PAPER NO. : **EMC/RCP/19/2005/CP07**

SUBJECT : **REVIEW OF THE SCOPE OF THE STUDY ON ACCURACY OF THE VERY SHORT TERM LOAD FORECAST**

FOR : **DECISION**

PREPARED BY : **SHASHANK SWAN**

VETTED BY : **PAUL POH LEE KONG**
SVP, MARKET ADMINISTRATION

DATE : **3 MAY 2005**

Executive Summary

The RCP considered a study of the accuracy of load forecast at its 16th meeting on 2 November 2004. While the RCP noted that no urgent review of the forecast methodology was required, it tasked the TWG to review the scope of the study. Subsequently, the TWG, at its seventh meeting on 17 February 2005 proposed the following revised scope of the study:

1. Compare the scheduled gross generation from the Real Time Schedule (RTS) against actual gross generation from SCADA data for 2004.
2. Compare the accuracy of the loss factor used in the MCE with the physical losses for a 12 month period; and
3. Analyse the need to change the existing generation forecast methodology to that of directly forecasting the load.

The results of the comparison between RTS values and gross generation values from SCADA data showed a mean absolute percentage forecast error of 0.82% compared to the value of 0.93% obtained using metered IEQ/Auxiliary load. Also, although the EMC noted possible errors in the metering data, the average physical loss factor of 0.43% is quite close to the existing loss factor of 0.6% set in the MCE. Further analysis indicates that there seems to be no compelling reason to change the existing generation forecast methodology to directly forecasting the load.

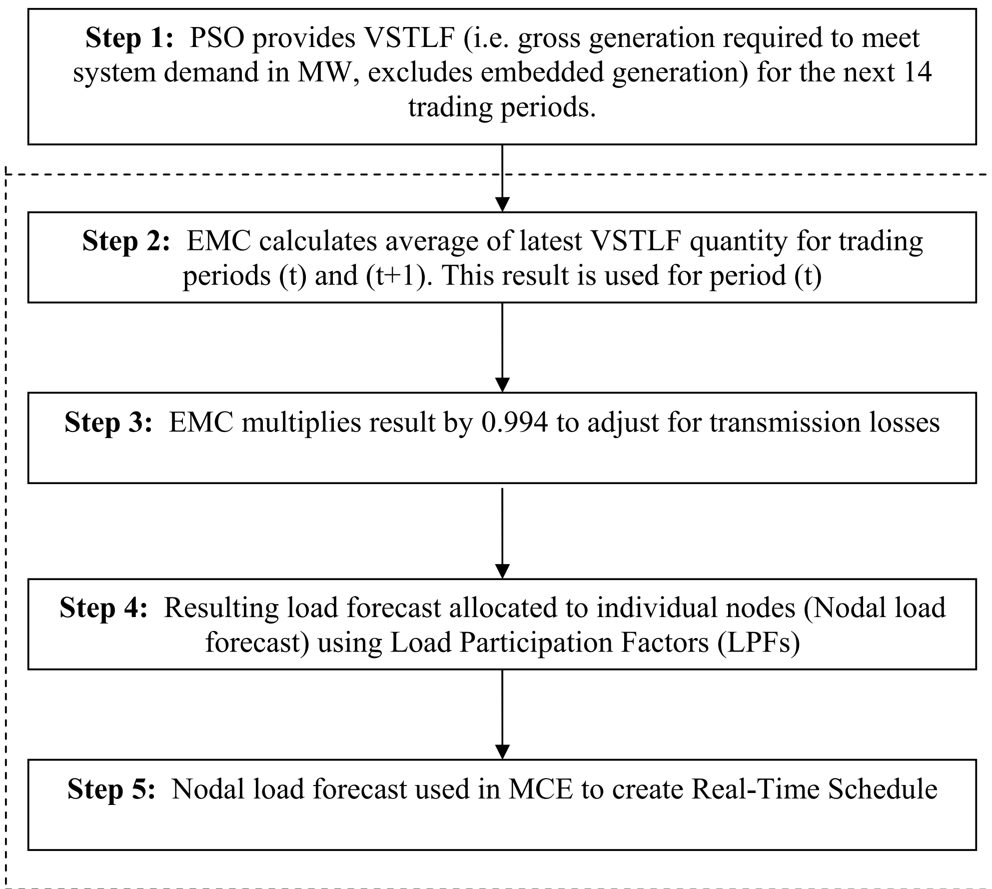
Therefore, the TWG recommends no change to use of existing generation forecast methodology as well as to the existing 0.6% loss factor used in the MCE. Further, the TWG proposes that the MSSL be requested to investigate the presence of negative losses inferred from the metering data for certain dispatch periods in 2004.

