ngVLA Memo # 18

Summary of the Science Use Case Analysis

R. Selina, E. Murphy, A. Erickson
NRAO

July 26th, 2017

Abstract

In the spring of 2017, the ngVLA Science Advisory Council (SAC) solicited Science Use Cases (SUC) from the astronomy community via the Science Working Groups (SWGs). These use cases were collated and parameterized to identify community needs on technical parameters of the ngVLA. A workshop was held in June 2017 to discuss the findings from the exercise and converge on both the key science missions and a supporting concept for the array. We summarize the technical findings from an initial analysis of the science use cases and the June 2017 Science Workshop.

1 Introduction

The development of the ngVLA concept is a top-down process – identifying the needs of the radio astronomy community in the coming decades, turning those needs into technical requirements, and then developing a concept to meet them. While this approach can sound linear, it is inherently iterative, since there are also programmatic constraints (cost, schedule) and technical feasibility limitations that inform the concept.

This analysis is in support of the ngVLA 'reference design' that will be the technical basis for a Decadal Survey 2020 proposal. The reference design will present the outline of a single concept that meets the key requirements, as defined by the radio astronomy community, for the facility. The reference design is not expected to represent an optimal design, and a more complete system design would be generated in the following years as part of the funded project design phase.

The development of the reference design is inherently time constrained by the decadal process, so the level of analysis must be constrained by available resources (time). This preliminary analysis will guide the reference design, but input should and shall continue to be welcome throughout the design phase of the project. The present intent is to capture major themes and features required by the user community and shape the reference design concept.

1.1 Science Advisory Council & Science Working Groups

The ngVLA Science Advisory Council (SAC) was formed in October2016. The SAC is the interface between the scientific community and NRAO, providing feedback and guidance directly to the next-generation

VLA (ngVLA) Project Office on issues that affect the scientific design for a U.S.-led, next-generation centimeter-to-millimeter wave interferometer.

The ngVLA SAC is a group of approximately twenty (20) members of the science community that represent a broad range of research interests and user requirements. The SAC has defined its own organizational structure, with chairs for each of four key ngVLA Science Working Groups (SWG) areas that include: Cradle of Life, Galaxy Ecosystems, Galaxy Formation, and Time Domain/Cosmology/Physics.

Each SWG in turn was tasked with developing a finite number of detailed ngVLA science cases. These cases were documented via a common template[1] provided by the ngVLA project office, which captured both its scientific rationale and associated science requirements.

SWGs used these summaries to prioritize recommendations and corresponding science requirements within their own area of specialization. The SAC then reconciled the various SWG recommendations, and is developing a prioritized list of science cases (Bolatto et. all, in prep.)

1.2 Science Use Case Capture

The Science Use Case Capture form consists of the following major elements:

- A narrative of the science goals.
- A narrative of the rationale, including a description of the measurements required.
- A series of tables that define the targets of the observations (see Table 1), the observational setup, the data products required, and any imaging considerations.
- A narrative of any additional functional or performance requirements or additional materials pertinent to the use case.

(A) 'TARGETS' OF OBSERVATIONS		
Type of observation (what defines a 'target')		Individual pointings per object
		Individual fields-of-view with multiple objects
		Mosaics of multiple fields of view
		Non-imaging pointings
Number of targets		
Position range of targets (RA/Dec.)		
Field of view (arcmin^2)		
Rapidly changing sky position? (e.g., comet, planet)		YES [details:]
		NO
Time Critical?		YES [details:]
		NO
Required rms (μJy/bm) [per km/s for lines]		
Peak brightness (μJy/bm)		
Expected polarized flux density (expressed as % of total)		

Table 1-Example table from the Science Use Case Template. This table captures requirements relevant to the targets of the observation.

[1] https://science.nrao.edu/futures/ngvla/documents/sciencecases/ngVLA_SciCase_Template.v1.01.docx

1.3 Use Case Parameterization

In parallel to the SWG and SAC ranking, the project aggregated the technical inputs from each use case to evaluate themes in the technical requirements of the array. A total of 75 use cases were received and included in this process.

The parameterization process consisted of three steps:

- Observational requirements definition and data normalization.
- Computation of supporting (derived) technical requirements.
- Analysis of the supporting technical requirements.

The definition and normalization process consisted of processing each use case and tabulating the observational requirements. The degree of development of the use cases varied, so there were often gaps that had to be filled, and nuance and variance in units had to be normalized to make the inputs machine readable.

Many science cases required multiple different observation modes (e.g. different frequency tunings, spectral line and continuum observations, etc.), and the requirements for each observation had to be parsed from the use case narratives and tables. Once the data was fully captured, the normalized inputs were sent to the original author for review.

