Prediction of the effect of impeller trimming on centrifugal pump performance using AI.

Supervisor : Prof. Mohamed farid Khalil

By :

mohammed twheed khater

Seif el eslam Ibrahim hassan

Project overview

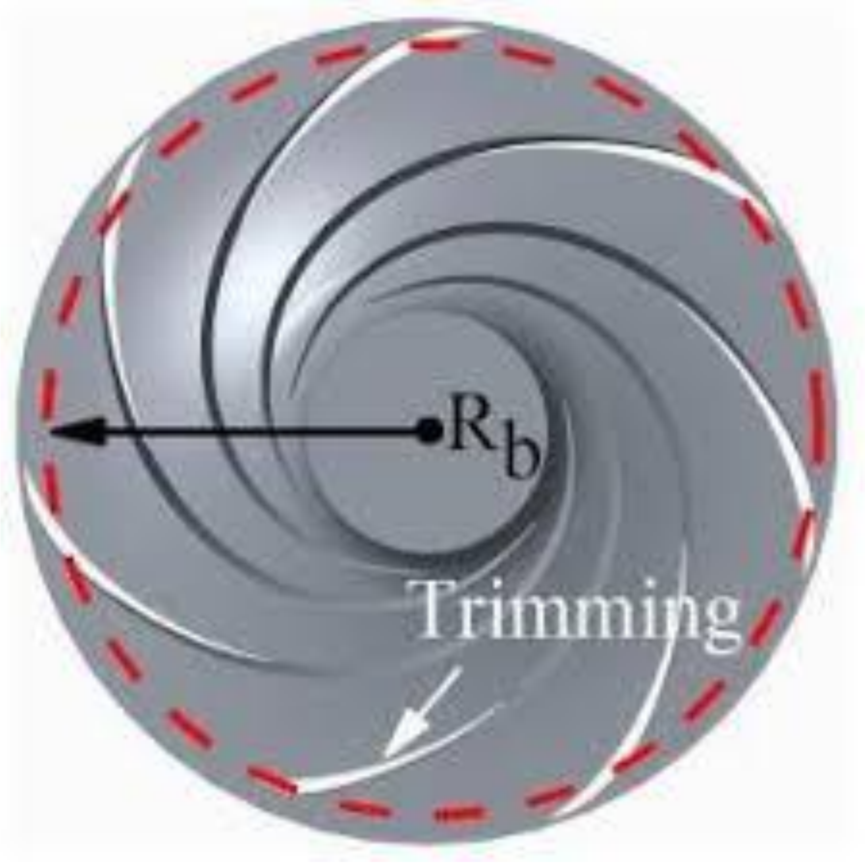
This project aims to optimize and validate the performance of neural networks applied in predicting the characteristics of pump with impeller trimming. The primary focus is on leveraging data-driven approaches to predict key metrics like flow rate, head, and power for different impeller diameters. The project consists of several MATLAB scripts and data files, which together facilitate the preparation, training, and optimization of neural networks for this purpose.

Main topics

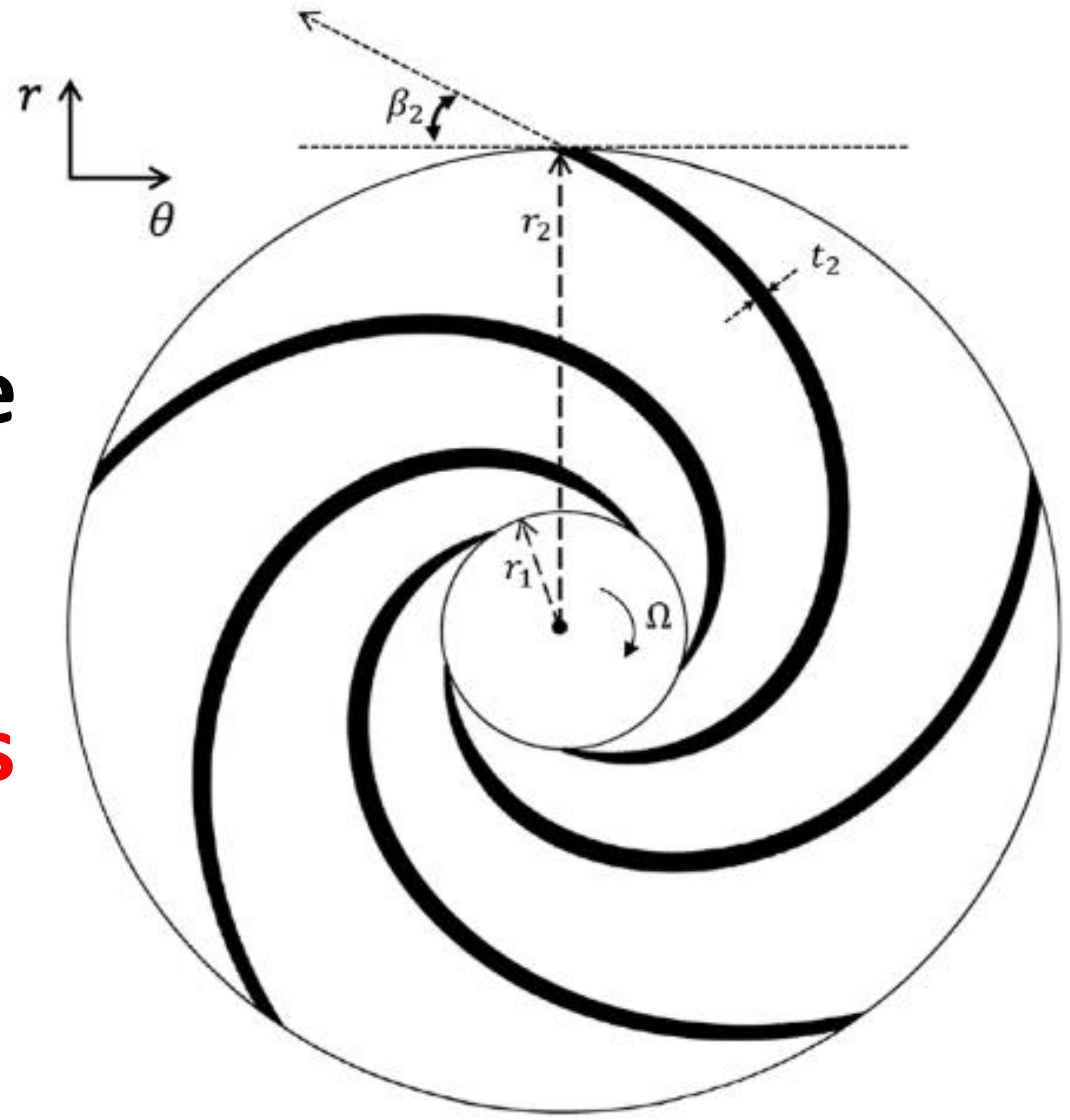
- What is trimming in centrifugal pump impeller ?
- Why trimming ?
- Energy impact
- Data Extraction or digitization
- Neural networks
- Genetic algorithm

What is trimming in centrifugal pump impeller

- It's a process of impeller diameter reduction



Trimming is not just limited to reducing the <u>outer radius r2</u> but could also reduce the <u>inner radius r1</u> but this is not considered here



(b)

Why Trimming?

- **Performance Optimization:**
  - Adjusting pump capacity to meet specific system requirements.
  - Reducing excess flow and pressure.
- **Energy Efficiency:**
  - Lowering energy consumption and corresponding air pollution.
  - Reducing operational costs.
- **Cost Savings:**
  - Extending the lifespan of the pump.
  - Decreasing maintenance and repair costs.

When to Consider Impeller Trimming

- System design changes.
- Over-sizing of the pump required for expected piping system with incomplete or precise technical information .
- Energy audit recommendations.

Environmental Impact of Trimming

Each 1 kW.hr energy saved at pump delivery side corresponds to 6 kW.hr reduction at power station input side.

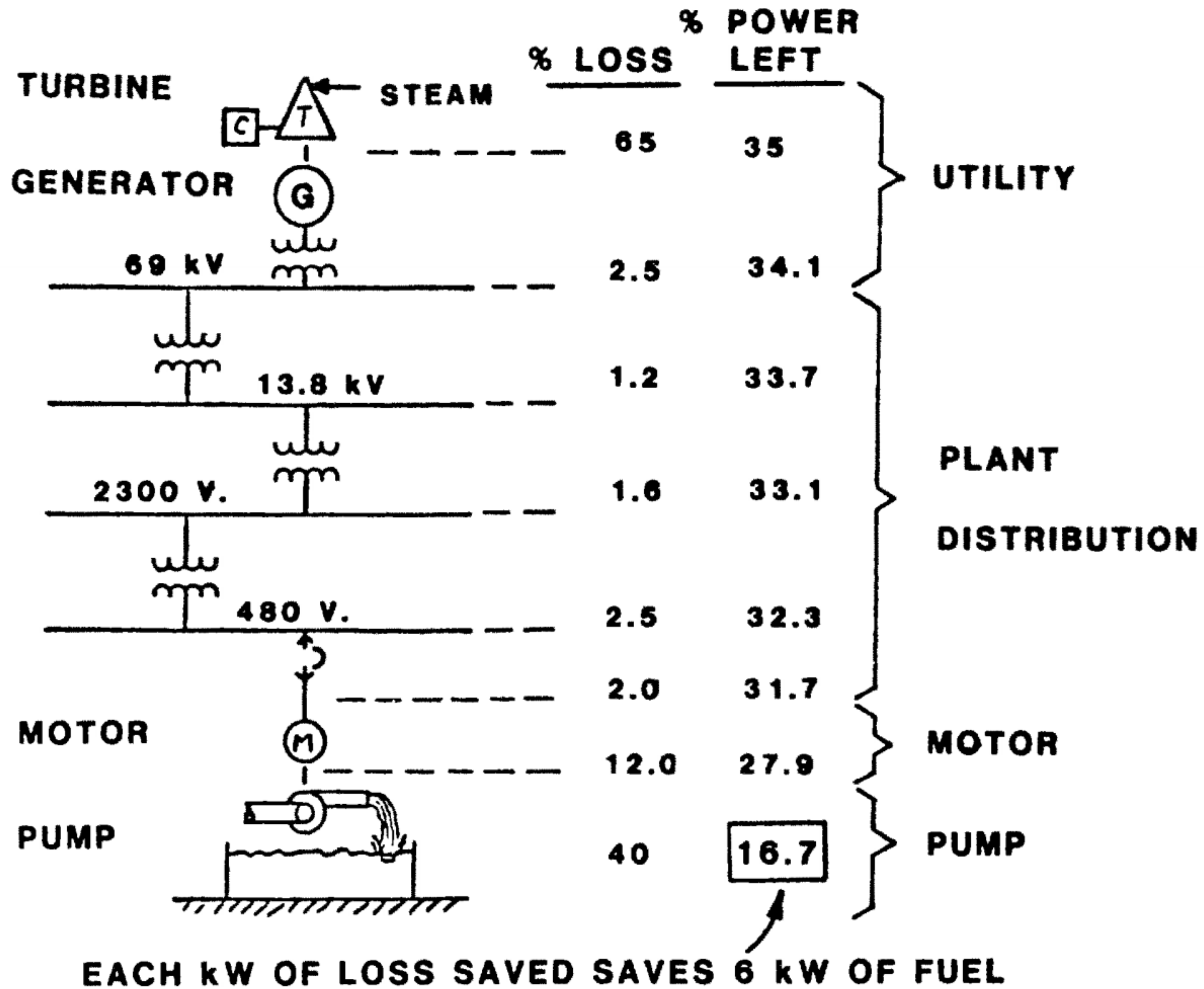


Fig. 2. Flow of energy and its typical losses through power system.

Why ANN application with impeller trimming;

1. Several formulaeare available for predicting the influence of the impeller trimming on the performance centrifugal pumps based on similarity , without confidence, like: [Stepanoff , 1957 ….Weme, 2018],for example;

$$\text{geometrical scaling}: \quad \frac{Q'}{Q} = \left(\frac{D_2'}{D_2}\right)^3 \quad \frac{H'}{H} = \left(\frac{D_2'}{D_2}\right)^2$$

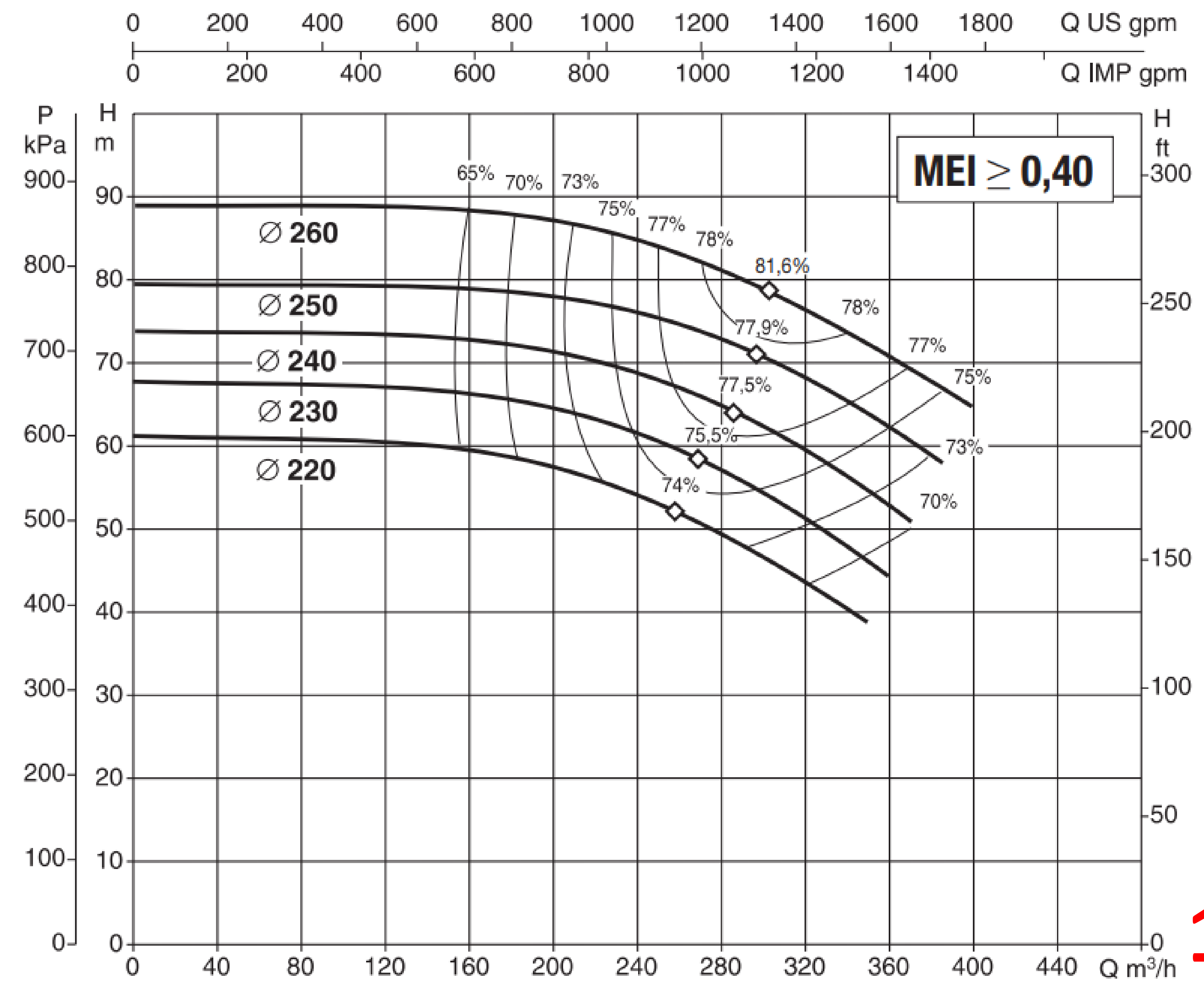
2. CFD methods are available but need comparison with experimental results (more expensive , time consuming and need computing stations)
3. Impeller trimming is generally performed in small steps because of the uncertainty in predicting the effect of trimming on the hydraulic performance.
4. ANN methods depend on the available pump manufacturer data ,with less time and cost.

Although the ANN is used as Black Box we are aiming to modified the internal structure to get ANN of more Accurate and precise predictions, How?

