Demand Projections and their Role in Power System Planning

21 April 2016

Morore MASHAO
Tel:+27 11 800 3822
Mobile: +27 78 457 1629
Email: Mashaomj@eskom.co.za

Outline

- Introduction
- Key Inputs
- Methodology
- Typical Outputs

Energy Forecast Assumptions

- Three main assumptions were used to perform the energy forecast:
- GDP Growth Rate (%)
- Electricity Intensity
- Dx, Tx and RSA Losses (%)
- Calculate Electricity Intensity as a form of sanity check and compare with countries with similar economic structure as RSA

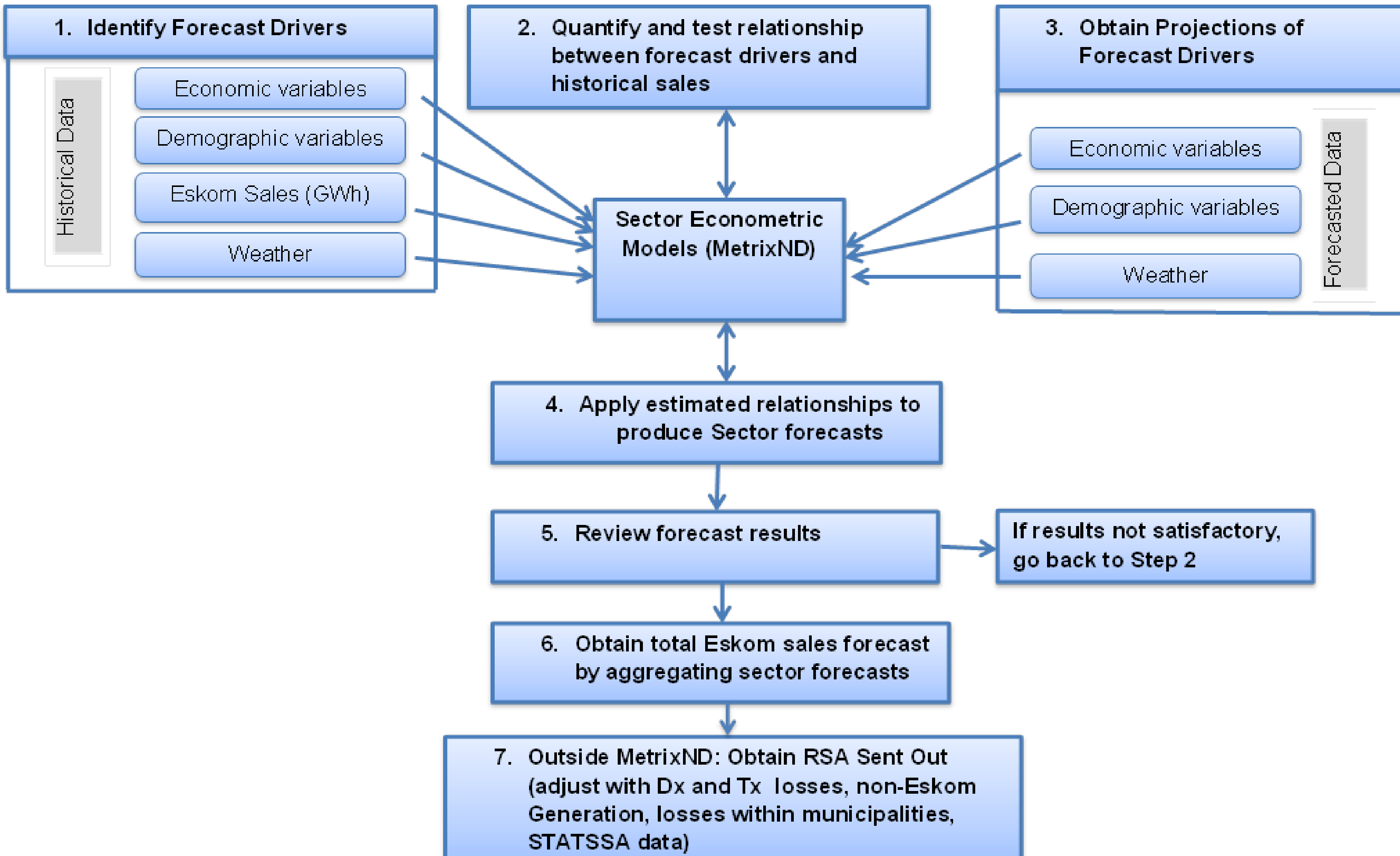
Bottom up approach

- A bottom up annual energy forecast is developed from individual forecasts of industry, mining, traction, commercial, residential and agricultural customers.
- Redistributors are not final consumers of electricity, so their sales are allocated to respective economic sectors using historical usage patterns of electricity within municipalities.
- Estimates of future international sales are obtained from SAE.
- DSM, Buybacks, IOS are added back to the historical energy sales data.
- A top down approach is also used to develop the energy forecast from a combination of electricity intensity, GDP and assumed system losses for the forecast period.

