Obtaining long-term forecasts of the key drivers of greenhouse gas emissions in South Africa

Introduction

Debates about the potential effects of climate change, the necessity for action, and the relative merits of different response strategies all inevitably make reference to what is expected to occur if we “do nothing” – meaning, loosely speaking, under policies not too different from those currently in place. Framing domestic climate change policies and national positions in global climate negotiations requires the best possible information about possible future outcomes. Defining this position is made by trading-off emission reductions and economic growth. Policy makers’ efforts in this regard are focused on achieving low carbon development at the lowest possible cost to the economy. This cost of mitigation is calculated as a difference in costs (defined in monetary units) between a baseline situation and a new one characterized by lower emissions (Hourcade 1996). For most developed countries the emission reduction is calculated relative to a benchmark date in the recent past. Usually in studies of developing countries, emission reduction is calculated in terms of a percentage reduction from an emission level in a baseline trajectory at a specified future date, often several decades in the futures. The definition of a baseline trajectory is problematic:

- Uncertainty: the baseline trajectory is driven by many uncertain factors
- Definition: which existing policies and effort are to be included in the baseline?
- Mitigation and Development: for developing countries issues of poverty, inequality, and education goals need to be traded-off against mitigation goals. Are development goals and aspirations met in the baseline?

The goal of the current project is to derive baseline forecasts of carbon dioxide and other GHG emissions for South Africa, from the present day to 2050. We use a mixed methodology that is innovative in some of its elements. Firstly, we forecast GHG emissions indirectly, by first obtaining forecasts of a number of key drivers of energy demand – population growth, economic growth, and various commodity prices. These forecasts are obtained using a combination of expert elicitation and a review of available local and international literature. All inputs require some further processing in order to take a number of independent sources – potentially measured over different timescales and with different frequencies – and obtain a single suitably fine-scaled forecast (generally an annual time series for each key driver). All forecasts are probabilistic in nature – that is, they include assessments of statistical uncertainty around the modal or most-likely trajectory.

The forecasts obtained from this process are used as inputs to a South African implementation of TIMES (the SATIM model, see URL), a partial equilibrium linear optimisation model that selects a mix of energy sources and technologies that meets the forecasted demand for useful energy at least cost. GHG emissions, as well as other relevant outcomes, are obtained as a a result of the optimisation model. Monte Carlo simulation used to generate 1000 possible trajectories from the probabilistic projections of each key driver of GHG emissions. These are assembled into input matrices, each of which combines a single set of projections for each of the key drivers, taking into account correlations between inputs (for example, population and economic growth are positively associated). Finally, the SATIM model operates deterministically on each of these input matrices, turning each one into an annual forecast of GHG emissions (and other relevant outcomes). By examining the set of all simulated GHG emission trajectories, we arrive at a probabilistic forecast of GHG emissions for South Africa for the period 2014 to 2050.

Long-term forecasting is a controversial topic. It is usually a highly complex task subject to enormous uncertainties. Its failings have been well documented. Many long-term forecasts turn out not just to be wrong, which is to be expected, but to be so wrong that the values that eventually occur lie outside of even the most extreme confidence intervals. This has led some to abandon attempts to quantify long-term forecasts and instead base strategic planning on robustness to a small number of qualitative scenarios. Nevertheless, in determining responses to climate change in the public sphere a quantitative forecasting approach remains popular, and in many instances it is almost impossible to avoid reference either explicitly or implicitly to baseline forecasts. The current project is an attempt to remain within the quantitative forecasting paradigm, but our outputs should be interpreted with the caution that must accompany all long-term forecasts.

Methodology

Definition of baseline scenario

Designating a set of conditions constituting an emissions baseline inevitably involves a degree of subjectivity. The lack of a definitive code for establishing national "baseline" conditions has been previously identified (Clapp & Prag, 2012). The same authors propose a set of guidelines for setting baselines, covering the following elements: start year and projection period, scope of emissions sources, assumptions related to key drivers of projections, treatment of domestic policy measures, modelling framework or methodology, uncertainty and sensitivity analysis, consultation and review, and updating procedures. In the interests of clarity and transparency we address each of these points below.

Start year and timeframe for emissions projections

Our baseline begins in 2014; projections are made to 2050.

Scope of emissions sources covered

At the heart of our baseline projections is SATIM, a sectoral energy model based on TIMES, a partial equilibrium linear optimisation model developed by ETSAP, one of the International Energy Agency's implementing agencies, and a successor to MARKAL. The SATIM model uses five demand sectors and two supply sectors – industry, agriculture, residential commercial and transport on the demand side, and electricity and liquid fuels on the supply side. Sectors are divided in turn into subsectors. The industry sector, for example, is divided into mining, iron and steel, chemicals, precious and non-ferrous metals, NMM products, food, beverage and tobacco, pulp and paper, and "other" subsectors. The level of detail for a sector depends on the relative contribution of the sector to total consumption and also on how much funding has been historically received for developing that sector in the model, but in general can be considered fairly comprehensive. Full details can be found in the report by the ERC Systems Analysis and Planning group (2013), available at http://www.erc.uct.ac.za/Research/Otherdocs/Satim/SATIM%20Methodology-v2.1.pdf. AFOLU and LULUCF emissions are currently not included in SATIM.
GHG emissions that are included in SATIM are CO_2, CH_4 (including fugitive emissions) and N_2O.

Assumptions related to key drivers for emissions projections

Based on knowledge of the underlying SATIM model, the following key drivers of GHG emissions were selected. These are shown in

Key driver	Units	How assessed
Population	People	Literature
GDP growth	%/year	Expert elicitation
GDP composition	% Tertiary	Expert elicitation
Global coal prices	2012 R/t	Literature
Global gas prices	2012 $/Mbtu	Literature
Global oil prices	2012 S/barrel	Literature
SA Coal prices	2012 R/t	Expert elicitation & further modelling
SA Gas prices	2012 $/Mbtu	Expert elicitation
Nuclear Costs, Lead Times and Availability	2012 $ OCC, years and %	Literature
PV costs	2012 $/W	Literature & further modelling
CSP costs and Capacity Credit	2012 $/W, %	Literature & further modelling
Hydro Imports	GW	Literature

Table 1.

Key driver	Units	How assessed
Population	People	Literature
GDP growth	%/year	Expert elicitation
GDP composition	% Tertiary	Expert elicitation
Global coal prices	2012 R/t	Literature
Global gas prices	2012 $/Mbtu	Literature
Global oil prices	2012 S/barrel	Literature
SA Coal prices	2012 R/t	Expert elicitation & further modelling
SA Gas prices	2012 $/Mbtu	Expert elicitation
Nuclear Costs, Lead Times and Availability	2012 $ OCC, years and %	Literature
PV costs	2012 $/W	Literature & further modelling
CSP costs and Capacity Credit	2012 $/W, %	Literature & further modelling
Hydro Imports	GW	Literature

Table 1: Selected key drivers of GHG emissions

For each of these drivers, our goal is to obtain probabilistic forecasts at time intervals of one year – that is, not only annual point forecasts of mean or modal "expected values", but also an assessment of the statistical uncertainty around each of those point estimates, expressed as a probability distribution. These forecasts, together with the assumptions underlying these forecasts, are described in detail in later sections.

Treatment of domestic climate policy measures

We define our baseline with no climate policy measures for South Africa, without necessarily imposing business as usual globally. That is, we include the possibility that global steps are taken to mitigate climate change but that, for whatever reasons, South Africa remains in a "business as usual" scenario, South Africa does not implement its 'Copenhagen pledge, that is the 34% deviation below BAU by 2020 and 42% by 2025. This primarily manifests in international commodity prices, which influence local prices particularly in the case of coal.

Modelling framework and/or projection methodology used

The methodology used to obtain forecasts depends on the nature of the key driver. For international commodity prices and for technologies in which South Africa can be expected to pay global prices (i.e. nuclear, PV, and CSP), a number of detailed long-term forecasts are available in the literature. We essentially used these forecasts verbatim, without additional input from expert elicitations. We describe the literature on which these forecasts are based in a section below. In the same vein, we used existing UN probabilistic population forecasts, which arguably represent the state-of-the-art in population forecasting practice.

Forecasts for the other key drivers (i.e. GDP growth, share of GDP claimed by the tertiary sector, domestic coal prices, domestic gas prices) are based on expert elicitations. This is largely because reliable literature sources were unavailable or the local nature of the information tipped the balance in favour of expert knowledge. Detailed semi-structured interviews were used to elicit qualitative information on possible future outcomes, followed by a quantitative assessment of ranges of possible values. We followed generally accepted best practice when assessing this information, using the protocol outlined below.

In order to keep the elicitation task manageable for experts, we assessed three-point probability distributions (minimum, mode, and maximum) at three distinct points-in-time (2020, 2035, 2050). Direct elicitation of annual probabilistic forecasts i.e. annual probability distributions, was not considered practically feasible and in any case would be subject to overwhelming anchoring biases. Even with this highly restricted elicitation goal, interviews took between 2 and 4 hours to complete, even after some preparatory work by experts before the interviews.

Information gathered using either literature searches or expert elicitation was rarely in a form that could be directly used by the SATIM. Some "post-processing" was invariably required. Operations included interpolation between the three key time-points in the case of elicited quantities, currency standardization, temporal discounting, and aggregation over sources. These too are described in detail in the sections that follow.

Uncertainty and sensitivity analysis

As is clear from the above, uncertainty is a fundamental component of our approach. All of our projections are probabilistic by nature: uncertainty in model inputs is explicitly captured and this is propagated into uncertainty in model outputs via Monte Carlo simulation.

Consultation and/or review

The current document, by proposing an approach and deriving baseline estimates from this approach, forms a key part of this review process. Several of our model inputs are derived through interviews with experts in an area – GDP growth and sectoral composition, coal prices, and gas prices. We follow standard best practice but have interviewed only a small number of experts in each field. Model inputs are freely available for review. The SATIM model is also well documented and has been used in a number of previous applications; it too is open to review and scrutiny.

Updating the baseline

At the present time no plans exist to update these particular baselines. Nevertheless, the methodology described here is flexible in this regard. Updates could be obtained with relatively little effort. Our external data sources are all well-established and thus can be expected to be available into the future.

Elicitation protocol

Our protocol is largely a summary of three commonly used protocols: the Stanford/SRI protocol, Morgan and Henrion's protocol, and the Wallsten/EPA protocol (Morgan & Henrion, 1990). The protocol is divided into five stages. Before the interview procedure, experts are sent a set of documents containing relevant background information. The interview procedure itself comprises three stages: establishing rapport with the expert; eliciting the expert's qualitative view of the problem, including factors influencing the outcome of interest; and eliciting the expert's probabilistic forecasts. After the interview, the elicited judgements and statements derived from them are sent to the expert to verify that they are both consistent and accurately reflect the expert's beliefs. Although these stages are executed in a sequential order, it should be stressed that some iteration between the stages can be expected.

Stage 1: Pre-elicitation

Prior to the interview, experts were asked to read three short documents: one summarizing the TIMES/MARKAL model (2 pages), one summarizing the available literature and points of view, for the quantity to be assessed (1-3 pages), and one summarizing the literature on heuristics and biases in probability assessment (5 pages). These documents are provided in the appendix.

Stage 2: Establishing rapport

The goal of this stage is simply to introduce the elicitation team to the expert and provide an overview of the reason for the elicitation and the underlying problem at hand.

Following a brief introduction of the team, we explained the TIMES/MARKAL model and the projections to be made, emphasising the uncertainty that exists around each of the key drivers to the model, and hence introduced the need for probabilistic forecasts. We then briefly noted the difficulties of long-term forecasts, emphasising that there is no "correct" answer to any of the elicitation questions, and that our main aim is to obtain judgments that reflect the expert's expressed beliefs, in particular the extent of their uncertainty, which may be large. Finally, experts were informed that if at any stage they felt truly unable or uncomfortable making numerical probability judgements, other qualitative elicitation techniques were available, although none ultimately made use of this option. Using the pre-interview document as a basis for discussion, experts were familiarised with the dangers of subjective probability assessment, particularly overconfidence, anchoring, and availability.

Stage 3: Qualitative elicitation of factors influencing key drivers

In this stage, experts were asked two main questions:

1. What factors, broadly speaking, influence the key driver on which their expertise was being sought?
2. How might these factors combine in the future, in particular to cause especially low or high values in the key driver?

The goal of the acclimatization stage is to get the expert to think critically about the problem at hand, and to identify, in a qualitative way, the important factors that should influence their later quantitative judgments, together with some assessment of what kinds of changes are possible.

The two questions above typically produced two distinct outcomes: firstly, a detailed qualitative description of the system relating to the key driver (the basic macro-economic model in the case of GDP growth, or the coal production and processing systems in the case of the coal price, for example); secondly, a form of "best-case" and "worst-case" scenario consisting of a qualitative storyline that might result in a particularly high or low value for the key driver. These views, which may well be at least partially constructed or modified as part of the process, form the justification for the later quantitative elicitation and as an audit trail for interested parties in the future.

Most texts emphasise the need to avoid a status quo bias by encouraging broad thinking and the consideration of alternate viewpoints. We repeatedly prompted experts to consider how outcomes other than the ones that they had already specified might arise, using questions such as "what might cause a sustained GDP growth rate of less than 1%?" if that had not yet been offered as a potential outcome.

Stage 4: Quantitative elicitation

This stage contains the formal assessment of probabilistic information. To keep the task manageable, we did not attempt to assess detailed probability distributions but rather assessed three-point (minimum, mode, maximum) distributions, modifying these where additional information (on intermediate quantiles, for example) was offered.

In order to avoid anchoring on central values, we began by asking the experts for extreme lower or upper values, although some experts insisted on starting with central values, and these requests were accommodated. All experts were more comfortable providing information first for 2015, then for 2030, and finally for 2050. Although this might lead later estimates to be biased towards 2015 values – which would, under normal conditions, be associated with overly narrow confidence intervals – again we felt that it would be counter-productive to force any other order.

In all cases, we attempted to combat overconfidence and overly narrow confidence intervals by asking the expert to think of scenarios that would result in values more extreme than the extreme values just given. Once ranges of extreme values had been given, we asked for a modal (most likely) value.

Stage 5: Post-elicitation verification

The aim of this stage is to present the expert with his or her elicited qualitative and quantitative judgements, to check that this reflects their views accurately, and to revise judgments as necessary. Ideally, this step would be conducted as part of the interview procedure but, because our interviews were already lengthy, we elected to send feedback to experts by email after the interview had been concluded. Feedback included a summary of their qualitative descriptions of the system and major influences of the key drivers and plots of the triangular probability density function obtained from their quantitative assessments.

Post-processing of elicited and external data sources

The results of our elicitation process or literature survey are probability distributions defined over values of the key drivers at specific time points, or mean values of the drivers over specified time periods. In order to convert these into simulated annual time series, we apply a small number of post-processing operations.

In the case of probability distributions defined directly over values of the key drivers (i.e. not means) we:

1. Use Algorithm 1 or Algorithm 2 to randomly generate values of the key driver at 2014, 2020, 2035, and 2050 by drawing from the appropriate distributions, taking into account any correlational information between periods.
2. Use Algorithm 3 to generate annual values consistent with the values generated in the previous step.

In the case of probability distributions defined over mean values of the key drivers we:

1. Use Algorithm 1 or Algorithm 2 to randomly generate mean values of the key drivers at 2014, 2020, 2035, and 2050 by drawing from the appropriate distributions, taking into account any correlational information between periods.
2. Use Algorithm 4 to generate annual values consistent with the means generated in the previous step.

Algorithm 1: Generating correlated random samples from arbitrary distributions

Suppose we wish to simulate N values from each of a set of P arbitrary distributions $F_1, F_2, \dots, F_P$, with the resulting simulated values having a correlation matrix $\mathbf{\Sigma}$. An algorithm for doing this to good approximation is:

1. Simulate N samples from a multivariate standard normal distribution with correlation matrix $\mathbf{\Sigma}$.
2. Convert the values generated in the previous step into probabilities by applying the univariate standard normal CDF to each of the P samples i.e. independently.
3. Simulate draws from the desired distributions by applying the inverse CDFs $F_1^{-1}, F_2^{-1}, \dots, F_P^{-1}$ to the probabilities generated in the previous step.

The simulated values preserve the desired correlation structure only approximately, because the transformation in step 2 preserves the rank order rather than the exact correlations. Nevertheless in general testing we found the approximation to be good, and given the inherent uncertainty in specifying the correlation in the first place, any errors introduced by the approximation are likely to be negligible.

Algorithm 2: Generating mean-reverting random samples from arbitrary distributions

For one of our series, experts indicated that sustained extremely high or low values over multiple time periods were implausible. Thus, extremely high or low values should be followed by periods of average (or at least, much less extreme high or low) values. This poses a problem for Algorithm 1 because the desired relationship is not easily expressed using correlation.

Suppose we wish to simulate N values from each of a set of P arbitrary distributions $F_1, F_2, \dots, F_P$, with the resulting simulated values having the "mean-reversion" property described above. An algorithm for doing this to good approximation is:

1. Simulate N samples from a multivariate standard normal distribution with correlation matrix $\mathbf{\Sigma}$.
2. Compute the ranks (in descending order) of the values in each of the samples generated in the previous step.
3. Simulate draws from the desired distributions by applying the CDFs $F_1^{-1}, F_2^{-1}, \dots, F_P^{-1}$ to uniformly-distributed probabilities.
4. Reorder each of the samples generated in the previous step so that their rank orders match those generated in step 2.

Note that in contrast to Algorithm 1, Step 3 returns uncorrelated values from the desired distributions, with the correlation structure being re-introduced (approximately) in Step 4. Again, the simulated values preserve the desired correlation structure only approximately, because the transformations in step 2 preserves the rank order rather than the exact correlations, and this is what is imposed in step 4. Again, any errors introduced by the approximation are likely to be negligible compared to the degree of imprecision inherent in the elicitation process.

Algorithm 3: Interpolating annual values between elicited values at specified time-points

Linear interpolation between the values generated by Algorithm 1 or 2 suffers from two drawbacks: it almost certainly underrepresents the year-to-year variability in the underlying time series, and the interpolated values between two time-points are constrained to lie between the values at those time-points. As a result, we use the following random walk algorithm that starts at $x_{p,1}$ and is guaranteed to end at $x_{p+1,1}$ after a fixed number of time steps τ (in our case, the number of years from time-points $p + 1$ to p):

1. Define a drift value $d_p = (x_{p+1,1} - x_p)/\tau$. This gives the average annual change needed to get from $x_{p,1}$ to $x_{p+1,1}$ in τ steps (note that this is the linear interpolator).
2. Select a step size Δ_p to measures a "typical" random change in annual values around the drift value d_p
3. For $t = 1$ to $\tau - 2$
   a. Generate three possible "moves"
      i. $x^*_{p,t+1} = x_{p,t} + d_p - \Delta_p$
      ii. $x^*_{p,t+1} = x_{p,t} + d_p$
      iii. $x^*_{p,t+1} = x_{p,t} + d_p + \Delta_p$
   b. For each possible move, calculate the distance between the terminal point $x_{p+1,1}$ and the sum of the proposed value $x^*_{p,t+1}$ and the drift that is still to be added in the remaining $(\tau - t)$ time periods

$$e^*_t = \left|x_{p+1,1} - \left(x^*_{p,t+1} + (\tau - t)d_p\right)\right|$$

      This quantity indicates the sum of remaining random steps needed to reach the target point $x_{p+1,1}$.
   c. Calculate the maximum change due to random steps that is possible in the $(\tau - t)$ time-steps that remain i.e. $L_t = (\tau - t)\Delta_p$
   d. Select one of the proposed moves at random, where the selection probabilities are given by $\theta^*_t = 1 - \max(0, (e^*_t/L_t - \epsilon)^2)$, where ϵ is a small constant that prevents the selection probability going to zero where $e^*_t = L_t$ exactly (i.e. where the move is still strictly permissible). Thus, where a proposed move leads to a point that is further away from the target than the maximum remaining changes that may occur, $e^*_t > L_t$ and the resulting selection probability θ^*_t will be zero. Moves become relatively less likely to be chosen as they approach this limit.

Note that we do not explicitly work out the final step from $x_{p,\tau-1}$ to $x_{p,\tau} = x_{p+1,1}$. In general, it will not be possible to reach $x_{p,\tau}$ from $x_{p,\tau-1}$ using only drift and the random change i.e. $x_{p,\tau} - x_{p,\tau-1} - d_j \neq \Delta_j$, but the above steps are sufficient for the final "random" change required to be smaller than Δ_j, which is sufficient for the purposes of our study.

Algorithm 4: Interpolating annual values between elicited means over specified time-periods

We use a similar algorithm to the one described in Algorithm 3, except in the case when mean values over a period have been specified, we have no values at specified time points to anchor on, so that some modification is necessary. In particular, what we have been given is effectively an instruction that the simulated annual values between time-points $p+1$ and p should collectively have the desired mean $\bar{x}_p$. Our "target" value is thus not the value $x_{p+1,1}$ but this mean $\bar{x}_p$. In other respects the algorithm is very similar to Algorithm 3.

We use the following random walk algorithm that starts at $x_{p,1}$ and is guaranteed to sum to $S_p = \bar{x}_p\tau$ after a fixed number of time steps τ (in our case, the number of years between time-points $p+1$ and p):

1. Let $\propto_p$ be the initial value of the time series. For the first period $p = 1$, $\propto_1$ is set to the current (2014) value of the key driver. For subsequent periods, $\propto_p = x_{p-1,\tau}$ i.e. periods are defined so that last year in period $p-1$ is the same as the first year in period p. This introduces a small degree of overlap between periods but allows for easy specification of initial values in later periods where $p > 1$. Given the fairly long periods we use, the generally smooth nature of changes in the mean, and the inherent imprecision in the input information, the relative effect of this modelling choice is in all likelihood negligible.
2. Define a drift value $d_p = 2(S_p - \propto_p \tau)/(\tau(\tau-1))$. Applying this drift will return an arithmetic series of length τ that starts from an initial value of $\propto_p$ and sums to S_p. Note that this, as before, is the linear interpolator.
3. Select a step size Δ_p to measures a "typical" random change in annual values around the drift value d_p
4. For $t = 1$ to $\tau - 1$
    a. Generate three possible "moves"
        i. $x^*_{p,t+1} = x_{p,t} + d_p - \Delta_p$
        ii. $x^*_{p,t+1} = x_{p,t} + d_p$
        iii. $x^*_{p,t+1} = x_{p,t} + d_p + \Delta_p$
    b. For each possible move, calculate the distance between the terminal sum S_p and the sum of the series that would be obtained if the proposed value $x^*_{p,t+1}$ was accepted and extended to the final time step τ using only the drift to modify it. This latter sum is obtained by adding together (a) the sum of the series thus far $S_t = x_{p,1} + x_{p,2} + \cdots + x_{p,t}$, (b) the proposed value $x^*_{p,t+1}$, and (c) the remaining sum that would be the contributed should $x^*_{p,t+1}$ with no further random steps (i.e. using only drift), given by $R^*_t = (\tau - t)x^*_{p,t+1} + (\tau - t)(\tau - t - 1)d_j/2$, the latter term being the cumulative sum of the remaining drifts to be added from the baseline $x^*_{p,t+1}$ (which, as it would appear $\tau - t$ times, contributes the first term). Finally, we obtain

    $$e^*_t = |S_p - S_t - x^*_{p,t+1} - R^*_t|$$

    This quantity, as before, indicates the sum of remaining random steps needed to reach the target point S_p and thus can be used to assess the feasibility of the proposed move.
    c. Calculate the maximum change due to random steps that is possible in the $(\tau - t)$ time-steps that remain i.e. $L_t = (\tau - t)\Delta_p$

d. Select one of the proposed moves at random, where the selection probabilities are given by $\theta_t^* = 1 - \max(0, (e_t^*/L_t - \epsilon)^2)$.

5. Set the final value in the series $x_{p,\tau}$ to the remainder necessary to precisely obtain the desired sum/mean i.e. $x_{p,\tau} = S_p - S_{\tau-1}$. This step is needed because the difference above will not in general be equal to the step size Δ_p (though the previous steps ensure that it will be less than Δ_p).

The SATIM model

SATIM – the South African Times Model – is a national large-scale energy model created and hosted by the Energy Research Centre at the University of Cape Town. This model was originally developped for the Long Term Mitigation Scenarios (LTMS) project but is now in its third generation. This section provides a brief overview of the model. Full details are provided in the report by the ERC Systems Analysis and Planning group (2013), available at http://www.erc.uct.ac.za/ Research/Otherdocs/Satim/SATIM%20Methodology-v2.1.pdf.

The economy of a nation or region consumes energy from a number of primary and secondary sources. This energy delivers services by means of a myriad of technologies large and small. A model of the demand for energy needs to capture this complex structure and thus these sources and technologies need to be organised in some logical way. The SATIM energy model is an attempt at just such a model. It is a parameterisation of TIMES, for the South African energy system. TIMES is a partial equilibrium linear optimisation model developed by ETSAP, one of the International Energy Agency's implementing agencies, and a successor to MARKAL.

The SATIM model is a stylized representation of the whole energy system, with an optimization step that selects the mix of supply-side technologies that meets the demand for final energy at least cost. The model includes economic costs, emissions, and a range of sector-specific constraints that can be applied at a point in time or cumulatively. A user interface provides a framework for both structuring the model and scenarios, and also for interpreting results. The model has proven useful in assessing the complex interrelationships between potential mitigation policies.

