AIAA SciTech Forum
January 3-7, 2022, San Diego, CA & Virtual
AIAA SCITECH 2022 Forum

10.2514/6.2022-1271

Estimating Core Body Temperature Under Extreme Environments Using Kalman Filtering

Ezequiel Juarez Garcia*
University of Florida, Gainesville, FL, 32611

David P. Ferguson†
Michigan State University, East Lansing, MI, 48824

Nicholas J. Napoli‡
University of Florida, Gainesville, FL, 32611

Extreme temperature or physiologically demanding environments often present in aerospace missions pose a high risk of heat stress to pilots and astronauts which can lead to heat illnesses and human performance decrements. This is particularly prevalent within military aircraft, where many flight research installations and airfields are located in hot arid desert or high-humidity tropical climates. The intense heat in these environments can exacerbate the severity of the heat stress already present in the pilot due other physiological and environmental stressors. To measure a key biological indicator of the level of heat stress, core body temperature, we propose a noninvasive method for subjects under extreme high thermal stress using heart rate and skin temperature measurements from mobile biosensors in real open-world environments. As an analog for pilots operating in extreme thermal environments, we utilize observations of professional race car drivers exposed to high thermal stress inside vehicle cockpits for several consecutive hours. The conditions experienced by the drivers not only incorporate the heat stress produced by the layered protective equipment but also the thermal stress from the operational environment and vehicle. A Kalman filter is designed to predict core body temperature utilizing linear models generated by the driver's heart rate and skin temperature sensors. The data obtained at 15 races from 4 different drivers was used to train the linear models and validate the Kalman filter. Ground truth core body temperature measurements obtained from ingestible core temperature capsules were used to measure the filter's performance. Despite the number of factors that negatively affected the quality of the biosensor data, such as stresssors due to the operational environment and vehicle (e.g., extreme ambient temperatures, g-forces, heat dissipation impairment), sensor dropouts and anomalies, and numerous physiological stressors imposed on the drivers (e.g., elevated heart rate, dehydration), our method was robust enough to be comparable to studies done in less complex environments.

I. Nomenclature

C	=	core body temperature reading
$\widehat{C}$	=	core body temperature reading estimate
H	=	heart rate reading
S	=	skin temperature reading
b_i	=	i-th coefficient used in the Kalman filter observation model
E_{RMSE}	=	root-mean-squared error

*Graduate Research Assistant in Human Informatics and Predictive Performance Optimization Lab, Department of Electrical and Computer Engineering, AIAA Student Member (e-mail: ejuarezgarcia@ufl.edu).

†Assistant Professor, Department of Kinesiology (e-mail: fergu312@msu.edu).

‡Assistant Professor and Director of Human Informatics and Predictive Performance Optimization Lab, Department of Electrical and Computer Engineering, AIAA Senior Member (e-mail: n.napoli@ufl.edu).

II. Introduction

Infrequent exposure to extreme temperature or physiologically demanding environments during aerospace missions can place undue thermal stress on pilots and astronauts due to lack of acclimation, among other factors. In particular, heat stress in pilots can, for example, stem from high physiological and cognitive workloads during complex mission tasks, heat dissipation impairment due to layered protective wear, radiating heat from the avionics and airframe, and higher temperatures in the cockpit compared to the ambient temperature [1]. Prolonged exposure to these conditions can hinder the body's ability to regulate core body temperature (thermoregulate) and lead to degradation in the cognitive performance of pilots [2], increasing the likelihood of an accident. The physiological demanding aspect of extreme environments comes naturally in military aviation due to the level of peak performance required from pilots. The other aspect of extreme environments, intense heat, does not only encompass high ambient temperature environments (e.g., arid deserts, high-humidity tropical climates) that airfields may be located in, but also heated environments due to other internal (e.g., psychophysiological) and external (e.g., vehicle, equipment) heat stressors acting on the pilot. These sources of heat stress and others are presented in Fig. 1.

Core body temperature is a key biological indicator of the effects of physiological stress and physical activity on human performance. Careful monitoring of core body temperature can prevent or mitigate physiological strain produced by heat stress [3]. This is especially important in environments that demand peak performance to operate within them or environments that subject personnel to high temperatures. These types of extreme environments can extend beyond the cockpit of an aircraft and into space missions [4, 5], boots on the ground military operations [6], and professional sports [7]. Consequently, insight into extreme aerospace environments can be gained by studying thermal environments in professional sports, such cycling and automobile racing, where participants can experience abnormally high body temperatures [8, 9]. Failure to track core body temperature and stay within thermoregulatory limits, increases the likelihood of diminished performance and harm to the individual.