Q-H pump curves for different diameters according to literature the *manufacturer only produces pump with largest diameter*, then trim it to get the lower diameters *(this much cheaper and efficient)*

ambient temperature: +40°C

= 2900 1/min

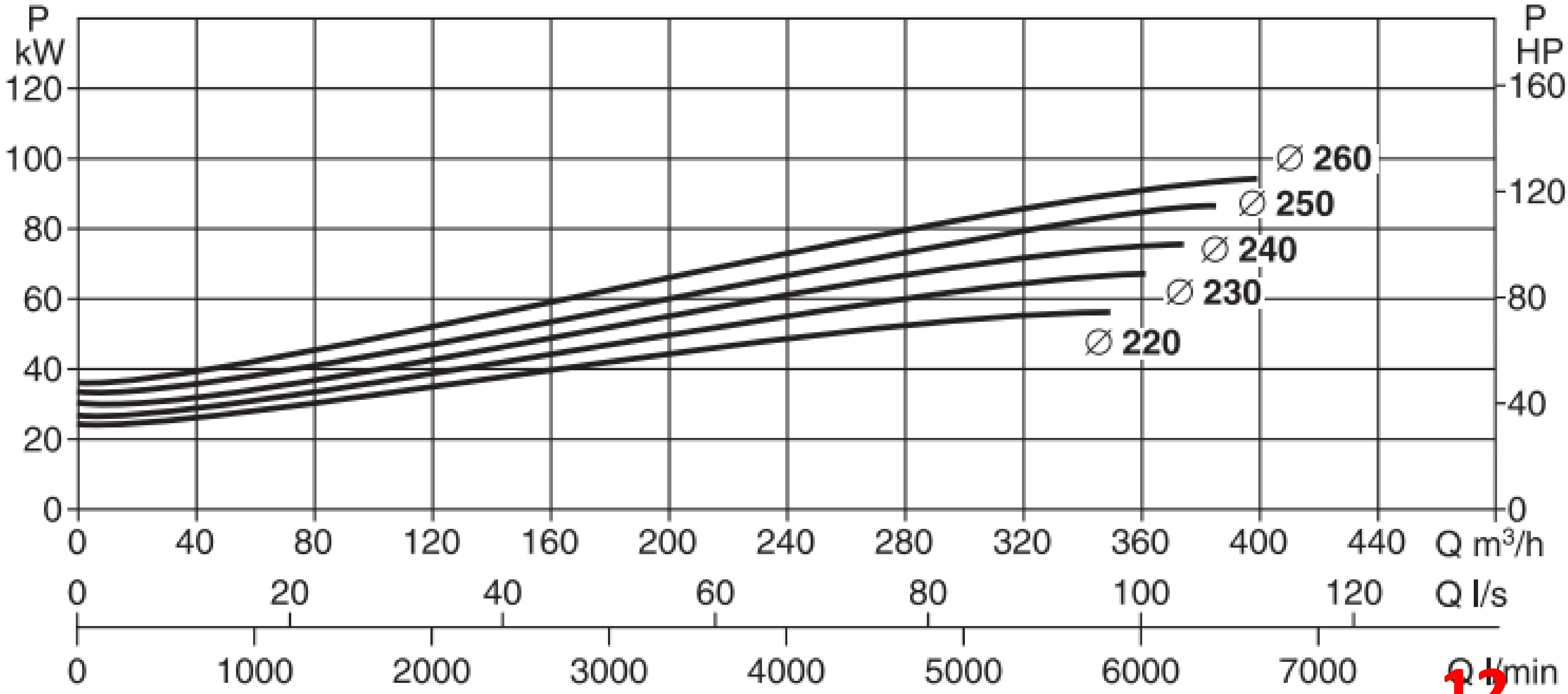


WATER•TECHNOLOGY

Headquarter:

PADOVA ITALY

Q-P Curves

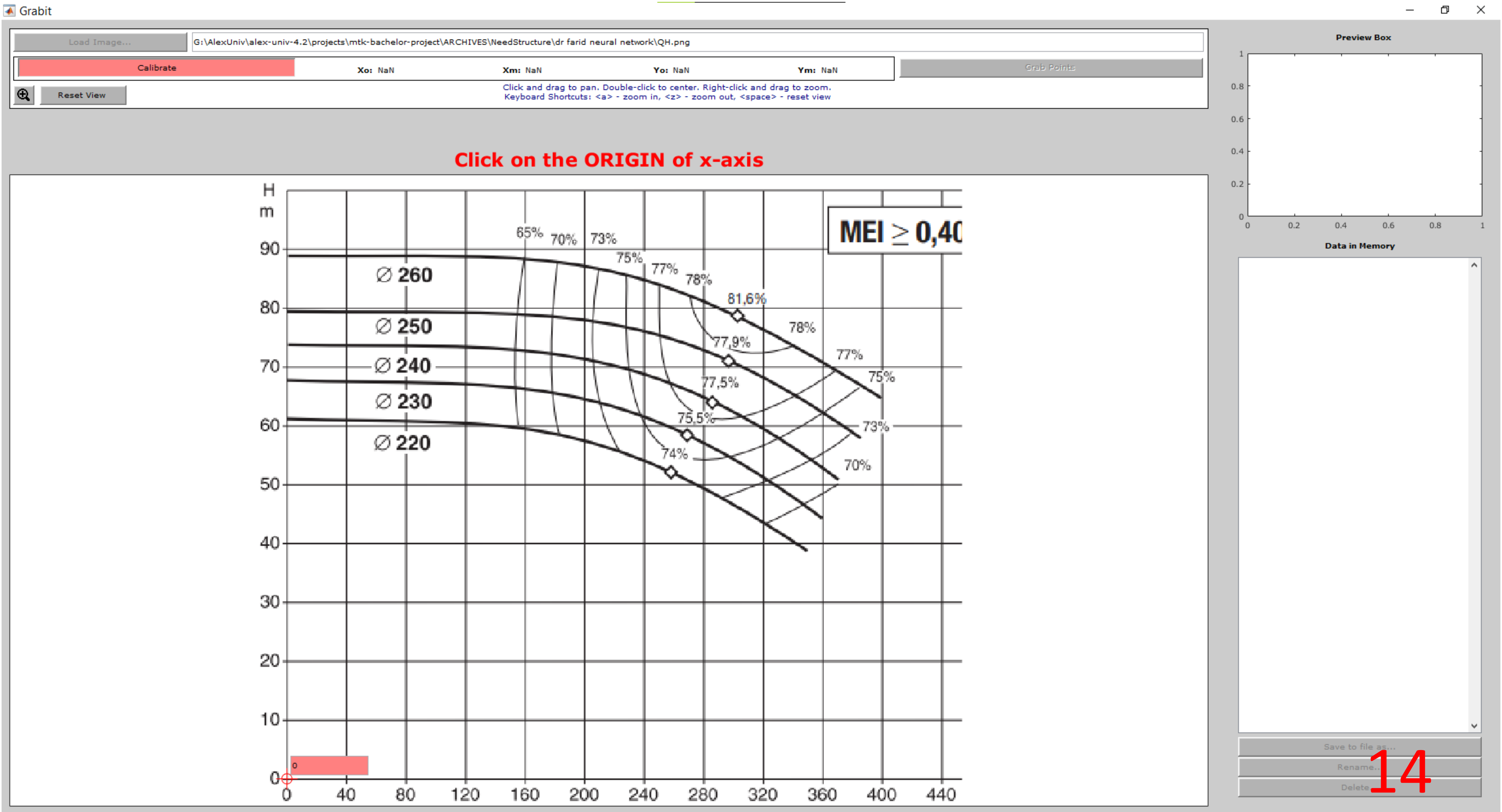


P
kW
120
100
80
60
40
20
0
P
HP
160
120
80
40
0
∅ 260
∅ 250
∅ 240
∅ 230
∅ 220
0
40
80
120
160
200
240
280
320
360
400
440
Q m³/h
0
20
40
60
80
100
120
Q l/s
0
1000
2000
3000
4000
5000
6000
7000
Q l/min

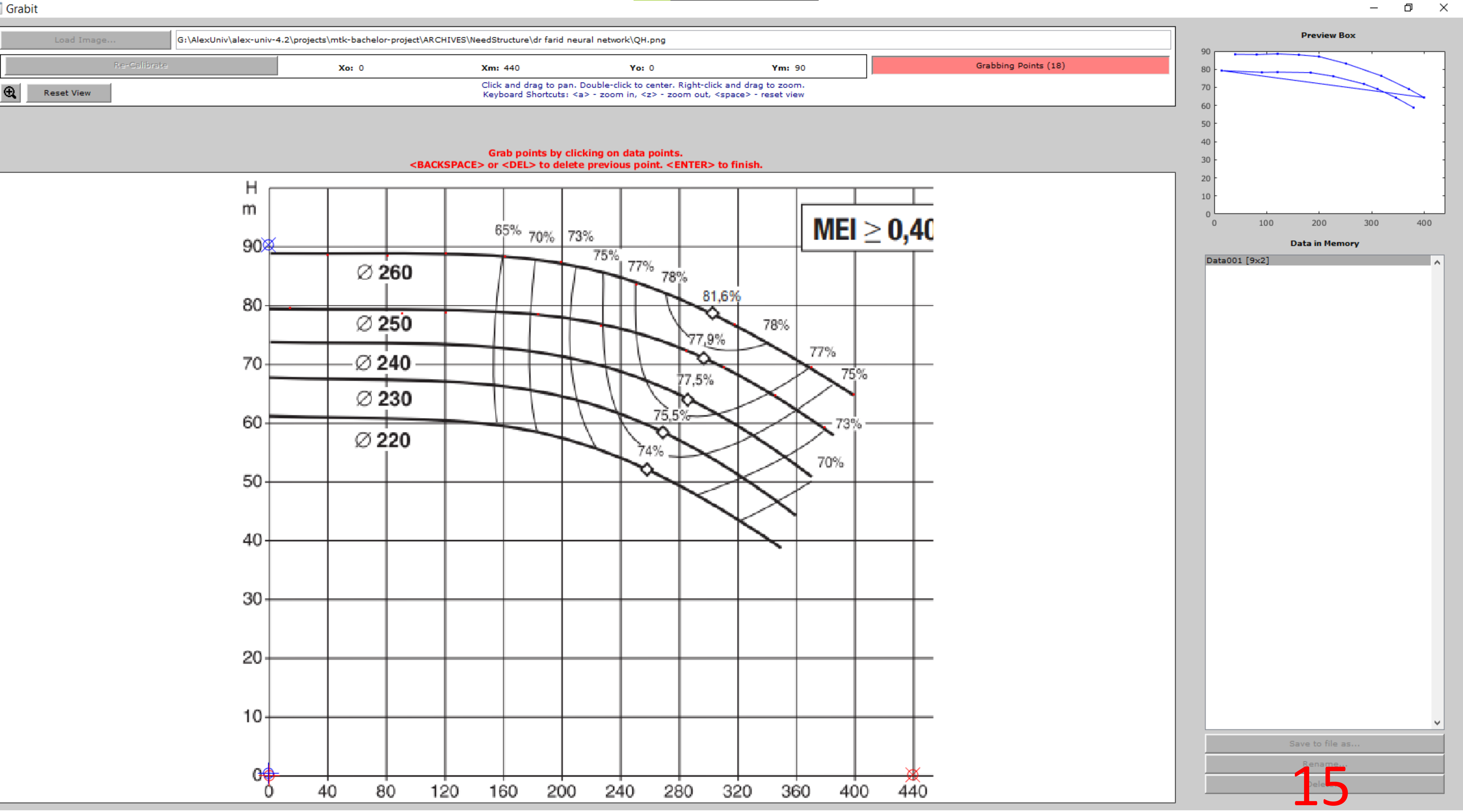
Digitization of Manufacturer Curves

- Extract data from catalogue to computable dataset in csv format to be imported then in MATLAB.
- Using a MATLAB tool found freely online called **garbit**
- **Or using WebPlotDigitizer** (could be used online or downloaded locally it has more features than **grabit**).

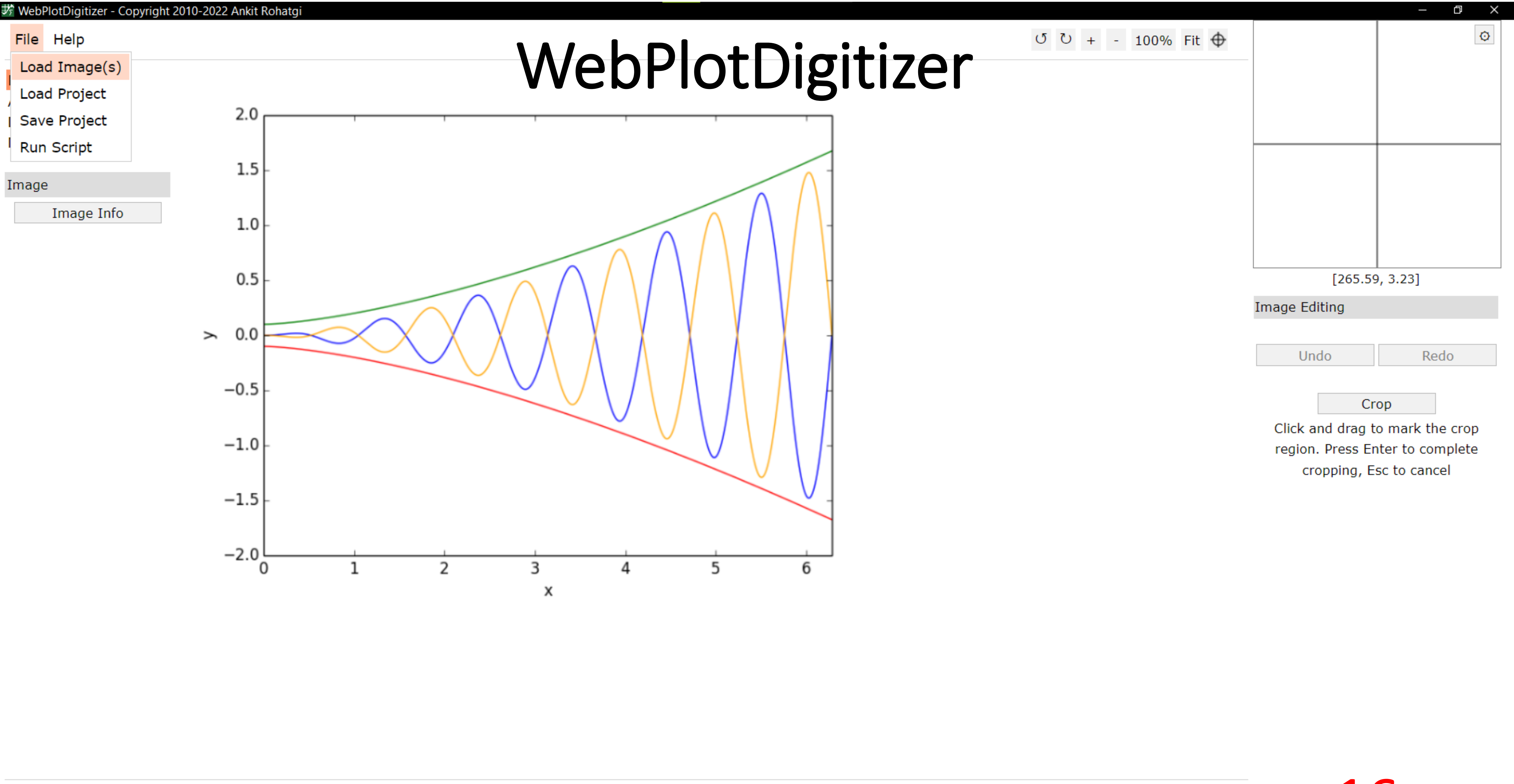
grabit

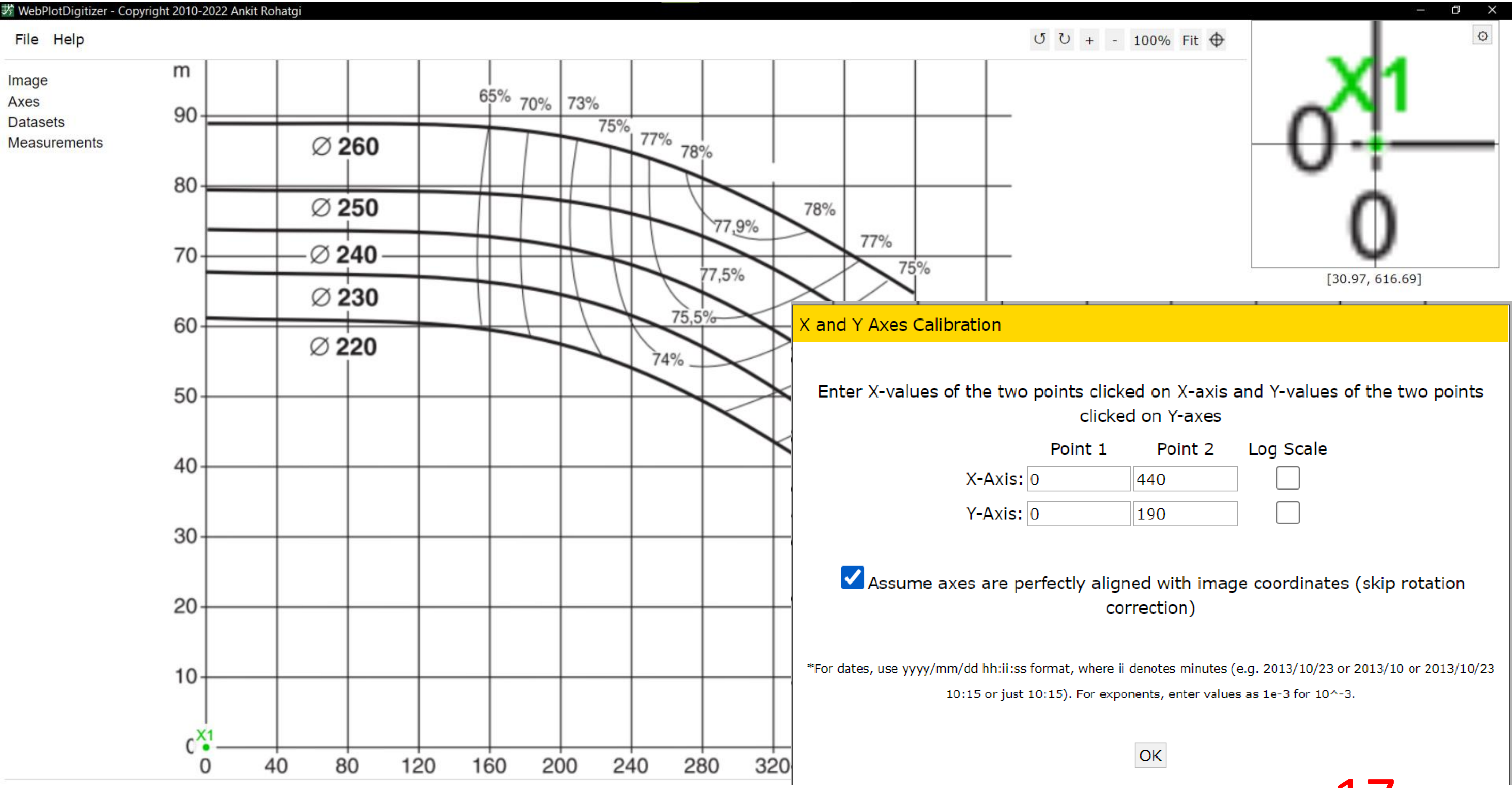


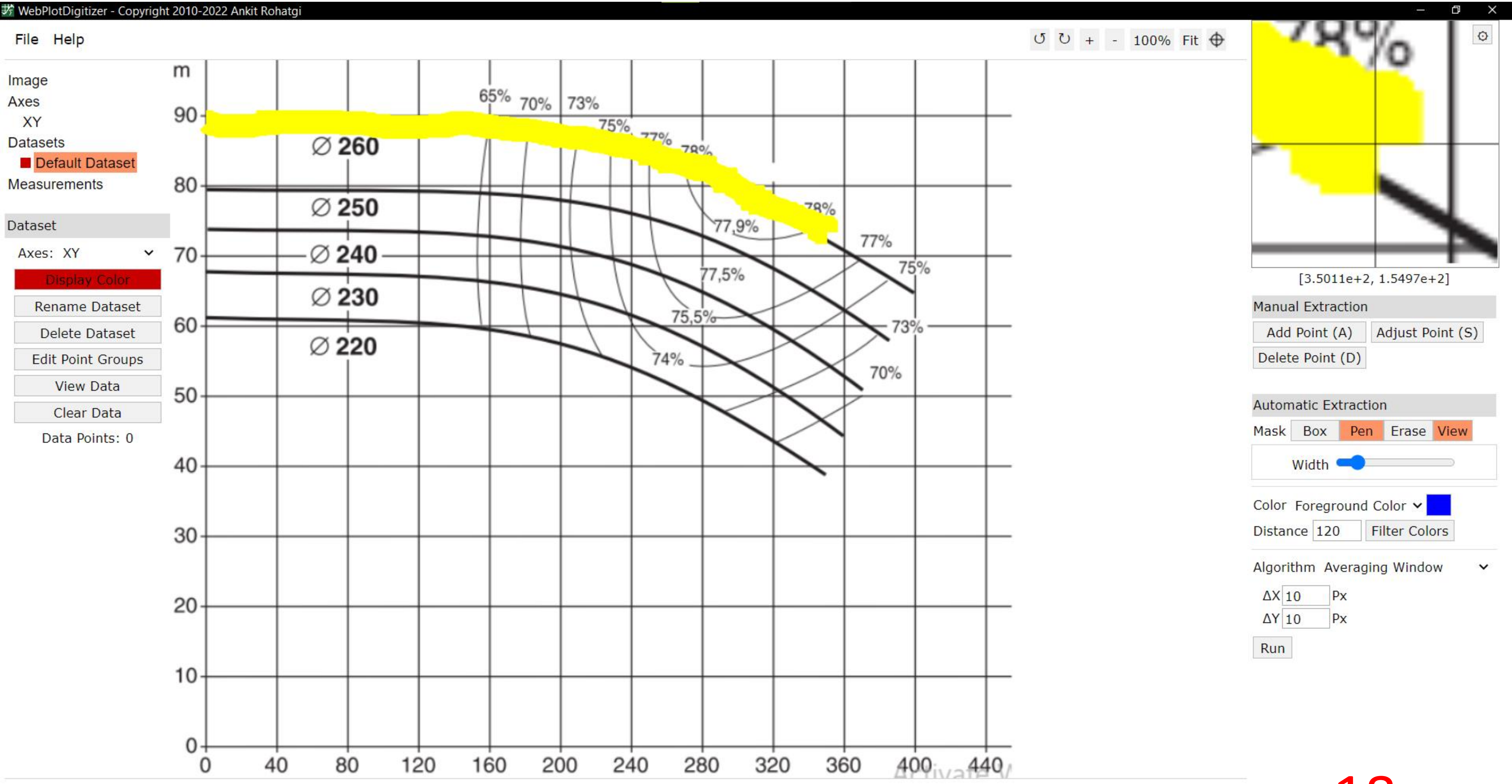
Grabit
Load Image...
G:\AlexUniv\alex-univ-4.2\projects\mtk-bachelor-project\ARCHIVES\NeedStructure\dr farid neural network\QH.png
Calibrate
Xo: NaN
Xm: NaN
Yo: NaN
Ym: NaN
Grab Points
Reset View
Click and drag to pan. Double-click to center. Right-click and drag to zoom.
Keyboard Shortcuts: <a> - zoom in, <z> - zoom out, <space> - reset view
Click on the ORIGIN of x-axis
Preview Box
Data in Memory
Save to file as...
Rename...
Delete
H
m
MEI ≥ 0,40
∅ 260
∅ 250
∅ 240
∅ 230
∅ 220
65%
70%
73%
75%
77%
78%
81,6%
78%
77,9%
77%
75%
77,5%
75,5%
73%
74%
70%
0 40 80 120 160 200 240 280 320 360 400 440
0 10 20 30 40 50 60 70 80 90