The SATIM model is fundamentally "sectoral", in that it organises the demand for energy by economic sector, and characterises the demand for energy in a sector by the energy services required by that sector. The SATIM model has five demand sectors and two supply sectors – industry, agriculture, residential commercial and transport on the demand side, and electricity and liquid fuels on the supply side. The level of detail for a sector depends on the relative contribution of the sector to total consumption and also on how much funding has been historically received or how much knowledge was available for developing that sector in the model. Thus the model for the Transport sector is quite detailed but that of the Agricultural sector is quite simplistically represented in SATIM, because in South Africa the Agriculture sector accounts for relatively small energy consumption and low emissions.

In SATIM, services supplied to each of the five sectors are driven by technologies that require energy, with the quantity of energy required depending on the efficiency of the technology. Useful energy (the energy service) is an exogenous model input disaggregated by energy end-use, for each demand sector. Final energy demand is determined endogenously using the assumed efficiencies of the least cost demand-side technologies selected by the model. The two supply sectors and primary energy sources must meet the sum of these demands, with the model optimizing the mix of supply-side technologies to meet the demand for final energy at least cost.

The SATIM model includes a number of parameters and general assumptions for each sector broadly covering: (a) the structure of the sector and its energy services as it impacts on the demand for energy; (b) the establishment of base year demand for energy in the sector; (c) technical and cost parameters of the technologies available to satisfy the demand for energy services currently and in the future; (d) the projection of future demand for energy services.

Alternate Model sensitivity analysis

The model setup and underlying assumptions within the model itself can affect the results. In this section, we describe the alternate models used to undertake a model sensitivity analysis in order to gauge how the model behaves under differing setups.

The Myopic model

Background

Optimisation models are usually run as perfect foresight models, that is – all information contained in the model is known to the 'central planner' or rather the solver for the entirety of the model horizon. However, this is not the case for real world decision makers where future commodity and technology prices are not certain and decisions that are made are done so while taking into account variables known only at the time of making decisions (Keppo & Strubegger, 2009). These decisions made with limited foresight have lasting effects on the overall energy system, they affect the choice of decision makers in the future as they become 'locked-in' to the system.

Myopic modelling is the method of simulating limited foresight - where the energy model optimises over a period which is shorter than the total model horizon and solves these periods subsequently sequentially? until the model reaches the end horizon. The myopic method allows the analyst to study the impacts of unforeseeable price shocks on the system, and to study the effects of incomplete information (Keppo and Strubegger, 2009; Hedenus, Johansson and Lindergren, 2013).

Method

The myopic optimisation model run is carried out by using the time-stepped model variant in TIMES. The procedure is described by way of an example taken from the TIMES manuals: in Figure 1 the model horizon is for 80 years, but the model optimises over 20 years and once complete, the model goes back 10 years and optimises for the next 20 year period. In this example the model reviews the optimisation path every 10 years.

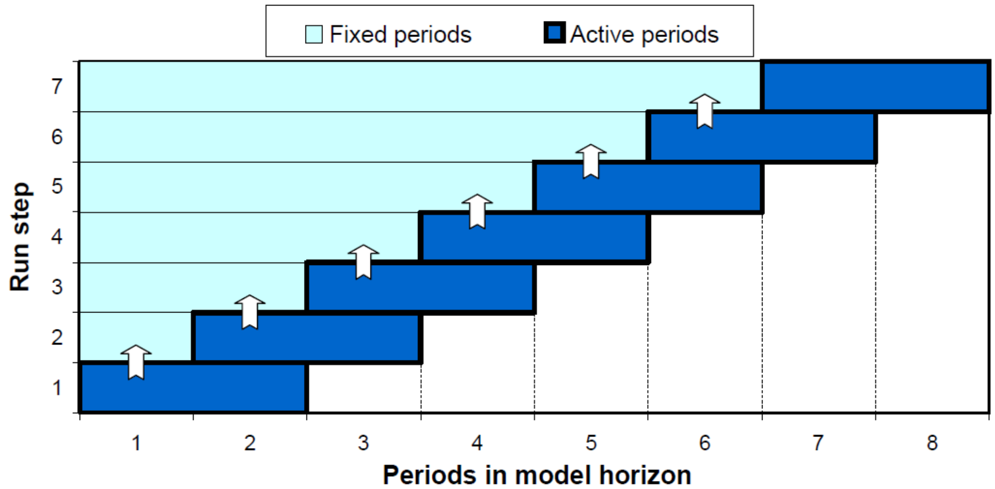

Figure 1: Overview of myopic process in TIMES. Figure taken from the TIMES user manual (Lehtila, 2011)

In this study, the model was set up to run as the Time-stepped model variant option available in ANSWER-TIMES with 10 years as the planning horizon and a 5 year overlap – essentially reviewing the planning every 5 years. The overall horizon to 2050 does not change. A review every 5 years one could argue would agree with real world practice - in South Africa the IRP2010 originally published in 2011 is in the final stages of being updated before being released ~~at the time of writing~~.

The Global Discount Rate

The global discount rate in the model? is set to 8% in line with the recent National Integrated Resource plan ~~efforts~~ (DOE 2011). The global discount rate however, is an important determinant as it affects how the technologies with high upfront capital costs (e.g. nuclear and renewables) compete with technologies with relatively low upfront costs but higher fuel costs over the life of the technology. For this reason Two alternate discount rates are used in the simulation:

- 5%, which would emulate a more "social" discount rate or where finance of energy projects is made relatively less costly.
- 11%, which would emulate the cost of finance faced by private investors.

Probabilistic projections of key drivers of GHG emissions

GDP growth

Overview

We conducted elicitation interviews with two experts on the subjects of GDP growth and the GDP sectoral distributions. Both experts preferred to think about GDP growth in terms of a mean growth rate (in %) over three intervals (2014-2020, 2020-2035, 2035-2050), rather than the annual growth rate in 2020, 2035, 2050. The elicited probability distributions thus covered possible values in the mean growth rate over these three periods.

Summary of qualitative discussions

Influential factors

Under the standard macroeconomic model, $GDPR = (TFP)K^{\alpha}L^{1-\alpha}$, where $GDPR$ is real GDP; K, L, E, and TFP are capital, labour, energy, and total factor productivity, and α is a parameters to be estimated. Factors influencing capital growth include investment rates (public and private), domestic and foreign savings. A useful back-of-envelope calculation is the incremental capital-output ratio, $ICOR = (I/GDPR)/\Delta GDPR$, where I is investment, so $I/GDPR$ is the investment rate. Currently, South Africa's investment rate is roughly 20%, giving an incremental capital-output ratio of roughly $ICOR = 0.2/0.035 = 5.7$. Thus, for a desired GDP growth rate of 6% for example, the investment rate would need to be (rearranging the subject of the formula above) in the region of $5.7 \times 0.06 = 34\%$ i.e. almost doubled. Factors influencing labour market growth include population growth, comprising domestic population growth and immigration. Factors influencing TFP include education and skills (quality of labour), technological innovation (quantity of labour), and good governance/strong institutions. Other factors affecting growth prospects included: the growth of other African economies, global economic growth, and SA's ability to export.

Trends and scenarios

In the period 2014-2020, there is no sign of a change in fundamentals that would shift SA out of its current GDP growth band of 2.0-3.5%. To shift the mean (over period 2014-2020) to 3.5% would require large improvements from the current base of around 2.4%, starting immediately – this is unlikely. There is not enough economic incentive (among business, labour, or government) to change the basic institutional agreements in place. Thus, TFP growth continues at 1-1.5% and there is a slow absorption of labour into the workforce. Although unlikely, it is plausible that the labour market could be substantially opened up by changes to government legislation. This would be most likely to happen at the skilled end of the job market. Capital growth rates are also unlikely to change dramatically, perhaps varying between 18% and 22%.

Over the medium term i.e. 2020-2035, large changes to the current GDP growth rate could occur. Over a 10-15 year period, mean investment rates of greater than 30% are certainly plausible, and have been achieved by, for example, China and Japan. Under a high growth scenario, current energy constraints would be removed and energy would potentially be sourced from, and shared with, neighbouring countries; the labour market would be expanded, particularly at the skilled end and predominantly (probably) from other African countries, by changes to legislation and/or incentives; and the world economy would have recovered from its current state and be in a "boom" scenario. Under a low growth scenario, the opposite would occur: current legislation prohibiting (or making very difficult) free flows of labour, energy constraints, low investment and low global economic growth. In setting limits to growth rates, it was noted that growth rates in excess of 6% have never been observed for long periods of time; while if growth rates drop much below 1.5% for any significant period of time political stability is seriously at risk.

In the long-term, the same general growth paths may occur as in the period 2020-2035. That is, there is a possibility of high average growth to be sustained over a decade, if this has not already occurred; or on the downside, for relatively low growth to be tolerated for a decade before civil action forces change. Both experts indicated that there should be some form of mean-reversion between the medium- and long-terms, such that, if there has been a period of either high or low growth in 2020-2035, the GDP growth rate returns to moderate levels; if there has been a period of moderate growth in 2020-2035, then GDP growth rate in 2035-2050 may move either up or down.

Quantitative forecasts

Figure 2 shows the probability distributions elicited from the two experts:

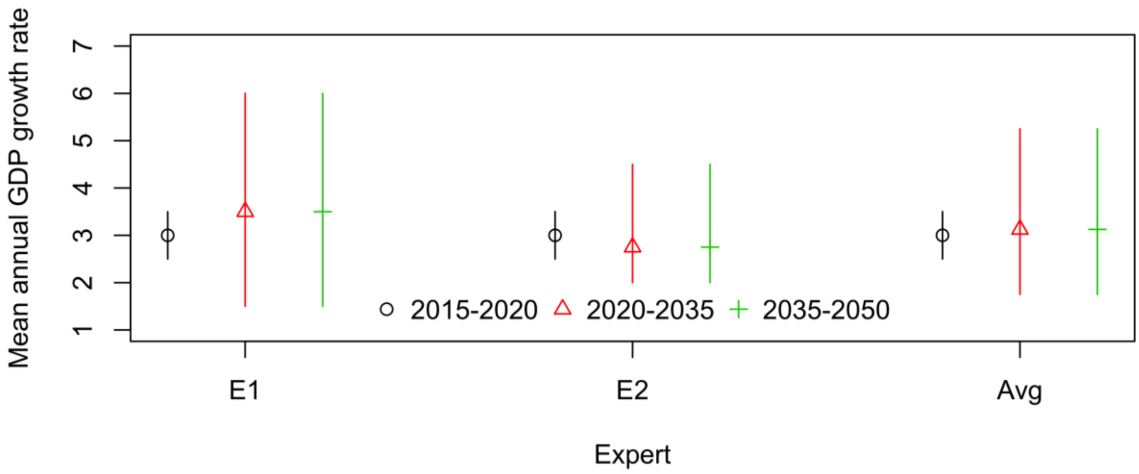

Figure 2: Elicited distributions for mean annual GDP growth rate (2 experts)

Post-processing

The experts' elicited distributions were combined by a simple averaging process, and converted into annual time series using Algorithm 2 and 4. When generating simulated mean values for the three time periods 2014-2020, 2020-2035, and 2035-2050, we impose a moderate negative autocorrelation between the extremeness of values obtained for 2020-2035 and 2035-2050, by setting $\Sigma_{23} = \Sigma_{32} = -0.4$. Sample trajectories generated by Monte Carlo simulation are shown in Figure 3. The green line shows the median trajectory, the red lines the 95% confidence interval and the blue lines the 80% confidence interval.

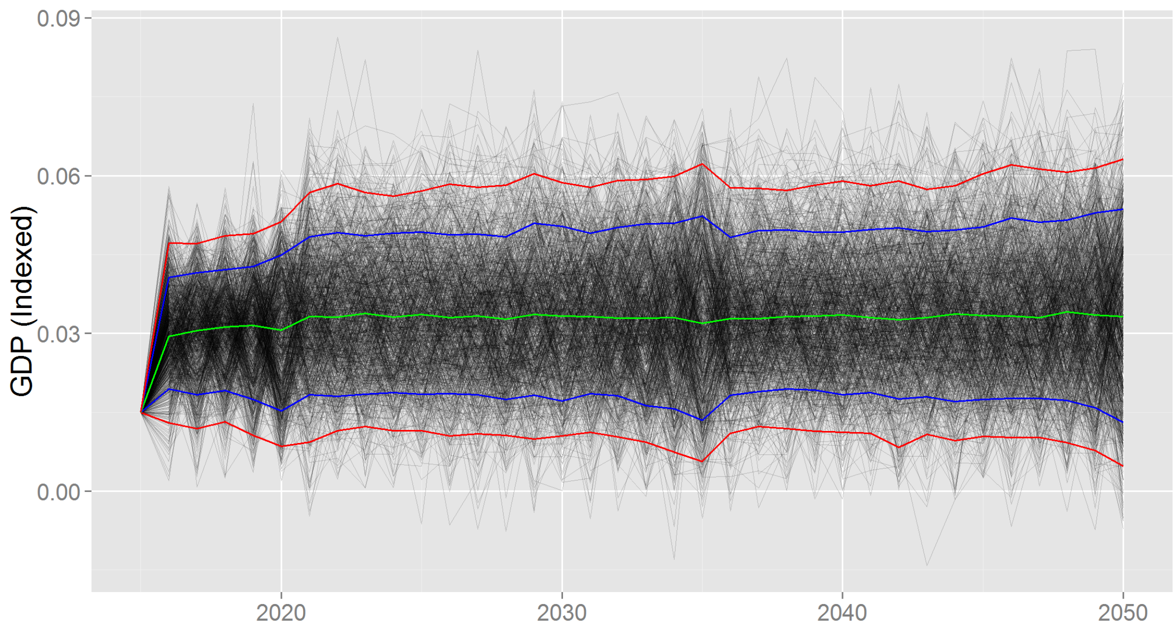

Figure 3: Probabilistic projections of annual GDP growth rate over the period 2014 – 2050.

Tertiary sector share of GDP

Overview

The SATIM model allows a distinction to be made between GDP growth rates in different sectors of the economy. For the purposes of our study, the primary distinction is between growth rates in the tertiary (service) sector and growth rates outside the tertiary sector (agriculture, mining, and manufacturing). We elicited probability distributions associated with the share of GDP provided by tertiary sector activities at the three key time-points 2020, 2035, 2050, from the same two experts from whom we elicited forecasts on GDP growth.

Summary of qualitative discussions

Influential factors

Traditional models of economic development propose that the main focus of a country's economy shifts from the primary sector, through the secondary sector, to the tertiary sector. This is effectively a result of mechanisation and development of industry, and then greater disposable incomes in a post-industrial society. It is sometimes called a "TPF[1]-led" path to growth, since TFP tends to increase sharply with the initial shift into manufacturing and industrial processes. However, there is now some debate about whether, perhaps as a result of increased globalisation, this traditional path to growth has been closed off. Recent growth

[1] TFP: Total Factor Productivity

success stories, particularly in resource-based economies like South Africa, have not seen an increase in manufacturing (e.g. Botswana).

South Africa is a resource economy with a relatively small manufacturing sector compared to countries with similar size economies. Drivers of change to its contribution include: labour strength (unions), currency strength, geographic positioning relative to major consumers (transport costs), and the existence of "value chains". Changes in the primary sector are largely functions of the resources (e.g. of discovered mineral deposits) and the degree of mechanisation in that sector. Increased tertiary sector contributions occur when incomes rise above the point required to satisfy basic material needs, and people can afford to spend more on, for example, education, health, and entertainment.

Trends and scenarios

The contribution of the agricultural sector is very unlikely to grow: agriculture is already highly mechanised, water resources and arable land are constrained. Were SADC to become more integrated, agricultural activities might become cheaper to perform outside South Africa. Primary sector contributions would fall as a result, mostly in favour of the tertiary sector, with the secondary sector, which benefits from the cheaper raw materials, also increasing its share.

The contribution of the mining sector is uncertain. There is current strong downward pressure on mining investment, due to strong unions, uncertainty about future government policies, and relatively better opportunities for South African mining companies abroad. However, there is at the same time the possibility of increased mining activity through fracking. Thus the fate of the secondary sector is uncertain – it may benefit from increases in mining activity were these to happen; or it may process existing raw materials more efficiently, but a decrease in contribution is also a distinct possibility.

The contribution of the tertiary sector to the South African economy is already at a high level, for a country of South Africa's size and stage of economic development. Thus while its contribution may rise somewhat, it is unlikely to see very large increases.

Both experts felt that large changes in the contributions made by the three sectors were unlikely. A large shift would be of the order of a 3% change in a sector's share per decade. Thus, over the roughly 3.5 decades until 2050, sectors could undergo a net change of, at most, about 10%. Currently, the tertiary sector, excluding transport services contributes around 65% of South Africa's GDP.

Quantitative forecasts

Figure 4 shows the probability distributions elicited from the two experts:

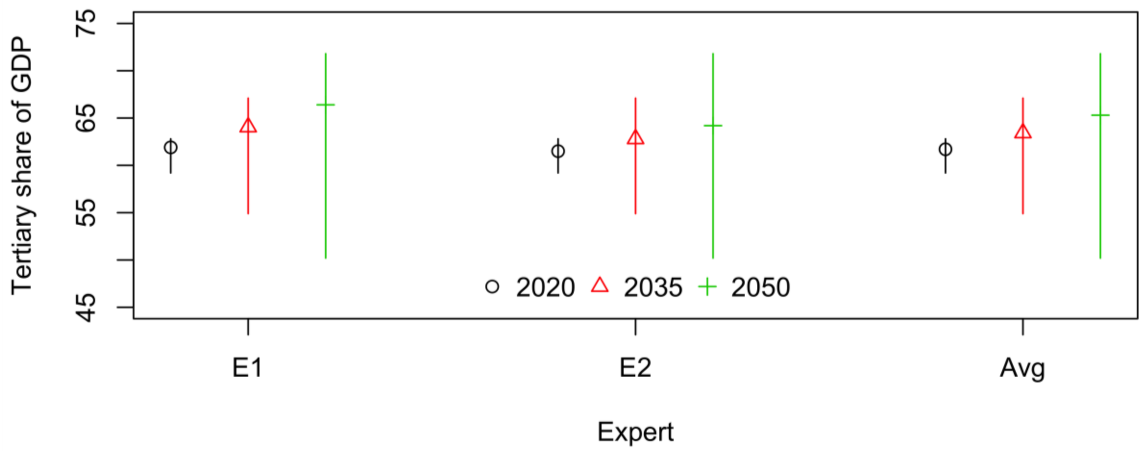

Figure 4: Elicited distributions for tertiary sector contribution to GDP (2 experts)

Post-processing

The experts' elicited distributions were combined by a simple averaging process, and converted into annual time series using Algorithm 1 and 3. When generating values for the three time points 2020, 2035, and 2050, we impose a moderate positive autocorrelation between the values obtained in consecutive periods, by setting off-diagonal correlations to 0.3 i.e. $\Sigma_{ij} = 0.3, \forall i \neq j$. Sample trajectories generated by Monte Carlo simulation are shown in Figure 5.

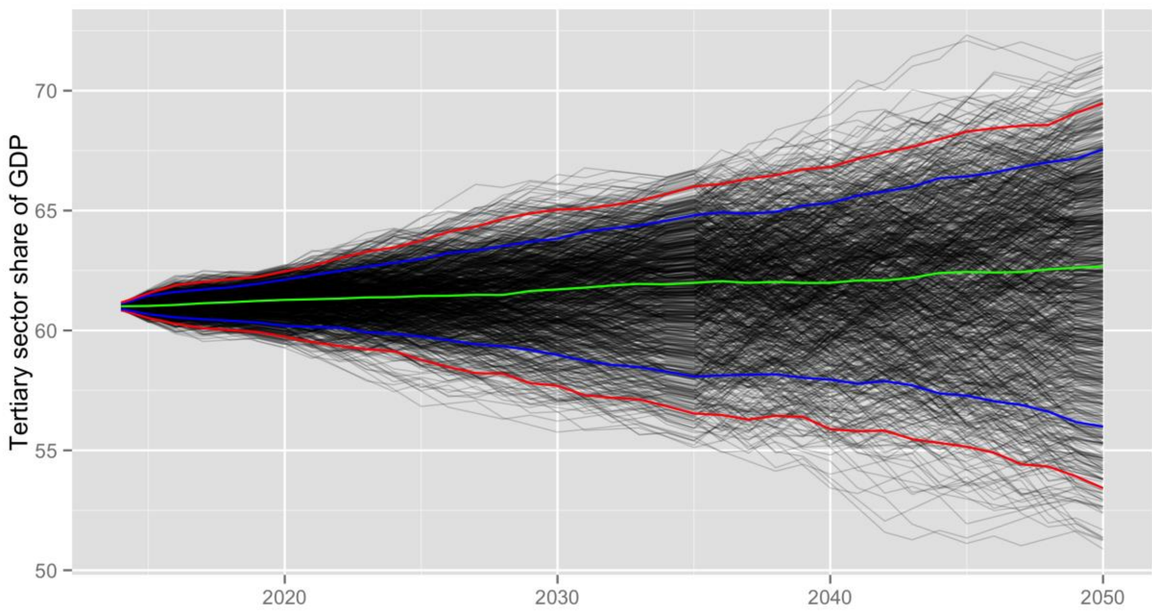

Figure 5: Probabilistic projections of tertiary share of GDP over the period 2014 – 2050.

Population growth

Overview

Models for producing probabilistic population projections have been recently developed specifically for use by the United Nation Population Division (Raftery, Li, Sevcikova, Gerland, & Heilig, 2012). We use the estimates generated by these models. For the sake of completeness we describe the methodology in detail below. (Alkema, et al., 2011)

Changes in a country's population are determined by a number of factors, but chiefly fertility and mortality. The approach employed by the UN comprises three main models: one for total fertility rate, from which trajectories of age-specific fertility rates are obtained; another estimates life expectancy at birth for females and males, which are also converted into trajectories of age- and sex-specific mortality rates; and a final model that converts the fertility and mortality trajectories into a trajectory of all population quantities of interest (e.g. total population, working age population, etc..).

Methodology

Fertility model

The model for total fertility rate (TFR) specifies that a country's TFR passes through three stages: a high fertility phase, a transition phase, and a post-transition low fertility phase. At the present time, the TFR in all countries has started to decline (Alkema, et al., 2011), so that Phase I is represented by historical data. The other two stages are each represented by their own statistical models. The starting points for Phase II and III are determined by deterministic rules. Starting points for Phase II are given by the most recent period with a local maximum fertility rate within 0.5 of the global maximum, provided this local maximum fertility exceeds 5.5. If it does not the start point is set to the beginning of the observation period. Two consecutive five-year periods of increasing TFR below a TFR of 2 children defines entry into Phase III.

In the transition phase, five-year changes (declines) in TFR are modelled using a double-logistic function

$$g(\boldsymbol{\theta}_c, f_{c,t}) = \frac{-d_c}{1 + exp\left(-2\ln(9)(f_{c,t} - \sum_i \Delta_{ci} + 0.5\Delta_{c1})\,/\Delta_{c1}\right)} + \frac{d_c}{1 + exp\left(-2\ln(9)(f_{c,t} - \Delta_{c4} - 0.5\Delta_{c3})/\Delta_{c3}\right)}$$

where $f_{c,t}$ is the current TFR i.e. in country c at time t, and $\boldsymbol{\theta}_c = (\Delta_{c1}, \Delta_{c2}, \Delta_{c3}, \Delta_{c4}, \mathrm{d}_c)$ is a set of country-specific parameters that determine the shape of $g(\boldsymbol{\theta}_c, f_{c,t})$. These parameters are estimated using a Bayesian hierarchical model, which assumes that country-specific parameters are drawn from probability distributions defined over all countries. These "world-level" probability distributions are updated using a country's historical data and Bayes' theorem, with the resulting posterior distributions conveying information about the country-specific parameter values. Parameter estimates can be obtained by drawing from the posterior distributions. The hierarchical model models changes in the fertility rate according to a random walk with drift i.e.

$$f_{c,t+1} = f_{c,t} - d_{c,t} + \varepsilon_{c,t+1}$$

where $d_{c,t}$ is the expected five-year decrement and $\varepsilon_{c,t}$ is a random distortion term. The drift term $d_{c,t}$ is given by $g(\boldsymbol{\theta_c}, f_{c,t})$ for $f_{c,t} > 1$, and is set to 0 if TFR drops below 1. The disturbance term is distributed $N(m_t, s_t^2)$ for $t = \tau_c$, the start of the fertility transition, and $N(0, \sigma(f_{c,t})^2)$ in the remainder of the phase. Gaussian world distributions are defined for each of the parameters in $\boldsymbol{\theta_c}$ (or suitable transformations of these parameters), suitably dispersed prior distributions specified for the other parameters above, and posterior distributions obtained using Markov Chain Monte Carlo.

In the post-transition phase, the TFR is modelled as a first-order autoregressive (AR) process, with its mean set to the approximate replacement fertility level of 2.1:

$$f_{c,t+1} = f_{c,t} + (1-\rho)(2.1 - f_{c,t}) + e_{c,t}$$

where ρ is the AR parameter and $e_{c,t} \sim N(0, s^2)$ is a random error term. These parameters (which are not country specific) are estimated using maximum likelihood).

