Probabilistic projections of baseline greenhouse gas emissions in South Africa to 2050

Bruno Merven[1], Bryce McCall[1], Ian Durbach[2,3], Harald Winkler[1] , Alison Hughes[1]
[1] *Energy Research Center, University of Cape Town, Rondebosch 7701*
[2] *Center for Statistics in Ecology, Environment, and Conservation, Department of Statistical Sciences, University of Cape Town, Rondebosch 7701*
[3] *African Institute for Mathematical Sciences, 6-8 Melrose Rd, Muizenberg 7945*

Keywords: forecasting, greenhouse gas emissions, E3 modelling, uncertainty

Abstract

We construct probabilistic projections of baseline carbon dioxide and other greenhouse gas emissions for South Africa to 2050. Our approach uses a mixed methodology. We obtain probabilistic projections of 11 key drivers of energy demand, using past literature, expert elicitation, and further modelling. These are randomly sampled and passed to an energy-economic model implementing TIMES for South Africa. Probabilistic projections of emissions are obtained as an output of this Monte Carlo simulation. Total emissions are expected to rise, slowly to 2030 and then more rapidly, but to fall per unit of GDP. Enormous uncertainties exist: 95% confidence intervals for total emissions are 450-475Mt CO_2 equivalent in 2020; 450-640Mt in 2035; and 465-1100Mt in 2050. Median projections are 465, 520, and 675Mt CO_2 equivalent per year in 2020, 2035, and 2050 respectively. Perhaps the key uncertainty in the setting of baseline GHG emissions in South Africa is the relative price of coal to gas, a result of the large share of emissions produced by electricity generation.

1. Introduction

Debates about the potential effects of climate change, the necessity for action, and the relative merits of different response strategies often refer to what is expected to occur if we "do nothing" – meaning, loosely speaking, under policies not too different from those currently in place (Azar, Lindgren & Andersson, 2013; O'Neill, Riahi & Keppo, 2010; Smith et al., 2000; Wright & Fulton, 2006). Domestic climate change policy, as well as national positions in global climate negotiations, for example, must strike a balance between reducing greenhouse gas (GHG) emissions and maintaining economic growth, particularly in developing countries. Mitigation costs, to take another example, are often calculated as the difference in monetary cost between a baseline situation and a new one characterized by lower emissions (Hourcade & Robinson, 1996).

The definition of an appropriate baseline trajectory is problematic (Clapp & Prag, 2012). Many factors influencing greenhouse gas emissions – population, for example – are subject to considerable external or aleatory uncertainty, and can only be predicted probabilistically. Deciding which policies and commitments to include in the baseline is also difficult, given the time-scales over which policies are implemented and historical differences between commitments and actions. For developing countries poverty, inequality, and education goals can not be traded-off against mitigation goals. The extent to which development goals should play a role in the baseline is contested.

The goal of the 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. We compute uncertainty around forecasted GHG emissions indirectly, by first obtaining forecasts of a number of key drivers of energy demand – population growth, economic growth, technology characterization and various commodity prices. These forecasts are obtained using a combination of expert elicitation (Anadon, Nemet & Verdolini, 2013), literature review and secondary sources (Raftery et al., 2012), and further modelling (Nemet, 2006; Masini & Franckl, 2013). 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), an energy-economic-environment 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 result of the optimisation model. Monte Carlo simulation is used to generate many 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. 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 to 2050.

The remainder of the paper elaborates on the approach summarized above, and reports the results obtained. Sections 2 and 3 describe our methodology and define the baseline scenario. Sections 4 and 5 report results obtained from the assessment of input variables and subsequent forecasts of GHG emissions respectively. Section 6 provides a discussion of the results and concludes the paper.

2. Methodology

2.1. Assessing uncertainty on key drivers of GHG emissions

We base our approach on the South African TIMES model (SATIM), a partial equilibrium linear optimisation model that selects a mix of energy sources and technologies 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/technologies 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 on each of the key input variables considered using a combination of (a) a review of the literature, (b) elicitation from national experts, (c) further modelling. 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, although in some cases (e.g. nuclear) with some adjustment for the local context. We describe the literature on which these forecasts are based, and any adjustments made, in the sections below. In the same vein, we used existing UN probabilistic population forecasts developed for the 2015 revision of the World Population Prospects (United Nations Department of Economic and Social Affairs, 2015), 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 (O'Hagan et al., 2006; Morgan & Henrion, 1990) when assessing this information, using the protocol outlined below.

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).
2. Establishing rapport: following introductions, we reviewed the pre-elicitation documents and the elicitation task, focussing on pitfalls of probability assessment (e.g. overconfidence, anchoring).
3. Qualitative elicitation of factors influencing key drivers: we asked experts to identify the important factors that might influence their later quantitative judgments and assess a small number of scenarios that might result in a particularly high or low value for the key driver.
4. Quantitative elicitation: to keep the elicitation manageable we assessed three points (minimum, mode, maximum) on each distribution, modifying these where additional information (on intermediate quantiles, for example) was offered. To avoid anchoring 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 2020, then for 2035, and finally for 2050. Although this might lead later estimates to be biased towards 2020 values – which would usually be associated with overly narrow confidence intervals – again we felt that it would be counter-productive to force any other order.

5. Post-elicitation verification: 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. Experts were asked to review their judgements and make adjustments where necessary.

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.

