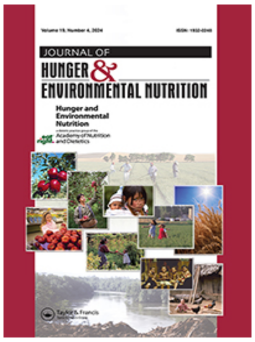

Taylor & Francis
Taylor & Francis Group

Journal of Hunger & Environmental Nutrition

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/when20

Predicting Food Distribution with Underlying Factors for a Hunger Relief Organization

Kehinde Odubela, Henry Ivuawuogu, Steven Jiang & Lauren Davis

To cite this article: Kehinde Odubela, Henry Ivuawuogu, Steven Jiang & Lauren Davis (15 Mar 2024): Predicting Food Distribution with Underlying Factors for a Hunger Relief Organization, Journal of Hunger & Environmental Nutrition, DOI: 10.1080/19320248.2024.2328353

To link to this article: https://doi.org/10.1080/19320248.2024.2328353

Published online: 15 Mar 2024.

JOURNAL OF HUNGER & ENVIRONMENTAL NUTRITION
https://doi.org/10.1080/19320248.2024.2328353

ORIGINAL RESEARCH

Predicting Food Distribution with Underlying Factors for a Hunger Relief Organization

Kehinde Odubela, Henry Ivuawuogu, Steven Jiang, and Lauren Davis

Department of Industrial and Systems Engineering, North Carolina A&T State University, Greensboro, NC, USA

ABSTRACT
Hunger relief organizations often estimate food demand using food distribution data. Leveraging Visual Analytics (VA) and historical data, we examine how underlying factors like unemployment, poverty rate, and median household income affect forecasts for aid recipients' food demand. Our study reveals that incorporating these factors enhances forecast accuracy. Visual Analytics empowers decision-makers to integrate field knowledge with computational insights, enabling more informed decisions. This innovative approach presents a valuable tool for charitable organizations to strategically improve forecasting precision in the dynamic landscape of hunger relief.

KEYWORDS
Food distribution; forecasting; underlying factors; visual analytics; hunger relief

Introduction

Food insecurity as defined by the United States Department of Agriculture (USDA) is the inability for an individual or household to have access to nutritious food. Individuals and households experience food insecurity when they have limited access to safe and nutritious food they need for an active, healthy life.[1;2] Food insecurity encompasses not only the physical availability of food but also considers factors such as accessibility, utilization, and stability. The Economic Research Service (ERS) of the (USDA) estimated that 87.2% of U.S. households were food secure throughout the entire year in 2022.[3] Thus, this percentage of the population always had access to enough food for an active, healthy life for all members of their households. However, the remaining 12.8% of households were food insecure, at least some time during the year. These include 5.1% that experienced very low food security, that is, one or more times, the food intake of the household members was reduced, and their eating patterns were disrupted. This is mostly because they lacked money and other resources for obtaining food due to poverty, unemployment, stagnant wages and rising cost of living.[3]

Food banks play a vital role in distributing nutritious food to those in need through dynamic distribution networks, facing complexities in supply

CONTACT Steven Jiang xjiang@ncat.edu Department of Industrial and Systems Engineering, North Carolina A&T State University, 1601 E Market St, Greensboro, NC 27411, USA

and demand uncertainties.[4–6;7;8] They also prioritize cultural sensitivity in ensuring that the foods available align with the diverse preferences of the communities they serve.[9] Capturing food demand data proves challenging, often relying on food distribution data as a proxy.[8] Decision makers grapple with planning equitable and efficient food distribution amidst these complexities, compounded by underlying factors such as poverty and unemployment.[10–12]

This study builds on existing research, focusing on forecasting food demand, an area with limited attention in the literature.[8] Utilizing a visual analytics approach, this study helps decision makers with understanding potential underlying factors influencing forecasts, providing a flexible and convenient tool. Additionally, manual configuration of forecast models using Python programs allows us to compare accuracy with visual analytics-generated forecasts.

In humanitarian logistics, research has predominantly focused on sudden-onset disaster relief, particularly in disaster operations management.[13;14;15] Resource allocation problems have been addressed using Mixed Integer Programming (MIP) and stochastic programming, emphasizing uncertainties in demand and supply.[16] Fewer studies explore uncertainties in slow-onset environments, such as in-kind food donations.[8] Davis et al. applied several forecasting algorithms to model the uncertainty inherent in material donations (supply), but the demand by aid recipients remains to be studied. This becomes even more imperative considering the several socioeconomic factors that might have a relationship with the food demand forecasts.[17] Odubela et al. characterized the behavior of food bank partner agencies and adopted a bottom-up approach to predict the number of persons served as a way to estimate the food demand. However, this study did not consider the underlying factors that might have an impact on food demand. Notably, our study bridges a gap by applying forecasting algorithms to model food demand and identify socioeconomic factors impacting forecasts, contributing to a more comprehensive understanding of food bank operations.

Time series forecasting, a common approach in humanitarian research, leverages past behavior to predict future events.[18] Studies have employed various techniques such as Empirical Mode Decomposition (EMD), Autoregressive Integrated Moving Average (ARIMA), and genetic algorithms to forecast demand post-disaster.[19;20] Additionally, a fuzzy logic-based algorithm has been applied for earthquake prediction.[21]

Visual analytics (VA) emerges as a powerful tool, combining human intuition with computational capabilities for data exploration.[22;23] Our study pioneers the use of VA to forecast food demand, enabling decision makers to interact with analytical models through visualization. This approach enhances the speed of data exploration and boosts decision maker confidence.[23]

Importantly, it allows for the exploration of underlying factors impacting forecasts, a unique contribution to the literature.

