Identifying Patterns in Fault Recovery Techniques and Hardware Status of Radiation Tolerant Computers Using Principal Components Analysis

Fereshteh Ramezani, Christopher Major, Colter Barney,
Justin Williams, Brock J. LaMeres, Bradley M. Whitaker
Electrical and Computer Engineering
Montana State University
Bozeman, Montana, USA

***Abstract*—Fault tolerant computers have been developed in recent years to operate in the harsh radiation environment of outer space. These computers employ multiple copies of soft processors in a reconfigurable hardware environment and can automatically repair faults caused by radiation strikes. However, during certain recovery procedures, data collection and processing can be halted, and valuable scientific data can be lost. In addition, current fault recovery procedures may inadvertently make the computer more susceptible to faults or errors, for example, by introducing voltage and temperature changes. Machine learning feature extraction algorithms have the potential to reduce data loss by identifying patterns related to computational fault mitigation and recovery techniques. In this work, we will gather telemetry data from RadPC: a reconfigurable, radiation tolerant computer that has been developed over the past 12 years by Montana State University to advance high performance space computing under varying environmental conditions. RadPC has recently been configured to provide regular telemetry data to measure and communicate the performance of the radiation-tolerant computing platform. Specifically, the telemetry data includes information about data memory integrity, faults experienced, and successful repairs; as well as various measurements including voltage, current, and temperature. While RadPC has been under development for some time, the developers have never searched the telemetry data for associations between fault recovery procedures and the physical state of the hardware itself (e.g., voltage and current levels of power supplies or internal temperature). In this work, the computer will be subject to synthetic faults—emulating radiation strikes that may occur in space—and perform standard recovery procedures. The tests will be performed with the RadPC on a high-altitude balloon flight as well as inside a temperature-controlled vacuum chamber, allowing for a range of controlled external environmental conditions. The collected telemetry data will be analyzed using PCA to detect patterns in the hardware status associated with fault recovery techniques. Identifying these patterns may lead to improved fault mitigation strategies that reduce the risk of subsequent faults by considering how recovery techniques affect the physical state of the hardware.**

I. Introduction

Fault tolerant computers have been developed in recent years to operate in the harsh radiation environment of outer space [1]. These computers employ multiple copies of soft processors in a reconfigurable hardware environment and can automatically repair faults caused by radiation strikes [2]. Over the past 12 years, Montana State University (MSU) has been developing a technology called RadPC to advance high performance space computing [3]. While RadPC has been under development for some time, the developers have never searched the telemetry data for associations between hardware state and fault recovery procedures. In this work, we analyze the data using principal componenets analysis (PCA), a feature extraction algorithm that identifies important patterns in data. Analyzing these patterns may reveal a link between hardware state and mitigation strategies.

Computing systems in space missions cannot operate as done on the surface of the earth due to the harsh radiation environment existing in space. On earth, computing systems are protected from the cosmic radiation by the planet's magnetic field. However, computers operating in space without shielding have no protection from radiation in space missions [4]. Therefore, data processing in space needs reliable approaches to prevent or mitigate radiation-based faults.

There are two classes of space radiation effects on computing platforms: single-event effects (SEE) and total ionizing doses (TIDs) [5]. Usually single event effects can be corrected by restarting the system. However these events can sometimes cause deep malfunction in the system. In modern integrated circuit design ($<$ 65 nm), TID is becoming more and more statistically unlikely to occur due to feature sizes with small process nodes [6]. While the smaller feature sizes are reducing the possibility of damage due to trapped charge, the probability of functionality interruption caused by high-energy particles is drastically increased. Thus, the need to mitigate the damaging effects from SEEs becomes of greater concern than TID in modern computing systems [7].