2.2. E3 modelling using SATIM

E3 modelling refers broadly to models that include energy, economic, and environmental components in a single model. These have been extensively used to address policy issues around mitigation and planning in response to potential climate changes (see e.g. Hedenus, Johansson, and Lindergren, 2013; O'Neill, Riahi, and Keppo I, 2010; Richels and Blanford, 2008; Rozenberg et al., 2010; Sassi et al., 2010)

SATIM – the South African Times Model – is an E3 model created and hosted by the Energy Research Centre at the University of Cape Town. This model was originally developed for the Long Term Mitigation Scenarios (LTMS) project but is now in its third generation. The SATIM energy model 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 (see e.g. Vaillancourt et al., 2008; McCollum et al., 2012).

The SATIM model is a stylized representation of the whole energy system, with an optimization step that selects the mix of 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 "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. 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) 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) demand projections for energy services.

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. We ran the SATIM model in both perfect foresight and myopic (ten-year planning horizon with five-year overlaps i.e. essentially reviewing the planning every five years) modes. The global discount rate, which affects how technologies with high upfront capital costs (e.g. nuclear and renewables) compete with other technologies with relatively higher fuel costs over the life of the technology, was varied at three levels: 8% (used in most recent national planning tasks), 5%, and 11%.

3. 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

GHG emissions that are included are CO_2, CH_4 (including fugitive emissions) and N_2O. 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. Our model thus covers GHG emissions for these sectors.

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 business-as-usual by 2020 and 42% by 2025. This primarily manifests in international commodity prices, which influence local prices particularly in the case of coal.

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 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.

4. Assessment of input variables

Based on knowledge of the underlying SATIM model, key drivers of GHG emissions were selected. These are shown in 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 are described in the following sections, while Figure 1 shows the probabilistic projections constructed for each input variable.

4.1. Population growth

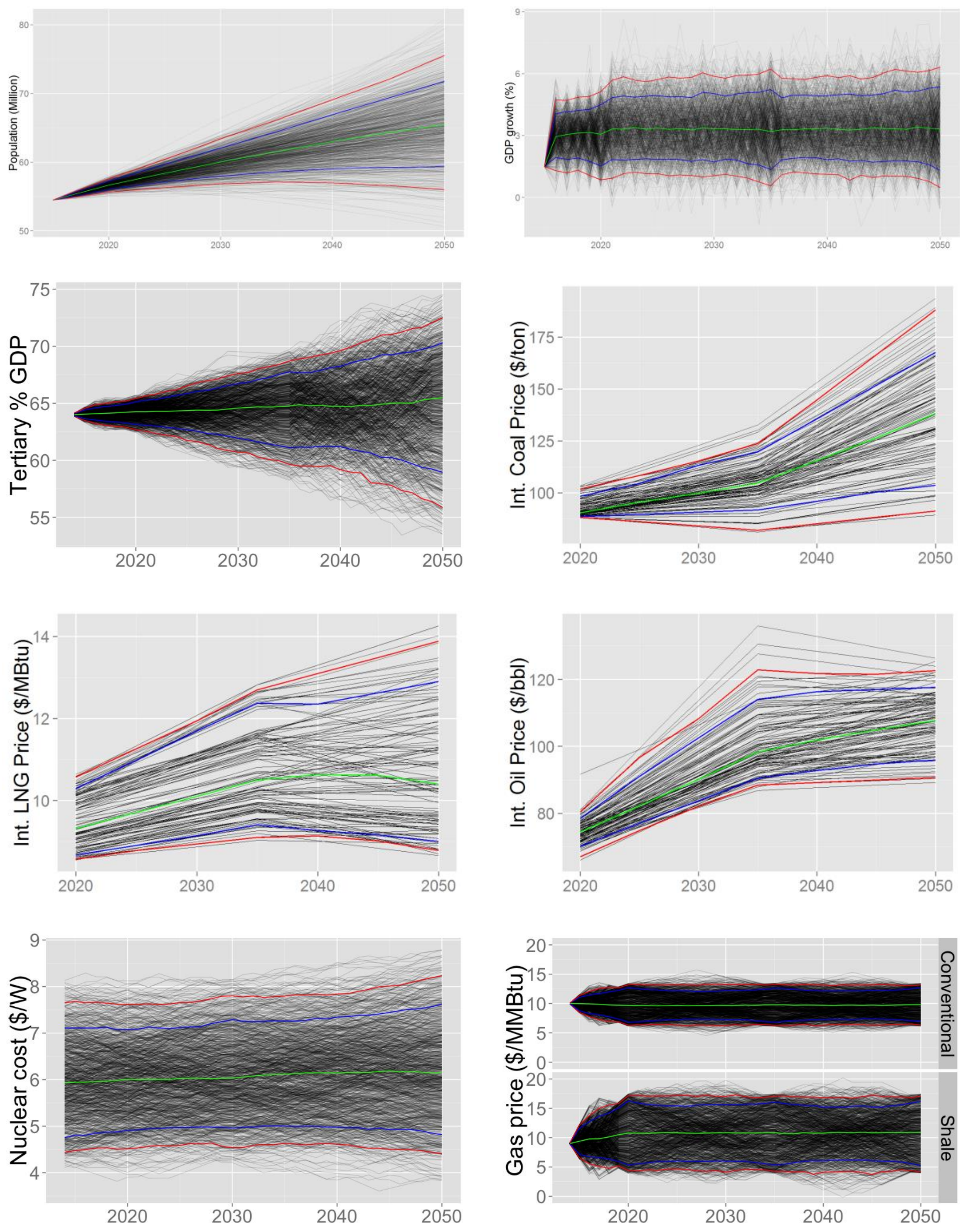
Models for producing probabilistic population projections have been recently developed specifically for use by the United Nation Population Division (Raftery, Li, Sevcikova, Gerland, & Heilig, 2012). Changes in a country's population are determined by a number of factors, but chiefly fertility and mortality. 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).

