Survey on economy-wide fuel consumption with emphasis on demand-side sectors

environmental affairs

Department:
Environmental Affairs
REPUBLIC OF SOUTH AFRICA

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

On behalf of:
Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety
of the Federal Republic of Germany

IMPRINT

Prepared for

Department of Environmental Affairs

Prepared by

Top Quartile - Project Services Africa
Energy Research Centre (ERC) University of Cape Town

Contributions by

University of Johannesburg (UJ)\City of Cape Town
City of Tshwane
City of Johannesburg
Western Cape Provincial Government
National Cleaner Production Centre (NCPC)
SASOL (Pty) Ltd
Airports Company of South Africa (SAPIA)
South African Petroleum Industry Association (SAPIA)
Statistics South Africa (StatsSA)
Human Sciences Research Council (HSRC)
Transnet
Sustainable Energy Africa (SEA)

Layout by

Twaai Design

Acknowledgements

The authors are thankful to the Department of Labour for their support and provision of data from the Vessels Under Pressure Registry. A special thanks the Department of Energy for support, provision of crucial information as well as contributing to the design of this study.

A special thanks in this regard to Jongikhaya Witi and Lungile Manzini for their leadership and contribution to the design of this project.

The authors would also like to thank the reviewers of the report.

Contact information

Department of Environmental Affairs
Environment House
473 Steve Biko Street
Arcadia
Pretoria 0001
South Africa

Tel: +27 12 399 9154 / Fax: +27 (0) 86 615 4321
Email: jwiti@environment.gov.za

Published in South Africa March 2016

Survey on economy-wide fuel consumption with emphasis on demand-side sectors

Final Report

environmental affairs
Department:
Environmental Affairs
REPUBLIC OF SOUTH AFRICA

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

On behalf of:
Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety
of the Federal Republic of Germany

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AACEI	Association for the Advancement of Cost Engineering International
ACSA	Airports Company of South Africa
AQM	Air Quality Management
BLNS	Botswana, Lesotho, Namibia and Swaziland
CCT	City of Cape Town
CIRED	International Agency for Research on Environment and Development
COJ	City of Johannesburg
CSIR	Council for Scientific and Industrial Research
CSV	comma separated value
CTCC	City of Cape Town
DEA	Department of Environmental Affairs
DoE	Department of Energy
DoL	Department of Labour
DWAF	Department of Water Affairs and Forestry
EB	energy balance
EECC	Environment and Climate Change Group (ERC)
ERC	Energy Research Centre
ESA	Energy Systems Analysis Group
e-Sage	Extended South African Computable General Equilibrium Model
GHG	greenhouse gas
HFO	heavy fuel oil
IEA	International Energy Agency
IP	illuminating paraffin
IPA	Independent Project Analysis
IPCC	Intergovernmental Panel on Climate Change
ISIC	International Standard Industrial Classification of all Economic Activities
ITLS`	Institute for Transport and Logistic Studies
LEAP	long range energy alternatives planning (model)
LFO	light furnace oil
LPG	liquefied petroleum gas
MJ	megajoule

MOOC	massive open online course
NAEIS	National Atmospheric Emission Inventory System
NAN	not a number
NCPC-SA	National Cleaner Production Centre of South Africa (CSIR)
OGC	Open Geospatial Consortium
PDRI	Project Definition Rating Index
PJ	petajoule
PMBOK	Project Management Body of Knowledge
PMI	Project Management Institute
PRINCE2	Projects in Controlled Environments, version 2
SAM	social accounting matrix
SANEDI	South African National Energy Development Institute
SAPIA	South African Petroleum Industry Association
SATIM	South African TIMES model
Stats SA	Statistics South Africa
SUT	supply and use table
SUT	supply and use tables
TIMES	The Integrated MARKAL-EFOM System
TJ	terajoules
TJ/a	terajoule per annum
UJ	University of Johannesburg
UNFCCC	United Nations Framework Convention on Climate Change
VAR	vector auto-regression (model)
VARX	vector auto-regression (model) with exogenous data
VECM	vector error correction model
WCPG	Western Cape Provincial Government

1. PREFACE

1.1 Introduction to Top Quartile and the Energy Research Centre

Top Quartile PMSA (www.topquartile.co.za) is a specialist professional services company that helps private, public and social sector clients develop and execute projects, build project capabilities and collaborate in the domains of energy, infrastructure development, and petro-chemicals. Through deep project expertise, a passion for society and collaboration we systematically develop, design and execute projects to create shared value. We combine project rigour with the flexibility required to advance confronting complex 'wicked' problems. Our approach is focused, collaborative and results oriented.

Our diverse and versatile team enables us to work in a variety of contexts and on issues across the sectors of business, government and civil society. Top Quartile employees and its partners have a track record for delivering projects successfully with collectively more than 80 years of relevant project management experience. Our project development offering includes supporting clients in the capacity of owner representative and / or independent principle agent, given our extensive experience in developing and managing projects through careful front-end loading and disciplined implementation. We have developed fit-for-purpose methodologies, tools and templates, incorporating leading project management practice (Project Management Body of Knowledge (PMBOK), Projects in Controlled Environments, version 2 (PRINCE2), Project Definition Rating Index (PDRI)) from benchmark organisations (such as the Project Management Institute (PMI), Independent Project Analysis (IPA), the Office of Government Commerce (OGC) and the Association for the Advancement of Cost Engineering International (AACEI)). These are complemented by our expertise in building quantitative models for rigorous scenario based analyses, always acting in clients' best interest to realise client objectives.

- Our service offering is built on our project expertise and our commitment to collective impact. We offer value through our primary services:

- ***Project Development:*** We assist clients to set up for success by evaluating and structuring ideas into bankable projects[1] and facilitate disciplined execution across the project lifecycle. At each project phase, we enable informed decision-making based on technical, economic, legal, social and environmental parameters.

- ***Project Capability Building:*** We build project capability through the provision of customised interventions and tools.

- **Cross-sector Collaboration:** We convene and facilitate diverse stakeholders to understand and confront tough challenges holistically.

Our client base includes billion dollar dual-listed companies, international development agencies, tertiary institutions, off-shore financiers and government departments. In addition to our established legal, financial and environmental associations, we have established associations with leading energy research institutions. Our name denotes our aspiration to consistently achieve the top 25 per cent of project performance outcomes.

The Energy Research Centre (ERC) at the University of Cape Town (www.erc.uct.ac.za) is one of South Africa's leading institutions involved in energy research, employing about 30 people from a variety of backgrounds. The ERC is a multi-disciplinary institution that conducts high quality, targeted and relevant research, as well as

[1] We conduct rigorous pre-feasibility, feasibility and development phase studies as well as due diligence assessments to establish project viability

offering postgraduate opportunities at the Masters and PhD levels. Within the ERC, the Energy, Environment and Climate Change (EECC) group researches the intersection between energy, local environment and global climate change. The research aims to contribute towards minimising impacts of energy use and production, from social, economic and environmental perspectives. The Energy Systems Analysis (ESA) Group undertakes research, analysis and training related to energy and climate change mitigation modelling, analysis and support. This includes issues related to energy demand, supply and infrastructure, and to collation, management and dissemination of energy data.

- The ESA group maintains a model of the South African energy system called the South African TIMES model (SATIM) developed over a number of years. There are a number of sub-models but the framework is centred around a so-called MARKAL type model on the TIMES platform. This is a partial equilibrium, linear, least-cost optimisation model that can capture the complex interactions in the entire energy system across and between sectors, and can be used to assess the least-cost combination of technologies and fuels required to meet future energy demand subject to constraints such as greenhouse gas emissions. The model includes the whole energy chain from primary fuel extraction, to transformation in refineries and power stations, to end use technologies supplying services such as lighting, heating and passenger travel. The methodology and assumptions for the model are published in detail in the public domain.

1.2 Executive summary

It is a United Nations Framework Convention on Climate Change (UNFCCC) requirement that each non-Annex I Party provides in its national inventory, on a gas-by-gas basis, estimates of anthropogenic emissions of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) by sources and removals by sinks. Furthermore the National Climate Change Response Policy requires of the Department of Environmental Affairs (in partnership with the South African Weather Service) to prepare a Greenhouse Gas (GHG) Emissions Inventory annually which will conform to the IPCC's 2006 and later guidelines (RSA 2011: 29). The said policy also stipulates that the inventory should include analyses of emissions trends, including detailed reporting on changes in emissions intensity in the economy and a comparison of actual GHG emissions against the benchmark national GHG emission trajectory range (RSA 2011).

South Africa is currently updating the GHG inventory to cover the period 2000–2010; however, accurate allocation of combustion fuels, including petroleum fuels and biomass based fuels to demand-side sectors is proving to be a challenge. The current source of information on the allocation of liquid fuels to demand-side sectors is outdated and, furthermore, it is basically limited to vehicle emissions, specifically petrol and diesel, with very limited to no information on disaggregation into the respective demand-side sectors and sub-sectors (SAPIA 2008). In light of these considerations this project was commissioned with the immediate objective being the disaggregation of combustion fuels (excluding coal) data according to the demand-side sectors and sub-sectors over the period 2000 to 2012.

In common with many other developing countries, data quality varies from poor to excellent and the unfolding data picture is blurred by the variation in data quality between the different data sources. In order to achieve the assignment objective of establishing the disaggregated

Data Quality / Usefulness

Source Classification	5	4	3	2	1
1	1. HSRC (DoE)$_{2012}$ LPG, wood, IP (Indirect) 2. Wessels Under Pressure (DoL)$_{Multi-Year}$ Coal, Fuel, Oil, Gas (Indirect)	3. DoE/SAPIA Fuel Sales$_{2007-2012}$ dsl, ptrl, LPG, IP 4. Tshwane AQM$_{Multi-Year}$ dsl, ptrl, LPG, wood/chrcl IP, HFO 5. CoJ AQM$_{Multi-Year}$ dsl, ptrl, LPG, wood/chrcl IP, HFO	7. NCPC$_{2007-2012}$ dsl, ptrl, LPG, wood/chrcl IP, HFO 8. CCT AQM$_{Multi-Year}$ dsl, ptrl, LPG, wood/chrcl IP, HFO 9. WCPG Energy Market$_{2013}$ dsl, ptrl, LPG, wood/chrcl IP, HFO	16. GHS$_{2002-2012}$ LPG, wood, IP (Indirect) 17. ACSA$_{2002-2012}$ Flight data (Jet Fuel) (Indirect) 18. DoT Marine Container Flows$_{2005-2010}$ (Indirect)	23. Sasol Natural Gas Sales$_{2000-2012}$ nat gas 24. SAPIA Fuel Sales$_{2002-2006}$ dsl, ptrl, LPG, IP 25. Stats SA SUTs$_{2002, 2007, 2011}$ (Indirect) 26. Eskom.co.za$_{2006-2015}$ Diesel
2a			10. Martin C - (MSc) Commercial LPG, IP, dsl, coal 11. Forestry South Africa$_{2000-2012}$ wood, chrcl	19. SEA SOE models dsl, ptrl, LPG, chrcl nat gas, IP	27. DWAF$_{2003}$ chrcl
2b		6. DoE EB$_{2007-2012}$ dsl, ptrl, LPG, wood, IP, HFO, LPG, nat gas	12. SANEA Energy Profile$_{2000}$ dsl, ptrl, LPG, wood IP, HFO, LPG, methane-rich gas 13. IEA EB$_{2008-2010}$ dsl, ptrl, LPG, wood, IP, HFO, LPG nat gas	20. DoE EB$_{2000-2006}$ dsl, ptrl, LPG, wood prf, HFO, LPG, nat gas	
3			14. ERC (SATIM)$_{2006}$ dsl, ptdsl, ptrl, LPG, wood	21. DoE Bottom-up$_{2002-2012}$ Wood 22. Hughes$_{1996, 2011}$ Wood	
4			15. Eskom (pers com)$_{2015}$ Fuel Oil		

(Figure 2: Pictorial assessment of useful data sources identified and their origin and level of quality/usefulness)

fuel consumption per industry sub-sector, an assessment of the data landscape had to be addressed first. Figure 2 in Section 3.2 (reproduced above) presents pictorially the main data sources and where they lie on the scale of certainty between published empirical measurements and model generated data (vertical axis) and on a general scale of quality/usefulness (horizontal axis). Given that this exercise will be repeated by others it is informative to establish what potential there is in the future for developing better data for the purposes of the GHG Inventory from the respective data sources. A brief overview of each of the main data sources, their strengths and weaknesses, how they were applied and what potential there is in the future for developing better data from that source is presented in Section 3.2.

None of the data sources could be considered accurate and authoritative enough at all sector levels, across all fuel carriers, spanning the full period from 2000 to 2012. A considerable level of uncertainty surrounds each source, at least in one or other aspect.

In light of the above, the data modelling approach commenced with a first iteration of the best available data from surveys of industry data, typically the Department of Energy's (DoE's) energy balances, the South African Petroleum Industry Association's (SAPIA's) aggregated and disaggregated data sets, the HSRC's residential survey data and the Energy Research Centre's (ERC's) SATIM model. In certain instances mediation of sources was necessary to inform the decision (in the absence of complete data) to either proceed with model based imputation, or apply econometric smoothing to the data, or to source further data, or a combination of the above. A key aspect of the approach was to adopt the hypothesis that the basic sources of data are the most accurate available, until sufficient evidence had been established to reject the null hypothesis.

In order to construct a consistent database that reflects reasonable accuracy per sector, the data sources had to be supplemented with imputation models and bottom-up models (where appropriate), and eventually conditioned with the sectoral profiles of petroleum product consumption indicated by the supply and use tables of Statistics South Africa (Stats SA). The uncertainty relating to some of the energy carriers, for example, liquefied petroleum gas (LPG), fuel oils, biomass, wood and charcoal, could only be addressed by specific bottom-up models and fragmented survey data. Industry statistics, energy system models and expert opinions provided substantiation for the bottom-up models. The improvement in certainty regarding consumption statistics in these cases is considered to be significant.

In the absence of representative sectoral consumption data, an alternative approach was followed in this assignment. Large producers or suppliers in the energy value chains often record their respective views on the consumption of their products. These were used to assess and track consumption profiles in sectors and sub-sectors. A few cases in point are discussed.

The 2011 National Climate Change Response Policy requires the Greenhouse Gas (GHG) Emissions Inventory to be prepared in conformance with the Intergovernmental Panel on Climate Change's (IPCC's) 2006 and later guidelines. With the IPCC 2006 category code definitions referencing the divisions, groups and classes of revision 3.1 of the International Standard Industrial Classification[2] of All Economic Activities (ISIC) (UN 2002), the ISIC Revision 3.1 categorisation system was adopted in this assignment, to facilitate future reporting to the International Panel on Climate Change.

It was found that a huge amount of disparity exists in the way that the data sources define sectors and sub-sectors. The sectoral divisions of consumption, if used by the data owners, are normally driven by in-house requirements. It is therefore not possible to obtain line-for-line matches between, for example, the energy balances and the specific sectors and sub-sectors defined by the IPCC guidelines. This observation has a significant impact on the consumption profiles to be obtained for the various sectors and sub-sectors as per the IPCC reporting guidelines. A mapping of specific data owner consumption categories to those of ISIC Revision 3.1 and IPCC 2006 thus had to be developed. This is reflected in Annexure 1 to the report.

Since the energy consumption data is essentially provided in a time series format it is possible to apply well-known econometric modelling techniques which will enable the

[2] The International Standard Industrial Classification of all Economic Activities (ISIC) is the international reference classification of productive economic activities. Its main purpose is to provide a set of clearly demarcated activity categories that can be utilized for the production of statistics according to such activities.

forecasting of consumption of specific energy carriers. This approach was also used to determine missing and doubtful historical values. For the purpose of this assessment MS Excel was employed to prepare the data for processing and to capture results for reporting. The statistical analyses were conducted in the programming language called "R". R is open source software which is well documented and the application of choice for many research projects at respected academic and commercial institutions.

The Department of Environmental Affairs (DEA) expressed the desire to see forecasts of liquid fuels consumption up

Table 1: Conversion factors used in this assignment

Energy carrier	Calorific Value	Units	Density (kg/l)	Intensity (MJ/kg)
Liquefied Petroleum Gas	26.7	MJ/l	0.5410	49.353
Paraffin Power	37.5	MJ/l	0.8130	32.841
Gas SASOL	41.0	MJ/m^3		
Diesel	38.1	MJ/l	0.839	31.824
Electricity	3.6	MJ/kWh		
Natural Gas	41.0	MJ/m^3		
Heavy Fuel Oil	41.6	MJ/l	0.984	27.134
Petrol	34.2	MJ/l	0.723	36.929
Paraffin Illuminating CSS (StatsSA) Data	37.0	MJ/l	0.788	33.883
Aviation Gas	33.9	MJ/l	0.730	36.575
Jet Fuel	34.3	MJ/l	0.793	33.670
Coal Eskom Average	20.1	MJ/kg		
Coal (General purpose)	24.3	MJ/kg		
Coal (Cooking)	30.1	MJ/kg		
Coke	27.9	MJ/kg		
Coke oven gas	17.3	MJ/m^3		
Blast furnace gas	3.1	MJ/m^3		
Refinery gas (estimate)	20.0	MJ/m^3		
Bagasse (wet)	7.0	MJ/kg		
Bagasse fibre (dry) 14.0 MJ/kg	14.0	MJ/kg		
Biomass (wood dry typical)	17.0	MJ/kg		
Coal gas (Sasol)	18.0	MJ/m^3		
Coal gas (Sasol - methane rich)	38.0	MJ/m^3		
Wood Charcoal	31.0	MJ/kg		

till 2025. In the econometric modelling approach the liquid fuels consumption data is supplemented with economic growth data for the various sector clusters in the standard reporting format adopted by central banks worldwide. The consumption data is considered endogenous while the economic data would be exogenous, namely, external to the set of liquid fuels under consideration and driving changes in the endogenous data. Selected confidence intervals (70%, 80% and 90%) for the respective fuel carriers are plotted for the period up to 2025.

The assignment also provides an assessment of confidence intervals for annual consumption per carrier per sub-sector over the study period.

The report incorporates a presentation and discussion of the study results as well as a reflection on the strengths and weaknesses of the assignment outcomes. It concludes with recommendations to facilitate future data quality and value and some thoughts on future work considerations.

Concluding remark: The current reality is that research and academic institutions, as well as other stakeholders in the energy and environmental management sectors, rather opt to use existing survey data than to conduct field surveys to expand the existing knowledge base on the sectoral distribution of energy consumption. As a consequence the quality of data is relatively low in South Africa, with quality data limited to a sub-set of fuel carriers and industry sub-sectors. Whereas this assignment has made significant strides towards addressing the resulting uncertainties in the data landscape, in the absence of more physical measurements and direct field survey work, the fuel consumption across demand side sectors will always contain a certain level of uncertainty.

Data summary: It is to be noted that the data reflected in the summary data table represents the best data obtained from base sources (SAPIA in particular), supplemented with alternative views (bottom-up estimations). In certain cases, for example, international jet fuel consumption, the base source data points are dropped in favour of the estimated consumption levels. This exchange is only executed when the base source data points are judged to be of inferior quality, or if there is a low level of associated certainty (see Section 4.5.7.2).

1.3 Units of measurement and conversion factors

In order to be consistent, all energy consumption levels recorded in this assignment are reported in terajoules per annum (TJ/a) in the final compilation of historical data. All forecasts, beyond 2012, are also provided in energy units (TJ/a).

It is to be noted that a number of consulted data sources post energy consumption in other native units, for example, kilolitres for volume based consumption. Similarly, many of the bottom-up assessments that had to be incorporated into this study also refer to native units of measure. These native units have been used, and are quoted, in the substantiating models which are submitted with this report.

In all cases where a transformation between native units and energy units is required, a standard list of conversion factors was employed. The standard conversion factors refer to those provided in the DoE's Energy Digest of 2009. A summary of the relevant calorific values related to the specific energy carriers appears in Table 1 on the previous page.

2. INTRODUCTION

2.1 Background

The ability of South Africa to achieve the objective of the 1992 United Nations Framework Convention on Climate Change (UNFCCC) to stabilise greenhouse gas (GHG) concentrations in the atmosphere at a level that would prevent dangerous human-induced interference with the climate system is highly dependent on accurate knowledge of emissions trends and on our collective ability to alter these trends. It is a UNFCCC requirement that in its national inventory each non-Annex I Party provides, on a gas-by-gas basis, estimates of anthropogenic emissions of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) by sources and removals by sinks (UN 2014: 21).

Furthermore the National Climate Change Response Policy requires the Department of Environmental Affairs (DEA) (in partnership with the South African Weather Service) to prepare a GHG Emissions Inventory annually which will conform to the IPCC's 2006 and later guidelines. The policy also stipulates that the inventory should include analyses of emissions trends, including detailed reporting on changes in emissions intensity in the economy and a comparison of actual GHG emissions against the benchmark national GHG emissions trajectory range.

South Africa is currently updating the GHG inventory to cover the period 2000–2010, however, accurate allocation of liquid fuels to demand-side sectors is proving to be a challenge.

In light of the above considerations this project was commissioned with the immediate objective being to disaggregate data for combustion fuels (excluding coal) according to the demand-side sectors and sub-sectors.

2.2 Study Objectives

This study aims primarily to detail the use of combustion fuels, including petroleum fuels and biomass based fuels, but excluding coal, in the demand sectors of the South African economy. The current source of information on the allocation of liquid fuels to demand-side sectors (SAPIA 2008) is outdated and furthermore is basically limited to vehicle emissions, specifically petrol and diesel, with very limited to no information on disaggregation into the respective demand-side sectors and sub-sectors.

With varying emission factors associated with the various demand-side sectors of the respective liquid and other combustion fuels, the splitting of these fuels into the various demand-side sectors is an important consideration. This is not only to meet reporting obligations to the International Panel on Climate Change (IPCC) and to ensure the completeness of the national GHG inventory but also, given the primary role of combustion fuels as a source of local air pollutants, it is also to improve air quality assessments, toxic emission inventories and the assessment of human exposure to pollutants. A combustion fuel such as wood, for example, may account for a small portion of the energy balance but its health impacts are profound given the high exposure levels to pollutants of consumers in the residential sector without access to modern energy. Petroleum fuels including liquid hydrocarbon fuels and liquefied petroleum gas (LPG) find their primary application as road transport fuels. This accounted for 73% of petroleum fuels energy consumption in the 2006 national energy balance. Their high energy density and transportability, however, means that they find use throughout the demand sectors of the economy to supply a range of energy services including heating, lighting, cooking and auto-generation of electricity for back-up and remote power supply. Emissions from non-road transport sectors such as aviation and rail can be expected to grow rapidly, the former directly driven by economic growth with the latter geared to a shift from road to rail due to rising oil prices and to some extent due to an increased reliance on diesel locomotives due to the electricity crisis. The quantities and trajectory of change of energy demand in such sectors is crucial data from a policy development perspective.

In light of the above considerations the following specific objectives were set for this assignment:

- disaggregate combustion fuels data according to the demand-side sectors and sub-sectors for the period 2000 to 2012;
- establish confidence intervals for annual consumption per carrier per sub-sector, over the study period;
- validate data sources;
- develop methodology to cater for the future incorporation of solid fuels (for example, coal);
- facilitate access to data sources to enable future maintenance and updating of the model;
- forecast study results to 2025 (late addition to scope) to facilitate IPCC reporting.

2.3 Scope (In/Out)

A detailed overview of the scope of this assignment is provided in Annexure 1, which is summarised in Table 2 below.

Table 2: Summary of assignment scope

In Scope	
Carriers	Diesel, Petrol (all grades), Fuel Oil, Liquefied Petroleum Gas, Biomass, Paraffin, Wood, Charcoal, Aviation Fuel
Boreal moist	***Agriculture*** • Agriculture, forestry & hunting ***Industry*** • Mining, food & tobacco, pulp & paper, chemicals, non-metallic minerals, iron & steel, non-ferrous metals, construction, other ***Commerce*** • Wholesale and retail trade and repair of motor vehicles and motor cycles • Wholesale trade, except for motor vehicles and motorcycles • Retail trade, except for motor vehicles and motorcycles ***Transport*** • Road, railways, international waterborne navigation, domestic waterborne navigation, international aviation, domestic aviation ***Residential***
Out of Scope	
Carriers	Coal
Sectors	***Industry*** • Manufacture of basic pharmaceutical products and pharmaceutical preparations, manufacture of rubber and plastic products, • Power generation • Petroleum refining • Manufacture of solid fuels and other energy industries

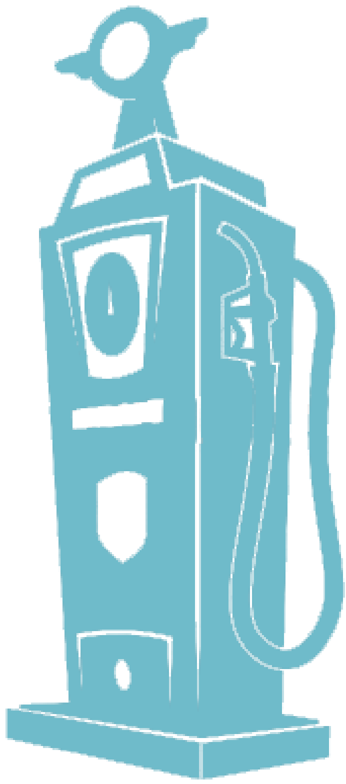
	• Water collection, treatment and supply • Sewerage • Waste collection, treatment and other waste management services • Warehousing and support activities for transportation • Postal and courier activities • Accommodation • Food and beverage services activities • Publishing activities • Motion picture, video and television programme production, sound recording and music publishing activities • Programming and broadcasting activities • Telecommunications • Computer programming, consultancy and related activities • Information services activities • Financial services activities, except insurance and pension funding • Insurance, reinsurance and pension funding, except compulsory social security • Activities auxiliary to financial service and insurance activities • Real estate activities • Legal and accounting activities • Activities of head offices; management consulting activities • Architectural and engineering activities; technical testing and analysis • Scientific research and development • Advertising and market research • Other professional, scientific and technical activities • Veterinary activities • Rental and leasing activities • Employment activities • Travel agency, tour operator, reservation service and related activities • Security and investigation activities • Services to buildings and landscape activities • Office administrative, office support and other business support activities • Public administration and defence, compulsory social security • Education • Human health activities • Residential care activities • Social work activities without accommodation • Creative, arts and entertainment activities • Libraries, archives, museums and other cultural activities • Gambling and betting activities • Sports activities and amusement and recreation activities • Activities of membership organisations • Repair of computers and personal and household goods • Other personal services activities • Undifferentiated goods-and services-producing activities of private households for own use. • Activities of extraterritorial organisations and bodies

While not part of the scope of this assignment, coal is incorporated as part of the combustion fuels assessment because it is a complement or substitute to the other fuels. In this context it is useful to include it given that inferences have to be made to understand how combustion fuels are used.

As reflected in Annexure 1, the IPCC classification system combines commercial and institutional buildings into a single category [1-A-4-a]. This contrasts with the ISIC Revision 3.1 approach of categorising institutional buildings into a number of sub-categories. Since the Stats SA supply and use tables, which closely emulate the ISIC Revision 3.1 classification system, provided values for most institutional buildings, it was decided to incorporate institutional buildings into the scope of the assignment to enable reconciliation of the respective mapping systems (Stats SA 2008–12; UN 2002).

2.4 Levels of detail

The following convention was adopted in terms of the level of detail on which reporting is done in this assignment. To the extent that quality data could be sourced at division level (as per ISIC Revision 3.1 definitions (UN 2002)) reporting is done at division level. However where source data was aggregated to section level reporting is done at section level.



2.5 Conditions

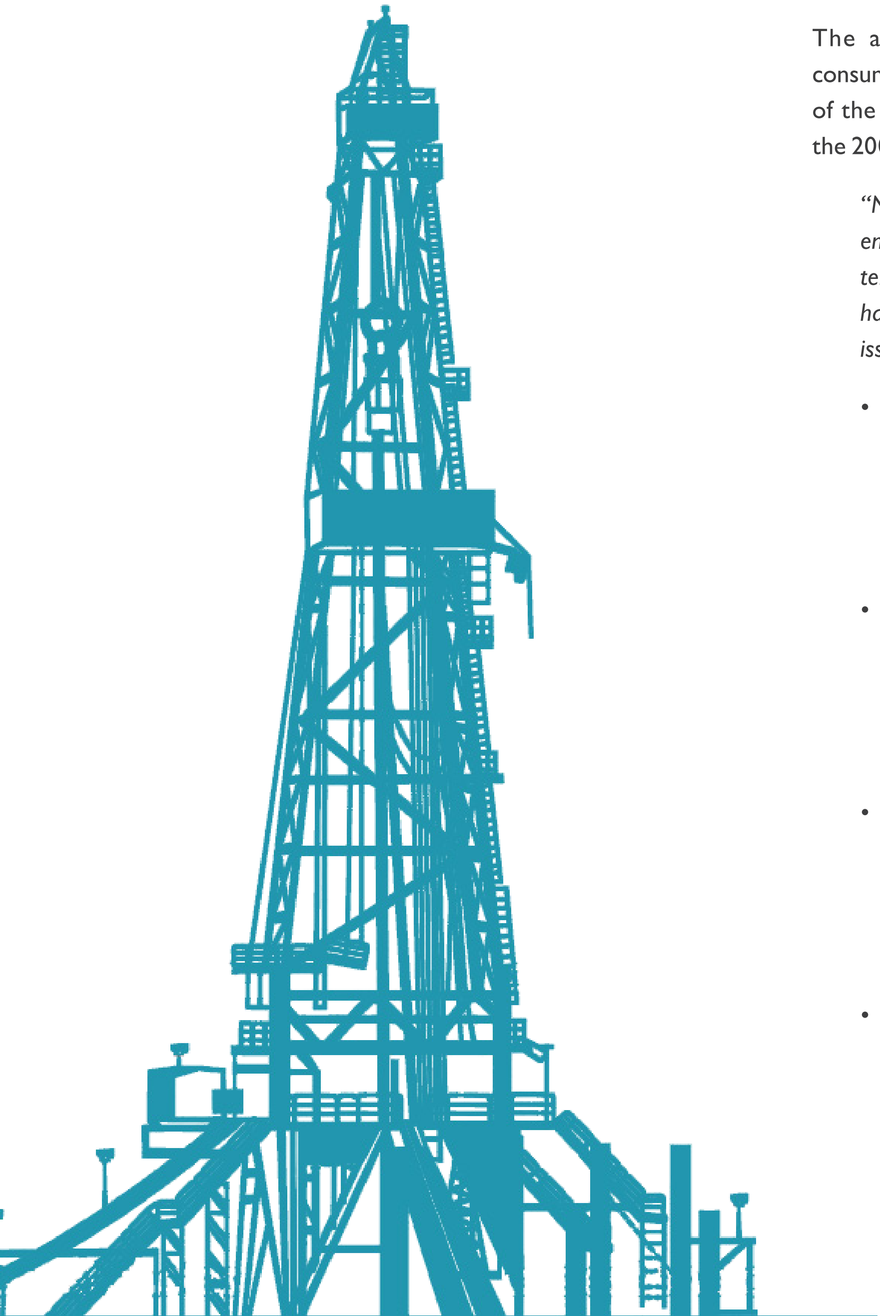
2.5.1 IPCC requirements

The National Climate Change Response Policy requires the DEA (in partnership with the South African Weather Service) to prepare a GHG emissions inventory annually which will conform to the IPCC's 2006 and later guidelines. The IPCC 2006 category code definitions reference the divisions, groups and classes of revision 3.1 of the International Standard Industrial Classification[3] (ISIC) of

[3] The International Standard Industrial Classification of all Economic Activities (ISIC) is the international reference classification of productive economic activities. Its main purpose is to provide a set of clearly demarcated activity categories that can be used for the production of statistics according to such activities.

all economic activities. To facilitate future reporting to the IPCC, the ISIC Revision 3.1 categorisation system was adopted in this assignment. Annexure 1 to this report provides a mapping that aligns the ISIC Revision 3.1 categorisation to the IPCC 2006 Guidelines for national greenhouse gas inventories.

The accumulation and verification of liquid fuels consumption is subject to the IPCC guidelines for coverage of the national inventories. The following extract from the 2006 IPCC Guidelines states the requirement clearly.



> *"National inventories should include greenhouse gas emissions and removals taking place within national territory and offshore areas over which the country has jurisdiction. There are, however, some specific issues to be taken into account:*
>
> - *Emissions from fuel for use on ships or aircraft engaged in international transport should not be included in national totals. To ensure global completeness, these emissions should be reported separately.*
> - *CO_2 emissions from road vehicles should be attributed to the country where the fuel is sold to the end user. The same allocation principle can be applied to other gases depending on the tier used to estimate emissions.*
> - *Fishing includes emissions from fuel used in inland, coastal and deep sea fishing. Emissions resulting from fuel used in coastal and deep sea fishing should be allocated to the country delivering the fuel.*
> - *The IPCC methodology for carbon stored in non-fuel products manufactured from fossil fuels or other non-biogenic sources of carbon takes into account emissions released from their production, use and destruction. Emissions are estimated at each stage when and where they occur, for example in waste incineration.*

- *Where CO_2 emissions are captured from industrial processes or large combustion sources, emissions should be allocated to the sector generating the CO_2 unless it can be shown that the CO_2 is stored in properly monitored geological storage sites as set out in Chapter 5 of Volume 2. Emissions from CO_2 captured for use, for example in greenhouses and soft drinks, and transported offsite should be allocated to the sector where the CO_2 was captured.*
- *CO_2 emissions from biomass combustion for energy are estimated and reported in AFOLU sector as part of net changes in carbon stocks.*
- *When reporting harvested wood products (HWP), countries can select any of the approaches reflected in Chapter 12 of Volume 4 for the AFOLU sector when estimating their emissions/removals from HWP (IPCC 2006a 8.4–8.5)."*

With regard to uncertainties in reported numbers, the IPCC states the following requirements:

> *"Uncertainty estimates are an essential element of a complete inventory of greenhouse gas emissions and removals. They should be derived for both the national level and the trend estimate, as well as for the component parts such as emission factors, activity data and other estimation parameters for each category. This guidance therefore develops a structured approach to estimating inventory uncertainty. It includes methods for:*
>
> *Determining uncertainties in individual variables used in the inventory (e.g., estimates of emissions from specific categories, emission factors, activity data);*
>
> - *Aggregating the component uncertainties to the total inventory;*
> - *Determining the uncertainty in the trend; and*
> - *Identifying significant sources of uncertainty in the inventory to help prioritise data collection and efforts to improve the inventory.*
>
> *While the methods outlined below are intended to estimate uncertainties for the national inventory, it is important to recognize that some uncertainties that are not addressed by statistical means may exist, including those arising from omissions or double counting, or other conceptual errors, or from incomplete understanding of the processes that may lead to inaccuracies in estimates developed from models.*
>
> *An uncertainty analysis should be seen, first and foremost, as a means to help prioritise national efforts to reduce the uncertainty of inventories in the future, and guide decisions on methodological choice. For this reason, the methods used to attribute uncertainty values must be practical, scientifically defensible, robust enough to be applicable to a range of categories of emissions by source and removals by sinks, methods and national circumstances, and presented in ways comprehensible to inventory users. A reference section is provided for more detailed and more theoretical information on topics discussed in this chapter.*
>
> *Quantitative uncertainty analysis is performed by estimating the 95 percent confidence interval of the emissions and removals estimates for individual categories and for the total inventory (IPCC 2006b: 3.6)."*

The above was used as a guideline for reporting uncertainty (or the inverse of certainty) in this study. However, due to the nature of the study, namely, the heavy dependence on desktop investigation and very limited direct field survey input, the uncertainties will be reported on a qualitative basis, supplemented by specific quantification where possible.

3. DATA LANDSCAPE

The readers of this report will likely need little convincing that the quality and quantity of data on energy and emissions from the transformation and consumption of energy in South Africa is problematic. In common with many other developing countries, data quality varies from poor to excellent and the unfolding data picture is blurred by the variation in data quality between the different data sources. Herein lies the frustration of the data landscape because in working through this partial data picture many questions arise and gaps have to be filled by analysis, modelling or assumption. Typically, such a situation begets much in the way of review and rework. In many cases promising avenues take up much time but yield little in the way of certainty. In order to achieve the assignment objective of establishing disaggregated fuel consumption per industry sub-sector and per fuel, an assessment of the data landscape had to be addressed first. This has been useful in indicating what is possible to establish from desktop research and what requires direct investment in measurement.