Once probabilistic projections of the TFR have been obtained, these are converted into age-specific fertility rates by multiplying the TFR by age-specific percentages from the 2010 edition of the UN World Population Prospects. These percentages are applied throughout the projection period.

Life expectancy model

The life expectancy model comprises three main elements. First, probabilistic projections of female life expectancy at birth are obtained, using a Bayesian hierarchical model similar in many respects to that used for TFR projections – that is, it is also a random walk with drift, the drift term is again is a double logistic function of current life expectancy. Second, probabilistic projections of male life expectancy at birth are generated conditional on the projections for females. Finally, both of these projections are converted into age-specific mortality rates. We discuss these elements in more detail below.

Similarly to the modelling of the TFR, female life expectancy at birth in country c at time t is assumed to follow a random walk with drift

$$l_{c,t+1} = l_{c,t} + g(\boldsymbol{\theta_c}, l_{c,t}) + \varepsilon_{c,t+1}$$

where the drift term $g(\boldsymbol{\theta_c}, l_{c,t})$ models the five-year gains in life expectancy and is given by the double logistic function

$$g(\boldsymbol{\theta}_c, l_{c,t}) = \frac{-k_c}{1 + exp\left(-A_1(l_{c,t} - \Delta_{c1} - A_2\Delta_{c2})/\Delta_{c2}\right)} + \frac{d_c}{1 + exp\left(-A_1(l_{c,t} - \sum_{i=1}^{3}\Delta_{ci} - A_2\Delta_{c4})/\Delta_{c3}\right)}$$

where $l_{c,t}$ is the current life expectancy (for females at birth in country c at time t), and $\boldsymbol{\theta_c} = (\Delta_{c1}, \Delta_{c2}, \Delta_{c3}, \Delta_{c4}, \mathrm{k}_c, \mathrm{z}_c)$ is a set of country-specific parameters that determine the shape of $g(\boldsymbol{\theta}_c, l_{c,t})$. A_1 and A_2 are constants. The disturbance term $\varepsilon_{c,t+1}$ is distributed as $N(0, \omega f(l_{c,t}))$ with $f(l_{c,t})$) a smooth, declining function of $l_{c,t}$. These parameters are estimated using MCMC with the "world-level" probability distributions and associated priors for the hierarchical model given in (Raftery et al. 2012).

To generate projections of male life expectancy at birth, the gap between the life expectancy of the two sexes was modelled as

$$G_{c,t+1} = \begin{cases} \beta_0 + \beta_1 l_{c,1953} + \beta_2 G_{c,t} + \beta_3 l_{c,t} + \beta_4 (l_{c,t} - 75)_+ + \varepsilon_{c,t+1}, & \text{if } l_{c,t} \leq M \\ \beta_5 G_{c,t} + \varepsilon_{c,t+1}, & \end{cases}$$

where M is the highest observed female life expectancy, $\varepsilon_{c,t} \sim t(0, \sigma_1^2, \nu_1)$ if $l_{c,t} \leq M$ and $\varepsilon_{c,t} \sim t(0, \sigma_2^2, \nu_2)$ This model (Raftery et al. 2012) is estimated by maximum likelihood (note that no parameters are country-specific).

Projections of female and male life expectancy at birth are converted into age-specific projections of mortality rates using a modification of the Lee-Carter method. The classic Lee-Carter method rewrites the log of the mortality rate for age class x at time t as

$$m_{x,t} = \propto_x + \beta_x \gamma_t + \varepsilon_{x,t},$$

where $\propto_x$, β_x, and γ_t are parameters to be estimated and $\varepsilon_{x,t}$ is a random disturbance. This parameterization is not unique in the sense that there are an infinite number of optimal solutions that will give the same forecasts. For this reason two identification constraints are added: various options are possible but commonly $\sum_t \gamma_t = 0$ and $\sum_x \beta_x = 1$ are used.

Given a set of historical age-specific mortality rates and imposing the constraints above, the parameters above can readily be estimated using a singular value decomposition of the centered age-specific log mortality rates. To forecast mortality rates into the future, one produces estimates of $\propto_x$, β_x, and γ_t using historical data over the observation period $t = 1, \ldots, T$. Then, assuming that $\propto_x$ and β_x are constant, forecasts of $m_{x,t}$ are obtained by forecasting $\hat{\gamma}_t$ using a random walk model with drift. It can be shown that the s-period ahead forecast of $\hat{\gamma}_t$ is given by

$$\hat{\gamma}_{T+s} = \hat{\gamma}_T + s\hat{\theta} + \sqrt{s}\xi_t$$

where $\hat{\theta} = (\hat{\gamma}_T - \hat{\gamma}_1)/(T-1)$ is the drift term and ξ_t is a random disturbance. Finally, the s-period ahead forecast of $m_{x,t}$ can be computed by substituting $\hat{\gamma}_{T+s}$ into the expression above i.e.

$$m^*_{x,T+s} = \hat{\alpha}_x + \hat{\beta}_x(\hat{\gamma}_T + s\hat{\theta}).$$

The modification to the Lee-Carter method estimates $\propto_x$, β_x, and γ_t using historical data as described above, but then when making forecasts does not use a random walk model to generate forecasts of $\hat{\gamma}_t$. Rather, a value of $\hat{\gamma}_t$ is chosen such that the resulting mortality rates $m^*_{x,t}$ generated using $\hat{\gamma}_t$, when aggregated over all age classes and converted to life expectancy at birth, give a result that fits the projected life expectancies generated previously. Note, for example, that for a particular forecast period t and assuming yearly age classes the forecasted life expectancy is given by $\sum_{x=1}^{\infty} x\left(1 - \exp\left(m^*_{x,t}\right)\right) = \sum_{x=1}^{\infty} x\,(1 - \exp(\propto_x + \beta_x\hat{\gamma}_t))$. Given a projected life expectancy L^*, the sole unknown parameter $\hat{\gamma}_t$ can be solved for. Forecasts of age-specific mortality rates follow directly.

It should be noted that the UN probabilistic projections of life expectancy for South Africa and other countries with high rates of HIV/AIDS is currently a work-in-progress, and uncertainty about the effect of the epidemic on life expectancy is not explicitly modelled at present. The UN's 2012 Revision of the World Population Prospects provides a single deterministic projection that incorporates the impact of HIV/AIDS on mortality, using a range of assumptions around potential improvements in life expectancy e.g. as a result of access to anti-retroviral treatments. Trajectories of life expectancy by sex constructed for the probabilistic projections are then adjusted in such a way as to ensure that the median trajectory for each country is consistent with this deterministic forecast.

Cohort component (aggregation) model

The age- and sex-specific fertility and mortality rates generated above are used to derive a full set of population forecasts using a cohort component projection method. This approach stratifies the population of each gender into a number of ascending age classes $x = 1, 2, \ldots, N$. Then number of gender g in an age class x at time $t + 1$ is given by

For the youngest age class $x = 1$

$$n_{g,1,t+1} = \sum_x B_{g,x,t} n_{g,x,t} + m_{g,x,t}$$

For the oldest age class $x = N$

$$n_{g,N,t+1} = S_{g,N-1,t} n_{g,N-1,t} + S_{g,N,t} n_{g,N,t} + m_{g,N,t}$$

While for the remaining intermediate age classes $x = 2, \ldots, N - 1$

$$n_{g,x,t+1} = S_{g,x-1,t} n_{g,x-1,t} + m_{g,x,t}$$

where $B_{g,x,t}$ is the number of offspring of gender g born to females in age class x at time t who are born in the t-th period and survive to time $t + 1$, divided by $n_{g,x,t}$, $S_{g,x,t}$ is the survival ratio for gender g in age class x at time t, and , $m_{g,x,t}$ is the net migration for gender g in age class x at time t. These equations are applied recursively to generate population projections.

Quantitative forecasts

Sample trajectories of total population generated by Monte Carlo simulation are shown in Figure 6.

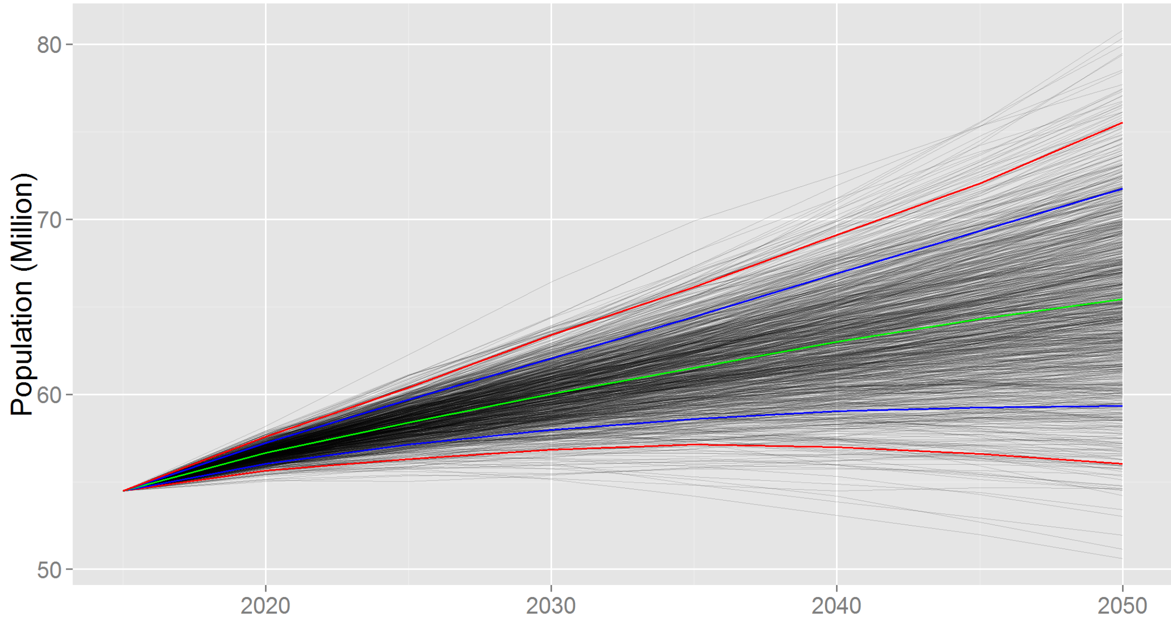

Figure 6: Probabilistic projections of total population over the period 2014 – 2050.

Global energy commodity prices

Overview

We construct trajectories for international coal, gas, and oil prices from two external sources. The IEA produces long-term forecasts of commodity prices as part of its World Energy Outlook, the most recent version of which was released in 2014. These forecasts, of coal, oil, and gas prices to 2050 under three mitigation scenarios, are perhaps the most widely-used long-term forecast of commodity prices, but there are no estimates of the uncertainty around the forecasts and are thus, on their own, they are unsuitable for our purposes. We therefore augment these values with distributions of coal, oil, and gas prices in 2020, 2035 and 2050 obtained from an application of IMACLIM-R, a hybrid energy-economic simulation model (Sassi, Crassous, Hourcade, Gitz, Waisman, & Guivarch, 2010). The IMACLIM-R data expresses commodity prices for 108 "baseline" scenarios and 108 "mitigation" scenarios, covering a range of assumptions on parameters values representing available technology, energy efficiency, lifestyle changes, and growth in labour productivity (Rozenberg, et al., 2010).

The Imaclim-R Data source

IMACLIM-R is a large-scale simulation model of the world economy comprising both static and dynamic components, described in detail in Sassi et al. (2010) and used in, for example, Rozenberg, et al. (2010). The model uses a static general equilibrium model to annually determine relative prices, wages, labour, quantities of goods and services, and value flows. Markets for production factors need not clear but goods markets are cleared by unique relative prices. These prices are determined by behaviours of agents, modelled in various sub-components of the full model.

The dynamic part of IMACLIM-R determines how the short-term constraints imposed on the static equilibrium model change over time. These changes determine the conditions under which static equilibrium will be computed in the next time step, and hence growth. Dedicated modules describe the dynamics of various sectors: fossil fuels, electricity generation, residential energy end-uses, transportation, agriculture, industry, and services.

The fossil fuel module describes the evolution of coal, oil, and gas prices, i.e. the trajectories that we use as inputs to our model, and how these prices are linked to extraction costs that are themselves related to cumulated reserves. Individual producers are modelled in detail, including resource discovery processes. Oil is divided into six subtypes according to production cost, with mark-up rates an increasing function of the ratio of current output to production capacity. Subtypes are exploited if they are sufficiently profitable. Regional imbalances in supply and consequent market power are explicitly modelled. The IMACLIM-R values express international energy prices for oil, gas and coal in 2020, 2035 and 2050, as indexes of 2010 prices, for 108 "baseline" scenarios and 108 "mitigation" scenarios.

Quantitative forecasts

		Business as usual		
		2020	2035	2050
Coal	IMACLIM-R (avg)	1.04/0.93	1.78/1.07	2.81/1.41
	IEA WEO 2015 (NP)	0.92	1.00	
	Wood Mackenzie	0.92	1.22	
	Adjustment factor	0.9	0.6	0.5
Gas	IMACLIM-R (avg)	1.10/1.27	1.30/1.43	1.37/1.45
	IEA WEO 2015 (NP)	1.33	1.49	
	Adjustment factor	1.15	1.1	0.9
Oil	IMACLIM-R (avg)	1.37/0.96	1.86/1.3	1.97/1.38
	IEA WEO 2015 (NP)	0.92	1.39	
	Adjustment factor	0.7	0.7	0.7

Table 2 shows mean commodity prices obtained from external data sources i.e. IMACLIM-R (indicated by the first of the two values in each cell of the table), the IEA WEO 2015 report, and the Wood-Mackenzie[2] forecasts. The IMACLIM-R forecasted prices are in some cases substantially larger than the IEA forecasts. As the IEA forecasts are both more widely used and more recent, we adjusted the mean IMACLIM-R values substantially in the direction of the IEA values. The sole exception is oil prices in 2020 under business as usual, where the lower estimate returned by IMACLIM-R was felt to perhaps be more plausible given recent trends in the oil price. Mean values from IMACLIM-R are subjectively adjusted to account for more recent information in the IEA WEO 2013 and other sources where available. Values in the IMACLIM-R

[2] Wood Mackenzie 2014. Johannesburg Coal Breakfast Briefing – Thermal Coal: Weathering the Storm.

rows denote indices before/after adjustment, with the multiplier used to make the adjustment provided

		Business as usual		
		2020	2035	2050
Coal	IMACLIM-R (avg)	1.04/0.93	1.78/1.07	2.81/1.41
	IEA WEO 2015 (NP)	0.92	1.00	
	Wood Mackenzie	0.92	1.22	
	Adjustment factor	0.9	0.6	0.5
Gas	IMACLIM-R (avg)	1.10/1.27	1.30/1.43	1.37/1.45
	IEA WEO 2015 (NP)	1.33	1.49	
	Adjustment factor	1.15	1.1	0.9
Oil	IMACLIM-R (avg)	1.37/0.96	1.86/1.3	1.97/1.38
	IEA WEO 2015 (NP)	0.92	1.39	
	Adjustment factor	0.7	0.7	0.7

Table 2: Mean international commodity (coal, gas, oil) prices, expressed as multiples of 2010 prices, under broad "business as usual" and "mitigation" scenarios.

Since we do not explicitly model international mitigation in the current project, we average over the two broad scenarios, "business as usual" and "with mitigation". We obtain commodity prices by applying the indices to 2010 prices: $75/ton for coal, $7.50/mbtu for gas, and $78/bbl for oil. Sample trajectories of commodity prices are obtained by sampling, with replacement, 1000 sets of prices (i.e. for 2020, 2035, and 2050) from the 108 scenarios, and linearly interpolate between the three time periods. The final trajectories are shown in Figure 7 to Figure 9.

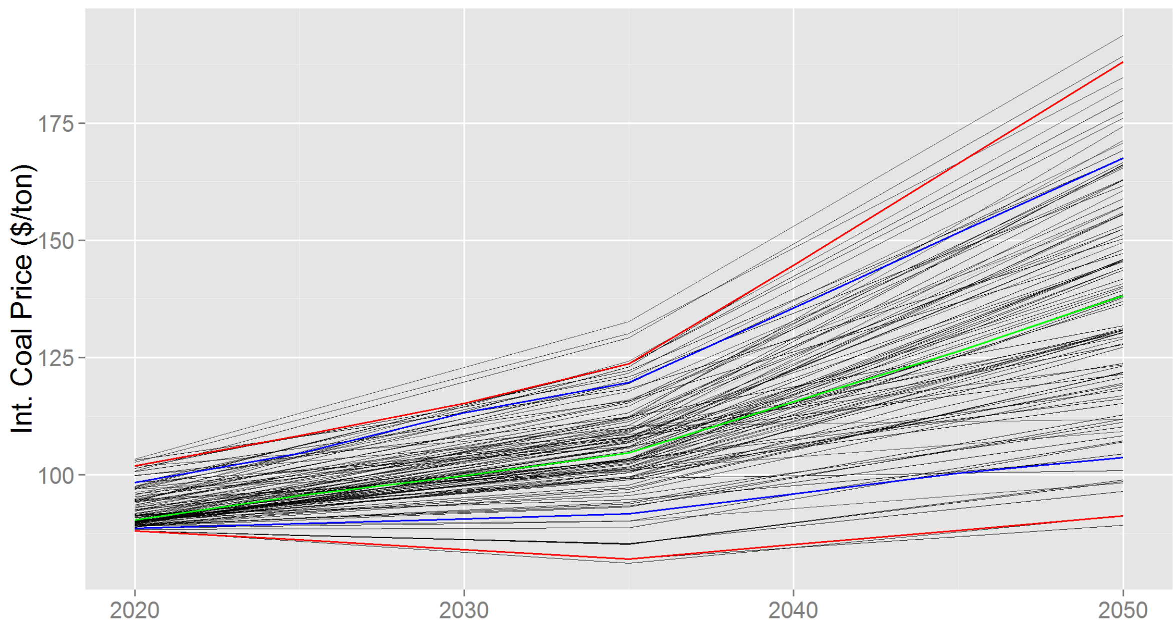

Figure 7: Probabilistic projections of international coal prices over the period 2020 – 2050.

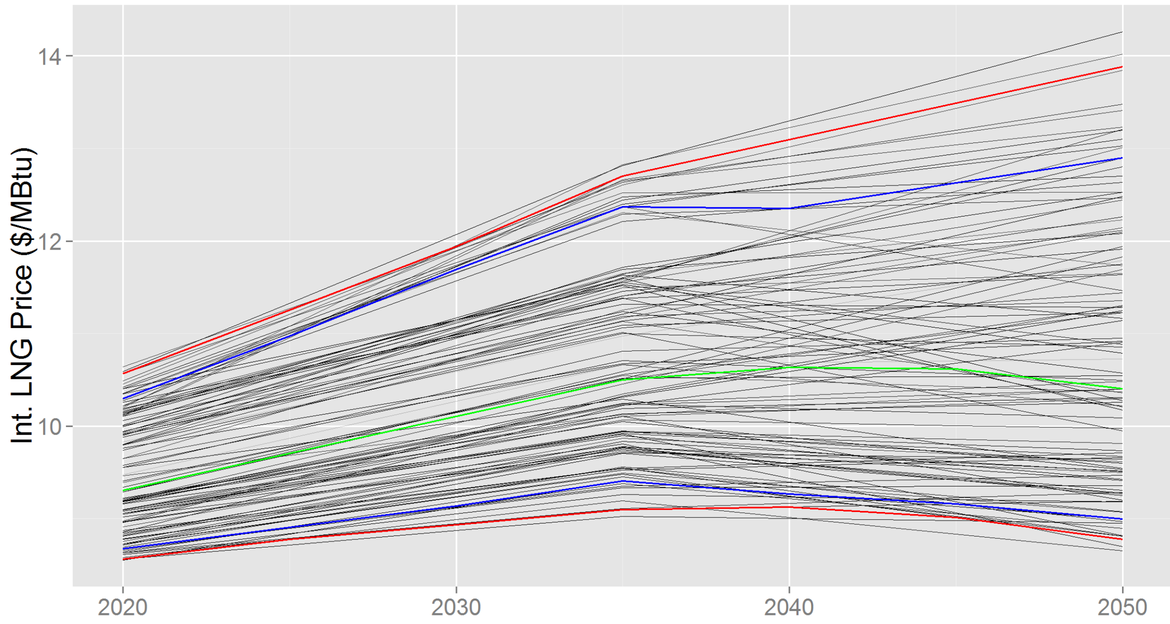

Figure 8: Probabilistic projections of international gas prices over the period 2020 – 2050.

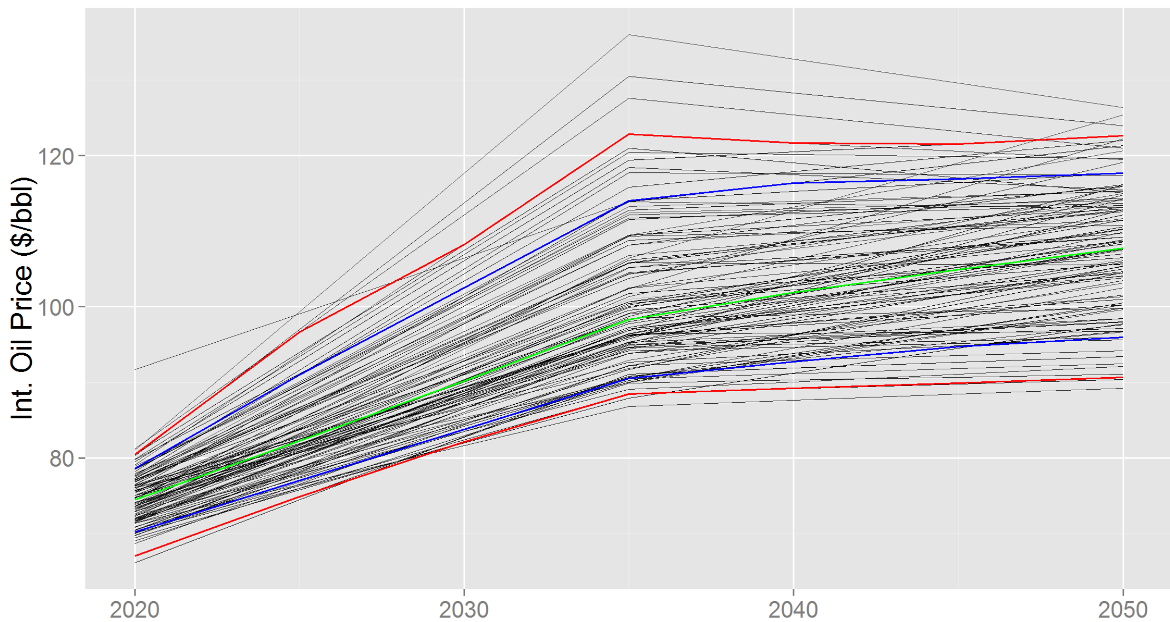

Figure 9: Probabilistic projections of international oil prices over the period 2020 – 2050.

Coal prices

Overview

We conducted elicitation interviews with four experts on the subject of future local coal prices. The elicited probability distributions thus covered possible values in the mean growth rate over these three.

Summary of qualitative discussions

Historical background and current context

Originally coalmines were tied geographically and economically to power stations. Transportation costs were small, and power stations were purpose-built to use the coal quality provided by the satellite mine (most of the coal in the Central Basin is of relatively low quality). Traditionally Eskom effectively assumed the risk of coal mining by issuing long-term contracts at a fixed rate. These were issued at cost of production plus a relatively small return on capital (historically around 9%). Later, the growth of an export market for high-grade (HG) coal led to the development of stand-alone export-focussed mines.

These two ostensibly independent systems – electricity generation using low-grade coal, and exports of high-grade coal – are in fact linked through the possibility of "washing" coal – a process that shifts the distribution of coal quality, with some proportion of coal becoming higher quality and some proportion becoming lower quality. The precise proportions and quality shifts are complex and mine-specific, but the net effect is that sometimes coal below export-grade is washed to export-grade, with the remaining lower-quality "middlings" still suitable for power generation.

already paired (i.e. with power stations) coalmines and export-based coalmines were strategically grouped, usually by rough geographical proximity, in a bid to optimise operations. The optimisation focussed on matching lower-grade middlings with power stations quality requirements and minimizing the associated transportation costs. The process included the renegotiation of long-term contracts with Eskom – again the agreed-upon price being the cost-of-production plus a relatively small mark-up.

Almost all of Eskom's long-term contracts end within the next 5-15 years. Traditionally, these contracts have made coal a relatively low-risk, low-return investment, at least in the local market. Now, with increased competition for investment from other industries and countries, global mining companies are demanding higher rates of return in order to make big capital investments – of the order of around 12%. India and China also use low-grade coal in their power stations, creating an export market (and hence competition) for coal traditionally destined for use by Eskom. Eskom's strategic dilemma is how to balance short-term contracts, which must pay a premium in order to compete with the low-grade export price and are thus both more expensive in the long run and subject to fluctuations in that market, and long-term contracts, which require huge amounts of capital to be committed.

