Scientific Data Flow and Array Simulation Analysis for the SKA1 Era

Authors: Guo Shaoguang, Lu Yang, An Tao, Lao Baoqiang, Xu Zhijun, Wu Xiaocong, Lü Weijia, Guo Shaoguang

Date: 2022-06-28T00:00:00+00:00

Abstract

As the next-generation radio telescope, the Square Kilometre Array (SKA) has undergone years of preparation, with Phase 1 (SKA1) commencing construction in July 2021. Upon formal operation, SKA1 is expected to generate 750 PB of scientific archived data annually, which will be stored at SKA Regional Centres worldwide for use by researchers. This paper conducts quantitative analysis of models at various stages including SKA observation stations, the Central Signal Processor, scientific data processing, and Regional Centres, using high-priority scientific observations of SKA1 as the primary basis to derive data flow assessments for each stage and computational requirements for scientific data processing. Taking the current SKA1-Low and SKA1-Mid arrays as examples, key factors affecting interferometric array configuration are summarized, including resolution, sensitivity, and UV coverage. Finally, OSKAR is used to simulate interferometric array data. Simulations of SKA1-Mid demonstrate system scalability and stability, while simulations of SKA1-Low on CSRC-P show that the design of the China SKA Regional Centre prototype has undergone thorough verification and optimization, yielding detailed computational requirements and specific information on data volumes. The demands of SKA for data processing, computation, storage, etc., will require joint advancement from electronics, communications, information technology, computer science, and interdisciplinary fields.

Full Text

Scientific Data Flow and Array Simulation Analysis for the SKA-1 Era

Guo Shaoguang[1,2]*, Lu Yang[1], An Tao[1,2], Lao Baoqiang[1,2], Xu Zhijun[1,2], Wu Xiaocong[1,2] & Lv Weijia[1,2]

[1]SKA Regional Centre Joint Lab, Shanghai Astronomical Observatory, Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Shanghai 200030, China
[2]SKA Regional Centre Joint Lab, Peng Cheng Laboratory, Shenzhen 518066, China

*Corresponding author, E-mail: sgguo@shao.ac.cn
Received: 2022-06-28; Accepted: 2022-07-xx
Supported by the SKA Special Program (No. 2020SKA0110300), National Key R&D Program (No. 2018YFA0404603), National Natural Science Foundation of China (Nos. 11873079, 12041301), and the Youth Innovation Promotion Association of Chinese Academy of Sciences (No. 2021258)

Keywords: Square Kilometre Array, data simulation, synthesis aperture, data format
PACS: 47.27.-i, 47.27.Eq, 47.27.Nz, 47.40.Ki, 47.85.Gj
doi: ??

1. Introduction

The Square Kilometre Array (SKA) represents an international mega-science project that will construct the world's largest and most sensitive radio telescope. After years of preparation, SKA Phase 1 (SKA1), which possesses 10% of the SKA's total receiving capacity, commenced construction in July 2021. SKA1 comprises two sites: SKA1-low in Australia and SKA1-mid in South Africa, covering frequencies from 50 MHz to 15.3 GHz.

SKA employs synthesis aperture technology to combine signals from array elements, providing an effective aperture equivalent to a one-kilometre radio telescope. The instrument will generate massive data volumes. For SKA1-mid, the Central Signal Processor (CSP) located near the array performs beamforming and correlation processing on 18 Tb/s of digitized data streams before transmission to a dedicated supercomputer in Cape Town for further processing. This data flow processing is critical to the overall SKA system design and poses enormous challenges for subsequent data management and processing. Current estimates suggest that transferring data from the Science Data Processor (SDP) in host countries to various SKA Regional Centres (SRCs) will require high-speed, stable intercontinental networks comparable to the Worldwide Large Hadron Collider Computing Grid (WLCG). To ensure stable and timely delivery of SKA1 data to regional centres, network bandwidth must reach at least 100 Gbps. Following formal operations, SKA1 is projected to distribute 710 petabytes annually to SRCs, requiring at least 25 PFlops of computing power. These requirements present two major technical challenges: providing sufficient computing power at low energy consumption for data stream processing, and distributing such vast data quantities to SRCs worldwide.