1. Introduction

The RCP had commissioned a *study of the accuracy of very short term load forecast* to address an issue raised by market participants (MPs) regarding the accuracy of the load forecast prepared by the PSO. The issue was raised as part of the yearly RCP work prioritisation plan for the year 2004. The study was discussed at the RCP at its meeting on 2 November 2004.

Briefly stated, the process of transforming a Very Short Term Load Forecast (VSTLF) to nodal load forecast, [which is then used as an input to generate real time schedule (RTS)] is shown in the figure below. VSTLF is a gross generation forecast (in MW) provided by PSO every period and is used in generating a RTS.

Figure 1



For that study, EMC used three months of data, ranging from January 2004 to March 2004, to assess the accuracy of the VSTLF (Step 1 in the figure above). The study compared the VSTLF quantity against the total metered generation quantity (IEQ) and measured a mean absolute percentage forecast error of 0.93%.

The RCP noted (at its meeting of 2 November 2004) that the results of the study were quite comparable to those of NEMMCO, Australia. While the results suggested no urgent need to improve the system-wide load forecast methodology, the RCP was of the view that the scope of the study could be reviewed for a more comprehensive study. The panel then tasked the TWG to carry out this review.

In response to the RCP's request, the TWG reviewed the scope of the study and asked EMC to conduct the following analysis:

1. Compare the scheduled gross generation from the Real Time Schedule (RTS) against actual gross generation from SCADA data for 2004.

2. Compare the accuracy of the loss factor of 0.6% with the physical losses for a 12 month period.

3. Analyse the need to change the existing generation forecast methodology to that of directly forecasting load.

On 12 April 2005, the TWG considered the results of the analysis presented in the following sections

2. Areas investigated

2.1 Compare the scheduled gross generation from the Real Time Schedule against actual gross generation form SCADA data for 2004

The TWG had noted that the data from the 1st Quarter of 2004 used in the earlier analysis may not be most appropriate for the study since this quarter has more holidays than other times of the year, which leads to less accurate load forecasting than usual.

The TWG further noted that the PSO's internal study of the accuracy of the VSTLF (when compared to SCADA generation data over 2003) indicated a mean absolute percentage error (MAPE) of about 0.6%. The TWG remarked that EMC's initial study had compared VSTLF quantities against the metered IEQ quantities plus auxiliary load. Thus the methodology used is likely to incorporate measurement error between two different ways of gathering data (EMS data versus the metering data). Therefore, the TWG decided that it would be more appropriate to compare scheduled gross generation from the RTS against the actual gross generation recorded in the SCADA data. As the scheduled generation in RTS is determined using the VSTLF (which itself is determined using the SCADA measurement), the TWG had agreed that such a study would be more appropriate to measure the accuracy of the load forecast.

Methodology

The following methodology was used to perform this analysis. The data used in the analysis ranges from 1 January 2004 to 31 December 2004.

1. The SCADA values (i.e. actual gross generation) provided by the PSO are average MW over the entire period T.

2. The RTS values (i.e. forecasted gross generation) produced by the market clearing engine (MCE) are *end of period* scheduled target MW for period T.

3. Therefore, the RTS values should first be converted to average values for correct comparison with SCADA values. This can be done by taking a simple average of start and end points of a period T in MW. This means taking an average of RTS values for periods T-1 and T (since RTS values for period T-1 is the end of that period value which is the same value for the start of the period value for period T).

4. Next step is to find the difference (i.e. RTS value – SCADA value). This provides the difference in the forecasted gross generation and the actual gross generation.

