VOLUME 8 | NUMBER 2 | 2024

DSIAC JOURNAL

DETECTING AND DEFENDING AGAINST MALICIOUS ATTACKS TO SHIP SENSORS

DSIAC JOURNAL

Volume 8 // Number 2 // 2024

Editor-in-Chief:
Aaron Hodges

Sr. Technical Editor:
Maria Brady

Graphic Designers:
Melissa Gestido, Katie Ogorzalek

The DSIAC Journal is a publication of the Defense Systems Information Analysis Center (DSIAC). DSIAC is a DoD Information Analysis Center (IAC) sponsored by the Defense Technical Information Center (DTIC) with policy oversight provided by the Office of the Under Secretary of Defense (OUSD) for Research and Engineering (R&E). DSIAC is operated by the SURVICE Engineering Company.

ISSN 2471-3392 (Print) // ISSN 2471-3406 (Online)

Distribution Statement A:
Approved for public release; distribution is unlimited.

On the Cover:
Digital Art Rendering *(Source: Kedek Creative [stock.adobe.com], YANKOVICH [stock.adobe.com], and Lunarts Studio [Canva]).*

DSIAC Defense Systems Information Analysis Center

ABOUT DSIAC

Who We Are

A DoD Information Analysis Center comprised of scientists, engineers, researchers, analysts, and information specialists.

What We Do

Generate, collect, research, analyze, synthesize, and disseminate scientific and technical information (STI) to DoD and federal government users and industry contractors.

Why Our Services

To eliminate redundancy, foster collaboration, and stimulate innovation.

DSIAC SERVICES

Subject Matter Expert (SME) Connections

Access to a network of experts with expertise across our technical focus areas.

Technical Inquiries (TIs)

Up to 4 hours of FREE research using vast DoD information resources and our extensive network of SMEs.

Specialized Task Orders

Research and analysis services to solve our customer's toughest scientific and technical problems.

Webinars & Events

Our webinars feature a technical presentation from a SME in one of our focus areas. We also offer key technical conferences and forums for the science and technology community.

STI Collection

Our knowledge management team collects and uploads all pertinent STI into DTIC's Research & Engineering Gateway.

Information Research Products

The Defense Systems Digest, state-of-the-art reports, journals, TI response reports, and more available on our website.

CONTACT DSIAC

IAC Program Management Office

8725 John J. Kingman Road
Fort Belvoir, VA 22060
Office: 571.448.9753

DSIAC Headquarters

4695 Millennium Drive
Belcamp, MD 21017-1505
Office: 443.360.4600
Fax: 410.272.6763
Email: contact@dsiac.org

DSIAC Technical Project Lead

Taylor Knight
4695 Millennium Drive
Belcamp, MD 21017-1505
Office: 443.360.4600

FEATURED ARTICLE

DETECTING AND DEFENDING AGAINST MALICIOUS ATTACKS TO SHIP SENSORS

By R. Glenn Wright

This article describes research to detect phenomena that may degrade or disrupt the performance of an otherwise fully functional sensor. Also examined are methods to potentially mitigate the effects of sensor degradation and develop effective countermeasures that enable naval vessels and unmanned vehicles to continue their missions with degraded capabilities.

IN THIS ISSUE

EMERGING RISKS IN SPACE FROM CHINA AND RUSSIA

BY DAVID CHEN

(PHOTO SOURCE: SHUTTERSTOCK)

SUMMARY

Emerging technologies include not only novel technical inventions but also repurposing existing technologies into new tactics, techniques, and procedures. In recent years, Russia and China have upended the way the space domain has operated for decades. For example, China has repurposed the Cold War-era concept of a "fractional orbital bombardment system" to deploy a novel hypersonic vehicle. Russia has deployed a test satellite in a program for placing a nuclear device in Earth's orbit. As the international terrestrial arms control treaty framework has steadily eroded, the risk of conflict in space has grown and transformed. Keeping ahead of the dynamics of disruption in the domain will require imagination and creativity. This article reviews the emerging risk factors in space enabled by new technological applications from China and Russia and recommends using a deeper understanding of adversaries' strategic thought as a foundational premise from which to develop multidomain technical, operational, and policy countermeasures.

AN EVOLVING BATTLESPACE

Recent events have indicated increasing volatility in space. Gen. Stephen Whiting, Commander of the U.S. Space Command, testified in early 2024 that "The [People's Republic of China] is moving breathtakingly fast in space. America must rapidly increase the timeliness, quality, and quantity of our critical national space and missile defense systems to match China's speed and maintain our advantage" [1].

On the quantity metric, China has developed its space capabilities rapidly and robustly. According to Maj. Gen. Greg Gagnon, Chief of Space Operations for Intelligence, China has launched more than 400 satellites in the past two years; but more importantly, "[These are] satellites that are designed with a proliferated architecture…An architecture that isn't designed for efficiency and cost effectiveness, [but] an architecture that's designed to go to war and sustain in war" [2]. As much of the Western Pacific and Indian Ocean becomes a de facto "weapons engagement zone" for U.S. forces, China is increasingly able to hold key assets at risk in any kinetic scenario.

Russia has also made clear that it intends to play a disruptive role in space. In May 2024, United States government (USG) officials, including Assistant Secretary of Defense for Space Policy John Plumb and Assistant Secretary of State for Arms Control, Deterrence, and Stability Mallory Stewart, offered the first official details about Russia's program for developing a nuclear detonation antisatellite (ASAT) weapon in Earth orbit [3]. Secretary Stewart focused on Russia's veto of a United Nations (UN) Security Council resolution reaffirming the international commitment to prohibit the deployment of weapons of mass destruction in space like nuclear weapons [4]. She also disclosed the existence of an on-orbit testbed satellite, which analysts quickly identified as Cosmos-2553, in an orbit at an unusually high altitude and within the inner Van Allen radiation belt (Figure 1) [5]. Russia's actions in space and at the UN indicate its

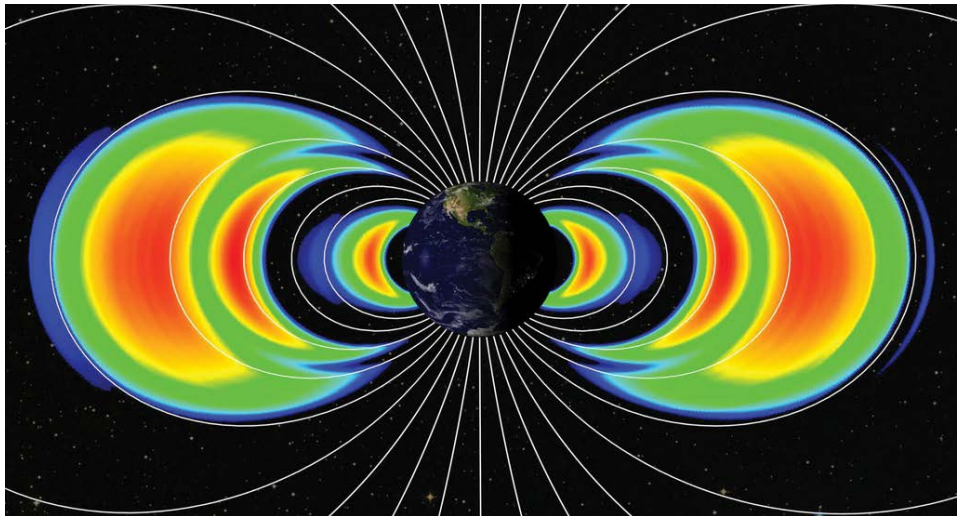


Figure 1. Van Allen Belts *(Source: National Aeronautics and Space Administration [NASA]).*

potential willingness to abrogate its treaty commitments on arms control in space.

AGENTS OF DISRUPTION

In recent years, China and Russia have challenged the predominant security environment in space using new technologies and novel deployments of legacy technologies. Since 2015, China has embarked on an organizational transformation of the Chinese People's Liberation Army (PLA), resulting in the creation of the PLA Rocket Force and the Strategic Support Force, the latter of which has recently been again reorganized into three functional forces—the Military Space Force (军事航天部队), the Cyberspace Force (网络空间部), and the Information Support Force (信息支援部) [6]. Combined with a sustained shipbuilding campaign as well as a robust space launch cadence, China has implemented a strategic plan to achieve expeditionary capabilities and extend the reach of its information and fire-support services into new theaters of operation. In the May 2024 "Joint Sword-2024A" (联合利剑-2024A) exercise [7], some of these capabilities were described by a PLA Navy military analyst as establishing a "firepower coverage network" (nicknamed a "firepower combo platter" since one could pick and choose the shooter from any domain or region). This network was established by weaving multidomain platforms together by "point, line, and surface" (点，线，面) into what Western analysts would term an "integrated reconnaissance-strike complex." This operational exercise helps illustrate how space capabilities are increasingly critical enablers for China's evolving operational doctrine and expanding force posture.

In space, China's disruptive role can be dated back to 2007. The first destructive ASAT test in a generation occurred when the People's Republic of China (PRC) destroyed a defunct weather satellite with a direct-ascent Dongneng-1 missile, generating a massive debris cloud [8]. In 2019, the PRC launched an experimental hypersonic vehicle through space that surprised analysts with its "fractional orbit" trajectory [9]. Distinct from a suborbital flight, a fractional orbit reaches orbital velocity, but the object decelerates to reenter the atmosphere before completing an orbit. China refused to comment on the hypersonic test, or a subsequent one, instead conflating it with a prior spaceplane flight [10]. The test was interpreted by analysts and experts in terms of strategic nuclear employment, echoing the rationale for the Soviet-era Fractional Orbital Bombardment System (FOBS).

Russia has also conducted kinetic ASAT tests in recent years. A provocative direct-ascent Nudol missile intercept in 2021 destroyed a defunct Cosmos satellite, generating debris that threatened the crew of the International Space Station, including two cosmonauts aboard at the time [11]. The crew sheltered in their return capsule lifeboats during the most dangerous initial passes through the debris cloud, and operations were halted temporarily [12]. By threatening its own personnel, Russia sent a costly signal to better demonstrate resolve to international observers. Additionally, Russia conducted apparent on-orbital ASAT tests using high-velocity projectiles [13]:

> In 2017, Cosmos 2521 released a subsatellite, Cosmos 2523, at a high velocity, possibly testing an orbital projectile weapon.… Cosmos 2542 potentially performed similar projectile weapons testing, launching an object with a relative velocity of between 140 and 186 meters per second.

> ***The first destructive ASAT test in a generation occurred when the People's Republic of China destroyed a defunct weather satellite with a direct-ascent Dongneng-1 missile, generating a massive debris cloud.***

Beyond kinetic strike capabilities, Russia has also advanced development of nuclear ASAT weapons in space. These developments demonstrate that Russia intends to remain a major space power capable of threatening international security. Given the asymmetric dependencies of its military operations on space-based assets vs. those of its adversaries, Russia may be particularly interested in acting as a spoiler in space.

RECONSIDERING CONCEPTS OF OPERATIONS (CONOPS)

The remainder of this article will examine CONOPS for employing a FOBS-hypersonic glide vehicle (HGV) hybrid system and nuclear detonation ASAT. Other developments in space, such as satellite remote proximity operations, non-Earth imaging from space (satellite-to-satellite imaging), cyber-electronic exploitation, directed energy, and more, all merit additional scrutiny. However, given scope constraints, the following discussion will focus on the FOBS-HGV and nuclear ASAT examples and their strategic implications.

FOBs and HGVs

The Soviet Union first developed the concept of operations for fractional orbit weapons systems in the 1960s. Article IV of the Outer Space Treaty (OST) commits states parties "not to place in orbit around the earth any objects carrying nuclear weapons… or station such weapons in outer space in any other manner" [14]. For compliance purposes, however, a nuclear weapon in space would not be considered in violation short of a full orbit. Hence, a fractional orbital bombardment system would remain in compliance with the letter of the treaty. The "Globalnaya Raketa-1" or "Global Missile-1" system, described by the North Atlantic Treaty Organization (NATO) as the "Fractional Orbital Bombardment System," would launch nuclear-armed intercontinental ballistic missiles (ICBMs) over southern polar regions (e.g., Figure 2), avoiding early generations of North American Air Defense ballistic missile early warning (BMEW) radars [15]. Yet, the original Soviet system quickly became obsolete due to the development of additional radar sites and BMEW satellites, such as the Defense Support Program series. Therefore, by the 1980s, the strategic logic of the FOBS no longer held significant competitive advantage over direct-strike ICBMs, leading the FOBS launchers to be "decommissioned under the terms of the SALT-2 treaty" [15].

The PRC's version of FOBS could serve a similar strategic role. U.S. Air Force Secretary Frank Kendall disclosed that the PRC system resembled the Soviet FOBS in its flight profile and asserted that China is "acquiring a first-strike capability" [16]. A hypersonic glide vehicle would be more difficult to track using satellite systems designed to warn against ICBM launches, but new sensors tailored for maintaining custody of hypersonic vehicles while in flight are in development. The Hypersonic and Ballistic Tracking Space Sensor system (Figure 3) will enter a multiyear on-orbit testing phase in 2024 [17]. Therefore, the window of competitive

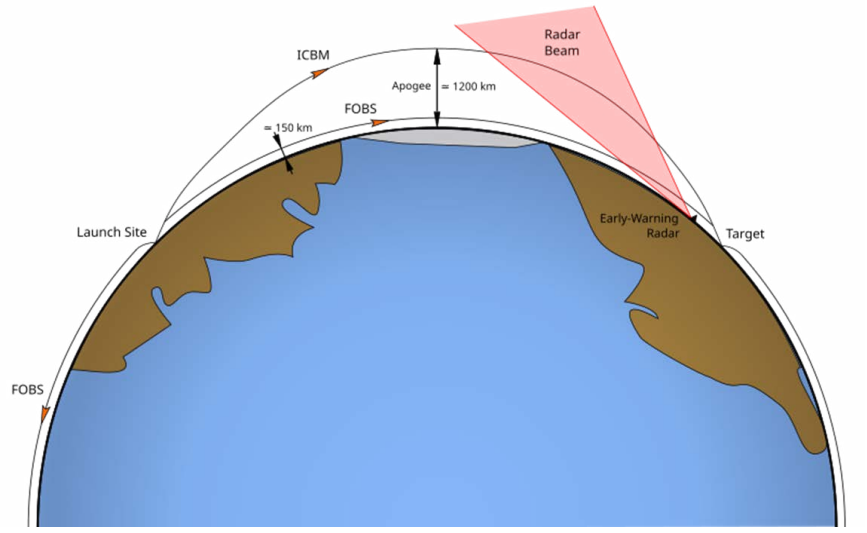


Figure 2. A Diagram Comparing Typical ICBM Trajectories With FOBS Trajectories *(Source: Heriberto Arribas Abato, Numerica Corporation [Creative Commons]).*

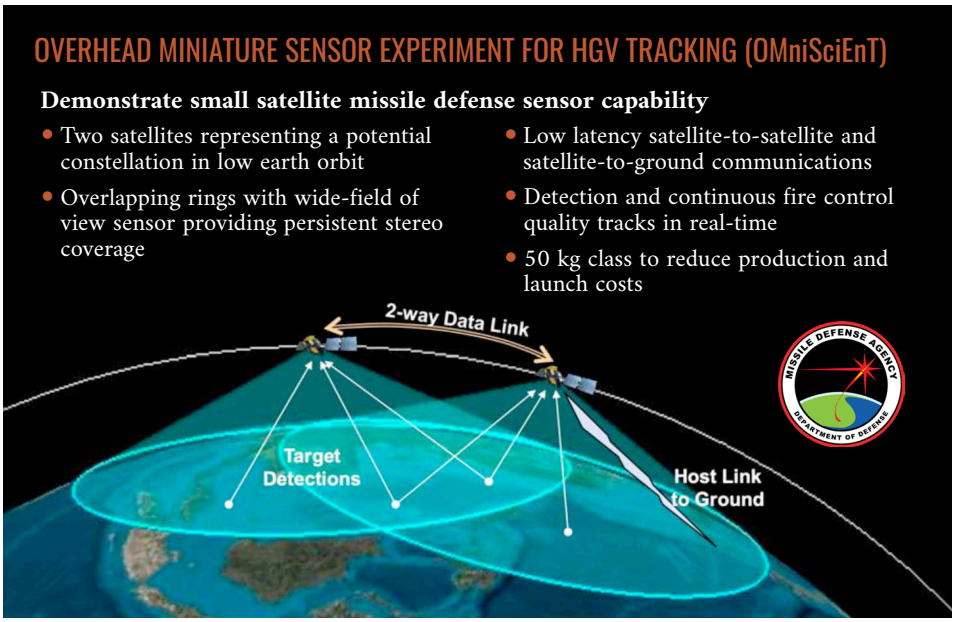


Figure 3. Hypersonic and Ballistic Tracking Space Sensor System *(Source: Missile Defense Agency)*.

advantage for a PRC nuclear FOBS like the Soviet version may be narrow. In addition, China's doctrine of "integrated strategic deterrence" mixes conventional and strategic deterrence "which could present severe disambiguation problems" for analysts trying to understand the CONOPS of such a platform [18]. China's policy of strategic ambiguity regarding nuclear and conventional deterrence creates deliberate uncertainties in contrast with the decades-long tradition of transparency and verification for nuclear weapons based on the international nuclear arms *control treaty framework*.

Other technical aspects of the system raise important questions about the employment possibilities of an integrated FOBS-HGV. To date, analysts have focused on the strategic utility of FOBS for a nuclear strike [19]:

> The use of an orbital bombardment system could increase PLA power projection capabilities against bases and territories globally, including targets in the 50 states. The use of an orbital bombardment system can complicate U.S. missile defenses by forcing the US to defend against joint and combined arms attacks from multiple directions.

Adding an HGV to FOBS introduces an additional element of uncertainty and agility. HGVs travel slower than ballistic reentry vehicles but are more difficult to intercept during the glide phase because they can maneuver. Still, they are not suited for every application. Travel at hypersonic speeds in the upper atmosphere generates a plasma envelope around the vehicle, degrading access to global command, control, communications, and computers intelligence, surveillance, and reconnaissance networks and creating high infrared and electromagnetic signatures. As a general feature of hypersonic flight, this makes HGVs most effective against fixed targets but largely impractical for mobile high-value assets like ships.

The 2019 PRC hypersonic test was groundbreaking not just because it entered an orbital trajectory but also because it demonstrated a capability no other nation has developed. After the hypersonic glide vehicle reentered the atmosphere, it released a secondary vehicle while traveling at hypersonic speeds, surprising USG observers: "Government scientists were struggling to understand the capability, which the US does not currently possess, adding that China's achievement appeared 'to defy the laws of physics'" [10]. While USG analysts' assessment of the hypersonic test flights remains necessarily closed, it is noteworthy that this element of the testing entered the public discourse. As the *Financial Times* explains [20]:

> Experts at DARPA, the Pentagon's advanced research agency, remain unsure how China managed to fire countermeasures from a vehicle travelling at hypersonic speeds.… Military experts have been poring over data related to the test to understand how China mastered the technology. They

> ***The 2019 PRC hypersonic test was groundbreaking not just because it entered an orbital trajectory but also because it demonstrated a capability no other nation has developed.***

are also debating the purpose of the projectile, which was fired by the hypersonic vehicle with no obvious target of its own, before plunging into the water.

This secondary payload could have been a multiple independently targetable, reentry vehicle warhead, a countermeasure, or something entirely different. In a nuclear context, reentry vehicles and countermeasures have been part of warhead development for decades. In a conventional context, the secondary vehicle may have addressed the fundamental limitation of hypersonic vehicles in fixing mobile targets by providing "target update" on a local scale.

The conventional use case aligns with evolving doctrinal discourse within the PRC in recent years. PRC doctrine has emphasized the role of multidomain warfare since at least 2013 [21]:

> Space and cyberspace increasingly constitute important battlefields after the traditional battlefields of land, sea, and air. A new type of five-dimensional battlespace of land, sea, air, space, and cyber is currently taking shape, which is wide in scope, hyper-dimensional, and combines the tangible and intangible. Battlefield control is moving from control of the land, sea, and air toward control of space and cyber.

"Multidimensional" includes both multidomain and multidirectional operations. This doctrine was demonstrated in the Joint Sword-2024A exercise in which multiple firepower platforms surrounded Taiwan to execute "joint firepower strikes" from multiple vectors [7]. PRC defense analysts have also advocated for using autonomy and speed to achieve decision dominance over a more formidable adversary. One theme in this discourse is developing the next generation of military doctrine to supplant "network-centric warfare" pioneered by the U.S. military. Hypersonic weapons play a key role in this analysis of future warfare, as described by one PLA analyst who proposes one potential successor concept as "energy-centric warfare" (能量中心战) [22]:

> "Energy-centric warfare" stresses increasing the speed of the link which is "attack": the specific way to do so is to develop new concept weapons such as near space hypersonic weapons, electromagnetic rail guns, and directed energy weapons, shortening the time between detection and destruction of a target.

This statement is made in the context of achieving a faster observe, orient, decide, act (OODA) loop kill chain to achieve effects before an adversary can react. Notably, the author contrasts this concept with U.S. concepts of operations, saying: "traditional 'network centric warfare' emphasizes delivery of information, whereas 'energy centric warfare' emphasizes delivery of firepower…whether information or firepower, both are categories of energy" [22].

Compared to the U.S. way of war where information dominance is assumed, PRC defense analysts posit that "information agility" is more important than information dominance. Per one PRC aerospace industry official, "Speed and agility are no longer most important. The key to winning air operations, electromagnetic operations, or cyber operations is 'information agility,' the priority and mobility of information" [23]. The argument is that flexible networks and built-in autonomy will enable the rapid delivery of effects needed to get inside an adversary's OODA loop. In contrast, there has been little evidence to indicate a shift in PRC nuclear doctrine toward a first-strike posture and FOBS-HGV role for such, even as Western analysts gravitate toward that possibility [24].

Under the PRC doctrinal concept of "integrated strategic deterrence," the potential use of a FOBS-HGV platform in a conventional strike role should be more fully considered, especially given the unexpected capabilities the PRC demonstrated in its flight testing.

If the PRC FOBS is not a new nuclear first-strike weapon but rather a conventional strike platform, the observed secondary munition could be part of a solution for fixing mobile targets like aircraft carriers, particularly under degraded or denied information access conditions. Integrating artificial intelligence capability, "intelligentization" (智能化) in PLA parlance, and providing local sensor data from the secondary vehicle could enable the warhead to rapidly find and fix a mobile target independent of mission controllers.

The 2015 book *Light Warfare* (光战争) described the advantage of speed in integrating detection and strike in this way: "As photonic weapons emerge, so will a genuine 'one-second kill' [capability], bringing about the true meaning of detect-and-destroy" [25]. *Light Warfare* proposed a thesis around photonic, electromagnetic, and hypersonic weapons as part of a hypervelocity suite of new-concept weapons that would supplant traditional firepower. An integrated FOBS-HGV system could represent such a development with the convergence of sensor and weapon ("查打一体"), integrating detection and destruction into a single platform operating on an independent kill chain.

China's uninhabited aerial vehicle or drone development programs connote the fusion of sensors and shooter into a single platform. Operationally, such a platform could take advantage of continued connectivity during the orbital phase of flight and rely on the deployable sensor platform to refine its targeting data when entering the terminal phase. Even the potential for such a weapon serves deterrence purposes. The ability to hold targets at risk using a variety of antiship cruise, ballistic, and hypersonic missiles at a variety of ranges is a core feature of "integrated strategic deterrence" [26]. Introducing a FOBS-HGV option would add new approach vectors for conducting coordinated saturation attacks against key targets like aircraft carriers. In a multivector attack, the defender must allocate scarce assets against the most pressing threat. A FOBS-HGV may be useful in drawing enough of these electronic and kinetic assets to allow other attack vectors to succeed.

