Take a deep breath

ML for a mechanical ventilator

Haritha Retnakaran, Martin Lützner, Christian Klingler

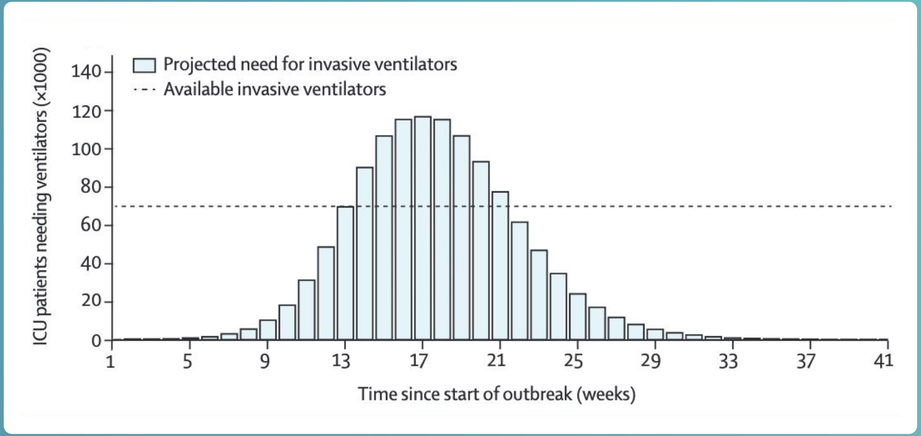
19-05-22

Our team

Haritha Retnakaran	Christian Klingler	Martin Lützner
		
Physicist Data Scientist Germany	Biochemist Data Scientist Freiburg	Theoretical Physicist Data Scientist Berlin

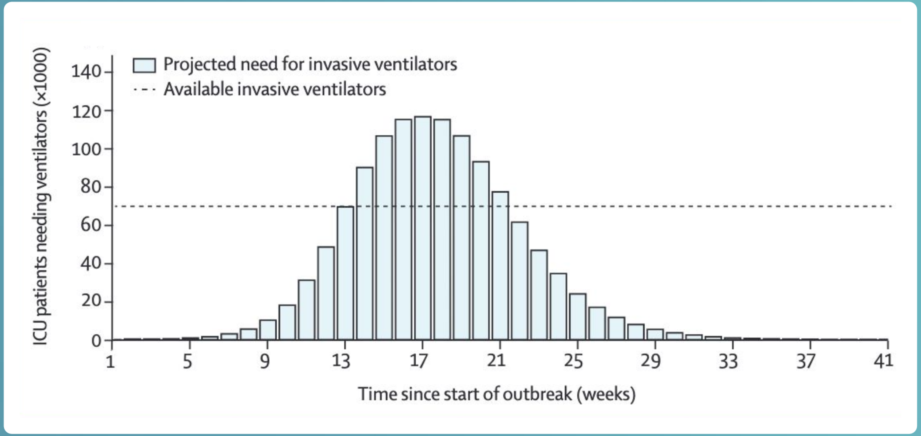
Medical situation and demand

Medical situation and demand



Wells C. R. et. al. The Lancet; 2020

Medical situation and demand



Wells C. R. et. al. The Lancet; 2020

- Rapidly-deployable ventilator need
- Low cost ventilator need
- “automatic” mode of action needed

