JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 34, No. 3, May–June 2011

Autonomous Exploration of a Wind Field with a Gliding Aircraft

Nicholas R. J. Lawrance* and Salah Sukkarieh†
University of Sydney, New South Wales 2006, Australia

DOI: 10.2514/1.52236

Soaring is the process of exploiting favorable wind conditions to extend flight duration. This paper presents an approach for simultaneously mapping and using a wind field for soaring with an unpowered aircraft. Previous research by the authors and others has addressed soaring in known wind fields. However, an adequate estimate of the wind field is required in order to generate energy-gain paths. Conversely, the exploration required to generate a useful map estimate requires energy. This work aims to address these problems simultaneously by attempting to maintain and improve a model-free wind map based on observations collected during the flight and to use the currently available map to generate energy-gain paths. Wind estimation is performed using Gaussian process regression. A path planner generates and selects paths based on energy efficiency and field exploration. The method is tested in simulation with wind fields consisting of single and multiple stationary thermal bubbles. The use of soaring flight with consistent improvement in map quality and accuracy is demonstrated in a number of scenarios.

Nomenclature

AR	=	wing aspect ratio, b^2/S
b	=	wingspan, m
C_D	=	drag coefficient
C_L	=	lift coefficient
$\mathbf{C}_a^i$	=	air relative to inertial frame transformation matrix
D	=	drag, N
d_{goal}	=	distance to global goal, m
E	=	energy, J
e	=	Oswald's efficiency factor
g	=	gravitational acceleration, m/s^2
J_w	=	spatial wind gradient matrix, s^{-1}
k	=	number of wind observations
L	=	lift, N
m	=	vehicle mass, kg
n	=	load factor, L/mg
$\mathbf{P}$	=	inertial position vector
R	=	reward function, J
r_{lift}	=	thermal lifting radius, m
S	=	wing area, m^2
V	=	vehicle speed, m/s
V_{core}	=	thermal maximum core lift velocity, m/s
W	=	windspeed, m/s
γ	=	climb angle, deg
ρ	=	air density, kg/m^3
ϕ	=	bank angle, deg
ψ	=	heading angle, deg

Subscripts

a	=	air relative
i	=	inertial
0	=	start of planning segment
1	=	end of planning segment

Presented as Paper 2010-8032 at the AIAA Guidance, Navigation, and Control Conference, Toronto, Ontario, Canada, 2–5 August 2010; received 4 September 2010; revision received 15 December 2010; accepted for publication 18 December 2010.

*Postgraduate Student, Australian Centre for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering. Student Member AIAA.

†Associate Professor, Australian Centre for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering.

I. Introduction

IN RECENT years, unmanned aerial vehicle (UAV) technology has improved considerably and UAVs have become more accessible for both commercial and military applications. However, endurance is often a major limitation of these vehicles due to the limited availability of onboard energy storage for propulsion.

Soaring is the process of an aircraft capturing energy in-flight from motions in the surrounding wind field. This is achieved using either static soaring (the process of using rising air to gain energy) or dynamic soaring (using wind gradients or distributions). Rising air can be due to ground temperature variations, which result in boundary-layer mixing or geographic features such as mountain ridges that deflect wind upward. Wind gradients are found in boundary layers over fixed surfaces and can be caused by geographic features and turbulence.

The physical size of small and micro UAVs means that these aircraft are generally less efficient (lower lift-to-drag ratio) than larger gliding aircraft. This, combined with limited carriage capacity, restricts the available space for onboard energy storage and limits flight duration. Soaring control could be applied to UAVs to extend mission duration and reduce carried energy requirements. The current research aims to provide a method of simultaneously collecting information about a wind field and using that information for soaring. However, the problem is complicated by the fact that usually only direct observations of the wind are available, and atmospheric structures vary over a wide range of spatial and temporal scales.

Early soaring research focused on the ability of birds to fly for long periods of time with minimum wing flapping [1–3]. More recent research has used computational optimization techniques to determine trajectories for dynamic soaring in specific conditions [4–7]. The development and increasing popularity of manned gliders has also inspired further consideration of soaring strategies. Recent research has started to analyze how a UAV could generate energy-gain paths with known wind information [8,9]. More reactive strategies have also been proposed, where knowledge of soaring mechanisms is used to determine and track the current optimal flight condition based on direct observations [10–12]. A further alternative is to attempt fitting a predefined wind model (such as a thermal model) to observations made in flight. If an energy-gain strategy is determined offline, it can be adapted for the learned model parameters and repeated during flight to produce energy-gain paths. Initial work by Allen [13] was further developed by Edwards and Silverberg [14], culminating in an autonomous glider competing against remote-controlled aircraft in a cross-country soaring challenge.

While previous soaring research has focused on the methods of soaring, more recent research has started to focus on how soaring

control could be integrated with other mission goals. Cutler et al. [15] demonstrated the potential for using ridge lift during a surveillance mission to provide continuous unpowered flight while maintaining a seeability requirement for the surveillance mission. Recent work has also focused on the higher-level planning required to search for and use drifting thermals in a wind field [16].

However, previous research has not focused on whether an aircraft could attempt to generate a model-free map of the local wind field while simultaneously using that information to generate energy-gain paths. This paper presents a suggested method for simultaneous exploration and exploitation of a wind field with limited prior information. A spatial wind map is generated from wind observations taken during the flight. This map is then used to plan paths that explore a predefined area of the wind field while using energy found in the field for soaring. An alternative approach to the mapping problem with a focus on the sensor requirements was suggested by Langelaan et al., where the wind is sampled without requiring specialized sensors for aerodynamic angle measurements (such as multihole pressure probes or wind vanes) [17].

Section II provides a brief discussion of the mechanisms of energy gain and presents the dynamic model used in simulation. Section III discusses the theory and justification for using Gaussian process (GP) regression to generate an estimate of the wind field. Section IV presents the planning architecture and discusses the target assignment algorithm and path planner. The simulation is described in Sec. V, and results are presented in Sec. VI. Concluding remarks are made in Sec. VII.

II. Soaring Dynamics

Soaring and minimum-energy flight strategies can be identified by analyzing the equations of motion of an aircraft in wind. This section briefly describes the dynamic model (Sec. II.A) and examines the mechanisms of soaring through analysis of the resulting energy equations (Sec. II.B). By analyzing the effect of wind on an aircraft model, it is possible to identify the conditions required to gain energy.

A. Glider Dynamic Model

The dynamic model used in the analysis and simulation for this work is an aerodynamic point mass model. The applied forces are the aerodynamic force (decomposed into lift L and drag D) and the weight force mg. The aerodynamic force is a function of the motion of the vehicle relative to the surrounding air and the physical properties of the aircraft (shape, size, and surface properties). Lift is defined as the component of the aerodynamic force acting perpendicularly to the air-relative velocity vector. Body force due to sideslip is not considered. Weight is the force due to gravity and is directed toward the center of mass of the Earth. In this case, a flat Earth model is assumed due to the relatively small scale of the aircraft and flight paths. Thus, there are two critical frames of reference: an inertial frame, which is fixed with respect to the ground (in this case, a north-east-down frame), and the air-relative frame, which is the aircraft motion relative to the surrounding air. Figure 1 illustrates the forces acting on a gliding aircraft in wind.

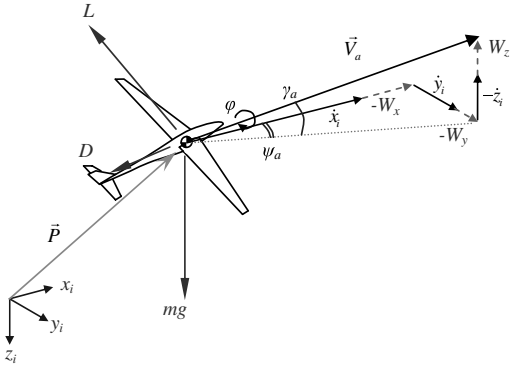


Fig. 1 Air-relative velocity and applied forces for a gliding aircraft.

Air-relative climb angle γ_a is the climb angle of the velocity vector with respect to the wind. Similarly, the air-relative heading ψ_a is the heading with respect to lateral wind. The bank angle ϕ is the rotation of the lift vector around the velocity vector. These angles and the velocity vectors are illustrated in Fig. 1. Wind is defined in inertial space and represents the motion of the atmosphere relative to the ground-fixed inertial frame. The air-relative-to-inertial-transformation matrix is denoted as $\mathbf{C}_a^i$ and is made up of the standard rotation transformation matrices denoted as $\mathbf{L_x}$, $\mathbf{L_y}$, and $\mathbf{L_z}$:

$$\mathbf{C}_a^i = \mathbf{L_z}(\psi_a)\mathbf{L_y}(\gamma_a)\mathbf{L_x}(\phi) \tag{1}$$

The air-relative velocity can be described in terms of the airspeed V_a and the heading and climb angles:

$$\mathbf{V}_a = \mathbf{C}_a^i \begin{bmatrix} V_a \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} V_a \cos\gamma_a \cos\psi_a \\ V_a \cos\gamma_a \sin\psi_a \\ -V_a \sin\gamma_a \end{bmatrix} \tag{2}$$

The drag coefficient is estimated using the common approximation where the effective drag coefficient C_D is the sum of parasitic $C_{D,0}$ and lift-induced $C_{D,i}$ drag components [18]. Induced drag is a function of the lift coefficient C_L, the aspect ratio AR, and the efficiency factor e:

$$C_D = C_{D,0} + C_{D,i} = C_{D,0} + \frac{C_L^2}{\pi \mathrm{AR} e} \tag{3}$$

Consider the case of locally spatially fixed linear wind gradients. Let $\mathbf{J_w}$ be the spatial wind gradients at a particular location:

$$\mathbf{J_w} = \begin{bmatrix} \frac{\partial W_x}{\partial x} & \frac{\partial W_x}{\partial y} & \frac{\partial W_x}{\partial z} \\ \frac{\partial W_y}{\partial x} & \frac{\partial W_y}{\partial y} & \frac{\partial W_y}{\partial z} \\ \frac{\partial W_z}{\partial x} & \frac{\partial W_z}{\partial y} & \frac{\partial W_z}{\partial z} \end{bmatrix} \tag{4}$$

By summing the applied forces (lift, weight, and drag) in the inertial frame and differentiating for acceleration, a set of dynamic equations can be obtained that describe the motion of a gliding aircraft. The wind field is stationary with locally linear spatial gradients. Assuming that roll rate is directly controlled, there are two free variables: climb angle rate ($\mathrm{d}\gamma_a/\mathrm{d}t$) and specific lift ($L/m$). The system can be solved by specifying one of these values as a control input. By specifying a climb angle rate, the resulting specific lift can be obtained [Eq. (5)]. However, physical limitations mean that the maximum specific lift is limited by maximum lift coefficient $C_{L,\max}$ and load factor constraints ($n_{\min}, n_{\max}$). In such cases, the lift is specified and Eq. (6) returns the climb angle rate. This effectively saturates climb angle rate commands to prevent stall and ensure commanded paths stay within physical limits. The resulting equations for the system are shown in Eqs. (5–9). Further information and derivation of the equations can be found from Lawrance and Sukkarieh [11]:

$$\frac{L}{m} = \frac{1}{\cos\phi}\left(V_a \frac{\mathrm{d}\gamma_a}{\mathrm{d}t} + g\cos\gamma_a - \begin{bmatrix} \cos\psi_a \sin\gamma_a \\ \sin\psi_a \sin\gamma_a \\ \cos\gamma_a \end{bmatrix}^T \mathbf{J_w}\dot{\mathbf{P}} \right) \tag{5}$$

$$\frac{\mathrm{d}\gamma_a}{\mathrm{d}t} = \frac{1}{V_a}\left(\frac{L}{m}\cos\phi - g\cos\gamma_a + \begin{bmatrix} \cos\psi_a \sin\gamma_a \\ \sin\psi_a \sin\gamma_a \\ \cos\gamma_a \end{bmatrix}^T \mathbf{J_w}\dot{\mathbf{P}} \right) \tag{6}$$

$$\frac{D}{m} = \frac{1}{2}\rho V_a^2 S C_{D,0} + \frac{L^2}{(1/2)\rho V_a^2 S \pi \mathrm{AR} e} \tag{7}$$

$$\frac{\mathrm{d}V_a}{\mathrm{d}t} = \frac{-D}{m} - g\sin\gamma_a - \begin{bmatrix} \cos\psi_a \cos\gamma_a \\ \sin\psi_a \cos\gamma_a \\ -\sin\gamma_a \end{bmatrix}^T \mathbf{J_w}\dot{\mathbf{P}} \tag{8}$$

$$\frac{\mathrm{d}\psi_a}{\mathrm{d}t} = \frac{1}{V_a \cos\gamma_a}\left(\frac{L}{m}\sin\phi + \begin{bmatrix} \sin\psi_a \\ -\cos\psi_a \\ 0 \end{bmatrix}^T \mathbf{J}_\mathbf{w}\dot{\mathbf{P}}\right) \tag{9}$$

