Guide to optimization model of crop flow

Mohammad Ahnaf Sadat

This document provides a comprehensive overview of the "optimization of crop flow". It is structured around various questions related to the topic, and we focus on addressing the following specific questions:

1. Why is the optimization of crop flow needed?
2. What expertise is required to develop mathematical models?
3. Describe the optimization model for the project.
4. How to obtain the optimal solution for the model?
5. Explain all the files associated with "optimization of crop flow."
6. What can be done in the future as extensions for the project?

1. Why is the optimization of crop flow needed?

Crops' prices vary significantly between counties and cities, creating opportunities to maximize profits by strategically selling them. When counties act alone, they aim to sell to the highest-paying counties with the consideration of transportation costs. However, if all counties follow this approach, it can lead to an oversupply, lowering prices and causing crop shortages in some areas.

Farmers (counties) need to adopt a collaborative strategy and explore optimization techniques to avoid such issues. This ensures a balanced distribution of crops and maximizes overall profit while avoiding market imbalances.

In this case, optimizing crop flow means determining the best amount of crops each county should transport to other counties.

2. What expertise is required to develop mathematical models?

Expertise in operation research and/or linear programming is necessary to develop the mathematical model. Although the book is old, the following provides a comprehensive guide on how to build the mathematical model using linear programming:

Hillier, F. S., & Lieberman, G. J. (2001). *Introduction to operations research*. McGraw Hill Companies, Inc.

Students in Industrial Engineering, Supply Chain Management, Computer Science, Operations Research, and Mathematics might be familiar with developing mathematical models.

3. Describe the optimization model for the project.

A mathematical optimization model typically consists of four main components: objectives, constraints, parameters, and variables. Below, we will describe the project's optimization model:

The project's optimization model aims to maximize the combined profit generated by the counties through adjustments in the crop distribution strategy. The model considers the demand and supply of harvested crops in each county, as well as the price each county is willing to pay. Please note that the production cost has been disregarded in the model's development, as it is unrelated to the distribution and does not impact the solution. The components of this optimization model are as follows:

Variables:

$$x_{i,j} = The\ quantity\ of\ products\ transported\ from\ county\ i\ to\ county\ j$$

$$R_j = The\ quantity\ of\ the\ products\ recieved\ in\ county\ j$$

Parameters:

$$N = Total\ number\ of\ counties$$

$$P_{conversion} = Metric\ tons\ to\ pounds\ conversion\ (2204.62)$$

$$P_i = Purchasing\ price\ of\ 1\ lb\ of\ crops\ in\ county\ i\ (\$/lb)$$

$$S_i = The\ harvested\ amount\ of\ crops\ in\ county\ i\ (Metric\ Tons)$$

$$D_i = The\ demand\ of\ crops\ in\ county\ i\ (Metric\ Tons)$$

$$C_{conversion} = The\ carrying\ cost\ of\ crops\ per\ unit\ weight\ per\ unit\ length$$

The $C_{conversion}$ for a Ford F-350 truck with a carrying capacity of 7640 lbs, a mileage of 15 miles/gallon, and a per gallon cost of $3.5 would be approximately $\frac{3.5}{15\times(\frac{7640}{2204.62})}\ \$/Metric\ Ton/Mile$.

$$c_{i,j} = The\ travelling\ distance\ between\ county\ i\ and\ j$$

The traveling distance has been calculated using the Euclidean distance formula based on the longitude and latitude coordinates. This calculation yields the distance in degrees. To convert this distance to miles, we simply multiply it by the conversion factor of 69 miles/degree.

Objective Function:

$$\max_{x} \sum_{i} P_{conversion} \times P_i \times R_i - \sum_{i,j} C_{conversion} \times x_{i,j} \times c_{i,j} \tag{1}$$

In expression (1), the first term represents the total revenue obtained by selling the crops, while the second term represents the total transportation cost.

The revenue generated from a county is calculated by multiplying the quantity of crops received in one county by the price ($P_{conversion} \times P_i$) it offers for those crops. This calculation is performed for each county, and the revenues from all counties are then summed to determine the total revenue.

The transportation cost is calculated based on the quantity of crops transported between counties multiplied by the transportation cost ($C_{conversion} \times c_{i,j}$). It is obtained by summing the costs for all destination and origin counties, resulting in the total transportation cost.

Constraints:

$$\sum_{i \in N} x_{i,j} = R_j \qquad \forall j \in N \qquad (2)$$

$$\sum_{j \in N} x_{i,j} \leq S_i \qquad \forall i \in N \qquad (3)$$

$$R_j \leq D_j \qquad \forall j \in N \qquad (4)$$

Equation (2) states that the sum of all transported quantities from all other counties to county *j* equals the quantity of crops received at county *j*.