Nuclear Detonation in Space

New concepts of operations for nuclear detonation in the upper atmosphere or in space are also reemerging (Figure 4). The 1963 Limited Test Ban Treaty, the 1967 Outer Space Treaty, and other arms control treaties were a direct consequence of the United States' 1962 Starfish Prime high-altitude nuclear test (Figure 5) and the realization that nuclear effects in space could lead to unintended consequences for global security and commerce [27].

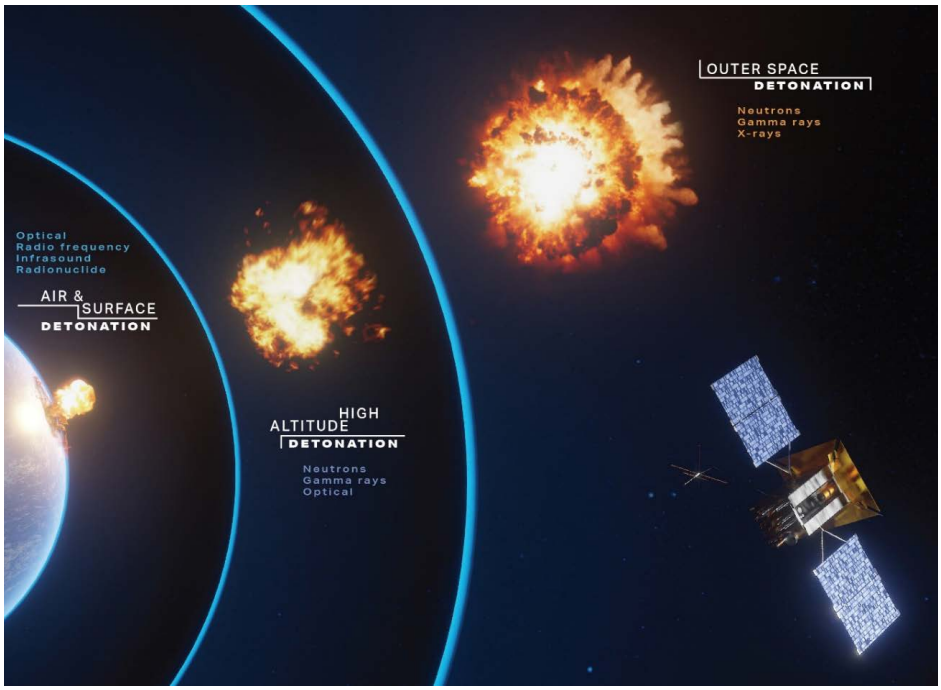


Figure 4. Depiction of Nuclear Detonation in Space at Different Altitudes *(Source: Los Alamos National Laboratories)*.

Figure 5. Starfish Prime Nuclear Test *(Source: [Top] U.S. Air Force 1352nd Photographic Group, Lookout Mountain Station and [Bottom] https://nuclearweaponarchive.org/Usa/Tests/Dominic.html).*

The norm against nuclear detonations in space now seems to be challenged.

Assistant Secretary Stewart described Russia placing a testbed satellite in an "unusual" orbit, later identified as Cosmos-2553, launching it in February 2022 to an apogee of 2,000 km, and reportedly carrying a synthetic aperture radar primary payload [28]. This region of space is at the upper limit of low Earth orbit (LEO) and passes through the inner Van Allen radiation belt. Russia's deployment of a testbed satellite within the inner Van Allen belt suggests a CONOPS for rapidly pumping the flux density of the inner belt, raising the radiation environment in LEO. Assistant Secretary Plumb suggested this possibility in his testimony before the House Armed Services Committee in May 2024, saying that most satellites in LEO are unhardened against radiation and their electronics would quickly degrade under elevated exposure [29]. Belt pumping could energize the magnetic south Atlantic anomaly as well as expand the volumetric space of the radiation belt, creating regions of high exposure as satellites repeatedly pass through. Such a weapon would be indiscriminate, with wide-ranging effects. Timed well, interactions with solar activity could exacerbate the effect. During and before its invasion of Ukraine, Russia has signaled an increasing willingness to employ nuclear weapons for tactical purposes.

In the initial phases of the Russia-Ukraine War in 2022, Russia repeatedly raised the possibility of using nuclear weapons to deter assistance flowing to Ukraine from NATO and other international partners. Longtime Russia analyst Masha Gessen posited that Vladimir Putin, with "grandiose" self-regard, desired to establish new nuclear doctrine for the world by raising the prospect of employing tactical nuclear strikes [30]. In 2018, Putin announced several novel-concept nuclear platforms, including the Tsirkon sea-launched hypersonic missile, the Avangard hypersonic HGV, the Kinzhal air-launched hypersonic missile, the Sarmat heavy ICBM-HGV, the Burvestnik nuclear-powered cruise missile, and the Poseidon nuclear-powered underwater drone/torpedo [31]. Some of these systems have passed initial operating capability and are in service. The Kinzhal has been employed in a conventional role in the war in Ukraine. In June 2024, Russia exercised tactical nuclear strike capabilities in the southern and eastern military districts and with Belarus, signaling potential "changes to Russia's nuclear doctrine" in the future [32]. A space-based nuclear weapon would align with these other destabilizing nuclear developments, even if stationing such a weapon in orbit would violate the Outer Space Treaty.

Russia has also demonstrated a willingness to attack assets critical to space-based services. Its invasion of Ukraine began with a multipronged cyberattack against ViaSat user terminals, which disabled receivers across Ukraine and the rest of Europe [33]. Russian operational planners clearly grasped the importance of space-enabled communications and sought to disable these services during the initial phase of the war. When the Starlink communications network stepped in to provide key distributed communications services to Ukraine, its network was also attacked by cyberelectronic means, possibly via "Tobol electronic warfare systems" or "the truck-mounted Tirada-2 system" [34]. Although proliferated satellite architectures are resilient against point

> ***Belt pumping could energize the magnetic south Atlantic anomaly as well as expand the volumetric space of the radiation belt, creating regions of high exposure as satellites repeatedly pass through.***

attacks, such as via kinetic weapons, they are more susceptible to area effect attacks, such as by cyberelectronic or nuclear means. Russia's frustration with the provision of satellite services to Ukraine during the war has been evident, and targeting Starlink is a prelude for countering proliferated architectures.

PRC researchers have also expressed anxiety about Starlink's resilience. They have proposed methods for countering not only that specific network but also other proliferated constellations, such as the disaggregated satellite networks being developed by the U.S. Space Development Agency (SDA) [35]. These lines of research include use of nuclear detonation devices to target LEO satellites. PRC researchers at the Northwest Institute of Nuclear Technology (西北核技术研究所) identified ways of modulating the shape and size of a nuclear radiological cloud in LEO by adjusting detonation altitude and yield [36]. Their findings suggest ways of using a nuclear explosion to ionize upper atmospheric particles to project a radioactive plume into the orbital trajectories of satellites passing within the area. Distinct from Russia's apparent CONOPS of Van Allen belt pumping, the PRC concept uses nuclear detonation atmospheric effects to produce a more targeted effect but one that would still disrupt the LEO regime in an indiscriminate manner.

Other recent academic papers published by PRC scholars also deal with the dynamics of high-altitude nuclear explosions. The Beijing Institute of Technology (北京理工大学) is a premier research institution for weapons research and development, supervised by the Ministry of Industry and Information Technology. Researchers from the Institute's School of Mechatronical Engineering (机电学院) and the Northwest Institute of Nuclear Technology recently modeled fission debris from high-altitude nuclear detonations, describing the types of energetic particles and their interaction with each other and the Earth's magnetic field to describe the fluid dynamics of the area of effect in three dimensions and over time [37]. Unlike the Russian concept of operations suggested by USG officials, this kind of atmospheric detonation would not violate the letter of the Outer Space Treaty. While such detonations would likely violate the Comprehensive Test Ban Treaty, China, the United States, and Russia have not ratified the treaty, and it remains pending as enforceable international law [38]. SDA Director Derek Tournear has addressed the risk to the Agency's strategy of proliferated architectures: "We know, obviously, that [nuclear detonation in space] would have a major impact on our architecture and our capabilities if something like that went off in space, that it would have a major impact on the world" [39]. Calling it a "black-swan event," Tournear tacitly acknowledged that there is no readily available solution to defending against such an attack.

CONCLUSIONS

Space is becoming increasingly unstable with China and Russia's introduction of new CONOPS. While certain technical capabilities like the Hypersonic and Ballistic Tracking Space Sensor system can mitigate such risks, the fundamental danger arises from miscalculation. Cold War archetypes for understanding these technologies may offer a starting point, but both potential adversaries have evolved their operational doctrine since that era. These technologies must be understood in the context of new strategic postures that emphasize the importance of decision-dominance, speed, and autonomy for China. For

Russia, the employment of new nuclear systems must be placed in the context of its leadership: "The problem is not so much that Putin is irrational; the problem is that there is a world in which it is rational for him to move ever closer to a nuclear strike, and most Western analysts cannot comprehend the logic of that world" [30]. If Putin's Russia aims to play a spoiler role in the global commons, there is the risk of it using its technical advantages in nuclear technology, space launch, and cyber operations to destabilize international security.

China's Xi Jinping has also served an outsized role in charting the course of China's military modernization, having directed the PLA to "strive to complete the centenary goal of army building" by 2027 [40]. As China's terrestrial conflicts become increasingly kinetic and the PLA's capabilities mature, the impetus to employ new tools to change the terms of confrontation will reach an inflection point. Even if national leadership changes, China and Russia have both embarked on a trajectory that could fundamentally change future space operations. The USG and outside observers must place these new technical capabilities in the context of adversaries' strategic doctrine and leadership direction to avoid mirror-imaging and limiting the scope of possibilities under which such weapons could be employed. ■

Even if national leadership changes, China and Russia have both embarked on a trajectory that could fundamentally change future space operations.

REFERENCES

[1] Erwin, S. "Space Force General Warns of 'Window of Vulnerability' in Satellite Defense." Space News, https://spacenews.com/space-force-general-warns-of-window-of-vulnerability-in-satellite-defense/, accessed on 31 May 2024.

[2] Decker, A. "Chinese Satellites are Breaking the U.S. 'Monopoly' on Long-Range Targeting." *Defense One*, https://www.defenseone.com/threats/2024/05/new-chinese-satellites-ending-us-monopoly-ability-track-and-hit-long-distance-targets/396272/, accessed on 3 June 2024.

[3] Harpley, U. L. "DOD Official Confirms Russia Is Developing an 'Indiscriminate' Space Nuke." *Air & Space Forces Magazine*, https://www.airandspaceforces.com/dod-official-russia-indiscriminate-space-nuke/, accessed on 3 June 2024.

[4] Center for Strategic and International Studies. "The Nuclear Option: Deciphering Russia's New Space Threat." https://www.csis.org/events/nuclear-option-deciphering-russias-new-space-threat, accessed on 3 June 2024.

[5] Strobel, W. P., D. Volz, M. R. Gordon, and M. Maidenberg. "Russia Launched Research Spacecraft for Antisatellite Nuclear Weapon Two Years Ago, U.S. Officials Say." *The Wall Street Journal*, https://www.wsj.com/politics/national-security/russia-space-nuke-launched-ukraine-invasion-c4aad62e, accessed on 4 June 2024.

[6] Bruzzese, M., and P. W. Singer. "Farewell to China's Strategic Support Force. Let's Meet Its Replacements." *Defense One*, https://www.defenseone.com/ideas/2024/04/farewell-chinas-strategic-support-force-lets-meet-its-replacement/396143/, accessed on 4 June 2024.

[7] "Military Report（军事报道）." *CCTV-7*, https://tv.cctv.com/2024/05/24/VIDELoPYch8iBTfBMgv6GqPh240524.shtml?spm=C28340.P3GbPoIN6ktz.Ei1cdgvmaht5.81, accessed on 24 May 2024.

[8] Pollpeter, K., E. Anderson, J. Wilson, and F. Yang. "China Dream, Space Dream: China's Progress in Space Technologies and Implications for the United States." U.S.-China Economic and Security Review Commission, https://www.uscc.gov/research/china-dream-space-dream-chinas-progress-space-technologies-and-implications-united-states, accessed on 3 June 2024.

[9] Sevastopulo, D., and K. Hille. "China Tests New Space Capability with Hypersonic Missile." *Financial Times*, https://www.ft.com/content/ba0a3cde-719b-4040-93cb-a486e1f843fb, accessed on 31 May 2024.

[10] Sevastopulo, D. "China Conducted Two Hypersonic Weapons Tests This Summer." *Financial Times*, https://www.ft.com/content/c7139a23-1271-43ae-975b-9b632330130b, accessed on 3 June 2024.

[11] Gohd, C. "Russian Anti-Satellite Missile Test Was the First of Its Kind." Space.com, https://www.space.com/russia-anti-satellite-missile-test-first-of-its-kind, accessed on 31 May 2024.

[12] Clark, S. "Station Resumes Normal Operations, But Risk From Russian ASAT Test Continues." *Spaceflight Now*, https://spaceflightnow.com/2021/11/18/station-resumes-normal-operations-but-russian-anti-satellite-test-poses-continued-risk/, accessed on 3 June 2024.

[13] Swope, C., et al. "Space Threat Assessment 2024." Center for Strategic and International Studies, https://aerospace.csis.org/wp-content/uploads/2024/04/240417_Swope_ SpaceThreatAssessment_2024.pdf, accessed on 6 June 2024.

[14] UN Office of Disarmament Affairs. "Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies, https://treaties.unoda.org/t/outer_space, accessed on 6 June 2024.

[15] Gyűrösi, M. "The Soviet Fractional Orbital Bombardment System Program." Air Power Australia, https://www.ausairpower.net/APA-Sov-FOBS-Program.html, accessed on 31 May 31, 2024.

[16] Tirpak, J. A. "Kendall: Modernize Now to Counter China." *Air & Space Forces Magazine*, https://www.airandspaceforces.com/kendall-modernize-now-to-counter-china/, accessed on 6 June 2024.

[17] U.S. Department of Defense. "MDA, SDA Announce Upcoming Launch of the Hypersonic and Ballistic Tracking Space Sensor and Tranche 0 Satellites." https://www.defense. gov/News/Releases/Release/Article/3676902/mda-sda-announce-upcoming-launch-of-the-hypersonic-and-ballistic-tracking-space/, accessed on 6 June 2024.

[18] Erickson, A. "China's Approach to Conventional Deterrence." *Modernizing Deterrence: How China Coerces, Compels, and Deters*, edited by Roy D. Kamphausen, National Bureau of Asian Research, https://www.nbr.org/publication/modernizing-deterrence-how-china-coerces-compels-and-deters/, p. 20, accessed on 7 June 2024.

[19] Pollpeter, K. "Coercive Space Activities: The View From PRC Sources." China Aerospace Studies Institute, https://www.airuniversity.af.edu/Portals/10/CASI/documents/Research/Space/2024-02-19%20Coercive%20Space%20Activities.pdf, February 2024.

[20] Sevastopulo, D. "Chinese Hypersonic Weapon Fired a Missile Over South China Sea." *Financial Times*, https://www.ft.com/content/a127f6de-f7b1-459e-b7ae-c14ed6a9198c, accessed on 7 June 2024.

[21] Shou, X. (寿晓松) (editor). "The Science of Military Strategy (军事战略学)." *Military Science Publishing House* (军事科学出版社), p. 96, 2013.

[22] Lan, S. (兰顺正). "New Concept for Future War – 'Big Energy-centric Warfare'" (未来战争新概念——"大能量中心战"), *National Defense Reference*, https://news.yibada.com/article--235454-1-1.html, accessed on 7 June 2024.

[23] "China Expected to Overtake West in Future Air Operations With Big Data, AI: Expert." *People's Daily*, http://www.chinadaily.com.cn/china/2017-7/04/content_29987630.htm, accessed on 11 June 2024.

[24] Chen, D. Personal communication with Arms Control Expert at the Middlebury Institute for International Studies, 17 October 2021.

[25] Hu, Y. (胡延宁), B. Li (李炳彦), and S. Wang (王圣良). *Light Warfare: New Trends in World Military Revolution* (光战争: 世界军事革命新趋势). PLA Publishing House (解放军出版社), 2015.

[26] Chase, M. S., and A. Chan. "China's Evolving Approach to 'Integrated Strategic Deterrence'." *RAND*, https://www.rand.org/pubs/research_reports/RR1366.html, accessed on 10 June 2024.

[27] King, G. "Going Nuclear Over the Pacific." *Smithsonian Magazine*, https://www.smithsonianmag.com/history/going-nuclear-over-the-pacific-24428997/, accessed on 10 June 2024.

[28] Hitchens, T. "Is Russia's Cosmos 2553 Satellite a Test for a Future Orbital Nuclear Weapon?" *Breaking Defense*, https://breakingdefense.com/2024/05/is-russias-cosmos-2553-satellite-a-test-for-a-future-orbital-nuclear-weapon/, accessed on 11 June 2024.

[29] House Armed Services Committee. Testimony of Asst. Sec. John Plumb. "FY25 Budget Request for National Security Space Programs." Testimony of Asst. Sec. John Plumb, https://armedservices.house.gov/hearings/str-hearing-fy25-budget-request-national-security-space-programs, accessed on 11 June 2024.

[30] Gessen, M. "Why Vladimir Putin Would Use Nuclear Weapons in Ukraine." *The New Yorker*, https://www.newyorker.com/news/our-columnists/why-vladimir-putin-would-use-nuclear-weapons-in-ukraine, accessed on 10 June 2024.

[31] Brookes, P. "Responding to Troubling Trends in Russia's Nuclear Weapons Program." The Heritage Foundation, https://www.heritage.org/sites/default/files/2021-03/BG3601.pdf, accessed on 10 June 2024.

[32] Falconbridge, G., and L. Kelly. "Russia Broadens Tactical Nuclear Weapons Drills." *Reuters*, https://www.reuters.com/world/europe/russia-says-its-non-strategic-nuclear-drills-involve-iskander-missiles-2024-06-12/, accessed on 12 June 2024.

[33] ViaSat. "KA-SAT Network Cyber Attack Overview." https://news.viasat.com/blog/corporate/ka-sat-network-cyber-attack-overview, accessed on 10 June 2024.

[34] Horton, A. "Russia Tests Secretive Weapon to Target SpaceX's Starlink in Ukraine." *The Washington Post*, https://www.washingtonpost.com/national-security/2023/04/18/discord-leaks-starlink-ukraine/, accessed on 10 June 2024.

[35] Ren, Y. (任远桢), et al. "星链计划发展现状与对抗思考 [The Development Status of Starlink and Its Countermeasures]." *Modern Defense Technology* (现代国防技术), vol. 50, no. 2, April 2022.

[36] Liu, L. (刘李), S. Niu (牛胜利), J. Zhu (朱金辉), Y. Zuo (左应红), and H. Xie (谢红刚). "Motion Characteristics and Laws of the Debris From a Near-Space Nuclear Detonation (临近空间核爆炸碎片云运动特征与规律研究)." *Nuclear Techniques* (核技术), vol. 45, no. 10, October 2022.

[37] Peng, G. (彭国良), and J. Zhang (张俊杰). "Hydro-Magneto-PIC Hybrid Model for Description of Debris Motion in High Altitude Nuclear Explosions (基于流体-磁流体-粒子混合方法的高空核爆炸碎片云模拟)." *Acta Physica Sinica* (物理学报), vol. 70, no. 18, 2021.

[38] CTBTO Preparatory Commission. "Ending Nuclear Tests." Comprehensive Test Ban Treaty Organization, https://www.ctbto.org/our-mission/ending-nuclear-tests, accessed on 10 June 2024.

[39] Decker, A. "Russian Space Nuke Wouldn't Alter U.S. Orbital-Network Plans, Space Force Says." *Defense One*, https://www.defenseone.com/threats/2024/02/russian-space-nuke-wouldnt-alter-us-orbital-network-plans-space-force-says/394509/, accessed on 11 June 2024.

[40] Song, X. (宋歆). "Sound the Charge and Fight Hard to Win the Tough Battle—Military Representatives Fervently Commit to Achieving the Struggle Goals of the 100th Anniversary of the Founding of the Army as Scheduled (吹响冲锋号打好攻坚战——军队代表委员热议确保如期实现建军一百年奋斗目标)." *Liberation Army Daily* (解放军报), http://www.81.cn/jwtt/16290484.html, accessed on 12 June 2024.

BIOGRAPHY

DAVID D. CHEN is a senior analyst focusing on aerospace, cyber, and cross-domain emerging technologies and China's military modernization. He previously held positions at BluePath Labs and CENTRA Technology, Inc., supporting the U.S. intelligence community for over 18 years. Fluent in Mandarin Chinese, he has studied in both China and Taiwan. Mr. Chen holds a master's degree in international affairs from the University of California San Diego's School of Global Policy and Strategy.

READ MORE

If you found this publication insightful and engaging, please check out our back issues on https://dsiac.dtic.mil. We also offer similar journals, covering the cyber and homeland security spheres, which you can find at https://csiac.dtic.mil and https://hdiac.dtic.mil.

Photo source: Helena Lopes (Canva)

A BIOINSPIRED SYSTEM to Autonomously Detect Tiny, Fast-Moving Objects in Infrared Imagery

BY CHRISTOPHE BOBDA, YONG-KYU YOON, SUDEEPTO CHAKRABORTY, SUHAS CHELIAN, AND SRINI VASAN (PHOTO SOURCE: U.S. NAVY AND QBERTSTUDIO [ADOBE.STOCK.COM])

SUMMARY

Autonomously detecting tiny, fast-moving objects emitting thermal radiation in the infrared is a challenging technical problem. In addition to being fast, these targets are often dim, small, and in the presence of clutter and occlusions. Conventional detection approaches require large size, weight, and power (SWaP) systems which may introduce substantial latencies. As such, the following will be explored in this article:

- An end-to-end system composed of scene simulation,
- Sensor capture from a novel, highly sensitive micro electromechanical systems (MEMS) microbolometer,
- A readout integrated circuit (ROIC) that uses a unique saliency computation to remove uninteresting image regions, and
- Deep-learning (DL) detection and tracking algorithms.

Simulations across these modules verify the advantages of this approach compared to conventional approaches. The system's ability is estimated to detect targets at less than 5 s after a fast-moving object enters a sensor's field of view. To explore low-energy implementations of these computer vision models, DL on commercial off-the-shelf (COTS) neuromorphic hardware is also discussed.

INTRODUCTION

The autonomous detection of small and rapidly moving aerial targets is a technically demanding task. Such targets include missiles and airplanes. Missiles are often smaller than airplanes and fly faster; they are thus more challenging to detect. Newer missiles are also more maneuverable than previous missiles, posing a new threat to existing defense systems. Although these objects emit detectable amounts of infrared energy visible far away, their speed makes it difficult to image and track them. At large distances, these objects appear tiny and faint, adding to the task's complexity. Additionally, these targets are often located in cluttered and occluded environments with other objects, such as slower moving airplanes, the sun, clouds, or buildings. Furthermore, robust detection in different environmental conditions, such as day, night, cloudy days, clear days, etc., poses more challenges.

Biological vision systems have become well adapted over millennia of evolution to ignore clutter and noise, detect motion, and compress visual information in a scene. On the other hand, conventional detection approaches may have difficulty with such a scene and would require increased complexity in hardware and/or software to filter out noise and alleviate nuisance factors while increasing target sensitivity with resulting inefficiencies in size, weight, power, and cost (SWaP-C).

Biological vision systems have become well adapted over millennia of evolution to ignore clutter and noise, detect motion, and compress visual information in a scene.

There have been several attempts to create bioinspired vision systems. For example, Scribner et al. [1] created a neuromorphic ROIC with spike-based image processing. Chelian and Srinivasa simulated image processing from retinal [2] and thalamic [3] circuits in the spiking domain under the Defense Advanced Research Projects Agency (DARPA) Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) program for noise suppression, ratios of spectral bands, and early motion processing. (Their work was informed by studies in the rate-coded domain [4, 5].) However, these works do not consider detection or tracking.