To address the radiation problem in space computing, Montana state University has been developing a fault-tolerant system called RadPC, which can respond to radiation-induced SEE faults and undergo repairs as necessary. RadPC is a radiation tolerant computing system (RTCS) designed to address the issue of space radiation threatening the integrity of computing hardware. It uses FPGA hardware and commercial off-the-shelf (COTS) components to provide a more cost-efficient solution for space missions that require reliable, affordable, radiation-tolerant processing systems for space exploration [8]. The platform uses a 4-modular redundant (4MR) processor architecture in the field-programmable gate array (FPGA) fabric to redundantly run a program in parallel and continue operation when a processor is faulted by induced radiation. This is conducted by a voter system that signals accompanying components to partially reconfigure damaged processing tiles and scrub memory systems for faults. The RadPC payload contains dosimeters to correlate the performance of the RadPC computer to the radiation environment it resides in. There is also a thermal management system within the payload to regulate the internal temperature of the experiments.

The goal of this work is to perform a novel analysis of telemetry data collected while the system was automatically correcting for synthetically injected faults in various environmental conditions. Investigating the relationship between fault correction and hardware state may lead to improved fault mitigation strategies as the system continues to be developed.

II. Methods

A. Balloon Flight Test

The purpose of the high-altitude balloon experiment was to test a fault recovery procedure for external memory in flight computers. The fault recovery procedure is part of a larger effort at Montana State University to develop RadPC. This balloon flight specifically focused on detecting and correcting errors in external data memory for RadPC. The experiment developed consisted of two identical payloads—each containing the RadPC computer, the memory fault recovery system, and interface electronics to the balloon system. The payloads were flown on the Raven Aerostar Thunderhead high-altitude balloon system. The Thunderhead Flight Control Unit (FCU) provided DC power to the payloads and served as a communication bridge to a ground station where telemetry data could be retrieved during flight. The system flew for approximately 50 hours and generated around 20,000 packets of telemetry data.

One packet was transmitted approximately every 10 seconds. Each telemetry packet contained information about fault injections, as well as voltages and currents from the various power supplies (5V, 3.3V, 2.5V, 1.8V, and 1.0V DC). The 5V supply was used as the source for the four downstream regulators supplies. The 2.5V rail powered the payload's external microcontroller and provided the general purpose input/output (GPIO) level for the onboard FPGA. The 1.8V and 1.0V rails supplied the core and internal memory voltages for the FPGA. The temperature monitoring circuitry provides two measured payload temperatures, one from the microcontroller, and the other from the FPGA.

Also within the telemetry data was state-of-health information about the RadPC computer system and the memory experiment. The overall goal of the RadPC development is to advance the Technology Readiness Level (TRL) of a flight computer that can provide increased reliability in the presence of space radiation. This is accomplished by testing the system's ability to recover from faults in different locations within the computer. The occurrence of faults caused by the radiation of concern in orbit are relatively rate (2-3 per day in Low Earth Orbit) and even more so on a balloon flight at lower altitudes. In order to stress the RadPC memory experiment while in flight, a background fault injection system continually altered data fields in the external memory of the computer, causing errors that were detected by the recovery system and repaired through a recovery sequence. The telemetry data contained an ongoing log of the number of faults injected manually, the number of faults detected (both manually-induced and radiation-induced), and the number of faults recovered from. This allowed for mission success in demonstrating the experiment's ability to recovery from faults even in the event that no faults occurred due to natural radiation.

B. Thermal Vacuum Test

A preliminary Thermal Vacuum (TVac) Test of the RadPC architecture was performed in preparation for a Lunar surface mission scheduled for 2022. The test comprised of monitoring power performance and software capabilities under conditions such as cold start, hot start, and steady state temperature. A thermal vacuum chamber on the campus of Montana State University was used to stress the system. The system was placed in pressure less than 10 micro-torr. Survival testing was done at -45 and 60 degrees Celsius, and thermal cycling was done on cycles ranging from -35 to 50 degrees C.

