



HAL
open science

Machine learning and control engineering: The model-free case

Michel Fliess, Cédric Join

► **To cite this version:**

Michel Fliess, Cédric Join. Machine learning and control engineering: The model-free case. Future Technologies Conference, FTC 2020, Nov 2020, Vancouver, Canada. pp. 258-278, 10.1007/978-3-030-63128-4_20 . hal-02851119v2

HAL Id: hal-02851119

<https://polytechnique.hal.science/hal-02851119v2>

Submitted on 6 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Machine learning and control engineering: The model-free case

Michel Fliess^{1,3} and Cédric Join^{2,3}

¹ LIX (CNRS, UMR 7161), École polytechnique, 91128 Palaiseau, France.
Michel.Fliess@polytechnique.edu

² CRAN (CNRS, UMR 7039), Université de Lorraine, BP 239,
54506 Vandœuvre-lès-Nancy, France.
Cedric.Join@univ-lorraine.fr

³ AL.I.E.N., 7 rue Maurice Barrès, 54330 Vézelize, France.
{michel.fliess, cedric.join}@alien-sas.com

Abstract. This paper states that Model-Free Control (MFC), which must not be confused with Model-Free Reinforcement Learning, is a new tool for Machine Learning (ML). MFC is easy to implement and should be substituted in control engineering to ML via Artificial Neural Networks and/or Reinforcement Learning. A laboratory experiment, which was already investigated via today’s ML techniques, is reported in order to confirm this viewpoint.

Keywords: Model-free control, intelligent proportional controllers, machine learning, reinforcement learning, supervised learning, unsupervised learning, artificial neural networks, half-quadrotor.

1 Introduction

The huge popularity today of *Machine Learning (ML)* is due to many beautiful achievements of *Artificial Neural Networks (ANNs)* (see, *e.g.*, [41, 42, 67]) and *Reinforcement Learning (RL)* (see, *e.g.*, [72]). Let us quote [63]: “Reinforcement Learning is the subfield of machine learning that studies how to use past data to enhance the future manipulation of a dynamical system. A control engineer might be puzzled by such a definition and interject that this precisely the scope of control theory. That the RL and the control theory communities remain practically disjoint has led to the co-development of vastly different approaches to the same problems. However, it should be impossible for a control engineer not to be impressed by the recent successes of the RL community such as solving Go [68].” Many concrete case-studies have already been investigated: see, *e.g.*, [3, 9, 13, 14, 16, 17, 21, 22, 32, 35, 40, 43–45, 47, 52–54, 56, 59–61, 70, 76, 77, 79, 80, 83, 85, 87–89]. Although those works are most promising, they show that ANNs and RL have perhaps not provided in this field such stunning successes as they did elsewhere.

Remark 1. The connection of RL with *optimal control* is known since ever (see, e.g., [15, 38, 63]). According to [48, 49] tools stemming from advanced control theory should enhance RL in general.

This communication suggests another route: *Model-Free Control (MFC)* in the sense of [25].

Remark 2. The meaning of *Model-Free Reinforcement Learning* is quite distinct. In model-free RL there is no transition probability distribution, *i.e.*, no *model* (see, e.g., [71, 72]). *Q-learning* is an example which has been used several times in control engineering.

MFC, which is easy to implement both from software [25] and hardware [33] viewpoints, leads to many acknowledged applications (see the bibliographies in [25], [6] for most references until 2018).⁴ The relationship with ML is sketched below.

Consider a system S with a single input u and a single output y . Under rather weak assumptions S may be approximated (see [25] and Section 2) by the *ultra-local model*:

$$\dot{y}(t) = F(t) + \alpha u(t) \quad (1)$$

where F encompasses not only the poorly known structure of S but also the disturbances. Since α is a constant that is easy to nail down (see Section 2.2), the main task is to determine the time-dependent quantity $F(t)$. A real-time estimate $F_{\text{est}}(t)$ is given thanks to a new parameter identification technique [26, 27, 69] by the following integral of the input-output data

$$F_{\text{est}}(t) = -\frac{6}{\tau^3} \int_{t-\tau}^t [(\tau - 2\sigma)y(\sigma) + \alpha\sigma(\tau - \sigma)u(\sigma)] d\sigma \quad (2)$$

where $\tau > 0$ is small. Formula (2), which ought to be viewed as a kind of *unsupervised learning* (see, e.g., [65]) procedure, takes into account the time arrow: the structure of S and especially the disturbances might be time-varying in an unexpected way. Moreover the unavoidable corrupting noises are attenuated by the integral, which is a low pass filter (see Remark 3). Associate the feedback loop [25]

$$u = -\frac{F_{\text{est}} - \dot{y}^* + K_P e}{\alpha} \quad (3)$$

where

- y^* is a reference trajectory,
- $e = y - y^*$ is the tracking error,
- $K_P \in \mathbb{R}$ is a gain.

⁴ Some applications are patented. Others have been published more recently (see, e.g., [1, 2, 7, 8, 18–20, 28–31, 36, 51, 55, 57, 58, 62, 64, 66, 73–75, 78, 81, 84, 86]).

It has been baptised *intelligent proportional controller* (*iP*), already some time ago [24]: this unsupervised learning permits not only to track the reference trajectory but also to limit the effects of the disturbances and of the poor system understanding (see Section 2.3 for further details). Note that ANNs and RL, and the corresponding methods from computer sciences and mathematics, are not employed.⁵

In order to support our viewpoint a lab experiment has been selected. A half-quadrotor is available to one of the authors (C.J.). Moreover quadrotors and half-quadrotors have been already examined via ANNs and RL: see, *e.g.*, [32, 40, 61, 79].⁶ Our results are not only excellent but also easy to obtain. The interested reader is invited to compare with the above references. It has been moreover shown [20] that the performances of MFC with respect to quadrotors are superior to those of PIDs (see, *e.g.*, [4, 5]).⁷

This communication is organized as follows. MFC is reviewed in Section 2. Section 3 discusses the lab experiment. Some concluding remarks may be found in Section 4.

2 MFC as a ML technique

2.1 The input-output system as a functional

Consider for notational simplicity a SISO system, *i.e.*, a system with a single control variable $u(t)$ and a single output variable $y(t)$, where $t \geq 0$ is the time. Even without knowing any “good” mathematical model we may assume that the system corresponds to a *functional* (see, *e.g.*, [37]), *i.e.*, a function of functions,

$$y(t) = \mathcal{F}(u(\mathbf{t}) \mid 0 \leq \mathbf{t} \leq t) \quad (4)$$

\mathcal{F} depends not only on initial conditions at $t = 0$, but also on the unavoidable disturbances.

2.2 The ultra-local model

It has been demonstrated [25] that, under mild assumptions, the input-output behavior (4) may be well approximated by the *ultra-local model*:

$$y^{(n)}(t) = F(t) + \alpha u(t)$$

where the order $n \geq 1$ of derivation is in all known examples equal to 1 or 2. In most concrete case-studies, $n = 1$. The case $n = 2$ arises, for instance, with weak frictions [25] (see, *e.g.*, [50] for a concrete case-study). Consider from now on only the case $n = 1$, *i.e.*, Equation (1), which works well in Section 3:

⁵ It is perhaps worth mentioning that some other fields of computer sciences might benefit from MFC (see, *e.g.*, [10, 34]).

⁶ Numerous other references do not use any traditional AI techniques.

⁷ *Proportional-Integral-Derivative* (*PID*) controllers ought to be regarded as the “bread and butter” of control engineering!

