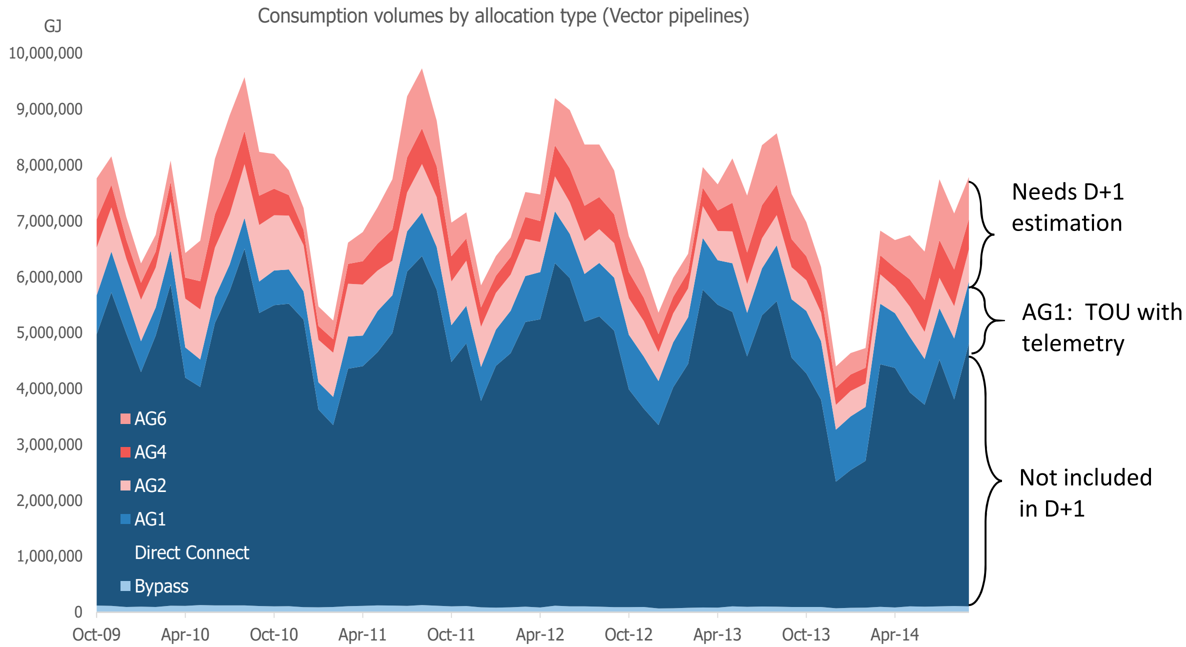
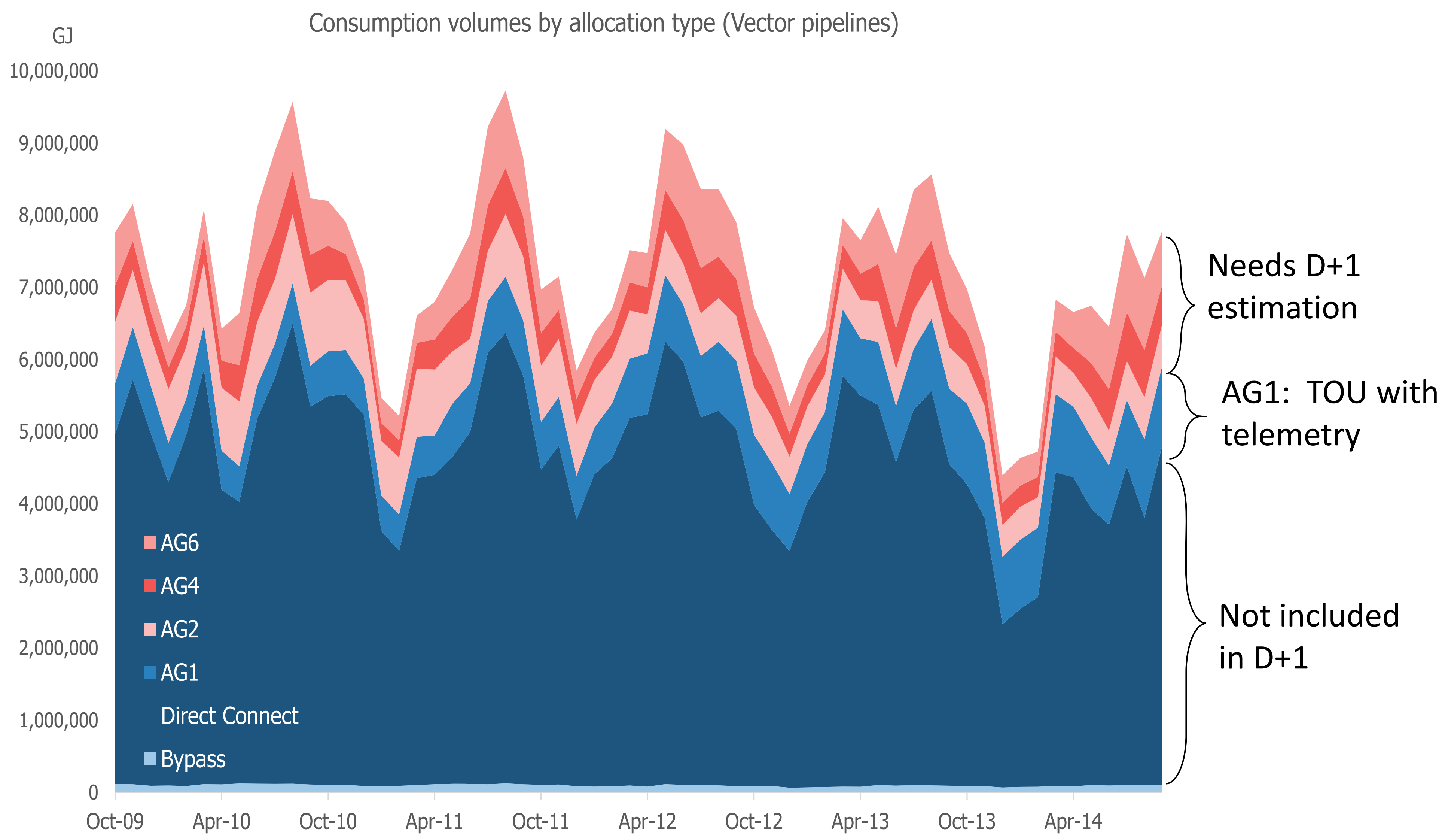
Developing a D+1 Allocation Model

D+1 context

- Balancing occurs for total gas quantity, of which allocated gas only makes up a small portion of total gas consumption.
  - Gas supplied to direct connect and bypass gates is known already
- So any errors in the D+1 allocation will be muted by other known gas
  - e.g. 5% error in allocated gas might translate to a 1-2% total gas error





Consumption volumes by allocation type (Vector pipelines)
GJ
10,000,000
9,000,000
8,000,000
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
Oct-09
Apr-10
Oct-10
Apr-11
Oct-11
Apr-12
Oct-12
Apr-13
Oct-13
Apr-14
AG6
AG4
AG2
AG1
Direct Connect
Bypass
Needs D+1 estimation
AG1: TOU with telemetry
Not included in D+1

When making the D+1 allocation, what is known?

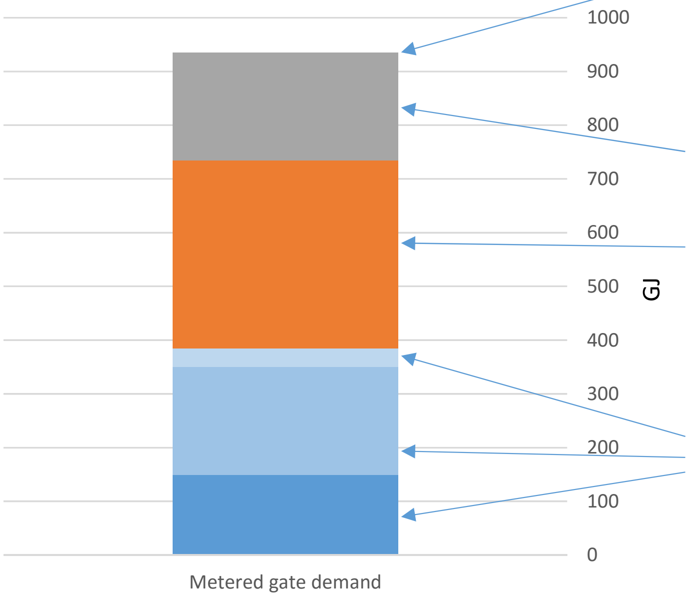
- D+1 attempts to estimate gas consumption for non-TOU-non-telemetry customers soon after the trading day using limited sources of data that are available close to real time.
  - Metered injection at each Maui TP welded point.
  - Metered injection at each Vector gas gate.
  - Allocation group 1 consumption at each Vector gas gate.
  - The number of ICPs for each retailer and allocation group at each Vector gas gate.
  - The AUFG factor for non-G1M gas gates.