As of the end of 2021, the world's fastest High Performance Computing (HPC) system was Fugaku, installed at the RIKEN Center for Computational Science in Kobe, Japan, with an HPL benchmark performance of 442 PFlops. This supercomputer employs Fujitsu's custom ARM A64FX processors and Tofu D interconnect technology for inter-node data transfer. Meanwhile, emerging network technologies such as Multi-Core Fiber (MCF) have enabled global internet bandwidth to reach 618 Tbps. While SKA1's computing and data transmission demands are modest compared to the fastest HPC systems and global internet infrastructure, they represent an exceptionally massive scale for a single scientific research domain. Moreover, SKA's diverse scientific research directions and observation modes result in highly complex scientific data processing pipelines. Conventional HPC systems cannot meet these diverse and customized requirements, necessitating heterogeneous hardware in SRCs to implement different processing pipelines for various scientific observations, classified according to data-intensive, compute-intensive, and storage-intensive characteristics.

This paper provides a detailed analysis of SKA1 data flow from array data reception to final scientific data delivery (including images, metadata, pulsar candidates, etc.) at SRCs. We present model-based estimates of data transmission rates and computing power requirements. Since SKA1 observation configurations are not yet finalized, current data flow assessments remain preliminary. Through quantitative description of SKA1 array layout principles and experimental testing of different arrays on the CSRC-P platform, we derive specific data flow characteristics. These results illustrate the overall SKA1 data flow landscape and its impact on CSRC-P computing, storage, and processing capabilities, offering recommendations and planning guidance for future CSRC-P data simulations.

2. SKA Data Flow

Radio signals from celestial sources arriving at the SKA telescope array are received by antennas, converted to digital signals by Analog-to-Digital Converters (ADCs), and then processed by the Central Signal Processor (CSP) to output visibility data. These data undergo preliminary pre-calibration to generate calibration data, image data, and other metadata. After deep processing, the data become accessible to scientific teams for subsequent research.

As shown in [Figure 1: see original paper], the entire data observation and processing workflow involves SKA1-low peripheral stations transmitting processed data through Remote Processing Facilities (RPF) and core stations through Central Processing Facilities (CPF) to the Science Operations Centre (SOC). SKA1-mid stations transmit directly via CPF to SOC, after which the Science Processing Centre (SPC) distributes data to SRCs across continents. SRCs then authorize data access for scientists and users.

Based on the current design baseline, detailed SKA1 parameters are presented in , where N_f represents the maximum number of frequency channels, which will not be exceeded in actual observations. For most SKA1-mid observation modes, this value is relatively small; for example, SKA-VLBI typically requires only 64K frequency channels.

SKA has established five Key Science Projects (KSPs): probing the cosmic dawn and dark ages, galaxy evolution and cosmology with dark energy, strong-field tests of gravity using pulsars and black holes, cosmic magnetism origins and evolution, and the cradle of life. These KSPs address unresolved questions in fundamental physics, astrophysics, and cosmology, promising major discoveries such as detecting the universe's first light. With SKA1 construction underway, important scientific themes not requiring the full array's sensitivity, resolution, or frequency coverage have been designated as High Priority Science Objectives (HPSOs). China has established a "2+1" strategic deployment ensuring two priority breakthrough areas and several distinctive research directions, launching two SKA science programs in 2020: the Epoch of Reionisation (EoR) project and the pulsar project. EoR will reproduce cosmic dawn using current SKA precursor array data while participating in large-area low-frequency surveys to statistically reveal the universe's overall structural characteristics during reionisation. The pulsar program includes Pulsar Search (PSS) and Pulsar Timing (PST), focusing on testing gravitational wave theories and detecting supermassive black hole merger events.

