



**‘AIDS Drugs for Africa!’– A case study examining transnational AIDS
treatment activism and the reduction of global antiretroviral prices from
1996 to 2001**

Eleyce M. Northcraft Kavli

NRTELE001

*A minor dissertation submitted in partial fulfillment of the requirements for the award of the
degree of Master of Social Science in International Relations*

Faculty of the Humanities

University of Cape Town

2012

Compulsory Declaration

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature: _____ Date: _____

University of Cape Town

Acknowledgements

To my supervisor Mr. Zwelethu Jolobe, thank you for supporting and encouraging my interest in transnational activism. In addition, I wish to thank you for giving me the freedom to explore this topic and guidance about how to approach my research. You have taught me to go with my instincts and to ‘own my own work,’ two lessons I will take with me outside of UCT.

To UCT, I wish to thank Professor Robert Schrire, Dr. Harry Stephan, Dr. John Akokpari, and Professor Annette Seegers for opening my eyes and mind allowing me to see and examine the complex world of international politics. I am extremely grateful. In addition, I would like to thank Joanne Polzin, Wahieda Gajjar, and Le-Anne Wolmerans for working so hard and keeping the Politics Department running.

To my family and friends, thank you for the love and support you have given me throughout my education at UCT. I would not have been able to do this without you.

To my husband Haakon, I am so thankful for your full-hearted support and belief in me and all my research ideas throughout the past two years. I have grown so much personally, professionally, and academically with you by my side and I could not have asked for a better partner to share in this journey.

Lastly, I wish to acknowledge that this research was driven by my personal interest and curiosity in the ability of individuals to create meaningful and sustainable change and therefore, wish to extend my gratitude to all those courageous individuals around the world who continually stand up and fight for change.

Abstract

Between September 2000 and February 2001, the global average price for one year of highly active antiretroviral therapy (HAART) per patient reduced from \$10,000 to less than \$350 dollars. This occurred after four years of intense opposition from the Pharmaceutical Research and Manufacturers Association (PhRMA) and the United States government, and is suggested to have been due to the strategic collective action of transnational AIDS treatment activists working around the world to improve access to life-saving antiretrovirals (ARVs). The most prominent transnational AIDS treatment activist organizations at this time were the Health Group Action Project (Health GAP), which included the AIDS Coalition to Unleash Power (ACT UP) and Consumer Project on Technology (CPT) in the United States and the Treatment Action Campaign (TAC) of South Africa. It is these organizations that are suggested, by both scholars and top pharmaceutical executives, to have been an important extra market force contributing to ARV price reductions between 1996, when HAART was discovered, and 2001. Therefore, this dissertation will examine this suggested causal relationship between transnational AIDS treatment activism and the reduction of global ARV prices specifically by explaining *how transnational AIDS treatment activism emerged in the United States and South Africa and contributed to the reduction of global ARV prices between September 2000 and February 2001.*

The current body of literature on transnational AIDS treatment activism, specifically concerning the strategies and impact of Health GAP and TAC, is simplistic and largely descriptive. Furthermore, current study of transnational activism is relatively under-theorized. As such, this dissertation has two research objectives: to provide a comprehensive and empirical account of transnational AIDS treatment activism and the international politics concerning ARV price reductions from 1996 to 2001 and to provide a critical analysis of current transnational activism theory. Therefore, this theoretical single case study will apply and test theoretical arguments developed by Sidney Tarrow in his book *The New Transnational Activism*, to examine the suggested causal relationship between transnational AIDS treatment activism and the reduction of ARV prices. In doing so, this research contributes to the literature in two ways. First, it argues that transnational AIDS treatment activism had a limited contribution to global ARV price reductions, and that its causal link is best understood in conjunction with broad and international structural changes. Second, this research highlights both strengths and limitations of Tarrow's theoretical arguments, which may be used for future theory development in the study of transnational activism.

List of Abbreviations

AAI	Accelerating Access Initiative
ACT UP	AIDS Coalition to Unleash Power
ADAP	AIDS Drug Assistance Programs
AIDS	Acquired Immune Deficiency Syndrome
ANC	African National Congress
ARV	Antiretroviral
AZT	Zidovudine
CARE	Comprehensive AIDS Resources Emergency Act (United States)
CBC	Congressional Black Caucus (United States)
CIPLA	Chemical, Industrial and Pharmaceutical Laboratories
CPT	Consumer Project on Technology
DA	Democratic Alliance
DAI	Drug Access Initiative
DOH	Department of Health
GEAR	Growth, Employment, and Redistribution Plan (South Africa)
GLOW	Gay and Lesbian Organization of the Witwatersrand
GPA	Global Program on AIDS
HAART	Highly active antiretroviral therapy
Health GAP	Health Group Action Project
HIV	Human Immuno-deficiency Virus
IGLHRC	International Gay and Lesbian Human Rights Commission
ILO	International Labor Organization
IMF	International Monetary Fund
IPAA	International Partnership Against AIDS in Africa
IPR	Intellectual property right
MAA	Mobilization Against AIDS
MAP	Multi-country AIDS Program
MCC	Medicines Control Council (South Africa)
MOH	Minister of Health
MSF	Médecins Sans Frontières
MTCT	Mother-to-child-transmission
MWT	Marxist Workers Tendency
NACOSA	National AIDS Convention of South Africa
NGCLE	National Coalition for Gay and Lesbian Equality

NGO	Non-governmental organization
OLGA	Organization of Gay and Lesbian Activists
PAC	Pan Africanist Congress
PACHA	Presidential Advisory Council on HIV/AIDS (United States)
PhRMA	Pharmaceutical Research and Manufacturers Association
PLWHA	People living with HIV and AIDS
PMA	Pharmaceutical Manufacturers Association
PWA	People with AIDS
R&D	Research and development
RDS	Revised Drug Strategy
SSA	Sub-Saharan Africa
TAC	Treatment Action Campaign
TAG	Treatment Action Group
TCT	Triple combination therapy
TRIPS	Trade Related Intellectual Property Rights
UN	United Nations
UNAIDS	Joint United Nations Program on HIV/AIDS
UNDP	United Nations Development Program
UNESCO	United Nations Educational, Scientific, and Cultural Organization
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNODC	United Nations Office on Drugs and Crime
UNPF	United Nations Population Fund
UNSC	United Nations Security Council
USD	United States dollar
USTR	United States Trade Representative
WFP	World Food Program
WHA	World Health Assembly
WHO	World Health Organization
WTO	World Trade Organization

List of Figures

Figure 1. Model of Global Issue Framing	16
Figure 2. Model of Internalization	17
Figure 3. Alternative Pathways of Transnational Diffusion	19
Figure 4. Model of Scale Shifting	20
Figure 5. The Boomerang Pattern	21
Figure 6. Model of Externalization	23
Figure 7. Typology of Transnational Coalitions	25
Figure 8. ARV Price Reductions 2000-2001	44
Figure 9. Diagram of Conjuncture	65

University of Cape Town

Table of Contents

Acknowledgements	iii
Abstract	iv
List of Abbreviations	v
List of Figures	vii
Chapter 1: Introduction	1
1.1 Research Objectives	4
1.2 Literature Review	4
1.2.1 ARV Prices, IPRs, and Transnational Political Contention	4
1.2.2 Transnational AIDS Treatment Activism	6
1.2.3 Health GAP and TAC	8
1.3 Methodology, Research Design, and Limitations	9
1.4 Empirical Findings and Theoretical Implications	11
1.5 Significance	11
1.6 Organization of the Study	12
Chapter 2: Analytical Framework - Tarrow's Transnational Activism	13
2.1 Domestically Rooted Transnational Activism	13
2.1.1 Four Key Factors and Concepts from Social Movement Theories	14
2.1.2 Domestic Processes: Global Issue Framing and Internalization	15
2.2 Coordinating Transnational Activism Across Borders	17
2.2.1 Transitional Processes: Diffusion and Scale Shifting	18
2.2.2 International Processes: Externalization and Transnational Coalition Formation	20
2.3 International Political Opportunities and Threats	26
2.3.1 Globalization, Internationalism, and Internationalization	26
2.4 The Analysis to Come	27
Chapter 3: Case Study - Transnational AIDS Treatment Activism and ARV Price Reductions from 1996 to 2001	28
3.1 Setting the Stage (1996 - 1997)	28
3.1.1 State Level: South Africa Up Against the US and PMA	28
3.1.2 International Level: Hope with HAART and UNAIDS	30
3.2 Global Frustration and Mobilization (1998)	32
3.2.1 State Level: The Emergence of Health GAP and TAC	33

3.2.2 International Level: Attempting to Bridge the Global Treatment Gap	34
3.3 Action for Africa (1999)	35
3.3.1 State Level: Government and Industry Under Pressure	36
3.3.2 International Level: Addressing AIDS and Development	39
3.4 Transnational Activists Working Together (January 2000 - August 2000)	41
3.4.1 State Level: Coordinated Efforts of TAC and Health GAP	41
3.4.2 International Level: Accelerating ARV Access	43
3.5 ARV Price Reductions (September 2000 - February 2001).....	44
3.6 Case Summary	45
Chapter 4: Analysis - Tarrow and Transnational AIDS Treatment Activism	47
4.1 Domestic Analysis of Transnational AIDS Treatment Activism.....	47
4.1.1 United States	48
4.1.2 South Africa	52
4.1.3 Domestic Impact of Transnational AIDS Treatment Activism on ARV Prices.....	57
4.2 Transnational Coordination of AIDS Treatment Activism.....	57
4.2.1 Diffusion and Scale Shifting	58
4.2.2 Externalization and Transnational Coalition Formation	59
4.2.3 Coordinated Impact on ARV Prices.....	61
4.3 International Opportunities and Threats	62
4.3.1 The WTO and the World Bank	62
4.3.2 UNAIDS and the WHO	63
4.3.3 International Political Shifts and Impact on ARV Prices.....	64
4.4 Conjuncture: From Above and Below	64
Chapter 5: Conclusion	66
5.1 Transnational AIDS Treatment Activism and ARV Price Reductions	67
5.2 Tarrow and Transnational AIDS Treatment Activism.....	69
5.3 Areas for Future Research	71
Bibliography	73

Chapter 1: Introduction

One of the earliest accomplishments in the global fight against the Human Immuno-deficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS) epidemic was the reduction of antiretroviral (ARV) prices to improve AIDS treatment access between 2000 and 2001. In 1996, one year of triple combination therapy (TCT) or highly active antiretroviral therapy (HAART) cost between \$10,000 and \$15,000 dollars (USD) per patient,¹ rendering life-saving AIDS treatment out of reach for 95% of people living with HIV and AIDS (PLWHA), in poor and developing countries such as those in Sub-Saharan Africa (SSA).² However, after four years of intense resistance from the Pharmaceutical Research and Manufacturers Association (PhRMA), which represents the world's largest multinational pharmaceutical corporations, and opposition from the United States (US) government,³ ARV prices began to reduce. Between September 2000 and February 2001, the global average price for one year of TCT per patient reduced to less than \$350 USD,⁴ marking a major turning point as one of the first critical barriers to AIDS treatment access had been addressed.⁵

Improving AIDS treatment access and the reduction of ARV prices is commonly attributed to *transnational AIDS treatment activism*, which refers to the mobilization of strategic collective action to improve ARV access “against external opponents or in favor of goals held in common with transnational allies.”⁶ During this period, from 1996 to 2001, the most prominent transnational AIDS treatment activist organizations working on the reduction of ARV prices were the Health Group Action Project (Health GAP) based in the US, which included the AIDS Coalition to Unleash Power (ACT UP) and the Consumer Project on Technology (CPT), and the Treatment Action Campaign (TAC) of South Africa. It is these organizations that top pharmaceutical executive, Kevin McKenna from Boehringer Ingelheim, refers to in an interview when he admitted, “activist organizations did persuade us to see the need for a middle ground between our need for returns on investment and the

¹ Raymond Copson, *CRS Issue Brief for Congress - AIDS in Africa* (Washington D.C.: Congressional Research Service, January 5, 2005), accessed October 28, 2011, <http://www.fas.org/man/crs/IB10050.pdf>, 9.

² Mary Schug, "Promoting Access to HIV/AIDS Pharmaceuticals in Sub-Saharan Africa within the Framework of International Intellectual Property Law," *Law and Inequality: A Journal of Theory and Practice* 19 (2001): 229-257.

³ Merrill Goozner, "Third World Battles for Aids Drugs - U.S. Firms Oppose Generic Licensing," *Chicago Tribune*, April 28, 1999, accessed October 24, 2011, http://articles.chicagotribune.com/1999-04-28/news/9904280067_1_compulsory-licensing-south-africa-aids-drugs; Lisa Richwine, "Gore Worked to Soften South Africa Health Law," *Reuters New Media*, April 16, 1999, accessed October 24, 2011, <http://www.aegis.com/news/re/1999/re990404.html>.

⁴ Carmen Perez-Cases et al., "Accessing ARVs: Untangling the Web of Price Reductions for Developing Countries," (Médecins Sans Frontières, 2001), accessed August 22, 2011, utw.msfaccess.org/downloads/311, 3.

⁵ Huntly Collins, "World's AIDS Fighters Plot New Strategy," *Philadelphia Inquirer*, July 11, 2000, accessed September 22, 2011, <http://search.ebscohost.com/login.aspx?direct=true&db=nfh&AN=KRBN200019301998&site=ehost-live>.

⁶ Sidney Tarrow, *The New Transnational Activism* (Cambridge, UK: Cambridge University Press, 2005), 43.

poor's need for medication."⁷ Furthermore, scholars, such as Nicoli Nattrass, claim these organizations were an important extra market force "profoundly" affecting ARV prices.⁸ Lastly, activists, such as Eric Sawyer from ACT UP and Nathan Geffen from TAC, maintain the work of their organizations were influential for improving AIDS treatment access and reducing ARV prices.⁹

It appears that a widely held perception exists in the current body of literature, which suggests there is a causal relationship between transnational AIDS treatment activism and the reduction of ARV prices for improving AIDS treatment access. This dissertation will examine this suggested causal relationship specifically by explaining *how transnational AIDS treatment activism emerged in the US and South Africa and contributed to the reduction of global ARV prices between September 2000 and February 2001.*

Transnational AIDS treatment activism "cuts across disciplines, is complex, under-theorized, and does not lend itself to neat theoretical explication."¹⁰ Furthermore, current study of transnational activism is relatively limited, as it is a rather new field of study and because current theories are under-developed. Therefore, this dissertation will contribute to the current body of literature by challenging existing theories of transnational activism by applying and testing theoretical arguments developed by Sidney Tarrow, in his recent book *The New Transnational Activism*. Tarrow's book, which James Richter claims, "provides one of the most comprehensive and nuanced studies of transnational activism currently available,"¹¹ is primarily concerned with explaining how transnational activism emerges, takes the shape it does, and is most effective. The main objective of Tarrow's book is to demonstrate and argue against the general claim that "globalization is responsible for the rise of transnational activism,"¹² because it does not and cannot "explain *how* people become involved in transnational activism; to be specific it fails to identify the opportunities and resources that they need to become active."¹³

⁷ Steven Friedman and Shauna Mottiar, "A Rewarding Engagement? The Treatment Action Campaign and the Politics of HIV/AIDS," *Politics Society* 33, no. 4 (2005): 540.

⁸ Nicoli Nattrass, "The (Political) Economics of Antiretroviral Treatment in Developing Countries," *Trends in Microbiology* 16, no. 12 (2008): 574-579.

⁹ Eric Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," in *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization*, eds. Benjamin Shepard and Ronald Hayduk (London: Verso, 2002), 88-102.; Nathan Geffen, *Debunking Delusions: The Inside Story of the Treatment Action Campaign* (Auckland Park: Jacana Media, 2010), 56.

¹⁰ Eduard Grebe, "Networks of Influence: A Theoretical Review and Proposed Approach to AIDS Treatment Activism," University of Cape Town Center for Social Science Research Unit - AIDS and Society Research Unit Working Paper, working paper, July 2008, accessed May 25, 2011. <http://www.cssr.uct.ac.za/sites/cssr.uct.ac.za/files/pubs/WP218.pdf>, 2.

¹¹ James Richter, "The New Transnational Activism – Sidney Tarrow," *Journal of Politics* 69, no. 1 (2007): 256.

¹² Tarrow, *The New Transnational Activism*, 5.

¹³ Simon Thompson, "Mapping Transnational Activism," *Democratija* 6 (2006): 116.

Tarrow explains that transnational activism is one form of *transnational political contention*, which refers to international conflicts that link individuals “to one another, to states and to international institutions.”¹⁴ More specifically, Tarrow defines *transnational activism* as the strategic collective action of individuals mobilized either against an external or international state or non-state actor, or for a cause or issue shared by activists in more than one country.¹⁵ By adopting and adapting theories of social movements and international politics, Tarrow proposes a number of hypotheses concerning how transnational activism emerges and is effective, three of which will be used in this study to form an analytical framework and be tested in their applicability for explaining the emergence and effectiveness of transnational AIDS treatment activism.

The first argument proposed by Tarrow and applied in this study concerns the idea that transnational activism is still very much rooted in domestic state politics, resources and networks, and has its most visible impact on domestic politics.¹⁶ To address this argument, this dissertation will examine how Health GAP and TAC emerged in the US and South Africa, the resources and opportunities they used, and assess their domestic impact in relation to global ARV price reductions. The second set of proposals used in this study are Tarrow’s domestic, transitional, and international processes which explain how transnational activism emerges and expands across borders to be effective.¹⁷ To test these processes, this study will examine the ways in which Health GAP and TAC came into contact with one another and began coordinating their efforts to reduce the prices of ARVs to improve AIDS treatment access. The third claim applied and tested by this study is Tarrow’s argument that changes in the international political structure, or *internationalization*, create new international political opportunities and threats, incentives, targets, and avenues for transnational activism to emerge and be most effective.¹⁸ To analyze this claim, this dissertation will trace international policy commitments, programs, and funding of international institutions for combating AIDS between 1996 and 2001, to examine how transnational AIDS treatment activism emerged and was effective in light of new international political opportunities.

Tarrow states he has neither “explained transnational activism nor predicted its outcomes,” but has merely delineated “processes that connect initial episodes of contention with responses and replications elsewhere and at different levels of the international system.”¹⁹ However, it is these processes and theoretical arguments that Jackie Smith claims, “can generate new attempts at theory-

¹⁴ Tarrow, *The New Transnational Activism*, 25.

¹⁵ *Ibid.*, 43.

¹⁶ *Ibid.*, 2-3.

¹⁷ *Ibid.*, 32-33.

¹⁸ *Ibid.*, 8.

¹⁹ *Ibid.*, 138.

building and hypothesis testing in the field,"²⁰ which this dissertation will do by examining transnational AIDS treatment activism and the reduction of global ARV prices from 1996 to 2001.

1.1 Research Objectives

This dissertation has two research objectives. The *first objective* is to provide a chronological and empirical account of transnational AIDS treatment activism in the US and South Africa and the international politics concerning ARV price reductions from 1996 to 2001. The *second objective* is to provide a critical analysis of Tarrow's theoretical arguments concerning transnational activism by testing his hypotheses to transnational AIDS treatment activism, thus highlighting strengths and weaknesses for future theory building in the field of transnational activism.

1.2 Literature Review

The availability and affordability of ARVs has been and continues to be a highly contentious international issue, which rests on the pharmaceutical industry's intellectual property rights (IPRs), granted and protected by state governments, and globally enforced through the World Trade Organization's (WTO) 1995 agreement on Trade Related Intellectual Property Rights (TRIPS).²¹ As such, the literature reviewed for this research falls within three categories: work on IPRs and access to medicines, literature on transnational AIDS treatment activism, and accounts of Health GAP and TAC. All three areas of study provide this research with an understanding of the problem (ARV prices and IPRs), the suggested explanation for ARV price reductions (transnational AIDS treatment activism under Health GAP and TAC), and how to examine transnational activism. However, considered in isolation each category of study fails to provide a comprehensive or critical account of when ARV prices and IPRs were a cause of major transnational political contention and when transnational AIDS treatment activism in the US and South Africa emerged and took action for the improvement of ARV access. Therefore, all three areas of study are limited in scope, rather simplistic, and largely descriptive. This section will provide a critical assessment of the literature used by this study and explain how this dissertation will contribute to the current body of literature by providing a comprehensive, theoretically grounded and causal analysis of transnational AIDS treatment activism and the transnational political contention surrounding ARV prices and IPRs.

1.2.1 ARV Prices, IPRs, and Transnational Political Contention

The first category of literature used by this study and reviewed here is on the issue of ARV prices and IPRs, which helps to explain the resistance from PhRMA and the US government to ARV price reductions. According to Mary Schug, in her article 'Promoting Access to HIV/AIDS Pharmaceuticals

²⁰ Jackie Smith, "Book Review: The New Transnational Activism," *The Canadian Journal of Sociology* 31, no. 4 (2006): 547.

²¹ Schug, "Promoting Access to HIV/AIDS Pharmaceuticals," 230-231.

in Sub-Saharan Africa Within the Framework of International Intellectual Property Law,” patents, as one form of IPRs granted by governments for pharmaceutical products in ARVs, gives the patent holder exclusive marketing rights for a specific period of time.²² This allows pharmaceutical companies to set high prices for medicines, which the industry argues are necessary in order to recoup extensive research and development (R&D) costs,²³ and to provide incentive for future research on life-saving medicines. In other words, the patent holder “generates monopoly rents (excess profits), and it is these profits that are supposed to provide incentive for engaging in research.”²⁴ Therefore, the reduction of ARV prices and loss in profits, or the manufacturing and acquiring of generic ARVs through parallel importation or compulsory licensing,²⁵ without permission from original patent holders poses a number of problems for the pharmaceutical industry.

The relationship between ARV prices and IPRs, as outlined by Schug, helps explain the reasons behind opposition to ARV price reductions, which began when the US applied trade pressure, at the request of PhRMA, on Brazil, Thailand, and South Africa, to deter them from generically manufacturing patented ARVs to protect the pharmaceutical industry’s exclusive marketing rights between 1996 and 2001.²⁶ Furthermore, it was on similar grounds that the Pharmaceutical Manufacture Association (PMA) of South Africa filed a lawsuit against President Nelson Mandela and the African National Congress (ANC) government in February 1998, which was due to the potential violation of South Africa’s obligation to TRIPS, as it had proposed to improve access to pharmaceuticals by circumventing IPRs to access affordable medicines.²⁷ Schug provides an insightful account of the international legal regimes and new policy commitments concerning IPRs and access to medicines, which this research uses for background information on the transnational political contention concerning ARV prices; however, Schug does not address transnational AIDS treatment activism.

Kenneth Shadlen’s article, ‘The Political Economy of AIDS Treatment: Intellectual Property and the Transformation of Generic Supply,’ also provides an examination of IPRs and issues of ARV production and supply, while accounting for transnational AIDS treatment activism. Shadlen acknowledges that the reduction of ARV prices between 2000 and 2001 was significant and that transnational activism was influential. However, Shadlen argues that these initial gains may not

²² Joseph Stiglitz, *Making Globalization Work* (London: Allen Lane, 2006), 107.

²³ Joseph DiMasi et al., "Cost of Innovation in the Pharmaceutical Industry," *Journal of Health Economics* 10, no. 2 (1991): 107-108.

²⁴ Stiglitz, *Making Globalization Work*, 108.

²⁵ Schug, "Promoting Access to HIV/AIDS Pharmaceuticals," 239.

²⁶ Raymond Smith and Patricia Siplon, *Drugs into Bodies: Global AIDS Treatment Activism* (Westport, Conn.: Praeger, 2006), 53.

²⁷ Nicoli Nattrass, *Mortal Combat: AIDS Denialism and the Struggle for Antiretrovirals in South Africa* (KwaZulu-Natal: University of KwaZulu-Natal Press, 2007), 50.

happen again given that the pharmaceutical industry must have new incentives to produce new ARVs, as resistance to older ARVs will occur in HIV and AIDS patients.²⁸ Shadlen's article provides a solid understanding of the problem this study addresses, however it does not account for how transnational activism contributed to the initial reduction in ARV prices, it merely describes certain events. Therefore, this dissertation will critically assess how transnational AIDS treatment activism emerged and identify the ways in which it contributed to early global ARV price reductions.

Lastly, Debora Halbert's book, *Resisting Intellectual Property*, focuses on the issue of ARVs and IPRs and explains that pharmaceutical patents are just one contentious issue concerning IPRs, US power and the WTO's TRIPS agreement. Halbert accounts for how transnational AIDS treatment activists began working on the issue of IPRs, but her examination is largely descriptive and does not employ an analytical framework to guide her study.²⁹ Therefore, by applying Tarrow's theoretical arguments to Halbert's descriptive account of transnational AIDS treatment activism this study will be a useful contribution, providing a foundation for future research.

In summary, literature on the contentious issue of ARV prices and IPRs provides this study with an understanding of the problem surrounding ARV price reductions. However, these three accounts, merely state that transnational activists were working on the issue of ARV prices and do not explain whether or how transnational AIDS treatment activism made an impact on ARV price reductions. This research and dissertation will provide a closer empirical look at the work and impact of Health GAP and TAC in relation to the transnational political contention concerning AIDS treatment access.

1.2.2 Transnational AIDS Treatment Activism

The second category of literature used by this study and reviewed here focuses on transnational AIDS treatment activism during the period from 1996 to 2001 and use older theories of transnational activism, which Tarrow's work builds upon, developed by Margaret Keck and Kathryn Sikkink in their 1998 book *Activists Beyond Borders: Advocacy Networks in International Politics*.³⁰

Susan Sell and Aseem Prakash examine transnational AIDS treatment activism, or what they term the Access Campaign, by applying and testing Keck and Sikkink's theoretical arguments about transnational advocacy networks. They argue against Keck and Sikkink's hypothesis, that transnational activism and advocacy networks are driven solely by ideational or normative interests.

²⁸ Kenneth Shadlen, "The Political Economy of AIDS Treatment: Intellectual Property and the Transformation of Generic Supply," *International Studies Quarterly* 51, no. 3 (2007): 559-581.

²⁹ Debora Halbert, *Resisting Intellectual Property* (New York: Routledge, 2005).

³⁰ Margaret Keck and Kathryn Sikkink, *Activists Beyond Borders : Advocacy Networks in International Politics* (Ithaca, N.Y.: Cornell University Press, 1998).

Rather, Sell and Prakash claim that transnational AIDS treatment activists had both normative (promoting human rights to access health care) and material interests (reducing ARV prices) and therefore examination of transnational activism must not only focus on normative goals and factors.³¹ As insightful as their article is, they do not mention TAC or events and developments in South Africa, making their article limited in scope. Secondly, their account extends beyond the period of this study and conflates certain events and policies, which weakens their causal explanation. Hence, this dissertation ensures a chronological order of events and developments in the US, South Africa, and internationally to provide a causal examination of the impact of transnational AIDS treatment activism.

Raymond Smith and Patricia Siplon also use Keck and Sikkink's framework to provide a comprehensive study of transnational AIDS treatment activism, by integrating key events, organizations, individuals and political outcomes from both the United States and South Africa in their book, *Drugs Into Bodies – Global AIDS Treatment Activism*. However, they acknowledge that they provide a US-based perspective and argue that it was the “merging of the US-based AIDS treatment activist movement into a larger transnational issue advocacy network” that contributed to specific political outcomes and ARV price reductions.³² Therefore, their book is also limited in scope as it focuses primarily on the work of US based organizations. To address this limitation, this research will provide a comprehensive and holistic examination by focusing equally on the work of Health GAP and TAC.

Thomas Olesen's account of transnational AIDS treatment activism is an insightful analysis, but is primarily concerned with explaining the abandonment of the South African PMA lawsuit in 2001. Olesen applies traditionally state-centric social movement theories and concepts to provide a critical assessment of how transnational activists won in the “international court of public opinion.”³³ However, Olesen's study does not account for changes in the international political structure and is rather narrow in scope as it is only concerned with domestic developments and issues in South Africa. Olesen concludes his study by stating that future research on “transnational campaign success” must account for “structural elements of explanations,”³⁴ which this study does by tracing changes in the international political structure by way of new international policy commitments, programs, and funding for addressing HIV and AIDS.

³¹ Susan Sell and Aseem Prakash, "Using Ideas Strategically: The Contest between Business and NGO Networks in Intellectual Property Rights," *International Studies Quarterly* 48, no. 1 (Mar. 2004): 153.

³² Smith and Siplon, *Drugs into Bodies*, 70.

³³ Thomas Olesen, "'In the Court of Public Opinion': Transnational Problem Construction in the HIV/AIDS Medicine Access Campaign, 1998-2001," *International Sociology* 21, no. 1 (2006): 5-30.

³⁴ Ibid.

In summary, the literature on transnational AIDS treatment activism does provide a platform for this study. However, as most research uses older theories of transnational activism and social movements, which Tarrow's theoretical arguments build upon, this study will provide a current analysis of transnational activist theories. Furthermore, it will provide a new and comprehensive examination of transnational AIDS treatment activism in the US and South Africa.

1.2.3 Health GAP and TAC

The third body of literature used by this research and reviewed here specifically focuses on TAC and Health GAP, and includes both personal accounts and academic articles on the work of both organizations. Literature about Health GAP primarily focuses on its member organizations, particularly the work of ACT UP. In *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization*, edited by Benjamin Shepard and Ronald Hayduk, a number of chapters discuss the involvement of ACT UP and its efforts in transnational AIDS treatment activism under Health GAP.³⁵ However, this book's personal accounts by activists are largely descriptive of direct action strategies and events, and provide a rather simple and overly deterministic assessment of their impact on improving AIDS treatment access. Furthermore, Geffen's personal account in his recently published book, *Debunking Delusions – The Inside Story of the Treatment Action Campaign*, provides important insights into TAC, but does not offer much detail about or account for, transnational AIDS treatment activism outside of South Africa. For the most part, Geffen's book focuses on South Africa's domestic AIDS policies and activist struggles after 2001, but it does provide useful information for this study.

Academic accounts of Health GAP and TAC typically describe fundamental elements of each organization and the strategies used in particular campaigns. For example, Mandisa Mbali provides an examination of the influences of TAC as "one of the most successful post-apartheid social movements."³⁶ Mbali explains that past activist networks, strategies, and resources used during the anti-apartheid struggle formed the foundation of TAC, which pressured the South African government to provide free ARVs for HIV-positive mothers in 2001 and forced the government to develop a National Treatment Plan in 2004.³⁷ However, Mbali does not provide a causal analysis or account for important international factors and connections with other transnational AIDS treatment activists, which this dissertation will do.

³⁵ Benjamin Shepard and Ronald Hayduk, eds., *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization* (London: Verso, 2002).

³⁶ Mandisa Mbali, "The Treatment Action Campaign and the History of Rights-Based, Patient-Driven HIV/AIDS Activism in South Africa," University of KwaZulu-Natal Center for Civil Society, 2005, accessed September 13, 2011, www.nu.ac.za/Ccs/files/RRreport_29.pdf, 21.

³⁷ Ibid.

Steven Friedman and Shauna Mottiar do however explain that TAC's "most strategically important alliance may have been the ones formed with international allies...It placed pressure on pharmaceutical companies because their head offices abroad feared being portrayed as unsympathetic to the poor."³⁸ Furthermore, they argue, "international activism remains a key resource for TAC,"³⁹ but provide little explanation as to how international alliances were formed, how efforts became coordinated, and how these alliances helped TAC be more effective. Steven Robins and Bettina von Lieres, conclude their article by stating that TAC "managed to convince international public opinion, and the [PMA], that their cause was undeniably right and just,"⁴⁰ but again, do not account for factors outside of TAC and South Africa. Eduard Grebe's, 'The Treatment Action Campaign's Struggle for AIDS Treatment in South Africa: Coalition-building Through Networks' provides the most instructive and analytical study of TAC and its international support by other transnational AIDS treatment activists. Grebe also draws upon the work of Keck and Sikkink to inform his study, but focuses on country specific outcomes in South Africa after global ARV price reductions in 2001.⁴¹

In summary, literature on transnational AIDS treatment activism based in South Africa and the US in relation to ARV price reductions appears narrow in scope and deterministic in its causal assessments. Studies tend to equate direct action or protests with specific outcomes and tend to mix and blend campaigns and events to suit research agendas and arguments. Furthermore, studies on TAC and Health GAP lack theoretical frameworks to examine these organizations in relation to international factors and political developments. Therefore, this dissertation will draw upon Tarrow's theoretical arguments in order to explain the emergence of transnational AIDS treatment activism, how activists in the US and South Africa began working together, and lastly how transnational AIDS treatment activism contributed to global ARV price reductions throughout 2000 and 2001.

1.3 Methodology, Research Design and Limitations

Given the empirical and theoretical research objectives of this dissertation, it uses the theoretical single case study method to provide a detailed account of transnational AIDS treatment activism by applying and testing Tarrow's theoretical arguments concerning how transnational activism emerges and is effective. This study primarily focuses on Health GAP and TAC in the US and South Africa from 1996 to 2001 in order to provide a causal examination of the relationship between transnational

³⁸ Steven Friedman and Shauna Mottiar, "Seeking the High Ground - the Treatment Action Campaign and the Politics of Morality," in *Voices of Protest*, Richard Ballard, Adam Habib and Imraan Valodia, eds. (KwaZulu-Natal: University of KwaZulu-Natal Press, 2006), 35.

³⁹ *Ibid.*, 36.

⁴⁰ Steven Robins and Bettina von Lieres, "Remaking Citizenship, Unmaking Marginalization: The Treatment Action Campaign in Post-Apartheid South Africa," *Canadian Journal of African Studies* 38, no. 3 (2004): 577.

⁴¹ Eduard Grebe, "The Treatment Action Campaign's Struggle for AIDS Treatment in South Africa: Coalition-Building through Networks." *Journal of Southern African Studies* 37, no. 4 (2011): 849-868.

AIDS treatment activism and the reduction of global ARV prices. Furthermore, these parameters were chosen because of clear time frames, geographical boundaries, and measurable outcomes.

This research uses primary and secondary literature for the case study. It relies upon academic work, personal activist accounts, and domestic and international governmental reports. It is designed to provide a complete narrative by triangulating sources “from three or more perspectives,”⁴² which according to Roger Pierce helps accomplish this. Furthermore, narrative accounts, according to Robert Yin, aid in explanation building, which this study attempts to do.⁴³ Yin states, “that narrative can be a useful tool for assessing causality in situations where temporal sequencing, particular events, and path dependence must be taken into account.”⁴⁴ It is worth noting that the value of narrative accounts is debated. However, narrative accounts are still considered to have “the obvious strength of allowing the analyst to show sensitivity to detail, process, conjuncture, and causal complexity.”⁴⁵

The theoretical single case study method is the examination “of a historical episode to develop or test historical explanations that may be generalizable to other events.”⁴⁶ This method however, has also received criticism, as scholars question its usefulness, however, single case studies are instructive because they can provide a detailed examination, contextual description, and be used to develop and test theories.⁴⁷ Furthermore, according to Alexander George and Andrew Bennett, case studies are “a useful means to closely examine the hypothesized role of causal mechanisms.”⁴⁸ The case study method allows for identifying “what conditions present in a case activate the causal mechanism,”⁴⁹ and remains a useful method for “assessing whether and how a variable mattered to the outcome.”⁵⁰ Lastly, George and Bennett argue that the most useful assessment of causality “concerns the relationship of a variable to conjunctions of variables that are themselves necessary and/or sufficient for an outcome.”⁵¹ By examining transnational AIDS treatment activism in relation to both domestic and international structural changes, which follows Tarrow’s argumentation, this research will be able to examine transnational activism in relation to other variables that may help explain how and whether transnational AIDS treatment activism contributed to the reduction of global ARV prices.

⁴² Roger Pierce, *Research Methods in Politics: A Practical Guide* (Los Angeles: Sage, 2008), 90.

⁴³ Robert Yin, *Case Study Research: Design and Methods*, 4th ed., vol. 5 (Los Angeles: Sage, 2009), 141.

⁴⁴ James Mahoney, "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis," *American Journal of Sociology* 104, no. 4 (Jan. 1999): 1164.

⁴⁵ *Ibid.*, 1168.

⁴⁶ Alexander George and Andrew Bennett, *Case Studies and Theory Development in the Social Sciences* (Cambridge, Mass.: MIT Press, 2004), 5.

⁴⁷ Yin, *Case Study Research: Design and Methods*, 47.

⁴⁸ George and Bennett, *Case Studies and Theory Development*, 19.

⁴⁹ *Ibid.*, 21.

⁵⁰ *Ibid.*, 25.

⁵¹ *Ibid.*, 26.

This study is limited in the following ways. First, it is limited because it only focuses on one outcome, the reduction of global ARV prices between September 2000 and February 2001. Furthermore, it is limited as it primarily examines transnational AIDS treatment activism in the US and South Africa. There are other transnational AIDS activist organizations around the world in developing countries, but they emerged after the period of this research, because countries such as Brazil and Thailand had government supported ARV programs from 1996.⁵² Lastly, this research is limited because it is reliant on subjective accounts and sources provided by individuals and organizations involved in the transnational activism focused on in this case study. This research uses information and documents provided by TAC, Health GAP, CPT, and ACT UP. Despite these limitations, which also give this study focus and clear parameters, this dissertation provides a comprehensive examination of transnational AIDS treatment activism and its contribution to global ARV price reductions.

1.4 Empirical Findings and Theoretical Implications

By drawing upon Tarrow's theoretical arguments to examine the suggested causal relationship between transnational AIDS treatment activism and the reduction of global ARV prices from 1996 to 2001, this dissertation makes empirical findings that have theoretical implications, and argues the following. First, the causal relationship and contribution of transnational AIDS treatment activism to global ARV price reductions is limited and is best understood in conjunction with separate and broader changes in the international political structure. Second, this research argues that Tarrow's hypothesis concerning the domestic origins, processes, and impact of transnational activism is instructive, and argues that his transitional and international processes, along with his conceptualization of internationalization, need further theoretical development.

1.5 Significance

Both the empirical findings and theoretical implications of this research are significant and contribute to the current literature on transnational activism and more specifically transnational AIDS treatment activism between 1996 and 2001. First, given that transnational AIDS treatment activism had a limited contribution to ARV price reductions and its impact is best understood in conjunction with broader changes in the international political structure, this research clarifies and illustrates specific ways transnational activism affected global ARV price reductions. Second, given the theoretical implications of this case study, which support and challenge Tarrow's theoretical arguments, current limitations of transnational activism theory are highlighted and therefore contribute to theory building by providing new areas for future research. In summary, this dissertation is significant because it uses and tests theory to provide a comprehensive examination of the effectiveness of transnational AIDS treatment activism in the US and South Africa, which is lacking in the current body of literature.

⁵² Natrass, "The (Political) Economics of Antiretroviral Treatment," 574.

1.6 Organization of the Study

This dissertation is organized into five chapters. This first chapter has introduced the objectives and the case study of this research. Additionally, it has outlined Tarrow's theoretical arguments, which form the analytical framework used and tested in this dissertation. Furthermore, this chapter has explained the methodology, research design and limitations, along with the significance of this study. The second chapter of this dissertation provides an analysis of Tarrow's theoretical arguments used here to explain the emergence and contribution of transnational AIDS treatment activism to global ARV price reductions. The third chapter provides a detailed and chronological account of the emergence of transnational AIDS treatment activism domestically in the US and South Africa, while examining international political developments for addressing the HIV and AIDS epidemic from 1996 to 2001. The fourth chapter employs Tarrow's theoretical arguments to examine the suggested causal relationship between transnational AIDS treatment activism and global ARV price reductions. Lastly, the fifth chapter summarizes the empirical findings and theoretical implications of this theoretical case study, and proposes areas for future research.

Chapter 2: Analytical Framework – Tarrow’s Transnational Activism

The aim of this chapter is to provide an analysis of Tarrow’s theoretical arguments concerning how transnational activism emerges and is effective, which inform the analytical framework developed in this study to examine transnational AIDS treatment activism. The first section of this chapter outlines Tarrow’s claims and concepts concerning the domestic origins, processes, and impact of transnational activism. The second section of this chapter examines the ways in which, according to Tarrow, transnational activism becomes coordinated in more than one country and extends across state borders to be most effective. The third section considers Tarrow’s argument regarding changes in the international political structure, which he claims creates new international political opportunities and threats that provide transnational activists new targets, resources, and avenues to be effective. Lastly, the final section of this chapter discusses the design of the case study in Chapter 3 and the analysis to come in Chapter 4.

2.1 Domestically Rooted Transnational Activism

The first theoretical argument developed by Tarrow and used by the analysis of this study concerns his hypothesis that transnational activism is still very much a product of domestic-state factors, politics, and processes, and “will have its most visible impact on domestic politics.”⁵³ Tarrow argues that individuals who engage in transnational activism rely upon domestic resources, networks, and domestic politics to form effective strategies and organizations. More specifically, Tarrow identifies individuals involved in transnational activism first as *rooted cosmopolitans*, which refers to people “who mobilize domestic and international resources and opportunities to advance claims on behalf of external actors, against external opponents, or in favor of goals they hold in common with transnational allies.”⁵⁴ *Transnational activists* fall within the broader category of rooted cosmopolitans, but are different because they are based “in specific national contexts, but who engage in contentious political activities that involve them in transnational networks of contacts and conflicts.”⁵⁵ Provided Tarrow’s domestic argument, this section outlines factors and concepts borrowed from traditionally domestic-orientated social movement theories⁵⁶ and domestic processes developed by Tarrow, that will help guide the domestic analysis of transnational AIDS treatment activism in the US and South Africa.

⁵³ Tarrow, *The New Transnational Activism*, 219.

⁵⁴ *Ibid.*, 29.

⁵⁵ *Ibid.*

⁵⁶ See Marc Edelman, "Social Movements: Changing Paradigms and Forms of Politics," *Annual Review of Anthropology* 30 (2001): 285-317, for an overview of social movement theories.

2.1.1 Four Key Factors and Concepts from Social Movement Theories

The first factor is the *mobilizing form and structure* of transnational activist organizations, which may be formal or informal, such as “temporary assemblies of challengers, to informal networks, to formal branches, clubs, and even military-like cells.”⁵⁷ Furthermore, mobilizing forms are a reflection of material resources (people, funding, information) and non-material resources (identity, shared history, experiences), which help explain how transnational activism emerges in the way that it does and is most effective.

The second factor and concept concerns *repertoires of contention*, meaning the various forms of contentious collective action, such as protests, boycotts, and sit-ins, counter summits, which may involve, “information diffusion, institutional access, and direct action,”⁵⁸ to be effective. Furthermore, according to social movement scholar Charles Tilly, repertoires of contention grow out of three kinds of factors: “a population’s daily routines and internal organization, the prevailing standards of rights and justice, and a population’s accumulated experience with collective action.”⁵⁹ Additionally, repertoires of contention emerge through active learning, in terms of “what people *know* about how to contend in various places and at different periods of history.”⁶⁰

The third factor and concept is *framing*, which refers to the construction of and dissemination of particular meanings and ideas relating to a specific contentious issue. “A frame is an interpretive schemata that simplifies and condenses the ‘world out there’ by selectively punctuating and encoding objects, situations, events, experiences, and sequences of actions within one’s present or past environments.”⁶¹ Moreover, frames are used to affect and produce “consensus and action mobilization,” and typically fall within three categories: 1) diagnostic framing of the problem activists are addressing, 2) prognostic framing of the solution activists want, and 3) motivational framing to increase support and mobilization.⁶²

The fourth factor and concept refers to *political opportunities and threats*, understood as “consistent – but not necessarily formal or permanent – dimensions of the political environment that provide

⁵⁷ Sidney Tarrow, *Power in Movement: Social Movements and Contentious Politics*, 2nd ed. (Cambridge, U.K.: Cambridge University Press, 1998), 124.

⁵⁸ Tarrow, *The New Transnational Activism*, 147.

⁵⁹ Charles Tilly, *The Contentious French* (Cambridge, Mass.: Belknap Press, 1986), 10.

⁶⁰ Tarrow, *The New Transnational Activism*, 102.

⁶¹ David Snow and Robert Benford, "Master Frames and Cycles of Protest," in *Frontiers in Social Movement Theory*, eds. Aldon D. Morris and Carol Mueller (New Haven, Conn.: Yale University Press, 1992), 137.

⁶² Scott Hunt, Robert Benford and David Snow, "Identity Fields: Framing Processes and the Social Construction of Movement Identities," in *New Social Movements - from Ideology to Identity*, eds. Enrique Larana, Hank Johnston and Joseph Gusfield (Philadelphia, Pa.: Temple University Press, 1994), 191.

incentives for collective action by affecting people's expectations for success or failure."⁶³ It is important here to note that the conceptualization of political opportunities in the study of social movements is highly debated. Jeff Goodwin, James Jasper and Jaswinder Khattra, provide an insightful critique of the concept. They argue that the concept is either too expansive, clumping any and all variables into a political opportunity, or too narrow and does not explain anything about contentious collective action.⁶⁴ Additionally, they argue there is a flawed structural bias in the way that political opportunities are understood. Political opportunities, as a catalyst for action, typically arise from a structural shift or change in policies, political leaders, or resources; however, the political structure by which activists engage with inherently implies a sort of stability or consistency not conducive for such changes.⁶⁵ Nevertheless, Tarrow's argues that political opportunities and threats "should not be understood as an invariant model inevitably producing social movements, but as a set of clues."⁶⁶

By using these four conceptualizations to domestically analyze transnational activism, it is possible to examine and understand how transnational activist organizations and contentious collective action emerges and takes the shape that it does. However, examining these factors alone leads to a more descriptive analysis and therefore must be taken into account with domestic processes that link domestic contention to international actors, pressures, and targets.

2.1.2 Domestic Processes: Global Issue Framing and Internalization

According to Tarrow, there are two domestic processes that link domestic and international actors, targets, and resources together in transnational activism: *global issue framing* and *internalization*. First, global issue framing, refers to "the mobilization of international symbols to frame domestic conflicts."⁶⁷ These symbols include the WTO, 'globalization' and the US, and are used to focus and solidify domestic grievances in relation to international political and economic issues. Tarrow argues that global issue framing may emerge in two ways, which both involve communication and convergence, as illustrated in the second stage of Figure 1 (on the following page). First, global issue framing may occur due to *structural equivalence*, which refers to similar domestic episodes of contention in response to shared and common external threats that are structural in nature, such as structural adjustment policies from the International Monetary Fund (IMF).⁶⁸ Second, global issue framing may occur due to *global thinking* of activists, which tentatively refers to the idea of "global

⁶³ Tarrow, *Power in Movement*, 76-77.

⁶⁴ Jeff Goodwin, James Jasper and Jaswinder Khattra, "Caught in a Winding, Snarling Vine: The Structural Bias of Political Process Theory," *Sociological Forum* 14, no. 1 (Mar. 1999): 31.

⁶⁵ *Ibid.*, 29.

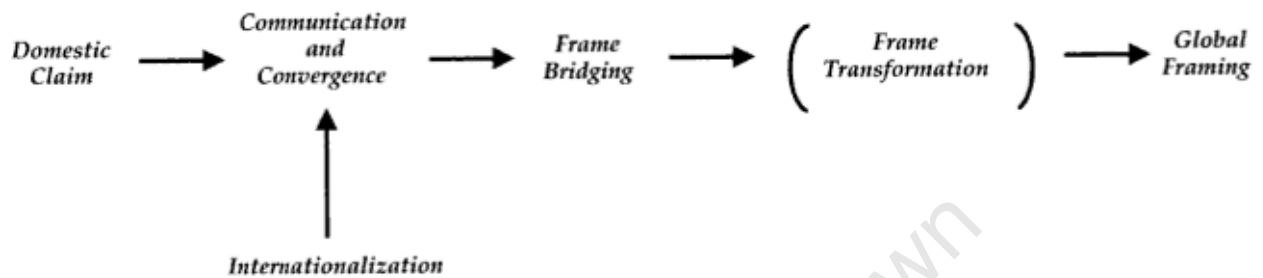
⁶⁶ Tarrow, *Power in Movement*, 20.

⁶⁷ Tarrow, *The New Transnational Activism*, 32.

⁶⁸ *Ibid.*, 64.

citizenship” of particular individuals who are well informed of global contentious issues and therefore frame domestic claims in alignment with contentious issues located elsewhere.⁶⁹ Lastly, Tarrow argues that global issue framing is a process that “can dignify, generalize, and energize activists whose claims are predominantly local, linking them symbolically to people they have never met and to causes that are distantly related to their own.”⁷⁰

Figure 1. Model of Global Issue Framing⁷¹



The second domestic process Tarrow identifies is *internalization*, “the migration of international pressures and conflicts into domestic politics and the triangular relationship that this creates among ordinary people, their governments, and international institutions.”⁷² In other words, internalization is a process by which external pressure, from foreign states or international institutions, affects domestic politics, which is then met by pressure from transnational activists that “employ contentious politics against international, state, or non-state actors on domestic ground.”⁷³ This leads to a “two-level game”⁷⁴ in which state governments enter into external/international negotiations as well as internal/domestic negotiations, resulting in either repression of domestic contentious claims, the offering of concessions, or brokering between domestic and international demands, which is illustrated in Figure 2 (on the following page).⁷⁵ Tarrow’s main argument is that when international or external pressure impinges on domestic politics and socio-economic policies, protests are more likely to be directed towards state governments.⁷⁶ An example of internalization Tarrow uses is when Spanish fishermen, upset with European Union (EU) fishing quotas and policies, targeted and embarrassed their own government to improve EU policies in their favor.⁷⁷ Furthermore, this process

⁶⁹ Ibid., 68-72.

⁷⁰ Ibid., 60.

⁷¹ Ibid., 63.

⁷² Ibid., 80.

⁷³ Ibid., 79.

⁷⁴ Robert Putnam, “Diplomacy and Domestic Politics: The Logic of Two-Level Games,” *International Organization* 42, no. 3 (1988): 427-460.

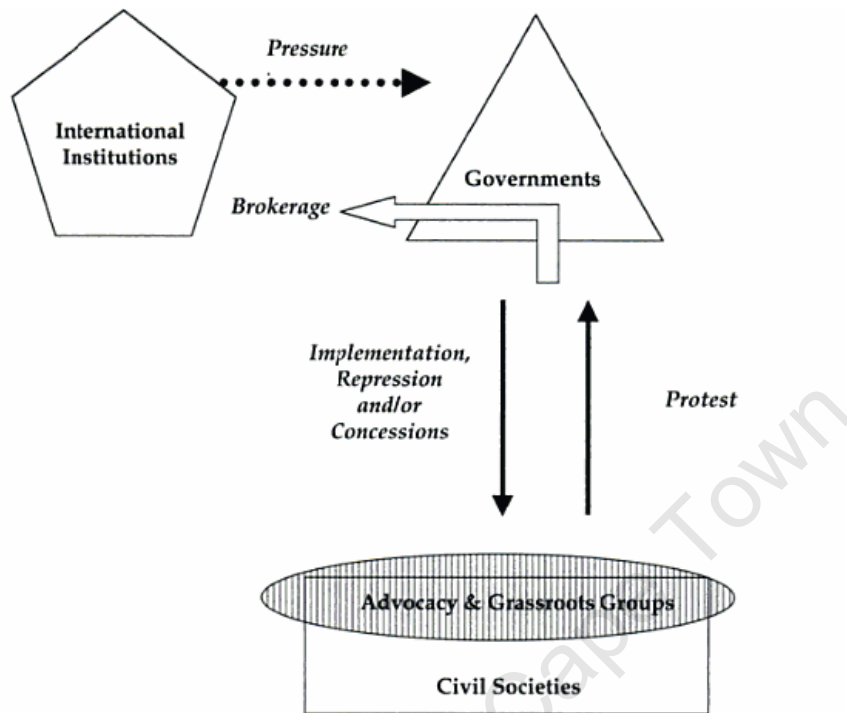
⁷⁵ Tarrow, *The New Transnational Activism*, 80.

⁷⁶ Ibid.

⁷⁷ Ibid., 84.

is evident in the case of transnational AIDS treatment activism in South Africa and will be discussed further in Chapter 4.

Figure 2. Model of Internalization⁷⁸



Both of Tarrow's domestic processes help explain how episodes of domestic political contention take on an international dimension. The first through framing a domestic struggle in terms of international symbols or targets and the second through protesting domestic policies influenced by international institutions. However, it is worth mentioning here that Tarrow's two processes have limitations. Both processes do not adequately account for non-state actors, such as multinational pharmaceutical corporations. This research suggests that this is problematic as the power and influence corporations have is complicated and also rooted in domestic politics and should therefore be taken into account. Nevertheless, by applying and testing Tarrow's domestically oriented arguments to the case of transnational AIDS treatment activism, this study will make empirical findings that highlight the strengths of his arguments.

2.2 Coordinating Transnational Activism Across Borders

The second set of theoretical arguments used in this study to analyze transnational AIDS treatment activism, concern Tarrow's transitional and international processes, which he claims explain the expansion, coordination, and effectiveness of transnational activism in more than one country. This

⁷⁸ Ibid., 81.

section examines each of Tarrow's processes and highlights strengths and weaknesses that are informative and problematic for this analysis of transnational AIDS treatment activism.

2.2.1 Transitional Processes: Diffusion and Scale Shifting

The first transitional process that links domestic political contention to external and international actors and environments is *diffusion*, which is "the transfer of claims or forms of contention from one site to the other."⁷⁹ Diffusion is the adoption of contentious claims, targets, and strategies from originators to adopters across the globe. In Figure 3 (on the following page), there are three different pathways of diffusion. First is *relational diffusion* through "well connected trust networks,"⁸⁰ which helps transfer crucial information, strategies, tactics, and resources "along established lines of interaction through the attribution of similarity."⁸¹ Established professional and personal relationships among like-minded activists form a strong pathway through which to mobilize external and additional support, however, this also makes campaigns heavily reliant on "pre-existing ties of those who pass on the message."⁸² According to Tarrow, transnational activism scholars focus primarily on relational diffusion, which in his view is too narrow because "in an age of massive immigration and cheap and easy transportation, information about collective action can spread through third parties, or brokers, who connect people who would otherwise have no contact with one another."⁸³ Therefore, Tarrow proposes two other pathways of diffusion: *non-relational diffusion* and *mediated diffusion*.

Non-relational diffusion refers to the transfer of claims or information through mainstream mass media, electronic communication and the internet. Furthermore, non-relational diffusion through the internet "has a greater reach, but its impersonal nature makes it impact thinner than that of relational diffusion."⁸⁴ Mediated diffusion occurs through "movement brokers," individuals or organizations that engage in *brokerage* "the connection of two unconnected sites by a third, which works through movement 'halfway houses', immigrants, or institutions."⁸⁵ An example of all three pathways of diffusion Tarrow cites is the adoption of Mohandas Gandhi's non-violent protesting from India, which was later used in the American civil rights movement under Dr. Martin Luther King Jr.⁸⁶

⁷⁹ Ibid., 32.

⁸⁰ Ibid., 101.

⁸¹ Ibid., 104.

⁸² Ibid., 208.

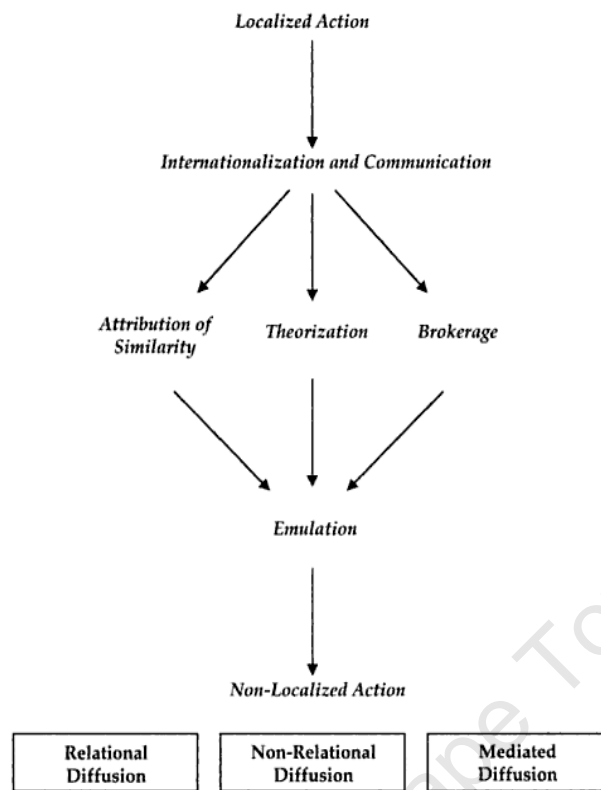
⁸³ Ibid., 105.

⁸⁴ Ibid., 208.

⁸⁵ Ibid., 104.

⁸⁶ Ibid., 106-108.

Figure 3. Alternative Pathways of Transnational Diffusion⁸⁷



In summary, the diffusion of domestic claims and localized action from one state to another occurs in a number of ways. However, examining these pathways and their potential impact on the emergence of transnational activism is difficult and frequently results in a thick description rather than a causal assessment of how transnational activism becomes coordinated in more than one country.

The second transitional process Tarrow identifies is *scale shifting*, meaning “the change in the number and level of coordinated contentious actions to a different focal point, involving a new range of actors, different objects, and broadened claims.”⁸⁸ According to Tarrow, scale shifting can occur both upwards, when “local action spreads outward from its origins” and downwards, “when a generalized practice is adopted at a lower level.”⁸⁹ An example of downward scale shifting given by Tarrow is the re-creation of the World Social Forum model, a counter-summit of the World Economic Summit in Davos, Switzerland, which emerged domestically in countries around the world as part of the global justice movement.⁹⁰

⁸⁷ Ibid., 105.

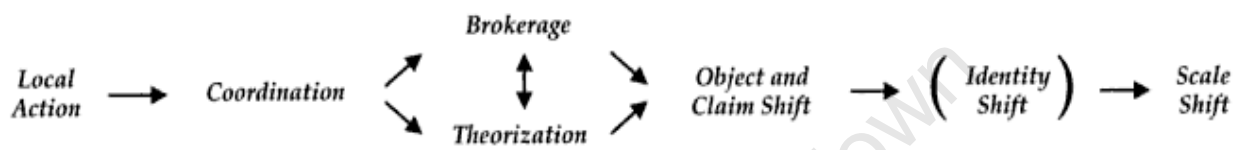
⁸⁸ Doug McAdam, John McCarthy and Charles Tilly, *Dynamics of Contention* (Cambridge, U.K.: Cambridge University Press, 2001), 331.

⁸⁹ Tarrow, *The New Transnational Activism*, 121.

⁹⁰ Ibid., 128-134.

Scale shifting, illustrated in Figure 4 below, begins with *coordination*, or organizing and planning of collective action by activists in more than one state. Then, diffusion through any or all three pathways occurs, resulting in *target/object and claim shifts* from the original location to another location. Meaning that domestic contentious claims became adopted and supported by transnational activists in an additional location.⁹¹ It is important to note that scale shifting involves “a move from familiar domestic structures of opportunity and constraint to new terrains and the need to forge new alliances with different allies against different opponents,”⁹² which is costly and difficult for individuals and organizations.

Figure 4. Model of Scale Shifting⁹³



While Tarrow’s process of scale shifting does outline steps concerning how political contention extends across borders and shifts to different levels in the international system, some aspects are vague and require further explanation. Especially as scale shifting directly builds upon diffusion and given the rather simple and implied linear projection of each of these processes from one area to another. Furthermore, Tarrow’s conceptualization of ‘levels’ and upwards or downwards scale shifting in the international system need further clarification, especially in relation to the position of the US in the international system. Jackie Smith argues, which this research agrees with, that research and analysis must be done “systematically about how US power affects transnational social change prospects and strategic options.”⁹⁴

2.2.2 International Processes: Externalization and Transnational Coalition Formation

Tarrow’s first international process is *externalization*, or “the vertical projection of domestic claims onto international institutions or foreign actors”⁹⁵ to gain external support. Tarrow explains that this process typically occurs when “domestic actors, frustrated at their inability to gain redress from their own governments,” actively search for “support of external allies”⁹⁶ to be more effective. It is important to note here that Tarrow’s externalization process is based upon the work of Keck and

⁹¹ Ibid., 122.

⁹² Ibid.

⁹³ Ibid., 123.

⁹⁴ Smith, “Book Review: The New Transnational Activism,” 548.

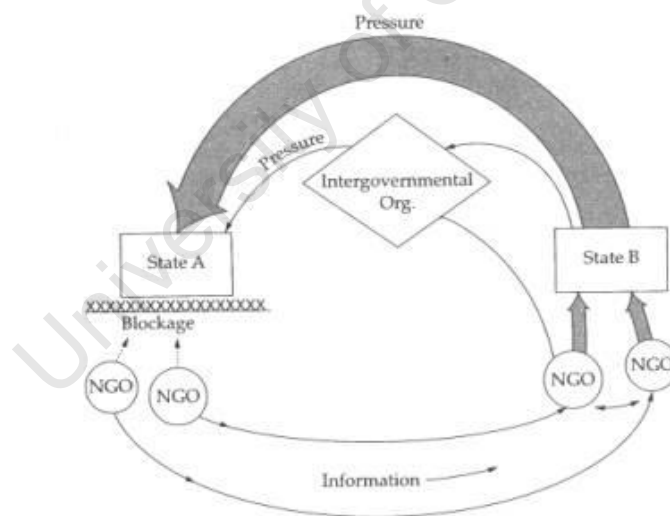
⁹⁵ Tarrow, *The New Transnational Activism*, 32.

⁹⁶ Ibid., 145.

Sikkink and their *boomerang pattern*, which Maria Rodrigues claims made transnational activism a “promising field of academic inquiry.”⁹⁷

The fundamental assumption of Keck and Sikkink’s theory is that international politics and all subsequent interaction between states, non-state actors, and international institutions is due to *networks*, defined as “forms of organization characterized by voluntary, reciprocal, and horizontal patterns of communication and exchange.”⁹⁸ Transnational advocacy networks, which transnational activism is a part of, are distinct because they “include those relevant actors working internationally on an issue, who are bound together by shared values, a common discourse, and dense exchanges of information and services.”⁹⁹ Keck and Sikkink argue that transnational advocacy networks emerge when individuals believe working together with transnational activists in other countries will help them be most effective, especially when “channels between domestic groups and their governments are blocked or hampered or where such channels are ineffective for resolving a conflict, setting into motion the ‘boomerang pattern’ of influence.”¹⁰⁰ The boomerang pattern, as illustrated in Figure 5, is most effective when transnational advocacy networks pressure and leverage more powerful Northern actors in order to place pressure on weaker and less powerful Southern state governments for a desired outcome.

Figure 5. The Boomerang Pattern¹⁰¹



Despite taking “the study of transnational contention to a new level of empirical specificity and theoretical sophistication,”¹⁰² there are limitations to Keck and Sikkink’s theoretical arguments.

⁹⁷ Maria Rodrigues, "Rethinking the Impact of Transnational Advocacy Networks," *New Global Studies* 5, no. 2 (2011): 1.

⁹⁸ Keck and Sikkink, *Activists Beyond Borders*, 8.

⁹⁹ *Ibid.*, 2.

¹⁰⁰ *Ibid.*, 12.

¹⁰¹ *Ibid.*, 13.

¹⁰² Tarrow, *The New Transnational Activism*, 146.

Firstly, Rodrigues comments on the assumption that organizations in Southern states benefit from or are empowered by Northern organizations. In her longitudinal comparison study of two transnational advocacy networks working on environmental sustainability in the Amazon regions of Brazil and Ecuador, Rodrigues illustrates that initial gains of local domestic activists were followed by negative outcomes and demobilization.¹⁰³ Rodrigues claims that the assumption that Southern activist organizations clearly benefit from resources and information from Northern organizations is shortsighted when examined over longer periods.¹⁰⁴ Secondly, Tarrow criticizes the boomerang pattern for being too simplistic and bilateral in the face of complex contentious politics involving many states and non-state actors.¹⁰⁵ Lastly, Keck and Sikkink's emphasis on shared values, norms, and ideas, the distinctive characteristic of transnational advocacy networks, limits the applicability of their framework. As mentioned in the literature review of Chapter 1, Sell and Prakash provide an insightful critique of Keck and Sikkink, which argues that advocacy networks are just like any other interest group and are driven by both material and ideational resources, goals, and strategies.¹⁰⁶ Therefore, analysis of transnational activism must be balanced giving attention to both ideational and material factors, which is what Mitchell Orenstein and Hans Schmitz claim Tarrow has done. They state that Tarrow's book moves "beyond the material/ideational divide in explaining social and political change,"¹⁰⁷ and that Tarrow charts "a pragmatic middle ground between ideational and interest-based analysis."¹⁰⁸

According to Tarrow, there are three important factors and steps involved in the process of externalization and are illustrated in Figure 6 (on the following page): domestic contexts, framing, and forms of collective action. The first factor is the domestic context and subsequent government reaction to domestic contention on a particular issue. Tarrow argues that governments may be repressive, unresponsive, or aid in the facilitation of externalizing contention by non-state actors when unable to act themselves.¹⁰⁹ The second factor Tarrow highlights is framing, he argues that domestic claims are uninteresting outside of a country unless the claims are framed to reach and appeal to a larger audience.¹¹⁰ Therefore, domestic claims must be re-framed to resonate and garner external support. The third factor in the externalization process is the type of direct action and/or strategies used to be most effective. According to Tarrow, there are three different forms of collective action; these include 1) the transmission and monitoring of information (the main focus of Keck and Sikkink)

¹⁰³ Rodrigues, "Rethinking the Impact of Transnational Advocacy Networks," 1-21.

¹⁰⁴ Ibid.

¹⁰⁵ Tarrow, *The New Transnational Activism*, 146.

¹⁰⁶ Sell and Prakash, "Using Ideas Strategically," 143.

¹⁰⁷ Mitchell Orenstein and Hans Schmitz, "The New Transnationalism and Comparative Politics," *Comparative Politics* 38, no. 4 (Jul. 2006): 485.

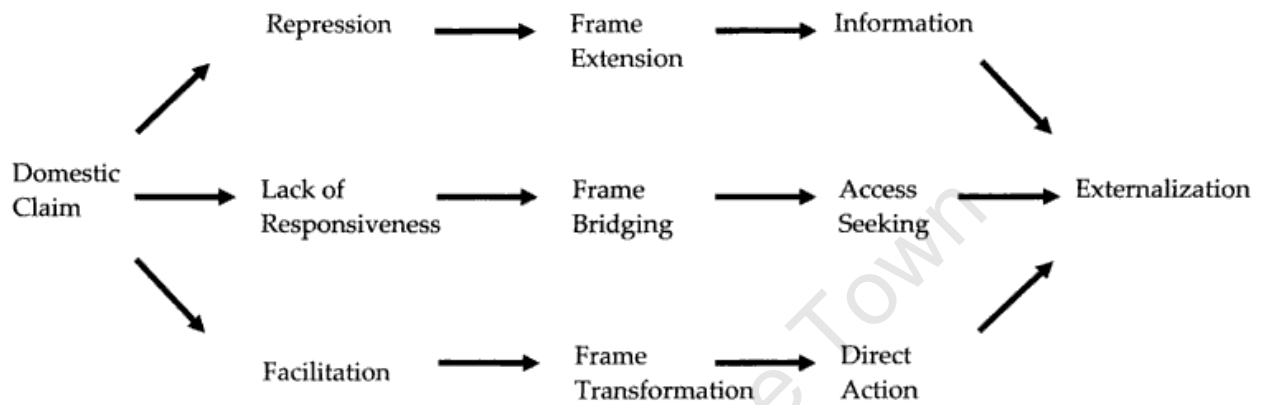
¹⁰⁸ Ibid., 494.

¹⁰⁹ Tarrow, *The New Transnational Activism*, 147.

¹¹⁰ Ibid.

2) access to international institutions, and 3) direct action and protests.¹¹¹ An example Tarrow cites for the process of externalization, is the work of NGOs in Chile during the 1970s, which monitored and reported mistreatment and human rights abuses that later strategically and successfully sought external support from international institutions and NGOs upholding human rights to intervene in Chile.¹¹²

Figure 6. Model of Externalization¹¹³



The major criticism here, again in agreement with Smith, is that externalization and upward scale shifting are very similar processes. Smith states that Tarrow’s “analytical distinctions seem to draw confusing and possibly misleading analytical boundaries.”¹¹⁴ It appears that the major difference between scale shifting and externalization is the additional emphasis placed on domestic contexts, which helps explain the beginning of the process, and collective action strategies, which form the end of the process. It is not clear if upward scale shifting and externalization analytically stand on their own, because any analysis would most likely include both contextual reasons behind collective action and the strategies employed. Therefore, distinguishing between externalization and scale shifting is not entirely useful to explain the emergence and coordination of transnational activism.

The second international process Tarrow identifies is the process of *transnational coalition formation*, which is “the horizontal formation of common networks among actors from different countries with similar claims.”¹¹⁵ Tarrow uses Margaret Levi and Gillian Murphy’s definition of coalitions, as “collaborative, means-oriented arrangements that permit distinct organizational entities to pool

¹¹¹ Ibid., 145.

¹¹² Ibid., 149-151.

¹¹³ Ibid., 148.

¹¹⁴ Smith, "Book Review: The New Transnational Activism," 547.

¹¹⁵ Tarrow, *The New Transnational Activism*, 32.

resources in order to effect change.”¹¹⁶ Essential for the emergence and creation of transnational coalitions, according to Levi and Murphy, are resources, framing, the establishment of trust, the ability to make credible commitments, and resolving tensions.¹¹⁷ Tarrow expands upon these factors by including three processes, which he claims account for the different forms of transnational coalitions.

First, Tarrow explains the process he calls *opportunity spirals*, which refers to changing domestic and international opportunity structures, how activist interpret these changes (as an opportunity or threat), strategize, and then act upon particular opportunities.¹¹⁸ Second is the process of *institutionalization*, meaning the centralization and creation of policies and parameters, which affect the sustainability of coalitions by ensuring *cooperative differentiation*, “maintaining a public face of solidarity towards their opponents while differentiating themselves in their relations with constituents.”¹¹⁹ Lastly, is the *socialization* process, meaning the “combination of discovery and solidarity that is experienced when people with very different backgrounds, languages, and goals encounter one another around a broad global theme.”¹²⁰

The factors and processes, mentioned above, contribute to Tarrow’s typology of transnational coalitions, illustrated in Figure 7 (on the following page). He argues there are *instrumental coalitions*, which are characterized by short-term cooperation and low levels of involvement, “around an occasional conjuncture of interest or program, but either drift apart or maintain purely formal ties after the issue that brought them together has dissipated.”¹²¹ *Event coalitions* are characterized by high levels of involvement during a short period of time, and “have the potential for future collaboration when they solder alliances among people who recognize their shared identities in the process of collective action.”¹²² *Federated coalitions* have a low degree of involvement over an extended period of time and are characterized by “long-term collaboration.”¹²³ Lastly, *campaign coalitions* “combine high intensity of involvement with long-term cooperation,”¹²⁴ and arguably provide sustainable results compared to the other transnational coalition forms. Tarrow uses the International Landmine Campaign during the 1990s, made up of NGOs, international institutions, and countries such as

¹¹⁶ Margaret Levi and Gillian Murphy, "Coalitions of Contention: The Case of the WTO Protests in Seattle," *Political Studies* 54, no. 4 (2006): 654.

¹¹⁷ Ibid.; Tarrow, *The New Transnational Activism*, 165.

¹¹⁸ Ibid., 176.

¹¹⁹ Ibid., 177.

¹²⁰ Ibid., 178.

¹²¹ Ibid., 168.

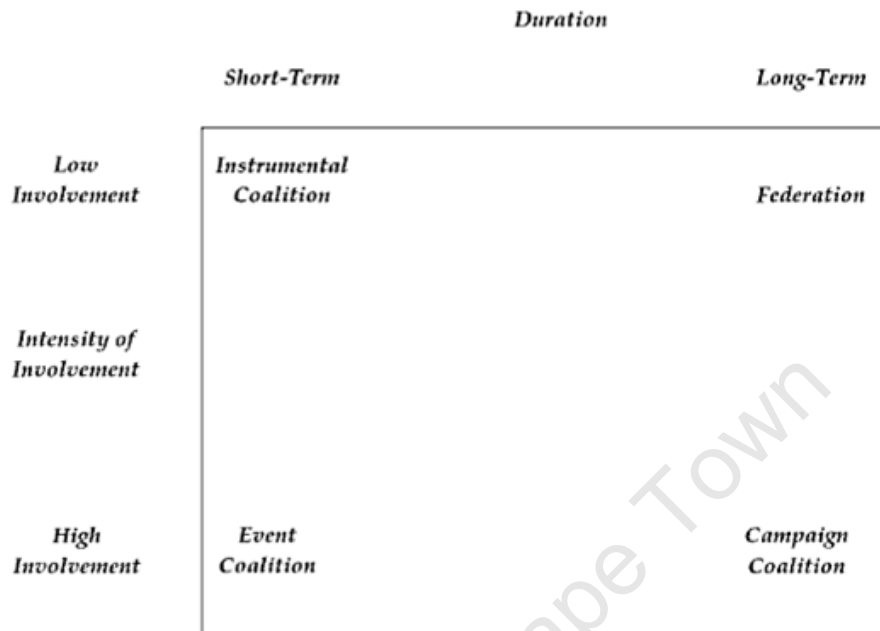
¹²² Ibid.

¹²³ Ibid.

¹²⁴ Ibid.

Canada and Norway, as an example of a successful transnational campaign coalition that mobilized long-term and effective support around the world.¹²⁵

Figure 7. Typology of Transnational Coalitions¹²⁶



Considering all four of Tarrow’s transitional and international processes, the process of transnational coalition formation is most instructive for this case study because it does not follow a linear projection and accounts for horizontal relationships made between similar goal-oriented organizations and will be discussed further in Chapter 4. In terms of the other three processes proposed by Tarrow, interesting factors emerge when considering this case, but the processes remain ambiguous and descriptive analytical tools. Critics, such as Simon Thompson agree and state that Tarrow’s book “is fundamentally a description of processes rather than a causal account of the phenomena it describes.”¹²⁷ Additionally, Richter claims that Tarrow’s models and processes are “skeletons without flesh.”¹²⁸ However, Carew Boulding states that Tarrow “offers a very solid foundation for future empirical work to rigorously test the hypotheses suggested here,”¹²⁹ which this research will do by examining and analyzing how transnational AIDS treatment activism emerged, became coordinated, and contributed to the reduction of global ARV prices in Chapter 3 and 4.

¹²⁵ Ibid., 173-175.

¹²⁶ Ibid., 167.

¹²⁷ Thompson, "Mapping Transnational Activism," 119.

¹²⁸ Richter, "The New Transnational Activism – Sidney Tarrow," 257.

¹²⁹ Carew Boulding, "Book Review: Tarrow, S. (2005). *The New Transnational Activism*. Cambridge, UK: Cambridge University Press," *Comparative Political Studies* 41, no. 6 (2008): 885.

2.3 International Political Opportunities and Threats

The third set of theoretical arguments proposed by Tarrow and used in the analysis of this study concern changes in the international political and economic structure, which according to Tarrow create new incentives, new focal points, expanded resources and new opportunities and threats for transnational activism. This section outlines Tarrow's conceptualizations and arguments, while highlighting points of concern for this theoretical case study.