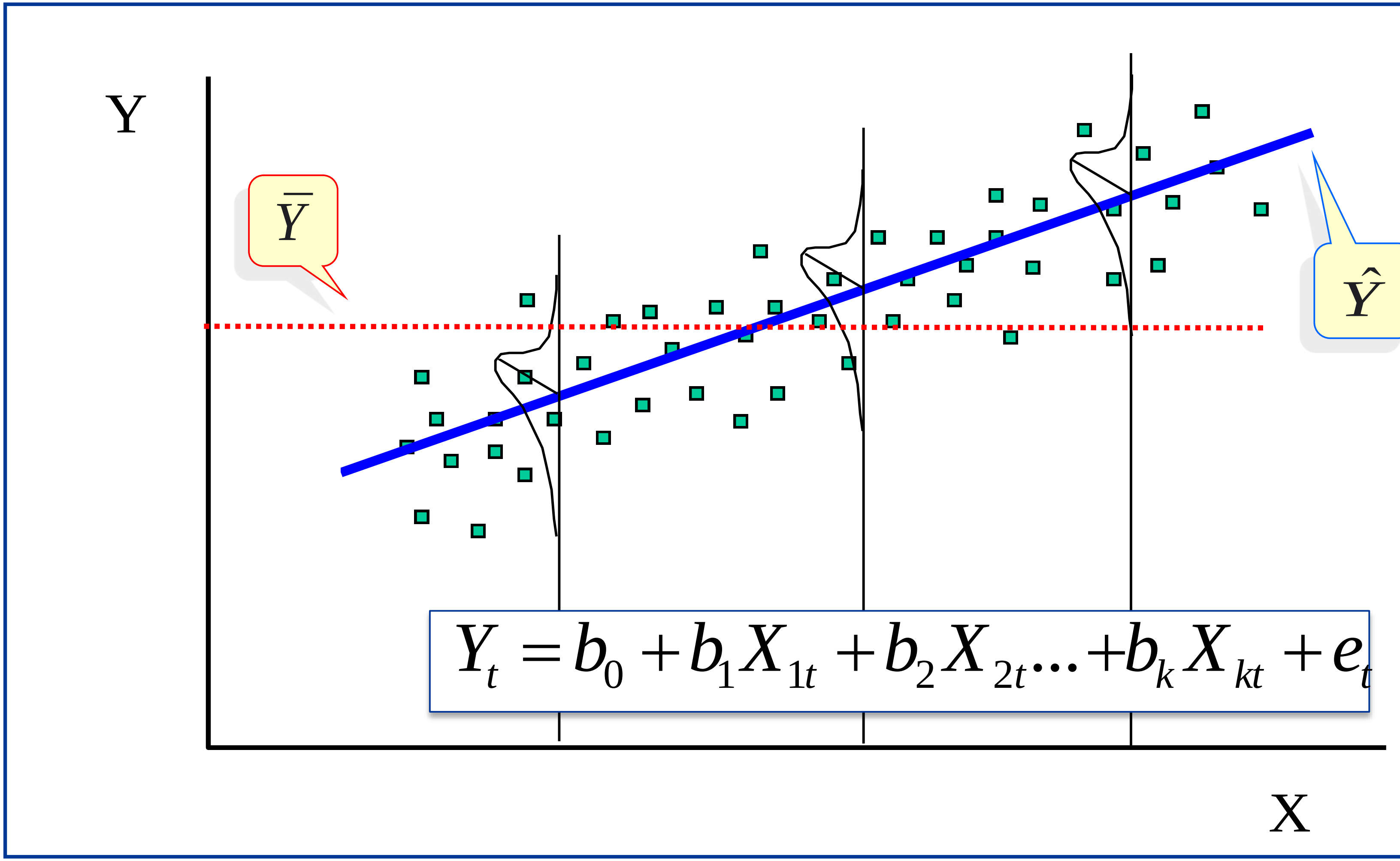
Top Down Energy Forecast Process

- Econometric regression models which are developed using MetrixND software to forecast annual sales of electricity in each of the economic sectors.
- Economic and demographic drivers which are used in the energy forecast include:
  - GDP, population,
  - Coal production index, PGM production index, gold ore treated,
  - Households – FCEH
  - Population
  - manufacturing production index, total mining production less gold index.
- Historical data of these drivers is obtained from publicly available sources, mainly StatsSA.
- Projections of these economic and demographic drivers are developed internally within Eskom using external data sets.
- Projections of these drivers are provided for four economic scenarios.

MetrixND ENERGY FORECAST PROCESS FLOW



REGRESSION MODELLING



The aim is to get a combination of independent drivers (X_k) which result in a good fit and also minimises the error.

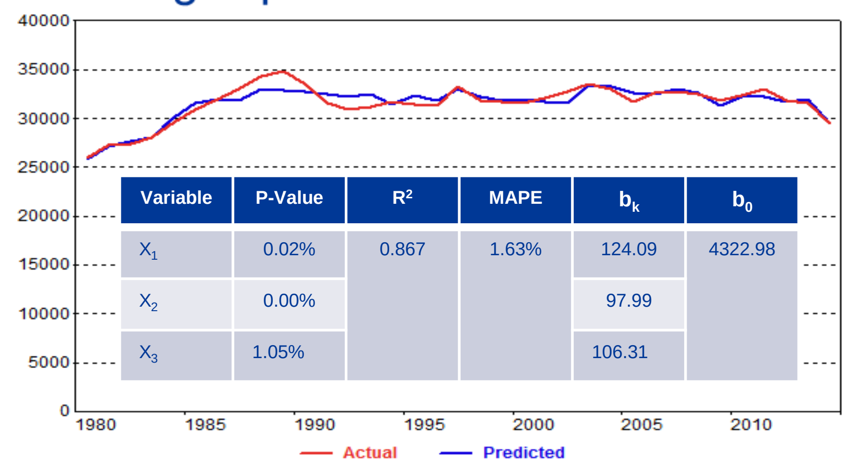
REGRESSION MODELLING

- Econometric regression models are used to forecast electricity consumption of Residential, Mining, Agriculture, Traction and Manufacturing & Commerce.
- $Y_t(GWh) = b_0 + b_1 X_{1t} + \mathrm{b}_2 \mathrm{X}_{2t} + \cdots + \mathrm{b}_k \mathrm{X}_{kt} + \mathrm{e}_t$
  - Y_t = Sales (GWh) in year t (dependent variable);
  - X_{kt} = economic or demographic drivers (independent variables);
  - b_0 = constant term;
  - b_1, b_2 …b_k = regression coefficients;
  - e_t = error term.
- The challenge is to obtain the best combination of drivers which produce an equation that will predict the history as close as possible, and satisfy some of the important statistical measures. The resulting equation is then used as a basis to predict future values of Y.