Artificially mimicking biological vision systems wherever feasible was explored in the current work to overcome the challenges to conventional systems described above. The imaging system includes everything from the optics taking in the scene to the final processor outputting target reports. The system components that would be implemented in hardware are a MEMS microbolometer, a ROIC which

uses a unique saliency computation to remove uninteresting image regions and increase overall system speed, and DL detection and tracking models.

The performance of each component and across the whole system is estimated via tools that include simulated images and videos. The feasibility of COTS neuromorphic hardware to implement DL with less energy than graphical processing units (GPUs) is also described. The bioinspired system's components are shown in Figure 1.

METHODS

There are five main thrusts to the design effort:

1. A scene simulation,
2. A MEMS microbolometer,
3. A ROIC that uses a unique saliency computation to remove uninteresting image regions and increase detection speed (referred to as Hierarchical Attention-oriented, Region-based Processing or HARP [6]),
4. DL detection and tracking models, and
5. Neuromorphic computing.

Additionally, end-to-end system evaluation and evaluation of each component are performed.

Scene Simulation

To detect and track tiny, fast objects in cluttered and noisy scenes, training data is needed. However, there is a scarcity of such datasets available to the public. For this reason, harnessing the power of synthetic datasets was started. The work of Park et al. [7], for example, illustrates this approach. Due to the scarcity of real hyperspectral images of contraband on substrates, synthetic hyperspectral images of contraband substances were created on substrates using infrared spectral data and radiative transfer models. For small and rapidly moving objects, a small publicly available dataset was used first because it had single-frame infrared targets with high-quality annotations, which can be used for detection modules. Animating targets for tracking modules was also explored. There are other infrared datasets of aerial targets such as unmanned aerial vehicles, but these tend to occupy more pixels per frame than the dataset used in the present work.

MEMS Microbolometer

According to Dhar and Khan [8], detection ranges are sensitive to temperature and relative humidity variations, and long-wave infrared (LWIR) ranges depend more upon these variations than mid-wave infrared (MWIR) ranges. Because of

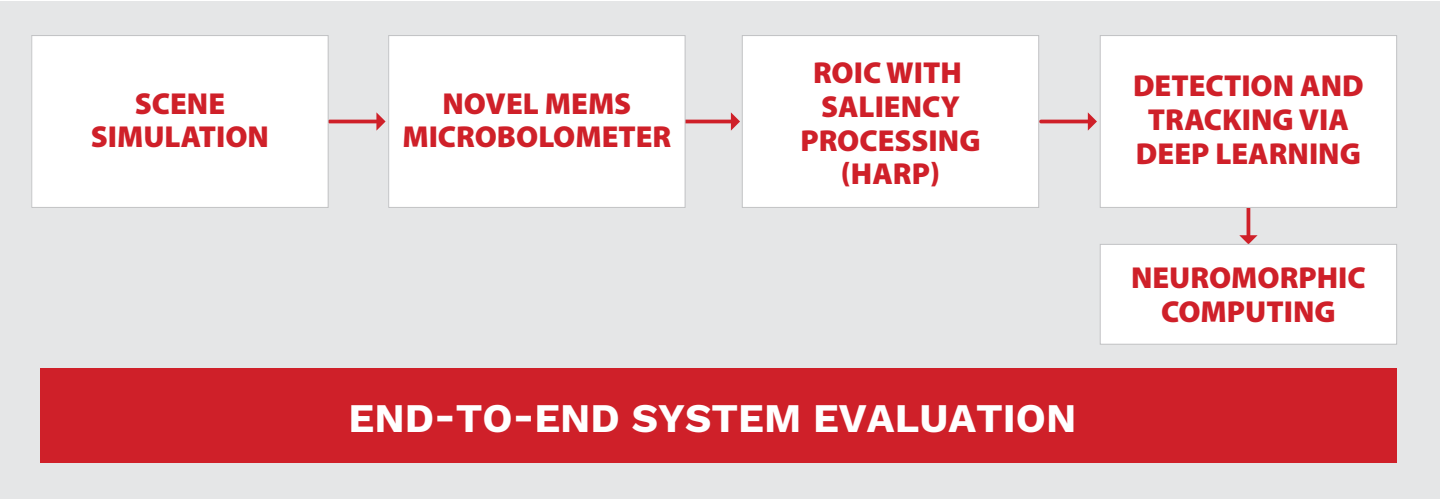


Figure 1. Bioinspired System for Autonomous Detection of Tiny, Fast Moving Objects in Infrared Imagery. Such a System Would Be Dramatically Smaller, Lighter, Less Power-Hungry, and More Cost-Effective Than Traditional Systems *(Source: C. Bobda, Y.-K. Yoon, S. Chakraborty, S. Chelian, and S. Vasan).*

this, on average, MWIR tends to have better overall atmospheric transmission compared to LWIR in most scenarios. Therefore, in this work, an MWIR MEMS microbolometer sensor that is highly selective in its spectral range was designed. Prior work in this area includes that of Dao et al. [9].

In a microbolometer, infrared energy strikes a detector material, heating it and changing its electrical resistance. This resistance change is measured and processed into temperatures which can be used to create an image.

There are commercially available non-MEMS microbolometers, but they have poorer sensitivity because thermal isolation is not as good as MEMS-based implementations. This is because in a MEMS device, there is a physical (e.g., air) gap between the detector and the substrate. The MEMS-based approach can increase the effective absorbing area of the sensor with complex structures and increase its responsivity. Yoon et al. [10] demonstrated multidirectional ultraviolet lithography for several complex three-dimensional (3-D) MEMS structures which can be used to create a MEMS microbolometer (Figure 2).

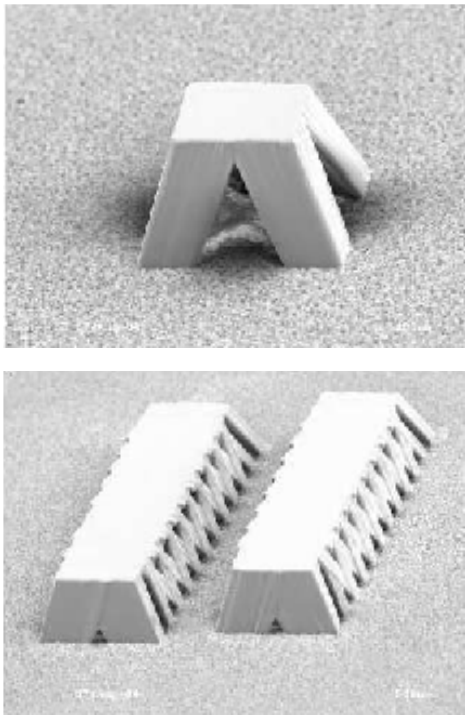


Figure 2. Previously Demonstrated Complex 3-D MEMS Structures Which Can Be Used for Microbolometers *(Source: Yoon et al. [10]).*

> ***The MEMS-based approach can increase the effective absorbing area of the sensor with complex structures and increase its responsivity.***

Hierarchical Attention-Oriented, Region-Based Processing (HARP)

Event-based HARP is a ROIC design that suppresses uninteresting image regions and increases processing speed. It was developed by Bhowmik et al. [6, 11]. The work draws inspiration from and is a simplified abstraction of the hierarchical processing in the visual cortex of the brain where many cells respond to low-level features and transmit the information to fewer cells up the hierarchy where higher-level features are extracted [12].

The main idea is illustrated in Figure 3a. Figure 3b shows an architecture diagram. In the first layer, a pixel-level processing plane (PLPP) provides early feature extraction such as edge detection or image sharpening. Several pixels are then grouped into a region. In the next stage, the structure-level processing plane (SLPP) produces intermediate features such as line or corner detection using a region processing unit (RPU). For a region, processing is only activated if its image region is relevant. Image relevance is computed based on several metrics, such as predictive coding in space or time, edge detection, and measures of signal-to-noise ratio (SNR). The RPU also sends feedback signals to the PLPP using an attention module. If image relevance is too low, pixels in the PLPP halt their processing using a clock gating method. Thus, like the Dynamic Vision System (DVS) [13], uninteresting image regions like static fields would not be processed and would save energy and time. On the other hand, unlike the DVS, HARP directly provides intensity information and could differentiate between extremely hot targets and moderately hot targets.

Finally, at the knowledge inference processing plane (NIPP), global feature processing such as with a convolutional neural network (CNN) is performed. NIPP implementations are described in the next subsection. Because only interesting image regions are processed—not all pixels—speed

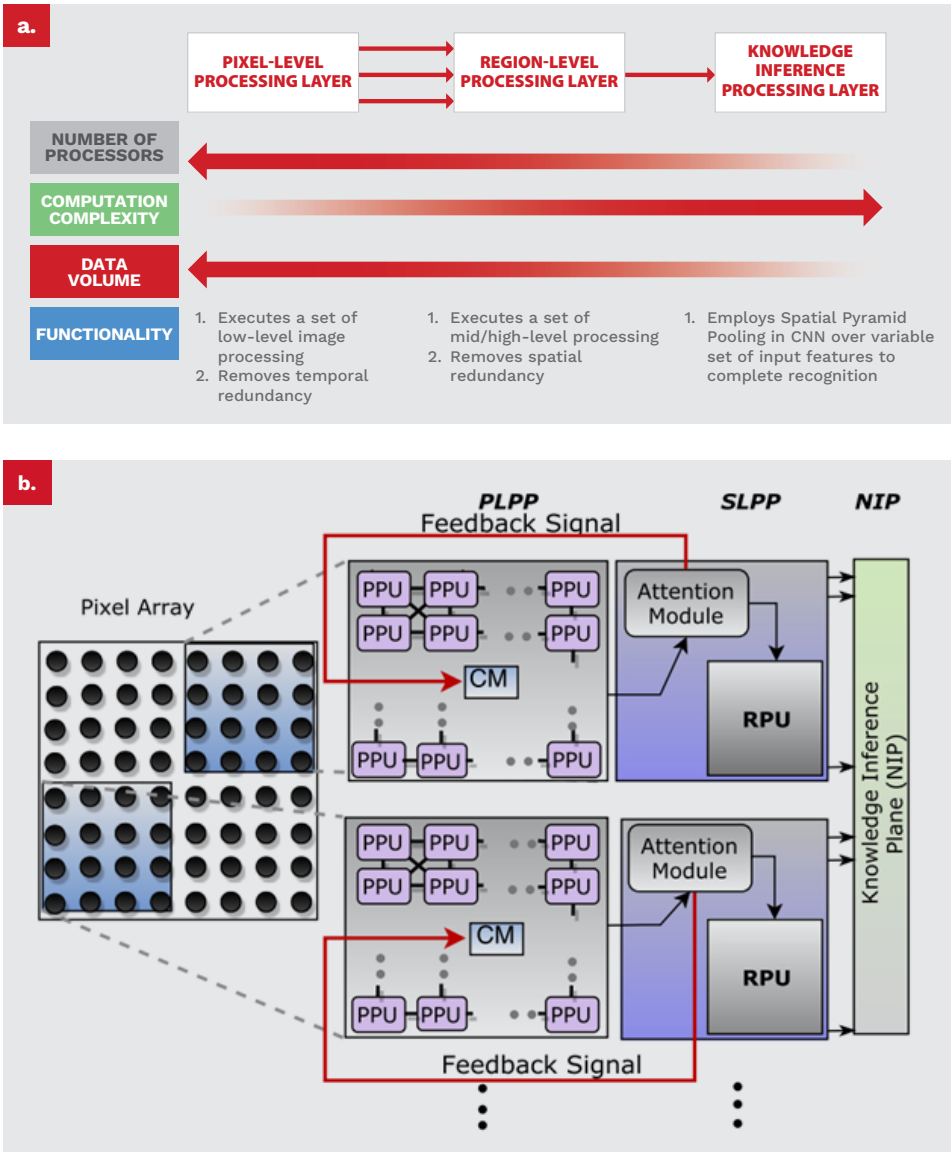


Figure 3. HARP Illustrated as (a) a Conceptual Diagram and (b) an Architecture Diagram. HARP Is Used to Remove Uninteresting Image Regions to Find Targets Faster *(Source: Bhowmik et al. [11, 12]).*

and power savings advantages are realized. Savings can be up to 40% for images with few salient regions. Because the hardware is parallel, latency is more or less independent of image size.

Detection and Tracking

Detection DL models such as You Only Look Once (YOLO) [14] and U-Net [15] were tried. Parameter searches over the number of layers, loss function (e.g., Tversky and Focal Tversky), and data augmentation (flipping and rotating images) to increase performance were also conducted.

For tracking, a fully convolutional Siamese network was chosen [16] on the videos derived from the previously mentioned, publicly available dataset. Siamese networks are relatively simple to train and computationally efficient. One input of the twin network starts with the first detection, and the other input is the larger search image. The twin network's job is to locate the exemplar image inside the search image. In subsequent frames, the first input is updated with subsequent detected targets.

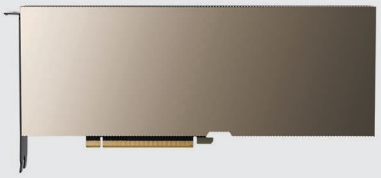
Neuromorphic Computing

Neuromorphic computing offers low SWaP-C alternatives to larger central processing unit (CPU)/GPU systems. DL detection and tracking modules can be run on neuromorphic computers to exploit these advantages. The alternatives are compared in Table 1.

Neuromorphic computing has been used in several domains, including computer vision and cybersecurity. In a computer vision project, classification of contraband substances using LWIR hyperspectral images has been demonstrated in a variety of situations, including

> ***Neuromorphic computing offers low SWaP-C alternatives to larger central processing unit (CPU)/GPU systems.***

Table 1. Comparison of CPU/GPU and Neuromorphic Compute Platforms. Neuromorphic Platforms Offer Lower SWaP-C *(Source for Left Image, Wikimedia; Right, Intel)*

CPU/GPU PLATFORMS	SWaP-C	NEUROMORPHIC PLATFORMS
NVIDIA A100 PCIE 4.0 DUAL SLOT		USB KEY FORM FACTOR FOR INTEL LOIHI 1
		
26.7 L x 11.2 H x 3.5 W	Size (cm)	5.1 L x 1.3 H x 0.6 W
1674	Weight (g)	~180
250	Power (W)	1
~$16,000	Cost (USD)	$50 (est.)

varied backgrounds, temperatures, and purities. Full-precision GPU and the BrainChip Akida compatible models gave promising results [7]. In a cybersecurity project, accurate detection of eight attack classes and one normal class was demonstrated in a highly imbalanced dataset. First-of-its-kind testing was chosen with the same network on full-precision GPUs and two neuromorphic offerings—the Intel Loihi 1 hardware and the BrainChip Akida 1000 [17]. Updates have since been made to this work, such as a smaller, more accurate neural network and the use of the BrainChip chip (not just a software simulator) and Intel's new DL framework Lava for the Loihi 2 chip [18, 19].

End-to-End System Evaluation

For end-to-end system evaluation, speed was the primary focus; however, power consumption was also estimated based on novel simulations or previous work. Each module had their own metrics, e.g., noise equivalent temperature difference (NETD) for the MEMS microbolometer or intersection over union (IoU) for detection and tracking.

RESULTS

Results from the five primary thrusts—scene simulation, MEMS microbolometer, HARP, detection and tracking, and neuromorphic computing—and end-to-end system evaluation are given. The system's ability is estimated to detect targets at less than 5 s after a fast-moving object enters a sensor's field of view. The system design would have a dramatically smaller SWaP-C envelope than conventional systems.

Scene Simulation

Images from a publicly available dataset were used for detection and tracking. Separate images were used for training vs. testing. Example images are shown in Figures 4 a and c. These infrared images have tiny targets in cluttered scenes; thus, they are a good starting point for the current application domain.

MEMS Microbolometer

COMSOL was used to simulate the microbolometer. A 10× improvement in NETD over existing vanadium oxide (VOx) microbolometers was achieved, and response times were approximately 3× less! This is because of the better thermal isolation of the MEMS microbolometer from its support structures and the unique choices of materials.

HARP

HARP was able to remove uninteresting image regions to decrease detection speed. In Figure 4, input images from a publicly available dataset and salient regions are shown in two pairs. The top pair retains the

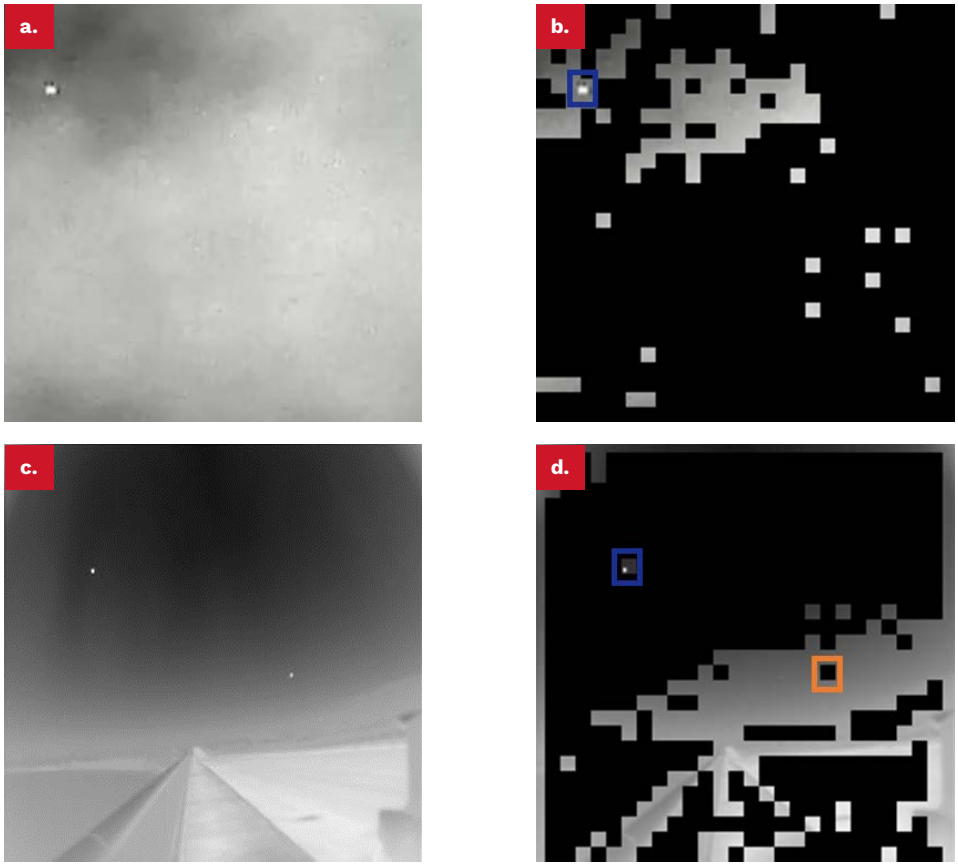


Figure 4. Infrared Input Image Regions (a and c) and Salient Regions (b and d) (Black Regions Are Removed by HARP). Image (a) Has a True Positive, and Image (d) Has a True Positive and False Negative (Blue Bounding Boxes Are Accurate Detections, and Orange Is a Missed Region) *(Source: github.com/YimianDai/sirst)*.

target—the bright region in the upper left. In the bottom pair, HARP retains one target in the upper left but omits another one to the right of the center of the image. This initial testing was carried out at a preliminary level. Further parameter tuning is necessary for more refined target extraction.

Detection and Tracking

For detection, U-Net worked better than YOLO networks probably because targets were tiny. U-net detected targets in different sizes and varied backgrounds. Across dozens of images, the average probability of detection was 90% and probability of false alarm was 12%. Further hyperparameter tuning will improve these results.

For tracking, a fully convolutional Siamese network method tracks targets in different sizes and varied backgrounds. Across about a dozen videos, the average IoU was 73%. Further hyperparameter tuning will lead to improved performance.

Neuromorphic Computing

In addition to prior work in neuromorphic computer vision to detect contraband materials [7], recent work from Patel et al. [20] implemented U-Net on the Loihi 1 chip. They achieved 2.1× better energy efficiency than GPU versions, with only a 4% drop in mean IoU. However, processing times were ~87× slower. Their code was not released, so it was not possible to use their implementation with the current dataset. The release of the more powerful BrainChip and Intel Loihi chips combined with other optimizations will decrease processing times. For example, preliminary pedestrian and vehicle detection experiments were conducted with a YOLO v2 model on the BrainChip AKD1000 chip using red-green-blue color images. With a test set of 100 images from the PASCAL Visual Object Classes dataset, there was a small decrease in mean average precision of 4% compared to full-precision GPU models, but processing times were ~3× slower.

End-to-End System Evaluation

The detection time was estimated to be 4 to 5 s once a faraway object entered the sensor's field of view. The following times were included: (1) microbolometer response and readout is estimated at 50 ms via COMSOL; (2) HARP at less than 1 s [12] (estimated at 8 μs to 12.5 ns, with an application-specific integrated circuit [ASIC] or field-programmable gate array [FPGA], respectively);

(3) detection and tracking at 2 s (DL algorithms operate at 30 fps or more); and (4) other latencies at 1 s.

For power, the system's estimated power draw is less than 10 mW for the microbolometer via COMSOL and 2 mW to 7 W for HARP with an ASIC or FPGA, respectively, based on Yoon et al. [10]. For detection and tracking, a tradeoff between power and speed is apparent. For targeting applications, GPUs are a better choice and would draw ~250 W, per Table 1. With a 50% margin of safety, the system would draw ~375 W. For less time-critical applications, neuromorphic processors are adequate and would draw ~1 W. With a 50% margin of safety, the system would draw ~1.5 W.

DISCUSSION

Overall, a bioinspired system to autonomously detect tiny, fast-moving objects in infrared imagery such as missiles and aircraft was presented. Detecting and tracking tiny fast-moving objects is vitally important to several U.S. Department of Defense customers. As threats continue to evolve, existing systems need updating and/or replacing. Unlike existing systems, the approach here offers considerable SWaP-C advantages of bioinspired computing in several stages. Civilian applications include tracking launch or reentry vehicles, which is of interest to the National Aeronautics Space Administration (NASA) as well as private sector companies. The current design encompasses every aspect from capturing the target through optics to the final processor outputting target reports. The system design effort has five main thrusts: (1) scene simulation, (2) a novel highly sensitive MEMS microbolometer, (3) HARP—a ROIC which uses a unique saliency computation to remove uninteresting image regions, (4) DL detection and tracking algorithms, and (5) neuromorphic computing. This system can detect targets within 5 s of a fast-moving object entering the sensor's field of view. Further enhancements are possible, and some improvements that can be realized in future work are described next.

Detecting and tracking tiny fast-moving objects is vitally important to several U.S. Department of Defense customers.

Scene Simulation

More realistic trajectories like those approximated by piecewise polynomial curves or physics-based models could be used here. Furthermore, the movement of several objects in the same scene could be simulated. This would present a significant challenge to tracking algorithms.

MEMS Microbolometer

Simulating and fabricating 10×10 modular pixel array samples with a 20-μm pitch would provide valuable performance characterization information. The geometry of the microbolometer and the fabrication processes can be optimized for performance. For example, thinner, longer legs would yield better sensitivity but must be compliant with manufacturing and operational constraints. Some studies in this area are referenced—Erturk and Akin [21] illustrate absorption as a function of microbolometer thickness, and Dao et al. [9] model resistivity as a function of temperature and fabrication process. Additional fabrication parameters include temperature, curing steps, and deposition speed. Simulation tools like COMSOL can be used to theoretically optimize design characteristics; however, an initial fabrication run is necessary to validate these tools.

Lastly, even though microbolometer technology provides uncooled infrared thermal detection, microbolometer performance is generally limited by low sensitivity, high noise, slow video speed, and lack of spectral content. Ackerman et al. [22] and Xue et al. [23] have demonstrated fast and sensitive MWIR photodiodes based on mercury telluride (MgTe) colloidal quantum dots that can operate at higher temperatures, including room temperatures. With sufficient maturing, this work may be used as the sensing device in lieu

of microbolometers. Cooling the microbolometer is another option to increase sensitivity.