Modellable

- Although actual metered data for non-AG1 customers is not available, the D+1 process can make reasonable estimates because gas consumption is predictable
- In combination with the available ‘real-time’ data, consumption for non-AG1 customers can be estimated using statistical models developed using additional historical information on:
  - Previous month's consumption
  - Number of customers
  - Time of year
  - Day of week

Modelling Approach

Explained Simply



- Start with metered injection at each Vector Tx gate
- If a G1M gate, model UFG for that gate
- Take off UFG-adjusted group 1 consumption
- Take off *modelled* group 2 consumption
- Aggregate residual gas for all gates in a pool
- Split the residual gas according to modelled proportions
  - Group 4 and 6 consumption is combined. The model produces a total non-TOU allocation

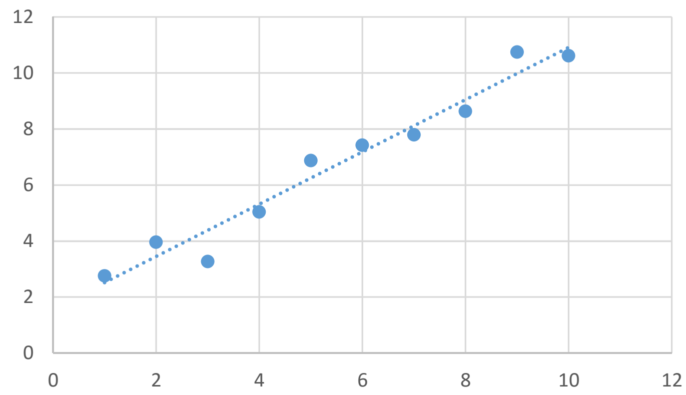
At heart of D+1 process are three models used to estimate consumption

Three models

- All models are produced using "R"
  - "R" is an open source statistical modelling and analysis package developed at University of Auckland and used worldwide

- They are multi-variable linear regression models

- The models use established relationships between parameters to predict consumption from future parameters



This is a simple linear regression model

- Models constantly updated so that they use the most up to date data available at the time

First model – Calculating UFG (only for G1M gates)

- A regression model predicts the UFG for every G1M gate for every month

Explanatory Variable	Why?
Month of Year	UFG can be seasonal
Month number (over time)	UFG can drift over time
AUFG from previous month	UFG from month to month can be highly correlated

- The UFG model isn't very complicated, but it's much better than setting UFG = 1
  - % error for group 1 using model is <1%.
  - % error using UFG=1 is >3%

Second model – Group 2 ICP consumption

- The group 2 model predicts consumption for every group 2 ICP for every day
- The model is weighted so that more recent observations have more effect on the D+1 allocation

Explanatory Variable	Why?
Month of Year	Consumption is often seasonal
Daily injection at gate	A large group 2 will often have a significant effect on total gate injection
Business day or not	Many group 2 ICPs have higher consumption on weekdays
Average ICP consumption for previous month	Average consumption often stays constant from month to month

Third model - Residual gas share

- The residual model predicts the *share* of residual gas to be allocated to every retailer for every pool for every day
- It is also weighted to favour more recent observations

Explanatory Variable	Why?
Month of year	Mass market consumption is seasonal
Average share of residual gas during previous month	Share of residual gas from month to month is highly correlated
Number of non-TOU ICPs	As a retailer gains more customers, its consumption will increase. Also, ICP data is much more up-to-date than consumption data
Retailer	Each retailer has its own mix of customers, which affects consumption.

Results

Comments and Details

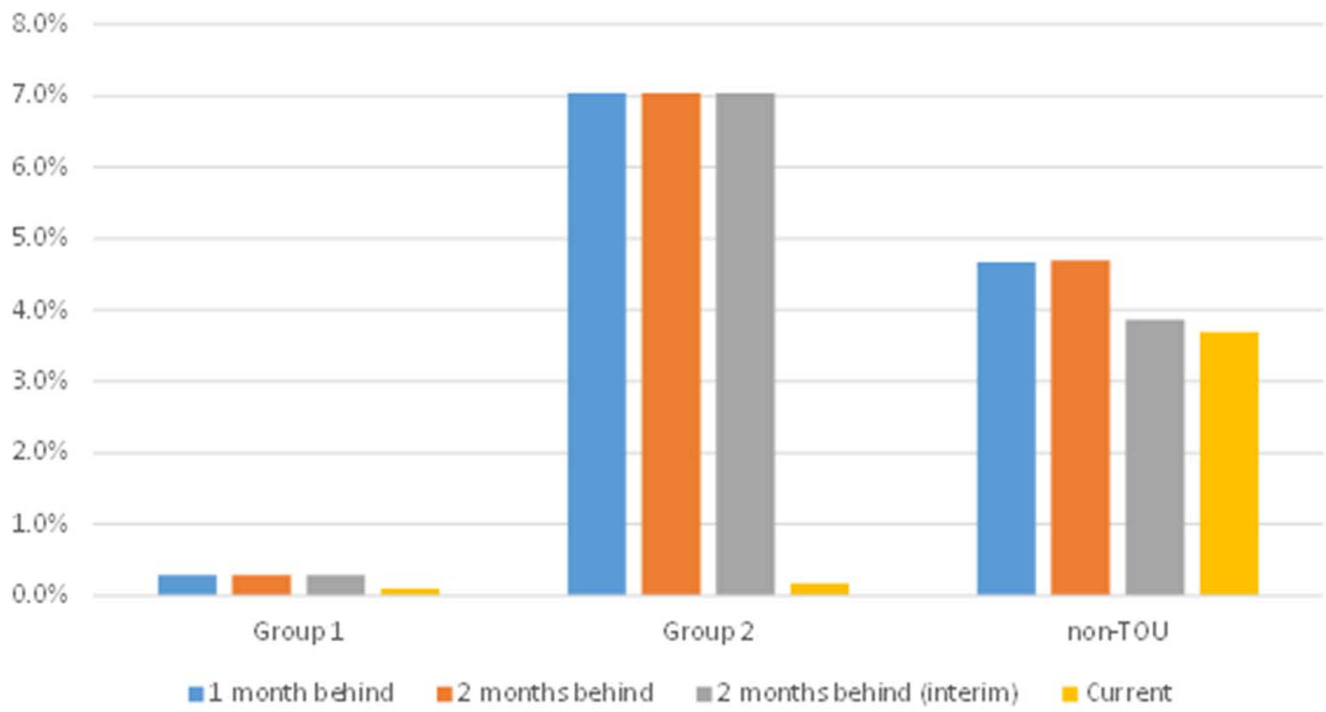
- Predictions are produced for a retailer's consumption at the pool level.
- Sometimes these predictions are also divided into allocation group 1, allocation group 2 and non-TOU consumption.
- There can by multiple retailers owned by the same parent company. Predictions are made for a parent company. (e.g. GEND, GENG and GEOL consumption are all included in the "GENG" prediction.)
- The results presented are for the period from October 2013 to September 2014.

Four types of result are calculated

- Three different approaches for using historical volume share information for calculating residual shares:
  - “1 month behind”. Using the previous month’s Initial allocation
  - “2 months behind” Using the month before the previous month’s Initial allocation. (Because for approx. the first week of a month, the previous month’s Initial allocation wouldn’t be available)
  - “2 months behind (interim)”. Model currently uses *Interim* allocation from two months previous. Not feasible in practice, but intended to give indication of what might be achievable if retailers submitted their *Initial* allocation later in the month.
- A "current" error is shown in most graphs. This is the difference between the current Initial allocation and the best known allocation (e.g. Interim or Final). It is a useful reference point to compare any D+1 errors.

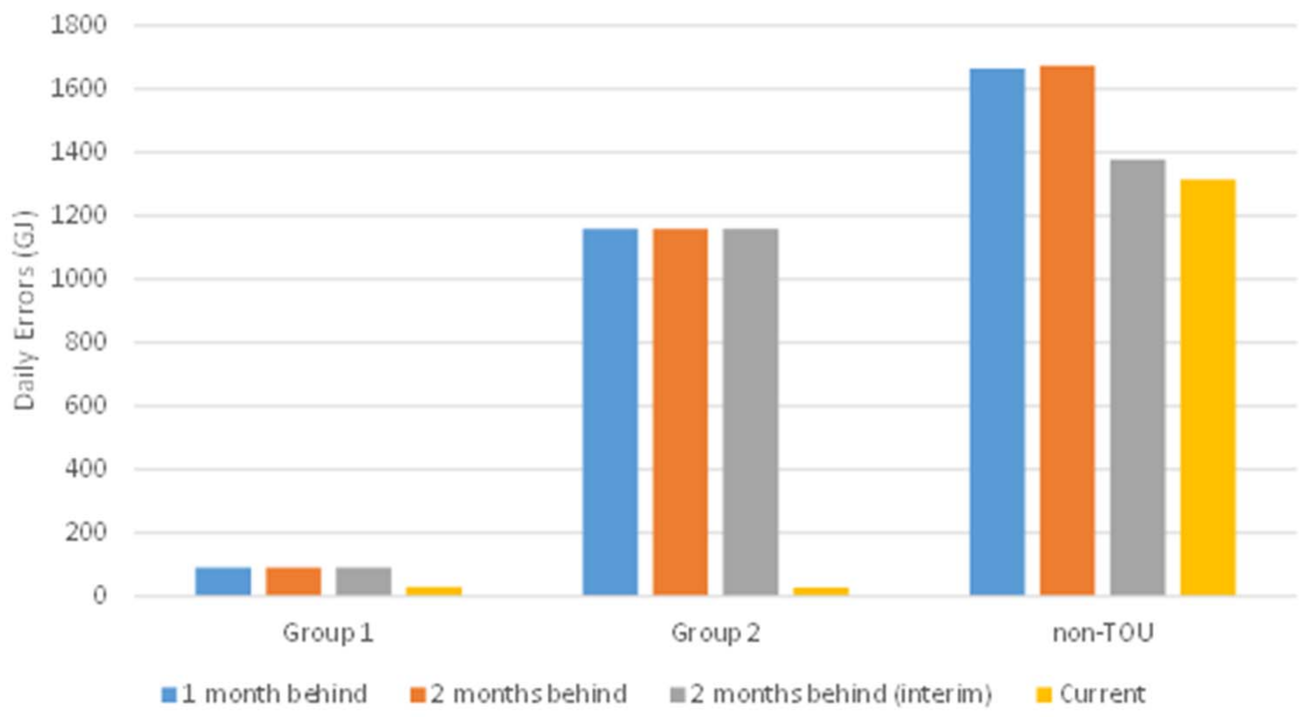
Percentage error by allocation group

- Group 1 errors are low
- Group 2 error are high. They have the highest percentage error of all the groups and they are much higher than Current
- Non-TOU errors are slightly higher than Current



Absolute error by allocation group

- The absolute group 2 error is less than for non-TOU, since there is less group 2 consumption