Minimise the error, as measured by MAPE (20% or less considered good).

Good fit, as measured by R^2, (R^2 =1, perfect fit).

Statistically significant drivers, as measured by P-Value, (5% or less is acceptable).



Variable	P-Value	R^2	MAPE	b_k	b_0
X_1	0.02%	0.867	1.63%	124.09	4322.98
X_2	0.00%			97.99	
X_3	1.05%			106.31	

DEEP DIVE INTO REGRESSION MODELLING

- The development of regression models involves exploring alternative possible combination of economic drivers.
- We use the following statistical measures:
  - **R^2:** This statistic measures how successful the fit is in explaining the variation of the data. High R^2 values (close to 1.0) indicate that the model explains the majority of the variation of the dependent variable.
  - **Adjusted R^2**: To show the impact of changes in the number of independent variables used in a model, an adjusted R-squared is used; therefore, the explained variation can be compared with the same dependent variable and different numbers of independent variables.
  - **MAPE**: Mean Absolute Percentage Error (MAPE) is the average of the absolute values of the percentage residuals. Any value below 20% is generally considered to be good enough.
  - **P-Value:** The statistical significance of the explanatory variables used in the model is measured by a P-Value. A variable with a P-Value of more than 5% is discarded.
  - **T-Statistic:** The T-statistic also measures statistical significance of the explanatory variables used in the model. A t-statistic (ignoring negative signs) of at least 2.0 would be required for a 95 percent level of confidence and 1.5 for a 90 percent level of confidence, depending upon the number of observations and variables used in the model.

ECONOMIC AND DEMOGRAPHIC DRIVERS

MINING	RESIDENTIAL	MANUFACTURING AND COMMRECE	AGRICULTURE	TRACTION
• Coal production index • Gold ore treated • PGM Production Index • Iron Ore Production Index	• Population • GDP	• Population • Manufacturing Production Index	• Log(FCEH)	• Coal Production Index • Iron Ore Production Index • Household Numbers