3.1 Energy carriers and their Uses

The goal of this study is essentially to disaggregate the consumption of all combustion fuels (except coal) that are used for non-road transport energy services by consuming sectors – based on public information assisted where possible with inference. While not part of the assignment scope, coal is part of the combustion fuels assessment because it is a complement to or substitutes for the other fuels. It is useful to include given that inferences have to be made to understand how combustion fuels are used.

Electricity is sometimes referred to as 'modern energy' but in many applications, particularly high temperature heating, it is more convenient, practical and cheaper to use direct combustion. It should be borne in mind that electricity in South Africa with its coal intensive generation fleet is just transformed heat energy from coal combustion. Transformation of electrical energy can be done very efficiently with an electric motor but electric heating can be far less efficient and resistance heaters can be slow to respond to control. This section therefore briefly reviews the uses of the combustion fuels assessed in this study.

3.1.1 Aviation gasoline

This is a high octane gasoline used by reciprocating spark ignition internal combustion engines propelling light aircraft, generally for recreational or crop spraying purposes.

3.1.2 Illuminating paraffin

Paraffin, also called kerosene is a highly refined, versatile, low sulphur, light to mid density distillate fuel (typical boiling range 150–250 °C) that finds a wide range of residential and industrial uses. Its consumption has steadily declined with electrification and its residential use is discouraged in many quarters because it is highly toxic if drunk by children and has been implicated in serious fires in informal settlements. Its industrial uses can be summed up as follows:

> *Extensively used in wick feed and pressure lamps, stoves, refrigerators and heaters. … Paraffin is also used as a solvent in the manufacture of certain paints, varnish, polish, insecticides and weed killers. Due to its purity, it also finds favour in the agricultural sector where it is well suited for grain drying and also satisfies a wide range of commercial and industrial heat fuel requirements such as process air heating, heating of liquids in tanks and firing gas turbines (Engen 2015).*

Paraffin is readily lit at ambient conditions without the dangerous volatility of gasoline and is therefore useful for starting combustion in burner systems designed to operate on heavier fuel oils once hot.

Power paraffin a slightly heavier cut of paraffin originally used as a low octane grade fuel for older agricultural tractors and now mostly used as a cleaning fluid and is, for the purposes of this study, assumed to be illuminating paraffin.

3.1.3 Jet fuel or jet kerosene

A highly refined paraffinic cut suitable for use in jet propulsion engines and gas turbines in aircraft and stationary applications by virtue of having a very low content of waxy hydrocarbons which could precipitate and cause damage or operational problems in these machines. The bulk of this fuel is used in the mass air transport business. The converted jet engines used as peaking plant by Eskom at Port Rex and Acacia also operate on jet kerosene and their use has increased markedly although absolute volumes are still small as shown in Figure 1.

3.1.4 Residual oils

In general these products are the high boiling point residue remaining after distillation of the petroleum fractions used to make fuels like diesel, paraffin and petrol. Such residual fuels may be too viscous to pump unless heated and can be high in sulphur content (+/- 3%) and as a result are generally significantly cheaper than refined fuels like paraffin and diesel. The bulk of non-marine residual oil sold would fall into the category of heavy fuel oil (HFO) but for the purposes of this study lighter fractions such as light furnace oil (LFO), fuels made from recycled oils and coal tar were considered in one broad group especially given that their supply chain and range of use is poorly understood in the public domain.

Engen describes HFO as follows:

> *Heavy Furnace Oil is often used in high volume industrial heat fuel applications such as steam and hot water boilers (other than domestic), a variety of furnaces and kilns as well as very large drying operations. Major applications are found in brickworks, the cement, ceramic and glass industry, metal and ore smelting, and the asphalt industry (Engen 2015)[.]*

Notably, residual oils provide a cheap and potentially more consistent alternative to coal. Falcon (2015) has pointed out that there are 6 000 coal boilers in South Africa many of which are 50 years old or older and, having been designed for higher grade coal, struggle to obtain viable coal supplies. It stands to reason that where coal is not available is regionally expensive or of uncertain quality, residual oils may find application for heating purposes, particularly in boilers.

3.1.5 Petrol

Petrol or gasoline is highly volatile and flammable and although chemically very variable is generally highly toxic (US DOE, 1991) and an aggressive solvent. It is therefore dangerous to use outside of internal combustion engines even as a workshop cleaning fluid. As a result, petrol does not find wide application outside of light vehicle road transport except as fuel for light machinery like lawnmowers and small generators, usually less than 5 kVA in power rating. The latter may account for a more significant but still very small share of consumption in the last two years with more regular load shedding but for the 2000–2012 window it was considered negligible. For the purposes of this study gasoline was considered to be 100% used in road transport.

3.1.6 Diesel

While most diesel is consumed in road transport, it finds extensive application in off-road vehicles and machinery used to remove and transport material like ore, soil or materials and products. It is also, in general, the fuel of choice for remote power production and for back-up power to the grid. Diesel by design is not readily flammable in its liquid form at ambient conditions and can therefore also be a useful burner fuel for heating or for stationary gas turbine propulsion, most extensively to fuel peaking plants operated by Eskom or the large metropoles. Pressure on the grid since 2007 shows a clear trend of increasing diesel use by Eskom fuelling the gas turbines at Ankerlig and Gourikwa and jet kerosene for the jet engines at Acacia and Port Rex.

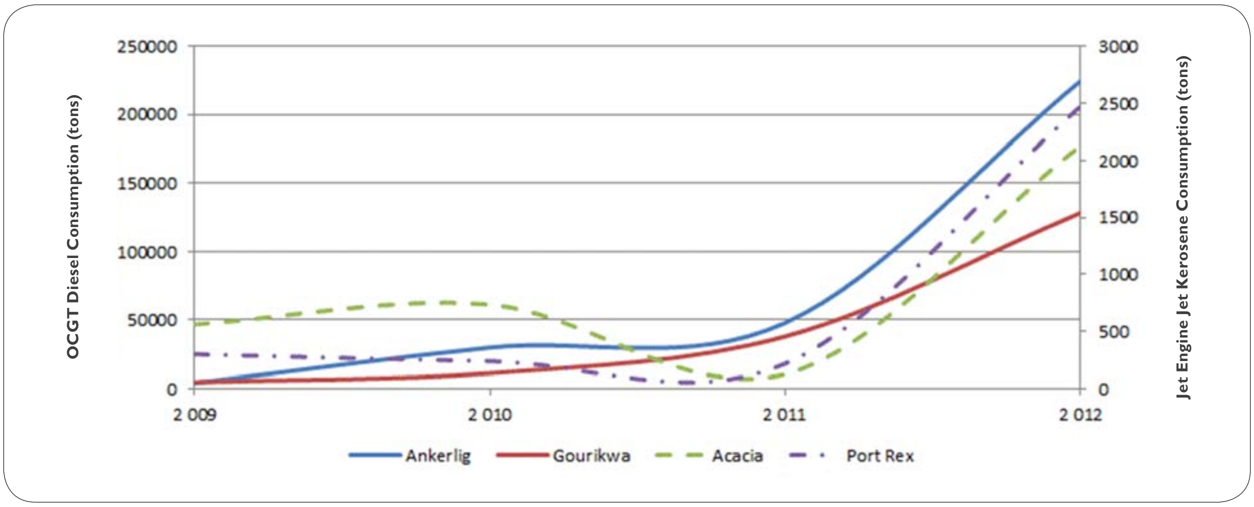


Figure 1: Growth in Eskom's diesel and jet kerosene consumption for peaking power 2009–2012

3.1.7 Liquid petroleum gas

Liquid petroleum gas (LPG) consists mainly of propane and butane, and is used in a wide variety of domestic, commercial and industrial applications. Engen describes the uses of LPG as follows:

> *LPG is used throughout the whole heating spectrum, satisfying all commercial and industrial requirements for space heating, heating of liquids in tanks, inert gas generation to hot water boilers and furnaces etc. LPG is also used for drying operations in the agriculture sector, paint, paper, food and glass industries, plus incineration and metal cutting. Due to its high Octane Number and level of purity, it is also suitable for modified petrol engines where it is favoured for its low engine exhaust emissions, especially in closed or confined spaces such as factory storerooms, cold storage rooms, ships holds, etc. Other factors such as utilisation convenience, ease of control, and portability allows LPG to compete economically with electricity in numerous commercial and industrial energy requirements (Engen 2015).*

While an attractive fuel, LPG is essentially a refinery by-product and an unfavourable tariff policy makes importation unattractive as can be seen from the very low values in the energy balances. Consequently supply can vary depending on refinery factors and seasonal demand which accounts for the ragged profile of the supply-side data for LPG.

3.2 Data source quality assessment

A great many documents and sources were retrieved and reviewed for this study. The majority of them were not directly useful or relevant. Given that this exercise will be repeated by others it is useful to review in this section the salient sources consulted and give a brief overview of their strengths and weaknesses, how they were applied and what potential there is in the future for developing better data for the purposes of the GHG inventory from that source. Figure 2 presents pictorially the data sources that were identified as being potentially useful and where they lie on the scale of certainty (vertical axis) between published empirical measurements (high) and model generated data (low) and on a general scale of quality/usefulness (horizontal axis) as regards this study. Tables 3 and 4 present the key to the certainty scale and the quality scale respectively. Please note, that some sources such as the raw data from the HSRC's *A survey of energy-related behaviour and perceptions in South Africa* and the Department of Labour's registration data for Vessels Under Pressure, could not be used in this study for practical reasons but are potentially very useful for future studies.

Data Quality / Usefulness

Source Classification	5	4	3	2	1
1	1. HSRC (DoE)$_{2012}$ LPG, wood, IP (Indirect) 2. Wessels Under Pressure (DoL)$_{Multi-Year}$ Coal, Fuel, Oil, Gas (Indirect)	3. DoE/SAPIA Fuel Sales$_{2007-2012}$ dsl, ptrl, LPG, IP 4. Tshwane AQM$_{Multi-Year}$ dsl, ptrl, LPG, wood/chrcl IP, HFO 5. CoJ AQM$_{Multi-Year}$ dsl, ptrl, LPG, wood/chrcl IP, HFO	7. NCPC$_{2007-2012}$ dsl, ptrl, LPG, wood/chrcl IP, HFO 8. CCT AQM$_{Multi-Year}$ dsl, ptrl, LPG, wood/chrcl IP, HFO 9. WCPG Energy Market$_{2013}$ dsl, ptrl, LPG, wood/chrcl IP, HFO	16. GHS$_{2002-2012}$ LPG, wood, IP (Indirect) 17. ACSA$_{2002-2012}$ Flight data (Jet Fuel) (Indirect) 18. DoT Marine Container Flows$_{2005-2010}$ (Indirect)	23. Sasol Natural Gas Sales$_{2000-2012}$ nat gas 24. SAPIA Fuel Sales$_{2002-2006}$ dsl, ptrl, LPG, IP 25. Stats SA SUTs$_{2002, 2007, 2011}$ (Indirect) 26. Eskom.co.za$_{2006-2015}$ Diesel
2a			10. Martin C - (MSc) Commercial LPG, IP, dsl, coal 11. Forestry South Africa$_{2000-2012}$ wood, chrcl	19. SEA SOE models dsl, ptrl, LPG, chrcl nat gas, IP	27. DWAF$_{2003}$ chrcl
2b		6. DoE EB$_{2007-2012}$ dsl, ptrl, LPG, wood, IP, HFO, LPG, nat gas	12. SANEA Energy Profile$_{2000}$ dsl, ptrl, LPG, wood IP, HFO, LPG, methane-rich gas 13. IEA EB$_{2008-2010}$ dsl, ptrl, LPG, wood, IP, HFO, LPG nat gas	20. DoE EB$_{2000-2006}$ dsl, ptrl, LPG, wood prf, HFO, LPG, nat gas	
3			14. ERC (SATIM)$_{2006}$ dsl, ptdsl, ptrl, LPG, wood	21. DoE Bottom-up$_{2002-2012}$ Wood 22. Hughes$_{1996, 2011}$ Wood	
4			15. Eskom (pers com)$_{2015}$ Fuel Oil		

Figure 2: Pictorial assessment of useful data sources identified and their origin and level of quality/usefulness

Table 3: Data source classification

1	Empirically observed
2a	Published, contacted source and established how data was derived.
2b	Published, uncertain on how data was derived (no methodology and way to validate)
3	Model derived data (Imputation)
4	Unpublished - personal communication

Table 4: Data quality classification

1	Sourcing approach, consistency and granularity of data excellent
2	Sourcing approach, consistency and granularity of data good
3	Sourcing approach, consistency and granularity of data reasonable
4	Sourcing approach, consistency and granularity of data poor
5	Sourcing approach, consistency and granularity of data inadequate

Table 5: Main data source references

Item	Study	Public Domain [Y/N]	Format	Reference
1	HSRC (DoE)$_{2012}$	N	Spreadsheet	Department of Energy, 2012; Department of Energy, 2013 (a);
2	Vessels Under Pressure (DOL)$_{\text{Multi-Year}}$	N	Unknown	N/A (see discussion below)
3	DoE/SAPIA Fuel Sales$_{2007-2012}$	N	Spreadsheet	SAPIA, n.d; Department of Energy, n.d.(b); Department of Energy, n.d. (c)
4	Tshwane AQM$_{\text{Multi-Year}}$	N	Spreadsheet	City of Tshwane, 2015
5	CoJ AQM$_{\text{Multi-Year}}$	N	Spreadsheet	City of Johannesburg, 2015
6	DoE EB$_{2007-2012}$	Y	Spreadsheet	Department of Energy, n.d. (a)
7	NCPC$_{2007-2012}$	N	Spreadsheet	National Cleaner Production Centre (NCPC), 2015
8	CCT AQM$_{\text{Multi-Year}}$	N	Spreadsheet	City of Cape Town, 2015
9	WCPG Energy Market $_{-2013}$	N	Spreadsheet	Visagie HJ, 2013 (b)
10	Martin C – (MSc) Commercial	Y	Thesis	Martin C. 2011
11	Forestry South Africa$_{2000-2012}$	Y	Spreadsheet	Forestry South Africa, 2015
12	SANEA Energy Profile$_{2000}$	N	Report	SANEA, 2003
13	IEA EB$_{2008-2010}$	Y	Spreadsheet	International Energy Agency (IEA), 2012
14	ERC (SATIM)$_{2006}$	N	Spreadsheet	Energy Research Centre, 2015
15	Eskom (pers com)$_{2015}$	N	Email	ESKOM, 2015
16	GHS$_{2002-2012}$	Y	Stata Files	Statistics South Africa (Stats SA), n.d. (a)

Table 5 continued...

Item	Study	Public Domain [Y/N]	Format	Reference
17	$ACSA_{2002-2012}$	N	Spreadsheet	University of Johannesburg, n.d.
18	DoT Marine Container $Flows_{2005-2010}$	Y	Report	National Department of Transport, 2011 (a), (b) and (c)
19	SEA SOE models	N	LEAP Models	Sustainability Energy Africa, 2015
20	DoE $EB_{2000-2006}$	Y	Spreadsheet	Department of Energy, n.d. (a)
21	DoE $Bottom\text{-}up_{2002-2012}$	N	Spreadsheet	Department of Energy, 2015 (b)
22	$Hughes_{1996, 2011}$	N	Spreadsheet	Hughes A. 2015
23	SASOL Natural Gas $Sales_{2000-2012}$	N	Spreadsheet	Sasol, 2015
24	SAPIA Fuel $Sales_{2000-2006}$	N	Spreadsheet	SAPIA n.d.; SAPIA 2000
25	Stats SA $SUTs_{2000, 2007, 2011}$	Y	Spreadsheet	Statistics South Africa (Stats SA), n.d. (b); Statistics South Africa (Stats SA), 2014; Statistics South Africa (Stats SA), 2015
26	$Eskom.co.za_{2006-2015}$	Y	Webpage	ESKOM, 2014
27	$DWAF_{2003}$	Y	Report	Department of Water Affairs and Forestry, 2003

The issues around quality for selected sources in the above figure are discussed in more detail below.

SAPIA fuel sales data 1999–2006, aggregate, by magisterial district and by market category

Until 2006 the supply side statistics for the petroleum industry were collected by one of the oil majors Chevron (formerly Caltex) and later the oil industry body, the South African Petroleum Industry Association (SAPIA). These statistics recorded the flows of the main petroleum products in South Africa and the Botswana, Lesotho, Namibia and Swaziland (BLNS) countries at quarterly intervals in three levels of detail:

- national aggregate
- disaggregated by magisterial district
- disaggregated by 'market category'

'Market category' includes some ISIC categories like mining, construction and marine fishing but also industry specific categories such as 'remainder of general trade', 'general dealers' and 'retail garages'. These last three are generally large for most commodities and therefore mapping to ISIC categories is not easily done. Market category data was also only made available to this study for petrol, diesel, LPG and illuminating paraffin. The magisterial district data includes other products.

This data is the official source for consumption of liquid fuels in South Africa and has been used directly to populate the national energy balance. It is therefore logical that it was adopted as the basis for the allocation to sector for this study. While flawed, it is the best and most complete supply-side source for consumption of liquid fuels in South Africa. This data set has been split in two for the purposes of assessing the data landscape in this section because after 2006 the quality of data dropped off sharply after the intervention of the Competition Commission in the data collection process citing anti-competitive behaviour. The Department of Energy had to, seemingly from some time in 2008, obtain the data directly from the major oil suppliers themselves and collate it. In 2012 an information exchange case involving the oil majors was referred to the Competition Tribunal by the Competition Commission who summed up their reason as follows:

> *The disaggregated sales information exchanged between oil companies in the case being referred here removed any element of surprise in strategic decision making and functioned as a reliable substitute to direct cartel interactions insofar as it made monitoring of rivals possible. This, together with the history of coordinated behaviour and other characteristics that exist in the petroleum industry, made achieving cartel outcomes post the exemption period possible. (Competition Commission, 2012)*

It is not clear to what extent the commission is aware of the widespread application of the disaggregated supply statistics in research, planning and environmental impact assessment and the trade-off of interests this implies or that the statistics while not actively published, were effectively public within the greater energy sector. Given the latter, it is hard to understand how the trade-off favours limiting information in an already challenged national data landscape. For the years 1999 and 2002–2006 the aggregate, magisterial and market category data was found to be consistent in terms of totals for each product.

SAPIA appears no longer to have market category data for 2000 and 2001. One of the study authors, however, had this data for 1999 in old project archives so trends could be interpolated to some degree. People in the energy and fuels business were widely contacted to try and track down legacy copies of original data for 2000 and 2001 but without success.

Contact: **Cornel Van Basten <cornel@sapia.co.za>**

Anton Moldan <Anton@sapia.co.za>

DoE / SAPIA fuel sales data 2007–2012, aggregate, by magisterial district and by market category

This dataset continues from the above but is of much poorer quality. Magisterial district and market category data for 2007 and 2008 was obtained from SAPIA and that for 2009–2012 was obtained from the Department of Energy.

The total consumption values of products between these two databases and the aggregate figures as published in SAPIA Annual Reports no longer agree from this point onwards as shown in Table 6.

It appears that the 2008 data was only for one quarter but this could not be confirmed as SAPIA stopped responding to questions after the initial supply of the data. Other issues observed include that the diesel consumption attributed to Transnet begins to deviate substantially from their own figures after 2005. This coincides with SAPIA taking over responsibility for data collation from Chevron (formerly Caltex) in that year and indicates that data quality problems started before Competition Commission intervention forced the DoE to take responsibility for fuel sales data collation in 2008.

SAPIA members and SAPIA had years of experience in collecting, collating and validating this data and when this responsibility passed to the DoE these skills had to be built up from scratch. This impact of the Competition Commission on the manner in which energy statistics was collected, resulting from the threat of heavy fines to many entities and industries, not just to SAPIA, carried through to the national energy balances which also deteriorated after 2006 as discussed below.

Table 6: Error between published aggregate consumption and totals from DoE/SAPIA market category database for diesel, IP, LPG and petrol 1999–2012

Year	Diesel gas oil - all grades	Illuminating paraffin	Liquid petroleum gas	Petrol - all grades
1999	-0.3%	-0.2%		-0.1%
2002	-0.5%	0.0%	0.8%	0.0%
2003	0.0%	-0.1%	0.0%	0.0%
2004	-0.1%	-0.1%	0.0%	0.0%
2005	0.0%	0.0%	0.0%	0.0%
2006	0.0%	0.0%	0.0%	0.0%
2007	-3.9%	1.7%	-5.1%	0.3%
2008	-75.6%	-64.9%	-77.2%	-74.3%
2009	9.2%	10.8%	5.9%	6.8%
2010	5.7%	9.4%	4.9%	4.5%
2011	6.1%	26.6%	-11.4%	5.4%
2012	6.4%	17.0%	-8.8%	5.3%

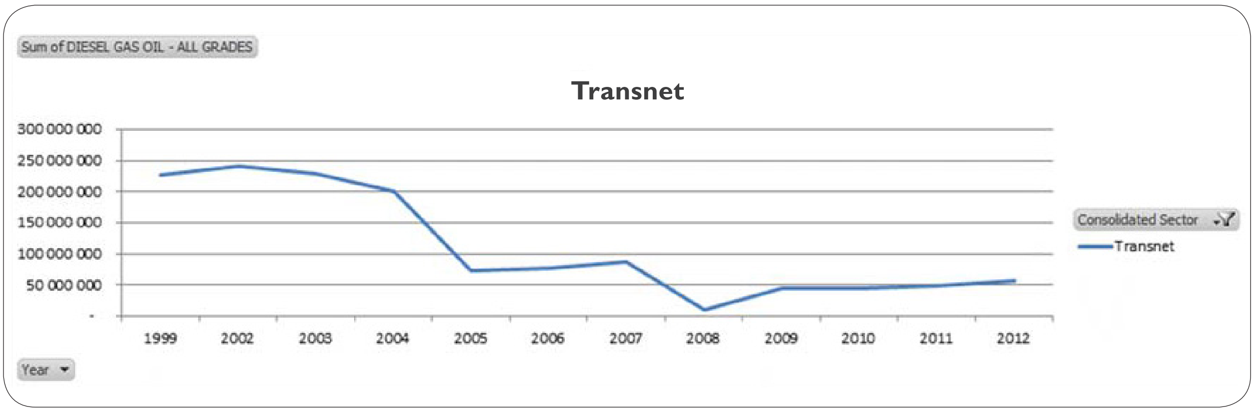


Figure 3: Trend of uncorrected SAPIA/DoE data for Transnet consumption of diesel

The supply-side data for liquid fuels after 2006 was therefore viewed to be of significantly lower quality for this study. The Competition Commission allowed an exemption for liquid fuels distribution data to be collected and shared again on 3 October 2011 extending to end December 2015 after an appeal that this was necessary to maintain supply.

It is not clear therefore what the liquid fuels data landscape will look like in the future.

Contact: **Ramaano Nembahe**
<Ramaano.Nembahe@energy.gov.za>

Department of Energy, energy balances 2000–2006

The national energy balance is the official account of energy supply, import, export, transformation and consumption by commodity and by sector including an abbreviated industry sub-sector breakdown. The energy balance is more than data, it is a high level model of the energy system derived from data, assumptions, simple equations and rules and forms the foundation of other more detailed models and plans in the sector.

It is, therefore of course, possible to derive a perfectly credible GHG inventory for the energy sector and non-process emissions from industry and the other demand sectors directly from a credible energy balance, allowing for small differences in reporting detail. The brief of this project made it clear, however, that the IPCC methodology advises that the inventory be developed in parallel to the energy balance which can then be used as a check and validation.

The lack of public data and problems with existing public data, however, means that this is not entirely practical in South Africa and, therefore, the energy balance formed the starting point for this study and where better information came to light this was substituted for energy balance values.

The deleterious effect of the Competition Commission's intervention in the effective collection and processing of energy sector statistics has been discussed with reference to SAPIA and petroleum industry statistics above. In common with these datasets there is a marked deterioration in the plausibility and consistency of energy balance statistics after 2006. This is being addressed by a current project in the DoE but for now the later balances remain problematic and for this study the 2000–2006 balances were considered more reliable.

This does not mean that questions do not arise with reference to this source and these include the residential consumption of coal and LPG, industrial and commercial LPG consumption, a plausible allocation of liquid fuels to industrial sub-sectors, the correct tracking of transformation in general and the absence of industrial biomass consumption.

Contact: **Ramaano Nembahe**
<Ramaano.Nembahe@energy.gov.za>

Department of Energy, energy balances 2007–2012

As mentioned above, the problems with later energy balances are in the process of being addressed by a project contracted to technology consultants Enerweb, an EOH company. The outputs for this project were not, however, available in time for this study. The problems with these later balances were serious enough to discount their use without careful assessment. These included sudden very large and implausible allocations of diesel to inland/domestic navigation (2009) and then to marine bunkers (2012), sudden and implausible increase of 'renewables and waste' (wood?) consumption (2010) and high coal allocation to the residential sector (2007–2009) despite electrification and indications that the primary use of coal domestically was dropping.

Enerweb's primary mandate was to improve the energy balance collation, processing and analysis systems but they also reviewed sources and the current state of data. It is to be hoped that the results of this valuable exercise are brought into the public domain and that all these initiatives build on each other in future to produce both a better energy balance and a better GHG inventory.

Contact: **Ramaano Nembahe**
<Ramaano.Nembahe@energy.gov.za>

Pranesh Ramjith
<Pranesh.Ramjith@enerweb.co.za>

Marc Stone
<Marc.Stone@enerweb.co.za>

International Energy Agency (IEA) energy balances 2006–2009

The IEA's balances generally derive from the national energy balance and its sources but in the case of South Africa some notable gaps are addressed, for instance biomass use in industry (for example, black liquor in the pulp and paper industry and bagasse by the sugar industry). These balances, therefore, provided another view on the thorny problem of industrial energy use but because they rely on the same public data sources, limited additional insight is provided.

These balances are not available at zero cost and the copies used in this study were legacy copies from previous studies. The contacts provided visited South Africa in 2015 on behalf of the IEA to investigate the state of energy statistics in South Africa.

Contact: **Justin Lacroix**
<justin.lacroix@hotmail.com>

Roger Purdue
<rogerpurdue@hotmail.com>

South African National Energy Association South African energy profile 2000

The South African National Energy Association's (SANEA's) energy sector overview for the year 2000 was notable for providing industry sub-sector fuel shares significantly different from the energy balances as shown by the following examples:

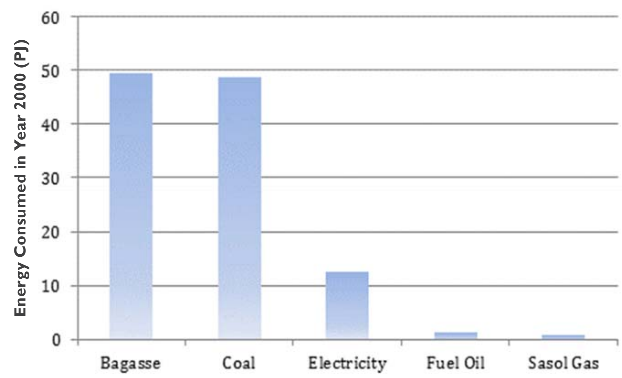


Figure 4: Fuel shares of energy consumed by the food, beverages and tobacco sector (2000) according to SANEA

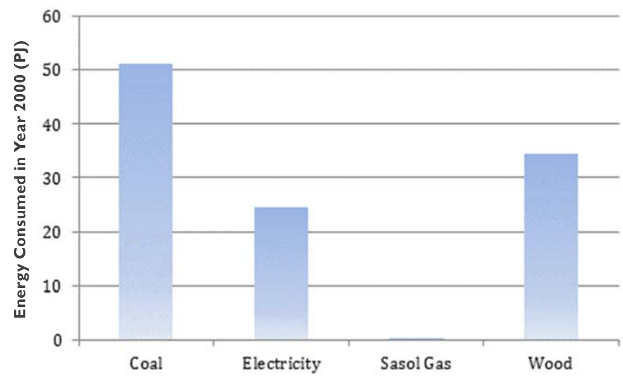


Figure 5: Fuel shares of energy consumed by the pulp and paper sector (2000) according to SANEA

This data was, however, only for the first year of the window and there was no description of the data collection or collation methodology in the source document.

Contact: **Zuerita Gouws**
<zuerita@mediainafrica.co.za>
(SANEA)

Marlene E van Rooyen
<marlene@25degrees.net>
(SANEA)

SASOL natural gas sales distribution data 2005–2015

Natural gas sales data for 2004–2015 was supplied by SASOL for the SASOL and Egoli distribution networks at a very high level of sub-sector detail (Sasol 2015). As such, although a relatively newly available energy carrier, its sectoral allocation likely has the highest quality data available of all carriers. Contributing to the data quality is the fact that there are only two distribution players. Data for 2000–2004 is unfortunately no longer available, however, prior to the supply of natural gas from Mozambique the industry was limited to hydrogen and methane rich gas from Sasolburg and Secunda.

Contact: **Rudi Hiestermann**
<Rudi.Hiestermann@sasol.com>

Statistics South Africa supply and use tables 2009–2012

The supply and use tables (SUTs) are published by STATSSA as part of the National Accounts and are the monetary flows associated with the supply of products by (supply table) and use of products by (use table) industrial sub-sectors at a high level of disaggregation equivalent to ISIC Revision 3.1. Supply and use tables are also available for 2002 but the level of disaggregation is much lower than those from 2009 onwards. While this level of disaggregation is potentially extremely useful for this study there are two major limitations in applying the data:

- The data is not available in physical units and prices, which may vary between sectors. Consequently prices have to be assumed to convert the data to physical units. This is less problematic with the highly regulated petroleum fuels than for coal.
- The product classifications are highly aggregated with all petroleum fuels lumped together under that description and gas aggregated with electricity.

In spite of these drawbacks the SUTs represent the only source that is an indicator of the relative consumption of petroleum fuels by industrial sub-sectors and was therefore used, in conjunction with some fuel share assumptions to break the industrial sector aggregate consumption of an energy carrier to sub-sector (ISIC Revision 3.1) level.

This work is ongoing as part of the development of the Energy Research Centre's economy-wide modelling framework called SATMGE[4] led by Tara Caetano and Bruno Merven. Lately they have been assisted by a post-doctoral researcher from the Alberto Luiz Coimbra Institute and Graduate School of Research and Engineering (COPPE), Federal University of Rio de Janeiro, named Romulo Ely. In response to this study he is hybridising a social accounting matrix (SAM) for South Africa which, being a balanced version of the national accounts showing the monetary flows between sectors and economic agents (households and government), is essentially a related data set to the SUTs. The hybridisation entails organising the data into a dual accounting system of monetary and physical units, the latter being of interest to this study. This is something COPPE specialises in due to their collaboration with the French Economics Institute, the International Agency for Research on Environment and Development (CIRED) and their IMACLIM economy-wide climate mitigation

[4] The abbreviation used for the linked model using the South African Times energy model (SATIM) and the Extended South African Computable General Equilibrium Model (e-SAGE) (Merven, Moyo, Stone et al. 2014)

impacts modelling system. CIRED are potential co-funders of ongoing work in this direction as they have built a version of IMACLIM for South Africa and are interested in continuing its development. An interesting follow on from this study might tie up the work done by Daniel Marais for this study with the work of these other researchers.

Contact: **Romulo Neves Ely**
<romulo@ppe.ufrj.br>
(COPPE – post-doc at ERC)

Tara Caetano
<tara.caetano@gmail.com>
(ERC)

Bruno Merven
<brunomerven@gmail.com>
(ERC)

Franck Lecocq
<lecocq@centre-cired.fr>
(CIRED)

Eskom liquid fuels consumption

Eskom publish their diesel and kerosene use online as part of their integrated report (Eskom 2014). This had to be adjusted from financial year to calendar year. The diesel / kerosene split is not given and whether the kerosene is jet kerosene or power paraffin is not disclosed but the relative production of the diesel fuelled Ankerlig and Gourikwa gas turbine power plants, and the kerosene fuelled Port Rex and Acacia jet engine power plants was forwarded in a personal communication from the system operator.

An estimate of heavy fuel oil consumption for starting up coal power plants (about 15,000 litres a month) was also forwarded in a personal communication.

Contact: **(System Operator) Keith Bowen**
<BowenKT@eskom.co.za>

Leslie Barker
<Barkerlf@eskom.co.za>

DoE Survey of Energy-Related Behaviour and Perception in South Africa

This survey has been undertaken twice (2012, 2013) by the Human Sciences Research Council (HSRC) on behalf of the DoE (DoE 2012; DoE 2013). Through direct surveys of a sample of households (3 004 in the 2013 study) covering all nine provinces the surveys explore how energy is being used, the adequacy of supply and the impact on people's lives of the availability of energy and energy services.

The only relevant part of the survey to this study is the measurement of what each household spends on a range of energy carriers including coal, paraffin, wood and LPG. Given, however, that the residential sector demand for these fuels was determined by bottom-up modelling for this study, this particular information was potentially extremely valuable. By assuming a price and converting to physical units it amounts to a direct, large sample observation of household energy intensity, as well as a household share of usage for each fuel. The latter is useful because the General Household Survey (GHS) data for fuel shares discussed below is only for primary users of a fuel.

The raw data received from the DoE in the form, apparently of a frequency distribution of monthly expenditure per fuel, was not consistent with the published report and was not accompanied by any clarifying remarks. While the published report itself remarked on the good agreement with GHS data this wasn't evident in the data received which was also not consistent with known rates of electrification or coal use. It seems likely that an error in the data extraction process is the root cause as the raw data couldn't be forwarded as it hasn't been anonymised. At the time of writing the DoE had not responded to requests for clarification. As a result this data could not be used directly and therefore has been located on the low end of the 'usefulness' scale in Figure 2.

Contact: **Machwene Molomo** **<Machwene.Molomo@energy.gov.za>**

Ramaano Nembahe **<Ramaano.Nembahe@energy.gov.za>**

Sustainable Energy Africa State of Energy of South African Cities and supporting LEAP models

Sustainable Energy Africa (SEA) recently published the State of Energy in South African Cities (SEA 2015) which presents, among other findings, the results of 18 energy systems models developed on the Stockholm Environment Institute's Long-range Energy Alternatives Planning (LEAP) platform. This type of model, in common with the ERC's SATIM model also discussed in this section requires exactly the same disaggregation of energy carriers by sector and sometimes sub-sector as was required for this study.

Petroleum fuels were allocated to sectors in fixed shares using a set of assumptions which are presented in Annexure 6. This adds another alternative view but doesn't essentially help to reduce uncertainty by much, which was the goal of this study, and therefore had limited use.

The energy intensities for cooking, space heating and water heating as used for the Cape Town and Polokwane models, the two available to the study, were, however, directly applied in the residential paraffin and LPG bottom-up models and as they were based on years of experience and hands-on modelling at city-level were extremely useful.

Contact: **Megan Euston-Brown** **<Megan@sustainable.org.za>**

Transnet, direct submission to GHG inventory 2010 and annual reports

This data was made available by the client at the start of the project and is Transnet's own written submission of their annual diesel consumption in response to the publication of the 2010 GHG inventory. It was used to substitute the SAPIA/DoE diesel allocation to Transnet rail which drops off sharply (and implausibly) after 2004.

University of Cape Town, Energy Research Centre, South African TIMES Model

The South African Times Model (SATIM) is a public domain national energy systems model developed by the University of Cape Town's Energy Research Centre (ERC) using the TIMES model generator. TIMES was developed under the auspices of the International Energy Agency's Energy Technology Systems Analysis Program (IEA-ETSAP), an international community operating under an IEA implementing agreement that uses long term energy scenarios to analyse energy and environmental parameters. TIMES is a partial equilibrium linear optimisation model capable of representing the entire energy system, tracking the flow of commodities (including energy, materials, emissions, demand services and water) through the system and determining the capital stock requirements for all technologies embodied in the system, including economic costs.