Based on experience with SKA precursor arrays including MWA, ASKAP, and LOFAR, data archived to SRCs includes not only data cubes but also optimized UV data from pipeline processing, intermediate products for calibration pipelines, UV gridded data, Point Spread Function (PSF) cubes, and multi-resolution weighting information.

The following data flow estimates are based primarily on HPSO scientific objectives and their observation time allocations, covering initial data generation, correlation processing, and pre-processing. Final scientific data also includes observations led by Principal Investigators (PIs), other scientific discoveries, and SKA-VLBI projects.

2.1 Observation Station End

After antennas receive raw analog signals, the signals are digitized. To improve temporal and frequency resolution, Finite Impulse Response (FIR) filtering and Fast Fourier Transform (FFT) operations are applied. For SKA1-low, beamforming is also performed for signal calibration and delay compensation.

Assuming N_{ant} antennas, each with N_{pol} polarization outputs, signal acquisition bandwidth $\Delta\nu$, and according to the Nyquist sampling theorem requiring a sampling rate of at least $2\Delta\nu$, the computational load using an N-order FIR is approximately $2N_{tap}$. Based on FFT algorithm complexity $O(n \log n)$, the number of FFT operations is $N \log N$. The total operations required per station after polyphase filtering are of the order:

$$2N_{tap}N_{ant}N_{pol}\Delta\nu + 2N_{bands}\log_2(2N_{band})N_{ant}N_{pol}\Delta\nu_{band}$$

For beamforming, calibration is achieved by providing correct delays to each antenna, primarily for high temporal and frequency resolution scientific applications such as pulsar studies and interplanetary scintillation. The SKA1-low Correlator BeamFormer (CBF) receives data from 512 stations and generates four Stokes parameters. However, the computational load for calibration is relatively small compared to FFT and FIR operations and is temporarily ignored in this evaluation. Therefore, the overall output data rate per observation station is approximately:

$$R_{station} \propto N_{beams}N_{bands}N_{pol}\Delta f_{band}S_{sample}$$

This expression shows that acquisition bandwidth and sampling bit depth significantly affect data flow from each station or array. Based on , the data rate from stations to CSP is estimated to be on the order of Tbps. The mid-frequency array comprises 197 telescopes (64 of which belong to MeerKAT), with a raw data output rate of approximately 2 TB/s. The low-frequency array comprises 131,072 antennas with a raw data output rate of 158 TB/s.

Correlation processing output depends primarily on integration time and channel number. Baseline Dependent Averaging (BDA) technology can reduce visibility data volume based on radio interferometer baseline distribution. Full-scale SKA1-low simulations indicate that BDA can reduce visibility data volume by approximately 50%–85% across different time intervals.

2.2 CSP

CSP performs correlation processing on station data, including delay compensation to eliminate fringe rotation caused by Earth's rotation, finer channelization for improved resolution, and bandpass filtering to correct polyphase filter effects. Channel granularity varies according to scientific objectives; here we assume N_{chan} channels.

CSP also performs Radio Frequency Interference (RFI) removal, though its computational load is small and temporarily excluded from this estimate. The CSP output data rate is primarily generated by pairwise correlation of stations:

$$R_{csp} \propto N_{beams}N_{chan}N_{stat}^2N_{tdump}$$

Since output data rate is proportional to the square of station count and channel number, the estimated input visibility data rate is 7.15 Tbps for SKA1-low and 6.4 Tbps for SKA1-mid.

For SKA1-mid, CSP divides input bandwidth from antennas into multiple frequency channels for cross-correlation processing. The input data stream consists of 133 SKA antennas (each with 100 Gbps output bandwidth) and 64 MeerKAT antennas (each with 40 Gbps output bandwidth), totaling approximately 57 Tbps input bandwidth. Imaging and pulsar processing employ different pipelines: imaging uses two-stage channelization with 512 coarse channels and 64,000 fine channels, while pulsar search provides approximately 75 kHz channelization across 300 MHz sub-bands, reducing time resolution to ~64 µs. Data are output at 6.4 Tbps.

