Code for ICRNN and proposed res\_RNN in Google Colab: <https://drive.google.com/file/d/1j2FxdwUojH6lWx8sfrNnzQgH6i6B4-4T/view?usp=sharing>

**L4DC: 1st December**

**ICML: around 9th - 26th January**

**COLT: around 10th February**

**UAI: around 17th February**

**NeurIPS: around 17th May**

**ICLR: around 28th September**

**AISTATS: around 13th October**

**ICRNN:**

* **Loss of representation power**
* **Vanishing gradient problem when the layer gets deeper**
* **Slow convergence speed**

**Setup:**

* Input\_baseline = [T, CA, Q, CA0]
* Input\_input\_convex = [T, CA, Q, CA0, -T, -CA, -Q, -CA0]
* Output = [T, CA]
* Need to update mean and std of input and output at MPC side
* 20 iterations for IPOPT to stabilize around the steady-state equilibrium
* Batch size 256, epochs 500, Adam optimizer, MSE loss

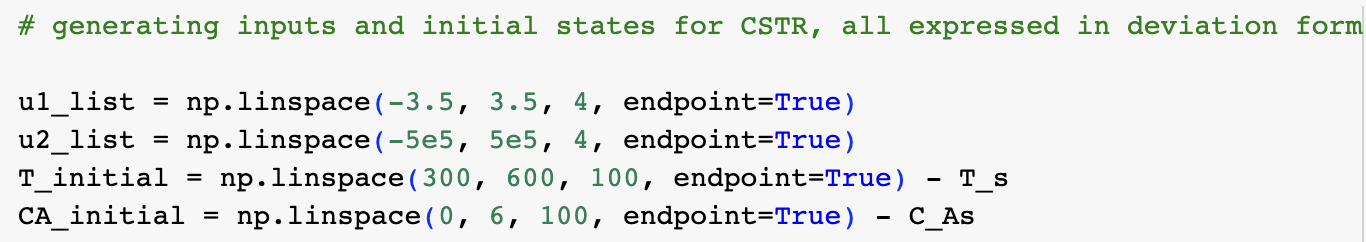


Fig. 1 Training data

**ICRNN architecture:**

1. RNN based
2. Input as [u.-u]
3. Non-negative weights
4. Parameterized residual connection
5. Convex and non-decreasing activation function such as ReLU, eLU
6. Convex loss function such as MSE