Grabit
Load Image...
G:\AlexUniv\alex-univ-4.2\projects\mtk-bachelor-project\ARCHIVES\NeedStructure\dr farid neural network\QH.png
Re-Calibrate
Xo: 0
Xm: 440
Yo: 0
Ym: 90
Grabbing Points (18)
Reset View
Click and drag to pan. Double-click to center. Right-click and drag to zoom.
Keyboard Shortcuts: <a> - zoom in, <z> - zoom out, <space> - reset view
Grab points by clicking on data points.
<BACKSPACE> or <DEL> to delete previous point. <ENTER> to finish.
H
m
MEI ≥ 0,40
∅ 260
∅ 250
∅ 240
∅ 230
∅ 220
65%
70%
73%
75%
77%
78%
81,6%
78%
77,9%
77%
75%
77,5%
75,5%
73%
74%
70%
Preview Box
Data in Memory
Data001 [9x2]
Save to file as...

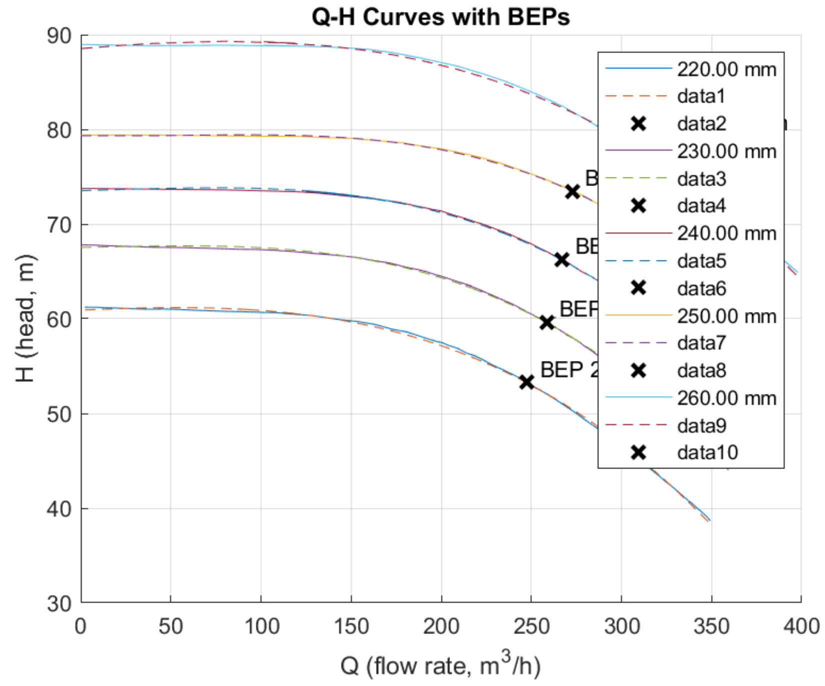
WebPlotDigitizer

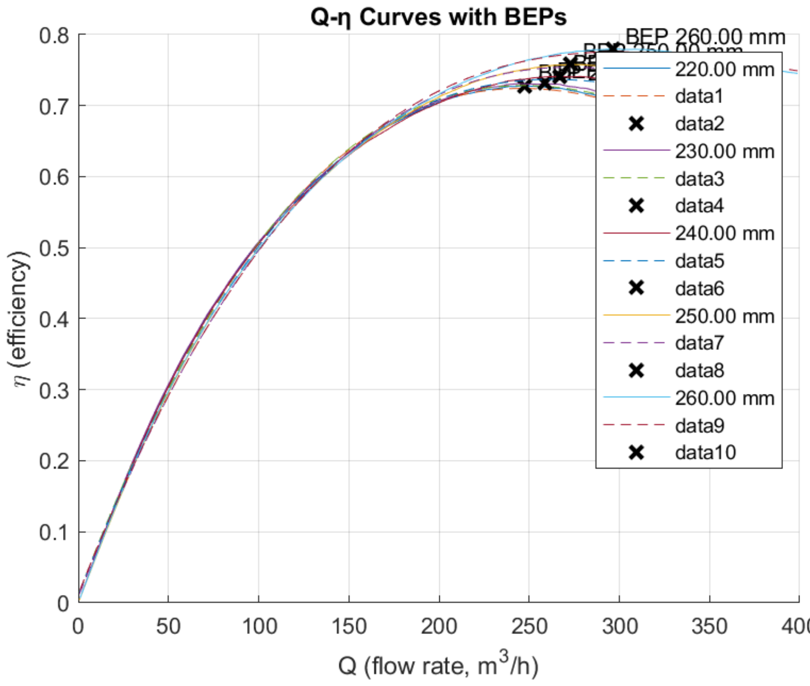


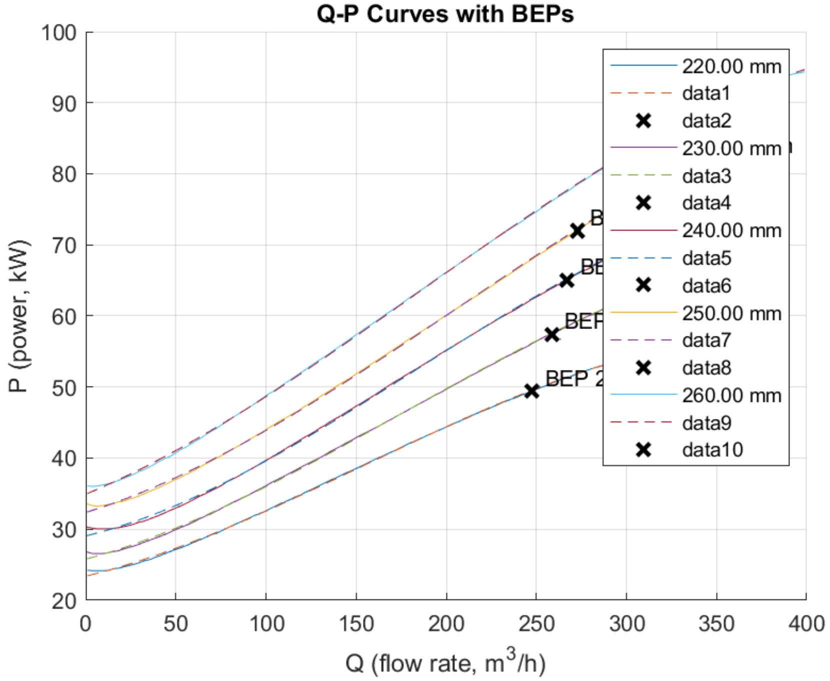

WebPlotDigitizer - Copyright 2010-2022 Ankit Rohatgi
File Help
100% Fit
Image
Axes
Datasets
Measurements
m
90
80
70
60
50
40
30
20
10
0
0 40 80 120 160 200 240 280 320
X1
65% 70% 73%
75% 77% 78%
78%
77,9%
77%
75%
77,5%
75,5%
74%
∅ 260
∅ 250
∅ 240
∅ 230
∅ 220
[30.97, 616.69]
X and Y Axes Calibration
Enter X-values of the two points clicked on X-axis and Y-values of the two points clicked on Y-axes
Point 1
Point 2
Log Scale
X-Axis: 0 440
Y-Axis: 0 190
Assume axes are perfectly aligned with image coordinates (skip rotation correction)
*For dates, use yyyy/mm/dd hh:ii:ss format, where ii denotes minutes (e.g. 2013/10/23 or 2013/10 or 2013/10/23 10:15 or just 10:15). For exponents, enter values as 1e-3 for 10^-3.
OK



WebPlotDigitizer - Copyright 2010-2022 Ankit Rohatgi
File Help
100% Fit
Image
Axes
XY
Datasets
Default Dataset
Measurements
Dataset
Axes: XY
Display Color
Rename Dataset
Delete Dataset
Edit Point Groups
View Data
Clear Data
Data Points: 0
m
90
80
70
60
50
40
30
20
10
0
0 40 80 120 160 200 240 280 320 360 400 440
∅ 260
∅ 250
∅ 240
∅ 230
∅ 220
65%
70%
73%
75%
77%
78%
78%
77,9%
77%
75%
77,5%
75,5%
73%
74%
70%
[3.5011e+2, 1.5497e+2]
Manual Extraction
Add Point (A)
Adjust Point (S)
Delete Point (D)
Automatic Extraction
Mask
Box
Pen
Erase
View
Width
Color Foreground Color
Distance 120
Filter Colors
Algorithm Averaging Window
ΔX 10 Px
ΔY 10 Px
Run



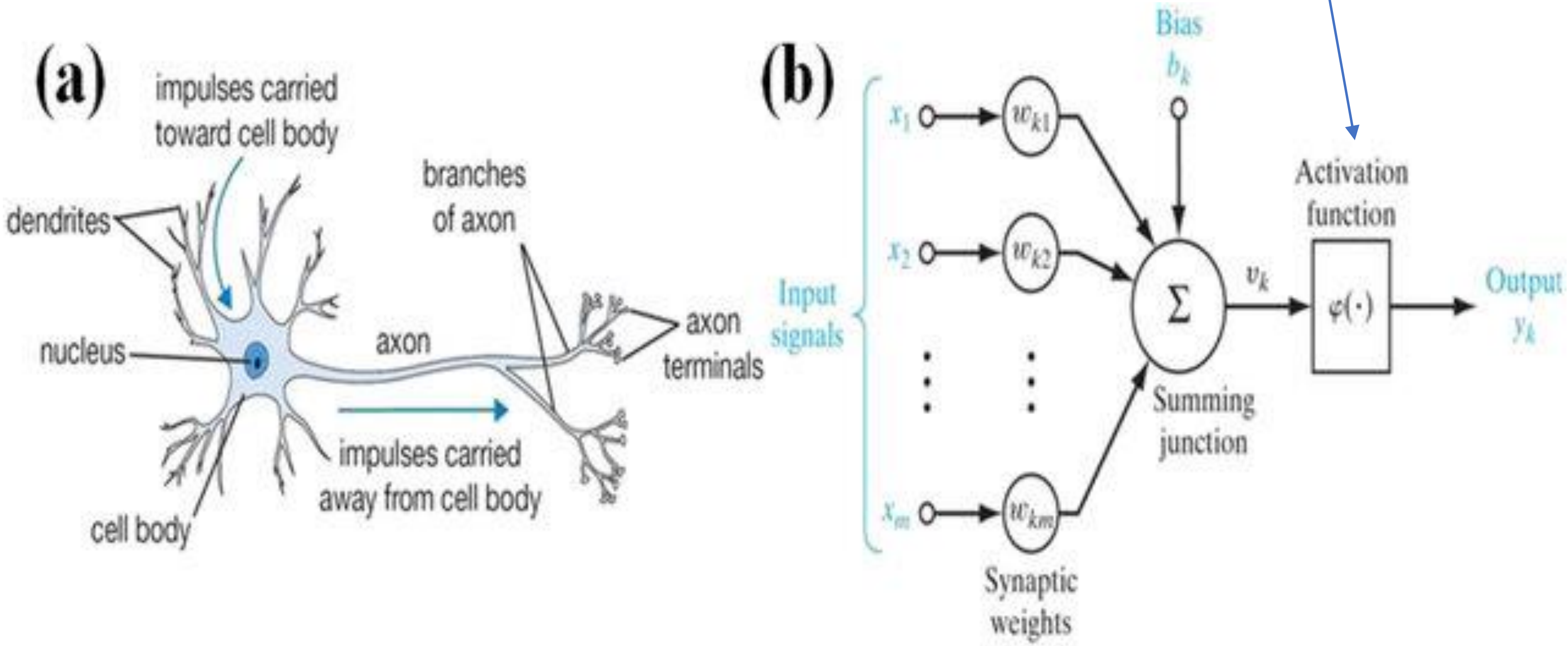




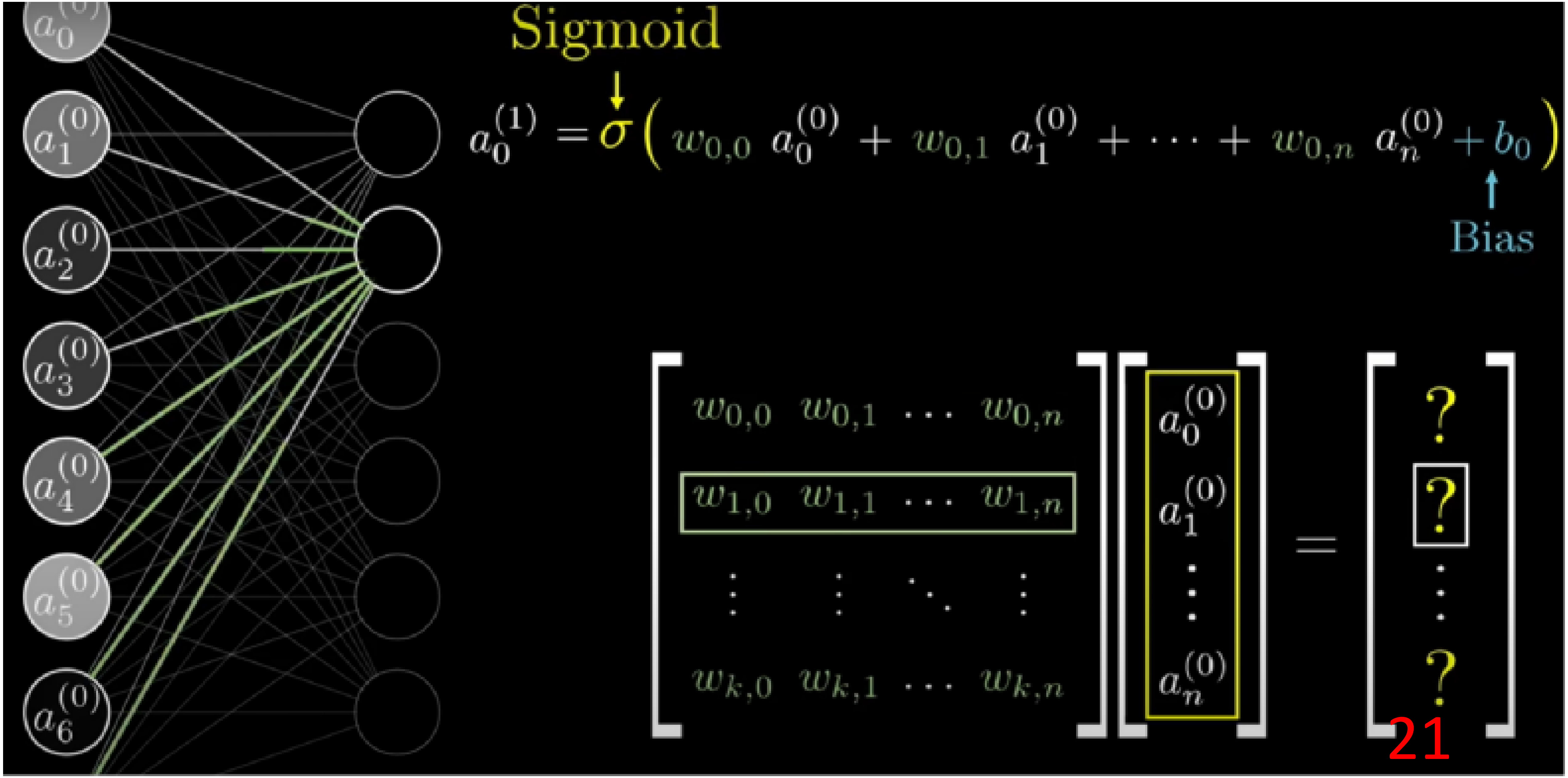
After data extraction here is a plot of it that fitted using MATLAB

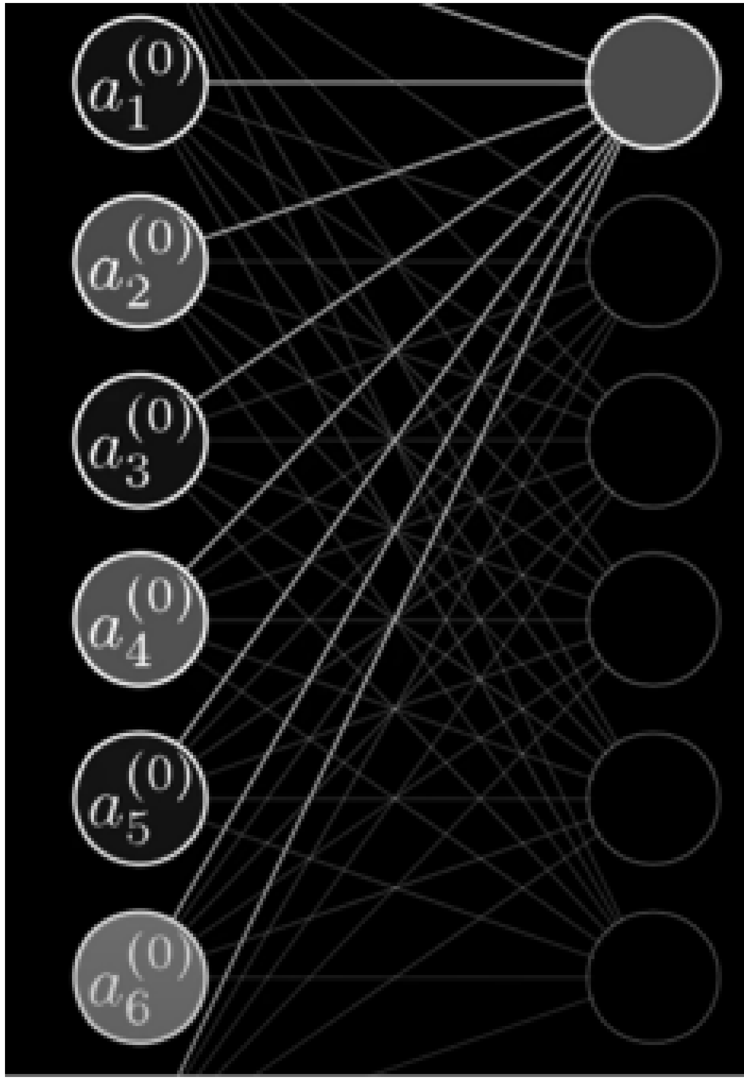
What is a Neural Network

A neural network is a computational model inspired by the way biological neural networks in the human brain process information. It consists **<u>of interconnected layers of nodes (neurons) that work together to recognize patterns, learn from data, and make decisions</u>**. Each node receives input, processes it using **<u>a mathematical function</u>**, and passes the output to the next layer. Neural networks are widely used in various applications, including image and speech recognition, natural language processing, and predictive analytics.



