Agriculture Planning and Food Demand

By Group Decisive

From the University of Toronto

1. Introduction

Agricultural development is often influenced by government policies such as specialized operations planning, production capacity estimation and farming regulations. Agricultural policies are used to ensure an adequate food supply, price stability, product quality and manage the use of agricultural land.

However, agricultural operations are also businesses that want to maximize profits. The population is the consumer of the product, and seeks consistent and affordable food to satisfy nutritional needs and dietary preferences. Governments aim to balance and optimize the supply and demand equilibrium [1].

Dietary preferences are influenced by several factors including tradition, food prices, and financial status. In developing countries where the economy is rapidly evolving and the middle class is expanding, dietary preferences, and therefore demand for certain foods is highly dynamic. [2]

Therefore, there is significant incentive for governments to understand and predict supply and demand. Inefficiencies in this system can result not only in direct financial losses but in health issues for the population [2].

Health issues can include micronutrient malnutrition and child stunting, as well as obesity and issues associated with being overweight.

1.1 Causes of Inefficiency

In many cases, agricultural policy is currently biased towards the staple grains - rice, wheat and maize. Staple grain-self-sufficiency was a major part of resolving food insecurity problem after the 2008 food price crisis [2]. Policy decisions after 2008 still define food security in terms of these staple grains, creating a disconnection between policy and nutritional needs and preferences. [2]

Policy is slow to respond to this disconnect, resulting in negative financial and nutritional outcomes [10]. The time lag between capturing the sub-optimality of a production plan and responding accordingly can lead to substantial inefficiency.

2. Problem Scope and Hypothesis

There is an inconsistency between existing agricultural policy and production requirements/consumer needs that is resulting in negative nutritional and financial outcomes. There is a need for a system to analyze and forecast supply and demand trends in order to guide policy.

Hypothesis: There is a significant discrepancy between agricultural production and dynamic consumer demand, caused by agricultural policy in BRICS countries.

To initially test this hypothesis and motivate the study, data from five developed and five developing countries was compared. Main crop categories studied include: meat, cereal and vegetables. Only the five developing countries were used for further assessment. A detailed rationale is provided in Section 3.1.

To test this hypothesis data from both developing and developed countries was used. Five major developing and five developed countries were analyzed and tested. Main crop categories discussed in this case include: meat, cereal and vegetables. However, only the five developing countries are used for further analysis and prediction. Detailed rationale is explained in Section 3.1.

3. Methodology and Discussion

3.1 Sampling analysis

A total of 10 randomly selected developing and developed countries are compared to test the hypothesis. Historical supply and demand data for each agricultural product demonstrates that the inefficiency (gap) between supply and demand is significantly smaller in developed countries compared to developing countries.

For example, Figure 1 shows the supply and demand relationship of all selected agricultural products in U.S. All products except cereal demonstrating equilibrium in supply-demand quantities. However, a direct comparison of historical supply and demand between major developing countries show large differences in Figure 2.

