

SUSTAINABLE BIOECONOMY
FOR ARID REGIONS



Biomass Supply Chain and
Transportation:
Instruction Manual



SYSTEMS AND INDUSTRIAL ENGINEERING
DEPARTMENT

By

Daniel A. ZUNIGA VAZQUEZ, Shunyu YAO, Neng FAN

February 20, 2023

Acknowledgments

This material is based upon funding provided by the

USDA-NIFA, Grant # 2017-68005-26867.

Any opinions, findings, conclusions, or recommendations expressed in this publication/work are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture. All authors state that there is no conflict of interest.

The authors would also like to acknowledge valuable input and assistance from the following individuals and organizations who kindly provided expertise, data, and/or feedback during the development of this instruction manual:

- Amy E. Landis (Colorado School of Mines)
- Blase Evancho (University of Arizona)
- Clark Seavert (Oregon State University)
- Evan Sproul (Colorado State University)
- Hailey M. Summers (Colorado State University)
- Jason C. Quinn (Colorado State University)
- Paul Gutierrez (New Mexico State University)
- Sita Khanal (New Mexico State University)
- Trent Teegerstrom (University of Arizona)
- VeeAnder Mealing (Colorado School of Mines)

All the data and source code in this manual can be found in the Github website: <https://github.com/omegayao/SBAR>.

Contents

1	Introduction	6
1.1	Overview	6
1.2	Need for a sustainable bio-economy for arid regions	8
1.3	Purpose of Report	8
1.4	Required Software and Installation Guides	8
2	Integrating Environmental and Social Impacts into Optimal Design of Guayule and Guar Supply Chains	11
2.1	Model Structure	12
2.2	GIS Data and Geoprocessing Tools	13
2.2.1	Importing Data	13
2.2.2	Customization of GIS Data	15
2.2.3	Distance Matrix and Printing Formats	16
2.3	Defined Parameters	18
2.3.1	Identifying processing facility candidates	20
2.3.2	CPLEX Parameter File Creation	21
2.4	Optimization Tools and Methods	23
2.4.1	CPLEX Optimizer Through a High Performance Computer	23
2.4.2	Reporting Outputs	24
2.5	Model Assumptions	24
2.6	Extended Results for the American Southwest	24
3	Optimal Production Planning and Machinery Scheduling for Semi-Arid Farms	27
3.1	Model Structure	28
3.2	GIS Data and Geoprocessing Tools	29
3.3	Defined Parameters	30
3.3.1	CPLEX Parameter File Creation	31
3.4	Optimization Tools and Methods	33
3.4.1	Reporting Outputs	33
3.5	Model Assumptions	34
4	Demand-Driven Harvest Planning and Machinery Scheduling	35
4.1	Model Structure	36
4.2	GIS Data and Geoprocessing Tools	37

<i>CONTENTS</i>	3
4.3 Defined Parameters	38
4.3.1 Input Parameter File Creation	39
4.4 Optimization Tools and Methods	41
4.4.1 Gurobi Optimizer Through a High Performance Computer	41
4.4.2 Reporting Outputs	43
4.5 Model Assumptions	44
5 Conclusions	45
Bibliography	47

List of Figures

1.1	BSCT overall structure	7
2.1	Biomass supply chain optimal design: model methodology	12
2.2	CropScape information	13
2.3	GIS data import	14
2.4	GIS data customization	15
2.5	GIS distance matrix and map output	16
2.6	Input parameter walkthrough: Step 1	21
2.7	Input parameter walkthrough: Step 2	22
2.8	Input parameter walkthrough: Steps 3 and 4	22
2.9	Running CPLEX optimizer in the HPC (Part 1)	23
2.10	Running CPLEX optimizer in the HPC (Part 2)	23
2.11	Arizona's GIS information	25
2.12	New Mexico's GIS information	25
2.13	Northwestern Texas' GIS information	26
3.1	Production planning and multi-machinery scheduling: model structure (part 1)	28
3.2	Production planning and multi-machinery scheduling: model structure (part 2)	28
3.3	Input parameter walkthrough: Step 1	31
3.4	Input parameter walkthrough: Step 2	32
3.5	Input parameter walkthrough: Steps 3 and 4	32
4.1	Harvest planning and machinery scheduling: model structure	36
4.2	Input parameter walkthrough: Step 1	39
4.3	Input parameter walkthrough: Step 2	40
4.4	Input parameter walkthrough: Steps 3 and 4	40
4.5	Running Gurobi optimizer in the HPC (Part 1)	41
4.6	Running Gurobi optimizer in the HPC (Part 2)	42
4.7	Running Gurobi optimizer in the HPC (Part 3)	42

List of Tables

2.1	Biomass Supply Chain Optimal Design: Sets and Indices	18
2.2	Biomass Supply Chain Optimal Design: Input Parameters	19
2.3	Biomass Supply Chain Optimal Design: Decision Variables	24
3.1	Optimal Production Planning and Scheduling: Sets and Indices	30
3.2	Optimal Production Planning and Scheduling: Input Parameters	30
3.3	Optimal Production Planning and Scheduling: Decision Variables	33
4.1	Harvest Planning and Machinery Scheduling: Sets and Indices	38
4.2	Harvest Planning and Machinery Scheduling: Input Parameters	38
4.3	Harvest Planning and Machinery Scheduling: Decision Variables	43

Chapter 1

Introduction

1.1 Overview

The Biomass Supply Chain and Transportation (BSCT) optimization tool has been developed by the Sustainable Bioeconomy for Arid Regions (SBAR) project in support with the U.S. Department of Agriculture (USDA) to evaluate scenarios for crop adoption and their scaled-up to profitable production.

For the economical, environmental, and social impacts analyses of the biomass supply chains for the adopted crops, BSCT uses stochastic scenarios to capture the uncertainties of the adoption rates of each field. Afterward, a stochastic optimization is deployed to identify optimal decisions for facility locations, transports from fields to facilities, and finally to customers, with a multi-objective approach to quantify the economic benefits (minimizing the costs of supply chains), environmental impacts (minimizing CO₂ equivalent greenhouse gas emissions), and social impacts (maximizing the local accrued jobs). A Geographic Information System is used for capturing field information and relevant factors, as well as identifying potential processing facility locations.

Once the optimal processing facility location and supplying farms have been identified, BSCT can perform a detailed analysis of the farm production planning and machinery scheduling can through its optimization model for perennial and non-perennial crops that maximizes the farmers' net present value. Such optimization model allows for multi-crop production planning, multi-machinery scheduling, and crop rotation. To address the water scarcity problem in arid and semi-arid regions, perennial and nonperennial low-water-use crops can be considered in the crop rotation. Based on the Geographic Information System, the optimal scheduling routes are determined and the irrigation water requirements are analyzed.

Fig. 1.1 presents the overall structure of BSCT. Step 1 focuses on defining the sets and input parameters for the integration of the economical, environmental, and social impacts into the design of the biomass supply chains. Step 2 identified the optimal processing facility locations and the expected production

from the supplying farms based on the optimization model. Step 3 defines the sets and input parameters for the production planning and machinery scheduling. Step 4 identifies the optimal production planning with crop rotation and multi-machinery scheduling through the optimization model. Two test scenarios are provided for the Steps 1 and 2 and two test scenarios are provided for Steps 3 and 4 as follows:

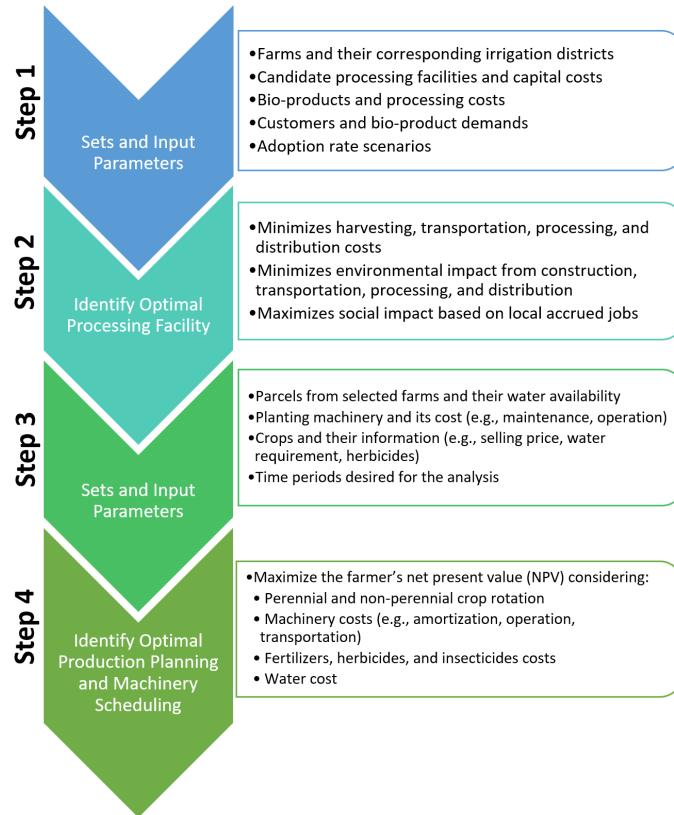


Figure 1.1: BSCT overall structure

- Guayule optimal processing facility location and expected farm production for Pinal and Maricopa counties in Arizona (refer to [1]).
- Guar optimal processing facility location and expected farm production for Dona Ana County in New Mexico (refer to [1]).
- Guayule optimal production planning and multi-machinery scheduling for Pinal County in Arizona (refer to [2]).
- Guar optimal production planning and multi-machinery scheduling for Cochran and Hockley counties in Texas (refer to [2]).

1.2. NEED FOR A SUSTAINABLE BIO-ECONOMY FOR ARID REGIONS8

- Guayule optimal harvest planning and machinery scheduling for Pinal County in Arizona (refer to [3]).

1.2 Need for a sustainable bio-economy for arid regions

Increasing competition in the agricultural sector is causing water to be a more scarce resource. In the U.S., approximately 80% of the consumptive water use is accounted for agriculture and it reaches up to 90% in the Western States [4]. Furthermore, frequent droughts and climatic change also increase the pressure for limiting the agricultural sector's water use [5]. The situation worsens in arid and semi-arid regions with low rainfall and high evaporation [6]. Therefore, ensuring efficient water use is an urgent problem to be solved [7]. To address this issue, farmers can gradually adopt low-water-use crops such as guar and guayule. These two crops are starting to create an interest for researchers and industries as a biomass resource [8].

Guayule and guar are low-water-use, drought-tolerant, and heat-resistant crops. Guayule is native to the Chihuahuan Desert, from southwestern U.S. regions to the northern Mexico, and it has been often treated as a source of natural rubber. Due to emerging economies in countries around the world, the demand for natural rubber has proliferated, and the guayule's natural rubber has been found to be an attractive commercial option as a domestic rubber source [9]. Guayule has even been declared as "a critical agricultural material in times of crisis" [10]. Similarly, guar is an important legume native from India and the source of guar gum. Partially due to its stabilizing and thickening properties, guar gum and its derivatives are essential in numerous applications such as paints, batteries, pharmaceuticals, ceramics, food, petroleum, cosmetics, textiles, and agriculture [11, 12]. BSCT is developed as an effort to achieve a scale-up to profitable production for guayule and guar.

1.3 Purpose of Report

BSCT will enable farmers and researchers assess the adoption of crops based economical, environmental, and social impact analysis. BSCT will also enable farmers evaluate the economical feasibility production planning and scheduling programs of the adopted crops. For such purpose, this instruction manual shows a step by step on the know-how to use BSCT and understand its results.

1.4 Required Software and Installation Guides

The following software is required by BSCT. Note that a license is required in several cases, while other are available for free. Download and detailed information links are provided for each case.

- **ArcGIS Pro**

ArcGIS Pro is a desktop Geographical Information System (GIS) software

developed by Esri.

Related links:

<https://www.esri.com/en-us/arcgis/products/arcgis-pro/resources>

- **Duo Mobile**

Secure access to work and personal, cloud and on-premises apps with one simple app.