- The time-dependent quantity F is not only encompassing the internal structure of the system, which may be poorly known, but also the disturbances. Write F_{est} its estimate which is derived in Section 2.4.
- The constant $\alpha \in \mathbb{R}$ is chosen by the practitioner such that the three terms in Equation (1) are of the same magnitude. A precise determination of α is therefore meaningless. In practice α is easily chosen via two possible approaches:
 - the absolute value $|\frac{\alpha u(t)}{y(t)}|$ is not too far from 1,
 - trial and error, *i.e.*, a kind of *supervised learning* (see, *e.g.*, [65]).

2.3 Intelligent proportional controllers

Close the loop in Equation (1) with the iP (3). Equations (1) and (3) yield

$$\dot{e} + K_P e = F - F_{\text{est}}$$

If the estimation F_{est} is “good”: $F - F_{\text{est}}$ is “small”, *i.e.*, $F - F_{\text{est}} \simeq 0$, then $\lim_{t \rightarrow +\infty} e(t) \simeq 0$ if $K_P > 0$. It implies that the tuning of K_P is quite straightforward.⁸ This is a major benefit when compared to the tuning of “classic” PIDs (see, *e.g.*, [4, 5]).

2.4 ML via the estimation of F

Any function, for instance F in Equation (1), may be approximated under a weak integrability assumption by a piecewise constant function (see, *e.g.*, [12]). The estimation techniques below are borrowed from [26, 27, 69].

First approach Rewrite Equation (1) in the operational domain (see, *e.g.*, [82]):

$$sY = \frac{\Phi}{s} + \alpha U + y(0)$$

where Φ is a constant. We get rid of the initial condition $y(0)$ by multiplying both sides on the left by $\frac{d}{ds}$:

$$Y + s \frac{dY}{ds} = -\frac{\Phi}{s^2} + \alpha \frac{dU}{ds}$$

Noise attenuation is achieved by multiplying both sides on the left by s^{-2} . It yields in the time domain the real-time estimate Formula (2) thanks to the equivalence between $\frac{d}{ds}$ and the multiplication by $-t$, where $\tau > 0$ might be quite small. This integral, which is a low pass filter, may of course be replaced in practice by a classic digital linear filter.

⁸ See, *e.g.*, in [74] an optimization procedure, which, in some sense, is closer to today’s viewpoint on ML.

Second approach Close the loop with the iP (3). It yields:

$$F_{\text{est}}(t) = \frac{1}{\tau} \left[\int_{t-\tau}^t (\dot{y}^* - \alpha u - K_P e) d\sigma \right]$$

Remark 3. *Noises*, which are usually described in engineering and, more generally, in applied sciences via probabilistic and statistical tools, are related here to quick fluctuations around 0 [23]: The integral of any such noise over a finite time interval is close to 0. The robustness with respect to corrupting noises is thus explained. See, *e.g.*, [11, 69] for concrete applications in signal processing where the parameter estimation techniques of [26, 27, 69] have been employed.

2.5 MIMO systems

Consider a multi-input multi-output (MIMO) system with m control variables u_i and m output variables y_i , $i = 1, \dots, m$, $m \geq 2$. It has been observed in [39] and confirmed by all encountered concrete case-studies (see, *e.g.*, [75]), that such a system may usually be regulated via m monovariable ultra-local models:

$$y_i^{(n_i)} = F_i + \alpha_i u_i \quad (5)$$

where F_i may also depend on u_j , y_j , and their derivatives, $j \neq i$.

3 Experiments: A half-quadrotor

3.1 Process description

Our half-quadrotor (see Fig. 1), called AERO, is manufactured by Quanser.⁹ Two motors driving the propellers, which might turn clockwise or not, are controlling the angular positions: the azimuth (horizontal) velocity and pitch (vertical) position of the arms. Outputs y_1 and y_2 are respectively measures of the azimuth velocity (rad/ms) and pitch position (rad). Write v_i , $i = 1, 2$, the supply voltage of motor i , where $-24\text{v} \leq v_i \leq 24\text{v}$ (volt). Measures and control inputs are updated each 10ms.

3.2 Control

It is clear that y_i , $i = 1, 2$, is mainly influenced by v_i . Equations (5) and (3) become

$$\dot{y}_i = F_i + \alpha_i u_i \quad (6)$$

$$u_i = -\frac{F_{i,\text{est}} - \dot{y}^* + K_{P,i} e_i}{\alpha_i} \quad (7)$$

The control variable u_i in Equation (6) is defined by

- if $u_i > 0$, then $v_i = 10 + u_i$,
- if $u_i < 0$, then $v_i = -10 + u_i$.

In Equations (6)-(7), set $\alpha = 0.001$, $K_{P,1} = 0.5$, $\alpha_2 = 5$, $K_{P,2} = 500$. Everything is programed in *C#* and stored in the server.

⁹ See the link <https://www.quanser.com/products/quanser-aero/>



Fig. 1. The Quanser AERO half-quadrotor

3.3 Experiments

Nominal half-quadrotor Consider two scenarios:

- scenario 1 – simple reference trajectory (see Fig. 3 and 4),
- scenario 2 – complex reference trajectory (see Fig. 5 and 6).

The tracking is excellent in both cases in spite of the rotating blades, the gyroscopic effects, and the frictions, which are all taken into account by F_i .

Adding a mass Fig. 2. shows that a mass of 4 grams is added. It is taken into account by F_i , $i = 1, 2$. There is no new calibration. Keep the previous scenarios:

- scenario 3 – simple reference trajectory (see Fig. 7 and 8),
- scenario 4 – complex reference trajectory (see Fig. 9 and 10).

The tracking does not deteriorate.



Fig. 2. Additive masses on the AERO

4 Conclusion

Of course further studies are needed in order to support the thesis of this paper. MFC might not be able to provide a satisfactory fault detection.¹⁰ The rôle of ANNs and RL might then be compelling (see, *e.g.*, [46]).

The epistemological connections of MFC with other existing approaches in AI (see, *e.g.*, [65]), Wiener's *cybernetics* and *expert systems* for instance, will be analyzed elsewhere.

¹⁰ See [25, 39] for fault accommodation.