Typical OUTPUT

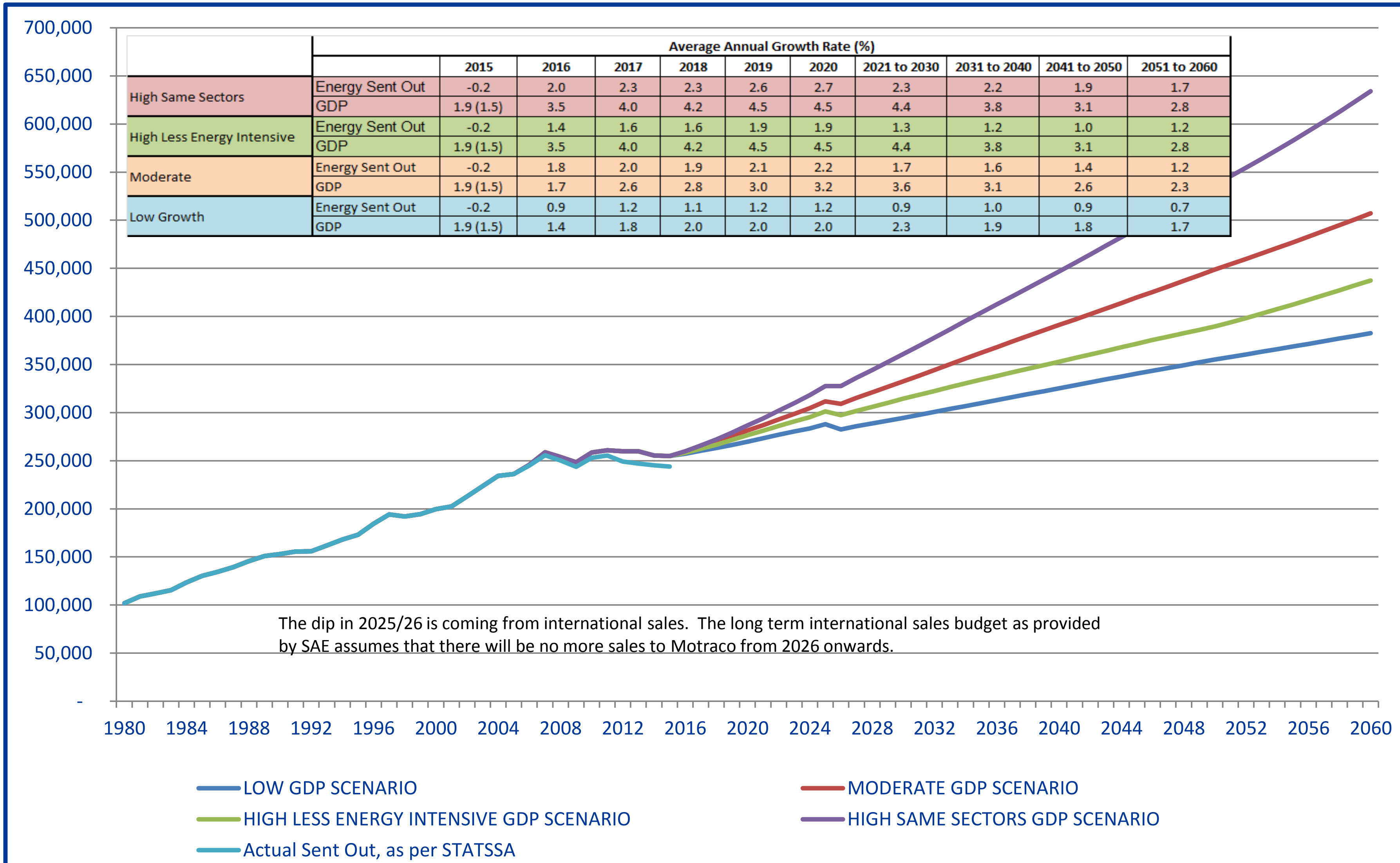
Mining Regression Model

Variable	Coefficient	StdErr	T-Stat	P-Value	Units	Definition
CONST	8012.335	2284.676	3.507	0.14%		Constant term
BaseCorrection.StartingYear	892.011	1155.790	0.772	44.63%		
EconomicData_HighSameSectors.Gold_ore_treated	86.930	14.166	6.137	0.00%	Mt	
EconomicData_HighSameSectors.Coal_production	150.601	31.314	4.809	0.00%	Index	
EconomicData_HighSameSectors.Iron_ore_production	52.974	14.439	3.669	0.09%	Index	
EconomicData_HighSameSectors.PGM_production	14.355	19.177	0.749	45.99%	Index	
Model Statistics			Forecast Statistics			
Iterations	1		Forecast Observations	0		
Adjusted Observations	36		Mean Abs. Dev. (MAD)	-		
Deg. of Freedom for Error	30		Mean Abs. % Err. (MAPE)	-		
R-Squared	0.867		Avg. Forecast Error	-		
Adjusted R-Squared	0.844		Mean % Error	-		
AIC	13.80		Root Mean-Square Error	-		
BIC	14.06		Theil's Inequality Coeffici	0.0000		
F-Statistic	38.994		-- Bias Proportion	0.00%		
Prob (F-Statistic)	0.0000		-- Variance Proportion	0.00%		
Log-Likelihood	-293.45		-- Covariance Proportion	0.00%		
Model Sum of Squares	164817,858.76					
Sum of Squared Errors	25360,334.32					
Mean Squared Error	845,344.48					
Std. Error of Regression	919.43					
Mean Abs. Dev. (MAD)	645.13					
Mean Abs. % Err. (MAPE)	1.98%					
Durbin-Watson Statistic	1.179					
Durbin-H Statistic	#NA					
Ljung-Box Statistic	10.62					
Prob (Ljung-Box)	0.0594					
Skewness	0.403					
Kurtosis	2.867					
Jarque-Bera	1.001					
Prob (Jarque-Bera)	0.6062					

THEN WE DO THE TOP DOWN FORECAST

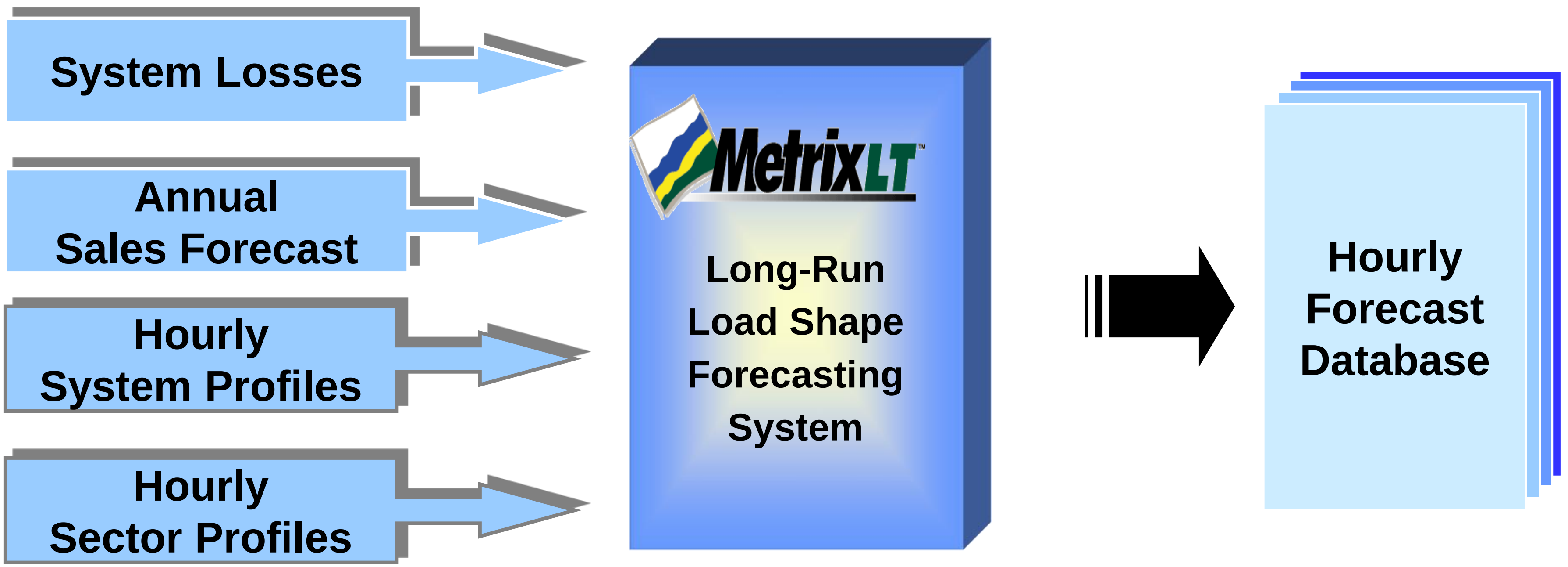
- A top down approach is a high level method which uses a combination of intensity and GDP to calculate the energy sent out.
- The results obtained from this method are used as a check against the results obtained from the bottom up approach.
- $Sent\ Out\ (GWh) = Intensity\ \left(\frac{kWh}{R}\right) \times GDP(Rm)$
- Once the total sent out it known, transmission losses and distribution losses are then systematically deducted from the sent out to arrive at the total sales.
- The total sales are then compared with the sum of the individual sector sales obtained by regression models. If needs be, adjustments are then made to individual sector sales to get an alignment of the two totals.