Related links:

<https://duo.com/product/multi-factor-authentication-mfa/duo-mobile-app>

<https://it.arizona.edu/documentation/install-duo-mobile-app-netid-plus#duo>

- **Cisco AnyConnect VPN**

Secure virtual private network (VPN) access.

Related links:

<https://www.cisco.com/c/en/us/products/security/anyconnect-secure-mobility-client/index.html>

<https://it.arizona.edu/service/ua-virtual-private-network-vpn>

- **PuTTY**

PuTTY is a free security shell (SSH) protocol and telnet client for Windows.

Related link:

<https://www.putty.org/>

- **WinSCP**

WinSCP is a free SSH File Transfer Protocol (SFTP) and File Transfer Protocol (FTP) client for Windows.

Related link:

<https://winscp.net/eng/index.php>

- **High Performance Computing (HPC)**

For an efficient solution of the large-scale models from BSCT, the use of a HPC is recommended. A service overview and guide for the University of Arizona HPC is provided in the following links:

Service overview: <https://it.arizona.edu/service/high-performance-computing>

User Guide: <https://public.confluence.arizona.edu/display/UAHPC/User+Guide>

- **CPLEX IBM ILOG CPLEX Optimization Studio** is an optimization software package.

Related link:

<https://www.ibm.com/analytics/cplex-optimizer>

- **Gurobi**

Gurobi Optimizer is a state-of-the-art solver for mathematical programming.

Related link:

<https://www.gurobi.com/>

Chapter 2

Integrating Environmental and Social Impacts into Optimal Design of Guayule and Guar Supply Chains

This research proposes a stochastic optimization to make optimal decisions for facility locations, transportations from fields to facilities, and finally to customers, with a multi-objective approach to quantify the economic benefits, environmental impacts, and social impacts. The adoption rates are integrated into the model as a random variable which allows the capture of the farmers' uncertainty to adopt guayule and guar in their fields. The quantified economic benefits include harvesting costs, transportation and distribution costs, production costs, and construction costs for the processing facilities. Environmental impacts are accounted for through CO₂ equivalent GHG emissions cost of the harvesting, production, transportation, distribution, and construction processes. Social impacts are quantified based on the local accrued jobs per county for harvesting, transportation, production, distribution, and construction. Additionally, a GIS analysis is integrated into the model's input parameters to identify potential candidates for processing facility location. Irrigation water requirements are analyzed through GIS irrigation district layers and their water availability. Independent weighted factors are considered within the quantification of the economic benefits, environmental impacts, and social impacts for sensitivity analysis purposes. This allows an informed decision making based on assigning higher weights on one or multiple economic, environmental, and social impacts as required. The proposed approaches are evaluated based on the cases of two areas: the Maricopa and Pinal counties in Arizona for the guayule supply chain, and the Dona Ana County in New Mexico for the guar supply chain. For full access to the model and input parameters, refer to [1].

2.1 Model Structure

To model the guayule and guar supply chains for the American Southwest, several economic, environmental, and social inputs are required as well as geographical information as shown in Fig. 2.1. To assess the economic impact, the inputs considered in the analysis include capital construction, harvesting, hauling, loading, unloading, transportation, processing, and distribution costs. Similarly, the environmental impact is evaluated through the equivalent greenhouse gas emissions of construction, transportation, processing, and distribution. As per the social impact, it is determined based on the local accrued jobs and the earnings per work hour.

Based on the number of farms, their acreage, the sales as defined by the Census of Agriculture 2017 [13], and discussions with industry experts, the farm groups cultivating cotton, grains, and oilseeds (CGO) crops are identified as the potential candidates to adopt guar and guayule [8]. The information of the farm groups such as the specific type of crop, the geographical location, and the acreage is obtained from USDA CropScape [14]. Candidates for processing facilities locations are identified using GIS layers for transportation, development, water resources, and the farm group's information. Through discussion with SBAR researchers, a 15% of adoption rate from the overall crops is defined as a baseline for the stochastic scenarios. The optimal routes are determined using the origin-destination cost matrix function of ArcGIS Pro 2.4.0.

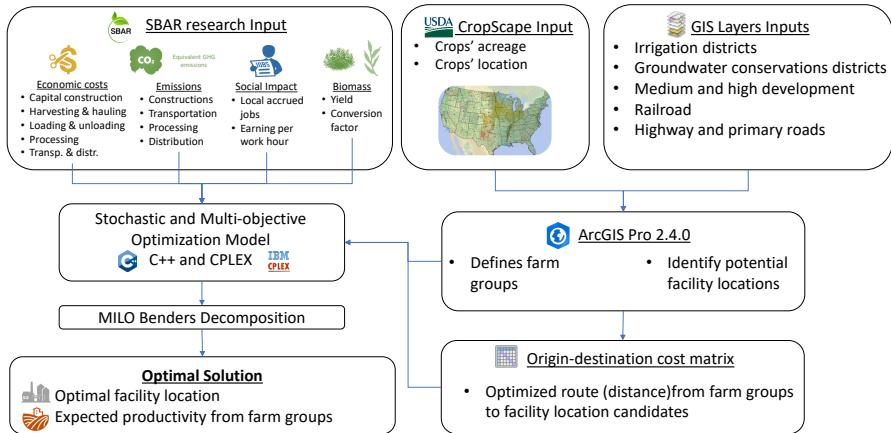


Figure 2.1: Biomass supply chain optimal design: model methodology

2.2 GIS Data and Geoprocessing Tools

2.2.1 Importing Data

The crops and farms information is obtained from USDA CropScape [14] as shown in Fig. 2.2. To download the information, select the polygon tool indicated in Step 1 and select the area of interest as marked in Step 2. In this example, New Mexico State is selected as the area of interest. Afterward, click on the download button indicated in Step 3 and a pop-up window will appear where you can select the years for which the information is desired (Step 4). The information will be downloaded in a raster format file.

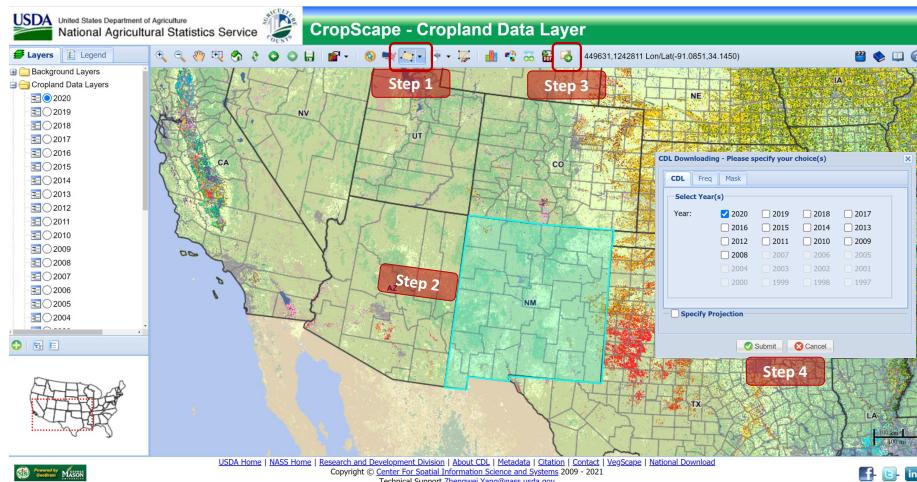


Figure 2.2: CropScape information

The next step is to import the raster file into ArcGIS Pro. For such purpose, please refer to Fig. 2.3. As shown in Fig. 2.3 a), to add the raster file, go to the Map menu, click on the Add Data drop down menu, and select the Data tool. You will need to provide the location of the saved raster file from CropScape. Once the raster file is uploaded, it will be overlaid on the map as presented in Fig. 2.3 b). From the Appearance menu (1), select the Symbology tool and the menu will appear on the right-hand side (2). The Symbology tool provides the crops' name and legend information. However, the raster file needs to be transformed to a fully editable version before proceeding with the analysis, in this case, the selected format is a polygon layer. The transformation process is presented in Fig. 2.3 c). First, select the Analysis menu (1), search for the Raster to Polygon tool (2), and complete the required information. Make sure to check the "Create Multipart Features" box to maintain the crop classification during the transformation. Finally select Run (3) to obtain the polygon layer. Fig. 2.3 d) shows an example where once the polygon layer was created, all crops were removed except CGO. Note how the Symbology changes, and

the classification is defined by gridcode (1) and the detailed information each polygon is also provided based on the gridcode (2).

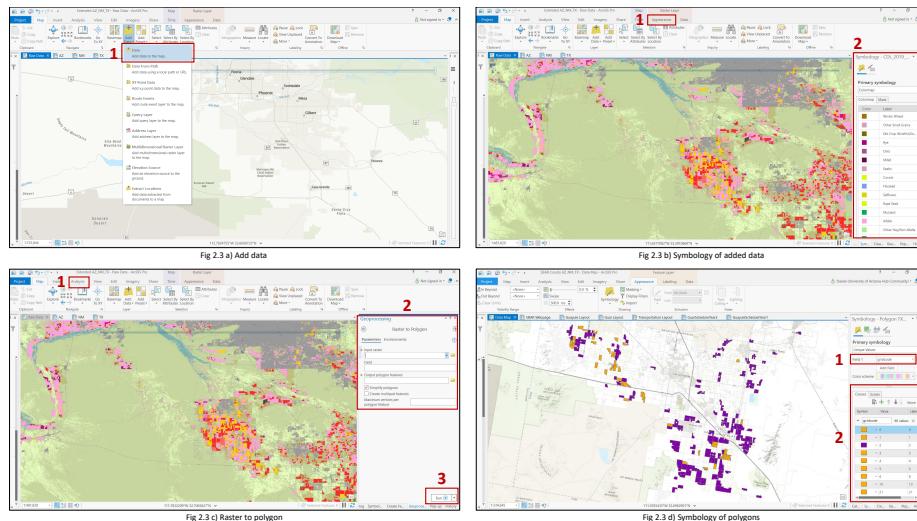


Figure 2.3: GIS data import

2.2.2 Customization of GIS Data

Once the polygon layer has been created and only the polygons from the farms of interest remain, the next step is to customize the layer as shown in Fig. 2.4. As shown in Fig. 2.4 a), more than one color range can be created for the polygon layer. For such purpose, right-click on the Map, select the Attribute Table, and it will appear on the bottom part of the screen (1). To add new attributes, select the button Add (2) and specify the characteristics for each attribute. In this example, three different color attributes are created (3) to ease switching between the color ranges for each polygon (farm). The specific color range is defined in the Symbology pane (4). Similarly, additional attributes need to be added such as the acreage and the geographic location of the polygon (farm) location as shown in Fig. 2.4 b). In this example, the attributes AreaAcre, X_Longitude, and Y_Latitude were created (1).

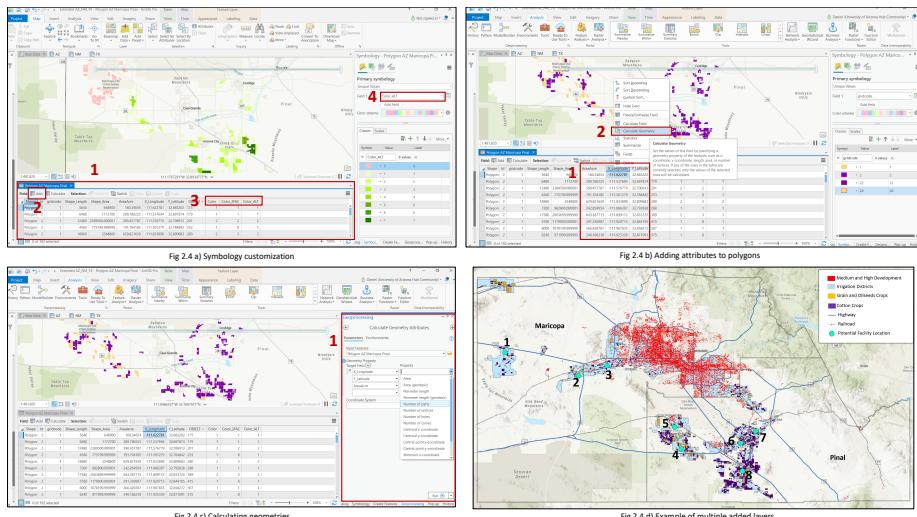


Figure 2.4: GIS data customization

The attributes will be created empty and the tool Calculate Geometry (2) needs to be selected to calculate each one. Fig. 2.4 c) shows the geoprocessing tool of Calculate Geometry Attributes (1), where you just need to assign the attributes of interest in the target fields and specify what property will be calculated for each attribute. Fig. 2.4 d) provides an example where additional layers have been added including main roads, development areas, and irrigation districts. Additional layers can be directly imported following the steps in Fig. 2.3 a).

