Legacy report: A High-resolution Interaction Based Approach to Modeling the Spread of Agricultural Invasive Species

# Team

* PI: Abhijin Adiga (UVA)
* UVA: Stephen Eubank, Madhav Marathe, Achla Marathe, Henning Mortveit, Anil Vullikanti, S S Ravi, Hannah Chungbaek, Srinivasan Venkatramanan
* Cirad-Biopass, Dakar, Senegal: Thierry Brévault, Andrea Apolloni and Anaïs Chailleux,
* INRA, France: Nicolas Desneux, Mateus Ribeiro de Campos

# Background

With rapid population growth, shrinking farm acreage and intensive agriculture, society has come to critically depend on long distance flows of agricultural commodities. While this phenomenon has led to availability of a variety of commodities round the year, it has also caused unprecedented disruptions to the native ecosystems by facilitating the spread of invasive alien species. Therefore, modeling them in all their complexity and identifying the environmental, economic, and social factors that drive them is critical to ensure food security, biodiversity, health, and economic stability. The overarching goal of this project was to develop a generic high-resolution computational modeling framework to study the spread of agricultural invasive species which will (1) accurately represent both the ecological and socioeconomic aspects of the spread (2) provide causal explanation to the dynamics, and (3) facilitate the evaluation of various intervention policies. In our work, we addressed this issue, and applied novel methods to a representative agropest: the tomato leafminer (*Tuta absoluta*). To this end, we developed a multi-theory, multi-layer network approach to simulate the spread dynamics of *T. absoluta.* It accounted for both the ecological and human-mediated pathways of spread. Specific aims of this project were to (i) identify and address gaps in the understanding of the pest biology, (ii) design & develop a multi-network agent-based model to capture the effect of trade and travel on the pest dynamics, (iii) develop network-based strategies for efficient surveillance and control, and (iv) assess the economic impact of damage due to spread.

The research contributions of the project have been published in diverse venues corresponding to biological research, crop protection, artificial intelligence, network science, and big data venues. The project has contributed two highly cited survey papers [Biondi et al. 2018, Campos et al. 2017] covering the state-of-art on the research on biology, spread, impact, and control of *T. absoluta*. Several postdoctoral fellows, graduate and undergraduate students have been mentored. Our work has been presented in several conferences and workshops held in many countries. A webinar was also organized as part of this project on the biology and spread of *T. absoluta.*

This report is organized as follows. First, we describe the research contributions including major results. This is followed by capacity development where mentoring and presentations under the project are covered. Finally, we provide a description of software contributions and data products coming out of the project.

# Research Contributions

## Modeling framework

One of the main contributions of this project is a multi-pathway network-based model to simulate the spread of invasive agricultural pests [McNitt et al. 2019]. ﻿Our model accounts for both self-mediated and human-mediated spread and encapsulates the spatial heterogeneity, temporal variations, and multi-scale nature of the propagation mechanisms. Calibrating such complex models is challenging. ﻿We developed a framework to parameterize and analyze the multi-pathway model with respect to ground truth by a novel application of popular supervised and unsupervised machine learning algorithms. Our approaches are motivated by recent research on machine learning surrogates for agent-based models and interpretable artificial intelligence. This model has also been extensively analyzed using graph theoretic techniques and experimentation [Adiga et al. 2022]. A preliminary version of the model was developed to analyze the spread in Nepal [Venkatramanan et al. 2020], which only accounted for long-distance human-mediated spread. All the models developed are generic in the sense that they can be easily adapted to other invasive pests as well as other epidemiological applications concerning food trade.

We studied the spread of the pest in many regions of interest. Our work on the spread in Nepal [Venkatramanan et al. 2020] revealed for the first time, a strong correlation between the spread over the trade network and observational data. Further, our analysis strongly indicates that given its dependency on vegetable imports and the pattern of agricultural produce flow unraveled by our model, Nepal is extremely vulnerable to attacks from pests and pathogens that can spread through trade. We identified regions of Nepal which are most likely to be negatively impacted by the pest. ﻿In our second work on Southeast Asian countries [McNitt et al. 2019], analysis with respect to historical invasion records suggests that even with modest self-mediated spread capabilities, the pest can quickly expand its range through domestic city-to-city vegetable trade. Monitoring high-consumption areas can help in early detection, and targeted interventions at major production areas can effectively reduce the rate of spread.

