

# Explorative forecasting of multidimensional data

Diana Domańska

Institute of Computer Science, University of Silesia, Poland



## Introduction

- ▶ Forecasting is related to decision making in many aspects of life.
- ▶ Intuitive method for imprecise, incomplete data is the theory of fuzzy sets.
- ▶ Multidimensional data – curse of dimensionality, fractional distance.

## The explorative forecast procedure

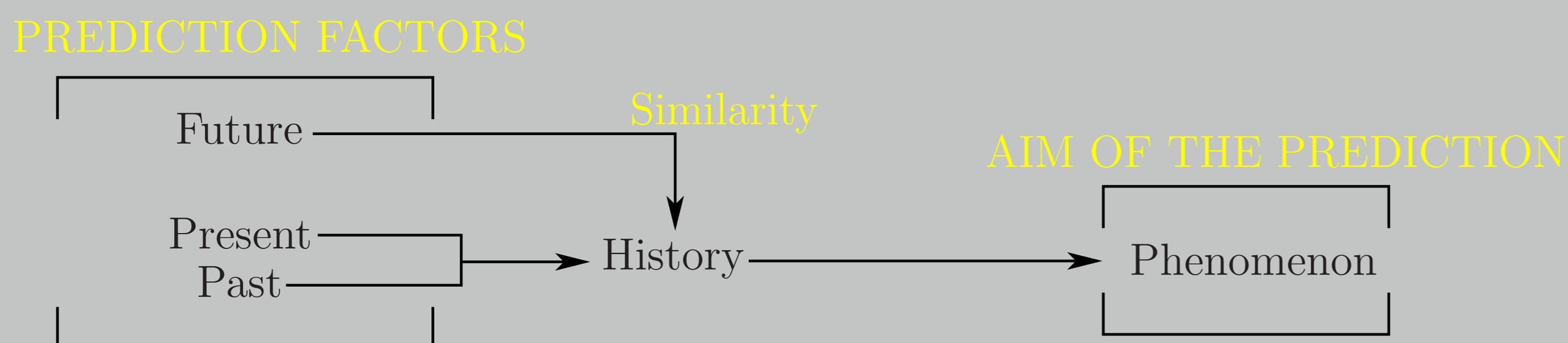


Figure : The explorative forecast procedure diagram.