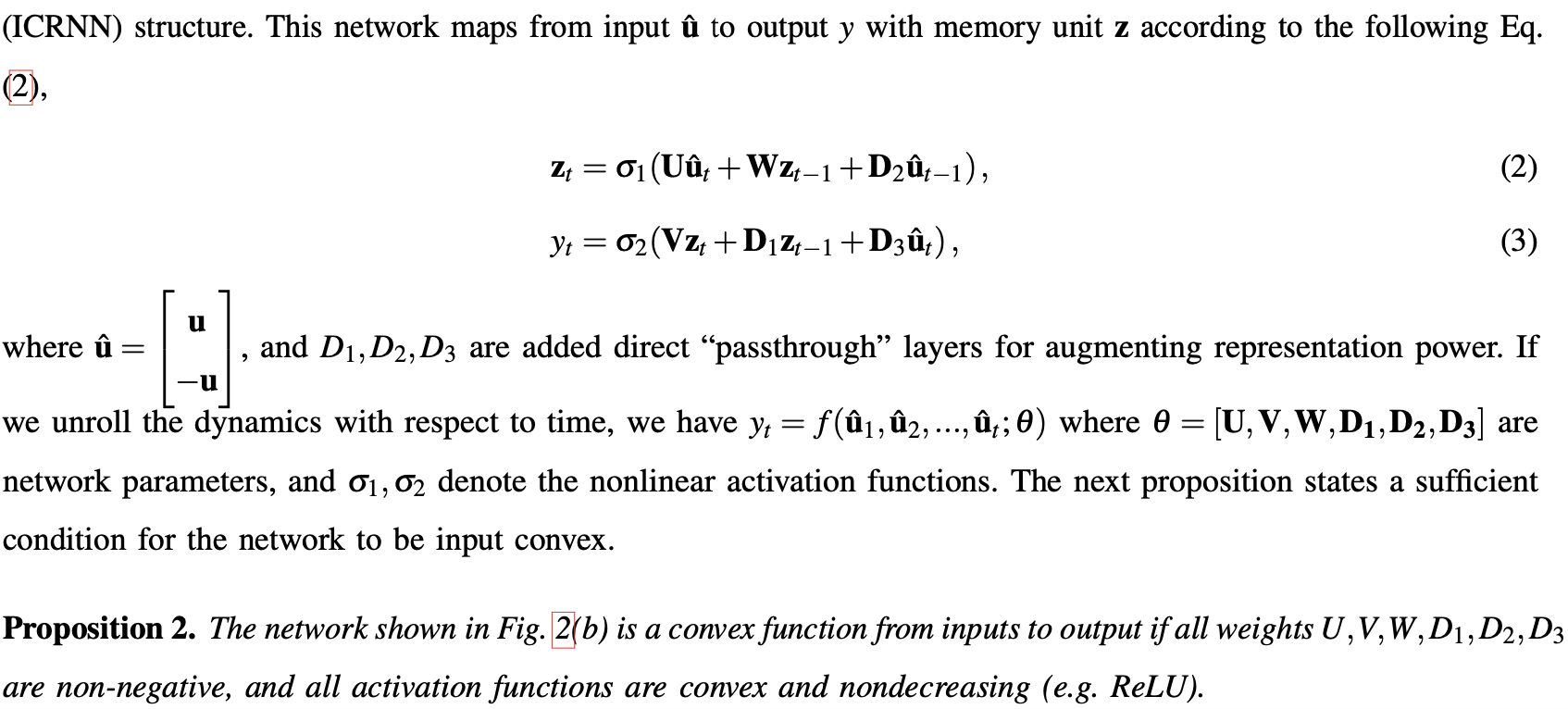


Fig. 2 ICRNN equations

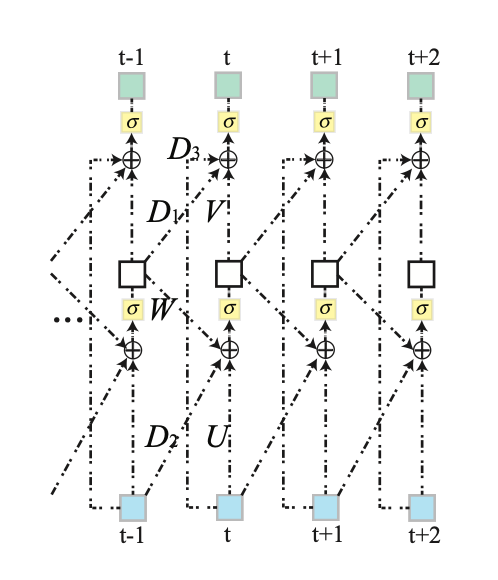


Fig. 3 ICRNN architecture

**Proposed residual\_LSTM architecture:**

1. Plain LSTM based, followed by dense layer for dimensionality reduction
2. Input as [u.-u]
3. Non-negative weights (i.e., [u,-u] as input makes sure non-negative weights will not restrict the hypothesis space)
4. Parameter-free skip connection from input to every deeper layer
5. Convex and non-decreasing activation function such as ReLU
6. Convex loss function such as MSE

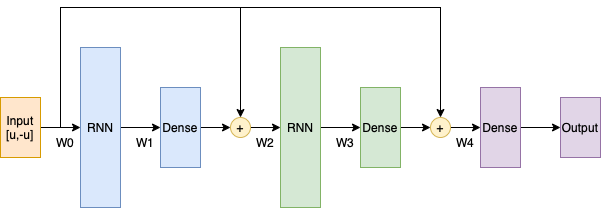


Fig. 4 residual\_LSTM

**Experiment A**

(T\_final = 0.005, T\_step = 0.0001, without Lyapunov-based constraint)

(Small region |CAi| < 0.1, |Ti| < 4)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 136.85  1642.20  (12 iterations) | 12,802 | 27,924 | [0.0140827,-1.223343] |
| 99.45  (20 iterations) | [0.0360961, -0.540501] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 82.37  988.44  (12 iterations) | 50,818 | 104,468 | [-0.041340,-1.034054] |
| 75.76  (20 iterations) | [0.0616584,-0.511369] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 98.92  989.20  (10 iterations) | 38,530 | 79,892 | [-0.067399,2.892405] |
| 120.66  (20 iterations) | [0.0839313,4.094705] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 31.77  349.47  (11 iterations) | 38,434 | 96,404 | [-0.065076,-0.534997] |
| 72.19  (20 iterations) | [-0.003731,-4.687944] |

