RC25443 (IRE1401-019) January 28, 2014
Computer Science

IBM Research Report

Using a Big Data Analytics Approach to Unlock the Value of Refrigeration Case Parametric Data

Niall Brady, Paulito Palmes
IBM Research
Smarter Cities Technology Centre
Mulhuddart
Dublin 15, Ireland

John Walsh
Tesco UK

Research Division
Almaden – Austin – Beijing – Cambridge – Dublin - Haifa – India – Melbourne - T.J. Watson – Tokyo - Zurich

Using a Big Data Analytics Approach to Unlock the Value of Refrigeration Case Parametric Data

Niall Brady B.E. (Elec) M.Sc   IBM Smart Cities Research Dublin, Ireland
Paulito Palmes (Ph.D) IBM Smart Cities Research Dublin, Ireland
John Walsh M.Eng, Tesco UK and Ireland Energy Manager

Email: bradynl@ie.ibm.com, palmesp@ie.ibm.com, John.Walsh@tesco.ie

Abstract :

Within the retail sector, refrigeration accounts from between 30% to 60% of a store's total electrical energy budget **(Ref 1 Tassou et al 2011)**, where the most commonly used method in providing the necessary cooling energy to refrigeration cases is through multiplex direct expansion systems. Here all shop floor refrigeration cases use direct expansion air-refrigerant coils that are connected to banks of system compressors located in a remote machine room normally in the back or on the roof of the store, along with the supporting air-cooled condenser rooftop units. As a consequence of the highly regulated nature of the refrigeration process itself, driven by business legal requirements, coupled with the need to maintain proper food quality levels, underlying fridge control parametric data is as a result, both very rich in content and readily accessible. In data volume terms with individual cases being sampled typically every 5 minutes, this is estimated to generate upwards of 70-100 million discrete pieces of relevant control point data for a typical large size store, in a single year. Therefore this paper, through a series of examples from data taken for actual stores from a large retailer, explores the value (economically and technically) of acquiring, harvesting, and applying big data aggregated statistical approaches to this large data set, to help the domain experts to deepen their knowledge of actual refrigeration case behaviour. It is shown that the followon energy savings from the knowledge gained from this data analytics approach can be significant, with one such project alone, relating to a defrost policy change, singularly capable of delivering over 2.5% saving in overall store energy usage. **(Ref 2 : ComputerScope Article 2013)**. It is further presented, that through a set of followon developed engineering driven key performance indicators (KPI's) taken from this readily available parametric data set, that this new insight will not only allow for additional energy saving through realtime case anomaly detection, but also has the potential to positively impact on the direction of future maintenance support models, supplementing traditional preventative/reactive methods with cost effective data analytics driven predictive maintenance approaches.

Keywords : Big Data Analytics, Low Temperature Refrigeration, Defrost, Energy Saving

1.0 Introduction

The basis of this paper publication has come from a collaboration project between IBM Research Labs in Dublin and Tesco Regional Energy Management team where a feasibility study was commenced in early 2013 around the effectiveness of applying big data **(Ref 3 Ward et al 2013)** analytics on the energy demand side of Tesco's case refrigeration estate in Ireland and the UK. Tesco has, over the past number of years, invested heavily in their IT infrastructure in support of their business operations, and while this has yielded significant returns already on the energy management side, it was felt that they had still not yet exploited the full value from their comprehensive data warehousing efforts to date. Therefore, the initial project objective was to identify and target a specific high volume data set within Tesco's data model where a big data analytics methodolgy could be applied to demonstrate to Tesco the business value and savings impact of applying such an approach.

Given that it is estimated that refrigeration alone accounts for nearly 9% worldwide consumption of energy **(Ref 4 Bertoldi et al 2003**), where in the US alone this equates to a commercial refrigeration energy footprint of the order of 300 Twh per year **(Ref 5 Goetzler et al 2009)**, one can see the obvious value in focusing project efforts in this area. The other main reason for choosing refrigeration was the availability and richness of the underlying fridge parametric datasets brought about by not only the highly controlled nature of the refrigeration process itself, but also the additional need for mandatory archiving of fridge data in support of food safety EU legislation **(Ref 6 EU Food Standard)**. It is estimated that for a typical store, anything from between 70 million to 140 million discrete records are generated annually, when archived 5 minute fridge parametric data is taken into account.

And while there are already well established methods and standards for performance testing of Refrigeration Cases **(Ref 7 EU Fridge Test Standard)**, the demonstrated descriptive statistical approach taken here in this paper, through the presentation of a series of base level sample case refrigeration KPI's, offers an equally effective way in conducting ongoing performance monitoring of the fridge estate, while additionally achieving a significant positive impact on energy usage and maintenance support levels within the estate.