These normalized and reviewed data inputs were then used to compute desired technical parameters for the array that would support the described observation. For example, a given angular resolution and center frequency could define the extent of the array required; the largest angular scale of interest could define the minimum baseline length needed; the spectral resolution and total bandwidth desired might be indicative of the number of channels required, etc.

The result of the use case parameterization exercise is a 195-line by 81-column matrix. The science requirements (split by observation) are included, subsequent technical requirements are estimated, and finally the performance implications of the current array concept are computed.

The matrix was parsed by a series of python scripts to compute relevant statistics and build histograms and other plots. Work will continue on refinements to the parameterization matrix to extract additional findings from the existing and any forthcoming parameterized use cases, but the initial findings are reported in Section 2.

2 Use Case Findings

Analysis of the use cases has helped to answer a number of technical questions and to define functional and performance requirements for the array. Major findings from the analysis are reported in this section.

The use cases are aggregated into three sets. ‘Drivers’ are the highest-ranked cases, by the SWGs and SAC, and represent the key science goals of the ngVLA. The ‘Related’ cases are associated with the drivers, while all other cases form the ‘Identified’ set.

2.1 Frequency Coverage

The distribution of use cases by frequency indicate broad support for all atmospheric windows in the frequency range of 1.2 GHz to 116 GHz. High-ranking (by the SWGs and SAC) cases are present at all bands, as can be seen in Figure 1.

Equally important was that all frequencies were requested at a wide range of angular scales, suggesting that the array should have a full complement of receivers on all antennas, as can be seen in Figure 2.

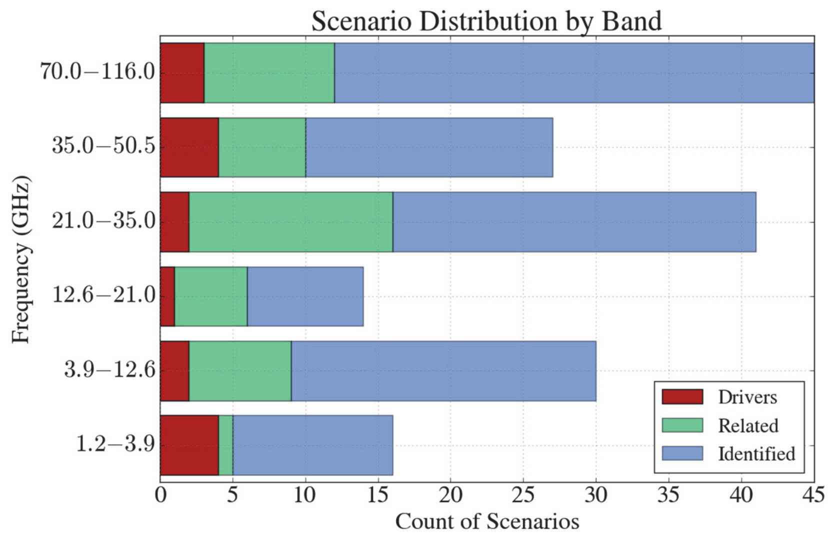


Figure 1 - Distribution of center frequency amongst observations. Concentrations of cases occur near 30 GHz and at 100 GHz, but there are also driving use cases extending to the 1.2GHz lower frequency bound.

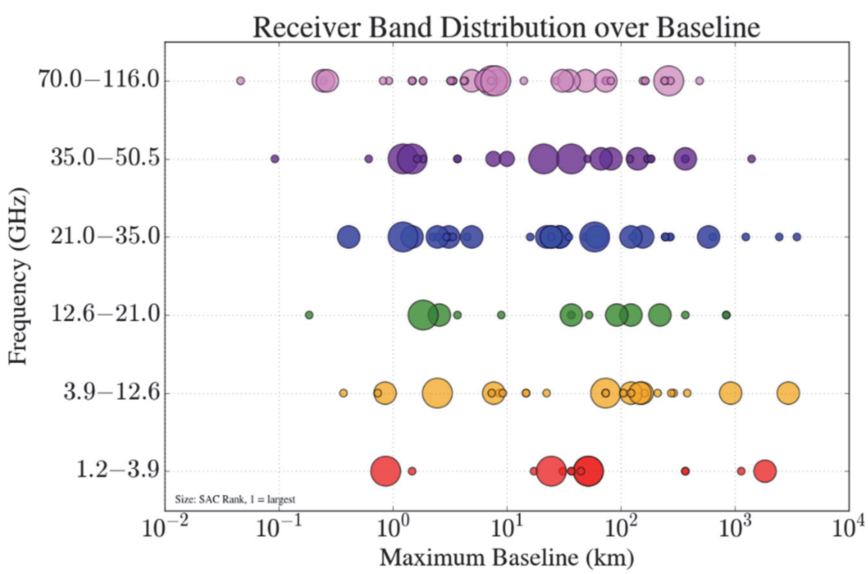


Figure 2 – Requested center frequency vs maximum baseline (to support the requested angular resolution). The size of the marker indicates SAC rank, with 'drivers' having larger diameters. Note that all frequencies are requested out to a few hundred kilometers, and lower frequency observations desire baselines out to a few thousand kilometers.