Correlation processing outputs visibility data with maximum resolution achieving 0.1 s integration time across all SKA1-mid baselines.

2.3 SDP and SRC Computing Load Requirements

SDP receives CSP output data for preliminary imaging pre-processing. While SKA's final imaging strategy remains undetermined, current precursor projects (LOFAR, MWA, ASKAP) indicate processing will primarily include calibration and iterative major/minor cycles for imaging and deconvolution, with major/minor cycles consuming most computational resources. After several iterations, SDP derives new calibration parameters using an improved sky model and initiates new calibration cycles to enhance the model further, ultimately creating final sky images by restoring sources from residual images.

During SKA1, SDP data from correlation processing includes primarily images and calibration data, typically without retaining visibility data—except for the SKA1-low EoR high-priority project, which requires visibility data for effective foreground removal. SKA1-mid SDP receives input data from CSP and Telescope Manager (TM), generating scientific and calibration products including continuum and spectral line cube imaging data, pulsar search candidate data, transient source detection data, single-dish intensity data, and atmospheric/ionospheric calibration data.

A major SKA challenge involves managing and processing data from diverse pipelines to produce high-dynamic-range images and other data products for scientific support. According to current SDP design, standard data products are defined per observation project. Imaging observations can generate maximum-resolution data products across three dimensions (two spatial, one spectral), polarization, and images. The most significant data volume reduction occurs during visibility-to-image processing.

Detailed observation data flow is shown in [Figure 1: see original paper], with product data flow illustrated in [Figure 2: see original paper]. Raw observation data generated at stations undergo preliminary processing at SDP to produce initial scientific data products, primarily imaging data (images or gridded visibility data) and other data. Imaging data size depends heavily on channel count, with final file size proportional to channel number. These data are distributed via intercontinental networks to sub-networks at each regional data center, where astronomers access them via internet connections. This processed data is termed Advanced Data Products (ADP), which can be provided to researchers through SKA regional centers or cloud solutions.

Among all SKA1 processing pipelines, imaging dominates computational demand and resources. As shown in [Figure 3: see original paper], EoR imaging and power spectrum observations occupying approximately 15% and 31% of SKA1-low observation time correspond to 31% and 55% of computational load, respectively. Non-imaging pipelines such as the Pulsar Search Pipeline (PSP) require relatively less computing, accounting for 39% of HPSO time on SKA1-low and 5% on SKA1-mid. For SKA1-low, the scientific data rate is approximately 3 Gbps without EoR UV visibility data, increasing to 25 Gbps when included. For SKA1-mid, the scientific data rate is approximately 9 Gbps.

The PSP requires only Real-time Calibration (RCAL), while continuum imaging additionally needs Iterative Self-Calibration (ICAL), Data Preparation pipeline producing Taylor-term images (DPrepA), and Data Preparation pipeline producing coarse channelized images (DPrepB). Spectral line imaging pipelines further require Data Preparation pipeline producing fine channelized images (DPrepC).

Weighted calculations from [Figure 3: see original paper] indicate final computing loads of at least 13.6 PFLOPS for SKA1-low and 11.5 PFLOPS for SKA1-mid, yielding a total SKA1 SRC computing requirement estimate of 25.1 PFLOPS.

3. SKA1 Array Layout

To achieve SKA's scientific objectives while balancing construction costs and technical feasibility, a telescope must be built that meets extreme performance requirements at a relatively controllable cost. This requires evaluating current scientific needs and technical parameters including baseline length, array size, antenna sensitivity, and power consumption. The final configuration confirmed a dual-site solution in Australia and South Africa.

