

# Estimation of Algorithm Execution Time Using Machine Learning

Author: Martin Buchta  
Supervisor: doc. Ing. Jiří Jaroš Ph.D.



## k-Wave

- ultrasound simulation
- distributed with MPI
- heavy usage of distributed 3D FFT

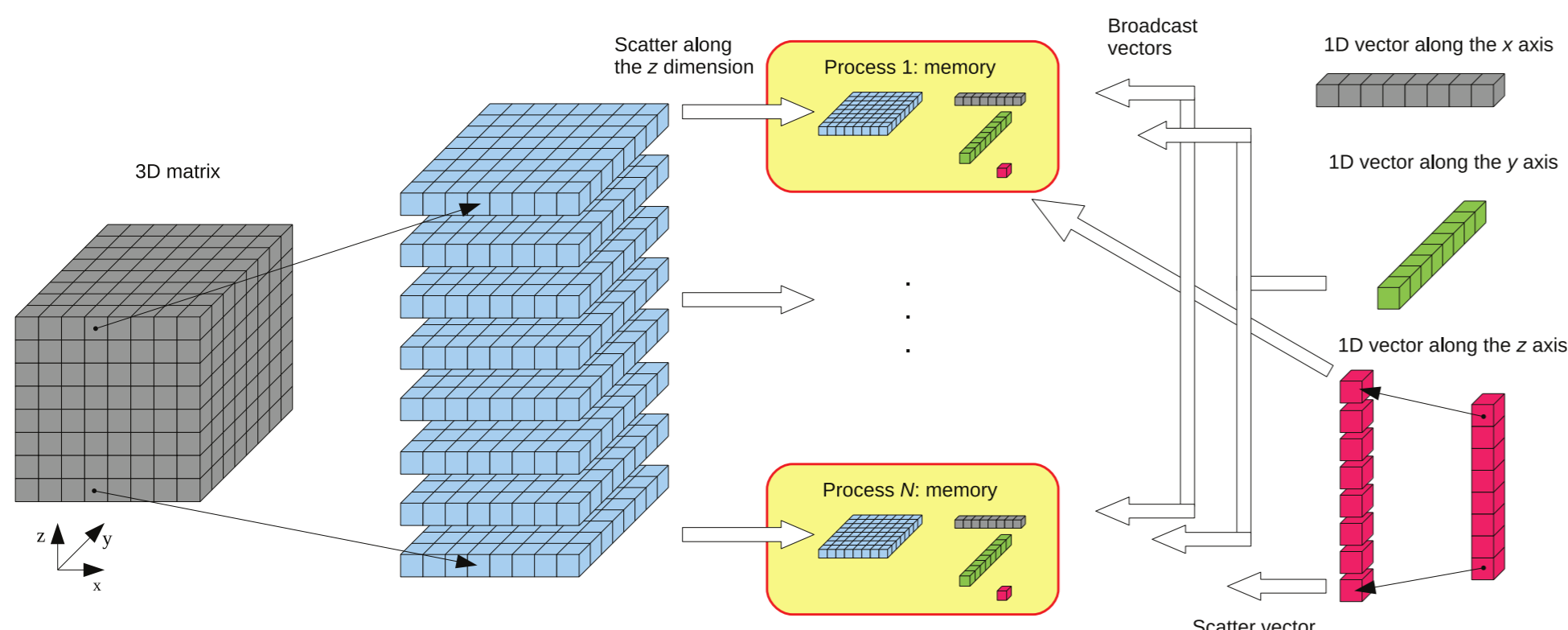


Image 1: Principle of distributing k-Wave simulation [1]

## Captured data

- in total 2080 runtime measurements on different domains
- captured dimensions, number of nodes, number of ranks and time per step
- nodes from 1 to 100, dimensions from  $512^3$  to  $3072^3$

