Modeling Technical Document

TCCC Indonesia Marketing Mix

Version 1.0

[1. Background 3](#_Toc514000909)

[2. Demand Drivers Edge (DDE) Modeling Module 3](#_Toc514000910)

[2.1. Model Scope 3](#_Toc514000911)

[2.1.1.1. Multiplicative: 3](#_Toc514000912)

[2.1.1.2. Bayesian Shrinkage 4](#_Toc514000913)

[2.1.1.2.1. Bayesian Priors 5](#_Toc514000914)

[2.1.1.2.2. Estimation 5](#_Toc514000915)

[2.1.2. Coefficient Bounds 6](#_Toc514000916)

[2.1.3. Measure Creation 6](#_Toc514000917)

[2.2. Model Form 7](#_Toc514000918)

[2.3. Model Measures 7](#_Toc514000919)

[2.4. Measure Transformations – 8](#_Toc514000920)

[3. Model Quality Criteria 9](#_Toc514000921)

[3.1. Face Validity: Coefficient Sign 9](#_Toc514000922)

[3.2. Face Validity: Elasticities 9](#_Toc514000923)

[3.3. Face Validity: Competitive Effects 9](#_Toc514000924)

[3.4. Face Validity: Volume Contribution 10](#_Toc514000925)

[3.5. Face Validity: Due To Change 10](#_Toc514000926)

[3.6. Face Validity: Year ago change 10](#_Toc514000927)

[3.7. Face Validity: Dummy Variables 10](#_Toc514000928)

[3.8. Statistical Validity: Model Fit 10](#_Toc514000929)

[3.9. Statistical Validity: Correction Factor 10](#_Toc514000930)

[3.10. Statistical Validity: Model Error 11](#_Toc514000931)

[3.11. Statistical Validity: Confidence 12](#_Toc514000932)

[3.12. Statistical Validity: Forecast Accuracy 12](#_Toc514000933)

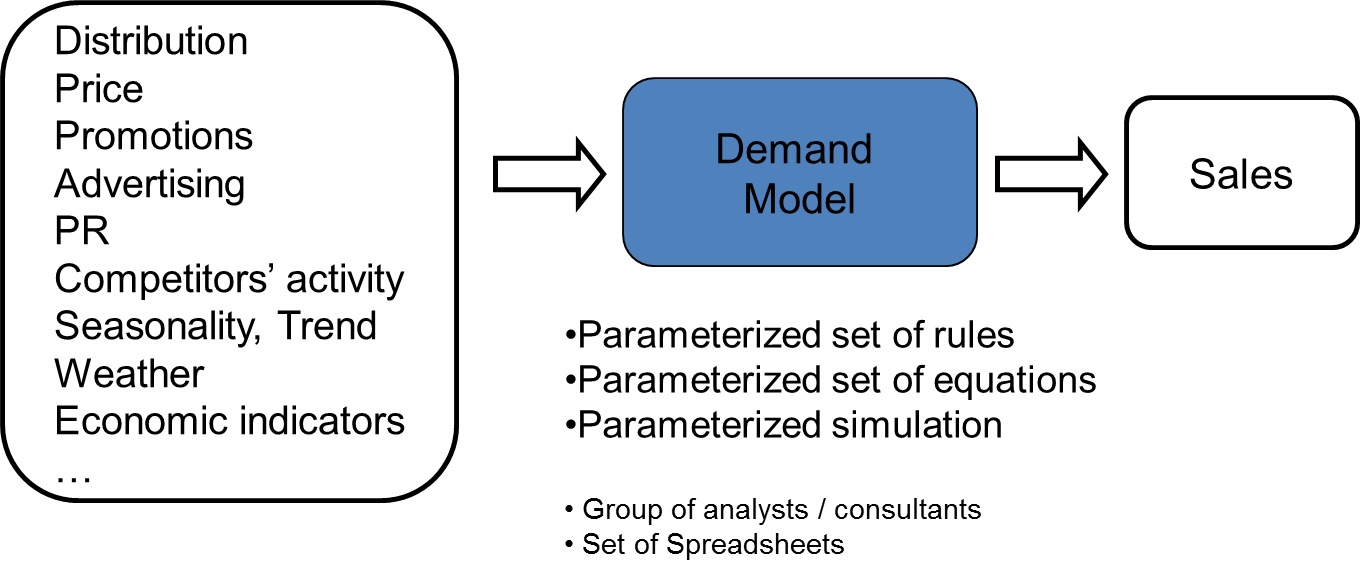
[3.13. Due-To Change YOY 12](#_Toc514000934)

[3.14. Media Effectiveness and Efficiencies 13](#_Toc514000935)

[3.15. Response Curves 13](#_Toc514000936)

# Background

MMx analysis is an analytical solution that estimates sales (Units, Volume, Revenues, Shipments) as a function of marketing activity and other measures describing the state of the market.



