

## Quadrotor simulation and trajectory tracking with wind-effect compensation using reinforcement learning

Zahra Darestani Farahani, Majid Sadedel\* 

Mechanical Engineering Department, Tarbiat Modares University, Tehran, Iran

### ARTICLE INFO

#### Article Type

Original Research

#### Article History

Received: December 24, 2025

Revised: February 20, 2026

Accepted: April 08, 2026

ePublished: May 15, 2026

### ABSTRACT

Today, with the increasing use of quadrotors in various missions, employing an efficient and effective control system has become essential. The aim of this research is to design an intelligent controller capable of achieving accurate trajectory tracking of a quadrotor in environments with varying wind conditions. To this end, the dynamic modeling of the quadrotor was first carried out. Wind modeling was also implemented by adding horizontal wind acceleration to the translational dynamic equations of the vehicle.

Next, to ensure precise position and attitude control of the quadrotor in the presence of wind disturbances, a hybrid control framework was designed, consisting of a baseline proportional–integral–derivative (PID) controller, a reinforcement learning–based gain tuner, and a disturbance observer along with its compensator. For adaptive tuning of the PID gains during trajectory tracking, the DDPG and TD3 reinforcement learning algorithms were utilized. To estimate wind intensity, an inverse dynamic model was employed as the observer. Furthermore, a feedforward compensation algorithm was developed to adjust the reference roll and pitch angles. By correcting the roll and pitch angles, the compensator also reduced position deviations along the x and y axes.

Finally, to evaluate the performance of the control framework, various experiments were conducted under no-wind conditions and under different wind intensities across multiple trajectories. The simulation results indicated that in environments with varying wind, adding the observer and compensator to a fixed-gain PID controller reduced the tracking error by 15% compared to the standalone PID controller. Additionally, PID control with reinforcement learning–based gain tuning, combined with the observer and compensator, reduced tracking error by 25% compared to the fixed-gain case. In the lemniscate trajectory, the DDPG algorithm performed 10% better than TD3; in the circular trajectory, the TD3 algorithm performed 5% better than DDPG; and in spiral trajectories, DDPG outperformed TD3 by 20%.

**Keywords:** Quadrotor, trajectory tracking, intelligent control, wind-effect compensation, reinforcement learning

### How to cite this article

Darestani Farahani Z, Sadedel M, Quadrotor simulation and trajectory tracking with wind-effect compensation using reinforcement learning. *Modares Mechanical Engineering*; 2026;26(07):547-558.

\*Corresponding author's email: [majid.sadedel@modares.ac.ir](mailto:majid.sadedel@modares.ac.ir)

\*Corresponding ORCID ID: 0000-0002-0285-8460



Copyright© 2025, TMU Press. This open-access article is published under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License which permits Share (copy and redistribute the material in any medium or format) and Adapt (remix, transform, and build upon the material) under the Attribution-NonCommercial terms.



## شبیه‌سازی و ردیابی مسیر کوادروتور با جبران‌سازی اثر باد به کمک یادگیری تقویتی

زهره دارستانی فراهانی، مجید ساده‌دل\*

گروه مهندسی مکانیک، دانشگاه تربیت مدرس، تهران، ایران

### چکیده

امروزه با توجه به افزایش روزافزون بکارگیری کوادروتورها در ماموریت‌های مختلف، بهره‌گیری از سیستم کنترلی کارآمد و مؤثر ضروری می‌باشد. هدف این پژوهش، طراحی کنترل‌کننده هوشمند برای دستیابی به ردیابی مسیرهای مختلف کوادروتور در محیط‌هایی با بادهای متغیر می‌باشد. به این منظور ابتدا به مدل‌سازی دینامیکی کوادروتور پرداخته شد. مدل‌سازی باد نیز با اضافه کردن شتاب باد در صفحه‌ی افقی به معادلات دینامیکی انتقالی پرنده انجام شد. سپس، با هدف کنترل دقیق موقعیت و جهت‌گیری کوادروتور در حضور اغتشاشات محیطی باد، چارچوبی ترکیبی طراحی شد که شامل یک کنترل‌کننده‌ی پایه تناسبی-انترگالی-مشتقی، یک تنظیم‌کننده‌ی ضرایب مبتنی بر یادگیری تقویتی و یک مشاهده‌گر اغتشاش به همراه جبران‌ساز آن بود. برای تنظیم تطبیقی ضرایب کنترل‌کننده PID حین ردیابی مسیر از الگوریتم‌های DDPG و TD3 یادگیری تقویتی بهره برده شد. جهت تخمین شدت باد از یک مدل دینامیک معکوس به عنوان مشاهده‌گر استفاده شد. همچنین یک الگوریتم جبران‌سازی پیش‌خور برای اصلاح زوایای مرجع رول و پیچ طراحی شد. جبران‌ساز با اصلاح زوایای رول و پیچ، انحراف موقعیت را نیز در جهت محورهای x و y جبران می‌کرد. در نهایت جهت ارزیابی عملکرد چارچوب کنترلی، آزمایش‌های مختلفی در شرایط بدون باد و با شدت‌های مختلف باد در مسیرهای مختلف پروازی انجام شد. نتایج شبیه‌سازی‌های انجام گرفته نشان دادند که در محیط‌هایی با بادهای متغیر، اضافه شدن مشاهده‌گر و جبران‌ساز به کنترل‌کننده PID با ضرایب ثابت ۱۵٪ خطا را نسبت به PID، و کنترل‌کننده PID تنظیم شده با الگوریتم‌های یادگیری تقویتی همراه با مشاهده‌گر و جبران‌ساز نسبت به حالت ضریب ثابت ۲۵٪ خطای ردیابی را کاهش داد. در مسیر بی‌نهایت الگوریتم DDPG، ۱۰٪ بهینه‌تر از الگوریتم TD3، در مسیر دایروی الگوریتم TD3 ۵٪ بهینه‌تر از الگوریتم DDPG، و در مسیرهای مارپیچ نیز الگوریتم DDPG، ۲۰٪ بهینه‌تر از الگوریتم TD3 عمل کرد.

### اطلاعات مقاله

#### نوع مقاله

مقاله پژوهشی

#### تاریخچه مقاله

دریافت: ۱۴۰۴/۱۰/۰۳

بازنگری: ۱۴۰۴/۱۲/۰۱

پذیرش: ۱۴۰۵/۰۱/۱۹

ارائه آنلاین: ۱۴۰۵/۰۲/۲۵

**کلیدواژه‌ها:** کوادروتور، ردیابی مسیر، کنترل هوشمند، جبران‌ساز اثر باد، یادگیری تقویتی

### نحوه ارجاع به این مقاله

دارستانی فراهانی زهرا، ساده‌دل مجید، شبیه‌سازی و ردیابی مسیر کوادروتور با جبران‌سازی اثر باد به کمک یادگیری تقویتی، مهندسی مکانیک مدرس. ۰۵۱-۰۵۴۷-۰۲۶(۰۷):۱۴۰۵

\*پست الکترونیکی نویسنده عهده‌دار مکاتبات: majid.sadedel@modares.ac.ir

\*شناسه ارکید نویسنده عهده‌دار مکاتبات: 0000-0002-0285-8460



پیشخور جهت پیش پردازش سیستم کنترل حرکت پهپادها و دو زیرسیستم ردیابی حالت‌ها برای حلقه موقعیت و جهت‌گیری توسعه دادند.

در [۱۰] پژوهشگران برای کنترل کوادروتور شش زیرسیستم مستقل تعریف کردند که هر کدام توسط یک عامل جداگانه کنترل می‌شد و به صورت ترتیبی آموزش داده می‌شدند. در [۱۱]، [۱۲]، [۱۳]، [۱۴]، [۱۵] نیز پژوهشگران به طراحی روش‌های هوشمند به وسیله الگوریتم‌های یادگیری تقویتی برای تنظیم و تخمین پارامترهای کنترل‌کننده PID در کوادروتورها پرداختند.

در [۱۶]، یک ساختار کنترلی ترکیبی که در آن کنترل‌کننده ارتفاع شامل یک MPC خطی به همراه جبران‌ساز گرانشی تطبیقی و کنترل‌کننده جهت‌گیری متشکل از دو حلقه کنترل PID برای موقعیت  $x$  و  $y$  و حلقه درونی شامل سه کنترل‌کننده PID با ضرایب تنظیم شده با الگوریتم DDPG برای زاویه‌های اوایلر بود.