The study holds merit for nonprofit hunger relief organizations, offering an interactive approach to identify factors influencing data and aiding decision-making. Our study fills a critical gap by integrating forecasting algorithms, visual analytics, and human expertise to model food demand for a hunger relief organization. This approach empowers decision makers to make informed and strategic distribution decisions in the face of uncertainties.

Methodology

Problem Background

The Food Bank of Central and Eastern North Carolina (FBCENC) is one of the seven food banks in North Carolina, all belonging to the Feeding America network – serving about a third of all the 100 counties in NC. The food bank faces the challenge of equitable distribution of food among the 34 counties, such that the amount of food distributed to each county – and agencies in the county – is proportional to the amount of need there.

The counties are reached through six food bank branch locations (warehouses) based on the service area under which they are categorized. FBCENC and its branch warehouses can receive donations from wholesale grocers, supermarkets, manufacturers, farmers and other individuals or groups. These donations are distributed through more than 700 partner agencies offering different programs such as kids' café, soup kitchens, children and elderly nutrition, food pantries and so on.

The food bank decision makers are required to determine the amount of food they expect to distribute to the counties based on available donations, since the resources available cannot completely satisfy the need. Hence, there is the need for improved forecasts. Presently, the food bank uses a fair share policy to equitably allocate food among the counties served. To assess the effectiveness of the fair share policy, Feeding America – the umbrella body of food banks in the U.S. – recommends that a county's pounds per person in poverty (PPIP) be a minimum of 75%.

To evaluate and augment the fair share policy, the food bank utilizes the Map the Meal Gap tool, a publication by Feeding America. This tool provides valuable insights through estimates of county-level and child food insecurity, as well as estimates of the food budget shortfall among food-insecure individuals. Map the Meal Gap leverages data on critical factors contributing to food insecurity, including unemployment, poverty, and various demographic and household characteristics.

While the fair share policy and Map the Meal Gap offer valuable perspectives, the challenge lies in forecasting food demand accurately.

Presently, the amount of food distributed to a county or branch, based on the fair share policy, serves as a proxy for demand. However, there exists an opportunity to improve forecasting accuracy by introducing underlying factors. This research endeavors to fill this gap by exploring the impact of these factors on the forecasts generated for food demand by a hunger relief organization.

Moreover, this study introduces a novel approach to aid decision-makers at hunger relief organizations – a visual analytics methodology. This approach allows decision-makers to interact with analytic models through visualization, eliminating the need for advanced programming skills or an in-depth understanding of complex statistical techniques. By incorporating this visual analytics approach, decision-makers gain a more accessible and flexible tool to navigate the intricacies of forecasting in the dynamic landscape of hunger relief.

Research Questions

To understand the nature of food demand at the food bank, this study seeks to answer the following research questions: (1) Considering the underlying structure of the demand data such as seasonality and trends, which forecasting model generates the best estimate for food demand? (2) What underlying socioeconomic factors have a relationship with the demand data? (3) What changes are made to the forecast by manipulating the factors?

Data Collection

To address the questions, six fiscal years of food distribution data (2006–2012) from FBCENC was gathered, focusing on each county's donation amounts. The data was pre-processed to include relevant fields: gross weight (pounds) and posting date. Food distribution data was used as a proxy for capturing food demand data as used in literature.[8] Historical records of socioeconomic factors were obtained to explore potential relationships, including county-level unemployment statistics from the U.S. Bureau of Labor Statistics (BLS) and poverty rates and median household income from the U.S. Census Bureau's Small Area Income Poverty Estimates (SAIPE) page for 2006–2012.

The data was aggregated by month and year, aiming to determine the most accurate forecasting model. The unemployment rates were sourced from the BLS (https://www.bls.gov/data/#unemployment), while poverty rates and median household income data were obtained from the USCB website (https://www.census.gov/programs-surveys/saipe/data.html). A summary of each measure is provided in Tables 1 and 2 displays sample data for one fiscal year.[24]

Table 1. Summary of data processing.

Measure (State and County Level)	Source	Measurement Unit	Aggregation	Period Considered
Food Demand - Gross Weight	FBCENC	Pounds	Monthly	July 2006 - June 2011
Unemployment Rate	BLS	Percentage	Monthly	July 2006 - June 2011
Poverty Rates	USCB			
Under Age 18		Percentage	Yearly	2006 – 2011
Ages 5–17		Percentage	Yearly	2006 – 2011
Under Age 5		Percentage	Yearly	2006 – 2011
Median Household Income	UCSB	Dollars	Yearly	2006 – 2011

Table 2. Sample data.

