Building Mega-Science: A Systems Engineering Tool for the Square Kilometre Array

Timothy M. Colegate #[1]

 # ICRAR / CIRA, Curtin University of Technology
GPO Box U1987
Perth Western Australia 6845
Australia
[1] `tcolegate@ivec.org`

***Abstract*—The Square Kilometre Array (SKA) will be the largest radio telescope in the world, with an aperture of up to one million square metres, due to be operational by 2022 at a cost estimated at 1.5 billion euros (2007). Designing a flexible instrument such as the SKA is a long-term task and requires a systems approach with inputs from both engineering and science specialists. Cost and performance modelling, and subsequent optimisation, is central to building the radio telescope. Curtin is taking a lead role in this process, and we present here a custom-developed systems engineering tool that is being used in the design phase of the SKA. We outline how such a tool is being used to illuminate the performance and cost trade-offs required for this complex mega-science project reliant on emerging technologies to achieve its scientific goals. We also present some simple design decisions resulting from these trade-offs, saving hundreds of millions of euros.**

I. INTRODUCTION

The Square Kilometre Array (SKA) will be the largest radio telescope in the world, requiring a large investment of scientific and engineering work. A global project, with an estimated cost of 1.5 billion euros (2007), the SKA involves 19 countries and a number of competing technology concepts. With nominally one square kilometre of collecting area at mid-frequencies, the core of the SKA will be located at one of the candidate sites - Western Australia or South Africa - selected as being scientifically suitable to host the radio telescope. The collecting area will not be a single dish such as the Arecibo or Parkes radio telescopes, but will be an array of thousands of coherently connected antennas, with large signal processing and science computing facilities. The SKA is being built to answer some of the most important problems in astronomy by observing the universe at radio frequencies of about 70 MHz to 30 GHz [1].

To maximise the scientific capability of the SKA for a fixed construction and operations cost, the telescope must be treated as a system, where the technological decisions and scientific performance of the telescope are interconnected. In this broad design space, Curtin is providing insight into the risk and consequences of these decisions. This paper outlines why the project needs a systems engineering approach and describes a custom-developed systems engineering tool, used to illuminate areas of the SKA design space that present challenges and risks.

The author's contribution to the Costing Tool is multi-faceted: software development including requirements analysis, code development and testing; creating system and subsystem models, as described in section V; promulgating trade-offs and cost exploration, as shown in section VI; and making the system view of the SKA accessible to engineers and scientists, the necessity of which is described in section III.

II. THE SKA: A MEGA-SCIENCE PROJECT

The specifications and design of the Square Kilometre Array will largely determine the discovery potential and science output of the telescope. Due to be operational by 2022, the SKA is still in its design phase, with a number of important decisions yet to be made. As a multi-purpose instrument with a diverse set of specifications, the generality of the telescope creates particular design challenges. It has been observed that the scientific reasons for which previous telescopes have been built are rarely what the telescopes become known for [2]. To capitalise on the unexpected discoveries that will arise, instrument flexibility is essential. Since the discovery of astronomical radio emission in the 1930s, improvements in radio telescopes has increased the discovery space of radio astronomy. This parameter space is not yet fully explored – the SKA will be another step forward.

Mega-science projects differ from many other mega-projects in that the system definition process requires careful thinking to ensure that early engineering decisions do not compromise the scientific discovery potential of the instrument. For the SKA, this is further complicated in that the choice of technology concepts are still being refined, and that the concepts have differing levels of maturity and carry different levels of risk. In a bid to avoid cost-overruns commonly associated with mega-science projects and to mitigate risk, a process of technology decision making and readiness evaluation is in place [1].

The SKA design has been refined over a number of years, through technology concept white papers and documents such as the Reference Design for the SKA and the Preliminary Specifications for the SKA (see www.skatelescope.org). From this, three receptor (antenna) technologies are being considered: parabolic reflectors with wide bandwidth, single-pixel feeds; parabolic reflectors with phased array feeds; and aperture arrays. The latter two use electrical beamforming to increase the field of view available to the telescope. Further design refinement is occurring in the current Preparatory Phase of the SKA (PrepSKA), a collaborative international system design effort – of which Curtin is a large contributor – due to end in 2012 [1]. Following this, detailed design work and construction will be undertaken.