C. Labeling Procedure

For investigating the associations between hardware state and fault recovery procedure, the telemetry packets of the two experiments were divided into two groups; the first group consisted of the packets with the incremented SEM fault counts and the second category were the rest of the packets collected during the normal procedure of RadPC. Table I shows the captured data by RadPC for the packets numbered 0x001d59 through 0x001d61. The highlighted line shows that an SEM fault was injected prior to transmitting packet 0x001d5e. This packet, as well as other packets corresponding to fault injections, are denoted as 'Recovery Packets.' The other packets are denoted as 'Standard Packets.'

Once an SEM injection error is recognized by RadPC, it tries to correct the fault and in case it succeeds to correct it, the SEM Correctable flag changes. By watching the corresponding telemetry packets, we hope to find a pattern that can explain the correlation between the fault recovery procedure and hardware state.

D. Analysis Procedure

Today, data sets are getting big and bigger in size and complexity that they need to be extracted and interpreted in a useful way. Principal Components Analysis, or PCA [9], is a dimensionality-reduction algorithm that is used to reduce the dimensions of the data by projecting it into a smaller dimension spanned by orthogonal 'principal components.' These components are selected to preserve most of the statistical

TABLE I
EXAMPLE OF TELEMETRY PACKET INFORMATION HIGHLIGHTING WHICH PACKETS WERE SELECTED AS 'INTERESTING.' EACH TIME THE INJECTION COUNT WAS INCREMENTED, THE RADPC WENT THROUGH ITS FAULT RECOVERY PROCEDURE. THE SUBSEQUENT TELEMETRY PACKET REFLECTS THE STATE OF THE HARDWARE WITHIN 10 SECONDS OF CORRECTING THE FAULT.

Packet Number	1.0 Voltage	1.0 Current	1.8 Voltage	1.8 Current	...	SEM Injection Count
0x001d59	1.0120	0.0603	1.7621	0.0922	...	0x0029
0x001d5a	1.0071	0.0612	1.7688	0.0833	...	0x0029
0x001d5b	0.9955	0.0625	1.7749	0.0765	...	0x0029
0x001d5c	1.0040	0.0544	1.7725	0.0838	...	0x0029
0x001d5d	1.0083	0.0610	1.7712	0.0787	...	0x0029
0x001d5e	1.0107	0.0573	1.7731	0.0878	...	0x002a
0x001d5f	1.0089	0.0585	1.7804	0.0784	...	0x002a
0x001d60	0.9967	0.0624	1.7664	0.0839	...	0x002a
0x001d61	0.9967	0.0618	1.7621	0.0904	...	0x002a

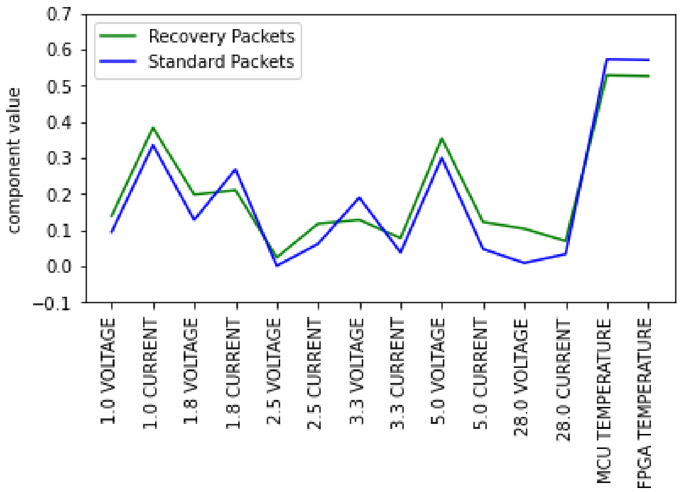

Fig. 1. Absolute value of first principal component as a function of features when analyzing the recovery packets and the standard packets obtained from the balloon test.

