Forecasts for electricity demand in South Africa (2010 – 2035) using the CSIR sectoral regression model

June 2010

Prepared for: Eskom
(as inputs into the IRP 2 process)

Document Reference Number: CSIR/BE/LQM/ER/2010/0044/B

Author Contact Details:
Project team leader: P Debba (012) 841-3421, **pdebba@csir.co.za**
Project team members: R Koen, JP Holloway, T Magadla,
M Rasuba, S Khuluse, CD Elphinstone

Table of Contents

List of Figures

1. Introduction

The CSIR was contracted by BHP Billiton to develop a set of forecasts for the national consumption of electricity in South Africa in 2003/4, and to revise the forecasts in 2005 and 2008 with updated data. Note that these forecasts were designed to determine the demand for electricity without considering constraints on the supply side. With the agreement of BHP Billiton, the CSIR has made available the use of these models to support Eskom in developing inputs for the IRP 2 process.

In order to produce a preliminary set of electricity demand forecasts for IRP 2, the same basic methodology as utilised in the development of the BHP Billiton models was followed. However, the models were revised to incorporate updated data and to make minor changes to align with the IRP 2 process. This document provides a set of national electricity consumption forecasts based on both data developed for the IRP 2 public participation process and these revised models. The forecasts are given in both tabular and graphical format.

In order to interpret these forecasts, it is necessary to understand more about the methodology followed to compile them. For simplicity, the term “CSIR model” will be used to refer to this revised BHP Billiton/CSIR model, which is really a collection of models that forecast the expected electricity consumption in various electricity sectors in order to obtain a total national forecast.

2. Methodology

The methodology followed to obtain the forecasts presented in this document consisted of two parts. The first part consisted of putting together the required datasets, and the second of compiling the models. However, some basic underlying assumptions are discussed briefly first.

2.1. Underlying assumptions

In accordance with the IRP 2 process and to align with information supplied by Eskom, the following was assumed:

- The long-term forecasts cover the total requirement for electricity to be generated in order to meet the needs of South Africa, including electricity imported, but excluding the needs of international customers and also excluding electricity consumed in power stations and auxiliary systems.
- The long-term forecasts cover the period 2010 to 2035.
- Forecasts in the form of estimated future values were developed on key parameters in order to feed into the forecasting process. Details of these are provided later in the document.
- The forecasts provided are based purely on statistical models (not judgmental models) and therefore rely on available quantitative data. Qualitative information on perceptions, risks and changes in the environment are not included in the models.
- Aspects that could potentially affect the consumption of electricity, but for which data were *not available* at the level of the chosen electricity sectors were not explicitly included in the models. Possible examples are electricity intensity and price elasticity.

2.2. Data selection and use

Data on historical consumption of electricity as well as potential "drivers" of electricity usage had to be collected and checked. Since the IRP 2 process provided forecasts for such potential "drivers", the main emphasis was on the collection and checking of the electricity consumption data.

2.2.1. Electricity consumption data

Information about both total consumption and consumption per sector had to be obtained for the modelling process.

2.2.2. Total consumption

Data on total national electricity consumption in GWh for South Africa from 1978 to 2006 were obtained from Statistics South Africa (StatsSA) from Statistical release P4141 – Generation and Consumption of electricity. This StatsSA data is reported as the total electricity available for distribution in South Africa which includes losses. Values are reported per month, and for the purpose of these forecasts, the monthly data was aggregated into a calendar year, i.e. January to December.

At the time of compiling the forecasts, data from StatsSA was available up to and including March 2010, but to correspond to other historical datasets, data was aggregated only up to the end of 2009.

2.2.3. Per sector consumption

Sectors could be defined as different types of customers, for example; domestic, mining, and so on. The reason for breaking up the demand into different sectors was because it was assumed that patterns of growth and decline in consumption would potentially differ between the various sectors, and that drivers may differ between the sectors. It was furthermore assumed that these patterns of usage in the individual sectors, especially the smaller sectors, may "cancel each other out" and no longer be visible within the patterns observed in the overall consumption. Attempts were therefore made to group together consumers with similar usage patterns, as represented by sectors. As a result, data on consumption patterns were required at a sector level.

Data from various sources had already been compared extensively during the process of developing the BHP Billiton forecasts. The following potential sources were identified, but none of these could provide a consistent set of per sector data from 1972:

- Statistics SA (some sectors, 1964 -1984)
- University of Johannesburg, then known as RAU (1972-1991) [1]
- Department of Energy, then known as DME (1998 Digest) (1990-1997) [2]
- Department of Energy, then known as DME (2000 Digest) (1992-2000) [3]
- Department of Energy, then known as DME (2006 Digest) (1992-2004) [4]
- Department of Energy, then known as DME (Spreadsheets) (1992-2005) [5]
- NERSA, previously called NER (1996-2005) [6]
- SA Energy Stats (Vol. 1 and Vol. 2) (1950-1993) [7 & 8]
- Eskom Annual Reports (1997-2009) [9]
- Eskom Statistical Yearbooks (1957-1996) [10]

Although the total consumption figures were roughly consistent between the various sources, at the individual sector level the various sources were not consistent. The CSIR team compared the different sources for each sector, and the most reliable data series to use for each sector was selected. These datasets chosen to represent each sector were checked by comparing the aggregated sectors to the Stats SA national consumption figures. The sectors selected by the CSIR corresponded to the categories used by the National Energy Regulator of South Africa (NERSA) that were also used by the Department of Energy, and reported to the International Energy Agency (see [11] for their 2007 statistics on electricity for South Africa, as well as [12] for a list of their data sources).

It should be noted that **neither the Department of Energy, nor NERSA have provided any data on electricity demand per sector since 2006**. Although the International Energy Agency lists the Department as one of their sources for the 2007 statistics they have published on South Africa, along with Eskom, they also indicate that they used internal estimates on this data. However, the CSIR had previously used a combination of NERSA, Department of Energy and Eskom data to obtain sector estimates for the more recent years. A similar technique was used to obtain estimates for 2007 – 2009 sector consumption data, based on Eskom's data published in their annual reports and using prior NERSA information relating to Eskom's percentage share. This was done in order to extend the "CSIR recommended" datasets per sector to the year 2009.

During 2010, the CSIR team also received a set of historical data per sector compiled by Johan Prinsloo for Eskom. Although the details regarding the compilation of each sector were not provided to the CSIR team, it was understood that the data were based on Eskom customer categories, but adjusted to RSA consumption in each sector by Johan using previous information pertaining to Eskom's share in each category.

The graphs in figures 1 – 5 provide an illustration of the various data sources consulted, and the relative differences between them. On all these graphs, the orange line labeled "ESKOM" refers to the sector data compiled by Johan Prinsloo, while the thick white line indicates the data considered by the CSIR team to be a reliable estimate for the sector. For some of the sectors, the CSIR recommended sets were quite close to that of the Eskom sectors, such as the mining sector. However, for the agricultural, transport and domestic sectors there were visible differences in either the size of the sector or the shape of the patterns. There did not seem to be an obvious explanation for the differences between the Eskom and the CSIR recommended sets for these sectors, since at the time of compiling this document complete details regarding the composition of the Eskom sectors were not known. Therefore the CSIR team decided to continue using the sector definitions that they had developed previously, rather than using the sector data provided by Johan Prinsloo.

Although the agricultural sector is a small sector, figure 1 indicates that the various sources differed quite widely on the pattern of consumption in this sector.

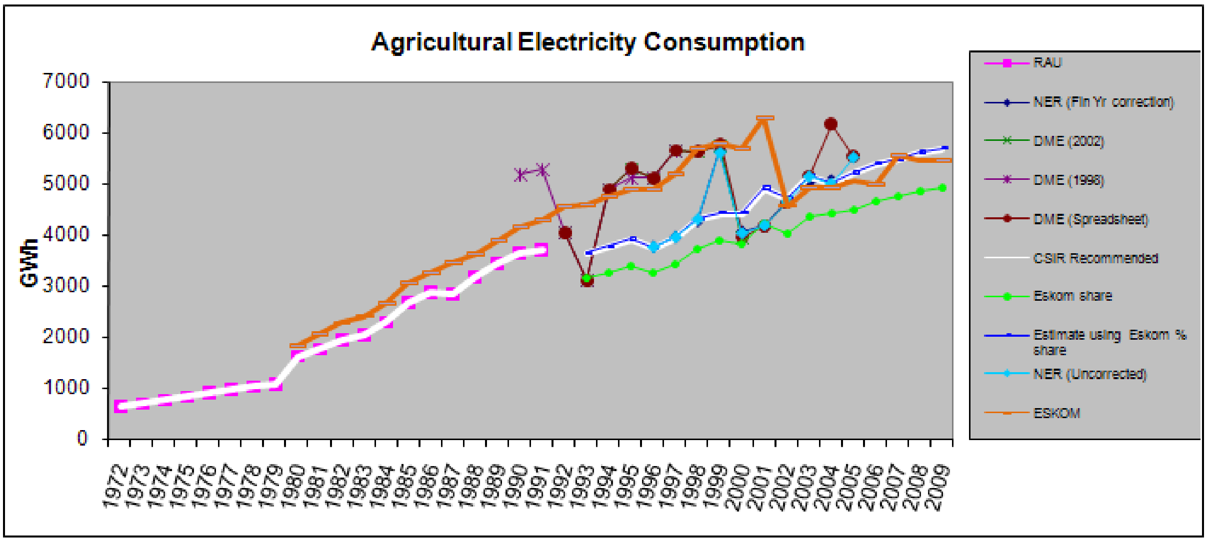


Figure 1 Comparing agricultural sector data between different sources

For the domestic sector, the overall patterns between some sectors coincided, but generally the consumption suggested by Johan for Eskom remained lower than other sources over the entire 1972 – 2009 period. It is assumed that the estimates for this sector are problematic due to the fact that most of the domestic consumers are supplied with electricity via the municipalities, i.e. they form a large part of the "redistributors" sector within Eskom sales. The green line on the graph in figure 2 (the one that is substantially lower than the others) indicates the direct sales from Eskom to domestic consumers.