MMx analysis is used for:

* Measuring the contribution of activities to sales; disentangling the effects that are happening simultaneously
  + Volume contribution from activity; translates into ROI
  + Explain volume differences over time
* Enhancing understanding of the market place
  + Measuring the impact of an activity; Price elasticity, promotion lift and strength of competitor’s impact
* Predicting sales for any marketing plan
  + Media budget simulation and optimization
  + Sales forecasting
* Monitoring of changes in the market place
  + Does my old model still comply with the current data?
  + How do models calibrated on old and new data differ?
  + What does it say about the effectiveness of activities?

# Demand Drivers Edge (DDE) Modeling Module

# Model Scope

We develop MMx models using our platform Demand Drivers Edge (DDE). DDE runs statistical models by blending POS, media, promotions, macro-economic indicators data and any other client specific inputs required to quantify the relationship on sales. For the Indonesia Marketing Mix we will be using the following model form:

# Multiplicative:



Pros:

* Larger relevant range
* Implicitly captures some interaction

Cons:

* May create some counter-intuitive interactions that may not meet face validity

# Bayesian Shrinkage

DDE allows users to develop Bayesian shrinkage models. Few other names used for Bayesian shrinkage models are mixed effects model, hierarchical regression model, multi-level model and variance components model with random effects. Bayesian shrinkage modeling approach is useful in the following cases:

* There are multiple cross sections and differences across these cross sections
* Relationship in a regression model varies by cross sections
* Practically not feasible to develop hundreds of models in time for business planning

Bayesian shrinkage:

* Fits a regression equation at the cross sectional level
* Allows estimates of the regression equation vary by cross section
* Uses cross sectional level measures to explain variation in the individual-level parameters
* Allows to test for main effects and interactions within and between levels

These models are very strict on the data related assumptions:

* Data needs to be normally distributed
* Means of the data are linear in terms of a certain set of parameters (fixed-effects parameters) -- same as standard linear model
* Variance and covariance of the data are structured and are in terms of a different set of parameters (covariance parameters) -- different from standard linear model
* Random effects parameters are normally distributed. This assumption serves to shrink outlier estimates toward the overall measures. Recent academic work has shown the advantages of using Bayesian Shrinkage Methodology over Market Level Models (Boatwright, McCulloch and Rossi 1999)

**Step I:** Build a model for each cross section separately using time series cross sectional data (Assume we have markets and channels   
within markets)

*yij = β*0*j + β*1i *Xij + εij……i = Markets j = Channels* ……………. (1)

**Step II:** Regression Parameters from step1 are then regressed on channel level data describing the markets process

*β*0*j = γ00*  + *γ01\* market size +*  u0j ……………. (2)

Substituting (2) in (1), we get

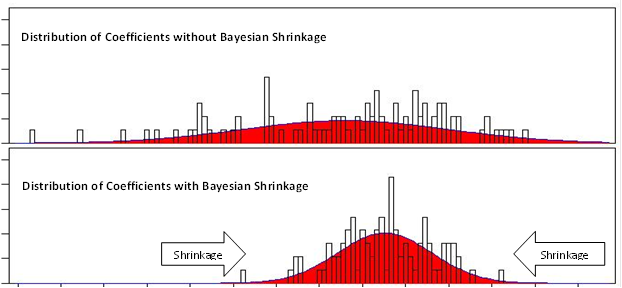
*yij = γ00 + γ01 Xij*  *+u 0j +εij*



**Fixed  
Effect**

**Random Effect**

The platform uses “R” statistical software to run Bayesian shrinkage model.

