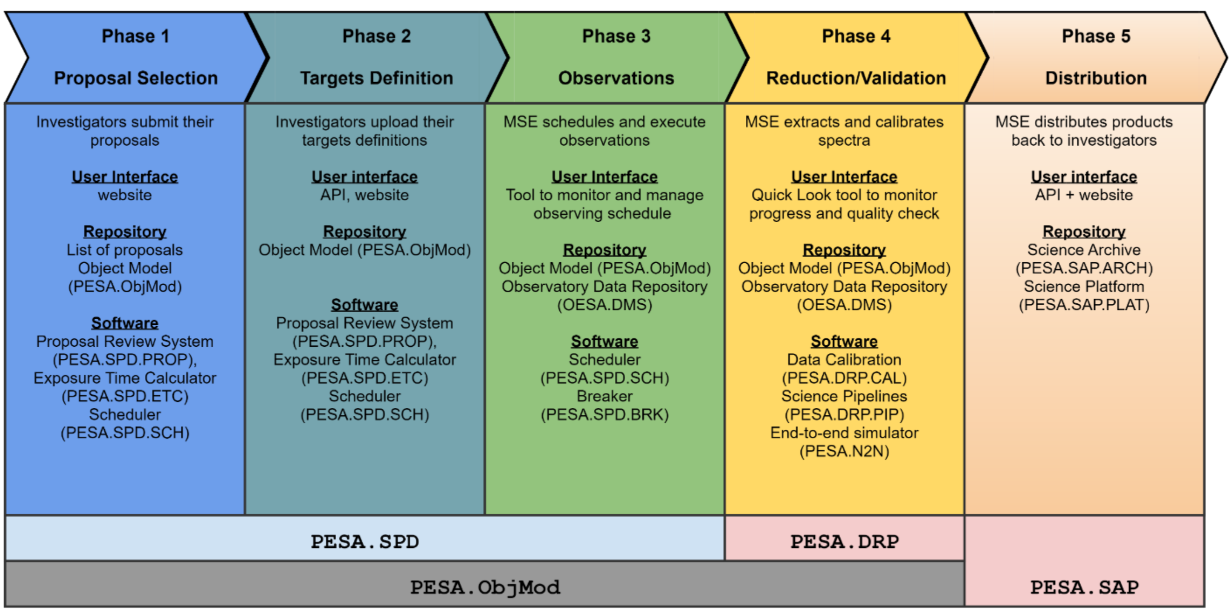

Figure 4 Summary of science operations workflow supported by PESA including the associated processes, user interfaces, repositories and software tools proposed.

Phase 4 Reduction/Validation: MSE scientific staff verify calibrated and reduced spectra generated by Data Calibration using the CAL interface for quick-look and QA process. CAL reduces 2D science frames and calibration frames from detector readouts into 1D spectra and update ObjMod target science information (Level 0, 1 & 2) with the associated engineering and environmental conditions information in OESA DMS as official MSE science data products. In addition, an End-to-End simulator is developed for the MSE scientific staff to facilitate CAL development and testing.

The survey and science teams may develop specialized reduction techniques as Science Pipelines, which are survey- or science program-specific, to refine and enhance the 1D spectra and form value added science products. However, PIP is not the responsibility of the MSE scientific staff, and its science products (Level 3 & 4) are separated from the ObjMod.

Phase 5 Distribution: After QA analysis, MSE scientific staff release Level 0, 1 & 2 science data products from Science Archive along with the associated metadata. Survey and science teams also release their Level 3 & 4 science data products from ARCH. Generally, users access all science data products via Science Platform with API or web interface according their data rights. In addition, API and website interface access are available for eligible users to access science information directly from ObjMod and ARCH.

In conclusion, we have defined the PESA products to be consistent with science operations and compatible with MSE's operations model. The architecture is designed to interface with all the other components of MSE, and to optimize the overall observation and data reductions processes. The PESA development will start as the Project enters the next development phase.