These equations can be integrated numerically with specified roll rate and either pitch rate or lift functions to simulate gliding flight in wind. In the simulations presented in this paper, the pitch rate is normally specified but rejected if the lift required would violate either a predetermined maximum lift coefficient or a load factor limit. In that case, the lift is specified and the resulting pitch rate is determined (effectively modeling a saturation limit of pitch rate command due to maximum lift requirements). This effectively models stall prevention and limit loading. However, it does assume that a controller exists and is able to command and track a specified roll rate and either the pitch rate or lift coefficient.

B. Soaring Energy

The dynamic equations in Sec. II.A can be examined in terms of the energy gained or lost to give an understanding of how wind contributes to the overall energy of the platform. The energy of a point mass can be described as the sum of gravitational potential and kinetic energy. Gravitational potential energy is defined with respect to the altitude above the Earth-fixed inertial frame $-z_i$ and is independent of wind. The kinetic energy requires more careful consideration, as the velocity of the aircraft with respect to the surrounding air determines the energy rate from aerodynamic force, but the air itself may be moving with respect to the ground-fixed inertial frame. We define the air-relative energy E_a as the aerodynamic energy of the vehicle with respect to the surrounding air (treating the air as a local inertial frame with respect to the aircraft):

$$E_a = -mgz_i + \frac{1}{2}mV_a^2 \tag{10}$$

Taking the time derivative of the air-relative energy and substituting the airspeed acceleration [Eq. (8)] yields the overall specific energy rate:

$$\frac{\dot{E}_a}{m} = -V_a\frac{D}{m} - gW_z - V_a\begin{bmatrix} \cos\psi_a\cos\gamma_a \\ \sin\psi_a\cos\gamma_a \\ -\sin\gamma_a \end{bmatrix}^T \mathbf{J}_\mathbf{w}\dot{\mathbf{P}} \tag{11}$$

This equation illustrates how a gliding aircraft can gain or lose air-relative energy from a wind field. The first term is the power loss due to drag. This is always an energy loss term, since the airspeed must be greater than zero.

The second term is the static soaring term; that is, it represents energy gained or lost from vertical motion of the air relative to the fixed inertial frame. This term is only affected by the vertical wind: positive values of W_z (downdrafts) result in energy loss and negative values (updrafts) result in energy gain. Since this term is only a function of vertical wind, flying at minimum sink (minimizing drag) will result in the fastest energy gain from rising air: a fact well known to glider pilots. This is also the only term that can result in energy gain from a uniform wind field ($\mathbf{J}_\mathbf{w} = \mathbf{0}$).

The third term is the dynamic soaring term and represents energy gained or lost due to wind gradients. This term is affected by airspeed, climb angle, and wind gradients. Equation (11) shows that energy gained or lost from windshear is proportional to the airspeed. In general terms, increased airspeed increases energy capture or loss rate and allows energy capture from lower magnitude wind gradients.

III. Gaussian Process Regression for Wind Mapping

An adequate estimate of the wind map is required to generate paths through the wind field. The map must be generated from the set of observations made during the flight. For the simulations in this research, it is assumed that the aircraft is capable of measuring its air-relative velocity. On a real aircraft, this is usually through measurements of airspeed and angles of attack and sideslip. It is also assumed that an Earth-fixed inertial measurement system is available, such as a the Global Positioning System/inertial navigation system, which estimates inertial position and speed. The difference between the air data and inertial solutions is due to wind, such that it should be possible to estimate wind whenever there are two coincident measurements from the air data and inertial systems.

The mapping problem is complicated by the fact that there is not often a static underlying wind field. Most wind fields are the result of a number of driving components, such as geostrophic pressure-driven wind at high altitude, shear layers, and boundary mixing at lower altitudes. Additionally, there is turbulence from many factors, such as ground features and random air motion (gusts). In the current research, we consider static fields made up of fixed features like thermal bubbles with an overlaid random gust model.

A number of alternative mapping solutions were considered for this problem. However, some limitations should be noted. The mapping method needs to provide an estimate of the wind at any given point from the set of previous observations to be useful in path estimation. The wind field is also effectively boundless, so the solution either needs to be constrained within a set of boundaries or be able to make predictions at any location. Finally, the high dimensionality of the problem effectively eliminates the use of discrete methods such as grid-based approximations, as there would be too many grid cells with reasonable resolution for an onboard computer with limited computational power and storage.

For this research, GP regression was selected to perform the map estimation. A relatively simple explanation of GP regression for the purposes of using the method in this context is provided in the Appendix. A much more comprehensive coverage of the theory and application of GPs is provided by Rasmussen and Williams [19]. GP regression has been used to model wind fields on meteorological scales, where the method is often known as kriging or best linear unbiased prediction [20]. Spatial modeling of winds is also of interest in the planetary boundary layer for wind power, and alternative methods such as neural networks [21] and genetic algorithm learning of a fuzzy model [22] have been used to predict the spatial and temporal variation of wind.

The primary advantage of GP regression is that it provides a variance estimate at each test point. This value represents the confidence of estimates drawn from the learned model and can be used by the planner to estimate the confidence in a proposed path. The variance estimate also allows the path planner to identify regions of low or noisy information in the field, which can be targeted for further exploration. However, this comes at the cost of computational complexity, which may be significant depending on the operations required. Retraining the model (learning the hyperparameters) is a relatively expensive operation, as each iteration of the optimization requires reevaluation and inversion of the covariance matrix [$\mathcal{O}(k^3)$, where k is the number of observations]. Querying the model for estimates is usually simpler, as it requires only a single inversion of the covariance matrix, which can be stored inverted for fast retrieval of additional test points. With a stored inverted covariance matrix, generating a mean and covariance estimate is a matrix multiplication operation of $\mathcal{O}(k^2)$. However, the covariance matrix also needs to be updated as new observations arrive, although there are existing methods to update the inverted version directly to reduce computational complexity [23,24]. This would allow online operation with limited computational power and storage, but for the current desktop simulation, this was considered unnecessary and should not affect performance significantly. Computational performance was not a primary consideration for the current research, but the authors believe that the use of GP regression with limited sample points could feasibly be used on an onboard computer with relatively minor changes that should not affect functionality or performance.

In short, GPs are used to estimate an underlying function (in this case, the wind field) using a set of observations made during the flight. In the current application, the wind field is assumed to be smooth through use of the common squared exponential radial basis function [Eq. (A6)] as the kernel function with a single fixed length scale l, such that the input dimensions are considered isotropic (displacements in all axes are treated equally). In the current application, GP regression is used to estimate the windspeed and

direction from a set of observations made during flight. The training set X is the set of inertial positions of the aircraft where observations were made. The wind observations are separated into the Cartesian coordinate components of the wind measured in the inertial frame (W_x, W_y, and W_z). A separate GP is used for each output dimension, but the hyperparameters are shared between the three models, which reduces computational complexity to the equivalent of a single model. This effectively means that the output space is also considered isotropic. Thus, the solution is invariant to rotation of the coordinate axes.

IV. Simultaneous Exploration and Exploitation of a Wind Field

The utilization of a known wind field was addressed in Sec. II, and the mapping was discussed in Sec. III. This section provides a method of combining these two actions to simultaneously explore and exploit the wind environment. The goal of the system is to explore a limited space in the environment: in this case, a box defined by axis limits in inertial space. The exploration task is defined as minimization of the variance estimate of grid points from inside the target box. However, the system starts the mission with insufficient energy to explore the space, so it must also collect enough energy during the flight to complete the mapping problem. This is achieved using a target assignment algorithm that directs the aircraft to either further explore the space or attempt to gain more energy. A system overview is illustrated in Fig. 2, and the following subsections detail the path planning and target assignment modules.

A. Path Planning

The path planning routine is based on previous research by the authors [11] that uses the energy equations to generate energy efficient paths through a known wind field. The path planner uses an expanding tree structure similar to a state lattice, but it is based on a set of control inputs (roll rate and pitch rate commands) for a fixed time horizon rather than target positions (which may be unreachable). A path segment is generated by numerically integrating the model presented in Sec. II from the starting position with the specified control sequences. Segments are ranked using a heuristic energy-based reward function. The best performing branches are propagated with a new set of control inputs, the resulting path segments are again ranked using the same reward function, and the process repeats for a predefined tree depth, as illustrated in Fig. 3. While this method is suboptimal, it generates achievable paths and allows exploration of the field within reachable states.

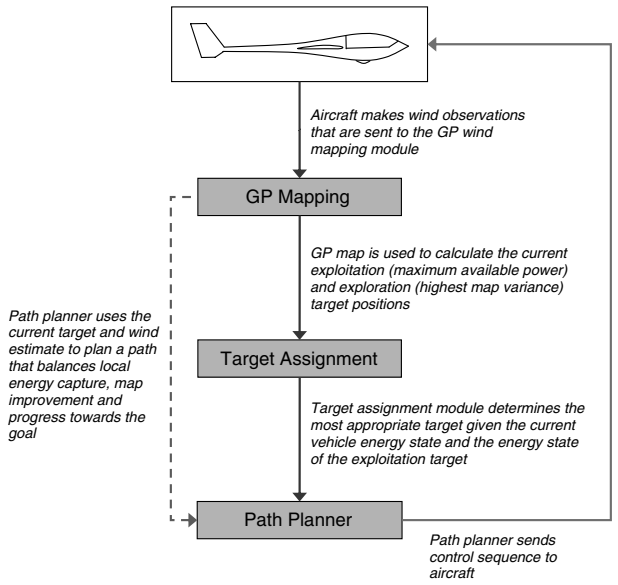


Fig. 2 System overview of simultaneous exploration and exploitation path planning architecture for a gliding UAV.

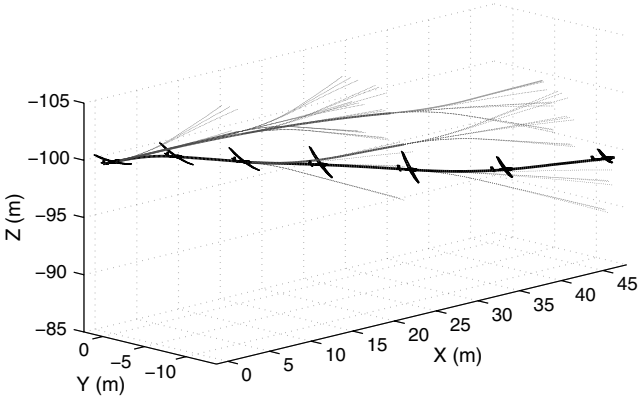


Fig. 3 Path planning tree for gliding flight with no wind.