Example of error calculation

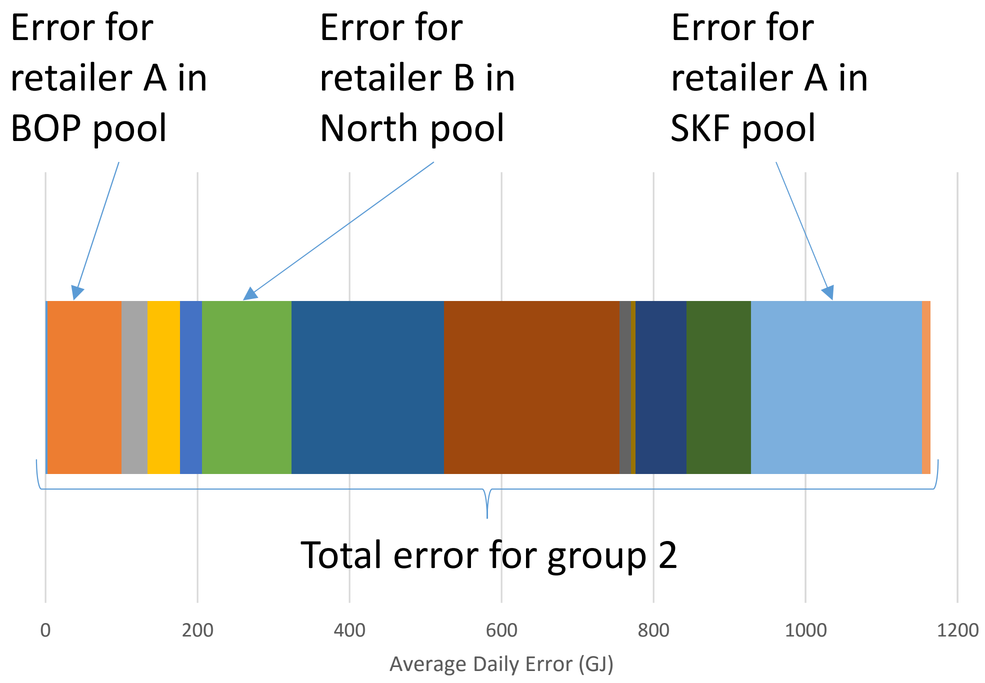
- If needed - here's an example of how the process works
- 3 example retailers are shown
- The percentage error for this group would be 11%
- Total D+1 and actual allocations sum to same amount (gate injection)

	D+1	Actual	Absolute Error	%age
Retailer A	800	1000	200	20%
Retailer B	5500	5000	500	10%
Retailer C	2700	3000	300	10%
Sum	9000	9000	1000	11%

← Weighted average

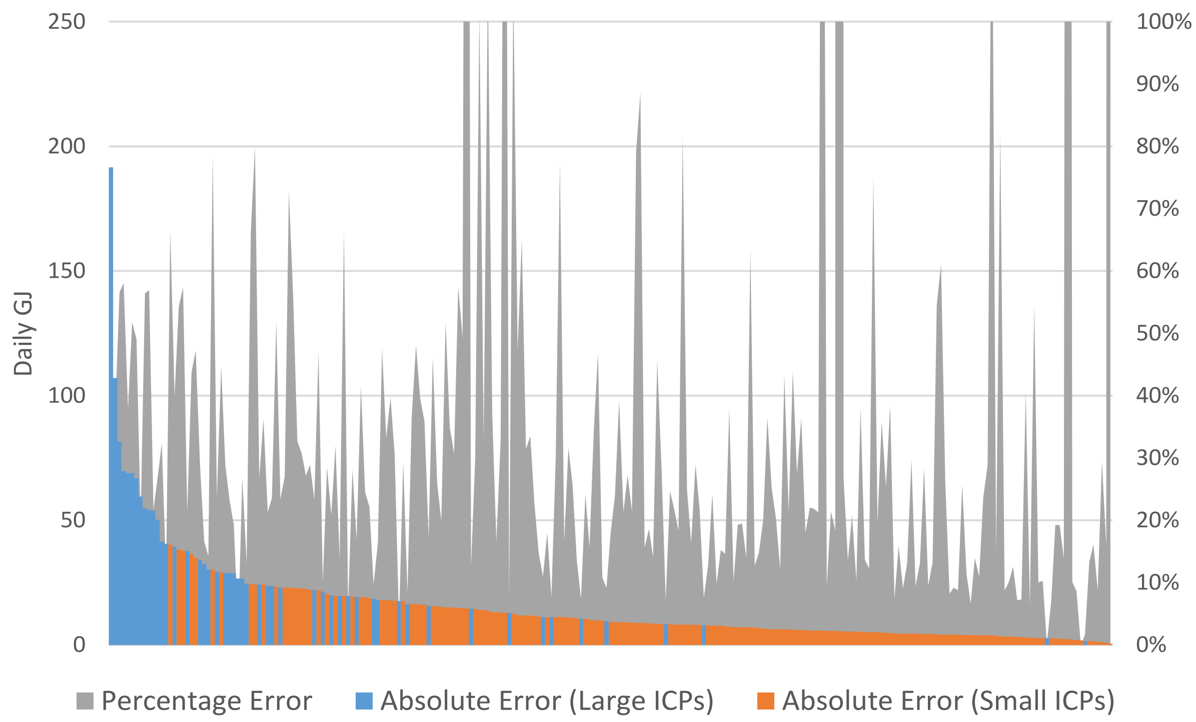
Error calculations in more detail

- A daily error is calculated for each of the ~60 combinations of retailer, group and pool (3 pools, 3 groups and 5-7 retailers)
- These combinations are grouped in different ways to show different things about the errors.
- The previous 2 slides grouped the errors by allocation group.
- The group 2 error on the previous slide was about 1200GJ
- This is simply the sum of the individual pool and retailer combinations for group 2
- The percentage error is the absolute error divided by the consumption in that group



Group 2 examples

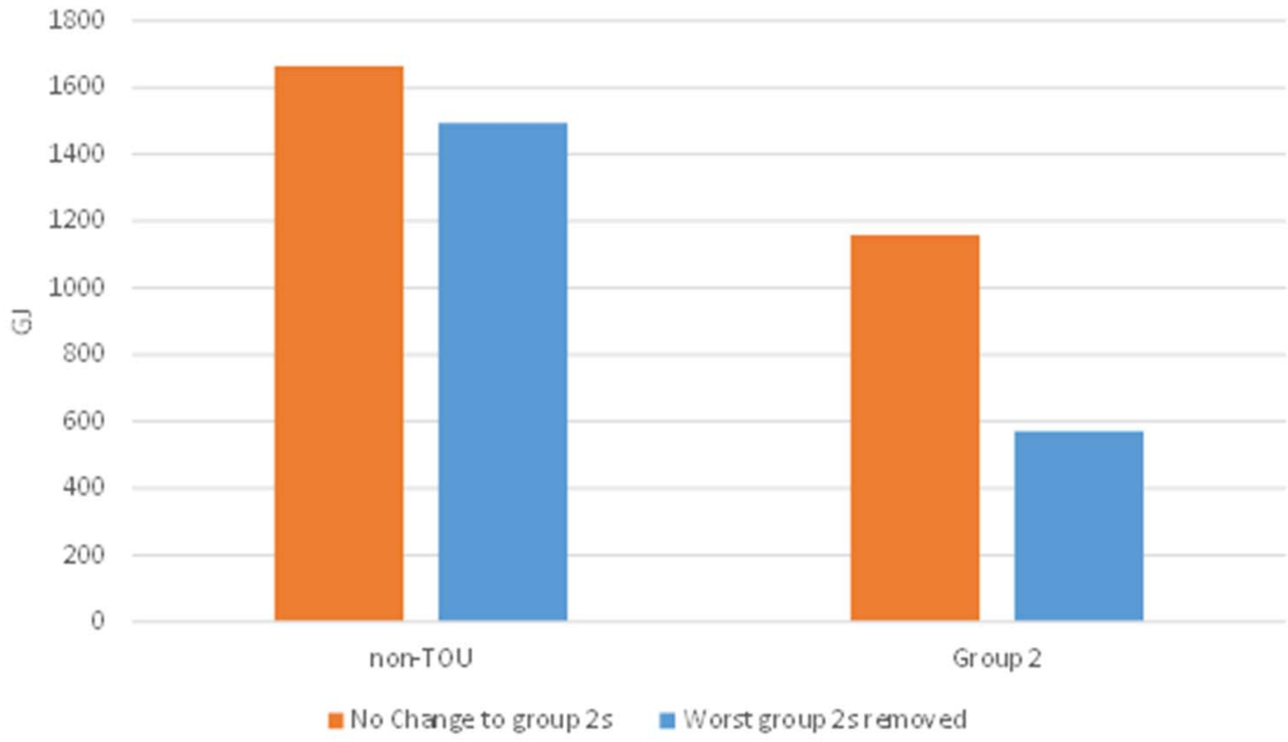
- Group 2 ICPs errors vary greatly
  - Ease of prediction depends on a number of factors
- In general, larger ICPs have larger error
  - But not always
- Often large errors are caused by an ICP starting or stopping
  - Advise allocation agent?
- Other times it is simply because they are very variable
  - Difficult to improve prediction without some sort of reporting



