

OIL PALM EXPANSION, FILOVIRUS SPILLOVER RISK,
AND OPTIMAL TAXATION IN AFRICA

by

PIT JONGWATTANAKUL

(Under the Direction of Susana Ferreira)

ABSTRACT

This dissertation comprises three integrated studies that examine the relationship between oil palm expansion and filovirus spillover risk in sub-Saharan Africa, combining remote sensing, spatial econometrics, and structural economic modeling.

Chapter 2 presents a novel methodology for generating annual estimates of oil palm plantation establishment using MODIS satellite imagery. The study builds upon existing Landsat-based methods by addressing temporal data gaps through a machine-learning classification framework that utilizes XGBoost and multiple vegetation indices. Although the model does not estimate land cover independently, it successfully detects bare-soil signals indicative of plantation establishment across 17 African countries from 2000 to 2020. Validation against the Descals et al. (2024) dataset shows high temporal and spatial consistency, capturing post-2013 expansion in both smallholder and industrial plantations.

Chapter 3 quantifies the epidemiological effects of oil palm expansion using a panel of 10,674 grid cells. Spatial regressions demonstrate that industrial plantations significantly increase the probability of filovirus spillovers. In contrast, smallholder plantations reduce risk at low to moderate densities but lose this protective effect when overly clustered. These

findings underscore the importance of plantation structure and scale in shaping disease ecology, suggesting that land-use configuration is a critical determinant of zoonotic emergence.

Chapter 4 incorporates these empirical insights into a spatial-dynamic bioeconomic model. The model captures household-level land allocation decisions under uncertainty and external health risks. Simulation results indicate that private land-use decisions diverge from the social optimum due to unpriced spillover externalities. A uniform tax of US\$15.30 per metric ton of crude palm oil is shown to reduce palm oil output by only 1.5% while internalizing the health risk externality and generating US\$86.70 million in public revenue and US\$112 million in annual health benefits. These results highlight the efficiency of modest fiscal instruments in realigning incentives toward socially desirable outcomes.

Collectively, the three chapters present a data-driven policy framework to mitigate zoonotic disease spillover risk while promoting agricultural growth, thereby demonstrating the feasibility of aligning public health and environmental goals through targeted economic interventions.

INDEX WORDS: Oil palm expansion, deforestation, MODIS, remote sensing, classification, bare soil detection, filovirus spillover, Ebola, Marburg, zoonotic disease, Africa, land-use change, smallholder oil palm plantations, industrial oil palm plantations, externalities, optimal taxation, environmental health, agricultural policy.

OIL PALM EXPANSION, FILOVIRUS SPILLOVER RISK,
AND OPTIMAL TAXATION IN AFRICA

by

PIT JONGWATTANAKUL

BEng, Chulalongkorn University, Thailand, 2007

MA, Thammasat University, Thailand, 2013

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2025

© 2025

Pit Jongwattanakul

All Rights Reserved

OIL PALM EXPANSION, FILOVIRUS SPILLOVER RISK,
AND OPTIMAL TAXATION IN AFRICA

by

PIT JONGWATTANAKUL

Major Professor:	Susana Ferreira
Committee:	Yukiko Hashida
	Shanjukta Nath
	John Paul Schmidt
	Alicia Peduzzi

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2025

DEDICATION

To my beloved parents, for their endless love and unwavering support.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Professor Susana Ferreira, for her invaluable guidance, patience, and continuous support throughout this research journey. Through her mentorship, I have learned what it means to be both an excellent advisor and a caring teacher. Her attentiveness and dedication will always be remembered and cherished.

I extend my heartfelt appreciation to my committee members: Yukiko Hashida and Shanjukta Nath, for their constructive comments and insights that significantly enhanced the quality of this work. Special thanks to John Paul Schmidt, who opened my eyes to the interdisciplinary connections between economics and ecology, particularly in statistical tools and methodological approaches used in ecological sciences. I am also grateful to Alicia Peduzzi, the remote sensing expert, for her valuable technical guidance.

I would like to acknowledge the Ebola Spillover Group for expanding my academic horizons by bridging economics with infectious disease research, particularly concerning filoviruses in Africa. My sincere thanks to Patrick R. Stephens, Nicole Gottdenker, John Drake, and Mekala Sundaram for their expert insights on Ebola and Marburg virus behavior, as well as bat ecology—knowledge areas that were crucial to my research. This group also provided financial support throughout my final three years of study.

I am deeply grateful to Thammasat University for their institutional support and for providing the fellowship that covered my living expenses and tuition fees, alleviating financial concerns that could have impacted my academic performance.

My appreciation extends to my Thai mentors who provided unwavering support and inspiration throughout my doctoral journey: Archanun Kohpaiboon, Supawat Rungsuriyawiboon, Nattapong Puttanapong, Orapan Nabangchang, Rawadee Jarungrattanapong, and the late Porphant Ouyyanont, whose guidance and legacy continue to inspire me.

I would like to thank my fellow graduate students. Although I may not have fully immersed myself in American culture—as most of my cohort were not American—I had the wonderful opportunity to learn about South Asian cultures, including those of India, Nepal, and Bangladesh. I particularly enjoyed the delicious food from their home kitchens.

My gratitude goes to all Thai students, seniors, and juniors at UGA who provided mutual support, friendship, and invaluable assistance in various matters. They also shared their experiences and guidance, which was essential for someone like me who had never lived abroad before, making my time here less lonely and more meaningful.

I am profoundly grateful to my parents, Prasit Jongwattanakul and Malee Jongwattanakul, for their unwavering belief in me, their constant encouragement, and their care for each other, which relieved me of worry and allowed me to focus on my studies. I also dedicate this work to the memory of my late uncle, Prakrit Jongwattanakul, with deep remembrance and affection.

Finally, my deepest appreciation to Nuntawan Duangpamorn for her understanding, encouragement, and support throughout this academic journey.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER	
1 INTRODUCTION	1
1.1 Oil Palm Expansion and Filovirus Risk in Africa	1
1.2 The Challenge of Sustainable Oil Palm Expansion	2
1.3 Filoviruses at the Human-Environment Interface	3
1.4 Research Objectives and Dissertation Structure	4
1.5 Broader Significance and Policy Relevance	5
1.6 References	6
2 EXTENDING THE TEMPORAL AND SPATIAL CLASSIFICATION OF OIL PALM PLANTATIONS IN AFRICA THROUGH REMOTE SENSING TECHNOLOGIES	12
2.1 Introduction	12
2.2 Methods	16
2.3 Results	30
2.4 Discussion	35
2.5 Conclusions	37

2.6	References	38
3	OIL PALM PLANTATIONS, DEFORESTATION, AND AFRICAN FILOVIRUSES	45
3.1	Introduction.....	45
3.2	Existing Evidence and Conceptual Framework.....	47
3.3	Data	53
3.4	Estimation Strategy	60
3.5	Results.....	63
3.6	Robustness Checks.....	65
3.7	Conclusions and Policy Implications.....	73
3.8	References	76
4	OPTIMAL SMALLHOLDING AND INDUSTRIAL OIL PALM PLANTATIONS: ACCOUNTING FOR THE FILOVIRUS SPILLOVER RISK IN AFRICA	83
4.1	Introduction.....	83
4.2	Theoretical Framework and Model.....	86
4.3	Numerical Simulation and Results.....	103
4.4	Conclusions and Policy Implications.....	113
4.5	References.....	116
5	CONCLUSIONS AND POLICY IMPLICATIONS	121

APPENDICES

A	Localities of Ebola and Marburg Spillovers	125
B	Econometric Derivation of Cost Function Parameters	126
C	Epidemiological Derivation of Spillover Risk Parameters	132

LIST OF TABLES

	Page
Table 2.1: Spectral Indices and Reflectance Measurements Used in Classification	20
Table 2.2: Distribution of Random Points Across Land Cover Categories	22
Table 2.3: Results of the classifications.....	27
Table 2.4: Cross-validation results for bare soil classification algorithms using MODIS data.....	28
Table 3.1: Descriptive statistics at cell level.....	53
Table 3.2: Main model estimates	64
Table 3.3: Three Alternative Spatial Definitions of Filovirus Spillover	67
Table 3.4: Combining Smallholder and Industrial Plantations	69
Table 3.5: Using an Alternative Oil Palm Dataset.....	70
Table 3.6: Individual Land-Use Effects on Filovirus Spillover.....	71
Table 4.1: Summary of Core Calibration Parameters	106
Table 4.2: Comparison of Baseline and Spillover-Inclusive Outcomes	112
Table A.1: Localities of Ebola and Marburg Spillovers	125
Table B.1: Summary Statistics for Key Variables	127
Table B.2: Regression Results for Smallholder Plantation Area	129
Table B.3: Regression Results for Industrial Plantation Area	129
Table C.1: Regression Results for Filovirus Spillover Probability.....	134

LIST OF FIGURES

	Page
Figure 2.1: Map of the 17 African countries included in this study	16
Figure 2.2: Workflow for the data preparation	23
Figure 2.3: Classification Workflow for Bare Soil Detection from 2000-2020	30
Figure 2.4: Timeline of Land Cover Transitions	31
Figure 2.5: Decision Tree for Plantation Year Estimation	32
Figure 2.6: Comparison of Smallholder Oil Palm Plantations	34
Figure 2.7: Comparison of Industrial Oil Palm Plantations.....	35
Figure 3.1: Ebola and Marburg Cases in Africa between 2001 – 2018	55
Figure 3.2: Creation of binary indicator FV_{kt}	56
Figure 3.3: Definition of each cell	66
Figure 4.1: Supply Curves	91
Figure 4.2: No Production.....	92
Figure 4.3: Industrial-Only Supply	93
Figure 4.4: Both Types Supplying	94
Figure 4.5: Social-Optimum Supply Function (All Produce).....	101
Figure 4.6: Optimal Uniform Unit Tax.....	103
Figure 4.7: International Crude Palm Oil Prices, 1997–2018.....	104

CHAPTER 1

INTRODUCTION

1.1 Oil Palm Expansion and Filovirus Risk in Africa

The global community stands at the nexus of agricultural expansion, environmental change, and public health risk. In tropical regions, rapid population growth has driven agricultural frontiers into ecologically sensitive landscapes (Gibbs et al., 2010; Hansen et al., 2013). While this process supports food security and economic growth, it also fosters novel interfaces between human populations and wildlife, facilitating the spillover of zoonotic diseases—pathogens that originate in animals but infect humans (Jones et al., 2008; Allen et al., 2017). The COVID-19 pandemic has underscored the potential of such spillovers to escalate into global crises, disrupting economies and societies worldwide (IPBES, 2020).

Africa exemplifies this challenge. With rich biodiversity, expanding agricultural economies, and recurrent zoonotic outbreaks, the continent is a critical setting for understanding the intersection of land-use change and disease emergence (Kilpatrick et al., 2017). Among the agricultural drivers of deforestation and landscape alteration, oil palm (*Elaeis guineensis*)—a native species increasingly cultivated in industrial forms—stands out. While its expansion has brought economic gains, the concurrent rise of filovirus outbreaks, including Ebola and Marburg, highlights an urgent need to evaluate the public health implications of agricultural transformation (Feldmann & Geisbert, 2011; WHO, 2016).

This dissertation investigates the ecological and epidemiological dimensions of oil palm expansion across Africa, emphasizing its relationship to filovirus spillovers. Through the integration of remote sensing, spatial epidemiology, and environmental economics, it aims to develop a policy-relevant understanding of how land-use patterns influence disease risks.

1.2 The Challenge of Sustainable Oil Palm Expansion

Oil palm cultivation presents both opportunity and risk in Africa's development trajectory. Traditionally grown by smallholders for local consumption, the crop has more recently become a focus of commercial agriculture due to its high yield per hectare—far surpassing that of alternative oilseeds (Rival & Levang, 2014). Between 2000 and 2020, the continent's oil palm area tripled, rising from approximately 1 million to 2.9 million hectares, with corresponding production growth from 121 to 414 million tons of fresh fruit bunches (FAOSTAT, 2020).

This expansion has generated significant economic benefits. It has created employment opportunities, increased household incomes, and contributed to both domestic food security and international trade (Meijaard et al., 2020). Yet these gains have been accompanied by environmental degradation, including deforestation, biodiversity loss, and elevated carbon emissions (Vijay et al., 2016; Austin et al., 2017). A less visible but equally pressing concern is the impact of landscape alteration on zoonotic disease emergence—an externality that is often neglected in development planning (Plowright et al., 2021).

Crucially, not all oil palm systems pose the same risks. Industrial monocultures differ ecologically from smallholder-dominated mosaics. The former are characterized by large, uniform tracts with intensive management, while the latter integrate oil palms with food crops and maintain heterogeneous land cover. These divergent systems influence habitat structure, biodiversity, and

the frequency of human–wildlife interactions in distinct ways (Descals et al., 2021; Gaveau et al., 2022), with potential implications for disease transmission dynamics.

1.3 Filoviruses at the Human–Environment Interface

Filoviruses—particularly Ebola and Marburg—are among the most lethal zoonoses known. Outbreaks of Ebola virus disease (EVD) have produced case fatality ratios as high as 90%, while Marburg virus disease (MVD) has caused mortality rates ranging from 20% to 90% (Feldmann & Geisbert, 2011; CDC, 2024). Beyond their health impacts, filovirus outbreaks disrupt healthcare systems, curtail economic activity, and generate long-term social costs (Huber et al., 2018).

Ecologically, fruit bats of the family *Pteropodidae* are suspected reservoirs of filoviruses. The Egyptian fruit bat (*Rousettus aegyptiacus*) is a confirmed host for Marburg virus, while other species—such as *Hypsignathus monstrosus*, *Epomops franqueti*, and *Myonycteris torquata*—have been implicated in Ebola virus ecology (Leroy et al., 2005; De Nys et al., 2018). These bats frequently forage in agricultural landscapes, including oil palm plantations, which offer food and roosting sites (Shafie et al., 2011; Oleksy et al., 2015). Consequently, oil palm monocultures may serve as ecological magnets that draw bats closer to human populations, increasing the risk of disease spillover (Leroy et al., 2009).

Agricultural intensification can also erode ecological buffers that regulate pathogen transmission, such as species diversity and trophic interactions (Wilcox & Ellis, 2006; Keesing et al., 2010). While previous studies have linked deforestation to filovirus outbreaks (Olivero et al., 2017; Rulli et al., 2017), the specific role of oil palm expansion—particularly the differential risks posed by distinct cultivation systems—remains poorly understood.

1.4 Research Objectives and Dissertation Structure

This dissertation addresses this knowledge gap through three interconnected chapters, each building sequentially from data generation to policy design. The first research component (Chapter 2) develops a novel classification method to map the expansion of industrial and smallholder oil palm plantations across 17 African countries from 2000 to 2020. Using MODIS satellite imagery and machine learning algorithms, the analysis provides high-resolution, annually updated maps that improve upon existing datasets in both temporal depth and spatial accuracy.

Chapter 3 then links these spatial patterns to zoonotic disease risks, specifically filovirus spillover events. Employing spatial panel econometric techniques, the analysis estimates the causal effects of different plantation types on spillover probability, controlling for deforestation, climate, and socioeconomic variables. The results reveal that industrial plantations are associated with increased risk, while smallholder systems exhibit a protective effect when maintained at moderate scales.

Building on these findings, Chapter 4 presents a bioeconomic model that integrates epidemiological externalities into land-use decision-making. The model simulates the effects of fiscal policy instruments—particularly a uniform tax on crude palm oil—on land allocation, output, and public health outcomes. Results suggest that modest taxation can internalize spillover risks without substantially reducing agricultural productivity or profitability.

Together, these chapters offer a comprehensive, interdisciplinary framework for understanding and managing the health-environment-agriculture nexus in the African context.

1.5 Broader Significance and Policy Relevance

The methodological and policy contributions of this dissertation extend beyond the African oil palm context. First, the remote sensing approach developed in Chapter 2 can be applied to other perennial crops or land-use transitions in data-scarce settings. Second, the econometric strategy used in Chapter 3 offers a generalizable framework for quantifying how agricultural systems influence zoonotic disease emergence. Third, the economic modeling in Chapter 4 demonstrates how environmental externalities can be systematically incorporated into fiscal instruments for land-use governance.

From a policy standpoint, these insights are timely. Many African nations are seeking to expand their agricultural sectors while grappling with recurrent epidemics. This research provides practical guidance for integrating disease risk into development planning through spatial targeting and economic incentives (Barbier et al., 2020; Dobson et al., 2020). It also supports One Health approaches (WHO, 2017) that emphasize the interdependence of human, animal, and environmental health.

In sum, this dissertation contributes to a growing literature that seeks to align economic development with ecological resilience and public health. By highlighting the risks and trade-offs embedded in land-use change, it offers a roadmap for designing agricultural policies that foster sustainable growth while minimizing unintended consequences.

1.6 References

- Allen, T., Murray, K. A., Zambrana-Torrel, C., Morse, S. S., Rondinini, C., Di Marco, M., Breit, N., Olival, K. J., & Daszak, P. (2017). Global hotspots and correlates of emerging zoonotic diseases. *Nature Communications*, 8(1), 1124. <https://doi.org/10.1038/s41467-017-00923-8>
- Amman, B. R., Nyakarahuka, L., McElroy, A. K., Dodd, K. A., Sealy, T. K., Schuh, A. J., Shoemaker, T. R., Balinandi, S., Atimmedi, P., Kaboyo, W., Nichol, S. T., & Towner, J. S. (2014). Marburgvirus resurgence in Kitaka Mine bat population after extermination attempts, Uganda. *Emerging infectious diseases*, 20(10), 1761–1764. <https://doi.org/10.3201/eid2010.140696>
- Austin, K. G., Mosnier, A., Pirker, J., McCallum, I., Fritz, S., & Kasibhatla, P. S. (2017). Shifting patterns of oil palm driven deforestation in Indonesia and implications for zero-deforestation commitments. *Land Use Policy*, 69, 41–48. <https://doi.org/10.1016/j.landusepol.2017.08.036>
- Barbier, E. B., Leshner, A., Bloomberg, M., Cement, A.-Y., Bank, T. W., Centre, T. C. D., Duke, N., Foundation, F., Foundation, G., Foundation, M. D., Fund, T. N. C., & Steyer, T. (2020). The economic value of natural capital and protected areas in Africa. *Nature Sustainability*, 3(7), 564–571.
- Bausch, D. G., Nichol, S. T., Muyembe-Tamfum, J. J., Borchert, M., Rollin, P. E., Sleurs, H., Campbell, P., Tshioko, F. K., Roth, C., Colebunders, R., Pirard, P., Mardel, S., Olinda, L. A., Zeller, H., Tshomba, A., Kulidri, A., Libande, M. L., Mulangu, S., Formenty, P., Grein, T., ... International Scientific and Technical Committee for Marburg Hemorrhagic

- Fever Control in the Democratic Republic of the Congo (2006). Marburg hemorrhagic fever associated with multiple genetic lineages of virus. *The New England journal of medicine*, 355(9), 909–919. <https://doi.org/10.1056/NEJMoa051465>
- Centers for Disease Control and Prevention. (2024, May 1). *About Marburg*. U.S. Department of Health & Human Services. <https://www.cdc.gov/marburg/about/index.html>
- Corley, R. H. V., & Tinker, P. B. (2015). *The oil palm* (5th ed.). Wiley-Blackwell. <https://doi.org/10.1002/9781118953297>
- De Nys, H. M., Kingebeni, P., Keita, A. K., Butel, C., Thaurignac, G., Villabona-Arenas, C....Peeters, M. (2018). Survey of Ebola Viruses in Frugivorous and Insectivorous Bats in Guinea, Cameroon, and the Democratic Republic of the Congo, 2015–2017. *Emerging Infectious Diseases*, 24(12), 2228–2240. <https://doi.org/10.3201/eid2412.180740>.
- Descals, A., Wich, S., Meijaard, E., Gaveau, D. L. A., Peedell, S., & Szantoi, Z. (2021). High-resolution global map of smallholder and industrial closed-canopy oil palm plantations. *Earth System Science Data*, 13(3), 1211–1231. <https://doi.org/10.5194/essd-13-1211-2021>
- Dobson, A. P., Pimm, S. L., Hannah, L., Kaufman, L., Ahumada, J. A., Ando, A. W., Bernstein, A., Busch, J., Daszak, P., Engelmann, J., Kinnaird, M. F., Li, B. V., Loch-Temzelides, T., pandemic prevention. *Science*, 369(6502), 379–381. <https://doi.org/10.1126/science.abc3189>
- FAOSTAT. (2020). *Food and Agriculture Organization of the United Nations statistical database*. <http://www.fao.org/faostat/en/#home>
- Feldmann, H., & Geisbert, T. W. (2011). Ebola haemorrhagic fever. *The Lancet*, 377(9768), 849–862. [https://doi.org/10.1016/S0140-6736\(10\)60667-8](https://doi.org/10.1016/S0140-6736(10)60667-8)

- Gaveau, D. L. A., Locatelli, B., Salim, M. A., Husnayaen, Manurung, T., Descals, A., Angelsen, A., Meijaard, E., & Sheil, D. (2022). Slowing deforestation in Indonesia follows declining oil palm expansion and lower oil prices. *PLOS ONE*, 17(3), e0266178. <https://doi.org/10.1371/journal.pone.0266178>
- Gibbs, H. K., Ruesch, A. S., Achard, F., Clayton, M. K., Holmgren, P., Ramankutty, N., & Foley, J. A. (2010). Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. *Proceedings of the National Academy of Sciences*, 107(38), 16732–16737. <https://doi.org/10.1073/pnas.0910275107>
- Han, B. A., Kramer, A. M., & Drake, J. M. (2016). Global patterns of zoonotic disease in mammals. *Trends in Parasitology*, 32(7), 565–577. <https://doi.org/10.1016/j.pt.2016.04.007>
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850–853. <https://doi.org/10.1126/science.1244693>
- Huber, C., Finelli, L., & Stevens, W. (2018). The economic and social burden of the 2014 Ebola outbreak in West Africa. *The Journal of Infectious Diseases*, 218(Suppl. 5), S698–S704. <https://doi.org/10.1093/infdis/jiy213>
- Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). (2020). *Workshop report on biodiversity and pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services* (Version 1.4). IPBES Secretariat. <https://doi.org/10.5281/zenodo.7432079>

- Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, 451(7181), 990–993. <https://doi.org/10.1038/nature06536>
- Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., & Ostfeld, R. S. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*, 468(7324), 647–652. <https://doi.org/10.1038/nature09575>
- Kilpatrick, A. M., Salkeld, D. J., Titcomb, G., & Hahn, M. B. (2017). Conservation of biodiversity as a strategy for improving human health and well-being. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1722), 20160131. <http://doi.org/10.1098/rstb.2016.0131>
- Leroy, E. M., Epelboin, A., Mondonge, V., Pourrut, X., Gonzalez, J.-P., Muyembe-Tamfum, J.-J., & Formenty, P. (2009). Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector-Borne and Zoonotic Diseases*, 9(6), 723–728. <https://doi.org/10.1089/vbz.2008.0167>
- Leroy, E. M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassanin, A., Yaba, P., Délicat, A., Paweska, J. T., Gonzalez, J.-P., & Swanepoel, R. (2005). Fruit bats as reservoirs of Ebola virus. *Nature*, 438(7068), 575–576. <https://doi.org/10.1038/438575a>
- Meijaard, E., Brooks, T. M., Carlson, K. M., Slade, E. M., Garcia-Ulloa, J., Gaveau, D. L. A., Lee, J. S. H., Santika, T., Juffe-Bignoli, D., Struebig, M. J., Wich, S. A., Ancrenaz, M., Koh, L. P., Zamira, N., Abrams, J. F., Prins, H. H. T., Sendashonga, C. N., Murdiyarso, D., Furumo, P. R., ... Sheil, D. (2020). The environmental impacts of palm oil in context. *Nature Plants*, 6(12), 1418–1426. <https://doi.org/10.1038/s41477-020-00813-w>

- Morse, S. S., Mazet, J. A. K., Woolhouse, M., Parrish, C. R., Carroll, D., Karesh, W. B., Zambrana-Torrel, C., Lipkin, W. I., & Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *The Lancet*, 380(9857), 1956–1965.
[https://doi.org/10.1016/S0140-6736\(12\)61684-5](https://doi.org/10.1016/S0140-6736(12)61684-5)
- Myers, S. S., Gaffikin, L., Golden, C. D., Ostfeld, R. S., Redford, K. H., Ricketts, T. H., Turner, W. R., & Osofsky, S. A. (2013). Human health impacts of ecosystem alteration. *Proceedings of the National Academy of Sciences*, 110(47), 18753–18760.
<https://doi.org/10.1073/pnas.1218656110>
- Oleksy, R., Racey, P. A., & Jones, G. (2015). High-resolution GPS tracking reveals habitat selection and the potential for long-distance seed dispersal by Madagascan flying foxes *Pteropus rufus*. *Global Ecology and Conservation*, 3, 678–692.
<https://doi.org/10.1016/j.gecco.2015.02.012>
- Olivero, J., Fa, J. E., Real, R., Márquez, A. L., Farfán, M. A., Vargas, J. M., Gaveau, D., Salim, M. A., Park, D., Suter, J., King, S., Leendertz, S. A., Sheil, D., & Nasi, R. (2017). Recent loss of closed forests is associated with Ebola virus disease outbreaks. *Scientific Reports*, 7(1), 14291. <https://doi.org/10.1038/s41598-017-14727-9>
- Plowright, R. K., Reaser, J. K., Locke, H., Woodley, S. J., Patz, J. A., Becker, D. J., Oppler, G., Hudson, P. J., & Tabor, G. M. (2021). Land use-induced spillover: A call to action to safeguard environmental, animal, and human health. *The Lancet Planetary Health*, 5(4), e237–e245. [https://doi.org/10.1016/S2542-5196\(21\)00031-0](https://doi.org/10.1016/S2542-5196(21)00031-0)
- Rival, A., & Levang, P. (2014). *Palms of controversies: Oil palm and development challenges*. Center for International Forestry Research (CIFOR).
<https://doi.org/10.17528/cifor/004860>

- Rulli, M. C., Santini, M., Hayman, D. T. S., & D’Odorico, P. (2017). The nexus between forest fragmentation in Africa and Ebola virus disease outbreaks. *Scientific Reports*, 7, 41613.
<https://doi.org/10.1038/srep41613>
- Shafie, N. J., Sah, S. A. M., Latip, N. S. A., Azman, N. M., & Khairuddin, N. L. (2011). Diversity pattern of bats at two contrasting habitat types along Kerian River, Perak, Malaysia. *Tropical Life Sciences Research*, 22(2), 13–22.
- Towner, J. S., Amman, B. R., Sealy, T. K., Carroll, S. A. R., Comer, J. A., Kemp, A., Swanepoel, R., Paddock, C. D., Balinandi, S., Khristova, M. L., Formenty, P. B. H., Albarino, C. G., Miller, D. M., Reed, Z. D., Kayiwa, J. T., Mills, J. N., Cannon, D. L., Greer, P. W., Byaruhanga, E., ... Nichol, S. T. (2009). Isolation of genetically diverse Marburg viruses from Egyptian fruit bats. *PLOS Pathogens*, 5(7), e1000536.
<https://doi.org/10.1371/journal.ppat.1000536>
- Vijay, V., Pimm, S. L., Jenkins, C. N., & Smith, S. J. (2016). The impacts of oil palm on recent deforestation and biodiversity loss. *PLOS ONE*, 11(7), e0159668.
<https://doi.org/10.1371/journal.pone.0159668>
- WHO Ebola Response Team. (2016). After Ebola in West Africa — Unpredictable risks, preventable epidemics. *New England Journal of Medicine*, 375(6), 587–596.
<https://doi.org/10.1056/NEJMs1513109>
- Wilcox, B. A., & Ellis, B. (2006). Forests and emerging infectious diseases of humans. *Unasylva*, 57, 11–18.
- World Health Organization. (2017). *One Health: questions and answers*. Retrieved July 14, 2025, from <https://www.who.int/news-room/questions-and-answers/item/one-health>

CHAPTER 2

EXTENDING THE TEMPORAL AND SPATIAL CLASSIFICATION OF OIL PALM PLANTATIONS IN AFRICA THROUGH REMOTE SENSING TECHNOLOGIES

2.1 Introduction

Oil palm (*Elaeis guineensis*) has become one of the most economically significant vegetable oil crops globally, accounting for approximately 35% of total vegetable oil production as of 2022 (PACRA, 2023, citing USDA). Southeast Asia—particularly Indonesia and Malaysia—continues to dominate global output. However, Africa, the species’ endemic region, is witnessing steady growth in production, currently contributing around 4% of global supply (Solidaridad, 2022). This expansion is driven in part by oil palm’s exceptional land-use efficiency: it yields substantially more oil per hectare than alternative oilseed crops such as soybean, rapeseed, or sunflower (Corley & Tinker, 2015). In addition, regional development strategies—such as the African Palm Oil Initiative—have promoted domestic production to reduce import dependency, diversify exports, and support rural livelihoods across Sub-Saharan Africa (World Economic Forum, 2022).

The expansion of oil palm cultivation across Africa has generated substantial socioeconomic benefits, particularly in terms of employment creation and rural income enhancement (Feintrenie, 2012; Nkongho et al., 2014). The sector provides livelihoods not only to farmers but also to numerous stakeholders along the value chain, including operators, transporters, and seed distributors, demonstrating significant potential for rural poverty alleviation (Nkongho et al., 2014; Feintrenie, 2012). In Ghana, for example, oil palm contributes up to 75% of total

household income for farmers engaged in its production, while countries such as Ghana and Nigeria have leveraged oil palm expansion to promote rural development and economic diversification, thereby reducing their dependence on traditional export commodities (Ofosu-Budu & Sarpong, 2013; Nkongho et al., 2014). This economic transformation has positioned oil palm as a key driver of regional development initiatives aimed at strengthening agricultural export capacity and improving rural livelihoods throughout Sub-Saharan Africa.

