The 9th International Conference on
Deep Learning in Computational Physics

July 4, 2025, Moscow, Russia

Finding optimal carbon dots synthesis parameters for quantitative analysis of components in multi-component aqueous solutions using machine learning*

Artem Guskov[1,2], Igor Isaev[2], Kirill Laptinskiy[2], Tatiana Dolenko[1,2], Sergey Dolenko[2]

[1] Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
[2] D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia

This study has been performed at the expense of the grant of Russian Science Foundation (project no. 22-12-00138-P)

Actuality

The problem of determination of concentrations of substances dissolved in water is very important for:

- Oceanology
- Ecological monitoring
- Control of industrial and waste waters

This problem is required to be solved in non-contact express mode with acceptable precision

Optical nanosensors

Widely used tool for analyzing multicomponent liquid media:

- Optical nanosensors based on **carbon dots** (CD)

Photoluminescence (PL) of CD is highly sensitive to changes in medium parameters, such as:

- pH
- Solution temperature
- Type and concentration of dissolved substances

Unique properties of CD

Optical properties:

- Intense and stable PL
- PL depends on wavelength of exciting radiation

Experimental simplicity:

- May be doped with various heteroatoms
- Surface may be easily functionalized to modify properties

Biological properties:

- Non-toxic
- Bio-compatible

Synthesis of CD

CD PL properties depend on the synthesis method and conditions

Common method for producing CD – hydrothermal synthesis

Parameters of synthesis that affect CD PL properties:

- Synthesis conditions (reaction time, temperature)
- Sets of precursors and their concentration ratios

CD properties that are affected:

- Wavelength of PL radiation
- Quantum yield of luminescence

Optimal synthesis parameters of CD

To successfully solve a concrete task,
carbon dots with specified optical properties are required

Key challenge – identifying optimal synthesis parameters of CD
that provide the necessary properties

To address this challenge, we considered creation
of an approximation model using machine learning

We use a neural network to approximate the dependence
of the error in determining concentrations of heavy metal salts
in water solutions on CD synthesis parameters

Task of the study

The task of this study is to identify optimal CD synthesis parameters that ensure high-precision determination of the concentration of specific heavy metal salts in aqueous solutions

Materials and methods

Method for producing CD – hydrothermal synthesis

Precursors – citric acid (CA) and ethylenediamine (EDA)

Synthesis conditions:

- Reaction time – from 30 to 360 minutes
- Temperature – from 100 to 200 °C

Total 74 types of CD with different synthesis parameters – reaction time, temperature of synthesis, CA:EDA concentration ratio – were obtained

Materials and methods

Heavy metal salts – $Co(NO_3)_2$, $Cu(NO_3)_2$

Salts concentration – from 0 to 6 mM with 0.67 mM step

Each type of CD was placed in water solutions of heavy metal salts with a fixed concentration of 5 mg/L

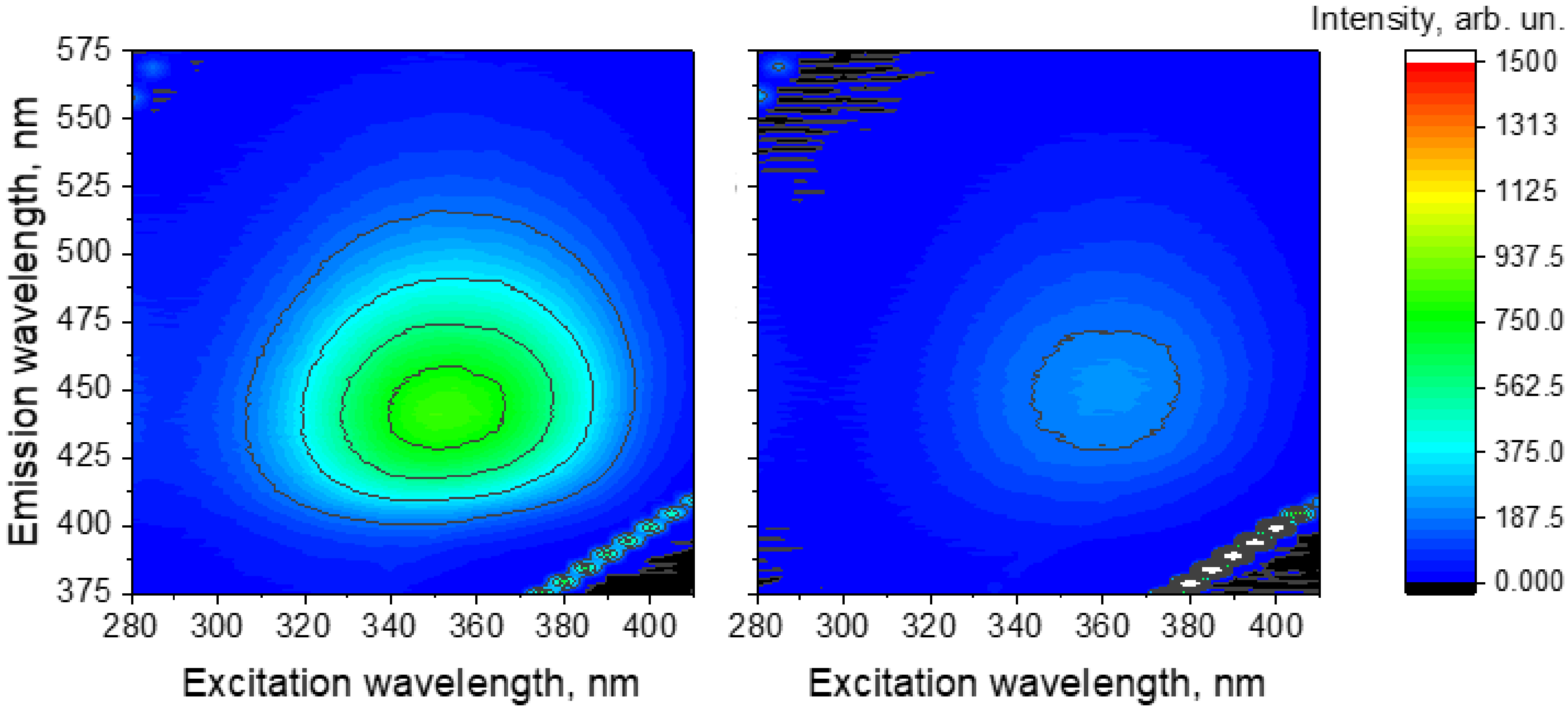
As a result, for each type of CD 100 aqueous solutions with different salts concentration were prepared

For each sample, 2D fluorescence spectrum was measured

2D fluorescence spectra

Aqueous solution of CD:

- *Left* – without salts
- *Right* – with salts $Co(NO_3)_2$ and $Cu(NO_3)_2$ of 6 mM each



Decomposing the main task

The main task is divided into two stages:

1. For each type of CD, solving the inverse problem of determining
   the concentrations of heavy metal salts in aqueous solutions
   from 2D fluorescence spectra of CD
   by machine learning methods
   Estimating the error

2. Building an approximation model:
   Error vs synthesis parameters

Stage 1 – concentrations from spectra

Data role	Data	Dimensionality
Input	2D fluorescence spectra	5 427 features
Output	Salt concentration	1 feature

For each type of CD a separate inverse problem was solved

Each separate task uses 100 patterns – spectra of 100 samples

These patterns were divided into training and test sets (80:20)

10-fold cross-validation was used

Machine learning algorithm – linear regression

Stage 1 – concentrations from spectra

№	Sythesis parameters			MAE, mM	
	CA:EDA	Temperature, °C	Time, min	$Co(NO_3)_2$	$Cu(NO_3)_2$
1	1	100	240	0.70	0.84
2	1	100	300	0.72	0.68
3	1	100	360	0.93	0.84
4	1	120	120	1.05	0.70
…	…	…	…	…	…
71	20	180	180	0.79	1.04
72	20	180	240	0.96	0.91
73	20	180	300	0.61	0.85
74	20	180	360	1.21	1.33

Statistics	MAE, mM		Relative MAE, %	
	$Co(NO_3)_2$	$Cu(NO_3)_2$	$Co(NO_3)_2$	$Cu(NO_3)_2$
AVG	0.77	0.72	12.9	12.1
MIN	0.46	0.42	7.7	7.1
MAX	1.25	1.4	20.8	23.3

Stage 2 – approximation model

Data role	Data	Dimensionality
Input	Synthesis parameters	3 features
Output	Error in determining salt concentration	1 feature

74 patterns – one pattern for each type of CD

Approximation model – neural network

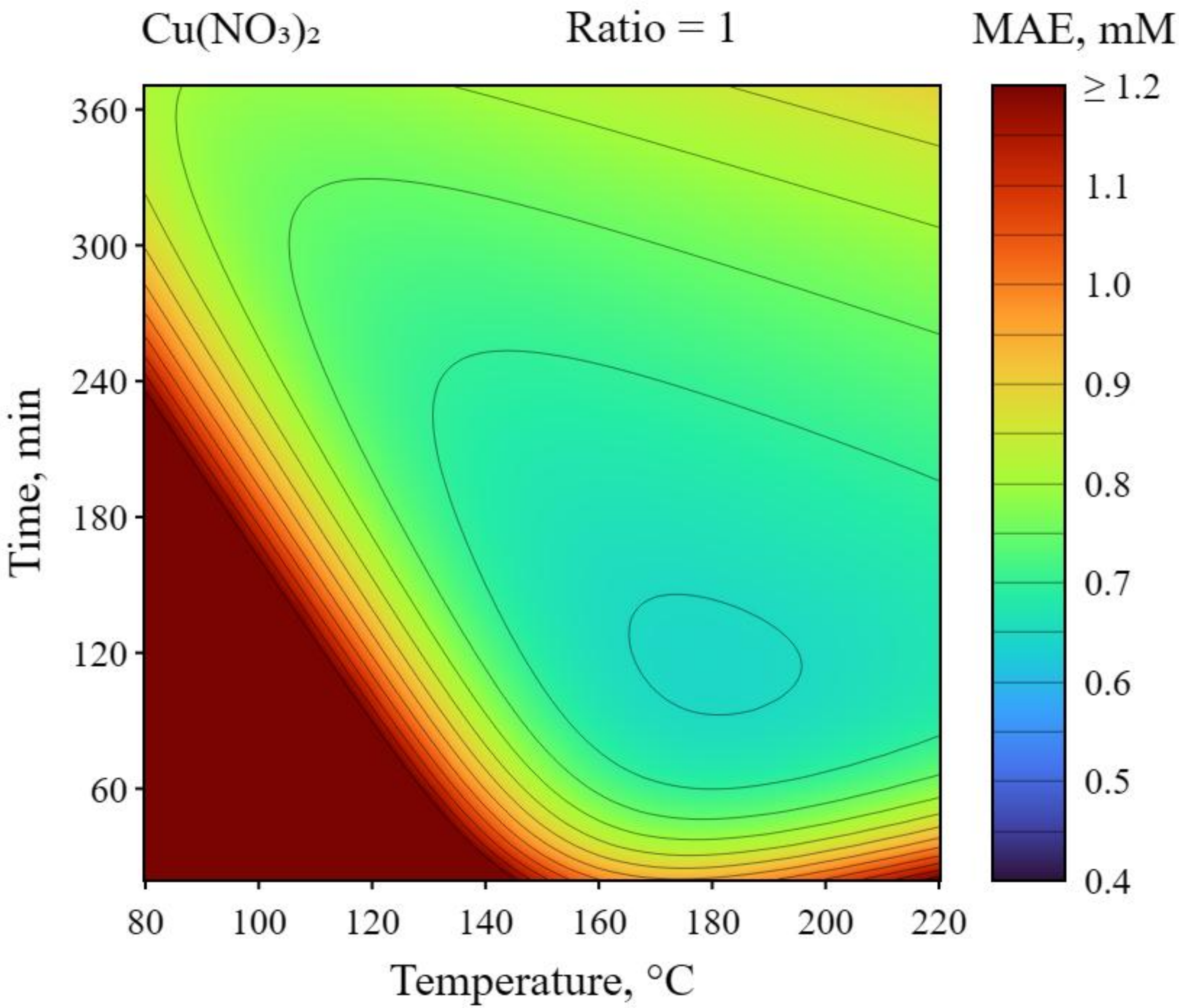
Stage 2 – approximation model

Parameters of the approximation model**:**

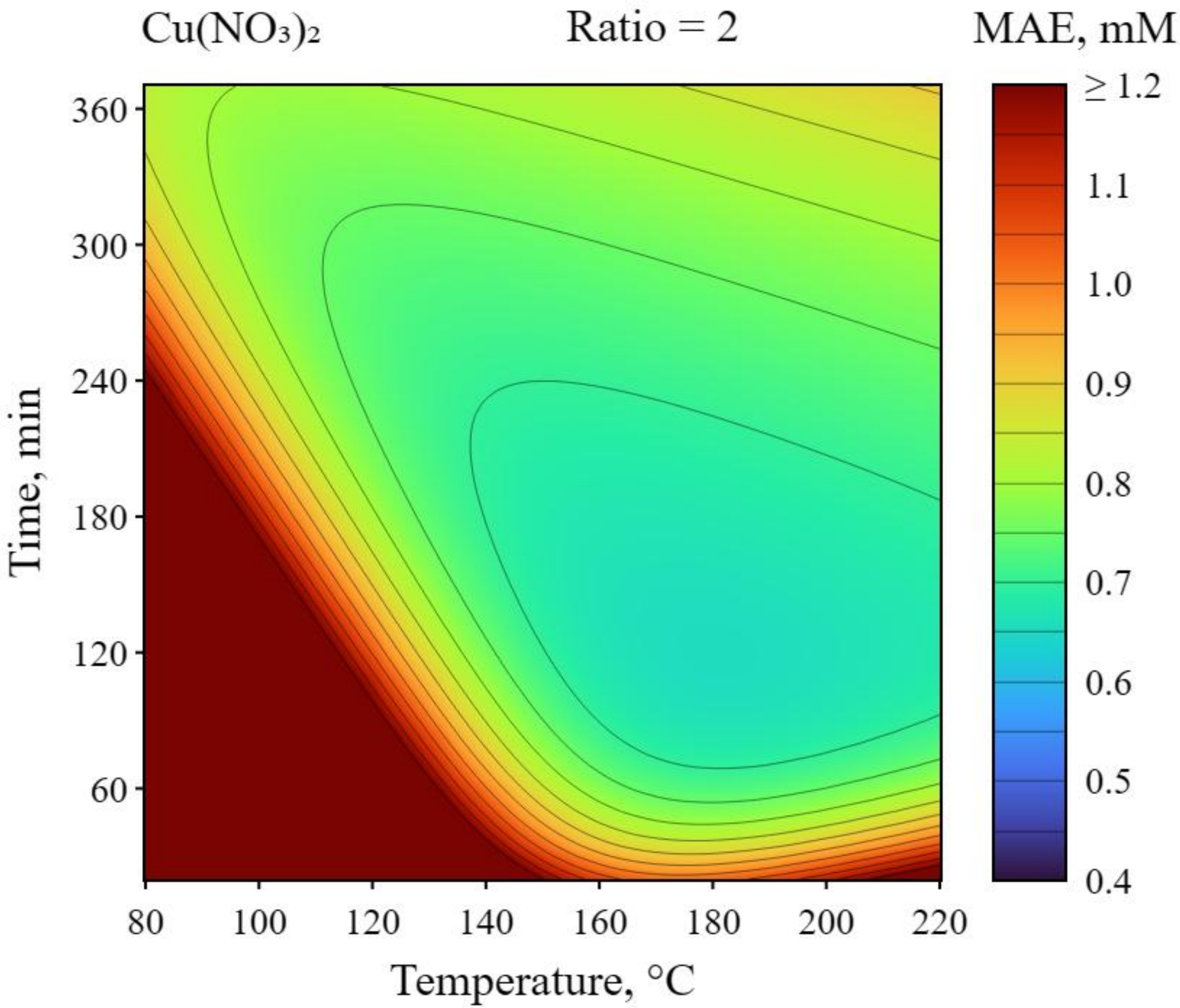
- Perceptron with 1 hidden layer
- 8 neurons in the hidden layer
- Activation function: logistic for the hidden layer

  linear for the output layer
- Optimization algorithm – SGD
- Learning rate – 0.1
- Momentum – 0.5