The heuristic reward function R used to rank paths is a linear sum of three energy-based components quantifying the estimated energy change, the progress toward the global goal, and a reward for traveling through unexplored regions. All components are expressed in terms of energy to maintain consistency.

The first component, R_{energy}, estimates the energy change during a segment [Eq. (12)]. The total energy change is a sum of the potential and air-relative kinetic energy changes between the initial and final conditions (denoted with subscripts 0 and 1, respectively). An energy estimate based on the instantaneous power at the end of the segment $\dot{E}_1$ is also included with a fixed weight $K_{\dot{E}}$ to favor terminal power for further energy gain in subsequent segments. This was demonstrated to perform better than considering energy change alone by Lawrance and Sukkarieh [11]:

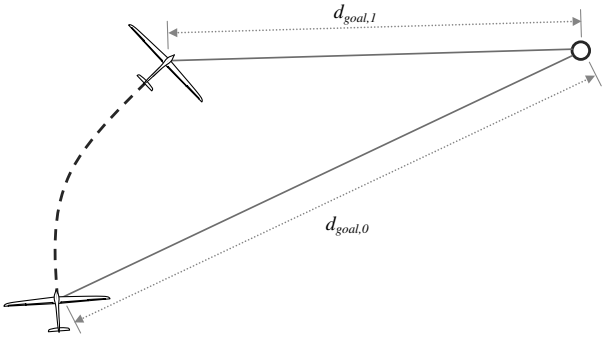
$$R_{\text{energy}} = -mg(z_{i,1} - z_{i,0}) + \tfrac{1}{2}m(V_{a,1}^2 - V_{a,0}^2) + K_{\dot{E}}\dot{E}_1\Delta t \quad (12)$$

Also considered is a navigation reward function, which rewards progress made toward a goal point determined by a higher-level planner. To be compatible with the energy reward, the navigation reward must also be expressed in terms of energy. The navigation reward is a discount on energy used to travel toward the goal. Figure 4 and Eq. (13) illustrate how the distance traveled toward the goal ($d_{\text{goal},0} - d_{\text{goal},1}$) is converted to a potential energy using the equivalent altitude from an estimate of the glide ratio $(L/D)_{\text{est}}$:

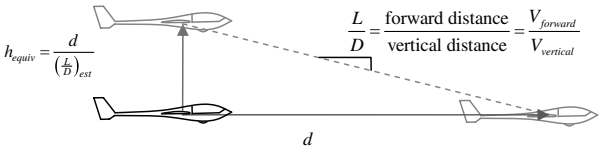
$$R_{\text{nav}} = mg\frac{d_{\text{goal},0} - d_{\text{goal},1}}{\left(\frac{L}{D}\right)_{\text{est}}} \quad (13)$$

While exploration of the entire field is facilitated by the global planner, the local planner can also take into account improvement of the local map by collecting sample points that minimize total variance. In smooth wind conditions, there are often path alternatives that are predicted to result in similar performance in terms of progress to the goal and energy efficiency, but they differ in the value of new information collected. Since the GP estimate of the current map variance is readily available, we propose using this information to select paths that pass through high variance regions to maximize the value of collected sample points.

To be compatible with the energy and navigation rewards presented thus far, the sampling reward function should be expressed in terms of the energy collected. However, the energy reward already estimates the energy that would be gained or lost given the current mean estimate of the wind field. The proposed solution is an energy heuristic based on the possible energy of a path segment: how much additional energy could be gained with favorable wind conditions over the mean estimate conditions. The possible additional energy is estimated by calculating the most favorable wind conditions within one standard deviation 1σ of the mean estimate. This uses the GP variance to reward exploration in poorly mapped (high variance) regions; but in well mapped (low variance) regions, the sampling



a) The distance traveled towards the goal is the difference between the distance to the goal at the beginning and end of the segment ($d_{goal,0} - d_{goal,1}$)



b) The equivalent altitude h_{equiv} represents the altitude that would be required to travel horizontal distance d at glide ratio $\left(\frac{L}{D}\right)_{est}$

Fig. 4 Estimation of the navigation reward function in terms of the equivalent energy of travel toward the goal.

reward becomes insignificant and the total reward is dominated by the energy and navigation components.

Calculating the correct gradient alignments and using the full set of dynamic equations to simulate each segment is computationally expensive. A faster alternative is to estimate the energy difference along the entire segment compared with the mean estimate case. From Eq. (10), additional wind energy must come from an increase in altitude and/or a change in effective airspeed. Thus, a reasonable method of estimating the possible energy gain would be to calculate the vertical displacement and airspeed changes along the length of the segment under favorable conditions. A simple longitudinal example is illustrated in Fig. 5.

Take a path estimate generated with the mean wind from point $p_0 = [x_0, y_0, z_0]^T$ to point $p_1 = [x_1, y_1, z_1]^T$. The corresponding airspeeds V_0 and V_1 are also known. Additionally, the wind variances at the start $[\sigma^2_{Wx,0}, \sigma^2_{Wy,0}, \sigma^2_{Wz,0}]^T$ and end $[\sigma^2_{Wx,1}, \sigma^2_{Wy,1}, \sigma^2_{Wz,1}]^T$ points are calculated from the GP map. Now consider the 1σ best-case favorable wind. In the z direction, this is always a lifting wind, so the best case is the mean wind plus $1\sigma_{W_z}$ vertical wind. The net effect is an increase of $\frac{1}{2}(\sigma_{Wz,0} + \sigma_{Wz,1})$ along the length of the segment. Increasing the total mean wind in the other directions does not increase energy; it only induces lateral motion. The maximum shear energy is always due to a maximum shear gradient, which will be opposing estimates at either end of the segment (with the direction determined by the heading of the aircraft), which results in a total variation of $\sigma_{Wx,0} + \sigma_{Wx,1}$ in the x direction and $\sigma_{Wy,0} + \sigma_{Wy,1}$ in the y direction along the length of the path segment. Thus, the most favorable 1σ wind change along the entire segment is

$$\Delta \mathbf{W} = \begin{bmatrix} \pm(\sigma_{Wx,0} + \sigma_{Wx,1}) \\ \pm(\sigma_{Wy,0} + \sigma_{Wy,1}) \\ \pm(\sigma_{Wz,0} - \sigma_{Wz,1}) \end{bmatrix} \tag{14}$$

Only the component of wind in the direction of flight will change the effective airspeed. This component can be calculated by taking the scalar projection of the wind vector in the flight direction. Let the airspeed at the end of the segment with favorable wind be V'_1:

$$\Delta V = V'_1 - V_1 = \Delta \mathbf{W} \cdot \hat{\mathbf{x}}' = \left(\begin{bmatrix} \sigma_{Wx,0} + \sigma_{Wx,1} \\ \sigma_{Wy,0} + \sigma_{Wy,1} \\ |\sigma_{Wz,0} - \sigma_{Wz,1}| \end{bmatrix} . \begin{bmatrix} |x'_1 - x_0| \\ |y'_1 - y_0| \\ |z'_1 - z_0| \end{bmatrix} \right) \Big/ \sqrt{(x'_1 - x_0)^2 + (y'_1 - y_0)^2 + (z'_1 - z_0)^2} \tag{15}$$

The additional airspeed results in an increase in total energy and is a function of the current airspeed. Letting the mean original airspeed along the segment be $\bar{V} = \frac{1}{2}(V_0 + V_1)$, the total kinetic energy increase is

$$\begin{aligned} \Delta E_k &= E_{k,1\sigma\text{favourable}} - E_{k,\text{mean}} \approx \tfrac{1}{2}m(\bar{V} + \Delta V)^2 - \tfrac{1}{2}m\bar{V}^2 \\ &= \tfrac{1}{2}m(2\bar{V}\Delta V + \Delta V^2) \end{aligned} \tag{16}$$

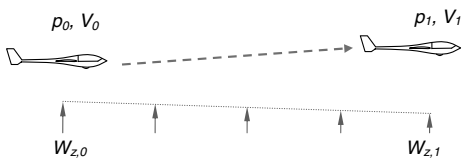
Gravitational potential energy is gained though the increased vertical component leading to additional altitude at the end of the segment:

$$\Delta E_p = E_{p,1\sigma\text{favourable}} - E_{p,\text{mean}} \approx \tfrac{1}{2}(\sigma_{Wz,0} + \sigma_{Wz,1})mg\Delta t \tag{17}$$

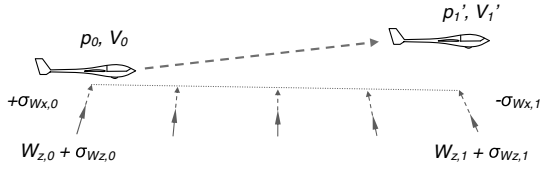
The corresponding total energy change is defined as the sampling reward function R_{sampling}:

$$\begin{aligned} R_{\text{sampling}} &= \Delta E_k + \Delta E_p = \tfrac{1}{2}(\sigma_{Wz,0} + \sigma_{Wz,1})mg\Delta t \\ &\quad + \tfrac{1}{2}m(2\bar{V}\Delta V + \Delta V^2) \end{aligned} \tag{18}$$

Note that this form of reward assumes an optimistic estimate of the wind. However, it is equally likely that the opposite case occurs, with 1σ opposing wind (downdraft, adverse wind gradient) resulting in a poorer energy performance than predicted. In this case, the exploration still provides utility, as improvement in the map and a consequent reduction in variance limits future exploration of that area. Also, the length scale of wind features is usually higher than the length of a planning segment, so the mapping component primarily acts in tie-break situations where there is similar energy performance but alternative paths could improve local mapping.



a) The original forward simulated path based on the mean wind estimate



b) Favorable wind conditions based on one standard deviation from the mean

Fig. 5 Additional air-relative energy for flight in favorable wind conditions.

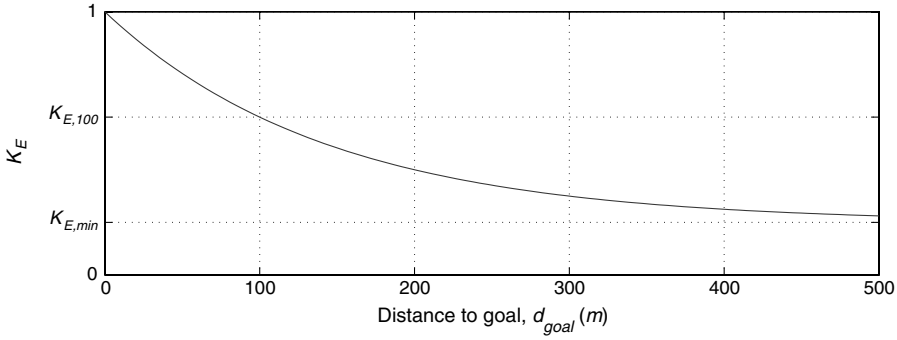


Fig. 6 Variation of energy weighting K_E as a function of distance to goal d_{goal} [Eq. (20)].