Stakeholder information

PVP

THE PEOPLE'S
VENTILATOR
PROJECT

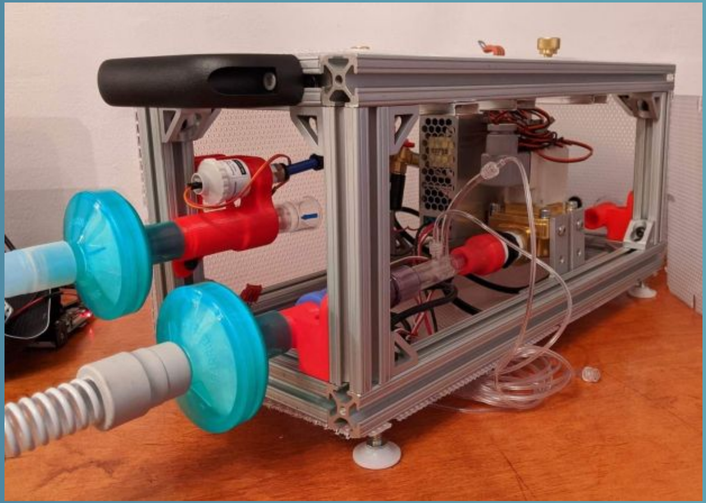
Stakeholder information

funded by
Princeton University

PVP
THE PEOPLE'S
VENTILATOR
PROJECT

Stakeholder information

funded by
Princeton University

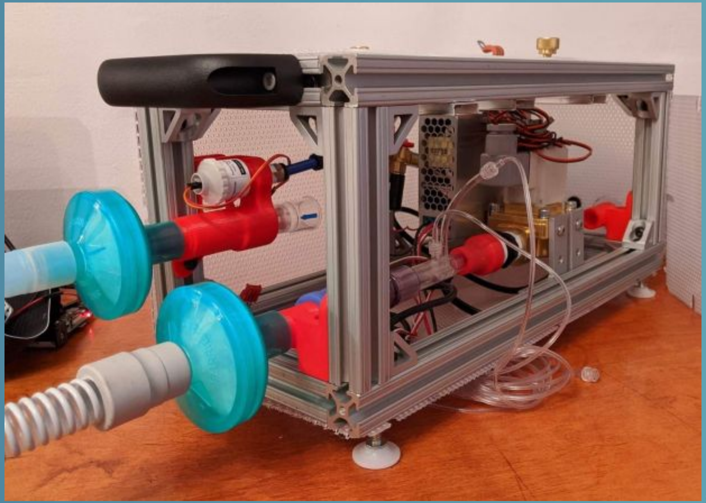


Fully automatic mechanical ventilator

Stakeholder information

funded by **Princeton University**

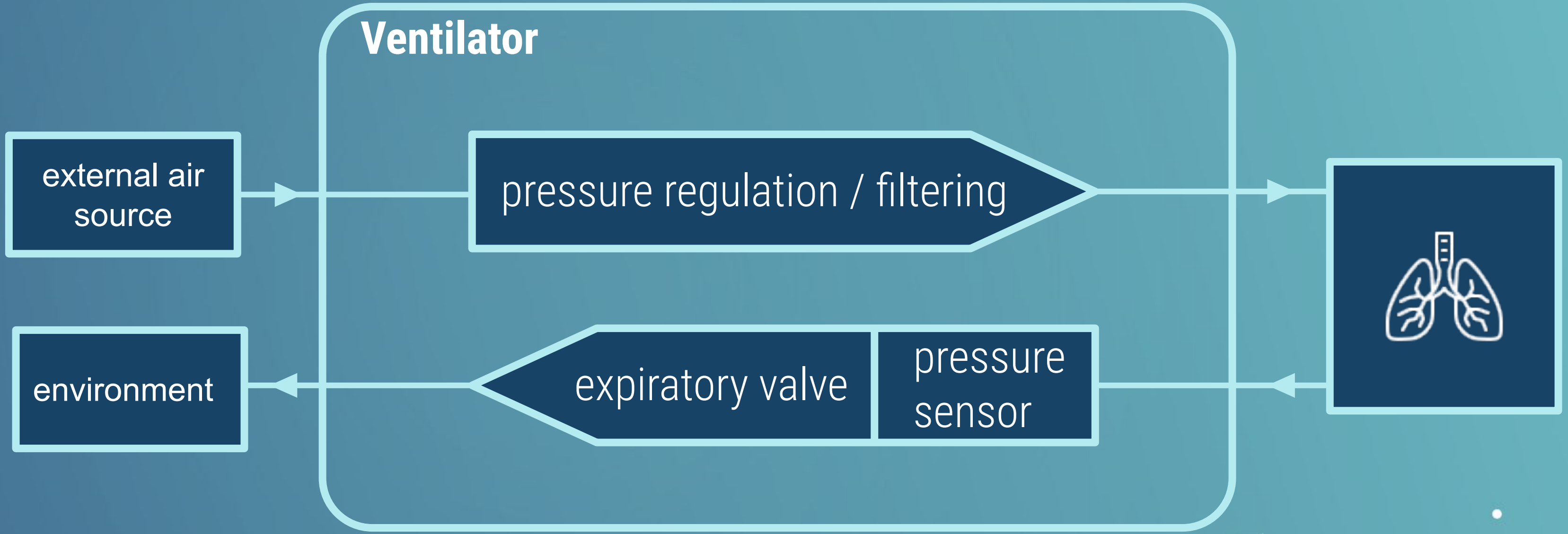
DIY for **< $1,300** by **single person** in **3 days**