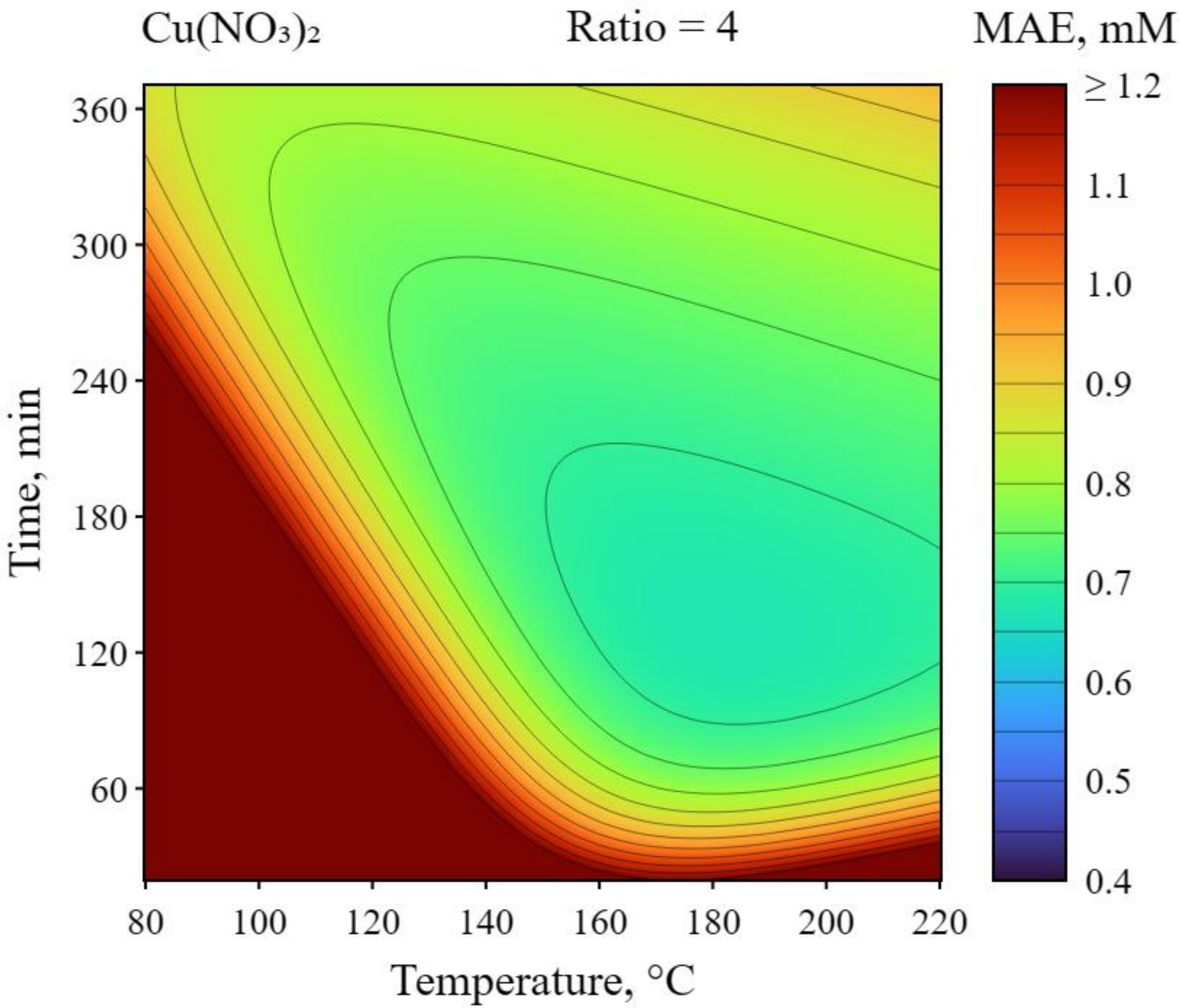
Stage 2 – approximation model



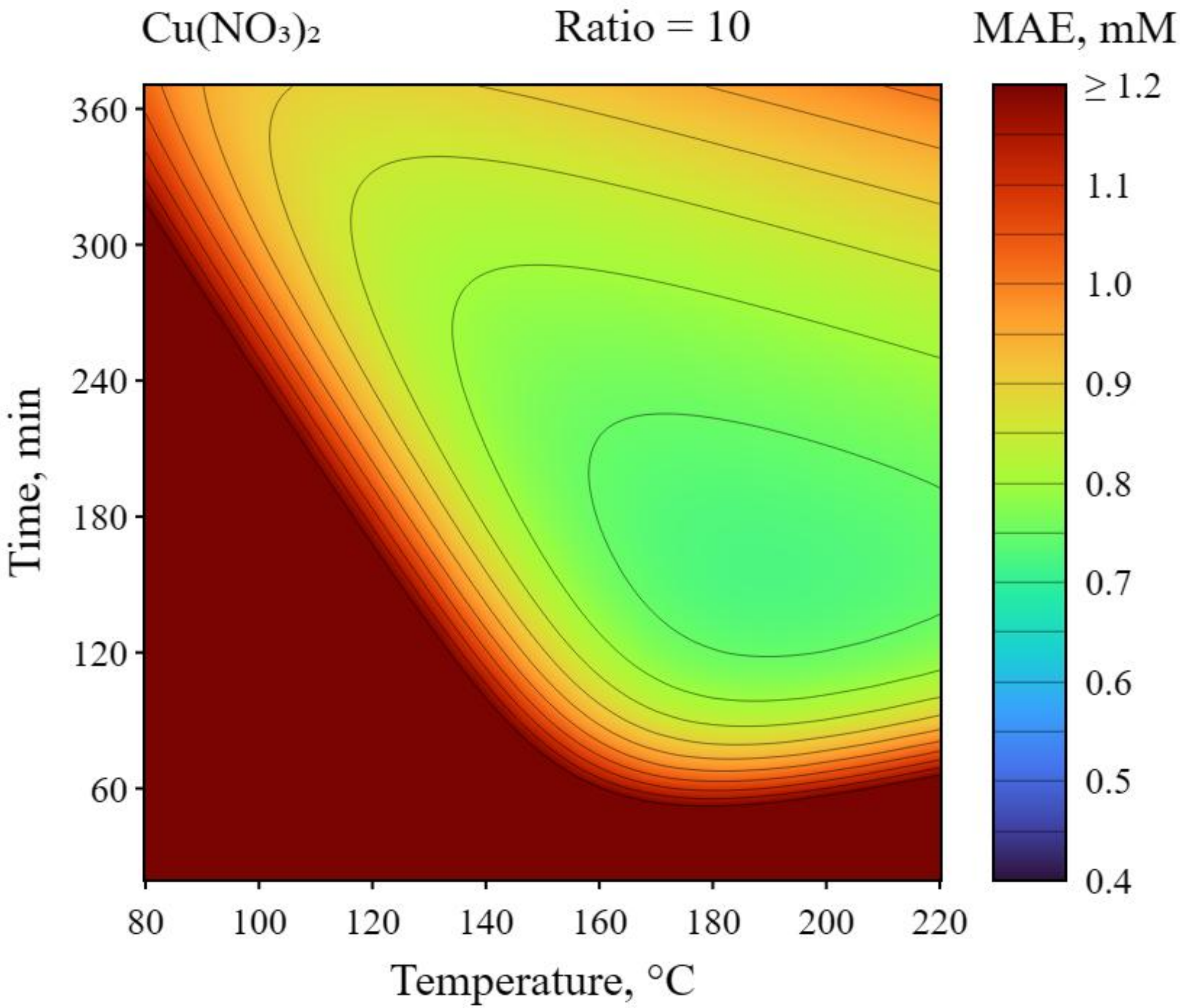
Stage 2 – approximation model



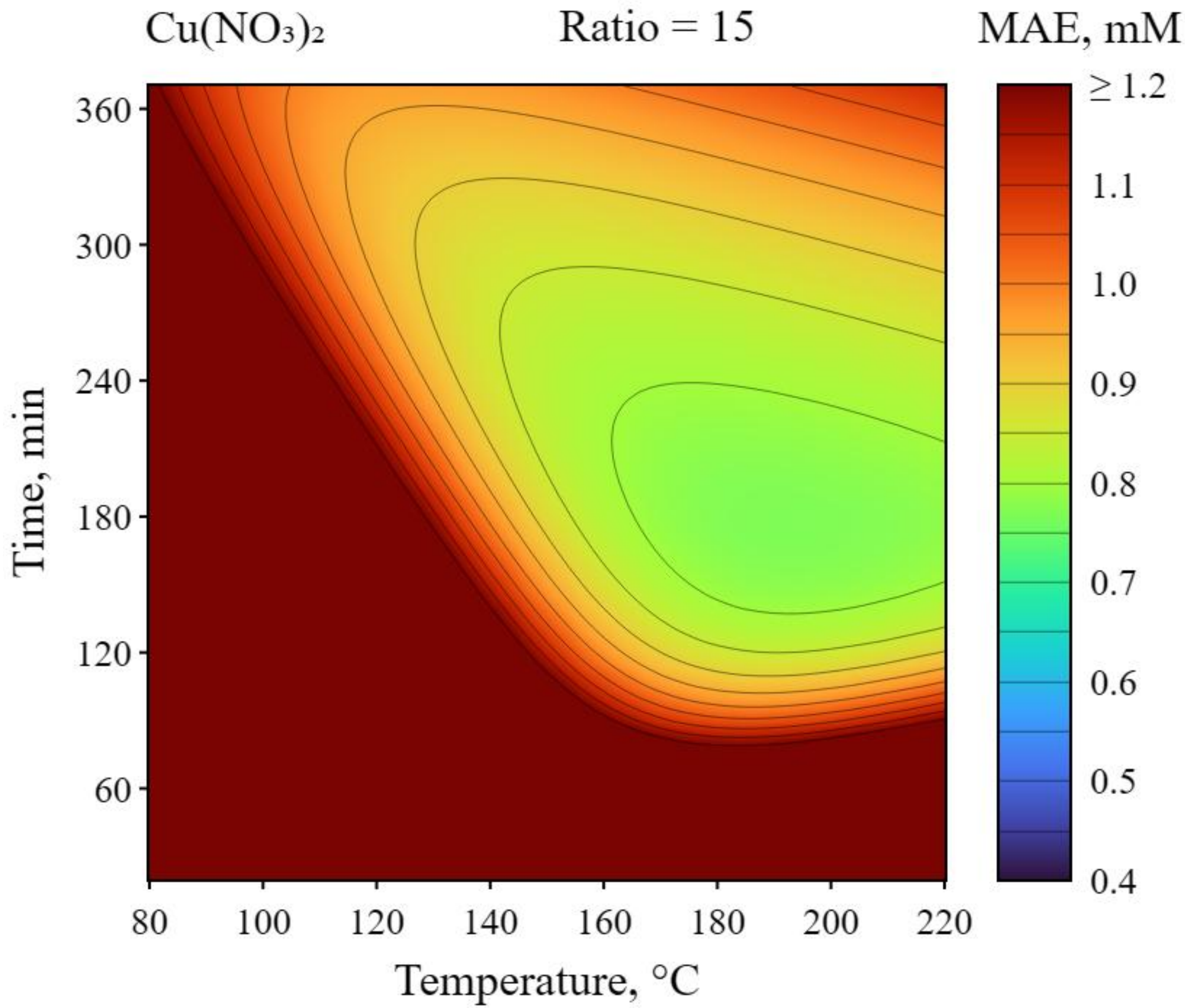
Stage 2 – approximation model



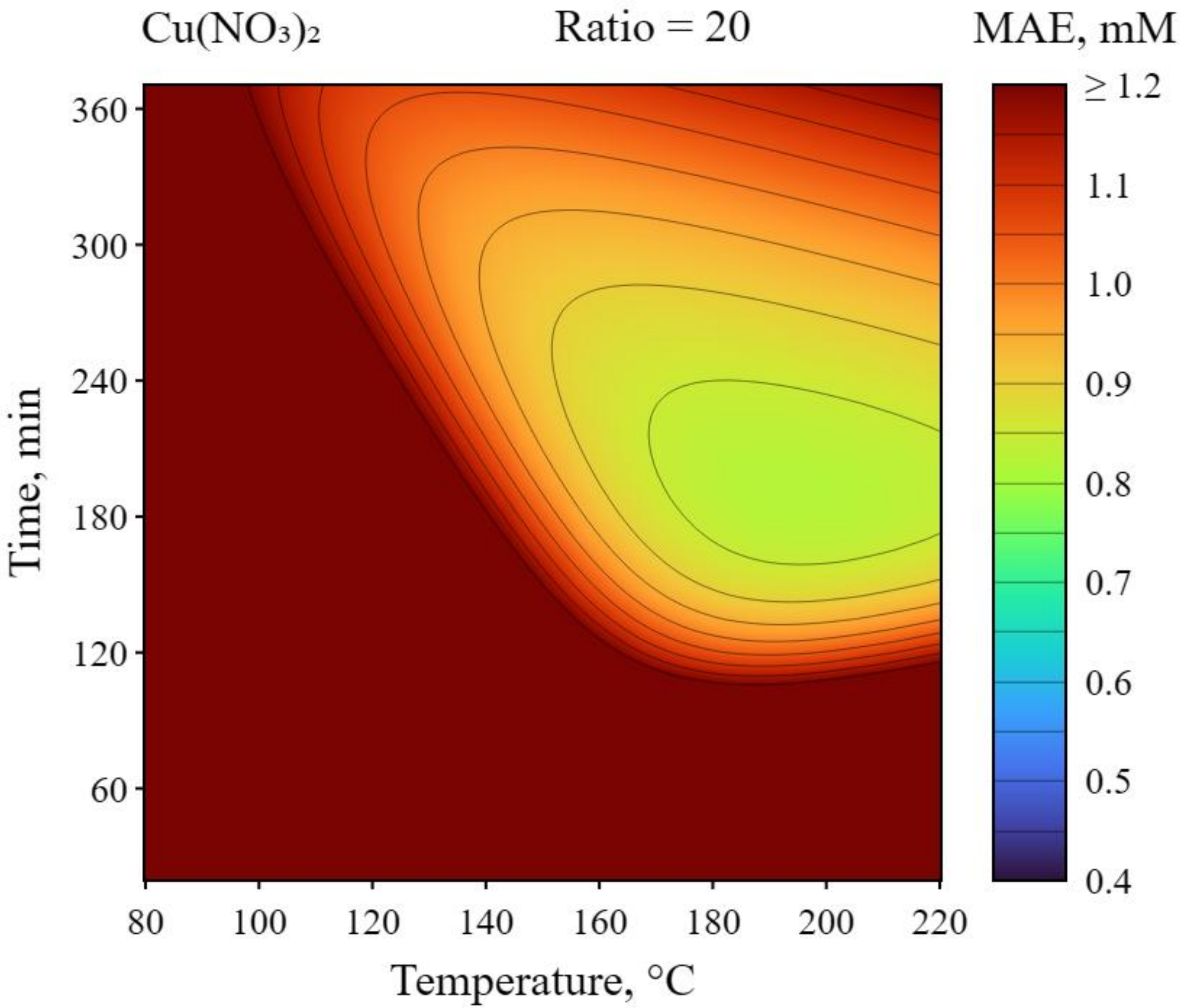
Stage 2 – approximation model



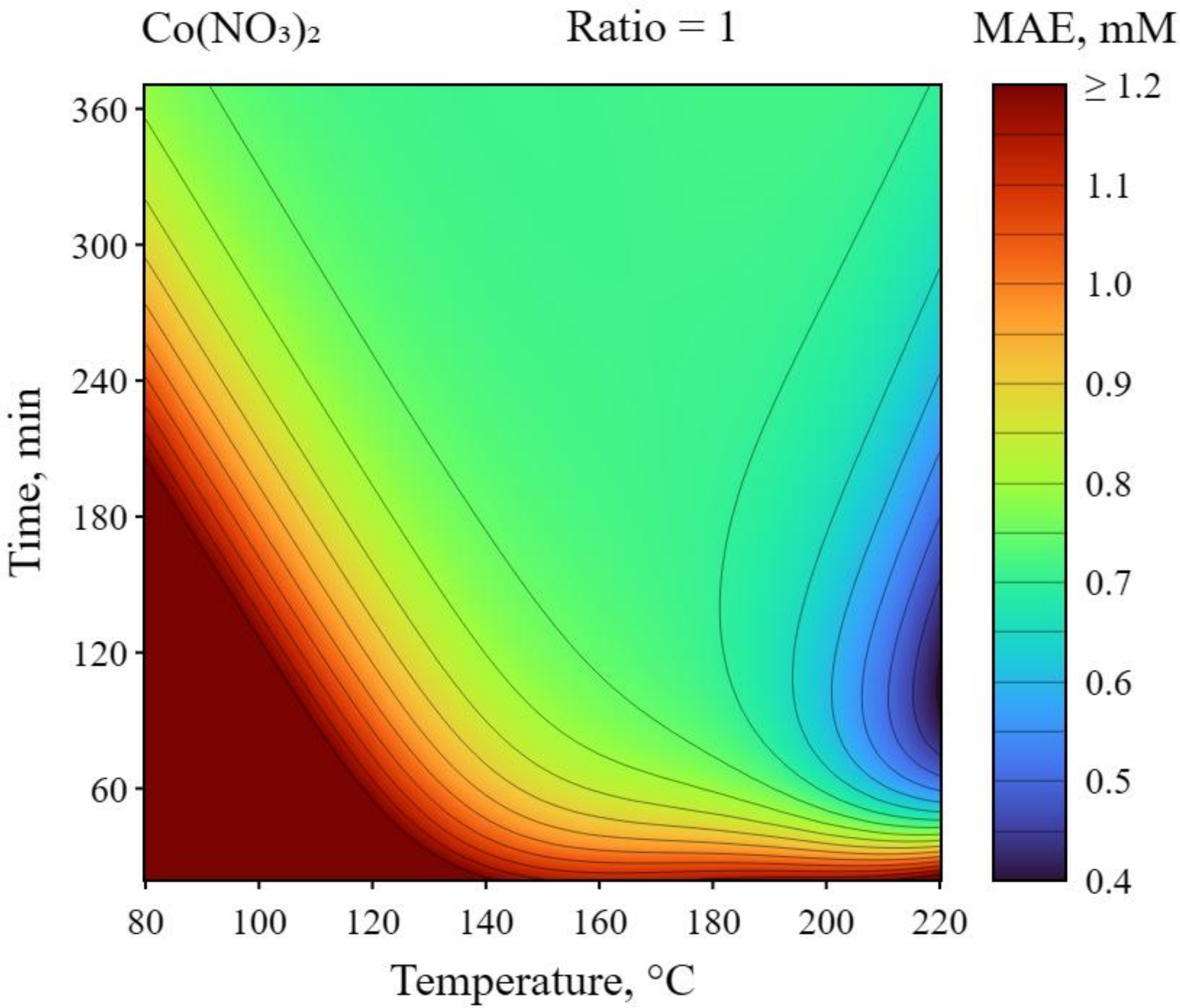
Stage 2 – approximation model



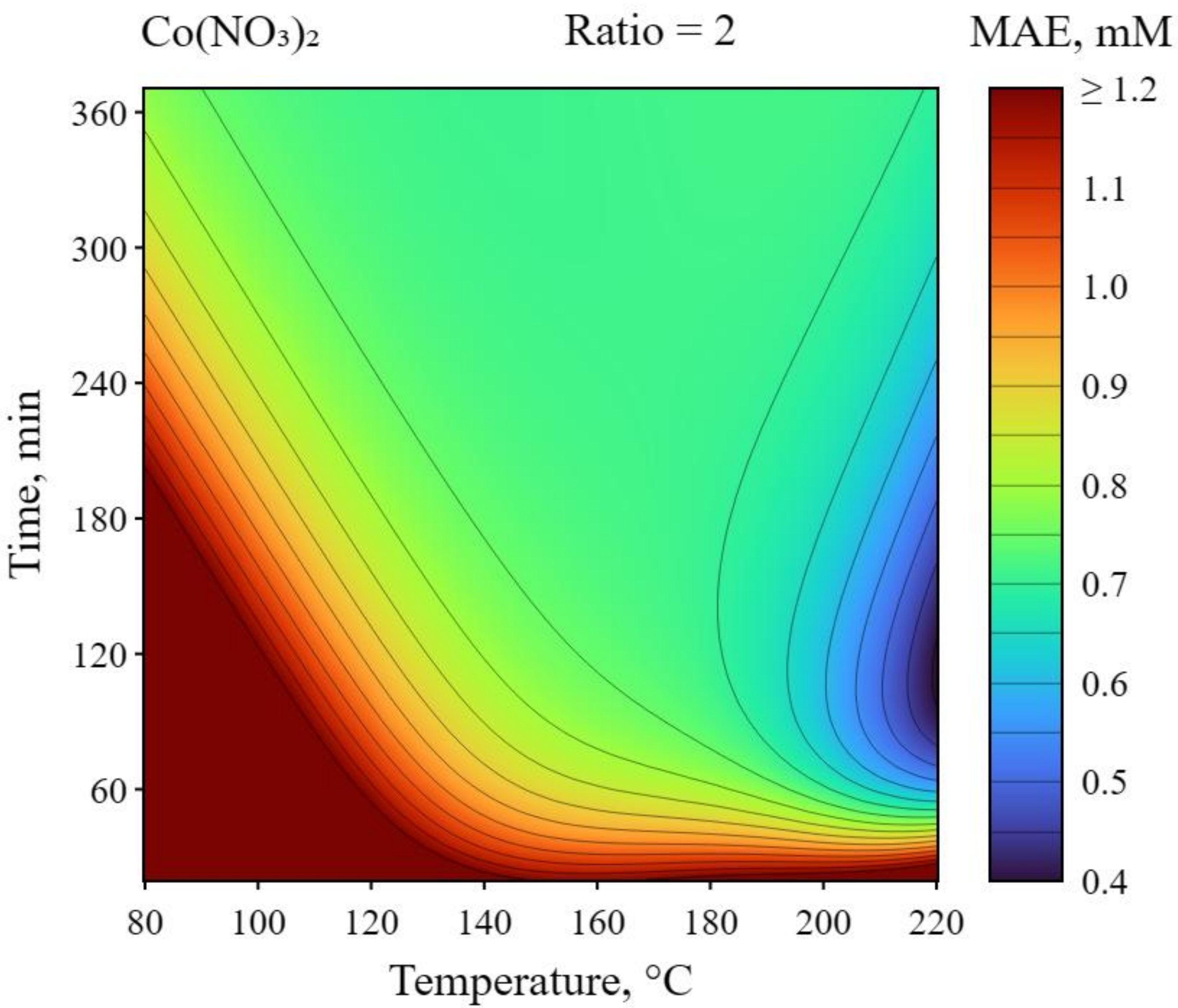
Stage 2 – approximation model



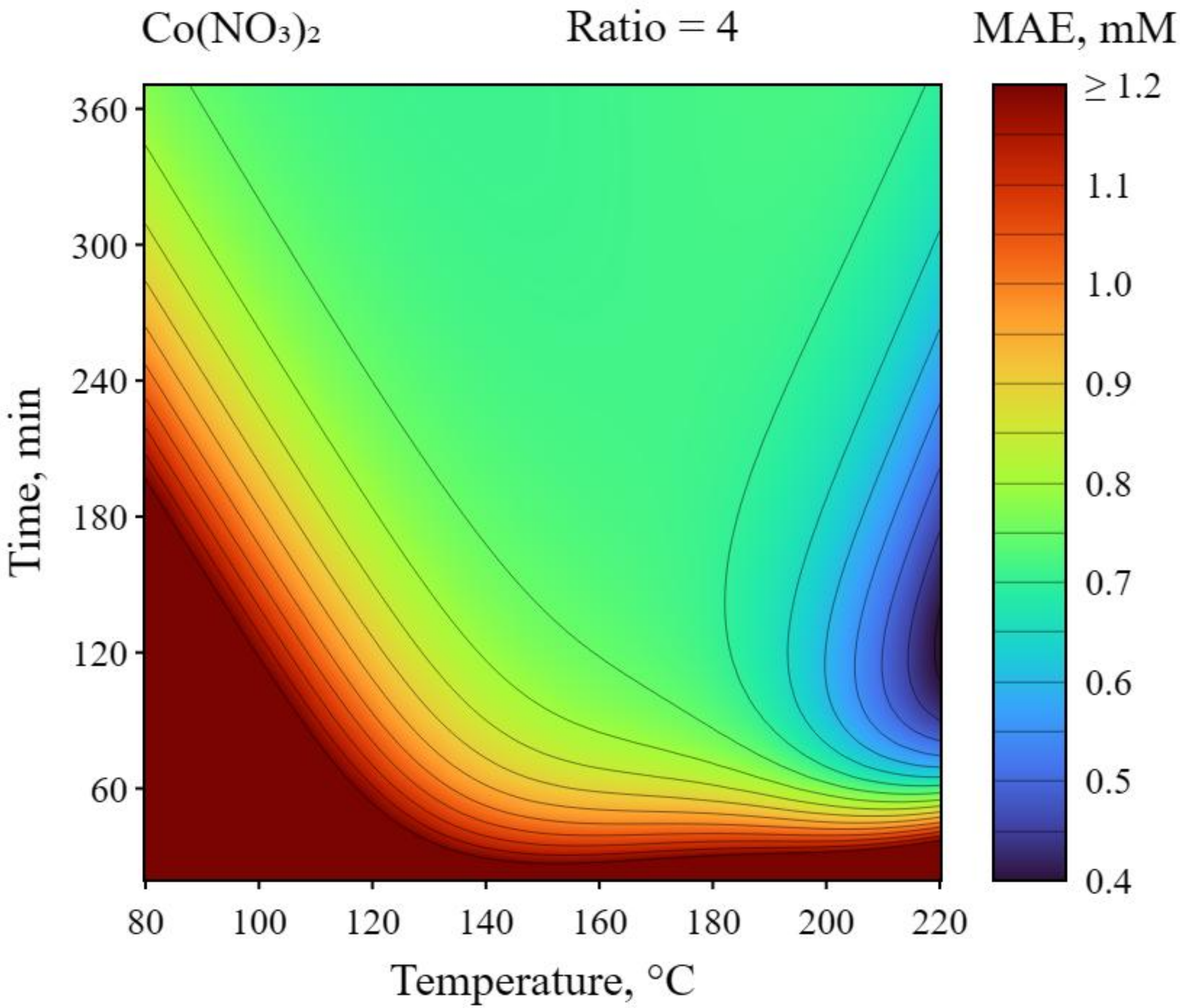
Stage 2 – approximation model



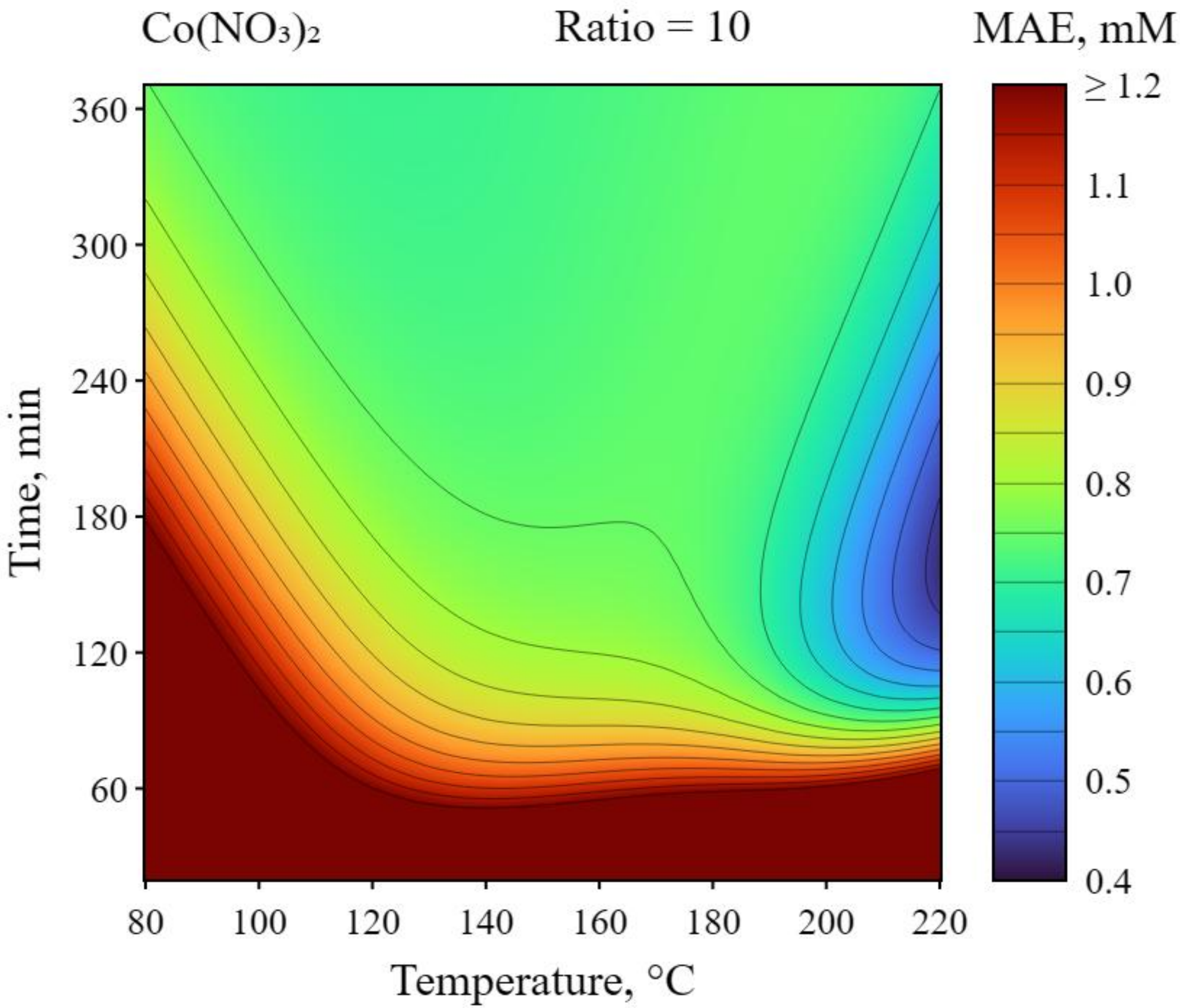
