Using Large Language Models to Automate Flight Planning under Wind Hazards

Amin Tabrizian[1], Pranav Gupta[2], Abenezer Taye[1], James Jones[3], Ellis Thompson[1],
Shulu Chen[1], Timothy Bonin[3], Derek Eberle[3], and Peng Wei[1]
[1] George Washington University, Washington, DC, USA
[2] Duke University, Durham, NC, USA
[3] MIT Lincoln Laboratory, Lexington, MA, USA

***Abstract*—We introduce a novel framework to automate flight planning processes using large language models (LLMs) to identify flight operator's preferences. Our framework integrates the recent advancements in LLM and prompt engineering, low-altitude wind hazard forecasts, flight mission energy estimation, and pre-departure strategic deconfliction. First, our approach begins with the forecast of wind hazard polygons to ensure safety in flight planning. Second, we generate a diverse set of candidate flight plans to avoid these wind hazard polygons. The flight plan features include total flight distance, cruising altitude, flight mission energy consumption, and number of waypoints. Third, human flight operator specifies their preferences through natural language prompts or plain words, which are fed to the LLM to extract and prioritize these features. Our framework then evaluates and scores each flight plan based on extracted user-defined preferences, recommending the flight plan that best matches the flight operator's needs. For the purpose of demonstration, we focus on a flight planning use case for an electric vehicle take-off and landing (eVTOL) aircraft in an advanced air mobility (AAM) mission in Dallas-Fort Worth area. Our simulation experiments show the effectiveness of this approach in generating personalized, safe, and efficient flight plans. To our best knowledge, this work is among the first attempts using LLMs to enable a human-centric flight planning automation for AAM operations.**

***Index Terms*—Flight planning, Large Language Model, Wind hazard forecast, Advanced air mobility**

I. Introduction

Safe and efficient flight planning is essential in air operations to ensure the safety of airspace users. Both the European Aviation Safety Authority (EASA) and the Federal Aviation Administration (FAA) forecast a significant rise in small unmanned aircraft system (sUAS) operations and electric vertical takeoff and landing (eVTOL) aircraft operations in the near future, respectively [1], [2]. With the anticipated high-density and highly-automated operations in sUAS package delivery and eVTOL passenger transportation, an automated and scalable flight planning system is much needed.

We propose a decision-support solution for flight planning automation powered by large language models (LLMs) to meet this end. Our approach keeps the human operator in the loop by identifying their flight mission preferences from their interactions with the LLM. The proposed concept minimizes the human operator workload by integrating mission planning information such as wind hazard forecasts, flight route generation, flight mission energy consumption estimation, and LLM. Human workload is further reduced by only providing natural language inputs when interacting with the automation.

In this paper, we focus on eVTOL flight planning in advanced air mobility (AAM) operations. We propose a framework that utilizes our customized A* algorithm for flight route generation avoiding wind hazards. We generate a diverse set of candidate flight plans with different features such as cruising altitudes, total distance, flight mission energy consumption, and number of waypoints. These flight plan features F, flight operation context S, and human operator's prompts P are then fed to the LLM to identify the human's preference of the flight mission. The LLM parses the operator prompt and outputs the preference weights. The preference weights are used to compute and rank the score function of each candidate flight plan. The flight plan with the highest score is selected to recommend to the human operator for their review, modification or confirmation.

This paper makes the following key contributions:

- We leverage LLMs to build a human-centric flight planning automation framework against wind hazards, making our team one of the first to use LLMs for personalized flight planning optimization in aviation operations.
- Our flight planning automation integrates our recent advancements in wind hazard forecasts, flight route generation, and flight mission energy estimation, enhancing the safety and efficiency of eVTOL flight planning.
- We set up simulation experiments with practical airspace data, wind forecast results, and realistic aircraft performance data in the Dallas-Fort Worth area, evaluating the system's performance under various wind hazard scenarios and different human operator prompts. We provide a comprehensive analysis of the recommended routes, highlighting the system's capability and limitations.

The remainder of this paper is organized as follows: Section II provides a review of the background and related work. In section III, we describe the wind hazards forecasting model. Section IV discusses the flight route generation process with our customized A* algorithm. Section V presents steps for estimating flight mission energy consumption for each generated flight plan. In section VI, we introduce a framework for automating the flight planning process with LLM. Simulation experiments and demonstrations are presented in section VII. Section VIII concludes the paper.