Furthermore, while the initial study focused on 8 stores covering the complete generic High Temperature (HT), and Low Temperature (LT) refrigeration case types, that subsequently generated over 500 million discrete records, this paper for simplicity focuses on selected examples within the store study group, and specifically on just one of these LT Freezer case types, namely the subgroup generically referred to as Full Glass Door Case (and within the figures labelled as FGDC). Confining the scope to a single subgroup has the additional benefit of allowing for proper peer-to-peer case and store-to-store performance comparisons given the similar nature and function of the LT cases within the store.

Equally, because this paper focuses on the output value of the big data aggregation exercise, there is no presented detail around the necessary data acquisition and data pre-processing steps of the project itself. Suffice to state that once the parametric raw data sources are properly located (or connection to the data source established), that the underlying time series data object set is verified and quality assured, and that the data processing sequence is appropriately scripted, the followon consumption and appropriate output KPI reporting of the initial raw data per store becomes a relatively quick exercise[1]

[1]Per KPI, a store's annualised 70 million per 5 minute records are fully preprocessed on an IBM X Series rack server with multicore Intel processors, within 10 seconds

The standard Parametric Data Object List chosen for the study that forms the basis of this sample KPI generation set and subsequent aggregation analysis is presented in Table 1 below.

KPI Description	Fridge Parametric Data Objects Required
KPI 1 Monthly Defrost Event Aggregation	Aggregation of daily defrost cycles based solely on DEFROST STATUS FLAG
KPI 2 Monthly Defrost Duration Estimate	Estimating of defrost duration based on DEFROST STATUS FLAG, AIR OFF (S4) temperature
KPI 3 Defrost Temperature Endpoint Attainment (using S5)	Determination of the Defrost End Condition based on DEFROST TERMINATION TEMPERATURE (S5)
KPI 4 Post Defrost Average Pull Down Duration Estimate	Estimating of post defrost Pull Down time using DEFROST STATUS FLAG , and AIR OFF (S4) temperature
KPI 5a Steady State Case S4 Monthly Average Monthly Temperature Control Performance	Use of aggregated AIR OFF (S4) temperature to establish proper case control behaviour, and help to identify anomalous cases
KPI 5a Steady State Case S3 Monthly Average Monthly Temperature Control Performance	Use aggregated AIR ON (S3) temperature to establish ongoing proper case behaviour , and help to identify anomalous cases

Table 1 : Sample Case Refrigeration KPI Summary

2.0 Analytics Methodology

The following section outlines the analytics approach taken through the presentation of a some sample KPI's and a series of illustrated examples that demonstrate the value and additional engineering insight gained through this KPI application when applied to the historical object data sets defined in Table 1, for some sample store datasets.

2.1 KPI 1 : Monthly Average Defrost Event Aggregation

***KPI 1 Monthly Average Defrost Event Aggregation** – This KPI is defined as the observed average number of daily defrost cycles within a month where defrost cycle is based on appropriate defrost status flag declarations within the provided standard case parametric set.*

The most basic and commonly used defrost control strategy for supermarket applications is the scheduled timed defrost. By aggregating all scheduled defrost events within a 12 month period for individual cases it is possible to quickly identify background defrost policy adherence and anomalous case behaviour. In Figure 1, in Store 1 below, examples of large variation in daily defrost averages can be seen to exist across cases, averaging from between 2 and 6 defrost cycles per day, prior to March 2013, from which point standardisation to 2 defrost cycles were adopted on an interim basis. Similarly for Store 2, in Figure 1, it can be seen the move to defrost cycle standardisation was executed in late 2012, where all cases went to a single defrost per day, and in fact has now finally been recommended as the standard defrost policy for all FGDC cases within the estate, again as a result of this project effort.