A weighted sum of the three specified reward functions is used to balance local energy capture, progress toward the goal, and local field exploration. In general, the goal points are approximations of a target area rather than specific points that need to be flown to. Near the goal, the capture of energy and new information is most important. Further from the goal, it is more important to travel toward the goal. A variable weighting is used between energy and navigation rewards. The sampling reward weighting is fixed, as it has equal significance throughout the flight. The variable weighting is $K_E \in [0, 1]$; lower K_E values give precedence to progress toward the goal, and higher K_E values favor local energy gains:

$$R = K_E R_{\text{energy}} + (1 - K_E) R_{\text{nav}} + K_{\text{sampling}} R_{\text{sampling}} \tag{19}$$

The value of K_E is determined by an exponential decay function, which assesses how near the goal the aircraft is and adjusts the gain accordingly. The value of K_E approaches a constant minimum value $K_{E,\min}$ at a distance far from the goal. As it approaches the goal, K_E increases, favoring energy gain over progress toward the goal. When the aircraft is at the goal, K_E is at its maximum value of one, and the navigation reward is neglected. To determine the decay rate, a second variable is introduced, $K_{E,100}$, which is the value of K_E at a distance of 100 m from the goal. In the current work, this value was selected offline at 0.6 but could be determined by estimating the size of thermals online. The equation and an illustration of the behavior of K_E as a function of distance to the goal d_{goal} are presented in Eq. (20) and Fig. 6, respectively:

$$K_E(d_{\text{goal}}) = K_{E,\min} + (1 - K_{E,\min}) \exp\left[\frac{d_{\text{goal}}}{100} \log\left(\frac{K_{E,100} - K_{E,\min}}{1 - K_{E,\min}}\right)\right] \tag{20}$$

The automatic balancing of the reward function allows the goal point to be set at an energy source (such as a thermal) or any other target while retaining the same path planning structure. In this way, the path planning module acts as the low-level guidance controller; given a goal location and the number of branches to search, it will return the control sequence required to produce an efficient path to the goal based on the current wind field estimate. In the current work, the path planning algorithm uses the GP map estimate to generate control sequences, which are carried out open loop by the simulated aircraft.

B. Target Assignment

The target assignment algorithm carries out the higher-level task of determining whether to pursue an exploitation (energy gain) or exploration (map improvement) goal. For each planning cycle, two goal points are defined. The exploration goal $\mathbf{P}_{\text{explore}}$ is defined as the point of maximum estimated uncertainty drawn from a regular three-dimensional grid of the target region. The energy goal $\mathbf{P}_{\text{exploit}}$ is defined as the point on the previous trajectory where maximum power gain was recorded. The target assignment algorithm selects between these goals each planning cycle.

Algorithm 1 Global target assignment algorithm

Each replanning cycle:
$E_{\text{travel}} = mgd_{\text{explore}}/(L/D)_{\text{est}}$
$E_{\text{explore}} = -mgz_{i,\text{explore}} + \frac{1}{2}mV^2_{\text{cruise}}$
$E_{\text{current}} = -mgz_{i,\text{current}} + \frac{1}{2}mV^2_{\text{current}}$
if $E_{\text{current}} \geq (E_{\text{explore}} + E_{\text{travel}})$ then
  $\mathbf{P}_{\text{target}} \leftarrow \mathbf{P}_{\text{explore}}$
else
  $\mathbf{P}_{\text{target}} \leftarrow \mathbf{P}_{\text{exploit}}$
end if

The goal of the mission is to reduce total map uncertainty over the entire region. However, depending on the energy state of the vehicle, the current exploration goal may be unreachable. An estimate of the energy required to reach the exploration goal E_{travel} is made by taking the distance to the goal d_{explore} and an estimate of the lift-to-drag ratio to determine the equivalent height and resulting gravitational potential energy required to travel that distance [as in Eq. (13)]. The energy required at the exploration goal E_{explore} is defined at the altitude of the target $-z_{i,\text{explore}}$ and a fixed speed (usually the cruise airspeed V_{cruise}). If the aircraft currently has more energy than the sum of the target energy and the travel energy, then we estimate that the aircraft can reach the exploration goal and the current goal is defined as the exploration goal. Otherwise, the current goal is defined as the exploitation goal, as summarized in Algorithm 1.

In this way, the aircraft either has enough energy to attempt to explore the most uncertain region of the map or returns to the previously best-known energy source to collect enough energy until exploration is possible. There are limits on the exploitation goal to prevent energy-gain attempts from insufficient sources (less than the sink speed), and the target rises in altitude with the aircraft to continue energy-gain flight.

V. Simulation Setup

The aircraft is simulated by numerical integration of the dynamic equations from Sec. II. The aircraft parameters are derived from an RnR SB-XC remote-controlled cross-country glider model, and relevant parameters are shown in Table 1. The wind is represented by a static field consisting of a number of thermal bubbles of various strengths and sizes, as described for each simulation in Sec. VI. Air flows around elliptical toroid sections: flowing inward and upward at the bottom of the torus and out again at the top. For each thermal, the maximum core lift speed V_{core} and lifting radius r_{lift} are specified. The vertical lift component V_z varies as a function of the radial distance from the center of the thermal r, as shown in Eq. (21). The mean lift across the lifting disk is $(4/\pi^2)V_{\text{core}} \approx 0.41V_{\text{core}}$. Note that the flow is conservative, such that the lifting volume of air for $r \in [0, r_{\text{lift}}]$ is the same as the sinking volume of air for $r \in [r_{\text{lift}}, 2r_{\text{lift}}]$. Further information can be found in [11] by Lawrance and Sukkarieh:

Table 1 Aerodynamic and geometric properties of the SB-XC glider model

Parameter	Value	Units	Explanation
$C_{D,0}$	0.012		Parasitic drag coefficient
b	4.32	m	Wingspan
S	0.957	m^2	Wing reference area
AR	19.54		Wing aspect ratio
e	0.85		Oswald's efficiency factor
m	5.44	kg	Vehicle mass
$n_{\max}$	2.0		Maximum load factor (positive)
$n_{\min}$	0		Minimum load factor (negative)
$C_{L,\max}$	1.2		Maximum lift coefficient
$d\phi/dt_{\max}$	30	deg/s	Maximum roll rate
$\gamma_{a,\max}$	50	deg	Maximum air-relative climb angle
$(L/D)_{\text{est}}$	30		Approximate glide ratio (for reward function)

$$V_z(r) = \begin{cases} V_{\text{core}} & r = 0 \\ \frac{r_{\text{lift}} V_{\text{core}}}{\pi r} \sin\left(\frac{\pi r}{r_{\text{lift}}}\right) & r \in (0, 2r_{\text{lift}}] \end{cases} \quad (21)$$

The GP mapping uses observation data collected during the flight by a simulated air data system. This system is simulated by taking the actual wind data from the simulation and adding unbiased Gaussian noise with standard deviation 0.1 m/s to represent measurement error. As explained further in the Appendix, the GP regression uses a square exponential covariance function with three hyperparameters: process length scale l, process variance σ_f^2, and noise variance σ_n^2. The GP hyperparameters were determined by flying 10 simulated trajectories through static wind fields with a random set of up to three thermal columns ranging in size from 25–100 m radius and core strength from 0–7 m/s. Each flight collected 300 data points, and the hyperparameters were trained using the MATLAB minimization routine fminunc to minimize the negative log marginal likelihood. The mean of the resulting hyperparameters was $l = 45$ m, $\sigma_f = 0.5$ m/s. and $\sigma_n = 0.09$ m/s. These values are used and remain fixed in the following simulations.

During each simulation presented in the results, at most, 150 data points are stored and used for regression to limit computational load. The simulated sensor system collects data at high frequency (10 Hz), so the maximum number of observations is reached relatively early in the flight. Observations must be discarded to maintain a database of the most useful observations. In the current implementation, the observations are ranked by the minimum distance to other observations in the set. The square distance is already calculated as part of the covariance function, so the results only require sorting. Observations that are too close together are removed and replaced with new observations. This results in a natural spatial sparsity of the data set representing the best coverage of the target space with the limited number of observations made along the flight trajectory.

The path planner uses the current wind field estimate generated by the GP mapping module to plan paths. Each planned segment is 1 s, with a total tree depth of five and replanning every 3 s. There are three scaled roll rate commands, up to a maximum bank of 45 deg, and three pitch rate commands (with load limit and maximum lift coefficient constraints), for a total of nine control options for each branch. The control sequence returned by the path planner is carried out open loop by the simulated aircraft with wind data drawn from the simulated wind field.

Target assignment is based on two targets returned from the GP map. The exploration target is the point of maximum variance based on a regular $8 \times 8 \times 5$ grid drawn from the target exploration region. The energy target is defined as the point at which the aircraft gained maximum power during the flight so far.

Also included in the final set of results is a gust model to simulate turbulence. The Dryden continuous turbulence model is used with specifications from MIL-F-8785C [25,26]. The turbulence is generated for a moderate level of turbulence at low altitude (less than 1000 ft), with mean windspeed at 20 ft altitude u_{20} of 15.23 m/s (30 kt). The turbulence is calculated at each time step for the current airspeed and altitude using Gaussian noise.

VI. Results

The simultaneous exploration and exploitation strategy is demonstrated in three simulated trajectories. Each simulation consists of a static wind field with a region of exploration defined. The aircraft begins the simulation with a small set of data already collected; this represents manual or autonomous flight before the autonomous soaring control is activated. These data provide a repeatable exploration path at the start of each simulation but are not necessary for autonomous soaring, and the control architecture does not require initial data.

This first simulation illustrates exploration of a region containing a single thermal near the corner of the region. The aircraft starts autonomous soaring at (0, 60, −200) heading in the direction of the positive x axis. There is a single thermal bubble centered at (350, 50, −200) with maximum vertical wind strength 3 m/s and a lifting radius of 100 m. The exploration box is defined by the limits $x \in [0, 400]$, $y \in [-100, 100]$, and $z \in [-250, -150]$. The simulation runs for 500 s. The resulting path is shown in Fig. 7. Linear dimensions are in meters.

Figure 8 shows the progress of the simulation from a top-down view. The straight path with a semicircular curve terminating at (0, 60, −200) indicates flight before autonomous soaring. Current sample points are indicated by the symbol +. The soaring controller is capable of using the energy in the thermal bubble to effectively explore the region. Initially, the aircraft does not pass through the thermal. During early exploration of the region, the thermal is discovered and used for the remainder of the flight. The general pattern of motion is to capture energy from the thermal and then use that energy to explore new regions. However, the energy reward limits the ability of the planner to use path segments that result in excessive local energy loss. While this is often beneficial, it can occasionally cause excessive detours, as seen in some of the longer loops extending outside the region of interest. In these cases, the airspeed is usually relatively high after exiting the thermal. A sharp

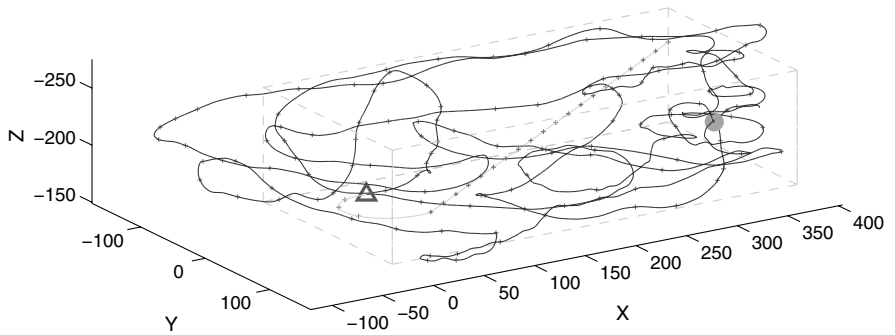


Fig. 7 Single thermal exploration at $t = 500$ s. Autonomous soaring flight starts at the triangle. There is a single thermal bubble with core vertical windspeed of 3 m/s illustrated by a filled circle.