However, this rapid agricultural expansion has significant environmental and social implications, including deforestation, biodiversity loss, land tenure conflicts, and disruption of ecosystem functions like carbon sequestration and hydrological regulation (Fitzherbert et al., 2008; Carlson et al., 2013; Koh & Wilcove, 2008; Cotula et al., 2009). Additionally, oil palm cultivation in Africa has raised important public health concerns due to its potential role in facilitating zoonotic diseases such as Ebola and Marburg virus spillovers. Fruit bats, identified as critical reservoir hosts for filoviruses, are particularly attracted to monoculture plantations due to their abundant food resources, thereby increasing potential human-wildlife interactions and associated disease risks (Alexander et al., 2015; Wallace et al., 2014).

Accurate spatiotemporal monitoring of oil palm plantation expansion is critical for addressing these ecological and public health challenges effectively. Specifically, understanding precisely when land areas were cleared and initially planted is essential for quantifying ecological impacts, modeling disease risks, and formulating sustainable land-use policies. Until recently, robust and continuous mapping of plantation establishment across Africa has been limited. However, Descals et al. (2024) significantly advanced this field by first using Sentinel-1 radar satellite imagery to classify industrial and smallholder oil palm plantations globally between 2016 and 2020. Subsequently, they applied Landsat imagery (1990–2020) and a single spectral index—

the Normalized Difference Water Index (NDWI)—to retrospectively detect bare soil events, thus identifying plantation establishment years.

Although this represents a major methodological advance, reliance on Landsat data introduces notable limitations. In particular, Landsat-7 imagery suffered a critical mechanical failure in May 2003, when the Scan Line Corrector (SLC) malfunctioned, causing significant data gaps (approximately 22% missing data per scene). These persistent gaps severely compromise temporal consistency and hinder accurate, continuous monitoring of plantation establishment events, especially in fragmented or rapidly changing landscapes characteristic of African agricultural contexts. Furthermore, dependence on a single vegetation index (NDWI) may be insufficient for reliably detecting bare soil in the diverse and heterogeneous landscapes of Africa, where complex seasonal vegetation dynamics, mixed land cover types, and varying soil moisture conditions complicate classification accuracy.

To overcome these methodological challenges, this chapter seeks to replicate and enhance the bare-soil detection approach employed by Descals et al. (2024) by replacing Landsat imagery with Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data. MODIS imagery offers substantial advantages for continuous temporal monitoring due to its daily revisit frequency and absence of data gaps, allowing reliable tracking of bare-soil signals indicative of initial land clearing. This improvement is particularly relevant given the significant temporal gaps resulting from Landsat-7's scan-line corrector malfunction.

In addition to utilizing MODIS data, this research extends the methodological framework by integrating multiple complementary vegetation indices rather than relying solely on NDWI. Specifically, we incorporate the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Soil-Adjusted Vegetation Index (SAVI), alongside basic reflectance

bands (Red, Near-Infrared [NIR], and Blue). These indices each capture distinct aspects of vegetation structure and soil characteristics: NDVI is sensitive to chlorophyll content, EVI better accounts for vegetation canopy structure, and SAVI corrects for soil brightness influences. Leveraging these multiple indices in an ensemble machine-learning classification workflow (Random Forest, CART, SVM, and XGBoost algorithms), our approach significantly improves the robustness and accuracy of bare-soil detection across diverse landscapes.

The resulting methodological enhancement provides more accurate and temporally continuous identification of plantation establishment events (bare-soil clearing) across 17 African countries, complementing the plantation classification maps (industrial vs. smallholder) developed by Descals et al. (2024). By precisely determining plantation establishment timing through MODIS-based bare-soil mapping, this research fills critical temporal data gaps and supports robust ecological, epidemiological, and economic analyses.

Ultimately, this study contributes to existing knowledge by: (1) providing a consistent, gap-free temporal dataset of plantation establishment events essential for detailed environmental and epidemiological impact assessments; (2) demonstrating the advantages of MODIS's continuous temporal coverage and multiple vegetation indices for accurate classification in complex tropical landscapes; and (3) delivering refined annual plantation establishment dates that support policymakers and stakeholders in targeted ecological conservation, agricultural planning, and zoonotic disease mitigation strategies. This enhanced methodological framework represents a significant advancement in remote sensing applications for monitoring perennial crop expansion, with broad applicability across diverse tropical agricultural systems.

2.2 Methods

2.2.1 Study area

This study encompasses 17 African countries selected based on their significant potential for oil palm cultivation and favorable agro-climatic conditions. These countries include Guinea, Sierra Leone, Liberia, Côte d'Ivoire, Ghana, Togo, Benin, Nigeria, Cameroon, Gabon, Republic of the Congo, Democratic Republic of the Congo, Central African Republic, Uganda, Burundi, Tanzania, and Guinea-Bissau (Figure 2.1).

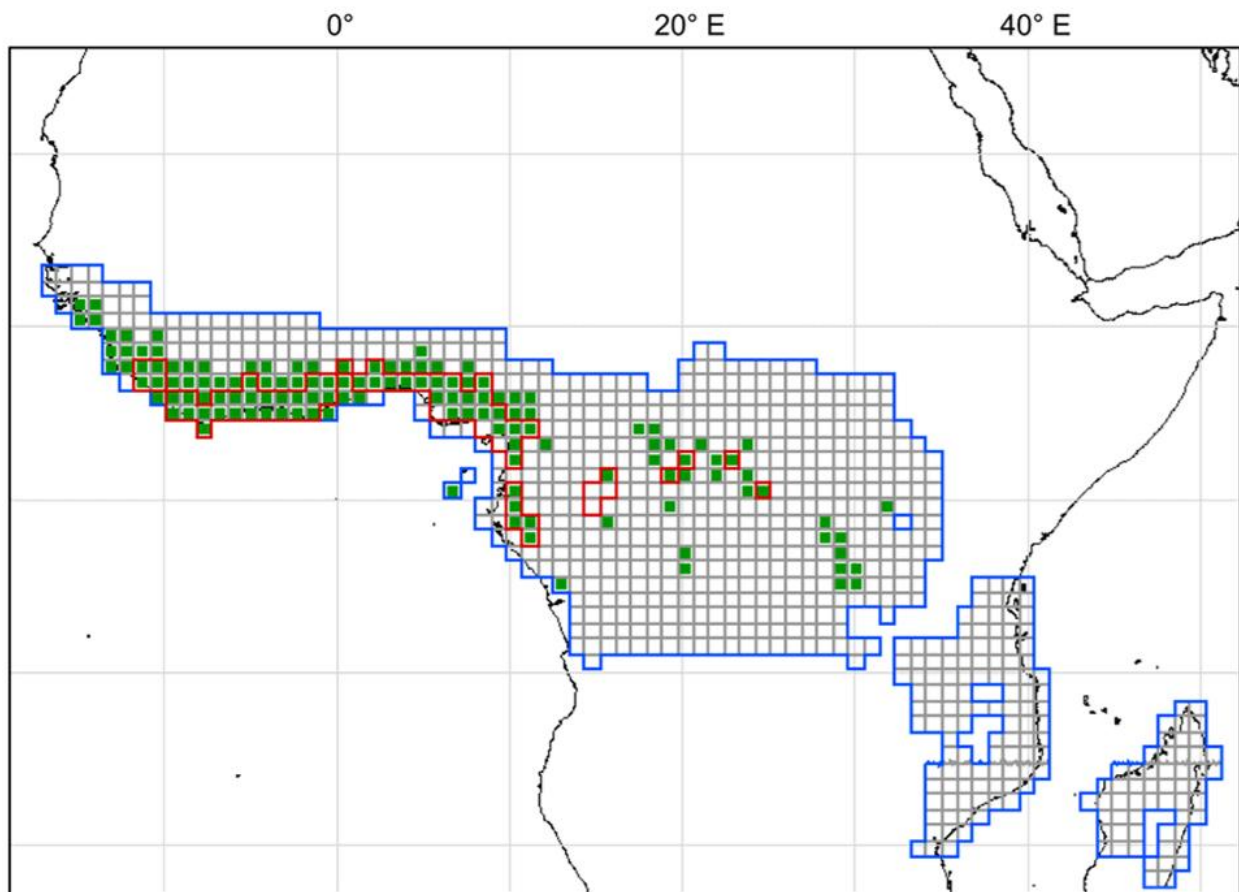


Figure 2.1: Map of the 17 African countries included in this study

Source: The Figure is from Figure 2 in the article of Descals et al., 2021

The selection criteria included suitability of climatic conditions, particularly rainfall patterns and temperature regimes conducive to oil palm growth; existing infrastructure facilitating

agricultural development and export; and active or emerging policy frameworks supportive of oil palm cultivation (Descals et al., 2021).

2.2.2 Data

Remote Sensing Data

This study utilizes MODIS (Moderate Resolution Imaging Spectroradiometer) data, strategically chosen for its reliable spectral information and consistent temporal coverage. Before explaining our specific approach, it's important to understand how satellite remote sensing works for land cover classification.

Satellite remote sensing captures electromagnetic radiation reflected from Earth's surface across different wavelengths of the spectrum. Various land cover types—forests, crops, bare soil, water—reflect light differently across these wavelengths, creating distinctive "spectral signatures" that can be measured and analyzed (Jensen, 2015). For example, healthy vegetation strongly absorbs blue and red light for photosynthesis while reflecting near-infrared light, whereas bare soil typically has higher reflectance in red wavelengths and lower reflectance in near-infrared. These differences in reflectance patterns allow researchers to distinguish between different land cover types using mathematical formulas that emphasize these contrasts (Tucker, 1979; Xue & Su, 2017).

MODIS offers continuous daily global coverage since 2000, making it uniquely suited for time-series analysis of land cover dynamics, particularly for tracking bare soil signals that indicate oil palm plantation establishment across African landscapes (Justice et al., 2002). Unlike Landsat, which experiences data gaps from the scan line corrector failure and has a 16-day revisit period that can be further compromised by cloud cover in tropical regions, MODIS provides consistent observations with minimal gaps (Roy et al., 2008).

The MODIS platform provides several key advantages for this research. With a spatial resolution of 250 meters, it offers an optimal balance between coverage area and detail, allowing for continental-scale analysis while still capturing significant land-use changes across the 17 African countries in our study area. While this resolution is coarser than Landsat's 30 meters, MODIS's higher temporal frequency compensates by providing more opportunities to capture cloud-free observations in tropical regions where persistent cloud cover often limits the utility of higher-resolution but less frequent imagery (Whitcraft et al., 2015).

The innovative aspect of our approach lies in the comprehensive multi-index strategy that addresses the limitations of single-index methods previously applied in African contexts. To understand this advantage, it's helpful to explain what vegetation indices are and how they're used in remote sensing. Vegetation indices are mathematical combinations of different spectral bands designed to enhance the signal of vegetation properties while minimizing background effects like soil brightness, atmospheric conditions, or viewing angle (Huete et al., 2002). Different indices are sensitive to different aspects of vegetation, such as chlorophyll content, canopy structure, or water stress.

Rather than depending on a single spectral index that might inadequately capture the complex land-use patterns in African landscapes, this study incorporates five key spectral indices and reflectance measurements to enhance classification accuracy. We selected these five MODIS indices and bands because they provide complementary information about land cover characteristics:

These spectral measurements include direct surface reflectance in three crucial bands: Red (sur_refl_b01), Near Infrared or NIR (sur_refl_b02), and Blue (sur_refl_b03). These raw reflectance values are particularly valuable for bare soil detection as exposed soils typically exhibit

distinctive reflectance patterns compared to vegetated areas (Barnes et al., 2003). Specifically, bare soils generally show higher reflectance in the red band and lower reflectance in the near-infrared band compared to vegetation. The blue band provides additional information useful for distinguishing between different soil types and conditions, as well as for atmospheric correction.

Additionally, the study utilizes two derived vegetation indices: the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). NDVI is calculated as $(\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})$ and is highly sensitive to the presence and density of green vegetation (Rouse et al., 1974). It capitalizes on the contrast between strong NIR reflectance and red light absorption by chlorophyll in healthy vegetation. EVI uses a more complex formula: $2.5 \times [(\text{NIR} - \text{Red})/(\text{NIR} + 6 \times \text{Red} - 7.5 \times \text{Blue} + 1)]$. By incorporating the blue band and correction coefficients, EVI reduces atmospheric influences and better captures vegetation variation in high-biomass regions where NDVI might saturate (Huete et al., 2002). The analysis incorporates these key spectral indices and reflectance measurements as illustrated in Table 2.1, which shows the five MODIS indices used in our analysis.

NDVI is widely used to detect rapid decreases in vegetation cover associated with land clearing, making it a valuable tool for monitoring deforestation and agricultural expansion. In contrast, EVI offers improved sensitivity to canopy structure and is less affected by atmospheric conditions, allowing for more accurate detection of subtle changes in vegetation (Huete et al., 2002). By combining both indices with direct reflectance measurements, researchers can more reliably distinguish between truly bare soils and areas with sparse or senescent vegetation (Jönsson & Eklundh, 2004).

Furthermore, MODIS's higher temporal resolution enables more precise detection of the often brief window when land is cleared for plantation establishment, a critical advantage for

accurate estimation of planting years. By capturing these temporally narrow bare soil signals consistently across all study years (2000-2020), our method identifies crucial land-use transition events that mark the beginning of oil palm cultivation (Verbesselt et al., 2010).

Table 2.1: Spectral Indices and Reflectance Measurements Used in Classification

Abbreviation	Spectral Index/band	Formula
sur_refl_b01	Red surface reflectance	Direct measurement
sur_refl_b02	NIR surface reflectance	Direct measurement
sur_refl_b03	Blue surface reflectance	Direct measurement
NDVI	Normalized Difference Vegetation Index	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$
EVI	Enhanced Vegetation Index	$2.5 \times (\text{NIR} - \text{RED}) / (\text{NIR} + 6 \times \text{RED} - 7.5 \times \text{BLUE} + 1)$

Auxiliary Data

Auxiliary datasets play a critical role in this study by providing complementary information that enhances the reliability of bare soil classification. The primary auxiliary dataset employed is the Global Land Use/Land Cover with Sentinel-2 (10 m) product developed by Zanaga et al. (2021), which offers high-resolution categorical information on bare/sparse vegetation and broader land cover types. This dataset serves as an independent reference layer that helps distinguish between temporary bare soil conditions and persistently sparse landscapes.

The Sentinel-2 derived land cover information is particularly valuable for addressing potential misclassifications in arid or seasonally dry regions where natural land cover might spectrally resemble bare soil. By incorporating this auxiliary data, the classification model gains additional contextual information about regional vegetation characteristics, thereby reducing commission errors in bare soil identification. The land cover categories consulted include tree cover, cropland, shrubland, grassland, and other relevant classifications as shown in Table 2.2.

Furthermore, the 10-meter resolution of this auxiliary dataset provides fine-grained land cover details that complement the coarser MODIS imagery, allowing for a more nuanced understanding of sub-pixel land cover heterogeneity within each MODIS pixel. While this multi-scale approach might introduce some noise in the classification process, it avoids the systematic data gaps associated with scan line errors that affect other satellite systems. The continuous and complete coverage of MODIS data ensures that every pixel in the study area has usable information for every time step in the 2000-2020 analysis period, a significant advantage for time-series analysis of land cover transitions.

Sample Point Data

The development of a comprehensive and representative sample dataset forms the foundation of the supervised classification approach employed in this study. Initially, 200,000 random points were generated across the vast geographic extent of the 17 African countries. Following a meticulous screening process to remove duplicates and invalid pixels within the MODIS grid system, approximately 189,247 points remained for subsequent analyses. This extensive dataset ensures broad spatial coverage and captures the diverse ecological conditions present across the study area.

These 189,247 points were strategically distributed across nine distinct land cover categories as detailed in Table 2.2, with tree cover constituting the largest portion (97,976 points), followed by grassland (33,291 points), shrubland (32,396 points), and cropland (21,908 points). Smaller but crucial categories include built-up areas (627 points), water bodies (1,115 points), herbaceous wetland (918 points), mangroves (331 points), and bare/sparse vegetation (685 points). This stratified distribution ensures adequate representation of major land cover types while maintaining sufficient samples for less common but environmentally significant categories.

Table 2.2 Distribution of Random Points Across Land Cover Categories

Land Cover Category	Number of Points	Percentage (%)
Tree cover	97,976	51.77
Shrubland	32,396	17.12
Grassland	33,291	17.59
Cropland	21,908	11.58
Built-up	627	0.33
Water	1,115	0.59
Herbaceous wetland	918	0.49
Mangroves	331	0.17
Bare / sparse vegetation	685	0.36
Total	189,247	100

To address the potential classification challenges posed by the relatively limited number of bare/sparse vegetation points (685) compared to other categories, an additional targeted subsampling was performed. Specifically, 600 points were randomly selected from the bare/sparse vegetation category and balanced with 600 points from other land covers. This balanced sampling approach significantly improves the classifier's ability to differentiate bare soil from other land cover types, which is essential for the subsequent identification of oil palm plantation establishment.

For each of these 1,200 balanced sample points, five critical MODIS indices (NDVI, EVI, Red, NIR, Blue) (Table 2.2) were extracted for the year 2000 via Google Earth Engine, establishing a robust spectral signature baseline for the beginning of the study period. The consistent temporal

coverage of MODIS ensures that spectral information is available for every sample point without the data gaps that would occur with systems affected by scan line errors. This complete temporal record is particularly valuable for tracking the transition from forest to bare soil to oil palm plantation, as it allows for the precise identification of when land clearing occurs—a critical indicator of plantation establishment.

The conclusion of this data preparation phase, as illustrated in Figure 2.2, establishes a solid foundation for the subsequent classification of bare soil and non-bare soil areas across the entire time series, enabling the tracking of oil palm expansion throughout the study region from 2000 to 2020.

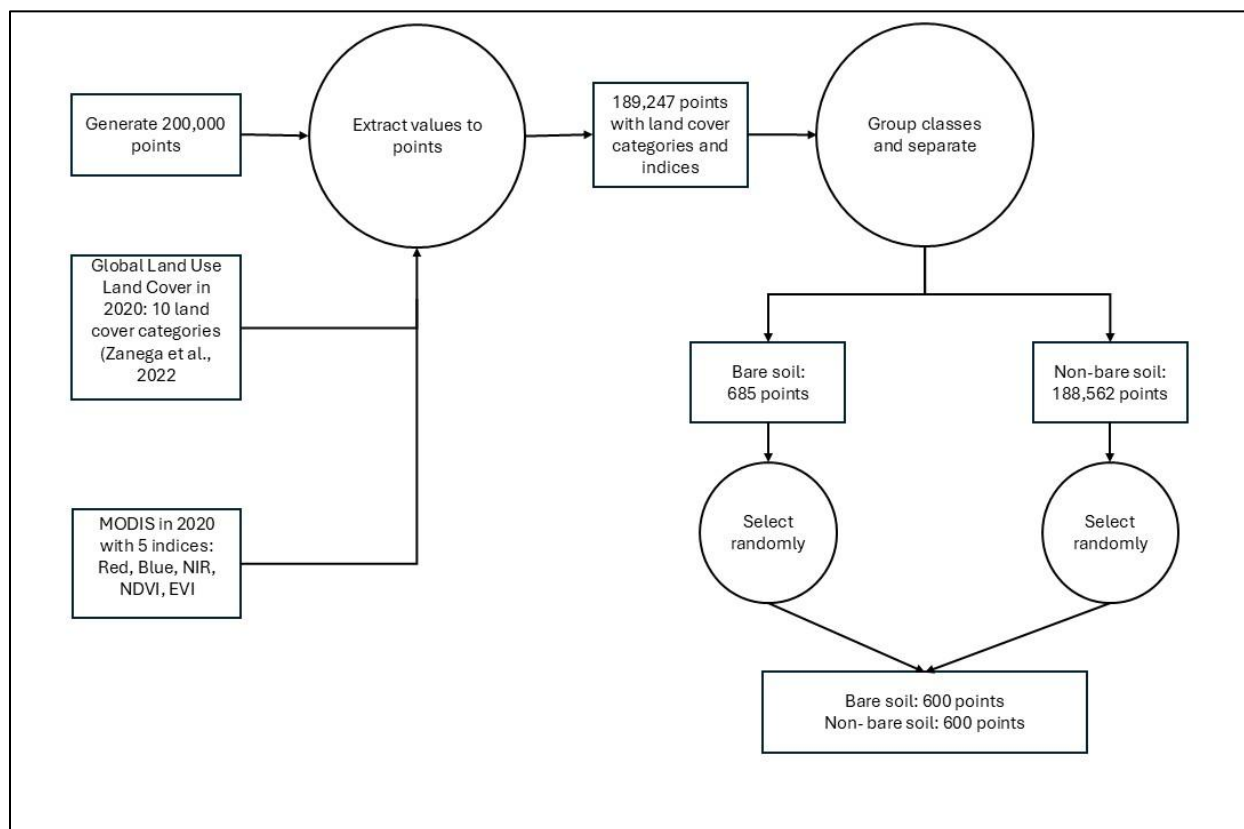


Figure 2.2 Workflow for the data preparation

2.2.3 Preprocessing

A critical step in optimizing machine learning algorithms involves the creation of well-curated training and validation datasets. This preprocessing phase is essential for ensuring that the classification models can effectively learn the spectral characteristics that distinguish bare soil from other land cover types across diverse African landscapes.

We followed the common 80:20 train–test split, a widely adopted practice in remote-sensing machine-learning studies that balances model training efficiency with robust out-of-sample evaluation. For example, Gómez et al. (2016) apply this partition when training a MODIS-based Random Forest model for burned-area mapping, reporting stable classification accuracies across multiple biomes. A similar approach is used by Vasilakos et al. (2020) in their ensemble classification of multitemporal Sentinel-2 imagery, and by Khan et al. (2023), who implement an 80% training and 20% testing division for a multi-branch deep learning framework applied to land scene classification. These studies highlight that using a smaller test set can inflate accuracy variance, while larger test sets can undermine the learning capacity of the model due to reduced training data. Guided by this evidence, we allocate 960 points (80%) to model training—sufficient to capture key spectral patterns—while reserving 240 points (20%) for independent testing to ensure a statistically meaningful accuracy assessment.

The training subset ($n=960$) is used to develop and calibrate the classification algorithms, enabling them to recognize the distinctive spectral signatures associated with bare soil conditions. Special attention is given to reducing potential biases in the training data by ensuring representative coverage across the study area's diverse ecological zones. This geographic

stratification is particularly important in the African context, where soil types, vegetation dynamics, and climatic conditions vary considerably across the 17 countries included in the study. For the testing subset ($n=240$), care is taken to maintain the same class balance as the training data (50% bare soil, 50% non-bare soil) while ensuring these points remain completely independent from the training process. This independence is crucial for obtaining unbiased estimates of classification accuracy and model performance. By reserving a significant portion of data (20%) for testing, the study can accurately assess how well the classification models generalize to unseen pixels in various parts of the study area.

The spectral information for each sample point includes the five MODIS-derived indices (Red, NIR, Blue, NDVI, and EVI) that form the foundation of the classification approach. During preprocessing, these spectral values undergo normalization to standardize their ranges, reducing the influence of extreme values and improving model stability. This normalization process is particularly important when working with multiple indices that operate on different scales.

Additionally, temporal consistency checks are performed to identify and address any anomalies in the time series data, such as missing values or artifacts from atmospheric interference. While MODIS data provides relatively consistent coverage, occasional cloud contamination can affect spectral readings. The preprocessing workflow includes cloud masking procedures and, where necessary, temporal interpolation to fill short gaps in the data record.

2.2.4 Classification

To map oil-palm and non-oil-palm classes across Africa's heterogeneous agro-ecological zones we compare four supervised machine-learning algorithms that are widely used in satellite remote-sensing studies, each offering a distinct bias–variance trade-off.

Random Forest (RF). RF is our primary classifier because its bootstrap aggregation of hundreds of decision trees copes well with the high dimensionality and multicollinearity typical of multispectral data (Breiman, 2001). Numerous land-cover studies report RF outperforms single-tree and parametric models while being almost immune to over-fitting (Belgiu & Drăguț, 2016; Gislason et al., 2006). The algorithm also ranks variable importance, helping identify which indices (NDVI, EVI, etc.) best discriminate bare soil from early-stage oil-palm canopy.

Classification and Regression Trees (CART).

Although less accurate than ensemble methods, CART's rule-based tree makes spectral thresholds transparent—an advantage when results must be communicated to local planners (Breiman et al., 1984). Its binary splits handle continuous and categorical bands without pre-scaling, and previous African studies show that CART can still achieve >80 % overall accuracy with modest training sets (Forkuor et al., 2019).

Support Vector Machine (SVM).

SVM is effective in high-dimensional feature spaces and excels when class boundaries are narrow or overlapping (Pal & Mather, 2005). Kernel functions let the hyper-plane warp around mixed pixels—useful where bare soil is interspersed with sparse ground-cover or crop residues. Meta-analyses report SVM competitive with RF on medium-resolution sensors, especially when training data are limited (Mountrakis et al., 2011).

eXtreme Gradient Boosting (XGBoost).

XGBoost combines gradient-boosted trees with regularization that tempers over-fitting and speeds computation (Chen & Guestrin, 2016). Recent remote-sensing applications show it rivals or surpasses RF for cropland mapping while handling missing observations from cloud-

contaminated time series (Zhong et al., 2019). Its sequential learning is well suited to capturing subtle soil-texture differences that vary across Africa’s savanna and forest belts.

All four algorithms are trained on identical feature stacks (Red, NIR, Blue, NDVI, EVI) and the same 80 % stratified sample, then evaluated on the withheld 20 %. This head-to-head design isolates algorithmic performance and reveals which spectral cues most reliably flag bare soil under diverse biophysical conditions.

2.2.5 Accuracy Assessment and Method Selection

We evaluated the four candidate algorithms with an independent 20% hold-out set (240 points) using standard categorical-map diagnostics recommended by Stehman and Foody (2019): overall accuracy (OA), producer's accuracy (PA), and user's accuracy (UA). OA—the share of all validation pixels that are labelled correctly—gives a headline figure but can mask class-imbalance effects, so we report PA and UA for the bare-soil class separately. PA measures omission error (how many true bare-soil pixels are missed); UA measures commission error (how many predicted bare-soil pixels are false alarms).

Table 2.3 Results of the classifications

		Random Forest	CART	XGBoost	SVM
OA (%)		79.17	74.58	79.58	79.58
	Bare soils	75.00	72.52	75.91	76.69
UA (%)	Others	85.00	77.06	84.47	83.18
	Bare soils	87.50	79.17	86.67	85.00
PA (%)	Others	70.83	70.00	72.50	74.17

XGBoost (XGB) achieves the highest overall accuracy (84.17%) and demonstrates excellent precision for bare soil detection with the highest UA (90.91%), meaning very few false bare-soil detections—crucial when bare soil is used as a proxy for new oil-palm clearing. Random Forest (RF) shows the most balanced performance across classes with strong detection of non-bare areas (PA = 91.67%). While CART achieves the highest sensitivity for bare soil detection (PA = 81.25%), its lower precision (UA = 76.47%) results in more false alarms. SVM provides consistent but not exceptional performance across all metrics.

Table 2.4 Cross-validation results for bare soil classification algorithms using MODIS data

	Random Forest	CART	XGBoost	SVM
OA (%)	79.17	74.58	79.58	79.58
5-Fold OA (%)	80.17 (± 2.15)	78.42 (± 2.73)	80.25 (± 2.12)	79.58 (± 1.26)

We checked that the performance differences were not an artifact of the particular split by running stratified 5-fold cross-validation on the 960 training points. XGBoost shows the highest cross-validated accuracy ($80.25\% \pm 2.12\%$) with excellent stability, followed closely by Random Forest ($80.17\% \pm 2.15\%$). SVM demonstrates the most consistent performance with the lowest standard deviation ($\pm 1.26\%$), while CART shows the highest variability across folds ($\pm 2.73\%$). Given XGBoost's superior accuracy, robust generalization, and computational efficiency for large-scale mapping tasks, it is retained as the study's primary classifier for annual bare-soil mapping and subsequent oil-palm expansion analysis.

2.2.6 Estimation of the planting year

Following the model selection phase, the XGBoost classifier—identified as the top performer with an overall accuracy of 80.25% and excellent cross-validation stability—was employed to generate annual bare-soil classifications across the entire study region from 2000 to

2020. This temporal sequence of bare soil maps serves as the foundation for estimating oil palm plantation establishment years throughout the 17 African countries.

The principle underlying this approach is that land clearing for new oil palm plantations manifests as bare or sparse vegetation for a discrete temporal window in satellite imagery. Unlike annual crops that may exhibit seasonal bare soil patterns, oil palm establishment follows a distinct trajectory: forest or other vegetation is cleared, creating a bare soil signature that persists for a relatively short period (typically 3-12 months) before young palms begin to establish vegetative cover. By pinpointing the first year when a pixel transitions to bare soil and subsequently maintains vegetation cover characteristic of oil palm, the study approximates the planting year with reasonable accuracy.

To implement this concept, a pixel-by-pixel temporal analysis was conducted across the 21-year MODIS time series. For each pixel identified as oil palm in the reference dataset from Descals et al. (2024), the complete temporal sequence of annual bare soil classifications was extracted. The appearance of bare soil in this sequence indicates potential land clearing activity, while the timing of this appearance provides a critical temporal marker for estimating plantation establishment.

The algorithm specifically searches for persistent bare soil signals that indicate systematic land clearing rather than ephemeral changes due to seasonal factors, agricultural rotation, or classification errors. This persistence criterion helps distinguish genuine plantation establishment from other land-use dynamics that might temporarily create bare soil conditions. In cases where multiple bare soil periods are detected over the 21 years, the most recent occurrence before continuous vegetation cover establishment is selected as the most likely planting year indicator.

The application of this methodology yields a comprehensive dataset that documents the year-by-year expansion of oil palm across the study region, providing unprecedented temporal detail on this important land-use change process in Africa. By combining the spatial extent of oil palm from high-resolution mapping with the temporal precision of annual MODIS-based bare soil detection, this approach achieves a synergy that overcomes the limitations of previous studies that provided only static snapshots of plantation distribution.

