ALMA - the first year of observations

Andreas Lundgren[a], Lars-Ake Nyman[a], Masao Saito[a], Baltasar Vila Vilaro[a], Gautier Mathys[a], Paola Andreani[b], John Hibbard[c], Sachiko K. Okumura[d], Ken'ichi Tatematsu[d], Bill Dent[a], Mark G. Rawlings[a], Eric Villard[a], and Lewis Ball[a]

[a]Joint ALMA Observatory, Alonso de Cordova 3107, Santiago, Chile,
[b]European Southern Observatory, Garching bei München, Germany,
[c]National Radio Astronomy Observatory, Charlottesville, Virginia, USA,
[d]National Astronomical Observatory of Japan, Mitaka, Japan

ABSTRACT

The Atacama Large Millimeter/submillimeter Array (ALMA) is a major new interferometer operated on Llano de Chajnantor at 5050 m altitude in the Chilean high Andes. This location is considered one of the world's outstanding sites for submillimeter astronomy.

ALMA is still under construction, but science observations has started already in what is commonly known as ALMA Early Science Cycle 0. The purpose of ALMA Early Science Cycle 0 is to deliver scientifically useful results to the astronomy community and to facilitate the ongoing characterization of ALMA systems and instrumentation as the capability of the array continues to grow. Early Science will continue through Cycle 1 and until construction and commissioning of ALMA is complete.

This publication aims to give an insight into the challenges we face operating telescope of this scale at Chajnantor, a plateau 4800–5100 meter above sea level in one of the driest places of earth. It also will also present statistics from the proposal submission, describe the path from an accepted proposal to a calibrated data product, and finally an outlook for the future.

Keywords: telescope, submillimeter, science operations, ALMA, interferometer

1. INTRODUCTION

The Atacama Large Millimeter/submillimeter Array (ALMA) is a new interferometer in the northern Atacama desert. There, both air pressure and humidity are unique. At a pressure of about 555 hPa and frequent relative humidities of below 15%, the atmospheric opacity permits to access the millimeter and submillimeter windows up to 200 μm (1.5 THz). When completed in 2013, ALMA will consist of an array of fifty 12-m antennas, with baselines up to 16 km, and the Atacama Compact Array (ACA) - an additional compact array of twelve 7-m, plus 4 total power (single dish) 12-m antennas that greatly enhance ALMA ability to image extended targets. It is designed to work at millimeter and sub-millimeter wavelengths, and will cover atmospheric windows from 31.5–950 GHz (9.5–0.32 mm).

ALMA is a partnership between North America, Europe, and East Asia in cooperation with the Republic of Chile. ALMA construction and operations are led on behalf of Europe by the European Organization for Astronomical Research in the Southern Hemisphere (ESO), on behalf of North America by the National Radio Astronomy Observatory (NRAO), which is managed by Associated Universities, Inc. (AUI), and on behalf of East Asia by the National Astronomical Observatory of Japan (NAOJ). The Joint ALMA Observatory (JAO), located in Santiago, Chile, provides the unified leadership and management of the construction, commissioning and operation of ALMA.

The observatory is physically divided in two geographically distinct sites: the Array Operations Site (AOS) at 5050 m altitude, and the Operations Support Facility (OSF) at 2900 m altitude. AOS hosts the antennas and the correlator, while the OSF is the base of the operations, with control room, laboratories, and dormitories. The two

Further author information: (Send correspondence to A.L.)
A.L.: E-mail: alundgre@alma.cl, Telephone: +56 2 4676360

sites are connected by a 28 km long road, wide enough to transport the 12-m diameter antennas. The computers at the two sites communicate via a fiber link bandwidth of 1 Gb/s (Gigabit/sec). See de Graauw et al. (2012),[1] Hills et al. (2012),[2] Bhatia et al. (2012),[3] and Lopez et al. (2012)[4] for more details on the site, instrumentation, and construction of ALMA. For more information on the ALMA science operations see Nyman et al. (2010)[5] and Rawlings et al. (2010).[6] More information can also be found on our websites `http://www.almaobservatory.org/` and `http://www.almascience.org/`.

We are currently in the first year of science observations using a partial array. This phase is called ALMA Early Science Cycle 0, and the capabilities are defined to be: an array of at least sixteen 12 m antennas, receiver bands 3, 6, 7 & 9 (wavelengths of about 3, 1.3, 0.8 and 0.45 mm), baselines from 18 m to 125 m (compact configuration) and from 36 m to 400 m (extended configuration), single field imaging and mosaics of up to 50 pointings, and a set of correlator modes that allow both continuum and spectral line observations.

In Cycle 0, the scientific observations are conducted on a best efforts basis in parallel with the ongoing construction, commissioning and verification of the whole ALMA system, a range of constraints apply during Cycle 0. The completion of the full array of 66 antennas with the full set of scientific capabilities will continue to be the highest priority. Early Science will not be allowed to delay unduly the construction of the full 66-antenna array, but nonetheless provides an important opportunity for first science from this cutting edge facility. Early Science will continue through Cycle 1 and until construction of the ALMA array is complete.

The proposal deadline for ALMA Cycle 0 was 30 June 2011, and on 30 September 2011, the JAO started Early Science observations (Cycle 0). ALMA Early Science Cycle 0 was initially expected to span 9 months, and it was anticipated that 500-700 hours of array time would be available for Early Science projects, but in early 2012 the time frame was extended to the end of 2012.

In this paper we present the capabilities in Cycle 0 (Section 2), the data archive (Section 3), the proposal process (Section 4), generation of Scheduling Blocks (Section 5), the Cycle 0 observations up to June 2012 (Section 6), data reduction (Section 7), and, finally, an outlook for the rest of Cycle 0 and Cycle 1 (Section 8).

2. OBSERVING CAPABILITIES IN CYCLE 0

The purpose of Early Science Cycle 0 is to deliver scientifically useful results to the astronomy community and to facilitate the ongoing characterization of ALMA systems and instrumentation as the capability of the array continues to grow.

2.1 Antennas

The nominal number of 12-m antennas for cycle 0 observations is 16, all of which are elements of the interferometer. No 7-m antennas or total power antennas are offered.

All ALMA antennas are designed to meet very stringent performance criteria, and to successfully operate under the extreme environmental conditions at the AOS, i.e. strong winds, large temperature ranges and gradients, strong solar irradiation and snow.

2.2 Antenna Configuration

There are two array configurations available for Cycle 0: Compact and Extended. The Compact configuration, with baselines from 18-125m, is designed to have high brightness-temperature sensitivity and is used for observations where extended structure is important. The Extended configuration (baselines 36-400m) has higher angular resolution and is more suited to the study of objects with higher surface brightness. The angular resolutions in the two configurations are summarized in Table 1. The angular resolution given is the FWHM of the synthesized dirty beam for a source in the declination range 0 to -40 degrees (outside this range of declination the beam becomes somewhat elongated in the North-South direction).

