sipri

THE IMPACT OF ARTIFICIAL INTELLIGENCE ON STRATEGIC STABILITY AND NUCLEAR RISK

Volume III

South Asian Perspectives

EDITED BY PETR TOPYCHKANOV

April 2020

STOCKHOLM INTERNATIONAL PEACE RESEARCH INSTITUTE

SIPRI is an independent international institute dedicated to research into conflict, armaments, arms control and disarmament. Established in 1966, SIPRI provides data, analysis and recommendations, based on open sources, to policymakers, researchers, media and the interested public.

The Governing Board is not responsible for the views expressed in the publications of the Institute.

STOCKHOLM INTERNATIONAL PEACE RESEARCH INSTITUTE
Signalistgatan 9
SE-169 72 Solna, Sweden
Telephone: + 46 8 655 9700
Email: sipri@sipri.org
Internet: www.sipri.org

The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk

Volume III
South Asian Perspectives

EDITED BY PETR TOPYCHKANOV

sipri

April 2020

Contents

Preface

The post-cold war global strategic landscape is currently in an extended process of being redrawn. Many different trends are in play here. Importantly, the underlying dynamics of world order are shifting with China's economic, political and military rise, Russia's reassertion of a great power role, and the disenchantment of the current administration of the United States towards the international institutions this country had a big hand in creating.

As a result of these trends, a binary Russian–US nuclear juxtaposition, a legacy of the old Soviet–US confrontation, is being gradually augmented by regional nuclear competitions. As the arms control framework that the Soviet Union and the USA created at the end of the cold war disintegrates, the commitment of the two states with the largest nuclear arsenals to pursue strategic stability through arms control and disarmament is in doubt to an unprecedented degree.

On top of this comes the impact of new technological developments. The world is undergoing a 'fourth industrial' revolution, characterized by rapid and converging advances in multiple technologies including artificial intelligence (AI), robotics, quantum technology, nanotechnology, biotechnology and digital fabrication. The question of how these technologies will be used has not yet been fully answered. It is beyond dispute, however, that nuclear-armed states seek to exploit these technologies for their national security.

The potential impact of these developments on strategic stability and nuclear risk has not yet been systematically documented and analysed. The SIPRI project 'Mapping the impact of machine learning and autonomy on strategic stability' is a first attempt to present a nuanced analysis of what impact the exploitation of AI could have on the global and regional strategic landscapes. This edited volume on South Asian perspectives is the third major publication of this two-year research project. The authors are experts from India, the Netherlands, Pakistan, Russia and Sri Lanka. This volume was preceded by volumes on Euro-Atlantic and East Asian perspectives and will be followed by a final report.

SIPRI commends this study to decision makers in the realms of arms control, defence and foreign affairs, to researchers and students in departments of politics, international relations and computer science, and to members of the general public who have a professional and personal interest in the subject.

Dan Smith
Director, SIPRI
Stockholm, April 2020

Acknowledgements

I express my sincere gratitude to the Carnegie Corporation of New York for its generous financial support of this project, and to the Pathfinder Foundation for co-hosting with SIPRI the third project workshop on 25–26 February 2019 in Colombo.

I am also indebted to all the scholars who participated in this workshop and agreed to contribute to this volume. The essays that follow are based on the presentations delivered at the workshop. The views expressed in the essays are those of the authors and should not be taken to reflect the views of SIPRI, the Carnegie Corporation or the Pathfinder Foundation.

I am also grateful to my SIPRI colleagues Dr Vincent Boulanin, Dr Lora Saalman and Pieter D. Wezeman for their comprehensive and constructive comments on the volume draft, and the SIPRI Editorial Department for its invaluable work.

Petr Topychkanov

Abbreviations

AI	Artificial intelligence
ASAT	Anti-satellite (weapon)
ATR	Automatic target recognition
BMD	Ballistic missile defence
CAIR	Centre for Artificial Intelligence and Robotics (of India)
CCW	(Convention on) Certain Conventional Weapons
CBM	Confidence-building measure
DRDO	Defence Research and Development Organisation (of India)
FCAS	Future Combat Air System
ICBM	Intercontinental ballistic missile
IHE	Insensitive high explosive
ISR	Intelligence, surveillance and reconnaissance
LAWS	Lethal autonomous weapon systems
MEMS	Microelectromechanical systems
nEM	Nanoenergetic material
NC3	Nuclear command, control and communications
N/MEMS	Nano- and microelectromechanical systems
NRRC	Nuclear Risk Reduction Center (of the United States)
R&D	Research and development
SPD	Strategic Plans Division (of Pakistan)
SRF	Strategic Rocket Forces (of Russia)
SSBN	Nuclear-powered ballistic missile submarine
UAV	Unmanned aerial vehicle
UN	United Nations
UUV	Unmanned underwater vehicle
WMD	Weapons of mass destruction

Executive Summary

The ongoing renaissance of artificial intelligence (AI) is reshaping the world. Just like many other developing countries, India and Pakistan—the two nuclear-armed states of South Asia—are exploring the subsequent opportunities for economic and social change. Their political leaders seem to prioritize civilian applications of AI over the military, and public attention reflects the political priorities. National efforts to militarize AI do not receive the same public coverage as civilian AI developments.

Meanwhile, according to the available open-source information, India and Pakistan are increasingly interested in the potential benefits of AI for defence and security. This might be one of the reasons why an expert debate on the opportunities and risks posed by the AI renaissance in the military realm has started in recent years. However, the debate suffers from large gaps, particularly in the emerging discussion on the potential impact of AI on strategic stability and nuclear risk in South Asia. This issue has been underexplored by scholars studying South Asia from both inside and outside the region.

This edited volume—which follows earlier volumes on Euro-Atlantic and East Asian perspectives—tries to fill the gaps in the scholarly debate on this important topic and to facilitate further regional debate. It is based on a workshop held in Colombo in February 2019. The eight expert contributors—from South Asia and around the world—reflect the variety of issues, approaches and views.

It is clear from a comparative study of the state of adoption of AI in South Asia that India and Pakistan are playing catch-up in the world competition on military AI. Compared to the United States, China and Russia, India's advances are modest, while Pakistan's are even less visible. One of the reasons seems to be under-resourcing and inefficiencies in defence research and state industries. These prohibit the development and adoption of emerging technologies within a reasonable time frame.

However, according to contributors from India and Pakistan, both countries are well aware of the strategic significance of AI. They see AI as one of many enablers of the mutual strategic balance. India must also take into consideration the role of AI in the military build-up of China, one of its long-term security concerns.

In assessing the strategic significance of AI, the expert contributors—regardless of their origin—agree that AI is a double-edged sword. On the one hand, AI could enhance nuclear command and control, early warning, intelligence surveillance and reconnaissance (ISR), and the physical security of nuclear capabilities, among other areas. In this way it would improve states' sense of security. On the other hand, the same advances could cast doubt on the survivability of their respective second-strike capabilities. This doubt would stimulate more aggressive nuclear postures that could increase nuclear risk.

There are several scenarios in which AI-enabled weapons could be involved in escalatory dynamics in South Asia. Given that there have been few military applications of AI in either India or Pakistan, the contributors do not endorse the view

that the use of AI systems could cause a nuclear war between India and Pakistan or between India and China—at least for the foreseeable future. However, most agree that the introduction of AI into the nuclear capabilities and postures of India and Pakistan could affect strategic stability in South Asia. For this reason, the majority of contributors support the idea that the states of South Asia should take steps now to reduce the nuclear risk.

The question of how to design those steps is more divisive. For some, the solution lies in the development of a legally binding international agreement that would limit the military use of AI. Others argue that elaborating regional transparency and confidence-building measures would be a more feasible option. A starting point in their view would be to establish a regional dialogue on nuclear doctrines and capabilities that would include a discussion on military AI. Given the success of several track 2 dialogues on security between China, India and Pakistan, such an initiative seems to be relatively realistic.

Introduction

1. Introduction

PETR TOPYCHKANOV

This volume explores the question of how artificial intelligence (AI) has an impact on strategic stability and nuclear risk in South Asia. At first glance, this is a narrow subject that has little significance for this region. When the region's main security challenges relate to nuclear and conventional forces and cross-border terrorism, the role of emerging technologies in nuclear weapon systems may seem marginal or at least premature. However, this perception is incorrect: information technology (IT), including AI, has had an impact on the South Asian countries' view of security and stability over several decades. More than 30 years ago at the Conference on Disarmament in Geneva, India presented its concerns regarding the combination of emerging technologies and nuclear weapons:

> The combination of the most recent advances in the field of electronics with the lethality of nuclear explosive power will prove to be deadly. At the same time, the functions of [reconnaissance], surveillance, target identification, kill assessment and evaluation are being re-designed to make greater use of all satellite and other sophisticated sensor technologies and data processing using fifth generation computers. With such deployments, the command and control systems stand in danger of becoming increasingly automated.[1]

At the time, neither India nor Pakistan was a nuclear-armed state. They subsequently crossed the nuclear threshold with tests of nuclear weapons in 1998. Today in South Asia, nuclear competition between India and Pakistan and between India and China have been gaining momentum. India and Pakistan are both expanding their nuclear weapon stockpiles and developing new delivery systems such as nuclear-powered ballistic missile submarines (SSBNs) and cruise missiles.[2] They are also developing emerging technologies such as unmanned vehicles and cyber-warfare. Military AI plays a significant role in many of these systems.

In this volume, the contributors describe current developments in the debates on AI technology in South Asia and explore connections between emerging technologies and nuclear weapon systems in India, Pakistan and those countries that are significant for the region—as the source of either threat or of technology and ideas. As in the other two volumes in this series, the contributors' analysis is built on three concepts: AI, strategic stability and nuclear risk as they relate to the region.[3] The discussions of strategic stability and nuclear risk tend to be specific to each region, but AI is common to all regions.

[1] Singh, K. N., Indian Minister of State for External Affairs, Statement in the plenary of the Conference on Disarmament, Indian Ministry of External Affairs, *Foreign Affairs Record*, vol. 34, no. 1 (Jan. 1988), p. 74.

[2] Kile, S. N. and Kristensen, H. M., 'Indian nuclear forces' and 'Pakistani nuclear forces', *SIPRI Yearbook 2019: Armaments, Disarmament and International Security* (Oxford University Press: Oxford, 2019), pp. 325–31 and pp. 332–37.

[3] Boulanin, V. (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, May 2019); and Saalman, L. (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. II, *East Asian Perspectives* (SIPRI: Stockholm, Oct. 2019).

The term 'artificial intelligence' has many meanings, having gone through many interpretations since the 1950s and being related to a broad spectrum of technologies and activities. In general, this volume follows the understanding of AI proposed in the first volume of this series (see box 1.1). The combination of AI and strategic stability as a topic of interest appeared in scholarly studies over 30 years ago.[4] However, only recently has the interest in this combination become wide in academic and expert communities. In most publications, the focus is on the countries that have achieved advances in military applications of AI and related technologies, such as the United States and China.[5] For South Asian studies, this combination is a relatively new object of research interest.[6]

The wide range of publications on strategic stability in South Asia shows diverse perspectives among scholars on the meaning of strategic stability, rather than a shared vision.[7] One of the issues actively debated in the literature is whether strategic stability as understood during the cold war applies to the relationships between nuclear-armed states in South Asia. If strategic stability is understood in terms of first-strike stability—when leaders on both sides have no reason to make the first nuclear strike and do not expect the first strike from the other side—then the concept may still apply to South Asia today. However, there is a broader understanding of strategic stability that reflects the realities of the post-cold war era.[8] This broader concept appears to have limited application in South Asia.

[4] Din, A. M. (ed.), *Arms and Artificial Intelligence: Weapons and Arms Control Applications of Advanced Computing* (SIPRI: Stockholm, 1987).

[5] Altmann, J. and Sauer, F., 'Autonomous weapon systems and strategic stability', *Survival*, vol. 59, no. 5 (2017), pp. 117–42; Geist, E. and Lohn, A. J., *How Might Artificial Intelligence Affect the Risk of Nuclear War?*, Perspectives series, 24 Apr. 2018 (RAND Corporation: Santa Monica, 2018); Fitzpatrick, M., 'Artificial intelligence and nuclear command and control', Survival Editor's Blog, International Institute for Strategic Studies, 26 Apr. 2019; Saalman, L., 'Fear of false negatives: AI and China's nuclear posture', *Bulletin of the Atomic Scientists*, 24 Apr. 2018; and Sharikov, P., 'Artificial intelligence, cyberattack, and nuclear weapons—a dangerous combination', *Bulletin of the Atomic Scientists*, vol. 74, no. 6 (2018), pp. 368–73.

[6] See e.g. Ahmad, K., 'Artificial intelligence and the future of warfare', Centre for Strategic and Contemporary Research, 2 Sep. 2018; Arif, S., 'Emerging trends of artificial intelligence in South Asia and its implications for Pakistan', *NUST Journal of International Peace and Stability*, vol. 2, no. 2 (July 2019), pp. 55–66; Khan, M., 'Impact of emerging technologies on South Asian strategic stability', Strategic Studies Institute Islamabad, 28 Dec. 2019; Kumar, A., 'AI–nuclear menace: Emerging trends', Centre for Land Warfare Studies, 13 Feb. 2019; Pant, A., *Future Warfare and Artificial Intelligence: The Visible Path*, Institute for Defence Studies and Analyses (IDSA) Occasional Paper no. 49 (IDSA: New Delhi, 2018); Pant, H. V. and Joshi, Y., 'Emerging technologies and India's nuclear deterrent', War Fare, Observer Research Foundation, 5 Feb. 2019; and Rajagopalan, R. P., 'Managing nuclear risks: The emerging technologies challenge', The Diplomat, 24 May 2019.

[7] E.g. Cheema, Z. I., *Indian Nuclear Deterrence: Its Evolution, Development, and Implications for South Asian Security* (Oxford University Press: Karachi, 2010); Ganguly, Š. and Kapur, P., *India, Pakistan, and the Bomb: Debating Nuclear Stability in South Asia* (Columbia University Press: Delhi, 2010); Gregory, S., *Rethinking Strategic Stability in South Asia*, South Asian Strategic Stability Institute (SASSI) Research Report no. 3 (SASSI: Bradford, Sep. 2005); Krepon, M. and Gagné, C. (eds), *The Stability–Instability Paradox: Nuclear Weapons and Brinksmanship in South Asia*, Stimson Centre Report no. 38 (Stimson Center: Washington, DC, June 2001); Sahgal, A., *Examining Efficacy of Strategic Stability in South Asia: An Analysis*, Sandia Report no. SAND2019-0177 (Sandia National Laboratories: Albuquerque, 1 Jan. 2019); and Topychkanov, P., *Nuclear Weapons and Strategic Security in South Asia*, Working Paper no. 3 (Carnegie Moscow Center: Moscow, 2011).

[8] Arbatov, A., Dvorkin, V., Pikaev, A. and Oznobishchev, S., *Strategic Stability after the Cold War* (Institute of World Economy and International Relations (IMEMO): Moscow, 2010), p. 8.

Box 1.1. Key definitions

Artificial intelligence

Artificial intelligence is a catch-all term that refers to a wide set of computational techniques that allow computers and robots to solve complex, seemingly abstract problems that had previously yielded only to human cognition.

Nuclear weapon systems

Nuclear weapon systems should be understood in the broadest sense. They include not only the nuclear warheads and the delivery systems but also all nuclear force-related systems such as nuclear command-and-control systems, early-warning systems, and intelligence, reconnaissance and surveillance systems. Relevant non-nuclear strategic weapons include long-range high-precision missiles, unmanned combat aerial vehicles (UCAVs) and ballistic missile defence systems.

Strategic stability

Strategic stability has many definitions. It is understood here as 'a state of affairs in which countries are confident that their adversaries would not be able to undermine their nuclear deterrent capability' using nuclear, conventional, cyber or other unconventional means.[a]

[a] Podvig, P., 'The myth of strategic stability', *Bulletin of the Atomic Scientists*, 31 Oct. 2012.

Source: Boulanin, V., 'Artificial intelligence: A primer', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, May 2019), p. 13.

Nuclear risk in South Asia has received less attention than strategic stability in the scholarly and expert literature. A possible reason is that the term 'nuclear risk' lacks explicit context because it is used both in the areas of nuclear weapons and non-proliferation and in relation to nuclear safety and security breaches. Researchers in India and Pakistan generally avoid exploring the possibilities of nuclear weapon use and nuclear safety and security breaches concerning their own countries because they may be seen as exposing vulnerabilities.

While this volume uses the term 'South Asia' freely, its scope is criticized in the research literature, especially in India and specifically in the context of nuclear weapons.[9] For Indian authors, this term, when applied to strategic stability and nuclear risk, does not encompass the spectrum of nuclear threats that India faces, the threat from China above all. Strictly speaking, strategic stability in South Asia more appropriately describes the situation of Pakistan and its India-centric nuclear doctrine; it is less appropriate to describe the Indian case of nuclear deterrence against China and Pakistan.

Within this context, this volume (and the project in general) combines these concepts and connects global and regional perspectives, technological and political dimensions, nuclear weapons and non-nuclear capabilities affecting balances between possible adversaries and, last but not least, the narratives in nuclear-armed states and states without nuclear weapons.

[9] E.g. Nayan, R., 'Nuclear India and the global nuclear order', *Strategic Analysis*, vol. 42, no. 3 (2018), p. 235.

Overview

This volume contains eight essays (chapters 2–9) grouped into three thematic parts followed by a final chapter (chapter 10) that summarizes the key conclusions drawn from the essays.

The impact of artificial intelligence on nuclear weapons and warfare

Part I explores the military and technological developments of AI that have the potential to change the face of war, focusing on India, Western countries and Russia.

In chapter 2, Maaike Verbruggen, a doctoral researcher at the Institute for European Studies in Brussels, provides an analysis of the different applications of AI in nuclear forces and non-nuclear technologies that might also affect strategic stability. Her focus is on the USA and its European allies. She argues that AI is becoming extensively integrated into SSBNs and nuclear-capable aircraft. In non-nuclear technologies—owned by the nuclear-weapon states and states without nuclear weapons alike—AI increases detection abilities, precision and speed. These advances, when applied to nuclear weapons, together risk speeding up crisis escalation.

In chapter 3, Kritika Roy, a research analyst at the Institute for Defence Studies and Analyses in New Delhi, explores the factors relating to the application of autonomy and machine learning in nuclear weapon systems, paying special attention to the case of India. She reports that, despite India's interest in integrating autonomy and machine learning into its nuclear and conventional weapon systems, this is unlikely to happen any time soon because of its risk-averse procurement processes.

In chapter 4, Dmitry Stefanovich, a research fellow at the Primakov Institute of World Economy and International Relations of the Russian Academy of Sciences, analyses Russia's applications of AI in its nuclear weapon systems. He argues that AI is an essential area for Russian military development, including nuclear forces. Russia uses or plans to use autonomous systems, machine learning and other related technologies in order to improve logistics and maintenance, early-warning systems, nuclear command, control and communications (NC3) systems, and combat capabilities.