در [۱۷]، چارچوبی با عنوان RLC-PID معرفی شده بود که در آن یک عامل یادگیری تقویتی عمیق از نوع DDPG به صورت برخط پارامترهای کنترل‌کننده PID را برای کنترل سرعت‌های عرضی تنظیم می‌کند، در حالی که سرعت فرود در محور  $z$  ثابت نگه داشته شده بود. همچنین در [۱۸]، مسئله‌ی اصلی که پژوهشگران بررسی کردند، چالش تنظیم دقیق پارامترهای کنترل‌کننده جهت‌گیری در شرایط واقعی پرواز بود. در این پژوهش از الگوریتم DDPG برای تنظیم پنج ضریب مرتبط با کنترل‌کننده داخلی PD در حلقه نگرش استفاده شده است. در [۱۹]، با بکارگیری روش مد لغزشی تطبیقی و در [۲۰]، با استفاده از کنترل‌کننده تناسبی-انتگرالی-مشتقی فازی مرتبه کسری مسئله ردیابی مسیر را در حضور اغتشاشات باد و عدم قطعیت‌های سیستم بررسی کردند.

## ۲-۱- نوآوری پژوهش

نوآوری این پژوهش در طراحی یک چارچوب کنترلی جامع است که دو جزء مکمل را در کنار یکدیگر به کار می‌گیرد. تنظیم تطبیقی ضرایب کنترل‌کننده PID با استفاده از الگوریتم‌های یادگیری تقویتی در کنار طراحی مشاهده‌گر اغتشاش و جبران‌ساز برای مقابله با اغتشاشات باد نوآوری این پژوهش است.

## ۲-۲- مدل‌سازی دینامیکی

در این بخش به مدل‌سازی دینامیکی یک کوادروتور با استفاده از معادلات نیوتن-اوایلر پرداخته می‌شود. کوادروتور دارای ۶ درجه آزادی و تنها ۴ ورودی کنترلی است. به دلیل ویژگی‌هایی چون غیرخطی بودن، ناپایداری ذاتی و کم‌عملگر بودن این ربات، مدل‌سازی دقیق آن اهمیت بالایی در طراحی کنترل‌کننده دارد. برای ساده‌سازی، چندین فرض فیزیکی در نظر گرفته شده و معادلات به دو دسته معادلات انتقالی و دورانی تقسیم می‌شوند.

برای مدل‌سازی دقیق ولی ساده، فرضیات زیر اعمال شده‌اند:

۱- کوادروتور به صورت صلب و با ویژگی‌های دینامیکی مشخص تعریف خواهد شد. (جرم، ابعاد...)

۲- ساختار کوادروتور نسبت به محورهای  $x$  و  $y$  متقارن فرض شده است.

۳- مرکز جرم با مبدا چارچوب مرجع بدنه هم‌مکان است.

۴- شرایط اولیه پرواز (موقعیت و سرعت) به طور دقیق مشخص هستند.

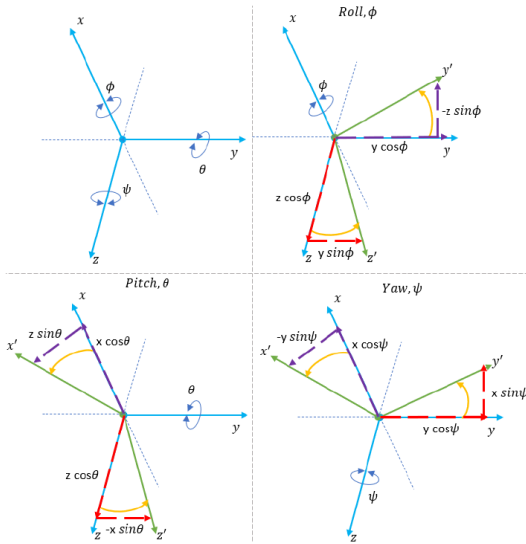
## ۱- مقدمه

در سال‌های اخیر استفاده از کوادروتورها توسط صنایع تجاری، موسسات تحقیقاتی و نظامی جهت نظارت برآتش سوزی جنگل‌ها [۱]، عملیات جستجو و نجات [۲]، بازرسی خطوط انتقال [۳]، حمل و نقل محموله [۴]، و نقشه برداری [۵]، مورد توجه پژوهشگران قرار گرفته است. از دلایل این موضوع می‌توان به وزن سبک‌تر، ساختار ساده‌تر، قابلیت مانورپذیری بالاتر، هزینه‌های تعمیر و نگهداری کمتر، کاهش زمان عملیاتی و ایمنی نسبت به هواپیماها و جنگنده‌ها اشاره کرد. حال نظر بر انجام عملیات‌ها بدون دخالت انسان، ایده برنامه‌ریزی و ردیابی مسیر به صورت بلادرنگ به جای کنترل از راه دور به ذهن می‌رسد. در بسیاری از مأموریت‌ها مسیر پروازی کوادروتور به طور دقیق و از پیش تعیین شده مشخص می‌شود، به همین دلیل ردیابی صحیح این مسیر اهمیت ویژه‌ای دارد. کوادروتور در محیط‌های بیرونی با اغتشاشات خارجی مختلفی از جمله وزش ناگهانی باد مواجه می‌شود. بادهای متغیر محیطی منجر به انحراف از مسیر مرجع و کاهش پایداری کوادروتور می‌شوند. انحراف از مسیر نه تنها باعث کاهش دقت مأموریت می‌شود بلکه می‌تواند به افزایش مصرف انرژی، کاهش ایمنی و حتی شکست کامل عملیات بیانجامد. بر همین اساس هدف این پژوهش، بهره‌گیری از ترکیب کنترل‌کننده تناسبی-انتگرالی-مشتقی با الگوریتم‌های یادگیری تقویتی عمیق و یک مشاهده‌گر اغتشاش به همراه جبران‌ساز آن برای دستیابی به دقت بالاتر در ردیابی مسیر است. این رویکرد پرنده را قادر می‌سازد در شرایطی که اغتشاشات باد حضور دارند نیز عملکردی پایدار و قابل اعتماد داشته باشد. کنترل‌کننده تناسبی-انتگرالی-مشتقی به دلیل قدرت کنترلی و قابلیت اطمینان بالا پرکاربردترین کنترل‌کننده در صنعت است. در محیط‌های پویا با شدت‌های مختلف باد، به دلیل ثابت بودن ضرایب و سیگنال‌های کنترلی تولید شده توسط کنترل‌کننده تناسبی-انتگرالی-مشتقی، عملکرد سیستم کاهش می‌یابد. ما برای بهبود عملکرد سیستم در محیط‌های واقعی از دو جز مکمل یادگیری تقویتی برای تنظیم ضرایب کنترل‌کننده تناسبی-انتگرالی-مشتقی به صورت تطبیقی و مشاهده‌گر اغتشاش جهت تخمین نیروهای اغتشاش و جبران‌ساز برای مقابله با این نیروهای تخمین زده شده استفاده شده است. در این پژوهش با ترکیب سازوکارهای یادگیری تقویتی و جبران‌ساز اغتشاش، تلاش شده تا ضمن کاهش خطای ردیابی و افزایش پایداری پرواز، سازگاری کوادروتور را در برابر تغییرات ناگهانی شدت باد افزایش داده شود.

## ۱-۱- پژوهش‌های پیشین

در [۶] پژوهشگران مسئله ردیابی مسیر کوادروتور را در یک راکتور تحت آموزش یادگیری تقویتی مورد بررسی قرار دادند. همچنین تابع پاداشی برای دستیابی به مسیر مطلوب به صورت مجذور فاصله اقلیدسی بین موقعیت فعلی و موقعیت بعدی که در مسیر تعیین شده است طراحی شده بود. محمدرضا ایزدی و همکارش در پژوهش خود برای تقویت استحکام کوادروتورها در ردیابی مسیر یک مشاهده‌گر اغتشاش با بهره بالا که اختلالات را به صورت بلادرنگ تخمین می‌زند، ارائه کرده است [۷].

در [۸] پژوهشگران با بهره‌گیری از مدل‌سازی دینامیکی کوادروتور و نیز مدل اغتشاشات باد درآیدن، یک کنترل‌کننده آشناری که در حلقه بیرونی، یک کنترل‌کننده PID برای کنترل موقعیت و در حلقه درونی، یک کنترل‌کننده مد لغزشی تطبیقی برای کنترل جهت‌گیری به کار گرفته شد. در [۹] پژوهشگران با یادگیری تقویتی یک کنترل‌کننده



شکل ۲ چرخش‌های اولیه کوادروتور [۲۱]

Fig.2 Initial rotations of the Quadcopter [21]

طبق معادلات نیوتن برای حرکت انتقالی در معادله ۲ مطابق مرجع ۲۱ داریم:

$$m\ddot{\xi} = m \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = -mgEz + U_1 RTEz \quad (2)$$

که در آن  $m$  جرم کوادروتور است و  $g$  شتاب گرانشی زمین است که مقدار آن در راستای منفی  $z$  در نظر گرفته می‌شود، چراکه جهت  $z$  مثبت به سمت بالا تعریف شده است.  $U_1 = \sum F_i$  مجموع نیروی رانش تولیدی توسط چهار موتور کوادروتور است.

