

Recuperación Óptima de Stall en Aviones Usando Aprendizaje Reforzado

Agustin Grillo, Gabriel Torre y Roberto A. Bunge
LINAR, Departamento de Ingeniería, Universidad de San Andrés
torreg@udesa.edu.ar



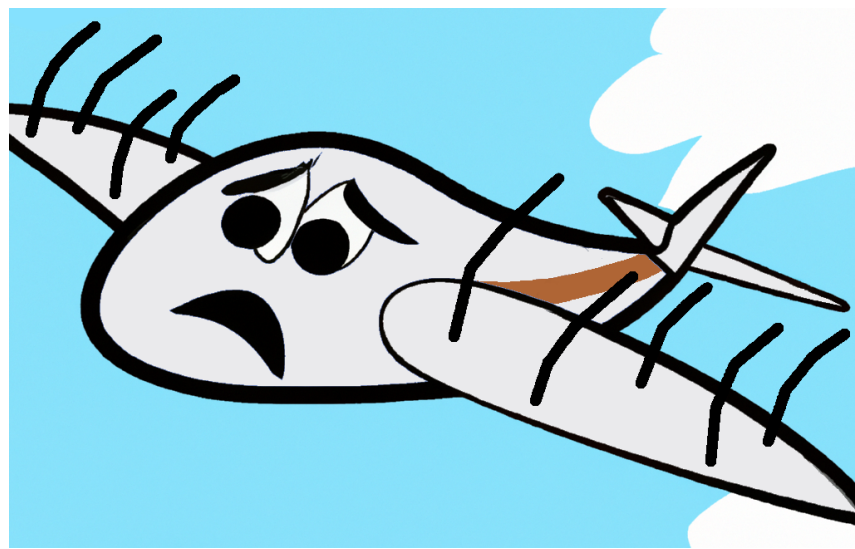
Motivación

El problema de recuperación de pérdida de sustentación en aviones es de suma importancia para la seguridad en vuelo.

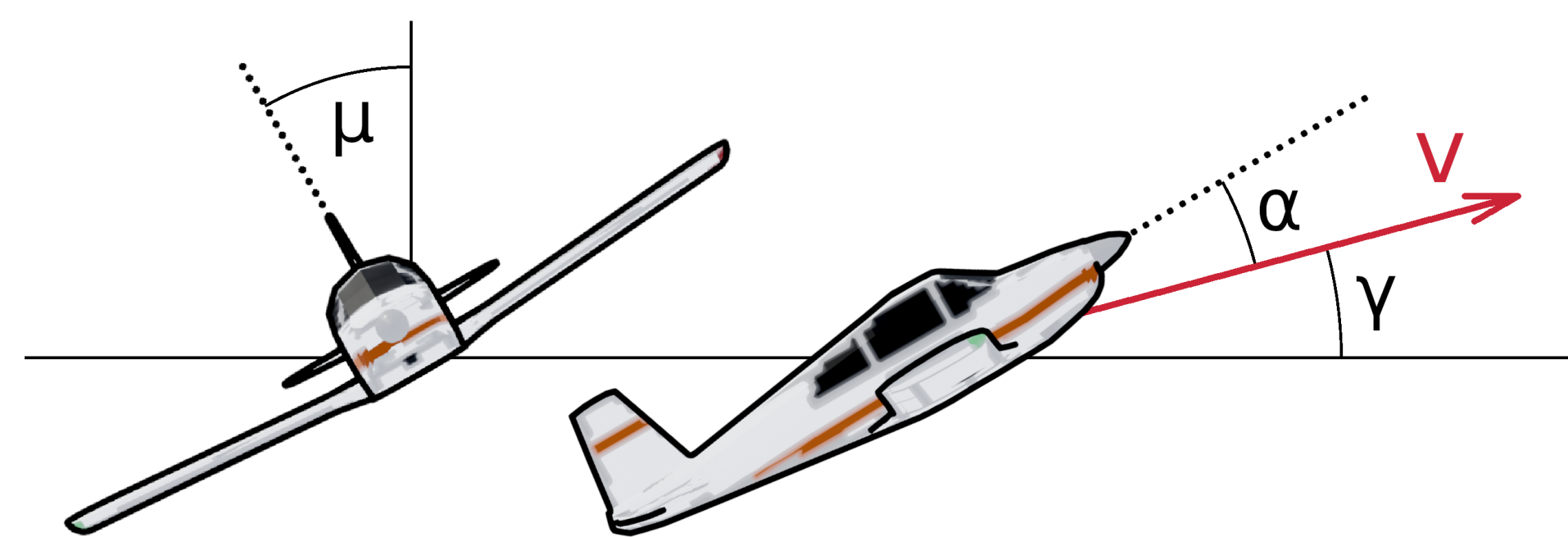
Buscamos las acciones clave que los pilotos deben tomar durante una pérdida de sustentación para minimizar la pérdida de altitud y restaurar un vuelo seguro.

