M $^{\text{odeling}}$ Mc $^{\text{omputing}}$ $_{\text{athematical}}$

Fuzzy controller, designed by reinforcement learning, for vehicle traction system application

Demkiv L. I.[1], Lozynskyy A. O.[1], Vantsevich V. V.[2], Gorsich D. J.[3],
Lytvyn V. V.[1], Klos S. R.[1], Letherwood M. D.[4]

[1]*Lviv Polytechnic National University,*
12 S. Bandera str., 79013, Lviv, Ukraine
[2]*University of Alabama at Birmingham,*
Birmingham, AL 35294, U.S.A.
[3]*US Army CCDC Ground Vehicle Systems Center,*
Warren, MI, U.S.A.
[4]*Alion Science and Technilogy, Ground Vehicle Systems Center,*
Warren, MI, U.S.A.

In this article, a fuzzy controller tuned by reinforcement learning is proposed. The developed algorithm utilizes a fuzzy logic theory and a reinforcement learning for fine-tuning parameters of the membership function for the fuzzy controller. Apart from the fuzzy controller developed, a fuzzy corrector of reference input (set-point) signal to the controller is applied. The fuzzy corrector changes the input (reference) signal of the system and takes into account an original reference input and type of external disturbances. Thus, the designed fuzzy control that is tuned by reinforcement learning is capable to ensure the stable, optimal, and safe performance of the system and takes into account external disturbances. To verify the performance of the proposed controller, the adaptive fuzzy controller tuned by reinforcement learning is applied to the mathematical model of a wheel locomotion module of an electric vehicle to advance a traction control system. Therefore, the effectiveness of the proposed adaptive fuzzy controller is proven through the simulation results.

1. Introduction

In a past few decades, various control techniques have been proposed, along with classical, optimal or intelligent control methods. Although the PID control is widely used for its simplicity and predictability, alternative methods are required to deal with complex/nonlinear systems or processes, external disturbances presence, to ensure desired control performance, etc. Indeed, a PID controller, with its known tuning process and computational simplicity, is characterized by its constant tuning parameters and, in general, cannot provide adaptivity or robustness to external disturbances and optimal control for nonlinear systems.

Due to the demand for optimality of the controller, optimal control methods, whose aim to find a control policy that optimizes a given criterion for a given dynamical system, might be an alternative that provides the robustness to the closed loop transients. However, optimal control problem requires solving the Hamilton–Jacobi–Bellman equation which usually cannot be obtained for nonlinear systems [1].

Therefore, in recent past, to control the nonlinear systems, the model predictive control (MPC) and intelligent control methods are widely investigated, including fuzzy logic control, learning control methods, etc. [2–7]. These control methods can be designed to provide adaptivity to external disturbances or system parameters change and, thus, design an adaptive control system. Adaptive control systems ensure the stability and performance of a closed-loop system in the presence of adverse conditions. Unlike robust control systems, they are not tuned into the worst-case scenario and they can constantly improve their performance through learning [8].

In this work, the fuzzy logic control is chosen because of its adaptive tuning capabilities, ability to deal with uncertainty and non-linearity while also its rapid operation [9–12]. Besides, a reinforcement learning algorithm, is utilized for tuning the parameters of membership function of the fuzzy controller, which minimizes the complexity of fuzzy controller tuning and synthesis. To further enhance control transients performance while presence of external disturbances, the fuzzy corrector of controller reference input signal is utilized. Thus, the whole control system which comprises fuzzy controller fine-tuned by reinforcement learning and fuzzy corrector is capable to provide a stable and optimal control policy which is adaptive to external disturbances.

In order to verify the developed control system, it is applied to a mathematical model of a wheel locomotion module of a vehicle that is one-quarter model of the vehicle. Specifically, the designed controller is synthesized for traction control of an off-road vehicle's wheel which is driven by an in-wheel electrical motor. In order to implement the control system for such vehicular application, the additional research is carried out to analyze the off-road dynamics of the wheel locomotion module and wheel/vehicle traction control system.

Improvement of the wheel/vehicle traction performance in off-road condition is a challenging task. Typically, the traction control algorithms are tuned and calibrated for asphalt road conditions that might cause issues when operating on different road/terrain conditions. The uncertain road conditions impact the performance of the control system significantly that opens the challenge for proper vehicle's parameters estimation and control [14]. For off-road vehicles, which commonly are heavy-vehicles for different applications such as agriculture, forestry, constriction, military, the essential factors of traction control system design are energy efficiency and mobility. The tire slippage optimization is usually connected with the aim of energy efficiency optimization of traction and main losses during traction control are due to high values of tire slippage. Mobility of the vehicle characterizes its ability to go or not to go in certain road/terrain conditions. The control methods for off-road traction control have to fulfill the task of energy efficient motion without immobilization operation in different/stochastic terrain conditions [13]. Therefore, an off-road traction control should provide fast and agile control action decisions that are adaptive to particular road/terrain conditions.

Authors in [15] also pay close attention to control methods, pointing out that this is a significant factor for the off-road traction control system. In the study, the authors put the problem of the efficiency and robustness of the traction system as one of the most important for off-road vehicles. In the current paper, authors want to draw attention to the fact that for off-road vehicles special treatment should be paid to adaptability rather than robustness, as off-road control systems must adapt as quickly as possible to rapid and stochastic terrain condition changes.

There are different studies dealing with the traction control systems for off-road mobility and energy efficiency improvement [16–18]. Most of them agreed that the adaptability of traction system for off-road conditions causes improvement in energy efficiency. In [16] the adaptive set-point value is implemented for traction control system. The authors stated the due to changes in terrain conditions, all the approaches with a fixed set-point are sub-optimal and might lead to low energy efficiency. An adaptive controller for off-road traction system designed taking into account the change of controller's set-point is proposed in [18].

The main contribution of this work is to introduce the adaptive fuzzy controller for stable and optimal performance control of dynamical systems and to verify the controller on the example of the wheel locomotion module. The contribution of the applied off-road traction controller is to address the

wheel angular velocity control taking into account the dynamics of e-motor and wheel while presence of an external stochastic uncertainties represented by wheel load. The key features are that the fine-tuning of the controller is carried out by reinforcement learning algorithm and the set-point for the controller is adaptive to different road/terrain conditions and implemented as a fuzzy corrector.

The reminder of the article starts with the background on reinforcement learning and fuzzy logic controller. The details on the control method synthesis are presented in Section 3. The control method proposed in this study is applied to the mathematical model of the wheel locomotion module described in Section 4.1 while the computational results of the proposed control implementation to the mathematical model are presented in Section 5.

2. Background

2.1. Reinforcement learning: Q-learning algorithm

Reinforcement learning is a subfield of machine learning. Learning-based applications in ground vehicles is an emerging and promising technique with a robust ability of the self-adaptation to changes of both system states and external disturbances. Even though some learning algorithms need to have an off-line learning phase, they still may be considered for real-time applications.

Reinforcement learning uses the data collected from past experiences in order to develop strategies for the future actions. In general, reinforcement learning algorithms and, in particular, Q-learning algorithms utilize a learning approach of "agent–environment" interactions to maximize a reward value that is observed from the environment (see Fig. 1). The environment herein is a control plant, i.e. a mathematical model of a physical object. The agent is an algorithm that analyses responses from the environment and produces control actions based on the acquired knowledge about the environment. The responses from the environment are represented by state and reward variables. The states are variables that are chosen to sufficiently characterize the environment and are selected by the algorithm designer. The reward variables should be scalars that are chosen to clearly depict a desired behavior of the environment by giving a positive reward in case of the desired behavior and a negative reward (punishment) for an undesired behavior. The term of a discounted cumulative reward means the sum of discounted immediate rewards over a simulation as shown by Eq. (1)

$$R_{cum} = \sum_{t=0}^{t=t_{sim}} \gamma^t \cdot r_t \tag{1}$$

where, the discount factor, γ, is assigned to range $[0; 1]$ to characterize earlier rewards as more valuable than latest ones, r_t is the reward at each simulation step, and t_{sim} is the time of simulation.
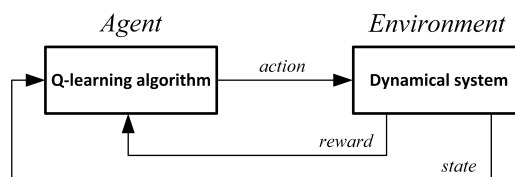

Fig. 1. Agent-environment interaction in Q-learning.