2.2 Sensitivity

The sensitivity required to support the use cases is shown in Figure 3 and Figure 4 for continuum and spectral line cases, respectively. The requirement is displayed as a cumulative histogram, with lower RMS values on the left side of the x-axis. Supporting 80% of continuum cases requires sensitivity of order 80nJy/bm in a reasonable period, while the spectral line cases require sensitivity of order 10uJy/bm/km/s.

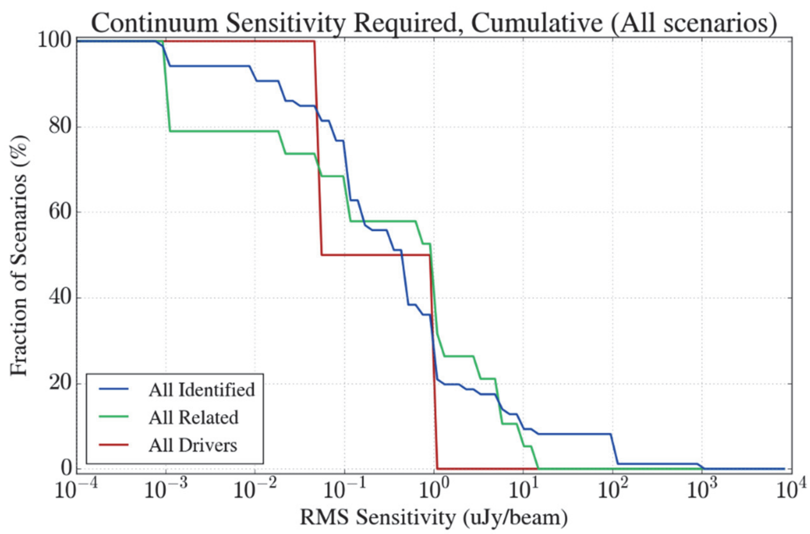


Figure 3 - Cumulative histogram of sensitivity required for the continuum cases. Lower noise figures are on the left of the plot. A continuum sensitivity of order 80nJy/bm is required in a 'reasonable' period to support the driving use cases and approximately 80% of all identified cases. A reasonable period might be of order ten hours.

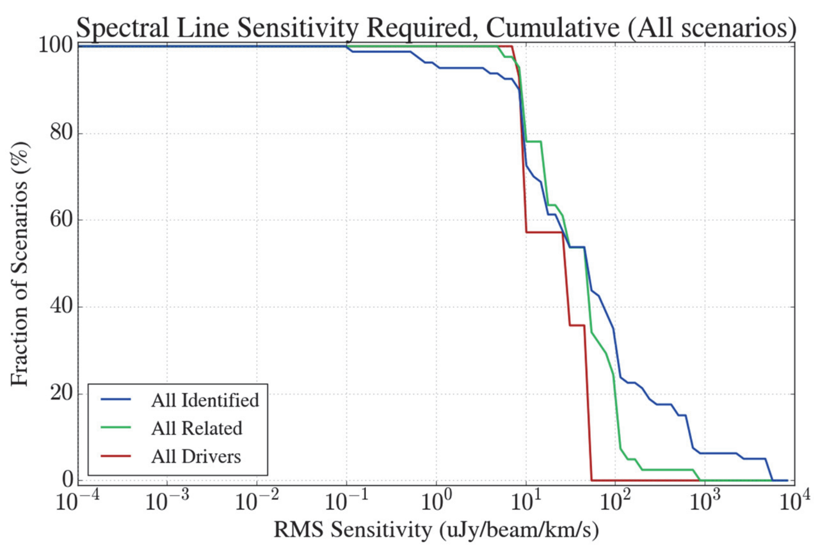


Figure 4 - Cumulative histogram of the required sensitivity to support spectral line cases. Lower noise figures are on the left of the plot. A spectral line sensitivity of order 10nJy/bm/km/s is required in a 'reasonable' period to support the driving use cases and approximately 80% of all identified cases, where a reasonable period might be of order ten hours.

