Sales Trend Extraction of Fast-Moving Consumer Goods

marketing, scanner data, time series, trend filtering, sparse modeling

Extended Abstract

Quantifying sales trend of products is essential for retailers and manufacturers to refine their marketing strategies. Through analysis of historical and current trends, retailers can make informed decisions regarding pricing, promotion, inventory management and shelf space allocation. Similarly, manufacturers can leverage this information for product development, advertising campaigns, competitive analysis, and distribution management.

In the fast-moving consumer goods (FMCG) industry, the analysis of sales trends has become increasingly important due to intense competition among brands. However, traditional methods such as year-over-year comparisons and moving averages are known to be imprecise, resulting in a delayed understanding of trends. This is particularly true due to the unique characteristics of time series data in FMCG sales, such as spikes, seasonality and abrupt changes in trend.

The social sciences have approached the analysis of FMCG sales data from explanatory perspectives. Research on the long-term effects of marketing mix has shown that product line length and distribution breadth have a stronger impact on brand sales than advertising and discounting [1]. Studies examining the effectiveness of TV advertising across 288 consumer packaged brands indicate that advertising has little effect on sales in general, leading to negative return on investment for most brands [5]. However, there has been little research on sales trend extraction methods that focus on scanner data within the fields of marketing and economics. Moreover, classical and prevailing trend extraction methods fail to accurately model abrupt changes in FMCG sales trend [2, 3].

To address this gap in research, we propose a method for trend extraction specifically tailored to FMCG sales. Our approach is a variation on ℓ_1 trend filtering, which can accurately model abrupt changes in trends [4]. The ℓ_1 trend filtering is a sparse modeling technique applied to time series data, assuming an underlying piecewise linear trend. Our method minimizes the following objective function:

$$
\begin{aligned}
f(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\omega}, \boldsymbol{s}) = \frac{1}{2}\sum_{t=1}^{n}(\log y_t - x_t - \mu_t - \omega_t - s_t)^2 \\
+ \lambda_1 \sum_{t=2}^{n-1} |x_{t-1} - 2x_t + x_{t+1}| + \lambda_2 \sum_{t=2}^{n} |\mu_t - \mu_{t-1}| + \lambda_3 \sum_{t=1}^{n} |\omega_t|
\end{aligned}
$$

where y_t denotes the sales of target brand at time t, x_t is the trend in time series, μ_t is the level of trend, ω_t is the spike and s_t denotes the seasonality effect, which is approximated by the Fourier series $s_t = \sum_{k=1}^{m}\left(a_k \cos\left(\frac{2\pi kt}{p}\right) + b_k \sin\left(\frac{2\pi kt}{p}\right)\right)$. The first term in the objective function measures the size of the residuals, the second term measures the smoothness of the trend slope by second-order differences, the third term measures the frequency of level shifts using first-order differences. The last term measures the frequency of irregular spikes.

In our experiments using synthetic and observed data, we set the regularization coefficients to $\lambda_1 = 10$, $\lambda_2 = 0.5$ and $\lambda_3 = 0.1$ in order to control the smoothness in the trend extraction. We also set the seasonality parameters to $m = 10$ and $p = 365.25/7 \approx 52.179$ to explain yearly seasonality.

We performed experiments on synthetic data by generating 1000 weekly sales time series, each spanning seven years, that exhibit abrupt changes in both slope and level of trend. In order to compare performance, we selected three baseline algorithms: the Hodrick-Prescott filter (HP) [3], Seasonal-Trend decomposition using LOESS (STL) [2], and the time series forecasting model implemented by Facebook (Prophet) [6]. Figure 1 illustrates a comparison of the baseline and proposed methods using a sampled time series. We calculated the Mean Absolute Error (MAE) by comparing the true trend in the synthetic data to the extracted trends. It is evident that the abrupt change in the level of trend between 2015 and 2016 is the major contributing factor to the differences in MAE. Table 1 presents the average MAE and Root Mean Squared Error (RMSE) of trend estimation methods obtained from 1000 randomly generated synthetic time series. Our trend extraction algorithm demonstrates significantly smaller errors in comparison to the HP, STL, and Prophet.

We apply our proposed sales trend extraction to observed data from the Nielsen RMS (Retail Measurement Services) retail scanner. The data are made available for academic research through a partnership between the Nielsen Company and the University of Chicago Booth School of Business. Our analysis focuses on sales data from a particular supermarket in the New York area spanning from 2013 to 2019. In order to exclude brands that were either launched or discontinued during this period, we select 5553 brands that generated sales every week. Figure 2 presents the results of trend extraction by our proposed method from 7-year weekly sales time series for a randomly sampled subset of brands. Similar to the results obtained from synthetic data, our algorithm successfully captures abrupt changes in the level of trend. Next, we compute the maximum value of $|x_{t-1} - 2x_t + x_{t+1}|$ and $|\mu_{t+1} - \mu_t|$ for all the target brands to investigate the presence of significant abrupt changes. The proportion of brands that exhibit more than 1% increase or decrease in trend slope is 5.0% and the proportion of brands that experience greater than 20% increase or decrease in trend level is 42.4%. Though these values vary with the use of different regularization parameters, it is clear that abrupt changes in trend occur with high frequency, and that modeling such shifts is imperative when analyzing FMCG sales trends.