در نهایت این معادله در راستای محورهای  $x, y, z$  به شکل معادلات ۳ تا ۵ نوشته می‌شود:

$$\ddot{x} = \frac{U_1}{m} (\cos\phi \sin\theta \cos\psi + \sin\phi \sin\psi) \quad (3)$$

$$\ddot{y} = \frac{U_1}{m} (\cos\phi \sin\theta \sin\psi - \sin\phi \cos\psi) \quad (4)$$

$$\ddot{z} = \frac{U_1}{m} (\cos\phi \cos\theta) - g \quad (5)$$

مدل دینامیک دورانی با استفاده از معادله نیوتن اوایلر برای اجسام صلب به صورت معادله ۶ مطابق مرجع ۲۱ تعریف می‌شود:

$$I\dot{\eta} = -\eta \times I\dot{\eta} - J_r(\eta \times Ez)\Omega d + \begin{bmatrix} U_2 \\ U_3 \\ U_4 \end{bmatrix} \quad (6)$$

در این معادله  $I = \text{diag}(I_{xx}, I_{yy}, I_{zz})$  ماتریس ممان اینرسی حول سه محور است.  $J_r$  ممان اینرسی روتورها و  $\Omega_r$  سرعت زاویه‌ای تجمعی روتورها است و  $U_2, U_3, U_4$  گشتاورهای تولیدشده توسط موتورها هستند. در نهایت، معادلات برای زوایای اوایلر به صورت معادلات ۷ تا ۹ زیر استخراج می‌شود:

$$\dot{\phi} = ((I_{yy} - I_{zz})\dot{\psi}\dot{\theta} + IU_2)/I_{xx} \quad (7)$$

$$\dot{\theta} = ((I_{zz} - I_{xx})\dot{\psi}\dot{\phi} + IU_3)/I_{yy} \quad (8)$$

$$\dot{\psi} = ((I_{xx} - I_{yy})\dot{\phi}\dot{\theta} + U_4)/I_{zz} \quad (9)$$

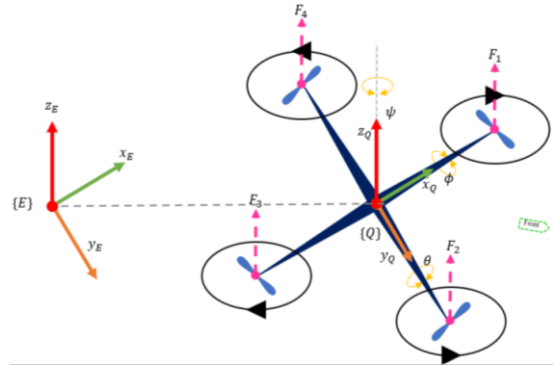
این فرضیات امکان استفاده از فرم ساده‌شده معادلات نیوتن-اوایلر را فراهم می‌کنند.

دو چارچوب مرجع تعریف می‌شود:

چارچوب زمین:  $\{X_E, Y_E, Z_E\}$

چارچوب بدنه:  $\{X_Q, Y_Q, Z_Q\}$

متغیرهای مکان نسبی در فضای زمین به صورت  $\xi = [x, y, z]^T$  و متغیرهای جهت‌گیری به صورت  $\eta = [\phi, \theta, \psi]^T$  تعریف می‌شوند. چرخش از چارچوب بدنه به زمین با استفاده از زوایای اوایلر و ماتریس‌های چرخش سه‌گانه به صورت  $R_\psi R_\theta R_\phi$  مدل می‌شود.



شکل ۱ چارچوب‌های مرجع کوادکوپتر [۲۱]

Fig.1 Quadcopter reference frame [21]

زوایای اوایلر، سه زاویه برای توصیف جهت‌گیری جسم صلب در سیستم مختصات می‌باشند. این زوایا رابطه میان دو چارچوب مرجع را بیان می‌کنند. با استفاده از این زوایا، مختصات یک نقطه از یک چارچوب مرجع به چارچوب مرجع دیگری انتقال داده می‌شود. این زوایا با نماد  $\phi, \theta, \psi$  نمایش داده و به ترتیب، زوایای رول، پیچ و یاء نامیده می‌شوند. بیان دوران بدنه حول محورهای مختصات سیستم، توسط این زوایا توسط معادله ۱ صورت می‌گیرد.

$$R(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix}$$

$$R(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & -\sin\theta \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

$$R(\psi) = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = R_z(\psi)R_y(\theta)R_x(\phi) \quad (1)$$

مقادیر پارامترهای مورد استفاده جهت بررسی مدل دینامیکی و انجام شبه‌سازی‌ها در جدول ۱ آورده شده است:

جدول ۱ پارامترهای فیزیکی پهپاد [۲۱]

Table 1 Physical parameters of the drone [21]

| Unit              | Value   | Symbol   |
|-------------------|---------|----------|
| Kg                | 0.8     | $m$      |
| m                 | 0.3     | $l$      |
| Kg.m <sup>2</sup> | 0.01567 | $I_{xx}$ |
| Kg.m <sup>2</sup> | 0.01567 | $I_{yy}$ |
| Kg.m <sup>2</sup> | 0.02834 | $I_{zz}$ |
| m.s <sup>2</sup>  | 9.81    | $g$      |

برای مدل‌سازی باد، شتاب‌هایی با عنوان  $a_{xw}, a_{yw}$  به معادلات دینامیکی انتقالی به شرح رابطه ۱۰ و ۱۱ مطابق مرجع ۲۴ اضافه شده‌اند:

$$\ddot{x} = a_{xm} + a_{xw} \quad (10)$$

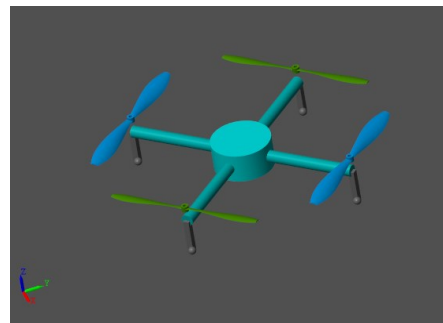
$$\ddot{y} = a_{ym} + a_{yw} \quad (11)$$

در این معادله  $a_{xm}, a_{ym}$  شتاب‌های ناشی از مدل اصلی دینامیک کوادروتور و  $a_{xw}, a_{yw}$  شتاب‌های ناشی از اثر باد می‌باشند.

برای این مدل‌سازی مفروضات زیر را در نظر گرفتیم:  
 ۱- مقدار باد اعمال شده در راستای محور z به سیستم برابر با صفر می‌باشد.

۲- برای کوادروتوری با ویژگی‌های دینامیکی پژوهش نباید شتاب باد بیشتر از ۳ m/s<sup>2</sup> شود زیرا پرنده سقوط خواهد کرد.

این مدل‌سازی به ما اجازه داد تا اغتشاشات محیطی در تحلیل پاسخ کوادروتور لحاظ شود.



شکل ۳ مدل‌سازی کوادروتور در متلب

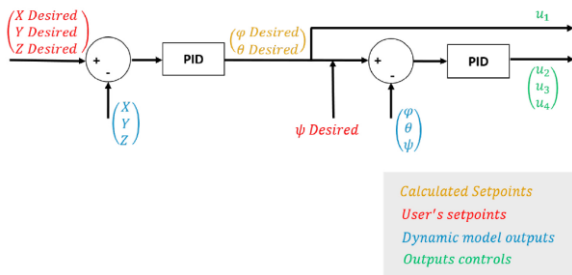
Fig.3 Quadrotor modeling in MATLAB

### ۳- ساختار کنترل‌کننده

در این پروژه، با هدف کنترل دقیق موقعیت و جهت‌گیری کوادروتور در حضور اغتشاشات محیطی نظیر باد، چارچوبی ترکیبی طراحی شده است که شامل یک کنترل‌کننده پایه‌ی تناسبی-انتگرالی-مشتقی، یک تنظیم‌کننده‌ی ضرایب مبتنی بر یادگیری تقویتی و یک مشاهده‌گر اغتشاش به‌همراه جبران‌ساز آن می‌باشد. در ادامه، به‌صورت جامع به توصیف هر بخش پرداخته می‌شود.