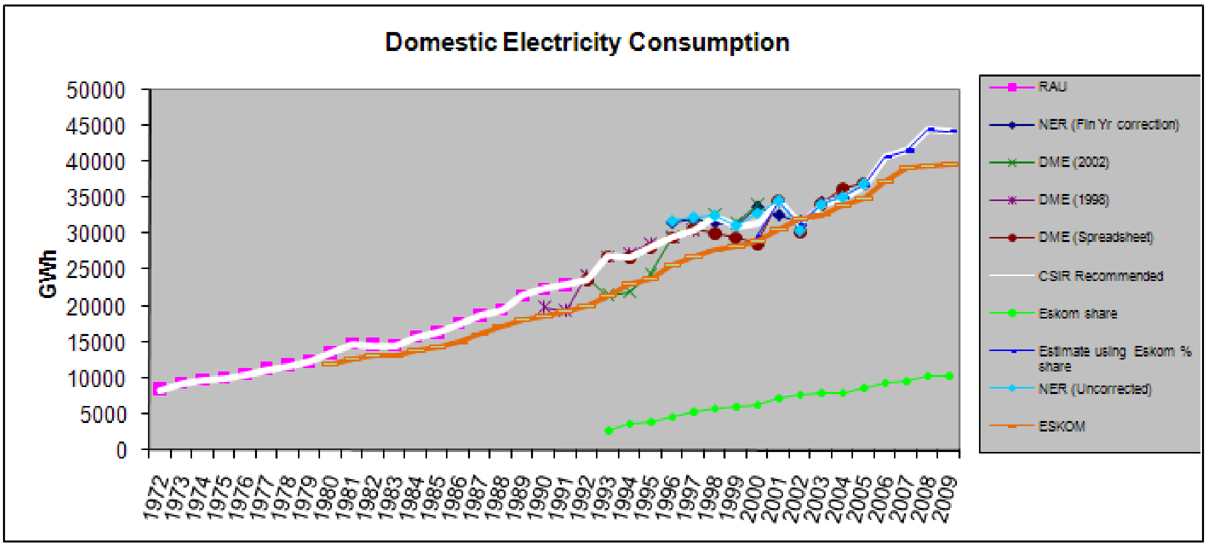


Figure 2 Comparing domestic sector data between different sources

Although most sources provide "Commerce" and "Manufacturing" sectors, definitions differ widely between them, and even within different years of the same source. Furthermore, most sources contain a "general" category, and the definition of this category is also not consistent. However, figure 3 shows that when data on "commerce", "manufacturing" and "general" sectors are combined for each of the various sources, the differences between the sectors are not that substantial.

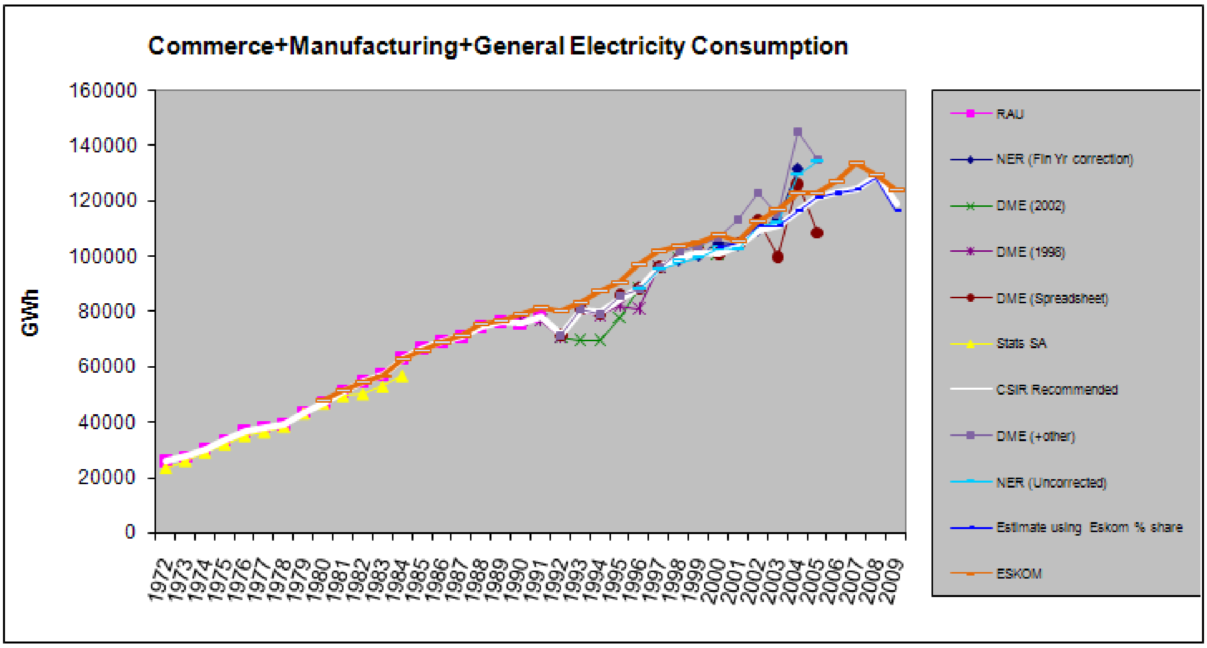


Figure 3 Comparing commercial and manufacturing sector data between different sources

Figure 4 shows that the various sources do not differ very much in terms of the pattern for the mining sector. Since Eskom supplies most of this sector directly, where differences occurred the Eskom values were considered to be a more reliable source.

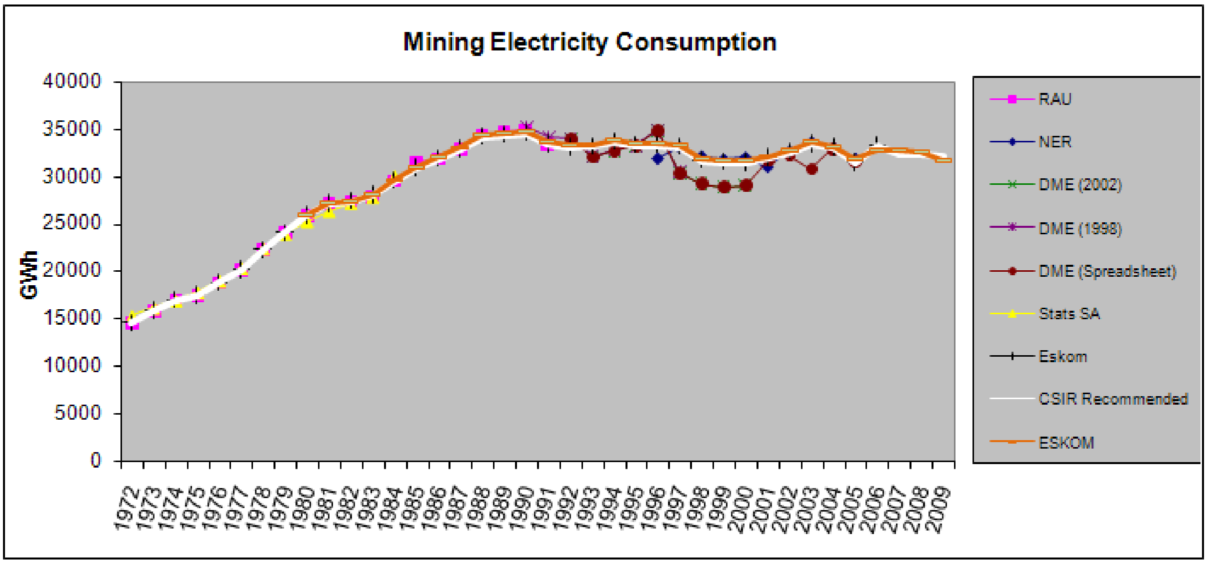


Figure 4 Comparing mining sector data between differnt sources

Figure 5 indicates that most sources showed a higher consumption in the transport sector than the estimates provided by Johan for this sector. The CSIR recommended data followed this higher pattern, but "smoothed" some of the fluctuations evident in the DME and NER data.

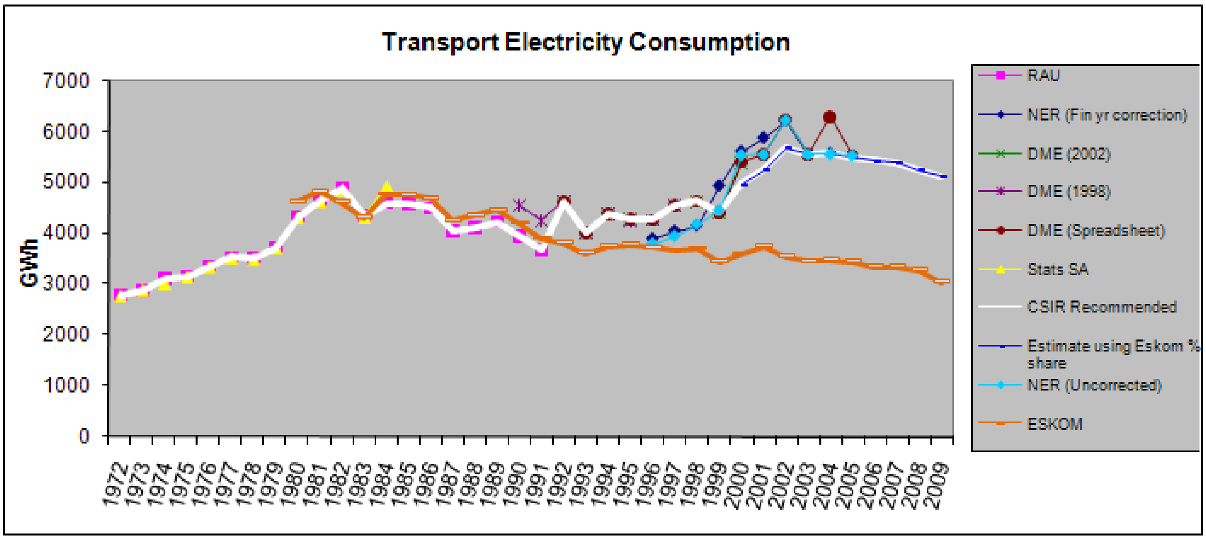


Figure 5 Comparing transport sector data between different sectors

The data used by the CSIR for each sector, namely the sources used and the data points excluded (where relevant), are summarised in table 1 below.