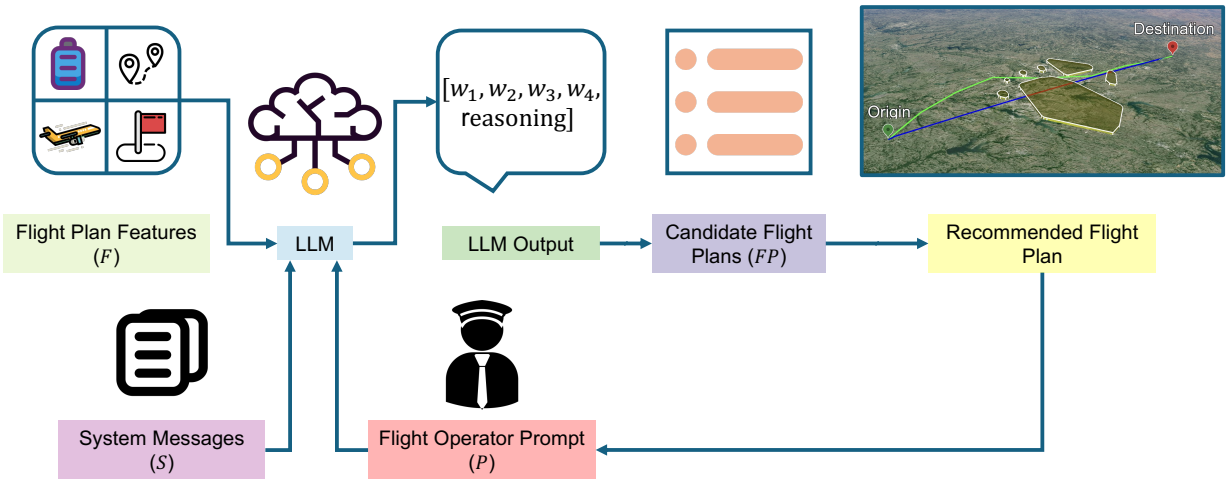

Figure 1: The proposed automation process of using a LLM for flight planning. The LLM interprets the flight operator prompt in natural language and assigns weights to the flight plan features such as total flight distance, cruising altitude, flight mission energy consumption, and number of waypoints. In addition, to better gain human operator's trust, the LLM provides reasoning and logic for its decision. The preference weights w_1, w_2, w_3, w_4 are used to compute and rank the scores of all candidate flight plans. The flight plan with the highest score is recommended to the flight operator for their review, modification and confirmation.

II. Background and Related Work

A. Motion Planning and Flight Route Generation

There is a rich body of literature for motion planning and route generation for robots, ground vehicles and aircraft. Rapidly-exploring Random Tree (RRT) is one of the most popular algorithms that has been extended to many variants. Our approach enhances the finite node concept presented by [3] to improve computational efficiency and memory usage. The rope pull method described in [4] optimizes the generated routes by eliminating unnecessary turns or waypoints through simulating a rope pull. Essentially, routes generated via RRT and RRT* [5] can be optimized by directly connecting points with a line of sight without intersecting obstacles. Similar to RRT methods, Probabilistic Road Map (PRM) [6] samples randomly from the configuration space. Unlike RRT, PRM constructs a practical graph for navigating a region, followed by a query phase using a path-finding algorithm to determine a valid route. Due to the random nature of node addition, RRT and PRM families are efficient and tractable, but they cannot guarantee the optimal solution. For dynamic environments where obstacles may move or change, D* and D*-Lite [7], [8] offer efficient solutions for these partially known environments. The A* search algorithm [9] is a path search and graph traversal algorithm with a worst-case complexity of $O(b^d)$, where b is the branching factor and d is the depth of the desired goal. It prioritizes exploring nodes that minimize $f(n) = g(n) + h(n)$, where $h(n)$ is the heuristic value which is commonly the direct distance from the current node Γ_n to the goal node Γ_{goal} and $g(n)$ is the shortest distance from the start node Γ_{start} to $\Gamma(n)$. Here n represents the index of a node.

B. Estimating Flight Mission Energy Consumption

The current body of literature exploring the power consumption of eVTOLs is limited; nevertheless, some prior works have addressed this issue in different contexts. Notably, [10] proposed a power consumption model for hybrid eVTOL aircraft covering the hover, climb, cruise, and descent phases of flight. Likewise, [11] developed a power consumption model for fixed-wing eVTOL aircraft. However, neither model accounts for detailed factors such as aircraft dynamics, wind conditions, and flight plans. Therefore, in this paper, we adopt a more realistic power consumption model developed by [12], which considers these crucial flight and aircraft characteristics.

Moreover, the power consumption model proposed by [12] is widely used in the AAM domain. For instance, [13] utilized this model to estimate charging demands at vertiports. Similarly, [14] applied the model to conduct hardware-in-the-loop battery prognostics for AAM operations. Additionally, [15] employed the model to perform energy-aware traffic management for AAM. In this paper we will adopt the methods in [12], [13] in our proposed flight planning framework. A detailed review of this power consumption model is presented in Section V.

C. Large Language Models

LLMs are transformer-based language models that typically have enormous parameters, numbering in the hundreds of billions. These models are capable of text generation and comprehension [16], which allows them to perform complex tasks. OpenAI has reached significant milestones with the development of GPT-3.5 and GPT-4 [17], showcasing substantial improvements in natural language understanding and generation. LLMs exhibit remarkable human-like capabilities and demonstrate strong proficiency in robotic manipulation [18], multi-modal understanding [19], and autonomous driving

[20]–[22]. In [20], a framework was introduced that leverages LLMs, combining a Reasoning and a Reflection module to enable the system to make decisions based on common-sense knowledge and continuously evolve. The integration of LLMs to mimic human-like decision-making in autonomous driving was explored in [21], enhancing situational understanding and interaction with human drivers. An LLM-based framework capable of performing close-loop autonomous driving in realistic simulators was proposed in [22]. According to our best knowledge, there is little research on using LLMs for flight planning in the aviation domain.