To measure core body temperature, traditional methods require the use of ingestible core temperature capsules or rectal probes. These invasive methods cannot be reused or retrieved after they have passed through the body. Moreover, if they have not sufficiently made their way down the digestive tract, ingestible core temperature capsules are known to be susceptible to changes in temperature due to the consumption of hot or cold beverages [10]. Thus, accuracy and precision of the measurement are constrained by the timing of the consumption of the ingestible sensor and the duration that the sensor stays within the digestive tract. This creates cost concerns and limitations, since several probes are typically required for each subject to maintain reliable and consistent core body temperature measurements.

Prior Work. Noninvasive methods for estimating core temperature which rely on Kalman filters, and its variants, have been proposed throughout the literature for different populations [11–14]. Though these studies address the estimation of biological indicators of heat stress, the over-reliance on controlled environments or specific exercises by some of them to simulate activity in extreme environments reduces the applicability of their methodology for aerospace missions or other extreme environment scenarios. Moreover, studies that focus on real world environments only tend to consider the heat stress build-up caused by protective equipment and not the added physiological stress imparted on the individual by complicated vehicle maneuvers (e.g., g-forces on pilots) and radiating heat from vehicle equipment. *This raises one fundamental research question, when compared to previous studies, can we achieve similar predictive performance of core body temperature in extreme environmental conditions using noninvasive mobile biosensors?*

Challenges. One of the most significant problems when studying thermal stress or other physiological systems within aerospace environments is the difficulty obtaining the approval and validation that the additional sensors added do not impact safety, especially in military aircraft [15]. These safety concerns range from the interference of sensitive electronic equipment to sensors causing physical obstructions during emergency ejections. Thus, simply adding additional human wearable sensors within an aircraft takes numerous exploratory safety tests causing experimental constraints in money and time. Furthermore, once human data collection is approved within the cockpit, an uncontrollable number of factors in this complex environment still exist that can affect the quality of the experimental data. These factors include physical vibrations of the equipment, electromagnetic interference, and multiple concurrent physiological stressors on the human. These physiological stressors are precisely the reason the algorithm approach is challenging. For example, while the pilots are being subjected to high thermal stress from temperatures inside and outside the cockpit, they also have to contend with intense g-force maneuvers, poor heat dissipation, regulated breathing, chest-wall constriction, and high cognitive workloads while being confined to their seating position for prolonged periods.

Insights. Due to the difficulty of measuring thermal stress within pilots, an experimental analog to mimic similar environmental and concurrent physiological conditions can be created through monitoring high-performance race car drivers. The conditions experienced by sports car drivers are perfect for studying thermoregulatory strain caused by the operation of performance demanding vehicles in heat stressed environmental conditions. Furthermore, similarly to

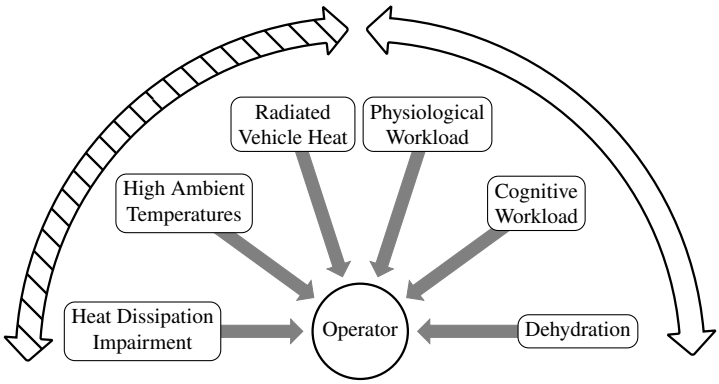


Fig. 1 Examples of external (dashed arrow) and internal (empty arrow) heat stressors experienced by pilots and drivers.

pilots, drivers also face the build-up of internal stress from using protective body equipment for several consecutive hours and the physiological stress imposed by the vehicle maneuvers. Radiating heat sources and temperatures upwards of 50°C [16] inside the cockpit further contribute to the thermal load on the driver.

Due to advancements in noninvasive wearable technology, the sensors in these devices are capable of measuring a plethora of biological signals such as heart rate, skin temperature, breathing rate, etc. To increase the accuracy of core temperature estimation, we incorporate other modalities such skin temperature readings into the estimation process. The use of skin temperature or other biological signals besides heart rate has not seen much attention outside medical settings, where, under stable environmental conditions, reliable estimates can be obtained.

Contributions. This study lays the foundation for potentially utilizing high performance race car driver data as an analog for human data collection within aerospace applications. Furthermore, we demonstrate the ability to estimate core temperature under the environmental conditions experienced by professional race car drivers which can then be later translated to aircraft pilots. Due to other environmental conditions causing concurrent physiological stressors, we gain algorithmic insights on how to improve the prediction of core body temperature for future work.