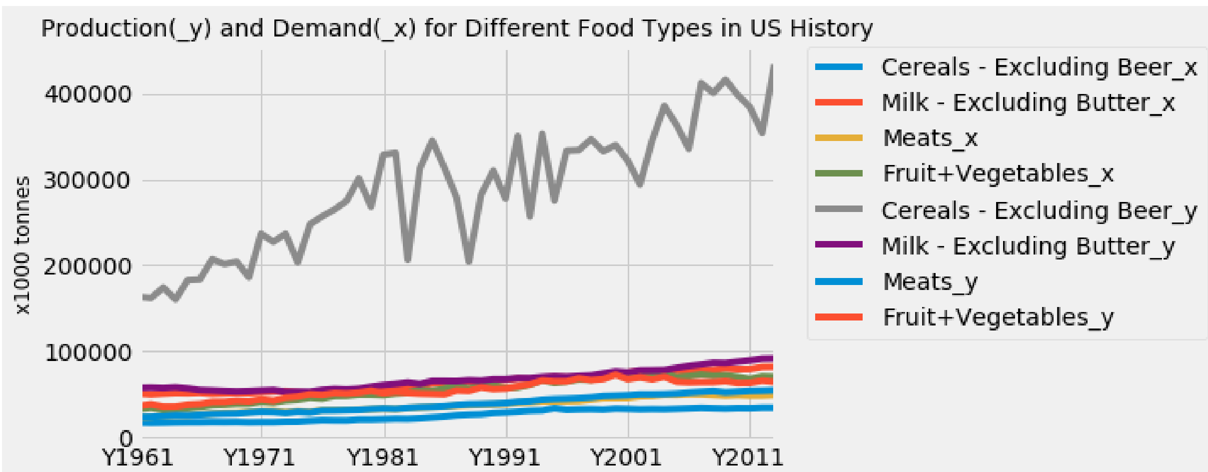


Fig. 1 Production and Demand for Different Food Types in US History

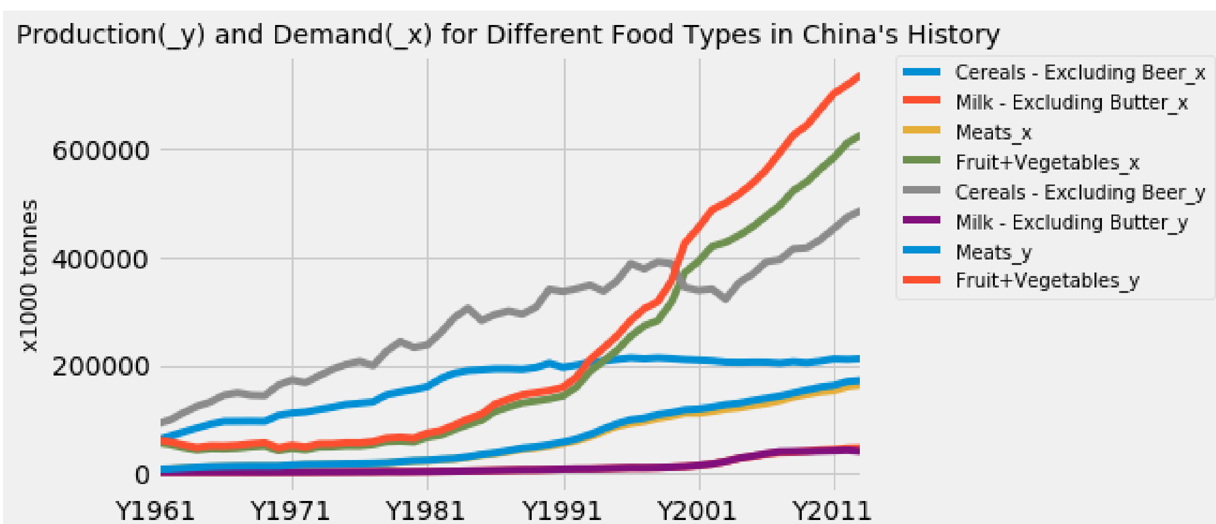


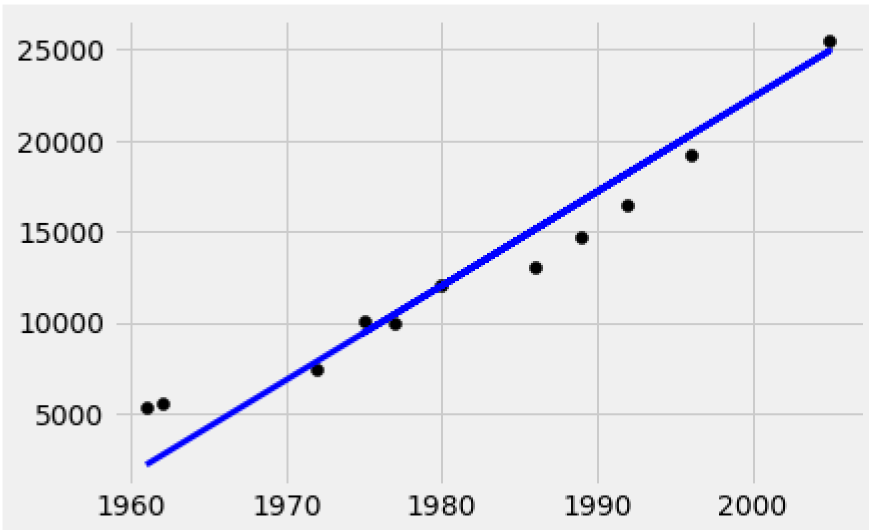
Fig. 2 Production and Demand for Different Food Types in China History