As such, it requires calibration in the base year at a high level of sectoral disaggregation. The current calibration year is 2006 necessitated by the deteriorating quality of the national energy balance in later years. Even in the case of 2006 a number of adjustments have been made based on bottom-up modelling and in the case of industry, some reconciliation with the supply and use tables. Four industry sub-sectors were disaggregated in SATIM during a CSIR funded project in 2013 which involved a measure of direct survey of the iron and steel, non-metallic minerals, pulp and paper and chemicals sectors. A number of researchers have worked on SATIM over several years and the calibration is constantly improved making it a useful view on sectoral disaggregation for one year.

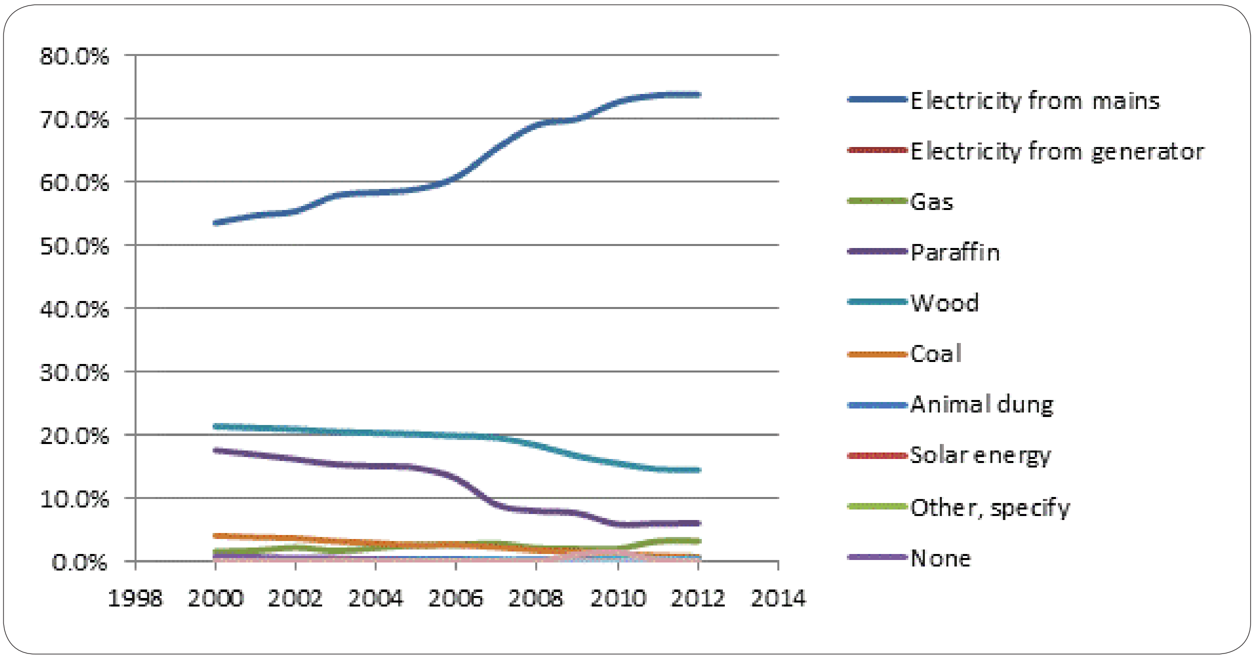


Figure 6: General Household Survey (GHS) data for the primary use share of household use of energy carriers for cooking

Contact: **Bruno Merven**
<brunomerven@gmail.com>

Alison Hughes
<alison.hughes@uct.ac.za>

General Household Surveys (GHS) 2002–2003, 2005–2012

The General Household Survey is published by Stats SA as the Statistical Release P0318 series. The anonymised raw data was made available in Stata files by Datafirst UCT. The introduction to the 2012 publication describes its purpose as follows:

> *The GHS is a household survey that has been performed annually by STATSSA since 2002. The survey was introduced to address a need identified by the Government of South Africa to determine the level of development in the country and to measure, on a regular basis, the performance of programmes and projects that were implemented to address these needs. The survey is specifically designed to measure multiple facets of the living conditions of South African households, as well as the quality of service delivery in a number of key service sectors (Stats SA 2013: 1).*

The share of households' primary use of energy carriers was used as an indirect source in this study. By assuming an energy intensity per household (averaged from the literature, SATIM and SEA's LEAP models) these shares could be used to estimate residential demand for energy carriers from the bottom-up. The share of secondary use had to be assumed or obtained from other sources.

As shown above, the GHS clearly shows a trend in electrification and a decline in domestic use of other carriers as a primary energy source. This trend is very useful in modelling residential demand or validating other sources.

Contact: **Lynn Woolfrey**
<lynn.woolfrey@uct.ac.za>
(Datafirst)

CSIR National Cleaner Production Centre of South Africa (NCPC-SA), Database of Firm Energy Audits, 2007–2012

This database was kindly forwarded by the NCPC-SA and includes the surveyed energy consumption of diesel, illuminating paraffin, LPG, petrol, fuel oil, biomass and wood/charcoal collected from 66 firms in seven provinces between 2007 and 2012 although most of the data is for 2012.

The data was digitised by hand from survey forms for this study and is still being debugged with two data points of the 66 making up 92% of the energy consumed in all the surveys. At the time of writing the plausibility of this had not yet been cleared up through correspondence with the NCPC-SA. However, this data and the city level data discussed below is most valuable as an ongoing input to building up a more realistic picture of industry sub-sector energy consumption, that will develop over time. Like many of the avenues followed up for this project the full value may only be realised after this project.

Contact: **Sashay Ramdharee**
<SRamdharee@csir.co.za>

Alfred Hartzenburg
<AHartzenburg@csir.co.za>

City of Tshwane, Air Quality Management, Large Combustion Source Licensing Database, accessed 2015

This is the database of registration data for large combustion sources compiled by the department responsible for air quality management at the City of Tshwane. It includes licensed consumption of coal, diesel, gas (LPG and natural gas) and electricity for 37 sites.

Paraffin use is not accounted for, and according to the data supplier, the value recorded is the license condition (maximum permitted consumption rate) and not consumption although their opinion is that the value is a good proxy for consumption. The sector and sub-sector is indicated for each site.

Contact: **Jan H. van den Berg**
<Janneman@TSHWANE.GOV.ZA>

City of Johannesburg, Air Quality Management, Large Combustion Source Licensing Database, accessed 2015

This is the database of registration data for large combustion sources compiled by the department responsible for air quality management at the City of Johannesburg. It includes licensed consumption of coal, diesel, electricity, fuel oil, LPG, natural gas and paraffin for 31 sites. The fuel level of disaggregation is good and the energy service or sector is given for many of the sites.

Contact: **Given Mbara**
<GivenMb@joburg.org.za>

Western Cape Provincial Government, Energy Demand Database for Cape Town and Atlantis, 2013

This data was compiled by a consultant for the Western Cape government using telephone and site visit surveys as part of their feasibility study into the supply of imported liquid natural gas (LNG) for Cape Town and Atlantis (Visagie 2013a). It covers all the large energy users of the Cape Town and Atlantis region and records the consumption of HFO, paraffin, LPG, waxy oil, coal and diesel for 170 industrial and commercial sites. The aggregate data by site is available in the project report but the data disaggregated by fuel in electronic format was kindly made available on request (Visagie 2013b).

This database is extremely useful as it records the energy service for which the fuel is used and is based on actual consumption. The results are revealing as to the use of paraffin and fuel oil by industry sectors like 'food and beverages' which is not recorded by the national energy balance.

Contact: **Johan Visagie**
<visagiehj@energybusiness.co.za>

City of Cape Town, Air Quality Management, Large Combustion Source Licensing Database, accessed 2015

This is the database of registration data for large combustion sources compiled by the department responsible for air quality management at the City of Cape Town. It includes licensed consumption of coal, coke, anthracite, diesel, electricity, fuel oil, LPG, gas (LPG?), paraffin, 50/50 blended paraffin and diesel, wood, wood waste and other waste for 576 industrial, commercial, residential and public sites. The data is average monthly consumption for one year (2014).

The datasets reviewed here are the most complete source/company level databases that were found for this study and, given how large the economies of the Cities of Johannesburg and Tshwane are, they are indicative of how partial the air quality management (AQM) datasets and the NCPC-SA dataset currently are. So while this is a promising avenue for the future, much measurement and survey work needs to be done which will be greatly helped by the various air quality management offices sharing data, techniques, experiences and skills.

Contact: **Ed Filby**
<Ed.Filby@capetown.gov.za>

Database of Vessels under Pressure, Department of Labour

Regulations governing industrial safety require that equipment including vessels under pressure over a certain size need to be registered with the Department of Labour's Directorate Electrical and Mechanical Engineering. This includes steam generators (boilers) and often also incinerators and other large industrial heating equipment and the data collected includes the type of fuel and capacity of the equipment (Department of Labour, 2009). This therefore presents an ideal opportunity to assess fuels used in heating from the bottom-up.

The data is collected regionally and resides in the nine provincial offices of the Directorate but unfortunately it became clear in the course of the study that it is not digitised and while the Directorate would like to digitise it, there is currently no active project to do so. The University of Stellenbosch have acquired limited data from this source for an undisclosed project but were unable to assist with more information. This promising source of data, therefore, in the end yielded nothing but may prove useful in the future if it is digitised.

Contact: **Jacob Malatse**
<Jacob.Malatse@labour.gov.za>
(Department of Labour)

Eugene Joubert
< ecjoubert@sun.ac.za >
(University of Stellenbosch)

4. METHODOLOGY

4.1 Study Design Approach

Figure 7 provides an overview of the study design approach. The first key step was to align with the study sponsor on the expected study outcomes and scope definition on a more granular level of detail.

To maximise the value derived from this assignment the 80/20 rule was applied to the sourcing of data with the majority of the effort addressed to the larger sectors and sub-sectors and imputation used more readily on the smaller sectors / sub-sectors. Given the limited resources assigned to this project no direct field survey work was conducted. The data sourcing approach incorporated a listing and assessment of existing data sources by demand sector and energy carrier as well as relevant contacts for further data collection and validation. Data validation was done through:

- interrogation of the underlying assumptions, scope and adopted survey approach by the respective survey data owners;
- comparison of data from multiple sources to establish the consistency and quality of data.

The sourced data had to be prepared for secure processing in the model execution stage. Besides converting the data into a format that would be acceptable for the model, it also entails the alignment of the data to the ISIC Revision 3.1 categorisation.

The data modelling step entailed the incorporation of best available survey data and key estimates per demand sector and energy carrier into the data template with a qualitative judgement / assessment on the quality of sourced data. The data landscape assessment, as described in Section 3, informed the data screening and prioritisation for incorporation into the data template.

The "modelling" process normally requires the preparation of data, the specification (configuration) of the model(s) to be used, the generation of the code required to execute the model(s) and the calibration of parameters that feature in the model(s). This is, however, a reiterative process, namely the quality of the outcome is continually checked against reality and expectations, and amendments applied to the model configuration and parameter calibration. Quality assurance can thus only be achieved by reviewing the resulting outcomes and then adjusting data, assumptions and the model configuration until the results become robust and meaningful.

The final phase entailed the final review, write-up, and professional packaging of the results in a report in line with the brief requirements.

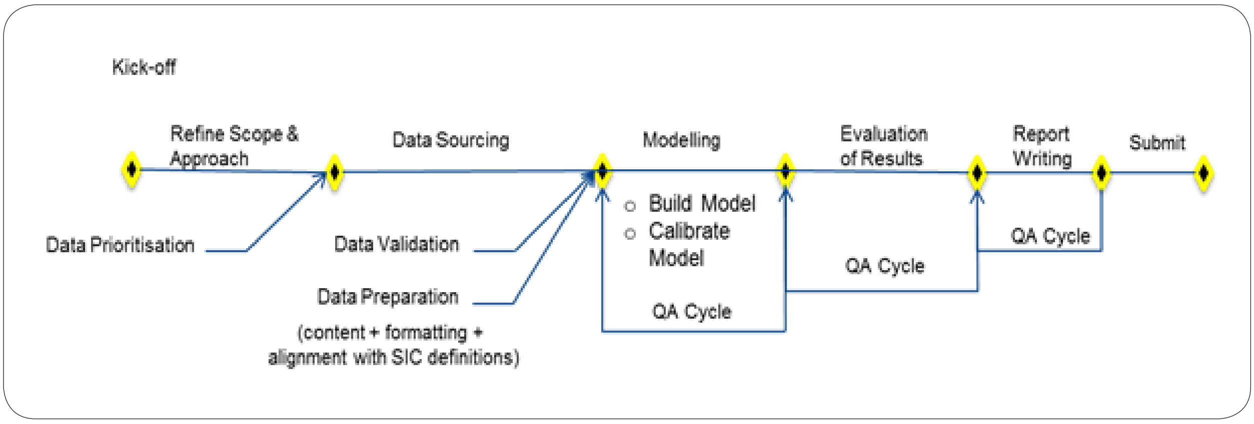


Figure 7: Study design approach

4.2 Data Modelling Approach

The initial outlook on the data modelling process specified four components as reflected in Figure 8. A first iteration of best available data was established from industry survey data, typically the DoE's energy balances, SAPIA's aggregated and disaggregated data sets, the HSRC's residential survey data and the ERC's SATIM model.[5]

In certain instances mediation of sources was necessary to inform the decision (in the absence of complete data) to:

- proceed with model based imputation, or
- apply econometric smoothing to the data, or
- to source further data, or
- a combination of the above.

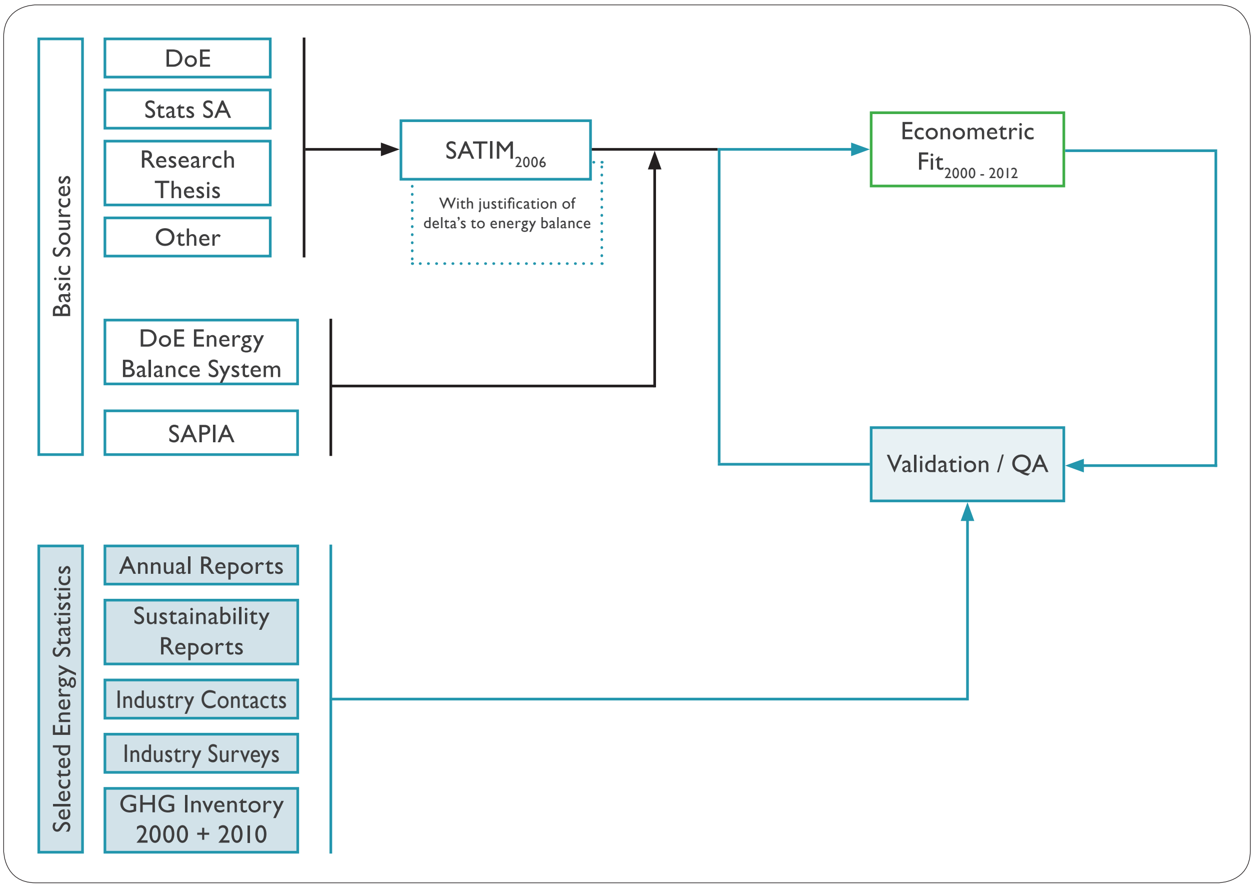


Figure 8: Initial outlook on data sourcing approach

[5] The ERC maintains a model of the South African energy system called SATIM (South African TIMES model) developed over a number of years. There are a number of sub-models but the framework is centred around a so-called MARKAL type model on the TIMES platform. This is a partial equilibrium, linear, least-cost optimisation model. It includes the whole energy chain from primary fuel extraction to transformation in refineries and power stations to end use technologies supplying services such as lighting, heating and passenger travel. As such the base year (2006) calibration requires the same level of disaggregation for the same energy carriers as required by this assignment.

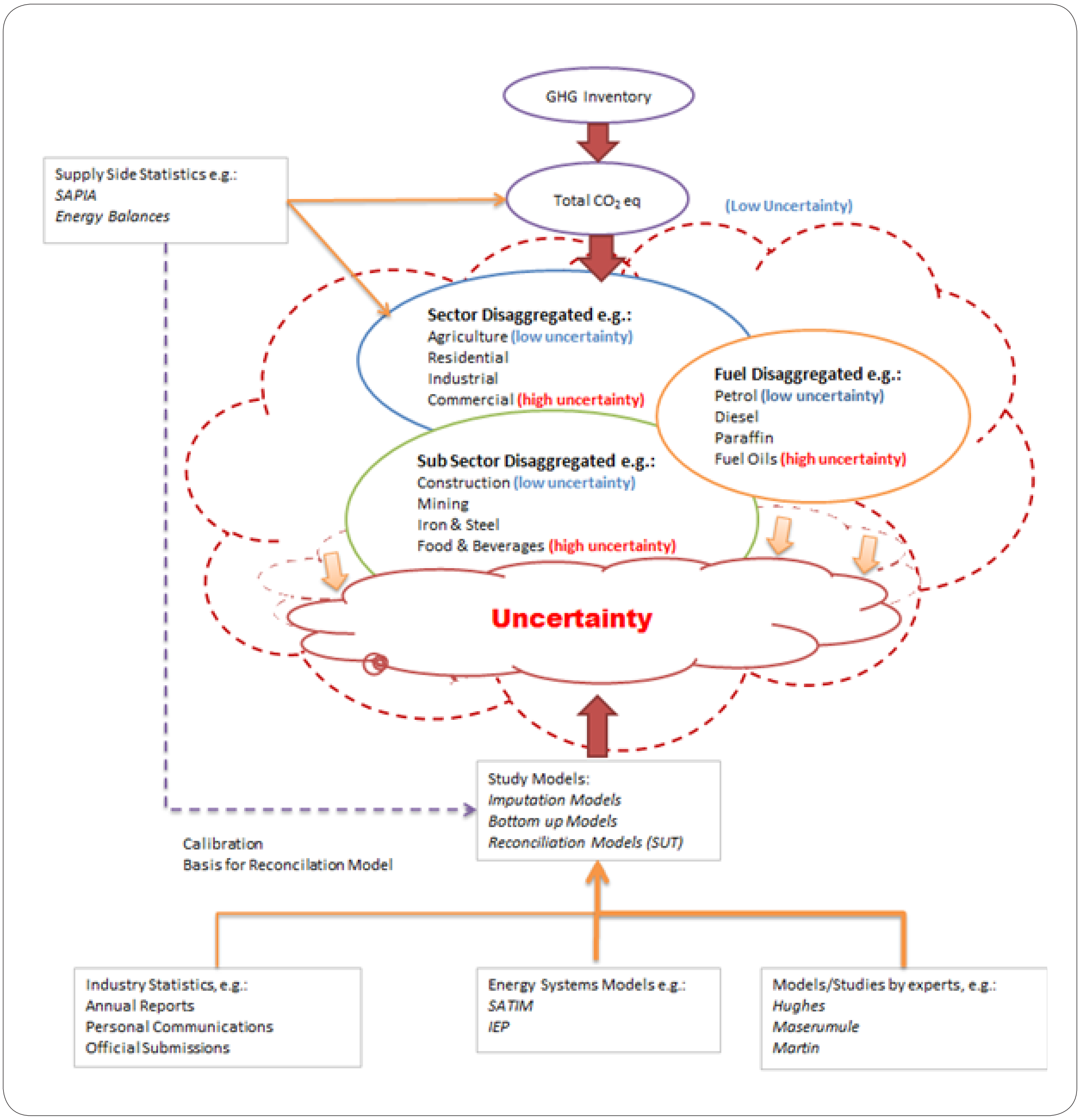


Figure 9: Reconciliation of views to address the issue of uncertainty

A number of spreadsheet based sub-models of varying complexity support SATIM mostly to establish the base year data and the share of demand and supply technologies of production and consumption. Some of these are analytical but most involve a mediation of sources and imputation of missing data or revision of primary sources that fail validation. Typically, for fuel use in industrial and commercial sub-sectors where the primary sources lack detail or are deemed unreliable, commodity flows from the national accounts as published by Stats SA form the basis for an estimate, adjusted by sector specific prices where they are known.

A key aspect of the approach is to adopt the hypothesis that the basic sources of data are the most accurate available, until sufficient evidence has been established to reject the null hypothesis.

Selected energy statistics sourced from annual reports, sustainability reports, industry contacts, as well as other available industry and public surveys were then used to challenge the null hypothesis and in doing so increased the robustness of the study results.

None of the data sources could be considered accurate and authoritative enough at all sector levels, spanning the full period from 2000 to 2012. A considerable level of uncertainty surrounds each source, in at least one or other aspect. However, a combination of the "best" of each does contribute towards a higher level of certainty, and thus a reduction in uncertainty. The diagram (Figure 9) depicts the outcome of the reconciliation that reduced the uncertainty level, from a qualitative, as well as a quantitative point of view.

It would be fair to state that the total CO_2 equivalent inventory can be assessed with reasonable accuracy at national level, as it can be substantiated by aggregate supply statistics. However, the disaggregation in terms of sectoral consumption and energy carriers cannot be propounded with the same level of confidence. One important constraint is the lack of sector categorisation detail that is inherent to all the data sources.

In order to construct a consistent database that reflects reasonable accuracy per sector, the data sources had to be supplemented with imputation models and bottom-up models (where appropriate), and eventually conditioned with the sectoral profiles of petroleum product consumption indicated by the supply and use tables of Stats SA.

The uncertainty relating to some of the energy carriers, for example, LPG, fuel oils, biomass, wood and charcoal, could only be addressed by specific bottom-up models and fragmented survey data. Industry statistics, energy system models and expert opinions provided substantiation for the bottom-up models. The improvement in certainty regarding consumption statistics in these cases is considered to be significant.

Finally, it needs to be stressed that uncertainty would never be eliminated by following the above approach; however, it can be reduced significantly.

4.3 Aggregate sources

It would be prudent to scrutinise the energy consumption statistics of large consumers representative of specific sectors or sub-sectors. This would be extremely useful in understanding the typical energy mix in particular sectors (or sub-sectors). However, access to such data in the public domain is sporadic and often unreliable. Detailed public recording in sustainability reports is also often found to be inconsistent and extremely rare.

In the absence of representative sectoral consumption data, an alternative approach was followed in this assignment. Large producers or suppliers in the energy value chains often record their respective views on the consumption of their products. These could also be used to assess and track consumption profiles in sectors and sub-sectors. A few cases in point are discussed here.

Apart from the DoE, which records energy consumption per selected sector, SAPIA also provides a breakdown of the consumption of a wide variety of their products in the South African market. The SAPIA data is provided not only on an aggregated basis, but also disaggregated per magisterial district or per predefined sector. It should be noted that SAPIA's sectoral split does not directly match that of the ISIC Revision 3.1 categorisation, but it can be mapped to the ISIC Revision 3.1 categorisation. Such a mapping is presented in Annexure 1.

The SAPIA aggregated and disaggregated data was judged to be more reliable and of higher quality than that reported in the DoE energy balances, especially for the period 2008 to 2012. Although the SAPIA aggregated data does not always differ substantially from the DoE energy balances, the disaggregated data do indicate significant differences.

The period 2008 to 2010 is of particular interest, as the economy at large suffered a significant recession and inflation in a rather short period of time. Such events normally reflect reasonably well in diesel consumption, albeit not to the same extent in the consumption of other energy carriers. In this instance, both the SAPIA data and the DoE energy balance data were found to be inconsistent. However, the actual fuel sales data published by the DoE does seem to corroborate the expected correlation with the economic turmoil at the time. It was therefore decided to accept the DoE fuel sales data for 2008–2009 as the most reliable, thereafter the SAPIA data proved to be the most reliable once again.

In the case of natural gas (and equivalents such as methane-rich gas and syngas) the data provided by the largest source in South Africa, namely Sasol, proved to be invaluable. Sasol has recorded a fairly detailed breakdown of consumption per sector, to a level of granularity somewhat similar to the ISIC Revision 3.1. Supplemented by detailed consumption data from Egoli Gas (as representative of distributors' consumption) an acceptable level of accuracy of consumption per sector could be obtained.

4.4 Consumption profiles for different sectors and sub-sectors

A huge amount of disparity exists in the way that the data sources define sectors and sub-sectors. The sectoral divisions of consumption, if used by the data owners, are normally driven by in-house requirements. For example, the DoE records the supply and use of energy according to international standards, and is satisfied with its compliance to the standards. However, the international standards do not require the level of granularity provided in the ISIC Revision 3.1 or IPCC 2006 guidelines. It is therefore not possible to obtain line-for-line matches between the energy balances and the specific sectors and sub-sectors defined by the IPCC guidelines. This observation has a significant impact on the consumption profiles to be obtained for the various sectors and sub-sectors as per the IPCC reporting guidelines.

The ISIC Revision 3.1 categorisation and IPCC guidelines for sectoral breakdown have been mapped and appear in Annexure 1 to this report. A mapping of specific data owner consumption categories to that of ISIC Revision 3.1 was thus required. The mapping is presented in the "Transform" sheets of supporting models to the summary table of data. In all cases, the ISIC Revision 3.1 categorisation provides for the highest level of granularity and can be "translated" into any other specific set of categories, as provided by data owners.

To assist in the mapping between the ISIC Revision 3.1 and the specific data owner categories, it was discovered that the supply and use tables of Stats SA (2008, 2009, 2010-2011, 2012) are of exceptional value. Although the first proper supply and use tables were published in 2005, and then omitted for 2006 and 2007, regular editions from 2008 onwards are obtainable from Stats SA. The 2005 edition tables were compiled at a granularity level higher than the ISIC Revision 3.1 categorisation. This was, unfortunately, not maintained in subsequent periods. The tables for 2008 to 2012 mimic an abridged version of the ISIC Revision 3.1 categories. However, a close matching of the supply and use tables and the ISIC Revision 3.1 categories could be obtained. These mappings also feature in the supporting models to the summary data table.

The consumption profiles for specific low consuming sectors and sub-sectors could thus be imputed by utilising the mapping between the ISIC Revision 3.1 and other categorisation frameworks. These can be found in the summary data table workbook, as well as in the supporting model workbooks that are provided with this report.

4.5 Econometric modelling

4.5.1 Purpose

Since the energy consumption data is essentially provided in a time series format it is possible to apply well-known econometric modelling techniques which will enable the forecasting of consumption of specific energy carriers. This approach can also be utilised to determine missing and doubtful historical values.

The improvement of the historical database will be driven by econometric analysis of trusted observed values from which expected values and potential errors can be estimated.

A convenient and useful approach to quantify the certainty levels of historic data would be to apply a simple exponential smoothing technique on the data. It should be emphasised, though, that this can only be a proxy for the metadata property which could be expressed as "subjective certainty". It is preferable to record the certainty (or uncertainty) of data simultaneously with the actual observation of consumption, but such metadata is not available for this study.

The forecasting of consumption by means of the econometric model(s) yields expected values as well as confidence intervals. The latter can be used to determine low and high scenarios for the future.

4.5.2 Software application selection

The choice of software for the econometric modelling is usually driven by cost, availability and ease of use. It is distinctly advantageous if a combination of software applications can be used to ensure the effective execution of the various steps in the modelling process, namely:

- data accumulation
- data preparation
- model specification
- model estimation
- reporting results
- inference of results

Software applications are particularly useful in the middle four steps, namely data preparation, model specification, estimation and results reporting.

For the purpose of this assessment it is recommended that MS Excel be employed to prepare the data for processing and to capture results for reporting. Excel is well suited for preparing data and for quick and easy reporting of analysis results.

The statistical analyses can be conducted in the programming language called "R". R is open source software which is well documented and the application of choice for many research projects at respected academic and commercial institutions. An elegant interface between R and Excel is also available as an open source application. The interface allows the specification and estimation of models to be conducted and controlled easily from Excel, while R runs in the background.

From a costing, availability and ease of use perspective, the combination of Excel and R is strongly recommended. Most enterprises run platforms with Excel as a standard application while R and the R-Excel interface can easily be downloaded from the internet at no cost. It would, however, be advisable in corporate environments to first check with the local IT department whether the adopted IT governance would allow the implementation of open source software on the general computing platform. Although R has been tested extensively and not found to affect the stability of common platform varieties, one needs to be guided by the in-house IT rules. If local governance prohibits the implementation of open source software such as R, it would be advisable to install the application (with Excel) on a stand-alone platform.

It is to be noted that R consists of a core computing platform which is supplemented by so-called downloadable "packages". The packages are modules that have been developed and peer-reviewed by participants in the R community of users. The packages are therefore well-trusted developments that can be downloaded freely and used for analytical purposes. The R code provided in Annexure 5 contains the instructions to download and install the required packages for use in the analyses.

A very useful introduction to R, as well as the use of econometrics in R, is provided by Kleiber and Zeileis (2008). This book is well written and should be easily readable by beginners in R. A more thorough introduction to the use of time series in R is provided by Cowpertwait and Metcalfe (2009). Those who are interested to learn more about econometrics are advised to consult the books by Gujarati (2003) and Enders (2004). Alternatively, the internet has many massive open online courses (MOOCs) and other short courses which can provide a very fruitful learning experience.

It is recommended that RStudio or RExcel be employed as user interfaces to R. Both these interfaces are open source and freely available from the internet. A simple Google search will allow the interested analyst to download the latest versions from the designated websites.

4.5.3 Modelling methodology

The liquid fuels consumption data is supplemented with economic growth data for the various sector clusters in the standard reporting format adopted by central banks worldwide. The consumption data is considered endogenous while the economic data would be exogenous, namely external to the set of liquid fuels under consideration and driving changes in the endogenous data.

In order to do forecasting of future consumption of the various energy carriers, the following procedure must be followed:

- fit an econometric model on the combined endogenous and exogenous data sets;
- extend the exogenous data to cover the full forecast period;
- use the extended exogenous data to generate forecasts of the endogenous variables.

The DEA expressed the desire to see forecasts of liquid fuels consumption up till 2025. Therefore the exogenous data set needs to be extended beyond the latest observation, which is for the first quarter of 2015.

A vector auto-regression (VAR) modelling approach is applied to forecast the quarterly economic data for the period 2015Q2 till 2025Q4. Although the economic drivers for the various sector clusters can be considered to be independent, the notion that primary clusters affect

Table 7: Extract of endogenous variable dataset

Year	agr_avg	agr_jet	agr_bio	agr_col	agr_dsl	agr_hfo	agr_ip	agr_lpg	agr_gas
2000	0	0	0	1 386.9	6 574.287	1 946.137	8 911.099	819.3061	0
2001	0	0	0	2 030.1	6 869.022	1 979.022	6 944.267	653.2925	0
2002	0	0	0	1 809	7 207.989	1 972.072	5 685.915	486.7042	0
2003	0	0	0	2 452.2	7 664.436	1 970.924	5 102.945	322.3964	0
2004	0	0	0	3 336.6	8 096.155	2 268.166	6 051.377	168.063	0
2005	0	0	0	281.4	8 565.165	2 249.09	5 807.157	136.8912	0
2006	0	0	0	562.8	9 188.651	2 493.028	5 115.251	354.2161	0
2007	0	0	0	394.5027	10 702.1	2 817.248	4 909.19	375.8953	0
2008	0	0	0	1 075.31	10 959	3 542.962	2 549.015	300.9066	0
2009	0	0	0	505.7763	9 958.708	4 239.427	2 616.358	210.0543	0
2010	0	0	0	1 339.464	10 732.57	3 988.006	2 983.361	179.6982	0
2011	0	0	0	7 081.069	11 844.92	4 620.06	3 102.349	124.072	0
2012	0	0	0	163.6542	11 884.5	5 797.05	2 610.884	15.20305	0

secondary clusters, which in turn affect tertiary clusters, is also well founded. Hence a "seemingly unrelated regression" technique is applied. This allows a system of variables to be considered in such a way that all variables (and lags thereof) affect one another to varying degrees. The VAR model provides this capability. The VAR approach was originally designed for sets of endogenous variables, but can be extended to accommodate exogenous variables (as will be done in our case) to become a VARX-model (VAR with exogenous data). VAR models are generally well suited for short term forecasting, but VARX models have the benefit of incorporating longer term exogenous data to yield longer term endogenous forecasts.

4.5.4 Data preparation

In order for R to import and interpret the data it is recommended that the data be prepared in comma separated value (CSV) format with the aid of Excel. The endogenous and exogenous data need to be prepared separately.

The endogenous data is obtained from the "pivots" sheet in the summary data table file. Note that the transposition of the compiled matrix is needed to feed into the R programme.

A typical view on the endogenous variables, namely the liquid fuels consumption data, is reflected in Table 7.

The data set reflects the accumulated energy consumption in terajoules (TJs) per annum for the period 2000 to 2012. The headings indicate the sector and the energy carrier, for example, "agr_col" indicates the consumption of coal in the agriculture sector, while "min_dsl" would signify diesel consumption in the mining sector. It is customary to prepare the time series per sector-energy carrier in columns to yield the time series from 2000 to 2012 per variable.

This Excel layout is then saved in .csv-format as "lfdata.csv" in the directory where R needs to pick it up. It is essential that the number format for all numbers in the table be set to "general" before saving the document.

The naming conventions depicted in Table 8 are applied to the endogenous data.

Table 8: Naming convention for endogenous data

Sector	Energy Carrier	Variable name
Agriculture	Aviation gasoline	agr_avg
	Aviation jet fuel	agr_jet
	Biomass	agr_bio
	Coal	agr_col
	Diesel	agr_dsl
	Heavy fuel oil	agr_hfo
	Illuminating paraffin	agr_ip
	Liquefied petroleum gas	agr_lpg
	Natural gas	agr_gas
	Petrol	agr_pet
Commercial & institutional	Aviation gasoline	cmi_avg
	Aviation jet fuel	cmi_jet
	Biomass	cmi_bio
	Coal	cmi_col
	Diesel	cmi_dsl
	Heavy fuel oil	cmi_hfo
	Illuminating paraffin	cmi_ip
	Liquefied petroleum gas	cmi_lpg
	Natural gas	cmi_gas
	Petrol	cmi_pet

Table 8 continued...