### ۳-۱- کنترل‌کننده‌ی تناسبی-انتگرالی-مشتقی

به‌منظور اطمینان از عملکرد صحیح کوادروتور در اجرای وظایف محوله، پیاده‌سازی یک کنترل‌کننده مؤثر بر سیستم امری ضروری است. در این پروژه، یک کنترل‌کننده‌ی تناسبی-انتگرالی-مشتقی به‌منظور کنترل متغیرهای دینامیکی کوادروتور شامل موقعیت و جهت‌گیری طراحی و شبه‌سازی شده است. قدرت کنترل، سادگی در پیاده‌سازی و سهولت تنظیم ضرایب، از جمله مزایای اصلی این نوع کنترل‌کننده است. در کوادروتور، این کنترل‌کننده وظیفه دارد با تنظیم سرعت چرخش هر یک از روتورها، وضعیت مطلوب سیستم را حاصل کند. به‌طور خلاصه، کنترل‌کننده‌ی تناسبی-انتگرالی-مشتقی اختلاف بین مقدار مرجع و مقدار اندازه‌گیری شده را دریافت کرده و براساس آن سیگنال کنترلی مناسب را تولید می‌کند. شکل ۴ نمودار بلوکی کلی سیستم کنترل مبتنی بر تناسبی-انتگرالی-مشتقی را نشان می‌دهد.



شکل ۴ شماتیک سیستم کنترل‌کننده پایه [۲۲]

Fig.4 Schematic of the baseline control system [22]

خطای سیستم به‌صورت  $e(t) = x_d(t) - x(t)$  تعریف می‌شود که در آن  $x_d(t)$  مقدار مطلوب و  $x(t)$  مقدار اندازه‌گیری شده است. خروجی کنترل‌کننده نیز طبق رابطه ۱۲ مطابق مرجع ۲۲ محاسبه می‌گردد:

$$U_z = K_p e(t) + K_i \int_0^t e(t) dt + K_d \frac{de(t)}{dt} \quad (12)$$

که در آن  $e(t)_z = z_d(t) - z(t)$  است.

برای کنترل زاویه‌های جهت‌گیری رول ( $\phi$ )، پیچ ( $\theta$ ) و یاء ( $\psi$ )، سه ورودی کنترلی به‌ترتیب با  $U_3, U_2, U_1$  در نظر گرفته شده‌اند. کنترل‌کننده‌ی تناسبی-انتگرالی-مشتقی مربوط به هر یک از این زوایا به‌شکل رابطه ۱۳ تعریف می‌گردد:

$$U_{2,3,4} = K_p e(t)_n + K_i \int_0^t e(t)_n dt + K_d \frac{de(t)_n}{dt} \quad (13)$$

که در آن  $n \in \{\phi, \theta, \psi\}$  و  $e(t)_n = n_d(t) - n(t)$  می‌باشد.

برای ردیابی مسیر در صفحه‌ی افقی ( $x, y$ ) به‌دلیل محدودیت ساختار کنترلی کوادروتور، نمی‌توان مستقیماً از ورودی  $U_1$  استفاده کرد. در عوض، موقعیت  $x$  و  $y$  از طریق ایجاد انحراف در زوایای رول و پیچ کنترل می‌شوند. فرضیات زاویه‌ی کوچک به کار گرفته شده است تا معادلات ساده‌تر گردند. برای محاسبه زوایای مطلوب رول و پیچ از کنترل‌کننده موقعیت، از روابط ۱۴ و ۱۵ استفاده می‌شود:

$$\phi_d = u_x \sin \psi_d - u_y \cos \psi_d \quad (14)$$

$$\theta_d = u_x \cos \psi_d + u_y \sin \psi_d \quad (15)$$

که در آن  $e_x, e_y, e_z, e_\psi$  خطای موقعیت در محورهای  $x, y, z$  و خطای  $yaw$  هستند.  $u_1, u_2, u_3, u_4$  سیگنال‌های کنترلی چهار ملخ هستند.  $I_{bounds}$  نیز تابع شاخصی است که اگر ورودی‌ها یا خطاها از حدود مجاز بیشتر شوند مقدار ۱ می‌گیرد و در غیر این صورت ۰، یعنی جریمه‌ی شدید برای نقض قیود در نظر گرفته شد. سپس برای پاداش مثبت یک ترم به صورت زیر اضافه شده است.

عامل پاداش  $+20$  دریافت می‌کند اگر تمام شرایط زیر برقرار باشند:

- $|e_x| \leq 0.5$
- $|e_y| \leq 0.5$
- $|e_z| \leq 0.5$
- $|e_\theta| \leq 0.2618 \text{ rad } (\approx 15^\circ)$
- $|e_\phi| \leq 0.2618 \text{ rad } (\approx 15^\circ)$

معیار طراحی تابع پاداش در این پژوهش بر پایه‌ی ترکیب دقت، کارایی انرژی و پایداری بنا شده است. به این ترتیب تابع پاداش دقت ردیابی مسیر را از طریق مربع خطاها، مصرف انرژی و روان بودن کنترل را از طریق مربع سیگنال‌های کنترلی، رعایت محدودیت‌ها را با جریمه سنگین و رسیدن به شرایط ایده‌آل را با پاداش مثبت بزرگ در نظر می‌گیرد [۲۳].

پارامترهای مورد استفاده در این شبیه‌سازی، در جدول ۲ شرح داده شده است.

جدول ۲ پارامترهای مورد استفاده در آموزش

Table 2 Parameters used in training

|         |   |
|---------|---|
| 256     | steps used in the update process        |
| 256     | Batch size                              |
| [64,64] | layers of the actor and critic networks |
| tanh    | Activation function                     |
| 0.0001  | Learning rate                           |
| 8500    | Total number of training steps          |

در ادامه طراحی کنترل‌کننده، به منظور افزایش پایداری و مقاوم‌سازی سامانه در برابر اغتشاشات بیرونی مانند باد، از چارچوبی شامل مشاهده‌گر اغتشاش و جبران‌ساز استفاده شده است.

### ۳-۳- مشاهده‌گر اغتشاش و جبران‌ساز

با توجه به این که در فرآیند آموزش عامل، اغتشاشات خارجی لحاظ نمی‌شوند، عملکرد کنترل‌کننده یادگیری تقویتی ممکن است در شرایط واقعی که با اغتشاشاتی نظیر باد، نویز حسگرها یا تغییرات مدل‌سازی مواجه است، کاهش یابد. به همین دلیل، یک سازوکار مکمل طراحی شده است که در آن نیروی اغتشاش تخمین زده می‌شود و سپس با اعمال جبران مناسب به ورودی‌های کنترلی، عملکرد کوادروتور بهبود می‌یابد. ایده اصلی مشاهده‌گر اغتشاش آن است که نیروی خارجی

که در آن  $u_x$  و  $u_y$  سیگنال‌های کنترلی موقعیت در راستاهای  $x$  و  $y$  هستند. ساختار تناسبی-انتگرالی-مشتقی برای این سیگنال‌ها به صورت رابطه 61 تدوین می‌گردد:

$$U_{x,y} = K_p e(t)_v + K_i \int_0^t e(t)_v dt + K_d(t) \frac{de(t)_v}{dt} \quad (16)$$

که در آن  $v \in \{x, y\}$  و  $v(t) = v_d(t) - v(t)$  می‌باشد. به منظور تنظیم این ضرایب، از روش یادگیری تقویتی استفاده شده است که مراحل آن به شرح زیر است.

### ۳-۲- یادگیری تقویتی

یادگیری تقویتی یکی از شاخه‌های اصلی یادگیری ماشین است که در آن یک عامل از طریق تعامل با محیط و دریافت پاداش، سیاستی بهینه برای انتخاب اعمال در شرایط مختلف می‌آموزد. هدف اصلی یادگیری تقویتی، پیشینه‌سازی پاداش تجمعی در طول زمان است، بدون آنکه مدل دقیقی از محیط در اختیار داشته باشد. این ویژگی، یادگیری تقویتی را به ابزاری توانمند برای کنترل سیستم‌های پیچیده، غیرخطی، و دارای پویایی‌های ناشناخته تبدیل می‌کند.