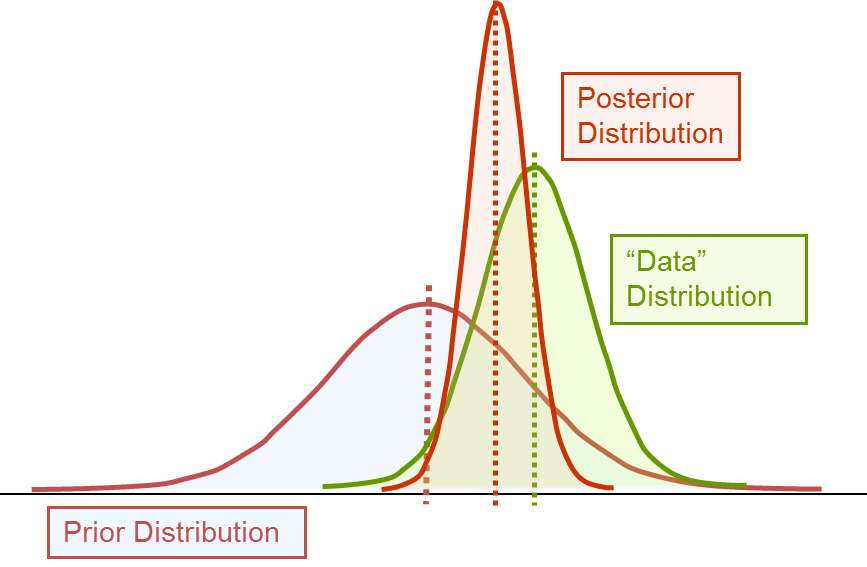


# Bayesian Priors

Often, the dataset at hand contains all the information that is available and needed to create an accurate demand model. Most clients have access to prior information that was distilled from other data sources and should be incorporated into demand models. Bayesian Estimation techniques are the best understood and most consistent tools to incorporate that information.

The platform allows use of Bayesian priors at the measure level to include in the model. Analytic Edge recommends using weighted priors derived by weighting past estimates and estimates from new data. These weighted priors must be within the distribution/standard deviation of the fixed effects estimates from the new data.





**Prior distribution** – probability tendency of an uncertain quantity, β, that expresses previous knowledge of β from, for example, a past experience, with the absence of some proof

* **Posterior distribution** – this distribution takes proof into account and is then the conditional probability of β. The posterior probability is computed from the prior and the likelihood function using Bayes’ theorem.
* **Posterior mean** – the mean of the posterior distribution
* **Posterior variance** – the variance of the posterior distribution
* **Conjugate priors** – a family of prior probability distributions in which the key property is that the posterior probability distribution also belongs to the family of the prior probability distribution

# Estimation

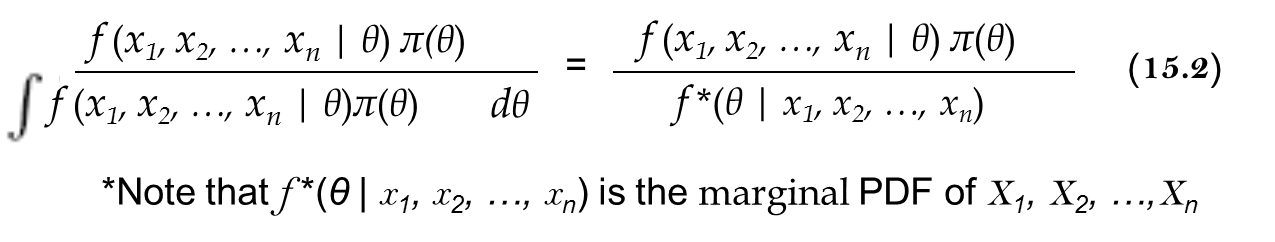
Let θ be an unknown parameter based on a random sample, *x1, x2, …, xn* from a distribution with *pdf/pmf f (x | θ).*

*Let π (θ) be the prior distribution of θ.*

*Let π \*(θ | x1, x2, …, xn)* be the posterior distribution.

\*\*Note that *π \*(θ | x1, x2, …, xn)* is the condition distribution of θ given the observed data, *x1, x2, …, xn.*

If we apply Bayes Theorem (Eq. 15.1), the **posterior distribution** becomes:



Two most commonly used methods of estimation are MAP (Maximum a Posteriori) and MCMC (Markov Chain Monte Carlo). MAP is similar to the MLE method of estimation with an addition of the prior term in the likelihood function. Estimation using MAP is difficult as figuring out gradient function for Newton Raphson are not always easy unless the assumption of conjugate priors is true (not always the case).