Table 1 Summary of electricity per sector data used

Sector	**Source**	**Excluded Years**
Agriculture	RAU data (1972-1991) Estimate from Eskom share (1993-1995, 2003-2009) NER data corrected using Eskom reports (1996-2002)	1992
Domestic	RAU data (1972-1991) DME Spreadsheets (1992-1998) NER data corrected using Eskom reports (1999-2002) Estimate from Eskom's % share (2003-2009)	
Commercial*	RAU data (1972-1991) DME Spreadsheets (1992-2000) NER data (2001, 2002) Estimate from Eskom's % share (2003-2009)	
Transport	RAU data (1972-1991) DME Spreadsheets (1992-1999) Estimate from Eskom's % share (2000-2009)	
Mining	Eskom Statistical Yearbooks (1972-1989) Eskom Annual Reports (1990-2009)	
Manufacturing*	RAU data (1972-1991) DME Spreadsheets (1992-2000) NER data including general category (2001, 2002) Estimate from Eskom's % share (2003-2009)	

*** Note that these two sectors were eventually combined into one because of difficulties with and changes in definitions between sources**

2.2.4. Data on losses

The various sources consulted by the CSIR differed both in terms of their definition of losses as well as the way they applied the loss percentages. Some sources "added on" losses after estimating each sector, while other sources "deducted" the losses in order to adjust estimates downwards.

The CSIR team decided to apply one total for losses (including both distribution and technical losses) to the combined total for all sectors in order to derive the total national consumption.

2.2.5. Data on drivers of electricity consumption

For data on potential “drivers” of electricity, historical data from various sources, but mainly the South African Reserve Bank and Stats SA, were obtained. The values for these “drivers” were compiled in consultation with Eskom as a basis for the forecasts. This ensured that a Eskom and the CSIR used a consistent basis for their electricity demand forecasts, for instance ensuring that GDP figures and the various indexes were based on the same year of standardisation. These “driver” forecasts were also provided by Eskom to the IRP 2 for the process of public participation.

2.3. Model selection

The methodology followed was to analyse data on electricity consumption as well as on aspects that describe general demographic and economic conditions which could conceivably influence consumption. Multiple regression modelling was chosen as the technique to be used for forecasting the annual consumption within the individual electricity sectors by relating such conditions (or “drivers”) to the demand in each sector.

2.3.1. How the individual sector models within the “CSIR model” were put together

A regression model “predicts” the value of one measure (variable) from that of another. There are various kinds of regression models. Simple regression predicts values from one other variable, for example when consumption is predicted from only the Gross Domestic Product (GDP). Multiple regression uses more than one variable for the prediction, for example when predicting consumption from both GDP and population. The “CSIR model” mostly used multiple regression. Regression models are well known and described in statistical literature, and many statistical software products are designed to do regression modelling. For the analysis for BHP Billiton, and the subsequent revisions, the SAS statistical software was used.

A *multiple regression model* can be expressed mathematically in the following way:

$$y = a_o + a_1 x_1 + a_2 x_2 + a_3 x_3 + ... + \varepsilon$$

where y represents the predicted quantity (called the predicted or dependent variable) and the x_i variables the different measures used for the prediction (called the predictor or independent variables). Regression modelling, simply put, is the process of finding the values of $a_0, a_1, a_2, a_3, \ldots$ (the *regression coefficients*) which would give the best prediction of y given the combination of x_i variables. The model allows for some “errors” in prediction due to random fluctuations, as represented by the *error* term ε.

The difficulty, from a practical point of view, is selecting a meaningful combination of x_i variables to use as predictor variables for the model. Not only must the predictor variables make *logical sense*, they must also give a *statistically valid* model.

In order to determine the most appropriate regression model for each sector model, various aspects were considered, the main ones being the following:

- The model must be an appropriate "fit" to the data, which means that no unnecessary driver variables must be included in the model, and that those variables that are included must provide a good overall prediction of the sectoral consumption values over the period of historical data available. The goodness of fit is measured with the R^2 (*correlation coefficient*) and adjusted R^2 measures: the closer the R^2 is to 1, the better the fit. Note that the adjusted R^2 measure deducts a "penalty" for each additional driver in the model from the R^2 value. Therefore a model with a high adjusted R^2 value implies a good fit from a model with a small number of variables used as drivers.

- When macro-economic variables are used as drivers it is often found that these variables are (strongly) related to each other. The statistical term for such interrelationships is *multi-collinearity*. This means that predictor variables included in the regression model are not only correlated with the electricity consumption, but also with one another. While the existence of multi-collinearity in a regression model might not mean that it is a "bad" model, multi-collinearity can be problematic in the interpretation of a model, since the model then contains some effects of the individual variables themselves, as well as some combined effect of the variables together. Also, multi-collinearity could cause problems when using the model for forecasts, since the model may not be reliable if these underlying relationships between the variables change over time or are not taken into account in future predictions. The amount of multi-collinearity contained within a specific regression model is measured with the *condition index* value – the lower the condition index, the better. A "rule of thumb" used in this modelling was that the condition index must be below 100 in order to indicate acceptable levels of multi-collinearity.

Unfortunately, it is very seldom that a model will be equally good on all criteria. Often, trade-offs must be made. For instance, it may be that one model does not have the largest R^2 value, but has a lower condition index than another model. For each of the individual electricity consumption sectors, a "best" model was selected by balancing the various criteria.

Although regression was chosen as the most appropriate statistical technique for this type of forecasting, it should be emphasised that the results from a regression model must be interpreted *purely as a statistically proven relationship*. A successful relationship between certain drivers and consumption *does not establish a cause and effect relationship* between them. This relationship merely describes how the observed patterns in the driver variables can "predict" the pattern in the consumption values.

2.3.2. The chosen regression models

The first step in the regression modelling was to identify a set of measures that could be considered as possible predictors for each electricity consumption sector. These measures were selected after discussions with various experts in electricity consumption forecasting. Once the possible sets of predictors were selected, regression modelling was used to determine the best regression models for forecasting consumption by sector.

The models were chosen to be statistically sound, to meet the various pre-set criteria, and to satisfy a logical understanding of the sector being forecasted. The CSIR team believes that the models are both a good fit to the collected historical data and appropriate for forecasting future demand.

Table 2 summarises the models used for each of the sector forecasts used within the "CSIR model". Data about the statistical fit of the models (summarised by the adjusted R^2), as well as the condition index that measures the amount of multi-collinearity present, are given for each model.

Table 2 Summary of regression models used per sector

Electricity sector	**Model used** (***Note: the "predictor variables" indicated in bold in each model***)	**Adjusted R^2**	**Condition index**
Agriculture	-62308+ 4905.54283×ln(**FCEH)** ***(Note: FCEH = Final Consumption Expenditure by Households, also called Private Consumption Expenditure or PCE)***	Adjusted R^2 = 0.96	N/A if only 1 variable in model
Transport	1664.13294 + 39.45163×**mining index**	Adjusted R^2 = 0.80	N/A if only 1 variable in model
Domestic	-19196 + 887.76683×**population** + 0.01598×**FCEH**	Adjusted R^2 = 0.98	CI = 38.5
Commerce & manufacturing	-332660 + 102198×ln(**population**) + 557.95182×**manufacturing index**	Adjusted R^2 = 0.98	CI = 97.6
Mining	5619.5596 + 78.93746×**platinum production index** + 120.94321×**coal production index** + 0.12249×**gold ore treated**	Adjusted R^2 = 0.91	CI = 36.8

It should be noted that the regression coefficients (numerical values linked to the predictor variables in the models in table 2) must not be interpreted too literally. For instance, in the model for the Commerce and Manufacturing sector consumption, the regression coefficients are -332660, 102198, and 557.95182. It cannot be said that population is a more important predictor than the manufacturing index just because of the magnitude of the regression coefficient of population (102198) compared to that of the manufacturing index (557.95). The magnitude of the regression coefficients is not only determined by the importance of the predictor variables, but also *by the units in which the predictors are measured*. If, for instance, the manufacturing index was changed from a base year of 2005 (as used in this model) to a 2010 base year, the value of the coefficient would change even though the statistical relationship remains unchanged.

2.3.3. Understanding the chosen regression models

The best way to understand the models is to look at the predictor variables included in each model. The predictor variables make sense in the models since they are associated with certain conditions that would be expected to affect consumption in a specific sector, but care must be taken *not to assume automatically that these are causal relationships*.

The reason for this is that most of the measures used in the model were obtained from publicly available sources and were not collected specifically for this process. They can be considered to be **proxy variables** for some of the "real" underlying drivers. For example, the amount of electricity consumed in the manufacturing sector is determined by the types of processes that require electricity, the electricity intensity of those processes and the number of times such processes are required during production processes in a month. However, data on all of these aspects are not available. Therefore, the "CSIR model" bases the consumption in this sector on the manufacturing index, which measures the volumes of goods produced by this sector, irrespective of how much electricity is used to produce each of the various types of goods. It is therefore not a direct measure of electricity demand, but one assumes that the patterns of production and the patterns of electricity usage will roughly correspond to each other, and this is picked up by the model.

Finally, note that the regression models were developed ***very specifically for the South African situation*** and were based on historical data collected for the South African situation only. At no point in this modelling procedure has an attempt been made to develop generic models for forecasting consumption – all models are based on historical data collected from public domain sources.