TIPICAL OUTPUT - RSA SENT OUT



		Average Annual Growth Rate (%)									
		2015	2016	2017	2018	2019	2020	2021 to 2030	2031 to 2040	2041 to 2050	2051 to 2060
High Same Sectors	Energy Sent Out	-0.2	2.0	2.3	2.3	2.6	2.7	2.3	2.2	1.9	1.7
	GDP	1.9 (1.5)	3.5	4.0	4.2	4.5	4.5	4.4	3.8	3.1	2.8
High Less Energy Intensive	Energy Sent Out	-0.2	1.4	1.6	1.6	1.9	1.9	1.3	1.2	1.0	1.2
	GDP	1.9 (1.5)	3.5	4.0	4.2	4.5	4.5	4.4	3.8	3.1	2.8
Moderate	Energy Sent Out	-0.2	1.8	2.0	1.9	2.1	2.2	1.7	1.6	1.4	1.2
	GDP	1.9 (1.5)	1.7	2.6	2.8	3.0	3.2	3.6	3.1	2.6	2.3
Low Growth	Energy Sent Out	-0.2	0.9	1.2	1.1	1.2	1.2	0.9	1.0	0.9	0.7
	GDP	1.9 (1.5)	1.4	1.8	2.0	2.0	2.0	2.3	1.9	1.8	1.7

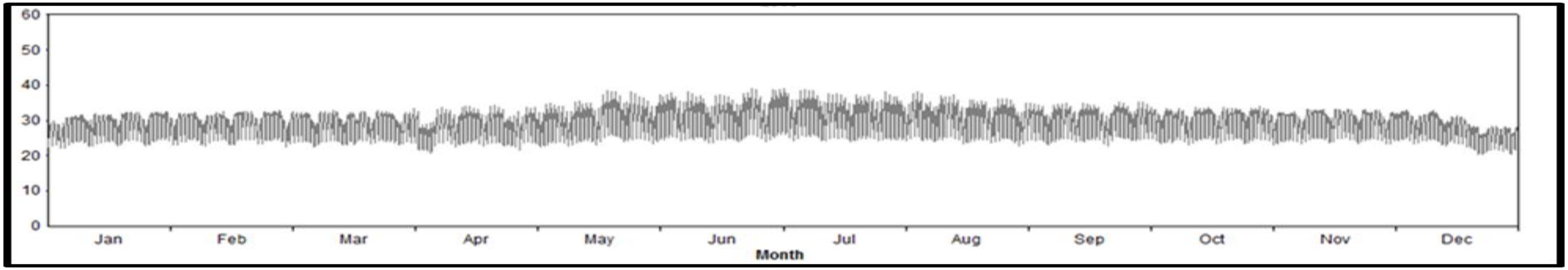
Process of Developing Long Term Hourly Forecast

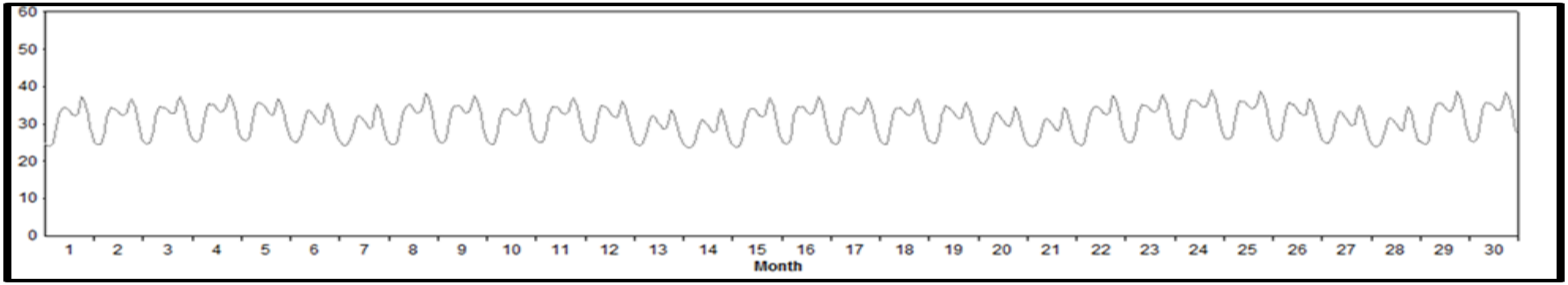


How do we compute the hourly peak demand

To obtain 8760 forecast from annual energy, peaks and shapes.