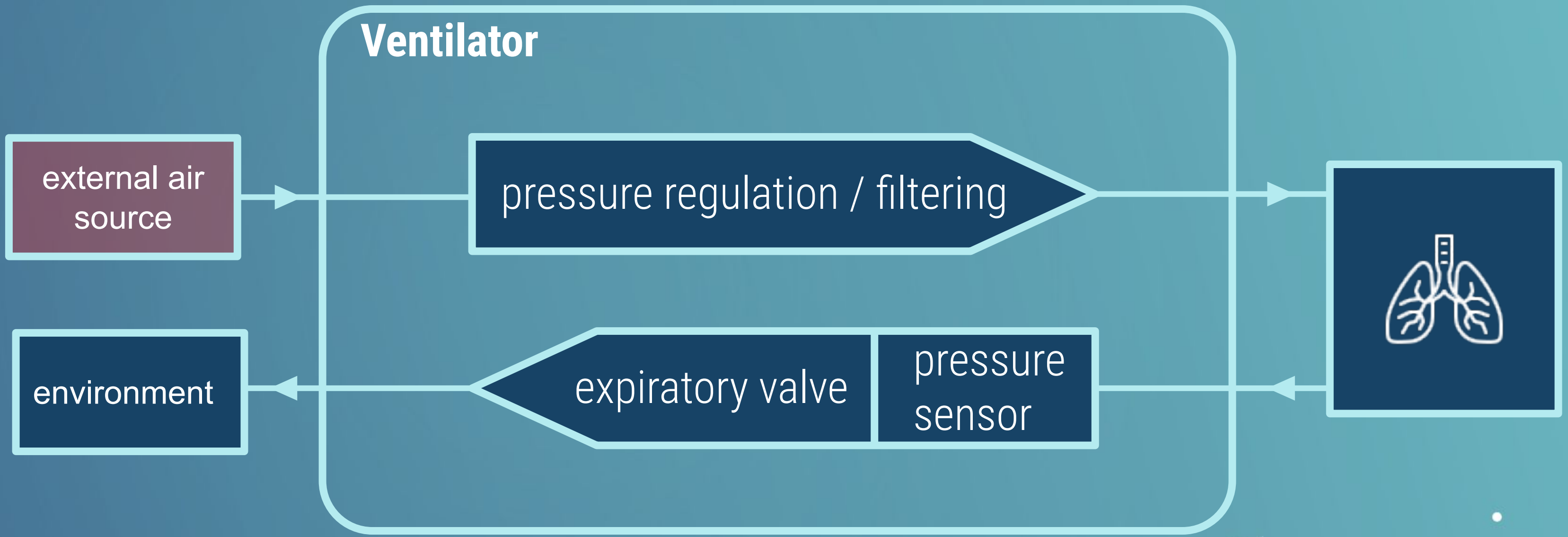
Fully automatic mechanical ventilator

Customer: **Health Care Services**

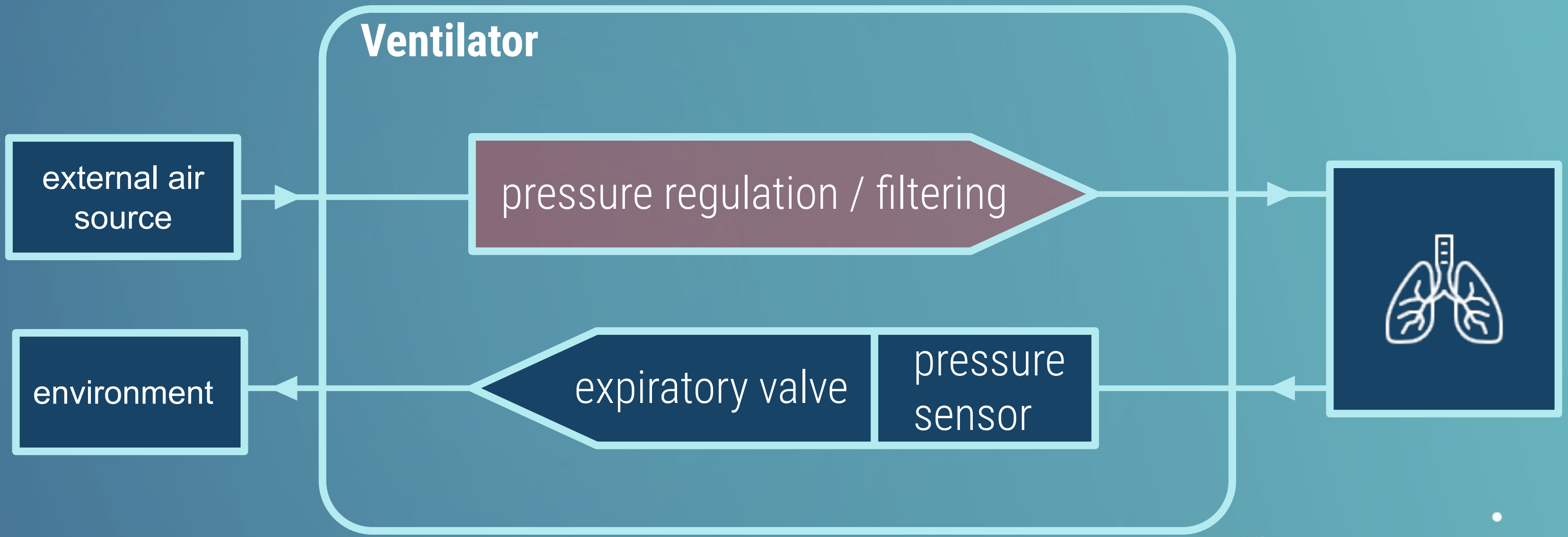