In chapter 5, Sanatan Kulshrestha, a retired rear admiral from India, examines the massive damage potential of advanced military technologies. He outlines developments in nanoenergetic materials, swarming and nanotechnology and their military applications and then describes Indian efforts to develop AI-driven technologies for military purposes. He argues that the combination of these technologies and AI will influence the development of existing strategic weapons and may make it possible to cause devastating damage similar to a nuclear strike without the accompanying radiation.

The impact of military artificial intelligence on strategic stability in South Asia

Part II comprises two essays that investigate the impact that the current or potential incorporation of AI into military systems—whether conventional or

nuclear—could have on strategic stability in South Asia, focusing on India and Pakistan.

In chapter 6, the present author, a senior researcher at SIPRI, describes the characteristics of strategic stability in South Asia and identifies the potential stabilizing and destabilizing effects of integrating AI into various aspects of nuclear weapon systems.

In chapter 7, Saima Aman Sial, a senior research officer at the Center for International Strategic Studies in Islamabad, examines the use of AI in Pakistan's nuclear command and control. She argues that the development of pre-emptive doctrines and the deployment of a variety of offensive and defensive AI-enabled strategic systems may have both positive and negative effects on stability and nuclear deterrence in South Asia. On the one hand, it could accentuate nuclear risk. On the other, it could play a stabilizing role by, for instance, improving situational awareness, enhancing early-warning systems and ensuring the credibility of nuclear deterrence. Overall, she offers cause for concern regarding integration of AI and machine learning in nuclear weapon systems.

Arms control and confidence-building measures in the area of artificial intelligence and nuclear weapons

The two essays of part III offer two different approaches to the question of prevention and mitigation of nuclear risk arising from the introduction of AI into nuclear weapon systems. The first approach is based on humanitarian and human rights laws and the second on confidence-building measures (CBMs).

In chapter 8, Yanitra Kumaraguru, a lecturer in the law faculty of the University of Colombo and coordinator of the Sri Lankan arm of the global Campaign to Stop Killer Robots, explains why the use of lethal autonomous weapon systems (LAWS) contravenes humanitarian and human rights laws. She outlines a proposed ban on the development, production and use of these systems in order to retain meaningful human control over the critical functions of weapons. She also outlines legal arms control instruments that may help to minimize the destabilizing effects of integrating AI into conventional and nuclear weapon systems.

In chapter 9, Malinda Meegoda, a research associate at the Lakshman Kadirgamar Institute in Colombo, offers a perspective from Sri Lanka, a non-nuclear weapon state whose security will be affected in the event of a nuclear crisis between China, India and Pakistan. He highlights the differences in the current South Asian nuclear security context compared to other adversarial geostrategic nuclear relationships, such as that between the Soviet Union and the USA during the cold war. He proposes a variety of CBMs that may partly help to manage crisis scenarios in South Asia, including the establishment of nuclear risk-reduction centres and the creation of an agreement concerning incidents at sea.

The volume concludes in chapter 10 with a summary of the key conclusions drawn from the essays. Notably, the chapter discusses the extent to which the contributors agree on the opportunities and risks that the AI revolution brings

to the field of nuclear weapon systems and strategic stability in the South Asian context.

Part I. The impact of artificial intelligence on nuclear weapons and warfare

The four essays in the first part of the volume explore the military and technological developments in artificial intelligence (AI) that have the potential to change the face of war. The authors explore four questions: What types of AI-driven application can be found in military weapon systems, including nuclear deterrence capabilities? How could recent advances in AI, machine learning and autonomy be used in capabilities for command and control, strategic offence and defence and non-nuclear deterrence? What place do India, Pakistan, Russia, and the United States and its European allies give to AI and other emerging technologies in their military modernization plans? Finally, what impact, both positive and negative, will these developments have on nuclear-armed states and their neighbours?

Maaike Verbruggen looks at the cases of the USA and some of its European allies in chapter 2. Dmitry Stefanovich answers these questions for the case of Russia in chapter 4. Kritika Roy and Sanatan Kulshrestha offer the Indian perspective on these questions in chapters 3 and 5, respectively. While chapter 4 focuses on the specific Russian case, chapters 2, 3 and 5 describe wider trends in the military research and development of AI and related emerging technologies.

PETR TOPYCHKANOV

2. The extensive role of artificial intelligence in military transformation

MAAIKE VERBRUGGEN

There is widespread interest around the world in developing artificial intelligence (AI) for military purposes. There are research projects to develop AI for almost every conceivable military application, from supporting non-combat operations to the strategic domain. AI is a general-purpose technology that does not stand alone but enhances or adds functionality when integrated into military systems. As such, the research and development (R&D) effort on military applications of AI is much more extensive than flagship efforts to conduct fundamental research on AI. In practice, R&D on AI is integrated into military R&D projects across the spectrum. Given the wide range of possible applications of AI, its impact on strategic stability goes far beyond its integration into nuclear weapons alone.

The use of AI in military systems is not a revolutionary change. Many applications, such as decision-making support of command and control, have been in use for decades. However, the functionality of AI is improving, and it will play an increasingly important role in daily operations. This poses well-documented risks such as over-trust and under-trust in machines, difficulties in switching between different levels of human involvement when calamities arise, vulnerabilities in cybersecurity, and biases in algorithms.[1] The risks of AI are not limited to its use in nuclear weapon systems, and AI in non-nuclear military technologies might also affect the nuclear domain.

To assess the impact of AI on strategic stability and nuclear risk, this essay uses a broad lens in examining the extensive role that AI plays in military transformation. It first (in section I) looks at applications of AI in delivery systems for nuclear weapons and then (in section II) turns to AI in non-nuclear military technologies.

I. AI in nuclear weapon delivery systems

In Western countries at the forefront of AI R&D for the military, AI plays an important role in improving and developing nuclear-capable aerial and naval forces. For example, France, the United Kingdom and the United States are all developing new classes of nuclear-powered ballistic missile submarine (SSBN): respectively the SNLE 3G (i.e. third-generation SSBN), the Dreadnought class

[1] Hawley, J. K., *Patriot Wars: Automation and the Patriot Air and Missile Defense System*, Voices from the Field series (Center for a New American Security: Washington, DC, 25 Jan. 2017); and United Nations Institute for Disarmament Research (UNIDIR), *Algorithmic Bias and the Weaponization of Increasingly Autonomous Technologies: A Primer*, UNIDIR Resources no. 9 (UNIDIR: Geneva, 2018).

and the Columbia class.[2] The Dreadnought is predicted to enter service in the late 2020s and the other two classes in the early 2030s.[3] AI will play a significant role in navigation and control of these submarines, as well as in improved underwater detection.[4]

The USA and France are also both investing in their nuclear-capable aircraft, where AI is used extensively as well. For example, the USA is developing the B-21 Raider strategic bomber, which is expected to enter into service around 2025, modernizing its B-2 strategic bomber and making the F-35 combat aircraft nuclear-capable, both of which are already in service.[5] Together with Germany, France announced plans to develop the Future Combat Air System (FCAS), a nuclear-capable combat aircraft expected to enter service around 2040, as part of its Next-Generation Weapon Systems plan.[6] AI is integrated in a host of different functions in these aircraft, including in navigation and control; taking over in emergency situations; acquiring, filtering and fusing data to present the pilot with the most relevant information; suggesting courses of action to execute missions; and coordinating with other platforms. The FCAS is designed to fly in a heterogeneous configuration with a swarm of unmanned aerial vehicles (UAVs), and both the B-21 and the FCAS will be optionally manned. However, both France and the USA have stated that, for now, all nuclear missions will be manned.[7]

II. AI in non-nuclear military technologies

AI will be used in many more military applications than nuclear weapons and their delivery systems. Analysis of the impact of AI on strategic stability, as well as any possible regulatory measures to mitigate adverse impacts, will need to take into account non-nuclear military technologies. The task is complicated by the fact that non-nuclear military technologies cover a wide range of weapon systems that are developed by both nuclear-armed and states without nuclear weapons. This section describes various non-nuclear applications of AI, grouped

[2] French National Assembly, National Defence and Armed Forces Commission, 'Audition de l'amiral Bernard-Antoine Morio de l'Isle, commandant des forces sous-marines et de la force océanique stratégique (ALFOST)' [Hearing of Admiral Bernard-Antoine Morio de l'Isle, commander of submarine forces and strategic ocean force (ALFOST)], 5 June 2019; Allison, G., 'A guide to the Dreadnought class ballistic missile submarine', *UK Defence Journal*, 24 Oct. 2017; and 'From attack submarines to spies: US Navy asks more of its underwater fleet', *Naval Technology*, 12 Dec. 2018.

[3] Vavasseur, X., 'Here is the first image of the French Navy next generation SSBN—SNLE 3G', *Navy Recognition*, 3 Oct. 2018; and Szondy, D., 'Rising tide: Submarines and the future of undersea warfare', *New Atlas*, 5 July 2017.

[4] Clark, B., *The Emerging Era in Undersea Warfare* (Center for Strategic and Budgetary Assessments: Washington, DC, 22 Jan. 2015); Mukherjee, T., *Securing the Maritime Commons: The Role of Artificial Intelligence in Naval Operations*, Observer Research Foundation (ORF) Occasional Paper (ORF: New Delhi, 16 July 2018).

[5] Evans, D. and Schwalbe, J., *The Long-Range Standoff (LRSO) Cruise Missile and Its Role in Future Nuclear Forces* (John Hopkins Applied Physics Laboratory: Laurel, ML, 2017).

[6] Everstine, B., 'French Air Force begins research into sixth generation aircraft', *Air Force Magazine*, 7 Feb. 2019.

[7] Sayler, K. and Scharre, P., 'The B-21 Bomber should be unmanned on day 1', Defense One, 31 May 2016; and Sprenger, S., 'With nukes in mind, French officials stake out must-haves for Franco-German warplane', *Defense News*, 15 Nov. 2018.

according to the benefit that AI can bring—detection abilities, precision, speed, decision-making and access to domains of warfare—and the implications of each application for strategic stability.

AI for detection

AI enhances military situational awareness capabilities. It enables military systems to sift through large amounts of data in order to find relevant information, to fuse together data from different sensors and to observe whether situations have changed. These capabilities increase the chances of detecting objects and activities in all domains of warfare.

One important application is in anti-submarine warfare. AI can aid in detecting objects underwater through more efficient control of capabilities, such as lower frequency active sonar, ambient noise and non-acoustic sensors (e.g. bouncing laser lights).[8] In additional, AI is beneficial for mapping the seabed and tracking currents, so the oceans become more transparent and accessible.[9] Another important application is in analysing satellite imagery; for example, to detect military construction, transport of mobile launchers or changes in military installations.[10]

Concealment is one of the primary methods for nuclear-armed states to guarantee the survivability of their nuclear weapons.[11] This is a key pillar of nuclear deterrence policies. Russia, the UK and the USA rely on SSBNs, which are hard to detect underwater, to ensure second-strike capability.[12] However, enhanced detection systems in the naval and aerial domains may erode concealment policies. A country might be more likely to launch a first strike if it feels confident that it can detect and destroy the nuclear assets of its adversary or if it is afraid that the adversary might destroy its assets in a 'use it or lose it' scenario.[13]

AI for precision

AI allows for targeting with higher precision. Improvements in selection and tracking of a target, including while both the weapon and the target are in motion, can greatly increase the lethality of existing conventional weapons. What such AI-enhanced conventional weapons lack in destructive power compared to nuclear weapons, they make up for in precision. This opens up their use in counterforce operations to target hardened nuclear launchers, which until now

[8] Wilson, J. R., 'Technology comes to bear on radar and sonar', *Military & Aerospace Electronics*, 1 Feb. 2017.

[9] Clark (note 4).

[10] Stewart, P., 'Deep in the Pentagon, a secret AI program to find hidden nuclear missiles', Reuters, 5 June 2018.

[11] Lieber, K. A. and Press, D. G., 'The new era of counterforce: Technological change and the future of nuclear deterrence', *International Security*, vol. 41, no. 4 (Apr. 2017), p. 9.

[12] Brixey-Williams, S., 'Will the Atlantic become transparent?', 3rd edn, British Pugwash, Nov. 2016.

[13] See e.g. Rickli, J.-M., 'The destabilizing prospects of artificial intelligence for nuclear strategy, deterrence and stability', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, 2019), pp. 91–98, p. 94.

needed nuclear weapons to breach their defences.[14] Increased precision against moving targets also enables the use of anti-submarine weapons at a greater distance than previously possible.[15] This raises the strategic value of conventional weapons (especially in convergence with other technologies such as hypersonics), thus altering the conventional–nuclear balance.

AI for speed

The third quality of AI is speed: machines empowered by AI can respond much faster than humans. An example is automatic target recognition (ATR), which has been used for decades in air defence systems such as the Aegis and Patriot systems.[16] However, while advances in AI will probably continue to make incremental improvements in existing air defence capabilities, especially in the realm of sensors, there are no indications that current advances in AI will fundamentally improve air defence to such an extent that it will alter strategic relations.[17] More important might be the effect of AI in conventional weapons. It is possible that at a certain point only autonomous weapons will be fast enough to respond to other autonomous weapons. This could make it necessary for all countries to develop such weapons, leading to a destabilizing arms race.

AI and military decision-making

The use of AI has implications for military decision-making. When integrated in military logistic planning software, AI can drastically reduce the time needed for force deployment, as demonstrated already in the 1990–91 First Gulf War.[18] This make military options available in an unprecedented short time frame. If the speed of conflict increases, there is less time for consultation and input from civilian decision makers. Military leaders might lose the chance to think things over carefully and let cooler heads prevail.

AI can be used to generate highly convincing fake images (e.g. of satellite imagery of the construction or transport of (mobile) launchers), fooling both militaries and the public alike.[19] This can lead to public outrage and bottom-up demand for retaliation, fuelled by large botnets spreading misleading information and creating the sense that the outrage is wider than it actually is. Even though AI is not the main cause, it can exacerbate the difficulty of navigating crises in

[14] Koblentz, G., *Strategic Stability in the Second Nuclear Age* (Council on Foreign Relations: New York, Nov. 2014).

[15] Clark (note 4).

[16] Scheer, J. A. and Holm, W. A., 'Introduction and radar overview', eds M. A Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar: Basic Principles* (SciTech Publishing: Raleigh, NC, 2010), pp. 46–47.

[17] Judson, J. 'Hyten: To address Russian and Chinese missile threats, it's all about the sensors', *Defense News*, 7 Aug. 2018.

[18] Hedberg, S. R., 'DART: Revolutionizing logistics planning', *IEEE Intelligent Systems*, vol. 17, no. 3 (May/June 2002), p. 81.

[19] Tucker, P., 'The newest AI-enabled weapon: "deep-faking" photos of the Earth', Defense One, 31 Mar. 2019.

the existing information ecosystem. AI might worsen the existing pressure on decision makers in crises.[20]

Both accelerated decision-making and decisions based on misinformation may lead to crisis escalation in an increasingly rapid manner, creating a risky strategic environment, potentially on a global level.

AI for access to domains of warfare

AI also opens up or enhances access to domains of warfare that were previously hard to access. The underwater domain is mentioned above as a domain that will become easier to penetrate thanks to AI. Other difficult domains include the polar regions and space. Thanks to AI, unmanned systems can operate in environments extremely hostile to humans, such as outer space or the seabed underneath and around the Arctic ice pack. Satellites can use AI to navigate orbits more easily, avoid space debris and other objects, and conduct real-time geospatial analysis.[21] Unmanned sensors in the Arctic can monitor surface, air and submarine traffic.[22] In any domain, AI can be used for planning missions and monitoring the state of systems and making repairs if needed. However, nuclear-armed states traditionally rely on the inaccessibility of these realms to secure their nuclear assets. Russia stations a significant part of its SSBN fleet in the Arctic, and satellites are part of the USA's nuclear command, control and communications (NC3) infrastructure. Improved accessibility increases the vulnerability of the nuclear assets stationed there.

AI also plays an important role in the domains of cyberwarfare and electromagnetic warfare, as all three technologies are highly convergent. AI enables malicious software to adapt to constantly updated cyberdefences and test new methods to penetrate adversarial systems. By now it is clear that cyber technologies have an important role in strategic stability. Cyber operations to take out air defence systems or missiles decrease the reliability of nuclear weapons. AI is also critical for the use of electromagnetic weapons, as it sifts through large amounts of data and helps to find the optimum frequency to operate in a crowded spectrum.[23] Since these weapons have non-kinetic effects, leaders might be more likely to use them if they believe that their use will be less escalatory than weapons with kinetic effects. One possible use is to disable command, control and communications conducted through satellites.[24] This increases the risk of entanglement between the conventional and nuclear realms. In the fog of war,

[20] *Three Tweets to Midnight: Nuclear Crisis Stability and the Information Ecosystem*, Policy Dialogue Brief (Stanley Foundation: Muscatine, IA, Feb. 2018).

[21] Erwin, S., 'Artificial intelligence arms race accelerating in space', SpaceNews, 3 May 2018.

[22] Keller, J., 'DARPA approaches industry for unmanned sensors to monitor Arctic land, sea, and air traffic', *Military & Aerospace Electronics*, 25 Feb. 2015.

[23] Pomerleau, M., 'New army AI is cutting through data-choked battlefields', C4ISRNET, 20 Dec. 2018.

[24] 'Source reveals tech details of new Russian anti-satellite warfare plane', Sputnik, 9 July 2018.

there will be much uncertainty about the intentions of cyber and electromagnetic attacks, creating a risk of retaliation with nuclear weapons.[25]

III. Conclusions

AI is not a stand-alone technology. It is highly convergent with other technologies, and its effects depend on integration into other systems. It has an incredibly wide range of applications, and it is more accessible than nuclear technology. The impact of AI on strategic stability thus goes far beyond its use in nuclear weapons or installations. The application of AI to non-nuclear technologies increases detection capabilities, making nuclear assets harder to hide and secure from a first strike. It also increases targeting precision in conventional weapons to enable use against SSBNs and hardened missile launchers. Moreover, it increases the speed of decision-making, which risks states racing up the conflict escalation ladder. AI also opens up new geographic and virtual domains of warfare, creating new vulnerabilities for nuclear assets and increasing the likelihood of retaliation.

[25] Acton, J. M., 'Escalation through entanglement: How the vulnerability of command-and-control systems raises the risks of an inadvertent nuclear war', *International Security*, vol. 43, no. 1 (Aug. 2018), p. 56.