Stage 2 – approximation model



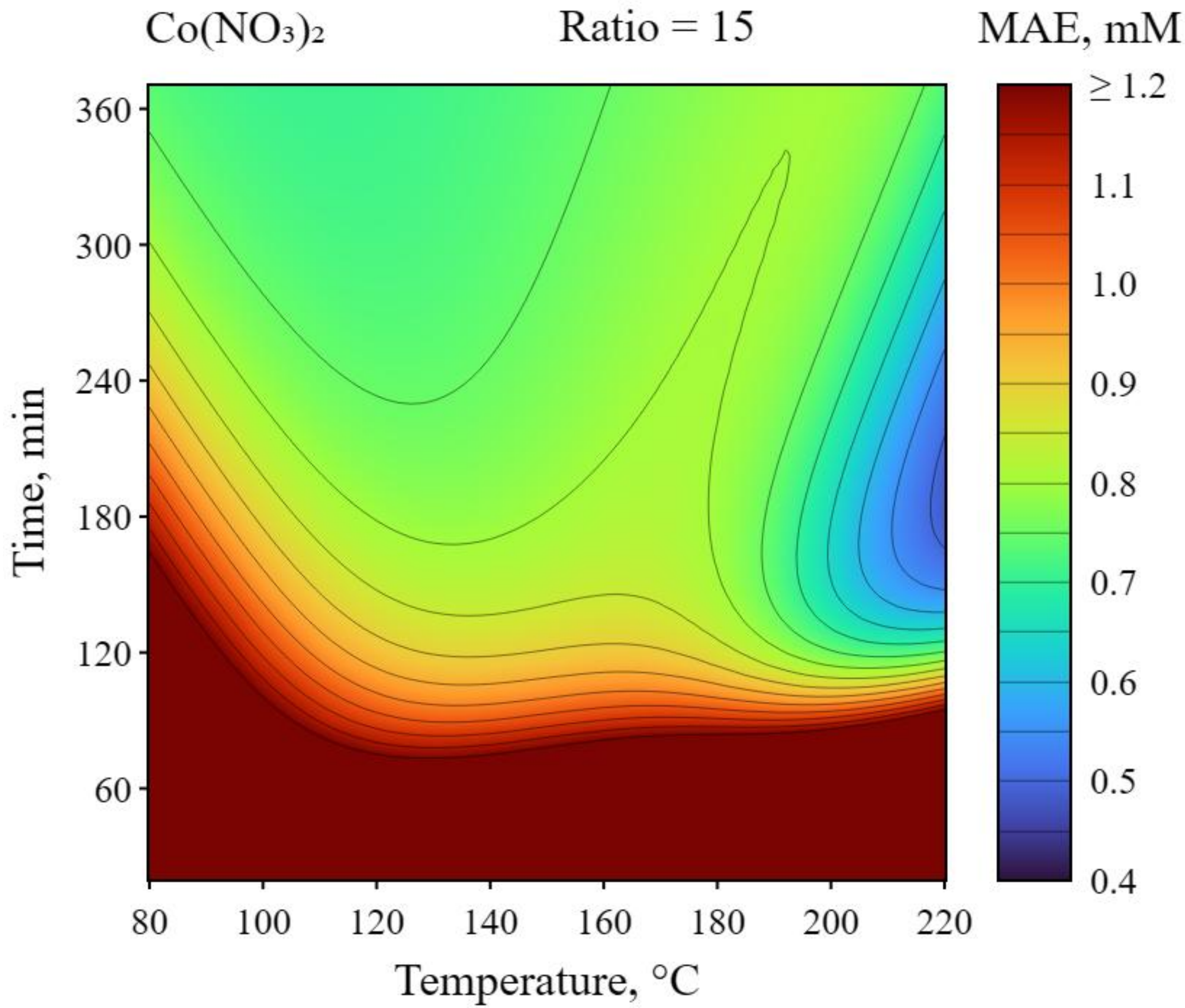
Stage 2 – approximation model



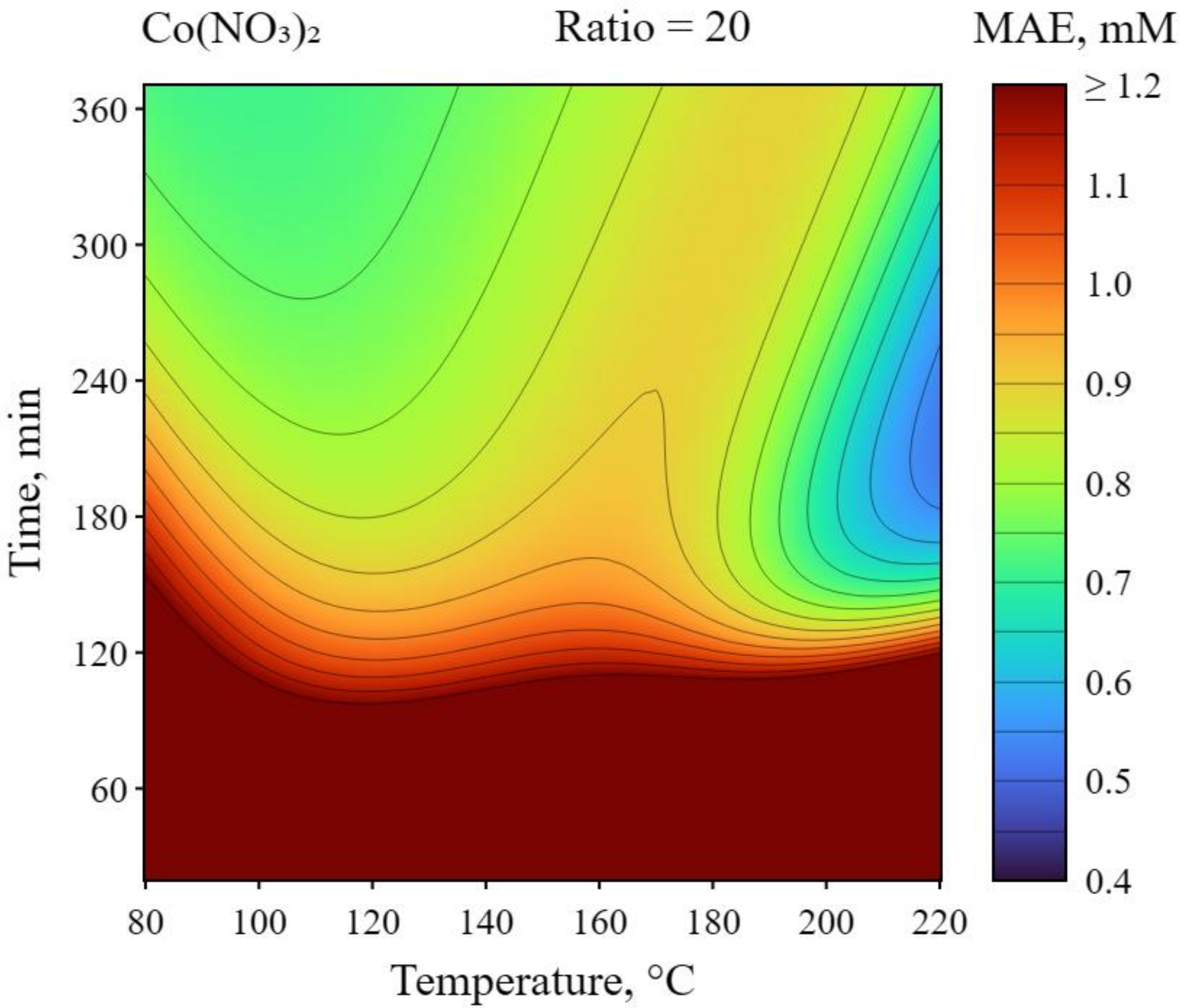
Stage 2 – approximation model



Stage 2 – approximation model



Stage 2 – approximation model



Conclusions

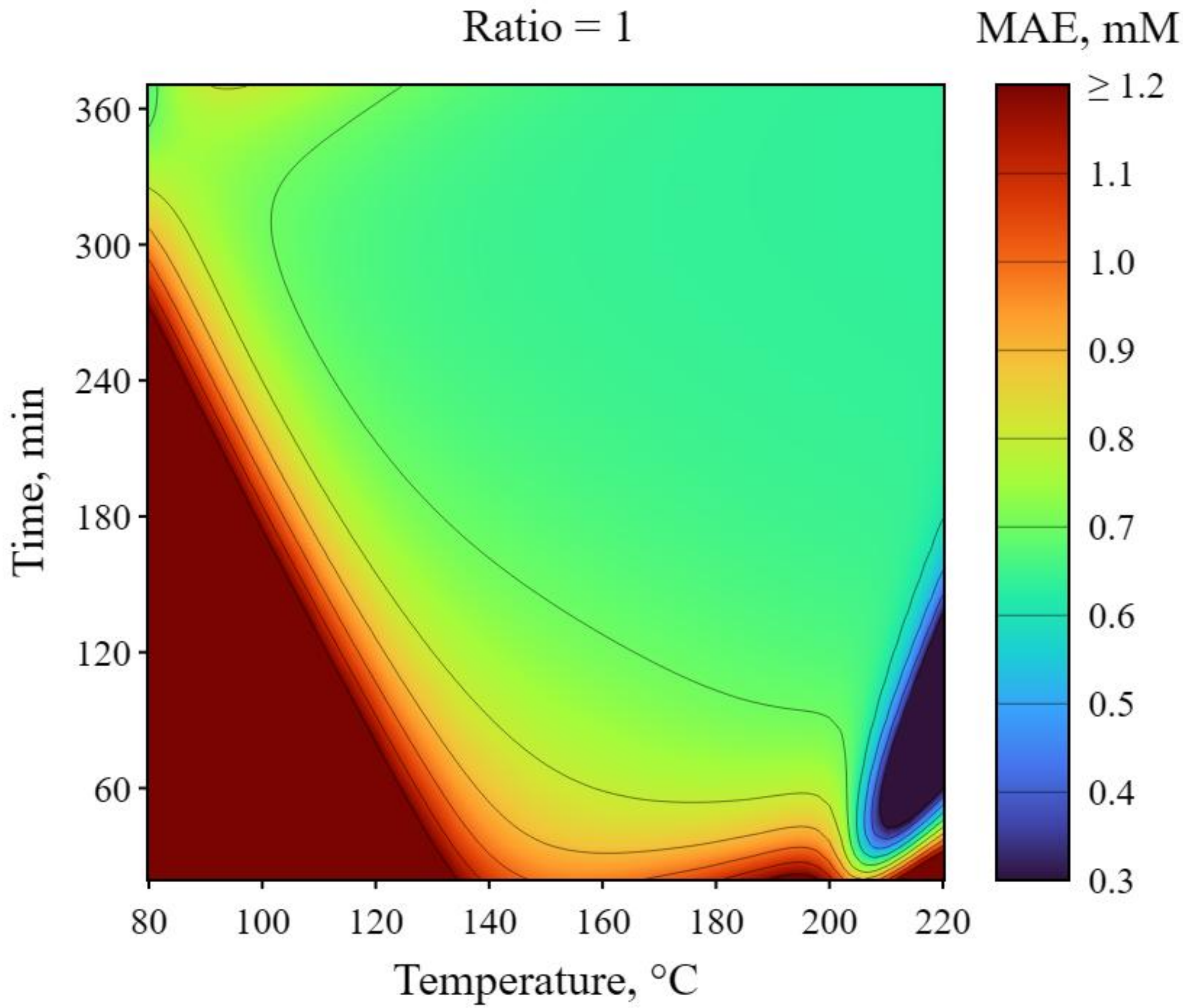
- To estimate the value of the error in determining the concentration of specific heavy metal salts in aqueous solutions at given synthesis parameters of carbon dots, a neural network approximator was created and optimized