Date	Food Demand (lbs.)	Poverty	Unemployment	Poverty-Under18	Poverty-Age 5–17	Poverty-Under5	Median Household Income
Jul-06	2853437	14.7	4.8	20.1	18.3	23.3	39797
Aug-06	2696531	14.7	4.9	20.1	18.3	23.3	39797
Sep-06	2325048	14.7	4.9	20.1	18.3	23.3	39797
Oct-06	2313834	14.7	4.9	20.1	18.3	23.3	39797
Nov-06	2717386	14.7	4.8	20.1	18.3	23.3	39797
Dec-06	2869578	14.7	4.7	19.5	18.3	23.3	39797
Jan-07	2916354	14.3	4.7	19.5	17.8	22.6	43513
Feb-07	2641102	14.3	4.6	19.5	17.8	22.6	43513
Mar-07	2887641	14.3	4.6	19.5	17.8	22.6	43513
Apr-07	2859369	14.3	4.6	19.5	17.8	22.6	43513
May-07	2760968	14.3	4.6	19.5	17.8	22.6	43513
Jun-07	2805675	14.3	4.6	19.5	17.8	22.6	43513

Forecasting Models

ARIMA Model

Yule introduced autoregressive (AR) and moving average (MA) models, expanded by.[25] ARIMA, by Gooijer and Hyndman,[26] connects successive observations in time series, using univariate models without independent variables and multivariate models with accurate forecasts of independent variables.[27]

In summary, for an autoregressive process, it can be expressed as:

$$\mathrm{Y_t} = \mathrm{a} + \theta_1 \mathrm{Y_{t-1}} + \theta_2 \mathrm{Y_{t-2}} + \ldots + \theta_p \mathrm{Y_{t-p}} + \mathrm{E_t} \tag{1}$$

$$\mathrm{Y_t} = \mathrm{b} + \theta_1 \mathrm{E_{t-1}} + \theta_2 \mathrm{E_{t-2}} + \ldots + \theta_p \mathrm{E_{t-p}} + \mathrm{E_t} \tag{2}$$

Equation (1) expresses the autoregressive process (AR), where p is the number of past observations and E_t is the random shock at time t. Equation (2) represents the moving average process (MA), with q as the number of past observations. Coefficients (ϴ) and constants (a and b) are estimated through repeated computation models.[27]

Estimates of the coefficients ϴ and constants (a and b) can be obtained by repeated computation models.[27] The ARIMA method assumes stationarity in time series, but real-world data often lacks this feature.[27] Pereira recommends removing trends, variance, and seasonality before applying ARIMA, with the option to reintroduce them later.[28] Enders suggests unit root tests, and determining autoregressive (p) and moving average (q) observations relies on examining ACF and PACF. Model adequacy is validated by assessing residuals, ensuring they resemble a normal distribution with zero mean and zero autocorrelation.

Exponential Smoothing Model

Exponential smoothing methods were first introduced in the works of[29–31] and.[32,33] first suggested a statistical basis for applying simple exponential smoothing after demonstrating optimal forecasts generated by the model for a random data with noise.[26] Gooijer & Hyndman have identified the best known exponential smoothing models including simple exponential smoothing (SES) for no trend and seasonality in the series, Holt's linear method for additive trend and no seasonality, and Holt-Winters' additive method for additive trend and seasonality. Trend-based methods excel with noticeable trends, identified by plotting the time series over time.[34]

$$Y_t = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + \ldots.. \tag{3}$$

In a simple exponential smoothing, an estimate of each one-step-ahead forecast is generated by repeatedly computing the Equation (3) below.[27] This method places more weight on the most recent observations by calculating a geometric sum of previous observations. The smoothing parameter (α) ranges from 0 to 1.

The Holt-Winters Method

The Holt-Winters method is a generalization of the exponential smoothing methods. It is very useful for forecasting a time series that has elements of trend and seasonality. In the additive Holt-Winters Method, the time series (Y_t) is composed of the trend (T_t), seasonality (S_t) and irregularity (I_t) components. It can be represented as in equation (4) below.

$$Y_t = T_t + S_t + I_t \tag{4}$$

The trend *T* is assumed to be linear while *S*, the seasonality component changes slowly over time. Irregularity *I* is due to a random shock or white noise. For the purpose of forecasting, the trend and seasonal components and trend gradient per time (G_t) can be estimated from the observations. Each of the trend, seasonality and gradient have associated smoothing parameters. The values of these parameters should be selected to minimize the mean square errors (MSE) of the one-step ahead forecast.[35]

SAS Visual Analytics

Several tools are available for bridging the gap between visualization and analytics domains. They allow business users create visual representations of their data and apply human judgments to make conclusions based on evidence, hypothesis or assumptions.[22] Tools like Tableau, SAS Visual Analytics (SAS VA), Board, and Microsoft Power BI facilitate seamless integration of visualization and analytics. As such, business users can decide what is important to focus on.[36]

SAS Visual Analytics (SAS VA) was chosen because of its ability to deal with huge data sets, automatically selecting the best chart (auto-charting) for visualizing the data and its advanced forecasting features. The forecasting feature in SAS VA can be applied by users to predict how their data trends into the future. It also automatically identifies and selects the best forecasting model that fits the data. The forecasting models currently available in SAS VA 7.4 include ARIMA, damped trend exponential smoothing, seasonal exponential smoothing, simple exponential smoothing, linear exponential smoothing and Winters method. SAS VA allows the user to manipulate the forecasted data to gain insights into how certain data items factor into the forecast. This involves finding underlying factors that have a relationship to the forecast which can be manipulated with scenario analysis.

