ansi_regression

March 30, 2018

1 ANSI Application analysis

#view_boxplot(rxnet_df)

```
In [1]: import numpy
        import pandas
        import matplotlib.pyplot as plotter
        from scipy.stats import pearsonr
        from sklearn.metrics import mean_squared_error, mean_absolute_error
In [2]: def view_boxplot(df):
            %matplotlib
            df.boxplot()
           plotter.show()
1.1 CPU data
In [3]: cpu_df = pandas.read_csv('data/ansi_fake_data/ansi_fake_data_cpu.csv', index_col='Time
In [4]: #cpu_df.columns
In [5]: #view_boxplot(cpu_df)
1.2 Network TX
In [6]: txnet_df = pandas.read_csv('data/ansi_fake_data/ansi_fake_data_network_tx.csv', index_
In [7]: #txnet_df.columns
In [8]: #view_boxplot(txnet_df)
1.3 Network RX
In [9]: rxnet_df = pandas.read_csv('data/ansi_fake_data/ansi_fake_data_network_rx.csv', index_
In [10]: #rxnet_df.columns
In [11]: rxnet_df = rxnet_df.clip(lower=0, upper=15000)
```

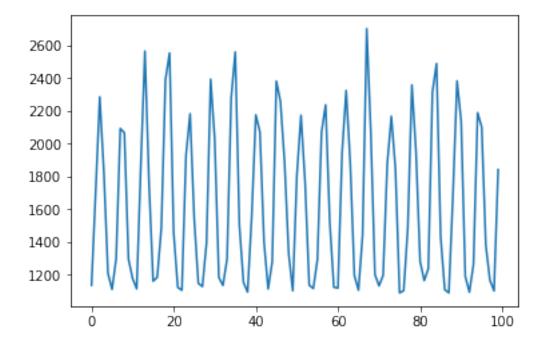
1.4 Disk IO data

```
In [12]: disk_df = pandas.read_csv('data/ansi_fake_data/ansi_fake_data_disk_io.csv', index_cole
In [13]: #disk_df.columns
In [14]: disk_df = disk_df.clip(lower=0, upper=4000)
         #view_boxplot(disk_df)
1.5 Context switching
In [15]: context_df = pandas.read_csv('data/ansi_fake_data/ansi_fake_data_context.csv', index_
In [16]: #context_df.columns
In [17]: context_df = context_df.clip(lower=0, upper=5000)
         #view boxplot(context df)
1.6 Seperate into proper dataframes for each node
In [18]: dframes = [cpu_df, txnet_df, rxnet_df, context_df, disk_df]
         node = \{\}
         for i in range(1,5):
             frames = []
             for dframe in dframes:
                 columns = list(filter(lambda x: f'bb{i}l' in x, dframe.columns))
                 frames.append(dframe[columns])
             node[i] = pandas.concat(frames, join='inner', axis=1).fillna(0)[:38200]
In [19]: for i in range(1,5):
             print(node[i].shape)
         print(node[1].columns)
(38200, 29)
(38200, 29)
(38200, 29)
(38200, 29)
Index(['cpu_value host bb1localdomain type_instance idle',
       'cpu_value host bb1localdomain type_instance interrupt',
       'cpu_value host bb1localdomain type_instance nice',
       'cpu_value host bb1localdomain type_instance softirq',
       'cpu_value host bb1localdomain type_instance steal',
       'cpu_value host bb1localdomain type_instance system',
       'cpu_value host bb1localdomain type_instance user',
```

'cpu_value host bb1localdomain type_instance wait',

```
'interface_tx host bb1localdomain instance lo type if_dropped',
 'interface_tx host bbllocaldomain instance lo type if_errors',
 'interface_tx host bb1localdomain instance lo type if_octets',
 'interface_tx host bb1localdomain instance lo type if_packets',
 'interface tx host bb1localdomain instance wlan0 type if dropped',
 'interface_tx host bb1localdomain instance wlan0 type if_errors',
 'interface tx host bb1localdomain instance wlan0 type if octets',
 'interface_tx host bb1localdomain instance wlan0 type if_packets',
 'interface_rx host bbllocaldomain instance lo type if_dropped',
 'interface_rx host bb1localdomain instance lo type if_errors',
 'interface_rx host bb1localdomain instance lo type if_octets',
 'interface_rx host bb1localdomain instance lo type if_packets',
 'interface_rx host bb1localdomain instance wlan0 type if_dropped',
 'interface_rx host bb1localdomain instance wlan0 type if_errors',
 'interface_rx host bb1localdomain instance wlan0 type if_octets',
 'interface rx host bb1localdomain instance wlan0 type if packets',
 'contextswitch_value host bb1localdomain type contextswitch',
 'disk_io_time host bb1localdomain instance mmcblk1 type disk_io_time',
 'disk_io_time host bb1localdomain instance mmcblk1boot0 type disk_io_time',
 'disk io time host bb1localdomain instance mmcblk1boot1 type disk io time',
 'disk_io_time host bb1localdomain instance mmcblk1p1 type disk_io_time'],
dtype='object')
```

1.7 Get data



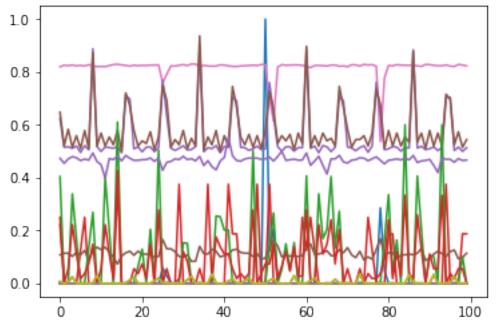
(152800, 29)

In [23]: #data = data[:,24]

1.8 Prepare scaler

1.9 Correrlation measurement

2 Prediction



```
In [28]: print(X.shape)
        LEN = X.shape[0]
        SPLIT = int(0.9*LEN)

        train_X = X[:SPLIT,:,:]
        val_X = X[SPLIT:SPLIT+1000,:,:]
        test_X = X[SPLIT+1000:,:,:]
(38200, 3, 29)
```

```
In [29]: X = train_X
         X = numpy.transpose(X, (1, 0, 2))
         \#X = X.reshape((-1,382,29))
         val_X = numpy.transpose(val_X, (1, 0, 2))
         test_X = numpy.transpose(test_X, (1, 0, 2))
         #val_X = val_X.reshape((-1,382,29))
         print(X.shape)
         print(val_X.shape)
(3, 34380, 29)
(3, 1000, 29)
In [30]: plotter.plot(test_X[0][:,25])
         plotter.show()
         \#test_X[0][:,25] = 0.0
         \#test_X[0][:,28] = 0.0
         plotter.plot(test_X[0][:,25])
         plotter.show()
         1.0
         0.8
         0.6
```

500

1000

0.4

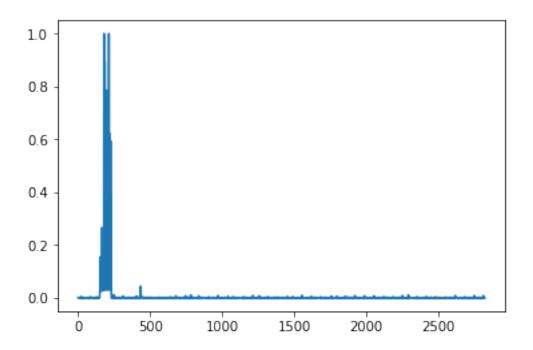
0.2

0.0

1500

2000

2500



```
In [31]: def flat_generator(X, tsteps = 5, ravel=1):
    i = 0

while True:
    batch_X = X[:,i:i+tsteps,:]
    batch_y = X[:,i+tsteps,:]

if ravel:
    batch_X = batch_X.reshape((batch_X.shape[0], -1))
    #print(batch_X.shape)
    #print(batch_y.shape)

yield batch_X, batch_y

i += 1
    if i > (X.shape[1] - tsteps - 1):
        i = 0
        continue
```

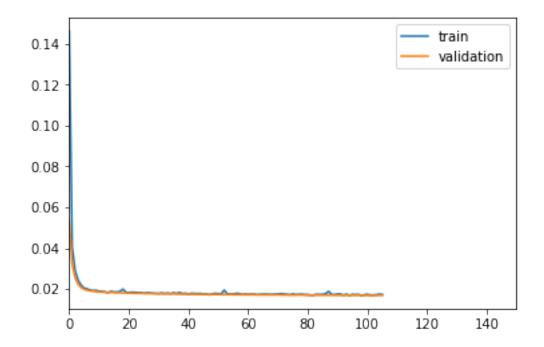
2.1 Flat models

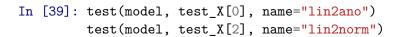
Using TensorFlow backend.

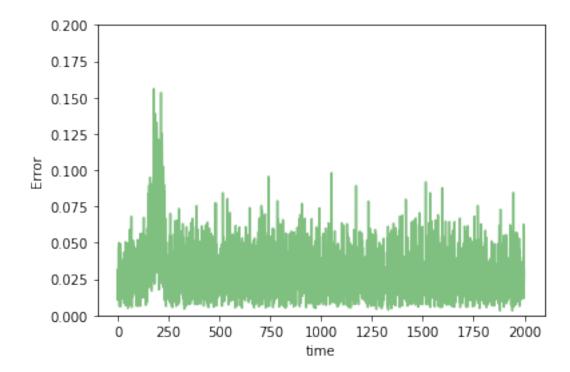
```
In [33]: def train(model, tgen, vgen, name="none"):
             estopper = EarlyStopping(patience=15, min_delta=0.0001)
             history = model.fit_generator(tgen, steps_per_epoch=1000, epochs=10000, callbacks
             plotter.plot(history.history['loss'],label='train')
             plotter.plot(history.history['val_loss'],label='validation')
             plotter.legend()
             plotter.xlim(0,150)
             plotter.show()
             plotter.savefig(f"{name}_train.png")
             print(history.history['loss'][-1])
In [34]: def test(model, dataset=test_X[0], ravel=1, write=0, name="none"):
             test_gen = flat_generator(numpy.array([dataset]), TIMESTEPS,0)
             error = []
             targets = []
             preds = []
             for i in range(2000):
                 _input,target = next(test_gen)
                 if i != 0:
                     #print(_input.shape)
                     _input = _input.squeeze()[1:,:]
                     #print(_input.shape)
                     _input = numpy.append(pred,_input, axis=0)[numpy.newaxis,:,:]
                     #print( input.shape)
                 targets.append(target.squeeze())
                 if ravel:
                     _input = _input.ravel()[:,numpy.newaxis].T
                 pred = model.predict(_input)
                 #print(target.shape)
                 #print(pred.shape)
                 preds.append(pred.squeeze())
                 error.append(mean_absolute_error(y_pred=pred, y_true=target))
             targets = numpy.vstack(targets)
             preds = numpy.vstack(preds)
             plotter.plot(error, 'g-', alpha=0.5)
             plotter.ylim(0,0.2)
             plotter.xlabel("time")
             plotter.ylabel("Error")
             plotter.show()
             plotter.savefig(f"{name}_testloss.png")
             error = numpy.array(error)
             print(numpy.mean(error))
             plotter.boxplot(error)
```

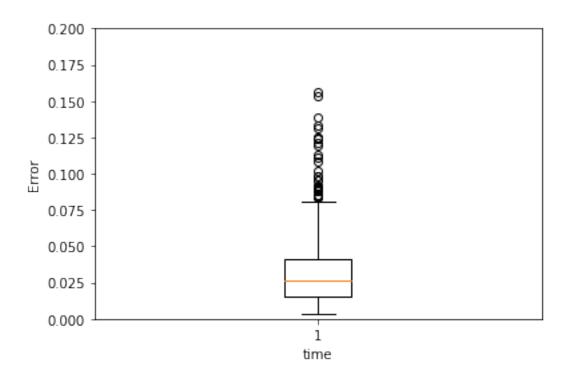
```
plotter.ylim(0,0.2)
plotter.xlabel("time")
plotter.ylabel("Error")
plotter.show()
plotter.savefig(f"{name}_boxplot.png")
if write:
    numpy.savetxt('loss.txt', numpy.array(error))
#print(error)
```

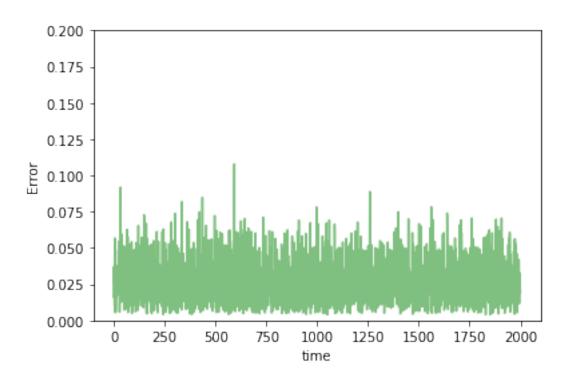
2.1.1 Linear Regression

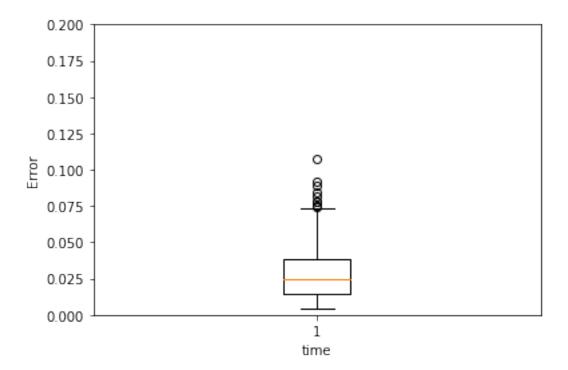










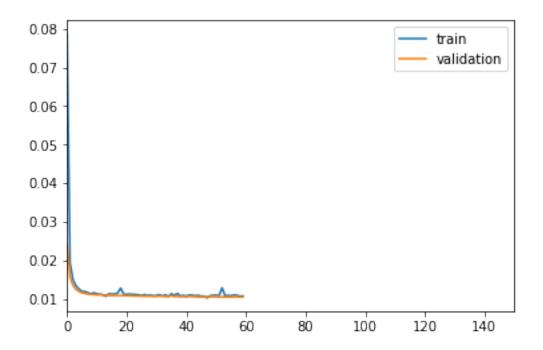


```
In [40]: TIMESTEPS = 5
        DIM = 29
        tgen = flat_generator(X, TIMESTEPS)
        vgen = flat_generator(val_X, TIMESTEPS)

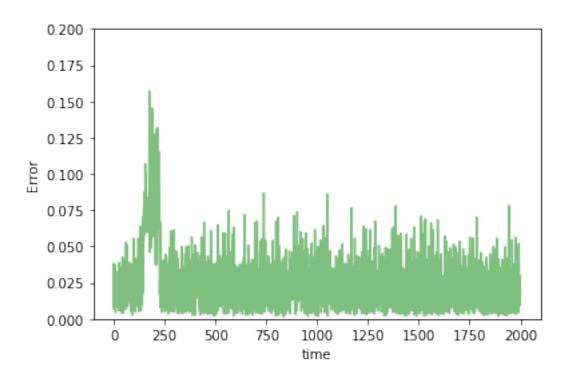
In [41]: input_layer = Input(shape=(TIMESTEPS*DIM,))
        output = Dense(DIM, activation='sigmoid')(input_layer)

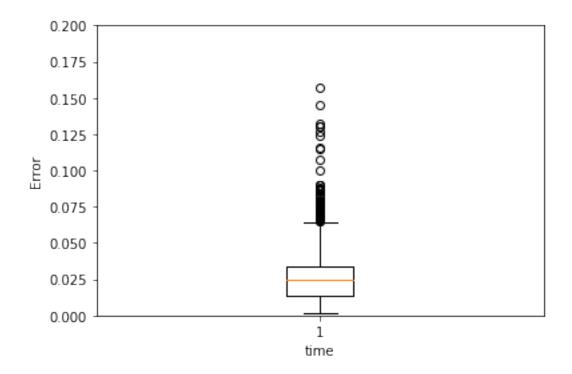
In [42]: model = Model(input_layer, output)
        model.compile(loss='mean_absolute_error', optimizer='adam', metrics=['mae'])

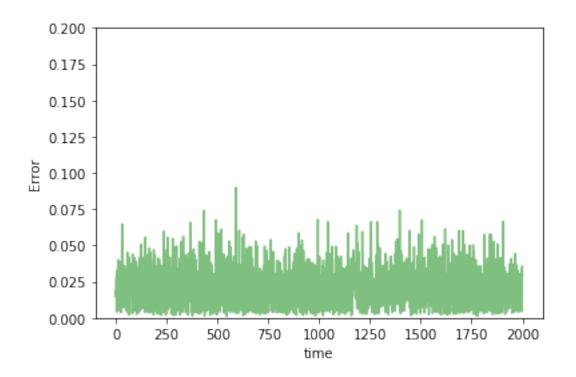
In [43]: train(model, tgen, vgen, name="lin5")
```

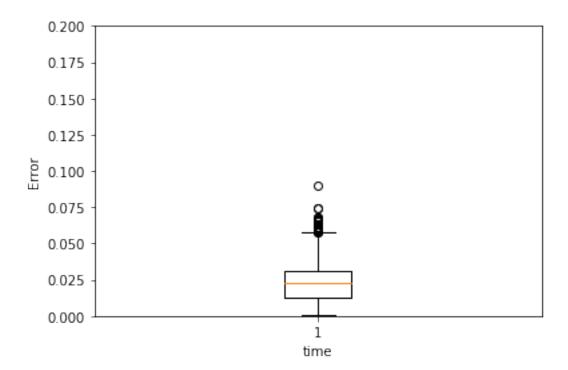


In [44]: test(model, test_X[0], name="lin5ano")
 test(model, test_X[2], name="lin5norm")







