EasyChair Preprint

№ 2097

# Contact center productivity improvement based on predictive analytics

Rishabh Tyagi and Sandeep Bhattacharya



December 5, 2019

# Contact center productivity improvement based on predictive analytics

Rishabh Tyagi
Symbiosis Institute of Business Management, Pune
rishabh.tyagi20@associates.sibmpune.edu.in

Sandeep Bhattacharya
Symbiosis Institute of Business Management, Pune
sandeepbhattacharya@sibmpune.edu.in

***Abstract:*** ***While the existing research has established various methods for contact center business optimization through Erlang models, this study aims at using predictive analytics to improve the performance of important KPIs measuring the effectiveness and efficiency of agents. This will help the contact center resource pool managers to take better informed decisions through scenario building while making future assumptions for metrics like occupancy of agents, shrinkage, Average Handle Time (AHT) and number of agents both at tactical and strategic level. The overall performance of a contact center is often measured through its ASA (Average Speed of Answer), abandonment rate and Service Level (SL). This study focuses on the predicting ASA as a function of call volumes, AHT, occupancy of agents, number of productive Full Time Equivalents (FTE) and Off Phone Activities % (OPA) and analyzing the impact of each of these parameters on ASA through sensitivity analysis. The sensitivity analysis can be transformed into a web interactive tool which can be used by staff planners for scenario building. They will not only be able to plan for meeting the demand (in load minutes) but also be able to keep the ASA within tolerance limits.***

***Keywords:*** **Average Speed of Answer, contact center, capacity planning, Time series analysis.**

## I. INTRODUCTION

In a contact center the mix of calls entering the system and the arrival pattern varies with time. The minute-by-minute call arrival shows stochastic variability. However, the daily, weekly and monthly call volumes show predictability. The service capacity in a contact center is not inventoried. Hence, the resource pool managers do capacity planning to have optimal level of service capacity which can meet the demand and match the variability in arrival rate. This will help in reducing the cost as well as keeping the waiting time (ASA) within acceptable limits. The *queuing* models decide how many agents will be present to handle the calls at an hourly level. The *scheduling* models determine which agents will work for which day(s) of a week for a given month. The *hiring* models determine how many agents need to be hired and trained at a monthly or quarterly level. The Work Force Management (WFM) systems often schedule the agents or 'Reps' to work in every 15 mins or 30 mins. The skill based routing ensures that the specialized calls are routed to full time and well trained agents while general calls are routed to part time or less trained agents. The existing research on capacity planning involves the usage of Erlang C (stationary system), Erlang B (system with busy signals) and Erlang A (system with abandonments) models along with enhancements to these models. The Erlang models provide the number of representatives required to meet the demand for a given arrival rate, service rate and service level. They also provide values for ASA. However, Erlang models operate with assumptions like fixed arrival rate, no shrinkage and no off phone activities of agents, thereby tend to underestimate the ASA values for a large contact center. The primitives used in queuing models vary systematically over the period of time. This holds true especially for arrival rate which shows regular trend and seasonality. The sources of systematic variation in call volumes and their impact on ASA should be analyzed. For example, a study by Gustafson (1982) incorporated the impact of "learning curve" of agents in service delivery model. As the agents gain experience, their service rate increases. Sze (1984) explains the "shift fatigue" among the agents i.e. initially the agents serve faster to get rid of overload but sustained overload reduces their productivity and increases the service time. There could be systematic variations because of annual events, for instance, new enrollments for scheme in the last quarter of the year. Hence there is a requirement of a model which can explain the interdependencies among the operational KPIs of the contact center and also able to capture the systematic seasonal variations in the call volumes, service rate and waiting time or ASA in a large contact center. This research work aims to identify the parameters that affect the ASA in a large contact center and also to predict ASA as a function of these parameters using multi-variate time series forecasting. Post modeling, sensitivity analysis has been done to measure the ASA movement (in seconds) with k % (k takes integral values from -10 to 10) in the input variables. This sensitivity analysis tool can be used by resource pool managers to adjust the input variables of the model as per the load demand and also keep the ASA within tolerance range (as per service level targets).