## Multi-pathway network construction and analysis

The multi-pathway simulation model requires a high-resolution network representation of links between geographic regions that enable multi-pathway spread of the invasive species. A critical component of this project has been the construction and analysis of realistic representations of production, trade, and consumption of agricultural crops that can be applied to epidemiological processes such as invasive species spread, food poisoning, and biological warfare [Adiga et al. 2021, McNitt et al. 2019]. Inferring or estimating commodity flow networks is a major challenge as there is hardly any data available on commodity-specific flows. Even if available, the spatial and temporal resolutions of such datasets are not adequate. Besides, the availability of data differs by country or study region, thus posing a hurdle to generalizing the construction framework. Since trade dynamics have become complex, traditional spatial interaction models might not be able to characterize commodity flows. It is important to incorporate knowledge of growing periods, whole-sale market locations, fine-resolution estimates of crop production, imports, and exports. Fusion of multi-type datasets misaligned in space and time and validation are major challenges. We developed a general framework to construct high-resolution temporal networks that capture the production, trade, and consumption of agricultural crops. We used several datasets (at grid-level, administrative unit-level, and qualitative) including crop production, growing seasons, market locations, trade, imports and exports, spatial data capturing human activities, and human population all fused together to obtain multi-scale networks with node and edge attributes.

We rigorously analyzed the resulting networks using statistical methods and structural analysis. Our results suggest that agricultural systems are increasingly vulnerable to attacks through trade of commodities due to their vicinity to regions of high demand and seasonal variations in production and flows.

## Surveillance and intervention algorithms

One of the main aims of this project was to develop effective strategies to control the spread of invasive species. Invasive species can be controlled by screening different regions (i.e., a node intervention in a network), or cutting flows (i.e., an edge intervention). Such interventions incur huge economic costs and many countries simply do not have the infrastructure to implement country-wide control. Besides, possible delays in discovery of the pest must be accounted for. Therefore, designing optimal interventions to control invasive species is a fundamental challenge in agriculture. We developed a stochastic optimization-based control algorithm based on the multi-pathway model [Sudhir 2021, Sambaturu 2021]. Given a budget on how many locations can be identified for control and time delay between introduction of the species and its discovery, our algorithm provides a near-optimal solution by identifying locations where interventions can be applied. The performance of our algorithm is consistently superior compared to popular baselines. We analyzed solution sets for different introduction scenarios in several Southeast Asian countries.

In another work, we analyzed international trade networks corresponding to four solanaceous crops obtained using the Food and Agricultural Organization trade database using Moore-Shannon network reliability. The objective was to identify groups of highly connected countries or clusters from the perspective of pest spread. Identifying such clusters help in surveillance and mitigation strategies. We presented a novel approach to identify important dynamics-induced clusters of highly connected nodes in a directed weighted network [Nath et al. 2018]. Our analysis shows that the structure and dynamics can greatly vary across commodities. However, a consistent pattern that we observe in these commodity-specific networks is that almost all clusters that are formed are between adjacent countries in regions where liberal bilateral trade relations exist. Our analysis of networks of different years shows that intensification of trade has led to increased size of clusters, which implies that the number of countries spared from the network effects of disruption is reducing. Finally, applying this method to the aggregate network obtained by combining the four networks reveals clusters very different from those found in the constituent networks demonstrating the importance of analyzing commodity specific flows.

## Economic impact analysis

﻿We evaluated the economic impact of *T. absoluta* in Nepal based on the projection of its spread from the initial infestation in Kathmandu using our spread models. Our analysis uses several data sources on tomato production, imports & exports, supply & demand elasticity to price, consumption, etc. To calculate a comprehensive total economic impact, we used the partial equilibrium approach. Estimates of the number of locations where the pest has spread, and the extent of damage was obtained using our spread models. The comprehensive economic impact analysis shows a social welfare loss of $22.4M and a price increase of 32%.