2.4. Track record of forecasts

The forecasts developed for BHP Billiton using this methodology were based on scenarios and therefore the "accuracy" of the forecasts depends to some extent on the "accuracy" of the scenarios. However, it is still possible to compare the numerical values of national consumption forecasts, from each scenario, to the actual values recorded for the national consumption.

The forecasts were first developed in 2003, and then updated twice: once in 2005 and once in 2008. It should be pointed out that at the time of the 2008 updates, data for 2007 consumption were not yet available, and therefore the forecasts were based on data that excluded the first load shedding episodes in 2007.

For the initial forecasts, as well as each of the updates, BHP Billiton provided four scenarios, ranging from an "overly positive" to a "very negative" scenario. It should be noted, however, that the negative scenarios very seldom included a negative GDP growth rate. In order to use short labels that are fairly clear, the scenarios are indicated with "original" for the initial forecasts, "Upd1" for the first set of updates and "Upd2" for the second set of updated forecasts. Within each of these sets, the forecasts are labelled "pos++" for the most positive scenario, "pos" for the scenario that was slightly less positive, "neg" for the relatively negative scenario and "neg- -" for the most negative scenario out of the set.

The graph in figure 6 illustrates how the past forecasts developed for BHP Billiton compared to the actual values recorded. (The graph refers to total national consumption.) Although it is difficult to compare the forecasts and the actual values directly, since the forecasts were developed to provide a range of both positive and negative scenarios, two aspects are immediately noticeable from the graph. Firstly, it does seem as if the actual values fell more or less within the various forecasted ranges, although the actual values seemed to be towards the lower end of the range of forecasts. The second very noticeable aspect is the way in which the 2008 and 2009 years fell below any of the forecasted values. This seems to indicate that the patterns observed over the past few years have not been typical of values experienced before, and therefore indicates the necessity of adjusting forecasts by bringing these values into the data on which the models are based.

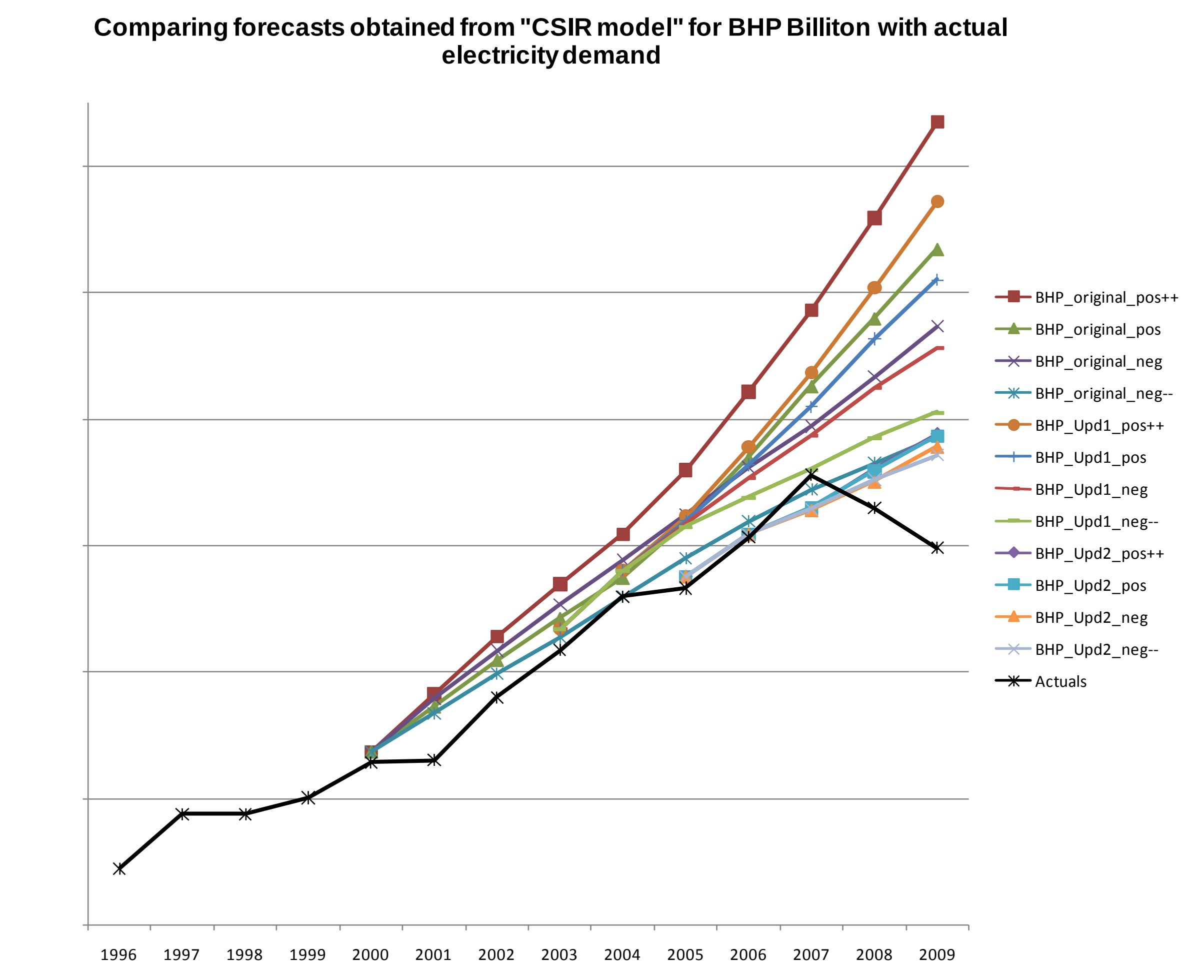


Figure 6 Track record of "CSIR model"

Table 3, below, summarises the percentage differences between the BHP Billiton forecasts and the actual recorded national consumption figures. This is provided since some of the lines on the graph in figure 6 overlap with each other, and therefore the differences are difficult to see for each of the individual forecasted scenarios. It is clear from this table that the forecasts are not very far from the actuals for most years, but that the differences become quite large over 2008 and 2009.

Table 3 Percentage differences between "CSIR model" forecasts and actual data

Year	BHP Billiton original model				BHP Billiton: 1st update				BHP Billiton: 2nd update			
	pos++	pos	neg	neg--	pos++	pos	neg	neg--	pos++	pos	neg	neg--
2001	5.3	4.3	4.9	3.7								
2002	4.7	2.8	3.6	1.8								
2003	4.9	2.4	3.3	0.9	1.6	1.6	1.6	1.6				
2004	4.4	1.3	2.6	-0.1	1.9	1.9	1.9	1.9				
2005	8.4	4.7	5.2	2.1	5.1	4.8	4.6	4.4	0.8	0.8	0.8	0.8
2006	9.9	5.4	4.8	1.1	6.1	4.9	4.0	2.7	0.2	0.2	0.2	0.2
2007	10.8	5.8	3.2	-0.9	6.7	4.5	2.6	0.4	-2.2	-2.2	-2.3	-2.2
2008	19.5	12.7	8.8	3.0	14.8	11.4	8.0	4.8	2.6	2.5	1.8	1.9
2009	29.3	20.6	15.2	7.7	23.8	18.5	13.8	9.4	7.9	7.7	7.0	6.7

3. Forecasting results

This section provides the forecasts obtained as inputs to the IRP 2 process, namely demand forecasts for national consumption of electricity for the period 2010 – 2034.

3.1. Forecasted driver values used

In the IRP, three growth scenarios were specified for use in forecasting, namely a “Low”, “Moderate” and “High” forecast. Expected values for the different driver variables were linked to each scenario. These scenarios represent the mechanism for introducing uncertainty regarding the future values of the drivers into the electricity forecasts. However, the differences between the scenarios related to economic variables, namely the expected GDP growth rates as well as the expected growth rate in the Final Consumption Expenditure of Households (FCEH) values. For other drivers, namely population size, as well as the relevant manufacturing and mining indexes, only one set of forecasts were available.

The models were used by inserting the expected future values for each of the driver variables into the relevant sectoral models in order to obtain a forecast for each sector. The sectoral forecasts were aggregated and then adjusted for losses in order to obtain a forecast for national consumption.

The following three tables summarise the “driver” values used for the forecasts. The first table (table 4) provides the driver values that did not change between scenarios, while the second one (table 5) provides the growth percentages on the macro economic variables that differed between the scenarios. Note that the annual data is provided as per calendar year, not per financial year, in both tables.

Table 4 Driver forecasts that were the same between different scenarios

Year	**Population (in millions)**	**Mining production index, excluding gold *(base year = 2005)***	**PGM mining index *(base year = 2005)***	**Coal mining index *(base year = 2005)***	**Gold ore treated**	**Line losses (% of sectoral total)**
2007	48.6	100.2	100.0	101.1	56362	11.86%
2008	49.1	96.5	90.7	103.1	54312	11.94%
2009	49.3	90.1	89.2	101.8	50469	12.02%
2010	49.9	93.2	94.6	102.8	49428	12.10%
2011	50.1	95.6	98.3	103.9	48409	12.17%
2012	50.3	97.6	102.3	104.4	47411	12.24%
2013	50.6	99.3	105.3	104.9	47411	12.31%
2014	50.8	101.0	108.5	105.4	47411	12.37%
2015	51.0	102.8	111.8	105.9	47411	12.42%
2016	51.2	104.8	115.1	106.5	47411	12.47%
2017	51.4	105.5	115.7	107.0	47900	12.52%
2018	51.6	106.2	116.3	107.5	48394	12.56%
2019	51.8	106.9	116.8	108.1	48892	12.61%
2020	52.0	107.7	117.4	108.6	49397	12.64%
2021	52.2	108.4	118.0	109.2	48898	12.68%
2022	52.4	109.0	118.6	109.7	48160	12.71%
2023	52.7	109.6	119.2	110.3	47918	12.74%
2024	52.9	110.2	119.8	110.8	47678	12.77%