## II. LITERATURE REVIEW

Capacity planning becomes complex in Quality and Efficiency Driven (QED) regime for multiple inter-connected contact centers, with cross-trained and geographically dispersed agents who attend to time-varying call loads from multiple types of customers. In QED regime, the delay in system is neither close to One (Efficiency regime) nor close to Zero (Quality regime). Work Force Management (WFM) software applications are designed to support the staff planners to find optimal service capacity. The contact centers often tradeoff accessibility of agents with their utilization. Higher the resource utilization, lesser is the accessibility. In some cases, the costs or revenues can directly be associated with the efficiency of the system. The higher waiting time or

ASA can lead to increased costs due to breach of Service Level (SL) targets. In case of "order taking" systems, higher abandonment or busy signals can lead to opportunity cost due to loss of sales (Andrews and Parsons, 1993; Aksin and Harker, 2003). Borst et al (2000) discusses the cost minimization approaches and constraint satisfaction which holds true in case of contact centers. In contact center, the upper management decides the service levels, and the resource pool managers have to defend their budgets for the given accessibility of agents and service levels. The capacity planning in contact centers follows bottom up approach (Buffa et al, 1976). As shown in Figure-1, call arrival at the intra-day level show stochastic variation while the daily and monthly level arrival rates some degree of predictability (due to seasonality and trend). Hence, in this research the prediction models have been created for daily and monthly level. At the lowest level in hierarchy (at half-hourly level), M/M/N (Erlang C) queuing model is applied to achieve stationary performance of system. It takes several assumptions like arrival rate follows Poisson distribution, service rate is exponentially distributed and system achieves steady state every time for a given half-hour interval. Calculations of Erlang C model are as follows. Let $\lambda_i$ be the arrival rate for $i^{th}$ interval, expected service time be E $[S_i]$ and service rate $\mu_i = E[S_i]^{-1}$ then load $(R_i) = \lambda_i / \mu_i$ and system occupancy $\rho_i = \lambda_i / N \mu_i$, for N agents. Here, N must be greater than $R_k$ to have steady state. So, at least $R_i$ Erlangs are required to meet the load demand. Erlang C formula for steady state is

$$C(N,R_i) \triangleq 1 - \frac{\sum_{m=0}^{N-1}(R_i^m/m!)}{\sum_{m=0}^{N-1}(R_i^m/m!)+(R_i^N/N!)(1/(1-R_i/N))} \quad (1)$$

Given that the arriving calls must wait, then the ASA calculation from Erlang C model is as follows

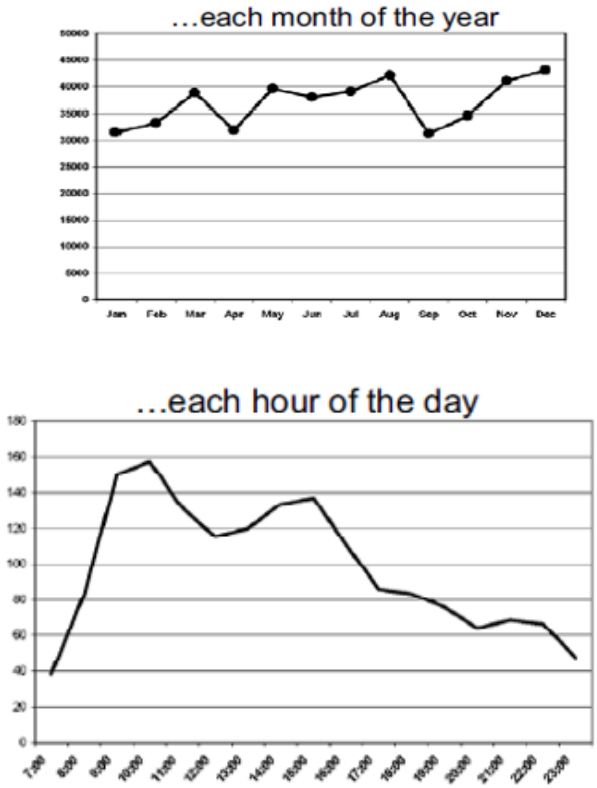
$$\begin{aligned} ASA \triangleq E[Wait] &= P\{Wait>0\}\cdot E[Wait\,|\,Wait>0] \\ &= C(N,R_i)\cdot\left(\frac{1}{N}\right)\left(\frac{1}{\mu_i}\right)\left(\frac{1}{1-\rho_i}\right), \end{aligned} \quad (2)$$

Figure-2 shows the relationship between the ASA and system occupancy ρ for a small contact center. The plot for aggregated data is in line with the relationship explained in (3). Hence, in this study occupancy of the agents has been taken as an input variable for the ASA prediction model. The Erlang B (M/G/N/N) system eliminates delays by "blocking". The number of telephone lines equal the number of representatives (Reps) unlike Erlang C system where the space (number of trunk lines) is infinite. One trades off delays with blocking in between Erlang B and Erlang C models. Increase in space can lead to delays but will reduce the busy signals. According to Feinberg (1990), if the number of telephone lines exceed the number of Reps by 10%, then system performance improves significantly. The Erlang A model (M/M/N/k+G queue) include busy signals with abandonment. In this model, a patience level for waiting is specified beyond which the customer drops the call. The special kind of Erlang A model (M/M/N/k+M queue) has exponential distribution of patience in call center (Garnett et al, 2002). Erlang A model demonstrates how a heavily loaded call center functions.