As this is the ongoing work, we plan to evaluate the performance of baseline and proposed models for observed data by applying time series cross validation that measures forecast error. we also intend to extend our model to detect signs of abrupt changes in future trends from time series of competing brands or sales in other regions.

References

1. Ataman, M. B., Van Heerde, H. J., & Mela, C. F. (2010). The long-term effect of marketing strategy on brand sales. *Journal of Marketing Research, 47*(5), 866-882.
2. Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition. *J. Off. Stat, 6*(1), 3-73.
3. Hodrick, R. J., & Prescott, E. C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, 1-16.
4. Kim, S. J., Koh, K., Boyd, S., & Gorinevsky, D. (2009). ℓ_1 trend filtering. *SIAM review, 51*(2), 339-360.
5. Shapiro, B. T., Hitsch, G. J., & Tuchman, A. E. (2021). TV advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica, 89*(4), 1855-1879.
6. Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician, 72*(1), 37-45.

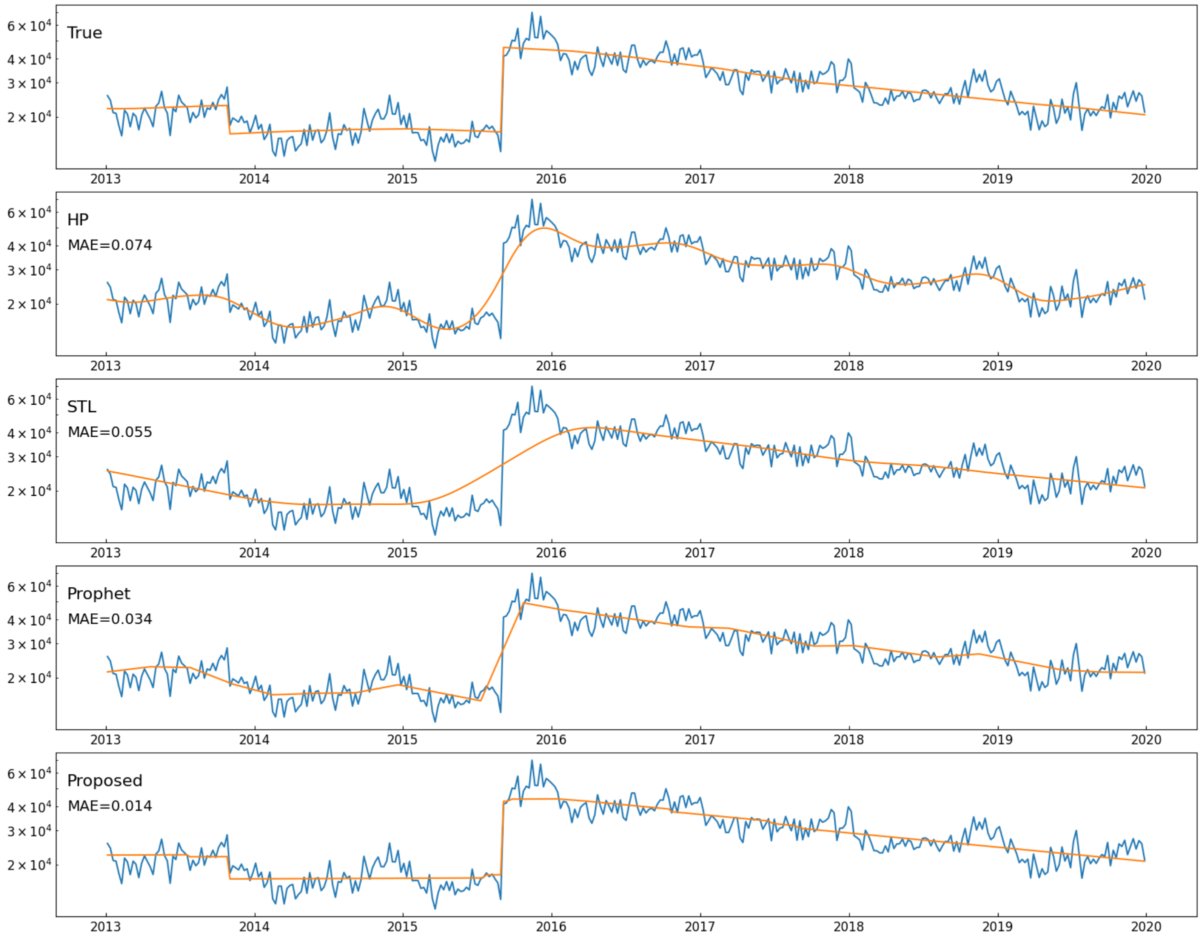

Fig 1. A sample of synthetic time series of weekly sales with its true trend and the comparison of sales trend extraction methods.