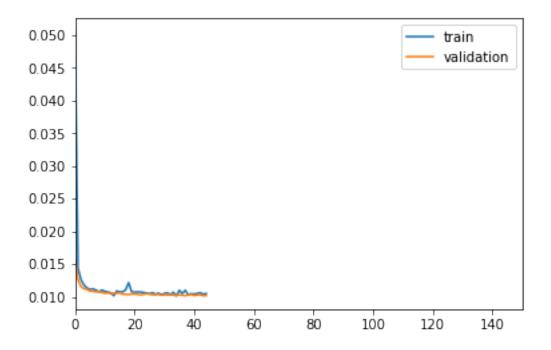


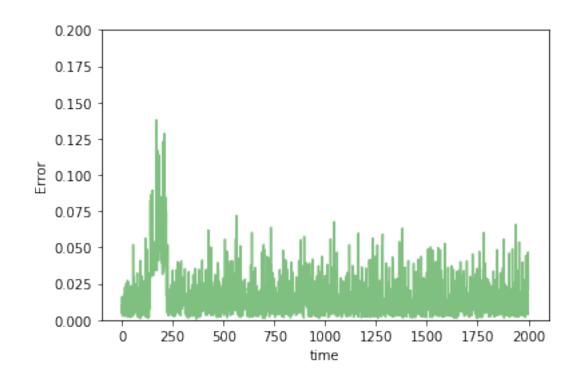
```
In [45]: TIMESTEPS = 10
        DIM = 29
        tgen = flat_generator(X, TIMESTEPS)
        vgen = flat_generator(val_X, TIMESTEPS)

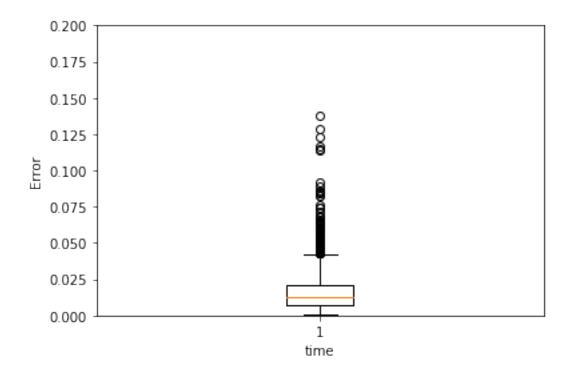
In [46]: input_layer = Input(shape=(TIMESTEPS*DIM,))
        output = Dense(DIM, activation='sigmoid')(input_layer)

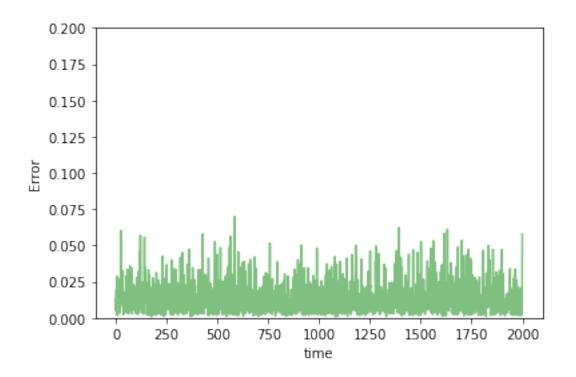
In [47]: model = Model(input_layer, output)
        model.compile(loss='mean_absolute_error', optimizer='adam', metrics=['mae'])

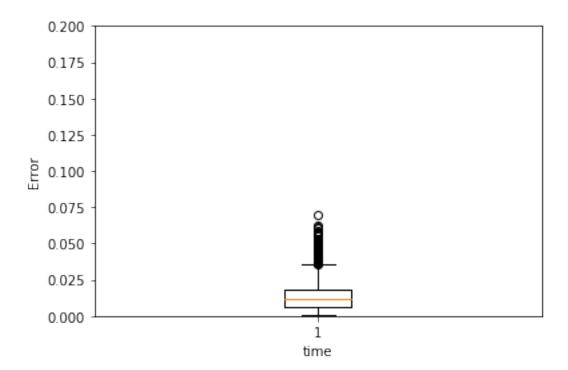
In [48]: train(model, tgen, vgen, name="lin10")
```









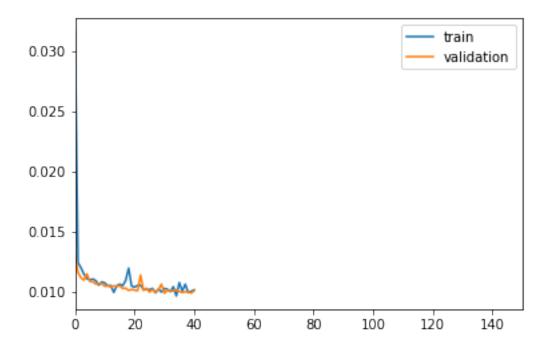


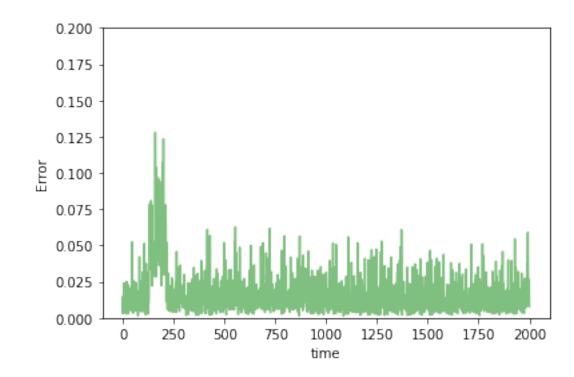
```
In [50]: TIMESTEPS = 20
        DIM = 29
        tgen = flat_generator(X, TIMESTEPS)
        vgen = flat_generator(val_X, TIMESTEPS)

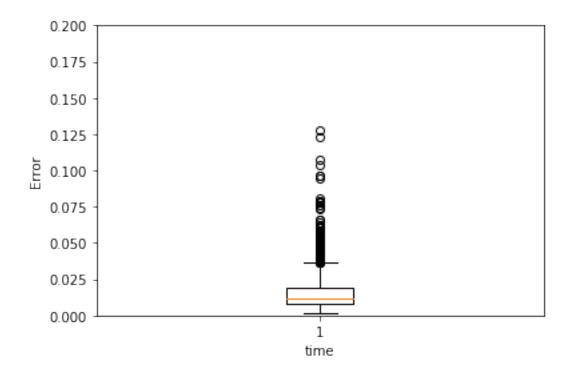
In [51]: input_layer = Input(shape=(TIMESTEPS*DIM,))
        output = Dense(DIM, activation='sigmoid')(input_layer)

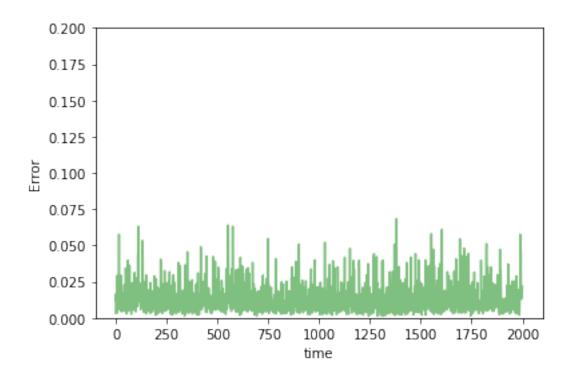
In [52]: model = Model(input_layer, output)
        model.compile(loss='mean_absolute_error', optimizer='adam', metrics=['mae'])

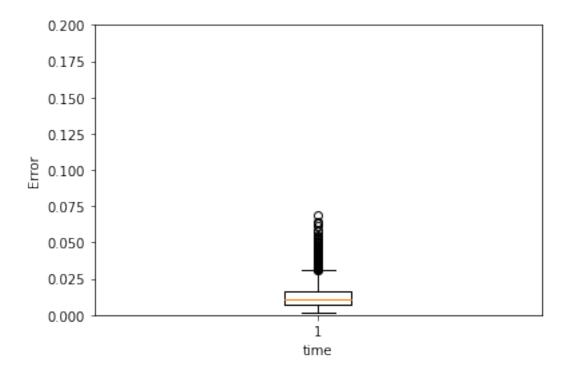
In [53]: train(model, tgen, vgen, name="lin20")
```









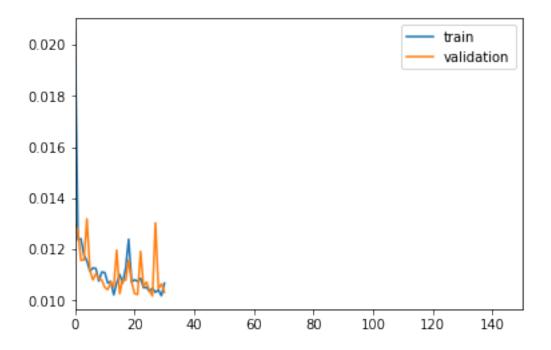


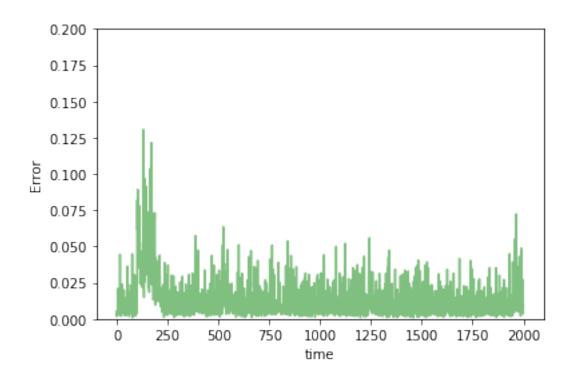
```
In [55]: TIMESTEPS = 50
        DIM = 29
        tgen = flat_generator(X, TIMESTEPS)
        vgen = flat_generator(val_X, TIMESTEPS)

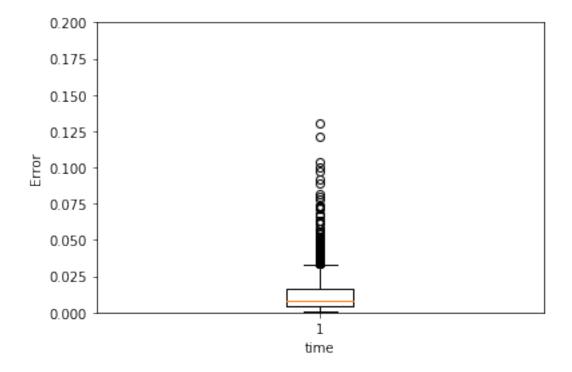
In [56]: input_layer = Input(shape=(TIMESTEPS*DIM,))
        output = Dense(DIM, activation='sigmoid')(input_layer)

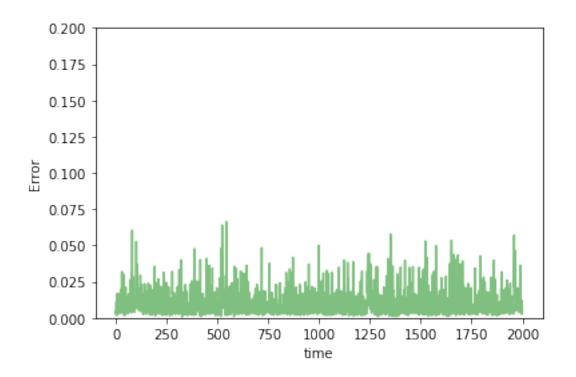
In [57]: model = Model(input_layer, output)
        model.compile(loss='mean_absolute_error', optimizer='adam', metrics=['mae'])

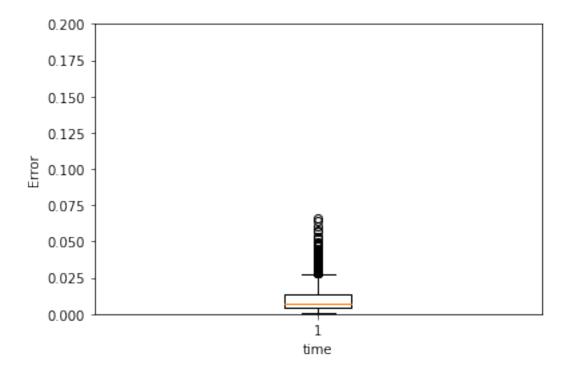
In [58]: train(model, tgen, vgen, name="lin50")
```