References

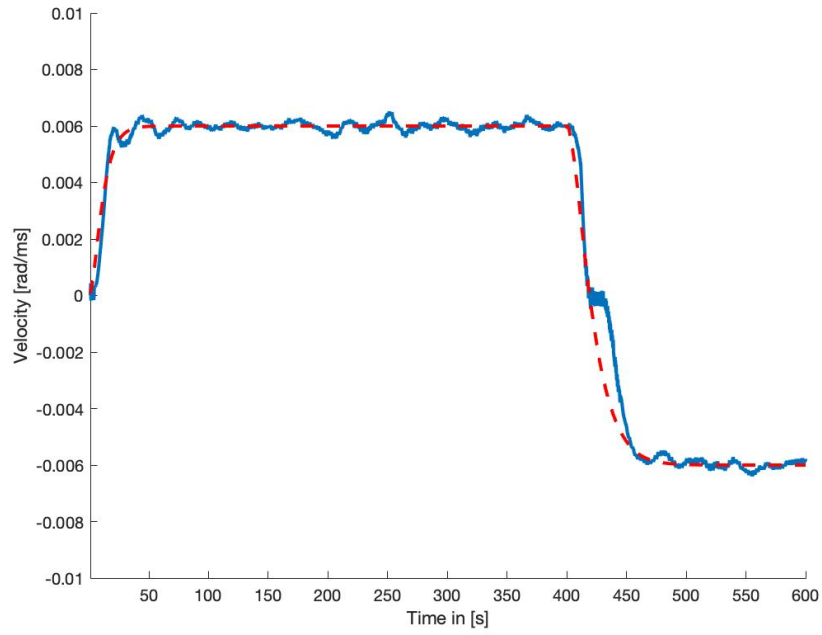
1. Abbaker, A.M.O., Wang, H., Tian, Y., 2020. Voltage control of solid oxide fuel cell power plant based on intelligent proportional integral-adaptive sliding mode control with anti-windup compensator. *Trans. Inst. Measur. Contr.*, 42, 116-130.
2. Abouaïssa, H., Chouraqui, S., 2019. On the control of robot manipulator: A model-free approach. *J. Comput. Sci.*, 31, 6-16.
3. Anderson, C.W., D.C. Hittle, D.C., Katz, A.D., Kretchmar, R.M., 1997. Synthesis of reinforcement learning, neural networks and PI control applied to a simulated heating coil. *Artif. Intell. Engin.*, 11, 421-429.
4. Åström, K.J., Hägglund, T., 2006. *Advanced PID Control*. Instrum. Soc. Amer..
5. Åström, K.J., Murray, R.M., 2008. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press.
6. Bara, O., Fliess, M., Join, C., Day, J., Djouadi, S.M., 2018. Toward a model-free feedback control synthesis for treating acute inflammation. *J. Theoret. Biology*, 448, 26-37.
7. Barth, J.M.O., Condomines, J.-P., Bronz, M., Hattenberger, G., Moschetta, J.-M., Join, C., Fliess, M., 2020. Towards a unified model-free control architecture for tail sitter micro air vehicles: Flight simulation analysis and experimental flights. *AIAA Scitech Forum*, Orlando.
8. Barth, J.M.O., Condomines, J.-P., Bronz, M., Moschetta, J.-M., Join, C., Fliess, M., 2020. Model-free control algorithms for micro air vehicles with transitioning flight capabilities. *Int. J. Micro Air Vehic.*, 12.
<https://doi.org/10.1177/1756829320914264>
9. Baumeister, T., Brunton, S.L., Kutz, J.N., 2018. Deep learning and model predictive control for self-tuning mode-locked lasers. *J. Opt. Soc. Am. B*, 35, 617-626.
10. Bekcheva, M., Fliess, M., Join, C., Moradi, A., Mounier, H., 2018. Meilleure élasticité “nuagique” par commande sans modèle. *ISTE OpenSci.*, 2, 15 pages.
<https://hal.archives-ouvertes.fr/hal-01884806/en/>
11. Beltran-Carbajal, F., Silva-Navarro, G., Trujillo-Franco, L.G., 2018. On-line parametric estimation of damped multiple frequency. *Elec. Power Syst. Res.*, 154, 423-452.
12. Bourbaki, N., 1976. *Fonctions d’une variable réelle*. Hermann. English translation, 2004: *Functions of a Real Variable*. Springer.
13. Brunton, S.L., Noack, B.R., Koumoutsakos, P., 2020. Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.*, 52, 477-508.
14. Bucci, M.A., Semeraro, O., Allauzen, A., Wisniewski, G., Cordier, L., Mathelin, L., 2019. Control of chaotic systems by deep reinforcement learning. *Proc. Roy. Soc. A*, 475, 20190351.
15. Buşoniu, L., de Bruin, T., Tolić, D., Kober, J., Palunko, I., 2018. Reinforcement learning for control: Performance, stability, and deep approximators. *Annual Rev. Contr.*, 46, 8-28.
16. Chen, J., Huang, T.-C., 2004. Applying neural networks to on-line updated PID controllers for nonlinear process control. *J. Process Contr.*, 14, 211-230.
17. Cheon, K., Kim, J., Hamadache, M., Lee, D., 2015. On replacing PID controller with deep learning controller for DC motor system. *J. Automat. Contr. Engin.*, 3, 452-456.
18. Clouatre, M., Thitsa, M., (2020). Shaping 800nm pulses of Yb/Tm co-doped laser: A control theoretic approach. *Ceramics Int.*,
<https://doi.org/10.1016/j.ceramint.2020.03.123>

19. Clouatre, M., Thitsa, M., (2020). Data-driven sliding mode control for pulses of fluorescence in STED microscopy based on Förster resonance energy transfer pairs. *MRS Advances*, <https://doi.org/10.1557/adv.2020.11>
20. Clouatre, M., Thitsa, M., Fliess, M., Join, C., 2020. A robust but easily implementable remote control for quadrotors: Experimental acrobatic flight tests. *Submitted*.
21. Dierks, T., Jagannathan, S., 2010. Neural network output feedback control of robot formations. *IEEE Trans. Syst. Man Cybern.*, 40, 383-399.
22. Duriez, T., Brunton, S.L., Noack, B.R., 2017. *Machine Learning Control – Taming Nonlinear Dynamics and Turbulence*. Springer.
23. Fliess, M., 2006. Analyse non standard du bruit. *C.R. Acad. Sci. Paris Ser. I*, 342, 797-802.
24. Fliess, M., Join, C., 2008. Intelligent PID controllers. 16th Med. Conf. Contr. Automat., Ajaccio. <https://hal.inria.fr/inria-00273279/en/>
25. Fliess, M., Join, C., 2013. Model-free control. *Int. J. Contr.*, 86, 2228-2252.
26. Fliess, M., Sira-Ramírez, H., 2003. An algebraic framework for linear identification. *ESAIM Contr. Optimiz. Calc. Variat.*, 9, 151-168.
27. Fliess, M., Sira-Ramírez, H., 2008. Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques. Garnier, H., Wang, L. (Eds): *Identification of Continuous-time Models from Sampled Data*, Springer, pp. 362-391.
28. Haddar, M., Chaari, R., Baslamisli, S.C., Chaari, F., Haddar, M., 2019. Intelligent PD controller design for active suspension system based on robust model-free control strategy. *J. Mech. Engin. Sci.*, 233, 4863-4880.
29. Han, S., Wang, H., Tian, Y., 2020. A linear discrete-time extended state observer-based intelligent PD controller for a 12 DOFs lower limb exoskeleton LLE-RePA. *Mech. Syst. Sign. Proc.*, 138, 106547.
30. Hatipoglu, K., Olama, H., Xue, Y., 2020. Model-free dynamic voltage control of a synchronous generator-based microgrid. *IEEE Innov. Smart Grid Techno. Conf.*, Washington.
31. Hong, Y., Yang, W., Jiang, B., Yan, X.-G., 2020. A novel multi-agent model-free control for state-of-charge balancing between distributed battery energy storage systems. *IEEE Trans. Emerg. Topics Comput. Intel.*, doi: 10.1109/TETCI.2020.2978434
32. Hwangbo, J., Sa, I., Siegwart, R., Hutter, M., 2017. Control of a quadrotor with reinforcement learning. *IEEE Robot. Automat. Lett.*, 2, 2096-2103.
33. Join, C., Chaxel, F., Fliess, M., 2013. “Intelligent” controllers on cheap and small programmable devices. 2nd Int. Conf. Contr. Fault-Tolerant Syst., Nice. <https://hal.archives-ouvertes.fr/hal-00845795/en/>
34. Join, C., Fliess, M., Chaxel, F., 2020. Model-Free Control as a Service in the Industrial Internet of Things: Packet loss and latency issues via preliminary experiments. 16th Med. Conf. Contr. Automat., Saint-Raphaël. <https://hal.archives-ouvertes.fr/hal-02546750/en/>
35. Kahn, S.G., Hermann, G., Lewis, F.L., Pipe, T., Melhuish, C., 2012. Reinforcement learning and optimal adaptive control: An overview and implementation examples. *Annu. Rev. Contr.*, 36, 42-52.
36. Kizir, S., Bingül, Z., 2019. Design and development of a Steward platform assisted and navigated transphenoidal surgery. *Turk. J. Elec. Eng. Comp. Sci.*, 27, 961-972.
37. Kolmogorov, A.N., Fomin, S.V., 1957 & 1961. *Elements of the Theory of Functions and Functional Analysis*, vol. 1 & 2 (translated from the Russian). Graylock.