2.3.1 Globalization, Internationalism, and Internationalization

According to Tarrow, shifts and changes in the international political structure by international institutions, like the World Bank, IMF, and United Nations (UN), affect the emergence and impact of transnational activism. Tarrow argues against the commonly held view that these changes or 'globalization' is the main driver behind the rise of transnational activism. Tarrow proposes that *globalization*, defined in rather narrow terms as "the increasing volume and speed of flows of capital and goods, information and ideas, people and forces that connect actors between countries"¹³⁰ merely provides new incentives and themes for transnational activism. Rather, according to Tarrow, *internationalism*, understood as "a dense, triangular structure of relations among states, non-state actors, and international institutions, and the opportunities this produces,"¹³¹ offers an operational framework for the emergence and impact of transnational activism. This dense and triangular structure is the product of what Tarrow argues and defines as *internationalization*, the process "through which the density of both horizontal and vertical ties expands and opportunities and threats are externalized,"¹³² by international institutions. Furthermore, Tarrow argues that internationalization "creates regular channels for communication and awareness of institutional similarities and differences among actors in different places," thus expanding and connecting the number of actors around the globe in new forms and creating new areas or venues for advancing contentious claims.¹³³ As such, tracing changes in the international political structure by way of international institutions is vital for understanding how transnational activism takes shape and makes an impact.

Tarrow's theoretical arguments concerning changes in the international political structure are essential for examining this case of transnational AIDS treatment activism from 1996 to 2001, as many new international political developments concerning HIV and AIDS took place at this time. However, Boulding argues that Tarrow "treats internationalization as a constant,"¹³⁴ and does not explain the complex interaction between structural changes and their impact on transnational activism.

¹³⁰Tarrow, *The New Transnational Activism*, 5.

¹³¹ Ibid., 25.

¹³² Ibid., 204.

¹³³ Ibid., 103.

¹³⁴ Boulding, "Book Review: Tarrow, S.(2005)," 886.

Additionally, Thompson claims that Tarrow “assumes but does not demonstrate that globalization and internationalization take the forms that he describes.”¹³⁵ As such, this dissertation will apply and test Tarrow’s arguments on what this research identifies as the *internationalization of AIDS* or the changing international structure for addressing the AIDS epidemic. In addition, it will examine the interaction and impact of these changes on transnational AIDS treatment activism and ARV price reductions. In doing so, this research will address both the empirical and theoretical objectives of this case study.

2.4 The Analysis to Come

This chapter has outlined and discussed three theoretical arguments developed by Tarrow, which form the foundation and framework of this dissertation’s case study in Chapter 3 and analysis in Chapter 4. The case study in Chapter 3 will be structured into domestic and international levels of analysis to account for both domestic and international politics, resources, and opportunities, which Tarrow argues are equally important for analyzing and explaining how transnational activism emerges and is effective. The analysis in Chapter 4 will be divided into three sections according to the three arguments analyzed here. First, given Tarrow’s argument that transnational activism is still very much a product of and has most impact on domestic state politics, the analysis will focus on domestic factors and contexts affecting transnational AIDS treatment activism in South Africa and the US between 1996 and 2001. Second, as Tarrow proposes two sets of transitional and international processes by which transnational activism becomes coordinated in more than one location, the analysis will test the applicability of Tarrow’s processes to explain the emergence and coordination between transnational activists in the US and South Africa. Third, provided that Tarrow argues that international political opportunities emerge due to changes in the international structure and are essential for explaining transnational activism, the analysis will examine how changes in international policies, programs, and funding affected the emergence and effectiveness of transnational AIDS treatment activism.

¹³⁵ Thompson, "Mapping Transnational Activism," 119.

Chapter 3: Case Study – Transnational AIDS Treatment Activism and ARV Price Reductions from 1996 to 2001

The aim of this chapter is to provide a chronological and empirical account of transnational AIDS treatment activism from 1996 to 2001. Specifically this chapter outlines the emergence, coordination, and impact of transnational activism in the US and South Africa under Health GAP and TAC, while accounting for changes in domestic and international policies to address the AIDS epidemic. This chapter is first divided into five sections by time-period and each section is further broken down into domestic-state and international levels of analysis. The last section of this chapter summarizes the case study and raises two points concerning the contribution of transnational AIDS treatment activism to the reduction of ARV prices between September 2000 and February 2001.

3.1 Setting the Stage (1996 – 1997)

The critical starting point for transnational AIDS treatment activism and the internationalization of AIDS was the discovery of HAART, as it was “the single most important shift in the history of the AIDS epidemic – transforming HIV infection from a ‘death sentence’ into a serious, but far more manageable chronic illness.”¹³⁶ Once the benefits of HAART, namely reducing morbidity, extending life, and preventing transmission of HIV became publicized and widespread in the developed world, individuals, NGOs, state governments, and international institutions slowly began to organize and mobilize to improve ARV access.¹³⁷ However, as the issue of ARV prices became more challenging, the pharmaceutical industry began to organize and mobilize to protect profits and IPRs within the US and South Africa. This period marks the beginning of when international institutions began to address AIDS from above, when individuals in developed and developing countries began to form strategies to access ARVs from below, and when the pharmaceutical industry began to pursue its own interests through domestic and international avenues.

3.1.1 State Level: South Africa Up Against the US and PMA

Since 1994, the newly elected African National Congress (ANC) government of South Africa faced many challenges undoing the legacy of apartheid. Specifically in terms of South Africa’s public health the government faced two major challenges between 1996 and 1997: the growing AIDS epidemic and an inadequate and distorted health system.¹³⁸ To address these issues a proposal was made for the

¹³⁶ Smith and Siplon, *Drugs into Bodies*, 44.

¹³⁷ Alan Berkman, "The Global AIDS Crisis: Human Rights, International Pharmaceutical Markets and Intellectual Property," *Connecticut Journal of International Law* 17, no. 2 (2002): 150.

¹³⁸ Robert Ostergard, "The Political Economy of the South Africa-United States Patent Dispute," *The Journal of World Intellectual Property* 2, no. 6 (1999): 875.

'Transformation of the Health System in South Africa' in April 1997.¹³⁹ The proposal outlined "the government's objectives in increasing access to health services, ensuring access to pharmaceuticals and strengthening disease prevention and health promotion."¹⁴⁰ Primarily, it was to correct "distortions of the apartheid years, where private sector health care was very expensive and the public sector health system charged prices in excess."¹⁴¹ Additionally, the proposal was made so that values enshrined in the new Constitution concerning equal rights to healthcare could be realized.¹⁴²

This proposal was immediately met by external pressure from the US government and was at the request of the PhRMA. According to CPT leader Jamie Love, Harvey Bale of PhRMA wrote a letter to the US Trade Representative (USTR) outlining a number of objections to the proposal.¹⁴³ Additionally, representatives from Johnson & Johnson, Bristol-Myers Squibb, and Merck met with Franklin Sonn, the South African Ambassador to the US, to discuss their objections.¹⁴⁴ Aldrige Cooper, Vice President of Johnson & Johnson, wrote to William Daley, the US Secretary of Commerce, claiming that the proposal would "have 'grave consequences for not only the US pharmaceutical industry, but all US direct investment in South Africa'."¹⁴⁵ As such, the US began to exert pressure on South Africa and urged the country not to proceed with the proposal in a bi-national commission between the US and South Africa, headed by Vice President Al Gore and South African Deputy President Thabo Mbeki.¹⁴⁶

Despite efforts and objections made by both the US government and pharmaceutical representatives, Mandela signed into law the Medicines and Related Substances Control Amendment Act (hereafter the Medicines Act) on 12 December 1997.¹⁴⁷ Depending on interpretation,¹⁴⁸ the Medicines Act declares that the Minister of Health (MOH) may determine when and how to obtain affordable medicines. Through parallel importation of patented medicines priced lower in other countries, or by

¹³⁹ Ibid., 877.

¹⁴⁰ Ibid.

¹⁴¹ Susan Cleary and Don Ross, "The 1998-2001 Legal Struggle between the South African Government and the International Pharmaceutical Industry: A Game-Theoretic Analysis," *Journal of Social, Political, and Economic Studies*. 27, no. 4 (2002): 450.

¹⁴² Jennifer Joni, "Access to Treatment for HIV/AIDS: A Human Rights Issue in the Developing World," *Connecticut Journal of International Law* 17, no. 2 (2002): 275-276.

¹⁴³ James Love, "Time-Line of Disputes Over Compulsory Licensing and Parallel Importation in South Africa," Consumer Project on Technology, accessed October 23, 2011, <http://www.cptech.org/ip/health/sa/sa-timeline.txt>.

¹⁴⁴ Ibid.

¹⁴⁵ Ibid.

¹⁴⁶ Smith and Siplon, *Drugs into Bodies*, 64.

¹⁴⁷ *Medicines and Related Substances Control Amendment Act, no. 90*, (December 12, 1997), accessed August 20, 2011, <http://www.info.gov.za/view/DownloadFileAction?id=70836>.

¹⁴⁸ See Cleary and Ross, "The 1998-2001 Legal Struggle," for an examination and interpretation of the Medicines Act and the South African government's intention of using it for ARVs.

issuing compulsory licenses to manufacture generic medicines domestically without permission from the original patent holder, both of which circumvent IPRs of the pharmaceutical industry and make medicines more affordable.¹⁴⁹

Shortly after the Medicines Act was signed, the PMA of South Africa, made up of 42 multinational pharmaceutical corporations, filed a lawsuit against the government of South Africa and Mandela.¹⁵⁰ According to the PMA's Notice of Motion, Section 10 of the Medicines Act was unconstitutional because the MOH was granted excessive power to overrule patent rights and, in so doing, violated South Africa's obligation to the WTO's agreement on TRIPS as "it only discriminates against patent owners in the pharmaceutical field."¹⁵¹ This lawsuit along with pressure from the US government marked the beginning of transnational political contention concerning ARV prices, IPRs and access to AIDS treatment, which would become a major mobilizing force and political opportunity for transnational AIDS treatment activism in the US and South Africa.

3.1.2 International Level: Hope with HAART and UNAIDS

During this period, the international political structure began to shift and focus on the issue of ARV treatment under the Joint United Nations Program on HIV/AIDS (UNAIDS) and Director Peter Piot, which officially began work in January 1996.¹⁵² It is important to note that international programs under the World Health Organization (WHO), such as the Global Program on AIDS (GPA), existed prior to the establishment of UNAIDS, which focused on dispelling myths surrounding the transmission of HIV, ending stigmatization, and improving state responses to the growing epidemic.¹⁵³ However, it became clear that a new international agency was required to gather accurate data on levels of infection, help countries understand the economic and political risks HIV and AIDS posed, assist in prevention, and above all coordinate efforts between international institutions and UN agencies.¹⁵⁴ As a result, UNAIDS was established after two years of intense planning¹⁵⁵ and was formed in collaboration with six cosponsors.¹⁵⁶ These included the WHO, United Nations Children's Fund (UNICEF), United Nations Development Program (UNDP), United Nations Population Fund

¹⁴⁹ Smith and Siplon, *Drugs into Bodies*, 52.

¹⁵⁰ Cleary and Ross, "The 1998-2001 Legal Struggle," 451.

¹⁵¹ *Ibid.*, 452.

¹⁵² Lindsay Knight, *UNAIDS: The First Ten Years 1996-2006* (Geneva: Joint United Nations Programme on HIV/AIDS, 2008), accessed September 1, 2011, http://data.unaids.org/pub/Report/2008/JC1579_First_10_years_en.pdf, 7.

¹⁵³ *Ibid.*, 15.

¹⁵⁴ *Ibid.*, 42.

¹⁵⁵ *Ibid.*, 27-46.

¹⁵⁶ Currently there are ten cosponsors, the International Labor Organization (ILO) joined in October 2001, the United Nations Office on Drugs and Crime (UNODC) joined in April 1999, the World Food Program (WFP) joined in October 2003, and lastly the Office of the United Nations High Commissioner for Refugees (UNHCR) joined in June 2004.

(UNPF), United Nations Educational, Scientific, and Cultural Organization (UNESCO), and the World Bank,¹⁵⁷ all of which had their own particular insights and reasons for addressing the epidemic.

In its early years, UNAIDS was attempting to establish itself as a credible and useful international agency and as an expert on HIV and AIDS, governmental policies and prevention to curb the epidemic. It encountered a number of hurdles, mainly securing funding and facilitating cooperation amongst cosponsors, but eventually it emerged as a leading institution working together with both state and non-state actors in the fight against AIDS and improving ARV access after three major international events and developments from 1996 to 1997.

The first event was the 11th International AIDS Conference in Vancouver, British Columbia, where HAART was officially announced in July 1996. Anne Winter, Communications Chief of UNAIDS, explained that it was difficult to persuade the organizers of the conference to allow Piot to speak, given that UNAIDS was new and relatively unknown.¹⁵⁸ However, Piot was eventually allowed to give a speech and announced that “bold action”¹⁵⁹ on all fronts were needed to address AIDS treatment inequality amongst developed and developing countries. This was the first time any international development official or UN agency publicly spoke about the need to change international and domestic policies to address the epidemic.¹⁶⁰ The 1996 conference “was undoubtedly a success for UNAIDS, and helped to position the organization as a key reference point.”¹⁶¹ Furthermore UNAIDS facilitated a shift in focus to poor developing countries and their inability to access new life-saving treatment.

The second important event and development, which UNAIDS used to further establish itself amongst businesses and political leaders and signifies a shift in the international political structure for addressing AIDS, was the annual meeting of the World Economic Forum in February 1997 in Davos, Switzerland. Sally Cowal, Director of External Relations for UNAIDS, attended the annual gathering one year earlier as part of a working group on healthcare, and made a strategic request for Mandela to attend alongside Piot in 1997. Cowal recounts, “we [would] never do better than this. There’s Peter on the podium with Mandela, and Mandela is saying to all these people, ‘if you don’t do something about AIDS, you can forget about development’.”¹⁶² This meeting led to the creation of the Global Business Council in October 1997 with both Mandela and Sir Richard Sykes, top executive of Glaxo

¹⁵⁷ Ibid., 30.

¹⁵⁸ Ibid., 60.

¹⁵⁹ Ibid.

¹⁶⁰ Ibid.

¹⁶¹ Ibid.

¹⁶² Ibid., 66.

Wellcome, in leadership positions.¹⁶³ This development highlights that UNAIDS was expanding horizontal and vertical ties and communication with both state and non-state actors very early on, which does reflect Tarrow's internationalization process.

The third important development for UNAIDS was the launching of a pilot program called the Drug Access Initiative (DAI) in November 1997.¹⁶⁴ This was the first collaboration between UNAIDS and brand-name pharmaceutical corporations, which resulted in ARV price reductions from \$10,000 to \$7,200 USD for one year of TCT for middle and low-income countries.¹⁶⁵ The program was divided into two parts. First, pharmaceutical corporations had to agree to reduce prices of ARVs and other medicines for AIDS-related illnesses. Second, UNAIDS was responsible for improving and adapting the health systems and infrastructure in four countries (Chile, Cote d'Ivoire, Uganda, Vietnam) "to ensure effective distribution," and demonstrate that individuals in resource poor developing countries could access and "use the medicines effectively."¹⁶⁶ The success of this program was crucial, as it demonstrated to the pharmaceutical industry that prices were the main barrier to access, not poor state infrastructure. Additionally, this marked the beginning of a long and heavily criticized relationship between UNAIDS and the pharmaceutical industry.

To summarize, the period from 1996 to 1997 laid the foundation for both transnational AIDS treatment activism and the reduction of ARV prices. The transnational political contention between the US and pharmaceutical industry against South Africa was a political opportunity and catalyst for mobilization in both countries, which will be discussed further in the next section and in Chapter 4. The discovery of HAART, the creation of UNAIDS, and the DAI marked the beginning of many changes in the international political structure, which also highlighted possible solutions to curb the epidemic with affordable ARVs.

3.2 Global Frustration and Mobilization (1998)

Domestically and internationally, 1998 was a year of disappointment, frustration, and setbacks in the fight against HIV and AIDS. The initial hope and positive outlook surrounding HAART gave way to the harsh reality of infection rates and conflict about ARV pricing and restrictive IPRs. There were many threats and hurdles for improving ARV treatment access, but these threats became catalysts for intense mobilization. This was the period that transnational activism for affordable ARVs and improving AIDS treatment access emerged simultaneously and independently from one another in the US, under Health GAP and Dr. Alan Berkman, and in South Africa under TAC and Zackie Achmat.

¹⁶³ Ibid., 68.

¹⁶⁴ Ibid., 69.

¹⁶⁵ Ibid.

¹⁶⁶ Ibid.

Additionally, this was the period that new international policies were proposed for protecting public health and access to medicines instead of protecting IPRs.

3.2.1 State Level: The Emergence of Health GAP and TAC

In 1998, tension between South Africa and the US was on the rise as members of Congress and PhRMA asked the USTR to place South Africa on its Special 301 Review. The Special 301 Review list, which was signed into law in 1988 under the Omnibus Trade and Competitiveness Act, is “designed to use the credible threat of unilateral retaliation” by the US government in order to “persuade trading partners to reform currently deficient intellectual property practices.”¹⁶⁷ PhRMA had argued that South Africa and its Medicines Act was a “test case” for countries that opposed US commitment to protect “all forms of American intellectual property, including pharmaceutical patents,”¹⁶⁸ and should be threatened with unilateral retaliation. As such, the USTR placed South Africa on its Special 301 Review on 01 May 1998, which officially marked a shift in US foreign policy towards South Africa.¹⁶⁹

This decision by the US created a political opportunity and target for mobilization and collective action for US transnational AIDS treatment activists. Mark Milano explains that he did not get involved in transnational AIDS treatment activism until this moment.

I was furious about the situation. This was different – this was an African government taking steps itself to provide HIV treatment in a very concrete way that would have a tremendous impact on the health of its people. And my government was preventing them from doing that. It was a situation ripe for action...I could tell my government: ‘Back off – you are not the cop of the world.’ I felt I could do something as an American that would help people around the world.¹⁷⁰

Milano’s statement highlights the importance of domestic political opportunities for transnational activism, as he explains that he needed clearer targets, in this case his own government, and more tangible strategies to help improve global ARV access. By the end of 1998, activists such as Milano slowly began to mobilize and Health GAP emerged with support from individuals and organizations, such as the AIDS Treatment Data Network, Search for a Cure, AIDS Treatment News, Essential Action, the International Gay and Lesbian Human Rights Commission (IGLHRC), Mobilization

¹⁶⁷ Judith Bello and Alan Holmer, " 'Special 301': Its Requirements, Implementation, and Significance," *Fordham International Law Journal* 13, no. 3 (1989): 259.

¹⁶⁸ Love, "Time-Line of Disputes."

¹⁶⁹ Office of United States Trade Representative, "USTR Announces Results of Special 301 Annual Review," (Washington D.C.: Office of United States Trade Representative, 1998), accessed November 11, 2011, http://keionline.org/sites/default/files/ustr_special301_1998.pdf. 1-33.

¹⁷⁰ Smith and Siplon, *Drugs into Bodies*, 65.

Against AIDS (MAA), and ACT UP and CPT.¹⁷¹ All of which provided technical, professional, and tactical expertise necessary to fight the US government and the pharmaceutical industry to improve AIDS treatment access around the world.

Meanwhile in South Africa, AIDS was rapidly spreading and for the first time on 01 December 1998, World AIDS Day, Mandela publicly spoke out and called for political commitment to fight the epidemic.¹⁷² Nine days later TAC officially launched on 10 December 1998, with its first campaign calling on the government to provide free ARVs for HIV-positive pregnant women to prevent mother-to-child transmission (MTCT).¹⁷³ In comparison to transnational AIDS treatment activism in the US, TAC formed because friends and family members were dying, nearly 600 people per day in South Africa, and little was being done to improve AIDS treatment access.¹⁷⁴ TAC emerged as a campaign for ARV access “by raising public awareness and understanding about issues surrounding the availability, affordability, and use of HIV treatments.”¹⁷⁵ It would soon grow from a handful of committed activists to a broad coalition of organizations in South Africa that would target governments and the pharmaceutical industry to improve ARV access in 1999 and eventually would reach out to other transnational AIDS treatment activists in 2000.

3.2.2 International Level: Attempting to Bridge the Global Treatment Gap

During 1998, UNAIDS began to broaden its support base by building partnerships, improving collaboration between cosponsors, reaching out to faith-based organizations, and facilitating horizontal or South-South cooperation of those working to alleviate HIV and AIDS.¹⁷⁶ Additionally, UNAIDS began work on the prevention of MTCT and improving epidemic statistics. Meanwhile, the World Health Assembly (WHA), the decision making body of the WHO, proposed the adoption of the Revised Drug Strategy (RDS), which sought to “ensure that public health rather than commercial interests have primacy in pharmaceutical and health policies.”¹⁷⁷ This indicated that the WHO was aware of the pharmaceutical industry, its lobbying power, and the threats being made to governments attempting to secure affordable and/or generic ARVs by the US. Furthermore, this policy proposal marked a shift in the international political structure in favor of improving AIDS treatment access.

¹⁷¹ Halbert, *Resisting Intellectual Property*, 99.

¹⁷² Knight, *UNAIDS: The First Ten Years*, 89.

¹⁷³ *Fighting for our Lives: The History of the Treatment Action Campaign 1998- 2010* (Cape Town: Treatment Action Campaign, 2010), accessed September 10, 2011, <http://www.tac.org.za/community/files/10yearbook/files/tac%2010%20year%20draft5.pdf>, 3.

¹⁷⁴ *Ibid.*, 4.

¹⁷⁵ Friedman and Mottiar, "Seeking the High Ground," 23; "About the Treatment Action Campaign," TAC, accessed September 20, 2011, <http://www.tac.org.za/community/about>.

¹⁷⁶ Knight, *UNAIDS: The First Ten Years*, 73-86.

¹⁷⁷ Love, "Time-Line of Disputes."

In July 1998, two years after the announcement of HAART, the 12th International AIDS Conference was held in Geneva, Switzerland under the theme 'Bridging the Gap'.¹⁷⁸ Thousands of activists, scientists, NGOs, PLWHA, and clinicians attended. UNAIDS presented the first official set of HIV surveillance data, supported by the WHO, Harvard University and the US Census Bureau.¹⁷⁹ Additionally, for the first time 180 country-specific HIV and AIDS fact sheets were published, which highlighted major differences in the epidemic between countries around the world. These positive developments in improving AIDS statistics were met by the negative reality that forecasts of infection rates were worse than expected.¹⁸⁰ Furthermore, Dr. Berkman, explained,

Sessions addressing AIDS poor and developing countries were largely ignored by scientists and clinicians from the wealthy nations...the pharmaceutical industry spent millions to entertain and influence doctors and researchers. The conference adjourned with no proposals on how to bridge the gap.¹⁸¹

Therefore, with no coordinated plan to 'bridge the gap' of ARV treatment access, individuals left the conference and devised their own plans within their own domestic and/or international working environments to improve AIDS treatment access.

To summarize, 1998 was a significant year for transnational AIDS treatment activism, as many domestic and international developments and events occurred that helped mobilize action and reflect Tarrow's theoretical arguments concerning domestic opportunities, targets, processes and the importance of international political opportunities. In the US, individuals mobilized domestically around and in response to US policy towards South Africa that interfered with AIDS treatment access. In South Africa, TAC emerged by pressuring the South African government to provide free ARVs for HIV-positive pregnant mothers, despite interference and pressure from external actors. Internationally, new policy proposals by the WHO were made to improve ARV access, but most importantly was the disappointment of the 12th International AIDS Conference, which helped mobilize transnational AIDS treatment activists.

3.3 Action for Africa (1999)

In 1999, transnational AIDS treatment activism was in full swing and major international policies and programs were being established. Domestically in the US, transnational activists under Health GAP were targeting and pressuring the US government to stop interfering with South Africa and its ability to access affordable ARVs. Domestically in South Africa, transnational activists under TAC were

¹⁷⁸ Halbert, *Resisting Intellectual Property*, 99.

¹⁷⁹ Knight, *UNAIDS: The First Ten Years*, 87.

¹⁸⁰ Ibid.

¹⁸¹ Berkman, "The Global AIDS Crisis," 149.

targeting both the South African and US government and the pharmaceutical industry. Internationally however, initiatives were made to specifically improve ARV access in Africa, just as the WHO revealed in its World Health Report that the number one cause of death in Africa was AIDS.¹⁸² By the end of 1999, the pharmaceutical industry began to discuss ARV price reductions.

3.3.1 State Level: Government and Industry Under Pressure

At the beginning of 1999, TAC targeted and pressured the South African government to provide free Zidovudine (AZT) to prevent MTCT. On 21 March, National Human Rights Day, TAC organized a 'Fast to Save Lives' across the country, by then TAC had "attracted significant media coverage and support" from "religious leaders, health professionals and even ANC officials."¹⁸³ Three days later MOH Dr. Nkosazana Dlamini-Zuma, "said that drug costs were the major obstacle to implementing a national program" for MTCT treatment.¹⁸⁴ Reacting to this TAC began to target specific pharmaceutical corporations, the PMA, and the US government for interfering with ARV access and prices, especially since South Africa was placed on the USTR Special 301 Review for a second year in a row.¹⁸⁵

In April 1999, TAC wrote a letter to Glaxo Wellcome, which asked "the company to respond to the national emergency posed by HIV/AIDS by selling AZT at cost price."¹⁸⁶ Additionally, TAC members protested outside Glaxo Wellcome headquarters in Midrand, but the company was unresponsive and refused "to say how much it cost to manufacture AZT."¹⁸⁷ According to TAC, it began targeting the US government in July when they protested, chanted, and sang outside US embassies and consulates around the country, "demanding an end to Washington's interference into South African affairs."¹⁸⁸ TAC's campaign against the PMA lawsuit began in September 1999 when activists "demonstrated outside the PMA offices in Johannesburg, calling for the unconditional withdrawal of suit."¹⁸⁹ In October 1999, TAC met with the Parliamentary Portfolio Committee for Health to explain why free AZT for HIV-positive pregnant women was essential, but suffered a huge setback when newly appointed President Thabo Mbeki "announced that AZT was a toxic drug,"¹⁹⁰ and explained that the government was concerned over the safety of ARVs not their affordability. This marked a turning point for TAC and for the epidemic in South Africa, but also for transnational

¹⁸² Knight, *UNAIDS: The First Ten Years*, 90.

¹⁸³ "History of TAC (1998-2001)," TAC, accessed September 22, 2011.

<http://www.tac.org.za/Documents/Other/tachist.pdf>, 9.

¹⁸⁴ Ibid.

¹⁸⁵ Love, "Time-Line of Disputes."

¹⁸⁶ "History of TAC (1998-2001)," 9.

¹⁸⁷ Ibid.

¹⁸⁸ Ibid., 12.

¹⁸⁹ Ibid.

¹⁹⁰ Ibid., 10.

AIDS treatment activism as HIV and AIDS scientists would mobilize and provide increased support against “Mbeki’s controversial stance on AIDS.”¹⁹¹ Thus, Mbeki provided new political opportunities for mobilization and threats to ARV access, but no major policy changes in favor of improving AIDS treatment access occurred in South Africa during 1999.

Meanwhile in the US, Health GAP members concerned about global AIDS treatment access worked in response to a number of political opportunities and threats throughout the year. First, Health GAP “focused [the] media’s attention to US trade policies and related government threats to withhold foreign development aid and restrict the ability of countries to trade with due to the patent infringement concerns of US-based drug companies.”¹⁹² In February 1999, activists drew positive attention in support of Representative Jesse Jackson Jr.’s Human Rights, Opportunity, Partnership, and Empowerment for Africa Act, which proposed to,

Cut off funding to any department or agency that sought through negotiation or otherwise, the revocation or revisions of any sub-Saharan African intellectual property or competition law or policy that is designed to promote access to pharmaceuticals or other medical technologies.¹⁹³

Additionally during this period, Sawyer of ACT UP explains that in early 1999 Health GAP “obtained a copy of what [they] called the smoking gun memo.”¹⁹⁴ The report outlined all efforts of the Clinton/Gore administration in support of the pharmaceutical industry in “their battle to prevent the South African government producing its own generic versions of AIDS and cancer medications.”¹⁹⁵ After accessing the memo activists shifted their efforts and specifically targeted Gore beginning in March 1999, by disrupting his appearance at a church picnic in New Hampshire.¹⁹⁶ Within one week members of Health GAP “received an invitation to meet with folks at the White House about [their] concerns.”¹⁹⁷ However, these discussions did not produce any significant policy changes and pressure on Gore increased throughout the year, especially in June 1999 when he officially launched his Presidential Campaign. Activists drove to Tennessee, New Hampshire, and New York, where they protested and picketed his speeches. Activists demanded ‘AIDS DRUGS FOR AFRICA!’ and plastered posters stating ‘Gore’s Greed Kills,’¹⁹⁸ which received “significant news coverage and

¹⁹¹ Ibid.

¹⁹² Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 95.

¹⁹³ Love, "Time-Line of Disputes."

¹⁹⁴ Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 98.

¹⁹⁵ Ibid.

¹⁹⁶ Ibid.

¹⁹⁷ Ibid., 99-100.

¹⁹⁸ "Health GAP Timeline: Direct Actions, Advocacy, and Response January 1999 - December 2000," Health GAP, accessed October 27, 2011, http://www.healthgap.org/hgap/Timeline_of_HGAP_1999-2000.html.

[threw] Gore into a frenzy.”¹⁹⁹ Additionally, a story was released by the Public Citizen campaign that exposed “Gore’s financial links with the pharmaceutical industry,” which “explained why he was taking a hard line with South Africa despite pressure from AIDS activists.”²⁰⁰

The second area or political opportunity activists became involved in was governmental working groups and negotiations about IPRs and access to medicines. Activists from Health Gap and members from CPT sought to educate Congressional leaders about the impact of IPRs on access to ARVs for developing countries. They convinced the Congressional Black Caucus (CBC) to write and question Gore about his US policies toward South Africa.²⁰¹ Furthermore, individuals attended a public debate on compulsory licensing and parallel imports led by the International Issues subcommittee of the Presidential Advisory Council on HIV/AIDS (PACHA) in June 1999.²⁰² Later on in the month, activists were invited to speak with Gore’s foreign policy spokesperson, Thomas Rosshirt, and the Director of White House Office of National AIDS Policy, Sandra Thurman.²⁰³ A few weeks later, the US government “announced that it would back off from its aggressive approach to South Africa, a move attributed directly to the pressure of activists on the administration in the campaign that year.”²⁰⁴ On 19 July 1999, Gore proposed the Leadership and Investment in Fighting an Epidemic initiative, with a heavy focus on AIDS in SSA.²⁰⁵

The last political opportunity and target of Health GAP was USTR Charlene Barshefsky leading up to the WTO’s trade negotiations in Seattle, Washington, where activists “were able to clarify that neo-liberal trade policies were doing little to help the poor or the sick around the world.”²⁰⁶ Jimmy McNulty, an ACT UP member, explains that everything “was timed and targeted carefully.”²⁰⁷ Members from Health GAP along with others from New York’s Fed Up Queers, “used a decoy to clear the entrance to Barshefsky’s office, where they locked themselves down and asked for a meeting with the trade rep. All this occurred within a federal building, on federal property close to the White House.”²⁰⁸ Benjamin Shepard explains, “that one hour was enough to force a change in federal trade policy,”²⁰⁹ which was officially announced by President Bill Clinton on 01 December 1999, World

¹⁹⁹ Benjamin Shepard, "Introductory Notes on the Trail from ACT UP to the WTO," in *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization*, eds. Benjamin Shepard and Ronald Hayduk (London: Verso, 2002), 14.

²⁰⁰ Halbert, *Resisting Intellectual Property*, 101.

²⁰¹ "Health GAP Timeline."

²⁰² Love, "Time-Line of Disputes."

²⁰³ Ibid.

²⁰⁴ Halbert, *Resisting Intellectual Property*, 101.

²⁰⁵ Copson, "CRS Issue Brief for Congress - AIDS in Africa," 12.

²⁰⁶ Halbert, *Resisting Intellectual Property*, 99.

²⁰⁷ Shepard, "Introductory Notes on the Trail from ACT UP to the WTO," 14.

²⁰⁸ Ibid.

²⁰⁹ Ibid., 15.

AIDS Day. Clinton explained, “the US will henceforth implement its healthcare and trade policies in a manner that ensures people in the poorest countries won’t go without medicine they so desperately need.”²¹⁰ Officially, Clinton’s Executive Order 13,155 on the ‘Access to HIV/AIDS Pharmaceuticals and Medical Technologies’ would be issued on 10 May 2000 and stated that the US would not use its Special 301 Review in “respect to any law or policy in beneficiary sub-Saharan African countries that promotes access to HIV/AIDS pharmaceuticals.”²¹¹ PhRMA responded to this by stating, “the President’s actions ‘set an undesirable and inappropriate precedent, by adopting a discriminatory approach to intellectual property laws, and focusing exclusively on pharmaceuticals’.”²¹² Nevertheless, this marked a major turning point for improving AIDS treatment access as US political support of its pharmaceutical industry weakened and shifted in support of global ARV treatment access by the end of 1999. This domestic development coincided with other international discussions and commitments by UNAIDS and the pharmaceutical industry to work on improving AIDS treatment access and reducing ARV prices.

3.3.2 International Level: Addressing AIDS and Development

In 1999, there were three important shifts in the international political structure for ARV access: the WHA approved the RDS, the World Bank’s new commitment to addressing AIDS, and the International Partnership against AIDS in Africa (IPAA) was established with support from UN General Secretary Kofi Annan. None of which however, directly involved or included the work of transnational AIDS treatment activism at this time.

Firstly, the newly approved RDS represented a shift as it urged states to safeguard public health, explore options under international agreements, and defend access to essential medicines.²¹³ Additionally, given that the RDS was partially a reaction to US pressure on South Africa, it required that the WHO “become involved in trade disputes involving pharmaceutical health policies.”²¹⁴ It is important here to note that the RDS was most likely approved because the US was not on the WHA Executive Board in 1999, as board membership rotates and the US is off every three years.²¹⁵ Nevertheless, this shift in international policy threatened the positions of the pharmaceutical industry and the US government towards South Africa and increased pressure to improve AIDS treatment access from above.

²¹⁰ Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 101.

²¹¹ Ibid.

²¹² Halbert, *Resisting Intellectual Property*, 102.

²¹³ Love, "Time-Line of Disputes."

²¹⁴ Ibid.

²¹⁵ Sell and Prakash, "Using Ideas Strategically," 163.

Secondly, in June 1999 the World Bank published a report entitled, 'Intensifying Action Against HIV/AIDS: Responding to a Development Crisis.'²¹⁶ This was the first time that HIV and AIDS was considered and written about as a serious threat to economic development. The report outlined 'four pillars' of which the World Bank's increased commitment to fighting HIV and AIDS would be based upon. The report stated the importance of advocacy by the institution, that it would frame AIDS as "a central development issue," and "increase and sustain an intensified response."²¹⁷ In addition, the report stated that the World Bank would increase resources and technical support for African countries, aid in prevention efforts, and help increase knowledge and data on HIV and AIDS to help governments design treatment programs.²¹⁸ However, the report noted that the World Bank's strategy would not include the provision of ARVs "because they are simply not affordable,"²¹⁹ but would work on ways to improve access to affordable ARV treatment. This report formed the basis of the World Bank's Multi-country AIDS Program (MAP) for Africa, which would be announced in 2000, and indicates that ARV prices were a major point of concern for the World Bank.

Thirdly, during 1999 the cosponsors of UNAIDS agreed to increase action towards SSA, specifically through the IPAA,²²⁰ which also inspired cosponsors to create new teams to focus specifically on AIDS in Africa within their own institutions. For example, the World Bank began to discuss increased funding for HIV and AIDS programs and created an AIDS Campaign Team.²²¹ UNICEF created a special task force and increased HIV and AIDS activities in most of its African country programs.²²² However, it would take the entire year to bring state and non-state actors together to officially establish IPAA, which was largely due to the personal involvement of Annan in the fight against AIDS in Africa. In June 1999, he gave his first public speech on the negative impacts of HIV and AIDS, which was "taking away Africa's future."²²³ Louise Frechette, the former UN Deputy Secretary-General, instrumental in Annan's involvement explained, "when the Secretary General [makes] AIDS a personal priority, it does reverberate around the world – most people don't have access to the Head of State, and it makes a huge difference."²²⁴ Therefore, when Annan brought together cosponsors, government leaders from the North and South, and top executives from the pharmaceutical industry to form a broad coalition of actors to fight AIDS in Africa at the beginning of

²¹⁶ "Intensifying Action Against HIV/AIDS in Africa: Responding to a Development Crisis," (The World Bank, 1999), accessed November 3, 2011, <http://siteresources.worldbank.org/AFRICAEXT/Resources/aidstrat.pdf>, 1-78.

²¹⁷ Ibid., 5.

²¹⁸ Ibid.

²¹⁹ Ibid., 49.

²²⁰ Knight, *UNAIDS: The First Ten Years*, 92.

²²¹ Ibid., 117-118.

²²² Ibid., 93.

²²³ Ibid., 94.

²²⁴ Ibid.

December 1999, the IPAA was firmly established. It was after this meeting, according to investigative journalist Barton Gellman, the pharmaceutical industry began to rethink ARV price reductions and take up Annan's "call to action."²²⁵

In conclusion, 1999 was a very important year for improving AIDS treatment access and the reduction of ARV prices. Transnational AIDS treatment activism appears to have had its most visible impact in the US under Health GAP as US foreign policy towards South Africa and its stance on IPRs and ARV access shifted. Furthermore, new international policies and commitments by international institutions for addressing AIDS treatment access emerged, changing the international political structure in alignment with the goals of transnational AIDS treatment activists. However, it is important to note that these new policy shifts occurred autonomously from the work of transnational activists. Nevertheless, shifts in state policies from below, which transnational AIDS treatment activism had a limited impact on, and shifts in international policies from above, began to place pressure on the pharmaceutical industry to reduce ARV prices.

3.4 Transnational Activists Working Together (January 2000 – August 2000)

As 1999 ended, after intense activist efforts in the US and South Africa and shifting domestic and international policies for addressing AIDS, 2000 was the year that ARV prices were negotiated and began to reduce. Meanwhile, a number of important developments took place while ARV price negotiations were held and eventually announced on 11 May 2000, with the Accelerating Access Initiative (AAI) that reduced prices by 90%.²²⁶ Transnational AIDS treatment activists in the US and South Africa began working together against pharmaceutical giant Pfizer and formed a transnational event coalition around the 13th International AIDS Conference in Durban, South Africa. This section examines these developments leading up to further ARV price reductions between September 2000 and February 2001.

3.4.1 State Level: Coordinated Efforts of TAC and Health GAP

In 2000, transnational activists in the US and South Africa began to coordinate their efforts in two important ways: pressuring Pfizer to reduce the price of fluconazole, a patented medicine to treat opportunistic infections,²²⁷ and organizing a global protest against ARV treatment inequality leading up to the 13th International AIDS Conference.

²²⁵ Barton Gellman, "A Turning Point that Left Millions Behind ; Drug Discounts Benefit Few while Protecting Pharmaceutical Companies' Profits," *The Washington Post*, December 28, 2000, accessed November 7, 2011, <http://www.lexisnexis.com/hottopics/lnacademic/?verb=sr&csi=8075&sr=HLEAD%28A+turning+point+that+left+millions+behind%29%2BAND%2BDATE%2BIS%2B2000>.

²²⁶ Collins, "World's AIDS Fighters Plot New Strategy."

²²⁷ Geffen, *Debunking Delusions: The Inside Story of the Treatment Action Campaign*, 50.

First, according to US activists, TAC requested that ACT UP's New York branch and Health GAP target Pfizer, as part of their Christopher Moraka Defiance Campaign in March 2000. At the time, Christopher Moraka, one of TAC's members, was suffering from thrush and eventually died on 24 July 2000 without receiving treatment.²²⁸ In the US, activists made an "unannounced and confrontational visit to [Pfizer's] CEO, as well as made threats to disrupt a Pfizer shareholder meeting."²²⁹ Eventually, Pfizer sent a letter to Mark Heywood of TAC, signed by Country Director George Flouty, stating that the company would provide free fluconazole to medical specialists in South Africa to treat those who could not afford the treatment.²³⁰ In South Africa, TAC members approached Pfizer and planned to increase competition by importing Biozole, a generic brand from Thailand. Achmat and Jack Lewis travelled to Thailand later in the year and purchased 3,000 capsules of Biozole. They crossed back into South Africa with the medicines and held a press conference to publicly announce what they had done.²³¹ The Democratic Alliance (DA) condemned TAC and later the Department of Health confiscated all the medicine. But news spread that 3,000 capsules of Biozole cost the same amount as 60 capsules of Pfizer's fluconazole, which increased public support for TAC and eventually led to the Medicines Control Council (MCC) allowing the importation of Biozole.²³²

Second, individuals from TAC, Health GAP and Médecins Sans Frontières (MSF), worked together to coordinate a mass march in July 2000 during the 13th International AIDS Conference in South Africa, which was the first time this conference was held in a Southern and developing country. In July, TAC and MSF worked together educating people about HIV and AIDS, treatment options, and ARV prices.²³³ Over 4,000 participated in the Global March for Treatment, even Winnie Madikizela-Mandela from the ANC and Patricia De Lille from the Pan Africanist Congress (PAC) joined.²³⁴ Reporters noted that activists were explaining to the media, "that the life-extending medicines are unaffordable in the Third World because western pharmaceutical companies want to exploit their patent rights."²³⁵ This global march and conference was the first major coordinated transnational event of transnational AIDS treatment activists from various countries and international NGOs. Furthermore, it was at this event that the demand for generic ARVs to create competition and drive patented ARV prices down further, was strongly articulated.

²²⁸ Ibid.

²²⁹ "Health GAP Timeline."

²³⁰ Gellman, "A Turning Point."

²³¹ Geffen, *Debunking Delusions: The Inside Story of the Treatment Action Campaign*, 50.

²³² Ibid.

²³³ Ibid., 55.

²³⁴ "History of TAC (1998-2001)," 10.

²³⁵ Collins, "World's AIDS Fighters Plot New Strategy."

3.4.2 International Level: Accelerating ARV Access

During 2000, a number of international events and policy shifts occurred for the improvement of AIDS treatment access. First, on 10 January 2000, the United Nations Security Council (UNSC) discussed AIDS in Africa as a human security issue and barrier to development for the continent. “Gore chaired the meeting and noted that ‘the activists were right,’ there is more the government must do regarding AIDS in Africa.”²³⁶ Piot explained that the UNSC discussions “opened so many doors, top leaders told me, [since] it was debated in the Security Council it must be a serious problem. Ridiculous, but I got that sort of response.”²³⁷ This further solidified international political commitment to reducing barriers to AIDS treatment access and placed added pressure on the pharmaceutical industry to reduce ARV prices given global concern and support for curbing the epidemic.

Second, as briefly mentioned above, UNAIDS and the pharmaceutical industry began negotiations about price reductions and improving the 1997 DAI program.²³⁸ Gro Harlem Brundtland, Director General of the WHO became involved in these discussions and publically proposed a global and holistic approach to effectively combat HIV and AIDS.²³⁹ Brundtland explained that US support was vital and cooperation of pharmaceutical companies, governments of developing countries, and financial institutions were all necessary to make essential medicines and technology readily available to those suffering from HIV and AIDS.²⁴⁰ After months of negotiations on 11 May 2000, one day after Clinton’s Executive Order was signed, five pharmaceutical corporations (Bristol Myers Squib, Merck & Co., Glaxo Wellcome, Boehringer Ingelheim and F. Hoffman-LaRoche) and UNAIDS announced the AAI. “They offered to sell their AIDS medicines to sub-Saharan Africa for prices well below what Americans pay for them,”²⁴¹ and prices were cut by almost 90%.²⁴² Many transnational AIDS treatment activists argued that these price reductions “were not enough to make the drugs affordable for many African governments, which spend under \$5 a year per person on health.”²⁴³ In response, transnational AIDS treatment activists demanded further price cuts and for governments to be allowed to access generic versions to drive ARV prices down further by adding competition to the supply.

²³⁶ “Health GAP Timeline.”

²³⁷ Knight, *UNAIDS: The First Ten Years 1996-2006*, 106.

²³⁸ *Ibid.*, 121-124.

²³⁹ Gro Harlem Brundtland, “Remarks at the WHO/Public Interest NGO Pharmaceuticals Roundtable, 3rd Meeting,” May 1, 2000, accessed August 17, 2011, http://www.who.int/director-general/speeches/2000/20000501_ngo.html.

²⁴⁰ Schug, “Promoting Access to HIV/AIDS Pharmaceuticals,” 232.

²⁴¹ Collins, “World’s AIDS Fighters Plot New Strategy.”

²⁴² *Ibid.*

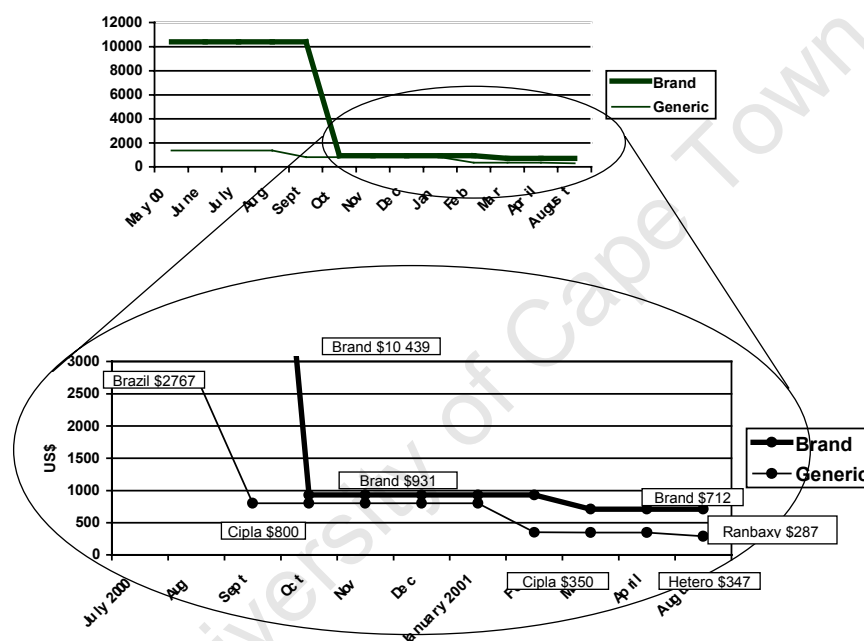
²⁴³ *Ibid.*

To summarize, during the first half of 2000 a number of events and policy commitments occurred at the same time. In May UNAIDS announced its AAI, which led to ARV price reductions by 90%.²⁴⁴ However, this reduction was inadequate for transnational AIDS treatment activists, which then began working together to show a united front against the AAI and demanded the generic supply of ARVs, which would be negotiated in the second half of 2000.

3.5 ARV Price Reductions (September 2000 – February 2001)

According to MSF and illustrated in the graph in Figure 8, global ARV prices reduced from \$10,000 to \$1,000 USD in September 2000 and further reduced to less than \$350 USD in February 2001.

Figure 8. ARV Price Reductions 2000-2001²⁴⁵



The first reduction of prices to less than \$1,000 USD is explained by the 90% price cuts of the AAI. However, the second reduction of prices too less than \$350 USD occurred after intense transnational AIDS treatment activist mobilization in July and after a number of new announcements and negotiations in September 2000. First, the World Bank's board approved and announced the MAP for Africa initiative in September 2000 with a budget of \$500 million USD.²⁴⁶ This was a huge step forward in improving AIDS treatment access and with this additional and increased support from the World Bank to the fight against AIDS, the pharmaceutical industry was under more pressure from above to reduce ARV prices.²⁴⁷ Furthermore, it is worth noting that one of MAP's key characteristics, which would create future political opportunities for activist organizations, was its direct funding of

²⁴⁴ Perez-Cases et al., "Accessing ARVs," 3.

²⁴⁵ Ibid.

²⁴⁶ Knight, *UNAIDS: The First Ten Years 1996-2006*, 117-118.

²⁴⁷ Sell and Prakash, "Using Ideas Strategically," 162.

NGOs, community organizations and the private sector.²⁴⁸ Additionally, during September, India's Chemical, Industrial and Pharmaceutical Laboratories (CIPLA) and MSF, at the urging of Love and CPT, began meeting to discuss the generic supply of ARVs with the pharmaceutical industry.²⁴⁹ It is important to explain here that the reason an Indian pharmaceutical company could produce generic ARVs as this time was because India did not have to be TRIPS compliant until 2005, and therefore was not in violation of the WTO.²⁵⁰ According to Gellman, CIPLA's top executive Yusuf Hamied offered, "to sell generic versions of patented medicines at 5 to 10 cents on the dollar, as a global public service."²⁵¹ Additionally, Hamied stated, "I represent the needs and the aspirations of the Third World...It is up to you, the international community, to grasp this opportunity...to alleviate the suffering of millions of our fellow men who are afflicted with HIV and AIDS."²⁵² Officially, CIPLA and MSF announced their agreement to produce and provide one year of TCT of generic ARVs for \$350 USD in February 2001, thus reducing the global average ARV prices to less than \$350 USD as illustrated in the graph on the previous page.²⁵³

3.6 Case Summary

By tracing both domestic and international political developments this chapter has provided a chronological and comprehensive account of transnational AIDS treatment activism during the period from 1996 to 2001. Firstly, this chapter illustrated that domestically in the US, the government was supporting the pharmaceutical industry and exerting pressure on South Africa, which created an opportunity for members of Health GAP, ACT UP, and CPT to exploit. By the end of 1999, after numerous protests, direct actions, and meetings with transnational AIDS treatment activists concerning IPRs and ARV access, US foreign policy towards South Africa shifted in support of improving ARV treatment access specifically for SSA. Secondly, this chapter explained that domestically in South Africa, the government and TAC were responding to international pressure from the US and pharmaceutical industry. However, once Mbeki began questioning the safety of ARVs both internal and external political threats and opportunities arose, which TAC acted upon and produced further international support. Ultimately, no major policy shifts occurred after TAC took up action during this period. Thirdly, this chapter highlighted that internationally the creation of UNAIDS, new policies of the WHO, and an increased interest by the World Bank for addressing AIDS, especially for SSA, represented an international structural shift, or the internationalization of AIDS. These developments, which have since created new avenues and opportunities for transnational

²⁴⁸ Hakan Seckinelgin, *International Politics of HIV/AIDS: Global Disease-Local Pain* (New York: Routledge, 2008), 60.

²⁴⁹ Sell and Prakash, "Using Ideas Strategically," 162.

²⁵⁰ Gellman, "A Turning Point."

²⁵¹ Ibid.

²⁵² Ibid.

²⁵³ Ibid.

activists, occurred simultaneously with the emergence and action of transnational AIDS treatment activists in the US and South Africa and were not directly capitalized upon by transnational activists to be most effective and reduce ARV prices.

From this domestic and international account, two important points emerge concerning the causal relationship between transnational AIDS treatment activism and the reduction of ARV prices. Firstly, transnational AIDS treatment activism appears to have had most impact on changing US policy to support the improvement of ARV access for SSA, weakening support for the pharmaceutical industry. Furthermore, the CIPLA and MSF deal that resulted in ARV price reductions to \$350 USD did occur after coordinated efforts and global protests, but there is little evidence to suggest that this was due to transnational AIDS treatment activism, especially because it involved intense and undisclosed negotiations, and because it was in CIPLA's own interest to increase its market share.²⁵⁴ Nevertheless, these negotiations were at the urging of specific transnational AIDS treatment activism, therefore, transnational AIDS treatment activism had a limited causal impact on global ARV price reductions. Secondly, major international policy shifts for addressing AIDS and improving access to ARV treatment by the WHO, UNAIDS, and the World Bank, occurred simultaneously and independently from transnational activism. Thus, it appears that the limited causal relationship between transnational AIDS treatment activism and the reduction of ARV prices is best understood in combination with broad structural changes and lends itself to a conjunctural explanation. From this case study, Chapter 4 will apply and test Tarrow's theoretical arguments to explain how transnational AIDS treatment activism in the US and South Africa emerged and how it had a limited contribution to the reduction of ARV prices.

²⁵⁴ Sell and Prakash, "Using Ideas Strategically," 162.

Chapter 4: Analysis – Tarrow and Transnational AIDS Treatment Activism

The aim of this chapter is to provide a comprehensive analysis of the emergence of transnational AIDS treatment activism and its limited contribution to the reduction of ARV prices from 1996 to 2001. It employs Tarrow's theoretical arguments, outlined in Chapter 2, to examine the causal relationship more closely and tests each of his arguments to add to the theoretical discussion of transnational activism. Provided that transnational activism, according to Tarrow, is very much a product of domestic state politics, this chapter first analyzes domestic origins, political opportunities and threats, and processes, through which transnational AIDS treatment activism emerged in the US and South Africa. Secondly, this chapter employs Tarrow's transitional and international processes by which claims and localized action extend across state borders; to examine how transnational AIDS treatment activism became coordinated globally. Thirdly, this chapter examines international political opportunities and threats that arose from changes in the international structure by the World Bank, the WTO, the WHO, and UNAIDS. Lastly, however, this research proposes that changes in the international political structure did not merely create new opportunities for transnational activism, for changes were in alignment with the goals of transnational AIDS treatment activism and provided separate and additional pressure on the pharmaceutical industry to improve ARV treatment access. This dissertation proposes that the reduction of ARV prices from September 2000 to February 2001 was due to a conjuncture of originally separate developments and processes, which emerged by way of transnational AIDS treatment activism from below and by way of international institutions from above.

4.1 Domestic Analysis of Transnational AIDS Treatment Activism

This section examines the domestic origins of transnational AIDS treatment activism in the US and South Africa, both of which are based in a history of gay rights and early AIDS activism from the 1970s to the 1990s and influenced strategies, resources, organizational forms, and framing efforts observed between 1996 and 2001. This section briefly discusses domestic political opportunities and threats, which affected how Health GAP and TAC emerged and designed strategies to pressure state governments and the pharmaceutical industry to improve AIDS treatment access by reducing ARV prices since 1998. In addition, this section illustrates that transnational AIDS treatment activism in the US and South Africa emerged simultaneously from similar rights-based origins, but within different contexts, in response to different political opportunities and threats, and through different domestic processes. Therefore, transnational AIDS treatment activism should not be assumed or examined as one entity, especially during this period, but should be understood as two separate and uncoordinated

processes that emerged in different locations and contributed differently to the reduction of ARV prices from below.

4.1.1 United States

Transnational AIDS treatment activism under Health GAP stems from a long history of gay rights and early AIDS activism from the 1970s and early 1980s in the US, all of which fought institutionalized discrimination.²⁵⁵ Initially, mobilization and organization surrounding AIDS began in gay communities in New York and San Francisco, as groups formed to discuss HIV and AIDS and to provide emergency care and psychological support to individuals and families with dying loved ones.²⁵⁶ These groups “developed quickly by drawing upon already existing friendships, neighborhoods, professional and other networks that predated the epidemic.”²⁵⁷ Additionally, these groups began to educate themselves about HIV and AIDS, symptoms, prevention, and began pooling resources in response to the growing epidemic.²⁵⁸ Eventually in 1983, the People with AIDS (PWA) self-empowerment movement began which fought against AIDS-related discrimination and sought to reclaim individual human rights. The movement produced seventeen principles, known as the ‘The Denver Principles,’ which outlined recommendations for the public and for PWA to deal with HIV and AIDS. Primarily, the principles called for increased support in the AIDS struggle, and urged PWA to become involved in all levels of AIDS decision-making.²⁵⁹ Moreover, the principles stated that individuals with AIDS were not victims or patients, that PWA deserved privacy, explanations of all medical procedures, and to “LIVE and die in dignity.”²⁶⁰ Additionally, the PWA movement challenged attitudes of doctors, nurses, and the government towards AIDS in hopes to work swiftly for a cure for the growing epidemic.²⁶¹

The PWA movement helped lay the foundation for ACT UP, an important member organization of Health GAP, which officially launched in 1987 under Larry Kramer, “the single most influential figure in AIDS politics of the 1980s.”²⁶² ACT UP is regarded as a very successful activist organization that has arguably influenced most activism in the US.²⁶³ However, ACT UP is primarily recognized for pressuring the pharmaceutical industry and the US government to improve HIV and

²⁵⁵ Mbali, "The Treatment Action Campaign," 7.

²⁵⁶ Smith and Siplon, *Drugs into Bodies*, 13-15.

²⁵⁷ *Ibid.*, 15.

²⁵⁸ *Ibid.*

²⁵⁹ Michael Callan and Dan Turner, "A History of the People with AIDS Self-Empowerment Movement," *Body Positive Magazine*, December 1997, accessed November 14, 2011, <http://www.thebody.com/content/art31074.html>.

²⁶⁰ *Ibid.*

²⁶¹ *Ibid.*

²⁶² Smith and Siplon, *Drugs into Bodies*, 18-19.

²⁶³ See Shepard and Hayduk, *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization*.

AIDS treatment. ACT UP is credited for improving AIDS research, increasing scientific journal publications, improving drug trial policies, and accelerating the drug approval process.²⁶⁴ Additionally, ACT UP was involved in promoting the Ryan White Comprehensive AIDS Resources Emergency Act, which included funding for AIDS Drug Assistance Programs (ADAPs) that improved access to medicine and made ARVs widely available for those with HIV and AIDS in the US since 1996.²⁶⁵ Most importantly however, ACT UP emerged as a “group of individuals committed to direct action to end the AIDS crisis,”²⁶⁶ with two main strategic forms: *zaps* and *actions*. Zaps addressed immediate issues and specific targets, and involved invading offices, sending postcards, letters and faxes, distributing factsheets, and picketing.²⁶⁷ Actions, on the other hand, increased public awareness, exposed inaction or information about targets, and made demands for change. Furthermore, actions involved a large number of activists and organizations, a planning and approval process, and extensive promotion.²⁶⁸ Both zaps and actions were helped by the fact that “so many members of ACT UP had backgrounds in visual arts, graphic design, theater, and other performing arts as well as in writing, editing, and advertising.”²⁶⁹ Maxine Wolfe of ACT UP explains that they “focused on what would stand out, what would show up,” they “did not just picket around the front of a building, which [was] totally boring; [they] broke into the building”²⁷⁰ to draw as much media attention as possible.

Some of ACT UP’s early zaps and actions, which were repeated in 1999, included invading pharmaceutical corporate offices, protesting pharmaceutical pricing policies in the New York Stock Exchange,²⁷¹ and scattering the ashes of those who died from AIDS on the White House lawn.²⁷² Additionally, ACT UP members became experts about HIV and AIDS, ARVs, and barriers to accessing treatment in order to debate with government, scientific, and pharmaceutical leaders, as well as answer media questions.²⁷³ Furthermore, ACT UP always prepared “fact sheets that included a list of demands that summarized workable solutions.”²⁷⁴ Despite ACT UP’s early success, between 1992 and 1996, AIDS activism declined as no treatment had been discovered.²⁷⁵ However, once

²⁶⁴ Steven Epstein, *Impure Science: AIDS, Activism, and the Politics of Knowledge* (Berkeley: University of California Press, 1996), 337-353.

²⁶⁵ Smith and Siplon, *Drugs into Bodies*, 44-45.

²⁶⁶ "Actions and Zaps," ACT UP New York, accessed November 14, 2011, <http://www.actupny.org/documents/newmem2.html>.

²⁶⁷ Ibid.

²⁶⁸ Ibid.

²⁶⁹ Smith and Siplon, *Drugs into Bodies*, 30.

²⁷⁰ Ibid.

²⁷¹ Ibid., 32.

²⁷² "ACT UP Capsule History 1992," ACT UP New York, accessed November 14, 2011, <http://www.actupny.org/documents/cron-92.html>.

²⁷³ Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 91.

²⁷⁴ Ibid.

²⁷⁵ Smith and Siplon, *Drugs into Bodies*, 36-37.

HAART was officially announced in 1996, AIDS activism revived and began to focus on global ARV access.

Health GAP emerged at the end of 1998 and included numerous organizations, which were mentioned in Chapter 3, most importantly however, was the new partnership formed with CPT, which Sawyer describes as “the missing link in terms of information to challenge drug company pricing and patents.”²⁷⁶ CPT provided crucial information and expertise concerning domestic and international IPR policies and provided avenues into the US government, as Presidential candidate, Ralph Nader, was a founding member of the organization.²⁷⁷ CPT was essential for identifying domestic political opportunities and threats, which were then capitalized upon by other ACT UP members with a history of effective direct action strategies.

Throughout 1999, Health GAP reacted to domestic opportunities and threats in order to criticize the pharmaceutical industry and the US government, draw media attention, increase support, and provide information to the public to improve AIDS treatment access and reduce ARV prices. First, Health GAP acted upon political opportunities presented by Gore’s Presidential campaign. Sawyer explains that attacking Gore on his campaign trail,

Had a major impact, due as much to the lack of one story, as to the presence of another. The media were already bored by a lackluster campaign by two candidates perceived to be equally dull. They were ripe for something interesting to write about, and we gave them the hook they were looking for.²⁷⁸

Gore’s involvement in trade discussions with South Africa were a threat to ARV treatment access, but this threat, accompanied by his Presidential campaign, provided an opportunity for activists to expose and highlight how the US was interfering with ARV treatment access. Secondly, Health GAP acted upon the political opportunity posed by the USTR’s Special 301 Review in relation to IPR controversy, the WTO, and the Seattle trade negotiations held in November 1999. The direct action of Health GAP and ACT UP members against Barshefsky, discussed in Chapter 3, was strategically planned and acted out in order to gain more support from other activist organizations planning to protest the WTO. Thirdly, activists in the US found an opportunity in the PMA lawsuit against South Africa to expose and inform the public about how the pharmaceutical industry was blocking access to affordable ARVs, which in turn created “public outrage,” and “led to widespread demands that the US government stop putting drug company profits before human lives.”²⁷⁹

²⁷⁶ Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 98.

²⁷⁷ Sell and Prakash, "Using Ideas Strategically," 162.

²⁷⁸ Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 99.

²⁷⁹ *Ibid.*, 95.

The direct actions, strategies, and framing efforts of Health GAP, ACT UP, and CPT in response to domestic political opportunities within the US, primarily reflect Tarrow's global issue framing process, which refers to the ways in which domestic claims and localized action use international symbols and information to frame domestic contentious politics in terms of a larger and international struggle.²⁸⁰ Furthermore, transnational AIDS treatment activism in the US does not reflect Tarrow's second domestic process of internalization, because activists were not responding to external or international pressure that they were directly affected by. Global issue framing was most evident in slogans used by Health GAP, such as 'AIDS Drugs for Africa!' and 'Stop Medical Apartheid!' that were chanted and plastered on posters to disrupt Gore's Presidential campaign.²⁸¹

According to Tarrow, global issue framing arises in two ways, through structural equivalence or global thinking.²⁸² Structural equivalence implies that individuals in more than one country experience similar external and structural pressure and in making domestic claims use this shared experience to create a feeling of solidarity and hope. However, in the case of AIDS treatment activism in the US, activists did not experience the same external and structural pressure from foreign governments, the WTO, or the pharmaceutical industry as activists experienced in South Africa. Yet, members of ACT UP Philadelphia, mostly African-American, joined the fight for global ARV access due to "feelings of ethnic solidarity as well as bonds of personal experience."²⁸³ Paul Davis explains,

One reason the global AIDS struggle has resonated so deeply with our members is because of the shared experience. Our members know what it means to have shoddy health care. Our members know what it means not to be able to afford life-saving drugs.²⁸⁴

Even though activists in the US did not experience the same external and structural pressure as in South Africa, some felt a common bond, and explained it was due to a feeling of shared identity and common experience. This helped garner support and frame domestic struggles against the US government in terms of a global and ongoing struggle for ARV access. However, it is also worth mentioning that this 'shared experience' might also be a reflection of what some scholars refer to as "survivor guilt,"²⁸⁵ that emerged and "led to anger on the part of some activists as they came to

²⁸⁰ Tarrow, *The New Transnational Activism*, 32.

²⁸¹ Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 99.

²⁸² Tarrow, *The New Transnational Activism*, 60.

²⁸³ Smith and Siplon, *Drugs into Bodies*, 61.

²⁸⁴ Ibid.

²⁸⁵ Ibid., 62.

understand that wealthy governments and pharmaceutical companies were²⁸⁶ standing in the way of improving ARV treatment access.

Secondly, Tarrow cautiously explains that global issue framing may arise from global thinking, which refers to individuals that connect and identify themselves beyond their immediate domestic-state surroundings to other areas from where they are primarily located.²⁸⁷ There is evidence to support this process in the US, as Sawyer, a key player in Health GAP and founder of ACT UP stated,

I am an international treatment activist today because I am angry that my friends in the developing world are dying due to lack of access to the treatments and health care I rely on to thrive with a fatal disease...I am an international treatment activist because I was born with a sense of humanity that tells me that every life matters and deserves to be fought for.²⁸⁸

Sawyer's statement provides evidence that the emergence of domestic claims and framing global access to ARVs in relation to Africa reflected some form of global thinking. Therefore, strategies, tactics, targets and frames that Health GAP used throughout 1999 to domestically pressure the US government and pharmaceutical industry do reflect Tarrow's global issue framing process.

In summary, the domestic origins of Health GAP are rooted in the strategic efforts and impact of early AIDS activism and ACT UP from 1987 to the early 1990s that fought against discrimination and the inaction of the US government and pharmaceutical industry in response to the epidemic. In 1998, Health GAP emerged by way of pre-existing networks, relationships, and strategies and in response to new political opportunities. It contributed to ARV price reductions through the process of global issue framing, which pressured the US government to change its stance and policies on ARV prices and protection of IPRs towards SSA, which Clinton announced on 01 December 1999.

4.1.2 South Africa

As in the US, transnational AIDS treatment activism in South Africa stems from a long history of gay rights and early AIDS activism, which took place during and formed part of the anti-apartheid struggle in the 1980s and early 1990s.²⁸⁹ Some refer to this as TAC's "struggle heritage,"²⁹⁰ which has influenced agendas, strategies, resources, and framing efforts of activists fighting for affordable ARVs

²⁸⁶ Ibid., 152.

²⁸⁷ Tarrow, *The New Transnational Activism*, 72.

²⁸⁸ Sawyer, "An ACT UP Founder 'Acts Up' for Africa's Access to AIDS," 88.

²⁸⁹ Friedman and Mottiar, "A Rewarding Engagement?," 511-565.; Grebe, "The Treatment Action Campaign's Struggle for AIDS Treatment," 849-868.; Steven Robins, "'Long Live Zackie, Long Live': AIDS Activism, Science and Citizenship After Apartheid," *Journal of Southern African Studies* 30, no. 3 (Sep. 2004): 651-672.

²⁹⁰ Grebe, "The Treatment Action Campaign's Struggle for AIDS Treatment," 852.

in South Africa. Initially, gay rights and early AIDS activists focused on fighting discrimination within the liberation movement of the ANC. Individuals, such as Simon Nkoli, and organizations, such as the Rand Gay Organization, the Gay and Lesbian Organization of the Witwatersrand (GLOW), and the Organization of Gay and Lesbian Activists (OLGA), and the National Coalition for Gay and Lesbian Equality (NGCLE), under the leadership of Achmat, began to mobilize within the ANC to ensure non-discrimination based on sexual orientation would be upheld. Nkoli “thought that gay activists had to ‘stand up and fight’ for their rights in the liberation movement, even if it meant courting ‘unpopularity’ with other anti-apartheid comrades.”²⁹¹ Ultimately, working within the ANC and in line with its own non-discrimination racial policies and human rights discourse, OLGA and GLOW lobbied for support and were able to “get the outlawing of discrimination on the grounds of sexual orientation included in the ANC’s Bill of Rights,”²⁹² which forms the basis of South Africa’s new democratic Constitution. By successfully upholding human rights, gay rights activists were able to turn their attention and efforts to the HIV and AIDS epidemic, and “to wider struggles such as TAC’s.”²⁹³

Mobilization against AIDS, during South Africa’s democratic transition, took many shapes and forms. Primarily, AIDS activists were trying to fight AIDS-related discrimination and did so by encouraging HIV-positive individuals to be open about their status in order to fight for human rights and non-discrimination enshrined in the new Constitution. In 1992, Edwin Cameron founded the AIDS Consortium Project and helped draft the Charter of Rights for People with AIDS and HIV, which sought to inform policy and the public about AIDS and discrimination.²⁹⁴ Additionally, in October 1992, the National AIDS Convention of South Africa (NACOSA) held its first conference on AIDS in South Africa, under the theme ‘South Africa United Against AIDS’ and incorporated anti-apartheid organizations, civil society organizations, businesses and governments.²⁹⁵ However, these early commitments and mobilization for AIDS soon faded behind more pressing political and economic issues, especially as there was no unifying voice fighting against AIDS or ARV treatment available at the time. It was not until the end of 1998, with the death of Nkoli due to the unavailability of ARVs, that TAC emerged and began demanding access to affordable ARVs in South Africa.

The influence of gay rights and early AIDS activism during the anti-apartheid struggle is embodied by TAC in a number of ways and was apparent in their direct action strategies in 1999. First, TAC’s struggle is one that uses human rights as its main framing mechanism, specifically Achmat explains,

²⁹¹ Mbali, "The Treatment Action Campaign," 11.

²⁹² Ibid.

²⁹³ Ibid., 10.

²⁹⁴ Ibid., 14.

²⁹⁵ Ibid., 16.

“it is the realization of the right to life and health for HIV positive people, as equal citizens.”²⁹⁶ Focusing on human rights, which was a main focus of the liberation movement and early gay rights activism, allowed TAC to fit in the post-apartheid era, especially because TAC strategically had to position itself as not opposed to the ANC government, but to its AIDS policies.²⁹⁷ The second aspect, influenced by the anti-apartheid struggle, is the organizational form of TAC. Both Achmat and Heywood have explained that their experience in the Marxist Workers Tendency (MWT) wing of the ANC led them to build the broadest coalition possible, which includes churches, unions, gay rights organizations, scientists, lawyers and policy makers.²⁹⁸ The third aspect is the strategies and repertoires of contention that were borrowed from the anti-apartheid struggle and used between 1998 and 2001 against the South African and US governments and the PMA. For example, TAC protests and events “commonly start with the right-fisted salute and the cry ‘Amandla!’ (power), which the audience answers with ‘Awethu!’ (to the people).”²⁹⁹ Furthermore, TAC members refer to one another as comrades and create new lyrics for old struggle songs, which Robins calls the “creative re-appropriation of locally embedded political symbols, songs and styles of the anti-apartheid struggle.”³⁰⁰ Despite TAC’s historical origins in gay rights and AIDS activism of the anti-apartheid struggle, TAC is still “a post-apartheid creature, which has used entirely new political and legal spaces created in post-apartheid South Africa.”³⁰¹ Therefore, it is important to examine the domestic political opportunities and threats that TAC faced in South Africa between 1998 and 2000, as it helps explain the emergence and impact of TAC.

TAC encountered many new political opportunities in South Africa between 1998 and 2000, which influenced how it emerged in relation to its early struggle history. Most importantly was the democratic ANC government, which “marked the moment when the leaders of South Africa’s anti-apartheid social movements entered the corridors of political power,”³⁰² and created new elite allies for TAC to work with. However, this transformation was also difficult as anti-apartheid activists were “absorbed into the post-apartheid government, thus leaving opponents of the government without a ‘voice’ with which to express or a mechanism to organize opposition.”³⁰³ Nevertheless, the democratization of South Africa, along with the adoption of South Africa’s Constitution, “one of the most progressive in the world for its inclusion of socio-economic rights, and the rights to gender

²⁹⁶ Ibid., 21.

²⁹⁷ Friedman and Mottiar, "Seeking the High Ground," 25.

²⁹⁸ Grebe, "The Treatment Action Campaign's Struggle for AIDS Treatment," 854.

²⁹⁹ Ibid., 852.

³⁰⁰ Robins, "'Long Live Zackie, Long Live'," 665.

³⁰¹ Mbali, "The Treatment Action Campaign," 19.

³⁰² Richard Ballard et al., "Introduction - from Anti-Apartheid to Post-Apartheid Social Movements," in *Voices of Protest*, eds. Richard Ballard, Adam Habib and Imraan Valodia (KwaZulu-Natal: University of KwaZulu-Natal Press, 2006), 1.

³⁰³ Ibid., 15.

equality and non-discrimination,³⁰⁴ created opportunities for TAC to frame their struggle for ARV access in terms of human rights enshrined in the Constitution.

On the other hand, TAC was confronted with many serious political threats, which made it difficult for TAC to make a domestic impact at this time. The first and most critical threat was the government's ambiguous and harmful AIDS and ARV policies, specifically under Mbeki and MOH, Manto Tshabalala-Msimang. There are many arguments concerning South Africa's AIDS policies and Mbeki's 'denialism,'³⁰⁵ but they are not of major concern here.³⁰⁶ Nevertheless, once Mbeki began questioning the safety and not the affordability of ARVs in 1999, TAC faced a major threat to ARV access and in turn had to re-strategize and re-focus its efforts from fighting the US government and pharmaceutical industry, to combating the South African government. The second threat, which is related to the first political threat, was the importance of economic growth and development in post-apartheid South Africa. Susan Cleary and Don Ross explain that South Africa's macroeconomic policy, outlined in the Growth, Employment and Redistribution (GEAR) plan, summarized South Africa's early commitment to increasing exports and foreign investment through "credibility of its stance in ongoing multilateral trade negotiations"³⁰⁷ in the WTO to improve economic growth. Cleary and Ross argue that South Africa repeatedly attempted to portray itself as a "model WTO citizen,"³⁰⁸ which restricted the ability of the government to negotiate for affordable or generic ARVs, and in turn provided a mobilizing force for TAC. Lastly, however, domestic threats posed by external actors, such as the PMA lawsuit and US pressure on South Africa, provided opportunities for TAC to target and campaign around. To summarize, transnational AIDS treatment activism in South Africa emerged by way of old strategies and networks in response to new domestic and international threats and opportunities in 1998.

In terms of Tarrow's domestic processes, transnational AIDS treatment activism in South Africa primarily reflects the process of internalization, rather than that of global issue framing as in the US. This is because TAC was formed to improve AIDS treatment access in South Africa and did not use international symbols to connect to the global struggle for ARV access. However, TAC is known for

³⁰⁴ Mbali, "The Treatment Action Campaign," 3.

³⁰⁵ Geffen of TAC defines an AIDS denialist as an individual that promotes one of the following views: "that HIV does not cause AIDS, that there is not a large AIDS epidemic in sub-Saharan Africa and that antiretrovirals cause more harm than good," and claims that Mbeki held all three views during his Presidency. See Geffen, *Debunking Delusions*, 2.

³⁰⁶ See Nicoli Nattrass, *The Moral Economy of AIDS in South Africa* (Cambridge: Cambridge University Press, 2004).; Anthony Butler, "South Africa's HIV/AIDS Policy, 1994–2004: How can it be explained?" *African Affairs* 104, no. 417 (2005): 591-614.; Robins, "Long Live Zackie, Long Live'," 661, for discussions on Mbeki's AIDS denialism.

³⁰⁷ Cleary and Ross, "The 1998-2001 Legal Struggle," 457.

³⁰⁸ *Ibid.*, 458.

framing access to AIDS treatment as a universal human right,³⁰⁹ but Heywood notes that TAC did not intend to work or address structural and international political issues interfering with ARV access.³¹⁰ Therefore, Tarrow's internalization process is best to explain how and why TAC began by domestically targeting the South African government, before Mbeki took office and despite activists knowing that the pharmaceutical industry was its toughest challenger. According to Tarrow, internalization occurs when external or foreign pressure influences domestic policy issues, which in turn cause activists to make domestic claims and target their own governments even when a threat or pressure is external.³¹¹ Evidence of this process is found in TAC's first campaign, which targeted and pressured the government to provide free ARVs for HIV-positive pregnant women, despite TAC members knowing that the pharmaceutical industry was the main actor to pressure for the reduction of ARV prices to improve AIDS treatment access. Thus, early protests and demonstrations by TAC members, reflect what Tarrow explains as *internalized protests*, which involve "contentious claims making in which," external or foreign actors are "the source or an indirect target of a protest by domestic actors, but the direct target of the action [is] either the state, its components, or other domestic actors present on its territory."³¹² However, as Mbeki began to pose a bigger threat to ARV treatment access in South Africa, compared to that of external pressure from the US, campaigning turned to these domestic and internal threats, which does not easily reflect Tarrow's processes of internalization or global issue framing during the period of this study.