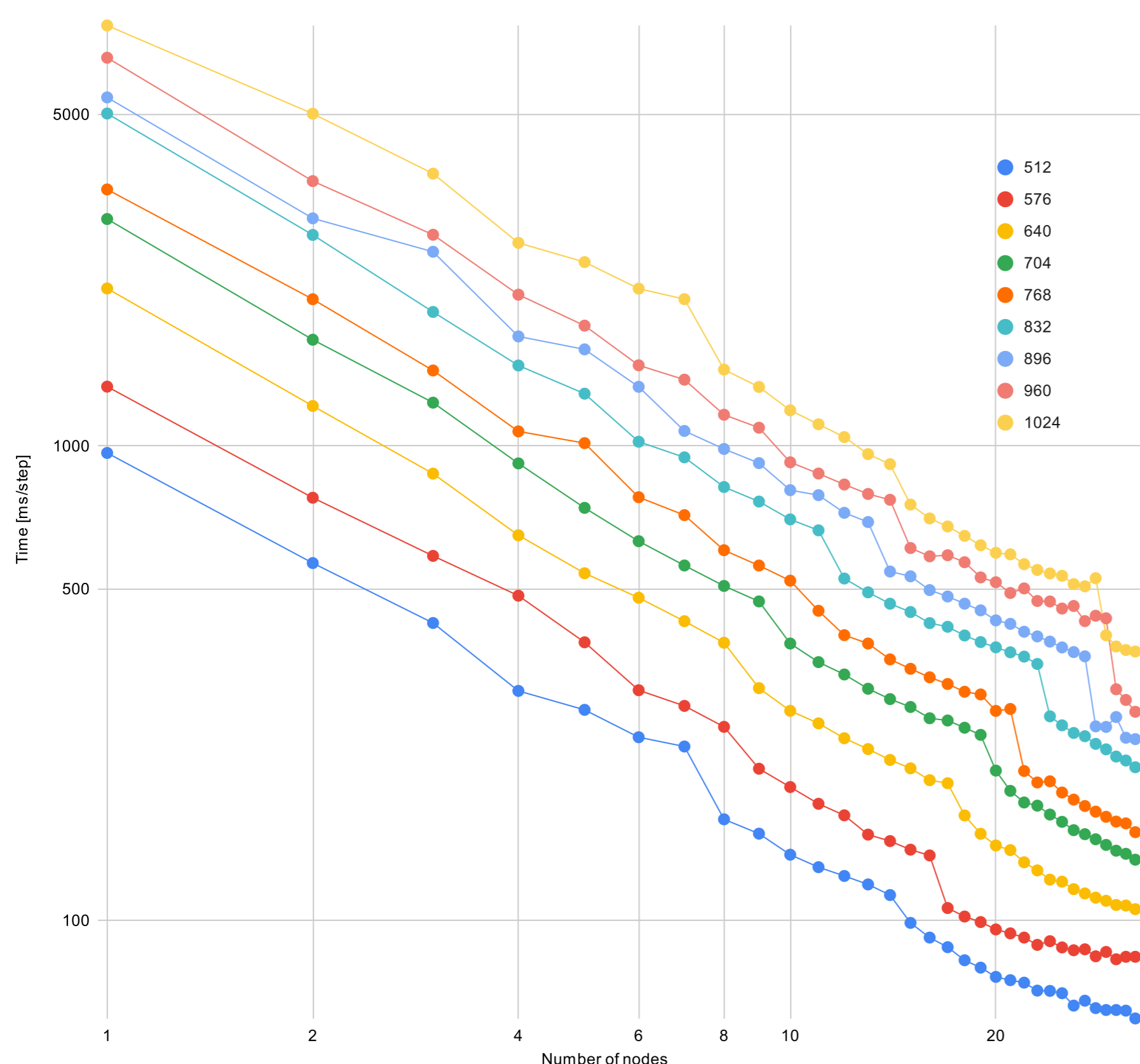


Image 2: Selected data captured

## Existing solutions

- predicting k-Wave MPI runtime using interpolation and splines, error on unknown space dimensions up to 15%
- manual estimation
- fixed prediction of 24 hours