Underlying Factors with Forecasting

Time series forecasting, predicting values based on historical records, can be enhanced by incorporating additional measures that influence the variable of interest.[37] SAS VA forecasting models integrate these measures during analysis, identifying their impact on the forecast. Non-influential measures can be removed, strengthening the model. An advanced feature in SAS VA updates forecast values automatically when new influencing measures are added.[37] Identified influential measures enable scenario analysis and goal seeking. SAS VA retains the best-fit forecasting model and allows users to manipulate underlying factors, visualizing forecast changes based on new values. Goal seeking permits users to set forecast values, exploring necessary changes in underlying factors. Accuracy in underlying factor values depends on their strong correlation with the forecasted variable.[37]

Forecast Model Evaluation

Data is split into two sets: the training set (July 2006 – June 2011) for building time series and forecast models, and the test set (July 2011 – June 2012) for assessing forecast accuracy using mean absolute percentage error (MAPE) with one-period-ahead forecasts ($T_0 = 12$). The choice of the split is based on careful consideration of the trade-off between model training and evaluation. Allocating about 80% of the data for training ensures that the models are exposed to a sufficiently large and diverse set of examples. This, in turn, allows the models to capture underlying patterns, and relationships, within the data. Reserving about 20% of the dataset for testing serves a critical purpose in the evaluation phase. This independent subset, untouched during the training process, acts as a robust benchmark to assess the generalization performance of the developed models. It provides a real-world simulation, allowing us to gauge how well the models are likely to perform on new, unseen data.

The careful split between training and testing sets helps mitigate the risk of overfitting, ensuring that the models do not merely memorize the training data but learn to generalize to new instances effectively. The evaluation on the testing set serves as a stringent test of the model's ability to extrapolate patterns learned during training to novel data, indicative of its real-world applicability.

One-period-ahead forecasts $\hat{Y}_{t+1|t}$ are generated for each of Y_1 to Y_t. The MAPE is determined according to the equation below:

$$\mathrm{MAPE} = T_0^{-1}\left[\sum_{t=T}^{T+T_0} \left|\frac{\hat{Y}_{t+1|t} - Y_{t+1}}{Y_{t+1}}\right] x100\right| \tag{5}$$

The mean absolute percentage error (MAPE) is employed to evaluate forecast improvement with each introduced underlying factor, specifically those significantly contributing to the forecast. A lower MAPE, when comparing forecasts for different factors, signifies improved accuracy by incorporating related information contributing to the forecasted food demand.

Results and Discussion

Underlying Factors of Food Demand and Time Series Plot

The correlation analysis reveals strong correlations (Pearson's $r > 0.8$) between food demand ("Demand") and underlying factors such as poverty and unemployment. However, median household income exhibits a moderate relationship with demand, approximately 0.4. The time series plot below (Figure 1) illustrates a consistent upward trend in food demand annually, with noticeable seasonality, particularly during holidays, reflecting increased food drives and donations to the food bank for distribution to those in need.

Forecast Model Results at the Food Bank Level

Without Underlying Factors

Utilizing Visual Analytics streamlines decision-making for non-experts by offering quick and user-friendly tools, as emphasized earlier. Leveraging SAS

Figure 1. Time series plot of monthly food demand in millions of pounds.

VA, the Additive Winters Method was identified as the optimal fit for the data before introducing underlying factors. This aligns with[35] endorsement of the Holt-Winters method for its robust forecasting performance. Particularly effective for time series with diverse components and subject to structural changes, the Holts-Winters method provides stable and accurate forecasts.

Figure 2 offers a detailed view of the monthly food demand forecast, spanning six intervals from the last month of the provided data. The forecast is accompanied by 95% confidence intervals, enhancing the precision of the predictions.

With Underlying Factors

Upon introducing underlying factors to the time series, the forecasting algorithm transitioned to the ARIMA model for optimal accuracy, following[8] recommendation for series exceeding 50 values. With our 60-data-point dataset spanning five fiscal years, SAS VA identified significant contributors using the ARIMAX model, an extension of ARIMA incorporating underlying factors.[38] It utilizes autocorrelation in residuals, akin to a multivariate regression model, selecting factors based on their statistical significance at a 5% level.

At the food bank level (34 counties combined), under-18 Poverty emerged as the sole contributor to the forecast, even with other factors present. However, excluding under-18 Poverty shifted the spotlight to median

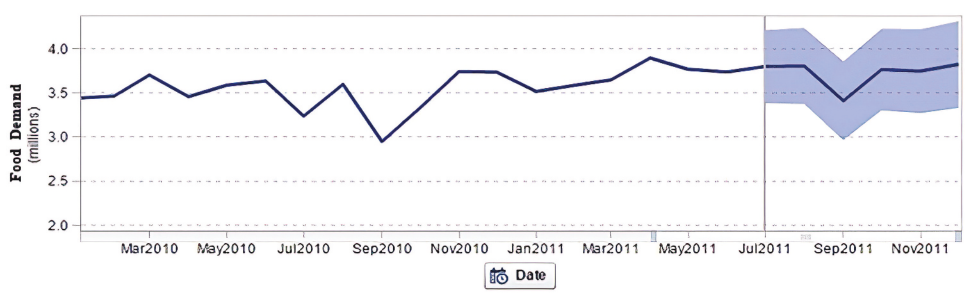

Figure 2. Closer view of food demand forecast – before introducing underlying factors.