Year	Population (in millions)	Mining production index, excluding gold *(base year = 2005)*	PGM mining index *(base year = 2005)*	Coal mining index *(base year = 2005)*	Gold ore treated	Line losses (% of sectoral total)
2025	53.1	110.8	120.4	111.4	47438	12.80%
2026	53.3	111.4	121.0	111.9	47200	12.82%
2027	53.5	112.0	121.6	112.5	46962	12.85%
2028	53.7	112.6	122.2	113.0	46726	12.87%
2029	53.8	113.2	122.8	113.6	46961	12.89%
2030	54.0	113.8	123.4	114.2	47197	12.91%
2031	54.2	114.5	124.0	114.7	47434	12.92%
2032	54.3	115.1	124.7	115.3	47673	12.94%
2033	54.5	115.7	125.3	115.9	47912	12.95%
2034	54.6	116.3	125.9	116.5	48153	12.97%

Table 5 Macro economic driver forecasts that differed between scenarios

Year	GDP growth rate % per scenario			FCEH growth rate % per scenario		
	Low	Moderate	High	Low	Moderate	High
2009	-1.7	-1.7	-1.7	5.52	5.5	5.5
2010	1.5	2.5	3.5	2.39	2.4	2.4
2011	2.7	3.7	4.7	-3.14	-3.1	-3.1
2012	3	4	5	-6.7	-5.3	-2.7
2013	3	4	5	0.3	2.3	4.7
2014	3	4	5	5.5	7.3	8.2
2015	3.5	4.5	5.5	3.1	5.1	5.7
2016	4	5	6	3.1	4.5	4.5
2017	4	5	6	3.5	5.1	5.6
2018	4	5	6	4.4	5.9	6.7
2019	4	5	6	4.6	6.3	6.8
2020	4	5	6	4.5	6.2	6.8
2021	3.9	4.9	5.9	4.7	6.2	6.8
2022	3.9	4.9	5.9	4.6	6.2	6.8
2023	3.8	4.8	5.8	4.5	6.0	6.8
2024	3.8	4.8	5.8	4.6	6.1	7.0
2025	3.8	4.8	5.8	4.3	5.8	6.6
2026	3.7	4.7	5.7	4.2	5.8	6.6
2027	3.7	4.7	5.7	4.5	6.0	6.6
2028	3.7	4.7	5.7	4.3	5.8	6.4
2029	3.6	4.6	5.6	4.4	5.9	6.4
2030	3.6	4.6	5.6	4.4	5.9	6.6
2031	3.5	4.5	5.5	4.2	5.7	6.4
2032	3.5	4.5	5.5	4.2	5.7	6.4
2033	3.4	4.4	5.4	4.3	5.5	6.3
2034	3.4	4.4	5.4	4.0	5.4	6.3

The information in table 6 and illustrated in figure 7, below, relate to scenarios for the Manufacturing Index. An initial set of forecasts were developed for the manufacturing production volume index in which a slow recovery of manufacturing production was anticipated after the recession, i.e. in which the index increased at a slow rate. However, there was a concern that the forecasted pattern may be too conservative, and that the industrial sector will recover more rapidly after the recession period. Therefore, it was considered prudent to create an alternative set of scenarios which included a more rapid recovery pattern for this index. The Manufacturing Index plays an important role in forecasting the commerce, manufacturing and general sector which is the biggest sector, and therefore has the biggest effect on total consumption of all the sectors. The forecast for this sector was therefore considered important, and consequently adding additional scenarios to allow for more flexibility in the forecasts for this sector was considered advisable. Three additional scenarios were therefore added in order to couple the “rapid recovery” scenario on the Manufacturing Index to each of the “low”, “moderate” and “high” growth scenarios.

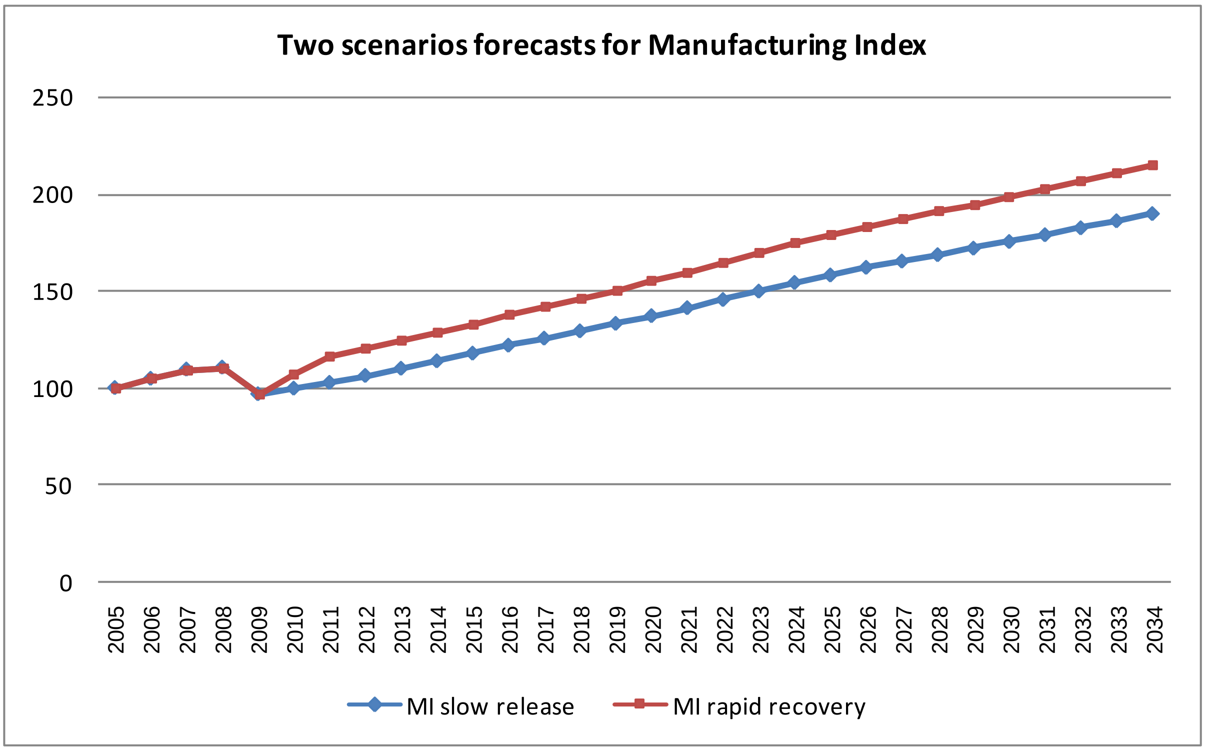


Figure 7 Manufacturing index

Table 6 Scenario values for the Manufacturing Index

Year	Manufacturing Index (slow increase scenario)	Manufacturing Index (rapid recovery scenario)
2007	109.6	109.6
2008	110.6	110.6
2009	96.8	96.8
2010	99.7	107.3
2011	102.7	115.9
2012	106.3	120.0
2013	110.0	124.2
2014	113.9	128.5
2015	117.8	133.0
2016	122.0	137.7
2017	125.6	141.8
2018	129.4	146.1
2019	133.3	150.4
2020	137.3	155.0

Year	Manufacturing Index (slow increase scenario)	Manufacturing Index (rapid recovery scenario)
2021	141.4	159.6
2022	145.6	164.4
2023	150.0	169.3
2024	154.5	174.4
2025	158.4	178.8
2026	162.3	183.2
2027	165.6	186.9
2028	168.9	190.6
2029	172.3	194.5
2030	175.7	198.3
2031	179.2	202.3
2032	182.8	206.4
2033	186.5	210.5
2034	190.2	214.7

3.2. Demand forecasts obtained

The forecasts obtained for each of the "low", "moderate" and "high" scenarios are listed in table 7.

Table 7 National electricity demand: historical data and forecasts

Year	Annual electricity demand forecasts (GWh) per scenario, provided per calendar year:			Electricity available for distribution in South Africa in GWh *(Totals obtained from Statistics SA)*
	Low	Moderate	High	
2005	222211	222211	222211	223255
2006	228730	228730	228730	231323
2007	236736	236736	236736	241170
2008	238987	238987	238987	235924
2009	229302	229302	229302	229599
2010	231916	232290	232477	N/A
2011	235227	236082	237007	N/A
2012	240093	241444	242720	N/A
2013	244531	246437	247884	N/A
2014	249066	251425	252930	N/A
2015	253841	256702	258445	N/A
2016	258838	262244	264337	N/A

Year	Annual electricity demand forecasts (GWh) per scenario, provided per calendar year:			Electricity available for distribution in South Africa in GWh *(Totals obtained from Statistics SA)*
	Low	**Moderate**	**High**	
2017	263386	267450	270038	N/A
2018	268046	272812	275748	N/A
2019	272858	278337	281657	N/A
2020	277793	284033	287776	N/A
2021	282720	289768	294209	N/A
2022	287780	295701	300801	N/A
2023	292931	301807	307459	N/A
2024	298212	308103	314433	N/A
2025	303241	314207	321186	N/A
2026	308232	320332	328012	N/A
2027	312869	326226	334578	N/A
2028	317608	332300	341527	N/A
2029	322451	338544	348650	N/A
2030	327395	344972	356110	N/A
2031	332351	351372	363666	N/A
2032	337282	357965	371514	N/A
2033	342245	364666	379391	N/A
2034	347311	371570	387514	N/A

The values obtained for these forecasts are illustrated graphically in figure 8, while the forecasts for the five sectors making up the total consumption is provided in five separate graphs in figure 9.