In summary, TAC's origins are based in gay rights and AIDS activism during the anti-apartheid struggle. TAC's organizational form, human rights framing, and strategies stem from a uniquely South African experience, which were re-worked in order to effectively respond to opportunities and threats posed by both external and internal factors from 1998 to 2000. The domestic process, specifically internalization, by which TAC emerged and attempted to contribute to the reduction of ARV prices differed from the US experience, and therefore, transnational AIDS treatment in South Africa must be examined separately from transnational AIDS treatment activism in the US. Lastly, however, this dissertation finds that transnational AIDS treatment activism in South Africa did not contribute, as clearly as in the US, to the reduction of ARV prices because no domestic policies or new commitments to improve AIDS treatment access were made by the South African government.

³⁰⁹ See Krista Johnson, "AIDS and the Politics of Rights in South Africa: A Contested Terrain," *Human Rights Review* 7, no. 2 (Jan, 2006): 115-129, for a discussion of the limitations about rights based frames in terms of socio-economic rights both used by the pharmaceutical industry and transnational AIDS treatment activists.

³¹⁰ Friedman and Mottiar, "A Rewarding Engagement?," 549.

³¹¹ Tarrow, *The New Transnational Activism*, 80.

³¹² *Ibid.*, 86.

4.1.3 Domestic Impact of Transnational AIDS Treatment Activism on ARV Prices

This domestic analysis of transnational AIDS treatment activism in the US and South Africa has contributed to the empirical and theoretical findings of this dissertation. First, by analyzing the origins and emergence of TAC and Health GAP it is clear that transnational AIDS treatment activism in the US and South Africa emerged from different contexts, responded to different political opportunities, and occurred through different processes. Therefore, transnational AIDS treatment activism in the US and South Africa should be understood as separate and uncoordinated process that emerged simultaneously and not be assumed as one entity until there is clear evidence of global cooperation and coordination, which is discussed in the following section. Second, the most evident domestic impact of transnational AIDS treatment activism is found in the US, as Health GAP influenced and worked with the US government to change its policy towards South Africa and ARV access, which ultimately weakened support and protection of the pharmaceutical industry. Furthermore, the strategies and processes used by Health GAP to accomplish this do reflect Tarrow's global issue framing process. As such, this research finds Tarrow's argument, that transnational activism is still very much a product of domestic politics and has its most visible impact on domestic ground, does apply in this case and is an instructive analytical tool. Especially as it was essential to examine the history behind transnational AIDS treatment activism in both countries in order to account for how strategies, tactics, and resources were used by both organizations to be most effective within their own contexts.

4.2 Transnational Coordination of AIDS Treatment Activism

This section examines the ways in which Health GAP and TAC began to coordinate their efforts by employing Tarrow's transitional and international processes, which explain how domestic claims and localized direct action within a specific state, extend across borders and become adopted within foreign states or international institutions. In addition, by tracing the processes involved in coordinating transnational activism across borders, it is possible to examine the impact of coordinated strategies and direct action of Health GAP and TAC on ARV price reductions. This analysis section highlights the most apparent limitations of Tarrow's theoretical arguments because transnational AIDS treatment activists in the US and South Africa had similar or aligned domestic claims from the very beginning. As such, coordination between Health GAP and TAC did not entirely require conscious or calculated strategic efforts to promote or gain external support for their causes or claims. Furthermore, it is worth noting that contentious issues and claims that are located within one state or region extend across borders unintentionally. For example, Berkman travelled to South Africa in 1998 during his fellowship at the Columbia School of Public Health, and explains that it was when he was in South Africa that he "began to see the outlines of what was going to happen around AIDS in

Africa³¹³ and decided to focus on improving global ARV treatment access. This highlights an important point for the case, as it shows that the adoption of certain causes and contentious issues in one location that are rooted in another location, do not necessarily reflect strategic coordination, especially in this case, as TAC had not begun to work on improving ARV access at this time. Nevertheless, it is worth examining the ways in which TAC and Health GAP began strategically working together in 2000, which created a strong unified and international demand for generic ARVs that was later fulfilled by CIPLA in February 2001.

4.2.1 Diffusion and Scale Shifting

According to Tarrow, diffusion and scale shifting are two transitional processes that take place in order for domestic claims and localized action to extend beyond state borders. Diffusion through relational, non-relational, and mediated pathways refers to “the transfer of claims or forms of contention from one site to the other.”³¹⁴ While scale shifting, which directly builds upon diffusion is “the change in the number and level of coordinated contentious actions to a different focal point, involving a new range of actors, different objects, and broadened claims.”³¹⁵ This section will apply and test Tarrow’s transitional processes to examine how Health GAP and TAC began working together.

As stated above, diffusion of claims or localized action from the US to South Africa or vice-a-versa was not entirely necessary for coordination across borders, or for that matter, strategically calculated. However, what is useful to examine here are Tarrow’s three pathways of diffusion, as they help explain the ways in which activists from South Africa and the US became in contact with one another, formed working relationships, and eventually began coordinating their efforts in 2000. In this case all three diffusion pathways are evident in the coordination and cooperation of Health GAP and TAC. Primarily, evidence points to the importance of mediated diffusion, that is the building and forming of new relationships and networks between activists by way of a third party or broker. This occurred when Eric Goemare of MSF advised Achmat to attend an MSF conference in Amsterdam in late 1999, where Achmat met US activists from ACT UP and CPT.³¹⁶ Furthermore, in early 2000, a mutual friend put Achmat in touch with Gregg Gonsalves and Mark Harrington of the Treatment Action Group (TAG), an offshoot of ACT UP, which evolved into a relational diffusion pathway.³¹⁷ Achmat traveled to the US in early 2000 to meet with Gonsalves and Harrington to form a working partnership, in which TAG, TAC, and MSF conducted HIV and AIDS literacy workshops as part of

³¹³ Smith and Siplon, *Drugs into Bodies*, 56.

³¹⁴ Tarrow, *The New Transnational Activism*, 32.

³¹⁵ McAdam, McCarthy and Tilly, *Dynamics of Contention*, 331.

³¹⁶ Grebe, "The Treatment Action Campaign's Struggle for AIDS Treatment," 863.

³¹⁷ *Ibid.*, 864.

TAC's Literacy Campaign in June and July 2000.³¹⁸ Furthermore, Achmat acknowledges that TAC only became exposed to information regarding parallel importation and compulsory licensing to circumvent IPRs when they came into contact with CPT through the internet and by email, which is an example of non-relational diffusion.³¹⁹ Therefore, it appears in this case that all three pathways of diffusion: relational, non-relational and mediated, helped connect transnational AIDS treatment activists from the US and South Africa, which subsequently laid the foundation for future transnational collaboration and is best explained and analyzed through Tarrow's transnational coalition formation process.

Tarrow's process of scale shifting, the coordination of collective action at a different level from where it originated,³²⁰ is the process this research finds most problematic as an analytical tool. According to Tarrow, scale shifting may occur upwards or downwards in the international system, but he does not explain or conceptualize high and/or low levels of the international system, especially in relation to the position of the US.³²¹ Scale shifting is problematic for this case because collective action, specifically under Health GAP, which went to South Africa to work alongside TAC in its Literacy Campaign, emerged at a different level than where it originated. However, collective action, specifically under TAC, which had similar claims for affordable ARV access, originated separately in South Africa. Therefore, Tarrow's process of scale shifting is not useful for this case, as any collective action in South Africa or the US may be interpreted or described as a scale shift. Furthermore, because Tarrow's conceptualization of levels in the international system is vague and does not account for US power and influence in the international system, it is not clear if the adoption or coordination of claims and collective action from the US to South Africa or from South Africa to the US reflect diffusion or upwards/downwards scale shifting.

4.2.2 Externalization and Transnational Coalition Formation

Tarrow explains that externalization and transnational coalition formation are two international processes that help expand transnational activism and connect transnational activists around the globe. Externalization refers to "the vertical projection of domestic claims onto international institutions or foreign actors,"³²² in order to gain support from external allies when state governments are unresponsive or repressive. Transnational coalition formation refers to the process through which the horizontal formation and expansion of similar goal-oriented networks among actors across state borders occur.³²³ Additionally, transnational coalitions vary in form and length depending on the

³¹⁸ Ibid.

³¹⁹ Friedman and Mottiar, "A Rewarding Engagement?," 547.

³²⁰ Tarrow, *The New Transnational Activism*, 32.

³²¹ Ibid.

³²² Ibid.

³²³ Ibid.

degree of involvement and time of involvement. This section will apply and test Tarrow's international processes to the case of transnational AIDS treatment activism, which contributes to the empirical findings of this research and have theoretical implications.

First, by applying Tarrow's externalization process to transnational AIDS treatment activism, only one case of externalization appears, from TAC in South Africa to Health GAP in the US. This is because there is little evidence suggesting that domestic claims from the US against the US government and the pharmaceutical industry were externalized to TAC in South Africa, or that TAC reached out to US activists to exert pressure on the South African government at this stage. Additionally, there is little evidence suggesting that the adoption of the WHO's RDS in 1999, which was against the US government and its support of the pharmaceutical industry's IPRs, was due to the externalization process from either country. Most evidence of the externalization of claims from South Africa to the US is found when TAC "reached out" to Health GAP to pressure Pfizer to reduce the cost of fluconazole in March 2000.³²⁴ This reflects Tarrow's externalization process because a specific claim that originated in South Africa that was not acted upon by the South African government extended to the US and resonated with US activists, which in turn pressured Pfizer through direct action. This is one clear example of externalization in the case of transnational AIDS treatment activism, however it could also be interpreted as scale shifting because it involved an increase in direct actions in more than one country. Hence, this research suggests that Tarrow's distinction between transitional and international processes, specifically scale shifting and externalization, are problematic for explaining early coordination efforts between TAC and Health GAP.

Second, by applying Tarrow's transnational coalition formation process, this research suggests that a transnational event coalition formed around the 13th International AIDS Conference in South Africa, which has since evolved into a transnational campaign coalition responding to new international opportunities and threats. An event coalition is characterized by a short duration period, such as the conference, with high intensity involvement, which occurred as thousands of demonstrators in South Africa and over 250 organizations around the world protested against the pharmaceutical industry and high ARV prices.³²⁵ Additionally, it was at this time when the transnational coalition began demanding the generic supply of ARVs in response to the AAI, which activists perceived as an international threat because the program reduced ARV prices by 90% but still left ARVs unaffordable for most developing countries. This reflects Tarrow's opportunity spirals, because transnational AIDS treatment activists were reacting to a new international political threat, began re-framing their claims, and re-strategizing to secure the generic supply of ARVs. Furthermore, scholars and activists from

³²⁴ "Health GAP Timeline."

³²⁵ Grebe, "The Treatment Action Campaign's Struggle for AIDS Treatment," 862.

TAC, Health GAP, and ACT UP, which “set the agenda for the global protests for access to medication,”³²⁶ note that organizing and preparing for this conference was an important and galvanizing period for transnational AIDS treatment activism around the world.³²⁷ The conference helped establish and solidify international partnerships that have since been sustained in the form of a transnational campaign coalition, which has been heavily involved in the fight against HIV and AIDS since 2000 and has responded to new international policies, resources and opportunities to directly help with ARV access. In conclusion, Tarrow’s transnational coalition formation process is the most instructive for this case because it helps explain how TAC and Health GAP, similar goal-oriented organizations formed horizontal ties to one another and campaigned together to improve ARV access. In addition, this process helps explain this case because transnational AIDS treatment activism did not emerge in a linear fashion, originating in one location that spread and was then adopted by another.

This section has employed Tarrow’s transitional and international processes in order to examine the ways in which transnational AIDS treatment activists in South Africa and the US began working together and eventually formed a transnational campaign coalition in 2000. The major criticism presented here is that each of Tarrow’s processes, by which transnational activism becomes globally coordinated in more than one country follows a rather linear projection. First, this is because each process traces direction action, contentious claims, and activism from one location to another. Second, this is because Tarrow’s processes build upon one another. Therefore, Tarrow’s distinctions between transitional and international processes are not entirely useful analytical tools on their own and require further development to account for the emergence of transnational activism in more than one location, unintentional diffusion, and levels in the international system in relation to US power.

4.2.3 Coordinated Impact on ARV Prices

Applying and testing Tarrow’s transitional and international processes to this case raised a number of issues concerning the causal relationship between transnational AIDS treatment activism and ARV price reductions. First, transnational AIDS treatment activism was largely uncoordinated between 1998 and 2000, as Achmat came into contact with US transnational AIDS treatment activists by the end of 1999. Efforts only became strategically coordinated in March 2000 by way of all three pathways of diffusion and primarily reflect Tarrow’s transnational coalition formation. However, this occurred after negotiations between UNAIDS and the pharmaceutical industry to reduce ARV prices by 90% were underway, and therefore the impact of coordinated efforts can only be traced to the period after the 13th International AIDS Conference in July, when global protests demanding generic ARVs were held. As such, the reduction of ARV prices to \$350 USD by CIPLA and MSF could have

³²⁶ Halbert, *Resisting Intellectual Property*, 97.

³²⁷ Smith and Siplon, *Drugs into Bodies*, 97.; Grebe, "The Treatment Action Campaign's Struggle for AIDS Treatment," 861.

been influenced by the global protests demanding generic ARVs, but there is little evidence to support this. Therefore, this research suggests that transnational AIDS treatment activism, specifically referring to the brokerage work of CPT and Love, had a limited contribution in the deal between CIPLA and MSF by helping coordinate meetings and discussions prior to the February 2001 announcement, which reduced ARV prices to less than \$350 USD.

4.3 International Opportunities and Threats

This chapter has analyzed the domestic origins and political opportunities by which transnational AIDS treatment activism emerged in the US and South Africa, and has examined the processes by which Health GAP and TAC began working together. This section examines international opportunities and threats, which according to Tarrow, arise from changes in the international political structure and influence strategies, resources, and targets of transnational activism and its overall effectiveness.³²⁸ The case study in Chapter 3 highlighted that from 1996 to 2001 the international political and economic structure was shifting to address AIDS. Therefore, the internationalization of AIDS by way of international health institutions, such as the WHO and UNAIDS, and international economic development institutions, specifically the World Bank and the WTO, created new opportunities and threats, which transnational activists, states, and the pharmaceutical industry reacted to and capitalized upon. However, by analyzing these changes in the international political structure, this research makes it most significant finding. Namely that international policy shifts, new commitments and resources to improve AIDS treatment access were not directly used by transnational activists at this time, but separately contributed to the reduction of ARV prices.

4.3.1 The World Bank and WTO

The WTO's 1995 agreement on TRIPS posed an international threat for ARV access, especially as it was used by both the US government and the pharmaceutical industry to pressure South Africa and other countries from producing or importing more affordable generic medicines. On the other hand, the WTO presented economic growth opportunities, which the South African government wanted to take advantage of and therefore did not want to appear "as an administration disrespectful of property rights, including intellectual property rights."³²⁹ Therefore, the WTO presented both opportunities and threats for various state and non-state actors during this period, but for the most part the WTO was a constant and consistent threat to improving ARV treatment access, as no major policy changes took place at this time.³³⁰

³²⁸ Tarrow, *The New Transnational Activism*, 8.

³²⁹ Cleary and Ross, "The 1998-2001 Legal Struggle," 458.

³³⁰ It is worth noting here that in November 2001, the WTO's Doha Declaration on TRIPS and Public Health was established to address the contentious issues concerning IPRs and pharmaceuticals, however there it is controversial and has not been used by countries to improve access to medicines. See Paul Vandoren and Jean

During this period, the World Bank offered new international political opportunities for transnational AIDS treatment activism and ARV access, because it was the first international economic institution to consider AIDS a risk to development and began to advocate and coordinate global responses to address the epidemic. The World Bank presented new opportunities by way of its \$500 million USD MAP program and contributed to a shift in the international economic structure by its new commitment to address AIDS as a threat to economic development in Africa. However, these new commitments were not created by or used by transnational activists at this time. Therefore, these changes must be understood, as autonomous and separate developments, which placed pressure on the pharmaceutical industry, as ARV prices were a major concern for the World Bank.

4.3.2 UNAIDS and the WHO

The creation of UNAIDS as a unifying international institution aiming to involve civil society organizations and PLWHA, presented opportunities for transnational AIDS treatment activism. However, during this period, the pharmaceutical industry and state governments had most access to UNAIDS, not transnational AIDS treatment activists. This is evident in the early history of UNAIDS and its initial DAI and subsequent AAI that reduced ARV prices in partnership with the pharmaceutical industry without involvement from TAC or Health GAP. Both the DAI and AAI laid the foundation for future ARV price reductions, country specific programs and were in alignment with the goals of transnational AIDS treatment activists. However, AAI was perceived as a threat by some transnational AIDS treatment activists and was subject to protest because prices were not reduced low enough. As a result, the generic supply of affordable ARVs was demanded. Therefore, UNAIDS presented both opportunities and threats for transnational AIDS treatment activism, but also directly contributed to ARV price reductions.

The WHO played an important role in shifting the international political structure with its RDS, which was also in alignment with the goals and claims of transnational AIDS treatment activists, especially as it stressed the importance of access to medicines over the importance of protecting IPRs. This new policy provided activists with a specific international commitment, which governments could be held accountable to and used to promote and improve ARV access. However, during this period little evidence suggests that transnational AIDS treatment activists used this or other international policy commitments to further their claims for affordable ARVs.

In summary, there were clear changes and shifts occurring in the international political and economic structure in favor of improving AIDS treatment access and presenting new opportunities for transnational activism. However, these shifts occurred simultaneously and independently from

Van Eckhaute, "The WTO Decision on Paragraph 6 of the Doha Declaration on the TRIPS Agreement and Public Health," *The Journal of World Intellectual Property* 6 (2003): 779-793, for more information.

transnational activist efforts in the US and South Africa, and were not capitalized or acted upon as new opportunities during this period. Therefore, these changes and shifts should be examined in combination with transnational activism and in terms of their combined contribution to ARV price reductions in 2000 and 2001.

4.3.3 International Political Shifts and Impact on ARV Prices

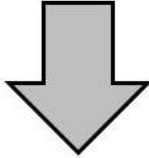
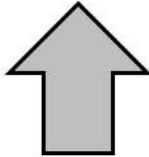
By examining shifts in the international political structure, both empirical and theoretical points were raised. Firstly, changes in the international political structure did occur, however as they happened at the same time that transnational AIDS treatment activists were working on the ground, little evidence suggests that these changes influenced transnational activism and its contribution to global ARV price reductions. Secondly, these international structural changes in policies, funding and commitments to AIDS treatment access directly affected ARV price reductions as seen in both the DAI and AAI. Lastly, from this analysis, Tarrow's conceptualization of internationalization, or changes in the international political structure, appears vague and does not account for changes that may be in alignment with the goals of transnational activism, which makes it difficult to assess causal relationships between transnational activism and desired outcomes.

4.4 Conjuncture: From Above and Below

By employing Tarrow's framework for analyzing transnational activism the first three sections of this analysis have made three important points. First, it illustrated that the emergence of transnational AIDS treatment activism in the US and South Africa reflects past activism, unique domestic opportunities and threats specific to each country, and that it emerged through different domestic processes. Secondly, it explained transnational AIDS treatment activism in both countries occurred simultaneously and was largely uncoordinated until 2000, when a transnational coalition formed. Thirdly, this analysis illustrated that the international political structure was shifting by way of new policy commitments to address AIDS, creating new opportunities and threats, but that these changes were more than just new opportunities for transnational activism, especially as they occurred simultaneously and were not directly acted upon during this period. Thus, separate and uncoordinated policy developments, domestically in the US and South Africa and internationally, all aiming to address the growing AIDS epidemic and improve ARV access, occurred simultaneously alongside transnational AIDS treatment activism during this period. Therefore, this research suggests that the combination of pressure to improve the affordability and availability of ARVs from below, by the US government and through the efforts of transnational AIDS treatment activists, and pressure from above, by international institutions, contributed to the reduction of ARV prices from September 2000 to February 2001. Please see Figure 9 (on the following page) for a diagram of conjuncture that

outlines major domestic and international policy changes, commitments, and direct action by Health GAP and TAC.

Figure 9. Diagram of Conjuncture

	Quarters	Q 1 (January - March)	Q 2 (April - June)	Q 3 (July - September)	Q 4 (October - December)
	Years				
International Pressure From Above 	1996-1997	•UNAIDS (1996)		•11 th Int. AIDS Conference (1996) – 'HAART'	•DAI launches (1997)
	1998	•WHA proposes RDS		•12 th Int. AIDS Conference – 'Bridging the Gap'	
	1999	•WHA approves RDS	•World Bank – 'Responding to a Development Crisis'		•IPAA launches
	2000	•UNSC Meeting - 'AIDS is a security issue'	•AAI launches	•13 th Int. AIDS Conference – South Africa •CIPLA negotiations •World Bank MAP announced	•CIPLA negotiations and announcement in February 2001
CONJUNCTURE	Global ARV prices reduced from \$10,000 to \$1,000 USD in September 2000 and to less than \$350 USD by February 2001				
 Domestic Pressure From Below	2000	•Health Gap and TAC begin to coordinate efforts for Moraka Defiance Campaign	•Health GAP and TAC target Pfizer in the US and SA •President Clinton's Executive Order issued	•Transnational event coalition forms and global protests	•TAC imports generic drugs from Thailand
	1999	•Health Gap campaigns for CARE •TAC protests for ARVs to prevent MTCT	•Health GAP targets Gore •SA on 301 Watchlist •Mbeki President	•Health GAP meets with Congress •TAC targets US government/PMA	•Health Gap storms USTR •Clinton Announcement
	1998	•PMA lawsuit filed against SA	•SA on 301 Watchlist		•Health GAP and TAC form
	1996-1997	•SA issue health proposal (1997)	•US and PMA pressure (1997)		•Medicines Act signed (1997)

In conclusion, this dissertation argues that “a conjunctural unfolding, of interactions of originally separately determined processes,”³³¹ led by transnational AIDS treatment activism in the US and South Africa, and by international institutions, contributed to the reduction of global ARV prices. Therefore, without broader structural changes by international institutions to combat AIDS, transnational AIDS treatment activism might not have been as effective as suggested in the literature because a conjunctural explanation emphasizes, “a particular combination of structural causes and events, in a particular time and place [that] may create unique outcomes that will not necessarily be repeated.”³³² As a result, this study provides support for Shadlen’s argument that highlights the “limits of transnational activist networks as enduring agents of change”³³³ for improving AIDS treatment access.

³³¹ Theda Skocpol, *States and Social Revolutions: A Comparative Analysis of France, Russia, and China* (New York: Cambridge University Press, 1979), 320.

³³² Jeffrey Paige, "Conjuncture, Comparison, and Conditional Theory in Macrosocial Inquiry," *American Journal of Sociology* 105, no. 3 (Nov. 1999): 782.

³³³ Shadlen, "The Political Economy of AIDS Treatment." 559.

Chapter 5: Conclusion

The aim of this dissertation was to provide an indepth examination of the suggested causal relationship between transnational AIDS treatment activism and the reduction of global ARV prices to improve AIDS treatment access between 1996 and 2001. By drawing upon theoretical arguments developed by Sidney Tarrow in his book *The New Transnational Activism* to explain *how transnational AIDS treatment activism in the US and South Africa emerged and contributed to the reduction of global ARV prices between September 2000 and February 2001*, this study had both empirical and theoretical objectives. First, it aimed to provide a chronological and comprehensive account of the emergence and work of Health GAP and TAC, and the international politics concerning ARV price reductions. Second, it aimed to provide a critical analysis of current transnational activism theory by employing and testing the applicability of Tarrow's theoretical arguments to the case of transnational AIDS treatment activism.

Given the research objectives of this dissertation, the theoretical single case study method was used to account for and examine the history of transnational AIDS treatment activism, domestic factors influencing strategies and targets, how efforts became coordinated internationally, and changes in the international political structure addressing AIDS during this period. The study used academic and personal accounts of transnational AIDS treatment activists in the US and South Africa in order to provide a comprehensive narrative of strategic events and outcomes. It was structured according to domestic and international levels of analysis to specifically highlight, domestic and international political opportunities and threats, which according to Tarrow are vital for analyzing how transnational activism emerges and is effective.

The introductory chapter of this dissertation provided a literature review of the work used by this study, which primarily fell into three broad categories: articles on the transnational political contention concerning ARVs and IPRs, literature on transnational AIDS treatment activism, and personal and academic accounts of TAC and Health GAP. Thus, Chapter 1 and the literature review provided an overview of the problem this research focused on (ARV prices and IPRs), the suggested explanation for ARV price reductions (transnational AIDS treatment activism, TAC, and Health GAP), and the framework and research design used to examine the suggested causal relationship.

The second chapter was an analysis of Tarrow's theoretical arguments concerning how transnational activism emerges and is effective, which formed the analytical framework used and tested in this study. Chapter 2 first outlined Tarrow's argument that transnational activism is very much rooted in domestic politics and has most impact on domestic politics. Additionally, it provided concepts and

factors borrowed from social movement theories that help examine transnational activist organizations. Secondly, the chapter explained Tarrow's transitional and international processes, which he argues help to trace how transnational activism extends across borders and becomes coordinated across the globe. Thirdly, the chapter discussed Tarrow's argument that changes in the international political structure create new opportunities and threats, resources, and avenues for transnational activists to act upon and be most effective. Lastly, the chapter outlined how Tarrow's arguments would form the framework for the case study and the analysis to come.

The third chapter of this dissertation was the case study of transnational AIDS treatment activism and ARV price reductions from 1996 to 2001. Chapter 3 was divided into five sections by time period and further separated into domestic and international levels of analysis to account for both domestic and international political opportunities to examine transnational activism. The chapter outlined important policy shifts in the US and South Africa, the emergence and direct action of Health GAP and TAC, new policy commitments by international institutions to address AIDS, and lastly established that ARV price reductions between September 2000 and February 2001 occurred after a combination of developments and events, related and unrelated to transnational AIDS treatment activism.

The fourth chapter of this study was an analysis of the emergence and impact of transnational AIDS treatment activism, which applied and tested Tarrow's theoretical arguments. The first section of Chapter 4 examined domestic factors, political opportunities, and processes of Health GAP in the US and TAC in South Africa. The second section used Tarrow's transitional and international processes to explain how Health GAP and TAC began working together in 2000 and highlighted the limitations of these processes. The third section analyzed changes in the international political structure, namely the creation of UNAIDS, and new policy commitments by the World Bank and the WHO, and its impact on transnational AIDS treatment activism and ARV price reductions. The last section outlined the main argument of this dissertation, that the limited impact of transnational AIDS treatment activism is best understood in conjunction with broader structural changes that also contributed to the reduction of ARV prices to improve AIDS treatment access.

In answering the central research question of this theoretical single case study, this dissertation has provided a critical explanation about the contribution of transnational AIDS treatment activism to the reduction of global ARV prices and has challenged Tarrow's theoretical arguments concerning transnational activism, which are both discussed in the following sections.

5.1 Transnational AIDS Treatment Activism and ARV Price Reductions

This dissertation provided a chronological and empirical account of transnational AIDS treatment activism, which emerged simultaneously in the US and South Africa in 1998. In addition, it examined

the transnational political contention between the US and South Africa concerning ARVs and IPRs, which was a strong mobilizing force for Health GAP and TAC, and accounted for changes in the international political structure for addressing the AIDS epidemic by way of the World Bank, the WHO, and UNAIDS from 1996 to 2001.

The case study in Chapter 3 explained that throughout 1999, Health GAP and TAC began targeting and pressuring their own governments and the pharmaceutical industry to improve AIDS treatment access and reduce ARV prices. Domestically in the US, Health GAP primarily targeted and criticized the US government for exerting pressure on South Africa for its 1997 Medicines Act. Members of Health GAP, namely from CPT and ACT UP, specifically targeted Gore and Barchefsky, and began working with Congress members to reverse US policy towards South Africa. Eventually these efforts led to the announcement made by Clinton on World AIDS Day 1999 that explained that the US would no longer interfere in AIDS treatment access in SSA.³³⁴ Domestically in South Africa, TAC began by pressuring the South African government to provide free ARVs for HIV-positive pregnant women, but soon focused its efforts and targeted the PMA and US government when the MOH claimed ARVs were too expensive in 1999.³³⁵ However, once Mbeki became President and began questioning the safety and not the affordability of ARVs, TAC re-shifted its efforts back to pressuring the South African government, but no official government policies for improving ARV access were made at this time. Thus, the evidence suggests that transnational AIDS treatment activism had a limited contribution to ARV price reductions by pressuring the US government to change its policy on ARV access and IPRs, which weakened support for the pharmaceutical industry and increased government pressure on the industry to reduce ARV prices.

Transnational AIDS treatment activism in the US and South Africa was largely uncoordinated until the end of 1999, when transnational activists from TAC and Health GAP came into contact with one another and began working together throughout 2000. Both organizations targeted and pressured Pfizer to reduce its prices for ARVs and organized protests and demonstrations around the world leading up to the 13th International AIDS Conference held in South Africa during July 2000. At this time ARV prices had reduced by 90% with the AAI of UNAIDS, but transnational AIDS treatment activists believed that this was still inadequate and began to demand the generic supply of ARVs. After these global protests were held, CIPLA, MSF, and CPT began negotiating and planning a program to provide one year of TCT of generic ARVs for \$350 USD. It is difficult to prove that transnational AIDS treatment activism directly contributed to this, but there is evidence to suggest that particular individuals involved in Health GAP, such as CPT leader Love, were influential in these

³³⁴ Sawyer, "An ACT UP Founder "Acts Up" for Africa's Access to AIDS," 101.

³³⁵ "History of TAC (1998-2001)," TAC, accessed September 22, 2011.
<http://www.tac.org.za/Documents/Other/tachist.pdf>, 9.

early discussions.³³⁶ Therefore, the contribution of transnational AIDS treatment activism to global ARV price reductions is limited, and is rooted in the domestic politics of the US and is evident after efforts became coordinated across the globe.

In addition to accounting for transnational AIDS treatment activism in the US and South Africa, the case study illustrated that during the period from 1996 and 2001, a number of international political and structural changes occurred. Primarily these included the DAI, AAI, and IPAA programs of UNAIDS, the RDS of the WHO, and the MAP of the World Bank, all of which were created to improve the global response to AIDS as a public health and development crisis, and targeted exorbitant ARV prices protected by IPRs. These policies and programs created new opportunities and resources for transnational AIDS treatment activism, but were not yet capitalized upon at this stage because they were established simultaneously as TAC and Health GAP were working on the ground. Therefore, this dissertation suggested that changes in the international political opportunity structure were more than just opportunities for transnational activism, as they directly contributed to ARV price reductions and were a major source of pressure on the pharmaceutical industry.

From this empirical and chronological account, this dissertation has suggested that transnational AIDS treatment activism had a limited contribution to global ARV price reductions between September 2000 and February 2001. It suggested secondly that changes in the international political structure that originated separately and simultaneously within UNAIDS, the WHO, and the World Bank also contributed to the reduction of global ARV prices. Therefore, the central argument presented in this dissertation was that a conjuncture of separately located processes from below, of which transnational AIDS treatment activism had a limited role, and above, by international institutions, contributed to global ARV price reductions.

5.2 Tarrow and Transnational AIDS Treatment Activism

This dissertation employed and tested the applicability of Tarrow's theoretical arguments to examine the causal impact of transitional AIDS treatment activism in Chapter 4 and to provide a critical analysis of transnational activism theory. The empirical findings of this research, as mentioned above, had theoretical implications that highlighted both strengths and weaknesses of Tarrow's hypotheses that will be summarized here and can form the basis for future research.

First, the analysis of this dissertation focused on Tarrow's claim that transnational activism is rooted in and has most impact on domestic state politics, by examining the domestic origins, factors, political opportunities and process of transnational AIDS treatment activism in the US and South Africa. This

³³⁶ Sell and Prakash, "Using Ideas Strategically," 162.

revealed that transnational activism, under Health GAP and TAC, was unique to each location, emerged through different processes in response to different domestic political opportunities and had most impact in different and domestic ways, thus highlighting and confirming Tarrow's main theoretical argument.

Secondly, the analysis applied and tested Tarrow's transitional and international processes, which explain and trace how transnational activism expands and becomes strategically coordinated in more than one country. However, because Health GAP and TAC emerged simultaneously in 1998 with similar goals to improve AIDS treatment access, and only began coordinating their efforts in 2000; Tarrow's processes were not entirely useful analytical tools. Primarily, this was because of Tarrow's weak conceptualization of levels in the international system in relation to US power and because his processes follow a rather linear projection. Both of Tarrow's transitional and international processes build upon one another and assume that contentious collective action and claims originate in one location and expand to another, which was not the case for transnational AIDS treatment activism. Nevertheless, Tarrow's diffusion pathways and transnational coalition formation process did help explain and examine how Health GAP and TAC began coordinating their efforts. Thus, this analysis highlighted both strengths and limitations of Tarrow's processes, which require further theoretical development.

Lastly, the analysis chapter examined international political opportunities and threats, which Tarrow argues emerge from changes in the international political structure, or through the process of internationalization, and affect strategies, resources, and the effectiveness of transnational activism. It was explained in Chapter 2 that Tarrow's conceptualization of internationalization has received criticism,³³⁷ which this research took into account and is in agreement with because Tarrow does not demonstrate how shifts in the international political structure interact with transnational activism. To be specific, Tarrow does not account for how changes may also directly affect or influence the desired goal of transnational activism, as witnessed in this case study. As such, further theoretical development must be done to account for the interaction and contribution of structural shifts by international institutions on the desired goals and outcomes of transnational activism, not just on the collective action strategies of transnational activism.

In conclusion, this research finds Tarrow's theoretical argument about how transnational activism is a product of and has most impact on domestic politics to be most instructive for this case, as it helped guide the research and highlighted important factors needed to assess the causal relationship between transnational AIDS treatment activism and the reduction of ARV prices. Tarrow's transitional and

³³⁷ Boulding, "Book Review: Tarrow, S.(2005)," 886.

international processes for the coordination of transnational activism in more than one country were useful guides to explain how Health GAP and TAC began working together, however this research challenges Tarrow's linear projection, for claims did not extend from one area to another as they emerged at the same time in the US and South Africa. Lastly, this research finds Tarrow's conceptualization of internationalization problematic, as this case finds changes in the international political structure were in alignment with the goals of transnational AIDS treatment activism and were not merely opportunities. Therefore, as this research tested and challenged three of Tarrow's hypotheses in relation to its empirical findings, this dissertation makes a significant contribution to theory building in the field of transnational activism.

5.3 Areas for Future Research

The empirical findings and theoretical implications of this dissertation open up many avenues for future research of transnational AIDS treatment activism and for theory building in study of transnational activism. Furthermore, given the research design, methodology, and limitations of this dissertation there are also other ways to expand and improve upon the study presented here.

First, as mentioned in Chapter 1, this research was limited by time frame (1996 to 2001), geographical boundaries (US and South Africa), and the number of transnational activist organizations (Health GAP and TAC) it examined. Therefore, future research could include examination of more transnational AIDS treatment activist organizations from different countries, trace international structural shifts over a greater period of time, and expand specific outcomes or dependent variables of study. However, it is critical to avoid weaknesses found in the literature on the effectiveness of transnational AIDS treatment activism by employing a sound theoretical framework, and by avoiding oversimplification and deterministic tendencies.

Second, given that this dissertation tested and challenged Tarrow's theoretical arguments, future research may include conceptualizing levels of the international system with theories concerning power and US hegemony in international politics. In addition, future research could be done on tracing the impact of changes in the international political structure on transnational activist strategies and desired outcomes. Lastly, given that this research found that transnational AIDS treatment activism contributed in a limited way and was one contributing factor in a broad international political shift addressing AIDS, it would be instructive to review this case in terms of macro causal analysis by using either nominal, ordinal or narrative strategies.³³⁸

³³⁸ See Mahoney, "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis," for possible research designs and methodologies.

In conclusion, this dissertation has highlighted specific and limited ways in which transnational AIDS treatment activism contributed to the reduction of global ARV prices from September 2000 and February 2001, and has illustrated that the broader field of study on transnational activism is in need of further theory development. As such, this dissertation has contributed to the current body of literature in two ways, by clarifying the suggested causal relationship between transnational AIDS treatment activism and the reduction of global ARV prices for improving AIDS treatment access and by illuminating the areas in which current transnational activism theory is weak. Therefore, future research may build upon this dissertation to provide further empirical and theoretical clarity on the effectiveness of transnational activism, which this research finds to be lacking in current literature and of increasing relevance given the current difficulty explaining the 2011 Arab Spring and global Occupy Wall Street protests.³³⁹

University of Cape Town

³³⁹ Joseph Stiglitz, "The Globalization of Protest," *Project Syndicate*, November 4, 2011, accessed December 15, 2011, <http://www.project-syndicate.org/commentary/stiglitz144/English>.

Bibliography

- "About the Treatment Action Campaign." TAC. Accessed September 20, 2011.
<http://www.tac.org.za/community/about>.
- "ACT UP Capsule History 1992." ACT UP New York. Accessed November 14, 2011.
<http://www.actupny.org/documents/cron-92.html>.
- "Actions and Zaps." ACT UP New York. Accessed November 14, 2011.
<http://www.actupny.org/documents/newmem2.html>.
- Ballard, Richard, Adam Habib, Imraan Valodia, and Elke Zuern. "Introduction - from Anti-Apartheid to Post-Apartheid Social Movements." In *Voices of Protest*, edited by Richard Ballard, Adam Habib and Imraan Valodia, 1-22. KwaZulu-Natal: University of KwaZulu-Natal Press, 2006.
- Bello, Judith, and Alan Holmer. "'Special 301': Its Requirements, Implementation, and Significance." *Fordham International Law Journal* 13, no. 3 (1989): 259-275.
- Berkman, Alan. "The Global AIDS Crisis: Human Rights, International Pharmaceutical Markets and Intellectual Property." *Connecticut Journal of International Law* 17, no. 2 (2002): 149-155.
- Boulding, Carew. "Book Review: Tarrow, S. (2005). *The New Transnational Activism*. Cambridge, UK: Cambridge University Press." *Comparative Political Studies* 41, no. 6 (2008): 883-886.
- Bruntland, Gro Harlem. "Remarks at the WHO/Public Interest NGO Pharmaceuticals Roundtable, 3rd Meeting." May 1, 2000. Accessed August 17, 2011. http://www.who.int/director-general/speeches/2000/20000501_ngo.html.
- Butler, Anthony. "South Africa's HIV/AIDS Policy, 1994-2004: How can it be explained?" *African Affairs* 104, no. 417 (2005): 591-614.
- Callan, Michael, and Dan Turner. "A History of the People with AIDS Self-Empowerment Movement," *Body Positive Magazine*, December 1997. Accessed November 14, 2011.
<http://www.thebody.com/content/art31074.html>.
- Cleary, Susan, and Don Ross. "The 1998-2001 Legal Struggle between the South African Government and the International Pharmaceutical Industry: A Game-Theoretic Analysis." *Journal of Social, Political, and Economic Studies*. 27, no. 4 (2002): 445-494.
- Collins, Huntly. "World's AIDS Fighters Plot New Strategy." *Philadelphia Inquirer*, July 11, 2000. Accessed September 22, 2011. <http://search.ebscohost.com/login.aspx?direct=true&db=nfh&AN=KRBN200019301998&site=ehost-live>.
- Copson, Raymond. "CRS Issue Brief for Congress - AIDS in Africa." Washington D.C.: Congressional Research Service, January 5, 2005. Accessed October 28, 2011.
<http://www.fas.org/man/crs/IB10050.pdf>. 1-16.
- DiMasi, Joseph, Ronald Hansen, Henry Grabowski, and Louis Lasagna. "Cost of Innovation in the Pharmaceutical Industry." *Journal of Health Economics* 10, no. 2 (1991): 107-142.
- Edelman, Marc. "Social Movements: Changing Paradigms and Forms of Politics." *Annual Review of Anthropology* 30, (2001): 285-317.

Epstein, Steven. *Impure Science: AIDS, Activism, and the Politics of Knowledge*. Berkeley: University of California Press, 1996.

Fighting for our Lives: The History of the Treatment Action Campaign 1998- 2010. Cape Town: Treatment Action Campaign, 2010. Accessed September 10, 2011.
<http://www.tac.org.za/community/files/10yearbook/files/tac%2010%20year%20draft5.pdf>.

Friedman, Steven, and Shauna Mottiar. "A Rewarding Engagement? The Treatment Action Campaign and the Politics of HIV/AIDS." *Politics Society* 33, no. 4 (2005): 511-565.

———. "Seeking the High Ground - The Treatment Action Campaign and the Politics of Morality." In *Voices of Protest*, edited by Richard Ballard, Adam Habib and Imraan Valodia, 23-44. KwaZulu-Natal: University of KwaZulu-Natal Press, 2006.

Geffen, Nathan. *Debunking Delusions: The Inside Story of the Treatment Action Campaign*. Auckland Park: Jacana Media, 2010.

Gellman, Barton. "A Turning Point that Left Millions Behind ; Drug Discounts Benefit Few while Protecting Pharmaceutical Companies' Profits." *The Washington Post*, December 28, 2000. Accessed November 7, 2011.
<http://www.lexisnexis.com/hottopics/lnacademic/?verb=sr&csi=8075&sr=HLEAD%28A+turning+point+that+left+millions+behind%29%2BAND%2BDATE%2BIS%2B2000>.

George, Alexander, and Andrew Bennett. *Case Studies and Theory Development in the Social Sciences*. Cambridge, Mass.: MIT Press, 2004.

Goodwin, Jeff, James Jasper, and Jaswinder Khattri. "Caught in a Winding, Snarling Vine: The Structural Bias of Political Process Theory." *Sociological Forum* 14, no. 1 (Mar. 1999): 27-54.

Goozner, Merrill. "Third World Battles for Aids Drugs - U.S. Firms Oppose Generic Licensing." *Chicago Tribune*, April 28, 1999. Accessed October 24, 2011. http://articles.chicagotribune.com/1999-04-28/news/9904280067_1_compulsory-licensing-south-africa-aids-drugs.

Grebe, Eduard. "Networks of Influence: A Theoretical Review and Proposed Approach to AIDS Treatment Activism." University of Cape Town Center for Social Science Research - AIDS and Society Research Unit Working Paper, working paper 218, July 2008. Accessed May 25, 2011.
<http://www.cssr.uct.ac.za/sites/cssr.uct.ac.za/files/pubs/WP218.pdf>.

———. "The Treatment Action Campaign's Struggle for AIDS Treatment in South Africa: Coalition-Building through Networks." *Journal of Southern African Studies* 37, no. 4 (2011): 849-868.

Halbert, Debora. *Resisting Intellectual Property*. New York: Routledge, 2005.

"Health GAP Timeline: Direct Actions, Advocacy, and Response January 1999 - December 2000." Health GAP. Accessed October 27, 2011. http://www.healthgap.org/hgap/Timeline_of_HGAP_1999-2000.html.

"History of TAC (1998-2001)." TAC. Accessed September 22, 2011.
<http://www.tac.org.za/Documents/Other/tachist.pdf>.

Hunt, Scott, Robert Benford, and David Snow. "Identity Fields: Framing Processes and the Social Construction of Movement Identities." In *New Social Movements - from Ideology to Identity*,

- edited by Enrique Larana, Hank Johnston and Joseph Gusfield, 185-208. Philadelphia, Pa.: Temple University Press, 1994.
- "Intensifying Action Against HIV/AIDS in Africa: Responding to a Development Crisis." The World Bank, 1999. Accessed November 3, 2011. <http://siteresources.worldbank.org/AFRICAEXT/Resources/aidstrat.pdf>. 1-78.
- Johnson, Krista. "AIDS and the Politics of Rights in South Africa: A Contested Terrain." *Human Rights Review* 7, no. 2 (Jan. 2006): 115-129.
- Joni, Jennifer. "Access to Treatment for HIV/AIDS: A Human Rights Issue in the Developing World." *Connecticut Journal of International Law* 17, no. 2 (2002): 273-280.
- Keck, Margaret, and Kathryn Sikkink. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca, NY: Cornell University Press, 1998.
- Knight, Lindsay. *UNAIDS: The First Ten Years 1996-2006*. Geneva: Joint United Nations Program on HIV/AIDS, 2008. Accessed September 1, 2011. http://data.unaids.org/pub/Report/2008/JC1579_First_10_years_en.pdf.
- Levi, Margaret, and Gillian Murphy. "Coalitions of Contention: The Case of the WTO Protests in Seattle." *Political Studies* 54, no. 4 (2006): 651-670.
- Love, James. "Appendix B: Time-Line of Disputes Over Compulsory Licensing and Parallel Importation in South Africa." Consumer Project on Technology. Accessed October 23, 2011. <http://www.cptech.org/ip/health/sa/sa-timeline.txt>.
- Mahoney, James. "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis." *American Journal of Sociology* 104, no. 4 (Jan. 1999): 1154-1196.
- Mbali, Mandisa. "The Treatment Action Campaign and the History of Rights-Based, Patient-Driven HIV/AIDS Activism in South Africa." University of KwaZulu-Natal Center for Civil Society Report, report 29, 2005. Accessed September 13, 2011. www.nu.ac.za/Ccs/files/RReport_29.pdf.
- McAdam, Doug, John McCarthy, and Charles Tilly. *Dynamics of Contention*. Cambridge, UK: Cambridge University Press, 2001.
- Medicines and Related Substances Control Amendment Act*. No. 90. (December 12, 1997). Accessed August 20, 2011. <http://www.info.gov.za/view/DownloadFileAction?id=70836>.
- Nattrass, Nicoli. *Mortal Combat: AIDS Denialism and the Struggle for Antiretrovirals in South Africa*. KwaZulu-Natal: University of KwaZulu-Natal Press, 2007.
- . *The Moral Economy of AIDS in South Africa*. Cambridge, UK: Cambridge University Press, 2004.
- . "The (Political) Economics of Antiretroviral Treatment in Developing Countries." *Trends in Microbiology* 16, no. 12 (2008): 574-579.
- Office of United States Trade Representative. "USTR Announces Results of Special 301 Annual Review." Washington D.C.: Office of United States Trade Representative, 1998. Accessed November 11, 2011. http://keionline.org/sites/default/files/ustr_special301_1998.pdf. 1-33.

- Olesen, Thomas. "In the Court of Public Opinion': Transnational Problem Construction in the HIV/AIDS Medicine Access Campaign, 1998-2001." *International Sociology* 21, no. 1 (2006): 5-30.
- Orenstein, Mitchell, and Hans Schmitz. "The New Transnationalism and Comparative Politics." *Comparative Politics* 38, no. 4 (Jul. 2006): 479-500.
- Ostergard, Robert. "The Political Economy of the South Africa - United States Patent Dispute." *The Journal of World Intellectual Property* 2, no. 6 (1999): 875-888.
- Paige, Jeffrey. "Conjuncture, Comparison, and Conditional Theory in Macrosocial Inquiry." *American Journal of Sociology* 105, no. 3 (Nov. 1999): 781-800.
- Perez-Cases, Carmen, Cecile Mace, Daniel Berman, and Julia Double. "Accessing ARVs: Untangling the Web of Price Reductions for Developing Countries." Médecins Sans Frontières, 2001. Accessed October 3, 2011. utw.msfaccess.org/downloads/311. 1-16.
- Pierce, Roger. *Research Methods in Politics: A Practical Guide*. Los Angeles: Sage, 2008.
- Putnam, Robert. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42, no. 3 (1988): 427-460.
- Richter, James. "The New Transnational Activism – Sidney Tarrow." *Journal of Politics* 69, no. 1 (2007): 256-257.
- Richwine, Lisa. "Gore Worked to Soften South Africa Health Law." *Reuters New Media*, April 16, 1999. Accessed October 24, 2011. <http://www.aegis.com/news/re/1999/re990404.html>.
- Robins, Steven. "'Long Live Zackie, Long Live': AIDS Activism, Science and Citizenship After Apartheid." *Journal of Southern African Studies* 30, no. 3 (Sep. 2004): 651-672.
- Robins, Steven, and Bettina von Lieres. "Remaking Citizenship, Unmaking Marginalization: The Treatment Action Campaign in Post-Apartheid South Africa." *Canadian Journal of African Studies* 38, no. 3 (2004): 575-586.
- Rodrigues, Maria. "Rethinking the Impact of Transnational Advocacy Networks." *New Global Studies* 5, no. 2 (2011): 1-21.
- Sawyer, Eric. "An ACT UP Founder "Acts Up" for Africa's Access to AIDS." In *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization*, edited by Benjamin Shepard and Ronald Hayduk, 88-102. London: Verso, 2002.
- Schug, Mary. "Promoting Access to HIV/AIDS Pharmaceuticals in Sub-Saharan Africa within the Framework of International Intellectual Property Law." *Law and Inequality: A Journal of Theory and Practice* 19, (2001): 229-257.
- Seckinelgin, Hakan. *International Politics of HIV/AIDS: Global Disease-Local Pain*. New York: Routledge, 2008.
- Sell, Susan, and Aseem Prakash. "Using Ideas Strategically: The Contest between Business and NGO Networks in Intellectual Property Rights." *International Studies Quarterly* 48, no. 1 (Mar. 2004): 143-175.

- Shadlen, Kenneth. "The Political Economy of AIDS Treatment: Intellectual Property and the Transformation of Generic Supply." *International Studies Quarterly* 51, no. 3 (2007): 559-581.
- Shepard, Benjamin. "Introductory Notes on the Trail from ACT UP to the WTO." In *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization*, edited by Benjamin Shepard and Ronald Hayduk, 11-16. London: Verso, 2002.
- Shepard, Benjamin, and Ronald Hayduk, eds. *From ACT UP to the WTO: Urban Protest and Community Building in the Era of Globalization*. London: Verso, 2002.
- Skocpol, Theda. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge, UK: Cambridge University Press, 1979.
- Smith, Jackie. "Book Review: The New Transnational Activism." *The Canadian Journal of Sociology* 31, no. 4 (2006): 546-549.
- Smith, Raymond, and Patricia Siplon. *Drugs into Bodies: Global AIDS Treatment Activism*. Westport, Conn.: Praeger, 2006.
- Snow, David, and Robert Benford. "Master Frames and Cycles of Protest." In *Frontiers in Social Movement Theory*, edited by Aldon Morris and Carol Mueller, 133-155. New Haven, Conn.: Yale University Press, 1992.
- Stiglitz, Joseph. *Making Globalization Work*. London: Allen Lane, 2006.
- . "The Globalization of Protest." *Project Syndicate*, November 4, 2011. Accessed December 15, 2011. <http://www.project-syndicate.org/commentary/stiglitz144/English>.
- Tarrow, Sidney. *Power in Movement: Social Movements and Contentious Politics*. 2nd ed. Cambridge, UK: Cambridge University Press, 1998.
- . *The New Transnational Activism*. Cambridge, UK: Cambridge University Press, 2005.
- Thompson, Simon. "Mapping Transnational Activism." *Democratija* 6, (2006): 115-119.
- Tilly, Charles. *The Contentious French*. Cambridge, Mass.: Belknap Press, 1986.
- Vandoren, Paul and Jean Van Eeckhaute. "The WTO Decision on Paragraph 6 of the Doha Declaration on the TRIPS Agreement and Public Health." *The Journal of World Intellectual Property* 6, (2003): 779-793.
- Yin, Robert. *Case Study Research: Design and Methods*. Applied Social Research Methods Series. 4th ed. Vol. 5. Los Angeles: Sage Publications, 2009.

Target Tracking using Multisensor Data Fusion for an Unmanned Aerial Vehicle Sense and Avoid System

by

Shuaib Omar

BSc (Eng) Electrical and Computer Engineering

A thesis submitted to the Department of Electrical Engineering,
University of Cape Town, in fulfilment of the requirements
for the degree of

Master of Science

at the

UNIVERSITY OF CAPE TOWN

Supervisor:

Dr. Simon Winberg



© University of Cape Town

June 2012

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in Engineering in the University of Cape Town. It has not been submitted before for any degree or examination in any other university.

Signature of Author:

Cape Town

29 May 2012

University of Cape Town

Abstract

The last decade has seen a proliferation of unmanned aerial vehicles (UAV) put to use in a myriad of applications. However, the true potential of UAVs lies in its timely adoption to civil airspace. For this to occur, UAVs need to exhibit increased levels of autonomy. The principal impediment in achieving this autonomy is the UAV's capability to sense and avoid. The aim of this dissertation is to develop an airborne threat tracking system using multisensor data fusion for a UAV sense and avoid system, termed the threat tracking unit (TTU).

This thesis entails an extensive study of relevant literature in the fields of target tracking, estimation and multisensor data fusion. Following this, a simulation environment was developed to test and qualify various iterations of the TTU.

Initial TTU designs entailed the design of single sensor tracking filters in order set a baseline with which the performance of multisensor tracking filters can be compared against. Subsequent to this, multisensor fusion architectures were investigated drawing on the shortcomings of their single sensor counterparts.

A centralised fusion architecture (FA1) demonstrated superior performance in both position and velocity tracking. However, this architecture possesses a single point of failure in flight-critical situations. Comparatively, a distributed fusion architecture (FA2) displayed adequate performance. This architecture, nevertheless, is characterised by increased robustness and lower computational load.

Finally, in order to further reduce computational complexity, a full dynamic feedback fusion architecture (FA3) was developed. Although FA3 displayed reduced tracking accuracy, feasible results were still obtained.

University of Cape Town

University of Cape Town

For Firdows, my dearest wife.

Acknowledgements

I would like to express my deepest gratitude to **The All Mighty** for granting me the strength to complete this massive task.

Thank you Dr. Simon Winberg for your constant support and supervision throughout this journey.

I am grateful to my parents and siblings for always encouraging me and wishing the best for me.

I acknowledge Tellumat for providing me with the financial aid to undertake this study. I would especially like to thank Chris Williams, Simon Norval and Johann Gerber. My gratitude goes to Eugene Julie for always being keen to bounce ideas off and provide me with valuable information.

Amir Patel, my studies and research would never have been the same without you working by my side. It has surely been an adventure. In reply to you I shall say, "Cheers!"

I would also like to thank all my friends.

Last, but most importantly, I am especially indebted to my wife, Firdows Omar. Without her this dissertation would not exist; there were times when I lost my way and hope and she pulled me back on the straight path. For that I am truly grateful. Thank you for your unwavering support and patience.

Nomenclature

ANEES	Average Normalised Estimation Error Squared
ATC	Air Traffic Control
BOT	Bearings-only Tracking
BRF	Body Reference Frame
CA	Collision Avoidance
DCM	Direction Cosines Matrix
EKF	Extended Kalman Filter
EO	Electro-Optic
FA	Fusion Architecture
FAA	Federal Aviation Administration
FC	Fusion Centre
FLIR	Forward Looking Infra-Red
FOR	Field of Regard
FOV	Field of View
IFR	Instrument Flight Rules
IRST	Infra-Red Search and Track
KF	Kalman Filter
LFM	Linear Frequency Modulation
LSE	Least Squares Estimation
MCS	Monte Carlo Simulations
MSDF	Multisensor Data Fusion
NED	North-East-Down
NMAC	Near Mid Air Collision
PF	Particle Filter
PLKF	Pseudo Linear Kalman Filter

RMSE	Root Mean Squared Error
SAA	Sense and Avoid
SAS	Sense and Avoid System
SS	Self Separation
TTU	Threat Tracking Unit
UAS	Unmanned Aerial System
UAV	Unmanned Aerial Vehicle
UKF	Unscented Kalman Filter
VFR	Visual Flight Rules

University of Cape Town

Table of Contents

CHAPTER 1: INTRODUCTION	16
1.1 BACKGROUND	16
1.2 STANDARDS FOR A SENSE AND AVOID SYSTEM	18
1.3 PROBLEM DESCRIPTION	27
1.4 SCOPE AND OBJECTIVES.....	28
1.5 THESIS OUTLINE.....	31
CHAPTER 2: LITERATURE REVIEW.....	33
2.1 MULTISENSOR DATA FUSION.....	34
2.1.1 An Introduction.....	34
2.1.2 Overview of Estimation Theory and Tracking Filters.....	46
2.2 TARGET TRACKING.....	54
2.2.1 Selecting a Tracking Coordinate System.....	56
2.2.2 Passive Bearings-Only Tracking.....	60
2.2.3 Tracking with Doppler Information	64
2.2.4 Tracking in the Presence of Multiple Sensors.....	73
2.3 SUMMARY	79
CHAPTER 3: METHODOLOGY.....	81
3.1 OVERVIEW OF RESEARCH STRATEGY.....	81
3.2 RESEARCH AND REQUIREMENTS DEFINITION.....	83
3.3 MODELLING AND SIMULATION.....	84
3.4 ALGORITHM DESIGN	84
3.5 TESTING THE THREAT TRACKING UNIT AND PERFORMANCE METRICS.....	85
3.5.1 Visual Evaluation	86
3.5.2 Filter Accuracy.....	86
3.5.3 Filter Credibility.....	87
3.5.4 Filter Consistency	87
CHAPTER 4: SIMULATION ENVIRONMENT	90
4.1 STATES OF INTEREST	91

4.2	UAV AND THREAT MODELS	92
4.3	FLIGHT SCENARIO	95
CHAPTER 5: TTU DESIGN		101
5.1	INTRODUCTION	101
5.2	COMMON FILTER PARAMETERS.....	103
5.3	EO CAMERA TRACKING FILTER.....	105
5.3.1	Results	109
5.3.2	Monte Carlo Simulations	111
5.3.3	Summary	112
5.4	RANGING RADAR TRACKING FILTER.....	112
5.4.1	Nonlinear Approach	115
5.4.2	Linear Approach	120
5.4.3	Summary	125
5.5	DOPPLER RADAR TRACKING FILTER	126
5.5.1	EKF Under Ideal Conditions	128
5.5.2	EKF with Correlated Measurement Errors.....	133
5.5.3	Summary	138
5.6	FUSION FILTER.....	138
5.6.1	Synchronising Sensor Sampling Periods	139
5.6.2	FA1	140
5.6.3	FA2	144
5.6.4	FA3	151
5.6.5	Summary	160
CHAPTER 6: DISCUSSIONS		161
6.1	SINGLE SENSOR TRACKING FILTERS	161
6.1.1	EO Camera.....	161
6.1.2	Radar	163
6.1.3	Single Sensor Summary.....	168
6.2	MULTISENSOR TRACKING FILTERS.....	170
6.2.1	FA1	170
6.2.2	FA2	171

6.2.3	FA3	173
6.3	COMPARISON OF FUSION ARCHITECTURES.....	175
6.4	SUMMARY	178
CHAPTER 7: CONCLUSIONS		179
7.1	RESEARCH OBJECTIVES: CONCLUSIONS	179
7.2	RECOMMENDATIONS FOR FUTURE WORK	181
REFERENCES.....		190

University of Cape Town

List of Figures

Figure 1.1: Predator UAV providing surveillance for aircraft carrier at sea [3].	17
Figure 1.2: High-level view of a sense and avoid system.	18
Figure 1.3: ADS-B in operation [7].	20
Figure 1.4: Layered relationship between cooperative and non-cooperative means of avoiding collisions [8]. CA, at the bottom, has the highest risk of collision as it is the UAV's last line of defence.	21
Figure 1.5: The eight SAA sub-functions.	22
Figure 1.6: Visual representation of the Collision Volume as well as the SS and CA Thresholds [2]. A UAV is said to be well clear of an intruder aircraft if the intruder is outside of the CA Threshold.	25
Figure 1.7: The FOR and FOV of a camera sensor with three allowable mechanical positions. Note FOV2 and FOV3 share -15° to 0° and 0° to 15° with FOV1, respectively.	26
Figure 1.8: Fusion of multiple sensor images. From left to right: a CT image; an MRI image; and the, more informative, fused image [10].	27
Figure 1.9: Thesis scope.	29
Figure 1.10: System overview. This thesis is concerned with the development of the TTU block. It is assumed that IMU and Actuator blocks exist inside the UAV picture.	30
Figure 2.1: Mapping from thesis objectives to fields of study and, consequently, headings in the Literature Review. On the right, the mappings of other objectives are illustrated for the sake of completeness.	34
Figure 2.2: Fighter aircraft with radar and IR sensors. MSDF is the enabling technology to fully reap the benefits of having multiple, redundant sensors as part of a system [11].	36
Figure 2.3: Cloud cover around the globe [23].	38
Figure 2.4: Centralised fusion architecture.	40
Figure 2.5: Distributed fusion architecture.	42
Figure 2.6: An example distributed fusion architecture application.	43
Figure 2.7: Single step Kalman filter overview obtained from page 219 of [39].	49
Figure 2.8: Airborne radar in action [11].	55
Figure 2.9: Ground-based radar target tracking [36].	56
Figure 2.10: Body axis coordinate system. Also known as the body reference frame (BRF) [49].	57
Figure 2.11: Relationship between BRF and a fixed earth frame, e.g. the NED system. [49]	58
Figure 2.12: Bearings-only problem geometry [62].	61
Figure 2.13: Target motion along observer LOS [62]. Angle theta remains constant in this scenario.	61
Figure 2.14: Using multiple UAVs to determine target position [62].	62
Figure 2.15: Pinhole camera model [63].	63
Figure 2.16: Projection of problem geometry in 2D.	64
Figure 2.17: Frequency modulated continuous wave (FMCW) signal transmitted by a radar. The range rate is determined by the frequency difference in transmitted and reflected signals.	65
Figure 2.18: Qingchao and Wenfei proposed fusion architecture [29].	76
Figure 2.19: Sensor fusion logical architecture implemented in [48]. Centralised fusion consisting of radar, IR and EO sensors.	78
Figure 2.20: EO camera target detection algorithm [48].	79
Figure 3.1: MBD approach.	83
Figure 4.1: Simulation environment to be modelled.	90
Figure 4.2: Euler angles illustration [79].	94
Figure 4.3: Inputs to the TTU.	96
Figure 4.4: UAV and Threat flight path shown in the NE plane.	97
Figure 4.5: Initial simulation setup procedure.	99
Figure 5.1: Relative range of flight scenario.	105
Figure 5.2: Relative azimuth in NED - truth and camera measured values.	107
Figure 5.3: Relative elevation in NED - truth and camera measured values.	107
Figure 5.4: Bearings-only tracking with an EKF.	108
Figure 5.5: Camera filter - estimated state error covariance of N, E and D components.	109
Figure 5.6: Camera filter - True and estimated threat position in N, E and D directions.	110
Figure 5.7: Camera filter - True and estimated threat velocity in N, E and D directions.	110

Figure 5.8: Camera filter - Estimation error in N, E and D directions.	111
Figure 5.9: Camera only filter RMSE in position and velocity.	111
Figure 5.10: Camera only filter ANEES.	112
Figure 5.11: Relative range in NED - truth and radar measured values.	114
Figure 5.12: Relative azimuth in NED - truth and radar measured values.	114
Figure 5.13: Relative elevation in NED - truth and radar measured values.	115
Figure 5.14: Nonlinear ranging radar tracking filter.	115
Figure 5.15: Radar nonlinear filter (EKF) – Estimated state error covariance of N, E and D position components.	116
Figure 5.16: Radar nonlinear filter (EKF) - True, estimated and measured relative motion range, azimuth and elevation in the BRF.	117
Figure 5.17: Radar nonlinear filter (EKF) – True and estimated threat position in N, E and D directions.	117
Figure 5.18: Radar nonlinear filter (EKF) – True and estimated threat velocity in N, E and D directions.	118
Figure 5.19: Radar nonlinear filter (EKF) – Position estimation error in N, E and D directions in a single flight scenario.	118
Figure 5.20: Radar nonlinear filter with no range rate measurements - RMSE in position and velocity.	119
Figure 5.21: Radar filter with no range rate measurements - ANEES.	119
Figure 5.22: Second range radar filter developed. A linear solution which involves measurement conversion to Cartesian coordinates.	120
Figure 5.23: Radar linear KF tracking in Cartesian coordinates – estimated state error covariance.	122
Figure 5.24: Radar linear KF tracking in Cartesian coordinates – True, estimated and measured relative range, azimuth and elevation in BRF.	122
Figure 5.25: Radar linear KF tracking in Cartesian coordinates – True and estimated position in N, E and D directions.	123
Figure 5.26: Radar linear KF tracking in Cartesian coordinates – True and estimated velocity in N, E and D directions.	123
Figure 5.27: Radar linear KF tracking in Cartesian coordinates – Estimation error in N, E and D directions in a single flight scenario.	124
Figure 5.28: Radar linear filter (KF) tracking with Cartesian coordinates - True and estimated threat position with no measurement covariance matrix transformation.	124
Figure 5.29: Radar linear filter tracking in Cartesian coordinates with no range rate measurements - RMSE in position and velocity.	125
Figure 5.30: Radar linear filter tracking in Cartesian coordinates with no range rate measurements - ANEES.	125
Figure 5.31: Relative range rate in NED - truth and radar measured values.	128
Figure 5.32: EKF with range rate measurements, assuming no correlation.	129
Figure 5.33: Radar EKF with range rate measurements – estimated state error covariance.	130
Figure 5.34: Radar EKF with range rate measurements – true and estimated threat position in N, E and D directions.	130
Figure 5.35: Radar EKF with range rate measurements – true and estimated threat velocity in N, E and D directions.	131
Figure 5.36: Radar EKF with range rate measurements – estimation error in N, E and D directions in a single flight scenario.	131
Figure 5.37: Radar EKF with range rate measurements – true, estimated and measured relative range, azimuth, elevation and range rate with respect to the BRF.	132
Figure 5.38: Radar EKF with range rate measurements – RMSE in position and velocity.	132
Figure 5.39: Radar EKF with range rate measurements - ANEES.	133
Figure 5.40: EKF with range rate measurements, assuming correlation exists. The EKF takes the range, azimuth, elevation and pseudo measurements as inputs. The pseudo measurement is formed using the range and range rate measurements.	134
Figure 5.41: Radar EKF with range rate measurements and correlation – estimated state error covariance.	135
Figure 5.42: Radar EKF with range rate measurements and correlation – true and estimated threat position in N, E and D directions.	135

Figure 5.43: Radar EKF with range rate measurements and correlation – true and estimated threat velocity in N, E and D directions.	136
Figure 5.44: Radar EKF with range rate measurements and correlation – estimation error in N, E and D directions in a single flight scenario.	136
Figure 5.45: Radar EKF with range rate measurements and correlation – RMSE in position and velocity.	137
Figure 5.46: Radar EKF with range rate measurements and correlation – ANEES.	137
Figure 5.47: Plot of radar and camera azimuth measurements illustrating their contrasting sampling periods. The camera provides measurements four times faster than the radar.	139
Figure 5.48: Fusion Architecture 3. Measurement fusion is performed in this centralised approach.	140
Figure 5.49: FA1 – state error covariance in N, E and D directions.	141
Figure 5.50: FA1 – true and estimated threat position in N, E and D directions.	141
Figure 5.51: FA1 – true and estimated threat velocity in N, E and D directions.	142
Figure 5.52: FA1 – estimation error in N, E and D directions in a single flight scenario.	142
Figure 5.53: FA1 – velocity estimation error in N, E and D directions in a single flight scenario.	143
Figure 5.54: FA1 centralised fusion - RMSE in position and velocity.	143
Figure 5.55: FA1 centralised fusion - ANEES.	144
Figure 5.56: Fusion Architecture 2. Both radar and camera local filter are EKFs.	144
Figure 5.57: FA2 - radar, camera and fused error covariances in N, E and D directions.	146
Figure 5.58: FA2 – true and estimated threat position in N, E and D directions.	146
Figure 5.59: FA2 – true and estimated threat velocity in N, E and D directions.	147
Figure 5.60: FA2 - radar, camera and fused position estimation errors in the N direction.	147
Figure 5.61: FA2 - radar, camera and fused position estimation errors in the E direction.	148
Figure 5.62: FA2 - radar, camera and fused position estimation errors in the D direction.	148
Figure 5.63: FA2 - radar, camera and fused velocity estimation errors in the N direction.	149
Figure 5.64: FA2 - radar, camera and fused velocity estimation errors in the E direction.	149
Figure 5.65: FA2 - radar, camera and fused velocity estimation errors in the D direction.	150
Figure 5.66: FA2 - RMSE in position and velocity.	150
Figure 5.67: FA2 - ANEES.	151
Figure 5.68: Fusion Architecture 3. It consists of a linear KF, an EKF and the Fusion Centre.	151
Figure 5.69: FA3 - radar, camera and fused error covariances in N, E and D directions.	155
Figure 5.70: FA3 – true and estimated threat position in N, E and D directions.	155
Figure 5.71: FA3 – true and estimated threat velocity in N, E and D directions.	156
Figure 5.72: FA3 - radar, camera and fused position estimation errors in the N direction.	156
Figure 5.73: FA3 - radar, camera and fused position estimation errors in the E direction.	157
Figure 5.74: FA3 - radar, camera and fused position estimation errors in the D direction.	157
Figure 5.75: FA3 - radar, camera and fused velocity estimation errors in the N direction.	158
Figure 5.76: FA3 - radar, camera and fused velocity estimation errors in the E direction.	158
Figure 5.77: FA3 - radar, camera and fused velocity estimation errors in the D direction.	159
Figure 5.78: FA3 - RMSE in position and velocity.	159
Figure 5.79: FA3 - ANEES.	160
Figure 6.1: Camera filter - True and estimated threat position in the NE plane.	162
Figure 6.2: Radar nonlinear filter (EKF) - True and estimated threat position in the NE plane.	164
Figure 6.3: Radar linear filter (KF) tracking with Cartesian coordinates - True and estimated threat position in the NE plane.	165
Figure 6.4: Radar EKF with range rate measurements - true and estimated threat position in the NE plane.	167
Figure 6.5: Radar EKF with range rate measurements and correlation - true and estimated threat position in the NE plane.	167
Figure 6.6: FA1 – true and estimated threat position in the NE plane.	171

List of Tables

Table 1.1: Field of Regard requirements for an on-board UAV sensor used by a sense and avoid system.....	25
Table 2.1: Advantages of MSDF.	39
Table 2.2: KF matrix dimensions.	50
Table 4.1: UAV and threat initial states.	96
Table 5.1: Designed tracking filters and the estimators on which they are built.	102
Table 5.2: EO camera measurement errors.	106
Table 5.3: Radar measurement errors.	113
Table 5.4: Doppler radar measurement errors.	128
Table 6.1: Single sensor tracking filter's RMSE in position and velocity and ANEES.....	169
Table 6.2: Performance comparison of FA3 and its two variants.....	175
Table 6.3: Measurement errors of Doppler Radar and EO Camera.	175
Table 6.4: Performance comparison of fusion architectures. Filter accuracies and credibility. The radar-only filter has been included in the table to easily compare the fusion with the single sensor performance.	176
Table 6.5: Position and Velocity RMSE when both sensor's angular measurement errors are the same and when the camera's are more accurate.	177

University of Cape Town

Chapter 1: Introduction

This dissertation aims to design an airborne tracking filter using multisensor data fusion. This filter is to be integrated into a tracking subsystem, termed the *threat tracking unit*, which forms part of a larger UAV sense and avoid system (SAS).

The remainder of this chapter concerns the research focuses, the research objectives and motivations for these.

1.1 Background

The ability to deploy unmanned aerial vehicles (UAVs) in a myriad of life-critical missions has seen its presence rapidly increase over the last decade [1]. Some of these UAV-suited applications are border patrol, battlefield surveillance (depicted in Figure 1.1), etc. Traditionally, UAV mission success would be dependent on some form of human involvement. Whether it is in direct remote-control (man-in-the-loop) [2] or assessment of data gathered; required human resources scale with mission criticality and number of UAVs available in said mission. Consider, for example, a military application where multiple UAVs are providing battlefield surveillance. This would require numerous operators to make inferences about the vast amounts of data arriving at the ground station. Therefore, in order to fully reap the advantages provided by UAVs, a need is to mitigate human involvement; gearing UAVs toward full autonomy.



Figure 1.1: Predator UAV providing surveillance for aircraft carrier at sea [3].

The principal impediment in achieving this full autonomy (and mitigation of human resources) is the UAV's capability to sense and avoid. This means that a UAV needs to gather information about its immediate surroundings (sense) and, based on this information, evaluate whether a potential threat [2] exists and determine the necessary action to be performed (avoid). This is analogous to the 'see and avoid' process used by pilots [4], when deciding whether an aircraft manoeuvre is required to maintain separation.

The possible applications of a fully autonomous UAV are vast. Therefore, the future sees a movement from the operation of UAVs in segregated airspace [5] to civil airspace; where UAVs collocate with other manned aircraft. This civil airspace integration presents many technical and operational challenges. The most important of which, again, is the UAV's sense and avoid (SAA) capabilities. Other integration challenges include: proper cohabitation with cooperative sensors such Air Traffic Control (ATC), TCAS and ADSB;

autonomous flight of UAVs must adhere to regulations under Visual Flight Rules (VFR) and Instrument Flight Rules (IFR); as well as other safety and security issues.

A UAV's sense and avoid capabilities can be divided into two high-level subsystems [2], as its name implies: sensory and avoidance; where the sensory component is responsible for threat detection and tracking. This high-level view is illustrated in Figure 1.2. The threats in this case, which to avoid, are other aircraft and obstacles in the UAVs field of view (FOV).

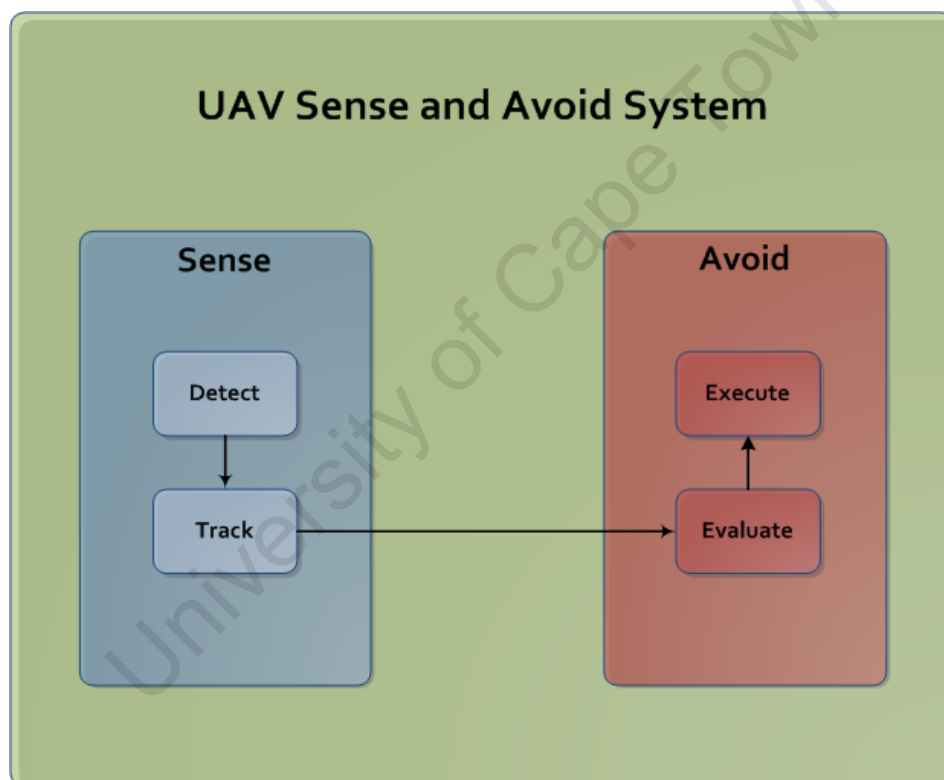


Figure 1.2: High-level view of a sense and avoid system.