در زمینه کنترل، یادگیری تقویتی می‌تواند به جای طراحی دستی قوانین کنترلی یا تنظیم پارامترها از طریق آزمون و خطا، به صورت خودکار و بهینه کنترل‌کننده را طراحی یا تنظیم کند. در مورد سیستم‌هایی همچون کوادروتور، که دارای دینامیک غیرخطی، جفت‌شده و متغیر با زمان هستند، یادگیری تقویتی مزایای بسیاری دارد.

برای مسئله حاضر، از الگوریتم‌های DDPG و TD3 بهره گرفته شده است. این الگوریتم‌ها، از نوع بازیگر-منتقد بوده و قابلیت کنترل در فضای پیوسته را دارا هستند. الگوریتم‌های بازیگر-منتقد از دو شبکه عصبی استفاده می‌کنند: شبکه بازیگر برای تولید اعمال، در اینجا، ضرایب کنترل‌کننده تناسبی-انتگرالی-مشتقی و شبکه منتقد برای ارزیابی کیفیت اعمال انتخاب‌شده بر اساس مقدار پاداش دریافتی می‌باشد.

در این پروژه، اعمال تولید شده توسط عامل یادگیری تقویتی عبارت‌اند از شش ضریب کنترل‌کننده تناسبی-انتگرالی-مشتقی که به صورت پیوسته در طول زمان تنظیم می‌شوند:

سه ضریب برای کنترل موقعیت‌ها:  $K_p^{pos}, K_i^{pos}, K_d^{pos}$

سه ضریب برای کنترل زوایا:  $K_p^{ori}, K_i^{ori}, K_d^{ori}$

عامل در هر گام زمانی با مشاهده خطاهای موقعیت و جهت‌گیری و تغییراتشان، ضرایب مناسب را تولید کرده و آن‌ها را به کنترل‌کننده تناسبی-انتگرالی-مشتقی اعمال می‌کند. سپس، بر اساس کیفیت پاسخ سیستم، یک مقدار پاداش محاسبه می‌شود که معیار یادگیری عامل است.

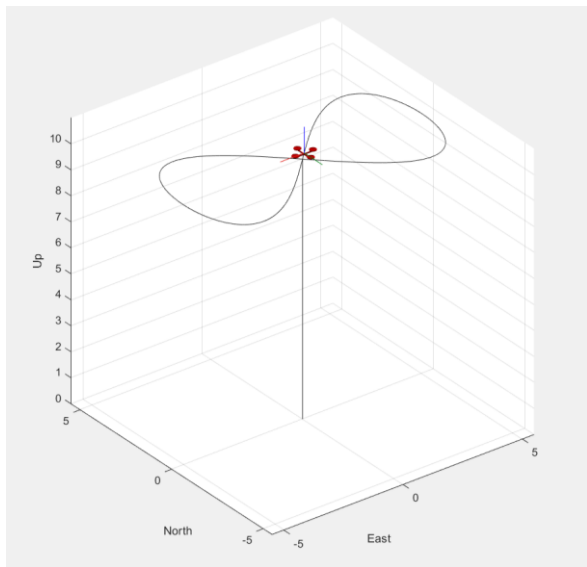
تابع پاداش طراحی شده، نقش بسیار مهمی در هدایت فرآیند یادگیری دارد. تابع پاداش، نقش بسیار مهمی در هدایت فرآیند یادگیری دارد. این تابع باید به گونه‌ای تعریف شود که عامل را به سمت رفتار مطلوب هدایت کند، یعنی کاهش خطای موقعیت و زاویه، با حداقل مصرف انرژی کنترلی به گونه‌ای که از حدود مجاز قیود دینامیکی عبور نکند. تابع پاداش در این پروژه به صورت رابطه ۱۷ مطابق مرجع ۲۳ تعریف شده است:

$$r(t) = -(e_x^2 + e_y^2 + e_z^2 + e_\psi^2 + 0.1(u_1^2 + u_2^2 + u_3^2 + u_4^2) + 10 I_{bounds}) \quad (17)$$

مورد نظر اضافه می‌شوند تا با تغییر جهت نیروی تراست، تأثیر نیروی اغتشاش خارجی خنثی شود. این ساختار موجب می‌شود که کنترل‌کننده بدون نیاز به آموزش مجدد، توانایی حفظ موقعیت و پایداری را در مواجهه با اغتشاشات به دست آورد [۲۴]. در ادامه به بررسی نتایج حاصل از شبیه‌سازی پرداخته می‌شود.

#### ۴- نتایج و بحث

برای ارزیابی جامع عملکرد کنترل‌کننده پیشنهادی، مجموعه‌ای از آزمایش‌های شبیه‌سازی طراحی شد که بارها و در شرایط متفاوت بدون باد و با باد در مسیرهای مختلف پروازی اجرا گردید تا اثر مستقیم روش‌های تنظیم ضرایب و جبران‌سازی بر توان کنترل و ردیابی مسیر مشخص شود.



شکل ۵ ردیابی مسیر بی‌نهایت در حضور باد

Fig. 5 Lemniscate trajectory tracking in the presence of wind

وارد بر کوادروتور در صفحه افقی را از طریق مقایسه شتاب واقعی اندازه‌گیری شده با شتاب پیش‌بینی شده توسط مدل اسمی تخمین بزند.

در حالت ایده‌آل، شتاب خالص کوادروتور در راستای محورهای بدنه، با توجه به نیروی تراست کلی و زوایای اوایلر محاسبه می‌شود. اجزای این شتاب به صورت روابط ۱۸ و ۱۹ تعریف می‌شوند.

$$a_x^{model} = \frac{u_1}{m} (\cos\phi \sin\theta \cos\psi + \sin\phi \sin\psi) \quad (18)$$

$$a_y^{model} = \frac{u_1}{m} (\cos\phi \sin\theta \sin\psi - \sin\phi \cos\psi) \quad (19)$$

حال با داشتن شتاب اندازه‌گیری شده می‌توان اختلاف آن را با مدل اسمی محاسبه کرده و به عنوان برآوردی از شتاب اغتشاش به صورت رابطه ۲۰ لحاظ نمود.

$$r_x = \ddot{x} - a_x^{model}, \quad r_y = \ddot{y} - a_y^{model} \quad (20)$$

که در معادلات  $\ddot{x}$  و  $\ddot{y}$  شتاب‌های واقعی در چارچوب مرجع اینرسی هستند.

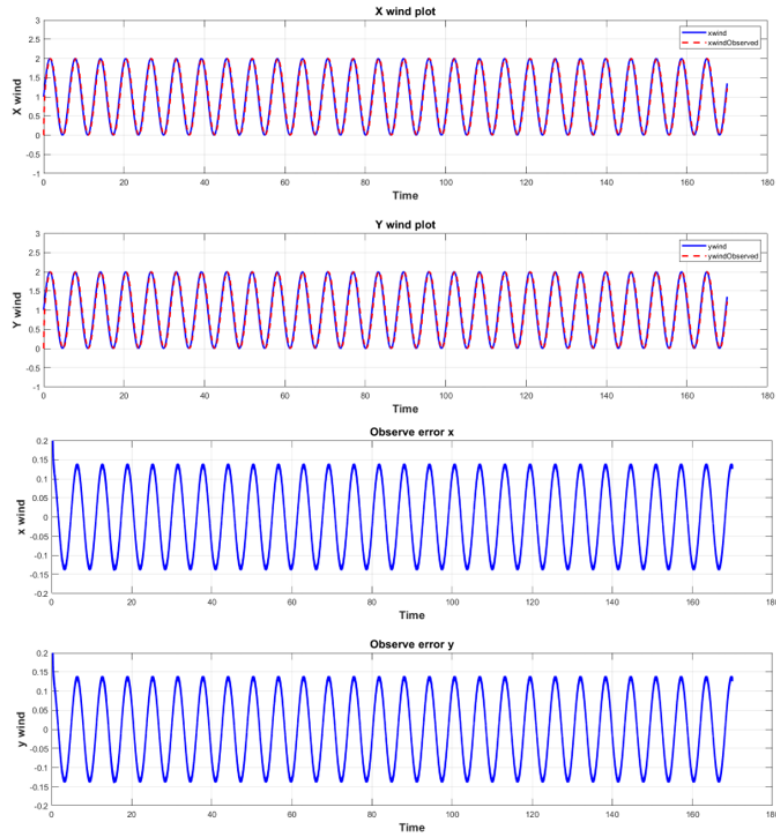
در نهایت، بردار اغتشاش برآورد شده به صورت رابطه ۲۱ تعریف می‌شود:

$$r = \begin{bmatrix} r_x \\ r_y \end{bmatrix} \approx \frac{1}{m} F_{ext} \quad (21)$$

که در آن  $F_{ext}$  نمایانگر نیروی خارجی مؤثر بر کوادروتور است. پس از تخمین بردار شتاب اغتشاش می‌توان از یک الگوریتم جبران‌سازی پیش‌خور ساده و مؤثر برای اصلاح زوایای مرجع رول و پیچ استفاده کرد. زوایای جبران‌ساز به صورت معادلات ۲۲ محاسبه می‌شوند.

