

1

Supplementary Material: Modeling 2 Commodity Flow in the Context of 3 Invasive Species Spread: Study of *Tuta* 4 *absoluta* in Nepal

5

6 **Organization.** Here, we will elaborate on the methods used to implement each module,
7 and provide additional results. Some of the content from the main paper is repeated for
8 continuity.

9

1 Flow network construction

10

1.1 Network construction

11 Regional markets serve as key locations facilitating agricultural commodity flow, hence it
12 makes sense to model the flow network with markets represented as nodes. We model
13 the flow of agricultural produce among markets based on the following premise: the total
14 outflow from a market is a function of the amount of produce in its surrounding
15 regions, and the total inflow is a function of the population to which it caters and the
16 corresponding per capita income. The main assumptions in this model are as follows. (i)
17 Imports and exports are not significant enough to influence domestic trade. (ii) Fresh
18 tomatoes are mainly traded for consumption. This motivates the use of population and
19 per capita income as indicators of tomato consumption in a given district. (iii) The higher
20 the per capita income, the greater the consumption.

21 The flows are estimated using a doubly constrained gravity model ([Kaluza et al., 2010](#);
22 [Anderson, 2011](#)). The flow F_{ij} from location i to j is given by

$$F_{ij} = a_i b_j O_i I_j f(d_{ij}) \quad (1)$$

23 where, O_i is the total outflow of the commodity from i , I_j is the total inflow to j , d_{ij} is the
24 distance to travel from i to j , $f(\cdot)$ is the *distance deterrence function*, and coefficients a_i
25 and b_j are computed through an iterative process to ensure flow balance.

Table 1: Datasets.

Description	Source	Resolution	Year
Population	Nepal Central Bureau of Statistics (http://cbs.gov.np/)	District/Town	2011
Per Capita Income	Nepal Central Bureau of Statistics (http://cbs.gov.np/)	District	2011
Tomato production	Nepal Ministry of Agricultural Development (MOAD) (http://moad.gov.np/)	District, Annual	2015
Production seasonality	iDE Nepal (http://idenepal.org/) and MOAD	Region, Monthly	2016
Major vegetable markets	MOAD Marketing Information System (http://www.agrimis.gov.np/)	Town	2017
Market distances	Google Maps, Distance matrix API	Market	2017
Tomato import/exports	Food and Agriculture Organization (FAOSTAT) (www.fao.org/faostat/)	Country, Annual	2013
Tomato consumption	FAOSTAT, MOAD	Country, Annual	2013
Flows to Kalimati market	Official website (kalimatinmarket.gov.np/)	District, Annual	2015
<i>T. absoluta</i> incidence reports	Nepal National Agriculture Research Council (http://narc.gov.np/), USAID IPM Innovation Lab and iDE Nepal	District/town	2017

Table 2: Notation and abbreviations.

Variables	Description
F_{ij}	Commodity flow from node i to j
O_i	Total outflow of commodity from node i
I_i	Total inflow of commodity into node i
d_{ij}	Distance between nodes i and j
$f(\cdot)$	deterrence function
β	Power-law exponent of gravity model
κ	Cutoff time of gravity model
γ	Per capita income parameter
σ	Gaussian parameter for spatial seeding
t	Time step for the spread model

26 **Seasonality of production** Due to altitude and temperature variations, the tomato
 27 production season varies across the regions of Nepal (see Figure ??). Production in the
 28 Mid Hills and High Hills is largely restricted to the summer months of June to November
 29 (referred to as season S1), while the Terai region produces during the winter months
 30 of December to May (referred to as season S2). As a result, we have two distinct flow
 31 networks, one for each season. We partitioned the districts into two groups: Mid Hills
 32 and High Hills belong to group 1, while the Terai districts belong to group 2. All districts
 33 belonging to group i were assigned their respective annual production for season S_i and
 34 zero for the other season.

35 **Market scope definition** The nodes of the flow network are the major markets, 69 in
 36 all, after merging markets that belong to the same town. Recall that the amount of
 37 production is specified at the district level. In order to obtain the production estimates
 38 at market level, we defined *market scope* as follows: The country’s map was overlaid
 39 by a grid cell of size $5km \times 5km$ and we constructed a Voronoi partition of these cells
 40 using market locations as centroids. This was motivated by the fact that tomato sellers
 41 and buyers will seek out the nearest market. We assumed uniform spatial distribution of

42 production within each district. Each grid cell was assigned a value of production in a
43 particular season proportional to the fraction of the area of the district covered by the
44 cell. The total outflow from the market is the sum of production of the grid cells assigned
45 to it for a particular season.

46 **Modeling consumption** We modeled the total inflow I_i into a market as a product of
47 the population catered to by the market and a function of the average per capita income
48 associated with the market η_i, η_i^γ , where γ is a tunable parameter. The population catered
49 to by the market was derived from district level population data and the market scope as
50 defined for production redistribution.

51 **Inter-market travel time** Owing to the diverse landscape of Nepal and varying road
52 conditions we used travel time by road instead of the geodesic or road distance between the
53 markets. We geolocated major vegetable markets using Google Maps. We then manually
54 embedded the market locations onto the Nepal road network, and constructed a planar
55 network by connecting the markets which have a direct route (without going through other
56 markets) between them. We also removed markets which were completely inaccessible
57 by road. We used Google Distance Matrix API¹ to compute travel times by road along
58 the edges of this planar network. This in turn yields a road network among the markets,
59 where the edges are weighted by their travel time. Distance between any two markets is
60 then obtained as the shortest travel time on the road network. The distance deterrence
61 function $f(d_{ij}) = d_{ij}^{-\beta} \exp(-d_{ij}/\kappa)$ combines power-law and exponential decay with d_{ij}
62 which can be controlled by the tunable parameters β , the power-law exponent, and κ , the
63 cutoff time.