1.2 Standards for a Sense and Avoid System

As an official set of standards for sense and avoid (SAA) systems did not exist, the Federal Aviation Administration (FAA) recognized the fact that this would become of paramount importance for the future of UAVs. They addressed this need, in 2008, by holding a series of workshops where experts

in collision avoidance (CA), airborne surveillance and UAV operation convened. There they discussed and documented the issues that become apparent when replacing the "see and avoid" ability of a pilot with that of the sense and avoid capability of a UAV. The workshops resulted in a clearly defined set of capabilities which a SAS should have in order to reliably, safely and autonomously avoid obstacles.

An SAA must exhibit an equivalent level of safety as that of manned aircraft. This must hold true in controlled¹ and uncontrolled airspace. The implication of this is that the SAS needs to communicate with ATC providers and/or other cooperative systems such as Traffic Alert and Collision Avoidance System (TCAS) II, Automatic Dependent Surveillance-Broadcast (ADS-B), etc. A cooperative aircraft is one that has a transponder² aboard and operating [2]. When collision avoidance is performed by an aircraft (manned or unmanned) that makes use of ATC management or one or more of the aforementioned systems, then that is termed cooperative collision avoidance.

TCAS II is a transponder based system which transmits its ownship³ position to other aircraft. It operates independently from ATC and also analyses the projected flight path of approaching aircraft and issues 'Resolution Advisories' to the pilot to resolve potential mid-air collisions [6]. TCAS is required internationally to be fitted in aircraft with more than 30 seats or weighing more than 15000 kg and is usually found on general aviation aircraft, but it only issues avoidance manoeuvres in the vertical plane.

In contrast to TCAS, ADS-B sees a move away from surveillance radar to a Global Navigation Satellite System (GNSS). An aircraft fitted with ADS-B is able to derive its precise position from the GNSS constellation, via a GNSS receiver [7]. This position is packaged along with other aircraft information such as speed, heading, altitude and flight number, and simultaneously transmitted to other ADS-B capable receivers as well as satellite

¹ Controlled airspace is that in which air traffic control (ATC) service is provided to flights.

² A transponder provides an electronic means of aircraft identification.

² A transponder provides an electronic means of aircraft identification.

³ The term "ownship" refers to "our own aircraft".

communications transceivers. The latter forwards the information on to ATC centres. Figure 1.3 illustrates this process.

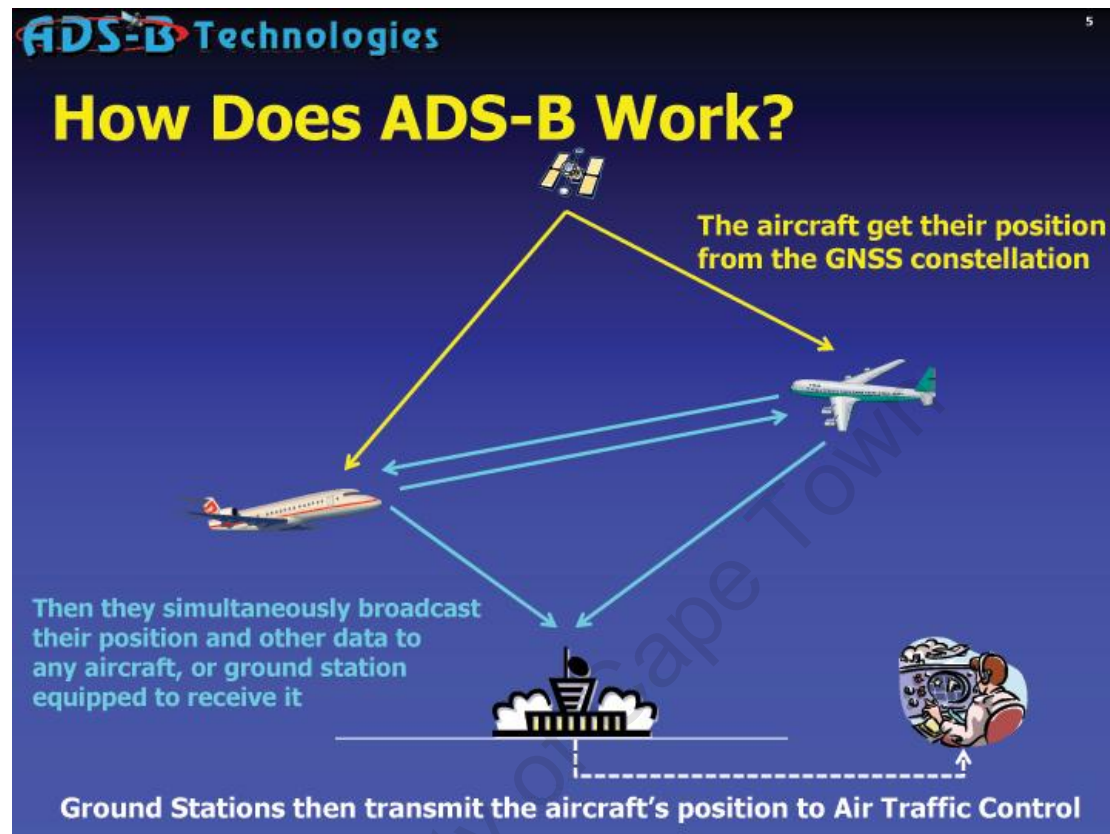


Figure 1.3: ADS-B in operation [7].

The FAA defines SAA as the capability of an Unmanned Aerial System⁴ (UAS) to remain well clear from and avoid collisions with other airborne traffic; this consequently results in the two main functions of a SAS: Self Separation (SS) and CA (see Figure 1.4). SS is the ability of the UAV to constantly remain well clear from other aircraft. It involves commanding small bank angles, whereas CA involves drastic last minute evasive manoeuvres. SS can be thought of as maintaining a distance so that a CA manoeuvre will not become necessary.

⁴ An Unmanned Aerial System is comprised of three components: a UAV; a wireless communications data link; and a ground station.

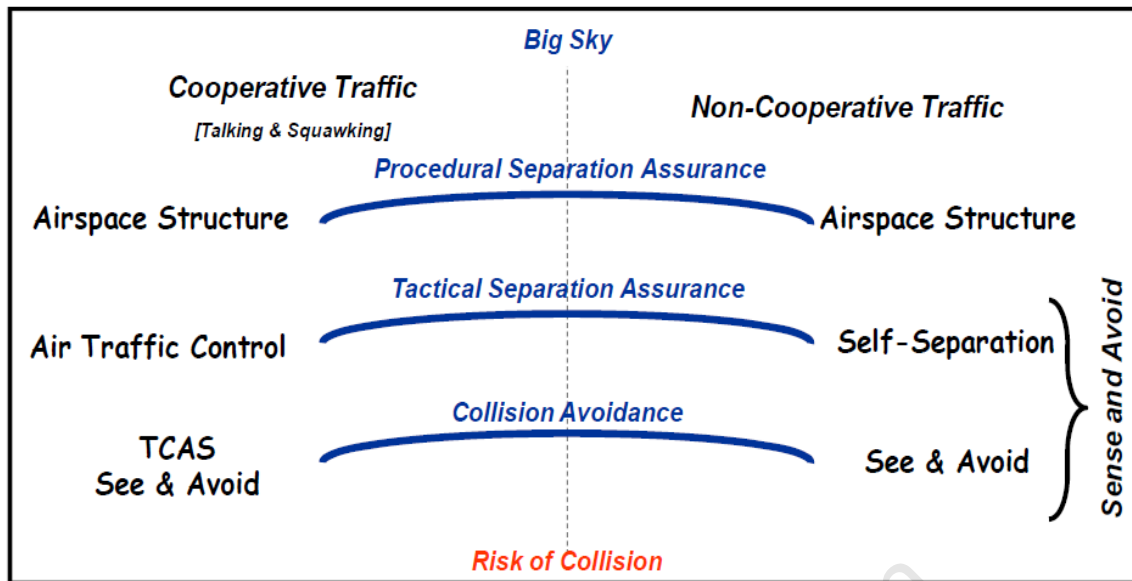


Figure 1.4: Layered relationship between cooperative and non-cooperative means of avoiding collisions [8]. CA, at the bottom, has the highest risk of collision as it is the UAV's last line of defence.

With reference to the aforementioned, a SAS is comprised of eight sub-functions:

- i. **Detect** – Determine presence of potential hazards.
- ii. **Track** – Estimated position and velocity (state) of a single intruder based on one or more surveillance reports.
- iii. **Evaluate** – Assess collision risk based on intruder and UAV states.
- iv. **Prioritize** – Determine which intruder tracks have met a collision risk threshold.
- v. **Declare** – Decide that action is needed.
- vi. **Determine Action** – Decide on what action is required.
- vii. **Command** – Communicate determined action.
- viii. **Execute** – Respond to the commanded action.



Figure 1.5: The eight SAA sub-functions.

Figure 1.5 shows the sub-functions on an encounter timeline. It brings to light the order of functions to be performed by a SAS from the moment of threat detection. Moreover, it shows the importance of limits (thresholds) in an encounter; here described as Minimum Detect Time and Minimum Response Time. These limits indicate that in order for a UAV to successfully evade a threat, headed on a collision course, there exists some fixed time that detection needs to occur. Furthermore, that the entire computational process of detection, tracking and evaluation as well as mechanical actuation of the determined avoidance manoeuvre needs to take place within a specified time limit⁵ (SAA Processing Time); this takes into account physical limitations and provides the UAV enough time with which to evade the target. Consider the following scenario; if any of the sub-functions take longer than their specified time (defined in the figure), the SAA Processing Time will be pushed

⁵ This last moment of action varies with different targets and is also dependent on ownship parameters.

back after the Minimum Response Time. The consequence of this, as one would assume, could lead to catastrophic results as the UAV will not be able to perform its avoidance manoeuvre in time to successfully evade the threat.

According to United States National Airspace System (NAS) regulations [2] a pilot must avoid collisions by remaining "well clear" from other aircraft. Translating this obscure term into an explicit specification for a SAS proved problematic for the team of experts; the term suggests that a pilot should use his subjective assessment when performing an evasive manoeuvre. Functionality that is clearly not able to be part of a sense and avoid system. The solution to this problem was to define a safety volume for the UAV as well as dynamic thresholds dependent on each aircraft being tracked in the UAV's FOV.

The volumes and thresholds that were defined are as follows:

- **Collision Volume.** A cylindrical volume around the UAV that travels with it and has a horizontal radius of 500 feet and a vertical height of 200 feet⁶. If an aircraft penetrates this volume then that is considered a Near Mid Air Collision (NMAC). An NMAC is potentially very dangerous, even if the two aircraft do not collide, as the UAV can be caught in the wake turbulence of the other aircraft and lose control.
- **Collision Avoidance Threshold.** Although illustrated in Figure 1.6 as a boundary, this is in actual fact a dynamic length of time that varies with for example, a threat's closure-rate (speed), distance and airspace class. It is the time from the moment a threat penetrates the Collision Avoidance Threshold to the moment of the two aircraft's closest point of approach (CPA) along their flight trajectories. The threshold is determined so that the UAV has enough time, once the threshold has been breached, to keep the threat out of its Collision Volume by performing an evasive manoeuvre.

⁶ That is 100 feet above and below the UAV.

- **Self Separation Threshold.** Similarly to the Collision Avoidance Threshold, the Self Separation Threshold is also a dynamic length of time that is affected by similar variables mentioned above. It is the time from the moment a threat penetrates the Self Separation Threshold to the moment of CPA along their flight trajectories. This threshold ensures that the UAV remains well clear from the threat aircraft, by performing a SS rather than a CA evasive manoeuvre. The term well clear is thus defined as a threat aircraft being outside of the Collision Avoidance Threshold.

The SS and CA thresholds, therefore, have the following effects, respectively:

- If a threat aircraft penetrates the SS Threshold, the UAV would perform an evasive SS manoeuvre (commanded by the SAS of course). By the definition of the threshold it allows the UAV enough time to perform this evasive manoeuvre while keeping well clear from the threat aircraft (outside of the Collision Avoidance Threshold).
- If a threat aircraft penetrates the CA Threshold, the UAV would perform an evasive CA manoeuvre. By the definition of the threshold it allows the UAV enough time to perform this last-minute evasive manoeuvre while preventing an NMAC (keeping the threat aircraft outside of the Collision Volume).

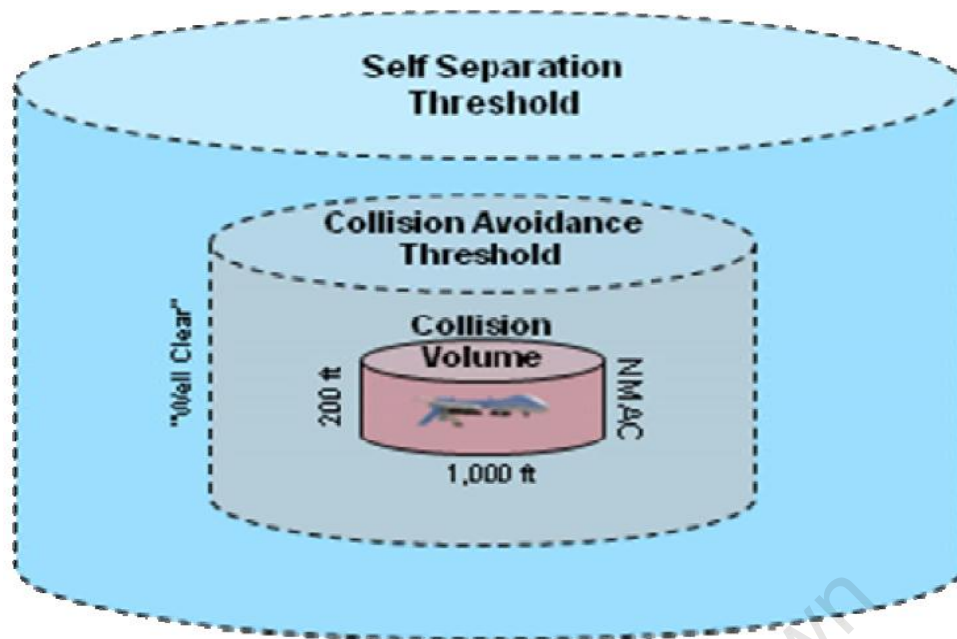


Figure 1.6: Visual representation of the Collision Volume as well as the SS and CA Thresholds [2]. A UAV is said to be well clear of an intruder aircraft if the intruder is outside of the CA Threshold.

A prerequisite for a SAS to make any inferences about its surrounds, is the presence of one or more sensors on-board the UAV platform. They provide the SAS with the capability to detect aerial threats and other obstacles in the UAV's path. Sensors commonly found on a UAV include electro-optic (EO) cameras, forward looking infra-red (FLIR)/infra-red search and track (IRST) sensors, and radars. Sensors on-board a UAV, which feed a SAS, must have a field of regard (FOR) of $\pm 110^\circ$ in azimuth (relative to longitudinal axis out of the nose of the UAV) and $\pm 15^\circ$ in elevation [9]. This is to ensure that sufficient coverage is provided to the SAS for reliable operation, especially during manoeuvres.

Table 1.1: Field of Regard requirements for an on-board UAV sensor used by a sense and avoid system.

Direction	Sensor Field of Regard	
Azimuth	-110°	+110°
Elevation	-15°	15°

A sensor's FOV is the volume in which it can detect targets while it is fixed in a single orientation. Whereas FOR is the overall FOV when the sensor occupies the extreme mechanical positions that is available to it (for example in a mechanically steered radar). The implication of this is that the FOV cannot be greater than the FOR. To explain this practically, let us consider a 2-Dimensional case where a sensor's FOV subtends a 30° arc. In the sensor's rest position, the longitudinal axis out of the centre of the lens is at 0° (see Figure 1.7). In this orientation, the FOV is $\pm 15^\circ$ (FOV1). Other than the rest position, there are another two orientations available: where the axis is centred at -15° and $+15^\circ$. At -15° the FOV is $-30^\circ - 0^\circ$ and at $+15^\circ$ the FOV is $0^\circ - 30^\circ$ (FOV2 and FOV3 respectively). Therefore the FOR for this scenario is $\pm 30^\circ$.

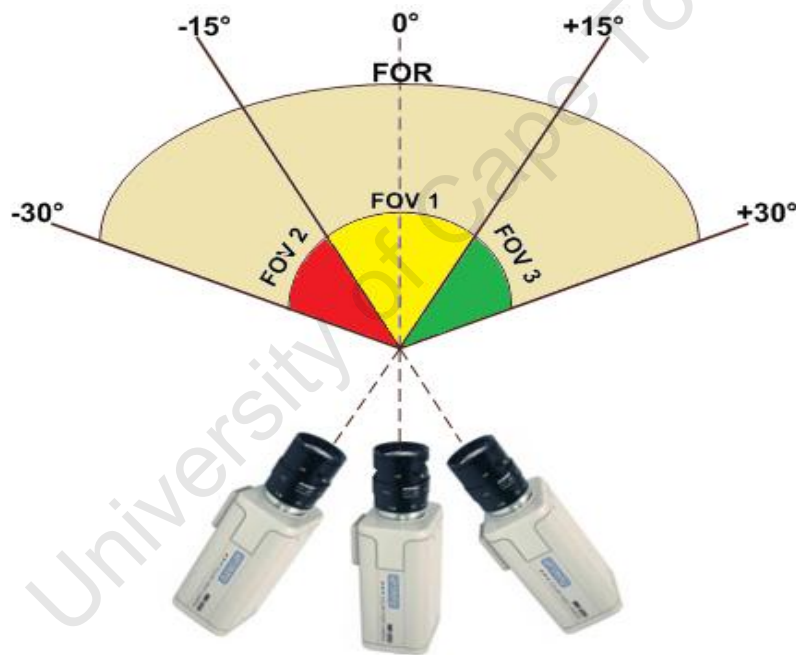


Figure 1.7: The FOR and FOV of a camera sensor with three allowable mechanical positions. Note FOV2 and FOV3 share -15° to 0° and 0° to 15° with FOV1, respectively.

In summary, the specifications (and rules and regulations) presented in this section have been developed so that a SAS which adheres to them will allow the smooth⁷ integration of UAVs into civil airspace. Moreover, it will realise the

⁷ Relatively smooth would probably be a better description as the words smooth and integration are, in practice, not usually married, as engineers would know all too well.

aviation community's dream of regular, fully autonomous UAV missions in civil airspace with risks to system safety properly mitigated while preserving existing air traffic flows [8].

1.3 Problem Description

As mentioned at the start of this dissertation, a robust and fully autonomous SAS is crucial for the future of UAVs. This SAS has the potential to become a reality because of the major advances actively being made in sensor technologies as well as computing power. Historically, the capabilities and functionality of systems were inherently limited by the technology on which they were built. This is no longer a limiting factor and researchers can begin to wholly utilise available resources.

Thus, through technological advancements, it is feasible for a SAS to make inferences on large amounts of data gathered from multiple sensors, allowing more accurate decisions to be made and the boundaries of UAVs to be pushed farther. The fusion of this multisensor data can be greatly beneficial to manned aircraft as well; where a single image is constructed from constituent sensor reports [10] (see Figure 1.8). This single, more meaningful "master sensor" report is what is presented to the pilot, opposed to multiple sensor reports, allowing him/her to focus more on the flying of the aircraft and less on complex sensor systems.

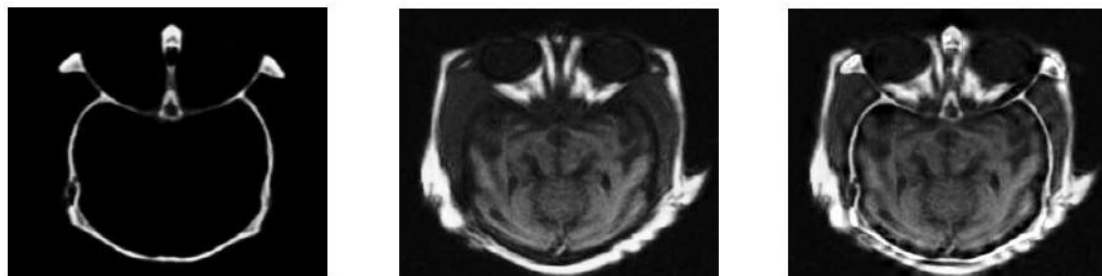


Figure 1.8: Fusion of multiple sensor images. From left to right: a CT image; an MRI image; and the, more informative, fused image [10].

The raison d'être of this dissertation is to present and discuss findings from the design and implementation of a threat tracking unit (here forth known as the TTU). In this project a UAV is equipped with both a passive EO camera as well as an active Doppler radar.

Kopp [11] is of the belief that sensor fusion techniques are 'at an early phase in their revolutionary life cycle'. Later in his article he expresses his view: 'It [sensor fusion] is another technology where the best is yet to come'. At the end of this journey, having analysed and solved the problems highlighted above, there would be considerable contributions to be made toward the inevitable realisation of a fully autonomous UAV SAS. Therefore, this research adds value to current research in a number of ways:

- By critically reviewing current trends in the chosen domain.
- Build on this current research and identify areas that are lacking.
- Address these gaps by critically examining and providing empirical solutions.

1.4 Scope and Objectives

The FAA's SAA Workshop [2] sets the scene for this dissertation. Figure 1.9 is a diagrammatic representation of the functionality required within a SAS. In this context, however, this thesis is only concerned with the development of the tracking functionality. Therefore this dissertation addresses the sub-function of **Track** and is illustrated by the blocked region in Figure 1.9.

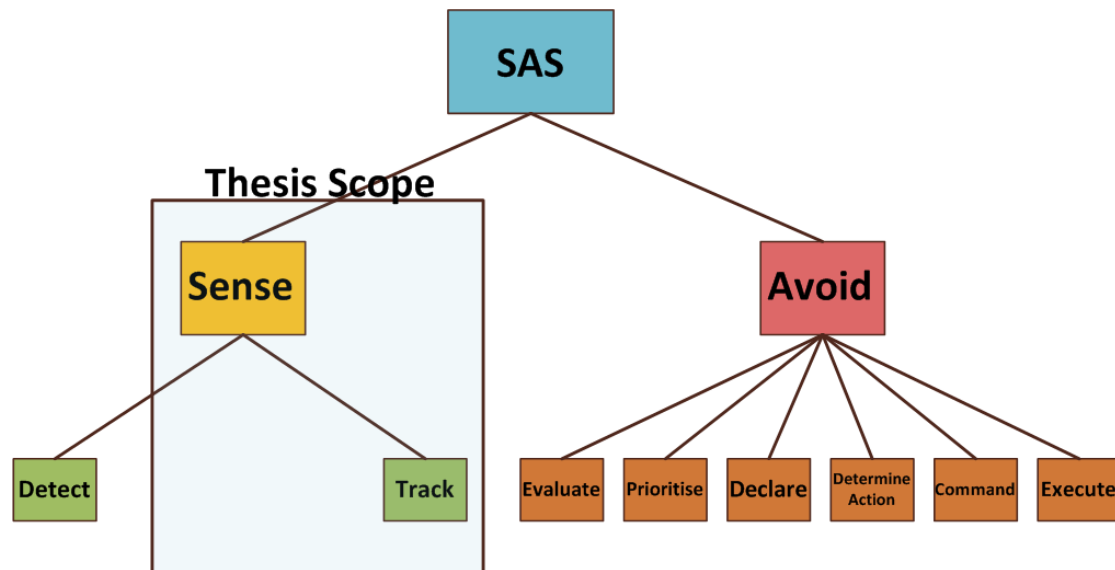


Figure 1.9: Thesis scope.

In this project a Doppler radar and an EO camera are available in order to perform target tracking. Thus, the overall aim of this research is to design an airborne tracking filter using multisensor data fusion. This leads to the following individual thesis objectives, with which to achieve this research aim:

1. Clarify the definition of multisensor data fusion.
2. To motivate the use the multiple, redundant sensors in a SAS.
3. As multisensor data fusion for target tracking is a non-trivial task, to provide a brief theoretical background on requisite tools and techniques in estimation and tracking.
4. Review current research on multisensor data fusion for target tracking.
5. Design and implement an airborne target tracking algorithm using multisensor data fusion.
6. Present performance metrics with which to evaluate the accuracy and performance of algorithms.
7. Demonstrate that the fusion filter more accurately tracks a target than a single sensor tracking filter.
8. Derive and discuss conclusions from the results acquired during algorithm development and knowledge gained from the literature study.

Through the completion of these objectives, the aim is to arrive at a suitable airborne tracking fusion architecture for a UAV SAS; of which the *Threat Tracking Unit* block, illustrated in Figure 1.10, will be composed.

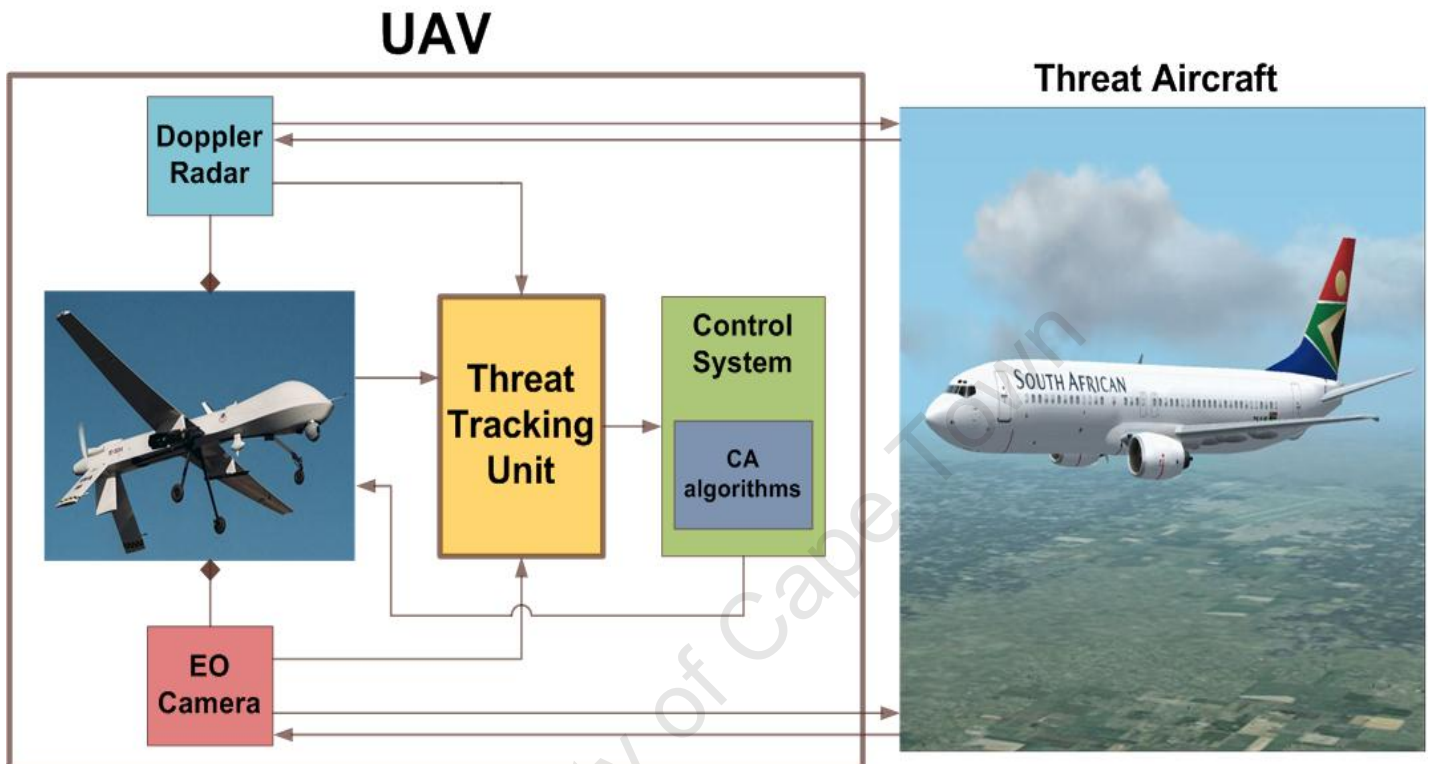


Figure 1.10: System overview. This thesis is concerned with the development of the TTU block. It is assumed that IMU and Actuator blocks exist inside the UAV picture.

Objective 1 to *objective 4* form the core of the Literature Review in Chapter 2. In order to achieve *objective 5*, the assumptions made during this study are:

- Two sensors are on-board the UAV platform: an EO camera and a Doppler radar.
- The EO camera provides its measurements to the SAS at a faster rate than the Doppler radar.
- A vision subsystem exists that detects the presence of other airborne threats and converts the detected image pixels to body-referenced target azimuth and elevation (using a pin-hole camera model).

- A radar subsystem exists that performs the necessary radar signal processing and converts the raw radar data to body-referenced target range, azimuth, elevation and range rate.
- Both sensors are fixed to the UAV centre of gravity.
- The UAV is modelled as a point mass travelling with constant velocity.
- The dynamics of threat aircraft trajectories obey a constant velocity model.
- Only one threat needs to be tracked at a time.
- There exists an Avoid subsystem whose inputs are the outputs from the tracker designed in this dissertation. The Avoid subsystem encompasses all the other SAA capabilities not in the scope of this thesis.
- Algorithms will not be tested on deployment hardware, but in simulation only. This is as a result of various risks present during flight tests.

1.5 Thesis Outline

Chapter 2 provides an extensive review of relevant literature pertaining to multisensor data fusion, estimation and target tracking. This chapter will form the foundation upon which the TTU algorithms will be designed.

Chapter 3 describes the design methodology adopted in this thesis.

Chapter 4 details the design of the simulation environment. This environment will be utilised extensively for the evaluation of the TTU algorithms.

Chapter 5 describes the detailed design of the TTU system. This entails the development of multiple airborne tracking filter architectures.

Chapter 6 consists of detailed comparisons of the various fusion architectures for the TTU system. These are discussed thoroughly and interpretations are made.

Chapter 7 concludes the thesis by unifying all the information obtained in the previous chapters.

University of Cape Town

Chapter 2: Literature Review

This literature review will clarify the term multisensor data fusion; motivate the need for multiple, redundant sensors and the advantages thereof; present and critically review present-day tracking filters⁸; and, when deemed necessary, provide a theoretical overview of mathematical tools and techniques. The literature review will draw to a close with a summary of topics covered and present the motivation for the algorithmic development to come.

The study within this review of literature focuses on objectives 1 to 4 as set out in the introductory chapter of this dissertation:

1. Clarify the definition of multisensor data fusion.
2. To motivate the use the multiple, redundant sensors in a SAS.
3. As multisensor data fusion for target tracking is a non-trivial task, to provide a brief theoretical background on requisite tools and techniques in estimation and tracking.
4. Review current research on multisensor data fusion for target tracking.
5. Design and implement an airborne target tracking algorithm using multisensor data fusion.
6. Present performance metrics with which to evaluate the accuracy and performance of algorithms.
7. Demonstrate that the fusion filter more accurately tracks a target than a single sensor tracking filter.
8. Derive and discuss conclusions from the results acquired during algorithm development and knowledge gained from the literature study.

Figure 2.1 illustrates how thesis objectives were translated into fields of study in the literature review. The reason for adopting this approach was to keep the literature review focused and to only discuss issues germane to satisfy the objectives of this thesis.

⁸ More emphasis is placed on the utilization of traditional tracking filters in multisensor tracking filters.

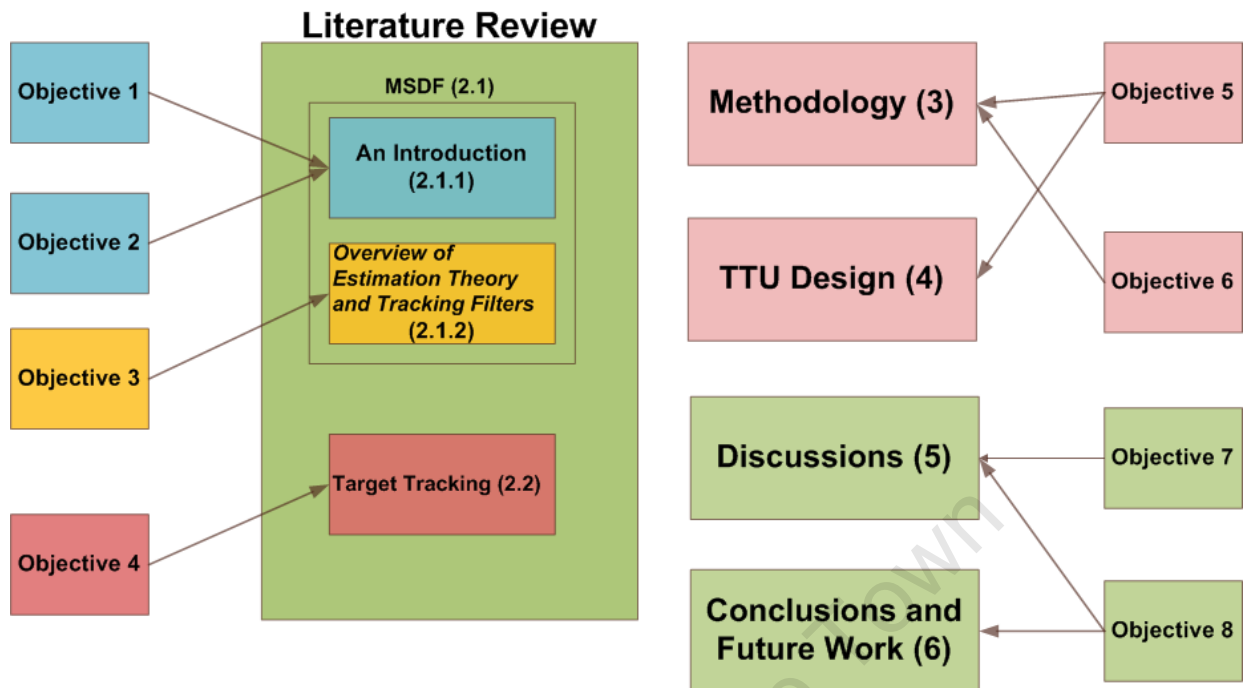


Figure 2.1: Mapping from thesis objectives to fields of study and, consequently, headings in the Literature Review. On the right, the mappings of other objectives are illustrated for the sake of completeness.

At the end of this literature review it is hoped that a thorough understanding of the literature and pertinent issues to target tracking are demonstrated. Furthermore, that the reader has a better understanding of these areas and that a clear justification for the algorithm development has emerged.

2.1 Multisensor Data Fusion

2.1.1 *An Introduction*

Multisensor data fusion (MSDF) seeks to combine data from multiple sensors to provide more meaningful information than from a single sensor [4]. This is analogous to the cognitive process evident in humans; we make decisions and draw conclusions based on a number of sensory inputs. MSDF may even provide information that is not possible with a single sensor. Moreover, MSDF allows for the detection, isolation and recovery of sensor failures [12] [13].

It is widely considered that Waltz and Llinas published the seminal text on MSDF in 1990; in which they explain that '[MSDF] involves significant integration of a number of research disciplines' [14]. Furthermore, data fusion spans a variety of applications. Some of these well suited applications are air-to-air and surface-to-air defence [15] [16], maritime surveillance [17] [18] and remote sensing [19] [20]. In [21] an airborne surveillance and tracking system is proposed which utilises MSDF; this system stands out from other designs, employing data fusion, as it consists of a particularly complex sensor suite, viz. a radar, an Identification Friend or Foe (IFF) system, an Electronic Support Measures (ESM) system, a Spotlight Synthetic Aperture Radar (SSAR) and a FLIR sensor. An example of an airborne platform with a similar sensor suite is illustrated in Figure 2.2. In order to understand the performance gains possible from utilising multiple, redundant sensors in a system, such as in [21], let us first consider the advantages of individual sensors.

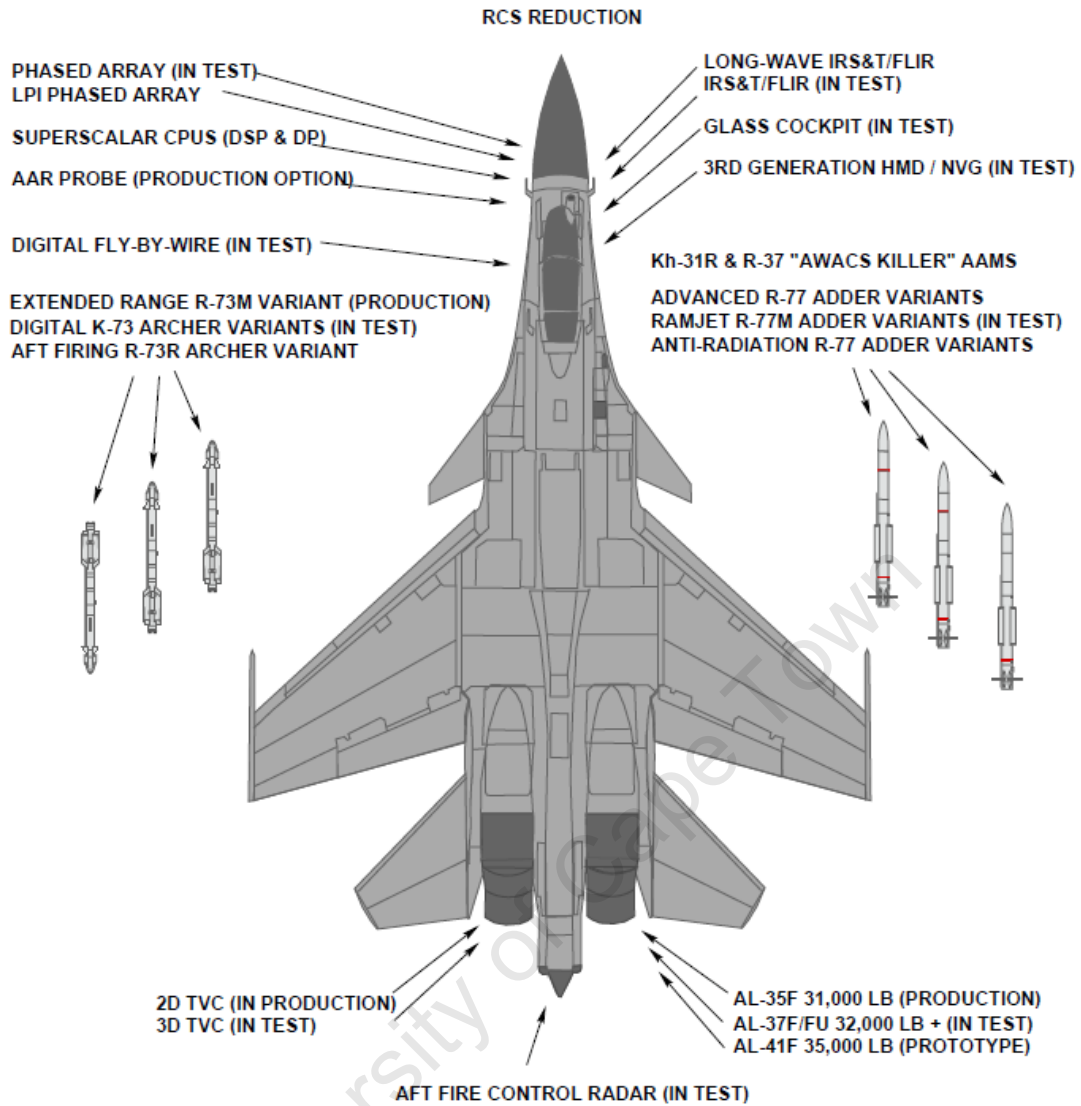


Figure 2.2: Fighter aircraft with radar and IR sensors. MSDF is the enabling technology to fully reap the benefits of having multiple, redundant sensors as part of a system [11].

For the purpose of this explanation (and to motivate our premise), let's review a radar and an EO sensor. By analysing their individual advantages and disadvantages, we would be in good stead in understanding the implications of their presence in a MSDF tracking system:

1. Radar is an active sensor⁹, whereas an EO sensor is a passive one.
2. An EO sensor is heavily dependent on atmospheric effects, including fog, rain and haze. [22]. This is especially relevant as the average

⁹ An active sensor is one that emits electromagnetic radiation.

global annual cloud cover is circa. 61 per cent [23] (see Figure 2.3).

Radar, on the other hand, is an all-day, all-weather sensor.

3. Sensors have inherent limitations on their sampling frequencies. An EO camera has a data rate that is faster than that of radar.
4. An EO sensor has greater angular resolution than radar. Radar, generally, has a greater effective range in which it can detect targets.
5. A radar measurement has a higher dimensionality than that of an EO sensor.

The effects of the pros and cons of each sensor, listed above, would have the following effects on a MSDF system:

1. The EO sensor would be able to detect targets but the radar would not during a fade in the radar return as a result of radar cross section scintillation [24], or other electronic countermeasures (ECM). This implies that when one subsystem stops operating that the overall tracking system can continue to be operational, thus making the tracking system more robust.
2. The tracking system would have all-day, all-weather operation. Weather, including cloud cover and rain, is a critical factor when selecting sensors.
3. An increased overall data rate results.
4. The system utilises the more accurate measurements from either sensor, improving overall system accuracy.
5. The addition of radar measurements leads to state observability [25].

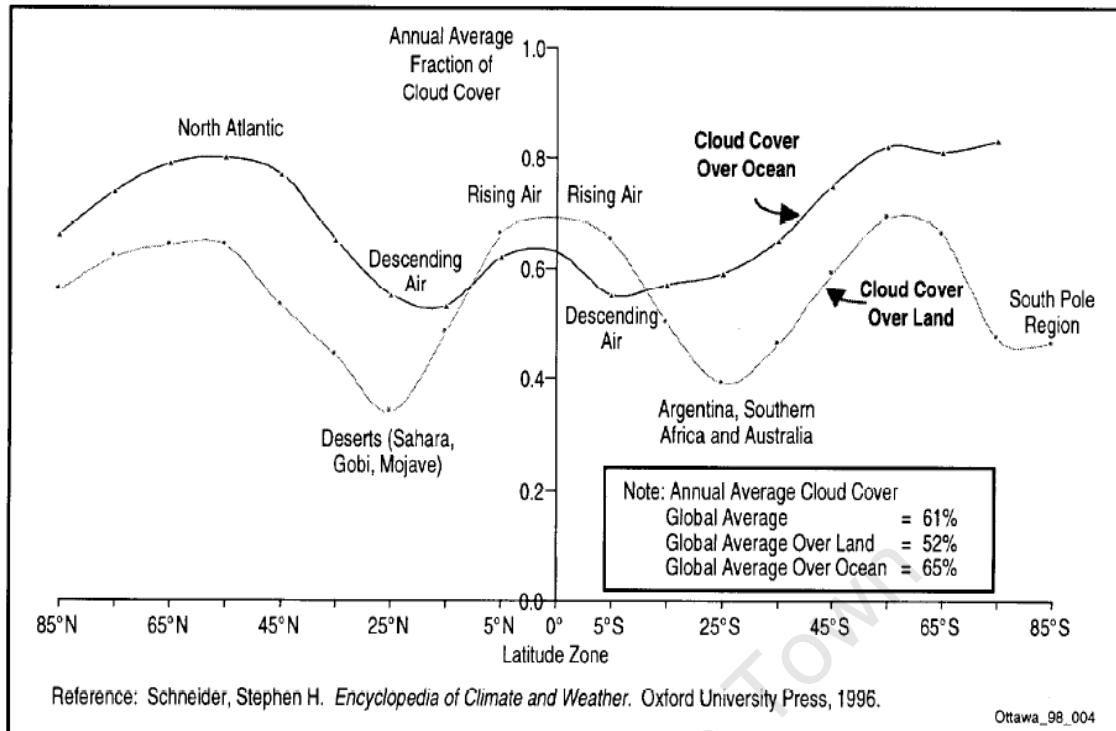


Figure 2.3: Cloud cover around the globe [23].

A further motivation for the need of MSDF is that, in our technologically-driven times, ECM systems are becoming more sophisticated and, therefore, the future will see more of the onus placed on passive sensors and algorithms. Moreover, advances in stealth technology, over the last two decades, have resulted in significant reductions in radar and IR signatures [11].

MSDF is the enabling technology with which the best attributes of both sensors, reviewed above, can be harvested. Therefore, in light of the above, the advantages of MSDF can be summarised in Table 2.1 below, adapted from [4].

Table 2.1: Advantages of MSDF.

Category of Advantage	Advantage	Operational Advantage
Improved system reliability	Multiple sensor systems have an inherent redundancy.	Graceful degradations.
Extended spatial coverage	That is, one sensor can look where another cannot.	Improved probability of detection.
Extended temporal coverage.	One sensor can detect when another cannot.	Robust system. E.g. less vulnerable to ECMs.

It is worth noting that, despite the advantages listed above, Hall and Steinberg [26] warn that the following precepts should be kept in mind when implementing a MSDF system:

- There is no substitute for a good sensor.
- Data fusion cannot correct errors in processing of individual sensor data.
- MSDF can result in poor performance if incorrect information about sensor performance is used.
- There is no such thing as a perfect data fusion algorithm.

There are three types of fusion architectures: centralised; distributed and hybrid. The architecture depends on the data that arrives at the fusion centre (FC) i.e. raw sensor data implies a centralised architecture (see Figure 2.4); processed data or state estimates implies a distributed architecture (see Figure 2.5); and a combination of data results in a hybrid architecture. Hybrid architectures are computationally intensive and not suitable for practical implementations. Therefore, they are not considered in the literature.

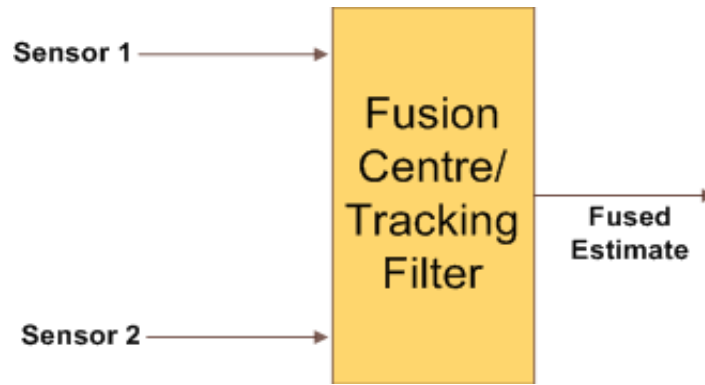


Figure 2.4: Centralised fusion architecture.

Centralised architectures imply raw sensor data arrives at the FC. Here there are three methods with which data can be fused in the FC:

- Measurement fusion method; based on minimising the mean square error [27].
- Parallel fusion [28].
- Sequential fusion.

Assume two sensors i and j with measurement equations z_k^i and z_k^j , respectively. When making use of the measurement fusion method, the fused measurement z_k^F at time instant k is then:

$$z_k^F = z_k^i + R_k^i (R_k^i + R_k^j)^{-1} (z_k^j - z_k^i). \quad (2.1)$$

Where R_k^i and R_k^j are the covariance matrices of z_k^i and z_k^j , respectively. The fused covariance matrix is determined by:

$$R_k^F = (R_k^{i^{-1}} + R_k^{j^{-1}})^{-1}. \quad (2.2)$$

These fused measurements are then tracked in order to produce an estimate of a target's state vector.

The second method of performing centralised fusion is parallel fusion and is the most intuitive one of the three as multiple sensors can be visualised as one. It involves combining all the observation vectors (from multiple sensors) and measurement matrices into larger composite ones. The composite observation becomes [28]:

$$z_k^F = \begin{bmatrix} z_k^i{}^T & z_k^j{}^T \end{bmatrix}^T. \quad (2.3)$$

With a measurement matrix:

$$H_k^F = \begin{bmatrix} H_k^i{}^T & H_k^j{}^T \end{bmatrix}^T. \quad (2.4)$$

And composite measurement covariance:

$$R_k^F = \text{diag} \left(\begin{bmatrix} R_k^i & R_k^j \end{bmatrix} \right). \quad (2.5)$$

These new composite matrices affect only the *corrector* stage of the Kalman Filter (KF) equations; and are used as such to track the state vector of a target. In principle, parallel fusion can accommodate any number of sensors. However, the computational resources required quickly multiply as sensors are added to the system. This point is evident by analysing the matrix dimensions of the composite matrices formed. Consider a state vector of dimension n , x observation vectors of dimension m^x with measurement matrices (H) of dimension $m^x \times n$. The composite observation vector will then have dimension $m = \sum_{i=1}^x m^x$ and the composite measurement matrix will

have dimension $m \times n$. As an example, if $n = 6, x = 2$ and $m^x = 3$ then the dimension of composite H is 6×6 . This increase in matrix dimensions is a challenge as H and R are involved in the computationally-expensive matrix inverse operation each time-step. Therefore, one can see that as sensors are added to the system the viability of this approach lessens.

The other method of performing centralised fusion is sequential fusion. Here each sensor's observation vector is sequentially used to update the state estimate. That implies that the KF *corrector* stage is repeated for as many times as there are sensors in the system; at each iteration using the state and covariance estimates produced from utilising the previous sensor's observation vector. Once all observation vectors have been processed the KF continues its normal filter operation. Sequential fusion is also very computationally expensive as it requires the computation of a new gain matrix and prediction every time-step for every observation from every sensor in the system.

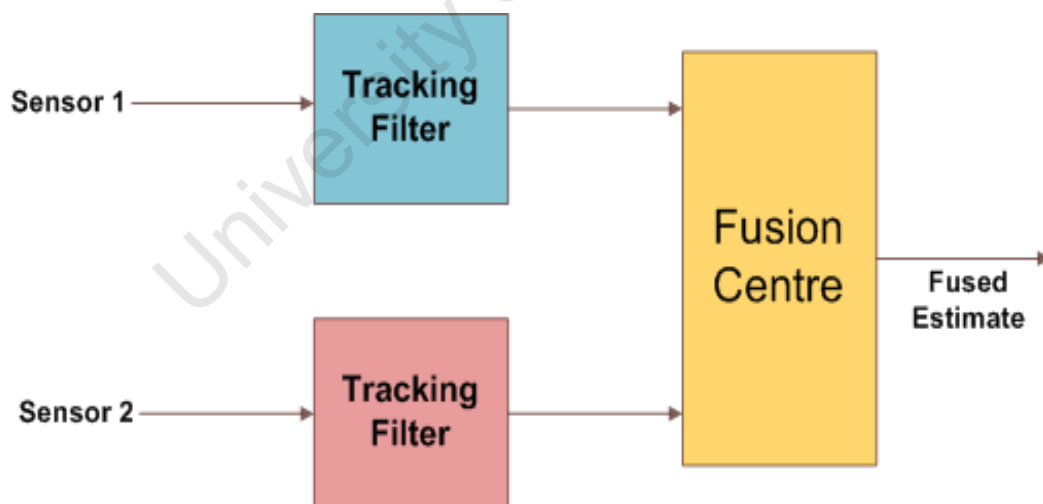


Figure 2.5: Distributed fusion architecture.

Distributed architectures, as its name suggests, are distributed in nature. There exists an independent tracking filter for each sensor in a system, illustrated in Figure 2.5 above. Each tracking filter produces its own local state

estimate and these state estimates are fused using the state vector method of fusion (an example of a distributed fusion architecture is illustrated in Figure 2.6). The fused covariance and state estimate are evaluated as follows [29]:

$$P_{k|k}^F = \left(P_{k|k}^{i-1} + P_{k|k}^{j-1} \right)^{-1} \quad (2.6)$$

and

$$\hat{X}_{k|k}^F = P_{k|k}^F \left(P_{k|k}^{i-1} \hat{X}_{k|k}^i + P_{k|k}^{j-1} \hat{X}_{k|k}^j \right). \quad (2.7)$$

Where $\hat{X}_{k|k}^i, \hat{X}_{k|k}^j$ are the posterior state estimates produced from each local tracking filter and $P_{k|k}^i, P_{k|k}^j$ are their covariance matrices.

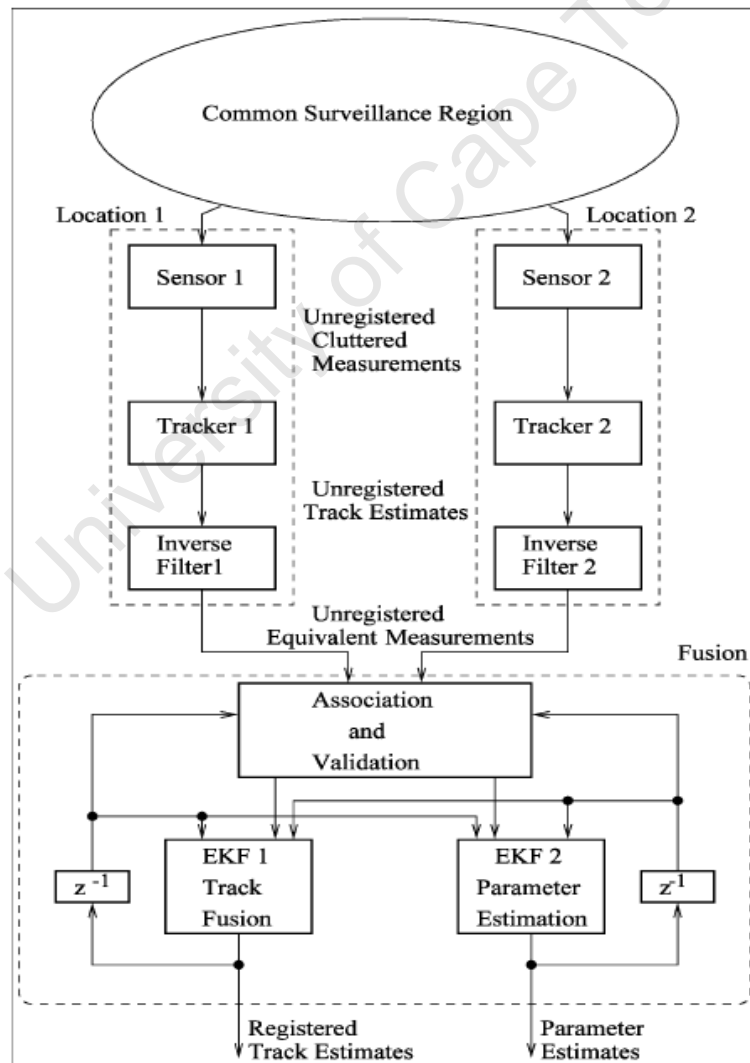


Figure 2.6: An example distributed fusion architecture application.

A common assumption made in distributed architectures is that local filter estimates are uncorrelated [29] [30] [31]. Multiple fusion architecture are developed in [30] to alleviate this problem. The reason for the aforementioned assumption is that when accounted for, there is a significant increase in filter complexity. The implications of a common process model used by each local filter are twofold: it first and foremost makes it feasible to fuse the individual state estimates; and second, it causes the individual state estimates to be correlated [32].

Equations (2.6) and (2.7) hold only under the aforementioned assumption. If local estimates are correlated, then this correlation needs to be accounted for in the filter in order to fuse the local estimates. This is accomplished by calculating the cross-covariance $P^{ij} = P^{ji^T}$ between the two state estimates and then using this to fuse them. The predicted cross-covariance is [32]:

$$P_{k|k-1}^{ij} = \Phi P_{k-1|k-1}^{ij} \Phi' + GQG'. \quad (2.8)$$

And the estimate cross-covariance is:

$$P_{k|k}^{ij} = [I - K_k^i H_k^i] P_{k|k-1}^{ij} [I - K_k^j H_k^j]^T. \quad (2.9)$$

Equations (2.12) and (2.13) provide a recursive means of calculating the cross-covariance between $\hat{X}_{k|k}^i$ and $\hat{X}_{k|k}^j$. The fused estimate and covariance matrix are determined by the following equations:

$$\begin{aligned} \hat{X}_{k|k}^F &= \hat{X}_{k|k}^i - \left(P_{k|k}^i \right. \\ &\quad \left. - P_{k|k}^{ij} \right) \left(P_{k|k}^i + P_{k|k}^j - P_{k|k}^{ij} \right. \\ &\quad \left. - P_{k|k}^{ij^T} \right)^{-1} \left(\hat{X}_{k|k}^j - \hat{X}_{k|k}^i \right) \end{aligned} \quad (2.10)$$

and

$$\begin{aligned}
 P_{k|k}^F = P_{k|k}^i - & \left(P_{k|k}^i \right. \\
 & \left. - P_{k|k}^{ij} \right) \left(P_{k|k}^i + P_{k|k}^j - P_{k|k}^{ij} \right. \\
 & \left. - P_{k|k}^{ij T} \right)^{-1} \left(P_{k|k}^i - P_{k|k}^{i T} \right).
 \end{aligned} \tag{2.11}$$

The reader should note that, according to the assumptions and objectives laid out in Chapter 1, the scope of this thesis is limited to single target tracking. As such, data association techniques are not included here. An excellent survey on data association for MSDF applications can be found in [4].

In summation of this introduction to MSDF, the choice of fusion architecture is dependent on the application at hand as well as the hardware and software limitations imposed on the system. Distributed fusion is a much less mature area of research compared to centralised fusion, which has seen much research [33]. Distributed architectures require less computational resources [34] and exhibits higher survivability¹⁰, but centralised architectures (through measurement fusion) result in more accurate tracking filters [35]. Furthermore, measurement fusion is the most widely used method when performing centralised fusion. An important consideration when selecting fusion methods is matrix dimensions; centralised fusion methods require that observation vectors from all the sensors, present in the system, have the same dimension and hence their measurement matrices. Whereas in distributed fusion, the state vector from all independent tracking filters are required to have the same dimensions.

¹⁰ The entire tracking system which employs a centralised fusion architecture has a single point of failure – the FC. The reader is reminded that in a centralised architecture the FC and tracking filter are one and the same.

2.1.2 Overview of Estimation Theory and Tracking Filters

Tracking using MSDF is a non-trivial task, which spans a number of engineering disciplines. In what follows, this subsection seeks to provide a background on mathematical techniques used in estimation and the realisation of tracking filters. The work presented in this section is used later in this thesis.

State estimation has its roots in least squares estimation (LSE). Invented in 1795 by Gauss to obtain the best estimate of orbital parameters for asteroids [4], LSE estimates the state vector using a batch processing technique where multiple measurements (over multiple sensor returns) are stored and then processed simultaneously [36]. This LSE method can be extended into a nonlinear case as well as a recursive form. The recursive LSE approach provides an updated state estimate after each sensor return, i.e. it does not store successive measurements. A least squares estimator is similar to a Kalman Filter in that they both minimise the expected mean squared error between the actual and estimated states; they are also dissimilar in that the LSE assumes deterministic state dynamics. As the reader may have deduced already, the recursive form of LSE has led to the innovation of the ubiquitous KF.

The KF is an extension to the least squares estimator; it incorporates process noise which allows for the modelling of random target motion. It is a linear estimator that, as long as the target dynamics and measurement noise are accurately modelled and the system is Gaussian, minimises the mean squared error. The KF provides a convenient measure of estimation accuracy through its state covariance matrix [25]. Furthermore, the KF has a number of useful features pertinent to MSDF applications [4]:

- A wide variety of measurement models (observation equation) and process/plant models (state equation) can be modelled.
- Quantitative evaluation of the contribution of each sensor to overall estimation accuracy.

- It is robust and fairly simple to implement.

In what follows, the KF algorithm is described and its equations are presented; the reader is directed to [37] for the detailed mathematical derivations of the KF equations and discussions of its applications. Consider a discrete-time target dynamics process which is modelled in the discrete-time Markov form [38]:

$$X_{k+1} = \Phi X_k + \Gamma w_k + u_{k+1|k}. \quad (2.12)$$

Where X_k is the $n - dimensional$ target state vector to be estimated. w_k is the zero-mean, white, Gaussian process noise¹¹ with covariance Q . Φ is the known state transition matrix, and $u_{k+1|k}$ is the known deterministic input.

Equation (2.12) describes the target dynamics in terms of a Markov process¹² represented by the state equation.

Sensor measurements arrive as linear combination of the state variables:

$$z_k = HX_k + v_k. \quad (2.13)$$

Where z_k is the $m - dimensional$ observation vector. v_k is the zero-mean, white, Gaussian measurement noise with covariance R . H is the $m \times n$ measurement matrix.

¹¹ An example of process noise when modelling a target is its random accelerations.

¹² A Markov process is a process in which its statistical representation in the future is determined solely by the present state, i.e. independent of the past.

Given Equations (2.12) and (2.13) the KF *predictor* and *corrector* stages are [25]:

a) Corrector

$$K_k = P_{k|k-1} H' (H P_{k|k-1} H' + R)^{-1} \quad (2.14)$$

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (z_k - H \hat{X}_{k|k-1}) \quad (2.15)$$

$$P_{k|k} = (I - K_k H) P_{k|k-1} \quad (2.16)$$

b) Predictor

$$\hat{X}_{k+1|k} = \Phi \hat{X}_{k|k} + u_{k+1|k} \quad (2.17)$$

$$P_{k+1|k} = \Phi P_{k|k} \Phi' + Q \quad (2.18)$$

Here K_k is the Kalman gain and I is the Identity Matrix. For future reference, the dimensions of the KF matrices are summarized in Table 2.2 below.

The *predictor* stage (also known as the time update) defined by Equations (2.17) and (2.18) serves to propagate in time the estimated state vector $\hat{X}_{k|k}$ and covariance $P_{k|k}$. The *corrector* stage (also known as the measurement update) defined by Equations (2.14) - (2.16), provides feedback into the system by incorporating the current measurement z_k into the *a priori* state estimate $\hat{X}_{k|k-1}$ to yield an improved *a posteriori* state estimate $\hat{X}_{k|k}$.

In (2.15) the factor $(z_k - H\hat{X}_{k|k-1})$ is known as the innovation (or residual) vector with a corresponding innovation covariance matrix defined by the factor $(HP_{k|k-1}H' + R)$. The innovation is an important quantity as it provides insight into the performance of the filter by giving us a measure of how well the filter is performing.

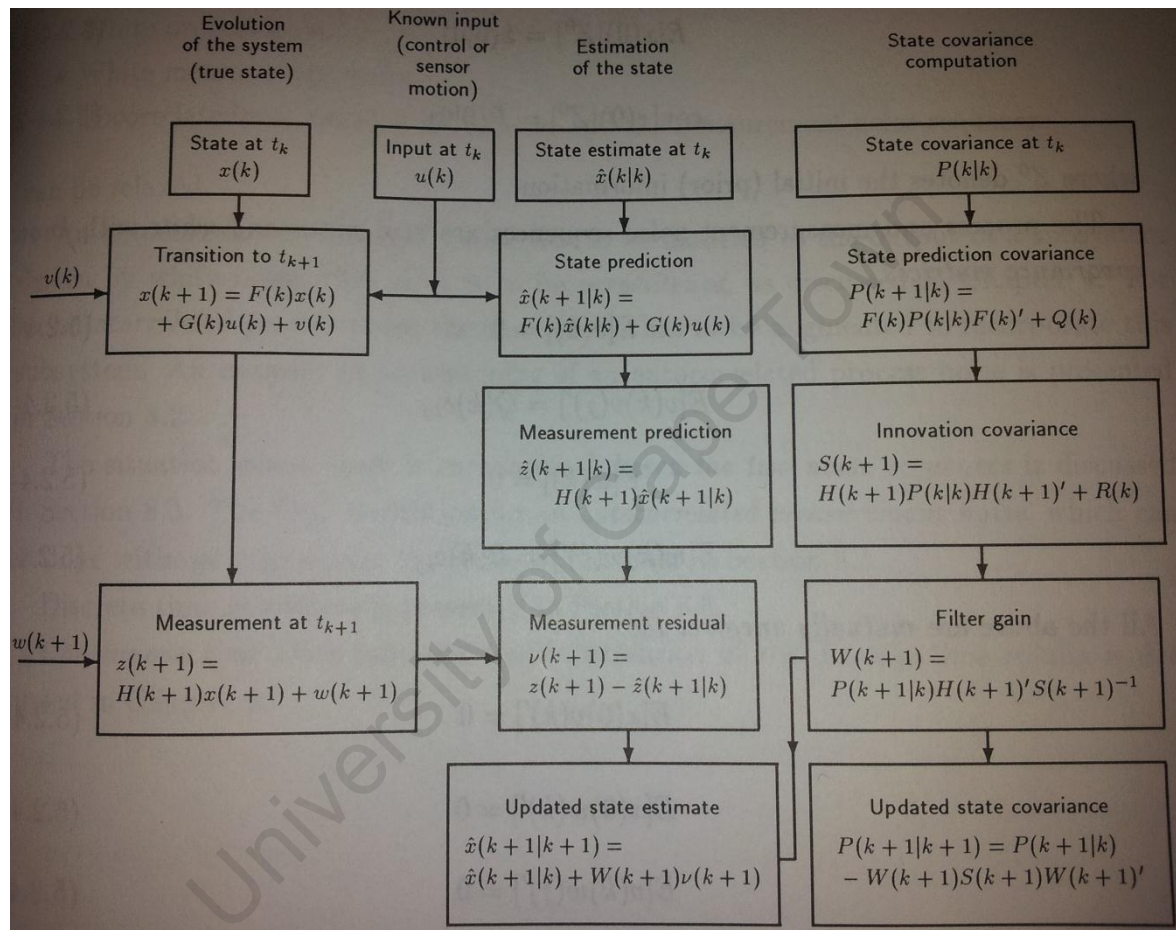


Figure 2.7: Single step Kalman filter overview obtained from page 219 of [39].

The Kalman gain in (2.14) tells us that as the measurement covariance R tends to zero, the real measurement is trusted more and, conversely, the predicted measurement $H\hat{X}_{k|k-1}$ is trusted less. This is in direct contrast to the way the *a priori* state covariance $P_{k|k-1}$ affects the innovation covariance; as $P_{k|k-1}$ tends to zero the real measurement is trusted less and, conversely, the predicted measurement is trusted more. Therefore the Kalman gain

dynamically weights either the real or predicted measurement when determining the posterior estimate, dependent on both the state and measurement covariance matrices and hence the innovation covariance. A graphical interpretation of the KF algorithm is displayed in Figure 2.7.

Table 2.2: KF matrix dimensions.

Matrix		Dimensionality
Name	Designator	
State vector	X	$n \times 1$
State covariance	P	$n \times n$
Process noise covariance	Q	$q \times q$
Transition	ϕ	$n \times n$
Noise	Γ	$n \times q$
Observation vector	z	$m \times 1$
Measurement	H	$m \times n$
Measurement noise covariance	R	$m \times m$

Before delving into the realm of nonlinear filters, the system designer has other options to handle nonlinear systems. One solution is the Pseudo Linear Kalman Filter (PLKF). Another option is a new filter developed in [40] (based on the KF); it accommodates nonlinear systems yet it does not require any form of linearization process. The PLKF transforms a nonlinear measurement equation into a more convenient linear form [25], negating the need for a nonlinear filter. Consider a sensor measurement as that in Equation (2.13),

but the measurement matrix H is replaced with $h(X)$ as the measurement equation is nonlinear. Through algebraic manipulation, the pseudo measurement is expressed as:

$$z_s = H(h(X))X = 0. \quad (2.19)$$

Where $H(h(X)) = H_s(z)$.

Then the standard KF equations (Equations (2.14) - (2.18)) are used to estimate the state vector with H_s replacing H , except for two notable differences:

- i. The residual in Equation (2.15) is replaced by the factor:

$$(0 - H_s(z)\hat{X}_{k|k-1}). \quad (2.20)$$

- ii. R_s replaces R , with:

$$R_s = gRg'. \quad (2.21)$$

Where g is found by taking the partial derivative of the pseudo measurement vector with respect to the actual measurement vector:

$$g = \frac{\partial z_s}{\partial z}. \quad (2.22)$$

When state dynamics or measurement models are no longer Gaussian in nature, standard KF equations do not hold true. That is the target dynamics or the measurement matrix is nonlinear. This implies that the transition matrix or the measurement matrix would be a nonlinear combination of the state variables. Under these conditions the Extended Kalman Filter (EKF) becomes

applicable. The EKF is a nonlinear extension of its linear counterpart: the KF; which is apparent from Equations (2.23) - (2.25). Historically speaking, the practical implementation of EKFs was limited by hardware platforms of the era. This is asserted by the fact that the EKF is significantly more computationally expensive than its linear equivalent. The *predictor* and *corrector* stages of the EKF are:

a) Corrector

$$K_k = P_{k|k-1} H_X(k)' (H_X(k) P_{k|k-1} H_X(k)' + R)^{-1} \quad (2.23)$$

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (z_k - h(\hat{X}_{k|k-1})) \quad (2.24)$$

$$P_{k|k} = (I - K_k H_X(k)) P_{k|k-1} \quad (2.25)$$

b) Predictor

$$\hat{X}_{k+1|k} = \Phi \hat{X}_{k|k} + u_{k+1|k} \quad (2.26)$$

$$P_{k+1|k} = \Phi P_{k|k} \Phi' + Q \quad (2.27)$$

It is evident that the *predictor* stages for the KF and the EKF are the same, but *corrector* stages differ. In the KF case the observation vector was a linear relation of the state variables, hence the measurement matrix was linear in nature. In this scenario the observation vector is a nonlinear function of the state vector, i.e. $h(X)$. According to the assumptions around which the EKF are built, the error dynamics can be accurately characterised by a linearized first-order Taylor series expansion and disregarding the second and higher-

order terms [41] [42]. Therefore, by performing a Taylor series expansion of this $h(X)$ and evaluating this about the predicted estimate, we arrive at the linearized measurement matrix [25]:

$$H_X(k) = \left. \frac{\partial h(X)}{\partial X} \right|_{X=\hat{X}_{k|k-1}} \quad (2.28)$$

In the preceding EKF example, it was assumed that the system to be estimated consisted of a nonlinear measurement process only. If the target dynamics were, instead, nonlinear, then only the *predictor* stage of the EKF would change. In this scenario the transition matrix is a nonlinear function of the state variables. The resulting *predictor* equations would be determined by the linearization of the transition matrix.

Through the linearization of the observation, or state models, errors are introduced into the system and, to quote Denham and Pines [43], ‘invalidates all properties of optimality and convergence of the KF’. Furthermore the computational processing of the EKF equations requires significantly more resources; this is evident from the juxtaposition of measurement matrices of the KF and EKF. In the KF the measurement matrix is constant, whereas in the EKF the linearized measurement matrix is evaluated at every time epoch. Furthermore, the predicted measurement calculation first involves evaluation and then matrix multiplication.

For highly nonlinear systems, there are other more complex filters that can be used. These filters seek to alleviate the inherent errors present in the EKF due to the linearization process, at the expense of even more computing resources. Some of these filters include used in tracking applications are: Iterated EKF, Second order EKF [44], Unscented KF (UKF) [44] [45] and Particle Filter [46]. Interacting Multiple Model (IMM) Filters are typically used to track manoeuvring targets. This is achieved through the use of multiple KF models, one for each target manoeuvre stage. Then the individual state

estimates and covariance matrices (from each KF) are combined through the use of a Markov model for the transition between manoeuvre stages [39].

The reader should note that in recent years it has been proven that the use of Equation (2.25) should be avoided [47]. According to Li and Jilkov [47], the reason for this is twofold: first, it invites horrible numerical problems; and second, it is theoretically valid only when the Kalman gain K is truly optimal (which, in practice, is rarely the case). The new equation is given as:

$$P_{k|k} = (I - K_k H_X(k))P_{k|k-1}(I - K_k H_X(k)) + K_k R K_k' \quad (2.29)$$

In summary, this subsection provided a clear definition of MSDF, conveyed some of its applications and motivated the need for multiple, redundant sensors. Following this, the reader was equipped with estimation and KF theory required in order to perform MSDF and realise the filters developed in Chapter 5. The above satisfies thesis objectives 3 and 4 laid out in section 1.4.

2.2 Target Tracking

Tracking is the process of determining a target's quantities of interest (state vector), including target position and velocity, under the influence of noisy sensor measurements. As outlined in Section 2.1.2, when a target's dynamic and measurement processes are assumed Gaussian then a linear filter, for example the Kalman filter, can be used to estimate the state vector of interest. Blackman and Popoli [25] state that, in this case, the "estimation of the mean target state and the associated covariance matrix is all that is required to define the probability density function (PDF) associated with the target position in state space". If, on the other hand, this Gaussian assumption is not accurate, such as when the target dynamics or measurement process is nonlinear, then a nonlinear filter becomes necessary. Blackman and Popoli

[25] go further where they declare that, in this case, it is better to propagate the PDF directly in target state space.

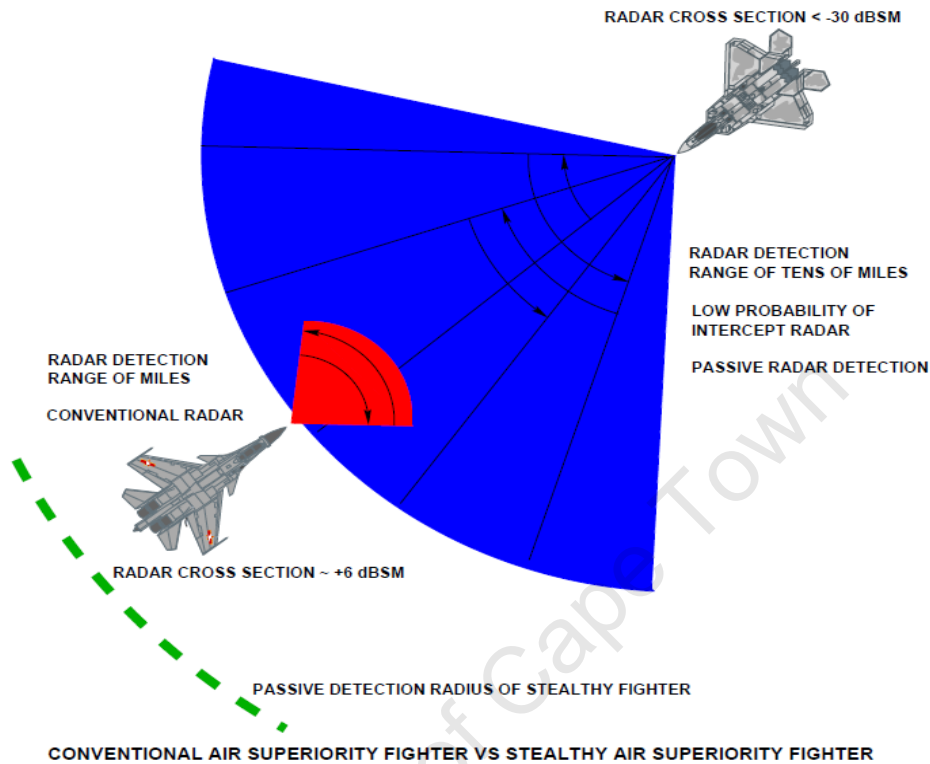


Figure 2.8: Airborne radar in action [11].

One of the first choices faced by the system designer when designing a tracking filter¹³ is selecting the coordinate system in which to track. Under the assumptions laid out in section 1.4 of this thesis, this literature study is constrained, for the most part, by the presence of these two sensors as part of the tracking system. As such, in order to track a target, one could use one of the following sensor configurations:

- Bearings-only (camera)
- Doppler radar
- Multiple sensors.

¹³ A tracking filter is so called as it **tracks** random variables in the state vector by **filtering** out noise from multiple sources. Moreover, the process of finding the best estimate from noisy data amounts to filtering out the noise.