## Explorative forecasting of multidimensional data

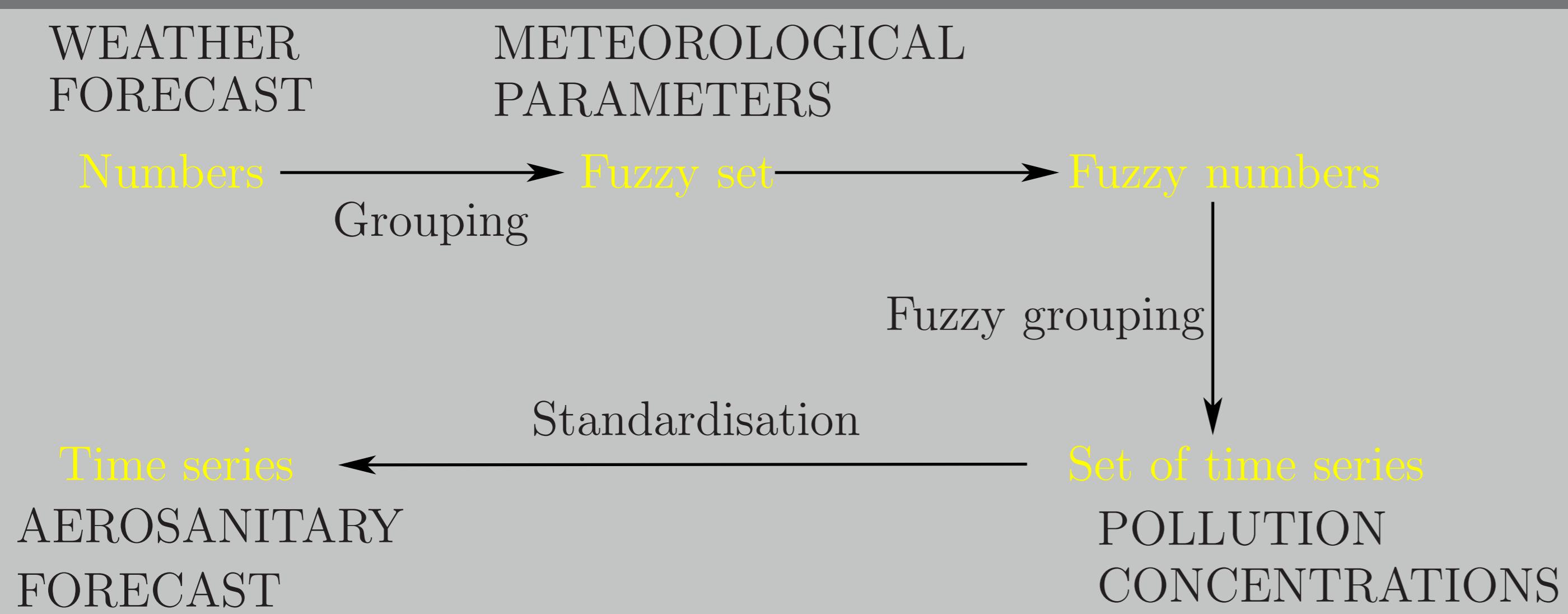


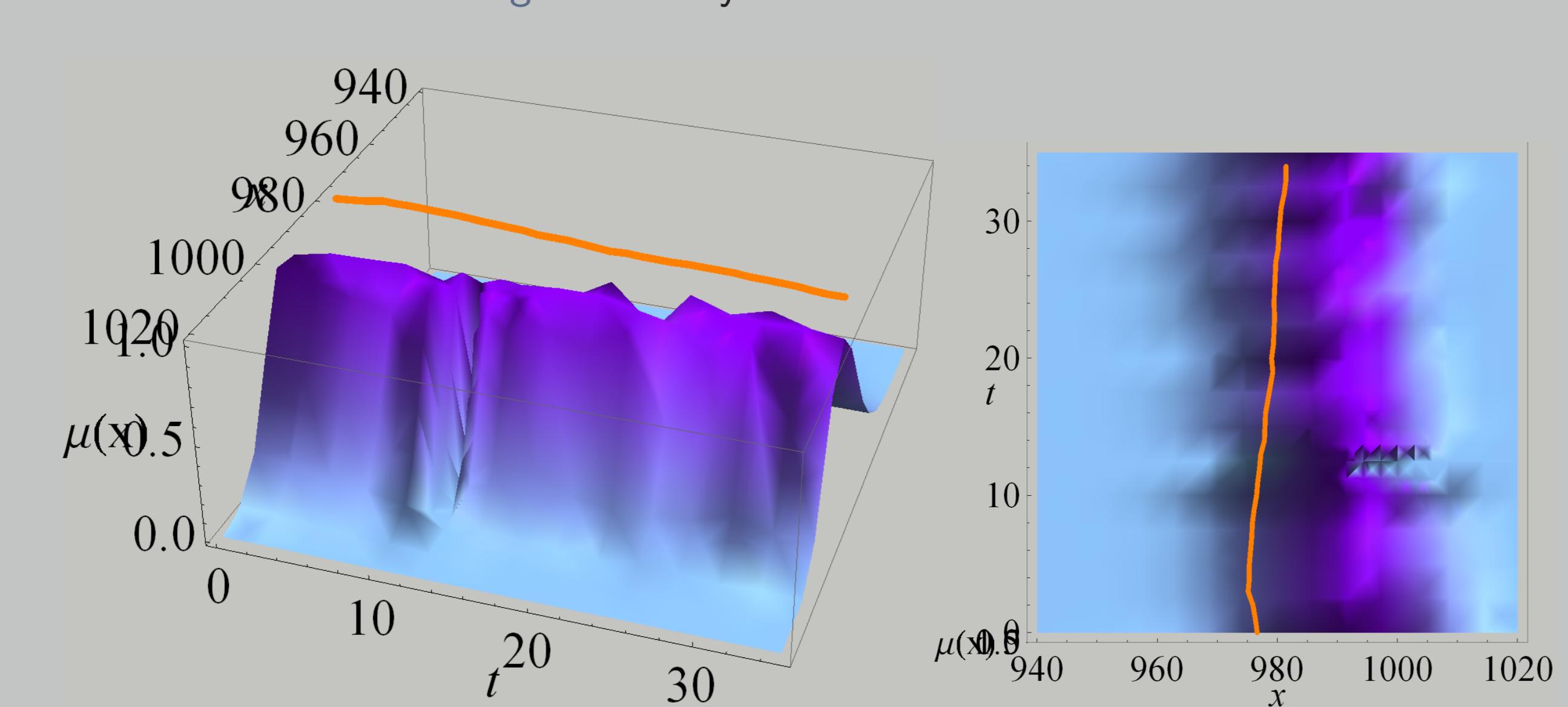