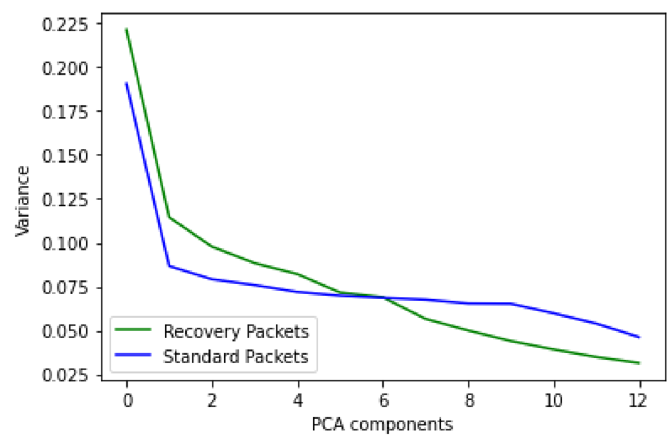

Fig. 2. Variance explained by each component when analyzing the recovery data and the standard data.

information in the data. PCA can be done in 5 steps explained as the following:

1) Standardize the dataset so that each of the features contributes to the analysis equally.
2) Compute the covariance matrix of the data in order to identify the correlations between the features. The covariance matrix is a symmetric matrix that provides the covariances of all possible pairs of the features.
3) Compute the eigenvectors and eigenvalues of the the covariance matrix. The eigenvectors are taken as principal components, and the corresponding eigenvalues indicate the amount of variance explained by that component.
4) Order the eigenvlaues and their corresponding eigenvectors in decreasing order based on the absolute value. This allows us to find the most significant principal components.
5) Choose the number of principal components to keep and project the data into the new reduced dimension.

For this experiment, 14 features of the telemetry packets, including voltage, current, and temperatures of the FPGA and microcontroller unit (MCU) were considered. Then PCA was applied to all three datasets: all the telemetry packets, the recovery packets, and the standard packets. Besides PCA, a normalized histogram, the mean, and the standard deviation of each feature was calculated for both standard and recovery packets.

III. RESULTS

To investigate the fault recovery procedure in RadPC, both data from the balloon test and the TVac test were used. After applying PCA to the recovery and standard packets from both tests, as well as the combined set of packets, 13 components were kept and different plots were made in order to be able to interpret the fault mitigation process.

For both experiments, the principal components calculated from the full dataset and those calculated from the set of standard packets were nearly identical. This is because the number of injected faults (≈100) was much smaller than the number of transmitted packets (≈20,000). For the remainder of the paper, we will discuss only the results obtained from the disjoint sets of recovery packets and standard packets. We also note that no radiation-induced faults were detected by the

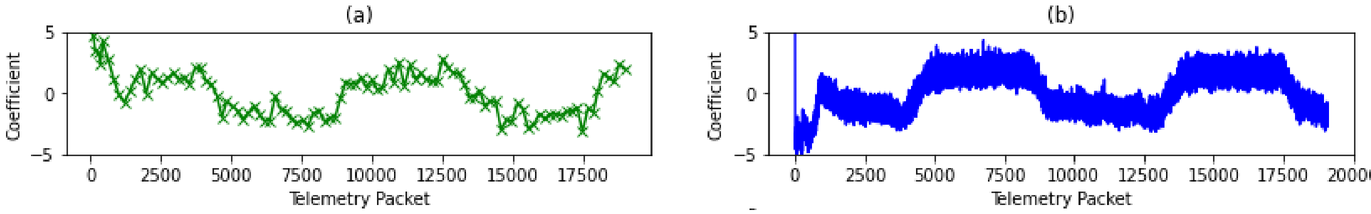

Fig. 3. Coefficient values as a function of Telemetry Packets analyzing the recovery data and the standard data. (a) First coefficient corresponding to the recovery data. (b) First coefficient corresponding to the standard data.