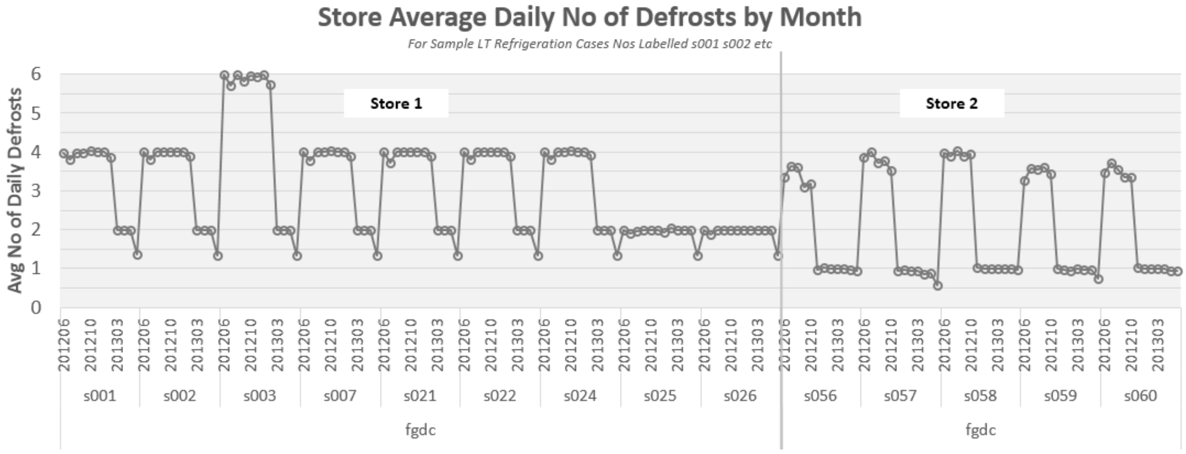

Figure 1: Monthly Defrost Events Average Summary Plot

2.1.1 KPI 1 : Estimating Monthly Defrost Policy Change Energy Saving Example

An additional advantage of using a KPI aggregation approach applied to the data set, is that when coupled with available actual instore energy submetering data, it is possible to both estimate and subsequently verify the savings potential of any changes made within the project. So taking the defrost event KPI and analysis as an example, it is possible to estimate the energy saving potential and subsequent impact of the unilateral LT Case defrost schedule policy change executed in Store 2 in November 2012, and as discussed in the previous section, and as observed in Figure 1 above.

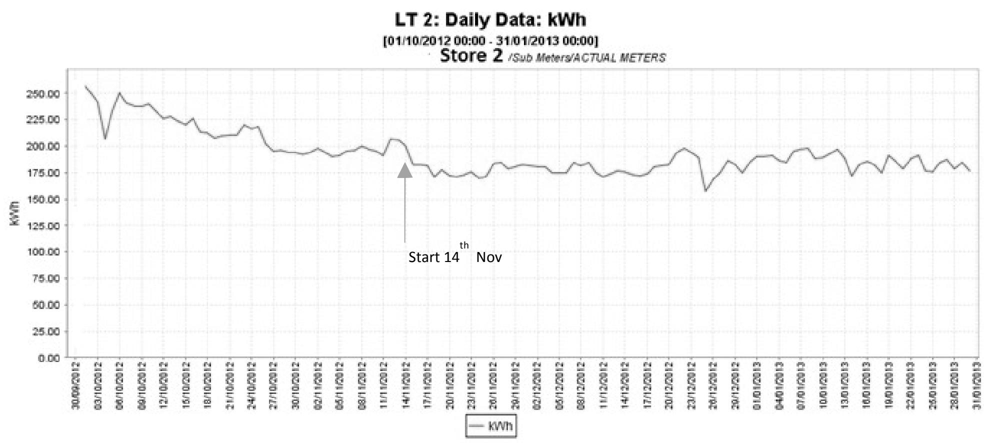

Figure 2(a) : Store 2 LT2 Compressor/Condenser Pack Metered Energy Snapshot

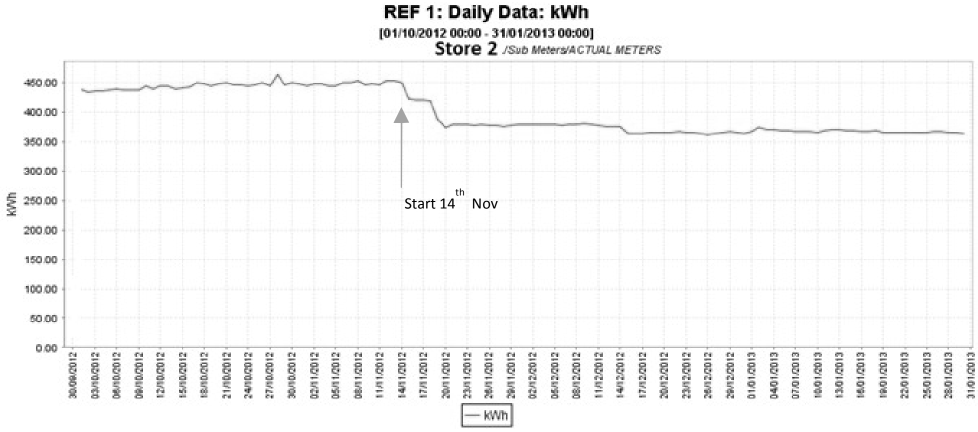

Figure 2(b) : Store 2 REF 1 Overall Refrigeration Case Metered Energy Snapshot

Figure 2 shows snapshots of the actual relevant daily store sub meter output plots around the defrost policy change period, with the relevant Supply Side Compressor/Condenser Pack labelled LT2 in Figure 2(a), and the relevant Demand Side case energy usage, labelled REF 1 in Figure 2(b), which meters all the case energy used within the store. In this Figure 2(b) it can be clearly seen the step change seen from the 14th of November which was exclusively due to the defrost electrical resistive heater load reduction on the LT cases, with the unilateral move from 4 to 1 daily defrost cycles.