## Solution

### 1) Symbolic regression

- trained with genetic algorithms, using HeuristicLab
- $10^6$  generations, each with 100 individuals
- 10% mutation probability

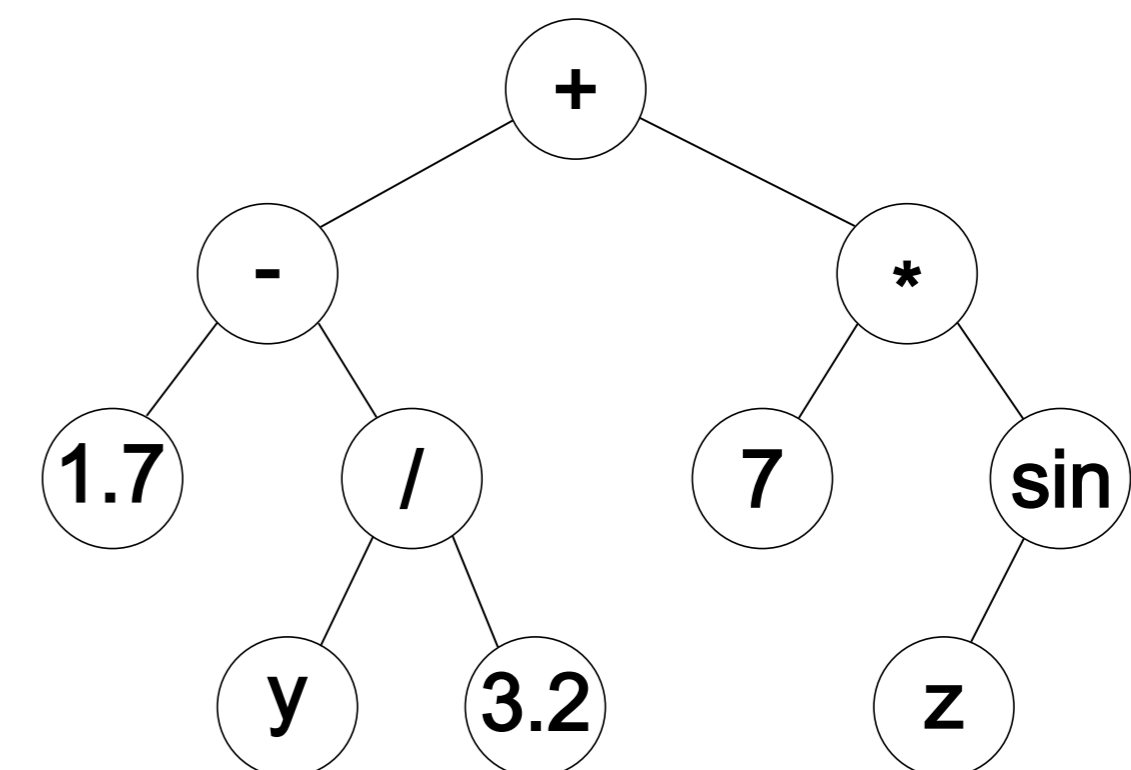


Image 3: Example symbolic regression tree

### 2) Neural network

- 6 hidden layers with 256 neurons and ReLU activation
- dropout layers
- data normalization
- early stopping

## Results

- average relative error

### a) suitable factor of axis sizes,

	Symbolic regression	Neural network
Train	5,58 %	5,75 %
Validation	11,26 %	21,71 %
Test	5,64 %	12,46 %

Table 1: Comparison of models

### b) non-filtered data

	Symbolic regression	Neural network
Train	13,7 %	9,35 %
Validation	9,31 %	10,33 %
Test	13,55 %	8,25 %