64 1.2 Construction workflow

65 Recall that nodes of the resulted network are major markets and directed weighted edges
66 represent the commodity flow from one market to another. Following Kaluza et al.
67 (2010); Anderson (2011), we used a gravity model to estimate the flows. The flow F_{ij}
68 from location i to j is given by $F_{ij} = a_i b_j O_i I_j f(d_{ij})$, where O_i is the total outflow of the
69 commodity from i , I_j is the total inflow to j , d_{ij} is the distance to travel from i to j , $f(\cdot)$
70 is the *distance deterrence function*. The coefficients a_i and b_j are computed through an
71 iterative process such that the total outflow and total inflow at each vertex agree with the
72 input values. The total outflow from each market i , O_i is the amount of produce that arrives
73 to the market in the specified season. The total inflow is the size of the population catered
74 to by the market times a function of the per capita income η_i, η_i^γ , where γ is a tunable
75 parameter. Here we use a general deterrence function: $f(d_{ij}, \beta, \kappa) = d_{ij}^{-\beta} \exp(-d_{ij}/\kappa)$,
76 where d_{ij} is the time taken to travel between markets i and j .

77 Therefore, to construct such a flow network, we need to estimate O_i and I_i for each
78 market i , as well as pairwise distance d_{ij} between market i and market j . Since the data
79 of population, per capita income, and tomato production are at the district level (see
80 Figure 1), whereas the nodes in the resulting network are markets, we need a mechanism

¹<https://developers.google.com/maps/documentation/distance-matrix/>

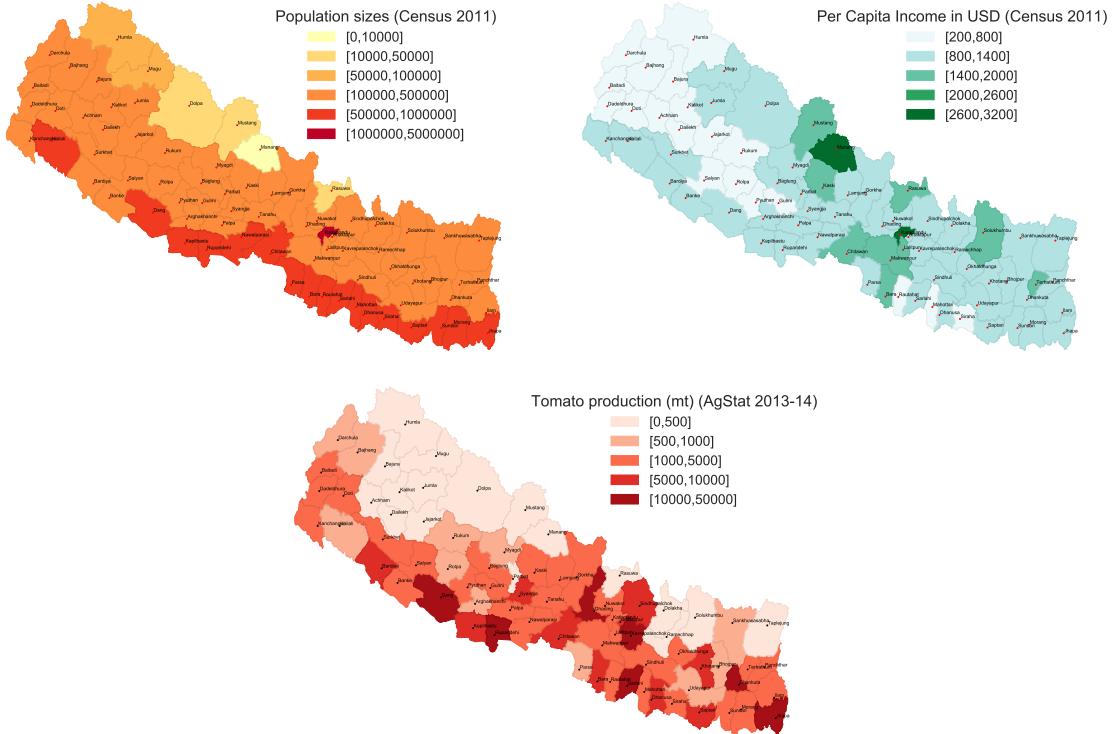


Figure 1: **District-level datasets.** (a) Population size. (b) Per capita income.
 (c) Tomato production.

81 to map the district level data to the individual markets. Following is a step by step
 82 description of the flow network construction pipeline.

83 **Step 1: Partition production data based on seasons.** The seasonal tomato production
 84 is shown in Figure 2b. Based on this we partitioned the districts into two groups: Mid
 85 Hills and High Hills belong to group 1, while the Terai districts belong to group 2. All
 86 districts belonging to group i were assigned their respective annual production for season
 87 S_i and zero for the other season.

88 **Step2: Estimate tomato consumption.** The total tomato consumption of a district is
 89 estimated by the size of its population times a function of the per capita income η , η^γ ,
 90 where γ is a tunable parameter.

91 **Step 3: Map district level data to individual markets.** The country's map was over-
 92 layed by a grid cell of size 25 sq.km. We constructed a Voronoi partition of these cells
 93 using node locations as centroids. We assumed uniform spatial distribution of production
 94 and population for each district. Each grid cell was assigned a value of production in that
 95 season (consumption) which was proportional to the fraction of the area of the district
 96 covered by the cell. Then, we assign each cell to its closest market and the total inflow
 97 (outflow) to the market is the sum of consumption (production) of the grid cells assigned

98 to it. See Figure 2a for the assignment of cells to each market.