Appendix A. Detailed Description of Observing Element Parameters

The **coordinates** should be precise enough to maximize the injection efficiency at the fiber input and will be specified in the GAIA reference coordinate system. All catalogs used by MSE will use that same reference frame.

The **ephemeris** should follow the same convention and have the same precision as that required for the coordinates. They should be provided for a period of time long enough to cover the duration of the program. Fast moving targets will not be suitable for MSE, although open-loop positioning of the fibers will enable observations of targets with apparent motion no greater than 0.5" per hour, TBC.

We expect that a given target will be observed with one spectrograph and in only one configuration, i.e., spectrograph settings, and hence each ObsEle will correspond to only one instrument setup. However, some observations might require the use of multiple spectral resolutions and coverages, e.g., both J and H band in the low resolution (LR); follow-up in high resolution (HR) of a previous observation in moderate resolution (MR). The observing programs will thus create several ObsEles for a given target, each with its own instrument setup.

Some observing programs might require **non-standard calibrations**. There will be a set of default calibrations established for each of the LR, MR, and HR spectrographs. Any calibration not covered by these defaults will be specified and justified with provisions provided by the PROP.

The observing programs will provide a **metric for validation** of the ObsEle, e.g., a SNR goal, along with a magnitude, at a given wavelength. The SNR goal will be defined for each target based on calculations done by the Exposure Time Calculator. The observed SNR will be measured by the Data Reduction and Pipelines for each target at the wavelength given by the program. There could be programs, however, for which the goal is to obtain a SNR value averaged over all the targets in a given field or over the entire sample. The PROP will allow such entries too.

The **program identification number** will identify the survey program to which the ObsEle belongs.

The **priority** within an observing program will help scheduling of the ObsEles. A simple integer number is enough to scale the priority. Internal priorities will be combined with global priorities set by MSE when allocating targets from multiple survey programs in the same field.

Some observations will require **time constraints**. For example, some targets might need to be observed repeatedly, others might need to be observed only within some specific windows, and yet others might need to be observed in sequence with a specific cadence. The PROP set will have the required capabilities to define those constraints.

Because MSE will provide data that have reached a given SNR, as defined and requested by the observing programs, the observations will be performed, by default, under any environment conditions, as long as the telescope time could be used efficiently. Therefore, **weather constraints** will need to be justified and accepted by the Survey Selection Committee.

Additional constraints may be allowed on the **settings** of the observations. For instance, investigators may want to observe a given target with the same fiber, or they may want to observe several targets at the same time. Those constraints will need to be justified and accepted by the Survey Selection committee.

ACKNOWLEDGEMENTS

The Maunakea Spectroscopic Explorer conceptual design phase was conducted by the MSE Project Office, which is hosted by the Canada-France-Hawaii Telescope. MSE partner organizations in Canada, France, Hawaii, Australia, China, India, and Spain all contributed to the conceptual design. The authors and the MSE collaboration recognize and acknowledge the cultural importance of the summit of Maunakea to a broad cross section of the Native Hawaiian community.

REFERENCES

1. Flagey N., “MSE Operations Concept Document”, MSE.PO.ENG.DOC-REQ-002, MSE project document (November 2017).
2. Flagey N., “MSE Injection Efficiency Budget Allocation Document”, MSE.PO.ENGSYS-SYS-009, MSE project document (November 2017).
3. Vermeulen T., Isani S., Withington K., Ho K., Szeto K., Murowinski R, “Observatory software for the Maunakea Spectroscopic Explorer”, *Software and Cyberinfrastructure for Astronomy*, Proc. SPIE. 9913 (July 2016)
4. Mcconnachie A., “MSE Science Calibration Plan”, MSE.PO.SCI.DOC-PLA-001, MSE project document (November 2017).
5. Marshall J., Petric A., Flagey N., “A design reference survey for the Maunakea Spectroscopic Explorer”, *Observatory Operations: Strategies, Processes, and Systems*, Proc. SPIE (December 2020)