Based on the observations in Graphs 1 and 2 as well as existing literature [sources] it can be suspected that the discrepancies between supply and demand in developing countries are caused by evolving economies, the growing middle class (increasing purchasing power), and changing diets compared to relatively stable incomes and diets in developed countries [1][3].

Therefore, developing countries are best suited for our analysis in resolving the inefficiency in food waste and sustainable sourcing problem by minimizing the supply-demand gap.

3.2 Data Analytics

At first glance, the data in Graphs 1 and 2 is not linearly correlated. However, by feature space transformation the data can be fitted using linear regression. This method can be illustrated by picturing a circle, which is a non-linear function on the Cartesian plane. By transforming from Cartesian to spherical coordinates, the circle can be described by a linear function. This method can apply here. Therefore, linear regression is applied and the results are highly accurate (avg 0.95 R^2 score).



Coefficients: [517.03879045]
Mean squared error: 2878651.08
Variance score: 0.92

As a brief summary of technical methodology, the data for each country and each food type is organized in input-output pairs, where input is year value and output is either production or demand quantity in units of 1000 tonnes. The data pairs are split into 80% training set and 20% testing set using standard procedure. Scikit-learn, a machine learning library in Python, automatically fits model on training data, and evaluates the model on testing data. Model accuracy is obtained in the form of R^2 score. Using the same model parameters learned from training, a prediction on next year quantity is interpolated.

3.3 Results Integration

The confidence interval of the predicted demand of agricultural products in China from section 3.2 can be calculated based on the following equation:

CI = x_mean +/- Z*STD

The confidence interval can give an upper bound and a lower bound for agriculture demand, which can be viewed as different scenarios for demand calculation, high demand scenario (demand = demand_mean + z*std), median (demand_mean), and low demand scenario (demand_mean - z*std).

For demonstration of results and ease of computation, later in the formulation of our stochastic program in 3.4, it is assumed that each of the scenarios captures the same probability (⅓). The predicted demand in each of the three different scenarios is determined by assuming that the prediction follows a normal distribution, then by dividing the area under the normal curve by three, and finding the exact middle point demand of each of the three areas. In that case, the Z score for calculating the confidence interval is estimated based on the middle point of the horizontal axis of each scenario, that is, z(-⅙)=-0.97, z(0)=0, and z(⅙)=0.97. Following this, the demand in each of the three scenario is determined by:

High demand = demand_mean + 0.97*std

Mid demand = demand_mean

Low demand = demand_mean - 0.97*std

3.4 Two-Stage Stochastic Modeling and Resource Optimization

Taking the demand prediction obtained from the data analysis section, a stochastic model is proposed to tailor recommended supply to accommodate possible uncertainties in predicted demand. The supply quantity can be viewed as the unknown and the cost of supply as the resource in classic stochastic modeling, where the model represents a set of decision variables x's that must be taken without full information regarding the changing equilibrium of supply and demand. In this case, the decision (variables) is designed to be the actual supply quantity of each type of agricultural product and then a second-stage set of resources decisions is designed to be the predicted demand based on linear regression [4].