5. Finally, divide the difference obtained in the previous step by SCADA value (i.e. actual gross generation) to establish a percentage error for each period.

$$\text{Hence, Percentage error} = \frac{\text{(RTS-SCADA)} \times 100}{\text{SCADA}}$$

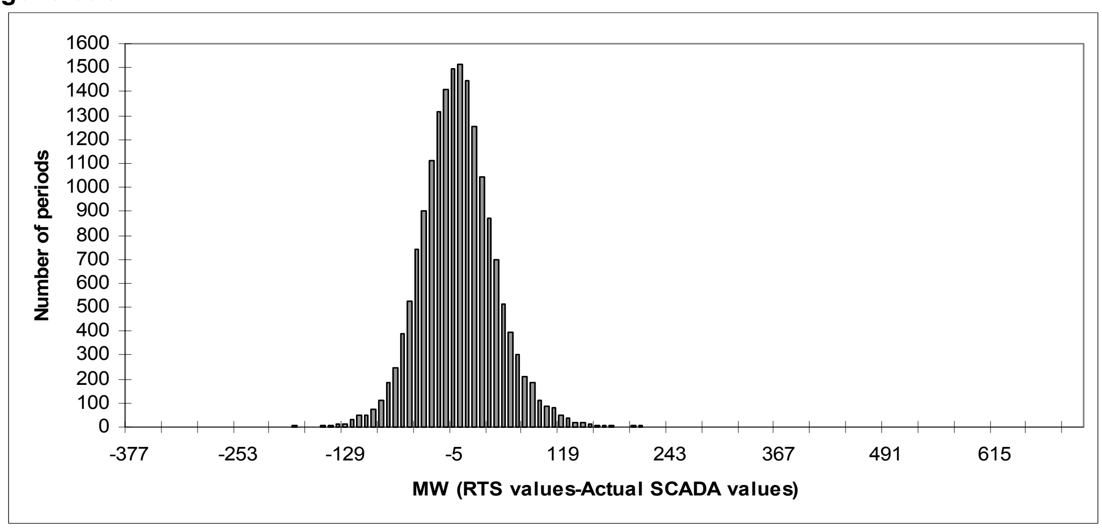
One must note that using this methodology has the following limitations

- The averaging process for RTS in step 3 assumes linear ramping in a period. However, in reality, the ramping between the start and end points of a period is likely to be non-linear. Thus, simple averaging may not produce the best results. However, given the limitation on data available to EMC for a period (i.e. values are available only for the start and end points of a period), this is the most suitable approach for this analysis.

- As noted in the earlier analysis, RTS scheduled values for a period T are produced by averaging VSTLF values for periods T and T+1. Further, the averaging process in step 3 produces an average of RTS values for periods T and T-1. Hence, this double averaging of two RTS values (i.e. values of the start and end of a period) is likely to introduce some errors in the final average MW for period T. For example, if the VSTLF values for periods T-1, T and T+1 are 5000, 5100, 5050 respectively, then the final average for comparison using this methodology for period T will be 5062.5 MW.

Results

Figure 1: Histogram of Difference in Forecasted gross generation and Actual gross generation



In Figure 1, the x-axis represents MW difference between the RTS value and actual SCADA value for total dispatched generation while the y-axis represents total number of periods for which a particular MW value was observed. For example, there are 1493 periods where the difference in RTS and actual SCADA is -5 MW. The histogram plots the error for all trading periods over the entire year of 2004. Please refer to Table 1 in Annex 1 for more details.

The analysis produced the following results:

Average	-3 MW
Standard Deviation	43.49
Median	-5 MW
Average Absolute	33 MW
Mean Absolute Percentage Error[1]	0.82%

[1] The mean absolute percentage error is the average of all the percentage errors, taken without regard to sign. This is a relatively standard measure of forecasting accuracy and avoids the problem associated with simple

Further, it was observed that only three periods had an error exceeding -210 MW and only two periods with errors above 210 MW. The result indicates a tight spread of all observations around the mean of -3 MW. On average, the VSTLF under-predicted actual total generation output by 3 MW.

The mean absolute percentage error, a statistic that is commonly used to measure forecast accuracy, is 0.82%. This means that the average absolute forecast error is 0.82% of the actual system load.

Despite the limitations of this methodology, this is a better result compared to the previous study done by EMC. The results here not only indicate that the forecast is more accurate than previously thought (the previous study showed a mean absolute percentage error of 0.93%), but the results are also more reliable due to the improved methodology used.

2.2 Assessing physical losses in the power system

The MCE uses a fixed loss factor of 0.6% to convert the gross generation forecast to a load forecast. Any difference between actual physical losses and this number will affect the accuracy of the nodal load forecast. This, in turn, could affect prices. Hence, it is important that the assumed losses match actual physical losses as closely as possible.

The current fixed loss factor of 0.6% of system load was derived using modeled power flows and cross checked with PSO using its EMS system.