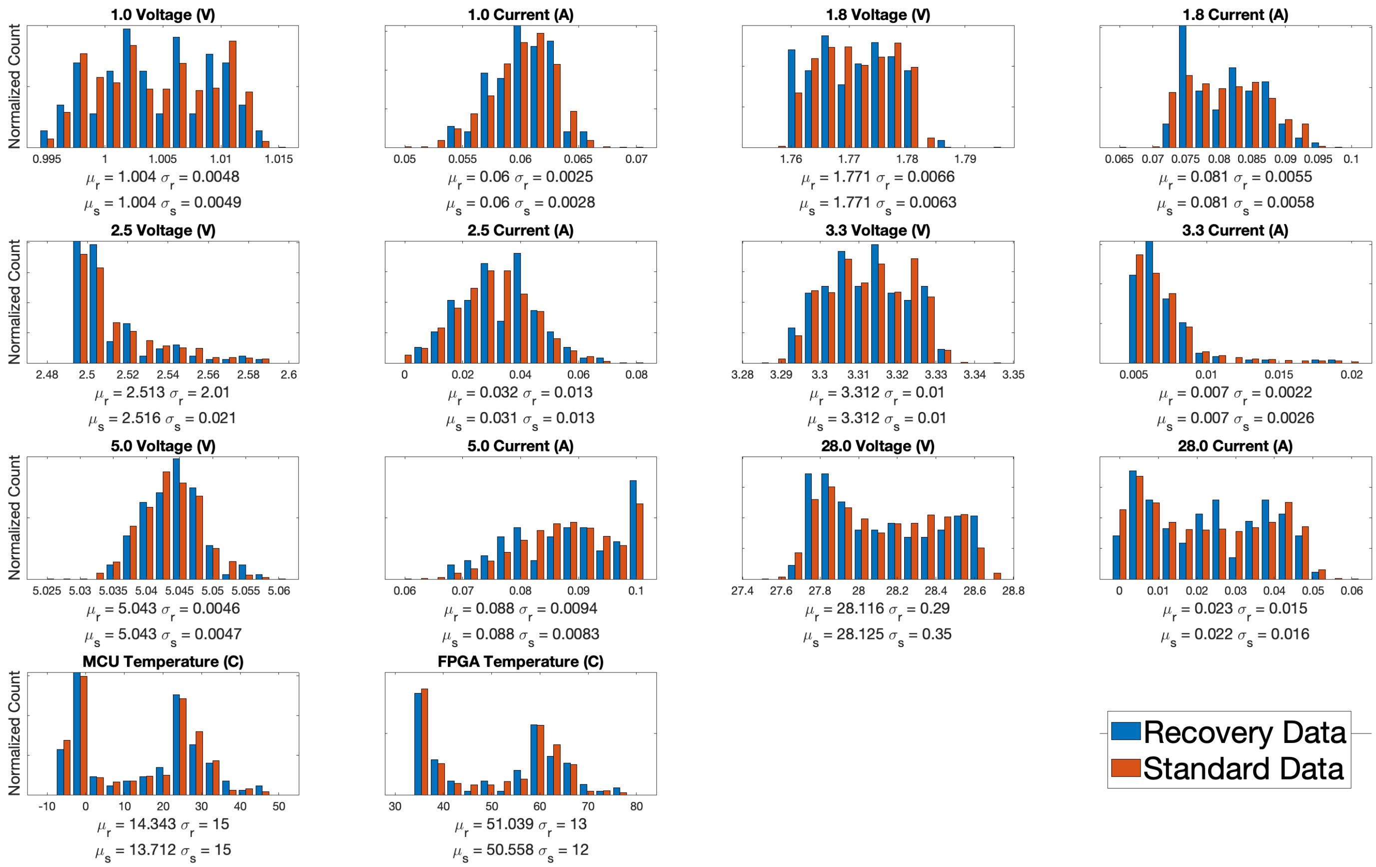

Fig. 4. Relative frequency (normalized histogram) of each measurement transmitted by the recovery (blue) telemetry packets and standard (red) telemetry packets. While the histograms for the two packet types are not identical, they are very similar. This indicates that the fault injection recovery procedure did not have a major effect on the state of the hardware.

system; for both the balloon test and the TVac test, all faults were synthetically injected.