Sector	Energy Carrier	Variable name
Industry (chemicals)	Aviation gasoline	chm_avg
	Aviation jet fuel	chm_jet
	Biomass	chm_bio
	Coal	chm_col
	Diesel	chm_dsl
	Heavy fuel oil	chm_hfo
	Illuminating paraffin	chm_ip
	Liquefied petroleum gas	chm_lpg
	Natural gas	chm_gas
	Petrol	chm_pet
Industry (coke & refined petroleum products)	Aviation gasoline	ref_avg
	Aviation jet fuel	ref_jet
	Biomass	ref_bio
	Coal	ref_col
	Diesel	ref_dsl
	Heavy fuel oil	ref_hfo
	Illuminating paraffin	ref_ip
	Liquefied petroleum gas	ref_lpg
	Natural gas	ref_gas
	Petrol	ref_pet
Industry (construction)	Aviation gasoline	cns_avg
	Aviation jet fuel	cns_jet
	Biomass	cns_bio
	Coal	cns_col
	Diesel	cns_dsl
	Heavy fuel oil	cns_hfo
	Illuminating paraffin	cns_ip
	Liquefied petroleum gas	cns_lpg
	Natural gas	cns_gas
	Petrol	cns_pet
Industry (food, beverages & tobacco)	Aviation gasoline	fbt_avg
	Aviation jet fuel	fbt_jet
	Biomass	fbt_bio
	Coal	fbt_col
	Diesel	fbt_dsl
	Heavy fuel oil	fbt_hfo
	Illuminating paraffin	fbt_ip
	Liquefied petroleum gas	fbt_lpg
	Natural gas	fbt_gas
	Petrol	fbt_pet

Table 8 continued...

Sector	Energy Carrier	Variable name
Industry (iron & steel)	Aviation gasoline	irs_avg
	Aviation jet fuel	irs_jet
	Biomass	irs_bio
	Coal	irs_col
	Diesel	irs_dsl
	Heavy fuel oil	irs_hfo
	Illuminating paraffin	irs_ip
	Liquefied petroleum gas	irs_lpg
	Natural gas	irs_gas
	Petrol	irs_pet
Industry (machinery)	Aviation gasoline	mac_avg
	Aviation jet fuel	mac_jet
	Biomass	mac_bio
	Coal	mac_col
	Diesel	mac_dsl
	Heavy fuel oil	mac_hfo
	Illuminating paraffin	mac_ip
	Liquefied petroleum gas	mac_lpg
	Natural gas	mac_gas
	Petrol	mac_pet
Industry (mining)	Aviation gasoline	min_avg
	Aviation jet fuel	min_jet
	Biomass	min_bio
	Coal	min_col
	Diesel	min_dsl
	Heavy fuel oil	min_hfo
	Illuminating paraffin	min_ip
	Liquefied petroleum gas	min_lpg
	Natural gas	min_gas
	Petrol	min_pet
Industry (non-ferrous metals)	Aviation gasoline	nfm_avg
	Aviation jet fuel	nfm_jet
	Biomass	nfm_bio
	Coal	nfm_col
	Diesel	nfm_dsl
	Heavy fuel oil	nfm_hfo
	Illuminating paraffin	nfm_ip
	Liquefied petroleum gas	nfm_lpg
	Natural gas	nfm_gas
	Petrol	nfm_pet

Table 8 continued...

Sector	Energy Carrier	Variable name
Industry (non-metallic minerals)	Aviation gasoline	nmm_avg
	Aviation jet fuel	nmm_jet
	Biomass	nmm_bio
	Coal	nmm_col
	Diesel	nmm_dsl
	Heavy fuel oil	nmm_hfo
	Illuminating paraffin	nmm_ip
	Liquefied petroleum gas	nmm_lpg
	Natural gas	nmm_gas
	Petrol	nmm_pet
Industry (power generation)	Aviation gasoline	pow_avg
	Aviation jet fuel	pow_jet
	Biomass	pow_bio
	Coal	pow_col
	Diesel	pow_dsl
	Heavy fuel oil	pow_hfo
	Illuminating paraffin	pow_ip
	Liquefied petroleum gas	pow_lpg
	Natural gas	pow_gas
	Petrol	pow_pet
Industry (pulp & paper)	Aviation gasoline	plp_avg
	Aviation jet fuel	plp_jet
	Biomass	plp_bio
	Coal	plp_col
	Diesel	plp_dsl
	Heavy fuel oil	plp_hfo
	Illuminating paraffin	plp_ip
	Liquefied petroleum gas	plp_lpg
	Natural gas	plp_gas
	Petrol	plp_pet
Industry (other)	Aviation gasoline	oth_avg
	Aviation jet fuel	oth_jet
	Biomass	oth_bio
	Coal	oth_col
	Diesel	oth_dsl
	Heavy fuel oil	oth_hfo
	Illuminating paraffin	oth_ip
	Liquefied petroleum gas	oth_lpg
	Natural gas	oth_gas
	Petrol	oth_pet

Table 8 continued...

Sector	Energy Carrier	Variable name
Residential	Aviation gasoline	res_avg
	Aviation jet fuel	res_jet
	Biomass	res_bio
	Coal	res_col
	Diesel	res_dsl
	Heavy fuel oil	res_hfo
	Illuminating paraffin	res_ip
	Liquefied petroleum gas	res_lpg
	Natural gas	res_gas
	Petrol	res_pet
Transport (domestic aviation)	Aviation gasoline	avd_avg
	Aviation jet fuel	avd_jet
	Biomass	avd_bio
	Coal	avd_col
	Diesel	avd_dsl
	Heavy fuel oil	avd_hfo
	Illuminating paraffin	avd_ip
	Liquefied petroleum gas	avd_lpg
	Natural gas	avd_gas
	Petrol	avd_pet
Transport (international aviation)	Aviation gasoline	avi_avg
	Aviation jet fuel	avi_jet
	Biomass	avi_bio
	Coal	avi_col
	Diesel	avi_dsl
	Heavy fuel oil	avi_hfo
	Illuminating paraffin	avi_ip
	Liquefied petroleum gas	avi_lpg
	Natural gas	avi_gas
	Petrol	avi_pet
Transport (domestic waterborne navigation)	Aviation gasoline	wwd_avg
	Aviation jet fuel	wwd_jet
	Biomass	wwd_bio
	Coal	wwd_col
	Diesel	wwd_dsl
	Heavy fuel oil	wwd_hfo
	Illuminating paraffin	wwd_ip
	Liquefied petroleum gas	wwd_lpg
	Natural gas	wwd_gas
	Petrol	wwd_pet

Table 8 continued...

Sector	Energy Carrier	Variable name
Transport (international waterborne navigation)	Aviation gasoline	wwi_avg
	Aviation jet fuel	wwi_jet
	Biomass	wwi_bio
	Coal	wwi_col
	Diesel	wwi_dsl
	Heavy fuel oil	wwi_hfo
	Illuminating paraffin	wwi_ip
	Liquefied petroleum gas	wwi_lpg
	Natural gas	wwi_gas
	Petrol	wwi_pet
Transport (railways)	Aviation gasoline	tra_avg
	Aviation jet fuel	tra_jet
	Biomass	tra_bio
	Coal	tra_col
	Diesel	tra_dsl
	Heavy fuel oil	tra_hfo
	Illuminating paraffin	tra_ip
	Liquefied petroleum gas	tra_lpg
	Natural gas	tra_gas
	Petrol	tra_pet
Transport (road)	Aviation gasoline	trd_avg
	Aviation jet fuel	trd_jet
	Biomass	trd_bio
	Coal	trd_col
	Diesel	trd_dsl
	Heavy fuel oil	trd_hfo
	Illuminating paraffin	trd_ip
	Liquefied petroleum gas	trd_lpg
	Natural gas	trd_gas
	Petrol	trd_pet

Note that the user can define his/her own names for the endogenous variables. Any deviation from the above may, however, affect the R code provided in Annexure 5, which is required to process the data. If name changes are applied, the R code will need to be amended. The analysis, as indicated in this report, can, however, still be performed.

It is to be noted that although the variable names in Table 8 follow the recognised convention, not all variables would necessarily be applicable or contain non-zero values. Nor would all of the above variables necessarily be included in the forecast modelling – the analyst can make a selection of only those that would be considered relevant.

Table 9: Extract of exogenous data

AFF	MQ	MAN	EGW	CONS	WRCA	TSC	FIREBS	GOV	CSPS	GDP	TAX-SUB
24 378	187 485	54 655	7 210	14 702	54 038	24 839	66 256	90 907	35 753	565 040	4 817
23 276	194 044	57 066	7 390	15 199	55 168	24 241	68 677	91 755	31 675	574 255	5 764
21 456	192 775	61 805	7 585	15 776	59 923	24 471	67 059	93 505	34 525	584 836	5 956
25 428	191 167	58 417	7 842	16 412	59 644	24 495	68 791	96 120	34 455	590 609	7 838
22 039	198 004	59 598	7 561	15 834	60 040	24 446	68 409	96 499	41 491	594 892	971
23 192	200 618	59 492	7 873	15 037	55 579	2 249.09	68 689	97 653	36 314	592 293	2 884
26 806	203 975	61 080	8 099	15 164	57 208	2 493.028	71 658	98 501	30 444	605 522	7 625
26 011	203 155	62 196	8 052	15 314	60 114	2 817.248	71 861	100 629	31 036	611 027	7 673
23 916	214 989	64 655	8 270	15 360	60 730	3 542.962	71 556	102 108	36 751	626 831	2 716
26 101	214 076	64 549	8 239	15 649	61 156	4 239.427	72 256	103 642	34 959	630 847	4 612
26 972	216 412	65 078	8 426	15 788	62 315	3 988.006	72 970	105 355	36 581	643 161	6 821
26 620	223 698	64 793	8 426	15 846	63 357	4 620.06	75 352	108 584	33 251	651 399	4 383
24 653	223 775	69 459	8 723	15 615	65 324	5 797.05	72 893	102 668	41 801	660 015	8 170
26 236	225 864	72 781	8 723	17 302	69 506	27 777	75 174	105 536	36 651	677 468	11 918

The exogenous data should be arranged in the structure as depicted in Table 9.

Note that the values reflect gross value added in constant 2010 rand values, and are reported quarterly from 1960 to the first quarter of 2015. Again the time series need to be presented column-wise.

The headings signify the following sectoral aggregations:

- AFF – agriculture, forestry and fishing
- MQ – mining and quarrying
- MAN – manufacturing industry
- EGW – electricity, gas and water
- CONS – construction industry
- WRCA – wholesale, retail, catering and accommodation
- TSC – transport, storage and communication
- FIREBS – finance, insurance, real estate and business services
- GOV – general government services
- CSPS – community, social and personal services
- GDP – gross domestic product
- TAXSUB – taxes and subsidies

The real gross value added per sector cluster, as provided by the South African Reserve Bank, have been selected as exogenous variables for two reasons, namely:

- It provides a good indication of the economic activity for each sector cluster, which contributes towards the total GDP of the country;
- As the generally accepted primary economic indicators, various long term views or forecasts of these values can be obtained from respectable sources of economic research.

It is, however, to be noted that in this assessment no long term forecasts of the exogenous variables will be sourced from other parties, but an econometric model will be used to generate the long term forecasts in situ.

The above should be prepared in Excel and saved as a unique file labelled "qexodata.csv" in the directory where R needs to pick it up.

The R code required to generate the long term forecasts for the exogenous data is provided in Annexure A5.2 (lines 1-244). In this section a VAR model is specified and fitted on observed data and then used to generate forecasts for the exogenous variables up to 2025–Q4. The output is written into the source directory under the filename "newExoData.csv".

The accumulated data for liquid fuels consumption is recorded on an annual basis, while the economic data obtained from the Reserve Bank database is provided on quarterly levels. It is suggested that the analyst disaggregate the liquid fuels data to quarterly levels as well. This can be achieved by applying the R code provided in Annexure A5.3 (lines 245-326). The output is written into the source directory in a file called "quarterlyData.csv".

The disaggregation of the liquid fuels data is driven by the singular fits of the endogenous variables on the set of exogenous data. Seasonal effects are therefore captured as determined by the relevant endogenous variables.

4.5.5 Forecasting model specification

As indicated in Section 4.5.3 it would be prudent to use the VARX modelling methodology to generate forecasts for the period 2013 to 2025 as requested by the DEA. In order to generate the best possible forecasts, it would be necessary to optimise the parameters of the model(s).

The aspects that normally need to be considered when specifying a VARX model are the following:

- the number of lags required for the endogenous variables
- the number of lags required for the exogenous variables
- the level of differencing required to allow proper estimation of the model

Another condition that can often improve the forecasting capability of the model is to know whether co-integration exists between two or more variables in the data set, and then to convert the VAR(X) model into a so-called vector error correction model (VECM) with or without exogenous variable input. Such models allow for internal correction of forecasts based on the estimated errors (residuals) that the model can generate.

The R code provided in Section A5.4 (lines 327–732) allows for the fitting of a simultaneous equation VECM model with exogenous variables; co-integration has been detected in the dataset and the preferred model would therefore be a VECM without differencing of the variables.

4.5.6 Model estimation

The prepared data is processed by applying a VECM model fit on the preferred historical data set which is discussed in Section 4.5.3. In order to avoid the condition of multi-collinearity that leads to singular matrices in the estimation process, the number of lags specified for the VECM model is adjusted (the value of K in the R code, line 470 in Section A5.4) to yield useful results (not "not-a-number" (NANs)). The default number of lags is 2.

Once the energy carriers have been analysed and forecasts generated, the distribution of energy carriers into the various sectors and sub-sectors can be estimated. It is, however, possible to have the R code prepare the forecasts for individual sector-carrier combinations. This can be achieved by changing the "lfagg.ts" series in line 469 (Section A5.4) to "lfendo.ts[,c("agr_col","min_col")]" where any selection of the sector-carriers can be added in the sequence indicated by "c(... , …)". The latter is a "combination" instruction in R. It is advised that the analyst always select at least two such variables to include, and not more than ten at a time. The R code considers the combination of sector-carrier series to all affect one another; it thus becomes a simultaneous equation problem which usually yields better forecasting results.

4.5.7 Results

4.5.7.1 Consumption forecasting

The following figures reflect quarterly forecasts for the various energy carriers that have been obtained from the R code in Section A5.4. The confidence intervals for the forecasts vary from 70% to 90%. Energy consumption amounts are measured in TJ/a.

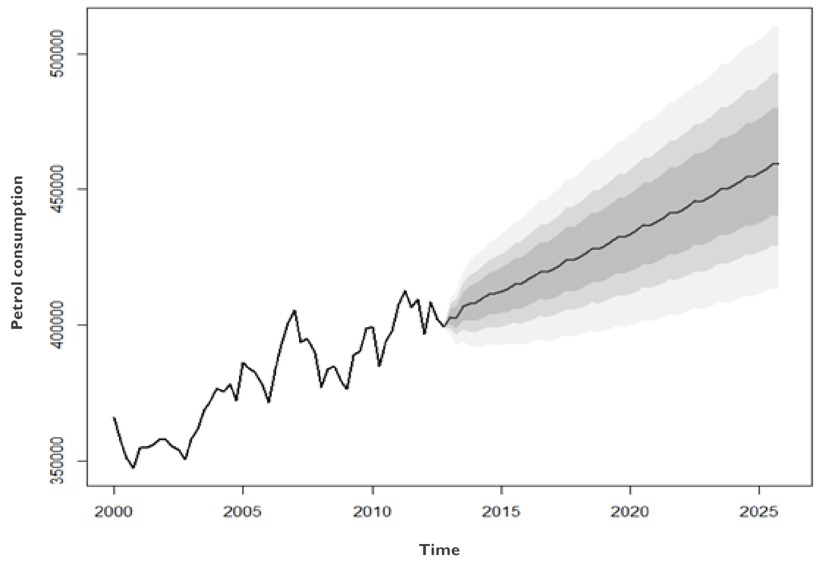


Figure 10: Petrol forecast with 70%, 80% and 90% confidence intervals

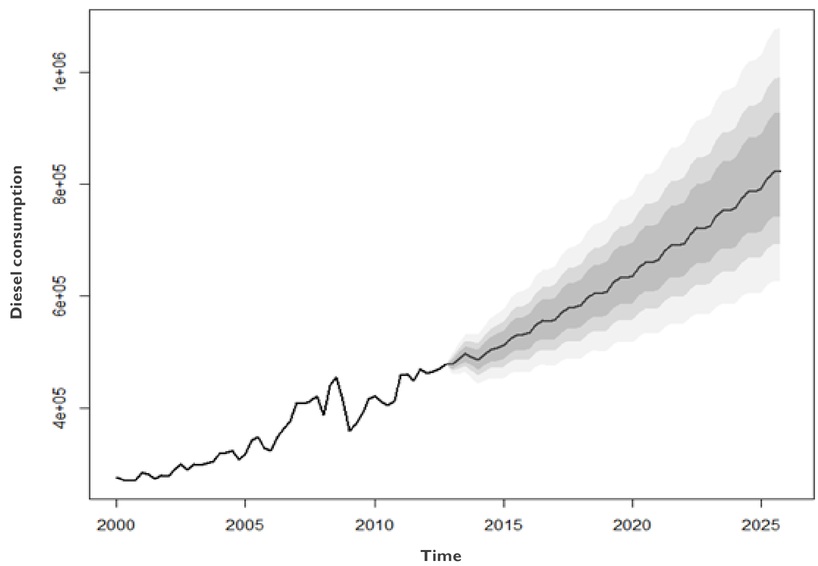


Figure 11: Diesel forecast with 70%, 80% and 90% confidence intervals

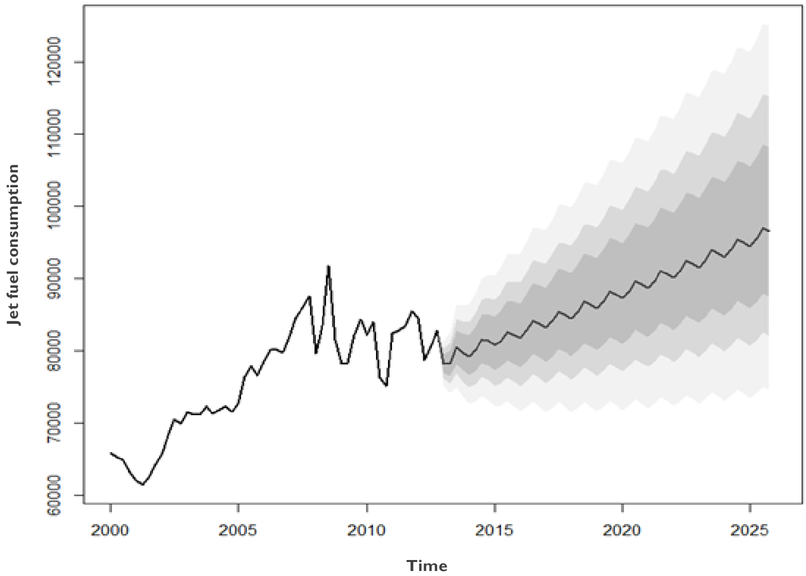


Figure 12: Jet fuel forecast with 70%, 80% and 90% confidence intervals

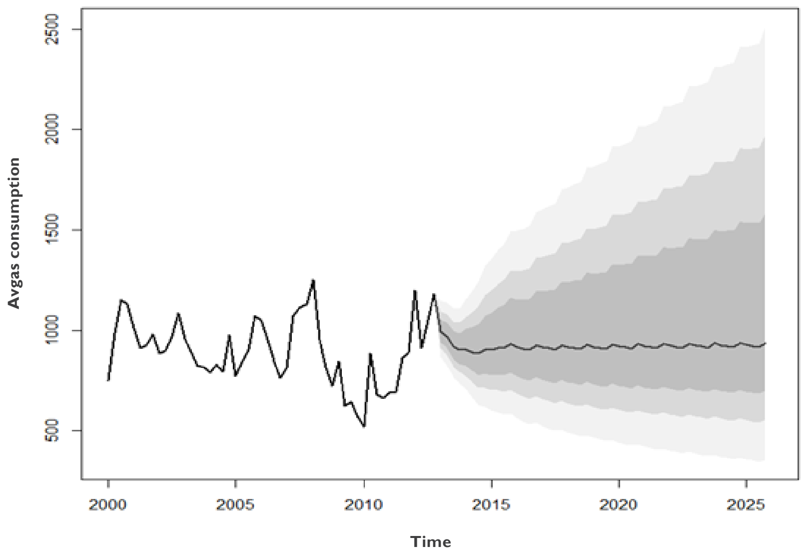


Figure 13: Avgas forecast with 70%, 80% and 90% confidence intervals

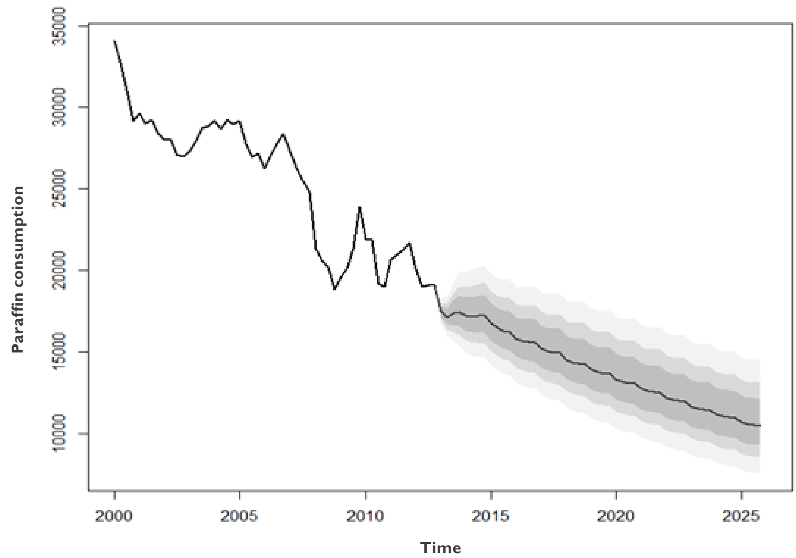


Figure 14: Paraffin forecast with 70%, 80% and 90% confidence intervals

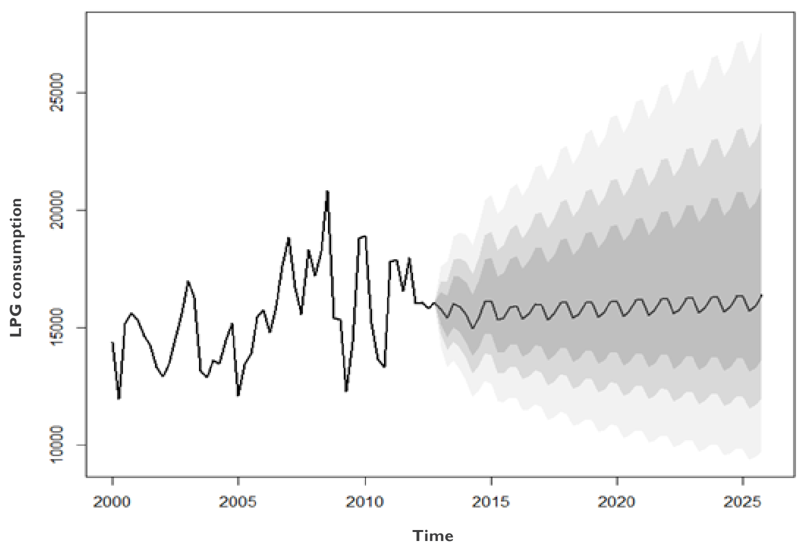


Figure 15: LPG forecast with 70%, 80% and 90% confidence intervals

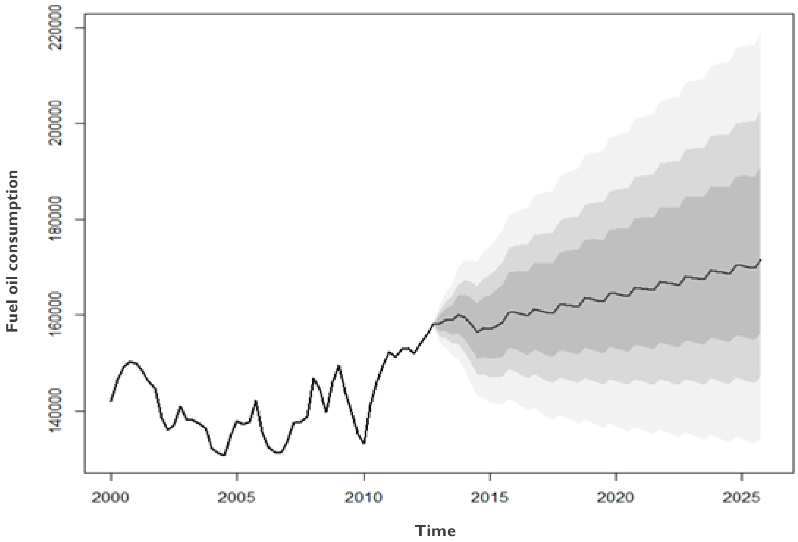


Figure 16: Fuel oil forecast with 70%, 80% and 90% confidence intervals

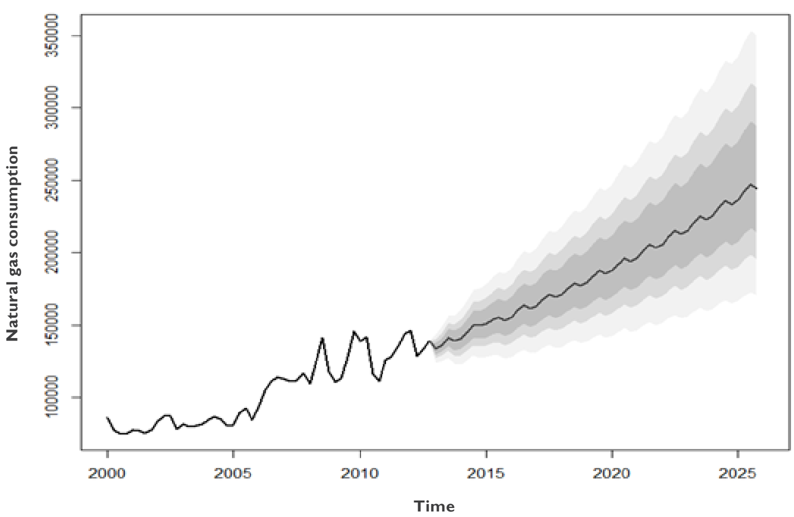


Figure 17: Natural gas forecast with 70%, 80% and 90% confidence intervals

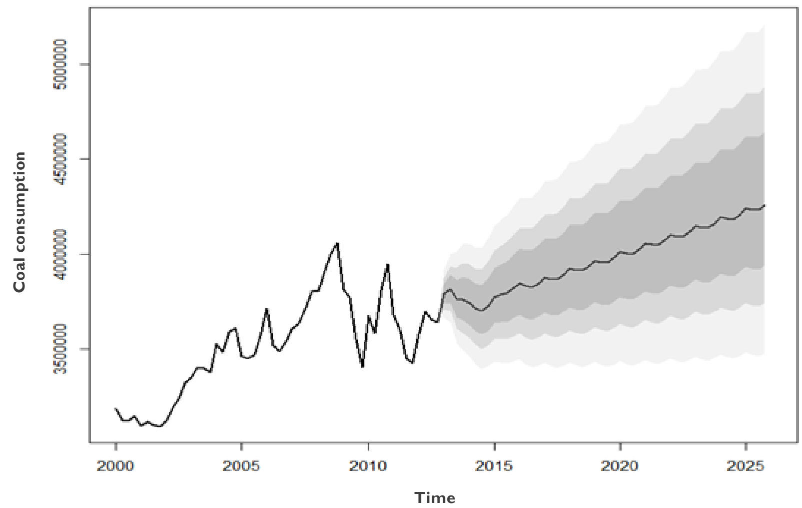


Figure 18: Coal forecast with 70%, 80% and 90% confidence intervals

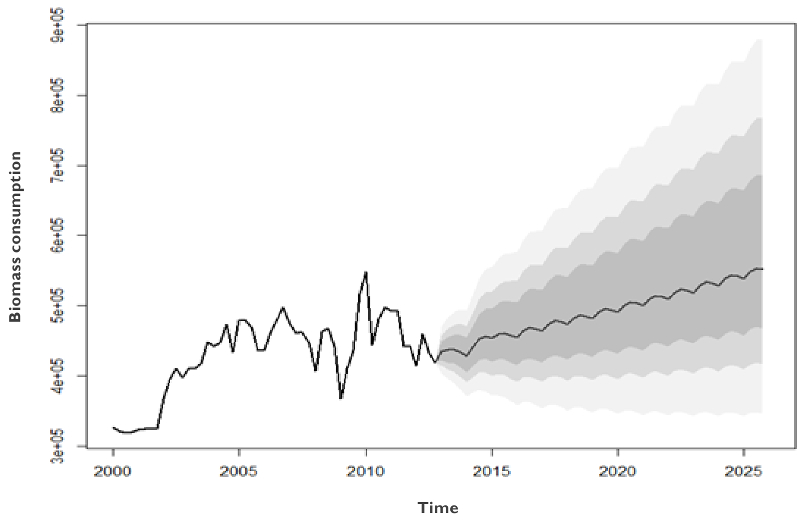


Figure 19: Biomass forecast with 70%, 80% and 90% confidence intervals

It is evident that the expected growth rate for diesel is significantly higher than that of petrol, and the forecast indicates that it would continue to outgrow petrol until 2025. The diesel use for power generation has become significant and its use would appear to be maintained in future.

The forecasts for jet fuel, paraffin and LPG corroborate the current views regarding future consumption growth. The use of illuminating paraffin is discouraged and hence the expected decline in consumption. The historic consumption of LPG, however, indicates high volatility and it is propagated into the future as indicated by the large confidence interval bands. Further work would be required on LPG consumption to reduce the huge volatility.

The forecast for avgas also indicates limited growth but with high volatility as can be observed from the confidence intervals.

The consumption forecast for fuel oil is somewhat surprising as one would expect this energy carrier to grow at a similar rate to the other carriers (except paraffin, of course). However, the limited marine fuel oil market in South Africa clearly has an impact on the strategy to refine deeper into the barrel, namely to convert as much black product as possible into white. The need for a higher white product recovery has a significant impact on the capital and operating expense requirements to run the additional cracking facilities at the refineries.

The forecast for natural gas starts from a rather low base and appears to gain popularity as an energy carrier. This is certainly the preferred carrier for the future, especially with the GHG emissions coming under stricter control with the imminent introduction of carbon taxes.

The forecast for coal consumption indicates a recovery by 2020 to about the levels achieved in 2008. This may, however, be affected significantly when a carbon tax is introduced.

Biomass consumption is surrounded by uncertainty and can only be assessed by bottom up estimates. The forecast does, however, indicate a slight increase in future, probably driven by population growth rather than economic activity.

It is very important to realise that the forecasts need to be interpreted in terms of the confidence levels. These ranges indicate the areas which can accommodate the growth path or evolution with a certain level of certainty. The potential upside, as well as the downside, needs to be considered when stating forecast views; it would be risky to advocate growth if only the expected values (solid lines in the graphs) are cited as evidence.

It is to be noted that although the above forecasts are reported graphically, the actual numbers used to plot the graphs can easily be extracted from R by simply instructing it to export to Excel or to write it out to disc in an Excel-formatted .csv-file.

The above forecasts only address the expectation of aggregate demand for the energy carriers. However, the demand of energy per sector or sub-sector can be projected in the forecasts, based on the current distribution patterns, or on trends in the distribution patterns. This will be based on the sector and sub-sector shares of total consumption as indicated in the comprehensive data table. Any additional strategic implications, which could affect future sectoral distribution patterns, would also be incorporated to generate the forecasts on a disaggregated level.

4.5.7.2 Certainty levels and potential ranges of historic data

The approach to deal with uncertainty is discussed in Section 4.2. The improvement in levels of certainty obtained from the approach is, however, often subjective; the confidence in consumption numbers are reflected qualitatively. However, an attempt to quantify certainty levels may be warranted. It may be useful to have some measure of quantification available which can be used to challenge the qualitative opinions on certainty and risk. The proposed approach here is to accept the qualitative views as the null hypothesis and only reject it if quantitative evidence provides enough justification for its rejection.

In order to be able to quantify certainty levels, it is noteworthy that certainty is generally assumed to be related to the inverse of variability (variance). Variance is associated with "risk", which is generally interpreted as being a function of probability and magnitude of undesired outcomes. The magnitude and the outcome probability are assumed to be orthogonal concepts. The measurement of magnitude may thus be applied to desired as well as undesired outcomes. It is, however, pertinent in this study to reflect on the concept of probability of specific outcomes. In our case we would like to associate the probability of a consumption value of any particular energy carrier in a given sector to be reflective of the "truth", or the expectation of it to be indicative of a certainty level, namely the larger the deviation of a number from its expected outcome, the lower the certainty level.

The probability of deviating from expectation can be obtained by applying some econometric smoothing to the data, in order to determine the expected level of sector-energy at any particular point in the time series. The deviation of the actual observation from the expected level would provide an indication of certainty. The R code in Section A5.5 (lines 733-783) can be used to generate certainty levels for all input data-points.

The certainty levels of the data reflected in Table 7 (reproduced below) would be as reflected in Table 10.

(Table 7: Extract of endogenous variable dataset)

Year	agr_avg	agr_jet	agr_bio	agr_col	agr_dsl	agr_hfo	agr_ip	agr_lpg	agr_gas
2000	0	0	0	1 386.9	6 574.287	1 946.137	8 911.099	819.3061	0
2001	0	0	0	2 030.1	6 869.022	1 979.022	6 944.267	653.2925	0
2002	0	0	0	1 809	7 207.989	1 972.072	5 685.915	486.7042	0
2003	0	0	0	2 452.2	7 664.436	1 970.924	5 102.945	322.3964	0
2004	0	0	0	3 336.6	8 096.155	2 268.166	6 051.377	168.063	0
2005	0	0	0	281.4	8 565.165	2 249.09	5 807.157	136.8912	0
2006	0	0	0	562.8	9 188.651	2 493.028	5 115.251	354.2161	0
2007	0	0	0	394.5027	10 702.1	2 817.248	4 909.19	375.8953	0
2008	0	0	0	1 075.31	10 959	3 542.962	2 549.015	300.9066	0
2009	0	0	0	505.7763	9 958.708	4 239.427	2 616.358	210.0543	0
2010	0	0	0	1 339.464	10 732.57	3 988.006	2 983.361	179.6982	0
2011	0	0	0	7 081.069	11 844.92	4 620.06	3 102.349	124.072	0
2012	0	0	0	163.6542	11 884.5	5 797.05	2 610.884	15.20305	0

Table 10: Extract of certainty levels associated with input data-points

Certainty	agr_avg	agr_jet	agr_bio	agr_col	agr_dsl	agr_hfo	agr_ip	agr_lpg	agr_gas
2000	21.2%	21.2%	21.2%	56.7%	44.6%	63.0%	51.5%	42.1%	21.2%
2001	14.8%	14.8%	14.8%	84.0%	96.7%	52.9%	62.3%	53.8%	14.8%
2002	7.2%	7.2%	7.2%	67.9%	38.5%	69.2%	9.1%	99.3%	7.2%
2003	5.2%	5.2%	5.2%	14.1%	28.3%	4.9%	6.0%	22.7%	5.2%
2004	10.6%	10.6%	10.6%	2.7%	19.0%	41.5%	12.9%	0.8%	10.6%
2005	12.4%	12.4%	12.4%	2.0%	19.8%	6.5%	4.2%	0.8%	12.4%
2006	7.7%	7.7%	7.7%	34.1%	60.4%	17.1%	8.3%	48.6%	7.7%
2007	7.2%	7.2%	7.2%	12.8%	0.1%	47.3%	2.2%	12.0%	7.2%
2008	5.9%	5.9%	5.9%	51.0%	0.4%	0.6%	0.1%	9.2%	5.9%
2009	5.2%	5.2%	5.2%	31.2%	0.3%	0.0%	0.9%	14.2%	5.2%
2010	4.1%	4.1%	4.1%	27.0%	9.3%	8.0%	94.8%	7.1%	4.1%
2011	6.3%	6.3%	6.3%	0.1%	33.6%	12.1%	21.3%	16.0%	6.3%
2012	15.8%	15.8%	15.8%	6.6%	91.2%	97.5%	77.7%	3.7%	15.8%

Table 11: Extract of upper limits for historical data-points (90% confidence)

Upper Limits	(90% confidence interval)							
	agr_avg	agr_jet	agr_bio	agr_col	agr_dsl	agr_hfo	agr_ip	agr_lpg
2000	0	0	0	8 296	6 771	2 061	10 182	1 212
2001	0	0	0	4 694	7 153	2 052	8 616	934
2002	0	0	0	3 146	7 580	2 091	7 482	768
2003	0	0	0	2 502	8 057	2 180	6 671	675
2004	0	0	0	2 230	8 569	2 317	6 048	610
2005	0	0	0	2 072	9 091	2 500	5 508	541
2006	0	0	0	1 926	9 602	2 730	5 002	460
2007	0	0	0	1 769	10 091	3 016	4 523	370
2008	0	0	0	1 625	10 558	3 371	4 080	284
2009	0	0	0	1 555	11 017	3 823	3 698	213
2010	0	0	0	1 665	11 499	4 417	3 410	162
2011	0	0	0	2 112	12 036	5 219	3 231	130
2012	0	0	0	3 168	12 632	6 305	3 145	109

Table 12: Extract of lower limits for historical data-points (90% confidence)

Upper Limits	(90% confidence interval)							
	agr_avg	agr_jet	agr_bio	agr_col	agr_dsl	agr_hfo	agr_ip	agr_lpg
2000	0	0	0	597	6 041	1 748	6 479	245
2001	0	0	0	689	6 582	1 820	6 196	291
2002	0	0	0	706	7 104	1 904	5 787	310
2003	0	0	0	640	7 594	2 001	5 278	295
2004	0	0	0	543	8 060	2 121	4 745	258
2005	0	0	0	465	8 520	2 277	4 260	218
2006	0	0	0	417	8 985	2 481	3 846	181
2007	0	0	0	397	9 458	2 746	3 499	149
2008	0	0	0	396	9 931	3 086	3 201	120
2009	0	0	0	398	10 384	3 510	2 925	93
2010	0	0	0	374	10 778	4 023	2 638	65
2011	0	0	0	310	11 076	4 629	2 324	40
2012	0	0	0	228	11 270	5 347	2 001	22

Note that the values in the green cells in Table 10 indicate high certainty, while those with a red background indicate low certainty.