It is important to create models for predicting future call arrival rates/volumes as a pre-cursor to personnel scheduling. As shown in Figure- 1, call arrival in the lower right panel shows stochastic variability while the call volumes when aggregated at a daily level or monthly level show fluidlike predictability (Mandelbaum et al, 2001). For forecasting of explanatory variables like call volumes it is important to consider factors like- day of week, time of day, week of month, quarter of the year, holidays, etc. Andrews and Cunningham (1995) used Auto Regressive Integrated Moving Average (ARIMA) for forecasting daily calls (for order purchase and inquiry) at L.L Bean. Though there exists research on predicting arrival rates brown et al 2002a, Massey et al 1996, Jongbloed and Koole 2001) still there is scope for further improvement by using covariates that are capable of capturing Poisson randomness.

The process of forecasting average waiting time (or ASA) is also critical. Hops and Sturgis (2001) explain how the service level constraint affect the point forecast of the delay or ASA. It also used simulation technique to test the robustness of the results from single-server systems on multi-server systems. Whitt (1999b) developed several First-Come-First -Serve (FCFS) systems to study the delays. There is scope for better estimation of ASA in time varying systems. This research work provides methodology for better ASA prediction through ML algorithms and ensemble method. It uses ARIMAX and feed forward neural network predictor for multi-variate time series forecasting and uses weighted average (based on accuracy of individual algorithm output) approach for ensembling. The ensemble model with the best accuracy is chosen. Sensitivity analysis is performed on the final model to observe the movement in ASA (in seconds) with change in model inputs. The prediction model and sensitivity analysis tool can be used for scenario building by changing input variables and thereby will help the resource pool managers in scheduling the agents at a monthly and weekly level.

**Number of calls arriving**



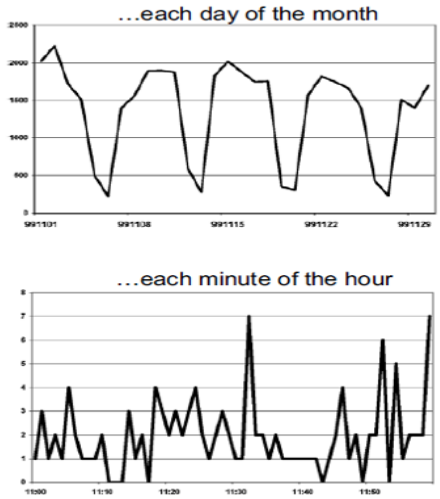



**Figure 1 - View of call arrival rates at different intervals (taken from Mandelbaum et al., 2001; Buffa et al., 1976)**

