CVaR-based Flight Energy Risk Assessment for Multirotor UAVs using a Deep Energy Model

Arnav Choudhry\textsuperscript{1}, Brady Moon\textsuperscript{2}, Jay Patrikar\textsuperscript{2}, Constantine Samaras\textsuperscript{1}, and Sebastian Scherer\textsuperscript{2}

Abstract—Energy management is a critical aspect of risk assessment for Uncrewed Aerial Vehicle (UAV) flights, as a depleted battery during a flight brings almost guaranteed vehicle damage and a high risk of human injuries or property damage. Predicting the amount of energy a flight will consume is challenging as routing, weather, obstacles, and other factors affect the overall consumption. We develop a deep energy model for a UAV that uses Temporal Convolutional Networks to capture the time varying features while incorporating static contextual information. Our energy model is trained on a real-world dataset and does not require segregating flights into regimes. We illustrate an improvement in power predictions by 29\% on test flights when compared to a state-of-the-art analytical method. Using the energy model, we can predict the energy usage for a given trajectory and evaluate the risk of running out of battery during flight. We propose using Conditional Value-at-Risk (CVaR) as a metric for quantifying this risk. We show that CVaR captures the risk associated with worst-case energy consumption on a nominal path by transforming the output distribution of Monte Carlo forward simulations into a risk space. Computing the CVaR on the risk-space distribution provides a metric that can evaluate the overall risk of a flight before take-off. Our energy model and risk evaluation method can improve flight safety and evaluate the coverage area from a proposed takeoff location.

The video and codebase are available at: [Video]\textsuperscript{a} | [Code]\textsuperscript{b}

I. INTRODUCTION

Limited battery capacity has been identified as one of the major hurdles in achieving large-scale commercial deployment of Uncrewed Aerial Vehicles (UAVs) \cite{1}. While there has been recent growth in improving battery energy density \cite{2}, \cite{3}, specific energy \cite{4}, and energy management systems \cite{5}, limited onboard battery capacity will always pose a significant risk to UAV operations. This risk stems from the difficulty in identifying and quantifying various factors that affect UAV energy consumption. The risk is also a function of the inherent uncertainty in the available estimates of these factors when determining the energy consumption. Traditionally, factors like prevailing wind conditions, air density, payload, UAV design and airspeed \cite{6} have been identified as considerably affecting UAV energy consumption. Therefore, a method that quantifies the flight risk based on an estimate of the UAV energy consumption given a certain set of conditions is the focus of this paper. To limit the scope, we specifically focus on multirotor UAV systems but the method is generalizable to any other platform.

While fixed-wing UAV systems benefit from decades-long research in energy models for fixed-wing crewed aircraft, multirotors can leverage some of the research in rotary-wing aircraft but generally lack such a comprehensive research background by the virtue of the recent development and more complicated physical interactions of UAVs.

A recent assessment \cite{7} found that while there are multiple attempts at formulating energy model for UAVs, the model estimates differ by a factor of 3-5. The paper attributes this discrepancy to the numerous assumptions in modeling the underlying dynamics of various UAV configurations. The authors of this work argue that the lack of a standardized empirical dataset makes it difficult to compare individual models. Broadly, the prevalent energy models can be grouped into two major categories \cite{8}. The first are the analytical models that refer to the first-principle based models that rely on established helicopter theories to calculate the energy consumption. The parameter values for these models are either calculated using historical data \cite{9} or by using a regression-based approach with empirical flight tests \cite{10}. One drawback of using analytical models is that they inherently make simplifying assumptions while formulating the equations. Rotor-Rotor interactions, crosswind effects, and...
the effect of higher order motions on the propulsion system are often ignored. On the other hand, end-to-end approaches completely avoid constructing any explicit models but directly use data-driven methods to regress parameters. The regression is often done by fitting predetermined polynomials to empirical data [11], [12]. Within this paradigm, Prasetya et al. [8] have presented an approach with multiple polynomial regressions to handle different maneuvers. While the results look promising, the proposed method is complicated and involves generating the movement labels in the dataset to identify various flight maneuvers. The aforementioned methods also do not explicitly take into account external factors like wind or air density which limits their accuracy. A deep learning based model is more expressive than a simple polynomial model. It has been shown recently [13] that compared to other temporal architectures, TCNs are more flexible, stable and less memory intensive.