III. Background

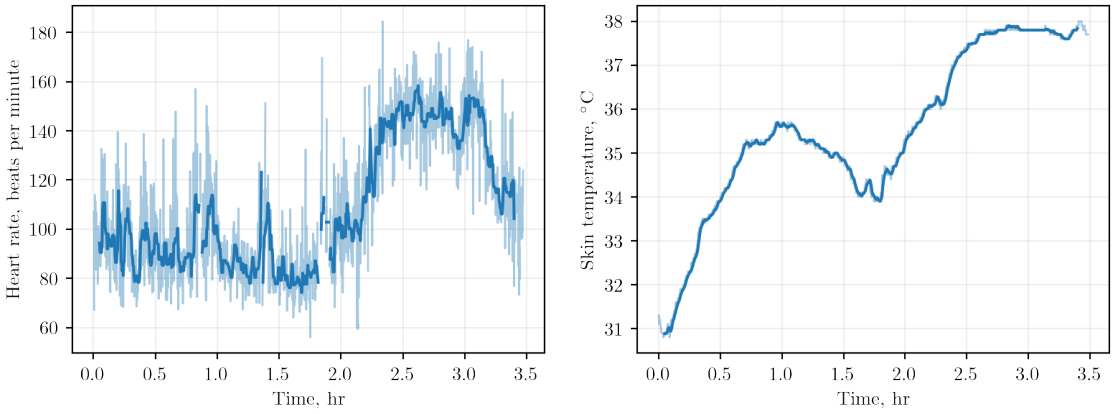
A. Participants

The data was collected from 4 male sports car drivers, between the ages of 25 and 42 years old, at 7 different raceways during the 2017 season of the IMSA Continental Tire SportsCar Challenge. Approval was obtained from the Institutional Review Board prior to the start of the study. Two of the participants were professional drivers and the other two were amateur drivers. All participants belonged to the same racing team and operated identical cars. Therefore, all drivers experienced relatively the same amount of physiological stressors of g-force, temperature, and other factors.

B. Data Collection

The drivers ingested an Equivital VitalSense core temperature probe and were equipped with the Equivital LifeMonitor ambulatory monitoring system (EQ02) prior to entering the race car. The EQ02 was strapped over the chest of the drivers and captured heart rate and skin temperature. Both heart rate and skin temperature readings were captured every 15 seconds simultaneously. The EQ02 was also responsible for receiving the measurement transmissions of the core temperature probes ingested by the drivers. The probes provided the ground truth measurements for core temperature. Each probe was single use and transmitted a temperature reading about every 15 seconds, with occasional dropouts and irregular sampling intervals.

A total of 28 different races were captured in the original dataset. During each race, two drivers, one professional and one amateur, were selected to pilot a race car, with the control of the car switching from one driver to the other at some point in the middle of the race. Data collection started a few minutes before the race and ended at a nonpredetermined time during or after the race. The length of the races ranged from 1.5 to 4 hours depending on the exact distance of the race. Additional details of the data collection process and races can be found in [17].



(a) (b)

Fig. 2 Example of the signal processing done on the (a) heart rate signal and (b) skin temperature signal. The light colored signal is the original and unfiltered signal. The darker signal has been processed.

C. Measurements Processing

The heart rate, skin temperature, and core temperature measurements were inspected manually to remove corrupted data and other errors. From the original 28 samples (i.e., races), 13 samples exhibited anomalies in the ground truth core temperature measurements obtained by the ingestible core temperature capsule. These anomalies included core temperature data dropouts for most of the data collection period and rapid downward changes in core temperature of more than 8°C per minute. We attribute the former anomaly to electromagnetic interference with the low frequency radio waves transmitted by the capsule [18], and the latter anomaly to the rapid consumption of cold fluids before the capsule had traveled sufficiently far into the digestive tract. As previously discussed, core temperature capsules are susceptible to these types of rapid changes in temperature, thus rendering some readings invalid for our study.

After we removed the samples that featured the aforementioned anomalies from the data, the number of samples decreased to 15. These samples were further filtered to remove periods before and after experiment where the EQ02 was no longer strapped to the chest of the drivers and reported false readings. Additionally, heart rate readings with a confidence value of less than 80% were removed. These confidence values were provided by the EQ02, one for each heart rate reading. Due to the signal dropouts in the temperature capsule, the timestamps of the core temperature readings did not necessarily match the timestamps of the heart rate and skin temperature readings. To produce the ground truth core temperature readings, we resampled the core temperature signals using linear interpolation to match the timestamps of the heart rate and skin temperature readings. To finalize the processing phase, the heart rate and skin temperature signals from each sample were passed through a 5-sample median filter to attenuate sharp spikes or outliers in the signals due to noise. No further filtering was done to the signals to properly capture the noise characteristics during the model learning stage of our methodology. An example of the aforementioned signal processing is shown in Fig. 2.