4.2. GDP growth and composition

We independently conducted elicitation interviews with two senior academic macroeconomists at the University of Cape Town. The elicitation team consisted of two analysts, one with a background in decision analysis and one working in energy modelling. Interviews took place in May and July 2014. Both experts were asked 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, and 2050. The elicited probability distributions thus covered possible values in the mean growth rate over these three periods.

Discussion with each expert lasted 1.5 – 2 hours, split roughly equally between two topics: factors influencing GDP growth in South Africa, and future trends and scenarios. Discussion around influential factors was remarkably consistent between experts and a number of common features were observed. This led quite naturally into the experts sketching some possible "high", "low" and "in-between" scenarios for GDP growth in South Africa over the three periods 2014-2020; 2020-2035; 2035-2050. These were assessed as qualitative, internally consistent stories involving, for example, changes in political policies, trading relationships, sectoral contributions, etc. Again the two experts showed a large degree of agreement in their qualitative scenarios, which had in common a perception of little or no change in the near future (2014-2020), limited prospects for the future but a hard floor beyond which GDP growth was felt to be unlikely to fall.

Two key properties expressed by both experts are (a) upper and lower bounds on growth, and (b) the mean-reverting nature of GDP growth rates. 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. Very high or low average growth rates may be sustained over a decade, but these would almost certainly return to more moderate levels in subsequent decades. Both experts felt that while the tertiary sector was likely to grow at the expense of either primary or secondary sectors, large changes – more than a 3% change in a sector's share per decade – were unlikely. 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.



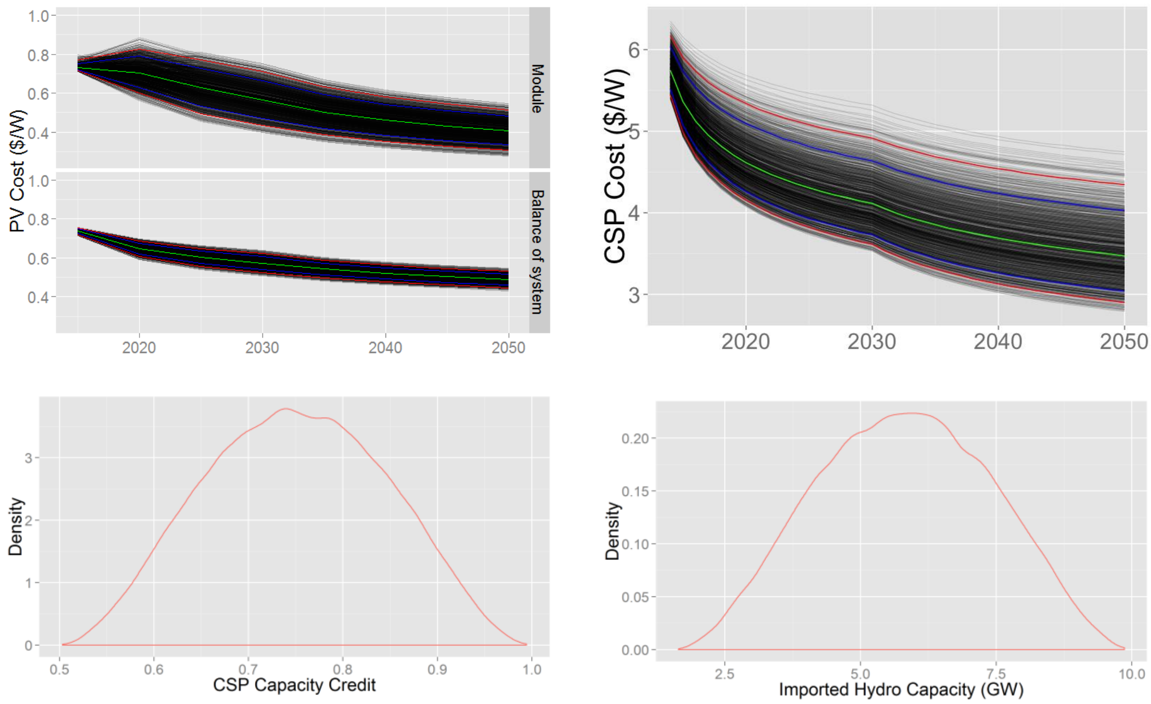


Figure 1: Probabilistic projections of E3 model inputs over the period 2015 to 2050.

4.3. Global energy commodity prices

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).

Table 2 shows mean commodity prices obtained from external data sources: IMACLIM-R (indicated by the first of the two values in each cell of the table), the IEA World Energy Outlook 2015 report (International Energy Agency, 2015), and the well-known Wood Mackenzie forecasts[1]. 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

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

2015 and other sources where available. Values in the IMACLIM-R rows denote indices before/after adjustment, with the multiplier used to make the adjustment provided below.

		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 interpolating between the three time periods.

4.4. Gas prices

We conducted elicitation interviews with two experts on the subject of gas prices. Prices depend primarily on the type and origin of the gas. The elicitation team consisted of two analysts, one with a background in decision analysis and one working in energy modelling. Interviews took place in May and July 2014. 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 and conventional deposits, 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.

Discussion with each expert lasted 2 – 3 hours, split roughly equally between two topics: factors influencing GDP growth in South Africa, and future trends and scenarios. 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 is perhaps the key determinant. This is a consequence of the flexibility of gas in its final uses, and the fact that it is a commodity traded for profit. Over the longer term, this means that prices are 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.