By applying the same methodology, the demand-supply program can be formulated in a similar way:

$$minimize\ (c_A x_A + c_B x_B + c_C x_C + c_D x_D + \tfrac{1}{3}(uncertainty\_upper) + \tfrac{1}{3}(uncertainty\_mean) + \tfrac{1}{3}(uncertainty\_lower))$$

$$subject\ to:\ x_A, x_B, x_C, x_D \geq 0 \qquad \forall y^-_{ABCD,i} \geq 0,\ \ \forall y^-_{ABCD,i} \geq 0$$

Scenario i:

$$x_A + y^-_{A,1} - y^+_{A,1} = D_{A,1}\ \ x_B + y^-_{B,1} - y^+_{B,1} = D_{B,1}\ \ x_C + y^-_{C,1} - y^+_{C,1} = D_{C,1}$$

$$x_D + y^-_{D,1} - y^+_{D,1} = D_{D,1}$$

Scenario 2:

$$x_A + y^-_{A,2} - y^+_{A,2} = D_{A,2}\ \ x_B + y^-_{B,2} - y^+_{B,2} = D_{B,2}\ \ x_C + y^-_{C,2} - y^+_{C,2} = D_{C,2}$$

$$x_D + y^-_{D,2} - y^+_{D,2} = D_{D,2}$$

Scenario 3:

$$x_A + y^-_{A,3} - y^+_{A,3} = D_{A,3}\ \ x_B + y^-_{B,3} - y^+_{B,3} = D_{B,3}\ \ x_C + y^-_{C,3} - y^+_{C,3} = D_{C,3}$$

$$x_D + y^-_{D,3} - y^+_{D,3} = D_{D,3}$$

$$uncertainty\_upper = c^-_{A,1} y^-_{A,1} + c^+_{A,1} y^+_{A,1} + c^-_{B,1} y^-_{B,1} + c^+_{B,1} y^+_{B,1} + c^-_{C,1} y^-_{C,1} + c^+_{C,1} y^+_{C,1} + c^-_{D,1} y^-_{D,1} + c^+_{D,1} y^+_{D,1}$$

$$uncertainty\_mean = c^-_{A,2} y^-_{A,2} + c^+_{A,2} y^+_{A,2} + c^-_{B,2} y^-_{B,2} + c^+_{B,2} y^+_{B,2} + c^-_{C,2} y^-_{C,2} + c^+_{C,2} y^+_{C,2} + c^-_{D,2} y^-_{D,2} + c^+_{D,2} y^+_{D,2}$$

$$uncertainty\_lower = c^-_{A,3} y^-_{A,3} + c^+_{A,3} y^+_{A,3} + c^-_{B,3} y^-_{B,3} + c^+_{B,3} y^+_{B,3} + c^-_{C,3} y^-_{C,3} + c^+_{C,3} y^+_{C,3} + c^-_{D,3} y^-_{D,3} + c^+_{D,3} y^+_{D,3}$$

Where xA, xB, xC, xD represent the normal supply required for product type A (Cereal), B (Milk), C (Meat), D (vegetable & fruits); yABCD,i+represent the possible surplus amount in each product type, and yABCD,i- represent the possible shortage amount in each product type.

The stochastic solution acts as a hedge against uncertainty by considering all possibilities of demand. It therefore generates a balanced production and purchasing strategy across the agriculture types. This highlights the differences between deterministic and stochastic solutions.

4. Proposed Solution and Outcome

Due to the large amount of data available, and to focus our solution, China is used as one live example to demonstrate the optimality of the proposed model. In order to compare the optimality, the team has showed two different approaches: 1. Net profit approach, 2. Value in Food loss approach.

For the first net profit approach, the net profit is calculated based on the formula below,

Net profit = sales revenue by selling the product - production costs estimated using the supply quantity

For the value in food loss approach, the value waste in food loss can be divided into two situations, supply produced is greater than demand, and vice versa. If the former overproduction exists, the value in waste loss is estimated to be the total production cost of the wasted food amount[(S-D)*Production cost]. If the latter case exists, the value in food loss is estimated to be the potential earning assuming the demand is satisfied [(D-S)*net profit].

4.1 Proposed Solution and Profitability Indicator

Using the optimized supply and demand, a simple net profit calculation gives that the optimized profit is computed to be $17.89 billion, which is significantly greater than the prediction based on historical figures without optimization.