The Maximum Scale given in Table 1 is the largest angular scale that can be observed effectively. Targets containing smoothly varying structure larger than this in both dimensions will be resolved out. This is the well known missing flux problem intrinsic to interferometry. The limit is taken to be 0.6 x (wavelength/baseline$_{min}$) - but this is only a guide. The Atacama Compact Array (7-m antennas and total power antennas) will eventually be used to measure larger scales, but are not available in Cycle 0.

Table 1. Cycle 0 Antenna configurations. Upper four rows apply to the compact configuration, and the lower four rows for the extended configuration.

band	Frequency[GHz]	Resolution [arcsec]	Max. scale [arcsec]
Compact			
3	100	5.3	21
6	230	2.3	9
7	345	1.55	6
9	675	0.80	3
Extended			
3	100	1.56	10.5
6	230	0.68	4.5
7	345	0.45	3
9	675	0.23	1.5

2.3 Receivers

ALMA bands 3, 6, 7 and 9 are available on all antennas in Cycle 0. For all bands both linear polarizations are received and processed separately. The receivers are based on SIS mixers and there are two types - dual-sideband (2SB), where the upper and lower sidebands are separated in the receiver and then processed separately, and double-sideband (DSB), where the sidebands are superimposed, but are separated in later processing. The frequency ranges, receiver types and measured noise temperature (over 80% of the tuning range) are given in Table 2.

Each antenna also has an amplitude calibration device (ACD) which employs ambient and hot loads to calibrate sky transmission. In addition, Water Vapor Radiometers (WVRs) are available on all 12-m antennas, and are used to measure atmospheric precipitable water vapor (PWV) along the line-of-sight. Corrections for phase errors due to fluctuations in PWV can then be applied to improve the coherence.

Table 2. Cycle 0 receiver characteristics.

band	Type	Frequency range[GHz]	T_{rx}[K]
3	2SB	84-116	40
6	2SB	211-275	40
7	2SB	275-373	75
9	DSB	602-720	120

2.4 Correlator

The correlator provides a set of spectral "windows" which can be used simultaneously. For Cycle 0, up to four simultaneous spectral windows are available, each of which must have the same bandwidth and resolution. The correlator can process both polarizations or all the resources can be used to analyze a single polarization. In Dual Polarization mode, the separate spectra for each linear polarization are normally combined to improve sensitivity. The correlator operates in two main modes - Time Domain Mode (TDM) or Frequency Division Mode (FDM). TDM provides modest frequency resolution and produces a relatively compact data set. It is used for continuum observations or for spectral line observations that do not require high spectral resolution. FDM gives high spectral resolution, but the data rate from the correlator is 32 times larger than TDM. Six FDM set-ups are available with different bandwidths and resolutions as listed in Table 3. For single-polarization operation, the number of channels is double that given in the table, and hence the channel spacing is halved.

The Bandwidth given in the table is the width of the spectrum processed by the digital correlator. The usable bandwidth is limited to about 1875 MHz by the anti-aliasing filter, which is ahead of the digitizer in the signal path. The Channel Spacing is the separation between data points in the output spectrum. The spectral resolution - i.e. the FWHM of the spectral response function - is larger than this by a factor that depends on the window function that is applied to the data in order to control ringing in the spectrum. For the default function - the "Hanning" window - this factor is 2, i.e. the effective spectral resolution is twice the channel spacing given in Table 3.

Table 3. Table 3: Cycle 0 Correlator Modes for Dual Polarization operation.

Bandwidth [MHz]	Channel Spacing [MHz]	No. of Channels	Mode
2000	15.6	128	TDM
58.6	0.0153	3840	FDM
117	0.0305	3840	FDM
234	0.061	3840	FDM
469	0.122	3840	FDM
938	0.244	3840	FDM
1875	0.488	3840	FDM

2.5 Calibration

Complex gain calibration is essential for all interferometric observations. It is accomplished by observing suitable calibrators - i.e. objects with small angular sizes and accurately known positions. So-called "Gain" calibration is used to remove time-dependent phase and amplitude fluctuations by regular observations of a nearby calibrator. Bandpass calibration is used to remove frequency-dependent variations across the receiver bandpass. The internal ACD is used to calibrate out differences in the atmospheric transmission towards the positions of the different calibrators. Absolute amplitude calibration is based on observations of objects of known flux, principally solar system objects. It is expected that the accuracy of the absolute amplitude calibration relative to these objects is better than 5% for Band 3. Calibration in the higher frequency bands will be less accurate. The goals are: better than 10% in bands 6 and 7, and better than 20% in band 9. Positional accuracy of better than 1/10th of a synthesized beam-width should be possible on sufficiently bright objects.

2.6 Observing Modes

Single field Interferometry is the standard observing mode in which the antennas track a single pointing position. When observing pointed mosaics, the antennas cycle around a series of pointing positions and the interferometric data are combined in post-processing to produce a "mosaic" image. The spatial resolution and largest angular scale is about the same as for a single field interferometry. Mosaics with up to 50 pointing positions are supported in Cycle 0.

3. ARCHIVE, DATA DISTRIBUTION, AND USERS SUPPORT

3.1 The ALMA Archive System

The purpose of the ALMA archiving subsystem is to provide services for persistent archiving and retrieval of observational data, observation descriptors (including scripts), images produced by the data processing pipeline and technical (device monitoring and logging) and environmental data. In addition the Telescope Monitoring and Configuration Data Base (TMCDB) is hosted by the archive. The ALMA archive consists of two parts - a front-end archive, which is used for storing details of observational scheduling blocks and related data needed to execute observations, and the back-end (science) archive, which stores the instrumental and processed data. The architecture is based on the NGAS system (New Generation Archive System) with Oracle technology for the metadata, as shown in Figure 1. The hardware used for each instance of the ALMA Archive has been assembled to guarantee high redundancy and 99.5% round-the-clock uptime. The system has also been designed to cope with 6.6 MB/s average data rates and 66.6 MB/s peak rates (for short periods of time). Copies of the Archive will be available over the World at each of the ALMA Regional Centers (ARCs). The locations selected for the ARCs are Garching (Germany) for Europe, Charlottesville (USA) for North America and Mitaka (Japan) for East Asia. In addition, there will be two Archives in Chile, one at the array control center (OSF) near San Pedro de Atacama and another at the Santiago Central Offices of the ALMA project (SCO). The flow of observational data will be from the OSF Archive to the SCO Archive and from there to the three ARC Archives (see Section 3.2).