The Q-learning algorithm is implemented by utilizing an action-value function update ($Q(s_t, a_t)$-update) that is also known as a value iteration method presented in [19]

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha\big(r_t + \gamma \max Q(s_{t+1}, A) - Q(s_t, a_t)\big) \tag{2}$$

where, $Q(s_{t+1}, A)$ is a matrix, each element of which represents a state-action value, s_t and s_{t+1} are state variables at time t and $t + 1$ respectively, a_t is the chosen action at time t, A is the vector of all possible actions, α is the learning rate parameter.

2.2. Fuzzy logic controller fundamentals

Fuzzy logic (FL) has been employed in variety of engineering systems, including ground vehicles. FL algorithms are usually characterized by their flexibility and ability to model nonlinear functions with a required accuracy while being tolerant to imprecise data. Regardless the absence of techniques for optimal tuning of fuzzy logic systems, they can be tuned for high precision and rapid operation by utilizing expert knowledge or learning algorithms.

Fuzzy logic defines a mapping of the input data into a scalar output, using fuzzy rules. Generally, a FL reasoning includes a fuzzification, inference, and defuzzification stages, which are shown in Fig. 2.
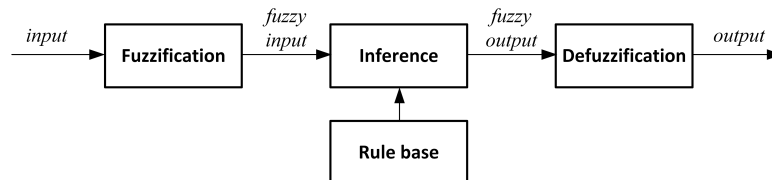

Fig. 2. Fuzzy logic system stages representation.

During the fuzzification, a fuzzy logic system utilizes input values to determine a degree to which the inputs belong to each of the fuzzy sets through membership functions. Fuzzy membership functions represent fuzzy term sets, which, in turn, represent the linguistic terms such as small, large, slow, fast, etc. Typically, linear (triangular, trapezoidal) and nonlinear (gaussian, sigmoidal) shapes of membership functions are utilized at the fuzzification stage as presented in Figure 3. In Figure 3a three membership functions (designated as μ_1, μ_2 and μ_3) are presented for each of the fuzzy set. For small inputs a R-trapezoidal membership function is used, triangular membership function is used for medium inputs and L-trapezoidal membership function is used for large inputs. Similarly, sigmoid membership functions (designated as μ_4 and μ_5) are used for fuzzy sets "slow" and "fast" respectively, and are demonstrated in Figure 3b.
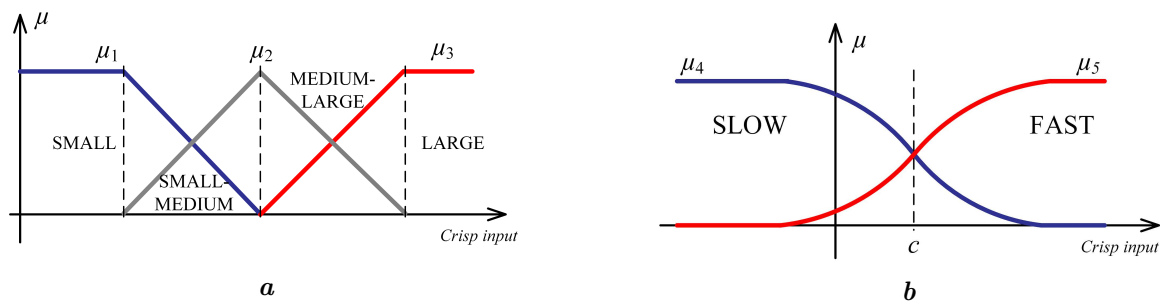

Fig. 3. Membership functions of the fuzzy logic controller.

During fuzzification, the degree of membership to particular fuzzy set (small, large, slow, fast, etc.) is assigned to every crisp input value. In Figure 3b the input signal is fuzzified by using sigmoid membership function as described as follows

$$\mu_4(\varepsilon, a, c) = \frac{e^{-a(\varepsilon-c)}}{1 - e^{-a(\varepsilon-c)}} \tag{3}$$

where, c is the parameter that specifies the value of the control error, ε, at which the value of membership function is equal to 0.5 and a specifies the parameter that defines the degree of the membership function slope at point $\mu_4(\omega_{err}, a, c) = 0.5$ (see Figure 3b). The second membership function is defined as $\mu_5 = 1 - \mu_4$.

The rule base for inference stage contains linguistic rules that are provided by experts. It is also possible to extract rules from numeric data. The inference phase defines the output fuzzy sets by

determining a degree, to which the antecedent is satisfied for each rule. In this study, the Takagi–Sugeno fuzzy model is utilized. The R^l rule of the rule base is described as follows

$$R^l: \text{IF } (|\text{Input}| \in A_m) \text{ THEN } u_{c_k} = f_k \cdot \sum_i k_i x_i \qquad (4)$$

where, Input is the input signal of the system, A_m is the set of terms of the input variable (e.g. large, small, etc.), x_i are system states, k_i are the feedback gains synthesized with the use the characteristic polynomial of the system, $f_k \cdot \sum_i k_i x_i$ is the output function of the rule, which, in this study, is a polynomial of a standard linear form that depends on the state vector of the system, $\bar{x}_i$. Therefore, to obtain fuzzy output of the rule, u_{c_k}, the characteristic polynomials (see Table 1) can be used. The ω_0 is a compound root of a closed loop system. The value of ω_0 depends on the system under investigation.

The crisp output of the FLC can be obtained by using the centroid defuzzification method as follows

$$u_c = \frac{\sum_k \mu_k \cdot u_{c_k}}{\sum_k \mu_k} \qquad (5)$$

where, u_c is the output of the system, μ_k is the rule's activation degree.

In this study, a fuzzy logic polynomials controller is synthesized for a dynamical system. The designed fuzzy logic controller (FLC) is tuned by applying a reinforcement learning algorithm as described in the following section.

3. An adaptive fuzzy controller design

The ultimate aim of this paper is to develop the controller that is able both to increase transients of a dynamical system without significant overshoot and to provide energy efficient behavior of the system. The developed controller should fulfill its goals for multiple sets of external disturbances. The last requirement is more challenging for environment with significant power losses of the system caused by interaction with the environment. Application of fuzzy theory opens possibilities for creating intelligent and adaptive control systems.

3.1. A polynomial fuzzy controller

As stated before, the aim of fuzzy polynomial controller is to control the output signal of the system concerning different external disturbances. In this study, it is proposed to synthesize the fuzzy controller which is able to control the output signal adaptively with respect to control input demand, i.e. engaging slow or fast transient dynamics if needed. Hence, the designed fuzzy polynomial controller ensures smooth switching between two subsystems (slow and fast), which characteristics polynomial are tuned to activate closed loop control system with different transients. The main feature of the fuzzy controller is to engage fast dynamics while output signal is far from the steady-state and engage slow dynamics while output signal approaches the steady-state. The details on synthesis of the developed fuzzy polynomial controller is presented onward.

The Tagaki–Sugeno type fuzzy logic controller is utilized in this study. The input (crisp value) of fuzzy logic controller is the absolute value of the error signal, ω_{err}, i.e. absolute of a difference between desired and actual value of the output signal of the system. This input signal is fuzzified by using sigmoid membership function as described earlier in Eq. (3).