- It was shown that it is possible to successfully determine the areas of optimal synthesis parameters of carbon dots that ensure the necessary accuracy in determining the concentrations of various components in multicomponent solutions

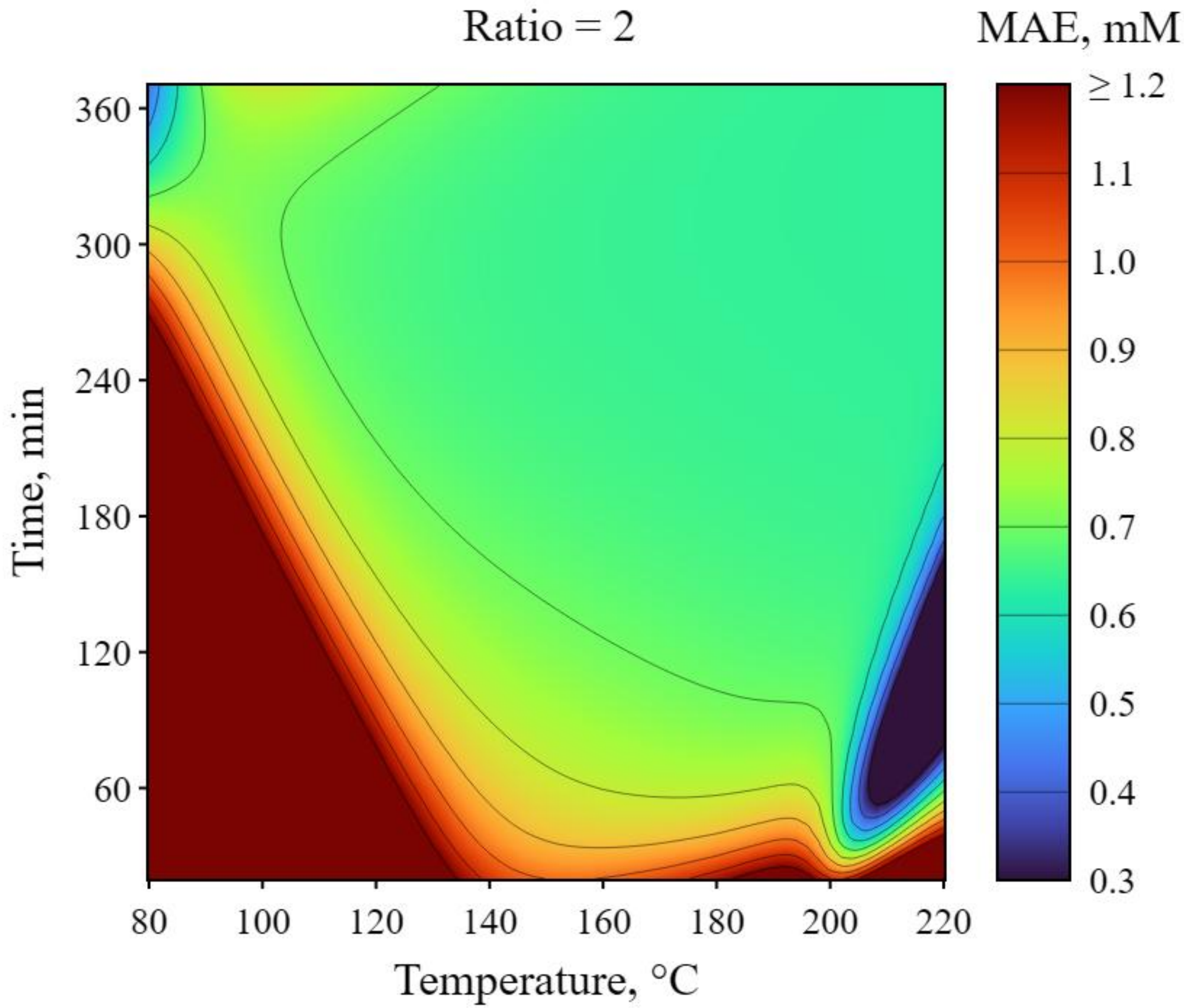
- Further improvement of the quality and stability of the obtained approximation model is required

Thank you for your attention

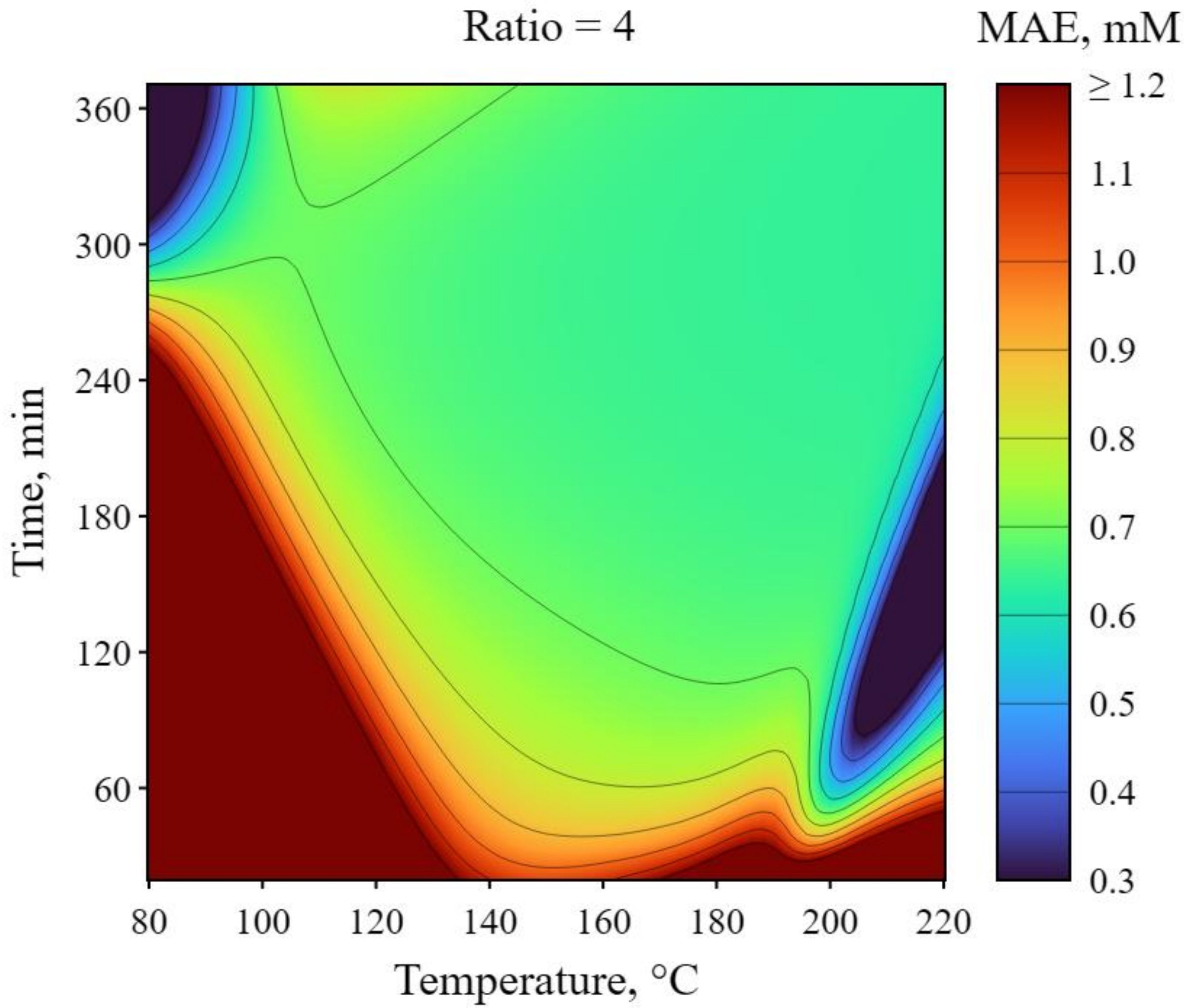
Stage 2 – approximation model – $Co(NO_3)_2$



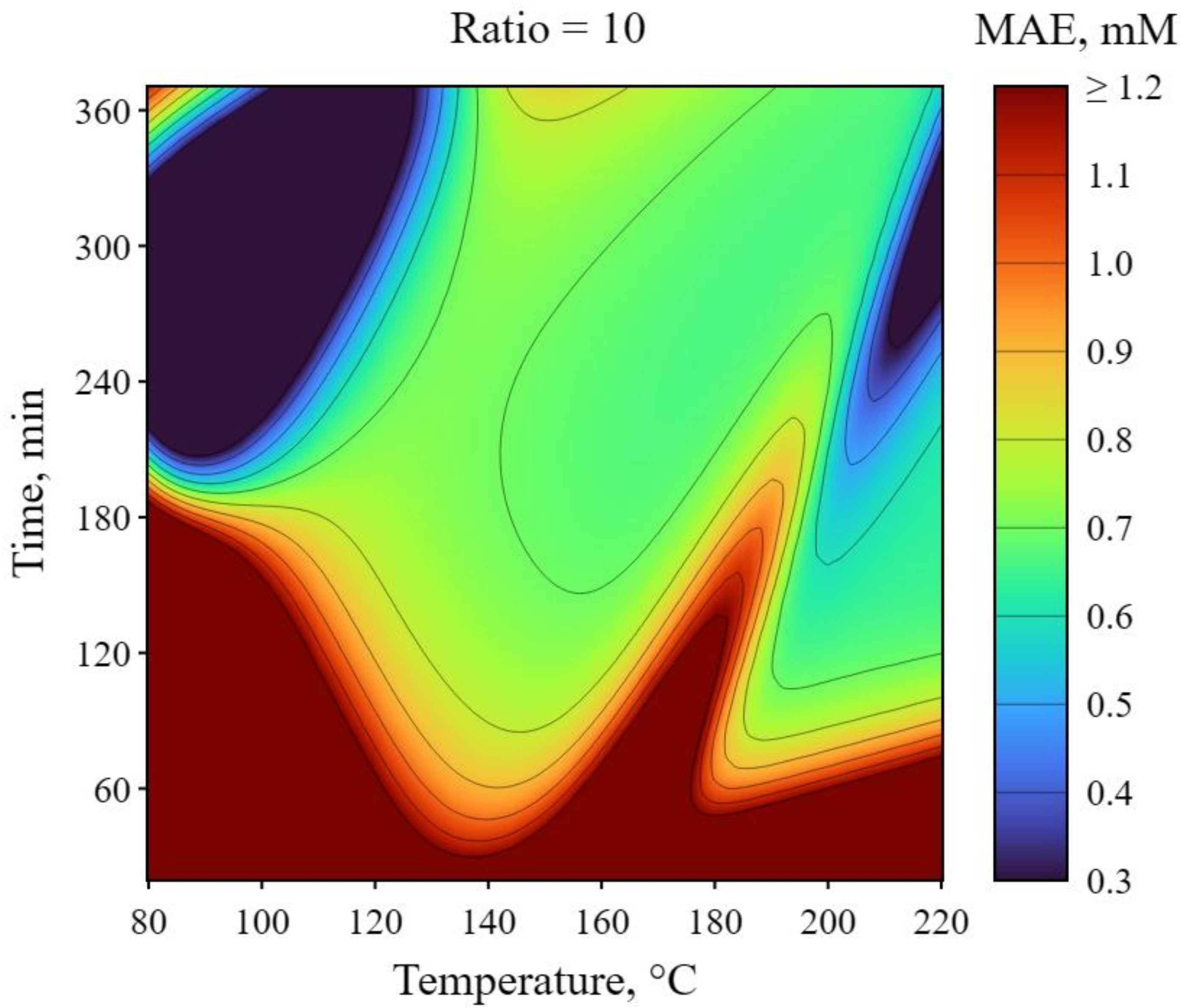
Stage 2 – approximation model – $Co(NO_3)_2$