Designing an instrument with demonstrated but not necessarily scientifically ready technology concepts contains implicit risk. To mitigate this, it is essential to understand what is achievable with current and future technologies, as these design decisions will affect the cost and capability of the SKA. The digital signal processing and computing power required to combine, process and analyse the cosmic signals needs new technologies and techniques, and a systems engineering approach is essential for structured, iterative science-engineering interactions.

III. SYSTEMS ENGINEERING FOR THE SKA

As in most other engineering projects, there is a strong link between SKA performance and cost. Modelling of the problem can provide many insights into the trade-offs to be made in the design space and gives the opportunity for optimisation of the problem.

Due to the complexity of the system, the simple linear approach of requirements gathering, design and implementation is insufficient to fully scope the problem. Ultimately, the SKA design will be governed by its cost. To optimise the design within the cost limitations requires high-level thinking where the science outcomes drive the engineering requirements, but the thinking is also informed by detailed performance and cost analysis. The Costing Tool (section V) takes this approach.

There are many high-level cost drivers that feed into the optimisation problem, some of which are outlined as follows:

- flexibility to cater for a wide range of science goals
- complex antenna placement and resulting image quality effects
- technology concept choices
- technology maturity and risk
- the operations versus capital cost of the instrument
- site-specific costs affected by geographical diversity

To enable a systems focused performance and cost exploration for the SKA, we have created a top-level description of the instrument based on the Reference Design, set out as a series of concept independent sub-systems. As shown in Fig. 1, these sub-systems follow the elemental signal path: the signal arrives at the antennas, is received and processed in real time so the incoming signal is appropriately observed by the telescope. Post-processing with high-performance computing (HPC) is then applied to the observation. Throughout the signal path, information is transported in either digital or analogue form; the problem essentially becomes one of data transport and processing, which is a more tractable problem to those outside of radio astronomy [3].

A systems engineering approach is necessary due to the interconnection of the SKA sub-systems, which are often geographically separate but which cannot be developed in isolation – for example, signal coherence must be maintained between sub-systems. To evaluate the system and explore the design space, the scientific requirements must be simplified to a set of key system performance indicators to drive the cost modelling. A component of PrepSKA's system design effort is to distil out these metrics through a set of case studies of science projects (Design Reference Mission) to define the envelope of technical specifications [1].

IV. A PROCESS FOR DESIGN AND COSTING

To provide guidance to the SKA design process, a performance and cost estimation tool has been developed to explore the trade-offs and boundaries that exist in the SKA design space. Cost and performance modelling is acknowledged as a major contributor to the SKA (preliminary) specification process as reported in [4]; such modelling is also central to the engineering design of the telescope. At a high-level, the specification document discriminates between various SKA technology implementations, whilst highlighting important science implications and trade-offs.

Given the complexity of the SKA, it is important to identify where the dominant costs could lie and work hard to reduce these. The tool uses the best information available at any given time for the cost estimates – this allows for the refinement of costs as technology develops and uncertainty reduces. The costs come from a variety of sources, such as new and existing instruments and knowledge, research and development programs, industry trends and industry quotes. The cost models are developed with bottom-up estimates where the detailed designs are available, and a top-down, parameterised approach for the costs that are better estimated from existing trends. As the designs of possible sub-systems for the SKA develop, this information can be used to update the cost models.

The process for design and cost estimation must allow for exploration and optimisation. For this reason, the cost models must be scalable. A series of science driven metrics can define performance requirements to drive the tool. When making design decisions for various technology concepts, technology and instrument specification trade-offs must be investigated. The Costing Tool provides the scope to illuminate areas where cost overruns could be large. Beyond the instrument itself, there are other project costs such as software, infrastructure, power, verification, commissioning and maintenance that need to be considered. The operating versus capital costs for different implementations must also be explored. The emerging power problem influences the operating costs, and we are beginning to investigate these trade-offs, with consideration to the optimum use of resources.