The rule base R^l of the proposed controller is described as follows from Eq. (4). To obtain the fuzzy output, u_{c_k}, that is the output function of the rule, the following standard polynomial forms are used: 4th order Binomial and 4th order Bessel polynomial forms (see Table 1).

And finally, the crisp output from the fuzzy output, u_{c_k}, the centroid defuzzification method is utilized as described in Eq. (5).

Polynomial standard linear form	2nd order characteristic polynomial coefficients $s^2 + a_1\omega_0 s + a_2(\omega_0)^2$	
	a_1	a_2
Binomial	2	1
Bessel	3	3
Butterworth	1.4142	1
	3rd order characteristic polynomial coefficients $s^3 + a_1\omega_0 s^2 + a_2(\omega_0)^2 s + a_3(\omega_0)$		
	a_1	a_2	a_3
Binomial	3	3	1
Bessel	6	15	15
Butterworth	2	2	1
	4th order characteristic polynomial coefficients $s^4 + a_1\omega_0 s^3 + a_2(\omega_0)^2 s^2 + a_3(\omega_0)^3 s + a_4(\omega_0)^4$			
	a_1	a_2	a_3	a_4
Binomial	4	6	4	1
Bessel	10	45	105	105
Butterworth	2.613	3.414	2.613	1

3.2. Reinforcement learning algorithm for tuning the parameters of membership function

Generally, the fuzzy controller requires an expert knowledge for its tuning or, at least, sufficient experience in fuzzy logic theory. Even in our case when there are only two subdomains and the tuning process can be shortened to matching two appropriate parameters of membership function that are a and c from Eq. (3), it could be hard to find the pair of them which would give the desired control system response. Even more complicated and more important is to construct the methodology of such parameter matching for other control applications. Bearing this in mind, it was decided to implement the algorithm that completes the fine-tuning of the parameters of the membership function. Therefore, in this study, the reinforcement learning algorithm is utilized for tuning the fuzzy polynomial controller.

The developed reinforcement learning algorithm is utilized to adjust the membership function parameters of the fuzzy logic controller exactly as described below. It should be stated that if the parameters of membership function would be changed at each time step of simulation (or at each sensor reading cycle during system operation) the unexpected behavior might appear due to system parameter change or emergence of unexpected external uncertainties. In this regard, the reinforcement learning algorithm that tunes FLC update its action once per episode of learning that, in our case, is one time per simulation which corresponds to the duration of acceleration process.

For the learning method developed, the Q-update formula is utilized from the Q-learning algorithm [19]. As the learning process is implemented once per episode, the environment state is the vector of three parameters that is observed at the end of each episode and is presented as follows

$$s_t = (|\max \varepsilon|, T_{set}, \varepsilon_{ss}) \tag{6}$$

where, $|\max \varepsilon|$ is the absolute value of the maximal overshoot of the output value, ε is the control error, T_{set} is the settling time, i.e. the time it takes to reach and stay within the range of $\pm 2\%$ of the reference signal (different percentage can be used depending on a kind of dynamical system, $\pm 2\%$ is one of a commonly used [20]) and ε_{ss} is the steady-state error at the end of the episode.

Actions: an action is the vector of the parameters of membership function (parameters a and c from sigmoidal membership function in Eq. (3)) that are applied to the fuzzy polynomial controller at the

each episode of learning. During the learning process, the range of change for each of the parameter a and c is from 0 to 10. Thereby, in this study, the finite number of actions are utilized and the complete action space is equal to 10 thousand of different actions (i.e. different a and c pairs).

The reward function after each episode of learning is based on the environment states and on the ITAE criterion for control systems analysis [9, 21] as follows

$$r_t = \begin{cases} 5000 - 100 \cdot |\max \varepsilon| - 100 \cdot T_{set} - 500 \cdot |\varepsilon_{ss}|, & \text{if } |\varepsilon| < 0.1 \text{ for } 1.5 \text{ sec,} \\ \dfrac{\text{ITAE}}{2} - 100 \cdot |\max \varepsilon| - 100 \cdot T_{set} - 500 \cdot |\varepsilon_{ss}|, & \text{if } |\varepsilon| < 0.5, \\ \text{ITAE} - 100 \cdot |\max \varepsilon| - 100 \cdot T_{set} - 500 \cdot |\varepsilon_{ss}|, & \text{else ,} \end{cases} \qquad (7)$$

Computer simulation results of the fine-tuned fuzzy polynomial controller by reinforcement learning algorithm are provided in Section 5.

3.3. A fuzzy corrector of a reference input to fuzzy controller

The aim of the designed fuzzy corrector is to alter the inappropriate reference input signal and to provide smooth input to fuzzy polynomial controller that is adaptive to external uncertainties. The details on the adaptive fuzzy corrector synthesis is provided onward.

In this paper, the fuzzy corrector is utilized to smooth the ramp input signal taking into account the reference input and type of external disturbances. Four different types of stochastic disturbances are studied, which characteristics parameters were taken from [13]. Thereby, the fuzzy corrector implementation stages correlate with the designed fuzzy polynomial controller from Section 3.1. At the first stage the fuzzification is performed by triangle membership function that divides the fuzzy set into four linguistic terms (subsets), namely, small (S), small-medium (SM), medium-large (ML) and large (L) as can be seen in Figure 3a.

The fuzzy corrector utilizes the Takagi–Sugeno fuzzy model, and the output of each rule in rule base which is also a function based on a polynomial of a standard linear form. For the fuzzy corrector synthesis, the output function of the fuzzy rule is defined according to the second order Butterworth polynomial form is utilized (see Table 1). Herein, the compound root, ω_0, corresponds to a certain type of road/terrain conditions and impacts the rate of change of the input signal.

The centroid defuzzification method (see Eq. (5)) is also utilized in the fuzzy corrector to obtain the crisp values of the smoothed reference signal of the input to, previously described, fuzzy polynomial controller.

When the fuzzy corrector is implemented, the whole adaptive controller is completed and fine-tuned (see Figure 4).
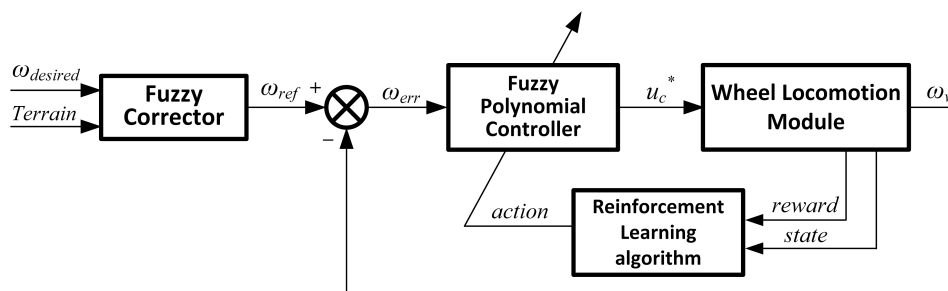

Fig. 4. Block diagram of the developed control system.

The approach described in this section may be applied to multiple types of dynamical systems. In this article, the application of the introduced approach to the wheel locomotion module is investigated onward. In Section 4 the mathematical model of the wheel locomotion module is introduced and in the following section the results of application of adaptive fuzzy logic controller to the case of one-quarter car model are presented.

4. Application of the adaptive fuzzy controller to the off-road traction system in a wheel locomotion module

As the off-road conditions are typically characterized by the stochastic external uncertainties presence and rapid changes of the road/terrain parameters it is important to provide the appropriate control actions in order to ensure safety, energy efficiency and mobility of vehicle motion. It is of particular importance for traction control systems (TCS) that are aimed to prevent immobilization and extra power losses during wheel/vehicle acceleration. This is commonly achieved by brake based controlling of the angular velocity of the wheel for the purpose of preventing significant tire slippage occurrence. Therefore, the adaptive control laws have the prospective advantages for application to the traction systems for off-road vehicles, particularly, when the road/terrain operation conditions might be changed rapidly and frequently. The ultimate goal of the adaptive control algorithms for off-road traction systems is to adapt control decision making to each of the working conditions and ensure the mobility performance and prevent energy losses during its operation.