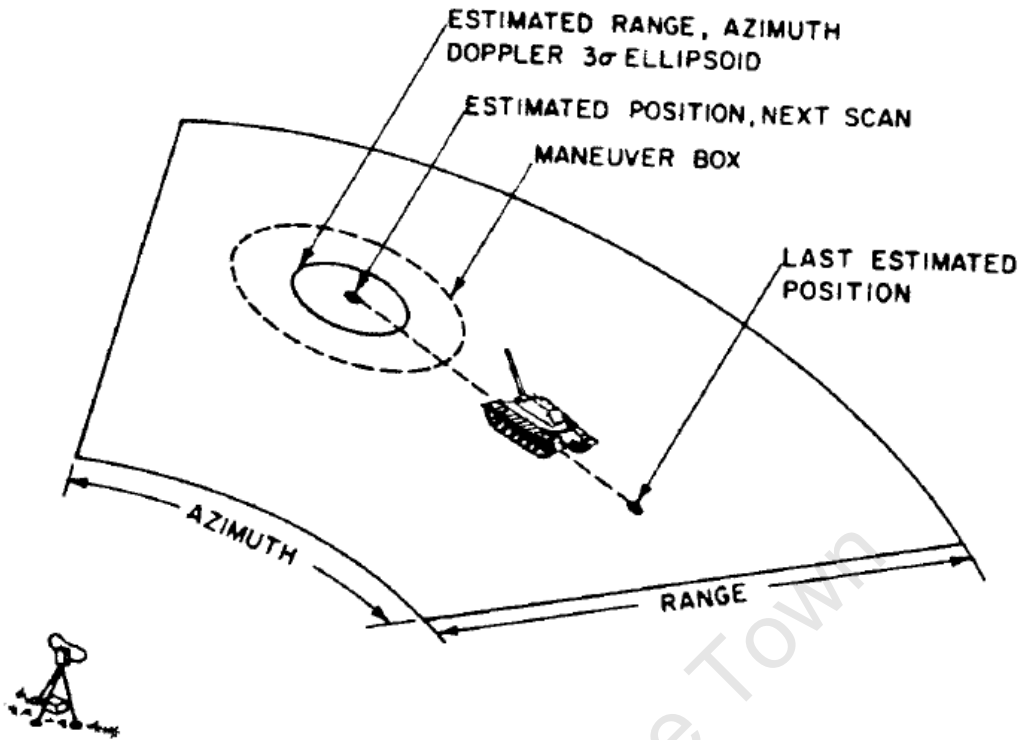


Figure 2.9: Ground-based radar target tracking [36].

In light of the above, this section will proceed in a logical fashion. It will provide an overview of tracking with each sensor leading up to, ultimately, tracking with multiple, redundant sensors.

2.2.1 *Selecting a Tracking Coordinate System*

All tracking systems require a frame of reference in which to track targets. The choice of which is not usually a straightforward task. It must be one that is meticulously thought out in order to make the most informed decision. Much like the selection of a fusion architecture (described in Section 2.1), the selection of a coordinate system in which to perform tracking is heavily dependent on the application at hand [25]. In the end of this section, it will be clear that the choice of tracking coordinate system plays a significant role in tracker performance.

The North-East-Down (NED) system is the best choice if multiple targets are to be tracked as a fixed inertial frame is necessary. Blackman [24] asserts this by stating that it is preferable to use a non-rotating or inertial coordinate system so that multiple target tracks can be processed with respect to the same fixed reference. An advantage of tracking in Body Reference Frame BRF (see Figure 2.10) is that it negates the need for a measurement conversion to the NED system. Furthermore, for a single target, the BRF system makes sense. However, Fasano [48] highlights the major disadvantage of tracking in BRF when he says, “relative motion in BRF includes attitude dynamics, which makes it more difficult to track, unlike its projection in NED, which only depends on relative position dynamics”.

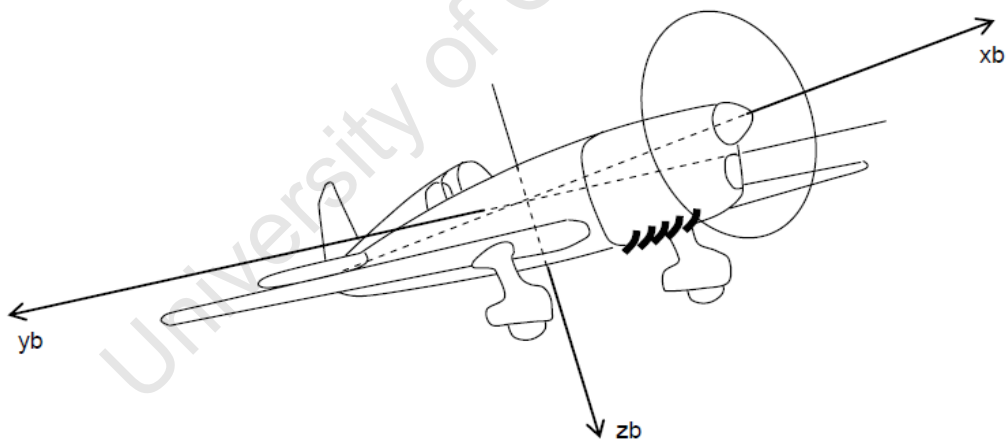


Figure 2.10: Body axis coordinate system. Also known as the body reference frame (BRF) [49].

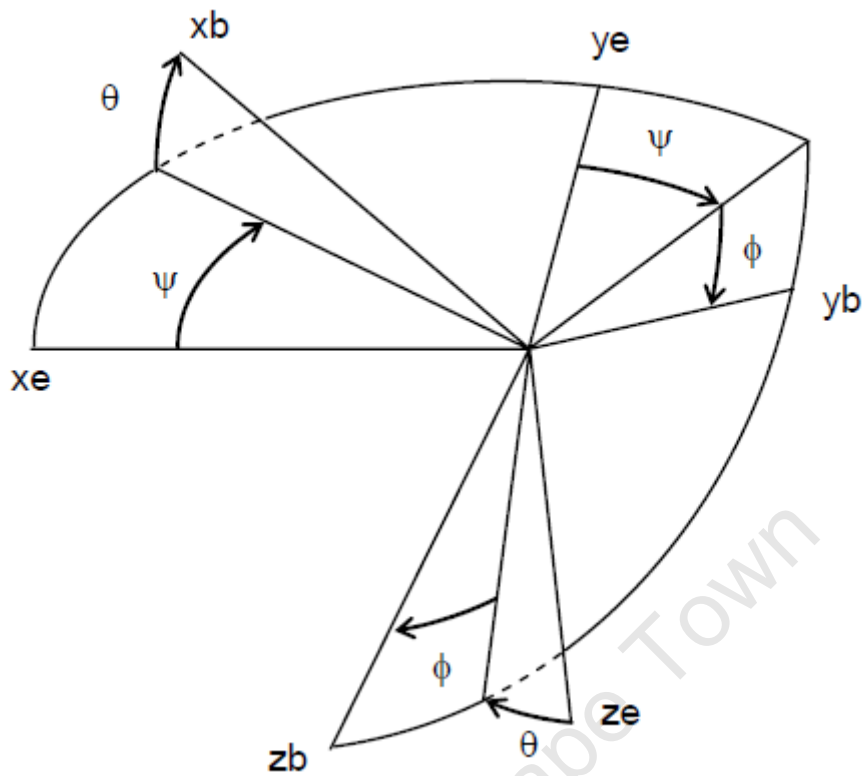


Figure 2.11: Relationship between BRF and a fixed earth frame, e.g. the NED system. [49]

Tracking is generally performed in either the geocentric North-East-Down reference frame or platform-centric Body Reference Frame (BRF)¹⁴. BRF has its origin at the aircraft centre of gravity and axes along longitudinal, lateral and vertical aircraft axes. The NED frame has the same origin as the BRF and its axes are determined by the north direction and down direction pointing to the centre of the Earth. The east axis is the direction perpendicular to the north and down axes. It is worth noting that the NED system is not strictly inertial for a moving platform. The reason for this is that platform axes are slowly changing their attitude (orientation in inertial space) as the platform moves over the Earth's surface. Blackman [25] affirms that these rotations are significant only when near the North Pole and, therefore, the NED system is considered an inertial one for moving platforms. As an example of tracking in

¹⁴ Tracking in BRF is often referred to as tracking in "body" or "body coordinates".

BRF, Pearson [50] presents a single target tracking system which tracks a target's position relative to an airborne sensor's pointing axes.

Four centuries ago Rene Descartes watched a fly walk across his ceiling and wondered how to capture its position [51]. This ultimately resulted in the birth of the Cartesian coordinate system. This is one of the coordinate systems with which to perform tracking; the other being spherical. The choice of Cartesian or spherical space for this purpose each come with their own set of caveats. For example, if a Cartesian frame is chosen then the measurement equation of the filter is nonlinear¹⁵. This is in contrast to the choice of a spherical frame where the target dynamic model is nonlinear, i.e. the state transition is a nonlinear function of the state variables. This nonlinearity, in the dynamics, results in the introduction of pseudo accelerations in the angular components (even for a constant velocity target model) [52]. However, for distant targets these pseudo accelerations are small [53]. Another example is when a near constant velocity model [54] is expressed in spherical space it is significantly more complex than its projection in Cartesian space [47].

It is considered that systems that make use of multiple, geographically separated sensors must use Cartesian coordinates [25]. In a technical report by Zollo and Ristic [53], they prove that tracking in Cartesian space is more accurate during non-maneuvring periods of the target trajectory; whereas tracking in spherical coordinates is, conversely, more accurate during manoeuvring periods.

When the Cartesian space is considered, tracking is commonly performed by a linear approximation of the measurement by taking the partial derivatives with respect to the target state and feeding this into the *corrector* stage of an EKF (see Section 2.1.2). Duan et al. [55] propose a method which reduces the nonlinearity of the measurement by forming a pseudo measurement which is the product of the range and range rate. To steer away from a nonlinear filter,

¹⁵ This is true in the case of airborne tracking.

an alternative is to convert the spherical sensor measurements to Cartesian space and then de-bias them [53] [48]. This allows the use of the linear KF for tracking.

2.2.2 Passive Bearings-Only Tracking

Monocular vision-based tracking has seen extensive research in the last decade [56] [57], especially in the field of UAV navigation. The reason for this proliferation is the increase in digital imagery technology as well as camera's light-weight nature makes them (visible or infra-red) suitable payloads for UAVs. Tracking with angles only to estimate a target's state vector is termed bearings-only tracking (BOT). As stated in Chapter 1, a passive sensor provides only angular measurements of a target, i.e. target azimuth and elevation. Therefore, the problem of BOT is a nonlinear one. As a result, EKFs [58] [59], UKFs [56] and PFs [60] are usually employed as the tracking filter.

A tracking system utilising only a passive angle-only sensor is not generally able to estimate the position of a target¹⁶ [25]. Despite this, it is possible to determine target states using multiple measurements (not in a single sensor observation) at different time instants. The problem geometry of classical BOT is illustrated in Figure 2.12. A further problem exists (see Figure 2.13) when a target's motion is purely radial, i.e. along the ownship line-of-sight (LOS). Here the target is moving, but the measured angle θ stays constant. A means of mitigating this effect is for the ownship (observer) to perform manoeuvres in order to take measurements when the target's motion is not radial [61]. This is only possible if the observer is manoeuvrable and is not generally considered an optimal solution.

¹⁶ In most cases single camera bearings-only tracking filters would diverge.

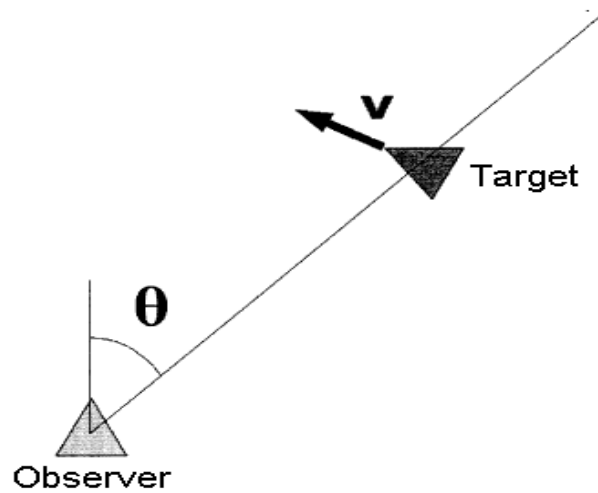


Figure 2.12: Bearings-only problem geometry [62].

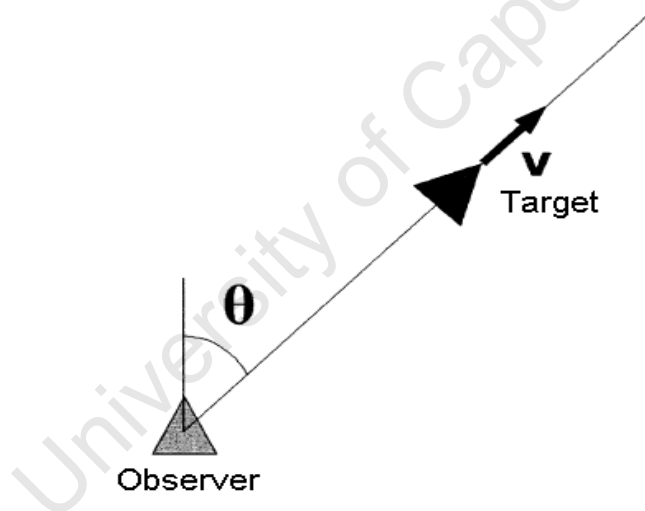


Figure 2.13: Target motion along observer LOS [62]. Angle theta remains constant in this scenario.

Notwithstanding this general consensus on bearings-only tracking, Bethke [62] proposes a solution to estimate the target state in a single sensor observation. This is achievable through the use of multiple, cooperative UAVs each equipped with its own EO camera. In order to determine the position and velocity of the target, multiple observations from multiple UAVs are

simultaneously combined (see Figure 2.14). This results in an unambiguous location of the target, regardless of target motion, found through the intersection of multiple rays from each observer to the target. Once the position is found, a linear KF filter is used to track the target's position and velocity. This solution, although novel and exciting, is not very practical. Yes, Bethke [62] presents acceptable results; but it is rare that so many resources would be available for a single mission. A less costly SAA solution would be more lucrative for industry.

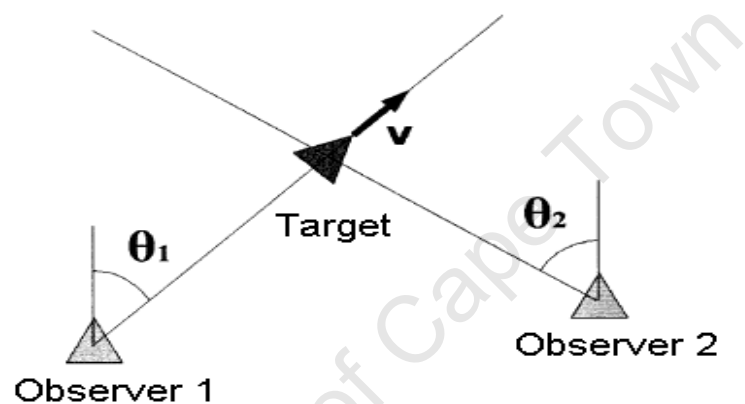


Figure 2.14: Using multiple UAVs to determine target position [62].

The following method describes an alternative to target tracking employing angles only measurements. It is technically not considered BOT, yet it does make use of angular measurements only. Watanabe [63] presents a tracking filter whose observation vectors are composed of detected pixels in the image plane.

In his system the existence of an image processing subsystem is assumed which detects potential threats in an image. Furthermore, the camera is assumed to be fixed to the centre of gravity of the UAV. Moreover, a pinhole¹⁷

¹⁷ For an excellent read on camera models, homographies and other image processing techniques, the reader is directed toward [86].

camera model is assumed as that in Figure 2.15, with the camera's X_C axis aligned with the principal (or optical) axis. Therefore the image plane is found at:

$$X_C = f. \tag{2.30}$$

Where f is the focal length of the camera. The origin of the camera frame is called the optical centre. The principal axis is the line originating from the optical centre, perpendicular to the image plane; and the point where they cut is the principal point.

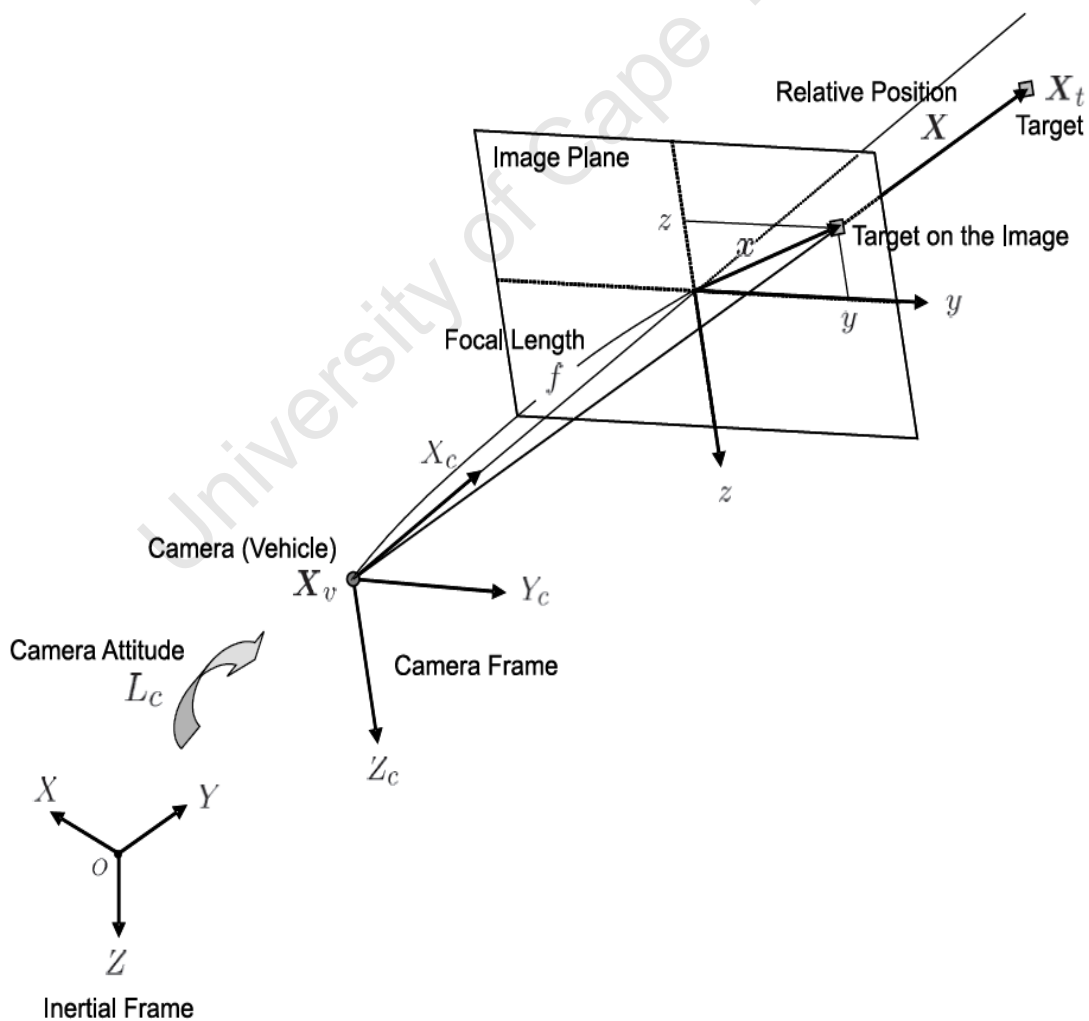


Figure 2.15: Pinhole camera model [63].

As a result of similar triangles (see Figure 2.16), the target position in the image at time step k is:

$$x_k = \begin{bmatrix} y_k \\ z_k \end{bmatrix} = \frac{f}{X_C^k} \begin{bmatrix} Y_C^k \\ Z_C^k \end{bmatrix}. \quad (2.31)$$

Following this, a transformation is performed so that estimation of the state vector may proceed in the inertial frame through the use of an EKF.

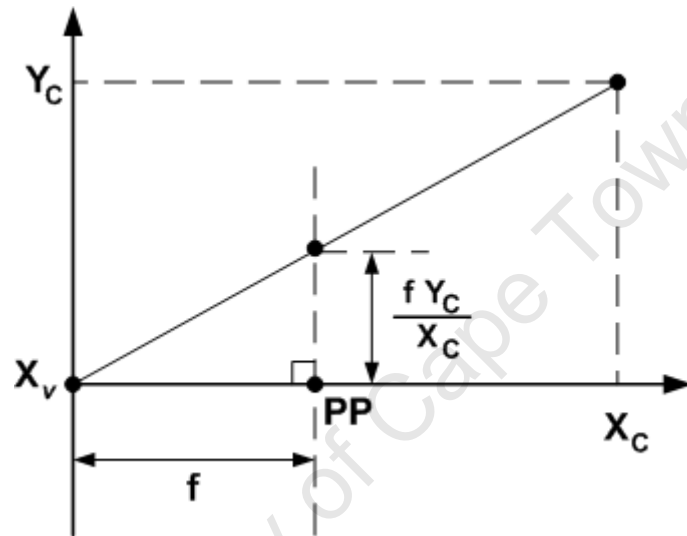


Figure 2.16: Projection of problem geometry in 2D.

Gupta [41] follow a similar approach as Watanabe [63], but performs tracking in the spherical coordinate system. Their reasons for this choice are similar to those laid out in Section 2.2.1. They state that the measurement equation is highly nonlinear in Cartesian space. By utilizing the EKF, the *corrector* stage results in a very complex Jacobian, which they cannot afford.

2.2.3 Tracking with Doppler Information

When an active Doppler radar is utilized in a tracking system, range rate (Doppler) information can be extracted from analyzing the frequency variation in the reflected signal (see Figure 2.17). One would assume that the inclusion of range rate measurements would result in better estimates produced by a

tracker, especially the velocity components. This is asserted by the results presented in [55], which indicate that tracking accuracy can be greatly improved with the use of Doppler information. This increased accuracy comes at the expense of increased filter complexity.

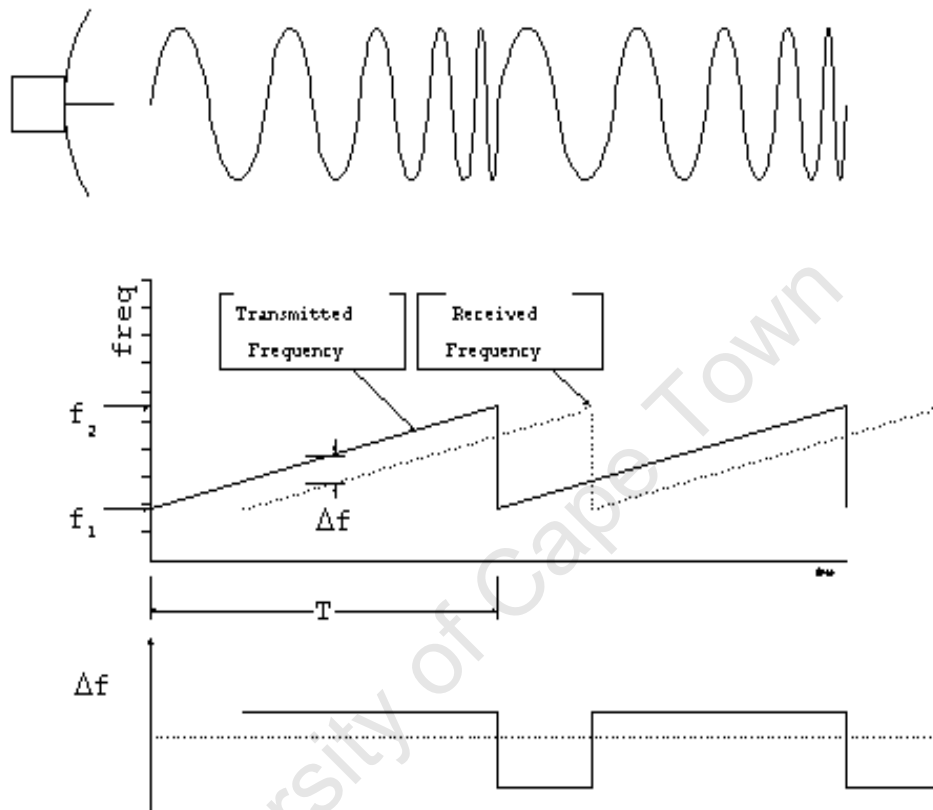


Figure 2.17: Frequency modulated continuous wave (FMCW) signal transmitted by a radar. The range rate is determined by the frequency difference in transmitted and reflected signals.

Notwithstanding the performance gain resulting from the use of Doppler information, range rate measurements are not always used in tracking filters. This is due to the measurements being highly nonlinear in Cartesian space. Moreover, large biases may be introduced in the posterior estimate [64].

Tracking with range rate measurements is further complicated in that measurement errors in range and range rate are correlated for certain radar waveforms [65]; with upsweep chirp Linear Frequency Modulation (LFM) waveforms there exists a negative correlation and with downsweep chirp LFM

waveforms there exists a positive correlation. The result of this can be seen in the radar measurement covariance matrix R with elements appearing off the main diagonal.

As a result of the above, there is not much focus in the literature on tracking with range rate measurements (especially so with regard to fusion – more on that in the subsequent section). Although complex, there are a number of methods in order to solve this problem. Duan et al. [44] describe that there are two ways of achieving this end:

- Tracking in mixed coordinates [65] [44] [42] [66].
- Tracking in Cartesian coordinates [38] [55] [67] [68].

Tracking in mixed coordinates implies that sensor measurements are used **directly** in the filter; that is the target's state vector and process noise are in Cartesian space while the sensor measurement and its measurement noise are in the sensor frame. This requires the linearization of the observation vector about the predicted estimate as described in subsection 2.1.2. Therefore, tracking is based on nonlinear filters such as an EKF. The disadvantage of this method is that the linearization process may introduce large errors in the mean and covariance of the posterior estimate. This equates to suboptimal filter performance and could result in the divergence of the tracking filter. Duan [55] state that tracking with mixed coordinates using an EKF usually results in poor tracking accuracy.

Bellotto and Hu [69] affirm the aforementioned point in [55] and state that, in general, better performance is achieved through the use of a UKF. The UKF predicts a state estimate by transforming an entire probability distribution. The difference between an EKF and a UKF is that the linearization process of the EKF is replaced with an unscented transformation. The unscented transformation, with carefully chosen weighted points¹⁸, captures the mean

¹⁸ These weighted points are also called sigma points.

and covariance of the probability distributions [69]. An alternative method of transforming probability distributions is to sample the distribution and then transform the samples separately; if the samples are drawn at random, then this approach is called Particle Filtering [70]. In contrast to PFs, UKFs make use of a smaller number of points and therefore make them less computationally intensive and, hence, more suitable to real-time applications [69].

A common assumption made when tracking with mixed coordinates and the availability of Doppler measurements is that measurement errors are uncorrelated. However, recent research proves that in reality this is not always the case [65] [44] [66]. As stated earlier, when tracking with Doppler information there exists a correlation in measurement errors between range and range rate. In an up-sweep chirp waveform the correlation coefficient ρ^{corr} is positive valued and implies negative correlation; whereas in down-sweep chirp, ρ^{corr} is negative in value and implies a positive correlation [66]. The general consensus in the literature is that when a negative correlation exists and is accounted for in the tracking filter, the result is increased tracking system performance; furthermore when ρ^{corr} is around -0.9 , significant reduction is evident in all the state estimation errors [65] [44] [66]. Bar-Shalom [66] extrapolates these findings by proving that tracking performance under positive correlation is worse than that when there is no correlation. However, Yuan et al. [65] build on the work presented in [66] and prove, via simulations, that tracking performance is not always worse in the positive correlation case as opposed to the no correlation case. In what follows, the assumption is that a radar exists and provides the tracking system with measurements of target range, azimuth, elevation and range rate denoted by ρ_k , θ_k , φ_k and $\dot{\rho}_k$, respectively.

When measurements errors in range and range rate are correlated the resulting measurement covariance is:

$$R = \begin{bmatrix} \sigma_\rho^2 & 0 & 0 & \rho^{corr} \sigma_\rho \sigma_{\dot{\rho}} \\ 0 & \sigma_\theta^2 & 0 & 0 \\ 0 & 0 & \sigma_\varphi^2 & 0 \\ \rho^{corr} \sigma_\rho \sigma_{\dot{\rho}} & 0 & 0 & \sigma_{\dot{\rho}}^2 \end{bmatrix}. \quad (2.32)$$

In order to decorrelate the measurement errors, let [44]:

$$\begin{aligned} L &= -[\rho^{corr} \sigma_\rho \sigma_{\dot{\rho}}] \times ([\sigma_\rho \quad \sigma_\theta \quad \sigma_\varphi])^{-1} \\ &= \begin{bmatrix} -\rho^{corr} \sigma_{\dot{\rho}} / \sigma_\rho & 0 & 0 \end{bmatrix}. \end{aligned} \quad (2.33)$$

Then:

$$B = \begin{bmatrix} I_3 & 0_{3 \times 1} \\ L & 1 \end{bmatrix}. \quad (2.34)$$

Where I_3 is the 3×3 identity matrix and $0_{3 \times 1}$ is a 3×1 vector of zeros.

By first performing Cholesky factorization [39] and then pre-multiplying Equation (2.34) on both sides of the observation vector, a new pseudo measurement results (while the range, azimuth and elevation remain unchanged):

$$\varepsilon_k = h_\varepsilon(X_k) + v_k^\varepsilon \quad (2.35)$$

$$\varepsilon_k = \frac{-\rho^{corr} \sigma_{\dot{\rho}} \rho_k}{\sigma_\rho} + \dot{\rho}_k \quad (2.36)$$

$$h_\varepsilon(X_k) = -\rho^{corr} \frac{\sigma_{\dot{\rho}}}{\sigma_\rho} \sqrt{x_k^2 + y_k^2 + z_k^2} + \frac{x_k \dot{x}_k + y_k \dot{y}_k + z_k \dot{z}_k}{\sqrt{x_k^2 + y_k^2 + z_k^2}} \quad (2.37)$$

The measurement noise of the new pseudo measurement is now uncorrelated with all other measurement noises and is still a zero-mean Gaussian white noise sequence with variance:

$$\sigma_\varepsilon^2 = (1 - \rho^{corr^2}) \sigma_{\dot{\rho}}^2. \quad (2.38)$$

In contrast to mixed coordinate tracking, tracking in Cartesian-only space necessitates a conversion of sensor measurements from the sensor frame¹⁹, in which they originate, to Cartesian space. This conversion is the basis of all Cartesian-only tracking. The conversion of measurements present a number of flaws [44]: converted measurement errors are state dependent; and, converted measurement sequences are not white anymore.

A means of mitigating this effect and consequently a method of tracking in Cartesian space with range rate measurements, is that performed by Duan et al. [55]. Their work follows on from that of [67]. Suchomski [67] presents a method of tracking in Cartesian space with a linear filter when a sensor is providing nonlinear measurements. This method extends that of prior research [38], which only operated in 2-dimensions, and realises a tracking filter that is based on debiased consistent converted measurements. Duan et al. [55] modify the former by including range rate measurements. Furthermore, through the formation of a pseudo measurement η_k that is the product of the range and range rate measurements:

¹⁹ Sensor measurements are generally either described in polar or spherical coordinates.

$$\eta_k = \rho_k \dot{\rho}_k = x_k \dot{x}_k + y_k \dot{y}_k + z_k \dot{z}_k. \quad (2.39)$$

The presence of this pseudo measurement reduces the nonlinearity of the measurement equation with respect to the state vector. The first step that needs to be performed is the conversion of radar measurements to Cartesian space:

$$x'_k = \rho_k \cos \theta_k \cos \varphi_k \quad (2.40)$$

$$y'_k = \rho_k \sin \theta_k \cos \varphi_k \quad (2.41)$$

$$z'_k = -\rho_k \sin \varphi_k \quad (2.42)$$

Following this, these converted measurements are debiased as follows:

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} x'_k \left[1 - \left(e^{-\sigma_\theta^2} e^{-\sigma_\varphi^2} - e^{-\frac{\sigma_\theta^2}{2}} e^{-\frac{\sigma_\varphi^2}{2}} \right) \right] \\ y'_k \left[1 - \left(e^{-\sigma_\theta^2} e^{-\sigma_\varphi^2} - e^{-\frac{\sigma_\theta^2}{2}} e^{-\frac{\sigma_\varphi^2}{2}} \right) \right] \\ z'_k \left[1 - \left(e^{-\sigma_\varphi^2} - e^{-\frac{\sigma_\theta^2}{2}} \right) \right] \end{bmatrix}. \quad (2.43)$$

Where σ_θ and σ_φ are the measurement errors in azimuth and elevation, respectively.

The debiased consistent converted measurement vector $[x_k \ y_k \ z_k \ \eta_k]^T$ has a covariance matrix R_k [67]:

$$R_k = \begin{bmatrix} & R_a & & \sigma_{x\eta} \\ & & & \sigma_{y\eta} \\ & & & \sigma_{z\eta} \\ \sigma_{x\eta} & \sigma_{y\eta} & \sigma_{z\eta} & \sigma_\eta^2 \end{bmatrix}. \quad (2.44)$$

Where R_a is defined as:

$$R_a = \begin{bmatrix} R_a^{xx} & R_a^{xy} & R_a^{xz} \\ R_a^{xy} & R_a^{yy} & R_a^{yz} \\ R_a^{xz} & R_a^{yz} & R_a^{zz} \end{bmatrix}. \quad (2.45)$$

Appendix A lists the full measurement covariance matrix and related equations when using the debiased consistent converted measurements KF method.

Another method of tracking in Cartesian coordinates is that presented in [68]. Wang et al. [68] propose a method of sequentially processing the range and range rate measurements. This correlates with the findings of Blackman and Popoli [25] as they state that the inherent nonlinearity in the system can be reduced by first processing the least nonlinear measurement first.

Zollo and Ristic present another tracking filter in [53]; this time tracking a manoeuvring target with spherical coordinates. The filter displays satisfactory tracking performance and is one that is found in a number of operational systems [71]. The system comprises three linear KFs; one for range/range rate, and three for the direction cosines²⁰ (DCM) in the north, east and down directions. The state vector of the range/range rate filter is:

²⁰ The pointing angle of sensors as well as aircraft attitude is generally described with direction cosines relative to the NED axis.

$$X_\rho = \begin{bmatrix} \rho \\ \dot{\rho} \\ \ddot{\rho} \end{bmatrix}. \quad (2.46)$$

And the state vector for the north filter is:

$$X_N = \begin{bmatrix} \Lambda_N \\ \dot{\rho}_N \\ \ddot{\rho}_N \end{bmatrix}. \quad (2.47)$$

Another two state vectors exist (not shown here) for the east and down directions. Λ_N , Λ_E and Λ_D are the direction cosines in the north east and down directions respectively. Components of target position in the north, east and down directions can be found as follows:

$$\rho_N = \rho \Lambda_N \quad (2.48)$$

$$\rho_E = \rho \Lambda_E \quad (2.49)$$

$$\rho_D = \rho \Lambda_D \quad (2.50)$$

The reader is urged to consult [53] for the remainder of the mathematics required for this scenario. Furthermore the derivation of the target dynamics used in [53] can be found in [24]. Once the direction cosines are known target angular position with respect to the NED system can be represented by *stabilised* azimuth η and *stabilised* elevation ε :

$$\eta = \tan^{-1} \left(\frac{\Lambda_E}{\rho \Lambda_N} \right) \quad (2.51)$$

and

$$\varepsilon = -\sin^{-1} \Lambda_D. \quad (2.52)$$

In summary, most of the literature reviewed above track targets in Cartesian space as it is generally considered that it yields more accurate filters but is

more computationally expensive [53]. However, these filters do not track in inertial NED, which would be most convenient for airborne platforms. Furthermore, it is assumed that the sensor, providing measurements, is stationary. Blackman and Popoli [25] claim that the use of acceleration states in the state vector is required, for satisfactory performance if a velocity measurement, such as range rate, is available. However, the literature described above, as well as the designs presented later in this thesis prove that that is not always the case. Moreover, tracking in mixed coordinates is considered to yield poor accuracy, but through fusion of multiple sensor's data we prove the contrary.

2.2.4 Tracking in the Presence of Multiple Sensors

Tracking a target while utilising measurements from multiple, redundant sensors is a fairly novel technique. There is definitely less literature on the matter compared to that of the proliferation on radar tracking. More specifically, there has been little discussion in the literature about fusing data from a Doppler radar and a passive angle-only sensor for tracking.

The presence of multiple sensors in a tracking system results in registration error [38]; a fundamental issue when considering the implementation of a MSDF system, yet no literature could be found which addresses the problem of time registration. All the literature that has been investigated makes the assumption that all sensors present in the system have the same sampling frequency.

Two types of registration error exist, namely spatial and temporal. Space registration addresses the fact that multiple sensors provide measurements with respect to their own sensor frame. The solution is found through pre-processing of the sensor data by performing the necessary coordinate transformations and conversions. This ensures that each sensor's

measurements are with respect to a single frame of reference. Time registration addresses the fact that measurement rates, from multiple independent sensors, are asynchronous, i.e. they have different sampling periods. In our case an EO camera provides measurements of:

$$z_k^C = [\theta_k^C \quad \varphi_k^C]^T. \quad (2.53)$$

At a faster rate than the radar provides its measurements:

$$z_k^R = [\rho_k^R \quad \theta_k^R \quad \varphi_k^R \quad \dot{\rho}_k^R]^T. \quad (2.54)$$

To solve this problem and to synchronise the measurements, data compression is performed. Data compression combines a number of observations (measurements) into a single observation that is a weighted average of the component observations; where more recent observations are weighted more heavily [25]. The combined measurement (z_{DC}^C) is found using a least squares curve fit to the N^{21} observations:

$$z_{DC}^C = \frac{1}{N} \sum_{i=1}^N z_i^C + \frac{12a}{(N-1)(N+1)} \sum_{i=1}^N \left(i - \frac{N+1}{2}\right) z_i^C. \quad (2.55)$$

Assuming the EO camera has a measurement covariance matrix R^C , the composite camera measurement covariance matrix R_{DC}^C is evaluated as:

$$R_{DC}^C = \frac{1}{N} (1 + 12a^2) R^C. \quad (2.56)$$

²¹ N is the ratio $\Delta T^R / \Delta T^C$.

The factor α in Equation (2.55) and (2.56) determines whether older or more recent measurements should be more heavily weighted. It lies in the region of $(0, 0.5)$ where a larger value more heavily weights recent observations.

When designing a MSDF system the estimation engineer needs to decide on the fusion architecture to make use of, namely centralised, distributed or hybrid. Centralised fusion combines multiple sensor observations (raw sensor data), whereas distributed fusion combines multiple tracking filters' states estimates. When a distributed architecture is employed, the main issues that need to be addressed are:

- Architectural structure. How sensor nodes are connected and how they share information.
- Tracking filters employed.
- Fusion algorithms. The way in which data is fused from each node.

These issues, consequently, are what differentiate researchers' MSDF systems.

Tracking with multiple sensors is then a problem involving the proper organisation and use of the tools and techniques discussed in Sections 2.1 and 2.2 thus far. There are a multitude of architectures and combinations of sensors found in MSDF literature. A subset of these includes:

- Distributed fusion of a Doppler radar and an IR sensor [29] [72].
- Fusion of radar and LIDAR [73].
- Fusion of radar, IR and LIDAR [74].
- Fusion of laser range finder (LRF) and EO camera [69].
- Fusion of radar and EO sensors [70].
- Fusion of EO camera, LIDAR and inertial measurement unit (IMU) sensors [75].

In [29], Qingchao and Wenfei propose a MSDF system that comprises two distributed local tracking filters: one EKF for the radar subsystem and one

EKF for the IR sensor subsystem. A pseudo measurement is formed by augmenting the radar and IR observation vectors; the range rate component is decoupled from the radar and concatenated to the IR observation vector. The advantage in creating this pseudo measurement is that measurement errors in range and range rate are not correlated anymore. Moreover, the increased dimensionality of the IR observation vector increases state observability. The fusion architecture is illustrated in Figure 2.18. The globally fused estimate and covariance is fed back to the IR subsystem only; termed partial feedback. The validity of this technique is affirmed by Zhu et al. in [76], where they state that by performing feedback, the accuracy of the local filter is improved. According to Qingchao and Wenfei [29] their reason for performing partial over full feedback is that, in so doing, the whole tracking system is not reliant on the fusion centre.

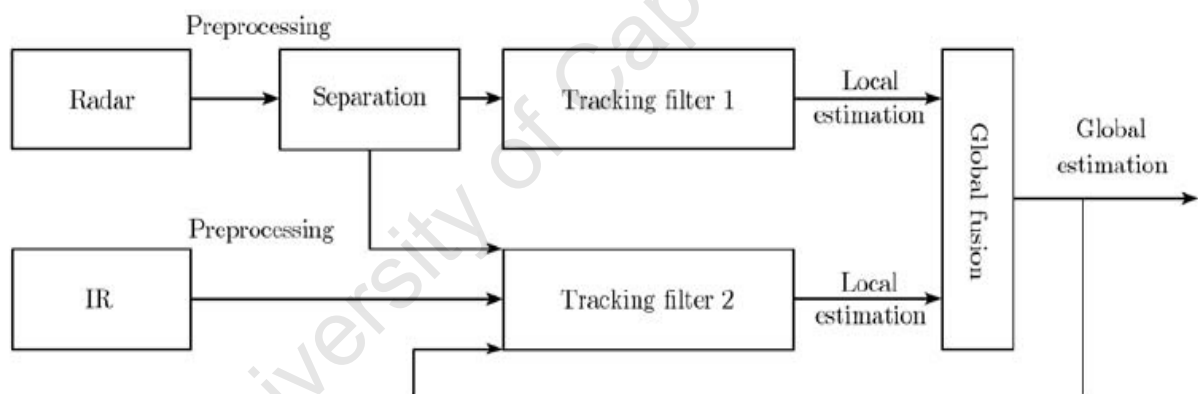


Figure 2.18: Qingchao and Wenfei proposed fusion architecture [29].

Once the local filters (Tracking filter 1 and Tracking filter 2) each produce their own local estimate, they are fused using Equations (2.6) and (2.7). Furthermore, two other variations of the architecture in Figure 2.18 are implemented: one variant where no range rate measurements are available and no separation occurs; and one variant with range rate measurements but no separation (radar observation vector used as is). They have ascertained, through simulations, that the variant using range rate measurements and no separation results in an error decrease of 21.68% compared to the no range

rate variant. Furthermore, their proposed architecture (Figure 2.18) sees a further 3.75% error reduction. However, this design does not take into account that radar and IR sensors operate at different frequencies. In order to account for this a technique such as data compression should be performed. Furthermore, it is assumed that the platform containing the sensors is stationary.

In [72] a similar design is presented as in [29], but with a few differences. Firstly, the radar does not measure relative radial velocity. Furthermore, instead of an EKF for the IR local filter, the PLKF method (see Section 2.1.2) is used to derive the filter. Local estimates are fused in a similar fashion to that of [29]. Following fusion, feedback is performed by setting the global estimate as the local estimates for the next iteration of the filter. However, in contrast to [29], here the information distribution coefficient β^i is introduced which proportionally feeds the global estimate back. Performance is compared to that of centralised fusion with radar and IR sensors as well as radar and IR only tracking. The results prove that the proposed distributed approach not only outperforms individual sensor tracking, but tracking performance is comparable to that of centralised fusion. Furthermore, the distributed architecture reduces the system computational load and exhibits good real-time performance.

Blanc et al. [73] also proposes a distributed MSDF system that consists of a LIDAR and a radar with EKFs as each local filter. However, they do not make the assumption as in [29] and [72] that local state estimates are uncorrelated. As a result Equations (2.8) to (2.11) are used in order to fuse the local estimates. Furthermore, Blanc et al. [73] presents a method of fusing local state estimates using a PF. Their PF fusion is based on: first, calculation of particle weights from different sensor measurements; and second, validation of the fused estimate by the local estimates. They conclude that the PF method and EKF (i.e. standard distributed fusion) exhibit similar performance, however the EKF approach is much less computationally expensive.

A centralised fusion architecture comprising of radar, IR and EO sensors is developed by Fasano in [48]. The proposed architecture (see Figure 2.19) has the radar as the primary sensor and IR and EO as aiding sensors; this is because only the radar is used to update a target's track status.

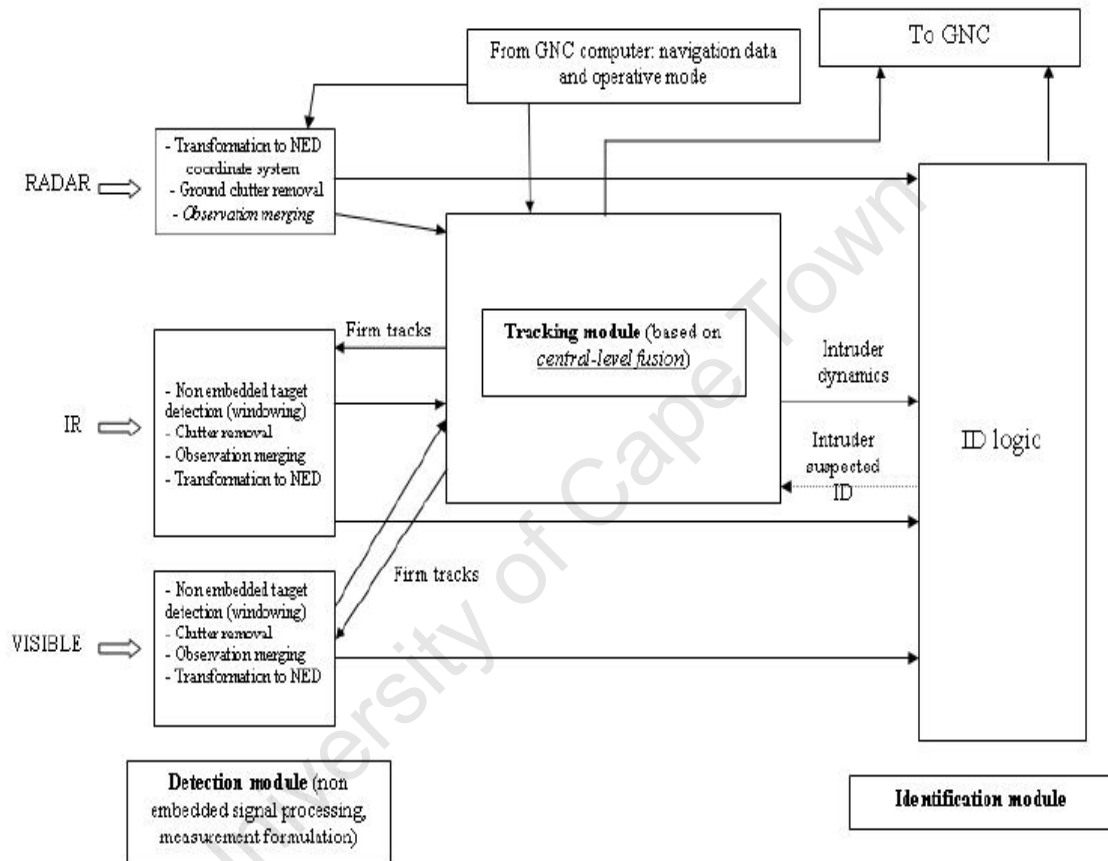


Figure 2.19: Sensor fusion logical architecture implemented in [48]. Centralised fusion consisting of radar, IR and EO sensors.

The EO camera uses information from the tracker in order to create a valid search window; the target's predicted estimate is utilised in order to define a region of interest in the image where valid target detection may occur. Once a target is detected, the necessary coordinate transformations are performed and the detected target position is transferred to the FC. This algorithm is illustrated in Figure 2.20. Finally centralised fusion is performed with

measurements from radar, EO and IR sensors in the FC in order to track a target.

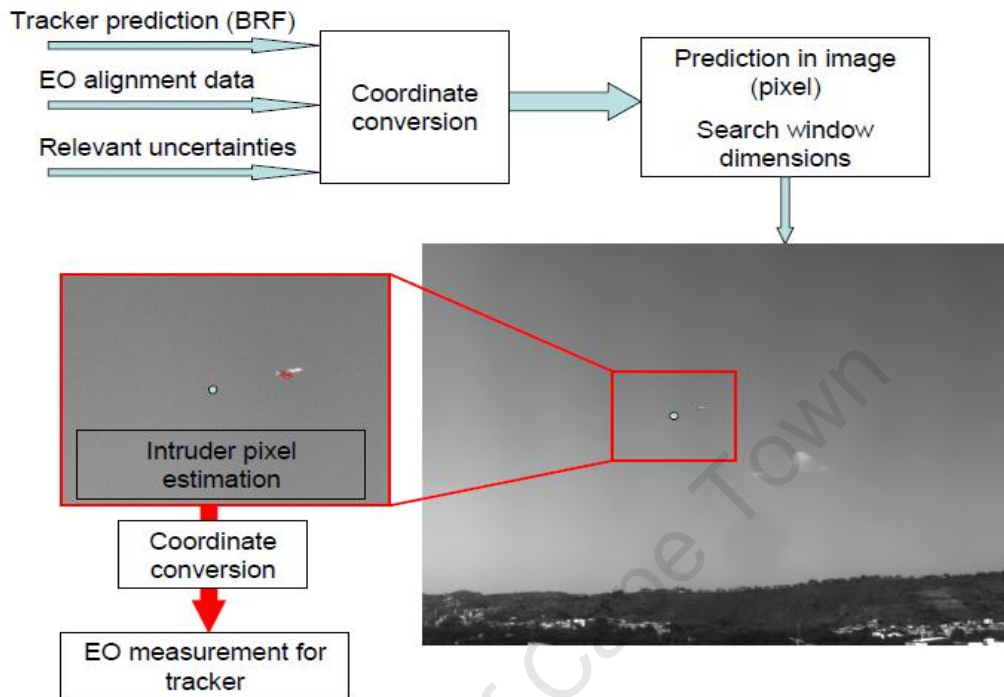


Figure 2.20: EO camera target detection algorithm [48].

In [77], Leibowicz et al. realise a system that consists of radar and ESM sensors. Their motive for fusion is not in search of improved filter accuracy, but rather to avoid alerting the enemy of one's presence. This is achieved by limiting the radar's activity at a specified time while allowing target tracking to continue by utilising only the passive sensor.

2.3 Summary

The review of literature stressed the need to have in place a tracking system that is accurate and robust enough for a UAV SAS, which would be a component in enabling its timely adoption into civil airspace. This vision will only be possible through the appropriate use of MSDF. Thus, it is evident from this literature review that target tracking using MSDF is by no means a trivial task. It is one that encompasses numerous techniques from various

engineering disciplines [14]. In summation, when designing a MSDF system the following needs to be addressed:

- Choice of tracking coordinate system.
- Choice of fusion architecture.
- Fusion algorithm selection.

In order to perform these tasks, a sound understanding of MSDF techniques and estimation fundamentals are required. First a definition of MSDF was produced, highlighting its similarity to the cognitive process evident in humans. The use of multiple, redundant sensors have many practical and operational advantages and these were discussed through the aid of an example. These advantages are: imperviousness to ECM; all-weather, all-day operation; improved overall tracking performance; and extended spatial and temporal coverage. That being said, Hall and Steinberg [26] warn that MSDF is not magic and that considerable effort is required to get the best performance out of a MSDF system.

Based on the provided review there is a need for research on target tracking with Doppler information in MSDF systems. Moreover, an even greater need for research on fusion architectures for a tracking subsystem that forms part of a UAV SAS; research that will be key for UAV's adoption in civil airspace.

To address this need, empirical research will be implemented. Specifically, such research will attempt to ascertain the most appropriate fusion architecture for a UAV SAS. As such multiple fusion architectures will be developed as well as radar and camera only trackers in order to bring to light the performance gains from performing MSDF. The next stage of this research will detail the methodology for the design of the fusion architectures, including details on the research strategy adopted, presentation of metrics with which to assess the designs and the development of the simulation environment.

Chapter 3: Methodology

The development of a tracking algorithm using MSDF is no easy task. Moreover, the development of an airborne one is considerably more complex. The focus of this chapter is to describe the strategy employed in this thesis in order to realise an airborne tracking filter. Sound engineering principles are adhered to and details of testing the proposed designs are discussed. This entails presenting performance metrics with which system performance can be ultimately measured. It also provides motivation for the components designed in Chapter 5..

3.1 Overview of Research Strategy

Chapter 2 identified a number of gaps in existing research: there is a lack of fusion literature which includes the use of a Doppler radar; the amount of literature on airborne tracking is limited; and, MSDF is a fairly new field of research which requires active research and development for its maturity.

The research strategy employed in this thesis is akin to the model based design (MBD) approach applied to research. A benefit of MBD is that it allows for the test and verification of designs early on in its life-cycle [78], when the cost of change is minimal; therefore, it has seen extensive use in the automotive and aerospace industries [78]. The MBD methodology is outlined below [79] and illustrated in Figure 3.1:

- **Research**

The first step in this approach is to perform a study of relevant literature.

- **Requirements Definition**

Based on the results and findings from the research phase, system requirements are defined.

- **Modelling the Process/Plant**

The process under investigation as well as the environment with which it interacts is modelled.

- **Algorithm Design**

This phase is responsible for the design and implementation of algorithms, using the information from the process model and its environment, which satisfy the outputs from the Requirements Definition phase.

- **Simulation**

Simulations are performed which test the algorithms in the modelled environment. Furthermore system evaluation is performed by defining performance metrics and assessing system performance based on them. At the end of the Simulation phase, the Algorithm Design phase can be performed again in order to increase system performance.

- **Deployment**

The last phase is the deployment on to target hardware. This involves integration with other algorithms or subsystems and, finally, placement on an embedded platform.

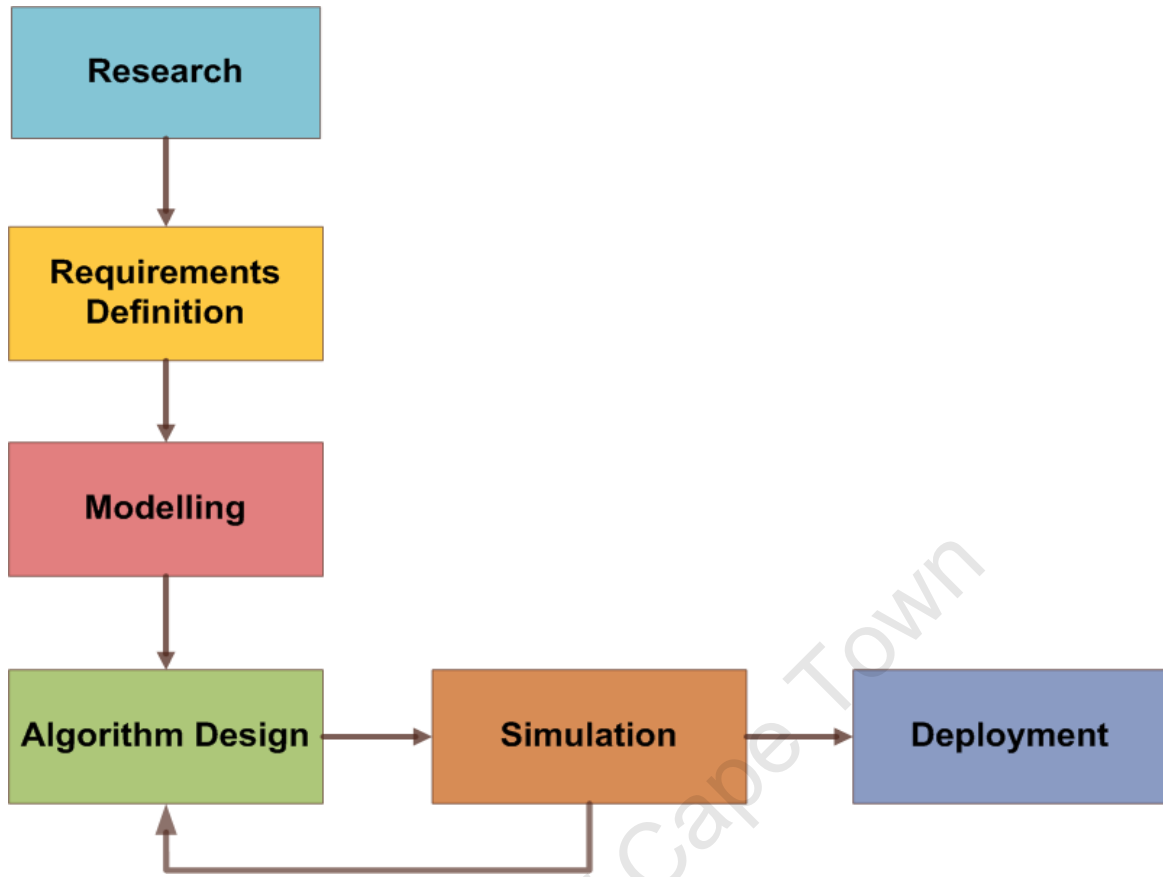


Figure 3.1: MBD approach.

The reason why this approach was adopted is twofold: first, it requires the modelling of the simulation environment early on, a very important aspect when dealing with dynamic systems. And second, in Chapter 5: multiple fusion architectures and filters are developed, which are assessed to determine the most suitable one for an airborne platform; this strategy works well for this iterative develop-simulate-test process. In what follows, how this MBD approach was applied to this thesis is detailed.

3.2 Research and Requirements Definition

A background study was performed in Chapter 1 which highlighted the need for the TTU. The problem description then justified this reasoning. Following this an exhaustive literature survey was performed in Chapter 2 which

reviewed current literature and various approaches in order to arrive at a solution to the problem at hand. The aforementioned points as well as the requirements from both [2] and [9], led to the definition of the thesis objectives detailed in section 1.4.

3.3 Modelling and Simulation

The modelling of the simulation environment is a crucial part in the design of the system. The reason for this is that the system developed is placed in this environment where it undergoes rigorous testing and verification. The plant is the scenario of airborne threat detection, from when a sensor on-board the UAV detects the presence of a threat until it is no longer in its FOV (i.e. the tracking period). This entails the mathematical modelling of both the UAV and threat dynamics in the inertial NED system. Furthermore sensors need to be modelled; however, sensor measurements appear in the BRF (as we assume that the sensor is at the centre of gravity of the UAV). Moreover sensor measurements are of relative²² dynamics. This measurement process as well as the mathematical modelling of UAV and threat dynamic models is addressed in section, which describes the simulation environment.

3.4 Algorithm Design

In order to design a suitable airborne multisensor tracking filter a sound foundation of tracking and filter performance, under various conditions, is required. As such, the approach taken to arrive at the final fusion architecture starts with the development of a fairly simple filter (with no sensor data fusion) on which incremental improvements and modifications are made. This allows verification and performance of each filter, which has been developed, to be analysed at every incremental stage. Moreover, the filters produced at each incremental stage become the building blocks for the final fusion architecture.

²² Relative dynamics is defined as the vector difference between threat and ownship (UAV) dynamics.

The design of the tracking filters and fusion architectures form the core of Chapter 5. Furthermore each of the designed filters shall track relative dynamics in the inertial NED system.

3.5 Testing the Threat Tracking Unit and Performance Metrics

This section details how each filter developed in Chapter 5 is to be properly tested. First the results from a single flight scenario are presented. This allows visual inspection to be performed to determine whether a filter is performing as expected. To determine by how much, Monte Carlo simulations are performed and a number of performance metrics presented, the result of which will be discussed on a quantitative level to assess their significance in Chapter 6. The following tests and metrics are presented in this section:

1. Visual evaluation.
2. Filter Accuracy Test
 - Root mean squared error (RMSE) in position.
 - Root mean squared error (RMSE) in velocity.
3. Filter Credibility Test
 - Average normalised estimation error squared (ANEES).
4. Tracking filter consistency test. This entails:
 - Innovation consistency test.
 - Innovation autocorrelation test.

To clarify the notions of filter consistency and efficiency, the following definitions are presented: a **consistent** filter is one that provides an increasingly accurate state estimate over time; and an **efficient** filter is one that minimises the MSE [37].

3.5.1 Visual Evaluation

Filter performance is firstly visually evaluated by presenting the estimation error of position and velocity components in NED. Intuitively, the closer these values are to zero, the more accurate the filter is. Furthermore, filter state covariance is illustrated which provides a measure of estimate uncertainty. Along the diagonal of the state covariance matrix is the principal uncertainties in each of the state vector elements. At each time-step with the arrival of a new measurement and the performance of the *corrector* stage, the state covariance is expected to decrease, and eventually reach steady-state.

3.5.2 Filter Accuracy

By performing large-scale Monte Carlo simulations (MCS), we have the ability to gain great insight into filter performance. The first metric determined during MCS is the RMSE in position and can be described as:

$$RMSE_k^{pos} = \sqrt{\frac{1}{N^{MC}} \sum_{i=1}^{N^{MC}} \left(\begin{array}{l} (N_k^i - \hat{N}_{k|k}^i)^2 + \\ (E_k^i - \hat{E}_{k|k}^i)^2 + \\ (D_k^i - \hat{D}_{k|k}^i)^2 \end{array} \right)}. \quad (3.1)$$

Where N^{MC} is the number of Monte Carlo runs; the vectors $[N_k^i \ E_k^i \ D_k^i]$ and $[\hat{N}_{k|k}^i \ \hat{E}_{k|k}^i \ \hat{D}_{k|k}^i]$ are the target's true position and position estimate in inertial NED, respectively. Similarly the velocity RMSE is:

$$RMSE_k^{vel} = \sqrt{\frac{1}{N^{MC}} \sum_{i=1}^{N^{MC}} \left(\begin{array}{l} (\dot{N}_k^i - \hat{\dot{N}}_{k|k}^i)^2 + \\ (\dot{E}_k^i - \hat{\dot{E}}_{k|k}^i)^2 + \\ (\dot{D}_k^i - \hat{\dot{D}}_{k|k}^i)^2 \end{array} \right)}. \quad (3.2)$$

The RMSE in position and velocity gives a physical measure of the accuracy of the tracking filter.

3.5.3 Filter Credibility

As stated in Chapter 2, the filter estimated state covariance $P_{k|k}$ provides a self-assessment of how well a filter is performing. However, this self-assessment is not always reliable due to a number of reasons. The level of reliability (or how true) the estimated covariance is, is termed filter credibility. A means of determining credibility is through the calculation of the ANEES.

The ANEES is determined by the following:

$$ANEES_k = \frac{1}{NMCn} \sum_{i=1}^{NMC} (X_{k|k}^i - \hat{X}_{k|k}^i)' P_{k|k}^{-1} (X_{k|k}^i - \hat{X}_{k|k}^i). \quad (3.3)$$

Where n is the dimension of X and $(X_{k|k}^i - \hat{X}_{k|k}^i)$ is the estimation error.

If the estimation error²³ is close to the estimated covariance, the ANEES will be approximately 1 and the filter is credible [80]. Moreover, the closer the ANEES is to 1, the more credible the filter. If the filter error covariance estimate is larger than the estimation error then the filter is said to be *pessimistic*, alternatively the filter would be deemed *optimistic*. This can be visualized by plotting the ANEES along with its 95% confidence interval. When the ANEES is above the 95% interval the filter is optimistic and when below, pessimistic.

3.5.4 Filter Consistency

The metrics defined above, viz. RMSE and ANEES require knowledge of **true** threat states. However, in contrast to simulation, practical real-world tracking systems are not aware of true data – there is no way that this data can be made available to the system. Therefore, evaluation of system performance in

²³ Estimation error is defined as the difference between the true state vector and its estimate (produced by the filter).

these pragmatic scenarios requires innovation analysis. In section 2.1.2 it was stated that the innovation/residual is an important factor in the KF and provides a measure of how well the filter is performing; the following two tests (presented in [81]) prove this:

- Innovation consistency test.
- Innovation autocorrelation²⁴ test (test for whiteness).

If the filter is performing correctly then the innovation sequence is zero mean and white with covariance S_k . Therefore, if the innovation sequence is consistent and white then the filter is consistent.

3.5.4.1 Innovation Consistency Test

The consistency test determines if the innovation is consistent. This test involves determining whether the innovation sequence is bounded by $\pm 2\sigma$, where:

$$\sigma = \sqrt{S_k}. \quad (3.4)$$

If approximately 95% of the values of the values lie within the $\pm 2\sigma$ bound, then the innovation sequence is consistent.

3.5.4.2 Innovation Autocorrelation Test

The autocorrelation test tests the innovation sequence for whiteness. Assume that the innovation sequence is denoted by v_k , then by performing MCS the autocorrelation is:

²⁴ From statistical theory, a white noise sequence is independent and identically distributed and, therefore, has zero autocorrelation [39].

$$r(\tau) = \frac{1}{N^{MC}} \sum_{k=0}^{N^{MC}-\tau-1} v_k^T v_{k+\tau}. \quad (3.5)$$

Equation (3.5) is then normalised by dividing $r(\tau)$ by $r(0)$ – this yields the normalised autocorrelation of the innovation. In order to prove whiteness the normalised autocorrelation must be randomly distributed about zero. To prove that the normalised autocorrelation is in fact random we first need to compute its $\pm 2\sigma$ bound. For large enough N^{MC} it can be assumed that $r(\tau)$ is zero mean with covariance $1/N$. Then the standard deviation is given by:

$$\sigma = \sqrt{1/N^{MC}}. \quad (3.6)$$

Now it can be verified whether the normalised autocorrelation has 95% of its values in the $\pm 2\sigma$ bound; if it does then the innovation sequence v_k is white.

If the innovation sequence was proved consistent and white from the above two tests, then the tracking filter is consistent.

Now that a platform for the evaluation of proposed designs has been established, thesis objective 6 has been satisfied; and the design of the TTU can commence.

Chapter 4: Simulation Environment

Modelling of the plant and the simulation environment needs to occur before the system can be designed and tested, as stated in section 3.2. Moreover, the TTU requires a simulation environment in which it can be tested to be developed. Therefore, this chapter presents the modelling of the simulation environment. This entails the mathematical modelling of the UAV and threat dynamic models as well as their generation in inertial space. Sensor measurements are generated in BRF and this process is explained. Moreover, using the above information, the flight scenario is described.

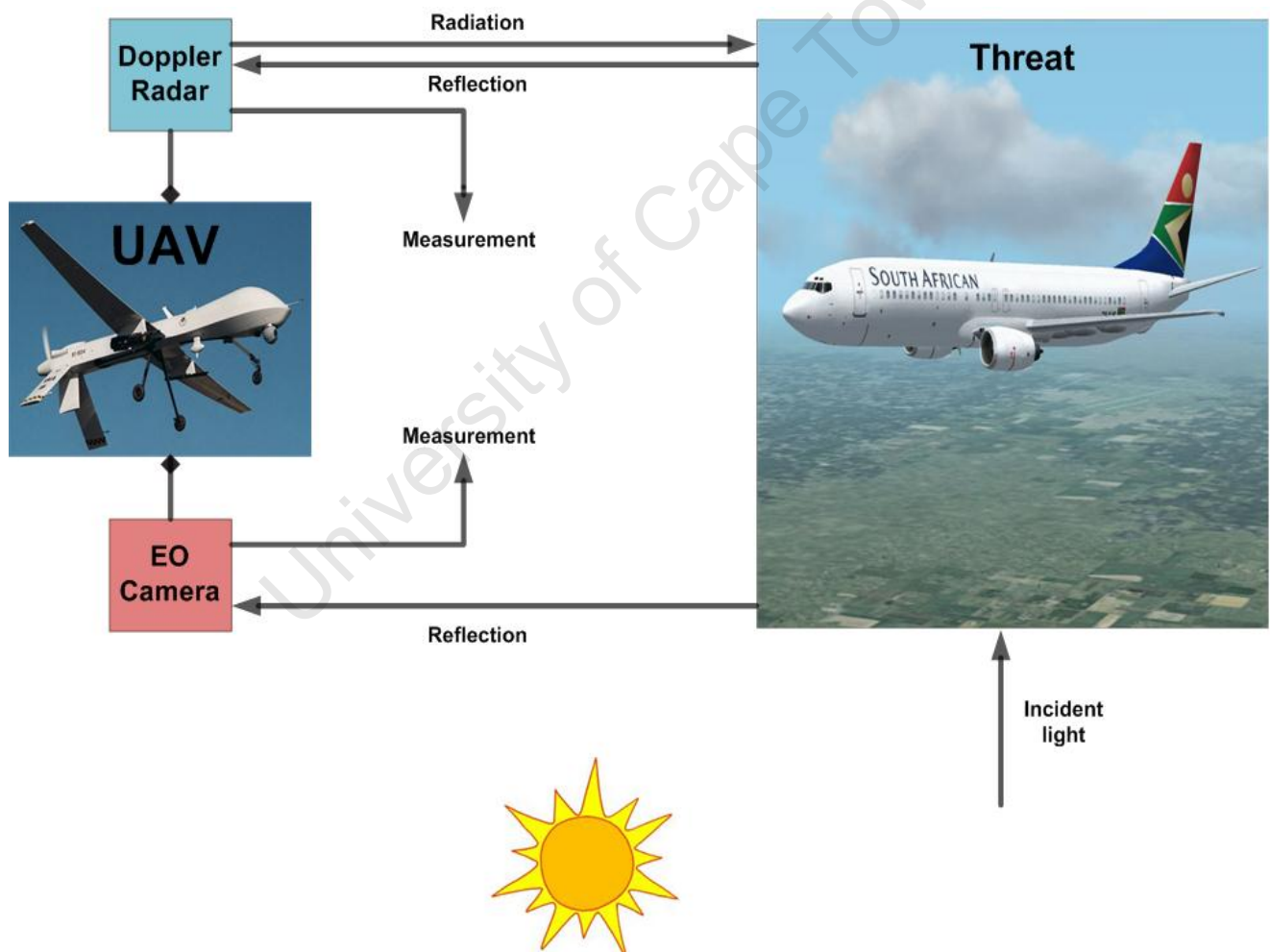


Figure 4.1: Simulation environment to be modelled.

4.1 States of Interest

The proper modelling of the simulation environment is critical to the success of the TTU design (see Chapter 3). First mathematical models for threat and UAV are presented. This information is used to generate threat and ownship trajectories in inertial space. Then the process of modelling Doppler radar and EO camera sensor measurements in inertial space is explained. Finally, the flight scenario used in the simulations is presented.

For the design of the TTU, the quantities of interest are the threat's position and velocity. The velocity components are of importance as the control system performing CA requires accurate threat velocity estimates [79] (see Figure 1.10). However, sensor **relative** measurements arrive at the TTU in the BRF with spherical coordinates.

The relationship between UAV, threat and relative motion is defined as:

$$X^{rel} = X^{threat} - X^{UAV} \quad (4.1)$$

and

$$V^{rel} = V^{threat} - V^{UAV}. \quad (4.2)$$

Where X and Y are position and velocity vectors respectively.

At this juncture it is possible to either select the BRF or NED frame for tracking. The simplest solution is to track relative motion in BRF, extract the threat's contribution (using Equations (4.1) and (4.2)) and then transform this to NED. However, things are not always simple; the literature review (in Chapter 2) revealed that relative motion in BRF depends on attitude (angular rates), unlike its projection in NED [48]. Therefore our objective is to track target motion in NED.

As stated before the measurements received from sensors are of relative motion in BRF. Therefore, we need to transform our sensor measurements to the NED frame. Once that transformation has been performed we have inertial **relative** measurements. The implication of using this inertial relative observation vector in the filter is that relative motion is tracked. Based on this notion, the state vector is chosen as:

$$X_k^{rel} = [N_k^{rel} \dot{N}_k^{rel} E_k^{rel} \dot{E}_k^{rel} D_k^{rel} \dot{D}_k^{rel}]^T. \quad (4.3)$$

Where N , E , D and \dot{N} , \dot{E} , \dot{D} are position and velocity respectively at time instant k in the inertial NED system.

4.2 UAV and Threat Models

The first issue that needs to be addressed (referring to Figure 4.1), is the dynamic motions of the UAV and threat. The modelling of aircraft motion in state space is a very important aspect of any tracking system. The most commonly used models are the Singer acceleration model [48] and the near constant velocity model [54]. The assumption made in Chapter 1 stated that the UAV and threat are modelled as 3D point masses, travelling with constant velocity. However, in reality, while travelling with constant velocity small accelerations affect the velocity. Therefore, the near constant velocity model was chosen to model both UAV and threat dynamics. This model adds white noise with small effects to the velocity, which accounts for unpredictable modelling errors [54] (small random accelerations).

The model presented in [54] presents a 2D model which has been extended to 3D. This extension adds perturbations to a platform's Down direction. The 3D near constant velocity model is presented in what follows.

A platform's dynamics is modelled in the discrete-time Markov form:

$$X_{k+1} = \Phi X_k + \Gamma w_k. \quad (4.4)$$

X_k is the platform's state vector. w_k is the zero-mean, white, Gaussian process noise (random accelerations) with covariance Q . Φ is the state transition matrix, given as:

$$\Phi = \begin{bmatrix} 1 & \Delta T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta T & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4.5)$$

Where ΔT is the sampling period.

According to the near constant velocity model, the process noise is:

$$w_k = [w_k^N \ w_k^E \ w_k^D]. \quad (4.6)$$

In [54] it is stated that w_k^N and w_k^E are the noisy accelerations along the N and E axes respectively, while w_k^D is the noisy velocity along the D axis. However, we wish to add noisy accelerations to the D axis as well. Therefore, the resulting process noise covariance matrix is:

$$Q = \begin{bmatrix} \Delta T^4/4 & \Delta T^3/2 & \Delta T^2/2 \\ \Delta T^3/2 & \Delta T^2/2 & \Delta T \\ \Delta T^2/2 & \Delta T & 1 \end{bmatrix}. \quad (4.7)$$

And:

$$\Gamma = \begin{bmatrix} \Delta T^2/2 & 0 & 0 \\ \Delta T & 0 & 0 \\ 0 & \Delta T^2/2 & 0 \\ 0 & \Delta T & 0 \\ 0 & 0 & \Delta T^2/2 \\ 0 & 0 & \Delta T \end{bmatrix}. \quad (4.8)$$

In order to transform position or velocity from BRF to NED, the UAV's attitude is first required. Euler 3-2-1 angles (Φ – roll, Θ - pitch, Ψ – yaw; see Figure 4.2) have been chosen to represent attitude as it is an intuitive method of describing attitude.

It is worth noting that the Euler 3-2-1 system suffers from a singularity as the pitch of a platform approaches $\pm 90^\circ$. This phenomenon would cause a tracking filter to become unstable. Despite this fact, Euler angles were chosen to describe the platform's attitude as pitching to $\pm 90^\circ$ is a highly unlikely event in normal flight.

Once the UAV attitude has been described with Euler angles [81], the transformation or direction cosine matrix (DCM) can be found which relates the two frames of reference.

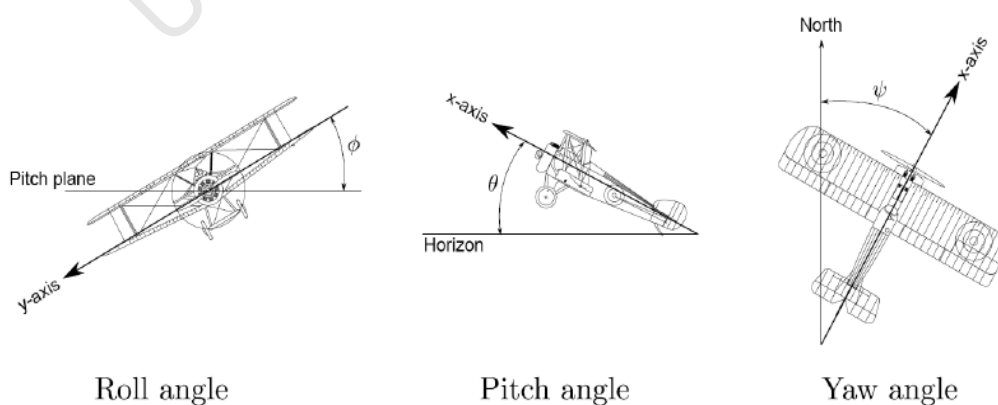


Figure 4.2: Euler angles illustration [79].