MCMC estimation works by taking samples from the posterior distribution of the data given the parameters.



Based on the prior information, MCMC finds out the shape of the distribution and outputs it as the parameter value. This brute force approach is useful when a balance is required between statistical and face validity. DDE uses Gibbs sampling algorithm assuming  is known

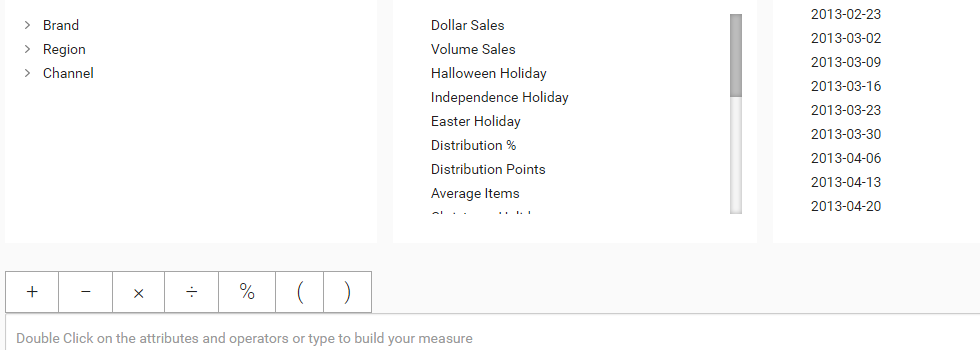
# Coefficient Bounds

DDE provides users bounds to the coefficient estimates. This feature is important to ascertain the face validity of the models. This will ensure that final model always has the right sign and acceptable limits for coefficients. Example - Price Elasticity for a pack size channel is between -0.25 to -4.0. For measures that are being modeled for the first time, Analytic Edge will brainstorm with the client and agree upon the transformation and expected limit for the measure.

# Measure Creation

Typically, before modeling, modelers may want to create measures that can be quickly tested while running iterations. Example – Total level TV by aggregating individual campaigns of TV; adding a week before or a week after holiday, etc. This comes handy when there is multi-collinearity and modelers try to address the challenge of statistical and face validity.

DDE allows users with the flexibility to create new measures by aggregating any of the existing measures across levels of cross sections

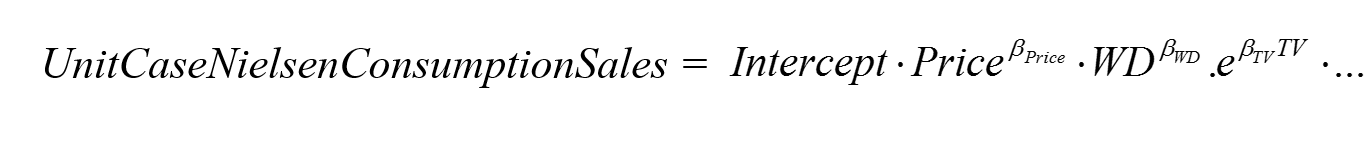


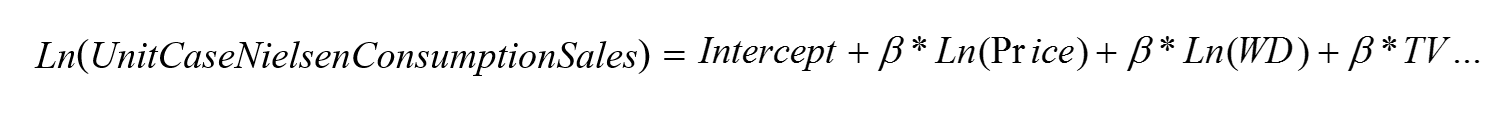
**Model Specifications – TCCC**

# Model Form

TCCC Indonesia requires price and distribution elasticities (in-store drivers) to be generated at the channel levels. For the remaining drivers, results will be reported at the total brand country level. This level of output is critical for TCCC Indonesia since it reduces the complexity while running scenarios for business planning and monitoring. Below is the model form:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Region | Week | Sales Brand 1 | Seasonality | Wt Price Brand 1 | Wt Trade Promotions Brand 1 | TV Brand 1 | Weather | Consumer Promotions Brand 1 |
| Region A | 1 | 10 |  |  |  |  |  |  |
| Region A | 2 | 11 |  |  |  |  |  |  |
| Region B | 1 | 12 |  |  |  |  |  |  |
| Region B | 2 | 13 |  |  |  |  |  |  |