Respiratory circuit



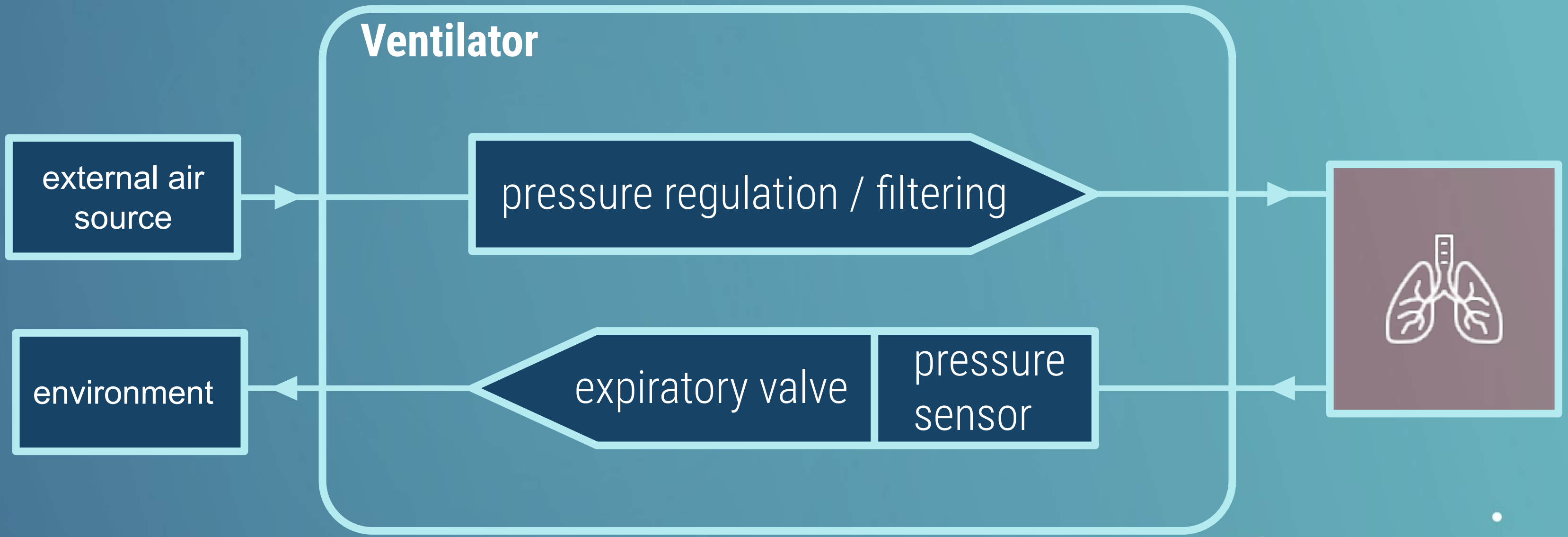
Respiratory circuit



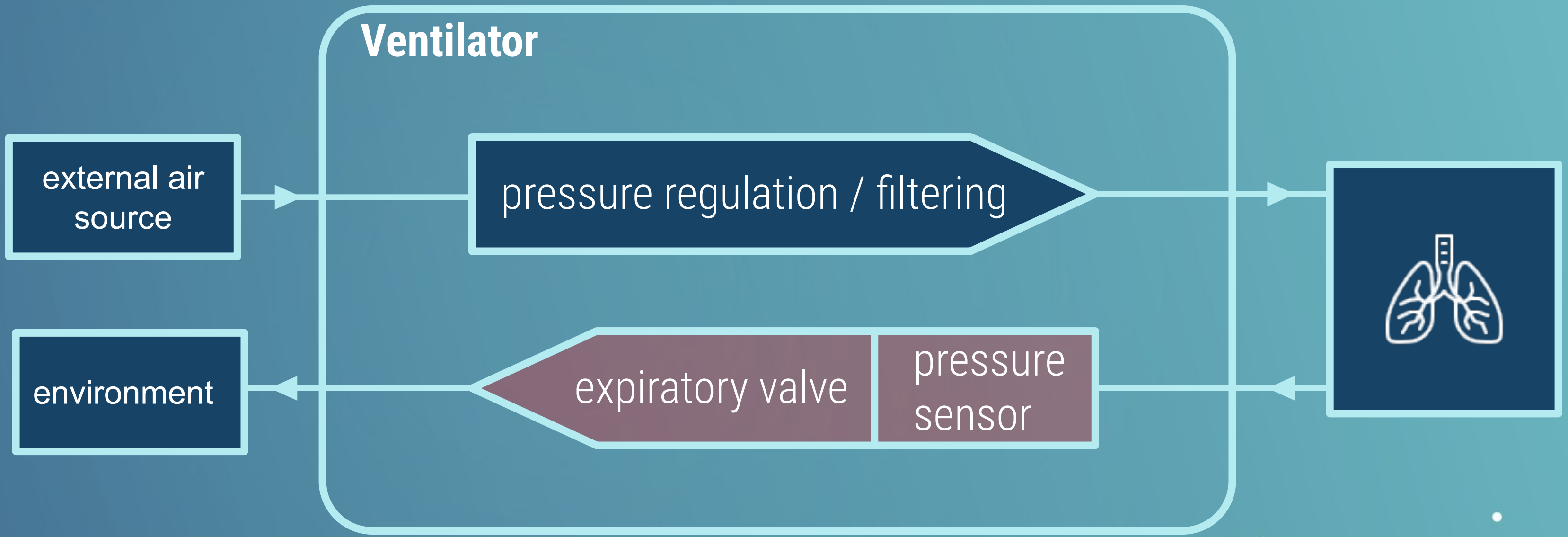