While risk assessment for aerial robots operating in real world is a relatively new research domain, techniques previously developed for performing risk assessments in the domain of applied finance [14], [15] can be leveraged to perform similar analysis for robot operations. Majumdar and Pavone [16] argue that the traditional risk metrics in robotics like mean-variance may misrepresent the risk associated with a stochastic process. They propose using more informative metrics like Conditional-Value-at-Risk (CVaR) which has been used in some previous works as an optimization constraint to limit the worst-case risk exposure in robotics applications [17], [18].

In this work, we propose the use of CVaR as a risk metric to quantify the risk associated with worst-case energy consumption. Focusing the analysis on the tail-end of the distribution gives a better idea of the associated worst-case energy use. The energy distribution is obtained by using a Temporal Convolutional Network (TCN) based energy consumption model. The model is trained, validated, and tested on a subset of the energy usage dataset collected using more than 10 hours of real-world flying time [19], [20]. A Monte-Carlo (MC) forward simulation setup is used to capture the stochasticity of the environment and the UAV dynamics model to capture their effects on the energy consumption. A CVaR value is then generated by mapping the stochasticity of the environment and the UAV dynamics model to a risk-space using a user-defined risk profile and battery capacity.

The paper is organized as follows: Section III provides details on the energy model and its results. Section IV provides details on the methods of calculating risk using the energy model and a Monte-Carlo setup. Section V gives two case studies to demonstrate applications of the risk method. Section VI presents the discussions and conclusions.

II. PROBLEM DEFINITION

Let $x_t \in \mathbb{R}^3 \times SO(3)$ represent the state of the UAV in inertial space at time $t$. Let $\xi : [0, t_f] \mapsto \mathbb{R}^3 \times SO(2)$ be a given nominal vehicle trajectory with yaw that the UAV follows using a controller $\pi : (x_t, \xi) \mapsto u_t$ under system dynamics as defined in Equation 1, where $d$ is the additive noise in system dynamics and $\kappa$ is the environmental context related to features like ambient wind.

$$\dot{x}_t = f(x_t, u_t, \kappa) + d$$ (1)

Given battery usable capacity $b$, random variable for energy consumption $e \in E$, where $E \in \mathbb{R}^1_+$ is energy-space, our objective is to provide a pre-flight energy-based risk assessment by calculating a suitable risk metric on a empirically derived risk distribution $p(r), r \in R$ where $R \in \mathbb{R}^1$ is the risk-space.

A. Approach

We address the problem of finding the risk distribution and the corresponding metric by first finding an energy distribution $J = \{e | \kappa, d\}$ and then transforming it into risk-space using a user-based risk profile $G(e, b) : E \mapsto R$. To do this, we first construct a power consumption model $Q(\eta_{t-\tau:t})$, where $\eta_{t-\tau:t}$ is a set of derived features from state history $x_{t-\tau:t}$ and $\kappa$. We then use $N$ MC forward simulations to find the energy distribution of a UAV while flying the trajectory $\xi$, under a pre-defined distribution of environmental noise $\kappa \sim N(\mu_\kappa, \Sigma_\kappa)$ (epistemic uncertainty) and a distribution of system dynamics noise $d \sim N(0, \Sigma_d)$ (aleatoric uncertainty), using Eq. 2.

$$e = \int_0^{t_f} Q(\eta_{t-\tau:t})dt$$ (2)

The MC generates a energy histogram with M bins which is then transformed from energy-space to a risk-space using the risk-profile $G(e, b)$ to approximate $\hat{p}_N(r)$. The risk-probability for $r$ in a given bin $R_i$ is given by Eq. 3.

$$\hat{p}_N(r) = \frac{M}{N} \sum_{i=1}^{N} I(G(e, b) \in R_i)$$ (3)

The risk-distribution can now be analysed using metrics like Conditional Value-at-Risk (CVaR) which the authors argue is a relevant metric in that it effectively captures the risk posed by the tail-end distribution. The CVaR value can then be used as part of a comprehensive evaluation that combines other risk factors such as risk to property and
people on the ground to aid in a go/no-go pre-flight decision. An overview of our approach can be seen in Figure 1.

III. ENERGY MODELING

This section details the construction of the power consumption model $Q$. In order to estimate the power consumed by the UAV, we introduce an end-to-end model that takes time varying features and invariant features, and estimates the power consumption of the UAV at each time step. This is done by framing the problem as a multivariate time-series regression problem.