HARP

Work with FPGAs and simulated ASICs is expected to continue. Furthermore, to better deal with dim moving objects, SLPPs can be tuned to be more sensitive to motion than contrast differences. Because saliency is the weighted sum of several feature vectors, this would make the weight of the former larger than the latter. This will also help disambiguate targets from clutter, such as slower moving airplanes or the sun.

Detection and Tracking

Single-frame detection techniques are used here. Alternatively, multiple frames can be leveraged to produce more accurate results, although it will also add latency. Another strategy could be to increase the integration time of the cameras so that targets would appear as streaks. These streaks would be larger features to detect and track.

Neuromorphic Computing

Work in the domain continues using BrainChip and Intel products in cybersecurity and other domains. BrainChip and Intel have released the second generation of their chips—seeing what improvements in speed can be gained from these products will be interesting.

End-to-End System Evaluation

More work would be helpful. Examples include calculating the SNR of the microbolometer; tabulating the results from HARP across several different images; and characterizing detection performance by scene type (e.g., day vs. night and cloudy vs. clear). For tracking, track length and uncertainty quantification can be explored. Furthermore, power, speed, and interface elements of hardware components could be further detailed.

CONCLUSIONS

By leveraging prior experience and the current work in synthetic data generation and developing MEMS devices, specialized ROICs, DL, and neuromorphic systems, this work can continue further and achieve new heights in MWIR imagery and autonomous detection and tracking in infrared imagery. This includes fast-moving pixel and subpixel object detection and tracking. Bioinspired computing can produce tremendous savings in SWaP-C, as was illustrated in Table 1 and described throughout this article.

In the short term, the microbolometer can be tested with pixel array samples, the HARP ROIC can be implemented on an FPGA, and the detection and tracking systems can be implemented on GPUs. In the long term, the microbolometer and ROIC can be bonded together as a flip chip and the GPUs replaced with neuromorphic ASICs. This will yield improvements in power and latency, allowing this system to be deployed on large or small platforms. ■

Bioinspired computing can produce tremendous savings in SWaP-C.

REFERENCES

[1] Scribner, D., T. Petty, and P. Mui. "Neuromorphic Readout Integrated Circuits and Related Spike-Based Image Processing. " The 2017 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, pp. 1–4, 2017.

[2] Chelian, S., and N. Srinivasa. "System, Method, and Computer Program Product for Multispectral Image Processing With Spiking Dynamics." U.S. Patent 9,111,182, 2015.

[3] Chelian, S., and N. Srinivasa. "Neuromorphic Image Processing Exhibiting Thalamus-Like Properties." U.S. Patent 9,412,051, 2016.

[4] Carpenter, G. A., and S. Chelian. "DISCOV (DImensionless Shunting COlor Vision): A Neural Model for Spatial Data Analysis." *Neural Networks*, vol. 37, pp. 93–102, 2013.

[5] Chelian, S. "Neural Models of Color Vision with Applications to Image Processing and Recognition." Boston University, https://www.proquest.com/openview/41ac5c06c80fffd0631e94c26245f00d/1, 2006.

[6] Bhowmik, P., M. Pantho, and C. Bobda. "HARP: Hierarchical Attention Oriented Region-Based Processing for High-Performance Computation in Vision Sensor." *Sensors*, vol. 21, p. 1757, 2021.

[7] Park, K. C., J. Forest, S. Chakraborty, J. T. Daly, S. Chelian, and S. Vasan. "Robust Classification of Contraband Substances Using Longwave Hyperspectral Imaging and Full Precision and Neuromorphic Convolutional Neural Networks." *Procedia Computer Science*, vol. 213, pp. 486–495, 2022.

[8] Dhar, V., and Z. Khan. "Comparison of Modeled Atmosphere-Dependent Range Performance of Long-Wave and Mid-Wave IR Imagers." *Infrared Physics & Technology*, vol. 51, pp. 520–527, 2008.

[9] Dao, T. D., A. T. Doan, S. Ishii, T. Yokoyama, O. H. S. Ørjan, D. H. Ngo, T. Ohki, A. Ohi, Y. Wada, C. Niikura, S. Miyajima, T. Nabatame, and T. Nagao. "MEMS-Based Wavelength-Selective Bolometers." *Micromachines*, vol. 10, p. 416, 2019.

[10] Yoon, Y. K., J. H. Park, and M. G. Allen. "Multidirectional UV Lithography for Complex 3-D MEMS Structures." *J. Microelectromech. Syst.*, vol. 15, pp. 1121–1130, 2006.

[11] Bhowmik, P., M. Pantho, M. Asadinia, and C. Bobda. "Design of a Reconfigurable 3D Pixel-Parallel Neuromorphic Architecture for Smart Image Sensor." The 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 673–681, 2018.

[12] Bhowmik, P., M. Pantho, and C. Bobda. "Bio-Inspired Smart Vision Sensor: Toward a Reconfigurable Hardware Modeling of the Hierarchical Processing in the Brain." *J. Real-Time Image Process.*, vol. 18, pp. 1–10, 2021.

[13] Lichtsteiner, P., C. Posch, and T. Delbrück. "A 128×128 120 dB 15 μs Latency Temporal Contrast Vision Sensor." *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.

[14] Redmon, J., S. Divvala, R. Girshick, and A. Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." The 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.

[15] Ronneberger, O., P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." The 2015 IEEE 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241, 2015.

[16] Xu, Y., Z. Wang, Z. Li, Y. Yuan, and G. Yu. "SiamFC++: Towards Robust and Accurate Visual Tracking With Target Estimation Guidelines." *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 7, pp. 12549–12556, 2020.

[17] Zahm, W., T. Stern, M. Bal, A. Sengupta, A. Jose, S. Chelian, and S. Vasan. "Cyber-Neuro RT: Real-Time Neuromorphic Cybersecurity." *Procedia Computer Science*, vol. 213, pp. 536–545, 2022.

[18] Zahm, W., G. Nishibuchi, S. Chelian, and S. Vasan. "Neuromorphic Low Power Cybersecurity Attack Detection." The 2023 ACM International Conference on Neuromorphic Systems, 2023.

[19] Zahm, W., G. Nishibuchi, A. Jose, S. Chelian, and S. Vasan. "Low Power Cybersecurity Attack Detection Using Deep Learning on Neuromorphic Technologies." *Journal of the Defense Systems Information Analysis Center*, to be published, 2024.

[20] Patel, K., E. Hunsberger, S. Batir, and C. Eliasmith. "A Spiking Neural Network for Image Segmentation." arXiv preprint arXiv:2106.08921, 2021.

[21] Erturk, O., and T. Akin. "Design and Implementation of High Fill-Factor Structures for Low-Cost CMOS Microbolometers." The 2018 IEEE Micro Electro Mechanical Systems (MEMS), pp. 692–695, 2018.

[22] Ackerman, M., X. Tang, and P. Guyot-Sionnest. "Fast and Sensitive Colloidal Quantum Dot Mid-Wave IR Photodetectors." *ACS Nano*, vol. 12, pp. 7264–7271, 2018.

[23] Xue X., M. Chen, Y. Luo, T. Qin, X. Tang, and Q. Hao. "High-Operating-Temperature Mid-Infrared Photodetectors via Quantum Dot Gradient Homojunction." *Light Sci. Appl.*, vol. 12, 2023.

BIOGRAPHIES

CHRISTOPHE BOBDA is a professor with the Department of Electrical and Computer Engineering at the University of Florida (UF). Previously, he joined the Department of Computer Science at the University of Erlangen-Nuremberg in Germany as a postdoc and was later appointed assistant professor at the University of Kaiserslautern. He was also a professor at the University of Potsdam and leader of The Working Group Computer Engineering. He received the Best Dissertation Award 2003 from the University of Paderborn for his work on synthesis of reconfigurable systems using temporal partitioning and temporal placement. Dr. Bobda holds a licence in mathematics from the University of Yaounde, Cameroon, and a diploma and Ph.D. in computer science from the University of Paderborn in Germany.

YONG-KYU YOON is a professor and a graduate coordinator with the Department of Electrical and Computer Engineering at UF, where his current research interests include microelectromechanical systems, nanofabrication, and energy storage devices; metamaterials for radio frequency (RF)/microwave applications; micromachined millimeter wave/terahertz antennas and waveguides; wireless power transfer and telemetry systems; lab-on-a-chip devices; and ferroelectric materials for memory and tunable RF devices. He was a postdoctoral researcher with the Georgia Institute of Technology, an assistant professor with the Department of Electrical Engineering at The State University of New York, Buffalo, and an associate professor at UF. He has authored over 200 peer-reviewed publications and was a recipient of the National Science Foundation Early Career Development Award and the SUNY Young Investigator Award. Dr. Yoon holds a Ph.D. in electrical and computer engineering from the Georgia Institute of Technology.

SUDEEPTO CHAKRABORTY has over 20 years of diverse experience in science and technology, including CNNs, image processing, computational electromagnetics, and meta-optics. He has worked on a wide variety of projects at Hewlett-Packard, Broadcom, Stanford, and Fujitsu labs. Dr. Chakraborty holds a Ph.D. in electrical engineering from Stanford University.

SUHAS CHELIAN is a researcher and machine-learning (ML) engineer. He has captured and executed projects with several organizations like Fujitsu Labs of America, Toyota (Partner Robotics Group), Hughes Research Lab, DARPA, the Intelligence Advanced Research Projects Agency, and NASA. He has over 31 publications and 32 patents demonstrating his expertise in ML, computer vision, and neuroscience. Dr. Chelian holds dual bachelor's degrees in computer science and cognitive science from the University of California San Diego and a Ph.D. in computational neuroscience from Boston University.

SRINI VASAN is the president and CEO of Quantum Ventura Inc. and CTO of QuantumX, the research and development arm of Quantum Ventura Inc. He specializes in artificial intelligence (AI)/ML, AI verification validation, ML quality assurance and rigorous testing, ML performance measurement, and system software engineering and system internals. Mr. Vasan studied management at the MIT Sloan School of Management.

DETECTING AND DEFENDING AGAINST MALICIOUS ATTACKS TO SHIP SENSORS

BY R. GLENN WRIGHT (PHOTO SOURCE: KEDEK CREATIVE [STOCK.ADOBE.COM], YANKOVICH [STOCK.ADOBE.COM], AND LUNARTS STUDIO [CANVA])

SUMMARY

This article describes research to detect phenomena that may degrade or disrupt the performance of an otherwise fully functional sensor. Also examined are methods to potentially mitigate the effects of sensor degradation and develop effective countermeasures that enable naval vessels and unmanned vehicles to continue their missions with degraded capabilities. An emphasis is placed on protecting sensors and sensor systems from cyberattacks and physical attacks aimed at accessing, changing, and/or destroying sensitive and essential information. Sensor degradation that may be attributed to natural causes is also considered, along with the importance of distinguishing between nefarious intent and natural occurrence in mission performance and execution.

BACKGROUND

Tactical advantages achieved by ships and unmanned vehicles are easily eroded by degrading sensor capabilities through jamming and exploiting vulnerabilities as well as environmental conditions inherent to sensor use that can compromise sensor performance, corrupt data, and reduce precision, functionality, utility, and overall effectiveness. Workarounds to prevail over such threats and accomplish mission objectives are effective only to the extent their characteristics can be accurately modeled and until adversaries modify their attack methods, forcing subsequent changes to countermeasures [1].

Methods described include the use of machine learning (ML) that can greatly assist in identifying attack characteristics from the frequency, time, and spatial-domain perspectives and train smart sensors to detect and overcome their effects. Fusion of complementary data from diverse sensors can help to maintain situational awareness under degraded sensor conditions. Deep-learning artificial intelligence (AI) also provides a means to enhance long-term sensor immunity and response to attack scenarios and harden future sensor designs against malicious activity and nefarious influences.

Current, completed research has demonstrated the effectiveness of this approach with various surface, undersea, and space-based sensors in an operational environment using a maritime testbed [2]. Conclusions reached, lessons learned, and applications to manned ships, unmanned and autonomous vehicles, and future technology directions are discussed.

SENSOR DEGRADATION DEFINED

Sensor degradation can be defined as deterioration or loss of performance in sensors over time due to various factors such as usage, environmental conditions, and aging. This generally involves gradual deterioration over relatively long periods of time from exposure to environmental conditions, including extreme temperatures and light, exposure to sea water and corrosive chemicals, mechanical wear of parts, electrical stress from excess voltage, current and electromagnetic interference, calibration drift, and material deterioration [3, 4]. Similar effects where partial sensor functionality remains may also occur rapidly and be permanent, long lasting, intermittent or periodic, or temporary, depending upon the cause. For external sensors, these include misalignment and partial blockage of view from being hit with debris as well as from large wave strikes, groundings, and allision that cause physical damage to sensors and transducers. Buildup of contaminants can occur from combustion of nearby vessel components, intense heat from fire, internal overheating caused by lack of ventilation, glancing strikes by laser weapons, and even the application of chemical agents designed to quickly render sensors blind or otherwise inoperable.

Internal sensor performance degradation can originate from failures within the various propulsion, electrical, hydraulic, communication, navigation, and other systems within the vessel or vehicle. A broad subset of navigation sensors was considered in this research, including radar, lidar, sonar, Automated Identification System (AIS), visible and infrared light, inertial navigation, and the Global Positioning System (GPS). Engineering sensors include vehicle temperature, pressure, flow, level, voltage, current, power, and other sensing and measurement characteristics commonly associated with these systems.

However, a critical cause of sensor degradation can be attributed to malicious activity on the part of adversaries from hacking and cyberattacks emulating many manmade and naturally occurring modes. This is the focus of this article.

SCOPE OF THE PROBLEM

Sensors on naval vessels extend the human senses to enhance internal and external situational awareness and supplement the intelligence of trained sailors and mission specialists

in performing their duties. When integrated into combat management systems (CMSs), these sensors provide multiple functions for detection, identification, command and control (C2), and decision-making. Unmanned and autonomous vehicle sensors provide onboard decision-making capabilities with comprehensive insight into states, conditions, and characteristics within the vehicle and in the surface and subsea operational environment. Natural events and adversarial activities resulting in physical destruction or damage that can render sensors useless and unable to perform their assigned functions are relatively easy to detect and identify. However, the subtleties of sensor degradation are often much more difficult to discern since high-resolution sensors and error detection and correction algorithms can compensate for degraded performance to a large degree.

Malicious actions can also significantly degrade sensor performance by tricking them into seeing things that are not there, not seeing things that are there, which are critical to mission success, and changing specific threat characteristics to disguise their true meaning and render them benign in appearance only.

Examples of successful attempts at sensor degradation by adversaries are many. Most notable was the infiltration of computer systems by the Stuxnet computer virus that changed sensor signals from centrifuges (as shown in Figure 1a) to make operational parameters appear nominal yet they were spinning far above their normal speed and tearing themselves apart [5]. Figures 1b and c illustrate where similar methods can be used to overcome safety features that restrict uncontrolled acceleration of a ship or vehicle engine, a radar antenna turning motor, motorized camera, satellite tracking antenna, or a gyroscope used to maintain heading.

Other examples include placing false echoes onto a radar screen [6], changing sonar target characteristics to obscure or disguise known threats [7], corrupting civilian and military layers of Electronic Chart Display and Information Systems (ECDISs) with false data and soundings to hide hazards to navigation [8], and manipulating AIS identities to misrepresent vessel identity, location, or intentions [9]. However, two of the more sinister threats include spoofing GPS signals on cue to ground a vessel in the middle of a critical waterway [10] and hijacking ship and vehicle controls [11, 12]. Properly timed and triggered, either of these scenarios could result in significant loss of life and catastrophic damage to vessels and critical infrastructure.

VULNERABILITY TO CYBERATTACKS

The heart of all modern naval and commercial vessels is the Integrated Bridge System (IBS), defined as a combination of interconnected systems allowing centralized access to sensor information or C2 from workstations,

a. Enriched Uranium Centrifuges

b. Ship Radar Antenna

c. Motorized Cameras

Figure 1. Examples of Sensor Motors That Can Be Compromised *(Source: [a] Alamy, Inc. and [b, c] R. Glenn Wright).*

with the aim of increasing safe and efficient ship's management by suitably qualified personnel [13]. Analogous systems exist onboard surface and underwater unmanned and autonomous vehicles. The Safety of Life at Sea Convention [14] states that IBS "shall be so arranged that failure of one sub-system is brought to immediate attention of the officer in charge of the navigational watch by audible and visual alarms and does not cause failure to any other sub-system. In case of failure in one part of an integrated navigational system, it shall be possible to operate each other individual item of equipment or part of the system separately" [15].

However, as previously mentioned, sensor degradation involves inhibiting or reducing the performance of an otherwise fully functional sensor without causing outright failure. This nuance will generally cause built-in-test and parametric testing methods to overlook characteristics not specifically defined as representing failure. Degraded sensors may also propagate errors throughout interconnected systems. Similar conditions exist for CMSs that acquire and display a ship's own sensor information with geographical data for C2, management, manipulation, and tactical information display.

IBSs and CMSs deliver unprecedented capabilities for vessel operation, navigation, and warfighting. However, these same network architectures also can make ships more vulnerable to cyberattacks via various flaws, potential exploits, and weaknesses in system hardware, software, administration, and organizational policies or processes [16]. With the increasing complexity of networked systems, comprehensive and thorough assessments of such vulnerabilities in terms of hardware, software, and human capabilities are essential.

The U.S. Navy is presently developing resources to take advantage of fleetwide connectivity through its Project Overmatch initiative. Integral to this is an Integrated Combat System (ICS) consisting of networked multidomain assets, including sensors, and comprising a common architecture across surface naval assets that all ships can pull from to conduct missions alone or in a group [17]. A February 2023 statement by RADM Fred Pyle at an American Society of Naval Engineers Conference in Arlington, VA, indicated that ICS will enable a decision-maker in the fleet, strike group, maritime operations center, or another ship to pair any sensor to any shooter [18]. Nevertheless, any such system can also pair any sensor to a wide range of hackers.

SENSOR SYSTEM PROTECTION FROM CYBERATTACKS

Much of the literature discussing cybersecurity has traditionally discussed the attack surface against which cyberattacks are made and the need to reduce this to as small a footprint as possible. One definition describes it as the total number of all possible entry points for unauthorized access into any system, including vulnerabilities and endpoints that can be exploited [19]. This also represents the entire area of a ship's networks and sensor systems susceptible to hacking. The smaller the attack surface, the easier it is to protect. However, naval system sensor infrastructure is already massive. As new technologies are introduced, the attack surface continues to expand. Also, with the increasing use of Internet of Ships (IoS) devices and sensors, the attack surface has expanded exponentially.

The U.S. Department of Defense (DoD) Zero Trust Strategy and Roadmap anticipates current and future cyberthreats and attacks that go beyond the traditional perimeter defense approach [20]. Rather than looking at the attack surface from a high level, in Zero Trust, the exact nature of what is needed to be

The U.S. Department of Defense Zero Trust Strategy and Roadmap anticipates current and future cyberthreats and attacks that go beyond the traditional perimeter defense approach.

protected is defined as the protect surface, which is the smallest possible reduction of the attack surface. It is defined based upon the protected data, application usage of sensitive data, asset vulnerability, and services that can be exploited to disrupt operations. The protect surface is orders of magnitude smaller than the overall attack surface and always knowable. Firewalls and other controls are moved as close as possible to the protect surface rather than the perimeter at the attack surface where it is decidedly further away from what needs to be protected. In this way, it is possible to determine what traffic moves in and out by a much smaller number of users or resources that need access to sensitive data or assets.

As with the attack surface, organizations must constantly monitor their protect surface to identify and block potential threats as quickly as possible. Theoretically, the smaller footprint makes this process more manageable. However, actual methods to model the protect surface are still being developed. Implementing distinct Zero Trust capabilities and activities is anticipated by 2027, and concerns continue on how this will be accomplished.

ACCESS AND CYBERATTACK INITIATION

Four attributes of sensors and sensor systems must be considered in determining how access may be achieved to defend against cyberattacks—sensor, system, and network hardware; software; human interfaces; and communications systems. In the past, sensor networks by their very nature were likely to have no more than a few active human users and real-time interactions with other information technology and operational technology systems. This is rapidly changing, as data sharing between multiple fleet and shore assets increases at exponential rates. Sensor operations should occur through one or more firewalls or equivalent technologies that can detect the legitimacy of access requests. They should also be well segmented and isolated by virtue of access being physically or electronically limited to a few fully vetted people for maintenance and system upgrades. Vulnerabilities encountered through outbound communications from a sensor network are also likely to be greatly minimized, as the destinations for sensor data and analytics products should be well defined and restricted.

Attacks are most likely to occur by insiders introducing malware through software and firmware updates. In the past, this has been accomplished using USB and other devices by maintainers through satellite and other communications systems, as well as "trusted" users whose end-point systems were compromised. Despite today's better operational procedures, viruses and other malicious code can still be passed along to computer server(s) and/or sensor controller(s) that interact with the server network and directly with the sensors. Security breaches can also occur by monitoring data communicated between sensors and their controllers and between the controllers to the server(s) using middleware that functions as a hidden layer between an operating system and sensor software applications. Middleware increases the possibility to insert malicious software that targets known vulnerabilities or models the sensor environment to devise methods on its own to detect and attack vulnerabilities within the network itself and/or the individual sensors operating within the network.

Sensor operations should occur through one or more firewalls or equivalent technologies that can detect the legitimacy of access requests.

Malware can be activated immediately after insertion or remain dormant for long periods until activated upon receipt of a stimulus or code. This stimulus may be predicated from a combination of unique operating conditions exhibited by an engine or other ship's machinery, a specific date and time, or a vessel's position defined

by latitude and longitude. Another method of initiating a cyberattack is by recognizing known unique visual, digital, acoustic, radio frequency, or directed energy signature(s) introduced external to the ship, which is then communicated through a ship's networks to sensor systems or the sensors themselves. This includes the medium being sensed, such as the air or water, to an onboard camera, radar, sonar, or other external sensors.

Essential to maintaining cybersecurity is the need to ensure conformance and interoperability of methods and processes to provide resilient and high-performance network capabilities that include service quality, prioritization, and avoiding service preemption from accidental and malicious sources. Cybersecurity technologies and products integrated onboard naval vessels and vehicles must be strictly validated to ensure communications and interoperability at the end point applications and the servers that manage these applications to overcome multiple vulnerabilities and opportunities to infiltrate, disrupt, and otherwise compromise sensor operations.

DISTINGUISHING BETWEEN THE NATURAL AND NEFARIOUS

Analysis of sensor data is accomplished using combinations of characteristics represented within frequency, time, and spatial domain sensor signal representations. Differences are detected through contrast with degrees of variation from nominal sensor operation as embodied within training data sets tailored to the data characteristics of specific sensor types and models. The training requirements presume a very large volume of data across the time, frequency, and spatial (image) domains over long periods, preferably with continuous learning so that the training data will continue to improve throughout the useful life of the sensor(s). This data must contain representative samples of nominal operating conditions for all sensors involved and be tailored to individual sensor functionality. Examples may include day, night, and twilight; calm to rough seas; fair to stormy weather; empty screens to ones filled with targets; and shallow to deep waters. This is an area where deep-learning AI takes a lead role to ensure that robust and comprehensive training is achieved with minimal false indications of degradation.

Sensor degradation caused by human agencies with nefarious intent include physical and cyberattacks that are also widespread and problematic. Many such events can often be diagnosed by a watchstander or technician who has gained a high level of expertise using onboard systems. However, watchstanders with learned knowledge of sensor theory but little practical experience in sensor operation will not necessarily be able to reliably discern degradation. More importantly, unmanned vessels and autonomous vehicles have even less ability to do so. Less obvious examples of sensor degradation include gradual loss of pixels in a camera over time that steadily reduce resolution and accuracy, the effects of precipitation and fog on light propagation and infrared night vision performance, and inadequately documented performance anomalies in new types of sensors recently introduced onto the bridge. Similar effects can be identified in engineering applications such as optical devices that become contaminated and grow cloudy, proximity sensors that shift from optimal alignment, and resistive transducers that age prematurely and exceed specifications due to environmental and other factors.