Table 2: Comparison of models on all data

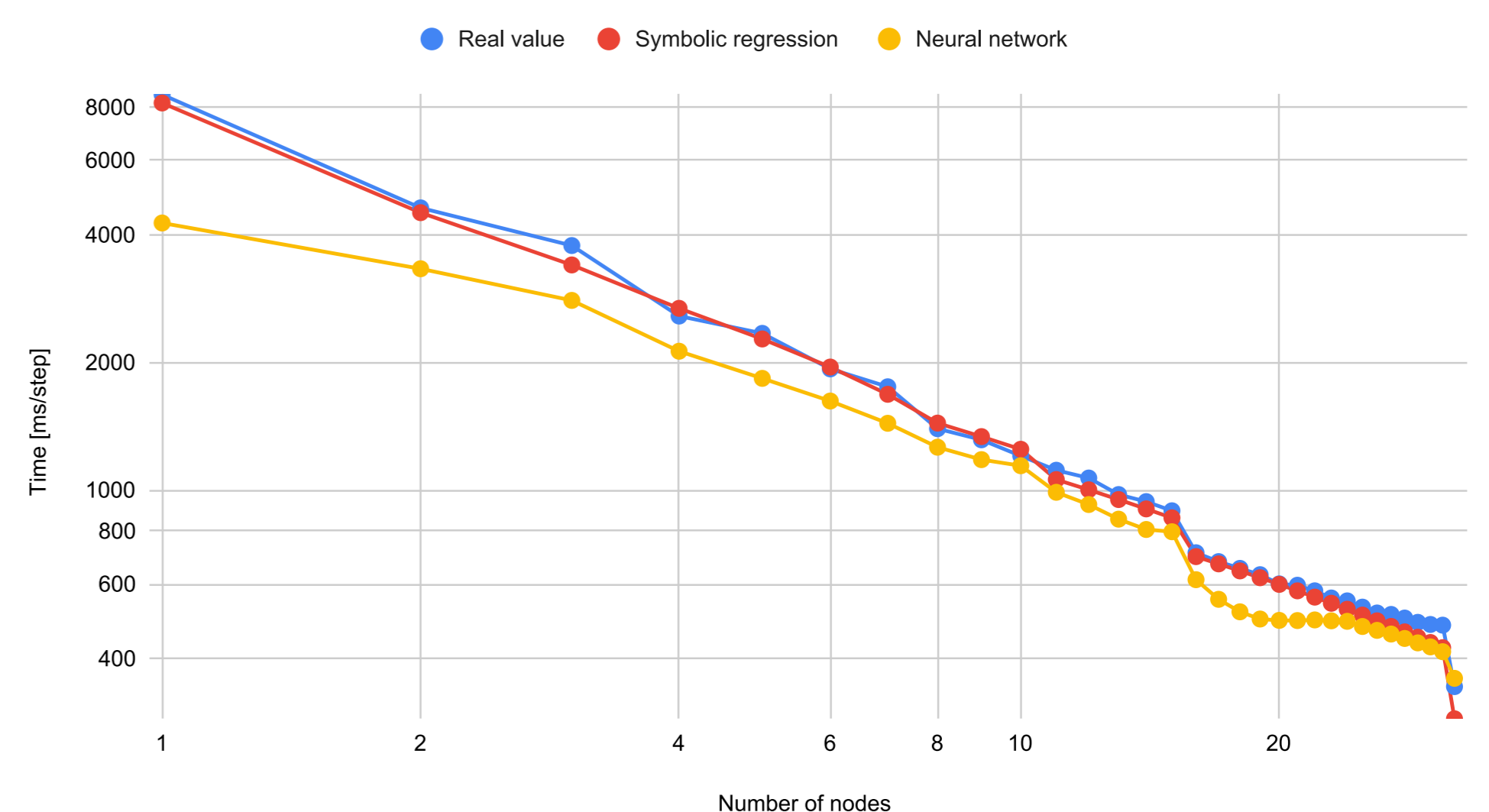


Image 4: Estimated values on domain  $1024^3$  