The DCM derived in [82] is:

$$DCM = \begin{bmatrix} C_\Psi C_\Theta & C_\Psi S_\Theta S_\Phi - S_\Psi C_\Phi & C_\Psi S_\Theta C_\Phi + S_\Psi S_\Phi \\ S_\Psi C_\Theta & S_\Psi S_\Theta S_\Phi + C_\Psi C_\Phi & S_\Psi S_\Theta C_\Phi - C_\Psi S_\Phi \\ -S_\Theta & C_\Theta S_\Phi & C_\Theta C_\Phi \end{bmatrix}. \quad (4.9)$$

Where $C_x = \cos x$ and $S_x = \sin x$

The DCM is used as follows:

$$X^{NED} = DCM \times X^{BRF}. \quad (4.10)$$

Where X^{BRF} is a 3×1 vector with respect to the BRF and X^{NED} is its projection in the inertial NED system.

4.3 Flight Scenario

Tracking of the threat is to be performed with respect to the NED system. Before the TTU can perform this function, the simulation environment needs to be set up. Furthermore, it needs to provide the TTU with all the required information in order to perform the tracking. In this context, the initial data required is illustrated in Figure 4.3. Therefore the following procedures need to be performed:

- Threat and UAV trajectory generation in NED.
- Determination of the relative dynamics in NED.
- UAV attitude calculation.
- Projection of the relative dynamics in BRF.
- Generation of radar and camera measurements.

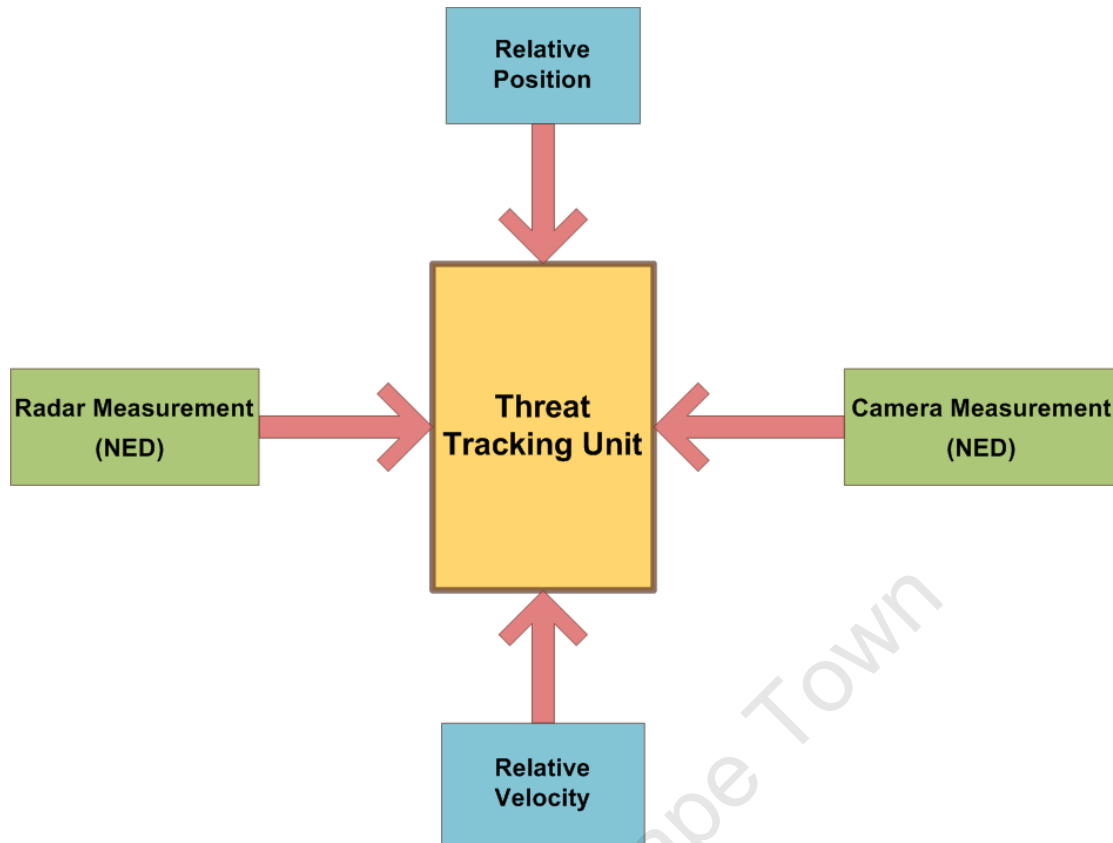


Figure 4.3: Inputs to the TTU.

For threat and UAV trajectory generation in NED we start by defining initial positions and velocities for both UAV and threat. That is defining two state vectors of Equation (4.3). The state vectors are then evolved in the state space with Equation (4.5) using the near constant velocity models. This involves generating the random vector w_k from a normal distribution with zero mean and covariance Q . For the tests that follow, the initial state vectors are presented in the table below.

Table 4.1: UAV and threat initial states.

State	Position (N, E, D)	Velocity ($\dot{N}, \dot{E}, \dot{D}$)
X_k^{UAV}	$[0\text{ m}, 0\text{ m}, 0\text{ m}]$	$[50\text{ m/s}, 0\text{ m/s}, 0\text{ m/s}]$
X_k^{threat}	$[30000\text{ m}, 0\text{ m}, 0\text{ m}]$	$[-100\text{ m/s}, 0\text{ m/s}, 0\text{ m/s}]$

The reader should note that the simulations performed in this thesis assumes VFR flight conditions as it has been proved that most NMACs (see Chapter 1) occur under these conditions. As a result the maximum speed of threats is limited to 150 m/s [79].

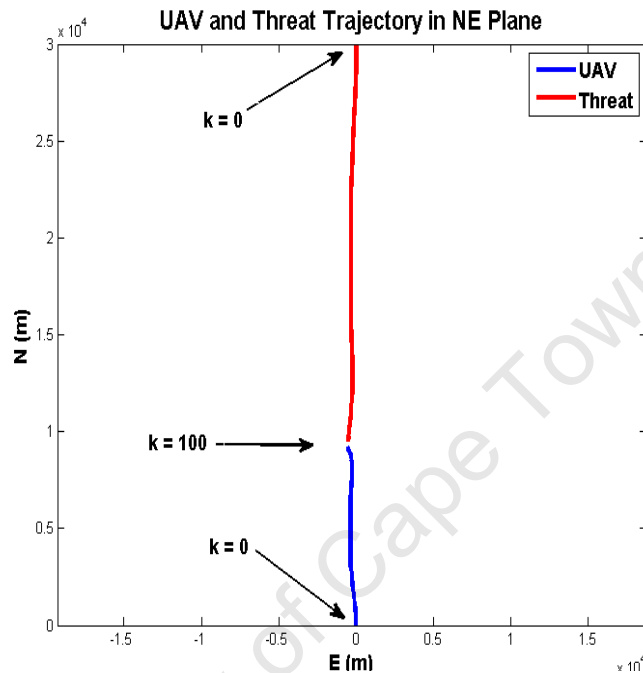


Figure 4.4: UAV and Threat flight path shown in the NE plane.

The trajectories illustrated in Figure 4.4 are the result of evolving the initial states listed in Table 4.1. The UAV starts at the origin of the NED system, at $k = 0$, flying due north; at the same time-step the threat is located 30 km north of the UAV, flying due south. There is no altitude separation between the two platforms and is considered a head-on collision scenario. The simulation is run until $k = 100$. Note that the effect of the process noise (near constant velocity model) is quite evident when inspecting both platforms' trajectories.

The next step is to find the relative position and relative velocity using the trajectories generated above and the relationships in Equations (4.1) and (4.2) respectively. Following this, the UAV attitude is determined and used to

transform the relative position and velocity to the BRF using Equation (4.10). The attitude is determined as follows:

$$\Psi = \tan^{-1} \left(\frac{\dot{E}}{\dot{N}} \right) \quad (4.11)$$

and

$$\Theta = -\tan^{-1} \left(\frac{\dot{D}}{\sqrt{\dot{N}^2 + \dot{E}^2}} \right). \quad (4.12)$$

With relative motion now modelled in the BRF, sensor measurements are formed by first sampling the true relative trajectory (according to sensor data rate); followed by a conversion to spherical space as that is system in which the sensor operates. Noise is then added akin to the method of determining the process noise, but with covariance R . The sensor measurements are now relative motion in spherical coordinates with respect to BRF; exactly how a sensor reports measurements. However, we require the sensor measurements in relative spherical with respect to the NED system. In order to transform the radar position measurements $(\rho_k^R, \theta_k^R, \varphi_k^R)$ to NED we could use the *DCM*, however this would first require a conversion to Cartesian coordinates which would introduce more errors into the measurements. The radar range rate $(\dot{\rho}_k^R)$ does not require a transformation as it is the same in NED as it is in its measure BRF.

In order to transform the camera measurements $(\theta_k^C, \varphi_k^C)$ to NED we cannot use the method described above. The conversion to Cartesian coordinates requires the range to the threat; of which the camera has no knowledge. Therefore another method is proposed that addresses the camera measurement deficiency and the potential injection of errors into the radar measurement. It allows the direct transformation from BRF to NED. First the direction cosines are found using sensor measurements:

$$\Lambda_1 = \cos \theta_k^{BRF} \cos \varphi_k^{BRF} \quad (4.13)$$

$$\Lambda_2 = \sin \theta_k^{BRF} \cos \varphi_k^{BRF} \quad (4.14)$$

$$\Lambda_3 = -\sin \varphi_k^{BRF} \quad (4.15)$$

Then the direction cosines along the N, E and D axes are determined with:

$$\begin{bmatrix} \Lambda_N \\ \Lambda_E \\ \Lambda_D \end{bmatrix} = DCM \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \\ \Lambda_3 \end{bmatrix}. \quad (4.16)$$

Now the stabilised azimuth and elevation angles can be determined by the method described by Zollo and Ristic [53] in Chapter 2 using Equations (2.48) to (2.50). Similarly, for the radar observation vector, the relative range with respect to the NED system is determined using Equations (2.51) and (2.52). In summation, the entire process of initialising required parameters is illustrated in Figure 4.5.

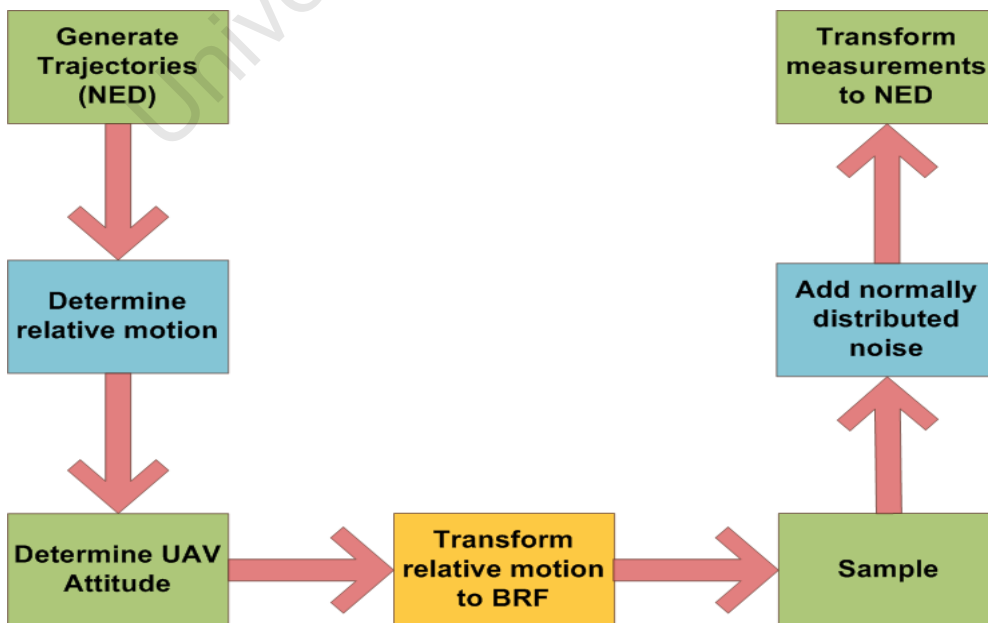


Figure 4.5: Initial simulation setup procedure.

A cross check with Figure 4.3 confirms that all components in generating the flight scenario have been addressed. Now that the TTU has all its required inputs, it needs to produce required outputs, viz. threat position and velocity estimates in the inertial NED system. In order to achieve this, multiple fusion architectures with their corresponding tracking filters are designed in Chapter 5 from which the best is selected for the final architecture.

University of Cape Town

Chapter 5: TTU Design

5.1 Introduction

This chapter describes the design of the TTU using MSDF. Thesis *objective 5* is to “Design and implement an airborne target tracking algorithm using multisensor data fusion”. However, before developing the fusion architectures using multiple redundant sensors, single sensor tracking filters are developed to illustrate their accuracies and set a baseline with which the filters resulting from MSDF can be evaluated against. Furthermore, the single sensor filters that are developed form the building blocks from which the fusion architectures are made up. Thus, this chapter begins by first developing EO camera and radar only tracking filters. Two cases for the radar only tracking filter are discussed: ranging radar and Doppler radar. In order to illustrate the filter complexity versus performance gain trade-off, nonlinear as well as linear ranging radar tracking filters are developed. As stated in Chapter 2, the complication in tracking with Doppler information is that measurement errors in range and rate are correlated [65]. However, a common approach is to disregard this fact. Therefore, two Doppler radar filters are developed: one which assumes measurement errors in range and range rate are not correlated; and the other assumes this correlation exists and is accounted for in the filter.

Following this, multiple fusion architectures and their respective tracking filters are designed. In order to evaluate the performance of the fusion architectures, Monte Carlo simulations are performed. The reason for multiple fusion architectures being developed is so that quantitative evaluation can occur and the best fusion architecture for airborne tracking using MSDF can be exposed; which will form the *Threat Tracking Unit* block in the system overview (Figure 1.10). This chapter ensures the accomplishment of thesis *objective 5*. Furthermore, the results from this chapter will see either the proving or disproving of the hypotheses set out in Chapter 3.

After a filter has been developed, it is integrated into the simulation environment described in Chapter 4 and then tested as set out in the Methodology. For each filter, first the results from a single filter run of k time-steps are presented. Thereafter Monte Carlo simulations are performed. The results of 500 independent Monte Carlo runs (N^{MC}) over k steps are then presented, illustrating the performance metrics detailed in section 3.5.

Table 5.1 lists all the tracking filters that are developed in this chapter and the estimators which they consist of.

Table 5.1: Designed tracking filters and the estimators on which they are built.

Design	Radar Local Filter	Camera Local Filter	Fusion Centre Filter
EO camera filter		EKF	
Nonlinear range radar filter	EKF		
Linear range radar filter	KF		
Uncorrelated Doppler radar filter	EKF		
Correlated Doppler radar filter	EKF		
FA1 (centralised)			EKF
FA2 (distributed)	EKF	EKF	
FA3 (distributed)	KF	EKF	

5.2 Common Filter Parameters

This section presents information that is required in the design of all the filters discussed in this chapter. In what follows, the superscripts R , C and F denote that the component belongs to either the radar, EO camera or fused subsystem, respectively.

For each filter the relative motion between UAV and threat is modelled in the discrete time Markov form:

$$X_{k+1}^R = \Phi X_k^R + \Gamma w_k, \quad (5.1)$$

and

$$X_{k+1}^C = \Phi X_k^C + \Gamma w_k. \quad (5.2)$$

Where the state vectors are:

$$X_k^R = X_k^C = X_k^F = [N_k^{rel} \dot{N}_k^{rel} E_k^{rel} \dot{E}_k^{rel} D_k^{rel} \dot{D}_k^{rel}]^T. \quad (5.3)$$

The state vectors contain the components of relative motion with respect to the inertial NED frame of reference.

Then the threat's position and velocity in NED are determined by the sum of the estimated state vector in Equation (5.3) and the UAV's current position and velocity – possibly provided by an IMU.

The state transition and noise matrices are given as:

$$\Phi = \begin{bmatrix} 1 & \Delta T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta T & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\Gamma = \begin{bmatrix} \Delta T^2/2 & 0 & 0 \\ \Delta T & 0 & 0 \\ 0 & \Delta T^2/2 & 0 \\ 0 & \Delta T & 0 \\ 0 & 0 & \Delta T^2/2 \\ 0 & 0 & \Delta T \end{bmatrix}$$

With $\Delta T^R = 2$ and $\Delta T^C = 1/2$. That is, the radar provides a new measurement every 2 s and the EO camera every 500 ms. Therefore, the camera provides measurements to the TTU four times faster than the radar does. This translates to radar and camera measurement frequencies of 0.5 Hz and 2 Hz, respectively. Note that cameras generally provide measurements at higher frequencies than this, however this value was chosen to account for the image processing time required to successfully detect a target.

The acceleration disturbances for both UAV and threat in the north, east and down directions are set to:

$$w_k = [0.05 \text{ m/s}^2 \ 0.05 \text{ m/s}^2 \ 0.05 \text{ m/s}^2]. \quad (5.4)$$

As all of the filters are performing airborne tracking, state estimates of relative motion are produced. Thus, the final step, in order to extract the states of interest (of the threat), is to add the UAV's position and velocity in NED to the state vector estimate. This then yields the estimate of only the threat position and velocity in the inertial NED system.

The relative range for the flight scenario under review is illustrated in Figure 5.1.

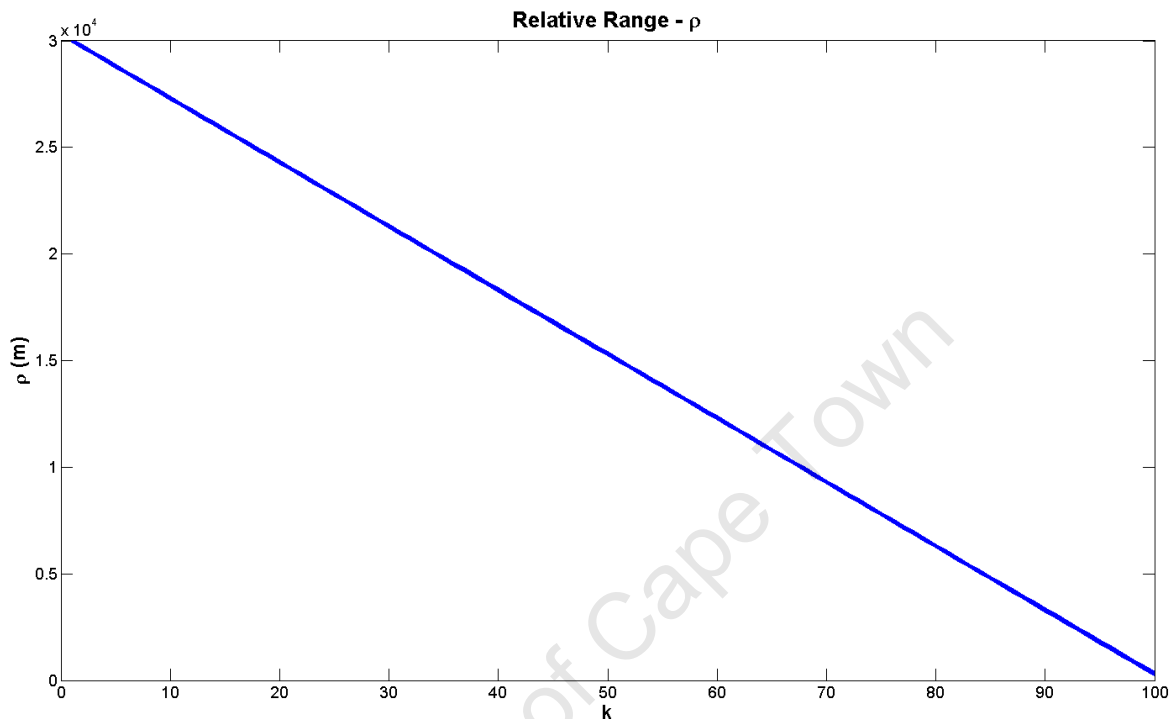


Figure 5.1: Relative range of flight scenario.

5.3 EO Camera Tracking Filter

In our progressive design structure, leading up to the development of tracking filters using multiple redundant sensors, the first tracking filter developed is that of the EO camera. As the relative motion dynamics has been described in Cartesian coordinates, tracking is performed with mixed coordinates; the filter uses the camera's measurements directly (raw). Therefore a nonlinear EKF is selected as the tracking filter.

With the EO camera measuring only target relative azimuth and elevation, its observation vector is:

$$z_k^C = [\theta_k^C \quad \varphi_k^C]^T + v_k^C. \quad (5.5)$$

Where v_k^C is the zero-mean, white, Gaussian camera measurement noise with covariance matrix R^C :

$$R^C = \begin{bmatrix} \sigma_\theta^{R^2} & 0 \\ 0 & \sigma_\varphi^{R^2} \end{bmatrix}. \quad (5.6)$$

Where σ_θ and σ_φ are the measurement errors in azimuth and elevation respectively; whose values are shown below.

Table 5.2: EO camera measurement errors.

Sensor	Measurement Error	Value
EO Camera	σ_θ^C	0.02 <i>radians</i>
	σ_φ^C	0.02 <i>radians</i>

The measurement errors tabulated in Table 5.2 imply camera measurement errors in azimuth and elevation of 1.15° . These small angular measurement errors are commonly found in EO cameras [48]. It will be evident in the following sections that the camera angular errors are a great deal smaller than those of the radar's. These accurate measurements can be confirmed by plotting the true relative azimuth and elevation against the camera measured values (see Figure 5.2 and Figure 5.3).

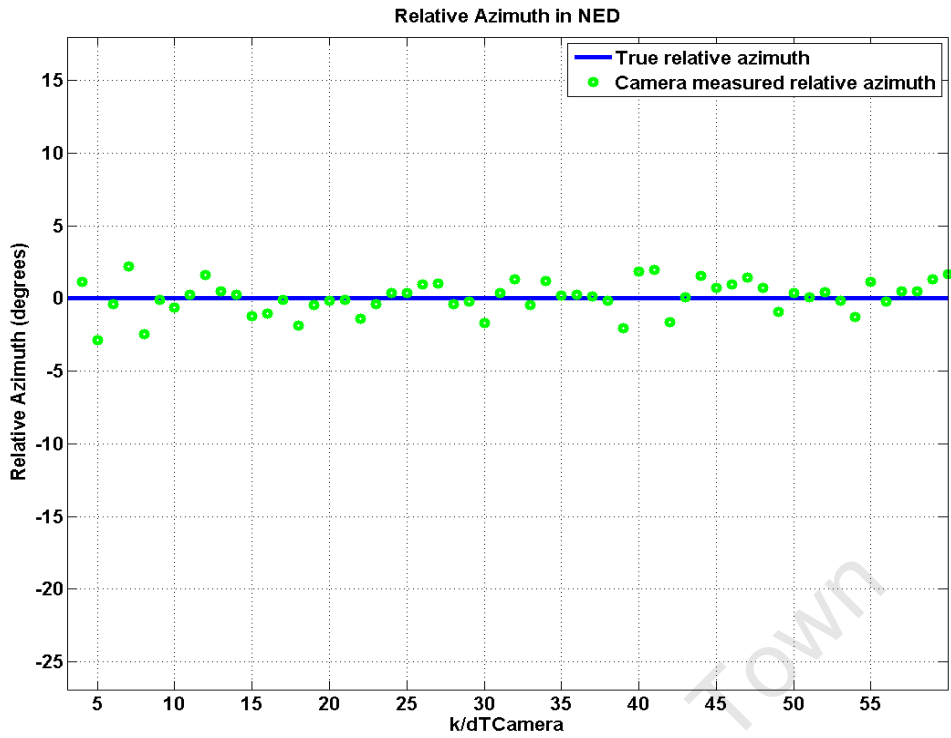


Figure 5.2: Relative azimuth in NED - truth and camera measured values.

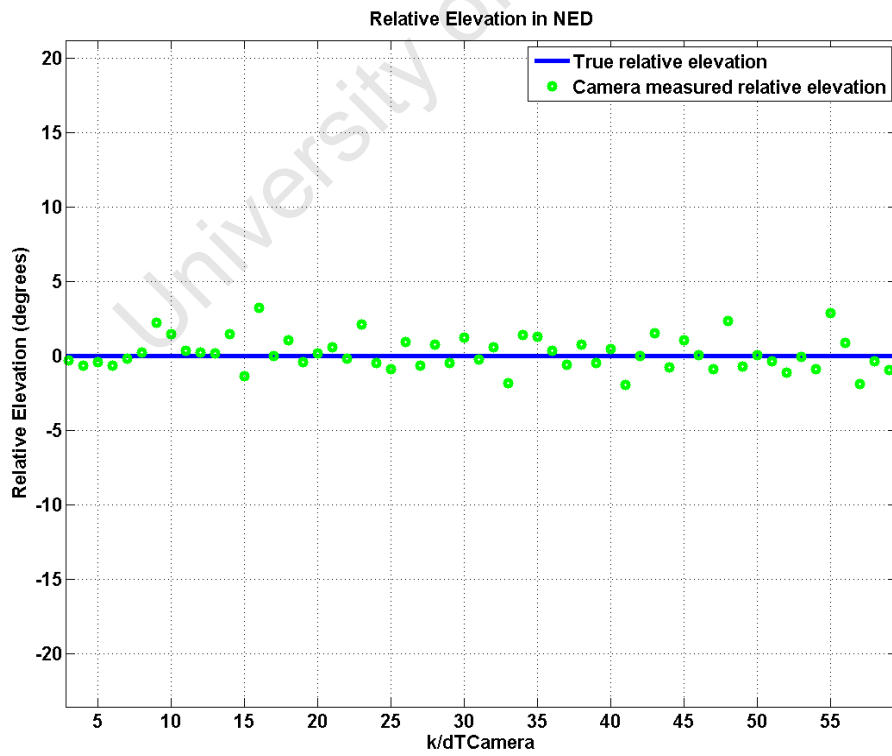


Figure 5.3: Relative elevation in NED - truth and camera measured values.

The first solution is tracking in mixed coordinates. This implies the camera's measurements must be modelled in the sensor frame as the platform's dynamics have been modelled in the Cartesian space.



Figure 5.4: Bearings-only tracking with an EKF.

Standard EKF equations are used in order to track the target. This involves first *prediction* as that in Equations (2.26) and (2.27), and *correction* performed by Equations (2.23) to (2.25).

In order to perform the update phase, the linearized measurement matrix is determined using a first-order Taylor series approximation around the current state estimate as follows:

$$h^c(X) = z_k^c = \begin{bmatrix} \tan^{-1}\left(\frac{E}{N}\right) \\ -\sin^{-1}\left(\frac{D}{\sqrt{N^2 + E^2 + D^2}}\right) \end{bmatrix} \quad (5.7)$$

$$H_X^c(k) = \left. \frac{\partial h^c(X)}{\partial X} \right|_{X=\hat{X}_{k|k-1}^c} \quad (5.8)$$

See Appendix A1 for the complete linearized measurement matrix.

Now the standard EKF equations are used to track the threat. Using the linearized measurement matrix the innovation covariance is calculated and hence the Kalman gain using Equation (2.23). Then the innovation and hence the state estimate is determined by Equation (2.24). Following this the state estimate covariance matrix is produced from Equation (2.25). The results from this filter are illustrated in the following section.

The reader should note that the EO camera was assumed to be able to detect the target at all times during the simulation; that is, at all target ranges. The reason for exaggerating the camera's detection range is twofold: first, so that the radar and camera filters can be easily compared against one another; and second, to make the positive effects of fusion more apparent. The camera's detection range is considerably smaller than the radar's, and, had it been taken into account, the full benefits of fusion would only be visible in the last few iterations of the simulation.

5.3.1 Results

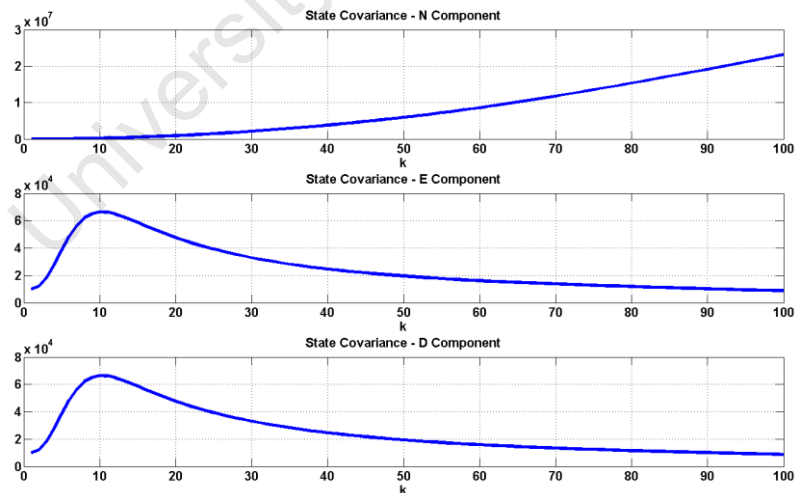


Figure 5.5: Camera filter - estimated state error covariance of N, E and D components.

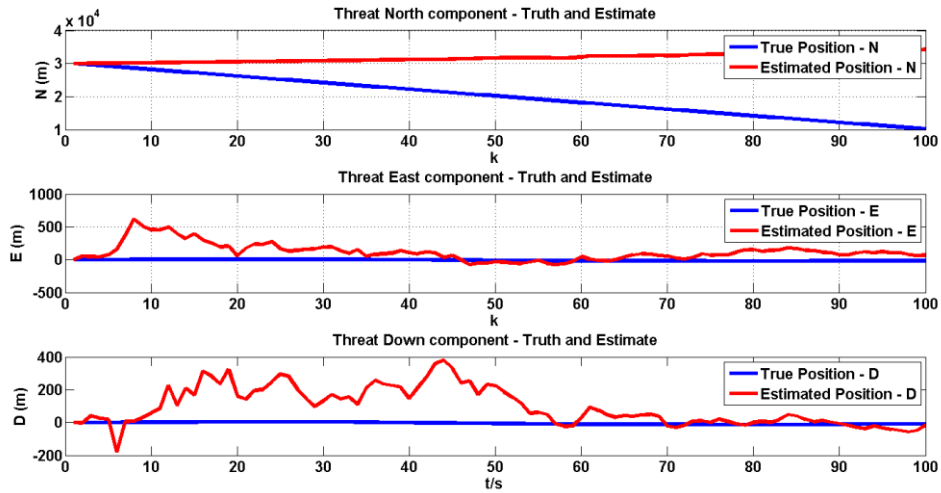


Figure 5.6: Camera filter - True and estimated threat position in N, E and D directions.

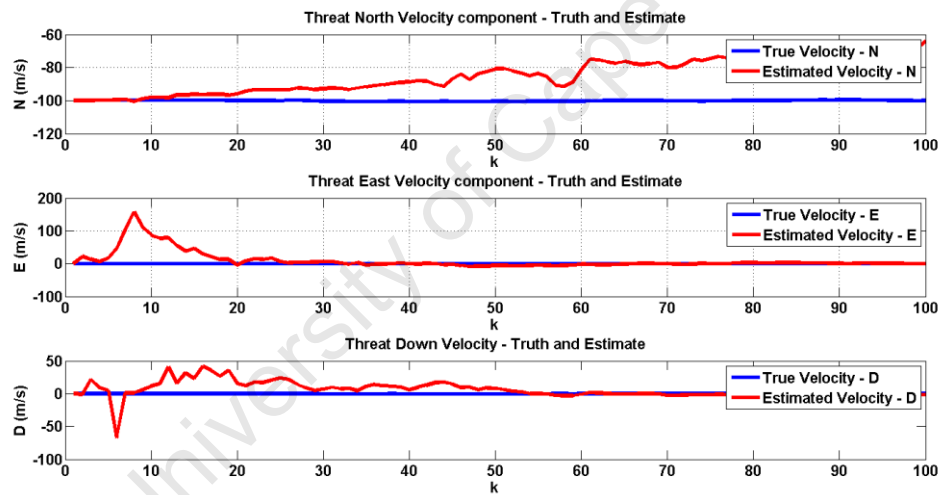


Figure 5.7: Camera filter - True and estimated threat velocity in N, E and D directions.

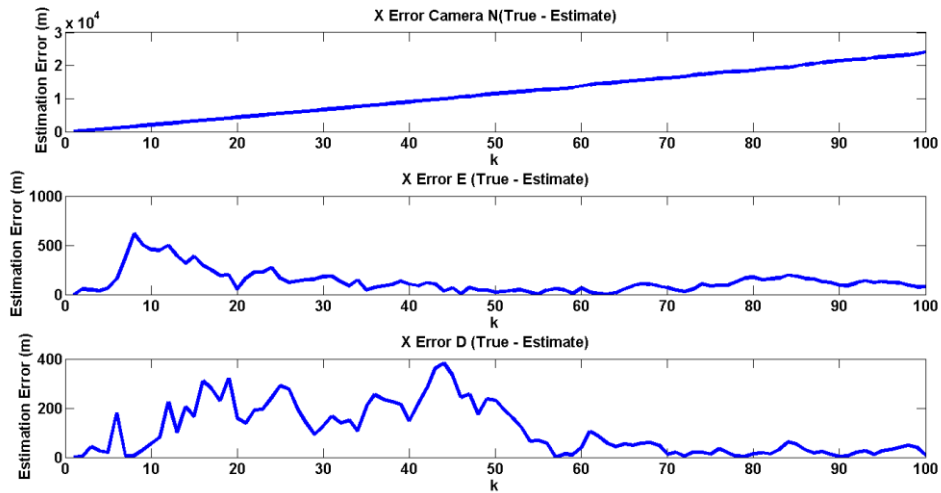


Figure 5.8: Camera filter - Estimation error in N, E and D directions.

5.3.2 Monte Carlo Simulations

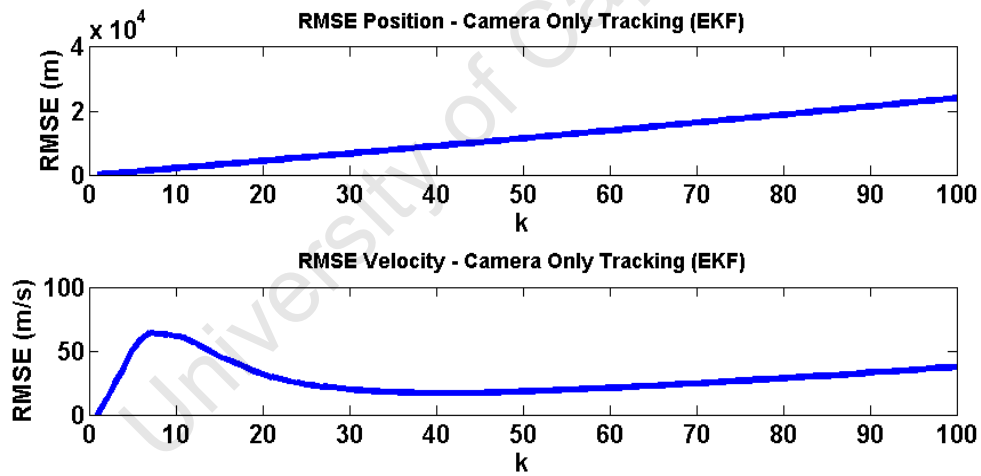


Figure 5.9: Camera only filter RMSE in position and velocity.

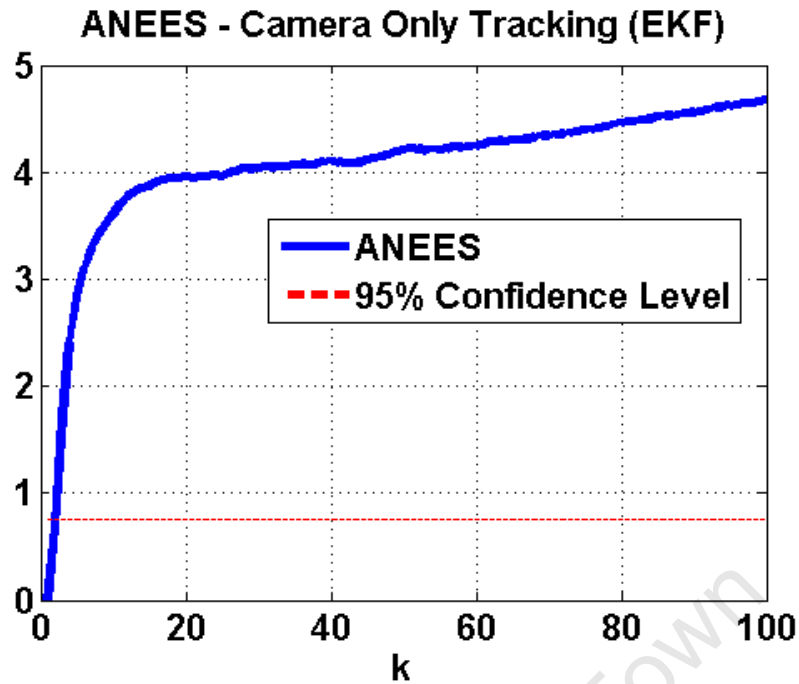


Figure 5.10: Camera only filter ANEES.

5.3.3 Summary

In summary, it is evident that the EO camera filter diverges. The results show that the threat position and velocity in the east and down directions are adequately tracked, but neither in the north direction. Therefore the EO camera tracking filter cannot suitably track the threat using angle-only information.

5.4 Ranging Radar Tracking Filter

For a ranging radar-only tracker we develop two filters. The first filter uses radar measurements exactly as they arrive from the radar, therefore necessitating the use of a nonlinear EKF. The second design utilises a linear filter to produce the state estimate; radar measurements are converted to Cartesian space and then debiased. Furthermore, the radar measurement noise matrix is projected from the sensor frame (BRF) to the Cartesian system.

The radar measures relative range, azimuth and elevation. Therefore the corresponding observation vector is:

$$z_k^R = [\rho_k^R \quad \theta_k^R \quad \varphi_k^R]^T + v_k^R. \quad (5.9)$$

Where v_k^R is the zero-mean, white, Gaussian radar measurement noise with covariance matrix R^R :

$$R^R = \begin{bmatrix} \sigma_\rho^{R^2} & 0 & 0 \\ 0 & \sigma_\theta^{R^2} & 0 \\ 0 & 0 & \sigma_\varphi^{R^2} \end{bmatrix}. \quad (5.10)$$

Where σ_ρ , σ_θ and σ_φ are the measurement errors in range, azimuth and elevation respectively; whose values are shown below.

Table 5.3: Radar measurement errors.

Sensor	Measurement Error	Value
Doppler Radar	σ_ρ^R	200 <i>m</i>
	σ_θ^R	0.3 <i>radians</i>
	σ_φ^R	0.3 <i>radians</i>

The measurement errors above are exaggerated and, in practical radars, would be significantly smaller. However, these values were chosen to illustrate the effects of MSDF when one sensor's measurements are very accurate and the other's not. The accurate sensor's measurements are that of the EO camera's listed in Table 5.2.

The effects of the radar's measurement errors are displayed in the following figures which contain the relative true and radar measured range, azimuth and elevation in NED respectively.

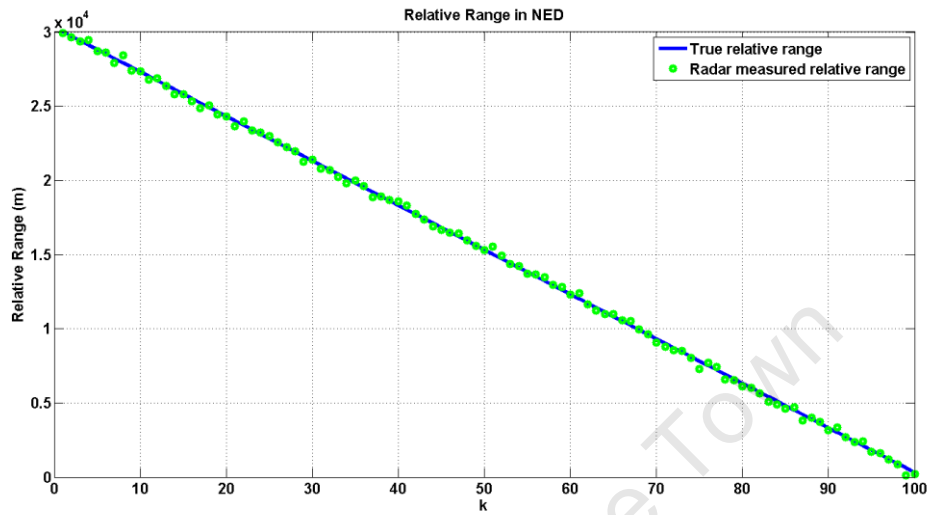


Figure 5.11: Relative range in NED - truth and radar measured values.

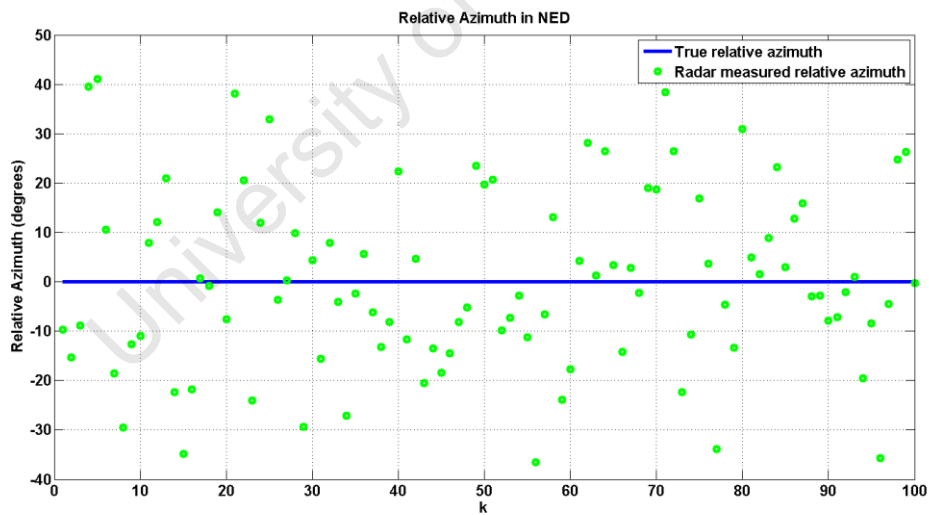


Figure 5.12: Relative azimuth in NED - truth and radar measured values.

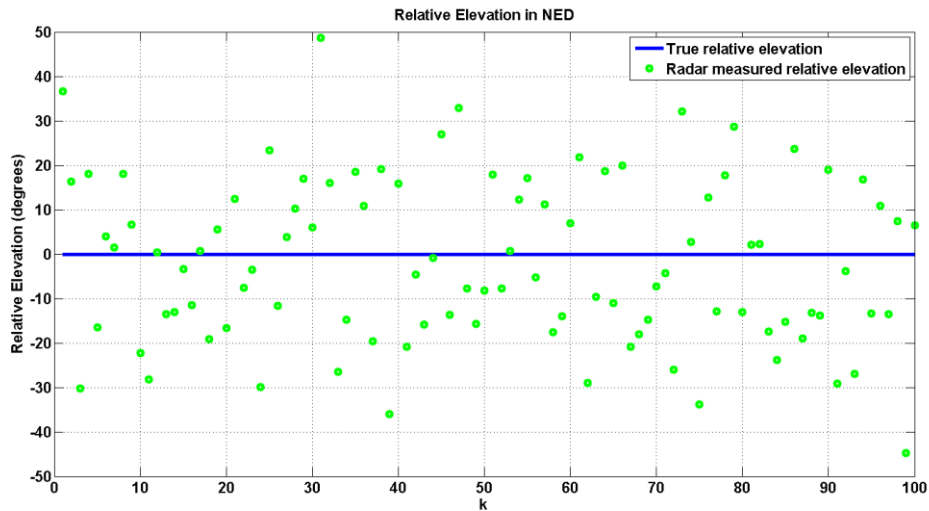


Figure 5.13: Relative elevation in NED - truth and radar measured values.

5.4.1 Nonlinear Approach

The first solution, much like our bearings-only tracker, involves tracking in mixed coordinates. As the radar's measurements are nonlinear combinations of the state variables, a nonlinear filter is required in order to perform tracking. We have selected the EKF in order to achieve this.



Figure 5.14: Nonlinear ranging radar tracking filter.

The linearized radar measurement matrix is found by calculating the Jacobian of the observation vector with respect to the state variables:

$$h^R(X) = z_k^R = \begin{bmatrix} \sqrt{N^2 + E^2 + D^2} \\ \tan^{-1}\left(\frac{E}{N}\right) \\ -\sin^{-1}\left(\frac{D}{\sqrt{N^2 + E^2 + D^2}}\right) \end{bmatrix}, \quad (5.11)$$

$$H_X^R(k) = \left. \frac{\partial h^R(X)}{\partial X} \right|_{X=\hat{X}_{k|k-1}^R}. \quad (5.12)$$

See Appendix A2 for the complete linearized measurement matrix.

Now we are able to produce state estimates using the standard EKF equations. The results of this filter are shown below.

5.4.1.1 Results

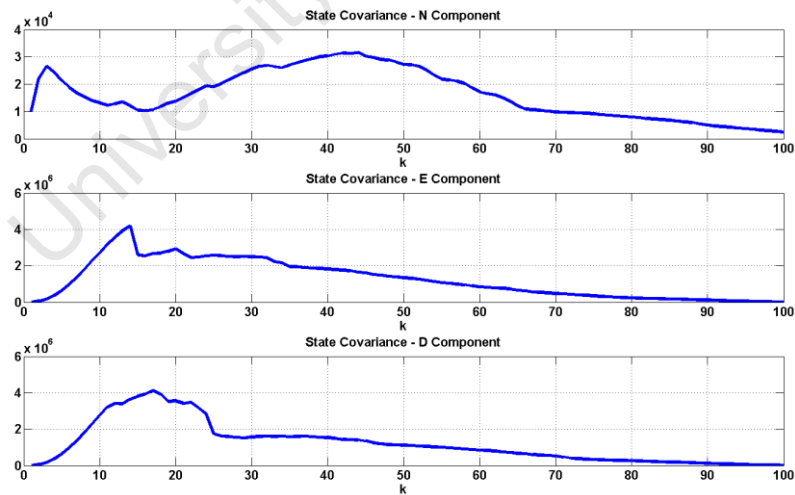


Figure 5.15: Radar nonlinear filter (EKF) – Estimated state error covariance of N, E and D position components.

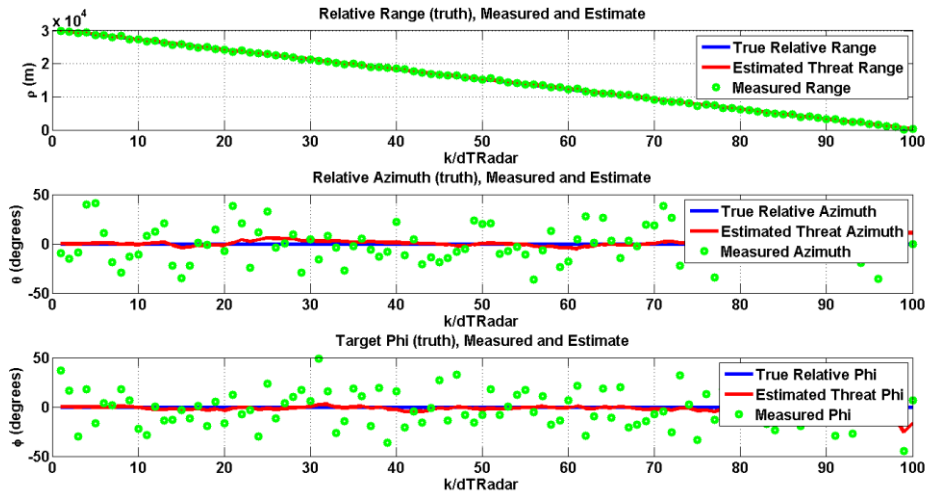


Figure 5.16: Radar nonlinear filter (EKF) - True, estimated and measured relative motion range, azimuth and elevation in the BRF.

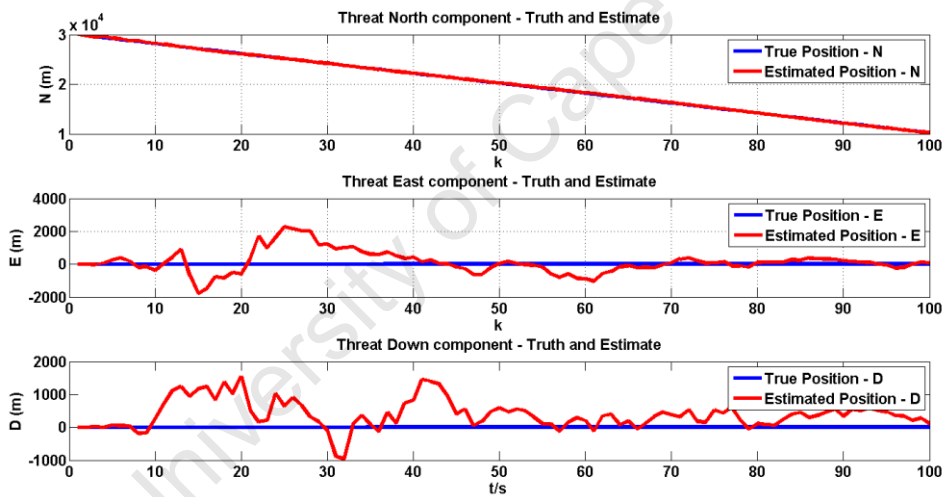


Figure 5.17: Radar nonlinear filter (EKF) – True and estimated threat position in N, E and D directions.

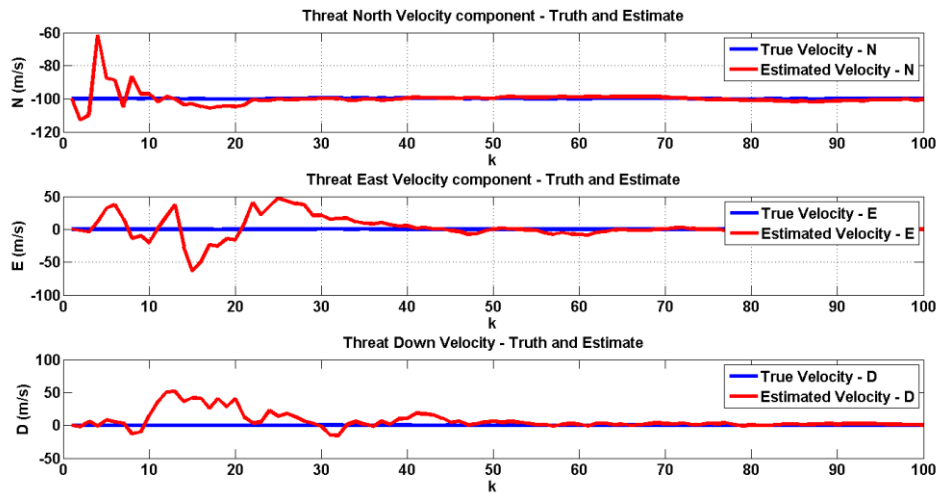


Figure 5.18: Radar nonlinear filter (EKF) – True and estimated threat velocity in N, E and D directions.

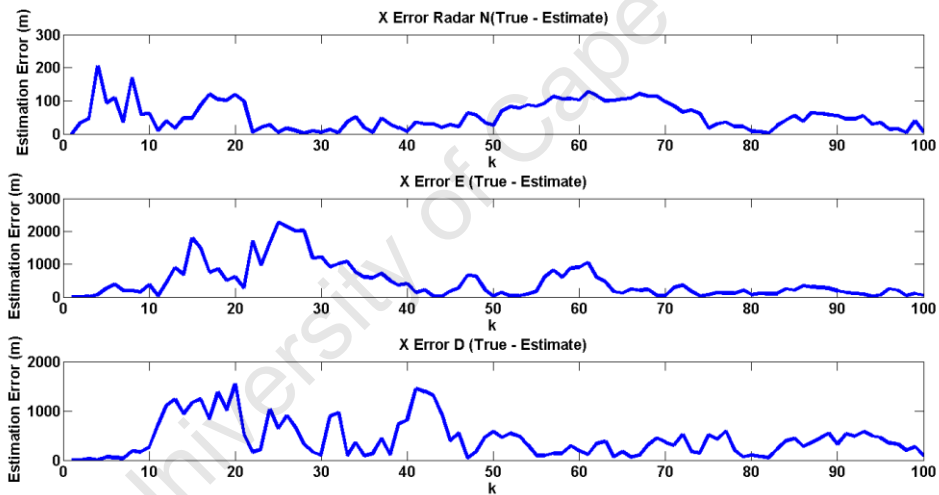


Figure 5.19: Radar nonlinear filter (EKF) – Position estimation error in N, E and D directions in a single flight scenario.

5.4.1.2 Monte Carlo Simulations

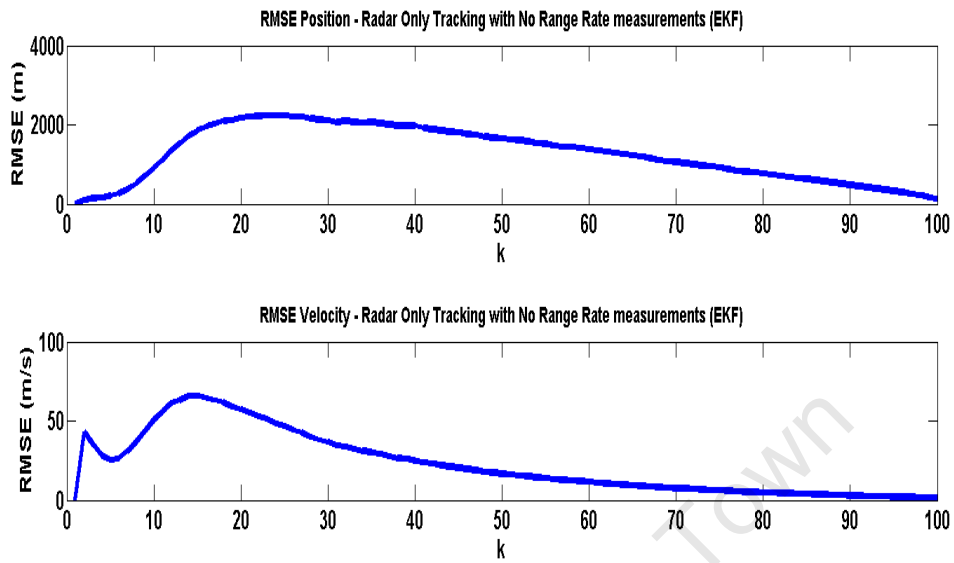


Figure 5.20: Radar nonlinear filter with no range rate measurements - RMSE in position and velocity.

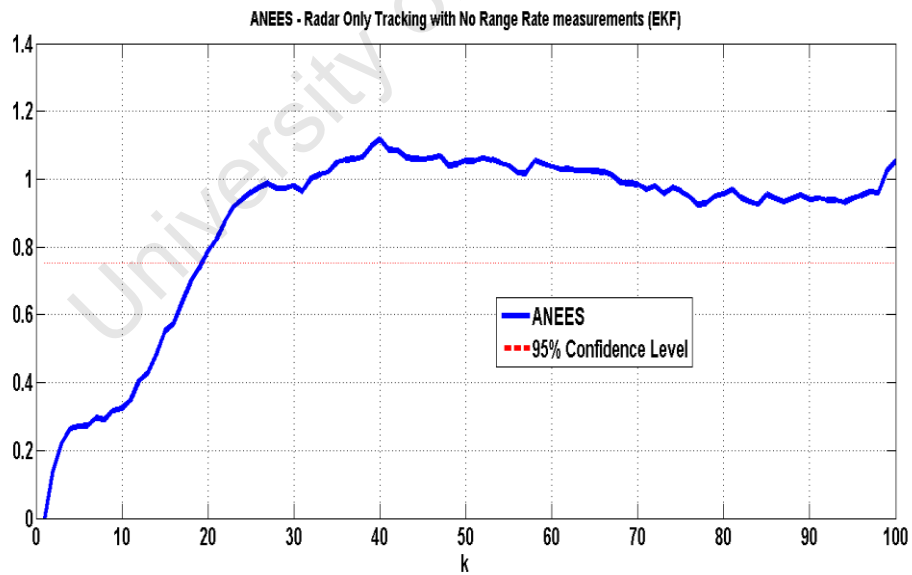


Figure 5.21: Radar filter with no range rate measurements - ANEES.

5.4.2 Linear Approach

The second filter developed, in contrast to all the filters presented thus far, is one that performs tracking in Cartesian coordinates only. First measurements are converted to Cartesian coordinates and then debiased. Concurrently the radar measurement matrix is determined in Cartesian space. These inputs feed a linear KF which utilise Equations (2.14) to (2.18) to produce estimates of relative position and velocity. The tracking filter is illustrated below.

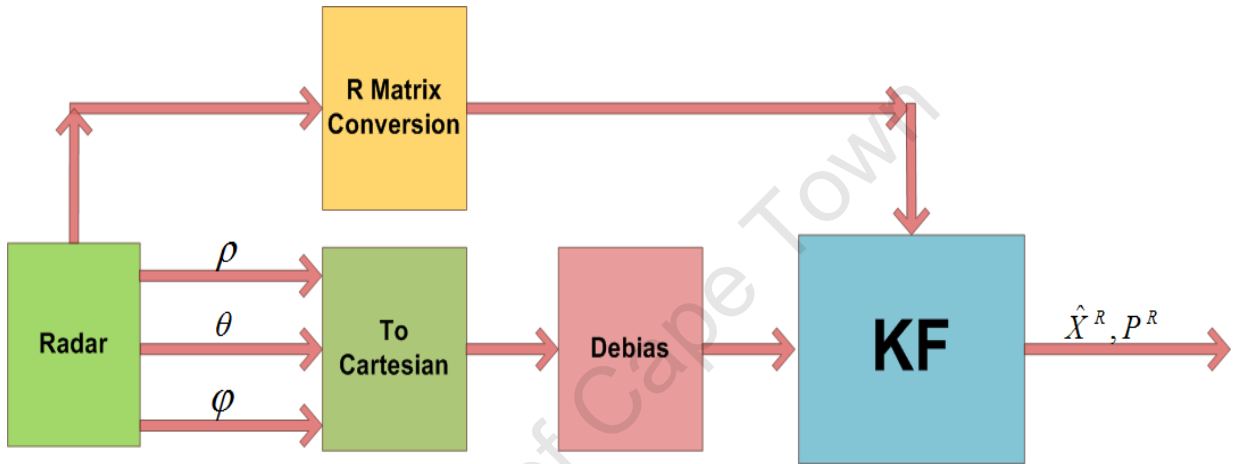


Figure 5.22: Second range radar filter developed. A linear solution which involves measurement conversion to Cartesian coordinates.

Radar relative measurements originate in spherical coordinates with respect to the BRF. However, in section 4.3 they were transformed to the NED system. Therefore the first step is to convert Equation (5.9) to Cartesian coordinates using the following equations:

$$N_k' = \rho_k \cos \theta_k \cos \varphi_k \quad (5.13)$$

$$E_k' = \rho_k \sin \theta_k \cos \varphi_k \quad (5.14)$$

$$D_k' = -\rho_k \sin \varphi_k \quad (5.15)$$

It was shown in Chapter 2 that converting measurements to Cartesian coordinates and using this in a tracking filter, leads to biased and inconsistent state estimates. Therefore, before tracking, the measurements need to be debiased. If measurements of range, azimuth and elevation are zero-mean, Gaussian and no correlations exist in measurement errors, then debiasing of the converted measurements is performed as follows [53]:

$$\begin{bmatrix} N_k \\ E_k \\ D_k \end{bmatrix} = \begin{bmatrix} N'_k \left[1 - \left(e^{-\sigma_\theta^2} e^{-\sigma_\varphi^2} - e^{-\frac{\sigma_\theta^2}{2}} e^{-\frac{\sigma_\varphi^2}{2}} \right) \right] \\ E'_k \left[1 - \left(e^{-\sigma_\theta^2} e^{-\sigma_\varphi^2} - e^{-\frac{\sigma_\theta^2}{2}} e^{-\frac{\sigma_\varphi^2}{2}} \right) \right] \\ D'_k \left[1 - \left(e^{-\sigma_\varphi^2} - e^{-\frac{\sigma_\theta^2}{2}} \right) \right] \end{bmatrix}. \quad (5.16)$$

Now that we have the converted observation from Equation (5.16) we need to transform R^R (the measurement covariance matrix in the radar local spherical frame) to the Cartesian frame as well.

By finding the Jacobian of Equations (5.13) to (5.15) with respect to the state vector in Equation (5.3), we arrive at:

$$A = \begin{bmatrix} \cos \theta_k^R \cos \varphi_k^R & -\rho_k^R \cos \theta_k^R \sin \varphi_k^R & -\rho_k^R \sin \theta_k^R \cos \varphi_k^R \\ \sin \theta_k^R \cos \varphi_k^R & -\rho_k^R \sin \theta_k^R \sin \varphi_k^R & \rho_k^R \cos \theta_k^R \cos \varphi_k^R \\ \sin \varphi_k^R & \cos \varphi_k^R & 0 \end{bmatrix}. \quad (5.17)$$

Since the transformation A from radar coordinates to Cartesian is a linear one, we can make the following statement about R^R in Cartesian space:

$$R^{RConv} = AR^R A^T. \quad (5.18)$$

Using the above information the standard KF equations are used to produce state estimates and its associated estimated state covariance. The results from this approach are illustrated below.

5.4.2.1 Results

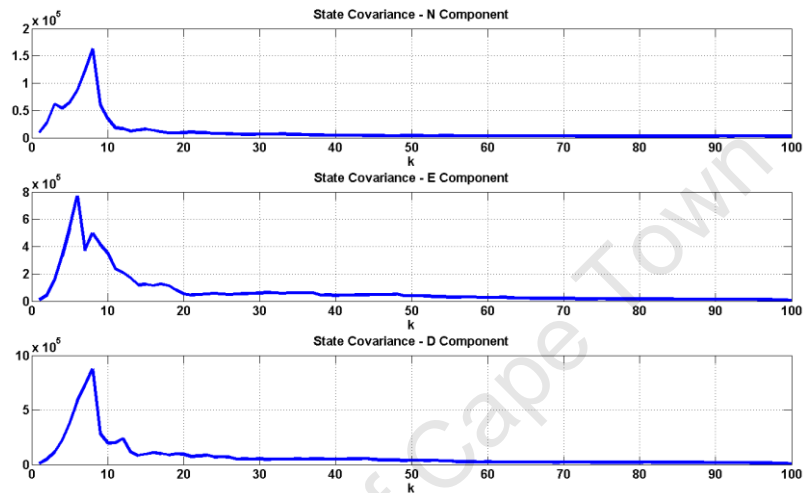


Figure 5.23: Radar linear KF tracking in Cartesian coordinates – estimated state error covariance.

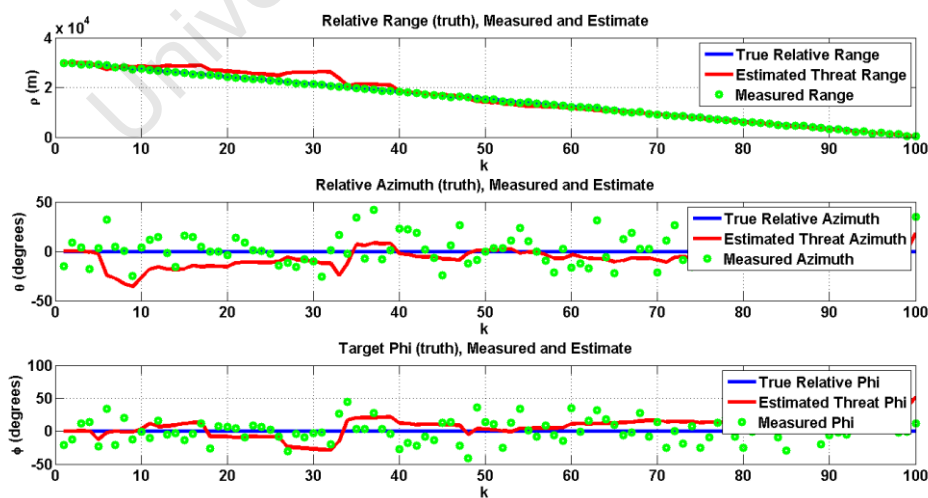


Figure 5.24: Radar linear KF tracking in Cartesian coordinates – True, estimated and measured relative range, azimuth and elevation in BRF.

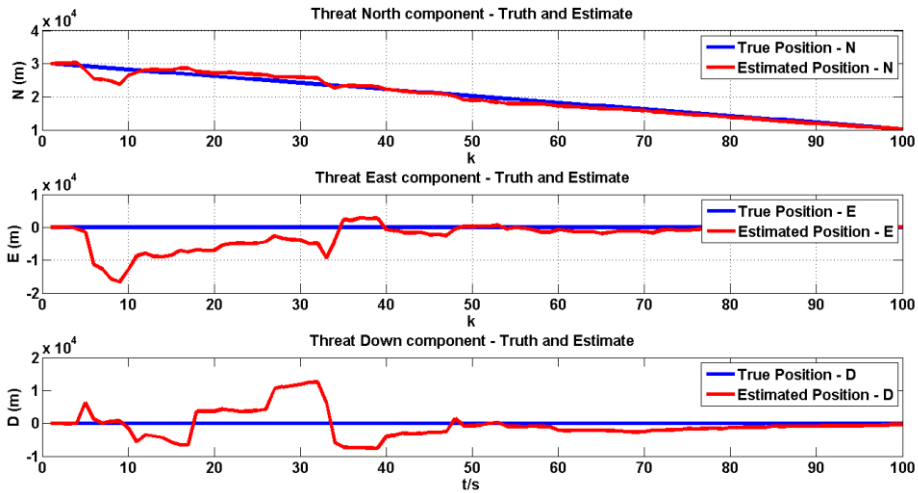


Figure 5.25: Radar linear KF tracking in Cartesian coordinates – True and estimated position in N, E and D directions.

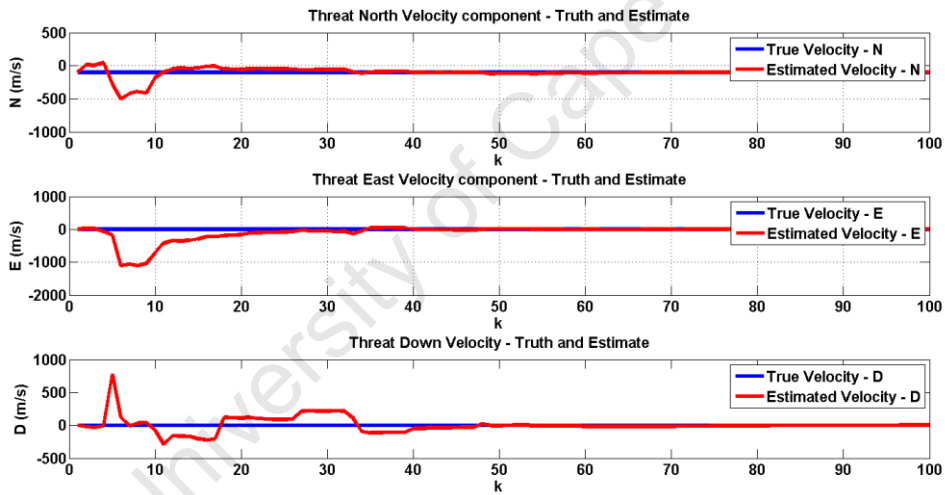


Figure 5.26: Radar linear KF tracking in Cartesian coordinates – True and estimated velocity in N, E and D directions.

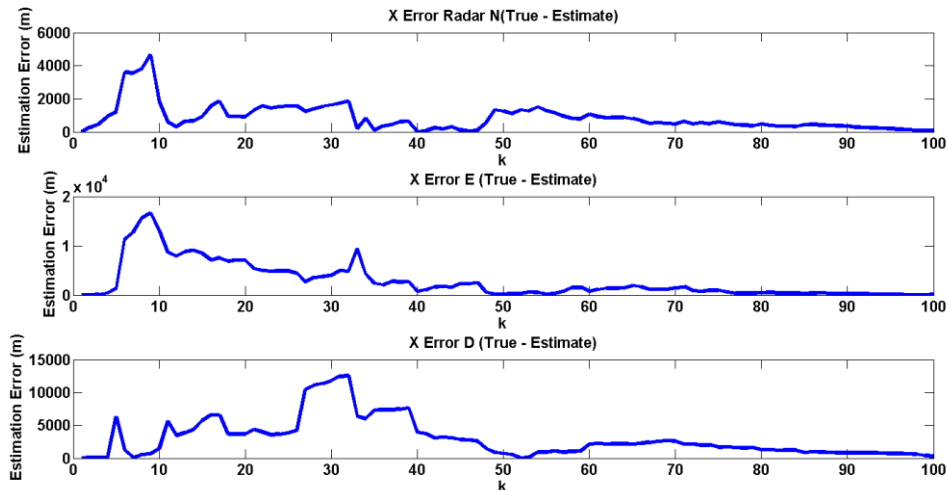


Figure 5.27: Radar linear KF tracking in Cartesian coordinates – Estimation error in N, E and D directions in a single flight scenario.

To illustrate the importance of the measurement covariance conversion to Cartesian space of the KF, performed in Equations (5.19) and (5.20), the NED threat position components are compared to the estimated values when no conversion is performed. The result is illustrated below.

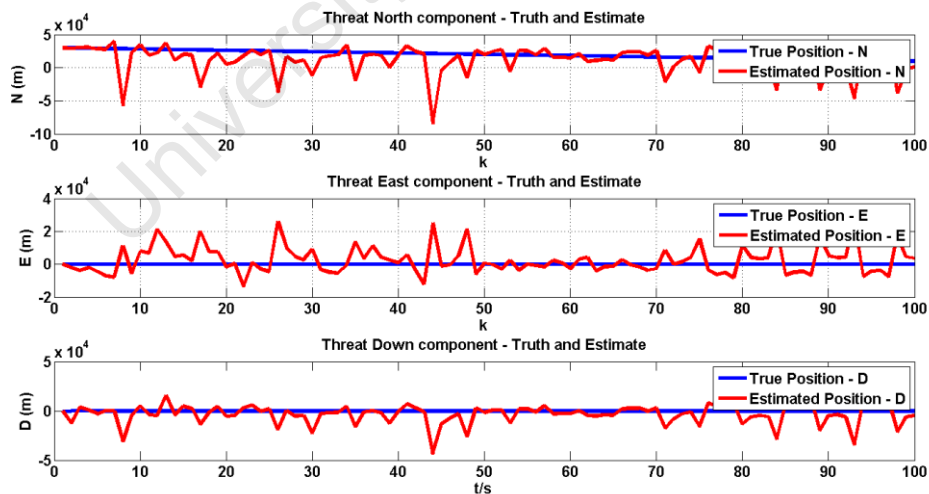


Figure 5.28: Radar linear filter (KF) tracking with Cartesian coordinates - True and estimated threat position with no measurement covariance matrix transformation.

5.4.2.2 Monte Carlo Simulations

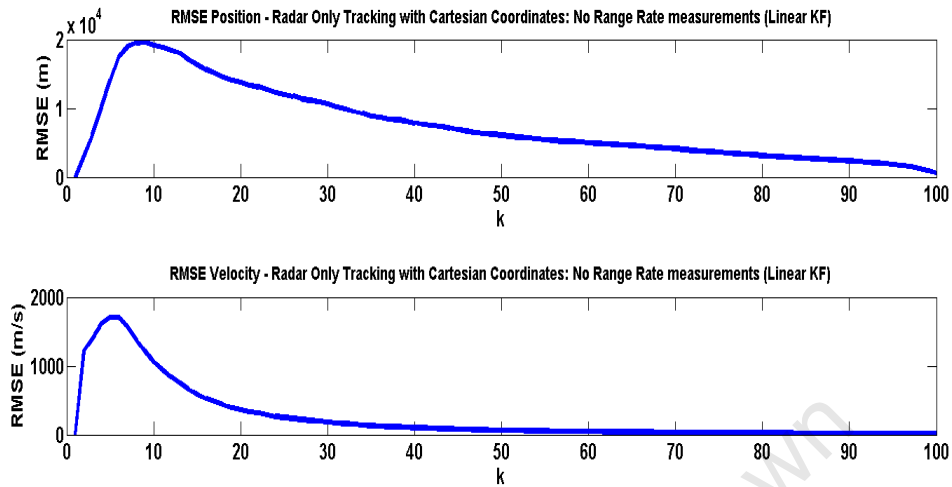


Figure 5.29: Radar linear filter tracking in Cartesian coordinates with no range rate measurements - RMSE in position and velocity.

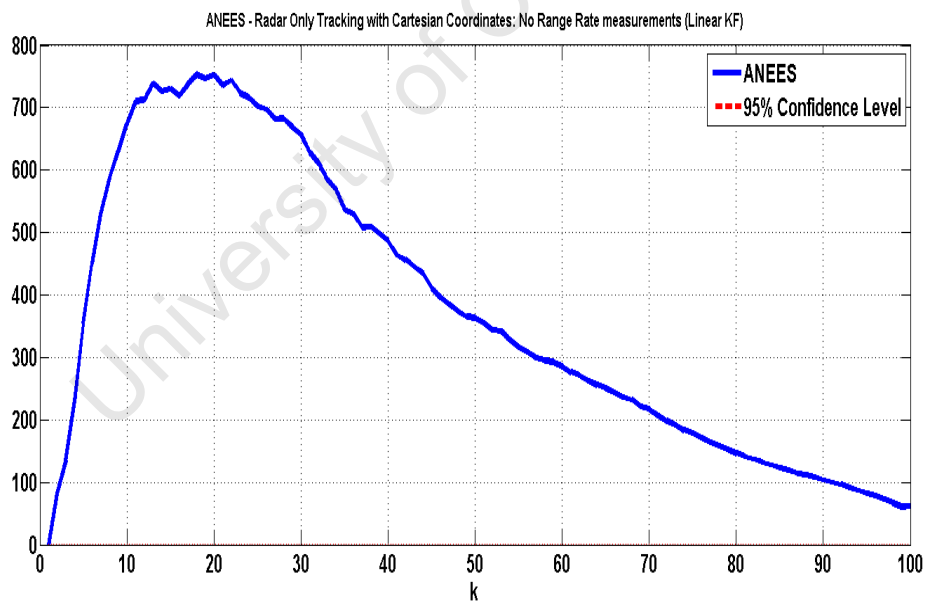


Figure 5.30: Radar linear filter tracking in Cartesian coordinates with no range rate measurements - ANEES.

5.4.3 Summary

In summary, both the process noise and measurement noises are set to high values in which the filters are required to track. It was shown that both the

nonlinear and linear range radar tracking filters are able to track the threat under these conditions. Threat position and velocity have been adequately tracked once the filter's steady state has been reached. However, the linear filter takes significantly longer to converge.

5.5 Doppler Radar Tracking Filter

Two filters with which to perform Doppler radar airborne tracking using range rate measurements are developed. The filters satisfy two cases according to the filter's knowledge of the existence of correlation between measurement errors in range and range rate:

- Case 1: the filter is unaware of the correlation between measurement errors in range and range rate. That is $\rho^{corr} = 0$. Furthermore, correlation is unaccounted for in the filter.
- Case 2: $\rho^{corr} \neq 0$ and the filter is aware and has accounted for this correlation.

It was discovered in Chapter 2 that in reality there exists a correlation in measurement errors between range and range rate. However, many designs in the literature make the assumption of Case 1, i.e. that the correlation does not exist. The reason for this is that when the correlation is accounted for the filter complexity increases. In what follows we present two Doppler radar tracking filters: first, a filter assuming ideal conditions, addressing Case 1; and second, a filter that accounts for the inherent correlation in measurement errors of range and range rate (Case 2). The reason for developing these two filters is to bring to light any performance differences resulting when correlation is accounted for, compared to when it is not.

Under Case 1 an EKF is used as the tracking filter allowing us to make use of the radar measurements directly. Therefore, tracking is performed in mixed coordinates. The second filter developed (Case 2) is an extension to the first with $\rho^{corr} = -0.9$ which implies negative correlation modelling an upsweep

LFM waveform. It has been proven in [66] that when ρ^{corr} is around -0.9 to significantly reduce state estimate errors, as opposed to positive correlation. Decorrelation is performed akin to the method presented in [44], described in subsection 2.2.3 of the literature review.