Complicating Eskom's contractual choices further, there is substantial political pressure to expand mine ownership beyond the big traditional mining companies to "juniors" with strong black empowerment initiatives – the mining charter requires 26% black ownership, while Eskom policy targets 51%. Smaller "juniors", however, pay more than the established large companies to borrow capital. This increased cost is passed to Eskom and, ultimately, results in higher electricity prices.

Influential factors

As will be clear from the previous background, the main uncertainty is how and over what time period Eskom attempts to resolve the strategic difficulties it now faces, in the form of sourcing and contracting coal, and how mining houses respond to these initiatives. Vast amounts of coal exist in South Africa, but what remains tends to be of low quality and expensive to mine. Mining companies will only mine coal if it is profitable to do so; these expenses will have to be borne by Eskom if the coal is to be mined at all. The Waterberg area of South Africa contains a vast amount of coal but mining operations here could be more difficult and expensive than they have been in the Central Basin (suggested costs run as high as R600/t for 22-23 MJ coal without capital and transport costs, and R850/t with these costs)

The main factors influencing the future price of coal are the costs of mining it, and the necessary return on capital: logistics (mainly transportation costs by road or rail); labour costs; energy inputs (in the form of diesel and electricity); capital expenditures and the associated required rates of return on capital; environmental and social costs (acid mine draining, royalties/licensing, carbon tax); the growth of the low-grade export market; and assorted "other" costs (water costs, engineering costs, replacement capital costs, employee housing costs, and equipment costs).

Each of these factors is a complex subsystem subject ~~itself~~ to major uncertainties regarding their futures. Transportation costs are a function of the geographic proximity of mines and power stations, the efficiency to which they are paired (the mines and the power plants), and the relative importance of the export market. Expansions to the rail network require large capital expenditure, so that these tend to occur in a small number of large increments. The timing of these is extremely difficult to predict. Currently labour costs are rising faster than inflation, which is unsustainable in the long run. In the Central Basin there is little opportunity for technological improvements to reduce labour costs, but some benefits may accrue if open cast mining in the Waterberg area occurs.

In the future Eskom is expected to need to offer in the region of 10-15% return on long term contracts to compete not only with the low-grade export market but also with copper mines and other investment opportunities. This number, however, is subject to the rate offered by other investments, and also to changes in risk introduced by, for example, political events. Environmental and social costs are expected to increase but the exact magnitude of this depends on unknown changes in legislature (for example, including a cash provision for certain types of mine rehabilitation not currently required, or increases in royalty fees, which are relatively low). Water costs are generally viewed as unimportant relative to other costs, although this position is not universally held.

Trends and scenarios

In the near-term (to 2020), any major changes in the domestic coal price are likely to be caused by (a) increases in the cost of capital on renegotiated contracts, once the current ones begin to expire; (b) legislation forcing mines to set aside cash provisions for environmental costs; (c) legislation increasing government royalties; (d) changes in the export price of low-grade coal; (e) transportation costs, expected to remain high but potentially increasing even further as a relatively larger proportion of coal is transported by road (any significant rail infrastructure would take longer to build). Labour costs are expected to stay relatively constant, increasing at 1-2% faster than inflation.

To 2035, substantial uncertainties exist around what will occur when the current long-term contracts expire, but the expectation is that producers will force Eskom into agreeing to a substantially higher return on capital than it has historically paid. Large infrastructure construction will be necessary to maintain supplies of coal to Central Basin stations, increasing transportation costs. Labour costs are uncertain but are expected to fall back to the inflation rate at some point, higher wages being unsustainable. By this time, it is likely that stronger environmental legislation will be in place, increasing these costs; and also that greater royalties will be extracted.

In the period 2035-2050, the effects of coal mining in the Waterberg, if it occurs, will be most keenly felt, and transportation costs increase further, unless power stations are built there. Labour costs, if they have not already come down, are expected to do so. Open cast mining in the Waterberg is likely to increase diesel costs (at least, with currently known technology). Mines that negotiated their contacts around 2030 will again be coming to the end of their contracts, and hence will be in the “low return on capital” part of their lifespan, although this will of course depend on the length of the signed contracts and will differ from mine to mine. Legislation on royalties and/or environmental and social costs is even more likely to have been implemented by this point in time.

Quantitative forecasts

Figure 10 shows the probability distributions elicited from the four experts:

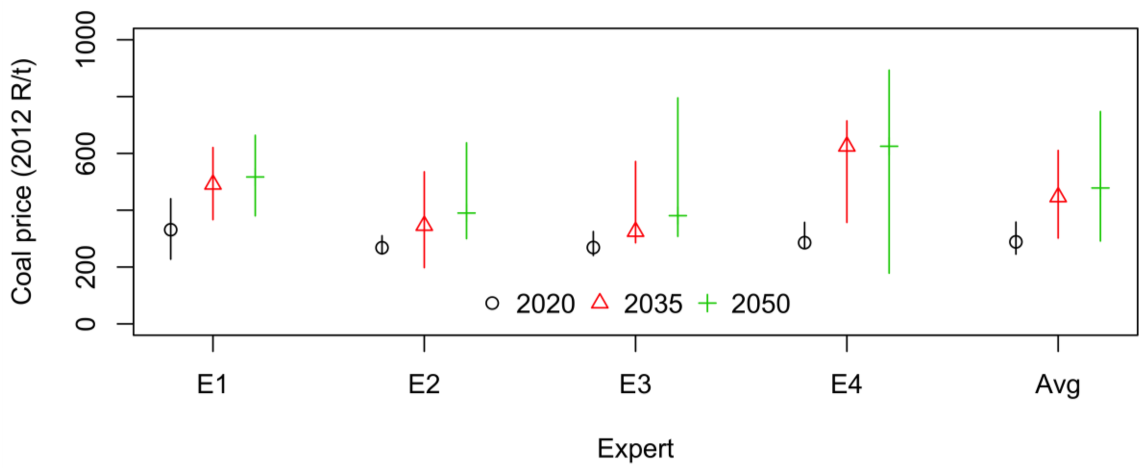

Figure 10: Elicited distributions for domestic coal prices (4 experts)

Post-processing

After reviewing the information gathered from the elicitation and literature, it was felt that simply having a single price for all the coal going to power plants in the model would be over simplistic and would miss some important factors such as:

- Some of existing power plants have ongoing fixed price contracts with mines that are close to the power plants (mine-mouth existing)
- There could potentially be cheaper coal in the Waterberg, which may or may not be worth transporting to the Central Basin, depending on the relative mining costs and the rail transport infrastructure
- The demand for coal from coal power plants depends on how competitive coal electricity is with other options for generating electricity
- The coal power plants relying on uncontracted coal would be competing with other domestic users of coal as well as global markets

The different supply options and their costs can be combined to form a crude supply cost curve for coal. The combined demand by coal power plants and other users when applied to the cost curve would determine the coal price. The supply cost curve was implemented in SATIM as shown in the diagram below.

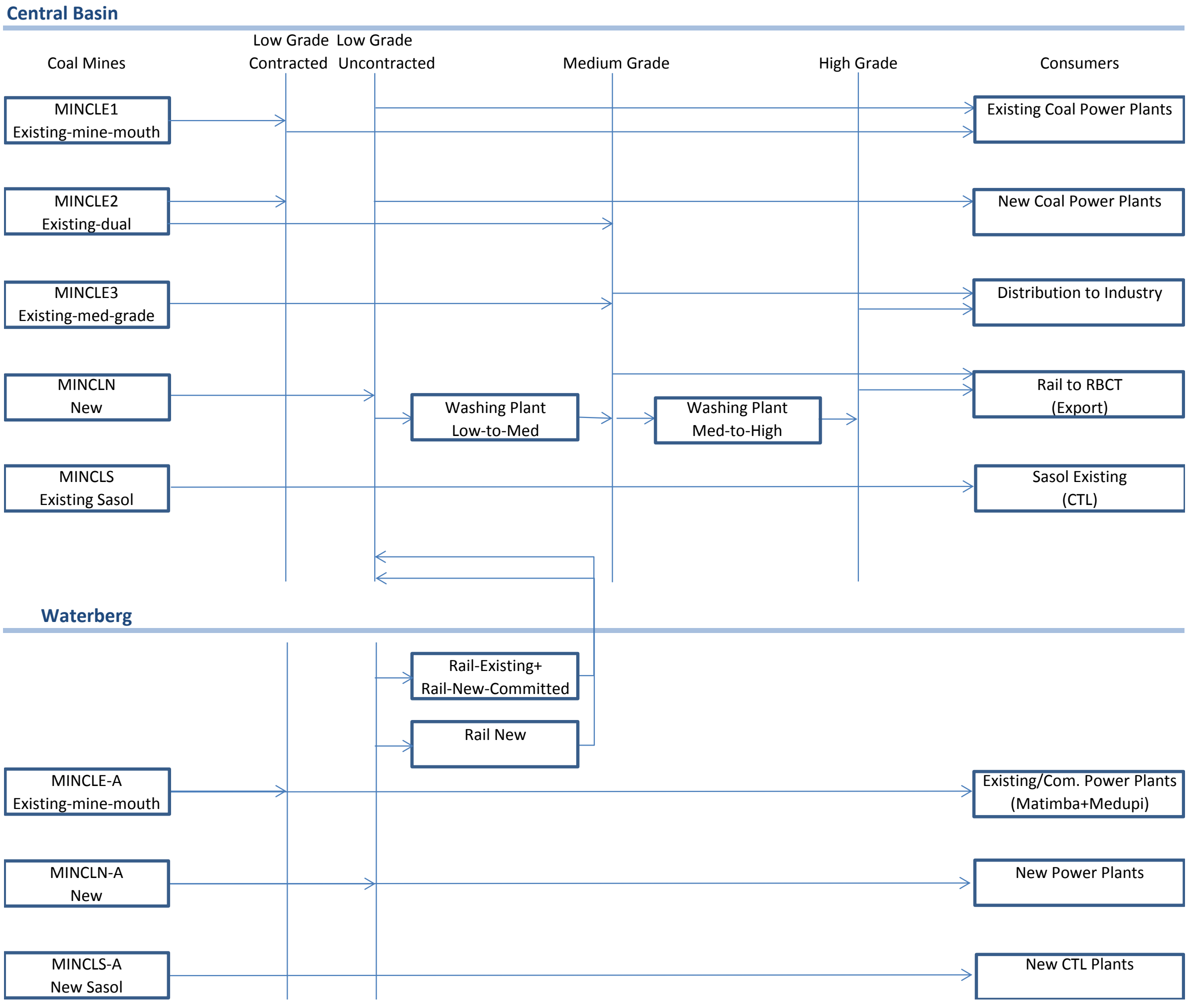

Figure 11 Implementation of Coal supply Curve in SATIM

The diagram shows 2 existing supply and one new supply route for power plants in the Central Basin:

- MINCLE1: Existing mine-mouth (conveyor link)
- MINCLE2: Existing (trucked/railed)
- MINCLN: New (trucked/railed)

The Waterberg has 2 supply routes:

- MINCLE-A: Existing mine-mouth (conveyor link)
- MINCLN-A: New (truck/rail/conveyor)

The supply to coal to liquids plants is modelled with two routes:

- MINCLS: Existing coal mines supplying Sasol's existing CTL plant in the Central Basin
- MINCLS-A: New coal mine to supply future CTL plants in the Waterberg

We also model one existing export grade producing route (MINCLE3) in the central basin. Discard coal, is modelled but is not shown in the diagram.
Over time the existing mines run out and the demand has to be met with production from new mines. The export grade and domestic high grade are then produced by washing coal produced by MINCLN. The existing rail lines from the Waterberg to the Central Basin and the export line to Richard's Bay as well as options to expand those lines are included.

The uncertainty range that was explored mainly concerns the new Central Basin and Waterberg low grade producing mines. The table and chart below show the cost breakdown per ton of 21MJ/kg for the aggregate of the existing mines in both regions as well as the extreme and median cases for both regions.

Table 3 Breakdown of parameters for the coal supply curve and associated uncertainty

	Central Basin					Waterberg						
	Con-veyor	Existing truck & rail	New truck & rail			Existing	New Surface			New Underground		
			L	M	H		L	M	H	L	M	H
Saleable Production Cost	188	200	200	200	232	106	132	281	595	298	364	592
Transport	1	100	84	100	100	1	10	17	23	10	17	23
Capital	46	46	59	59	59	23	27	27	27	68	68	68
Return on Capital	24	24	129	161	211	33	59	77	96	148	194	241
Acid mine drainage	0	0	10	30	50	0	10	30	50	10	30	50
Total	**259**	**370**	**482**	**550**	**652**	**163**	**238**	**432**	**790**	**534**	**672**	**975**

The acid mine drainage rehabilitation costs range from 10 to 50 R/ton based on literature (Golder Associates,2010) and the expert interviews. The return on capital ranges between 10% and 15% also based on literature (Macquarie) and interviews. The capital costs for the central basin are based on a capital intensity assumption of around 1800 R/ton of washed product (including cost of washing plant), and a life of 30 years for the mine. The capital intensity for a new surface mine in the Waterberg is assumed to be around 800 R/ton is based on the Resgen Boikarabelo project.

The saleable production cost is a function of labour, energy inputs and other running costs per ton mined, and the stripping ratios and washing plant yields. The large uncertainties in the Waterberg region is due to the uncertainty around the stripping ratios and washing yields.

The transport cost is a function of the how the coal is transported from the mine to the power plant (conveyor/rail/road) and unit cost of transport of each mode. We assume a higher share of non-conveyor transport in the Central Basin assuming that the new mines will not be located near power plants. The price of diesel is an important factor for road and this is endogenous to the model. The assumed ranges for mining and transport are shown in the table below.

Table 4 Detailed assumptions on uncertainty on the stripping, washing and transport parameters

Coal mining assumptions	Central Basin					Waterberg						
	Conveyor	Existing truck & rail	Conveyor			Existing truck & rail	Surface			Underground		
			L	M	H		L	M	H	L	M	H
Stripping Ratio	2	1.8	1.8	1.8	1.8	0.6	0.6	1.6	2.5	4.5	4.8	5
Washing Yield	80%	70%	70%	70%	70%	50%	40%	33%	25%	70%	60%	50%
Transport assumptions (Share of total coal transported from mine to power plant)												
Conveyor	100%	0%	0%	0%	0%	100%	75%	58%	40%	75%	58%	40%
Rail	0%	21%	40%	21%	21%	0%	13%	21%	30%	13%	21%	30%
Road	0%	79%	60%	79%	79%	0%	13%	21%	30%	13%	21%	30%

The resulting average prices for the Central Basin, Waterberg and combined is shown below. Figure 13 below shows the result of averaging the costs sampled from the ranges described above weighted by the production (shown in Figure 13) of each supply route for each corresponding cost scenario. The weighted average matches the combined elicited values shown in Figure 10 quite well.

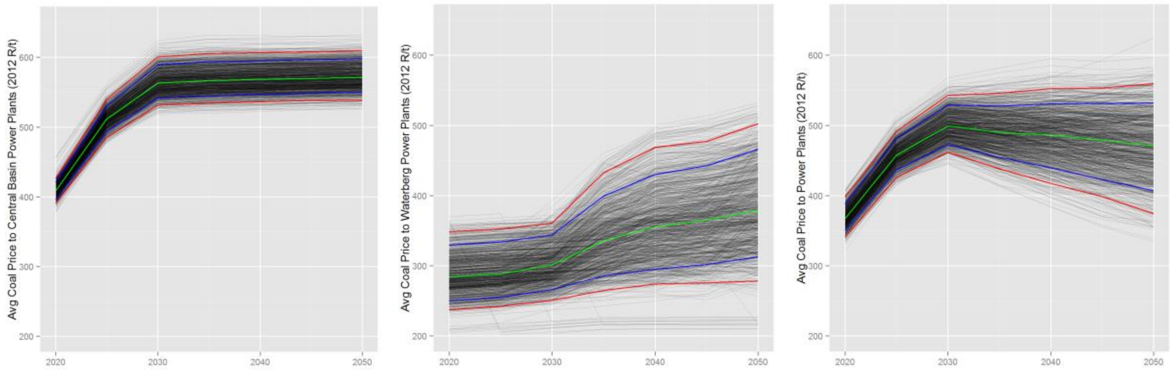

Figure 12 Average coal price seen by coal power plants

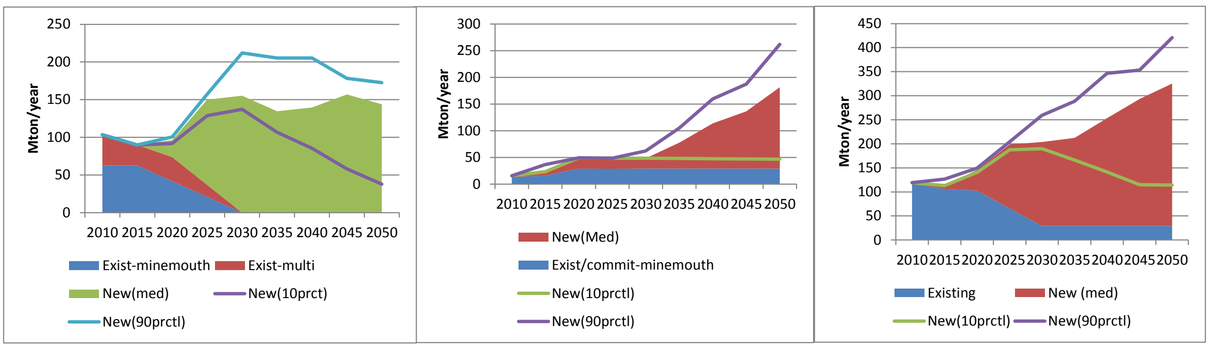

Figure 13 Production Range for different supply routes

Gas prices

Overview

We conducted elicitation interviews with two experts on the subject of gas prices. Prices depend primarily on the type and origin of the gas. An initial discussion constructed 8 reference categories: conventional gas deposits, unconventional deposits (shale, coal-bed methane), gas imported from SADC countries (by pipeline, by LNG terminal or LNGT, by floating storage regasification unit or FSRU), and gas imported from outside the SADC region (by LNGT or FSRU). The final model included only shale gas, conventional deposits, and LNG imports, so that we focus on these categories here. The elicited probability distributions cover possible prices of these different gas types in the reference years 2020, 2035, and 2050.

Summary of qualitative discussions

Influential factors

The main factors influencing production costs, and hence gas prices, are the "raw" costs of extracting the gas at the wellhead, transportation costs, and the costs of building the related infrastructure. Each of these major costs is influenced by a number of factors. In addition, when referring to the price of gas for electricity production, the price of gas for alternate uses (e.g. heating) is perhaps *the* key determinant.

Wellhead costs depend primarily on the size, location and geology of deposits, in particular whether the deposit is inland or offshore. Transportation costs depend on whether the gas is transported by pipeline, LNG terminal, or FSRU. Infrastructure for gas potentially demands massive capital investment, so that the determinants of these costs must include the complexities of capital arrangements. Pipeline costs depend on the installed capacity (e.g. 24- or 36-inch pipelines), which is itself a function of expected demand. The cost of this capital investment is typically recovered over the lifetime of the pipelines via tariff structures, which are a function of whether pipes are running at or near full capacity. When pipelines are underutilised per-unit transport costs rise to defray capital costs.

Other price determinants are essentially a consequence of the flexibility of gas in its final uses, and the fact that it is a commodity traded for profit. Gas is priced using netback pricing – pricing based on subtracting the costs of bringing a resource to the marketplace from all the revenues generated by that resource – and hence sales of gas in one market (e.g. electricity generation) depend on the value that that gas can fetch in alternate markets (e.g. industrial processes, heating). Moreover gas companies must compete with other investment possibilities for available capital. Over the longer term, this means that prices are (somewhat) self-regulating. If prices drop dramatically, there is an effective oversupply of capital, some of which will be withdrawn and invested elsewhere. Demands for return on capital are in turn influenced by perceived risk and hence by government policy and political instability.

Trends and scenarios

Current strategic planning around gas in South Africa centers on the development of large shale deposits in South Africa and even larger deposits of conventional gas in Mozambique. Opinions are that, given the size of the available deposits and the current political climate, shale deposits are highly likely to be developed; though uncertainty exists as to what extent.

Southern Mozambican gas is currently transported via pipeline but this is almost at capacity and expansions are planned. Two LNG trains are planned for 2020, with more LNG trains expected by 2035. Much larger deposits exist in Northern Mozambique but these are unlikely to be transported by pipeline as distances are too great. This possibility becomes even more remote if shale gas is developed in South Africa. Transportation via LNG train to Richard's Bay is more likely. Significant infrastructure and hence capital expenditure would be needed for either option. Development of this option would require substantial regional co-operation, but the possibility of this seems, currently, limited. The Mozambican government has a share of ownership of the NM deposits, but this is shared with gas companies who are driven primarily by profits. More likely, South Africa will have to compete with other customers for Mozambican gas on the open market.

Conventional gas deposits in South Africa has thus far been limited in their scale and impact, and this is expected to continue. Coal-bed methane is an expensive, location-dependent source with high environmental and water treatment costs. Although potentially quicker to develop than shale, it will probably not be an important prospect for South Africa.

Uncertainty ranges on gas price forecasts are relatively constant through time, in contrast to other commodities like oil and coal. This is because the gas price is effectively set at the marginal producer's cost of production (including capital). This is very likely to be a conventional off-shore gas well with a LNG liquefaction plant attached. Crucially, this step in the supply curve is very large in terms of global volumes, both now and in the long term, and the costs of production for these units are more or less homogeneous. Short-term fluctuations can of course occur because of imbalances between supply and demand, but in general these trends will be relatively short-lived (i.e. a few years).

Quantitative forecasts

Figure 14 shows the probability distributions elicited from the two experts. We have quantified this assessment using an average of the expert's assessment of the different gas types at [3.9,17.5] for shale and [6.8, 13.7] for imported LNG.

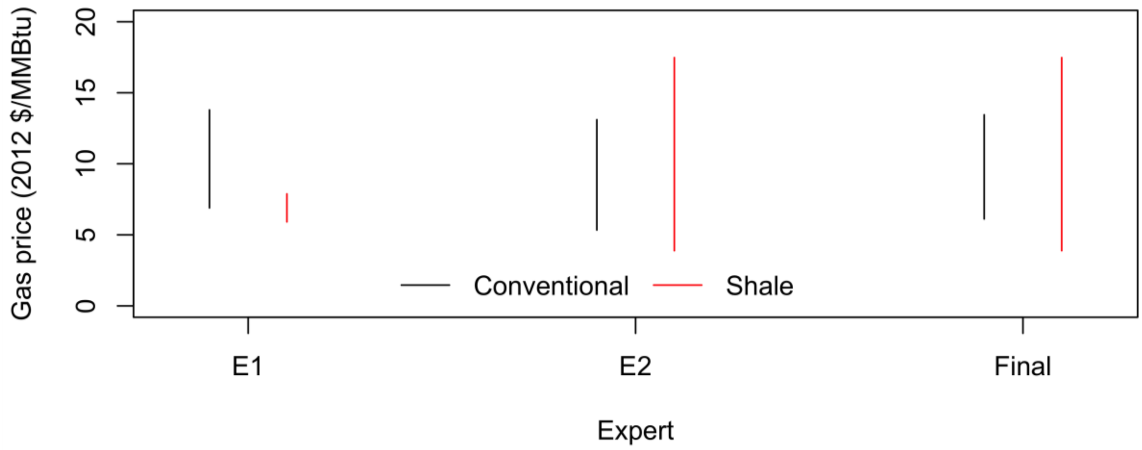

Figure 14: Elicited distributions for domestic conventional gas and domestic shale gas (2 experts)

Post-processing

Between-expert agreement is particularly low for shale gas prices. We suspect that Expert 1's assessed price range for shale takes into account factors endogenous to the SATIM model. In particular the upper bound of $8/MMBtu appears to assume that at higher prices shale gas would be superseded by other gas types and thus not available locally. These are plausible causal explanations, but they can also be accounted for within SATIM. We therefore preferred to use the broader uncertainty bounds provided by Expert 2 in the case of shale gas.

The aggregated distribution was converted into annual time series using Algorithm 1 and 3. When generating values for the three time points 2020, 2035, and 2050, we impose a moderate positive autocorrelation between the values obtained in consecutive periods, by setting off-diagonal correlations to 0.3 i.e. $\Sigma_{ij} = 0.3, \forall i \neq j$. Sample trajectories generated by Monte Carlo simulation are shown in Figure 15.