with suitable factor

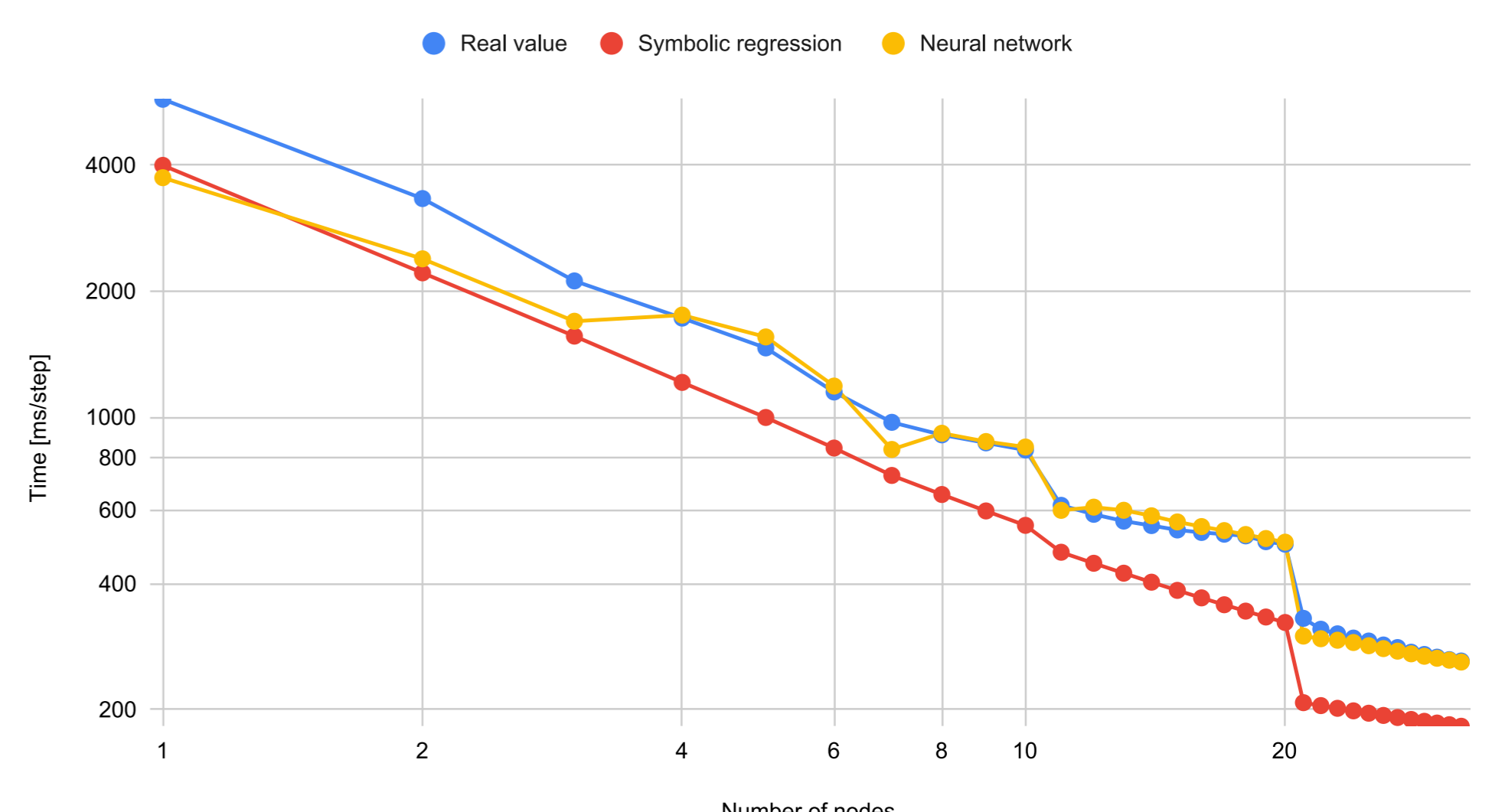


Image 5: Estimated values on domain  $736^3$  with not suitable factor