2.2.3 Distance Matrix and Printing Formats

After customizing the polygon layer, the next information to calculate is the distance matrix (Fig. 2.5), i.e., the distance between origins and locations, or in this case, farms and candidate processing facilities. Fig. 2.5 a) shows how to access the Origin-Destination Cost Matrix through the Analysis menu. Once selected, a new and empty Cost Matrix layer will be created. To add the origins and destinations, the tools indicated in Fig. 2.5 b), i.e., Import Origins and Import Destinations (1) can be used. After imported, they will appear in the content pane under the created Cost Matrix layer (2). A transportation layer can be selected to determine the distances or use the ESRI extensive transportation information (e.g., streets, major and minor roads, and highways) by default. After importing the origins and destinations, click the Run button (1).

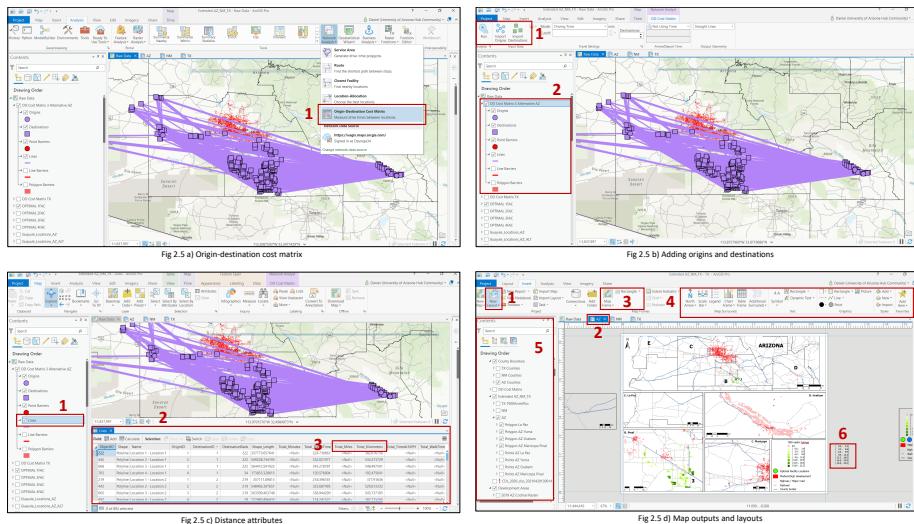


Figure 2.5: GIS distance matrix and map output

Results will appear in form of straight lines as presented in Fig. 2.5 c) from each origin to every destination. To see detailed information on the results, right-click the line checkbox (1) and open its Attribute table (2). The Attribute table will provide the distance in kilometers or miles. Note that the straight lines only indicate relationship and the calculated distance is bases on the transportation layer, i.e., the roads and highways. Finally, to create print the results, a map layout can be created as shown in Fig. 2.5 d). To do so, select the Inset Menu and click on the New Layout (1). Once a layout has been selected, you can navigate back to the previous map using the tabs indicated in (2). Multiple frames can be added to a single map layout to show multiple maps together (3). The example provided includes 5 different map frames. Additional information such as legends, scale bars, and north arrows can be automatically created (4)

if desired. If a layer is not being displayed, make sure its checkbox is marked in the Content Pane (5). Backup information or alternative map frames can be stored outside of the map layer (6) for ease access in future analysis.

2.3 Defined Parameters

The sets and indices are presented in Table 2.1 and Table 2.2 defines the required parameters for planning, harvesting, production, transportation, and distribution, as well as for the environmental and social impacts. The abbreviations of the units are defined as follows: acre (ac), CO₂ equivalent (CO₂ eq), cubic meter (m³), hour (hr), kilometer (km), U.S. dollar (\$), million U.S. dollar (M\$), metric tonne (t), and thousand metric tonne (Kt).

Table 2.1: Biomass Supply Chain Optimal Design: Sets and Indices

\mathcal{I}	Set of irrigation districts, indexed by i .
\mathcal{F}	Set of farms, indexed by f .
\mathcal{F}^i	Set of farms in water irrigation district i .
\mathcal{B}	Set of processing facilities, indexed by b .
\mathcal{P}	Set of bio-products, indexed by p .
\mathcal{C}	Set of customers, indexed by c .
Ω	Set of adoption rate scenarios, indexed by ω .

Table 2.2: Biomass Supply Chain Optimal Design: Input Parameters

$\theta^1, \theta^2, \theta^3$	Weighted factor for the economic, environmental, and social objectives, respectively.
$Prob(\omega)$	Probability of adoption rate scenario ω .
C_b^{cap}	Capital construction costs for processing facility b (\$).
K	Number of planning processing facilities.
Max_b	Minimum annual capacity of processing facility b (t).
Min_b	Maximum annual capacity of processing facility b (t).
C^H	Crop's harvesting and hauling cost (\$/t).
C^L	Crops's loading and unloading cost (\$/t).
Y_f	Yield for crop per acre of farm f (t/ac).
$A_f^{(\omega)}$	Area of supply of farm f for adoption rate ω (ac).
$P_f^{(\omega)}$	Crop productivity of farm f for adoption rate ω , e.g., $P_f^{(\omega)} = Y_f \times A_f^{(\omega)}$ (t).
W_i	Water availability in irrigation district i (m^3).
U_f	Irrigation water use by farm f (m^3/t).
C^P	Processing cost of biomass (\$/t).
θ_p	Conversion factor for bio-product p .
D_{pc}	Demand of bio-product p for customer c (t).
C^T	Transportation cost for biomass (\$/ $t \times km$).
C_p^D	Distribution cost for bio-product p (\$/t).
L_{fb}	Distance from farm f to processing facility b (km).
μ_b	Construction emission factor for processing facility b (CO_2 eq).
β	Transportation emission factor (CO_2 eq/ $t \times km$).
γ_p	Distribution emission factor for bio-product p (CO_2 eq/ $t \times km$).
δ_p	Processing emission factor for bio-product p (CO_2 eq/t).
ε	Emission price (\$/ CO_2 eq).
UB^{EnvImp}	Emissions impact upper-bound (\$).
$E_b^B, E_f^H, E_f^T, E_b^P, E_b^D$	Earning per work hour for construction of processing facility b harvesting farm f , transportation from farm f , production of bio-product from processing facility b , and distribution of bio-product from processing facility b , respectively (\$/hr).
H^T, H_b^D	Number of work hours required for transportation and distribution of bio-product from processing facility b , respectively ($hr/t \times km$).
H^H, H_b^P	Number of work hours required for harvesting and production of bio-product from processing facility b , respectively (hr/t).
H^B	Number of work hours required for processing facility construction (hr).
LB^{SocImp}	Social impact lower-bound (\$).

2.3.1 Identifying processing facility candidates

To identify the processing facility location candidates, a GIS analysis can be performed with information from irrigation districts, CGO crops, medium/high development, highway, and railroad layers which are overlapped to identify potential facility locations and farms. The GIS layers such as the development levels, county boundaries, highways, railroads, and irrigation districts are collected from online databases. Distances between farms, processing facility location candidates, and customers are calculated using ArcGIS Pro as shown in Fig. 2.5. The criteria used to identify the processing facility location candidates are as follows:

- (a) Accessibility to biomass supplies, close to feasible planting fields.
- (b) Within a 2-mile straight line distance to major roads.
- (c) Within a 2-mile straight line distance to the railway.
- (d) Within a 2-mile straight line distance of the pipeline/major water bodies.
- (e) Avoid the developed areas (high intensity and medium intensity areas).

2.3.2 CPLEX Parameter File Creation

A Microsoft Excel Macro-Enabled file is provided to create the input parameter file required for the CPLEX optimization software (refer to the file MultiObjective_Data_Original.xlsxm) Fig. 2.6 to Fig. 2.8 will guide you through the four steps in the xlsm file. Fig. 2.6 highlights the instruction for the 4 steps, where step 1 is just to provide the desired name for the output TXT file, which will be used as an input for CPLEX. Step 1 additionally requires to fill the cardinality of the sets defined in Table 2.1, e.g., the number of farms, number of irrigation districts, and number of candidate processing facilities. Once all the information in Step 1 has been completed, Step 2 just requires to click the Restructure Table button and the spreadsheet will restructure based on the provided cardinality of the sets as shown in Fig. 2.7.

Integrating environmental and social impacts into optimal design of guayule and guar supply chains

Input Data Workbook

Instructions:

1. Fill in the file name and the cardinality of the sets, i.e., the number of each type of elements → **TXT File Name**
2. Click the restructure table button → **Restructure Table**
3. Fill in the restructured table (all cells with light green color)
4. Click save txt file button → **Save .txt File**

Note: The .txt file is saved in the workbook folder and overwrites it if the file already exists

DO NOT CHANGE the cells' colors

Data Reference

Data Reference	Description	Units	Model Reference
Auxiliary parameter for sensitivity analysis	(No.)	Ver	1
Number of farms to be considered	(No.)	F	
Number of irrigation districts	(No.)	I	
Number of biomass processing facility candidates	(No.)	B	
Number of bio-products	(No.)	P	
Number of customers	(No.)	C	
Number of stochastic scenarios	(No.)	OMEGA	

Figure 2.6: Input parameter walkthrough: Step 1

Once all of the information of the restructured table has been completed, i.e., every cell in green color as shown in Fig. 2.8, Step 4 is to click the Save .txt File button and a text file with the assigned name will be created within the same folder of the .xlsm file.

Note: **Do not use the original .xlsm file.** Since the file uses macros for the restructuring, changes cannot be easily undone. It is better to create a copy of the .xlsm file to work in and leave the original as a backup. Additionally, **the color of the cells must not be modified** since the macros are color coded. If an update in a parameter is required (except for the sets cardinalities) the values can be changed directly and the file saved again with the Save .txt File button to update the text file.

2.3. DEFINED PARAMETERS

22

Integrating environmental and social impacts into optimal design of guayule and guar supply chains

Input Data Workbook

Instructions:

1. Fill in the file name and the cardinality of the sets, i.e., the number of each type of elements → **TXt File Name** → **MultiData**
2. Click the restructure table button → **Restructure Table**
3. Fill in the restructured table (all cells with light green color)
4. Click save txt file button → **Save .txt File**

Note: The .txt file is saved in the workbook folder and overwrites it if the file already exists

DO NOT CHANGE the cells colors

Data Reference	Description	Units	Model Reference							
Auxiliary parameters for sensitivity analysis	(No.)	Ver	1							
Number of farms to be considered	(No.)	F	160							
Number of irrigation districts	(No.)	I	18							
Number of biomass processing facility candidates	(No.)	B	8							
Number of bio-products	(No.)	P	2							
Number of customers	(No.)	C	6							
Number of stochastic scenarios	(No.)	OMEGA	10							
Probability of adoption rate for stochastic scenario w	(0,1)	Prob	1 2 3 4 5 6 7 8 9 10							
Weighted factor for the economic objective	(0,1)	theta_1								
Weighted factor for the environmental objective	(0,1)	theta_2								
Weighted factor for the social objective	(0,1)	theta_3								
Crop's harvesting and hauling cost	(\$/metric tonne)	CH								
Crop's loading and unloading cost	(\$/metric tonne)	CL								
Transportation cost for biomass (\$/metric tonne x km)		CT								
Number of planning processing facilities	(No.)	K								
Irrigation district per farm (enumerative list form)	(No.)	F_I								
Yield for crop per acre of farm (metric tonne / acre)		Y_I								
Irrigation water use by each farm (m³/metric tonne)		U_I								
Farm area to adopt per stochastic scenario	(acres)	Aw_I								
Farm No.		F_I Y_I U_I Aw_I	1 2 3 4 5 6 7 8 9 10							
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										