## Thermal biology of *T. absoluta*

Most prior studies of *T. absoluta’s* thermal biology were based on populations from tropical regions. These are not particularly suitable in explaining its invasion of large areas of the Palearctic. In this work led by INRA [de Campos et al. 2021], extensive experimentation was conducted to study the thermal requirements, age-specific fertility of *T. absoluta* female and existence of facultative diapause. Our studies show that *T. absoluta* has lower thermal threshold and upper thermal threshold (and , respectively) than South American populations used in prior studies ( and ). Age-specific life tables were used to estimate the effects of temperature on its demographic parameters. Demographic parameters estimated at different temperatures show that T. absoluta has a rapid developmental time and can reproduce at temperatures close to the development threshold. Diapause in *T. absoluta* had not been characterized prior to our study. We found facultative diapause in pupae developing from larvae exposed to relatively low temperatures (i.e., 2 and 5 °C) and short-day length for different exposure periods. The strength of diapause was measured as an increase in post-treatment developmental times of pupae (i.e., degree days) that on average were 2.45–3-fold greater than of pupae reared at favorable temperatures. A lower developmental threshold and a facultative diapause increase the invasive potential of *T. absoluta* in temperate areas. Knowledge of this thermal biology is essential for predicting the potential geographic spread of this pest and to develop management and control strategies.

## Biological invasion risk assessment using Physiologically based demographic models

The capacity to assess invasion risk from potential crop pests before invasion of new regions globally would be invaluable, but this requires the ability to predict accurately their potential geographic range and relative abundance in novel areas. This may be unachievable using de facto standard correlative methods as shown for the South American tomato pinworm *T. absoluta* as very little or no data is available in many instances. This particularly applies to regions that are yet to be invaded by the pest. Its global invasive potential was not identified until after rapid invasion of Europe, followed by Africa and parts of Asia where it has become a major food security problem on solanaceous crops. Early prospective assessment of its potential range is possible using physiologically based demographic modeling that would have identified knowledge gaps in *T. absoluta* biology at low temperatures. Physiologically based demographic models (PBDMs) realistically capture the weather-driven biology in a mechanistic way allowing evaluation of invasive risk in novel areas and climes including climate change. PBDMs explain the biological bases for the geographic distribution, are generally applicable to species of any taxa, are not limited to terrestrial ecosystems, and hence can be extended to support ecological risk modeling in aquatic ecosystems. PBDMs address a lack of unified general methods for assessing and managing invasive species that has limited invasion biology from becoming a more predictive science. In this work, a mechanistic physiologically-based demographic model (PBDM) that captures the weather-driven biology of the moth is developed to explain and predict prospectively its invasiveness in the western Palearctic and in uninvaded areas such as the USA and Mexico under extant and climate change scenarios [Ponti et al. 2021].

## Geographic variation of host preference

We studied oviposition acceptance (proportion of females accepting a given plant as host for laying egg), oviposition preference (number of eggs laid by females on a given host plant) and performance (offspring development estimated as survival from egg to adult) of two *T. absoluta* populations originating from France (FRA) and Senegal (SEN) on six solanaceous plants (tomato, eggplant, Ethiopian eggplant, potato, sweet pepper and pepper) [Serigne, et al. 2019]. This effort was led by CIRAD (Senegal). The ovipositional behavioral pattern differed between the two populations; the SEN population showed higher oviposition acceptance on Ethiopian eggplant and sweet pepper than the FRA population. In addition, SEN population showed higher oviposition preference toward sweet pepper and potato than the FRA population. By contrast, the FRA population showed higher preference toward tomato and eggplant than the SEN population. The two populations of *T. absoluta* performed best on tomato (the preferred host plant) and showed similar decreasing trend in performance when comparing the two populations on the various other host plants. For both populations, performance on solanaceous plant species was closely related to ovipositional response of females to these plants. The differences observed between the two populations may indicate an ongoing differentiation in the host range of *T. absoluta* in the two invaded areas, possibly due to the abundance of these alternative host crops in Senegal at a period when tomato crops are scarce. This work was integrated into our multi-pathway models to estimate the establishment suitability given the volume of production of different host crops in that location.