3. Rationales for introducing artificial intelligence into India's military modernization programme

KRITIKA ROY

The blend of autonomy and artificial intelligence (AI) is known to multiply the effectiveness of any weapon system. The growing popularity of intelligent autonomous systems as a vital area of research and development (R&D) is demonstrated by China's aim of becoming world leader in the field of AI and the renewed push by the United States of the technology as its Third Offset Strategy.[1]

India is cognizant of the promise of AI and machine learning technologies. Although the country adopted a more civilian approach to its national AI strategy by focusing on how to 'leverage AI for economic growth, social development and inclusive growth', India is now facilitating R&D to move towards an AI-driven military ecosystem.[2] In June 2018 the Ministry of Defence's AI Task Force submitted its final report, which included the recommendation that India become 'a significant power of AI in defence specifically in the area[s] of aviation, naval, land systems, cyber, nuclear, and biological warfare'.[3] Although there are no clear indications at this time that machine learning will be directly integrated into mission-critical systems, such integration cannot be ruled out in future military development. For example, India is using its growing biotechnology infrastructure to support biodefence R&D, including the development of countermeasures ranging from protective equipment via pharmaceuticals to vaccines.[4] Integrating AI within the biotechnology sector for detection, diagnosis and decontamination measures (e.g. using unmanned ground vehicles and robots in contaminated zones) may aid in strengthening India's biodefence architecture.

This essay looks first (in section I) at factors relevant to the application of autonomous intelligence generally, then (in section II) at developments in India to explore its potential application in military-specific areas, and finally (in section III) at regional considerations.

[1] Cadell, C. and Jourdan, A., 'China aims to become world leader in AI, challenges U.S. dominance', Reuters, 20 July 2017; and Thomas-Noone, B., 'US playing catch-up as technology advantage erodes', *The Strategist*, 17 July 2018. See also Saalman, L., 'Exploring artificial intelligence and unmanned platforms in China', ed. L. Saalman, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. II, *East Asian Perspectives* (SIPRI: Stockholm, Oct. 2019), pp. 43–47; and Stoutland, P., 'Artificial intelligence and the modernization of US nuclear forces', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, 2019), pp. 63–67.

[2] National Institution for Transforming India (NITI Aayog), *National Strategy for Artificial Intelligence: #AIforall*, Discussion paper (NITI Aayog: New Delhi, June 2018), p. 7.

[3] Indian Ministry of Defence, Press Information Bureau, 'AI task force hands over final report to RM', Press release, 30 June 2018.

[4] Nanjappa, V., 'India has never pursued an offensive bio weapons programme', Rediff, 21 Dec. 2012.

I. Factors relevant to the application of autonomy and machine learning

There are several factors 'central to the debate' on the application of autonomy and machine learning within weapon systems.[5] The term 'autonomy' in relation to machines has various definitions, but the general understanding is that it means 'the ability of a machine to perform an intended task without human intervention using interaction of its sensors and computer programming environment'.[6] 'Machine learning' is defined as 'an approach to software development that consists of building a system that can learn and then teaching it what to do using a variety of methods'.[7] Both are AI applications, which can be used in combination.

Degree of human involvement

In discussions about autonomous weapons, the position of human decision-making—'in the loop', 'on the loop' or 'out of the loop'—is hotly debated.[8] Machine learning provides a machine with the ability to perform intended tasks without human intervention in ways that go beyond simple stimulus–response automation but does not yet give the machine the capability to make complex decisions like a human. In other words, this technology has yet not matured to the extent that it could match the depth, robustness, flexibility and ethical dimensions of human cognition and intuition that inform human decision-making.

For instance, autonomy and machine learning can be integrated into offensive missile systems to enhance their performance. However, allowing these systems to autonomously take a decision also means allowing room for error, as in the incident of 1983 when Lieutenant Colonel Stanislav Petrov of the Soviet Union decided to follow his own instincts rather than the machine's information, declaring a 'false alarm' and thus averting a nuclear crisis.[9]

Offensive and defensive tasks

Another major factor in integrating autonomy and machine learning within any weapon system is whether the system is meant to perform offensive tasks or

[5] This section takes its headings from the list of central factors in Gill, T. D., 'Unmanned and autonomous weapons: Have they left the law behind?', Presentation, University of Amsterdam and Netherlands Defence Academy, slide 10.

[6] Boulanin, V. and Verbruggen, M., *Mapping the Development of Autonomy in Weapon Systems* (SIPRI: Stockholm, Nov. 2017), p. vii. See also Boulanin, V., 'Artificial intelligence: A primer', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, 2019), pp. 13–25.

[7] Boulanin and Verbruggen (note 6), p. 16.

[8] Tucker, P., 'Report: Weapons AI increasingly replacing, not augmenting, human decision making', Defense One, 26 Sep. 2016.

[9] Irving, D., 'How artificial intelligence could increase the risk of nuclear war', RAND Corporation, 24 Apr. 2018. See also Topychkanov, P., 'Autonomy in Russian nuclear forces', ed. Boulanin (note 6), pp. 68–75, box 8.2.

defensive tasks. This is a potent question as this consideration would influence an opponent's strategic posture and the overall strategic stability of the region.

For example, making offensive weapon systems fully autonomous may induce a sense of fear in the opponent, who could then believe that making a pre-emptive strike is the best defence.[10] At the same time intelligent autonomous systems could be used in missile defence systems so that tracking and immediate retaliation would be in line with the scale of the attack.

Type of target to be engaged

Another factor relevant to the introduction of autonomy and machine learning into weapon systems is the type of target; that is, does the weapon system target military platforms, incoming missiles or the local populace? The level of integration of autonomy within a system needs to be appropriate to the target.

For instance, interception of missiles needs an immediate retaliatory response and so an autonomous weapon system could act as a positive force. However, if the system is used to target human populations, there have to be effective safety measures in place, as one small glitch could lead to massive fatalities.

Type of strategic environment

Another important factor is the strategic environment of the state that is integrating autonomy and machine learning within its weapon systems—that is, whether the state is politically volatile or stable, how secure its strategic assets are, whether the civil government is dominant or military rule prevails.

Autonomous weapon systems can be customized for different kinds of government. For instance, a military regime could use such a system to tighten its grip on restive populations. A weak state may be tempted to adopt the technology not only to strengthen its own power, but also to equip proxies such as insurgent or terrorist groups. Additionally, some states could provide autonomous weapon systems to terrorists, which could use them to wreak havoc.

Predictability of how the system would perform

With growing advances in technologies, high-speed systems are increasingly able to accomplish more complex tasks in more complex setups. If machine learning is applied in these systems, they may have the ability to modify their 'behaviour' in ways not foreseen by their developers.

For example, bots have been used in stock trading for quite some time as they are known to execute a trade in milliseconds and even microseconds. On 6 May 2010 the Dow Jones Industrial Average inexplicably experienced a series of drops that led to an estimated loss of US$1 trillion in the US stock market. Although

[10] Klare, M. T., 'Autonomous weapons systems and the laws of war', *Arms Control Today*, vol. 49, no, 2 (Mar. 2019).

the bot system was seen as an efficient and intelligent capital-allocation machine, when these bots began to interact they created unexpected behaviours leading to a 'flash crash'.[11] Imagine a similar scenario in a battleground using autonomous weapon systems—any unexpected behaviour could lead to massive fatalities.[12]

The difficulty in predicting ahead of time the boundaries of behaviour of intelligent autonomous systems makes it a crucial consideration in applying machine learning and autonomy to weapon systems.

II. Military applications of autonomy and machine learning in India

India's military modernization plans take a long-term perspective.[13] India is a growing economy and its objective has always been to use its resources optimally in a cost-effective manner. Implementing autonomy and machine learning in its military systems could have the potential to change how traditional armed forces operate, from training and logistics management, via the command chain to force deployment.

Indeed, substantial work has already been done in deploying AI in the civilian sector, with several Indian companies having built considerable expertise in this area.[14] It would, therefore, not require much effort to transfer the technology, knowledge and expertise already present in the civilian sector to meet military needs, at least at the preliminary level.

India's Defence Research and Development Organisation (DRDO), which looks into the military needs of the country, has done some work in the area but still has a long way to go. It has established the Centre for Artificial Intelligence and Robotics (CAIR), a dedicated laboratory for AI-based research. As the name suggests, CAIR's research focuses mainly on AI, robotics and intelligent control systems.[15] A remarkable achievement of CAIR is the AI techniques for Net Centric Operations (AINCO) project, a customized 'suite of technologies for creation of knowledge base, semantic information reception and handling, interference reasoning and event correlation'.[16] There is also a proposal to develop a series of intelligent robots for intelligence, surveillance and reconnaissance (ISR) purposes. These would have the capability to navigate in semi-structured environments with the ability to sense roadblocks and real-time feedback. Other CAIR initiatives include the Multi-Agent Robotics Framework (MARF) project, which includes different types of robot such as the Wall Climber and the Snake,

[11] Salmon, F. and Stokes, J., 'Algorithms take control of Wall Street', *Wired*, 27 Dec. 2010.

[12] Horowitz, M. C., 'Artificial intelligence and nuclear stability', ed. Boulanin (note 6), pp. 80–83.

[13] Dutta, A., 'India's defence modernisation: Challenges and prospects', *Indian Defence Review*, 7 July 2016.

[14] Sachitanand, R., 'Here's why Indian companies are betting big on AI', *Economic Times* (New Delhi), 10 Feb. 2019.

[15] On CAIR see also chapter 5 in this volume.

[16] Chakravorty, P. K., 'Artificial intelligence and its impact on the Indian armed forces', *Indian Defence Review*, 5 May 2017.

and research into such applications as image recognition for target detection and classification.[17]

Some areas where the application of autonomy and machine learning currently benefits India's military services are described below along with probable areas for future application.[18] Despite India's advances in these areas, the integration of autonomy and machine learning within mission-critical systems will still take many years and will probably need another revolution in AI. This is because India's established procedures are risk-averse, making adoption of new military technologies relatively slow. Moreover, autonomy and machine learning are seen more as collaborative ways to enhance the effectiveness of existing systems than as comprehensive solutions.

Intelligence, surveillance and reconnaissance capabilities

ISR capabilities are a key element of situational awareness. In this regard, autonomy and machine learning could be used for simultaneous collection and processing of data from various sources. This would be extremely useful in real-time monitoring of data in a network-centric environment, facilitating dynamic visualization of actual occurrences on the ground. India's armed forces have already begun using unmanned aerial vehicles (UAVs) in reconnaissance, border security and maritime patrol. The DRDO has successfully designed and developed many versatile UAVs including Nishant, Lakshya and Rustom.[19] Jointly with IdeaForge, the DRDO has developed Netra, a mini UAV quadcopter used for surveillance and reconnaissance operations, which has an autonomous navigation and guidance system.[20]

ISR data collected and collated using various means could be analysed using autonomy to provide much deeper insights, either to augment human capabilities or give humans more time to focus on important aspects of decision-making. CAIR is developing a system called the Command Information and Decision Support System (CIDSS) that would facilitate storage, retrieval, processing (filtering, correlation, fusion) and visualization of tactical data, and provide effective decision support to commanders at the requisite time.[21]

Early-warning and control systems

Early-warning and control systems are vital strategic instruments for detecting incoming threats. Autonomy and machine learning could be used in boosting the detection capabilities of extant early-warning systems.

[17] Chakravorty (note 16).

[18] Reddy, R. S., 'How AI can help the Indian Armed Forces', LiveMint, 5 Mar. 2018.

[19] Indian Defence Research and Development Organisation (DRDO), 'Unmanned aircraft systems and technologies', *Technology Focus*, vol. 18, no. 6 (Dec. 2010), p. 1.

[20] Bhardwaj, V., 'DRDO Netra mini unmanned aerial vehicle (UAV), quadcopter, Indian Armed Forces', AerMech.IN, 23 Oct. 2015.

[21] Indian Defence Research and Development Organisation (DRDO), 'Major products', [n.d.].

Himshakti is an integrated electronic warfare system developed by the DRDO. It can be used for surveillance, analysis, interception, direction finding, position fixing, signal intelligence and jamming of all communication and radar signals while protecting electronic assets in the battlefield.[22]

Protection from cyberthreats

Autonomy and machine learning can be employed to automate threat detection and facilitate in-built capabilities to counter cyberthreats to the strategic assets of a country.[23] For example, the technology could be used to patch a country's own vulnerabilities while exploiting those of its opponent.

Resource management and logistics support

Autonomy and machine learning can be used in resource management and inventory tracking to enable the creation of one central location to send signals when inspections are needed and to indicate precisely which parts require repair. This central location can also automatically send signals to a local production cell to indicate other needs (e.g. fuel requirements) of a unit. Autonomously carrying out such activities could allow personnel to be engaged in more important tasks.

Predictive maintenance

The process of fault detection and diagnosis could be carried out by autonomous systems to indicate wear and tear and also to keep track of servicing and other requirements in weapon systems to ensure that they are mission-ready at all times. In this context, CAIR has developed robots for non-destructive testing of composite parts of its light combat aircraft, the HAL Tejas.[24]

Simulations

Autonomy and machine learning could be employed to run simulations and war-games for training and R&D purposes.[25]

III. Regional impacts

India is flanked by two other nuclear-armed states: China and Pakistan. It is wary of the ongoing modernization of China's existing nuclear forces and of Pakistan's steadily growing arsenal, which now features tactical nuclear weapons as well as

[22] 'Himshakti EW: India indigenous electronic warfare system', Indian Defence Update, 4 Aug. 2018.

[23] Seth, V., 'Artificial intelligence based cyber antivirus technology', Startup@IITD, Indian Institute of Technology Delhi, [n.d.].

[24] Chakravorty (note 16).

[25] Pant, A., 'Internet of things centricity of future military operations', *Journal of Defence Studies*, vol. 13, no. 2 (Apr.–June 2019), p. 37.

conventional weapons.[26] Additionally, the deepening of the strategic partnership between these states is a growing concern for India.[27]

However, the region is under the shadow of the stability–instability paradox (especially in the India–Pakistan case).[28] This highlights that, although the possession of nuclear weapons has stopped an all-out war between the countries, low-intensity armed conflict or limited war cannot be ruled out even in the foreseeable future.[29] Any new development in the region—be it the development of small conventional arsenals, UAVs or information warfare—disrupts regional stability. Thus, the introduction of autonomy and machine learning into military systems adds to the deep-rooted instability in the region.

If India facilitates the application of autonomy and machine learning in mission-critical systems such as nuclear weapon systems or missile systems, then the impact on regional stability could range from positive to negative.

On the positive side, these technologies may aid in improving military capabilities by providing better information, enhancing decision-making capability and increasing the speed of engagement. With further advances, the technologies could also facilitate the monitoring of nuclear weapon-related development, conduct verification operations, and detect cyberattacks or any third-party interceptions.

On the negative side, these technologies may at the same time pose a severe threat. The same detection capabilities could lead to an arms race or may give a false perception of adversaries' capabilities, leading to pre-emptive strikes in the region. There is also a credible threat from human spoofing attacks or data poisoning, considering data is the central ingredient in machine learning.[30]

IV. Conclusions

India today follows a technological revolution in military affairs model where on the one hand the country has set procedures and weapon systems in place, while on the other hand there is dedicated military R&D of autonomy and machine learning within weapon systems. The general consensus within the country has been that AI has the potential to have a transformative impact on national security and provide military superiority. The 2018 report of the AI Task Force

[26] Bommakanti, K. and Kelkar, A., 'China's military modernisation: Recent trends', Observer Research Foundation (ORF) Issue Brief no. 286, Mar. 2019; and Ahmed, M., 'Pakistan's tactical nuclear weapons and their impact on stability', Regional Voices on the Challenges of Nuclear Deterrence Stability in Southern Asia, Carnegie Endowment for International Peace, 30 June 2016.

[27] Press Trust of India, 'China to boost military cooperation with Pakistan: Report', *Economic Times* (New Delhi), 12 July 2018.

[28] The 'stability–instability paradox' is the inverse relationship between the probability of nuclear and conventional military conflict—where there are two nuclear powers, generally the likelihood of nuclear conflict declines as the risk of conventional war increases. Similarly, as the likelihood of nuclear conflict increases, the risk of conventional war declines.

[29] Kumar, A., 'Theories of deterrence and nuclear deterrence in the subcontinent', ed. E. Sridharan, *The India-Pakistan Nuclear Relationship: Theories of Deterrence and International Relations* (Routledge India: New Delhi, 2018), chapter 6.

[30] Avin, S. and Amadae, S. M., 'Autonomy and machine learning at the interface of nuclear weapons, computers and people', ed. Boulanin (note 6), pp. 105–18.

of the Indian Ministry of Commerce and Industry states that national security imperatives need the development of 'technology based force multipliers' and that 'areas where AI based systems could be usefully deployed' are 'AI based cyber-attack mitigation and counter attack systems', adaptive communication systems, autonomous surveillance and combat systems and 'multi-sensor data fusion based decision-making systems'.[31] However, India's risk-averse military procurement process will hinder this transformation.

As the technology advances and India takes steps to slowly and steadily apply autonomy and machine learning to strengthen its defensive and offensive capabilities, the stability of the region may experience some turbulence where adversary states, such as China and Pakistan, also work in the same direction to build up deterrence capability. However, these technological developments are not solely responsible for the instability of the region.

[31] Task Force on Artificial Intelligence, *Report of Task Force on Artificial Intelligence* (Ministry of Commerce and Industry: New Delhi, Mar. 2018), p. 25.

4. Artificial intelligence advances in Russian strategic weapons

DMITRY STEFANOVICH

The military is historically an area of scientific advances and early adoption of cutting-edge technologies. A set of new technologies known under the umbrella term 'artificial intelligence' (AI) is currently being developed on a large scale, including systems related to nuclear and other strategic forces. The Russian military is no exception: AI is an important area for Russian military research and development (R&D), including in nuclear forces, and AI-related technologies are already being widely applied. Machine learning, autonomous systems and other related technologies are used or planned to be used in order to improve military logistics and maintenance, early-warning systems; nuclear command, control and communications (NC3), and warfighting capabilities.

The Russian military–political leadership also stresses the importance of this domain: in December 2018 Russian President Vladimir Putin emphasized 'digital technologies and artificial intelligence, robotization, and unmanned systems' as the major field of 'the qualitative development agenda of our Armed Forces'.[1] This is the reality today and the only way to reduce the chance of inadvertent escalation between different actors seems to be to limit any ambiguity in the communication of their intentions.

It is crucial to distinguish the use of, on the one hand, AI in decision-making processes and implementation of those decisions (at the management or headquarters level) and, on the other hand, AI as part of on-board control systems for weapons and other military equipment. Those two domains have different kinds and extents of fail-safe controls and measures to ensure that mistakes are avoided.

Before looking at how AI is applied in practice in Russia (in section II), this essay starts (in section I) by describing how AI is defined in Russian military contexts.

I. Definition of AI in Russia

Russia's *Strategic Rocket Forces Encyclopaedia* offers a definition for AI in military affairs (posted not later than in 2011) as 'a field of research in which models, systems, and devices imitating human intellectual activity (perception and logical inference) in warfare are developed'.[2] It divides the three main areas of AI research into knowledge-based systems, neural systems and heuristic search systems.