In addition, the computed food waste using the proposed model is largely reduced by 2.78% compared to the sub-optimal method. A detailed savings in food loss can be found in the Table 1 and 2 below,

Category	Proposed Optimal Supply (million ton)	Predicted Demand (million ton)	Gap Value/Loss in Waste ($ mil)	Net Profit ($ bil)
Cereal	159.53	249.38	$ 387,110,284.83	$ 6.05
Milk	31.92	31.94	$ 8,463,112.78	$ 4.05
Meats	249.37	159.55	$ 117,491,339.26	$ 4.67
Vegetables & Fruits	34.65	34.73	$ 4,927,699.40	$ 3.13
Total	475.47	475.60	**$ 517,992,436.26**	**$ 17.89**

Table 1. Predicted Net Profit and Value in Loss with Proposed Optimal Solution

Category	Predicted Supply (million ton)	Predicted Demand (million ton)	Gap Value/Loss in Waste *$ mil)	Net Profit ($ bil)
Cereal	488.59	249.38	$ 159,123,354.52	$ 2.05
Milk	35.48	31.94	$ 786,185.16	$ 5.18
Meats	165.00	159.55	$ 12,301,842.32	$ 20.22
Vegetables & Frui	660.67	34.73	$ 361,118,186.90	-$ 34.02
Total	1349.74	475.60	**$ 533,329,568.90**	**-$ 6.57**

Table 2. Predicted Net Profit and Value in Loss without Optimization

The total profit earned from implementing the solution is simply the difference in net profit between the two shown methods. Comparing the two results, the proposed solution can help the country gain more than $24.46 billion (total profit from solution = net profit using the proposed method - net profit of the predicted trend).

If viewing the implementation of the model as a technology project, the expense of this solution consists of spending in model development, front end (user interface) development and end testing. The market pricing for such project and roles can be found below:

4.4 Business Solution

The implementation of the solution could be viewed as a large-scale technical business project, which has many associated costs, especially if it will be implemented on a national level. The costs of the project include IT development (database and model), creating a user interface, testing, administration, and the ongoing costs for administration and system maintenance. Continuous research and development will also be a cost, as the model should be refined with new data. This could include the costs of conducting surveys to obtain current supply and demand data.

To obtain an estimate for the IT development costs, the team interviewed the IT Outsourcing Department of Deloitte to inquire about specific pricing. It is assumed that the development of the program will occur in all 663 cities in China. In comparison to similar projects, a project of this scale would take 15 000 to 18 000 hours per region. This data has been used to estimate pricing, taking into account the upper and lower bound values as well as the median.

The actual pricing is estimated as below:

		Hours Required for 3 Scenarios			Total Spending (Hours*Price)		
Role	Hourly Price	High (hr)	Median (hr)	Low (hr)	High	Median	Low
IT Analyst	$ 150.00	11934000	10939500	9945000	$ 1,790,100,000	$ 1,640,925,000	$ 1,491,750,000
Consultant	$ 500.00	11934000	10939500	9945000	$ 5,967,000,000	$ 5,469,750,000	$ 4,972,500,000
Developer	$ 300.00	11934000	10939500	9945000	$ 3,580,200,000	$ 3,281,850,000	$ 2,983,500,000
				Total ($bil)	$ 11.34	$ 10.39	$ 9.45

Table 3. IT Technology Development Cost Estimation*

*Pricing estimation based on Deloitte IT Outsourcing standard

In addition to the IT development costs, government administration and continuous monitoring of operations would be required. These costs are estimated based on historical ranges of government spending in 43 developing countries.

Surveying costs are estimated using historical census costs for the past 50 years in China as the solution will be implemented across all regions of China. For each sub-group of the cost regions defined, the upper and lower bounds as well as the median costs are used to capture the range of possibilities.