V. THE SKA COSTING AND DESIGN TOOL

The complex estimation process required us to custom-develop a costing and design tool which treats the SKA as system. Complementary approaches to performance and cost analysis are presented in SKA Memos 92 [5] and 111 [6]. The main software tool is a collaborative international effort, which

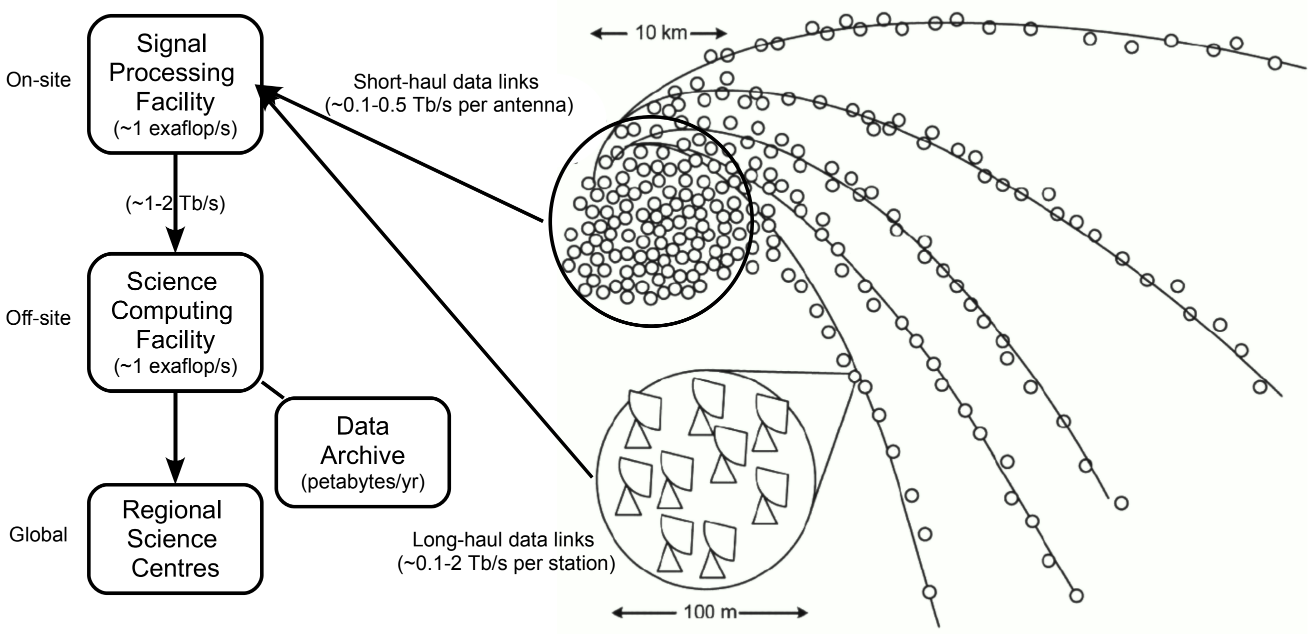


Fig. 1. A conceptual representation of the SKA array configuration and the data flow through the system. The circles represent clusters of antennas, the boxes indicate the digital signal processing facilities and the arrows are the data links between these sub-systems. The numbers are representative to give an indication of the scale of the project (adapted from [1]). One exaflop is 10^{18} floating point operations.

incorporates major aspects of all approaches to cost estimation and embodies the following philosophies:

- Flexible, extensible tool capable of trade-off exploration and optimisation
- Expandable architecture and scalable designs
- Presentation of a common (sub)-system view
- Signal path and geographical array zone analysis
- Uncertainty and contingency analysis
- Engagement with the wider community, and input from domain experts