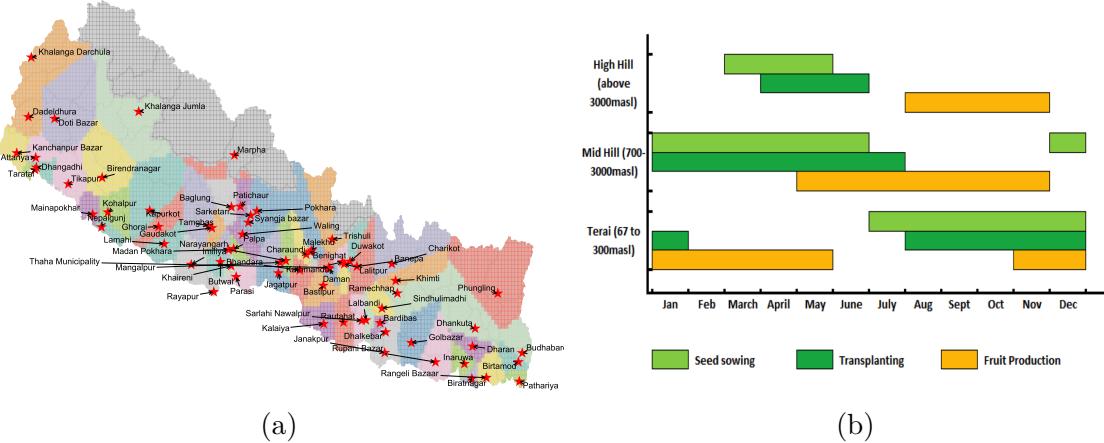


Figure 2: **Assigning seasonal node attributes.** (a) Market scope: Assignment of cells to individual markets. We have excluded six districts: Humla, Mugu, Dolpa, Mustang, Manang, and Rasuwa (shown in grey color), since they have low population/production and are disconnected by the road network. (b) Altitude-induced production cycle for tomato. We divided the year into two parts, season S1 (June to November) and season S2 (December to May).

99 **Step 4: Constructing road network and estimating distances.** We first manually
100 identified pairs of nodes which were directly connected by road (without any other node in
101 between) and used the Google Distance Matrix API [Google \(2017\)](#) to computed the d_{ij} s
102 (in minutes). Then we applied Dijkstra's shortest path algorithm for weighted graphs to
103 compute pairwise travel time between markets.

Step 5: Estimate gravity model coefficients and calculate flows. The scaling factors a_i and b_j are obtained by iteratively solving the system of equations

$$a_i = \frac{1}{\sum_j b_j I_{ij} f(d_{ij})},$$

$$b_j = \frac{1}{\sum_i a_i O_{ij} f(d_{ij})}.$$

104 Kaluza et al. [Kaluza et al. \(2010\)](#) show that the iterative process converges to fixed values
105 of a_i and b_j . There is a tolerance factor which enables faster convergence at the cost of
106 accuracy of these parameters, and in turn the flow. We set the tolerance factor to 0.01.

107 2 Dynamic stochastic model for spread

108 2.1 Spread Dynamics

109 We develop a discrete-time SI (Susceptible-Infected) epidemic model on directed weighted
110 networks ([Pastor-Satorras et al., 2015](#)) to model pest dispersal. Each node is either

111 susceptible (free from pest) or infected (pest is present). Henceforth, we use the term
 112 “infected” for a node or a region frequently to imply *T. absoluta* infestation at that location.
 113 A node i in state I infects each of its out-neighbors j in the network with probability
 114 proportional to the flow F_{ij} at each time step t . The infection probabilities are obtained
 115 by normalizing flows globally: $\lambda_{ij} = \frac{F_{ij}}{\max_{i,j} F_{ij}}$. The model is based on two assumptions:
 116 (i) an infected node remains infected and continues to infect its neighbors, and (ii) the
 117 chance of infection is directly proportional to the volume traded. Considering the fact
 118 that Nepal was ill-prepared for this invasion and the lack of effective intervention methods,
 119 (i) is a fair assumption. Historically, *T. absoluta* has spread rapidly in regions where
 120 tomato trade has been the highest (parts of Europe and Middle-East for example) thus
 121 motivating assumption (ii).

Let $P_S(i, t, f_0)$ denote the probability that node i remains uninfected (i.e., susceptible)
 by time t given the initial condition f_0 which assigns probability of infection at time
 step $t = 0$ to each node. In general, computing P_S is difficult. Efficient methods have
 been proposed to estimate this probability (Lokhov et al., 2014). Here, we adopt the
 $dynamic message passing algorithm$ by Lokhov et al. (Lokhov et al., 2014), summarized
by the following equations.

$$\begin{aligned}
 P_S^{i \rightarrow j}(t+1) &= P_S(i, 0, f_0) \prod_{k \in \delta i \setminus j} \theta^{k \rightarrow i}(t+1) \\
 \theta^{k \rightarrow i}(t+1) &= \theta^{k \rightarrow i}(t) - \lambda_{ki} \phi^{k \rightarrow i}(t) \\
 \phi^{k \rightarrow i}(t) &= (1 - \lambda_{ki}) \phi^{k \rightarrow i}(t-1) \\
 &\quad - [P_S^{k \rightarrow i}(t) - P_S^{k \rightarrow i}(t-1)]
 \end{aligned} \tag{2}$$

122 In the above equations, λ_{ki} is the infection probability across edge (k, i) , and θ, ϕ are
 123 intermediate messages used to update the node states. Finally, the quantity of interest
 124 $P_S(i, t, f_0)$, the probability that node i remains uninfected (i.e., susceptible) till time t is
 125 given as:

$$P_S(i, t+1, f_0) = P_S(i, 0, f_0) \prod_{k \in \delta i} \theta^{k \rightarrow i}(t+1)$$

126 Note that for any given t , $P_S(i, t, f_0) + \gamma_i^t(f_0) = 1$, and hence the entire evolution of the
 127 epidemic on the network is captured by $P_S(i, t, f_0), \forall i, t$ given the initial condition f_0 .