2.3 Antenna Baselines

The use cases suggest that an array spanning from 1 km to approximately 250 km is required to support the identified driving use cases. The tail of the distribution extends out to continental scales. The distribution of array extent required to support the use cases is shown in Figure 5.

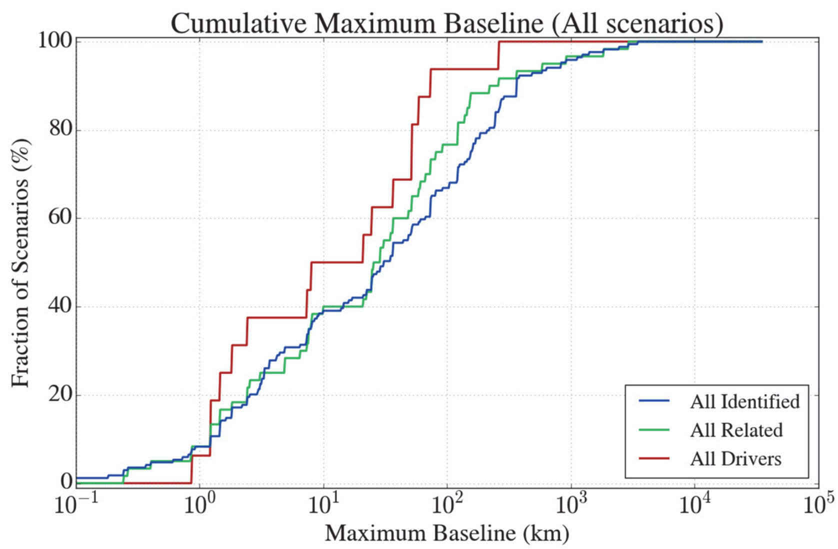


Figure 5 - Cumulative histogram of maximum baseline required. Approximately 35% of driving cases can be supported with a compact core of 2-3km extent. 60% can be supported with an array of tens of km in extent (roughly the envelope of the current VLA). The remaining driving cases extend out to a few hundred km in extent. For all identified cases, there is a long tail out to continental scales.

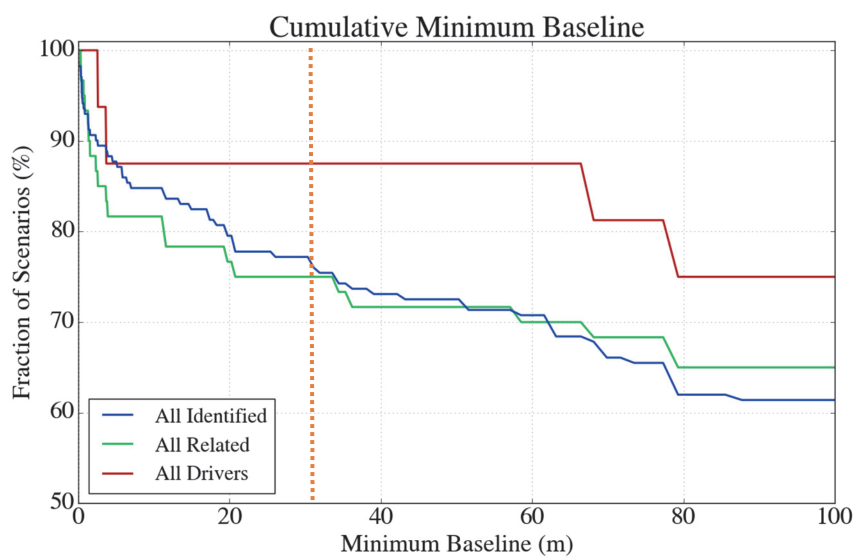


*Figure 6 - Cumulative histogram of the minimum baseline required to recover the largest angular scale of interest. Approximately 75% of cases can be supported by a 1.75*D spacing between 18m diameter dishes, shown as a dashed vertical line. The remaining 25% of cases require a large single dish, or short spacing array and total power antennas.*

The shortest spacings of the array are dictated by the largest angular scale of interest. A cumulative histogram of the shortest baseline required is shown in Figure 6. Approximately 75% of identified cases are supported by a minimum spacing of 30m, which corresponds to the shortest spacing achievable with 18m offset apertures. Recovering the full range of scales for all cases suggests a need for a short spacing array plus a total power mode, or a single dish to recover total flux.

The distribution of surface brightness sensitivity requirements suggests a need for brightness sensitivity over a range of scales that largely follows a B_{MAX}^2 relationship. This is shown in Figure 7 and Figure 8 for spectral line and continuum cases, respectively.