In multiplicative model, we deal with a non-linear equation by taking a log on both sides and make the model linear in “parameters”. The above equation then becomes:

* Dependent measure is at the brand and total country level
* βs are different for each measure in the model
* Other drivers are trade, media, and macro indicators, competition and other base indicators
* Price Elasticities will be reported at the SKU and channel level.
* The impact of distribution will be reported at a channel level

****

* No weights will be applied to drivers that don’t change by pack size

Bayesian shrinkage model will be developed to derive price and distribution elasticities at the channel level.

# Model Measures

Typically, measures to be modeled can be grouped into the following categories:

* **Execution**
* Price & In-Store Promotion
* CPI adjusted Average Price
* Feature Only / Display Only / Feature & Display or any other Trade Activity/Discount variables
* Distribution
* Items Per Store
* Coolers
* Forward Stock
* Weighted Distribution/Numeric Distribution
* Out of Stock
* **Marketing**
* TV GRP’s
* Radio Spends
* Print Spends
* OOH Spends
* Digital Impressions
* Consumer Promotions
  + - Sampling
* Competition
  + Weighted Distribution/Numeric Distribution
  + Items Per Store
  + Price Index on Target Brand
  + Media
* **Calendar**
  + Seasonality
  + Consumption Days
  + Major Holidays/Events
* **Macro Environment**
  + Temperature
  + Precipitation
  + Real Private Consumption
  + # Tourists
* **Category Headwinds/ Category Perception**
  + Future Consumption Intent/CATA/BLS

**Assumptions**

* Price: If it is a new/seasonal/dis-continued pack, there will be blanks. AE will use the average of first four weeks of data available to impute the blanks.
* Media: Saturation parameter of 0.82. Short Term half life of 1.5 months and Long-Term half life of 6 months.
* Weights have been computed using retail sales from 2017.
* Digital Impressions converted to GRPs by dividing by the population
* Non TV Spend converted to GRPs by dividing spend by TV Cost/GRP

# Measure Transformations –

|  |  |
| --- | --- |
| **Measure** | **Transformation** |
| Distribution | Weighted Distribution in % |
| IPS | LN(ItemsPerStore) – Double Log transformation, Negative Exponential |
| Price | LN(Price/ CPI) (Weighted by Channel / Brand weights if Competition Price Factors) |
| Forward Stock | FS/100 (Weighted by Brand Weights Relative to Target) |
| Temperature | (Temperature –Temperature Cycle); (Temperature-Temperature Cycle)\*Temp Cycle Norm |
| Precipitation | (Precip-Precip Cycle)/No of Days Cycle |
| Unemployment | (UnemploymentRate/100) |
| Tourists | ln(#Tourists) |
| TV GRP’s | Delta transformation |
| Non TV (Radio, OOH, Print) Spend | Delta transformation |
| Digital Impressions | Delta transformation |
| Sampling | Delta Transformation |

Analytic Edge MMx platform has many other advanced transformations built-in. These transformations are documented in the marketing science literature and widely used across clients. Following are a few in-built transformations in DDE:

* **Direct** – “Direct” indicates the measure will be used as is in the model without any transformation.
* **Log** – “Log” indicates the natural log of the measure will be used in the model.
* **Lag** – “Lag” indicates the lagged values of the measure will be used in the model. Lag transformation expects user to provide the periodicity to take the lag i.e. one input parameter. If the time is in weeks, user has to provide the number of weeks by which the measure will be lagged. Input parameter values should be greater than 0.
* **Moving Average** – It indicates moving average values of the measure will be used in the model. It takes one input parameter i.e. Period and the value should be greater than 0.
* Backward Weighted Average – It indicates backward moving average of the measure will be weighted by response measure and used in the model. It takes one input parameter i.e. Period and the value should be greater than 0.
* **Forward Weighted Average** – It indicates backward moving average of the measure will be weighted by response measure and used in the model. It takes one input parameter i.e. Period and the value should be greater than 0.
* **Ad-Stock** – Ad-Stock takes one input parameters viz. Decay. Decay parameter takes values between 0 and 1 and the standard values used are 0.1, 0.2, ---- 0.9. Decay parameter value of 0.1 means 10% of the impact of advertisement is observed during the time of execution and remaining 90% of the impact of advertisement is carried forward to remaining weeks.