The areas of specific interest to Russia's Strategic Rocket Forces (SRF) are decision support systems, intelligent systems and weapons (on-board control systems), expert systems and automation. In this context, an 'expert system'

[1] President of Russia, 'Defence Ministry Board meeting', 18 Dec. 2018.

[2] [Artificial intelligence in the military], *Strategic Rocket Forces Encyclopaedia* (Russian Ministry of Defence: Moscow, [n.d.]), (in Russian, author translation).

means a set of software tools that implements AI methods based on knowledge. The expert system accumulates knowledge from a subject area within a specific knowledge model (e.g. production, network, frame) and uses this data to bring new knowledge, solve practical intellectual problems and explain the course of its decision. The components of an expert system include a knowledge base, a linguistic processor for user communications, a solver that implements an inference engine, knowledge-acquisition software, and a provider of explanations of the course taken and the result of a problem-solving process.

II. Applications of AI in Russia's nuclear weapon systems

Warnings and orders

The critical area for autonomous analysis and decision preparation is early-warning and related systems. In this case, major tasks for AI-related technology are incoming threat assessment and damage forecasting. This technology may help to understand the scope of an attack, its origins and its possible intentions, and to swiftly develop an appropriate response scenario, including retaliation.

In retaliation scenarios, NC3 systems become crucial. Machine learning and related technologies will provide decision-making support, including counter-manoeuvring of assets and strike plan optimization. Real-time updates, sensor fusion and other modern solutions help to improve the quality of battle management.[3]

Autonomous transmission and execution of orders win the Perimetr system (called Dead Hand in Western sources) remains a possibility, although a theoretical one.[4] At present, humans remain in the decision-making loop at all times, so it may be appropriate to see this system with signal missiles as another layout of NC3. However, a fully automated launch process is indeed technologically feasible, and its elements may be in place already. In case strategic stability deteriorates rapidly (e.g. if the United States deploys intermediate-range and shorter-range missiles in Europe), the decision to pre-delegate launch authority to Perimetr may once more be a possibility.[5]

Intelligent logistics

Two recent examples of AI-related military research are available in the domains of logistics and maintenance systems.

[3] RIA Novosti, [Monitoring system for national defence management centre], Geodesy Research Institute, [n.d] (in Russian).

[4] See Borrie, J., 'Cold war lessons for automation in nuclear weapon systems' and Topychkanov, P., 'Autonomy in Russian nuclear forces', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, 2019), pp. 41–52 and pp. 68–75.

[5] Odnokolenko, O., [Colonel General Victor Esin: 'If the Americans finally deploy their missiles in Europe, we will have to replace the launch under attack doctrine with the doctrine of pre-emptive strike'], *Zvezda*, 8 Nov. 2018 (in Russian).

In 2015 a group of researchers proposed a model of an automated control system of logistics support (including at wartime) for SRF units armed with road-mobile intercontinental ballistic missiles (ICBMs).[6] This model could increase management quality, enhance communications resilience (including under electronic warfare attacks) and decrease the volume of transmitted information. The researchers offered a mathematical model of material and technical support during preparation and operations that expands applied methods for improvement and modernization of control system elements.

In 2018 another research group proposed and tested a neural network-based model as a solution for predicting the remaining operation time of SRF equipment.[7] This model aims to determine the significance and calculate the weights of diagnostic signs and could be implemented for different models of weapons and for military and special equipment in the SRF. The model went through a full-scale experiment (although on non-military equipment). Given the significant number of legacy systems in SRF service, including previous generation of ICBMs and communication equipment, the model may eventually become an advantageous maintenance solution.

Automated warfare

Targeting is the domain where machine learning and related technologies may be currently most useful in warfare. Intelligent systems that process images can identify an object as a particular type of target, its precise location and its vulnerabilities. At the battle planning stage, such targeting can lead to optimization of allocation and yields of nuclear warheads.

With the help of AI, on-board controls of re-entry vehicles may achieve higher precision and better manoeuvrability. Such R&D is currently carried out with regard to tactical cruise missiles.[8] Post-boost vehicles are likely to use similar but hardened technologies. Hypersonic glide vehicles (e.g. the Avangard missile system with a winged warhead that was expected to reach initial operational capability in December 2019[9]) experience plasma build-up around the glider during endoatmospheric flight, which affects the ability to send and receive signals to and from the vehicle. This severely reduces the capability of flight control by both external means and internal sensors. Sophisticated on-board control and guidance systems seem an appropriate solution.

Another way for AI to contribute to automated warfare is through missile defence penetration aids in general, including smart decoys and trajectory-shaping

[6] Isaev, A. V. et al., [Model for automated control system for material support of military units and formations of the SRF in the development of materiel and technical support system of the armed forces of the Russian Federation], *Nauka i Voennaya Bezopasnost*, no. 3 (2015), pp. 59–65 (in Russian).

[7] Gaivoronsky, O. V., Kartunin, D. N. and Wojciechowski, I. A., 'Model of determining the significance and calculation of weight coefficients of diagnostic signs for forecasting residual resource of complex technical system with uneven development resource', Paper presented at the 42nd Academic Space Conference, Moscow, 23–26 Jan. 2018.

[8] Ramm, A. and Litovkin, D., 'Self-learning cruise missiles will appear in 2050', *Izvestia*, 11 Aug. 2017.

[9] TASS, 'Over 30 Yars, Avangard ICBMs to assume combat duty in Russia next year', 19 Dec. 2018.

capabilities. The main challenges for the successful interception of an incoming threat are to define whether it is a real warhead or a dummy and to estimate its trajectory. A dummy warhead enhanced with AI may make the dummy behave like an actual delivery vehicle, while AI on a real warhead can make random changes of flight path to prevent interception.

Underwater warfare is another critical domain in which vehicle autonomy is essential because of challenges to communication with command centres. A signal may not reach its intended destination or it may be intercepted by an adversary and so reveal the location of the vehicle or other underwater object. Poseidon (Status-6) class unmanned underwater vehicles (UUVs) will enter service with the Russian Navy relatively soon: tests have been successful and the first carrier will undertake sea trials in 2020.[10] Despite popular belief that the Poseidon UUV may be used for nuclear delivery, it may instead be used as a situational awareness tool, or for conventional tasks such as precision minelaying. AI seems to be a useful solution for any of those tasks given the challenges of underwater warfare, effectively turning this new weapon into an underwater force multiplier. However, given the fact that many countries currently pursue 'heavy' UUVs, the ways in which these systems may interact requires specific research, especially because of the importance of sea-based deterrents (e.g. nuclear-powered ballistic missile submarines, SSBNs) in the force postures of most nuclear weapon states.

III. Conclusions

AI elements are already used by Russia's SRF to support decisions made by humans in the field. But there is also the threat that, in a military situation, human analysis will be inadvertently replaced by AI 'thinking': the views and assessments of operators and commanders may be shaped by machine conclusions. Even though a person remains in the cycle of assessing the situation, making decisions and deploying weapons, these actions will be based on the data provided exclusively by the machines.

Similarly, AI is already being used in guidance and control systems of missiles and anti-missile systems, and in other systems such as UUVs. Machine learning and autonomy lead to higher survivability, precision and penetrating capabilities of weapons, making them a tempting direction for military R&D.

Unlike the classic cold war balance between two superpowers, the current polycentric world nuclear order (or disorder) faces greater ambiguity, often intentional. Nuclear-armed states often avoid transparency in their internal procedures for nuclear command and control, targeting and, obviously, warhead allocation. The situation will probably remain this way as long as nuclear deterrence remains a working concept and is employed as an instrument in military–political affairs.

[10] TASS, 'Russia floats out first nuclear sub that will carry Poseidon strategic underwater drones', 23 Apr. 2019. See also Hwang, I. and Kim, J., 'The environmental impact of nuclear-powered autonomous weapons', ed. L. Saalman, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. II, *East Asian Perspectives* (SIPRI: Stockholm, Oct. 2019), pp. 86–80.

However, with advances in science and technology, the situation may become increasingly unstable.

Regional and global partners and adversaries need regular assessments of Russian nuclear capabilities—and they can make mistakes that will affect threat perceptions. Moreover, there is a chance that advances by different actors in nuclear force-related domains may lead to levels of insecurity sufficient to fuel a full-scale AI arms race or even nuclear escalation out of fear of losing deterrence capabilities. To avoid mistakes with far-reaching consequences, nuclear modernization and doctrines, including the role of machine learning and autonomy, must be a permanent topic of bilateral and multilateral discussions.

5. The Indian perspective on the massive damage potential of advanced military technologies

SANATAN KULSHRESTHA

Advocacy for denuclearization has gained momentum in justifying global prohibitions on the production, retention or use of nuclear weapons. But it has not deterred the quest for development of powerful conventional weapons that may perhaps cause as much damage as the primitive atomic bomb, although without its radiation effects. Nor has it deterred the quest to incorporate emerging technologies into nuclear arsenals to make them safer and more reliable to operate. In the context of denuclearization, it is just as important to scrutinize developments in conventional weapons lest such developments lead to massive damage to humankind and nature.

The few years have seen unprecedented advances in technologies that in times of war would have an impact on the world as a whole. Such technologies are now within reach of many states, including India, and would allow such a state to sidestep the nuclear option and its attendant radiation fallout while causing just as much devastation to the adversary.

This essay focuses first (in section I) on technologies that may affect how warfare is conducted in the future and then (in section II) looks at India's application of these and other technologies to enhance its conventional military capabilities.

I. Technologies that may change future wars

Advances in nanoenergetic materials

The study of synthesis and fabrication of energetic materials or composites at nanoscale is known as nanoenergetics. Examples of nanoenergetic materials (nEMs) are metal oxides such as aluminium–copper(II) oxide and metal–metal composites such as aluminium–titanium. In the near future, nEMs may form the basis of materials used in a wide range of military systems.[1] Induction of nano-enabled energetic systems with controlled energy release is the focus of current research at such institutes as the United States Naval Academy, the US Navy's Naval Surface Warfare Centers and the University of Maryland.[2]

In simple terms, nEMs perform better than conventional materials because of their much larger surface area, which increases the speed of reaction and enables a larger energy release in a much shorter time. A heterogeneous mixture of a metal (the fuel, e.g. aluminium) and a metal oxide (the oxidizer, e.g. molybdenum

[1] Kaste, P. J. and Rice, B. M., 'Novel energetic materials for the Future Force: The army pursues the next generation of propellants and explosives', *AMPTIAC Quarterly*, vol. 8, no. 4 (2004), p. 89.

[2] Kavetsky, R. et al., 'Energetic systems and nanotechnology: A look ahead', *International Journal of Energetic Materials and Chemical Propulsion*, vol. 6, no. 1 (2007); and Kavetsky, R., 'The navy's program in nanoscience and nanotechnology: A look ahead', US Office of Naval Research, 2004.

oxide), both of nanoscale dimension, results in a class of high-reaction-rate metastable intermolecular composites called nano-thermites or super-thermites. Aluminium–molybdenum oxide, aluminium–Teflon and aluminium–copper(II) oxide have been researched for military use.[3] The addition of super-thermites (nano-aluminium based) to existing compositions has shown an immediate increase in explosive power. The use of nanosized materials in explosives has increased safety and insensitivity by as much as 30 per cent or more without affecting reactivity. It is predicted that nEMs would provide the same explosive power at mass up to two orders of magnitude less than current explosive systems.[4] While nanosizing of high explosives leads to an increase in their explosive power and a decrease in their sensitivity to external forces, it also decreases their thermal stability. The shelf life of such explosives could, therefore, be reduced; however, some patents reveal that this problem has also been resolved technically.[5]

It is expected that nEMs will replace conventional explosives and provide existing conventional weapons with explosive powers higher in magnitude by a factor of two, with enhanced safety because of lower sensitivity to external stimulation by at least 30 per cent.[6] Further, research at the University of Texas, USA, in 2012 established that nEMs could be encapsulated in integrated micro-electromechanical systems (MEMS), which include microelectronic controlling, sensing, diagnostic and processing integrated circuits.[7] Application-specific thrust impulses, thrust-vectoring and continuous thrust can be ensured by micro-thrusters and their arrays.

Military applications of swarming

New classes of extremely precise and lethal small or micro-scale weapon system are already in development. These systems have been scaled down by at least two orders of magnitude from current systems, creating space for the possible paradigm shift from bigger and fewer to smaller and numerous holdings of weapons. This advance heralds the era of swarm warfare—that is, assault by swarms of unmanned aerial vehicles (UAVs) 'made up of cooperative, autonomous robots that react to the battlefield as one'.[8] UAV swarms armed with nanoenergetic warheads as well as other nEM-integrated MEMS can be deployed in a multitude of missions such as strike, jamming, reconnaissance and saturation assault.

[3] Miziolek, A., 'Nanoenergetics: An emerging technology area of national importance', *AMPTIAC Quarterly*, vol. 6, no. 1 (spring 2002).

[4] Yarbrough, A., *The Impact of Nanotechnology Energetics on the Department of Defense by 2035*, Air War College Research Report (Air University: Maxwell Air Force Base, AL, 17 Feb. 2010).

[5] E.g. 'Thermal enhanced blast warhead', US Patent no. US 2012/0227613, 13 Sep. 2012.

[6] Rossi, C., 'Two decades of research on nano-energetic materials' (Editorial), *Propellants, Explosives, Pyrotechnics*, vol. 39, no. 3 (June 2014), pp. 323–27.

[7] Martirosyan, K., Hobosyan, M. and Lyshevski, S. E., 'Enabling nanoenergetic materials with integrated microelectronics and MEMS platforms', Proceedings of the IEEE Conference on Nanotechnology, 20–23 Aug. 2012; and Martirosyan, K. and Lyshevski, S. E., 'MEMS technology microthrusters and nanoenergetic materials for micropropulsion systems', Proceedings of the IEEE Conference on Methods and Systems of Navigation and Motion Control, 9–12 Oct. 2012.

[8] Scharre, P., 'How swarming will change warfare', *Bulletin of the Atomic Scientists*, vol. 74, no. 6 (2018).

As regards swarming, artificial intelligence (AI) has achieved demonstrable success. For example, one research group has developed 'swarm-enabling technology for multi-robot systems' that exhibits behaviours that include perimeter defence, aggregation, leader–follower, search and exploration, and heading consensus.[9] The technology was 'achieved by combining a modular and transferable software toolbox with a hardware suite composed of a collection of low-cost and off-the-shelf components' and is designed to be 'ported to a relatively vast range of robotic platforms—such as land and surface vehicles—with minimal changes and high levels of scalability'.[10] This low-budget, scalable approach makes swarm warfare accessible and adaptable to a number of military situations.

Swarm warfare could be as devastating and damaging as a nuclear weapon onslaught but without the radiation hazard and could displace tactical nuclear weapons from the battlefield. These smart, precise and lethal weapons have emerged in a grey zone between conventional and nuclear options. As such they pose a danger that needs debate and scrutiny.

Nanotechnology for improved sensors

Sensor technology is another area in which AI and nanotechnology are converging, with possible military applications. For example, 'theoretical and computational modelling already use algorithms to depict the ideal structure of a material, determine its energy and properties, and its interaction in different environments' so it is a 'natural progression' to enhance this modelling with AI.[11] Another example is scanning probe microscopy, used for imaging and measuring nanoscale surfaces at atomic height or to manipulate atoms and molecules, although it has traditionally suffered from resolution problems. AI in the form of advanced neural networks 'leads to a much more efficient imaging system'.[12]

Nanotechnology for low-yield nuclear options

There is an ongoing quest to develop very low-yield nuclear explosives that could be used as controlled micro-explosion sources for nuclear bombs as well as other weapons, if compact fusing mechanisms were available. It received a further impetus when it was found that it was more practical to design a micro-fusion explosive than a micro-fission device. Currently, this research forms the main thrust area at nuclear weapon laboratories in France and the USA.[13]

Nuclear weapon packages include fission or fusion material that is enriched in a sophisticated process but requires highly complex initiating components such

[9] Chamanbaz, M. et al., 'Swarm-enabling technology for multi-robot systems', *Frontiers in Robotics and AI*, vol. 4 (Apr. 2017), article 12.

[10] Chamanbaz et al. (note 9).

[11] Critchley, L., 'The convergence of AI and nanotechnology', *Nano*, 22 Aug. 2018.

[12] Critchley (note 11).

[13] Hambling, D., 'Darpa's handheld nuclear fusion reactor', *Wired*, 6 July 2009; and Badziak, J., 'Laser nuclear fusion: Current status, challenges and prospect', *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 60, no. 4 (Dec. 2012).

as arming and safety devices, and ancillaries for fusing and initiating a nuclear reaction. These should be controllable, safe and remain extremely reliable until the last possible moment of political decision-making (possibly with the incorporation of AI, even after launch and up to the instant before it reaches and hits the target). These critical criteria favour use of least failure, redundant devices incorporating nanotechnology and MEMS. The explosive train of a nuclear warhead contains an insensitive high explosive (IHE) that is initiated by a small sensitive initiator. These are kept misaligned before arming as a safety precaution and are aligned with the IHE using a nano- and microelectromechanical system (N/MEMS) device. There are many IHEs in a nuclear warhead that are brought into alignment by as many N/MEMS devices and individual detonators. These devices thus form a critical component of the safety and reliability chain in nuclear weapons. Nanotechnology is being increasingly used in better materials for capacitors, integrated circuits, high accelerations and temperature-resistant components, which together further enhance the possibility of greater safety and therefore new utility roles for nuclear weapons in the military.

The USA's Sandia National Laboratories has the credit for building the most complicated nuclear safety mechanism, the Micro Guardian, and its upgrades.[14] This mechanism ensures that the nuclear weapon does not detonate until a predefined sequence of events is complete. The availability of such devices and the fact that they have improved the resistance to failure of critical components in fusing, arming, detonators and neutron generators by many magnitudes, have spurred research into the next generation of fusion-based nuclear weapons.

These devices (IHE plus N/MEMS initiators plus tiny amounts of fission material) would not weigh more than a few kilograms, and the output could be equivalent to fractions of a tonne of TNT up to tens of tonnes. They use fission material in tiny quantities, thus resulting in negligible radioactive fallout. Such warheads are being considered for use in precision-guided munitions. Currently, there is no mechanism in place that restricts using nanotechnology to this end. The possessor of such weapons would be able to not only unleash a swarm of conventional weapons but also carry out a devastating assault without breaching the taboo of the first strike.

II. Research and development of emerging technologies for military applications in India

The above discussion shows how emerging technologies such as nanotechnology in tandem with AI are reshaping the landscape of conventional weapons and making them nearly as devastating as nuclear weapons, albeit without the ravaging onslaught of radiation hazards.

[14] Burroughs, C., 'Tiny "Micro Guardian" promises to safeguard nuclear weapons in big way', *Sandia Lab News*, vol. 51, no. 1 (15 Jan. 1999).