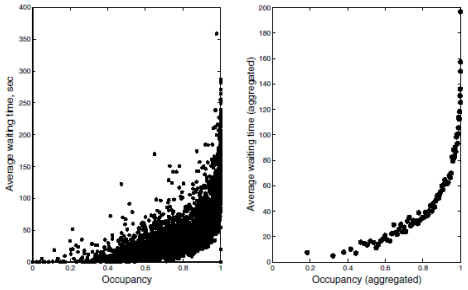



**Figure 2-Relationship between system occupancy and ASA for Raw and Aggregated data (from Brown et al, 2002a)**

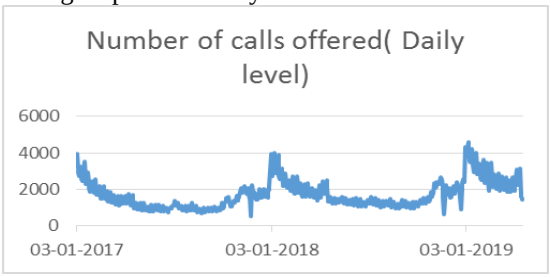
| Month | Average of NCO | Average of Queue Sec | AHT(in Sec) | Average of Occupancy | Average of OPA | Average of Prod_FTE | Average of ABA% |
|---|---|---|---|---|---|---|---|
| Jan-17 | 2677 | 37171 | 444 | 70% | 14% | 75.5 | 0.76 |
| Feb-17 | 1813 | 10578 | 442 | 60% | 20% | 73.9 | 0.36 |
| Mar-17 | 1372 | 3599 | 437 | 52% | 22% | 72.1 | 0.29 |
| Apr-17 | 1279 | 3057 | 413 | 48% | 22% | 71.4 | 0.33 |
| May-17 | 967 | 3324 | 407 | 53% | 25% | 68.3 | 0.29 |
| Jun-17 | 947 | 5013 | 387 | 60% | 26% | 63.2 | 0.48 |
| Jul-17 | 1030 | 14522 | 413 | 86% | 23% | 57.8 | 0.68 |
| Aug-17 | 892 | 11460 | 414 | 77% | 24% | 51.5 | 0.69 |
| Sep-17 | 904 | 3888 | 406 | 56% | 25% | 51.2 | 0.28 |
| Oct-17 | 1299 | 17710 | 424 | 75% | 26% | 51.2 | 0.78 |
| Nov-17 | 1802 | 25489 | 442 | 69% | 23% | 65.5 | 0.75 |
| Dec-17 | 1730 | 21651 | 435 | 66% | 21% | 79.4 | 0.57 |
| Feb-18 | 2101 | 54729 | 399 | 77% | 25% | 78.4 | 1.09 |
| Mar-18 | 1768 | 38527 | 408 | 66% | 30% | 76.2 | 1.12 |
| Apr-18 | 1739 | 43142 | 398 | 63% | 27% | 73.0 | 0.90 |
| May-18 | 1350 | 17998 | 381 | 56% | 29% | 70.4 | 0.75 |
| Jun-18 | 1234 | 24726 | 387 | 71% | 34% | 66.2 | 0.96 |
| Jul-18 | 1268 | 26535 | 384 | 78% | 32% | 61.9 | 0.96 |
| Aug-18 | 1138 | 17650 | 370 | 71% | 33% | 59.5 | 0.84 |
| Sep-18 | 1125 | 20725 | 375 | 72% | 25% | 56.4 | 1.00 |
| Oct-18 | 1389 | 30400 | 404 | 72% | 25% | 52.5 | 1.12 |
| Nov-18 | 2065 | 150493 | 430 | 87% | 25% | 59.9 | 2.98 |
| Dec-18 | 1996 | 167857 | 476 | 87% | 27% | 68.2 | 3.47 |
| Jan-19 | 3492 | 303091 | 428 | 83% | 16% | 76.9 | 3.66 |
| Feb-19 | 2637 | 97923 | 418 | 80% | 21% | 79.5 | 1.44 |
| Mar-19 | 2177 | 37429 | 410 | 63% | 24% | 77.5 | 0.75 |
| Apr-19 | 2262 | 139717 | 428 | 72% | 21% | 72.8 | 2.59 |

## III. RESEARCH METHODOLOGY

An ensembled model was created to predict the ASA as a function of call volume, AHT, occupancy of agent, off phone activity percentage and number of agents available during a specific time period.

### 3.1. Population and Sample

Sample data was collected for the period of 3 years (Jan -2017 to Apr-2019) from a contact center run by a US based company. The data was collected at an intra-day level (at 30 mins interval). 80% of the dataset was used as training set and remaining 20% was used as test set. Data related to Number of Calls Offered (NCO), Average Handle Time (AHT), Occupancy % of agents, Number of productive Full Time Equivalents (FTEs), Off Phone Activities (OPA), queue seconds, abandonment seconds and abandonment % (ABA). The following table shows the average values for primitive variables from the sample data. The table has been created by rolling up the daily level data to monthly level.