# Model Quality Criteria

Model quality criteria can be thought of in two broad buckets: statistical validity, and face validity.

Face validity criteria are:

1. Data transformations based on mental models and adhering to global best practices
2. Correct signs on coefficients
3. Elasticities or volume response rates within reasonable ranges
4. Interactions and competitive effects within reasonable ranges
5. Volume contributions within reasonable ranges
6. Change to year ago due to errors of reasonable size
7. Change to year ago due to effect sizes sensible
8. Minimal use of dummy variables and trend terms

Statistical validity criteria are:

1. Low MAPES
2. High R2
3. Small Correction Factor
4. Alignment of predicted to actual, no odd jumps or 0s in predicted volume
5. High confidence in coefficients either via heavy reliance on Bayesian Priors or via high t-stats
6. Forecast accuracy

The statistical criteria are necessary, but not sufficient for model quality. When not all criteria can be met, face validity is valued over   
statistical validity.

# Face Validity: Coefficient Sign

Correct signs on coefficients

The platform has the ability to set bounds for every measure being modeled. This will ensure that final model always has the right coefficients. For measures that are being modeled for the first time, Analytic Edge will brainstorm with the client and agree upon the transformation and expected sign for the measure.

# Face Validity: Elasticities

Elasticities or volume response rates within reasonable ranges

The platform has the ability to set lower and upper bounds for every measure being modeled. This will ensure that final model always has coefficients that are within acceptable limits (For e.g. Price Elasticity for a pack size channel is between -0.25 to -4.0). For measures that are being modeled for the first time, Analytic Edge will brainstorm with the client and agree upon the transformation and expected limit for the measure.

For Coke, Analytic Edge will use the best practice ranges and responses listed out in the model quality criteria v13.0 doc and report the outputs   
as desired.

# Face Validity: Competitive Effects

Interactions and competitive effects within reasonable ranges

For each of the measures used to evaluate the impact of competitive activity, the elasticity or volume lift per increment in the competitive activity will be compared to fair share expected volume change to judge face validity. In addition, for any competitive products new in the marketplace, or those which have grown substantially over the model fit time period, implied cannibalization can be computed and checked for reasonableness. For e.g. Total of volume hit in unit cases should be reasonable relative to fair share and never larger than the new brand volume gain itself.

# Face Validity: Volume Contribution

Volume contributions within reasonable ranges

Analytic Edge will provide a complete atomic decomposition table in Excel including the volume contributions and due-to change reports in power-point. Analytic Edge will maintain a repository of all outputs and overtime build a norm for the respective client and country.

# Face Validity: Due To Change

Change to year ago due to errors of reasonable size

Due-To errors must be small for the business to trust the MMx predictions. Exactly how much error is “small enough” varies depending on the situation. For Coke, we will try and follow the below criteria subject to the acceptable limits for a specific country:

* Error in a Due To is smaller than the total volume change unless the total volume change is less than 0.5%
* Error in a Due-To is smaller than most of the named measure effects
* Error in a Due-To should be less than 2.0%, ideally, especially if model is free of dummy variables

# Face Validity: Year ago change

Change to year ago due to effect sizes sensible

Key measures from baseline and all measures from incremental will be shown in the Due-To YOY change output. Analytic Edge will follow all best practices to not over fit the model and ensure the model error in the Due-To chart and %change values are logically explainable.

# Face Validity: Dummy Variables

Minimum use of dummy variables

MMx analysis is highly recommended for business planning purpose. Analytic Edge does not believe in over beautifying the model with the use of dummy variables and will do so only after the sign off from the client and ensuring all other measures in the model meet the statistical and face validity criteria.