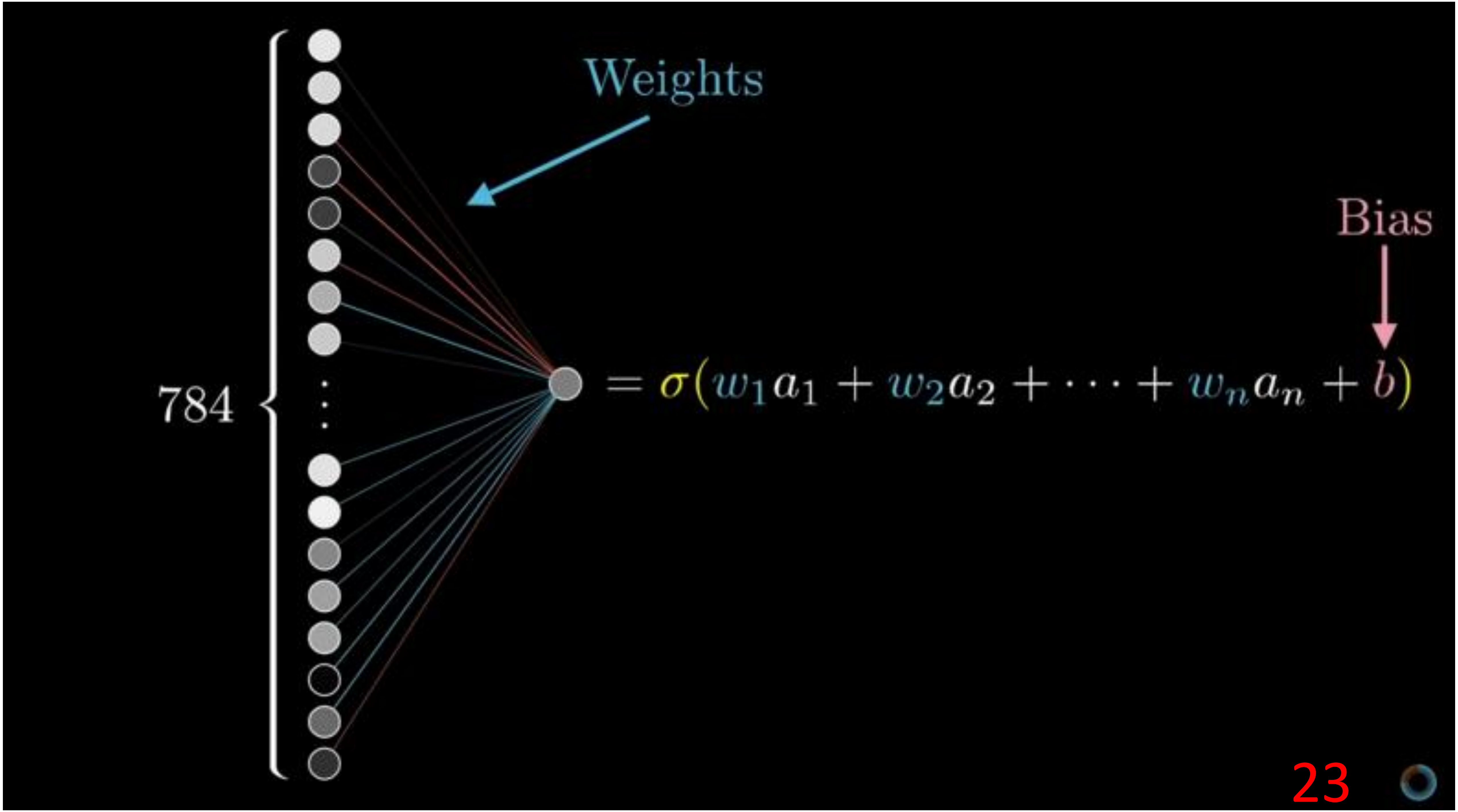
How neural network works

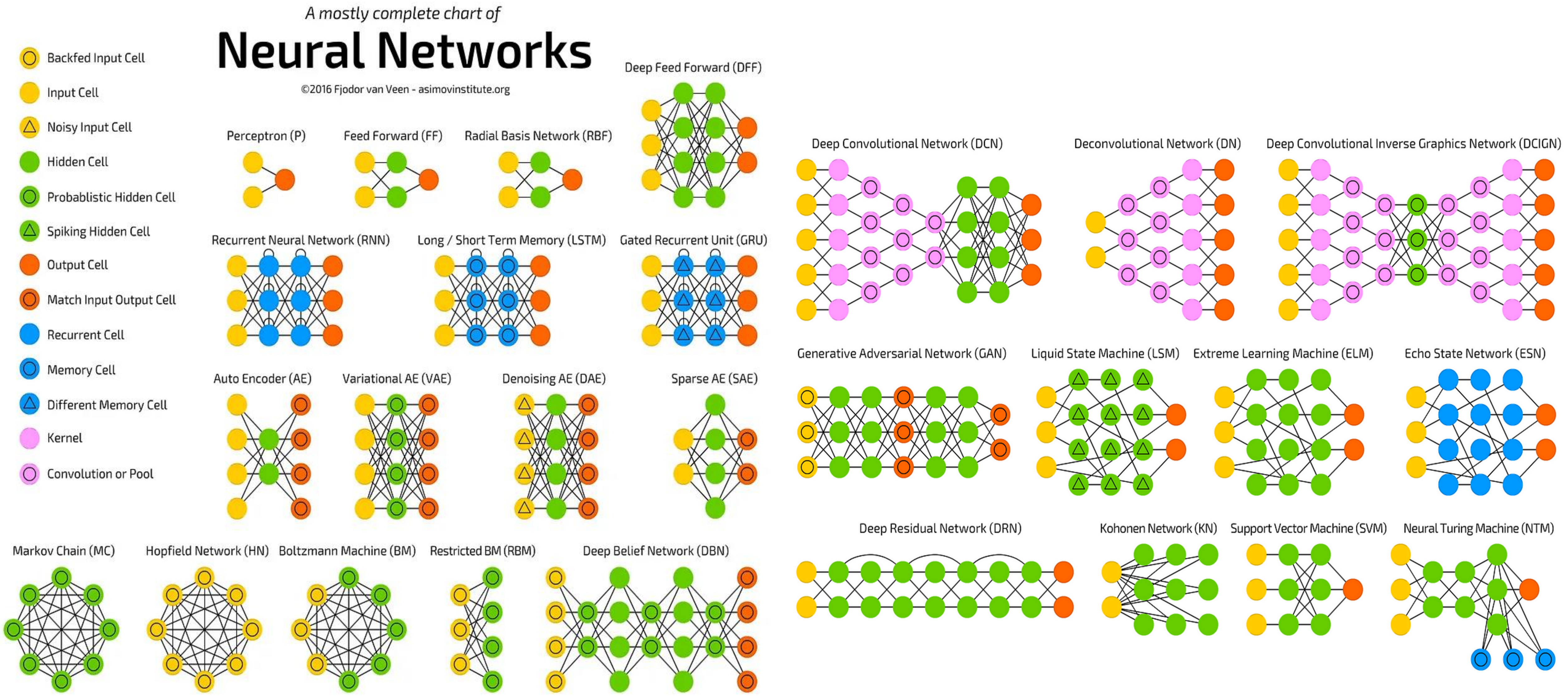




$$\sigma\left(\mathbf{W}\mathbf{a}^{(0)} + \mathbf{b}\right)$$

$$\sigma\left(\begin{bmatrix} w_{0,0} & w_{0,1} & \cdots & w_{0,n} \\ w_{1,0} & w_{1,1} & \cdots & w_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{k,0} & w_{k,1} & \cdots & w_{k,n} \end{bmatrix} \begin{bmatrix} a_0^{(0)} \\ a_1^{(0)} \\ \vdots \\ a_n^{(0)} \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{bmatrix}\right)$$


Weights
Bias
784
$= \sigma(w_1 a_1 + w_2 a_2 + \cdots + w_n a_n + b)$

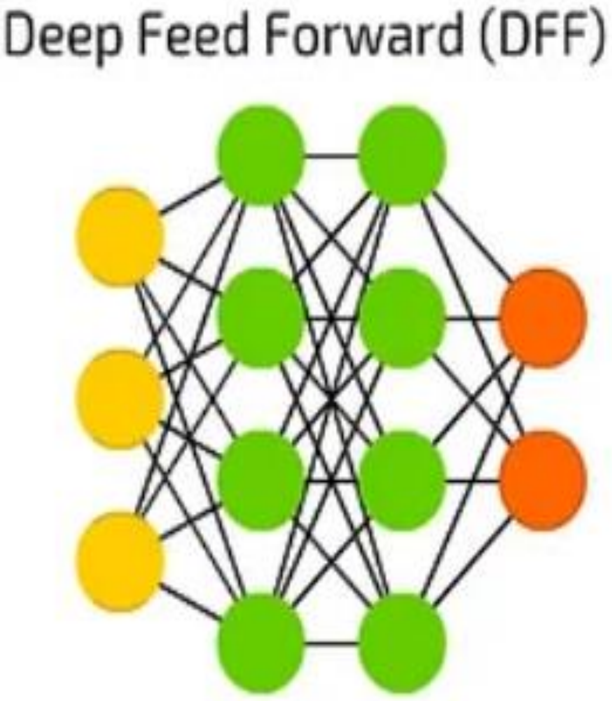

A mostly complete chart of
Neural Networks
©2016 Fjodor van Veen - asimovinstitute.org
Backfed Input Cell
Input Cell
Noisy Input Cell
Hidden Cell
Probablistic Hidden Cell
Spiking Hidden Cell
Output Cell
Match Input Output Cell
Recurrent Cell
Memory Cell
Different Memory Cell
Kernel
Convolution or Pool
Perceptron (P)
Feed Forward (FF)
Radial Basis Network (RBF)
Deep Feed Forward (DFF)
Recurrent Neural Network (RNN)
Long / Short Term Memory (LSTM)
Gated Recurrent Unit (GRU)
Auto Encoder (AE)
Variational AE (VAE)
Denoising AE (DAE)
Sparse AE (SAE)
Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)
Deep Convolutional Network (DCN)
Deconvolutional Network (DN)
Deep Convolutional Inverse Graphics Network (DCIGN)
Generative Adversarial Network (GAN)
Liquid State Machine (LSM)
Extreme Learning Machine (ELM)
Echo State Network (ESN)
Deep Residual Network (DRN)
Kohonen Network (KN)
Support Vector Machine (SVM)
Neural Turing Machine (NTM)

Now this arise the question what architecture should we use for out dataset ??!!

Based on our pump curves dataset we will work only with the deep feed forward NN

But we still need to determine :

1. Number of hidden layers
2. Number of hidden neurons in each layer
3. Activation function
4. Training algorithm



Genetic algorithm

Basic Steps of a Genetic Algorithm

A genetic algorithm (GA) is an optimization technique inspired by the process of natural selection. Here are the basic steps of a genetic algorithm:

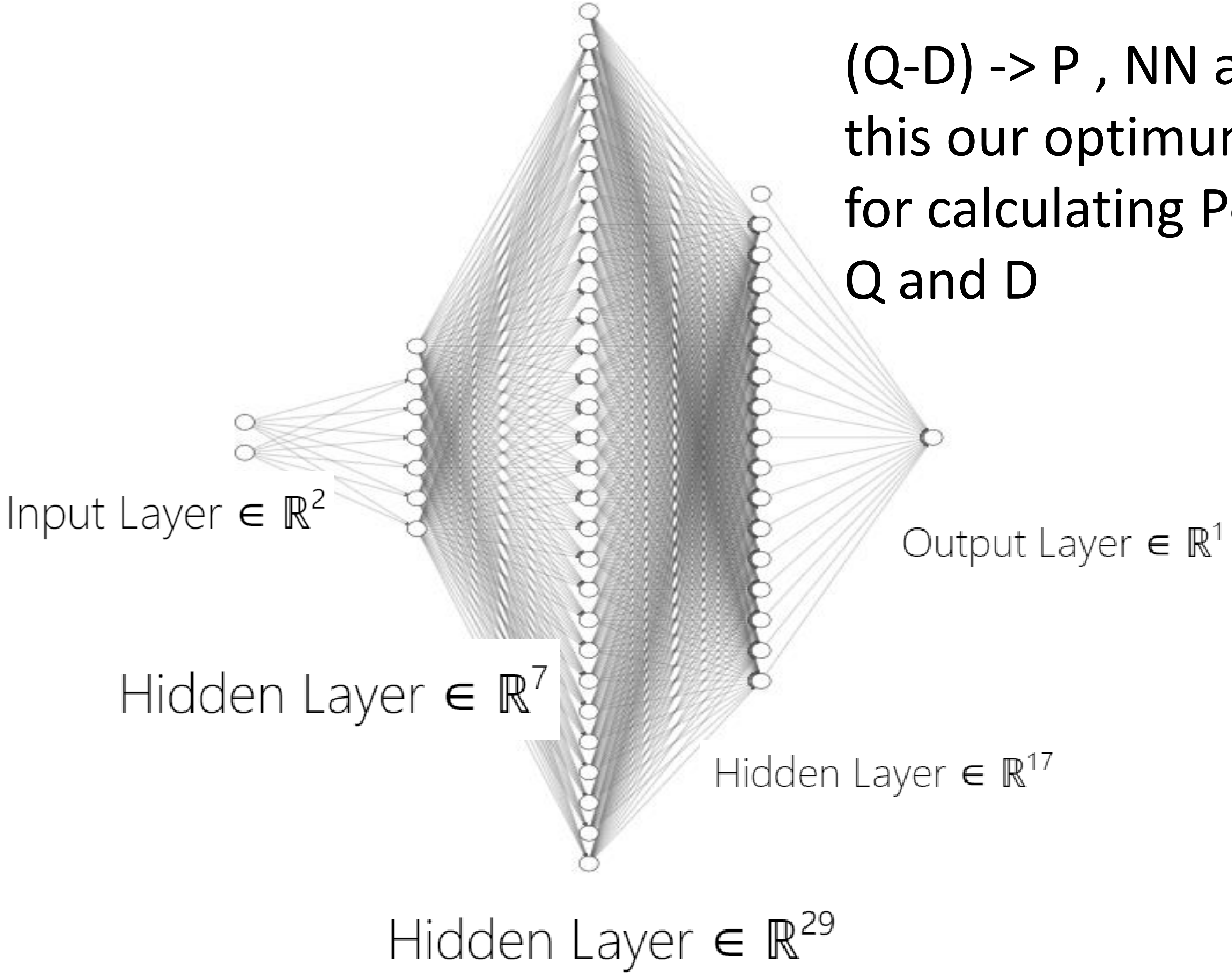
- **Initialization**:
  - Generate an initial population of individuals (solutions) randomly.
  - Each individual is represented by a chromosome (a set of parameters).
- **Evaluation**:
  - Evaluate the fitness of each individual in the population using a fitness function.
- **Selection**:
  - Select individuals for reproduction based on their fitness. Higher fitness individuals have a higher chance of being selected.
- **Crossover (Recombination)**:
  - Combine pairs of parents to produce offspring (children). This is done by swapping parts of the parents' chromosomes.
- **Mutation**:
  - Introduce small random changes to some individuals' chromosomes to maintain genetic diversity within the population.
- **Replacement**:
  - Replace the current population with the new generation of individuals.
- **Termination**:
  - Repeat the evaluation, selection, crossover, and mutation steps until a stopping criterion is met, such as a maximum number of generations or a satisfactory fitness level

Genetic algorithm

Basic Steps of a Genetic Algorithm

- **Initialization:**
- **Evaluation:**
- **Selection: Crossover (Recombination):**
- **Mutation:**
- **Replacement:**
- **Termination:**

(Q-D) -> P , NN architecture this our optimum found NN for calculating Power from Q and D



$$R^2 = \frac{\sum_{m=1}^{M}\left(y_m - \overline{y}\right)^2 - \sum_{m=1}^{M}\left(y_m - \hat{y}_m\right)^2}{\sum_{m=1}^{M}\left(y_i - \overline{y}\right)^2}$$

(Coefficient of Determination)