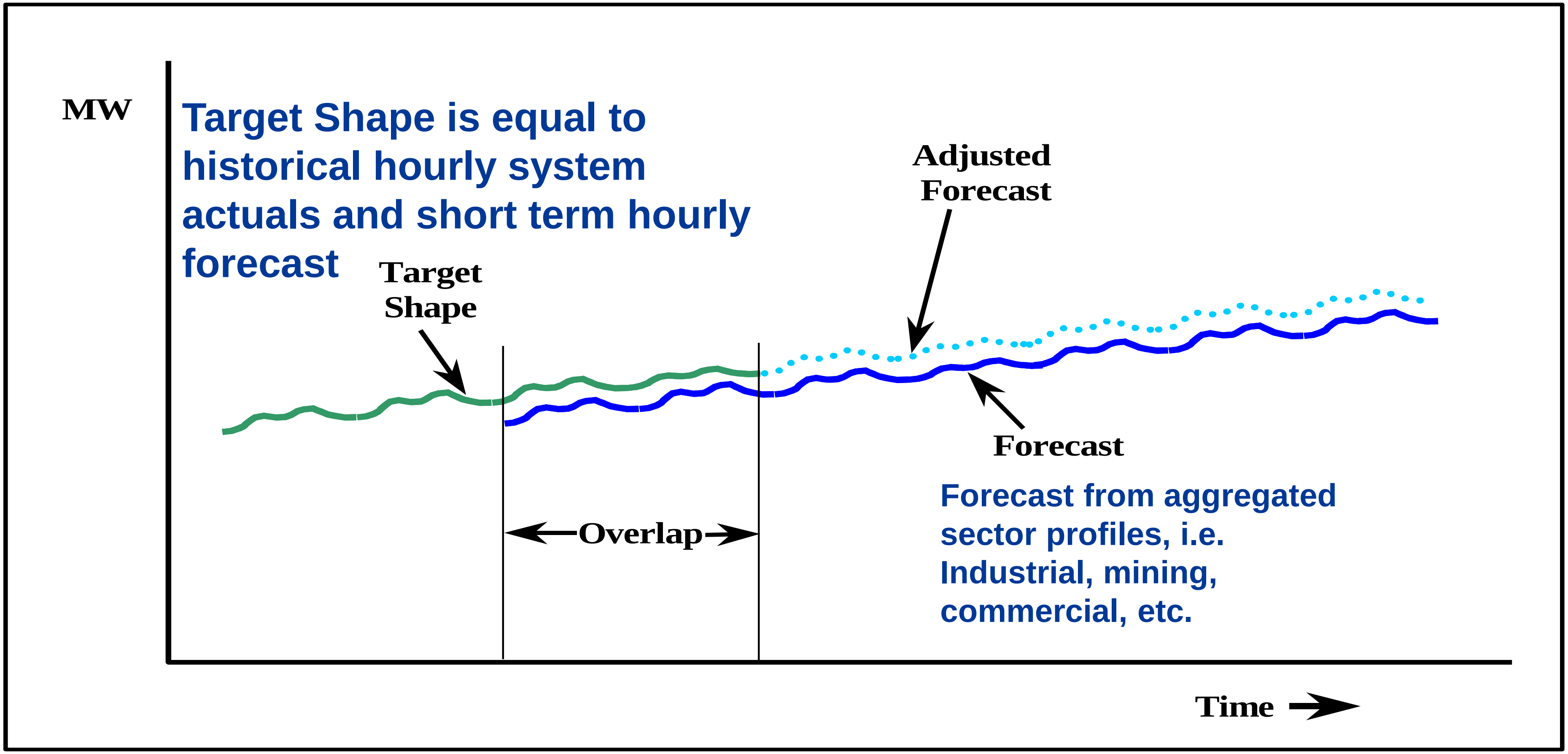
Hourly Forecast for Power System Planning





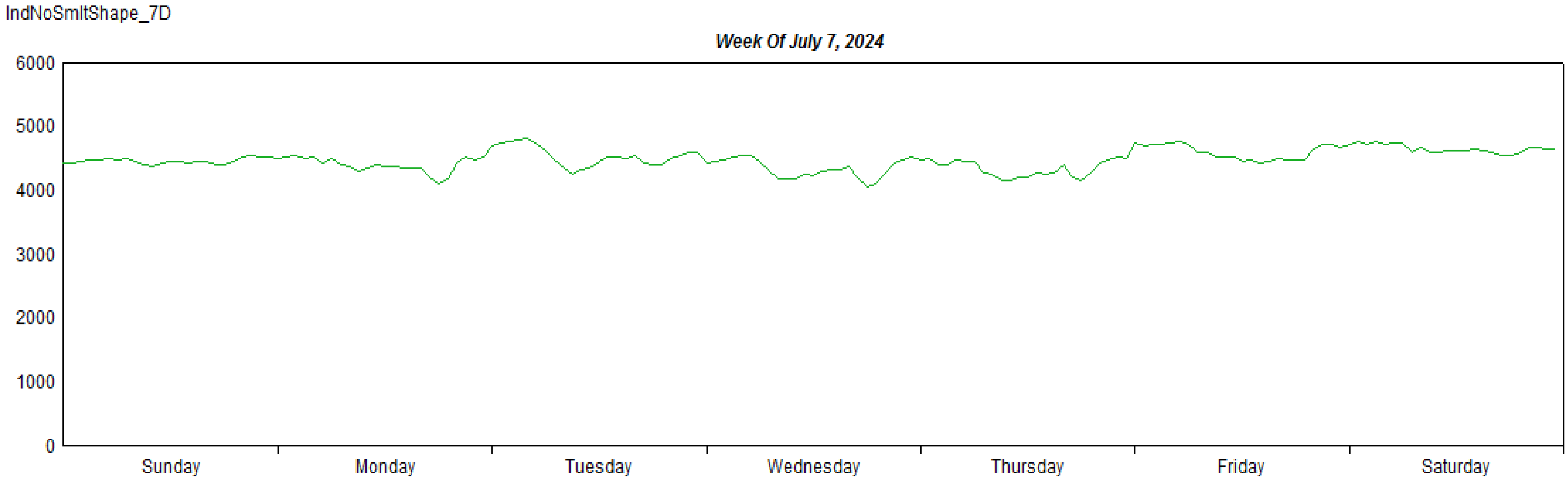
- We use sector profiles reconciled with the system profile to obtain an hourly shape forecast that reflects the changing mix of electricity consumption structure.