Respiratory circuit



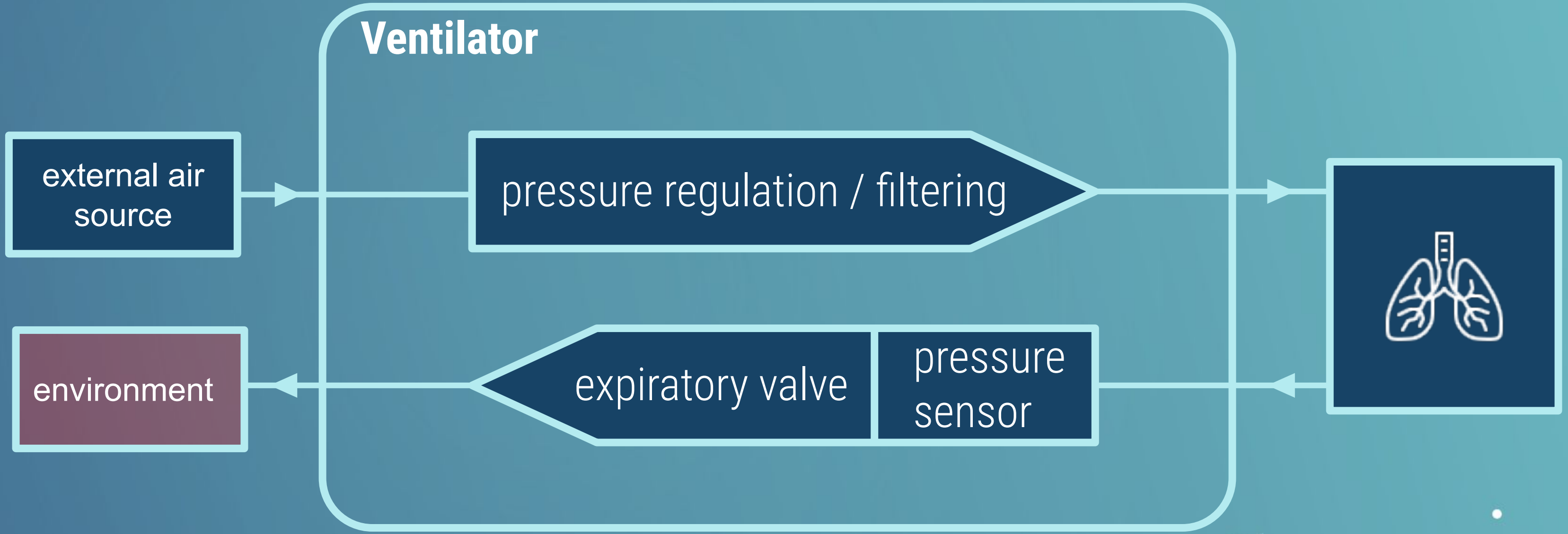
Respiratory circuit



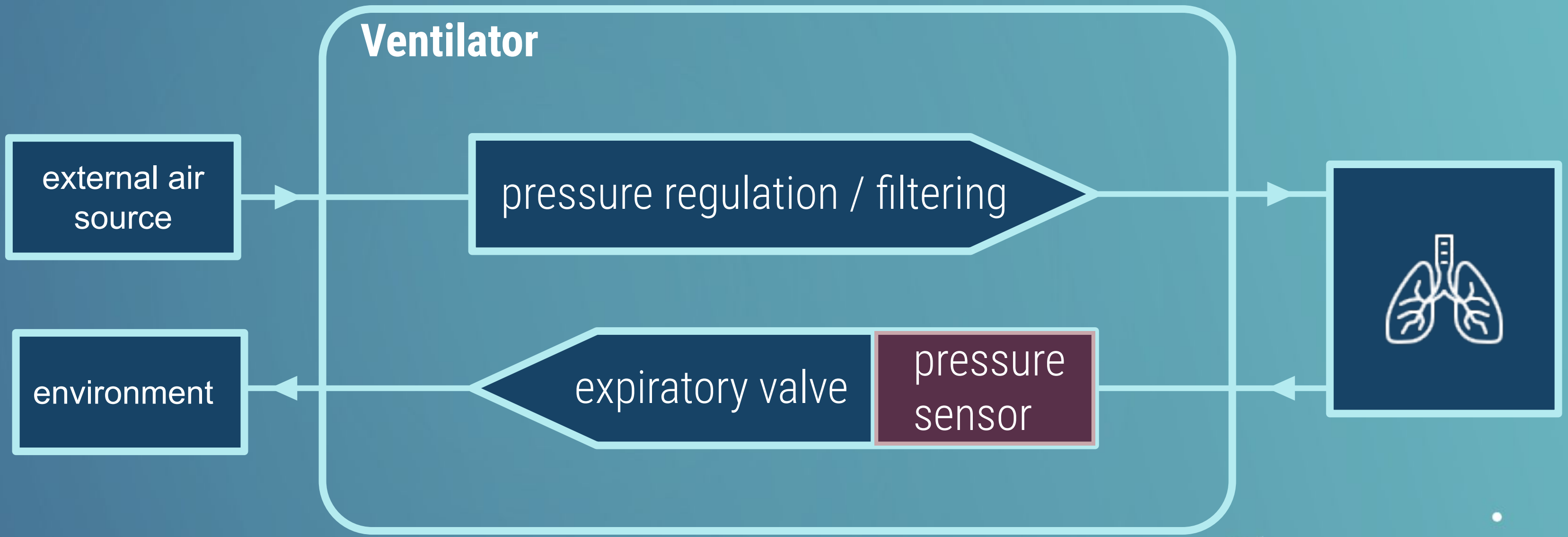
Respiratory circuit