Methodology

The actual physical losses in the transmission system can be calculated and compared against the fixed loss factor used by the MCE. This comparison will show how accurately modeled losses match actual physical losses. The physical losses can be measured by comparing metered injection into the grid with metered off-take using the following formula:

Loss Q in MWh = (Injection in MWh – Withdrawal in MWh) +/- Intertie flows in MWh
= (∑IEQ) - (∑WEQ) +/- Intertie flows
= (∑ IEQ + Intertie import) – (∑WEQ+ Intertie export)

Therefore, loss factor C = (Loss Q / IEQ) x 100

By using this formula, a data set can be generated for a suitable duration (e.g.12 months for all the periods) and measured losses can be compared with the fixed loss factor of 0.6%.

Results

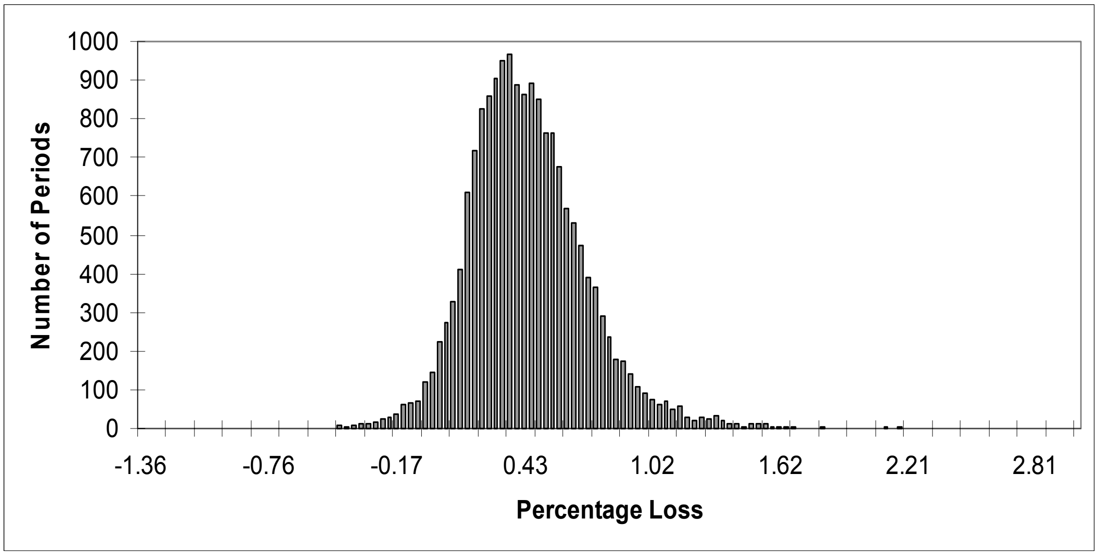
For this analysis, the most recent 12 month period was chosen, i.e. from 1 March 2004 to 28 February 2005. It was found that there were 654 periods where the losses were negative (this means that gross withdrawal exceeds gross injection[2], after adjusting for intertie flows). The average negative loss was found to be 4 MW with a maximum of about 56 MW. Such observations highlight a possible anomaly in the metering data received by the EMC, which should be investigated further by the MSSL.

averages where positive and negative errors cancel each other out, giving a potentially misleading picture of forecast accuracy.

[2] However, as the tolerance allowed for accuracy of various classes of meters under the Metering Code is +/- 0.5% to +/- 2.5%, this tolerance level could possibly explain the negative losses observed.

Consequently, the analysis presented here does not provide any firm conclusion but only serves as a guide to the accuracy of the fixed 0.6% loss factor. The analysis shows the following result

Figure 2: Histogram of physical losses as percentage of IEQ



The x-axis in Figure 2 represents percentage loss factor C in the system for the period between 1 March 2004 and 28 February 2005. The y-axis represents the number of periods a specific loss factor occurred during this time. For example, there were 965 periods with a percentage loss of 0.36%. Please refer to Table 2 in Annex 1 for more details.

The analysis produced the following results:

Mean	0.43%
Standard deviation	0.22%
Median	0.40%

Comparing the mean (0.43%) with the fixed loss percentage of 0.6% for a 5000 MW demand, the difference in losses for the entire system is only 8.5 MW. To assess accuracy of loss factor more reliably, it is proposed that the MSSL investigate the negative losses explained earlier. Subsequently, the following course of action should be adopted.

1) If the MSSL is able to rectify the negative losses issue, then a revised data set without negative losses can be used to repeat the analysis performed under this section and a conclusion can then be drawn.