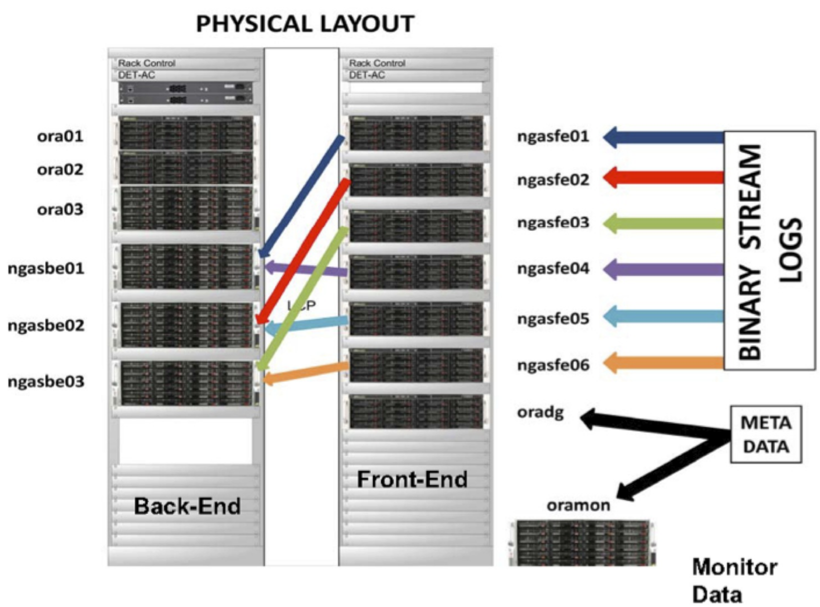

Figure 1. Archive design (front-end & Back-end) at the OSF to store metadata and raw and monitor data.

The archive at the OSF is designed to provide up to a year of temporary storage for the scientific and monitoring data. The rest of Archives are intended to be permanent copies of the data. The process of copying the data to any of the archives involves a replication of the metadata (support data) and a mirroring of the bulk data (ASDM and FITS files), see Figure 2. These two processes do not need to be simultaneous and they can use different routes (i.e., media delivery and Internet). Oracle stream technology will be used for the copies of the metadata.

3.2 Data Flow

Data flow in ALMA spans a large geographically distributed system using the ALMA Archives as nodes (see Figure 3). For the science user the process starts with proposal submission (see Section 4). Each submitted project is assigned a unique project code that will identify it during its whole lifecycle. A dedicated server is

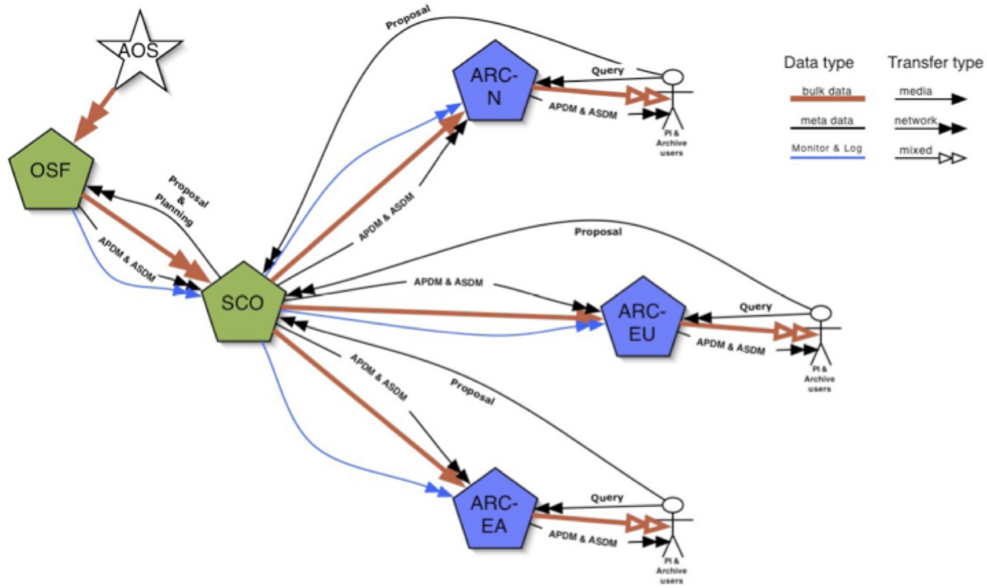

Figure 2. Data flow from the AOS down to the ARCs.

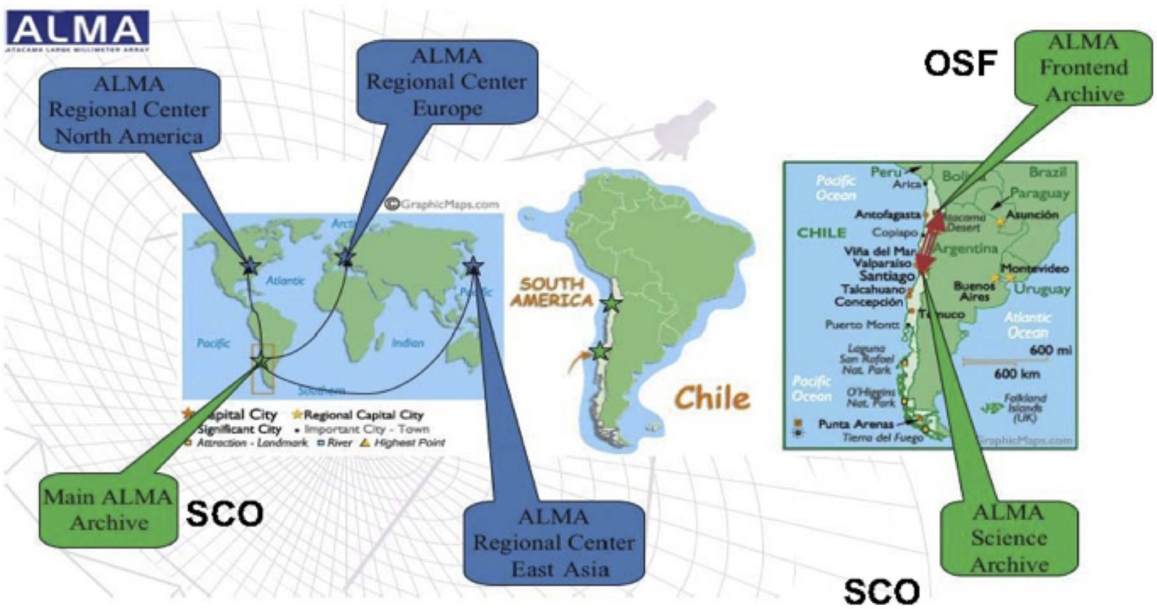

Figure 3. Location of the ALMA archives.

in charge of the basic checking of the submitted information and the allocation of this identifier, together with the requests for creation of the basic associated metadata in the main ALMA Archive located in the SCO in Santiago, Chile. The proposal metadata is used during the Proposal Review process by the Review Committee members that add comments and ranks. For all the approved projects XML files with all the relevant parameters needed for observations are then created (Phase 2) and stored in the ALMA Archive (SCO) as Scheduling Blocks (hereafter SBs). These SBs are then replicated to the Archive at the OSF, where the observations are executed.