As a result, both experts gave essentially constant uncertainty through time, in contrast to other commodities like oil and coal. This was motivated by the gas price being, in effect, set at the marginal producer's cost of production, including capital costs. 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.

Current strategic planning around gas in South Africa centers on the development of large shale deposits within South Africa and even larger deposits of conventional gas in Mozambique. Opinions indicated 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 the extent. With respect to Mozambican gas, South Africa will have to compete with other customers for Mozambican gas on the open market. Conventional gas deposits in South Africa have thus far been limited in their scale and impact, and both experts felt that this was likely to continue.

When eliciting quantitative estimates, one expert indicated that conventional gas prices should be priced against a “next best alternative”. We have quantified this assessment using an average of the expert’s assessment of other competing gas types at [3.9, 17.5] for shale and [6.8, 13.7] for imported LNG.

4.5. Coal prices

South Africa relies heavily on coal, which currently supplies approximately 80% of energy needs. Forecasts of potential coal prices are therefore particularly important. We performed a set of elicitation interviews with four coal experts drawn from a range of backgrounds – private, public, and research. The elicitation team consisted of three analysts, one with a background in decision analysis, one working in energy modelling, and one researching institutional arrangements in the coal industry. Interviews took place between April and August 2014. During the course of these interviews and subsequent modelling, it became apparent that eliciting an exogenous coal price was impossible – the values our experts had assessed were conditioned on assumptions, primarily about the demand for coal, which were endogenous to the SATIM model. We therefore constructed a probabilistic supply cost curve, using input costs already elicited from experts as well as additional assumptions and external sources. The combined demand by coal power plants and other users, endogenously determined within SATIM, could then be applied to the reconstructed cost curve to determine the coal price.

We constructed the coal supply curve by decomposing the cost of coal into a set of input factors elicited as part of the expert interviews: the costs of mining it, 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); and assorted "other" costs (water costs, engineering costs, replacement capital costs, employee housing costs, and equipment costs. These are summarized in Table 3. Costs are given separately for two regions in South Africa: the Central Basin, which is where all current power stations and coal mines are situated; and the Waterberg, a relatively remote area of South Africa containing vast and unexploited deposits of mostly low-quality coal.

	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**

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

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 range of possible prices in the Waterberg region is due to uncertainty regarding stripping ratios and washing yields. Transport costs are a function of 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 Table 4.

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%

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

The resulting average prices for the Central Basin, Waterberg and combined is shown below. Figure 3 below shows the result of averaging the costs sampled from the ranges described above weighted by the production (shown in Figure 3) of each supply route for each corresponding cost scenario. The weighted average matches the combined elicited values quite well.

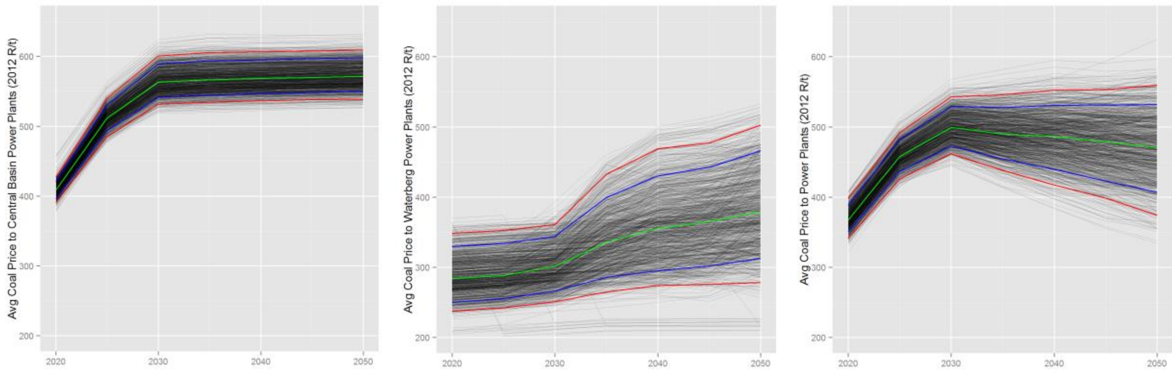


Figure 2 Average coal price seen by coal power plants

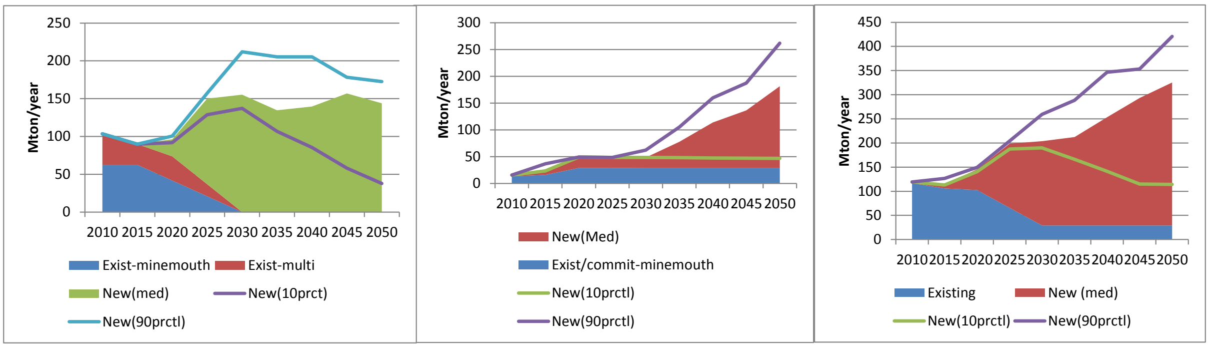


Figure 3 Production Range for different supply routes

4.6. Solar investment costs

Solar technologies are relatively young, and further advances are 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.