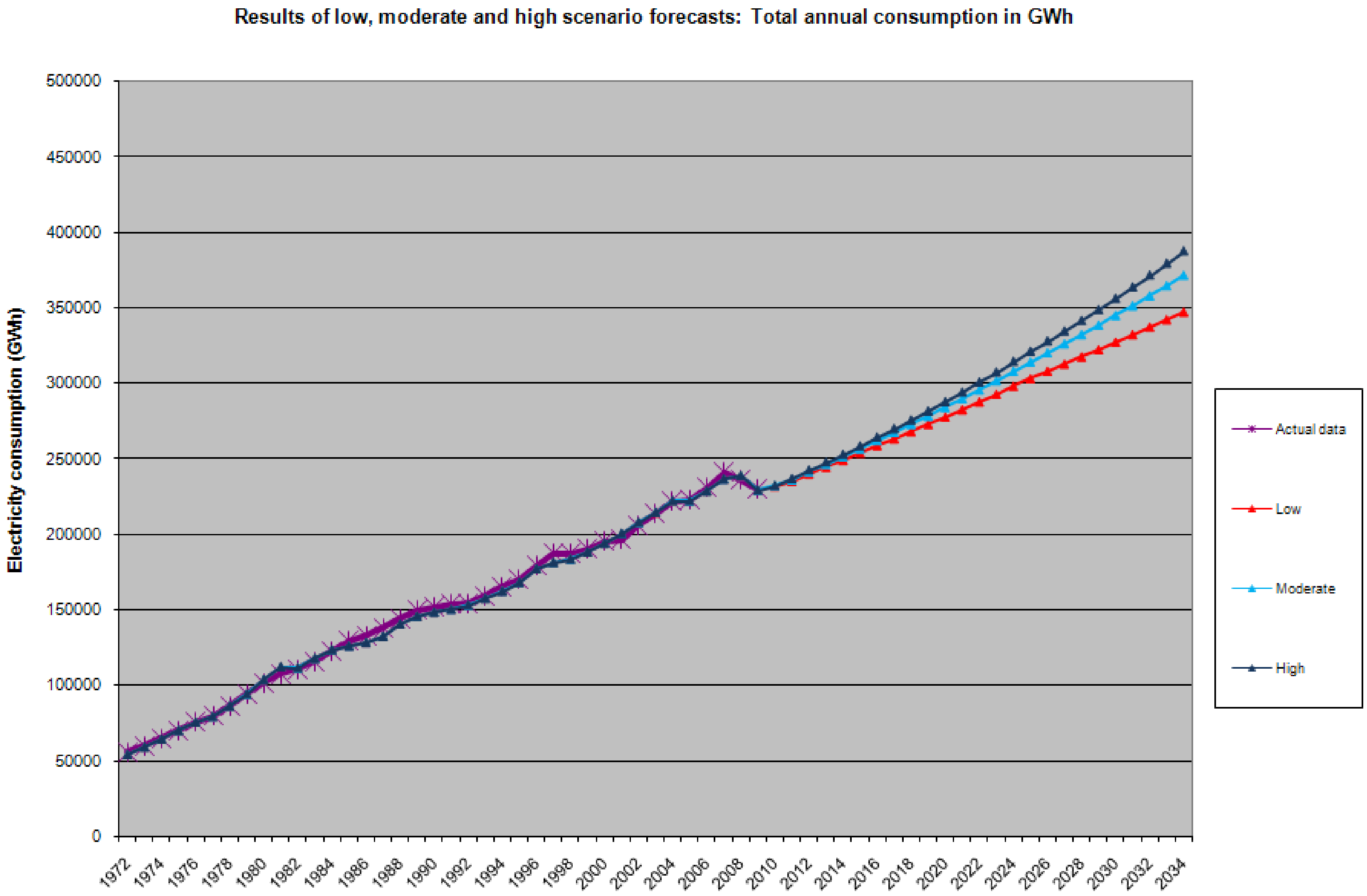
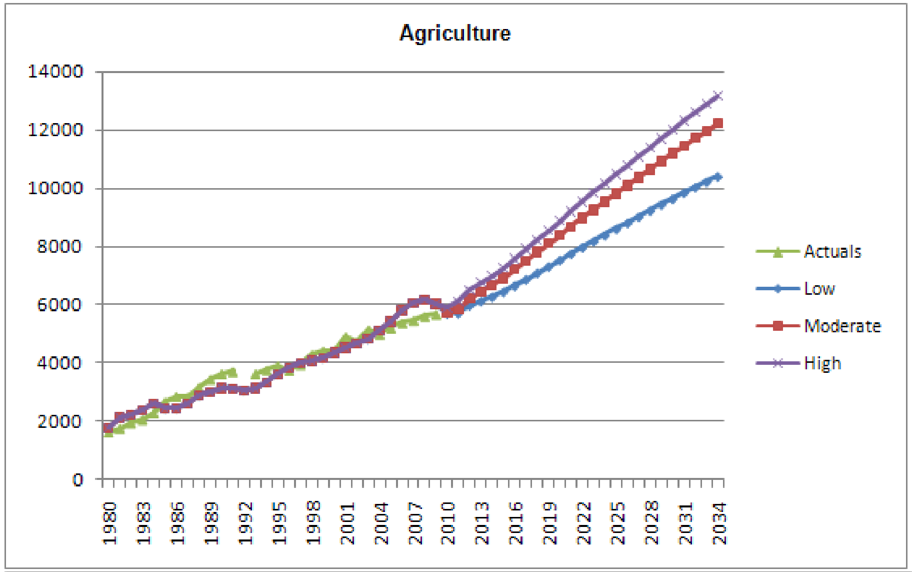
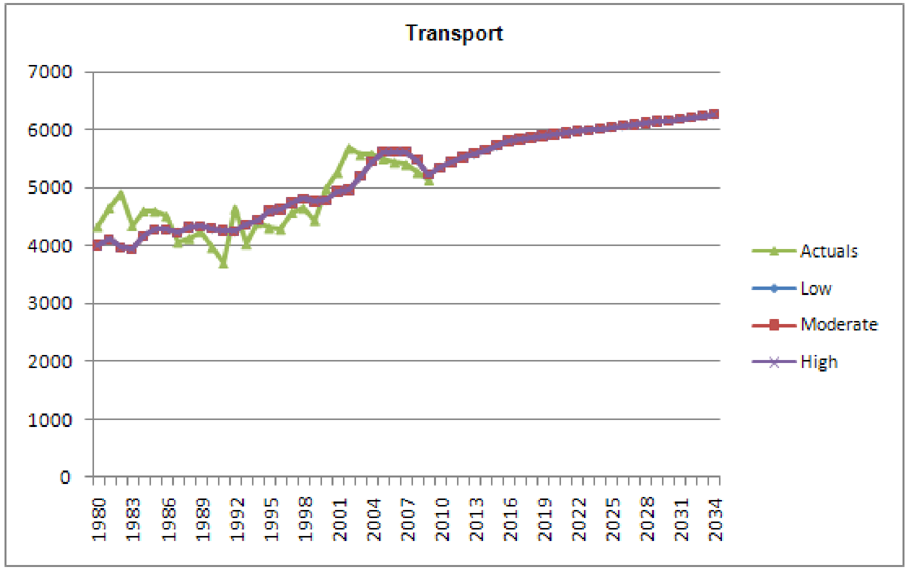
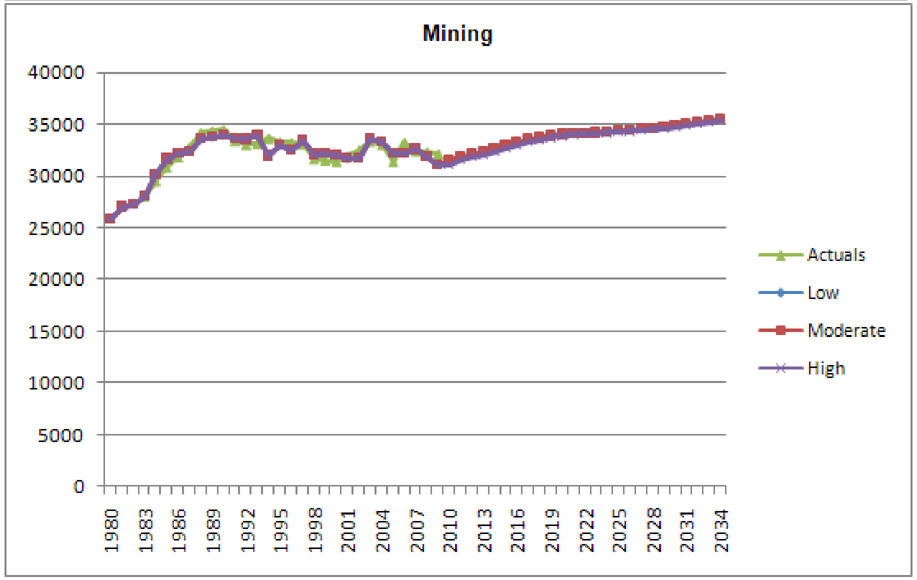
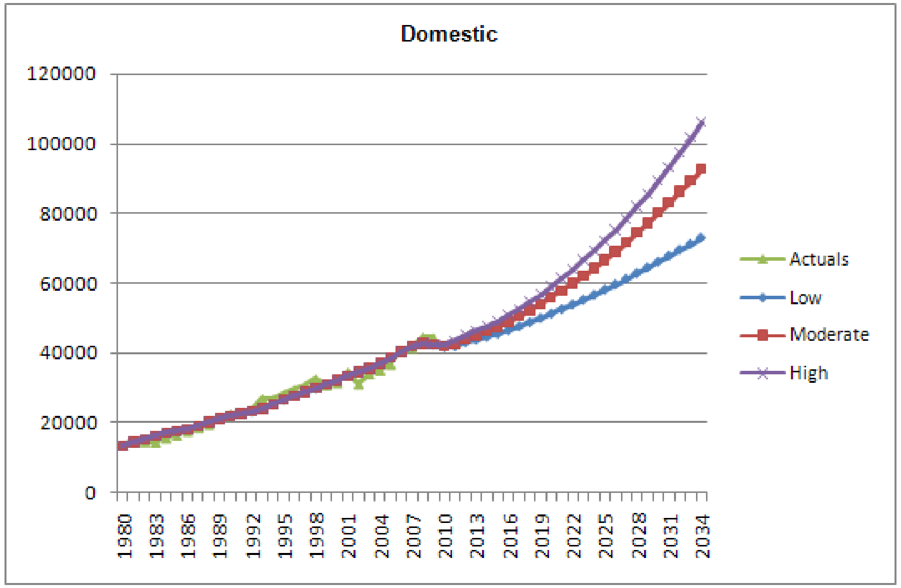