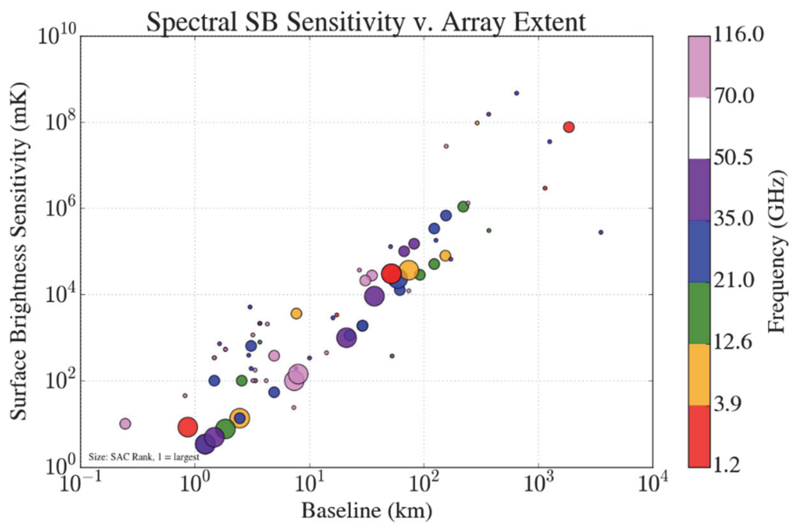


Figure 7 – Surface brightness sensitivity required for spectral line observations as a function of array extent. The distribution follows a B_{MAX}^2 relationship, with sensitivity of a few mK at 1km scales, and of order a few hundred Kelvin at 100km scales.

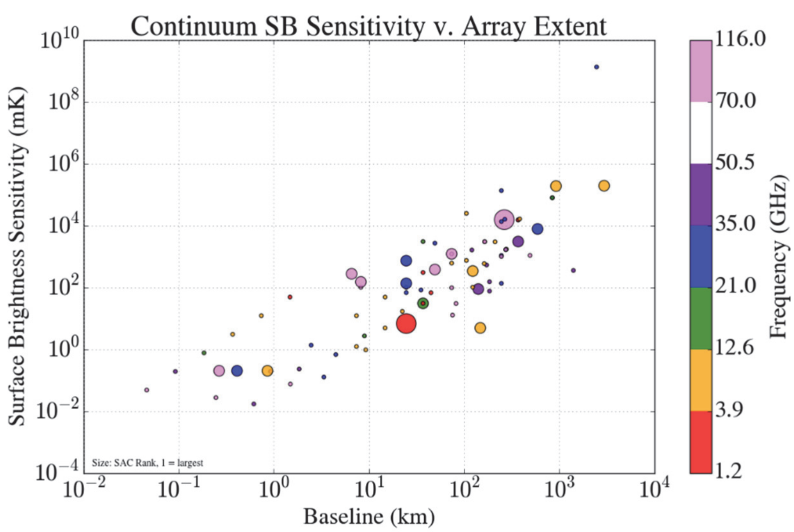
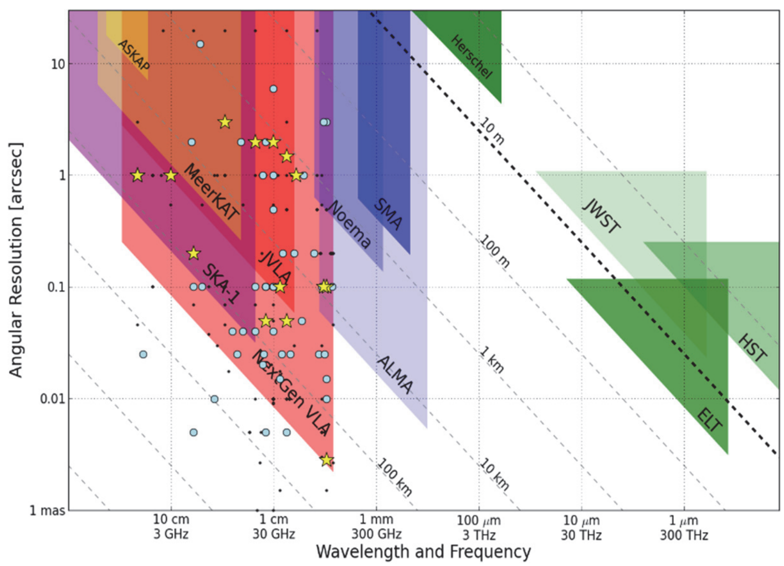


Figure 8 – Continuum surface brightness sensitivity required as a function of angular extent. The community's requests also follow a B_{MAX}^2 relationship, with sensitivity of 0.1mK at 1km and of order 1K at 100 km supporting a majority of cases.



v25

Figure 9 – Distribution of use cases compared to the angular resolution of the ngVLA. Driving use cases are shown as stars, with related use cases as blue circles, and all remaining identified cases as black dots. A large number of cases are within angular scales accessible by other arrays, but require significantly more sensitivity than those arrays can provide on 1000hr time scales. The ngVLA use cases are demanding angular resolution and/or sensitivity that are presently unavailable.