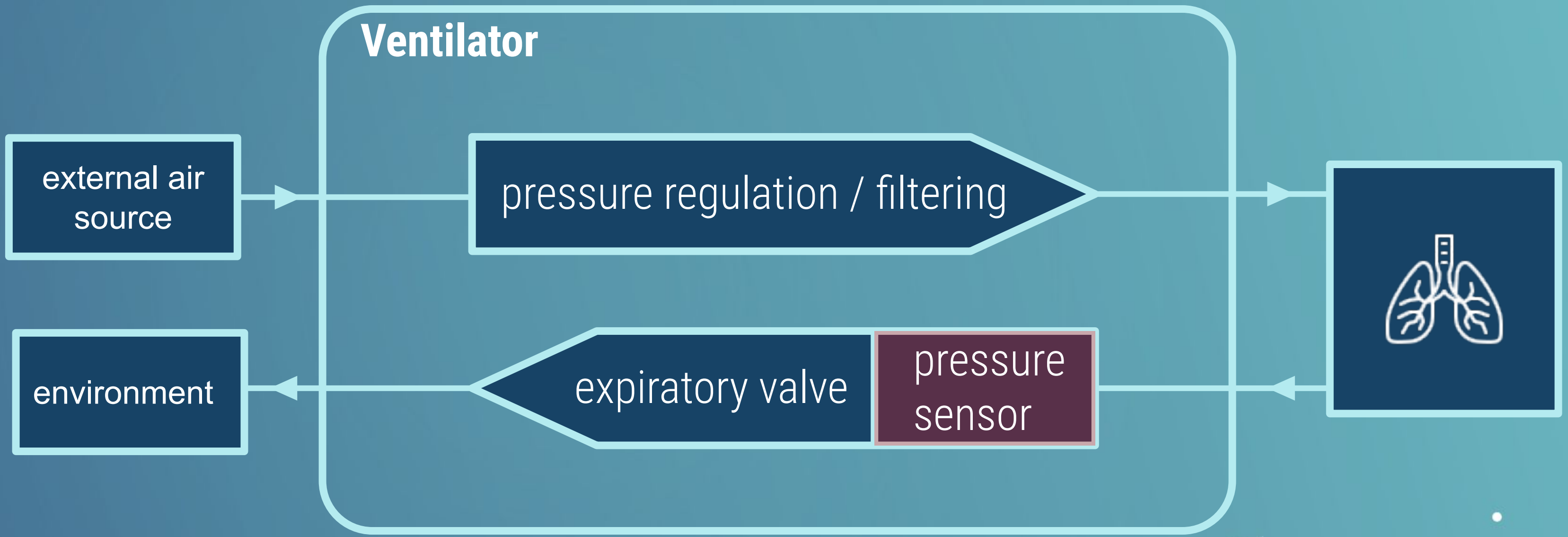
Respiratory circuit



Respiratory circuit



Respiratory circuit



Error metric:
Mean absolute error (MAE)