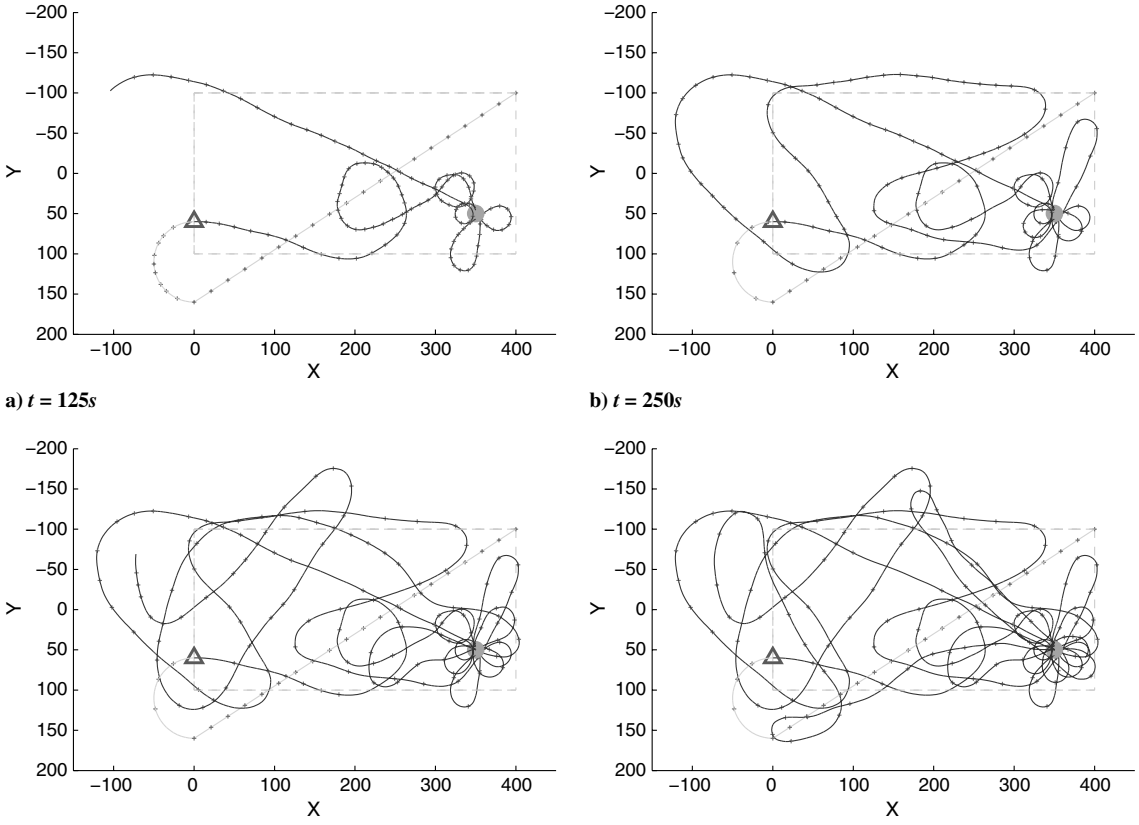
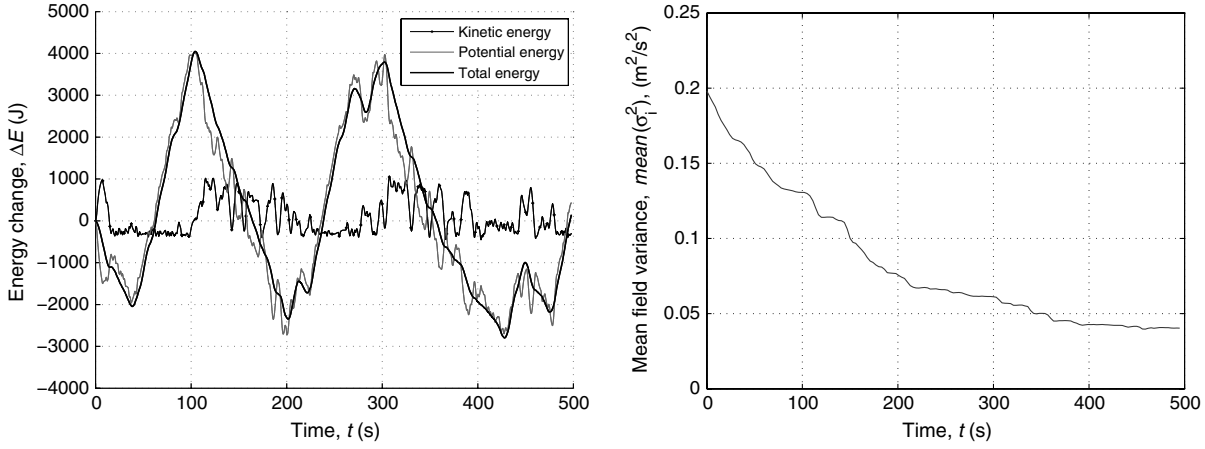


Fig. 8 Exploration history of a wind field with a single thermal. Stored sample positions are indicated by + symbols.

turn at high speed would result in rapid energy loss, so the planner continues to select segments that move further from the goal but are locally energy efficient. Eventually, the navigation reward becomes more significant and/or the aircraft slows down enough so that a turn can be made efficiently. This highlights a limitation of the local planning strategy that could be improved with an increased planning horizon. Similarly, large changes in commanded altitude often occur as the exploration goal is moved to a new target. Since lower altitude necessarily means a loss in energy, the planned paths tend to traverse longer distances to minimize energy loss rate rather than diving toward the target with rapid energy loss. This is generally favorable and allows the sampling reward to have a more significant effect to maximize the benefit of data points collected between exploration targets.

The mean field variance is the mean of the variance estimates taken at grid points throughout the region of interest. Both the energy change and the change in mean variance are shown in Fig. 9. The energy variation illustrates the planner alternating between capturing energy and exploring the region. This can be seen in the correlation between the two plots. Initially, the aircraft is mostly flying in unexplored regions, resulting in a rapid reduction in variance. However, as the aircraft returns to the energy source to capture more energy ($t = 200$–300 s), the new observations are close to previous ones and do not result in significant reduction of the field variance.



a) Energy change during flight

b) Mean variance estimate during flight (mean field variance is mean of variance estimates drawn from regular 8x8x5 grid of target exploration region)

Fig. 9 Energy and estimated variance during single thermal exploration.

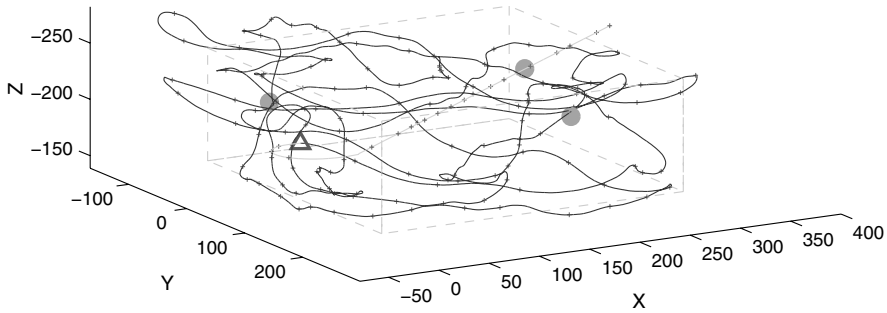
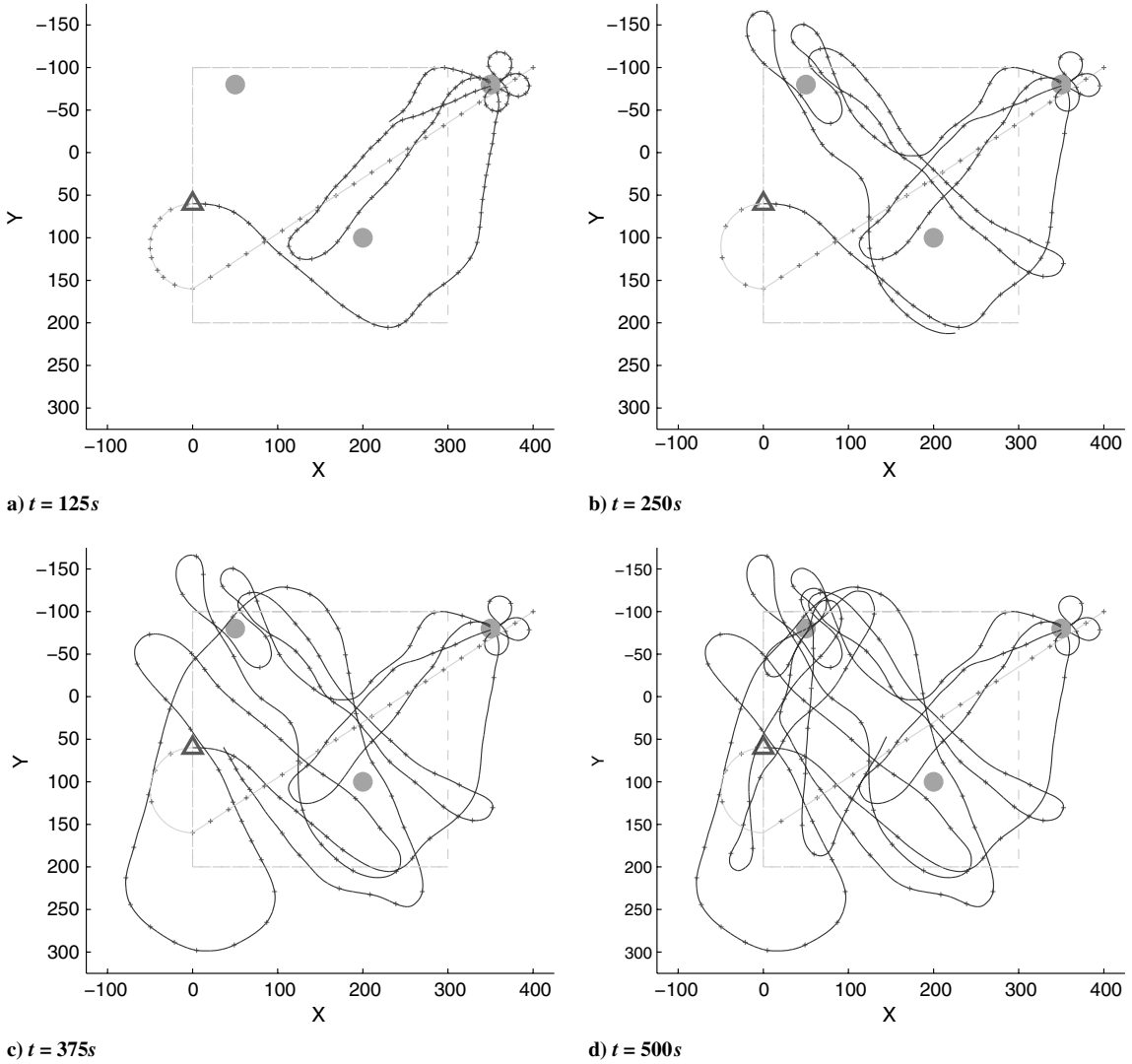


Fig. 10 Multiple thermal exploration at $t = 500$ s. There are three thermal bubbles centered at (200, 100, −250), (350, −80, −150), and (50, −80, −200) with V_{core} of 2, 3, and 5 m/s, respectively.

After enough energy is captured, the aircraft continues to explore the field, resulting in further reduction of the variance.

In general, the variance shows an asymptotic decline in overall variance toward a minimum value. This represents the fact that there are a limited number of sample points and, due to the nature of GPs, there is necessarily some uncertainty in the regions between the sample points. Overall uncertainty could be further reduced by increasing the number of sample points and allowing additional time for further exploration.