128 The initial configuration f_0 is chosen to mimic a spatially dispersed seeding scenario.
 129 We first select a *central* seed node, and then use a Gaussian kernel with parameter σ
 130 around the seed node to assign initial infection probabilities for neighboring markets. A
 131 market at a geodesic distance d from the seed is assigned the infection probability $e^{-\frac{d^2}{2\sigma^2}}$.
 132 The kernel accounts for factors such as uncertainty in determining the pest location, the
 133 possibility of spread of the pest through natural means, as well as interactions between
 134 these markets.

135 The message passing approach for simulating the SI epidemic model is adapted from
 136 Lokhov et al. (2014). The framework uses two cavity messages $\theta^{i \rightarrow j}(\cdot), \phi^{i \rightarrow j}(\cdot)$ which are
 137 exchanged across each edge in the network. The initial conditions are set as follows:

$$\begin{aligned}
 \theta^{i \rightarrow j}(0) &= 1 \\
 \phi^{i \rightarrow j}(0) &= P_I(i, 0, f_0)
 \end{aligned}$$

¹³⁹ where $P_I(i, 0, f_0)$ is the initial seeding probability for node i obtained using spatial
¹⁴⁰ seeding given the initial condition f_0 . The closed set of recursion rules are given by:

$$\begin{aligned} \text{141} \quad P_S^{i \rightarrow j}(t+1) &= P_S(i, 0, f_0) \prod_{k \in \delta i \setminus j} \theta^{k \rightarrow i}(t+1) \\ \text{142} \quad \theta^{k \rightarrow i}(t+1) - \theta^{k \rightarrow i}(t) &= -\lambda_{ki} \phi^{k \rightarrow i}(t) \\ \phi^{k \rightarrow i}(t) &= (1 - \lambda_{ki})(1 - \mu_k) \phi^{k \rightarrow i}(t-1) - [P_S^{k \rightarrow i}(t) - P_S^{k \rightarrow i}(t-1)] \end{aligned}$$

¹⁴³ In the above equations λ_{ki} is the infection probability across edge (k, i) and μ_k is the
¹⁴⁴ recovery probability for node k . For an SI model, nodes remain infected (infested in our
¹⁴⁵ case), and never recover. Thus, $\mu_k = 0$.

¹⁴⁶ Finally, the quantity of interest $P_S(i, t, f_0)$, the probability that node i remains uninfected
¹⁴⁷ (i.e., susceptible) till time t is given as:

$$P_S(i, t+1, f_0) = P_S(i, 0, f_0) \prod_{k \in \delta i} \theta^{k \rightarrow i}(t+1)$$

¹⁴⁸ Note that for any given t , $P_S(i, t, f_0) + P_I(i, t, f_0) = 1$, and hence the entire evolution of
¹⁴⁹ the epidemic on the network is captured by $P_S(i, t, f_0), \forall i, t$ given the initial condition f_0 .

¹⁵⁰ 3 Monitoring and distribution of *T. absoluta*

¹⁵¹ Several organizations are involved in the monitoring of *T. absoluta* spread in Nepal. Our
¹⁵² sources are primarily NARC, USAID IPM-IL, ENBAITA, and Agricare Pvt. Ltd. IPM-IL
¹⁵³ works through iDE Nepal. On May 3, 2016, *T. absoluta* was officially reported by NARC's
¹⁵⁴ entomology division, Khumaltar, Lalitpur. During the first quarter of 2016, farmers
¹⁵⁵ from Kathmandu, Bhaktapur, and Kavre districts reported concern about the new pest
¹⁵⁶ attacking their tomato plants. A team from IPM-IL visited these sites and collected the
¹⁵⁷ moth and larva of the pest. Lures from Pest Control India (PCI) were installed in these
¹⁵⁸ infested fields. Samples of the trapped larva and adults were sent to the School of Life
¹⁵⁹ Science, Arizona State University in June, 2016. Results came positive for *T. absoluta*.

¹⁶⁰ In a preliminary assessment from May to June, 2016 heavy outbreaks of *T. absoluta*
¹⁶¹ were reported from 15 Village Development Committees (VDC) of Kathmandu, 9 VDC
¹⁶² of Bhaktapur, 6 VDC of Lalitpur, and 3 VDC of Kavre district. The pheromone trap
¹⁶³ installed in 1 VDC of Dhading, 3 VDC of Kaski, 3 VDC of Banke, 4 VDC of Surkhet, 2
¹⁶⁴ VDC of Jhapa, and 1 VDC of Sunsari district showed no sign of the pest. Since April 25,
¹⁶⁵ 2017, more incidence of *T. absoluta* have been reported from additional districts: Chitwan,
¹⁶⁶ Kaski, Palpa, Syangja, Surkhet, Banke, Saptari, and Kailali.

¹⁶⁷ 4 Economic impact

¹⁶⁸ The notations used in this section are given in Table 3. The total economic impact or the
¹⁶⁹ change in social welfare is the sum of change in consumers' and producers' surplus. The
¹⁷⁰ change in consumers' surplus is given by

$$\Delta CS = - \int_{P_1}^{P_2} \chi P^{-\eta} dP = - \frac{\chi}{1-\eta} P^{1-\eta} \Big|_{P_1}^{P_2}$$

171 where P_1 and P_2 are the old and new price respectively.