Most use cases require improvements in sensitivity, angular resolution, or both. This is evident in Figure 9, which overlays the required resolution of the observation on a plot comparing the resolution of the ngVLA to other existing and planned arrays. A number of cases explore the angular resolution available with the ngVLA, while others are within the angular resolutions achievable with the JVLA and other instruments. In the latter cases, high surface brightness or point source sensitivity is required to support the observation.

2.4 Field of View

The field of view requirements for the use cases vary, with of order 10% of the identified cases consisting of large area mappings that demand survey speed, and the remaining 90% of cases achievable with either small mosaics or single pointings. As shown in Figure 10, with an 18m aperture, of order 60% of identified cases are single pointings.

All high-ranking driving use cases can be accommodated with either single pointings or small mosaics, as shown in Figure 11. Sensitivity should therefore be prioritized over survey speed, but the selected aperture size should be balanced since more apertures (a larger N array) are desirable for image fidelity.

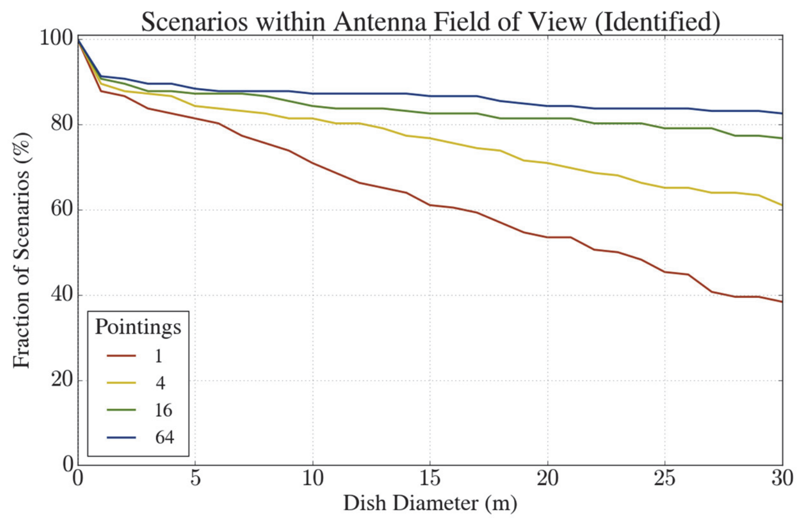


Figure 10 – Fraction of observations' desired field of view that fit within 1, 4, 16 or 64 pointings as a function of antenna diameter.

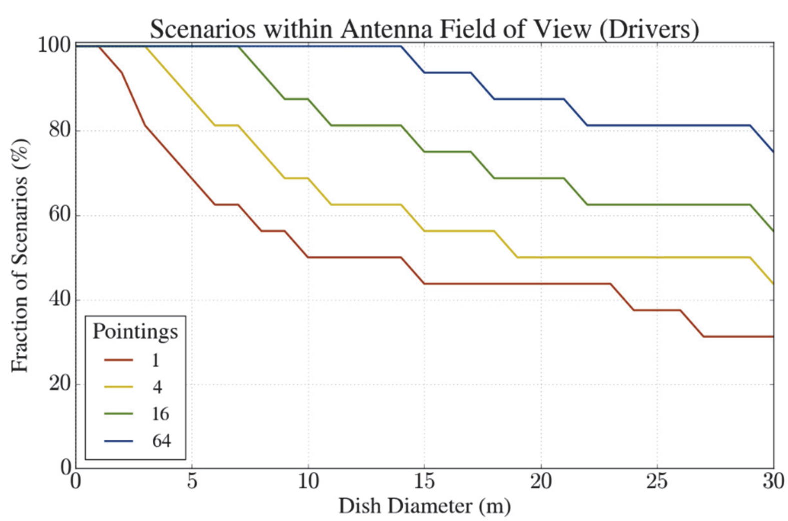


Figure 11 – Fraction of 'driving' use case observations with a field of view that can be imaged with 1, 4, 16 or 64 pointings. At a 14m aperture, all fields fit within 64 pointings, and 50% are single pointings. For an 18m aperture, of order 90% fit within the same mosaic size.