4.1. Wheel locomotion module dynamics

As a concept of modularity in design, a wheel locomotion module is the assembly unit of the 4x4 electric vehicle with the individual e-motors. Such wheel locomotion module can be considered as a one quarter of the vehicle that comprises the suspension, e-drive, steering, breaking system and the wheel itself. In this research, the wheel locomotion module is a quarter of the 9-ton off-road vehicle. It is represented by a nonlinear mathematical model not taking into account the lateral behavior of the wheel. The wheel locomotion module's dynamics includes equations for representing the module's rotational, longitudinal and vertical dynamics. Modeling of dynamic processes in a heavy off-road vehicle has multiple peculiarities which are outlined below with corresponding references to the relevant literature for the interested reader.

In this study, an electric drive-line of the module is represented by the rotational dynamics model that includes rotational inertial of the e-motor and the rotational inertial of the wheel (see [22]), and can be represented by the following differential equations

$$\begin{cases} \dfrac{dI_m}{dt} = \dfrac{1}{L_a} \cdot \left(-R_a \cdot I_m - k_{emf} \cdot \omega_m + u_c \cdot \dfrac{k_{bat}}{u_{c_{max}}} \right), \\ \dfrac{d\omega_m}{dt} = \dfrac{k_r \cdot I_m - k_{rg} \cdot T_s}{J_m}, \\ \dfrac{dT_s}{dt} = k_s \cdot (k_{rg} \cdot \omega_m - \omega_w), \\ \dfrac{\omega_w}{dt} = \dfrac{1}{J_w}(T_s - T_{wload} - B_{fw} \cdot \omega_w) \end{cases} \tag{8}$$

where, I_m is the motor electric current, R_a is the armature resistance, L_a is the armature inductance, k_{emf} is the back EMF constant, k_{bat} is the maximum output of the PWM battery, u_c and $u_{c_{max}}$ are the control input and maximum control input respectively, ω_m is the angular velocity of the e-motor, J_m is the rotational inertia that includes the inertia of the motor and inertia of the gear reduction set that connects the motor with the wheel, k_t is the DC motor constant, k_{rg} is the speed ratio of the gear set, T_s is the elastic torque of the gear set, k_s is the rotational stiffness of the gear set, ω_w is the angular velocity of the wheel, J_w is the rotational inertia of the wheel, T_{wload} is the wheel load torque that comes from stochastic terrain, and B_{fw} is the viscous friction coefficient of the tire.

The 4th equation of the system (8) reflects the longitudinal dynamics of the wheel as follows. The wheel load torque, acting on the wheel module, is derived from the power balance in the contact patch between the tire and its surface of motion according to the following formula from [13]

$$T_{wload} = F_x r_w^0 = (R_x + F_{frame_x} + D_a + F_a) \cdot r_w^0 \tag{9}$$

where, F_x is the is the wheel circumferential force also known as wheel longitudinal force and r_w^0 is the tire rolling radius in the driven mode, i.e. when there is no driving torque applied to the wheel (more details on driven/driving/braking/free/neutral wheel power-loading modes can be found in [9, 23–25], R_x is the rolling resistance that is computed from the Bekker–Wong model [24, 26], F_{frame_x} is the vehicle frame force, D_a is the air drag force, $F_a = M \cdot \frac{dV_x}{dt}$ is the force needed to increase the kinetic energy of a wheel in translational motion, M is the module gross mass.

It should be noted that in addition to the dependence of wheel load torque or the wheel circumferential force, F_x, on longitudinal tire/terrain characteristic, such as rolling resistance, air drag, longitudinal inertia forces, etc.; the force F_x is dependent on the wheel normal dynamics described by the wheel normal reaction force, R_z, that can be defined from Eq. (10), showing a nonlinear dependence of the wheel circumferential force on the tire slippage (more details are presented in [13])

$$F_x = \mu_{px} R_z \left(1 - e^{-ks_\delta}\right) \tag{10}$$

where, μ_{px} is the peak friction coefficient also known as road grip, k is an empirical factor that depends on the properties of the tire and the surface, R_z is the tire normal reaction and s_δ is the tire slippage. Thus, Eqs. (9)-(10) represents the modeling of the longitudinal and normal dynamics of the wheel locomotion module.

In order to model the dynamics of wheel module motion, the stochastic modeling of road/terrain parameters (see [27]) for asphalt, meadow, soil, snow is performed using the data from [13]. Corresponding results are presented in Section 5.

4.2. Application of the off-road traction control

For the purposes of measuring the accuracy and performance of the traction control, the tire slippage can be computed during the wheel acceleration process with the aim of the formula that is reasonable for off-road applications

$$s_\delta = \frac{V_t - V_x}{V_t} = \frac{\omega_m r_w^0 - V_x}{\omega_m r_w^0} = 1 - \frac{r_w}{r_w^0} \tag{11}$$

where, V_t is the theoretical longitudinal velocity at zero slippage, V_x is the actual velocity of the wheel's center, ω_w is the angular velocity of the wheel and r_w is the tire rolling radius in the driving mode.

Apart from tire slippage analysis, a wheel mobility index is introduced in [28] for online computation of mobility performance of the wheel and is presented as follows

$$WMI = 1 - \frac{F_x}{R_z \mu_{px}} \tag{12}$$

here, WMI is the wheel mobility index.

These two variable are used for the verifying the performance of the controller designed in this article.

5. Results and discussion

Although the developed fuzzy polynomial controller is synthesized for linear model of the wheel locomotion module that includes only rotational dynamics (presented in Eq. (8)), further, the controller is applied to the nonlinear wheel locomotion model with rotational, longitudinal and normal dynamics modelled (see Eqs. (8), (10)). Then, the fuzzy corrector together with the fine-tuned fuzzy polynomial controller form the off-road traction control system of the wheel locomotion module.

In this section, the simulation results of the developed adaptive polynomial fuzzy controller are presented during each step of the control system development process. The logic flow of the presented simulation results in this section is constructed in the same order as controller synthesis is presented in Section 3.

Firstly, it should be noted that the four different road/terrain conditions are simulated using stochastic road/terrain parameters. Thus, each of the simulated road/terrain profile is character-ized by stochastic variation of road height, road grip coefficient (peak friction coefficient) and rolling resistance coefficient. The stochastic variation of these road/terrain parameters is performed basing on the road profile generation method from [27] and is presented in Figure 5. The Figures 5a–d show the road height variation for each of for road/terrain types which is presented as a function of distance in meters. It can be seen that the road height for asphalt road is rather small (up to 20 mm) and does not have steep fluctuations with respect to distance traveled, however, for soil terrain condition the fluctuations are much steeper and the road height can reach up to 100 mm. Figure 5e presents the peak friction coefficient (road grip) μ_{px} variations with respect to road distance. It is clear that the peak friction coefficient is quite different for each of road/terrain types that correlates with the data from [13, 26]. In addition, the fluctuations of peak friction coefficient are more obvious for road snow conditions (see curve 4 in Figure 5e). Figure 5f shows the rolling resistance coefficient f variation with respect to distance where the lowest resistance as expected can be observed for asphalt road. Hence, these profiles of road/terrain parameters are utilized for modeling of the wheel locomotion module run-ning on four road/terrain types: asphalt road, meadow, soil and snow road. The controller verification has to be conducted for each of the road/terrain types.

As it was mentioned above, for the linear model of the locomotion module (see Eq. (8)) the fuzzy polynomial controller was synthesized. It should be noted, that the designed fuzzy polynomial con-troller is dependent on the system states values that in this study are simulated by the ideal sensors, however, it can be extended for using the state observers that are being developed for the wheel locomotion module in [29].