172 To calculate the change in producers' surplus, we first determine the new supply function
173 for each district which is the sum of unaffected plus affected supply i.e. $\sum_i \beta_i P^\theta l_i (1 - z_i) + \sum_i (1 - h) \beta_i (vP)^\theta l_i z_i$.

174 The new price must satisfy the equilibrium condition i.e.

$$f\chi P^{-\eta} = \sum_i \beta_i P^\theta l_i (1 - z_i) + \sum_i (1 - h) \beta_i (vP)^\theta l_i z_i$$

176 where the left-hand side is the domestic demand and the right-hand side is the sum of
177 unaffected supply and the affected supply in the domestic market.

178 Next we calculate the producers' surplus from P_1 to P_2 for each polygon i as

$$\begin{aligned} \Delta PS_i &= \int_0^{P_2} \beta_i P^\theta l_i (1 - z_{i,2}) dP + \int_0^{P_2} (1 - h) \beta_i (vP)^\theta l_i z_{i,2} dP \\ &\quad - \left(\int_0^{P_1} \beta_i P^\theta l_i (1 - z_{i,1}) dP + \int_0^{P_1} (1 - h) \beta_i (vP)^\theta l_i z_{i,1} dP \right) \\ &= \frac{\beta_i}{1 + \theta} l_i z_{i,2} P_2^{1+\theta} - \frac{\beta_i}{1 + \theta} l_i z_{i,1} P_1^{1+\theta} \\ &\quad + \left(\frac{\beta_i}{1 + \theta} l_i z_{i,2} (1 - h) v^\theta P_2^{1+\theta} - \frac{\beta_i}{1 + \theta} l_i z_{i,1} (1 - h) v^\theta P_1^{1+\theta} \right) \end{aligned}$$

179 where $z_{i,1}$ and $z_{i,2}$ correspond to no invasion and invasion scenarios.

180 To derive the total economic impact, we sum up the changes in consumers' and producers'
181 surplus i.e. $-\frac{x}{1-\eta} P^{1-\eta}|_{P_1}^{P_2} + \sum_i \left(\frac{\beta_i}{1+\theta} l_i z_{i,2} P_2^{1+\theta} - \frac{\beta_i}{1+\theta} l_i z_{i,1} P_1^{1+\theta} + \left(\frac{\beta_i}{1+\theta} l_i z_{i,2} (1 - h) v^\theta P_2^{1+\theta} - \frac{\beta_i}{1+\theta} l_i z_{i,1} (1 - h) v^\theta P_1^{1+\theta} \right) \right)$. For the actual economic impact, we instantiate the parameters
182 by assuming $h = 0.25$, $v = 0.2$, original price $P_1 = 400$ (\$/ton), $\eta = -0.7$, $\theta = 0.5$ and
183 $f = 0.94$. β_i is represented by the Yield of district i (i.e. $y_i l_i$) divided by P_1^θ . The
184 parameter values used here have been taken from the literature FAO (2016); Bajracharya
185 et al. (2016); USDA (2012); of Nepal; Khidr et al. (2013).

187 References

- 188 James E Anderson. The gravity model. *Annual Review of Economics*, 3(1):133–160, 2011.
- 189 Ajaya Shree Ranta Bajracharya, Ram Prasad Mainali, Binu Bhat, Sanjaya Bista,
190 PR Shashank, and NM Meshram. The first record of South American tomato leaf miner,
191 Tuta absoluta (Meyrick 1917)(Lepidoptera: Gelechiidae) in Nepal. *J. Entomol. Zool.
192 Stud*, 4:1359–1363, 2016.
- 193 FAO. Production and trade. <http://www.fao.org/faostat/en/#data>, 2016.
- 194 Google. Distance Matrix API. [https://developers.google.com/maps/documentation/
distance-matrix/](https://developers.google.com/maps/documentation/
195 distance-matrix/), 2017.

Table 3: Table of Notations.

Variable	symbol
Demand elasticity to domestic price	η
Yield	y
Fraction of demand met by domestic production	f
Market price	P
Tomato cultivation area	l
Supply elasticity to domestic price	θ
Proportion of the area affected	z
Proportion lost due to pest invasion	h
Increased cost of production due to control measures	v
Fraction of domestic demand met by domestic supply	f
Supply function	βP^θ
Demand function	$\chi P^{-\eta}$
Unaffected supply	$\beta P^\theta(1 - z)$
Affected supply	$(1 - h)\beta(vP)^\theta z$