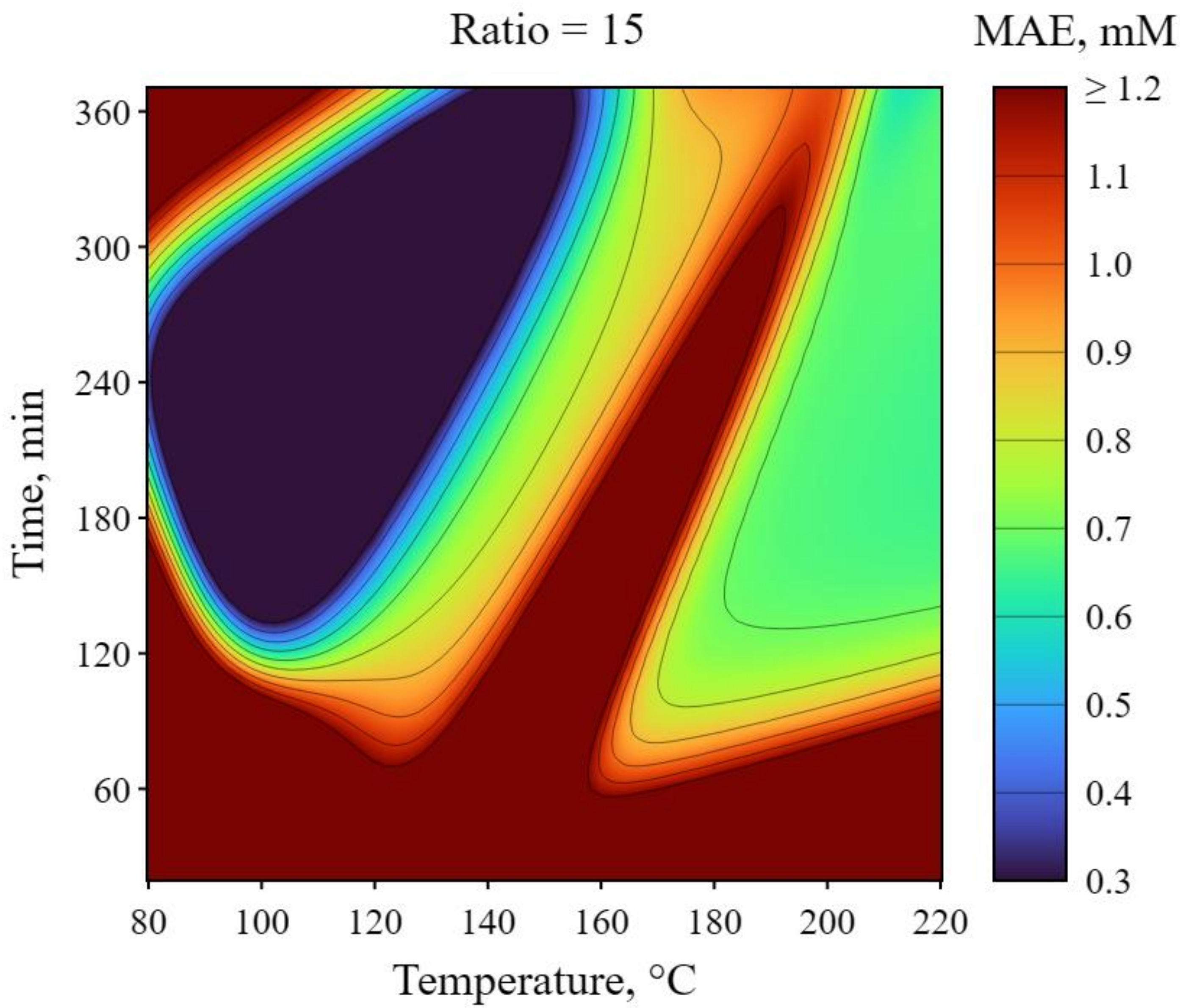
Stage 2 – approximation model – $Co(NO_3)_2$



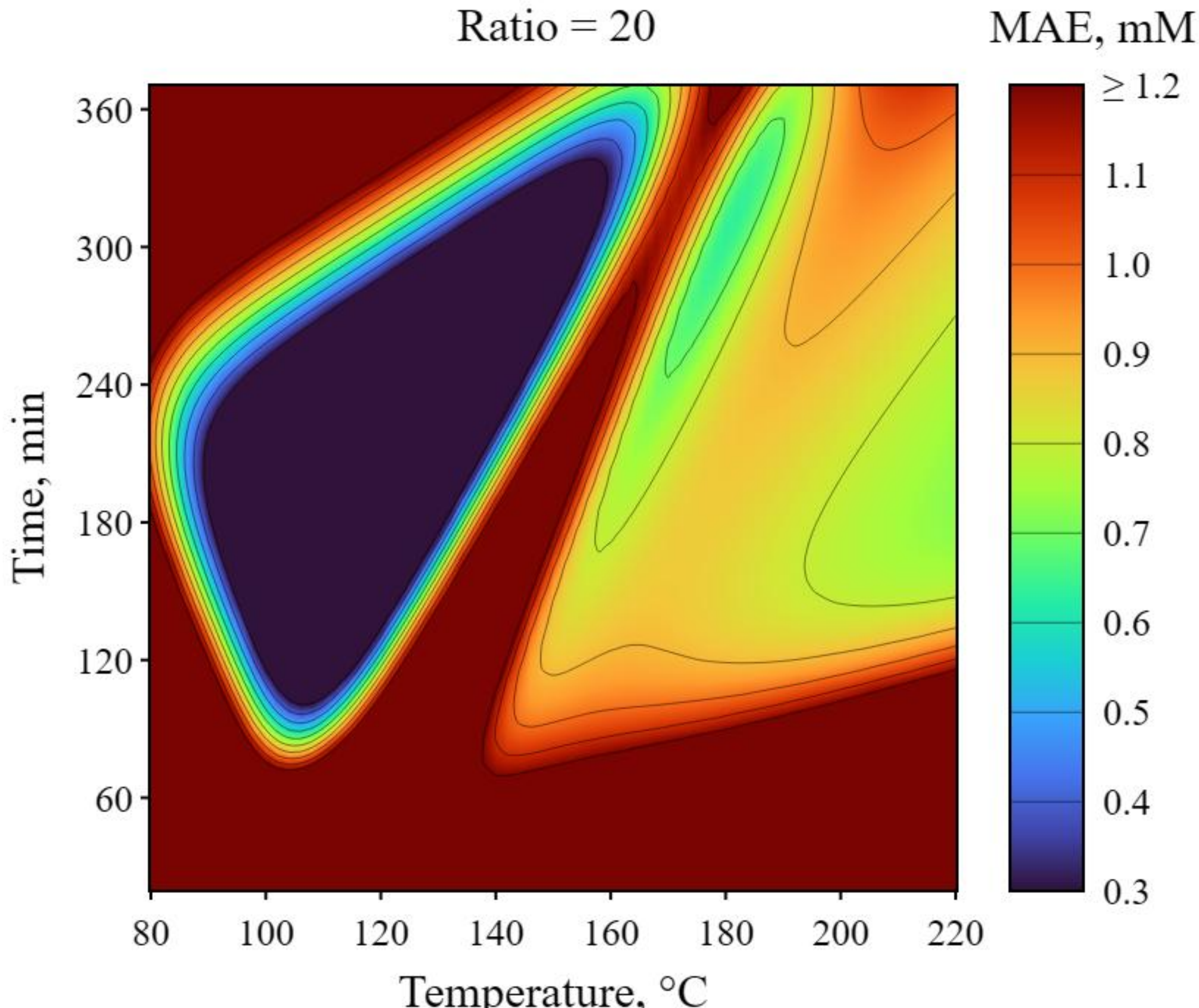
Stage 2 – approximation model – $Co(NO_3)_2$



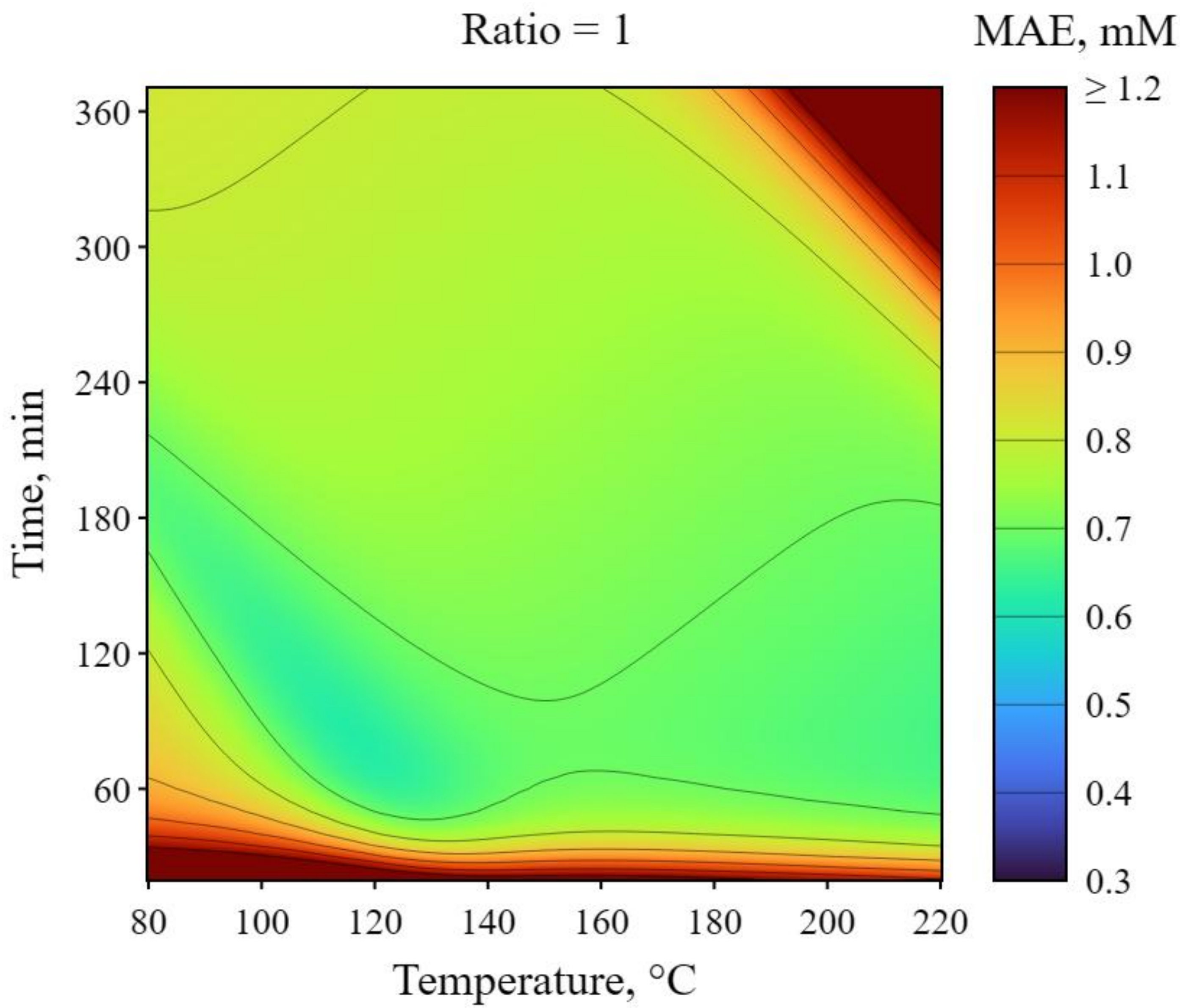
Stage 2 – approximation model – $Co(NO_3)_2$



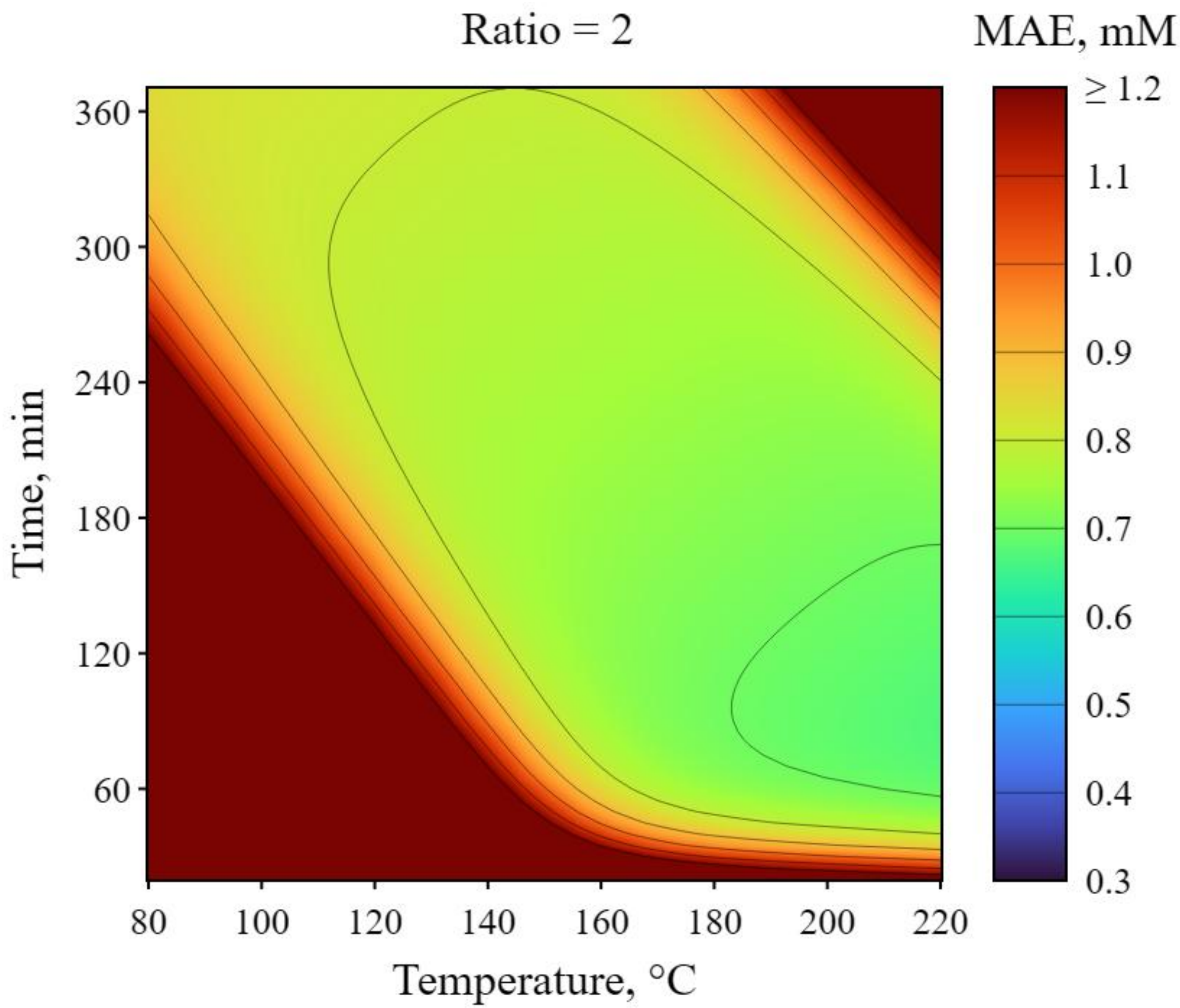
Stage 2 – approximation model – $Co(NO_3)_2$



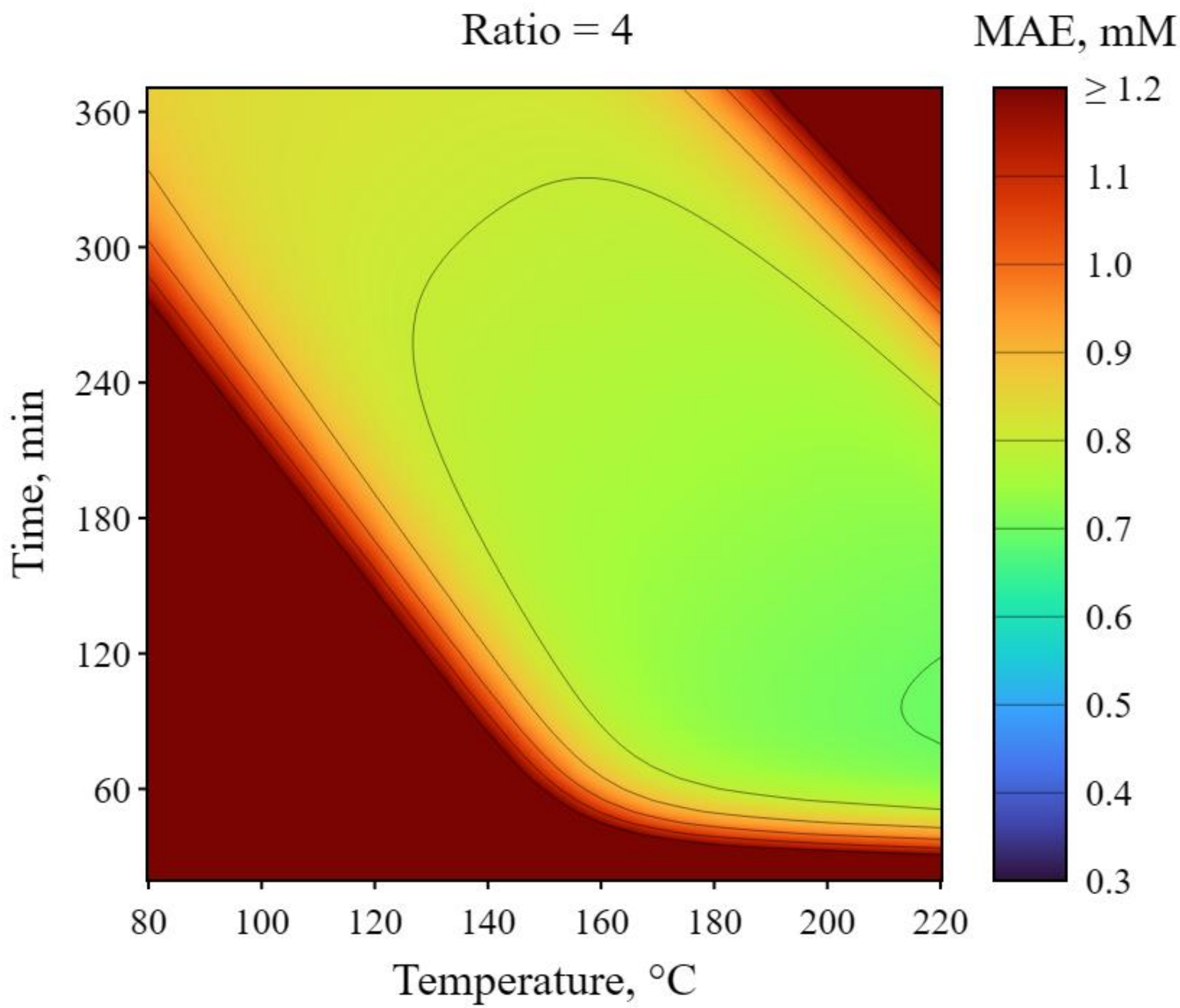
Stage 2 – approximation model – $Cu(NO_3)_2$