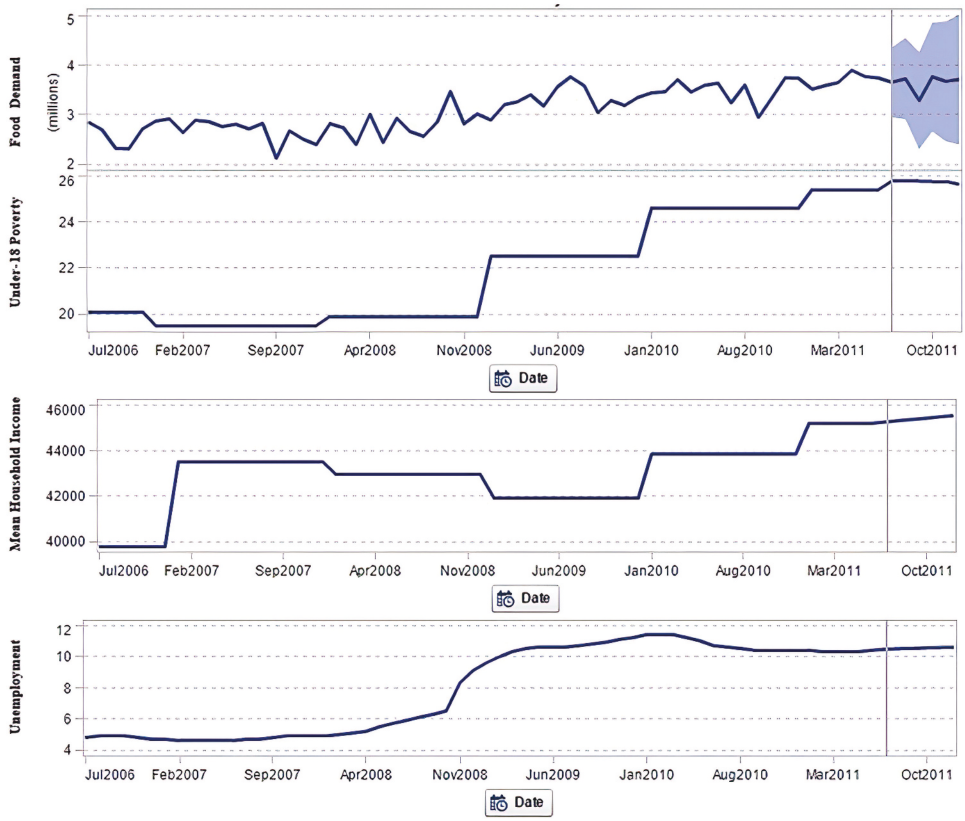

Figure 3. Forecast of food demand with under-18 poverty introduced, median household income introduced, and unemployment introduced.

household income as the sole contributor, indicating no joint contribution of factors in this context. SAS VA permits multiple contributors, but, in this application, it showcased a single significant factor at a time. Removing median household income revealed unemployment as the sole contributor to the forecast, as depicted in Figure 3 below.

Forecast Model Results at the County Level

The analysis above was repeated with data from Wake County, the largest county among the 34 counties served by the food bank, aimed to utilize readily available socioeconomic data at the county level. Wake County, the second most populous in the state, received 9–18% of FBCENC's distributed food and averaged 13% monthly.[39]

Similar to the food bank-level analysis, introducing underlying factors to the time series led to an update of the forecasting algorithm to the ARIMA model for precision. Unemployment emerged as a significant contributor.

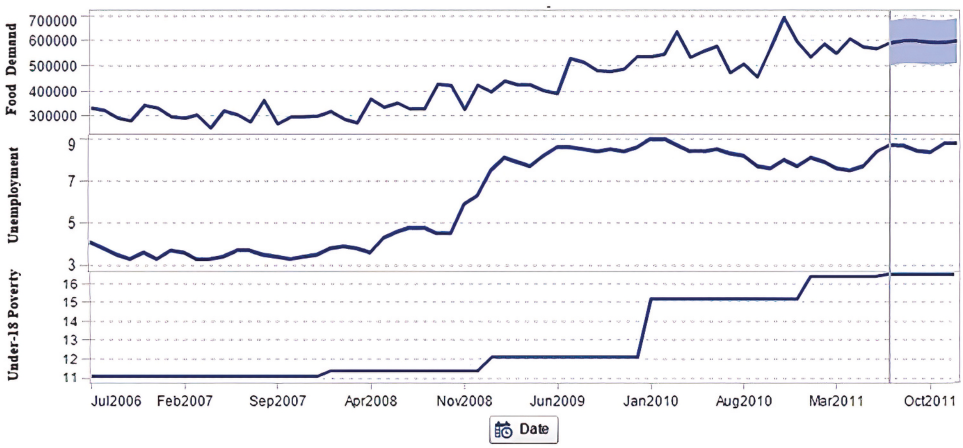

Figure 4. Forecast of food demand by Wake County (in pounds) – unemployment and under-18 poverty introduced.

Subsequently, under-18 poverty, not median household income, made a significant contribution. Notably, unlike the food bank-level forecast, unemployment and under-18 poverty jointly contributed to the county forecast, showcasing a unique pattern.

It's noteworthy that Wake County's median household income exceeds state and national averages, holding the state's highest in 2017.[40] Figure 4 presents the obtained forecast.

Forecast Model Accuracy

While the visual analytics tool automatically selects the best forecasting model, it's crucial to ensure its accuracy, especially for users unfamiliar with forecasting or programming. The mean absolute percentage error (MAPE) was used to assess accuracy, comparing tool-generated forecasts to test data. Tables below display MAPE for food bank and county level forecasts. A lower MAPE indicates the forecast is more accurate relative to the test data. Results indicate underlying factors contributed to improvements of accuracy for a 6-month forecast, particularly median household income for a 12-month forecast. Food banks, facing dynamic conditions, should cautiously plan for the next six months, as forecasts projecting too far ahead may be less accurate, as seen in Table 3.