SKA array layout and design prioritize KSP observation requirements, increasing baseline length to improve spatial resolution and employing a core array to enhance UV coverage and sensitivity. Design considerations include frequency range, sensitivity, observation bandwidth, polarization capability, collecting area, baseline length, and processing capacity. Key factors include:

- **Resolution:** Meeting high-priority science resolution requirements, which depends on frequency configuration. Higher frequencies achieve higher spatial resolution more easily. For example, a 150 km array at 8 GHz achieves 0.0629 arcsecond resolution.
- **UV Snapshot Coverage:** The two-dimensional baseline vector distribution determines instantaneous spatial frequency coverage, which should maximally fill the UV plane. In spiral array layouts, snapshot coverage is determined by spiral rotation angle, antenna positions along spiral arms,

and arm count. [Figure 4: see original paper] and [Figure 5: see original paper] show adjusted spiral angles and antenna positions to provide good UV coverage.

- **Core Region:** The core region provides sufficient sensitivity for pulsar search and HI line observations.

3.1 SKA1-mid

SKA1-mid operates at 0.35–15.3 GHz in South Africa's Karoo desert, focusing on high-sensitivity observations of pulsars, 21-cm neutral hydrogen, and continuum emission. The telescope consists of 133 offset Gregorian antennas of 15 m diameter plus 64 antennas of 13.5 m diameter from MeerKAT. SKA1-mid features a 1 km diameter core array, a randomly distributed 3 km two-dimensional array, and three spiral arms with maximum baseline of 150 km, as shown in [Figure 4: see original paper].

3.2 SKA1-low

SKA1-low operates at 50–350 MHz at the Murchison Radio-astronomy Observatory in Western Australia, focusing on cosmic reionisation, pulsars, and exoplanetary systems. SKA1-low comprises 131,072 log-periodic antennas divided into 512 stations of 256 antennas each. Of these, 296 stations are configured in a central core region, with the remaining 216 deployed along three spiral arms reaching a maximum baseline of 65 km, as shown in [Figure 5: see original paper].

Current aperture synthesis array technology has relatively mature calibration and imaging algorithms, but software and algorithm development is still required to support SKA1-mid's data volume and achieve better dynamic range with reduced human intervention.

4. Experimental Testing

This section presents OSKAR simulations for SKA1-low and SKA1-mid, describing simulation methodology, execution, computational time, and resulting data file sizes.

OSKAR, developed by the University of Cambridge, is SKA simulation software that generates visibility data based on the Radio Interferometer Measurement Equation (RIME). It simulates station beams and interferometric array data, generating binary visibility data or Measurement Set (MS) format, converting FITS files to sky models, and adding noise to visibility data.

The simulation strategy employed sky models with 9, 100, 1,024, and 10,000 sources; observation durations from 1 second to 24 hours (1 s, 6 s, 60 s, 600 s, 1,200 s, 2,400 s, 3,600 s, 7,200 s, 14,400 s, 28,800 s, 43,200 s, 54,000 s, 72,000 s, 86,400 s); default single frequency channel; SKA1-mid observation frequency at 1.4 GHz with 1 kHz bandwidth; and SKA1-low observation frequency at 200 MHz with 30 MHz bandwidth.

4.1 SKA1-mid Simulation Testing

Due to moderate computational requirements, SKA1-mid simulations used two GPU models to evaluate system scalability and stability: NVIDIA Tesla K40M and Tesla K20X. The Tesla K20X (Kepler architecture) features 2,688 cores, 6 GB memory, 249.6 GB/s bus bandwidth, and 1,312 GFLOPS double-precision theoretical performance. The Tesla K40M (Kepler architecture) features 2,880 cores, 12 GB memory, 288.4 GB/s bus bandwidth, and 1,682 GFLOPS double-precision theoretical performance. As shown in [Figure 6: see original paper], [Figure 7: see original paper], and [Figure 8: see original paper], K40M demonstrates improved performance over K20X, though the enhancement is modest due to their architectural similarity. K40M maintains advantages with lower time consumption in multi-point simulations.