Store 2 Summary	**Pre Defrost Policy Change** Monthly Average	**Post Defrost Policy Change** Monthly Average	**Policy Change Average Monthly Saving**	**Estimated Kwh Usage per Cycle**
Metered Indirect Supply Side Compressor Pack LT 2 Load (Kwh)	7,221	5,570	1,652	1.22[1]
Metered REF 1 Store Case Demand Side includes Resistive Heater LT Case Load (Kwh)	13,394	11,437	1,958	1.45[2]
Avg Monthly No of LT Defrost Cycles	1,117	442	675[3]	
Estimated Annualised Defrost Energy usage per LT case per defrost cycle (MWh)	**0.97**			
	Estimated	*Actual Metered*		
2013 Annualised LT Defrost Policy Change % Saving of Overall Store Energy Usage	**2.5%**	**4.7%**		

[1] *Not all the demand side compressor pack saving can be accredited to the defrost policy change alone, and other uncertainties including natural demand reduction in colder during winter months need to be considered therefore conservatively only 50% of observed saving is taken.*

[2] *REF 1 is metererd energy usage for complete store case estate but as seen in Figure 2 difference is exclusively attributed to the LT Pack defrost policy change spread over two packs LT1 and LT2, so 50% of observed saving is taken for LT2 cases.*

[3] *Move from 4 cycles to 1 cycle except for certain Freezer Room evaporators in the Pack that remained at 4 cycles post defrost policy change*

Table 2 : Defrost Policy Change Energy Savings Calculations

And so, it is possible, as is summarised in Table 2, to make energy saving estimates from the defrost policy change by combining the defrost cycle KPI aggregation (pre and post policy change) and relevant available underlying submeter data as previously presented in Figure 2. Furthermore, data driven estimates of energy usage per defrost cycle could now be established beyond previous published model driven defrost cycle energy usage estimates **(Ref 8 Fricke et al 2010)**, where the overall total (direct and indirect) annualised electrical energy usage footprint per defrost cycle for the store's LT 2.64 meter long cases, was estimated to be just under 1 Mwh per cycle per year.

And again from Table 2, it can be seen that by rolling up the subsequent energy savings from the defrost policy change to the store level, it is estimated to yield a conservative overall 2.5% reduction in store energy usage, and this compares favourably to the actual confirmed annualised store metered 2013 energy reductions of 4.5%, where separate additional energy saving initiatives executed within the year contributed to the overall impressive year-on-year energy savings.

2.2 KPI 2 : Monthly Average Defrost Duration Estimate

***KPI 2 Monthly Defrost Duration Estimate** – This KPI is based on defrost flag analysis and estimating the time period from defrost flag transition (time between transition defrost flag going 0 to 1 at start of defrost, and transition of defrost flag going 1 to 0)*

By defining this KPI, and attempting to aggregate defrost duration estimates over the months, it is possible to establish baseline case defrost durations, that allow not only for within store case-to-case comparison and detection of anomalous case behaviour, but also to detect store-to-store variation, as can be seen in Figure 3 in Store 1 and in Store 2, where significant differences in average defrost duration estimates are clearly apparent. Additionally among other valuable uses for this KPI, it is possible, by plotting monthly defrost durations aggregations, to detect defrost average duration drifts over time, where detecting such continued month-to-month average time degradation could be an early indicator of either a defrost heating problem or possible ice buildup conditions in the case evaporator coil.

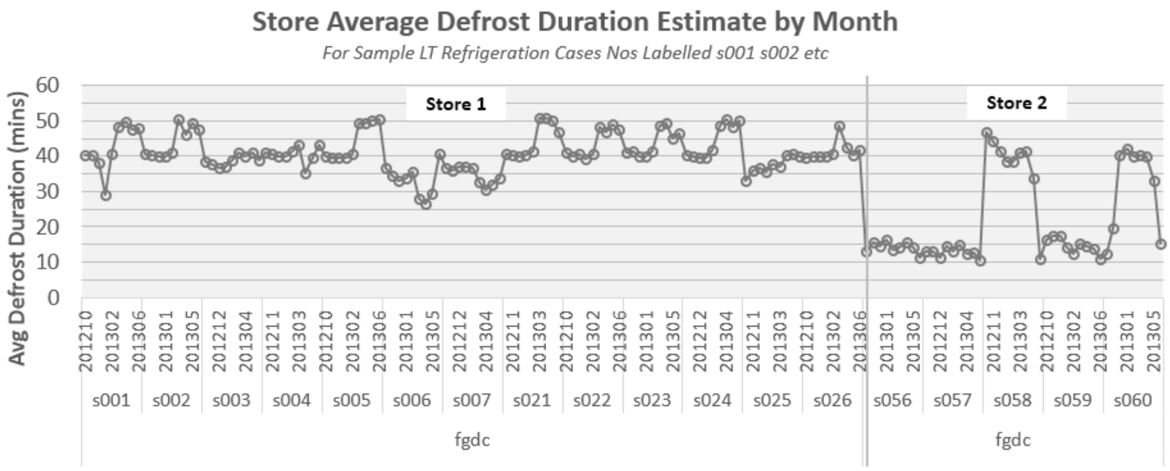

Figure 3 : KPI 2 Defrost Duration Summary Plot

2.3 KPI 3 : Defrost Termination Setpoint Attainment using S5 Temperature Sensor

KPI 3 Defrost Termination Setpoint Attainment *– This is an ongoing post project KPI development and is defined for now as a simple daily assessment of individual cases achieving their S5 termination temperatures within a store.*