Moving ICPs from group 2 to group 1

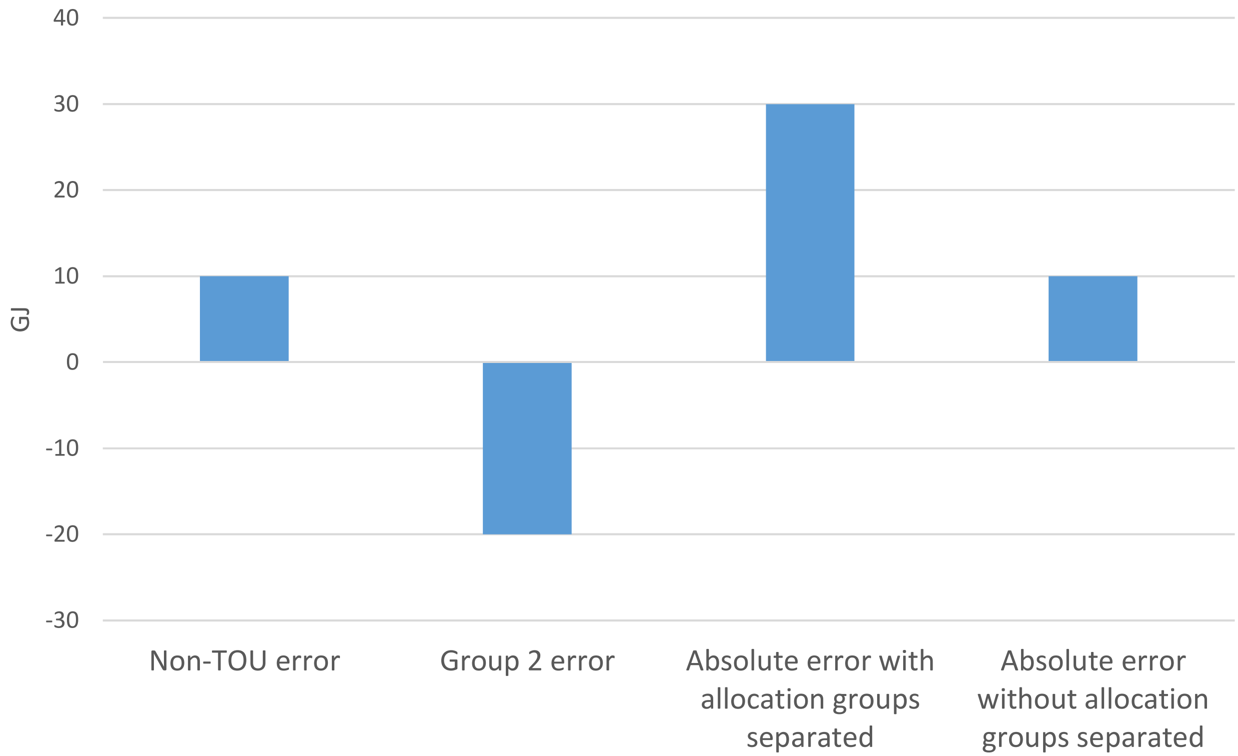
- If the 50 ICPs with the largest absolute errors were switched from group 2 to group 1, this would reduce total group 2 error by 50%

- This would also reduce non-TOU error as residual gas would be known with more accuracy



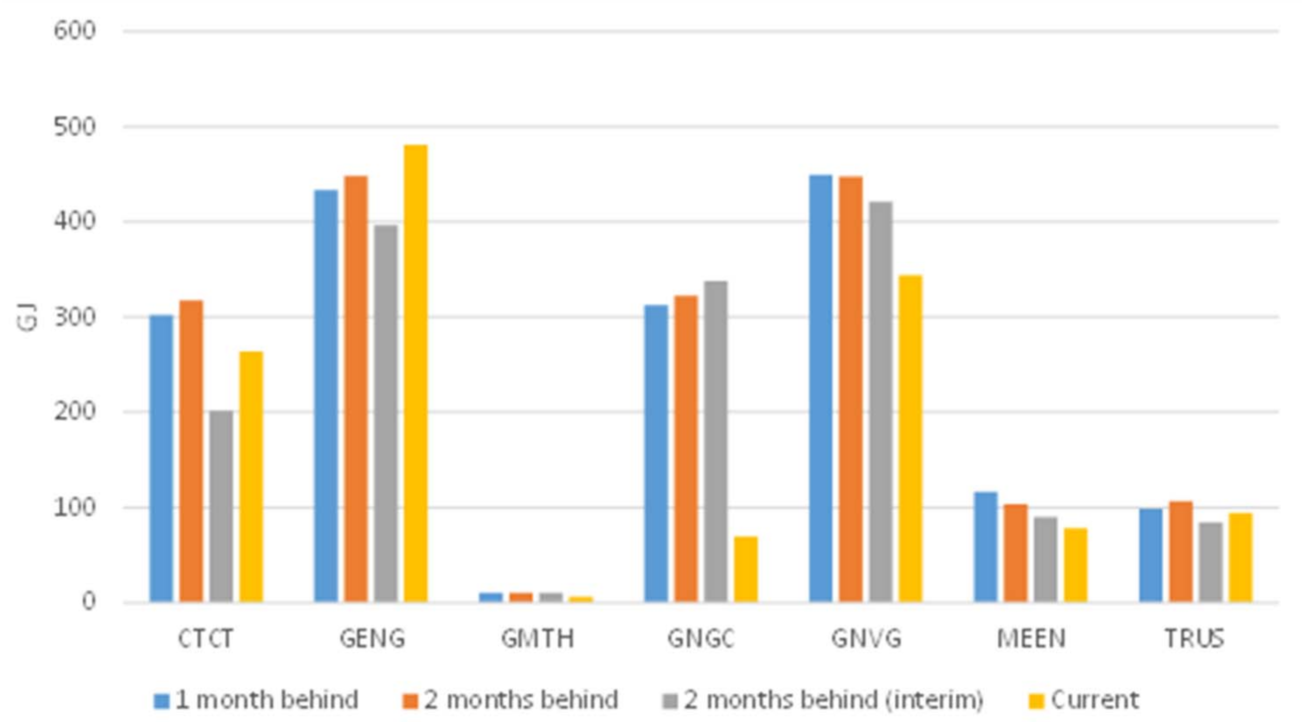
Two different grouping methods

- For most of the results that follow, the errors are not distinguished by allocation group.
- "Overs and unders" cancel out
- This is because the primary purpose of the D+1 allocation is to allow retailers to balance their nominations at a pool level



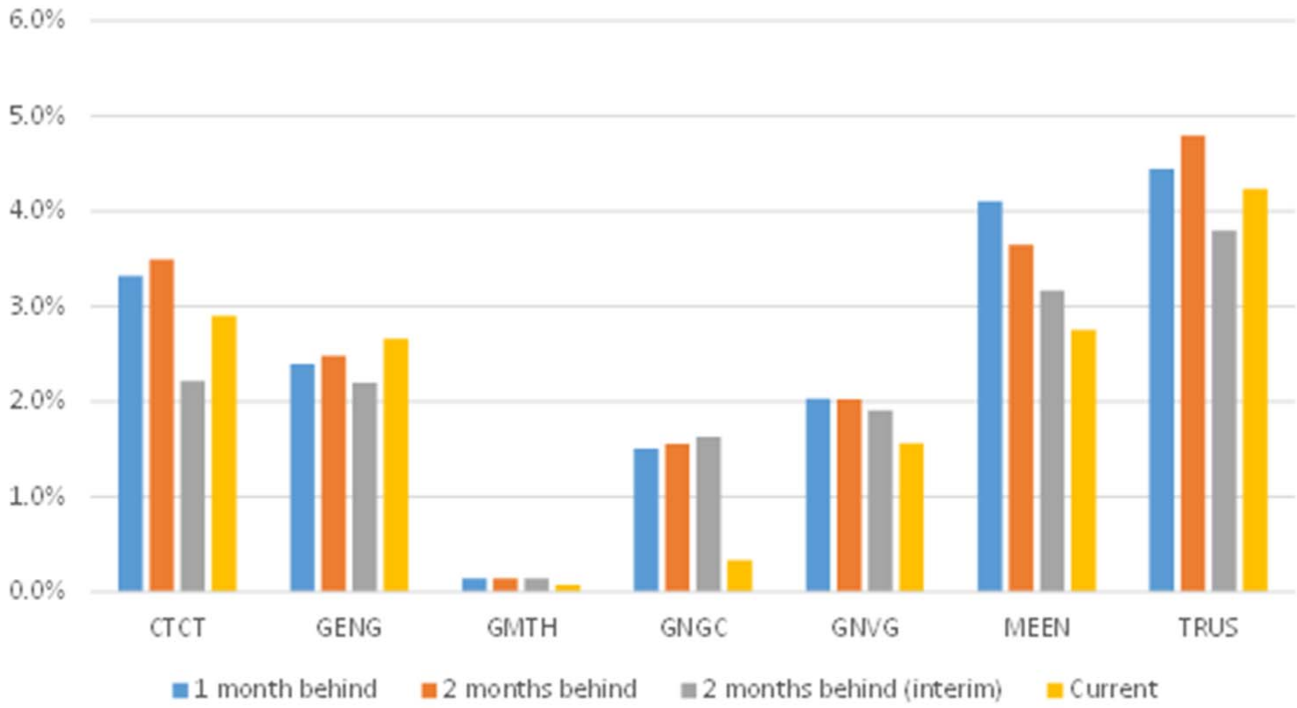
Absolute error by retailer

- This shows the error for retailers across all allocation groups
- The primary determinant of the size of a retailer's total error is the size of the retailer's consumption
- i.e. larger retailers have larger error



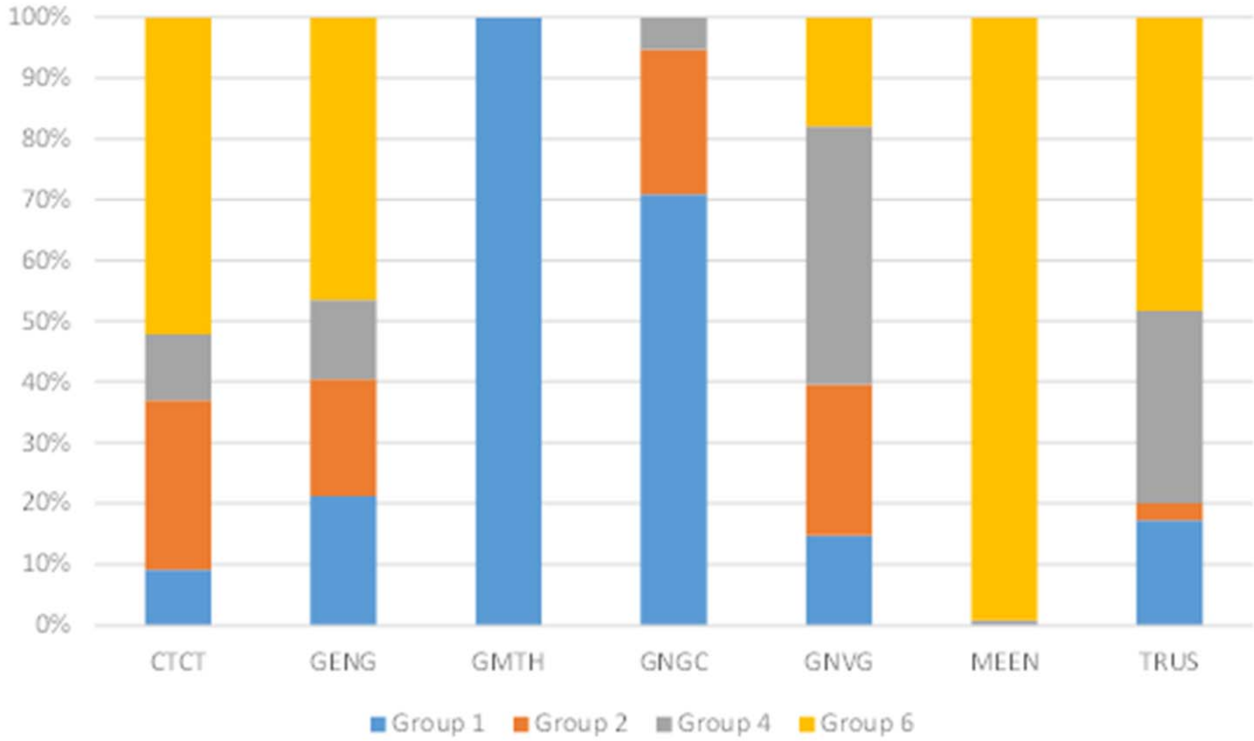
Percentage error by retailer

- Percentage errors are highly influenced by the amount of group 1 consumption a retailer has
- Group 1 error is very low, so retailers with a large amount of group 1 consumption have lower error overall