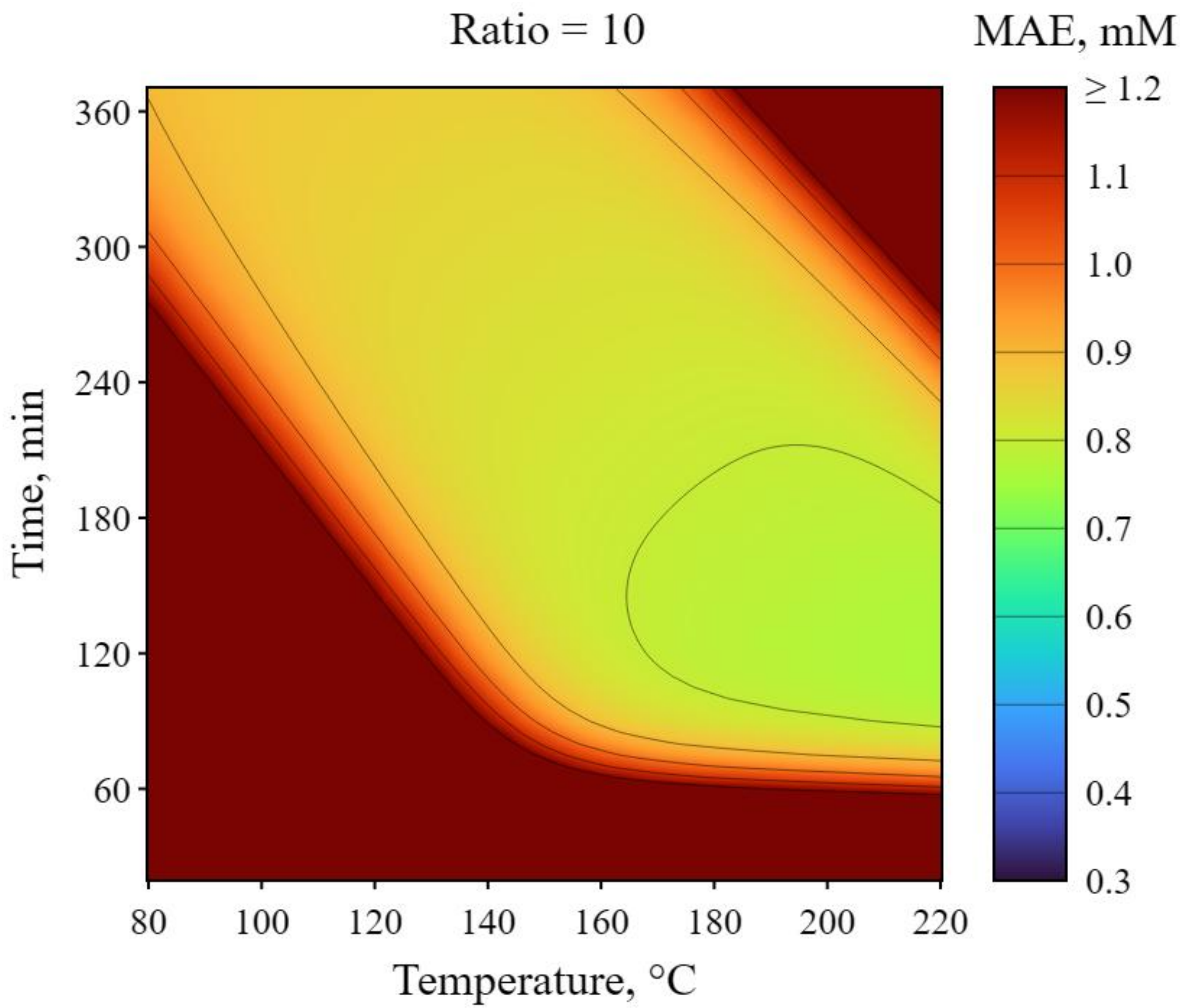
Stage 2 – approximation model – $Cu(NO_3)_2$



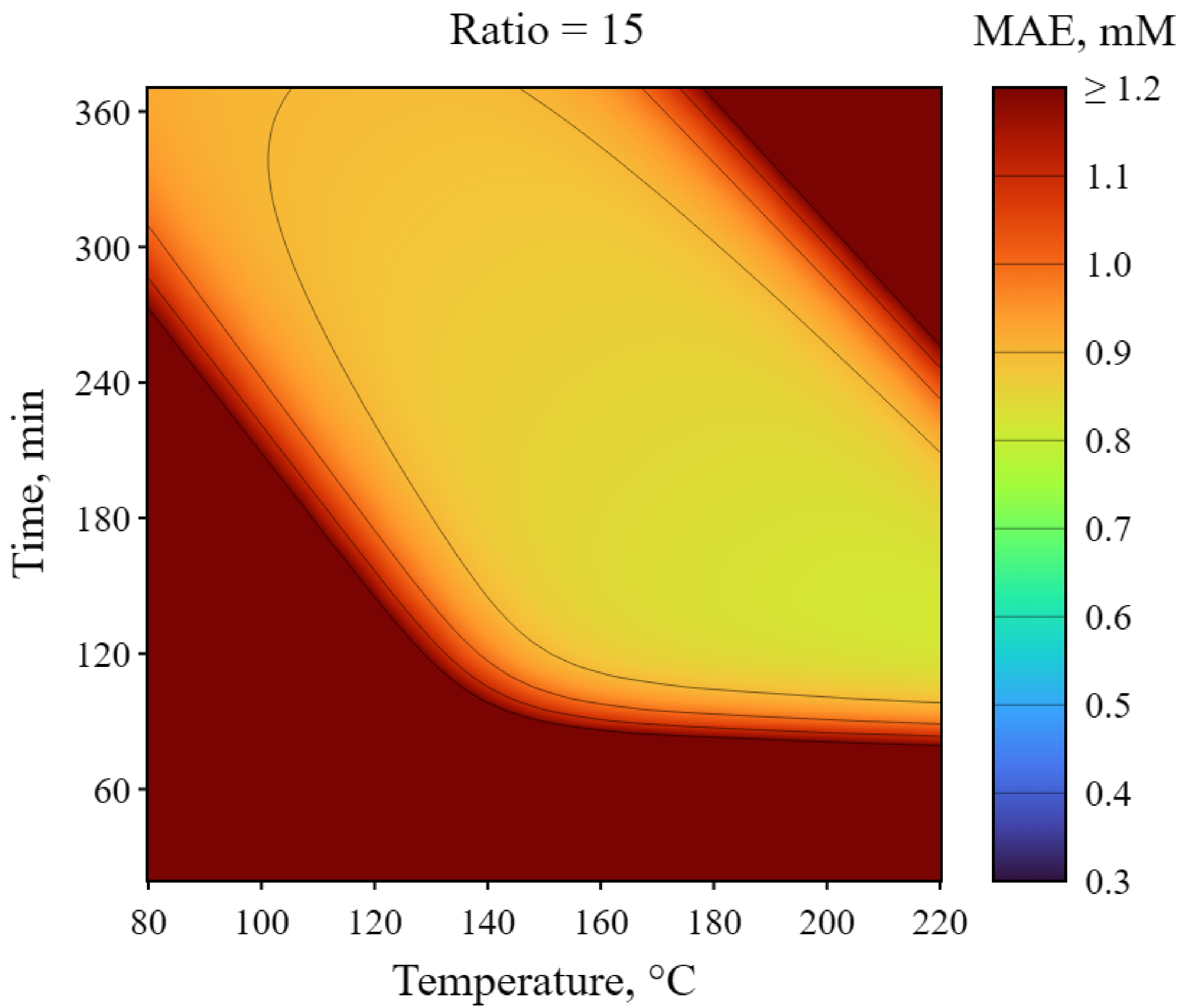
Stage 2 – approximation model – $Cu(NO_3)_2$



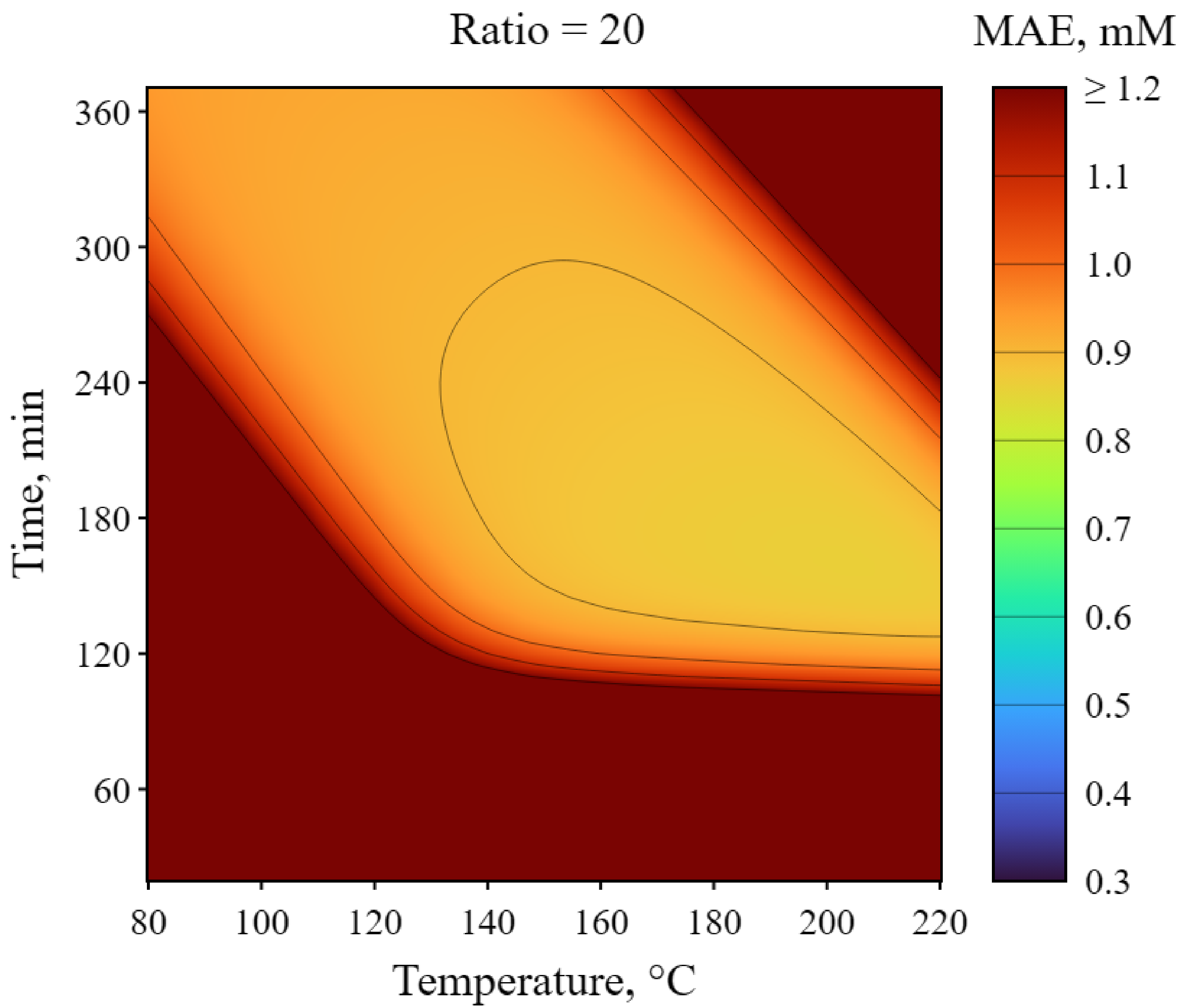
Stage 2 – approximation model – $Cu(NO_3)_2$



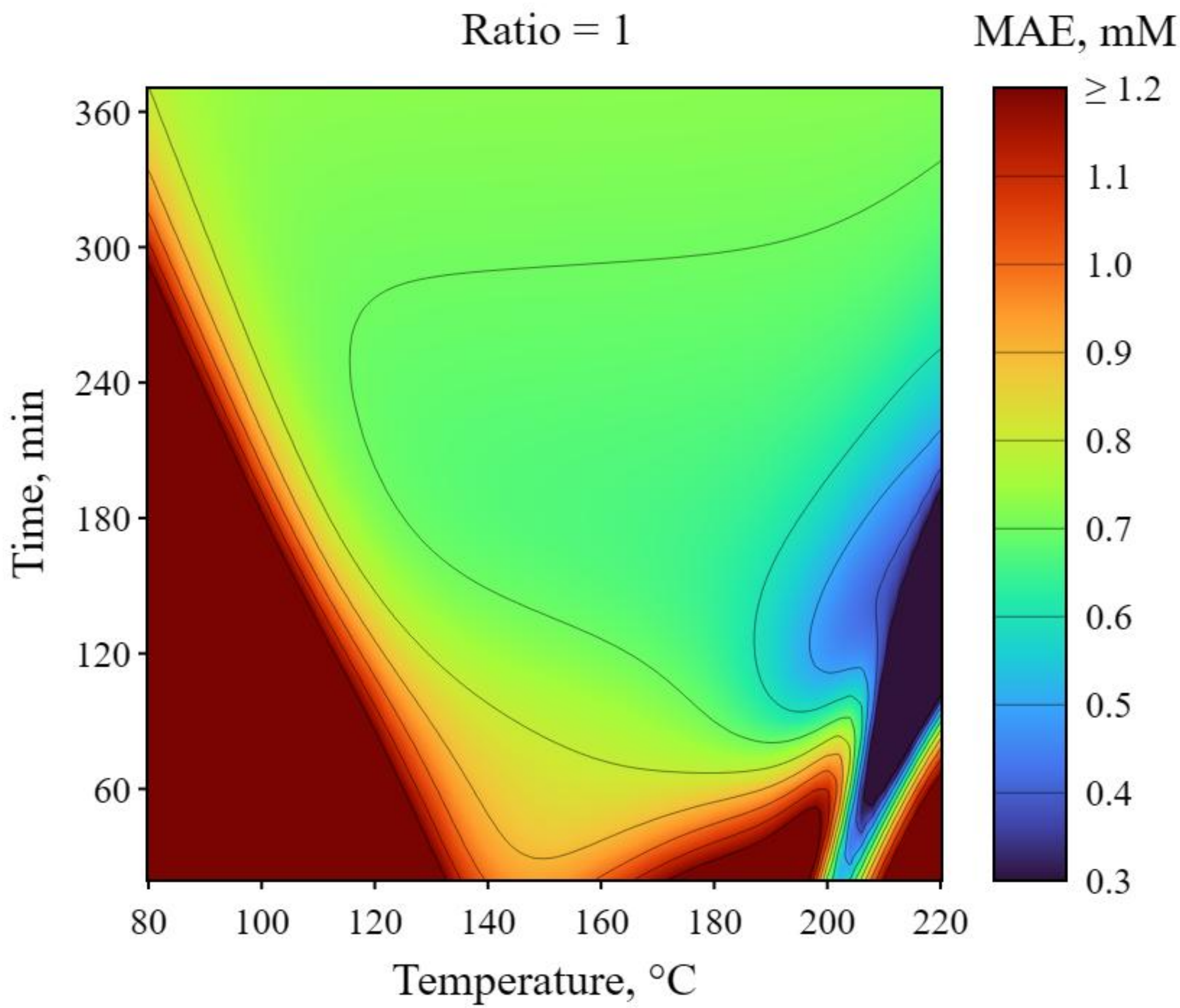
Stage 2 – approximation model – $Cu(NO_3)_2$



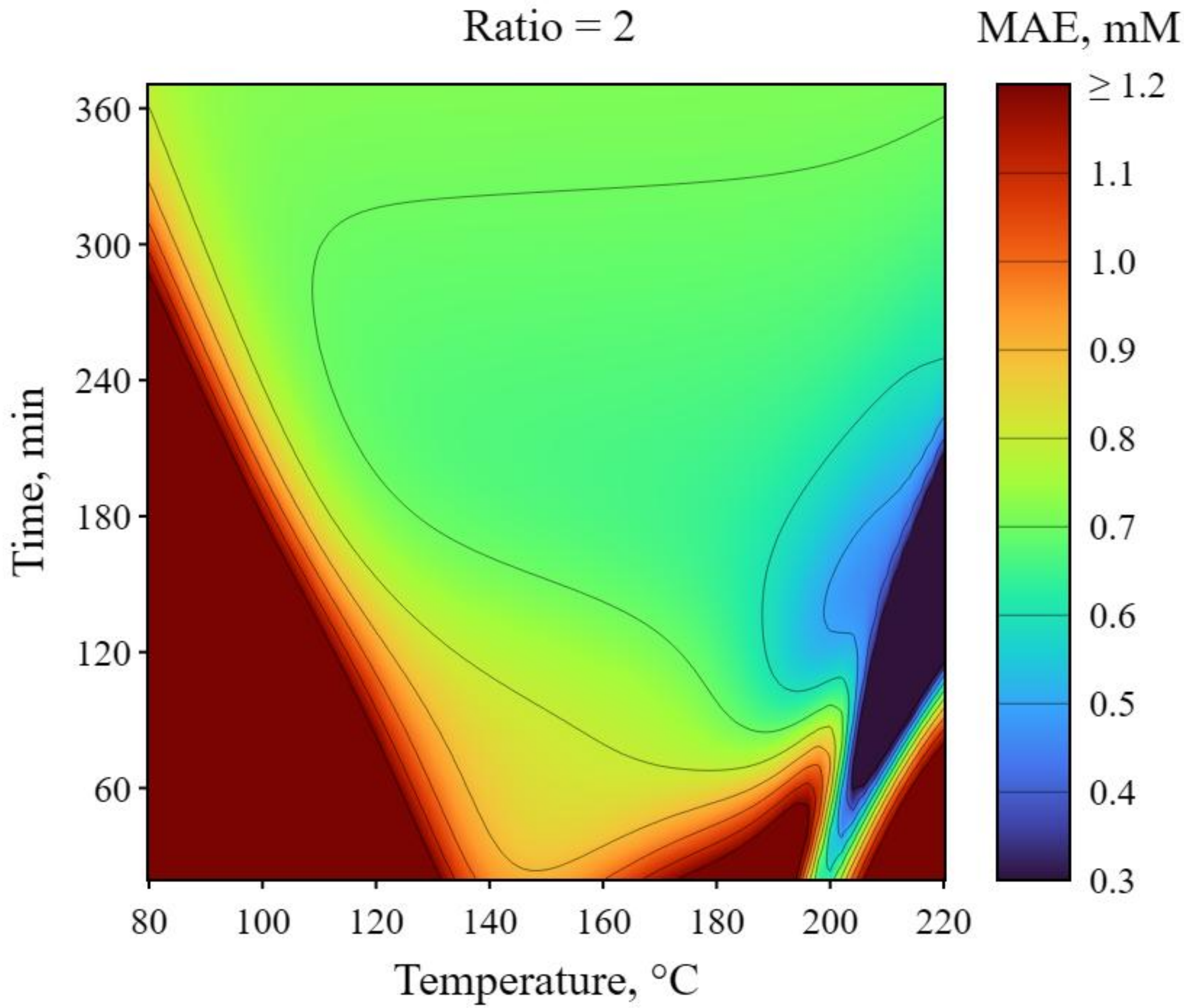
Stage 2 – approximation model – $Cu(NO_3)_2$