4.2 SKA1-low Simulation Testing

Due to SKA1-low's large number of stations and high computational demands, simulations were conducted on the China-SRC platform (CSRC-P) equipped with NVIDIA Tesla V100 SXM2 GPUs. The V100 (Volta architecture) features 5,120 cores, 32 GB memory, 897 GB/s bus bandwidth, and 7,834 TFLOPS double-precision theoretical performance. SKA1-low simulations show that for observation durations under 2,400 s, simulation time remains under 100 seconds, demonstrating V100's substantial computational advantages.

As shown in , final simulation data size correlates weakly with sky model source count, with simulation time remaining relatively consistent. Simulation time and final visibility data size depend primarily on observation duration. By setting observation frequency channels, file size is proportional to this value. For SKA1-low simulations using the maximum frequency channels from , visibility data generation reaches approximately 2.048 GB per second, or about 2 TB/s.

5. Summary and Outlook

SKA will undoubtedly be one of the largest radio observational facilities ever constructed. Through model-based estimation, analysis, and simulation of SKA1's complete system data flow, the Tbps data streams, petabyte-scale archival storage requirements, and observation complexity pose the most severe challenges for data management, processing, computing, storage, and networking. This discussion of the entire data flow provides system-level support for the future operation of China's SKA Regional Centre. The project will integrate radio astronomy technology with modern information technologies including communications and computer science. SKA's ultra-large data flow will fundamentally transform astronomical research methods, requiring revolutionary improvements and breakthroughs in traditional framework design for successful operation.

The SKA1 phase demands interdisciplinary expertise spanning electronics, computer technology, and information science. SKA will particularly drive advances in information technology, astronomical software, and communications technology. The authors express gratitude to reviewers and those who contributed to this work. This research utilized resources from the China SKA Regional Centre prototype, funded by the National Key R&D Program for Large Scientific Facilities (2018YFA0404603).

References

1. An T, Wu X P, Hong X Y. SKA data take centre stage in China. *Nat Astron*, 2019, 3: 1030–1030
2. Guo S G, Zheng X Y, Mao Y F, et al. Scheme and Prospect of the SKA Big Data Transferring (in Chinese). *E-science Technology & Application*, 2018, 9(3): 3–13
3. Swart G P, Dewdney P E. Highlights of the SKA1-Mid Telescope architecture. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2022, 8(1): 011021
4. Quinn P, van Haarlem M, An T, et al. SKA Regional Centres White Paper v1.0. Technical Report, Square Kilometre Array
5. Guo S G, An T, Xu Z J, et al. Progress and Prospect of transcontinental high-speed data transmission at SKA Regional Center in China (in Chinese). *ChinaXiv*:T202206.00291 (2022)
6. Jongerius R, Wijnholds S, Nijboer R, et al. An End To End Computing Model for the Square Kilometre Array. *Computer*, 2014, 47(9): 48–54
7. An T, Wu X C, Lao B Q, et al. Status and progress of China SKA Regional Centre prototype. *arxiv*:2206.13022
8. Quinn P, Axelrod T, Bird I, et al. Delivering SKA science. In: *Proceedings of Advancing Astrophysics with the Square Kilometre Array (AASKA14)*. Giardini Naxos, 2015
9. An T. Science opportunities and challenges associated with SKA big data. *Science China: Physics, Mechanics & Astronomy*, 2019, 62(8): 121–126
10. Hollitt C, Johnston-Hollitt M, Dehghan S, et al. An Overview of the SKA Science Analysis Pipeline. In: *Astronomical Data Analysis Software and Systems XXV*. San Francisco: Astronomical Society of the Pacific, 2017. 367–370
11. Lao B Q, Zhang Y K, An T, et al. Software Platform on China SKA Regional Center Prototype System (in Chinese). *ChinaXiv*:202206.00173
12. Xu Z J, An T, Guo S G, et al. A machine learning dataset for FRB detection in raw data (in Chinese). *ChinaXiv*:T202206.00321
13. Wei J W, Zhang C F, Zhang Z L, et al. Parallel optimization of the pulsar search pipeline on China SKA Regional Centre Prototype (in Chinese). *ChinaXiv*:T202206.00297