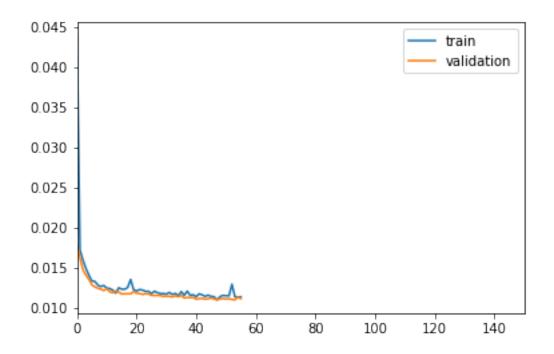


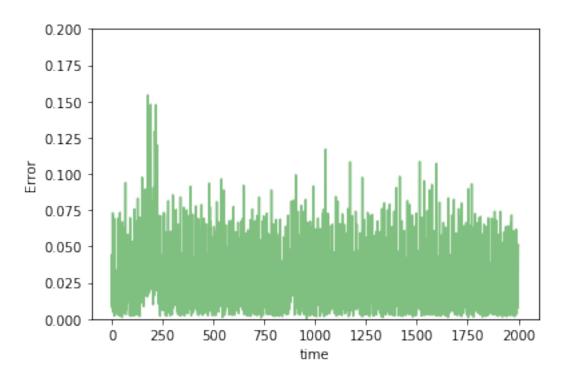


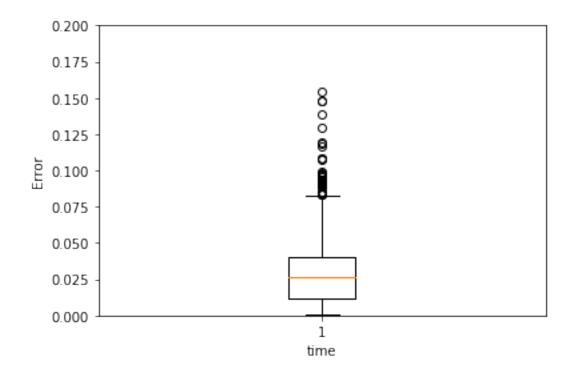


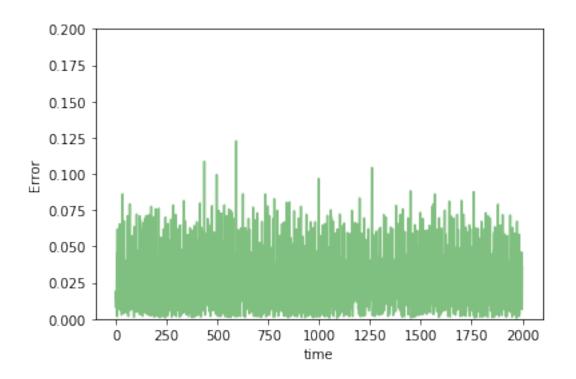


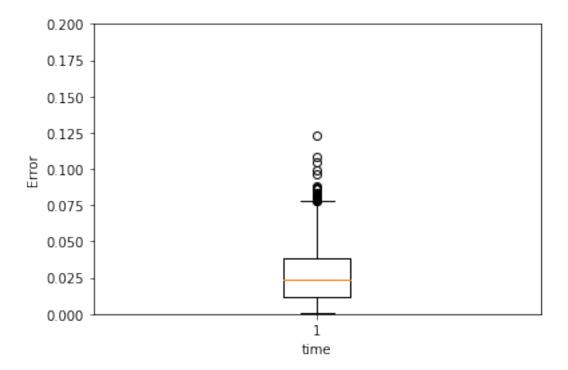
2.1.2 NN with 1 hidden layer

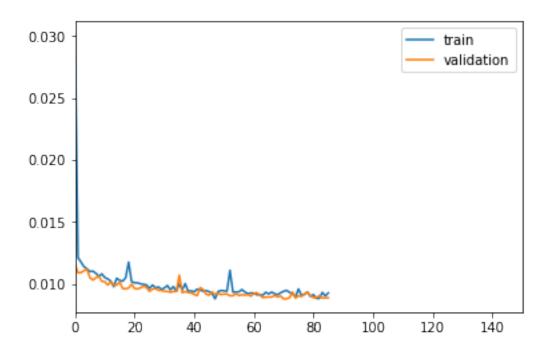


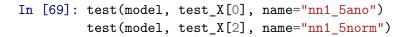


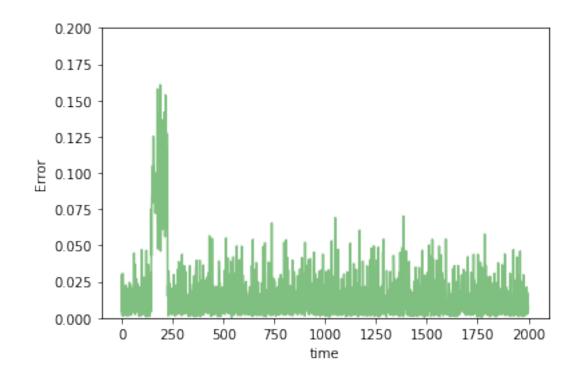


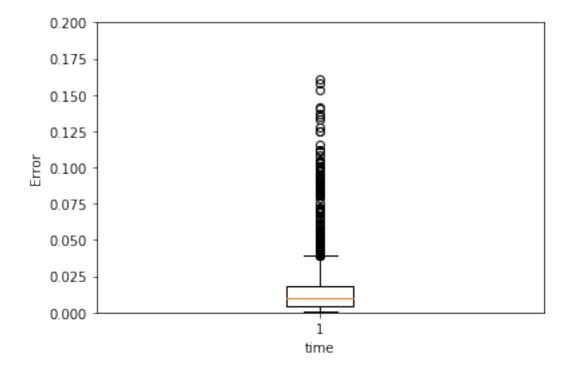


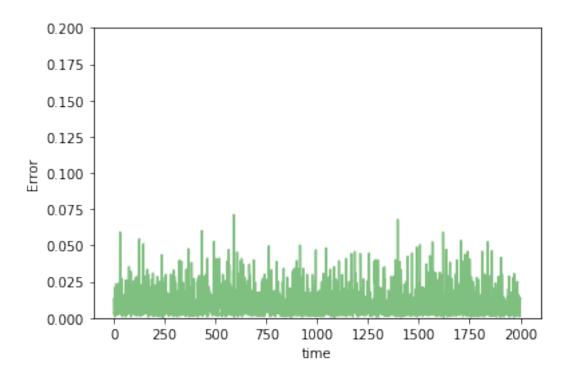


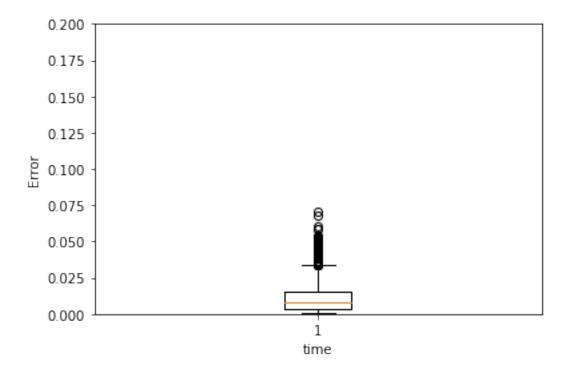


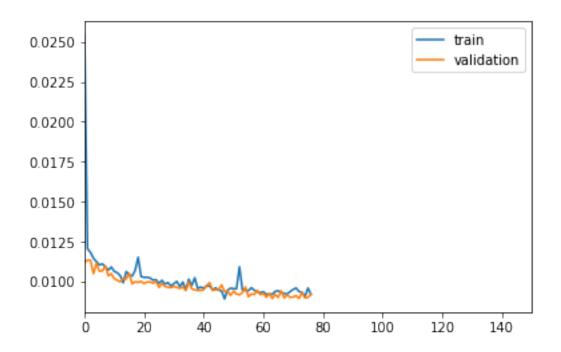


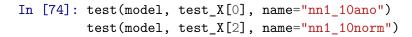


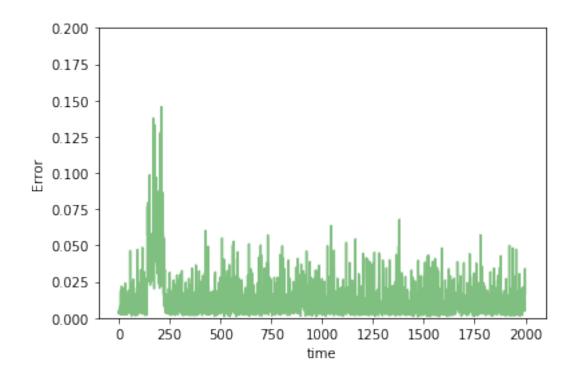


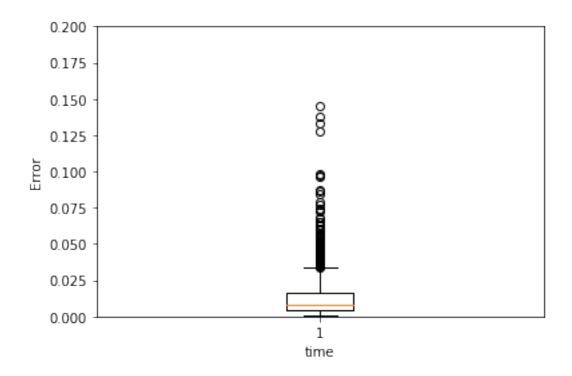


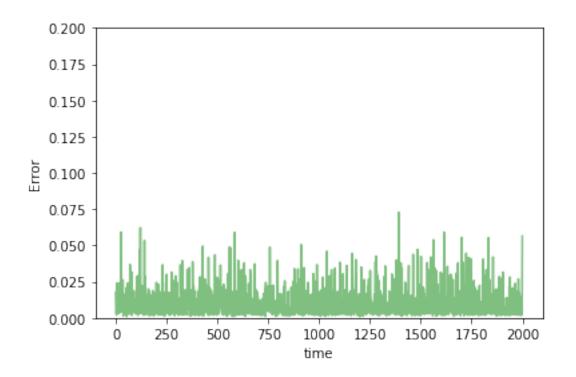


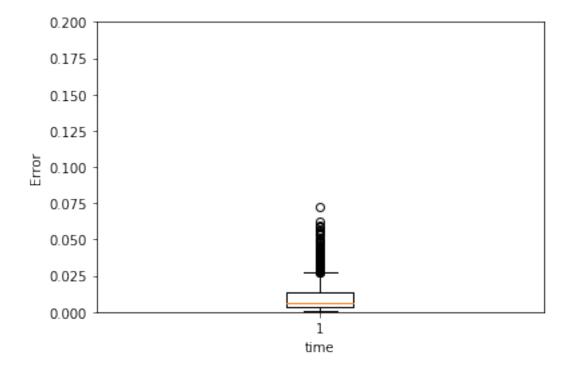


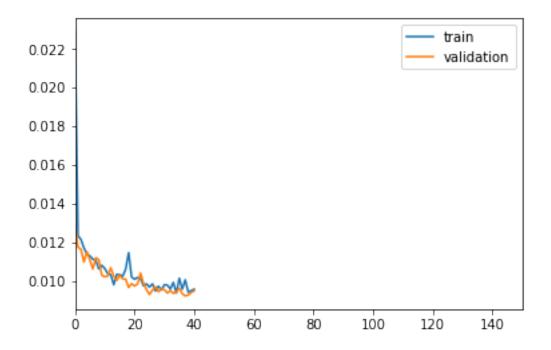


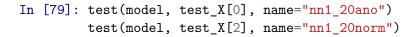


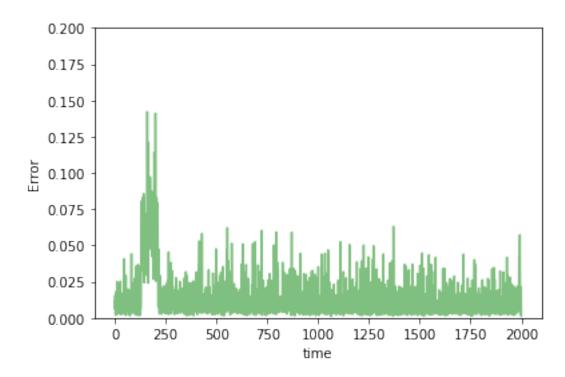


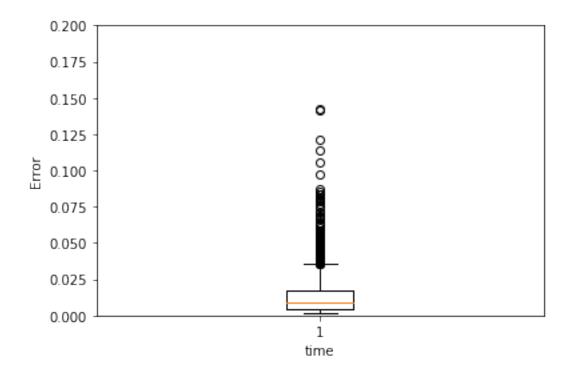


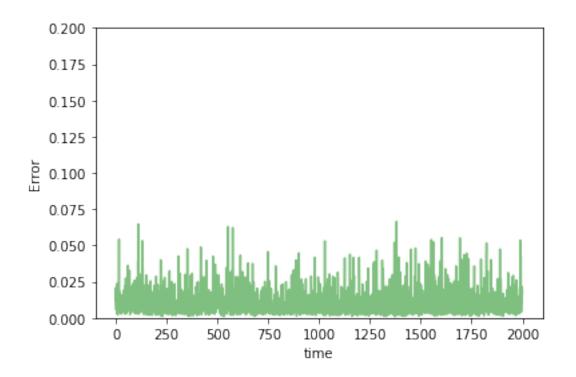


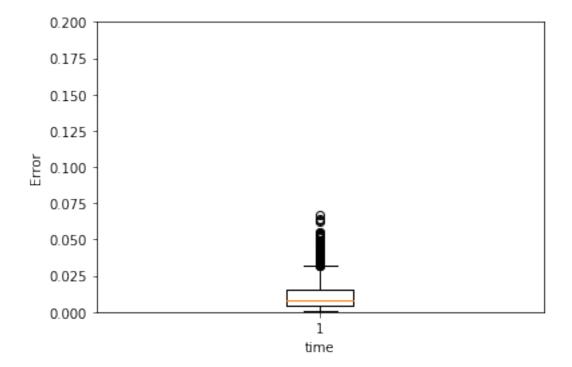


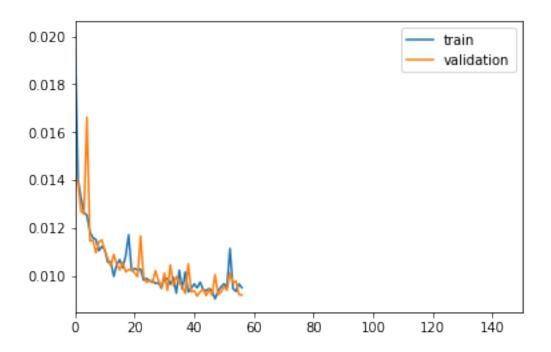


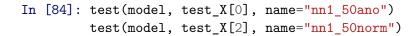


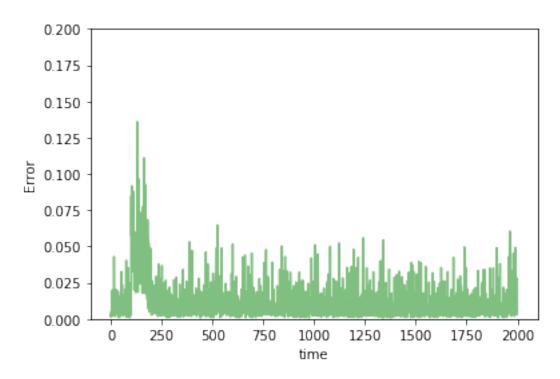


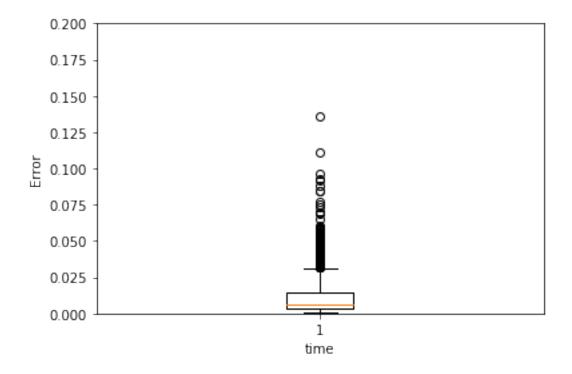


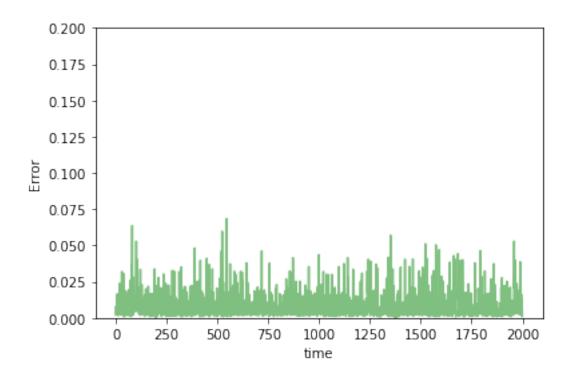


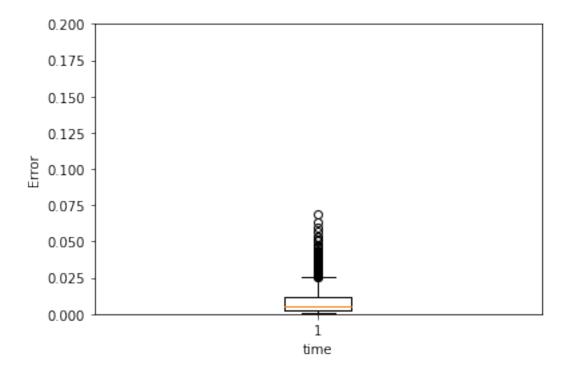




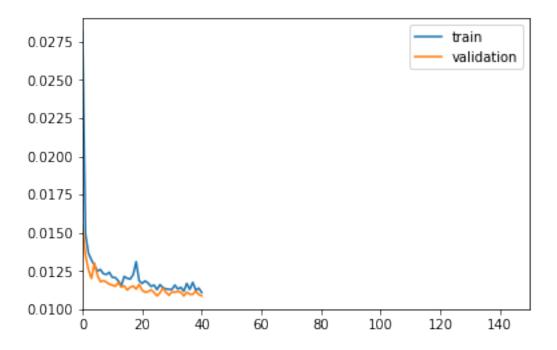


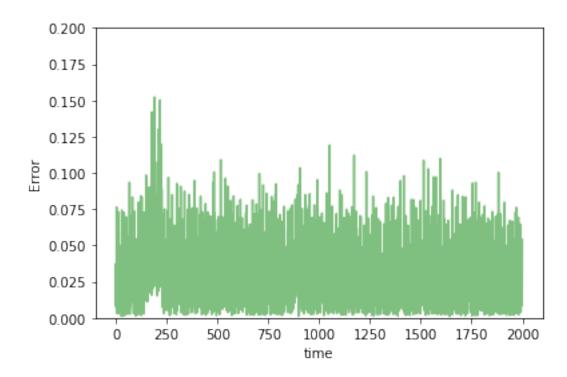


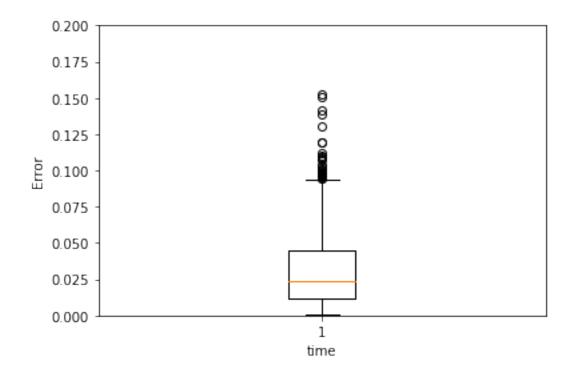


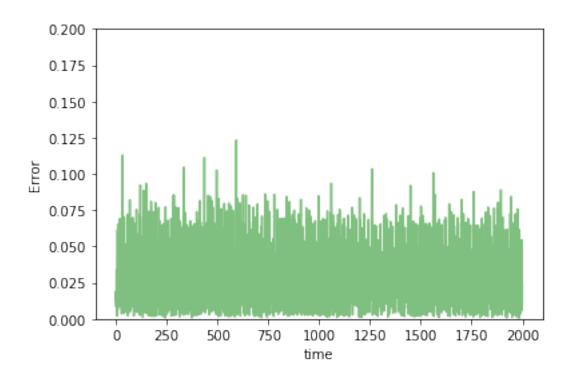


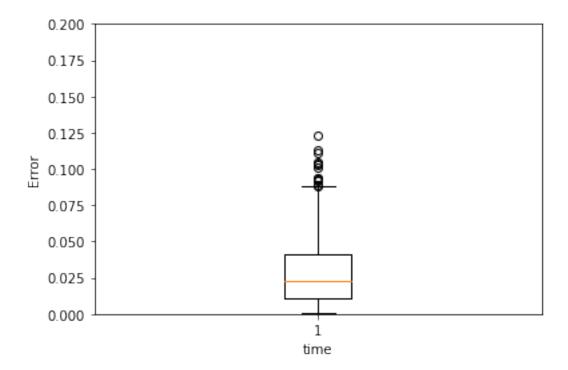
2.1.3 NN with 2 hidden layers