3 Workshop Findings

A workshop was held on the New Mexico Tech campus in Socorro, NM, on June 26-29 to further develop the ngVLA science program. Approximately 150 people participated in the workshop and provided feedback on the design and science case.

On the final day of the workshop, the ngVLA Science Working Groups (SWGs) were asked to evaluate the relative importance of various design alternatives to their key science goals. This provided an opportunity to confirm or modify our initial findings. Each SWG held a break-out session, followed by co-chair reports of their findings.

The following were points of convergence at the workshop:

- The array should have full 1.2 to 116 GHz frequency coverage at all angular scales accessible with the array.
- Sensitivity is a driver for almost all cases, and the spectral line cases are the most demanding. The down-select for front end concepts should primarily be based on a sensitivity metric that is irrespective of bandwidth.
- The array configuration should be fixed, rather than reconfigurable. There was a general preference for fixed array configurations for a number of reasons, such as the budget allocation to total collecting area, reconfiguration cycle time impact on data turn-around, etc.
- The array configuration should have a distribution of collecting area that provides brightness temperature sensitivity at a wide range of angular scales, from approximately 10 to 1000 mas.
- Surface brightness sensitivity should be prioritized over angular resolution. Most of the driving use cases could be accommodated with 150-200km maximum baselines, but there is certainly interest in the 200km to 1000km intermediate scales too, so long as there is enough sensitivity at these finer angular scales.
- The array concept should have single pixel feeds only. Support for multi-pixel options waned once cost was a consideration. There was some support for shaping the antennas less radically to permit future W-band FPAs, but not if it costs more than a few percent in sensitivity.
- Given limited development resources, there was support for proceeding with an 18m homogeneous array concept. This concept appeared to meet the majority of community requirements.
- There are a number of key science cases that would benefit from short-spacing information. It is clear that a mechanism to recover larger scale structures and/or total flux should be available, but there was no consensus on how best to provide this capability (existing or new single dish, short spacing array, etc.)
- Flexibility in the front-end system and correlator will be important. The system needs to support mixed modes (such as a high spectral resolution window in a continuum observation), tunability to both ends of the band simultaneously, etc.
- Sub-array capabilities should be provided, e.g., for multiple phased-arrays performing time-domain research concurrent with imaging observations. A minimum of ten sub arrays should be supported by the system.

Additional input from the workshop is summarized in the Science Use Case Compatibility Matrix shown in the Appendix. Note that a low ranking for a feature does not indicate that it is expendable, rather that the particular use case is not dependent on that feature.

4 Open Questions & Future Work

While there was a significant amount of consensus expressed at the meeting, a few issues were also identified as open questions. These issues require additional study in the development of the reference design:

- Recovering large-scale structure / total flux.
- Phase calibration strategy & operational overheads.
- Longest baseline in the connected array: 300 km vs 1000 km.

The large-scale structure and phase calibration questions are primarily technical. The needs are relatively clear, but the preferred concepts to meet the needs are not. Technical studies should be performed to evaluate the available options.

The 300km to 1000km baseline question requires further scientific input on the merits (and trade-offs) of extending the array concept to these scales. This question is currently being addressed by the SAC, who are in the process of evaluating the merits of the additional science enabled by higher angular resolution within the array concept.

Additional analysis of the use cases is also desirable. In particular, the array time required to support each use case could be better estimated. A reference observing program could then be constructed that would support a majority of the identified science over a 10-20 year period.

Data rates and compute capacity requirements could also be estimated for the reference observing program, providing insight into the data post-processing system design. Where science cases do not require the full array, possible pairings in sub-arrays can be assessed to inform the correlator system requirements.

5 Acknowledgments

We would like to thank the Science Advisory Council for leading the use case capture effort, and the Science Working Groups for documenting the science use cases. The response to this call was overwhelming - it demonstrated the community’s need for this instrument and provided valuable insight into the emerging ngVLA concept.