Total installed capacity

The 2014 IEA Energy Technology Perspectives report (International Energy Agency, 2014, 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 unscaled PV capacity as $\mathcal{B}(2,3)$ and CSP capacity as $\mathcal{B}(3,3)$.

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%.

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.

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 the choice of parameters for the respective beta distributions. These investment costs, calculated for "baseline" PV (utility with no tracking) and CSP (parabolic trough with 6-hour storage capacity) technologies, are converted into costs for other technologies by multiplying these by a factor held fixed at their current values i.e. at the present-day (2014) cost ratios,

4.7. Nuclear costs

Anadon, Nemet, & Verdolini (2013) report responses from 67 US and European 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. These assessments 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 South African government's Integrated Resource Plan (Department of Energy, 2013). The final distributions used are: Tri(4109, 5800, 8200) in 2010; Tri(4301, 5942, 8269) in 2030; and Tri(4119, 5942, 8528) in 2050.

4.8. Hydro Imports

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 1.

4.9. Post-processing of inputs

Several input variables required additional processing in order to transform them into forms suitable to pass to the SATIM model i.e. annual time series. Two main types of post-processing were required. Firstly, where we have collected information at only a small number of time-points, we generate probabilistic projections for each remaining time-point by first simulating randomly from the existing probability distributions, taking into account any desired inter-temporal relationships, and then by simulating values between elicited time points, again taking into account any inter-temporal information gathered as part of the elicitation or modelling process. Different simulation approaches may be required for the first step depending on the nature of the inter-temporal information. In the appendix we describe two algorithms that can be used when inter-temporal information is expressed as a matrix of correlations (Algorithm 1) or as a process of mean-reversion (Algorithm 2). We use linear interpolation, with or without the addition of a noise component, to simulate the remaining values, although other approaches are possible.

Where information has been collected from a number of sources (e.g. by elicitation from several experts), we combine these sources using a linear opinion pool (French, 2011). This is equivalent to simulating projections from each source in proportion to the desired weight of that source. The result is a linear opinion pool for projections.

Finally, we assemble the projections independently generated for each input variable into input matrices, each of which combines a single set of projections for each of the key drivers. In doing so, we need to account for correlations between inputs (for example, population and economic growth are positively associated), but assessing the full correlation matrix is difficult. Experts exist with subject areas, but the assessment of inter-variable correlations requires an extremely broad and deep knowledge, encompassing all the input variables. We therefore place a moderate positive correlation of 0.3 between population and GDP growth; the international commodity prices obtained from IMACLIM-R are already correlated in that they are drawn from the same set of scenarios. Apart from these inter-relationships input matrices are generated by sampling independently from the input distributions.

5. Probabilistic projections of baseline GHG emissions

Our main results are shown in Figure 4. 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 (Figure 4a). 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.

Our results show that a no climate policy scenario has wide ranges of GHG emissions, with median projections rising throughout, but moderately: from 420Mt CO2-equivalent per year in 2010 to 520Mt per year in 2035 and 675Mt in 2050. Median projections should be interpreted cautiously. For example, the median projection of 520 Mt CO2-equivelent in 2035 is within the 'peak, plateau and decline' (PPD) trajectory range in national climate policy[2], which is 398 to 614 Mt CO2-eq for 2035. The range in our modelling projects GHG emissions from 450Mt and 640Mt 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.

Per capita emissions are also expected to rise, though by less than absolute emissions (not shown here). Substantial uncertainty again exists in the forecasts, particularly beyond 2030. Our results indicate that 95% of trajectories lie between 8t and 8.4t per capita in 2020; between 7.6t and 10t per capita in 2035; and between 7.6t and 15.3t per capita in 2050. Median per capita emissions remain roughly the same as present-day values of 8t per capita until 2035, after which they rise steadily to just over 10t per capita in 2050.

[2] Government of the Republic of South Africa, National Climate Change Response White Paper, October 2011