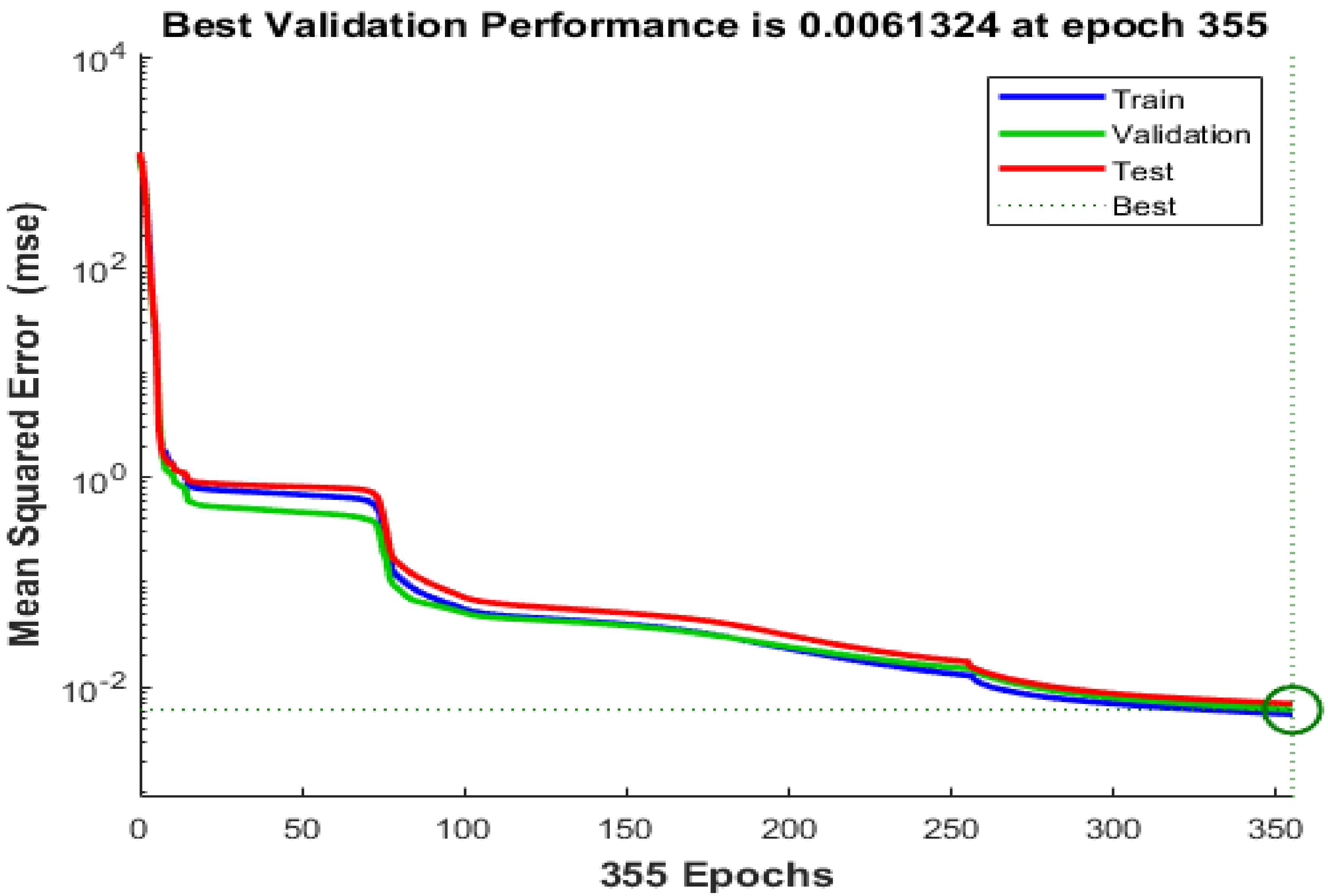
$$MSE = \frac{\sum_{m=1}^{M}\left(y_m - d_m\right)^2}{M}$$

(Mean Squared Error)

It is a measure of the proportion of the variance in the dependent variable (y) that is predictable from the independent variables. It indicates the goodness of fit of a regression model. Where:

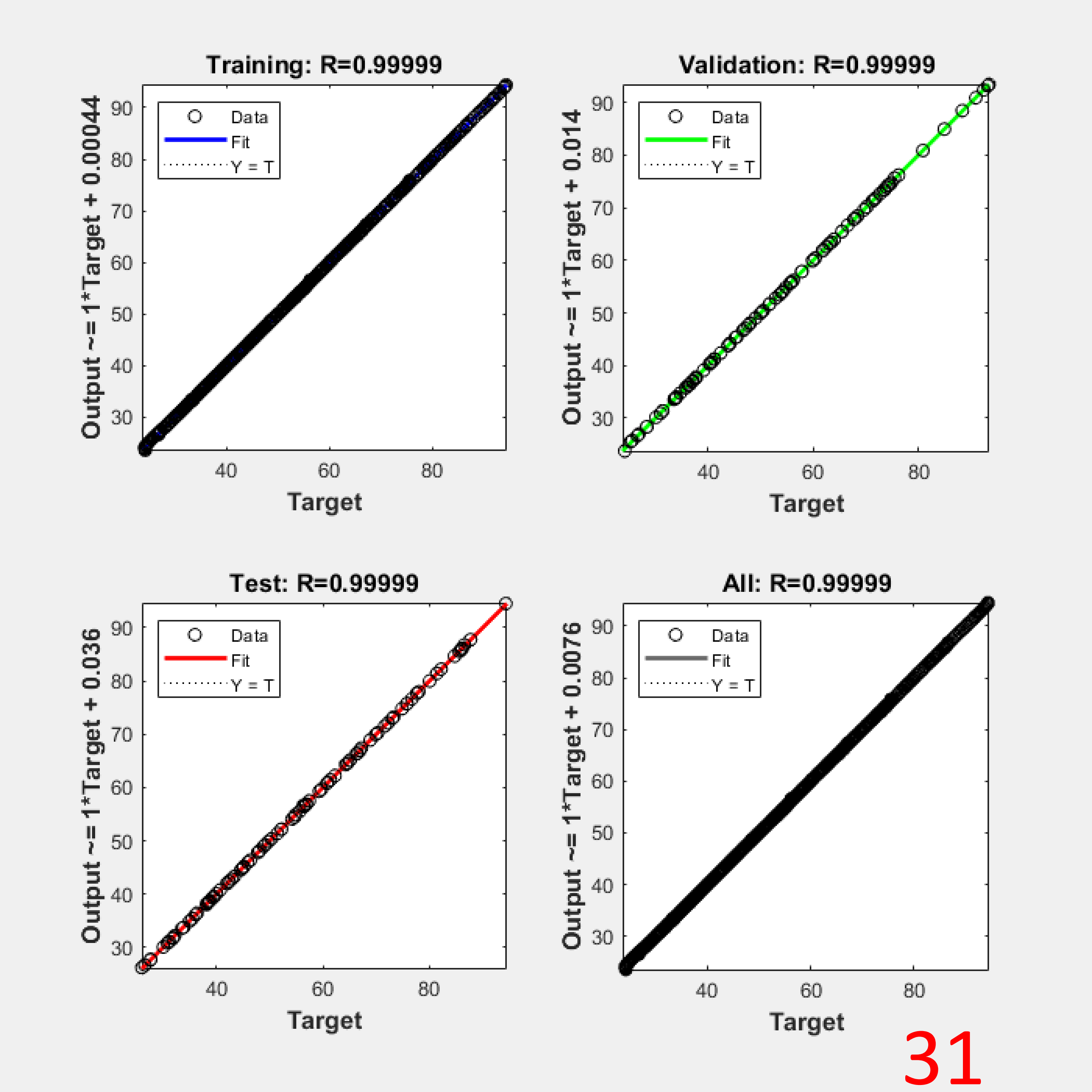
- y_m: Actual value of the dependent variable for the m-th observation.
- $\hat{y}_m$: Predicted value of the dependent variable for the m-th observation.
- $\overline{y}$: Mean of the actual values of the dependent variable.
- M: Number of observations.

(Q-D) -> P , NN performance



(Q-D) -> P , NN Regression

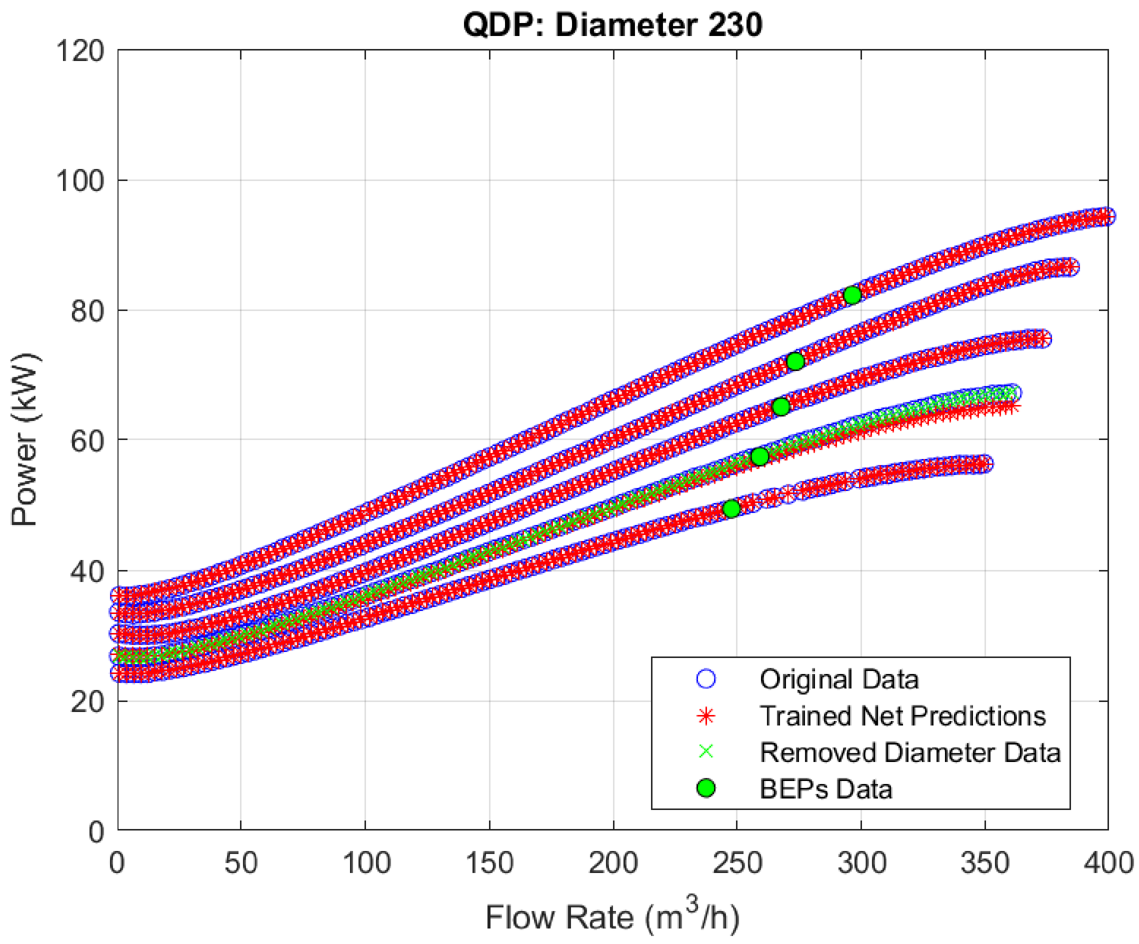
Comparison between predicted and actual values

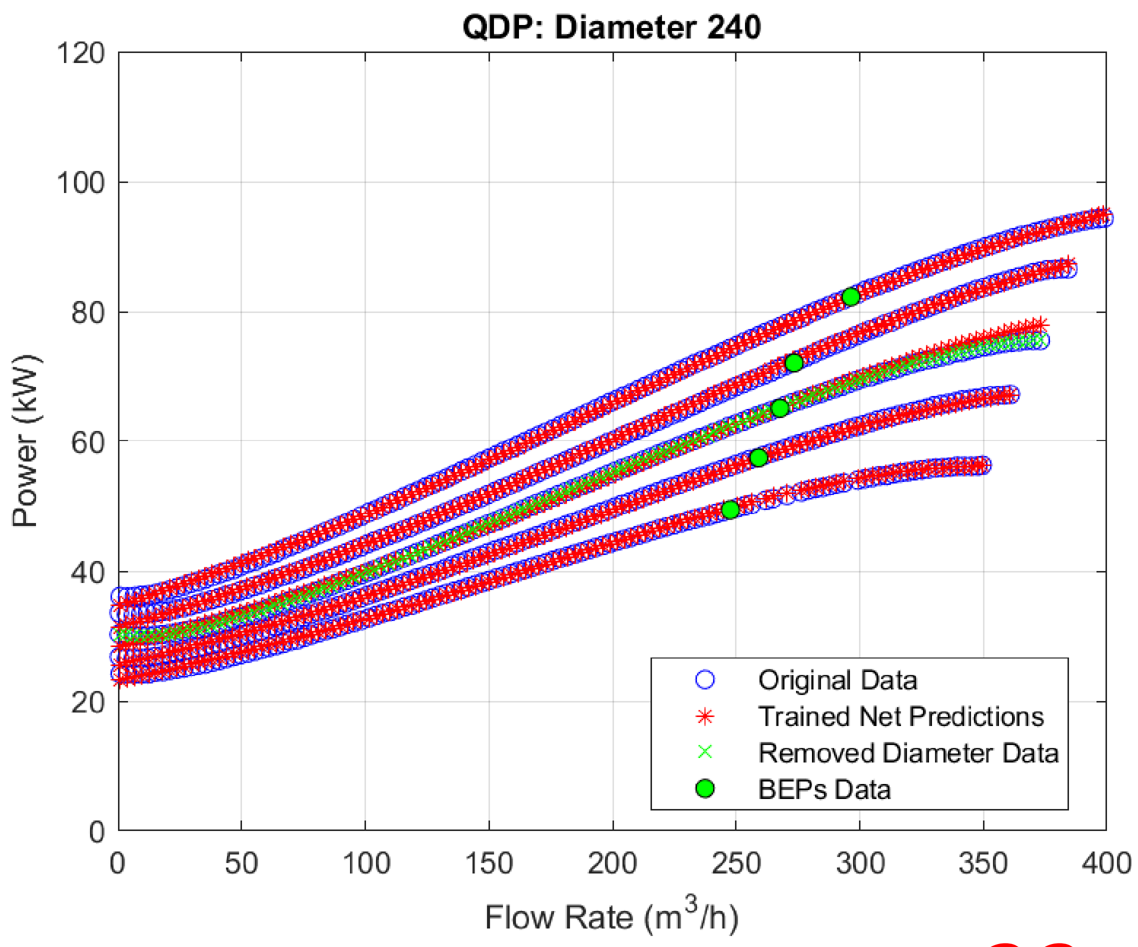


(Q,D) -> P , NN Statistics with complete diameters curves removed from the training dataset

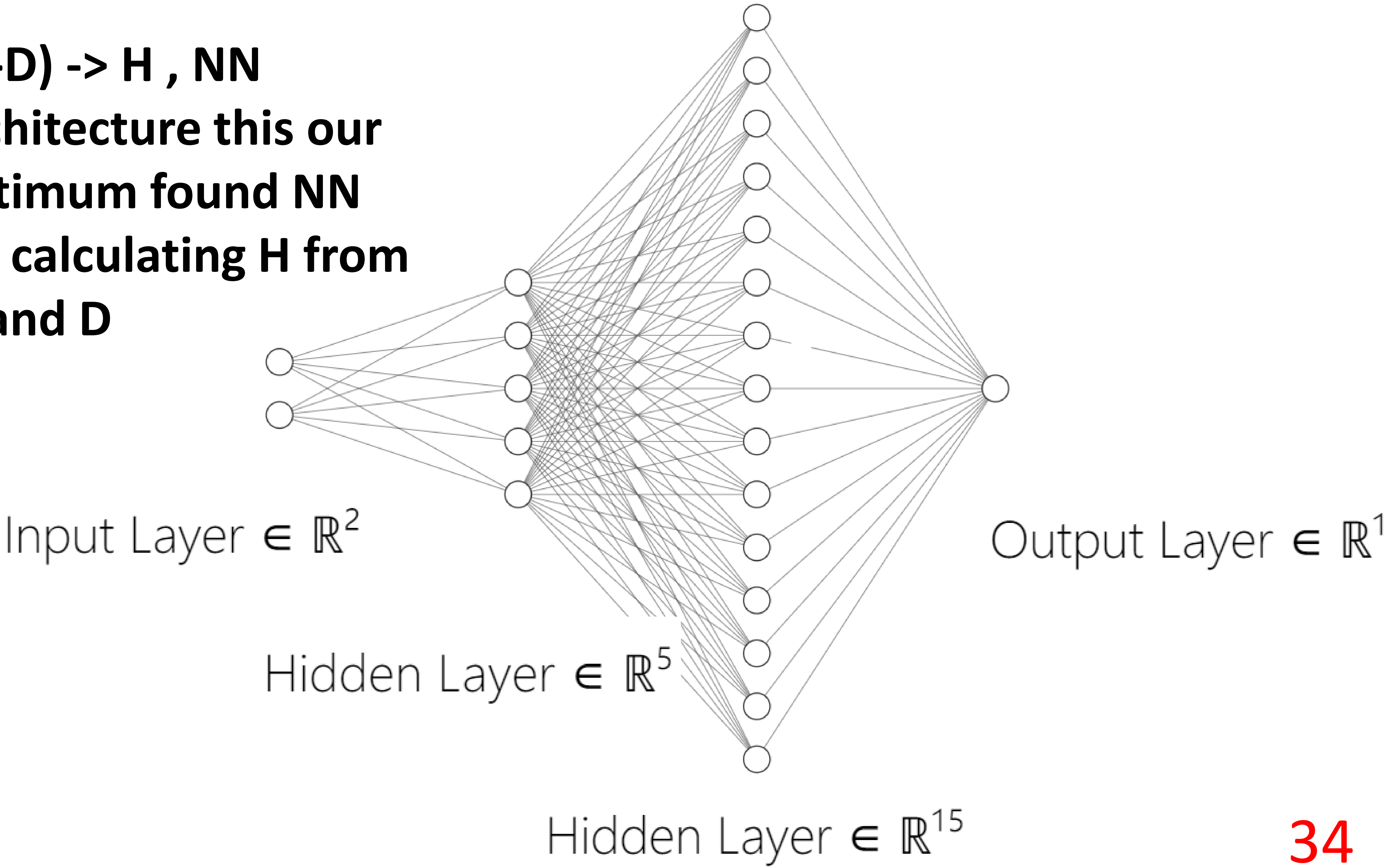
DiameterRemoved	AvgMSE	TrainPerformance	ValPerformance	TestPerformance
NaN	0.006147457	0.005572869	0.006132388	0.007010981
220	0.005916987	0.004854676	0.003499901	0.009904145
230	0.001361164	0.001132582	0.001203953	0.001855751
240	0.084458559	0.101945378	0.06412288	0.078944159
250	1.245301068	1.161826211	1.373474012	1.24072159
260	0.957587821	0.916760554	0.978026414	0.99741805

These numbers represent the error calculated as difference between the actual and predicted values