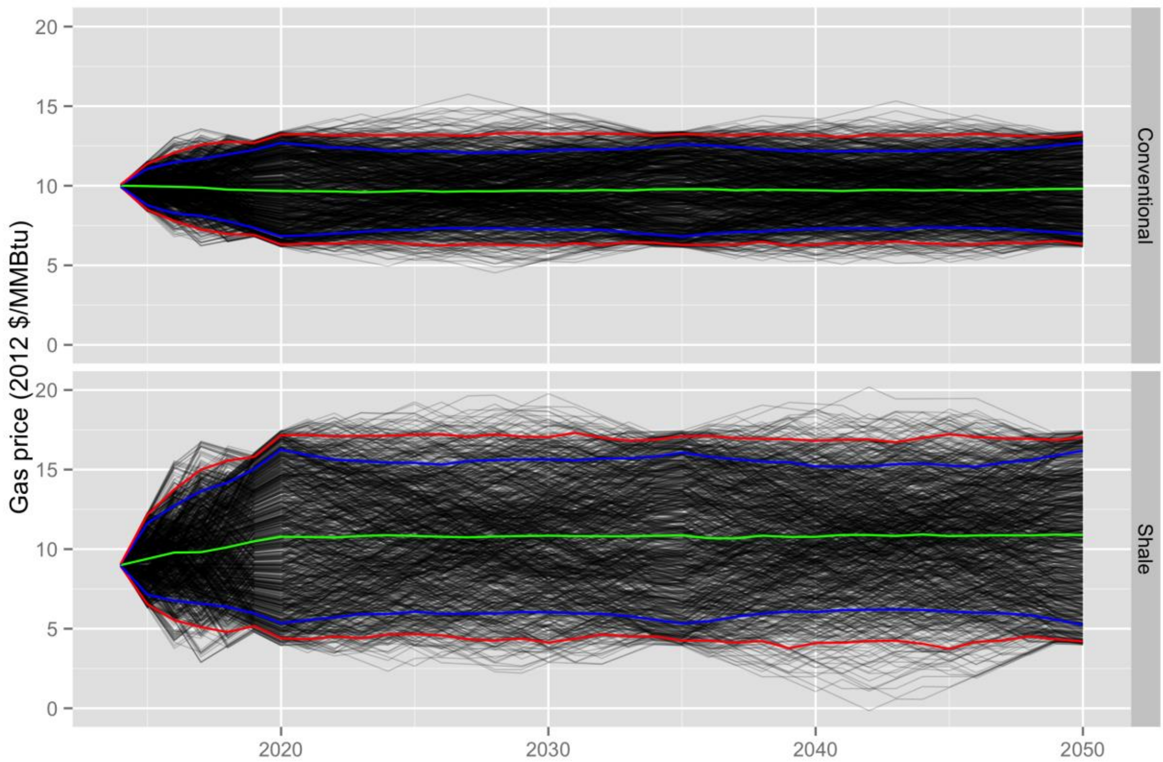

Figure 15: Probabilistic projections of conventional and shale gas prices over the period 2014 – 2050.

Solar (PV and CSP) investment costs

Overview

Solar technologies are relatively young, with further advances generally expected to lead to lower costs. We generate possible overnight investment costs using a simple learning model, using the following four-stage approach.

1. For a baseline solar technology, simulate the total installed capacity at 2030 and 2050, using distributions obtained from external sources.
2. Simulate a learning rate over the period 2014-2030 and 2030-2050, using historical learning rates with some additional uncertainty added.
3. Use standard learning models scaled to a benchmark of 2010 solar costs to calculate investment costs over the period 2014-2050.
4. Calculate investment cost trajectories for other solar technologies by scaling the costs of the baseline technology according to current price differentials.

These steps are described in detail below. Since calculations for PV and CSP are very similar, we treat these areas together in this section.

Methodology

Total installed capacity

The 2014 IEA ETP Report (Table 4.3, p148) gives expected values of total installed capacity of PV and CSP in 2030 and 2050 under 2-degree (with or without high renewable activity), 4-degree, and 6-degree scenarios. We use the 4-degree and 2-degree (without high renewables) estimates as lower and upper bounds for our forecasts. These are given, together with estimates of recent installed capacity, in Table 5.

	PV			CSP		
	2014	**2030**	**2050**	**2013**	**2030**	**2050**
4DS	176	602	1813	3.4	40	185
2DS	176	1927	4626	3.4	155	646

Table 5: Estimates of total installed capacity of solar technologies (GW)

We model total installed capacity in 2030 and 2050 as a beta distribution scaled to lie between the bounds given in Table 5. The beta distribution allows for a flexible modelling of constrained random variables. We model PV capacity as $\mathcal{B}(2,3)$ and CSP capacity as $\mathcal{B}(3,3)$. The resulting scaled beta distributions for total installed PV and CSP capacity are shown in Figure 16 and Figure 17 respectively.

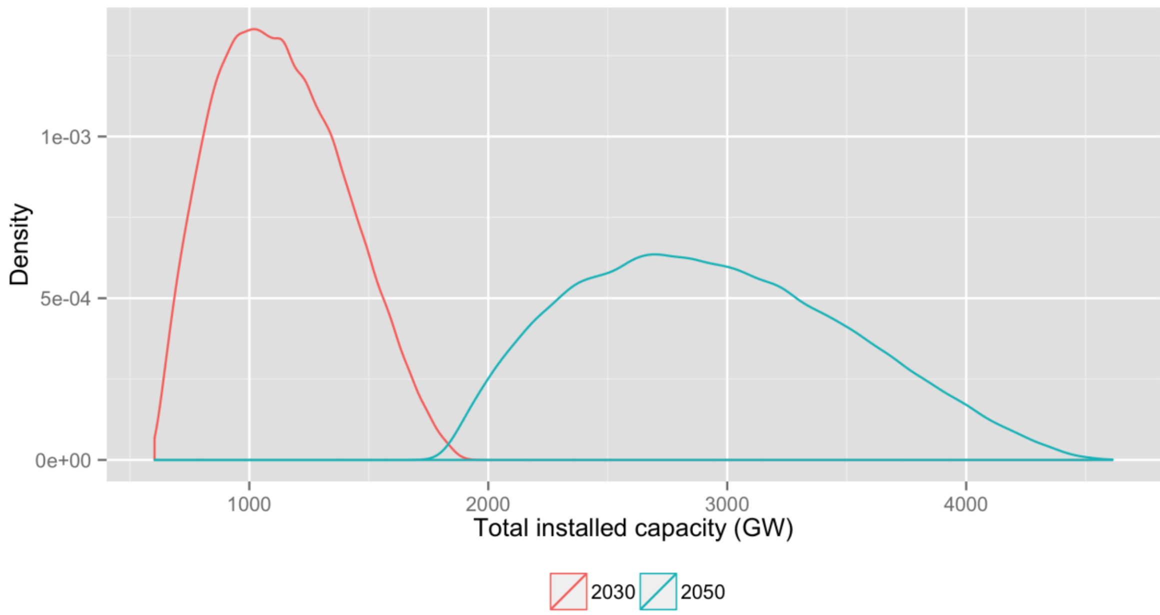

Figure 16: Probability distributions used to simulate total installed capacity of PV in 2030 and 2050, under different mitigation scenarios. Note the log-scaling used on the horizontal axis.

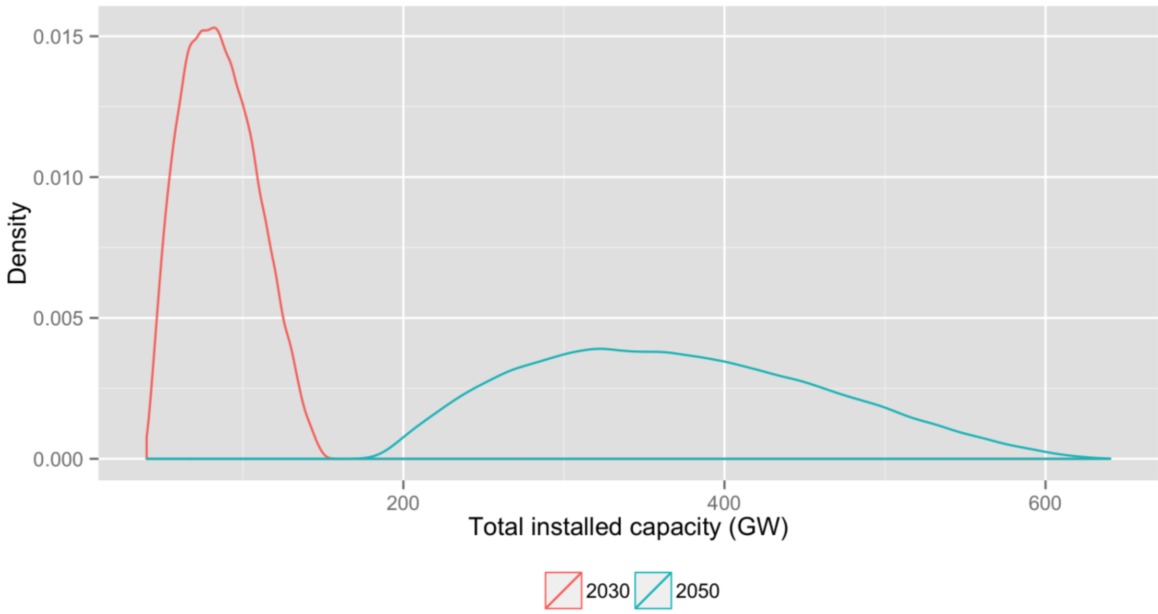

Figure 17: Probability distributions used to simulate total installed capacity of CSP in 2030 and 2050, under different mitigation scenarios. Note the log-scaling used on the horizontal axis.

Learning rates

Learning rates for CSP are simulated from a beta distribution $\mathcal{B}(4,2)$ scaled to lie between 5% and 11%. This gives a symmetric distribution centered on 9%, with 95% of the probability mass lying between 5.7% and 10.5%.

Learning rates for PV modules are simulated from a beta distribution $\mathcal{B}(3,3)$ scaled to lie between 18% and 25%. This gives a symmetric distribution with a median of 21.4%, and 95% of the probability mass lying between 19.3% and 23.6%. Learning rates for PV balance-of-system are simulated from a beta distribution $\mathcal{B}(3,3)$ scaled to lie between 9% and 12.5%. This again gives a symmetric distribution with a median of 10.8%, and 95% of the probability mass lying between 9.6% and 11.9%.

The resulting scaled beta distributions for learning rates on PV and CSP are shown in Figure 18.

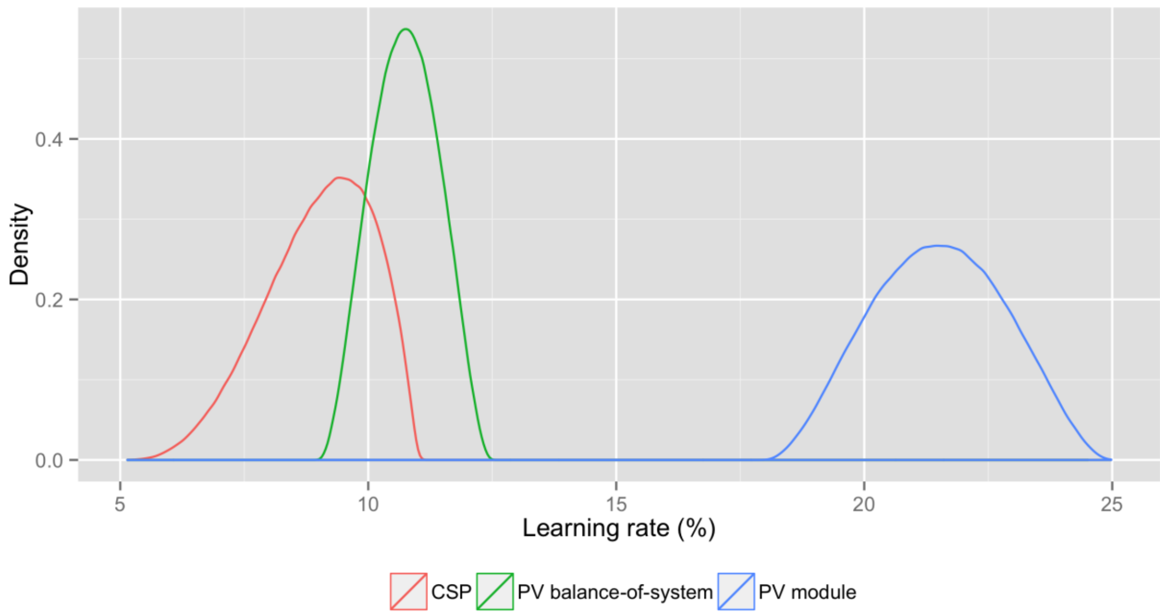

Figure 18: Probability distributions used to simulate learning rates of CSP and PV components.

Investment costs for baseline solar technologies

Investment costs Y are calculated as a function of total installed capacity C using a standard learning model

$$Y = Y_0 \left(\frac{C}{C_0}\right)^{\log_2(1-b)}$$

where Y_0 and C_0 are investment costs and total installed capacity at some baseline period and b is the learning rate.

For CSP, there is no long-term history with which to reliably estimate learning rates and identify baseline periods. We therefore initialise our learning rate using the most recent information available to us, the empirical project costs of a CSP plant constructed in South Africa in 2013, giving starting values of $Y_0 = 6.42$ and $C_0 = 3.4$.

The situation for PV is somewhat more complex. Considerable historic data exists for PV module costs, from which a learning rate of 20% has been estimated. Prior to 2000, cost predictions made using a learning model above with $b = 0.2$ matched observed costs almost exactly, but since then costs have decreased both slower and faster than predicted by learning alone in 2013-2014. Currently, costs are substantially lower than what the standard learning model would predict, but we consider this to be unsustainable, as it is largely due to oversupply and Chinese government subsidies. We therefore assume that the trajectory of observed costs will return to the trajectory predicted by the learning with $b = 0.2$, but that the time taken to achieve this return is uncertain.

We operationalize this by forming a weighted average of the learning model above (predictions made using $b = 0.2$ and a starting point of, for example, $Y_0 = 1.39$ and $C_0 = 70.0$) and the current PV module cost of $Y_* = 0.8$. The final cost is then given by

$$Y = \omega Y_0 \left(\frac{C}{C_0}\right)^{\log_2(1-b)} + (1-\omega)Y_*$$

where ω is a linear weighting function taking on the value $\omega = 0$ at the current installed capacity and $\omega = 1$ at some uncertain future time τ (i.e. the year in which the learning curve is rejoined). Noting that the learning model with $b = 0.2$ predicts costs below the current costs $Y_* = 0.8$ at a capacity of $C = 390$ GW, the predicted capacity in 2020. We generate τ from a beta distribution $\mathcal{B}(3,3)$ scaled to lie between 2016 and 2027. The resulting distribution is shown in Figure 19.

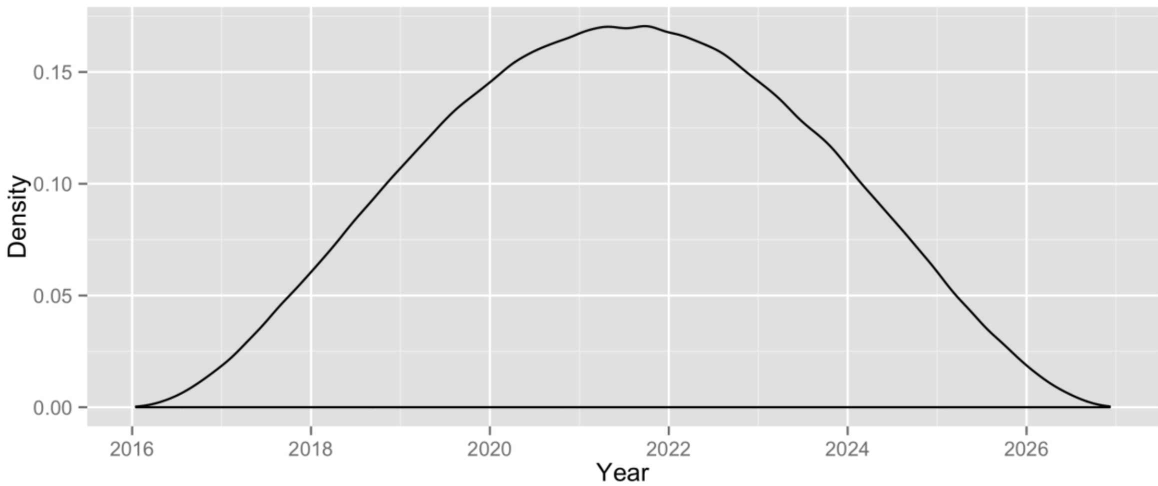

Figure 19: Probability distributions used to simulate the total installed capacity at which PV module costs regain the historical learning curve.

For PV balance-of-system costs, current costs are $Y_0 = 0.93$ at a total installed capacity of $C_0 = 130$. Learning rates for the balance-of-system costs, however, are expected to be substantially lower than historical learning rates for module costs, as reflected in Figure 18.

Investment costs for other solar technologies

The investment costs calculated in the previous step are for "baseline" PV and CSP technologies: specifically, for PV, Utility with no tracking and, for CSP, a parabolic trough with 6-hour storage capacity. Costs for other technologies are calculated as a multiple of the costs of the baseline technologies. The multipliers are fixed at their current values i.e. the present-day (2014) cost ratio, shown in Table 6 and Table 7.

PV Technology	Cost multiplier
ERSOLPCF-N (fixed axis)	1
ERSOLPCT-N (tracking axis)	1.16
ERSOLPRC-N (commercial rooftop)	1.30
ERSOLPRR-N (residential rooftop)	2.39

Table 6: Investment cost ratios of the balance of plant of the different PV technologies, relative to the baseline technology, ERSOLPCF-N.

CSP Technology	Cost multiplier
Parabolic trough with 6h storage	1
Parabolic trough with 0h storage	0.58
Central tower with 3h storage	0.74
Parabolic trough with 3h storage	0.79
Central tower with 6h storage	0.88
Central tower with 9h storage	1.01
Central tower with 12h storage	1.15
Parabolic trough with 9h storage	1.20
Central tower with 14h storage	1.24

Table 7: Investment cost ratios of different CSP technologies, relative to the baseline technology, a parabolic trough with 6-hour storage capacity.

Quantitative forecasts

Sample trajectories of investment costs for baseline PV and CSP technologies generated by Monte Carlo simulation are shown in Figure 20 to Figure 21.

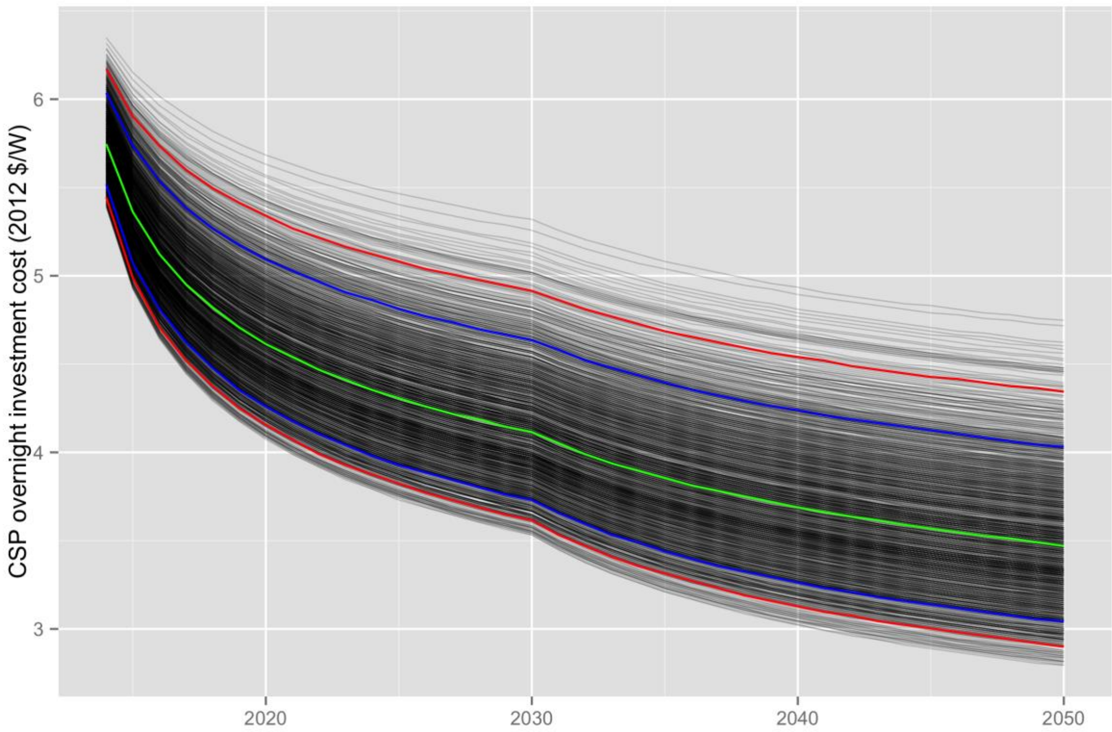

Figure 20: Probabilistic projections of CSP overnight investment costs for the baseline technology over the period 2014 – 2050.

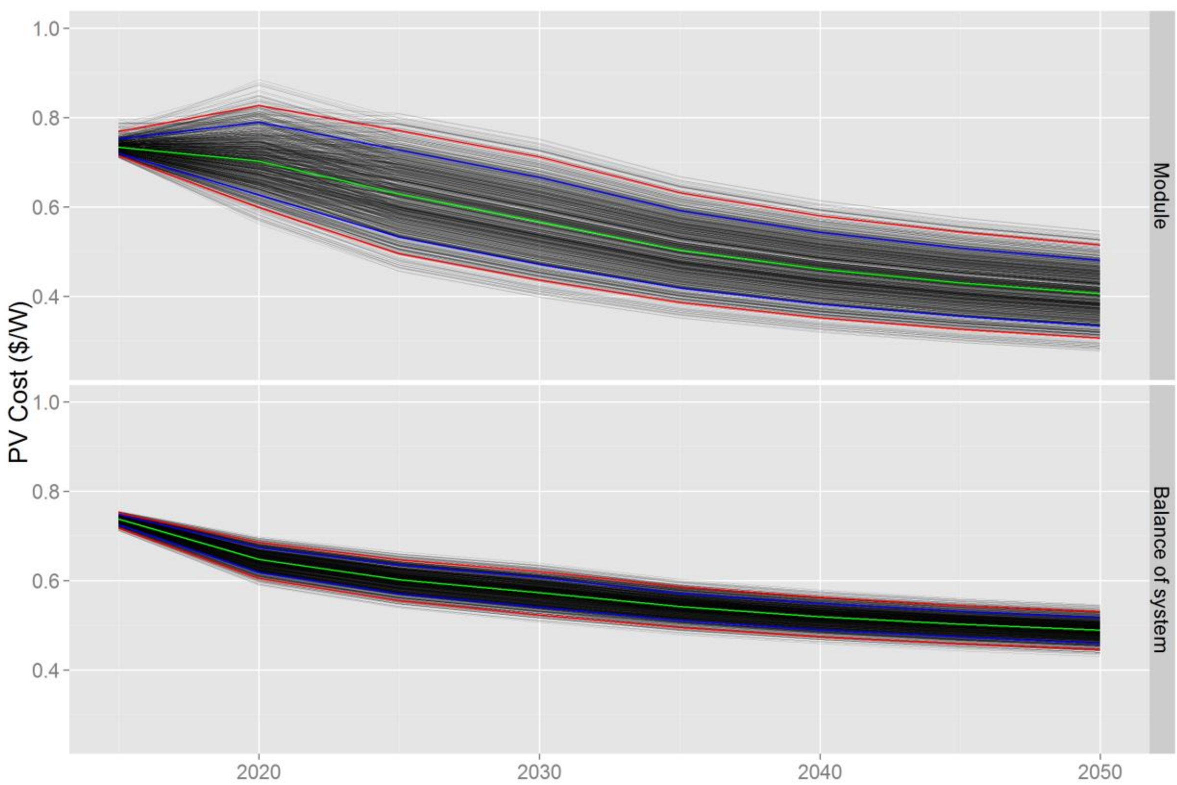

Figure 21: Probabilistic projections of overnight investment costs for the baseline PV technology over the period 2014 – 2050.

Nuclear costs

Overview

Anadon, Nemet, & Verdolini (2013) report responses from 67 experts about the future costs of nuclear power. Experts provided medians and 10% and 90% percentiles of expected overnight capital costs in 2010 and 2030 for Generation III/III+ reactors under business-as-usual investment in R&D. Estimates were also obtained for other reactor types under different R&D investment scenarios, but these estimates were not elicited from all experts and R&D investment scenarios were set to expert-specific “desired” levels and are thus difficult to standardise across experts. We therefore did not use these additional assessments.

Quantitative estimates

Assessments made by US and European experts cannot be directly used as estimates of nuclear costs in South Africa due to different material and labour costs. However, as relatively few nuclear facilities are built worldwide we might expect future trends and uncertainties in costs to be roughly comparable between countries that adopt similar regulation around nuclear facilities. We therefore standardised each expert’s assessment by expressing their judgments relative to their 2010 median assessments. That is, their 2010 median judgments were set to 100, and all other judgments were calculated relative to this baseline.

Through this transformation we found that, on average, experts’ 10% percentile assessments were 75% of their 2010 median assessments in 2010 and 78% of their 2010 median assessments in 2030. Experts’ median percentile assessments were 102% of their 2010 median assessments in 2030 (and of course 100% in 2010). Experts’ 90% percentile assessments were 133% of their 2010 median assessments in 2010 and 135% of their 2010 median assessments in 2030.

It is thus clear that experts express relatively little change in uncertainty ranges between 2010 and 2030, and this might well be reasonably extrapolated to 2050. Conservatively though, we made our 2050 assessments 10% more uncertain than 2030, giving a ratio of 0.76 for the 10% percentile, 1.02 for the median, and 1.38 for the 90% percentile.