IV. Methods

A. Kalman Filter

The Kalman filter used in this study is broken down into two steps: a prediction and an update step. In the prediction step, the core temperature evolves according to the *time update model*:

$$\widehat{C}_k = C_{k-1} + w_k^c, \tag{1}$$

where C is the core temperature, the subscript k denotes the time step, and $w^c \sim \mathcal{N}(0, \sigma_c^2)$ is zero-mean Gaussian noise with variance σ_c^2. The hat denotes a value before the update step. The uncertainty in the core temperature prediction is captured by the state variance

$$\widehat{p}_k = p_{k-1} + \sigma_c^2. \tag{2}$$

In the update step, when a measurement of heart rate or skin temperature is provided to the Kalman filter, the following *observation model* is used to predict the measurement:

$$\widehat{y}_k = b_1 \widehat{C}_k + b_0 + w_k^y, \tag{3}$$

where y denotes either a heart rate measurement H or a skin temperature measurement S, and $w^y \sim \mathcal{N}(0, \sigma_y^2)$. The coefficients b_1 and b_0 of the model are obtained from linear regression. Using the observation model, we can write the other equations used in the Kalman filter. The Kalman gain is calculated using

$$K_k = \frac{\widehat{p}_k b_1}{b_1^2 \widehat{p}_k + \sigma_y^2}. \tag{4}$$

The final two equations in the Kalman filter update the state prediction C_k and its variance p_k:

$$C_k = \widehat{C}_k + K_k(y_k - \widehat{y}_k) \tag{5}$$

$$p_k = \widehat{p}_k(1 - K_k b_1), \tag{6}$$

B. Model Coefficients Calculation

The 15 races available after processing the original data were partitioned to create training and validation sets using k-fold cross-validation. The decision to use this form of cross-validation was to avoid reducing the number races further if we performed a percentage split of the data. Moreover, the collection of the time series data for each driver and raceway into a separate archive made each race easily amenable to act as a fold. During each iteration of the cross-validation procedure, 14 folds were used to learn the model coefficients in Eqs. (1) and (3), and remaining fold was used for validation. The coefficients in the observation models are calculated using linear regression with an ordinary least squares loss function. We justify our use of linear models due to the robustness of the Kalman filter under nonideal conditions. Studies such as [19, 20] have used nonlinear models for heart rate and skin temperature, though with minimal gain in performance. With the use of more complex models with a variant of the Kalman filter, careful consideration must be taken to avoid overfitting.

The variance σ_c^2 of the noise used in the time update model is obtained by performing a Gaussian distribution fit of the core temperature differences $C_k - C_{k-1}$. The variance of the distribution fit is set equal to σ_c^2. This assumes that the core temperature differences are normally distributed with zero mean, which is an assumption made based on empirical evidence in similar studies [13]. The variance σ_y^2 for a sensor measurement in the observation model is obtained differently. After calculating the coefficients of the model using linear regression, the standard deviation σ_y is taken as the RMSE between the readings predicted by the linear model and the actual readings. From this, we obtain σ_h^2 and σ_s^2 for heart rate and skin temperature, respectively.

C. Performance Metric and Summary

To measure the performance of our method, we compute the RMSE between the true set of core temperature readings $\{C_k\}_{k=1}^N$ provided by the core temperature probe, and the set of estimated readings $\{\widehat{C}_k\}_{k=1}^N$:

$$E_{\text{RMSE}} = \sqrt{\frac{\sum_{k=1}^N \left(C_k - \widehat{C}_k\right)^2}{N}}$$

A summary of our Kalman filter approach for estimating core body temperature is given in Algorithm 1.