III. Wind Hazards Forecasting Model

A deterministic forecast was generated with Version 4.2 of the National Center for Atmospheric Research's (NCAR's) open-sourced Weather Research and Forecasting (WRF) model [23] to provide high spatial-resolution gridded wind information. The model simulates the weather conditions by assuming initial and boundary conditions and solving a set of partial differential equations. WRF produces a range of weather data, including temperature, wind, precipitation, humidity, and many other products over a full 3-dimensional volumetric grid ranging out to an assigned forecast horizon. An adapted version of WRF was developed by MIT Lincoln Laboratory run on the Lincoln Laboratory Supercomputer Center (LLSC) [24], a high-performance computing cluster, to accelerate the generation of high-resolution forecasts.

A 100 m grid spacing was used to capture abrupt changes in wind conditions. By projecting at this granularity, the WRF forecasts are able to model a significant variability within the wind field due to turbulence within the boundary layer that may not be accounted for by higher-resolution forecasts such as the High-Resolution Rapid Refresh Model (HRRR). Kernel density estimation was applied to assess the probability that the wind speed exceeded a certain value based on the information in the forecast. A thresholding process was then used to identify regions of airspace where such conditions occurred. The Density-Based Spectral Clustering of Applications and Noise (DBSCAN) algorithm was then used to consolidate adjacent polygon regions within close proximity of one another [25].

IV. Flight Route Generation

For flight route generation, we customized the A* algorithm, which is commonly used in route planning tasks to achieve a near optimum. Other path planning or trajectory planning algorithms are also compatible and can be used in our framework.

A. Point Expansion

The environment where our A* algorithm is implemented is not restricted to a grid or discrete map. We traded off the enhanced efficiency of a pre-defined grid of nodes for a given airspace region, which reduces the overall memory usage of the search.

We constructed a graph during the node expansion phase, with introduced nodes adhering to specific rules. To minimize sharp-angled turns in our flight plan, nodes are generated only within a finite arc defined by a radius r and an angle α, with the center of the arc extending the line $\Gamma_n \rightarrow \Gamma_{n+1}$ as depicted in Figure 2. To limit the search space, nodes inserted within a certain distance of already explored nodes are considered the same. In such cases, the $g(n)$ values are compared, and if the inserted node has a lower $g(n)$ value, the explored node's value is updated and re-expanded to ensure route optimality.

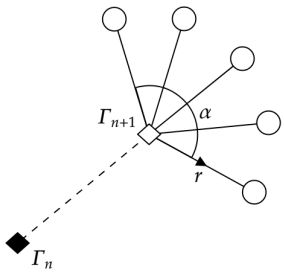

Figure 2: New nodes are iteratively created along the arc subtended by the angle α, with the center of the arc extending the path from the current node Γ_n to the next node Γ_{n+1}. Here n represents the index of a node.

For any previously explored points, the heuristic $h(n)$ needs to be calculated only once. As described in Equation 2, the heuristic function does not consider any information about previously visited nodes. Since the heuristic function is the most computationally demanding aspect of this approach, calculating it just once significantly improves computation speed.

During the expansion phase, nodes are classified as either valid or invalid. A valid node exists within the airspace region and has a valid connection to its parent node. A valid connection is one that does not violate airspace constraints or intersect with any obstacles. In this paper, the obstacles are the wind hazard polygons. Conversely, a node with an invalid connection to its parent node that intersects with airspace constraints or a wind hazard polygon. To clarify, an invalid node is any node located within or too close to a wind hazard polygon or outside the airspace region or map boundaries. Nodes with invalid connections are handled differently; instead of being excluded from expansion, they are immediately added to the closed list. This approach allows for pre-calculation and storage of the heuristic, reducing the number of calculations needed if the node is reached again via a safe route.

B. Heuristic Function

The heuristic function $h(n)$ employed in our customized A* algorithm modifies the technique suggested in [26]. This approach initially determines the angle from the node (Γ_n) to all vertices of the wind hazard polygon Z_j, as shown in Equation 1. The forecasted wind hazard polygons could be concave, posing a challenge for the route-generation algorithm.

$$\Theta(\Gamma_n, Z_j) = \{\theta(\Gamma_n, Z_j^i) - \theta(\Gamma_n, \Gamma_{\text{goal}}) \,|\, \forall i \in Z_j\} \tag{1}$$

This results in a positive value if the point is to the right of the

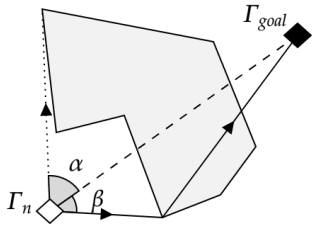

Figure 3: The heuristic for a point with an obstructed line of sight due to a non-convex wind hazard polygon involves calculating the distance to the nearest extreme point on the obstructing polygon with the smallest angle, followed by the distance to the goal. This path is denoted by β, while the path α is not used since the distance $\theta_\alpha > \theta_\beta$.