14. Wei J W, Zhang C F, Lao B Q, et al. Optimization of parallel processing of Square Kilometre Array low frequency imaging pipeline (in Chinese). *ChinaXiv*:T202206.00292
15. Braun R, Bourke T L, Green J A, et al. Advancing astrophysics with the square kilometre array. In: *Proceedings of Advancing Astrophysics with the Square Kilometre Array (AASKA14)*. Giardini Naxos, 2015
16. Dewdney P. SKA1 System Baseline Design. Technical Report, SKA Organisation, 2016
17. Operational Model for inclusion of SKA in Global VLBI, JIV-ERIC and Square Kilometre Array Phase 1 Project, Submission date: 15/6/2020
18. Carrilli C L, Rawlings S. Science with the Square Kilometer Array: Motivation, Key Science Projects, Standards and Assumptions. *New Astronomy Reviews*, 2004, 48(11): 979–984
19. Bourke T, Braun R, Fender R, et al. Advancing Astrophysics with the Square Kilometre Array (AASKA14)
20. Whitney A, Booler T, Bowman J, et al. The Murchison Widefield Array (MWA): Current Status and Plans. In: *American Astronomical Society, AAS Meeting #218. Bulletin of the American Astronomical Society*, 2011, Vol. 43
21. Haarlem M P V, Wise M W, Gunst A W, et al. LOFAR: The Low Frequency Array. *Astronomy & Astrophysics*, 2013, 556(7)
22. Jonas J L. MeerKAT as an SKA Pathfinder. In: *Proceedings of Panoramic Radio Astronomy: Wide-field 1-2 GHz research on galaxy evolution*. Edited by G. Heald and P. Serra. Groningen, 2009
23. Johnston S, Bailes M, Bartel N, et al. Science with the Australian Square Kilometre Array Pathfinder. *Publications of the Astronomical Society of Australia*, 2007, 24(4): 174–188
24. Ford D, Bolton R C, Colegate T, et al. The SKA Cost/Performance Tool: A Hierarchical SKA Modelling Tool. In: *Proceedings of Wide Field Astronomy & Technology for the Square Kilometre Array*. Chateau de Limelette, 2009
25. Cotton W D. Special problems in imaging. In: *Synthesis Imaging*, ed. R. A. Perley, F. R. Schwab, & A. H. Bridle, 1986
26. Cotton W D. Special Problems in Imaging. In *Astronomical Society of the Pacific Conference Series*, Vol. 180, *Synthesis Imaging in Radio Astronomy II*, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley, 1999. 357–370
27. Deng Q W, Wang F, Deng H, et al. Performance evaluation of baseline-dependent averaging based on full-scale SKA1-LOW simulation. *Research in Astronomy and Astrophysics*, 2022, 22(4): 045014
28. Natarajan S, Barbosa D, Barraca J P, et al. SKA Telescope Manager (TM): status and architecture overview. In: *Proceedings of the SPIE*, 2016, 9913: 991302
29. Torchinsky S A, van Ardenne A, van den Brink-Havinga, et al. SKA Data Flow and Processing. In: *Proceedings of Wide Field Astronomy & Technology for the Square Kilometre Array*. Chateau de Limelette, 2009
30. Alexander P, Bregman J A, Faulkner A J. SKA Data Flow and Processing.

In: *Proceedings of Wide Field Astronomy & Technology for the Square Kilometre Array*. Chateau de Limelette, 2009
31. Spekkens K, Chiang C, Kothes R, et al. Final Report to LRP Panel: the Square Kilometre Array
32. Broekema P C, van Nieuwpoort R V, Bal H E. The square kilometre array science data processor. Preliminary compute platform design. *Journal of Instrumentation*, 2015, 10(7): C07004
33. Grainge K, Alachkar B, Amy S, et al. Square Kilometre Array: The radio telescope of the XXI century. *Astron Rep*, 2017, 61: 288–296

Note: Figure translations are in progress. See original paper for figures.