Table I. Initial condition as [CAi, Ti] = [-1.5, 70] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 113.57  1589.98  (14 iterations) | 12,802 | 27,924 | [0.0528272,0.902580] |
| 93.12  (20 iterations) | [-0.055302,-2.334252] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 87.01  1131.14  (13 iterations) | 50,818 | 104,468 | [0.0637462,1.411909] |
| 63.32  (20 iterations) | [-0.023376,-3.574177] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 123.61  1483.32  (12 iterations) | 38,530 | 79,892 | [0.0648305,1.428434] |
| 87.09  (20 iterations) | [0.0279148,1.877667] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 80.25  1123.5  (14 iterations) | 38,434 | 96,404 | [0.0344987,-0.336749] |
| 75.87  (20 iterations) | [0.0045401,2.467877] |

Table II. Initial condition as [CAi, Ti] = [1.5, -70] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 119.33  1073.97  (9 iterations) | 12,802 | 27,924 | [-0.092670,1.411086] |
| 82.80  (20 iterations) | [0.0631080,-0.595821] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 139.85  1398.50  (10 iterations) | 50,818 | 104,468 | [-0.024903,-1.794260] |
| 85.51  (20 iterations) | [0.0179093,-1.295458] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 121.84  1340.24  (11 iterations) | 38,530 | 79,892 | [-0.083878,1.331406] |
| 120.21  (20 iterations) | [-0.032944,2.976239] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 47.54  570.48  (12 iterations) | 38,434 | 96,404 | [0.0052048,-3.126283] |
| 91.65  (20 iterations) | [-0.046544,-1.473470] |

Table III. Initial condition as [CAi, Ti] = [-1.25, 50] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 136.59  1639.04  (12 iterations) | 12,802 | 27,924 | [0.0480228,0.677339] |
| 117.54  (20 iterations) | [0.0700910,-4.349927] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 51.90  570.91  (11 iterations) | 50,818 | 104,468 | [0.0464287,-1.677691] |
| 47.35  (20 iterations) | [-0.028926,-0.008207] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 136.91  1369.11  (10 iterations) | 38,530 | 79,892 | [0.0858702,3.580197] |
| 120.58  (20 iterations) | [0.0970791,0.319159] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 67.59  743.52  (11 iterations) | 38,434 | 96,404 | [0.0649186,2.963727] |
| 87.93  (20 iterations) | [0.0006854,-0.279112] |

Table IV. Initial condition as [CAi, Ti] = [1.25, -50] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 75.01  600.06  (8 iterations) | 12,802 | 27,924 | [0.0110165,1.608246] |
| 64.41  (20 iterations) | [-0.064366,0.995813] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 83.68  753.12  (9 iterations) | 50,818 | 104,468 | [-0.021464,1.978281] |
| 60.58  (20 iterations) | [-0.036084,2.685399] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 39.47  394.71  (10 iterations) | 38,530 | 79,892 | [0.0037248,3.202584] |
| 71.18  (20 iterations) | [-0.034498,5.162849] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 30.82  246.57  (8 iterations) | 38,434 | 96,404 | [-0.047488,0.0514859] |
| 34.32  (20 iterations) | [0.0507985,2.332750] |

Table V. Initial condition as [CAi, Ti] = [-1, 60] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 100.22  1102.44  (11 iterations) | 12,802 | 27,924 | [-0.011292,0.811972] |
| 79.58  (20 iterations) | [-0.061833,-2.094542] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 56.17  673.99  (12 iterations) | 50,818 | 104,468 | [-0.001956,-0.313665] |
| 47.98  (20 iterations) | [0.0719977,-5.574035] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 77.42  774.20  (10 iterations) | 38,530 | 79,892 | [0.0771076,-0.448549] |
| 152.27  (20 iterations) | [0.0266666,-1.498481] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 54.70  492.30  (9 iterations) | 38,434 | 96,404 | [0.0496803,0.354156] |
| 57.12  (20 iterations) | [0.0685088,-3.435497] |

Table VI. Initial condition as [CAi, Ti] = [1, -60] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 51.50  412.01  (8 iterations) | 12,802 | 27,924 | [-0.044653,0.483079] |
| 36.04  (20 iterations) | [0.0490480,1.962824] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 25.31  202.45  (8 iterations) | 50,818 | 104,468 | [-0.099121,0.405274] |
| 40.34  (20 iterations) | [-0.049597,-2.436385] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 51.33  513.29  (10 iterations) | 38,530 | 79,892 | [0.0153827,-0.033929] |
| 106.75  (20 iterations) | [-0.001051,0.609182] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 19.21  153.70  (8 iterations) | 38,434 | 96,404 | [-0.093098,0.460468] |
| 25.16  (20 iterations) | [-0.077169,4.040569] |