01

Data Analysis

Data structure

Breath ID

base

- Time steps (TS)
- Input flow (IF)
- output valve
- Resistance (R)
- Compliance (C)

engineered

- total air volume
- Time-shifted IF

Data structure

Breath ID

base

Time steps (TS)

Input flow (IF)

output valve

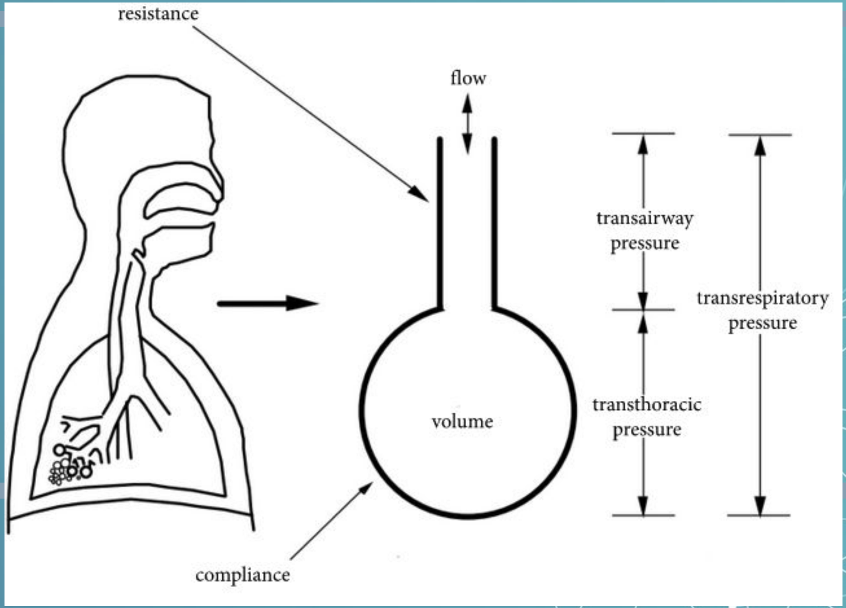
Resistance (R)

Compliance (C)

engineered

total air volume

Time-shifted IF



Tran A.S. et. al. Math. Problems in Engineering; 2022

Data structure

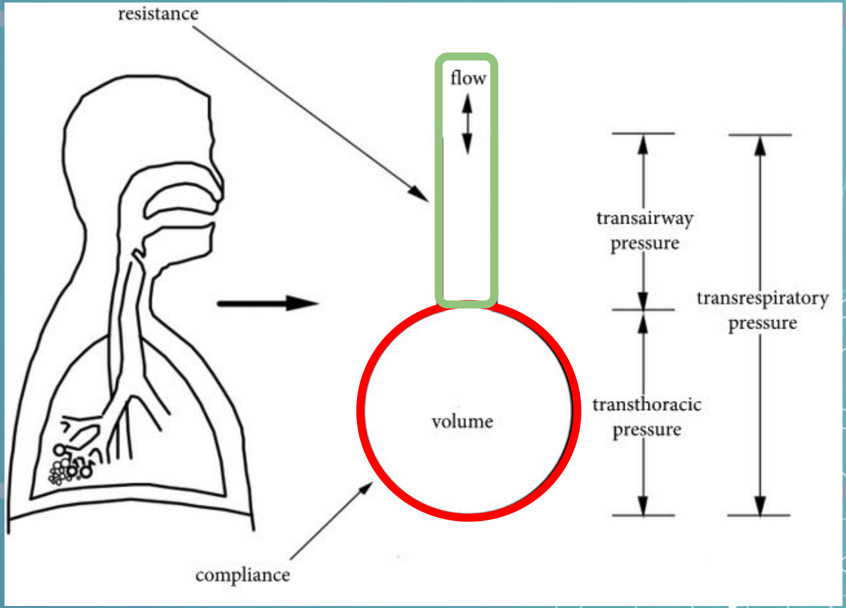
Breath ID

base

- Time steps (TS)
- Input flow (IF)
- output valve
- Resistance (R)
- Compliance (C)