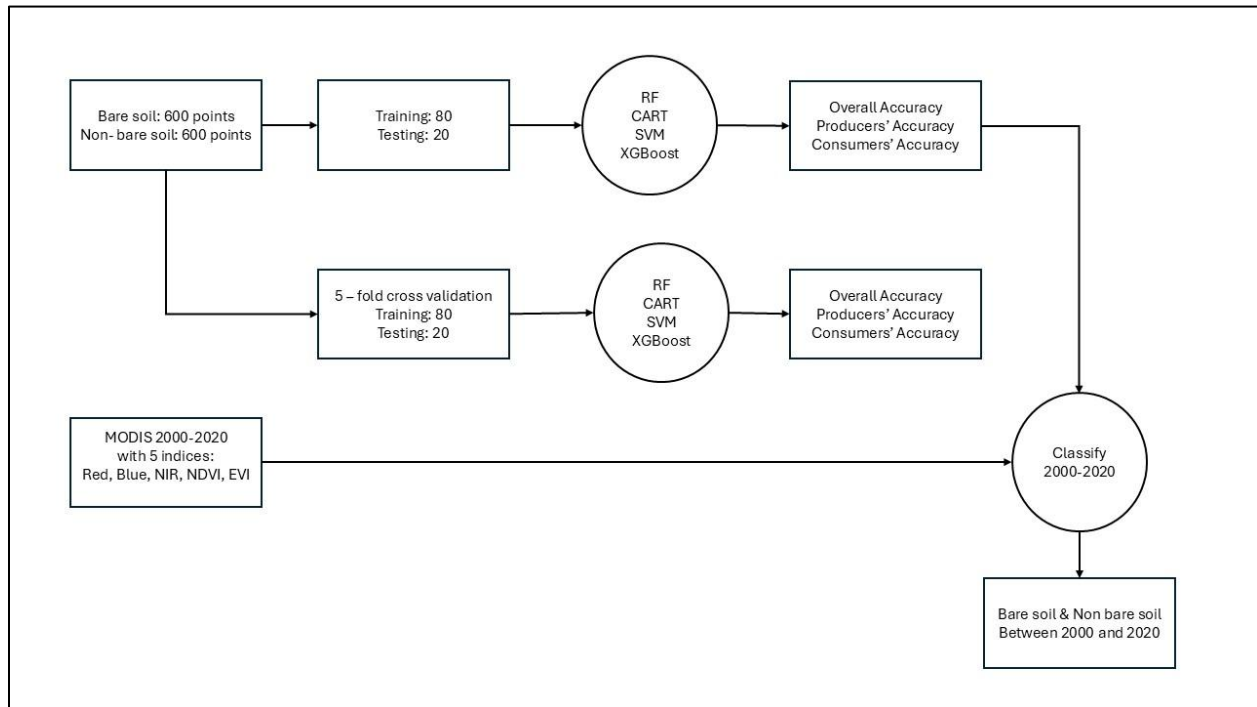


Figure 2.3: Classification Workflow for Bare Soil Detection from 2000-2020

2.3 Results

2.3.1 African oil palm extent

A central objective of this study is to estimate when individual plots transition into oil palm plantations—whether industrial or smallholder—across the 17 African countries under

investigation. The methodology leverages the temporal sequence of bare soil classifications derived from MODIS data to identify likely planting years for oil palm across the continent.

Figure 2.4 provides a schematic visualization illustrating the analytical approach for determining plantation establishment years. In this example, bare soil is detected between 2009 and 2011, indicated by the yellow blocks in the timeline. Outside these intervals, the pixel is classified as either forest (green blocks before 2009) or an existing oil palm plantation (green blocks after 2011). This distinctive temporal signature—forest followed by bare soil followed by plantation—creates a recognizable pattern that can be systematically identified across millions of pixels throughout the study area.

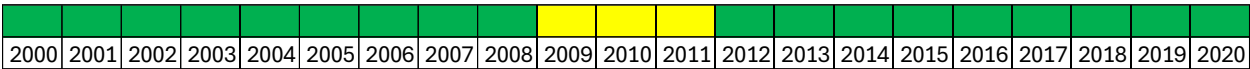


Figure 2.4: Timeline of Land Cover Transitions

Under the proposed logic, the latest instance of bare soil detection (in this case, 2009-2011) is treated as the operative period of plot establishment. This approach recognizes that land clearing for oil palm typically occurs as a discrete event, after which the plantation remains in place for decades. The years following 2011 are therefore classified as oil palm plantation, while the period before 2009 represents the pre-conversion land cover, predominantly forest in this case.

The detection of this temporal transition pattern is particularly important for distinguishing newly established plantations from existing agricultural lands or naturally sparse vegetation. By focusing specifically on the forest-to-bare soil-to-plantation sequence, the methodology effectively isolates oil palm expansion from other land use dynamics that might create temporary bare soil conditions, such as annual crop rotations or seasonal vegetation changes.

Figure 2.5 illustrates the algorithmic workflow implemented to determine planting years for each pixel. The process begins by checking whether the pixel is identified as oil palm in the

reference dataset from Descals et al. (2024). If so, the algorithm initiates a backward-looking temporal analysis starting from 2020. For each year, it examines whether the pixel was classified as bare soil. When bare soil is detected, the subsequent year is designated as the plantation starting year. If the analysis reaches 2000 without detecting bare soil, the pixel is classified as established before the study period.

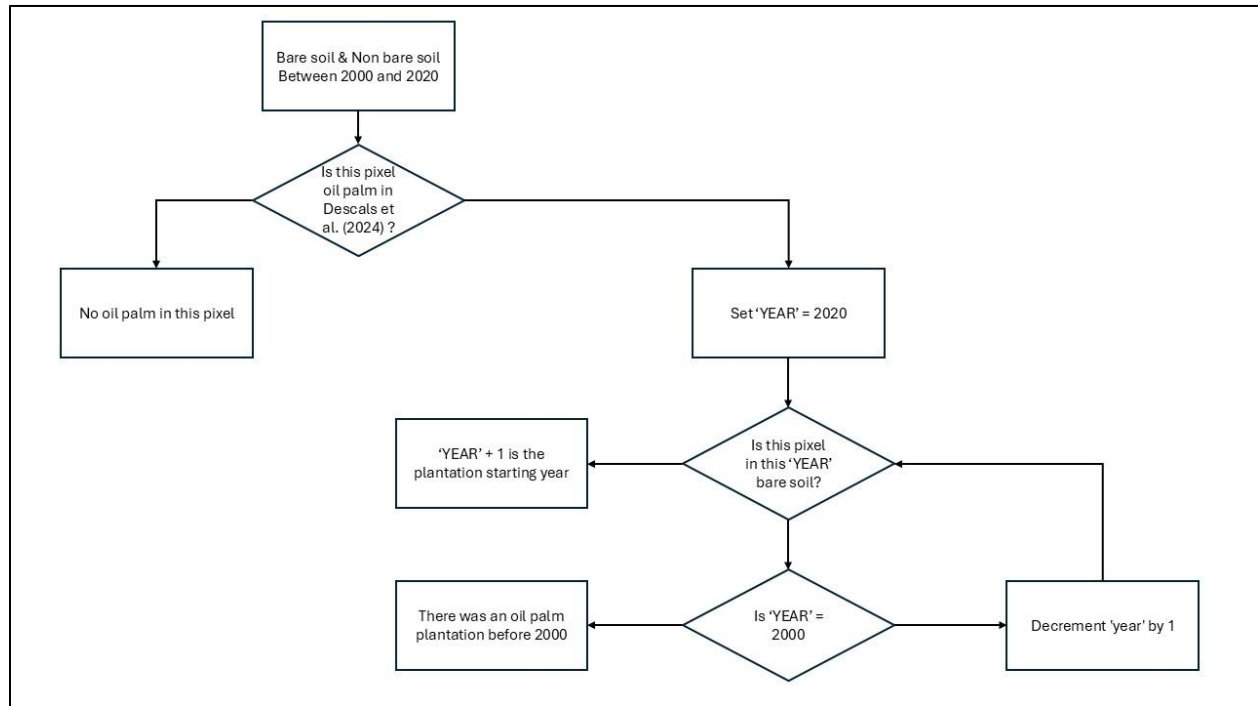


Figure 2.5: Decision Tree for Plantation Year Estimation

This systematic decision tree approach was applied consistently across the entire MODIS time series (2000–2020) for all pixels identified as oil palm in the reference dataset. By overlaying the temporal bare soil classifications with the spatial delineation of industrial and smallholder oil palm footprints, the study systematically identifies and maps the planting years for both plantation types across the 17 countries. The resulting chronological dataset provides unprecedented insight into the spatial and temporal dynamics of oil palm expansion throughout Africa during the 21st century.

The color-coded timeline visualization in Figure 2.3 and the logical flow diagram in Figure 2.4 together demonstrate how this analytical approach translates complex time-series data into interpretable information about land use history. By identifying these temporal signatures across the landscape, the study creates a comprehensive historical record of when and where oil palm has expanded across Africa.

2.3.2 Validation

Figures 2.6 and 2.7 juxtapose our MODIS-based annual estimates of oil-palm area with the series reported by Descals et al. (2024), distinguishing smallholder and industrial plantations from 2000 to 2020. Across both production systems the two datasets portray a similar, monotonic expansion, reinforcing confidence in the temporal pattern recovered by the bare-soil chronology. For smallholder plantings (Figure 2.6) the two curves begin at just under 100 000 ha in 2000 and rise steadily over the study period. Between 2000 and 2005 our estimates exceed those of Descals et al. by roughly 10–15 per cent; the lines converge during 2006–2012 and then diverge again, with our series exhibiting a steeper ascent after 2013. By 2020 the difference between the two totals is approximately five percentage points ($\approx 540,000$ ha versus $\approx 520,000$ ha).

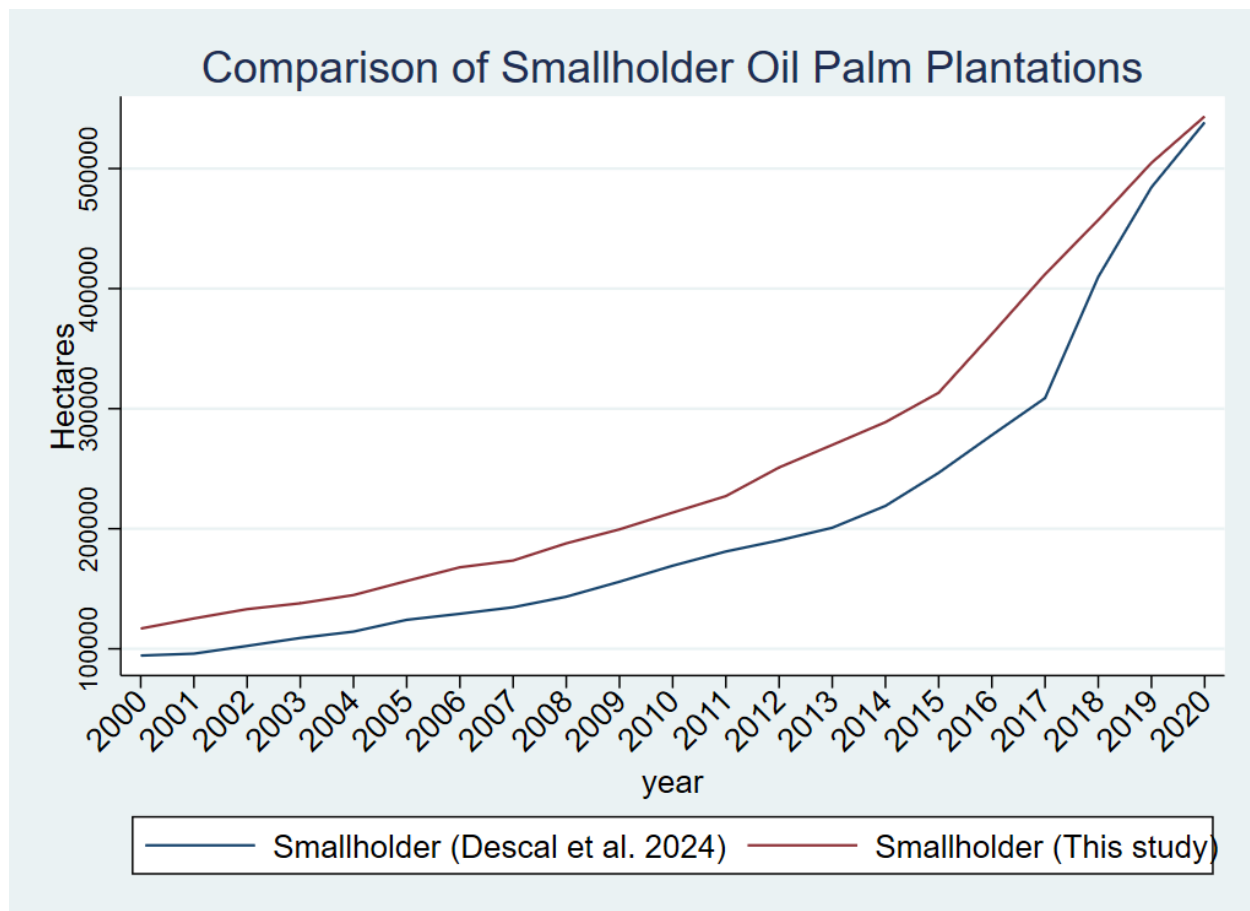


Figure 2.6: Comparison of Smallholder Oil Palm Plantations

For industrial farms (Figure 2.7) both time series show gradual growth through the early 2000s. From 2006 onward our estimates increase more sharply, culminating in about 410,000 ha in 2020, compared with roughly 390,000 ha in Descals et al. Although the gradients differ, the absolute gap never exceeds 25,000 ha in any year, and the two trajectories remain within the same order of magnitude throughout.

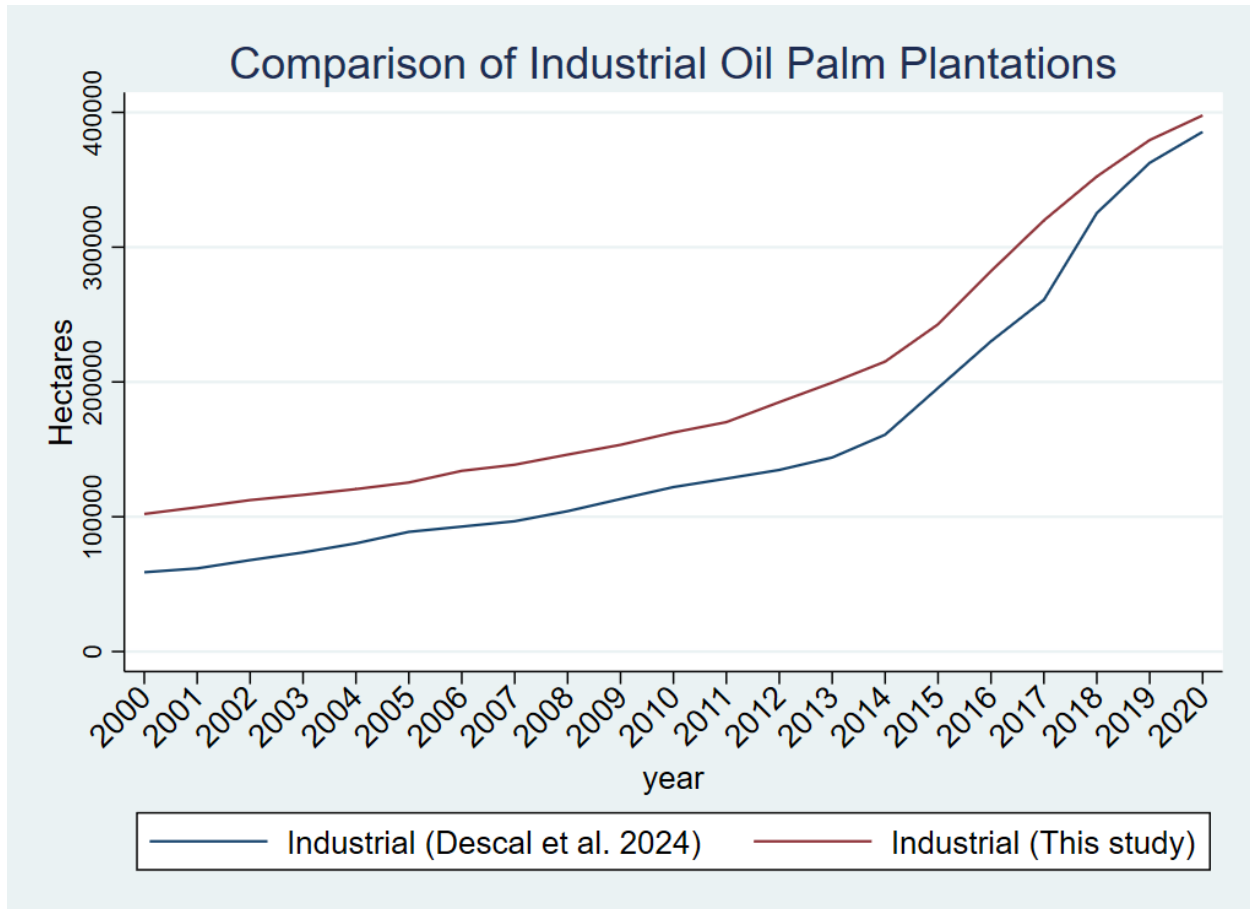


Figure 2.7: Comparison of Industrial Oil Palm Plantations

2.4 Discussion

This study provides a detailed temporal and spatial analysis of oil palm plantation expansion across 17 African countries using MODIS data and advanced classification techniques. The employment of XGBoost as the primary classifier proved particularly effective in accurately detecting bare soil transitions associated with plantation establishment. A comprehensive assessment of the model's performance indicated its strength in managing high-dimensional, complex satellite imagery data, especially important given the heterogeneous land cover types prevalent in the study region.

The observed expansion of smallholder plantations, particularly noticeable from around 2013 onward, reveals critical insights into agricultural patterns and rural development dynamics. The accelerated growth during this period likely reflects both favorable regional agricultural policies and the incremental nature of smallholder cultivation practices. The nuanced temporal resolution of MODIS imagery enabled capturing these subtle yet crucial annual changes, previously difficult to detect using coarser or temporally limited datasets.

Our validation against Descals et al. (2024) further confirms the robustness of our approach, but also highlights differences arising from methodological choices, particularly in terms of capturing ephemeral land-use transitions. While Descals et al. relied on higher spatial resolution imagery, their less frequent observations potentially overlook smaller or fragmented clearings typically associated with smallholders. Our methodology thus complements existing research by providing improved temporal continuity, filling important gaps in historical plantation mapping.

Nevertheless, several limitations of our approach must be acknowledged. The 250-meter resolution of MODIS imagery introduces inherent mixed-pixel effects, complicating accurate detection in highly fragmented landscapes. Future research should integrate higher spatial resolution imagery, such as Sentinel-1 and Sentinel-2, to address these scale issues. Additionally, expanding the scope to incorporate advanced deep learning methods, which have shown promising results in capturing intricate land-cover dynamics, could significantly refine plantation establishment estimations.

2.5 Conclusions

This study presents a robust methodological framework for continuous, large-scale monitoring of oil palm plantation expansion in Africa, utilizing MODIS satellite imagery and advanced machine-learning techniques. Over the two-decade period analyzed (2000–2020), substantial growth in oil palm plantations, particularly among smallholders, highlights significant shifts in agricultural land use with profound implications for rural development, ecosystem management, and public health policies.

The application of the XGBoost classifier significantly enhanced the accuracy and reliability of plantation establishment detection, demonstrating superior performance in complex, multi-dimensional classification tasks. By providing a consistent and detailed temporal dataset, this research contributes valuable insights for policymakers and stakeholders aiming to balance agricultural growth with ecological sustainability and public health concerns.

Future work should prioritize integrating higher-resolution imagery and innovative classification methods to further improve mapping precision. Doing so will facilitate more targeted and informed agricultural, conservation, and public health strategies, ensuring sustainable land management practices in the rapidly evolving African landscapes.

2.6 References

- Alexander, K. A., Sanderson, C. E., Marathe, M., Lewis, B. L., Rivers, C. M., Shaman, J., Drake, J. M., Lofgren, E., Dato, V. M., Eisenberg, M. C., & Eubank, S. (2015). What factors might have led to the emergence of Ebola in West Africa? *PLOS Neglected Tropical Diseases*, 9(6), e0003652. <https://doi.org/10.1371/journal.pntd.0003652>
- Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R. J., Li, H., & Moran, M. S. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In P. C. Robert, R. H. Rust, & W. E. Larson (Eds.), *Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, USA, 16–19 July 2000* (pp. 1–15). American Society of Agronomy.
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315139470>
- Carlson, K. M., Curran, L. M., Asner, G. P., Pittman, A. M., Trigg, S. N., & Adeney, J. M. (2013). Carbon emissions from forest conversion by Kalimantan oil palm plantations. *Nature Climate Change*, 3(3), 283–287. <https://doi.org/10.1038/nclimate1702>

- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery.
<https://doi.org/10.1145/2939672.2939785>
- Corley, R. H. V., & Tinker, P. B. (2015). *The oil palm* (5th ed.). Wiley-Blackwell.
<https://doi.org/10.1002/9781118953297>
- Cotula, L., Vermeulen, S., Leonard, R., & Keeley, J. (2009). *Land grab or development opportunity? Agricultural investment and international land deals in Africa*. International Institute for Environment and Development (IIED). <https://www.iied.org/12561iied>
- Descals, A., Gaveau, D. L. A., Wich, S., Szantoi, Z., & Meijaard, E. (2024). Global mapping of oil palm planting year from 1990 to 2021. *Earth System Science Data*, 16(5), 5111–5129.
<https://doi.org/10.5194/essd-16-5111-2024>
- Descals, A., Wich, S., Meijaard, E., Gaveau, D. L. A., Peedell, S., & Szantoi, Z. (2021). High-resolution global map of smallholder and industrial closed-canopy oil palm plantations. *Earth System Science Data*, 13(3), 1211–1231. <https://doi.org/10.5194/essd-13-1211-2021>
- Feintrenie, L. (2012, April 23–26). *Transfer of the Asian model of oil palm development: From Indonesia to Cameroon*. Paper presented at the Annual World Bank Conference on Land and Poverty: Land governance in a rapidly changing environment, Washington, DC, United States.
- Fitzherbert, E. B., Struebig, M. J., Morel, A., Danielsen, F., Brühl, C. A., Donald, P. F., & Phalan, B. (2008). How will oil palm expansion affect biodiversity? *Trends in Ecology & Evolution*, 23(10), 538–545. <https://doi.org/10.1016/j.tree.2008.06.012>

- Forkuor, G., Dimobe, K., Serme, I., & Tondoh, J. E. (2017). Landsat-8 vs. Sentinel-2: Examining the added value of Sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GIScience & Remote Sensing*, 55(3), 331–354.
<https://doi.org/10.1080/15481603.2017.1370169>
- Gaveau, D. L. A., Locatelli, B., Salim, M. A., Husnayaen, Manurung, T., Descals, A., Angelsen, A., Meijaard, E., & Sheil, D. (2022). Slowing deforestation in Indonesia follows declining oil palm expansion and lower oil prices. *PLOS ONE*, 17(3), e0266178.
<https://doi.org/10.1371/journal.pone.0266178>
- Gislason, P. Ó., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, 27(4), 294–300.
<https://doi.org/10.1016/j.patrec.2005.08.011>
- Gómez, C., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, 55–72. <https://doi.org/10.1016/j.isprsjprs.2016.03.008>
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850–853.
<https://doi.org/10.1126/science.1244693>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)

- Jensen, J. R. (2015). *Introductory digital image processing: A remote sensing perspective* (4th ed.). Pearson.
- Jönsson, P., & Eklundh, L. (2004). TIMESAT—A program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30(8), 833–845.
<https://doi.org/10.1016/j.cageo.2004.05.006>
- Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P., & Morisette, J. T. (2002). An overview of MODIS land data processing and product status. *Remote Sensing of Environment*, 83(1–2), 3–15.
[https://doi.org/10.1016/S0034-4257\(02\)00084-6](https://doi.org/10.1016/S0034-4257(02)00084-6)
- Khan, S. D., & Basalamah, S. (2023). Multi-Branch Deep Learning Framework for Land Scene Classification in Satellite Imagery. *Remote Sensing*, 15(13), 3408.
<https://doi.org/10.3390/rs15133408>
- Koh, L. P., & Wilcove, D. S. (2008). Is oil palm agriculture really destroying tropical biodiversity? *Conservation Letters*, 1(2), 60–64. <https://doi.org/10.1111/j.1755-263X.2008.00011.x>
- Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259.
<https://doi.org/10.1016/j.isprsjprs.2010.11.001>
- Nkongho, R. N., Feintrenie, L., & Levang, P. (2014). Strengths and weaknesses of the smallholder oil palm sector in Cameroon. *OCL*, 21(2), D207.
<https://doi.org/10.1051/ocl/2013043>
- Ofosu-Budu, K., & Sarpong, D. (2013). Oil palm industry growth in Africa: A value chain and smallholders study for Ghana. In A. Elbehri (Ed.), *Rebuilding West Africa's food potential* (pp. 349–389). FAO & IFAD.

PACRA. (2023). *Edible oil sector overview: Pakistan and global trends* (February 2023).

Pakistan Credit Rating Agency. Retrieved from

https://www.pacra.com/view/storage/app/Edible%20Oil%20-%20PACRA%20Research%20-%20Feb%2723_1675958400.pdf

Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing.

International Journal of Remote Sensing, 26(5), 1007–1011.

<https://doi.org/10.1080/01431160512331314083>

Rival, A., & Levang, P. (2014). *Palms of controversies: Oil palm and development challenges*.

Center for International Forestry Research (CIFOR).

<https://doi.org/10.17528/cifor/004860>

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation

systems in the Great Plains with ERTS. *NASA Special Publication*, 351, 309-317.

Roy, D. P., Borak, J. S., Devadiga, S., Wolfe, R. E., Zheng, M., & Descloitres, J. (2002). The

MODIS land product quality assessment approach. *Remote Sensing of Environment*,

83(1–2), 62–76. [https://doi.org/10.1016/S0034-4257\(02\)00087-1](https://doi.org/10.1016/S0034-4257(02)00087-1)

Solidaridad. (2022). *Palm Oil Barometer 2022*. Retrieved July 14, 2025, from

[https://www.solidaridadnetwork.org/wp-content/uploads/2022/09/Palm-Oil-Barometer-](https://www.solidaridadnetwork.org/wp-content/uploads/2022/09/Palm-Oil-Barometer-2022_solidaridad.pdf)

[2022_solidaridad.pdf](https://www.solidaridadnetwork.org/wp-content/uploads/2022/09/Palm-Oil-Barometer-2022_solidaridad.pdf)

Stehman, S. V., & Foody, G. M. (2019). Key issues in rigorous accuracy assessment of land

cover products. *Remote Sensing of Environment*, 231, 111199.

<https://doi.org/10.1016/j.rse.2019.05.018>

- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- Vasilakos, C., Kavroudakis, D., & Georganta, A. (2020). *Machine Learning Classification Ensemble of Multitemporal Sentinel-2 Images: The Case of a Mixed Mediterranean Ecosystem*. *Remote Sensing*, 12(12), Article 2005. <https://doi.org/10.3390/rs12122005>
- Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114(1), 106–115. <https://doi.org/10.1016/j.rse.2009.08.014>
- Vijay, V., Pimm, S. L., Jenkins, C. N., & Smith, S. J. (2016). The impacts of oil palm on recent deforestation and biodiversity loss. *PLOS ONE*, 11(7), e0159668. <https://doi.org/10.1371/journal.pone.0159668>
- Wallace, R. G., Gilbert, M., Wallace, R., Pittiglio, C., Mattioli, R., & Kock, R. (2014). Did Ebola emerge in West Africa by a policy-driven phase change in agroecology? Ebola's social context. *Environment and Planning A: Economy and Space*, 46(11), 2533–2542. <https://doi.org/10.1068/a46263>
- Whitcraft, A. K., Becker-Reshef, I., Killough, B. D., & Justice, C. O. (2015). Meeting Earth observation requirements for global agricultural monitoring: An evaluation of the revisit capabilities of current and planned moderate resolution optical Earth observing missions. *Remote Sensing*, 7(2), 1482–1503. <https://doi.org/10.3390/rs70201482>
- World Economic Forum. (2022). *How African palm oil can boost livelihoods and protect forests*. Retrieved from <https://www.weforum.org/agenda/2022/11/how-african-palm-oil-can-boost-livelihoods-and-protects-forests/>

- Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, 2017, 1353691.
<https://doi.org/10.1155/2017/1353691>
- Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.-E., ... Arino, O. (2021). *ESA WorldCover 10 m 2020 v100 (Version v100)* [Data set]. Zenodo.
<https://doi.org/10.5281/zenodo.5571936>
- Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. *Remote Sensing of Environment*, 221, 430–443. <https://doi.org/10.1016/j.rse.2018.11.032>

CHAPTER 3

OIL PALM PLANTATIONS, DEFORESTATION, AND AFRICAN FILOVIRUSES

3.1 Introduction

African filoviruses—Ebola virus disease (EVD) and Marburg virus disease (MVD)—remain among the deadliest infectious threats worldwide. Historical outbreaks of EVD have produced case fatality ratios of 25 %–90 % (Feldmann & Geisbert, 2011; WHO, 2021), while MVD outbreaks have ranged from 20 % to 90 % (Bausch et al., 2006; CDC, 2024). The 2004–2005 Marburg outbreak in Angola vividly illustrated this lethality, killing 227 of 252 confirmed patients (Towner et al., 2006). Even larger in scale, the 2014–2016 West African Ebola crisis claimed 11,325 lives—518 of them healthcare workers—and wiped an estimated US \$53 billion from the combined GDP of Guinea, Liberia, and Sierra Leone (WHO Ebola Response Team, 2016; Huber et al., 2018).

Filoviruses spill over to people through complex ecological networks involving reservoir species and susceptible mammalian hosts. Within the Pteropodidae family, the Egyptian fruit bat (*Rousettus aegyptiacus*) has been established as a confirmed reservoir for Marburg virus, supported by consistent detection of viral RNA and antibodies in wild populations (Amman et al., 2012; Towner et al., 2009). In contrast, the reservoir ecology of Ebola virus remains less definitive, with several bat species—*Hypsignathus monstrosus*, *Epomops franqueti*, and *Myonycteris torquata*—proposed as potential reservoirs based on serological and molecular evidence, though conclusive identification is still lacking (Leroy et al., 2005; De Nys et al., 2018). Nonhuman primates, including gorillas, chimpanzees, and duikers, frequently serve as intermediate

amplifying hosts, with numerous human index cases traced to direct contact with or consumption of infected carcasses (Leroy et al., 2004a).