direct line $\Gamma_n \rightarrow \Gamma_{\text{goal}}$, and a negative value otherwise. From this set, the extremes are determined, identifying the points furthest left and right of the line $\Gamma_n \rightarrow \Gamma_{\text{goal}}$. This is achieved by finding the minimum and maximum values in $\Theta(\Gamma_n, Z_j)$. The heuristic, shown in Equation 2 and illustrated in Figure 3, is the distance between Γ_n and the extreme vertex Z_j^*, whose absolute angle is the smallest from the line $\Gamma_n \rightarrow \Gamma_{\text{goal}}$, plus the distance from Z_j^* to Γ_{goal}. This value is further multiplied by ω, where $1.1 \leq \omega \leq 1.5$, which enhances the A* algorithm's performance by prioritizing exploration of points closer to the goal.

$$h(n) = \omega \left(d(\Gamma_n, Z_j^*) + d(Z_j^*, \Gamma_{\text{goal}}) \right) \tag{2}$$

V. Estimating Flight Mission Energy Consumption

In this section, we provide a review of background knowledge and formulation of the energy consumption model for an eVTOL aircraft. We first introduce the point mass guidance model of eVTOL aircraft as presented in [12]. Following this, we describe the energy consumption model used to estimate the total energy required for the flight mission.

A. Aircraft Model

The eVTOL aircraft lateral flight dynamics model can be shortly described as:

$$\dot{x}_t = \xi(x_t, u_t) \tag{3}$$

In this model, $x_t \in \mathbb{R}^2$ and $u_t \in \mathbb{R}^3$ represent the state and control input of the aircraft at time t, respectively. This model can capture the cruise phase of the flight mission. The states of the model are $x = [\lambda, \tau, V, \psi]$, where λ is the latitude, τ is the longitude, V is the true airspeed, and ψ is the heading angle w.r.t. north. The control inputs of the model are the net thrust (T), the rotor tip-path-plane pitch angle (θ), and the rotor tip-path-plane roll (bank) angle (ϕ).

$$\frac{dV}{dt} = \frac{T\cos\phi\sin\theta - D}{m} \tag{4}$$

$$V\frac{d\psi}{dt} = \frac{T\sin\phi}{m} \tag{5}$$

$$(R_{\text{earth}} + h)\frac{d\lambda}{dt} = V\cos\psi \tag{6}$$

$$(R_{\text{earth}} + h)\cos\lambda\frac{d\tau}{dt} = V\sin\psi \tag{7}$$

The multi-rotor aircraft is assumed to have four rotors [27] and the net thrust (T) generated by the aircraft is a sum of thrusts generated by each motor (T_{rotor}). Therefore, $T = 4T_{\text{rotor}}$, and parasite drag (D) on the multi-rotor eVTOL is calculated as $D = 1.1984\frac{\rho V^2}{2}$, where ρ is the density of air.

B. Power Consumption Model

The power consumption model we use to compute the mission-level energy requirement is based on momentum theory for rotorcraft, adopted from [12] and [28]. Below, we summarize the main components of this model. Assuming quasi-steady flight, the instantaneous power required for forward cruise flight at a constant altitude is the sum of induced power, parasite power, and profile power:

$$P_{\text{required}} = P_{\text{induced}} + P_{\text{parasite}} + P_{\text{profile}} \tag{8}$$

The induced power (P_{induced}) is the total of the induced power losses from each rotor $(P_{\text{induced rotor}})$. Therefore, P_{induced} is:

$$P_{\text{induced}} = \sum_{n=1}^{4}(P_{\text{induced rotor}})_n = \kappa\sum_{n=1}^{4}(T_{\text{rotor}})_n(\nu_i)_n, \tag{9}$$

where κ is the induced power correction factor and ν_i is the induced velocity. The parasite power loss (P_{parasite}), which represents the power required to propel the aircraft forward at a constant altitude, is given by:

$$P_{\text{parasite}} = TV\sin\alpha, \tag{10}$$

where α is the angle of attack between the air stream and the rotor disk (tip-path-plane). The profile power loss (P_{profile}) is given as:

$$P_{\text{profile}} = \frac{\rho A_{\text{rotor}}(\Omega R)^3\sigma C_{\text{d mean}}F_P}{8}, \tag{11}$$

where Ω is the rotational velocity of the rotor blades, σ is the thrust-to-weight solidity ratio, $C_{\text{d mean}}$ is the mean blade drag coefficient, and F_P is a function accounting for the increase in blade section velocity due to rotor edgewise and axial speed.

Finally, we can write the instantaneous power required P_{required} in forward flight as:

$$P_{\text{required}} = \kappa\sum_{n=1}^{4}(T_{\text{rotor}})_n(v_i)_n + \frac{\rho A_{\text{rotor}}(\Omega R)^3\sigma C_{\text{d mean}}F_P}{8} + TV\sin\alpha \tag{12}$$

Table I: Performance data of the eVTOL aircraft [27]

Parameter (Unit)	Symbol	Value
Rotor radius (m)	R	4.0
Rotor aria (m^2)	A_{rotor}	50.26
Mass (kg)	m	$2,940$
Thrust-to-weight solidity ratio	σ	0.055
Mean blade coefficient	$C_{\text{d mean}}$	0.0089
Blade section velocity increase function	F_P	0.97
Correction factor	κ	1.75
Rotational velocity (rad/sec)	Ω	30.12
Maximum power (kw)	P_{max}	494.25