Figure 2.7: Input parameter walkthrough: Step 2

Integrating environmental and social impacts into optimal design of guayule and guar supply chains

Input Data Workbook

Instructions:

1. Fill in the file name and the cardinality of the sets, i.e., the number of each type of elements → **TXt File Name** → **MultiData**
2. Click the restructure table button → **Restructure Table**
3. Fill in the restructured table (all cells with light green color)
4. Click save txt file button → **Save .txt File**

Note: The .txt file is saved in the workbook folder and overwrites it if the file already exists

DO NOT CHANGE the cells colors

Data Reference	Description	Units	Model Reference
Auxiliary parameters for sensitivity analysis	(No.)	Ver	1
Number of farms to be considered	(No.)	F	160
Number of irrigation districts	(No.)	I	18
Number of biomass processing facility candidates	(No.)	B	8
Number of bio-products	(No.)	P	2
Number of customers	(No.)	C	6
Number of stochastic scenarios	(No.)	OMEGA	10
Probability of adoption rate for stochastic scenario w	(0,1)	Prob	1 2 3 4 5 6 7 8 9 10
Weighted factor for the economic objective	(0,1)	theta_1	0.15 0.33 0.27 0.08 0.07 0.01 0.03 0.04 0.01 0.00
Weighted factor for the environmental objective	(0,1)	theta_2	0.33
Weighted factor for the social objective	(0,1)	theta_3	0.33
Crop's harvesting and hauling cost	(\$/metric tonne)	CH	29.7
Crop's loading and unloading cost	(\$/metric tonne)	CL	5.7
Transportation cost for biomass (\$/metric tonne x km)		CT	0.12
Number of planning processing facilities	(No.)	K	1
Irrigation district per farm (enumerative list form)	(No.)	F_I	
Yield for crop per acre of farm (metric tonne / acre)		Y_I	
Irrigation water use by each farm (m³/metric tonne)		U_I	
Farm area to adopt per stochastic scenario	(acres)	Aw_I	
Farm No.		F_I Y_I U_I Aw_I	1 2 3 4 5 6 7 8 9 10
1	13	12.140508 406.39	244 204 345 395 445 496 546 596 647 697
2	13	12.140508 406.39	592 714 836 958 1080 1202 1324 1446 1568 1690
3	13	12.140508 406.39	150 170 190 210 230 250 270 290 310 330
4	13	12.140508 406.39	278 335 392 449 507 564 621 678 736 793
5	13	12.140508 406.39	148 179 209 240 270 301 331 362 392 423
6	13	12.140508 406.39	700 845 989 1134 1278 1423 1567 1712 1856 2001
7	13	12.140508 406.39	537 705 900 1097 1295 1494 1692 1890 2088 2286
8	10	12.140508 406.39	218 261 308 352 397 442 487 532 577 622
9	10	12.140508 406.39	857 1034 1210 1387 1564 1741 1918 2094 2271 2448
10	10	12.140508 406.39	486 540 600 657 714 771 828 885 942 1000
11	10	12.140508 406.39	233 250 330 378 426 474 522 571 619 667
12	10	12.140508 406.39	59 71 84 96 109 120 132 145 157 169
13	10	12.140508 406.39	472 570 667 764 862 950 1057 1154 1252 1349
14	9	12.140508 406.39	73 96 113 130 148 165 183 201 218 237
15	9	12.140508 406.39	332 401 469 538 606 675 743 812 880 940
16	14	12.140508 406.39	109 131 154 176 199 221 244 266 289 311
17	14	12.140508 406.39	366 442 517 593 668 744 819 895 970 1046
18	4	12.140508 406.39	223 240 257 300 318 334 351 369 386 405
19	4	12.140508 406.39	163 197 231 265 296 332 366 400 433 467
20	4	12.140508 406.39	207 250 293 335 378 421 464 506 549 592
21	4	12.140508 406.39	198 239 279 320 361 402 443 483 524 569

Figure 2.8: Input parameter walkthrough: Steps 3 and 4

2.4 Optimization Tools and Methods

2.4.1 CPLEX Optimizer Through a High Performance Computer

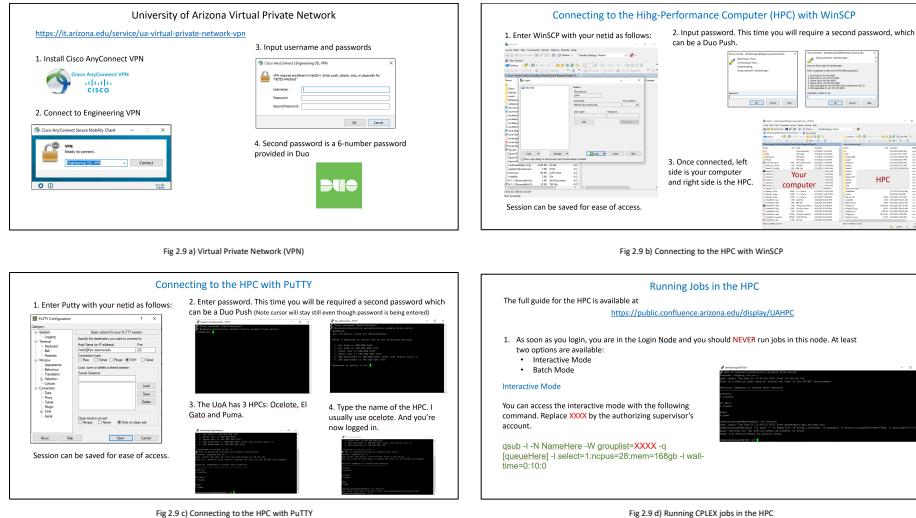


Figure 2.9: Running CPLEX optimizer in the HPC (Part 1)

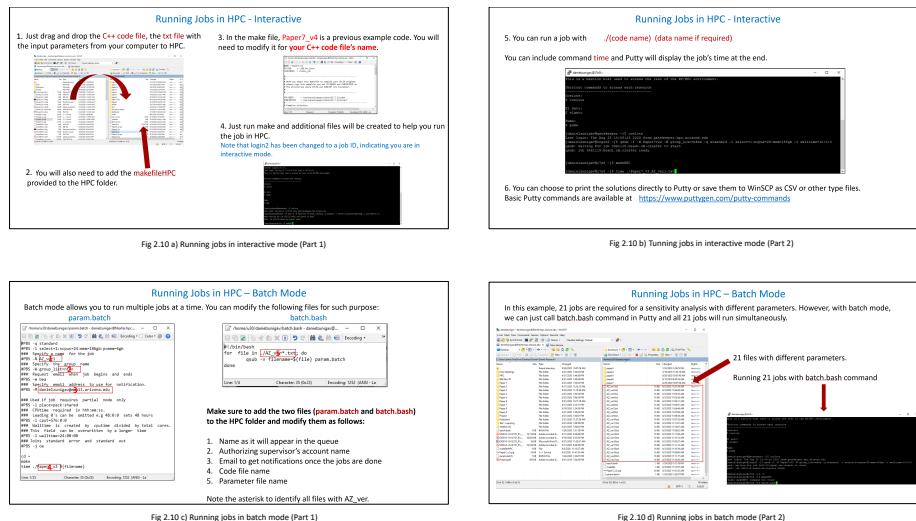


Figure 2.10: Running CPLEX optimizer in the HPC (Part 2)

2.4.2 Reporting Outputs

Table 2.3: Biomass Supply Chain Optimal Design: Decision Variables

x_b	Binary variable that is 1 if processing facility b is added and 0 otherwise (\mathbf{x} is the vector form).
cap_b	Capacity of processing facility b (\mathbf{cap} is the vector form).
$u_{fb}^{(\omega)}$	Commodity flow of biomass from farm f to processing facility b (\mathbf{u} is the vector form) (metric tonnes).
$v_{pbc}^{(\omega)}$	Commodity flow of bio-product p from processing facility b to customer c (\mathbf{v} is the vector form) (metric tonnes).

2.5 Model Assumptions

Model assumptions includes that certain information about the probability distribution for the stochastic scenarios for the adoption rate is known beforehand. Additionally, a single crop is considered per farm or farm group and capital costs do not consider interest rates.

2.6 Extended Results for the American Southwest

This research an extension of Sections 2.1-2.5 to the American Southwest. To identify potential farm groups for the adoption of guayule and guar, all of the farms groups within this region are evaluated based on their type of crop, location, water availability, and acreage. A 15% base of the overall crops is defined for the stochastic scenarios with log-normal distribution probabilities. This research explores further the contributions from the model of Chapter 2 by considering multiple processing facility location candidates. A sensitivity analysis is performed for multiple optimal facility locations based on the economic, environmental, and social impacts of the guayule and guar supply chains. Three case studies are presented as follows: i) Arizona for guayule adoption (Fig 2.11), ii) New Mexico for the adoption of guar (Fig 2.12), and iii) Texas for guar adoption (Fig 2.13).