The tuning of the fuzzy polynomial controller is performed by the reinforcement learning algorithm as described in Section 2.2. Thus, Figure 6 shows the transients of fine-tuned fuzzy polynomial con-troller for four different simulated wheel load torques, T_{wload}. Such arbitrary chosen wheel load torques are used while tuning the fuzzy polynomial controller to roughly represent different stochasticity of wheel load from road/terrain. As can be seen from the transients in Figure 6, the tuned fuzzy controller is capable to ensure the specified angular velocity at each of the wheel load torques simulated.

Then, the tuned by reinforcement learning algorithm fuzzy polynomial controller for linear model is applied to the nonlinear model of the wheel locomotion module which includes rotational, longitudinal and normal dynamics of the wheel (see Eqs. (8), (10)). This fuzzy polynomial controller was verified on four aforementioned road/terrain types, namely, asphalt road, meadow, soil and snow road. The transient processes of the obtained control system are presented in Figure 7. It was found out from these transients that, despite the output signal of the system is controlled at the specified value of input signal, the fast ramp input signal causes inappropriate high e-motor current values I_m (see Figure 7), and even if these currents could be provided the tire slippage and wheel mobility loss would occur during early phase of acceleration (can be seen further in Figures 10–11). One of the reasons of such behavior is that the inappropriate (raw) input was accepted by the controller. This problem may be solved by adding corresponding term to the reward function. However, this would make the learning process much longer and complicated. Therefore, in this paper a fuzzy corrector was applied to solve this problem. Typically, the reference wheel velocity signal, which is either when the acceleration pedal or upper level controller (in case of autonomous vehicle), increases following ramp trajectory and provides an input to velocity controller.

Therefore, a control input signal correction algorithm is applied based on the fuzzy logic corrector that is described in Section 3.3. As the developed fuzzy corrector is dependent on (i) character of the received input signal itself (ramp input) and (ii) type of the road/terrain condition, it produces the different control inputs for the fuzzy polynomial controller for different types of terrain (see Figure 8).

Finally, the traction controller is developed using integration of fuzzy polynomial controller tuned by reinforcement learning algorithm with the fuzzy corrector for controller input signal. The comparison of the simulation results for four road/terrain conditions with and without control input signal correction
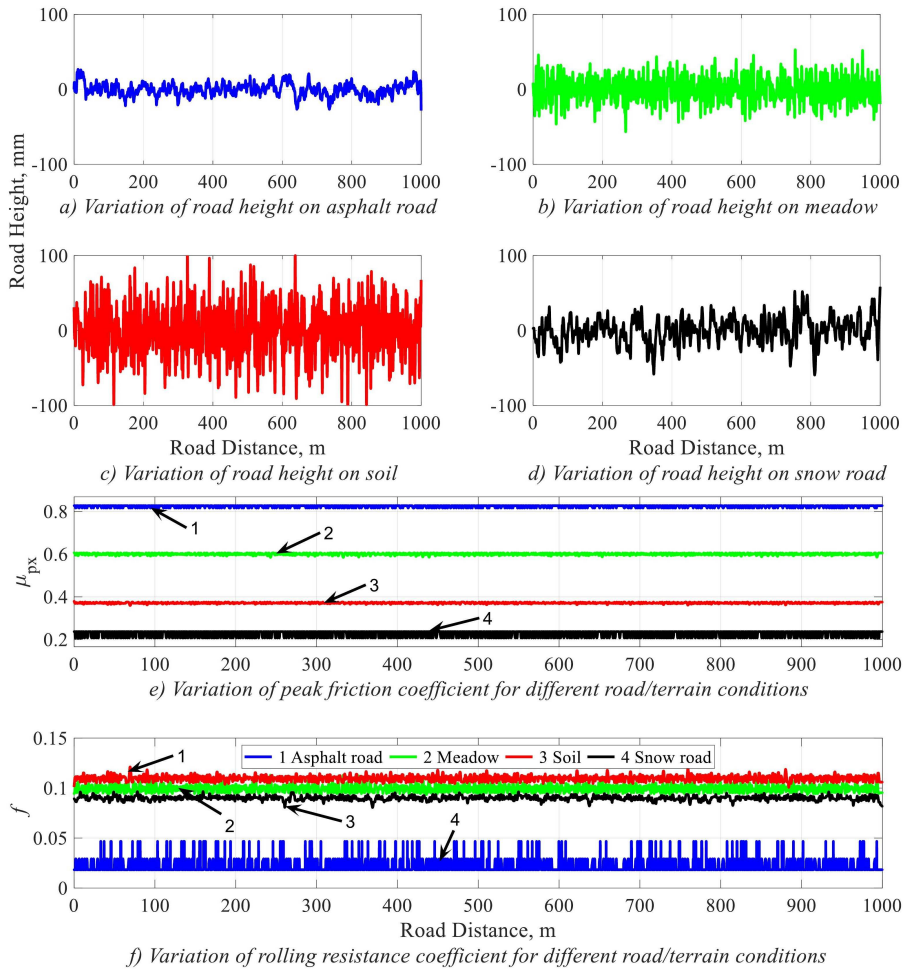
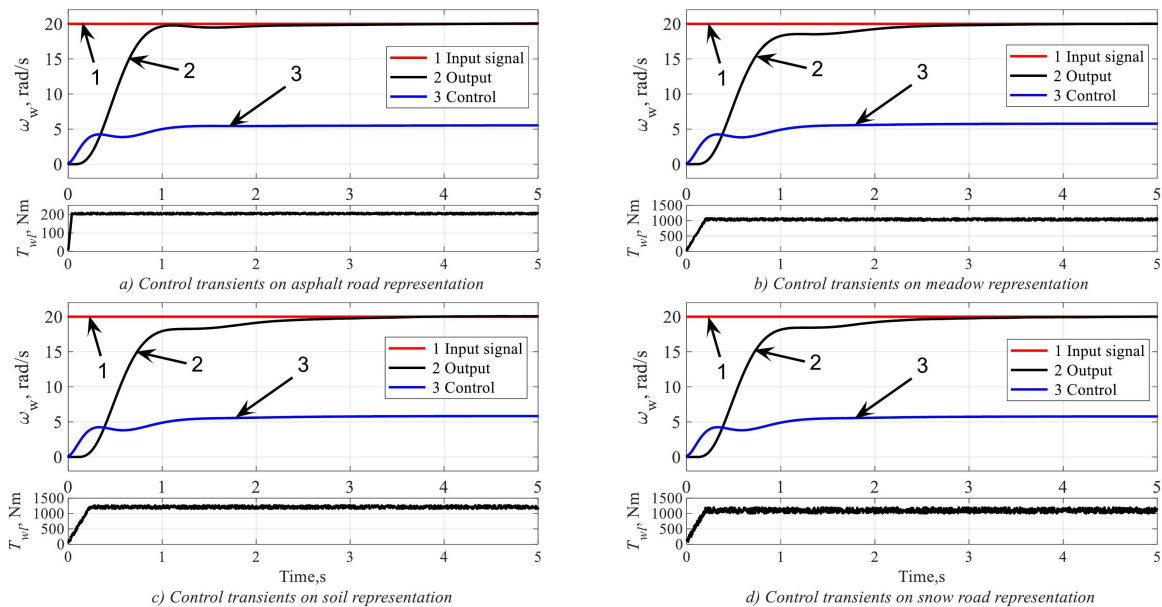**Fig. 5.** Stochastic variation of the road/terrain parameters.