Landscape change has amplified these transmission pathways. Spillover events cluster where ecological disruption intensifies human contact with bats and primates (Pigott et al., 2014; Alexander et al., 2015). Deforestation consistently elevates Ebola risk by altering wildlife community composition, fragmenting habitat, and drawing people deeper into formerly intact forests (Olivero et al., 2017); metrics of forest fragmentation likewise track infection hotspots (Rulli et al., 2017). Reduced understory density may eliminate natural barriers that hinder pathogen spread (Walsh et al., 2009).

Conversion of cleared forests to oil palm plantations creates an additional and often overlooked human-wildlife interface. Plantation landscapes provide fruit bats with abundant food and thermally favorable roosting sites (Shafie et al., 2011), while wide trail networks and evenly spaced palms facilitate their movement. Historical case studies underscore the risk: In Congo, bat colonies thrived in an abandoned plantation where local bat hunting preceded a documented Ebola outbreak (Leroy et al., 2009). Industrial monoculture estates may pose an even greater danger than smallholder mosaics: by sharply reducing biodiversity and simplifying ecological communities, they erode biotic controls that normally dampen pathogen transmission (Wilcox & Ellis, 2006; Perfecto & Vandermeer, 2010). In contrast, smallholder or polyculture systems, which retain forest patches and crop diversity, can dilute reservoir populations across more heterogeneous habitats (Kremen & Miles, 2012; Wallace et al., 2016).

Although deforestation has been rigorously linked to filovirus emergence (Olivero et al., 2017; Rulli et al., 2017), the specific contribution of oil palm expansion has not been systematically quantified. To fill this gap, we assemble a spatial panel of 10,674 grid cells

across Africa (2001–2018) and estimate how plantation development—disaggregated into smallholder versus industrial regimes—influences the probability of zoonotic spillover. This approach allows us to identify whether and under what production systems oil palm growth exacerbates the risk of future Ebola and Marburg spillovers.

3.2 Existing Evidence and Conceptual Framework

3.2.1 Existing Evidence on Land-Use Change and Filovirus Spillover

A growing body of literature underscores the tight coupling between environmental disturbances—particularly deforestation—and the emergence or re-emergence of filoviruses, including Ebola and Marburg (Olivero et al., 2017; Rulli et al., 2017). Forest fragmentation and habitat alteration can disrupt ecological communities in ways that heighten pathogen transmission, such as by altering the distribution and abundance of virus reservoirs and intermediate hosts (Walsh et al., 2009). These disruptions increase the interface between humans and wildlife, creating new pathways for spillover events. Recent forest losses have been significantly correlated with higher Ebola spillover risk in West and Central Africa (Alexander et al., 2015; Pigott et al., 2014), while secondary or fragmented forest patches may harbor elevated densities of bats or primates, further amplifying the potential for human contact with infected hosts (Olivero et al., 2017; Rulli et al., 2017).

The rapid expansion of oil palm cultivation has emerged as an increasingly critical factor in zoonotic disease ecology, operating synergistically with deforestation processes. Both smallholder and large-scale industrial plantations reshape landscapes in ways that can facilitate filovirus transmission. Plantations often attract fruit bat species by offering readily available feeding and roosting sites (Shafie et al., 2011), while their simplified vegetation structure, including wide trails and uniform spacing, can facilitate bat movement between roosts and

foraging grounds (Leroy et al., 2009). Large-scale industrial plantations have drawn particular scrutiny for their role in reducing biodiversity and homogenizing local ecosystems (Wilcox & Ellis, 2006; Wallace et al., 2016). This ecological simplification weakens natural regulatory mechanisms that might otherwise suppress disease propagation (Perfecto & Vandermeer, 2010) and may lead to increased densities of reservoir species (Kremen & Miles, 2012). By contrast, smallholder or polyculture plantation systems can retain more biodiversity, integrating forest remnants and crop diversity that potentially diffuse reservoir populations and reduce concentrated contact with humans (Kremen & Miles, 2012; Wallace et al., 2016).

Beyond land-use factors, broader socioeconomic and cultural practices fundamentally condition filovirus spillover risk. Bushmeat hunting and consumption—particularly of primates and duikers—can provide direct routes for viral transmission when infected animals are handled or consumed (Leroy et al., 2004a). Population growth, land tenure changes, and economic pressures may further encourage communities to encroach on formerly intact forest areas or expand agricultural frontiers, thus amplifying exposure to potential viral reservoirs. Concurrently, urbanization patterns and infrastructural development (e.g., roads, markets) can bring once-isolated wildlife populations into greater contact with humans, sometimes accelerating or amplifying outbreaks (Pigott et al., 2014).

3.2.2 Conceptual framework: land-use change and filovirus spillover

This framework integrates existing evidence with theoretical mechanisms to explain how deforestation and oil palm expansion modulate filovirus spillover risk through distinct but interconnected pathways.

Deforestation and Habitat Disruption

A growing body of literature underscores the tight coupling between environmental disturbances—particularly deforestation—and the emergence or re-emergence of filoviruses, including Ebola and Marburg (Olivero et al., 2017; Rulli et al., 2017). Forest fragmentation and habitat alteration can disrupt ecological communities in ways that heighten pathogen transmission, such as by altering the distribution and abundance of virus reservoirs and intermediate hosts (Walsh et al., 2009).

These disruptions increase the interface between humans and wildlife, creating new pathways for spillover events. Recent forest losses have been significantly correlated with higher Ebola spillover risk in West and Central Africa (Alexander et al., 2015; Pigott et al., 2014), while secondary or fragmented forest patches may harbor elevated densities of bats or primates, further amplifying the potential for human contact with infected hosts (Olivero et al., 2017; Rulli et al., 2017).

Forest loss reduces the spatial barriers that traditionally limit contact among wildlife species and between wildlife and humans (Olivero et al., 2017). By disturbing primary forest habitats, deforestation may drive bat populations into edge or agricultural areas, thereby increasing the probability of human–bat encounters and potential pathogen transmission. These habitat modifications can create novel ecological niches that favor certain reservoir species while displacing others, potentially concentrating viral hosts in human-modified landscapes (Brooks et al., 2019). This pattern aligns with Schmalhausen's Law, which posits that organisms under stress often become more vulnerable to secondary stressors, potentially contributing to higher viral loads or increased shedding among reservoir species (Kareiva et al., 1993; Fox & Reed, 2011).

Oil Palm Plantations as Amplification Sites

The rapid expansion of oil palm cultivation has emerged as an increasingly critical factor in zoonotic disease ecology, operating synergistically with deforestation processes. Both smallholder and large-scale industrial plantations reshape landscapes in ways that can facilitate filovirus transmission. Plantations often attract fruit bat species by offering readily available feeding and roosting sites (Shafie et al., 2011), while their simplified vegetation structure, including wide trails and uniform spacing, can facilitate bat movement between roosts and foraging grounds (Leroy et al., 2009).

Where forests are converted to oil palm, the type of management regime—smallholder versus industrial—plays a pivotal role in shaping disease emergence risk. Smallholder or polyculture systems often retain forest fragments or include mixed cropping regimes that preserve some level of biodiversity (Kremen & Miles, 2012). This greater ecological complexity can help disperse reservoir hosts over a broader area, potentially diminishing the intensity of human–bat contact.

In contrast, large-scale industrial plantations often represent homogenous landscapes with minimal natural habitat left intact. Such monocultures can concentrate wildlife populations, particularly fruit bats, within or near plantation zones for food and roosting. Large-scale industrial plantations have drawn particular scrutiny for their role in reducing biodiversity and homogenizing local ecosystems (Wilcox & Ellis, 2006; Wallace et al., 2016). This ecological simplification weakens natural regulatory mechanisms that might otherwise suppress disease propagation (Perfecto & Vandermeer, 2010) and may lead to increased densities of reservoir species (Kremen & Miles, 2012). Moreover, the regular spatial arrangement of oil palm rows can facilitate bat movement, increasing opportunities for spillover events (Shafie et al., 2011; Leroy et al., 2009).

Socioeconomic Drivers

Socioeconomic conditions, particularly systemic poverty and institutional fragility, create complex pathways that heighten the risk of filovirus spillover events. In regions characterized by limited economic diversification and weak governance structures, communities often face stark choices between immediate survival needs and long-term health security.

Beyond land-use factors, broader socioeconomic and cultural practices fundamentally condition filovirus spillover risk. Bushmeat hunting and consumption—particularly of primates and duikers—can provide direct routes for viral transmission when infected animals are handled or consumed (Leroy et al., 2004a). While often framed primarily as subsistence activity, bushmeat hunting frequently serves as a crucial income source rather than solely for direct consumption (Cawthorn & Hoffman, 2015; de Merode et al., 2004). Indeed, wild meat is often considered a luxury rather than necessity, highlighting the economic rather than nutritional drivers of such practices (Friant et al., 2020).

Population growth, land tenure changes, and economic pressures may further encourage communities to encroach on formerly intact forest areas or expand agricultural frontiers, thus amplifying exposure to potential viral reservoirs. This economic pressure manifests in increased resource extraction activities within forested areas, as households seek supplementary income sources through activities that intensify human-wildlife contact (Leach, 2015).

These economic vulnerabilities typically coincide with institutional weaknesses that further amplify spillover risks. Areas experiencing socio-political instability or armed conflict show heightened vulnerability to zoonotic disease emergence, as health surveillance systems collapse and populations are displaced into forested areas (Bausch & Schwarz, 2014; Benedicta et

al., 2022). Conflict zones can simultaneously accelerate deforestation and wildlife exploitation while diminishing capacity for outbreak detection and response.

Concurrently, urbanization patterns and infrastructural development (e.g., roads, markets) can bring once-isolated wildlife populations into greater contact with humans, sometimes accelerating or amplifying outbreaks (Pigott et al., 2014).

Climatological and Environmental Conditions

Temperature and precipitation patterns shape the distribution, reproduction, and migration of known or suspected reservoir species. Climatic fluctuations can also affect the phenology of fruiting trees in plantations and adjacent forest remnants, potentially influencing bat movement and roosting behaviors (Schmidt et al., 2017). These environmental conditions may modulate the temporal and spatial patterns of human-wildlife contact, creating seasonal variations in spillover risk.

In this integrated framework, filovirus spillover emerges as a complex phenomenon shaped by the dynamic interplay of ecological disruption, agricultural transformation, socioeconomic conditions, and environmental factors. The framework highlights how deforestation and oil palm expansion function as primary drivers that restructure landscapes and wildlife habitats, while socioeconomic vulnerabilities and climatic patterns modulate the frequency and intensity of human-wildlife interactions.

3.3 Data

The dataset was constructed using a comprehensive spatial framework that integrates information on filovirus spillover occurrences, deforestation, oil palm plantation areas, and socioeconomic and environmental factors across Africa between 2001 and 2018.

3.3.1 Spatial Framework and Overall Dataset Structure

We base our spatial structure on the Prio-grid dataset developed by Tollefsen et al. (2022). This grid divides the African continent into subnational units of $0.5^\circ \times 0.5^\circ$ latitude and longitude—cells that measure roughly $55 \text{ km} \times 55 \text{ km}$ at the equator, tapering in size toward higher latitudes. In total, 10,667 grid cells are included, covering 51 African countries. Each cell is assigned a unique identifier, allowing us to merge data from multiple sources while preserving geographical alignment.

Table 3.1 provides a summary of key descriptive statistics for 192,132 cell-year observations drawn from these African countries between 2001 and 2018. As illustrated, filovirus events are extremely rare, with a mean rate of just 0.0001301 per cell-year, corresponding to 0.013% (SD = 0.0114). Despite their rarity, these events have profound epidemiological and socioeconomic implications, underscoring the need to understand their drivers.

Table 3.1 Descriptive statistics at cell level

Variables	Observations	Mean	Standard deviation	Median
Filovirus Spillover Events	192,132	0.0001301	0.0114063	0.0000000
Ratio of forest loss	192,132	0.0110544	0.0348801	0.0000000
Ratio of smallholder oil palm plantations	192,132	0.0000714	0.0011456	0.0000000
Ratio of industrial oil palm plantations	192,132	0.0000554	0.0008405	0.0000000
Mean of Night Light per hectare	192,132	0.009544	0.0323872	0.0000000
Mean of Population per hectare	192,132	0.3277523	1.173621	0.0552002
Mean of temperature (°C)	192,132	24.43532	3.950397	24.75833
Mean of rainfall (mm)	192,132	654.427	612.8877	491.2

Source: Authors' computation

3.3.2 Filovirus Spillover Locality

While the Centers for Disease Control and Prevention (CDC) has extensive records on Ebola outbreaks (e.g., case counts, mortality), they do not provide precise geolocations of these events. Consequently, we rely on Sundaram et al. (2024) for Ebola spillover data and Filion et al. (2023) for Marburg spillover data—both of which include accurate geocoordinates and the month in which initial spillover occurred as shown in Figure 3.1.

Our analytical period (2001–2018) encompasses 18 Ebola spillover events and 6 Marburg spillover events. It is important to note that filovirus outbreaks have historically been confined to the African continent, which justifies our geographical focus. As evident from Figure 1, these events display a distinct geographical pattern, predominantly occurring in countries within the equatorial belt of Africa, including the Democratic Republic of Congo, Uganda, Guinea, and Angola. This concentration around the equator corresponds with the natural habitat range of suspected reservoir species, particularly fruit bats.

Temporally, we observe fluctuations in outbreak frequency, with clusters of events in 2007-2008 and 2012-2014, suggesting potential cyclical patterns that may correlate with ecological or climatic factors. This distribution pattern informs our subsequent robustness tests, where we conduct additional analyses limited to the equatorial latitudes used in Pigott et al. (2014).

A complete listing of filovirus outbreak localities is provided in Appendix A.

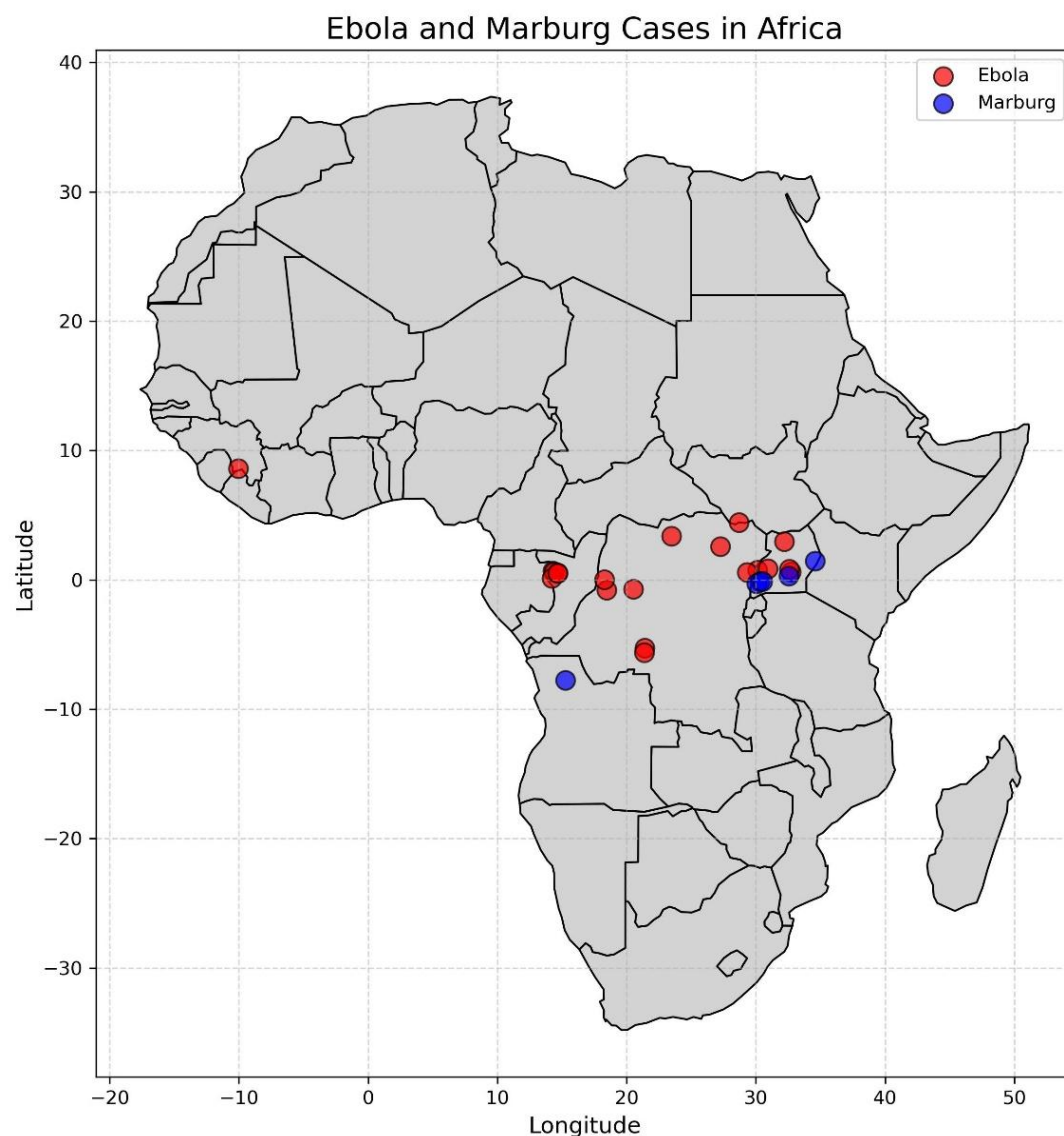


Figure 3.1: Ebola and Marburg Cases in Africa between 2001 – 2018

Source: Sundaram et al. (2024) and Fillion et al. (2023)

For each cell-month observation, we define a binary indicator FV_{kt} equal to 1 if at least one filovirus event occurs in cell k during year t , and 0 otherwise. Recognizing that the straw-colored fruit bat (*Eidolon helvum*)—a known reservoir for filoviruses (Okawa et al., 2015)—routinely travels 10–40 km nightly and has been documented to cover up to 88 km (Richter & Cumming,

2008; Fahr et al., 2015), we further extend the indicator to include the eight adjacent grid cells around each reported spillover point as shown in Figure 3.2.

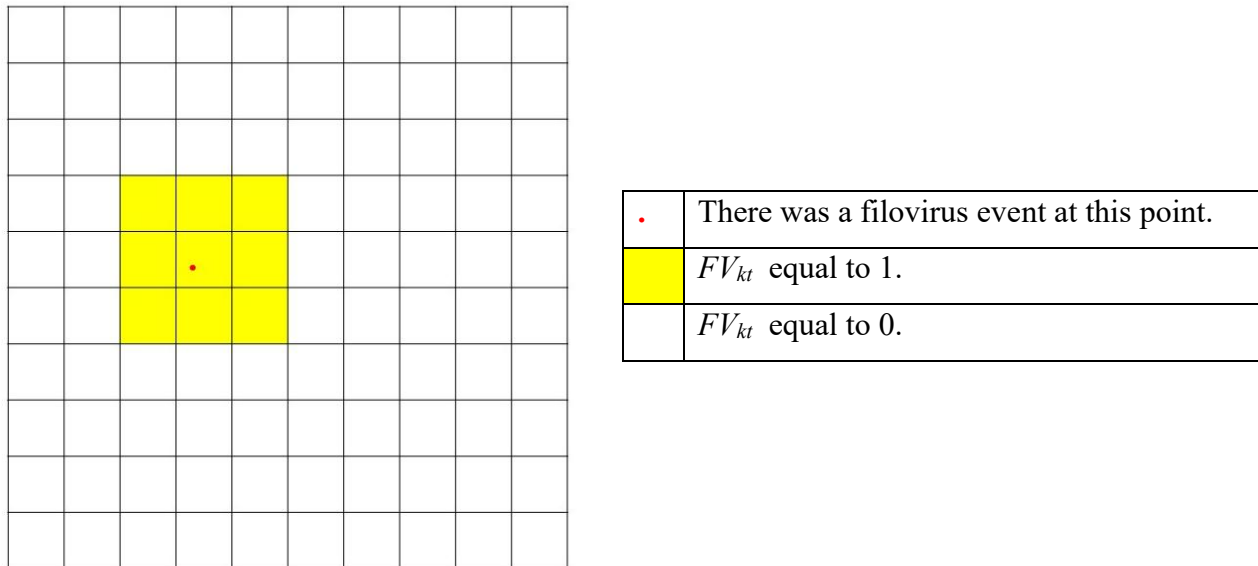


Figure 3.2: Creation of binary indicator FV_{kt}

This approach allows us to capture potential spillover effects in the immediate vicinity of reported outbreaks, accounting for both bat mobility and potential human movement in these regions.

3.3.3 Deforestation

We employ the Global Forest Change dataset (Hansen et al., 2013) for annual estimates of forest cover change at a 30-meter resolution using Landsat imagery. Data for 2000–2023 are accessible, but we focus on annual forest loss from 2001 to 2018. We process these layers in Google Earth Engine, aggregating forest loss within each 0.5° Prio-grid cell. To ensure comparability across latitudes—where cells vary in actual land area—we normalize the total forest

loss area by the cell's land area, yielding a standardized deforestation ratio (forest loss area / total cell area) in a grid cell.

As shown in Table 3.1, deforestation varies considerably across the African continent. The average annual forest loss ratio is approximately 1.11% (SD = 0.0349), but with a median of zero, revealing a stark contrast between areas experiencing significant deforestation and regions exhibiting negligible levels of forest disturbance. This spatial heterogeneity is particularly pronounced in the Congo Basin and West African forests, where most filovirus events have been documented.

3.3.4 Smallholder and Industrial Oil Palm Plantations

To examine how different types of oil palm operations might influence spillover risk, we compile annual classifications of oil palm plantations into two distinct categories—smallholder and industrial—for each grid cell from 2001–2018.

This analysis addresses a significant data limitation in African oil palm research. Unlike Southeast Asian plantations that have been extensively mapped, African oil palm datasets were extremely scarce until recently. The development of advanced remote sensing techniques has only recently enabled continental-scale classification of plantation types. While Descals et al. (2024) provides a global classification using Sentinel-1 data spanning 2016–2021 with back-estimation to 1990, that dataset faces potential accuracy issues from Landsat 7's scan line corrector failure after May 2003, which created systematic data gaps in the imagery.

To overcome these scan line limitations, we developed an alternative approach using MODIS satellite data. MODIS provides uninterrupted daily global coverage from 2000 onward, ensuring consistent temporal monitoring of land cover changes across Africa. This continuous data stream proves especially valuable for detecting brief windows of bare soil exposure that

characterize new plantation establishment—events often missed by sensors with lower temporal resolution or compromised by data gaps.

Our methodological framework, detailed in Chapter 2, integrates multiple spectral indices including the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and raw reflectance bands (Red, Near-Infrared, and Blue) to capture the diverse biophysical characteristics of African landscapes. These combined indices enhance discrimination between land cover types, particularly in heterogeneous and seasonally dynamic environments.

For classification, we employed an ensemble of machine learning algorithms, ultimately selecting XGBoost as our primary classifier based on its superior cross-validation performance (overall accuracy: $80.25\% \pm 2.12\%$). This approach proved especially reliable for identifying industrial plantations, which exhibit uniform canopy structure and spatial arrangement, compared to the more fragmented and heterogeneous smallholder systems.

The resulting annual maps distinguish between smallholder and industrial plantation regimes, with areas normalized by grid cell size to yield spatially explicit ratios for epidemiological analysis. We validated our MODIS-based classifications through cross-referencing with Descals et al. (2024), confirming result stability across different measurement approaches. While our methodology offers significant improvements in temporal continuity and classification accuracy, mixed-pixel effects in highly fragmented landscapes may still limit precision in detecting smallholder plantations.

After classification, we normalize plantation areas by total cell area, yielding ratios of smallholder and industrial oil palm coverage per grid cell. Table 3.1 reveals that both smallholder (mean = 0.0000714) and industrial (mean = 0.0000554) plantation coverage appears minimal when averaged across all cells, yet high standard deviations indicate considerable geographic

concentration. This pattern aligns with known oil palm cultivation hotspots in West and Central Africa, which notably overlap with areas of higher filovirus incidence.

3.3.5 Socioeconomic and Environmental Variables

To capture the socioeconomic and environmental contexts that may influence filovirus spillover risk, we incorporate data from the AfroGrid dataset (Schon & Koren, 2022). This dataset integrates multiple geospatial sources within the Prio-grid framework (Tollefsen et al., 2012), ensuring compatibility with our study's spatial and temporal structure.

Nighttime Light per Hectare

Nighttime light intensity serves as a widely recognized proxy for economic development, infrastructure quality, and urbanization (Chen & Nordhaus, 2011; Michalopoulos & Papaioannou, 2013). Data for this variable come from two sources: the Defense Meteorological Satellite Program (DMSP) for earlier years and the Visible Infrared Imaging Radiometer Suite (VIIRS) for later years.

We normalize nighttime light intensity by each cell's land area, yielding a measure of nighttime light per hectare. While nighttime lights have proven valuable for capturing economic activity, recent studies by Doll et al. (2006) and Chen & Nordhaus (2019) note potential limitations in rural areas where economic activities may not generate significant illumination. Nevertheless, the variable remains useful for identifying areas with infrastructure development that might influence disease surveillance capabilities.

Population per Hectare

Population density reflects the distribution and concentration of human populations within each grid cell. We derive these estimates from the WorldPop project (WorldPop, 2021), which provides spatially disaggregated, annually updated population counts for Africa. Like nighttime

light intensity, population counts are normalized by the cell's land area to yield population per hectare.

As seen in Table 3.1, population density shows considerable variation (mean = 0.3277 persons per hectare, SD = 1.17), reflecting Africa's complex settlement patterns, from dense urban centers to sparsely populated rural areas. This variation likely influences human-wildlife contact patterns and consequently spillover risk.

Temperature and Precipitation

Both temperature and precipitation data are sourced from the Climate Research Unit Time-Series (CRU TS) dataset (Harris et al., 2020). Following the approach used in Pigott et al. (2014), we include these variables to account for environmental conditions that influence reservoir species distribution and behavior. Temperature plays a critical role in shaping the habitats of virus reservoir species such as bats and primates (Han et al., 2016), while precipitation patterns affect the structure and composition of ecosystems.

The inclusion of these climate variables allows us to control for their effects when assessing the impact of deforestation and oil palm expansion. By including cell fixed effects in our econometric model (see Section 3.4), we leverage the year-to-year variation in these variables within each cell to identify their influence on spillover risk.

Our inclusion of both socioeconomic and environmental controls allows us to isolate the specific effects of land-use changes (deforestation and oil palm expansion) from broader contextual factors that might independently influence filovirus spillover dynamics.

3.4 Estimation Strategy

To analyze the determinants of filovirus spillover events, we employ a linear probability model (LPM) as our baseline specification. While nonlinear models like logit or probit are

common alternatives for binary outcomes, the LPM offers several advantages in our context, including straightforward interpretation of coefficients, computational efficiency with high-dimensional fixed effects, and flexibility in handling interaction terms for heterogeneity analysis (Angrist & Pischke, 2009).

Our dependent variable is a binary indicator representing spillover occurrence in a specific location during a given year. The probability of a filovirus spillover in grid cell k at time t is modeled as follows:

$$\Pr[FV_{kt}=1] = \alpha_1 ForestLoss_{kt} + \alpha_2 Small_{kt} + \alpha_3 Industrial_{kt} + \alpha_4 X_{kt} + \gamma_k + \delta_{it} + \varepsilon_{kt}$$

In this equation, FV_{kt} is a binary indicator equal to 1 if a filovirus (Ebola or Marburg) spillover event occurs in grid cell k or its eight adjacent cells at time t and 0 otherwise. $ForestLoss_{kt}$ represents the ratio of deforested area to total cell area in grid cell k at time t . $Small_{kt}$ and $Industrial_{kt}$ capture the ratios of smallholder and industrial oil palm plantation areas to total cell area, respectively. X_{kt} is a vector of time-varying control variables measured at the grid cell level, including nighttime light intensity, population density, mean annual temperature, and annual rainfall. We also include γ_k , which denotes grid cell fixed effects controlling for time-invariant characteristics at the cell level, and δ_{it} , which represents country-by-year fixed effects accounting for time-varying national factors. Finally, ε_{kt} is the error term.

Our empirical strategy addresses potential sources of bias through several channels. First, we include cell-specific fixed effects to control for unobserved, time-invariant characteristics at the grid cell level. These fixed effects absorb any stable geographical features such as elevation, slope, soil quality, historical land-use patterns, and distance to rivers or other natural boundaries. They also account for baseline ecological conditions that might influence both land-use decisions and disease dynamics. By employing a within-cell identification strategy, we effectively compare each cell to itself over time, isolating the impact of temporal changes in our key independent

variables while holding geography constant. This approach is particularly crucial when studying spatial phenomena where location-specific factors might confound the relationship between land use and disease emergence.

We also incorporate country-by-year fixed effects to account for time-varying factors at the national level. These include policy changes affecting forest management or agricultural development, national economic conditions that might influence land-use decisions, public health interventions and disease surveillance capabilities, annual fluctuations in climate patterns that might affect vector populations, and conflict or political instability that could disrupt both land management and health systems. By including these fixed effects, we absorb any nation-wide shocks or trends that might simultaneously influence land-use practices and disease dynamics. This two-way fixed effects structure helps mitigate omitted variable bias by accounting for both spatial and temporal unobserved heterogeneity (Wooldridge, 2010). The country-by-year fixed effects also address potential concerns about reporting bias, as outbreak detection may vary with a country's surveillance capabilities, which can change over time.

Given that our dependent variable is binary and extremely rare (mean = 0.00013), we take specific steps to address the challenges associated with rare events analysis. While King and Zeng (2001) demonstrate that logistic regression may underestimate the probability of rare events, other econometric studies suggest that linear probability models with fixed effects can remain consistent even with rare binary outcomes (Greene, 2004).

While linear probability models have known limitations, including the possibility of predicted probabilities outside the $[0,1]$ interval and heteroskedasticity, this specification offers several advantages for our analysis. The coefficients are directly interpretable as marginal effects on the probability of spillover occurrence. For instance, α_1 represents the change in spillover

probability associated with a one-unit increase in forest loss ratio. The model readily accommodates our extensive fixed effects structure without encountering the incidental parameters problem common to nonlinear specifications (Lancaster, 2000). Additionally, the linear framework allows for straightforward interpretation of interaction terms in subsequent specifications exploring heterogeneous effects. To address heteroskedasticity concerns, we employ robust standard errors throughout our analysis. Additionally, we verify that predicted probabilities from our main specifications remain predominantly within the unit interval, mitigating concerns about the linear functional form.