Figure 8 Forecasts for national consumption of electricity using the "CSIR model"


Agriculture
14000
12000
10000
8000
6000
4000
2000
0
1980 1983 1986 1989 1992 1995 1998 2001 2004 2007 2010 2013 2016 2019 2022 2025 2028 2031 2034
Actuals
Low
Moderate
High


Transport
7000
6000
5000
4000
3000
2000
1000
0
1980 1983 1986 1989 1992 1995 1998 2001 2004 2007 2010 2013 2016 2019 2022 2025 2028 2031 2034
Actuals
Low
Moderate
High


Mining
40000
35000
30000
25000
20000
15000
10000
5000
0
1980 1983 1986 1989 1992 1995 1998 2001 2004 2007 2010 2013 2016 2019 2022 2025 2028 2031 2034
Actuals
Low
Moderate
High


Domestic
120000
100000
80000
60000
40000
20000
0
1980 1983 1986 1989 1992 1995 1998 2001 2004 2007 2010 2013 2016 2019 2022 2025 2028 2031 2034
Actuals
Low
Moderate
High

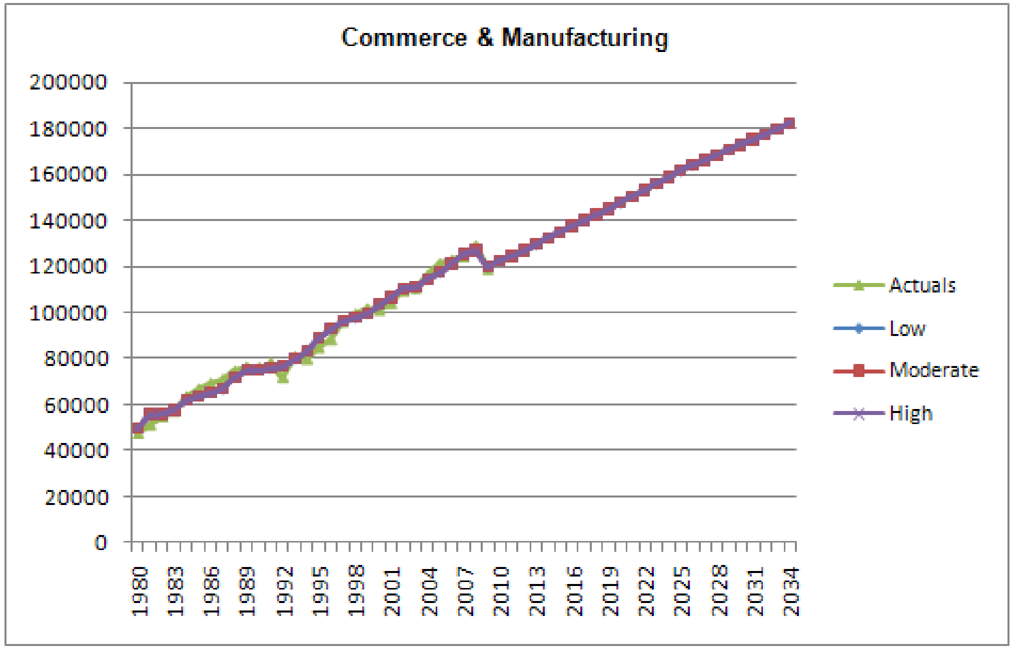


Figure 9 Forecasted values for the 5 electricity sectors

3.3. Alternative forecasts

Using the alternative set of forecasts for the Manufacturing Index, as explained in section 3.1, three additional scenario forecasts were developed. These are presented in table 8 and illustrated in the graph in figure 9.

Table 8 Alternative forecasts for national electricity demand, using "Eskom sectors"

Year	**Annual electricity demand forecasts (GWh) per scenario, provided per calendar year:**			**Electricity available for distribution in South Africa in GWh *(Totals obtained from Statistics SA)***
	Low + rapid increase on MI	**Moderate + rapid increase on MI**	**High + rapid increase on MI**	
2005	222211	222211	222211	223255
2006	228730	228730	228730	231323
2007	236736	236736	236736	241170
2008	238987	238987	238987	235924
2009	229302	229302	229302	229599
2010	236684	237056	237259	N/A
2011	243515	244378	245326	N/A
2012	248665	250010	251315	N/A
2013	253424	255327	256802	N/A
2014	258264	260598	262131	N/A
2015	263368	266222	267974	N/A
2016	268684	272083	274178	N/A
2017	273564	277617	280185	N/A
2018	278503	283261	286181	N/A
2019	283660	289120	292424	N/A
2020	288888	295136	298859	N/A
2021	294177	301235	305630	N/A
2022	299567	307503	312561	N/A
2023	305098	313960	319581	N/A
2024	310738	320631	326939	N/A
2025	316089	327069	334032	N/A
2026	321367	333497	341159	N/A

Year	Annual electricity demand forecasts (GWh) per scenario, provided per calendar year:			Electricity available for distribution in South Africa in GWh *(Totals obtained from Statistics SA)*
	Low + rapid increase on MI	**Moderate + rapid increase on MI**	**High + rapid increase on MI**	
2027	326269	339646	348012	N/A
2028	331284	345998	355251	N/A
2029	336392	352517	362650	N/A
2030	341613	359249	370407	N/A
2031	346874	365955	378276	N/A
2032	352112	372818	386464	N/A
2033	357369	379839	394661	N/A
2034	362741	387018	403090	N/A

The values obtained for these forecasts are illustrated graphically in figure 10.

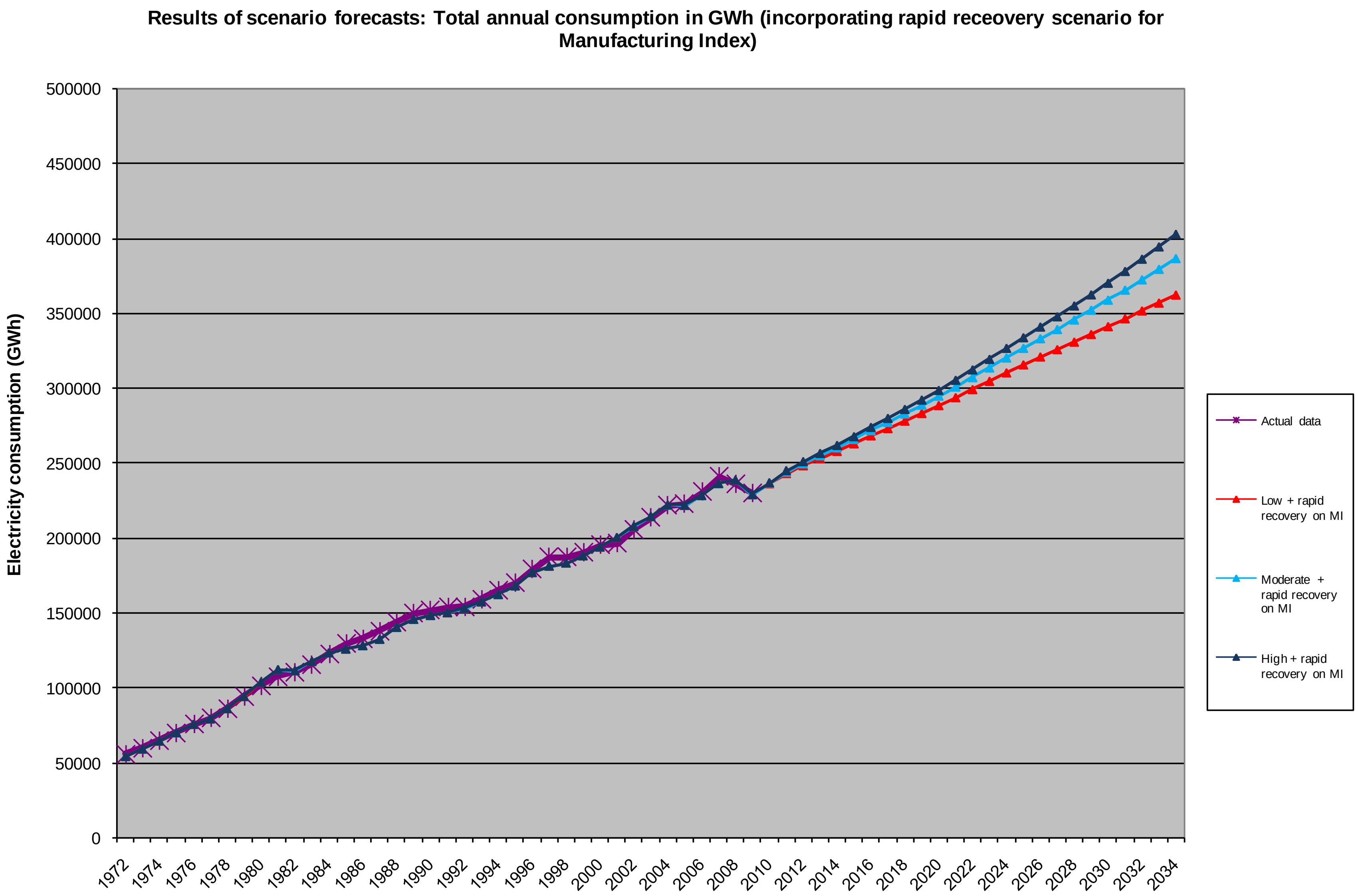


Figure 10 Alternative forecasts using an alternative set of forecasts for the Manufacturing Index

4. Final remarks: applicability of model results

The "CSIR model" forecasts the national demand for electricity at a macro level, based on data relating to macro-level economic and demographic indicators. This basic approach can be contrasted to other methods used for forecasting electricity demand in order to assess the relative advantages and disadvantages of the various approaches with respect to one another. Firstly, the "CSIR model" is compared to the approach followed internally by Eskom to produce long-term forecasts, and then to other approaches such as econometric modeling. Finally, these comparisons lead to some general comments about the need for reliable data on which to base electricity demand forecasting.