The availability of S5 Defrost Temperature Sensor data was considered critical to understanding defrost performance within the LT cases. Not all inspected LT FGDC cases were found to have S5 probes correctly positioned[1], or reporting externally (and as a consequence there was limited archived data for analysis for the project). In fact over the project lifetime, one of the significant findings of the work was that, in general, incorrect S5 probe positioning leads to excessive energy usage and case downtime scenarios due to unpredicted evaporator coil icebuild up events.

Figure 4 below shows the average of the daily defrost termination temperature for Store1 and while some cases show varying daily termination temperatures probably related to sampling frequency, there is a clear pattern observable in the separation between cases achieving average defrost termination temperatures of greater than 5 degrees, and those closer to 0 degree C average termination temperatures. Subsequent physical inspection of the evaporator coils show good correlation between S5 defrost setpoint attainment and condition of the evaporator coil with respect to ice buildup. In Figure 4 for cases s001, s003, and s005 where S5 termination temperatures never exceeded 0 degrees, followup physical inspections were found to show significant icebuild on the coils, or minimally show signs of commencement of ice buildup around the S5 probe area.

[1]*Best practice state that S5 probe is positioned in the coldest part of the evaporator* ***(Ref 9 KER Therm Solutions 2012)*** *– by extension because most cases are made up of multiple coils, the single case S5 probe is best positioned as close to the Expansion Valve inlet of the case as possible. This approach allows for better early warning of potential ice-up conditions, where the first signs of ice-up will occur at the coldest part of the case, that being towards the Expansion Valve inlet side of the case.*

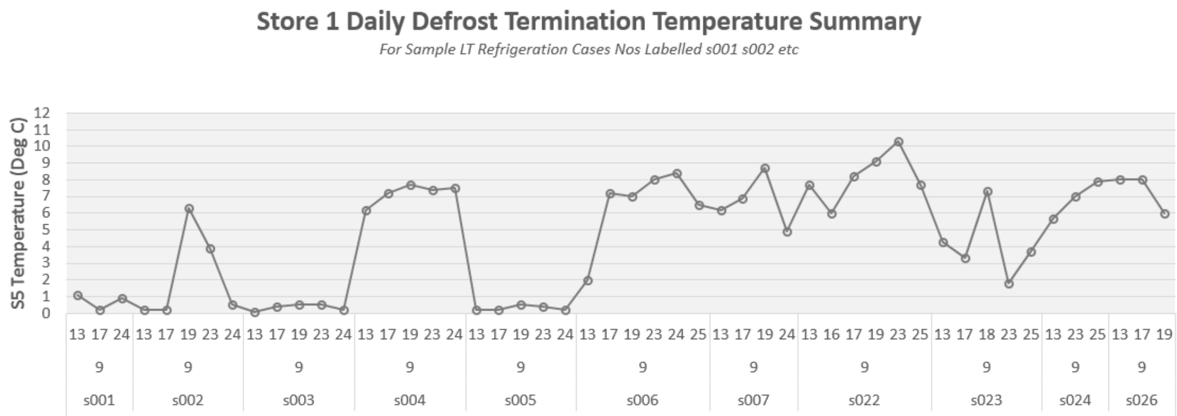

Figure 4 : Defrost Temperature Setpoint Attainment Summary Plot

2.3.1 KPI 3 : Monitoring Ongoing Reference Case Defrost Behaviour Example

Within the project, as a followon exercise, a reference case s022, in Store 1, which was initially detected and found physically as having significant ice buildup on its evaporator coil, went through a complete maintenance program cycle in preparation of reference case declaration. In this exercise the case coils were fully de-iced, and its S5 probe mounting and position validated as per previous commentary. From there the case defrost termination temperature was closely monitored, and in parallel had regular physical inspections of the coil for the following 3 month period for signs of significant ice buildup on the coils. Figure 5 shows that ongoing plot of Case s022 daily defrost termination temperatures and setpoint attainment analysis, where the continued regular physical coil inspections confirmed the expected continued ice free conditions, and thus validating the criticality of S5 data, and the good correlation between high S5 temperature attainment and proper ice free condition of the evaporator coil.