A. Balloon Test

In Fig. 1, we have plotted the absolute value of each element of the first PCA vector as a function of input feature for both groups of telemetry packets. It is interesting to note that the component magnitudes for the recovery and standard packets are almost identical. This tells us that the distributions between recovery and standard telemetry packets are similar, since the combination of features that explains the most variance is the same for both sets. This, in turn, tells us that the fault recovery techniques do not drastically change the fundamental state of the hardware.

Fig. 2 shows the Percentage of variance explained by each of the selected components (variance ratio) for both data packets achieved from the balloon test. The plots shows that the components for both packets have similar variance patterns. However, it is important to note that the first principal component obtained from the recovery packets explain 22.14% of the variance in that dataset, while the first principal component from the standard packets only explain 19.07% of the variance. This indicates that there is more variability in set of the standard packets, which is to be expected since the size of that collection is much larger. Importantly, the first

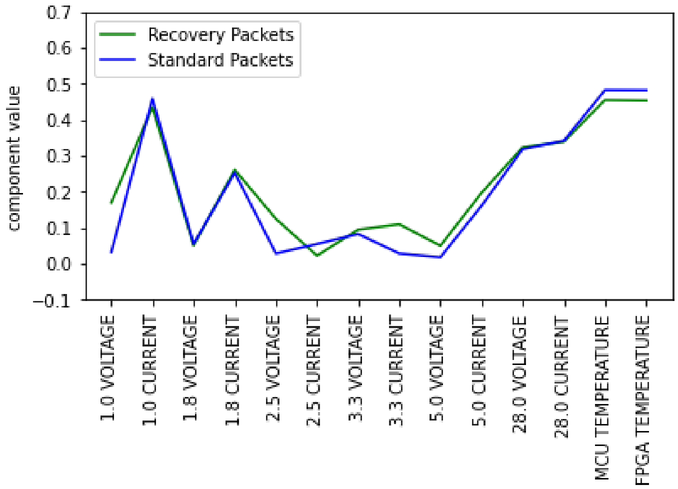

Fig. 5. Absolute value of the first principal component as a function of features when analyzing the recovery packets and the standard packets obtained from the TVac test.

component explained more than twice the amount of variance than any other component, indicating that a significant amount of statistical information is accounted for in this single component.

The coefficients relating the principal components back to the original feature space of 14 measurements is shown in Fig. 3. As can be seen in subplots (a) and (b), the patterns associated with the first component are nearly the same. (They are inverted because one of the principal components was almost exactly the negative of the other.) There is no clear pattern in the remaining coefficient usage for either dataset. However, in future work, individual component coefficients and reconstructed telemetry packets may be more closely examined to possibly reveal important patterns.

We have also plotted the normalized histogram of each measurement transmitted by the recovery and standard packets (Fig. 4). The mean and standard deviation associated with both packet types are presented for each feature below their respective distributions. Looking at the histogram figures, it can be seen that these values are consistent between the recovery packets and the standard packets. Even though there are several locations where the distributions appear to be different (e.g. 2.5 current plot near the middle), the differences are not large enough or common enough to make an effective discriminator between the standard and recovery packets. To this end, we can conclude that the standard and recovery packets come from the same statistical distribution, which indicates the fault recovery procedure implemented by RadPC does not significantly affect the hardware status of the system.

In summary, we have investigated three pieces of evidence to determine if the RadPC recovery procedure affects the state of the hardware: PCA, 1st- and 2nd-order statistics, and feature distributions. Small differences in all three analyses indicate that standard packets and recovery packets come from the same statistical distribution. This indicates that recovery procedure does not significantly affect the health of the hardware.

B. TVac Test

Similar findings as reported previously were discovered when analyzing telemetry packets obtained during the TVac test. The variance pattern and coefficient utilization for the TVac test match those observed during the balloon test. Again, the TVac test also showed that the first principal component calculated from both datasets was nearly identical—see Fig. 5. (This was the same behavior observed during the flight test and described in Fig. 1.) Feature statistics and distributions were also similar between the recovery and standard telemetry packets. For brevity, we do not include figures of the time-series coefficient values or histograms associated with the TVac test, but the results were similar to those obtained during the balloon test.