National investment in emerging technologies

The Indian Government appreciates that nanotechnology is 'a knowledge-intensive and "enabling technology" which is expected to influence a wide range of products and processes with far-reaching implications for national economy and development'.[15] Accordingly, a Mission on Nano Science and Technology (Nano Mission) was launched in May 2007, with the Department of Science and Technology as the nodal agency for its implementation.

India has also initiated development in the field of AI through the National Institution for Transforming India (NITI Aayog), which is a policy think tank formed on 1 January 2015 after the closure of the Planning Commission. The NITI Aayog's discussion paper on India's national strategy for AI recognizes that AI has the potential to be disruptive but that it also 'presents opportunities to complement and supplement human intelligence and enrich the way people live and work'.[16]

India also allocated 30.7 billion rupees ($462 million at current rates) in 2018 for its Digital India programme, which is an initiative to promote AI, machine learning, 3D printing and other digital technologies.[17]

Military research and development on AI

The strategic implications of AI from the perspective of national security was studied by an AI Task Force established by the Ministry of Defence and comprised of multiple stakeholders including the government, services, academia, industry, professionals and start-ups.[18] This AI Task Force looked at AI development in India generally but also specifically in the context of military needs. Among the recommendation in its final report were that (*a*) India should become 'a significant power of AI in defence' especially in 'aviation, naval, land systems, cyber, nuclear, and biological warfare', for both defensive and offensive needs, including counter-AI needs; (*b*) specific policy and institutional interventions are required to 'regulate and encourage . . . robust AI based technologies for [the] defence sector'; and (*c*) the government should work with start-ups and commercial industry on using AI 'for defence purposes'.[19] The task force also considered AI in relation to lethal autonomous weapon systems (LAWS) in the air, on the ground and underwater for both human-in-the-loop and human-on-the-loop scenarios; simulated war games and training (a key area where AI can play a crucial role in training

[15] Indian Department of Science and Technology, 'Nano Mission', [n.d.].

[16] National Institution for Transforming India (NITI Aayog), *National Strategy for Artificial Intelligence: #AIforall*, Discussion paper (NITI Aayog: New Delhi, June 2018), p. 5.

[17] Indian Ministry of Finance, Press Information Bureau, 'Highlights of Budget 2018–19', Press Release, 1 Feb. 2018.

[18] Indian Ministry of Defence, Press Information Bureau, 'AI task force hands over final report to RM', Press release, 30 June 2018.

[19] Indian Ministry of Defence (note 18).

the forces in a simulated environment); unmanned surveillance; cybersecurity; intelligence and reconnaissance; and aerospace security.[20]

The Centre for Artificial Intelligence and Robotics (CAIR) of the Defence Research and Development Organisation (DRDO) researches specific areas of AI for the Indian armed forces.[21] Its projects include (*a*) multipurpose robots 'including industrial grade capability robots and futuristic research oriented robotic platforms'; (*b*) a comprehensive data-mining toolbox containing data-mining algorithms, for use 'in different problem spaces'; (*c*) a decision support system (DSS) framework, which is 'completely driven by knowledge base maintained as ontologies' and includes algorithms like multi-criteria decision-making (MCDM), swarm algorithms, game-theoretic approaches to resource allocation, and search algorithms; (*d*) a semantically enabled service-oriented architectural framework; and (*e*) AI algorithms for path planning, simultaneous localization and mapping (SLAM), object detection and recognition, and task coordination for mobile platforms.[22]

Robotic sentinels

India faces two neighbours with unresolved border disputes and active insurgencies in many districts that have put its military and paramilitary forces under a veil of constant threat. There is thus a need for India to develop weapons to defend its vast land border, coastline and assets in space, with minimal risk to its forces. This could be achieved by using robotic sentinels that can respond effectively and neutralize the arising threats—a kind of LAWS. India has stressed that technology such as that being developed for LAWS has both peaceful and military uses.[23]

Robotic sentinels already in existence include the SGR-A1 robots and the Super aEgis II. The SGR-A1 robots, developed jointly by Samsung Techwin and Korea University, can automatically detect intruders walking over the border and could technically fire without the help of a human.[24] The Super aEgis II, developed by South Korean firm DoDaam, is an automated turret originally designed with an auto-firing system.[25] A robotic sentinel under development by Kalashnikov as one of 'a range of products based on neural networks' is a 'fully automated combat module' that can identify and shoot at its targets.[26]

[20] Pandit, R., 'India now wants artificial intelligence-based weapon systems', *Times of India*, 21 May 2018.

[21] See also chapter 3 in this volume.

[22] Defence Research and Development Organisation, 'Major products', [n.d.].

[23] Verma, D. B. V., Permanent Mission of India to the Conference on Disarmament, Statement at the CCW Informal Meeting of Experts on Lethal Autonomous Weapons Systems, Geneva, 17 Apr. 2015.

[24] 'Future tech? Autonomous killer robots are already here', NBC News, 15 Aug. 2011.

[25] Parkin, S., 'Killer robots: The soldiers that never sleep', BBC Future, 16 July 2015. See also Boulanin, V., 'The future of machine learning and autonomy in nuclear weapon systems', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, 2019), pp. 53–62, p. 60.

[26] Tucker, P., 'Russian weapons maker to build AI-directed guns', Defense One, 14 July 2017.

III. Conclusions

The above discussion highlights the aspects of nanotechnology and AI research and development that directly impinge on the war-waging capability of states. It shows that it will soon be feasible to deploy a conventional bomb with an nEM warhead or to engage in swarm warfare, either of which can cause devastating damage of the proportions of a primitive atomic bomb without the accompanying radiation. The collaborative nature of scientific studies and experimentation permits sharing of knowledge and thus leads to a wider proliferation of conventional weapon technologies. There is thus a need to study current technological developments in conventional weapons and to consider instituting international safeguards to cover developments that have the potential to cause massive damage.

Part II. The impact of military artificial intelligence on strategic stability in South Asia

The second part of this volume shifts the focus entirely to the South Asian context. The two authors address the following four questions: What are the specific risks associated with the use of artificial intelligence (AI), machine learning and autonomy in nuclear deterrence systems in South Asia? On which aspects of the strategic stability relations in the region are AI, machine learning and autonomy likely to have an impact? How might the military use of AI, machine learning and autonomy have an impact on the threshold for nuclear use? Finally, what kinds of crisis might be triggered by AI, machine learning and autonomy in this region?

In chapter 6 the present author bases his responses on analysis of the concept of strategic stability and AI-enhanced capabilities, which, when in use, might trigger escalatory scenarios among China, India and Pakistan. In chapter 7 Saima Aman Sial explores the positive and negative effects on instability and nuclear deterrence in South Asia of the development by Pakistan of pre-emptive doctrines and the deployment of a variety of offensive and defensive AI-enabled strategic systems.

PETR TOPYCHKANOV

6. Artificial intelligence and strategic stability in South Asia: New horses for an old wagon?

PETR TOPYCHKANOV

Strategic stability is an old concept that has found new interpretations in the context of the nuclear-armed states of South Asia. In this context, the future introduction of artificial intelligence (AI) into nuclear weapon systems could have both positive and negative impacts for strategic stability. But it can be questioned whether the new interpretations of strategic stability and the impacts of AI differ in substance from traditional concepts. That is, does the old wagon need new horses?

This essay examines the main characteristics of strategic stability in South Asia (in section II) by comparing it with the dynamics of this concept during the cold war (described in section I). This comparison is then used (in section III) to indicate the probable impacts of AI on the escalation ladder in South Asia, and the positive and negative consequences for strategic stability.

I. The concept of strategic stability in the cold war

The concept of strategic stability and the related principle of nuclear deterrence appeared during the cold war. In one of the first studies of strategic stability, John D. Steinbruner described it as a characteristic of deterrence based on mutually assured destruction.[1] During the cold war both the Soviet Union and the United States had growing nuclear capabilities and security concepts of nuclear weapon use, and both countries were ready for the consequences of the massive use of these weapons. For this reason, strategic stability was almost always assessed through the prism of the ratio of each side's strategic armaments, both offensive and defensive.

Strategic stability in the cold war era consisted of several critical elements.[2] First, acceptance of the idea of mutually assured destruction made the probability of nuclear war low. Despite their different mixes of strategic forces, the USA and the USSR implicitly accepted the idea of strategic parity. Second, both sides agreed to create a process to control the numbers of overtly offensive nuclear weapons in each other's arsenals and thereby prevent an unconstrained arms race. Third, development of redundant second-strike capabilities on both sides made it impossible for either side to realistically consider a first strike that would leave the adversary unable to respond to an attack. Fourth, a system of communications could be activated during confrontations and crises to prevent escalation into

[1] Steinbruner, J. D., 'National security and the concept of strategic stability', *Journal of Conflict Resolution*, vol. 22, no. 3 (Sep. 1978), pp. 411–28, p. 411.

[2] Russell, J. A., *Strategic Stability Reconsidered: Prospects for Escalation and Nuclear War in the Middle East*, Institut Français des Relations Internationales (IFRI) Proliferation Paper no. 26 (IFRI Security Studies Center: Paris, 2009), pp. 19–20.

a conflict. Fifth, confidence-building measures (CBMs) helped create a more cooperative political atmosphere. Sixth, both countries accepted that competition, conflict and rivalry could all co-exist in the interstate relationship.

However, changes in the political and military environment after the end of the cold war caused the concept of strategic stability to evolve, as reflected in two joint statements of 1990 and 1994. According to the Soviet–US statement of 1990, strategic stability is the status of strategic relations between the two powers in which there are no incentives for a first strike.[3] As a stabilizing principle, the statement placed emphasis 'on removing incentives for a nuclear first strike, on reducing the concentration of warheads on strategic delivery vehicles, and on giving priority to highly survivable systems'. Underpinning the statement was the mutual understanding that the purpose of a first strike was to prevent or significantly weaken the enemy's strike capabilities. The operational plans of the first counterforce strike were to achieve maximum destruction of the enemy's strategic forces, including its command, control and communications systems.

According to the joint Russian–US statement of 1994, Russia and the USA agreed to ensure an indefinite and unconditional extension of the 1968 Treaty on the Non-Proliferation of Nuclear Weapons (Non-Proliferation Treaty, NPT), a conclusion of negotiations for the Comprehensive Nuclear-Test Ban Treaty (CTBT) at the earliest possible date and a proposed global prohibition on the production of fissile materials for nuclear weapons (fissile materials cut-off treaty, FMCT).[4] The 1994 statement also related to mutual CBMs, the control, accounting and physical protection of nuclear materials, and cooperation in the development and fielding of effective theatre missile defence systems.[5]

These two official statements demonstrated the expansion of the concept of strategic stability beyond nuclear deterrence. Today this concept is further expanded to also include the proliferation of weapons of mass destruction (WMD) and related technologies; terrorism and nuclear terrorism; regional armed conflicts, with unpredictable escalation of military operations and expansion of conflict; drug trafficking; organized crime; and climate change and environmental threats.[6]

[3] Soviet–United States Joint Statement on Future Negotiations on Nuclear and Space Arms and Further Enhancing Strategic Stability, 1 June 1990, Washington, DC.

[4] Russian–United States Joint Statement on Strategic Stability and Nuclear Security, 29 Sep. 1994, Washington, DC; Treaty on the Non-Proliferation of Nuclear Weapons (Non-Proliferation Treaty, NPT), opened for signature 1 July 1968, entered into force 5 Mar. 1970, INFCIRC/140, 22 Apr. 1970; and Comprehensive Nuclear-Test-Ban Treaty (CTBT), opened for signature 24 Sep. 1996, not in force. Negotiation of an FMCT are still ongoing. See International Panel on Fissile Materials, 'Draft fissile material (cutoff) treaty, or FM(C)T', 5 Feb. 2009.

[5] Development of ballistic missile defence (BMD) was limited at that time by the Soviet–US Treaty on the Limitation of Anti-Ballistic Missile Systems (ABM Treaty), signed 26 May 1972, entered into force 3 Oct. 1972, not in force from 13 June 2002, *United Nations Treaty Series*, vol. 944 (1974), pp. 13–17.

[6] The most recent example of a wide interpretation of the concept of strategic stability is the Chinese–Russian Joint Statement on Strengthening Global Strategic Stability in the Modern Era, Moscow, 5 June 2019 (in Russian). See also Chinese Ministry of Foreign Affairs, 'Assistant foreign minister Zhang Jun publishes a signed article on jointly strengthening global strategic stability between China and Russia', 12 June 2019.

II. Strategic stability in South Asia

The peace, security and very survival of the South Asian subcontinent depend on the robustness of nuclear deterrence and strategic stability.[7] In relations between India and Pakistan, the outcomes of several dynamics will have a serious impact on strategic stability, including the state of conventional military and nuclear weapon capabilities, the arms race in both fields, and the impact of conventional military asymmetry on deterrence and strategic stability. Equally important are the management and resolution of India–Pakistan disputes over contentious issues of vital interest to each other, the state of political and diplomatic normalcy, adherence to security agreements and CBMs, and improvement of commercial, economic and cultural relations.[8]

Comparing key elements of strategic stability in contemporary South Asia with relations between the USA and the USSR in the cold war era, it is possible to find both differences and similarities.

Mutually assured destruction

The first common element is acceptance of the idea of mutually assured destruction, which makes the risk of nuclear war low, and which includes an implicit acceptance of the idea of strategic parity—despite the presence of a different mix of strategic forces. Two concepts are essential aspects of this element: credible minimum deterrence, whereby a state possesses the minimum number of nuclear weapons needed to deter an enemy from attacking; and no first use, which is a promise by a state not to use nuclear weapons unless first attacked by an enemy's nuclear weapons.

Minimum deterrence and no first use policies

Both India and Pakistan have declared that they will adhere to credible minimum deterrence policies. However, the meaning of the policy is different for each country, although both are linked. For India, the main goal is to prevent the use of WMD by any other state; for Pakistan, the goal is to prevent a critical war in which India uses WMD and conventional weapons against it.

In the case of India, minimum nuclear deterrence requires (*a*) sufficient, survivable and operationally prepared nuclear forces; (*b*) a robust command-and-control system; (*c*) capable intelligence and early-warning capabilities; (*d*) comprehensive planning and training for operations in line with the strategy; and (*e*) the will to employ nuclear forces and weapons.[9] A few issues are unclear in India's minimum deterrence policy. First is how the credibility of India's nuclear forces can be achieved without becoming a maximum deterrent. In attempting

[7] Cheema, Z. I., *Indian Nuclear Deterrence: Its Evolution, Development, and Implications for South Asian Security* (Oxford University Press: Karachi, 2010), p. 436.

[8] Cheema (note 7), p. 436.

[9] Rajain, A., *Nuclear Deterrence in Southern Asia: China, India and Pakistan* (Sage: New Delhi, 2005), p. 229.

to increase the credibility and effectiveness of its nuclear weapons as a deterrent, India's nuclear doctrine fails to limit itself to minimum deterrence. The second unclear issue is related to India's no-first-use obligation. Since India does not currently possess effective second-strike capabilities (e.g. submarine-launched ballistic missiles), and since it is actively developing its ballistic missile defence (BMD), many experts doubt that India adheres strictly to the no-first-use policy.[10]

In the case of Pakistan, its minimum deterrence policy is subject to changes in circumstances. In other words, this principle cannot be defined according to static numbers of nuclear weapons. Instead, deployment patterns change according to risks of pre-emption and interception.

Strategic parity

In the context of strategic parity, it is important to emphasize that both optimists and pessimists among nuclear experts agree that nuclear weapon build-up in South Asia will not lead to the deliberate outbreak of large-scale war in the region. Neither Indian nor Pakistani leaders wish to initiate a conflict that could end in a nuclear exchange with disastrous consequences. The difference between these two camps of South Asian experts lies in their assessment of the possibility that catastrophic conflict could occur even though neither state intends to start a nuclear war. The pessimists believe that nuclear exchange is likely to occur, especially if the repeated Indian–Pakistani confrontations and the still underdeveloped state of their nuclear control and early-warning systems are taken into account.[11] The optimists argue that nuclear disaster in South Asia remains highly unlikely, in particular through the practice of lowered combat preparedness during peacetime—that is, an 'operationally dormant' state of nuclear arsenals.[12]

Agreements on processes to prevent arms race

Another element of cold war strategic stability is agreement between the states to create a process for controlling the numbers of overtly offensive nuclear weapons in each other's arsenal and thereby prevent an unconstrained arms race. However, there is no such agreement between India and Pakistan. Neither country is interested in having another country involved in controlling its nuclear arsenals. The reasoning of both states, according to Indian and Pakistani government officials, are similar: first, the capability of either state to build nuclear weapons is more or less clear to the other state; and second, India and Pakistan both adhere

[10] Khalid, I., 'Nuclear doctrine: Ramifications for South Asia', *South Asian Studies*, vol. 27, no. 2 (July–Dec. 2012), p. 319.

[11] Ganguly, Š. and Kapur, S. P., *India, Pakistan, and the Bomb: Debating Nuclear Stability in South Asia* (Columbia University Press: New Delhi, 2010), p. 85.

[12] Chufrin, G. et al., 'South Asia', eds A. Arbatov and V. Dvorkin, *Nuclear Weapons after the Cold War* (Carnegie Moscow Center: Moscow, 2008), p. 336.

to minimum nuclear deterrence policies, and so are not interested in any nuclear competition or arms race.[13]

Regarding nuclear control, the interests of India and Pakistan also diverge somewhat. India is more concerned about controlling the nuclear arsenal of China than that of Pakistan (although China shows no interest in exchanging data on nuclear weapons with India). Thus, while Pakistan would like to enter into an agreement with India on nuclear arms control, India is unlikely to seek such an agreement with Pakistan.

Agreements on confidence-building measures

The situation is slightly better in the area of CBMs, with communications systems that could be activated during confrontations and crises to prevent escalation in the event of a conflict. India and Pakistan have a number of such agreements: (*a*) a 1988 agreement not to attack each other's nuclear facilities, which requires an exchange of lists of their respective nuclear installations on 1 January every year; (*b*) a 2005 agreement to shift the hotline for the directors general of military operations to a fibre-optic link; (*c*) a 2005 missile test pre-notification agreement; and (*d*) a 2007 agreement to reduce the risk of accidents and unauthorized use of nuclear weapons. It is crucial to emphasize that none of these agreements has any verification mechanism.

There was a window of opportunity to develop CBMs during the 2004–2008 Composite Dialogue between India and Pakistan. However, after the Mumbai terrorist attack in 2008, this dialogue was frozen by India, which accused Pakistan of supporting terrorists and using them against India.

III. The arrival of AI in regional escalatory dynamics

The nuclear-armed states of South Asia and its neighbourhood—China, India and Pakistan—are interested in developing AI technologies for military purposes.[14] With various budgets and scales of research and development (R&D), these three states are each exploring military applications of AI in areas of strategic significance, including command-and-control, early-warning, BMD, and intelligence, surveillance and reconnaissance (ISR) systems; unmanned underwater vehicles (UUVs) and unmanned aerial vehicles (UAVs); and electronic warfare and cyberwarfare.