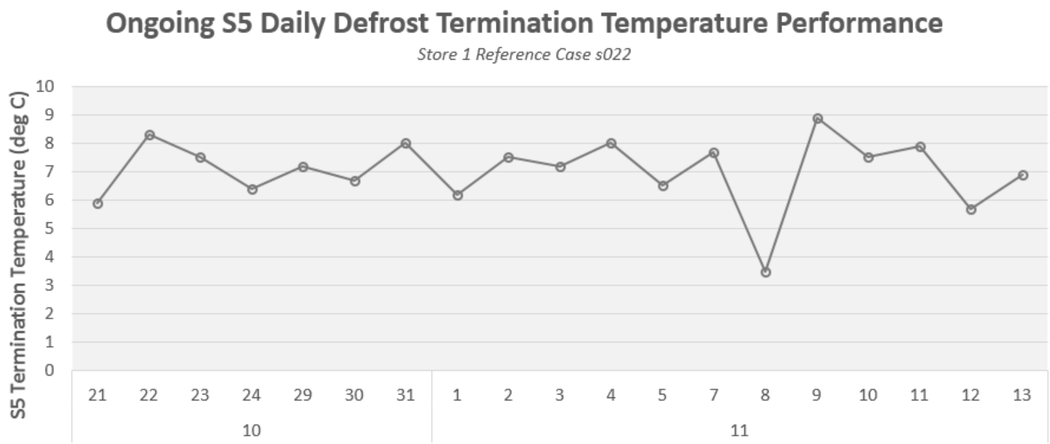

Figure 5 : Defrost Temperature Setpoint Attainment Summary Plot

2.4 KPI 4 : Post Defrost Average Pull Down Duration Estimate

***KPI 4 Post Defrost Average Pull Down Duration Estimate** – This developed KPI is a summary of recovery or pull down time for the case after defrost, and is based on an estimate of time from end of defrost status condition to S4 achieving an arbitrary -20 deg C*

This KPI was considered a good indicator of overall evaporator coil condition and case performance, where any deterioration of case aggregate recovery time (or in fact in the rate of recovery) observed over time, can be an early warning of coil efficiency degradation. Figure 6 below shows the varying pull down time monthly aggregations for Store 1 and Store 2 respectively, where it can be clearly seen that case s004 in Store 1 shows continued poor aggregate monthly pull down time with respect to its peers, and where case s060 in Store 2 showed ongoing month-on-month deterioration in aggregate pull down times. Both these anomalous cases were subsequently found to have significant coil ice buildup on physical examination showing the positive correlation between poor pull down times and ice buildup levels on the coils, and thus helping to validate again the potential value of the generated KPI with respect to case performance degradation indicators.

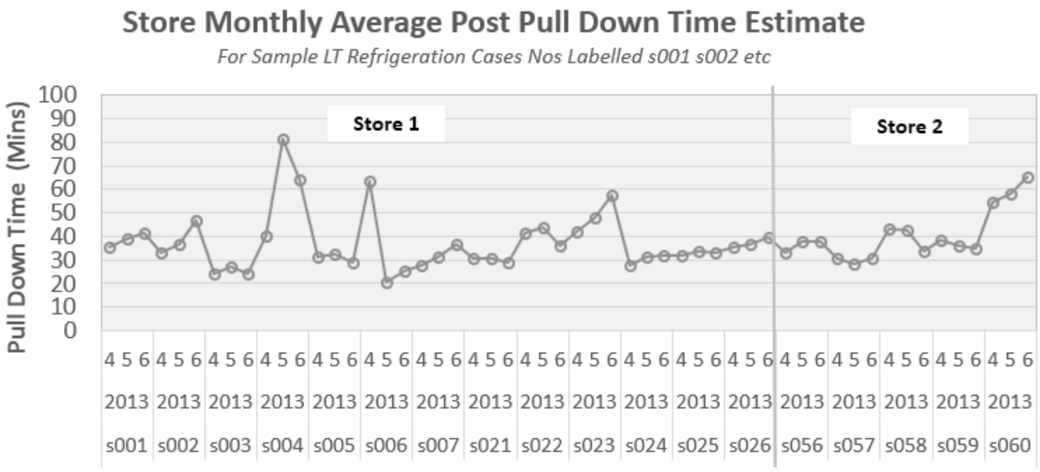

Figure 6 : Post Defrost Pull Down Time Summary Plot

2.5 KPI 5 : Steady State Case Monthly Temperature Control Performance

KPI 5a and KPI 5b Steady State Case Monthly Average Temperature Control Performance *– This KPI is simply an aggregation of the S4 Air Off (KPI 5a), and S3 Air On (KPI 5b) temperatures over the month during the defined steady state period, from post defrost end of pull down to beginning of next defrost cycle.*

While this subset of KPI's are built around the fundamental temperature control performance metrics for the cases, as can be seen from some examples given in Figures 7 and 8 below, they can still be very useful indicators of continued case performance when aggregated over time, to help identify anomalous case behaviour with respect to energy usage and possible energy saving opportunities within the estate.