V. Results

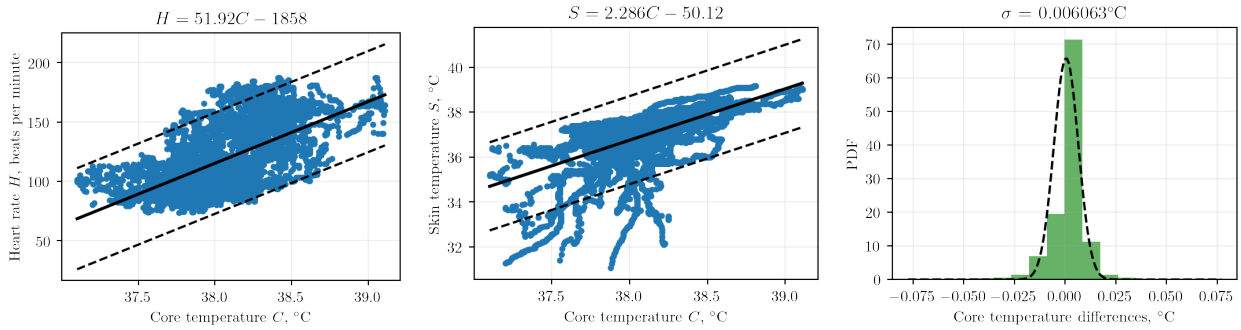
The model fits obtained for the observation models Eq. (3) are shown in Fig. 3(a) & (b). The linear regression coefficients for heart rate and skin temperature are summarized in Table 1 for the first training set. The variances used in the observations models are also provided in the table. Other training sets exhibited similar coefficients. A normalized histogram and a Gaussian distribution fit of the core temperature differences is provided in Fig. 3(c). The distribution is for core temperature differences 15 seconds apart. This is equal to the time step duration of the Kalman filter. The variance of the fit is provided in Table 1 for the first training set. The linear fit coefficients for the time update model are omitted because they do not apply to the model.

Algorithm 1 Core temperature estimation using Kalman filter

Inputs: C, H, S regression coefficients and variances; known initial C, H readings $\{H_k\}_{k=1}^{N}$; S readings $\{S_k\}_{k=1}^{N}$
Outputs: core temperature estimates $\{\widehat{C}_k\}_{k=1}^{N}$
for each time step k **do**
  predict core temperature using (1) & (2) $\longleftarrow \widehat{C}_k,\ \widehat{p}_k$
  for each observation (y_k, σ_y^2) **do**
    compute Kalman gain using (4) $\longleftarrow K_k$
    update core temperature estimate using (5) & (6) $\longleftarrow C_k,\ p_k$
  end for
end for

Table 1 Training Set 1: Coefficients and variances for time update (C) and observation models (H, S)

Model	b_1	b_0	σ^2
C	—	—	3.676×10^{-5}
H	51.92	-1858	453.2
S	2.286	-50.12	0.962



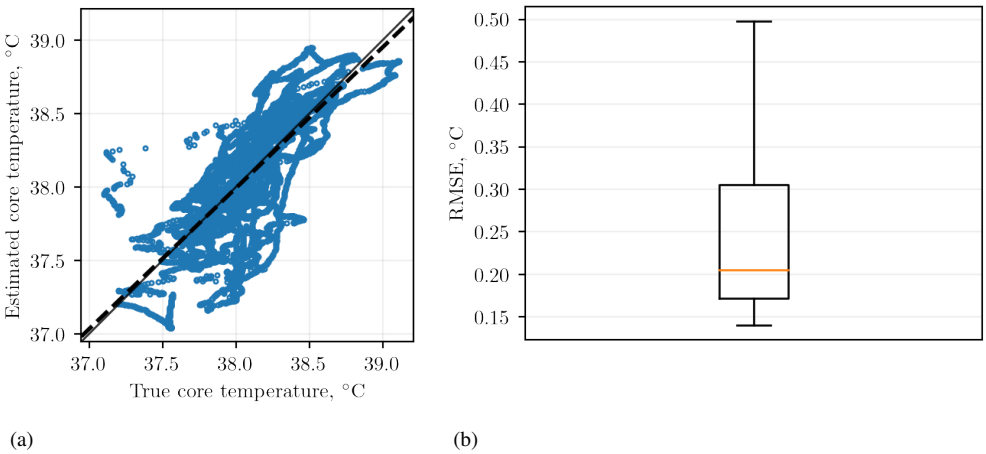
(a) (b) (c)

Fig. 3 Training Set 1: Linear regression for (a) heart rate and (b) skin temperature. For (a) and (b), the linear trend is shown in solid black and 2σ bounds are shown in dashed black. (c) Normal distribution fit for core temperature differences.

The results of using the Kalman filter across all validation sets are shown in Fig. 4 and in Fig. 5 in the Appendix. The true vs. estimated core temperature readings are shown in Fig. 4. Ideally, the majority of the data should lie along the identity line, with minimal deviation from it. The performance achieved by the Kalman filter on the validation sets was an $E_{\text{RMSE}} = 0.249 \pm 0.0955$°C (mean ± standard deviation, SD), assuming a normal distribution of the error. The error spread can be visualized in Fig. 4(b). In the Appendix, Fig. 5 provides the Kalman filter core temperature estimates of all validation sets.