# Statistical Validity: Model Fit

MAPE’s and R2’s

Higher MAPEs are expected for more volatile series, lower volume brands, and for less than ideal data situations. As with MAPES, it is not realistic to make a hard rule about minimum required R2s. One might expect higher R2 for more volatile series and for high volume series; lower R2s are expected for low volume brands, and for less than ideal data situations. MAPE and R2 will be computed on unit cases (not logged) and the goals are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Brand Size** | **Granularity of Periods** | **Yearly MAPE** | **Quarterly MAPE** | **Monthly MAPE** | **R2** |
| Large, >= 20 mil Unit Cases | Weekly | 1.5% | 3% | 6% | 85% |
| Small | Weekly | 4% | 5% | 8% | 80% |

Face validity takes precedence over statistical validity. A high R2 and low MAPES could be achieved with extensive use of dummy variables but such a model would not be suitable for scenario planning purposes. The reports will highlight the MAPE for the last year and compare it with the previous modeling year periods if available.

# Statistical Validity: Correction Factor

Small Correction Factor

Correction Factor is required to correct for the bias introduced because of moving from a log scale to raw scale and occasionally due to some errors in the model fit. It is the sum of predicted volume divided by the sum of actual volume over the time period of the model fit. It will be multiplied with the predicted volume and corrected predicted volume will be reported. Correction factor computed over all model fit time periods is expected to be 1% or less.

To reduce correction factor, Analytic Edge will adapt the following approach:

* Visually inspect the actual versus predicted plot. Look for **spikes** in actual volume where the error is large: these could indicate a promotion or other one-time marketing or execution activity that is left out of the model, or not accounted for appropriately in the current iteration for the model
* Visually inspect the actual versus predicted plot. Look for **runs** in the residuals, i.e. long stretches of time where all the errors are in the same direction. This could indicate a missing driver or data problem
* Visually inspect the actual versus predicted plot; brainstorm with the client and try to find a missing driver or alternative data transformation on existing drivers that can reduce the correction factor

# Statistical Validity: Model Error

Actual, Predicted, Error Time Series

Analytic Edge will visually inspect actual versus predicted plots for all brands individually. There will be a close look for:

* Good alignment of the actual to the predicted volume
* Residuals that appear to be random without runs of positive or negative values or particular patterns, e.g. always missing on high volume periods.
* No odd jumps in the predicted volume, which can be caused by incorrect data or variable transformations

The Actual versus Predicted reporting deliverable is an Excel file at the time period level with raw data plus a pivot table and chart output





# Statistical Validity: Confidence

Coefficient Confidence

In regression models, P-values indicate the probability that a coefficient is not different from zero. P-values are less relevant when Bayesian statistical approaches are used. Modelers may sometimes fix a coefficient to a prior and in such cases there will not be any P-value reported from the Bayesian solving algorithm

Analytic Edge will provide the following during the model review:

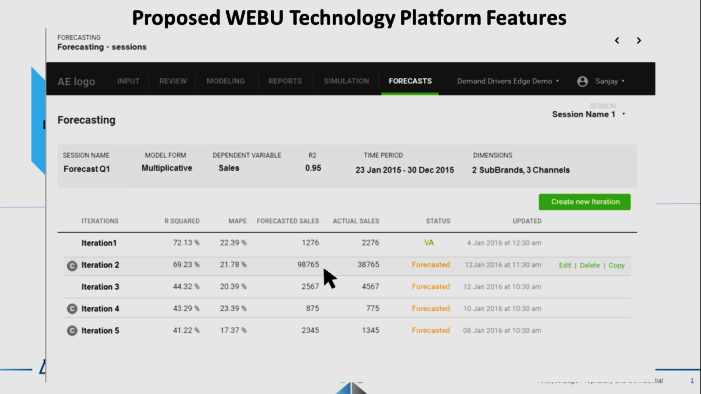
1. Coefficient, P-values and VIF for every measure in the model
2. Highlight of the coefficients that were forced to the bounds due to data related input

# Statistical Validity: Forecast Accuracy

Forecasting Accuracy

MMx analysis should not be looked at as a one off analysis. It should be periodically calibrated and used to estimate volume and profit under alternative scenarios and to recommend business actions needed to achieve volume and profit goals.

The platform offers a business planning module. This module offers the ability to load incremental data, generate scenarios & forecasts, make assumptions and track the model performance on a periodic basis. The platform will output the difference between actual sales and forecasted sales and report the forecast accuracy. The goal is to scrutinize this forecast accuracy periodically and take corrective measures to ensure model is able to capture as many data driven nuances as feasible.