Furthermore, descriptive statistical KPIs manipulating variants of the available S3 and S4 data, can also be considered here, like using Calculated Product Temperature (CPT which uses an S3 and S4 temp ratio, normally 50:50 to calculate an actual shelf product temperature) aggregation, to highlight any possible food health and safety considerations, and delta T, the difference in the S3 and S4 aggregated temperature to highlight case energy usage anomalies, with respect to in store case peer-to-peer temperature control performance, have also shown to be of value.

2.5.1 KPI 5a - S4 Air Off Average Temperature Analysis Example

S4 Air Off is the temperature of the air leaving the evaporator coil that is subsequently circulated through the case for product refrigeration at the appropriate setpoint temperature. For KPI 5a, plotting of case S4 monthly average temperatures, as in Figure 7 below, for Stores 2 and 3, one can clearly see both the large store-to-store monthly average temperature variation, as well as in store case-to-case variation. A significant portion of this variation is due to the fact that S4 is one of the few critical control levers available to the Fridge Engineer to make overall temperature adjustments to the freezer environment (and quickest to resolve warmer control temperature issues) when responding to high temperature alarm callouts, which causes cases to drift from case policy temperature setpoints over time. Having the ability through this type of analysis to quickly identify anomalous cases, for example s009, and s016 in Store 3 which are clearly operating at too cold an S4 temperature, represent significant energy wastage events and saving opportunity when identified and resolved in a timely manner, demonstrating just one example of the value of such a KPI when applied to the S4 Air Off monthly aggregated data
set.