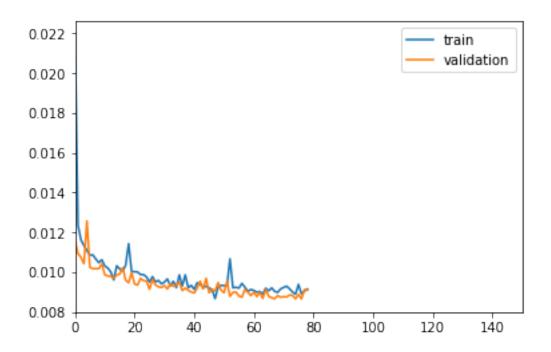


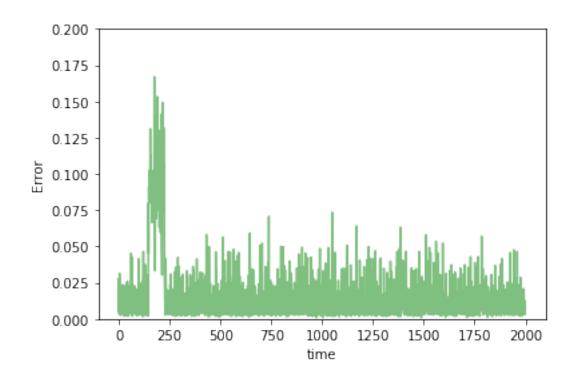
```
In [90]: TIMESTEPS = 5
        DIM = 29
        tgen = flat_generator(X, TIMESTEPS)
        vgen = flat_generator(val_X, TIMESTEPS)

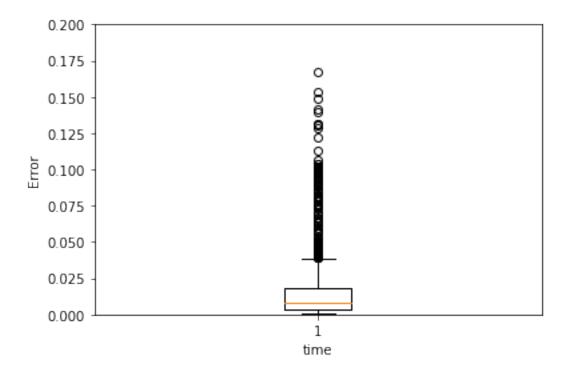
In [91]: input_layer = Input(shape=(TIMESTEPS*DIM,))
        hidden = Dense(500, activation='relu')(input_layer)
        hidden = Dense(100, activation='relu')(hidden)
        output = Dense(DIM, activation='sigmoid')(hidden)

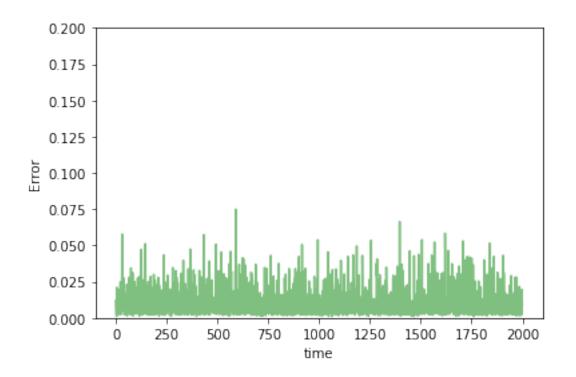
In [92]: model = Model(input_layer, output)
        model.compile(loss='mean_absolute_error', optimizer='adam', metrics=['mae'])

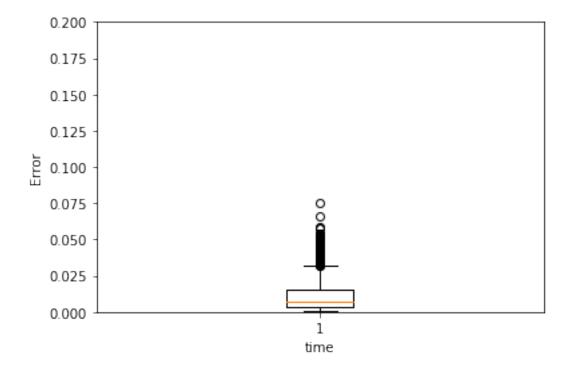
In [93]: train(model, tgen, vgen, name="nn2_5")
```









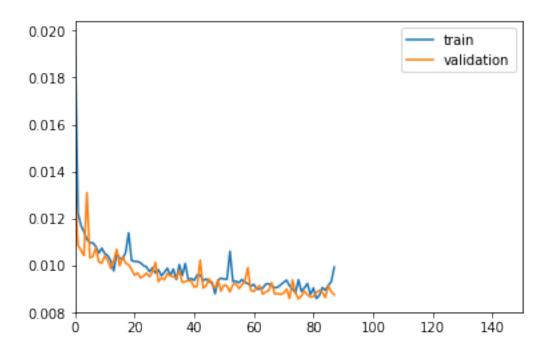


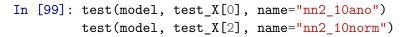
```
In [95]: TIMESTEPS = 10
        DIM = 29
        tgen = flat_generator(X, TIMESTEPS)
        vgen = flat_generator(val_X, TIMESTEPS)

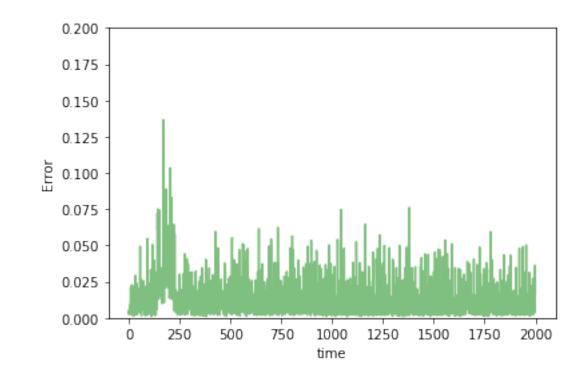
In [96]: input_layer = Input(shape=(TIMESTEPS*DIM,))
        hidden = Dense(500, activation='relu')(input_layer)
        hidden = Dense(100, activation='relu')(hidden)
        output = Dense(DIM, activation='sigmoid')(hidden)

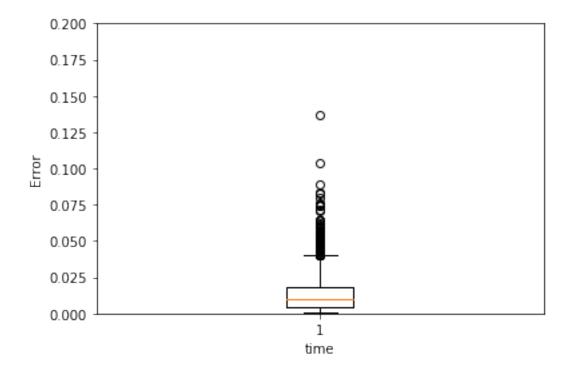
In [97]: model = Model(input_layer, output)
        model.compile(loss='mean_absolute_error', optimizer='adam', metrics=['mae'])

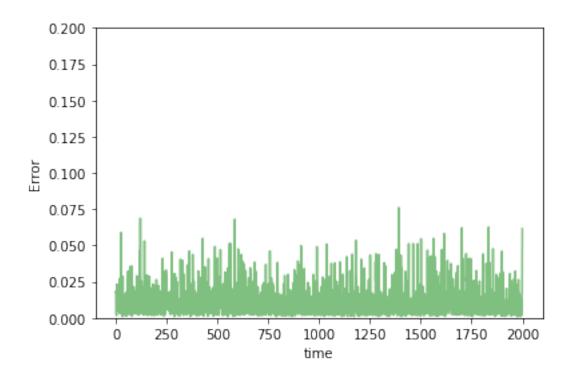
In [98]: train(model, tgen, vgen, name="nn2_10")
```