The radar measures relative range, azimuth, elevation and range rate. Therefore the corresponding observation vectors are:

$$z_k^R = [\rho_k^R \quad \theta_k^R \quad \varphi_k^R \quad \dot{\rho}_k^R]^T + v_k^R. \quad (5.19)$$

Where v_k^R is the zero-mean, white, Gaussian radar measurement noise with covariance matrix R^R :

$$R^R = \begin{bmatrix} \sigma_\rho^{R^2} & 0 & 0 & 0 \\ 0 & \sigma_\theta^{R^2} & 0 & 0 \\ 0 & 0 & \sigma_\varphi^{R^2} & 0 \\ 0 & 0 & 0 & \sigma_{\dot{\rho}}^{R^2} \end{bmatrix}. \quad (5.20)$$

However, as stated in Chapter 2, measurement errors in range and range rate are correlated. Therefore, in contrast to Equation (5.20), the measurement covariance matrix becomes:

$$R^R = \begin{bmatrix} \sigma_\rho^{R^2} & 0 & 0 & \rho^{corr} \sigma_\rho^R \sigma_{\dot{\rho}}^R \\ 0 & \sigma_\theta^{R^2} & 0 & 0 \\ 0 & 0 & \sigma_\varphi^{R^2} & 0 \\ \rho^{corr} \sigma_\rho^R \sigma_{\dot{\rho}}^R & 0 & 0 & \sigma_{\dot{\rho}}^{R^2} \end{bmatrix}. \quad (5.21)$$

Where the range and range-rate measurement errors are correlated with ρ^{corr} . And σ_ρ , σ_θ , σ_φ and $\sigma_{\dot{\rho}}$ are the measurement errors in range, azimuth, elevation and range rate respectively; whose values are shown below.

Table 5.4: Doppler radar measurement errors.

Sensor	Measurement Error	Value
Doppler Radar	σ_ρ^R	200 m
	σ_θ^R	0.3 radians
	σ_φ^R	0.3 radians
	$\sigma_{\dot{\rho}}^R$	50 m/s

The effect of the range rate measurement error is displayed in the following figures which contain the relative true and Doppler radar measured range rate with respect to the NED system.

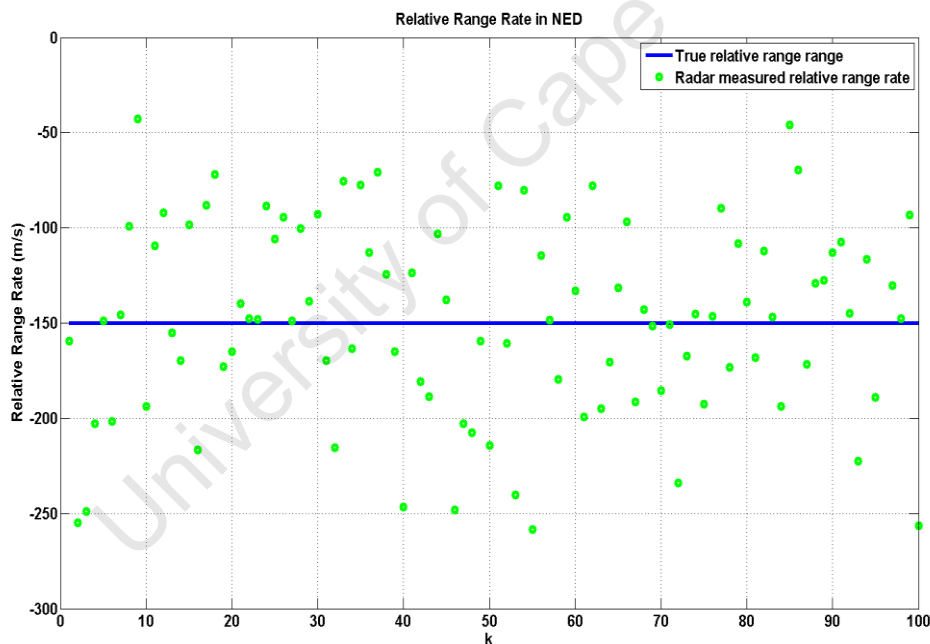


Figure 5.31: Relative range rate in NED - truth and radar measured values.

5.5.1 EKF Under Ideal Conditions

The first Doppler radar tracking filter developed is one that assumes correlation in measurement error do not exist. Range rate measurements are highly nonlinear in Cartesian space and therefore, an EKF is chosen as the tracking filter.

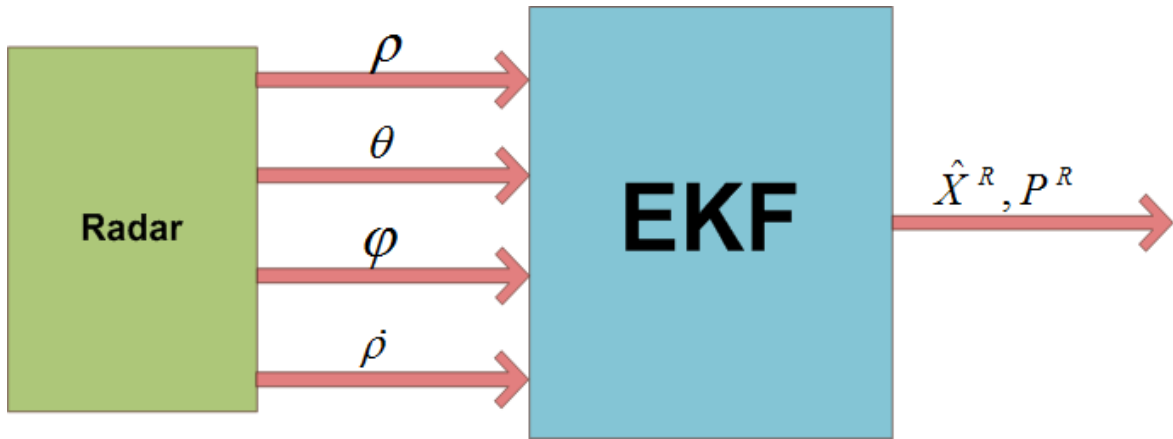


Figure 5.32: EKF with range rate measurements, assuming no correlation.

The filter design is the same as that of the ranging radar described in subsection 5.4.1 above. However, the inclusion of range rate information results in a new observation vector and linearized measurement matrix:

$$h^R(X) = z_k^R = \begin{bmatrix} \sqrt{N^2 + E^2 + D^2} \\ \tan^{-1}\left(\frac{E}{N}\right) \\ -\sin^{-1}\left(\frac{D}{\sqrt{N^2 + E^2 + D^2}}\right) \\ \frac{N\dot{N} + E\dot{E} + D\dot{D}}{\sqrt{N^2 + E^2 + D^2}} \end{bmatrix}, \quad (5.22)$$

$$H_X^R(k) = \left. \frac{\partial h^R(X)}{\partial X} \right|_{X=\hat{X}_{k|k-1}^R}. \quad (5.23)$$

See Appendix A3 for the complete linearized measurement matrix.

As before, at this point the standard EKF equations produce state and covariance estimates. The results of this filter with $\rho^{corr} = 0$ are shown below.

5.5.1.1 Results

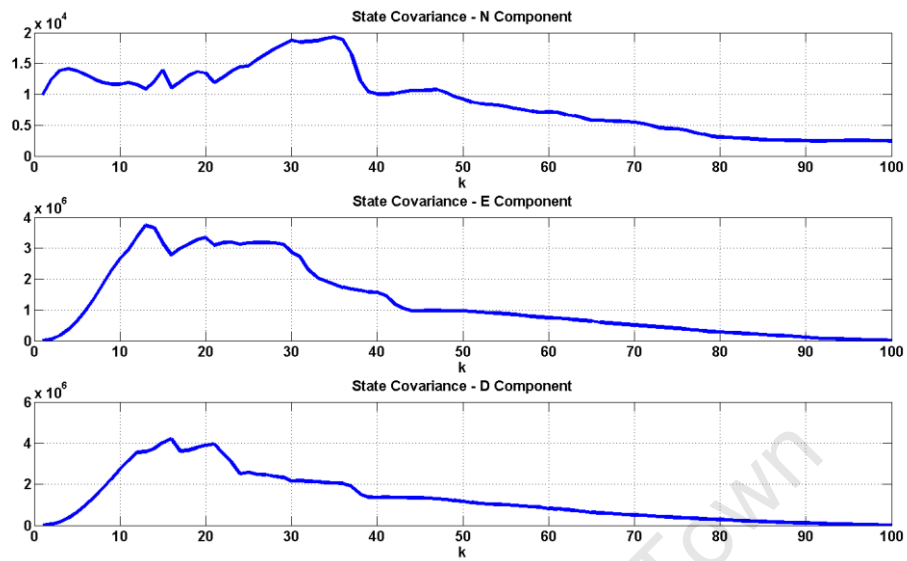


Figure 5.33: Radar EKF with range rate measurements – estimated state error covariance.

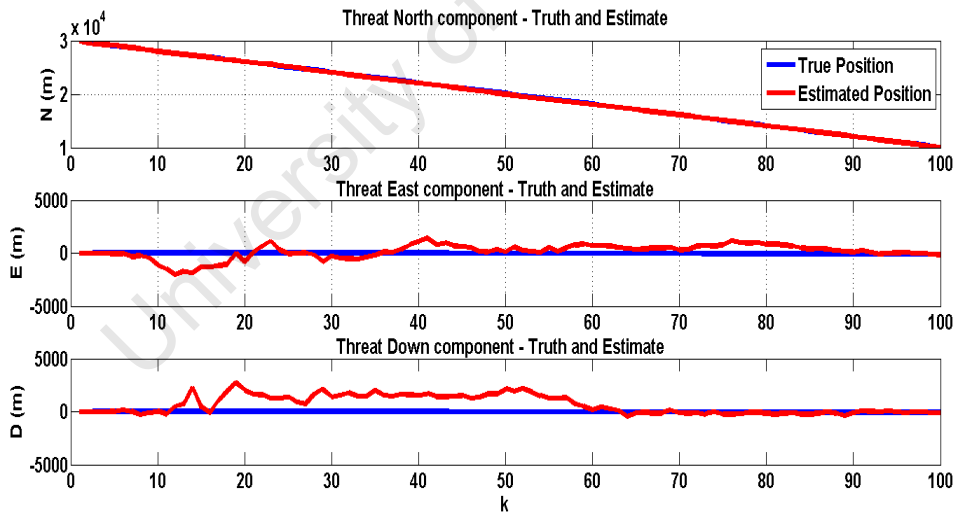


Figure 5.34: Radar EKF with range rate measurements – true and estimated threat position in N, E and D directions.

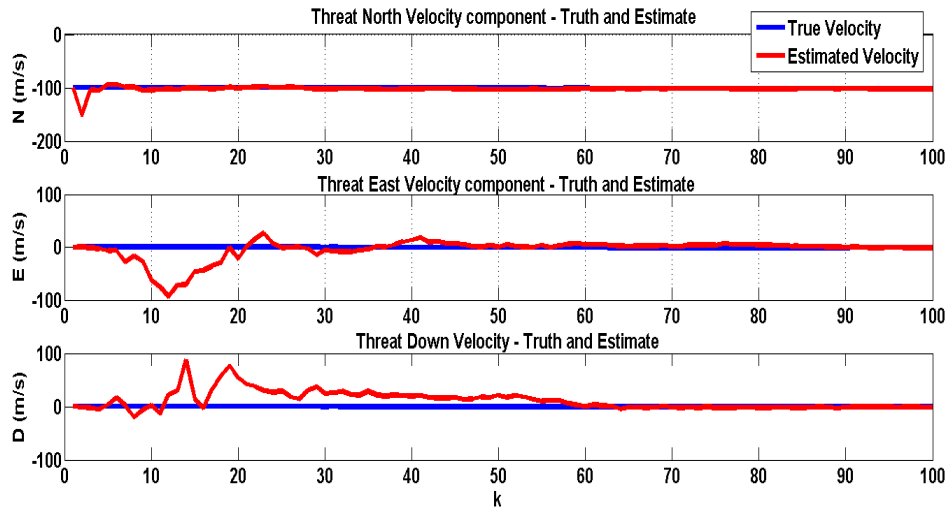


Figure 5.35: Radar EKF with range rate measurements – true and estimated threat velocity in N, E and D directions.

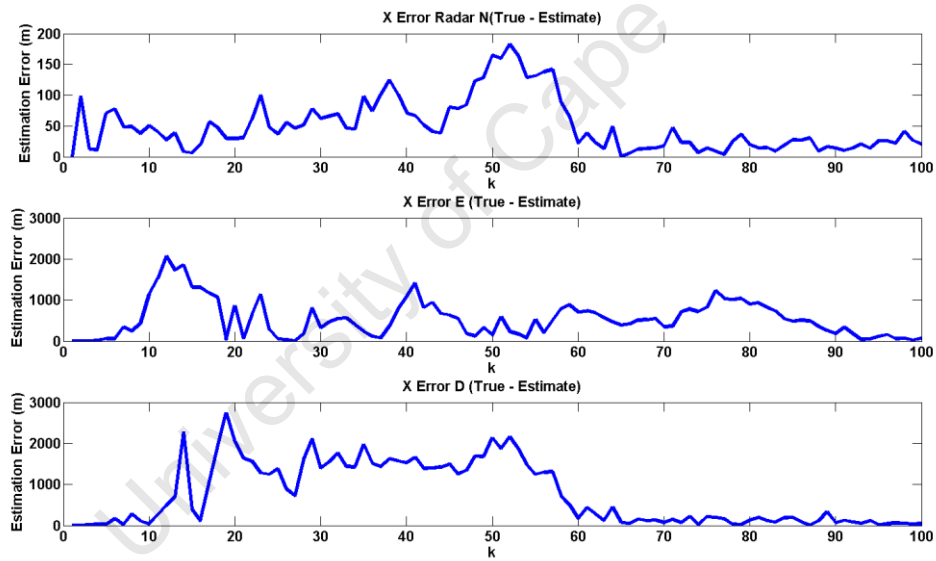


Figure 5.36: Radar EKF with range rate measurements – estimation error in N, E and D directions in a single flight scenario.

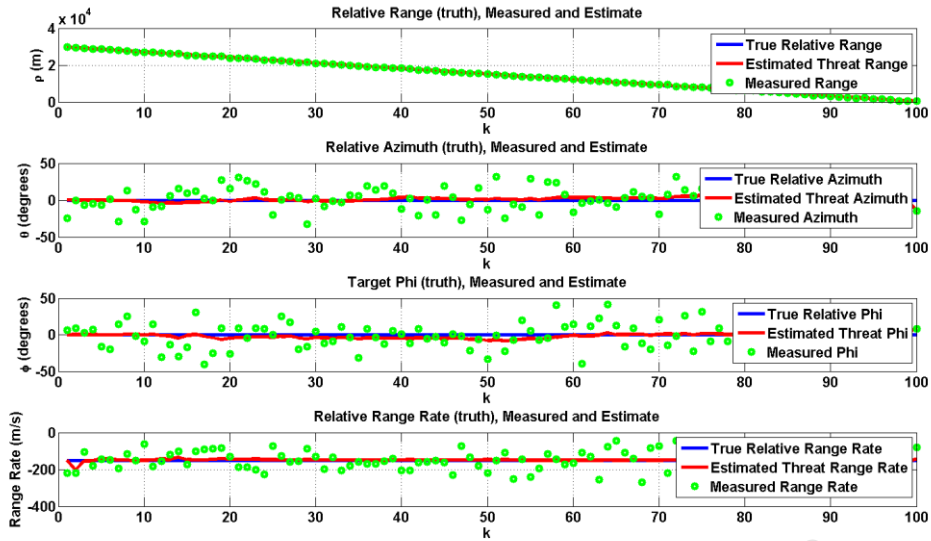


Figure 5.37: Radar EKF with range rate measurements – true, estimated and measured relative range, azimuth, elevation and range rate with respect to the BRF.

5.5.1.2 Monte Carlo Simulations

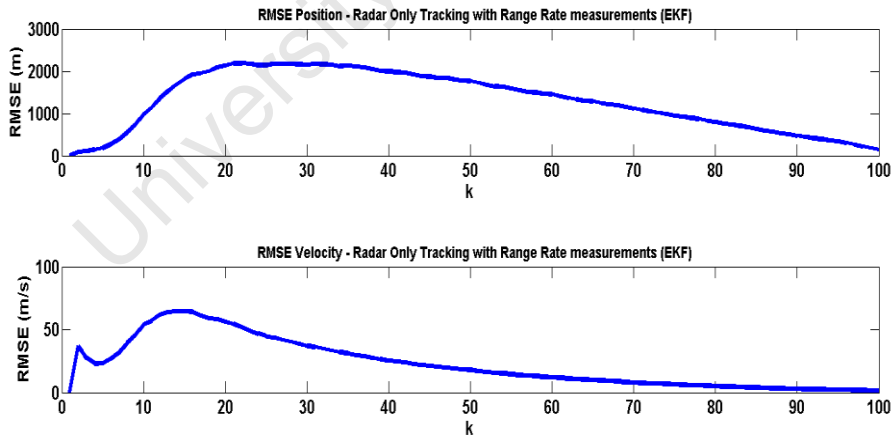


Figure 5.38: Radar EKF with range rate measurements – RMSE in position and velocity.

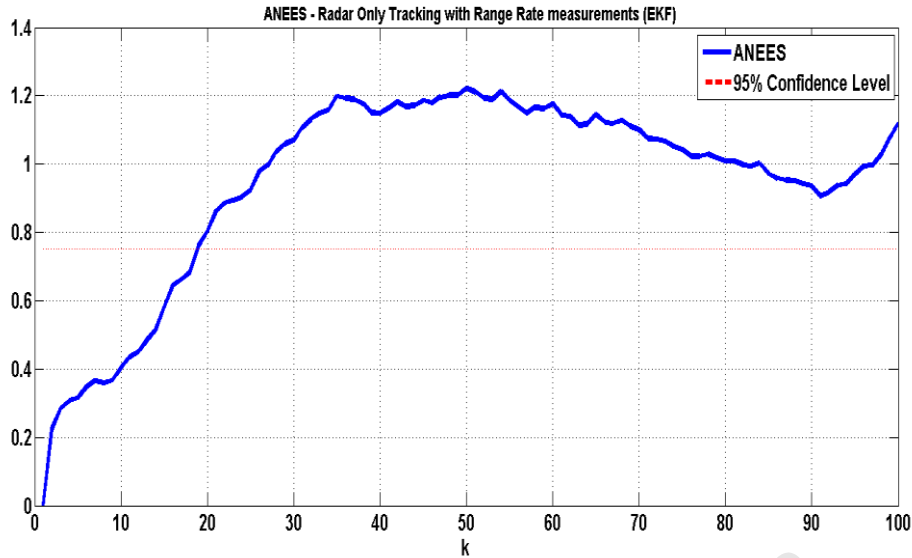


Figure 5.39: Radar EKF with range rate measurements - ANEES.

5.5.2 EKF with Correlated Measurement Errors

In Chapter 2 methods were presented for decorrelation and it was discovered that these filters, which account for the correlation, are more computationally intensive. However, there was no evidence which compared the performance between the two. Therefore a filter which accounted for the correlation was developed.

This filter assumes that measurement errors in range and range rate are correlated with $\rho^{corr} = -0.9$. In order to address this issue, the filter is designed to decorrelate the errors through the formation of a pseudo measurement [55], as discussed in the literature review (see section 2.2.3). This method is similar to the debiased consistent converted measurements method [55] also discussed in the literature review, however tracking is performed in Cartesian coordinates only. The method proposed by Duan et al. [55] was preferred as the conversion to Cartesian coordinates introduces further measurement errors.

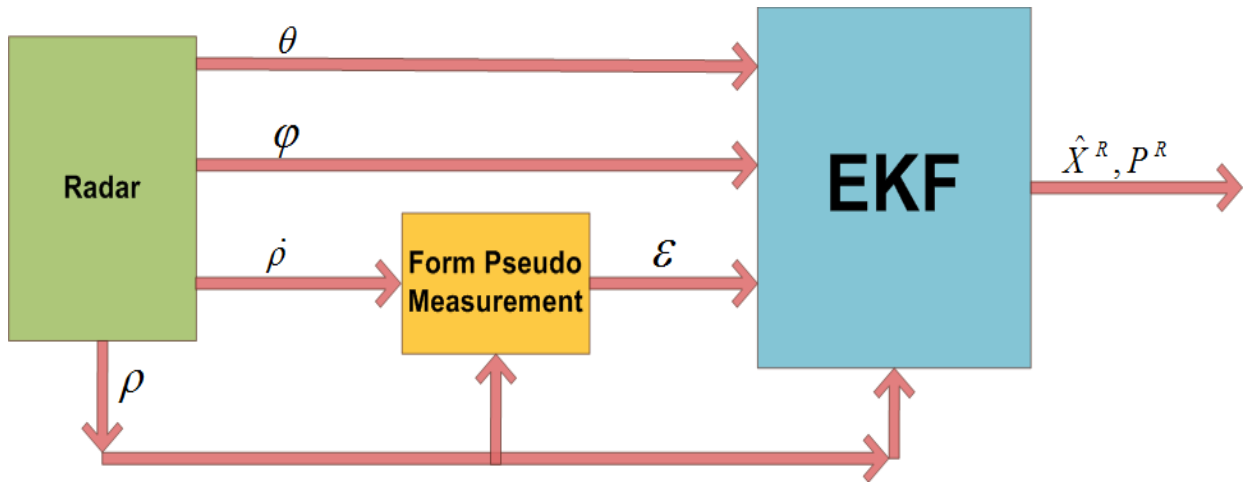


Figure 5.40: EKF with range rate measurements, assuming correlation exists. The EKF takes the range, azimuth, elevation and pseudo measurements as inputs. The pseudo measurement is formed using the range and range rate measurements.

The pseudo measurement formed is:

$$\varepsilon_k = \frac{-\rho^{corr} \sigma_\rho \rho_k}{\sigma_\rho} + \dot{\rho}_k,$$

with covariance:

$$\sigma_\varepsilon^2 = (1 - \rho^{corr2}) \sigma_\rho^2.$$

The new radar observation vector is now:

$$z_k^R = [\rho_k^R \quad \theta_k^R \quad \phi_k^R \quad \varepsilon_k^R]^T + v_k^R. \quad (5.24)$$

Then the linearized measurement matrix is the same as in 5.5.1, and through the use of the EKF equations yield state and covariance estimates. The results of the correlated filter are presented below.

5.5.2.1 Results

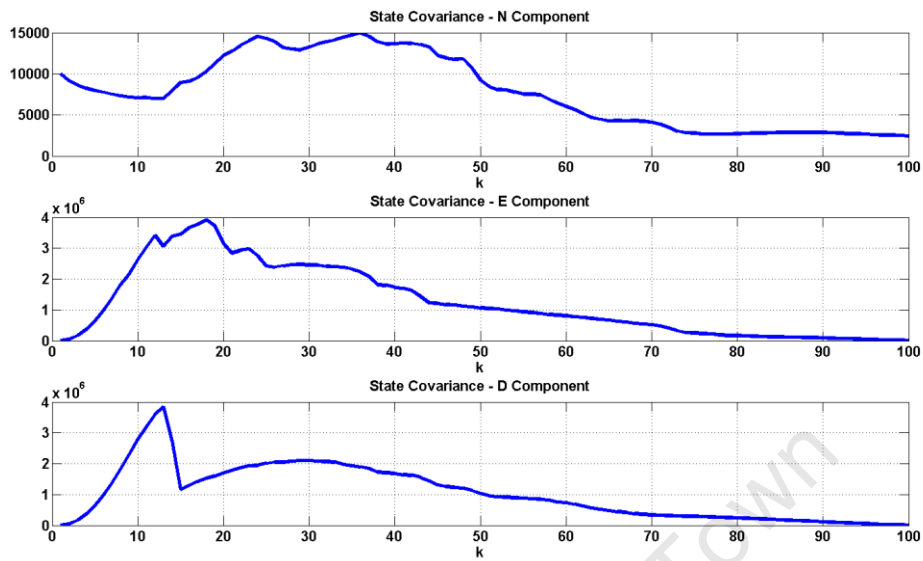


Figure 5.41: Radar EKF with range rate measurements and correlation – estimated state error covariance.



Figure 5.42: Radar EKF with range rate measurements and correlation – true and estimated threat position in N, E and D directions.

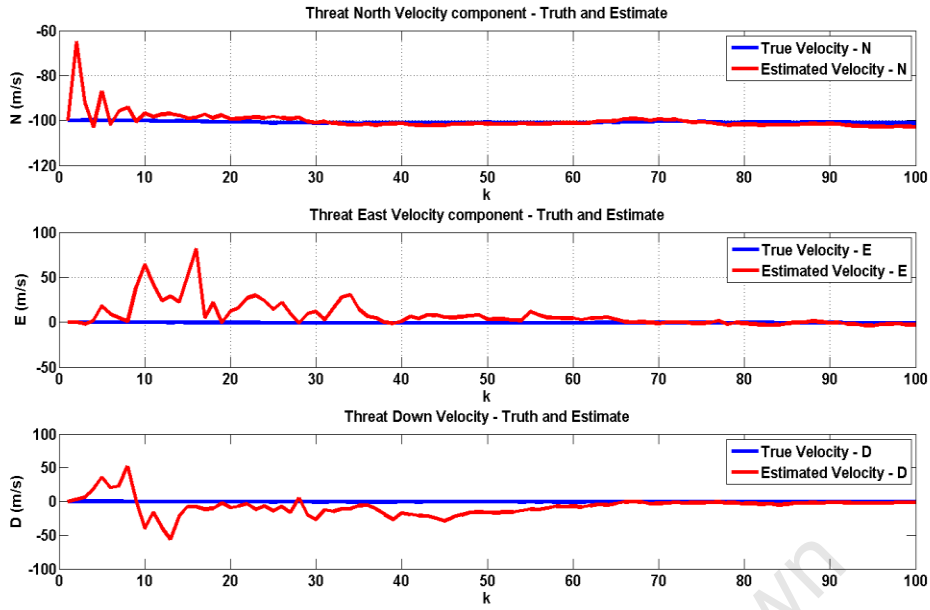


Figure 5.43: Radar EKF with range rate measurements and correlation – true and estimated threat velocity in N, E and D directions.

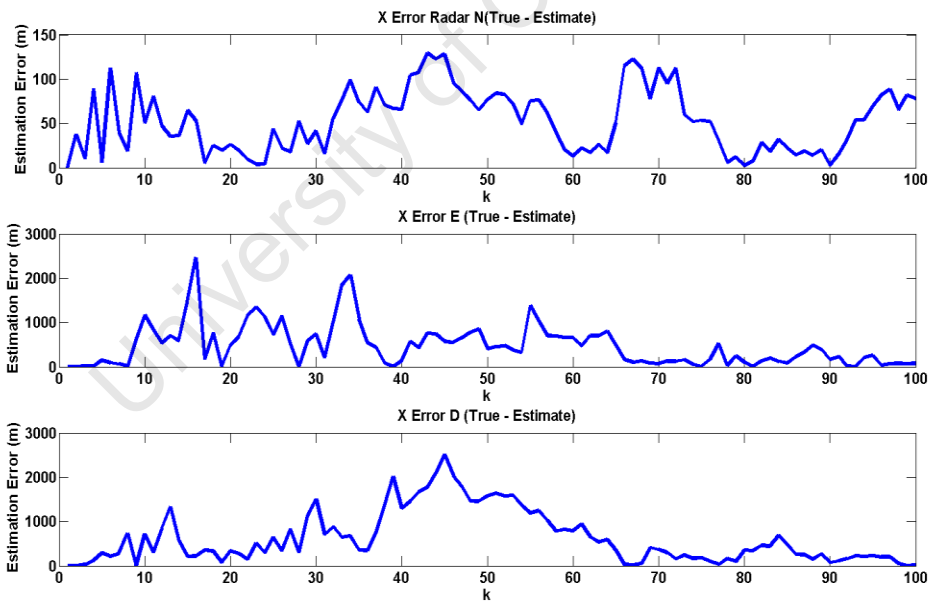


Figure 5.44: Radar EKF with range rate measurements and correlation – estimation error in N, E and D directions in a single flight scenario.

5.5.2.2 Monte Carlo Simulations

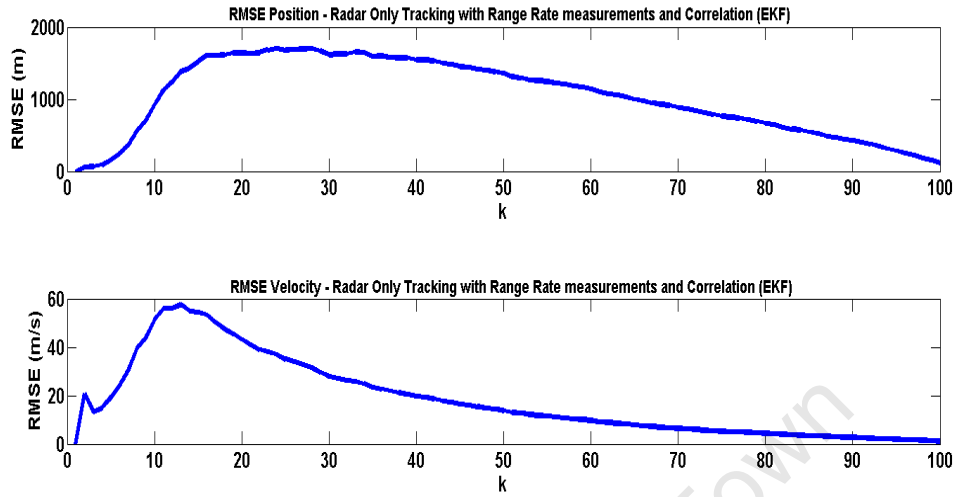


Figure 5.45: Radar EKF with range rate measurements and correlation – RMSE in position and velocity.

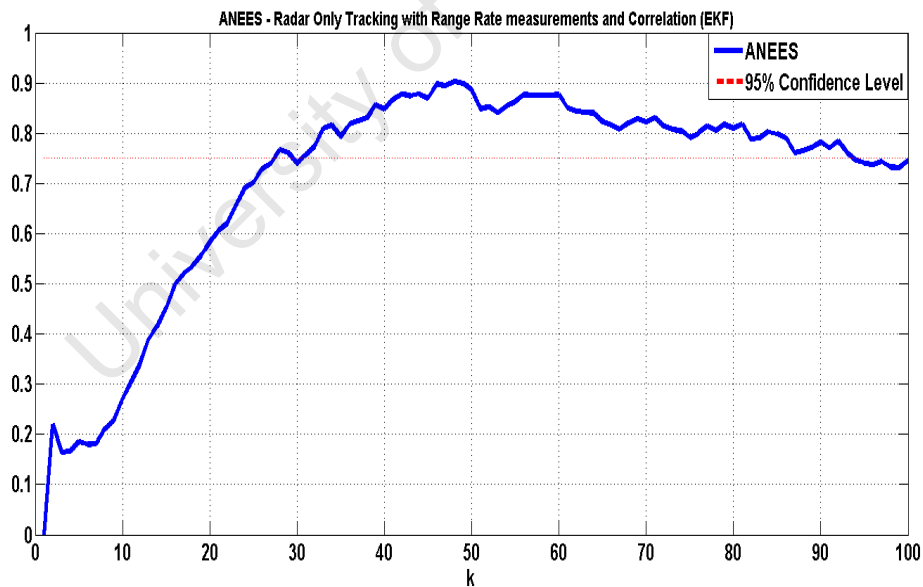


Figure 5.46: Radar EKF with range rate measurements and correlation – ANEES.

5.5.3 Summary

In summary, both tracking filters (Case 1 and Case 2) are able to track the threat under the high noise conditions with the use of range rate measurements; although not particularly well. The range rate measurements increase the accuracy of the velocity estimates in both filter cases. In Case 1, the position tracking error is quite large, until the filter reaches the steady state towards the end of the simulation. This is as a result of the linear approximation of the measurement equation. Case 2 displays similar large position estimation errors. However, the steady state is reached sooner than Case 1 and the estimation error is significantly reduced, as can be seen in Figure 5.42.

5.6 Fusion Filter

One of the objectives of this thesis, stated in Chapter 1, is to demonstrate that a fusion filter more accurately tracks the threat than a single sensor tracking filter. The single sensor filters developed in the previous sections of this chapter satisfy half of this objective. Therefore, in what follows, the design of tracking filters using the measurements from multiple, redundant sensors are detailed.

The TTU is composed of a fusion filter (or fusion architecture). Three fusion architectures are designed, named FA1-FA3, to perform data fusion using an EO camera and a Doppler radar with range rate measurements as set out in Chapter 1. The tracking filters developed thus far form the building blocks from which the fusion architectures are built up with. This section describes all three architectures and presents their results so that, ultimately, a performance comparison can take place to identify the fusion architecture to be used in the TTU.

For a UAV SAS a centralised architecture would, at first thought, make most sense, as sensors and computing platforms are all on-board the UAV.

However, future SAA systems might make use of a network of distributed sensors (for example a network of aiding ground-based radars). Furthermore, current computing platforms provide the potential for running multiple independent filters in parallel. Therefore, distributed fusion architectures have been researched and developed in this section.

5.6.1 Synchronising Sensor Sampling Periods

As stated in section 2.1, the EO camera operates at a higher data rate than the radar. This implies that for each radar measurement received, the TTU would have received $\Delta T^R / \Delta T^C$ many camera measurements. This occurrence can be seen in Figure 5.47. In order to account for this and to synchronise the tracking filter with measurements from radar and camera, data compression is performed as explained in subsection 2.2.4. Data compression is performed in all fusion architectures.

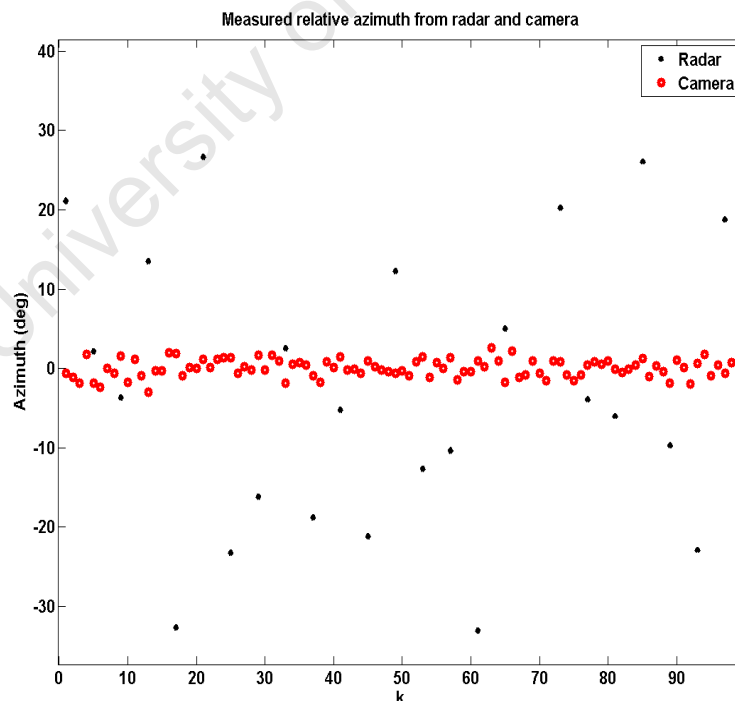


Figure 5.47: Plot of radar and camera azimuth measurements illustrating their contrasting sampling periods. The camera provides measurements four times faster than the radar.

5.6.2 FA1

FA1 is based on a centralised architecture – raw sensor data, from both radar and camera, arrives at the FC which are fused. This fused measurement is then tracked to produce an estimate of the state vector.

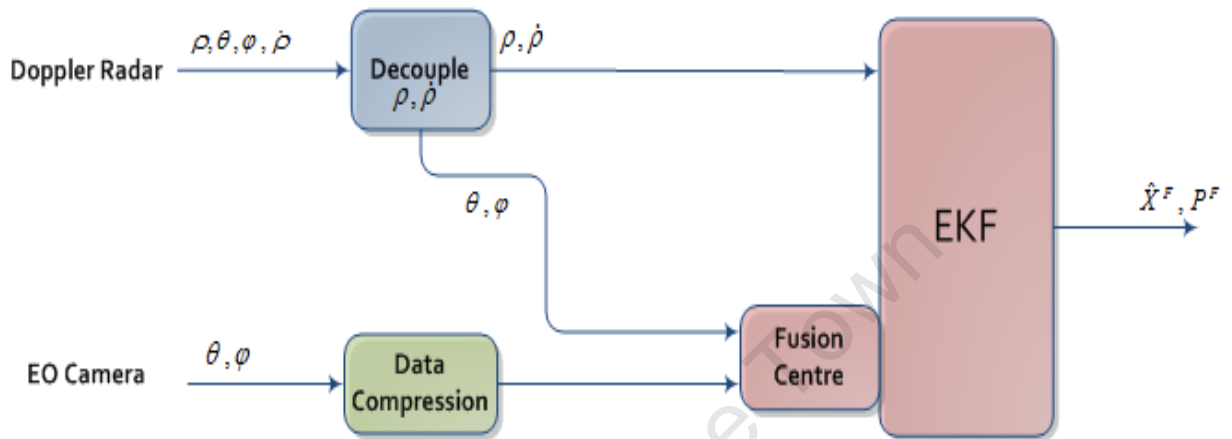


Figure 5.48: Fusion Architecture 3. Measurement fusion is performed in this centralised approach.

Data compression is performed on the camera observations to synchronise its frequency with that of the radar's data rate. Then the range and range rate measurements are temporarily decoupled from the radar observation vector. The reason for doing this is so that the angular components (threat azimuth and elevation) from the radar and camera can be used. The FC fuses this redundant data by utilising measurement fusion, identified in the Literature Review, using Equations (2.1) and (2.2).

The fused measurement and its covariance are then augmented by placing the previously decoupled range and range rate data (and their covariance) back into the fused matrices. These fused-augmented matrices are then input into an EKF for tracking. The EKF implementation used in FA1 is that which was developed in 5.5.1.

5.6.2.1 Results

The results from running FA1 in the simulation environment over 100 discrete time-steps are presented below.

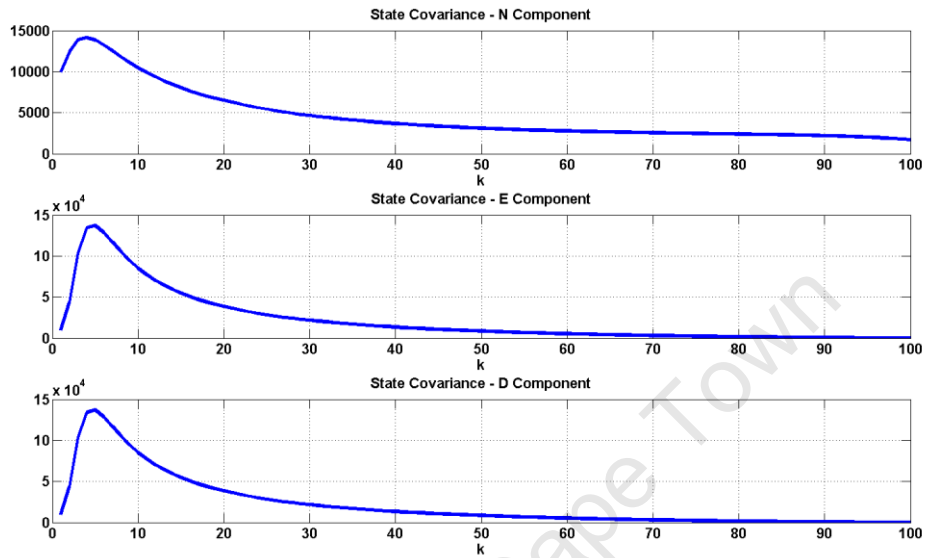


Figure 5.49: FA1 – state error covariance in N, E and D directions.

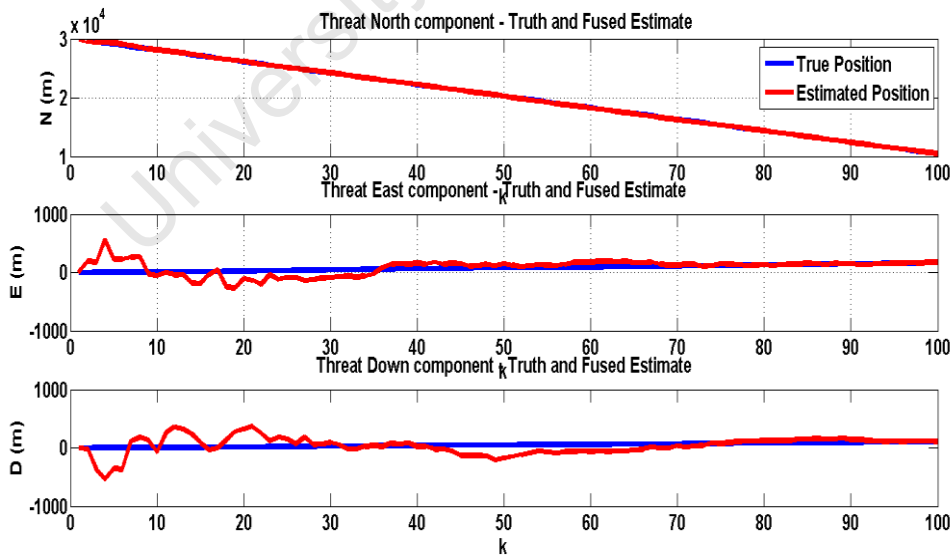


Figure 5.50: FA1 – true and estimated threat position in N, E and D directions.

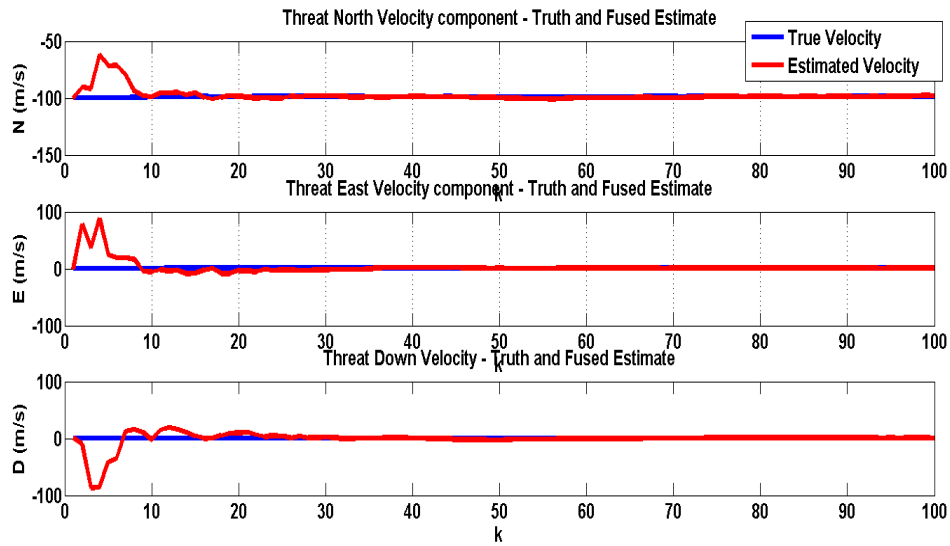


Figure 5.51: FA1 – true and estimated threat velocity in N, E and D directions.

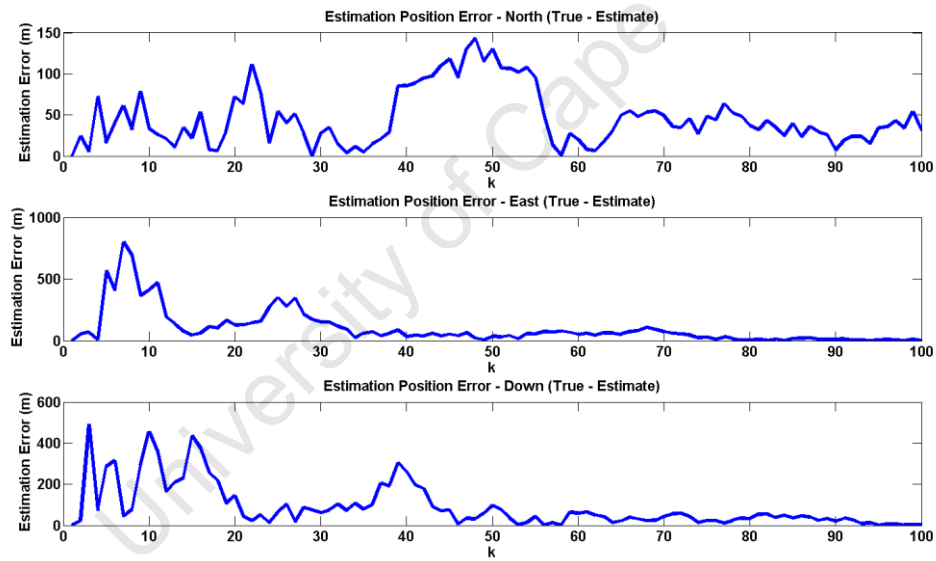


Figure 5.52: FA1 – estimation error in N, E and D directions in a single flight scenario.

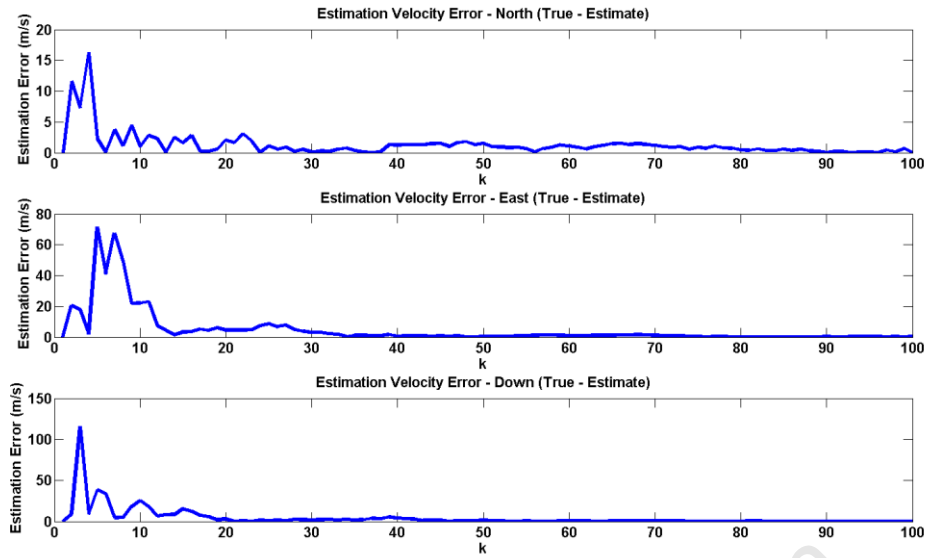


Figure 5.53: FA1 – velocity estimation error in N, E and D directions in a single flight scenario.

5.6.2.2 Monte Carlo Simulations

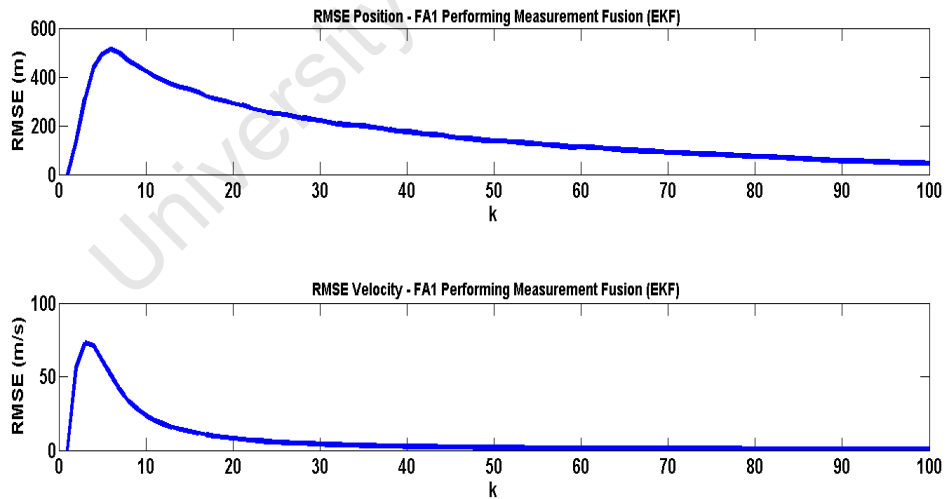


Figure 5.54: FA1 centralised fusion - RMSE in position and velocity.

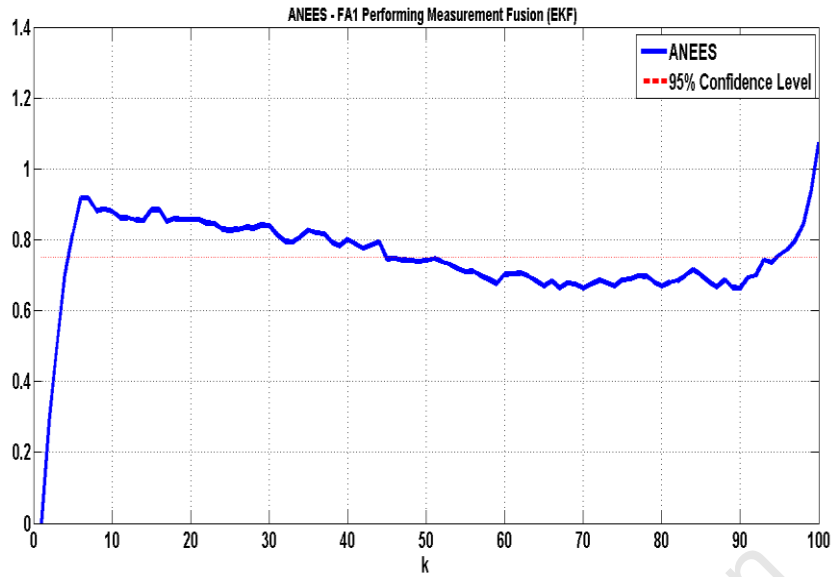


Figure 5.55: FA1 centralised fusion - ANEES.

5.6.3 FA2

FA2 is a distributed fusion architecture (see Section 2.1.1) employing state vector fusion with partial feedback. The measurements from radar and camera are used directly in the target state estimation and therefore FA2 consists of two EKF – one for each local filter. FA2 is illustrated in Figure 5.56.

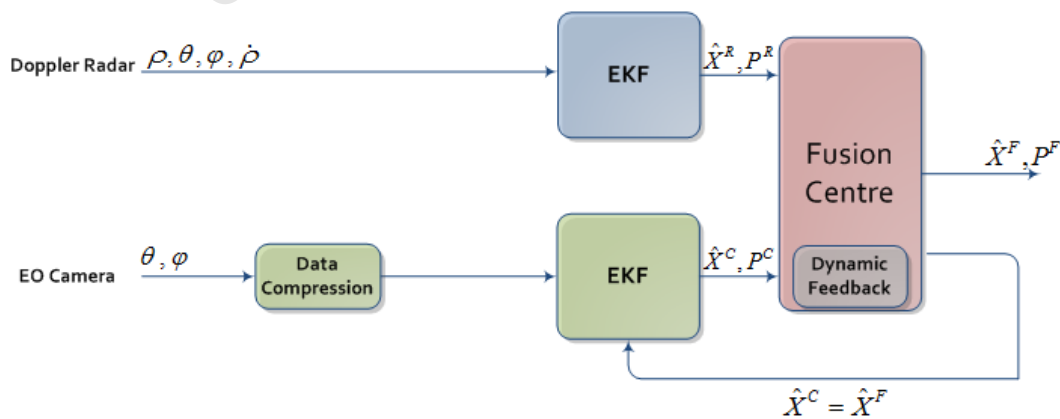


Figure 5.56: Fusion Architecture 2. Both radar and camera local filter are EKFs.

As sensor measurements are nonlinear with respect to the state variables and that, in FA1, they are not manipulated in any way, two EKFs are employed for both radar and camera local filters.

Centralised architectures produce the globally optimal estimate, by minimising the theoretical MSE (an efficient estimator). Distributed architectures can achieve this through the technique of feedback. This entails setting the local estimate to the globally fused estimate. In FA2 we set the camera local filter's state estimates to the globally fused one:

$$\hat{X}_{k|k}^C = \hat{X}_{k|k}^F. \quad (5.25)$$

The reason for doing this is twofold: first so that the architecture does not have a single source of failure, i.e. the whole system does not rely on the FC. If the FC were to fail during operation, then the radar local filter is still able to function correctly and perform tracking. Second, the feedback helps the camera's deficiency in being able to track with its angular only measurements.

The radar local filter is that as designed in Section 5.5.1 and the camera local filter is that designed in Section 5.3.

5.6.3.1 Results

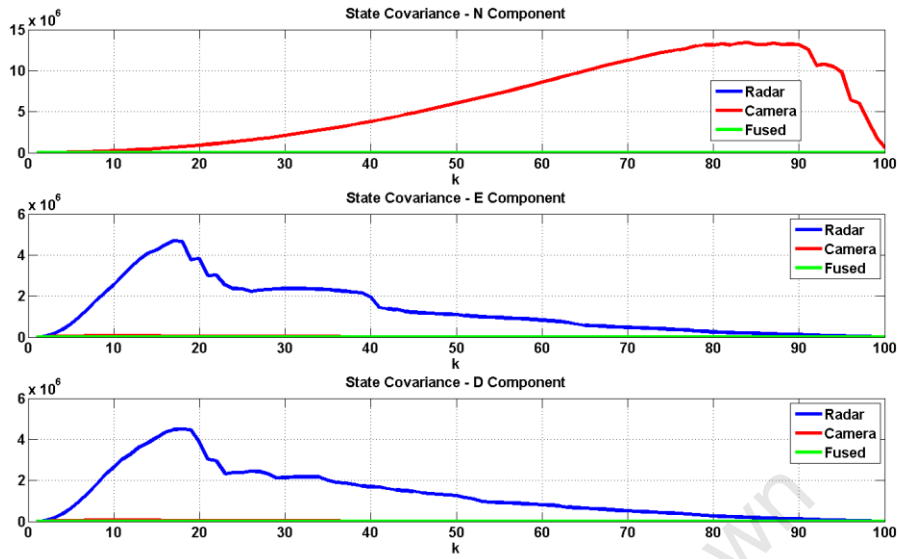


Figure 5.57: FA2 - radar, camera and fused error covariances in N, E and D directions.

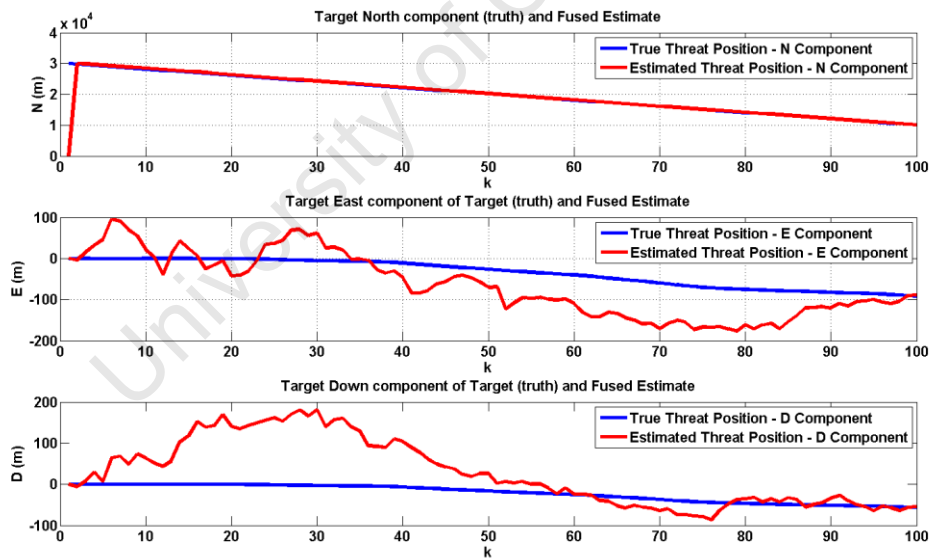


Figure 5.58: FA2 – true and estimated threat position in N, E and D directions.

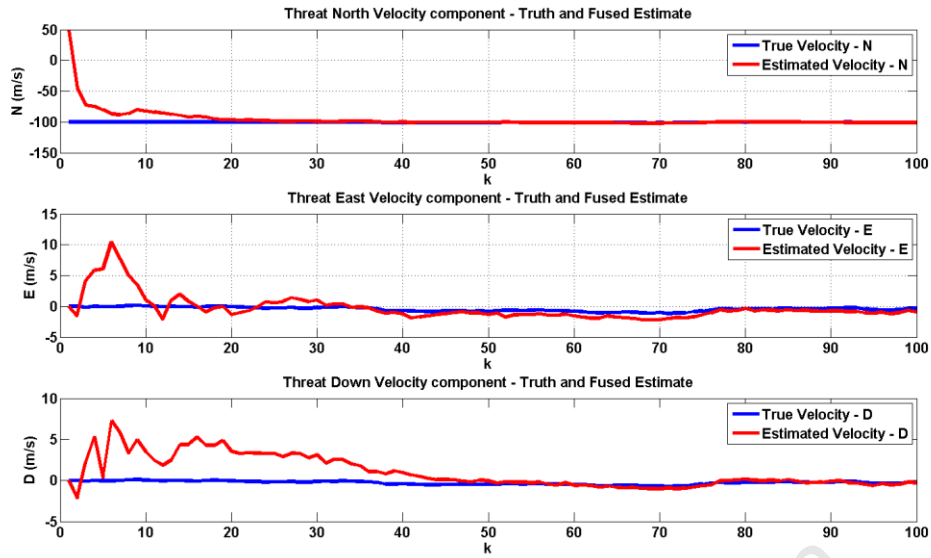


Figure 5.59: FA2 – true and estimated threat velocity in N, E and D directions.

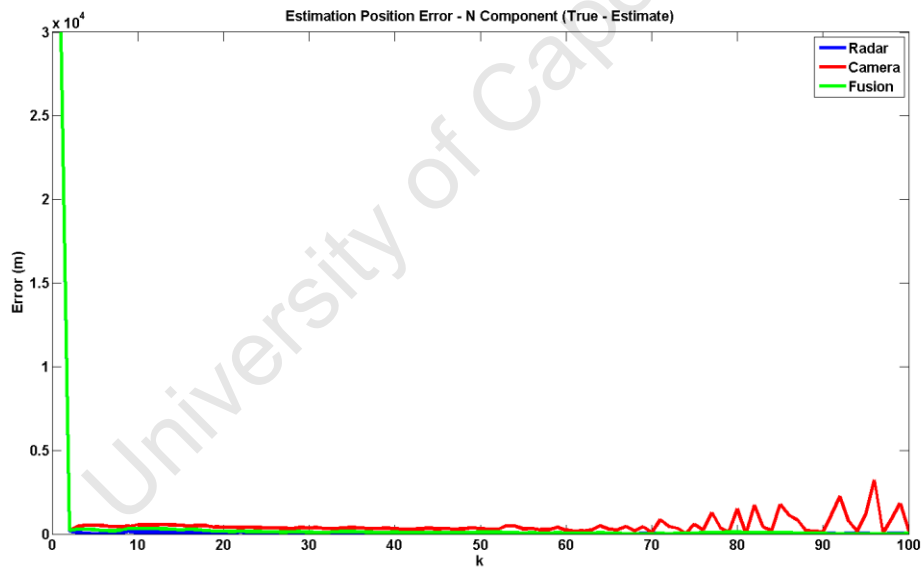


Figure 5.60: FA2 - radar, camera and fused position estimation errors in the N direction.

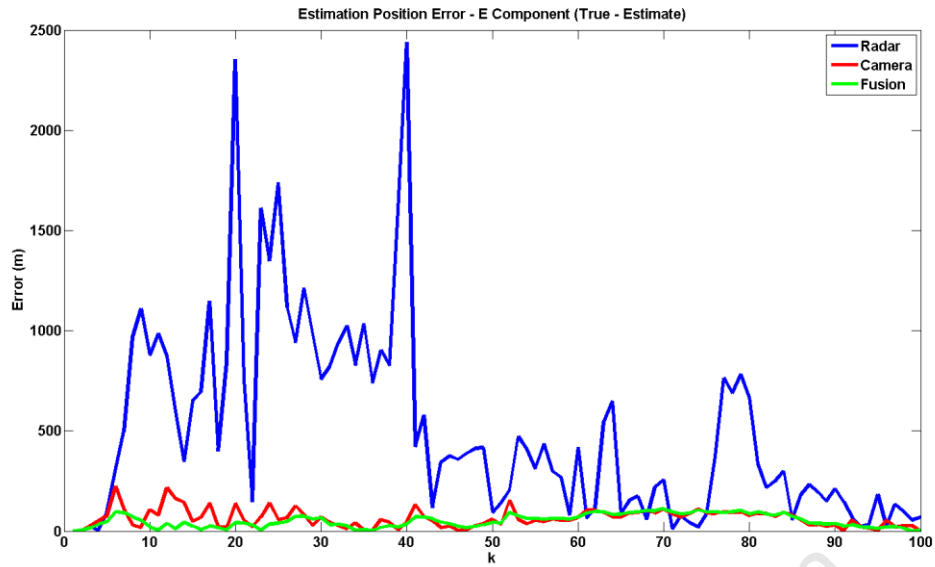


Figure 5.61: FA2 - radar, camera and fused position estimation errors in the E direction.

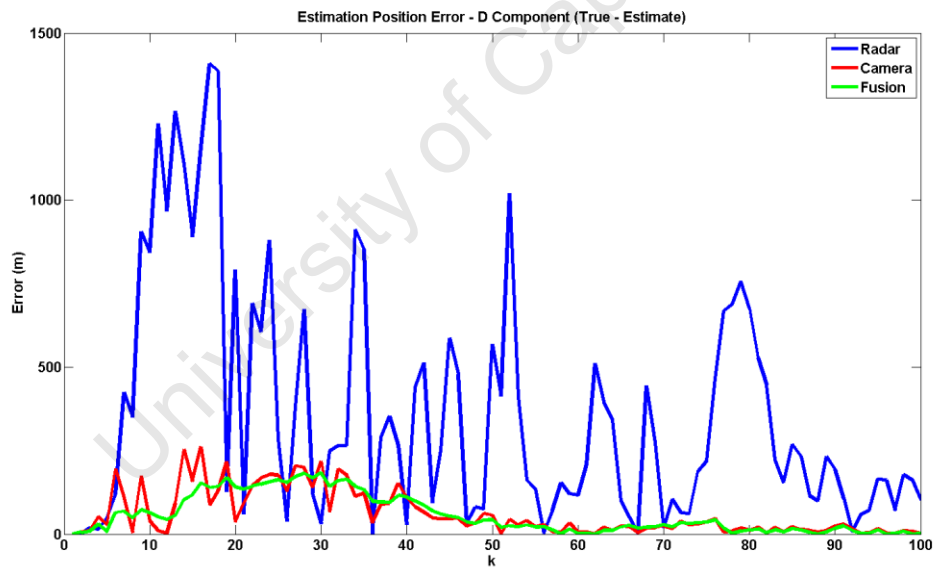


Figure 5.62: FA2 - radar, camera and fused position estimation errors in the D direction.

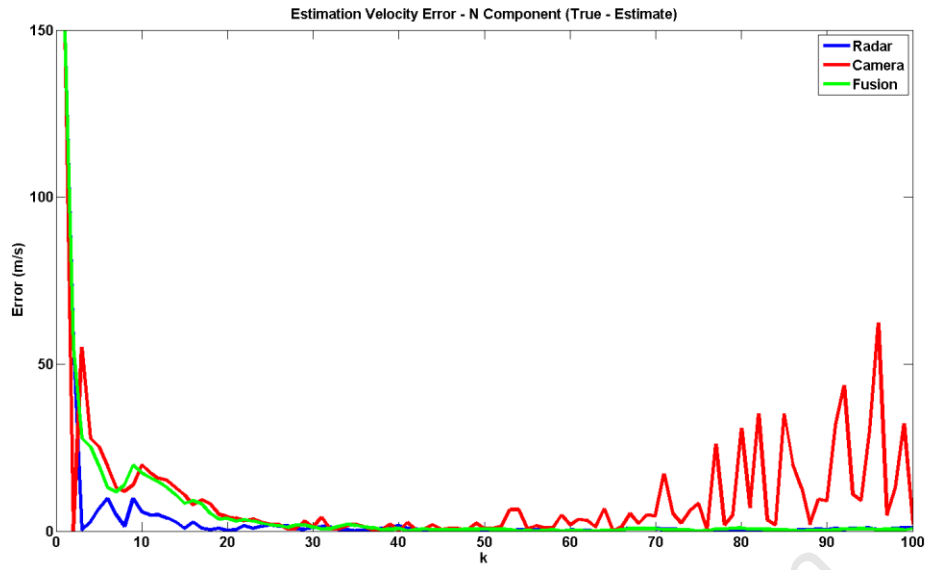


Figure 5.63: FA2 - radar, camera and fused velocity estimation errors in the N direction.

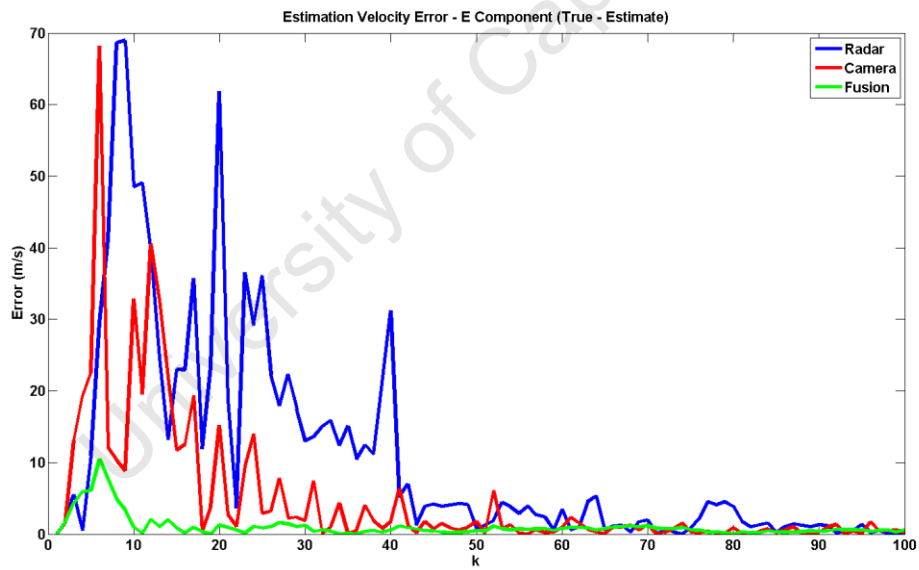


Figure 5.64: FA2 - radar, camera and fused velocity estimation errors in the E direction.

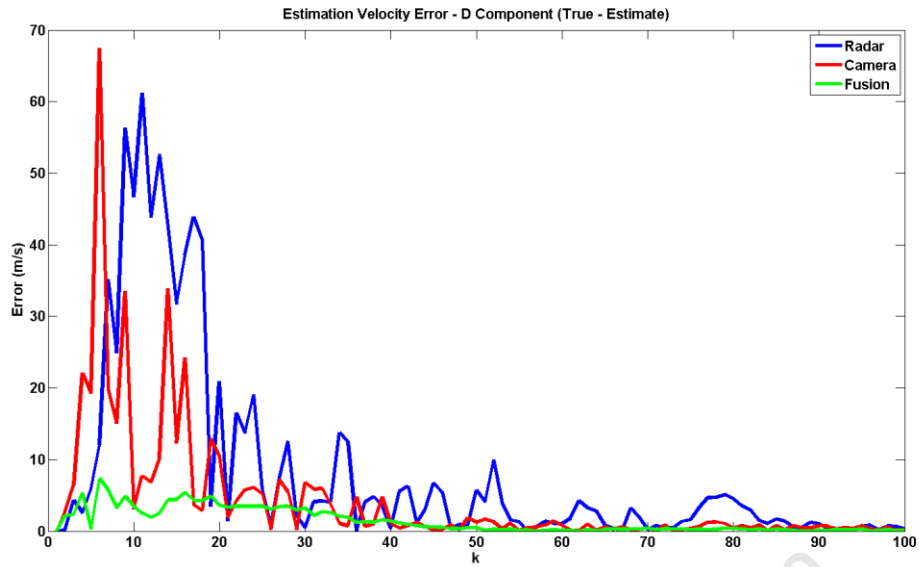


Figure 5.65: FA2 - radar, camera and fused velocity estimation errors in the D direction.

5.6.3.2 Monte Carlo Simulations

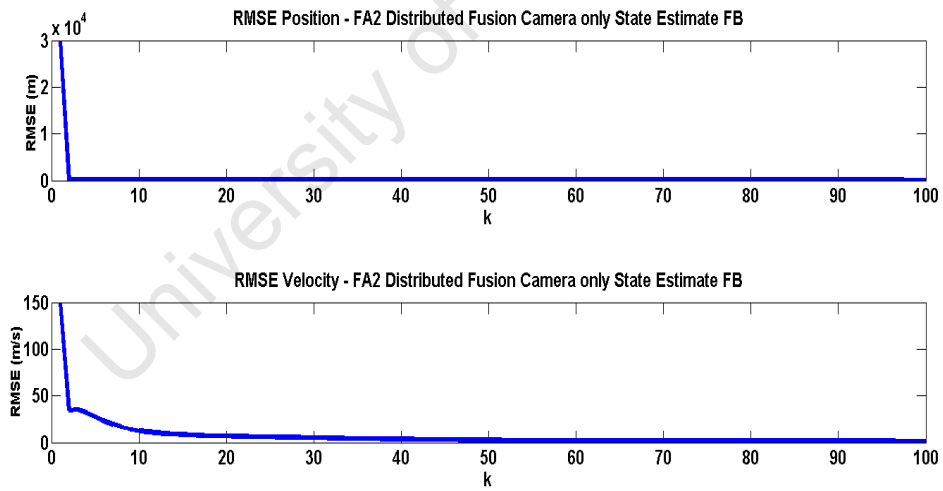


Figure 5.66: FA2 - RMSE in position and velocity.

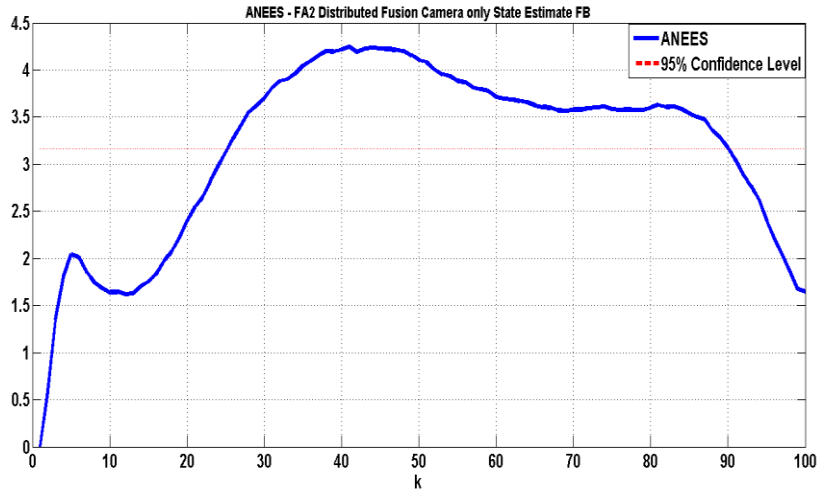


Figure 5.67: FA2 - ANEES.

5.6.4 FA3

Following the fusion system design in [29], we adopt a similar approach albeit with a few modifications. FA3 is another distributed fusion architecture employing state vector fusion with dynamic feedback. However, in contrast to FA2, Radar and camera pseudo measurements are formed that allow us to make use of a linear KF for the radar local filter and an EKF for the EO camera local filter.

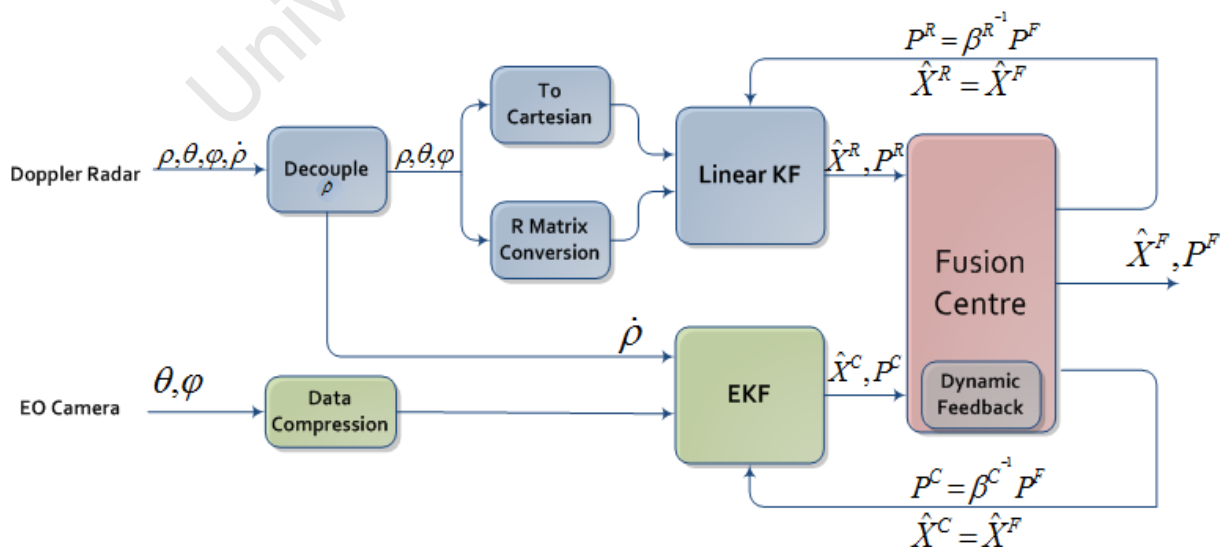


Figure 5.68: Fusion Architecture 3. It consists of a linear KF, an EKF and the Fusion Centre.

FA3 was developed in an attempt to reduce the computational load of FA2, which is comprised of two EKFs. In order to achieve this end, radar and camera pseudo measurements were formed; thus making it possible to replace the nonlinear radar local filter in FA2 with a linear KF.

First range rate measurements are decoupled from the radar camera observation vector and then concatenated with the camera observation vector. This process is illustrated in Figure 5.68. The result of this is the formation of new radar and camera observation vectors in Equation (5.26) and (5.27): radar (pseudo) measurements consists of only relative range azimuth and elevation; and the new camera measurements consists of relative azimuth, elevation and range rate. Therefore a linear KF filter and an EKF are used for the radar and camera local filters respectively:

$$z_k^{PR} = [\rho_k^R \theta_k^R \varphi_k^R]^T + v_k^R \quad (5.26)$$

and

$$z_k^{PC} = [\theta_k^C \varphi_k^C \dot{\rho}_k^R]^T + v_k^C. \quad (5.27)$$

In contrast to FA2, FA3 assumes that the correlation of errors in range and range rate exists and therefore needs to be accounted for. However, through the formation of the pseudo measurements the correlation between range and range rate does not exist anymore as they are a part of separate tracking filters. Therefore, unlike the filter developed in section 5.5.2, no mathematical techniques are required to account for the correlation. Furthermore, this solution sees a decrease in system complexity as a linear filter is used for the radar local filter.