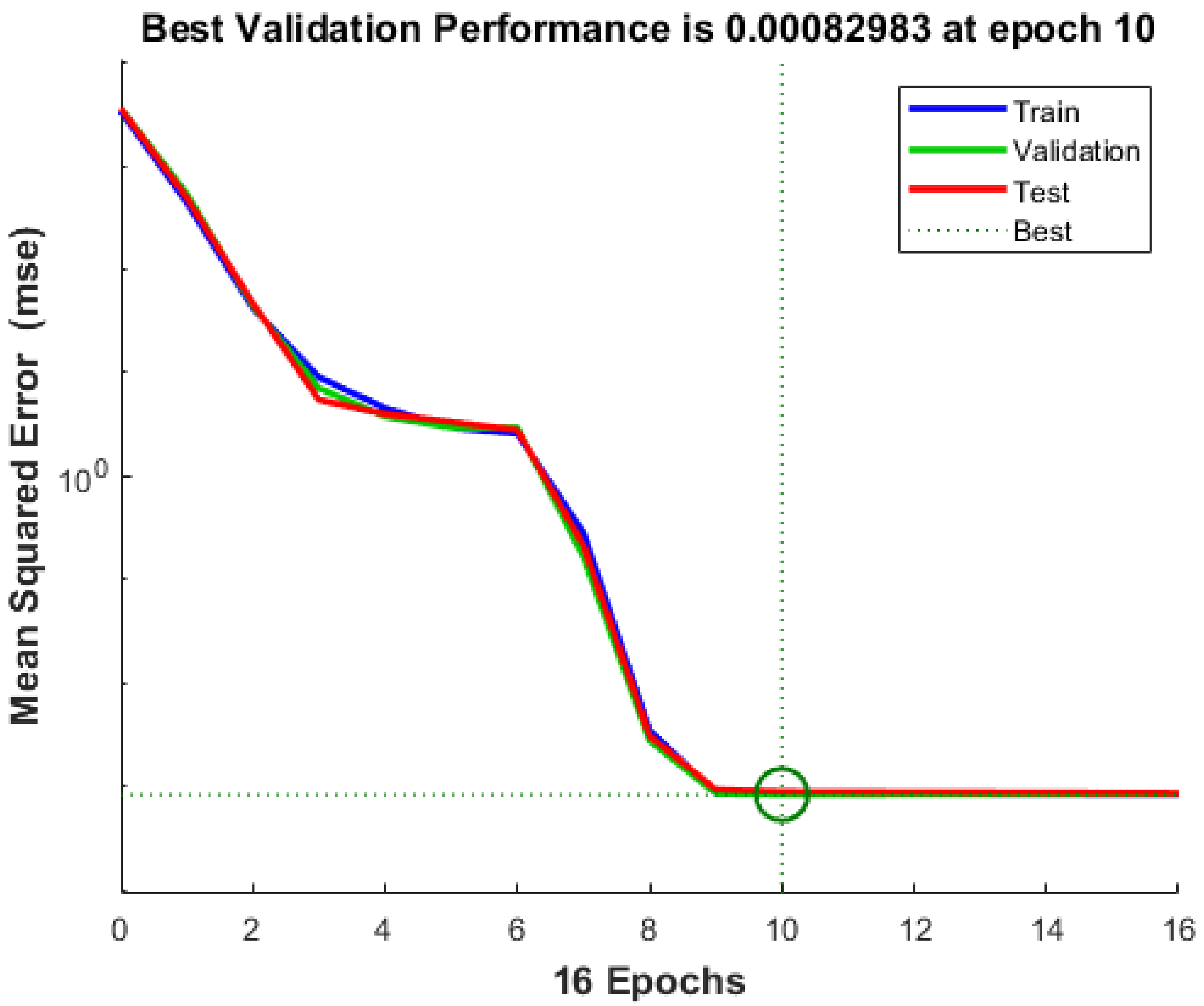

QDP: Diameter 230
Power (kW)
Flow Rate (m^3/h)
0
20
40
60
80
100
120
0
50
100
150
200
250
300
350
400
Original Data
Trained Net Predictions
Removed Diameter Data
BEPs Data


QDP: Diameter 240
Power (kW)
Flow Rate (m^3/h)
0
20
40
60
80
100
120
0
50
100
150
200
250
300
350
400
Original Data
Trained Net Predictions
Removed Diameter Data
BEPs Data

(Q-D) -> H , NN architecture this our optimum found NN for calculating H from Q and D

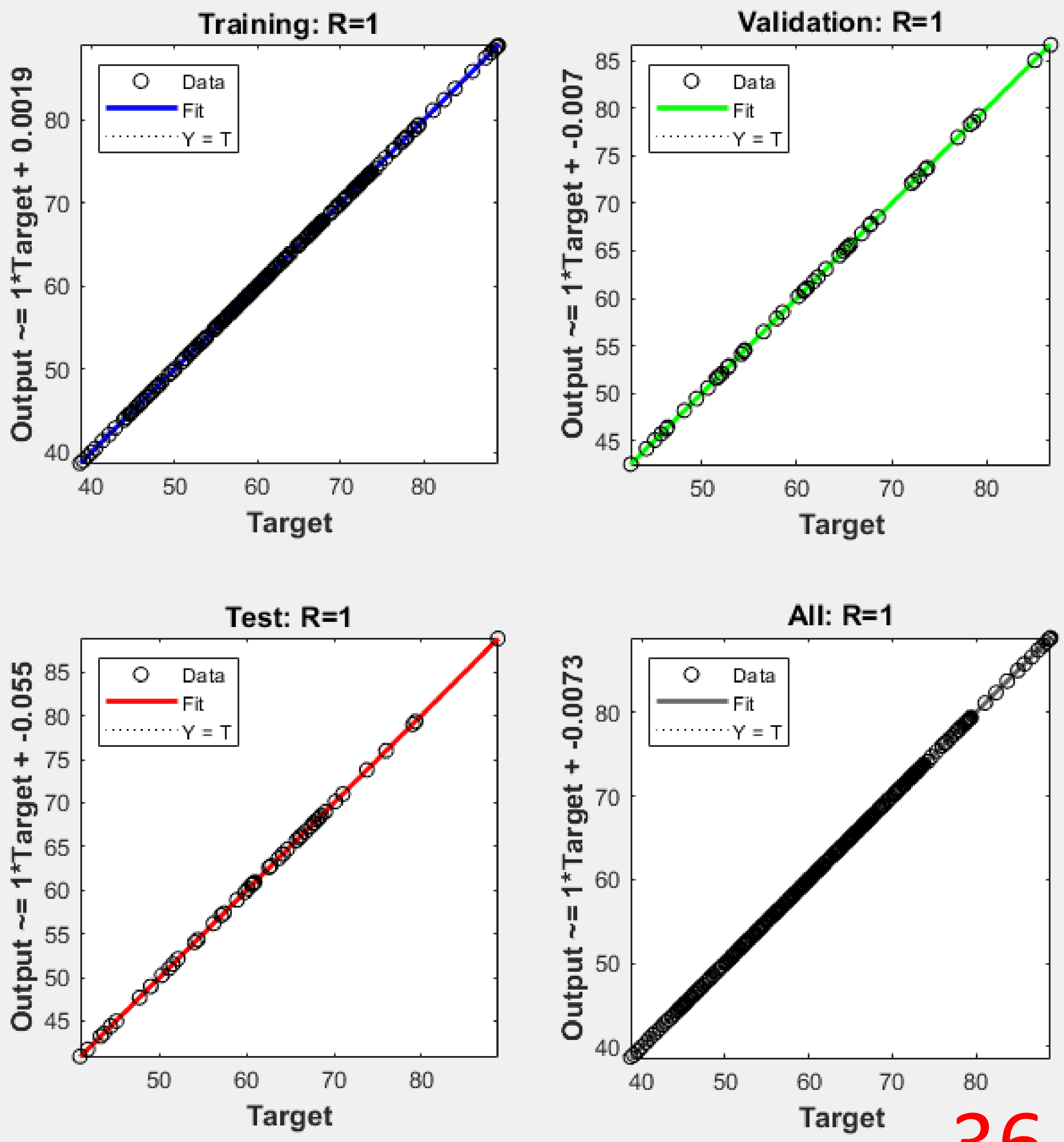


(Q-D) -> H , NN performance



(Q-D) -> H , NN Regression

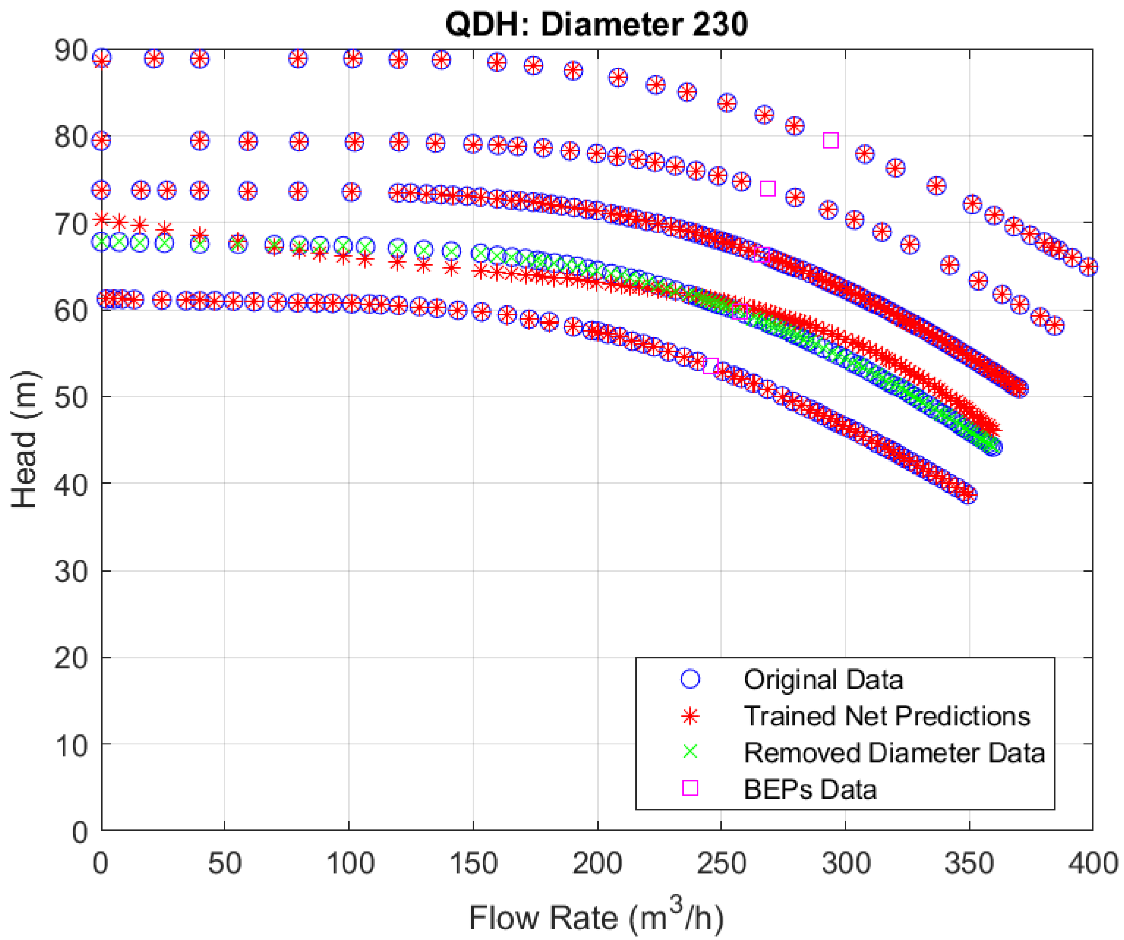
Comparison between predicted and actual values

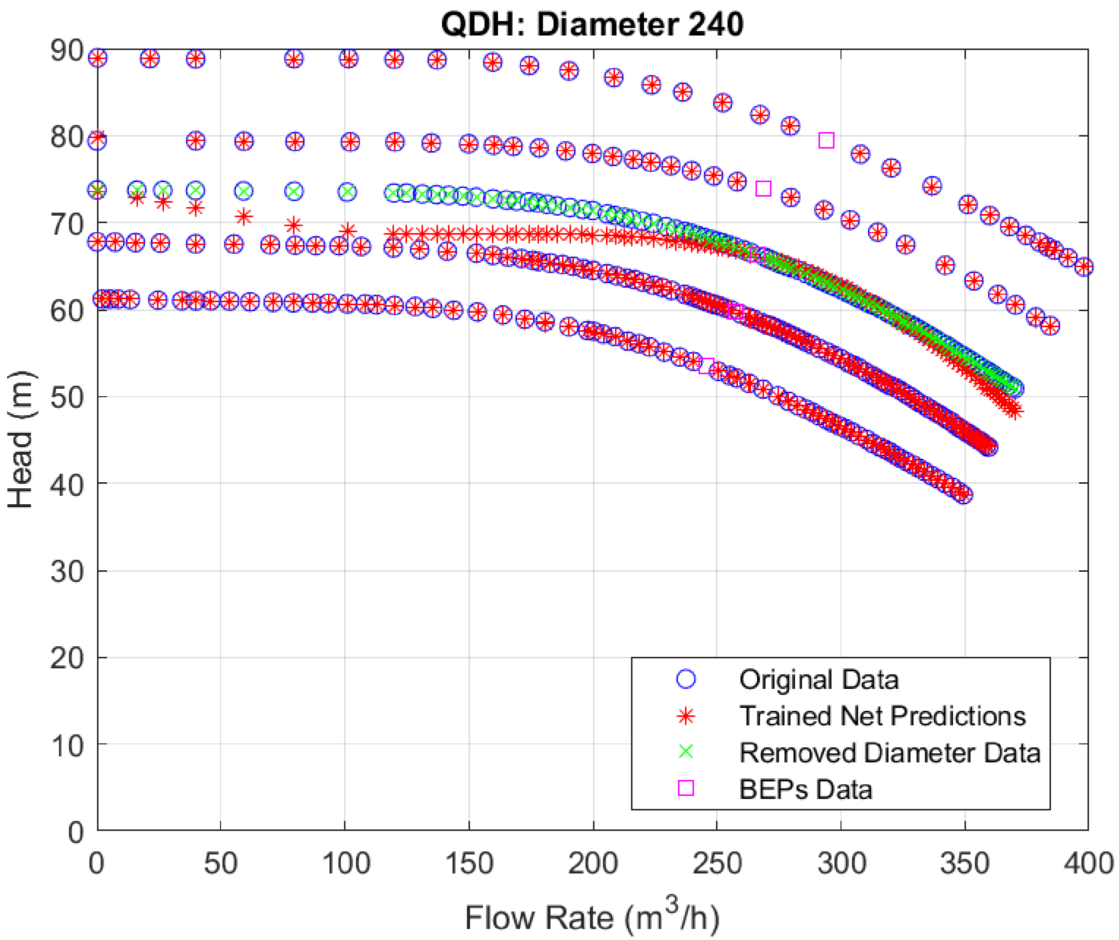


(Q,D) -> H , NN Statistics with complete diameters curves removed from the training dataset

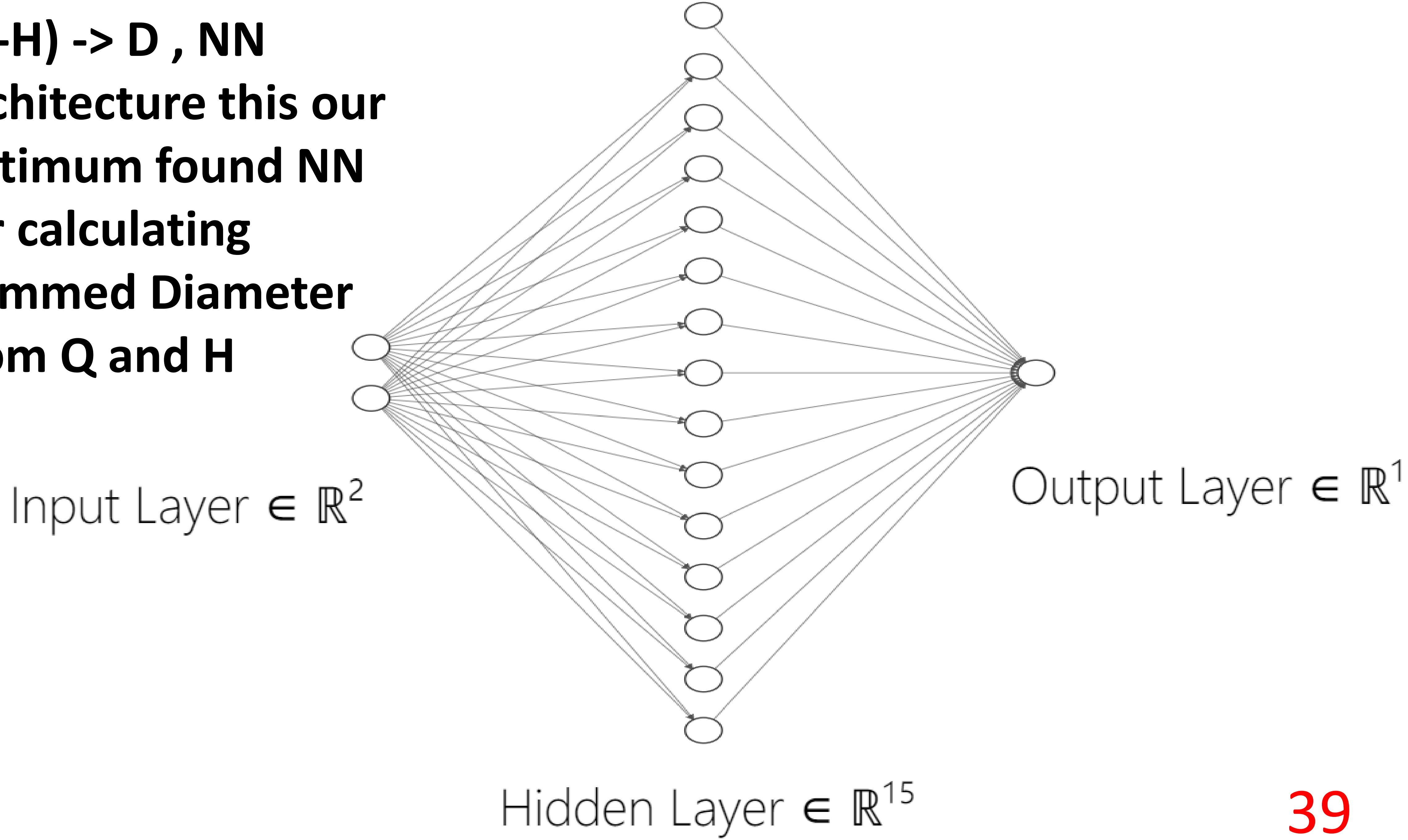
DiameterRemoved	AvgMSE	TrainPerformance	ValPerformance	TestPerformance	MSEDeletedDiameter	MSEBEPs	Score
NaN	0.0008532	0.000846493	0.00082983	0.000886313	NaN	0.000648145	0.00048133
220	0.0009923	0.000665467	0.00098963	0.001479805	16.94746424	3.741905792	7.527934268
230	0.0023287	0.000856961	0.001350094	0.005473421	3.361787196	0.123855919	1.371469264
240	0.0024057	0.001075105	0.001385032	0.005388948	4.145278926	0.003002364	1.660667427
250	0.000928	0.00078843	0.000869935	0.001192634	69.02117353	8.566865515	29.32230344
260	0.0008203	0.000762084	0.000848833	0.000877525	61.10410746	0.051287715	24.4522501