At this stage, an armed conflict based on AI use is unlikely between either India and China or India and Pakistan because the level of AI integration into military solutions is still immature. But given the tensions in the regions of Kashmir, the

[13] Indian officials, Interviews with author, New Delhi, 5 Oct. 2011; and Pakistani officials, Interviews with author, Islamabad, 3 Oct. 2011.

[14] On India see chapters 3 and 5 in this volume; on Pakistan see chapter 7 in this volume; and on China see Saalman, L., 'Exploring artificial intelligence and unmanned platforms in China', ed. L. Saalman, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. II, *East Asian Perspectives* (SIPRI: Stockholm, Oct. 2019), pp. 43–47.

Himalayas and the Indian Ocean, and also the space competition between India and China, such armed conflicts cannot be excluded in the future.

The possible scenarios of such conflicts can be seen as steps up the escalation ladder.[15] First is disagreement and rivalry, where the sides are preoccupied with pre-crisis manoeuvring and military signalling. Next is 'hot' warfare, including covert operations or surgical strikes. The next step is conventional war, with force mobilization and a precautionary nuclear alert. This is followed by major conventional war, with complete mobilization, conventional attacks and air-strikes against military targets, and a full nuclear alert. Finally comes nuclear weapon use, including various nuclear options from demonstrative or signalling use to all-out countervalue strikes. However, nuclear escalation might not necessarily follow these steps in South Asia. For instance, in previous crises the nuclear signalling could have happened at the hot warfare stage or even prior to it.[16] Some experts are concerned about the potential use of nuclear weapons, mainly tactical weapons, at the initial stages of a possible conflict between India and Pakistan.[17]

Today, it is hard to see a central role for military AI in the majority of these scenarios since the introduction of AI into strategic weapons by China, India and Pakistan is in its early stages. Autonomous systems offer a supporting capability, rather than playing a decisive role. Nevertheless, it is worth considering the role of AI in this escalation ladder as it becomes more and more probable while China, India and Pakistan are obtaining early-warning and BMD systems, advanced missile technologies, including hypersonic weapons, and potentially dual-capable combat UAVs and UUVs. In some cases it might take decades to obtain these capabilities, but in other cases it might be sooner. For instance, China is building space-based early warning with Russian support, and India is following the same path.[18]

Military AI potentially has specific roles in each of the steps of the escalation ladder described above. During the period of disagreement and rivalry, autonomous ISR systems would help decision makers to assess the actions of the adversary and the operational deployment of dual-capable offensive weapons, including autonomous ones, would be an instrument of military signalling. During the hot warfare stage, autonomous ISR would support covert operations and combat UAVs would play a key role in offensive actions. In the conventional and major conventional war stages, the functions of ISR and offensive capabilities remain the same. At the same time, at these stages the conflicting sides would rely more significantly on nuclear early warning and strategic command and control. Finally, early warning, BMD, nuclear command and control, and autonomous components of the nuclear arsenal would be fully enacted during nuclear weapon

[15] Jones, R., *Nuclear Escalation Ladders in South Asia* (US Defense Threat Reduction Agency: Ft. Belvoir, VA, Apr. 2011), pp. 14, 17, 23.

[16] Yusuf, M., 'Banking on an outsider: Implications for escalation control in South Asia', *Arms Control Today*, vol. 41, no. 5 (June 2011).

[17] Hooey, D., 'Pakistan's low yield in the field: Diligent deterrence or de-escalation debacle', *Joint Force Quarterly*, vol. 95, no. 4 (2019), pp. 34–45, pp. 40–41.

[18] President of Russia, 'Valdai Discussion Club session', 3 Oct. 2019; and Sputnik, 'India launches military surveillance satellite to track enemy radar', Space Daily, 2 Apr. 2019.

use. However, the level of the ongoing R&D in China, India and Pakistan means that there can be no decisive role for fully automated strategic command and control in the foreseeable future.

Two aspects of military AI should be considered as more immediate risks to strategic stability in South Asia. First, given immature development of autonomous ISR, early-warning and BMD capabilities, there is a high risk of false alarms from these systems. Faulty reports from a country's early-warning, ISR and BMD sensors, in close proximity to an adversary, might be considered to be valid and lead to a response by preventive or pre-emptive uses of nuclear forces. Second, the development of dual-capable autonomous platforms by China, India and Pakistan may provoke one of the countries to fear a surprise nuclear attack if one or both of the others were to deploy such a platform.

While the advent of military AI has a destabilizing potential, it also has the potential to diminish nuclear risk. For instance, AI-enabled satellite imagery and remote sensing may help China, India and Pakistan to interpret each other's actions correctly. These technologies may help to prevent inadvertent escalations via a cooperative aerial monitoring agreement like the 1992 Treaty on Open Skies in Europe and North America.[19]

IV. Conclusions

At least two critical points arise from the discussion above.

First, the nuclear weapon policy of any nuclear-armed state should be based on political goals rather than technological advances. That is, a state should have constant goals, based on a cross-party consensus, and the integration of new technologies into weapons and military equipment should comply with these goals. India and Pakistan lack such clarity about their nuclear doctrines. These countries plus China lack the transparency and undisrupted mutual communication between them that may help them to avoid armed conflicts.

Second, given the absence of CBMs or arms control between India, Pakistan and China in the nuclear and conventional areas, the advent of AI into the military domain may create an even less predictable and stable situation for South Asia.

[19] Wise, J., *Satellite Imagery, Remote Sensing, and Diminishing the Risk of Nuclear War in South Asia*, Special Report no. 434 (United States Institute of Peace: Washington, DC, Nov. 2018), p. 9; and Treaty on Open Skies, opened for signature 24 Mar. 1992, entered into force 1 Jan. 2002.

7. Military applications of artificial intelligence in Pakistan and the impact on strategic stability in South Asia

SAIMA AMAN SIAL

Soon after the advent of nuclear weapons, which altered the dynamics of war and maintenance of an uneasy peace, Henry Kissinger wrote:

> In Greek mythology, the gods sometimes punished man by fulfilling his wishes too completely. It has remained for the nuclear age to experience the full irony of this penalty. Throughout history, humanity has suffered from a shortage of power and has concentrated immense effort on developing new sources and special applications of it. It would have seemed unbelievable even fifty years ago that there could ever be an excess of power that everything would depend on the ability to use it subtly and with discrimination.[1]

In an age of rapidly advancing technology, the challenge still remains how to use emerging technologies like artificial intelligence (AI), machine learning and autonomy, both 'subtly and with discrimination', such that their use serves security and peace rather than undermines it.

There is hardly any published or publicly available literature on the subject of the application AI in nuclear weapon systems in nuclear-armed states, and the case of Pakistan is no different. The available literature explores the possible military applications of AI and machine learning, such as for early warning, target acquisition, remote sensing and cyber forensics, and as a decision-making aide. Focusing on integration of AI in nuclear weapon systems, the literature shows that during the cold war the Soviet Union and the United States showed great interest in AI as a tool for making the decision-making process 'more agile' and for providing decision makers more time to consider their responses.[2] However, both states were equally cognizant of the importance of not assigning the 'higher-order assessments and launch decisions' to AI systems.[3] The USSR did develop a fully automated command-and-control system, Perimetr (or Dead Hand), but it was only to be activated in the exceptional circumstance of a decapitation strike.[4]

This essay seeks to fill the gap in the literature for the case of Pakistan. It continues (in section I) how Pakistan could integrate AI into its nuclear command-and-control system. It then (in section II) considers the implications of military AI for strategic stability in South Asia.

[2] Boulanin, V., 'AI and nuclear weapons—promise and perils of AI for nuclear stability', AI & Global Governance, United Nations University, Centre for Policy Research, 7 Dec. 2018.

[3] Boulanin (note 2).

[4] Borrie, J., 'Cold war lessons for automation in nuclear weapon systems' and Topychkanov, P., 'Autonomy in Russian nuclear forces', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, May 2019), pp. 41–52 and pp. 68–75.

I. AI and Pakistan's nuclear command-and-control system

Pakistan's nuclear command-and-control system is assertive.[5] That is, the prime minister is the chair of the National Command Authority (NCA) and the Strategic Plans Division (SPD) serves as its secretariat. Moreover, to ensure assertive control over the nuclear weapons, the commanders of Pakistan's strategic forces, which have custody of the delivery systems, do not have custody of warheads, which are centrally controlled by the SPD.[6]

Some scholars place more emphasis on the control of the nuclear weapons (i.e. the prevention of their use) than on command (which relates to their actual use). Thus, control should take precedence over command, as the loss of control could inadvertently lead to a nuclear war.[7] Pakistan's nuclear command-and-control system follows this principle closely. Owing to the scrutiny of Pakistan's political security following the terrorist attacks on the USA of 11 September 2001, its nuclear command-and-control system has developed a security culture that takes control seriously. Moreover, geographical separation of the warheads from the delivery vehicles also acts as a control against accidental or unintended use.[8]

In nuclear command-and-control systems there is a constant tension between the weapons always being available when required by a decision maker and never being available for unauthorized use, popularly known as the 'always–never dilemma'.[9] Pakistan balances this tension through 'a robust strategic command and control apparatus designed to ensure tight negative use control during peacetime and prompt operational readiness (positive control) at times of crisis and war'.[10]

Pakistan has a nuclear posture of recessed deterrence as its nuclear forces are not deployed. However, the recent decision to develop nuclear weapons for sea deployment will change the posture to a ready arsenal, as separation of nuclear warhead and delivery system will no longer be technically feasible.[11] Bringing warheads and delivery systems together reduces launch times and increases readiness levels.

The geographical contiguity of India and Pakistan creates serious challenges for the survivability of nuclear forces because proximity makes them harder to hide and easier to strike with nuclear and conventional means. In contrast to the long

[5] Ahmed, M., 'Pakistan's tactical nuclear weapons and their impact on stability', Regional Voices on the Challenges of Nuclear Deterrence Stability in Southern Asia, Carnegie Endowment for International Peace, 30 June 2016.

[6] Salik, N., *Learning to Live with the Bomb, Pakistan: 1998–2016* (Oxford University Press: Karachi, 2017), p. 164.

[7] See e.g. the comments of Rajesh Basrur quoted in Salik (note 6), p. 139.

[8] Basrur (note 7), p. 140.

[9] Feaver, P. D., 'Command and control in emerging nuclear nations', *International Security*, vol. 17, no. 3 (winter, 1992/93), p. 163.

[10] Lavoy, P. R., 'Pakistan's nuclear posture: Security and survivability', Nonproliferation Policy Education Center, 21 June 2007.

[11] Pakistani Inter Services Public Relations, 'Pakistan conducted another successful test fire of indigenously developed submarine launched cruise missile Babur having a range of 450 kms', Press Release no. PR-125/2018-ISPR, 29 Mar. 2019.

missile flight times—amounting to several minutes—in the Russian–US context, the flight times of missiles in South Asia are extremely short. Pakistan's fear of a pre-emptive first strike from India and possible developments in this regard mean that it sees a need for technologies that can enhance the capability of its early-warning systems and assist in compressing the time frame for decision-making.

Much like other nuclear-armed states, Pakistan is unlikely to give complete autonomy to nuclear platforms. However, AI and machine learning can assist in developing better situational awareness about strategic assets during peacetime as well as in crisis. For example, machine learning can have applications in such nuclear-related capabilities as automatic target recognition (ATR), early warning and decision support systems. In this regard, Pakistan has

> established a National Command Center (NCC), which has a fully automated Strategic Command and Control Support System (SCCSS) that enables the decision makers at the NCC to have round the clock situational awareness of all strategic assets during peacetime and especially in times of crisis. As per the official statements, all deployments/ employments would be centrally monitored and controlled by the NCC.[12]

Furthermore, AI systems are already in use at strategic organizations for security purposes, such as for access control using thumb impressions, digital retina scans and facial recognition software to restrict access to sensitive high-security installations.[13]

II. AI benefits and perils for strategic stability in South Asia

Implications of military applications of AI

Any technology is neither good nor bad in itself. Its particular application in nuclear weapon systems determines the effect that it generates; that is, whether it strengthens or undermines deterrence. An example is the use of AI for war games. In nuclear strategizing, all propositions about how escalation may be played out in a crisis or war are untested, since no nuclear war has yet been fought. Therefore, feeding a data set into an AI system of how a nuclear war or crisis may play out would be difficult. However, AI has many more military applications than war games.

AI is becoming more relevant for each component of complex deterrence architectures, especially given its potential to significantly enhance intelligence, surveillance and reconnaissance (ISR) capabilities.[14] The integration of AI into nuclear weapon systems could enhance the accuracy and reaction times, and hence overall performance, of strategic offensive and defensive systems. The impact on strategic stability may be either positive or negative.

[12] Ahmed (note 5).

[13] The present author has visited the SPD several times and observed the use of different applications for access control.

[14] Gasser, P., Loss, R. and Reddie, A., 'Assessing the strategic impact of artificial intelligence', Workshop summary, Center for Global Security Research, 20–21 Sep. 2018, p. 3.

An instability-inducing impact of AI could be in the field of precision and counterforce targeting. Advances in precision targeting, along with enhanced ISR capabilities in the adversary's arsenal, could create challenges for a secure second strike and increase fears of a pre-emptive strike. Even an increased ability to target strategic assets with conventional precision-strike capabilities, enhanced through AI, would heighten instability, especially in crisis situations.[15] Another example is autonomy applied to unmanned underwater vehicles (UUVs) to be used in antisubmarine warfare; submarines would no longer be the ultimate deterrents guaranteeing state survival.[16] In South Asian deterrence this proposition may still be in the future, but given the pace of technology transfers to India by the leading arms exporters, it is a threat to watch out for.[17]

Cyberwarfare is another domain where AI could affect strategic stability. Machine learning could help to improve the performance of early-warning radars for signals processing and could be used for situational awareness in space. AI could help cyber-defenders monitor intrusions and detect anomalies. These improvements in cybersecurity to protect critical command, control and communications systems could potentially strengthen regional and strategic stability. However, the same cyber means could be used to inject misinformation into the system using 'deep fakes' through social media platforms as well.[18] This can have serious crisis initiation and escalation implications for a region such as South Asia. The crisis after the attack on Pulwama in Indian-administered Kashmir on 14 February 2019 demonstrated how information manipulation can lead to escalation.[19]

Implications for nuclear force postures in South Asia

The challenges raised by applications of AI to nuclear systems, coupled with compressed decision-making time, would have implications for nuclear force postures taken by Pakistan and its adversary, India. Stable mutual deterrence requires opposing nuclear powers to have credible, survivable nuclear forces. Until now, an assured second-strike capability has been considered to be a means of ensuring stability of deterrence. Pakistan uses concealment and mobility as a means of survivability of its nuclear arsenal.[20] For Pakistan, the integration of

[15] Sial, S. A., 'To use or not to use: India's fractured NFU', *South Asian Voices*, 20 Mar. 2017.

[16] See e.g. Rickli, J.-M., 'The destabilizing prospects of artificial intelligence for nuclear strategy, deterrence and stability', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, May 2019), pp. 91–98, p. 94; and chapter 2 in this volume.

[17] Wezeman, S. T. et al., 'Developments in arms transfers, 2017', *SIPRI Yearbook 2018: Armaments, Disarmament and International Security* (Oxford University Press: Oxford, 2018), pp. 204–206.

[18] Chesney, R. and Citron, D., 'Deepfakes and the new disinformation war: The coming age of post-truth geopolitics', *Foreign Affairs*, Jan./Feb. 2019. See also Avin, S. and Amadae, S. M., 'Autonomy and machine learning at the interface of nuclear weapons, computers and people', ed. Boulanin (note 16), pp. 105–18.

[19] Yusuf, M. W., 'The Pulwama crisis: Flirting with war in a nuclear environment', *Arms Control Today*, vol. 49, no. 4 (May 2019).

[20] Sial, S. A., 'Rationalizing Pakistan's quest for a sea-based deterrent force', *Pakistan Politico*, Oct. 2018, pp. 30–31.

AI into India's nuclear weapon systems would make maintaining an adequate second-strike force harder. Furthermore, it would in turn lead Pakistan to invest in either robust defensive systems or increased mobile launch systems that can avoid detection.

An overall assessment of the nature of the strategic stability in South Asia shows that India's doctrine towards counterforce and pre-emption has shifted. Recent developments in the Indian arsenal complement the overall trends. India is investing in developing and deploying integrated multilayered ballistic missile defence (BMD) systems. In 2019 India approved a deal worth US$5.2 billion for five regiments of the Russian-made S-400 Triumf air defence system.[21] It has two domestic programmes underway for low- and high-altitude ballistic missile interception: the Advance Air Defence and Prithvi Defence Vehicle, respectively.[22] On 27 March 2019 India tested anti-satellite (ASAT) weapon technology targeting one of its own satellites in lower earth orbit.[23] This capability means that in future India will be able to target an enemy's military satellites as well as long-range strategic missiles. The Indian space agency, the Indian Space Research Organisation (ISRO), has launched 104 satellites from a single rocket, thus also demonstrating a capability to carry several different warheads on a ballistic missile (multiple independently targetable re-entry vehicles, MIRVs).[24]

India is currently enhancing its strategic and conventional precision-strike weapons for strikes deep within an adversary's territory for offensive pre-emptive counterforce. Examples include development of cruise missiles for its land, sea and air forces, deployment of cruise missiles such as the BrahMos on Su-30 aircraft, and extending the range of the BrahMos and Nirbhay cruise missiles.[25]

These recent developments in the Indian arsenal create new pressures for Pakistan and so for strategic stability. Consider a scenario in which a Pakistani military satellite is targeted by India using an ASAT weapon, leading to a crisis situation. The Indian ASAT test demonstrated the capability to target Pakistan's long-range nuclear missiles as well as its military satellites. Combine this with the advanced ATR capability and enhanced ISR, and the challenge to deterrence stability becomes more acute. This could lead the South Asian nuclear postures to shift from deterrence to readiness, reducing the amount of time needed for a decision to make the first strike, in turn leading to crisis instability.

A caveat here is the discrepancy between perception of an adversary's capability and reality. That is, postures in adversarial nuclear relationships may well be shaped by the perception of an adversary's integration of AI rather than knowledge

[21] 'Russia, India sign $5bn deal for S-400 air defence system', *Dawn*, 5 Oct. 2018.

[22] 'India successfully tests interceptor missile', *Dawn*, 1 Mar. 2017.

[23] Agence France-Presse, 'India claims to shoot down satellite, join "space superpowers"', *Dawn*, 27 Mar. 2019.

[24] Reuters, 'India launches record 104 satellites in one go', *Dawn*, 15 Feb. 2017.