Sensor degradation caused by human agencies with nefarious intent include physical and cyberattacks that are also widespread and problematic.

Physical and cyberattack characteristics can display distinctly different and unique sensor data and signal characteristics than naturally

occurring phenomena. In such cases, identification is possible by including attack data in training data sets. However, instances such as increasing degradation of individual camera pixels caused by malware leading to resolution loss can also be determined through fusing and analyzing complementary sensors' data as well as individual sensor instruction sets, much of which can be accomplished in real time.

The capability to identify specific causes of sensor degradation requires a significant investment in further training using the specific characteristics of degraded sensor signals with one or more of the three available domain signal representations. For example, visual camera images representing spatial domain imagery illustrate the occurrence of specific causes of sufficiently different degradation to enable positive identification based upon their unique characteristics. Specific degradation classes and types can then be created for training using ML techniques.

Two different classes of degradation can be defined as being partial or full based upon the extent of the affected sensor surface area. Partial degradation affects only part of the sensor image or reduced/changed spectral frequency range and content of sensor signals, while full degradation generally affects 50% or more of the sensor image. Types of degradation for imaging sensors include damage, obstruction, obscurant, smoke, fog, and precipitation. Identifying the causes of degradation based only upon the unique characteristics presented from spatial domain perspectives may be sufficient to accurately and consistently identify degradation types and sources.

Other cases may require additional insight to determine proper classification and type. In such instances, unique characteristics associated with each degradation event may also exist within the frequency domain. The properties of each causal agent may filter specific frequencies of light entering, for example, into a camera and possibly introduce new frequencies not naturally present. These interdomain correlations can help to provide greater specificity in identification. Examining time domain characteristics can help identify the exact circumstances of degradation initiation (whether it occurred quickly or gradually), the extent of degradation, and whether it continues unchanged or if the sensor is recovering from the effects of the degradation agent through evaporation, dislocation, or another scenario.

The capability to identify specific causes of sensor degradation requires a significant investment in further training using the specific characteristics of degraded sensor signals with one or more of the three available domain signal representations.

THE FACES OF SENSOR DEGRADATION

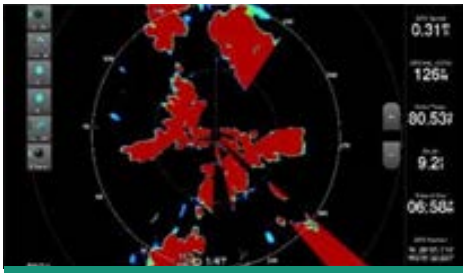
Many of the effects of sensor degradation cannot be readily discerned by human operators, as training may not be available for visual clues that appear on their displays (spatial domain) and the characteristics of degradation may only appear and be detected in the frequency and/or time domain representations. However, many such examples exist and are shown in Figure 2. Detecting these sensor degradation examples and many others not shown was achieved with high reliability and reproducibility [20].

Camera images shown in Figures 2a and b depict two of many types of natural and manmade degradation that include fog, heavy precipitation, physical damage, chemical obscurants, and physical obstructions. These are just a few of the many images obtained under different weather conditions at different times of day used in testing image degradation detection capabilities. Imagery obtained from drone aircraft operating off the

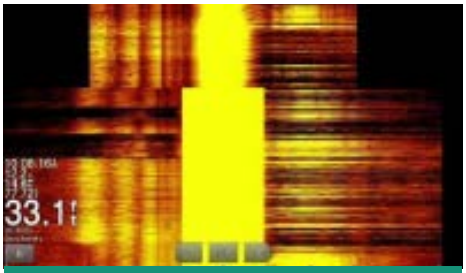
a. Degradation Due to Fog

b. Chemical Obscurant

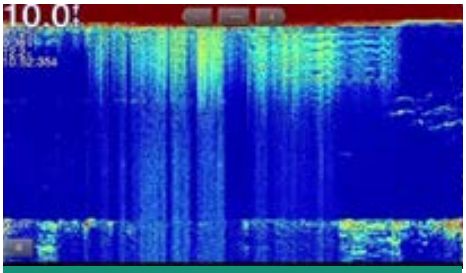
c. Precipitation Degradation



d. Electromagnetic Interference



e. Nonlaminar Flow



f. Environmental Interference

g. Physical and Laser Interference

h. Weather Sensor Disruption

i. WX Satellite Interruption

Figure 2. Examples of Degradation of Various Shipboard Sensors *(Source: Wright [20])*.

research vessel also provide live video, audio, and other sensor telemetry data feeds amenable to using sensor degradation detection technology.

Figure 2c illustrates degraded imagery obtained from an infrared camera. Sensor degradation was accomplished because of heavy rain showers. The degree of degradation encountered was generally proportional to shower intensity. Degradation was also achieved using chemical obscurants, physical obstructions, and conditions like smoke and fog. Additional degradation was accomplished using many of the same techniques performed for visual camera degradation, with similar results.

Figure 2d illustrates degraded radar imagery representative of different types of degradation that takes advantage of various combinations of events, timing, and signal characteristics to facilitate detection. The degraded image depicts the result of electromagnetic interference targeting the heading sensor the radar system depends on for proper geospatial orientation. The primary effect of this experiment was to cause the radar display to jitter and gyrate wildly, resulting in small-to-extreme distortion of the displayed imagery. Similar results were obtained from both digital and analog radars from different manufacturers. Many degraded states not shown were also detected.

Figures 2e and f represent two of the many different types of degradation that may be experienced by the various types of sonar sensors likely to be

found on conventional and unmanned vessels. These sensors included side-scan, traditional echosounders, forward-looking navigation, and chirp sonars produced by various commercial manufacturers.

An example of lidar sensor degradation is shown in Figure 2g, which represents an image spanning 360°, with the vessel centered at the middle. Nominal operation is illustrated, where there is a natural blind spot on the navigation light array in the lower left field of view. The imagery in the red circle illustrates the area that has been degraded because of an obstruction placed in the laser propagation path that causes a shadow effect where detail is lost. Additional degradation was attempted by inflicting strikes on the sensor with several low-power (3–30 mW) lasers utilizing different wavelengths (450–905 nM). Results included shadows casted from the direction of origin, interference lines, and occasional display blinking.

Figure 2h shows an example of onboard weather sensor degradation performed by introducing external stimulus and situations that would interfere with normal operation. Illustrated are the effects of splashed and streams of water that resulted in maximum readings in wind speed and relative humidity, abnormal temperature readings, and other measurements uncharacteristic and unrepresentative of the physical environment. Anomalies were readily detected and included water and air temperature, humidity, barometric pressure, wind speed, and direction during a clear, calm day, with steady temperatures and without wind.

Figure 2i illustrates reception degradation of satellite transmissions for weather and other onboard services achieved by antenna shielding and disrupting the reception path between the satellite and the vessel. The loss of signal resulted in the degradation, which was easily detected. Note that in the absence of a signal, the satellite weather display became blank. However, the AIS overlay continued without interruption.

FUTURE OPPORTUNITIES AND TECHNOLOGIES

There are many opportunities to create multiple solutions that enable rapid development and dynamic update of accurate and comprehensive reduced-order models of sensor and sensor system architectures. This is especially important when dealing with complex vessel network infrastructure to detect critical system elements that would make them susceptible to model-based attacks. Novel AI-based and future quantum methods will help autonomously and rapidly build and update these models to represent classes of sensors and sensor network system functionalities that span multiple, different ship and unmanned vehicle infrastructure configurations and defend against a multitude of diverse threat vectors. Such methods must minimize human-guided domain knowledge and training requirements and feature dynamic development, adaptation, and reconfiguration to overcome vulnerabilities exposed by destabilization cyberattacks. Order-of-magnitude reduction in development time over existing methods is needed, along with enhanced resilience to broad-based attacks and reduced brittleness to previously unknown cybertactics and a broad range of natural phenomena such tactics can mimic.

One focus of the 2023 DoD Cyber Strategy is developing and applying new technologies to expand cybercapabilities and prioritizing technologies to confound malicious cyberactors and prevent them from achieving their objectives in and through cyberspace [21]. These DoD capabilities implement the priorities of the National Cybersecurity Strategy to defend critical infrastructure and disrupt and dismantle threat actors [22]. This includes Zero Trust

Novel AI-based and future quantum methods will help autonomously and rapidly build and update these models to represent classes of sensors and sensor network system functionalities.

architectures and their associated cybersecurity technologies, advanced endpoint monitoring capabilities, tailored data collection strategies, enhanced cyberforensics, automated data analytics, and systems that enable network automation, network restoration, and network deception. Assisting in this implementation is the recently launched Cyber Operational Readiness Assessment (CORA) Program of the Joint Force Headquarters–Department of Defense Information Network (JFHQ-DODIN) to harden information systems, reduce the attack surface of their cyberterrain, and enhance a more proactive defense [23]. CORA implementation for ships and unmanned vehicles is under the U.S. Fleet Cyber Command, which serves as the Navy's component command to the U.S. Cyber Command.

Quantum computing is also envisioned to equip the Navy with more secure communications networks, more advanced sensors, and faster threat detection and response that comes with them to improve navigation and result in smarter autonomous systems and more accurate modeling and simulation [24]. The Naval Research Laboratory and Naval Information Warfare Center (NIWC) Pacific have established the Naval Quantum Computing Program Office, where quantum subject matter experts across all 14 naval warfare centers can collaborate on quantum applications for the DoD [25]. ■

Quantum computing is also envisioned to equip the Navy with more secure communications networks, more advanced sensors, and faster threat detection and response.

REFERENCES

[1] Eckstein, M. "U.S. Navy Updating Tactics for Sensors, Weapons Based on Houthi Attacks." *Defense News*, https://www.defensenews.com/newsletters/2024/02/14/us-navy-updating-tactics-for-sensors-weapons-based-on-houthi-attacks/, 13 February 2024.

[2] Wright, R. G. "In-Stride Detection of Sensor Degradation." GMATEK, Inc., final report, Revision: A, Contract: N6833520G2005, Naval Sea System Command, Washington, DC, 28 April 2022.

[3] Honeywell. "What Can Cause Sensor Readings to Drift?" Honeywell, Charlotte, NC, article no. 000005197, https://sps-support.honeywell.com/s/article/What-can-cause-sensor-readings-to-drift, 25 October 2023.

[4] Ahmad, M. "3 Basic Facts About Automotive Sensor Degradation." *Electronic Design News*, https://www.edn.com/three-basic-facts-about-automotive-sensor-degradation/, 1 October 2020.

[5] Kushner, D. "The Real Story of Stuxnet." *IEEE Spectrum*, https://spectrum.ieee.org/the-real-story-of-stuxnet, 26 February 2013.

[6] Walmor, C. L., Jr., C. Coreixas de Moraes, C. E. P. de Albuquerque, R. C. Santos Machado, and A. Oliveira de Sá. "A Triggering Mechanism for Cyber-Attacks in Naval Sensors and Systems." *Sensors*, vol. 21, no. 8, p. 3195, 4 May 2021.

[7] Jiang, J., et al. "Bio-Inspired Covert Active Sonar Strategy." *Sensors*, vol. 18, no. 8, p. 2436, https://doi.org/10.3390/s18082436, 26 July 2018.

[8] MFame Editor. "Cyber Risk of ECDIS Operational Technology Onboard Ships." https://mfame.guru/cyber-risk-of-ecdis-operational-technology-onboard-ships/, 25 October 2021.

[9] Iphar, C., C. Ray, A. Napoli, P.-Y. Martin, and A. Bouju. "Multi-Domain Assessments in AIS Falsification Cases." Maritime Big Data Workshop, CMRE, https://www.cmre.nato.int/maritime-big-data-workshop-home/maritime-big-data-workshop-presentations/1202-multi-domain-assessments-in-ais-falsification-cases/file, May 2018.

[10] Bhatti, J., and T. E. Humphreys. "Hostile Control of Ships via False GPS Signals: Demonstration and Detection." University of Texas, preprint article in *NAVIGATION*, vol. 64, no. 1, https://radionavlab.ae.utexas.edu/images/stories/files/papers/yacht.pdf, 7 May 2017.

[11] Solnør, P., Ø. Volden, K. Gryte, S. Petrovic, and T. I. Fossen. "Hijacking of Unmanned Surface Vehicles: A Demonstration of Attacks and Countermeasures in the Field." *Journal of Field Robotics*, Wiley, vol. 39, pp. 631–649, 10 February 2022.

[12] Vinnem, J. E., and I. B. Utne. "Risk From Cyberattacks on Autonomous Ships." *Safety and Reliability-Safe Societies in a Changing World*, DOI:10.1201/9781351174664-188, June 2018.

[13] International Maritime Organization (IMO). Safety of Life at Sea Convention (SOLAS) Convention, chapter V, regulation 19, paragraph 6, 2020.

[14] International Maritime Organization. "Integrated Bridge Systems." Maritime Safety, https://www.imo.org/en/OurWork/Safety/Pages/IntegratedBridgeSystems.aspx, accessed on 25 April 2024.

[15] Fortinet. "Network Security Vulnerabilities." Sunnyvale, CA, https://www.fortinet.com/resources/cyberglossary/network-security-vulnerability, accessed on 23 April 2024.

[16] Pyle, F. "Surface Warfare SITREP." American Society of Naval Engineers, Combat Systems Symposium, https://www.navalengineers.org/Symposia/TSSCSS23/CSSProceedings, 2 February, 2023.

[17] Eckstein, M. "How the U.S. Navy Is Creating the 'Nirvana of One Combat System'." *Defense News*, https://www.defensenews.com/naval/2023/02/08/how-the-us-navy-is-creating-the-nirvana-of-one-combat-system/, 8 February 2023.

[18] Fortinet. "What Is an Attack Surface?" https://www.fortinet.com/resources/cyberglossary/attack-surface, accessed on 30 April 2024.

[19] Office of Prepublication and Security Review. "DOD Zero Trust Strategy." https://dodcio.defense.gov/Portals/0/Documents/Library/DoD-ZTStrategy.pdf, 21 October 2022.

[20] Wright, R. G. "Ship Sensors: Conventional, Unmanned and Autonomous." Chapter 10, *Vessel Sensor Degradation*, Taylor and Francis, Oxford, UK, ISBN: 978-1-032-45621-8, February 2024.

[21] U.S. DoD. "2023 DoD Cyber Strategy." https://media.defense.gov/2023/Sep/12/2003299076/-1/-1/1/2023_DOD_Cyber_Strategy_Summary.PDF, 12 September 2023.

[22] The White House Washington. "National Cybersecurity Strategy." https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Cybersecurity-Strategy-2023.pdf, 1 March 2023.

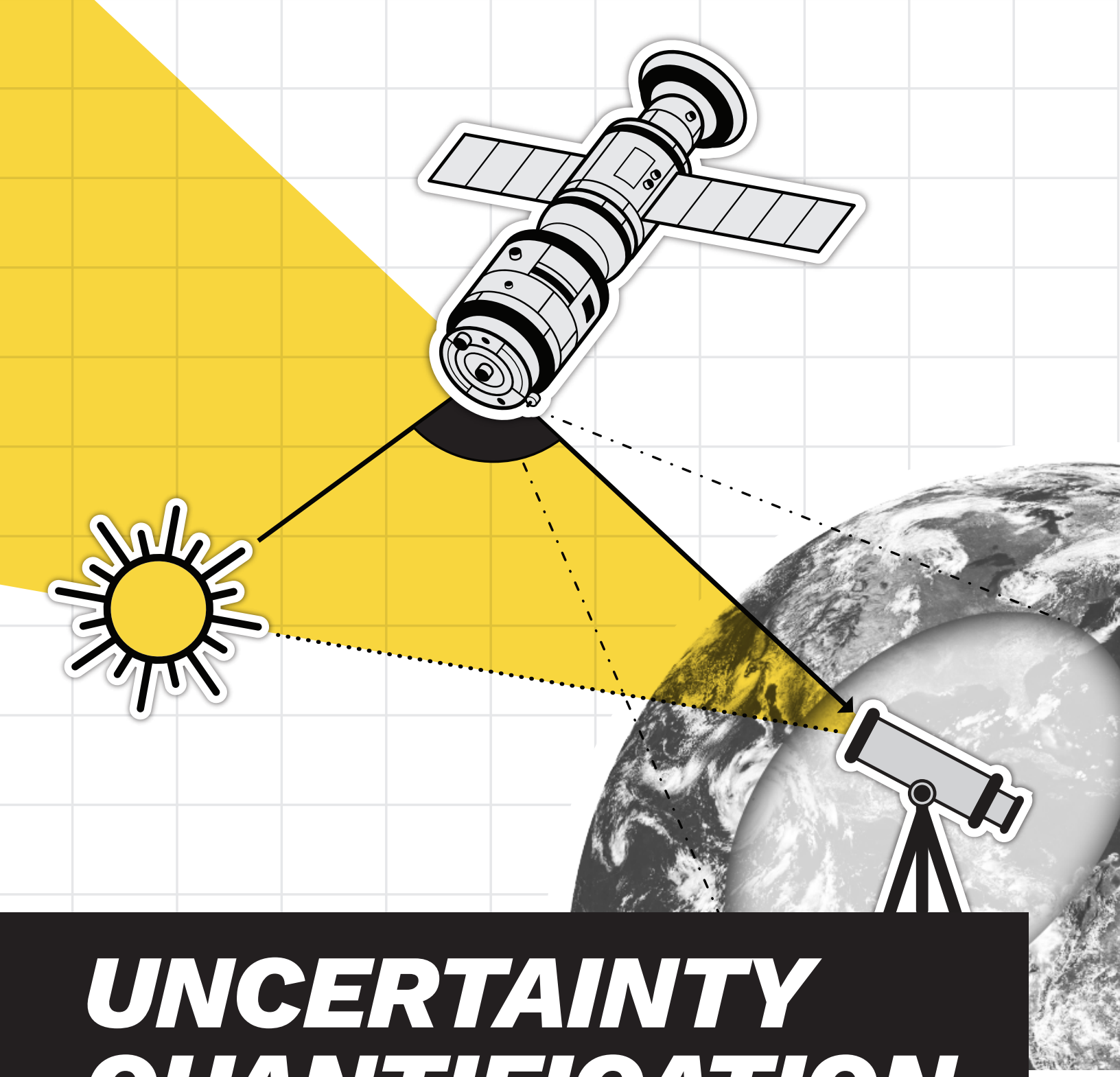
[23] Mavica, S. "JFHQ-DODIN to Officially Launch Its New Cyber Operational Readiness Assessment Program." *DOD News*, https://www.defense.gov/News/News-Stories/Article/Article/3691583/jfhq-dodin-officially-launches-its-new-cyber-operational-readiness-assessment-p/, 1 March 2024.

[24] Piedfort, M. "NIWC Pacific and Its Partners Are Building a Quantum Navy." *Defense Visual Information Distribution Service*, https://www.dvidshub.net/news/441514/niwc-pacific-and-its-partners-building-quantum-navy, 29 March 2023.

[25] Rombado, L. "The Quantum Future of Naval Warfare." *Proceedings of the U.S. Naval Institute*, vol. 148, no. 2, p. 1428, https://www.usni.org/magazines/proceedings/2022/february/quantum-future-naval-warfare, February 2022.

BIOGRAPHY

R. GLENN WRIGHT has over 45 years of experience in aerospace, maritime, and medicine. He has led projects associated with sensor-based systems that include surface vessel, autonomous underwater, and marine remotely operated vehicles; meteorological and oceanographic data systems; and surface navigation. He is a member of the National Academies of Science, Engineering, and Medicine Transportation Research Board committees on maritime safety and AI; an expert contributor to the International Hydrographic Organization Crowd Sourced Bathymetry Working Group; a master mariner; and an author of books on autonomous ships and their sensors and more than 100 journal articles and conference papers. Dr. Wright holds a B.S. in electrical engineering from the New Jersey Institute of Technology, an M.S. in computer science from the Polytechnic Institute of New York University, and a Ph.D. in maritime affairs from the World Maritime University in Malmö, Sweden.



UNCERTAINTY QUANTIFICATION

TO DETECT RESIDENT SPACE OBJECT ANOMALIES

BY IMÈNE R. GOUMIRI, AMANDA L. MUYSKENS, BENJAMIN W. PRIEST, ROBERT E. ARMSTRONG, AND J. LUC PETERSON (PHOTO SOURCE: MAICONFZ, 8FPIXS, ZABI JOSE AND GRAPHIXMANIA [CANVA])

SUMMARY

Anomaly detection, whereby signals are scanned for behavior changes, is a common problem in many applications. As data streams grow in complexity and volume, automated methods for anomaly detection, especially those based on machine learning (ML), are increasingly attractive. However, many real-world defense applications like space domain awareness (SDA) are challenging environments for these algorithms because sensors can be noisy, datasets can be large, and observations can have unpredictable gaps. Furthermore, many automatic detections are limited in their actionable utility because they lack a meaningful measure of uncertainty and confidence. Without uncertainty analysis, it can be difficult for ML models and their recommendations to be trustworthy.

This article will describe a scalable ML algorithm, based on Gaussian processes (GPs), to address many of these concerns. An example from SDA is used that detects and identifies resident space object (RSO) trajectories and behaviors. It shows how the algorithm MuyGPs can accurately identify RSO anomalies from real-world data and provide uncertainty in those detections. While immediately useful for SDA applications, the MuyGPs framework is flexible, fast, and general, making it an attractive option for a diverse set of real-world missions whenever abnormal behavior needs to be identified and acted upon.

INTRODUCTION

Anomaly detection, the process of identifying unusual patterns or events, is an essential tool for monitoring the vast and dynamic space environment. By effectively detecting anomalies in RSO trajectories or behavior, valuable insights can be gained into maneuvers, malfunctions, or potential threats. However, applications to real-world data face significant hurdles. Measurement noise and gaps in the data from faulty sensors to external phenomena like the weather can mask true anomalies or trigger false alarms. Furthermore, the amount of data expected to grow dramatically in the near future due to the increase in satellite launches and the proliferation of low-cost cameras that can easily collect large samples of data can overwhelm traditional processing techniques, making it difficult to identify anomalies in a timely manner. This motivates a need for automatic methods for detecting anomalies in large, noisy, and potentially sparse data sets.

Light curves of RSOs are time series of observed brightness that provide a wealth of information. The brightness of an RSO depends on several components, including size, shape, reflective material, distance from observer, and solar phase angle. A light curve is also sensitive to dynamic quantities like spin-rate, tumbling motion, or maneuvers. A sudden change in the brightness of a satellite could indicate that it has been damaged or has had its performance degraded. Several studies have successfully demonstrated the use of light curves to infer RSO properties [1–6]. They are also good examples of large, sparse, and noisy datasets, which is why they are used as a realistic benchmark for demonstrating the methods described here.

> ***Anomaly detection, the process of identifying unusual patterns or events, is an essential tool for monitoring the vast and dynamic space environment.***

GPs provide a nice framework for working with such sparse and noisy datasets, as they typically give accurate predictions while requiring less training data than more popular ML techniques and naturally incorporating uncertainty into their predictions. But standard GP models are typically limited to small training datasets (less than ~10,000 observations). To overcome these limitations, a scalable GP model called MuyGPs was employed to efficiently predict large datasets faster and more accurately than competing GP methodology [7]. It

is ideal for the future of this SDA problem, where efficient predictions are needed in very short time spans and with minimal computing power.

An approximate GP-based method for SDA anomaly detection and object classification is presented. This expands on the previous results of Goumiri et al. [8, 9], where it was shown that GPs can effectively be used to interpolate and forecast RSO light curves. This article will expand upon these models to demonstrate a method for detecting anomalies in various scenarios and a workflow that can be used to classify unknown RSOs based on their light curves.

The anomaly detection workflow consists of predicting known data using a GP model comparing predictions with actual measurements considering the predicted uncertainty. This model includes a heteroscedastic noise model, wherein each observation has a previous, separate noise derived from on-sensor measurements taken at the time of observation as well as an anisotropic kernel assumption. The details will be described further in the Methodology section.