These numbers represent the error calculated as difference between the actual and predicted values

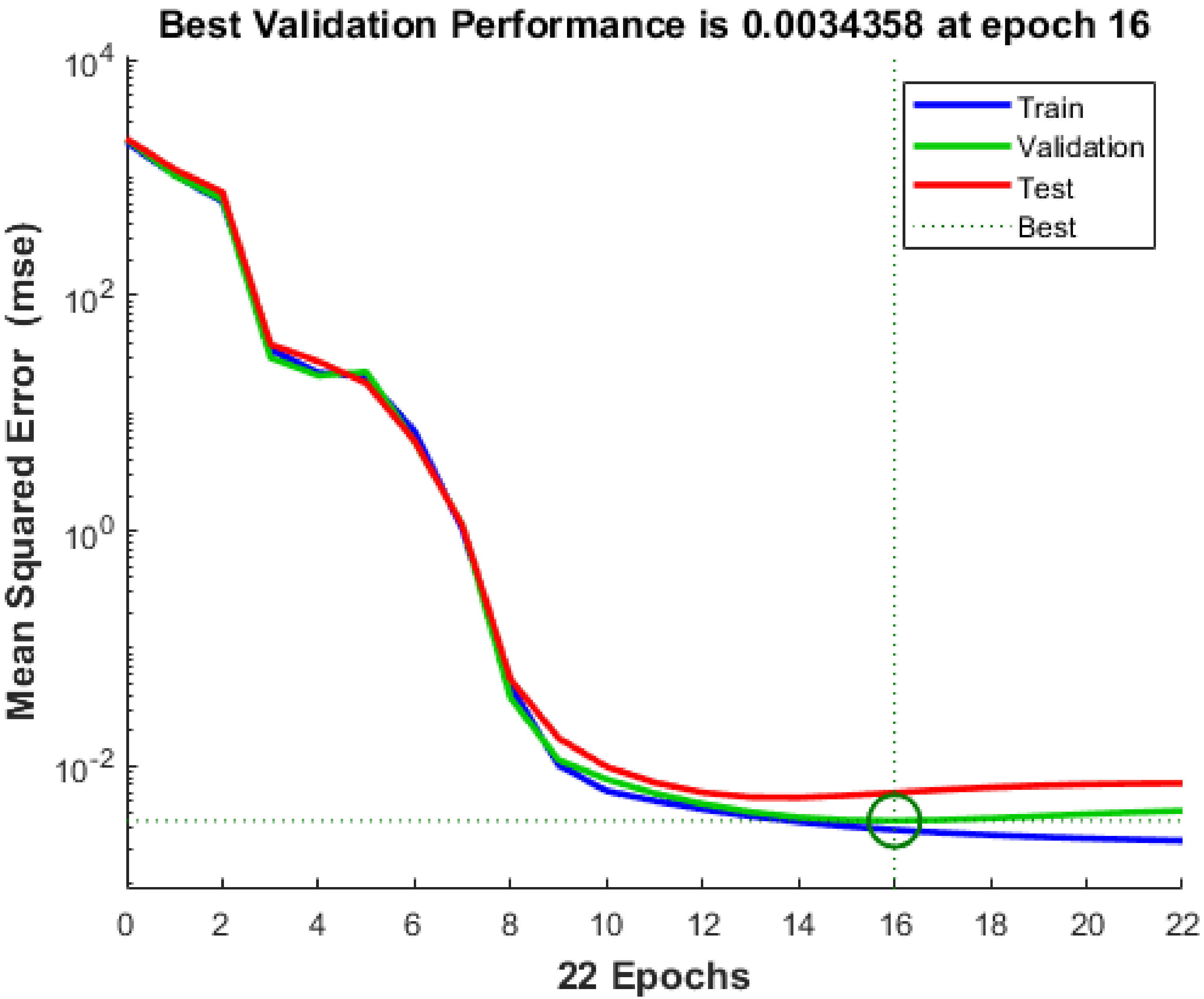

QDH: Diameter 230
Head (m)
Flow Rate (m3/h)
Original Data
Trained Net Predictions
Removed Diameter Data
BEPs Data


QDH: Diameter 240
Head (m)
Flow Rate (m3/h)
Original Data
Trained Net Predictions
Removed Diameter Data
BEPs Data

(Q-H) -> D , NN architecture this our optimum found NN for calculating Trimmed Diameter from Q and H

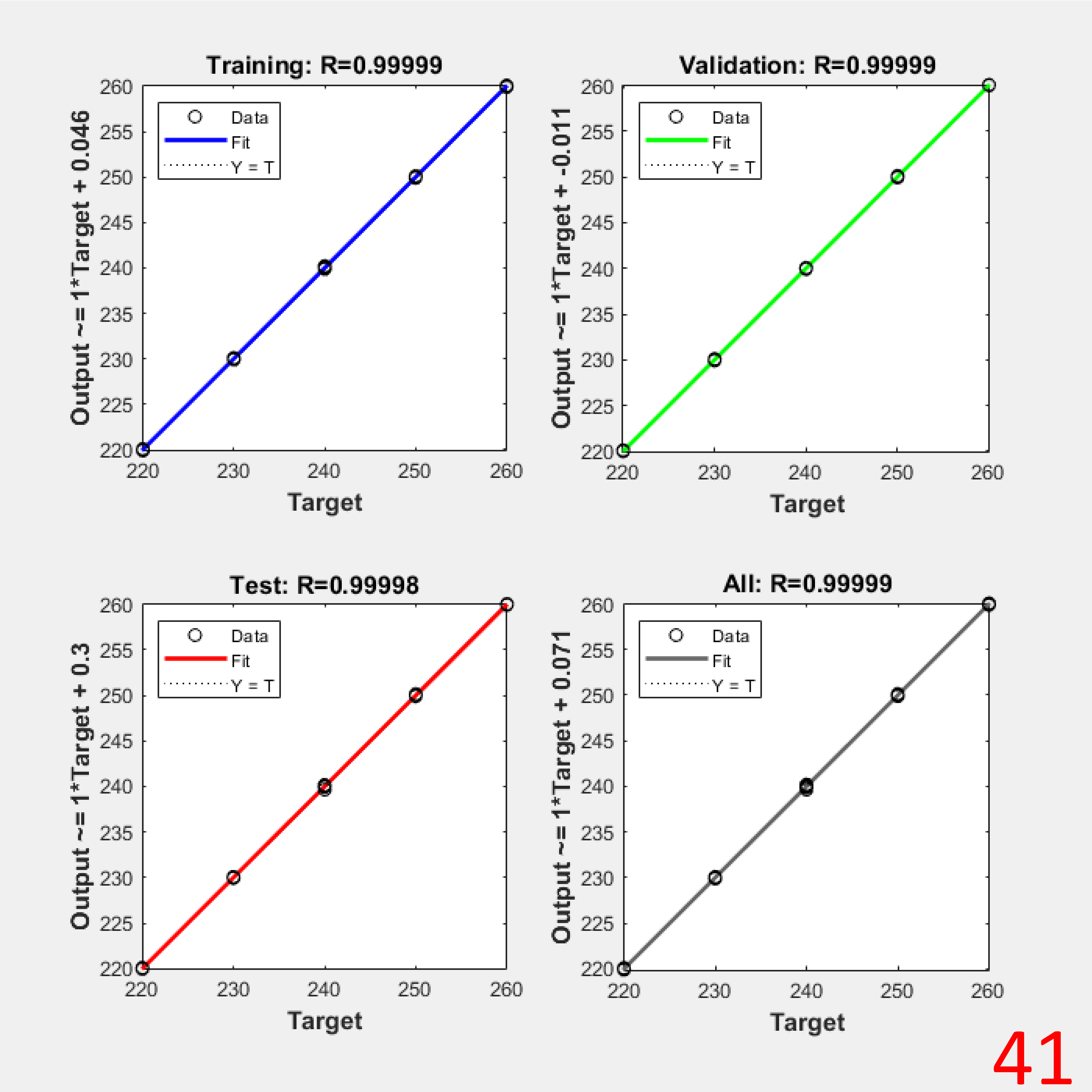


(Q,H)->D ,NN Performance



(Q-H) -> D , NN Regression

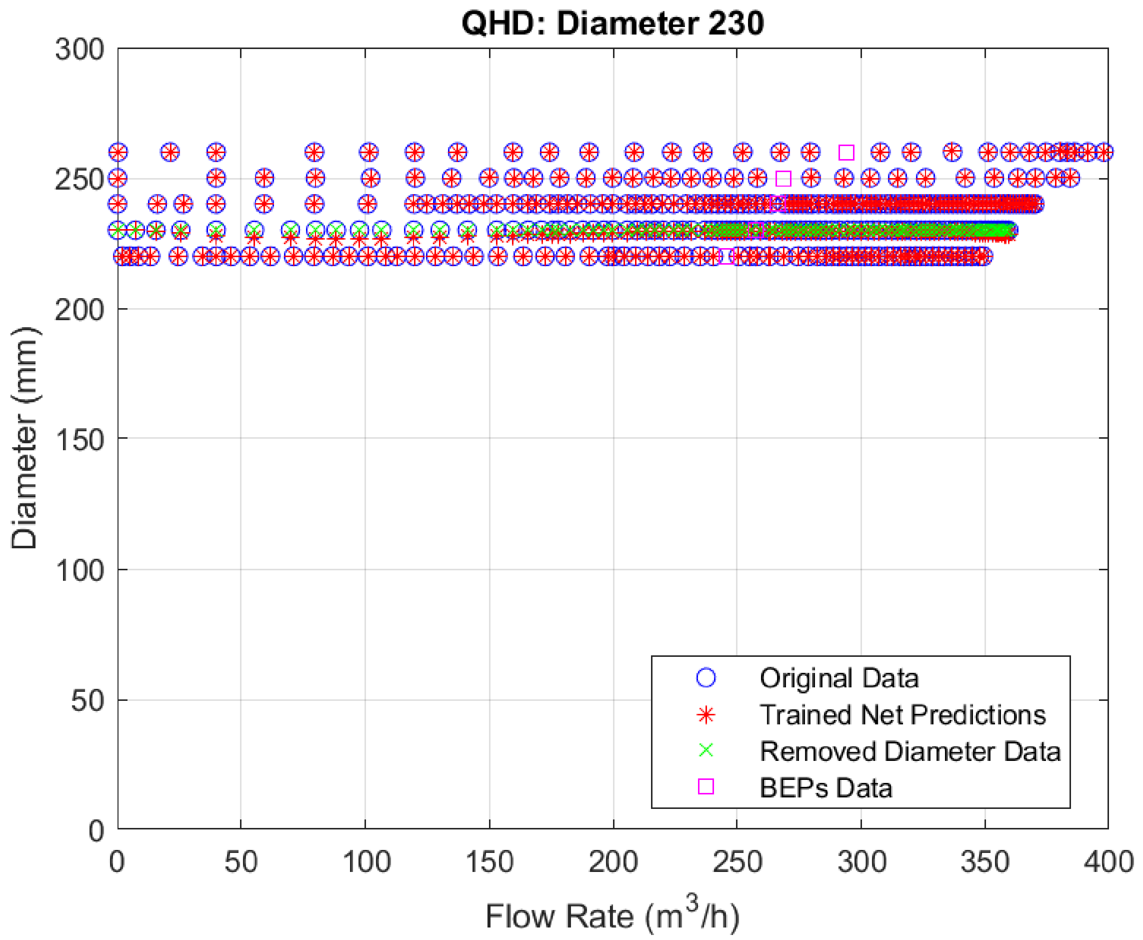
Comparison between predicted and actual values



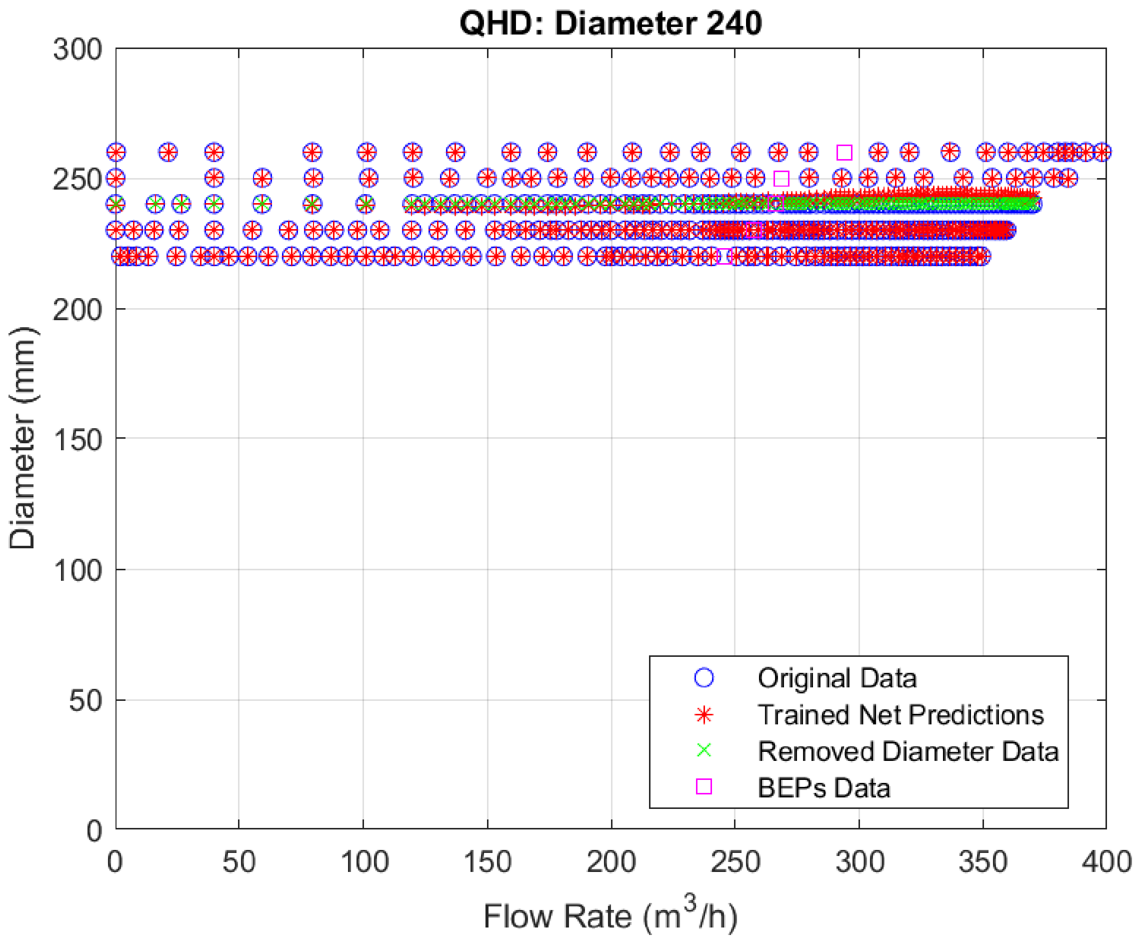
(Q,H) -> D , NN Statistics with complete diameters curves removed from the training dataset

DiameterRemoved	AvgMSE	TrainPerformance	ValPerformance	TestPerformance	MSEDeletedDiameter	MSEBEPs	Score
NaN	0.000853152	0.000846493	0.00082983	0.000886313	NaN	0.000648145	0.00048133
220	0.000992341	0.000665467	0.00098963	0.001479805	16.94746424	3.741905792	7.527934268
230	0.002328746	0.000856961	0.001350094	0.005473421	3.361787196	0.123855919	1.371469264
240	0.002405723	0.001075105	0.001385032	0.005388948	4.145278926	0.003002364	1.660667427
250	0.000927968	0.00078843	0.000869935	0.001192634	69.02117353	8.566865515	29.32230344
260	0.000820265	0.000762084	0.000848833	0.000877525	61.10410746	0.051287715	24.4522501

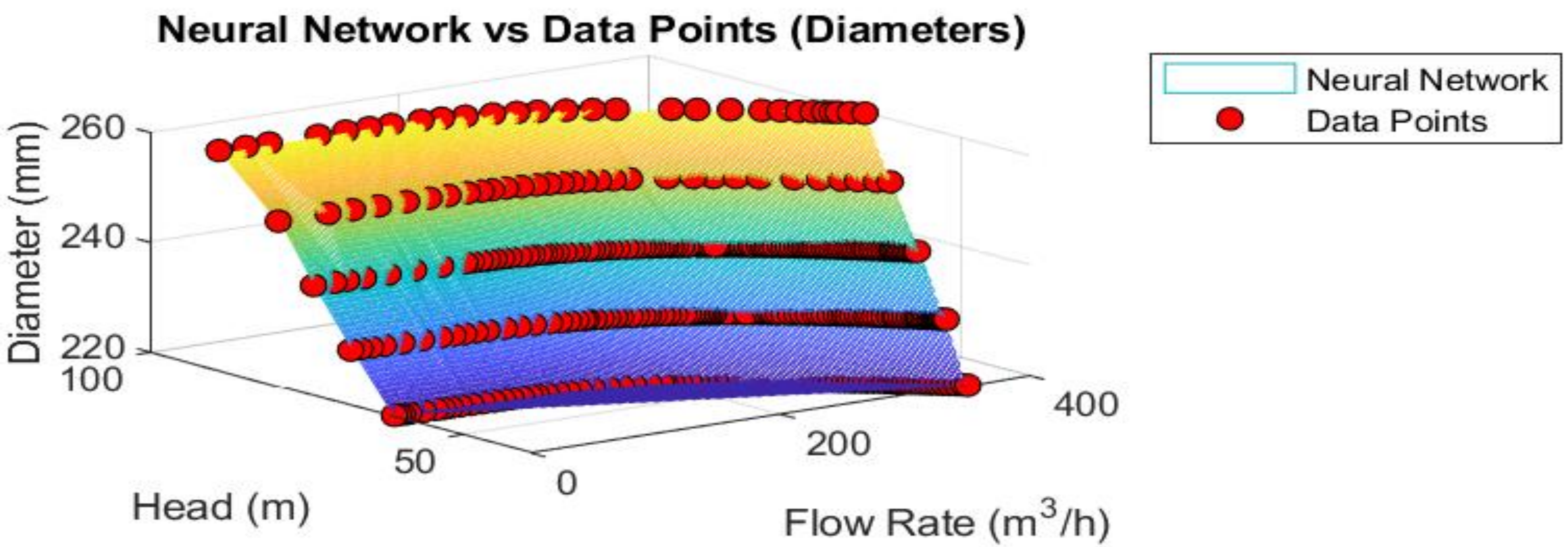
These numbers represent the error calculated as difference between the actual and predicted values