At the time of the observations, the parameters in the SB that is selected for execution are passed to the Control software that carries out the observation. Data coming from the correlator (binary) located at the AOS Technical Building is sent to the OSF via a fiber optic link (1-10Gbs). This rate of transfer is larger than the speed of data ingestion into the OSF Archive. To prevent data loss, a Frontend Archive has been added to the OSF Archive (see above), which is in charge of buffering the data writes into the Backend Archive. Together with the binary data (bulk data), the system is creating associated metadata tables that, among other things, will reference the locations of the binary data on the NGAS system. The metadata information is gathered from several external(weather information, monitoring data, etc) and internal (pointing models, etc) sources by the Data Capture software. This software converts all the information into XML tables with fully synchronized metadata streams and the logical links to other tables and the bulk data, and sends them to the OSF Archive. The format of these tables and links is specified by the ALMA Science Data Model (ASDM). Each execution of an SB results in an ExecutionBlock associated with an ASDM file, which contains all the data and metadata relevant to that execution. Additional information regarding the conditions of the observations, logs, quality assurance information, etc are also stored in the Archive via the ShiftLog Tool software.

Copies of the data in the OSF Archive into the SCO Main ALMA Archive are done using buffered stacks of requests for the bulk data and Oracle streams technology for the metadata. The latter is usually synchronized within minutes of data archiving into the OSF Backend Archive. For the bulk data, the current rate of transmission is ≈ 60 Mbs and it is expected to grow in the near future. The data are then replicated to all the regional Archives (ARC Archives) using available international networks. For the ARC in Europe, networks belonging to the university consortium REUNA are used from Santiago to Brazil. From there, undersea cable lines to Europe are used. For the East Asian ARC, the network is connected through the USA by using R&E networks including SINET, RedCLARA, and REUNA. For the USA, the existing network for transfer of LSST Telescope data from Chile to Chicago (USA) is used.

Once all the data needed to achieve one of the science goals specified in a proposal are gathered, the data are reduced (see Section 7). All sites (SCO and ARCs) use high-end fast data reduction clusters with large storage disk arrays using the LUSTRE file system model. Data are ingested into the data reduction clusters from the local Archive copy and reduced there. The final data products to be delivered to the PIs are packaged in the clusters (tar files) and delivered to the PIs. For data reductions carried out at the SCO, the tar files are ingested back into the Main Archive and then replicated to all the ARC Archives. Personnel in the ARCs check the products and releases the data to the PI via ftp servers. For data reduced at an ARC, the data are released to the PIs via the ftp servers and also copied (using VPN rsync) to a server at SCO. Personnel at SCO ingest the tar files into the Main Archive and they are replicated from there to all the ARC Archives. Data product replication progress is checked from SCO and the ARCs and the contents of the tar files are verified with checksums and internal README files for additional redundancy. Additional external copies of the data sent from the Main Archive are also generated on serial tapes directly from the data reduction cluster. The end result of the data flow process is that there are copies of the original and processed data and all the associated metadata in four locations around the world, giving optimum redundancy in case of failures in one of the nodes.

In addition to science data (and supporting information) the OSF Archive also stores monitoring/engineering data that is used to troubleshoot issues, trend analysis and optimization procedures, scheduling of tasks, etc. These data are not copied further to any other Archives down the data flow line.

3.3 Users support provided by the ARCs

The ARCs are providing user support over the full life cycle of an ALMA observation. Astronomers are getting assistance with proposal preparation, refining observing scripts, and processing and analysing data. Indeed, the key user services provided by the ARC include:

- Preparing and distributing the Call for Proposals
- Writing documentation and training materials to help with proposal preparation and submission
- Providing and supporting the Observing Tool, preparing the observing scripts on behalf of the PIs
- Offering ALMA training workshops and tutorials
- Staffing the Helpdesk

- Archiving and distributing ALMA data
- Providing data reduction support, documentation, and cookbooks
- Developing and maintaining user-oriented observing and analysis tools such as the ALMA sensitivity calculator, the spectral line catalog (Splatalogue), and observing simulators (e.g. SIMdata)
- Hosting visits on-site at the ARC nodes for user support

The ALMA Helpdesk is the main resource for users seeking help from their assigned ARC. The Helpdesk is where users can ask general questions, report problems or bugs in any ALMA-related software, request information about their observing programs, and make arrangements to visit one of the ARCs. Users must be registered with the ALMA Science Portal to submit a Helpdesk ticket.

4. PROPOSAL SUBMISSION AND PROPOSAL REVIEW PROCESS

A call for proposals for ALMA Early Science Cycle 0 was published on March 30, 2011, with a submission deadline of June 30. The astronomical community responded enthusiastically: 919 unique proposals were received. Their distribution across the four ALMA science categories was as follows:

1. Cosmology and the high redshift universe: 20%;
2. Galaxies and galactic nuclei: 27%;
3. ISM, star formation/protoplanetary disks and their astrochemistry, exoplanets: 40%;
4. Stellar evolution, the Sun and the solar system: 13%.

These proposals were assessed by 8 ALMA Review Panels (ARP). Each ARP was composed of 6 Scientific Assessors (7 for Category 4) whose combined expertise covered the range of topics relevant to one of the four scientific categories. There were two ARPs for each of Categories 1 and 2, three for Category 3, and one for Category 4. The Chair and the Deputy Chair of each ARP served on the ALMA Proposal Review Committee (APRC). The APRC Chair did not belong to any ARP. Hence the total number of Science Assessors contributing to the Cycle 0 Proposal Review Process was 50.

The proposals were assessed on the basis of the overall scientific merit of the proposed investigation and its potential contribution to the advancement of scientific knowledge, as well as on the extent to which the planned observations demonstrated and exploited the ALMA Cycle 0 capabilities.

In order to keep the workload of the panels to a manageable level, science assessments took place in two stages. For Stage 1 review, each proposal was assigned to four members of one of the ARPs of its science category, who each gave it a score. The individual assessor's scores were combined to compute a mean preliminary score. A single ranked list of all proposals was built and the top 70% of the proposals proceeded to Stage 2, including proposals for which the standard deviation of the individual Stage 1 scores was much greater than average.

These proposals were submitted to technical assessment by 25 ALMA staff members, and feasibility issues were flagged to the ARPs. Only 27 of the assessed proposals were found to be technically infeasible, a very positive result that indicates the quality of the Observing Tool software, the training and awareness sessions conducted by the ALMA Regional Centers, and the documentation provided to proposers.