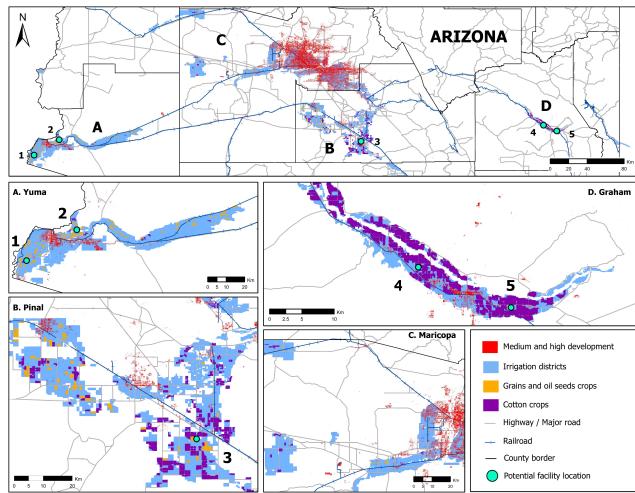


Figure 2.11: Arizona's GIS information

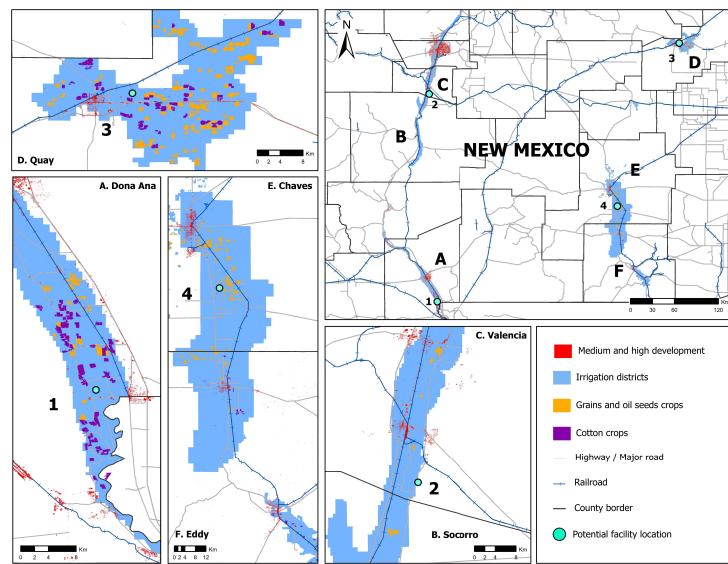


Figure 2.12: New Mexico's GIS information

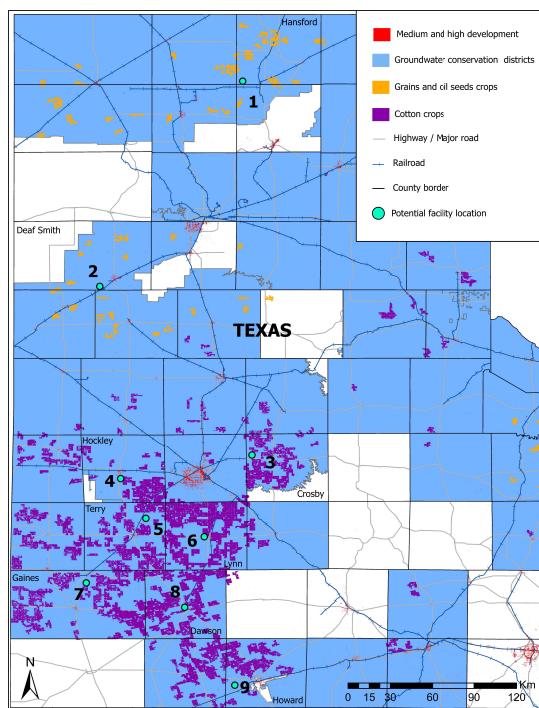


Figure 2.13: Northwestern Texas' GIS information

Chapter 3

Optimal Production Planning and Machinery Scheduling for Semi-Arid Farms

This research proposes an integer linear optimization (ILO) model to identify optimal decisions for production planning and machinery scheduling of perennial and non-perennial crops that maximize the farmers' NPV. Key differences between this paper and previous work include the integration of both, multi-crop planning and multi-machinery scheduling into a single model while enabling crop rotation within a defined set of time periods. In contrast with the available literature, this paper differentiates multi-crop production between perennial and non-perennial crops throughout a planning horizon (e.g. a five or ten years planning). Additionally, the optimal multi-machinery scheduling is identified for each time period based on the perennial or non-perennial crop requirements. The revenue is determined based on the sales of each crop and the total costs include the machinery loan amortization, operational costs (labor, repairs, maintenance, fuel, and lube), machinery transportation costs, fertilizers, herbicides, insecticides, seeds, and irrigation water costs. For scheduling purposes, the optimal routes between the parcels are determined based on ArcGIS Pro tools. Irrigation water requirements and availability are analyzed through GIS irrigation district layers. The resulting model yields a complex large-scale ILO problem. To assess the model, two case studies for low-water-use crops are developed, one for the guar crop considering parcels located in Cochran and Hockley counties, Texas, and one for the guayule crop considering parcels located in Pinal County, Arizona. For full access to the model and input parameters, refer to [2].

3.1 Model Structure

One of the goals of production planning is to identify the crops to be planted in each parcel to achieve the maximum economic benefit. Likewise, one of the goals of machinery planting scheduling is to determine the sequence of parcels to work that minimizes the machinery transportation costs. The structure of the proposed model is presented in Fig. 3.1 and Fig. 3.2. Here, the optimized production planning is performed for c crops in p parcels, and the optimized machinery scheduling is determined for m machinery sets of each crop based on p parcels.

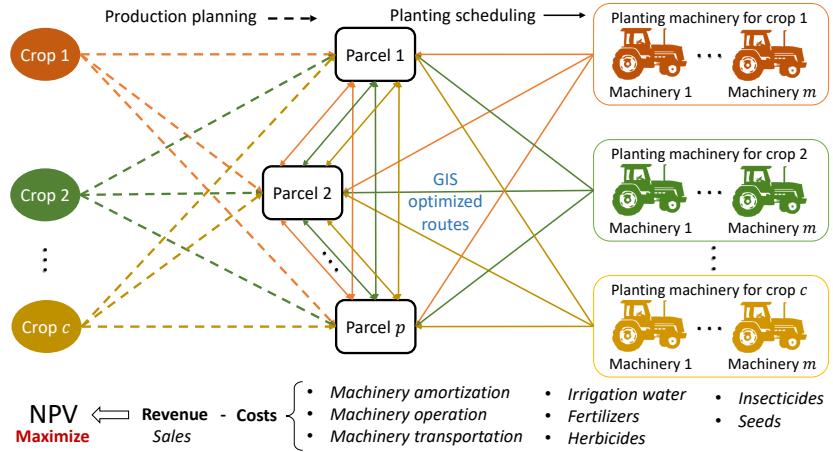


Figure 3.1: Production planning and multi-machinery scheduling: model structure (part 1)

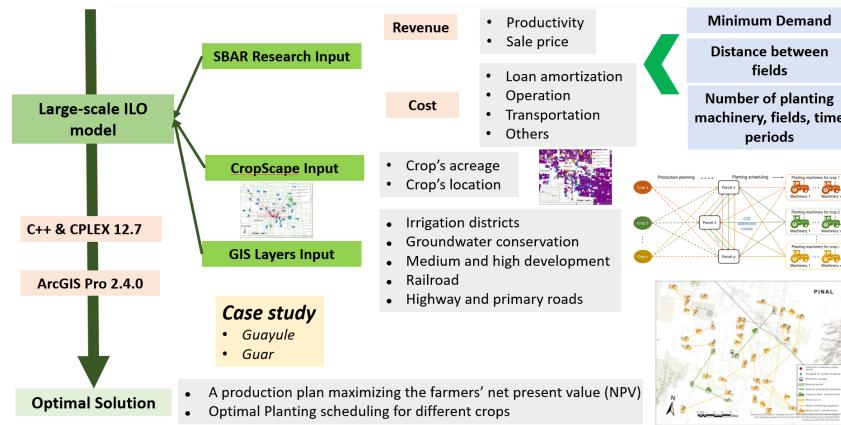


Figure 3.2: Production planning and multi-machinery scheduling: model structure (part 2)

3.2 GIS Data and Geoprocessing Tools

Please refer to [2.2](#) for detailed instructions on how to import, process, an analyze the required GIS data.

3.3 Defined Parameters

The sets and indices are presented in Table 3.1 and Table 3.2 defines economic, planting, and water parameters required for the model. The abbreviations of the units are defined as follows: acre (ac), CO₂ equivalent (CO₂ eq), cubic meter (m³), hour (hr), kilometer (km), U.S. dollar (\$), million U.S. dollar (M\$), metric tonne (t), and thousand metric tonne (Kt).

Table 3.1: Optimal Production Planning and Scheduling: Sets and Indices

\mathcal{C} :	Set of crops, indexed by c .
\mathcal{I} :	Set of irrigation districts, indexed by i .
\mathcal{M} :	Set of planting machinery, indexed by m , $M = \mathcal{M} $.
\mathcal{P} :	Set of parcels, indexed by p and q , $P = \mathcal{P} $.
\mathcal{R} :	Set of route distances between parcels, indexed by r .
\mathcal{P}_i :	Set of parcels in irrigation district i .
\mathcal{T} :	Set of time periods, indexed by t , $T = \mathcal{T} $.

Table 3.2: Optimal Production Planning and Scheduling: Input Parameters

$C_c^{A,t}$:	Loan amortization of machinery required for planting crop c for time t (\$).
C_c^O :	Machinery operational cost (including labor) during planting of crop c (\$/acre).
C_c^{fhi} :	Fertilizer, herbicide, and insecticide cost of crop c , e.g. N ₂ , P, prowl (\$/acre).
C_c^T :	Transportation cost of planting machinery for crop c (\$/mile).
C^I :	Irrigation water cost (\$/m ³).
C^{IL} :	Irrigation labor cost (\$/acre).
C_c^S :	Seed cost of crop c (\$/acre).
D_c^t :	Minimum demand of crop c for time period t (t).
S_c :	Sale price of crop c (\$/t).
IR^{Sale} :	Sale price inflation rate (unitless).
IR^{Cost} :	Costs inflation rate (unitless).
α :	Discount rate (unitless).
T_c :	Time periods required for crop c (i.e. length of harvest season).
N_c^R :	Maximum number of sequential plantings (harvest length season) of crop c before crop rotation.
Y_{cp} :	Yield for crop c per acre of parcel p (t/acre).
A_{cp} :	Acres for crop c of parcel p (acre).
P_{cp} :	Productivity for crop c of parcel p , e.g. $P_{cp} = Y_{cp} \times A_p$ (t).
$D(p, q)$:	Distance d between (p,q) parcels (miles).
M^{Cap} :	Capacity of planting machinery per length of harvest season (acres).
$MaxP$:	Maximum number of parcels the planting machinery can work on.
W_c^t :	Irrigation water use per acre for crop c in time period t (m ³ /acre).
W_{cp}^F :	Maximum irrigation water capacity for crop of parcel p (m ³).
W_i^I :	Water availability in irrigation district i (m ³).

3.3.1 CPLEX Parameter File Creation

A Microsoft Excel Macro-Enabled file is provided to created the input parameter file required for the CPLEX optimization software (refer to the file Product-Plan-Data-Original.xlsxm) Fig. 3.3 to Fig. 3.5 will guide you through the four steps in the xlsm file. Fig. 3.3 highlights the instruction for the 4 steps, where step 1 is just to provide the desired name for the output TXT file, which will be used as an input for CPLEX. Step 1 additionally requires to fill the cardinality of the sets defined in Table 3.1, e.g., the number of crops, number of irrigation districts, and number of time periods. Once all the information in Step 1 has been completed, Step 2 just requires to click the Restructure Table button and the spreadsheet will restructure based on the provided cardinality of the sets as shown in Fig. 3.4.

Data Reference	Description	Units	Model Reference
	Auxiliary parameter for sensitivity analysis	(No.)	Ver
	Number of crops	(No.)	C
	Number of irrigation districts	(No.)	I
	Number of planting machinery	(No.)	M
	Number of parcels (or farms)	(No.)	P
	Number of time periods	(No.)	T

Figure 3.3: Input parameter walkthrough: Step 1

Once all of the information of the restructured table has been completed, i.e., every cell in green color as shown in Fig. 3.5, Step 4 is to click the Save .txt File button and a text file with the assigned name will be created within the same folder of the .xlsm file.

Note: **Do not use the original .xlsm file.** Since the file uses macros for the restructuring, changes cannot be easily undone. It is better to create a copy of the .xlsm file to work in and leave the original as a backup. Additionally, **the color of the cells must not be modified** since the macros are color coded. If an update in a parameter is required (except for the sets cardinalities) he values can be changed directly and the filed saved again with the Save .txt File button to update the text file.

3.3. DEFINED PARAMETERS

32

Optimal Production Planning and Machinery Scheduling for Semi-Arid Farms									
Input Data Workbook									
Instructions:									
1. Fill in the file name and the cardinality of the sets, i.e., the number of each type of elements								TXT File Name ProductionData	
2. Click the restructure table button  Restructure Table									
3. Fill in the restructured table (all cells with light green color)									
4. Click save txt file button  Save .txt File									
Note: The .txt file is saved in the workbook folder and overwrites it if the file already exists									
DO NOT CHANGE the cell's colors									
Data Reference	Description	Units	Model Reference						
	Auxiliary parameter for sensitivity analysis Number of crops Number of irrigation districts Number of planting machinery Number of parcels (or farms) without Dummy Parcel Number of time periods	(No.)	Ver C I M P T	1 2 1 7 37 10					
	Minimum demand of crop c for time period t	(metric tonnes)	Dt_c						
	Loan amortization of machinery required for planting crop c for time t	(\\$)	CAT_c						
	Machinery operational cost (including labor) during planting of crop	(\\$ / acre)	CO_c						
	Fertilizer, herbicide, and insecticide cost of crop c, e.g. N2, P, prowl	(\\$ / acre)	Chfi_c						
	Transportation cost of planting machinery for crop c	(\\$ / mile)	CT_c						
	Irrigation water cost Irrigation labor cost	(\\$ / m³) (\\$ / acre)	CI CIL						
	Seed cost of crop c	(\\$ / acre)	CS_c						
	Sale price of crop c	(\\$ / metric tonne)	S_c						
	Sale price inflation rate Costs inflation rate Discount rate	(unitless)	IRSale IRCost α						

Figure 3.4: Input parameter walkthrough: Step 2

Optimal Production Planning and Machinery Scheduling for Semi-Arid Farms									
Input Data Workbook									
Instructions:									
1. Fill in the file name and the cardinality of the sets, i.e., the number of each type of elements								TXT File Name ProductionData	
2. Click the restructure table button  Restructure Table									
3. Fill in the restructured table (all cells with light green color)									
4. Click save txt file button  Save .txt File									
Note: The .txt file is saved in the workbook folder and overwrites it if the file already exists									
DO NOT CHANGE the cell's colors									
Data Reference	Description	Units	Model Reference						
	Auxiliary parameter for sensitivity analysis Number of crops Number of irrigation districts Number of planting machinery Number of parcels (or farms) without Dummy Parcel Number of time periods	(No.)	Ver C I M P T	1 2 1 7 37 10					
	Minimum demand of crop c for time period t	(metric tonnes)	Dt_c						
	Loan amortization of machinery required for planting crop c for time t	(\\$)	CAT_c						
	Machinery operational cost (including labor) during planting of crop	(\\$ / acre)	CO_c						
	Fertilizer, herbicide, and insecticide cost of crop c, e.g. N2, P, prowl	(\\$ / acre)	Chfi_c						
	Transportation cost of planting machinery for crop c	(\\$ / mile)	CT_c						
	Irrigation water cost Irrigation labor cost	(\\$ / m³) (\\$ / acre)	CI CIL						
	Seed cost of crop c	(\\$ / acre)	CS_c						
	Sale price of crop c	(\\$ / metric tonne)	S_c						
	Sale price inflation rate Costs inflation rate Discount rate	(unitless)	IRSale IRCost α						

Figure 3.5: Input parameter walkthrough: Steps 3 and 4

3.4 Optimization Tools and Methods

Please refer to [2.4](#) for detailed instructions on how to access the CPLEX optimization solver through the HPC.