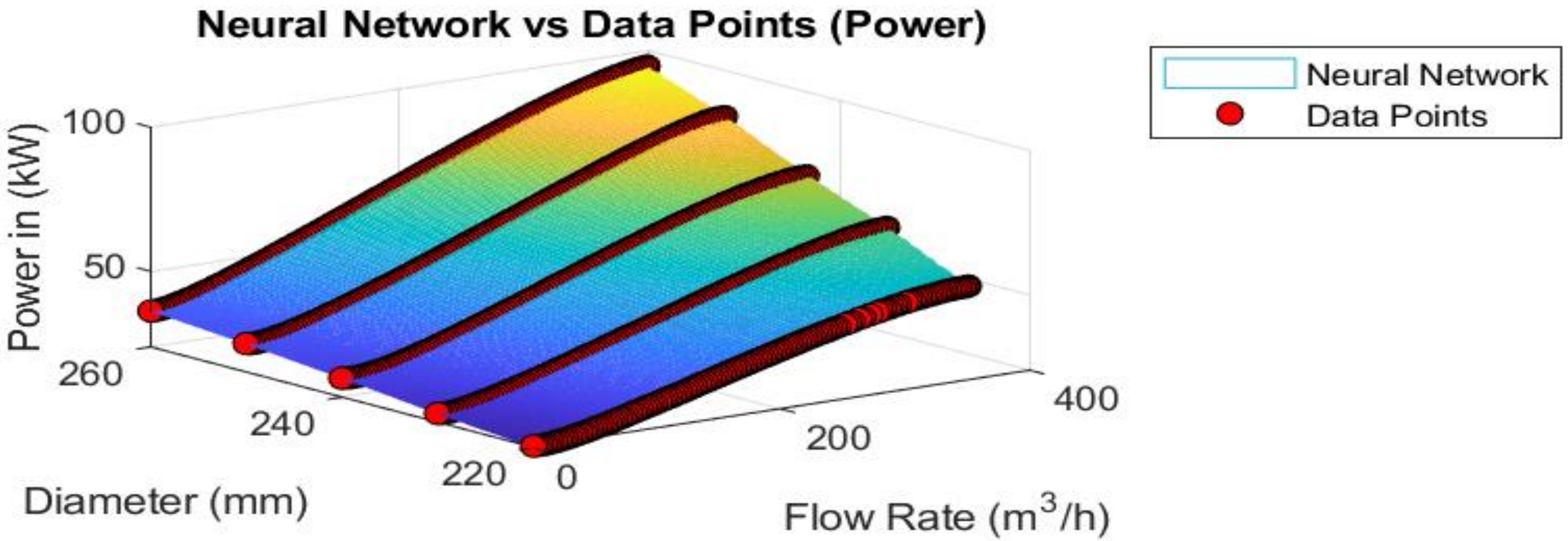


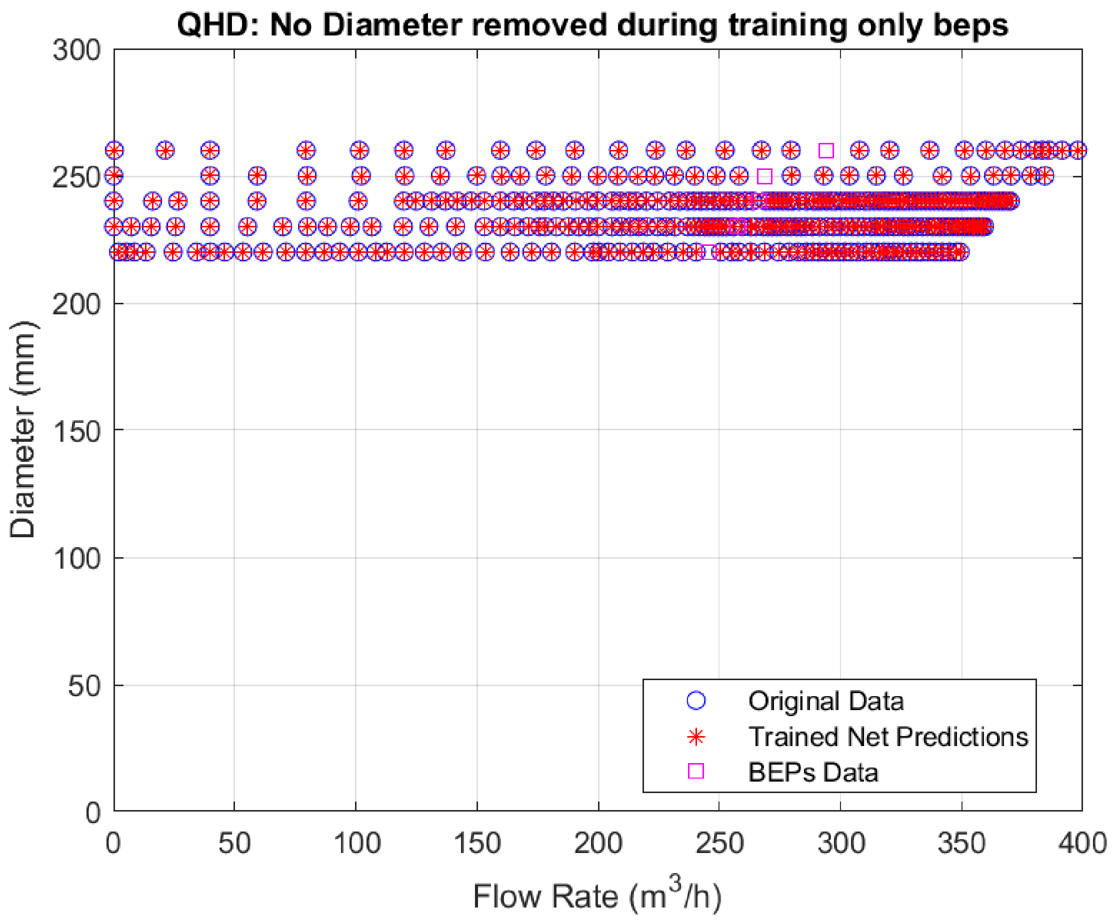
QHD: Diameter 230
300
250
200
150
100
50
0
Diameter (mm)
0
50
100
150
200
250
300
350
400
Flow Rate (m3/h)
Original Data
Trained Net Predictions
Removed Diameter Data
BEPs Data

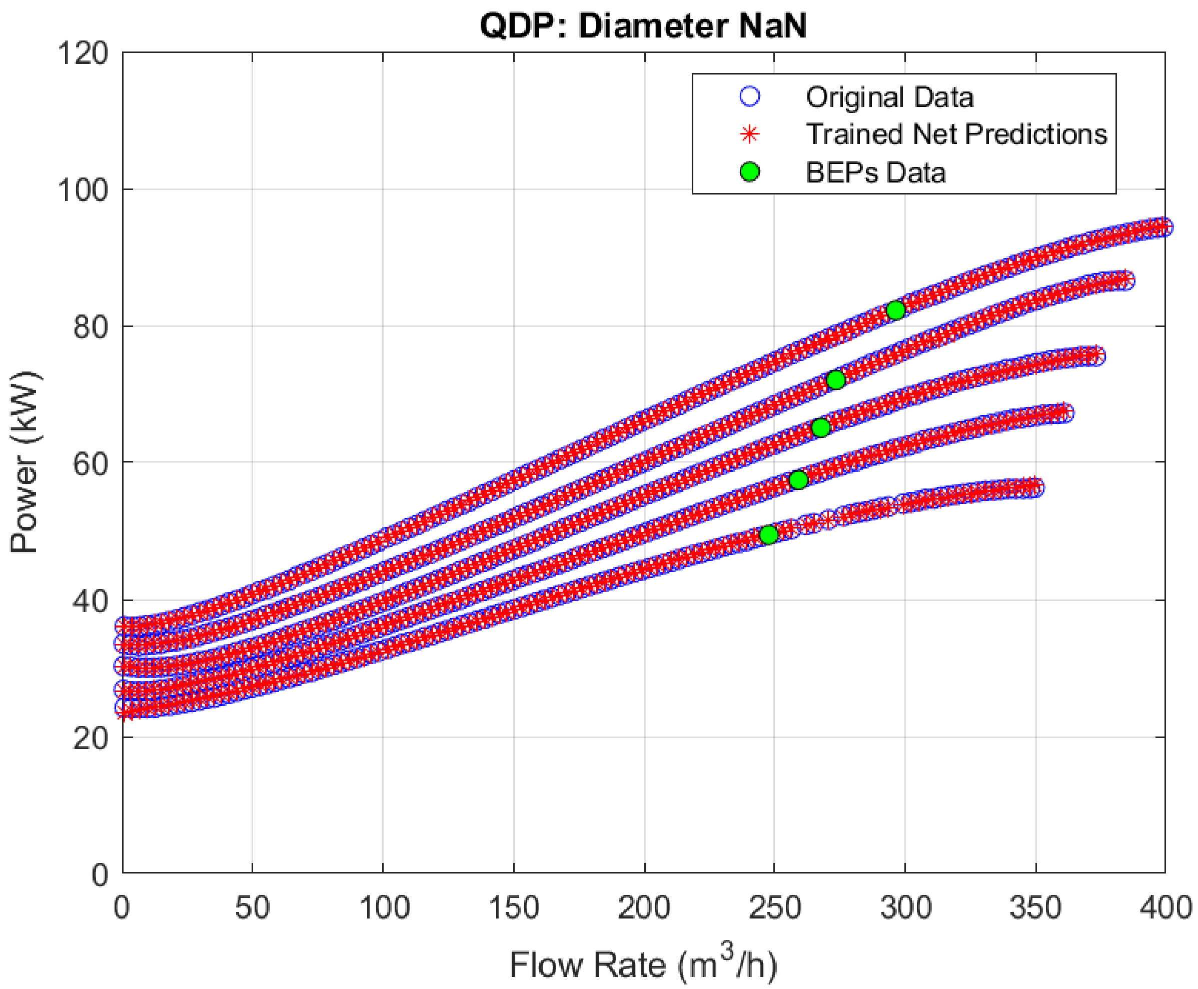


QHD: Diameter 240
300
250
200
150
100
50
0
Diameter (mm)
0
50
100
150
200
250
300
350
400
Flow Rate (m3/h)
Original Data
Trained Net Predictions
Removed Diameter Data
BEPs Data


Neural Network vs Data Points (Diameters)
Neural Network
Data Points
Diameter (mm)
260
240
220
100
50
Head (m)
0
200
400
Flow Rate (m^3/h)


Neural Network vs Data Points (Power)
Neural Network
Data Points
Power in (kW)
100
50
260
240
220
Diameter (mm)
0
200
400
Flow Rate (m^3/h)


QHD: No Diameter removed during training only beps
Diameter (mm)
Flow Rate (m3/h)
Original Data
Trained Net Predictions
BEPs Data


QDP: Diameter NaN
Power (kW)
Flow Rate (m3/h)
Original Data
Trained Net Predictions
BEPs Data

Comparison between NN and traditional trimming method

Constant Area Scaling

Here we compared the neural network diameter prediction given (Q,H) to the prediction we get from **Constant Area Scaling method according to :**

$$\frac{H'}{H} = \left(\frac{D_2'}{D_2}\right)^2 \qquad \frac{Q'}{Q} = \frac{D_2'}{D_2}$$

%error (constand area scaling)	% error in (neural network)
0.702223662	0.04664809
0.511913672	0.013840085
1.332875307	0.000273148
4.44273456	0.032058884

$$\text{percent\_error} = \left|\frac{d_{\text{trimmed}} - d_{\text{act}}}{d_{\text{act}}}\right| \cdot 100\%$$

CONCLUSIONS

1. **The presented study indicated that GA-ANN can be used in predicting the performance of centrifugal pumps with impeller trimming precisely and accurately ,than the other available methods.**
2. **The application of ANN Method needs less time and less cost in order to get the optimum impeller size and corresponding energy saving.**

Future Work

Applications of ANN will be studied for prediction the centrifugal pump performance using different industrial fluids like highly viscous liquids and non-Newtonian slurry Fluids.

QHD Results:

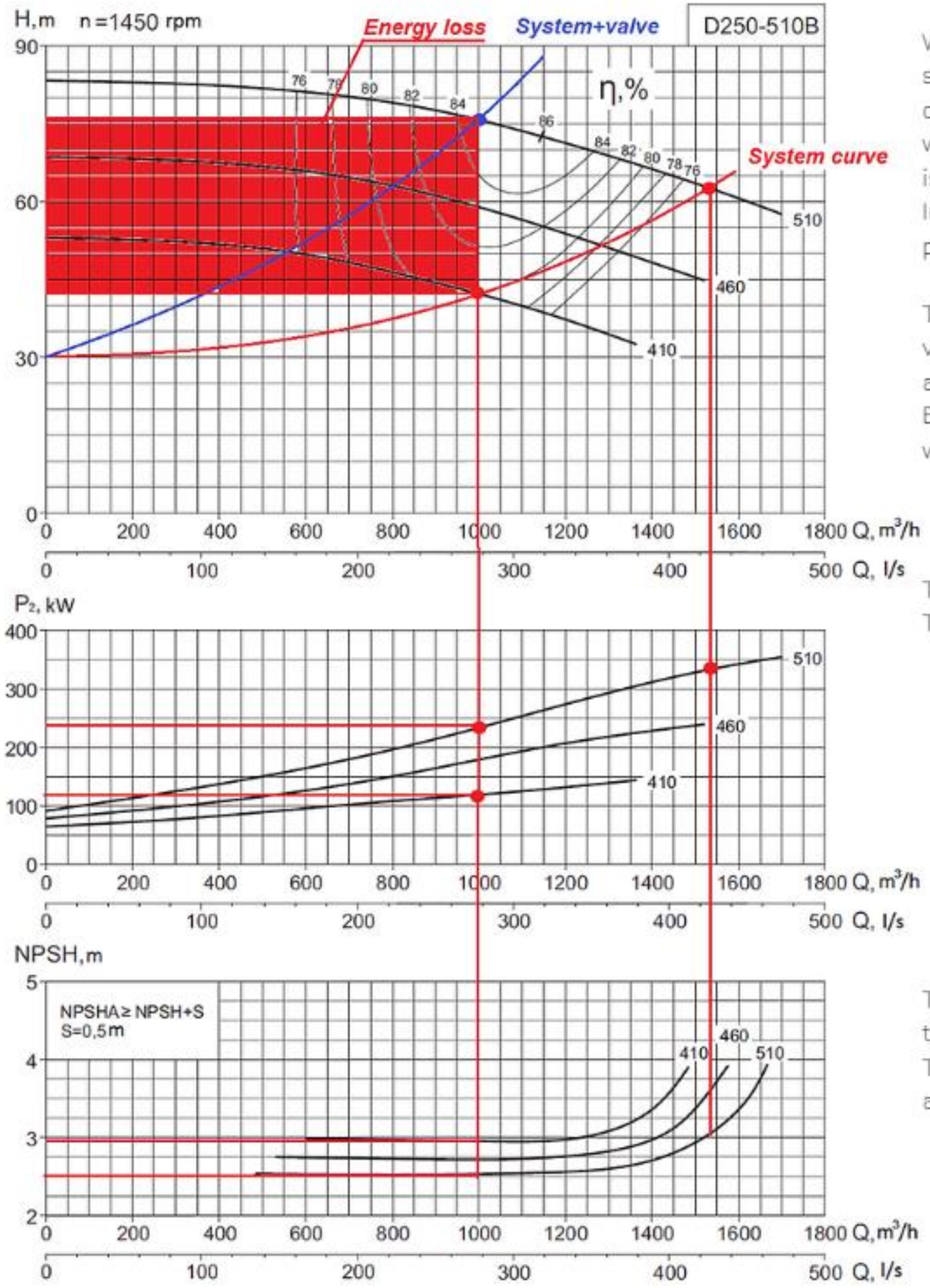
DiameterRemoved	AvgMSE	TrainPerformance	ValPerformance	TestPerformance	MSEDeletedDiameter	MSEBEPs	Score
NaN	0.0059276	0.0027995	0.0085836	0.0078929	NaN	0.0012612	0.0034439
220	0.002961	0.0010262	0.003807	0.0049843	22.836	4.245	9.9852
230	0.0027679	0.0016405	0.0020061	0.005189	1.4926	0.03873	0.60671
240	0.0032114	0.0016348	0.0019197	0.0068285	4.5879	0.28366	1.8944
250	0.0020779	0.001477	0.0032896	0.001756	1.9477	0.048723	0.7896
260	0.0037689	0.0029197	0.0060066	0.0027841	16.887	0.024424	6.761

QDP Results:

DiameterRemoved	AvgMSE	TrainPerformance	ValPerformance	TestPerformance	MSEDeletedDiameter	MSEBEPs	Score
NaN	0.0061475	0.0055729	0.0061324	0.007011	NaN	0.014941	0.0057486
220	0.005917	0.0048547	0.0034999	0.0099041	290.19	88.169	133.71
230	0.0013612	0.0011326	0.001204	0.0018558	0.55901	0.048094	0.23393
240	0.084459	0.10195	0.064123	0.078944	0.48975	0.076482	0.24203
250	1.2453	1.1618	1.3735	1.2407	0.99093	2.2473	1.3488
260	0.95759	0.91676	0.97803	0.99742	46.133	11.336	21.119

QDH Results:

DiameterRemoved	AvgMSE	TrainPerformance	ValPerformance	TestPerformance	MSEDeletedDiameter	MSEBEPs	Score
NaN	0.00085315	0.00084649	0.00082983	0.00088631	NaN	0.00064815	0.00048133
220	0.00099234	0.00066547	0.00098963	0.0014798	16.947	3.7419	7.5279
230	0.0023287	0.00085696	0.0013501	0.0054734	3.3618	0.12386	1.3715
240	0.0024057	0.0010751	0.001385	0.0053889	4.1453	0.0030024	1.6607
250	0.00092797	0.00078843	0.00086994	0.0011926	69.021	8.5669	29.322
260	0.00082026	0.00076208	0.00084883	0.00087752	61.104	0.051288	24.452



When the pump with impeller 510 mm is selected the duty point will be on the cross of **H-Q curve** and **System curve** with a flow 1540 m3/h, but required flow is 1000 m3/h.
In addition, the pump consumes extra power and overloads the electric motor.

To meet system requirements, a control valve is used. Operating point shifts to a flow 1000 m3/h.
But control valve throttle a pump head what cause the **Energy loss**.

The difference in shaft power is 120 kW.
The saving potential is very substantial.

The difference in NPSH between fullsize and trimmed impeller is 0.5 m.
The difference is necessary to take into account.