All ARPs met face-to-face in Santiago to discuss the proposals that proceeded to Stage 2. Following its discussion, each proposal was assigned a single, final ARP score, which was the arithmetic mean of the scores assigned by each member of the relevant ARP and submitted by a secret vote. The proposals reviewed by the ARP at Stage 2 were ranked according to their ARP score. The resulting ranked list was the final product of the ARP meeting. It was complemented by additional recommendations as the ARP deemed appropriate. In particular, when two or more proposals involved duplication of observations, the ARP(s) reviewing them issued recommendations about their respective priorities, based on their scientific merits.

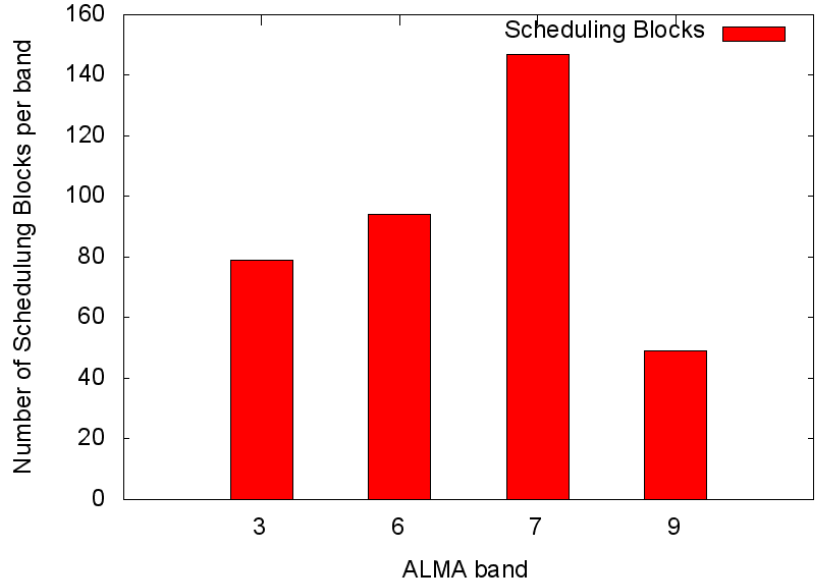

Figure 4. The number of SBs in the top 112 accepted proposals as function of ALMA band.

The APRC met face-to-face immediately after the ARP meetings, to review the ARP results and prepare a single ranked list of all proposals. It issued final recommendations for treatment of duplicated observations, based on the ARP input.

The project completion probabilities were estimated at this stage for communication to proposers. Highest priority was assigned to the top-ranked proposals of each region until their cumulative execution time exceeded the respective regional share of the available science time for the cycle (500 hours), taking into account the technical feasibility of the proposals. Further down the regional rankings, proposals were flagged as "fillers" until the cumulative execution time exceeded 1.5 times the regional share of the available science time. Filler projects are observed only if the conditions do not allow any of the highest priority projects to be executed. The remaining proposals were deemed very unlikely to be observed.

The Directors Council made the final decision about proposal selection, based on the APRC recommendations. The outcome of the process was communicated via email to all PIs.

Ultimately, 112 proposals were identified as having the highest priority for completion, and a further 52 as filler projects. The formal amount of observing time required for execution of all the highest priority proposals, estimated with the Observing Tool software, was 506.4 hours. Fillers were scheduled for a total amount of 242.7 hours of estimated observing time. Both groups of projects were shared across the regions in the agreed proportions of array time based on the shares of the partners' and of the host country. The higher-priority projects covered a wide range of topics; they were distributed across the four ALMA scientific categories proportional to the demand.

The total amount of requested time in all submitted proposals, according to the OT estimate, was 4292 hours, so that the overall oversubscription of ALMA in Cycle 0 (that is, the ratio of the time that would have been required for execution of all submitted proposals to the time needed to complete the highest-priority projects) was 8.5. This is similar to the ratio of the number of submitted proposals to that of highest-priority proposals, 8.2.

The 112 highest priority projects comprised 371 Scheduling Blocks (SBs, see Section 5), which about half were intended for compact configuration and half for extended configuration. Most of these SBs in the group were for band 7 (148 out of 371), followed by band 6 (95), band 3 (79), and band 9 (49). The distribution on sky was also uneven. In order to see how this would affect scheduling, the RA of the sources were binned per project,

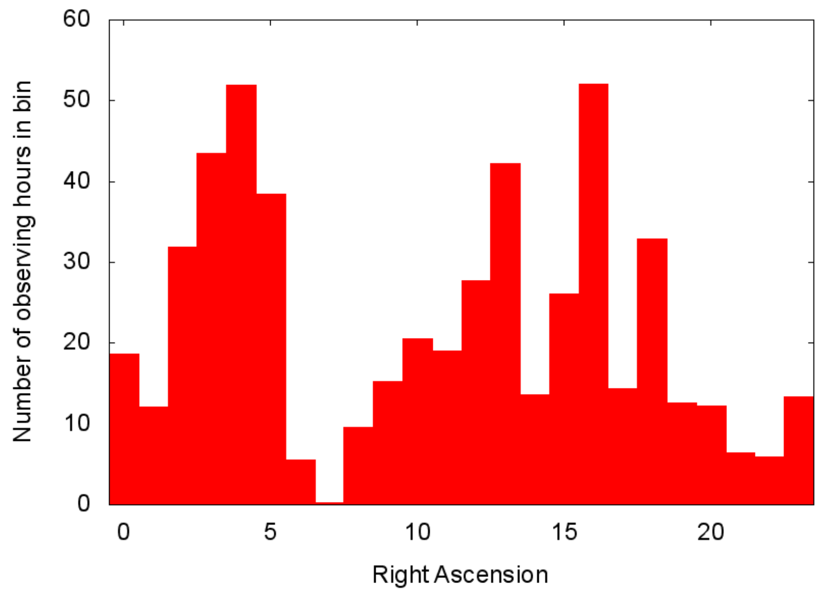

Figure 5. Total estimated requested observing time as function of Right Ascension.

each bin was divided by the total number of sources, and multiplied by 4.5, which was the average number of hours per project, and then the bins were summed over the projects. The result (see Figure 5) showed that there was a strong demand on the RA range 2-5 hours, and a pronounced demand on the RA range 12-18.

5. SCHEDULE BLOCK GENERATION

Each set of observations is coded as Scheduling Blocks or SBs. Each SB includes scripts for the actual observations (e.g. standard observing modes offered by the Observatory, such as Single Point Interferometry of Mosaicing) and also all the additional information needed to carry out the observations (source properties, spectral setup, integration times, calibrators, etc).