The RSO classification workflow has two major modeling stages. First, since the training data is collected at sparse and irregularly spaced locations, a model like that in Goumiri et al. [9] is utilized to interpolate light curve observations to a consistent set of observing times over each night. The second major step in this workflow is to utilize these interpolated observations as training data to classify the light curves.

This article outlines the ML workflow useful for future SDA anomaly detection and classification problems based on light curve observations. The Methodology section describes the data and GP methods, including descriptions of the heteroscedastic and anisotropic model features and the reference frame interpolation used for preparing data for classification. The Results section details how these methods work in practice on forecasting and anomaly detection tasks as well as a classification task that requires interpolated data. Finally, the Conclusions section explains the implications of the results and future directions for this work.

METHODOLOGY

Light Curve Data

A ground-based, commercial off-the-shelf camera pointed at geostationary orbits can accumulate a large amount of RSO data for relatively little cost. However, there are limitations working from the ground, such as changing lighting conditions, variable sensor availability, and poor or missing observations resulting from bad weather. Light curves from these objects will routinely include regions of noisy data and large gaps that must be addressed if the data are to be used for training ML models.

In this analysis, a dataset provided by Dave Monet is used that consists of 43 known satellites measured from a single camera in Flagstaff, AZ, between September 2014 and September 2018 [10]. This is the same dataset used in previous papers by Goumiri et al. [8, 9]. Each measurement consists of a measured time, position on the sky, and brightness and error in a filter close to the Johnson V-band. There are ~500,000 data points per object, resulting in a dataset of ~6.5 million data points. This analysis was limited to 13 satellites labeled as "Nominal" in the dataset (see the Interpolation With GPs subsection), interpolated to denoise and complete missing data (Figure 1). These were selected because they have data over the full period and appear to be active satellites.

GPs

GPs are powerful tools for modeling and predicting noisy data that can be used to learn nonlinear relationships between features and responses. These responses can be a continuous variable, such as the brightness of the light curve at a particular time, or discrete labels, such as the identification of light curve observations with individual RSOs or classes of object or orbital regimes. GPs for regression and classification tasks are used in this work.

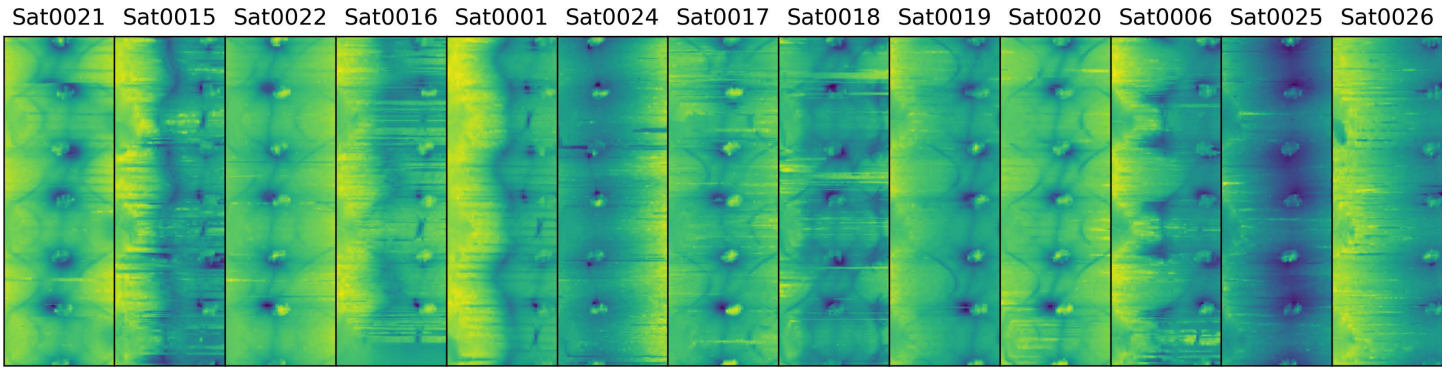


Figure 1. Nominal Light Curves: Magnitude Is Plotted as a Function of Days (Rows) and Times of Day (Columns) for a Four-Year Duration *(Source: Goumiri et al. [9])*.

> ***GPs are powerful tools for modeling and predicting noisy data that can be used to learn nonlinear relationships between features and responses.***

A GP parameterizes a collection of random variables such that any subset of these variables is a jointly multivariate Gaussian with a known mean function $m_\vartheta(\cdot)$ and covariance function $k_\vartheta(\cdot, \cdot)$, where ϑ is a set of hyperparameters that control the behavior of the GP. $m_\vartheta(\cdot)$ is usually taken to be 0, without a loss of generality. As such, the joint probability distribution of any set of GP outputs can be completely specified by its mean vector and covariance matrix, and the conditional (posterior) distribution of predictions can be estimated by applying Bayes' rule.

GPs naturally incorporate uncertainty into their predictions, and model decisions including the form of $k_\vartheta(\cdot, \cdot)$ have a natural statistical interpretation. These features are advantageous compared to other popular ML techniques such as deep neural networks (DNNs), although uncertainty quantification and interpretation of DNNs is an active area of research. The covariance matrix of a GP specifies the degree of correlation between different outputs. Being able to quantify the uncertainty of predictions without training specifically for it is important for many applications.

Goumiri et al. initially explored the GP modeling of RSO light curves [8]. That work used isotropic and homoscedastic assumptions in its model, where a covariance was taken to be a function of Euclidean distance in the input space and the noise of each measurement was assumed to be i.i.d. Gaussian. However, the resulting posterior variances were misspecified, as the uncertainties around the relatively stable measurements taken in the middle of the night were the same width as the much noisier measurements taken near dusk or dawn. This analysis attempts to correct these model discrepancies to more accurately measure the uncertainties associated with RSO light curve prediction.

The model in Linares and Furfaro [5] assumed measurement noise for observations was drawn i.i.d. from $\mathcal{N}(0, \varepsilon)$. That is, given time dimensions $X = (\mathbf{x}_1, \ldots \mathbf{x}_n)^T$, the response $Y(X)$ is modeled as

$$Y(X) = (Y(\mathbf{x}_1), \ldots, Y(\mathbf{x}_n))^T \sim \mathcal{N}(\mathbf{0}, \sigma^2 (K_\vartheta(X, X) + \epsilon I_n)), \quad (1)$$

where N is the multivariate Gaussian distribution, $\mathbf{0}$ is the n-dimensional zero vector, σ^2 is a variance scaling term, I_n is the $n \times n$ identity matrix, and $K_\vartheta(X, X)$ is an $n \times n$ positive definite, symmetric covariance matrix between the elements of X controlled nonlinearly through kernel function $k_\vartheta(\cdot, \cdot)$ with hyperparameters ϑ.

As noted in Goumiri et al. [8], using a homoscedastic noise model implies that the prior noise is constant over all magnitudes. However, in practice, the data at high magnitudes tend to have more variance than the data collected at lower magnitudes. The model considered in this analysis uses a vector of noise prior variances $\boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon}_i \sim \mathcal{N}(0, \boldsymbol{\varepsilon}_i)$ for each observation i. These variance priors are taken from measurements collected from the sensor at the same time as the corresponding magnitude. This changes the prior in equation 1 to

$$Y(X) = (Y(\mathbf{x}_1), \ldots, Y(\mathbf{x}_n))^T \sim \mathcal{N}(\vec{0}, \sigma^2 (K_\vartheta(X, X) + diag(\boldsymbol{\varepsilon}))), \quad (2)$$

where diag($\boldsymbol{\varepsilon}$) is the diagonal matrix with $\boldsymbol{\varepsilon}$ along the diagonal.

The prior model is further generalized by utilizing an anisotropic distance model. It used the isotropy assumption that $k_\vartheta(\mathbf{x}, \mathbf{z}) = p_\vartheta(\|\mathbf{x}-\mathbf{z}\|^2 / \ell^2)$ for some function $p_\vartheta(\cdot)$ and a single length scale parameter ℓ. However, embedding time into multiple dimensions imposed an arbitrary relationship between displacement in time of day and displacement in the day-of-observation interval sensitive to the units of the observations and how the data were normalized. While the initial exploration disregarded this detail, this relationship is modeled here by considering functions of the form $k_\vartheta(\mathbf{x}, \mathbf{z}) = p_\vartheta(\Sigma_{j=1\ldots d} (\mathbf{x}_j-\mathbf{z}_j)^2 / \ell^2))$ for the same function $p_\vartheta(\cdot)$ and separate length scale parameters ℓ_j.

Explicitly modeling length scales along each feature dimension creates less sensitivity to data normalization choices and makes the covariance scales along each dimension clear.

Interpolation With GPs

Because light curves are often noisy and intermittent and many applications require complete and smooth data, interpolation is key to being able to denoise and further process light curve data. Goumiri et al. [8] demonstrated that GPs were very performant at interpolating light curves provided that the data were properly embedded in a low-dimensional space reflecting their periodicity. Using two-dimensional (2-D) embedding where each data point was replaced by a 2-D vector was described. The first dimension represented the time of day as a real number in the [0, 1] interval. Another dimension represented the day of the observation period (four years) as an integer, starting at zero on the first day of observation, and further normalized to the interval [0, 1]. As was customary, the constant mean was also subtracted from the responses and added back to the predictions.

MuyGPs with a Matérn kernel and a fixed smoothness hyperparameter of $\nu = 0.2$ was used to implement the new models. This smoothness hyperparameter was selected to ensure that predicted curves had the desired degree of smoothness and were fixed to limit the cost of optimizing other

Because light curves are often noisy and intermittent and many applications require complete and smooth data, interpolation is key to being able to denoise and further process light curve data.

hyperparameters. Isotropic models have one additional hyperparameter, the length scale ℓ, that is set to $\ell = 1$. Anisotropic models have two length scale hyperparameters, ℓ_1 and ℓ_2, that are set to $\ell_1 = 0.1$ and $\ell_2 = 1$. The median of the measured magnitude error was used as the homoscedastic noise ϵ and a polynomial fit of order 2 of that same measured magnitude error to smooth the heteroscedastic noise $\boldsymbol{\varepsilon}$.

Classification With GPs

To demonstrate an example of a task that requires properly interpolated data rather than raw time series of magnitudes, GPs are used as tools for identifying unseen light curves based on a database of known ones. This classification task maps vectors of magnitudes to light curves labels and hence requires that the magnitudes for all different light curves be interpolated to a common set of time points. Interpolating the light curves using the methods described in the

Interpolation With GPs subsection on a regularly-spaced grid of time points (5 min apart) spanning the longest common interval of time to all light curves in the dataset (about four years) was chosen. That common interval guarantees strictly interpolation and not extrapolation since extrapolation is susceptible to larger errors and the use of a regularly-spaced grid for times makes plotting easier (see Figure 1). To incorporate the uncertainty provided by the GPs during the interpolation phase into the classification, one-day chunks from the posterior distributions are sampled, and each sampled chunk of interpolated magnitudes forms a feature vector. Each feature vector is then associated with a one-hot encoded label with zero mean, meaning light curve i has a vector label that is all $-1/N$ except for component i, which is $(N-1)/N$, where $N = 13$ is the number of light curves in the dataset.

To make the task realistic, the first three years are reserved for training the classifier, an isotropic GP with homoscedastic noise similar to the one described in the Interpolation With GPs subsection. The fourth year is used as the test data.

The training and test data are interpolated independently. Predictions are then compared to actual measurements to determine the accuracy. This is a realistic SDA scenario since it forecasts the future and can be readily applied given historical data.

RESULTS

The sensors used to capture light curves were calibrated for the night sky. Unsurprisingly, the images produced at dawn or dusk, when the sky became brighter, were noisier than images produced in the middle of the night. Figure 2 shows an example of the magnitude of a light curve for one day of missing data (Day 445 of Sat0019). At the start of night, the captured magnitude is high and noisy. Then, as the night progresses, the magnitude drops and so does the noise. Finally, as dawn approaches, the magnitude gets higher and noisier again.

However, a homoscedastic noise model cannot capture that natural variation of the noise level with the magnitude. The left pane of Figure 2 shows how a homoscedastic model underestimates the noise at dawn and dusk and overestimates it in the middle of the night. The pink-shaded region represents the 95% confidence interval, where 95% of the data is

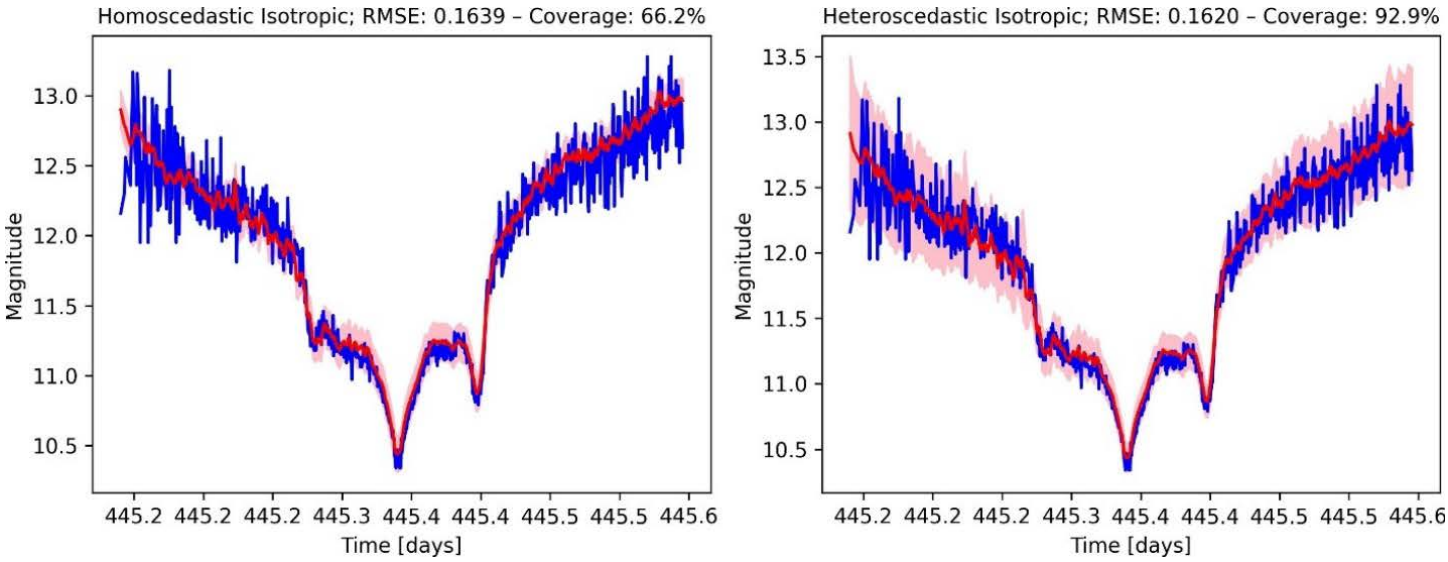


Figure 2. A Homoscedastic Model (Left) vs. a Heteroscedastic Noise Model (Right) Comparing GP Prediction (Red) With Measurement (Blue) *(Source: Goumiri et al. [9])*.

expected to lie. The coverage is about 66%, meaning that 66% of the points are within the shaded region overall. Ideally, the coverage should be close to 95%. Near the edges, many points lie outside the region, visibly more than 5%. At the center, 100% of the points are within the region.

On the other hand, a heteroscedastic noise model can consider the heterogeneity of the variance. While the relation between the noise and the magnitude could be learned, e.g., using a GP, this dataset contained measurement errors obtained from the sensors for each data point, so they were used to fit a quadratic function from the noise model. The right pane of Figure 2 shows how a heteroscedastic model produces a magnitude-dependent variance that properly reflects the uncertainty at any time of the night. The coverage is close to 95% and, as importantly, homogeneously correct throughout the night.

Figure 3 is like Figure 2 but shows the effect of using an anisotropic deformation model. The left pane shows an example of the magnitude of a light curve during a single night as predicted by an anisotropic model with homoscedastic noise. The right pane shows the same prediction using an anisotropic model with heteroscedastic noise. The pink-shaded region is the 95% confidence interval of the prediction. The root means square error is lower (better) using the anisotropic models, although the difference is marginal. This is because the scaling of the input's two dimensions was carefully optimized to get a reasonably isotropic distribution of nearest neighbors. The effect might be more pronounced in problems with greater scale differences across dimensions of the feature space.

One important selling point for using GPs in space applications is their ability to predict full posterior distributions, providing insight into the confidence of the model for each prediction. Having an objective measure of confidence can be useful in computing probabilities and help in decision-making. It is also possible to use the posterior distribution to detect anomalies. For instance, GPs can be used to predict known measurements and compare the posterior mean to the measurements given the posterior standard deviation. Points where the difference consistently exceeds the standard deviation can be flagged as anomalies and further analyzed. Figure 4 shows an example of anomalies detected by the GP for a certain period for light curve Sat0024 that can sometimes uncover maneuvers (top). Sat0024 exhibits such anomalies in the fourth year (yellow-shaded regions). Analyzing the corresponding orbital elements obtained from Spacetrack (NORAD ID 29155; bottom) reveals discontinuities consistent with a maneuver (pink-shaded region) that could explain the second anomaly.

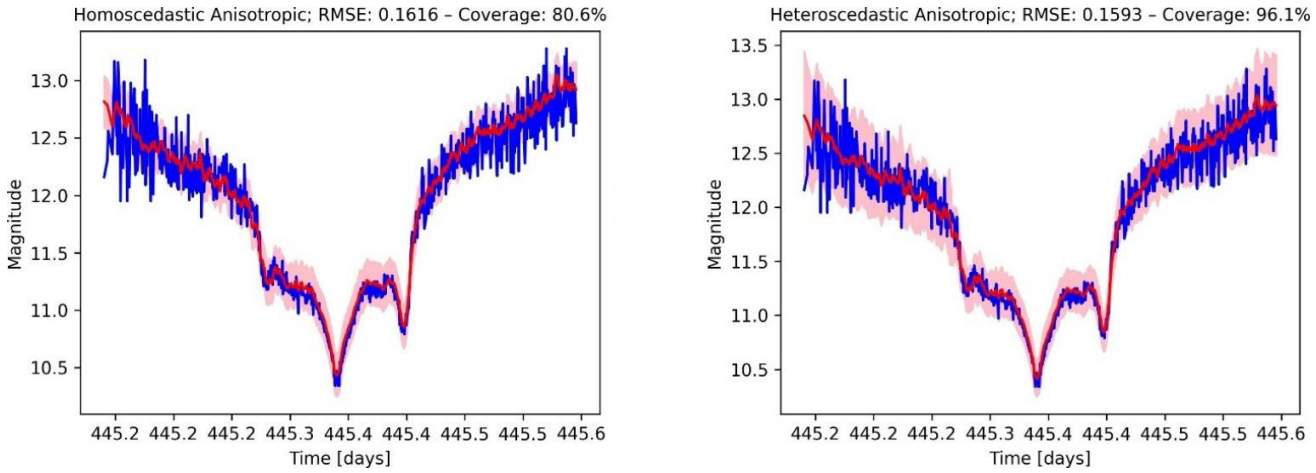


Figure 3. Anisotropic Models Comparing GP Prediction (Red) to Measurement (Blue) for One Day of Missing Data (Day 445 of Sat0019) *(Source: Goumiri et al. [9]).*

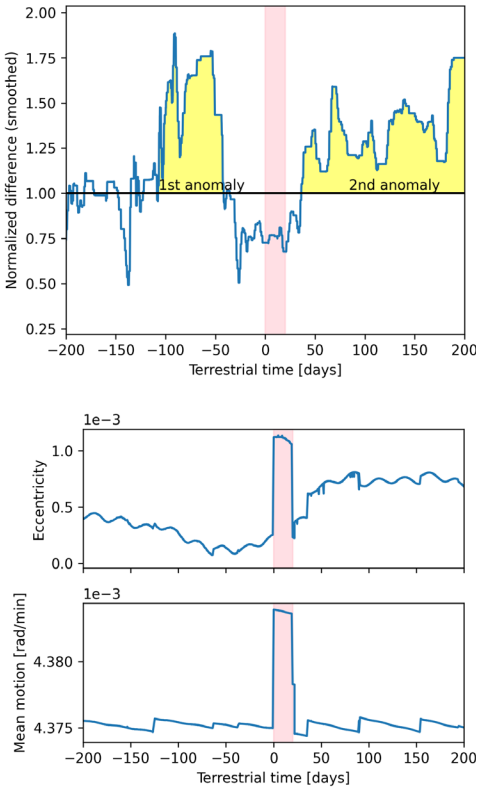


Figure 4. Comparing GP Predictions (Posterior Mean) to Measurements and Normalizing the Difference by the Posterior Standard Deviation *(Source: Goumiri et al. [9])*.

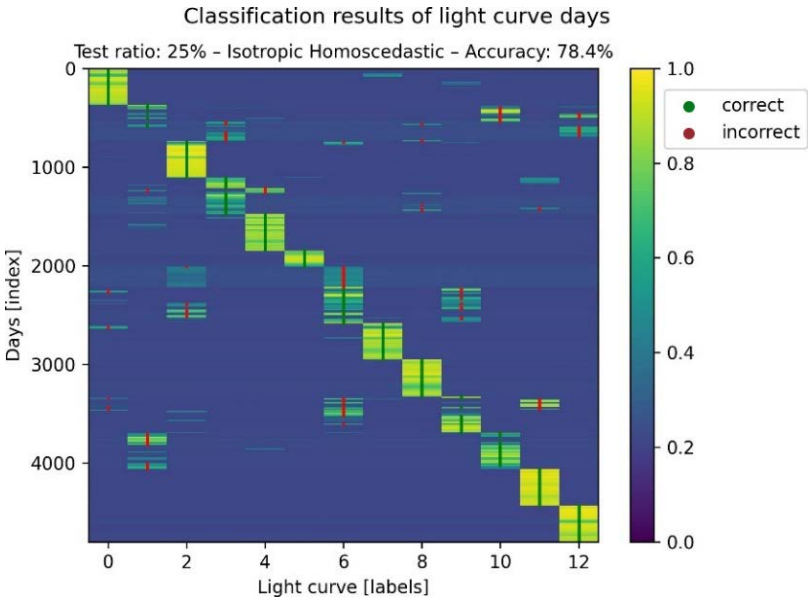


Figure 5. Classification Results *(Source: Goumiri et al. [9])*.

Lastly, to demonstrate a realistic task that requires preprocessing raw light curves, a classifier (also powered by MuyGPs) was built and trained on the first three years of the dataset. This was set up to identify the light curve it belongs to among the 13 trained on from a single day of unseen data from the fourth year. Every single day of the last year of data is classified this way, and the results are reported in Figure 5.

Using the anisotropic model with heteroscedastic noise to interpolate the light curves yields a slightly better classification accuracy. The overall accuracy is 78%, meaning that the provenance of 78% of the 13 × 365 test days is correctly identified. For certain light curves, the accuracy is close to 100%, while for others, it is lower. Each row is a test day, ordered by time and light curve, and each column represents one light curve. The colors represent the one-hot encoded prediction results—vectors of probabilities of belonging to each of the 13 light curves. On top of that, the green and red dots represent correct and incorrect predictions, respectively, and are positioned on the predicted point's light curve. There are several potential reasons for the misclassifications, but they are mainly due to issues in the interpolation step since a different classifier, a random forest, on the same data was tested and resulted in similar results. What happens in the interpolation phase is that some of the satellites either maneuver or rotate. This causes cyclic variations every couple of days that interfere with the nearest neighbor selection process and lead to jumps in the predicted data. This, in turn, leads to misclassification in the classification phase.

CONCLUSIONS

A novel method for efficiently processing and classifying light curve data while inherently quantifying uncertainty was presented. This capability is crucial for SDA, as the volume of data continues to rise due to advancements in sensor technology and the growing number of satellites and debris. While automation offers significant advantages, its effectiveness ultimately relies on the confidence assigned to its results. GPs address this challenge by naturally incorporating uncertainty

quantification through their mathematical framework, bypassing the potential pitfalls of learning-based approaches. This inherent trust in GP predictions is further enhanced by the approach that leverages the full posterior distributions generated by the model to create representative training data for the classifier, effectively incorporating uncertainty into the classification process.