The DEA expressed the desire to have upper and lower limits defined for the historical data-points. This can be associated with a specific confidence interval around the expected value per input cell. For the purpose of this assessment, a 90% confidence interval has been selected. The values in Table 11 indicate an extract of the upper limits, and those in Table 12 the lower limits, of the specific historical data-points.

Any other two-tail probability may be selected for a confidence interval; the value can be introduced as the value to variable "cii" in line 749 of the R code in Section A5.5.

Note that the R code generates output files that contain the complete sets of certainty levels, as well as upper and lower limits associated with all data points that feature in the endogenous variable input-file.

5. INSIGHTS ARISING FROM A RECONCILIATION OF SOURCES

Industry energy usage is particularly poorly understood because of the diverse types of users, the different energy carrier prices in different regions, particularly for coal, and the fact that a comprehensive survey has not been conducted since 1975. As discussed above in Section 3, some firm/plant level data was available from the cities of Cape Town (CTCC), Tshwane and Johannesburg (CoJ), the Western Cape Provincial Government (WCPG) and the CSIR's National Cleaner Production Centre of South Africa(NCPC-SA). This was combined in a database with national level views from the SATIM model, the DoE energy balances 2000–2012, the IEA energy balances 2000–2012 and the 2003 SANEA study for the year 2000. This database is included in this study in the accompanying spreadsheet "database of sources" referred to in Annexure 3. The results of a comparative analysis provide some interesting insights into where the energy balance in particular probably needs review. At this time Cape Town is overwhelmingly overrepresented in the firm level data and so the views obtained cannot be said to be conclusive but provide a useful foundation on which to build.

A comparison of the energy carrier share of the DoE and IEA energy balances is itself quite useful as shown in Figures 20 and 21.

Aside from smoother trends the IEA energy balances include small but significant allocations to biomass and fuel oil for industry not evident in the DoE energy balances. A comparison of average energy carrier shares for all the sources tends to support the IEA's view as shown in Figure 22, even allowing that the sources don't cover the same time window.

The alternative national sources SATIM, SANEA and IEA are in broad agreement on biomass share to industry and three of the firm level sources Tshwane and the two Cape sources, the City of Cape Town and the Western

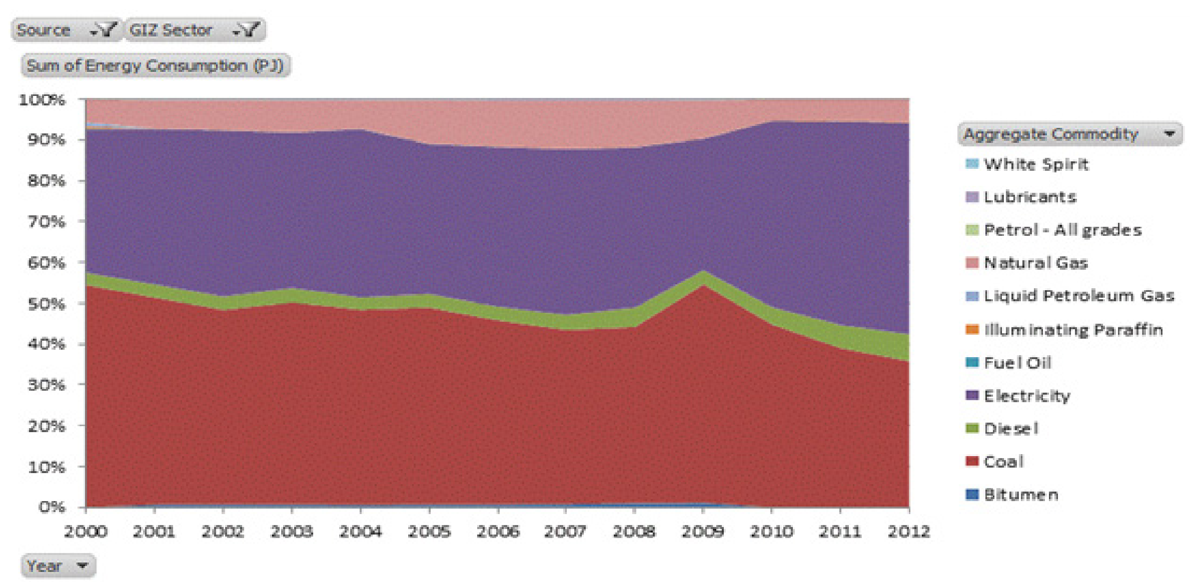


Figure 20: Energy carrier shares in the industry sector - DoE energy balances 2000–2012

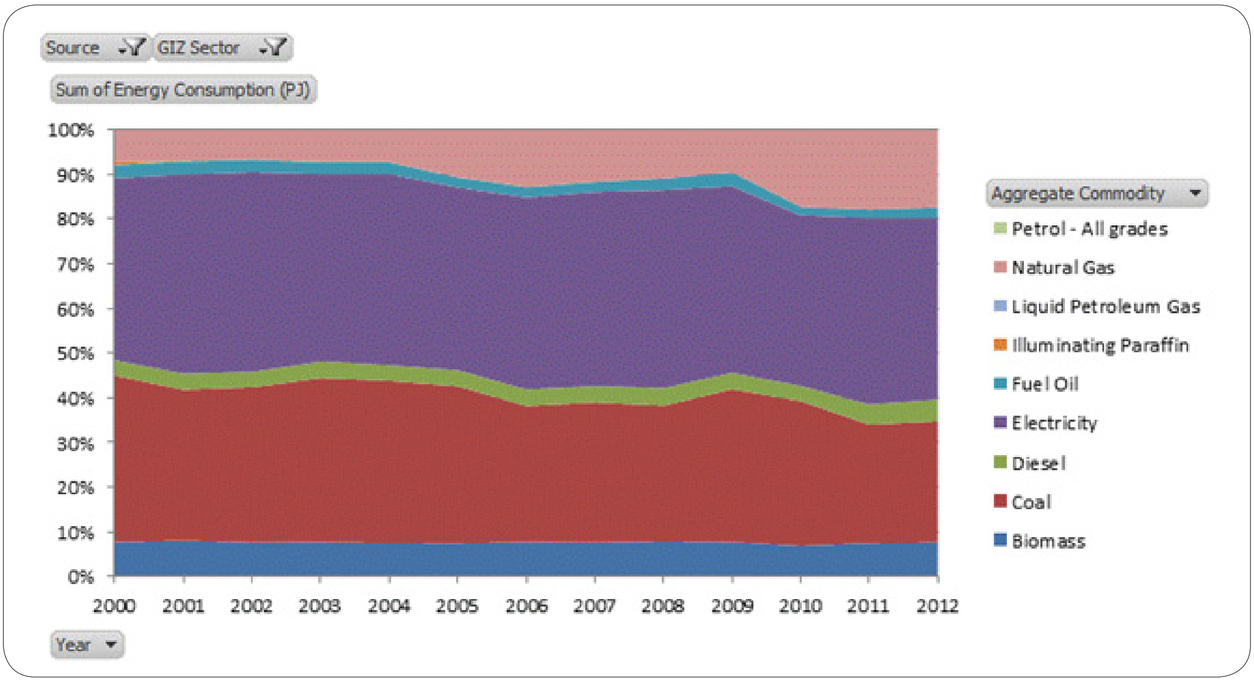


Figure 21: Energy carrier shares in the industry sector - IEA energy balances 2000–2012

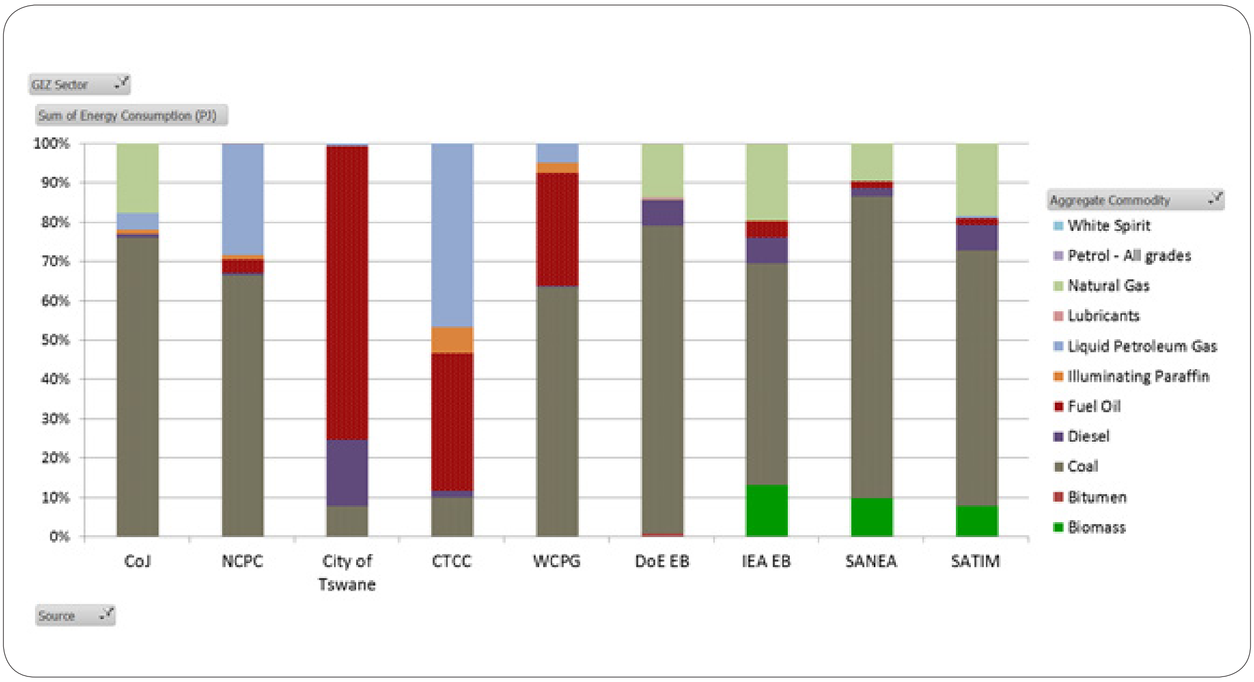


Figure 22: Energy carrier shares in the industry sector – all sources (average for all years available)

Cape Provincial Government, showed high shares of fuel oil. This supports the assessment that fuel oil use for heating and steam generation is likely to be significant in industry for practical reasons, namely that coal quality may be problematic inland and that coal prices may be high at the coast. Drilling down to sub-sector level for the firm level data lumped together as one data set totalling 650 data points reveals an interestingly diverse picture as shown in Figure 23.

Again we see that fuel oil demand from sub-sectors like food, beverages and tobacco, non-ferrous metals and non-metallic minerals is not reflected in these energy balances or in the IEA sub-sector shares shown in Figure 24. Paraffin and LPG also appear to be more widely used than is reflected in the aggregate sources.

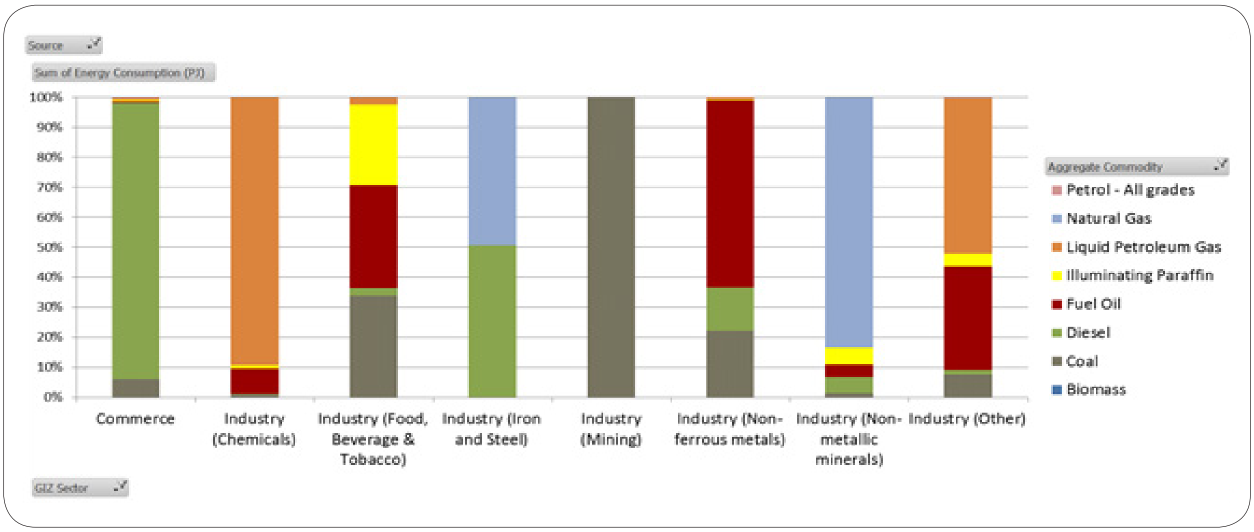


Figure 23: Energy carrier shares in industry sub-sectors – firm level sources combined

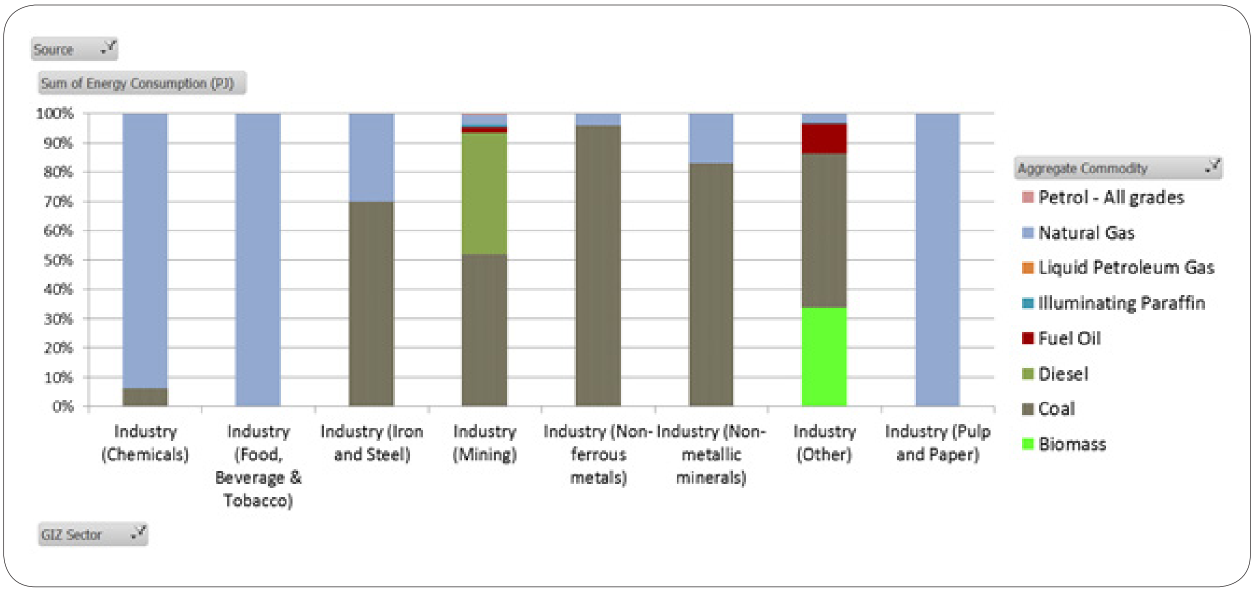


Figure 24: Energy carrier shares in industry sub-sectors – IEA energy balances averaged excluding electricity

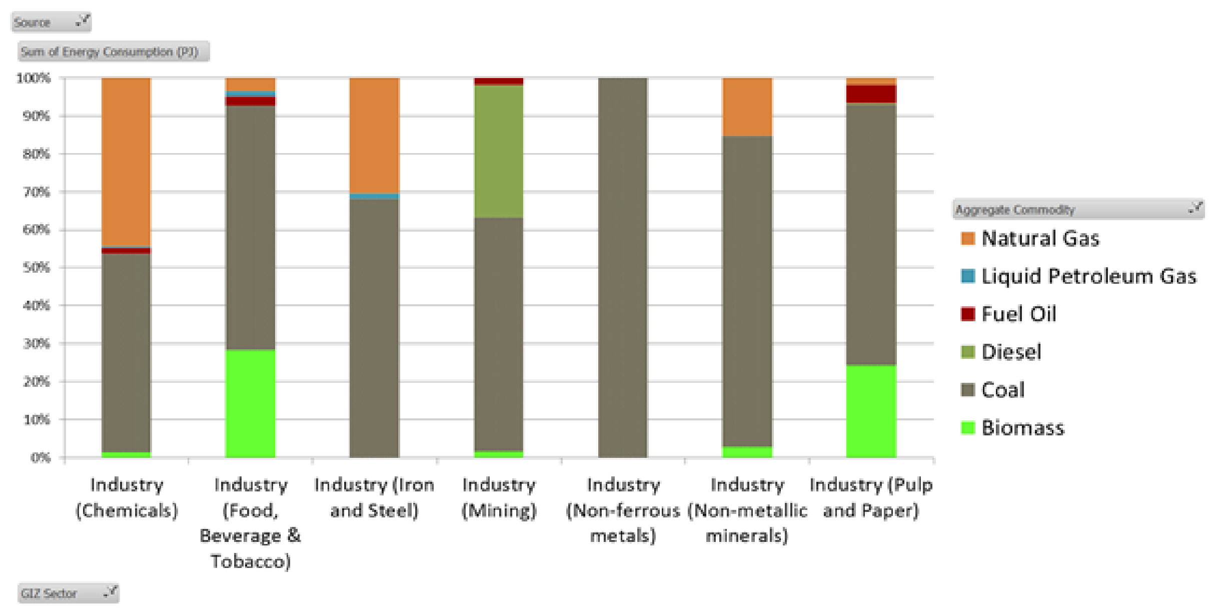


Figure 25: Energy carrier shares in industry sub-sectors – SATIM Model 2006

A further interesting comparison can be made with the SATIM model where the biomass allocation to industry is split over the food, beverages and tobacco sub-sector (combined heat and power (CHP) production from sugar-cane bagasse) and the pulp and paper sub-sector (CHP production from black liquor).

A comparison of absolute numbers across sources also yielded some insights as shown below for fuel oil.

Table 13: Comparison of fuel oil consumption in industry and electricity production for data sources

Data Source	Type	Year	Consumption (PJ)
City of Tswane[1]	Firm Level	updated 2015	22.89
CoJ[1]	Firm Level	updated 2015	0.00
CTCC[1]	Firm Level	updated 2015	5.31
WCPG[2]	Firm Level	2013	5.36
NCPC-SA[3]	Firm Level	2009–2012	0.31
ESKOM[4]	Firm Level	2015	10.65
DoE EB	National	2006	0.17
IEA EB	National	2006	21.77
SATIM	National	2006	3.90
SANEA	National	2000	13.7

[1] These are license values and not consumption values but are likely to be similar if one assumes this is based on an initial assessment of consumption
[2] Telephone survey
[3] On-site surveys of 61 firms in 7 provinces
[4] Approximately 21,000 tons/month to support coal-fired boilers (ESKOM 2015)

The two Cape Town sources agree well which, along with the Eskom consumption, at a minimum supports the notion that the sole allocation of fuel oil not used in marine bunkers to commerce in the National Energy Balance is incorrect. Interestingly, the City of Tshwane database (only 37 firms) alone equals the IEA fuel oil allocation to industry but caution should be exercised in drawing definitive conclusions from the firm level data without more validation, which was still ongoing at the time of writing. It was observed that very few data points, possibly outliers, accounted for a disproportionate share of consumption as shown in Figure 26:

Of the 650 observations in the firm level data, less than 1% accounted for 50% of the energy consumed and less than 5% accounted for 80% of the energy consumed. Interestingly, nearly 8% of the firms assessed consumed over 200 TJ/a, the threshold at which firms will have to report on energy management plans and GHG emissions under new regulations and they accounted for nearly 90% of the energy consumed (DoE 2015). The firm level data required several enquiries and issues over outliers, units and conversions remain far from resolved with the possible exception of the Cape Town and Western Province databases which are more established.

The insights gained from the reconciliation of sources was taken into account in forming a view, where required, in the econometric and imputation modelling but did not, except for a few discreet cases like industrial consumption of biomass, form a direct input to the final database. This exercise does, however, point to where further work should be done and has created a foundation for building on city-level and firm survey level data which promise the most certain route to high quality data for future greenhouse gas inventories.

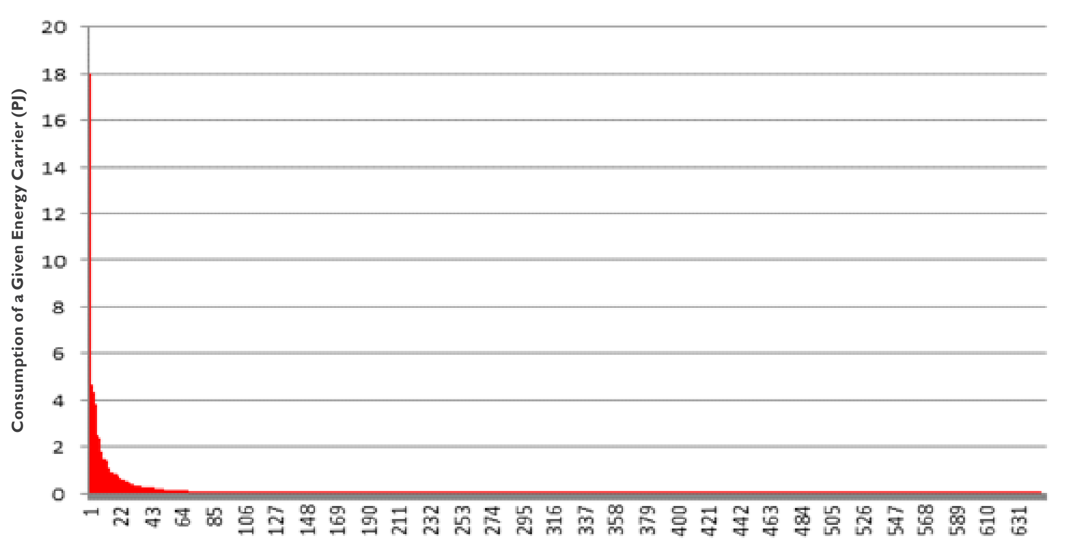


Figure 26: Distribution of plant level energy consumptions of energy carriers in city and regional datasets

6. STRENGTHS AND WEAKNESSES

Reflecting on the assignment outcomes the following are considered specific strengths of the work conducted:

- Having established a mapping framework among the respective industrial classification systems (namely ISIC, the IPCC, Stats SA's supply and use tables, SAPIA and the DoE) enables future translation of results from one system to another.

- A rather comprehensive review of the survey data from different stakeholders to establish which data sources represent the best and most consistent data for incorporation into the summary data table.

- Quite a few small but significant questions around fuel demand have been assessed by bottom-up models, for example, the domestic/international aviation kerosene split, residential wood demand, biomass in general and fuel oil use by internal/domestic navigation between ports.

- While the firm level data from cities and the NCPC-SA sample is small it is a good starting point, the contacts have been established and, furthermore, some interesting initial views emanate from it. For instance we can quickly see that the energy balance is wrong regarding zero HFO use in the food and beverages subsector.

- An imputation model that deals with noise and data gaps has been developed and can be an ongoing project in the DEA.

The following are considered weaknesses of the assignment outcome that can be addressed through future action.

- The accuracy levels of the bottom-up models can be improved through further calibration processes, refinement of energy service energy intensities and assessment of the numbers of secondary users of a fuel.

- The fuel oil consumption figures in the energy balance lack consistency (year on year) and the energy balance figures don't reflect the sample of industry data points in our possession. Some further work is required on this.

- Related to the above point, it wasn't clear if the magisterial district data from SAPIA – our only disaggregated source for fuel oil – is capturing the residual oil distributed by secondary suppliers like FFS Refiners. The aggregate figures seem a bit implausible once Eskom's consumption is accounted for. There is no metadata for supply-side liquid fuel statistics in the public domain which makes validation and interpretation difficult. Without an industry insider's understanding of supply it's difficult to know if some industrial consumption may not be in the very large allocation to marine bunkers in the energy balance, which speaks to the next point about interaction with SAPIA.

- Regrettably, a constructive dialogue with SAPIA wasn't an outcome of this study although they readily supplied data. There was, however, no feedback on the interpretation of statistics or data problems so it was, for instance, not possible, other than by assumption, to ascertain exactly what the oil companies mean by their "market category" descriptions. Given the central role of this data in the study it detracts from what could have been a fuller understanding of the numbers with more collaboration.

- Firm level data, which is required to accurately disaggregate to ISIC Revision 3.1 sub-sector level, is very sparse except for the Cape Town Air Quality Management (AQM) database which has 574 data entries, compared to 37 for the Tshwane AQM database and 31 for the City of Johannesburg AQM database. A further sample of 66 data entries was obtained from the NCPC-SA energy audit database.

The implication is that until more measurements are taken we will not understand industrial energy use statistically. The picture may however change quickly with the envisaged Energy Management Plan and GHG emission reporting regulations for sites consuming more than 200 TJ (DoE 2015).

- To overcome the above, the study used Stats SA's supply and use tables (SUTs) to establish ISIC Revision 3.1 level data, but the fuel share by subsector is essentially by assumption. While the relative monetary flows from sectors to "petroleum products" supply is a useful guide, the statistics are aggregated and going beyond this becomes an exercise in calibration and reconciliation with other information where available. Furthermore, inconsistencies are apparent in the SUTs. Expenditure on petroleum fuels use by the food product manufacturing subsector, for example, is very low compared to post and telecommunications. A recent study (WWF, 2013), however, showed that processed food is the most carbon intensive freight commodity. This may be because the food sector contracts out most of its transport but given the current view of the accounts, using the SUTs alone to create a subsector view will result in under-allocation to at least the food product manufacturing subsector. This was partially addressed in the supporting data models by imposing a view on each sector as to whether it would typically use a certain fuel or not.

7. PRESENTATION OF RESULTS

7.1 DEA official layout

At the inception of this assignment, the DEA provided an Excel table that indicated the current view on energy carrier consumption per sector (and a few sub-sectors) for the period 2000 to 2012. It was agreed that the final data generated in this assignment would be compiled in a similar tabular format for further use by the DEA. This has been achieved and forms a key part of the final report submission. A data table that summarises the assessed consumption per energy carrier per sector (at ISIC Revision 3.1 disaggregation level) is supplemented by a number of supporting models that will enable the DEA to follow the build-up of arguments that culminated in the compilation of the final data table.

7.1.1 Excel table

The historic fuel consumption data in tabular form is submitted with this report as a separate Excel-workbook. It is henceforth referred to as the "summary data table".

It is to be noted that the data reflected in the summary data table represents the best data obtained from base sources (SAPIA in particular), supplemented with alternative views (bottom-up estimations). In certain cases, for example, for international jet fuel consumption, the base source data points are dropped in favour of the estimated consumption levels. This exchange is only executed when the base source data points are judged to be of inferior quality, or if the data points have a low level of associated certainty (see Section 4.5.7.2).

Since the layout of the summary data table requires the breakdown of consumption to the detailed categorisation level required by the IPCC, the table appears rather sparsely populated. Moreover, the tabular format is not particularly useful when frequent analyses require the extraction of summaries of the data. For that purpose the summary table data has been converted into a database (also submitted with this report) which can easily be exploited in Excel pivot tables. An example of such a pivot table is used to extract the liquid fuels consumption data for use in the econometric forecasting discussed in Section 4.5.

7.1.2 Supporting models

The summary data table mentioned in Section 7.1.1 is compiled from a set of energy carrier tables that can be viewed in the "Summary Data Table" workbook submitted. However, in order to avoid clutter it was decided not to include further levels of substantiating data in the energy carrier sheets. Instead, the substantiating data has been compiled in a set of supporting model workbooks. The supporting models are arranged in terms of the energy carriers that feed into the carrier tables in the "Summary Data Table" workbook.

Each supporting model reflects the DoE energy balance numbers for the particular carrier, the aggregated SAPIA consumption data (compiled from magisterial district datasets), other estimations (typically bottom-up calculations) and the transformation matrices which provide the mapping between the ISIC Revision 3.1 categorisation, the supply and use tables, the DoE energy balances and other sectoral categorisations provided by the original data owners. In addition, the standard calorific values for the fuels are also included in each supporting model.

It is to be noted that the supporting model workbooks report the consumption of the various energy carriers in terms of native units, for example, kilolitres per annum. The "Summary Data Table" workbook, however, reflects the energy consumption per sector (sub-sector) in energy units, i.e. terajoules per annum.

In order to update the "Summary Data Table" workbook mentioned in Section 7.1.1, it is required that the analyst opens both the "Summary Data Table" workbook and the supporting models on the same PC/laptop and then checks in each of the energy carrier sheets that the sourcing of the supporting data is done appropriately. This may require some manual intervention from the analyst but can be completed in a few seconds. Once the links between the "Summary Data Table" workbook and the supporting models are established, the files can each be saved on the analyst's PC for future use. However, if the database format of the "Summary Data Table" workbook needs to be refreshed, the analyst would need to apply reverse pivoting techniques to construct a new database (see www.youtube.com/watch?v=N3wWQjRWkJc for a simple explanation on how to do this). The reverse pivoting technique freezes the database values; it would thus not be dynamically linked to the values in the "Summary Data Table".

Apart from the submitted supporting models, an Excel add-in (loess.xla) is also provided in the pack. An analyst would need to load it on their PC when scrutinising the supporting models. In certain cases it was necessary to apply econometric smoothing to the data by executing the LOESS-function in Excel, which, unfortunately is not in the Excel core list of functions but needs to be loaded separately as an .xla add-in. This function applies a polynomial curve-fit to a number of neighbouring data-points in order to generate the econometric smoothed values. It is advised that the smoothing technique only be applied on the sectoral shares of consumption levels and not on the original consumption data. The econometric smoothing enables the imputation of sparse consumption tables, as well as the determination of certainty levels of observed data-points. This technique is applied in the R code mentioned in Section 4.5.7.2.

7.2 Discussion of results

Although specific observations relating to the final database are discussed in Section 5 in this report, it would be prudent to highlight a few salient points regarding the obtained results.

In general the quality of the consumption data, namely, in terms of consistency and availability at the desired sector and sub-sector levels, was found wanting. As indicated in Section 3.2 there are portions of available consumption data that are acceptable, but sector classifications change over time and the methodology applied to allocate aggregated supply data to sector levels (and sub-sectors) does not appear to be consistent. The final outcome is therefore a collage of data sections that are judged to be the "best" available. Supplemented by fragmented surveys, bottom-up models and imputation, the reconciled result is a database which is considered less uncertain than its parts. It is, however, important to realise that the capability to reduce the uncertainty associated with consumption of specific energy carriers at sector levels is limited when based on a desktop survey approach. The only way to break through the remaining uncertainty barrier is to collate specific data from targeted market surveys and to balance the obtained data against available aggregated sources.

As indicated in Section 4.3, the SAPIA aggregated and disaggregated data was judged to be better than that reported in the energy balances of the DoE, especially for the period 2008 to 2012. Although the SAPIA aggregated data does not always differ substantially from the DoE energy balances, the disaggregated data do indicate significant variances. The period 2008 to 2010 is of particular interest, as the economy at large suffered a significant recession and inflation in a rather short period of time. Such events normally reflect reasonably well in diesel consumption, albeit not to the same extent in the consumption of other energy carriers. In this instance, both the SAPIA data and DoE energy balance data was found to be inconsistent. However, the actual fuel sales data published by the DoE does seem to corroborate the expected correlation with the economic turmoil at the time. The SAPIA fuel sales data would be the preferred aggregated source.

Although the mediocre quality of data from sources that are meant to be reputable, for example the DoE, is lamentable, a number of good data sources do warrant exploitation in future studies. The consumption profiles in the commercial and residential sectors in particular, can be supplemented from energy studies in the large metropoles and focused studies conducted by the National Cleaner Production Centre of South Africa (NCPC-SA).

The disaggregated Sasol gas data portraying the landscape of natural gas (and equivalents) consumption deserves praise as being the most consistent source, albeit only available for the period 2005 to 2012. It does provide the best indication of gas consumption at the ISIC Revision 3.1 sector levels.

The aggregated flight data provided by ACSA proved to be extremely useful in determining the ratio of international versus domestic aviation fuel consumption. It formed a crucial part in the bottom-up assessment that resulted in a significantly different ratio to that used in the recent past. The aggregated flight data does not reveal any specific information regarding airline activities but does provide sufficient information on aircraft categories for which publicly available fuel efficiencies can be obtained. In order to exploit this source in future the DEA would be advised to contact the Institute for Transport and Logistic Studies (Africa) at the University of Johannesburg. In the event that such data is not available, the DEA could use the estimated factors in Table A4.2 for the period 2011 to 2012 – the ratio of international versus domestic consumption appears to have reached a plateau at 75:25 and is likely to remain there for the foreseeable future.

The econometric forecasts of energy carrier consumption, addressed in Section 4.5, provide some insight into the future for estimating the GHG inventory for South Africa. However, it should be borne in mind that the forecasts are only as good as the data used in estimating the models. The methodology provided in Section 4.5 can be applied to future updates to the database. It should, however, be noted that the forecasts also depend on the specification of the econometric models employed. A specific constraint encountered in this assignment was the fact that the available data was not sufficient at disaggregated levels to enable forecasts per energy carrier per sector. The data had to be consolidated on a national level, but allowed for the split per energy carrier. The level of detail in the forecasts is therefore not ideal, but can be used in imputation models to estimate future consumption per sector. The accuracy of such estimates will improve in future as more data becomes available at the desired (sub-)sector levels.

8. FUTURE WORK PROGRAMME

8.1 Current developments

On 5 June 2015 – in Government Gazette 38857, Notice Number 541 – the Minister of Environmental Affairs published for public comment the draft National Greenhouse Gas (GHG) Emission Reporting Regulations under the National Environmental Management: Air Quality Act 39 of 2004 (NEM: AQA). The purpose of the draft regulations is to introduce a single national reporting system for the reporting of GHG emissions which will be used:

- to inform policy formulations;
- for the Republic of South Africa to meet its obligations under the United Nations Framework Convention on Climate Change and any other international treaties to which it is bound; and
- to establish and maintain a National Greenhouse Gas inventory (DEA 2015: 9).

Annexure 1[6] of the draft regulations lists the activities resulting in GHG emissions which must be reported to the competent authority[7] and includes specified activities in the energy, fuel supply, industrial processes, agricultural, forestry and other land use and waste sectors.