VI. Discussion

Despite the violation in the linear model assumption, the linear regression results in Fig. 3 show that the linear fits for heart rate and skin temperature capture the general trends in the data to a great extent. This is in agreement with the studies discussed in the prior work that also use linear models to relate core body temperature to heart rate and other modalities. That being said, upon closer inspection, the imperfections in the fits become more noticeable. For example, the 2σ (±2 SD) bounds for the heart rate fit, shown in Fig. 3(a), leave out more heart rate measurements in the core temperature range of 37.75–38.5°C than in the ranges found at the extremes of the data points. These data points may include elevated heart rate readings caused by g-force oxygen deprivation as the driver begins to get acclimated to the race and the core temperature has not reached peak value (greater than 38.5°C). The large and uneven distribution of the



(a) (b)

Fig. 4 Kalman Filter Results: (a) True core temperature readings from the validation set plotted against the estimated readings. The linear trend of the data is shown as a thick dashed line. The identity line is shown as a thin solid line. (b) RMSE distribution of the validation sets.

data points in this range contributes to the rapid overestimation and, in particular, underestimation that occurs when starting out near these core temperature values, especially when the only core temperature reading assumed to be known (i.e., provides a correction) is the one at the beginning of the estimation. Examples of the rapid deterioration of the core temperature estimate due to underestimation can be found in the subplots for Validation Sets 1, 4, 8, and 10 in Fig. 5.

Similarly, the skin temperature fit shown in Fig. 3(b) also violates the linear relationship between skin temperature and core temperature below a core temperature of 38.25°C, as indicated by the branching out of the data points in many directions. We attribute the branching out to the "warm-up" period exhibited by the skin temperature sensor in the EQ02 sensor suite. During the warm-up period, which varies in length between races and drivers, the skin temperature readings climb from a "cold start" value to a value that reflects the approximate skin temperature at the sensor location. Beyond this skin temperature value, the relationship between skin and core temperature becomes more linear. The warm-up period is caused by the inhibition of heat dissipation through perspiration by the insulated race suit on the drivers [21]. This can potentially lead to severe dehydration through excessive sweating if the driver does not hydrate adequately. The warm-up period can be visualized in Fig. 2(b), although to a lesser degree than in other samples, as the skin temperature climbs from approximately 31°C to 34°C in the span of half an hour, before climbing again at the start of the race near the 2-hour mark. The skin temperature readings collected during the warm-up period must be carefully processed, as simply discarding them negatively impacts the performance of the Kalman filter, possibly due to creation of a less generalizable model, albeit a more linear one.

Further analysis of the estimates produced by the Kalman filter across all validation sets provides us with invaluable insight for future improvements to our experimental design. For example, the group of outliers present above the main linear cluster was a direct consequence of the rapid decrease in core temperature at the end of Validation Set 2, shown in Fig. 5. Rapid changes in core temperature, also exhibited in the other validation sets to a lesser extent, are not tracked by the Kalman filter very well due to the higher trust that is placed on the time update model compared to the observation models. Thus, the Kalman filter reacts slowly (i.e., it is not as responsive) to rapid changes in core body temperature. Additionally, as previously discussed, large underestimation errors tend to occur near the core temperature range of 37.75–38.5°C at the beginning of the estimation process. A complementary explanation to this peculiarity is explained by the variance in the baseline heart rate readings that the drivers exhibited at the beginning of each race, explaining the poor fit in the lower heart rate extreme in Fig. 3(a). These differences in baseline heart rate contributed to the poor performance of the Kalman filter prior start of the race, and the lift in performance that occurred during and after the race, where the heart rate and core temperature model is shown to have a more pronounced linear, yet still unevenly distributed, relationship due to the driver's elevated core body temperature. Relatedly, the heart rate model resulted in large overestimation errors when the true core body temperature was not sufficiently high enough (e.g., greater than 38.25°C), as shown in Validation Set 13, Fig. 5.

In spite of the imperfect linear fits and the influence of numerous stressors on the readings, the Kalman filter we

used resulted in an RMSE of 0.249 ± 0.0955°C with a median of 0.205°C as shown in Fig. 4 . This error is comparable to the RMSE obtained in [19] of 0.23 ± 0.08°C when validated on U.S. Marines on patrol missions in 42–47°C weather in Iraq, and an improved model proposed in [13] with an RMSE of 0.28°C. We note that the authors in the former study only relied on heart rate measurements, used a quadratic fit for their nonlinear heart rate and core temperature model, and used training data from 17 U.S. Army soldiers during field training in 24–36°C humid weather. The authors in the latter study included a mixed population of men and women, with 17 of the 25 volunteers being the U.S. Army soldiers from the former study, and used a sigmoid function for their nonlinear model. Thus, a true direct performance comparison between our study and others is difficult due to the use of different populations for model training. Moreover, even with more complex models and estimation methods, uncontrollable variables in the data collection process will result in similar performance across studies, especially when the behavior of heat stress indicators is greatly affected by the environment and other psychophysiological factors.