3.5 Results

The empirical results from the linear probability models (LPM) examining the determinants of filovirus spillover events are presented in Table 3.2. Initially, in the simplest specification without fixed effects (Column 1), the ratio of forest loss demonstrates a strong positive and statistically significant relationship with spillover probability. However, this association diminishes substantially upon the introduction of grid-cell fixed effects (Column 2) and becomes statistically insignificant once country-by-year fixed effects are incorporated (Columns 3–5). The attenuation of these coefficients and the subsequent loss of significance indicate that the initial observed relationship primarily captured time-invariant, location-specific characteristics rather than dynamic temporal changes within individual cells.

Table 3.2 Main model estimates

Estimator Dependent variable	LPM Was there a FV spillover?					
	(1)	(2)	(3)	(4)	(5)	(6)
Ratio of forest loss	0.03475*** (0.00600)	0.01433* (0.00747)	0.00460 (0.00766)	0.00431 (0.00764)	0.00449 (0.00766)	-0.01767 (0.01451)
Ratio of forest loss ²						0.06353 (0.04527)
Ratio of smallholder oil palm plantations	-0.15059*** (0.04093)	-0.04060 (0.02672)	-0.06966* (0.03601)	-0.09827** (0.04561)	-0.09598** (0.04475)	-0.29294** (0.12437)
Ratio of smallholder oil palm plantations ²						2.18405** (0.98021)
Ratio of industrial oil palm plantations	-0.11874 (0.09426)	-0.01964 (0.08751)	0.21982** (0.11102)	0.21430* (0.11300)	0.21952* (0.11327)	0.52499** (0.26484)
Ratio of industrial oil palm plantations ²						-7.56173 (4.91112)
Socioeconomic Controls	No	No	No	Yes	Yes	Yes
Environmental Controls	No	No	No	No	Yes	Yes
Country x year fixed effects	No	No	Yes	Yes	Yes	Yes
Cell fixed effects	No	Yes	Yes	Yes	Yes	Yes
Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Observations	192,006	192,006	192,006	192,006	192,006	192,006
Cells	10,667	10,667	10,667	10,667	10,667	10,667

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, standard errors in parentheses.

The effects of smallholder oil palm plantations on spillover risk reveal a consistently negative relationship across most model specifications. Although initially not statistically significant with only grid-cell fixed effects (Column 2), the negative association becomes significant when country-by-year fixed effects and additional socioeconomic and environmental controls are included (Columns 3–5). Specifically, in the fully controlled model (Column 5), smallholder plantation coverage exhibits a statistically significant negative relationship with spillover risk. This result suggests that smallholder plantations, characterized by landscape heterogeneity and ecological complexity, potentially mitigate the conditions conducive to filovirus transmission.

Introducing quadratic terms in column 6 uncovers nonlinear relationships for smallholder plantations, revealing a statistically significant U-shaped pattern. Specifically, the linear term is negative, indicating reduced spillover risk at lower plantation coverage levels, whereas the quadratic term is positive, suggesting a potential increase in risk at higher plantation densities.

Notably, however, observed plantation densities rarely reach levels at which the risk-reduction effect reverses. Thus, within the typical observed range, smallholder plantations predominantly exert a protective effect against filovirus spillover.

Conversely, industrial oil palm plantations exhibit a distinctly positive and statistically significant association with spillover risk in models incorporating fixed effects and additional covariates (Columns 3–5). This positive relationship becomes particularly robust and pronounced when country-by-year fixed effects and socioeconomic and environmental controls are included, underscoring the role of industrial plantation regimes in elevating filovirus spillover risk. The nonlinear specification (column 6) further reinforces these findings, showing a consistent positive linear effect for industrial plantations, while the quadratic term remains statistically insignificant. This outcome suggests that the spillover risk increases steadily with industrial plantation density, aligning with ecological hypotheses highlighting the role of biodiversity loss and ecological simplification in facilitating pathogen transmission.

Overall, these results illustrate significant heterogeneity in spillover risks associated with different land-use management systems. Industrial monoculture systems consistently enhance the risk of filovirus spillovers, whereas smallholder systems demonstrate a generally protective effect. These findings support the ecological theory suggesting that landscape heterogeneity and biodiversity preservation within smallholder systems may effectively reduce zoonotic transmission risks compared to ecologically simplified industrial plantations.

3.6 Robustness Checks

This section presents a comprehensive set of sensitivity analyses that establish the reliability of our main findings across alternative specifications, variable definitions, and data sources. These robustness checks systematically address potential concerns regarding the stability

of our findings and confirm that the differential effects of industrial versus smallholder oil palm plantations on filovirus spillover risk are not artifacts of particular methodological choices.

3.6.1 Varying the Spatial Definition of Filovirus Spillover

One potential concern with our analysis is that the geographical delineation of spillover events might influence our findings. The baseline model codes a filovirus spillover in cell k if an event occurs either in that cell or within its eight surrounding cells (i.e., a 9-cell buffer). This approach acknowledges the mobility of fruit bat species—particularly the straw-colored fruit bat (*Eidolon helvum*), which routinely travels 10-40 km nightly and has been documented to cover up to 88 km (Fahr et al., 2015; Abedi-Lartey et al., 2016). However, to ensure our results are not sensitive to this particular spatial definition, we examine two alternative configurations:

1 Cell Definition: A stricter metric where $FV_{kt} = 1$ only if a spillover occurs directly in cell k .

13 Cells Definition: An expanded buffer that includes the eight adjacent cells plus four additional cells along the middle facets of each side, yielding 13 total cells.

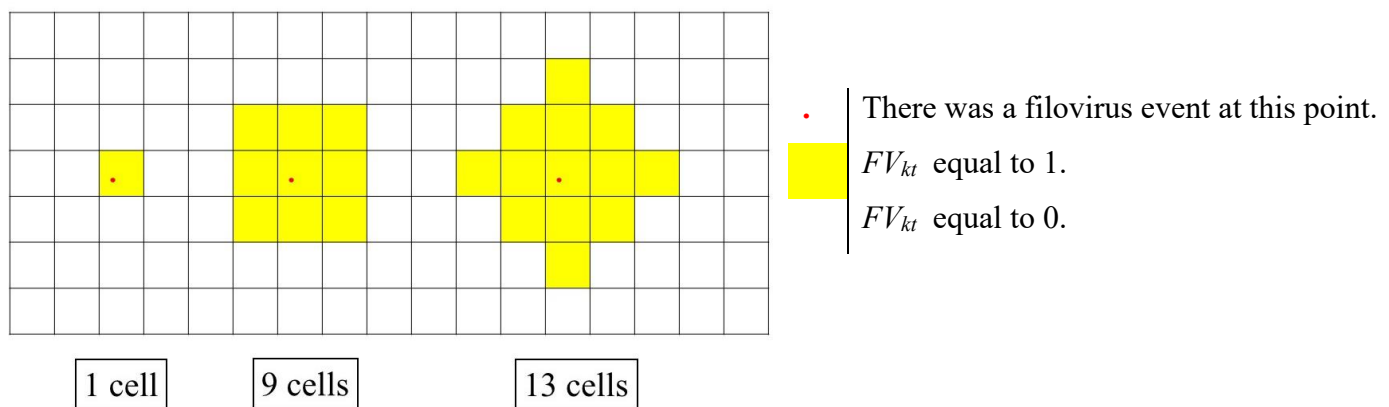


Figure 3.3: Definition of each cell

Table 3.3 presents robustness checks employing alternative spatial definitions of filovirus spillover events to verify the sensitivity of our results. Initially, our baseline definition (9-cell buffer) codes a filovirus spillover event as occurring if it happens within a grid cell or its eight surrounding cells. This definition accounts for the typical mobility of reservoir species, particularly fruit bats, which frequently travel substantial distances. To assess robustness, we also employ two alternative spatial definitions: a stricter 1-cell definition, which records spillovers only within the exact grid cell of occurrence, and a broader 13-cell definition, encompassing spillovers occurring in the cell plus the adjacent eight cells and four additional surrounding cells.

Table 3.3 Three alternative Spatial Definitions of Filovirus Spillover

Estimator Dependent variable	LPM		
	Was there a FV spillover?		
	1 cell	9 cells	13 cells
Ratio of forest loss	-0.00389 (0.00386)	-0.01767 (0.01451)	-0.02441 (0.01498)
Ratio of forest loss ²	0.01487 (0.01397)	0.06353 (0.04527)	0.07136 (0.04536)
Ratio of smallholder oil palm plantations	-0.04400 (0.03630)	-0.29294** (0.12437)	-0.36675** (0.15032)
Ratio of smallholder oil palm plantations ²	0.33642 (0.28267)	2.18405** (0.98021)	2.76774** (1.18488)
Ratio of industrial oil palm plantations	0.02314 (0.06742)	0.52499** (0.26484)	0.67729** (0.31650)
Ratio of industrial oil palm plantations ²	-0.13648 (1.14322)	-7.56173 (4.91112)	-9.18197 (5.95370)
Control variables	Yes	Yes	Yes
Country x year fixed effects	Yes	Yes	Yes
Cell fixed effects	Yes	Yes	Yes
Cluster	Yes	Yes	Yes
Observations	192,132	192,132	192,132

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, standard errors in parentheses.

Results under the strictest 1-cell definition indicate no statistically significant relationships between smallholder or industrial oil palm plantations and spillover risk. The lack of significance in this restrictive spatial definition suggests it inadequately captures broader ecological processes and interactions at the landscape scale, particularly the movement and habitat use of highly mobile reservoir species such as fruit bats.

In contrast, results from both the 9-cell and the more inclusive 13-cell definitions provide robust support for the initial findings. Industrial oil palm plantations consistently exhibit

statistically significant positive associations with spillover risk, whereas smallholder plantations demonstrate robust negative relationships. The strength and statistical significance of these relationships notably increase with the broader 13-cell definition, emphasizing that ecological interactions influencing spillover risks extend beyond individual cells, reflecting landscape-scale processes.

These robustness checks underscore the importance of considering appropriate ecological scales when modeling zoonotic spillovers. The findings emphasize that broader spatial definitions better capture the critical ecological dynamics, highlighting the role of industrial plantations in amplifying spillover risks and the potential protective effects of smallholder plantations. This reinforces the validity of employing moderate to expansive spatial buffers in epidemiological studies on zoonotic disease emergence.

3.6.2 Combining Smallholder and Industrial Plantations

Table 4 presents additional robustness checks to evaluate whether aggregating smallholder and industrial oil palm plantations into a single category affects the observed relationships with filovirus spillover risk. When both plantation types are combined, the resulting oil palm variable is statistically insignificant across all spatial definitions—1-cell, 9-cell, and 13-cell buffers. This aggregated analysis contrasts sharply with earlier results, where disaggregated plantation types exhibited distinct and statistically significant effects.

Table 3.4 Combining Smallholder and Industrial Plantations

Estimator Dependent variable	LPM		
	Was there a FV spillover?		
	1 cell	9 cells	13 cells
Ratio of forest loss	-0.00390 (0.00386)	-0.01773 (0.01451)	-0.02449 (0.01497)
Ratio of forest loss ²	0.01487 (0.01396)	0.06351 (0.04524)	0.07134 (0.04533)
Ratio of oil palm plantations	-0.01014 (0.02721)	0.06197 (0.08892)	0.09810 (0.10371)
Ratio of oil palm plantations ²	0.06045 (0.18914)	-0.44410 (0.63468)	-0.65539 (0.74103)
Control variables	Yes	Yes	Yes
Country x year fixed effects	Yes	Yes	Yes
Cell fixed effects	Yes	Yes	Yes
Cluster	Yes	Yes	Yes
Observations	192,132	192,132	192,132

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, standard errors in parentheses.

The lack of significant findings for the aggregated oil palm variable highlights the importance of distinguishing between different plantation management regimes. Combining both plantation types masks their divergent ecological impacts, effectively neutralizing their respective positive and negative influences on the risk of spillover. Industrial plantations, characterized by ecological simplification and reduced biodiversity, significantly enhance spillover risk, while smallholder plantations, featuring greater landscape complexity and biodiversity, tend to mitigate risk.

These findings highlight a critical methodological insight: treating heterogeneous land-use categories as homogenous entities may obscure important epidemiological relationships. Consequently, future research and policy recommendations should carefully differentiate between plantation types to accurately reflect their distinct ecological roles and potential implications for the emergence of zoonotic diseases.

3.6.3 Using an Alternative Oil Palm Dataset

Table 3.5 presents a robustness analysis utilizing an alternative dataset from Descals et al. (2024) to verify the stability and generalizability of our primary findings. This dataset utilizes Sentinel-1 radar data and historical Landsat imagery to classify smallholder and industrial oil palm

plantations, offering a comparative assessment of plantation dynamics from an independent source.

Table 3.5 Using an Alternative Oil Palm Dataset

Estimator Dependent variable	LPM		
	Was there a FV spillover?		
	1 cell	9 cells	13 cells
Ratio of forest loss	-0.00387 (0.00386)	-0.01787 (0.01453)	-0.02458 (0.01499)
Ratio of forest loss ²	0.01484 (0.01397)	0.06372 (0.04529)	0.07152 (0.04537)
Ratio of smallholder oil palm plantations (Descals et al., 2024)	-0.03677 (0.03006)	-0.15444 (0.11027)	-0.21565 (0.13550)
Ratio of smallholder oil palm plantations ² (Descals et al., 2024)	0.27283 (0.21968)	1.01028 (0.78944)	1.44668 (0.97590)
Ratio of industrial oil palm plantations (Descals et al., 2024)	-0.05255 (0.08576)	0.54241 (0.49036)	0.60700 (0.45232)
Ratio of industrial oil palm plantations ² (Descals et al., 2024)	1.93783 (1.99766)	-10.22546 (12.28666)	-9.73837 (10.78549)
Control variables	Yes	Yes	Yes
Country x year fixed effects	Yes	Yes	Yes
Cell fixed effects	Yes	Yes	Yes
Cluster	Yes	Yes	Yes
Observations	192,132	192,132	192,132

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, standard errors in parentheses.

Results from analyses based on this alternative dataset generally align with our initial findings in terms of directionality. Specifically, industrial plantations consistently exhibit positive associations with filovirus spillover risk, while smallholder plantations show negative relationships. However, these coefficients do not reach statistical significance in this alternative dataset. This attenuation in significance is likely due to increased measurement noise associated with methodological limitations in the Descals et al. dataset, such as temporal misclassification and reduced efficacy in capturing dynamic land-use changes.

Despite the absence of statistical significance, the persistent directional relationships corroborate our primary ecological hypotheses and support the qualitative robustness of our conclusions. This robustness check highlights the critical importance of precise measurement and high-frequency temporal resolution in epidemiological studies, reinforcing the methodological advantage of employing reliable, high-quality datasets such as MODIS imagery. These findings

underscore the validity of our main conclusions regarding the contrasting ecological roles of smallholder and industrial oil palm plantations in influencing zoonotic spillover risks.

3.6.4 Individual Land-Use Impact Analysis

To examine the isolated effects of each land-use type, we conducted separate regression analyses across 9 grid cells for deforestation, smallholder oil palm plantations, and industrial oil palm plantations. These individual analyses reveal distinct patterns in their relationships with filovirus spillover risk (Table 3.6).

For deforestation alone (column 1 and 2), neither linear nor quadratic terms exhibit statistical significance, suggesting a negligible individual impact of deforestation on spillover events within the analytical framework employed. Although deforestation has frequently been associated with elevated zoonotic risks in prior studies, the lack of statistical significance here implies that isolated forest loss may not sufficiently capture the complexity of ecological dynamics influencing filovirus spillovers, particularly when socioeconomic, environmental, and spatiotemporal controls are comprehensively incorporated.

Table 3.6 Individual Land-Use Effects on Filovirus Spillover

Estimator Dependent variable	LPM Was there a FV spillover?					
	(1)	(2)	(3)	(4)	(5)	(6)
Ratio of forest loss	0.00128 (0.00281)	-0.00392 (0.00387)				
Ratio of forest loss ²		0.01489 (0.01398)				
Ratio of smallholder oil palm plantations			-0.01038 (0.00994)	-0.02913 (0.02765)		
Ratio of smallholder oil palm plantations ²				0.22643 (0.22755)		
Ratio of industrial oil palm plantations					0.00522 (0.02755)	-0.00571 (0.05820)
Ratio of industrial oil palm plantations ²						0.33518 (1.04933)
Socioeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Environmental Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country x year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Cell fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Observations	192,006	192,006	192,006	192,006	192,006	192,006
Cells	10,667	10,667	10,667	10,667	10,667	10,667

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, standard errors in parentheses.

The regressions isolating smallholder oil palm plantations (Columns 3 and 4) similarly show statistically insignificant coefficients, though the direction of the relationships aligns consistently with expectations from ecological theory. Specifically, the negative signs observed in both linear and quadratic specifications suggest a potential protective effect of smallholder plantation landscapes, attributable to their greater ecological heterogeneity and biodiversity. The quadratic specification further hints at a possible U-shaped relationship, indicating diminishing protective effects at higher plantation densities; however, these effects remain statistically inconclusive. This evidence underscores the role smallholder plantation management might play in moderating zoonotic spillover risk, though it also highlights the complexity of definitively capturing these dynamics through separate regressions alone.

For industrial oil palm plantations (Columns 5 and 6), the regression results likewise do not achieve statistical significance when analyzed independently. Nevertheless, the direction of the coefficients, predominantly positive, is consistent with theoretical expectations that industrial monoculture systems, characterized by ecological simplification and reduced biodiversity, could amplify zoonotic spillover risks. The absence of statistical significance in this isolated analysis suggests that the impacts of industrial plantations on spillover risk might be more accurately represented when considered within a broader ecological and land-use context, indicating possible interaction effects with other landscape features not captured here.

In summary, the separate regression analyses highlight that isolating single land-use categories—deforestation, smallholder plantations, or industrial plantations—does not fully capture the complexity or statistically robust relationships previously identified in combined regressions. These findings reinforce the importance of modeling spillover dynamics within integrated frameworks that simultaneously account for multiple interacting land-use factors.

Consequently, future analyses and policy recommendations should continue to emphasize differentiated management practices and consider synergistic effects between distinct land-use categories to effectively mitigate zoonotic disease risks.

3.7 Conclusions and Policy Implications

This chapter aimed to quantify the relationship between land-use transitions—specifically deforestation, smallholder oil palm expansion, and industrial oil palm expansion—and the risk of Ebola or Marburg virus spillover events across tropical Africa. Employing a linear-probability framework with a 0.5° grid-cell panel dataset spanning 2001-2018 and incorporating comprehensive socioeconomic and climatic controls, the analysis isolated within-cell temporal variations. This empirical approach yielded robust insights into how distinct oil palm plantation systems differentially influence filovirus spillover risk.

Key findings emerged from this analysis. First, deforestation alone, once controlling for persistent grid-cell and country-year fixed effects, was not significantly associated with filovirus spillover events. Although forest loss has commonly been linked to increased zoonotic risks in earlier literature, the absence of statistical significance here suggests that the complexity of ecological dynamics influencing spillover events requires considering subsequent land management practices beyond mere deforestation.

Second, industrial oil palm plantations were consistently and significantly associated with increased spillover probabilities. The quantitative evidence demonstrated that a one-percentage-point increase in the share of industrial oil palm plantations correlated with approximately a 0.22 percentage-point rise in spillover risk. This relationship intensified when considering broader spatial buffers reflective of bat mobility ranges, highlighting the critical role industrial-scale

monocultures may play in facilitating pathogen transmission through simplified habitats and reduced biodiversity.

Third, contrastingly, smallholder oil palm plantations exhibited a significant negative relationship with spillover risk, suggesting that these heterogeneous landscapes, which typically maintain higher biodiversity and ecological complexity, may mitigate zoonotic transmission. Sensitivity analyses employing alternative spatial scales and different datasets consistently reinforced the protective role of smallholder plantations, underscoring the ecological resilience inherent in more diversified agricultural systems.

These empirical findings contribute to broader debates within ecological and agricultural economics by clarifying that plantation management regimes, rather than oil palm cultivation per se, primarily drive variations in zoonotic spillover risks. The analysis supports ecological theories that emphasize the risk-amplifying potential of simplified landscapes compared to heterogeneous agricultural mosaics.

Given these results, targeted policy implications arise directly from this study. Policymakers should prioritize steering industrial-scale plantation developments away from ecologically sensitive areas, especially regions recognized as habitats and corridors for bat populations. Spatially explicit environmental assessments and strategic planning can help minimize epidemiological externalities associated with large-scale monoculture plantations. Additionally, supporting smallholder oil palm cultivation through technical assistance, market access initiatives, and financial incentives can simultaneously promote rural development objectives while effectively mitigating spillover risks.

Finally, this research highlights the necessity for integrated policy frameworks that explicitly consider zoonotic risk factors in agricultural expansion decisions. Coordinated efforts

across the agricultural, environmental, and public health sectors can significantly enhance disease surveillance capabilities, allowing for proactive interventions at critical plantations. Thus, promoting landscape-level planning that incorporates ecological resilience emerges as an essential step toward sustainable agricultural development that safeguards both human health and biodiversity.

3.8 References

- Abedi-Lartey, M., Dechmann, D. K. N., Wikelski, M., Scharf, A. K., & Fahr, J. (2016). Long-distance seed dispersal by straw-coloured fruit bats varies by season and landscape. *Global Ecology and Conservation*, 7, 12–24. <https://doi.org/10.1016/j.gecco.2016.03.005>
- Alexander, K. A., Sanderson, C. E., Marathe, M., Lewis, B. L., Rivers, C. M., Shaman, J., Drake, J. M., Lofgren, E., Dato, V. M., Eisenberg, M. C., & Eubank, S. (2015). What factors might have led to the emergence of Ebola in West Africa? *PLOS Neglected Tropical Diseases*, 9(6), e0003652. <https://doi.org/10.1371/journal.pntd.0003652>
- Amman, B. R., Nyakarahuka, L., McElroy, A. K., Dodd, K. A., Sealy, T. K., Schuh, A. J., Shoemaker, T. R., Balinandi, S., Atimmedi, P., Kaboyo, W., Nichol, S. T., & Towner, J. S. (2014). Marburgvirus resurgence in Kitaka Mine bat population after extermination attempts, Uganda. *Emerging Infectious Diseases*, 20(10), 1761–1764. <https://doi.org/10.3201/eid2010.140696>
- Bausch, D. G., Nichol, S. T., Muyembe-Tamfum, J. J., Borchert, M., Rollin, P. E., Sleurs, H., Campbell, P., Tshioko, F. K., Roth, C., Colebunders, R., Pirard, P., Mardel, S., Olinda, L. A., Zeller, H., Tshomba, A., Kulidri, A., Libande, M. L., Mulangu, S., Formenty, P., Grein, T., Leirs, H., Braack, L., Ksiazek, T., Zaki, S., Bowen, M. D., Smit, S. B., Leman, P. A., Burt, F. J., Kemp, A., & Swanepoel, R. (2006). Marburg hemorrhagic fever associated with multiple genetic lineages of virus. *New England Journal of Medicine*, 355(9), 909–919. <https://doi.org/10.1056/NEJMoa051465>
- Centers for Disease Control and Prevention. (2024, May 1). *About Marburg*. U.S. Department of Health & Human Services. <https://www.cdc.gov/marburg/about/index.html>

- Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21), 8589–8594. <https://doi.org/10.1073/pnas.1017031108>
- De Nys, H. M., Kingebeni, P. M., Doshi, P., Ntamponi, J. P., Wiley, M. R., Makiala-Mandanda, S., ... & Bjork, A. (2018). Survey of Ebola viruses in frugivorous and insectivorous bats in Guinea, Cameroon, and the Democratic Republic of the Congo, 2015–2017. *Emerging Infectious Diseases*, 24(12), 2228–2240. <https://doi.org/10.3201/eid2412.180740>
- Descals, A., Gaveau, D. L. A., Wich, S., Szantoi, Z., & Meijaard, E. (2024). Global mapping of oil palm planting year from 1990 to 2021. *Earth System Science Data*, 16(5), 5111–5129. <https://doi.org/10.5194/essd-16-5111-2024>
- Elvidge, C. D., Baugh, K. E., Zhizhin, M., Hsu, F. C., & Ghosh, T. (2017). VIIRS night-time lights. *International Journal of Remote Sensing*, 38(21), 5860–5879. <https://doi.org/10.1080/01431161.2017.1342050>
- Fahr J, Abedi-Lartey M, Esch T, Machwitz M, Suu-Ire R, Wikelski M, et al. (2015) Pronounced Seasonal Changes in the Movement Ecology of a Highly Gregarious Central-Place Forager, the African Straw-Coloured Fruit Bat (*Eidolon helvum*). PLoS ONE 10(10): e0138985. <https://doi.org/10.1371/journal.pone.0138985>
- Feldmann, H., & Geisbert, T. W. (2011). Ebola haemorrhagic fever. *The Lancet*, 377(9768), 849–862. [https://doi.org/10.1016/S0140-6736\(10\)60667-8](https://doi.org/10.1016/S0140-6736(10)60667-8)
- Filion, A., Sundaram, M., & Stephens, P. R. (2023). Preliminary Investigation of Schmalhausen’s law in a directly transmitted pathogen outbreak system. *Viruses*, 15(2), 310. <https://doi.org/10.3390/v15020310>

- Greene, W. H. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal*, 7(1), 98–119.
<https://doi.org/10.1111/j.1368-423X.2004.00123.x>
- Han, B. A., Kramer, A. M., & Drake, J. M. (2016). Global patterns of zoonotic disease in mammals. *Trends in Parasitology*, 32(7), 565–577.
<https://doi.org/10.1016/j.pt.2016.04.007>
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850–853.
<https://doi.org/10.1126/science.1244693>
- Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7(1), 109.
<https://doi.org/10.1038/s41597-020-0453-3>
- Hayman, D. T., Bowen, R. A., Cryan, P. M., McCracken, G. F., O'Shea, T. J., Peel, A. J., Gilbert, A., Webb, C. T., & Wood, J. L. (2013). Ecology of zoonotic infectious diseases in bats: Current knowledge and future directions. *Zoonoses and Public Health*, 60(1), 2–21.
<https://doi.org/10.1111/zph.12000>
- Huber, C., Finelli, L., & Stevens, W. (2018). The economic and social burden of the 2014 Ebola outbreak in West Africa. *The Journal of Infectious Diseases*, 218(Supplement_5), S698–S704. <https://doi.org/10.1093/infdis/jiy213>

- Kremen, C., & Miles, A. (2012). Ecosystem services in biologically diversified versus conventional farming systems: Benefits, externalities, and trade-offs. *Ecology and Society*, 17(4), 40. <http://dx.doi.org/10.5751/ES-05035-170440>
- Leroy, E. M., Epelboin, A., Mondonge, V., Pourrut, X., Gonzalez, J. P., Muyembe-Tamfum, J. J., & Formenty, P. (2009). Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector-Borne and Zoonotic Diseases*, 9(6), 723–728. <https://doi.org/10.1089/vbz.2008.0167>
- Leroy, E. M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassanin, A., Yaba, P., Délicat, A., Paweska, J. T., Gonzalez, J.-P., & Swanepoel, R. (2005). Fruit bats as reservoirs of Ebola virus. *Nature*, 438(7068), 575–576. <https://doi.org/10.1038/438575a>
- Leroy, E. M., Rouquet, P., Formenty, P., Souquière, S., Kilbourne, A., Froment, J.-M., Bermejo, M., Smit, S., Karesh, W., Swanepoel, R., Zaki, S. R., & Rollin, P. E. (2004a). Multiple Ebola virus transmission events and rapid decline of central African wildlife. *Science*, 303(5656), 387–390. <https://doi.org/10.1126/science.1092528>
- Ogawa, H., Miyamoto, H., Nakayama, E., Yoshida, R., Nakamura, I., Sawa, H., Ishii, A., Thomas, Y., Nakagawa, E., Matsuno, K., Kajihara, M., Maruyama, J., Nao, N., Muramatsu, M., Kuroda, M., Simulundu, E., Changula, K., Hang'ombe, B., Namangala, B., ... Takada, A. (2015). Seroepidemiological prevalence of multiple species of filoviruses in fruit bats (*Eidolon helvum*) migrating in Africa. *The Journal of Infectious Diseases*, 212(Suppl 2), S101–S108. <https://doi.org/10.1093/infdis/jiv063>
- Olivero, J., Fa, J. E., Real, R., Márquez, A. L., Farfán, M. A., Vargas, J. M., Gaveau, D., Salim, M. A., Park, D., Suter, J., King, S., Leendertz, S. A., Sheil, D., & Nasi, R. (2017). Recent

- loss of closed forests is associated with Ebola virus disease outbreaks. *Scientific Reports*, 7, 14291. <https://doi.org/10.1038/s41598-017-14727-9>
- Perfecto, I., & Vandermeer, J. (2010). The agroecological matrix as alternative to the land-sparing/agriculture intensification model. *Proceedings of the National Academy of Sciences*, 107(13), 5786–5791. <https://doi.org/10.1073/pnas.0905455107>
- Pigott, D. M., Golding, N., Mylne, A., Huang, Z., Henry, A. J., Weiss, D. J., Brady, O. J., Kraemer, M. U. G., Smith, D. L., Moyes, C. L., Bhatt, S., Gething, P. W., Horby, P. W., Bogoch, I. I., Brownstein, J. S., Mekaru, S. R., Tatem, A. J., Khan, K., & Hay, S. I. (2014). Mapping the zoonotic niche of Ebola virus disease in Africa. *eLife*, 3, e04395. <https://doi.org/10.7554/eLife.04395>
- Richter, H. V., & Cumming, G. S. (2008). First application of satellite telemetry to track African straw-coloured fruit bat migration. *Journal of Zoology*, 275(2), 172–176. <https://doi.org/10.1111/j.1469-7998.2008.00425.x>
- Rulli, M. C., Santini, M., Hayman, D. T. S., & D’Odorico, P. (2017). The nexus between forest fragmentation in Africa and Ebola virus disease outbreaks. *Scientific Reports*, 7, Article 41613. <https://doi.org/10.1038/srep41613>
- Schmidt, J. P., Park, A. W., Kramer, A. M., Han, B. A., Alexander, L. W., & Drake, J. M. (2017). Spatiotemporal Fluctuations and Triggers of Ebola Virus Spillover. *Emerging infectious diseases*, 23(3), 415–422. <https://doi.org/10.3201/eid2303.160101>
- Schon, J., & Koren, O. (2022). Introducing AfroGrid, a unified framework for environmental conflict research in Africa. *Scientific Data*, 9(1), 116. <https://doi.org/10.1038/s41597-022-01198-5>