5.1 PRP Results and Staff Training: the Phase 2 Group

As part of the conclusion of the ALMA Proposal Review Process (PRP), information regarding the decisions taken as to which ALMA proposals were accepted is automatically written back to the projects in the ALMA Science Archive (ASA). For Cycle 0, the projects most likely to be observed were identified, factoring in the Right Ascension distribution of the science field sources, the receiver bands to be used, the overall PRP ranking of the project, the Executive affiliation of the proposers and so on. In accordance with such factors, the Cycle 0 projects were grouped into "batches" based on relative urgency. The projects in each batch were assigned, along with appropriate preparation deadlines, to ALMA staff members at the JAO and ARCs trained in Phase 2 (Program) preparation. The training of these ALMA staff members took the form of intensive programs ("Phase II boot camps") lasting several days, specifically designed to turn users of the ALMA OT into Phase 2 experts. The growing list of ALMA staff having undergone such training has come to be known within the project as the "P2G" or Phase 2 Group. An additional long-term aim of these training programs is to have as many Phase 2 experts distributed across the project as possible, so that they could then further distribute knowledge at each of the ALMA sites through local training sessions.

5.2 The ALMA OT at Phase 2

As at the Phase 1 (Proposal) stage, the ALMA Observing Tool (OT) is used to generate the necessary components of the project that will actually be executed on the telescope. The assigned P2G member launches the OT and retrieves the project from the Science Archive. At this stage, the OT presents the project in the Phase 2

“Program” tab by default (although the original proposal is still available for read-only examination in the Phase 1 “Proposal” tab). The initial Phase 2 version of the project contains a copy of all of the Phase 1 Science Goals, which can then be further modified by the P2G member, if so mandated during the PRP.

Once the P2G member is satisfied that the Science Goals are all up to date and conform to the OT automatic validation criteria, the option to “Generate SBs from the Selected Goal” is then selected. At this point, the OT applies a set of standard principles, previously decided by ALMA staff, to generate standard-specification XML Scheduling Blocks inside parent ObsUnitSets (at the simplest level, an ObsUnitSet may be regarded as a hierarchical “folder” structure into which one or more related SBs may be grouped). Only a few observing modes were offered for Cycle 0, and the associated automatic SB generation process was rudimentary: each Cycle 0 Science Goal yielded a single ObsUnitSet containing a single SB. As additional capabilities become available to ALMA proposers in future Cycles the number and complexity of observing modes offered will increase, requiring additional structural sophistication. During Cycle 1, for example, a single Science Goal may generate one SB for the 12-m array and an additional SB for the ACA, if necessary.

5.3 General SB Design Philosophy and Its Implications

Ideally, any automatically generated SB should exactly match the needs of the Science Goal provided as input, and this is a long-term goal for ALMA. The precise methods to be used for tailored automatic SB generation are, however, still very much under development. For Cycle 0, therefore, two fundamentally opposite philosophies were originally considered when establishing the desired contents of an automatically generated SB: the “minimalist” and “maximalist” approaches. The former is based on the premise that the OT should only generate the absolute minimum number of necessary components (e.g. types of calibrators, observing group structure, etc.) common to all projects, and that that all additional components needed for a given project would subsequently be manually added by the P2G member. This is a simpler approach for software development, but requires that the P2G member have a greater level of technical knowledge. In contrast, the “maximalist” approach is based on the premise that the OT should generate SBs that contain all possible components that might be needed for any Cycle 0 observation. This approach simplifies the P2G editing work at the expense of increased software complexity. For reasons of simplicity, a standardized approach to observing was adopted for Cycle 0, and since many of the P2G members were also still learning all the details of SB structure, the decision was taken for Cycle 0 to adopt the “maximalist” approach, particularly since the long-term goal was increased automation.

5.4 Routine Phase 2 Tasks

For Cycle 0, the adoption of the “maximalist” approach described above has made much of the P2G work an exercise in the removal of unnecessary extra components for a given project. At this early stage in ALMA life, however, additional SB editing work has always been required. This additional editing work can be divided into two clear categories. The first of these addresses the current need for manually selected calibrators. For future observing Cycles, ALMA is expected to make full use of calibrators that are automatically selected during the actual execution of SBs on the telescope, but this option is not yet available. Consequently, the assigned P2G member has been required to manually select and insert suitable calibrator candidates during initial project preparation, and subsequently liaise with the project support scientist to implement changes and updates as necessary. The second type of additional P2G work encountered during Cycle 0 has involved the implementation of a set of evolving P2G best practices. This list of best practices includes guidelines on (e.g.) grounds for the inclusion of additional calibrators, optimal cycle times for repeated observations and details of workarounds needed to address the various temporary technical limitations of a telescope still actively undergoing construction and testing.

To each project an ARC contact scientist has been assigned, whose task is to discuss with the PI the observing setup of the SBs and get his/her approval. Once the project SBs have been readied by the assigned P2G member, all such modifications are documented in both the OT project and a JIRA issue tracking ticket created for the purpose. The P2G member then uses this JIRA ticket thread to thoroughly describe to the contact scientist, so that the contact scientist can then discuss the contents of the Phase 2 project with the Principal Investigator of the project and generate any feedback needed. When all parties are satisfied that the project is ready for execution, the scientist sets the project status to “Ready” to indicate that the project is ready for execution.

Occasionally, a change request may also subsequently be filed by the project PI to the Observatory. If such a change request is granted, then the assigned P2G member updates the project accordingly.

6. OBSERVATIONS

6.1 Observing time and staffing

In Cycle 0, observing is conducted in blocks normally lasting 7 nights (5 nights in the first 5 months), and the staff astronomers have access to the telescope from 17:00 to 09:00 (Chilean Local Time), except Wednesdays, when part of the time is allocated to computer regression tests, and the start of the telescope time is delayed to 22:00.

Each observing block is staffed with (a minimum of) 6 astronomers, plus a lead astronomer, and 3 telescope operators. The staff is distributed equally over three shifts: Morning (06:00-16:00), Day (13:00-23:30), and Night (22:30-06:30). The astronomers come from JAO, the ARCs and ESO Chile.

The observing period starts following the conclusion of Phase 2. The Astronomer on Duty (AoD) searches the ALMA Archive for SBs on the basis of a number of criteria, including current weather conditions, suitability for the current antenna configuration, project ranking, etc. The selected SB is passed to the array for execution and run as observing sequences. The SBs are executed using the Scheduler subsystem. The low level commands required by the hardware are sent from the Control Room to the AOS via dedicated optical fibers to the OSF.

All successfully-executed SBs are marked as having been run. Any SBs that fail to complete (for whatever reason) are flagged as "Suspended" and are subsequently examined. Any SBs with intrinsic problems are identified as such, repaired by ALMA staff and re-admitted to the queue as ready for execution.