VII. Conclusion

The performance of our method can be further improved through the use of higher orders models for heart rate and skin temperature, although these models tend to yield only slightly better performance. This can be achieved by performing informed feature extraction on the predictor signals (e.g., heart rate, ECG, etc.) to create new predictor signals. Furthermore, by identifying key states in these new signals, piecewise linear models amenable to Kalman filter, or Kalman filter variants, can be used to maintain simplicity. Relatedly, the performance of our method can also be improved by carefully studying the effects of other stressors that are characteristic of pilots and drivers, such as dehydration and g-forces. Nonetheless, the core temperature estimation method we proposed in such an unparalleled and complex environment (e.g., closed cockpits with radiating heat and layered body equipment) compared to other open world environments, such as cycling, was able to produce comparable performance when the supplementary physiological features and sensors were added to the Kalman filter. This work showed that the use of a linear Kalman filter is robust to the poor linear regression fits not strictly due to heat stress but rather dynamic and sporadic changes within the physiology due to g-force stresses, and mental workload and fatigue during driving, similar to flight. As the movement for physiological monitoring sensor systems for pilots continually develops [15], this will support more advanced algorithms that require additional physiological sensors and features for improved monitoring of thermal stress and other human performance decrements that occur in flight [22–24].

Appendix

The following appended figures discussed in the results section, provide a graphical view of the error that is quantitatively explained using the RMSE value across all 15 validation sets. This provides a tertiary view to understand the Kalman filter core temperature estimates across all the validation sets, shown in Fig. 5.

References

[1] Schminder, J., Gårdhagen, R., Nilsson, E., Storck, K., and Karlsson, M., "Development of a Cockpit-Pilot Model for Thermal Comfort Optimization During Long-Mission Flight," *AIAA Modeling and Simulation Technologies Conference*, 2016.

[2] Shetty, J., Lawson, C. P., and Shahneh, A. Z., "Simulation for temperature control of a military aircraft cockpit to avoid pilot's thermal stress," *CEAS Aeronautical Journal*, Vol. 6, No. 2, 2015, pp. 319–333.

[3] Frank, A., Belokopytov, M., Shapiro, Y., and Epstein, Y., "The cumulative heat strain index – a novel approach to assess the physiological strain induced by exercise-heat stress," *European Journal of Applied Physiology*, Vol. 84, No. 6, 2001, pp. 527–532.

[4] Stahn, A. C., Werner, A., Opatz, O., Maggioni, M. A., Steinach, M., Ahlefeld, V. W. V., Moore, A., Crucian, B. E., Smith, S. M., Zwart, S. R., Schlabs, T., Mendt, S., Trippel, T., Koralewski, E., Koch, J., Choukèr, A., Reitz, G., Shang, P., Röcker, L., Kirsch, K. A., and Gunga, H.-C., "Increased core body temperature in astronauts during long-duration space missions," *Scientific Reports*, Vol. 7, No. 1, 2017, p. 16180.

[5] Werner, A., and Gunga, H.-C., "Monitoring of Core Body Temperature in Humans," *Stress Challenges and Immunity in Space: From Mechanisms to Monitoring and Preventive Strategies*, edited by A. Choukèr, Springer International Publishing, 2020, pp. 477–498.

Fig. 5 Kalman filter results across all 15 validation sets.

[6] Valk, P., and Veenstra, B., "Military Performance and Health Monitoring in Extreme Environments," *Research and Technology Organisation Meeting Proceedings of NATO*, Sofia, Bulgaria, 2009.

[7] Blanchfield, J. E., Hargroves, M. T., Keith, P. J., Lansing, M. C., Nordin, L. H., Palmer, R. C., Louis, S. E. S., Will, A. J., Scherer, W. T., and Napoli, N. J., "Developing Predictive Athletic Performance Models for Informative Training Regimens," *Systems and Information Engineering Design Symposium (SIEDS)*, 2019, pp. 1–6.

[8] Racinais, S., Moussay, S., Nichols, D., Travers, G., Belfekih, T., Schumacher, Y. O., and Periard, J. D., "Core temperature up to 41.5ºC during the UCI Road Cycling World Championships in the heat," *British Journal of Sports Medicine*, Vol. 53, No. 7, 2019, pp. 426–429.