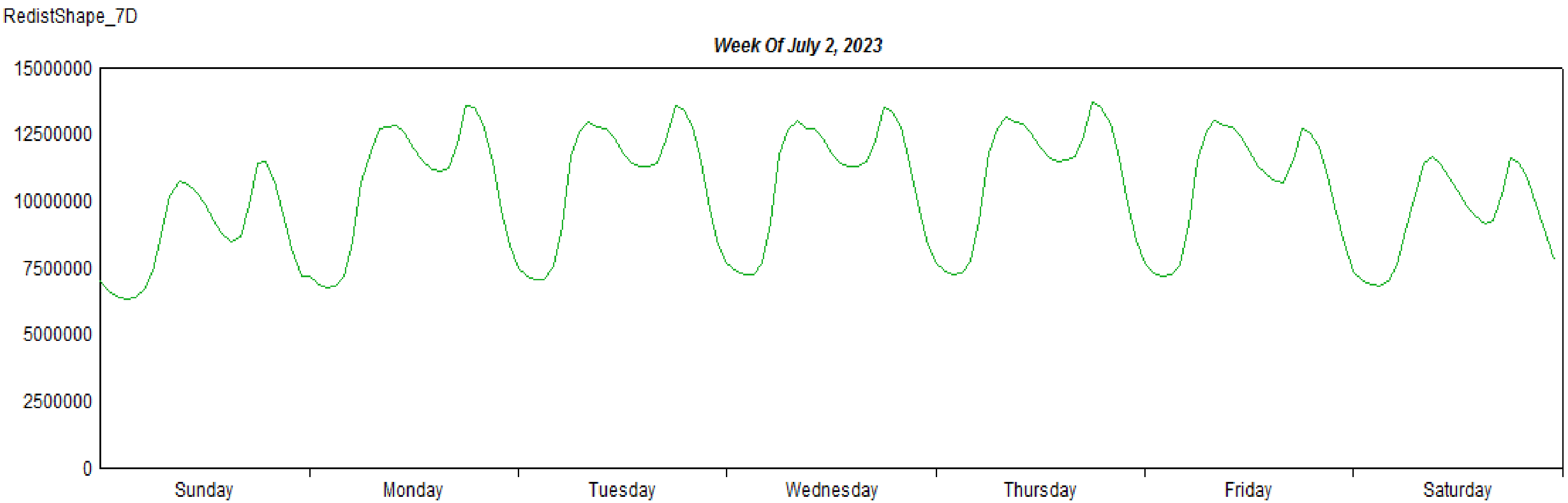
How we solve Forecasting Problems: Basic Schematic Representation of the Methodology



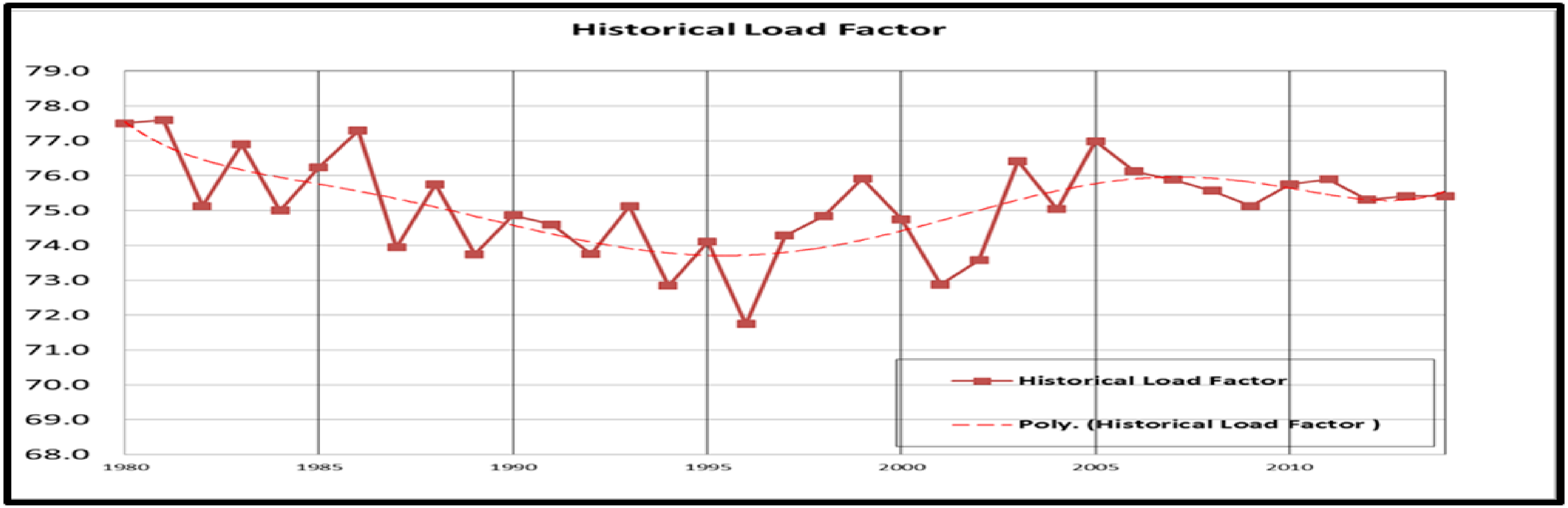
Target Shape and Moderate Sector Forecast