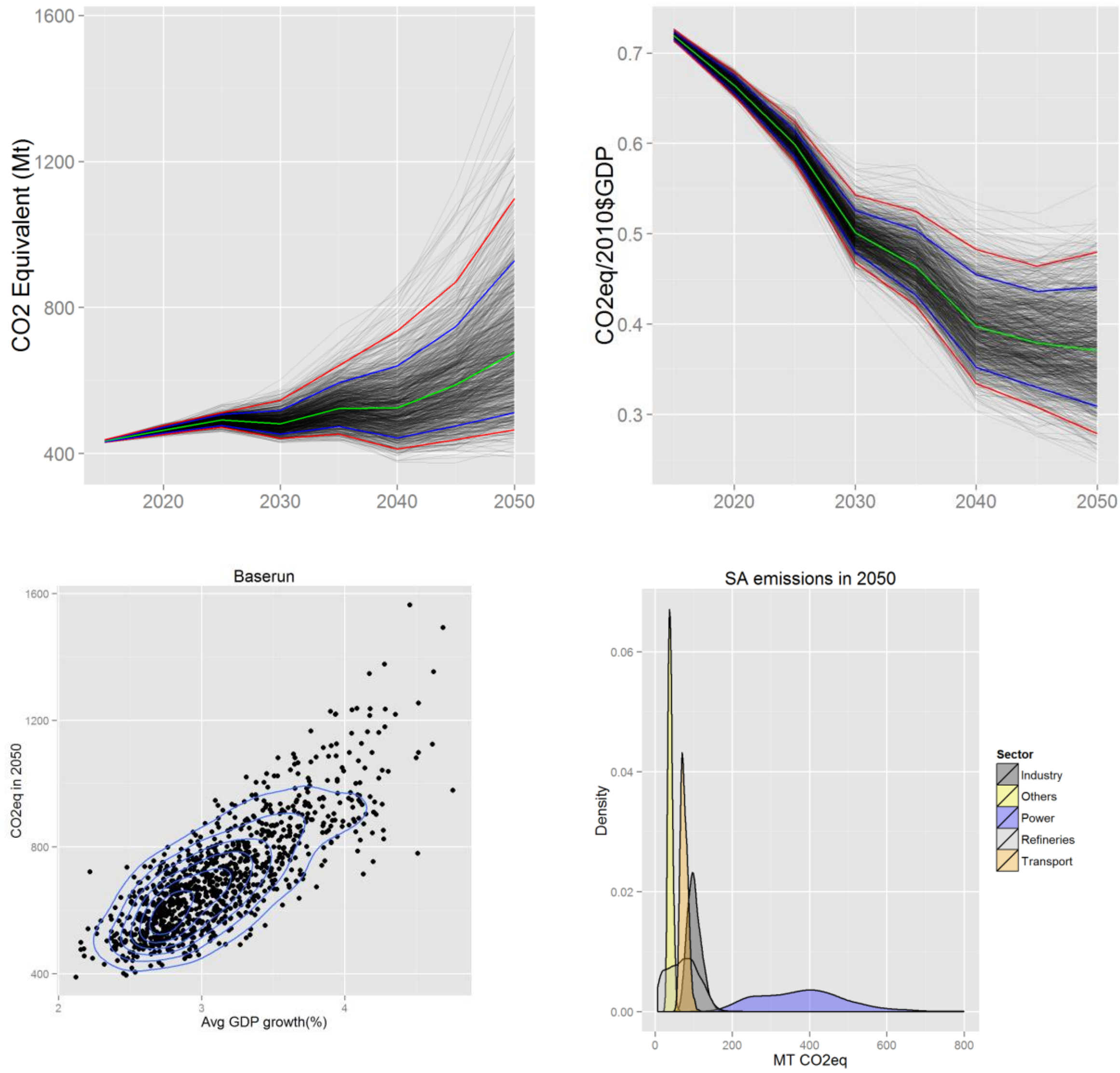


Figure 4: Probabilistic projections of CO_2 equivalent produced by South Africa over the period 2015 to 2050: (a) total GHG emissions, (b) GHG emissions per unit GDP, (c) scatterplot showing relationship between GDP growth and GHG emissions at 2050, (d) histogram showing GHG emissions contributed by each sector in 2050.

Emissions intensity, that is GHG emissions per unit of GDP, falls consistently and approximately linearly throughout the forecast period (Figure 4b). Nevertheless substantial uncertainty still exists, particularly after 2030. Our results indicate that 95% of trajectories lie between 0.65 and 0.68kg/$GDP in 2020; between 0.42 and 0.52kg/$GDP in 2035; and between 0.28 and 0.48kg/$GDP in 2050. Median forecasts are for CO2 emissions per dollar of GDP to drop from current day levels of 0.66kg/$GDP to 0.46kg/$GDP in 2035 and 0.37kg/$GDP in 2050. The strong positive relationship between GDP growth and GHG emissions is shown in Figure 4c.

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 (Figure 4d).

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. These are summarized in Figure 5. The primary uncertainty is the extent to which gas replaces coal in the production of electricity (Figure 5a and Figure 5b). 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.

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 (Figure 5c). 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 18% in 2050, although the median projection is for CSP to contribute little or nothing (less than 1%) throughout the forecast period (Figure 5d).