¿Qué es el Stall en Aviones?

La entrada en pérdida, conocida como *stall* en inglés, ocurre cuando la sustentación de un perfil aerodinámico cae abruptamente, lo que puede provocar una peligrosa caída del avión, similar a perder el control de un auto en una ruta resbaladiza.



Sistema de Referencia



p : tasa de alabeo, *roll rate*
 q : tasa de cabeceo, *pitch rate*

¿Y el Spin?

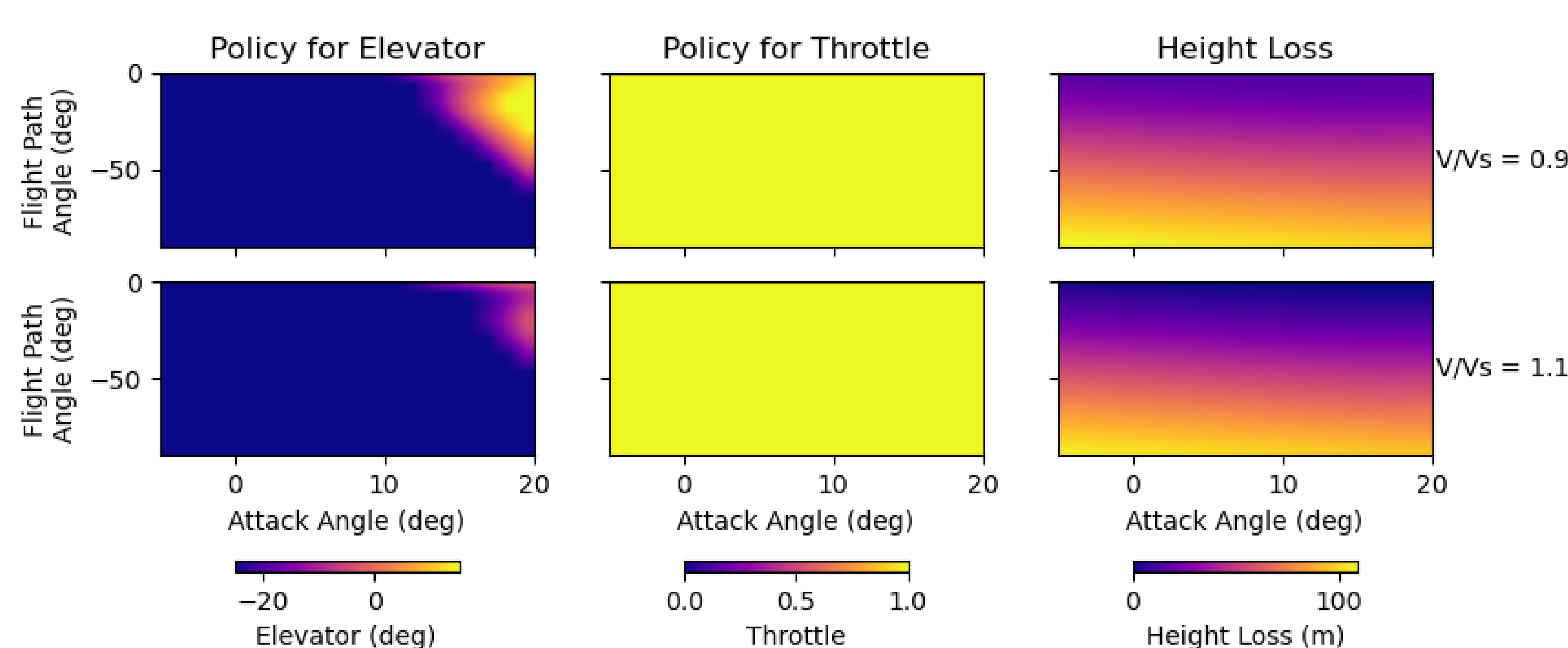
La barrena o *spin*, es un *stall* prolongado, en el cual el avión cae con la nariz baja, girando alrededor de su eje vertical, similar a un sacacorchos. Este tipo de maniobras representa una porción considerable de los accidentes fatales en avionetas.