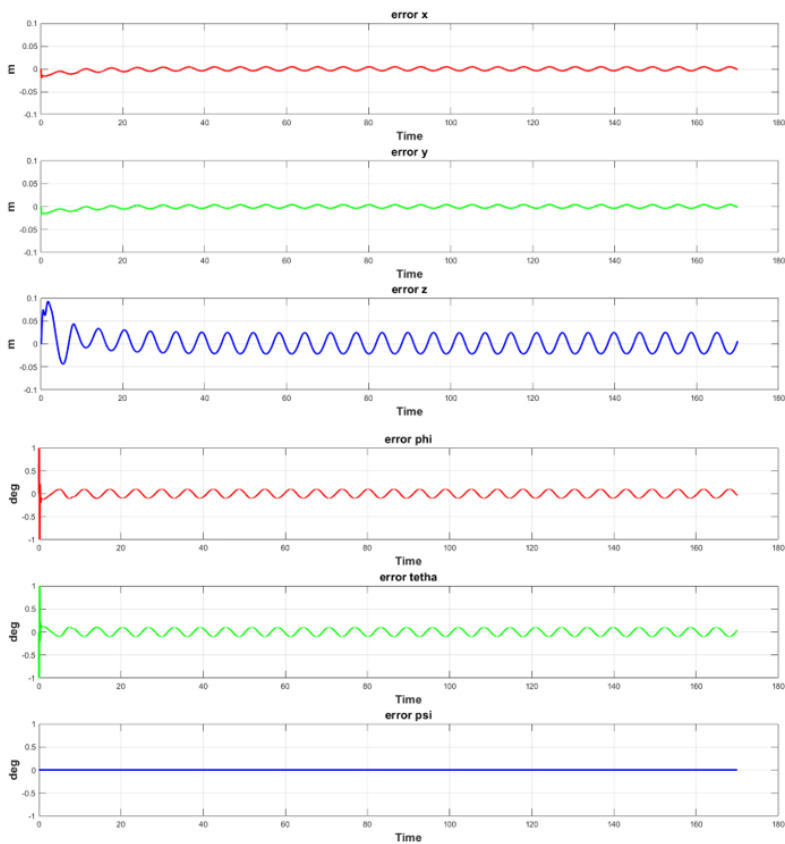
$$\phi_{ff} = \frac{a_y^{ext}}{g}, \quad \theta_{ff} = -\frac{a_x^{ext}}{g} \quad (22)$$

که  $g$  شتاب گرانشی زمین و  $\theta_{ff}$  و  $\phi_{ff}$  زوایای اصلاح شده برای جبران اغتشاش می‌باشند. در این مرحله، زوایای جبران‌ساز به زوایای هدف



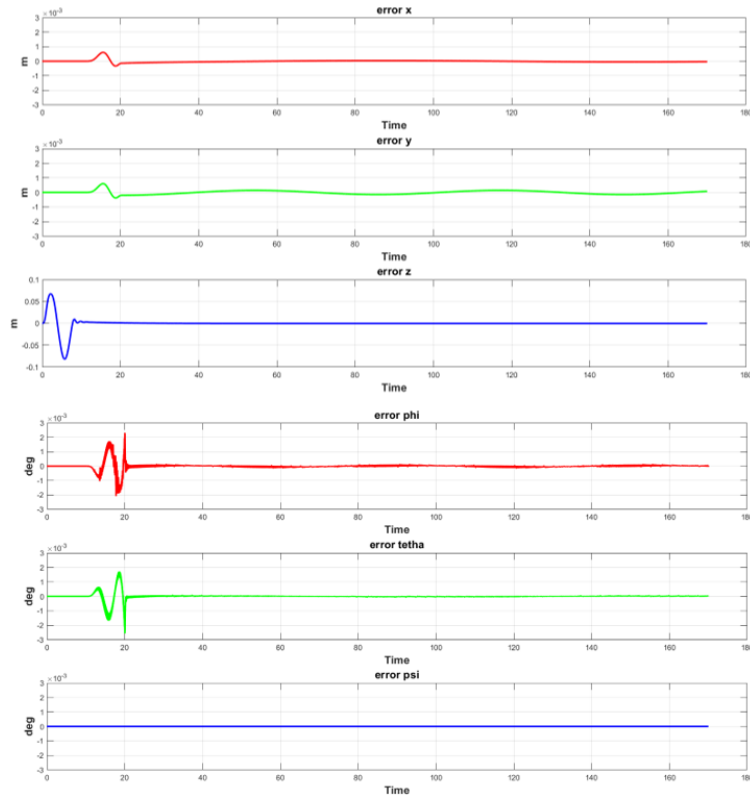
شکل ۶ نمودار باد سینوسی و خطای تخمین مشاهده‌گر

Fig. 6 Sin wind profile and observer estimation error



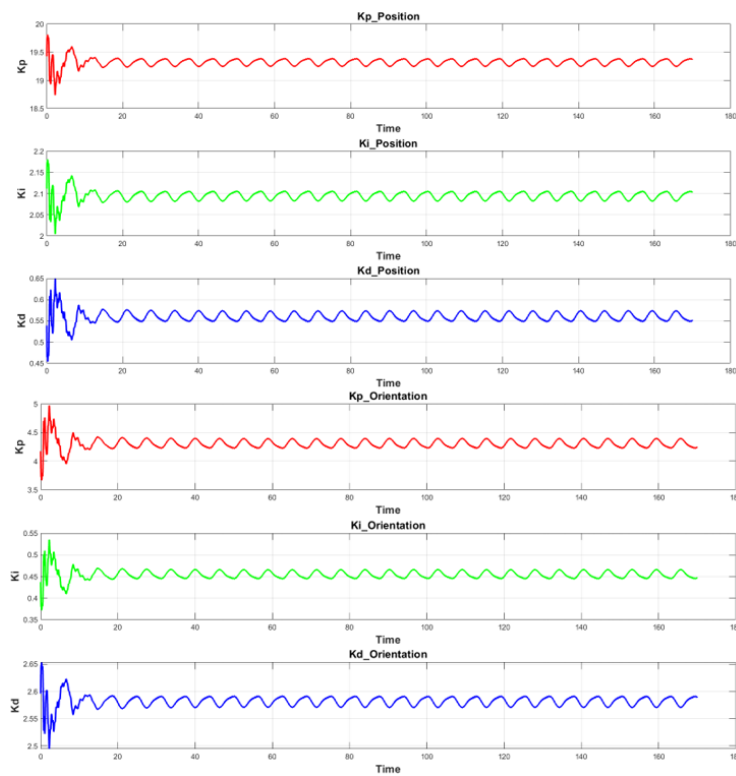
شکل ۷ نمودار خطای ردیابی مسیر در حضور باد با PID ثابت

Fig. 7 Trajectory tracking error in the presence of wind using fixed PID



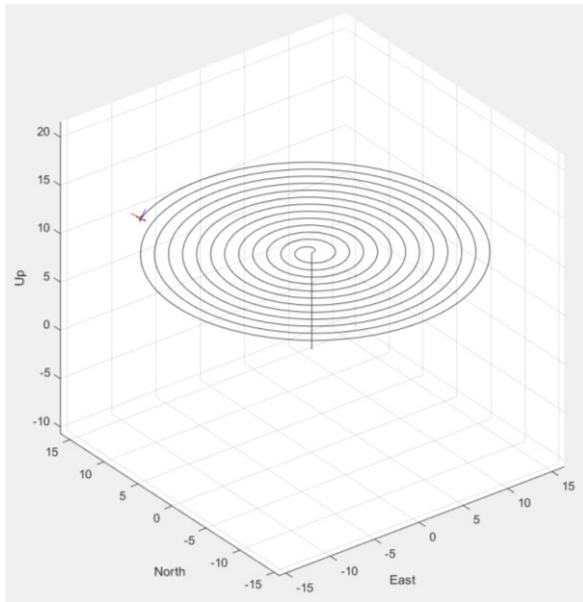
شکل ۸ نمودار خطای ردیابی مسیر در باد با PID-DDPG