C. Assumptions and Constraints

In this paper, we consider the following two mission requirements and path constraints for each flight mission. First, the net vertical force on the eVTOL aircraft is zero, as we focus on the cruise phase.

$$T\cos\phi\cos\theta = mg \tag{13}$$

Second, the trajectories that the aircraft follow to fly between waypoints are governed by the great-circle trajectory:

$$\begin{aligned}&(V\sin\psi)(\sin\lambda_2\cos\lambda_1 - \sin\lambda_1\cos\lambda_2\cos(\tau_2-\tau_1))\\ &-\ (V\cos\psi)(\sin(\tau_2-\tau_1)\cos\lambda_2) = 0\end{aligned} \tag{14}$$

D. Energy-optimal Guidance Law Design

Having established the power consumption model and performance requirements, we now focus on determining the guidance law (u_t) that lead to minimum energy consumption for a flight mission. To achieve this, we formulate the following optimal control problem, aiming to minimize mission-level energy consumption while incorporating the aircraft dynamic model as constraints.

$$\min_{\forall t\in[0,T_f]} \int_0^{T_f} (P_{\text{required}})dt \tag{15}$$

subject to

$$\begin{cases}\dot{x}_t = \xi(x_t, u_t),\\ T\cos\phi\cos\theta - mg = 0,\\ (V\sin\psi)(\sin\lambda_2\cos\lambda_1 - \sin\lambda_1\cos\lambda_2\cos(\tau_2-\tau_1))\\ \qquad -\ (V\cos\psi)(\sin(\tau_2-\tau_1)\cos\lambda_2) = 0\end{cases} \tag{16}$$

Solving the optimal control problem described above yields the optimal energy consumption at a fixed altitude and cruise speed. Therefore, to estimate the optimal energy consumption for the generated candidate flight plans from Section IV, this optimal control problem must be solved for each candidate flight plan by varying cruise speeds. In this paper, we utilized an open-source C++ library called PSOPT [29] to numerically solve the nonlinear programming problem using a sparse nonlinear programming solver.

The resulted minimum energy for each generated candidate flight plan will be listed as the estimated total energy consumption in Table I. More detailed flight mission energy consumption model will be explored in our future work by adding the energy consumption of climbing phase and descending phase of the flight.

VI. Automating the Flight Planning Process with Large Language Model (LLM)

In this section we describe the framework for using the LLM to automate flight planning under wind hazards. A diverse set of candidate flight plans is generated (see Section IV) given the forecasted wind hazard polygons (see Section III). These flight plans are characterized by various features, including total flight distance (*D*), cruising altitude (*A*), flight mission energy consumption (*E*), and number of waypoints (*W*). Each of these features can significantly impact the flight plan's suitability depending on the flight operator's specific preferences.

To capture the flight operator preferences, our framework (see Figure 1) allows the operator to input prompts specifying their desired features in a flight plan. These prompts can be expressed in natural language, such as, "I prefer a flight plan with the lowest energy consumption," or "The route should avoid high altitudes." To ensure the LLM behaves consistently and aligns with specific user needs, we use system messages S to provide the flight operation context to the LLM, regulate the behavior of the LLM assistant, and provide specific instructions, enhancing the overall user interaction experience [30]. These instructions will help the LLM assistant to better understand its role, what its task is with a given prompt, and how it should generate a response as a flight operator assistant. For instance here is an example of a system message: "You are an expert on determining the best flight plan flying from origin to destination. Given an input question, determine the weights for these features of a flight plan."

The LLM, in this work OpenAI's GPT-4o interprets the flight operator prompt and turns it into a structured query. Let P represent the operator prompt or human input in natural language, $\text{LLM}(P, F, S)$ the model output of the prompt, and F the set of flight plan features D, A, E, W. The LLM model processes the input to prioritize the features according to the operator's preferences:

$$[w_1, w_2, w_3, w_4, \text{reasoning}] = \text{LLM}(P, F, S) \tag{17}$$

Subsequently, our framework evaluates the candidate flight plans based on the identified flight operator preferences. Each candidate flight plan FP_i is scored based on how well it aligns with the operator's specified preference. The scoring function Ψ is formulated as:

$$\Psi(FP_i) = -(w_1\tilde{D}_i + w_2\tilde{A}_i + w_3\tilde{E}_i + w_4\tilde{W}_i), \tag{18}$$

where w_1, w_2, w_3, and w_4 are the preference weights obtained from the LLM's interpretation of the flight operator prompt, reflecting the importance of each feature. These weights should add up to one. $\tilde{X}$ represents the standardized feature X by removing the mean and scaling to unit variance.

The flight plan with the highest score, indicating the best match to the flight operator's preferences, is then suggested and displayed to the operator. This recommendation system leverages the powerful natural language understanding capabilities of the LLM to ensure that the suggested flight plan aligns closely with operator preferences.