IV. Conclusion

The results from applying PCA to telemetry data collected from the balloon and TVac experiments show that the first component for both recovery and standard packets have almost the same distribution, Together with the results of mean, variance, and normalized histogram between the two types of packets, we can conclude that the fault-mitigation technique used in RadPC does not significantly affect the hardware state.

Ongoing work includes exploration of other pattern recognition methods like sparse coding to investigate the correlation between fault recovery technique and hardware state of fault tolerant computers (RadPC). In addition to training linear methods to determine patterns between telemetry data and current faults, we plan to identify associations between telemetry data and future faults, or the future state of the hardware. Because of the time-varying nature of this task, we plan to use a deep Long Short-Term Memory (LSTM) neural network [10] to identify these associations. While using the deep network will make it impossible to directly map associated outcomes to individual telemetry features, it is likely to identify patterns in the time series data that would be infeasible to analyze with linear feature extraction methods.

Another avenue of future work could be to repeat these experiments with more frequent fault injection. Alternatively, a telemetry packet could be sent immediately after a fault is detected. This would reduce the average time between fault injection and receiving the telemetry information, allowing for an analysis of the immediate effects in hardware status.

This work was supported by the NASA Flight Opportunities Program under CAN# 80NSSC20K0107.

References

[1] T. M. Lovelly, T. W. Wise, S. H. Holtzman, and A. D. George, “Benchmarking Analysis of Space-Grade Central Processing Units and Field-Programmable Gate Arrays,” *J. Aero. Info. Sys.*, vol. 15, no. 8, pp. 518-529, 2018.

[2] R. Glein, F. Rittner, A. Becher, D. Ziener, J. Frickel, J. Teich, and A. Heuberger, “Reliability of Space-Grade vs. COTS SRAM-Based FPGA in N-Modular Redundancy,” *2015 NASA/ESA Conf. on Adap. Hard. and Sys*, pp. 1-8, 2015.

[3] B. J. LaMeres and C. Gauer, "Dynamic reconfigurable computing architecture for aerospace applications," *2009 IEEE Aerospace Conference,* 2009.

[4] J. L. Barth, C. S. Dyer, and E. G. Stassinopoulos. "Space, atmospheric, and terrestrial radiation environments," *IEEE Transactions on Nuclear Science,* vol. 50, no. 3, pp. 466-482, 2003.

[5] C. Claeys, E. Simoen, M.-A. Trauwaert, and J. Vanhellemont, "Electrical and Physical Characterization of Proton and Cf-252 Bulk Radiation Damage," *MEC Report*, p. 352, 2002.

[6] H. J. Barnaby, "Total-ionizing-dose effects in modern CMOS technologies," *IEEE Transactions on Nuclear Science*, vol. 53 no. 6, pp. 3103-3121, 2006.

[7] C. M. Major, A. Bachman, C. Barney, S. Tamke, and B. J. LaMeres, "RadPC: A Novel Single-Event Upset Mitigation Strategy for Field Programmable Gate Array–Based Space Computing." *Journal of Aerospace Information Systems*, vol 18, no. 5, pp. 280-288, 2021.

[8] C. Major, B. J. LaMeres, D. Klumpar, L. Springer, J. Sample, S. Tamke, R. Meuchel, C. Barney, A. Bachman, and J. Davis, "Overview of the Upcoming RadPC-Lunar Mission." *2021 IEEE Aerospace Conference*, 2021.

[9] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems* vol. 2, nos. 1-3, pp. 37-52, 1987.

[10] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep Learning with Long Short-Term Memory for Time Series Prediction," *IEEE Communications Magazine*, vol. 57, no. 6, pp. 114-119, 2019.