Fig. 8 Trajectory tracking error under wind disturbance using DDPG-tuned PID



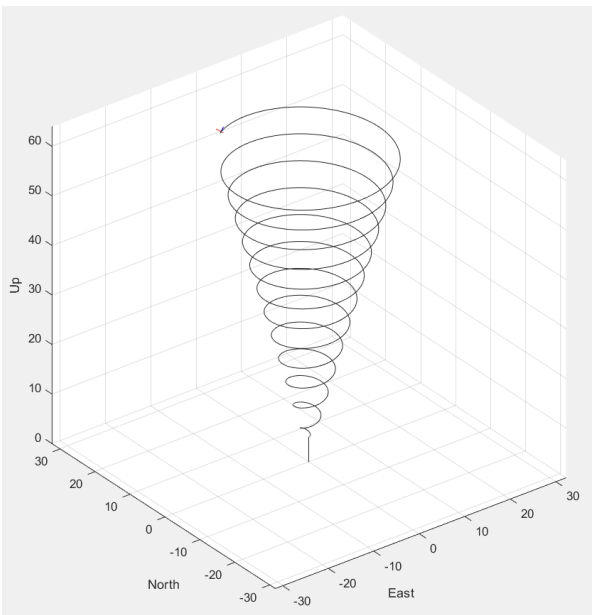
شکل ۹ نمودار ضرایب تنظیم‌شده توسط عامل DDPG

Fig. 9 PID gains tuned by the DDPG agent



شکل ۱۱ ردیابی مسیر مارپیچ دو بعدی در حضور باد

Fig. 11 2D Helix trajectory tracking in the presence of wind

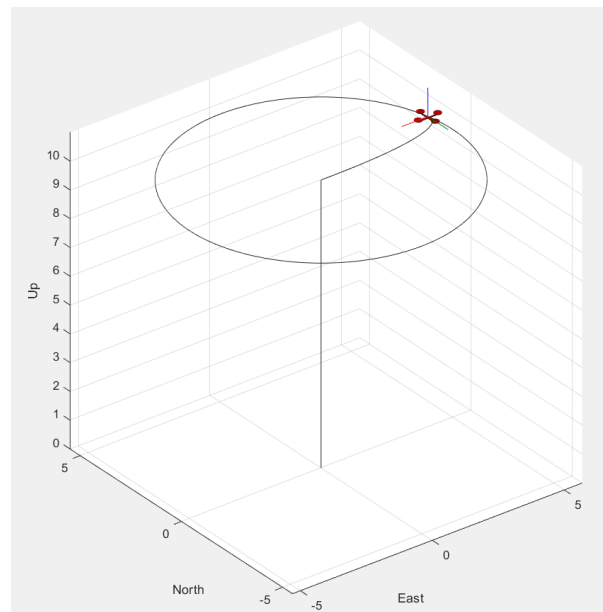


شکل ۱۲ ردیابی مسیر مارپیچ سه بعدی در حضور باد

Fig. 12 3D Helix trajectory tracking in the presence of wind

در مسیر بی‌نهایت، مقایسه مقادیر RMS نشان می‌دهد که الگوریتم DDPG در کاهش خطای موقعیت و زاویه نسبت به TD3 عملکرد بهتری داشته است. به‌طور تقریبی DDPG توانسته است مقادیر RMS موقعیت را حدود ۱۵٪ درصد کمتر از TD3 نگه دارد و در مؤلفه‌های زاویه‌ای نیز حدود ۱۰٪ بهبود نسبت به TD3 ارائه دهد. این امر بیانگر توانایی بیشتر DDPG در یادگیری دقیق‌تر سیاست کنترلی برای مسیرهای نوسانی و پیچیده است. در مسیر دایره‌ای اما نتایج متفاوتی مشاهده شد. در این حالت، TD3 در کاهش RMS موقعیت حدود ۵٪ نسبت به DDPG برتری دارد. در این مسیر با نوسانات ملایم و مرجع

در محورهای موقعیت مقادیر RMS در حالت کنترل‌کننده تناسبی-انتگرالی-مشتقی با ضرایب ثابت برابر با تقریباً  $0.00257$ ،  $0.00267$  و  $0.00953$  است. هنگامی که جبران‌ساز اغتشاش به سیستم افزوده می‌شود، این مقادیر به ترتیب به  $0.00177$ ،  $0.00171$  و  $0.00631$  کاهش پیدا می‌کنند. به‌طور میانگین، این کاهش‌ها بیانگر بهبود حدود ۳۴ درصدی در عملکرد موقعیت‌یابی هستند. این بهبود ناشی از عملکرد مؤثر جبران‌ساز در تخمین و حذف مؤلفه‌های اغتشاشی در راستاهای افقی و عمودی است. در مرحله بعد، با به‌کارگیری عامل یادگیری تقویتی DDPG برای تنظیم ضرایب کنترل‌کننده تناسبی-انتگرالی-مشتقی، مقادیر RMS در همان محورهای موقعیت به ترتیب به  $0.00161$ ،  $0.00143$  و  $0.00362$  کاهش یافتند. در مقایسه با حالت پایه، این مقادیر نشان‌دهنده‌ی کاهش خطا به میزان تقریبی ۴۵ تا ۵۰٪ و در مقایسه با حالت دارای جبران‌ساز، حدود ۲۰٪ بهبود بیشتر هستند. در مؤلفه‌های زاویه‌ای، روند مشابهی مشاهده می‌شود. خطاهای RMS از حدود  $0.00068$  و  $0.00063$  در حالت اولیه، به  $0.00063$  و  $0.00060$  در حالت دارای جبران‌ساز، و سپس به  $0.00015$  و  $0.00014$  در حالت تنظیم‌شده توسط عامل یادگیری تقویتی کاهش یافته‌اند. این روند حاکی از بهبود نزدیک به ۸۰ درصد در پایداری زاویه‌ای پس از اعمال یادگیری تقویتی است. فهرست علائم در این سناریو، باد هم به‌صورت پله‌ای و هم به‌صورت سینوسی و هم به‌صورت رندوم به سیستم وارد شد تا عملکرد در انواع مختلف اغتشاش سنجیده شود. نتایج نشان دادند که این ترکیب قادر است در برابر هر نوع اغتشاش پایداری مناسبی ایجاد کند و پرنده را با حداقل خطا در مسیر مرجع نگه دارد. علاوه بر این، ردیابی مسیرهای پیچیده‌تر شامل مسیر دایره‌ای و دو نوع مسیر مارپیچ نیز بررسی شد و در تمامی آن‌ها عملکرد پایدار و دقیق مشاهده شد.



شکل ۱۰ ردیابی مسیر دایره‌ای در حضور باد

Fig. 10 Circular trajectory tracking in the presence of wind

optimization approach. Future Generation Computer Systems, 2024. DOI: [10.1016/j.future.2024.05.007](https://doi.org/10.1016/j.future.2024.05.007)

[3] Zhang, R., et al., Reactive UAV-based automatic tunnel surface defect inspection with a field test. Automation in Construction, 2024. DOI: [10.1016/j.autcon.2024.105424](https://doi.org/10.1016/j.autcon.2024.105424)

[4] Fan, Y., et al., Research and verification of trajectory tracking control of a quadrotor carrying a load. Applied Sciences, 2022. DOI: [10.3390/app12031036](https://doi.org/10.3390/app12031036)

[5] Nwilag, B.D., A.E. Eyoh, and C.E. Ndehedehe, Digital topographic mapping and modelling using low altitude unmanned aerial vehicle. Modeling Earth Systems and Environment, 2023. DOI: [10.1007/s40808-022-01677-z](https://doi.org/10.1007/s40808-022-01677-z)

[6] Yu, C., et al., Trajectory tracking control of an unmanned aerial vehicle with deep reinforcement learning for tasks inside the EAST. Fusion Engineering and Design, 2023. DOI: [10.1016/j.fusengdes.2023.113894](https://doi.org/10.1016/j.fusengdes.2023.113894)

[7] Izadi, M. and R. Faieghi, High-gain disturbance observer for robust trajectory tracking of quadrotors. Control Engineering Practice, 2024. DOI: [10.1016/j.conengprac.2024.105854](https://doi.org/10.1016/j.conengprac.2024.105854)

[8] Zeng, Z., et al. Modeling and Control for Quadrotor Aircraft Subject to Wind Gusts. IEEE, 2023. DOI: [10.23919/CCC58697.2023.10240862](https://doi.org/10.23919/CCC58697.2023.10240862)

[9] Xiong, H. and Y. Zhang, Reinforcement learning-based formation-surrounding control for multiple quadrotor UAVs pursuit-evasion games. ISA transactions, 2024. DOI: [10.1016/j.isatra.2023.12.006](https://doi.org/10.1016/j.isatra.2023.12.006)