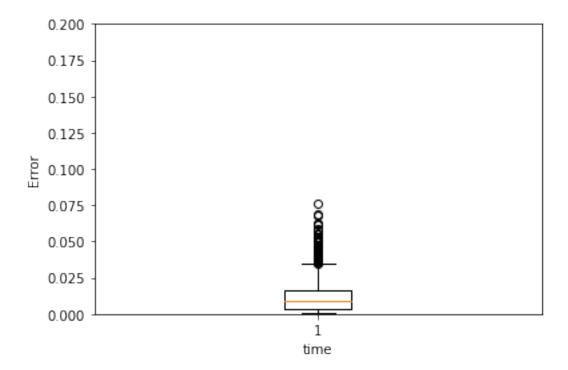


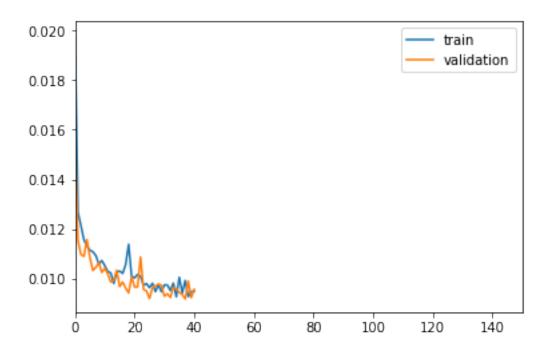


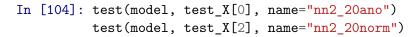


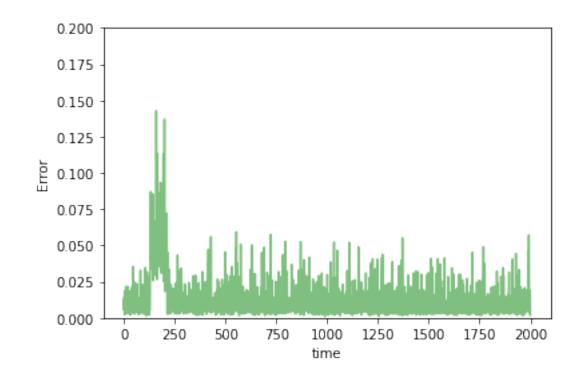


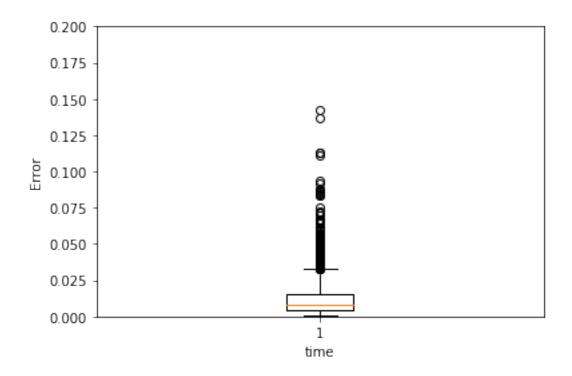


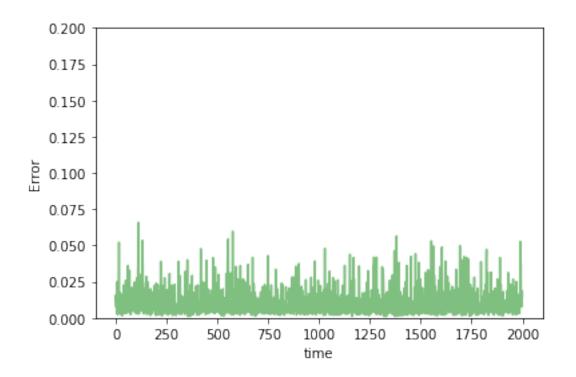


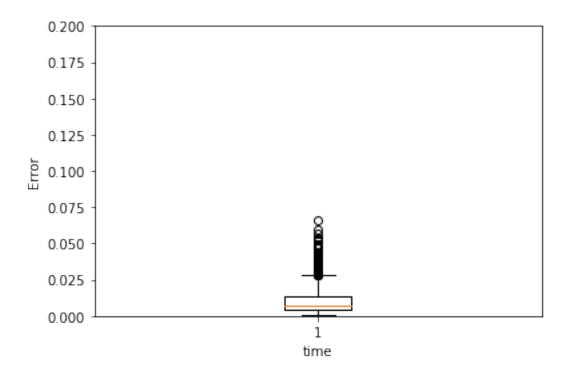


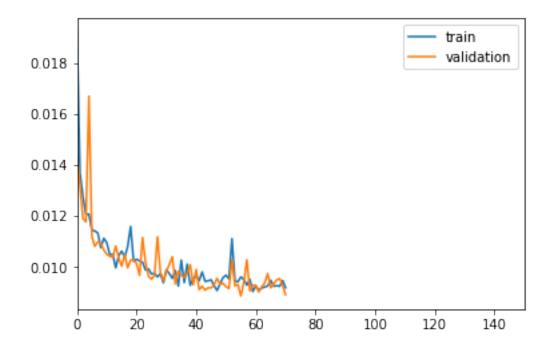


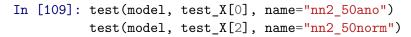


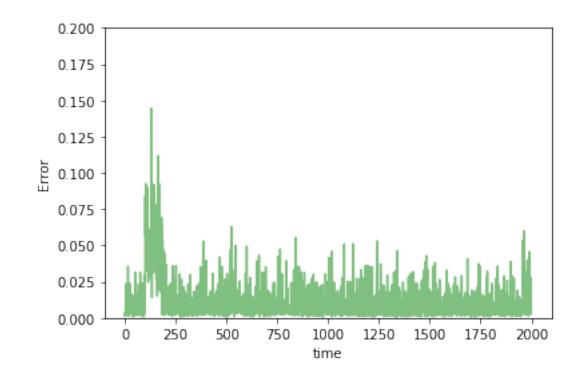


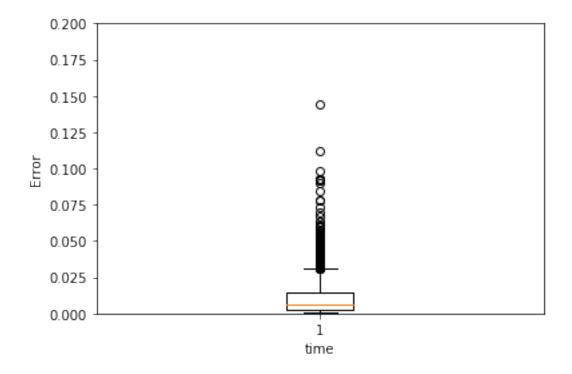


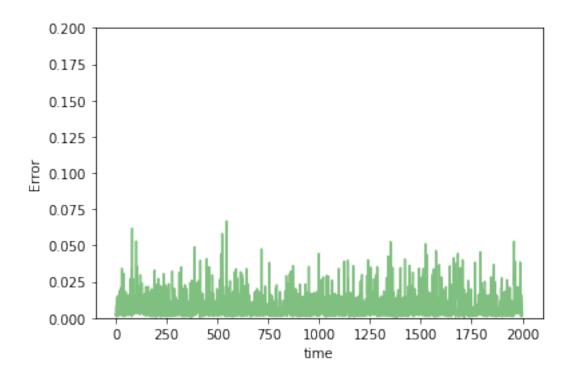


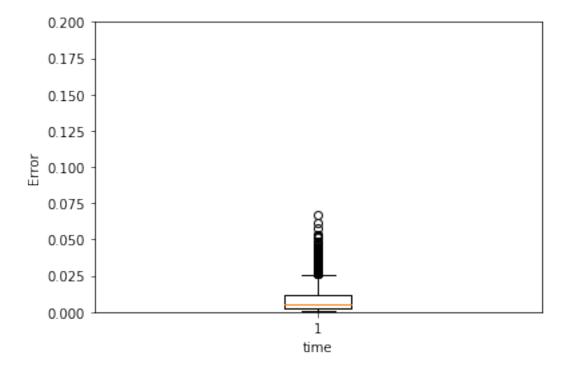




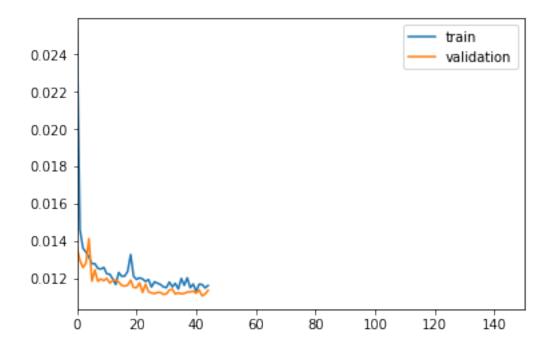


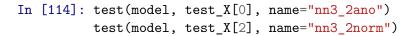


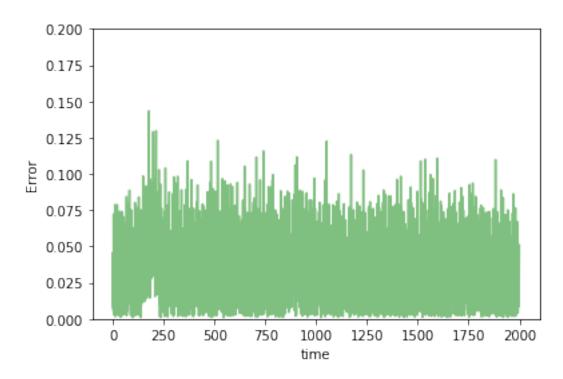


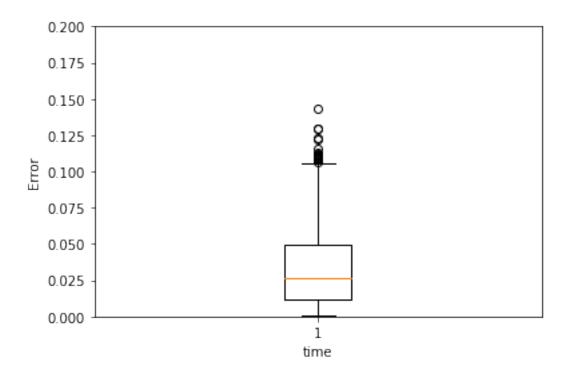


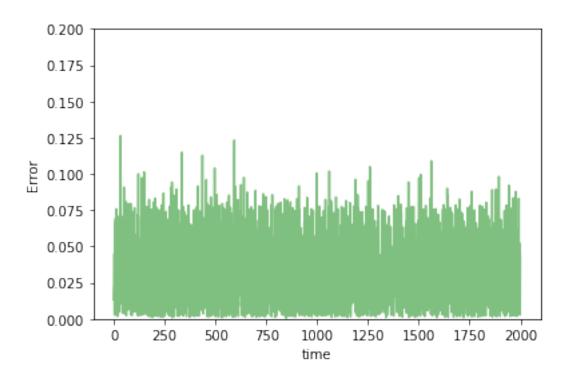
2.1.4 NN with 3 hidden layers

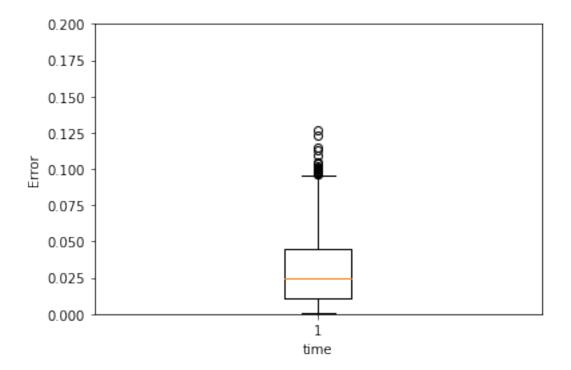


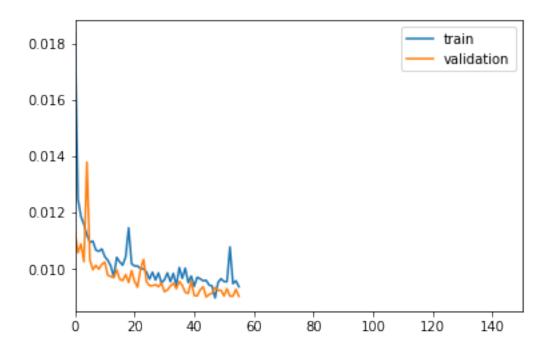


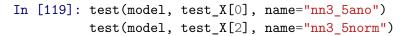


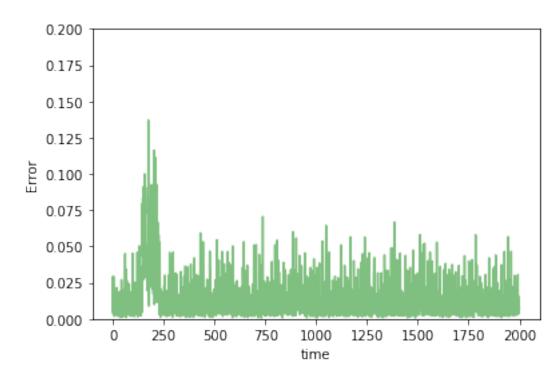


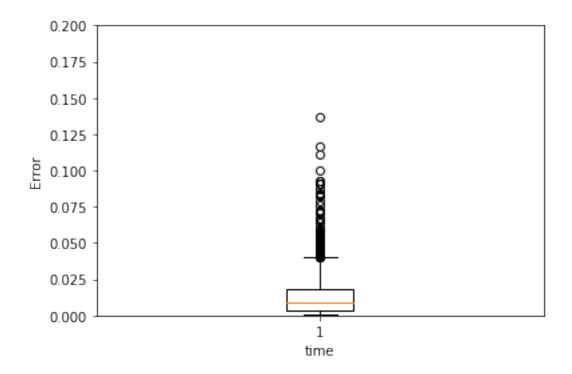


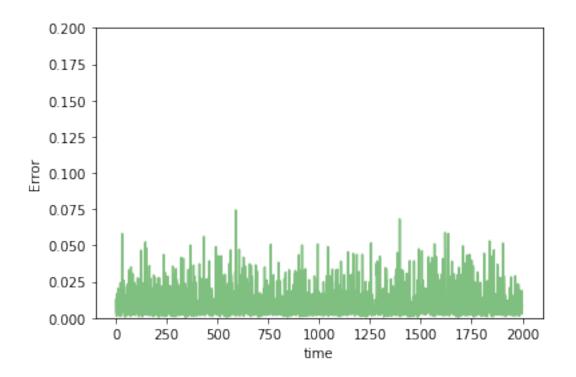


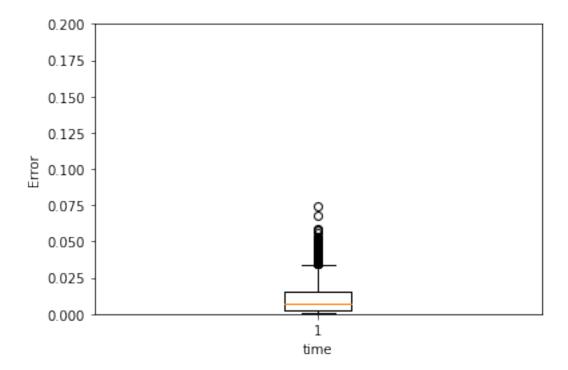


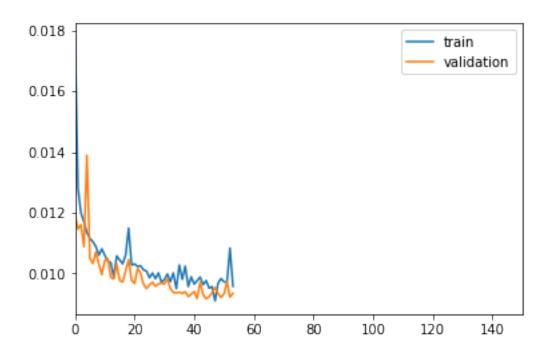


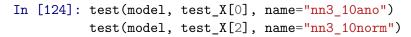


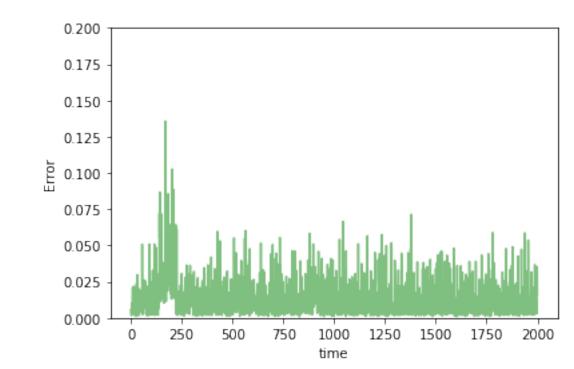


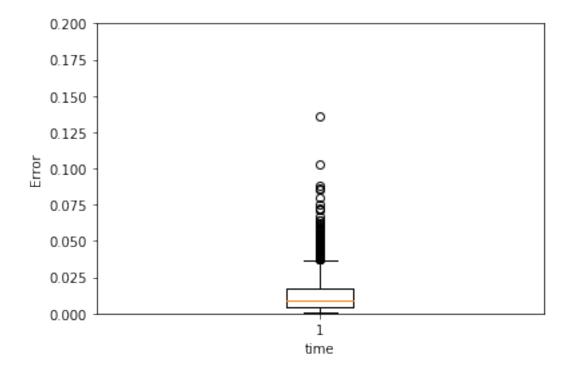


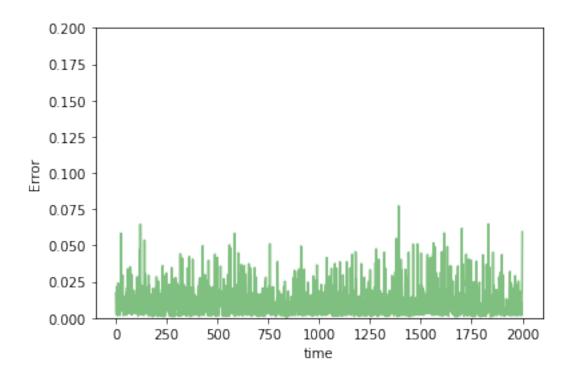


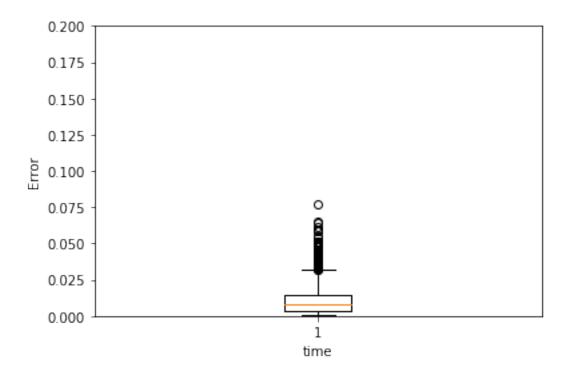