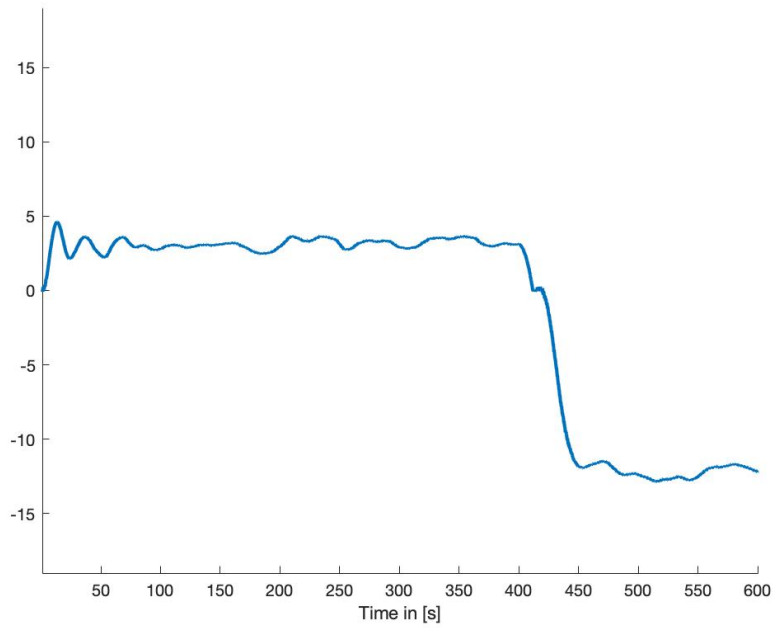
38. Kiumarsi, B., Vamvoudakis, K.G., Modares, H., Lewis, F.L., 2018. Optimal and autonomous control using reinforcement learning: A survey. *IEEE Trans. Neural Netw. Learn. Syst.*, 29, 2042-2062.
39. Lafont, F., Balmat, J.-F., Pessel, N., Fliess, M., 2015. A model-free control strategy for an experimental greenhouse with an application to fault accommodation. *Comput. Electron. Agricul.*, 110, 139-149.
40. Lambert, N.O., Drew, D.S., Yaconelli, J., Levine, S., Calandra, R., Pister, K.S.J., 2019. Low-level control of a quadrotor with deep model-based reinforcement learning. *IEEE Robot. Automat. Lett.*, 4, 4224-4230.
41. Le Cun, Y., 2019. *Quand la machine apprend*. Odile Jacob, 2019.
42. LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature*, 521, 436-444.
43. Li, S., Zhang, Y., 2018. *Neural Networks for Cooperative Control of Multiple Robot Arms*. Springer.
44. Lucia, S., Karg, B., 2018. A deep learning-based approach to robust nonlinear model predictive control. *IFAC PapersOnLine*, 51-20, 511-516.
45. Luo, B., Liu, D., Huang, T., Wang, D., 2016. Model-free optimal tracking control via critic-only Q-learning. *IEEE Trans. Neural Netw. Learn. Syst.*, 27, 2134-2144.
46. Lv, F., Wen, C., Bao, Z., Liu, M., 2016. Fault diagnosis based on deep learning. *Amer. Contr. Conf.*, Boston.
47. N. Ma, G. Song, H.-J. Lee. Position control of shape memory alloy actuators with internal electrical resistance feedback using neural networks. *Smart Mater. Struct.*, 13, 777-783, 2004.
48. Matni, N., Proutiere, A., Rantzer, A., Tu, S., 2019. From self-tuning regulators to reinforcement learning and back again. *58th Conf. Decis. Contr.*, Nice, 2019.
49. Matni, N., Tu, S., 2019. A tutorial on concentration bounds for system identification. *58th Conf. Decis. Contr.*, Nice.
50. Menhour L., d'Andréa-Novel, B., Fliess, M., Gruyer, D., Mounier, H., 2018. An efficient model-free setting for longitudinal and lateral vehicle control: Validation through the interconnected Pro-SiVIC/RTMaps. *IEEE Trans. Intel. Transp. Syst.*, 19, 461-475.
51. Michailidis, I.T., Schild, T., Sangi, R., Michailidis, P., Korkas, C., Fütterer, J., Müller, D., Kosmatopoulos, E.B., 2018. Energy-efficient HVAC management using cooperative, self-trained, control agents: A real-life German building case study. *App. Ener.*, 211, 113-125.
52. Miller III, W.T., Sutton, R.S., Werbos, P.J., 1990 (Eds). *Neural Networks for Control*. MIT Press.
53. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. *Nature*, 518, 529-533.
54. Moe, S., Rustand, A.M., Hanssen, K.G., 2018. Machine learning in control systems: An overview of the state of the art. M. Bramer, M. Petridis (Eds): *Artificial Intelligence XXXV, Lect. Notes Artif. Intel.* 11311, pp. 250-264, Springer.
55. N'Doye, I., Asiri, S., Aloufi, A., Al-Awan, A., Laleg-Kirati, T.-M., 2020. Intelligent proportional-integral-derivative control-based modulating functions for laser beam pointing and stabilization, *IEEE Trans. Contr. Syst. Techno.*, 28,1001-1008.
56. Nicol, C., Macnab, C.J.B., Ramirez-Serrano, A., 2008. Robust neural network control of a quadrotor helicopter. *Canad. Conf. Elec. Comput. Engin.*, Niagara Falls, 2008.