Table 3. MAPE of the food bank level forecast.

	Pure forecast*	Median household income	Under-18 poverty	Unemployment
6 months forecast	8.16	6.33	6.72	6.78
12 months (one period) forecast	8.79	8.71	8.97	8.97

*Pure forecast: without underlying factors.

Table 4. MAPE of the county level forecast – Wake County.

	Pure forecast	Under 18 poverty & Unemployment	Unemployment
6 months forecast	15.39	20.33	20.59
12 months (one period) forecast	12.94	16.30	15.27

The county study yielded similar results. Wake County exhibited substantial variability in food demand, resulting in higher prediction errors at the county level compared to the food bank level. Surprisingly, the forecast had a lower MAPE before introducing underlying factors, despite the visual analytics tool identifying significant relationships between food demand, under-18 poverty, and unemployment in Wake County. This supports the notion that centralized forecasts are more accurate than decentralized ones. Table 4 shows the MAPE for the county level forecast.

Forecast Model Comparison

Analyzing time series data requires consideration of correlated observations and errors from consecutive periods. Ignoring these correlations can result in significant errors when analyzing time-dependent data. SAS VA prevents such errors for users without statistical or programming knowledge by providing a straightforward way to use advanced forecasting features. It employs various algorithms, parameters, and underlying factors, selecting the combination that yields the best forecast.

To assess its performance, SAS VA with time series models built from scratch using Python were compared. Further research is needed to evaluate the effectiveness of models generated by VA tools. The steps involved for each model development are described below. Further research is required to assess the effectiveness of models generated through VA tools.

Holt Winters' Exponential Smoothing

This model, suitable for univariate time series with trend and/or seasonal components, operates akin to the ARIMA family by generating forecasts through a weighted sum of previous observations. Given our data's trend and seasonal components, this method is a valuable choice. Utilizing a seasonal period of 12 months, the time series was decomposed. The forecast used the same data fed to the VA tool, with five years of historical data (fiscal years 2006–2011) as the training dataset and one year (2011–2012) as the test dataset for prediction. MAPE measured accuracy, resulting in an 8.72% MAPE for a 12-month forecast and a more accurate 6.48% for a 6-month forecast, aligning with expectations as longer forecast horizons tend to degrade model performance. Figure 5 provides a decomposition of the time series, unveiling distinct features. Notably, it highlights a discernible upward trend over time, accompanied by evident seasonality,

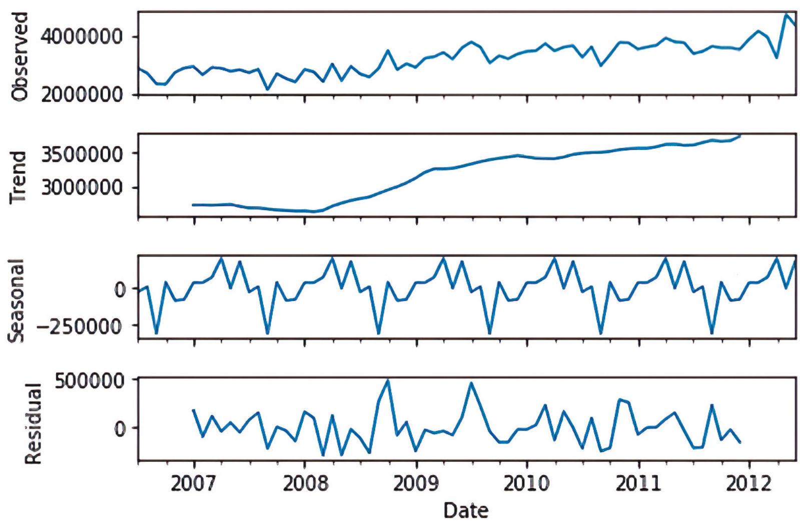

Figure 5. Decomposition of the time series showing trend, seasonality and noise.

showcasing consistent patterns year after year. The residual section represents the differential component derived after extracting both the seasonal and trend elements from the observed data.

ARIMA Model

The five years of historical data were split into training and validation sets (60:40 ratio), reserving the sixth fiscal year as the unknown test set. To ensure stationarity, the Augmented Dickey-Fuller (ADF) test confirmed the 1-lag differenced time series lacked a time-dependent structure. Conducting a grid search for ARIMA hyperparameters on our small dataset of 60 observations, the goal was to identify the optimal combination of p, d, and q using root mean square error (RMSE) as the performance metric. Employing walk-forward validation, ARIMA (1,1,0) emerged as the optimal configuration. Trained on the entire dataset using a rolling forecast, the one-period forecast had a 7.78% MAPE, while the 6-month forecast improved to 3.85%.

Table 5 compares the MAPE for food distribution forecasts generated by the VA tool and those manually configured. Notably, the manually configured models consistently outperformed the VA tool's choice, likely due to the researcher's ability to iterate and fine-tune model parameters for optimal performance.

Table 5. Model comparison using MAPE (%).

	VA Tool	Holt-Winters*	ARIMA*
6 months forecast	8.16	6.48	3.85
12 months (one period) forecast	8.79	8.72	7.78

*Manually configured model.