In the event that the draft regulations are approved, any person conducting an activity listed in Annexure 1 will be obliged to register on the National Atmospheric Emission Inventory System (NAEIS) within 30 days after the commencement of these regulations or the commencement of a listed activity. In terms of regulation 7(1) Category A data providers must submit the total greenhouse gas emissions arising from each of the activities as set out in Annexure 1 to these regulations for the preceding calendar year, to the NAEIS by 31 March of each year. The reporting of combustion emissions apply to individual combustion installations with a 10 Megawatt energy capacity or more. The draft regulations require proper record keeping of emissions data, the verification of information collected and supplied, and on-site verification of emissions by a competent authority once every two years. The draft regulations also make provision for significant fines and prison sentences in the event that a person is convicted of an offence in terms of the regulation (DEA 2015).

In terms of the National Energy Act, 2008 (Act 34 of 2008) the Department of Energy is mandated to collect and analyse energy and fuel use data, as well as information on energy efficiency improvements. To this end, on 27 March 2015 the Department of Energy released draft regulations entitled "Draft regulations regarding Registration, Reporting on Energy Management and Submission of Energy Management Plans" (DoE 2015). The purpose of the regulations is to ensure routine reporting of energy consumption for companies consuming above 180 TJ per annum and energy management plans for companies

---

6 The list of activities for which GHG emissions must be reported to a competent authority include:
- Energy: – all fuel/petroleum and electricity-related activities, manufacturing and construction sectors, the fugitive emission from fuels, spontaneous combustion and burning coal dumps, oil and natural gas and any other emissions from energy industries)
- Industrial processes and production use – mineral and chemical production, the metal industry, non-energy products from fuels and solvents use, electronics industry, product uses as substitutes of ozone depleting substances and refrigeration and air conditioning
- Agriculture, forestry and other land use
- Waste sector – solid waste disposal, biological treatment of solid waste, incineration and open burning of waste and waste water treatment and discharge
- Transport – civil aviation, cars, railways, water-borne navigation and other transport means including pipelines and off-road (DEA 2015: 14–18)

7 The National Inventory Unit based at the National Department of Environmental Affairs

consuming above 400 TJ per annum in order to support monitoring of energy efficiency targets of the National Energy Efficiency Strategy post 2015. The Energy Efficiency Target Monitoring System will be used to collect, analyse and report on the data from companies which meet the thresholds outlined in the draft regulations for energy reporting. Currently tools for collecting data from energy intensive industries such as mining and metal production have been piloted.

Assuming that the above two sets of regulations will be promulgated into law, it may significantly enhance future sourcing of data to enable the Republic of South Africa to meet its obligations under the UNFCCC and any other international treaties to which it is bound.

8.2 Recommendations to facilitate future data quality and value

The current reality is that research and academic institutions as well as other stakeholders in the energy and environmental management sectors rather opt to use existing survey data than to conduct field surveys to expand the existing knowledge base on the sectoral distribution of energy consumption. As a consequence the quality of data is relatively low in South Africa, with quality data limited to a sub-set of fuel carriers and industry sub-sectors. Some recommendations on how to address data quality going forward are provided:

- In the event that the draft regulations referred to in Section 8.1 come into effect, but don't yield the anticipated outcomes, a repeat of the 1975 Bennett Energy Use in Industry Survey can be considered. The survey was characterised by the sourcing of energy data through a single team – working systematically through a number of surveys and in the process developing a very specific expertise / competency in this regard. A further notable characteristic of the Bennett Study that contributed to its success was the rather informal approach adopted, which facilitated access to detailed operational data that would typically be stifled by overly formal approaches – involving frequent official communication from senior government officials (Bennett 1975). Some of the benefits resulting from this approach would be:
  - Provides an independent view on the data;
  - Provides access to the data of companies with energy output levels below the 180 TJ per annum threshold proposed in the draft Department of Energy regulation.
  - Facilitates access to a more granular level of data, which may enable the establishment of an expert system that would facilitate future on-site verification of emissions by the competent authority as set out in the draft (GHG) emission reporting regulations.

It is suggested that the starting point for conducting these industry surveys would be the industry research reports published by the company Who Owns Whom. These reflect company profiles of over four thousand South African companies, categorised into approximately 300 industry sub-sectors, aligned with the five digit standard industrial classification code (SIC Code) level. Each report incorporates the contact details of all the significant players in the specific industry sub-sector – representing eighty per cent of the specific sub-sector's production output.

- It is suggested that consideration be given to the establishment of a single independent data custodian for energy and environmental data. Potential candidates are the CSIR, South African National Energy Development Institute (SANEDI) or Stats SA. It is anticipated that the following benefits will result from this approach:
  - Consistent use of the same data across government departments;
  - Facilitate improved inter-departmental flow of data;
  - Less biased relationship with industry.
- A focused initiative to systematically improve cooperation between stakeholders will go a long way to address current household and industry data priorities of national strategic importance. Relatively small adjustments to current Stats SA household survey questionnaires can provide useful quantitative substantiation of household energy utilisation, which is currently lacking.
- It was pointed out in Section 4.3 of this report that there is a huge amount of disparity in the way that different data sources categorise data. Consideration should be given to convert all government department reporting systems to conform to a single data classification system, such as ISIC Revision 3.1, or an aggregation thereof. This will facilitate future data validation and interchangeability and fast-track the journey towards quality data reporting.
- Once the draft regulations referred to in Section 8.1 come into effect, huge volumes of data will be submitted to the DoE and DEA respectively. It is recommended that clear strategies and procedures be developed to address the following aspects:
  - What will the data be used for;
  - How will the data be managed, validated and secured;
  - How to manage data access, namely what level of access will be given to different stakeholder groups, for example, the general public, trade and industry associations, local, provincial and national government departments.
  - Decide on a data repository structure – informed by the data access management philosophy.

Compliance with the regulations will place a burden on industry. Without well thought through data management strategies the potential value that will be unlocked from the sourced data will be compromised.

8.3 Future work considerations

The resources allocated to this assignment facilitated a reduction in the level of uncertainty of the overall data landscape as depicted in Figure 9. Should the need arise to further reduce data uncertainty the following suggestions should be considered:

- As far as the sourcing of fuel consumption data is concerned it was found that engaging with major fuel producers yielded a better return on effort compared to fuel consumer entities. It seems that there is a sufficient marketing and/or strategic planning objective for fuel producers to track fuel sales rather meticulously at a granular level. The effort to replicate these statistics, following a bottom-up approach of engaging with consumer organisations is at least one order of magnitude higher.
- The fuel oil consumption figures in the energy balance were found to be wanting. They lack

consistency (year on year) and the energy balance fuel oil consumption figures don't reflect the sample of industry data points in our possession. It is recommended that further work be done to validate the fuel oil consumption figures. According to the Pressure Equipment Regulations of the South African Occupational Health and Safety Act, 1993 (Act No. 85 of 1993) (DoL 2009), no user may use a steam generator unless such user is in possession of a certificate of registration for the said equipment. Consequently the Department of Labour has a database reflecting all boilers installed in South Africa – including the capacity and energy source for each boiler. Unfortunately the database comprises a decentralised paper-based system and given the limited resources allocated to this assignment it was not possible to make use of this comprehensive data source in this assignment. Consideration should be given to launch a project, together with the Department of Labour, to migrate the paper-based system to an electronic database – perhaps starting with the Gauteng provincial office as a first phase.

- There is quite a range of niche mid- and heavy-distillate products that were necessarily aggregated for this study. For instance all the fuel oil grades of different qualities (sulphur level) and densities including marine bunker fuels, heavy fuel oil (HFO) and light fuel oil (LFO) were aggregated although they can differ markedly in price and application. Similarly the various grades of diesel for industrial and marine applications were aggregated with automotive diesel. In the scale of the larger energy system the volumes of these fuels consumed, with the exception of marine bunker fuel, are generally dwarfed by petrol and automotive diesel and thus these distinctions have little effect on the aggregate picture. They are however important for understanding the liquid fuels landscape and the potential for substitution, and future work should attempt a disaggregation of these elements and an understanding of the price differences and supply chain. The investigations undertaken for this study suggest that some of the residual oil allocation to marine bunkers in the national energy balance is actually used in a range of industrial activities, including the 10–12 PJ of fuel oil used to support coal combustion in power stations.

- Romulo N. Ely, a Brazilian economist specialising in working up input / output and general equilibrium models will be doing his post-doctoral work at the ERC from November 2015 onwards. Consideration should be given to proposing a project to develop a calibrated industrial sub-sector liquid fuels disaggregation model – based on the Stats SA supply and use tables and other sources.

- Limited success was achieved in finding quality alternative data sources for the commercial sector to validate the figures in the energy balance allocated to this sector. It is furthermore anticipated that the commercial sector will be impacted the most by the current electricity load curtailment programme. It is therefore suggested that further work be done to validate the fuel consumption figures in the commercial sector. The approach followed by Caroline Martin (2013) could be considered, but incorporating some field survey work for restaurants, schools and hospitals.

9. REFERENCES

9.1 List of articles

The articles referenced in this assignment are reflected below. An Excel-based version of the list of articles, incorporating some additional metadata fields, is referenced in Annexure 2 of the report.

Bennett K. (1975). Energy Utilisation in South Africa. PhD Thesis, Energy Research Institute, University of Cape Town.

Bouwer G. 2002. The role of supply and use tables in South Africa as a tool for economic analysis. In The Fourteenth International Conference on Input-Output Techniques from 10th – 15th October 2002, Montreal, Canada. Montreal, p. 33.

British American Tobacco Plc, 2010. *Sustainability report 2010, Balanced scorecard for South Africa*, viewed 26 August 2014, from http://www.bat.com/group/sites/UK__9D9KCY.nsf/vwPagesWebLive/DO9DCL3P?opendocument.

British American Tobacco Plc, 2011a. *Sustainability report 2010*, viewed 26 August 2014, from http://www.bat.com/group/sites/UK__9D9KCY.nsf/vwPagesWebLive/DO9DCL3P?opendocument.

British American Tobacco Plc, 2011b. *Sustainability report 2011*, viewed 26 August 2014, from http://www.bat.com/group/sites/UK__9D9KCY.nsf/vwPagesWebLive/DO9DCL3P?opendocument.

British American Tobacco Plc, 2011c. *Sustainability report 2011*, viewed 26 August 2014, from http://www.bat.com/groupfs/sites/BAT_8NXDKN.nsf/vwPagesWebLive/DO8FAG29/$file/Full_sustainability_report_2011.pdf.

British American Tobacco Plc, 2012. *Sustainability: Why it matters. Sustainability summary 2012*, viewed 26 August 2014, from http://www.bat.com/group/sites/UK__9D9KCY.nsf/vwPagesWebLive/DO9DCL3P?opendocument.

British American Tobacco Plc, 2013. *Value shared. Sustainability summary 2013*, viewed 26 August 2014, from http://www.bat.com/group/sites/UK__9D9KCY.nsf/vwPagesWebLive/DO9DCL3P?opendocument.

Carbon Protocol of South Africa, n.d. Progress on the South African Carbon Tax, viewed 26 August 2014, from http://carbonprotocol.org/progress-on-the-south-african-carbon-tax/.

Carec, n.d., *Annual report 2010*, viewed 26 August 2014, from http://www.carec.org/.

Central Statistical Service, 1993. *Standard Industrial Classification of all Economic Activities (SIC), Fifth Edition*, Pretoria, Report No. 09-90-02.

Cermalab, 2014. *Survey of the status of the informal clay brick making sector. Eastern Cape, South Africa, 2014, viewed 26 August 2014*, from http://www.swisscontact.org/fileadmin/documents/Projekte/Informal_clay_brick_making_sector.pdf.

City of Cape Town, 2011. *State of energy and energy futures report*, Cape Town, viewed 26 August 2014, from https://www.capetown.gov.za/en/EnvironmentalResourceManagement/publications/Documents/State_of_Energy_+_Energy_Futures_Report_2011_revised_2012-01.pdf.

City of Cape Town, 2015. City of Cape Town AQM Data: Spreadsheet of inventory of fuel consumption of combustion sources in the City of Cape Town as of 23 June 2015; kindly forwarded by Ed Filby in personal communication – September 2015.

City of Johannesburg, 2015. City of Johannesburg AQM Data: Spreadsheet of licensed fuel consumption limits for large combustion sources in the city of Johannesburg; kindly forwarded by S. Davis (ICLEI) in personal communication and supplied to him by G Mbara, Department of Environment and Infrastructure Services, City of Johannesburg.

City of Tshwane, 2015. City of Tshwane AQM Data: Spreadsheet of licensed fuel consumption limits for large combustion sources in the city of Tshwane; kindly forwarded to the Department of Environmental Affairs in personal communication by S Mzondi, Environmental Management Services Department, City of Tshwane.

Comair Limited, n.d. *Comair Limited integrated annual report 2013*, viewed 26 August 2014, from http://www.comair.co.za/Media/Comair/files/2013/Comair-Limited-Intergrated-Annual-Report-2013.pdf.

Competition Commission, 2012. Competition Commission press release, 24 October 2012.http://www.compcom.co.za/assets/Uploads/AttachedFiles/MyDocuments/Competition-Commission-refers-a-case-of-collusion-against-oil-companies.pdf

Council for Scientific and Industrial Research (CSIR), 2013. *10th annual state of logistics survey for South Africa: 2013* N. Viljoen, ed., Pretoria: Council for Scientific and Industrial Research (CSIR), viewed 26 August 2014, from http://www.kpmg.com/ZA/en/IssuesAndInsights/ArticlesPublications/Management-Consulting/Documents/10th_SoL_Bold_Steps_Forward_web.pdf.

Cowpertwait P.S.P. and Metcalfe A.V. 2009. *Introductory Time Series with R*, Berlin: Springer Science and Business Media.

DEA, 2014. *Mitigation report: South Africa's Greenhouse Gas (GHG) Mitigation Potential Analysis*, Pretoria, viewed 20 August 2015, from https://www.environment.gov.za/sites/default/files/docs/mitigationreport.pdf.

DEA, 2015. *National Environmental Management: Air Quality Act, 2004 (Act No. 39 of 2004): Draft National Greenhouse Gas Emission Reporting Regulations*, Pretoria, viewed on 20 September 2015, from http://www.gpwonline.co.za/Gazettes/Gazettes/38857_5-6_EnvirAff.pdf.

Department of Energy, 2010. Digest of South African energy statistics 2009 4th ed., Directorate: Energy Information Management Process Design and Publication, Pretoria: Department of Energy.

Department of Energy, 2011. *South African energy synopsis 2010*, Pretoria: Directorate: Energy Information Management, Department of Energy. Available at http://www.energy.gov.za/files/media/explained/2010/South_African_Energy_Synopsis_2010.pdf

Department of Energy, 2012. *A survey of energy-related behaviour and perceptions in South Africa: The residential sector*, Pretoria, viewed 26 August 2014, from http://www.energy.gov.za/files/media/Pub/Survey of Energy related behaviour and perception in SA - Residential Sector - 2012.pdf.

Department of Energy, 2013 (a). *A survey of energy-related behaviour and perceptions in South Africa: The residential sector*, Pretoria, viewed 26 August 2014, from http://www.energy.gov.za/files/media/Pub/DoE-2013-Survey-of-EnergyRelated-Behaviour-and-Perception-in-SA.pdf; data for household expenditure of fuels in excel spreadsheet forwarded by personal communication with M. Molomo on 29 September 2015

Department of Energy, 2013 (b). Draft 2012 Integrated Energy Planning Report. , p.160, viewed 26 August 2014, from http://www.energy.gov.za/files/IEP/IEP_Publications/Draft-2012-Integrated-Energy-Plan.pdf.

Department of Energy, 2013 (c). Transformation of the gas sector. Presentation by Chief Director: Hydrocarbons Policy dated 13 August 2013.

Department of Energy, 2015 (a). *National Energy Act, 2008 (Act No. 34 of 2008): Published for public comment: Draft Regulations regarding Registration, Reporting on Energy Management and Submission of Energy Management Plans*, Pretoria, viewed on 20 September 2015, from http://www.gpwonline.co.za/Gazettes/Gazettes/38614_27-3_Energy.pdf.

Department of Energy, 2015(b). Derivation of Biomass for Residential Sector Use with Calculations, Spreadsheet Model forwarded in personal communication from R Maserumule on 1 October 2015.

Department of Energy, n.d. (a). Energy Balances: 2000–2012, viewed 26 August 2015, from http://www.energy.gov.za/files/energyStats_frame.html

Department of Energy, n.d. (b). Liquid fuels annual aggregated national production and consumption: 2005–2012, viewed 26 August 2015, from http://www.energy.gov.za/files/energyStats_frame.html.

Department of Energy, n.d. (c). Liquid Fuels Sales for Petrol, Diesel, LPG and Paraffin by Magisterial District and Trade Category: 2009–2013, Spreadsheet provided in personal communication from R Nembahe on 25th March 2015.

Department of Environmental Affairs (DEA), 2013. *GHG Inventory for South Africa 2000-2010*, viewed 26 August 2014, from https://www.environment.gov.za/sites/default/files/docs/greenhousegas_invetorysouthafrica.pdf.

Department of Labour (DoL), 2009, Occupational Health and Safety Act (85/1993): *Pressure equipment regulations*, Pretoria, viewed on 25 October 2015, from http://us-cdn.creamermedia.co.za/assets/articles/attachments/22439_reg_734.pdf.

Department of Mineral Resources: Directorate: Mineral Economics, 2014. *South Africa's mineral industry 2012 - 2013*. SAMI, Pretoria.

Department of Minerals and Energy, 2002. *South African national energy database energy prices*, Pretoria, viewed 26 August 2014, from http://www.energy.gov.za/files/energyStats_frame.html.

Department of Minerals and Energy, 2005 (b). *Digest of South African energy statistics*, Pretoria: Directorate: Energy Planning and Development, Department of Minerals and Energy, viewed 26 August 2014, from http://www.energy.gov.za/files/media/explained/2005%20Digest%20PDF%20version.pdf.

Department of Minerals and Energy, 2006. *Digest of South African energy statistics 2006*, Directorate: Energy Planning and Development, Pretoria: Department of Minerals and Energy. Available at http://www.energy.gov.za/files/policies/energyPlanning/Draft-2012-Integrated-Energy-Planning-Report-an-Executive-Summary.pdf

Department of Minerals and Energy, Eskom, Energy Research Centre University of Cape Town, 2002. *Energy outlook for South Africa: 2002*, Pretoria, viewed 26 August 2014, from http://www.gov.za/sites/www.gov.za/files/DME_energy_outlook_23012004_0.pdf.

Department of Minerals: Directorate: Mineral Economics, 2005. *South Africa's mineral production and sales 1985-2004*. Report R50/2005, Pretoria, viewed 26 August 2014, from http://s3.amazonaws.com/zanran_storage/www.dme.gov.za/ContentPages/18713960.pdf.

Department of Water Affairs and Forestry, 2003. *Roundwood Supply and Demand to 2030 A sub-Sector Analysis*, August 2004. Available athttp://www.daff.gov.za/doaDev/sideMenu/ForestryWeb/webapp/files/Roundwoodscr.pdf

Enders W. 2004. *Applied Econometric Time Series*, 2nd ed., Wiley.

Energy Research Centre, 2013. *Assumptions and Methodologies in the South African TIMES (SATIM) Energy Model. Version 2.1*, University of Cape Town. Cape Town, viewed 26 August 2014, from http://www.erc.uct.ac.za/Research/Otherdocs/Satim/SATIM%20Methodology-v2.1.pdf.

Energy Research Centre, 2015. SATIM Results for 2006 and 2010 by Industrial Sector and Fuel, Spreadsheet forwarded in personal communication by B Merven on 24 August 2015

Engen, 2015. *Product Handbook*. [Webpage] viewed 07 December, 2015 from http://www.engen.co.za/home/apps/content/products_services/fuels/product_handbook.aspx

ESKOM 2014. Supplementary and divisional report: 10-year statistical overview. Eskom Holdings SOC Limited. [Web page]. Viewed 07 December 2015 from, http://integratedreport.eskom.co.za/supplementary/app-overview.php

ESKOM 2015, Estimated Monthly Fuel Oil Consumption to Support Coal-fired Power Generation. Personal Communication from L Barker via email, 28th September 2015.

Falcon R 2015. Unpublished Notes from Oral Presentation at Round-table discussion: Water and Coal-Based Power Generation, Water Research Commission, Pretoria, Friday, 28 August 2015.

Forestry South Africa, 2015. Oct 2015 - South African Forestry & Forest Products Industry Facts - 1980 - 2013 (excel version).

Gujarati D.N. 2003, *Basic Econometrics*, 4th ed., McGraw-Hill.

Haw M. and Hughes A. 2007. *Clean Energy and development for South Africa: Background data*, Cape Town, viewed 26 August 2014, from http://www.erc.uct.ac.za/Research/publications/07Haw-Hughes Clean energy & development - 1.pdf.

Heiberger M.H. and Neuwirth E. 2009. *R Through Excel*, Springer.

Hughes A. 2015. Personal Communication – Bottom-up Model of Residential Wood-Use for 1996, 2006 and 2011, Excel spreadsheet, Energy Research Centre, University of Cape Town.

Institute for Statistics and Mathematics n.d. The Comprehensive R Archive Network, Institute for Statistics and Mathematics, WU (Wirtschaftsuniversität Wien), viewed 23 October 2015, from http://CRAN.R-project.org.

Intergovernmental Panel on Climate Change (IPCC), 2010. *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. Viewed 07 December from https://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf

International Energy Agency (IEA), 2011. *Energy statistics of non-OECD countries*, 2011 Edition, Paris, available from: http://www.oecd-ilibrary.org/energy/energy-statistics-of-non-oecd-countries-2011_9789264104013-en.

International Energy Agency (IEA), 2012. Extended energy balances 2000-2012. Spreadsheet format energy balances acquired by authors for previous projects, Available for purchase at: http://www.iea.org/statistics/relateddatabases/worldenergystatisticsandbalances/

IPCC, 1996. *Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories: Reference manual (Volume 3)*. Viewed 05 December 2015, from http://www.ipcc-nggip.iges.or.jp/public/gl/invs6.html

IPCC, 2000. *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. Viewed 21 August 2015, from http://www.ipcc-nggip.iges.or.jp/public/gp/english/gpgaum_en.html.

IPCC, 2006 (a). *2006 IPCC Guidelines for National Greenhouse Gas Inventories:* Volume 1 General guidance and reporting: Chapter 8 Reporting guidance and tables. Viewed 26 August 2014, from http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol1.html.

IPCC, 2006 (b). *2006 IPCC Guidelines for National Greenhouse Gas Inventories:* Volume 1 General guidance and reporting: chapter 3 Uncertainties, p.3.6. Viewed 26 August 2014, from http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol1.html.

IPCC, 2013. Fifth Assessment Report - Climate Change 2013, viewed 26 August 2014, from https://www.ipcc.ch/report/ar5/

Kabacoff R. I. 2014.Quick-R: accessing the power of R. Web page viewed 23 October 2015 from http://www.statmethods.net/

Kleiber C. and Zeileis A, 2008. *Applied Econometrics with R*, Springer.

Lozynsky Y. Neelis M. Blinde P. (Ecofys); Lewis Y. Cohen B. Van der Merwe A.B. and Patel I. (The Green House) 2014. *Emissions intensity benchmarks for the South African Carbon Tax: Technical support study*, Utrecht: Ecofys. Available at http://www.treasury.gov.za/publications/other/GHG_Emissions_Intensity_Benchmarks_for_SA_Carbon_Tax.pdf

Martin C. 2011. Establishing energy benchmarks for commercial buildings in the City of Cape Town. Masters thesis. University of Cape Town. Available from http://uctscholar.uct.ac.za/PDF/55776_Martin,%20C.pdf

Martin C. 2013. Generating low-cost national energy benchmarks: A case study in commercial buildings in Cape Town, South Africa. *Energy and Buildings*, 64, 26–31. Available from http://www.erc.uct.ac.za/Research/publications/13Martin-lowcost_national_energy_benchmarks.pdf

Merven B. Moyo A. Stone A. Dane A. and Winkler H. 2014. *Socio-economic implications of mitigation in the power sector including carbon taxes in South Africa*. Working paper for CDKN project on linking sectoral and economy-wide models. Energy Research Centre, University of Cape Town, Cape Town. Available at http://www.erc.uct.ac.za/Research/publications/14-Merven-etal-Socioeconomic_implications.pdf

Merven B. Stone A. Hughes A. and Cohen B. 2012. *Quantifying the energy needs of the transport sector for South Africa: a bottom-up model*. Energy Research Centre, University of Cape Town, Cape Town: SANEDI. Available at http://www.erc.uct.ac.za/Research/publications/12-Merven-etal_Quantifying_energy_needs_transport%20sector.pdf

Molefe B. Transnet Group Chief Executive comments on the draft National Greenhouse Gas Inventory: 2000 - 2010. Formal correspondence addressed to DEA. 2014.

Mondi, 2004. *Mondi in South Africa: Stakeholder report 2004*, viewed 26 August 2014, from http://www.mondigroup.com/uploads/mondi_sa_stakeholder_report2004_671.pdf.

Napp T. 2013. Investigating the potential for gas switching in manufacturing. Presentation by Dr Tamaryn Napp at Stakeholder Consultation Workshop: Planning for an Emerging Natural Gas Economy for South Africa, 17 October 2013.

National Cleaner Production Centre (NCPC), 2015. Database of Plant Level Energy Audits by Fuel, Sector and Province, Spreadsheet forwarded by personal communication from S Ramdharee on 14 October 2015.

National Department of Transport, 2011 (a). Part 1: Growth of a South African Maritime Transport Industry. In *SA Maritime Transport Sector Study*. Pretoria, National Department of Transport, p. 50. Available from http://www.transport.gov.za/Portals/0/Maritime/Part%201_Final_270711yfx.pdf

National Department of Transport, 2011 (b). Part 2: Coastal and Short-Sea Shipping. In *SA Maritime Transport Sector Study*. Pretoria, National Department of Transport, p. 34. Available from http://www.transport.gov.za/Portals/0/Maritime/Part2_Final_280711yfx.pdf

National Department of Transport, 2011 (c). Part 3: Transhipments. In *SA Maritime Transport Sector Study*. Pretoria, National Department of Transport, p. 38. Available from http://www.transport.gov.za/Portals/0/Maritime/Part%203_Final_280711yf.pdf

Palmer Development Consulting (PDC) and Science Consultancy Enterprises (SCE) 2003. Review of the effectiveness of energy subsidies and related taxation policies in South Africa. Report completed for National Treasury and Department of Minerals and Energy. Pretoria: National Treasury and Department of Minerals and Energy. Viewed 05 December 2015 from http://www.pdc1.co.za/energy.html#_Toc144022368

Price Waterhouse Coopers, 2012. The gas equation: An analysis of the natural gas industry in South Africa. Cape Town, PWC, p.40. Available at http://www.pwc.co.za/en/assets/pdf/the-gas-equation-june-2012.pdf

Republic of South Africa (RSA), 1994. *Maritime Zones Act, 1994 (Act No. 15 of 1994)*, Pretoria, Government Printer, viewed 20 August 2015, from http://www.plato.org.za/pdf/legislation/Maritime%20Zones%20Act%2015%20of%201994.pdf

Roberts E. and Wentzel M. 2006. Household paraffin consumption in four areas: Benoni, Galeshewe, Gugulethu and Lady Grey. *Journal of Energy in Southern Africa* 17(2):23–26. Available at http://www.erc.uct.ac.za/jesa/volume17/17-2jesa-roberts.pdf

RSA, 2011, *National Climate Change Response Whitepaper*, viewed 12 December 2015, from https://www.environment.gov.za/sites/default/files/legislations/national_climatechange_response_whitepaper.pdf

RSA, 2014, *South Africa's first biennial update report*, viewed 26 August 2015, from http://unfccc.int/resource/docs/natc/zafbur1.pdf

Rypdal K. 2000. *Background paper to the IPCC Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*, viewed 21 August 2015, from http://www.ipcc-nggip.iges.or.jp/public/gp/bgp/2_5_Aircraft.pdf

SAPIA 2000, Fuel Sales Statistics Disaggregated by Product, Magisterial District and Trade Category 1999. Spreadsheet forwarded on behalf of SAPIA by Mr Peens of the then Caltex Oil in 2000 for academic purposes (see Stone and Bennett referenced below)

SAPIA, 2008. *Petrol and diesel in South Africa and the impact on air quality*, viewed 20 August 2015, from http://www.sapia.co.za/publications/special-interest.html

SAPREF, 2014. *Sustainability report 2013: A report to our stakeholders*, viewed 20 August 2015, from http://www.sapref.co.za/Media/Default/Page/SAPREF%20FINAL%20WEB.pdf

Sasol, 2011. *Sustainable development report: Pursuing responsible growth*, viewed 20 August 2015, from http://www.sasol.co.za/sites/default/files/content/files/11431_Sasol_SD2011_1328013983285.pdf

Sasol, 2012. *Sasol sustainable development report 2012*, viewed 20 August 2015, from http://www.sasol.co.za/sites/default/files/content/files/SASOL_SDR%25202012_1353934690274.pdf

Sasol, 2015. Data for Natural Gas Sales by SASOL and Egoli Gas 2005 – 2015, Spreadsheet forwarded in personal communication by R Hiestermann on 15 September 2015

Scorgie Y. Annegarn H. and Burger L. 2004. *Fund for research into industrial development growth and equity (FRIDGE): Study to examine the potential socio-economic impact of measures to reduce air pollution from combustion – Final Report. Task 4a. Supplementary report: Quantification of Environmental Benefits Associated with Fuel Use Interventions.*

Scorgie Y. Annegarn H. and Burger L. 2004. *Fund for research Into industrial development growth and equity (FRIDGE): Study to examine the potential socio-economic impact of measures to reduce air pollution from combustion - Final Report. Part 1 Report : Definition of air pollutants associated with combustion processes.*

Scorgie Y. Annegarn H.J. and Burger L. 2004 (a). *Fund for research Into industrial development growth and equity (FRIDGE): Study to examine the potential socio-economic impact of measures to reduce air pollution from combustion - Final Report. Task 4B Report: Socio-economic impacts associated with reducing air pollution from combustion.*

Scorgie Y. Annegarn H.J. and Burger L. 2004 (b). *Fund for research Into industrial development growth and equity (FRIDGE): Study to examine the potential socio-economic impact of measures to reduce air pollution from combustion - Final Report. Task 5A&B. Report: Information and timing on supply-side measures.*

South African National Energy Association (SANEA) 2003. South African Energy Profile 2003. Melville, South Africa, No longer available from SANEA – electronic copy held by Energy Research Centre, University of Cape Town.

South African Petroleum Industry Association (SAPIA), n.d. Fuel Sales Statistics Disaggregated by Product, Magisterial District and Trade Category 2002–2008. Spreadsheets forwarded in personal communication with C van Basten on 29 June 2015.

Statistics South Africa (Stats SA), 2013. Energy 2002-2012: *In-depth analysis of the General Household Survey data*, Report No. 03-18-04, Pretoria, viewed 20 August 2015, from http://oldwww.lib.uct.ac.za/governmentpublications/govpub-news/ghs-series-v-energy-2002-2012-in-depth-analysis-of-the-general-household-survey-data/

Statistics South Africa (Stats SA), n.d. (a). Data Files in STATA Format for the General Household Surveys 2002 – 2012, downloaded from Datafirst Portal, University of Cape Town: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/central

Statistics South Africa (Stats SA), n.d. (b). Supply & use tables for South Africa 2008 – 2009, spreadsheet forwarded in personal communication by B Merven, Energy Research Centre (ERC), University of Cape Town on 5 August 2015.

Statistics South Africa, 2012, *Standard Industrial Classification of all Economic Activities, Seventh Edition*, Pretoria, Report No. 09-90-02. Available at http://www.statssa.gov.za/classifications/codelists/Web_SIC7a/SIC_7_Final_Manual_Errata.pdf

Statistics South Africa, 2013. General Household Survey 2012. Statistical Release P0318, Statistics South Africa, Pretoria. Viewed 07 December 2015 from http://www.statssa.gov.za/publications/P0318/P03182012.pdf

Statistics South Africa, 2014. Report-04-04-02 - Input-output tables for South Africa, 2010 and 2011. From: http://www.statssa.gov.za/?page_id=1854&PPN=Report-04-04-02&SCH=5769

Statistics South Africa, 2015. Report-04-04-02 - Input-output tables for South Africa, 2012. From: http://www.statssa.gov.za/?page_id=1854&PPN=Report-04-04-02

Stone A., Bennett K., 2001, *A bulk model of emissions from South African diesel commercial vehicles*, Energy Research Centre, University of Cape Town, from: http://www.erc.uct.ac.za/Research/publications-pre2004/01Stone-Bennett_Diesel_emissions.PDF

Sustainability Energy Africa, 2015. State of Energy in South African Cities, Cape Town, South Africa, LEAP models kindly made available on request.

Sustainable Energy Africa (SEA), 2015. *State of Energy in South African Cities*, Sustainable Energy Africa, Cape Town (supporting LEAP Models for Cape Town and Polokwane were also used in addition to the public report). Available at http://www.sustainable.org.za/uploads/files/file80.pdf

Tait L. Merven B. and Senatla M. 2013. *Investigation into current and future roles of paraffin in South Africa outcomes*, Cape Town, viewed 26 August 2014, from http://www.erc.uct.ac.za/Research/publications/13Tait-etal_Paraffin_in_SA.pdf.

Taviv R. et al., 2008, *Population of the LEAP system to model energy futures in South Africa*, South African National Energy Research Institute (SANERI).

Transnet, 2006. *Transnet annual report 2005*, Johannesburg, Transnet viewed 26 August 2014, from http://www.transnet.net/InvestorRelations/TransnetRep.aspx.

Transnet, 2008. *Transnet Annual Report 2007: Strategic Intent*, Johannesburg, Transnet, viewed 26 August 2014, from http://www.transnet.net/InvestorRelations/TransnetRep.aspx.

Transnet, 2011. *Transnet Limited annual report 2010, Johannesburg, Transnet*. Available at http://www.transnet.net/InvestorRelations/TransnetRep.aspx

Transnet, 2012. Presentation to the Portfolio Committee on Economic Development: Transnet infrastructure spending plan – R300bn over 7 years. , p.29, viewed 26 August 2014, from http://pmg-assets.s3-website-eu-west-1.amazonaws.com/docs/120508transnet_1.pdf.

Transnet, 2013. *Transnet sustainability report 2012*, viewed 26 August 2014, from http://www.transnet.net/InvestorRelations/TransnetRep.aspx.

Transnet, 2014. *Transnet integrated report 2013*, Johannesburg, Transnet. Available at http://www.transnet.net/InvestorRelations/TransnetRep.aspx

Transnet, 2014a. *Freight demand forecast 02*, viewed 26 August 2014, from http://www.transnet.net/BusinessWithUs/LTPF%202012/1.LTPF%202014_Chapter%2002__Freight%20Demand_Final%20Proof_Sept%202014.pdf.

Transnet, 2014b. Transnet SOC LTD [Registration No. 1990/0009000/30] Request for information [RFI] for the supply of electricity, viewed 26 August 2014, from http://www.transnet.net/Tenders/TC/Corporate%20Tenders/RFI%20for%20Electricity%20Supply%20Specimen.pdf.

United Nations, 1992. United Nations Framework Convention on Climate Change. New York: United Nations. Viewed 07 December 2015 from https://unfccc.int/resource/docs/convkp/conveng.pdf

United Nations, 2002. International Standard Industrial Classification of All Economic Activities (ISIC), Revision 3.1, Department of Economic and Social Affairs, Statistics Division, United Nations, New York, viewed 26 August 2014, from http://unstats.un.org/unsd/publication/seriesm/seriesm_4rev3_1e.pdf.

United Nations, 2014. Handbook on Measurement, Reporting and Verification for Developing Country Parties. United Nations Framework Convention on Climate Change. Bonn: United Nations Climate Change Secretariat. Viewed on 12 December 2015 from http://unfccc.int/files/national_reports/annex_i_natcom_/application/pdf/non-annex_i_mrv_handbook.pdf

United States Department of Energy (US DOE), 1991. Assessment of Costs and Benefits of Flexible and Alternative Fuel Use in the U.S. Transportation Sector Technical Report Seven: Environmental, Health, and Safety Concerns. United States Department of Energy, Office of Policy, Planning and Analysis, Washington, DC 20585, October 1991. Available at http://www.osti.gov/scitech/servlets/purl/6360489

University of Johannesburg n.d., Data for Aircraft Movement. Made available by the Airports Company South Africa (ACSA) for research purposes and forwarded with permission for use in this study.