57. Plumejeau, B., Delprat, S., Keirsbulck, L., Lippert, M., Abassi, W., 2019. Ultra-local model-based control of the square-back Ahmed body wake flow. *Phys. Fluids*, 31, 085103.
58. Qin, Z.-C., Xin, Y., Sun, J.-Q., 2020. Dual-loop robust attitude control for an aerodynamic system with unknown dynamic model: algorithm and experimental validation. *IEEE Access*, 8, 36582-36594.
59. Qu, S.T., 2019. Unmanned powered paraglider flight path control based on PID neural network. *IOP Conf. Ser. Mater. Sci. Eng.*, 470, 012008.
60. Rabault, J., Kuchta, M., Jensen, A., Réglade, U., Cerardi, N., 2019. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *J. Fluid Mech.*, 865, 281-302.
61. Radac, M.-B., Precup, R.-E., Roman, R.-C., 2017. Model-free control performance improvement using virtual reference feedback tuning and reinforcement Q-learning. *Int. J. Syst. Sci.*, 48, 1071-1083.
62. Rampazzo, M., Tognin, D., Pagan, M., Carniello, L., Beghi, A., 2019. Modelling, simulation and real-time control of a laboratory tide generation *Contr. Eng. Pract.*, 83, 165-175.
63. Recht, B., 2019. A tour of reinforcement learning: The view from continuous control. *Annu. Rev. Contr. Robot. Autonom. Syst.*, 2, 253-279.
64. Rocher, V., Join, C., Mottelet, S., Bernier, J., Rechdaoui-Guerin, S., Azimi, S., Lessard, P., Pauss, A., Fliess, M., 2018. La production de nitrites lors de la dénitrification des eaux usées par biofiltration - stratégie de contrôle et de réduction des concentrations résiduelles. *J. Water Sci.*, 31, 61-73.
65. Russel, S., Norvig, P., 2016. *Artificial Intelligence – A Modern Approach* (3rd ed.). Pearson.
66. Sancak, C., Yamac, F., Itik, M., Alici, G., 2019. Model-free control of an electro-active polymer actuator. *Mater. Res. Expr.*, 6, 055309.
67. Sejnowski, T.J., 2020. The unreasonable effectiveness of deep learning in artificial intelligence. *Proc. Nat. Acad. Sci.*, <https://doi.org/10.1073/pnas.1907373117>
68. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D., 2016. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529, 484-489.
69. Sira-Ramírez, H., García-Rodríguez, C., Cortès-Romero, J., Luviano-Juárez, A., 2014. *Algebraic Identification and Estimation Methods in Feedback Control Systems*. Wiley.
70. Stalph, P., 2014. *Analysis and Design of Machine Learning Techniques* Springer.
71. Sugiyama, M., 2015. *Statistical Reinforcement Learning – Modern Machine Learning Approaches* CRC Press.
72. Sutton, R.S., Barto, A.G., 2018. *Reinforcement Learning* (2nd ed.). MIT Press.
73. Ticherfatine, M., Zhu, Q., 2018. Fast ferry smoothing motion via intelligent PD controller. *J. Marine. Sci. App.*, 17, 273-279.
74. Villagra, J., Join, C., Haber, R., Fliess, M., 2020. Model-free control for machine tool systems. 21st World IFAC, Berlin. <https://hal.archives-ouvertes.fr/hal-02568336/en/>
75. Wang, Y., Li, H., Liu, R., Yang, L., Wang, X., 2020. Modulated model-free predictive control with minimum switching losses for PMSM drive system. *IEEE Access*, 8, 20942-20953.

76. Wang, H., Li, S., Tian, Y., Aitouche, A., 2016. Intelligent proportional differential neural network control for unknown nonlinear system. *Stud. Informat. Contr.*, 25, 445-452.
77. Wang, Y., Velswamy, K., Huang, B., 2018. A novel approach to feedback control via deep reinforcement learning. *IFAC PapersOnLine*, 51-18, 31-36.
78. Wang, Z., Wang, J., 2020. Ultra-local model predictive control: A model-free approach and its application on automated vehicle trajectory tracking. *Contr. Eng. Pract.*, 101, 104482.
79. Waslander, S.L., Hoffmann, G.M., Jung Soon Jang ; Tomlin, C.J., 2005. Multi-agent quadrotor testbed control design: integral sliding mode vs. reinforcement learning. *IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Edmonton.
80. Wu, Y., Song, Q., Yang, X., 2007. Robust recurrent neural network control of biped robot. *J. Intell. Robot. Syst.*, 49, 151-169.
81. Yang, H., Liu, C., Shi, J., Zhong, G., 2019. Development and control of four-wheel independent driving and modular steering electric vehicles for improved maneuverability limits. *SAE Tech. Paper*, 2019-01-0459.
82. Yosida, K., 1984. *Operational Calculus* (translated from the Japanese). Springer.
83. Zhang, Y., Ding, S.X., Yang, Y., Li, L., 2015. Data-driven design of two-degree-of-freedom controllers using reinforcement learning techniques. *IET Contr. Theory Appl.*, 9, 1011-1021.
84. Zhang, J., Jin, J., Huang, L., 2020. Model-free predictive current control of PMSM drives based on extended state observer using ultra-local model. *IEEE Trans. Indus. Electron.*, doi: 10.1109/TIE.2020.2970660
85. Zhang, X., Li, M., Ding, H., Yao, X., 2019. Data-driven tuning of feedforward controller structured with infinite impulse response filter via iterative learning control. *IET Contr. Theory Appl.*, 13, 1062-1070.
86. Zhang, Y., Liu, X., Liu, J., Rodriguez, J., Garcia, C., 2020. Model-free predictive current control of power converters based on ultra-local model. *IEEE Int. Conf. Indust.Techno.*, Buenos Aires.
87. Zhang, X., Wang, H., Tian, Y., Peyrodie, L., Wang, X., 2018. Model-free based neural network control with time-delay estimation for lower extremity exoskeleton. *Neurocomput.*, 272, 178-188.
88. Zhang, X.-M., Wei, Z., Asad, R., Yang, X.-C., Wang, X., 2019. When does reinforcement learning stand out in in control? A comparative study on state representation. *npj Quantum Inf.*, 5. <https://doi.org/10.1038/s41534-019-0201-8>
89. Zhu, L., Ma, J., Wang, S., 2019. Deep neural networks based real-time optimal control for lunar landing. *IOP Conf. Ser. Mater. Sci. Engin.*, 608, 012045.

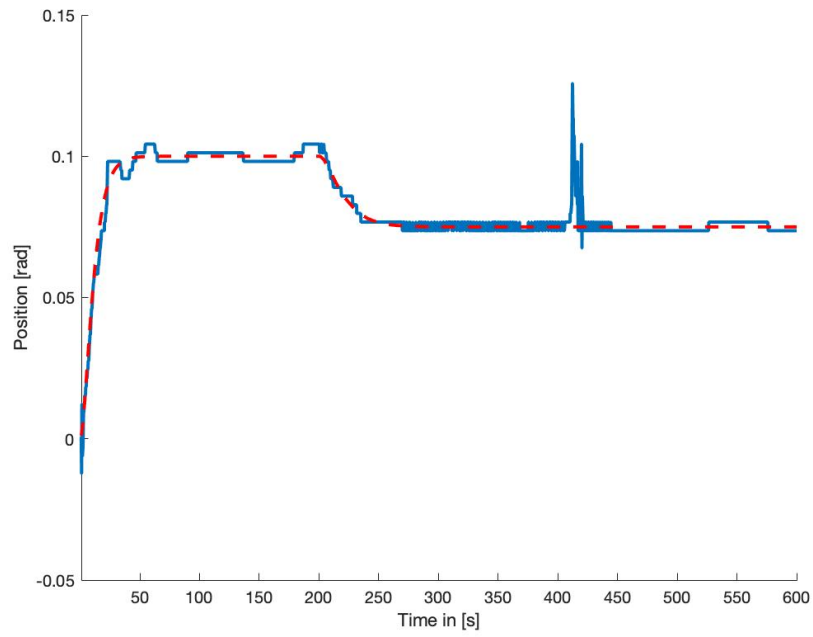


(a) Azimuth velocity (blue —), reference trajectory (red - -)