Appendix

ngVLA Sci/Tech Workshop
Science Use Case Compatibility Matrix

	Parameter	CoL/PF3: Characterizing Planet-Disk Interactions	CoL/AC5:Prebiotic Chemistry	CoL/SS6: Giant Planet Atmospheres	New: Formation of Solar System Analogs (subsumes PF3?)	GEco/NGA2: Atomic Hydrogen in the Local Universe	GEco/NGA8: Parsec-Scale Cold Gas Structure Across the Whole Local Galaxy Population	GEco/NGA3: Radio Continuum Emission from Galaxies: An Accounting of Energetic Processes	GFor/HiZ1: The Molecular Gas Budget	GFor/HiZ5: Mapping High-z CO Gas	GFor/HiZ7: Deep Continuum Surveys	TdCP1: Galactic Center Pulsars (Gravity Probes)	TdCP9: High Precision Astrometry	TdCP: Explosive Transients (FRB, TDE, GRB, Mdwarfs, GWEM...)
Feature Comparison														
Functional Capabilities & Features	Phased Array	0	0	0	0	0	0	0	0	0	0	6	6	1
	VLBI Recording	0	0	0	0	0	0	0	0	0	0	3	6	4
	Full Polarization Synthesis	3	3	5	5	4	4	6	1	1	6	6	6	6
	High Accuracy Linear Pol Measurements (0.1% Purity)	3	2	4	5	3	3	6	0	0	4	3	2	4
	High Accuracy Circular Pol Measurements (0.1% Purity)	0	0	4	2	3	3	4	0	0	2	3	2	4
	Accurate Autocorrelation (Total Power) Products	0	0	0	0	2	2	2	0	0	0	0	0	0
	Total Power Single Dish / Zero Spacing Element	0	2	4	0	6	6	6	2	0	0	0	0	0
	Short Spacings (< 30 m)	0	4	5	2	6	6	6	0	2	0	0	0	0
	Reconfigurability	0	0	0	0	2	2	2	0	1	0	0	0	0
	> 3 SubArrays	0	0	0	0	0	0	0	0	0	0	5	3	4
	1000 Spectral Channels	2	2	5	2	6	6	6	6	6	5	6	5	6
	> 64k Correlator Spectral Channels	0	6	3	6	0	5	0	0	0	0	0	0	4
	Mosaic / On-The-Fly Mosaic	0	1	6	0	6	6	6	6	0	6	0	2	5
Array Performance	Point Source Senstivity: 10x JVLA, ALMA	6	6	4	6	6	6	6	6	6	6	6	6	6
	Surface Brightness Sensitivity: 5x JVLA, ALMA	6	6	4	6	6	6	6	4	6	5	0	0	2
	Survey Speed: 5x JVLA, ALMA	0	0	5	0	6	6	6	6	0	6	0	5	5
Baselines	1 km	3	4	6	3	6	6	6	6	4	4	4	0	4
	3 km	4	6	6	4	6	6	6	4	6	6	6	0	6
	30 km	5	5	4	6	6	6	6	0	6	6	4	4	6
	300 km	6	4	0	6	2	2	6	0	0	2	3	6	5
	1000km	0	0	0	0	0	0	1	0	0	0	3	6	5
	4000km	0	0	0	0	0	0	1	0	0	0	0	6	5
RF / FE Features	High RF Dynamic Range (e.g., Solar)	0	0	0	0	0	0	0	0	0	0	0	0	0
	< 1 GHz	0	1	1	0	0	0	3	0	0	1	4	1	3
	1-4 Ghz	2	4	5	3	6	4	6	0	0	5	6	4	6
	4-11 GHz	4	5	5	4	0	6	6	2	2	6	6	6	6
	11-50 GHz	6	6	6	6	0	6	6	6	6	6	6	6	6
	70 - 115 Ghz	6	6	6	6	0	6	6	6	6	5	2	4	4
	Future Multi-Pixel 70-115 GHz Option	0	0	5	0	0	4	4	3	0	3	0	0	1
	Circular Polarization Front End	0	0	0	0	2	2	2	0	0	0	0	0	0
	> 2:1 BW Ratio	4	4	3	4	0	4	5	4	2	4	5	4	5
	> 20GHz BW	5	5	4	6	0	5	4	4	1	5	0	4	5
Time Domain Features	Rapid Response Time (1 minute)	0	0	0	0	0	0	0	0	0	0	0	0	6
	msec-scale Search Capabilities	0	0	0	0	0	0	0	0	0	0	5	0	6
	< msec Search Capabilities	0	0	0	0	0	0	0	0	0	0	5	0	5
Other	Ability to do continuum and spectral line at same time	5	6	6	6									
	Flux density scale calibration accuracy	3	3	6	3									
	Bandpass calibration accuracy	2	5	5	5									
	Multiple Beamforming (>10)											5	3	0
	Atmospheric Phase Calibration Strategy					0	6	6						
Concepts - Comparison														
Concept A	18m Homogeneous Array	6	5	5	6	6	6	6	5	6	6	4	4	4
Concept B	13/26m Heterogeneous Array	6	6	6	6	0	0	0	6	5	4	3	3	3

KEY:

	Features	Concept Fit
0	Not Required	Poor Fit
1		
2		
3	Desirable	Moderate Fit
4		
5		
6	Essential	Strong Fit