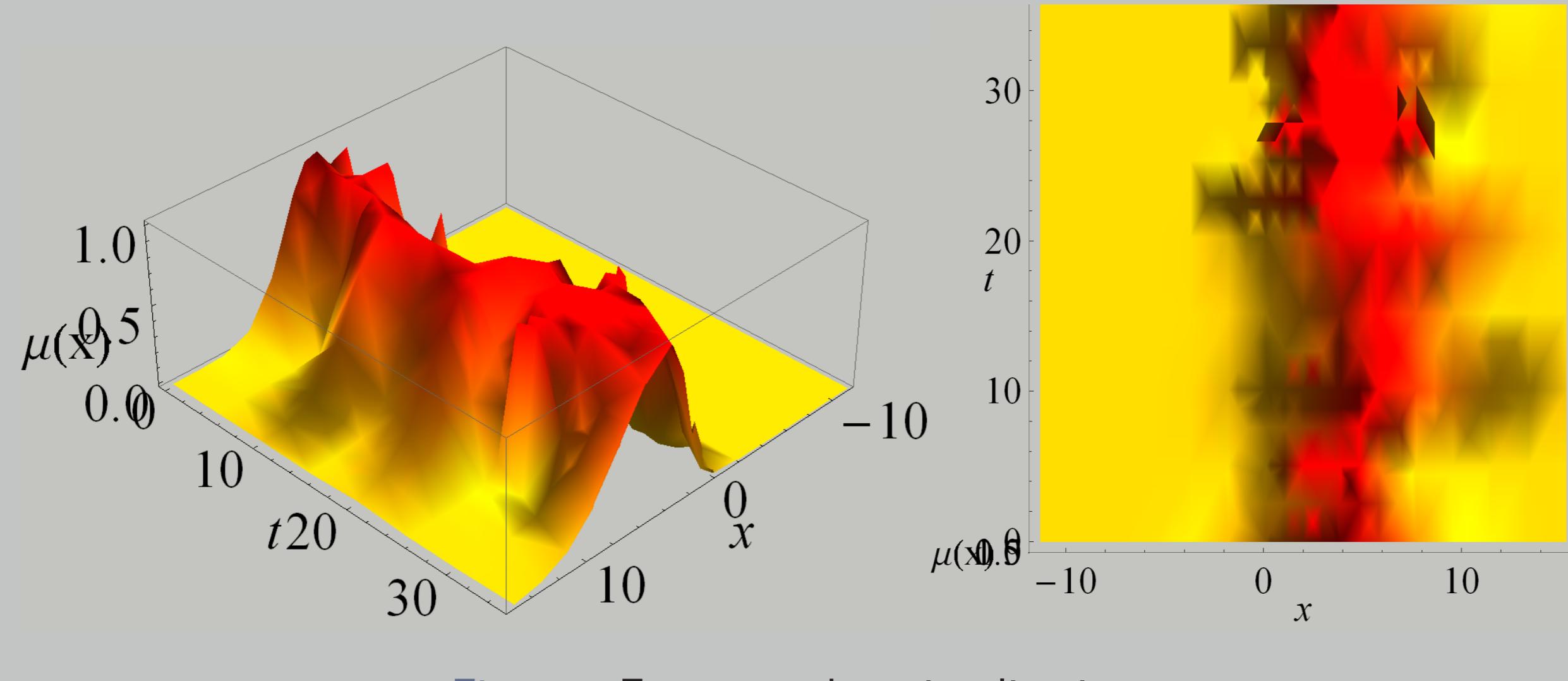
Figure : Explorative forecasting of multidimensional data.