Fig. 6. Transients of the linear rotational model of the wheel locomotion module with fuzzy polynomial controller tuned by reinforcement learning algorithm.
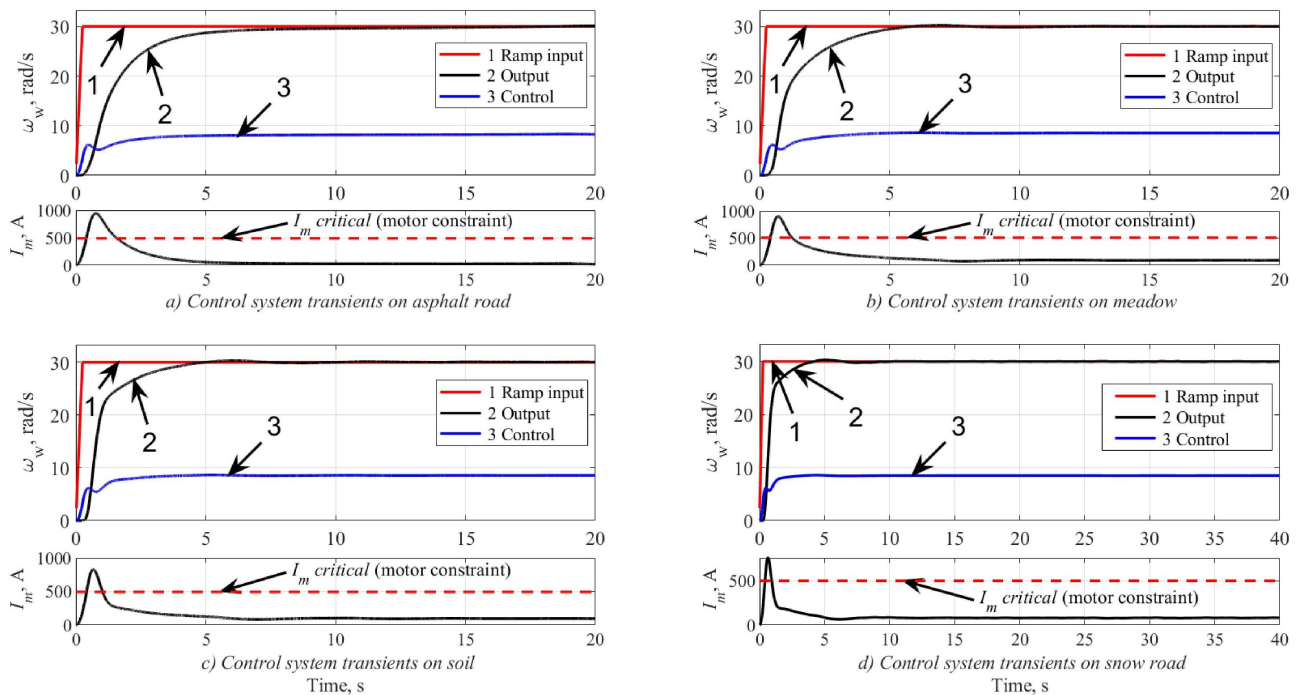
Fig. 7. Transients of the nonlinear model of the wheel locomotion module with fuzzy polynomial controller tuned by reinforcement learning algorithm.

is presented in Figure 9. From the simulation results it can be seen that the acceleration time on asphalt road is approximately equal to acceleration time using ramp input (see Figure 9a). On other road/terrain conditions, namely, meadow, soil and snow road, the acceleration takes more time when using a fuzzy corrector. However, it should be noted that in this case the e-motor current I_m is in safe ranges (see Figure 9), and, furthermore, the safety, energy efficiency and mobility is ensured during acceleration with utilizing fuzzy corrector, which is the primary goal of traction control system. The comparison of tire slippage and wheel mobility index during the whole acceleration process is presented in Figures 10–11. The values of tire slippage and wheel mobility index can be estimated in real-time using the Eqs. (11)–(12).

Wheel mobility index WMI is the value in percentage that was introduced in [28] to determine the margins of mobility performance of the wheel in different road/terrain conditions and at every particular moment of time. In general, 0% of wheel mobility index means the wheel have lost mobility (is spinning or skidding) and is out of control at the particular moment that could cause critical motion situations. In Figure 10, it can be seen that the wheel mobility index without utilizing the corrected input to fuzzy polynomial controller drops down to 0% for all types of the road/terrain conditions (Figure 10a–d causing the loss of mobility and unsafe motion conditions. However, while utilizing correction of the control input, the controller maintains the wheel mobility margin is away from 0%, thus, improve the safety and mobility performance of the traction control system.

Fig. 8. Results of fuzzy correction of the reference input signal to fuzzy polynomial controller.

Tire slippage values, s_δ, are also used for the developed controller verification. Tire slippage can vary from 0 to 1 for acceleration (1 means fully slipping and 0 means no slipping). Figure 11 shows that the tire slippage during acceleration with fuzzy correction of control input prevent the unsafe situation
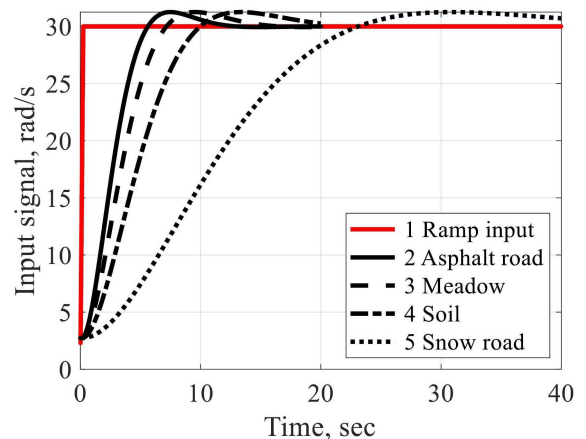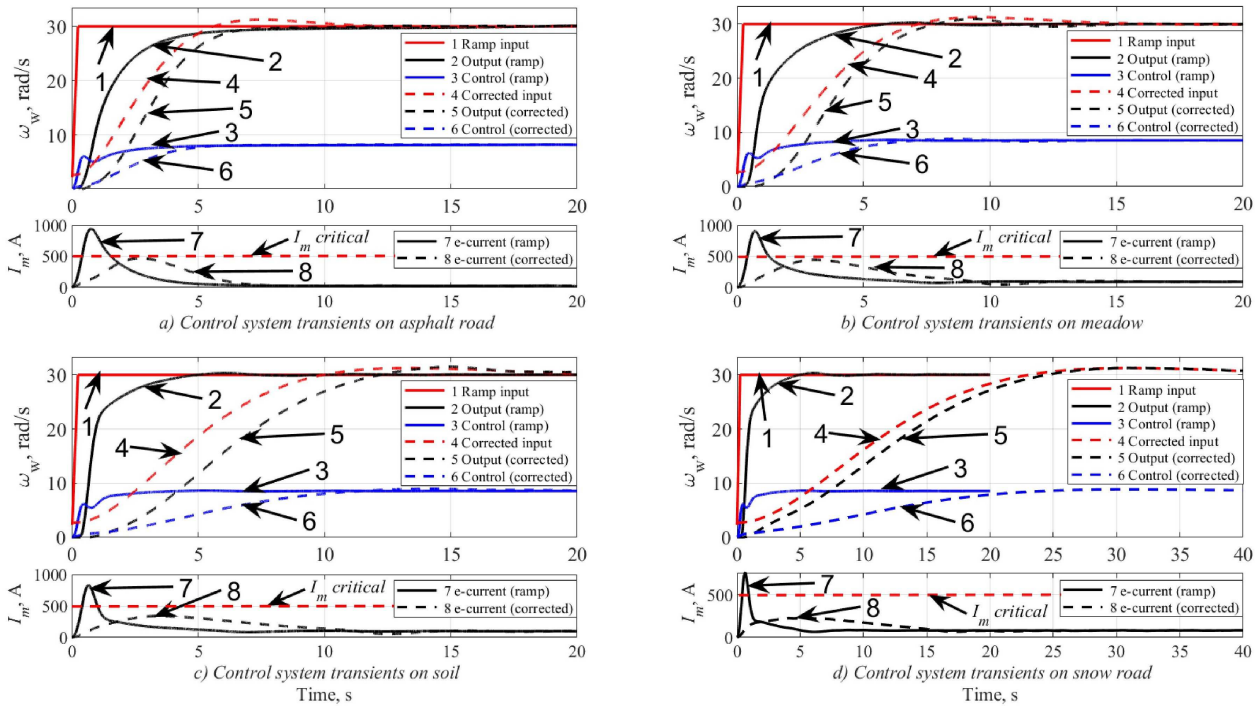
Fig. 9. Comparison of the control system transients with and without fuzzy correction of the reference input on: *a*) asphalt road; *b*) meadow; *c*) soil; *d*) snow road.
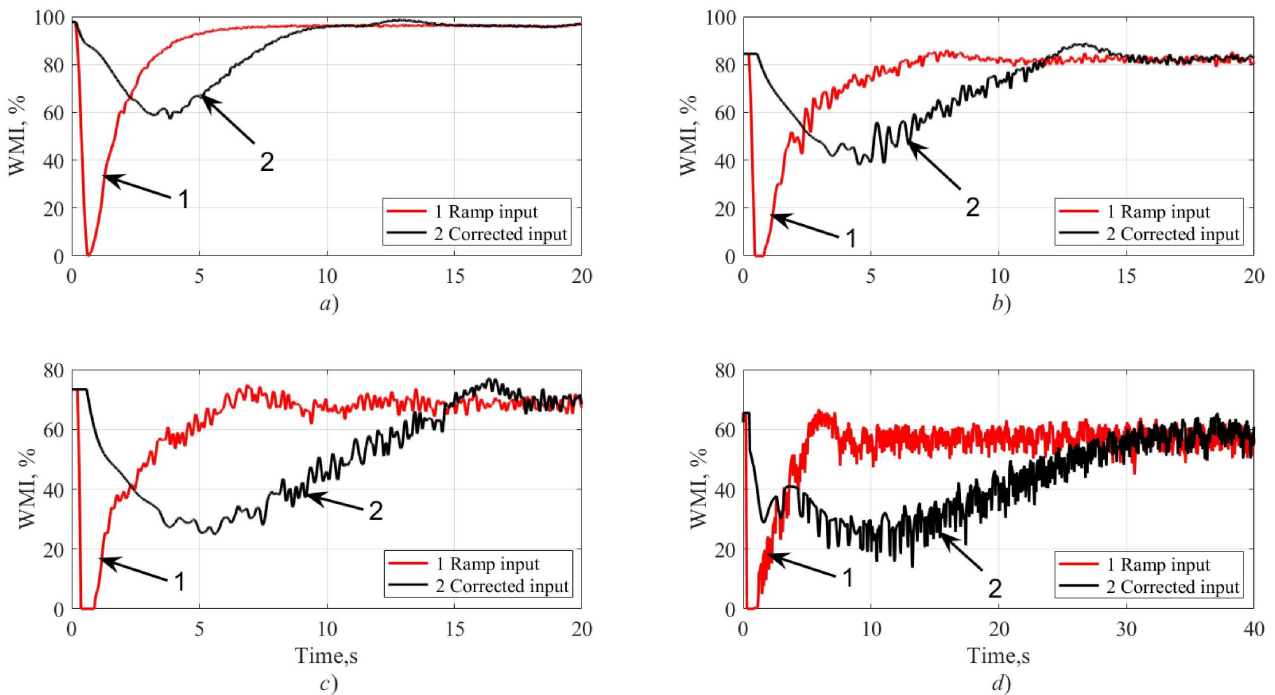