## Biological control

The predatory species, *Nesidiocoris tenuis* (Heteroptera: Miridae), was collected, for the first time in Senegal, from tomato fields infested with *T. absoluta*. Laboratory feeding bioassay confirmed a potential of *N. tenuis* as a biocontrol agent of *T. absoluta* [Sylla et al. 2016]. This predator is commercially produced as a biological control agent worldwide. It is known as an effective natural enemy for controlling whiteflies, thrips, leafminers, aphids, mites, and eggs of lepidopteran pests. We monitored the spontaneous colonization and establishment of *N. tenuis* and other pest natural enemies of non-sprayed open field crops and net houses with two different mesh sizes. Results showed that the predator colonized the open field and was able to colonize the net house with the bigger mesh size, but too late to protect efficiently the crop. However, net houses were interesting at first to delay crop infestation by *T. absoluta* moths.

## Seasonal variations in *T. absoluta* distribution

Insect pest populations exhibit seasonal dynamics in response to changes in resource availability or other environmental factors such as climatic conditions, natural enemies, and intra- or interspecific competition. Understanding such dynamics is critical for developing effective integrated pest management strategies. The objective of the present study was to identify factors driving the seasonal decline of the tomato leafminer, *T. absoluta* (Meyrick), in the shifting landscape of a vegetable-growing area in Senegal. Several tomato fields were monitored for the number of *T. absoluta* adults caught in pheromone traps and for the incidence of larvae [Sylla et al. 2018]. The surface of solanaceous host crops, climatic conditions, and abundance of natural enemies were also monitored. A drastic decline in *T. absoluta* abundance was observed during the rainy season. The decrease in resource availability, especially tomato crops, in the surrounding landscape of monitored fields was the main factor affecting the population dynamics during the rainy season. However, alternative host crops such as eggplant and Ethiopian eggplant, can provide reservoirs for residual populations of the pest. For applied purpose, it would be interesting to focus the management efforts on residual populations during the rainy season, to make more difficult the reconstitution of populations during the dry season.

## Commodity flow data collection and analysis for Senegal

For Senegal, through extensive surveys seasonal production data at province-level and trade at city and market level were collected in 2017 [Mohamed 2017]. This data was applied in the construction and validation of tomato commodity flow for this region [Adiga et al. 2021]. It also contributed to our understanding of commodity flow dynamics. The sources of data for this study have been multiple. Data on tomato production at the department and district levels from national agencies and companies. In a second step, surveys were conducted with all types of actors in the sector (producers, traders) to obtain commodity flow data. This included interviews with various actors in wholesale markets in different regions of Senegal with prominent production and consumption of vegetables.

# Capacity building

Over the duration of the project, several postdoctoral fellows, graduate, and undergraduate students have worked on the topics corresponding to this project. This work has appeared as part of several PhD and masters thesis. One undergraduate student has received the “2022 Louis T. Rader Undergraduate Research” award from the Computer Science Department in UVA. A webinar “New Approaches to Control the South American Tomato Leaf Miner *T. absoluta*” informed the audience on new research in biology, ecology, and spread of the pest. Invited talks were delivered in several universities and agencies.

# Software and Data Products

A large amount of code was generated during the project. This corresponds to implementation of the simulator (multi-pathway model), algorithms (intervention algorithm, spatial interaction models, clustering, community detection, network reliability, Bayesian inference of invasion source, etc.), downloading, cleaning, transforming, and storing data, network analysis, and visualization. Several platforms were used to implement the simulator, analyze data and simulation outcomes. The primary language used was Python, followed by the R statistical package. The project used Sqlite and PostgresSql for management of data resulting several scripts and Python interfaces in these languages. For running the experiments, we have extensively used the HPC platform Rivanna, which uses SLURM workload manager. Much of our experiment design scripts are written using Shell scripts on LINUX platform. We have also used software such as NetLogo for simulations. Most of the software (including simulator code, clustering, and network analysis) is available in two Github repositories linked to two published papers [Adiga et al. 2022, McNitt et al. 2019], which have public access.