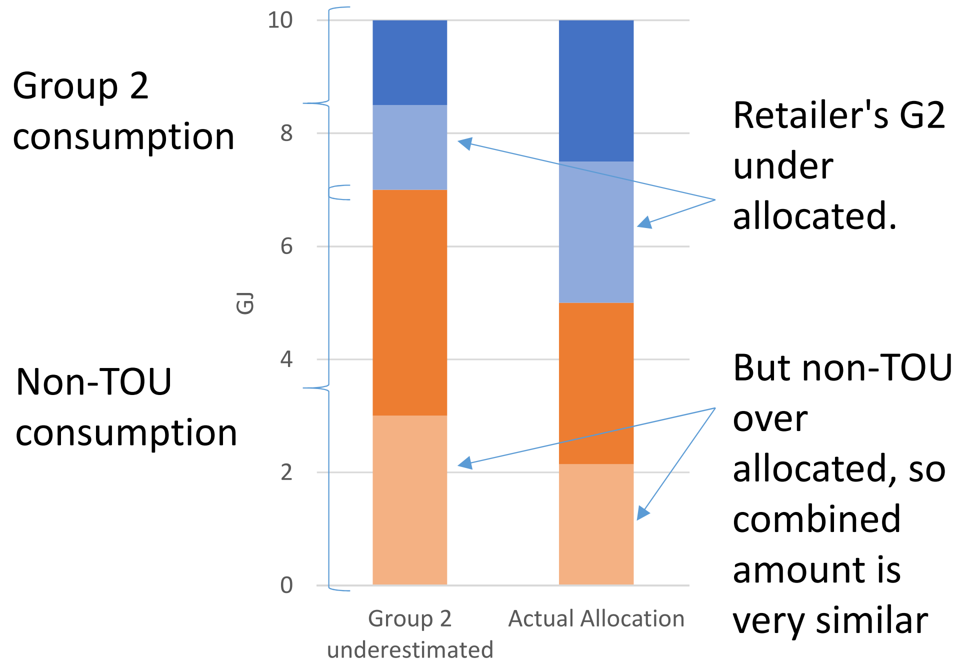
Share of retailer consumption in each group

- GMTH has almost entirely group 1 consumption
  - As such, it has very low error
- GNGC also has a large percentage of group 1 consumption
  - But its error is not significantly lower than other retailers
  - This is because of its non-AG1 consumption being dominated by the harder-to-predict AG-2 group, and doesn't have non-TOU group to 'offset' errors



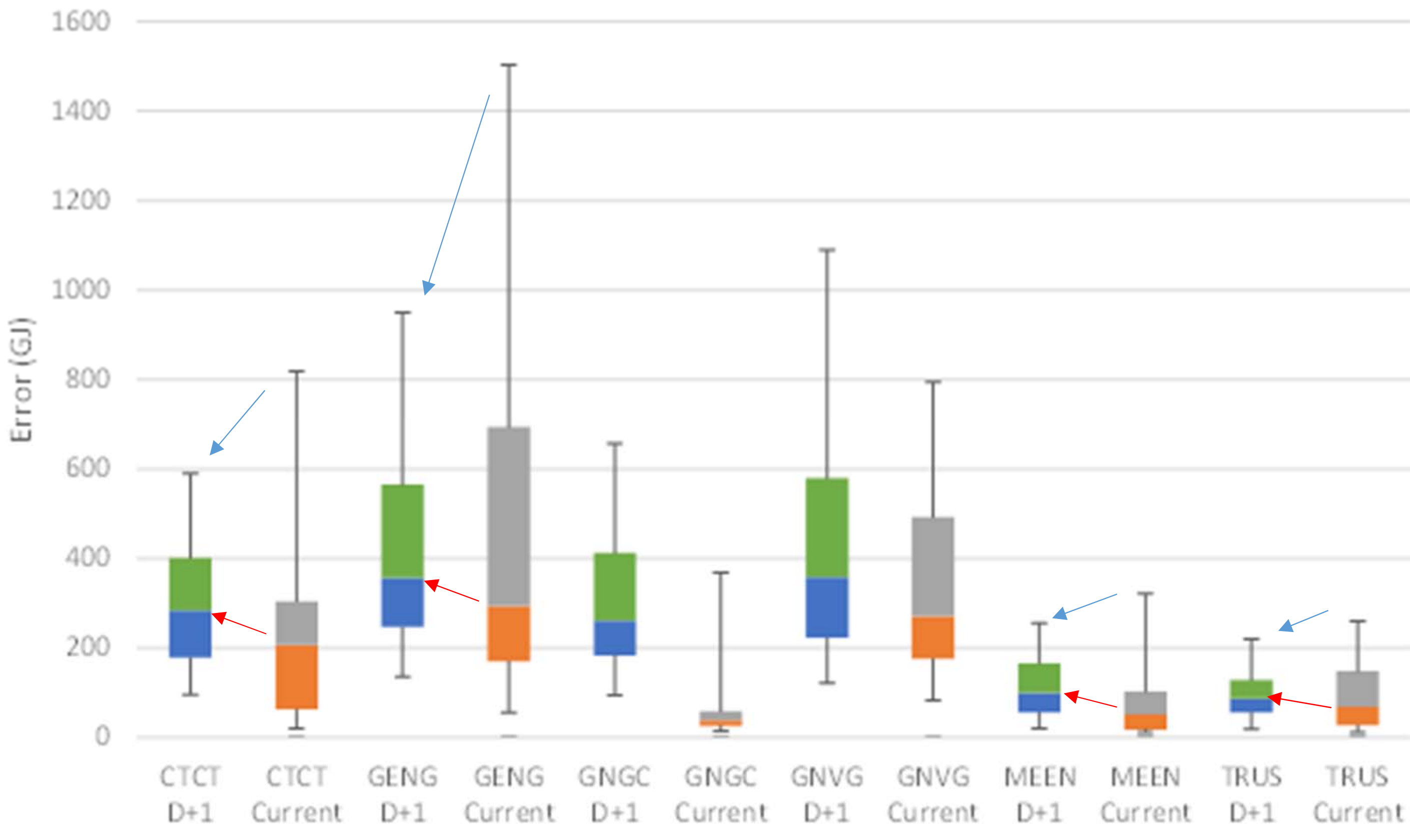
Offset

- Since the gas injected at a gate and group 1 consumption is known, the total group 2 and non-TOU consumption is also known
- If the model underestimates group 2 consumption, then it will necessarily overestimate non-TOU consumption. In this way, the modelling process is somewhat self correcting.
- Retailers that have both group 2 and non-TOU consumption will benefit from this
  - GNVG does well!



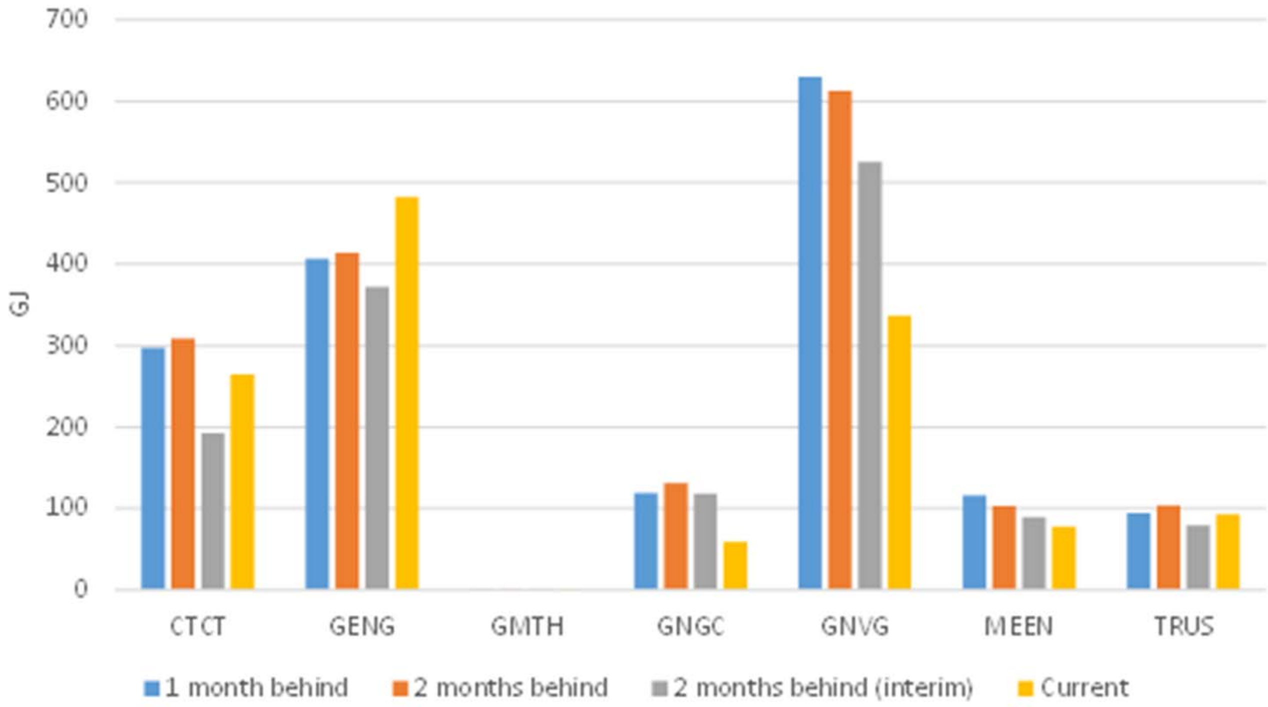
Distribution of errors for retailers - all pools