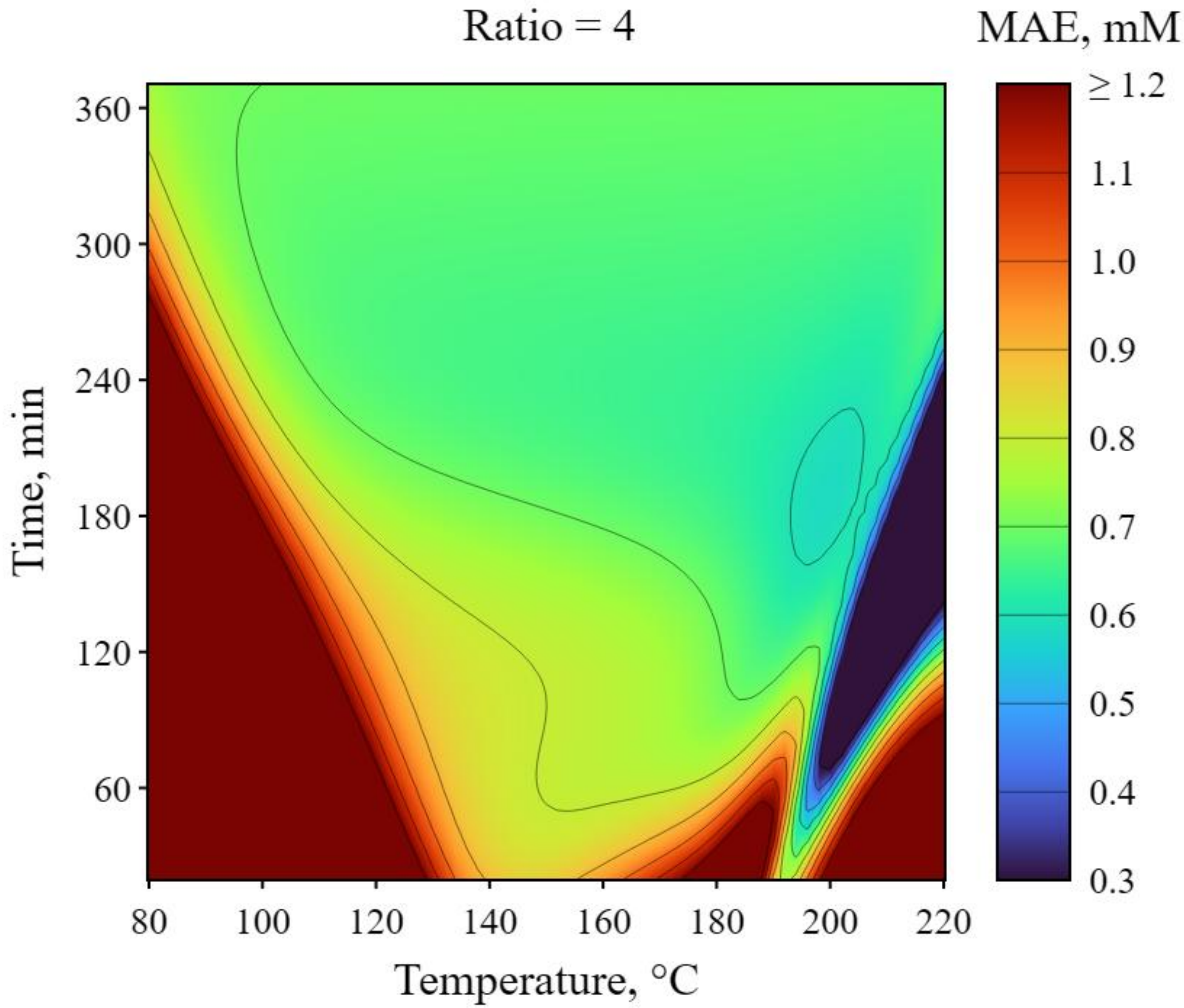
Stage 2 – approximation model – $Co(NO_3)_2$



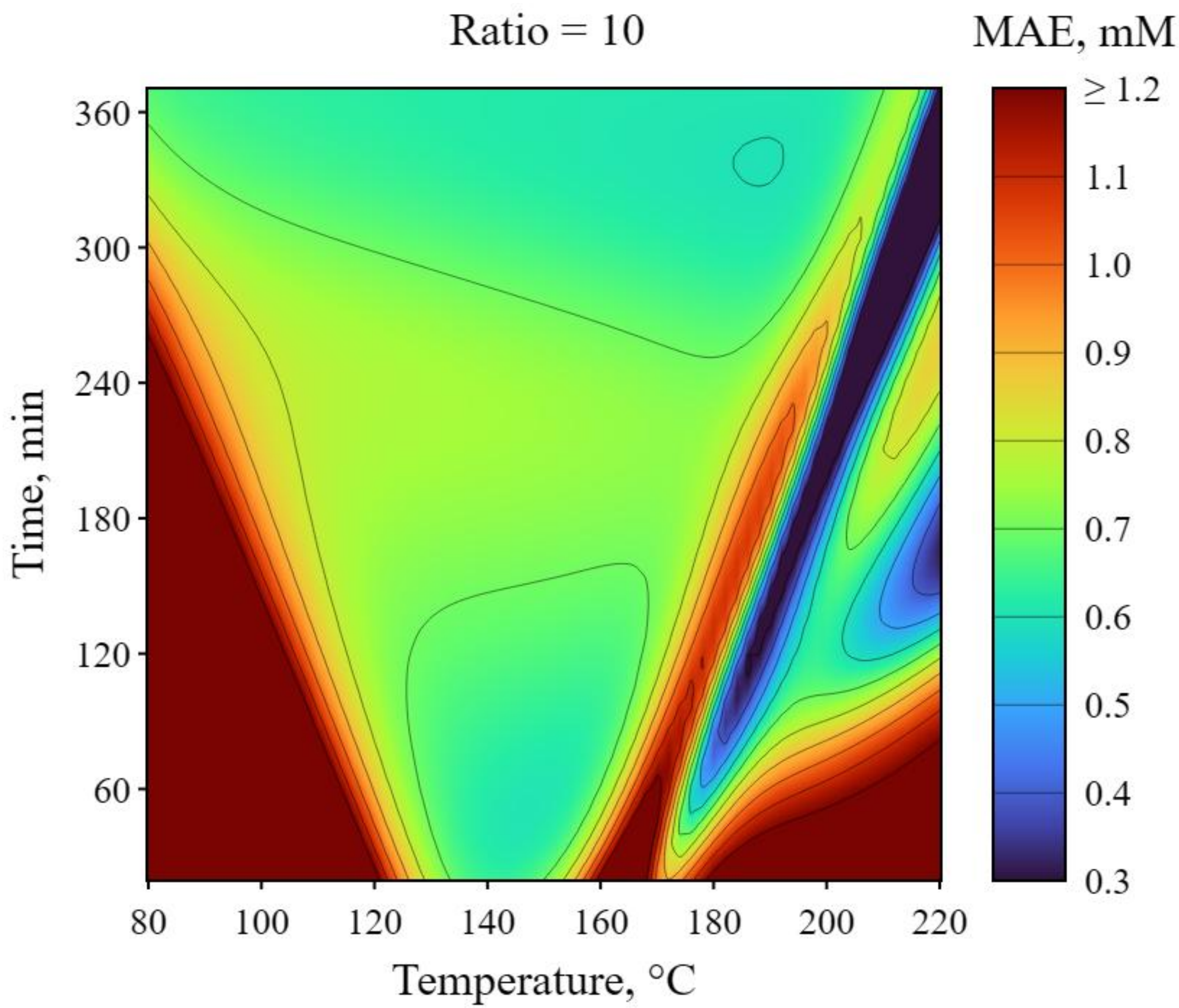
Stage 2 – approximation model – $Co(NO_3)_2$



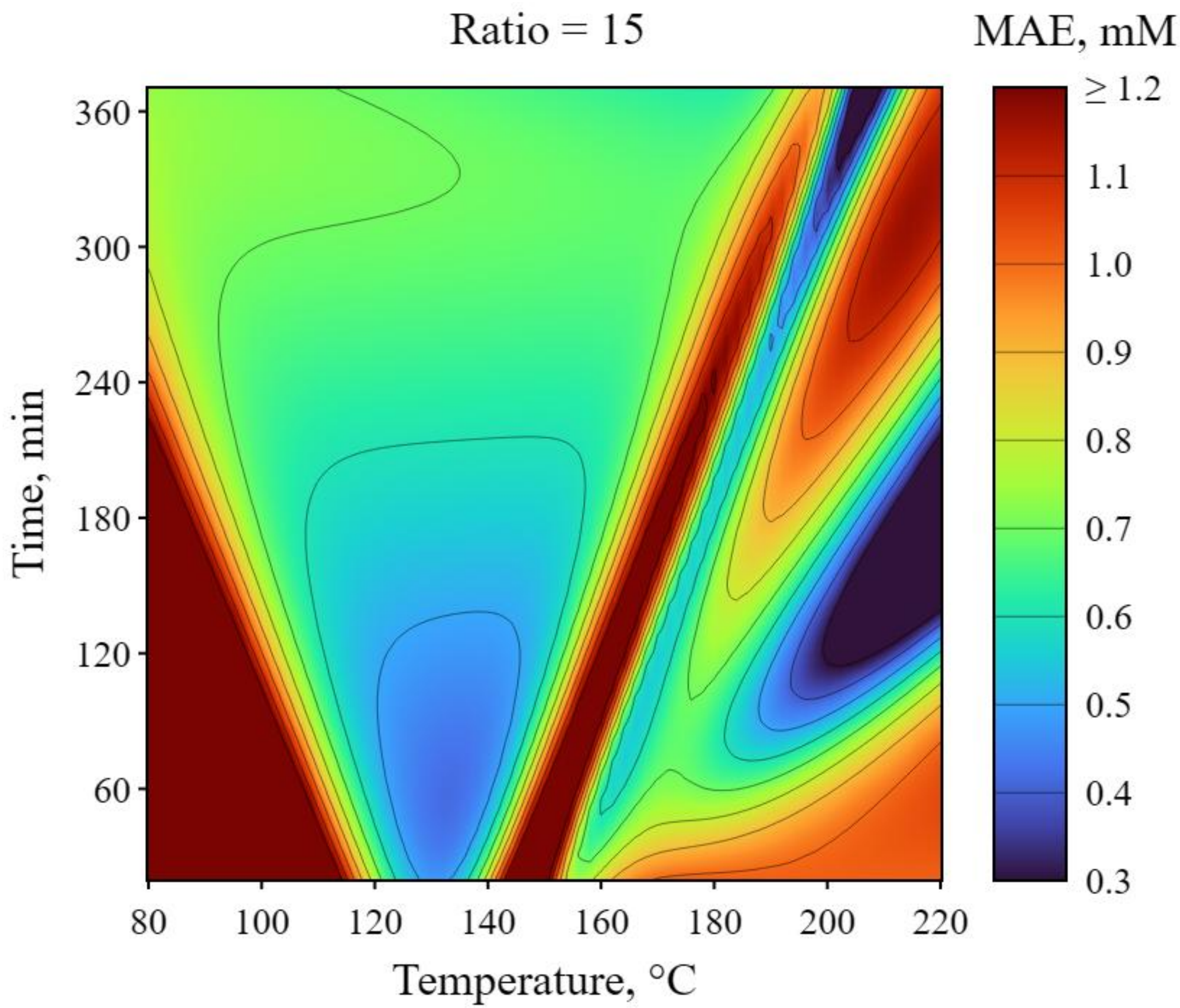
Stage 2 – approximation model – $Co(NO_3)_2$



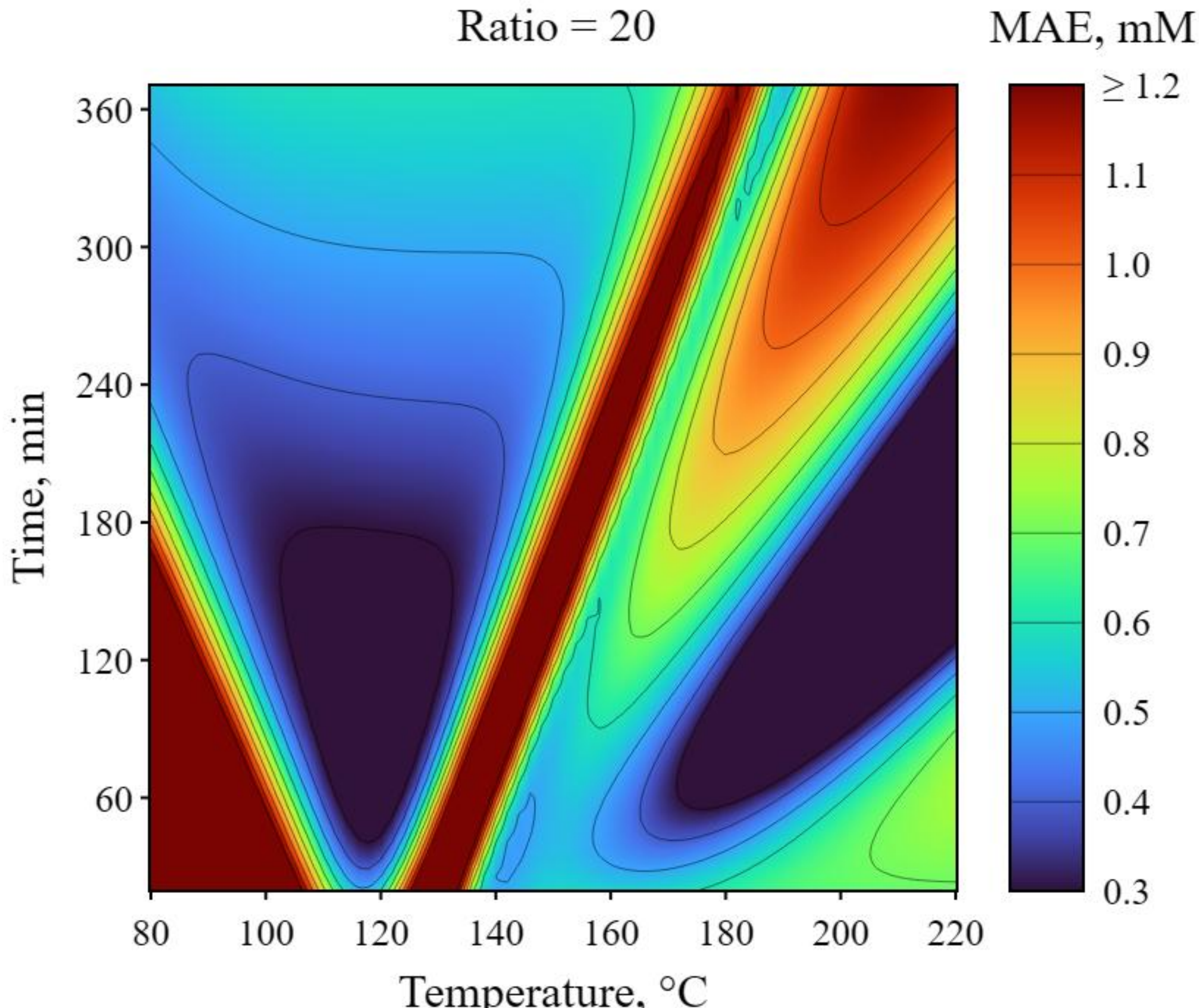
Stage 2 – approximation model – $Co(NO_3)_2$