Visagie H.J. 2013 (a). Pre-feasibility report for the importation of natural gas into the Western Cape with specific focus on the Saldanha Bay–Cape Town corridor. Department of Economic Development and Tourism, Western Cape Government, Cape Town. Available from: https://www.westerncape.gov.za/assets/departments/economic-development-tourism/lng_pre-feasibility_study_2013.pdf

Visagie HJ. 2013 (b). Cape Town-Atlantis: Survey of Energy Consumption of Large Users Disaggregated by Fuel, Spreadsheet made available by personal communication on 25th May 2015

Winkler H. Afstad T. and Howells M. 2005. *South African energy policies for sustainable development. Phase 2. Final report*. Energy Research Centre, University of Cape Town, viewed 26 August 2014, from http://www.erc.uct.ac.za/Research/publications/05Winkler-Alfstad-Howells%20Energy%20Policies%20Phase%202.pdf.

World Wildlife Fund (WWF) South Africa (2013), *Low Carbon Frameworks: Transport , Understanding freight emissions*. WWF Brief, Cape Town, World Wildlife Fund. Available at http://awsassets.wwf.org.za/downloads/wwf_freight_emmisions_brief_1.pdf

9.2 List of contacts

Name	Organisation	Designation	Contact Number	E-mail address
Mr Mark Govender	Airports Company SA	Corporate Specialist: Airline Development	+27 11 723 1474	Mark.Govender@airports.co.za
Prof J Walters	University of Johannesburg	HOD: Department of Transport and Supply Chain Management; Director: Institute for Transport and Logistics Studies (Africa)	+27 11 559 2088	jwalters@uj.ac.za
Xolile Mabusela	Department of Energy	Director: Energy Efficiency Initiatives	+27 12 406 7697 +27 79 369 3711	Xolile.Mabusela@energy.gov.za
Carla Mould	Who Owns Whom	Chief Operations Officer	+27 11 280 0880	carlam@wow.joburg
Cornel Van Basten	SAPIA	Economic Regulation Advisor	+27 (0)78 894 4169	cornel@sapia.co.za
Anton Moldan	SAPIA	Environmental Adviser	+27 (0)82 801 6215	Anton@sapia.co.za
Ramaano Nembahe	Department of Energy	Deputy Director: Energy Data Analysis and Dissemination	+27 12 406 7540 +27 82 338 8953	ramaano.nembahe@energy.gov.za
Pranesh Ramjith	ENERWEB	Technical Lead – Energy Balance Project		Pranesh.Ramjith@enerweb.co.za
Marc Stone	ENERWEB	Project Manager – Energy Balance Project		Marc.Stone@enerweb.co.za
Justin Lacroix	IEA	Energy Balance Specialist		1justin.lacroix@hotmail.com
Roger Purdue	IEA	Energy Balance Specialist		rogerpurdue@hotmail.co
Rudi Hiestermann[8]	SASOL	Senior Manager: Policy and Regulation Southern Africa Energy - Strategy, Technology and Regulation		Rudi.Hiestermann@sasol.com
Keith Bowen	ESKOM	Senior Planner System Operator		BowenKT@eskom.co.za
Leslie Barker	ESKOM	Contact for HFO consumption		Barkerlf@eskom.co.za
Romulo Neves Ely	LIMA Institute, COPPE, Federal University of Brazil	post-doc at ERC		romulo@ppe.ufrj.br
Tara Caetano	ERC, UCT	Researcher		tara.caetano@gmail.com
Bruno Merven	ERC, UCT	Senior Researcher		brunomerven@gmail.com

8 Special reference needs to be made to the quality of data provided by Rudi Hiestermann, portraying the landscape of natural gas (and equivalents) consumption, closely aligned to the ISIC Rev 3.1 sector levels.

Name	Organisation	Designation	Contact Number	E-mail address
Franck Lecocq	CIRED	Directeur	+33 143947321	lecocq@centre-cired.fr
Machwene Molomo	DoE			Machwene.Molomo@energy.gov.za
Megan Eus-ton-Brown	Sustainable Energy Africa	Project Manager City Energy Support Unit		Megan@sustainable.org.za
Alison Hughes	ERC, UCT	Senior Researcher		alison.hughes@uct.ac.za
Lynn Woolfrey	Datafirst, UCT	Data Centre Manager	+27 21 650 5707 +27 72 391 3777	lynn.woolfrey@uct.ac.za
Sashay Ramdharee	NCPC, CSIR	Project Manager: Energy Systems Optimisation	+27 12 841 3915 +27 74 254 3488	SRamdharee@csir.co.za
Alfred Hartzenburg[9]	NCPC, CSIR	National Project Manager - Industrial Energy Efficiency Project	+27 21 658 3982 +27 82 779 2871	AHartzenburg@csir.co.za
Jan H. van den Berg	City of Tshwane, AQM	Functional Head: Environmental Information Management	+27 12 358 8708 +27 82 555 9242	Janneman@TSHWANE.GOV.ZA
Given Mbara	City of Johannesburg, Dept of Environment and Infrastructure Services	Senior Professional Officer		GivenMb@joburg.org.za
Johan Visagie	ENERGYBUSINESS	Consultant to WCPG Gas Feasibility Study		visagiehj@energybusiness.co.za
Jim Petrie	Western Cape Govern-ment, Dept of Econom-ic Development and Tourism	Director: Energy and Green Economy	+27 21 483 9422	Jim.Petrie@westerncape.gov.za
Ed Filby	City of Cape Town, Air Quality Management	Senior Professional Officer		Ed.Filby@capetown.gov.za
Jacob Malatse	Directorate Electrical & Mechanical Engineer-ing, Department of Labour	Director	082 774 0320	Jacob.Malatse@labour.gov.za
Eugene Joubert	Centre for Renewable and Sustainable Energy Studies (CRSES) Department of Me-chanical and Mecha-tronic Engineering	Research Engineer		ecjoubert@sun.ac.za

9 Special reference needs to be made to the responsiveness of the NCPC team that on very short notice mobilised a team to mine through their audit reports to provide valuable fuel consumption data on sector aggregated level as input to this assignment. Our appreciation is extended to Alf Hartzenburg, Sashay Ramdharee, Sandile Khumalo and Anna Mothapo.

ANNEXURE 1 SIC, ISIC, IPCC AND ASSIGNMENT CATEGORISATION MAPPING

Refer to electronic file: Ann 1_SIC Classification_20151030_s1.1

ANNEXURE 2 LIST OF DATA SOURCES

Refer to electronic file: Ann 2-List of Data Sources_20151117_s1.1

ANNEXURE 3 SUMMARY DATA TABLE & SUPPORTING MODELS

Refer to electronic files:

Ann 3-Summary data table_20151117_s2.0

Ann 3-Database (supplement to Summary data table)_20151117_s2.0

Ann 3-Supporting models_20151117_s2.0

Ann 3-Database of Sources_20151119_s1.0

ANNEXURE 4 NOTES ON JET FUEL CONSUMPTION IN SOUTH AFRICA

A4.1 Background and problem statement

The amount of jet fuel consumed for domestic and international transport purposes, as a ratio to total consumption, is reflected as follows in the DOE energy inventories:

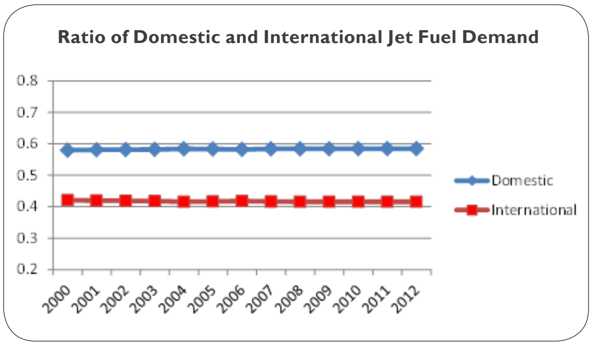


Figure A4.1: Ratio of international and domestic consumption as indicated by the DOE

The split of domestic to international consumption is evidently approximately 58:42.

The DOE has apparently always collected this data through the Fuel Sales Volume questionnaires which are submitted by all the major oil companies on a monthly basis. The totals are then simply aggregated across all the individual oil company data sets. The DOE has never had any reason to doubt the accuracy of this assessment and thus has never challenged the oil companies (SAPIA) about it.

In order to provide a meaningful account of the GHG emissions for South Africa, the DEA needs to have a view on how much aviation fuel is consumed locally and how much would be assigned to international transport.

In a paper that was published as background to the IPCC's Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories: Reference Manual (Volume 3), Rypdal (2000) mentions that the main difficulty in providing good estimates for reporting on aviation emissions lies in the distribution of fuel between domestic and international use. Rypdal (2000) further states:

> *In the current reporting IPCC Guidelines emissions from aircraft are reported as one separate item. No distinction is made between altitudes. Reporting shall include emissions from fuel used by all civil domestic passenger and freight traffic inside a country. This includes emissions from all stages of the flights (take-off, climb, cruise, descent and landing). Emissions from ground operations and stationary combustion are reported elsewhere. Reporting would become more transparent if a distinction was made between emissions in various altitudes (1000 meters).*

Emissions from aircraft bunkers are reported as a separate item (memo item). This includes emissions from fuel sold to aircraft to be used for all international aviation in the reporting country. These emissions are not to be reported as part of national totals.

It is particularly important to document the origin of the fuel use data and explain how the split between national and international aviation has been made (Rypdal 2000: 101).

The IPCC thus recommend that it would be good practice to test the outcome of a Tier 1 estimate of emissions (normally based on sales to consumers) against that obtainable from a Tier 2 assessment which is more detailed and considers the landing and take-off information associated with aircraft movement. If energy statistics and surveys do not provide accurate enough assessments of the split between international and domestic aviation fuel use, it is imperative to also consider a bottom-up assessment based on physical aircraft movement.

In this analysis an approach has been adopted whereby aircraft movement is considered for a bottom-up assessment, albeit that the data has not been obtained from individual airlines but that all the data originated from the Airports Company South Africa (ACSA). It is recognised that the three major ACSA airports process more than 85% of all passenger traffic and more than 80% of all scheduled aircraft movements. The ratio of unscheduled flights to scheduled flights amounts to about 50%. The ACSA data therefore provide a very useful proxy for the total impact in South Africa.

A4.2 Proposed methodology

The aggregate consumption of jet fuel and aviation gasoline as obtained from SAPIA and the DOE is ostensibly accurate enough as the application of jet fuel is limited to the aviation market, while the production of jet fuel is well monitored at the refineries. Exports and imports are also well recorded. In particular, the SAPIA sales figures are considered to be the accurate source for the period 2002 to 2006, while the DOE surveys are considered sufficiently accurate for the period 2009 to 2012.

The split between international and domestic consumption is the target of this investigation. If the split is assessed with sufficient accuracy, the split in consumption can be accepted as reasonable. The data recorded by ACSA on flights to and from its airports could be utilised to determine a bottom-up assessment of the split. Although the database contains information on all the ACSA airports, the three major airports, namely OR Tambo International in Johannesburg, Cape Town International and King Shaka International Airport in Durban were considered in this estimation. As stated above, these major airports process the vast majority of aircraft and passengers and can be assumed to be a representative sample of the universe of scheduled (and unscheduled) aircraft and passenger movements.

The international/domestic split requires information regarding aircraft type, fuel efficiency per aircraft type, seating capacity and flight distance, in addition to the number of recorded departures from South Africa to international destinations. Only departures would require the loading of fuel on South African territory.

A simplified view can thus be derived, based on the average flight distance ratio, average capacity per flight, flight frequencies and average fuel efficiency. These factors would apply to domestic and international flights and can be used to determine the split quite conveniently. Wide-bodied aircraft with higher capacity used on international routes would typically require more fuel for landing, take-off and cruising than the smaller equivalents used on domestic routes. The above ratios would therefore provide for a simple approach to determine the split.

A4.3 Data preparation

ACSA has a database with accurate recordings of all arrivals and departures since 2003. It contains, amongst other information, the type of aircraft, maximum seating capacity, as well as actual passengers and/or freight loading. Load factors can therefore be obtained from the data quite easily. It also records all timestamps and destinations for departures.

It is noteworthy that the ACSA data for 2003 was only recorded for a 9 month period. For the purpose of this analysis, the useful data spanned the period 2004 to 2012. This was sufficient to check for consistency in the ratios that were ultimately determined.

The aircraft types recorded in the ACSA database span a much wider variety than those quoted by the IPCC in their manual on "Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories" (2000). The seating capacity, fuel capacity and maximum range for all aircraft were therefore taken from the internet. In this regard the Wikipedia-site is quite useful. Detail specifications per aircraft can be obtained and typically resemble the example for the Boeing 777 series in Table A4.1.

Fuel efficiency in units of "litres per 100 km per seat kilometre" could be calculated from the above specifications and were compared to published data (for the selected number of aircraft types) on, amongst other sources, http://en.wikipedia.org/wiki/Fuel_economy_in_aircraft.

Table A4.1: Specifications by model (Sourced from https://en.wikipedia.org/wiki/Boeing_777 , viewed on 21 August 2015)

Model	777-200	777-200ER	777-200LR	777 Freighter	777-300	777-300ER	777-8	777-9
Cockpit crew	Two							
Seating capacity, typical	314 (3-class) 400 (2-class) 440 (maximum)			N/A (cargo)	386 (3-class) 451 (2-class) 550 (maximum)		353 (3-class)	427 (3-class)
Length	209 ft 1 in (63.7 m)				242 ft 4 in (73.9 m)		228 ft 2 in (69.5 m)	250 ft 11 in (76.5 m)
Wingspan	199 ft 11 in (60.9 m)		212 ft 7 in (64.8 m)		199 ft 11 in (60.9 m)	212 ft 7 in (64.8 m)	Unfolded: 235 ft 6 in (71.8 m) Folded: 212 ft 8 in (64.8 m)	
Wing sweepback	31.64°						TBA	
Tail height	60 ft 9 in (18.5 m)		61 ft 1 in (18.6 m)		60 ft 8 in (18.5 m)		64 ft 6 in (19.7 m)	
Cabin width	19 ft 3 in (5.87 m)						19 ft 7 in (5.97 m)	
Seat width	18.5 in (47.0 cm) in 9 abreast standard economy or 17.2 in (43.7 cm) in 10 abreast high density economy						18 in (45.7 cm) in 10 abreast standard economy	
Fuselage width	20 ft 4 in (6.20 m)							
Maximum cargo capacity	5,720 cu ft (162 m³) 32× LD3			23,051 cu ft (653 m³) 37× pallets	7,640 cu ft (216 m³) 44× LD3		40× LD3	48× LD3
Operating empty weight	297,300 lb (134,800 kg)	304,500 lb (138,100 kg)	320,000 lb (145,150 kg)	318,300 lb (144,400 kg)	353,800 lb (160,500 kg)	370,000 lb (167,800 kg)	TBA	415,000 lb (188,241 kg)
Maximum landing weight	445,000 lb (201,840 kg)	470,000 lb (213,180 kg)	492,000 lb (223,168 kg)	766,800 lb (347,800 kg)	524,000 lb (237,680 kg)	554,000 lb (251,290 kg)	TBA	557,000 lb (252,651 kg)
Maximum takeoff weight (MTOW)	545,000 lb (247,200 kg)	656,000 lb (297,550 kg)	766,000 lb (347,500 kg)	766,800 lb (347,800 kg)	660,000 lb (299,370 kg)	775,000 lb (351,500 kg)		
Typical cruise speed	Mach 0.84 (560 mph, 905 km/h, 490 knots) at a cruise altitude of 35,000 ft (11,000 m)						TBA	
Maximum speed	Mach 0.89 (590 mph, 950 km/h, 512 knots) at a cruise altitude of 35,000 ft (11,000 m)						TBA	
Maximum range (with maximum payload)	5,240 nmi (9,700 km; 6,030 mi)	7,730 nmi (14,300 km; 8,900 mi)	9,500 nmi (17,600 km; 10,900 mi)	4,970 nmi (9,200 km; 5,720 mi)	6,005 nmi (11,100 km; 6,910 mi)	7,830 nmi (14,500 km; 9,010 mi)	8,700 nmi (16,100 km; 10,000 mi)	7,600 nmi (14,100 km; 8,750 mi)
Takeoff distance at MTOW (at sea level, ISA)	8,000 ft 2,440 m	11,100 ft 3,380 m	9,200 ft 2,800 m	9,300 ft 2,830 m	10,600 ft 3,230 m	10,000 ft 3,050 m	TBA	
Maximum fuel capacity	31,000 US gal (117,348 L)	45,220 US gal (171,177 L)	47,890 US gal (181,283 L)		45,220 US gal (171,177 L)	47,890 US gal (181,283 L)	TBA	
Service ceiling	43,100 ft (13,140 m)							
Engine (×2)	PW 4077 RR 877 GE90-77B	PW 4090 RR 895 GE90-94B	GE90-110B1 GE90-115B1		PW 4098 RR 892 GE90-92B/-94B	GE90-115B1	GE9X	
Thrust (×2)	PW: 77,000 lbf (342 kN) RR: 76,000 lbf (338 kN) GE: 77,000 lbf (342 kN)	PW: 90,000 lbf (400 kN) RR: 93,400 lbf (415 kN) GE: 93,700 lbf (417 kN)	GE -110B: 110,100 lbf (490 kN) GE -115B: 115,300 lbf (512 kN)		PW: 98,000 lbf (436 kN) RR: 93,400 lbf (415 kN) GE: 92,000/93,700 lbf (409 kN/418 kN)	115,300 lbf (512 kN)	105,000 lbf (470 kN)	

Flight distances were obtained from the following website: http://www.travelmath.com/flying-distance/from/Cape+Town,+South+Africa/to/Antarctica. The distance per flight route could therefore be assigned to the recorded routes in the ACSA database.

A4.4 Analysis results

A compilation of the resulting estimates is presented in Table A4.2. This reflects jet fuel consumption only, excluding aviation gasoline.

The split between international and domestic consumption varied between 68:32 (2002) and 76:24 (2012), giving an average of 73:27 with an increasing trend in international consumption evident. These splits are significantly different from those recorded by the DOE (42:58, international to domestic).

It is to be noted that the figures for the period 2002 to 2003 could not be determined from the ACSA data but are estimated stochastically, based on patterns emerging for the period 2004 to 2012.

The ACSA data provides sufficient detail to distinguish between international (INT), domestic (DOM), and regional (REG) consumption. ACSA classifies regional flights as those servicing the SACU countries, namely the southern African customs union of Namibia, Lesotho, Swaziland and Botswana. However, for the purposes of this analysis, the regional flight information is combined with that for the international category. As is the case for international flights, the regional flights also depart from South African territory to land in foreign countries and according to the IPCC guidelines the "regional" flight statistics should therefore be combined with the international statistics. Thus, although the estimates reported in Table A4.2 reflect the ACSA distinction between international, domestic and regional categories, the final ratio estimates reflect the combination of international and regional consumption, against that of domestic use.

The observed trend in the split appears to be levelling off, as can be seen from Figure A4.2. It would appear that future splits will remain at about 75:25 for some time.

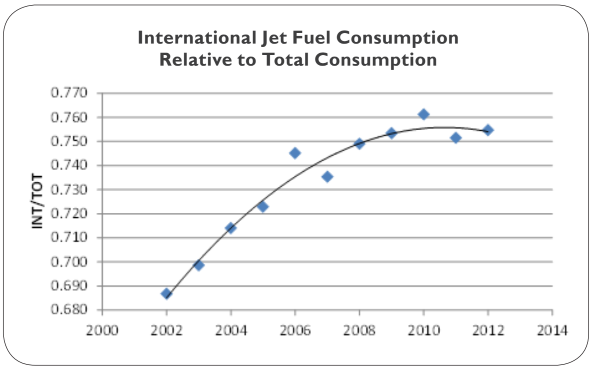


Figure A4.2: Trend of international jet fuel consumption relative to total

As indicated in Section A4.2 the above international: domestic splits are applied in this report on jet fuel consumption. It is evident from the above bottom-up estimation that it is impossible to maintain the split as reported by the DOE. It would appear that the respondents to the Fuel Sales surveys do not necessarily base their responses on detailed audits.

Although estimated averages are reported in Table A4.2 and the international : domestic ratios are based on average calculations, these figures are naturally descriptive summaries of outcomes with associated statistical distributions. In this analysis it was found that the reported averages are somewhat biased towards the lower end in terms of international and regional consumption distributions, while that for domestic consumption appear to be fair. It is, however, suggested that the simplified view, based on averages, be used for this assessment of the international : domestic split. As it stands, the outcome of the bottom-up split assessment already suggests that the DOE splits be amended. Detailed estimates, with added complexity incorporating probability distributions, do not contribute significantly to the accuracy, while requiring significantly more effort to produce.

Table A4.2: Estimates of factors contributing to split calculation

Year / Units	Departures			Route distance			Seating capacity			Fuel efficiency			Ratio	
	INT	DOM	REG	INT	DOM	REG	INT	DOM	REG	INT	DOM	REG	INT/ TOT	DOM/ TOT
	Count	Count	Count	km	km	km	Count	Count	Count	l/100km/ ask	l/100km/ ask	l/100km/ ask		
2002	19 737	140 823	11 594	6 234	885	706	303	138	74	2.981	3.029	3.863	0.685	0.315
2003	21 980	141 894	11 804	5 837	892	715	308	134	72	2.926	3.050	3.875	0.695	0.305
2004	24 516	142 021	11 421	5 747	883	684	319	139	72	2.915	3.068	4.359	0.714	0.286
2005	25 340	142 228	11 445	5 803	866	701	316	140	73	2.896	3.054	4.249	0.723	0.277
2006	27 017	134 042	10 944	5 696	855	687	314	140	78	2.911	3.041	3.960	0.745	0.255
2007	30 619	148 982	10 955	5 530	864	677	304	142	70	2.951	3.023	4.187	0.735	0.265
2008	31 581	146 371	11 411	5 204	869	656	314	141	80	2.946	3.049	3.876	0.738	0.262
2009	32 497	137 572	11 920	5 311	871	677	302	141	92	2.983	3.066	3.656	0.753	0.247
2010	34 559	127 881	11 433	5 145	899	699	299	147	99	3.003	3.016	3.372	0.761	0.239
2011	43 592	137 900	11 947	5 165	876	719	301	146	95	2.984	3.054	3.393	0.751	0.249
2012	33 955	128 157	11 477	5 154	886	751	300	152	97	2.945	2.978	3.354	0.755	0.245
Mean	28 763	138 897	11 487	5 530	877	697	307	142	82	2.949	3.039	3.831	0.732	0.268
Std Dev	5 293	6 739	332	363	13	26	7	5	11	0.035	0.026	0.355	0.025	0.025

A4.5 Future approach

The above assessment would only be possible if the detailed flight data, as recorded by ACSA, could be used. ACSA is, understandably, a bit reluctant to provide route data to outside parties which could reveal the involvement and strategies of specific airlines.

The University of Johannesburg (UJ), however, has an agreement with ACSA allowing it to use the detailed dataset to generate publications in exchange for updating a set of forecasting models which were produced for ACSA. The publications are, of course, always subject to ACSA approval. No disaggregated airline information may, under this agreement, be divulged to third parties. The above assessment was based on the data that UJ currently has at its disposal and this publication was approved by ACSA.

It is recommended that the above simplified methodology be employed by the DEA to update future splits between international and domestic consumption of jet fuel. Access to the data could be arranged through a memorandum of understanding and non-disclosure agreements with ACSA.

Alternatively, the DEA may approach the Institute for Transport and Logistic Studies (ITLS) at the University of Johannesburg to conduct this assessment on an annual basis. Such an arrangement will not require any additional legal agreements with ACSA, while still maintaining the confidentiality of detailed airline information.

ANNEXURE 5 R CODE FOR DATA PREPARATION AND FORECAST MODELLING

A5 General comments regarding the R code used in this assignment

A5.1 The interfaces to R

Please note that the R code presented in this section would be good for use in the RStudio interface.

It is equally simple to use RExcel as the interface to R (R can then run in the background). As indicated in the comment blocks at the beginning of the code sections it can simply be copied and pasted into an Excel spreadsheet in the file containing the input data. Rather than importing the data, as would normally be done in RStudio, the statements defining the "loadpath" and reading of the data from a .csv-file can be ignored while the data may be loaded straight from Excel. It can be loaded directly into memory by using the RExcel ("PUT") instructions. Kindly ensure that the variable names that are assigned through RExcel are the same as those listed below – the rest of the code below can then be used in RExcel exactly as it would have been used in RStudio.

A5.2 R Code for preparation of the exogenous data set:

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
##                                                                                              ##
##                   Copy this code into Excel and run from RExcel, if needed                   ##
##                                                                                              ##
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #

##===========================================================================
##
## STARTING POINT OF THE RUN:
##
## R Code for preparation of the exogenous data set
##
## Prepare environment for R processing
##---------------------------------------------------------------------------
## If the necessary packages below have not yet been loaded into the R instance
## used, then uncomment the statements below (remove "#" at beginning of each
## statement) and allow packages to be installed.
##---------------------------------------------------------------------------
#Install.packages("vars")
#Install.packages("tseries")
#Install.packages("tempdisagg")
#Install.packages("urca")
#Install.packages("forecast")

## Load packages required for this section of the analysis
##---------------------------------------------------------------------------
```

```
library(tseries)
library(vars)

####################################################################################
##
## Define the directory where data is located and import the data into R memory
##------------------------------------------------------------------------------
## NOTE: This section is not required when using RExcel
##
## Simply define exogenous data-frame as "exo.df" if code is run from REXcel
##------------------------------------------------------------------------------
loadPath = "H:/"              # ... or appropriate location on the analyst's PC
exo.df = read.csv(file=paste(loadPath, "qexodata.csv", sep=""), header=TRUE)
##
###################################################################################

## Convert observed data to logarithms
##------------------------------------------------------------------------------
exo.df2 = log(exo.df)

## Convert the dataset into time series format
##------------------------------------------------------------------------------
exo.ts = ts(exo.df2, start=c(1960,1), end=c(2015,1), frequency=4)
exo.ts2 = window(exo.ts[,1:10], start=c(1980,1), frequency=4)
## Test for number of lags to be included in VAR-model
##------------------------------------------------------------------------------
## This section was used in the initial analysis.  However, if it needs to be
## reviewed, please uncomment the following statements
##------------------------------------------------------------------------------
#var.test.dif = VARselect(diff(exo.ts2), lag.max=8, type="const")
#var.test.dif

## Fit VAR-model
##------------------------------------------------------------------------------
exo.var.dif = VAR(diff(exo.ts2), p = 5, type = "const")

##
## Uncomment the following statements if the results need to be investigated
##------------------------------------------------------------------------------
#summary(exo.var.dif)
#plot(exo.var.dif)

## Test for stability
##------------------------------------------------------------------------------
## This section was used in the initial analysis.  However, if it needs to be
## reviewed, please uncomment the following statements
##------------------------------------------------------------------------------
#roots(exo.var.dif)
```

```
#exo.stabil.dif = stability(exo.var.dif, type = "OLS-CUSUM", dynamic = FALSE)
#plot(exo.stabil.dif)

## Test for normality of residuals
##-----------------------------------------------------------------------------
## This section was used in the initial analysis.  However, if it needs to be
## reviewed, please uncomment the following statements
##-----------------------------------------------------------------------------
#exo.norm.dif = normality.test(exo.var.dif)
#exo.norm.dif

## Test for no serial correlation in residuals
##-----------------------------------------------------------------------------
## This section was used in the initial analysis.  However, if it needs to be
## reviewed, please uncomment the following statements
##-----------------------------------------------------------------------------
#exo.ser.dif = serial.test(exo.var.dif, type = "BG")
#exo.ser.dif

## Test for no arch-effect in residuals
##-----------------------------------------------------------------------------
## This section was used in the initial analysis.  However, if it needs to be
## reviewed, please uncomment the following statements
##-----------------------------------------------------------------------------
#exo.arch.dif = arch.test(exo.var.dif)
#exo.arch.dif

## Run restricted coefficients model
##-----------------------------------------------------------------------------
exo.res.dif = restrict(exo.var.dif, method = "ser", thresh = 1.2817) # … at 80% CI
##-----------------------------------------------------------------------------
## This section was used in the initial analysis.  However, if it needs to be
## reviewed, please uncomment the following statements
##-----------------------------------------------------------------------------
#exo.res.dif$restrictions
#exo.var2.dif = Acoef(exo.res.dif)
#exo.var2.dif
#plot(exo.res.dif)

## Generate forecast till 2025Q4 with 90% confidence interval
##-----------------------------------------------------------------------------
exo.prd.dif = predict(exo.res.dif, n.ahead = 10*4+3, ci = 0.9)
##-----------------------------------------------------------------------------
## This section was used in the initial analysis.  However, if it needs to be
## reviewed, please uncomment the following statements
##-----------------------------------------------------------------------------
#head(exo.prd.dif)
#plot(exo.prd.dif)
```

```
#fanchart(exo.prd.dif, plot.type = “single”, colors = 3)

## Capture expected levels, low and high scenarios
##-----------------------------------------------------------------------------
AFF.lev = exp(diffinv(exo.prd.dif$fcst$AFF[,”fcst”], lag=1, differences=1,
                  xi=exo.ts2[141,”AFF”]))
AFF.lwr = exp(diffinv(exo.prd.dif$fcst$AFF[,”lower”], lag=1, differences=1,
                  xi=exo.ts2[141,”AFF”]))
AFF.upr = exp(diffinv(exo.prd.dif$fcst$AFF[,”upper”], lag=1, differences=1,
                  xi=exo.ts2[141,”AFF”]))
AFF.plus = c(exo.df[,”AFF”],AFF.lev[2:44])
AFF.plus.lwr = c(exo.df[,”AFF”],AFF.lwr[2:44])
AFF.plus.upr = c(exo.df[,”AFF”],AFF.upr[2:44])
#---------------------------------------------
MQ.lev = exp(diffinv(exo.prd.dif$fcst$MQ[,”fcst”], lag=1, differences=1,
                  xi=exo.ts2[141,”MQ”]))
MQ.lwr = exp(diffinv(exo.prd.dif$fcst$MQ[,”lower”], lag=1, differences=1,
                  xi=exo.ts2[141,”MQ”]))
MQ.upr = exp(diffinv(exo.prd.dif$fcst$MQ[,”upper”], lag=1, differences=1,
                  xi=exo.ts2[141,”MQ”]))
MQ.plus = c(exo.df[,”MQ”],MQ.lev[2:44])
MQ.plus.lwr = c(exo.df[,”MQ”],MQ.lwr[2:44])
MQ.plus.upr = c(exo.df[,”MQ”],MQ.upr[2:44])
#---------------------------------------------

MAN.lev = exp(diffinv(exo.prd.dif$fcst$MAN[,”fcst”], lag=1, differences=1,
                  xi=exo.ts2[141,”MAN”]))
MAN.lwr = exp(diffinv(exo.prd.dif$fcst$MAN[,”lower”], lag=1, differences=1,
                  xi=exo.ts2[141,”MAN”]))
MAN.upr = exp(diffinv(exo.prd.dif$fcst$MAN[,”upper”], lag=1, differences=1,
                  xi=exo.ts2[141,”MAN”]))
MAN.plus = c(exo.df[,”MAN”],MAN.lev[2:44])
MAN.plus.lwr = c(exo.df[,”MAN”],MAN.lwr[2:44])
MAN.plus.upr = c(exo.df[,”MAN”],MAN.upr[2:44])
#---------------------------------------------
EGW.lev = exp(diffinv(exo.prd.dif$fcst$EGW[,”fcst”], lag=1, differences=1,
                  xi = exo.ts2[141,”EGW”]))
EGW.lwr = exp(diffinv(exo.prd.dif$fcst$EGW[,”lower”], lag=1, differences=1,
                  xi=exo.ts2[141,”EGW”]))
EGW.upr = exp(diffinv(exo.prd.dif$fcst$EGW[,”upper”], lag=1, differences=1,
                  xi=exo.ts2[141,”EGW”]))
EGW.plus = c(exo.df[,”EGW”],EGW.lev[2:44])
EGW.plus.lwr = c(exo.df[,”EGW”],EGW.lwr[2:44])
EGW.plus.upr = c(exo.df[,”EGW”],EGW.upr[2:44])
#---------------------------------------------
CONS.lev = exp(diffinv(exo.prd.dif$fcst$CONS[,”fcst”], lag=1, differences=1,
                   xi=exo.ts2[141,”CONS”]))
CONS.lwr = exp(diffinv(exo.prd.dif$fcst$CONS[,”lower”], lag=1, differences=1,
```

```
                  xi=exo.ts2[141,”CONS”]))
CONS.upr = exp(diffinv(exo.prd.dif$fcst$CONS[,”upper”], lag=1, differences=1,
                  xi=exo.ts2[141,”CONS”]))
CONS.plus = c(exo.df[,”CONS”],CONS.lev[2:44])
CONS.plus.lwr = c(exo.df[,”CONS”],CONS.lwr[2:44])
CONS.plus.upr = c(exo.df[,”CONS”],CONS.upr[2:44])
#--------------------------------------------
WRCA.lev = exp(diffinv(exo.prd.dif$fcst$WRCA[,”fcst”], lag=1, differences=1,
                  xi=exo.ts2[141,”WRCA”]))
WRCA.lwr = exp(diffinv(exo.prd.dif$fcst$WRCA[,”lower”], lag=1, differences=1,
                  xi=exo.ts2[141,”WRCA”]))
WRCA.upr = exp(diffinv(exo.prd.dif$fcst$WRCA[,”upper”], lag=1, differences=1,
                  xi=exo.ts2[141,”WRCA”]))
WRCA.plus = c(exo.df[,”WRCA”],WRCA.lev[2:44])
WRCA.plus.lwr = c(exo.df[,”WRCA”],WRCA.lwr[2:44])
WRCA.plus.upr = c(exo.df[,”WRCA”],WRCA.upr[2:44])
#--------------------------------------------
TSC.lev = exp(diffinv(exo.prd.dif$fcst$TSC[,”fcst”], lag=1, differences=1,
                 xi=exo.ts2[141,”TSC”]))
TSC.lwr = exp(diffinv(exo.prd.dif$fcst$TSC[,”lower”], lag=1, differences=1,
                 xi=exo.ts2[141,”TSC”]))
TSC.upr = exp(diffinv(exo.prd.dif$fcst$TSC[,”upper”], lag=1, differences=1,
                 xi=exo.ts2[141,”TSC”]))
TSC.plus = c(exo.df[,”TSC”],TSC.lev[2:44])
TSC.plus.lwr = c(exo.df[,”TSC”],TSC.lwr[2:44])
TSC.plus.upr = c(exo.df[,”TSC”],TSC.upr[2:44])
#--------------------------------------------
FIREBS.lev = exp(diffinv(exo.prd.dif$fcst$FIREBS[,”fcst”], lag=1, differences=1,
                   xi=exo.ts2[141,”FIREBS”]))
FIREBS.lwr = exp(diffinv(exo.prd.dif$fcst$FIREBS[,”lower”], lag=1, differences=1,
                   xi=exo.ts2[141,”FIREBS”]))
FIREBS.upr = exp(diffinv(exo.prd.dif$fcst$FIREBS[,”upper”], lag=1, differences=1,
                   xi=exo.ts2[141,”FIREBS”]))
FIREBS.plus = c(exo.df[,”FIREBS”],FIREBS.lev[2:44])
FIREBS.plus.lwr = c(exo.df[,”FIREBS”],FIREBS.lwr[2:44])
FIREBS.plus.upr = c(exo.df[,”FIREBS”],FIREBS.upr[2:44])
#--------------------------------------------
GOV.lev = exp(diffinv(exo.prd.dif$fcst$GOV[,”fcst”], lag=1, differences=1,
                 xi=exo.ts2[141,”GOV”]))
GOV.lwr = exp(diffinv(exo.prd.dif$fcst$GOV[,”lower”], lag=1, differences=1,
                 xi=exo.ts2[141,”GOV”]))
GOV.upr = exp(diffinv(exo.prd.dif$fcst$GOV[,”upper”], lag=1, differences=1,
                 xi=exo.ts2[141,”GOV”]))
GOV.plus = c(exo.df[,”GOV”],GOV.lev[2:44])
GOV.plus.lwr = c(exo.df[,”GOV”],GOV.lwr[2:44])
GOV.plus.upr = c(exo.df[,”GOV”],GOV.upr[2:44])
#--------------------------------------------
CSPS.lev = exp(diffinv(exo.prd.dif$fcst$CSPS[,”fcst”], lag=1, differences=1,
```