A comprehensive pricing table is generated as following,

Cost Type	Total Spending ($bil) High	Median	Low
IT Development	$ 11.34	$ 10.39	$ 9.45
Government Administration/Monitoring* (Estimated by historical avg of Government Spending in 43 Developing Countries)	$ 5.30	$ 3.95	$ 2.60
Survey Spending* (Estimated by historical Census Cost)	2.20	$ 1.45	$ 0.70
Total	$ 18.84	$ 15.79	$ 12.75

Table 4. Comprehensive Expense for Solution Implementation [5][6]

Based on the estimated cost range, the project ROI can be computed as follow,

Total Expense for Implementatio	High	18.84	Total Profit from Solution($bil) *(Optimal Net Profit - Predicted Net Profit)	24.46	Net Return (Net Profit from solution - Total Expense)	High	5.62
	Median	15.79				Median	8.67
	Low	12.75				Low	11.71

Table 5. Summary of Data Points Computed (Net profit from solution = Optimal Net Profit Calculated - Predicted Return Calculated)

ROI	High	29.85	Payback Period (Years)	High	3.35
	Median	54.88		Median	1.82
	Low	91.88		Low	1.09

Table 6. ROI and Payback Period Summary

In conclusion, from a business perspective, the proposed solution can help government significantly reduce agricultural waste. As shown in the tables, in the lower bound cost scenario, the ROI reaches 91.88 and the minimum payback period is 1.09 years, and at the upper bound cost scenario, the minimum ROI is 29.85 with a maximum payback period of 3.35 years. The median case falls in between the two extremes.

5. Limitation and Future Work

Although this analysis demonstrates the motivation for improved production planning and our model provides recommendations for improvement, it has substantial limitations in its current form.

The model uses data provided by the FAO on food consumption to quantify and predict the demand. However this demand can be inherently influenced by current food availability and cost. For example, the price of a product consumer & demand could be inflated due to limited supply, resulting in lower consumption. Alternative methods such as surveying could improve their demand prediction. Consumer data from grocers could also improve predictions and incorporate geographical differences.

The current stochastic model is only able to predict one period. Expanding the model for longer time periods could help make long-term predictions. Comparing supply and demand directly could overlook frictional factors such as inflation, consumer price index, and import/export patterns. The impact of these factors would need to be evaluated.

- Improve accuracy on demand prediction by surveying on how residence's nutrition preference and daily need
- Expand the current stochastic model from one period to multi-period so as to make predictions on future ongoing periods
- Consider more frictional factors into the model such as inflation, consumer price index, import and export

6. Conclusion

The proposed stochastic model generates an implementation ready solution. We have shown that decreasing inefficiencies in supply and demand is profitable, will reduce food loss, and have positive nutritional outcomes.

By using this model, governments will have more predictive capacity to make policy decisions and reduce the lag time between identifying inefficiency and implementing solutions.

The proposed stochastic modeling solution to capture all

- Implementation ready solution
- Profitable
- Contribute to sustainable sourcing by reducing food loss

Appendix

A. Commodity Price of Agriculture Product Estimated Based on Market Price in China

Category	Price per ($/million ton)	Cost per ($/million ton)
Cereal	$ 1,385.30	$ 665.21
Milk	$ 1,867.47	$ 222.27
Meats	$ 3,599.02	$2,254.31
Vegetables & Fruits	$ 1,179.87	$ 576.92

Reference

[1] http://www.gov.mb.ca/mr/plups/pdf/prg_agri.pdf

[2] Agricultural policy and nutrition outcomes – getting beyond the preoccupation with staple grains Prabhu Pingali 1

[3] “Eating well with Canada's Food Guide.” Eating well with Canada's Food Guide - Eat Right Ontario, www.eatrightontario.ca/en/Articles/Canada-s-Food-Guide/Eating-well-with-Canada-s-Food-Guide.aspx

[4] Introduction to Linear Optimization, Dimitris Bertsimas and John N. Tsitsiklis

[5] http://www1.worldbank.org/publicsector/pe/pfma07/OPMReview.pdf

[6] china statistical databse http://www.stats.gov.cn/tjsj/pcsj/