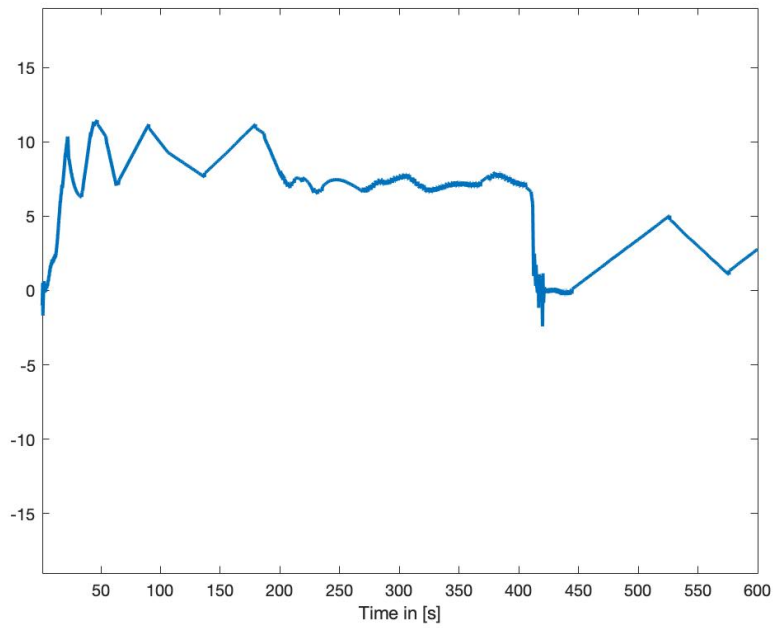


(b) Control u_1

Fig. 3. Scenario 1: Azimuth

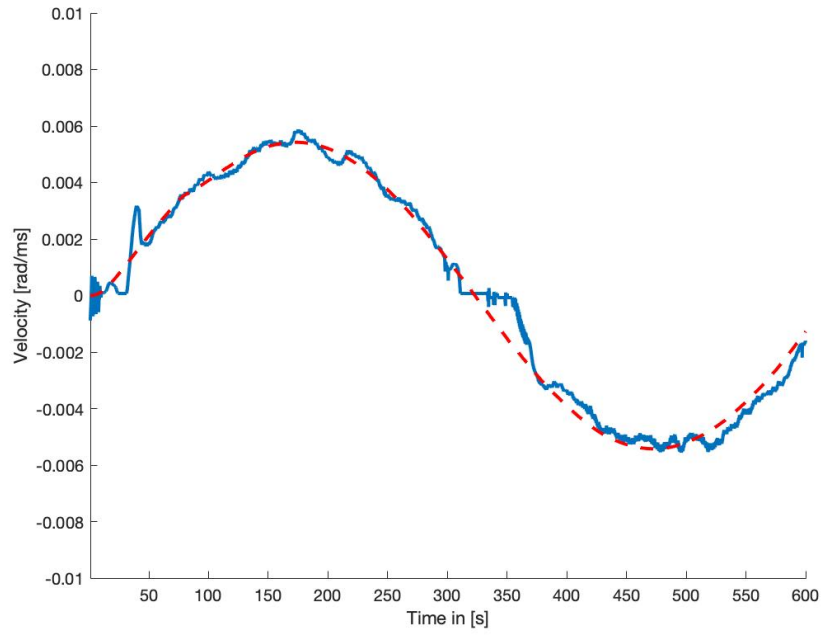


(a) Pitch position (blue --), reference trajectory (red --)

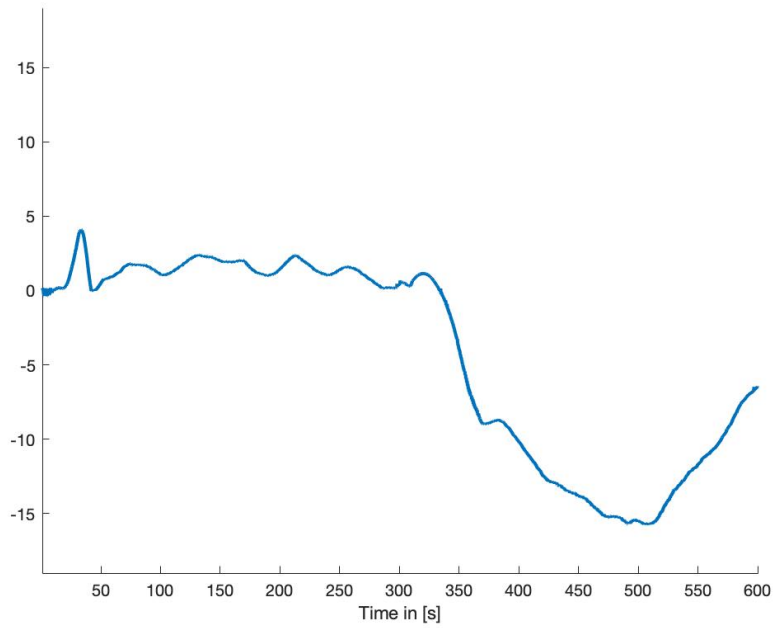


(b) Control u_2

Fig. 4. Scenario 1: Pitch

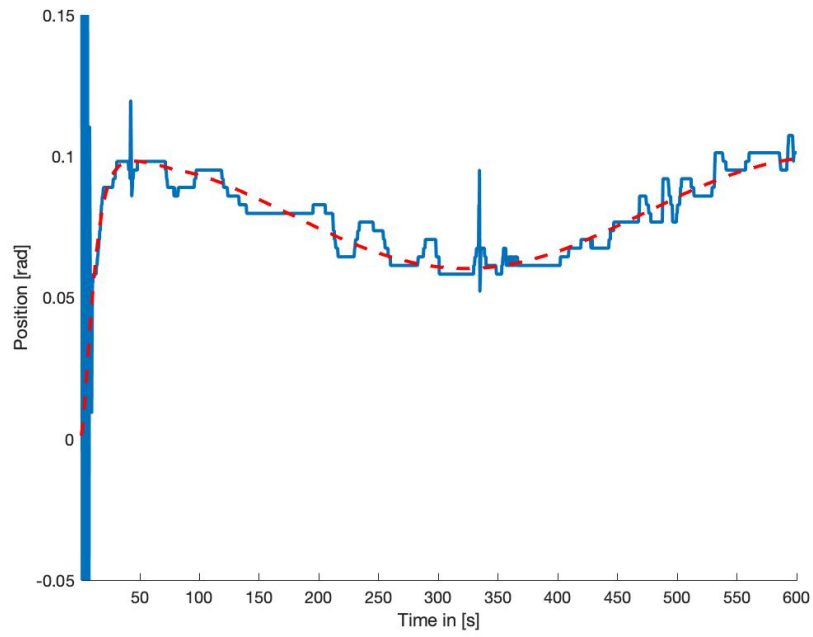


(a) Azimuth velocity (blue —), reference trajectory (red - -)