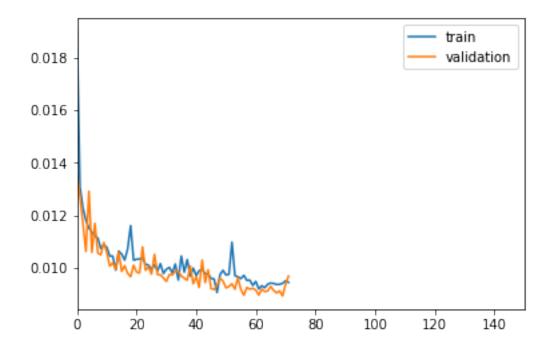


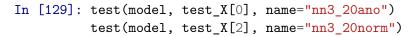


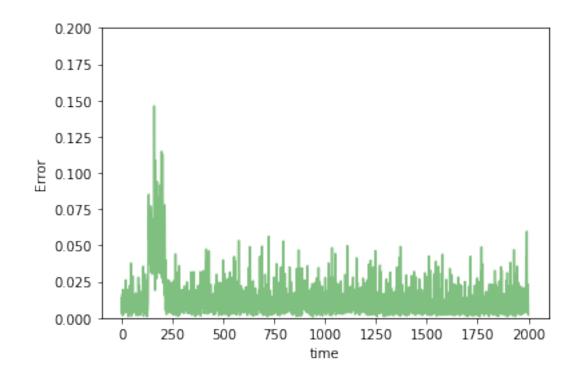


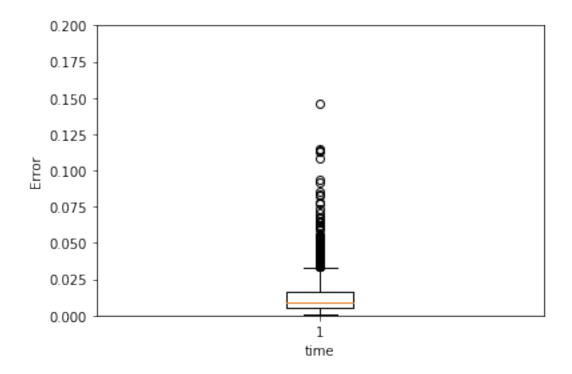


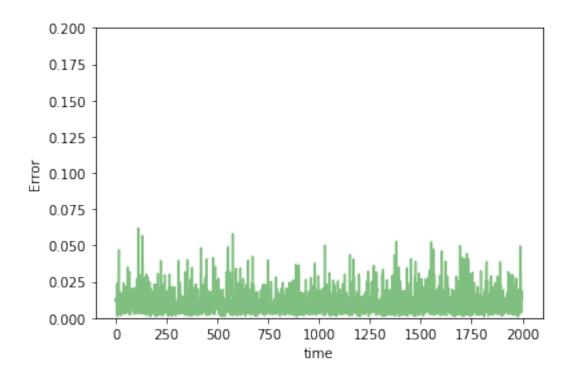


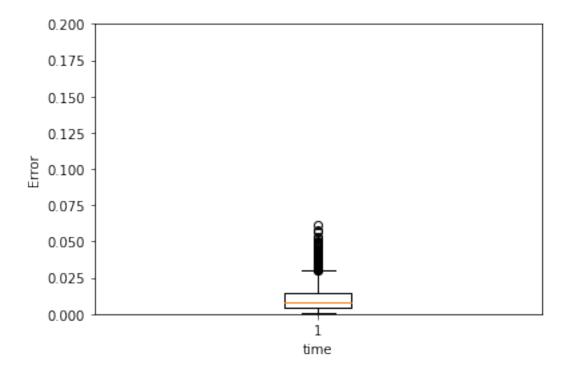


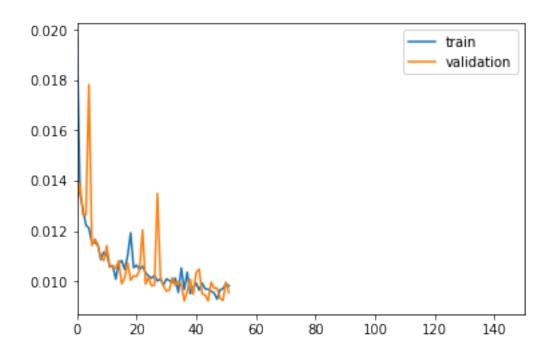


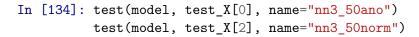


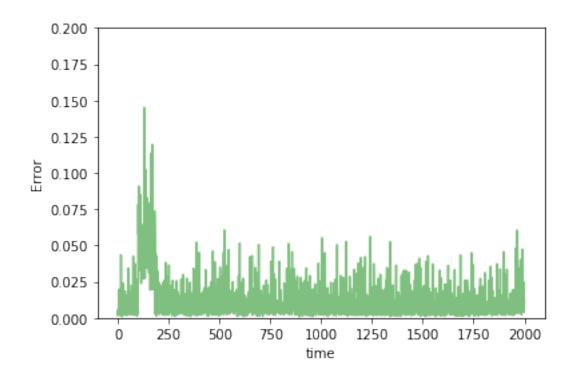


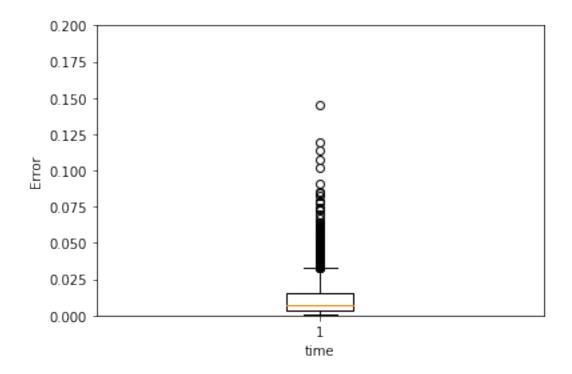


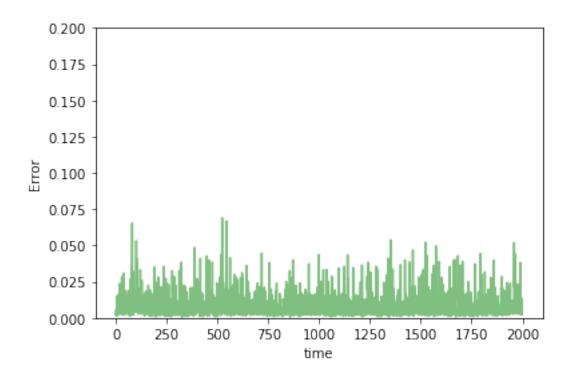


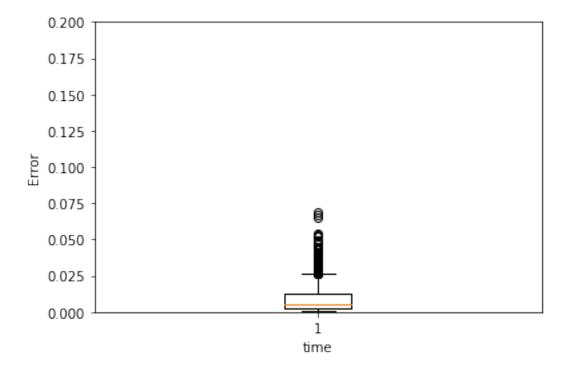




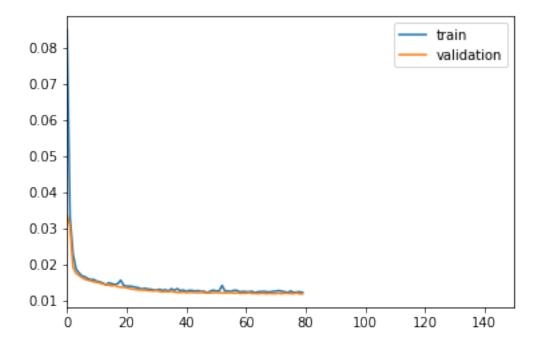


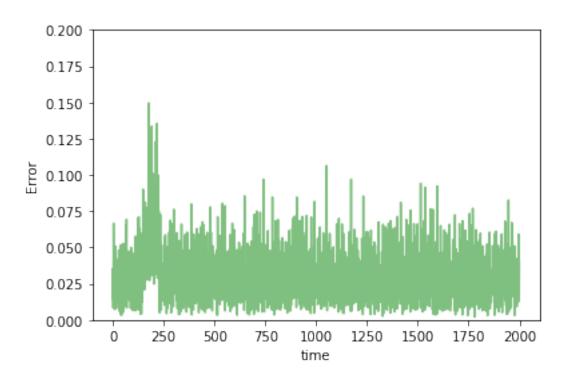


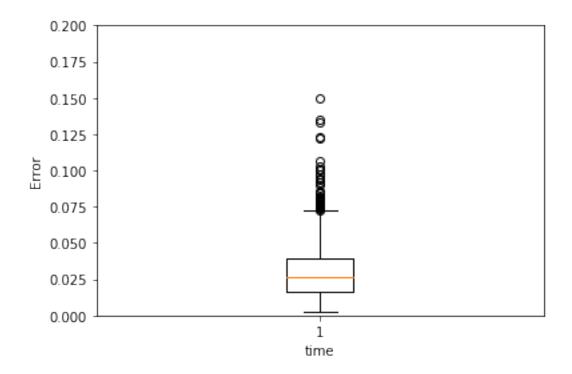


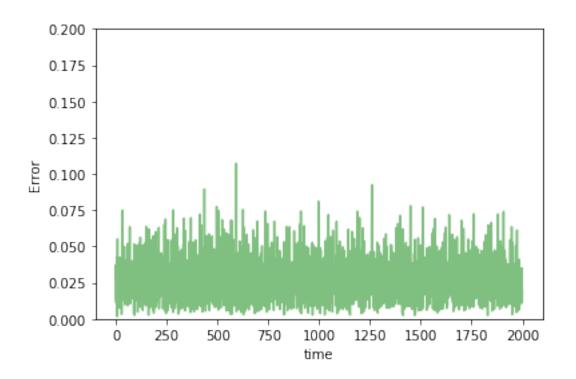


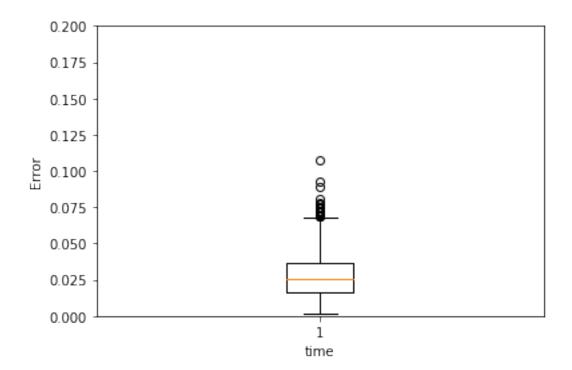
2.1.5 RNN with 1 GRU layers

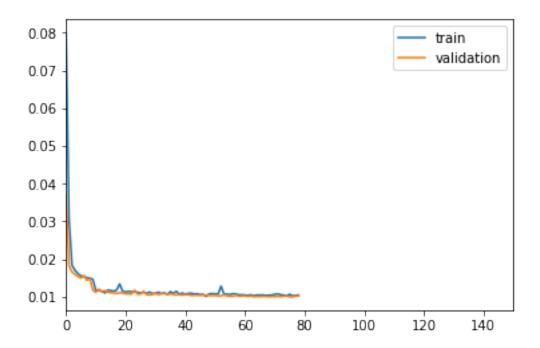


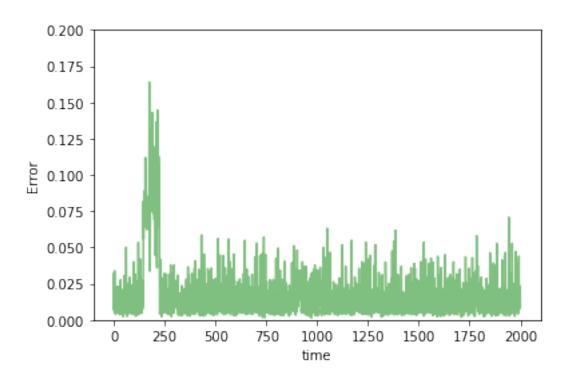


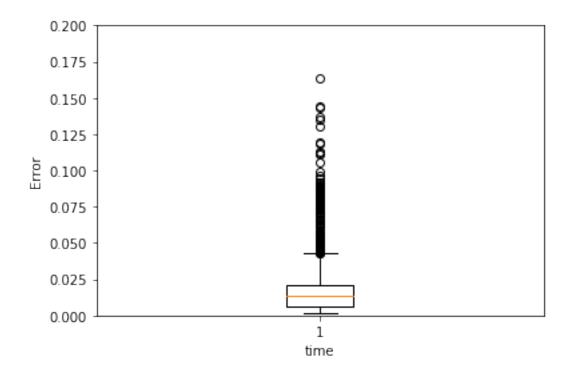


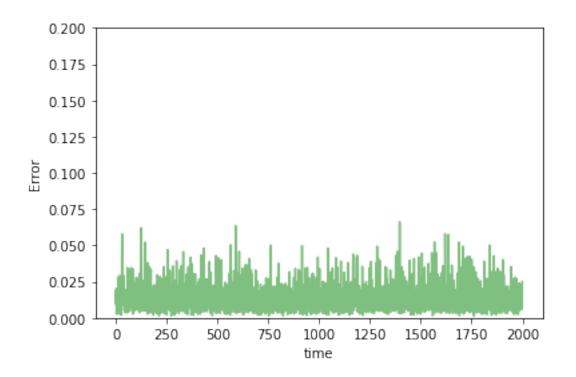


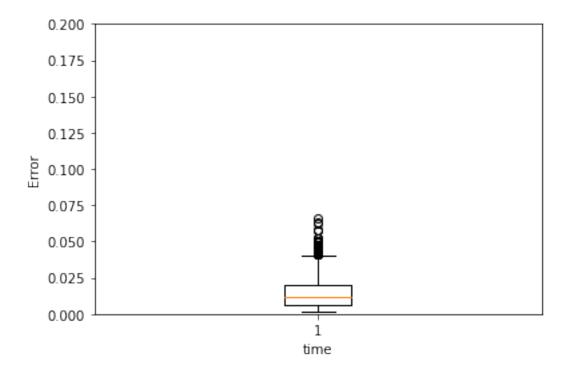


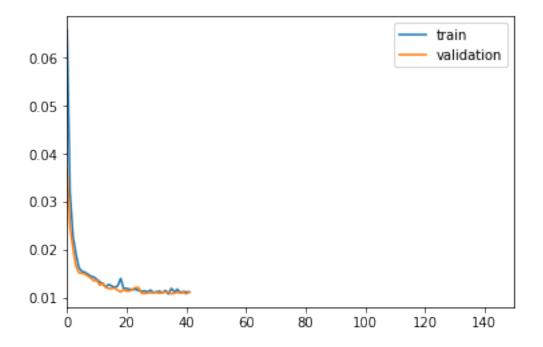


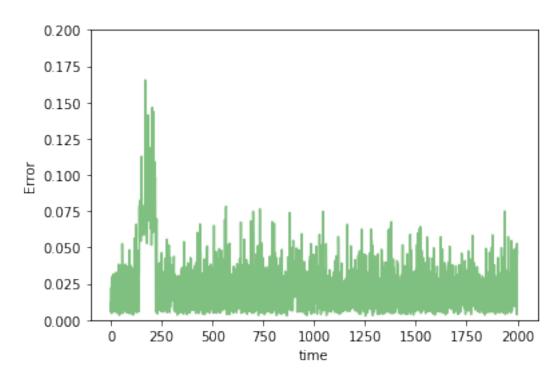


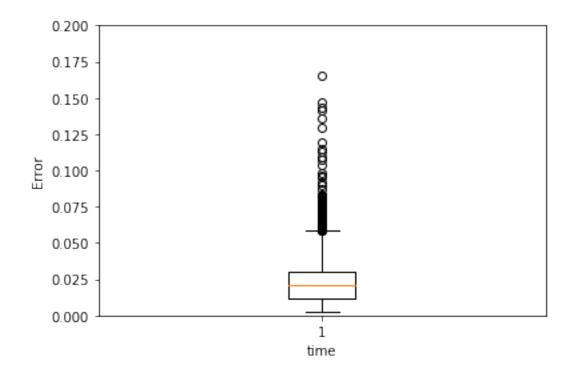


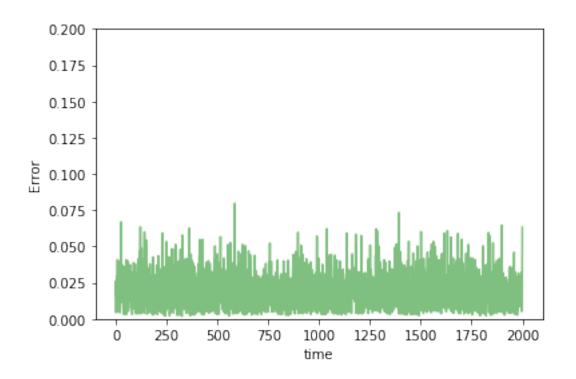


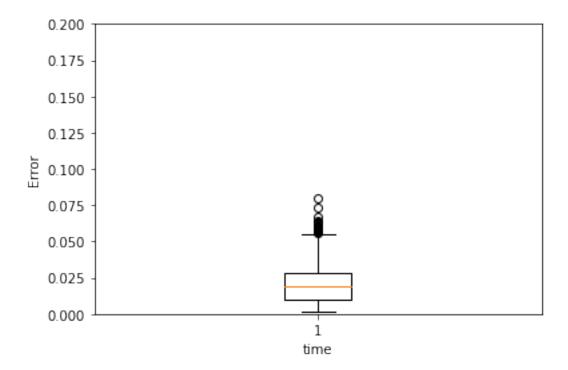


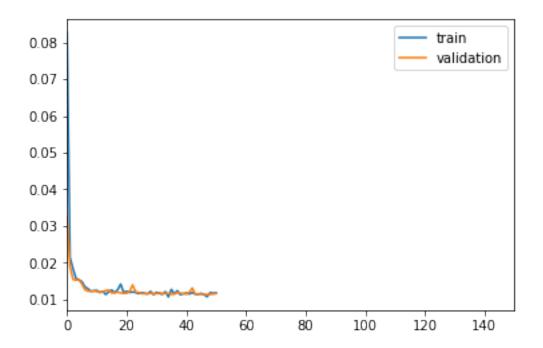


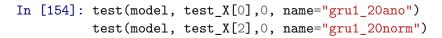


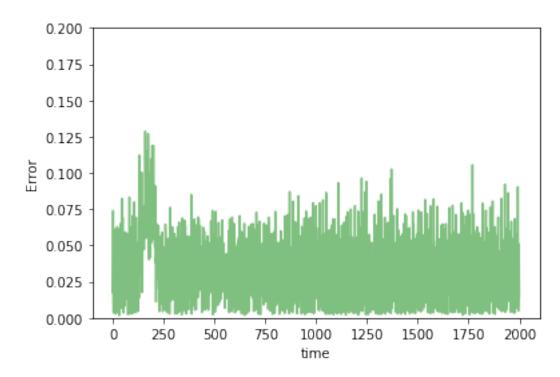


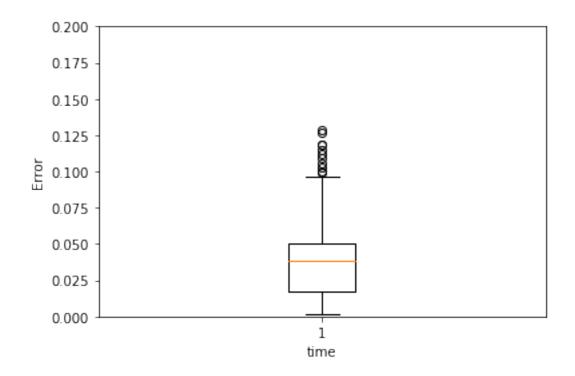