- Shafie, N. J., Sah, S. A., Latip, N. S., Azman, N. M., & Khairuddin, N. L. (2011). Diversity pattern of bats at two contrasting habitat types along Kerian River, Perak, Malaysia. *Tropical Life Sciences Research*, 22(2), 13–22.
- Sundaram, M., Dorado, M., Akaribo, B., Fillion, A., Han, B. A., Gottdenker, N. L., Schmidt, J. P., Drake, J. M., & Stephens, P. R. (2024). Fruit–frugivore dependencies are important in Ebolavirus outbreaks in Sub-Saharan Africa. *Ecography*, e06950. <https://doi.org/10.1111/ecog.06950>
- Tollefsen, A. F., Strand, H., & Buhaug, H. (2022). Prio-grid: A unified spatial data structure. *Journal of Peace Research*, 49(2), 363–374. <https://doi.org/10.1177/0022343311431287>
- Towner, J. S., Amman, B. R., Sealy, T. K., Carroll, S. A., Comer, J. A., Kemp, A., Swanepoel, R., Paddock, C. D., Balinandi, S., Khristova, M. L., Formenty, P. B. H., Albarino, C. G., Miller, D. M., Reed, Z. D., Kayiwa, J. T., Mills, J. N., Cannon, D. L., Greer, P. W., Byaruhanga, E., Farnon, E. C., Atimnedi, P., Okware, S., Katongole-Mbidde, E., Downing, R., Tappero, J. W., Zaki, S. R., Ksiazek, T. G., Nichol, S. T., & Rollin, P. E. (2009). Isolation of genetically diverse Marburg viruses from Egyptian fruit bats. *PLoS Pathogens*, 5(7), e1000536. <https://doi.org/10.1371/journal.ppat.1000536>
- Towner, J. S., Khristova, M. L., Sealy, T. K., Vincent, J. P., Erickson, B. R., Bawiec, D. A., ... & Nichol, S. T. (2006). Marburgvirus genomics and association with a large hemorrhagic fever outbreak in Angola. *Journal of Virology*, 80(13), 6497–6516. <https://doi.org/10.1128/JVI.00069-06>
- Wallace, R. G., Kock, R., Bergmann, L., Gilbert, M., Hogerwerf, L., Wallace, R., & Holmberg, M. (2016). The dawn of Structural One Health: A new science tracking disease emergence

- along circuits of capital. *Social Science & Medicine*, 129, 68–77.
<https://doi.org/10.1016/j.socscimed.2014.09.047>
- Walsh, P. D., Bermejo, M., & Rodriguez-Teijeiro, J. D. (2009). Disease avoidance and the evolution of primate social connectivity: Ebola, bats, gorillas, and chimpanzees. In H. Huffman & C. Chapman (Eds.), *Primate parasite ecology: The dynamics and study of host–parasite relationships* (pp. 183–198). Cambridge University Press.
- WHO Ebola Response Team. (2016). After Ebola in West Africa — Unpredictable risks, preventable epidemics. *New England Journal of Medicine*, 375(6), 587–596.
<https://doi.org/10.1056/NEJMs1513109>
- Wilcox, B. A., & Ellis, B. (2006). Forests and emerging infectious diseases of humans. *Unasylva*, 57, 11–18.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.
- WorldPop. (2021). Population counts and densities for Africa. *WorldPop Project*.
<https://www.worldpop.org/>

CHAPTER 4

OPTIMAL SMALLHOLDING AND INDUSTRIAL OIL PALM PLANTATIONS: ACCOUNTING FOR THE FILOVIRUS SPILLOVER RISK IN AFRICA

4.1 Introduction

Palm oil is one of the world's most economically significant crops, widely used in processed foods, soaps, cosmetics, and biodiesel. According to the Food and Agriculture Organization (FAO, 2024), global production of oil palm fresh fruit bunches (FFB) has more than tripled, increasing from approximately 120 million tons in 2000 to over 400 million tons in 2020. Similarly, the area cultivated for oil palm has expanded from about 10 million hectares to nearly 29 million hectares over the same period (FAO, 2024; Corley & Tinker, 2016).

The expansion of oil palm plantations in Africa has significantly contributed to economic growth in a region where approximately one-third of the population lived below the poverty line as of 2019 (World Bank, 2022). Over 20 African countries currently cultivate oil palm on nearly 6 million hectares of land, providing critical employment and income opportunities for rural communities (World Economic Forum, 2022). In Cameroon, smallholder oil palm farming generates annual wages of approximately \$1,281 per hectare per household for farmers practicing intercropping, substantially boosting rural incomes (Ayompe et al., 2021). A study from Malawi found that oil palm production simultaneously provides cooking oil for household consumption, creates local employment, and increases cash income for indigenous farming households (Mweta et al., 2025). The industry's development potential is substantial, with estimates suggesting up to 22 million hectares in West and Central Africa could be converted to oil palm plantations in

coming years, potentially transforming regional economies if managed sustainably (World Economic Forum, 2016). Palm oil production also contributes to poverty alleviation in many African countries, helping fulfill Sustainable Development Goal 1 by providing sustainable livelihoods (Ayompe et al., 2021).

Despite its economic importance, the rapid expansion of oil palm plantations has raised significant environmental and public health concerns. The primary ecological issues include deforestation, biodiversity loss, and the increased risk of zoonotic disease spillovers (Faust et al., 2018; Wilkinson et al., 2018). Recent ecological studies suggest a possible link between oil palm plantation expansion and the spread of Filoviruses, specifically Ebola and Marburg viruses, through interactions involving fruit bats, which are identified as primary reservoir hosts crucial to filovirus transmission cycles (Alexander et al., 2015; Leroy et al., 2009). Oil palm plantations offer fruit bats abundant food sources and refuge from heat stress, thereby increasing the likelihood of viral spillover events in human populations, particularly among plantation workers who come into frequent contact with these bats (Shafie et al., 2011; Wallace et al., 2016). Monoculture plantations, in particular, are hypothesized to be more attractive to bats, thereby presenting a heightened risk compared to polyculture plantations (Wallace et al., 2016).

The 2014-2016 West African Ebola epidemic exemplifies the profound socioeconomic and health impacts of zoonotic disease outbreaks in the region. This crisis reduced the combined GDP of Guinea, Liberia, and Sierra Leone by approximately US \$2.2 billion while causing 28,600 confirmed and probable cases with 11,325 deaths (World Bank Group, 2014; WHO, 2016). Ebola's exceptional virulence—characterized by an average case-fatality rate of approximately 50% with historical outbreaks ranging from 25% to 90%—makes it particularly devastating (WHO, 2025). Subsequent outbreaks have demonstrated similarly high mortality rates despite advances in

medical countermeasures: the 2018-2020 Democratic Republic of the Congo outbreak resulted in 3,481 cases and 2,299 deaths despite vaccine availability, while Uganda's 2022 Sudan-ebolavirus episode produced 164 cases with 55 deaths, representing a case-fatality rate of approximately 47% (WHO, 2020, 2023). These recurring epidemics highlight the substantial economic vulnerability associated with zoonotic diseases, underscoring how public health emergencies can rapidly undermine regional development gains and impose significant long-term costs on affected economies.

Despite the clear economic and public health relevance, few studies have integrated economic modeling explicitly to analyze the relationship between environmental change and infectious disease spillovers. Notable exceptions include Barbier (2021), who developed an economic model examining the costs and benefits associated with habitat transformation, explicitly incorporating zoonotic disease transmission from wildlife to humans. Such frameworks underscore the need to balance the economic benefits derived from habitat modifications, including agricultural expansion, against the substantial risks posed by zoonotic disease spillovers (Albers et al., 2020).

Building upon these insights, this study develops an economic model explicitly designed to examine decision-making processes underlying oil palm plantation expansion. The model distinguishes between monoculture and polyculture plantations to derive separate marginal private cost curves reflecting differing agricultural practices and associated ecological risks. Additionally, we quantify the marginal external costs of filovirus spillovers for each plantation type, integrating epidemiological variables such as spillover probabilities, expected numbers of infected individuals, with societal willingness to pay to reduce mortality risk. Using this framework, we numerically simulate the effectiveness of taxation policies intended to internalize these

externalities, comparing unregulated market outcomes to socially optimal solutions. Ultimately, our analysis aims to enhance understanding of the economic trade-offs associated with palm oil cultivation in Africa and inform more proactive and sustainable agricultural expansion and disease management policies.

4.2 Theoretical Framework and Model

4.2.1 Set up and Assumptions

This section develops an economic model that captures both the private production incentives and environmental externalities associated with crude palm oil (CPO) production in Africa. The model distinguishes between two producer types with fundamentally different ecological characteristics: smallholder oil palm plantations (S) and industrial oil palm plantations (L). To maintain analytical tractability while preserving the essential economic mechanisms, the model builds on five core assumptions:

Assumption 1: Divergent Cultivation Systems

Industrial oil palm plantations (L) employ monoculture systems, converting large contiguous areas into uniform stands of high-yielding oil palm varieties. In contrast, smallholder farms (S) typically adopt polyculture approaches, integrating oil palms with other crops and maintaining more diverse landscape elements (Carrère, 2011). This structural difference drives divergent ecological impacts: monocultures create extensive homogeneous habitats with synchronized fruiting patterns that attract fruit bats (Alexander et al., 2015; Wallace et al., 2016). Polyculture systems, with their mixed species composition, irregular canopy structures, and retained forest fragments, might dilute bat visitation rates and reduce human-bat contact opportunities.

Assumption 2: Price-Taking Behavior

African CPO producers account for a relatively small share of global supply (6.3 % in 2023, FAOSTAT), so each farm believes it cannot influence P . Profit maximization is therefore equivalent to choosing output q_i (or, equivalently, cultivated area A_i) such that marginal cost equals the exogenous price.

Assumption 3: Leontief production technology with fixed land–labor ratio

Oil-palm production is subject to strict biophysical and managerial constraints: every hectare must be pruned, fertilized, weeded, and—most labor-intensively—harvested at 7- to 10-day intervals (Corley & Tinker 2016). Field surveys confirm that estates and smallholders therefore allocate workers in *fixed* proportions to cropped areas (Kubitza & Krishna 2020). We capture this one-to-one linkage with a Leontief production function in which output is limited by the scarcest of three inputs—land (A_i), hired labor (L_i), and a bundle of other inputs:

$$q_i = \beta_i \min\{A_i, \alpha_i L_i, \bar{K}_i\}, i \in \{S, L\} \quad (1)$$

where β_i is the productivity or physical yield, α_i fixes the land-labor ratio. The composite $\bar{K}_i$ lumps together non-labor inputs that are either proportional to area (fertilizer, pesticide) or sunk (mill equipment) and will therefore be folded into the cost function below.

Assumption 4: Increasing Marginal Land-Preparation Costs

The per-hectare cost of establishing new plantations increases as cultivation expands into less favorable locations—areas with poorer soils, challenging topography, or greater distance from existing infrastructure. We model this through a linear marginal preparation cost function:

$$CP_i(A_i) = a_i + b_i A_i$$

Integrating from 0 to A_i yields a quadratic total preparation cost:

$$CP_i(A_i)A_i = \int_0^{A_i} (a_i + b_i A_i) dA_i = a_i A_i + \frac{b_i A_i^2}{2} \quad (2)$$

Economically, the parameter a_i embodies a first-mover advantage: when $a_i < 0$, it implies that the very first hectares converted are unusually cheap because they lie on prime land—fertile soils, gentle topography, and close to existing roads or milling facilities—so small expansions can be undertaken at minimal cost. By contrast, $b_i > 0$ captures the increasing marginal difficulty of land preparation: as cultivation radiates outward, plots become progressively less favorable (poorer soils, steeper terrain) and farther from infrastructure, driving up per-hectare clearing, drainage, and transport expenses.

Assumption 5: Competitive labor market

Producers hire workers at the prevailing wage w . The Leontief technology fixes labor requirements at $L_i = q_i/(\alpha_i \beta_i)$, yielding total wage costs that scale linearly with output.

Combining these elements, the annual total cost function for producer type i is:

$$TC_i = FC_i + a_i A_i + \frac{b_i A_i^2}{2} + w L_i \quad (3)$$

where

- FC_i is fixed costs, non-area-dependent outlays such as perennial tree establishment, road upgrades, certification fees, and management salaries;
- $a_i A_i + \frac{b_i A_i^2}{2}$ is the land-preparation component;
- $w L_i$ is the hired-labor cost.

Using the yield relation $q_i = \beta_i A_i$ and labor requirement $L_i = q_i/(\alpha_i \beta_i)$, equation (3) can be rewritten in output terms, which sets the stage for deriving marginal and firm-level supply functions, and ultimately the aggregate market supply in Section 4.2.2.

4.2.2 Supply curves

This section derives the supply functions for smallholder (S) and industrial (L) oil palm producers, capturing their respective production decisions under the technological and cost constraints established in Section 4.2.1.

Firm-Level Supply Functions

Each producer maximizes profit given the world price P . The profit function is given by

$$\pi_i = Pq_i - FC_i - a_i A_i - \frac{b_i A_i^2}{2} - wL_i, i \in \{S, L\} \quad (4)$$

where A_i is the cultivated area, L_i the hired labor (at wage w), and the quadratic term captures rising land-preparation costs. In equation (4), “ i ” indexes the type of plantation, smallholder or industrial, $i \in \{S, L\}$. For simplicity we assume representative firms within each type of plantation.

Under the Leontief technology, output relates to cultivated area through $q_i = \beta_i A_i$, while labor requirements follow $L_i = q_i / (\alpha_i \beta_i)$. Substituting these relationships into the profit function and differentiating with respect to output q_i yields the first-order condition for profit maximization:

$$P = MC_i(q_i) = \frac{b_i}{\beta_i^2} q_i + \frac{a_i}{\beta_i} + \frac{w}{\alpha_i \beta_i} \quad (5)$$

This condition states that each producer expands output until marginal cost equals the market price. Solving for q_i provides the firm's supply function:

$$q_i = \frac{\beta_i^2 P}{b_i} - \frac{\beta_i}{b_i} \left(a_i + \frac{w}{\alpha_i} \right) \quad (6)$$

Since $b_i > 0$ for both producer types, these supply functions exhibit positive price responsiveness, in other words, production expands as world prices increase.

Break-even prices and market participation

Each supply function crosses zero output at a critical break-even price where revenue exactly covers marginal cost:

$$P_i^{BE} = \frac{a_i + \frac{w}{\alpha_i}}{\beta_i} \quad (7)$$

Production begins only when the market price exceeds this threshold. Based on stylized facts of oil palm cultivation in Africa, we anticipate that $P_L^{BE} < P_S^{BE}$ because industrial estates typically have higher yield per hectare ($\beta_L > \beta_S$) and more efficient labor utilization ($\alpha_L > \alpha_S$). This ordering implies that industrial producers enter the market at lower price points than smallholders, reflecting their productivity advantages and economies of scale.

Aggregate market supply

Combining the firm-level supply functions yields a piecewise, "kinked" market supply function that reflects sequential market entry as prices rise:

$$Q(P) = \begin{cases} 0, & P < P_L^{BE} \\ q_L(P), & P_L^{BE} \leq P < P_S^{BE} \\ q_L(P) + q_S(P), & P \geq P_S^{BE} \end{cases} \quad (8)$$

This aggregate supply function captures three distinct market regimes:

1. No production ($P < P_L^{BE}$): When prices fall below the industrial break-even threshold, no production occurs as neither producer type can cover marginal costs.
2. Industrial-only supply ($P_L^{BE} \leq P < P_S^{BE}$): As prices rise above the industrial break-even point but remain below the smallholder threshold, only industrial estates operate, leveraging their cost advantages.

3. Dual-producer supply ($P \geq P_S^{BE}$): Once prices exceed the smallholder break-even level, both producer types actively participate in the market, with their relative contributions determined by their respective supply elasticities.

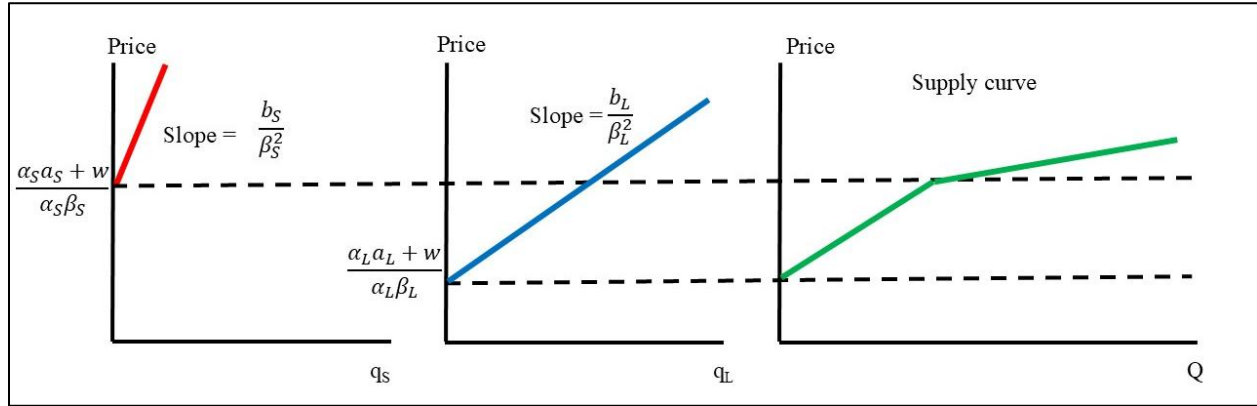


Figure 4.1: supply curves

This sequential entry pattern aligns with observed market dynamics in African palm oil producing regions, where industrial producers often establish operations before smallholder sectors develop (Reference). The model thus captures not only the volume response to price changes but also the evolving composition of production across different cultivation systems—a critical distinction when considering zoonotic spillover risks that vary by producer type.

4.2.3 Market Equilibrium

Building on the supply curves developed in Section 4.2.2, this section characterizes the unregulated market equilibrium that emerges when producers respond to the exogenous world price without internalizing spillover externalities. This equilibrium establishes the benchmark against which we will evaluate policy interventions.

Equilibrium Conditions

Under Assumptions 1-5, each producer takes the world price P^* as given and selects output to equate marginal cost with this price. The resulting market equilibrium depends critically on how P^* compares to the break-even thresholds P_L^{BE} and P_S^{BE} (equation 8).

These thresholds partition the equilibrium into three possible regimes, each with distinct production patterns and welfare implications.

Regime-Specific Equilibria

1. No production ($P^* < P_L^{BE}$)

When the world price falls below the industrial break-even threshold, neither producer type covers its marginal costs. The market-clearing quantities are:

$$Q_S^0 = Q_L^0 = 0$$

This corner solution represents periods of extreme price depression in global palm oil markets, during which new plantation development halts and existing operations may be temporarily abandoned or converted to alternative crops.

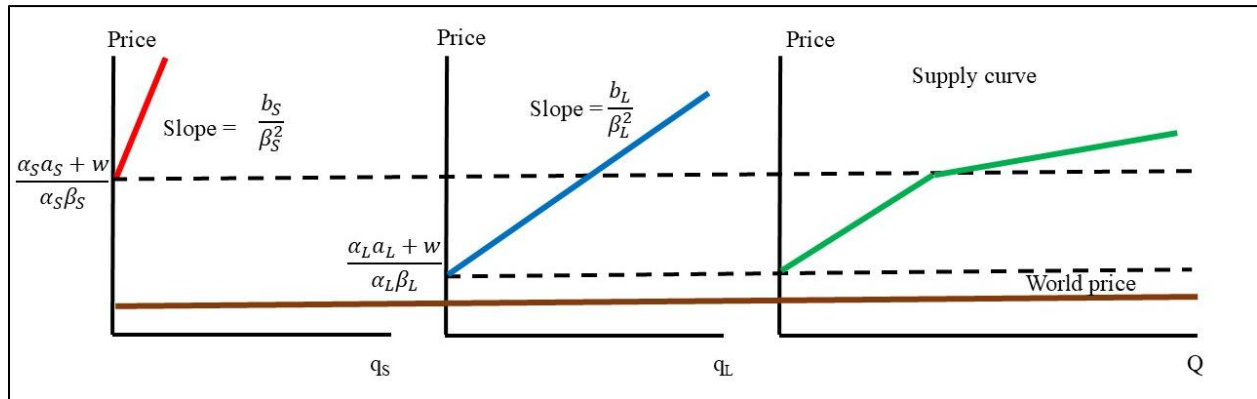


Figure 4.2: No production

2. Industrial-only supply ($P_L^{BE} \leq P^* < P_S^{BE}$)

As prices rise above the industrial break-even point but remain below the smallholder threshold, only industrial plantations operate profitably. Their equilibrium output is:

$$Q_L^0 = q_L(P^*) = \frac{\beta_L^2 P^*}{b_L} - \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right)$$

while smallholder production remains dormant:

$$Q_S^0 = 0$$

This intermediate regime illustrates how industrial producers' technological and scale advantages enable them to maintain production even during periods of moderate price compression, while smallholders remain excluded from the market.

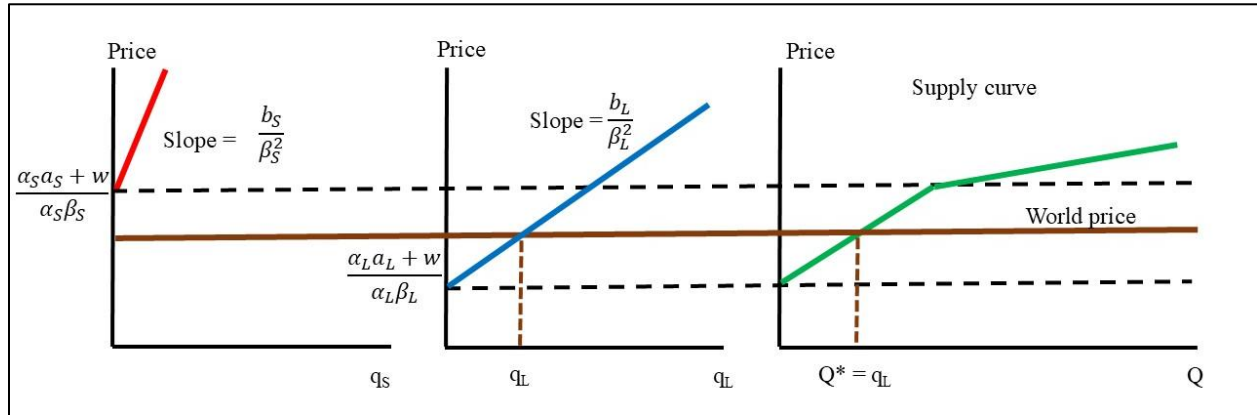


Figure 4.3: Industrial-only supply

3. Both types supplying ($P^* \geq P_S^{BE}$)

Once the world price exceeds the smallholder break-even threshold, both producer types participate in the market. Their respective equilibrium quantities are:

$$Q_S^0 = \frac{\beta_S^2 P^*}{b_S} - \frac{\beta_S}{b_S} \left(a_S + \frac{w}{\alpha_S} \right), Q_L^0 = \frac{\beta_L^2 P^*}{b_L} - \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right)$$

yielding a total regional supply of:

$$Q^0 = Q_S^0 + Q_L^0$$

This regime represents the most common market state in periods of stable or rising palm oil prices, with both producer types actively contributing to aggregate supply

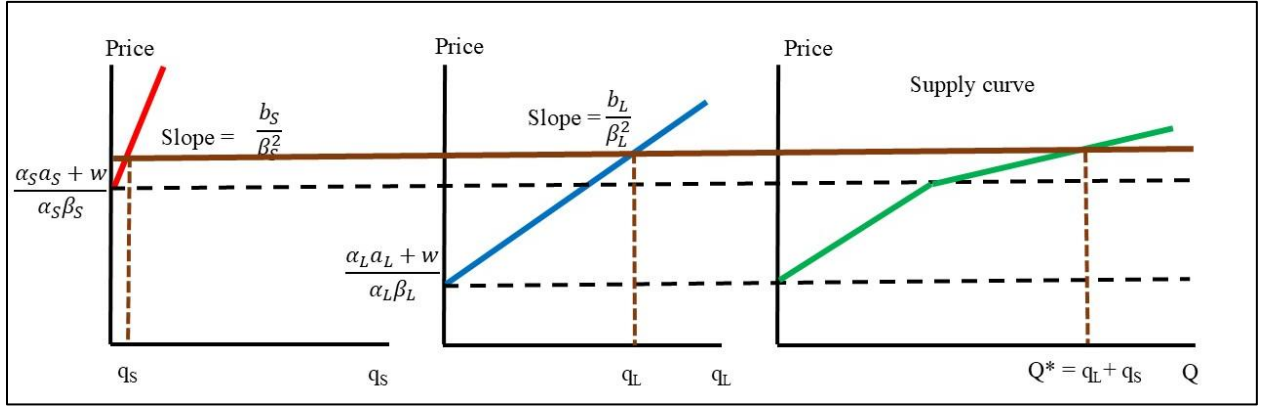


Figure 4.4: Both types supplying

Consolidating these regime-specific outcomes, the unregulated market equilibrium can be expressed as a piecewise function:

$$Q^0 = Q(P^*) = \begin{cases} 0, & P^* < P_L^{BE} \\ \frac{\beta_L^2 P^*}{b_L} - \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right), & P_L^{BE} \leq P^* < P_S^{BE} \\ \left(\frac{\beta_L^2}{b_L} + \frac{\beta_S^2}{b_S} \right) P^* - \left[\frac{\beta_S}{b_S} \left(a_S + \frac{w}{\alpha_S} \right) + \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right) \right], & P^* \geq P_S^{BE} \end{cases} \quad (9)$$

Given Africa's status as a price-taker in global palm oil markets, P^* is determined exogenously, and Q^0 reflects the unconstrained, unregulated allocation of land, labor, and output across producer types

Implications for Land Allocation and Ecological Impact

This unregulated equilibrium has significant implications for both economic welfare and ecological risk. In Regime 3, the most relevant for contemporary African palm oil markets, the relative contribution of each producer type to total output is determined by their respective supply elasticities, which in turn depend on technological parameters (β_i, α_i) and cost structures (a_i, b_i)

Crucially, this equilibrium makes no provision for the differential spillover risks associated with each production system. Therefore, these costs remain unpriced in the unregulated equilibrium, leading to systematic undervaluation of the ecological services provided by more diversified production systems.

4.2.4 External Cost of Filovirus Spillover

In the unregulated market equilibrium described in Section 4.2.3, producers make decisions based solely on private costs and benefits, disregarding the potential public health externalities associated with filovirus transmission. This section develops a framework to quantify these external costs, enabling their integration into a social cost function and subsequent policy analysis.

Conceptualizing Externality

Filoviruses (Ebola and Marburg virus) represent significant public health threats with high mortality rates, substantial economic impacts, and complex ecological dynamics. Recent epidemiological evidence suggests that fruit bats serve as natural reservoir hosts for these viruses (Leroy et al., 2005), with oil palm plantations potentially creating ecological conditions that facilitate spillover to humans (Olivero et al., 2017). According to the previous chapter, this risk

appears to vary systematically between cultivation systems, with monoculture industrial plantations providing more attractive habitat for bat aggregation than polyculture smallholder systems.

The externality arises because individual producers have no market incentive to factor the epidemiological consequences of their land-use choices into production decisions, yet society ultimately bears the costs. During the 2014-2016 West African Ebola epidemic, the direct health burden reached 28 646 confirmed, probable, and suspected cases and 11 323 deaths across Guinea, Liberia, and Sierra Leone (WHO, 2016). The indirect economic fallout was equally stark: the World Bank estimated a regional GDP shortfall of US \$2.2 billion in 2014 and projected losses of up to US \$32.6 billion for 2014-2015 under a high-transmission scenario (World Bank.,2014), while mobile-phone surveys showed that nearly half of Liberia's workforce was no longer employed by November 2014 (World Bank.,2014b)

Quantifying Spillover Risk

To incorporate this externality into our economic framework, we model the total external cost (TEC) for each producer type as a function of cultivated area, reflecting how plantation expansion influences ecological conditions and human-wildlife interfaces that facilitate virus transmission. Let A_i denotes the total hectares under smallholder (S) and industrial oil palm cultivation (L). We introduce the quadratic term because Chapter 2's empirical spillover model revealed a statistically significant polynomial relationship between plantation area and filovirus risk. Therefore, allowing A_i to enter as both a linear and a squared term in equation (10). The total external cost for producer type $i \in \{S, L\}$ is:

$$TEC_i = (\rho 1_i \times A_i + \rho 2_i \times A_i^2) \times WTP \times N \quad (10)$$

where:

- $\rho 1_i$ and $\rho 2_i$ are empirically estimated coefficients capturing the linear and quadratic components of spillover probability per hectare.
- WTP represents society's willingness-to-pay per statistical life (VSL). We adopt and weight the benchmark VSL from Viscusi and Masterman (2017 which translates a marginal reduction in spillover probability into its expected mortality-cost equivalent by valuing the prevention of one "statistical" death.
- N denotes the expected number of infections in a spillover event (based on historical outbreak data)

Marginal External Cost Functions

Differentiating TEC_i with respect to area A_i yields the marginal external cost (MEC), which we decompose into two components for analytical clarity:

$$MEC_i(A_i) = (\rho 1_i \times WTP \times N) + (2\rho 2_i \times A_i \times WTP \times N) \quad (11)$$

We define:

$$MEC1_i(A_i) = \rho 1_i \times WTP \times N, \quad MEC2_i(A_i) = 2\rho 2_i \times A_i \times WTP \times N$$

The constant term $MEC1_i$ represents the baseline per-hectare external cost regardless of scale, while the area-dependent term $MEC2_i$ captures potential non-linearities in transmission dynamics. If $\rho 2_i > 0$, the marginal external cost increases with plantation area, reflecting accelerating risk as ecological transformations intensify or as human-wildlife interfaces expand. Conversely, if $\rho 2_i < 0$, marginal risk might decline with scale, potentially due to economies of scale in disease surveillance or changes in ecological interactions at larger scales.