Table VII. Initial condition as [CAi, Ti] = [-0.75, 40] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 99.89  599.36  (6 iterations) | 12,802 | 27,924 | [0.0670101,3.376633] |
| 75.90  (20 iterations) | [0.0437870,0.989792] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 69.62  487.34  (7 iterations) | 50,818 | 104,468 | [-0.014574,-1.431967] |
| 93.79  (20 iterations) | [-0.002609,-1.464011] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 98.93  593.56  (6 iterations) | 38,530 | 79,892 | [0.0098412,-2.301585] |
| 146.69  (20 iterations) | [0.0907552,-6.251444] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 63.99  447.92  (7 iterations) | 38,434 | 96,404 | [0.0704775,-2.196346] |
| 40.70  (20 iterations) | [-0.063286,-0.374912] |

Table VIII. Initial condition as [CAi, Ti] = [0.75, -40] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 92.17  1013.88  (11 iterations) | 12,802 | 27,924 | [-0.055117,2.0156574] |
| 74.10  (20 iterations) | [-0.003525,2.379394] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 128.41  1155.67  (9 iterations) | 50,818 | 104,468 | [-0.006396,0.496525] |
| 99.34  (20 iterations) | [0.0538478,3.933221] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 103.92  1247.03  (12 iterations) | 38,530 | 79,892 | [-0.023025,3.513124] |
| 125.97  (20 iterations) | [0.0853728,2.135955] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 37.26  409.82  (11 iterations) | 38,434 | 96,404 | [-0.077975,-1.639386] |
| 41.96  (20 iterations) | [-0.004831,-0.406343] |

Table IX. Initial condition as [CAi, Ti] = [-1.3, 60] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 135.88  1494.64  (11 iterations) | 12,802 | 27,924 | [0.0836928,-0.470164] |
| 118.19  (20 iterations) | [0.0618424,-1.979901] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 81.76  981.10  (12 iterations) | 50,818 | 104,468 | [0.0875188,1.469819] |
| 94.60  (20 iterations) | [-0.011410,0.986210] |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 119.56  1195.59  (10 iterations) | 38,530 | 79,892 | [0.0148144,1.089583] |
| 151.52  (20 iterations) | [0.0785579,0.203054] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 77.34  1005.42  (13 iterations) | 38,434 | 96,404 | [0.0458329,-0.265385] |
| 71.63  (20 iterations) | [-0.063441,2.330419] |

Table X. Initial condition as [CAi, Ti] = [1.3, -60] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| [CAi, Ti] | Distance to the small region | Plain RNN (model26) | Plain LSTM (model27) | ICRNN (model28) |
| --- | --- | --- | --- | --- |
| [-1.5, 70] | 66.015 | 4.699 | 2.828 | 2.831 |
| [1.5, -70] | 66.015 | 1.415 | 1.007 | 1.320 |
| [-1.3, 60] | 56.013 | 2.474 | 2.820 | 3.043 |
| [1.3, -60] | 56.013 | 1.487 | 0.976 | 1.189 |
| [-1, 60] | 56.007 | 2.434 | 3.054 | 1.601 |
| [1, -60] | 56.007 | 2.239 | 1.369 | 1.573 |
| [-1.25, 50] | 46.014 | 1.883 | 2.451 | 2.349 |
| [1.25, -50] | 46.014 | 2.204 | 0.768 | 1.841 |
| [-0.75, 40] | 36.006 | 2.681 | 1.317 | 3.340 |
| [0.75, -40] | 36.006 | 1.338 | 1.088 | 1.325 |
| average | NA | 2.29 | 1.71 | 2.04 |

Table XI. Speedup for ICLSTM

**Experiment B** (T\_final = 0.005, T\_step = 0.0001, with Lyapunov-based constraint) (Small region |CAi| < 0.1, |Ti| < 3)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 2201.86  (14 iterations) | 12,802 | 27,924 | [-0.027767,2.630091] |
| 2498.41  (20 iterations) | [0.0457713,2.547107] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1708.84  (14 iterations) | 50,818 | 104,468 | [-0.083279,-0.998804] |
| 2348.18  (20 iterations) | [-0.050779,0.809626] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model32) | relu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1204.34  (12 iterations) | 38,530 | 79,892 | [0.0724973,1.476156] |
| 2058.10  (20 iterations) | [-0.006946,1.781468] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 1018.79  (12 iterations) | 38,434 | 96,404 | [-0.099335,-1.071199] |
| 1686.06  (20 iterations) | [-0.065839,-0.760956] |

