

Food Price Forecasting and Alert for Price Spikes

1 Overview

The price forecast and alert tool is a state-of-the-art approach for humanitarian aid professionals, scholars, students and anyone else interested in the food security implications of staple food price volatility. It is composed of two distinct methodologies for price forecasting and for detecting price spikes based on WFP's Alert for Price Spikes (ALPS) indicator. They are designed to inform WFP's internal operational programming related to cash, vouchers and local procurement while triggering early warning and further food security and market analysis, including monitoring.

2 How the price forecasting works

"Time present and time past Are both perhaps present in time future And time future contained in time past." T.S. Eliot, Four Quartets

Attempting to forecast future prices is by definition a task prone to error. With this caveat, we have created this forecasting tool to offer a well-informed understanding of future price developments in the markets that WFP monitors.

The tool tests different approaches: ARIMA models, and the Seasonal Holt-Winters (SHW) multiplicative model. The best model is the one with the lowest Root Mean Square Error:

$$RMSE = \sqrt{\frac{1}{T} \sum_t^T \epsilon_t^2},$$

which represents the size of the error ϵ for a given time t , measured as the difference between the actual and the estimated price, that is

$$\epsilon_t = p_t - \hat{p}_t.$$

It is measured in the same unit as the original data, applying a greater penalty on large forecasting errors. If both the ARIMA and the SHW approaches fail to produce results, we rely on forecasts using a third approach: a Moving Average Smoother corrected with the Grand Seasonal Index to account for seasonality.

ARIMA (p, d, q): The class of models called ARIMA can be used to forecast stationary time-series. The aim of having stationary time-series is to remove autocorrelation from the forecast errors by adding lags of differenced series and/or of the forecast errors. We adopt the following notation:

- p is the autoregressive term (AR);
- d is the number of non-seasonal differences, namely the differences needed to make the series integrated, i.e. stationary (I);
- q is the moving average term (MA).

An automatic procedure scans our price database, performing the following steps:

1. We ensure the time-series are at least three years long and have less than 30 percent of values missing both for the entire time-series and for the last three years of data. We only show time-series with a minimum of price variability, given by two indicators: the coefficient of variation, as described by, $cv = \sigma/\bar{p}$, and the highest price variation, as described by

$$pv = \max(p_t^H - p_t^L) / p_t,$$

where σ is the standard deviation of prices, $\bar{p}$ is the average price, and p_t^H and p_t^L are respectively the highest and the lowest price observation. For both the coefficient of variation and the price variation, we control for the entire time

series and the last three years of data, allowing the tool to show all time series that have at least one of the four indicators above an arbitrarily defined threshold. Specifically, for *cv* we set a threshold of 15 percent for the entire time-series, and 10 percent for the last three years. For *pv*, we define a threshold of 50 percent for the entire time-series, and 30 percent for the last three years.

2. We linearly interpolate missing values. Note that normally, less than 5 percent of the data is interpolated in this step.
3. We determine the number of autoregressive terms (p) to be included in the model. Akaike’s Information Criterion (AIC) is used to make the selection.
4. We check whether the model with lags p determined above also has a statistically significant trend.
5. We test for stationarity of the model resulting from steps 3 and 4 (i.e. lags and trend) using both the Augmented Dickey Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) unit root tests. If the time-series is not stationary, steps 6 and 7 are performed on a differentiated time-series to obtain stationarity ($d = 1$), otherwise with d equal to 0.
6. We create ARIMA models using p , d , and the trend where applicable, and we iteratively test the number of lagged forecast errors q from 0 to 3.
7. The choice of the best ARIMA (p, d, q) model is eventually made using AIC statistics. We acknowledge the limitation derived from running a thorough statistical routine for all the market-commodity pairs in our price database, without doing more in-depth market-specific price analyses, including controlling for structural breaks.

Seasonal Holt-Winters (SHW): This method is a variant of exponential smoothing, which is a procedure for continually revising a forecast considering more recent data. We apply the SHW multiplicative method, using a forecast equation and three-level equations – one for the level, one for the trend, and one for the seasonal component.

Intuitively, this method averages backward past observations, attaching more importance to the most recent ones. It also introduces a trend estimate that changes over time and a seasonal component. The major limitations are as follows: *i*) the trend can dominate the forecasts after a short period; *ii*) the missing values need to be estimated; and *iii*) the outliers strongly affect the forecasting. This is a widely used approach for short-term forecasts and normally provides the best forecasts of the three approaches proposed here.

Moving Average Smoother corrected with Grand Seasonal Index changes:

This method equally weights past (and future) observations by deriving the (centred) moving average (CMA) in a rolling window of 13 months, looking backwards and forwards. We compute the seasonal indices dividing each observation by its CMA value. We then compute the grand seasonal index (GSI), which is the mean by month of the seasonal indices. Finally, we compute the monthly changes of the GSI. Monthly forecasts are therefore based on the previous known price adjusted by its value times the corresponding GSI change.

3 How the Alert for Price Spikes works

"The natural price, therefore, is, as it were, the central price, to which the prices of all commodities are continually gravitating. Different accidents may sometimes keep them suspended a good deal above it, and sometimes force them down even somewhat below it." A. Smith, The Wealth of Nations

Food price crises are correlated with food security crises. The early detection of rising prices supports decision-making and early action. The Alert for Price Spikes (ALPS) allows humanitarian decision-makers and food security analysts to be aware of the latest abnormal changes in food prices and to take further monitoring or food security programming decisions. ALPS is an indicator that monitors the extent to which a local food commodity market experiences unusually high food prices. It shows abnormal price levels of selected staple foods. It compares the level of monthly food prices (both actual and forecast) against estimated seasonal trends and categorizes them as *normal*, *stress*, *alert* or *crisis*.

The ALPS allows comparison between markets over the same period, or study of the same market over different periods. The [methodology](#) is based on [empirical work](#) undertaken at the Centre of Research and Studies on Economic Development (CERDI). We retain only those foods with at least three years long time series and have less than 30 percent of values missing both for the entire time series and for the last three years of data. We only show time series with a minimum of price variability, given by two indicators: the coefficient of variation, as described by, $cv = \sigma/\bar{p}$, and the highest price variation, as described by

$$pv = \max(p_t^H - p_t^L) / p_t,$$

where σ is the standard deviation of prices, $\bar{p}$ is the average price, and p_t^H and p_t^L are respectively the highest and the lowest price observation. For both the coefficient of variation and the price variation, we control for the entire time series and the last three years of data, allowing the tool to show all time series that have at least one of the four indicators above an arbitrarily defined threshold. Specifically, for cv we set a threshold of 15 percent for the entire time-series, and 10 percent for the last three years. For pv , we define a threshold of 50 percent for the entire time-series, and 30 percent for the last three years.

References

- [1] Araujo C., Araujo-Bonjean C., and S. Brunelin (2012). Alert at Maradi: Preventing Food Crises by Using Price Signals, in World Development, Vol. 40(9), pp-1882-1894.
- [2] Holt, C.C. (1957). Forecasting Trends and Seasonals by Exponentially Weighted Averages, Carnegie Institute of Technology, Pittsburgh ONR memorandum no. 52.
- [3] Maddala G.S., and I.-M. Kim (1998). Unit Roots, Cointegration and Structural Change, Cambridge University Press.
- [4] WFP (2014). Calculation and Use of the Alert for Price Spikes (ALPS) Indicator, Technical Guidance Note, Analysis and Nutrition Service.