For other key drivers, we assess estimates such that it is “very unlikely” that more extreme values occur. Without specifying the precise percentile, we suggest that the resulting judgments are more extreme than the 10/90% percentiles used here. We therefore introduce a simple mechanism for making the judgments more extreme: before applying the transformation above, we first expand the range of each expert’s assessed judgments by assuming a triangular distribution with the specified percentiles, and extrapolating this distribution to its minimum and maximum values. This results in a triangular distribution with parameters (0.71, 1, 1.41) in 2010, (0.74, 1.02, 1.43) in 2030, and (0.71, 1.02, 1.47) in 2050.

Finally we apply these distributions to the most recent estimate of overnight investment cost in South Africa, the $5800/kW (2012 dollars) given in the 2013 update to the IRP (DOE 2013). The final distributions used are:

Year	Minimum	Mode	Maximum
2010	4109	5800	8200
2030	4301	5942	8269
2050	4119	5942	8528

Table 8: Parameters of constructed triangular distributions for overnight investment cost of Gen. III nuclear facility (2012$/kW)

Sample trajectories of nuclear costs generated by Monte Carlo simulation are shown in Figure 22.

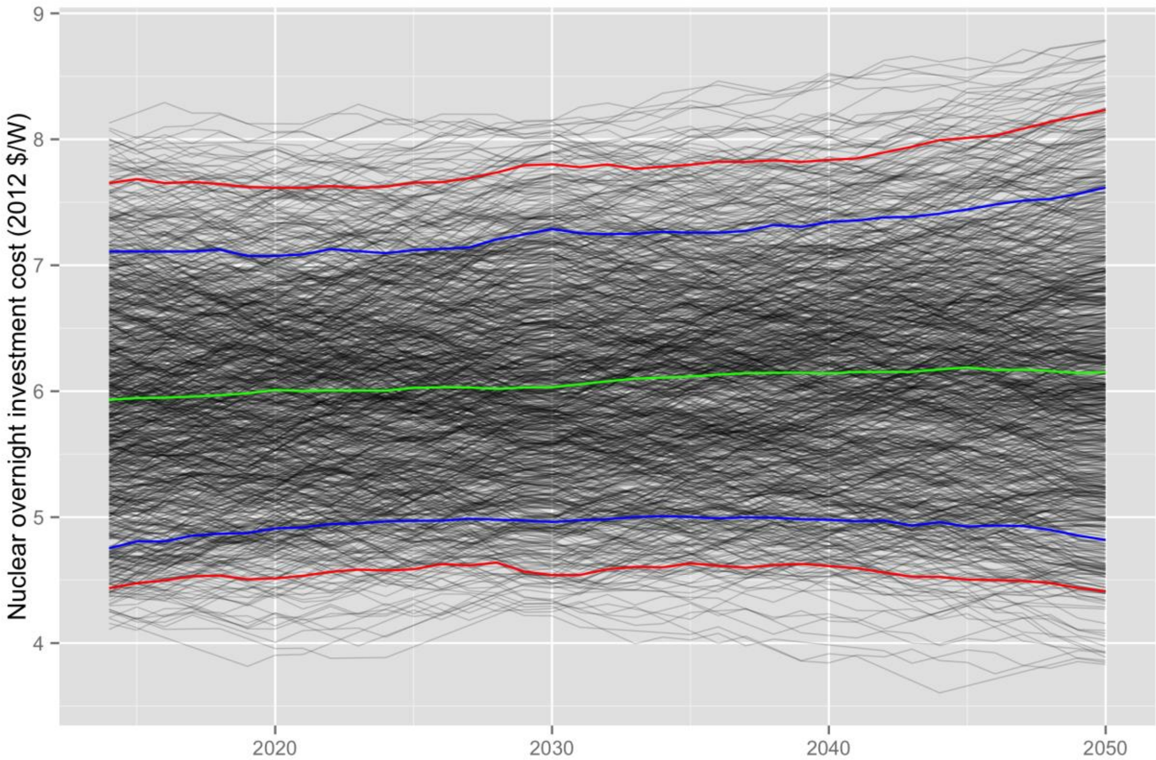

Figure 22: Probabilistic projections of overnight investment costs of a Gen III/III+ nuclear facility over the period 2014 – 2050.

Imported Hydro

The Southern African Power Pool distributes electricity throughout the region via major infrastructure corridors. A number of regional hydro import projects have been identified in the recent IRP (DOE 2011) and IRP update (DOE 2013). Given recent developments around Grand Inga an additional 3.6 GW is considered, parameterised as per (SNEL et al. 2011). The distribution assumed for imported hydro is shown in Figure 23.

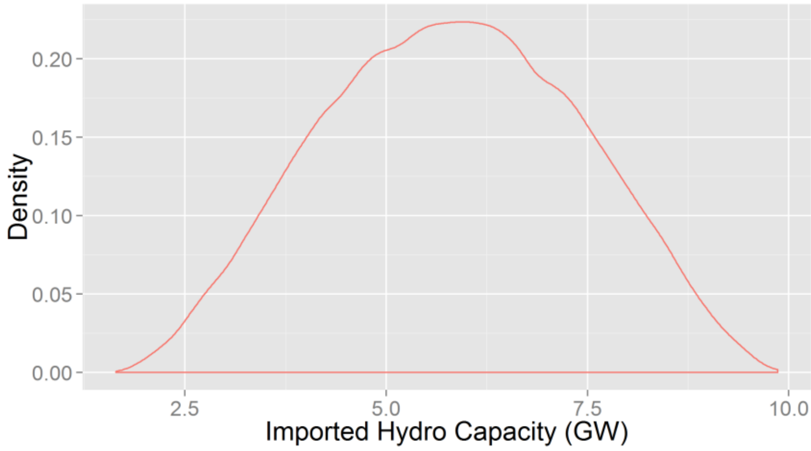

Figure 23 Assumed distribution for imported hydro

Probabilistic projections of baseline GHG emissions

The probabilistic projections described in the previous section are passed as inputs to the SATIM-F energy model. Each combination of input trajectories i.e. consisting of a single sampled trajectory for each of the key input variables, is passed to SATIM and results in an output trajectory for a number of output variables of interest – most importantly GHG emissions but also related quantities such as the proportion of electricity production supplied by each fuel source, electricity and other prices, *etc*.. We show results both under the assumption of perfect foresight (using an 8% discount rate) and myopic foresight (using a ten-year time interval, with five-year overlaps).

GHG emissions

Figure 24 to Figure 26 provide our main results, respectively showing baseline GHG emissions as a raw quantity, in megatons of CO_2 equivalent, as the quantity of emissions per capita, and as the quantity of emissions per unit of GDP.

The median baseline projection is for CO_2 emissions in South Africa to rise slowly to 2030, followed by a period of rapidly increasing emissions from 2030 to the end of the forecasting period, 2050. Between 2015 and 2030, median emissions rise from 420Mt per year in 2010 to 550Mt per year in 2025, falling slightly to around 470Mt in 2030. Between 2030 and 2050 emission levels rise 30% in absolute terms to 650Mt.

While the median baselines may provide a useful single summary, the projections in Figure 24 clearly show the substantial uncertainty around emissions estimates, particularly in the period after 2030. There is relatively little variability between emissions trajectories between now and 2030: in 2020 95% of the trajectories lie between 420Mt and 460Mt, and in 2030 95% of the trajectories lie between 420Mt and 550Mt. Beyond 2030 uncertainty rapidly increases. In 2035 a 95% prediction interval for emissions is for them to lie between 420Mt and 640Mt; this widens still further to between 420Mt and 1000Mt in 2050.

The overall trend is thus for emissions to rise slowly to 2030; thereafter the most likely outcome is for emissions to continue to rise, with substantial uncertainty about the rate of this rise. In most cases, the post-2030 rise in emissions is greater than pre-2030, giving the appearance of an exponential growth in emissions. In a number of simulated trajectories the post-2030 growth in emissions levels is indeed remarkable – note, for example, the higher of the two blue lines in Figure 24, which shows the trajectory such that there is a 10% chance of emissions exceed the values indicated by this trajectory. The general increase, however, is not guaranteed under a baseline scenario. In as many as 10% of our simulated trajectories there is no growth or even a reduction in emissions from 2030.

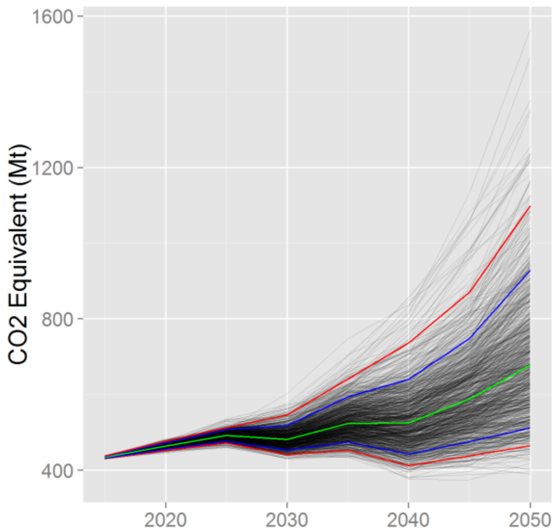

Figure 24: Probabilistic projections of CO_2 equivalent produced by South Africa over the period 2010 – 2050, under perfect foresight.

Per capita emissions are also expected to rise, though less than absolute emissions. Substantial uncertainty exists in the forecasts, particularly beyond 2030. Median per capita emissions remain roughly the same as present-day values of 8t of CO_2 equivalent per capita until 2035, after which they rise steadily to just over 10t per capita in 2050. As with absolute emissions the distribution of per capita emissions is slightly skewed to the right, so that values in the right tail i.e. relatively large per capita emissions, tend to be further from the median, and hence more extreme, than values in the left tail i.e. relatively small per capita emissions. The upper extreme of possible per capita emission would appear to be around 8.4t per capita in 2020, 10t per capita in 2035, and 15t per capita in 2050. Lower extremes are 8t per capita, 7.6t per capita, and 7.6t per capita in 2020, 2035, and 2050 respectively.

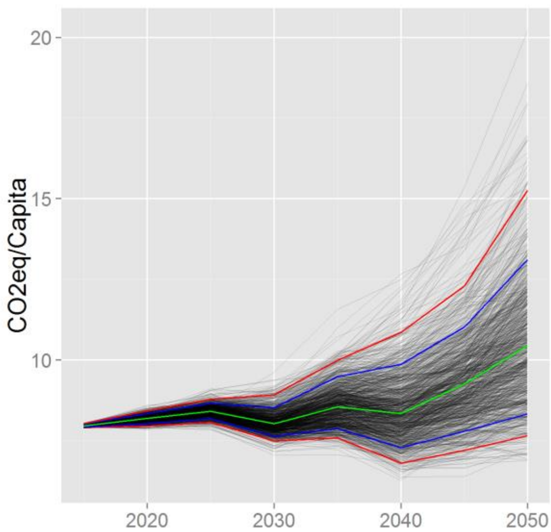

Figure 25: Probabilistic projections of CO_2 equivalent per capita produced by South Africa over the period 2010 – 2050, under perfect foresight.

GHG emissions per unit of GDP fall consistently throughout the forecast period, as a result of both increased activity in the tertiary sector and reduced use of coal for industrial activities. Decreases in CO_2 emissions are close to linear over most of the period, and subject to slightly less uncertainty than the two GHG emission indicators considered above. Nevertheless substantial uncertainty still exists, particularly after 2030. Median forecasts are for CO_2 emissions per dollar of GDP to drop from present levels of 0.66kg/$GDP to 0.46kg/$GDP in 2035 and 0.37kg/$GDP in 2050; 95% prediction intervals around the median projections are (0.65; 0.68) in 2020, (0.42, 0.52) in 2035, and (0.28, 0.48) in 2050.

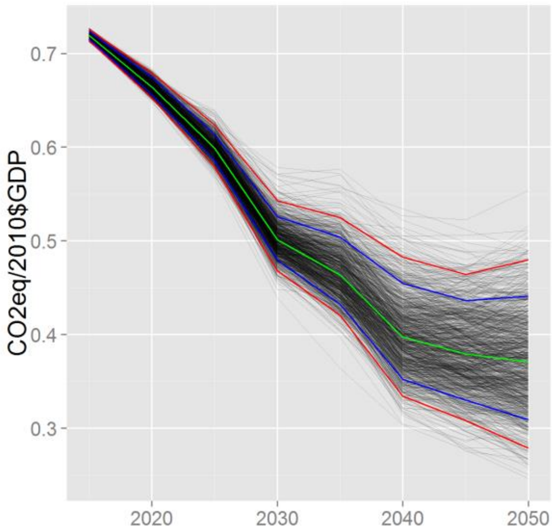

Figure 26: Probabilistic projections of CO_2 equivalent per unit of GDP produced by South Africa over the period 2010 – 2050, under perfect foresight.

Changes in GHG emissions are strongly associated with economic growth. However, the strength of the relationship is perhaps not as strong as might be expected. Figure 27 shows the nature of this relationship in more detail, plotting GDP growth against CO_2 emissions. Conditional distributions of CO_2 at different levels of GDP growth can be examined by taking vertical cross-sections through the scatterplots – these distributions show substantial variability across all levels of GDP growth. For example, at an average GDP growth rate of 3.5%, emissions levels (ignoring outliers) might be anywhere between 450 and 950Mt of CO_2 equivalent in 2050.

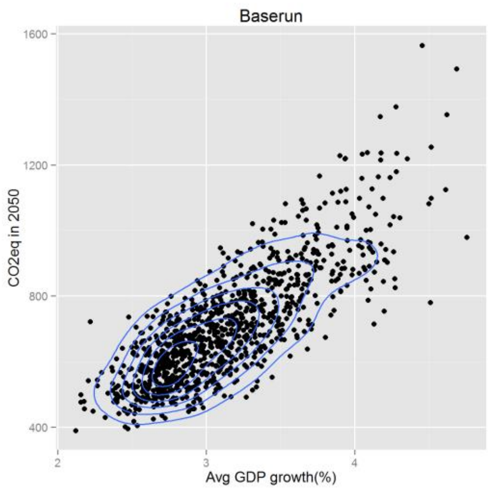

Figure 27: Scatterplots showing relationships between average GDP growth and CO_2 production in 2050, under perfect foresight. Individual points represent pairs of GDP growth/CO_2 values i.e. the values of GDP and CO_2 in a single time period (five years). Contours plot lines of equal probability.

Finally, we show the contributions of different sectors to GHG emissions in 2050. These are of course uncertain, so that Figure 28 shows possible values in the form of probability distributions. Similar figures (not shown here) are available for other points in time. Two noteworthy features of Figure 28 are that the power sector accounts for the majority of emissions, with other sectors contributing smaller but still significant amounts; and that emissions from the power sector, are subject to much more uncertainty than emissions from other sectors.

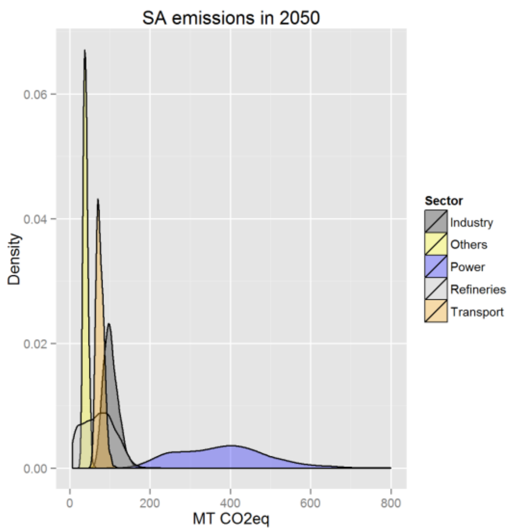

Figure 28: Histograms showing possible distributions of sectoral contributions to overall CO_2 equivalent production in South Africa in 2050, under perfect foresight.

Contributions of fuel types

The previous section has shown that uncertainty around baseline GHG emissions in South Africa is largely due to uncertainties around GHG emissions in the power sector i.e. electricity generation. These uncertainties relate in turn to the relative mix of fuels used to satisfy South

Africa's demand for power. This section describes our results on these aspects – probabilistic projections of the composition of fuels used for electricity production.

Figure 29 to Figure 32 show the projected share of electricity production generated by each of five technologies: coal, gas, nuclear, and solar (PV and CSP technologies).

Figure 29 and Figure 30 show that the primary source of uncertainty in setting a single baseline trajectory for GHG emissions in South Africa is the extent to which gas replaces coal in the production of electricity. Under a median trajectory, coal is expected to remain by far the dominant fuel source for electricity production. Although it declines from its current contribution of 85% to around 70% in the period 2025-2030 from the contribution of regional hydro projects, this latter level is maintained for the remainder of the forecast period. In the median gas trajectory, gas is hardly exploited at all and its contribution to electricity generation remains below 5%. However, enormous uncertainties exist around these median projections.

The most important aspect of this uncertainty is that the distribution of coal's contribution is skewed to the left while the distribution of gas' contribution is skewed to the right. Thus while the median trajectory for coal predicts that it contributes 70% of South Africa's electricity, there is at least a one-in-ten chance that its contribution is less than 25%. Similarly while the median trajectory for gas predicts that its share is below 5%, there is at least a one-in-ten chance that its contribution is above 60%.

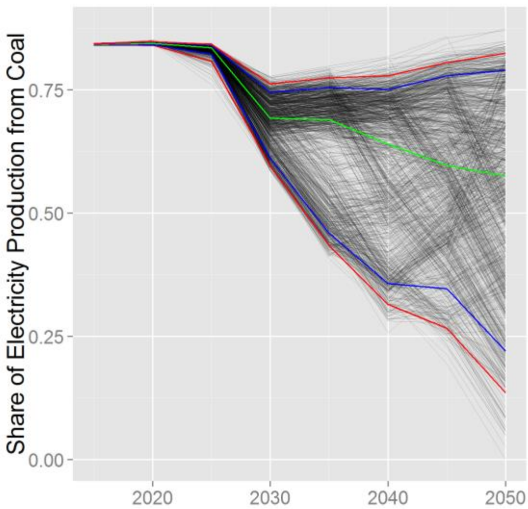

Figure 29: Probabilistic projections of the share of electricity production in South Africa contributed by coal over the period 2010 – 2050, under perfect foresight.

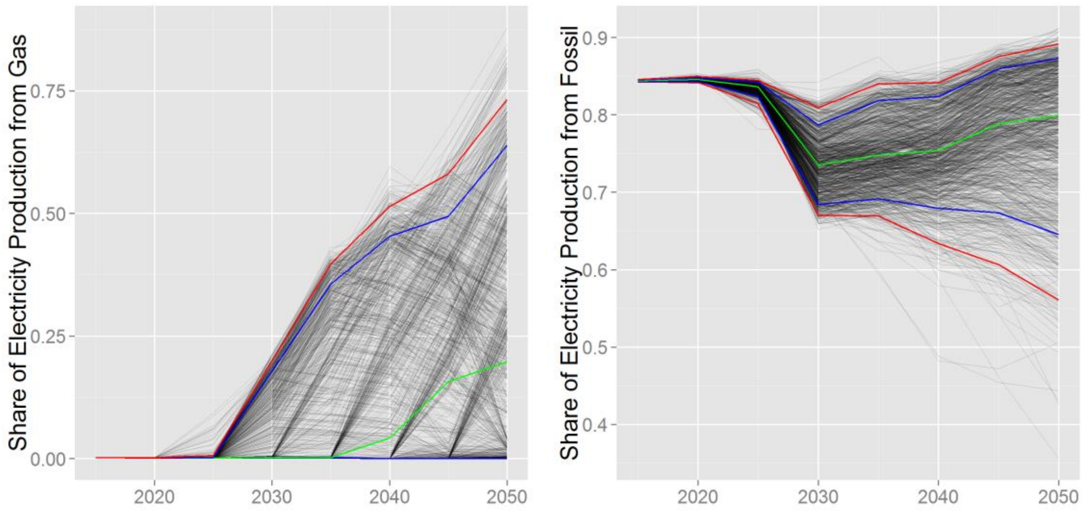

Figure 30: Probabilistic projections of the share of electricity production in South Africa contributed by gas (left) and by fossil: coal + gas (right) over the period 2010 – 2050, under perfect foresight.

The primary uncertainty around shale gas is whether and when it is exploited at all. Most of the trajectories in Figure 30 suggest that if gas is taken up at all it quickly becomes, if not the dominant fuel source for electricity generation, then at least a significant contributor.
The combined coal and gas share is around 70% in 80% of the cases as shown on the right.

Figure 31 shows that under baseline assumptions, PV technologies do not become major contributors to South African electricity production. The median projections are for nuclear to decline as a proportion-of-total, as no new plants are built, and for PV to increase marginally but remain a minor contributor. Relatively little uncertainty exists around these projections.

Concentrated solar power (CSP) in contrast is subject to substantial uncertainty. Under a median projection, it contributes little or nothing to electricity production. However at each modelled time-point from 2030 i.e. each five-year interval from 2030, there is a reasonable chance that the technology becomes a significant contributor from that time-point on. Thus for example there is a 10% chance that CSP is taken up in 2035, and if this happens it increases its contribution of electricity generation to around 10% in 2050. Extreme scenarios see CSP contributing as much as 20% of South Africa's electricity needs, but these should be regarded as extreme (i.e. one-in-a-thousand) events.

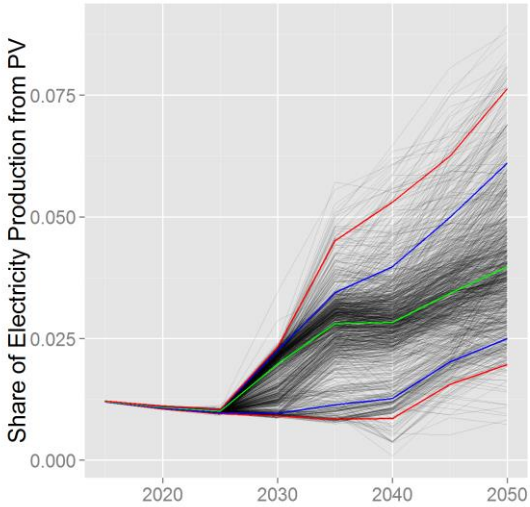

Figure 31: Probabilistic projections of the share of electricity production in South Africa contributed by PV technologies over the period 2010 – 2050, under perfect foresight.

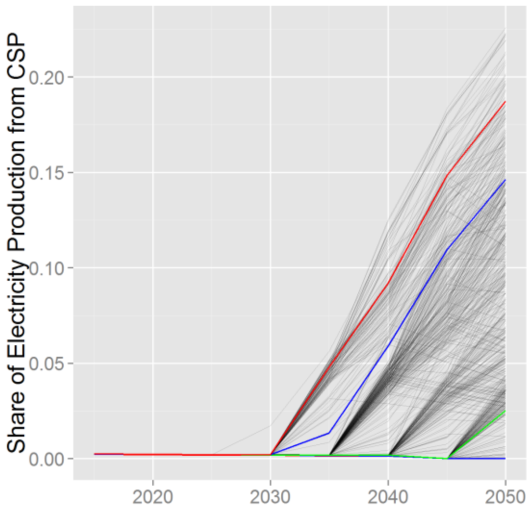

Figure 32: Probabilistic projections of the share of electricity production in South Africa contributed by CSP technologies over the period 2010 – 2050, under perfect foresight.

Alternate Model sensitivity analysis

Myopic model results

Figure 33 shows the results for the base model (perfect foresight) compared to the myopic model (limited foresight) with 10 year periods and 5 year overlaps. The myopic model has only marginally higher emissions compared to the perfect foresight model, and exhibits very similar variance in the emissions in 2050.

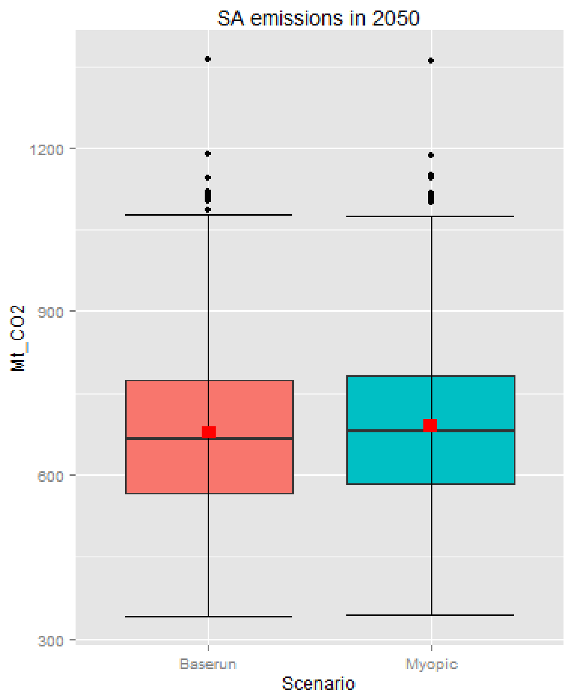

Figure 33: The CO2 emissions box plot for the Baserun model with perfect foresight (left), and the Model with limited foresight (right). Red dots represent the mean.

This slight increase in emissions is the result of slightly higher coal share of electricity production in 2050 as indicated in Figure 34 below.