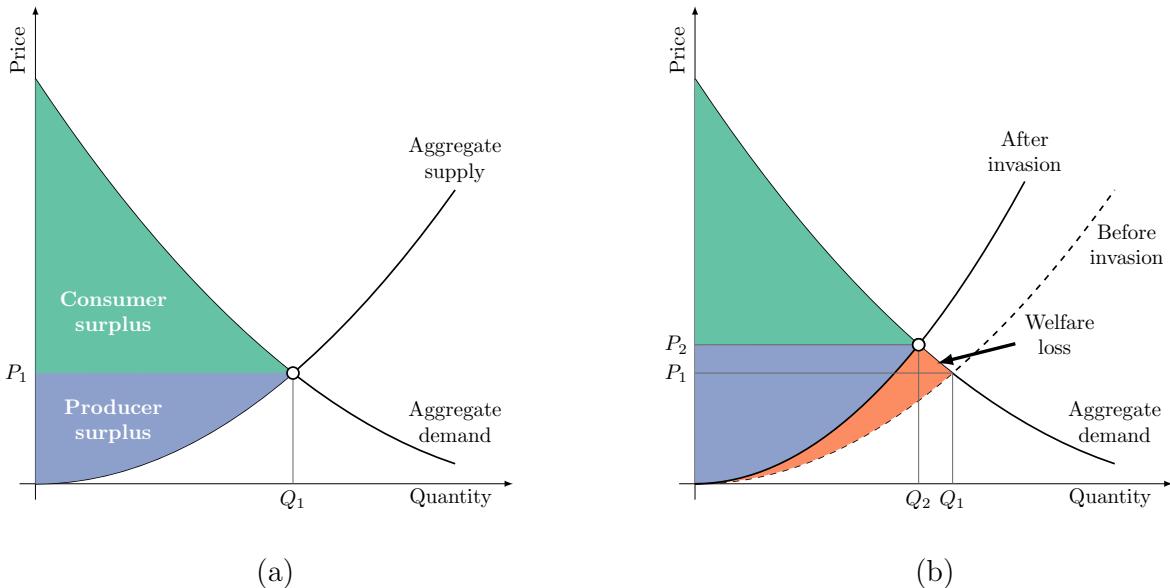


Figure 3: **Partial equilibrium model.** Fig (a) shows the original demand and supply curve, and the consumers' and producers' surplus as given by the green and blue shaded areas respectively. P_1 and Q_1 are the original equilibrium price and quantity respectively. Fig (b) shows the proportional shift in supply curve after *T. absoluta* invasion. This results in a higher equilibrium price and a lower equilibrium quantity. The updated consumers' and producers' surplus are shown in the same colors. The orange area shows the welfare loss to the society.

¹⁹⁶ Pablo Kaluza, Andrea Kölzsich, Michael T Gastner, and Bernd Blasius. The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48):

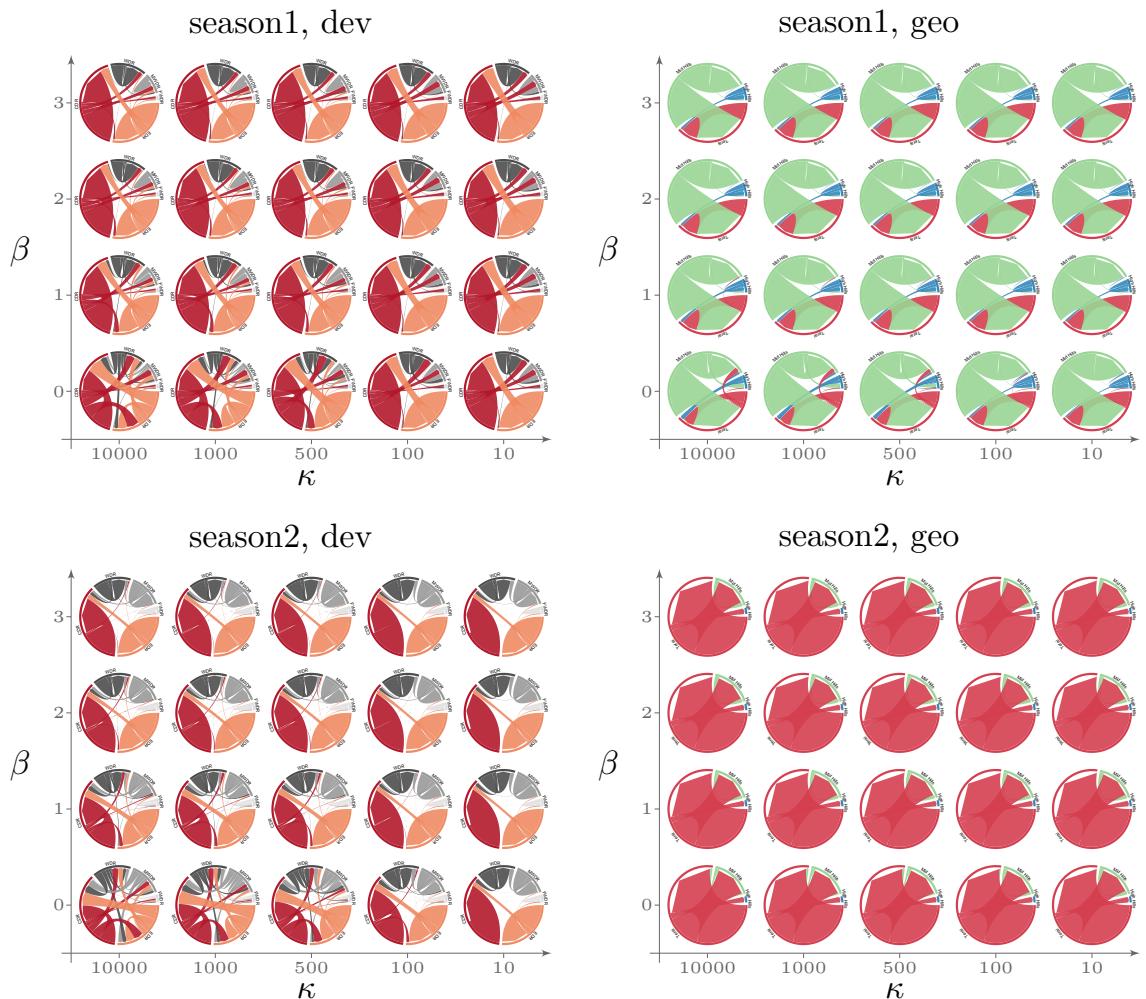


Figure 4: **Sensitivity of the flow networks to β and κ .** We note that as β is increased and κ is decreased, the interregional flow decreases significantly, particularly between the Development Regions. However, the general trends are preserved. For example, in season S1, we see the east to west flow. In season S2, we note that CDR is a big sink, and has significant flow entering from WDR and EDR.