Soluciones Tradicionales

La FAA recomienda una recuperación específica para situaciones de *spin*, considerando el estrés del piloto y entonces limitando cambios en los controles. Sin embargo, un algoritmo no tiene estas restricciones y puede aprovechar de manera óptima todos los controles disponibles.

Resultados

Analizamos varios tipos de stall, como la recuperación de un stall simétrico con motor, donde la política aprovecha la sustentación adicional en altos ángulos de ataque para minimizar la pérdida de altura.



Políticas para la recuperación de un stall simétrico con motor para un $q = 0$ deg/s

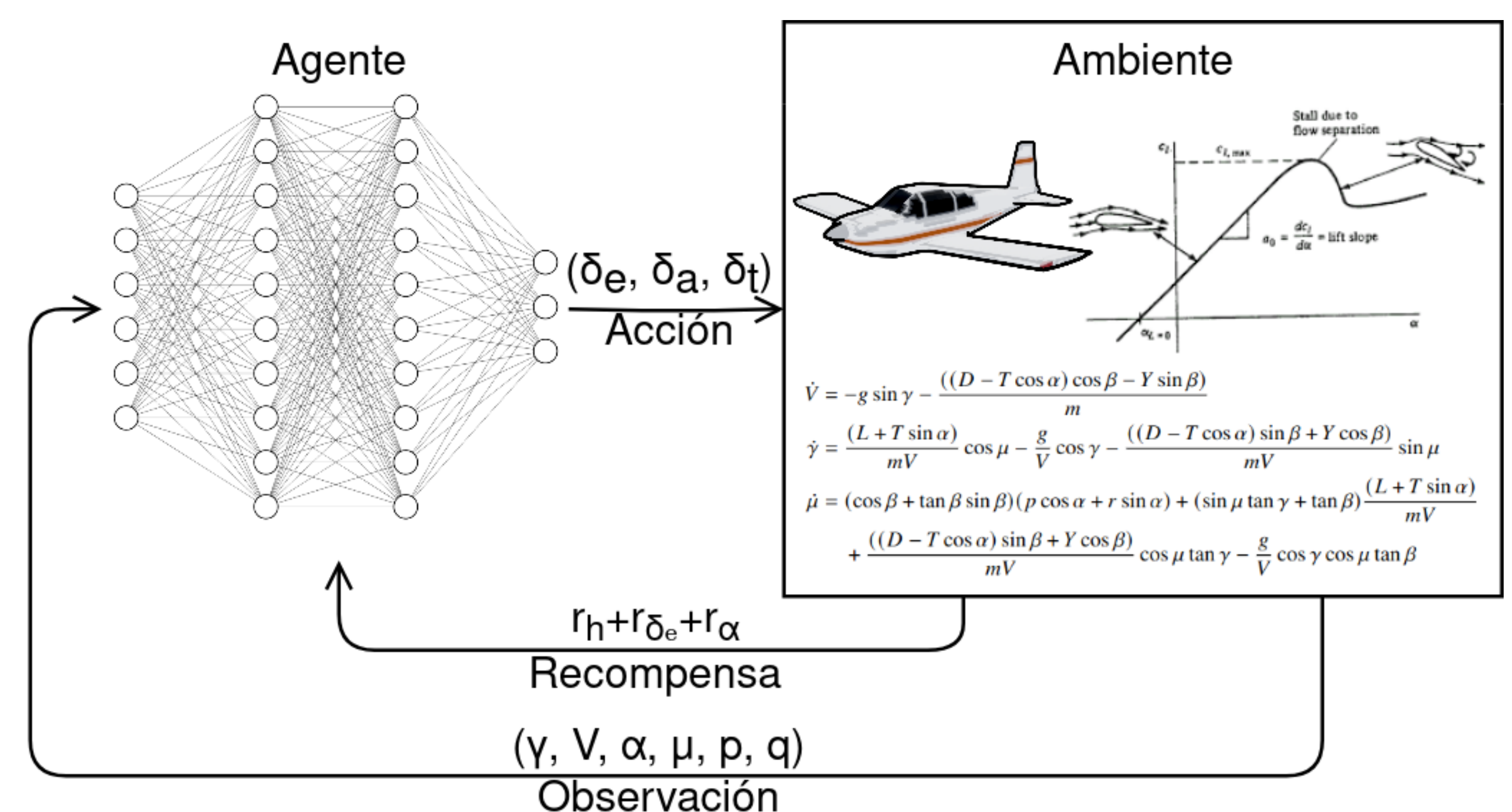
Método

El problema a resolver es minimizar la altura perdida, podemos expresarlo como:

$$h^*(s_0) = \left\{ \min_{a(t), t_f} \int_0^{t_f} -V \sin \gamma dt \mid \gamma(t_f) = 0, \dot{s} = f(s, a), s(0) = s_0 \right\}$$