Table I. Initial condition as [CAi, Ti] = [-1.5, 70] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 2554.50  (19 iterations) | 12,802 | 27,924 | [0.0411258,-1.332709] |
| 2719.43  (20 iterations) | [0.0293038,-1.019970] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1938.92  (17 iterations) | 50,818 | 104,468 | [0.0027784,0.845425] |
| 2086.21  (20 iterations) | [-0.072278,3.837983] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1733.70  (16 iterations) | 38,530 | 79,892 | [0.0044364,2.191300] |
| 2069.91  (20 iterations) | [0.0184103,0.249431] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 1538.70  (14 iterations) | 38,434 | 96,404 | [0.0513659,2.483461] |
| 2066.54  (20 iterations) | [-0.024548,-1.373013] |

Table II. Initial condition as [CAi, Ti] = [1.5, -70] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 1412.78  (10 iterations) | 12,802 | 27,924 | [-0.079042,0.855165] |
| 2533.91  (20 iterations) | [0.0081195,3.992507] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1076.55  (9 iterations) | 50,818 | 104,468 | [-0.074693,-1.470040] |
| 1823.08  (20 iterations) | [0.0611639,-0.305645] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1114.71  (11 iterations) | 38,530 | 79,892 | [-0.052887,2.488515] |
| 2584.81  (20 iterations) | [0.116540, -2.528756] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 649.06  (11 iterations) | 38,434 | 96,404 | [0.0305285,-1.388787] |
| 1793.56  (20 iterations) | [0.0201740,-1.191542] |

Table III. Initial condition as [CAi, Ti] = [-1.3, 60] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 2194.86  (17 iterations) | 12,802 | 27,924 | [0.0097205,1.575460] |
| 2642.85  (20 iterations) | [-0.280913,13.975776] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1552.72  (15 iterations) | 50,818 | 104,468 | [-0.033197,-2.777269] |
| 2085.37  (20 iterations) | [0.0454866,-2.587032] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1403.33  (16 iterations) | 38,530 | 79,892 | [0.0237879,0.166506] |
| 1821.21  (20 iterations) | [0.171360, -4.579784] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 1512.86  (12 iterations) | 38,434 | 96,404 | [0.0957296,1.016236] |
| 2144.93  (20 iterations) | [0.0599186,2.669514] |

Table IV. Initial condition as [CAi, Ti] = [1.3, -60] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 1259.34  (11 iterations) | 12,802 | 27,924 | [-0.066267,1.635576] |
| 1882.70  (20 iterations) | [0.0434776,-3.278165] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1450.84  (9 iterations) | 50,818 | 104,468 | [-0.082766,0.105347] |
| 2243.49  (20 iterations) | [-0.016487,1.414223] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1065.07  (9 iterations) | 38,530 | 79,892 | [-0.050298,0.730584] |
| 2393.50  (20 iterations) | [0.0909495,-1.835805] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 937.44  (9 iterations) | 38,434 | 96,404 | [-0.002844,0.962998] |
| 1495.94  (20 iterations) | [-0.028350,-1.053358] |

Table V. Initial condition as [CAi, Ti] = [-1, 55] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 1767.98  (14 iterations) | 12,802 | 27,924 | [0.0829862,-1.243341] |
| 2314.09  (20 iterations) | [0.0545986,-0.270692] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1106.84  (11 iterations) | 50,818 | 104,468 | [0.0970259,-1.809445] |
| 1649.57  (20 iterations) | [-0.057738,1.194908] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1037.69  (10 iterations) | 38,530 | 79,892 | [0.0565395,-1.614210] |
| 1409.22  (20 iterations) | [0.0187195,1.236004] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 986.35  (10 iterations) | 38,434 | 96,404 | [0.0435459,1.050671] |
| 1295.00  (20 iterations) | [-0.032430,-0.756762] |

Table VI. Initial condition as [CAi, Ti] = [1, -55] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 1312.55  (12 iterations) | 12,802 | 27,924 | [-0.041036,1.918220] |
| 2496.72  (20 iterations) | [0.0582662,1.889237] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1327.58  (11 iterations) | 50,818 | 104,468 | [-0.017141,-1.632236] |
| 1715.98  (20 iterations) | [-0.035410,1.184625] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1246.42  (12 iterations) | 38,530 | 79,892 | [-0.091065,0.534862] |
| 2104.60  (20 iterations) | [0.0041309,2.996182] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 993.71  (12 iterations) | 38,434 | 96,404 | [-0.082386,0.543442] |
| 1784.25  (20 iterations) | [0.0059496,1.885579] |

Table VII. Initial condition as [CAi, Ti] = [-1.25, 50] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 1974.38  (16 iterations) | 12,802 | 27,924 | [-0.065011,1.396168] |
| 2446.16  (20 iterations) | [-0.010526,0.255689] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1854.80  (13 iterations) | 50,818 | 104,468 | [0.0278772,-2.425189] |
| 2328.72  (20 iterations) | [-0.056251,-1.843715] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1353.39  (15 iterations) | 38,530 | 79,892 | [0.0240004,-0.721269] |
| 1888.54  (20 iterations) | [0.100793, -1.095490] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 1199.94  (12 iterations) | 38,434 | 96,404 | [0.0761132,0.112490] |
| 1605.96  (20 iterations) | [-0.036167,2.956739] |