3.4.1 Reporting Outputs

Table 3.3: Optimal Production Planning and Scheduling: Decision Variables

u_{mc} :	Binary variable that is 1 if planting machinery m for crop c is added and 0 otherwise (\mathbf{u} is the vector form).
w_p :	Integer variables that denote the parcel p position on the planting machinery tour, (\mathbf{w} is the vector form).
x_{cp}^t :	Binary variable that is 1 if crop c is planted in parcel p for time period t and 0 otherwise (\mathbf{x} is the vector form).
$y_{(p,q)c}^t$:	Binary variable that is 1 if planting machinery has been assigned for crop c from parcel p to parcel q in time period t and 0 otherwise (\mathbf{y} is the vector form).

3.5 Model Assumptions

Model assumptions are specified as follows: i) the length of harvest season must be equal or greater than two time periods, $T_c \geq 2$. For example, if the length of the harvest season of a crop is one year, it needs to be modeled as two semesters; ii) the model assumes a single machinery storage location, i.e., all the planting machinery must depart from the same storage location and return to it after its work is finished; and iii) if the acreage of a parcel exceeds the planting capacity of the machinery M^{Cap} , the parcel needs to be divided into multiple parcels to ensure the planting machinery capacity is greater.

Chapter 4

Demand-Driven Harvest Planning and Machinery Scheduling

This research proposes a large-scale mixed-integer linear optimization (MILO) model to identify optimal decisions for harvest planning and machinery scheduling that minimize the total costs. This work is an extension of previous research, considering many factors such as periodic harvest, different demands, time-window qualification, multi-machinery scheduling, resource limitations and late penalties to obtain an optimal long-term harvest plan. The goal of this optimization model is to address three significant problems when making a harvest plan: When to harvest? How to harvest? How much to harvest? To address these problems, the long-term harvest plan is decomposed into several harvest seasons. And in each harvest season, time-window variables are used to indicate the best harvest time and decision-making variables are also introduced to represent the optimal harvest scheduling. The objective function of this MILO model is defined as the total cost (including rent cost, operation cost, transportation cost and penalty cost), and model constraints consist of harvesting requirements, resource demands, harvesting capacity, time-window and transportation constraints. When solving the optimization model, we obtain a 10-year harvest plan for the crops as well as the optimal harvest scheduling during each harvest season. To validate the model, 37 fields located in Pinal County, Arizona are developed for the numerical experiments. For full access to the model and input parameters, please refer to [3].

4.1 Model Structure

To model the harvest planning and machinery scheduling, several inputs and geographical information are required, as shown in Fig. 4.1. The research input including various costs, diverse characteristics of crops, different harvesting demands is incorporated in this MILO model. Most of the data comes from previous research [1, 2] and some parameters are determined by researchers and industry experts from SBAR project [8]. The proposed optimization model is implemented in Python 3.9.6 using Gurobi 9.5.0 with default settings. And ArcGIS Pro 2.9.2 is used to present the GIS information and results of optimal machinery scheduling.

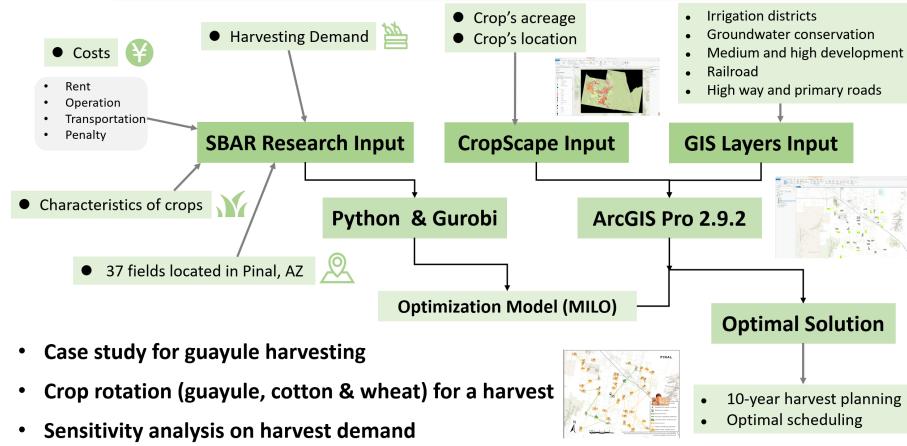


Figure 4.1: Harvest planning and machinery scheduling: model structure

4.2 GIS Data and Geoprocessing Tools

Please refer to [2.2](#) for detailed instructions on how to import, process, an analyze the required GIS data.

4.3 Defined Parameters

The sets and indices are presented in Table 4.1 and Table 4.2 defines several parameters required for the model.

Table 4.1: Harvest Planning and Machinery Scheduling: Sets and Indices

F :	Set of fields, indexed by f
H :	Set of harvesters, indexed by h
T :	Set of time periods, indexed by t (e.g. 1st week of harvesting)
K :	Set of harvest seasons, indexed by k (e.g. fall of the second year)

Table 4.2: Harvest Planning and Machinery Scheduling: Input Parameters

C_h :	Ownership/Rent cost of harvester h (\$/harvest season)
C^o :	Unit machinery operational cost (including labor, repairs, maintenance, fuel, and lube) during harvesting one field (\$/acre)
C^r :	Unit route/transportation cost (\$/mile)
$d_{(i,j)}$:	Distance between field i and j (mile)
C_f^p :	Penalty cost incurred for not harvesting field f on time (\$/(acre*period))
R_f^k :	Ready date of field f for harvest season k
D_f^k :	Due date of field f for harvest season k
A_f :	Acres for field f (acre)
M_f :	Maximum number of machines that can harvest field f in a given time period
M^{Cap} :	Harvesting capacity of one harvester per length of harvest season (acre)
I_f^k :	Binary values indicate whether field f should be harvested in the harvest season k

4.3.1 Input Parameter File Creation

A Microsoft Excel Macro-Enabled file is provided to created the input parameter file required for the Gurobi optimization software (refer to the file HarvestPlan_data_Original.xlsxm) Fig. 4.2 to Fig. 4.4 will guide you through the four steps in the xlsxm file. Fig. 4.2 highlights the instruction for the 4 steps, where step 1 is just to provide the desired name for the output TXT file, which will be used as an input for Gurobi. Step 1 additionally requires to fill the cardinality of the sets defined in Table 4.1, e.g., the number of fields, number of harvesters, number of harvest seasons, and number of time periods. Once all the information in Step 1 has been completed, Step 2 just requires to click the Restructure Table button and the spreadsheet will restructure based on the provided cardinality of the sets as shown in Fig. 4.3.

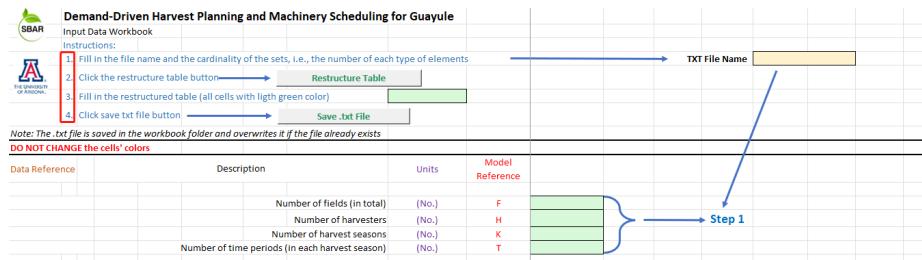


Figure 4.2: Input parameter walkthrough: Step 1

Once all of the information of the restructured table has been completed, i.e., every cell in green color as shown in Fig. 4.4, Step 4 is to click the Save .txt File button and a text file with the assigned name will be created within the same folder of the .xlsxm file.

Note: **Do not use the original .xlsxm file.** Since the file uses macros for the restructuring, changes cannot be easily undone. It is better to create a copy of the .xlsxm file to work in and leave the original as a backup. Additionally, **the color of the cells must not be modified** since the macros are color coded. If an update in a parameter is required, the values can be changed directly and the file saved again with the Save .txt File button to update the text file.

4.3. DEFINED PARAMETERS

40

**Demand-driven harvest planning and machinery scheduling for Genetic
Harvest Scheduling**

1. Fill in the file name and the cardinality of the sets, i.e., the number of each type of elements
2. Click the minimize title button
3. Click save net file button
4. Click save net file button
User: The .jar file is located in the working folder and executes it if the checkbox is checked
DO NOT CHECK THE FILE VALUES

Description	Unit	Mode Reference
Number of fields (in total)	(No.)	F
Number of harvest sessions	(No.)	K
Number of time periods (in each harvest session)	(No.)	T
Distance between fields (in km)	(km)	R
Acre of each field	(Acre)	A
Ready date of one field	(period)	RJ
Due date of one field	(period)	DJ
Indicate which fields needs to be harvested in one harvest session	(Acre)	I
Distance between 2 fields (in km)	(km)	RJ