The data from SBs that complete execution immediately undergo some preliminary inspection and are subjected to QA0 (level 0 Quality Assurance). QA0 consists of real-time/semi-real-time monitoring of calibration data during its acquisition, and the calibration summaries at the end of an SB, plus various system monitoring data. This allows the monitoring of the integrity of the whole signal path, from the atmosphere, through antenna and front-end issues, down to the back-ends. The QA0 parameters used monitor possible atmospheric effects, antenna issues, front-end, back-end and various connectivity issues. The QA0 parameters are used by the AoD to decide whether a given dataset has been obtained under satisfactory conditions or has to be re-observed. If the collected data for an SB pass QA0, then that execution of the SB is flagged as successful. If the collected data fail QA0, then this is noted, and the cause of the failure investigated and resolved.

6.2 Experiences in Early Science Cycle 0

During the first 5 months, Early Science Cycle 0 observations were conducted in blocks of 5 days every two weeks. While the quality of the data being collected was excellent, the completion rate of projects was lower than planned. By the end of January 2012 approximately 50% of the anticipated observing time for Cycle 0 had been used and we estimate that roughly 25% of the highest priority observations had been completed. The reason was a mix of difficulties with both software and hardware, and weather. While consistent with the "best efforts" basis of ALMA Early Science this is lower than we had hoped.

In order to increase the likelihood that most PIs of the highest priority Cycle 0 projects will receive scientifically valuable data sets, the Cycle 0 observing period was extended until the end of 2012. Furthermore, during the additional period of Cycle 0 observing the fraction of ALMA time used for scientific operations was increased from the initial fraction of 33% to around 50% of the total available array time.

In February, there was a planned break. The basic reason is that the weather conditions on Chajnantor are not suitable for observations, and the time was instead be focused on construction activities, upgrades and antenna reconfiguration. However, in 2012, the Altiplanic winter brought with it an unusually abundant precipitation both at the OSF and the AOS, which had a severe negative impact on the ALMA operations. The Altiplanic winter is a meteorological phenomenon in which the jet stream reverses and brings moist air from the east to this usually extremely arid site. Generally it occurs sporadically between December and February. The unusually long period of poor weather in 2012 delayed the restart of the observations until late March.

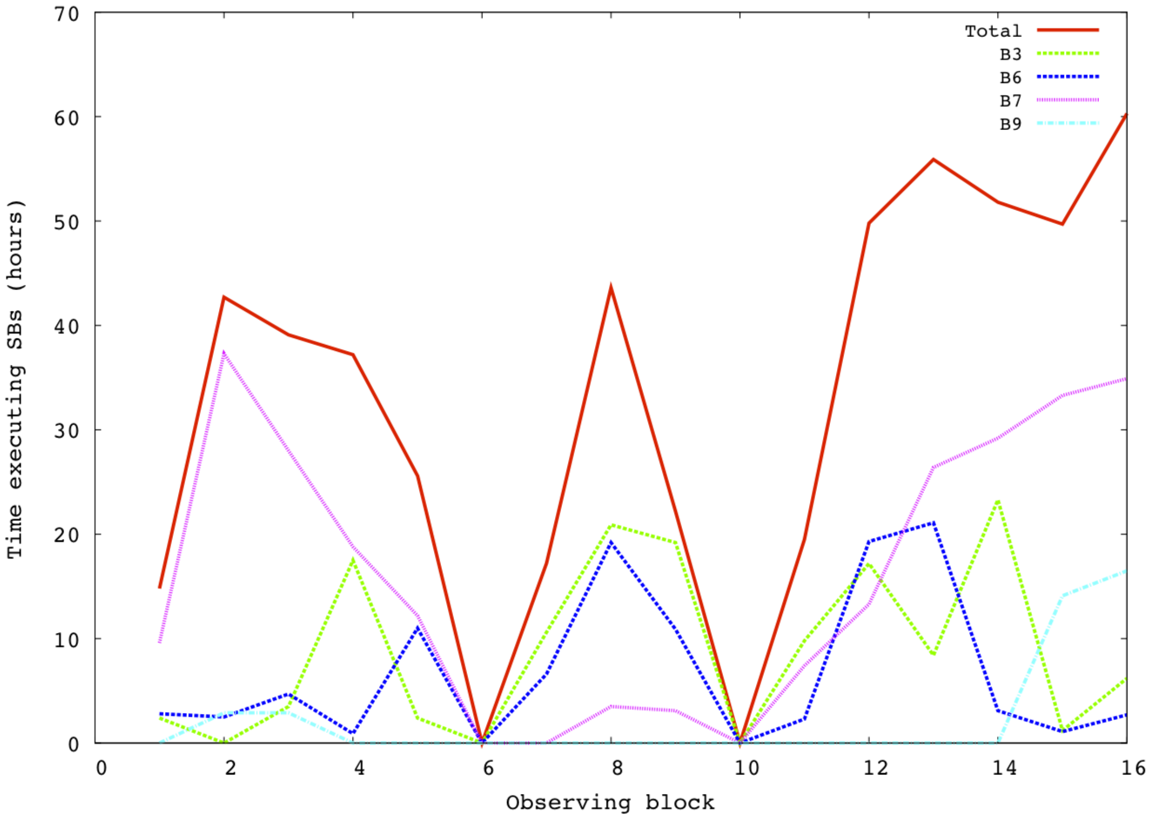

Figure 6. Schedule Block execution time (in hours) as function of observing block. Observing block 1 started on September 30, 2011, and observing block 16 on May 30, 2012. Block 6 was lost due to a technical problem, and block 10 was lost due to residual effects of the Altiplanic winter, which also severely hampered block 11. In Block 15 the weather had improved so much that we could start executing SBs for band 9.

In April, when the effects of Altiplanic winter gradually subsided, the amount of observing time increased. The effect of the decreasing precipitable water vapor (PWV) also allowed observations at higher frequencies (band 7 and Band 9).

Figure 6 shows the Schedule Block execution time as function of observing block from the start of Cycle 0 in September 2011, to the beginning of May 2012.

7. DATA REDUCTION

Data reduction is the next step after data acquisition. When ALMA is fully operational, the output of data reduction will be images that principal investigators (PIs) can immediately use to do science. ALMA is the first mm-interferometry observatory that has set the requirement to produce science quality products, in order to be accessible also to astronomers without any radio-interferometry background. This requires the development of dedicated software tools, capable of handling a large variety of observations, science goals and reduction techniques. As a result, it is a requirement of ALMA to standardize observations, which helps developing reduction recipes.