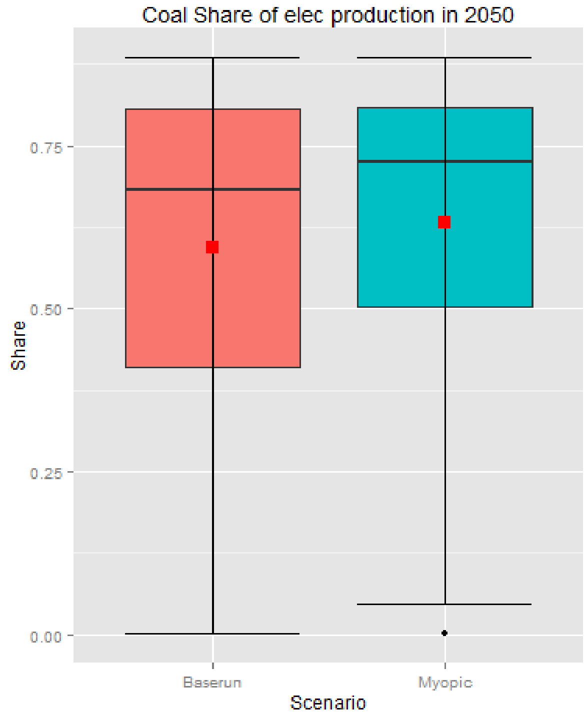

Figure 34: Coal share of electricity production in the perfect foresight model and the myopic (limited foresight) model compared.

The distribution of emissions by each sector is given in the figures below. The results show that there is very little difference in the pattern between the two variations of the model.

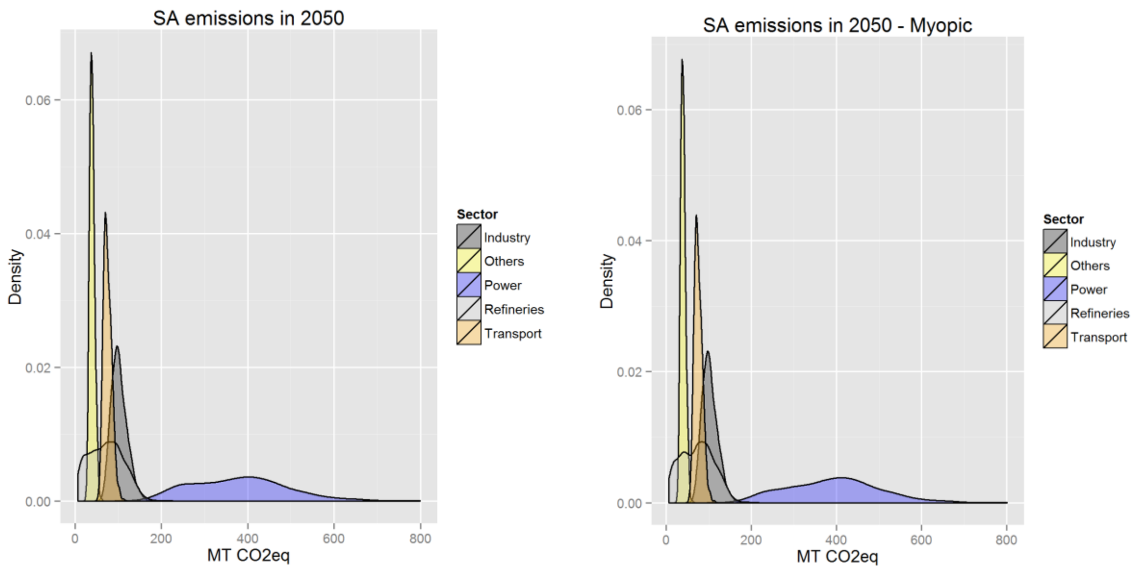

Figure 35: The density plot of emissions by sector for the base model (perfect foresight) in 2050 (left), and for the Myopic model (limited foresight) – right.

Discussion

The results between the perfect foresight and myopic models are very similar. The myopic model exhibited slightly higher emissions compared to the perfect foresight model, as the model invested slightly more in coal power. Also notable is the spread of coal contribution to the electricity sector is smaller than in the perfect foresight model.

However, it should be noted that in general, a myopic model would most likely show larger changes where prices fluctuate widely or where policy interventions result in price changes (such as carbon taxes). In this model the prices of commodities do not exhibit a large enough fluctuation to induce significant changes between the perfect foresight and myopic models. Further work is needed to test the model under policy scenarios which alter commodity prices or other constraints such as a carbon price that changes over time.

Alternate Discount rates

To test the models sensitivity to the discount rate which is set to 8% in the base run, two other alternative discount rates where analysed and compared: a low discount rate scenario of 5%, and a scenario for a higher rate of 11%.

Discount rate results

Figure 36 shows the total CO2 emissions in 2050 for the low and high discount rates compared to the original base run model.

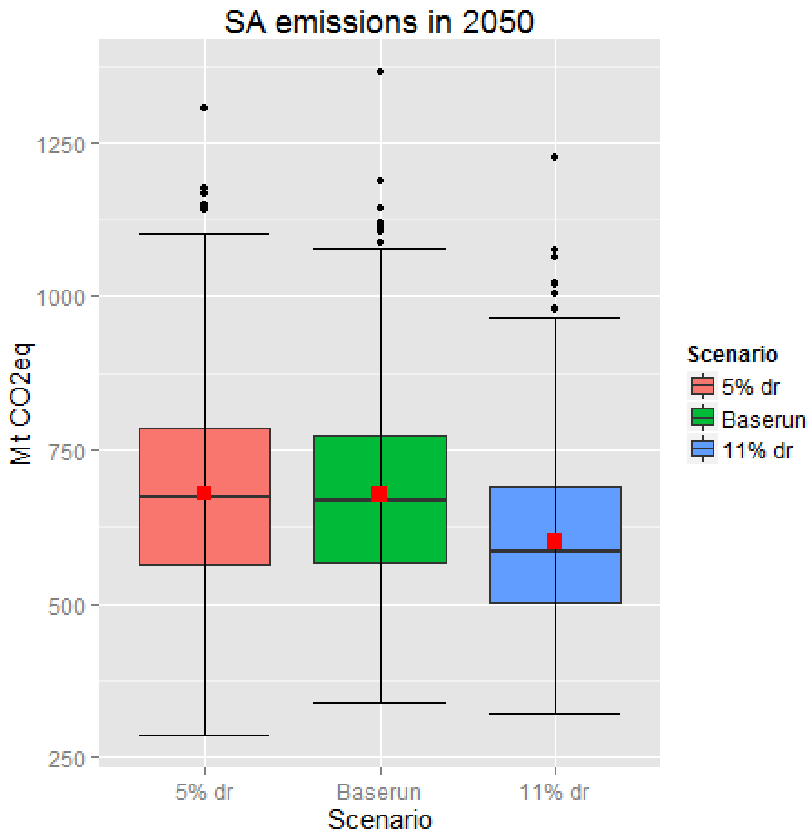

Figure 36: Comparison of CO2 emissions in 2050 for the discount rate sensitivity analysis

Noticeably the higher discount rate of 11% has generally lower emissions in CO2, despite the fossil fuel share of electricity production being higher than the other two scenarios as indicated in Figure 37. The lower emissions despite higher fossil fuel share in electricity production is a result of a higher use of gas in the model as indicated in plot 2 of Figure 37 below.

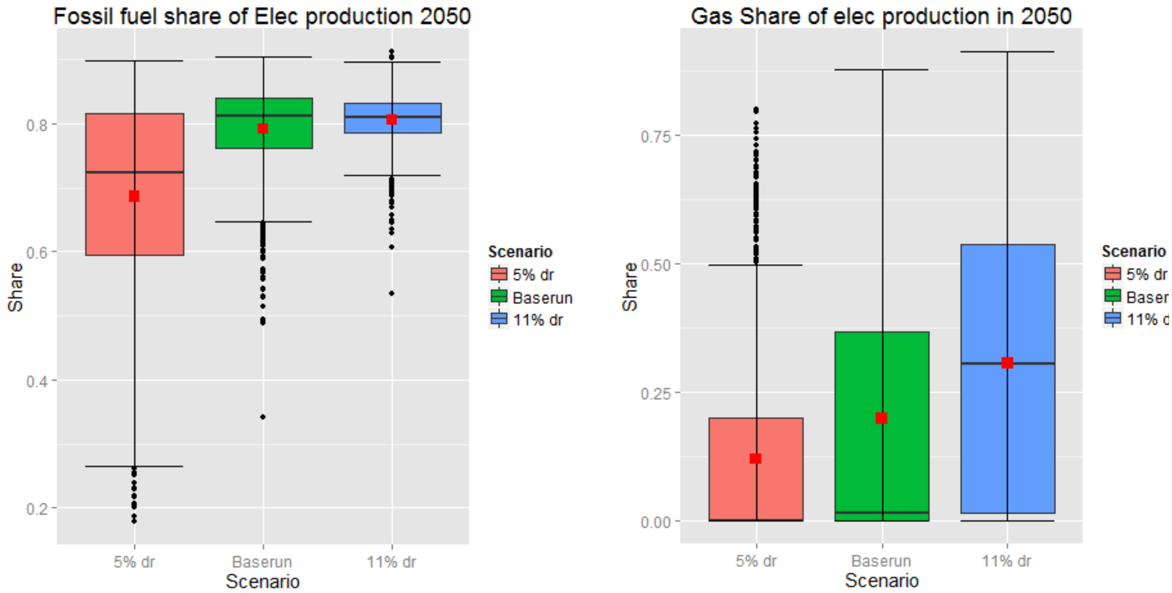

Figure 37: The share of fossil fuels (left) and gas (right) in electricity for the model using 5%, baserun 8%, and 11% (red, green, blue plots).

The increased use of gas to generate electricity at the higher discount rate is a result of both Shale gas and LNG competing well with coal as indicated in LCOE's in Figure 38 below.

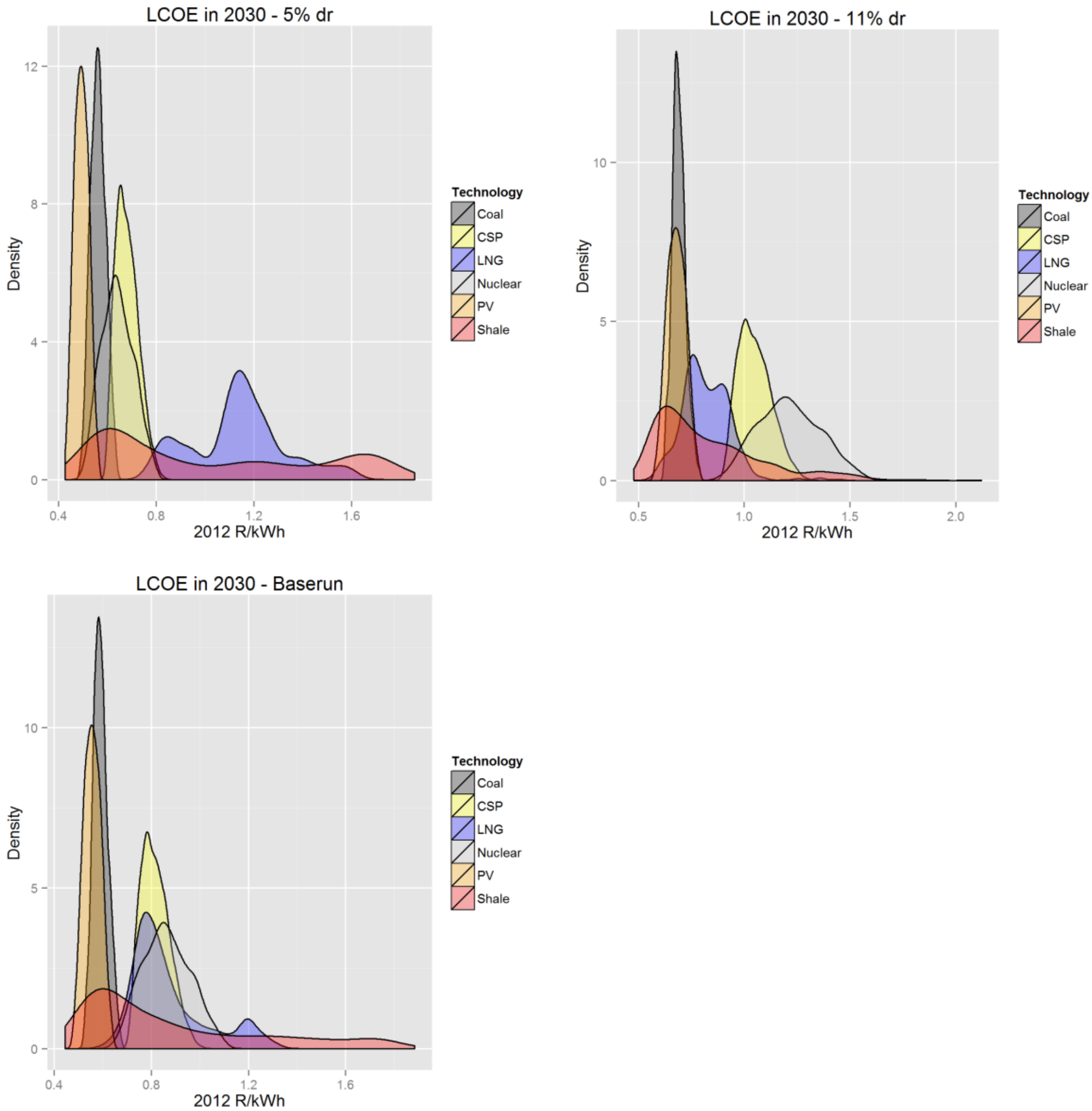

Figure 38: The density plots of the LCOE's for each of the main electricity generating technologies in the year 2030. Top left is 5% discount rate model, top right is the 11% discount rate and bottom figure is the baserun model (8% discount) model. Note that density plots have an area of 1 for each technology, and please note the scales on the y axis are not uniform between these figures.

Also of interest is the emissions by sector in 2050, presented in Figure 39. Of note is that refineries are most likely to drop out almost entirely from the contributions to emissions by 2050 in a higher discount rate scenario.

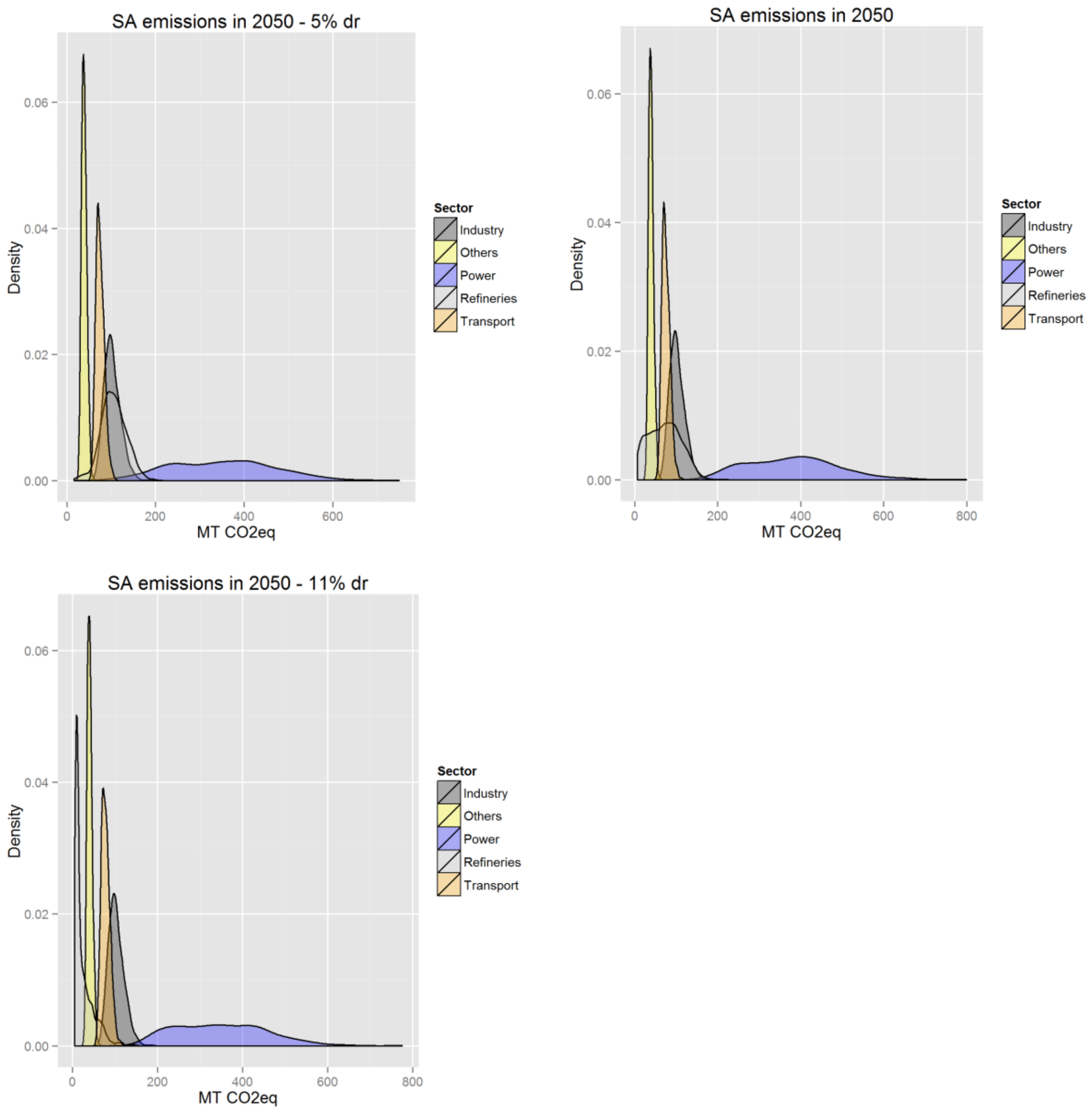

Figure 39: The density plots of emissions contributions by sector for South Africa in 2050, in the 5% discount rate model top left, the baserun (8% discount rate) in top right, and 11% discount rate model bottom.

The 11% discount rate has effect of lowering the CO_2 emissions in the country as a result of more use of gas power and refineries contributing very little toward CO_2 emissions as compared to the baserun model of 8% discount rate.

A lower discount rate of 5% lowers the overall share of fossil fuels in the electricity generation system by 2050 and increases the spread for fossil fuel share of the power sector. While the contribution to emissions from the power sector is generally lower, the contribution to emissions from refineries is higher as indicated in Figure 39.

Discussion and conclusions

Summary of approach

The objective of the project on which this report in based is to construct a baseline projection for GHG emissions in South Africa to 2050, taking into account the inevitable uncertainty that must accompany such projections. To this end the following approach was followed:

- We base our approach on the South African TIMES model (SATIM), a partial equilibrium linear optimisation model that selects a mix of energy sources to meet a given demand for useful energy at least cost. GHG emissions are obtained as output of the optimisation model. In doing so, we abstract the task of assessing uncertainty about GHG emissions into the easier tasks of assessing uncertainty about (a) energy demand and (b) which fuels are used to meet this demand.
- Uncertainty about energy demand is in turn decomposed into uncertainty about various determining factors, specifically population growth, economic growth, and differing growth rates across economic sectors. Uncertainty about the fuel mix used to meet energy demand is decomposed into uncertainty about the prices of the various fuel sources (e.g. coal and gas) and the costs of energy technologies (e.g. renewables in the form of PV and CSP).
- We assess uncertainty about each of the eight key input variables using a combination of (a) a review of the literature, (b) elicitation from national experts, (c) further modelling. External sources are used for quantities such as population growth, international commodity prices, and local fuel prices that are expected to remain tightly coupled to international prices (e.g. nuclear), for which comprehensive probabilistic forecasts already exist and are widely used. Where these are unavailable (e.g. for local economic growth and local commodity prices) we obtained probabilistic forecasts from a small number of local experts, using elicitation procedures drawn from established best practices.
- In cases where quantities were assessed by expert elicitation, a degree of additional modelling is required to bring these into a form suitable to be used as inputs to SATIM. Primarily this involved interpolation between key time-points used in the elicitation, since eliciting full projections from experts was not possible.
- Probabilistic inputs are passed to SATIM. Each combination of input trajectories results, deterministically, in a set of trajectories for each output of interest: primarily GHG emissions but also related quantities such as how those emissions are distributed between sectors and electricity prices. The approach we follow is a Monte Carlo simulation. Taken as a whole, the set of 1000 possible input trajectories results in a set of 1000 possible output trajectories, from which distributional outputs can easily be obtained.
- An alternate model formulation is explored through the use of the optimisation using myopic foresight instead of perfect foresight, and different discount rates.

The primary limitations of our approach are the following:

- *SATIM dependency*: our approach depends heavily on the underlying SATIM model, in terms of how energy inputs are linked to energy outputs. The model has been developed over a number of years specifically for the South African context, and is perhaps the most comprehensive model of national energy production and consumption available at the current time. Nevertheless the model assumes that fuels are selected so as to minimize cost, which may not reflect the complexities of decisions taken in a world in which other socio-political pressures exist, especially on the demand side. Fundamental or extreme changes to the system from climate impacts e.g. dramatic rises in sea level, are not taken into account by SATIM.
- *Independence of input variables*: Although autocorrelation within each key input variable is modelled explicitly, correlations between input variables, except in the case of the international fuel prices for coal gas and oil, and in the case of population and GDP, are assumed to be zero. That is, we sample independently when constructing combinations of input variables. The difficulty in this regard is simply finding experts with sufficient knowledge to assess these correlations. Experts exist with subject areas, but the assessment of inter-variable correlations requires an extremely broad and deep knowledge, encompassing all the input variables. The exceptions here are international commodity prices (coal, gas, oil), for which correlational information is available.
- *Biases in human judgment*: Although highlighted several times, it is worth repeating that long-term forecasts are notoriously fallible, and susceptible in particular to biases anchoring these forecasts to the current status quo. Quite simply, in many applications of long-terms the observed reality turns out to be well outside of expected bounds. Thus, any long-term forecast must be interpreted and used with caution.

Summary of key results

Our primary results are as follows:

1. Most baseline projections of CO_2 emissions in South Africa rise slowly to 2030, followed by a period of more rapid increase of emissions from 2030 to the end of the forecasting period, 2050. Enormous uncertainty exists around the precise quantity of emissions, however, particularly after 2030. Our results indicate that 95% of trajectories lie between 445Mt and 475Mt CO_2 equivalent in 2020; between 415Mt and 635Mt in 2035; and between 420Mt and 1000Mt in 2050. The median projection is for emissions of CO_2 equivalent to rise from 420Mt per year in 2010 to 500Mt per year in 2035 and 670Mt in 2050. Full projections are given in Figure 24.
2. Our results show that a no climate policy scenario has wide ranges of GHG emissions, but with median projections rising throughout, but moderately: from 420Mt CO2-eq per year in 2010 to 500Mt per year in 2035 and 670Mt in 2050. Median projections should be interpreted cautiously. For example, the median projection of 500 Mt CO2-eq in 2035 is within the 'peak, plateau and decline' (PPD) trajectory range in national climate policy (RSA 2011), which is PPD range is 398 to 614 Mt CO2-eq for 2035. The range in our modelling projects GHG emissions from 415Mt and 635Mt in 2035. We emphasise the high level of uncertainty in absolute emission projections, especially further into the future, after 2030. It is more advisable to consider ranges, than the median values.

3. Per capita emissions are also expected to rise, though by less than absolute emissions. Substantial uncertainty again exists in the forecasts, particularly beyond 2030. Our results indicate that 95% of trajectories lie between 8t and 9t per capita in 2020; between 7t and 11t per capita in 2035; and between 6.5t and 17t per capita in 2050. Median per capita emissions remain roughly the same as present-day values of 8.5t per capita until 2035, after which they rise steadily to just over 10t per capita in 2050. Full projections are given in Figure 25.

4. Emissions intensity, that is GHG emissions per unit of GDP, falls consistently and approximately linearly throughout the forecast period. Nevertheless substantial uncertainty still exists, particularly after 2030. Our results indicate that 95% of trajectories lie between 0.66 and 0.69kg/$GDP in 2020; between 0.39 and 0.55kg/$GDP in 2035; and between 0.26 and 0.47kg/$GDP in 2050. Median forecasts are for CO2 emissions per dollar of GDP to drop from current day levels of 0.67kg/$GDP to 0.47kg/$GDP in 2035 and 0.38kg/$GDP in 2050. Full projections are given in Figure 26.

5. Uncertainty around baseline GHG emissions in South Africa is largely due to uncertainties around GHG emissions in the power sector i.e. electricity generation. In nearly all projections, electricity production accounts for the majority of GHG emissions, but the precise quantity of emissions is subject to enormous uncertainty, substantially more than emissions in any other sectors. These results are shown in Figure 28.

6. Uncertainty around GHG emissions due to electricity production relate in turn to the relative mix of fuels used to satisfy South Africa's demand for power. The primary uncertainty is the extent to which gas replaces coal in the production of electricity. Our results show that 95% of trajectories indicate that coal contributes between 45% and 75% and gas between 0% and 38% of electricity produced in 2035, from their current shares of 85% and 0% respectively, and that coal contributes between 15% and 85% and gas between 0% and 75% in 2050. That is, almost anything can happen: although unlikely, gas may almost entirely usurp coal as the main source of South Africa's electricity. Median coal shares decline from 85% in 2020 to 70% in 2030 and remain at this level to 2050. Median gas shares remain near zero throughout the forecast period. Full projections are given in Figure 29 and Figure 30.