Furthermore, the predicted uncertainty serves as a powerful tool for anomaly detection. By identifying data points that deviate significantly from the model's expected range of magnitudes established through prior observations, these anomalies are flagged as potential indicators of malfunctions or maneuvers. Correlating them with additional signals can further strengthen the case for investigation, enabling either manual or automated follow-up analysis. This approach extends beyond the realm of light curve data; its applicability transcends time-series datasets and can be effectively applied to a broader range of scientific inquiries. By incorporating uncertainty quantification into the analysis pipeline, defense analysts can automate classification and gain valuable operational insights from potential outliers that might hold significant intelligence value.

While light curve data can be inherently noisy and incomplete, this analysis also highlighted the challenges posed by significant variations in the data itself. ML methods rely heavily on training data that accurately reflect the test data. In this case, some light curves exhibited year-over-year or even day-over-day variations due to maneuvers or rotations, which could lead to misclassifications. However, these variations were often captured by the uncertainty inherent in the posterior distributions of the interpolated light curves. Interestingly, comparing GPs and random forest classifiers for the classification task revealed similar performance and error rates. This suggests that the core challenge lies in effectively interpolating the data to a common grid and not in the classification itself.

The MuyGPs method showcases its potential beyond light curve classification. It is particularly well-suited for applications that involve large datasets, often exceeding the capabilities of traditional GPs. But since it is inherently a GP, it also proves valuable when the available data volume might not be sufficient for training other ML techniques such as DNNs. While light curves served as the demonstration here, MuyGPs broadly applies and works best under two key conditions. First, a reasonable overlap between training and test data distributions is crucial for optimal performance. Second, the data dimensionality should ideally be manageable, either inherently (less than 500 dimensions) or through dimensionality reduction techniques like principal component analysis (PCA). As prior work demonstrated, PCA can even enhance accuracy and performance [11].

Looking ahead, one potential area of improvement for MuyGPs lies in addressing the challenges posed by variations in light curve data. Currently, nearest neighbor calculations for interpolation rely on isotropic normalization of day and time of day, which can introduce inconsistencies due to the arbitrary scaling between dimensions. Investigating methods to incorporate anisotropic scaling during neighbor selection or even iteratively updating the nearest neighbor index regarding learned deformation parameters could lead to improved accuracy in handling these variations in future work. ■

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory (LLNL) under Contract DE-AC52-07NA27344. Funding for this work was provided by LLNL laboratory-directed research and development grant 22-ERD-028.

NOTE

This document was prepared as an account of work sponsored by an agency of the U.S. government (USG). Neither the USG nor Lawrence Livermore National Security, LLC, nor any of their employees make any warranty, expressed or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by tradename, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the USG or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the USG or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

REFERENCES

[1] Dianetti, A. D., and J. L. Crassidis. "Space Object Material Determination From Polarized Light Curves." AIAA Scitech 2019 Forum, p. 0377, 2019.

[2] Furfaro, R., R. Linares, and V. Reddy. "Space Objects Classification via Light-Curve Measurements: Deep Convolutional Neural Networks and Model-Based Transfer Learning." AMOS Technologies Conference, Maui Economic Development Board, 2018.

[3] Furfaro, R., R. Linares, and V. Reddy. "Shape Identification of Space Objects via Light Curve Inversion Using Deep Learning Models." AMOS Technologies Conference, Maui Economic Development Board, Kihei, Maui, HI, 2019.

[4] Jia, B., K. D. Pham, E. Blasch, Z. Wang, D. Shen, and G. Chen. "Space Object Classification Using Deep Neural Networks." The 2018 IEEE Aerospace Conference, pp. 1–8, 2018.

[5] Linares, R., and R. Furfaro. "Space Object Classification Using Deep Convolutional Neural Networks." The 19th International Conference on Information Fusion (FUSION), pp. 1140–1146, 2016.

[6] Linares, R., M. K. Jah, J. L. Crassidis, and C. K. Nebelecky. "Space Object Shape Characterization and Tracking Using Light Curve and Angles Data." *Journal of Guidance, Control, and Dynamics*, vol. 37, no. 1, pp. 13–25, 2014.

[7] Muyskens, A., B. Priest, I. Goumiri, and M. Schneider. "MuyGPs: Scalable Gaussian Process Hyperparameter Estimation Using Local Cross-Validation." arXiv preprint arXiv:2104.14581, 2021.

[8] Goumiri, I. R., A. M. Dunton, A. L. Muyskens, B. W. Priest, and R. E. Armstrong. "Light Curve Completion and Forecasting Using Fast and Scalable Gaussian Processes (MuyGPs)." Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS), 2022.

[9] Goumiri, I. R., A. L. Muyskens, B. W. Priest, and R. E. Armstrong. "Light Curve Forecasting and Anomaly Detection Using Scalable, Anisotropic, and Heteroscedastic Gaussian Process Models." Technical report, Lawrence Livermore National Laboratory, Livermore, CA, 2023.

[10] Monet, D. Personal communication. Dataset (2014–2018) from Flagstaff, AZ, 2022.

[11] Goumiri, I. R., A. L. Muyskens, M. D. Schneider, B. W. Priest, and R. E. Armstrong. "Star-Galaxy Separation via Gaussian Processes with Model Reduction." arXiv preprint arXiv:2010.06094, 2020.

BIOGRAPHIES

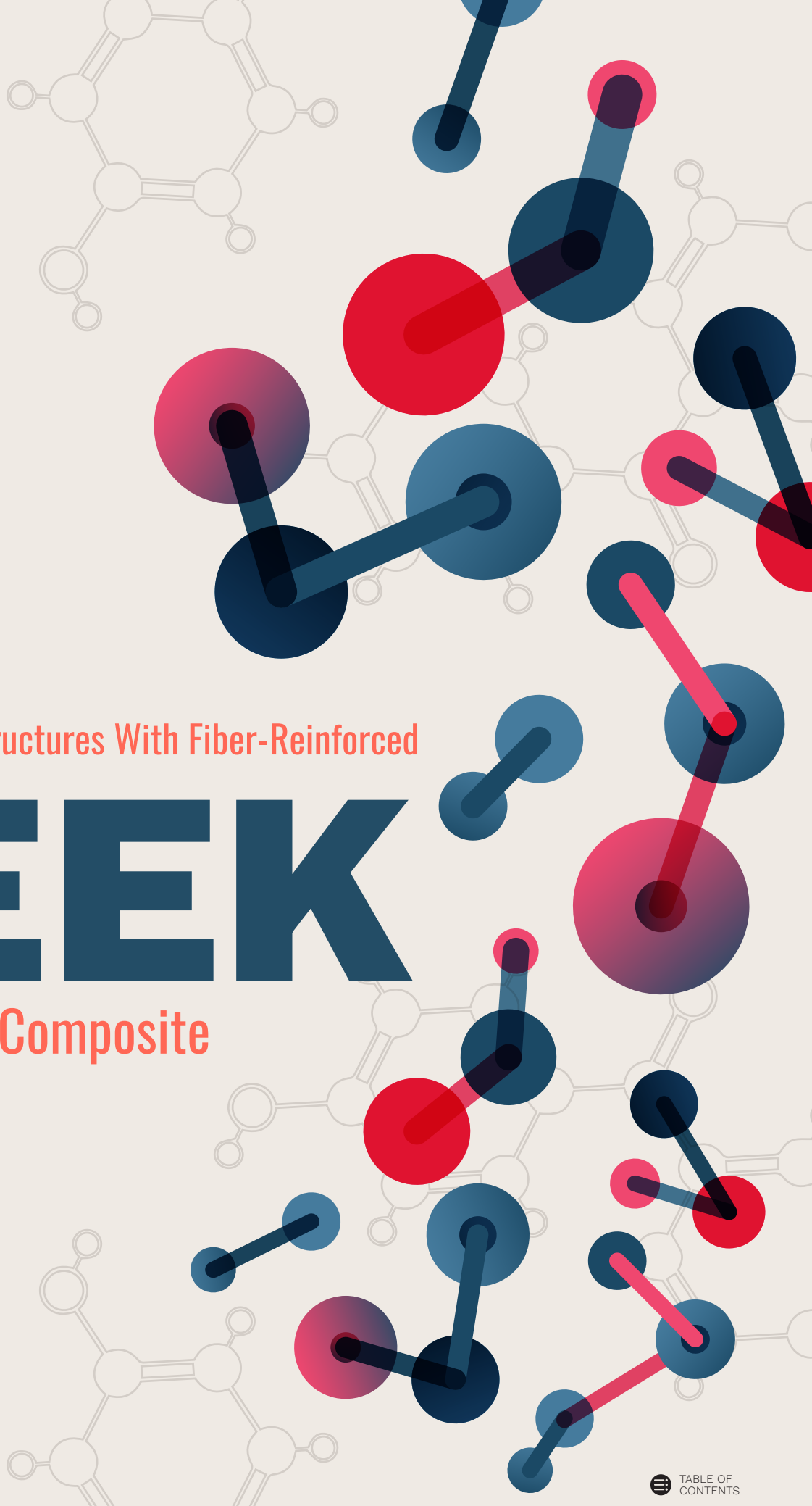
IMÈNE GOUMIRI is a staff scientist in the Computational Engineering Group at LLNL working on applications of Gaussian processes at scales ranging from atomic structures to astronomy. Her expertise includes control theory, plasma physics, and numerical analysis. Dr. Goumiri holds an M.A. and Ph.D. in mechanical and aerospace engineering from Princeton University, an M.Eng. in mathematical and mechanical modeling from the Polytechnic Institute of Bordeaux, and an M.S. in mathematics, statistics, and economics (with a specialty in modeling, computation, and environment) from the Université de Bordeaux 1.

AMANDA MUYSKENS is a staff member in the Applied Statistic Group at LLNL, where she leads the MuyGPs project that has developed novel methods for scalable, nonstationary Gaussian processes for high-performance computing and the Data Science Summer Institute. Her expertise includes surrogate modeling, Gaussian process models, computationally efficient ML, uncertainty quantification, and statistical consulting. Dr. Muyskens holds bachelor's degrees in mathematics and music performance from the University of Cincinnati and an M.S. and Ph.D. in statistics from North Carolina State University.

BENJAMIN (MIN) WESLEY PRIEST is a postdoctoral researcher at the Center for Applied Scientific Computing, LLNL, focusing on scalable graph analytics, high-performance computing, and scalable statistics and ML. Min has developed novel algorithms for Gaussian process estimation, distributed subspace embeddings and sketching for massive graphs, and published in conferences and journals. Dr. Priest holds a Ph.D. in engineering from Dartmouth College.

ROBERT ARMSTRONG is a staff scientist at LLNL working on image processing algorithms for large astronomical surveys and focusing on the upcoming Rubin Legacy Survey of Space and Time. His research interests involve using survey data to understand open questions in physics, including dark energy, dark matter, and aspects of the early universe; Bayesian statistics; ML; and high-performance computing. He has performed research for the Dark Energy Survey and the Hyper Suprime-Camera Collaboration.

JAYSON "LUC" PETERSON is the associate program leader for data science within the Space Science and Security Program at LLNL, where he oversees several data science and space projects. He has worked on modeling and simulation, experimental design, digital engineering, uncertainty quantification, verification and validation, data analytics, high-performance computing, and ML in various applications from nuclear fusion to COVID-19 response. Dr. Peterson holds a B.A. in physics and science, technology, and society from Vassar College and an M.S. and a Ph.D. in astrophysical sciences (plasma physics) from Princeton University.



Designing Primary Structures With Fiber-Reinforced

PEEK

Thermoplastic Composite

BY HARRY R. LUZETSKY

(PHOTO SOURCE: ANTISHOCK [123RF.COM] AND SERGEISHIMANOVICH [SHUTTERSTOCK.COM])

INTRODUCTION

Polyetheretherketone (PEEK) is a semicrystalline thermoplastic that was invented in November 1978 and brought to market in 1981 by Imperial Chemical Industries, which later became Victrex. As developed, this material possessed a maximum 48% degree of crystallinity. The composite material was introduced a year later as APC-1 with a 52% fiber volume, and it was optimized to yield APC-2 with a 63% fiber volume. The fiber adhesion properties of APC-2 resulted in superior impact and crack resistance compared to APC-1, as well as existing epoxy-based composites. The birth of the fiber-reinforced PEEK composite coincided with the push for composites on various aircraft applications to enhance performance while reducing weight. Properties associated with this material were very alluring to structures designers due to enhanced mechanical properties over epoxy fiber-reinforced composites and many metals. Improvements to these properties include the following:

> ***The birth of the fiber-reinforced PEEK composite coincided with the push for composites on various aircraft applications to enhance performance while reducing weight.***

- Extreme toughness/damage tolerance/survivability
- 150% ballistic damage tolerance improvement over aluminum
- >100% better impact resistance over toughened epoxy
- Enhanced energy absorption
- Improved damping
- Superior environmental/chemical/solvent resistance
- Excellent fatigue resistance
- Low water absorption
- No refrigeration/out-time/exotherm considerations
- Recyclable
- Melt processable (no cure chemistry, long soak times)
- No toxicity/hazardous chemical issues
- Excellent wear resistance
- Low friction coefficient
- Stable glass transition temperature—even under hot/wet conditions
- Hydrolytic stability (polymer dependent)

Despite the identified advantages associated with fiber-reinforced PEEK composites, designers have been slow to accept the material for several issues, including the following:

1. Lack of robust material properties and/or design/allowables to support designing structures and support quality control activities in a production environment.
2. Material fiber variability impact on mechanical properties and design.
3. Lack of processing options that support translation of developed designs into structural realities.
4. Ill-perceived elevated material and process costs (related to elevated processing temperatures and pressures) and limited material usage (globally) that directly impacts material costs and availability.

Since their inception, research and development activities have quickly resolved these concerns; however, the focus of designers has largely remained on epoxy-based, fiber-reinforced composites. This is largely due to their familiarity with the systems and the large investment in equipment, such as autoclaves and large freezers. Using proper design philosophy, one system does not take precedence over the other, but the one that is best suited for the application should be selected. To design primary structures with fiber-reinforced PEEK thermoplastic composite, a basic understanding of the issues that have plagued full acceptance of the material is required.

MATERIAL PROPERTIES

Since the development of the material, sufficient data have been collected to support the development of B-basis allowables based on the directions provided in the Composites Materials Handbook-17 (CMH-17). B-basis design allowables (90% probability with 95% confidence) can and have been recreated from data used in

previous Army experimental programs. Once allowables are established, they are continuously updated as new data become available. The development of these allowables is possible by pooling data of different graphite fiber systems per CMH-17 guidelines. A large database of material information is required for establishing design allowables where the key is repetition of results. Typically, each resin and fiber combination requires a new set of design allowables. For small data populations, the result of any basis value calculation strongly depends on the sample size. Smaller sample populations are less costly to test, but this approach is limited because as the population size decreases, so does the calculated B-basis value. Not only does the estimated B-basis value increase with larger sample sizes, but as the one sigma limits illustrate, the expected variation in estimated value significantly decreases.

With data pooling, the IM7 fiber is superior to the AS4 and IM6 fibers in which it is grouped. Two possible effects from adding AS4 or IM6 data to IM7 data are described in Table 1.

Sample design allowables for the IM7/PEEK thermoplastic are illustrated in Table 2. These allowables have since been updated. However, due to their conservatism, they are still viable for developmental activities considering additional data may be necessary to address identified material or application data gaps.

Table 1. Effects on B-basis From Grouping Data *(Source: R. Luzetsky)*

CONDITION OF RAW DATA	EFFECT ON B-BASIS
AS4 or IM6 data very similar to IM7 data	A raising of the B-basis value due to an increased confidence in IM7 average.
AS4 or IM6 data dissimilar to IM7 data (higher or lower)	A lowering of the IM7 B-basis value due to batch variation and, hence, low confidence in IM7 average. The B-basis will be lowered, even if the AS4 or IM6 values raise the overall average.

Table 2. Sample Design Allowables for IM7/PEEK *(Source: SURVICE Engineering)*

PROPERTY	AVERAGE	B-BASIS	MINIMUM	MAXIMUM
E11 (MSI)	23.8	19.4	21.3	26.5
E22 (MSI)	1.30	1.13	1.18	1.32
G12 (MSI)	0.843	0.527	0.726	1.026
X_T (KSI)	383.3	306.0	332.5	419.2
X_C (KSI)	151.2	100.1	133.3	188.7
Y_T (KSI)	11.1	10.1	10.4	11.6
Y_C (KSI)	27.8	23.0	25.3	30.7
S_S (KSI)	24.8	21.8	22.9	26.4
$E11_T$ (ε in/in)	8,414	7,320	7,396	9,072
$E11_c$ (ε in/in)	4,681	4,130	4,400	4,963

Since design allowable development can be a costly proposition, teaming with a company that has vetted design allowables provides a way to move forward with a prototype design so a particular application can be evaluated without the expenditure necessary to collect and develop a separate set of design allowables.

Once material design properties are established, composite lamination theory and finite-element modeling can be used to develop a design for a given application. The focus is to develop properties that meet and/or exceed the requirements for the component being developed. This is accomplished in conjunction with the capability provided by the fabrication process. Fortunately, for fiber-reinforced PEEK thermoplastic materials, a myriad of processes has been developed to provide configuration versatility for the designer.

MATERIAL FIBER VARIABILITY

Numerous fibers are compatible with the PEEK resin system, creating great design versatility. In addition to traditional graphite and glass fibers (E and S2), nickel-coated carbon fibers have been demonstrated. Recently, a multifunctional form of IM7/PEEK

composite has been developed that exhibits structural properties and high levels of electromagnetic (EM) shielding. This material form involves integrating expanded copper mesh within the laminate construct and has exhibited the potential to create EM shielding levels exceeding 100 dB. Any time a new fiber type is introduced to the PEEK matrix, a series of tests, including shear, is implemented to determine any impact to the design allowables. With both the nickel-coated carbon and expanded copper mesh, deleterious effects were found to be minimal. When properly integrated, they can yield EM properties for the entire composite laminate greater than that of the base composite material (IM7/PEEK). The ability to select various fiber forms supports the ability to tailor material properties to a specific structural application.

FABRICATION PROCESSES AND TECHNIQUES

Proper design with IM7/PEEK needs to consider the fabrication process. For IM7/PEEK, the available processes include autoclave, press or compression forming, double-diaphragm forming, in-situ tape placement, continuous compression molding (CCM), three-dimensional (3-D) printing with fiber-reinforced PEEK thermoplastic, and overmolding.

Autoclave

Autoclave processing requires using special high-temperature materials (vacuum bagging and sealing tape). A laminate stack is created by using a soldering iron to tack individual plies to one another. The laminate stack is aligned over a tool, and once the appropriate forming temperature is achieved, vacuum and pressure are applied to cause the laminate stack to form over the tool. Limitations exist in the potential shapes, as complex curvatures would be subject to an uncontrolled fiber wash, which would directly impact structural response characteristics.

Press or Compression Forming

Like autoclave processing, a laminate stack is created using a soldering iron to tack the individual plies to one another to form the stack. The laminate stack is placed over a tool in a high-temperature, pressure-platen press. The laminate stack is heated to the appropriate temperature, and the press is used to force the laminate stack either into or over the tool, as illustrated in Figure 1. There are some curvature and size limitations due to fiber wash during the application of pressure. With this process and the additive nature of the material, it

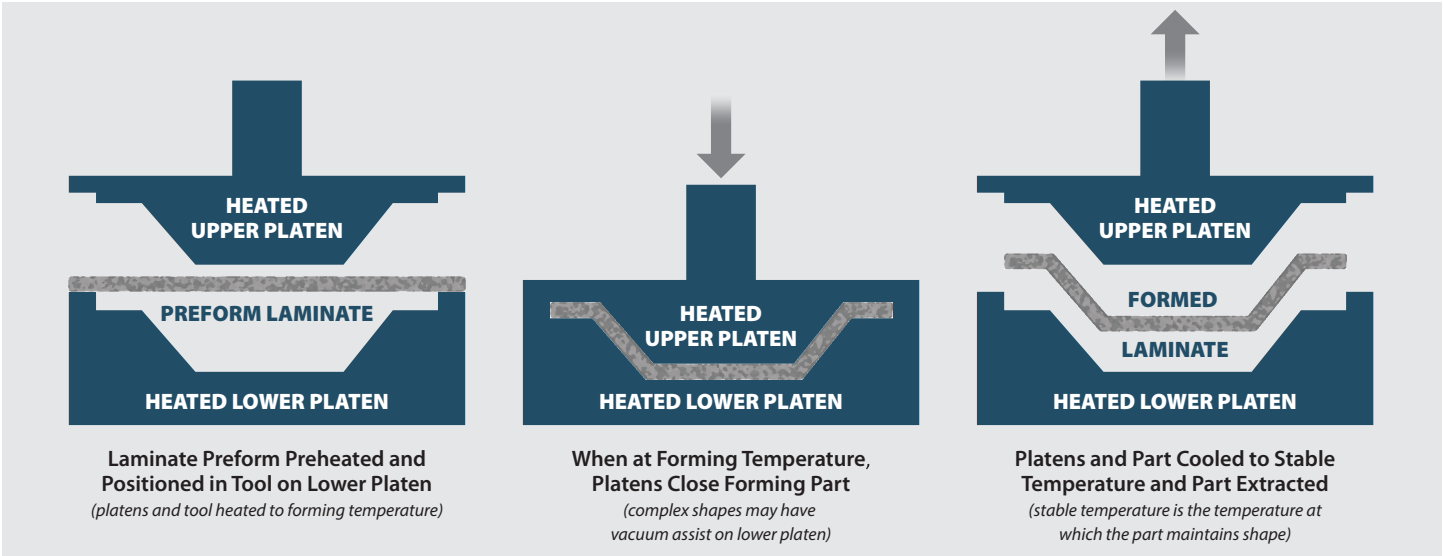


Figure 1. Press or Compression Molding Process *(Source: H. R. Luzetsky)*.

is possible to form subelements and combine them into a larger unified structure with the press operations, as illustrated in Figure 2.

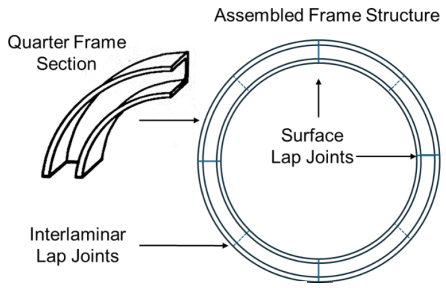


Figure 2. Segmented Frame Press Consolidated From Sections *(Source: H. R. Luzetsky)*.

Double-Diaphragm Forming

Double-diaphragm forming is used for forming structures with integrated reinforcements, such as integrated reinforcing beads (Figure 3).

In this forming process, the laminate stack is placed between two sheets of superplastic aluminum suspended over a tool in a pressure vessel, as illustrated in Figure 4. The sandwiched structure is heated to the appropriate forming temperature. Once the temperature is reached, the tool is raised to contact the laminate structure and pressure is applied to form the material blank over the tool, followed by cooling. The aluminum forming sheets are required, serving as rigid caul plates for the operation while conforming to the underlying tool. These sheets are removed, revealing the formed composite structure. Due to the expense of the expendable materials, this is a process reserved for closed 3-D-reinforcing structures like that in Figure 3; it can be accomplished in a separate tool that does not require an autoclave for the vacuum and pressurization.