Fig. 10. Wheel mobility index estimations during acceleration on: *a*) asphalt road, *b*) meadow, *c*) soil and *d*) snow road.

of large slipping on all of the road/terrain types (see Figure 11*a–d*) while maintaining the tire slippage at safety level even for such severe road/terrain conditions as soil and snow road.

Therefore, the computational results show that the designed adaptive fuzzy polynomial controller tuned by reinforcement learning algorithm ensures lower values of tire slippage during acceleration while also preventing wheel mobility losses that is the ultimate goal of the off-road traction control system.
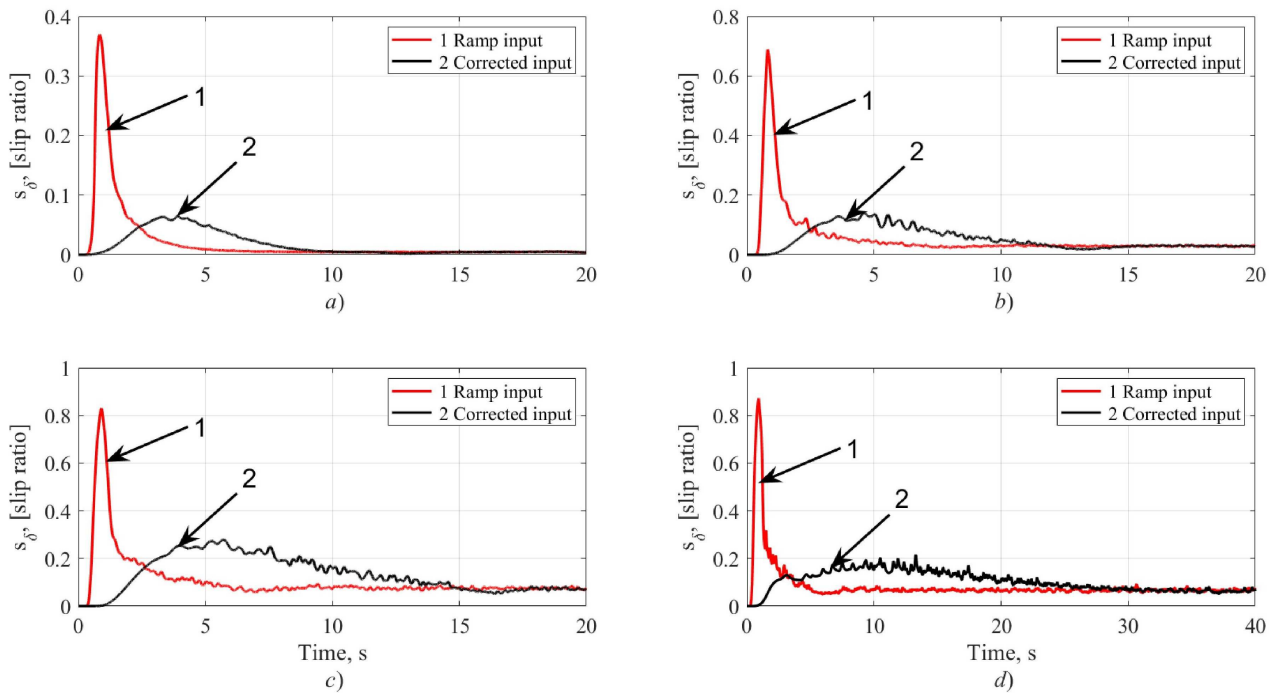
Fig. 11. Tire slippage estimation during acceleration on: *a*) asphalt road, *b*) meadow, *c*) soil and *d*) snow road.

6. Conclusions

In this study, the adaptive fuzzy polynomial controller tuned by reinforcement learning with the fuzzy input corrector is presented. This controller ensures the stable, optimal, and safe performance of the system and takes into account external disturbances. Stability of the controller is ensured by utilization of fuzzy logic controller that is designed to switch between the subsystems that ensure stable closed loop transients. Optimality, with respect to the reward function, is ensured by utilization of reinforcement learning. External disturbances influence the input signal of controller via fuzzy corrector ensuring smooth and, hence, energy-efficient transients.

The performance of the introduced controller is verified on the example of the wheel locomotion module. Four different types of stochastic road/terrain conditions are modeled for the purpose of controller verification in off-road conditions. The proposed controller does not influence acceleration time on firm terrains. However, the acceleration time is automatically adjusted by it to maintain safety and energy efficiency on loose soils and snow roads which coincides with the behavior of the experienced driver. The designed controller is capable to significantly reduce tire slippage and increase wheel mobility during the acceleration phase.