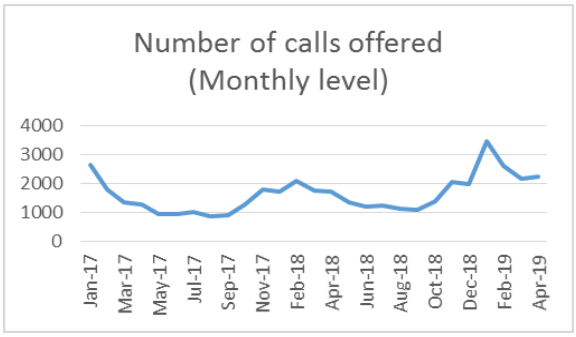




**Figure 3- Call Arrival distribution at daily and monthly level for sample data**

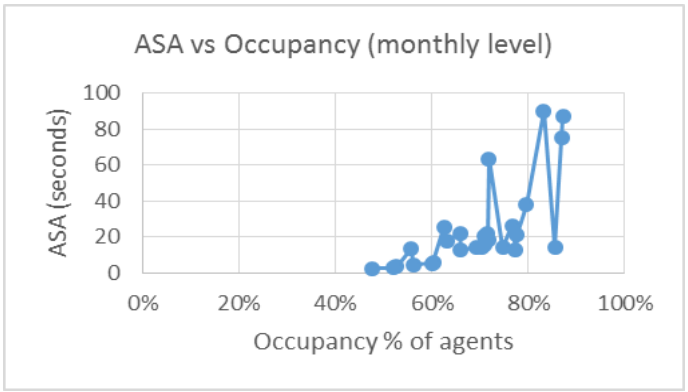



**Figure 4- Relationship between system occupancy and ASA at monthly level for sample data**

### 3.2. Theoretical framework

3.2.1. *Pre-Modeling*

Rather than predicting ASA based on historical data, queue seconds were predicted using predictive modeling. The predicted queue seconds were divided by Number of Calls Handled (NCH) to get the ASA values. For a given NCO and ABA%, $NCH = NCO - (NCO * ABA)/100$.

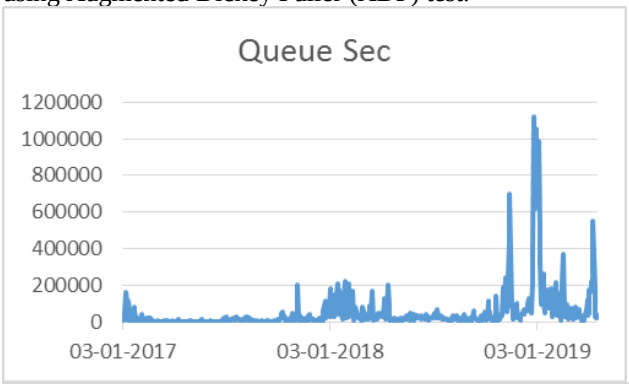
*Model Inputs:* - The model establishes the following relationship

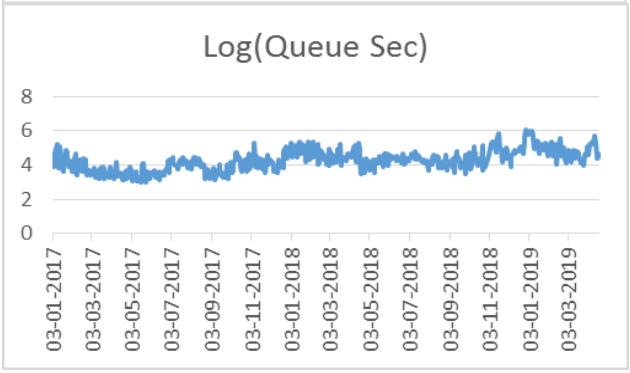
$$Queue\ seconds = f(NCO, AHT, ProdFTE, Occupancy\%, OPA\%)$$

NCO-Number of Calls Offered
AHT-Average Handle Time (Talk time + Hold time +Wrap time)
ProdFTE- Number of Productive Full Time Equivalents.
Occupancy % - Percentage of the logged in time an agent is busy in handling calls.
OPA % - Percentage of the logged in time an agent is busy in necessary off phone activities like taking down notes, training, breaks, etc.
Queue seconds – The total waiting time in seconds during a particular period.

3.2.2 *Modeling*

*Stationarity:* - The queue seconds time series data was transformed using Box-cox transformations to make it stationary. Stationarity of the queue second data was tested using Augmented Dickey Fuller (ADF) test.




```
        Augmented Dickey-Fuller Test

data:  hist_data$dfq
Dickey-Fuller = -4.3518, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary
```

**Figure 5- Actual Queue seconds, Transformed Queue seconds and result of ADF test on transformed Queue Seconds**

*Multi-Collinearity:* - Multi-Collinearity check was done by finding pair-wise correlation between independent variables.

*Hyper-parameter tuning:* - For ARIMAX algorithm, the order of hyper-parameters (p, d, q) was decided using Grid-Search method. A grid was created where in $0 \leq p \leq 4$ ,$0 \leq d \leq 2$, $0 \leq q \leq 4$ and through multiple iterations the best fit model with optimal hyper-parameters and the best prediction accuracy was chosen.

For Neural Network (NN) predictor, the number of hidden layers were chosen through Grid-Search.

*Seasonal Dummies for ARIMAX model:* - Seasonal dummies like time of the date, day of week, week of month, month, quarter of year and year were included in external regressor parameter of ARIMAX model. This will help in capturing the daily, weekly and monthly seasonality more effectively.