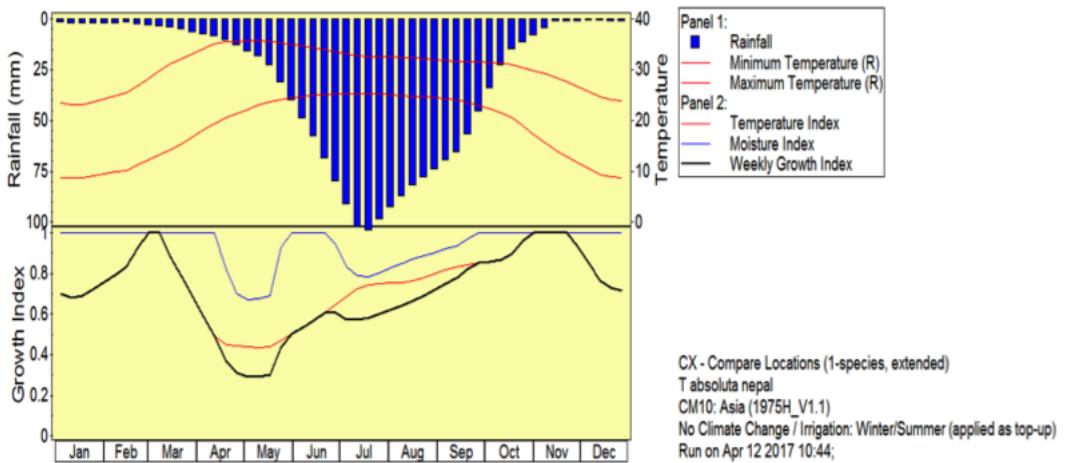


Figure 5: **Growth index of $T. absoluta$ in Kathmandu, Nepal.**

198 1093–1103, 2010.

199 AA Khidr, SA Gaffar, S Maha, A Nada, A Taman, A Fathia, and A Salem. New approaches
 200 for controlling tomato leafminer, *Tuta absoluta* (Meyrick) in tomato fields in Egypt.
 201 *Egyptian Journal of Agricultural Research*, 91(1):335–348, 2013.

202 Andrey Y Lokhov, Marc Mézard, Hiroki Ohta, and Lenka Zdeborová. Inferring the origin
 203 of an epidemic with a dynamic message-passing algorithm. *Physical Review E*, 90(1):
 204 012801, 2014.

205 Govenment of Nepal. Kalimati Fruits and Vegetable Market Development Board.
 206 kalimatimarket.com.np/daily-price-information.

207 Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro
 208 Vespignani. Epidemic processes in complex networks. *Reviews of modern physics*, 87
 209 (3):925, 2015.

210 USDA. Commodity and Food Elasticities. <https://data.ers.usda.gov/reports.aspx?ID=9464>, 2012.

212 **5 Flow validation and sensitivity analysis**

213 **Flow validation:** The unavailability of sample data on seasonal trade of tomato crop
 214 makes it challenging to calibrate and validate the flow network model. The only data that
 215 is available is the yearly data on the volume of tomato arriving from each district to the
 216 largest wholesale market of Nepal, Kalimati (located in Kathmandu). In Figures 7d–7f,
 217 we compare this data with the network flows. Given a set of network parameters (β, κ, γ) ,
 218 we obtained the inflow from a particular district to Kathmandu as follows: we combined