[10] Han, H., et al., Cascade flight control of quadrotors based on deep reinforcement learning. IEEE Robotics and Automation Letters, 2022. DOI: [10.1109/LRA.2022.3196455](https://doi.org/10.1109/LRA.2022.3196455)

[11] Rezaei, A., A. Talaeizadeh, and A. Alasty, Self-Tuning PID Controller for an Arm-Angle Adjustable Quadrotor Using an Actor-Critic-Based Neural Network. IEEE, 2024. DOI: [10.1109/ICIS64839.2024.10887514](https://doi.org/10.1109/ICIS64839.2024.10887514)

[12] Sonmez, S., et al. Reinforcement learning based pid parameter tuning and estimation for multirotor uavs. IEEE, 2024. DOI: [10.1109/ICUAS60882.2024.10557032](https://doi.org/10.1109/ICUAS60882.2024.10557032)

[13] Hoover, R.J., W. Wu, and K. Shimada. Applying Reinforcement Learning to PID Flight Control of a Quadrotor Drone to Mitigate Wind Disturbances. IEEE, 2024. DOI: [10.1109/ICARA60736.2024.10553186](https://doi.org/10.1109/ICARA60736.2024.10553186)

[14] Lu, S., Y. Li, and Z. Liu. Quadrotor Control using Reinforcement Learning under Wind Disturbance. IEEE, 2023. DOI: [10.1109/CCDC58219.2023.10326621](https://doi.org/10.1109/CCDC58219.2023.10326621)

[15] Qingqing, Z., et al. A PID gain adjustment scheme based on reinforcement learning algorithm for a quadrotor. IEEE, 2020. DOI: [10.23919/CCC50068.2020.9188426](https://doi.org/10.23919/CCC50068.2020.9188426)

[16] Barzegar, A. and D.-J. Lee, Deep reinforcement learning-based adaptive controller for trajectory tracking and altitude control of an aerial robot. Applied Sciences, 2022. DOI: [10.3390/app12094764](https://doi.org/10.3390/app12094764)

[17] Wu, L., et al., Deep reinforcement learning with corrective feedback for autonomous uav landing on a mobile platform. Drones, 2022. DOI: [10.3390/drones6090238](https://doi.org/10.3390/drones6090238)

[18] Sönmez, S., et al., Reinforcement learning based prediction of pid controller gains for quadrotor uavs. IEEE, 2025. DOI: [10.3390/drones9080581](https://doi.org/10.3390/drones9080581)

[19] Ali Mottahedi, Ali Akbar Zade, Adaptive robust sliding mode control of quadrotor in the Presence of Wind Disturbance. Mme Modares, 2016. DOI: <https://dor.isc.ac/dor/20.1001.1.10275940.1395.16.12.4.1.6>

یکنواخت. هر دو الگوریتم رفتار پایداری از خود نشان داده‌اند، ولی پاسخ TD3 نرم‌تر و نزدیک‌تر به مسیر مرجع بوده است. در مسیرهای مارپیچ، تفاوت بین دو الگوریتم بارزتر است. مقادیر RMS موقعیت در DDPG حدود ۲۰٪ کمتر از TD3 بوده و در مؤلفه‌های زاویه‌ای نیز نوسانات کاهش محسوسی داشته‌اند. این نتایج نشان می‌دهند که DDPG توانسته است در مسیرهای دارای تغییرات سریع و پیچیدگی بالا، سازگاری بهتری با دینامیک سیستم برقرار کند و خطاهای گذرا را سریع‌تر میرا نماید.

## ۵- جمع‌بندی و پیشنهادات

نتایج شبیه‌سازی‌ها به خوبی نشان دادند که رویکردهای مختلف کنترلی تفاوت معناداری در عملکرد ردیابی و مقاومت در برابر اغتشاش ایجاد می‌کنند. در مواجهه با باد سینوسی، کنترل‌کننده PID با ضرایب ثابت دچار بیشینه انحراف قابل توجهی شد. افزودن مشاهده‌گر و جبران‌ساز به این کنترل‌کننده با ضرایب ثابت توانست بیشینه انحراف را حدود ۱۰٪ کاهش دهد. این مقایسه نشان داد که حتی در شرایطی که از یادگیری تقویتی استفاده نشود، به‌کارگیری یک جبران‌ساز مبتنی بر مشاهده‌گر می‌تواند به‌طور چشمگیری عملکرد سیستم را در برابر اغتشاشات باد متغیر بهبود دهد. در ترکیب نهایی که شامل کنترل‌کننده PID با ضرایب تنظیم شده توسط عامل DDPG به همراه مشاهده‌گر و جبران‌ساز بود، بهترین عملکرد حاصل شد. در این حالت، خطای ردیابی مسیر تحت شرایط بادهای سینوسی در حدود ۲۵٪ کمتر از کنترل‌کننده PID با ضرایب ثابت بود. در نهایت، مقایسه نسبی بین الگوریتم‌های یادگیری تقویتی در ردیابی مسیرهای مختلف نشان داد که در مسیر بی‌نهایت الگوریتم DDPG، ۱۰٪ بهینه‌تر از الگوریتم TD3 عمل کرد. در مسیر دایروی الگوریتم TD3، ۵٪ بهینه‌تر از الگوریتم DDPG عمل کرد. در مسیرهای مارپیچ نیز الگوریتم DDPG، ۲۰٪ بهینه‌تر از الگوریتم TD3 عمل کرد. به‌طور کلی، جمع‌بندی نتایج بیانگر آن است که هر یک از اجزای افزوده شده شامل یادگیری تقویتی برای تنظیم ضرایب کنترل‌کننده PID و جبران‌ساز مبتنی بر مشاهده‌گر برای مقابله با اغتشاشات باد، به‌طور مستقل موجب بهبود عملکرد می‌شوند، اما ادغام آن‌ها در یک چارچوب یکپارچه، بهترین عملکرد ممکن را فراهم می‌آورد.

با وجود نتایج رضایت‌بخش حاصل از طراحی و پیاده‌سازی چارچوب ترکیبی کنترل همچنان زمینه‌هایی برای توسعه‌ی بیشتر این سامانه وجود دارد. توسعه مدل باد از حالت شتابی به سرعتی، گسترش فضای مشاهدات عامل و اضافه کردن یک جریمه مربوط به تغییرات سرعت کوادروتور پیشنهادهای ما برای بهبود و گسترش پروژه می‌باشند.

**تأییدیه اخلاقی:** محتویات علمی این مقاله حاصل پژوهش نویسندگان است و در هیچ نشریه ایرانی و غیرایرانی منتشر نشده است.

**تعارض منافع:** هیچ تعارض منافعی برای اظهار وجود ندارد.

## منابع

[1] Hu, J., et al., Fault-tolerant cooperative navigation of networked UAV swarms for forest fire monitoring. Aerospace Science and Technology, 2022. DOI: [10.1016/j.ast.2022.107494](https://doi.org/10.1016/j.ast.2022.107494)

[2] Naouri, A., et al., Maximizing UAV fog deployment efficiency for critical rescue operations: A multi-objective

- [22] Abdelhay, S. and A. Zakriti, Modeling of a quadcopter trajectory tracking system using PID controller. *Procedia Manufacturing*, 2019. DOI: 10.1016/j.promfg.2019.02.253
- [23] Mohanty, A. and E. Schneider, Tuning of an Aircraft Pitch PID Controller with Reinforcement Learning and Deep Neural Net. 2019. DOI: [91355065/26643693.pdf](https://doi.org/10.1016/j.promfg.2019.02.253)
- [24] Pi, C.-H., W.-Y. Ye, and S. Cheng, Robust quadrotor control through reinforcement learning with disturbance compensation. *Applied Sciences*, 2021. DOI: [10.3390/app11073257](https://doi.org/10.3390/app11073257)
- [20] Farhad Parivash, and Ali Qasemi, trajectory tracking control of Quadrotor using Fractional-Order Fuzzy PID controller in the Presence of Wind Disturbance. *Mme Modares*, 2018. DOI: <https://dor.isc.ac/dor/20.1001.1.10275940.1397.18.8.22.3>
- [21] Baharuddin, A.d. and M.A.M. Basri, Trajectory tracking of a quadcopter uav using pid controller. *ELEKTRIKA-Journal of Electrical Engineering*, 2023. DOI: [10.11113/elektrika.v22n2.440](https://doi.org/10.11113/elektrika.v22n2.440)