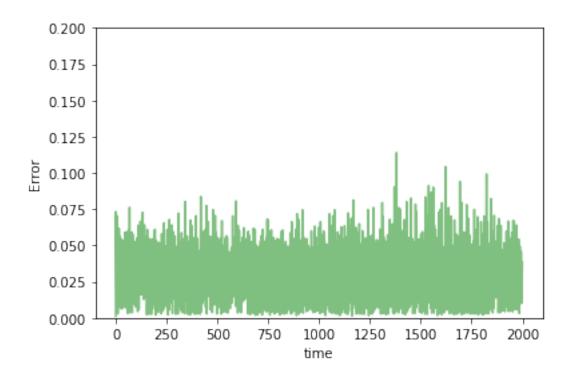


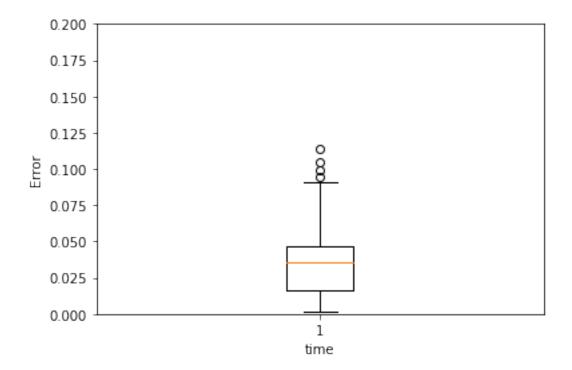


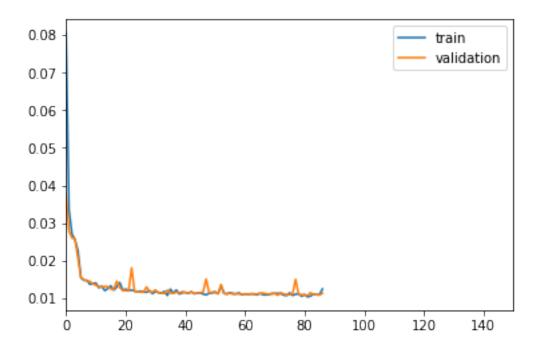


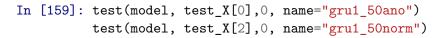


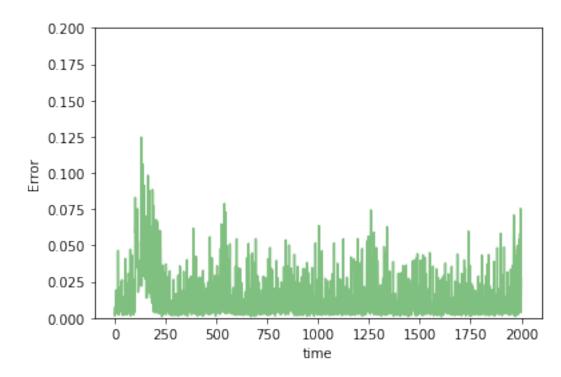


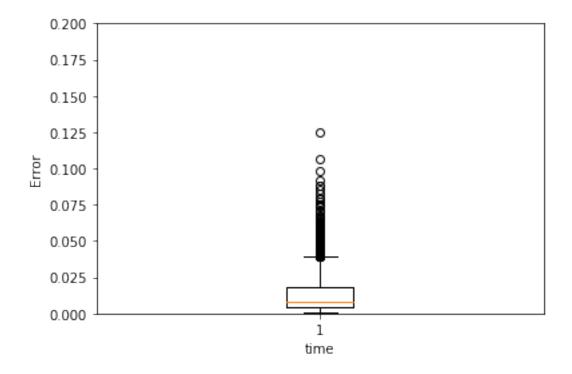


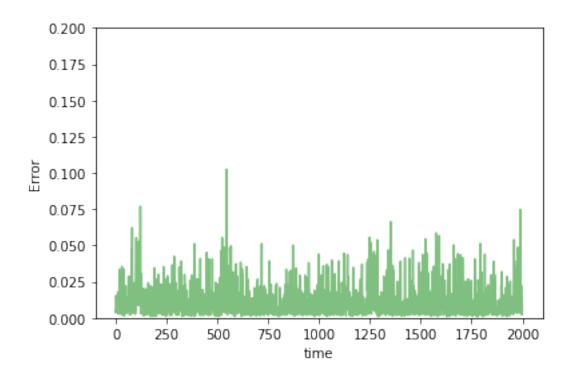


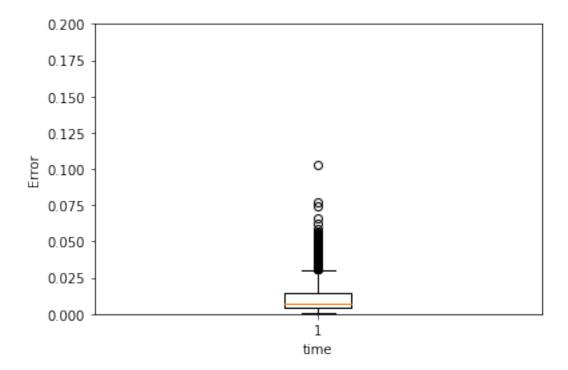




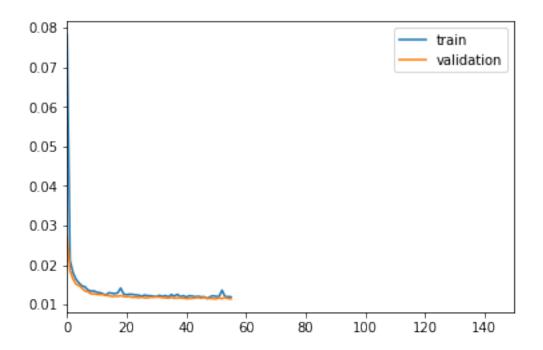


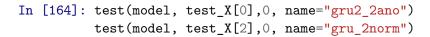


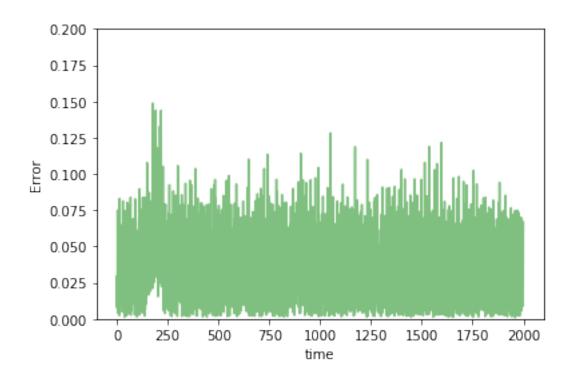


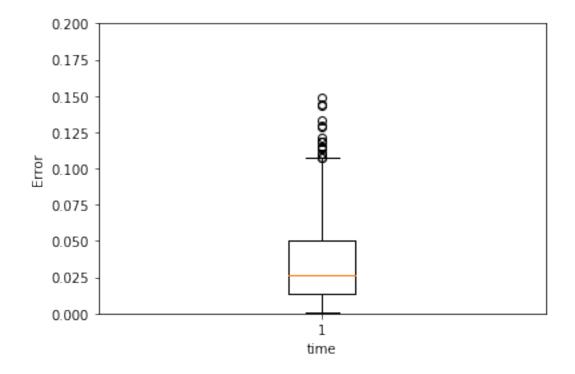


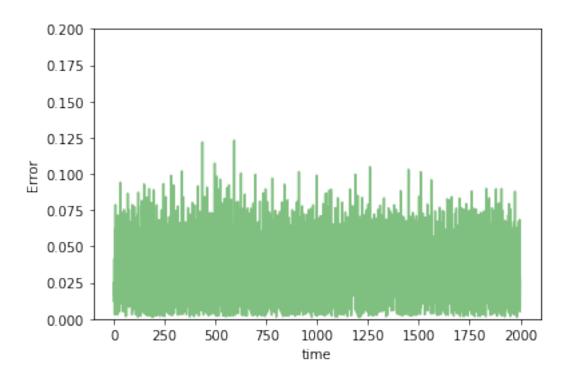
2.1.6 RNN with 2 GRU layers

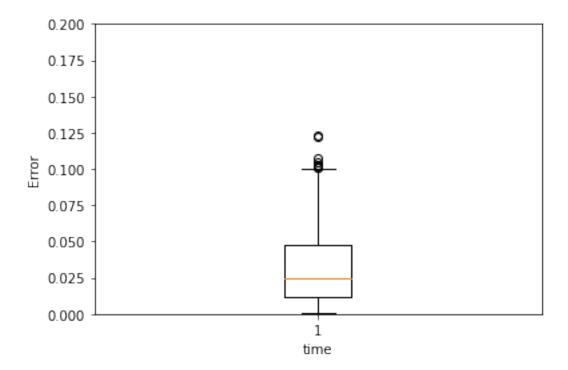


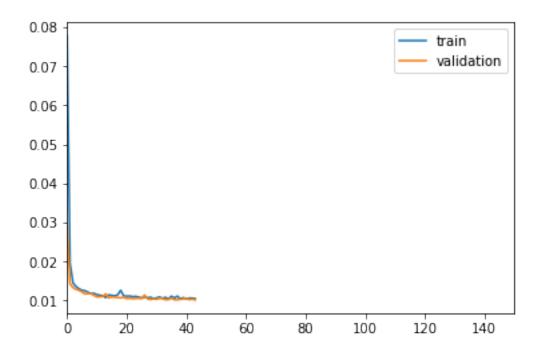


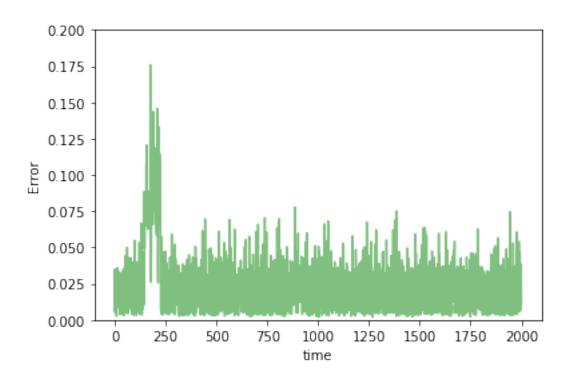


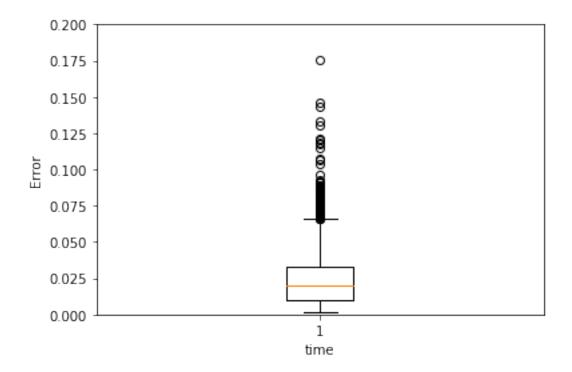


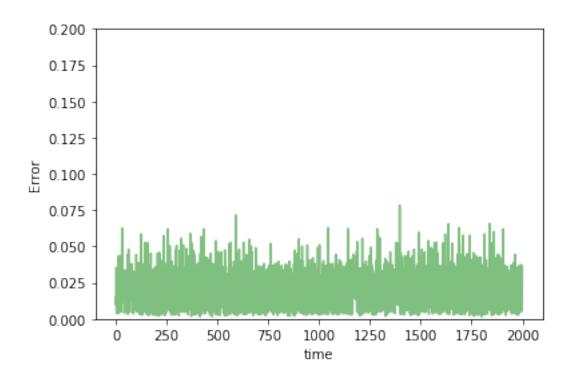


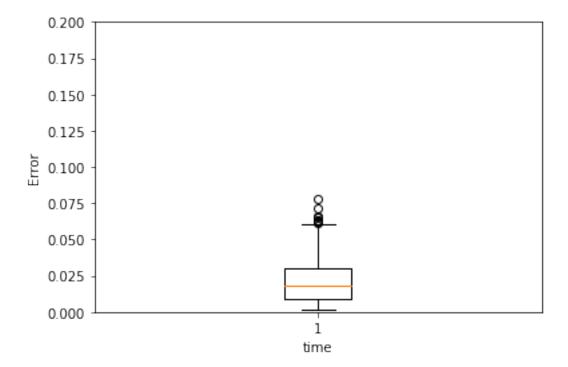


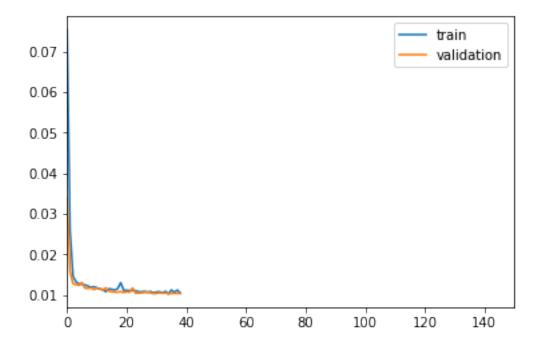


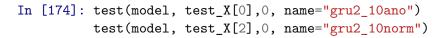


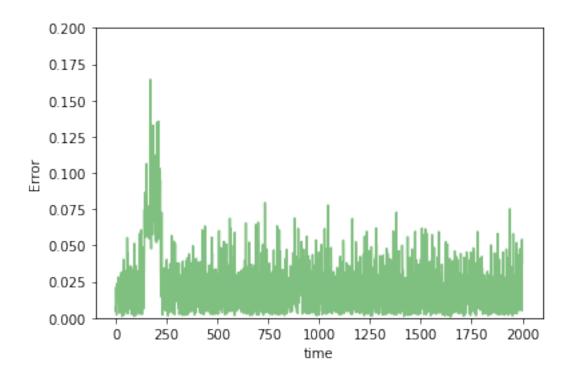


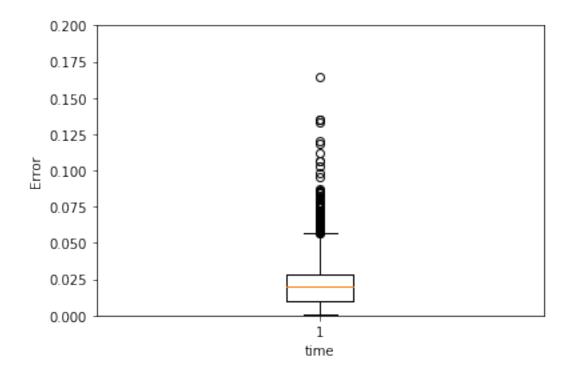


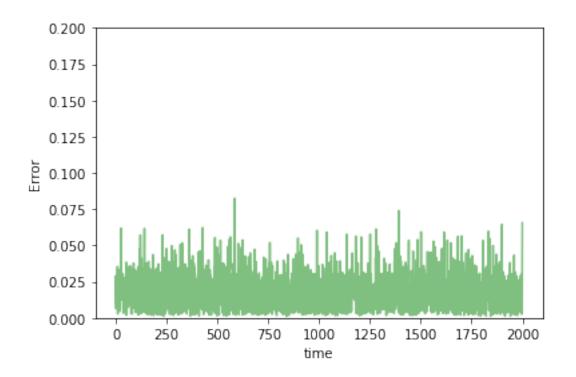


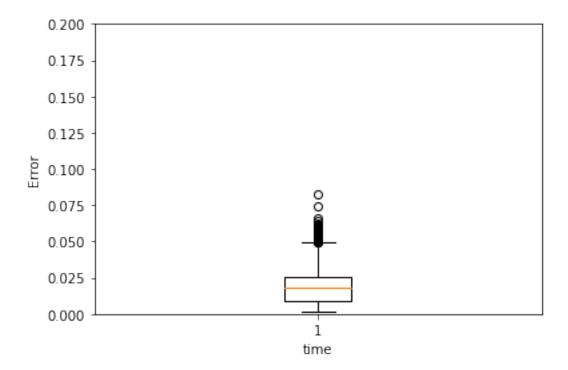


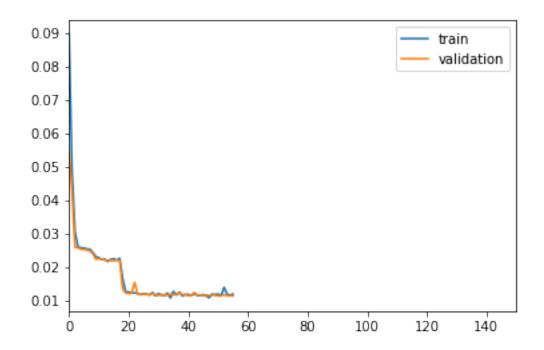


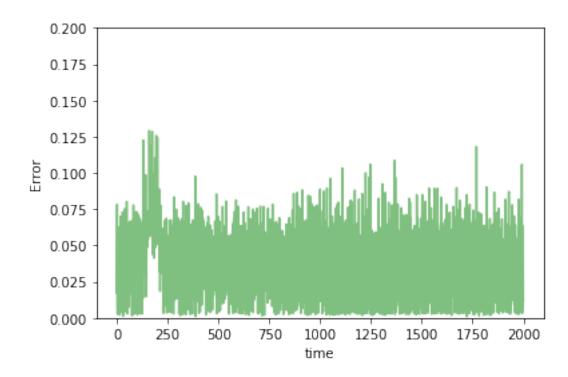


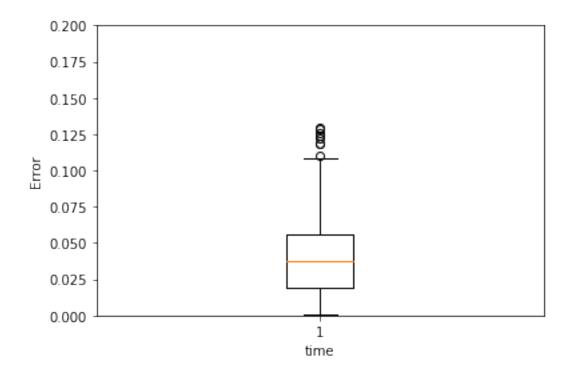


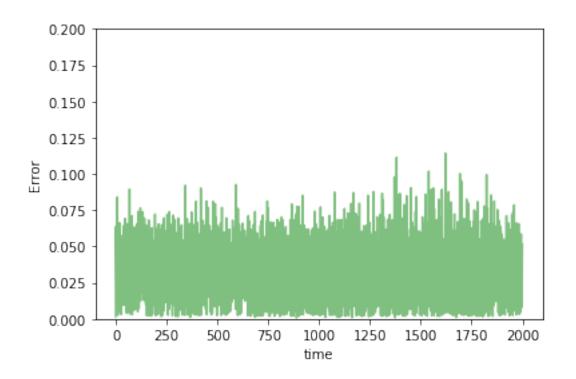


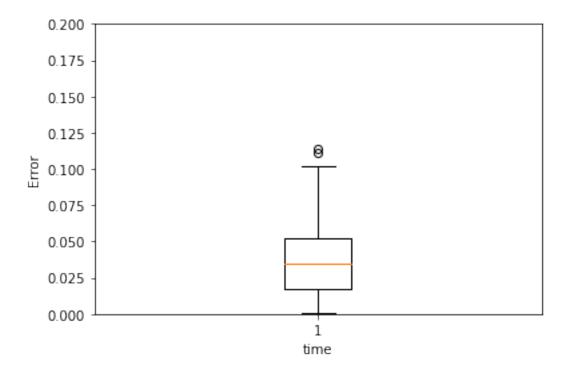


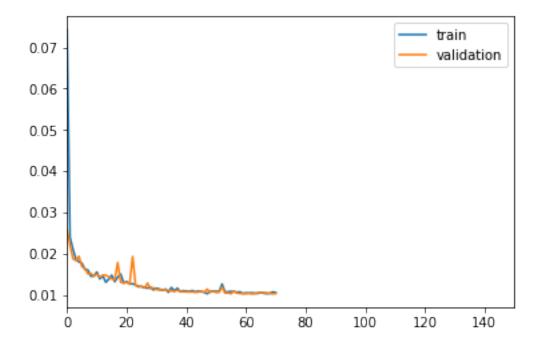