- Average errors don't tell the full story
- For many retailers, compared with current Initial error, median error is higher in the D+1 process, but large errors occur less often
  - May be more useful for balancing purposes



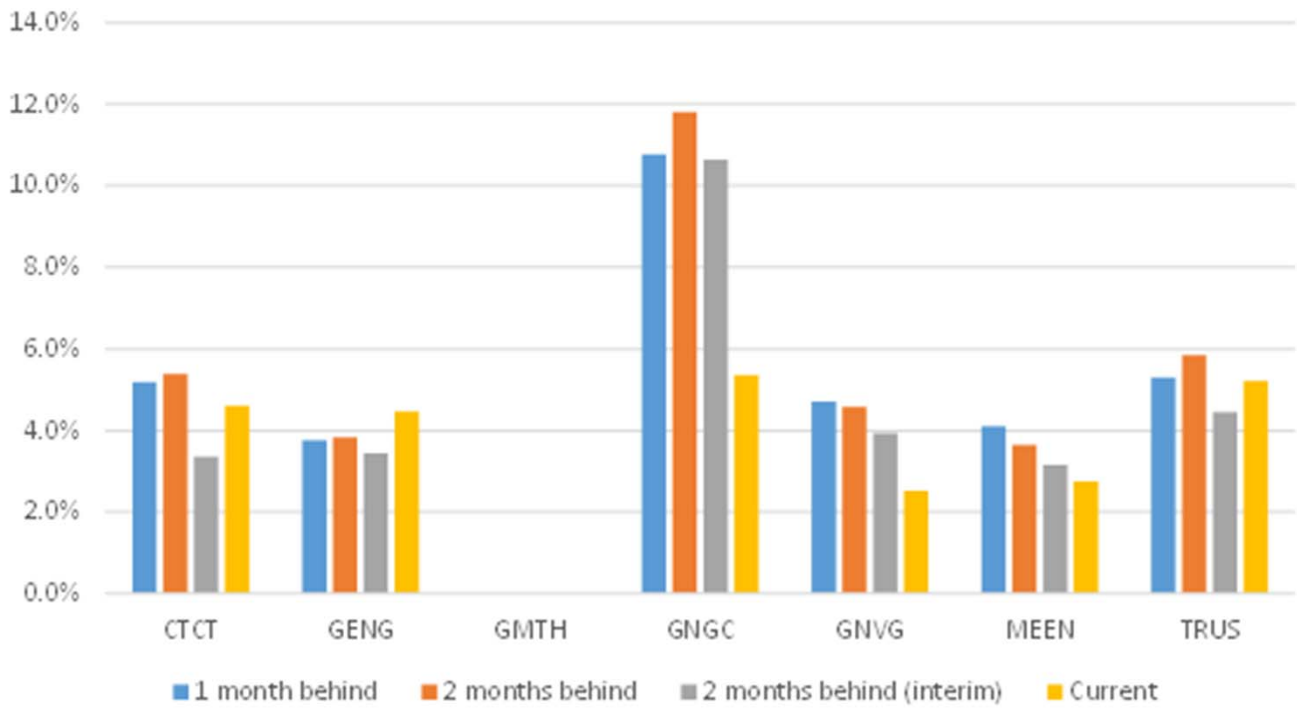
Non-TOU absolute error by retailer

- Looking at Non-TOU only errors isolates any differences in the mass market modelling process
- Just like for the "all consumption" errors, the size of a retailer's error is mostly determined by the retailer's consumption
- For some retailers (CTCT, GENG, TRUS) D+1 could be better than current. (And would almost certainly be the case if large AG2s move to AG1.)



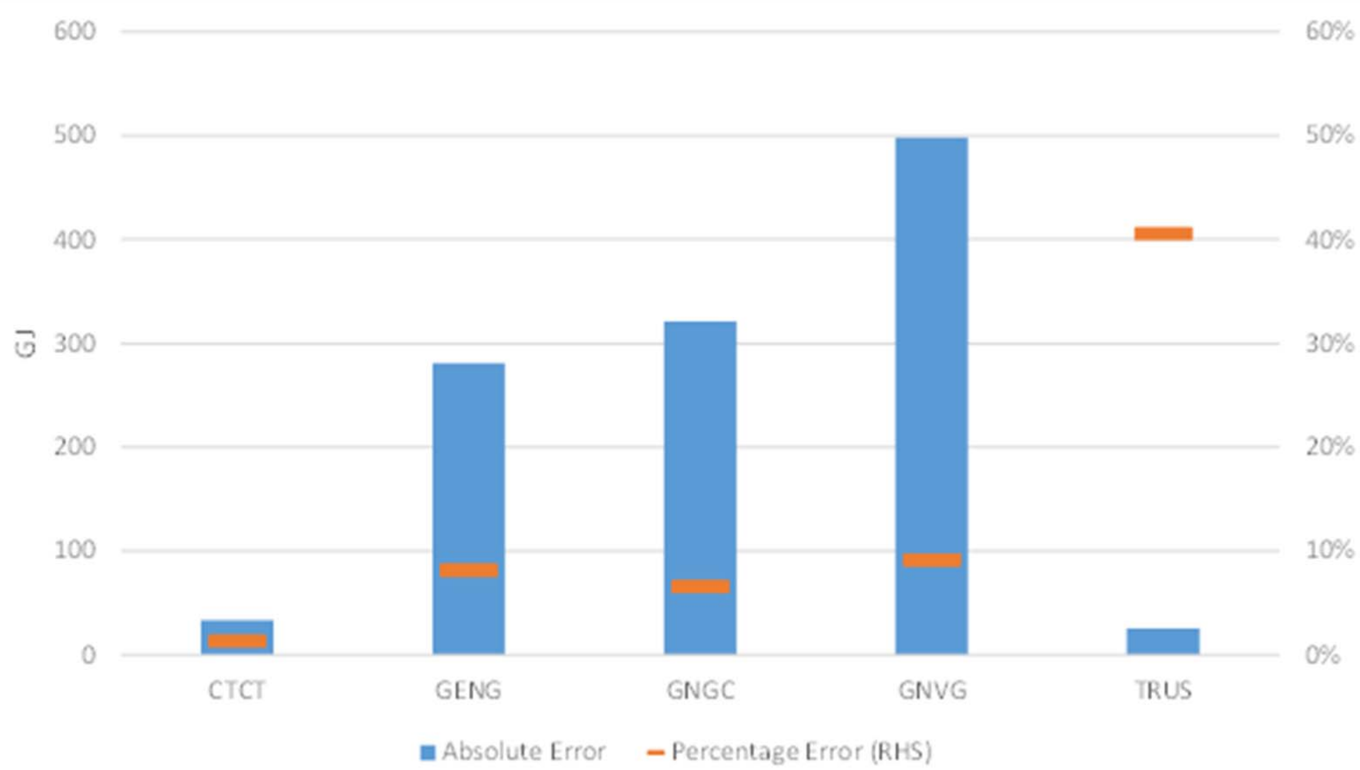
Non-TOU percentage error by retailer

- However, there are also significant differences between percentage errors.
- GNVG and MEEN's current error is lower than others
  - Monthly readings
- GNGC's error is much higher
  - Higher proportion of difficult to model group 4 consumption
  - Small in absolute terms



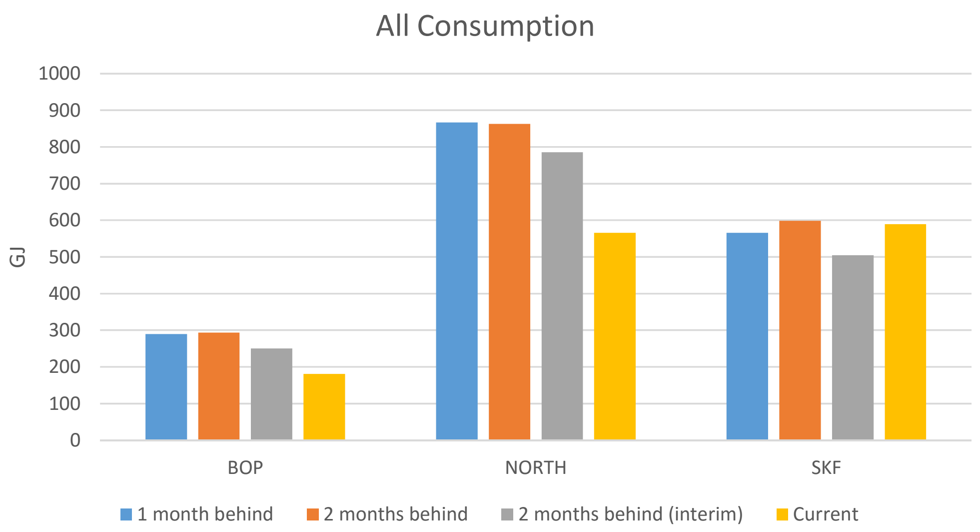
Group 2 results by retailer

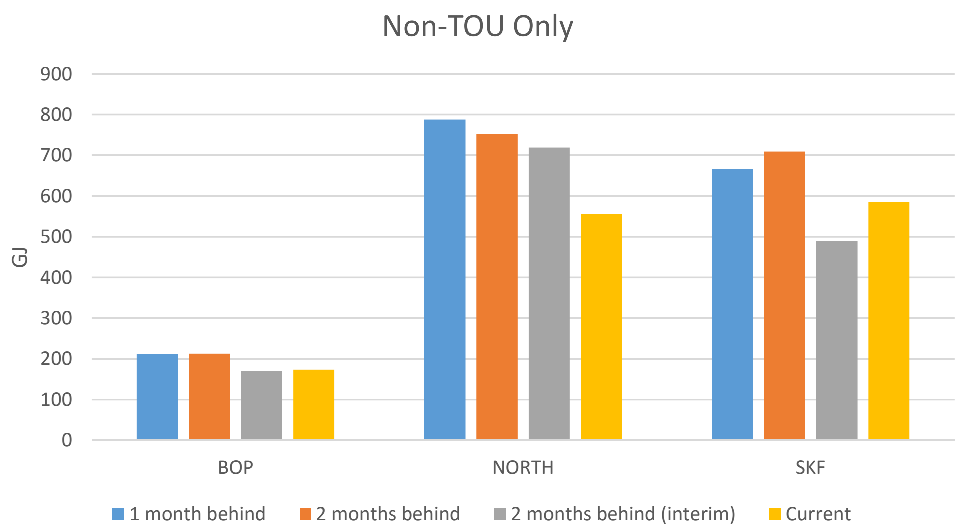
- Largest 3 retailers are similar
- CTCT has slightly lower group 2 consumption
- And significantly lower percentage error
  - Something special about CTCT group 2 ICPs?
- TRUS has very small amount of group 2 consumption