VII. Simulation Results and Demonstration

We focus on eVTOL aircraft operation and flight planning in AAM use case. The proposed framework utilizes practical wind hazard forecasts from MIT Lincoln Laboratory in Dallas-Fort Worth area. The wind hazard polygons were at four different altitudes $(1000, 1500, 2500,$ and 3000 ft). The origin vertiport and destination vertiport are colored in red and green respectively, as illustrated in Figure 4. We varied the distance limit parameter between each pair of waypoints in our customized A* algorithm described in Section IV, enabling the generation of flight plans with different numbers of waypoints (W) and total distance (D). We also ran the flight route generation algorithm at different altitudes (A) to provide more diverse candidates. Each candidate route was provided to the procedure described in Section V to estimate the total energy consumption (E).

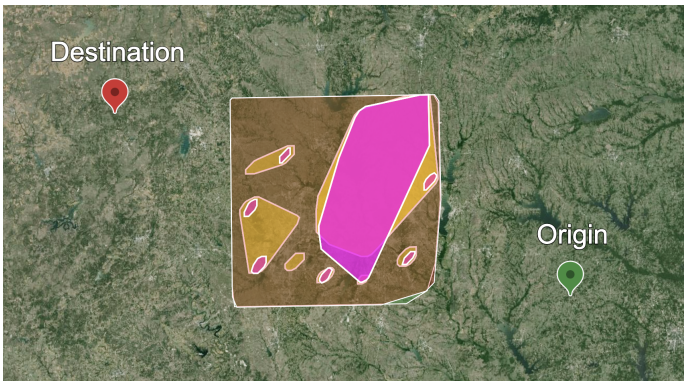

Figure 4: Demonstration of wind hazard polygons: a flight planning example in Dallas-Fort Worth airspace showing four different wind hazard polygons at various altitudes: 1000 ft (green), 1500 ft (yellow), 2500 ft (red), and 3000 ft (magenta). Brown area is the overlap of 1000 and 2500 ft altitudes.

All experiments were conducted on a MacBook Pro equipped with an Intel 2GHz i5 processor and 32GB of memory. The GPT-4o API was used to obtain preference weights based on the flight operator prompts.

A summary of the generated candidate flight plans is presented in Table II. An example of a generated candidate flight plan is illustrated in Figure 5. This table provides an overview of the candidate flight plans generated under different conditions, highlighting the flexibility and adaptability of the system in handling various wind hazard scenarios and different flight operator needs.

Additionally, Table III showcases examples of flight operator interactions with the LLM and the preference weights. These

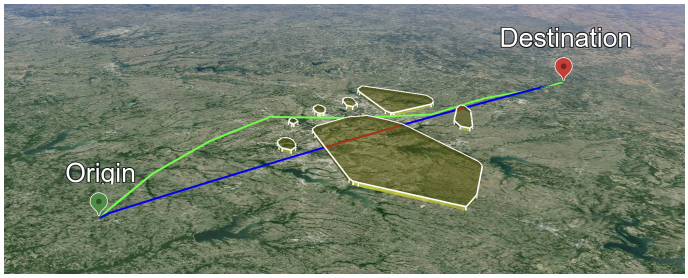

Figure 5: Demonstration of flight route generation: a flight planning example given a set of wind hazard polygons at 1500 ft. The blue route is the point-to-point flight plan overlapping with the wind hazard polygon, where the intersection segment is highlighted in red. The route is a generated candidate flight plan provided by the customized A* algorithm.

examples illustrate how flight operator inputs were understood and parsed into specific weights, which in turn influenced the flight plan sorting and selection process. In all of the examples, the LLM was able to successfully assign proper weights based on flight operator preferences. In Table III we avoided adding the reasoning part from the LLM for better demonstration purposes. But they are all part of LLM's output and available to show based on human flight operator's need. For instance, the reasoning for the weights assignment of the prompt "I want to minimize fuel use and have a moderately smooth flight." is "The user prioritizes minimizing fuel use, so energy consumption (w_3) has the highest weight. A moderately smooth flight is related to maintaining an optimal altitude, giving it a moderate weight (w_2). Distance and number of waypoints are not mentioned, so their weights are set to zero."

Here are some remarks and observations from our prompt engineering with the LLM: Since flight plan 7 has the lowest distance and energy consumption it has been recommended in all of the flight operator's prompts that are related to either of these features. There are some words that convey an ambiguous meaning even for a human interpreter. In those cases, some of the assigned weights might not be reproducible. For instance, it is not clear if the term "smooth flight" is related to having a smaller number of waypoints or flying at lower altitudes. In this case, for some prompts, the LLM assigned a high weight to altitude and to the number of waypoints for some others.

Limitations: Although the flight operator preference is completely satisfied through the resulting preference weights, the balance between the weights is not consistent at each run. This could be improved by setting a fixed random seed in order to have more consistent answers from the LLM.