Step 1

Step 2

Step 3

Step 4

Step 5

Step 6

Step 7

Step 8

Step 9

Step 10

Step 11

Step 12

Step 13

Step 14

Step 15

Step 16

Step 17

Step 18

Step 19

Step 20

Step 21

Step 22

Step 23

Step 24

Step 25

Step 26

Step 27

Step 28

Step 29

Step 30

Step 31

Step 32

Step 33

Step 34

Step 35

Step 36

Step 37

Step 38

Step 39

Step 40

Step 41

Step 42

Step 43

Step 44

Step 45

Step 46

Step 47

Step 48

Step 49

Step 50

Step 51

Step 52

Step 53

Step 54

Step 55

Step 56

Step 57

Step 58

Step 59

Step 60

Step 61

Step 62

Step 63

Step 64

Step 65

Step 66

Step 67

Step 68

Step 69

Step 70

Step 71

Step 72

Step 73

Step 74

Step 75

Step 76

Step 77

Step 78

Step 79

Step 80

Step 81

Step 82

Step 83

Step 84

Step 85

Step 86

Step 87

Step 88

Step 89

Step 90

Step 91

Step 92

Step 93

Step 94

Step 95

Step 96

Step 97

Step 98

Step 99

Step 100

Step 101

Step 102

Step 103

Step 104

Step 105

Step 106

Step 107

Step 108

Step 109

Step 110

Step 111

Step 112

Step 113

Step 114

Step 115

Step 116

Step 117

Step 118

Step 119

Step 120

Step 121

Step 122

Step 123

Step 124

Step 125

Step 126

Step 127

Step 128

Step 129

Step 130

Step 131

Step 132

Step 133

Step 134

Step 135

Step 136

Step 137

Step 138

Step 139

Step 140

Step 141

Step 142

Step 143

Step 144

Step 145

Step 146

Step 147

Step 148

Step 149

Step 150

Step 151

Step 152

Step 153

Step 154

Step 155

Step 156

Step 157

Step 158

Step 159

Step 160

Step 161

Step 162

Step 163

Step 164

Step 165

Step 166

Step 167

Step 168

Step 169

Step 170

Step 171

Step 172

Step 173

Step 174

Step 175

Step 176

Step 177

Step 178

Step 179

Step 180

Step 181

Step 182

Step 183

Step 184

Step 185

Step 186

Step 187

Step 188

Step 189

Step 190

Step 191

Step 192

Step 193

Step 194

Step 195

Step 196

Step 197

Step 198

Step 199

Step 200

Step 201

Step 202

Step 203

Step 204

Step 205

Step 206

Step 207

Step 208

Step 209

Step 210

Step 211

Step 212

Step 213

Step 214

Step 215

Step 216

Step 217

Step 218

Step 219

Step 220

Step 221

Step 222

Step 223

Step 224

Step 225

Step 226

Step 227

Step 228

Step 229

Step 230

Step 231

Step 232

Step 233

Step 234

Step 235

Step 236

Step 237

Step 238

Step 239

Step 240

Step 241

Step 242

Step 243

Step 244

Step 245

Step 246

Step 247

Step 248

Step 249

Step 250

Step 251

Step 252

Step 253

Step 254

Step 255

Step 256

Step 257

Step 258

Step 259

Step 260

Step 261

Step 262

Step 263

Step 264

Step 265

Step 266

Step 267

Step 268

Step 269

Step 270

Step 271

Step 272

Step 273

Step 274

Step 275

Step 276

Step 277

Step 278

Step 279

Step 280

Step 281

Step 282

Step 283

Step 284

Step 285

Step 286

Step 287

Step 288

Step 289

Step 290

Step 291

Step 292

Step 293

Step 294

Step 295

Step 296

Step 297

Step 298

Step 299

Step 300

Step 301

Step 302

Step 303

Step 304

Step 305

Step 306

Step 307

Step 308

Step 309

Step 310

Step 311

Step 312

Step 313

Step 314

Step 315

Step 316

Step 317

Step 318

Step 319

Step 320

Step 321

Step 322

Step 323

Step 324

Step 325

Step 326

Step 327

Step 328

Step 329

Step 330

Step 331

Step 332

Step 333

Step 334

Step 335

Step 336

Step 337

Step 338

Step 339

Step 340

Step 341

Step 342

Step 343

Step 344

Step 345

Step 346

Step 347

Step 348

Step 349

Step 350

Step 351

Step 352

Step 353

Step 354

Step 355

Step 356

Step 357

Step 358

Step 359

Step 360

Step 361

Step 362

Step 363

Step 364

Step 365

Step 366

Step 367

Step 368

Step 369

Step 370

Step 371

Step 372

Step 373

Step 374

Step 375

Step 376

Step 377

Step 378

Step 379

Step 380

Step 381

Step 382

Step 383

Step 384

Step 385

Step 386

Step 387

Step 388

Step 389

Step 390

Step 391

Step 392

Step 393

Step 394

Step 395

Step 396

Step 397

Step 398

Step 399

Step 400

Step 401

Step 402

Step 403

Step 404

Step 405

Step 406

Step 407

Step 408

Step 409

Step 410

Step 411

Step 412

Step 413

Step 414

Step 415

Step 416

Step 417

Step 418

Step 419

Step 420

Step 421

Step 422

Step 423

Step 424

Step 425

Step 426

Step 427

Step 428

Step 429

Step 430

Step 431

Step 432

Step 433

Step 434

Step 435

Step 436

Step 437

Step 438

Step 439

Step 440

Step 441

Step 442

Step 443

Step 444

Step 445

Step 446

Step 447

Step 448

Step 449

Step 450

Step 451

Step 452

Step 453

Step 454

Step 455

Step 456

Step 457

Step 458

Step 459

Step 460

Step 461

Step 462

Step 463

Step 464

Step 465

Step 466

Step 467

Step 468

Step 469

Step 470

Step 471

Step 472

Step 473

Step 474

Step 475

Step 476

Step 477

Step 478

Step 479

Step 480

Step 481

Step 482

Step 483

Step 484

Step 485

Step 486

Step 487

Step 488

Step 489

Step 490

Step 491

Step 492

Step 493

Step 494

Step 495

Step 496

Step 497

Step 498

Step 499

Step 500

Step 501

Step 502

Step 503

Step 504

Step 505

Step 506

Step 507

Step 508

Step 509

Step 510

Step 511

Step 512

Step 513

Step 514

Step 515

Step 516

Step 517

Step 518

Step 519

Step 520

Step 521

Step 522

Step 523

Step 524

Step 525

Step 526

Step 527

Step 528

Step 529

Step 530

Step 531

Step 532

Step 533

Step 534

Step 535

Step 536

Step 537

Step 538

Step 539

Step 540

Step 541

Step 542

Step 543

Step 544

Step 545

Step 546

Step 547

Step 548

Step 549

Step 550

Step 551

Step 552

Step 553

Step 554

Step 555

Step 556

Step 557

Step 558

Step 559

Step 560

Step 561

Step 562

Step 563

Step 564

Step 565

Step 566

Step 567

Step 568

Step 569

Step 570

Step 571

Step 572

Step 573

Step 574

Step 575

Step 576

Step 577

Step 578

Step 579

Step 580

Step 581

Step 582

Step 583

Step 584

Step 585

Step 586

Step 587

Step 588

Step 589

Step 590

Step 591

Step 592

Step 593

Step 594

Step 595

Step 596

Step 597

Step 598

Step 599

Step 600

Step 601

Step 602

Step 603

Step 604

Step 605

Step 606

Step 607

Step 608

Step 609

Step 610

Step 611

Step 612

Step 613

Step 614

Step 615

Step 616

Step 617

Step 618

Step 619

Step 620

Step 621

Step 622

Step 623

Step 624

Step 625

Step 626

Step 627

Step 628

Step 629

Step 630

Step 631

Step 632

Step 633

Step 634

Step 635

Step 636

Step 637

Step 638

Step 639

Step 640

Step 641

Step 642

Step 643

Step 644

Step 645

Step 646

Step 647

Step 648

Step 649

Step 650

Step 651

Step 652

Step 653

Step 654

Step 655

Step 656

Step 657

Step 658

Step 659

Step 660

Step 661

Step 662

Step 663

Step 664

Step 665

Step 666

Step 667

Step 668

Step 669

Step 670

Step 671

Step 672

Step 673

Step 674

Step 675

Step 676

Step 677

Step 678

Step 679

Step 680

Step 681

Step 682

Step 683

Step 684

Step 685

Step 686

Step 687

Step 688

Step 689

Step 690

Step 691

Step 692

Step 693

Step 694

Step 695

Step 696

Step 697

Step 698

Step 699

Step 700

Step 701

Step 702

Step 703

Step 704

Step 705

Step 706

Step 707

Step 708

Step 709

Step 710

Step 711

Step 712

Step 713

Step 714

Step 715

Step 716

Step 717

Step 718

Step 719

Step 720

Step 721

Step 722

Step 723

Step 724

Step 725

Step 726

Step 727

Step 728

Step 729

Step 730

Step 731

Step 732

Step 733

Step 734

Step 735

Step 736

Step 737

Step 738

Step 739

Step 740

Step 741

Step 742

Step 743

Step 744

Step 745

Step 746

Step 747

Step 748

Step 749

Step 750

Step 751

Step 752

Step 753

Step 754

Step 755

Step 756

Step 757

Step 758

Step 759

Step 760

Step 761

Step 762

Step 763

Step 764

Step 765

Step 766

Step 767

Step 768

Step 769

Step 770

Step 771

Step 772

Step 773

Step 774

Step 775

Step 776

Step 777

Step 778

Step 779

Step 780

Step 781

Step 782

Step 783

Step 784

Step 785

Step 786

Step 787

Step 788

Step 789

Step 790

Step 791

Step 792

Step 793

Step 794

Step 795

Step 796

Step 797

Step 798

Step 799

Step 800

Step 801

Step 802

Step 803

Step 804

Step 805

Step 806

Step 807

Step 808

Step 809

Step 810

Step 811

Step 812

Step 813

Step 814

Step 815

Step 816

Step 817

Step 818

Step 819

Step 820

Step 821

Step 822

Step 823

Step 824

Step 825

Step 826

Step 827

Step 828

Step 829

Step 830

Step 831

Step 832

Step 833

Step 834

Step 835

Step 836

Step 837

Step 838

Step 839

Step 840

Step 841

Step 842

Step 843

Step 844

Step 845

Step 846

Step 847

Step 848

Step 849

Step 850

Step 851

Step 852

Step 853

Step 854

Step 855

Step 856

Step 857

Step 858

Step 859

Step 860

Step 861

Step 862

Step 863

Step 864

Step 865

Step 866

Step 867

Step 868

Step 869

Step 870

Step 871

Step 872

Step 873

Step 874

Step 875

Step 876

Step 877

Step 878

Step 879

Step 880

Step 881

Step 882

Step 883

Step 884

Step 885

Step 886

Step 887

Step 888

Step 889

Step 890

Step 891

Step 892

Step 893

Step 894

Step 895

Step 896

Step 897

Step 898

Step 899

Step 900

Step 901

Step 902

Step 903

Step 904

Step 905

Step 906

Step 907

Step 908

Step 909

Step 910

Step 911

Step 912

Step 913

Step 914

Step 915

Step 916

Step 917

Step 918

Step 919

Step 920

Step 921

Step 922

Step 923

Step 924

Step 925

Step 926

Step 927

Step 928

Step 929

Step 930

Step 931

Step 932

Step 933

Step 934

Step 935

Step 936

Step 937

Step 938

Step 939

Step 940

Step 941

Step 942

Step 943

Step 944

Step 945

Step 946

Step 947

Step 948

Step 949

Step 950

Step 951

Step 952

Step 953

Step 954

Step 955

Step 956

Step 957

Step 958

Step 959

Step 960

Step 961

Step 962

Step 963

Step 964

Step 965

Step 966

Step 967

Step 968

Step 969

Step 970

Step 971

Step 972

Step 973

Step 974

Step 975

Step 976

Step 977

Step 978

Step 979

Step 980

Step 981

Step 982

Step 983

Step 984

Step 985

Step 986

Step 987

Step 988

Step 989

Step 990

Step 991

Step 992

Step 993

Step 994

Step 995

Step 996

Step 997

Step 998

Step 999

Step 1000

Figure 4.4: Input parameter walkthrough: Steps 3 and 4

4.4 Optimization Tools and Methods

4.4.1 Gurobi Optimizer Through a High Performance Computer

A simple way to access the Gurobi optimization solver through the HPC is using the Open OnDemand website (<https://ood.hpc.arizona.edu/>), which is an NSF-funded open-source HPC portal and is available for users and provides web browser access for interfacing with HPC. Fig. 4.5 to Fig. 4.7 will guide you through the three steps to run Python with Gurobi optimizer through HPC.

After entering the OnDemand website, go to “My Interactive Sessions” → “Jupyter Notebook”. Once you’ve entered all your details, click “Launch” at the bottom of the page. This will create a job request for Jupyter Notebook in the HPC. After that, you can use Jupyter Notebook to view, edit, and upload/download your files. In the Jupyter Notebook, find the corresponding ipynb file and click “Run All” under the “Cell” button. Then the optimization model will be implemented by Python with Gurobi optimizer through HPC (see Fig. 4.7).