3.3.3. *Post- Modeling*

The results from both ARIMAX and NN predictor are collected and aggregated at daily, weekly and monthly level. Mean Absolute Percentage Error [ $MAPE = (Actuals - Forecast) * 100/Actuals$ ] is calculated at each level. The outputs from both the algorithms are combined using weighted average in order to reduce MAPE.

## IV. RESULTS AND DISCUSSION

For model training, data from January 2016 to December 2018 was used. The data from January 2019 to March 2019 was used as test data for evaluating model accuracy.

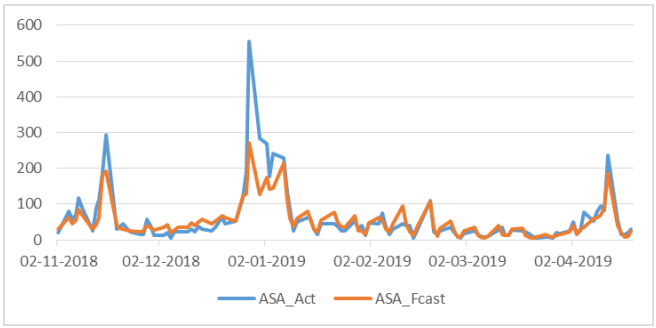



**Figure 6- Daily level comparison between ARIMAX ASA Forecast and ASA actuals for test data. MAPE-36%.**

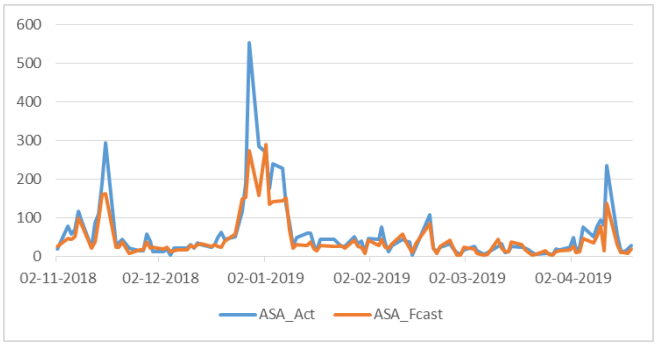



**Figure 7- Daily level comparison between Neural Network ASA Forecast and ASA actuals for test data. MAPE-42%.**

Ensembled Model is created by taking weighted average of ARIMAX output (64%) and NN predictor output (36%). This helps in reducing the daily level MAPE to 32%

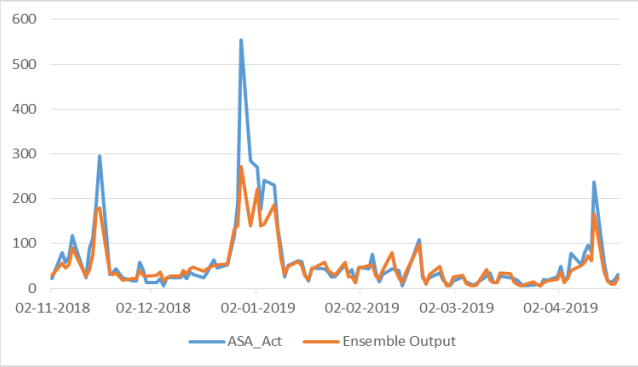



**Figure 8- Daily level comparison between Ensembled Model ASA Forecast and ASA actuals for test data. MAPE-32%. All ASA values are in seconds.**

| Jan | ASA actuals | ASA Forecast | Deviation |
|---|---|---|---|
| Mon | 123 | 125 | 1% |
| Tue | 72 | 70 | 2% |
| Wed | 105 | 73 | 43% |
| Thu | 61 | 57 | 6% |
| Fri | 99 | 76 | 30% |
| | | MAPE | 17% |

| Feb | ASA actuals | ASA Forecast | Deviation |
|---|---|---|---|
| Mon | 42 | 72 | 41% |
| Tue | 67 | 62 | 9% |
| Wed | 28 | 24 | 18% |
| Thu | 10 | 14 | 26% |
| Fri | 33 | 40 | 17% |
| | | MAPE | 22% |

| Mar | ASA actuals | ASA Forecast | Deviation |
|---|---|---|---|
| Mon | 23 | 31 | 27% |
| Tue | 21 | 12 | 68% |
| Wed | 12 | 9 | 36% |
| Thu | 13 | 9 | 40% |
| Fri | 17 | 18 | 9% |
| | | MAPE | 36% |