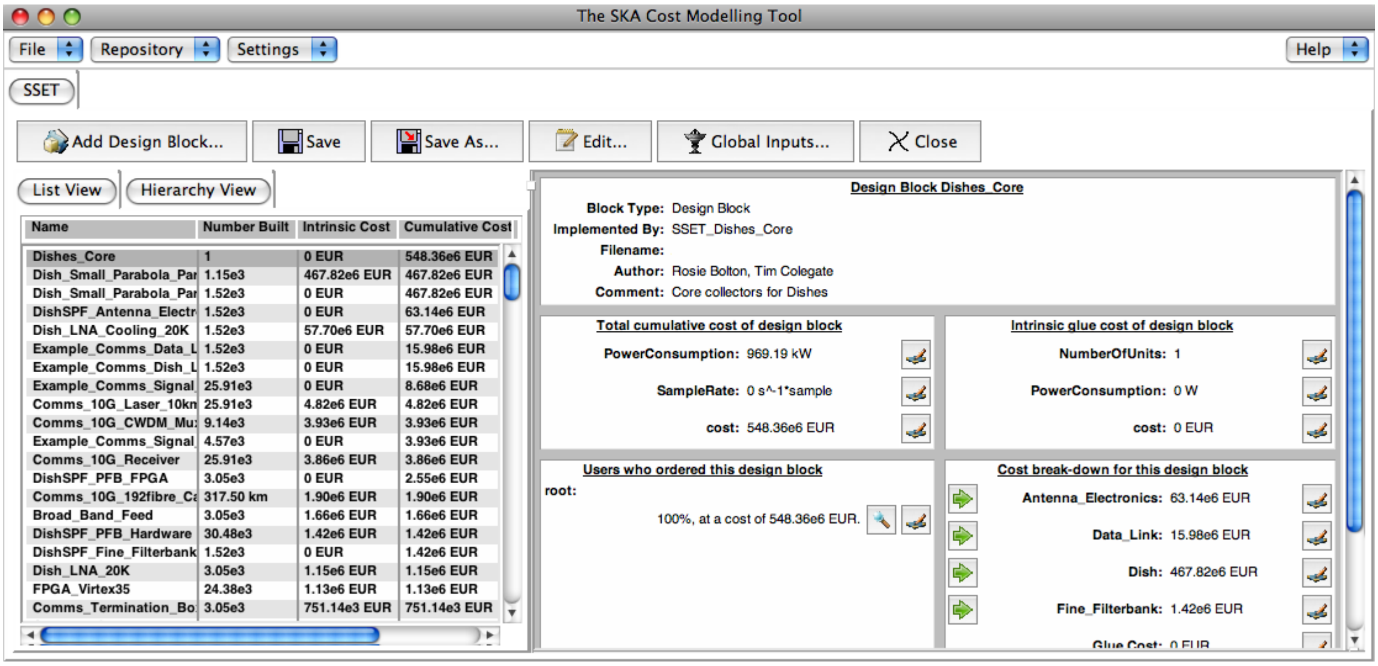
The tool needs to be extensible so new cost drivers can be incorporated. Although monetary costing is the primary goal, the tracking of power consumption and data transmission rates are two such examples of where the tool is being extended.

The Costing Tool is an interactive calculation engine with a mixed source of inputs, written in Python to allow quick development and prototyping. The tool is structured with sharp divisions. It contains a costing engine which does the mathematical cost and performance calculations. This engine acts upon models of the SKA telescope (telescope designs) to calculate their cost and performance characteristics. The design is held as a series of hierarchically interconnected sub-systems, containing component cost information, performance information and cost scaling models [7]. The engine is accessed via a user interface, providing accessibility to the telescope designs and allowing for cost trade-off exploration, examples of which are in section VI.

A screenshot of the tool is shown in Fig. 2. The components and sub-systems (also called design blocks) are listed on the left-hand side. These, in combination with some top-level global inputs and an index file, contain the all the information to model the system. The components and design blocks are stored as a simple database of XML files with snippets of Python code in the design blocks for the performance calculations. The buttons on the right-hand side give further information on the sub-system and lets the user drill down the system hierarchy.

There are subtleties in the cost estimation process that become important for a large project such as the SKA, with a long development time and a reliance on technological development. It is overly simplistic to take quotes for costs and to use these to develop a cost estimation. The Costing Tool implements a number of strategies to allow for a more complex cost estimation. Costs are broken down into non-recurring engineering (NRE) and manufacturing costs to reflect the research and development costs of some of the technologies.

The SKA is relying on the exponential improvement of digital technologies for many of the performance gains, especially in the areas of data transport, signal processing and computing. This is reflected in the tool, where the costs are time-dependent, following a scaling such as inflation, Moore's law or a custom designed scaling. The cost of the sub-systems are extrapolated using these scaling models, and will vary depending on the date they are purchased and built.