Typical load factor profiles used in the modelling



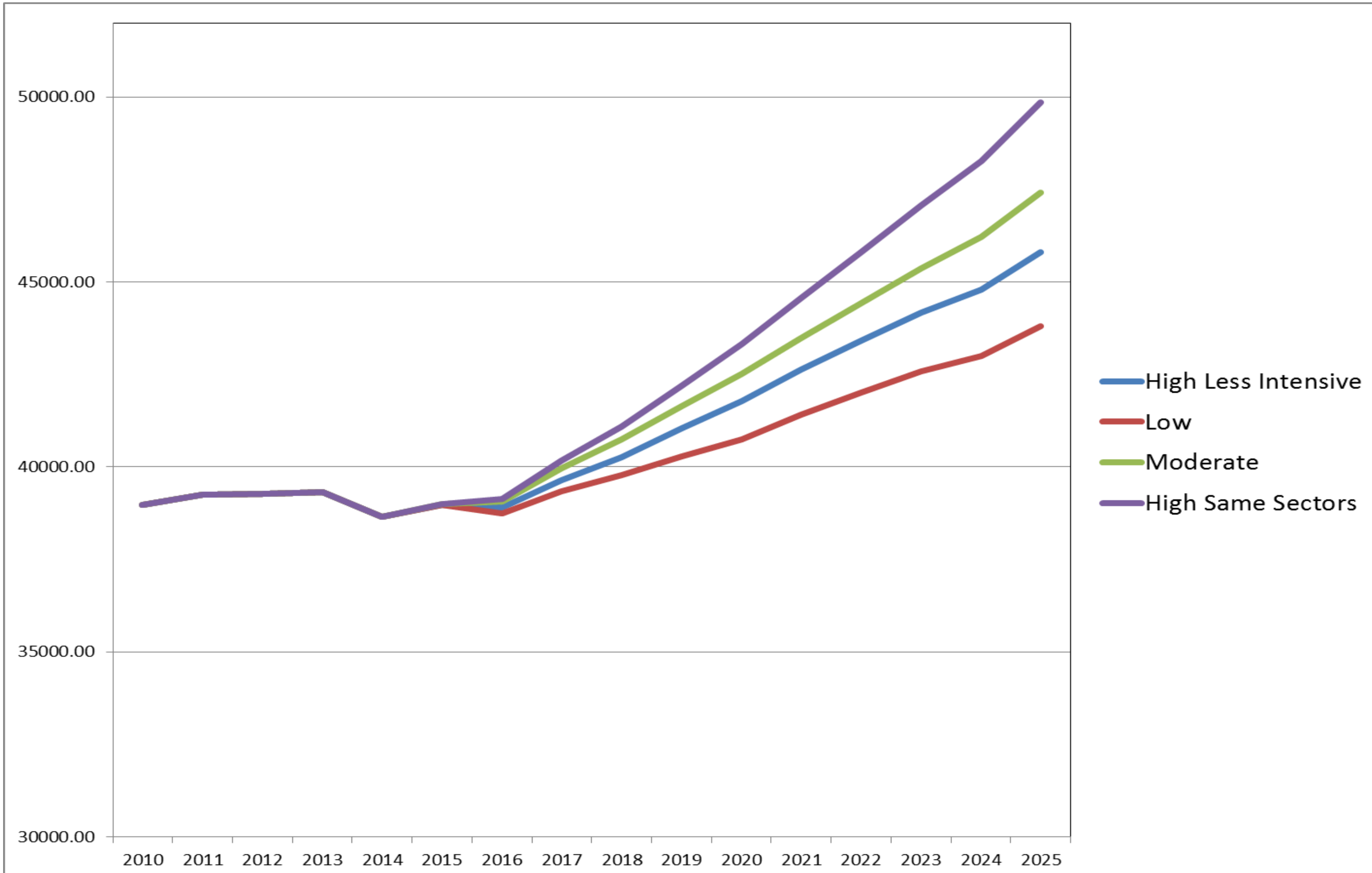


Compute System load factor as a sanity check



- Historical RSA has a high load factor due to the high load factor sectors such as Industry dominating in the customer mix of the energy forecast.
- For long term forecasting we analyse the load factor going forward as it gives us an indication of whether the load is becoming peakier or flatter. This helps us to indicate which sectors are the cause of the change of the overall profile.

Typical Results of Peak Demand Forecast (MW)



Thank you