Algorithm	discrete_PPO	DDPG	DDPG	DDPG	DDPG
LR-Actor	3e-4	3e-4	1e-4	1e-3	1e-4
LR-Critic	1e-3	1e-3	1e-3	1e-3	1e-3
miniBatchSize	32	64	32	32	32
Exp. Buffer(DDPG		1e6	1e6	1e6	1e6
Discount Factor	0.96	0.99	0.99	0.99	0.99
Std. Deviation(DDPG)		0.2	0.05	0.1	0.05
Std. Dev. Decay(DDPG)		1e-6	1e-6	1e-6	1e-6
NN-layer width	100	100	100	100	100
Exp. Horizon(PPO)	(512)				
Epochs(PPO)	3				
EntropyLossWeight(PPO)	0.02				
Clip Factor(PPO)	0.2				
GAE(PPO)	0.95				
Episode Steps	512	512	512	512	512
Reward Weights:					
survival	0.1	0.1	0.1	0.1	0.1
s_x	1	1	1	1	1
v_x	1	1	1	1	1
v_y	-1	-1	-1	-1	-1
height_diff	-1	-1	-1	-1	-1
EnergyCons.	-0.1	-0.025	-0.025	-0.025	-0.025
Early Episode Termination:	-50	-50	-50	-50	0
Notes:	Energy penalty to high, agent does not move at all	Initially good learning, then flatlining	no converge toward high reward, just fluctuations		Here, episodes do not terminate prematurely, so no -50. Very slow training
Results:					
Episodes	1300	8500	9775	4357	180
Max. reward	< -50	<150	146	120	10

Algorithm	discrete_PPO	DDPG	DDPG	discrete_PPO	DDPG
LR-Actor	1e-4	3e-4	3e-4	1e-4	3e-4
LR-Critic	1e-3	1e-3	3e-4	1e-4	3e-4
miniBatchSize	32	16	16	128	16
Exp. Buffer(DDPG		1e6	1e6		1e6
Discount Factor	0.99	0.99	0.99	0.997	0.99
Std. Deviation(DDPG)		0.05	0.05		0.05
Std. Dev. Decay(DDPG)		1e-6	1e-6		1e-6
NN-layer width	100	100	100	200	100
Exp. Horizon(PPO)	(512)			(512)	
Epochs(PPO)	3			3	
EntropyLossWeight(PPO)	0.02			0.01	
Clip Factor(PPO)	0.2			0.02	
GAE(PPO)	0.95			0.95	
Episode Steps	512	512	512	512	512
Reward Weights:					
survival	0.1	0.1	0.1	0.1	0.1
s_x	1	1	1	1	1
v_x	1	1	1	1	1
v_y	-1	-1	-1	-1	-1
height_diff	-1	-1	-1	-1	-1
EnergyCons.	-0.025	-0.025	-0.025	-0.025	-0.025
Early Episode Termination:	-50	-50	0	0	0
Notes:	Episodes terminate again !		no -50, but episodes can terminate prematurely again	no -50, but termination	"
Results:					
Episodes	4981	19637	3361		11734
Max. reward	< -50	164	160		179

Algorithm	DDPG	DDPG	DDPG	DDPG	DDPG
LR-Actor	3e-4	3e-4	3e-4	3e-4	3e-4
LR-Critic	3e-4	3e-4	3e-4	3e-4	3e-4
miniBatchSize	16	128	16	16	316
Exp. Buffer(DDPG	1e6	1e6	1e6	1e6	1e6
Discount Factor	0.99	0.98	0.99	0.99	0.99
Std. Deviation(DDPG)	0.3	0.1	0.05	0.05	0.05
Std. Dev. Decay(DDPG)	1e-6	1e-6	1e-6	1e-6	1e-6
NN-layer width	100	200	100	100	100
Exp. Horizon(PPO)					
Epochs(PPO)					
EntropyLossWeight(PPO)					
Clip Factor(PPO)					
GAE(PPO)					
Episode Steps	512	512	512	512	256
Reward Weights:					
survival	0.1	0.1	0.1	0.1	0.1
s_x	1	1	1	1	1
v_x	1	1	5	50	50
v_y	-1	-1	-1	-1	-1
height_diff	-1	-1	-1	-1	-1
EnergyCons.	-0.025	-0.025	-0.025	-0.025	-0.025
Early Episode Termination:	0	0	0	0	0
Notes:	"	"	"	PC crashed overnight	Highest reward, but changed reward weights, still slightly better than all before
Results:					
Episodes	11430	8485	7707		> 16000
Max. reward	154	167	197		> 410

Algorithm	DDPG	DDPG
LR-Actor	3e-4	3e-4
LR-Critic	3e-4	1e-3
miniBatchSize	16	128
Exp. Buffer(DDPG	1e5	1e5
Discount Factor	0.995	0.99
Std. Deviation(DDPG)	0.05	0.05
Std. Dev. Decay(DDPG)	1e-6	1e-6
NN-layer width	100	128
Exp. Horizon(PPO)		
Epochs(PPO)		
EntropyLossWeight(PPO)		
Clip Factor(PPO)		
GAE(PPO)		
Episode Steps	256	256
Reward Weights:		
survival	0.1	0.1
s_x	1	1
v_x	50	50
v_y	-1	-1
height_diff	-1	-1
EnergyCons.	-0.025	-0.025
Early Episode Termination:	0	0
Notes:	"	"
Results:		
Episodes	14220	
Max. reward	365	