Integrated Social Cost Function

To fully account for both private production costs and public health externalities, we construct the marginal social cost (MSC) function for each producer type by adding the external cost components to the private marginal cost derived in Section 4.2.2:

Since output and area are linked through the yield parameter ($q_i = \beta_i A_i$), this integrated cost function provides the basis for determining the socially optimal allocation of production across plantation types. By explicitly differentiating the external costs associated with each production system, the model can identify efficiency-enhancing policy interventions that account for the heterogeneous nature of spillover risk.

Empirical Calibration of Risk Parameters

The parameterization of spillover risk coefficients ρ_{1i} and ρ_{2i} draws on emerging epidemiological evidence regarding the ecological mechanisms of filovirus transmission. Wallace et al. (2016) suggest that industrial monocultures may create more favorable conditions for bat aggregation through synchronized fruit availability and simplified canopy structures that facilitate roosting. In contrast, smallholder polycultures appear to maintain more diverse ecological structures that potentially diffuse bat concentrations and reduce human-wildlife contact opportunities.

The full empirical derivation of these risk parameters is detailed in Appendix B, drawing on statistical analyses that relate historical filovirus outbreak locations to plantation characteristics while controlling for confounding variables. These empirically grounded parameters enable the model to capture the nuanced relationship between cultivation practices and disease risk, providing a foundation for evidence-based policy interventions.

In Section 4.2.5, we will leverage this integrated social cost function to derive the socially optimal equilibrium that internalizes these externalities, setting the stage for evaluating potential policy instruments in Section 4.2.6.

4.2.5 Social Optimum

In contrast to the unregulated market equilibrium, a social planner would incorporate both private production costs and the external costs of filovirus spillover when determining the optimal allocation of resources. This section derives the socially optimal equilibrium, providing a normative benchmark against which policy interventions can be evaluated.

The Social Planner's Problem

The social planner seeks to maximize net social welfare by choosing cultivated areas (A_S, A_L) that equate marginal social cost with the world price P for each producer type $i \in \{S, L\}$. Combining the private marginal cost function from Section 2.2 with the external cost components from Section 4.2.4 yields the comprehensive marginal social cost:

$$MSC_i(q_i) = \frac{b_i}{\beta_i^2} q_i + \frac{a_i}{\beta_i} + \frac{w}{\alpha_i \beta_i} + (\rho 1_i + 2\rho 2_i \times A_i) \times WTP \times N \quad (13)$$

Given the technological relationship $q_i = \beta_i A_i$, we can express this cost function in terms of output:

$$MSC_i(q_i) = \frac{b_i}{\beta_i^2} q_i + \frac{a_i}{\beta_i} + \frac{w}{\alpha_i \beta_i} + (\rho 1_i + 2\rho 2_i \times \frac{q_i}{\beta_i}) \times WTP \times N \quad (14)$$

Socially Optimal Output

Setting price equal to marginal social cost provides the planner's first-order condition:

$$P = \frac{b_i}{\beta_i^2} q_i + \frac{a_i}{\beta_i} + \frac{w}{\alpha_i \beta_i} + (\rho 1_i + 2\rho 2_i \times \frac{q_i}{\beta_i}) \times WTP \times N \quad (15)$$

Solving for the social-optimum output q_i^{SO} :

$$q_i^{SO} = \frac{\beta_i^2 [P - \frac{a_i}{\beta_i} - \frac{w}{\alpha_i \beta_i} - (\rho 1_i \times WTP \times N)]}{b_i + 2\rho 2_i \times \beta_i \times WTP \times N} \quad (16)$$

This expression reveals several important insights about how spillover risk affects the socially optimal allocation:

1. Baseline risk effect: The term $(\rho 1_i \times WTP \times N)$ in the numerator functions as an additional fixed cost per unit of output, effectively raising the break-even price and reducing optimal production levels.
2. Area-dependent risk effect: When $\rho 2_i > 0$, the denominator term $(2\rho 2_i \times \beta_i \times WTP \times N)$ augments the slope of the marginal cost curve, making supply less elastic and further constraining optimal output.
3. Differential impacts across producer types: If spillover risk parameters differ between industrial and smallholder plantations as suggested by ecological evidence, the socially optimal allocation will shift production toward the system with lower external costs, potentially altering the composition of aggregate supply.

For any given price P , if both $b_i > 0$ and $\rho 2_i > 0$, then $q_i^{SO} < q_i^0$ —internalizing spillover risk contracts each producer's supply curve leftward relative to the unregulated equilibrium. The magnitude of this contraction depends on the specific risk parameters, valuation of health impacts, and productivity characteristics of each production system.

Social-Optimum Supply Function

Aggregating across smallholder and industrial plantations, the social-optimum supply function takes the form:

$$Q^{SO}(P) = \begin{cases} 0, & P < P^{SO}, \\ q_L^{SO}(P), & P_L^{SO} \leq P < P_S^{SO}, \\ q_S^{SO}(P) + q_L^{SO}(P), & P \geq P_L^{SO}, \end{cases} \quad (17)$$

where each break-even price P_i^{SO} is the value of P that makes $q_i^{SO} = 0$ in the optimal output equation. These socially optimal break-even thresholds exceed their unregulated counterparts by an amount proportional to the baseline marginal external cost:

$$P_i^{SO} = P_i^{BE} + \rho 1_i \times WTP \times N \quad (18)$$

This upward shift reflects the additional social costs that must be covered for production to generate maximum net social benefits.

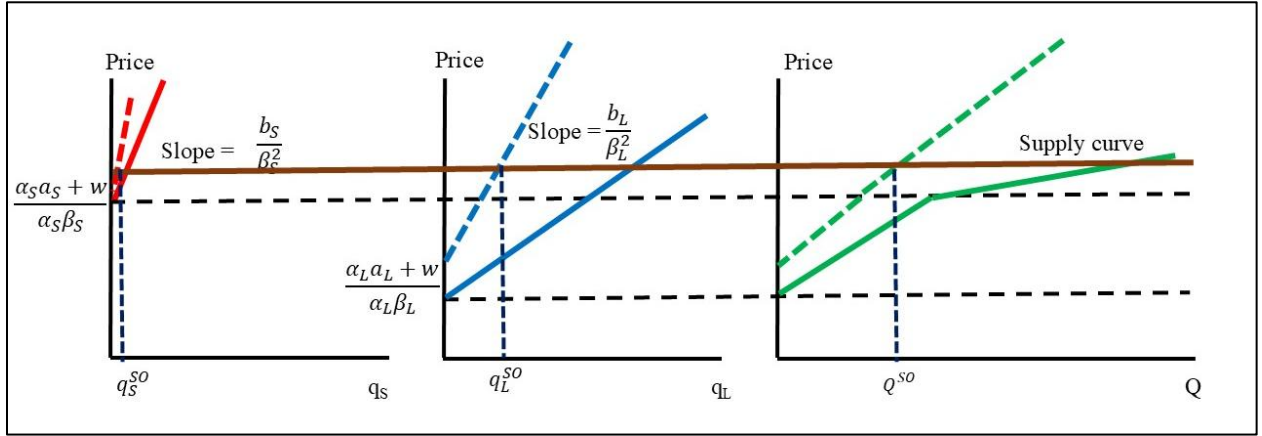


Figure 4.5 Social-Optimum Supply Function (all produce)

4.2.6 Policy Instrument: Optimal Uniform Unit Tax

While the social planner's solution in Section 4.2.5 provides a theoretical benchmark, implementing differentiated policies for each producer type presents significant practical challenges. In practice, crude palm oil marketed internationally is functionally identical regardless of its production system origins. Once processed and entering the supply chain, oil from smallholder polycultures cannot be distinguished from that produced in industrial monocultures without elaborate and costly traceability systems. This fungibility creates a fundamental

implementation constraint: any workable policy must operate without requiring verification of producer type at the point of taxation.

Thus, instead of attempting to implement producer-specific Pigouvian taxes (which would require perfect monitoring of production sources), we propose a single, uniform specific tax t (USD per ton) applied equally to all crude palm oil production. The objective is to identify the tax rate that, when applied to the world price, induces private producers to collectively supply the socially optimal quantity of output.

Mathematically, the optimal uniform tax must satisfy:

$$Q^0(P - t) = Q^{so}(P) \quad (19)$$

That is, the aggregate private supply at the net-of-tax price $(P - t)$ should equal the planner's target supply at the world price P . This condition ensures that the tax adjusts the effective price signal to align private incentives with the social optimum.

Substituting the linear supply functions derived in equations 6 and 16, the tax condition becomes:

$$q_S^{so}(P) + q_L^{so}(P) = \frac{\beta_S^2}{b_S} + \frac{\beta_L^2}{b_L}(P - t) - \left[\frac{\beta_S}{b_S} \left(a_S + \frac{w}{\alpha_S} \right) + \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right) \right] \quad (20)$$

Solving for t yields:

$$t = P - \left(\frac{Q^{so}(P) + \left[\frac{\beta_S}{b_S} \left(a_S + \frac{w}{\alpha_S} \right) + \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right) \right]}{\left(\frac{\beta_S^2}{b_S} + \frac{\beta_L^2}{b_L} \right)} \right) \quad (21)$$

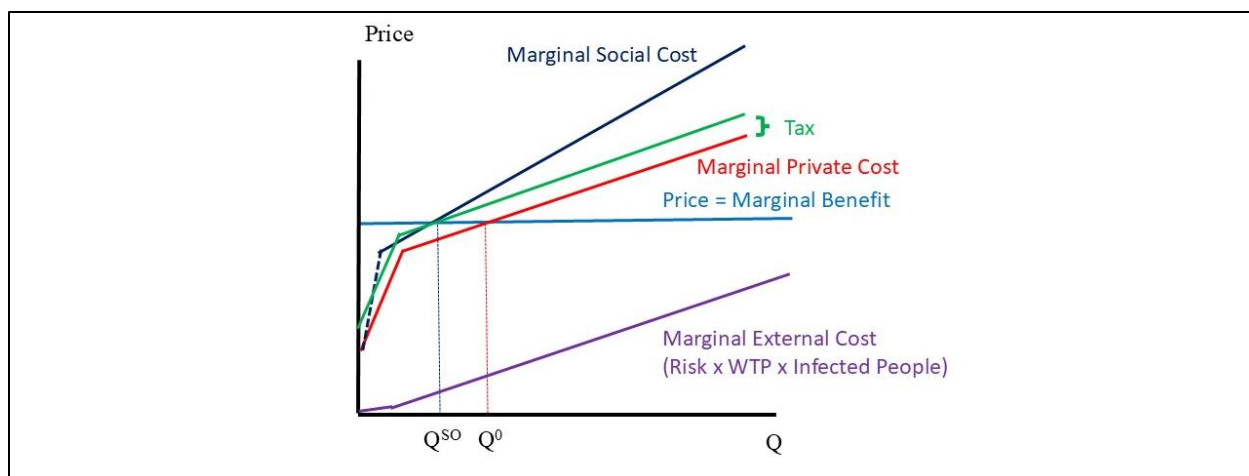


Figure 4.6: Optimal Uniform Unit Tax

This formula provides a computable tax rate based on observable market parameters and the socially optimal output levels derived from the spillover risk model.

4.3. Numerical Simulation and Results

This section operationalizes the theoretical framework developed in Section 2 through empirical parameterization and quantitative analysis. We calibrate the model using real-world data from Africa and simulate both the unregulated market equilibrium and the socially optimal allocation under the uniform tax policy.

4.3.1 Data & Parameterization

Our model calibration integrates multiple data sources, including remote sensing estimates, agricultural statistics, economic parameters, and epidemiological risk assessments. Table 1 summarizes key parameter values and sources.

Economic and Agricultural Parameters

World Market Price (P): To represent the long-run price signal that guides planting decisions, we use the international crude-palm-oil (CPO) benchmark averaged over 1997–2018,

reported by the IMF and archived in the Federal Reserve’s FRED database (series: *PPOILUSDM*) as shown in Figure 4.7. The 22-year arithmetic mean is \$589.68 per ton.

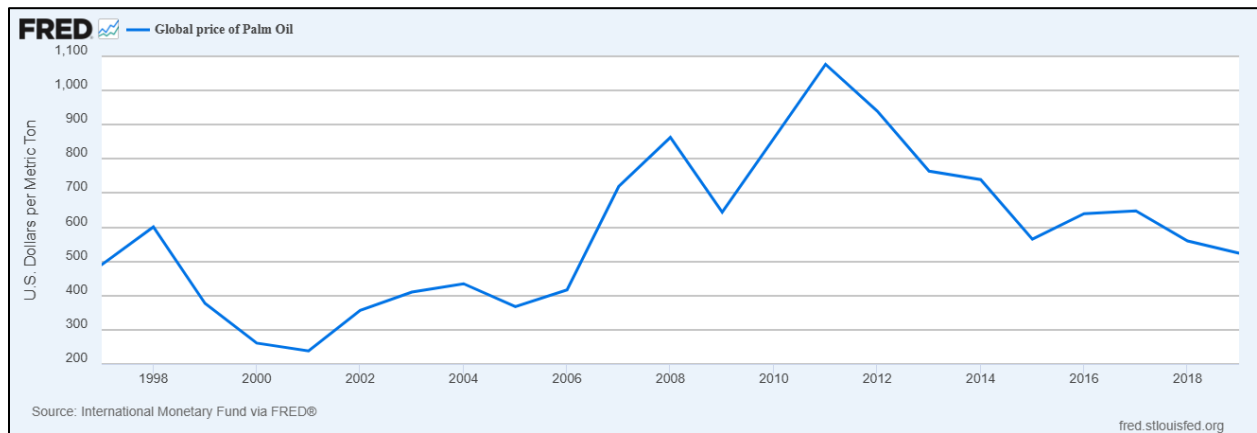


Figure 4.7: International Crude Palm Oil Prices, 1997–2018

Source: International Monetary Fund via FRED

Productivity (β_i): Drawing on field studies by Ayompe et al. (2021), we specify differential productivity parameters: 7 tons/hectare/year for smallholder polycultures (β_S) and 20 tons/hectare/year for industrial monocultures (β_L). This study conducted extensive field research in Ghana, West Africa—a region where smallholders account for over 60% of oil palm cultivation—systematically measuring yields across both smallholder and industrial systems.

Land-Labor Ratio (α_i): Based on detailed labor utilization studies by Zapata-Hernández et al. (2024), we parameterize the labor efficiency as 7.1 hectares/worker for smallholders (α_S) and 11.3 hectares/worker for industrial producers (α_L). These parameters capture the greater labor intensity of smallholder systems, which typically employ less mechanization and more diversified management practices.

Wage (w): The rural agricultural wage is set at €1.60 per day (Carrère, 2010), converting to \$432.96 per year assuming 260 working days. This wage parameter reflects the prevailing labor market conditions in oil palm producing regions of Africa.

Cost Function Parameters

Land-Preparation Cost Parameters (a_i, b_i): These crucial parameters were derived econometrically through ordinary least squares (OLS) estimation of harvested area regressed on three-year lagged world price, capturing the dynamic planting response to price signals. This lag structure reflects the biological reality that planting decisions are influenced by prices prevailing approximately three years earlier, given the preparation, financing, and establishment phases of oil palm cultivation.

The empirical analysis (detailed in Appendix B) yields the following parameter estimates:

- For smallholders: $a_S = -2,612.79$, $b_S = 0.02608$
- For industrial producers: $a_L = -9,396.20$, $b_L = 0.10684$

The negative intercept terms a_i indicate first-mover advantages where initial hectares have unusually low preparation costs due to favorable conditions. The positive slope parameters b_i capture increasing marginal costs as expansion proceeds into less favorable terrain.

Spillover Risk Parameters

Risk Coefficients ($\rho 1_i, \rho 2_i$) These parameters quantify the relationship between plantation area and filovirus spillover probability, derived from the empirical analysis presented in Chapter 2 and further detailed in Appendix B. The estimated values are:

- For smallholders: $\rho 1_S = -9.58 \times 10^{-8}$, $\rho 2_S = 2.53 \times 10^{-13}$
- For industrial producers: $\rho 1_L = 9.47 \times 10^{-8}$, $\rho 2_L = 0$

Notably, these parameters suggest contrasting risk profiles between production systems. Smallholder polycultures exhibit an initially negative baseline risk (perhaps due to diversified landscapes serving as ecological buffers) but a strong positive quadratic term indicating accelerating risk at larger scales. In contrast, industrial monocultures demonstrate a positive linear

risk component but no significant quadratic effect, suggesting a constant marginal increase in risk with expansion.

Health Impact Valuation

Value of a Statistical Life (WTP): Following Viscusi & Masterman (2017), we adopt a region-appropriate value of \$107,000 per statistical life, reflecting the income-adjusted willingness-to-pay to prevent mortality in lower-income countries.

Expected Infections (N): Based on historical epidemic data, we set $N = 30,000$ expected infections per spillover event, approximating the scale of the 2014-2016 West African Ebola outbreak.

Parameter Summary

Table 4.1 summarizes the core calibration parameters, providing the empirical foundation for our subsequent analysis:

Table 4.1: Summary of Core Calibration Parameters

Parameter	Symbol	Value	Source	Notes
World price	P	\$589.68 / ton	FAOSTAT (2020)	Global market price for palm oil
Productivity (smallholder)	β_S	7 tons / ha / year	Ayompe et al. (2021)	Average yield for smallholder plantations in Cameroon
Productivity (industrial)	β_L	20 tons / ha / year	Ayompe et al. (2021)	Average yield for large industrial plantations in Cameroon
Land-labour ratio (S)	α_S	7.1 ha / worker	Zapata-Hernández et al. (2024)	Fixed labor input requirement for smallholders
Land-labour ratio (L)	α_L	11.3 ha / worker	Zapata-Hernández et al. (2024)	Fixed labor input requirement for industrial plantations
Wage	w	\$432.96 per year	Carrère (2010)	Annual rural agricultural wage (not inflation-adjusted)
Marginal cost intercept of smallholder	a_S	-2612.78 / ha	Appendix B	Fixed component of smallholder marginal cost function
Marginal cost slope of smallholder	b_S	0.02608 / ha ²	Appendix B	Rate of increase in marginal cost with area for smallholders
Marginal cost intercept of industrial	a_L	-9396.20 / ha	Appendix B	Fixed component of industrial marginal cost function
Marginal cost slope of industrial	b_L	0.10684 / ha ²	Appendix B	Rate of increase in marginal cost with area for industrial plantations
Spillover linear of smallholder	ρ_{1S}	-9.58×10^{-8}	Appendix C	Linear term in smallholder disease spillover function
Spillover quadratic of smallholder	ρ_{2S}	2.53×10^{-13}	Appendix C	Quadratic term in smallholder disease spillover function
Spillover linear of industrial	ρ_{1L}	9.47×10^{-8}	Appendix C	Linear term in industrial disease spillover function
Spillover quadratic of industrial	ρ_{2L}	0	Appendix C	No quadratic spillover effect for industrial plantations
VSL (WTP)	WTP	\$107,000 / life	Viscusi & Masterman (2017)	Weighted value of statistical life for low-income countries
Expected infections	N	30,000 persons	2014–16 Ebola epidemic data	Peak West-Africa epidemic size

Source: Author’s compilation

4.3.2 Baseline Equilibria

This section analyzes the unregulated market equilibrium—the "business-as-usual" scenario where producers maximize profits without internalizing the external costs of filovirus spillover. This baseline serves as the reference point against which we will evaluate the welfare gains from policy intervention.

Under the unregulated scenario, both smallholder and industrial producers respond to the prevailing world price of crude palm oil, $P^0 = \$589.68$ per ton

This price represents the long-term average that drives current planting decisions, which will reach full production after a lag. The parameterization of the theoretical model, as explained in Appendix B, recognizes these lags.

Firm-Level Output

Applying the calibrated parameters from equation 6 to the linear supply functions, we compute equilibrium output levels:

Smallholder Production (S):

$$q_S^0 = \frac{\beta_S^2}{b_S} P^0 - \frac{\beta_S}{b_S} \left(a_S + \frac{w}{\alpha_S} \right) \approx 1,792,587 \text{ tons}$$

Industrial Production (L):

$$q_L^0 = \frac{\beta_L^2}{b_L} P^0 - \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right) \approx 3,959,376 \text{ tons}$$

These calculations represent the steady-state equilibrium output levels that emerge when producers respond to sustained price signals over multiple planting cycles. The significant contribution of industrial plantations—approximately 69% of total regional production despite covering only 44% of the total oil palm area—reflects their substantially higher productivity (20 tons/ha versus 7 tons/ha for smallholders).

Aggregate Market Output

The total unregulated market supply combines production from both systems:

$$Q^0 = q_S^0 + q_L^0 \approx 5,751,963 \text{ tons}$$

This baseline output reflects the long-run market response to current economic incentives without accounting for filovirus spillover externalities. It represents the culmination of planting decisions made approximately six years prior, which have now reached mature production levels.

Land Allocation

The cultivated area for each producer type follows directly from output and productivity parameters:

Smallholder Area:

$$A_S^0 = \frac{q_S^0}{\beta_S} = \frac{1,792,587}{7} \approx 256,463.7 \text{ hectares}$$

Industrial Area:

$$A_L^0 = \frac{q_L^0}{\beta_L} = \frac{3,959,376}{20} \approx 198,233.7 \text{ hectares}$$

Total Area:

$$A^0 = A_S^0 + A_L^0 \approx 454,053 \text{ hectares}$$

This distribution of cultivated area represents the cumulative result of planting decisions over multiple years, with each annual cohort at a different stage of maturity. The substantial proportion of land under industrial monoculture (43.6% of total area) creates extensive habitats potentially favorable to fruit bat aggregation, while the smallholder polyculture systems maintain more heterogeneous landscapes that may modulate disease transmission dynamics.

Labor demand

The Leontief technology specification ties labor requirements directly to output through fixed land-labor ratios:

Smallholder Labor:

$$L_S^0 = \frac{q_S^0}{\alpha_s \beta_s} = \frac{1,792,587}{7 \times 7.1} \approx 36,068 \text{ workers},$$

Industrial Labor:

$$L_L^0 = \frac{q_L^0}{\alpha_l \beta_l} = \frac{3,959,376}{20 \times 11.3} \approx 17,519 \text{ workers},$$

Total Labor:

$$L^0 = L_S^0 + L_L^0 \approx 53,587 \text{ workers}.$$

These employment figures highlight the substantial rural livelihoods supported by oil palm cultivation. Notably, smallholder systems generate approximately 67% of total employment despite accounting for only 31% of production, underscoring their greater labor intensity and potential socioeconomic benefits beyond raw output value.

4.3.3 Spillover-inclusive Optimum

This section develops the socially optimal allocation that accounts for both private production costs and the external costs of filovirus spillover. By comparing this social optimum with the unregulated baseline, we can quantify the efficiency gains achievable through policy intervention.

The Social Planner's Target

To implement the social planner's preferred allocation in a decentralized market context, we first compute the socially optimal output levels for each producer type based on the theoretical framework in Section 4.2.5. We then determine the uniform tax rate that, when applied to the world price, will contract private supply to match the social optimum.

Computing Socially Optimal Outputs

Using equation (16) from Section 4.2.5, we calculate each producer's socially optimal output at the prevailing world price $P^0 = \$589.68$ per ton:

Smallholder Oil Palm Plantations:

$$q_S^{SO} = \frac{\beta_S^2 [P^0 - \frac{a_S}{\beta_S} - \frac{w}{\alpha_S \beta_S} - (Risk1_S \times WTP \times N)]}{b_S + 2Risk2_S \times \beta_S \times WTP \times N} \approx 1,765,006 \text{ tons}$$

Industrial Oil Palm Plantations:

$$q_L^{SO} = \frac{\beta_L^2 [P^0 - \frac{a_L}{\beta_L} - \frac{w}{\alpha_L \beta_L} - (Risk1_L \times WTP \times N)]}{b_L + 2Risk2_L \times \beta_L \times WTP \times N} \approx 3,900,992 \text{ tons}$$

These calculations incorporate both the private marginal costs and the marginal external costs of spillover internalized by the social planner. The resulting output levels reflect the trade-off between the economic benefits of palm oil production and the public health risks associated with filovirus transmission.

Aggregate Social-Optimum Supply

Summing across producer types gives the planner's total preferred output:

$$Q^{SO} = q_S^{SO} + q_L^{SO} \approx 1,765,006 + 3,900,992 = 5,665,998 \text{ tons}$$

This socially optimal target is 85,965 tons (1.49%) below the unregulated baseline of 5,751,963 tons. While this contraction may appear modest in percentage terms, it represents a significant reduction in spillover risk when concentrated in high-risk production zones.

Deriving the Uniform Tax Rate

To implement the social-optimum allocation through market mechanisms, we need to identify the uniform tax rate that aligns private incentives with social welfare. Following the methodology in Section 4.2.6, we first characterize the aggregate private supply function as:

$$Q^0(x) = Ax - B,$$

where

$$A = \frac{\beta_S^2}{b_S} + \frac{\beta_L^2}{b_L}$$

$$B = \frac{\beta_S}{b_S} \left(a_S + \frac{w}{\alpha_S} \right) + \frac{\beta_L}{b_L} \left(a_L + \frac{w}{\alpha_L} \right)$$

The optimal uniform tax must satisfy:

$$Q^0(P^0 - t) = Q^{SO}(P)$$

Substituting the slope–intercept form,

$$A(P^0 - t) - B = Q^{SO} \Rightarrow t = P^0 - \frac{Q^{SO} + B}{A}$$

Plugging in our calibrated values for $P^0 = 589.68$, $Q^{SO} = 5,665,998$ the computed A and B , we obtain:

$$t^* \approx \$15.29 \text{ per ton}$$

This tax rate represents the "weighted average" of the marginal external costs across producer types, with weights determined by their respective supply elasticities. At this rate, producers face an effective price of $P^0 - t^* = \$574.39$ per ton, which elicits exactly the socially optimal aggregate quantity of 5.67 million tons.

Table 4.2 provides a comprehensive comparison of key baseline and spillover-inclusive outcomes:

Table 4.2: Comparison of Baseline and Spillover-Inclusive Outcomes

Metric	Baseline (3.2)	Spillover-Inclusive Optimum (3.3)	Absolute Change	% Change
Smallholder output q_S	1,792,587 t	1,765,006 t	–27 581 t	–1.55 %
Industrial output q_L	3,959,376 t	3,900,992 t	–58 384 t	–1.47 %
Total output Q	5,751,963 t	5,665,998 t	–85 965 t	–1.49 %
Smallholder area A_S	256,084 ha	252,144 ha	–3 940 ha	–1.54 %
Industrial area A_L	197,969 ha	195,050 ha	–2 919 ha	–1.48 %
Total area A	454,053 ha	447,194 ha	–6 859 ha	–1.49 %
Smallholder labor L_S	36,068 workers	35,513 workers	–555 workers	–1.54 %
Industrial labor L_L	17,519 workers	17,261 workers	–258 workers	–1.47 %
Total labor L	53,587 workers	52,774 workers	–813 workers	–1.52 %

Source: Author's calculation

Welfare Analysis

Imposing the \$15.29 per-ton uniform tax moves the sector from the unregulated outcome to the spillover-inclusive optimum. Three numbers give the scale of the change.

Tax Revenue Generation

This represents the direct fiscal impact of the policy, calculated as the per-unit tax multiplied by the total quantity of palm oil produced after the tax is implemented:

$$\text{Tax Revenue} = t^* \times Q^{SO} \approx \$15.29 \times 5,665,998 \approx \$86.7 \text{million annually}$$

Producer Surplus Reduction

This value represents the difference between producer surplus in the baseline scenario and the tax scenario. For linear supply curves with our calibrated parameters, this equals approximately

\$91.5 million, representing the economic burden placed on producers. Therefore, the tax imposes costs on producers through reduced prices and quantities:

$$\text{Producer Surplus Reduction} = \$91.5 \text{ million}$$

Public Health Benefits

This benefit is calculated by multiplying the reduction in spillover risk (derived from the area changes and risk coefficients) by the expected infections (30,000) and the value per statistical life (\$107,000). The resulting value of approximately \$112.2 million represents the expected social benefit from reduced disease burden. Thus, the contraction in oil palm area reduces the probability of filovirus spillover events:

$$\text{Public Health Benefits} = \$112.2 \text{ million}$$

Net Social Welfare Effect

The net welfare gain is the difference between the public health benefits and the producer surplus reduction:

$$\text{Net Social Welfare Effect} = \$112.2 - \$91.5 = \$20.7 \text{ million}$$

This positive value indicates that the tax policy creates more benefits than costs from a societal perspective, representing the efficiency gain from internalizing the spillover externality.

The calculations demonstrate that while producers bear a substantial cost from the tax policy, the public health benefits outweigh these costs, resulting in a net welfare improvement of approximately \$20.7 million.

4.4. Conclusions and Policy Implications

This chapter developed a spatially explicit bioeconomic model to assess the optimal allocation of land between smallholder and industrial oil palm plantations in the presence of zoonotic externalities, specifically filovirus spillover risk. Building on empirical evidence from

Chapter 3, the model integrates both the private production incentives and the public health costs associated with different plantation types. The findings demonstrate that while industrial oil palm plantations contribute substantially to output, they also impose disproportionately higher spillover risks relative to smallholder systems. Conversely, smallholder plantations, particularly at moderate densities, provide a more favorable trade-off between economic returns and epidemiological safety. These risk differentials have profound implications for land-use efficiency and public welfare, suggesting that the composition of plantation systems matters as much as the total area under cultivation.

A central policy implication of the analysis is the efficacy of a uniform tax on crude palm oil production in internalizing the public health externality. Simulation results indicate that a modest tax—approximately US\$15.30 per metric ton—can reduce overall production by less than 2% while generating substantial social returns. Specifically, the tax yields US\$86.7 million in annual revenue and achieves US\$112 million in expected reductions in spillover-related health costs. This suggests that relatively minor fiscal adjustments can yield disproportionately large social gains when they are well-calibrated to epidemiological risk. Importantly, these results hold across a range of elasticity and parameter assumptions, underscoring the robustness of the proposed intervention.