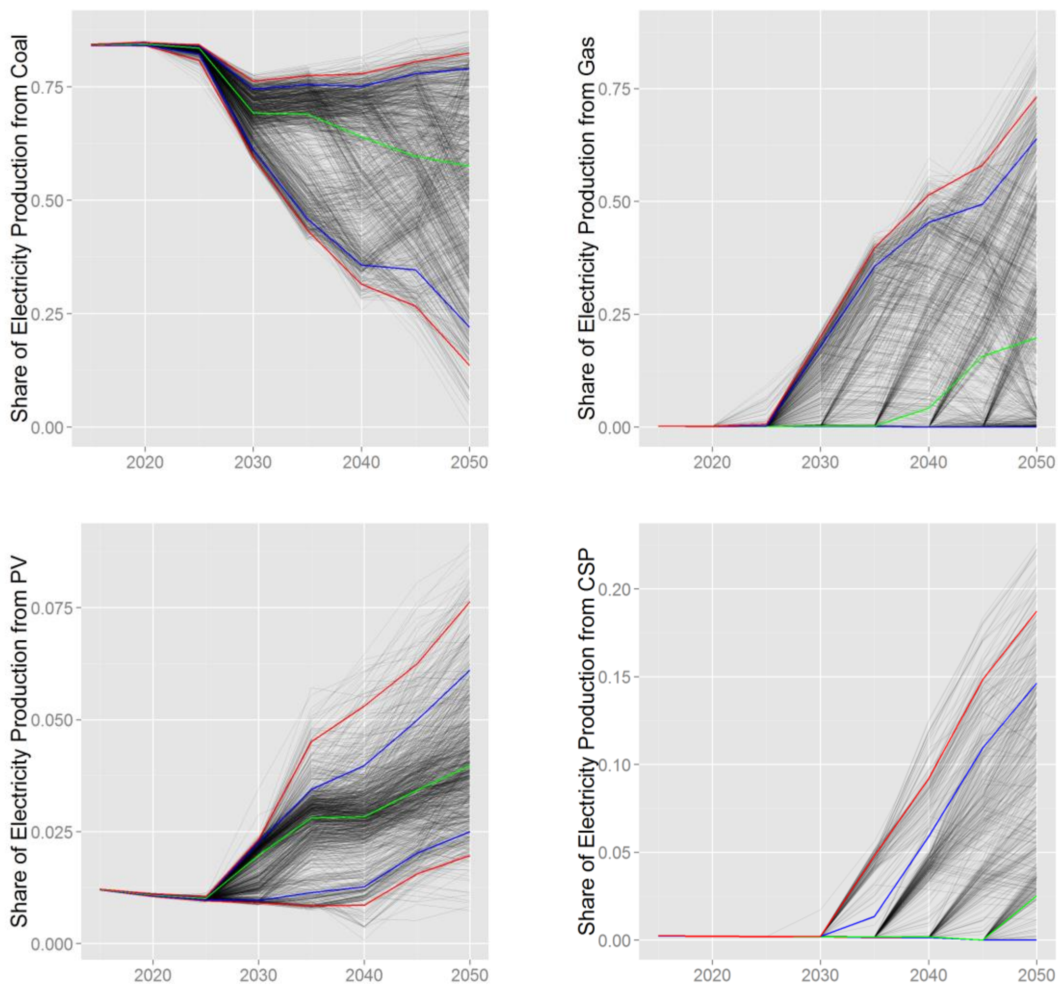


Figure 5: Probabilistic projections of the share of electricity production in South Africa contributed by (a) coal, (b) gas, (c) PV, (d) CSP over the period 2010 to 2050.

Figure 6 shows the sensitivity of our main results to global discount rate assumptions. We show results at 2050 under various model assumptions; the same conclusions are drawn if other periods are examined. GHG emissions are extremely robust to the assumption of perfect foresight; emissions obtained from a myopic model show no qualitative differences and are thus not shown here. Emissions decrease marginally at higher discount rates (Figure 6a), even as the fossil fuel share of electricity production increases (Figure 6b). This is a direct result of gas replacing coal as a fuel source (Figure 6c). The increased use of gas to generate electricity at the higher discount rate is a result of both shale gas and LNG being competitively priced in relation to coal and other fuel sources (Figure 6f), whereas this does not occur at lower discount rates (Figure 6d and e).

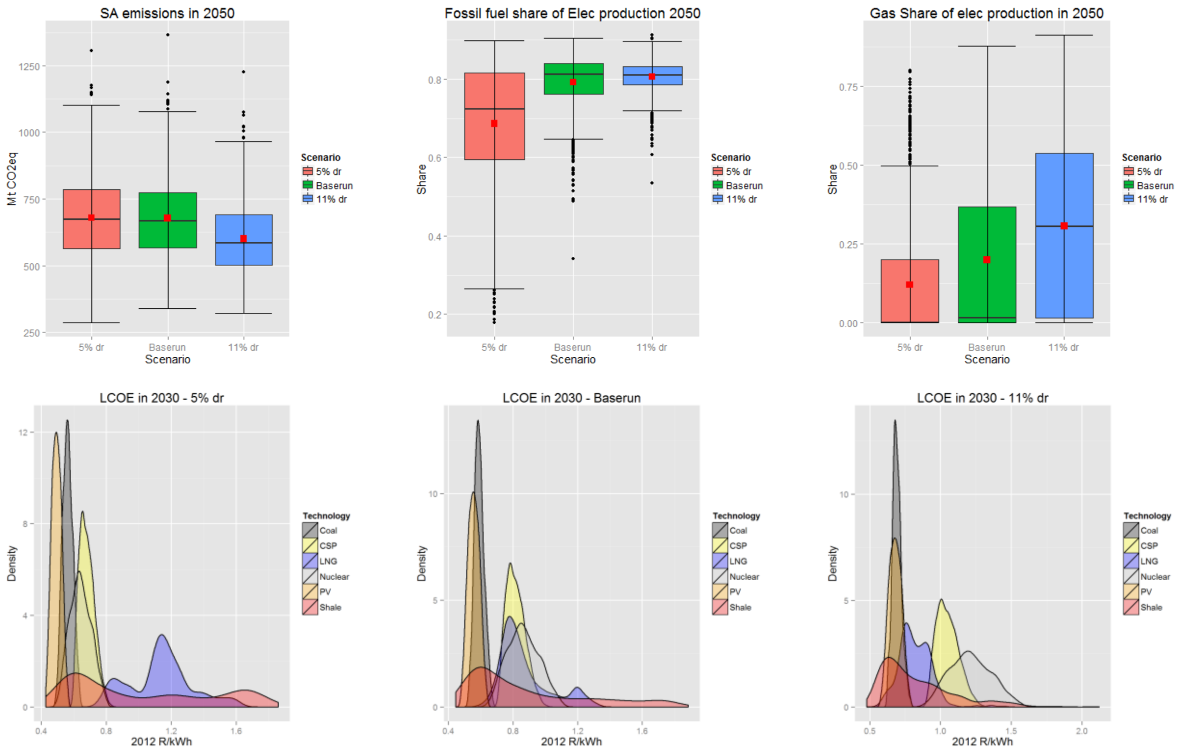


Figure 6: Sensitivity of results to discount rate assumptions.

6. Discussion and conclusions

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.

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. 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 they 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 this point, our forecasts of GHG emissions are also 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 becomes available. This also motivates for the regular updating of baseline projections. Our projections are based in places on the assessments obtained from only a small sample of experts; enlarging this sample provides another avenue for further work. 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.