Conclusion

In conclusion, our study employed SAS VA, a Visual Analytics (VA) tool, to assess the impact of underlying factors on food distribution forecasts, serving as an estimation for demand. While the tool identified the best-fitting time series model, particularly showcasing the significance of under-18 poverty, median household income, and unemployment rates at the food bank level, the observed improvements were more prominent in short-term forecasts (6 months) than in longer-term predictions (12 months). Interestingly, for the selected county, manual configuration outperformed VA tool selections, underscoring the importance of iterative parameter optimization. Our findings, aligned with previous studies on socioeconomic factors, highlight variations in contributors to demand forecasts across different organizational levels. Despite these nuances, our research supports the efficacy of VA tools like SAS VA, offering a user-friendly approach for time series model construction. The showcased scenario analysis and goal-seeking options empower decision-makers to integrate empirical knowledge with analytical insights, enhancing forecast accuracy and potentially optimizing food bank operations.

Limitations and Future Research

In acknowledging the complexity of forecasting food demand, this study selected various socioeconomic factors, recognizing that additional variables could further enhance forecast accuracy. Future research endeavors may explore demographics, homeownership rates, labor force participation in specific industries, geographical factors, inflation, consumer price index, natural disasters, and more. This comprehensive expansion aims to provide a more nuanced understanding and improved predictive capabilities, incorporating insights derived from recent data on food bank operations.

Moreover, the current research primarily focuses on food demand forecasts at the food bank level, leaving room for future investigations to extend their scope to the branch/network level. Planned explorations encompass diverse food types and factors influencing donations, contributing to strategic planning for food purchases and donations at a more localized level. It is worth noting that our current measure of food distributed may potentially underestimate true demand, highlighting an avenue for future research to explore innovative methodologies for a more accurate estimation of demand within a food bank's service area. Additionally, considering the dynamic nature of forecasting models, there is acknowledgment that other, potentially superior models might exist and could be explored in subsequent studies.

Acknowledgements

We would like to express our sincere appreciation to the leadership of the Food Bank of Central and Eastern North Carolina (FBCENC) for their valuable input to this project. This research was partially supported by NSF grant (#1718672) titled PFI: BIC - Flexible, equitable, efficient, and effective distribution (FEEED), NSF grant (#1735258) titled NRT: Improving strategies for hunger relief and food security using computational data science, and NSF grant (#2100855) titled:" EiR: Human Centered Visual Analytics for Evidence Based Decision Making in Humanitarian Relief.".

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the NSF [1718672], NSF (1735258) and NSF (2100855).

ORCID

Kehinde Odubela http://orcid.org/0000-0003-2972-8970
Steven Jiang http://orcid.org/0000-0002-6937-927X

References

1. Campbell CC. Food insecurity: a nutritional outcome or a predictor variable? *J Nutr*. 1991;121(3):408–415. doi:10.1093/jn/121.3.408.
2. Coleman-Jensen A, Gregory C, Singh A. Household food security in the United States in 2013. *USDA-ERS Economic Research Report*. 2014;173. doi:10.2139/ssrn.2504067.
3. Rabbitt MP, Hales LJ, Burke MP, Coleman-Jensen A. Household food security in the United States in 2022 (report No. ERR-325). U.S. Department of Agriculture, Economic Research Service. 2023. doi:10.32747/2023.8134351.ers.
4. Holmes E, Fowokan A, Seto D, Lear SA, Black JL. Examining food insecurity among food bank members in Greater Vancouver. *J Hunger Environ Nutr*. 2019;14(1–2):141–154. doi:10.1080/19320248.2018.1465001.
5. Popielarski JA, Cotugna N. Fighting hunger through innovation: evaluation of a food bank's social enterprise venture. *J Hunger Environ Nutr*. 2010;5(1):56–69. doi:10.1080/19320240903575160.
6. Delpish R, Jiang S, Davis L, Odubela K (). A visual analytics approach to combat confirmation bias for a local food bank. In *Advances in Human Error, Reliability, Resilience, and Performance: Proceedings of the AHFE 2018 International Conference on Human Error, Reliability, Resilience, and Performance July 21-25, 2018, Loews Sapphire Falls Resort at Universal Studios*, Orlando, Florida, USA: Springer International Publishing; 2019:13–23.
7. Ivuawuogu H, Jiang S, Davis L. Clustering Partner Agencies for a Local Food Bank. In *Proceedings of 2023 Industrial and Systems Engineering Annual Conference*, New Orleans, LA, USA; 2023.