Table VIII. Initial condition as [CAi, Ti] = [1.25, -50] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 1787.51  (15 iterations) | 12,802 | 27,924 | [0.0639645,0.567098] |
| 1962.95  (20 iterations) | [0.0730626,-7.203713] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1174.78  (8 iterations) | 50,818 | 104,468 | [-0.019111,0.392983] |
| 1648.14  (20 iterations) | [-0.059004,1.842775] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1329.34  (8 iterations) | 38,530 | 79,892 | [0.0036530,2.742700] |
| 2123.51  (20 iterations) | [0.0389703,-1.000908] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 705.49  (9 iterations) | 38,434 | 96,404 | [-0.014181,1.170480] |
| 1800.22  (20 iterations) | [0.0213923,-0.995376] |

Table IX. Initial condition as [CAi, Ti] = [-0.75, 40] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neuron | input | MSE | Average runtime/  MPC iteration in second | Total parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model26) | tanh | No constraints | 2 | 64 | [u] | 2.7492e-05 | 1923.91  (12 iterations) | 12,802 | 27,924 | [0.0676054,-2.709203] |
| 2775.63  (20 iterations) | [0.0609468,-1.386020] |
| Plain LSTM (model27) | tanh | No constraints | 2 | 64 | [u] | 1.7523e-06 | 1584.45  (10 iterations) | 50,818 | 104,468 | [0.0286076,-0.877192] |
| 2233.71  (20 iterations) | [-0.018987,5.772907] |
| ICNN  (model33) | relu | Non-negative | 2 | 64 | [u,-u] | 8.1487e-05 | ( iterations) | 15,828 | 31,572 |  |
| (20 iterations) |  |
| ICRNN (model28) | elu | Non-negative | 2 | 64 | [u,-u] | 4.7460e-04 | 1090.99  (10 iterations) | 38,530 | 79,892 | [-0.031107,0.705749] |
| 1995.55  (20 iterations) | [-0.025111,0.136750] |
| ICLSTM (model29) | relu | Non-negative | 2 | 64 | [u,-u] | 2.4660e-06 | 586.78  (9 iterations) | 38,434 | 96,404 | [0.0550468,-0.413228] |
| 1588.16  (20 iterations) | [-0.029601,-3.012348] |

Table X. Initial condition as [CAi, Ti] = [0.75, -40] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| [CAi, Ti] | Distance to the small region | Plain RNN (model26) | Plain LSTM (model27) | ICNN (model33) | ICRNN (model28) |
| --- | --- | --- | --- | --- | --- |
| [-1.5, 70] | 67.015 | 2.161/53.730% | 1.677/40.381% |  | 1.182/15.407% |
| [1.5, -70] | 67.015 | 1.660/39.765% | 1.260/20.641% |  | 1.127/11.248% |
| [-1.3, 60] | 57.013 | 2.177/54.058% | 1.659/39.709% |  | 1.717/41.773% |
| [1.3, -60] | 57.013 | 1.451/31.073% | 1.026/2.567% |  | 0.928/-7.805% |
| [-1, 55] | 52.008 | 1.343/25.561% | 1.548/35.386% |  | 1.136/11.983% |
| [1, -55] | 52.008 | 1.792/44.210% | 1.122/10.886% |  | 1.052/4.948% |
| [-1.25, 50] | 47.014 | 1.321/24.292% | 1.336/25.149% |  | 1.254/20.275% |
| [1.25, -50] | 47.014 | 1.645/64.540% | 1.546/35.306% |  | 1.128/11.338% |
| [-0.75, 40] | 37.006 | 2.534/60.532% | 1.665/39.947% |  | 1.884/46.929% |
| [0.75, -40] | 37.006 | 3.279/69.501% | 2.700/62.966% |  | 1.859/46.216% |
| average | NA |  |  |  |  |

Table XI. Speedup/Percentage decrease for ICLSTM

**Appendix A**

**Miscellaneous results**

Experiment A

(T\_final = 0.01, T\_step = 0.0001)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 167.92 | 12,802 | 30,376 | [0.088298, -0.049764] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 136.72 | 50,818 | 106,408 | [-0.009543,-6.99692] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 135.44 | 38,530 | 82,427 | [0.11038, 5.60968] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 132.42 | 38,434 | 116,584 | [0.025038, -4.038777] |
| 1dCNN (model24) | relu | No constraints | 2 | 64 | [u,-u] | 6.0135e-05 |  | 73,640 | 349,544 |  |
| res\_1dCNN (model25) | relu | No constraints | 2 | 64 | [u,-u] | 8.5084e-05 |  | 106,984 | 519,016 |  |