**Figure 9-Weekly level comparison between actual ASA and Forecasted ASA (from Jan-2019 to Mar-2019) for ensembled model. All ASA values are in seconds.**

$$MAPE_{weekly} = (APE_{mon} + APE_{Tue} + APE_{wed} + APE_{Thu} + APE_{Fri})/5$$

| Month | ASA actuals | ASA Forecast | Deviation |
|---|---|---|---|
| Jan | 90 | 78 | 16% |
| Feb | 38 | 44 | 13% |
| Mar | 17 | 17 | 3% |
| | | MAPE | 11% |

**Figure 10- Monthly level comparison between actual ASA and forecasted ASA for ensembled model**

$$MAPE_{monthly} = APE_{jan} + APE_{feb} + APE_{mar}$$

***Sensitivity Analysis***

To do sensitivity analysis using ensembled model, monthly baseline assumption data is created from the *forecasted values* of these primitives from Jan-2019 to Mar-2019. In the baseline data, NCO, AHT and number of Agents are taken as monthly average values. The occupancy of the agents and OPA% are assumed to be 75% and 20% respectively for all three months in the monthly baseline data.

| | Monthly Baseline Assumptions | | |
|---|---|---|---|
| Metrics/ Months | Jan | Feb | Mar |
| Fcast NCO | 5890 | 8874 | 7930 |
| Fcast AHT | 457 | 578 | 496 |
| Assumed Agents | 264 | 286 | 305 |
| Assumed Occupancy% | 76% | 76% | 76% |
| Assumed OPA% | 20% | 20% | 20% |
| Fcast ASA(seconds) | 21 | 42 | 45 |

**Figure 11-Monthly Baseline Assumptions from forecasted data for call center primitive variables**

| | Jan_ASA_delta | | Feb_ASA_delta | | Mar_ASA_delts | |
|---|---|---|---|---|---|---|
| Metrics | 10% | -10% | 10% | -10% | 10% | -10% |
| NCO | 18 | -12 | 35 | -12 | 28 | -15 |
| AHT | 13 | -8 | 19 | -26 | 17 | -13 |
| # Agents | -16 | 21 | -28 | 29 | -14 | 22 |
| Occ %( changed from 76%) | -19 | 15 | -18 | 11 | -12 | 16 |
| OPA% (changed from 20%) | 15 | -18 | 16 | -17 | 9 | -12 |

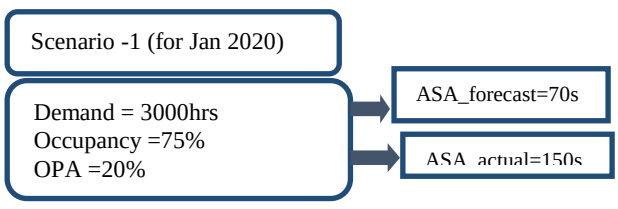
**Figure 12-Sensitivity Analysis matrix created using monthly baseline assumption and ensembled model**

*Reading the Sensitivity Matrix- example*

If NCO for the month of Jan changes by +10% over the monthly baseline value for Jan-2019 (5890 calls), keeping rest other variables (AHT, occupancy, number of agents and OPA%) as constant , then the ASA will move by +15 seconds.

***Practical application of Sensitivity Analysis***

The above sensitivity analysis matrix can be converted into a tool where all the model inputs –NCO, AHT, number of agents, occupancy% and OPA% can be changed from -10% to +10% (for instance) and the impact of change in any of these primitives on ASA (in seconds) can be captured. This type of sensitivity analysis tool is of great value for resource pool managers for scenario building. They can adjust the model inputs while making monthly staff plans to not just meet the demand (load mins = NCH * AHT) but will also be able to keep the ASA within tolerance range (say -10s to +10s) as per service level target.



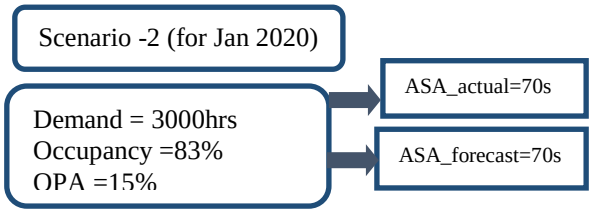



**Figure 13-Sample Scenario Building for meeting demand as well as keeping actual ASA close to forecasted ASA values**

## V. CONCLUSION

For a large contact center, the productivity can be achieved through business optimization. It involves capacity planning to meet the call load, scheduling and rostering of representatives. The resource pool managers have to make assumptions about the contact center primitives like number of Reps, AHT, occupancy, and shrinkage of Reps. But, they often miss out acceptable ASA limit because of the lack of proper ASA forecasting technique. A contact center's performance is often expressed in terms of Average Speed of Answer (ASA) and abandonment percentage. Through ensemble model used in this study, the contact center would be able to predict ASA more accurately at a daily, weekly and monthly levels. It will also help in understanding the impact of call volumes, occupancy of agents, shrinkage/Off Phone Activities (OPA) and number of productive FTEs on the ASA values. The sensitivity analysis tool developed using the ensemble model will help the resource pool managers in scenario building and arriving at optimal values of key call center primitives.