8. Davis LB, Jiang SX, Morgan SD, Nuamah IA, Terry JR. Analysis and prediction of food donation behavior for a domestic hunger relief organization. *Int J Prod Econ*. 2016;182:26–37. doi:10.1016/j.ijpe.2016.07.020.
9. Batey L, DeWitt E, Brewer D, Cardarelli KM, Norman-Burgdolf H. Exploring food-based cultural practices to address food insecurity in rural Appalachia. *Health Educ Behav*. 2023;50(4):529–537. doi:10.1177/10901981231175360.
10. Rose D. Economic determinants and dietary consequences of food insecurity in the United States. *J Nutr*. 1999;129(2):517S–520S. doi:10.1093/jn/129.2.517S.
11. Lê Q, Auckland S, Nguyen HB, Murray S, Long G, Terry DR. *The Socio-Economic and Physical Contributors to Food Insecurity in a Rural Community*. SAGE Open; 2015. doi:10.1177/2158244014567401.
12. Olabiyi OM, McIntyre L. Determinants of food insecurity in higher-income households in Canada. *J Hunger Environ Nutr*. 2014;9(4):433–448. doi:10.1080/19320248.2014.908450.
13. Hoyos MC, Morales RS, Akhavan-Tabatabaei R. Or models with stochastic components in disaster operations management: a literature survey. *Computers & Industrial Engineering*. 2015;82:183–197. doi:10.1016/j.cie.2014.11.025.
14. Hu SL, Han CF, Meng LP. Stochastic optimization for joint decision making of inventory and procurement in humanitarian relief. *Computers & Industrial Engineering*. 2017;111:39–49. doi:10.1016/j.cie.2017.06.029.
15. Kawashima S, Morita A, Higuchi T. Emergency stockpiling of food and drinking water in preparation for earthquakes: evidence from a survey conducted in Sendai city, Japan. *J Hunger Environ Nutr*. 2012;7(2–3):113–121. doi:10.1080/19320248.2012.704661.
16. Escudero LF, Garín MA, Monge JF, Unzueta A. On preparedness resource allocation planning for natural disaster relief under endogenous uncertainty with time-consistent risk-averse management. *Computers & Operations Research*. 2018;98:84–102. doi:10.1016/j.cor.2018.05.010.
17. Odubela K, Jiang S, Davis L. Estimating true demand at Hunger Relief Organizations with predictive modeling. *J Hunger Environ Nutr*. 2022;1–20. doi:10.1080/19320248.2022.2061885.
18. Shumway RH, Stoffer DS, Olkin I, Casella G, Fienberg S. *Time Series Analysis and Its Applications*. New York: Springer- Verlag; 2000.
19. Xu X, Qi Y, Hua Z. Forecasting demand of commodities after natural disasters. *Expert Syst Appl*. 2010;37(6):4313–4317. doi:10.1016/j.eswa.2009.11.069.
20. Wang Z-X. A genetic algorithm-based grey method for forecasting food demand after snow disasters: an empirical study. *Natural Hazards*. 2013;68(2):675–686. doi:10.1007/s11069-013-0644-8.
21. Aydin I, Karakose M, Akin E. The prediction algorithm based on fuzzy logic using time series data mining method. *World Acad Sci Eng Technol*. 2009;51:91–98.
22. Thomas JJ, Cook KA, Ed. *Illuminating the Path: The R&D Agenda for Visual Analytics*. Richland, Washington: National Visualization and Analytics Center; 2005.
23. Keim DA, Ward MO. Visual data mining techniques. In: Berthold M Hands D, eds. *Intelligent Data Analysis: An Introduction*. Berlin: Springer; 2002: 2–27.
24. Yule GU. On the method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions Of The Royal Society London, Series A*. 1927;226:267–298.
25. Box GEP, Jenkins GM. *Time series analysis: forecasting and control*. Revised ed. San Francisco: Holden Day; 1970.
26. De Gooijer JG, Hyndman RJ. 25 years of time series forecasting. *International Journal Of Forecasting*. 2006;22(3):443e473. doi:10.1016/j.ijforecast.2006.01.001.

27. Pereira A. Performance of time-series methods in forecasting the demand for red blood cell transfusion. *Transfusion*. 2004;44(5):739–46. doi:10.1111/j.1537-2995.2004.03363.x.
28. Enders W. *Applied Econometric Time Series*. New York: Wiley; 2004.
29. Brown RG (1959). Statistical forecasting for inventory control. New York7 McGraw-Hill.
30. Brown RG. *Smoothing, forecasting and prediction of discrete time series*. Englewood Cliffs, NJ: Prentice-Hall; 1963.
31. Holt CC. Forecasting seasonality and trends by exponentially weighted averages. O.N.R. Memorandum 52/1957, Carnegie Institute of Technology. *International Journal Of Forecasting*. 1957;20:5–13. Reprinted with discussion in 2004.
32. Winters PR. Forecasting sales by exponentially weighted moving averages. *Management Science*. 1960;6(3):324–342. doi:10.1287/mnsc.6.3.324.
33. Muth JF. Optimal properties of exponentially weighted forecasts. *J Am Statistic Assoc*. 1960;55(290):299–306. doi:10.1080/01621459.1960.10482064.
34. Nahmias S. *Production and Operations Analysis*. New York: McGraw-Hill; 2009.
35. Chen C. Robustness properties of some forecasting methods for seasonal time series: a monte carlo study. *Int J Forecasting*. 1997;13(2):269–280. doi:10.1016/S0169-2070(97)00014-9.
36. Scanaxrio Analysis SAS. Wikipedia. Retrieved November 22, 2018. https://en.wikipedia.org/wiki/Scenario_analysis#cite_note-1.
37. Aanderud T, Collum R, Kumpfmiller R. SAS® visual analytics tricks we learned from reading hundreds of SAS® community posts. In *SAS Global Forum Proceedings, Paper 1010-2017*, Orlando, FL; 2017.
38. Peter Ď, Silvia P. ARIMA vs. ARIMAX – which approach is better to analyze and forecast macroeconomic time series? In *Proceedings of the 30th International Conference Mathematical Methods in Economics*, Karviná, Czech Republic; 2012:136–140.
39. U.S. Bureau of Labor Statistics. Local Area Unemployment Statistics, (https://data.bls.gov/map/MapToolServlet). accessed 2018-10-24.
40. U.S. Census Bureau. Small area income and poverty estimates (SAIPE), (https://www.census.gov/data-tools/demo/saipe/saipe.html). accessed 2018-10-24.