[1] Zhang Y., Li S., Liao L. Near-optimal control of nonlinear dynamical systems: A brief survey. Annual Reviews in Control. **47**, 71–80 (2019).

[2] Grune L., Pannek J. Nonlinear Model Predictive Control. Communications and Control Engineering. Springer, Cham (2017).

[3] Bououden S., Chadli M., Karimi H. R. A Robust Predictive Control Design for Nonlinear Active Suspension Systems. Asian Journal of Control. **18**, 122–132 (2016).

[4] Shin Y. C., Xu C. Intelligent systems: modeling, optimization, and control. CRC press (2009).

[5] Cervantes J, Yu W., Salazar S., Chairez I. Takagi–Sugeno Dynamic Neuro-Fuzzy Controller of Uncertain Nonlinear Systems. IEEE Transactions on Fuzzy Systems. **25** (6), 1601–1615 (2017).

[6] Vantsevich V., Lozynskyy A., Demkiv L., Klos S. A Foundation for Real-Time Tire Mobility Estimation and Control. Proc. 19th International and 14th European-African Regional Conference of the ISTVS, Budapest, Hungary (2017).

[7] Dawei M., Yu Z., Meilan Z., Risha N. Intelligent fuzzy energy management research for a uniaxial parallel hybrid electric vehicle. Computers and Electrical Engineering. **58**, 447–464 (2017).

[8] Arabi E., Gruenwald B. C., Yucelen T., Nguyen N. Intelligent fuzzy energy management research for a uniaxial parallel hybrid electric vehicle. International Journal of Control. **91** (5), 1195–1208 (2018).

[9] Dorf R. C., Bishop R. H. Modern control systems. Pearson (2016).

[10] Behrooz F., Mariun N., Marhaban M. H., Radzi M., Amran M., Ramli A. R. Review of control techniques for HVAC systems—nonlinearity approaches based on Fuzzy cognitive maps. Energies. **11** (3), 495 (2018).

[11] Lozynskyy A., Demkiv L. Application of dynamic systems family for synthesis of fuzzy control with account of non-linearities. Advances in electrical and electronic engineering. **14** (5), 543–550 (2016).

[12] Demkiv L. I. Research of dynamic system with unstable subsystem that has one root in the right half-plane. Mathematical modeling and computing. **1** (2), 156–162 (2014).

[13] Andreev A. F., Kabanau V., Vantsevich V. Driveline systems of ground vehicles: theory and design. CRC Press (2010).

[14] Lozynskyy A. O., Demkiv L. I., Vantsevich V. V., Borovets T. V., Gorsich D. J. An estimation accuracy of state observers under uncertain initial conditions. Mathematical modeling and computing. **6** (2), 320–332 (2019).

[15] Savitski D., Schleinin D., Ivanov V., Augsburg K., Jimenez E., He R., Barber P. Improvement of traction performance and off-road mobility for a vehicle with four individual electric motors: driving over icy road. Journal of Terramechanics. **69**, 33–43 (2017).

[16] Osinenko P. V., Geissler M., Herlitzius T. A method of optimal traction control for farm tractors with feedback of drive torque. Biosystems engineering. **129**, 20–33 (2015).

[17] Kim J., Lee J. Traction-energy balancing adaptive control with slip optimization for wheeled robots on rough terrain. Cognitive Systems Research. **49**, 142–156 (2018).

[18] Addison A., Vacca A. Real-Time Parameter Setpoint Optimization for Electro-Hydraulic Traction Control Systems. Proc. 15th Scandinavian International Conference on Fluid Power, Linköping, Sweden. **144**, 104–114 (2017).

[19] Sutton R. S., Barto A. G. Reinforcement learning: An introduction. MIT press (2018).

[20] Tay T. T., Mareels I., Moore J. B. High performance control. Springer Science and Business Media (2012).

[21] Lozynskyy A., Demkiv L. Synthesis of multicriteria controller by means of fuzzy logic approach. Advances in Fuzzy Systems. **2014**, Article ID 758207 (2014).

[22] Vantsevich V. V., Lozynskyy A., Demkiv L., Holovach I. Fuzzy logic control of agile dynamics of a wheel locomotion module. Dynamics of Vehicles on Roads and Tracks 1: Proc. 25th International Symposium on Dynamics of Vehicles on Roads and Tracks, Rockhampton, Queensland, Australia. CRC Press (2018).

[23] Chudakov E. A. Theory of Automobile. State Publishing House of Machine-Building Literature, Moscow, Russia (1950), (in Russian).

[24] Bekker M. G. Introduction to Terrain-Vehicle Systems. Michigan University Ann Arbor (1969).

[25] Kutzbach H. D., Bürger A., Bottinger S. Rolling radii and moment arm of the wheel load for pneumatic tyres. Journal of Terramechanics. **82**, 13–21 (2019).

[26] Wong J. Y. Terramechanics and off-road vehicles. Elsevier (1989).

[27] Gray J. P., Vantsevich V. V., Opeiko A. F., Hudas G. R. A Method for Unmanned Ground Wheeled Vehicle Mobility Estimation in Stochastic Terrain Conditions. Proc. 7th Americas Regional Conference of the ISTVS, Tampa, Florida, USA (2013).

[28] Gray J. P., Vantsevich V. V., Overholt J. L. Indices and Computational Strategy for Unmanned Ground Wheeled Vehicle Mobility Estimation and Enhancement. Proceedings of the ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 6A: 37th Mechanisms and Robotics Conference. Portland, Oregon, USA. August 4–7, 2013. ASME Paper No. DETC2013-12158 (2014).

[29] Vantsevich V., Gorsich D., Lozynskyy A., Demkiv L., Borovets T. State Observers for Terrain Mobility Controls: A Technical Analysis. Uhl T. (eds) Advances in Mechanism and Machine Science, IFToMM WC 2019, Mechanisms and Machine Science. **73**, Springer, Cham (2019).

Нечіткий регулятор, синтезований методом навчання з підкріпленням, для застосування у антибуксувальній системі автомобіля

Демків Л. І.[1], Лозинський А. О.[1], Ванцевич В. В.[2], Горсіч Д. Дж.[3], Литвин В. В.[1], Кльось С. Р.[1], Лезервуд М. Д.[4]

[1]*Національний університет "Львівська політехніка",*
вул. С. Бандери, 12, 79013, Львів, Україна
[2]*Університет Алабами у Бірмінгемі,*
м. Бірмінгем, штат Алабама, AL 35294, США.
[3]*Центр систем наземного транспорту армії США,*
Воррен, штат Мічиган, США.
[4]*Центр систем наземного транспорту, Наука і техніка Аліон,*
Воррен, штат Мічиган, США.

У цій статті представлено нечіткий регулятор, що налаштовується методом навчання з підкріпленням. Розроблений алгоритм використовує теорію нечіткої логіки та методи навчання з підкріпленням для підбору параметрів функцій належності нечіткого регулятора. Крім цього, імплементовано нечіткий задавач інтенсивності вхідного сигналу (сигналу завдання) нечіткого регулятора. Нечіткий задавач інтенсивності змінює вхідний сигнал регулятора враховуючи оригінальне значення вхідного сигналу та тип зовнішніх збурень у системі. Таким чином, розроблена система керування з нечітким регулятором налаштованим за допомогою методу навчання з підкріпленням забезпечує стабільну, оптимальну та безпечну роботу системи, та враховує зовнішні збурення в системі. Для перевірки роботи запропонованого методу керування, його було синтезовано до математичної моделі колісного модуля електроавтомобіля, щоб покращити антибуксувальну систему транспортного засобу. Ефективність розробленої системи керування на базі нечіткого регулятора підтверджено результатами імітаційного моделювання.

Ключові слова: *навчання з підкріпленням, нечіткий регулятор, антибуксувальна система, мобільність транспортного засобу.*