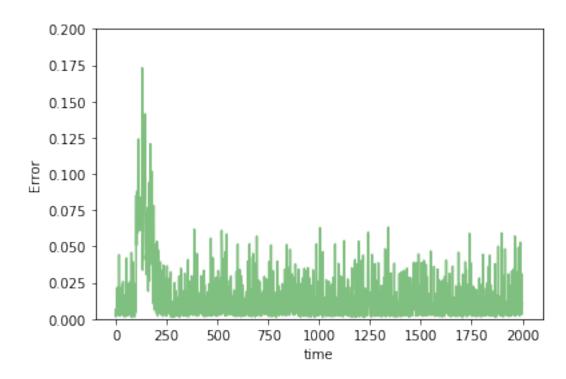


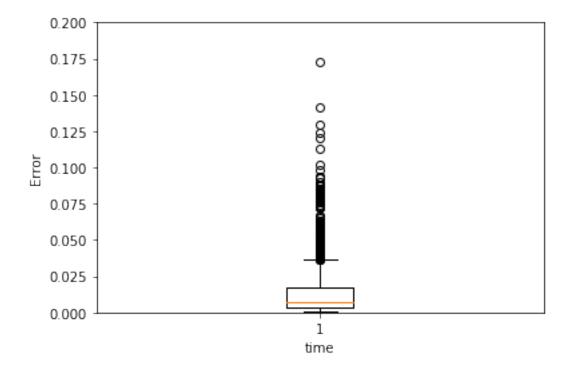


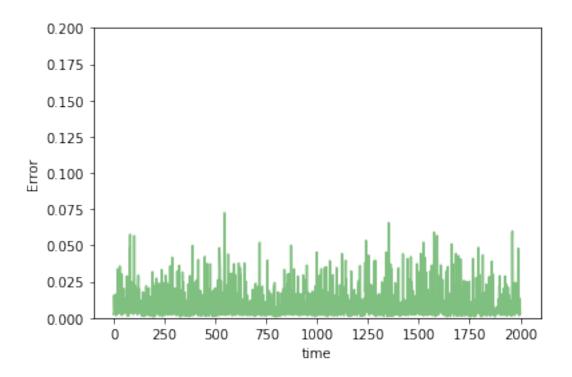


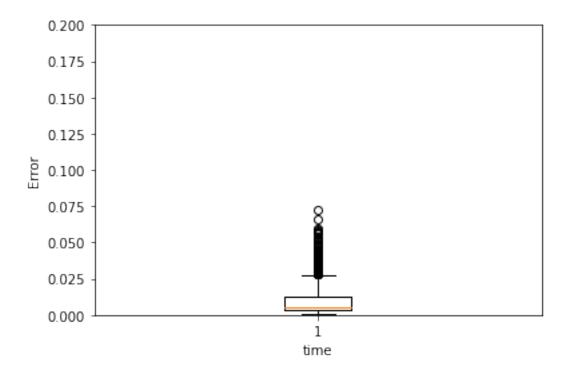




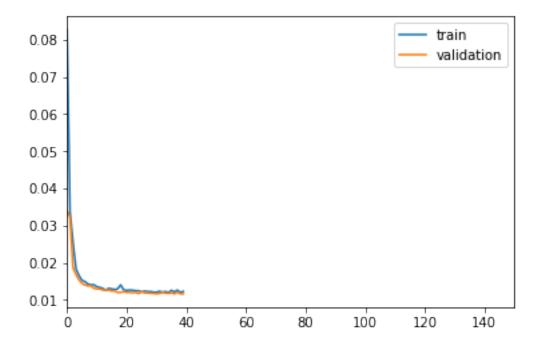


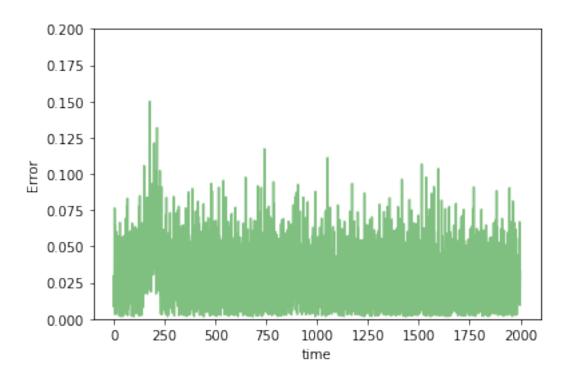


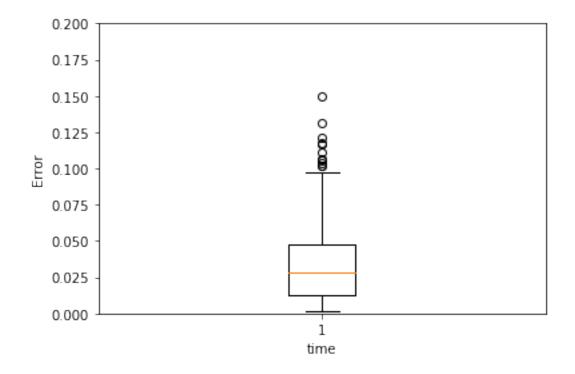


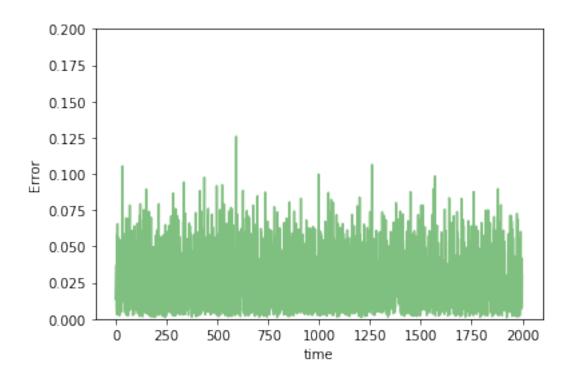


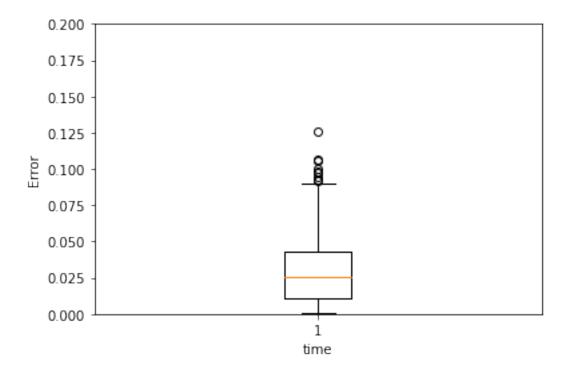
2.1.7 RNN with 3 GRU layers

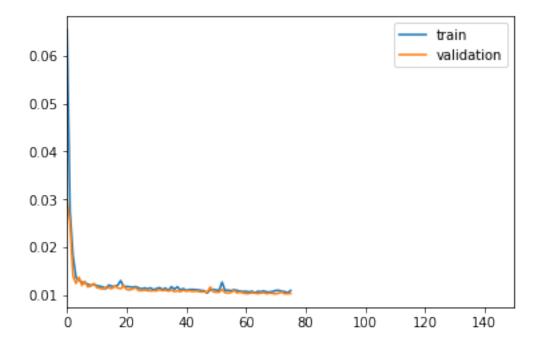


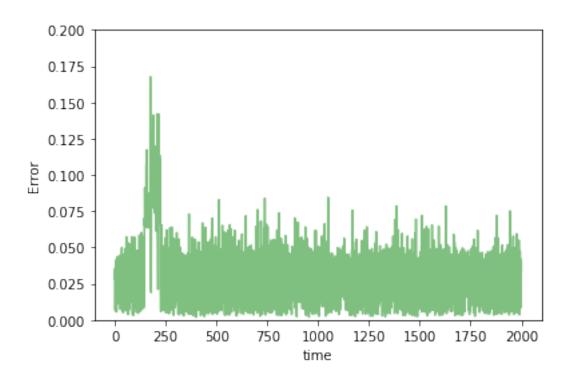


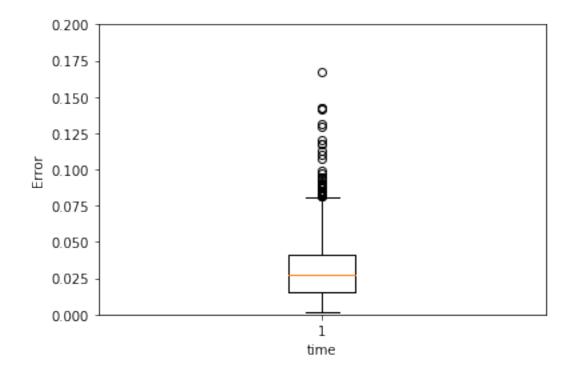


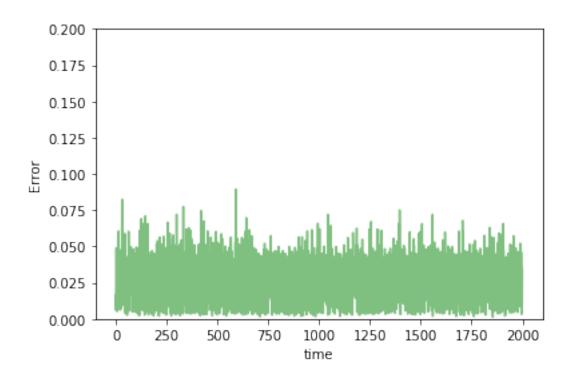


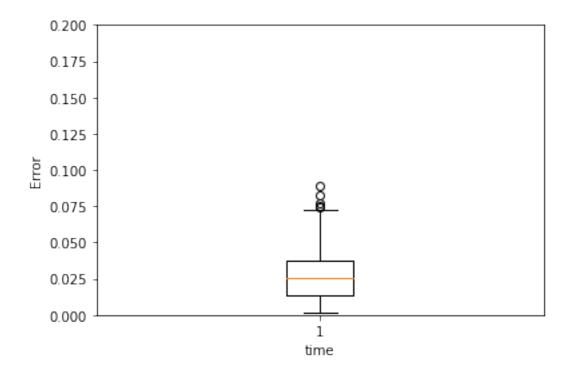


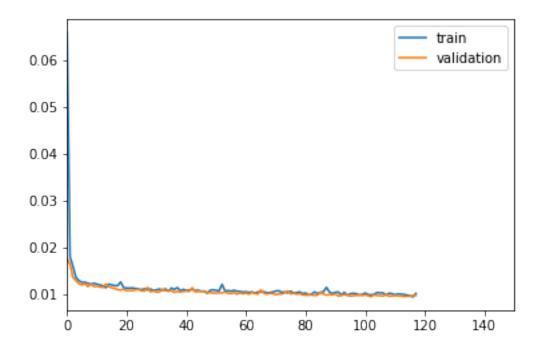


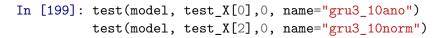


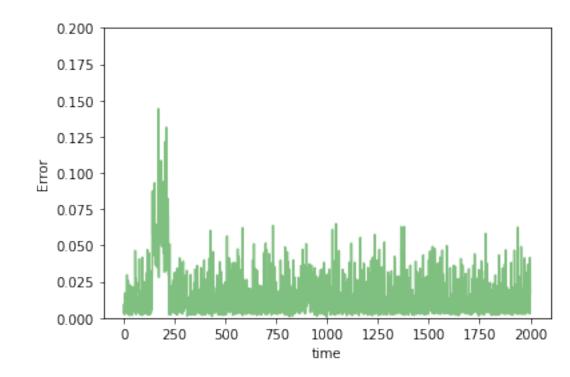


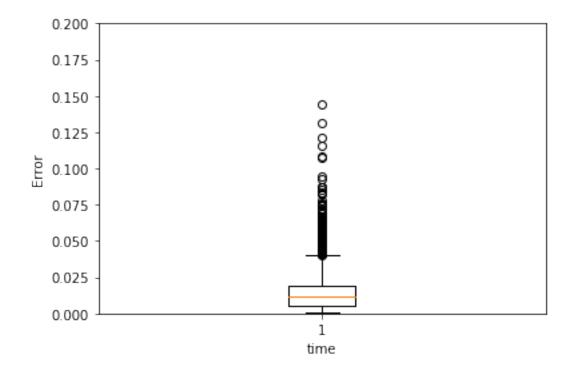


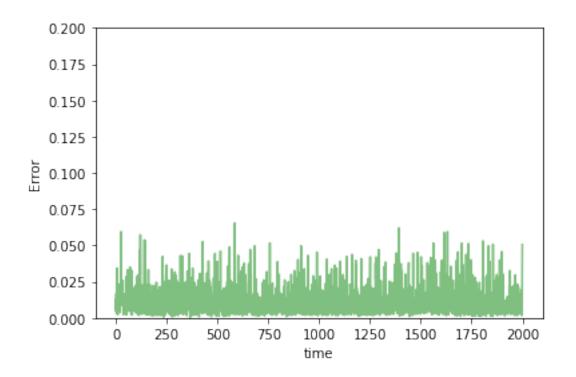


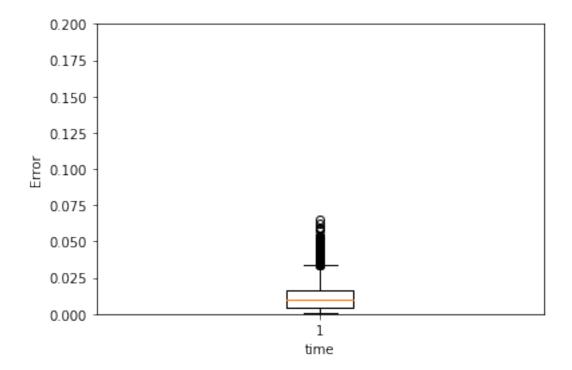


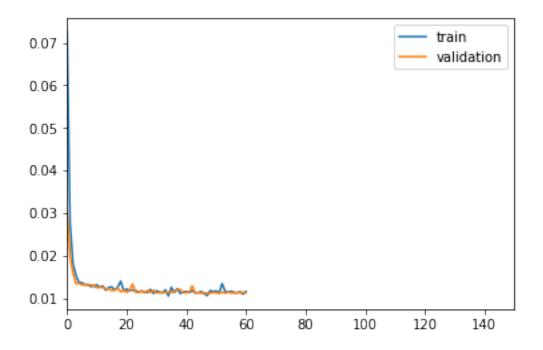


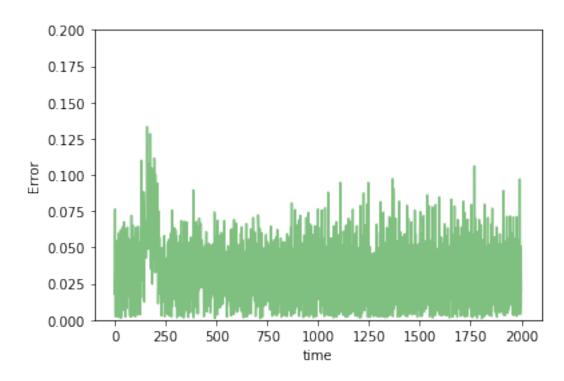


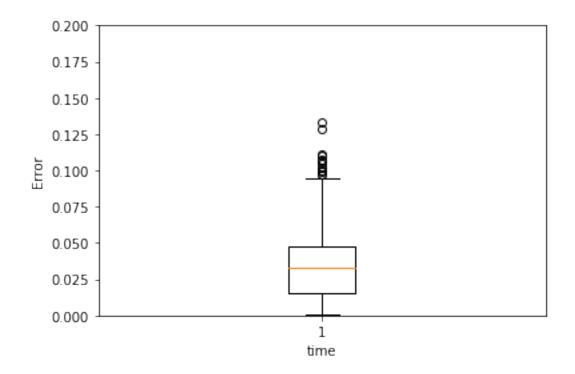


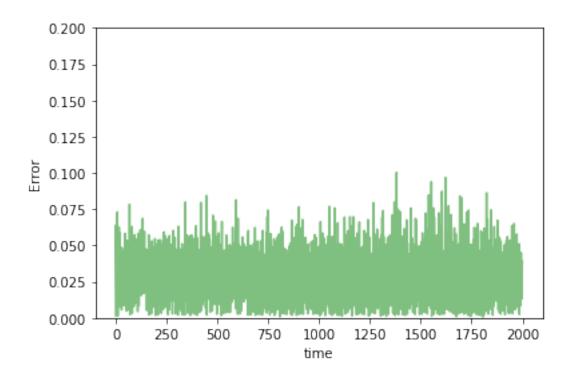


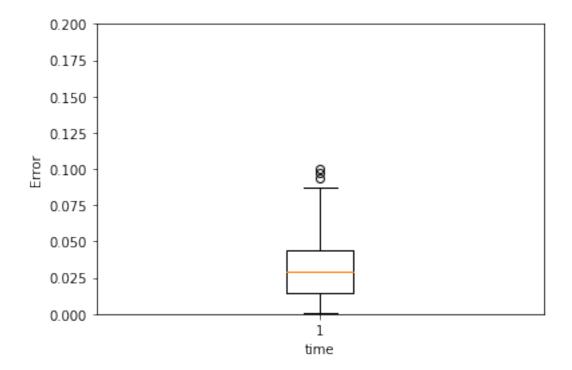


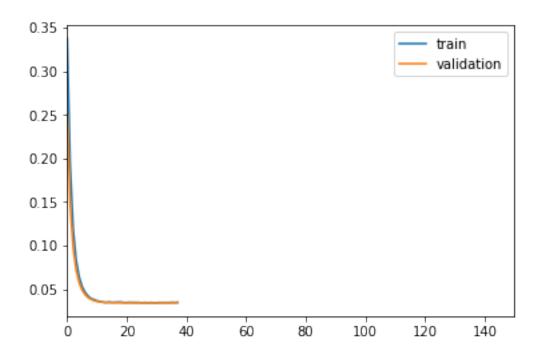


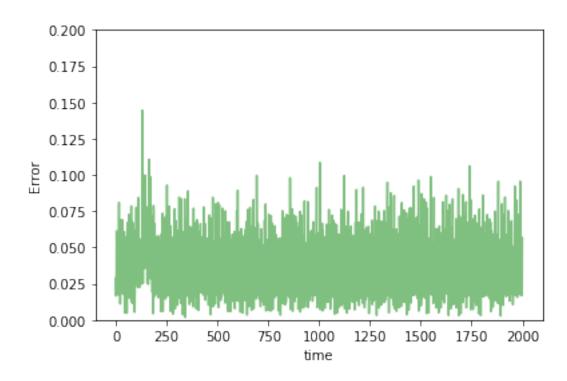


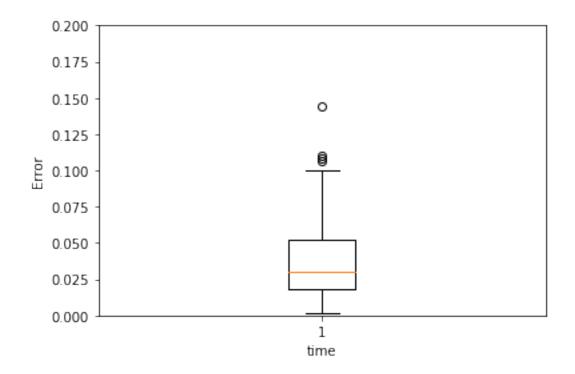


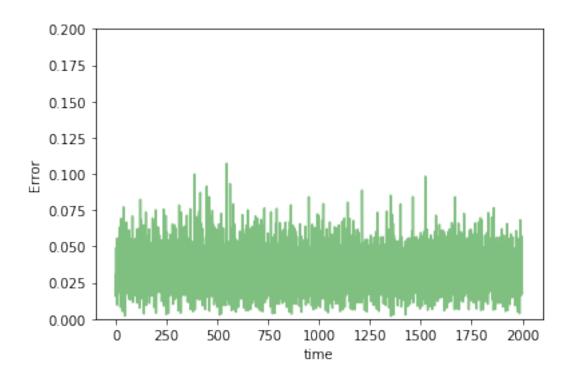


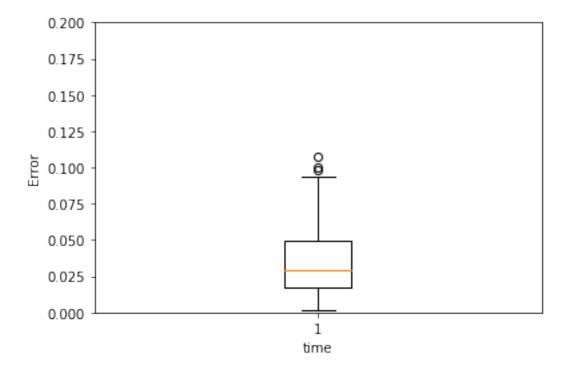












2.1.8 RNN with 4 GRU layers dim compression.