Name	Number Built	Intrinsic Cost	Cumulative Cost
Dishes_Core	1	0 EUR	548.36e6 EUR
Dish_Small_Parabola_Par	1.15e3	467.82e6 EUR	467.82e6 EUR
Dish_Small_Parabola_Par	1.52e3	0 EUR	467.82e6 EUR
DishSPF_Antenna_Electr	1.52e3	0 EUR	63.14e6 EUR
Dish_LNA_Cooling_20K	1.52e3	57.70e6 EUR	57.70e6 EUR
Example_Comms_Data_L	1.52e3	0 EUR	15.98e6 EUR
Example_Comms_Dish_L	1.52e3	0 EUR	15.98e6 EUR
Example_Comms_Signal_	25.91e3	0 EUR	8.68e6 EUR
Comms_10G_Laser_10kn	25.91e3	4.82e6 EUR	4.82e6 EUR
Comms_10G_CWDM_Mu	9.14e3	3.93e6 EUR	3.93e6 EUR
Example_Comms_Signal_	4.57e3	0 EUR	3.93e6 EUR
Comms_10G_Receiver	25.91e3	3.86e6 EUR	3.86e6 EUR
DishSPF_PFB_FPGA	3.05e3	0 EUR	2.55e6 EUR
Comms_10G_192fibre_Ca	317.50 km	1.90e6 EUR	1.90e6 EUR
Broad_Band_Feed	3.05e3	1.66e6 EUR	1.66e6 EUR
DishSPF_PFB_Hardware	30.48e3	1.42e6 EUR	1.42e6 EUR
DishSPF_Fine_Filterbank	1.52e3	0 EUR	1.42e6 EUR
Dish_LNA_20K	3.05e3	1.15e6 EUR	1.15e6 EUR
FPGA_Virtex35	24.38e3	1.13e6 EUR	1.13e6 EUR
Comms_Termination_Bo	3.05e3	751.14e3 EUR	751.14e3 EUR

Design Block Dishes_Core

Block Type: Design Block
Implemented By: SSET_Dishes_Core
Filename:
Author: Rosie Bolton, Tim Colegate
Comment: Core collectors for Dishes

Total cumulative cost of design block

PowerConsumption: 969.19 kW
SampleRate: 0 s^-1*sample
cost: 548.36e6 EUR

Intrinsic glue cost of design block

NumberOfUnits: 1
PowerConsumption: 0 W
cost: 0 EUR

Users who ordered this design block

root:
100%, at a cost of 548.36e6 EUR.

Cost break-down for this design block

Antenna_Electronics: 63.14e6 EUR
Data_Link: 15.98e6 EUR
Dish: 467.82e6 EUR
Fine_Filterbank: 1.42e6 EUR
Glue Cost: 0 EUR

Fig. 2. A screenshot of the Costing Tool showing the core part of the SKA filled with parabolic dishes. The left half shows quantities and costs of sub-systems (design blocks) and their components, the right half contains information about the top-level sub-system (SSET_Dishes_Core) and its children sub-systems.

VI. PRACTICAL APPLICATIONS OF THE TOOL

A strong driver for using a custom written tool is the exploration of trade-offs in the SKA design. The tool can be used by expert scientists and engineers to ask high-level questions, without the need for them to fully understand the complex underlying models. Such questions might be:

- How much cheaper will it be if we buy the computing n years later?
- If Moore's law fails us, what will be the cost?
- What components must we improve to reduce cost?

Here we present some examples for which the tool has been applied to inform the decision making.

A. Cost versus Dish Diameter

For the parabolic dish antenna technology concepts, the cost of construction for the dishes (except at small diameters) is the major cost of the SKA. At small diameters the electronics and signal processing costs become dominant and at large diameters the civil engineering challenges of building a large steerable dish cause a considerable increase in cost per unit area. Fig. 3, which complements those in [4], is an example of one of the many trade-offs to be made in the SKA design space. This plot is a breakdown of SKA costs for a telescope composed of parabolic dishes with wide-band single-pixel feeds. It shows the cost curve that results from the dish diameter versus cost trade-off. In this plot there is a shallow minimum of dish diameter between about 12 – 20 m.