[9] Pineault, L., and Ferguson, D. P., "The biology of automobile racing, more than just turning left," *The Science of Motorsport*, Routledge, 2018, pp. 1–10.

[10] Wilkinson, D. M., Carter, J. M., Richmond, V. L., Blacker, S. D., and Rayson, M. P., "The Effect of Cool Water Ingestion on Gastrointestinal Pill Temperature," *Medicine & Science in Sports & Exercise*, Vol. 40, No. 3, 2008.

[11] Buller, M. J., Tharion, W. J., Duhamel, C. M., and Yokota, M., "Real-time core body temperature estimation from heart rate for first responders wearing different levels of personal protective equipment," *Ergonomics*, Vol. 58, No. 11, 2015, pp. 1830–1841.

[12] Welles, A. P., Xu, X., Santee, W. R., Looney, D. P., Buller, M. J., Potter, A. W., and Hoyt, R. W., "Estimation of core body temperature from skin temperature, heat flux, and heart rate using a Kalman filter," *Computers in Biology and Medicine*, Vol. 99, 2018, pp. 1–6.

[13] Looney, D. P., Buller, M. J., Gribok, A. V., Leger, J. L., Potter, A. W., Rumpler, W. V., Tharion, W. J., Welles, A. P., Friedl, K. E., and Hoyt, R. W., "Estimating Resting Core Temperature Using Heart Rate," *Journal for the Measurement of Physical Behaviour*, Vol. 1, No. 2, 2018, pp. 79–86.

[14] Eggenberger, P., MacRae, B. A., Kemp, S., Bürgisser, M., Rossi, R. M., and Annaheim, S., "Prediction of Core Body Temperature Based on Skin Temperature, Heat Flux," *Frontiers in Physiology*, Vol. 9, 2018.

[15] Napoli, N., Harrivel, A., and Raz, A., "Improving Physiological Monitoring Sensor Systems for Pilots," *Aerospace America*, Vol. 12, 2020.

[16] Carlson, L. A., Ferguson, D. P., and Kenefick, R. W., "Physiological strain of stock car drivers during competitive racing," *Journal of Thermal Biology*, Vol. 44, 2014, pp. 20–26.

[17] Barthel, S. C., Buckingham, T. M., Haft, C. E., Bechtolsheimer, J. E., Bechtolsheimer, T. A., and Ferguson, D. P., "A Comparison of the Physiological Responses in Professional and Amateur Sports Car Racing Drivers," *Research Quarterly for Exercise and Sport*, Vol. 91, No. 4, 2020, pp. 562–573.

[18] Byrne, C., and Lim, C. L., "The ingestible telemetric body core temperature sensor: a review of validity and exercise applications," *British Journal of Sports Medicine*, Vol. 41, No. 3, 2007, pp. 126–133.

[19] Buller, M. J., Tharion, W. J., Cheuvront, S. N., Montain, S. J., Kenefick, R. W., Castellani, J., Latzka, W. A., Roberts, W. S., Richter, M., Jenkins, O. C., and Hoyt, R. W., "Estimation of human core temperature from sequential heart rate observations," *Physiological Measurement*, Vol. 34, No. 7, 2013, pp. 781–798.

[20] Laxminarayan, S., Rakesh, V., Oyama, T., Kazman, J. B., Yanovich, R., Ketko, I., Epstein, Y., Morrison, S., and Reifman, J., "Individualized estimation of human core body temperature using noninvasive measurements," *Journal of Applied Physiology*, Vol. 124, No. 6, 2018, pp. 1387–1402.

[21] Barthel, S. C., and Ferguson, D. P., "Cockpit Temperature as an Indicator of Thermal Strain in Sports Car Competition," *Medicine and Science in Sports and Exercise*, Vol. 53, No. 2, 2021, pp. 360–366.

[22] Stephens, C., Kennedy, K., Napoli, N., Demas, M., Barnes, L., Crook, B., Williams, R., Last, M. C., and Schutte, P., "Effects on task performance and psychophysiological measures of performance during normobaric hypoxia exposure," *19th International Symposium on Aviation Psychology*, 2017, p. 202.

[23] Napoli, N. J., Adams, S., Harrivel, A. R., Stephens, C., Kennedy, K., Paliwal, M., and Scherer, W., "Exploring cognitive states: Temporal methods for detecting and characterizing physiological fingerprints," *AIAA SciTech Forum*, 2020, p. 1193.

[24] Napoli, N. J., Demas, M., Stephens, C. L., Kennedy, K. D., Harrivel, A. R., Barnes, L. E., and Pope, A. T., "Activation complexity: A cognitive impairment tool for characterizing neuro-isolation," *Scientific Reports*, Vol. 10, No. 1, 2020, pp. 1–20.