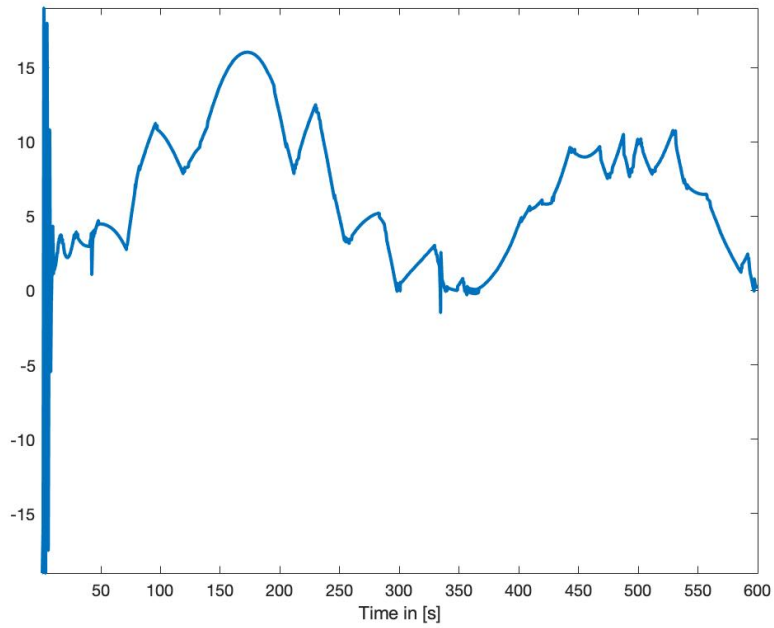


(b) Control u_1

Fig. 5. Scenario 2: Azimuth

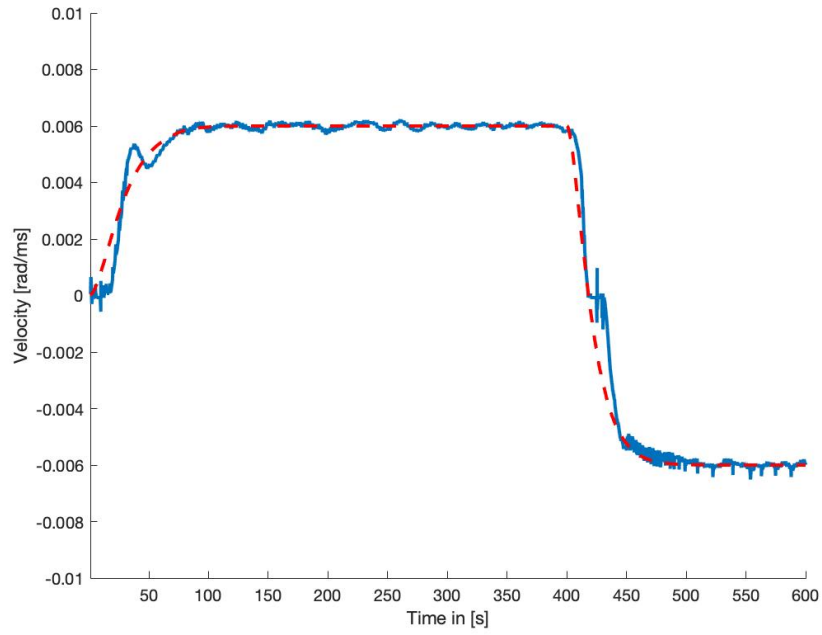


(a) Pitch position (blue —), reference trajectory (red - -)

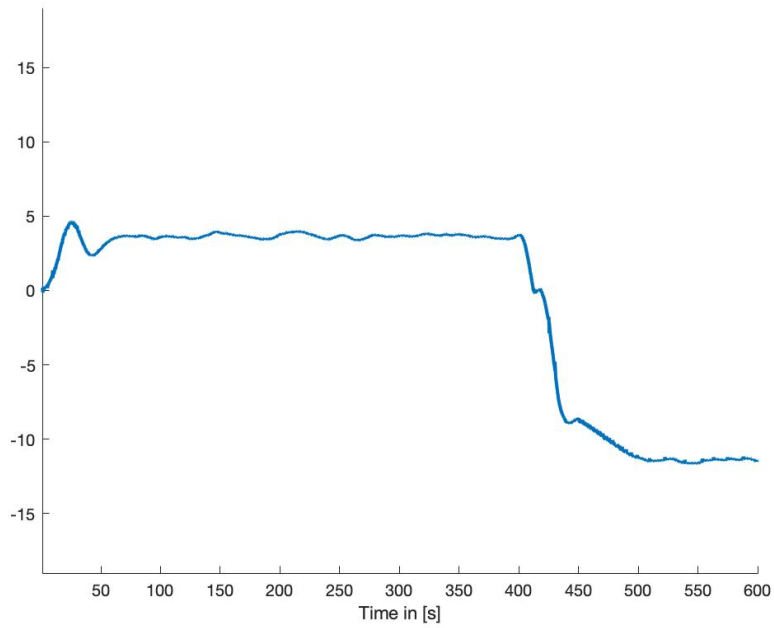


(b) Control u_2

Fig. 6. Scenario 2: Pitch

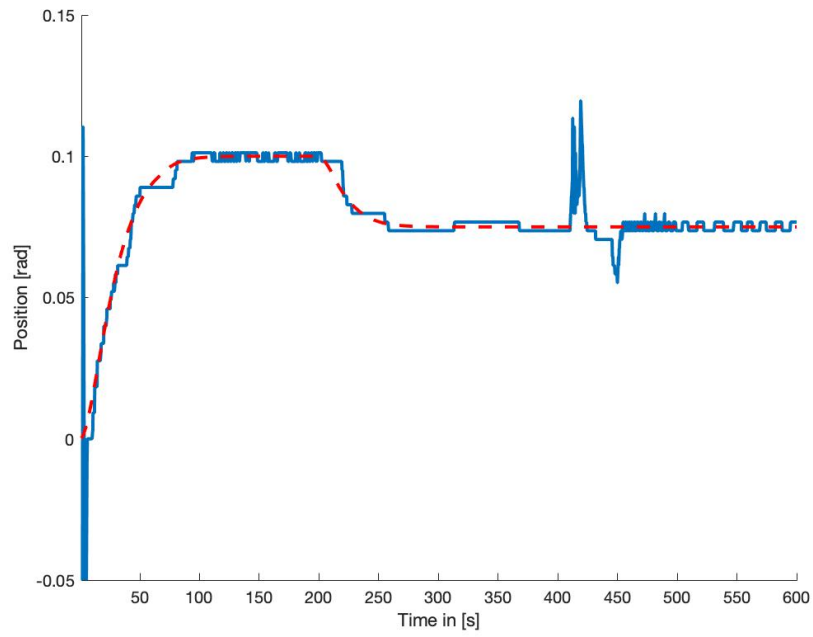


(a) Azimuth velocity (blue --), reference trajectory (red --)