Absolute error by pool

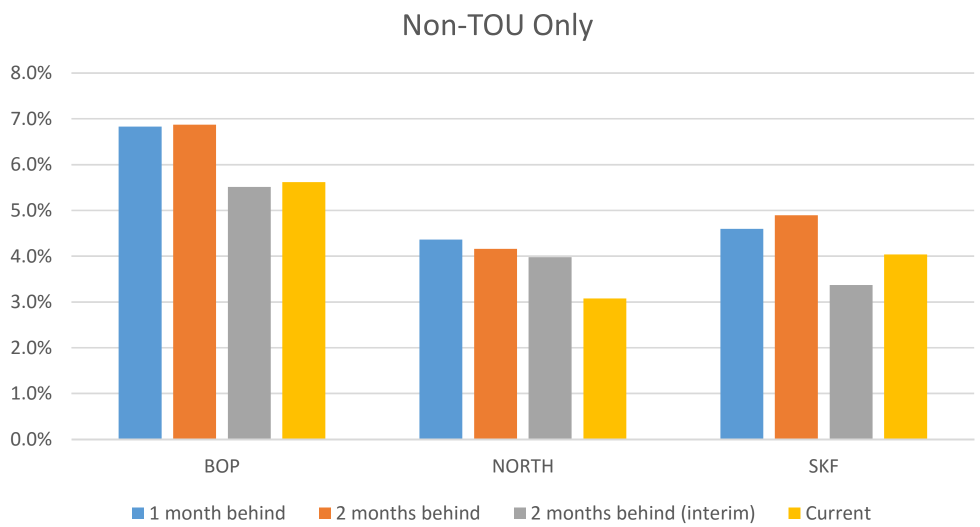
- The primary determinant of the size of a pool's error is the size of the pool
- i.e. larger pools have larger errors

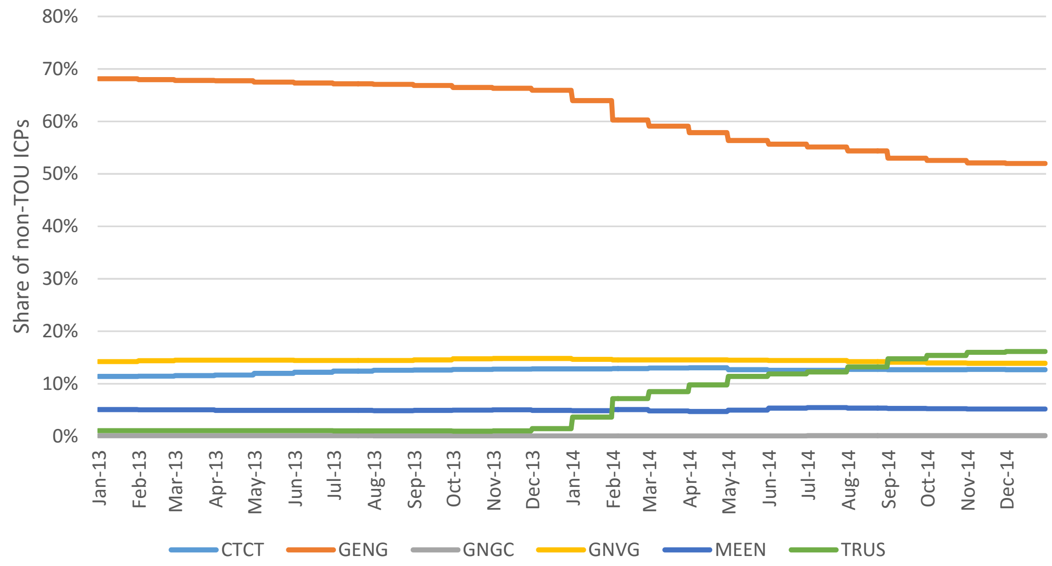




Percentage error by pool

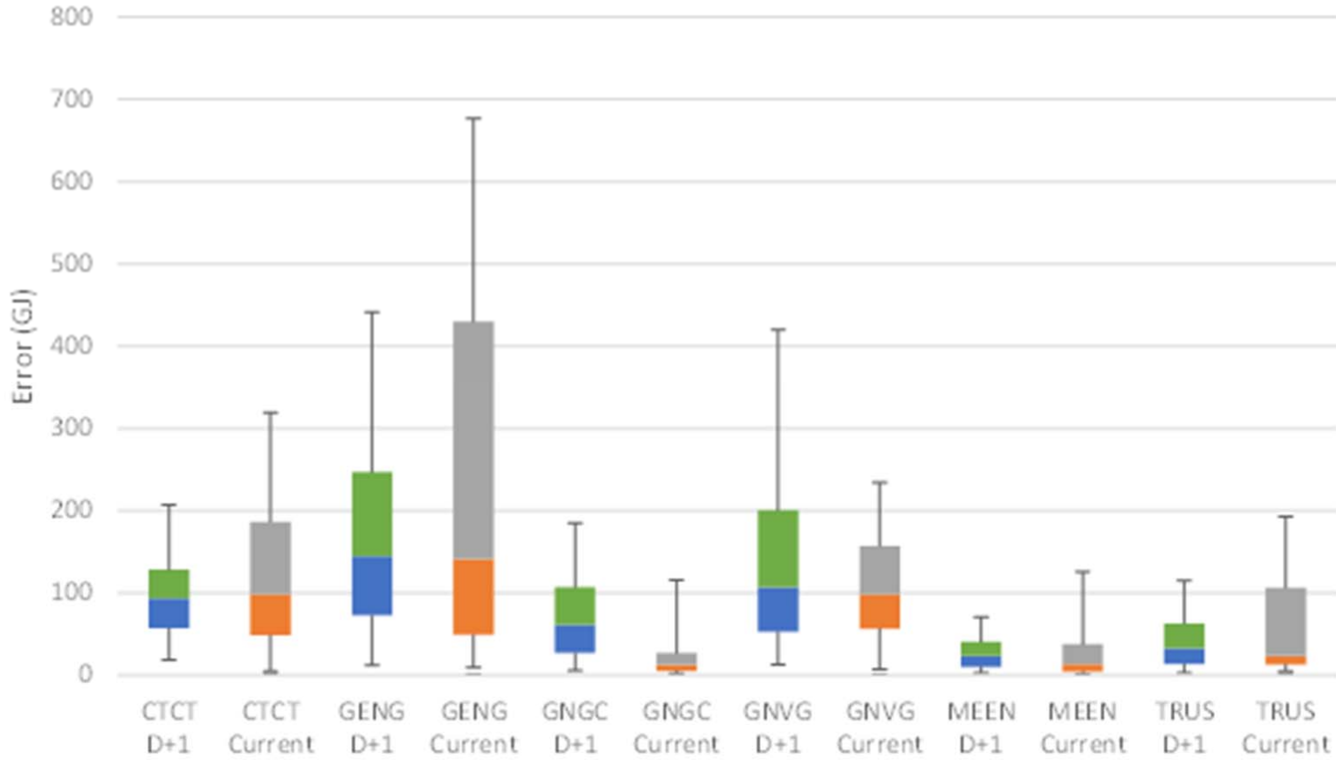
- For both non-TOU only and "all consumption" the BoP pool has the highest error, while the North pool has the lowest
  - Only non-TOU is shown, but "all consumption" is similar
- BoP error is higher because of the increase in ICP share for TRUS