Table I. Initial condition as [CAi, Ti] = [-1.5, 70] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 167.49 | 12,802 | 30,376 | [-0.034254,-4.13065] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 164.55 | 50,818 | 106,408 | [-0.002047,8.36754] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 166.04 | 38,530 | 82,427 | [0.032410,1.93434] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 143.80 | 38,434 | 116,584 | [0.091862,3.097277] |

Table II. Initial condition as [CAi, Ti] = [1.5, -70] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 192.43 | 12,802 | 30,376 | [-0.12820, 4.40521] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 154.84 | 50,818 | 106,408 | [-0.14999, 4.89174] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 195.75 | 38,530 | 82,427 | [0.046350, 2.55342] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 151.32 | 38,434 | 116,584 | [0.032933, -6.79117] |

Table III. Initial condition as [CAi, Ti] = [1, -25] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 179.72 | 12,802 | 30,376 | [0.033532, -7.88183] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 147.64 | 50,818 | 106,408 | [0.001176, -3.05411]] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 183.87 | 38,530 | 82,427 | [-0.007052,3.28557] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 139.28 | 38,434 | 116,584 | [0.043838,4.25872] |

Table IV. Initial condition as [CAi, Ti] = [-0.5, 30] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 189.26 | 12,802 | 30,376 | [-0.061026,-0.88658] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 124.38 | 50,818 | 106,408 | [0.074389, -1.76465] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 157.16 | 38,530 | 82,427 | [0.017750, -2.14142] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 156.58 | 38,434 | 116,584 | [0.081925, -1.95383] |

Table V. Initial condition as [CAi, Ti] = [-1, 50] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 211.28 | 12,802 | 30,376 | [-0.072199,-1.77597] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 103.68 | 50,818 | 106,408 | [-0.095680,0.50534] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 183.41 | 38,530 | 82,427 | [0.13476, 1.93122] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 143.66 | 38,434 | 116,584 | [-0.12689, -5.81186] |

Table VI. Initial condition as [CAi, Ti] = [0.75, -40] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 231.68 | 12,802 | 30,376 | [0.087484, -8.034599] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 131.54 | 50,818 | 106,408 | [0.035713,3.88300] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 169.49 | 38,530 | 82,427 | [-0.048183,1.31287] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 174.39 | 38,434 | 116,584 | [-0.053101,0.597122] |

Table VII. Initial condition as [CAi, Ti] = [-0.75, 15] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 123.51 | 12,802 | 30,376 | [-0.13078, 0.30905] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 160.32 | 50,818 | 106,408 | [0.10347, -6.27058] |
| ICRNN (model14) | elu | Non-negative | 2 | 64 | [u,-u] | 0.0020 | 140.66 | 38,530 | 82,427 | [0.085642,0.85767] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 119.62 | 38,434 | 116,584 | [-0.10643, -4.78938] |

Table VII. Initial condition as [CAi, Ti] = [0.5, -25] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (HORIZON=2)

Experiment B

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 144.36 | 12,802 | 30,376 | [-0.084230, -4.80835] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 134.19 | 50,818 | 106,408 | [-0.017932, -2.75213] |
| ICRNN (model13) | relu | Non-negative | 2 | 64 | [u,-u] | 0.0017 | 144.44 | 38,530 | 81,960 | [-0.030071,7.84450] |
| residual\_RNN (model3) | relu | Non-negative | 2 | 64 | [u,-u] | 3.6951e-05 | 150.77 | 10,402 | 60,520 | [-0.059110, 7.98947] |
| residual\_RNN (model19) | tanh | Non-negative | 2 | 64 | [u,-u] | 4.1668e-05 | 134.62 | 10,402 | 60,520 | [-0.20291, 8.30165] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 130.30 | 38,434 | 116,584 | [-0.14106, -5.058874] |
| residual\_LSTM (model21) | tanh | Non-negative | 2 | 64 | [u,-u] | 1.4391e-06 | 122.98 | 38,434 | 116,584 | [0.076810, -9.78527] |
| skip\_LSTM (model2) | tanh | No constraints | 2 | 64 | [u,-u] | 2.9629e-06 | 123.18 | 38,343 | 116,584 | [-0.17013, 3.10009] |