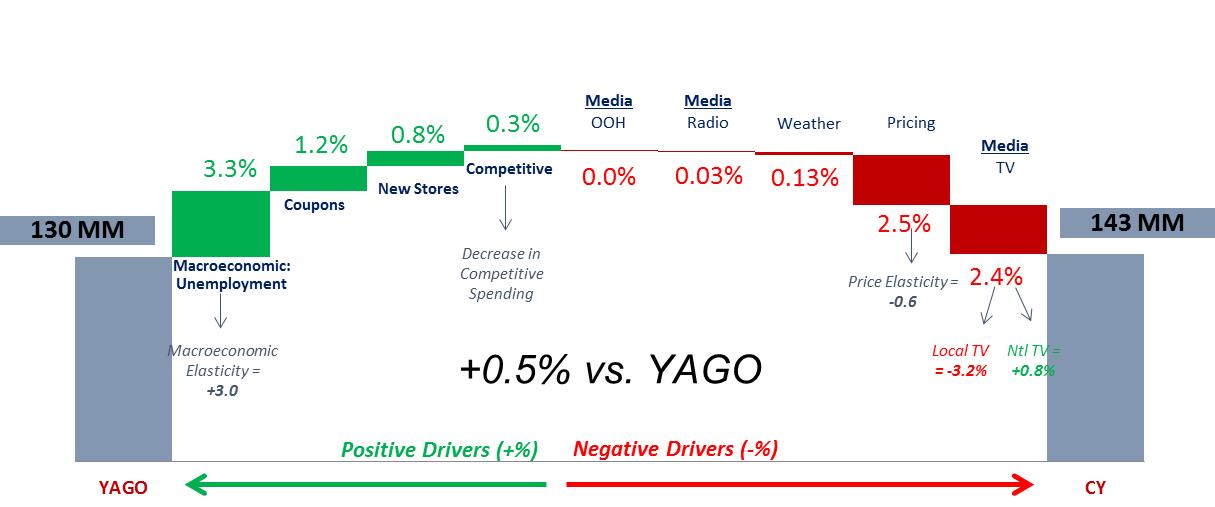


# Due-To Change YOY

Contribution from incremental measures is mathematically easier by adjusting for correction and using subtractive approach. However, for baseline measures, the math usually does not add up in multiplicative models. So, contribution from baseline measures is usually never reported. Instead, due-to change YOY is reported and specific treatment is applied to calculate due-to for baseline measures.

Math used for base measures due-to change calculation:

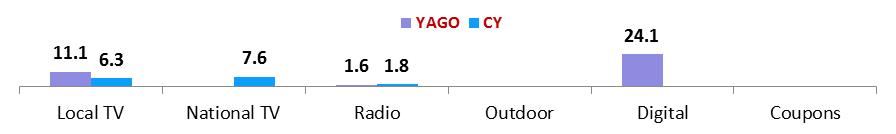
* User chooses CY and YAGO period
* For each time period in the “CY” period, % change in baseline measure from the YAGO baseline measure is calculated and summed across   
  all weeks
* In many cases, this %change in each measure between the CY and YAGO does not add up to 100%. In such cases, to align the waterfall chart correctly, the gap or additional sales should be grouped under the bucket “Model Error”. It can be increasing or decreasing based on the adjustment required
* The models have to do a good job of keeping model error as low as possible
* 

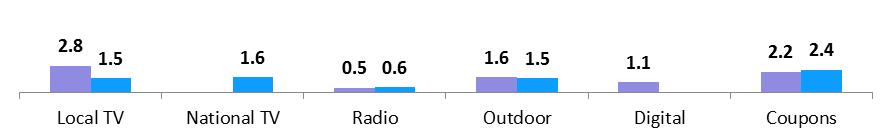


# Media Effectiveness and Efficiencies

Effectiveness is calculated as Incremental sales divided by the support for the respective media measure. Efficiency is calculated as

Efficiency is calculated as Incremental Revenues divided by spend for the respective media measures. In addition, the report also provides ROI which is calculated as Efficiency times gross profit margin





# Response Curves

Response Curves have implications in planning media execution strategy. They help understand the current spend versus return level and also indicates the optimal range of spend to obtain an optimal return before diminishing returns sets in.