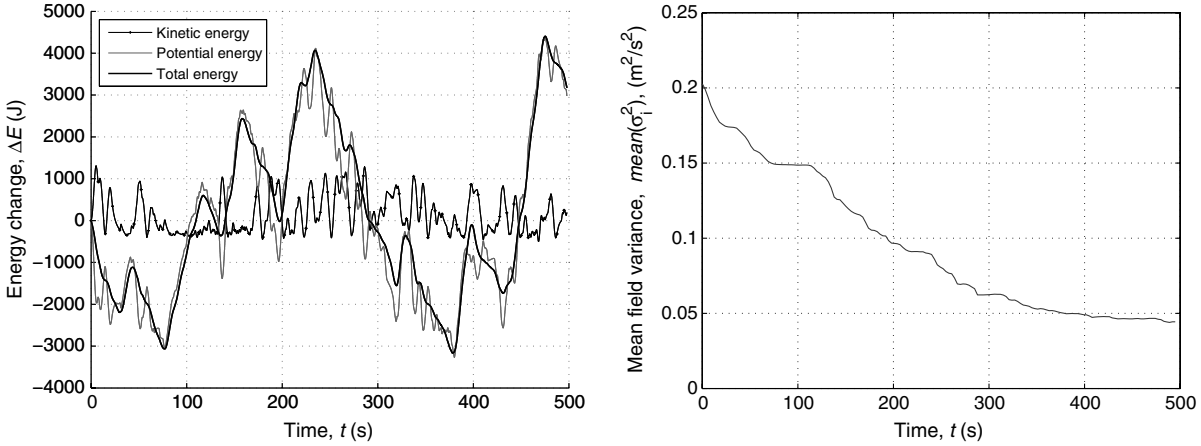
The second simulation result demonstrates autonomous soaring and mapping over a larger region with three thermal bubbles. In this case, there are two small thermals located at (200, 100, −250) and (350, −80, −150) with V_{core} of 2 and 3 m/s and $r_{\text{lift}} = 50$ and 75 m, respectively. There is also a larger, more powerful thermal with $r_{\text{lift}} = 100$ m and $V_{\text{core}} = 5$ m/s located at (50, −80, −200). The exploration box is $x \in [0, 300]$, $y \in [-100, 200]$, and $z \in [-250, -150]$. The simulation time is 500 s and the resulting path is shown in Figs. 10 and 11.



a) $t = 125s$ **b) $t = 250s$**

c) $t = 375s$ **d) $t = 500s$**

Fig. 11 Exploration history of a wind field with three thermals. The thermals are centered at (200, 100, −250), (350, −80, −150), and (50, −80, −200) with V_{core} of 2, 3, and 5 m/s, respectively.



a) Energy change during flight **b) Mean variance estimate during flight**

Fig. 12 Energy and estimated variance multiple thermal exploration.

These results demonstrate the ability of the system to handle multiple energy sources. When autonomous control begins the only known energy source is the 3 m/s thermal at (350, −80, −150). This is used during the first two traversals of the field. At $t \approx 130$ s, the larger thermal is found. For the remainder of the flight, the planner can gain enough energy from short traversals through the strong thermal to continue exploration (see Fig. 12). This demonstrates that the method can use different strength energy sources, with considerably less circling required in the strong thermal than the smaller thermals.

A similar pattern in energy use and variance reduction is shown in the multiple thermal case as in the single thermal case. Again, the controller cycles through energy collection and exploration phases. The later cycles are shorter due to the high power available from the large thermal. The results show a fairly steady reduction in overall variance to a similar level seen in the single thermal case. This is a good result given the larger size of the exploration region. Also worth noting is the use of the smallest thermal during flights through that region. Despite the smallest thermal never acting as an energy target, the low-level controller still makes use of the available energy for paths crossing through or near the thermal bubble. This demonstrates the utility of the reward function in generating local paths, which use any available energy.

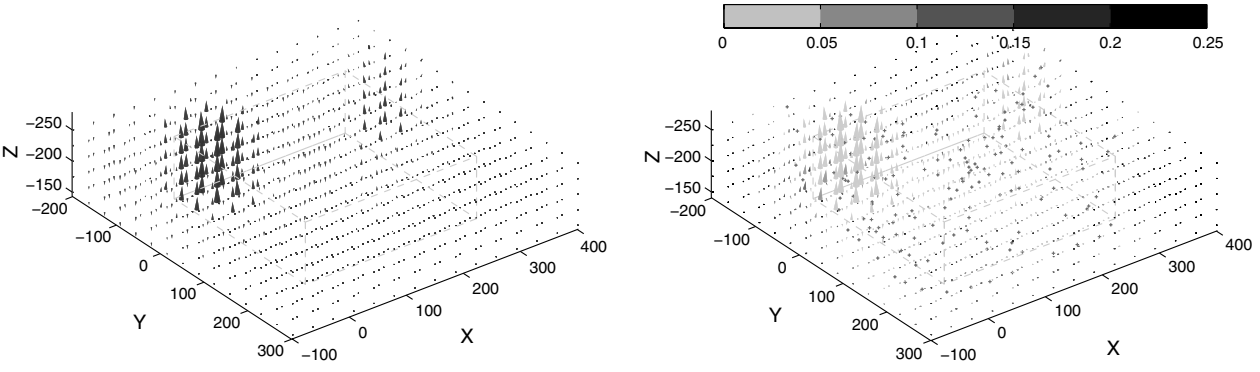
Figure 13 illustrates the estimate of the wind field at the end of the multithermal simulation. This is provided to demonstrate that the exploration is effective at mapping the complete wind field. The GP regression estimates for the thermal core velocities are 1.22, 2.57, and 5.26 m/s for the 2, 3, and 5 m/s thermals, respectively. Also captured is the toroidal shape of the thermals and good estimation of the inflow and outflow in the smaller thermals (the larger thermal is too tall for this to be noticeable). This represents an accurate estimate of the wind field and demonstrates the ability of the GP mapping to account for a number of structures simultaneously, which can be difficult for model-based methods that have to separately identify each structure. Of further note is that retraining the model with the final observation set yields optimal hyperparameters of $l = 45.05$, $\sigma_f = 0.489$, and $\sigma_n = 0.055$, very similar to the values used in simulation. This suggests that retraining should be possible during flight, since a gradient descent would be able to track the optimal solution as the hyperparameters vary slowly from a good initial estimate.

The final results present a more difficult soaring case to demonstrate some of the limitations of the current method. In the final scenario, the wind field is made up of a sinusoidal wave and a single thermal. The wave field is generated using a sinusoidal variation along the x axis:

$$W_x = -V_{\text{wave}} \cos\left[\frac{\pi}{4}\cos\left(\frac{2\pi x}{\lambda}\right)\right] \tag{22}$$

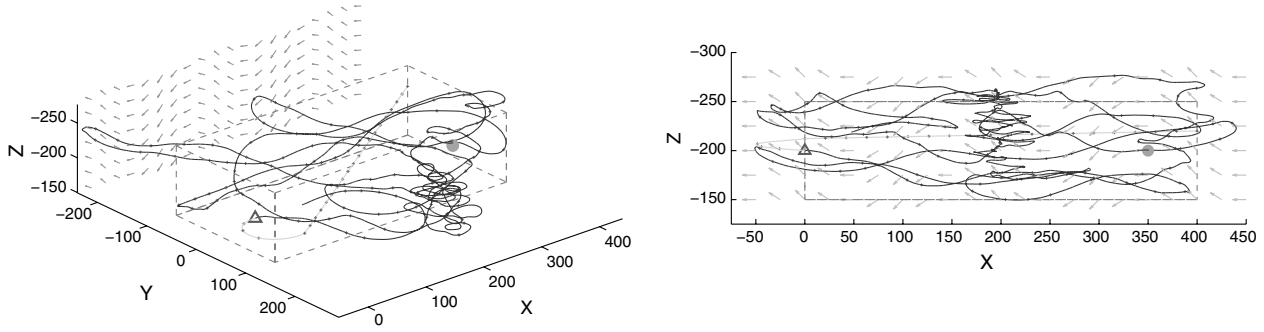
$$W_z = -V_{\text{wave}} \sin\left[\frac{\pi}{4}\cos\left(\frac{2\pi x}{\lambda}\right)\right] \tag{23}$$

The wind strength V_{wave} is 2 m/s, and the wavelength λ is 200 m. The maximum lift is 1.41 m/s, which is enough for energy-gain flight (minimum sink is ∼0.60 m/s straight and level at $V = 15$ m/s) but in a relatively narrow region. The thermal is centered at (350, 50, −200) as in the first set of results but with $V_{\text{core}} = 2.5$ m/s and a reduced lifting radius of $r_{\text{lift}} = 75$ m. The resulting path is shown in



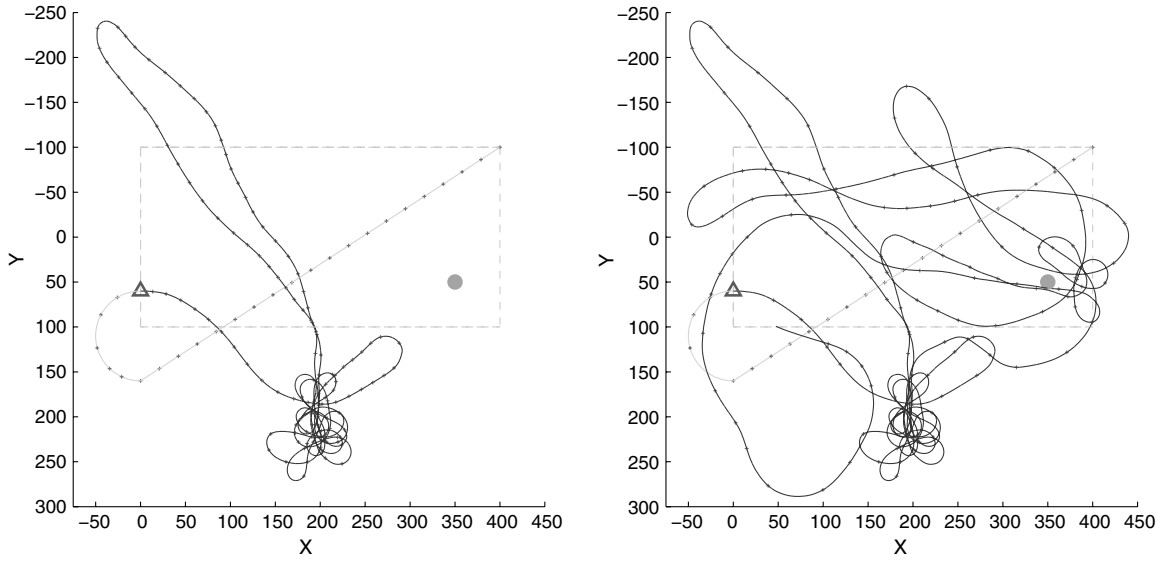
a) Actual wind field **b) Final field estimate**

Fig. 13 Final estimate of wind field from multithermal flight. The cone shading in Fig. 13b represents the wind variance estimate at each sample point, σ_W^2 in $(\text{m/s})^2$.



a) Oblique view b) Side (xz) view

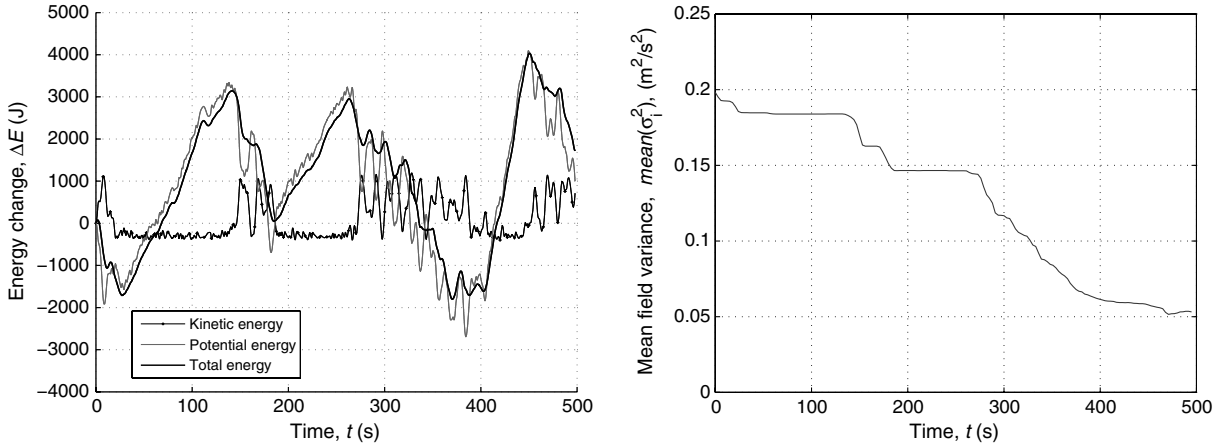
Fig. 14 Exploration of a wave and thermal wind field. The wave is shown by the vector field projected onto the xz plane. The thermal bubble has $V_{\text{core}} = 2.5$ m/s, and the location is illustrated by the filled circle.