7.1 The ALMA Quality Assurance process

The goal of ALMA Quality Assurance (QA) is to deliver to the PIs a reliable final data product that has reached the desired control parameters outlined in the science goals, that is calibrated to the desired accuracy and free of calibration or imaging artifacts. To be more efficient in detecting problems, ALMA QA has been divided into several stages that mimic the main steps of the data flow. The broad classification of this multi-layered QA approach is:

QA0: is a near-real-time verification of data quality. It deals with rapidly-varying performance parameters (at scales of a scheduling block, or SB, execution length or shorter, such as e.g. phase fluctuations, antenna gain, offset pointing, and focus.) and thus it has to be performed at the time of the observations. Assessment is performed by AoDs (Astronomers on Duty). See also Section 6.1

QA1: includes slowly varying (timescales longer than a week, such as e.g. baseline measurements, delays, all-sky pointing, and focus curves) array performance parameters. They will all be measured by AoDs executing standard calibration SBs created as specified by the Calibration Plan.

QA2: Data Reduction, see Sections 7.2 and 7.3

QA3: is post-reduction evaluation of the data products delivered to the PIs, triggered by PIs (or ARC staff).

For more details, see `http://almascience.eso.org/call-for-proposals/technical-handbook`

7.2 QA2 for Cycle 0

The exact QA2 metrics for the science pipeline are still being developed. They will benefit very much from the experience gained through reducing Cycle 0 data. In the mean time, we have developed a collection of routines that try mainly to identify outlier solutions in the calibration tables, and residuals in the calibrated data. As for the imaging, the most important criteria for the science goals is the RMS in the final image.

In the future, most of the data reduction workflow will be automated. When the observations are completed and the data are successfully ingested into the main archive, they will be exported, calibrated and imaged, their quality will be assessed, in particular with respect to the PI's science goals, the various products from the reduction will be packaged, ingested into the final archive, and announced as ready for delivery. Most of this workflow will eventually be covered by the ALMA science pipeline, which has recently entered a phase of testing and commissioning.

The reduction of ALMA data, especially submillimeter data, is an exercise that requires some specific experience. On the other hand, the pipeline will not go into operation before Cycle 1. So for Cycle 0, it was decided that the ALMA astronomers themselves would calibrate and image the data manually, and assess the data quality.

7.3 Typical data reduction workflow for Cycle 0

For Cycle 0, most observations are relatively simple, so a standard reduction recipe can be defined. Below are the main steps of the workflow, starting with the import of the data from the archive and finishing with the ingestion of the science products into the archive.

- Import data into reduction software
- A priori and QA0 flagging
- Phase correction based on data from the Water Vapor Radiometers (WVRs)
- System temperature (Tsys) calibration
- Antenna position correction (antpos) calibration
- Application of the WVR, Tsys and antpos calibration tables
- Split of the science data
- Initial flagging (shadowing, atmospheric lines)

- Putting a model for the spatially-resolved calibrators
- Bandpass calibration
- Gain calibration (phase and amplitude)
- Application of the bandpass and gain calibration tables
- Flux equalization/bootstrapping across observations
- QA2 of calibration
- Imaging
- QA2 of imaging
- Packaging of data and products

The software used for reduction is called CASA (Common Astronomy Software Applications). It is the official software for ALMA data, and is being developed by an international consortium of scientists at institutions around the world. It uses a Python interface, and provides a collection of low and high-level routines. The science pipeline is being built upon CASA, and will be distributed with it in the form of specialized routines.

CASA is a powerful tool for radio-interferometry reduction. Something important that the pipeline routines will add is the optimum choice of parameter values. In the interim, until the official pipeline is functional, we have developed a script generator that creates Python script with CASA commands already filled in with some nominal values ALMA staff use these script making changes to the parameters if necessary.

8. OUTLOOK

8.1 Continuing observations in Cycle 0

With the completion of observing block 16 (finished on June 6, 2012), ALMA has reached half of the time that will be devoted to observations of Cycle 0 projects.

With the increasing number of antennas it will be possible to schedule observations using the extended or compact configuration during the latter part of Cycle 0. Given the current efficiencies we expect to deliver scientifically valuable data sets to most PIs of the highest priority Cycle 0 projects, and to some PIs of filler projects.

8.2 Capabilities in Cycle 1

While the highest priority of the ALMA project continues to be the completion of the full 66-antenna array, Cycle 1 observations will provide a significant opportunity for science from this unique world-class facility.

The ALMA Early Science Cycle 1 capabilities will include:

- Thirty-two 12-m antennas in the 12-m Array (twice that of Cycle 0)
- 12-m Array configurations with maximum baselines ranging from 160 m to 1 km (pp to twice that of Cycle 0)
- Nine 7-m antennas, and two 12-m antennas for total power (single dish) observations, in the Atacama Compact Array (ACA), for use in combination with 12-m Array observations for imaging of extended structures (not available in Cycle 0)
- Receiver bands 3, 6, 7 & 9 - wavelengths of about 3, 1.3, 0.8 and 0.45 mm (same as in Cycle 0)
- Single field imaging and mosaics of up to 150 pointings (three times that of Cycle 0)
- Continuum and spectral line observations (same as in Cycle 0)

Preparations for Cycle 1 is well underway, including preparation of user documentation, updates to the Science Portal, several tests of the Observing Tool and proposal submission, tests of the Ph1M (the software tool that is used in the proposal review process) and preparations for the proposal review process.

The deadline for submission of Notices of Intent (NoI) was May 15. In total 772 NoIs were received. The executive distribution was 355 from EU, 205 from NA, 149 from EA, 6 from EA+NA, 30 from Chile and 27 others. A document providing the NoI statistics is being prepared.

The Cycle 1 Call for Proposals was issued successfully on May 31, and the deadline for proposal submission is 15:00 UT on July 12, 2012. It is anticipated that 800 hours of 12-m Array time and up to 800 hours of ACA time will be available for Cycle 1 projects.

For more details for the call for proposal, see `http://www.almascience.org/call-for-proposals/`

ACKNOWLEDGMENTS

The authors would like to express their gratitude to all that have worked together to realize this amazing project. No one mentioned, and no one forgotten.

REFERENCES

[1] de Graauw, T. et al., "ALMA construction and early science," in [*SPIE Conference Series*], *SPIE* **8444** (2012).

[2] Hills, R. et al., "ALMA commissioning and science verification ," in [*SPIE Conference Series*], *SPIE* **8444** (2012).

[3] Bhatia, R. et al., "ALMA commissioning and science verification ," in [*SPIE Conference Series*], *SPIE* **8449** (2012).

[4] Lopez, B. et al., "Assembly, integration, and verification (AIV) in ALMA: series processing of array elements," in [*SPIE Conference Series*], *SPIE* **8449** (2012).

[5] Nyman, L. et al., "ALMA science operations," in [*SPIE Conference Series*], *SPIE* **7737** (2010).

[6] Rawlings, M. et al., "ALMA science operations and user support: software," in [*SPIE Conference Series*], *SPIE* **7737** (2010).