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. Uncertainty in CO_2 per GDP is much narrower than uncertainty regarding the absolute level of CO_2 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.

While the uncertainty explored here focuses on supply technologies and fuel prices, 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. Incorporating demand-side uncertainty is an important area for future research, and one that is likely to widen uncertainty bounds further, particularly on the lower end of the distribution of GHG emissions. Related to this point, 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 makes a number of assumptions that may not hold over the long-term. Fundamental or extreme changes to the system from climate impacts e.g. increase in temperature, are not taken into account by SATIM.

References

Azar, C., Lindgren, K., & Andersson, B. 2003. Global energy scenarios meeting stringent CO2 constraints – cost effective fuel choices in the transportation sector. *Energy Policy 31,* 961-976.

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

Craig, P., Gadgil, A., & Koomey, J. (2002). What can history teach us? A retrospective examination of long-term energy forecasts for the United States. *Annual Review of Energy and the Environment, 27(1),* 83-118.

Department of Energy. (2013). Integrated resource plan for electricity 2010-2030 – update report 2013. Department of Energy, South Africa. Available from http://www.doe-irp.co.za/content/IRP2010_updatea.pdf.

French, S. (2011). Aggregating expert judgement. Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas, 105(1), 181-206.

Hedenus, F., Johansson, D., & Lindergren, K. (2013). A Critical Assessment of Energy-Economy-Climate Models for Policy Analysis, *Journal of Applied Economics and Business Research, 3(2)*, 118-132.

Hourcade, J-C., & Robinson, J. Mitigating factors: assessing the costs of reducing GHG emissions. (1996). Energy Policy, 24(10), 863-873.

International Energy Agency. (2013). World Energy Outlook 2013. International Energy Agency. Available from http://www.worldenergyoutlook.org/publications/weo-2013.

International Energy Agency. (2014). Energy Technology Perspectives 2014. International Energy Agency. Available from http://www.iea.org/etp2014.

Masini, A., & Franckl, P. (2003). Forecasting the diffusion of photovoltaic systems in southern Europe: A learning curve approach. *Technological Forecasting and Social Change, 70(1),* 39-65.

McCollum, D., Yang, C., Yeh, S., & Ogden, J. Deep greenhouse gas reduction scenarios for California – Strategic implications from the CA-TIMES energy-economic systems model. (2012). Energy Strategy Reviews, 1(1), 19-32.

Nemet, G. (2006). Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy, 34(17),* 3218—3232.

O'Neill, B.C., Riahi K., & Keppo I. 2010. Mitigation implications of mid-century targets that preserve long-term climate policy options. *Proceedings of the National Academy of Sciences*, *107(3)*, 1011-1016.

O'Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, P., Jenkinson, D., Oakley, J., & Rakow, T. (2006). *Uncertain judgements: eliciting experts' probabilities.* John Wiley & Sons, Chichester.

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, 109(35)*, 13915-13921.

Richels, R.G, & Blanford G.J. 2008. The value of technology advance in decarbonizing the US economy, *Energy Economics 30*, 2930-2946.

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.

Smith, P., Powlson, D., Smith, J., Falloon, P., & Coleman, K. (2000). Meeting Europe's climate change commitments: quantitative estimates of the potential for carbon mitigation by agriculture, Global Change Biology, 6(5), 525-539.

United Nations Department of Economic and Social Affairs. (2015). World Population Prospects: The 2015 Revision. United Nations Publications.

Vaillancourt, K., Labriet, M., Loulou, R., & Waaub, J-P. (2008). The role of nuclear energy in long-term climate scenarios: An analysis with the World-TIMES model. Energy Policy, 36(7), 2296-2307.

Wright, L., & Fulton, L. (2005). Climate change mitigation and transport in developing nations. Transport Reviews, 25(6), 691-717.

Appendix A: Details of post-processing of elicited data sources

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 $F_1, F_2, \ldots, F_P$, 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 $F_1^{-1}, F_2^{-1}, \ldots, F_P^{-1}$ 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

Series in which extremely high or low values should be followed by periods of average (or at least, much less extreme high or low) values pose a problem for Algorithm 1 because the desired relationship is not easily expressed using correlations defined between the original variables. We use the following procedure:

1. Simulate N independent values from each of the desired distributions $F_1, F_2, \ldots, F_P$ using standard methods. For each t for which mean-reversion should occur from t to $t+1$:
2. For each of the values simulated from F_t (denoted $x_{it}, i = 1, \ldots, N$) compute the "extremeness" of the observation, as measured by its absolute difference to the sample mean $\bar{x}_t$ i.e. $e_{it} = |x_{it} - \bar{x}_t|$, and rank order these in descending order.
3. Allocate each observation x_{it} into a "mean-reverting" group with probability inversely proportional to the rank of e_{it}.
4. Allocate each of the N_{MR} observations in the mean-reverting group to intermediate positions in the rank order of simulated values from F_{t+1}. This can be done in several ways. We identified the N_{MR} most central values in $\boldsymbol{x}_{t+1}$, and then reorder these so that the correlation between the ranks of $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$ is some desired value ρ, using Algorithm 1. This is done so that there is still some positive correlation between time periods within the set of observations that are subjected to mean reversion. That is, a very large observation at time t will tend to be less extreme, but still above the median, at $t+1$.
5. All observations not in the mean reverting group are allocated to the remaining $N - N_{MR}$ ranks at $t+1$ by reordering these so that the correlation between the ranks of the remaining values of $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$ is some desired value ρ, again using Algorithm 1.