a) $t = 250s$ b) $t = 500s$

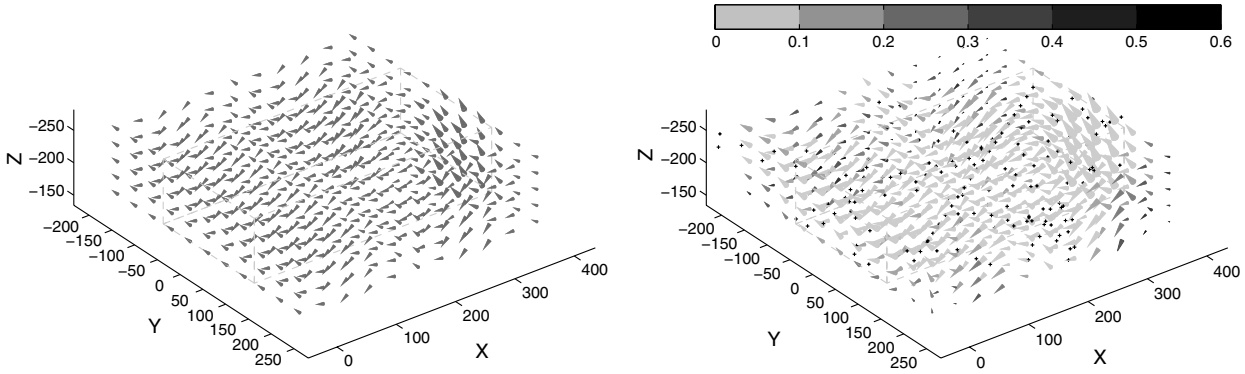
Fig. 15 Exploration history of a wave and thermal field. The location of the thermal bubble is indicated by the filled circle.

Fig. 14. In this scenario, the thermal is not found until late in the exploration ($t \approx 400$ s). Although exploration of the field initially heads toward the thermal, energy is lost in travel due to relatively poor estimation of the (initially) unknown sinusoidal field. However, it travels close enough, such that the (0, −100, −250) corner of the box becomes the next exploration target. Also, there is just enough energy in the rising regions of the sinusoidal oscillation to gain altitude. However, the constant flow in the negative x direction makes energy gain difficult and relatively slow. Thus, the planner spends a significant amount of time attempting to gain energy from the rising air in the sinusoid (see Fig. 15a). During this time, there is no reduction in field variance, as the aircraft is outside the exploration



a) Energy change during flight b) Mean variance estimate during flight

Fig. 16 Energy and estimated variance during wave and thermal exploration.



a) Actual wind field **b) Final field estimate**

Fig. 17 Final mean estimate of the wind field from wave and thermal flight. The cone shading in Fig. 17b represents the wind variance estimate at each sample point, σ_W^2 in $(m/s)^2$.

region (see Fig. 16). Eventually, it recovers enough to explore the region around the $x = 0$ face, and the information target moves to the opposing ($x = 400$) face. However, the energy lost during exploration needs to be recovered again, and the planner spends more time recovering energy from the rising areas of the wave flow. Finally, at $t \approx 400$ s, the thermal is found with significantly higher lift and becomes the new energy target. During the last 100 s, further exploration is possible due to the increased energy capture from the thermal.

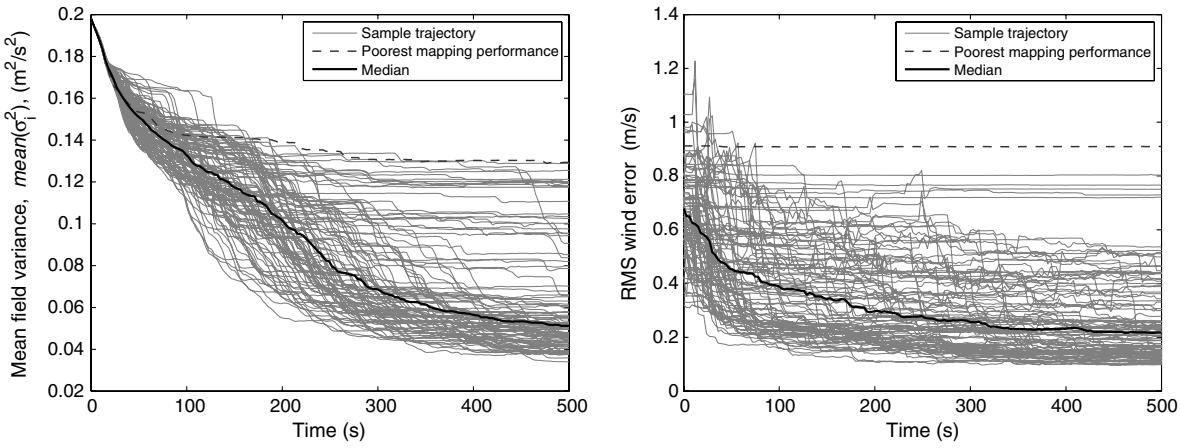
These results demonstrate the limitations of the current method. In the final scenario, the aircraft does not carry out the mission particularly efficiently. This is largely due to the nature of the target assignment, where the exploration target is defined as the single point that will yield the most information but is not necessarily the most useful one in terms of continuous field exploration. Similarly, the energy target is initially outside the target exploration region due to a noisy sample that yielded the highest estimated lift there, despite there being equal energy regions inside the exploration box. In the illustrated scenario, the planner selects a high-energy target at the top corner of the region early in the exploration, requiring significant energy capture from the relatively limited energy available from the wave flow. Collecting sufficient energy to travel to the target takes a long time, with very little exploration performed. Both the energy target and information target are single points, and their selection can be biased by poor sampling or noisy observations, leading to a reduction in exploration performance. The resulting wind map at $t = 500$ s (Fig. 17) has unexplored regions and significantly higher variance than in the previous cases.

To demonstrate the robustness of the method, a set of 100 randomized trials was performed. Each trial is a 500 s simulation for a flight through a random thermal field. We define the thermal volumetric flow rate Q_{thermal} as the volume of air moving upward through a horizontal plane through the center of a thermal. The volumetric flow is dependent on the core vertical speed V_{core} and lifting radius r_{lift}, as shown in Eq. (24). Note that the thermal model used is conservative, so there is an equal volume of air moving downward through the same plane in the sinking air region outside the thermal core:

$$Q_{\text{thermal}} = \tfrac{4}{\pi} r_{\text{lift}}^2 V_{\text{core}} \tag{24}$$

For the randomized simulations, there are between two and four thermal bubbles placed randomly in the exploration region, with a total volumetric flow equivalent to a single thermal bubble of 2.5 m/s vertical wind and 100 m lifting radius (total $Q_{\text{thermal}} = 3.18 \times 10^4$ m^3/s). The core speeds range from 0 to 7 m/s, and lifting radii range from 40 to 100 m. The resulting time history of the mapping variance estimate and mapping accuracy are shown in Fig. 18. The accuracy is the rms map error taken by comparing the wind estimate with the actual wind field at each time step over the regular $8 \times 8 \times 5$ grid used for target assignment. The median field variance at $t = 500$ s across 100 trials was 0.051 $(m/s)^2$, and the median rms error was 0.218 m/s.

The simulation was completed in all cases such that the planner never had to cancel the mission due to lack of energy. While most trials significantly reduced the rms map error, some did not complete the mapping task adequately. The poorest performing case (shown by the dashed line in Fig. 18) consisted of two thermals with core speeds of 1.05 and 6.91 m/s and lifting radii of 54.7 and 64.1 m,



a) Mean variance estimate **b) RMS mapping error**

Fig. 18 Variance estimate and mapping error over 100 simulated flights of 500 s each through random thermal fields.

respectively. The weak thermal was near the presampled path and, consequently, already relatively well mapped at the start of autonomous flight. After initial exploration, the aircraft spent the remaining flight attempting to collect enough energy from the known small thermal to continue exploration. However, due to the low lift and small radius, the aircraft gained energy very slowly and often lost energy when it left the lifting core and made turns in the surrounding sinking air attempting to return to lifting air. This meant that there was limited improvement in mapping accuracy due to mapping error being dominated by the unmapped second thermal. This demonstrates the conservative approach of the current control scheme, given that no prior information on the wind field is available. An improved target assignment algorithm may need to be more flexible in selecting exploration targets to prevent extended attempts at using weak energy sources to reach a high-energy exploration target.

VII. Conclusions

This paper presented a strategy for simultaneous exploration and exploitation of a wind field for a gliding UAV. A novel application of GP regression is used to build a model-free spatial wind map from point observations of the wind taken during the flight. The resulting map provides a continuous estimate of the wind field velocity as well as an estimate of the uncertainty (variance) across the map. A global planner facilitates exploration of the map by defining a global target each planning cycle. The global target is selected based on the current energy of the aircraft and is either an exploration target (the position with maximum estimated variance) or an exploitation target (the position at which maximum power gain was recorded). A low-level planner uses an energy-based heuristic to identify control actions that maximize local energy capture, travel toward the global goal, and reduce local map uncertainty. The system was tested in simulation for static wind fields containing thermal bubbles. The results demonstrate robust simultaneous exploration and exploitation of an unknown wind field, with the glider model effectively exploring a fixed region while using energy sources found in the field.

Appendix: Gaussian Process Regression

GP regression is a nonparametric Bayesian machine learning method of regression analysis. The following provides a brief explanation for GP regression used in the current context. The notation used is generally consistent with that used by Rasmussen and Williams [19], and further information can be found there. As in most regression techniques, the goal is to characterize the underlying function from a finite set of observations $\mathbf{y} = \{y_i\}_{i=1}^n$, where $y_i \in \mathbb{R}$ taken at locations in the input space $X = \{\mathbf{x_i}\}_{i=1}^n$, where $\mathbf{x_i} \in \mathbb{R}^d$. The observations are assumed to be drawn from the set of actual function values $f(\mathbf{x})$ with additive zero-mean Gaussian noise with variance σ_n^2:

$$y = f(\mathbf{x}) + \epsilon \tag{A1}$$

$$\epsilon \sim \mathcal{N}(0, \sigma_n^2) \tag{A2}$$

The goal of the GP regression analysis is to learn a function f, which is the underlying function drawn from the input data. However, GP regression does not return an analytical function but rather returns a distribution of functions from which we can draw estimates of the function value at any selected test point $\mathbf{x}_*$.