[25] Clary, C. and Narang, V., 'India's counterforce temptations: Strategic dilemmas, doctrine and capabilities', *International Security*, vol. 43, no. 3 (Winter, 2018/19), pp. 25–31.

of the actual level of integration.[26] This issue is further compounded by the secrecy surrounding the level of integration of AI in nuclear-related systems.

It is unlikely that AI would be accorded to role of making the decision to launch in Pakistan. The political leadership would retain the authority to make the final launch decision, despite the shorter reaction times.

III. Conclusions

The use of AI, machine learning and autonomy to enhance capabilities of precision munitions, BMD, nuclear submarines, unmanned vehicles, airborne warning and control systems, and ASAT weapons, coupled with dangerous postures such as pre-emptive counterforce, will only increase strategic instability and crisis risks in South Asia. It is too early to make definitive judgements about the long-term effects of AI on strategic stability in the region; nonetheless, whatever the level of advances in technology, human control in strategic decision-making and command-and-control systems should never be relinquished.

[26] See e.g. Saalman, L., 'The impact of artificial intelligence on nuclear asymmetry and signalling in East Asia', ed. L. Saalman, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. II, *East Asian Perspectives* (SIPRI: Stockholm, Oct. 2019), pp. 103–108.

Part III. Arms control and confidence-building measures in the area of artificial intelligence and nuclear weapons

The two contributors to this third part of the volume offer their views on the arms control and confidence-building measures (CBMs) that are necessary to address the challenges of introducing artificial intelligence (AI) into nuclear weapon systems.

They answer two principal questions: How can earlier lessons learned on arms control and CBMs be applied to design a comprehensive response to destabilizing trends associated with AI, machine learning and autonomy? What could be recommended for the South Asian context?

The two chapters complement each other, offering different approaches to prevention and mitigation of nuclear risk arising from military AI. Yanitra Kumaraguru in chapter 8 argues for the need to introduce a multilateral ban on the development, production and use of lethal autonomous weapon systems (LAWS). Based on the experiences of the Soviet Union or Russia and the United States, Malinda Meegoda in chapter 9 explores various bilateral and trilateral formats that are feasible for South Asia.

PETR TOPYCHKANOV

8. A pre-emptive ban on lethal autonomous weapon systems

YANITRA KUMARAGURU

The development of artificial intelligence (AI) and robotics continues in both civilian and military contexts. In the latter arena, several weapons already possess elements of automation. An automated weapon must, however, be distinguished from an autonomous weapon system.[1] Unlike automated weapons, autonomous weapon systems do not retain meaningful human control in their operation and remove humans from a meaningful place in the decision-making loop.[2] The scope of this essay is limited to the latter—specifically lethal autonomous weapon systems (LAWS). It provides (in section I) a non-exhaustive discussion on the problems inherent in the operation of LAWS in order to show (in section II) the need for a ban on their development, production and use. It echoes the sentiment expressed by companies working in AI and robotics in their open letter to the parties to the 1981 Convention on Certain Conventional Weapons (CCW Convention): 'We do not have long to act. Once this Pandora's box is opened, it will be hard to close.'[3]

I. Problems inherent to lethal autonomous weapon systems

Crossing ethical and moral lines

In delegating the decision to kill to a machine, LAWS cross ethical and moral lines. Machines programmed to make decisions according to the computations of algorithms lack moral agency to take decisions concerning human life and death. Despite being subject to numerous groundbreaking technological developments, LAWS still lack the distinctly human qualities of empathy, reason, intellect and compassion that are necessary to exercise moral and prudential judgement on the battlefield. Weapons that select and engage targets without meaningful human involvement also undermine human dignity. Machines are unable to appreciate the value of human life and therefore cannot feel the gravity of its loss.[4] They also lack the innate reluctance and internal struggle that humans are forced to confront in deciding to harm or kill another human being.[5]

The notion of delegating crucial components of the targeting cycle to a weapon further distances humans from the gravity of the decision to exercise force and engage in violence. Entrusting these decisions to a machine and not a human removes the moral burden that would otherwise immediately precede and follow

[4] Human Rights Watch (HRW) and International Human Rights Clinic (IHRC), *Making the Case: The Dangers of Killer Robots and the Need for a Pre-emptive Ban* (HRW: New York, 9 Dec. 2016), pp. 22–23.

[5] Human Rights Watch and International Human Rights Clinic (note 4), pp. 25–26.

such a decision, thereby carrying with it the potential to reduce the threshold for violence and warfare.[6]

Non-compliance with humanitarian and human rights laws

In removing meaningful human control from critical junctures in the weapon targeting cycle, such as selection and engagement, those who deploy LAWS fail to abide by international humanitarian law and human rights law. The ability of LAWS to comply with these laws must also be considered. Programmed to compute their choices, LAWS will fail in attempts to comply with the key principles of distinction and proportionality on the battlefield when encountering several of the infinite combinations of possibilities that humans, even with their more nuanced understanding of the concepts, grapple with.[7] International humanitarian law continues, for example, to struggle with streamlining determinations of whether someone is directly participating in hostilities and thus deprived of protection as a civilian. These judgements are not those that can be left to a machine and its algorithms with no meaningful human control. Consider the example of a child who picks up a weapon lying on the edge of a road, or that of a person stepping in and out of direct participation in hostilities. Use of LAWS in warfare would also undermine the precautionary principle in international humanitarian law.[8]

There is also the need to consider the potential use of LAWS in situations of law enforcement (e.g. riot control and counterterrorism efforts) that have not yet crossed the threshold of armed conflict under international humanitarian law. The use of LAWS in these contexts poses grave human rights concerns in the form of arbitrary killings in violation of the right to life, an impediment to the right to a remedy and the undermining of human dignity.[9]

Human rights and humanitarian concerns are further aggravated by the fact that the algorithms controlling LAWS and the data relied on are informed and influenced by the prejudices, biases and perceptions of their human designers.[10] Thus, human prejudices and perceptions concerning but not limited to race, religion and gender may not only be ingrained but also exacerbated in the operation of LAWS.[11]

[6] Kumaraguru, Y., 'Unimpeded development of science and technology at what cost?', *CCW Report*, vol. 6, no. 11 (4 Sep. 2018).

[7] On the ability of autonomous weapons to conform to the principles of distinction and proportionality see Human Rights Watch and International Human Rights Clinic (note 4), p. 5 and p. 8.

[8] International Panel on the Regulation of Autonomous Weapons (iPRAW), *Focus on Computational Methods in the Context of LAWS*, 'Focus on' Report no. 2 (iPRAW: Berlin, Nov. 2017).

[9] Human Rights Watch and International Human Rights Clinic (note 4), pp. 18–20.

[10] Kumaraguru (note 6); and Buranyi, S., 'Rise of the racist robots: How AI is learning all our worst impulses', *The Guardian*, 8 Aug. 2017.

[11] On such biases see Saalman, L., 'The impact of artificial intelligence on nuclear asymmetry and signalling in East Asia', ed. L. Saalman, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. II, *East Asian Perspectives* (SIPRI: Stockholm, Oct. 2019), pp. 103–108.

Lack of legal accountability

The complexity of programming and deploying LAWS stumbles yet again in relation to securing legal accountability for violations of both human rights law and international humanitarian law. As autonomy increases, machines become able to carry out critical functions of the targeting cycle without meaningful human intervention. This in turn makes it harder to establish intention in any one of several instances of 'decisions' and to hold a perpetrator to account, given the degree of independence exercised by the weapon.[12] This is compounded by the unpredictability of the machine's conduct stemming from either the machine's design and degree of independence or its interactions with the environment.[13]

Since the machine itself cannot be held responsible, despite its autonomous nature, the question of whom the wrongful conduct must be attributed to is a matter of contention. Would responsibility lie with the manufacturers of the weapon? Its designers? The programmers? The commanding military officer? Or the officer who deployed the weapon?[14]

Fallibility and vulnerability

Autonomous weapons are not infallible. Despite their complexity, computational systems used to control LAWS are not devoid of limitations and points of failure. When these systems are joined in a sequential manner to bring the autonomous weapon into operation, an error in one step of the sequence or single computational method could rebound and snowball throughout the sequence, bringing with it devastating consequences.[15] It must be remembered that such weapons could also reach the wrong hands or cause catastrophic impacts if subject to cyberattack.

II. A ban on the development, production and use of autonomous weapons

For the reasons outlined above, it is clear that LAWS should not be developed, produced or stockpiled, nor their use threatened, either within or outside situations of armed conflict. A pre-emptive ban on the development, production and use of LAWS remains the most effective path forward in this regard and has a precedent

[12] Davison, N. and Giacca, G., 'Background paper prepared by the International Committee of the Red Cross', International Committee of the Red Cross (ICRC), *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, Expert meeting, Versoix, Switzerland, 15–16 Mar. 2016 (ICRC: Geneva, Aug. 2016), pp. 81–82.

[13] Goussac, N., 'Safety net or tangled web: Legal reviews of AI in weapons and war-fighting', Humanitarian Law and Policy, International Committee of the Red Cross, 18 Apr. 2019.

[14] United Nations, Human Rights Council, Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, Christof Heyns, A/HRC/23/47, 9 Apr. 2013, para 77.

[15] International Panel on the Regulation of Autonomous Weapons (note 8), p. 19.

in the 1995 protocol that pre-emptively banned blinding laser weapons.[16] While discussions on the issue of autonomous weapons have been underway since 2014 within the seemingly appropriate forum of the CCW Convention, states parties to the convention must urgently move towards a negotiating mandate if the issue is to be acted on in an effective manner.[17] However, because concrete progress may often be impeded by a handful of states because of the consensus rule within the CCW framework, the process might have to be taken outside the CCW regime to formulate an effective ban that meets the urgency of the situation.[18] Such a ban would ensure that all weapon systems that are designed and manufactured with elements of automation still retain meaningful human control over critical functions in warfare.[19]

Mandatory retention of meaningful human control over critical functions such as target selection and engagement would not just prohibit the operation of weapon systems with no human intervention but would also prohibit the operation of weapon systems where humans provide only nominal approval to proceed, in response to the machine's suggestions, without a sufficient understanding of the context.[20]

Meaningful human control would instead include, but not be limited to, the human operator making the necessary legal and ethical assessments in the context of having a clear understanding of the situation, access to necessary information and sufficient time to make decisions. This control should enable effective oversight of the mission, with the human operator able to interact with and monitor the weapon system as well as monitor the relevant environment. It is essential that the human operator is able to intervene immediately before the use of force by the machine. The human control exercised should also be sufficient to secure accountability for decisions made and actions taken, and the consequences of both.[21]

Tailored in this manner, the proposed ban is not overly broad. It acknowledges the role of robotics and AI in a military context, as long as that role is within the confines of morality and legality and allows for accountability. But it does not in any way hamper the development of technology for civilian or peaceful purposes.

[16] Additional Protocol to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects (Protocol IV entitled Protocol on Blinding Laser Weapons), opened for signature 13 Oct. 1995, entered into force 30 July 1998, *United Nations Treaty Series*, vol. 1380 (1984).

[17] Human Rights Watch (HRW), Campaign to Stop Killer Robots, Statement to the Convention on Conventional Weapons Group of Governmental Experts on Lethal Autonomous Weapons Systems, Geneva, 27 Mar. 2019.

[18] Klare, M., 'US, Russia impede steps to ban "killer robots"', *Arms Control Today*, vol. 48, no. 8 (Oct. 2018).

[19] On the verification of such a ban see Gubrud, M., 'Can an autonomous weapons ban be verified?', International Committee for Robot Arms Control, 14 Apr. 2014.

[20] Moyes, R., 'Meaningful human control over individual attacks', International Committee of the Red Cross (note 12), pp. 46–52.

[21] International Panel on the Regulation of Autonomous Weapons (note 8); International Panel on the Regulation of Autonomous Weapons (iPRAW), *Focus on the Human-Machine Relation in LAWS*, 'Focus on' Report no. 3 (iPRAW: Berlin, Mar. 2018); and PAX, *Killer Robots: What Are They and What Are the Concerns?* (PAX: Utrecht, [n.d.]).

9. Autonomous weapons in the South Asian context: Risks and countermeasures

MALINDA MEEGODA

At the end of the lecture 'Can digital computers think?', Alan Turing remarked that 'the attempt to make a thinking machine will help us greatly in finding out how we think ourselves'.[1] This assertion has most probably guided developers in the realm of artificial intelligence (AI) with much-publicized events starting in the 1990s with IBM's chess-playing supercomputer Deep Blue, and up to the AlphaGo software program of Google's Deep Mind project.[2] However, it remains to be seen if AlphaGo's victory over Lee Sedol, a master player of the game go, actually marks a genuine paradigm shift in the way human cognition can be mimicked by computing processes. This uncertainty is a key driver of the scepticism and apprehension about, and opposition to, the deployment of lethal autonomous weapon systems (LAWS).

This uncertainty is amplified when held against international humanitarian law guidelines that emphasize the need to distinguish between enemy targets and civilians (the distinction principle), the maintenance of proportionality of attacks and responses, and the use of precautions against attacks on civilians.[3] A main point of contention is not the structure of the weapons themselves and their effects but the method of their use. Despite the major global military powers issuing tepid cautionary statements, it is unlikely that they will stall the progress in development of autonomous weapons. The prospects of LAWS being true threat multipliers through entanglement with weapons of mass destruction (WMD) merits serious attention being given to the ways in which such weapons can be developed and deployed in the future. This is particularly critical in a highly volatile region such as South Asia, which includes three nuclear powers (China, India and Pakistan) engaged in geostrategic competition.[4] To address this, this essay looks first (in section I) at the potential impacts of autonomous weapons in the South Asian context. It then (in section II) proposes confidence-building measures (CBMs) to mitigate the related nuclear risk.

I. Autonomous weapons and the South Asian context

The discussion on development of autonomous weapons remains nascent, so it is opportune to outline some of the risks associated with the integration of nuclear weapons and autonomous systems. Both China and India currently possess weapon

[1] Turing, A., 'Can digital computers think?', Lecture, BBC Radio, 15 May 1951, Turing Digital Archive.

[2] Zastrow, M., 'Humans strike back: How Lee Sedol won a game against AlphaGo', *New Scientist*, 14 Mar. 2016.

[3] Piccone, T., 'How can international law regulate autonomous weapons?', Order From Chaos, Brookings Institution, 10 Apr. 2018.

[4] Fihn, B., 'It's time to disrupt nuclear weapons', *TechCrunch*, 10 Mar. 2019.

systems that have some form of autonomy. India recently acquired the S-400 Triumf air defence system, which has the ability to track targets independently.[5] Although this system is designed to be operated under human supervision, like other air defence platforms, the degree of autonomy in such systems is likely to increase with further modernization.[6]

The South Asian nuclear security context is distinct from other adversarial geostrategic nuclear relationships. Some of the key differences are the close geographical proximity of the adversaries, the entanglement of border disputes and the presence of violent non-state actors. In contrast, the cold war between the Soviet Union and the United States took place in the context of a bipolar strategic and ideological competition. Moreover, the greater distance separating the two powers made some of the crises more manageable because of longer response times. In the geographic context of South Asia, where ballistic missile flight times are under 10 minutes and conventional military interaction takes place along an extended border, the dangers of integrating autonomous systems into conventional and nuclear weapons could prove to be catastrophic.[7]

Another major difference is that, during the cold war, both superpowers managed alliances with third-party states while managing internal stability. That is, neither the USA nor the USSR experienced terrorist attacks on its own soil perceived to be coming from its great power rival. Although there are charges that nuclear-armed states have an undue influence on certain third-party states in South Asia, these partnerships have not developed to the same extent as the nuclear umbrella arrangements of the cold war. In addition, security experts often characterize the India–Pakistan relationship especially as one ruled by the stability–instability paradox, where states are willing to trade high levels of strategic stability for low-level instability. The question then is how much instability any of these actors is willing to tolerate.

While a potential arms race between states to develop and deploy autonomous weapons in the future is a concern in South Asia, perhaps of more immediate concern is the use of autonomous cyberweapons.[8] The prevalence of non-state actors engaging in violent struggles in South Asia adds new security risks, particularly in the cyber realm. For example, it is foreseeable that a non-state actor could realistically acquire a cyberweapon such as the 'Stuxnet virus' and use it to disrupt civilian or military facilities. Since cyberweapons have an attribution problem when deployed, such an attack in a South Asian context could be misconstrued as a deliberate attack from a state actor and escalate into a nuclear crisis.[9]

[5] Boulanin, V. and Verbruggen, M., *Mapping the Development of Autonomy in Weapons Systems* (SIPRI: Stockholm, Nov. 2017).

[7] Khan, Z., 'A restraint regime on MIRV flight-testing in South Asia', Stimson Center, 7 Nov. 2018.

[8] United Nations Institute for Disarmament Research (UNIDIR), *The Weaponization of Increasingly Autonomous Technologies: Autonomous Weapon Systems and Cyber Operations*, UNIDIR Resources no. 7 (UNIDIR: Geneva, 2017).

[9] Mussington, D., 'Strategic stability, cyber operations and international security', Centre for International Governance Innovation, 9 Apr. 2019.

II. Confidence-building measures and nuclear risk reduction

The objectives of CBMs are to prevent the outbreak of war and escalation in a crisis; to increase trust so as to avoid escalation; and to enhance early-warning systems and predictability. In a broad sense, CBMs can be grouped into four categories: communication measures, constraint measures, transparency measures and verification measures. CBMs can serve to lay the foundation for agreeing on acceptable norms of behaviour for states as well as building confidence and trust in order to avoid miscalculation and conflict escalation.[10] They can also represent initial steps towards discussions on arms control and finding common ground for understanding future cyberthreats in a crisis or warlike situation, including how to protect strategic assets and critical civilian infrastructure.

It is a difficult task, however, to imagine the sorts of CBM that can be adopted from other arms control regimes to regulate the use of autonomous weapons when the world is still struggling to agree on what autonomous weapons constitute. When discussing different types of measure it is important to be clear about what those measures aim to prevent or at least mitigate. In this regard, measures that instil strategic restraint in offensive cyber operations that have the potential to cause physical damage and harm must be the priority. However, CBMs also need to generally work and be sustainable. For this, CBMs need to be legally binding, rather than politically binding, as the latter 'help afford India and Pakistan the latitude to skirt proper implementation'.[11]

Proposed measures for South Asia

There is no shortage of proposals for CBMs and nuclear risk-reduction measures for South Asia. Proposals include all the items listed in the memorandum of understanding that accompanied the Lahore Declaration.[12] Newer proposals include incidents at sea agreements, nuclear risk-reduction centres and action items that can be undertaken unilaterally, such as decoupling warheads from delivery vehicles, keeping delivery vehicles unfuelled and sharing test-flight information. However, without making significant progress on the region's border disputes, even some of the most meaningful risk-reduction measures could come unstuck at a moment's notice.

Nuclear risk-reduction centres and incidents at sea agreements are two particularly effective CBMs.