VIII. Conclusion

In this paper, we presented an innovative approach for flight planning automation that integrates wind hazard forecasts, flight route generation, flight mission energy consumption estimation, and LLM. By predicting wind hazard polygons, we ensured the safety and efficiency of generated flight plans. The generated flight plans were evaluated based on their features and then

Table II: Candidate flight plans with their features

Flight Plan Index	Total Distance (km)	Cruising Altitude (ft)	Energy Consumption (MJ)	No. of Waypoints
1	251.96	1000	1108.57	10
2	279.60	1000	1230.18	11
3	280.27	1000	1233.11	13
4	277.69	1000	1221.78	15
5	261.10	1500	1148.77	9
6	249.01	1500	1095.56	10
7	246.48	1500	1084.43	11
8	248.16	1500	1091.84	14
9	280.99	2500	1236.27	10
10	278.97	2500	1227.38	11
11	278.07	2500	1223.45	13
12	277.01	2500	1218.78	15
13	252.77	3000	1112.14	9
14	251.96	3000	1108.57	10
15	251.84	3000	1108.04	12
16	250.46	3000	1101.97	14

(D) Total Distance, (A) Cruising Altitude, (E) Energy Consumption, and (W) No. of Waypoints

Table III: Examples of flight operator prompts, resulted preference weights obtained by the LLM, and the selected flight plan (FP) with highest score

Explicit Prompts					
Prompt from the flight operator:	$w_1(D)$	$w_2(A)$	$w_3(E)$	$w_4(W)$	Selected FP
Focus on minimizing energy consumption. Distance is secondary.	0.2	0	0.8	0	7
Prioritize distance. Altitude and energy consumption are less important.	0.6	0.15	0.15	0.1	7
Prioritize minimizing energy consumption. Waypoints and distance are less important.	0.1	0	0.7	0.2	7
Optimize for the shortest distance. Altitude and energy consumption are flexible.	1	0	0	0	7
Implicit Prompts					
Smooth flight is my main concern. Other factors are less critical.	0	1	0	0	1
I prefer a smooth flight with fewer stops, even if it takes a bit longer.	0.2	0	0	0.8	13
I want to minimize fuel use and have a moderately smooth flight.	0	0.4	0.6	0	1
A smooth, efficient flight with minimal stops is ideal.	0.2	0	0.5	0.3	7

prioritized based on a flight operator's preferences. The LLM's capability to interpret user prompts allowed for personalized flight plan recommendations, aligning closely with user-defined preferences. Our experimental results demonstrated the effectiveness of this approach in optimizing flight plans not only for safety and efficiency but also in accordance with individual operator needs.

While the current study provides a solid foundation for human-centric flight planning automation, there are several avenues for future research. One potential direction is to enhance the LLM's ability to interpret more complex and nuanced user prompts, possibly incorporating feedback loops to continuously refine the model's accuracy. Additionally, expanding the feature set to include dynamic weather conditions, real-time air traffic data, and environmental impact could further optimize flight plan recommendations.

REFERENCES

[1] C. Lee, B. Bae, Y. L. Lee, and T.-Y. Pak, "Societal acceptance of urban air mobility based on the technology adoption framework," *Technological Forecasting and Social Change*, vol. 196, p. 122807, Nov. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0040162523004924

[2] Federal Aviation Administration, "FAA aerospace forecast fiscal years 2020-2040," https://shorturl.at/uFIU8, 2020, (accessed Oct. 24, 2022).

[3] B. Tong, Q. Liu, and C. Dai, "A RRT*FN Based Path Replanning Algorithm," in *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Chengdu, China, Dec. 2019, pp. 1435–1445, iSSN: 2381-0947. [Online]. Available: https://ieeexplore.ieee.org/document/8997746

[4] L. Petit and A. L. Desbiens, "RRT-Rope: A deterministic shortening approach for fast near-optimal path planning in large-scale uncluttered 3D environments," in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, VIRTUAL, Oct. 2021, pp. 1111–1118, iSSN: 2577-1655. [Online]. Available: https://ieeexplore.ieee.org/document/9659071

[5] S. Karaman and E. Frazzoli, "Incremental Sampling-based Algorithms for Optimal Motion Planning," May 2010, arXiv:1005.0416 [cs]. [Online]. Available: http://arxiv.org/abs/1005.0416

[6] L. Kavraki, P. Svestka, J.-C. Latombe, and M. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 4, pp. 566–580, Aug. 1996. [Online]. Available: http://ieeexplore.ieee.org/document/508439/

[7] A. Stentz, "Optimal and efficient path planning for partially-known environments," in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. San Diego, CA, USA: IEEE Comput. Soc. Press, 1994, pp. 3310–3317. [Online]. Available: http://ieeexplore.ieee.org/document/351061/

[8] S. Koenig and M. Likhachev, "D*Lite," in *Eighteenth National Conference on Artificial Intelligence*. USA: American Association for Artificial Intelligence, 2002, p. 476–483.

[9] P. E. Hart, N. J. Nilsson, and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, Jul. 1968. [Online]. Available: https://ieeexplore.ieee.org/document/4082128

[10] T. Donateo and H. Çinar, "Conceptual design and sizing optimization based on minimum energy consumption of lift-cruise type eVTOL aircraft powered by battery and fuel cell for urban air mobility," *Journal of Physics: Conference Series*, vol. 2385, no. 1, p. 012072, Dec. 2022. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/2385/1/012072

[11] S. Sripad and V. Viswanathan, "The promise of energy-efficient battery-powered urban aircraft," *Proceedings of the National Academy of Sciences*, vol. 118, no. 45, p. e2111164118, Nov. 2021. [Online]. Available: https://pnas.org/doi/full/10.1073/pnas.2111164118

[12] P. Pradeep *et al.*, "Wind-Optimal Trajectories for Multirotor eVTOL Aircraft on UAM Missions," in *AIAA Aviation Forum 2020*. Virtual Event: American Institute of Aeronautics and Astronautics, Jun. 2020. [Online]. Available: https://arc.aiaa.org/doi/10.2514/6.2020-3271

[13] A. Taye, P. Wei, and J. Jones, "Energy demand analysis for eVTOL charging stations in urban air mobility," in *AIAA AVIATION Forum 2024*, Las Vegas, NV, 2024, p. 0532.