The covariance function $k(\mathbf{x}, \mathbf{x}')$ defines the variance between two input points. A covariance function must be a positive semidefinite function. There is a large range of common covariance functions that determine the properties (such as smoothness, periodicity, stationarity, etc.) of the function set generated by the GP. A common set of covariance functions are radial basis functions, which are based only on the distance between the two input points. In the case of a Euclidean space and distance measurement, this is also often an isotropic distance (input dimensions have equal weight) such that $k(\mathbf{x}, \mathbf{x}')$ is a function of r, where $r = |\mathbf{x} - \mathbf{x}'|$. In this case, the covariance is high (they are strongly correlated) for points close together and low (weakly correlated) for points far apart. The covariance function is used to generate a positive semidefinite covariance matrix $K(X, X)$, which represents the covariance relationship between all points in the input space. To include observation noise, as specified in Eq. (A1), the final covariance is

$$\text{cov}\,(X, X) = K(X, X) + \sigma_n^2 I \tag{A3}$$

GP regression uses the information expressed in the covariance matrix to generate a Gaussian distribution of functions. Any point on the function space can be tested, and the return is a Gaussian distribution output at that point defined by a mean value and covariance. The mean value at a sample location is actually a weighted sum of the training points, with the weightings determined via the covariance function. Without covering the mathematics in depth, the estimated mean value $\bar{f}_*$ and covariance $\text{cov}(f_*)$ for a set of test points X_*, training points X, observations Y, and covariance function k are shown as follows:

$$\bar{f}_* = \mathbb{E}[f(X_*)|X, y, X_*] = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}\mathbf{y} \tag{A4}$$

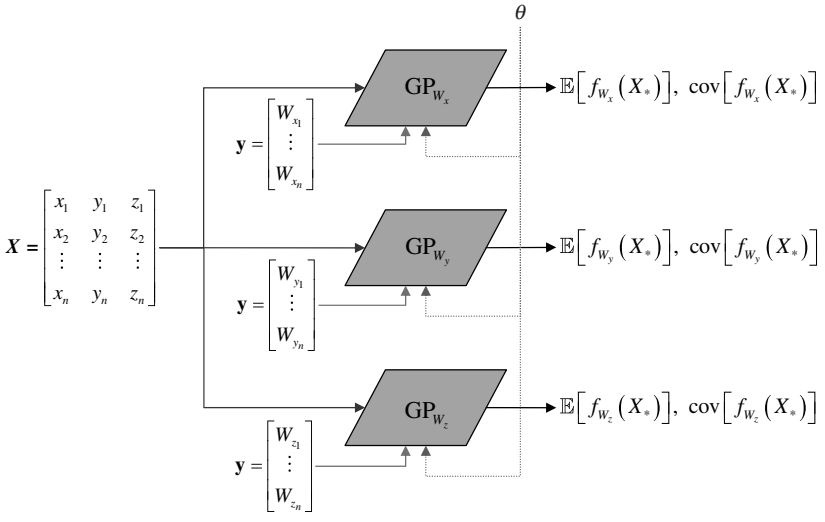


Fig. A1 Structure of GP regression over Cartesian coordinates of wind.

$$\text{cov}(f_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} K(X, X_*) \tag{A5}$$

Selection of an appropriate covariance function is critical, as it determines the set of basis functions from which the solution set is drawn. In this paper, we use the squared exponential covariance [Eq. (A6)], a radial basis function. It is a commonly used covariance function that generates smooth outputs and, in this case, has zero mean, is isotropic (input dimensions are equally weighted), stationary (function only of relative displacement, not of the absolute values of the input points), and nondegenerate (infinitely differentiable):

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(\frac{-|\mathbf{x} - \mathbf{x}'|^2}{2l^2}\right) \tag{A6}$$

As can be seen from the equation, there are a number of variable parameters that affect the values returned by the covariance function. These are known as hyperparameters θ, and training the model is the process of learning the most appropriate set of hyperparameters from the training data. In the case of a squared exponential covariance function, the set of hyperparameters is the length scale l, the signal variance σ_f^2, and the sensor noise variance σ_n^2.

Selection of the hyperparameters affects the output mean and variance estimates. Since the effect of the hyperparameters on the resulting estimate is not always intuitive, we require some measure of determining whether the selected hyperparameters result in a good fit. A common method of estimating the fitness of a candidate set of hyperparameters is the marginal likelihood. In a Bayesian model comparison, the marginal likelihood represents the probability of obtaining the training observations y, given the training points X, the set of hyperparameters θ, and the current model M. The weights $\mathbf{w}$ are marginalized out to give the marginal likelihood:

$$p(\mathbf{y} \mid X, \theta, M) = \int p(\mathbf{y} \mid X, \mathbf{w}, M) p(\mathbf{w} \mid \theta, M)\, d\mathbf{w} \tag{A7}$$

This integral can be evaluated over a Gaussian model set, meaning that (the log of) the marginal likelihood is relatively easily to calculate. Although an in-depth discussion of the marginal likelihood is beyond the scope of this work, it is sufficient to say that maximizing the marginal likelihood represents a fit based on a balance of minimum model complexity and remaining near the training data. Very complex models tend to fit the data well near the training points but are poor predictors away from the training data (overfitting), whereas simple models may not be capable of capturing the information in the training data and consequently assume a simple underlying function and very noisy observations (underfitting). In the current application, the GP is trained using an optimization routine (in this case, MATLAB's fminunc or a simple gradient descent function), which varies the hyperparameters to minimize the negative log of the marginal likelihood. The hyperparameters are learned offline from a number of sample trajectories in similar conditions. The trained model is then used for the simulated test flight, and the hyperparameters remain fixed during the simulation. In the current implementation, the hyperparameters are shared across the three input dimensions such that only one inverse covariance matrix is required for estimation at target points. The structure is illustrated in Fig. A1.

Acknowledgments

This work is supported by the Australian Research Council (ARC) Centre of Excellence programme, funded by the ARC, the New South Wales State Government, and the Australian Centre for Field Robotics.

References

[1] Rayleigh, L., "The Soaring of Birds," *Nature*, Vol. 27, No. 701, 1883, pp. 534–535. doi:10.1038/027534a0

[2] Tucker, V. A., and Parrott, G. C., "Aerodynamics of Gliding Flight in a Falcon and Other Birds," *Journal of Experimental Biology*, Vol. 52, No. 2, 1970, pp. 345–367.

[3] Pennycuick, C., "Soaring Behaviour and Performance of Some East African Birds, Observed From a Motor-Glider," *Ibis*, Vol. 114, No. 2, 1972, pp. 178–218. doi:10.1111/j.1474-919X.1972.tb02603.x

[4] Weimerskirch, H., Guionnet, T., Martin, J., Shaffer, S. A., and Costa, D. P., "Fast and Fuel Efficient? Optimal Use of Wind by Flying Albatrosses," *Proceedings of the Royal Society of London, Series B: Biological Sciences*, Vol. 267, No. 1455, 2000, pp. 1869–1874. doi:10.1098/rspb.2000.1223

[5] Zhao, Y. J., "Optimal Patterns of Glider Dynamic Soaring," *Optimal Control Applications and Methods*, Vol. 25, No. 2, 2004, pp. 67–89. doi:10.1002/oca.739

[6] Zhao, Y. J., and Qi, Y. C., "Minimum Fuel Powered Dynamic Soaring of Unmanned Aerial Vehicles Utilizing Wind Gradients," *Optimal Control Applications and Methods*, Vol. 25, No. 5, 2004, pp. 211–233. doi:10.1002/oca.744

[7] Sachs, G., "Minimum Shear Wind Strength Required for Dynamic Soaring of Albatrosses," *Ibis*, Vol. 147, No. 1, 2005, pp. 1–10. doi:10.1111/j.1474-919x.2004.00295.x

[8] Chakrabarty, A., and Langelaan, J. W., "Energy Maps for Long-Range Path Planning for Small- and Micro- UAVs," AIAA Guidance, Navigation and Control Conference, Chicago, IL, AIAA Paper 2009-6113, 2009.

[9] Lawrance, N. R., and Sukkarieh, S., "A Guidance and Control Strategy for Dynamic Soaring with a Gliding UAV," *IEEE International Conference on Robotics and Automation*, Kobe, Japan, IEEE Publ., Piscataway, NJ, 2009, pp. 3632–3637.

[10] Lissaman, P., "Wind Energy Extraction by Birds and Flight Vehicles," 43rd AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, AIAA Paper 2005-0241, 2005.

[11] Lawrance, N. R., and Sukkarieh, S., "Wind Energy Based Path Planning for a Small Gliding Unmanned Aerial Vehicle," AIAA Guidance, Navigation and Control Conference, Chicago, IL, AIAA Paper 2009-6112, 2009.

[12] Langelaan, J. W., "Gust Energy Extraction for Mini and Micro Uninhabited Aerial Vehicles," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 2, 2009, pp. 464–473. doi:10.2514/1.37735

[13] Allen, M. J., "Autonomous Soaring for Improved Endurance of a Small Uninhabited Air Vehicle," 43rd AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, AIAA Paper 2005-1025, 2005.

[14] Edwards, D. J., and Silverberg, L. M., "Autonomous Soaring: The Montague Cross-Country Challenge," *Journal of Aircraft*, Vol. 47, No. 5, 2010, pp. 1763–1769. doi:10.2514/1.C000287

[15] Cutler, M. J., McLain, T. W., Beard, R. W., and Capozzi, B., "Energy Harvesting and Mission Effectiveness for Small Unmanned Aircraft," AIAA Guidance, Navigation and Control Conference, Toronto, AIAA Paper 2010-8037, 2010.

[16] Bower, G. C., Flanzer, T. C., Naiman, A. D., and Saripalli, S., "Dynamic Environment Mapping for Autonomous Thermal Soaring," AIAA Guidance, Navigation and Control Conference, Toronto, AIAA Paper 2010-8031, 2010.

[17] Langelaan, J. W., Alley, N., and Neidhoefer, J., "Wind Field Estimation for Small Unmanned Aerial Vehicles," AIAA Guidance, Navigation and Control Conference, Toronto, AIAA Paper 2010-8177, 2010.

[18] Anderson, J. D., Jr., *Fundamentals of Aerodynamics*, 4th ed., Series in Aeronautical and Aerospace Engineering, McGraw–Hill, New York, 2007, pp. 404–430.

[19] Rasmussen, C. E., and Williams, C. K., "Gaussian Processes for Machine Learning," *Adaptive Computation and Machine Learning*, MIT Press, Cambridge, MA, 2006, pp. 7–29.

[20] Handcock, M. S., and Wallis, J. R., "An Approach to Statistical Spatial-Temporal Modeling of Meteorological Fields," *Journal of the American Statistical Association*, Vol. 89, No. 426, 1994, pp. 368–378. doi:10.2307/2290832

[21] Alexiadis, M., Dokopoulos, P., and Sahsamanoglou, H., "Wind Speed and Power Forecasting Based on Spatial Correlation Models," *IEEE Transactions on Energy Conversion*, Vol. 14, No. 3, 1999, pp. 836–842.

doi:10.1109/60.790962
[22] Damousis, I., Alexiadis, M., Theocharis, J., and Dokopoulos, P., "A Fuzzy Model for Wind Speed Prediction and Power Generation in Wind Parks Using Spatial Correlation," *IEEE Transactions on Energy Conversion*, Vol. 19, No. 2, 2004, pp. 352–361. doi:10.1109/TEC.2003.821865
[23] Smola, A. J., and Bartlett, P., "Sparse Greedy Gaussian Process Regression," *Advances in Neural Information Processing Systems 13*, MIT Press, Cambridge, MA, 2001, pp. 619–625.
[24] Csató, L., and Opper, M., "Sparse On-Line Gaussian Processes," *Neural Computation*, Vol. 14, No. 3, 2002, pp. 641–668. doi:10.1162/089976602317250933
[25] "Military Specification: Flying Qualities of Piloted Airplanes," U.S. Department of Defense, MIL-F-8785C, Nov. 1980.
[26] Gage, S., "Creating a Unified Graphical Wind Turbulence Model from Multiple Specifications," AIAA Modeling and Simulation Technologies Conference and Exhibit, Austin, TX, AIAA Paper 2003-5529, 2003.