A. Energy Data Description

We use a publicly available dataset [19], [20] which contains power and state measurements for 209 flights on a DJI Matrice M100 drone. 195 of these flights were performed while following a fixed triangular trajectory (see Figure 2b) with varying altitudes, speeds, and payloads, while the remainder of 10 flights (referred to as random flights) followed a different flight path (see Figure 2c) for unique testing data. We split the flights as 60:20:20 for the train, validation and test sets of flights, while making sure that the 10 random path flights were only in the test set. We used a total of 114 training flights (4 hours and 4 minutes flight time), 38 validation flights (1 hour and 24 minutes flight time) and 37 test flights (1 hour and 23 minutes flight time).

B. Energy Model Architecture

Figure 3 provides an overview of the deep energy model. We construct and split $\eta_i$ as the time varying features $\eta^v_i$ and time invariant features $\eta^i_i$. The time varying features $\eta^v_i = \{v_x, v_z, \alpha_t\}$ are airspeed $v_i$, its components in body-frame $v_x$, and $v_z$, vertical speed, and the angle of attack $\alpha_t$. The time-invariant features $\eta^i = \{\rho, m\}$ are the density of air $\rho$, and the mass of the payload $m$. A vector $\Phi_i$ containing temporal embedding of features is obtained from the temporal encoder.

The contextual information is then added to $\Phi_i$ by concatenating time-invariant features to create an input $\Phi'_i = \Phi_i \oplus \eta^i$, containing both temporal as well as contextual information. $\Phi'_i$ is then fed into a dense layer to estimate the power consumption of the UAV at time $t$. There are two reasons we separated the time-varying and fixed inputs. The first is we noticed that separating the inputs improves the convergence rates, and second is that it makes the model modular.

In order to encode the time-varying information, we chose a TCN-based [21] architecture. In the interest of comparative analysis, we also tested LSTMs [22] for the temporal modeling. For the TCNs, we specifically used the implementation as described in [13]. TCNs have been shown to perform at least at or better than LSTMs in a variety of temporal modeling tasks with added advantages including lower memory requirements, flexible receptive field sizes, and stable gradients.

C. Model Experiments

We perform a hyperparameter search for both TCNs and LSTMs. For TCNs, we vary the number of filters in each causal convolution unit between $\{16, 32, 64, 128\}$, the kernel size between $\{2, 3, 4\}$, the number of residual block stacks between $\{1, 2\}$, and the number of layers between $\{3, 4, 5, 6\}$. We use the Adam optimizer with a learning rate of 0.002 and make use of exponential dilation [13]. We selected the top two models LSTM and TCN models based on their performance on the validation set and evaluate them on the test set. The results of which are in Table I and are further discussed in Section III-E.

*The angle of attack is approximated to pitch angle as the dataset does not include explicit angle of attack information.
TABLE I
ENERGY MODEL RESULTS

| Model   | Random Flights | Test Flights | # params |
|---------|----------------|--------------|----------|
|         | MAPE(%)        | RE(%)        | MAPE(%)  | RE(%)    |
| Liu [10]| 14.05          | 9.16         | 12.70    | 7.41     | 5        |
| b-LSTM  | 13.58          | 10.35        | 11.47    | 7.57     | 84,649   |
| s-LSTM  | 10.38          | 7.38         | 9.06     | 5.12     | 76,073   |
| b-TCN   | 13.36          | 7.56         | 9.13     | 5.52     | 5,225    |
| s-TCN   | 10.36          | 7.38         | 9.06     | 5.12     | 76,073   |

D. Metrics

We use the mean absolute percentage error (MAPE) to measure the performance of the models on the power estimation task. The MAPE for a flight $i$, is defined in Eq. 4, where $t = [0, t_f]$, and $y_t^{(i)}$ and $\hat{y}_t^{(i)}$ are the true and predicted value of power for flight $i$ at time step $t$, respectively.

$$\text{MAPE}_i = \frac{1}{t_f} \sum_{t=0}^{t_f} \left| \frac{y_t^{(i)} - \hat{y}_t^{(i)}}{y_t^{(i)}} \right| \times 100\% \quad (4)$$

For measuring the model performance on the energy estimation task, we first divide the flight into sections, $s \in \{1, \ldots, S_t^{(i)}\}$, based on the yaw angle. We then calculate the energy consumption in each section, $e_s^{(i)}$, and use these to calculate the adjusted relative error for flight $i$, $\text{RE}_i$ in Eq. 5, where $\hat{e}_s^{(i)}$ is the predicted energy for each section.

$$\text{RE}_i = \frac{1}{S_t^{(i)}} \sum_{s=1}^{S_t^{(i)}} \left| \frac{e_s^{(i)} - \hat{e}_s^{(i)}}{e_s^{(i)}} \right| \times 100\% \quad (5)$$

This metric forces the model to accurately predict total energy consumed in all portions of the flight and not have errors from different parts of a flight cancel out.