2) If, however, the MSSL is unable to rectify the negative losses issue, then it is proposed that the existing fixed loss factor of 0.6 % be maintained. This is because no firm conclusion can be drawn about the variance of physical losses from the fixed loss factor used because of the presence of negative losses. In addition, although the TWG decided that it is not suitable to compare MCE model losses with physical losses as such a comparison does not take into account losses due to reactive power, the mean for of the MCE calculated losses is 0.56% for the month of December 2004. There is thus no reason to believe that fixed 0.6% is not suitable.

2.3 Is using generation forecast methodology more appropriate for NEMS than using directly forecasted load?

It was noted earlier that some jurisdictions (such as NEMMCO, Australia) directly forecast the system demand at the zonal level (rather than forecasting the generation needed). Therefore taking into account the local conditions, the TWG had decided that this area should be investigated further to assess the need for a change in methodology.

Theoretically, forecasting load directly is a more efficient approach. This is because a direct load forecast would need less manipulation of data to produce a dispatch schedule. However, the accuracy of the current forecast methodology (mean absolute percentage error 0.82%) is good and compares favorably with other jurisdictions such as NEMMCO, NE ISO and New Zealand. It uses about 35 points for data collection and forecasting. If load were to be forecasted directly, then load would need to be forecast using approximately 400 points in the Singapore system. Auxiliary load would need to be added to directly forecasted load to ensure sufficient gross generation dispatch.

In addition, Singapore’s weather pattern is quite predictable and stable. The load pattern is also predictable and non-volatile. The variation in load from one period to another is not sudden but gradual. These conditions mean that a simple forecast methodology can deliver good forecast accuracy.

The existing method is less costly to operate than a direct load forecast as it adds up generation of some 35 units as opposed to adding up power flow values of some 400 load points in Singapore. Implementing a system to forecast load directly using the SCADA system would have significant cost implications[3]. Changing the methodology to directly forecast the load would not only involve significant system modification for the PSO[4] and EMC but could be more prone to errors than the existing system due to its higher complexity. Forecasting load directly would involve collection and manipulation of a larger set of data (400 as compared to 35 data points).

Therefore, implementing a direct load forecast method would involve incurring significant costs. The benefits of any change are uncertain as it is unclear how much a nodal load forecast would improve the forecast accuracy. Given that the accuracy of the current forecast methodology (mean absolute percentage error of 0.82%) performs well and compares favorably with other jurisdictions[5], there is no compelling reason to change the current forecast methodology.

3. TWG Deliberations

The first area of the study was to assess the overall accuracy of the existing forecast methodology by comparing gross generation scheduled in RTS with the actual gross generation recorded by SCADA measurement for each period of 2004. Although the methodology employed had some limitations, the TWG noted that given the data available, it was the most suitable approach. The results not only ratified results of the previous study done by EMC but also compared very well with PSO’s internal study on accuracy of load forecast (which was noted as approximately 0.6%). Thus, regarding RCP’s query about the impact of a 1% forecast error on system security, the TWG was of the view that there was no adverse impact on system security as sufficient regulation is scheduled in every period.

---

[3] A ball park estimate by PSO for enhancing EMS systems to implement this is in the range of $1 million
[4] System capability to forecast auxiliary loads will have to be newly built.
[5] For NEMMCO (5 minute dispatch) it is 0.45%- 0.8%; for NEISO (1 hour dispatch) it is 2%-3% and for NZEM (30 minute dispatch) it is upto 3%

In the second area of the study, the physical loss in the transmission system was compared with the fixed loss factor of 0.6% used in the MCE. The TWG noted that no firm conclusion could be drawn from the results because of the presence of negative losses. The TWG recommends that the MSSL investigate the negative losses issue. It further recommends that the RCP agree that the current factor of 0.6% be maintained unless the EMC is able to repeat the analysis with revised metering data and subsequently recommends a change in the loss factor.

Finally, the last area of the study assessed the need (in the Singapore context) for changing the current generation forecast methodology to a direct load forecast. The TWG noted that implementing a direct load forecast would involve significant costs, without the certainty of improved accuracy. Therefore, given that the accuracy of the current generation forecast methodology is relatively high and compares favourably with other jurisdictions, the TWG was of the view that there is no compelling reason to change the current generation forecast methodology. As such, the TWG recommends that there be no change.