Table I. Initial condition as [CAi, Ti] = [1.5, -70] with MPC iteration of 200 and tolerance of 1e-5 for non-convex MPC

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in seconds | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 137.15 | 12,802 | 30,376 | [-0.080204,-2.53547] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 136.28 | 50,818 | 106,408 | [0.13559, -2.38774] |
| ICRNN (model13) | relu | Non-negative | 2 | 64 | [u,-u] | 0.0017 | 97.26 | 38,530 | 81,960 | [0.16120, 7.41330] |
| residual\_RNN (model3) | relu | Non-negative | 2 | 64 | [u,-u] | 3.6951e-05 | 152.42 | 10,402 | 60,520 | [-0.050908,-4.099978] |
| residual\_RNN (model19) | tanh | Non-negative | 2 | 64 | [u,-u] | 4.1668e-05 |  | 10,402 | 60,520 |  |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 |  | 38,434 | 116,584 |  |
| residual\_LSTM (model21) | tanh | Non-negative | 2 | 64 | [u,-u] | 1.4391e-06 | 148.39 | 38,434 | 116,584 | [-0.034999,-2.24819] |
| skip\_LSTM (model2) | tanh | No constraints | 2 | 64 | [u,-u] | 2.9629e-06 |  | 38,343 | 116,584 |  |

Table II. Initial condition as [CAi, Ti] = [-1.5, 70] with MPC iteration of 200 and tolerance of 1e-5 for non-convex MPC

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in seconds | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 172.07 | 12,802 | 30,376 | [-0.056542, 1.51577] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 147.95 | 50,818 | 106,408 | [0.044504,1.92701] |
| ICRNN (model13) | relu | Non-negative | 2 | 64 | [u,-u] | 0.0017 | 147.25 | 38,530 | 81,960 | [0.013725,1.62016] |
| residual\_RNN (model3) | relu | Non-negative | 2 | 64 | [u,-u] | 3.6951e-05 |  | 10,402 | 60,520 |  |
| residual\_RNN (model19) | tanh | Non-negative | 2 | 64 | [u,-u] | 4.1668e-05 |  | 10,402 | 60,520 |  |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 |  | 38,434 | 116,584 |  |
| residual\_LSTM (model21) | tanh | Non-negative | 2 | 64 | [u,-u] | 1.4391e-06 | 172.28 | 38,434 | 116,584 | [-0.016491,1.89112] |
| skip\_LSTM (model2) | tanh | No constraints | 2 | 64 | [u,-u] | 2.9629e-06 |  | 38,343 | 116,584 |  |

Table III. Initial condition as [CAi, Ti] = [1, -25] with MPC iteration of 200 and tolerance of 1e-5 for non-convex MPC

| model | activation | weights | No. of layer | No. of hidden neurons/layer | input | MSE | Average runtime/MPC iteration in second | Total trainable parameters | No. of FLOPs | Final MPC model output [CA,T] |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Plain RNN (model1) | tanh | No constraints | 2 | 64 | [u] | 3.7448e-05 | 41.21 | 12,802 | 30,376 | [-0.069601,-2.34537] |
| Plain LSTM (model6) | tanh | No constraints | 2 | 64 | [u] | 9.4855e-07 | 23.75 | 50,818 | 106,408 | [-0.035143,3.70806] |
| ICRNN (model13) | relu | Non-negative | 2 | 64 | [u,-u] | 0.0017 | 61.09 | 38,530 | 81,960 | [-0.073363,6.95117] |
| residual\_RNN (model3) | relu | Non-negative | 2 | 64 | [u,-u] | 3.6951e-05 | 80.14 | 10,402 | 60,520 | [-0.002560,5.71404] |
| residual\_RNN (model19) | tanh | Non-negative | 2 | 64 | [u,-u] | 4.1668e-05 | 24.28 | 10,402 | 60,520 | [0.18081, -7.16307] |
| residual\_LSTM (model20) | relu | Non-negative | 2 | 64 | [u,-u] | 1.0835e-06 | 28.95 | 38,434 | 116,584 | [-0.10684,-5.86027] |
| residual\_LSTM (model21) | tanh | Non-negative | 2 | 64 | [u,-u] | 1.4391e-06 | 45.32 | 38,434 | 116,584 | [-0.092220,-5.016106] |
| skip\_LSTM (model2) | tanh | No constraints | 2 | 64 | [u,-u] | 2.9629e-06 | 40.27 | 38,343 | 116,584 | [-0.13311, 2.08048] |

Table V. Initial condition as [CAi, Ti] = [1.5, -70] with MPC iteration of 200 and tolerance of 1e-5 for convex MPC (num\_horizon = 1)