E. Energy Model Results

Based on their performance on the validation set, we selected the top two LSTM and TCN models. The b-TCN model is the bigger TCN model with 64 filters, kernel size 2, 1 residual block stack, and 5 layers. The s-TCN model is the smaller TCN model with 16 filters, kernel size 2, 1 residual block stack, and 5 layers. Both b-LSTM and s-LSTM, had 3 hidden layers with 0 dropout and were trained using RMSprop optimizer. b-LSTM had 64 hidden units, whereas s-LSTM had 16 hidden units.

From Table I, we see that the TCNs consistently outperform the LSTMs on the power as well as energy estimation tasks. Additionally, note that even though the s-TCN is about 93% smaller than the b-TCN, the performance of both is similar. This shows that the TCN architecture lends itself better to model compression and hence might be better for edge applications by using less memory and faster inference. However, it should be noted that the s-TCN takes considerably more number of epochs to train. Figure 4 shows time-series plots of the best and worst flights for the b-TCN. We see that the model is able to follow the peaks and troughs of power consumption.

IV. RISK MODELING

A. Wind Modeling

An important factor in predicting energy consumption for a UAV is accounting for the effects of the wind field. In order to accurately simulate the energy usage of a given trajectory, an accurate wind model must be included in the environmental context $\kappa$ for the system dynamics of Equation 1. For this work, wind is broken down into two components: constant wind and turbulence.

The constant wind is the averaged wind field which is temporally static but spatially varying for a given environment. Given an accurate model of the environment, the current wind field can be found using the wind inlet angle $\Theta$ and magnitude $U$ by using CFD. CFD has been shown to be an effective method in calculating the low altitude urban wind flow patterns for UAV flight [23]–[26]. However, wind fields are not temporally constant because the inlet conditions are not steady and therefore need to be defined as a distribution. The mean and standard deviations of the inlet angle and magnitude are given as $\mu_\kappa = \{\mu_\Theta, \mu_U\}$ and $\Sigma_\kappa = \text{diag}(\sigma_\Theta, \sigma_U)$ respectively and can come from historical data spanning over years or even an hour depending on the application.

In addition to the constant wind, turbulence is added using the Dryden turbulence model [27]. Previous works [28]–[31] have shown that this model can be adapted for simulation of wind turbulence for small UAV.

B. Monte Carlo Simulations

Given $\mu_\kappa$, $\Sigma_\kappa$, and $\Sigma_\xi$, we run $N$ MC simulations to create a set of possible vehicle trajectories given the distribution of conditions. The simulated vehicle follows the trajectory $\xi$ with the controller $\pi$ and the dynamics defined in the previous section.

C. Risk Modeling

As defined in Section II, $G(e, b)$ represents the user-defined operation risk profile. In this work, as a test case, $G(e, b)$ can be chosen as in Eq. 6.

$$G(e, b) = \exp \left( \frac{\gamma}{(\max(b - e, \lambda))} \right) - 1 \quad (6)$$
where $\gamma, \lambda > 0$ are constants. Such a choice of $G(e, b)$ takes into account the fact that as the energy consumption nears the battery capacity, the risk increases exponentially until it reaches a limit value.

The MC simulations defined in the previous section can be used to generate the distribution for $e$ which in turn can be used to generate the risk probability density function $p(r)$ by normalizing the output of $G(e, b)$. From here on we follow the formulation proposed by Rockafellar et al. [15] to define Conditional Value-at-Risk (CVaR).

The Conditional Value-at-Risk at level $\nu$ or $\text{CVaR}_\nu$ gives the expected risk of the worst $(1 - \nu)$ quantile of the distribution $G(e, b)$. Intuitively $\text{CVaR}_\nu$ gives us an idea on the behaviour of the worst $(1 - \nu)$ quantile cases by providing the expected risk in the cases that are “beyond” Value-at-Risk at level $\nu$ ($\text{VaR}_\nu$) risk.

$$\text{CVaR}_\nu = \frac{1}{1 - \nu} \int_{G(e, b) \geq \text{VaR}_\nu} G(e, b)p(r)dr \quad (7)$$