```
                  xi=exo.ts2[141,”CSPS”]))
CSPS.lwr = exp(diffinv(exo.prd.dif$fcst$CSPS[,”lower”], lag=1, differences=1,
                  xi=exo.ts2[141,”CSPS”]))
CSPS.upr = exp(diffinv(exo.prd.dif$fcst$CSPS[,”upper”], lag=1, differences=1,
                  xi=exo.ts2[141,”CSPS”]))
CSPS.plus = c(exo.df[,”CSPS”],CSPS.lev[2:44])
CSPS.plus.lwr = c(exo.df[,”CSPS”],CSPS.lwr[2:44])
CSPS.plus.upr = c(exo.df[,”CSPS”],CSPS.upr[2:44])
#---------------------------------------------
expExo = cbind(AFF.plus,MQ.plus,MAN.plus,EGW.plus,CONS.plus,WRCA.plus,TSC.plus,
            FIREBS.plus,GOV.plus,CSPS.plus)
lwrExo = cbind(AFF.plus.lwr,MQ.plus.lwr,MAN.plus.lwr,EGW.plus.lwr,CONS.plus.lwr,
            WRCA.plus.lwr,TSC.plus.lwr,FIREBS.plus.lwr,GOV.plus.lwr,
            CSPS.plus.lwr)
uprExo = cbind(AFF.plus.upr,MQ.plus.upr,MAN.plus.upr,EGW.plus.upr,CONS.plus.upr,
            WRCA.plus.upr,TSC.plus.upr,FIREBS.plus.upr,GOV.plus.upr,
            CSPS.plus.upr)

colnames(expExo) = colnames(lwrExo) = colnames(uprExo) = c
(“AFF”,”MQ”,”MAN”,”EGW”, “CONS”,”WRCA”,”TSC”,”FIREBS”,”GOV”,”CSPS”)

##
## Write the output to disc
write.csv(expExo, file=paste(loadPath, “newExoData.csv”))

## Uncomment the following statements if these value sets also need to be written
##   to disc
##----------------------------------------------------------------------------
#write.csv(lwrExo, file=paste(loadPath, “lwrExoData.csv”))
#write.csv(uprExo, file=paste(loadPath, “uprExoData.csv”))
##
##=========================================================================
```

A5.3 R Code for disaggregating the endogenous data set:

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
##                                                                                                                                    ##
##                                 Copy this code into Excel and run from RExcel, if needed                                           ##
##                                                                                                                                    ##
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

```
#===========================================================================================
##
## STARTING POINT OF THE RUN:
##
## R Code for disaggregating the endogenous data set
##
## Run this section only after the long term exogenous data forecasts (A5.2) have
## been completed
##
##-----------------------------------------------------------------------------
## Load packages required for this section of the analysis
##-----------------------------------------------------------------------------
library(tseries)
library(tempdisagg)

##########################################################################################
##
## Define the directory where the endogenous data is located and import to memory
##-----------------------------------------------------------------------------
## NOTE: This section is not required when using RExcel
##
## Simply define liquid fuels dataframe as "lf.df0" if code is run from REXcel, and
## Define exogenous dataframe as "exodata.df" if code is run from REXcel
##-----------------------------------------------------------------------------
loadPath = "H:/"        # ... or appropriate location on the analyst's PC
lf.df0 = read.csv(file=paste(loadPath, "lfdata.csv", sep=""), header=TRUE)
exodata.df = read.csv(file=paste(loadPath, "newExoData.csv", sep=" "), header=TRUE)
##
##########################################################################################

## Select relevant section
##-----------------------------------------------------------------------------
exodata.df = exodata.df[,2:11]

## Convert everything to logs
##-----------------------------------------------------------------------------
lf.df0 = lf.df0 + 1e-6   # ... to prevent log-function in next step to produce NaNs
lf.df2 = log(lf.df0)
exodata.df2 = log(exodata.df)
```

```
## Convert datasets to time series format
##------------------------------------------------------------------------------
lf.ts = ts(lf.df2, start = 2000, frequency = 1)
exodata.ts = ts(exodata.df2, start = c(1960,1), frequency = 4)
exodata.ts2 = window(exodata.ts, start = c(2000,1), end = c(2012,4), frequency = 4)
exodata.ts3 = window(exodata.ts, start = c(2013,1), frequency = 4)

## Define outlier removal function
##------------------------------------------------------------------------------
remove_outliers <- function(x, na.rm = TRUE, ...) {
  qnt <- quantile(x, probs=c(.25, .75), na.rm = na.rm, ...)
  H <- 1.5 * IQR(x, na.rm = na.rm)
  y <- x
  y[x < (qnt[1] - H)] <- qnt[1] - H
  y[x > (qnt[2] + H)] <- qnt[2] + H
  y
}

## Interpolate annual data for liquid fuels onto quarterly basis, as a function of
##   exogenous database
##------------------------------------------------------------------------------
lfendo <- matrix(rep(NA, nrow(exodata.ts2)*ncol(lf.ts)), nrow=13*4)
for (i in 1:length(colnames(lf.ts))){
  lfendo[,i] = t(predict(td(lf.ts[,i] ~ exodata.ts2[,"AFF"] + exodata.ts2[,"MQ"]+
                  exodata.ts2[,"MAN"] + exodata.ts2[,"EGW"] +
exodata.ts2[,"CONS"] +
                  exodata.ts2[,"WRCA"] + exodata.ts2[,"TSC"] +
exodata.ts2[,"FIREBS"] +
                  exodata.ts2[,"GOV"] + exodata.ts2[,"CSPS"],
conversion = "average")))
  lfendo[,i] = log(remove_outliers(exp(lfendo[,i])))
}
lfendo.ts = ts(lfendo, start = c(2000,1), frequency = 4)
colnames(lfendo.ts) <- colnames(lf.ts)

## Combine endogenous and exogenous data-files and write to disc
##------------------------------------------------------------------------------
lfdata.ts = exp(cbind(lfendo.ts, exodata.ts2))
colnames(lfdata.ts) = c(colnames(lfendo.ts),colnames(exodata.ts2))
write.csv(lfdata.ts,file = paste(loadPath,"quarterlyData.csv"))
##
##==============================================================================
```

A5.4 R Code for fitting the appropriate models and forecasting liquid fuels consumption:

```
#################################################################
##                                                             ##
##       Copy this code into Excel and run from RExcel, if needed       ##
##                                                             ##
#################################################################
```

```
##===========================================================================
##
## STARTING POINT OF THE RUN:
##
## R Code for fitting appropriate models and forecasting liquid fuels consumption
##
## IMPORTANT TO NOTE:-
##
## Run this section only after disaggregation of annual data (A5.3) and after the
## long term forecasts for the exogenous variables (A5.2) have been completed
##
##-----------------------------------------------------------------------------
## Load packages required for this section of the analysis
##-----------------------------------------------------------------------------
library(tseries)
library(forecast)
library(vars)
library(urca)

##############################################################################
##
## Define the directory where the endogenous data is located and import to memory
##-----------------------------------------------------------------------------
## NOTE: This section is not required when using RExcel
##
## Simply define liquid fuels dataframe as “lf.df0” if code is run from REXcel, and
## Define exogenous dataframe as “exodata.df” if code is run from REXcel
##-----------------------------------------------------------------------------
#loadPath = “H:/”        # ... or appropriate location on the analyst’s PC
#lf.df0 = read.csv(file=paste(loadPath, “lfdata.csv”, sep=“”), header=TRUE)
#exodata.df = read.csv(file=paste(loadPath, “newExoData.csv”, sep=” “),
header=TRUE)
##
##############################################################################
#exodata.df = exodata.df[,2:11]

# Convert datasets to time series format
##-----------------------------------------------------------------------------
#lf.df = ts(lf.df0, start = 2000, frequency = 1)
```

```
#exodata.ts = ts(exodata.df2, start = c(1960,1), frequency = 4)
#exodata.ts2 = window(exodata.ts, start = c(2000,1), end = c(2012,4), frequency =
4)
#exodata.ts3 = window(exodata.ts, start = c(2013,1), frequency = 4)

## Re-assign data for ease of use
##---------------------------------------------------------------------------------
lf.df = exp(lfendo.ts)
exodata.dfs = expExo

## Data preparation - summation per energy carrier
##---------------------------------------------------------------------------------
coal    =
rowSums(lf.df[,c(“agr_col”,”min_col”,”chm_col”,”ref_col”,”cns_col”,”fbt_col”,

                  “irs_col”,”mac_col”,”nfm_col”,”nmm_col”,”oth_col”,”gen_col”,

                  “plp_col”,”cmi_col”,”res_col”,”avd_col”,”avi_col”,”wwd_col”,
                                          “wwi_col”,”tra_col”,”trd_col”)])
biomass  =
rowSums(lf.df[,c(“agr_bio”,”min_bio”,”chm_bio”,”ref_bio”,”cns_bio”,”fbt_bio”,

                  “irs_bio”,”mac_bio”,”nfm_bio”,”nmm_bio”,”oth_bio”,”gen_bio”,

                  “plp_bio”,”cmi_bio”,”res_bio”,”avd_bio”,”avi_bio”,”wwd_bio”,
                                          “wwi_bio”,”tra_bio”,”trd_bio”)])
petrol   =
rowSums(lf.df[,c(“agr_pet”,”min_pet”,”chm_pet”,”ref_pet”,”cns_pet”,”fbt_pet”,

                  “irs_pet”,”mac_pet”,”nfm_pet”,”nmm_pet”,”oth_pet”,”gen_pet”,

                  “plp_pet”,”cmi_pet”,”res_pet”,”avd_pet”,”avi_pet”,”wwd_pet”,
                                          “wwi_pet”,”tra_pet”,”trd_pet”)])
diesel   =
rowSums(lf.df[,c(“agr_dsl”,”min_dsl”,”chm_dsl”,”ref_dsl”,”cns_dsl”,”fbt_dsl”,

                  “irs_dsl”,”mac_dsl”,”nfm_dsl”,”nmm_dsl”,”oth_dsl”,”gen_dsl”,

                  “plp_dsl”,”cmi_dsl”,”res_dsl”,”avd_dsl”,”avi_dsl”,”wwd_dsl”,
                                          “wwi_dsl”,”tra_dsl”,”trd_dsl”)])
paraffin = rowSums(lf.df[,c(“agr_ip”,”min_ip”,”chm_ip”,”ref_ip”,”cns_ip”,”fbt_ip”,
                  “irs_ip”,”mac_ip”,”nfm_ip”,”nmm_ip”,”oth_ip”,”gen_ip”,
                  “plp_ip”,”cmi_ip”,”res_ip”,”avd_ip”,”avi_ip”,”wwd_ip”,
                  “wwi_ip”,”tra_ip”,”trd_ip”)])
fueloil  =
rowSums(lf.df[,c(“agr_hfo”,”min_hfo”,”chm_hfo”,”ref_hfo”,”cns_hfo”,”fbt_hfo”,

                  “irs_hfo”,”mac_hfo”,”nfm_hfo”,”nmm_hfo”,”oth_hfo”,”gen_hfo”,
```

```
                          “plp_hfo”,”cmi_hfo”,”res_hfo”,”avd_hfo”,”avi_hfo”,”wwd_hfo”,
                                                  “wwi_hfo”,”tra_hfo”,”trd_hfo”)])
lpg     =
rowSums(lf.df[,c(“agr_lpg”,”min_lpg”,”chm_lpg”,”ref_lpg”,”cns_lpg”,”fbt_lpg”,

                          “irs_lpg”,”mac_lpg”,”nfm_lpg”,”nmm_lpg”,”oth_lpg”,”gen_lpg”,

                          “plp_lpg”,”cmi_lpg”,”res_lpg”,”avd_lpg”,”avi_lpg”,”wwd_lpg”,
                                                  “wwi_lpg”,”tra_lpg”,”trd_lpg”)])
natgas   =
rowSums(lf.df[,c(“agr_gas”,”min_gas”,”chm_gas”,”ref_gas”,”cns_gas”,”fbt_gas”,

                          “irs_gas”,”mac_gas”,”nfm_gas”,”nmm_gas”,”oth_gas”,”gen_gas”,

                          “plp_gas”,”cmi_gas”,”res_gas”,”avd_gas”,”avi_gas”,”wwd_gas”,
                                                  “wwi_gas”,”tra_gas”,”trd_gas”)])
jetfuel  =
rowSums(lf.df[,c(“agr_jet”,”min_jet”,”chm_jet”,”ref_jet”,”cns_jet”,”fbt_jet”,

                          “irs_jet”,”mac_jet”,”nfm_jet”,”nmm_jet”,”oth_jet”,”gen_jet”,

                          “plp_jet”,”cmi_jet”,”res_jet”,”avd_jet”,”avi_jet”,”wwd_jet”,
                                                  “wwi_jet”,”tra_jet”,”trd_jet”)])
avgas    =
rowSums(lf.df[,c(“agr_avg”,”min_avg”,”chm_avg”,”ref_avg”,”cns_avg”,”fbt_avg”,

                          “irs_avg”,”mac_avg”,”nfm_avg”,”nmm_avg”,”oth_avg”,”gen_avg”,

                          “plp_avg”,”cmi_avg”,”res_avg”,”avd_avg”,”avi_avg”,”wwd_avg”,
                                                  “wwi_avg”,”tra_avg”,”trd_avg”)])

lfagg.dfs =
cbind(petrol,diesel,jetfuel,paraffin,lpg,fueloil,natgas,coal,avgas,biomass)

## Convert data to logarithms
##-------------------------------------------------------------------------------
lfagg.dfs = lfagg.dfs + 1e-6
lfagg.df2 = log(lfagg.dfs)
exodata.df2 = log(exodata.dfs)

## Convert dataset to time-series format
##-------------------------------------------------------------------------------
lfagg.ts = ts(lfagg.df2, start = c(2000,1), frequency = 4)
exodata.ts = ts(exodata.df2, start = c(1960,1), frequency = 4)
exodata.ts2 = window(exodata.ts, start = c(2000,1), end = c(2012,4), frequency = 4)
exodata.ts3 = window(exodata.ts, start = c(2013,1), frequency = 4)
```

```
## Select confidence interval and forecast period length
##-------------------------------------------------------------------------------
ci = 0.9
fper = 4*13      # number of quarters ahead
## Preferred analysis: consider full data set
##-------------------------------------------------------------------------------
coint = ca.jo(lfagg.ts,
            K=2, spec =”longrun”, season=4, ecdet=”trend”)
lf.var = vec2var(coint, r=5)
lf.pred = predict(lf.var, n.ahead = fper, ci = ci, dumvar = exodata.ts3)
Forecast.lev = cbind(lf.pred$fcst$petrol[,”fcst”],lf.pred$fcst$diesel[,”fcst”],
                 lf.pred$fcst$jetfuel[,”fcst”],lf.pred$fcst$paraffin[,”fcst”],
                 lf.pred$fcst$lpg[,”fcst”],lf.pred$fcst$fueloil[,”fcst”],
                 lf.pred$fcst$natgas[,”fcst”],lf.pred$fcst$coal[,”fcst”],
                 lf.pred$fcst$avgas[,”fcst”],lf.pred$fcst$biomass[,”fcst”])
Forecast.upr = cbind(lf.pred$fcst$petrol[,”upper”],lf.pred$fcst$diesel[,”upper”],
 lf.pred$fcst$jetfuel[,”upper”],lf.pred$fcst$paraffin[,”upper”],
                 lf.pred$fcst$lpg[,”upper”],lf.pred$fcst$fueloil[,”upper”],
                 lf.pred$fcst$natgas[,”upper”],lf.pred$fcst$coal[,”upper”],
                 lf.pred$fcst$avgas[,”upper”],lf.pred$fcst$biomass[,”upper”])
Forecast.lwr = cbind(lf.pred$fcst$petrol[,”lower”],lf.pred$fcst$diesel[,”lower”],
 lf.pred$fcst$jetfuel[,”lower”],lf.pred$fcst$paraffin[,”lower”],
                 lf.pred$fcst$lpg[,”lower”],lf.pred$fcst$fueloil[,”lower”],
                 lf.pred$fcst$natgas[,”lower”],lf.pred$fcst$coal[,”lower”],
                 lf.pred$fcst$avgas[,”lower”],lf.pred$fcst$biomass[,”lower”])
colnames(Forecast.lev)=colnames(Forecast.upr)=colnames(Forecast.lwr)=c(“petrol”,
 “diesel”,”jetfuel”,”paraffin”,”lpg”,”fueloil”,”natgas”,”coal”,”avgas”,”biomass”)
Forecast.lev = rbind(exp(lfagg.ts[52,]),exp(Forecast.lev))
Forecast.upr = rbind(exp(lfagg.ts[52,]),exp(Forecast.upr))
Forecast.lwr = rbind(exp(lfagg.ts[52,]),exp(Forecast.lwr))
Forecast.lev.ts = ts(Forecast.lev, start=c(2012,4), frequency=4)
Forecast.upr.ts = ts(Forecast.upr, start=c(2012,4), frequency=4)
Forecast.lwr.ts = ts(Forecast.lwr, start=c(2012,4), frequency=4)
## Generate confidence interval bands
##-------------------------------------------------------------------------------
Forecast.lwr.ts.sd = (Forecast.lev.ts-Forecast.lwr.ts)/qnorm(ci)
Forecast.lwr.ts.90 = Forecast.lwr.ts
Forecast.lwr.ts.80 = Forecast.lev.ts-Forecast.lwr.ts.sd*qnorm(0.8)
Forecast.lwr.ts.70 = Forecast.lev.ts-Forecast.lwr.ts.sd*qnorm(0.7)
```

```
colnames(Forecast.lwr.ts.70) = colnames(Forecast.lwr.ts.80) =
colnames(Forecast.lwr.ts.90) = colnames(Forecast.lev.ts)

Forecast.upr.ts.sd = (Forecast.upr.ts-Forecast.lev.ts)/qnorm(ci)
Forecast.upr.ts.90 = Forecast.upr.ts
Forecast.upr.ts.80 = Forecast.lev.ts+Forecast.upr.ts.sd*qnorm(0.8)
Forecast.upr.ts.70 = Forecast.lev.ts+Forecast.upr.ts.sd*qnorm(0.7)
colnames(Forecast.upr.ts.70) = colnames(Forecast.upr.ts.80) =
colnames(Forecast.upr.ts.90) = colnames(Forecast.lev.ts)

## Print graphs with forecasts
##-----------------------------------------------------------------------------
ts.plot(ts.union(exp(lfagg.ts[,”petrol”]),Forecast.lev.ts[,”petrol”],
            Forecast.upr.ts.70[,”petrol”], Forecast.upr.ts.80[,”petrol”],
            Forecast.upr.ts.90[,”petrol”], Forecast.lwr.ts.70[,”petrol”],
            Forecast.lwr.ts.80[,”petrol”], Forecast.lwr.ts.90[,”petrol”]),
      main=”Petrol forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”Petrol consumption”, xlim = c(2000,2026),
      lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”petrol”]),rev(index(Forecast.lev.ts[,”petrol”])))
Y1 = c(Forecast.upr.ts.70[,”petrol”],rev(Forecast.lwr.ts.70[,”petrol”]))
Y2 = c(Forecast.upr.ts.70[,”petrol”],rev(Forecast.upr.ts.80[,”petrol”]))
Y3 = c(Forecast.lwr.ts.70[,”petrol”],rev(Forecast.lwr.ts.80[,”petrol”]))
Y4 = c(Forecast.upr.ts.80[,”petrol”],rev(Forecast.upr.ts.90[,”petrol”]))
Y5 = c(Forecast.lwr.ts.80[,”petrol”],rev(Forecast.lwr.ts.90[,”petrol”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”diesel”]),Forecast.lev.ts[,”diesel”],
            Forecast.upr.ts.70[,”diesel”], Forecast.upr.ts.80[,”diesel”],
            Forecast.upr.ts.90[,”diesel”], Forecast.lwr.ts.70[,”diesel”],
            Forecast.lwr.ts.80[,”diesel”], Forecast.lwr.ts.90[,”diesel”]),
      main=”Diesel forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”Diesel consumption”, xlim = c(2000,2026),
      lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”diesel”]),rev(index(Forecast.lev.ts[,”diesel”])))
Y1 = c(Forecast.upr.ts.70[,”diesel”],rev(Forecast.lwr.ts.70[,”diesel”]))
Y2 = c(Forecast.upr.ts.70[,”diesel”],rev(Forecast.upr.ts.80[,”diesel”]))
Y3 = c(Forecast.lwr.ts.70[,”diesel”],rev(Forecast.lwr.ts.80[,”diesel”]))
Y4 = c(Forecast.upr.ts.80[,”diesel”],rev(Forecast.upr.ts.90[,”diesel”]))
Y5 = c(Forecast.lwr.ts.80[,”diesel”],rev(Forecast.lwr.ts.90[,”diesel”]))
```

```
polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”jetfuel”]),Forecast.lev.ts[,”jetfuel”],
            Forecast.upr.ts.70[,”jetfuel”], Forecast.upr.ts.80[,”jetfuel”],
            Forecast.upr.ts.90[,”jetfuel”], Forecast.lwr.ts.70[,”jetfuel”],
            Forecast.lwr.ts.80[,”jetfuel”], Forecast.lwr.ts.90[,”jetfuel”]),
      main=”Jet fuel forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”Jet fuel consumption”, xlim = c(2000,2026),
      lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”jetfuel”]),rev(index(Forecast.lev.ts[,”jetfuel”])))
Y1 = c(Forecast.upr.ts.70[,”jetfuel”],rev(Forecast.lwr.ts.70[,”jetfuel”]))
Y2 = c(Forecast.upr.ts.70[,”jetfuel”],rev(Forecast.upr.ts.80[,”jetfuel”]))
Y3 = c(Forecast.lwr.ts.70[,”jetfuel”],rev(Forecast.lwr.ts.80[,”jetfuel”]))
Y4 = c(Forecast.upr.ts.80[,”jetfuel”],rev(Forecast.upr.ts.90[,”jetfuel”]))
Y5 = c(Forecast.lwr.ts.80[,”jetfuel”],rev(Forecast.lwr.ts.90[,”jetfuel”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”paraffin”]),Forecast.lev.ts[,”paraffin”],
            Forecast.upr.ts.70[,”paraffin”], Forecast.upr.ts.80[,”paraffin”],
            Forecast.upr.ts.90[,”paraffin”], Forecast.lwr.ts.70[,”paraffin”],
            Forecast.lwr.ts.80[,”paraffin”], Forecast.lwr.ts.90[,”paraffin”]),
      main=”Paraffin forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”Paraffin consumption”, xlim = c(2000,2026),
      lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”paraffin”]),rev(index(Forecast.lev.ts[,”paraffin”])))
Y1 = c(Forecast.upr.ts.70[,”paraffin”],rev(Forecast.lwr.ts.70[,”paraffin”]))
Y2 = c(Forecast.upr.ts.70[,”paraffin”],rev(Forecast.upr.ts.80[,”paraffin”]))
Y3 = c(Forecast.lwr.ts.70[,”paraffin”],rev(Forecast.lwr.ts.80[,”paraffin”]))
Y4 = c(Forecast.upr.ts.80[,”paraffin”],rev(Forecast.upr.ts.90[,”paraffin”]))
Y5 = c(Forecast.lwr.ts.80[,”paraffin”],rev(Forecast.lwr.ts.90[,”paraffin”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)
```

```
ts.plot(ts.union(exp(lfagg.ts[,”lpg”]),Forecast.lev.ts[,”lpg”],
             Forecast.upr.ts.70[,”lpg”], Forecast.upr.ts.80[,”lpg”],
             Forecast.upr.ts.90[,”lpg”], Forecast.lwr.ts.70[,”lpg”],
             Forecast.lwr.ts.80[,”lpg”], Forecast.lwr.ts.90[,”lpg”]),
      main=”LPG forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”LPG consumption”, xlim = c(2000,2026),
      lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”lpg”]),rev(index(Forecast.lev.ts[,”lpg”])))
Y1 = c(Forecast.upr.ts.70[,”lpg”],rev(Forecast.lwr.ts.70[,”lpg”]))
Y2 = c(Forecast.upr.ts.70[,”lpg”],rev(Forecast.upr.ts.80[,”lpg”]))
Y3 = c(Forecast.lwr.ts.70[,”lpg”],rev(Forecast.lwr.ts.80[,”lpg”]))
Y4 = c(Forecast.upr.ts.80[,”lpg”],rev(Forecast.upr.ts.90[,”lpg”]))
Y5 = c(Forecast.lwr.ts.80[,”lpg”],rev(Forecast.lwr.ts.90[,”lpg”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”fueloil”]),Forecast.lev.ts[,”fueloil”],
             Forecast.upr.ts.70[,”fueloil”], Forecast.upr.ts.80[,”fueloil”],
             Forecast.upr.ts.90[,”fueloil”], Forecast.lwr.ts.70[,”fueloil”],
             Forecast.lwr.ts.80[,”fueloil”], Forecast.lwr.ts.90[,”fueloil”]),
      main=”Fuel oil forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”Fuel oil consumption”, xlim = c(2000,2026),
      lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”fueloil”]),rev(index(Forecast.lev.ts[,”fueloil”])))
Y1 = c(Forecast.upr.ts.70[,”fueloil”],rev(Forecast.lwr.ts.70[,”fueloil”]))
Y2 = c(Forecast.upr.ts.70[,”fueloil”],rev(Forecast.upr.ts.80[,”fueloil”]))
Y3 = c(Forecast.lwr.ts.70[,”fueloil”],rev(Forecast.lwr.ts.80[,”fueloil”]))
Y4 = c(Forecast.upr.ts.80[,”fueloil”],rev(Forecast.upr.ts.90[,”fueloil”]))
Y5 = c(Forecast.lwr.ts.80[,”fueloil”],rev(Forecast.lwr.ts.90[,”fueloil”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”natgas”]),Forecast.lev.ts[,”natgas”],
             Forecast.upr.ts.70[,”natgas”], Forecast.upr.ts.80[,”natgas”],
             Forecast.upr.ts.90[,”natgas”], Forecast.lwr.ts.70[,”natgas”],
             Forecast.lwr.ts.80[,”natgas”], Forecast.lwr.ts.90[,”natgas”]),
      main=”Natural gas forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”Natural gas consumption”, xlim = c(2000,2026),
```

```
        lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”natgas”]),rev(index(Forecast.lev.ts[,”natgas”])))
Y1 = c(Forecast.upr.ts.70[,”natgas”],rev(Forecast.lwr.ts.70[,”natgas”]))
Y2 = c(Forecast.upr.ts.70[,”natgas”],rev(Forecast.upr.ts.80[,”natgas”]))
Y3 = c(Forecast.lwr.ts.70[,”natgas”],rev(Forecast.lwr.ts.80[,”natgas”]))
Y4 = c(Forecast.upr.ts.80[,”natgas”],rev(Forecast.upr.ts.90[,”natgas”]))
Y5 = c(Forecast.lwr.ts.80[,”natgas”],rev(Forecast.lwr.ts.90[,”natgas”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”coal”]),Forecast.lev.ts[,”coal”],
                Forecast.upr.ts.70[,”coal”], Forecast.upr.ts.80[,”coal”],
                Forecast.upr.ts.90[,”coal”], Forecast.lwr.ts.70[,”coal”],
                Forecast.lwr.ts.80[,”coal”], Forecast.lwr.ts.90[,”coal”]),
        main=”Coal forecast with 70% - 80% - 90% confidence intervals”,
        ylab=”Coal consumption”, xlim = c(2000,2026),
        lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”coal”]),rev(index(Forecast.lev.ts[,”coal”])))
Y1 = c(Forecast.upr.ts.70[,”coal”],rev(Forecast.lwr.ts.70[,”coal”]))
Y2 = c(Forecast.upr.ts.70[,”coal”],rev(Forecast.upr.ts.80[,”coal”]))
Y3 = c(Forecast.lwr.ts.70[,”coal”],rev(Forecast.lwr.ts.80[,”coal”]))
Y4 = c(Forecast.upr.ts.80[,”coal”],rev(Forecast.upr.ts.90[,”coal”]))
Y5 = c(Forecast.lwr.ts.80[,”coal”],rev(Forecast.lwr.ts.90[,”coal”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”avgas”]),Forecast.lev.ts[,”avgas”],
                Forecast.upr.ts.70[,”avgas”], Forecast.upr.ts.80[,”avgas”],
                Forecast.upr.ts.90[,”avgas”], Forecast.lwr.ts.70[,”avgas”],
                Forecast.lwr.ts.80[,”avgas”], Forecast.lwr.ts.90[,”avgas”]),
        main=”Avgas forecast with 70% - 80% - 90% confidence intervals”,
        ylab=”Avgas consumption”, xlim = c(2000,2026),
        lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”avgas”]),rev(index(Forecast.lev.ts[,”avgas”])))
Y1 = c(Forecast.upr.ts.70[,”avgas”],rev(Forecast.lwr.ts.70[,”avgas”]))
Y2 = c(Forecast.upr.ts.70[,”avgas”],rev(Forecast.upr.ts.80[,”avgas”]))
Y3 = c(Forecast.lwr.ts.70[,”avgas”],rev(Forecast.lwr.ts.80[,”avgas”]))
```

```
Y4 = c(Forecast.upr.ts.80[,”avgas”],rev(Forecast.upr.ts.90[,”avgas”]))
Y5 = c(Forecast.lwr.ts.80[,”avgas”],rev(Forecast.lwr.ts.90[,”avgas”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

ts.plot(ts.union(exp(lfagg.ts[,”biomass”]),Forecast.lev.ts[,”biomass”],
            Forecast.upr.ts.70[,”biomass”], Forecast.upr.ts.80[,”biomass”],
            Forecast.upr.ts.90[,”biomass”], Forecast.lwr.ts.70[,”biomass”],
            Forecast.lwr.ts.80[,”biomass”], Forecast.lwr.ts.90[,”biomass”]),
      main=”Biomass forecast with 70% - 80% - 90% confidence intervals”,
      ylab=”Biomass consumption”, xlim = c(2000,2026),
      lty=c(1,1,0,0,0,0,0,0), lwd=2)

X = c(index(Forecast.lev.ts[,”biomass”]),rev(index(Forecast.lev.ts[,”biomass”])))
Y1 = c(Forecast.upr.ts.70[,”biomass”],rev(Forecast.lwr.ts.70[,”biomass”]))
Y2 = c(Forecast.upr.ts.70[,”biomass”],rev(Forecast.upr.ts.80[,”biomass”]))
Y3 = c(Forecast.lwr.ts.70[,”biomass”],rev(Forecast.lwr.ts.80[,”biomass”]))
Y4 = c(Forecast.upr.ts.80[,”biomass”],rev(Forecast.upr.ts.90[,”biomass”]))
Y5 = c(Forecast.lwr.ts.80[,”biomass”],rev(Forecast.lwr.ts.90[,”biomass”]))

polygon(X,Y1, col=gray(.5,alpha=0.5), border = FALSE)
polygon(X,Y2, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y3, col=gray(.7,alpha=0.5), border = FALSE)
polygon(X,Y4, col=gray(.9,alpha=0.5), border = FALSE)
polygon(X,Y5, col=gray(.9,alpha=0.5), border = FALSE)

##========================================================================
```

A5.5 R Code for quantification of certainty of liquid fuels consumption database:

```
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
##                                                                                                                              ##
##                                    Copy this code into Excel and run from RExcel, if needed                                  ##
##                                                                                                                              ##
# # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # # #
```

```
##==============================================================================
##
## STARTING POINT OF THE RUN:
##
## R Code for quantification of certainty of input data
##
## Estimate by applying local polynomial regression fitting
##------------------------------------------------------------------------------
## Define the directory where the endogenous data is located and import to memory
##------------------------------------------------------------------------------
loadPath = "H:/"        # ... or appropriate location on the analyst's PC
lf.df0 = read.csv(file=paste(loadPath, "lfdata.csv", sep=""), header=TRUE)
lf.df0 = lf.df0 + 1e-6

## Select the confidence interval level
##------------------------------------------------------------------------------
cii = 0.9               # ... adjust confidence interval here

## Estimate expected values with LOESS-function and determine p-values of
deviations
##------------------------------------------------------------------------------
lfdata.ts0 = ts(lf.df0, start = 2000, frequency = 1)
lf.cert = lf.fit = lf.se = lf.z = matrix(NA, nrow = nrow(lf.df0),ncol =
ncol(lf.df0))
for (i in 1:ncol(lf.df0)){
  lf.dat = log(lfdata.ts0[,i])
  lf.lo = loess(lf.dat~time(lf.dat), span = 5,
         control = loess.control(surface="interpolate"))
  lf.lo.pred = predict(lf.lo, se = TRUE)
  lf.fit[,i] = lf.lo.pred$fit
  lf.se[,i] = lf.lo.pred$se.fit
  lf.z[,i] = (lf.dat-lf.lo.pred$fit)/lf.se[,i]
  lf.cert[,i] = 2*pt(-abs(lf.z[,i]), df = lf.lo.pred$df)
}
colnames(lf.fit) = colnames(lf.se) = colnames(lfdata.ts0)
lf.fit.cert = ts(exp(lf.fit), start=2000, frequency=1)
lf.upr.cert = ts(exp(lf.fit+qnorm((1+cii)/2)*lf.se), start=2000, frequency=1)
lf.lwr.cert = ts(exp(lf.fit-qnorm((1+cii)/2)*lf.se), start=2000, frequency=1)
```

```
#lf.cert = 2*pt(-abs(lf.z), df = nrow(lfdata.ts0)-1)
colnames(lf.cert) = colnames(lfdata.ts0)
lf.cert = ts(lf.cert, start = 2000, frequency = 1)

## Write results to disc
##------------------------------------------------------------------------------
write.csv(lf.cert,"H:/Input_certainty.csv")        # ... or appropriate destination
write.csv(lf.fit.cert,"H:/Input_fit.csv")
write.csv(lf.upr.cert,"H:/Input_upr.csv")
write.csv(lf.lwr.cert,"H:/Input_lwr.csv")

##==============================================================================

```

ANNEXURE 6 SUSTAINABILITY ENERGY AFRICA (SEA) LIQUID FUEL SECTORAL DISAGGREGATION ASSUMPTIONS FOR MUNICIPALITIES

The following are the published assumptions made by SEA for their city models presented in their State of Energy in South African Cities (SEA 2015) except for one (not identified) for which SAPIA/DoE Market Category data was available:

- All petrol consumption was assigned to the transport sector.
- All diesel consumption was assigned to the transport sector, although an unknown proportion of diesel is used for stationary combustion in the commercial and industrial sectors. Diesel use at Eskom's Ankerlig power plant, based in the Cape Town municipal area, was subtracted from Cape Town's diesel use figure to avoid double-counting (the fuel is used to generate electricity, which is accounted for in the electricity use data).
- All heavy fuel oil consumption was assigned to the industrial sector.
- Paraffin use was assigned entirely to the residential sector due to the data age and uncertainty about where to apportion the 30% of paraffin not consumed by households: according to a 2003 National Treasury Report (PDC & SCE 2003), households consume over 70% of paraffin.
- LPG consumption was split 25% residential, 25% commercial and 50% industrial use, based on LPG allocations in the Cape Town Long-term Mitigation Scenarios (LTMS) work, which was based on interviews with LPG suppliers. There have been no detailed studies on LPG use in the country. More research is required in the future.
- All jet fuel and aviation gasoline was assigned to the transport sector. As was done in the case of Ankerlig, the use of jet fuel at Eskom's Acacia power plant (also situated within the Cape Town municipal boundaries) was subtracted from total jet fuel use.

Environment House
473 Steve Biko
cnr Steve Biko and Soutpansberg Road
Arcadia
Pretoria, 0083
South Africa

Postal Address
Department of Environmental Affairs
P O Box 447
Pretoria
0001

Publishing date: March 2016

www.environment.gov.za