Aprendizaje Reforzado

Usamos aprendizaje reforzado debido a su escalabilidad en problemas de alta dimensionalidad, superando a otros enfoques previos en ese aspecto. Empleamos el algoritmo PPO, ampliamente aceptado como estándar de facto por su eficacia en control continuo.



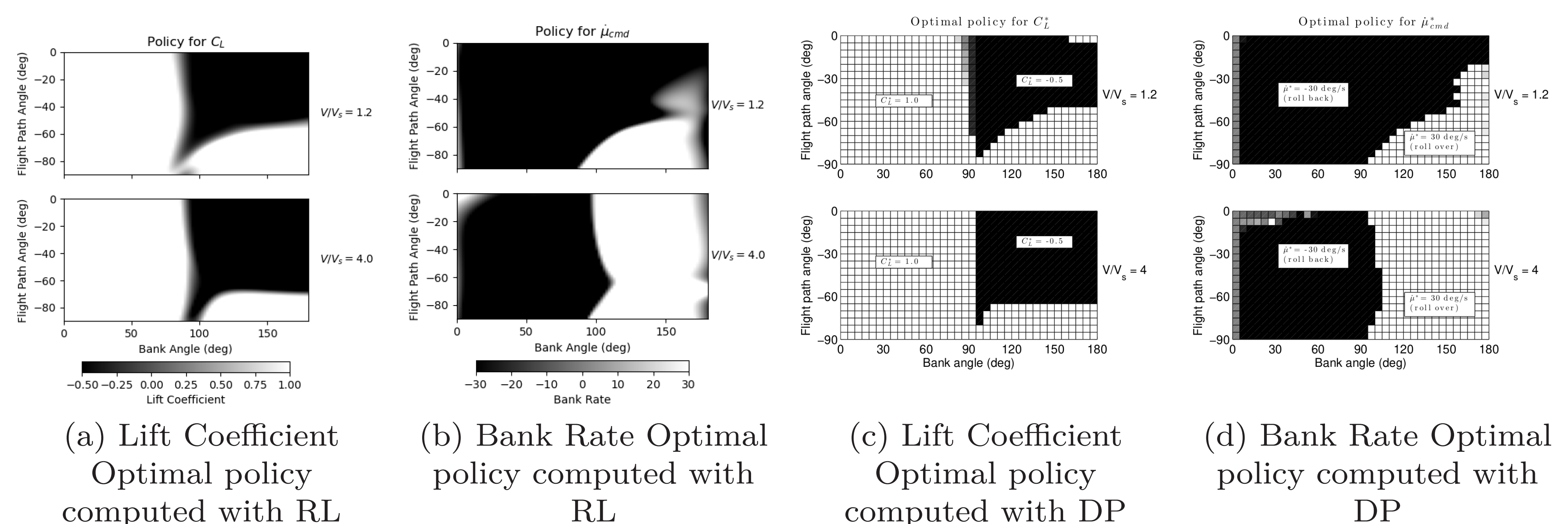
Esquema del entorno de entrenamiento de la policy

Función de Recompensa

La función de recompensa incluye señales adicionales al objetivo final de optimización para evitar problemas como *reward hacking* y acelerar la convergencia.

Validación

Para validar la optimalidad de nuestro método, analizamos un sub-problema resuelto de manera exacta con programación dinámica debido a su baja dimensionalidad.



Comparación de nuestros resultados (a) y (b) con los de [Bunge et al. 2018] (c) y (d). Destaca la estabilidad numérica en los bordes a diferencia de métodos de grillado.

Discusión

- Controlar la velocidad del motor (δ_t) durante la recuperación reduce la pérdida de altitud.
- El algoritmo PPO muestra eficacia en determinar acciones óptimas para reducir la pérdida de altitud.
- El estudio sugiere que el algoritmo podría extenderse a problemas más complejos como el de recuperación de *spin*.