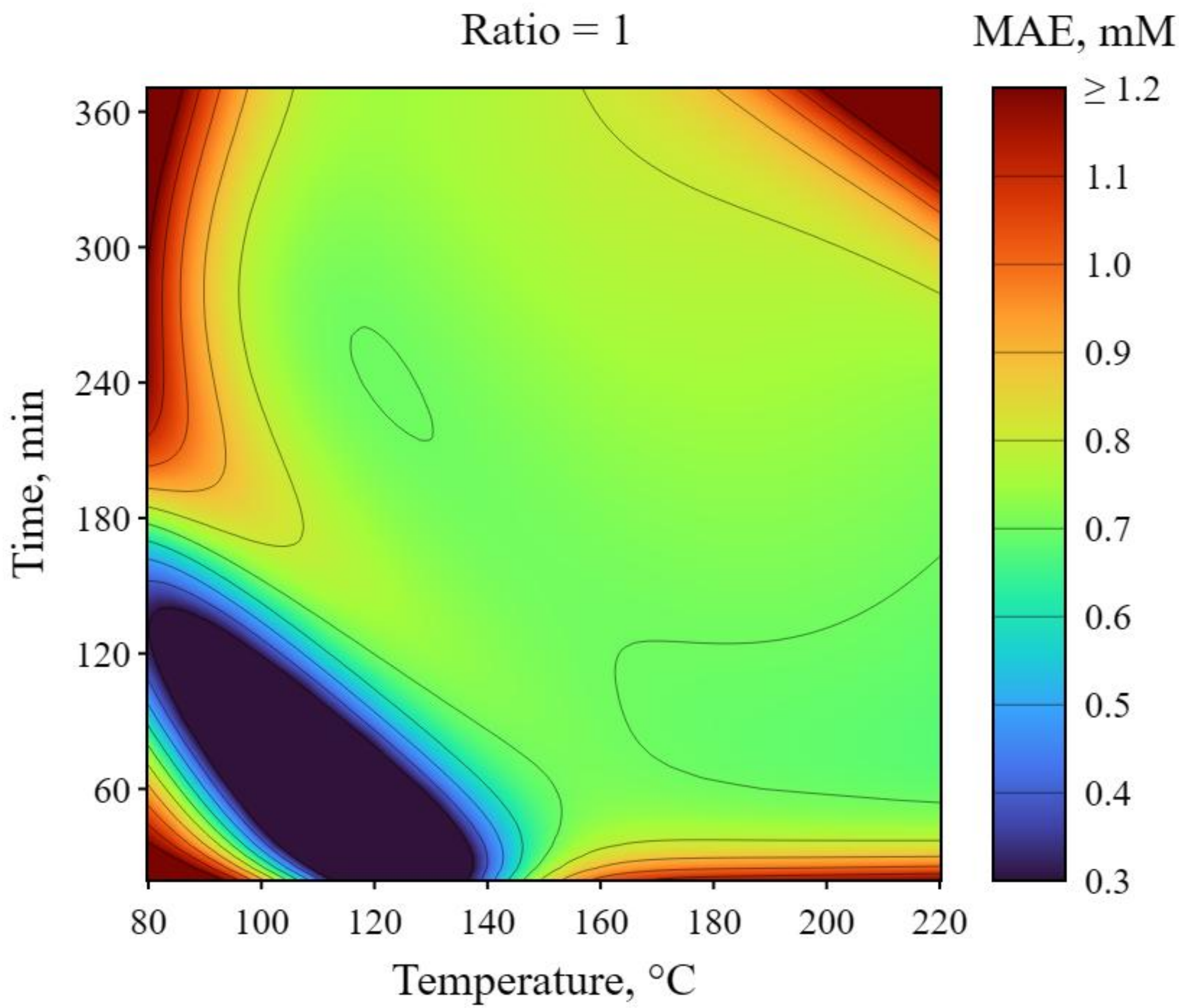
Stage 2 – approximation model – $Co(NO_3)_2$



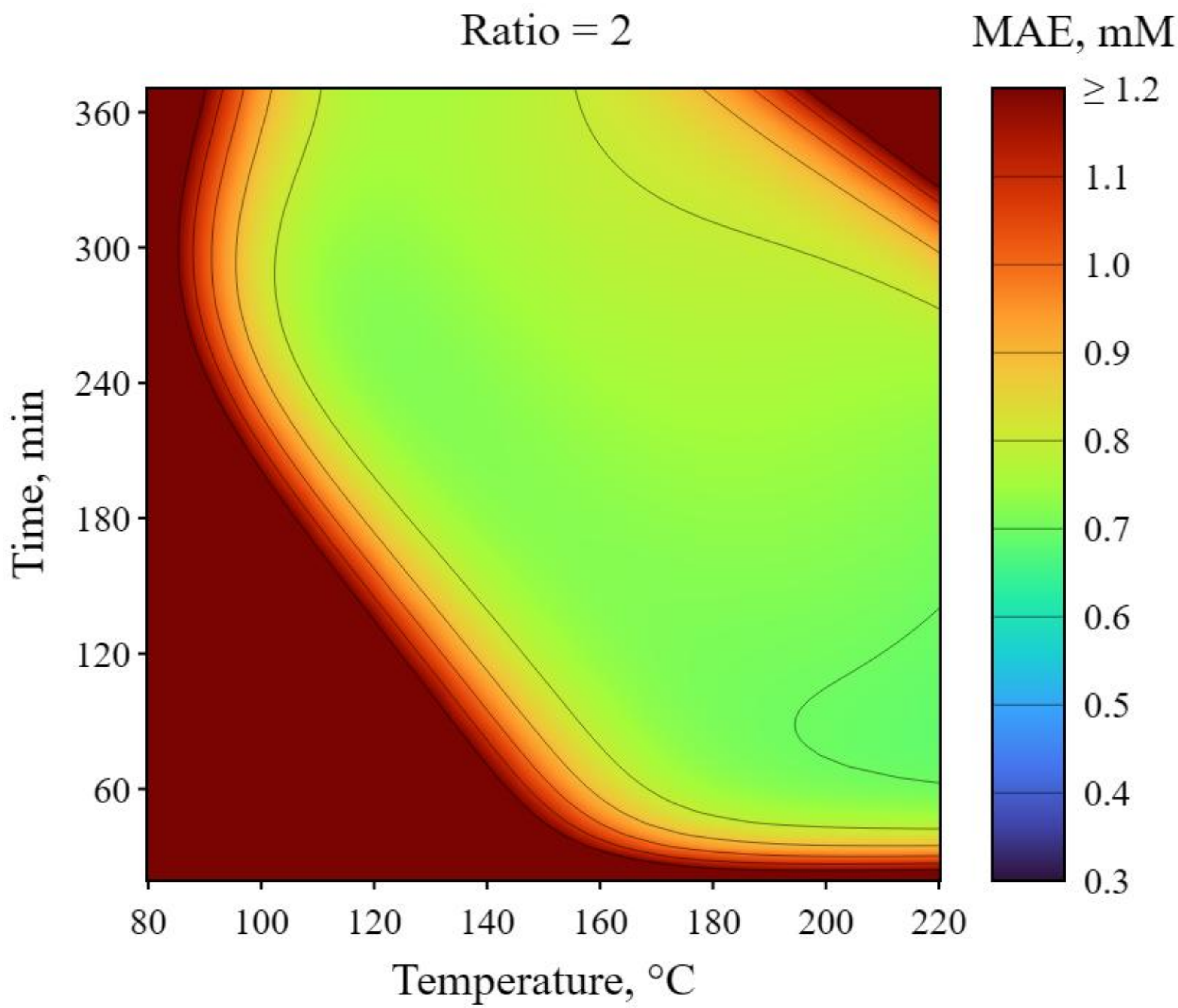
Stage 2 – approximation model – $Co(NO_3)_2$



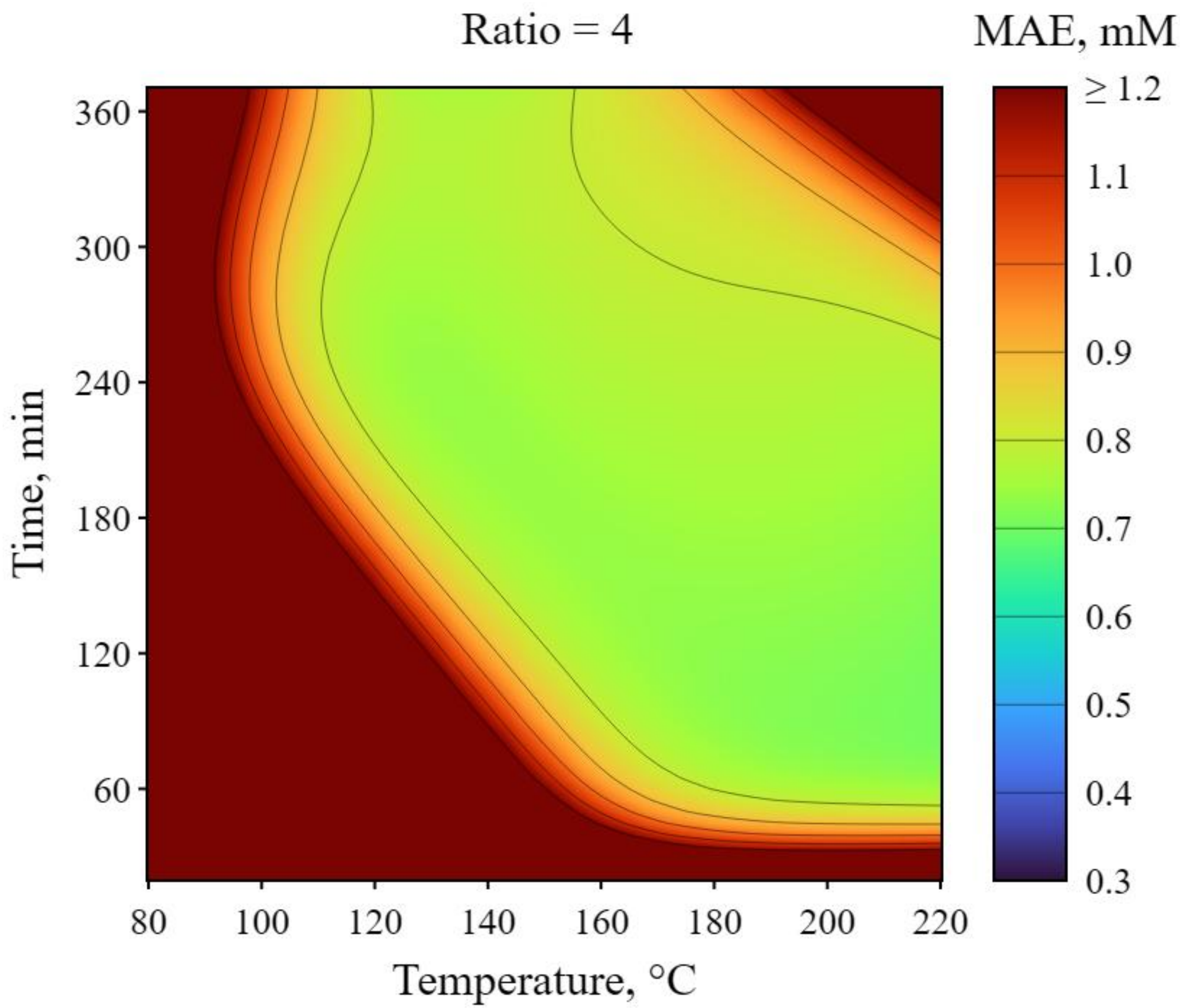
Stage 2 – approximation model – $Cu(NO_3)_2$



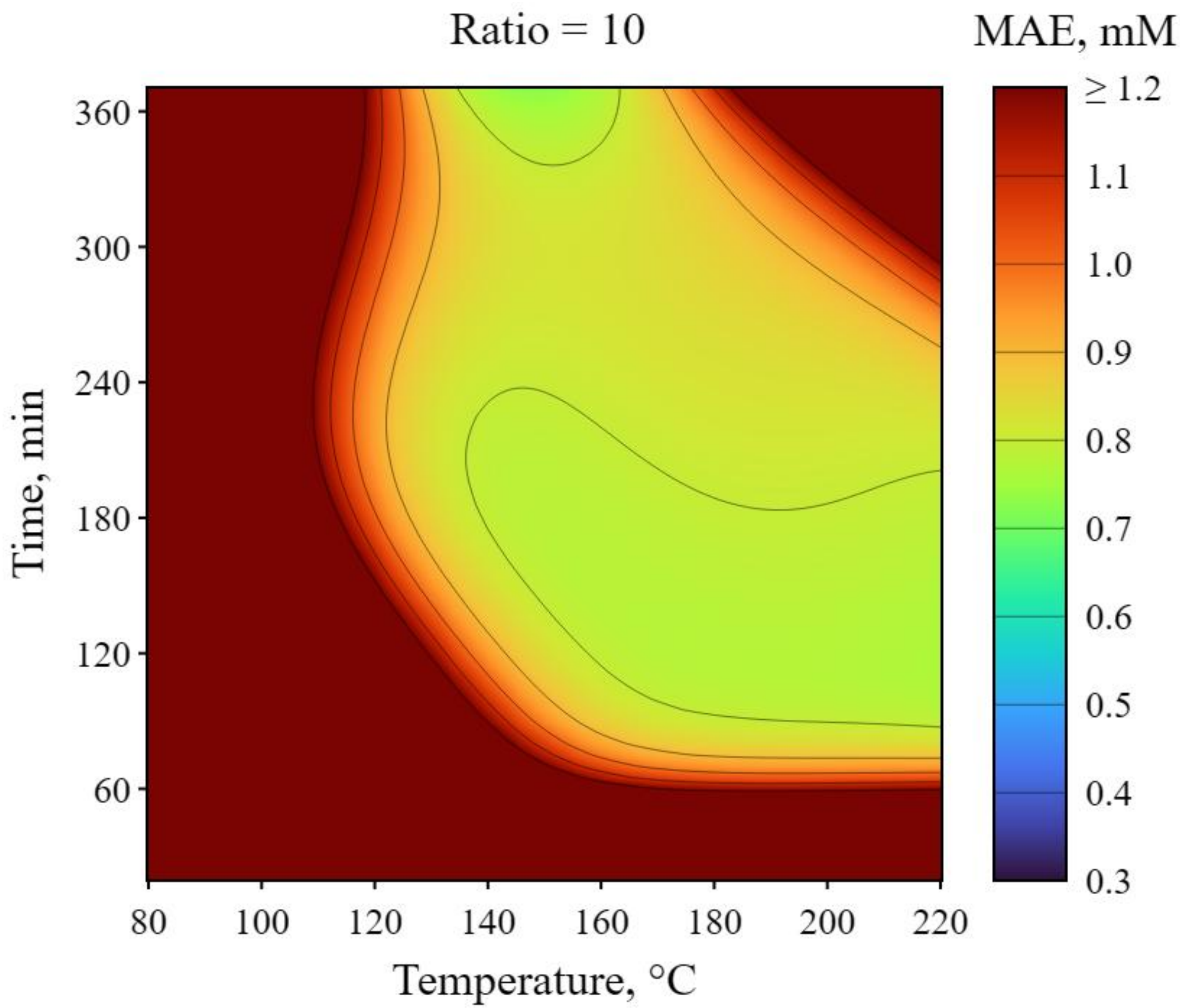
Stage 2 – approximation model – $Cu(NO_3)_2$



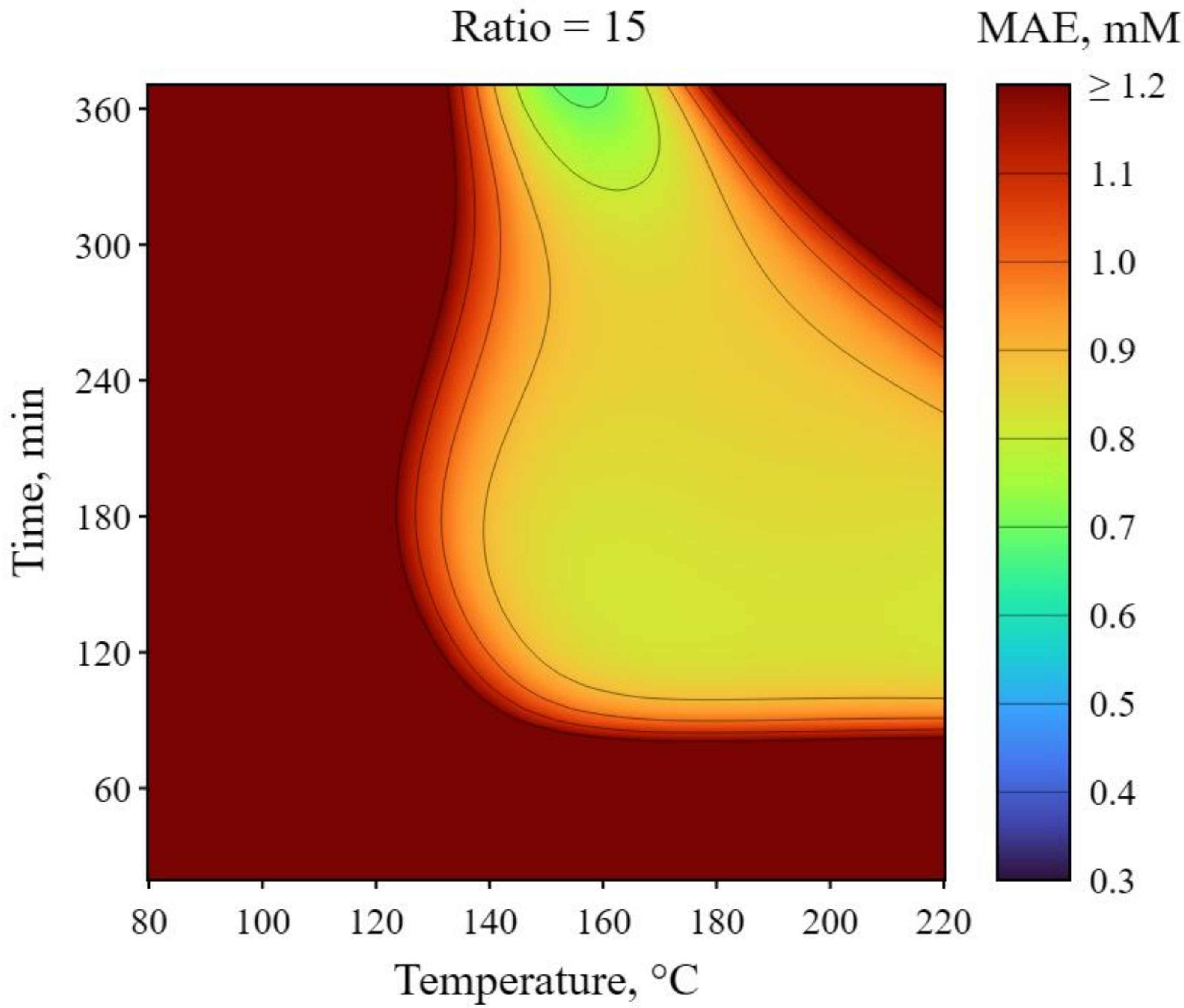
Stage 2 – approximation model – $Cu(NO_3)_2$



Stage 2 – approximation model – $Cu(NO_3)_2$



Stage 2 – approximation model – $Cu(NO_3)_2$



Stage 2 – approximation model – $Cu(NO_3)_2$