In contrast to FA2, full feedback is performed in FA3, i.e.:

$$\hat{X}_{k|k}^R = \hat{X}_{k|k}^C = \hat{X}_{k|k}^F. \quad (5.28)$$

Furthermore the estimated state covariance is proportionally fed back based on the information distribution coefficient β [72]:

$$P_{k|k}^R = \beta^{R-1} P_{k|k}^F, P_{k|k}^C = \beta^{C-1} P_{k|k}^F. \quad (5.29)$$

Where:

$$\beta^R + \beta^C = 1. \quad (5.30)$$

Instead of fixing the radar and camera information distribution coefficients, the coefficients are dynamically updated. The FC determines whether the uncertainties in radar or camera local filter estimates are higher. Initially the radar and camera distribution coefficients are set to 0.6 and 0.4 respectively. If the principal uncertainty in the radar local filter's estimate is greater than that of the camera, then β^R is more heavily weighted. If the camera filter's uncertainty is greater, then β^C is more heavily weighted. This alteration to the values of the coefficients is performed at each time-step just prior to the feedback step of Equation (5.29).

The linear KF (radar local) used in this architecture was developed in section 5.4.2. However, the EKF (camera local filter) has not yet been as it is only used in FA3. As before, the camera measurement matrix is determined by finding the Jacobian of Equation (5.27) with respect to the state variables in Equation (5.3) and evaluated around the current posterior estimate:

$$h^{PC}(X) = z_k^{PC} = \begin{bmatrix} \tan^{-1}\left(\frac{E}{N}\right) \\ -\sin^{-1}\left(\frac{D}{\sqrt{N^2 + E^2 + D^2}}\right) \\ \frac{N\dot{N} + E\dot{E} + D\dot{D}}{\sqrt{N^2 + E^2 + D^2}} \end{bmatrix}, \quad (5.31)$$

$$H_X^{PC}(k) = \left. \frac{\partial h^{PC}(X)}{\partial X} \right|_{X=\hat{X}_{k|k-1}^F}. \quad (5.32)$$

See Appendix A4 for the complete linearized measurement matrix.

Two other variants of FA3 have also been developed in order to quantitatively present the effects of state feedback, covariance feedback and the dynamic weighting of the information distribution coefficients. Therefore the developed variants are:

- a) FA3a – FA3 with the FC providing no feedback at all to local filters.
- b) FA3b – full feedback (to radar and camera), i.e. state and covariance with constant gain coefficients.

The results of FA3 are presented below.

5.6.4.1 Results

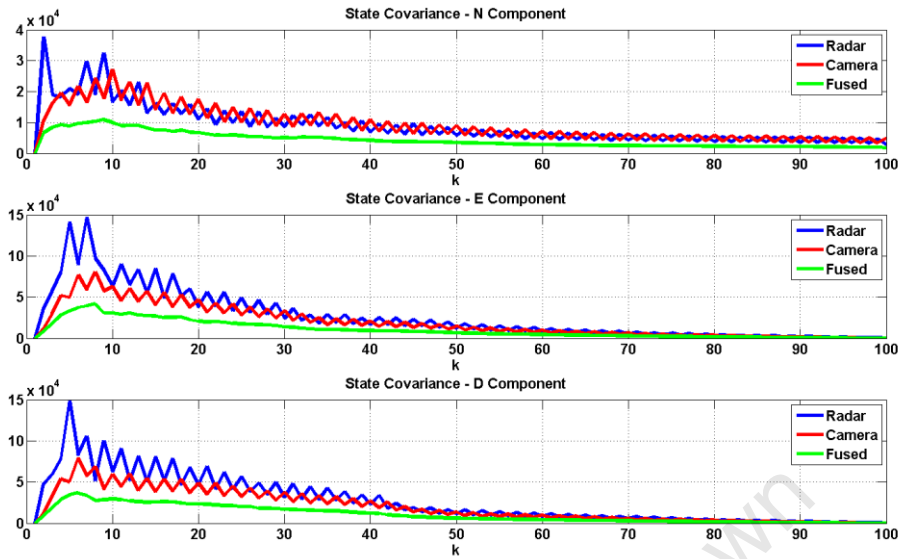


Figure 5.69: FA3 - radar, camera and fused error covariances in N, E and D directions.

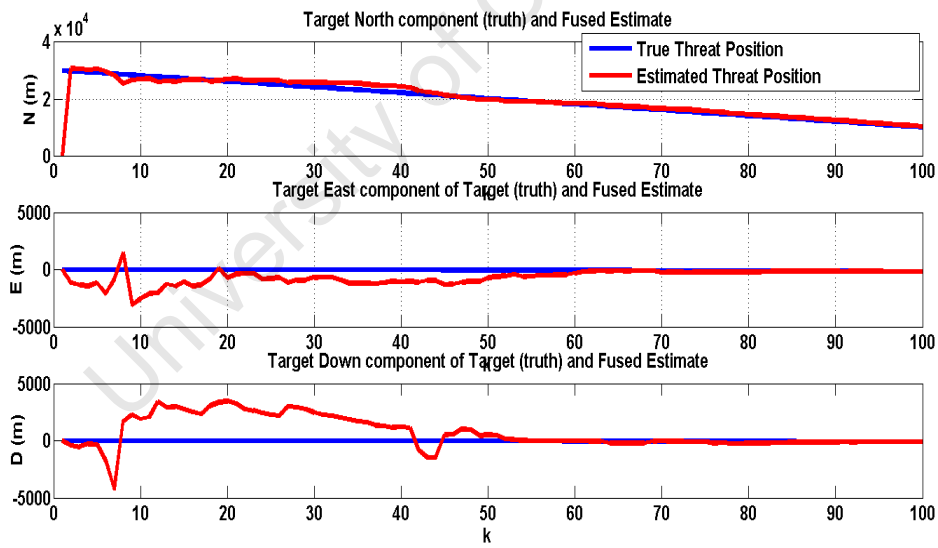


Figure 5.70: FA3 – true and estimated threat position in N, E and D directions.

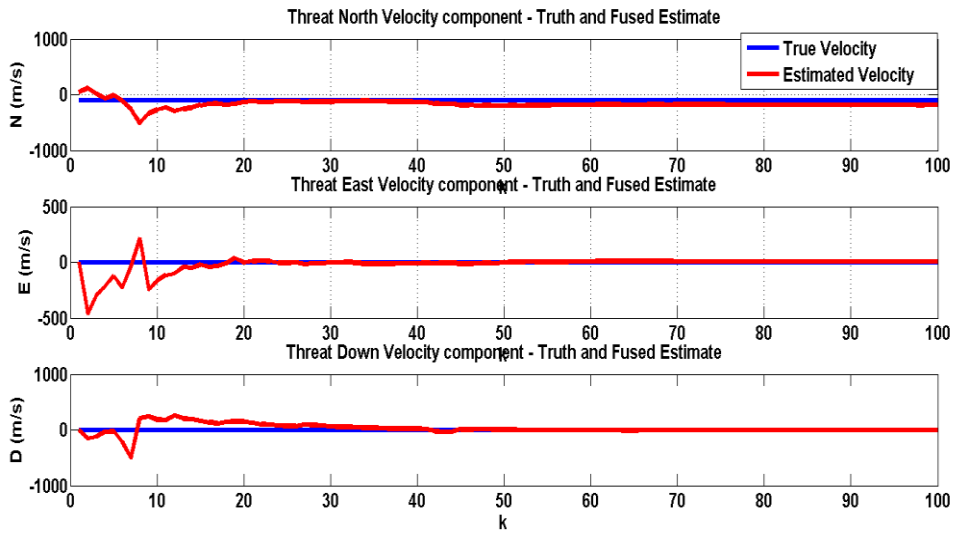


Figure 5.71: FA3 – true and estimated threat velocity in N, E and D directions.

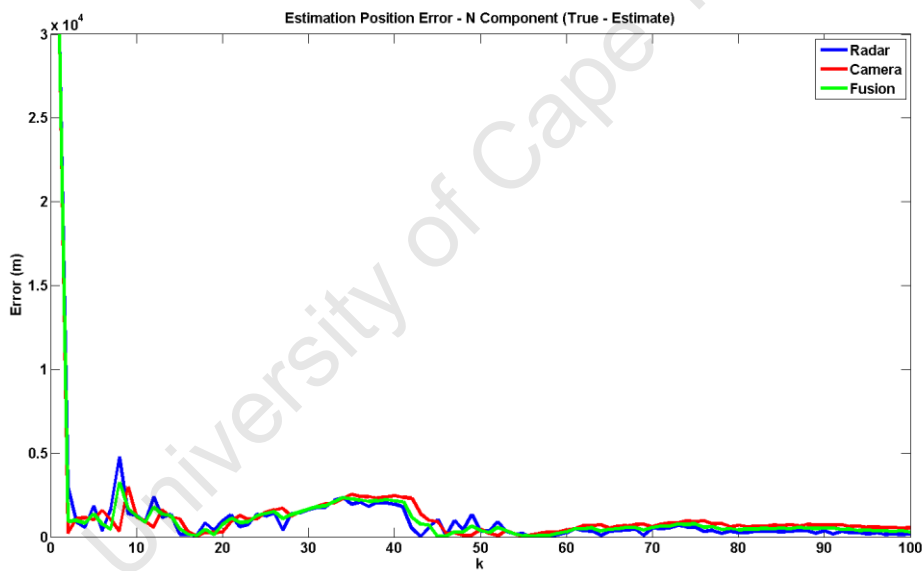


Figure 5.72: FA3 - radar, camera and fused position estimation errors in the N direction.

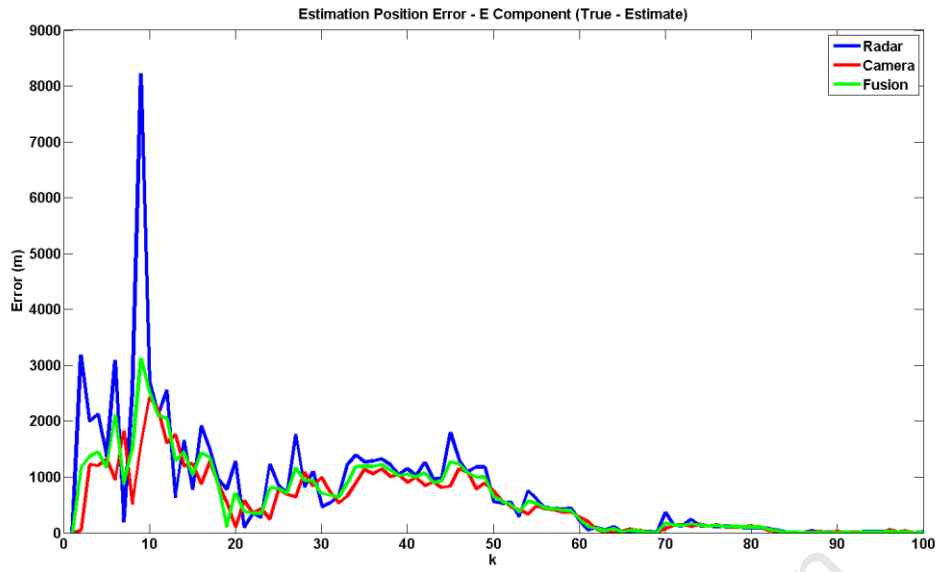


Figure 5.73: FA3 - radar, camera and fused position estimation errors in the E direction.

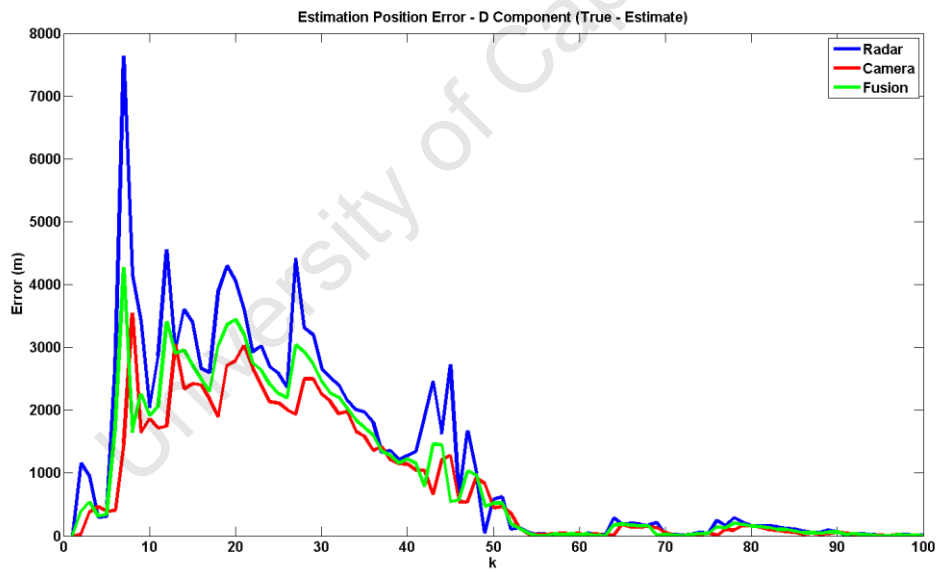


Figure 5.74: FA3 - radar, camera and fused position estimation errors in the D direction.

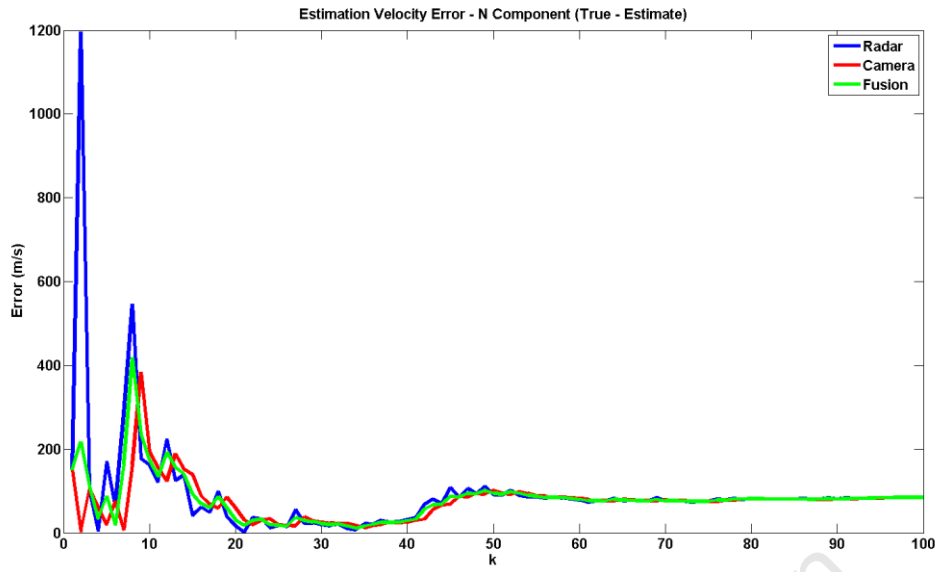


Figure 5.75: FA3 - radar, camera and fused velocity estimation errors in the N direction.

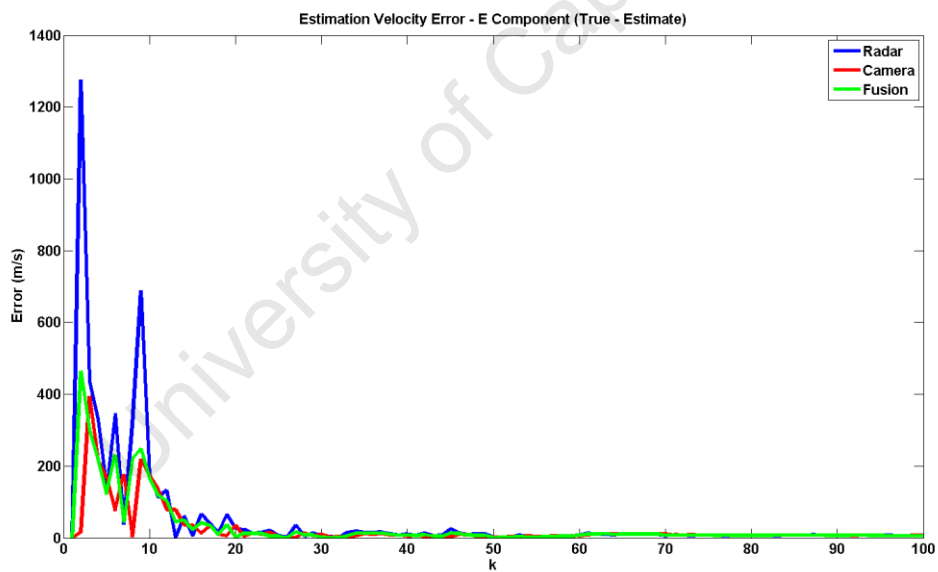


Figure 5.76: FA3 - radar, camera and fused velocity estimation errors in the E direction.

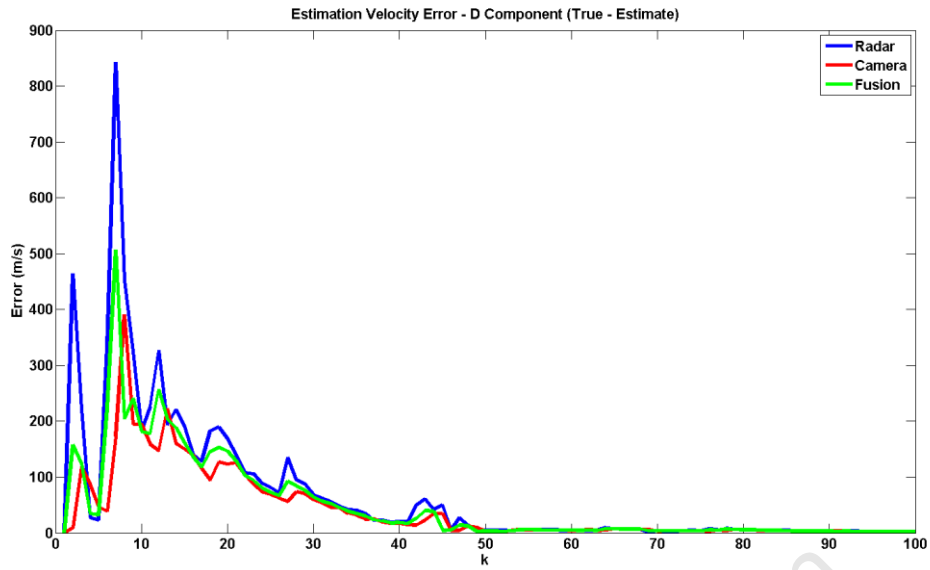


Figure 5.77: FA3 - radar, camera and fused velocity estimation errors in the D direction.

5.6.4.2 Monte Carlo Simulations

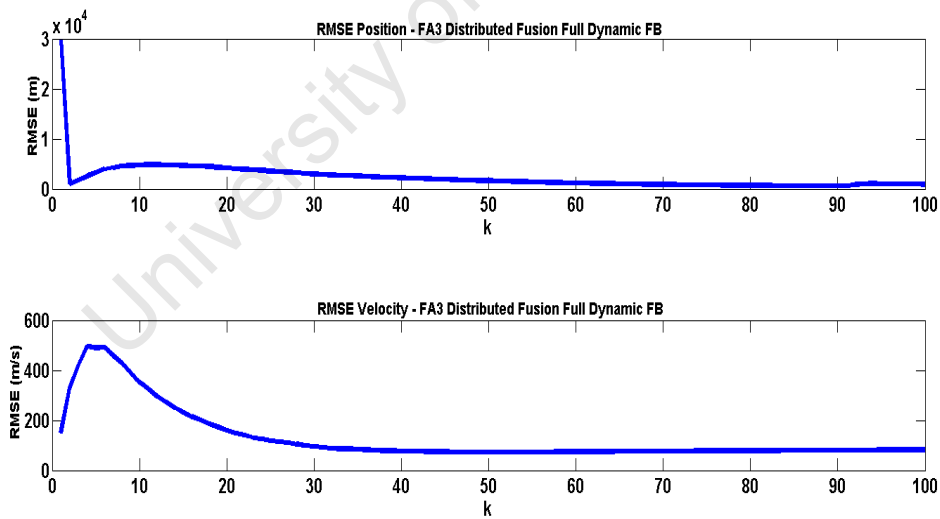


Figure 5.78: FA3 - RMSE in position and velocity.

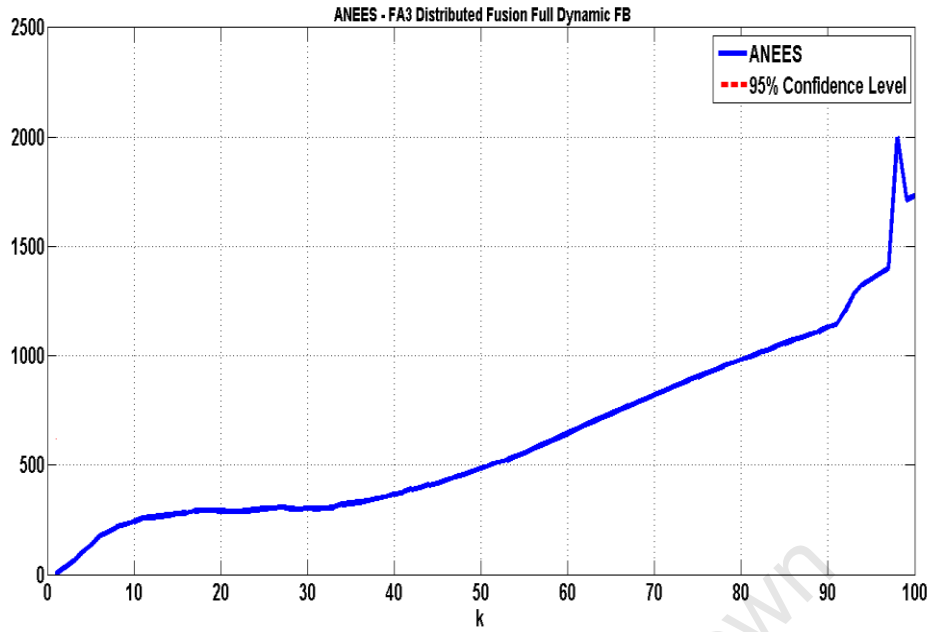


Figure 5.79: FA3 - ANEES.

5.6.5 Summary

In summary, FA1 displayed the most accurate state estimates. FA2 displayed slightly less accurate results. Furthermore, FA3 was substantially less accurate than FA1. These results will be discussed in detail in the following chapter.

Chapter 6: Discussions

This chapter discusses the results of the tracking filters developed in Chapter 5. Furthermore comparisons between like filters are performed.

The comparisons are made between the following airborne tracking filters:

- Range radar linear and nonlinear variants.
- Doppler radar with and without correlation.
- FA1, FA2 and FA3.

6.1 Single Sensor Tracking Filters

The single sensor tracking filters of EO camera, range radar and Doppler radar, presented in sections 5.3, 5.4 and 5.5 respectively, were developed so that their results would form a baseline with which the multisensor filters can be compared against. In addition to this, to prove that filters employing MSDF are more accurate than their single sensor counterparts.

6.1.1 EO Camera

It is evident from Figure 5.5 that the EO camera-only filter diverges, despite its very accurate angular measurements. This is because our state vector is not observable with only the angular measurements provided by the EO camera. This lack of observability makes the camera-only filter, on its own, unable to track the motion of the threat. In fact, the effects of this can be seen by illustrating the true threat motion versus the filter estimated motion in the NE plane.

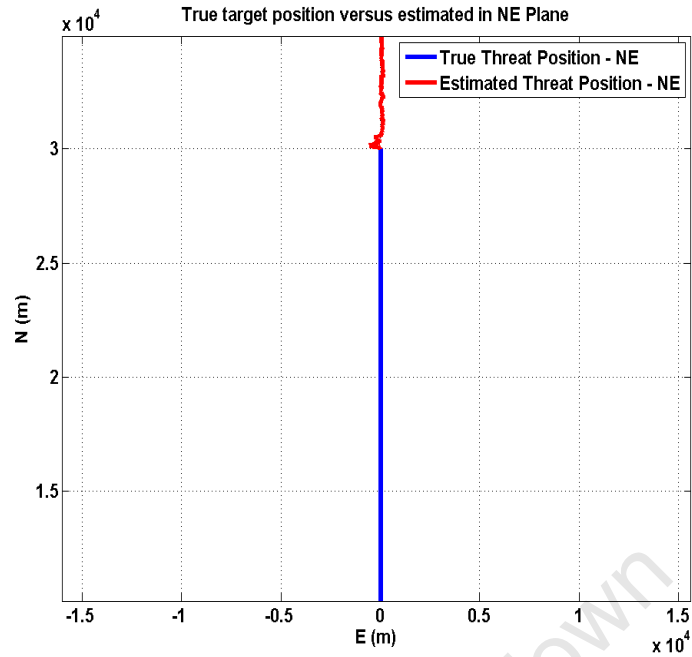


Figure 6.1: Camera filter - True and estimated threat position in the NE plane.

The filter starts tracking the threat correctly, but soon after the filter estimates that the threat is flying due north when it is actually travelling due south. This is the case even though the threat's east and down position components are adequately tracked as seen in Figure 5.6. This makes sense as the position of the threat at every discrete time is a 3-vector of position in the north, east and down directions respectively. Therefore if the N component is estimated incorrectly, as this filter does, then it implies that the threat is in an entirely different position than it actually is.

The MCS further confirms this behaviour, Figure 5.9 shows that the RMSE in position (RMSEP) is 20 km.

It is worth noting that, other than the camera tracking filter, all other filters converge to the steady state.

6.1.2 Radar

6.1.2.1 Without Range Rate Measurements

Two radar tracking filters were developed in sections 5.4.1 and 5.4.2 respectively: first a nonlinear filter based on an EKF and the other based on a linear KF.

The nonlinear filter was developed first as it is the most intuitive – one that uses the measurements directly from the radar; determines the linearized measurement matrix; computes the Kalman gain; and performs state and covariance estimation. Nonlinear filters are, however, computationally intensive. Therefore, the linear approach was developed to investigate its performance relative to that of the EKF based approach.

It can be seen that from the true and estimated position and velocity components, the nonlinear radar filter is adequately tracking the threat, once the steady state has been reached, in Figure 5.17 and Figure 5.18. This is reflected by projecting the true and estimated position in the NE plane in Figure 6.2. Furthermore, as the radar provides measurements in the sensor frame (BRF), the relative range, azimuth and elevation measured by the radar is contrasted against the threat's same components in the BRF in Figure 5.16. It is evident that the EKF filters out the radar sensor's noise to produce accurate estimates.

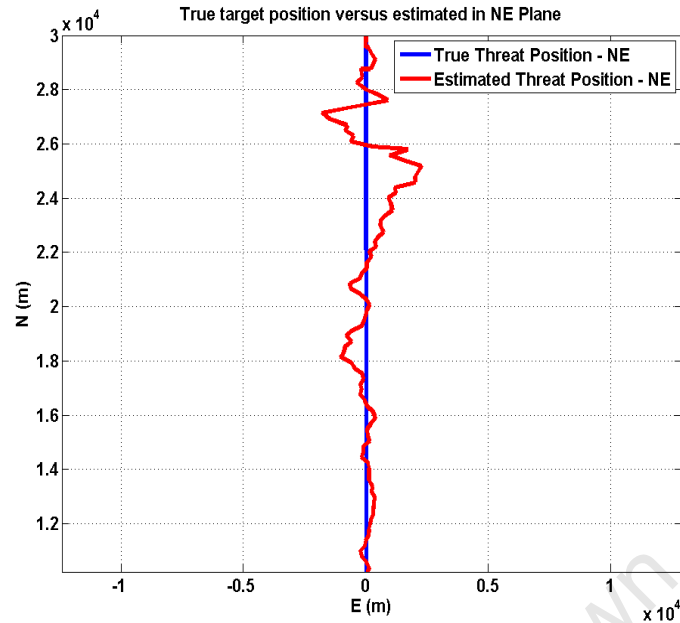


Figure 6.2: Radar nonlinear filter (EKF) - True and estimated threat position in the NE plane.

Similar conclusions can be drawn from the linear KF, whose estimated position and velocity accurately track the true threat motion in Figure 5.25 and Figure 5.26. However, the earlier estimates produced by the filter are quite far from the truth. This is attributed to the introduction of errors as a result of the coordinate transform to Cartesian coordinates. Considering that this conversion, in our simulation environment, is the last in line of many transformations and conversions (see Chapter 4), these errors eventually compound resulting in the above observation. Furthermore, as explained at the beginning of this chapter, the radar measurement error covariance matrix is set to reflect high noise in the radar measurements, and this matrix is also transformed to Cartesian space. This effect can easily be seen in Figure 6.3. This projection introduces further errors. In spite of this, the filter is able to recover from these effects and eventually produce accurate estimates.

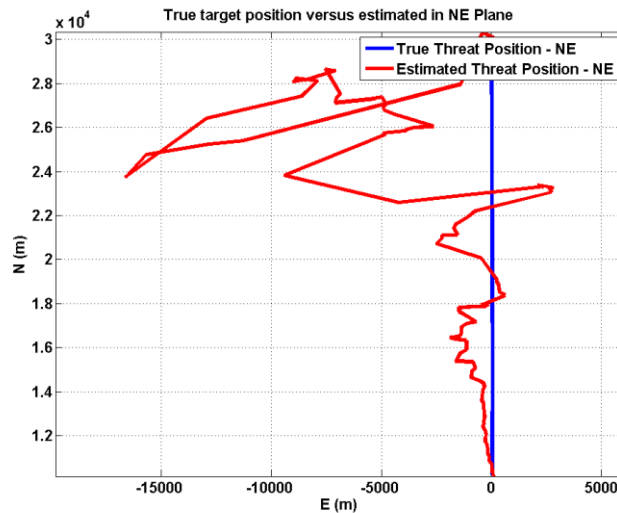


Figure 6.3: Radar linear filter (KF) tracking with Cartesian coordinates - True and estimated threat position in the NE plane.

The RMSE of the EKF in Figure 5.20 and KF in Figure 5.29 shows that both filters are accurate as the estimation error reaches an acceptable level. Moreover, it is evident that the error decreases every time-step. Therefore both filters are *consistent*, that is the filters produce increasingly accurate estimates over time.

The state error covariance of the nonlinear EKF and linear KF in Figure 5.15 and Figure 5.30 show that both filters converge. However, the KF reaches the steady-state at around $k = 40$, a lot sooner than the EKF. At this point the filter believes it is producing accurate state estimates and therefore more heavily weights the predicted measurement (opposed to the real measurement) in the estimation of the state vector; a consequence of the dynamic Kalman gain calculation.

That being said, care should be taken when interpreting the error covariance. The error covariance is the filter's self-assessment of how well it believes it is performing. The EKF's and KF's ANEES are illustrated in Figure 5.21 and Figure 5.30 respectively. The ANEES provides a measure of how true the filter produced error covariance is. It is clear that the EKF is credible as the

ANEEES is approximately equal to 1. However, the KF is non-credible. Therefore the estimated KF state covariance matrix is not a reflection of the truth. In fact, the filter is optimistic; it believes the estimation error is smaller than it actually is. This is apparent from the filter's estimated error covariance in Figure 5.23. Note that this does not mean that the estimates produced by the filter are not accurate, but rather that the filter is not aware of how accurate (or inaccurate) the estimates actually are.

6.1.2.2 With Range Rate Measurements

Two radar tracking filters with range rate measurements were developed in Chapter 5: one where a correlation between measurement errors in range and rate exists and the other where the correlation does not exist. Both filters make use of an EKF due to the high nonlinearity of the range rate measurement in Cartesian space; recall that these filters track with mixed coordinates. That is, the relative motion dynamics is modelled in Cartesian coordinates, whereas the radar measurements are in spherical sensor coordinates.

It is evident from the true and estimated position and velocity components, that both filters are adequately tracking the threat once the steady state has been reached, when correlation does not exist and when accounted for in Figure 5.34, Figure 5.35 and Figure 5.42, Figure 5.43 respectively. This is reflected by projecting the true and estimated position in the NE plane of both filters in Figure 6.4 and Figure 6.5.

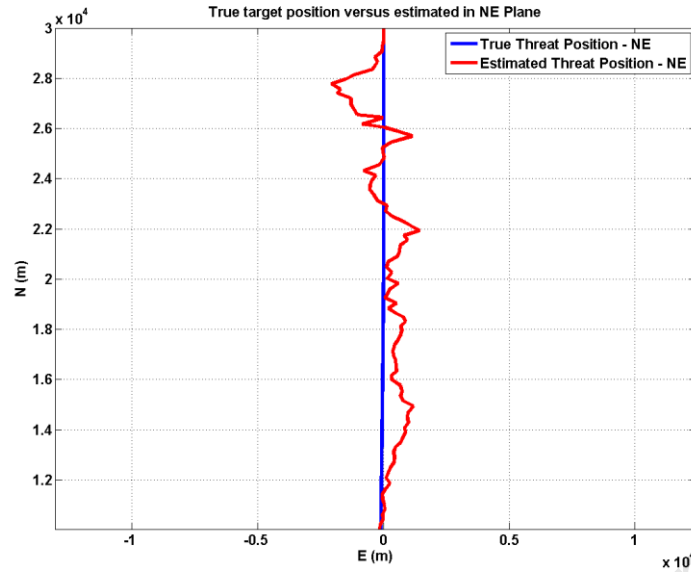


Figure 6.4: Radar EKF with range rate measurements - true and estimated threat position in the NE plane.

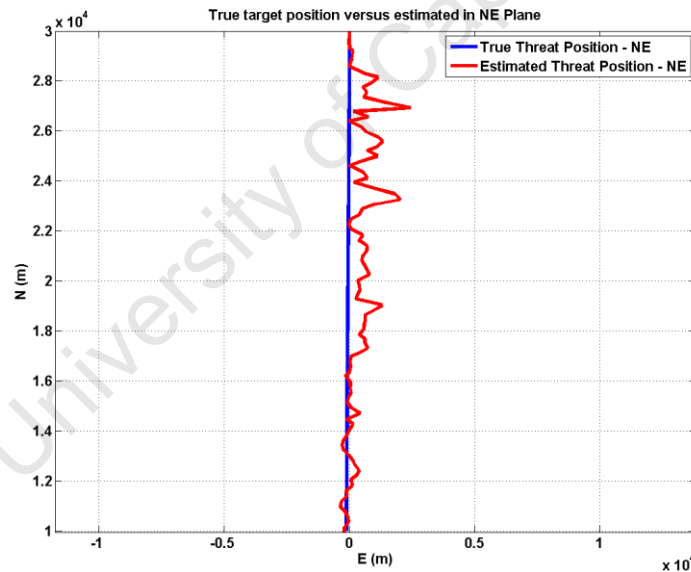


Figure 6.5: Radar EKF with range rate measurements and correlation - true and estimated threat position in the NE plane.

It is interesting to note that range rate estimates are very stable and more accurate for decreasing distance, i.e. as k moves towards 100 (relative range decreases from the start of the simulation). The range rate is illustrated in Figure 5.37.

The RMSE of both filters in Figure 5.38 and Figure 5.45 decrease over time and therefore they are consistent. Furthermore, the RMSE illustrates that their performance is very similar to the nonlinear range radar filter. Chapter 2 stated that the inclusion of range rate measurements greatly increases tracking performance [55]. However, in Chapter 5, it has been proved that under high noise conditions, the inclusion of range rate measurements produces performance similar to the filter where the range rate is excluded. Furthermore, velocity errors are reduced minimally, however the minimisation occurs at a sooner time than filters without velocity measurements. Note that the filter which accounts for correlation is the most computationally intensive filter.

The ANEES in Figure 5.39 and Figure 5.46 characterises that both filters are credible. Therefore, both filter's self-assessments are reliable. However, the filters are underestimating the actual estimation error.

6.1.3 Single Sensor Summary

It was proved in subsection 6.1.1 that the EO camera is not able to estimate the required state vector of a threat. This shortcoming is attributed to the fact that the camera provides only relative angular measurements.

When range rate measurements are not available in a radar tracking filter, tracking can be performed by using a nonlinear filter or a linear one by converting the measurements to Cartesian space. Despite the increased accuracy and shorter settling time exhibited by the nonlinear range radar filter shown above, it is not desirable to make use of a nonlinear filter. It is significantly more computationally intensive than its linear counterpart which was proved to yield similar performance. For our application the credibility of the filter is not of paramount importance, but rather the accuracy of a threat's state estimates. Therefore, the linear range radar filter tracking with Cartesian coordinates only is preferred when no range rate measurements are available.

Two filters were discussed that performed tracking with Doppler information: one that assumed no correlation between measurement errors in range and range rate; and the other filter assumed and accounted for the existence of the correlation by forming a pseudo measurement which decorrelated the errors. These two filters exhibited similar performance, although the latter is more computationally intensive. Furthermore, in comparison to the range radar filters, they provide increased tracking accuracy. Table 6.1 provides the RMSE in position and velocity as well as the ANEES for all the single sensor tracking filters developed.

Table 6.1: Single sensor tracking filter's RMSE in position and velocity and ANEES.

Tracking Filter	Position RMSE (<i>m</i>)	Velocity RMSE (<i>m/s</i>)	Credibility	
			Credible?	ANEES
EO camera	Divergent	37.21	No	N/A
Radar KF (Cartesian only)	576.5	12.8	No	N/A
Radar EKF	134.6	1.5	Yes	1.05
Radar EKF with range rate measurements	136.77	1.5	Yes	1.1
Radar EKF with range rate measurements and correlation	105	1.2	Yes	0.74

In summation, all the radar filters that were developed are able to adequately track the threat. However, the camera filter is not able to. It would be a tragedy if the availability of the accurate camera measurements would go to

waste. The next section discusses how the advantages of each sensor are utilised in the fusion architectures developed.

6.2 Multisensor Tracking Filters

Thesis objective 5, outlined in Chapter 1, required the design and implementation of an airborne tracking filter using MSDF. However, the approach adopted in this thesis was to develop multiple airborne tracking filters using MSDF. The reason for this was to investigate the performance of different approaches; to ultimately arrive at the best solution and consequently the selection for use in the TTU. Thus, three FAs were developed. FA1 is a centralised architecture employing measurement fusion. FA2 is a distributed architecture and was the first attempt at mitigating the camera-only tracking filter's deficiencies, while making good use of its accurate measurements. This was achieved through partial feedback of the fused estimate to the camera local filter. FA3 is built on FA2 by introducing state covariance feedback, using dynamically weighted information distributed coefficients, and full feedback. Furthermore, FA3 was developed in order to reduce to the computational requirements by FA2.

In spite of the fact that the radar EKF with range rate measurements and correlation accounted for, exhibited the best performance of all the single sensor tracking filters; the radar EKF which did not account for the correlation was preferred for use in the fusion architectures. The reason for this is that it requires less computing resources, especially since it forms part of a larger fusion/tracking system.

6.2.1 FA1

FA1 is the first FA developed and exhibits similarities to the single sensor filters as it contains a single tracking filter only. Theoretically, centralised fusion provides the most accurate filters. Indeed, the results from FA1 reflect

very accurate estimation in Figure 5.50 to Figure 5.54. This accurate estimation can be visualised by plotting the threat's motion in the NE plane with its filter-produced estimate (see Figure 6.6).

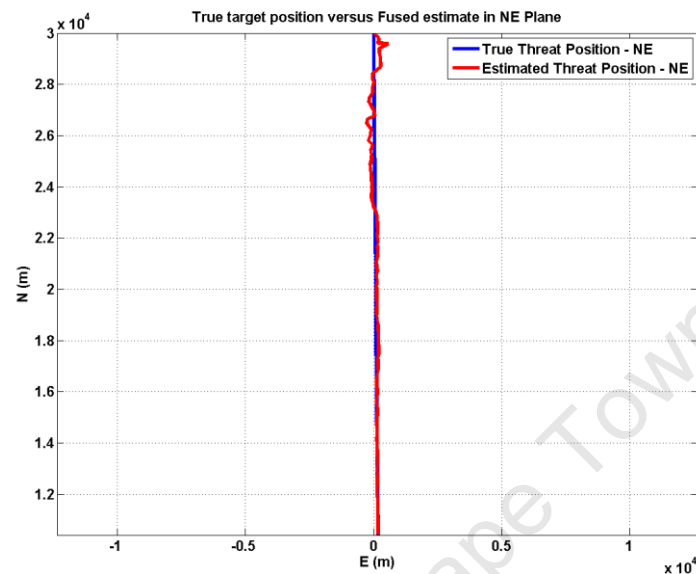


Figure 6.6: FA1 – true and estimated threat position in the NE plane.

The covariance in Figure 5.49 reaches the steady-state around $k = 40$. But unlike the single sensor KF tracking in Cartesian coordinates, this filter self-assessment is very reliable. The conclusion about the filter's reliability of its estimated error covariance is based on the evaluation of its ANEES in Figure 5.55; which shows a value of approximately 1. Therefore FA1 is highly credible. Furthermore the ANEES tells us that when $k < 40$ the filter is optimistic and pessimistic thereafter. This is what is expected as the estimation error starts out high and, when the filter settles, becomes very low.

6.2.2 FA2

FA2 is the first distributed architecture developed. It consists of two EKFs: one for the radar local filter and one for the camera local filter.

The state error covariance in Figure 5.57 provides insight into the working of this distributed architecture. In section 5.3 it was proved that the camera filter diverges; more specifically, the North component. This effect occurs again in FA2. However, very interestingly, towards the end of the simulation, the camera's North error covariance starts converging. This is as a result of the performance of partial feedback; where the local camera state estimate is set to the globally fused estimate, after radar and camera state vector fusion has been performed. This provides a means of mitigating the camera's deficiencies, while the overall tracking system reaps the benefits from the accurate camera angular measurements.

Figure 5.57 also displays the fused and radar local filter's error covariance. It is clear that both reach the steady-state. Through fusion, the fused state error covariance is minimised and stays constant practically throughout the simulation.

Figure 5.60 to Figure 5.62 and Figure 5.63 to Figure 5.65 shows the estimation error of radar, camera and fusion in position and velocity respectively, of a single flight scenario (simulation run). The radar yields the expected error; the same as that of the single sensor radar EKF developed in subsection 5.5.1. The errors in the states produced by the camera, however, are very low. The reason for this is since the performance of partial feedback enables the camera filter to track the threat, the camera's highly accurate angular measurements lead to minimal errors in state estimation. Furthermore, the fused estimation error is even lower. It can be seen that as a result of fusion, the fused estimation error is less than both the radar and camera local filters'.

The RMSE in Figure 5.66 shows that FA2 produces very accurate estimates of the threat's state vector. In fact the RMSE in position settles at 94 m: more accurate than all the single sensor filters. The reason for this is that fusion is performed. However, the camera only provides angular measurements and

should not greatly affect the errors in range. Despite this, we see a decrease in error which is a result of the coupling between range and angles in the spherical – Cartesian equations used by the EKF. Lastly, the ANEES in Figure 5.67 shows that, although not as close to unity as FA1, FA2 is credible.

6.2.3 FA3

FA3 is the result of a few modifications to FA2. The radar local filter in FA3 is a linear KF tracking with Cartesian coordinates only. The use of a nonlinear filter is possible as pseudo measurements were formed by decoupling the range rate measurement from the radar observation vector. At the FC, full feedback is performed, viz. to radar and camera local filters. Furthermore, the fused error covariance matrix is fed back to both local filters using dynamically weighted information distribution coefficients.

In contrast to FA2, the camera local filter does not diverge at any point during the simulation. Figure 5.69 shows the fused as well as radar and camera local filter's estimated error covariance. In particular, this figure highlights the effects of performing dynamic feedback. Dynamic feedback causes the local filter's error covariance matrix to be brought closer to the fused error covariance matrix. At each time-step it is determined which local filter believes it is producing worse estimates. Then that filter's information distribution coefficient is increased. This results in the sawtooth shape seen in the figure and causes both local filters error covariance to converge to the fused error covariance.

Figure 5.72 to Figure 5.74 and Figure 5.75 to Figure 5.77 shows the estimation error of radar, camera and fusion in position and velocity respectively, of a single flight scenario. The result of performing full feedback as well as error covariance feedback is that estimation errors from each local filter (and consequently the FC) are sufficiently close to one another.

The RMSE in Figure 5.78 shows the tracking filter's accuracy, which settles at about 900 m in position. This error is fairly high in comparison to the other fusion architectures developed and is attributed to the conversion involved in tracking with Cartesian coordinates only in the radar local filter, as well as the high process noise of the system. As stated earlier, FA3 was developed in order to reduce the computational load of FA2 and achieves this through the replacement of the nonlinear EKF in FA2 with a linear KF. Therefore the trade-off exists between computational load and tracking filter accuracy.

Figure 5.79 shows the ANEES of FA3. It is clear that its value is far from unity. Therefore the filter is non-credible. In fact it is highly optimistic and the estimated state error covariance matrix is not reliable. Note that despite this the estimates produced by FA3 are still reliable.

Recall that two variants of FA3 were developed to illustrate the importance of feedback and the dynamic weighting of the distribution coefficients. The four variants that were developed are as follows:

- a) FA3a – FA3 with the FC providing no feedback at all to local filters.
- b) FA3b – full feedback (to radar and camera), i.e. state and covariance with **constant** gain coefficients.

Thus, the table below lists the RMSE in position and velocity as well as the ANEES for FA3 and its four variants.

Table 6.2: Performance comparison of FA3 and its two variants.

Fusion Architecture	Position RMSE (<i>m</i>)	Velocity RMSE (<i>m/s</i>)	Credibility	
			Credible?	ANEES
FA3	940	82.5	No	N/A
FA3a	2530	36.8	No	N/A
FA3b	1900	101	No	N/A

FA3a is clearly the least accurate of the FA3 variants. Thus it quantitatively proves that through the performance of feedback more accurate filters result. In FA3b full feedback is performed but with constant information distribution coefficients. Therefore with feedback (FA3b), we see a decrease of 25% in position error from that of no feedback (FA3a); and a further decrease 50% with full dynamic feedback (FA3).

6.3 Comparison of Fusion Architectures

Before the results from the FAs are quantitatively presented, the measurement errors of the EO Camera and Doppler radar are restated from their respective design sections in Chapter 5.

Table 6.3: Measurement errors of Doppler Radar and EO Camera.

Sensor	Measurement Error	Value
Doppler Radar	Range	200 <i>m</i>
	Azimuth	17.19°
	Elevation	17.19°
	Range Rate	50 <i>m/s</i>
EO Camera	Azimuth	1.15°
	Elevation	1.15°

The following table summarises the accuracy and credibility of the fusion architectures developed.

Table 6.4: Performance comparison of fusion architectures. Filter accuracies and credibility. The radar-only filter has been included in the table to easily compare the fusion with the single sensor performance.

Tracking Filter	Position RMSE (<i>m</i>)	Velocity RMSE (<i>m/s</i>)	Credibility	
			Credible?	ANEES
FA1	45	0.9	Yes	1.07
FA2	93.8	1.36	Yes	1.64
FA3	940	82.5	No	N/A
Radar EKF	136.77	1.5	Yes	1.1

It has been proved that all three fusion architectures can accurately track an airborne threat. Moreover, in-line with thesis objective 7, that FA1 and FA2 does so more accurately than all of the single sensor tracking filters.

FA1 is clearly the most accurate architecture from Table 6.4; this was expected as centralised fusion is employed in FA1. However there is only a 50 *m* difference in the distributed FA2's position error. This could possibly be reduced further through performing dynamic feedback. The attempt at reducing the computational load of FA2 resulted in FA3. Although FA3 is able to track the threat, its accuracy is a lot worse than the other FAs. Its position error is ~ 1 *km*; which is too high for our application, but might be adequate for certain non-critical applications where minimising the computational load is of utmost importance.

By inspecting the original sensor measurement errors in Table 6.3, we see that the radar's measurement errors in range and range rate are 200 *m* and

50 m/s respectively. Through fusion both FA1 and FA2 produces position and velocity errors which are considerably more accurate than the sensor's measurement errors.

To quantify the effect on fusion tracking accuracy when one of the sensors measurement errors are more accurate, Table 6.5 lists, for each architecture, the position and velocity RMSE when the radar's angular measurement accuracies are the same as the camera's and when the camera's is more accurate by a factor of about 10.

Table 6.5: Position and Velocity RMSE when both sensor's angular measurement errors are the same and when the camera's are more accurate.

Architecture	Angle Measurement Erros			
	$\sigma_{\theta}^R, \sigma_{\varphi}^R, \sigma_{\theta}^C, \sigma_{\varphi}^C = 0.3rad$		$\sigma_{\theta}^R, \sigma_{\varphi}^R = 0.3rad,$ $\sigma_{\theta}^C, \sigma_{\varphi}^C = 0.02rad$	
	$RMSE^{pos}$ (m)	$RMSE^{vel}$ (m/sec)	$RMSE^{pos}$ (m)	$RMSE^{vel}$ (m/sec)
Radar-only EKF	136.77	1.5	136.77	1.5
FA1	95.5	1.1	45	0.9
FA2	135.2	1.8	93.8	1.36
FA3	970	87.5	940	82.5

Table 6.5 makes it clear that the fusion of accurate angular measurements from the camera improves the overall tracking performance.

Considering the above information, albeit FA1 is the most accurate of the architectures, FA2 is the best compromise between filter accuracy,

computational load and robustness. Therefore it has been selected as the best candidate for use in the TTU.

6.4 Summary

This chapter provided a discussion of the results derived from the simulation of single sensor tracking filters and the fusion architectures developed. The comparison between the three fusion architectures showed that FA1 is the most accurate filter for an airborne tracking system. Furthermore, that the difference in accuracy exhibited by FA2 in comparison to FA1 is negligible.

It is a known fact that centralised architectures yield the best performance, yet distributed architectures are not without merits: robustness - the FC is not a single point of failure in the system as in the centralised architecture; and they are less computationally intensive.

Taking this into account, it was concluded that FA2 is the superior architecture for an airborne threat tracking system and hence the TTU.

In summation, it has been proved that:

- All three FAs can reliably track a threat.
- Centralised FA1 is the best performing tracking filter.
- The Distributed FA2's accuracy is also very good.
- Both FA1 and FA2 are more accurate than any single sensor tracking filter.
- FA3 is not nearly as accurate as FA1 and FA2, but it does have a lower computational burden of the lot.

Chapter 7: Conclusions

The overall aim of this research was to design an airborne tracking filter using MSDF. This chapter will summarise the findings from the work performed in this thesis to achieve this end and offer conclusions based on the findings. Additionally, recommendations for future research will be discussed, with regards to the progression of this study.

7.1 Research Objectives: Conclusions

This thesis focussed on the design of an airborne tracking filter using MSDF. MSDF fuses data from multiple sensors to provide more meaningful information than from a single sensor. The advantages and disadvantages of radar and camera sensors were analysed, which made it clear that by combining the two a superior sensor would result. A progressive design structure was adopted, where first single sensor tracking filters were developed, followed by multisensor tracking filters. Statistical analysis of the filters was performed based on numerical simulations.

Three fusion architectures were developed in order to perform target tracking while making use of the measurements from multiple, redundant sensors. The architectures were termed FA1, FA2 and FA3.

FA1 was a centralised architecture performing measurement fusion, whereas FA2 and FA3 were distributed architectures employing state vector fusion. It was proved that all fusion architectures could reliably track an airborne threat under high measurement and process noise.

As this thesis is concerned with airborne tracking, the fusion architectures tracked relative motion between UAV and threat. It was decided that tracking was to be performed in the inertial NED system. This choice stemmed from the fact that relative motion in the BRF depends on attitude dynamics,

whereas the motion's projection in NED does not. Thus tracking in NED is considerably less complex than that in the BRF.

It was proved that FA1 with measurement fusion was the most accurate and the most credible tracking filter. The design of FA2 was the first attempt at mitigating the camera tracking filter's deficiency – its inability to track a threat using high accuracy angular measurements. This was successfully achieved through partial feedback and proved that FA2 was very accurate in tracking the threat in both position and velocity.

FA3 built on FA2 by introducing state covariance feedback, using dynamically weighted information distributed coefficients. Another extension was full feedback – the global fused estimate was fed back to both radar and local filters. Furthermore, FA3 was developed in order to reduce to the computational requirements by FA2. This was achieved through the decoupling of the range rate measurements from the radar observation vector, which resulted in the formation of radar and camera pseudo measurements.

The new radar pseudo measurement allowed the use of a linear KF, thereby reducing system complexity. However, the cost of this reduced system complexity is the trade-off with filter accuracy. Therefore, although requiring the least computational resources, FA3 was the worst performing architecture.

Single sensor tracking filters were developed to quantify the performance gain resulting from MSDF. These included:

- EO camera filter. An EKF was used as the angular measurements are nonlinear with respect to the state variables.
- A nonlinear and a linear radar filter. The nonlinear filter was the EKF; and the linear a KF tracking with Cartesian coordinates only. The sensor measurements were converted to Cartesian space and then debiased, which allowed the use of a linear filter.

- Two Doppler radar filter with range rate measurements. The one was designed under the assumption that no correlation exists between measurement errors in range and range rate; and the other based on the assumption that the correlation exists.

Furthermore, a simulation environment was created in which the developed filters were simulated. This simulation environment generated the UAV and threat aircraft's motion in the inertial NED system as well the Doppler radar and EO camera sensor measurements with respect to the BRF. The pre-processing step to filtering was the transformation of the sensor measurements to the NED system. This was achieved using direction cosines.

The application of MSDF demonstrated large improvements in tracking accuracy. Despite FA1 being the most accurate architecture, FA2 was assessed to be a good compromise between filter accuracy and credibility, computational load and robustness. Therefore it has been selected as the tracking filter to make up the TTU; and in-turn be integrated with the overall UAV SAS.

7.2 Recommendations for Future Work

Although the fusion architectures developed in this work accurately track a threat in inertial space, in practical applications a number of challenges exist that are not addressed in this thesis. For example, in civil airspace the UAV is likely to be surrounded by multiple aircraft. Thus, a natural extension to the filters developed would be to account for multiple targets. This would entail adding data association techniques into the filter to pair observations of target's to existing tracks. Once an observation has been discriminated against a target, tracking of the target's state vector can occur.

Another challenge not addressed in this thesis is that of manoeuvring targets. Therefore the extension to this work to add the ability to track manoeuvring threats is recommended. The widely used method of tracking manoeuvring targets makes use of the IMM filter; where multiple KF models are used for each manoeuvre stage. Then the estimates from each KF are combined with the use of a Markov model. Furthermore, to simulate the threats more accurately a Singer acceleration model could be used.

Improved sensor models are also recommended; where the sensor's FOV is taken into account. That is measurements are only returned if a target is in the sensor's FOV. Furthermore, noise should be added to the measurement model to make it more realistic in that measurements do not always arrive at the tracking filter.

Other than the sensor's FOV, the detection ranges of the sensors should also be incorporated into the fusion algorithm. If an EO camera can generally detect a target within 3 km, then data fusion should only occur when a target is within 3 km of the UAV. This can be further extended by realising the fact that the radar is not able to reliably detect a target at very close ranges. Therefore, the fusion algorithm will decide the appropriate filter, depending on the distance to the target.

This thesis deals with target tracking using simulated data. In order to validate the performance of the fusion filters with practical measurements, it is recommended that a real-world test be conducted. However, practical data for this application would be particularly hard to come by; for example, performing flight tests. An intermediate step towards that end could be a laboratory setup where a 3D graphics flight simulation is generated on one computer. The simulation should be generated from the UAV's first-person perspective. A camera-in-the-loop system should then be set up on another computer, which performs image processing algorithms and feeds the TTU with real sensor data.

Finally, the TTU is to be integrated alongside control algorithms performing CA functions; where the TTU provides the control system with threat position and velocity estimates. The amalgamation of these two subsystems sees the realisation of a UAV SAS.

University of Cape Town

Appendix A1: EO Camera Linearized Measurement Matrix

$$H_X^C(k) = \begin{bmatrix} H_{11} & 0 & H_{13} & 0 & 0 & 0 \\ H_{21} & 0 & H_{23} & 0 & H_{25} & 0 \end{bmatrix} \quad (\text{A.1})$$

Where:

H_{11}	$\frac{-E}{N^2 + E^2}$
H_{13}	$\frac{N}{N^2 + E^2}$
H_{21}	$\frac{N \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{23}	$\frac{E \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{25}	$\frac{-(N^2 + E^2)}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$

Appendix A2: Radar Linearized Measurement Matrix

$$H_X^R(k) = \begin{bmatrix} H_{11} & 0 & H_{13} & 0 & H_{15} & 0 \\ H_{21} & 0 & H_{23} & 0 & 0 & 0 \\ H_{31} & 0 & H_{33} & 0 & H_{35} & 0 \end{bmatrix} \quad (\text{A.2})$$

Where:

H_{11}	$\frac{N}{\sqrt{N^2 + E^2 + D^2}}$
H_{13}	$\frac{E}{\sqrt{N^2 + E^2 + D^2}}$
H_{15}	$\frac{D}{\sqrt{N^2 + E^2 + D^2}}$
H_{21}	$\frac{-E}{N^2 + E^2}$
H_{23}	$\frac{N}{N^2 + E^2}$
H_{31}	$\frac{N \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right) (N^2 + E^2 + D^2)^{3/2}} \right)}$
H_{33}	$\frac{E \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right) (N^2 + E^2 + D^2)^{3/2}} \right)}$
H_{35}	$\frac{-(N^2 + E^2)}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right) (N^2 + E^2 + D^2)^{3/2}} \right)}$

Appendix A3: Radar Linearized Measurement Matrix With Range Rate Measurements

$$H_X^R(k) = \begin{bmatrix} H_{11} & 0 & H_{13} & 0 & H_{15} & 0 \\ H_{21} & 0 & H_{23} & 0 & 0 & 0 \\ H_{31} & 0 & H_{33} & 0 & H_{35} & 0 \\ H_{41} & H_{42} & H_{43} & H_{44} & H_{45} & H_{46} \end{bmatrix} \quad (\text{A.3})$$

Where:

H_{11}	$\frac{N}{\sqrt{N^2 + E^2 + D^2}}$
H_{13}	$\frac{E}{\sqrt{N^2 + E^2 + D^2}}$
H_{15}	$\frac{D}{\sqrt{N^2 + E^2 + D^2}}$
H_{21}	$\frac{-E}{N^2 + E^2}$
H_{23}	$\frac{N}{N^2 + E^2}$
H_{31}	$\frac{N \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{33}	$\frac{E \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{35}	$\frac{-(N^2 + E^2)}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{41}	$\frac{\dot{N}E^2 - N\dot{E}E + \dot{N}D^2 - N\dot{D}D}{(N^2 + E^2 + D^2)^{3/2}}$
H_{42}	$\frac{N}{\sqrt{N^2 + E^2 + D^2}}$

H_{43}	$\frac{\dot{E}N^2 - \dot{N}EN + \dot{E}D^2 - E\dot{D}D}{(N^2 + E^2 + D^2)^{3/2}}$
H_{44}	$\frac{E}{\sqrt{N^2 + E^2 + D^2}}$
H_{45}	$\frac{\dot{D}N^2 - \dot{N}DN + \dot{D}E^2 - E\dot{E}D}{(N^2 + E^2 + D^2)^{3/2}}$
H_{46}	$\frac{D}{\sqrt{N^2 + E^2 + D^2}}$

University of Cape Town

Appendix A4: Camera Pseudo Linearized Measurement Matrix

$$H_X^C(k) = \begin{bmatrix} H_{11} & 0 & H_{13} & 0 & 0 & 0 \\ H_{21} & 0 & H_{23} & 0 & H_{25} & 0 \\ H_{31} & H_{32} & H_{33} & H_{34} & H_{35} & H_{36} \end{bmatrix} \quad (\text{A.4})$$

Where:

H_{11}	$\frac{-E}{N^2 + E^2}$
H_{13}	$\frac{N}{N^2 + E^2}$
H_{21}	$\frac{N \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{23}	$\frac{E \times D}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{25}	$\frac{-(N^2 + E^2)}{\left(\sqrt{\left(\frac{1 - D^2}{N^2 + E^2 + D^2} \right)} (N^2 + E^2 + D^2)^{3/2} \right)}$
H_{31}	$\frac{\dot{N}E^2 - N\dot{E}E + \dot{N}D^2 - N\dot{D}D}{(N^2 + E^2 + D^2)^{3/2}}$
H_{32}	$\frac{N}{\sqrt{N^2 + E^2 + D^2}}$
H_{33}	$\frac{\dot{E}N^2 - \dot{N}EN + \dot{E}D^2 - E\dot{D}D}{(N^2 + E^2 + D^2)^{3/2}}$
H_{34}	$\frac{E}{\sqrt{N^2 + E^2 + D^2}}$
H_{35}	$\frac{\dot{D}N^2 - \dot{N}DN + \dot{D}E^2 - E\dot{E}D}{(N^2 + E^2 + D^2)^{3/2}}$

H_{36}	$\frac{D}{\sqrt{N^2 + E^2 + D^2}}$
----------	------------------------------------

University of Cape Town

References

- [1] Department of Defence - United States of America, "FY2009-2034 Unmanned Systems Integrated Roadmap," 2009.
- [2] Federal Aviation Administration, "Sense and Avoid (SAA) for Unmanned Aircraft Systems (UAS)," SAA Workshop, Final Report 2009.
- [3] Net Resources International. Army Technology. [Online]. <http://www.army-technology.com/projects/rq1-predator/rq1-predator3.html>
- [4] D.L. Hall and S.A.H. McMullen, *Mathematical techniques in multisensor data fusion.*: Artech House Publishers, 1992.
- [5] T. Hutchings, S. Jeffryes, and SJ Farmer, "Architecting UAV sense & avoid systems," in *IET*, 2007, pp. 1-8.
- [6] Federal Aviation Administration. (2010, August) FAA. [Online]. <http://adsb.tc.faa.gov/TCAS.htm>
- [7] ADS-B Technologies. (2011) ADS-B Technologies. [Online]. <http://www.ads-b.com/>
- [8] The MITRE Corporation, "Airspace Integration Alternatives for Unmanned Aircraft," in *AUVSI Unmanned Systems Asia-Pacific*, Singapore, 2010.
- [9] NATO Naval Armaments Group, "Sense Aand Avoid Requirements for Unmanned Aerial Vehicle Systems," Joint Capability Group On Unmanned Aerial Vehicles, 2008.
- [10] A.B. Hamza, Y. He, H. Krim, and A. Willsky, "A multiscale approach to pixel-level image fusion," *Integrated Computer Aided Engineering*, vol. 12, no. 2, pp. 135-146, 2005.
- [11] C. Kopp, "Sensors, Weapons and Systems for a 21st Century Counter-Air Campaign," , 1999, pp. 15-16.
- [12] S. Reece, S. Roberts, C. Claxton, and D. Nicholson, "Multi-sensor fault recovery in the presence of known and unknown fault types," in *IEEE*,

2009, pp. 1695-1703.

- [13] T.G. Lee, "Centralized Kalman filter with adaptive measurement fusion: its application to a GPS/SDINS integration system with an additional sensor," *International Journal Of Control Automation And Systems*, vol. 1, pp. 444-452, 2003.
- [14] E. Waltz, J. Llinas, and others, *Multisensor data fusion.:* Artech House Norwood, MA, 1990, vol. 685.
- [15] WJ Harlin and DA Cicci, "Ballistic missile trajectory prediction using a state transition matrix," *Applied mathematics and computation*, vol. 188, no. 2, pp. 1832-1847, 2007.
- [16] GR Zabriskie, CA Cummings, and EA Rogers, "Multisensor fusion using a new procedure for ballistic path estimation," in *IEEE*, 1994, pp. 521-528.
- [17] A. Gad and M. Farooq, "Data fusion architecture for maritime surveillance," in *IEEE*, vol. 1, 2002, pp. 448-455.
- [18] Y. Fischer and A. Bauer, "Object-oriented sensor data fusion for wide maritime surveillance," in *IEEE*, 2010, pp. 1-6.
- [19] O. Hellwich, "An alternative paradigm for data evaluation in remote sensing using multisensor data fusion," in *IEEE*, vol. 1, 1999, pp. 299-301.
- [20] ASP Bhogall et al., "Extraction of forest attribute information using multisensor data fusion techniques: A case study for a test site on Vancouver island, British Columbia," in *IEEE*, vol. 2, 2001, pp. 674-680.
- [21] E. Shahbazian, L. Gagnon, J.R. Duquet, M. Macieszczak, and P. Valin, "Fusion of imaging and non-imaging data for surveillance aircraft," *Sensor Fusion: Architecture, Algorithms, and Applications, SPIE Aerosense*, vol. 97, pp. 179-189, 1997, In [], a system is proposed that makes of FLIR,.
- [22] S.G. Narasimhan and S.K. Nayar, "Vision and the atmosphere," in *ACM*, 2008, p. 69.

- [23] The Boeing Company, "Sensor Data Fusion Technologies for Precision Strike," in *RTO SCI Symposium on Sensor Data Fusion and Integration of the Human Element*, Ottawa, Canada, 1998.
- [24] S.S. Blackman, "Multiple-target tracking with radar applications," *Dedham, MA, Artech House, Inc., 1986, 463 p.*, vol. 1, 1986.
- [25] S.S. Blackman and R. Popoli, *Design and analysis of modern tracking systems.*: Artech House Norwood, MA, 1999, vol. 685.
- [26] D.L. Hall, "Dirty secrets in multisensor data fusion," DTIC Document, 2001.
- [27] JA Roecker and CD McGillem, "Comparison of two-sensor tracking methods based on state vector fusion and measurement fusion," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 24, no. 4, pp. 447-449, 1988.
- [28] Q. Gan and C.J. Harris, "Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 37, no. 1, pp. 273-279, 2001.
- [29] W. Qingchao and W. Wenfei, "Tracking method based on separation and combination of the measurements for radar and IR fusion system," *Journal of Systems Engineering and Electronics*, vol. 20, no. 2, pp. 241-246, 2009.
- [30] X. Tian and Y. Bar-Shalom, "Sequential track-to-track fusion algorithm: exact solution and approximate implementation," , vol. 6969, 2008, p. 696910.
- [31] H.Z. Qiu, H.Y. Zhang, and H. Jin, "Fusion algorithm of correlated local estimates," *Aerospace science and technology*, vol. 8, no. 7, pp. 619-626, 2004.
- [32] Australian Centre for Field Robotics, University of Sydney, Multi Sensor Data Fusion, 2001, Course Notes.

- [33] M.E. Liggins et al., "Distributed fusion architectures and algorithms for target tracking," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 95-107, 1997.
- [34] L.Y. Pao, "Distributed multisensor fusion," *Am. Inst. of Aeronautics and Astronautics*, 1994.
- [35] Carol Anway-Wiese, Adrienne Miller, and Stephen Beardslee, "The Effects of Centralized Vs. Distributed Data Fusion on Mission Outcome in a Cruise Missile Defense Scenario," in *6th International Command and Control Research and Technology Symposium*, 2001.
- [36] E. Brookner and J. Wiley, *Tracking and Kalman filtering made easy*.: Wiley Online Library, 1998.
- [37] D. Simon, *Optimal state estimation: Kalman, H [infinity] and nonlinear approaches*.: John Wiley and Sons, 2006.
- [38] Y. Bar-Shalom, "Multitarget-multisensor tracking: advanced applications," *Norwood, MA, Artech House*, 1990, 391 p., vol. 1, 1990.
- [39] Y. Bar-Shalom and X.R. Li, "Estimation and tracking- Principles, techniques, and software," *Norwood, MA: Artech House, Inc*, 1993., 1993.
- [40] S.J. Julier and J.K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Spie Bellingham, WA*, vol. 3, 1997, p. 26.
- [41] A. Gupta, "Nonlinear Geometric and Differential Geometric Guidance of UAVs with Vision Sensing for Reactive Obstacle Avoidance," DTIC Document, 2010.
- [42] P.R. Mahapatra and K. Mehrotra, "Mixed coordinate tracking of generalized maneuvering targets using acceleration and jerk models," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 36, no. 3, pp. 992-1000, 2000.
- [43] W.F. Denham and S. Pines, "Sequential estimation when measurement function nonlinearity is comparable to measurement error(Sequential linear estimation when measurement function is nonlinear, not

considering system dynamics, applied to space trajectories)," *AIAA; AIAA JOURNAL*, vol. 4, pp. 1071-1076, 1966.

- [44] Z. Duan, X.R. Li, C. Han, and H. Zhu, "Sequential unscented Kalman filter for radar target tracking with range rate measurements," in *IEEE*, vol. 1, 2005, pp. 8--pp.
- [45] M. Lei and C. Han, "Sequential nonlinear tracking using UKF and raw range-rate measurements," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 43, no. 1, pp. 239-250, 2007.
- [46] S. Du, Z. Shi, W. Zang, and K. Chen, "Using interacting multiple model particle filter to track airborne targets hidden in blind Doppler," *Journal of Zhejiang University-Science A*, vol. 8, no. 8, pp. 1277-1282, 2007.
- [47] X.R. Li and V.P. Jilkov, "A survey of maneuvering target tracking—Part III: Measurement models," , vol. 4473, 2001, pp. 423-446.
- [48] G. Fasano, "Multisensor based Fully Autonomous Non-Cooperative Collision Avoidance System for UAVs," University of Naples, Faculty of Engineering, 2008.
- [49] W. Premerlani and P. Bizard, "Direction cosine matrix IMU: theory," *DIY DRONE: USA*, pp. 13-15, 2009.
- [50] JB Pearson and E.B. Stear, "Kalman filter applications in airborne radar tracking," *Aerospace and Electronic Systems, IEEE Transactions on*, no. 3, pp. 319-329, 1974.
- [51] W. Kuhn, "Semantic reference systems," *International Journal of Geographical Information Science*, vol. 17, no. 5, pp. 405-409, 2003.
- [52] GV Trunk, "Automatic detection, tracking, and sensor integration," DTIC Document, 1988.
- [53] S. Zollo and B. Ristic, "On the choice of the coordinate system and tracking filter for the track-while-scan mode of an airborne pulse Doppler radar," 1999.

- [54] X. Rong Li and V.P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 39, no. 4, pp. 1333-1364, 2003.
- [55] Z. Duan, C. Han, and X.R. Li, "Sequential nonlinear tracking filter with range-rate measurements in spherical coordinates," in *Citeseer*, vol. 1, pp. 599-605.
- [56] Y.J. Zhang and G.Z. Xu, "Bearings-only target motion analysis via instrumental variable estimation," *Signal Processing, IEEE Transactions on*, vol. 58, no. 11, pp. 5523-5533, 2010.
- [57] K. Dogancay and others, "Reducing the bias of a bearings-only TLS target location estimator through geometry translations," , 2004, pp. 1123-1126.
- [58] K. Radhakrishnan, A. Unnikrishnan, and KG Balakrishnan, "Bearing only Tracking of Maneovuring Targets Using a Single Coordinated Turn Model," *International Journal of Computer Applications IJCA*, vol. 1, no. 1, pp. 29-36, 2010.
- [59] S. Sadhu, M. Srinivasan, S. Mondal, and TK Ghoshal, "Bearing only tracking using square root sigma point Kalman filter," in *IEEE*, 2004, pp. 66-69.
- [60] M.S. Arulampalam, B. Ristic, N. Gordon, and T. Mansell, "Bearings-only tracking of manoeuvring targets using particle filters," *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 2351-2365, 2004.
- [61] S.C. Nardone and V.J. Aidala, "Observability criteria for bearings-only target motion analysis," *Aerospace and Electronic Systems, IEEE Transactions on*, no. 2, pp. 162-166, 1981.
- [62] Brett Bethke, "Persistent Vision-Based Search and Track Using Multiple UAVs," Massachusetts Institute of Technology, MSc Thesis 2005.
- [63] Y. Watanabe, "Stochastically optimized monocular vision-based navigation and guidance," Georgia Institute of Technology, 2008.

- [64] D.F. Bizup and D.E. Brown, "The over-extended kalman filter-don't use it!," , vol. 1, 2003, pp. 40-46.
- [65] X. Yuan, C. Han, and Z. Duan, "Performance analysis and correlation selection with Doppler measurements," in *IEEE*, pp. 373-379.
- [66] Y. Bar-Shalom, "Negative correlation and optimal tracking with Doppler measurements," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 37, no. 3, pp. 1117-1120, 2001.
- [67] P. Suchomski, "Explicit expressions for debiased statistics of 3D converted measurements," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 35, no. 1, pp. 368-370, 1999.
- [68] JG Wang, T. Long, and PK He, "A New Method of Incorporating Radial Velocity Measurement into Kalman Filter," *Proceedings of signal Processing*, no. 18, pp. 414-416, 2002.
- [69] N. Bellotto and H. Hu, "Multisensor-based human detection and tracking for mobile service robots," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 39, no. 1, pp. 167-181, 2009.
- [70] E. Richter, R. Schubert, and G. Wanielik, "Radar and vision based data fusion-advanced filtering techniques for a multi object vehicle tracking system," in *IEEE*, 2008, pp. 120-125.
- [71] A. Farina and F.A. Studer, "Radar data processing," 1985.
- [72] Y. Xu and X. Liang, "Information fusion for radar/infrared compound seeker based on federated filter," *International Journal of Digital Content Technology and its Applications*, vol. 5, no. 4, 2011.
- [73] C. Blanc, L. Trassoudaine, and J. Gallice, "EKF and particle filter track-to-track fusion: a quantitative comparison from radar/lidar obstacle tracks," in *IEEE*, vol. 2, 2005, pp. 7--pp.
- [74] C. Blanc, L. Trassoudaine, Y. Le Guilloux, and R. Moreira, "Track to track fusion method applied to road obstacle detection," in *Citeseer*, 2004.

- [75] N. Chumerin and M.M. Van Hulle, "Cue and Sensor Fusion for Independent Moving Objects Detection and Description in Driving Scenes," *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, vol. 3, pp. 161-180, 2008.
- [76] Y. Zhu, Z. You, J. Zhao, K. Zhang, and X.R. Li, "The optimality for the distributed Kalman filtering fusion with feedback," *Automatica*, vol. 37, no. 9, pp. 1489-1493, 2001.
- [77] I. Leibowicz, P. Nicolas, and L. Ratton, "Radar/ESM tracking of constant velocity target: comparison of batch (MLE) and EKF performance," in *IEEE*, vol. 1, 2000, pp. TUC2--3.
- [78] Arkadiy Turevskiy, Stacey Gage, and Craig Buhr, "Model-Based Design of a New Light-weight Aircraft," 2007.
- [79] Amir Patel, "UAV Collision Avoidance: A Specific Acceleration Matching Approach," University of Cape Town, Cape Town, MSc Thesis 2011.
- [80] Z. Zhao and X.R. Li, "Probabilistic model distortion measure and its application to model-set design of multiple model approach," in *IEEE*, vol. 2, 2004, pp. 2146-2150.
- [81] Ian Reid, "Estimation II ," University of Oxford, Lecture Notes 2001.
- [82] Iain Peddle, "Aircraft Dynamics," University of Stellenbosch, Course Notes 2009.
- [83] Y. Gao, W.J. Jia, X.J. Sun, and Z.L. Deng, "Self-tuning multisensor weighted measurement fusion Kalman filter," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 45, no. 1, pp. 179-191, 2009.
- [84] W. Khawsuk and L.Y. Pao, "Decorrelated state estimation for distributed tracking of interacting targets in cluttered environments," in *IEEE*, vol. 2, 2002, pp. 899-904.
- [85] D.L. Hall and S.A.H. McMullen, *Mathematical techniques in multisensor data fusion.*: Artech House Publishers, 2004.

- [86] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision.*: Cambridge Univ Press, 2000, vol. 2.
- [87] GSW Klein and TW Drummond, "Tightly integrated sensor fusion for robust visual tracking," *Image and Vision Computing*, vol. 22, no. 10, pp. 769-776, 2004.
- [88] S. You and U. Neumann, "Fusion of vision and gyro tracking for robust augmented reality registration," in *IEEE*, 2001, pp. 71-78.
- [89] T. Auer and A. Pinz, "The integration of optical and magnetic tracking for multi-user augmented reality," *Computers & Graphics*, vol. 23, no. 6, pp. 805-808, 1999.

University of Cape Town