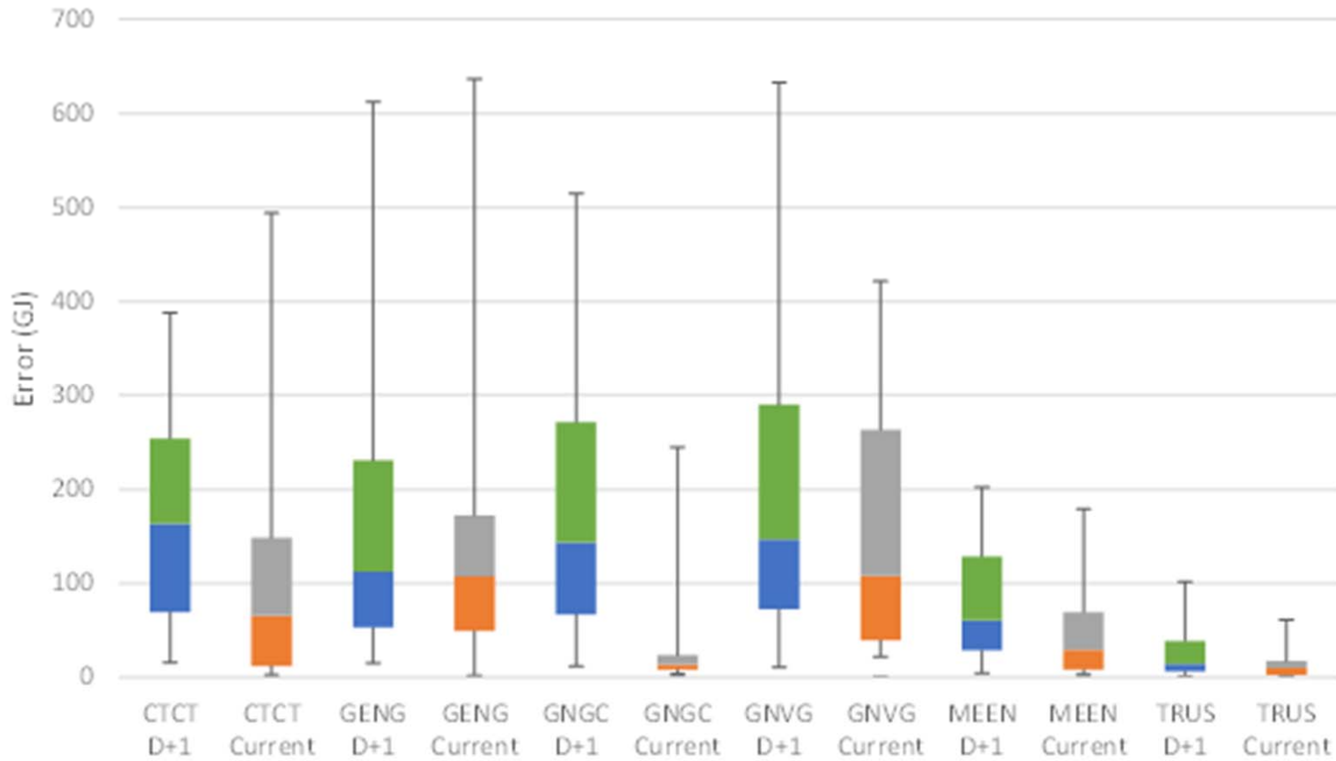
Distribution of errors for retailers - SKF

- In general individual pool graphs have similar characteristics to all pools graph
  - Any differences are highlighted
- TRUS errors are lower
  - TRUS market share does not change significantly in SKF



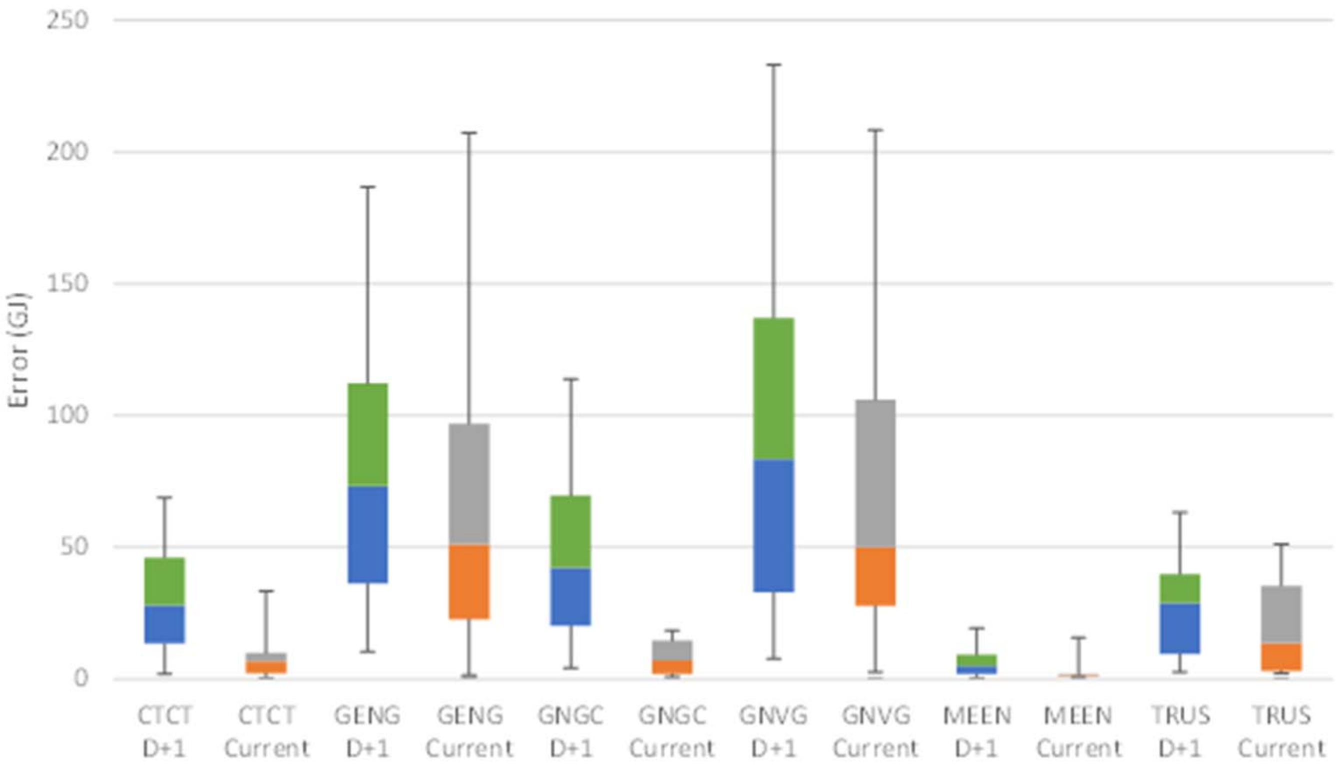
Distribution of errors for retailers - North

- GNGC error is higher than for other pools, which reflects high group 2 consumption
- TRUS error is also higher, which is caused by TRUS increasing its market share



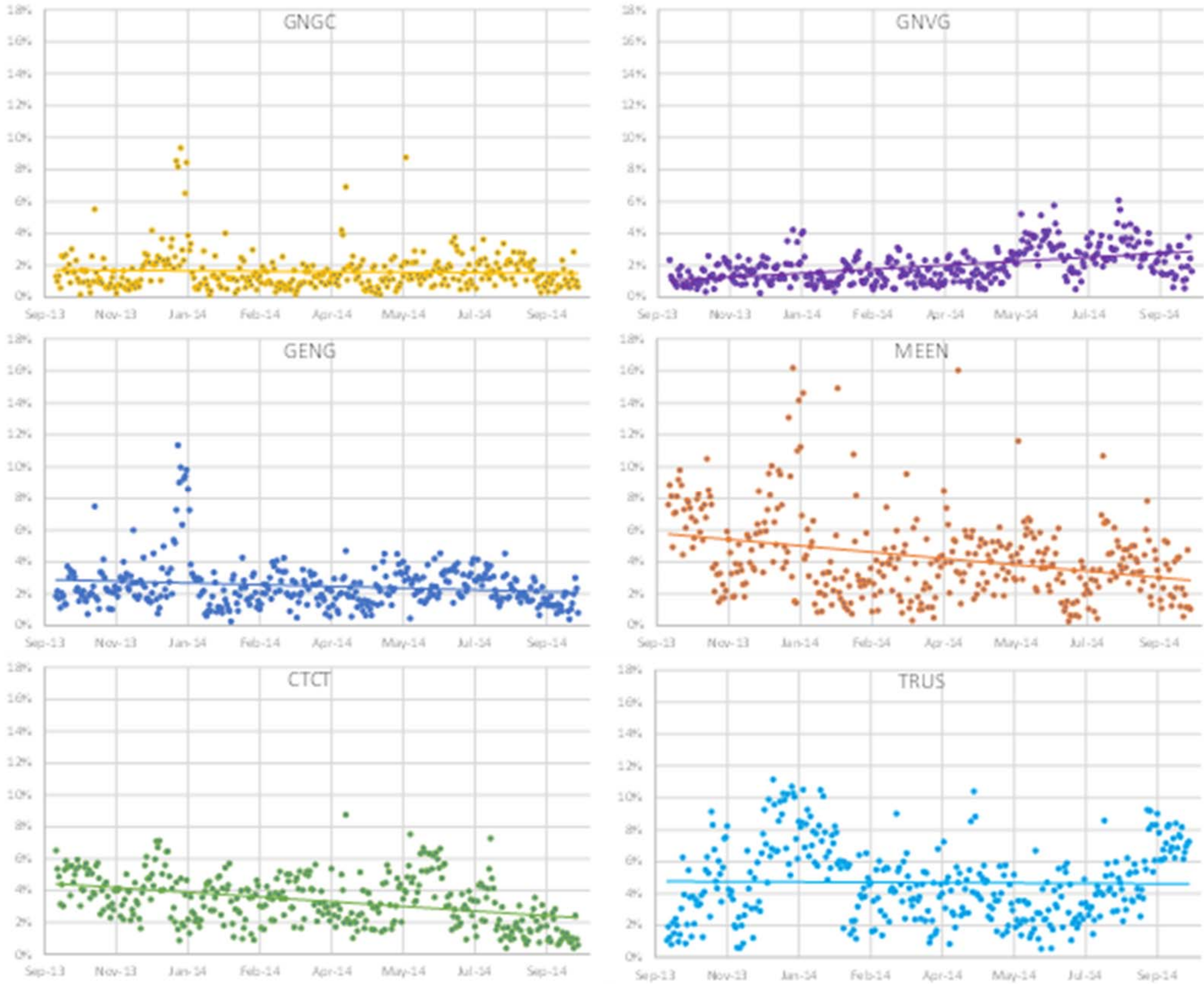
Distribution of errors for retailers - BoP

- The D+1 prediction for MEEN is worse than the current allocation.



Daily D+1 Errors

- MEEN and CTCT appear to be getting better
- GNVG appears to be getting worse
- There are worse errors in December and over Easter
  - Harder to predict because variation in whether businesses operate during such periods



Conclusions

- In general, the D+1 process produces allocations that have a similar error to the current Initial allocation
  - Group 2 ICPs are modelled more poorly – an inherent problem due to inherent variability in AG2 consumption patterns, and the fact that the current Initial allocation has actual meter reads
  - non-TOU consumption is modelled with a similar level of accuracy
  - Combined group error is also similar
  - GNGC is notably worse, due to its large amount of group 2 consumption, and because most of its non-TOU consumption is group 4 (which is inherently harder-to-predict than group 6)
- The more immediate information provided by a D+1 allocation probably outweighs any small increase in error when daily balancing is required
  - Improved nominations facilitated by D+1 should result in reduced incidence of excess pipeline balance