Equation (3) asserts that the total transportation from a specific county must be less than or equal to the quantity of crops harvested in that county.

Equation (4) asserts that the quantity of crops received at county j must be less than or equal to the demand of that county.

4. How to obtain the optimal solution for the model?

To solve the mathematical model, the Gurobi Optimizer is utilized through Python. While Gurobi is a paid software, it offers a 1-year license for academic use, which can be renewed annually for free.

For those interested in exploring a basic optimization model implemented in Python with Gurobi, a reference example can be found in the following link:

https://www.gurobi.com/documentation/current/quickstart_mac/cs_example_mip1_py.html

This example provides a starting point for using Gurobi to solve optimization problems.

5. Explain all the files associated with "optimization of crop flow".

We have three scripts for this project:

i. opti.py

This file presents a simple outline of the optimization model. To adjust the number of counties (N), we can modify the variable "nc".

In this script, the counties' location, demand, and harvested units are generated randomly.

The script generates the optimal distribution of the crop flow, which can be found in "results.csv". Additionally, it displays a graph illustrating the solution.

ii. **opti_Data_fed.py**

This script is identical to the previously mentioned "opti.py", with the sole difference that it retrieves data from the file "Subset_data.csv" instead of generating it randomly.

To utilize this script, we need to ensure that "Subset_data.csv" adheres to the following precise structure:

County Name	Demand	Supply	Price	Lattitude	Longitude
Boone	4.711	2.5954	1.76	42.08817	-93.9878
Benton	4.552	3.3618	2.61	42.13872	-92.0665
Plymouth	4.5467	3.6827	2.08	42.77667	-96.1527
Bremer	4.472	2.7171	2.16	42.76108	-92.3814
Washington	3.9961	3.1515	2.45	41.3478	-91.7539

The script produces the optimal crop flow distribution in two different formats:

"results.csv": This file contains the optimal distribution represented in matrix format.

"power_bi_data.csv": In this file, the script provides only the distributions that involve transferring harvested crops to a different destination. The data is formatted to be compatible with PowerBI's visualization plugin "flow map."

Additionally, the script displays a graph illustrating the solution.

iii. opti_Data_fed_profit_retained_ratio.py

This script is identical to the previously mentioned " opti_Data_fed.py "; however, we only get the variable profit_retained as output in this file.
In this context, we define the profit_retained as the ratio between the actual profit generated with the current strategy and the highest possible profit that could be achieved if there were only one supplier and all the receivers.

To obtain the highest possible profit for each county, we have defined the mathematical model as follows:

Objective Function:

$$\max_{x} \sum_{j\in N} P_{conversion} \times P_j \times x_{selected\_county,j} - \sum_{j\in N} C_{conversion} \times x_{selected\_county,j} \times c_{selected\_county,j}$$

Constraints:

$$\sum_{i\in N} x_{i,j} = R_j \qquad \forall j \in N$$

$$\sum_{j\in N} x_{i,j} \leq S_i \qquad \forall i \in N$$

$$R_j \leq D_j \qquad \forall j \in N$$

In the updated version of the mathematical model, the objective function has been modified to focus on maximizing the profit of a specific county, rather than the combined total profit of all counties. To achieve this, the model needs to be run multiple times, precisely equal to the number of counties. Each run of the model will be dedicated to optimizing the profit for one county, and the process needs to be repeated for all counties individually.

Later with the solution obtained in the combined strategy (original model), we can calculate the profit for each county as follows:

$$\sum_{j\in N} P_{conversion} \times P_j \times x_{selected\_county,j} - \sum_{j\in N} C_{conversion} \times x_{selected\_county,j} \times c_{selected\_county,j}$$

And with these two values, we can find out the ratio.

6. What can be done in the future as extensions for the project?

The followings are a few of the things that can we can consider in the future:

I. We can consider the real-world distance from Google Maps rather than the Euclidean distance.

II. In this model, we assigned a lower cost to partial truck loads compared to full truck loads. For future improvements, we can consider only integer numbers of trucks for travel, and the cost will be directly proportional to the number of trucks, regardless of whether they are fully or partially loaded. This type of problem can be modeled using integer linear programming.

III. We have the option to assign different weights to the profit of each county, aiming to balance the retained profit among all the counties in the combined strategy.

IV. We can integrate "where to plant the crops" into this crop flow optimization model.