4.1. Relative advantages and disadvantages of the "CSIR model" vs Eskom official forecasting process

Eskom has for many years produced long-term forecasts for electricity demand, mostly for internal planning purposes. The methodology for these forecasts was developed by Johan Prinsloo, and has been slightly refined and modified over the years to link up well with Eskom needs and requirements.

Although the CSIR team has not extensively analysed or audited the forecasting process, we have had a number of meetings with Johan Prinsloo and Koos Potgieter in order to understand the methodology that is being employed to generate the forecasts. In essence, it is a "bottom-up" approach, forecasting the needs of individual customers, customer groupings and regional development initiatives, and then aggregating these forecasts up to a national level. Putting together the forecasts require a large amount of time and expertise, as well as data about historical and planned activities of customers, and information on specific needs within a region in terms of either expansion or decline. It also requires large amounts of customer data internal to Eskom (most of these probably confidential), as well as judgement regarding how to use these patterns for extrapolation into the future. However, once the forecasts have been completed, it contains a large amount of detail and can be broken down in any number of ways.

The "CSIR model", in contrast, follows more of a "top-down" approach and does not require any customer-level data. In fact, the "CSIR model" requires no information from Eskom other than what it publishes in its Annual Reports and makes publicly available regularly, i.e. does not require customer-specific data. It combines data from Eskom with other public domain sources and tries to model the demand by using as few as possible economic and demographic indicators. One obtains forecasts for electricity usage from forecasts of these indicators, and a different set of indicators can be obtained very quickly from a new set of forecasts indicator variables. The advantage is that one can more easily test different scenarios with the "CSIR model", but the disadvantage is that the "CSIR model" cannot give information regarding the detailed breakdown of these forecasts than would be possible with the Eskom methodology.

The CSIR team is of the opinion that Eskom has developed a forecasting methodology that maximises the amount of information that can be obtain out of its internal data and systems, and that it is probably is the best way to carry out a bottom-up approach. However, it approaches the forecasting problem from a completely different point of view than the "CSIR model". Therefore, forecasts emanating from the Eskom methodology and those emanating from the "CSIR model" cannot be compared at any level other than the national level, i.e. on the overall forecasted values emanating from both methodologies. The two methods can never be completely consistent and attempts to align them closer to each other would not be meaningful, as it may remove the comparative advantages these methods have over one another. However, one would want the overall results from the two approaches to be fairly comparable at the national level. Ideally therefore, the methods should be used in parallel, with the “CSIR model” used to test macro-level changes and the Eskom method used to provide the required details down to a customer level as well as providing the required inputs into the operational needs of Eskom.

4.2. Relative advantages and disadvantages of the “CSIR model” vs econometric modelling

Media attention to issues of Eskom’s potential inability to provide enough electricity to satisfy demand, as well as the proposed price increases have sparked some interest from analysts, who have put forward alternative methods to forecast electricity demand. Two examples of these were Lizelle van Wyk from Exxaro [13] and Roula Inglesi [14] who both provide a set of forecasts that attempt to incorporate price elasticity into the modelling of electricity demand. Both of them used econometric methods to produce the forecasts, in contrast to the “CSIR model” that used a statistical approach.

Although there is no real difference between the techniques used in statistics and those used in econometrics, econometrics applies statistical techniques to economic theory. In other words, an econometric method uses statistical techniques to test a specific economic theory, in these cases specifically the price elasticity of electricity demand. The statistical approach followed by the “CSIR model” assumed no specific underlying theory, but incorporated all available public domain data. The “CSIR model” also did not specifically exclude any potential drivers of electricity demand theoretically. Any aspect could be included in the model as long as it can be quantified, including aspects such as electricity intensity or elasticity of price or income. However, reliable data for these factors, at the level of electricity demand per sector, could not be found and therefore these were not explicitly included in the “CSIR model”.

The CSIR team is of the opinion that it is not possible to model price elasticity successfully at the national level only, such as was done by Inglesi and Van Wyk. Although the work by these authors provides a good starting point for econometric modelling, price elasticity cannot be assumed to have an equal effect across all electricity sectors, just as the average price is not the same across all these sectors. Instead, a sectoral approach such as that followed by Ziramba [15] may provide a more accurate picture of the effect of price or per capita income on demand.

Furthermore, although it is possible to argue that price will affect demand, especially given the dramatic price increases proposed by Eskom, there does not seem to be sufficient historical data to allow a thorough investigation into this aspect. Such a dramatic price increase has never occurred in the past, and therefore there is no historical data available to use as a basis for forecasting its effect quantitatively. The slight dip in the domestic sector in 1999 has been previously attributed to price increases by some data sources but beyond this no further data on price elasticity is available.

Finally, it can be argued that although the “CSIR model” does not **explicitly** use various variables that could have a big effect on the demand for electricity, it does model many aspects **implicitly** by modeling the overall patterns. In the absence of reliable, sustainable sources of publicly available information on many aspects that drive demand, models that implicitly take into account the combined effect of these aspects in an overall pattern may be the best way to use the available data.

4.3. The importance of reliable data

It is of the utmost importance to have “good” forecasts for the demand for electricity so that it can guide planning of electricity supply. Although no quantitative forecasts can be expected to be entirely accurate, one can put some measure of trust in forecasts which are carried out using acceptable forecasting principles and are based on accurate information.

Another way to improve forecasts would be to use different forecasting approaches, and then to compare results between the different methods in order to highlight areas of uncertainty or identify potential improvements to the various methodologies. If the different methods seem to provide similar forecasts, then that provides a further measure of trust in the reliability of the forecasts. Therefore, when applying processes to derive additional forecasts, such as from the “CSIR model” or from econometric approaches, the aim should not be to replace Eskom’s forecasts, but rather to provide additional support and functionality to Eskom’s forecasting methods and results.

However, **it is of major concern that there is so little data on electricity usage from sources other than Eskom**. If other methodologies are to be used to supplement the Eskom forecasts, then ideally those methodologies should not simply be applied to Eskom’s internal data, but should make use of a range of different sources, including sources outside Eskom. During the work CSIR did for BHP Billiton, NERSA and the Department of Energy collected data, but neither of these two sources have published survey information on sectoral electricity use since 2006. The forecasts reported in this document were based on estimates of sectoral consumption using data from Stats SA and Eskom, and extrapolating some past trends. While such estimates should be good enough for the purpose of producing the forecasts contained in this document, it is not ideal to continue modelling on estimated rather than actual patterns.

To improve both forecasting as well as policy development regarding electricity usage in South Africa, figures on Eskom sales and national electricity consumption will not be sufficient, since it does not capture enough details about consumption patterns of important usage groupings. Modelling electricity demand at the sector level provides much more informative and accurate results than doing models at the national level, since many of the patterns observed in individual sectors disappear when aggregated into total consumption. In addition, sectoral demand information will become increasingly important if any study on the effect of price increases on electricity demand is to be carried out in the short to medium future. Attention should be paid by agencies such as NERSA, Department of Energy or Stats SA to produce a reliable, accurate and sustainable source of such information.

5. References

[1] CJ Cooper and DJ Kotze, **Energy Projections for South Africa Vol 8,** 1992, Institute for Energy Studies, Rand Afrikaans University (RAU), (Appendix A).
This reference contains data on electricity consumption per sector and data on economic indicators for South Africa for the period 1972-1991.

[2] CJ Cooper, **Digest of South African Energy Statistics 1998,** Department of Minerals and Energy (DME), (section 6).
This reference contains data on electricity consumption per sector in South Africa for the period 1990-1997.

[3] **Digest of South African Energy Statistics 2002,** Department of Minerals and Energy (DME), (section 6).
This reference contains data on electricity consumption per sector in South Africa for the period 1992-2000.

[4] **Digest of South African Energy Statistics 2006,** Department of Minerals and Energy (DME) (section 6).
This reference contains data on electricity consumption per sector in South Africa for the period 1992-2004.

[5] **Energy Balance Spreadsheets,** Department of Minerals and Energy (DME).
This reference contains a detailed breakdown of data on electricity consumption per sector in South Africa for the period 1992-2000.

[6] **Electricity Supply Statistics,** 1996, 1998-2001, National Electricity Regulator (NER).
This reference contains annual electricity supply data and electricity sales per sector.

[7] **South African Energy Statistics 1950-1989, No 1,** 1990, National Energy Council (NEC), (chapter 4).
This reference contains sectoral consumption of electricity for the period 1950-1989.

[8] **South African Energy Statistics 1950-1993, No 2,** 1995, Department of Minerals and Energy Affairs (DMEA) and Eskom Marketing Intelligence, (chapter 4).
This reference contains sectoral consumption of electricity for the period 1950-1993.

[9] **Eskom Annual Reports,** 1990-2009, Eskom.
Annual data on electricity consumption in the mining sector was of interest in this reference.

[10] **Eskom Statistical Yearbook,** 1996, Eskom.
This reference contains data on electricity consumption per mining activity and for the total mining sector for the period 1972-1996.

[11] Website of the International Energy Agency (iea), at the specific address http://www.iea.org/stats/electricitydata.asp?COUNTRY_CODE=ZA, accessed May 2010

[12] Website of the International Energy Agency (iea), at the specific address http://wds.iea.org/wds/pdf/documentation_wedbes.pdf, accessed May 2010

[13] L van Wyk, **Regression Analysis to Estimate South Africa's Electricity Demand**, 17 November 2009, Presented at the *"Forecasting in an ever changing environment"* Conference organised by Eskom.

[14] R Inglesi, **Aggregate electricity demand in South Africa: Conditional forecasts to 2030**, 2010, *Applied Energy*, Vol 87, pp 197 – 204

[15] E Ziramba, **The demand for residential electricity in South Africa**, 2008, *Energy Policy*, Vol 36, pp 3460 - 3466