From a policy design perspective, the findings advocate for differentiated treatment of plantation types in agricultural and fiscal planning. While blanket restrictions on oil palm expansion may conflict with development goals, targeted policies that discourage risky forms of production—such as large-scale monocultures in high-risk zones—can achieve both economic and epidemiological objectives. In practice, this could include integrating spillover risk assessments into land-use zoning, offering preferential subsidies for mixed-cropping systems, and using tax

instruments to shift incentives toward more sustainable production configurations. These policy tools, when combined with improved surveillance of land-use transitions and disease outbreaks, provide a practical pathway to harmonize agricultural expansion with One Health principles.

The chapter also contributes methodologically by embedding spatially disaggregated epidemiological risk into a formal economic model, allowing for more precise estimation of socially optimal land allocations. This approach improves upon aggregate bioeconomic models by explicitly accounting for landscape heterogeneity and the nonlinearity of spillover dynamics. It also reinforces the broader policy insight that environmental externalities—when spatially concentrated and non-marginal—require equally granular policy responses. Thus, land-use governance in filovirus-endemic regions must consider not only the total agricultural area but also the ecological configuration and production methods employed.

In summary, this chapter demonstrates that modest, well-targeted policy interventions can realign private land-use decisions with social welfare objectives, significantly reducing zoonotic spillover risks without imposing prohibitive costs on agricultural development. As Africa continues to expand its oil palm frontier, policy frameworks that explicitly account for zoonotic externalities will be critical for ensuring that development is both economically viable and epidemiologically safe. Future research should aim to refine these estimates using localized data on wildlife host densities, human-wildlife contact patterns, and compliance responses to fiscal instruments, thereby enabling even more effective and context-specific policy design.

4.5 References

- Albers, H. J., Lee, K. D., Rushlow, J. R., & Zambrana-Torrel, C. (2020). Disease risk from human-environment interactions: Environment and development economics for joint conservation-health policy. *Environmental & Resource Economics*, 76(4), 929–944. <https://doi.org/10.1007/s10640-020-00449-6>
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Augeraud-Véron, E., Fabbri, G., & Schubert, K. (2021). Prevention and mitigation of epidemics: Biodiversity conservation and confinement policies. *Journal of Mathematical Economics*, 93, 102484. <https://doi.org/10.1016/j.jmateco.2021.102484>
- Ayompe, L. M., Schaafsma, M., & Egoh, B. N. (2021). Towards sustainable palm oil production: The positive and negative impacts on ecosystem services and human wellbeing. *Journal of Cleaner Production*, 278, 123914. <https://doi.org/10.1016/j.jclepro.2020.123914>
- Barbier, E. B. (2021). Habitat loss and the risk of disease outbreak. *Journal of Environmental Economics and Management*, 108, 102451. <https://doi.org/10.1016/j.jeem.2021.102451>
- Bernstein, A. S., Ando, A. W., Loch-Temzelides, T., Vale, M. M., Li, B. V., Li, H., Busch, J., Chapman, C. A., Kinnaird, M., Nowak, K., Castro, M. C., Zambrana-Torrel, C., Ahumada, J. A., Xiao, L., Roehrdanz, P., Kaufman, L., Hannah, L., Daszak, P., Pimm, S. L., & Dobson, A. P. (2022). The costs and benefits of primary prevention of zoonotic pandemics. *Science Advances*, 8(5), eabl4183. <https://doi.org/10.1126/sciadv.abl4183>
- Carrere, R. (2013). *Oil palm in Africa: Past, present and future scenarios – 2013 update* (WRM Series on Tree Plantations No. 15). World Rainforest Movement.

<https://www.wrm.org.uy/publications/oil-palm-in-africa-past-present-and-future-scenarios-2013-update>

- Corley, R. H. V., & Tinker, P. B. H. (2016). *The Oil Palm* (5th ed.). Wiley-Blackwell.
- Gaveau, D. L. A., Locatelli, B., Salim, M. A., Husnayaen, Manurung, T., Descals, A., Angelsen, A., Meijaard, E., & Sheil, D. (2022). Slowing deforestation in Indonesia follows declining oil palm expansion and lower oil prices. *PLOS ONE*, 17(3), e0266178. <https://doi.org/10.1371/journal.pone.0266178>
- Gouel, C., & Balint, P. J. (2014). The role of storage in the stabilisation of commodity prices. *American Journal of Agricultural Economics*, 96(4), 1011–1028.
- Papenfus, M. I. (2002). *Investing in Oil Palm: An Analysis of Independent Smallholder Oil-Palm Adoption*. Working Paper, University of Washington.
- Espinosa, R., Tago, D., & Treich, N. (2020). Infectious diseases and meat production. *Environmental & Resource Economics*, 76(4), 1019–1044. <https://doi.org/10.1007/s10640-020-00484-3>
- Faust, C. L., McCallum, H. I., Bloomfield, L. S. P., Gottdenker, N. L., Gillespie, T. R., Torney, C. J., Dobson, A. P., & Plowright, R. K. (2018). Pathogen spillover during land conversion. *Ecology Letters*, 21(4), 471–483. <https://doi.org/10.1111/ele.12904>
- Food and Agriculture Organization of the United Nations (FAO). (2024). *FAOSTAT Statistical Database*. Retrieved May 13, 2025, from <https://www.fao.org/faostat/>
- International Monetary Fund, Global price of Palm Oil [PPOILUSDM], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/PPOILUSDM>, May 1, 2025

- King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137–163. <https://doi.org/10.1093/oxfordjournals.pan.a004868>
- Kubitza, C., Krishna, V. V., Klasen, S., Kopp, T., Nuryartono, N., & Qaim, M. (2024). Labor displacement in agriculture: Evidence from oil palm expansion in Indonesia. *Land Economics*, 100(3), 547–567. <https://doi.org/10.3368/le.100.3.122122-0109R1>
- Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2), 391–413. [https://doi.org/10.1016/S0304-4076\(99\)00044-5](https://doi.org/10.1016/S0304-4076(99)00044-5)
- Leroy, E. M., Epelboin, A., Mondonge, V., Pourrut, X., Gonzalez, J. P., Muyembe-Tamfum, J. J., & Formenty, P. (2009). Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector-Borne and Zoonotic Diseases*, 9(6), 723–728. <https://doi.org/10.1089/vbz.2008.0167>
- Mukpo, A (2017) In Liberia, a battered palm oil industry adjusts to new rules. Mongabay Series: Global Forests, Global Palm Oil, URL <https://news.mongabay.com/2017/05/in-liberia-a-battered-palm-oil-industry-adjusts-to-new-rules/> (Accessed February 2023)
- Mweta, D. E., Chiona, M., & Banda, K. (2025). Socio-economic benefits and challenges confronting oil palm production among indigenous rural farmers in Karonga District, Malawi. *Frontiers in Sustainable Food Systems*, 9, 1473991. <https://doi.org/10.3389/fsufs.2025.1473991>
- Shafie, N. J., Sah, S. A. M., Latip, N. S. A., Azman, N. M., & Khairuddin, N. L. (2011). Diversity pattern of bats at two contrasting habitat types along Kerian River, Perak, Malaysia. *Tropical Life Sciences Research*, 22(2), 13–22.
- TFA (Tropical Forest Alliance) (2023). Africa Palm Oil Initiative. URL <https://www.tropicalforestalliance.org/> (Accessed January 2023)

- Viscusi, W. K., & Masterman, C. J. (2017). Income elasticities and global values of a statistical life. *Journal of Benefit-Cost Analysis*, 8(2), 226–250. <https://doi.org/10.1017/bca.2017.12>
- Wallace, R. G., & Wallace, R. (Eds.). (2016). *Neoliberal Ebola: Modeling disease emergence from finance to forest and farm*. Springer.
- Wilkinson, D. A., Marshall, J. C., French, N. P., & Hayman, D. T. S. (2018). Habitat fragmentation, biodiversity loss and the risk of novel infectious disease emergence. *Journal of the Royal Society Interface*, 15(149), 20180403. <https://doi.org/10.1098/rsif.2018.0403>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.
- World Bank. (2014a). *The economic impact of the 2014 Ebola epidemic: Short- and medium-term estimates for West Africa*. Retrieved from <https://www.worldbank.org/en/region/afr/publication/the-economic-impact-of-the-2014-ebola-epidemic-short-and-medium-term-estimates-for-west-africa>
- World Bank. (2014b, November 19). *Nearly half of Liberia's workforce no longer working since start of Ebola crisis* [Press release]. Retrieved from <https://www.worldbank.org/en/news/press-release/2014/11/19/half-liberia-workforce-no-longer-working-ebola-crisis>
- World Bank. (2017). *The economic impact of the 2014 Ebola epidemic: Short and medium-term estimates for West Africa*. Retrieved from <https://www.worldbank.org/en/region/afr/publication/the-economic-impact-of-the-2014-ebola-epidemic-short-and-medium-term-estimates-for-west-africa>
- World Bank. (2022). *Poverty and shared prosperity 2022: Correcting course*. Retrieved from <https://openknowledge.worldbank.org/handle/10986/37739>

- World Economic Forum. (2016, May 17). *Africa Palm Oil Initiative: Balancing economic development with conservation*. Retrieved from <https://www.weforum.org/stories/2016/05/africa-palm-oil-initiative-balancing-economic-development-with-conservation/>
- World Economic Forum. (2022, November 10). *How African palm oil can boost livelihoods and protect forests*. Retrieved from <https://www.weforum.org/stories/2022/11/how-african-palm-oil-can-boost-livelihoods-and-protects-forests/>
- World Health Organization. (2016, June). *Ebola outbreak 2014–2016 – West Africa*. Retrieved from <https://www.who.int/emergencies/situations/ebola-outbreak-2014-2016-West-Africa>
- World Health Organization. (2020, July 3). *Ebola outbreak 2018–2020 – North Kivu & Ituri, Democratic Republic of the Congo*. Retrieved from <https://www.who.int/emergencies/situations/Ebola-2019-drc->
- World Health Organization. (2023, January 11). *Ebola outbreak 2022 – Uganda (Sudan virus disease)*. Retrieved from <https://www.who.int/emergencies/situations/ebola-uganda-2022>
- World Health Organization. (2025, May 1). *Ebola disease* (Fact sheet). Retrieved from <https://www.who.int/news-room/fact-sheets/detail/ebola-disease>
- Zapata-Hernández, A., Ruiz-Álvarez, E., Arias, N., Mosquera-Montoya, M., & Cooman, A. (2024). Adoption of mechanization alternatives in oil palm crops in the Colombian Orinoquía natural region. *OCL*, 31, 10. <https://doi.org/10.1051/ocl/2024008>

CHAPTER 5

CONCLUSIONS AND POLICY IMPLICATIONS

This dissertation aims to illuminate how agricultural expansion—specifically, the rapid growth of oil palm cultivation—reshapes ecological systems and, in turn, alters the probability of zoonotic disease emergence in tropical Africa. By integrating advances in remote sensing, spatial econometrics, and bio-economic modeling, the research presents a coherent narrative that progresses from the detection of landscape change to the measurement of its epidemiological consequences and, ultimately, to the design of instruments capable of internalizing the attendant external costs. The findings collectively demonstrate that sustainable agricultural development must be evaluated not only in terms of its contributions to economic growth but also in light of its ecological and public health ramifications.

The first major contribution lies in the creation of a scalable monitoring architecture for oil palm landscapes. Leveraging daily MODIS imagery and an XGBoost classification pipeline, Chapter 2 reconstructs annual planting dates for more than 450,000 ha of smallholder and industrial oil palm plantations across seventeen African countries. This approach overcomes the data-gap limitations of Landsat, offering a temporally continuous record that supports robust longitudinal analysis. The resulting dataset provides a critical empirical platform for both environmental assessments and disease-risk modeling, enabling researchers and policymakers to observe land-use trajectories in near real time.

Building on this geospatial foundation, Chapter 3 employs a high-resolution panel of 10,667 grid cells (2001–2018) to estimate the heterogeneous epidemiological footprints of

different production systems. The econometric results reveal that industrial monocultures are positively and significantly associated with Ebola and Marburg spillover risk, whereas diversified smallholder mosaics exhibit either neutral or mildly protective effects. Deforestation alone is not a reliable predictor once the post-conversion land use is specified, underscoring the importance of distinguishing among plantation types when evaluating zoonotic vulnerability. These findings advance the literature by demonstrating that uniform treatment of plantation land cover obscures crucial ecological heterogeneity.

Chapter 4 integrates these empirical insights into a spatially explicit bio-economic model that values land-use choices in the presence of disease-related externalities. Simulations suggest that a uniform excise of approximately US\$ 15 per metric ton of crude palm oil would shift production only marginally below baseline levels, while yielding a net social welfare gain of roughly US\$ 20 million per year after accounting for reduced expected public health losses. This result illustrates that well-calibrated fiscal instruments can nudge the sector toward a socially optimal allocation of land, even when the administrative capacity to enforce more granular regulations is limited.

Several policy implications emerge from this body of evidence. First, land-use governance should move beyond blanket deforestation bans to a differentiated strategy that embeds spillover-risk metrics in environmental impact assessments. Zoning regulations that steer large-scale monocultures away from bat-migration corridors while supporting diversified smallholder systems are likely to deliver both ecological and public health benefits. Second, fiscal policy offers a pragmatic lever for internalizing externalities: a single, risk-adjusted excise on crude palm oil is administratively simpler than farm-type-specific taxes yet remains economically efficient if updated periodically with new risk estimates. Allocating a portion of the resulting revenues to

community One-Health programs and remote-sensing surveillance would reinforce the link between revenue generation and risk mitigation.

Third, development banks and agricultural ministries should expand credit lines and technical assistance for agro-diverse smallholder models, as these systems not only reduce human-wildlife contact intensity but also generate higher rural employment per tonne of output. Fourth, effective intervention requires cross-sectoral coordination: agriculture, environment, and health agencies need integrated data platforms that couple MODIS-based land-conversion alerts with veterinary and human disease surveillance. Finally, voluntary market mechanisms can complement state action. Updating sustainability standards—such as the Roundtable on Sustainable Palm Oil—to include zoonotic-risk indicators and mandating public disclosure of plantation boundaries and planting years would harness consumer and investor pressure in favor of safer production practices.

Like all empirical endeavors, the present research is not without limitations. The 250-m resolution of MODIS may underestimate heterogeneity in fragmented landscapes, suggesting that future work should fuse higher-resolution Sentinel-2 or commercial imagery to refine classification accuracy. Moreover, spillover risk is proxied here by land-cover aggregates; integrating serological surveys or bat-roost telemetry would help unpack causal mechanisms. Finally, the tax simulations assume full compliance; agent-based models that incorporate informal milling, cross-border leakage, and strategic land-clearing could offer more realistic forecasts of policy effectiveness.

Notwithstanding these caveats, the dissertation contributes a multidisciplinary framework for balancing agricultural growth with ecological resilience and human well-being. By demonstrating that the configuration—not merely the extent—of oil palm expansion determines

zoonotic outcomes, it underscores the need for nuanced, risk-aware land-use policies in the tropics. Implemented together, differentiated zoning, modest uniform taxation, smallholder-centered incentives, and enhanced surveillance constitute a feasible blueprint for aligning private incentives with collective health objectives as Africa's agricultural frontier continues to advance.

Appendix A: Localities of Ebola and Marburg Spillovers

Table A.1: Localities of Ebola and Marburg Spillovers

year	type	Latitude	Longitude
2000	Ebola	2.94998	32.19997
2001	Ebola	0.67705	14.28902
2002	Ebola	0.62049	14.37774
2002	Ebola	0.13418	14.20981
2003	Ebola	0.56015	14.65732
2004	Ebola	4.43149	28.7054
2004	Marburg	-7.7639	15.25855
2005	Ebola	0.494444	14.67861
2007	Marburg	-0.13065	30.30894
2007	Ebola	-5.25956	21.40954
2007	Ebola	0.7706	30.13041
2008	Marburg	-0.2772	30.052
2008	Ebola	-5.63674	21.37481
2011	Ebola	0.62415	32.73669
2012	Marburg	-0.11667	30.5
2012	Ebola	2.57874	27.27105
2012	Ebola	0.86599	30.92654
2012	Ebola	0.83175	32.58253
2013	Ebola	8.6226	-10.0642
2014	Marburg	0.312232	32.55874
2014	Ebola	-0.71387	20.53024
2017	Ebola	3.390009	23.46741
2017	Marburg	1.466845	34.57683
2018	Ebola	0.57	29.32
2018	Ebola	-0.7987	18.4471
2020	Ebola	0.032569	18.28119

Source: Sundaram et al. (2024) and Filion et al. (2023)

Appendix B: Econometric Derivation of Cost Function Parameters

This appendix details the econometric analysis used to derive the land preparation cost parameters (a_i, b_i) for smallholder and industrial oil palm plantations in Africa. These parameters are critical components of the theoretical model presented in Chapter 4, as they determine the shape of the marginal cost functions and, consequently, the supply responsiveness of each producer type. The starting point is equation (6) describing the supply function, which combined with the technological relationship $q_i = \beta_i A_i$, allows us to express the cultivated area as a function of prices:

$$A_i = \frac{\beta_i P}{b_i} - \frac{1}{b_i} \left(a_i + \frac{w}{\alpha_i} \right) \quad (\text{B1})$$

This can be rewritten as a linear regression model:

$$A_i = \gamma_i + \delta_i P + \varepsilon_i \quad (\text{B2})$$

where:

$$\gamma_i = -\frac{1}{b_i} \left(a_i + \frac{w}{\alpha_i} \right), \quad \delta_i = \frac{\beta_i}{b_i}, \quad \text{and } \varepsilon_i \text{ is the error term.}$$

We estimate this regression separately for $i=S, L$, using data on annual world crude palm oil price and plantation areas for smallholder systems and industrial plantations.

B.1 Plantation Area Data

Data on oil palm plantation areas were derived from remote sensing classification of satellite imagery covering major palm oil-producing countries in Africa. The dataset distinguishes between smallholder polyculture systems and industrial monoculture plantations based on spectral signatures, plantation geometry, and canopy structure characteristics. Annual data cover the period 2000-2020, representing the total hectares under each cultivation system across the study region.

Price Data

World crude palm oil (CPO) prices were obtained from the International Monetary Fund's Primary Commodity Price System, covering the period 1997-2020 (to allow for lagged effects). These represent monthly average prices in USD per metric ton, which were converted to annual averages to match the temporal resolution of the plantation area data. All prices were adjusted for inflation using the US Consumer Price Index to ensure comparability across years.

Summary Statistics

Table B.1: Summary Statistics for Key Variables

Variable	Observation	Mean	S.D.	Min	Max
Smallholder area (ha)	21	256463.7	129994.2	116828	543456
Industrial area (ha)	21	198233.7	94379.31	102100	397754
World CPO price (\$/ton)	24	590.125	216.1134	238	1077

Source: Author's calculation

B.2 Temporal Dynamics of Production Decisions

Investment in oil-palm is governed by long biological and institutional lags. Seedlings take three to four years before their first harvest and do not reach full yield until years 7–8 (Corley & Tinker 2016). In many African smallholder settings, land titling, clearing and road construction add further delays. We therefore model a two-stage lag structure:

$$P_t \stackrel{\text{Decision}}{=} A_{t+3} \stackrel{\text{Maturation}}{=} q_{t+6} \quad (\text{B3})$$

- Decision lag (price $\rightarrow$ area). Producers observe the world CPO price P_t and decide how much land to convert.
- Establishment lag (area realization). Land preparation and planting turn those decisions into a measurable plantation area A_{t+3} roughly three years later.

- Maturation lag (area $\rightarrow$ output). New stands begin commercial harvesting about three years after planting, so output appears as q_{t+6}

Although large Indonesian oil palm plantations with pre-approved land banks have been shown to react within a single year to price shocks (Gaveau et al. 2022), such rapid adjustment is less typical for smallholders and frontier regions in Sub-Saharan Africa, where tenure formalities and physical clearing slow the conversion process. Our three-year lag thus captures the modal adjustment horizon for these producers while remaining consistent with option-value evidence that price volatility can postpone planting by several years (Papenfus 2002; Gouel & Balint 2014).

Empirically, Appendix B regresses annual changes in plantation area on lagged world prices and finds the strongest, most significant coefficient at the three-year lag—validating the temporal sequence in (B3). The resulting coefficients (γ_0, γ_1) map directly into the structural cost parameters:

$$A_{i,t+3} = \frac{\beta_i P_t}{b_i} - \frac{1}{b_i} \left(a_i + \frac{w}{\alpha_i} \right) \quad (\text{B4})$$

allowing us to recover a_i and b_i while respecting the biological and institutional realities of oil-palm investment.

B.3 Estimation Results

We estimate the relationship between plantation area and world prices using ordinary least squares (OLS) with robust standard errors to account for potential heteroskedasticity. For each producer type (smallholder and industrial), we test models with contemporaneous prices and lags of one, two, and three years.

Smallholder Oil Palm Plantations

Table B.2: Regression Results for Smallholder Plantation Area

Estimator Dependent variable	OLS The number of smallholder oil palm plantation areas in Africa			
	(t)	(t-1)	(t-2)	(t-3)
World price	153.99856 (92.04620)			
Lag 1 year of world price		175.24449** (81.12571)		
Lag 2 years of world price			218.31859*** (71.50724)	
Lag 3 years of world price				268.36878*** (69.91335)
constant	163367.87017** (62404.26030)	152935.89825** (57508.24911)	126688.85641** (46557.09504)	97832.15910** (38738.50924)
Robust	Yes	Yes	Yes	Yes
Observations	21	21	21	21

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, robust standard errors in parentheses.

Industrial Oil Palm Plantations

Table B.3: Regression Results for Industrial Plantation Area

Estimator Dependent variable	OLS The number of industrial oil palm plantation areas in Africa			
	(t)	(t-1)	(t-2)	(t-3)
World price	95.63906 (67.18199)			
Lag 1 year of world price		112.53940* (60.08252)		
Lag 2 years of world price			146.41132** (53.85448)	
Lag 3 years of world price				187.19135*** (52.74343)
constant	140417.62777*** (45862.68805)	131749.72245*** (41922.77362)	111202.64436*** (34320.78073)	87585.79983*** (28752.82965)
Robust	Yes	Yes	Yes	Yes
Observations	21	21	21	21

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, robust standard errors in parentheses.

B.4 Model Selection and Interpretation

For both plantation types, the statistical significance of the price coefficient increases consistently as the lag length increases from zero to three years. The contemporaneous price model (t) shows no significant relationship for either plantation type. The significance improves progressively with one-year lag ($p < 0.05$ for smallholders, $p < 0.1$ for industrial), two-year lag ($p < 0.01$ for smallholders, $p < 0.05$ for industrial), and reaches its strongest level with the three-year lag ($p < 0.01$ for both types).

The coefficient magnitude also increases with lag length, from 153.99 to 268.37 for smallholders and from 95.64 to 187.19 for industrial plantations. This pattern indicates that producers respond more strongly to price signals observed three years prior to the observed plantation area, which aligns with the biological and economic realities of oil palm cultivation. The three-year lag captures the time required for financing arrangements, land preparation, seedling procurement, planting, and initial establishment before plantations appear in satellite imagery as mature stands. Concurrently, the constant term decreases as the lag increases, suggesting that a larger portion of the variation in plantation area is explained by the lagged price variable.

This strong relationship with the three-year lag validates our theoretical framework's temporal sequence (equation B3): $P_t \rightarrow A_{t+3} \rightarrow q_{t+6}$

The results provide empirical confirmation that oil palm producers in Africa make planting decisions based on price signals observed three years prior to plantation establishment, reflecting the realities of perennial crop investment dynamics.

B.5 Derivation of Cost Function Parameters

Having identified the three-year lag model as most appropriate, we now derive the cost function parameters required for our theoretical model.

From the regression results and our theoretical framework, we have:

- For smallholders: $\gamma_S = 97,832\$$ and $\delta_S = 268.369\$$
- For industrial plantations: $\gamma_L = 87,586\$$ and $\delta_L = 187.191\$$

Using the relationships:

$$\gamma_i = -\frac{\beta_i}{b_i} \left(a_i + \frac{w}{\alpha_i} \right), \quad \delta_i = \frac{\beta_i^2}{b_i}$$

We can solve for the cost parameters:

$$b_i = \frac{\beta_i}{\delta_i}$$

$$a_i = -\frac{\gamma_i b_i}{\beta_i} - \frac{w}{\alpha_i}$$

Substituting the known values from table 1 yields:

$$a_S = -2612.78 / \text{ha}, b_S = 0.02608 / \text{ha}^2, a_L = -9396.20 / \text{ha}, b_L = 0.10684 / \text{ha}^2$$

Appendix C: Epidemiological Derivation of Spillover Risk Parameters

C.1 Introduction

This appendix details the methodology used to derive the filovirus spillover risk parameters ($\rho1_i, \rho2_i$) for smallholder and industrial oil palm plantations in Africa. These parameters quantify the relationship between plantation area and spillover probability, which are crucial components of the external cost functions in our theoretical model (Section 4.2.4).

The analysis draws on spatiotemporal data associating historical filovirus outbreak locations with land-use characteristics, including oil palm cultivation systems. By linking gridded outbreak data with remote sensing measurements of land cover change, we estimate how different plantation types contribute to spillover probability while controlling for confounding environmental and socioeconomic factors.

C.2 Data Sources and Methodology

C.2.1 Spatial Grid Construction

We employed a standard grid cell approach used in some research, dividing the study region into uniform 55 km $\times$ 55 km grid cells. This resolution balances granularity of ecological measurement with the spatial uncertainty inherent in retrospective outbreak data. For each grid cell, we compiled:

- Historical filovirus spillover events (primarily Ebola and Marburg virus outbreaks)
- Land cover classification from remote sensing, including:
 - Forest cover and forest loss
 - Smallholder oil palm plantations (polyculture systems)
 - Industrial oil palm plantations (monoculture systems)
- Environmental covariates (elevation, precipitation, temperature, etc.)

- Socioeconomic indicators (population density, poverty indices, healthcare access, etc.)

C.2.2 Spatial Unit Conversion

An important consideration in our analysis is the conversion between grid cell measurements and hectare-level risk parameters required for our economic model. Each grid cell covers:

- Area per grid cell: $55 \text{ km} \times 55 \text{ km} = 3,025 \text{ km}^2$
- Converting to hectares: $3,025 \text{ km}^2 \times 100 \text{ ha/km}^2 = 302,500 \text{ hectares}$

Since our regression analysis uses a neighborhood approach that considers a 3×3 grid cell window (capturing potential spillover dynamics across adjacent cells), the effective analysis area becomes:

- Total analysis area: $302,500 \text{ hectares} \times 9 = 2,722,500 \text{ hectares}$

This spatial framing allows us to express incremental changes in oil palm area as ratios of the total landscape:

- Ratio of 1-hectare increase: $1 \text{ hectare} \div 2,722,500 \text{ hectares} = 0.000000367$

These conversion factors are critical for translating the regression coefficients (based on ratios) into the hectare-level risk parameters used in our economic model.

C.3 Empirical Specification

We employed a linear probability model (LPM) to estimate the relationship between land-use characteristics and spillover probability:

$$\Pr[FV_{kt} = 1] = \alpha_1 \text{ForestLoss}_{kt} + \alpha_2 \text{Small}_{kt} + \alpha_3 \text{Industrial}_{kt} + \alpha_4 X_{kt} + \gamma_k + \delta_{it} + \varepsilon_{kt}$$

Where:

- FV_{kt} is a binary indicator equal to 1 if a filovirus (Ebola or Marburg) spillover event occurs in grid cell k or its eight adjacent cells at time t and 0 otherwise.
- ForestLoss_{kt} represents the ratio of deforested area to total cell area in grid cell k at time t .

- $Small_{kt}$ and $Industrial_{kt}$ capture the ratios of smallholder and industrial oil palm plantation areas to total cell area, respectively.
- X_{kt} is a vector of time-varying control variables measured at the grid cell level:
 - Nighttime light intensity (lumens per hectare)
 - Population density (people per hectare)
 - Mean annual temperature ($^{\circ}\text{C}$)
 - Annual rainfall (mm)
- γ_k denotes grid cell fixed effects, controlling for time-invariant characteristics at the cell level.
- δ_{it} represents country-by-year fixed effects, accounting for time-varying national factors.
- ε_{kt} is the error term.

C.4 Estimation Results

Table C.1: Regression Results for Filovirus Spillover Probability

Estimator Dependent variable	LPM Was there a FV spillover?
Ratio of forest loss	0.00457 (0.00767)
Ratio of forest loss ²	
Ratio of smallholder oil palm plantations	-0.26118** (0.10994)
Ratio of smallholder oil palm plantations ²	1.87659** (0.87512)
Ratio of industrial oil palm plantations	0.25779** (0.11678)
Ratio of industrial oil palm plantations ²	
Socioeconomic Controls	Yes
Environmental Controls	Yes
Country x year fixed effects	Yes
Cell fixed effects	Yes
Cluster	Yes
Observations	192,006
Cells	10,667

The regression results reveal distinct risk profiles for different plantation types:

1. **Smallholder Oil Palm Plantations:** Exhibit a significant negative linear term (-0.26118) and a significant positive quadratic term (1.87659). This pattern suggests that at low densities, smallholder polyculture systems may actually reduce spillover risk (perhaps by providing ecological buffers or alternative food sources for wildlife), but at higher concentrations, the risk increases quadratically.
2. **Industrial Oil Palm Plantations:** Show a significant positive linear term (0.25779) with no significant quadratic term. This indicates that industrial monocultures consistently increase spillover risk in proportion to their areal extent, without the initial buffering effect observed in smallholder systems.

C.5 Derivation of Risk Parameters

To convert the regression coefficients to the risk parameters needed for our economic model, we apply the spatial conversion factors derived in Section C.2.2:

- For smallholders:

$$Risk1_S = -0.26118 \times 0.000000367 = -9.58 \times 10^{-8}$$

$$Risk2_S = 1.87659 \times 0.000000367^2 = 2.53 \times 10^{-13}$$

- For industrial producers:

$$Risk1_L = 0.25779 \times 0.000000367 = 9.47 \times 10^{-8}$$

$$Risk2_L = 0$$