Table 1. Average MAE and RMSE for trend extraction in 1000 randomly generated synthetic time series of weekly sales. The numbers in parentheses represent the standard deviation.

	HP	STL	Prophet	Proposed
MAE	0.078 (0.029)	0.057 (0.028)	0.042 (0.018)	**0.024 (0.007)**
RMSE	0.101 (0.032)	0.096 (0.046)	0.068 (0.028)	**0.036 (0.015)**

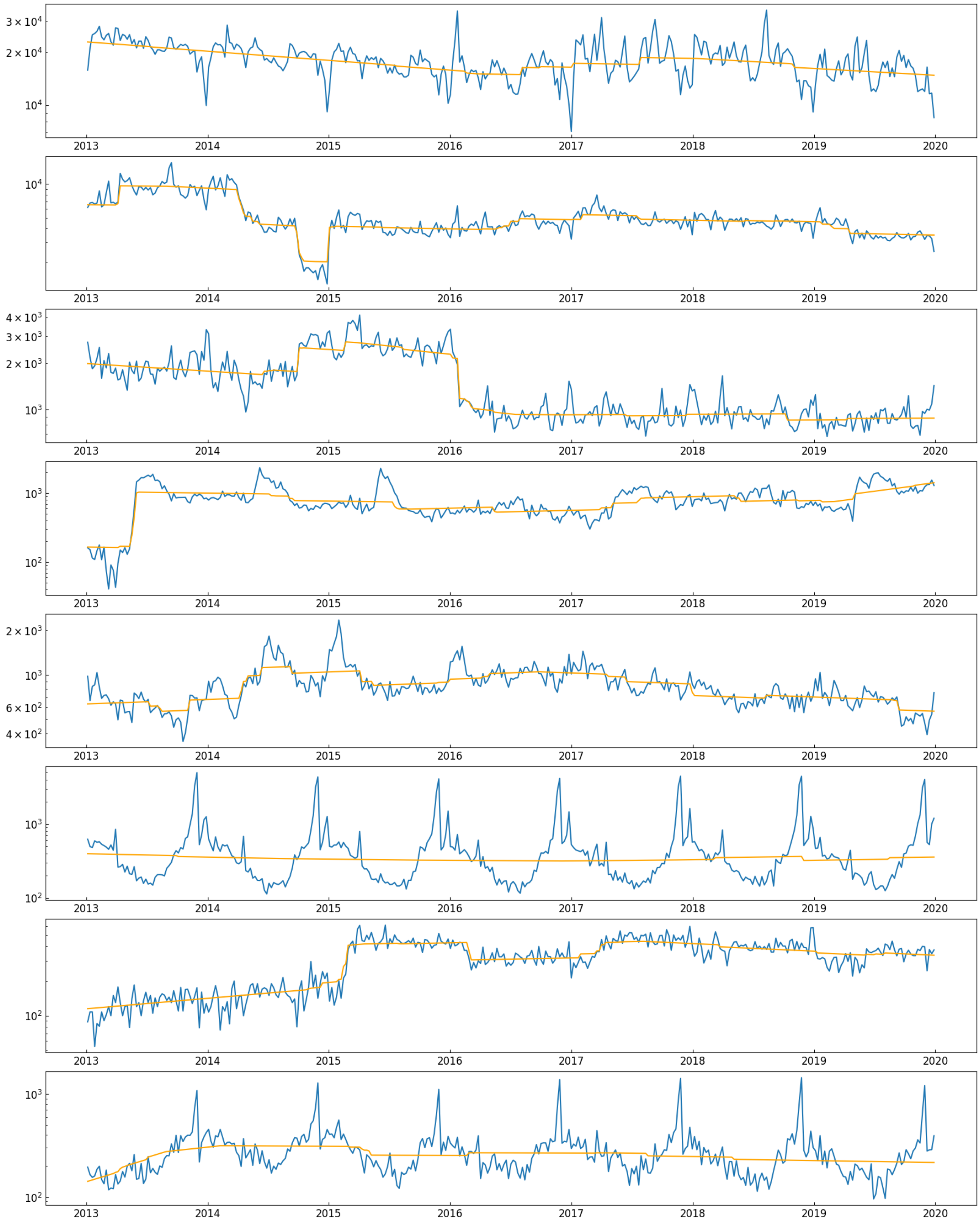

Fig 2. Results of trend extraction by our proposed method for randomly sampled real brands sales from Nielsen dataset. The blue line represents the actual sales, and the orange line represents the extracted trends.