> ***Double-diaphragm forming is used for forming structures with integrated reinforcements, such as integrated reinforcing beads.***

In-situ Tape Placement

With in-situ tape placement, an additive manufacturing process, localized application of heat and pressure with a compaction roller (as illustrated in Figure 5) mitigates processing with an autoclave or oven. It is considered an "on-the-fly" process in which the part is fully consolidated as the raw material is being put into place. The raw material is first heated using a heat source (hot gas torch stream or laser), and then it is

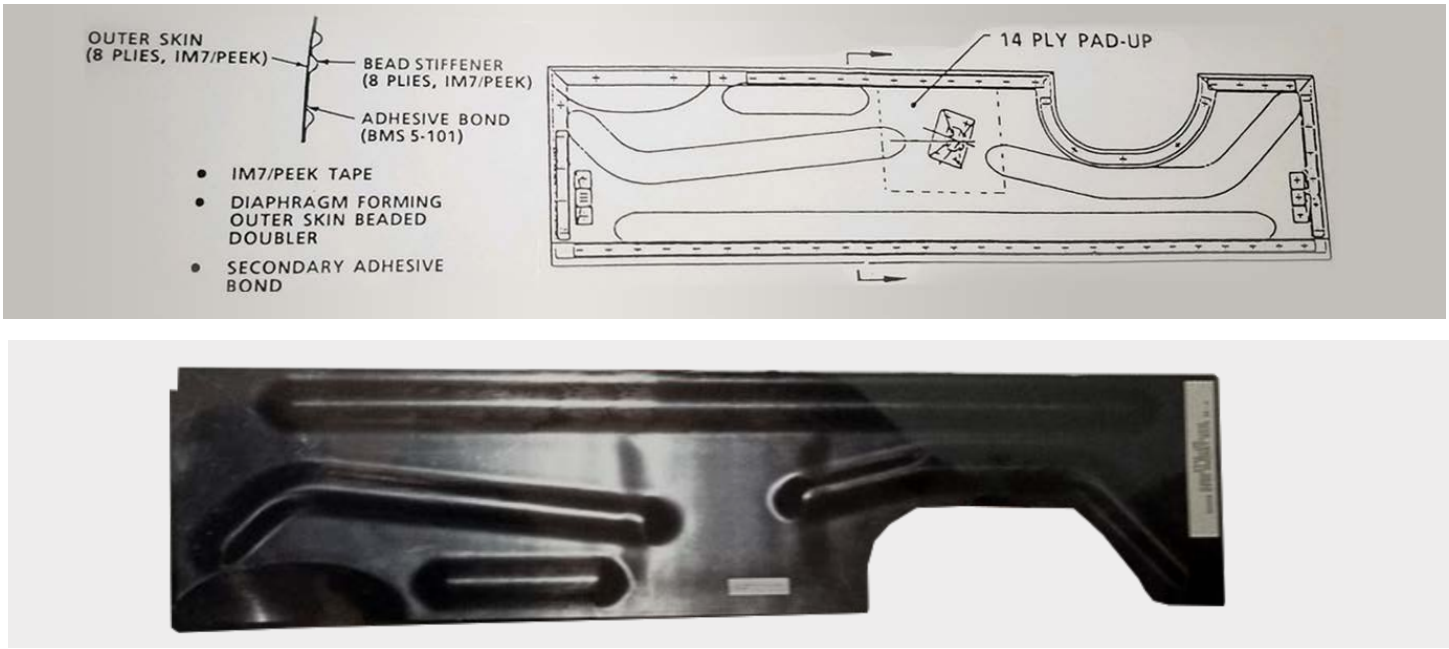
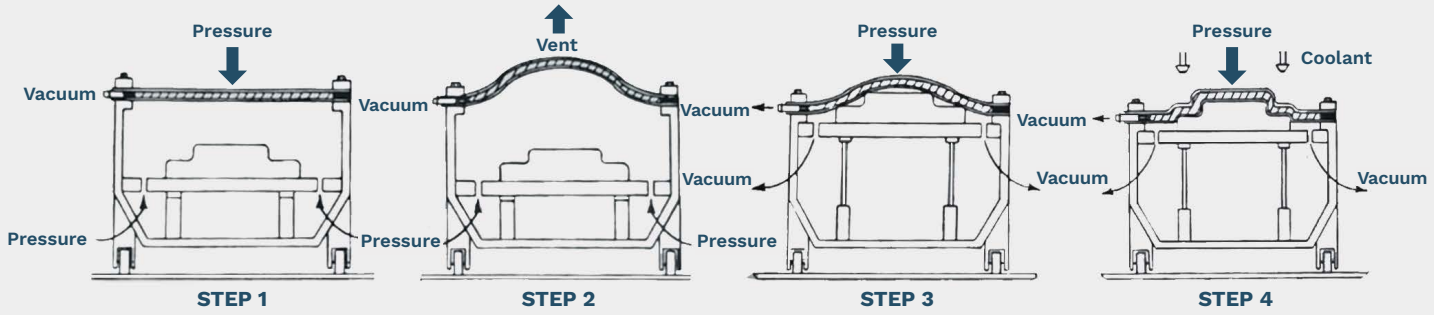


Figure 3. Prototype Landing Gear Door Formed From Double-Diaphragm Forming Process *(Source: H. R. Luzetsky)*.



STEP 1: Install tack-welded laminate pack between diaphragm materials and draw vacuum between diaphragm sheets combined with the application of pressure and heat on both sides of the laminate to consolidate laminate.

STEP 2: Once forming temperature is reached, vent the pressure from above the laminate stack and increase the pressure within the tool chamber, causing a billowing of the preform laminate stack (preform with diaphragm skins).

STEP 3: Run the heated tool up into the laminate stack while applying pressure external to the stack, simultaneously drawing vacuum from the tool chamber and drawing the laminate stack down over the tool.

STEP 4: Continue to pull vacuum in combination with external pressure while applying coolant to reduce temperature. Once laminate has cooled lower tool, remove the laminate stack and separate the diaphragm material from the part.

Figure 4. Diaphragm-Forming Process Using Superplastic Aluminum as Forming Sheets *(Source: H. R. Luzetsky)*.

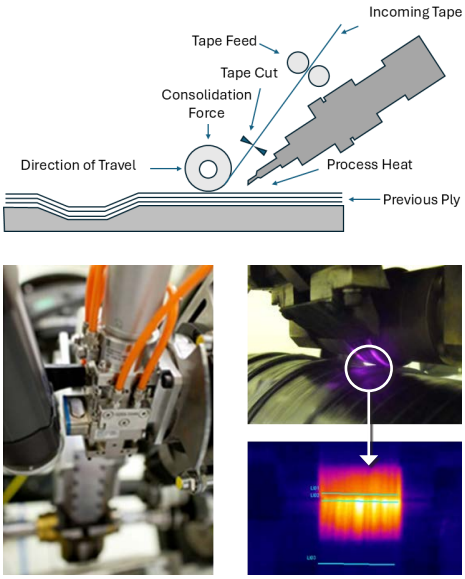


Figure 5. In-situ Automated Fiber Placement (AFP) Process With Laser Heat Source *(Source: TSS/Albany)*.

consolidated/compacted with a rigid steel roller. There are no intermediate debulking steps or postprocessing needed. This process precludes the need for an external heat or pressure source to consolidate the composite laminate and is not limited by size restrictions. Since consolidation occurs with a compaction roller, the potential for fiber wrinkling is mitigated. With the in-situ tape placement additive manufacturing process, localized application of heat and pressure with a compaction roller allows for versatile fiber orientations supporting discrete fiber placement tailoring and optimizing structural properties.

The laser heating source is a recent modification to the process that provides faster processing, improved efficiency, more control, and real-time temperature logging. The process affects a smaller heat-affected zone, thus reducing the chance of damaging a part by limiting the number of times it is heated and cooled.

This process also supports unlimited geometric considerations regarding structure size and wall thicknesses since it mitigates exotherm concerns typical for processing thick epoxy composites. Continuous, unidirectional, graphite fiber-reinforced PEEK thermoplastic is ideally suited for the in-situ fiber placement process (Figure 6). The raw material is delivered as a prepreg tape, meaning that the raw fibers are preimpregnated with the thermoplastic resin. This is done offline at the material supplier's facility. Because the material is a thermoplastic, there is no shelf life and no "out-time" issues (i.e., the raw material can be stored at room temperature for an indefinite amount of time). This is different from thermoset materials, which require relatively strict storage temperatures and have a limited shelf life.

In addition to process advancements, improvements in the raw tape material

Figure 6. IM7/PEEK Thermoplastic Tape *(Source: TSS/Albany)*.

have resulted in reduced voids in the prepreg tape. This translates to a lower void content in the finished product. The material in tape form has a void content <1%, which is carried through to the final part configuration.

Due to the additive nature of the in-situ tape placement process, it is used for no-fastener assembly. This process involves: (1) fabricating individual reinforcements, (2) creating a segmented breakaway tool, (3) loading the tool with the reinforcing structures that are flush with the tooling outer surface, and (4) overwinding the entire structure fusing the reinforcing structures to the skin and yielding a coconsolidated structural assembly with no fasteners. Removal of the breakaway tooling releases the assembled structure to the next operation. A typical breakaway tool is shown in Figure 7, and the resulting coconsolidated structure with no fasteners is shown in Figure 8.

Using this process with a proprietary tape placement head, expanded copper mesh has been discretely placed within

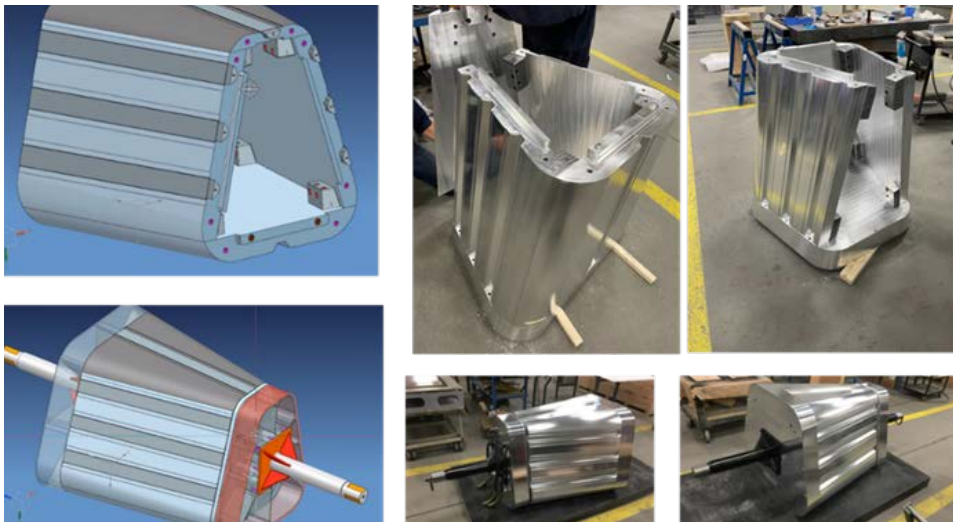


Figure 7. Breakaway Tool for Reinforced Fuselage Structure *(Source: H. R. Luzetsky)*.

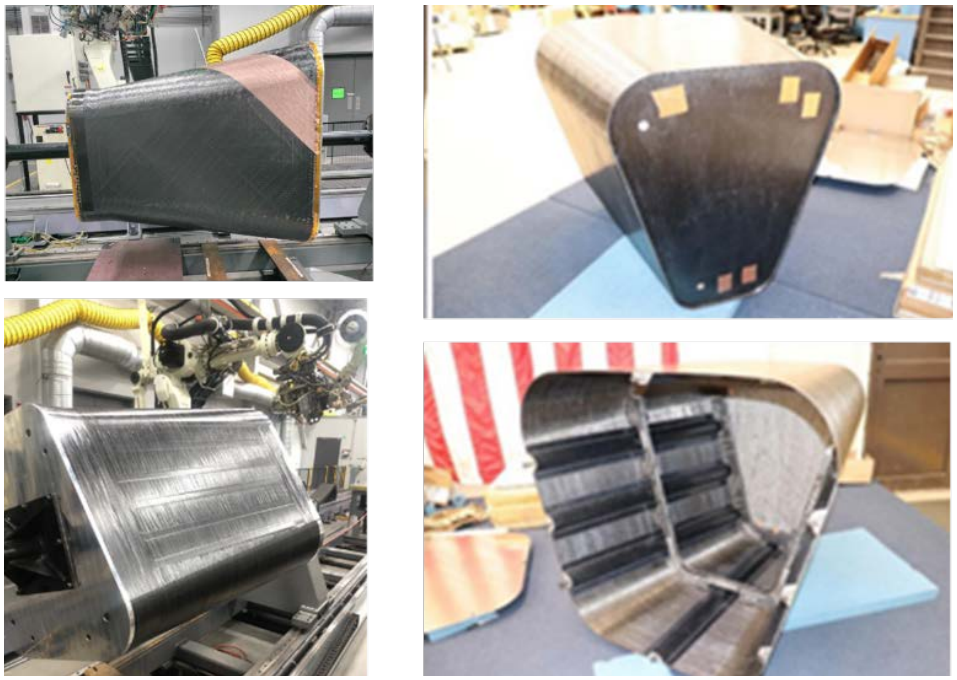


Figure 8. No-Fastener Assembly Structure *(Source: H. R. Luzetsky)*.

the laminate. This has resulted in a multifunctional composite form that provides both structural and EM shielding properties. A prototype unmanned aerial system (UAS) fuselage shown in Figure 9 exhibits EM shielding effectiveness between 60 and 110 dB, depending on frequency. The multifunctional material was developed by integrating expanded copper mesh into an IM7/PEEK laminate through the in-situ tape placement process. Its orientation within the laminate is critical to developing the mechanical and EM shielding properties as well as contributing to the overall physical properties (e.g., environmental resistance, damage tolerance, and durability).

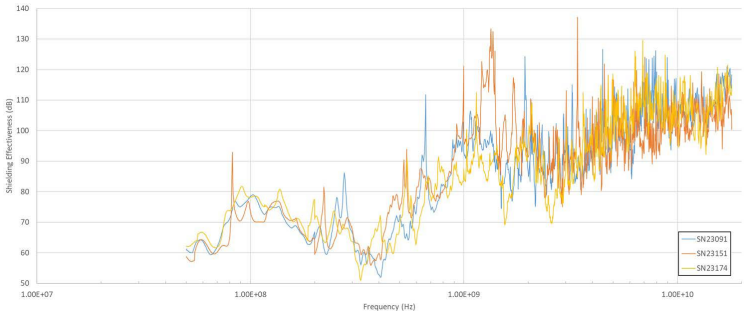


Figure 9. Multifunctional UAS Structure With Integrated EM Shielding and Lightning Protection *(Source: H. R. Luzetsky)*.

CCM

The basic CCM process (illustrated in Figure 10) begins with the feed creel, which is loaded with as many rolls of semifinished thermoplastic composite input material as required to form the specified laminate. Angled plies not readily spooled are cut and butt-welded edge to edge or tacked together to form a multilayer panel of a specified stacking sequence. This welded laminate is then loaded onto the feed creel. A hydraulic feed mechanism pulls the input material into stainless steel "preformers" prior to entry into a shaped die. The partially formed material is brought to flow temperature and pulled into a compression molding press with steel dies for the desired shape profile. The continuous consolidated shaped profile exits the press area and is automatically cut and stacked.

Figure 10. CCM Process Flow *(Source: TSS/Albany)*.

3-D Printing With Fiber-Reinforced PEEK Thermoplastic

The application of carbon/PEEK 3-D printing was explored for developing complex curvature structural sections to reduce cost and take advantage of integral reinforcement not possible with standard composite fabrication techniques. A Desktop Metal FiberTM 3-D printer was used for this program (Figure 11). This technology features two printheads—one for 3-D printing chopped, fiber-reinforced filament and another for laying down continuous fiber prepreg tape. While the continuous fiber tape process, dubbed micro automated fiber placement, permits fabrication of parts with up to 60% fiber volume fraction, the technology is limited to flat/nearly flat geometries. On the other hand, the chopped fiber filament process, which offers up to 30% fiber volume fraction and improved mechanical properties compared to standard thermoplastics, is nearly unlimited when it comes to geometry.

The application of carbon/PEEK 3-D printing was explored for developing complex curvature structural sections to reduce cost and take advantage of integral reinforcement not possible with standard composite fabrication techniques.

With 3-D printing, a complexity is developed in a structure that is

Figure 11. Desktop Metal 3-D Printer *(Source: TSS/Albany and Desktop Metal Proto3000 Fibre Brochure).*

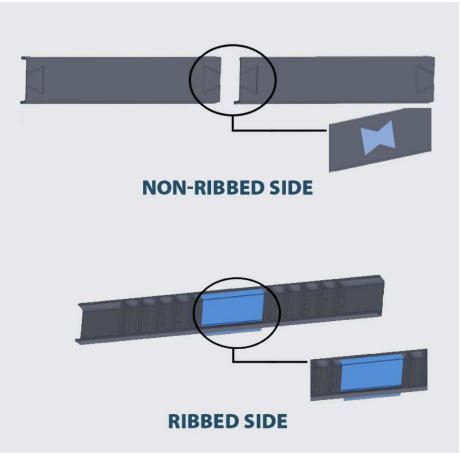


Figure 13. Frame Section Assembly Connections *(Source: H. R. Luzetsky).*

not possible with other processing techniques. The complexity of the structure compensates for property reduction associated with the 3-D process using chopped carbon fibers. However, there are advancements in the process that use continuous fibers to raise the mechanical properties compared to those associated with other processes (i.e., in-situ tape placement, press, and autoclave). This is illustrated with a frame that was developed with this process, shown in Figure 12. The frame was constructed of rib-reinforced sections like those shown in Figure 12 using the 3-D printing process. Each frame section had a connection joint included in the 3-D print. Each section aligns with one another, with their positions located/secured with a butterfly joint and joined with a splice plate on either side (see Figure 13). The butterfly joint and splice plates are secured to the frame sections with an epoxy adhesive. These joints are flush with the frame web area, and the plates are placed on either side of the web area, effectively tripling the thickness. In addition, the splice plates are ~3 inches long to completely cover the butterfly joint. The frame consists of 10 sections joined to one another with butterfly joints and splice plates, as shown in Figure 14.

Overmolding

Overmolding is a process by which a part is created with two or more materials combined into a single unit or component. An example is illustrated in Figure 15, where a fastener insert is molded into the component needing a point of connection such as that required for an access door or some other type of structure that normally would be secondarily attached to the composite structure. This process eliminates secondary operations that can drive costs from an installation and maintenance perspective.

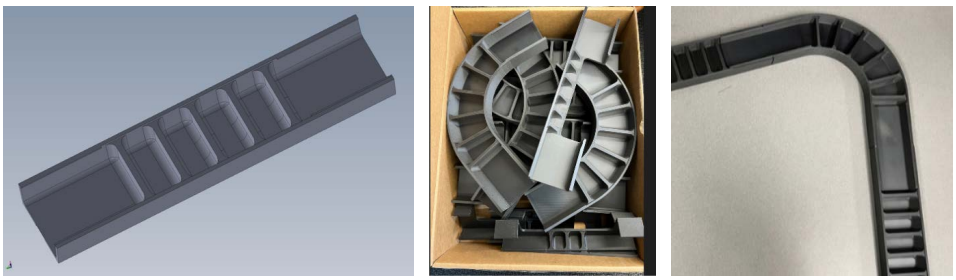
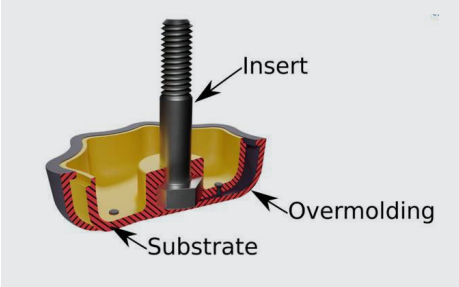
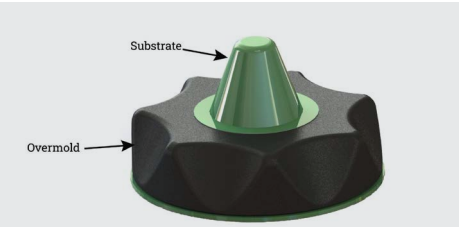


Figure 12. With 3-D printing, a complexity is developed in a structure that is not possible with other processing techniques *(Source: H. R. Luzetsky).*

> ***With 3-D printing, a complexity is developed in a structure that is not possible with other processing techniques.***

Figure 14. Assembled Frame Showing Ribbed Side (Left) and Reversed Side (Right) *(Source: H. R. Luzetsky).*





Overmolded Metal to Composite Example

Overmolded B-Bracket for Aircraft Storage Bin

Figure 15. Examples of Overmolded Components *(Source: CW CompositesWorld [1] and IQS Directory [2]).*

MATERIAL AND PROCESS COSTS

A carryover from the initial development is a continuing misconception that the use of fiber-reinforced PEEK costs more than an epoxy composite. This may seem to be true on the surface; however, accurate cost accounting must include life-cycle costs and consider costs associated with the volume of material used throughout global industry. When these elements are considered, the costs are comparable to an epoxy composite and may be significantly less. The lower costs are driven by the superior mechanical properties, which support a thinner structure and therefore less weight and material. With the reduction in weight comes enhanced performance and lower fuel costs. In addition, improved damage tolerance reduces maintenance costs,

A carryover from the initial development is a continuing misconception that the use of fiber-reinforced PEEK costs more than an epoxy composite.

which directly impact life-cycle costs. Finally, making full use of the material and its fabrication processes supports a no-fastener assembly through applying coconsolidation to fuse individual components to a base substrate. This process eliminates assembly costs, which contribute significantly to lower system costs. Finally, studies have shown that as the quantity of material used increases industry-wide, the raw material costs decrease to a level comparable to that associated with epoxy-based composite materials.

CONCLUSIONS

PEEK thermoplastic composite is suitable for applying to a primary structure. With proper design and fabrication techniques, the resulting structures can exhibit the following characteristics:

1. Enhanced performance, including damage tolerance and mechanical properties.
2. Reduction in weight for the same performance levels attributed to

the reduction or elimination of fasteners simplifying assembly.

3. Thinner gage materials required due to the improved mechanical and damage tolerance properties.
4. Lower life-cycle costs, which account for enhanced performance and reduced maintenance requirements due to improved damage tolerance and environmental resistance.
5. Reduction in structural fabrication costs associated with elimination of fasteners, out-of-autoclave processing, and reduction in parts' count.

One size does not fit all, and the fabrication process must be an integral part of the design process to optimize structural properties while minimizing costs and weight. ■

REFERENCES

[1] CW CompositesWorld. "Overmolded Hybrid Parts Open New Composites Markets." Accessed on 15 June 2024.

[2] IQS Directory. "Plastic Overmolding: What Is It? Processes, Types, Grades." iqsdirectory.com, accessed on 15 June 2024.

BIOGRAPHY

HARRY "RICK" LUZETSKY is a subject matter expert at the SURVICE Engineering Company, with more than 40 years of experience in composites and more than 30 years of experience in survivability. With a specific expertise in design, test, and research and development, he has helped develop and assess survivability features for numerous aircraft and has been active in composite design for vehicle performance and survivability improvements. He is the lead engineer for SURVICE's role in developing the thermoplastic drive shaft and a coauthor of a pending patent on an advanced fuel containment technology and fiber-reinforced structural composite faraday cage enclosure for electronics. Mr. Luzetsky holds a B.S. in materials engineering from Drexel University.

DSIAC Defense Systems Information Analysis Center

TECHNICAL INQUIRY SERVICES

FOUR FREE HOURS

Research within our 10 focus areas available to academia, industry, and other government agencies. Log in to https://dsiac.dtic.mil to submit your inquiry today.

TECHNICAL AREAS

Survivability & Vulnerabilty
Advanced Materials
Autonomous Systems
Non-Lethal Weapons
Weapons Systems
Military Sensing
Directed Energy
Energetics
RMQSI
C4ISR

Photo source: 123rf.com, SURVICE Engineering, U.S. Marine Corps, U.S. Navy, and U.S. Army

DSIAC JOURNAL

The Defense Systems Information Analysis Center (DSIAC) is a component of the U.S. Department of Defense's (DoD's) Information Analysis Center (IAC) enterprise, serving the defense enterprise of DoD and federal government users and their supporting academia and industry partners.

HTTPS://DSIAC.DTIC.MIL

CONNECT WITH US ON SOCIAL MEDIA