During the project several public data sources were accessed including but not limited to, FAOSTAT, NASS, WorldClim, LandScan, GPCC, FAF, USDA NASS, and Mapspam. Many datasets were obtained from by accessing databases, government reports of countries, research articles, and from collaborators. A subset of such data sources is listed in our papers [McNitt et al. 2021, Venkatramanan et al. 2020, Adiga et al. 2021]. The data collected can be accessed from public Github repositories [McNitt et al. 2019]. Important contributions of this project are derived datasets including set of commodity flow networks and gridded production and consumption data generated from various data sources using statistical and machine learning techniques. These are domestic tomato trade networks for most South and Southeast Asian countries and Senegal. Some of these data sources have been made public. Others will be made public through future publications that are in preparation.

# References

Adiga, Abhijin, et al. "Network Models and Simulation Analytics for Multi-scale Dynamics of Biological Invasions." *Frontiers in big Data* 5 (2022).

Adiga, Abhijin, et al. "Realistic Commodity Flow Networks to Assess Vulnerability of Food Systems." *International Conference on Complex Networks and Their Applications*. Springer, Cham, 2021.

Biondi, Antonio, et al. "Ecology, worldwide spread, and management of the invasive South American tomato pinworm, Tuta absoluta: past, present, and future." Annual review of entomology 63 (2018): 239-258.

Campos, Mateus R., et al. "From the Western Palaearctic region to beyond: *Tuta absoluta* 10 years after invading Europe." Journal of Pest Science 90.3 (2017): 787-796.

Campos, Mateus R., et al. "Thermal biology of Tuta absoluta: demographic parameters and facultative diapause." Journal of Pest Science 94.3 (2021): 829-842.

McNitt, Joseph, et al. "Assessing the multi-pathway threat from an invasive agricultural pest: Tuta absoluta in Asia." *Proceedings of the Royal Society B* 286.1913 (2019): 20191159.

Mohamed, I.A.: Cartographie des flux commerciaux de la tomate au S´en´egal: reconstitution des fili`eres, mod´elisation des flux et rˆ0le dans la propagation du ravageur *Tuta absoluta* (2017)

Nath, Madhurima, et al. "Using network reliability to understand international food trade dynamics." *International Conference on Complex Networks and their Applications*. Springer, Cham, 2018.

Ponti, Luigi, et al. “Biological invasion risk assessment of Tuta absoluta: mechanistic versus correlative methods.” Biological Invasions 23.12 (2021): 3809-3829.

Sambaturu, Prathyush. Controlling Epidemics on Networks Using Stochastic Optimization Techniques. University of Virginia, Computer Science - School of Engineering and Applied Science, PHD (Doctor of Philosophy), 2022.

Sudhir, Manisha. Controlling Diffusion on Multi-Pathway Spatial Networks: Application to Biological Invasions. University of Virginia, Computer Science - School of Engineering and Applied Science, MS (Master of Science), 2021, 2021, doi.org/10.18130/7pc8-me44.

Sylla, S., et al. "First record of Nesidiocoris tenuis (Reuter)(Heteroptera: Miridae), as a predator of the tomato leaf miner, Tuta absoluta (Meyrick)(Lepidoptera: Gelechiidae), in Senegal." (2016).

Sylla, Serigne, et al. "Geographic variation of host preference by the invasive tomato leaf miner Tuta absoluta: implications for host range expansion." Journal of Pest Science 92.4 (2019): 1387-1396.

Sylla, Serigne, et al. "Seasonal decline of the tomato leafminer, T uta absoluta, in the shifting landscape of a vegetable‐growing area." Entomologia Experimentalis et Applicata 166.8 (2018): 638-647.

Venkatramanan, S., et al. "Modeling commodity flow in the context of invasive species spread: Study of Tuta absoluta in Nepal." *Crop Protection* 135 (2020): 104736.