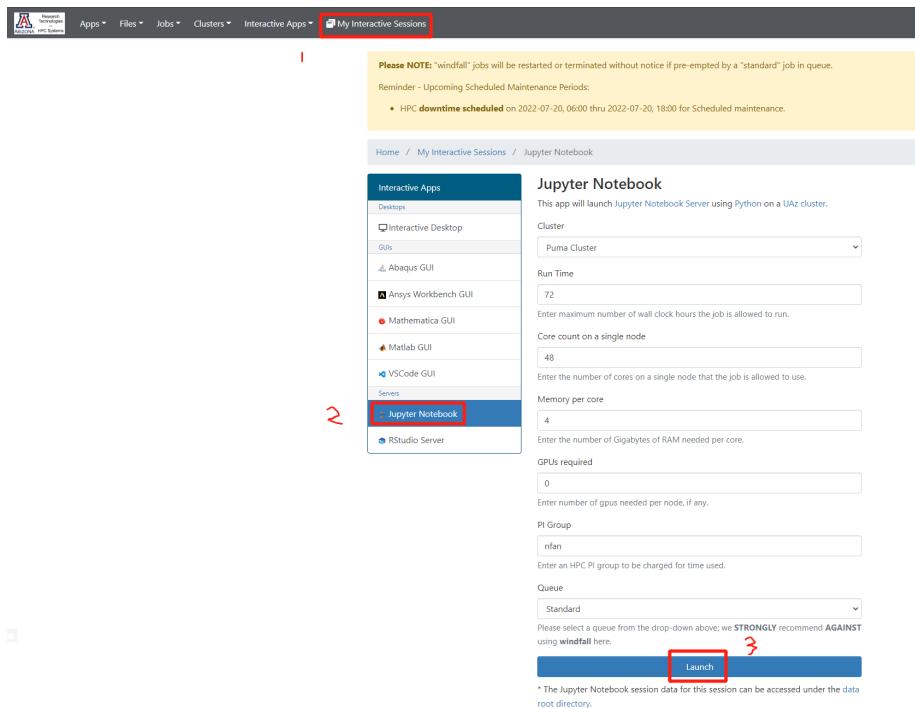


Figure 4.5: Running Gurobi optimizer in the HPC (Part 1)

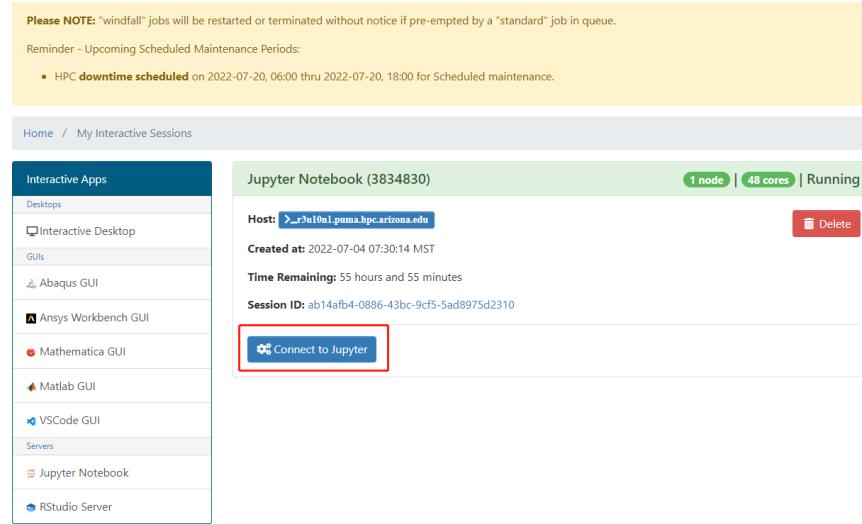


Figure 4.6: Running Gurobi optimizer in the HPC (Part 2)

```

jupyter harvest_v10 Last Checkpoint: 13 hours ago (autosaved)
File Edit View Insert Cell Kernel Widgets Help Trusted Logout
In [1]: # import modules
import numpy as np
from gurobipy import *

# Sets and indices
F = list(range(int(f.readline()) + 1)) # fields
H = list(range(1, int(f.readline()) + 1)) # harvesters
K = list(range(1, int(f.readline()) + 1)) # harvest seasons
T = list(range(int(f.readline()) + 1)) # time periods

## Parameters
C_h = float(f.readline())/2 # rent cost: divided by 2 to make the conversion from per year to per harvest season
if C_h < 0.001:
    C_h = 1 # add a small epsilon to avoid redundant harvester when C_h = 0
C_o = float(f.readline()) # operational cost
C_r = float(f.readline()) # transportation cost
C_penalty = float(f.readline()) # penalty cost (derive from average gross profit)

M = int(f.readline()) * np.ones(len(F)) # maximum number of harvesters that can harvest one field in a given period
H = M.tolist()
M_cap = float(f.readline()) # harvesting capacity

loc = int(f.readline()) # facility location & 29 is the default facility location
A = [0] + list(map(float, f.readline().split())) # acres for fields & 29 is the default facility location
R = [] # ready date
D = [] # due date
for k in K:
    R.append([0, k]) # (0, k) = (1, k)
for i, ss in enumerate(f.readline().split(), start=1):
    R[i, k] = int(ss)
for i, ss in enumerate(f.readline().split(), start=1):
    D[i, k] = int(ss)

ins = []
I = np.zeros((len(F), len(K) + 1)) # indicator
for k in K:
    ins[k] = list(map(int, f.readline().split()))
    for idx in ins[k]:
        I[idx, k] = 1 # indicate which fields need to be harvested during season k

# distance matrix
d = np.zeros((len(F), len(F)))
for i in F[1:]:
    for j, ss in enumerate(f.readline().split(), start=1):
        d[i, j] = int(ss)

```

Figure 4.7: Running Gurobi optimizer in the HPC (Part 3)

4.4.2 Reporting Outputs

Table 4.3: Harvest Planning and Machinery Scheduling: Decision Variables

$u_h^k \in \{0, 1\}$:	$u_h^k = 1$ if harvester h is used in harvest season k
$x_{ft}^k \in \{0, 1\}$:	$x_{ft}^k = 1$ if field f is harvested at period t in harvest season k
$y_{ht,(i,j)}^k \in \{0, 1\}$:	$y_{ht,(i,j)}^k = 1$ if harvester h moves from field i to j at the end of period t of harvest season k
$z_{hft}^k \in \{0, 1\}$:	$z_{hft}^k = 1$ if harvester h for field f is used at period t in harvest season k
$\alpha_{hft}^k \in [0, 1]$:	Harvesting fraction for field f using harvester h in period t of harvest season k
$s_f^k \in \{0, \dots, T \}$:	Harvest starting period for field f in harvest season k
$e_f^k \in \{0, \dots, T \}$:	Harvest ending period for field f in harvest season k

4.5 Model Assumptions

Model assumptions are specified as follows: 1) Suppose a time period is one week, and guayule is harvested in each field in a time window $[R_f^k, D_f^k]$ during the harvest season k . 2) Once a field is available, it should be harvested before the due date. Otherwise, there may exist some penalties for not harvesting on time. 3) Guayule in all fields should be completely harvested over the harvest season. 4) Suppose there exists only one warehouse/facility in this harvest area, and all machinery departs from the warehouse/facility at the beginning of a harvest season and returns to it at the end of the harvest season. 5) Machinery moves between fields at the end of the period in order to fully utilize productive time during the next working period.

Chapter 5

Conclusions

This instruction manual summarizes the development and current capabilities of the BSCT model. The instruction manual focuses on three major sections. One on a stochastic optimization model to achieve optimal decisions for facility locations, transportations from field to facilities, and finally to customers, with a multi-objective approach to quantify the economic benefits, environmental, and social impacts. This model is developed based on the adoption requirements of the two desert-dwelling crops guar and guayule and modifications may be required to assess the adoption of other crops. To capture the uncertainty of the adoption rates from each farm, stochastic scenarios were applied. The model and algorithm were tested on a case study for two desert-dwelling crops, guayule in Arizona for Maricopa and Pinal counties, and guar in New Mexico for Dona Ana County (please refer to [1]).

The second major section focuses on an optimization model to achieve optimal operations for farm production planning and machinery scheduling for perennial and non-perennial crops considering multi-crop planning, multi-machinery scheduling, and crop rotation within a defined set of time periods. Two case studies are developed, one for the guar crop (non-perennial) considering parcels located in the Cochran and Hockley counties, TX, and one for the guayule crop (6-year perennial) considering parcels located in Pinal County, AZ (please refer to [2]).

The third major section focuses on an optimization model to obtain optimal harvest planning and machinery scheduling under a series of factors like periodic harvest, different demands, resource limitations, etc. In addition to satisfy different types of requirements, this optimization model also aims to find the best decision-making strategy in order to minimize the total cost. Case studies for three crops located in Pinal County, Arizona are developed to validate the model (please refer to [3]).

We consider that BSCT will provide a fundamental framework for further facility location, production and scheduling optimization research. Future studies may consider complementing our model with optimization approaches for

production planning and machinery scheduling in semi-arid regions to maximize the different stakeholders' benefits. Additional research to increase the sale value of guar and guayule may be required for their adoption in the Southwestern U.S. agriculture. Rainy weather may affect the harvesting scheduling due to level of moisture requirements for the processing facilities. Therefore, further weather analysis considering rain risk during the harvesting periods can be considered in future research. Additionally, a water life-cycle-analysis per crop can be performed to identify farming regions potential water availability risks which may lead to consider higher adoption rates for guayule and guar within those regions.

Bibliography

- [1] Daniel A. Zuniga Vazquez, Ou Sun, Neng Fan, Evan Sproul, Hailey M. Summers, Jason C. Quinn, Sita Khanal, Paul Gutierrez, VeeAnder Mealing, Amy E. Landis, Clark Seavert, Trent Teegerstrom, and Blase Evancho. Integrating environmental and social impacts into optimal design of guayule and guar supply chains. *Computers & Chemical Engineering*, 146:107223, 2021.
- [2] Daniel A. Zuniga Vazquez, Neng Fan, Trent Teegerstrom, Clark Seavert, Hailey M Summers, Evan Sproul, and Jason C Quinn. Optimal production planning and machinery scheduling for semi-arid farms. *Computers and Electronics in Agriculture*, 187:106288, 2021.
- [3] Shunyu Yao, Neng Fan, Clark Seavert, and Trent Teegerstrom. Demand-driven harvest planning and machinery scheduling for guayule. *Operations Research Forum*, 4(1):1–25, 2023.
- [4] USDA. United states department of agriculture. irrigation & water use, 2019. <https://www.ers.usda.gov/topics/farm-practices-management/irrigation-water-use> (accessed December 2, 2020).
- [5] Mehran Homayounfar, Sai Hin Laiand, Mehdi Zomorodian, Ali Reza Sepaskhah, and Arman Ganji. Optimal crop water allocation in case of drought occurrence, imposing deficit irrigation with proportional cutback constraint. *Water Resources Management*, 28:3207 – 3225, 2014.
- [6] G. Niu, Y.P. Li, G.H. Huang, J. Liu, and Y.R. Fan. Crop planning and water resource allocation for sustainable development of an irrigation region in china under multiple uncertainties. *Agricultural Water Management*, 166:53 – 69, 2016.
- [7] Chenglong Zhang, Mo Li, and Ping Guo. An interval multistage joint-probabilistic chance-constrained programming model with left-hand-side randomness for crop area planning under uncertainty. *Journal of Cleaner Production*, 167:1276 – 1289, 2017.
- [8] SBAR. Sustainable bioeconomy for arid regions. project objectives, 2017. <https://sbar.arizona.edu/our-goals/project-objectives> (accessed July 21, 2020).

- [9] Griffin M. Bates and Katrina Cornish. Rapid and complete removal of guayule (*parthenium argentatum*) leaves by cryodefoliation, and freeze and thaw induction of rubber particle coagulation. *Industrial Crops and Products*, 125:491 – 495, 2018.
- [10] Daniel C. Ilut, Paul L. Sanchez, Terry A. Coffelt, John M. Dyer, Matthew A. Jenks, and Michael A. Gore. A century of guayule: Comprehensive genetic characterization of the guayule (*parthenium argentatum* a. gray) usda germplasm collection. *bioRxiv*, 2017.
- [11] Abdulraheim Hasan and Manar El-Sayed Abdel-Raouf. Applications of guar gum and its derivatives in petroleum industry: A review. *Egyptian Journal of Petroleum*, 27, 05 2018.
- [12] Archana George, Priyanka A. Shah, and Pranav S. Shrivastav. Guar gum: Versatile natural polymer for drug delivery applications. *European Polymer Journal*, 112:722 – 735, 2019.
- [13] USDA. United states department of agriculture. census of agriculture, 2018. <https://www.nass.usda.gov/AgCensus/> (accessed January 28, 2020).
- [14] NASS. National agricultural statistics service. cropscape - cropland data layer, 2019. <https://nassgeodata.gmu.edu/CropScape/> (accessed January 28, 2020).