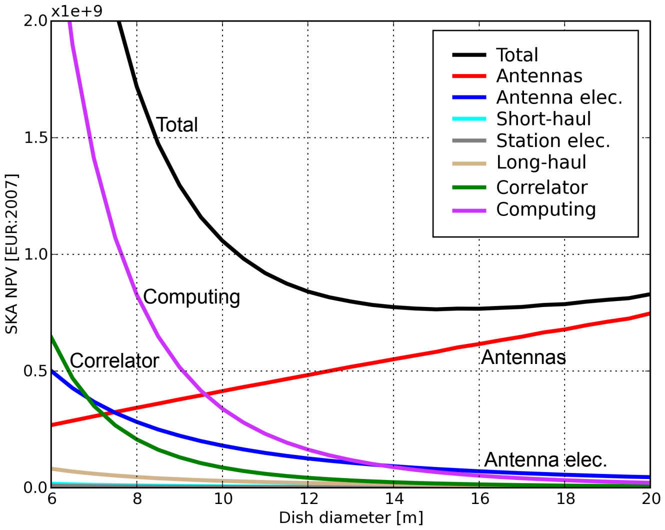


Fig. 3. Net Present Value cost versus Dish Diameter for the parabolic dish with single-pixel feed antenna technology concept. The effective area is held constant at $350\,000\,m^2$ with $T_{sys} = 35\,K$. The survey speed (a product of sensitivity and telescope field of view) is $2.73 \times 10^8\, deg^2.m^4/K^2$ at a dish diameter of 6 metres. This reduces to $2.47 \times 10^7\, deg^2.m^4/K^2$ for a 20 metre dish diameter, hence the smaller dish diameters are faster at surveying the sky, at a cost of increased digital processing. The correlator and computer purchase date is 2015.

In this example, the total SKA cost for 10 m dishes is about 1 billion euros. In comparison, the cost of 15 m dishes is about 800 million euros, a difference of 200 million euros. The non-linear scaling of the digital components (antenna electronics, short and long-haul data links, correlator and computing) is due to two factors. For dishes which are half the diameter, the collecting area is one-quarter and hence the number of required dishes is increased by a factor of four. Additionally, the sub-system costs scale as either n_{dish} or n_{dish}^2 . The square is because the signal from every dish has to be correlated with every other dish. Each of these correlated signals needs to be processed, hence the exponential increase of the correlator and computing costs as diameter reduces. This trade-off influenced the selection of the 15 m dish diameter in the preliminary specifications for the SKA [4].

B. Cost as a Function of Time

Digital technology advancements creates a time-dependent axis in the SKA design space. Fig. 4 is an accompanying plot to Fig. 3 and demonstrates how the timeline for purchasing the digital components (correlation and computing hardware) has a big impact on SKA design optimisations, including dish diameter. In the first plot, the correlator and computing purchase date is set to the year 2015, while in the second it is 2020. Here we have generalised the computing advancement using Moore's law with a cost halving period of 2 years, and the effect is to move the shallow minimum of dish diameter to about 10 – 15 m. Here, the cost difference between 10 and 15 m is negligible, however between 12 m and 20 m the 150 million euros difference is significant.

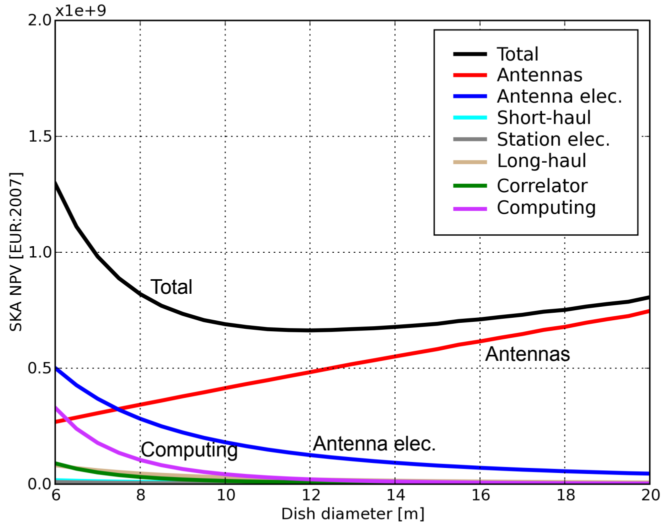


Fig. 4. Net Present Value cost versus Dish Diameter for the parabolic dish with single-pixel feed antenna technology concept, with the same specifications as Fig. 3 except that the correlator and computer purchase date is 2020.