7. Nuclear and PV technologies do not become major contributors to South African electricity production. Most projections are for nuclear to decline as a proportion-of-total, as no new plants are built, and for PV to increase marginally but remain a minor contributor. Concentrated solar power (CSP) is subject to substantial uncertainty. Our results show that 95% of trajectories indicate that CSP contributes between 0% and 5% of electricity produced in 2035, and between 0% and 15% in 2050, although the median projection is for CSP to contribute little or nothing (less than 1%) throughout the forecast period. Full projections are given in, Figure 31 and Figure 32.

8. The uncertainty in the baseline seems robust to the different model formulations, in myopic vs perfect foresight and the different discount rates, where little variation is observed.

Policy implications

Baseline forecasts play an important role in strategic planning around responses to climate change, providing inputs into discussions around fair allocations among countries and responsibility for mitigation actions. Drawing out the detailed policy implications of the baseline projections provided here is beyond the scope of the current project, but it is not difficult to see the challenges that South Africa faces in this regard. South Africa is in a fairly unusual situation: it depends heavily on coal for power generation, and under "business as usual" policies such as assumed here, this dependency is projected to continue for some time, perhaps to 2050. At the same time it has a population that is projected to grow substantially in size, coupled with relatively modest economic growth. Improving the living conditions of a substantial proportion of the population is likely to be a challenge, even under a "business as usual" dependency on coal.

Further research

The current project is far from a definitive statement of South Africa's baseline projection for GHG emissions, rather it should be seen as a first step along this process. The following areas are perhaps the most effective areas to direct future efforts:

- Our results show that uncertainty increases exponentially over time, and that for some key variables (e.g. the relative mix of coal and gas for electricity production) almost anything can happen over a period of decades. Median projections can be calculated, but are no prediction of the future. We stress again that this uncertainty is "baseline" uncertainty i.e. under the assumption of relatively unchanging policies. Assessments incorporating policy uncertainty will be even more variable. As a result, baselines need to be regularly updated, perhaps at intervals of no more than five years.

- Related to the above point, our forecasts of GHG emissions are based on forecasts of key input variables, obtained from a variety of sources. These too will change over time, and updates should be incorporated into the baseline forecasts when information on changes in input variables become available. This also motivates for the regular updating of baseline projections.

- The uncertainty in CO_2 per GDP is much narrower than uncertainty regarding the absolute level of CO2 emissions. The reason for the narrower range is that the variability caused by different GDP growth scenarios is partly taken away. This might make emissions intensity an attractive metric for mitigation commitments. The absolute result in future GHG emissions would still, however, be subject to uncertainty about the GDP projection assumed at the time.

- Our projections are based in places on the assessments obtained from only a small number of experts. Eliciting information from a greater number of experts would provide a greater degree of confidence in the results.

- Our results indicate that perhaps the key uncertainty in the setting of baseline GHG emissions in South Africa is the relative price of coal to gas. This is due to the large share of emissions from electricity generation. Particular emphasis could be placed on modelling these two quantities.

- The uncertainty explored focuses on supply technologies and fuel prices, but uncertainties also exist in the future cost and performance of demand technologies such as advanced air-conditioning and electric cars, as well as the uncertainty in the costs of the supporting distribution infrastructure required for the mass uptake of new fuels for South Africa, such as natural gas in the transport, residential and commercial sectors and electricity in the transport sector.

- The tool developed in this project was only used to look at a no climate policy scenario. However, with very little further modification it could be used to look at a whole host of climate policies ranging from CO2 prices, or "no more coal power" policies, or even looking for other robust climate policies, which would be very interesting to do.

- More detailed sensitivity analysis could be carried out to determine the more sensitive parameters and where more effort could be put to try and reduce if possible the uncertainty.

- As mentioned, our model assumes independence between most of the key input variables. Further work might explore the effect that correlations between input variables might have on results, as well as effective ways to measure these complex correlations.

Bibliography

Alkema, L., Raftery, A., Gerland, P., Clark, S., Pelletier, F., Buettner, T., et al. (2011). Probabilistic projections of the total fertility rate for all countries. *Demography, 48*, 815-839.

Anadon, L., Nemet, G., & Verdolini, E. (2013). The future costs of nuclear power using multiple expert elicitations: effects of R&D and elicitation design. *Environmental Research Letters, 8*(3), 034020.

Clapp, C., & Prag, A. (2012). *Projecting emissions baselines for national climate policy: Options for guidance to improve transparency.* OECD. Paris: OECD/IEA.

Morgan, M., & Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge: Cambridge University Press.

Raftery, A., Li, N., Sevcikova, H., Gerland, P., & Heilig, G. (2012). Bayesian probabilistic population projection for all countries. *Proceedings of the National Academy of Sciences*, 13915-13921.

Rozenberg, J., Hallegatte, S., Vogt-Schilb, A., Sassi, O., Guivarch, C., Waisman, H., et al. (2010). Climate policies as a hedge against the uncertainty on future oil supply. *Climatic Change, 101*(3), 663-668.

Sassi, O., Crassous, R., Hourcade, J.-C., Gitz, V., Waisman, H., & Guivarch, C. (2010). Imaclim-R: a modelling framework to simulate sustainable development pathways. *International Journal of Global Environmental Issues, 10*(1), 5-24.

SNEL, AECOM and EDF. 2011. Etude de Dévelopement du site Hydroélectrique d'Inga et des Interconnexions életriques associées [Study of the Development of Hydroelectric Sites of Grand Inga and Associated Electrical Sites]. Kinshasa, Democratic Republic of Congo: Société Nationale d'Electricité, AECOM and Electricité de France.

Keppo, I. & Strubegger, M., 2009, *Implications of limited foresight and sequential decision making for long-term energy system planning: an application of the myopic MESSAGE model*, International Institute for Applied Systems Analysis, Luxenburg, Austria.

F. Hedenus, D. Johansson, K. Lindergren, *A critical Assessment of Energy-Economy-Climate Models for policy Analysis*, Journal of Applied economic and Business research, 2013

Republic of South Africa 2011. National Climate Change Response White Paper. Government Gazette No. 34695, Notice 757 of 2011. Pretoria, Department of Environmental Affairs. http://www.greengazette.co.za/documents/national-gazette-34695-of-19-october-2011-vol-556_20111019-GGN-34695.pdf

Appendices

Pre-elicitation documentation

Introduction to probability

As this project will be assessing your knowledge or beliefs in a probabilistic form, we should first spend a moment defining what a probability is. Simply put,

The probability of an event is a measure of how likely it is to occur

A probability of 0 means that the event is absolutely certain not to occur, while a probability of 1 means that the event will occur with absolute certainty. The central value, 0.5, means that the event is as likely to occur than not to occur.

A fundamental law of probability is that sum of the probabilities across all distinct outcomes must equal 1. By "distinct" we mean all outcomes that cannot occur together – the term for this is *mutually exclusive.* In the case of an event, either the event does occur or it does not. These are the only two possible outcomes, and their probabilities must therefore sum to 1.

Often, we deal with quantities that can take on a range of values, rather than a single event. Whereas an event either does or does not occur, a quantity can take on any of a range of possible values.

An uncertain quantity is called a random variable

If the random variable can take on any value within some range, it is called *continuous*. Take, for example, the total weight of coal reserves that remain on Earth. This quantity could in theory take on any positive value, although clearly it would be very many tons. The South African population, on the other hand, can only take on certain distinct values (any whole number), but cannot take on values between these. Random variables like this are called *discrete*.

For continuous random variables, a fact that surprises many people encountering probability for the first time is that the probability associated with any one particular value is zero. This is because there are in theory an infinite number of possible values that could occur, so that the probability of any single outcome occurring must be vanishingly small – zero.

Although we cannot express the probability that a random variable takes on a single value, we can assess the probability that it lies between two distinct values. Often, it is easier to just assess the probability that the random variable is less than some value (which we label x), for a number of possible x values, and to work out the probability for any interval by doing some simple calculations. For example, if you think that there's a probability of 0.7 that your friend weighs less than 70kg, and a probability of 0.3 that they weigh less than 60kg, then clearly the probability that they weigh between 60kg and 70kg is 0.7 – 0.4 = 0.3.

Mathematically, we write the statement that "the random variable is less than some value x" as $\Pr[X \leq x]$. Random variables are usually denoted with capital letters (so X is a placeholder for the random variable we are interested in), while lowercase letters indicate specific values that the random variable can take on. In fact, the difference between "event-based" and "interval-based" probabilities is not as big as it may seem. We can easily think of the outcome $X \leq x$ as an "event" – either the random variable is less than or equal to x, or it is greater than x.

If we assess $\Pr[X \leq x]$ over all possible values of x, then we end up with a function called the cumulative distribution function (or CDF) of X.

The cumulative distribution function or CDF is a critical piece of information that describes, in full detail, what values are more or less likely to occur.

In practice, it will be impossible to assess the CDF of X for all values of x, so we usually do the assessment for a few x values and make some assumptions about what happens between these values. This is in fact what will happen during the elicitation process you will be guided through.

It is now time to take a closer look at uncertainty, and the kinds of uncertainties that are amenable to modelling with probabilities. We'll mainly talk about two different interpretations of probability: one called frequency probability, the other personal (or subjective) probability.

The frequency definition of probability says that the probability of an event is the proportion of times it occurs in a long sequence of repeated trials

Most introductions to probability start with a frequency interpretation of probability, which says that the probability of an event is the proportion of times that it occurs in a long sequence of repeated trials. For example, the probability of a coin landing heads-up is 0.5, because that is the proportion we would expect to observe if we flipped the coin many times.

This focus on frequency probabilities is unfortunate, because it excludes much of the uncertainty that we experience in daily life. Often, we feel uncertain not because we are faced with a fundamentally random process like a coin flip that can be repeated many times. We feel uncertain because we face a once-off, unrepeatable situation about which we have limited knowledge. This process may not even be random; the answer may be potentially knowable, at least in theory. It is just that, in our current state, we do not know what that answer is.

In cases like these, the frequency interpretation of probability breaks down, and if we want to make probability statements at all (which we do) we must use a different interpretation of probability called personal or subjective probability.

The personal definition of probability says that a probability of an event represents someone's degree of belief that the event will occur.

In this interpretation probability represents someone's degree of belief in an uncertain statement – like the occurrence of an event. The same basic rules apply – probabilities must be between zero and one, must sum to one, and so on – but we are allowed to make reference to a broader class of events, including once-off events for which repeated experimentation is not possible.

Many people who have been exposed to some probability training are initially uncomfortable making personal probability judgments.

This is particularly true of scientists, for whom "subjectivity" is often viewed negatively, with associations of being insensitive to evidence and sensitive to irrelevant emotions through hidden biases and prejudices. While such feelings can be difficult to set aside, they are based on a rather limited view of judgment in general, and elicitation in particular. There are a number of things to bear in mind when participating in an elicitation process:

There is no objectively "correct" answer to almost any elicitation question

Often, the people asked to give their opinions experience discomfort because they do not know what will happen – they don't know the "correct" answer. We wish to be clear that your opinion has been sought because there is *no* single accepted answer to the question asked of you. Any answer is necessarily *yours* – although it can and should be informed by the available evidence, it is ultimately *your* opinion that counts. The range of answers that you give, and the probabilities that you associate with them, should reflect *your* degree of uncertainty in the outcome you have been asked to assess.

Your opinion may change during the elicitation process

There is a broad consensus from research on human judgement that elicited probabilities are at least partly constructed during the elicitation process. Probabilities, like many other kinds of judgements, are not pre-formed quantities that exist in someone's mind waiting to be "read off" by a trained assessor. The way in which questions are asked and answered can and does change the answers given. We will discuss these and other findings from research on the psychology of human judgement in a companion document. We stress once again that there is no objectively "correct" answer to almost any elicitation question, and that your opinion may and should change as you consider and weigh up different sources of evidence.

The objective of good elicitation is to eradicate negative biases and to assist you in a rational assessment of your own knowledge and experience

The fact that there is no "correct" answer doesn't mean that any answer is equally good – a common criticism of "subjectivity". Indeed there is a clear consensus about what makes probability judgements "good", the most important of which are that they should accurately reflect the true beliefs of the person being assessed while remaining consistent with the basic laws of probability. There is also good agreement on what kinds of assessment procedures have the best chance of achieving these aims, as we also describe in an accompanying document. The assessment process we follow closely follows the best available practices.

Practical elicitation nearly always involves variables that are uncertain because of limited knowledge rather than fundamental randomness

Finally, the use of personal probability is an unavoidable aspect of the kind of work we are doing. The very fact that peoples' opinions are being sought almost guarantees that the process in question is uncertain at least partly because knowledge is limited rather than because the process is fundamentally random. It would not make much sense, for example, to consult an "expert" about a coin flip. Although putting a number on your judgments may be uncomfortable, the alternative is to leave these judgements vague and unspecified. This makes them far more open to manipulation and far less likely to be useful inputs to later policy discussions and decision-making.

Introduction to thinking about probability

This document gives a short, informal summary of what is currently known about how people go about thinking about uncertainty, and in particular how they make probability judgements. By probability judgements we mean informal statements about how likely or unlikely various types of events are, as well as the kinds of explicit judgments about numerical probabilities that you will make later on. Our aim in this summary is to highlight some common pitfalls that people fall into when thinking about probabilities. As we'll discuss, knowing about these pitfalls can help to reduce them (but, unfortunately, rarely to avoid them completely!).

We hope that by reading this document the probabilities you end up giving will accurately reflect your true beliefs.

Up until the mid-1960's the general view was that people were relatively good at translating their personal experiences and observations into probability judgements. When asked to estimate descriptive statistics like an average, probability, or proportion, people generally did this with reasonable accuracy. But within a decade this view had been almost entirely shattered and replaced by a very different one. Using a series of simple but extremely convincing experiments, researchers found that

When people judge probabilities or estimate uncertain quantities, they employ a variety of simple strategies that can sometimes lead to systematic errors

A number of points are worth making early on here. Firstly, just because judgements *can* be flawed does not imply that they always *will* be flawed. Indeed, much of the subsequent research on judgements has focused on identifying those conditions that differentiate good and bad judgements, and this research forms the backbone of good elicitation practice. Second, the kinds of errors we talk about are "systematic", meaning that they generally operate in a single direction, rather than "random" errors scattered around a point. This is important, because while random error is largely something we just have to live with, systematic errors can sometimes be avoided or corrected for once they have been identified. And thirdly, "simple strategies" should not be viewed negatively. We all operate with limited time and information. Our capacity for processing information is also limited, compared to say what a computer can do. We have thus evolved a number of short-cut or approximate strategies, called *heuristics*, for judging the nature of a risk.

Heuristics are efficient – they can operate with limited time and information, and can usually be relied upon to give an answer that is "good enough", even if it is not the best that could be found with unlimited resources.

Heuristics are in fact rather remarkable. They allow us to operate in a highly-pressurized world, and to largely make the right judgments and decisions. But sometimes, they get it wrong.

The systematic errors associated with the use of heuristics are called "biases"

Most of the early research into probability judgment involved identifying those situations in which errors do arise. This work was largely the product of two psychologists, Amos Tversky and Daniel Kahneman. They used a series of simple but highly convincing experiments to demonstrate a number of biases, and proposed heuristics to explain these biases. Their work came to known as the heuristics-and-biases research program, and it has since become enormously influential in psychology, economics, and business. In 2002 Kahneman was awarded the Nobel Prize in Economics (Tversky had passed away in 1996, and the Nobel prize is not awarded posthumously) "for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty". In the rest of this document we review some of the key messages that have emerged from the heuristics-and-biases research.

When people make intuitive probability judgments, they often rely on a combination of three heuristics, called availability, anchoring-and-adjustment, and representativeness. We discuss each in turn.

People often judge the probability of an event by how easily specific instances of that event come to mind. This is called judgement by "availability".

A bit of thought reveals both why this heuristic is generally good, and why it can fail. Firstly, it will usually be easier to recall events that occur more often in reality than ones that occur only occasionally. This supports the use of the heuristic. But there are also elements that affect how easily an event comes to mind, and these may be entirely unrelated to the frequency with which it occurs. Events that have occurred recently, or that are particularly evocative or attention-grabbing, tend to be more memorable, and therefore judged as more probable, than they really are. For example, risks associated with shark attacks or terrorist bombings tend to be overestimated, while more "mundane" risks like flu or car accidents tend to be underestimated (outside of the holiday season, when the risk of car travel is probably overestimated). When judging the likelihood of an event, it is important to ask yourself if your estimate is being swayed by features, like the vividness or recency of any information you have gathered, that should not influence a probability judgement.

People often make estimates by quickly creating or using an initial "guess" and then adjusting this to get a final answer. This is called judgment by "anchoring and adjustment".

Again, this is a seemingly reasonable strategy when time and cognitive abilities are limited. An initial guess is often relatively easy to make, and can be suggested either by how the problem is stated, or by the result of a quick, partial computation. Some more careful thought can then go into adjusting the initial guess upward or downward. However, a number of studies have shown that the formation of the initial guess (called the "anchor") can be quite easily and strongly manipulated, even by clearly arbitrary actions. Compounding this problem, when people adjust away from the initial anchor, they usually adjust by too little, so that the final estimate is biased towards (is too close to) the anchor.

Some examples of biases arising from the anchoring-and-adjustment heuristic have become justifiably famous. In one experiment, two groups of people saw a roulette wheel spin, and then were asked whether the percentage of United Nation members that were African nations was greater or less than the number that came up on the wheel. They then were asked to give an exact estimate of this percentage. The experimental "trick" was that the roulette wheel was manipulated to land on number 10 for one group, and number 65 for the other. The group who saw the wheel land on 10 gave a mean estimate of 25%; the groups who saw the wheel land of 65 gave a mean estimate of 45%! In another study, one group was asked whether Mahatma Gandhi died before or after age 9, or another group whether he died before or after age 140. After this the groups estimated the age at which he died. The mean estimate in the first group was age 50; in the second group it was 67! Even though the anchors are clearly random in the first example, and clearly incorrect in the second, they still strongly influenced peoples' judgments.

Probably the most harmful effect of anchoring-and-adjustment on elicitation is that when people are asked to give range estimates that cover what is likely to occur, these ranges that they give tend to be too narrow.

Many studies have documented this phenomenon. For example, when asked for 90% ranges (so that the "correct" answer should lie outside the given range in only 10% of cases), the expressed ranges covered the correct answer in only 57% of cases. When asked for 99.9% ranges, the expressed ranges covered the correct answer in only 85% of cases. This effect, called "overconfidence", happens because people first think of a central, moderate, or most likely value. They then adjust this up and down, but (as before) insufficiently, giving range estimates that are too narrow.

Completely avoiding anchoring-and-adjustment is nearly impossible, but when making probability assessments, it is vital to ask yourself critically whether the heuristic is having an undue influence on your estimate, especially when ranges or full distributions are involved. Think about what, if any, baseline or status quo you are using as an anchor, and whether this is a reasonable thing to do. Also, try to explicitly think of conditions under which more extreme values than your current range estimates might occur – if you can think of these conditions fairly easily, your range may be too narrow.

People often judge the probability of an event by the degree to which it is similar to a larger group of occurrences or the process that generated it. This is called judgment by "representativeness".

The representativeness heuristic is perhaps the most difficult to pin down conceptually – it is closely related to the stereotypes and hidden assumptions that we all carry with us. In the classic demonstration of representativeness, people were given the following description of a person named Steve:

> “Steve is very shy and withdrawn, invariably helpful, but with little interest in people. A meek and tidy soul, he has a need for order and structure, and a passion for detail.”

People judged Steve much more likely to be a librarian than a salesperson, because the descriptive “fits” the stereotype of a librarian. But think about how many salespeople there are in South Africa, and how many librarians there are. Is it really likely that, despite there being (probably) millions of salespeople and only a few hundred librarians in South Africa, Steve is more likely to be a librarian? What is happening here is that people are ignoring the base rates of the two professions when they make their judgements, even though this is a critical piece of information, far more diagnostic than the simple two-sentence summary of Steve.

Ignoring base rates is a common bias associated with representativeness, but the broader problem is that representativeness induces us to give too much weight to fairly weak qualitative information. Descriptions alluding to stereotypes or assumptions are usually weakly predictive, at best. Usually this overweighting of qualitative descriptions comes at the expense of much stronger quantitative information, like base rates, sample sizes, and the like.

Biases due to representativeness creep in whenever we allow qualitative descriptions to have more weight than they really deserve. This is an especially difficult problem to identify, because it is rarely perfectly clear how much weight a description should get – whether it is highly accurate, or not at all, or something inbetween. The balance of evidence suggests though that overweighting qualitative evidence is much more of a problem than underweighting it. Finding a balance is difficult, but again it is important to evaluate the quality of the information you are basing your judgment on, especially if you are making a judgment based on how much it “sounds like” something else. Try to imagine circumstances in which an event may be quite different from the phenomenon that it seems similar to. If it is easy to imagine such conditions, the apparent similarity is probably only superficial, and weakly diagnostic at best.

The final bias we describe here does not arise so much from a heuristic as it is a general psychological trait. That is,

We tend to seek mainly evidence that confirms our beliefs. We often do not seek out evidence that would contradict or disprove our beliefs, or find reasons to downweight or ignore this evidence when we do come across it.

This tendency ignores the fact that contradictory (or “disconfirming”) evidence is logically far more powerful than supporting (or “confirming”) evidence. In logic a single contradiction is enough to prove a belief false, while no number of supporting pieces of evidence can conclusively “prove” that a belief is true – there is always the possibility that some contradiction may be found later on.

A classic demonstration of our tendency to seek confirming evidence is to show people the set of four cards below, each of which have a letter on one side and a number on the other.

Suppose someone claims: “If a card has a vowel on one side, then it has an even number on the other side”. Which of the cards should one turn over to test this claim? In the original study, 45% of people chose “A and 4”, 33% chose “just A”, and 5% chose “A and 7”. Before reading further, you may want to decide which cards you would choose.

It is fairly clear that the “A” card should be turned over: if there is an even number on its other side we have some confirming evidence, if there is an odd number then we have disproved the claim. But most people do not see at first that we must also turn over the “7” card – if we find a vowel, we have disproved the claim. The key difference between the “A” and “7” card is that the “7” card cannot provide any confirming evidence, only disconfirming. The “4” card here is irrelevant because even if there is a consonant on the other side that would not disprove the claim (since the claim says nothing about what should appear on the other side of a consonant card).

In reality things are rarely so clear-cut that beliefs or theories can be conclusively proved or disproved, but it is important to bear in mind our tendency to seek out and pay attention to evidence that confirms our current beliefs. Try to actively think of possible counter-arguments to your opinions, and to find information that may support these opposing points of view. Although you may not change your mind, paying attention to a range of evidence is likely to improve your knowledge of the problem at hand, and reduce any overconfidence that might otherwise creep in.

Introduction to the SATIM model

The information that you provide during the elicitation process will be used to generate a range of plausible inputs to a large-scale energy model created and hosted by the Energy Research Centre at the University of Cape Town. This model is known as SATIM – the South African Times Model – and was originally created for the Long Term Mitigation Scenarios (LTMS) project but is now in its third generation. This document provides a brief overview of the model

The economy of a nation or region consumes energy from a number of primary and secondary sources. This energy delivers services by means of a myriad of technologies large and small. A model of the demand for energy needs to capture this complex structure and thus these sources and technologies need to be organised in some logical way. The SATIM energy model is an attempt at just such a model. It is based on TIMES, a partial equilibrium linear optimisation model developed by ETSAP, one of the International Energy Agency's implementing agencies, and a successor to MARKAL.

The SATIM model is a stylized representation of the whole energy system, with an optimization step that selects the mix of supply-side technologies that meets the demand for final energy at least cost.

The model includes economic costs, emissions, and a range of sector-specific constraints that can be applied at a point in time or cumulatively. A user interface provides a framework for both structuring the model and scenarios, and also for interpreting results. The model has proven useful in assessing the complex interrelationships between potential mitigation policies.

The SATIM model is fundamentally "sectoral", in that it organises the demand for energy by economic sector, and characterises the demand for energy in a sector by the energy services required by that sector.

The SATIM model uses five demand sectors and two supply sectors – industry, agriculture, residential commercial and transport on the demand side, and electricity and liquid fuels on the supply side. The level of detail for a sector depends on the relative contribution of the sector to total consumption and also on how much funding has been historically received for developing that sector in the model. Thus the model for the Transport sector is quite detailed but that of the Agricultural sector is quite simplistically represented in SATIM, because in South Africa the Agriculture sector accounts for relatively small energy consumption and low emissions.

In SATIM, services supplied to each of the five sectors are driven by technologies that require energy, with the quantity of that energy supply depending on the efficiency of the technology.

Useful energy is an exogenous model input disaggregated by energy carrier, for each demand sector. Final energy demand is determined endogenously using the assumed efficiencies of the least cost demand-side technologies selected by the model. The two supply sectors and primary energy sources must meet the sum of these demands, with the model optimizing the mix of supply-side technologies to meet the demand for final energy at least cost.

The SATIM model includes a number of parameters and general assumptions broadly covering, for each sector: (a) the structure of the sector and its energy services as it impacts on the demand for energy; (b) the establishment of base year demand for energy in the sector; (c) technical and cost parameters of the technologies available to satisfy the demand for energy services currently and in the future; (d) the projection of future demand for energy services. Several of these are the focus of the current elicitation process.