## Data for explorative forecasting pollution concentrations

	APFM – (Air Pollution Forecasting Model)	e-APFM – (extended-Air Pollution Forecasting Model)
Input	-weather forecast, meteorological parameters, pollution concentration -information about wind direction -information about the meteorological stations	
Output	Time series of pollution concentrations for 72 hours with one hour time step	

## Other applications

- ▶ data visualization



## Other applications

- ▶ improvement of selected attribute values forecasts i.e. humidity

Synoptic station	Range	MAE error between the measured and the predicted humidity	fuzzy number using center of gravity	fuzzy number using middle of maximum
Bielsko-Biała	+ 1 day	13.10	10.51	10.96
	+ 2 day	13.50	10.51	10.97
	+ 3 day	13.69	12.53	13.42
Katowice	+ 1 day	14.28	12.62	14.08
	+ 2 day	15.04	12.86	14.29
	+ 3 day	15.71	14.54	16.60

## Experiments - forecast for 3 days with an hour time step

PM<sub>10</sub> pollution range (2011-2012) for selected monitoring stations and methods with an hour time step.

Station	Range	Mean	Standard deviation
APFM, e-APFM [Złoty Potok]	[0.00, 242.00]	27.45	20.79
m. optim. NN 22 (NN) [Zagrzeb]	-	28.55	24.49

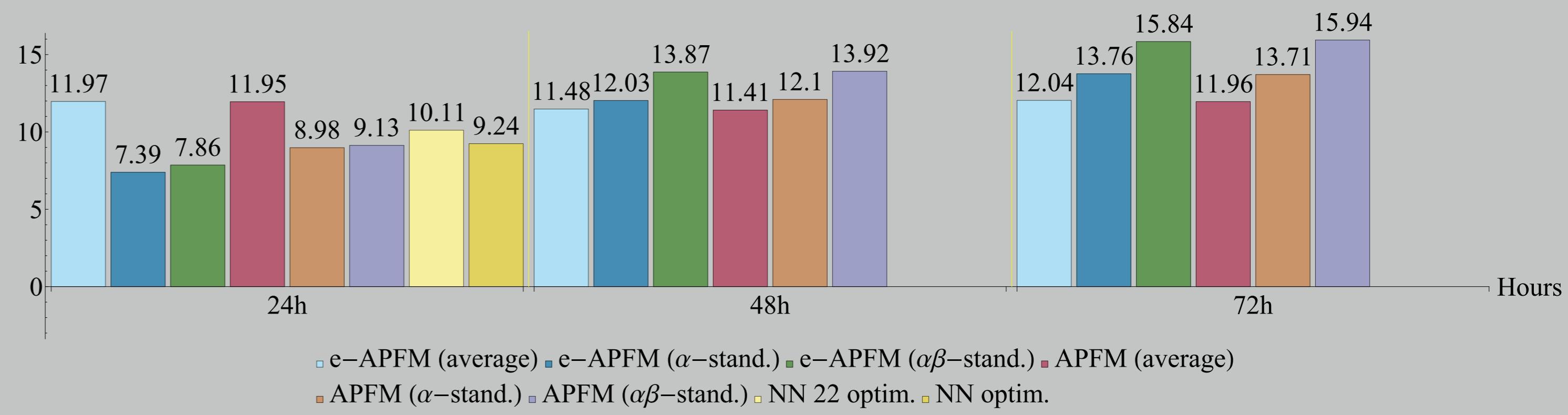


Figure : Results of forecasting with a time step equal to one hour for the APFM, e-APFM methods and methods from the literature.

## Experiments - forecast for 3 days with an average daily time step

PM<sub>10</sub> pollution range (2011-2012) for selected monitoring stations and methods with an average daily time step.

Station	Range	Mean	Standard deviation
APFM, e-APFM [Złoty Potok]	[3.54, 178.58]	27.51	17.73
distance-based geographic model [Istanbul area F-U-S (K-Y-S)]	[9.00, 206.00]	53.00	32.00

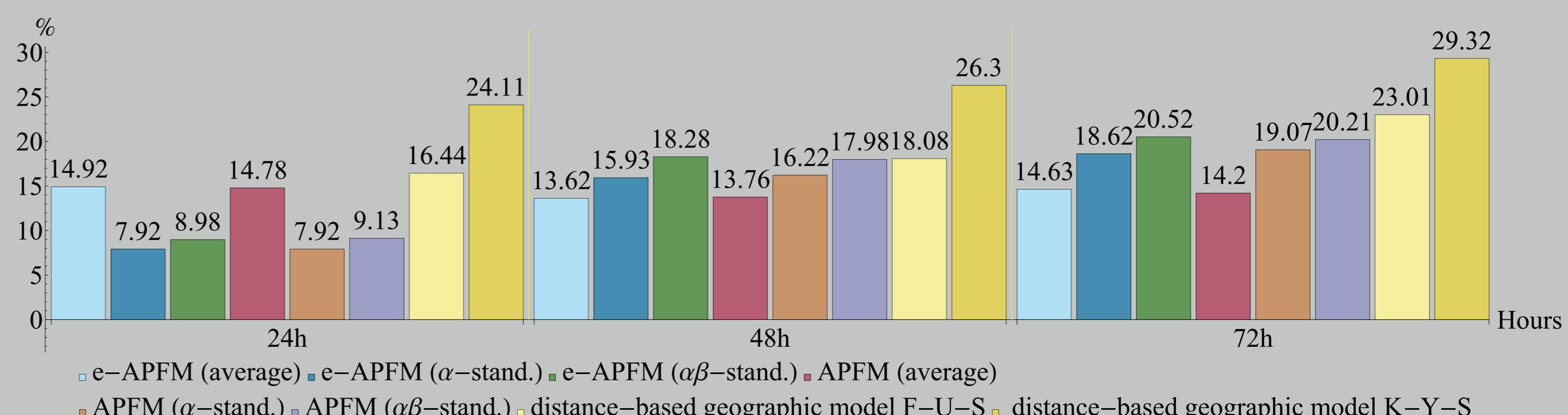


Figure : Results of daily average forecasting for the APFM, e-APFM methods and methods from the literature.

## Future plans

- ▶ Forecasting high values of pollution concentrations.
- ▶ Using Delaunay's triangulation to forecast pollution not only for station, but also for area.
- ▶ Multi fuzzy numbers.