C. Power Estimation

The extensibility of the Costing Tool allows us to easily incorporate information about power. The first step we are undertaking is to consider each sub-system to understand the peak and average continuous power loads created by the sub-system. Some sub-systems can be broken down into components and have their power tracked in the the same fashion as component costs by summing the components in the system. Other sub-systems in the more generalised computing domain need the power requirements to be parameterised as a unit cost such as watts per flop (floating point operation). With this information, we can assign an electricity unit cost and determine an operational cost for the instrument.

VII. CONCLUSION

In the SKA design process, the Costing Tool has been useful to provide insights into some of the problems that may arise, and highlight some important trends and trade-offs. The tool has played an important role in developing a systems view of the SKA, by allowing domain experts to understand the effects of engineering changes on the system as a whole, and thus the telescope's scientific performance. As the sub-system models are improved, this will continue, especially in the digital signal processing, computing and power domains, allowing the capital versus operational cost trade-offs to be considered in further detail.

The opportunity for savings of millions of euros indicates that the system-level performance and cost exploration capability of the tool will assist major technology decision-making. However, there will also be a gradual transition where the high-level assumptions and "black-box" sub-systems are replaced with more detailed designs. It is possible that these will be stored in some form of database-style structure, giving the flexibility for established project management and systems engineering tools to access this information, ensuring that the Costing Tool will provide many more insights in the continuing SKA design process.

ACKNOWLEDGEMENT

The author thanks his supervisor P. J. Hall for comments, and the SKA Program Development Office for its support. Colleagues at the CSIRO and the University of Cambridge have provided valuable advice, together with complementary software and modelling contributions.

REFERENCES

[1] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, 2009.

[2] P. Wilkinson, K. Kellermann, R. Ekers, J. Cordes, and T. J. W. Lazio, "The exploration of the unknown," *New Astronomy Reviews*, vol. 48, no. 11-12, pp. 1551–1563, 2004.

[3] P. J. Hall, "The Square Kilometre Array: An international engineering perspective," *Experimental Astronomy*, vol. 17, no. 1, pp. 5–16, 2004.

[4] R. T. Schilizzi, P. Alexander, J. M. Cordes, P. E. Dewdney, R. D. Ekers, A. J. Faulkner, B. M. Gaensler, P. J. Hall, J. L. Jonas, and K. I. Kellermann, *Preliminary Specifications for the Square Kilometre Array*, SKA Memo 100, 2007. [Online]. Available: http://www.skatelescope.org/PDF/memos/100_Memo_Schilizzi.pdf

[5] A. P. Chippendale, T. M. Colegate, and J. D. O'Sullivan, *SKAcost: a tool for SKA cost and performance estimation*, SKA Memo 92, 2007. [Online]. Available: http://www.skatelescope.org/PDF/memos/memo_92.pdf

[6] R. C. Bolton, A. Faulkner, P. Alexander, S. A. Torchinsky, A. van Ardenne, P. Wilkinson, M. de Vos, L. Bakker, S. Garrington, G. Harris, T. Ikin, M. Jones, D. Kant, D. Kettle, R. McCool, P. Patel, and J. Romein, *SKADS benchmark scenario design and costing 2 (The SKA Phase 2 AA scenario)*, SKA Memo 111, 2009. [Online]. Available: http://www.skatelescope.org/PDF/memos/111_Memo_Bolton.pdf

[7] D. Ford, R. C. Bolton, T. M. Colegate, P. Alexander, and P. J. Hall, *The SKA Costing and Design Tool*, SKADS Technical Memo 23, 2009. [Online]. Available: http://www.skads-eu.org/PDF/SKAsim.pdf