[14] C. S. Kulkarni, P. Pradeep, and G. B. Chatterji, "Simulation Studies for an Urban Air Mobility Aircraft using Hardware-In-Loop Experiments," in *AIAA AVIATION 2022 Forum*. Chicago, IL & Virtual: American Institute of Aeronautics and Astronautics, Jun. 2022. [Online]. Available: https://arc.aiaa.org/doi/10.2514/6.2022-3648

[15] A. Taye, S. Chen, and P. Wei, "Energy-aware strategic traffic management for urban air mobility," in *to appear in AIAA SciTech Forum 2025*, Orlando, FL, 2025.

[16] W. X. Zhao *et al.*, "A Survey of Large Language Models," Nov. 2023, arXiv:2303.18223 [cs]. [Online]. Available: http://arxiv.org/abs/2303.18223

[17] OpenAI *et al.*, "GPT-4 Technical Report," Mar. 2024, arXiv:2303.08774 [cs]. [Online]. Available: http://arxiv.org/abs/2303.08774

[18] D. Driess *et al.*, "PaLM-E: An Embodied Multimodal Language Model," Mar. 2023, arXiv:2303.03378 [cs]. [Online]. Available: http://arxiv.org/abs/2303.03378

[19] B. Li *et al.*, "SEED-Bench-2: Benchmarking Multimodal Large Language Models," Nov. 2023, arXiv:2311.17092 [cs]. [Online]. Available: http://arxiv.org/abs/2311.17092

[20] L. Wen *et al.*, "DiLu: A Knowledge-Driven Approach to Autonomous Driving with Large Language Models," Feb. 2024, arXiv:2309.16292 [cs]. [Online]. Available: http://arxiv.org/abs/2309.16292

[21] D. Fu *et al.*, "Drive Like a Human: Rethinking Autonomous Driving with Large Language Models," Jul. 2023, arXiv:2307.07162 [cs]. [Online]. Available: http://arxiv.org/abs/2307.07162

[22] W. Wang *et al.*, "DriveMLM: Aligning Multi-Modal Large Language Models with Behavioral Planning States for Autonomous Driving," Dec. 2023, arXiv:2312.09245 [cs]. [Online]. Available: http://arxiv.org/abs/2312.09245

[23] W. C. Skamarock *et al.*, "A Description of the Advanced Research WRF Model Version 4," National Center for Atmospheric Research, Boudler, CO, Tech. Rep., Mar. 2019, Technical Note TN–556+STR.

[24] A. Reuther *et al.*, "Interactive Supercomputing on 40,000 Cores for Machine Learning and Data Analysis," in *2018 IEEE High Performance extreme Computing Conference (HPEC)*. Waltham, MA: IEEE, Sep. 2018, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8547629/

[25] J. C. Jones, T. Bonin, and E. Mitchell, "Evaluating Wind Hazards for Advanced Air Mobility Operations," in *AIAA AVIATION 2023 Forum*. San Diego, CA and Online: American Institute of Aeronautics and Astronautics, Jun. 2023. [Online]. Available: https://arc.aiaa.org/doi/10.2514/6.2023-4104

[26] S. Razzaq, C. Xydeas, M. E. Everett, A. Mahmood, and T. Alquthami, "Three-Dimensional UAV Routing With Deconfliction," *IEEE Access*, vol. 6, pp. 21 536–21 551, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8345589/

[27] C. Silva, W. R. Johnson, E. Solis, M. D. Patterson, and K. R. Antcliff, "VTOL Urban Air Mobility Concept Vehicles for Technology Development," in *2018 Aviation Technology, Integration, and Operations Conference*. Atlanta, Georgia: American Institute of Aeronautics and Astronautics, Jun. 2018. [Online]. Available: https://arc.aiaa.org/doi/10.2514/6.2018-3847

[28] P. Pradeep and P. Wei, "Energy-Efficient Arrival with RTA Constraint for Multirotor eVTOL in Urban Air Mobility," *Journal of Aerospace Information Systems*, vol. 16, no. 7, pp. 263–277, Jul. 2019. [Online]. Available: https://arc.aiaa.org/doi/10.2514/1.I010710

[29] V. M. Becerra, "Solving complex optimal control problems at no cost with PSOPT," in *2010 IEEE International Symposium on Computer-Aided Control System Design*, Yokohama, Japan, Sep. 2010, pp. 1391–1396, iSSN: 2165-302X. [Online]. Available: https://ieeexplore.ieee.org/document/5612676

[30] OpenAI, "Chat Completions API Guide," https://platform.openai.com/docs/guides/text-generation/chat-completions-api, accessed: 2024-06-21.