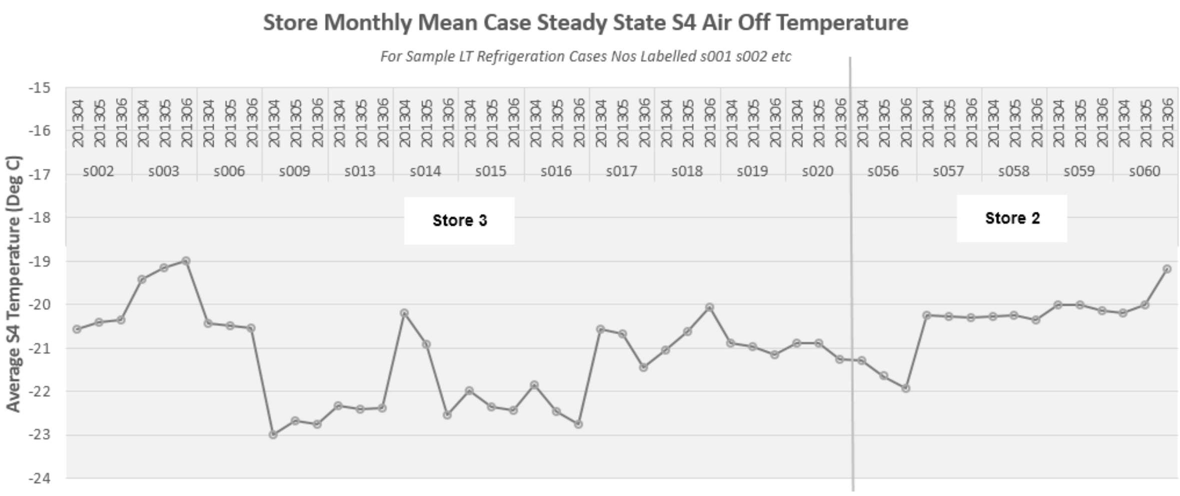

Figure 7 : S4 Air Off Monthly Average Temperature Summary Plot

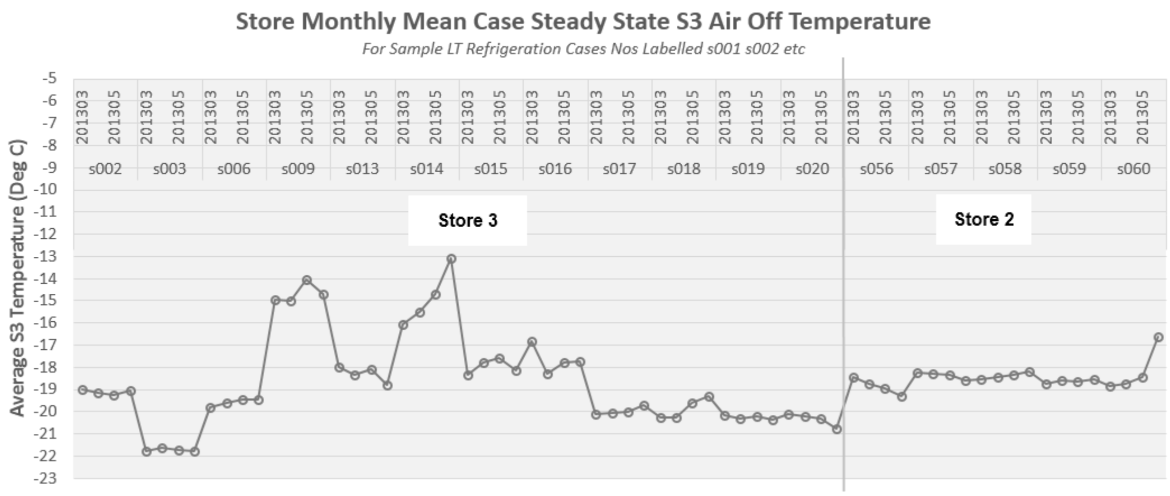

Figure 8 : S3 Air On Average Temperature Summary Plot

2.5.2 KPI 5b - S3 Air On Average Temperature Analysis Example

S3 Air On Temperature which represents the returned circulated air within the LT FGDC case is an indirect indicator of the amount of cooling demand present within the case, with warmer S3 values an indicator of increased case demand (due to possible product restock, or prolonged door opening times, or excessive air ingress due to a door seal problem). Unlike S4, S3 is not directly controllable and as a result can be effectively used as a useful indicator of overall case environmental conditions. For example drifting in S3 aggregate values over time as observed with case s014 in Store 3, and case s060 in Store 2 in Figure 8 above, can be indicators of air circulation problems within the cases, where there may be restriction of airflow across the coil (due to a possible fan failure or ice build up) which causes warmer S3 values. This in can in turn can accelerate case iceup conditions as can be seen when comparing comparable monthly S3 and S4 values in Figure 7 and Figure 8 for case s014 in Store 3, whereas expected S4 values are being pulled downwards excessively to compensate for the warmer S3's in order for the case to maintain ongoing setpoint temperature control. Another likely cause for warmer S3 temperatures over time in the LT FGDC cases may be due to excessive door opening times, although given the observed normal behaviour of this activity **(Ref 10 Fricke et al 2011)**, it is unlikely to be the cause of any systemic movement in average S3 temperatures.

3.0 Conclusions

There is significant value to be gained from the harvesting and analysing of the readily available standard refrigeration case parametric data, as a means of validating of case refrigeration performance, and to help in identifying case anomalies against standard policies and control strategies within the refrigeration case estate.

The potential for overall energy, CO2, and cost savings can be very substantial indeed when one considers just the impact of implementing one of the project recommendations based on single defrost cycles, which yielded a 2.5% overall energy saving, and that if applied to the wider retail sector refrigeration estate, has the potential in the US alone to save multiple Terrawatthours annually.

And not only will this deeper knowledge yield significant energy savings in the short term, but can also lead to a positive wider impact on future fridge control strategies. For example, by implementing appropriate best practice positioning of S5 defrost sensor, and making the data available near realtime, will lead to improved case performance, and implementation of more effective and consistent defrost strategies across the estate.

Also it has been shown that with when appropriately defined KPI's are applied to the underlying refrigeration parametric dataset it is possible in many cases to accurately predict impending refrigeration case failures with respect to one of the biggest causes of freezer case downtime, and associated loss of sales, that of excessive ice buildups within the evaporator coils. Having the ability to predict and prevent failures and subsequent loss of sales, through this low cost KPI analysis approach should be considered a useful addition to augment current maintenance support strategies.

4.0 References

1. Tasou et al, “Energy Consumption and Conservation in Food”, Applied Thermal Engineering, Vol 31, Issues 2-3, Feb 2011

2.Goodwin, B. “Tesco uses big data to cut cooling costs by up to €20m” Online ComputerWeekly Journal Publication (http://www.computerweekly.com/news/2240184482/), May 2013

3. Ward et al “Undefined By Data: A Survey of Big Data Definitions”, Cornell University Library, September 2013

4. Bertoldi et al “The European End-Use Energy Efficiency Potential and the Needed Policies”. European Commission, Presentation at Fondazione Eni Enrico Mattei, Milan, Italy, December 2003

5. Goetzler et al, “Cooling Energy Savings Opportunities in Commercial Refrigeration”, Navigant Consulting Inc Final Report to US DOE, September 2009

6. European standard Jan 2002 EC Standing Committee on the Food Chain and Animal Health Regulation (EC) No 178/2002

7. European standard, EN ISO 23953-2:2005. Refrigerated display cabinets - Part 2: Classification, requirements and test conditions , 2005

8. Fricke B et al “Demand Defrost Strategies in Supermarket Refrigeration Systems” Interim Report Retail Energy Alliance, October 2011

9. KER Therm Solutions Missouri USA, “Refined Coil Temperature Sensor Location” Website Publication (http://www.ke2therm.com)/files/Q-1-3a_Sensor_location.pdf), May 2012

10. Fricke B. et al “ Energy Use of Doored and Open Vertical Refrigerated Display Cases on Excessive Door Opening” International Refrigeration and Air Conditioning Conference Purdue, July 2010