Conclusion

In summary, the TWG was satisfied with the methodology, analysis and results of the expanded study.

4. TWG Recommendations

TWG recommends that the RCP

1. **Agree** that the analysis presented in this paper for the accuracy of load forecast ratifies the results of previous study conducted by the EMC;
2. **Agree** that the current fixed loss factor of 0.6% be maintained unless EMC is able to repeat this analysis and recommends otherwise;
3. **Agree** that the MSSL be requested to investigate the negative losses issue ; and
4. **Agree** that there is no need to change the current methodology of using forecast of generation to forecast of load.

Annex 1

Table 1

MW (Diff)	Frequency	MW (Diff)	Frequency	MW (Diff)	Frequency
-377	1	36	869	450	0
-369	0	45	699	458	0
-361	0	53	511	466	0
-352	0	61	395	474	0
-344	0	69	304	483	0
-336	0	78	213	491	0
-327	0	86	185	499	0
-319	0	94	113	507	0
-311	0	102	87	516	0
-303	0	111	78	524	0
-294	0	119	50	532	0
-286	0	127	37	540	0
-278	0	135	19	549	0
-270	0	144	17	557	0
-261	0	152	15	565	0
-253	1	160	9	574	0
-245	0	168	8	582	0
-237	1	177	8	590	0
-228	0	185	2	598	0
-220	1	193	2	607	0
-212	0	202	4	615	0
-203	3	210	4	623	0
-195	0	218	1	631	0
-187	4	226	0	640	0
-179	2	235	0	648	0
-170	3	243	0	656	0
-162	2	251	0	664	0
-154	5	259	0	673	1
-146	5	268	0	681	0
-137	10	276	0	689	0
-129	15	284	0	698	0
-121	32	292	0	706	0
-113	47	301	0		
-104	47	309	0		
-96	75	317	0		
-88	109	326	0		
-79	187	334	0		
-71	246	342	0		
-63	387	350	0		
-55	524	359	0		
-46	743	367	0		
-38	903	375	0		
-30	1115	383	0		
-22	1313	392	0		
-13	1406	400	0		
-5	1493	408	0		
3	1516	416	0		
11	1446	425	0		
20	1252	433	0		
28	1041	441	0		

Table 2

%age loss	Frequency	%age loss	Frequency	%age loss	Frequency
-1.36	1	0.23	825	1.82	3
-1.33	0	0.26	858	1.85	1
-1.29	0	0.29	906	1.88	1
-1.26	0	0.33	952	1.92	1
-1.23	1	0.36	965	1.95	1
-1.19	0	0.39	890	1.98	0
-1.16	0	0.43	863	2.01	1
-1.13	0	0.46	892	2.05	1
-1.09	0	0.49	852	2.08	0
-1.06	0	0.53	762	2.11	3
-1.03	0	0.56	762	2.15	1
-0.99	0	0.59	675	2.18	3
-0.96	1	0.63	567	2.21	1
-0.93	0	0.66	531	2.25	0
-0.90	0	0.69	474	2.28	0
-0.86	1	0.72	389	2.31	0
-0.83	0	0.76	366	2.35	0
-0.80	1	0.79	290	2.38	0
-0.76	1	0.82	237	2.41	0
-0.73	2	0.86	178	2.44	0
-0.70	0	0.89	176	2.48	0
-0.66	1	0.92	140	2.51	1
-0.63	0	0.96	109	2.54	0
-0.60	2	0.99	91	2.58	0
-0.56	1	1.02	73	2.61	0
-0.53	0	1.06	64	2.64	0
-0.50	1	1.09	69	2.68	0
-0.47	1	1.12	49	2.71	0
-0.43	10	1.15	58	2.74	0
-0.40	5	1.19	28	2.78	0
-0.37	9	1.22	22	2.81	0
-0.33	13	1.25	28	2.84	0
-0.30	12	1.29	23	2.87	0
-0.27	17	1.32	33	2.91	0
-0.23	23	1.35	20	2.94	0
-0.20	28	1.39	11	2.97	0
-0.17	36	1.42	12		
-0.13	63	1.45	3		
-0.10	68	1.49	12		
-0.07	71	1.52	11		
-0.04	120	1.55	11		
0.00	145	1.58	6		
0.03	224	1.62	4		
0.06	274	1.65	5		
0.10	329	1.68	6		
0.13	409	1.72	2		
0.16	612	1.75	2		
0.20	717	1.78	0		