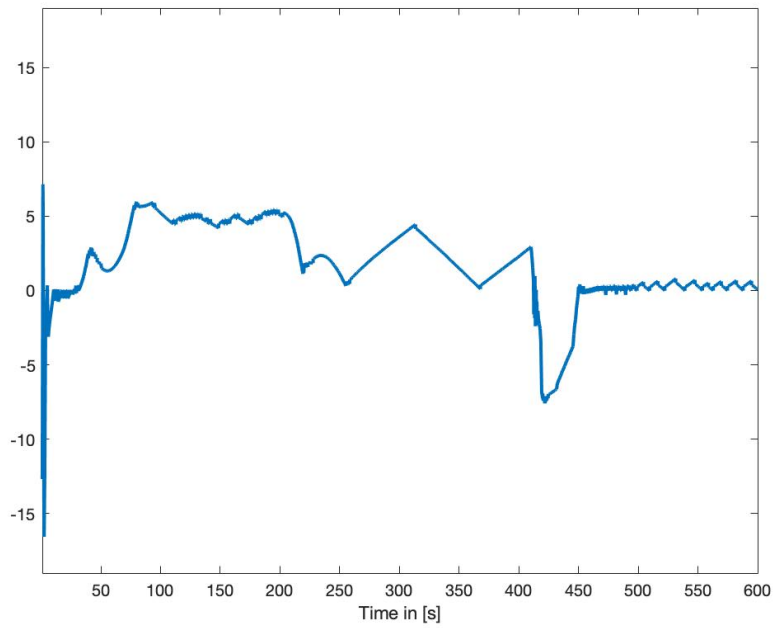


(b) Control u_1

Fig. 7. Scenario 3: Azimuth

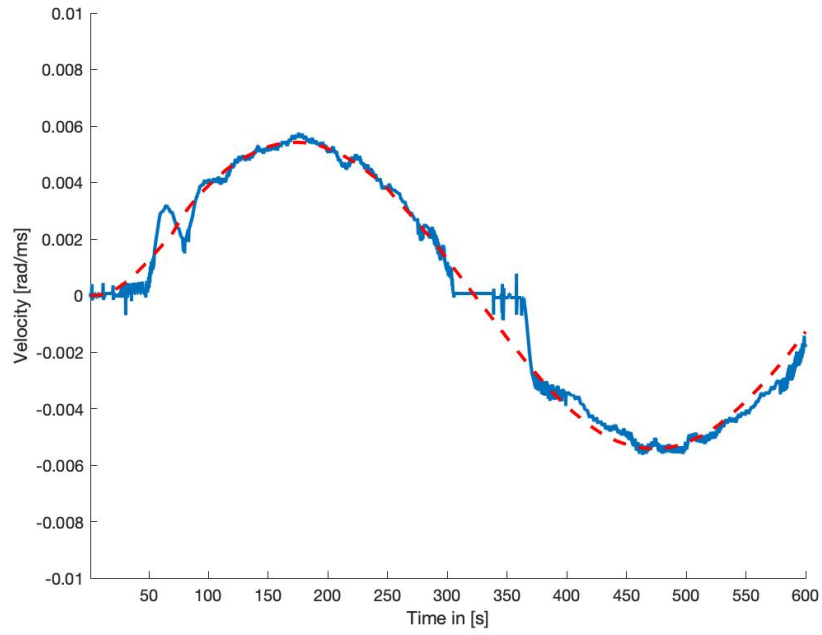


(a) Pitch position (blue --), reference trajectory (red --)

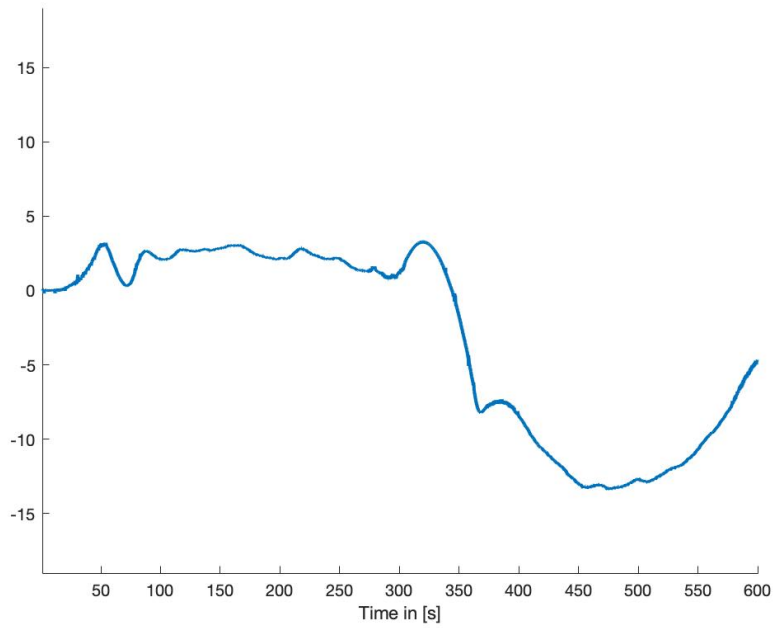


(b) Control u_2

Fig. 8. Scenario 3: Pitch

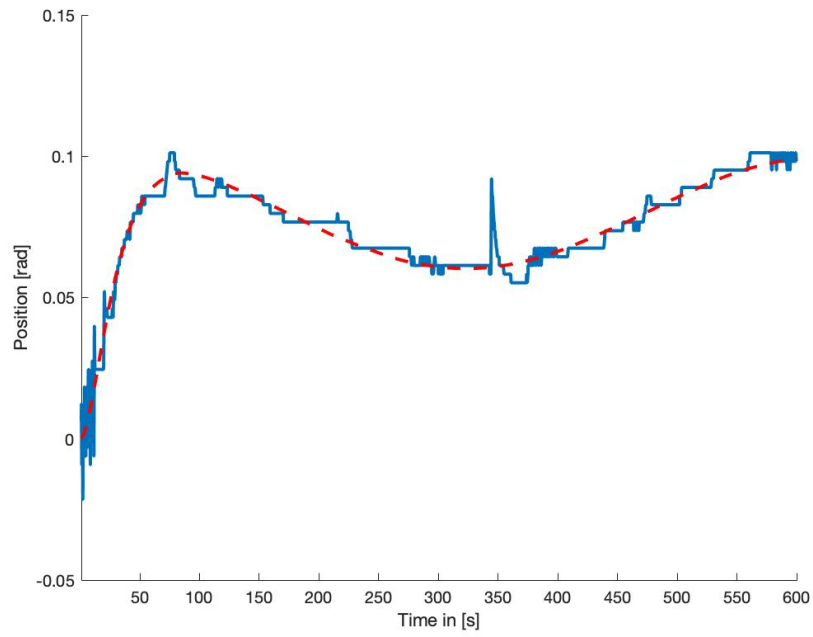


(a) Azimuth velocity (blue —), reference trajectory (red - -)

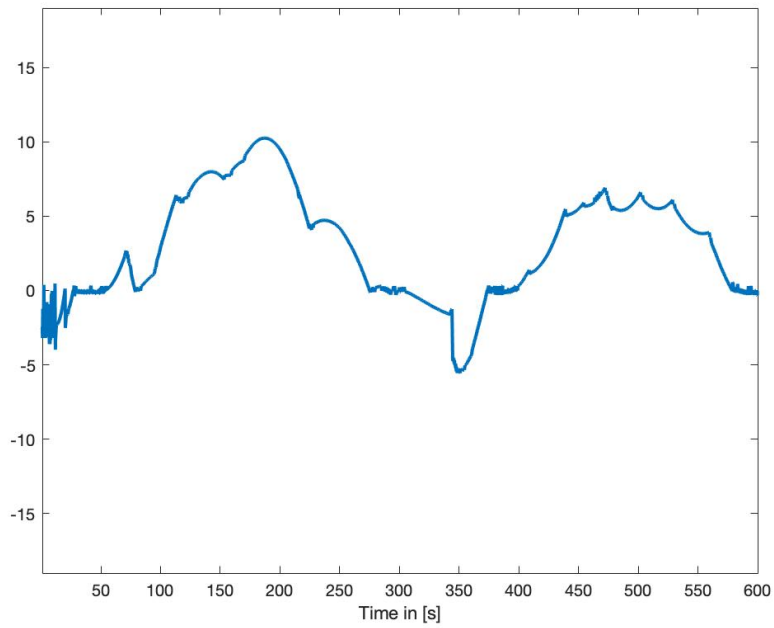


(b) Control u_1

Fig. 9. Scenario 4: Azimuth



(a) Pitch position (blue --), reference trajectory (red --)



(b) Control u_2

Fig. 10. Scenario 4: Pitch