While the mathematical superiority of CVaR as a coherent risk metric [16] is useful in formulating convex CVaR optimization problems [15], it also helps to intuitively see how $\text{VaR}_\nu$ and $\text{CVaR}_\nu$ help provide information about a particular risk distribution profile. Figure 5 shows a deliberately constructed example that highlights how the different metrics—mean, $\text{VaR}_\nu$, and $\text{CVaR}_\nu$—capture the risk profile. The three figures show three different energy MC outputs in red. These outputs are then transformed into risk space by using equation 6 with coefficients $\gamma = 6.93, \lambda = 10, b = 40$ J. This in turn gives us $p(r)$ or the risk $p.d.f$ in blue. $\text{VaR}_\nu$ and $\text{CVaR}_\nu$ values are calculated at $\nu = 0.95$. As can be seen, while intuitively it is clear that moving to the right of the three examples, the MC energy output relates to a higher risk flight because of a longer tail, both the mean and $\text{VaR}_\nu$ are not able to capture this behaviour.

### V. CASE STUDIES

In order to illustrate potential applications of the presented work, two case studies will be given. The first case study shows how the energy risk for a proposed flight trajectory would be evaluated and could be used before the fly or no-fly command is given. The second case study is an evaluation of all potential flight paths from a given takeoff location to evaluate the coverage area for a proposed takeoff location.

#### A. Case 1:

Before a UAV takes off to carry out a nominal path, it would be prudent to evaluate the risk of depleting the battery during flight. Using our proposed method of evaluating risk, the pilot in command could calculate the distribution of necessary energy to carry out the flight path and its $\text{CVaR}_\nu$. The environmental conditions for the MC simulations would model the conditions during the proposed flight window.

To demonstrate this case, real environmental data is used from a test site with around ten buildings ranging in heights of 5 to 35 meters. This area is similar to a typical urban environment with non-uniform wind fields. The area is around 300 x 250 meters. The wind field model was calculated using CFD, and the wind inlet conditions were captured using a weather station on the tallest building in the area. The window of time for this test was February 19, 2020 from 4:20-4:22 PM. The inlet conditions were $\mu_x = \{-2.53^\circ, 3.14 m/s\}$, $\Sigma_x = \text{diag}(28.47^\circ, 1.55 m/s)$

Figure 6 shows the 1000 MC simulations and the predicted amount of energy for the trajectory. The mean energy consumption was 34.86 kJ. With $b = 369.36$ kJ, $\lambda = 92.34$, $\gamma = 64$, and $\nu = .95$, the $\text{CVaR}_\nu$ value is 0.2117. This value reflects how risky this proposed flight path is and would be compared against the allowable risk determined by the operator or company.

#### B. Case 2:

In many applications of a UAV, the vehicle will take off and return to the same location every time. In these cases it could be useful to evaluate the coverage area for that particular location. This knowledge could influence the placement of launching location and optimize the number of base stations necessary to ensure full coverage of a region.

The testing environment for this case is a 3300 x 2130 meter area of a major metropolitan city. This city has a higher density of buildings compared to Case 1. The wind field model for this city came from a high resolution CFD
Fig. 6. Monte Carlo simulations for Case 1. Each flight is represented with a pink line and is shaded according to the total energy consumption. The distribution of the total energy consumption for the flights is shown in the histogram.

Fig. 7. Contour plot of the CVaR_\eta values for Case 2 with the city map overlay.

Fig. 8. Contour plot of the CVaR_\nu for Case 2 when the usable battery capacity is low. The risk values increase quickly as the UAV travels away from its takeoff location.

VI. Conclusions

This work presents a novel, open-source, state-of-the-art, deep-learning based energy model trained on a publicly available dataset to quantify the energy risk for multirotor UAVs. MC forward simulations are used to find a energy distribution which is transformed into a risk-space. CVaR is used to conduct a risk assessment on the risk-space distribution. Case studies using real world data highlight how the proposed framework can be applied.

While the proposed deep energy model is generalizable across various multirotor configurations, we only show results using a quadrotor model. The results for the energy model are also only tested in typical delivery-style missions and do not account for aggressive, evasive, or acrobatic maneuvers. Even with these limitations, we demonstrate that CVaR-based risk assessment can be used for energy consumption analysis of multirotor systems.

ACKNOWLEDGMENT

This work is supported by the U.S. Department of Energy (Grant DE-EE0008463). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE1745016.

The authors would like to thank Bastian Wagner for his help with the implementation of the analytical model and Cherie Ho for her help with the MC simulations.