## VI. ABBREVIATIONS

| | |
|---|---|
| NCO | Number of Calls Offered |
| NCH | Number of Calls Handled |
| AHT | Average Handle Time (in seconds) |
| ASA | Average Speed of Answer (in seconds) |
| OPA | Off Phone Activity |
| Prod_FTE | Number of Productive Full Time Equivalents |
| ABA | Average Abandonment percentage |

## VII. REFERENCES


1. Ak¸sin, O. Z., P.T. Harker, 2003. Capacity sizing in presence of common shared resource: dimensioning an inbound call-center. *Eur. J. Oper. Res*. **147** 464-483.
2. Brown, R. G. 1963. *Smoothing, Forecasting and Prediction of Discrete Time Series*. Prentice-Hall, Englewood Cliffs, NJ.
3. Andrews, B., S. M. Cunningham. 1995. L.L. Bean improves call center forecasting. *Interfaces* **25(6)** 1-13.
4. Atar, R., A. Mandelbaum, M. Reiman. 2002. Scheduling a multi class queue with many i.i.d. servers: asymptotic optimality in heavy traffic, working paper, Technion, Haifa, Israel
5. Borst, S.C., P. Seri 2000. Robust algorithms for sharing agents with multiple skills. Working paper, CWI, Amsterdam, Netherlands.
6. Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, L. Zhao. 2002b. Empirical analysis of a network of retail-banking call centers. Work in progress, The Wharton School, University of Pennsylvania, Philadelphia, PA.
7. Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, L. Zhao. 2002a. Statistical analysis of a telephone call center: A queueing science perspective. Working paper, The Wharton School, University of Pennsylvania, Philadelphia, P.A
8. Buffa, E.S., M.J. Cosgrove, B.J. Luce. 1976. An integrated work shift scheduling. *Decision Sci*. **7** 620-630.
9. Chen, H., D. D. Yao. 2001. *Fundamentals of Queueing Networks*. Springer-Verlag, New York.
10. Eick, S. G., W. A. Massey, W. Whitt. 1993a. The physics of the M*t*/G/_ queue. *Oper. Res.* **4**1 731–742.
11. Eick, S. G., W. A. Massey, W. Whitt 1993b. The M*t*/G/_ queue with sinusoidal arrival rates. *Management Sci.* **39** 241–252
12. Erlang, A.K. 1948. On rational determination of number of circuits. E. Brockmeyer, H.L.
13. Halstrom, A. Jensen, eds. *The Life and works of A.K. Erlang*. The Copenhagen Telephone company, Copenhagen, Denmark
14. Evenson, A., P.T. Harker, F. X. Frei. 1998. Effective call center management, evidence from financial services. Working paper 99-25-B, Wharton Financial Institution Center, University of Pennsylvania, Philadelphia, P.A
15. Feinberg, M. A. 1990. Performance characteristics of automated call distribution systems. *Proc. IEEE GLOBECOM '90*, San Diego, CA. IEEE, New York, 415–419.
16. Gustafson, H.W. 1982. Force loss cost an*alysis. In W.H. Mobley Employee turnover: causes, consequences and control*. Adison Wesley, Reading, M.A.
17. Garnett, O., A. Mandelbaum., 2001. An Introduction to skill based routing and its operational complexities
18. Jongbloed, G., G.M Koole. 2001. Managing uncertainties in call centers using Poisson mixtures. *Appl. Stochastic Models in Bus. Indus.* ***17*** *307-318*
19. Jelenkovic, P., A. Mandelbaum, P. Momcilovic. 2002. The GI/D/*N* queue in the QED regime. In preparation. Jennings, O. B., A. Mandelbaum, W. A. Massey, W. Whitt. 1996. Server staffing to meet time-varying demand. *Management Sci.***42** 1383–1394.
20. Quinn, P., B. Andrews, H. Parson 1991. Allocating telecommunication resources at L.L. Bean, Inc. *Interfaces* **21** 75-91.
21. Sze, D.Y. 1984. A queuing model for telephone operator staffing. *Oper. Res 32 229-249*