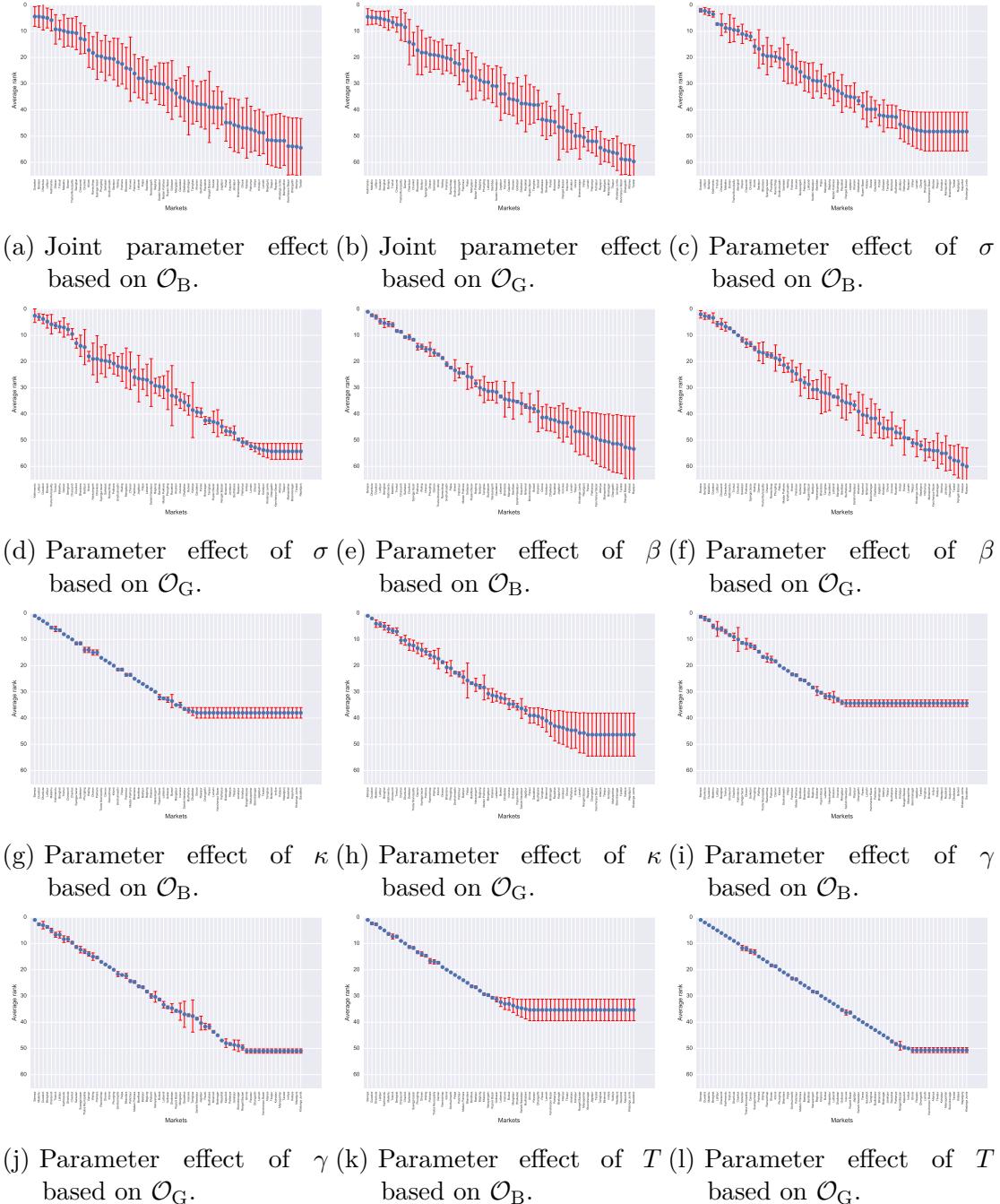
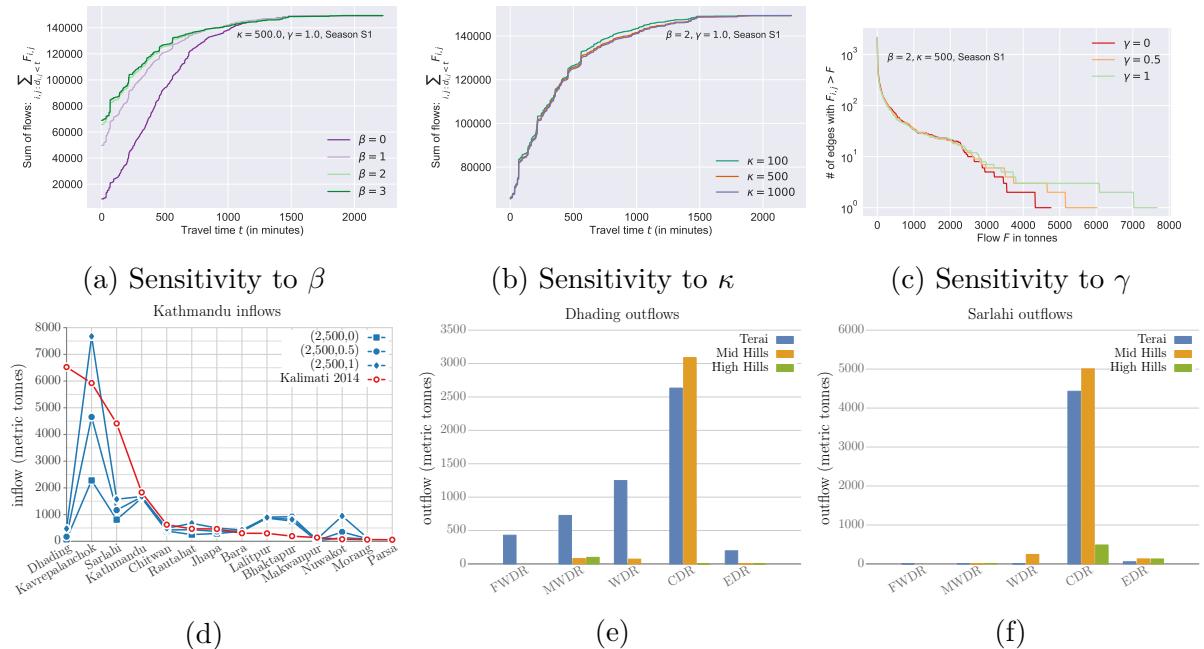


Figure 6: **Parameter effects on stability of individual market ranks.** Blue dots represent mean rank for each market and error bars are standard deviation.

219 the weights of all edges of the corresponding network with destination node “Kathmandu”
 220 and source nodes belonging to that district.

221 As seen in Figure 7d, for γ values between 0.5 and 1, the flows from the networks are
 222 comparable to the Kalimati data except for two districts: Dhading (the top contributor)
 223 and Sarlahi (third highest). Upon further investigation we find that Dhading, which is a
 224 major producer west of Kathmandu, serves the Mid Hills and Terai regions of the Central
 225 Development Region in the flow networks (Figure 7e). While the gravity model predicts
 226 that these flows will be directly delivered to these regions, in reality, it is possible that
 227 Dhading’s produce is routed through Kalimati market as there are several traders from
 228 Dhading registered in the Kalimati market². As for Sarlahi, even though there is little
 229 inflow to Kalimati market in the flow networks, other markets in the Kathmandu valley
 230 (belonging to Bhaktapur and Lalitpur districts) receive significant flows from Sarlahi
 231 (Figure 7f), which could, as in the previous case, be routed through Kalimati market.
 232 These issues highlight some of the limitations of the gravity model, which do not account
 for real-world trader dynamics.



233 Figure 7: Sensitivity analysis and flow validation

²<http://mrsmp.gov.np/files/download/tomato%20book.pdf>