[10] Sultan, A., *Universalizing Nuclear Nonproliferation Norms: A Regional Framework for the South Asian Nuclear Weapon States* (Palgrave Macmillan: Cham, 2019).

[11] Krepon, M. et al. (eds), *Global Confidence Building: New Tools for Troubled Regions* (St Martin's Press: New York, 1999), p. 176.

[12] Indian Ministry of External Affairs, Lahore Declaration, 2 Feb. 1999.

Nuclear risk-reduction centres

The US Nuclear Risk Reduction Center (NRRC) and the Russian National Centre for Nuclear Risk Reduction (NCNRR) constitute a successful CBM that uses the communication, transparency and verification elements.[13] In 2013 cybersecurity risks were added to the NRRC's scope. From a regional perspective, there is a dire need for the three nuclear powers in South Asia to establish a triangular CBM process that also includes the creation of nuclear risk-reduction centres.

Incidents at sea agreements

The 1972 Soviet–US Incidents at Sea Agreement aimed primarily at avoiding collisions between the two sides' vessels, along with other steps that could be taken to navigate high-traffic areas.[14] Such a successful CBM merits consideration for the major powers in South Asia. Such agreements will be a necessity as navies of the region engage in underwater swarming military exercises that could cause inadvertent escalation. Also, although the acoustic signatures of current nuclear-powered ballistic missile submarines (SSBNs) are somewhat detectable, the development of greater stealth capabilities of subsurface vessels will heighten maritime security concerns for all parties in the region.[15]

III. Conclusions

One of the key challenges for security analysts, particularly those working on arms control issues in South Asia, is how to interpret political and military signals within a coherent framework. The USA is virtually alone among nuclear powers in the way its military doctrinal framework is openly discussed, published and debated. The closest thing that India has to a nuclear doctrine is the 2003 version that contains all the characteristics that are familiar: no first use, credible minimum deterrence, massive nuclear retaliation.[16] Pakistan and, to a lesser extent, China remain even more opaque about their nuclear posture and doctrine.

In South Asia this strategic ambiguity is compounded by the uncertainty surrounding the development of autonomous weapons, regional instability and the current lack of effective CBMs. It is in this context that ever more sophisticated weaponry could result in strategic misjudgements through human error, overzealous commanders and, with the advent of autonomous weapons, a machine error. The case of Lieutenant Colonel Stanislav Petrov, who served at a Soviet early warning facility and who chose not to act on the false alarm of the automated Oko nuclear missile early-warning system, illustrates not only the

[13] US Department of State, 'United States Nuclear Risk Reduction Centre (NNRC)', [n.d.].

[14] Soviet–US Agreement on the Prevention of Incidents On and Over the High Seas, signed and entered into force 25 May 1972, *Nations Treaty Series*, vol. 852 (1972), pp. 151–54.

[15] 'Nuclear navies in South Asia', Arms Control Wonk podcast, 12 Nov. 2018.

[16] Narang, V., 'Five myths about India's nuclear posture', *Washington Quarterly*, vol. 36, no. 3 (2013); and Indian Ministry of External Affairs, 'The Cabinet Committee on Security reviews operationalization of India's nuclear doctrine', Press release, 4 Jan. 2003.

fallibility of automated systems, but also the importance of human judgement when it comes to key decision-making within highly sensitive command-and-control structures.[17] Even a single false positive appearing in such an automated system could eventually result in catastrophe—only sheer luck has prevented such an occurrence thus far. Therefore, it is of great strategic and humanitarian importance that military strategists, government officials and policymakers chart a path towards global restriction and regulation of the development and deployment of autonomous weapons.

[17] Topychkanov, P., 'Autonomy in Russian nuclear forces', ed. V. Boulanin, *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. I, *Euro-Atlantic Perspectives* (SIPRI: Stockholm, May 2019), p. 70.

Conclusions

10. The opportunities and risks of artificial intelligence for strategic stability in South Asia

PETR TOPYCHKANOV

This edited volume is the third in a trilogy that explores regional perspectives and trends related to the impact that recent advances in artificial intelligence (AI) could have on nuclear risk and strategic stability. It brings together the views of eight experts from South Asia and around the world who participated in a workshop in Colombo in February 2019.

This concluding chapter explores the opportunities and risks of AI for strategic stability in South Asia. It begins (in section I) by summarizing the state of adoption of military AI in the region. It then (in section II) assesses the overall strategic impact of AI. Specific arms control initiatives and confidence-building measures (CBMs) that may mitigate this impact are described (in section III) before the chapter and the volume conclude (in section IV) with a brief remark on the South Asian dialogue on military uses of AI.

I. The state of adoption of the military AI in South Asia

The accomplishments of India and Pakistan in the adoption of military AI are modest. They cannot compete with the significant advances in the militarization of AI achieved by the United States, Russia and China. The different levels of public access to information on military programmes mean that more can be said about India's programmes than those of Pakistan. Open sources on Pakistan do not provide a complete picture of national efforts to militarize AI. In contrast, India's state and independent media give a sense of the scale of military programmes related to AI.

As the two contributors from India—Kritika Roy (chapter 3) and Sanatan Kulshrestha (chapter 5)—explain, the central institution behind India's militarization of AI is the Centre for Artificial Intelligence and Robotics (CAIR) of the Defence Research and Development Organisation (DRDO). However, although CAIR has been working on military applications of AI for strategic systems for at least the past two decades, neither Roy nor Kulshrestha name any product of CAIR that is directly related to nuclear forces. While Roy mentions several examples of intelligence, surveillance and reconnaissance (ISR) and electronic warfare systems, most are still under development and have no clear connection to India's nuclear forces. She offers two explanations for this. First, because there remain many reliability and control problems with today's AI, India is likely to be cautious about the integration of AI into its weapons and military equipment. The second explanation is the Indian defence procurement process, which is characterized by a risk-aversion that causes long delays in weapon acquisition.

As Kulshrestha reports, the Indian Ministry of Defence's AI Task Force has recommended that the government works with India's many private companies of different sizes—including start-ups—that are interested in developing AI for military purposes. India's future efforts to militarize AI are likely to focus on the non-state sector because it might become more efficient in making AI products for the Indian armed forces than the state-owned arms-production industry.

The Pakistani case also requires further research to understand the scope of national efforts to militarize AI. Saima Aman Sial argues (in chapter 7) that her country does not differ from other nuclear-armed states, where 'There is hardly any published or publicly available literature on the subject of the application AI in nuclear weapon systems'. However, Pakistan does differ in the degree of information available on research and development (R&D) programmes on military AI and the state of adoption of these technologies. For this reason, analysis of military AI in Pakistan has to rely on indirect indicators.

Sial mentions several nuclear weapon-related systems that are enabled by AI, including the fully automated Strategic Command and Control Support System (SCCSS). However, the developers of these systems and the level of AI adoption evade analysis. But this does not mean that there is no focus on military AI in Pakistan. On the contrary, the Pakistani armed forces are apparently exploring AI-related opportunities, building national capacity for developing AI, and introducing it into weapons and military systems. However, there is less public transparency in these efforts than in India's development of military AI.

II. The impact of AI on strategic stability in South Asia

The idea that AI is a double-edged sword as far as strategic stability is concerned is prevalent in this volume, especially in the essays by Maaike Verbruggen (chapter 2), Roy (chapter 3), Kulshrestha (chapter 5), the present author (chapter 6) and Sial (chapter 7).

Sial describes the double-edged sword as manifested in South Asia from the Pakistani perspective. On the one hand, AI-driven technologies could supposedly play a critical role in each component of the nuclear deterrence capabilities, especially by enhancing ISR and early-warning capabilities. Military applications of AI could also enhance the accuracy and reaction time of strategic offensive and defensive weapons. On the other hand, AI may fuel instability, especially in crises, by challenging the survivability of an adversary's second-strike capabilities. The AI-enhanced ability to target strategic assets with conventional high-precision weapons could increase the fear of a first strike.

The introduction of military AI may risk the lowering of the nuclear threshold. Enhanced precision and autonomy, both enabled by AI, may revitalize the role of tactical nuclear weapons. Taking into account the lack of experience in arms control and CBMs for tactical nuclear weapons, such a revitalization would have a destabilizing effect for any region, including South Asia.

There are also difficulties specific to the South Asian context. Significant sections of the borders between China, India and Pakistan are disputed. There

have been many armed conflicts and crises between Pakistan and India and between China and India. During the most recent crisis between India and Pakistan, in February 2019, both sides used dual-capable combat aircraft. The close geographic proximity of the two sides leaves only a few minutes for situation assessment, decision-making and taking military action.

These difficulties partly explain the types of AI-enabled capability that India and Pakistan might plan to obtain. As Kulshrestha notes, for military operations in disputed areas, India could use 'robotic sentinels that can respond effectively and neutralize the arising threats'. As he rightly notes, these would minimize risk to the Indian armed forces. However, the practice of bringing autonomous systems, especially lethal autonomous weapon systems (LAWS), to such disputed areas could lower the threshold for the use of force in at least two ways.

First, given the early stage of development of autonomous ISR and early-warning capabilities in India and Pakistan, there is a high risk of false alarms from these systems. In the geographical conditions of South Asia, which leave no time to double-check data, these alarms might be considered valid and responded to by use of force.

Second, when one side deploys such autonomous platforms in a disputed area, the other side might consider it as aggression or as preparations for attack. Hence, misperceptions and incorrect calculations about AI-enabled weapons could cause military responses between the South Asian opponents.

Furthermore, Roy and Sial mention the possibility of AI and machine learning being used to bring about these misperceptions via poisoning data, creating deep fakes and other actions in cybersphere. However, the question of which actors would be interested in using cyber means to provoke a war between nuclear-armed states remains unanswered. During the workshop in Colombo, which laid the ground for this volume, some participants expressed doubts about terrorists' using AI and machine learning in this way.

Alongside the negative impacts of AI on strategic stability in South Asia, there are positive impacts, covered by Roy, Sial and the present author. As these authors hint in their essays, these impacts might be achieved through unilateral, bilateral or multilateral efforts. Unilaterally, military AI could aid in providing better ISR and early warning, enhancing decision-making capability, increasing the safety and reliability of nuclear arsenals, and improving cybersecurity. In bilateral and multilateral contexts, AI-enabled solutions could be part of transparency, confidence-building and verification mechanisms. This seems to be feasible from a technological perspective. However, the political context of the relationships between China, India and Pakistan mean that there is no expectation of such measures in the nearest future.

III. Arms control and confidence-building measures in the area of AI and nuclear weapons

The question of controlling the application of AI in the military domain has various dimensions. From the technological perspective, there is no mechanism in place

that restricts the application or deployment of AI-driven military developments. The history of nuclear arms control includes no example of banning or controlling technologies during their R&D stage. Usually technologies have only been banned or controlled after testing, deployment and use, when the ways in which they could be limited have become apparent.

The problem becomes even more difficult in the growing grey zone between conventional and nuclear options because of the introduction of AI to both conventional and nuclear weapon systems. The contributors thus draw attention to developing various types of transparency and CBM, as well as traditional arms control.

Banning lethal autonomous weapon systems

Yanitra Kumaraguru (chapter 8) recommends an international pre-emptive ban on the development, production and use of LAWS. Such a ban would not only cover weapon systems without any human control but also those where human control is nominal; that is, based on the machine's suggestions, and without sufficient context for ethical decision-making. Kumaraguru argues that this ban will not hamper the development of technology for civilian or peaceful purposes in any way, while allowing R&D of AI for military purposes to continue within an ethical framework.

However, the question of how to verify a pre-emptive ban is unanswered. Kumaraguru does not explore this question at length, but it should be discussed further. A ban might not need verification at the initial stage, as shown by the 2017 Treaty on the Prohibition of Nuclear Weapons. Some measures that may help to verify the ban include pre-notifications by states of the deployment of LAWS and limitations on these deployments to agreed locations. Similarly, while Malinda Meegoda (chapter 9) recommends global restriction and regulation of autonomous weapon systems, she finds it hard to formulate the CBMs needed when there is no agreed definition of these systems.

In the South Asian context, a ban on R&D and production of LAWS is not attractive for either India or Pakistan. It is difficult for these two countries to achieve a common understanding of threats associated with LAWS without a mutual dialogue. Hence, the initial regional steps towards a prohibition could be a doctrinal dialogue with a specific focus on military AI and increased transparency and CBMs. These less ambitious steps have the potential to clear the path for future limitations on LAWS.

Transparency and communication

Transparency is a core CBM. However, as the present author notes (in chapter 6), South Asia's three nuclear-armed states—China, India and Pakistan—currently lack transparency and undisrupted mutual communication. As an initial step, they could each take simple, unilateral transparency measures, such as being more open about their R&D projects on military AI. The level of transparency

described by Verbruggen (in chapter 2) in the Euro-Atlantic context offers a ready example for South Asia.

In addition to simple unilateral steps towards transparency, Meegoda (chapter 9) argues that the three nuclear powers need to establish a triangular CBM process. As well as bilateral or trilateral transparency, he proposes such measures as nuclear risk-reduction centres and an 'incidents at sea' agreement. The nuclear risk-reduction centres would be modelled on those of Russia and the USA. Doctrinal dialogue on nuclear weapon issues may help the three countries to realize the need for regional nuclear risk-reduction centres. The establishment of such permanent portals for transparency, CBMs and nuclear risk mitigation will shrink the grounds for misperceptions related to nuclear weapon policies. The 'incidents at sea' agreement would also be modelled on that between Russia and the USA for the prevention of vessel collisions and procedures to de-escalate a crisis. In South Asia, where the region's navies engage in underwater swarming military exercises that could inadvertently cause escalation, such an agreement should also cover the use of autonomous weapons at sea.

Regional dialogues

Dmitry Stefanovich (chapter 4), the present author (chapter 6) and Sial (chapter 7) each recommend the establishment, as a CBM, of a regional dialogue on nuclear modernization and doctrines, including the role of AI. China, India and Pakistan have already engaged in various track 2 dialogues in bilateral and trilateral formats. China and India also have an annual defence and security dialogue, which met for the ninth time in 2019. However, the three countries continue to skirt around the subject of nuclear weapons and emerging technologies at such meetings.

IV. Final remarks

South Asian countries are still at the early stages of adoption of military AI. However, foreseeable advances could be destabilizing as they could affect these states' sense of security. It is evident from this volume and the project in general that there is a need for a regional dialogue on AI and nuclear risk. Such a dialogue could help South Asian countries avoid misperception of each other's capabilities in this area. Because China, India and Pakistan are involved in nuclear deterrence relations, the dialogue on military AI might be a separate chapter of a comprehensive dialogue on strategic stability.

About the authors

Sanatan Kulshrestha (India) was a rear admiral of the Indian Navy until his retirement in 2011. During more than 34 years of service, he specialized in quality assurance in the Indian Naval Armament Service, with key appointments at Naval Command Headquarters, defence research establishments and Indian Ordnance Factories, before finally becoming the Director General of Naval Armament Inspection. Kulshrestha has been engaged in the study of strategic aspects of emerging technologies with specific relevance to national security issues. He graduated from Jodhpur University with a Gold Medal in physics and has a doctorate in international studies from Jawaharlal Nehru University, New Delhi. He speaks frequently at international conferences and contributes regularly to defence journals on maritime issues and defence technology.

Yanitra Kumaraguru (Sri Lanka) lectures at the Faculty of Law of the University of Colombo. She also serves as coordinator of the Sri Lanka Campaign to Stop Killer Robots, under the Forum for Disarmament and Development. Her research interests revolve around questions of humanitarian and human rights laws. She is currently reading for a master's degree in law at Harvard Law School, United States.

Malinda Meegoda (Sri Lanka) is a research associate at the Lakshman Kadirgamar Institute in Colombo. He is a graduate of the International Cooperation and Conflict Studies programme of the University of Saskatchewan, Canada. His research interests include nuclear weapons, arms control, disarmament and non-proliferation, and international development. His publications include articles appearing in a number of major Sri Lankan newspapers and magazines such as the *Sunday Times*, the *Daily Financial Times* and *Lanka Monthly Digest*.

Kritika Roy (India) is a research analyst at the Cybersecurity Centre of Excellence at the Institute for Defence Studies and Analyses (IDSA) in New Delhi. She is also an assistant editor of IDSA's *CBW Magazine*. Roy is an engineering graduate (electronics and communication) and has a master's degree in geopolitics and international relations from Manipal Academy of Higher Education. Her research interests revolve around the emergence of new technologies, including artificial intelligence (AI) and the issues of countering the threat of weapons of mass destruction in the age of disruptive technologies. Her latest book is *Advances in ICT and the Likely Nature of Warfare* (Routledge, 2019).

Saima Aman Sial (Pakistan) is a senior research officer at the Center for International Strategic Studies (CISS) in Islamabad. She is also the associate editor of the biannual journal *CISS Insight*. Sial has been a visiting research fellow at the James Martin Center for Nonproliferation Studies, USA, the Nonproliferation Education and Research Center, South Korea, the Henry L. Stimson Center, USA,

and the Sandia National Laboratories, USA. She has a master's degree in strategic studies from the National Defence University, Islamabad.

Dmitry Stefanovich (Russia) is a research fellow at the Center for International Security at the Primakov Institute of World Economy and International Relations (IMEMO) of the Russian Academy of Sciences. He is also a visiting research fellow at the Arms Control and Emerging Technologies project of the Institute for Peace Research and Security Policy at the University of Hamburg (IFSH). Since 2016 he has been a Russian International Affairs Council (RIAC) expert and an independent military observer. Stefanovich graduated from the National Research Nuclear University MEPhI in Moscow with a specialist degree in international relations, focusing on international scientific and technological cooperation. His research interests revolve around the disruptive technologies that are challenging strategic stability and the logic behind the development and deployment of new strategic weapons.

Petr Topychkanov (Russia) is a senior researcher with the SIPRI Nuclear Disarmament, Arms Control and Non-proliferation Programme. He works on issues related to nuclear non-proliferation, disarmament, arms control and the impact of new technologies on strategic stability. Prior to joining SIPRI in 2018, Topychkanov was a senior researcher at the IMEMO Center for International Security. From 2006 to 2017 he was a fellow with the Carnegie Moscow Center's Nonproliferation Program. He has a doctorate in history from the Institute of Asian and African Studies, Moscow State University. His recent publications include 'US–Soviet/Russian dialogue on the nuclear weapons programme of India', *Strategic Analysis* (May 2018).

Maaike Verbruggen (the Netherlands) is a doctoral researcher at the Institute for European Studies of Vrije Universiteit Brussel, where she is investigating the drivers of and obstacles to military innovation in AI. Her other research interests include arms control, the arms trade and emerging military technologies. Her recent publications focus on the challenges to 'spin-in' of AI, the impact of civilian innovation on the development of lethal autonomous weapon systems (LAWS) and the challenges to human control over swarms. Before undertaking her doctoral research, Verbruggen worked at SIPRI.