engineered

- total air volume
- Time-shifted IF



Tran A.S. et. al. Math. Problems in Engineering; 2022

Data structure

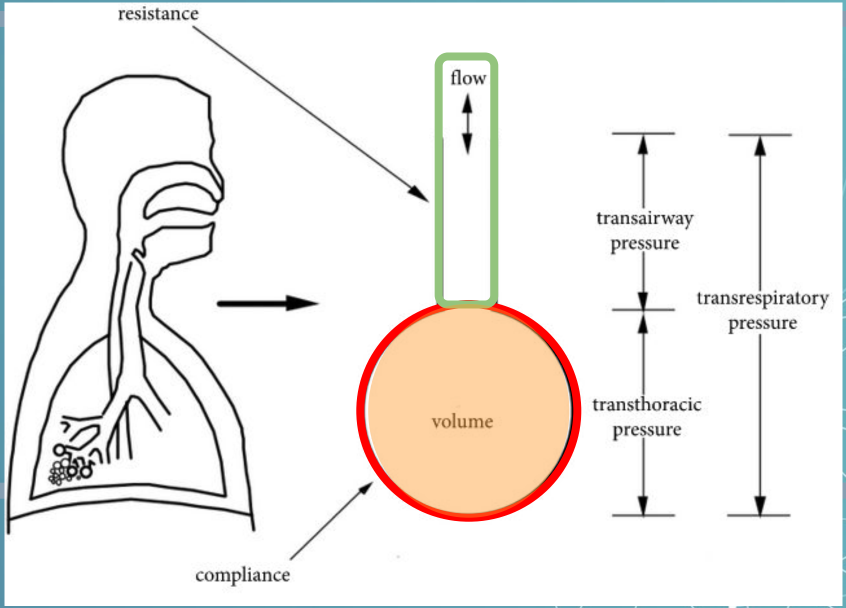
Breath ID

base

- Time steps (TS)
- Input flow (IF)
- output valve
- Resistance (R)
- Compliance (C)

engineered

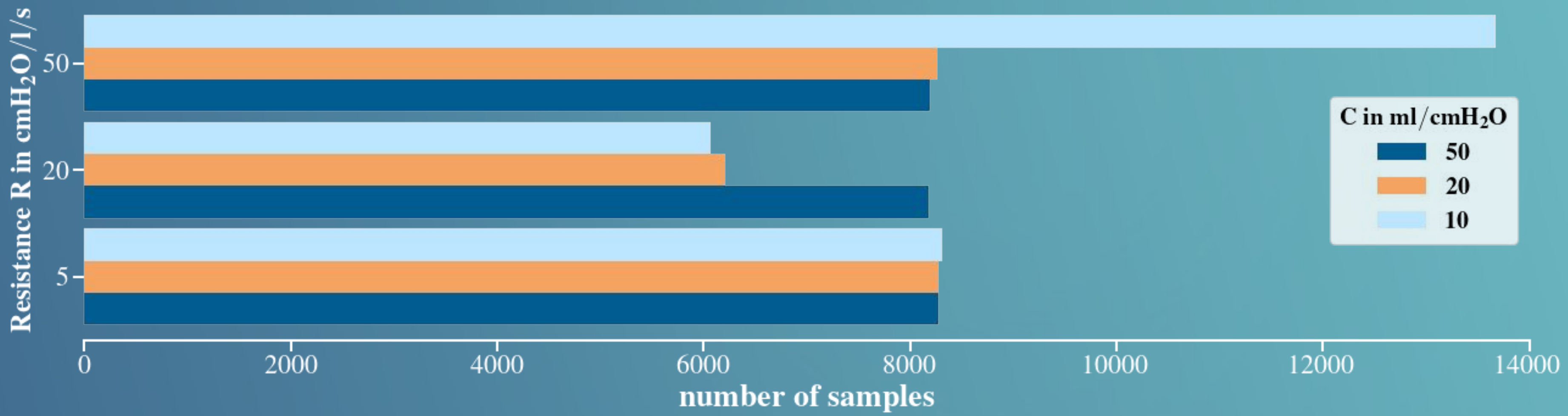
- total air volume
- Time-shifted IF



Tran A.S. et. al. Math. Problems in Engineering; 2022

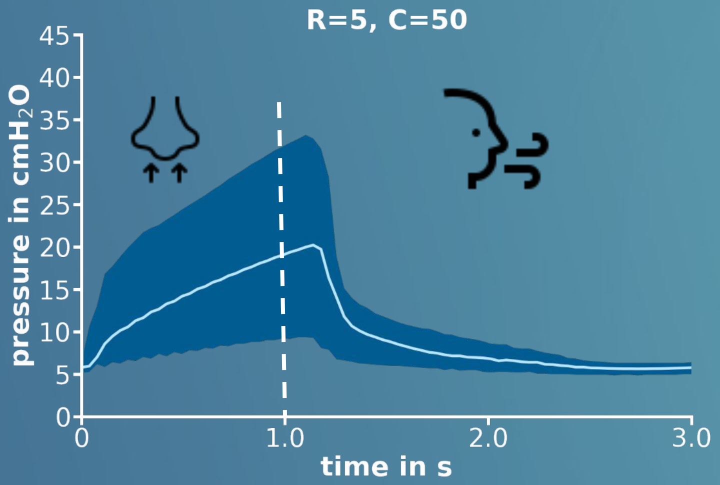
Distribution of data

- 74500 sample breaths with 9 different parameter combinations
- More samples for unhealthy lung parameters



Breathing profiles

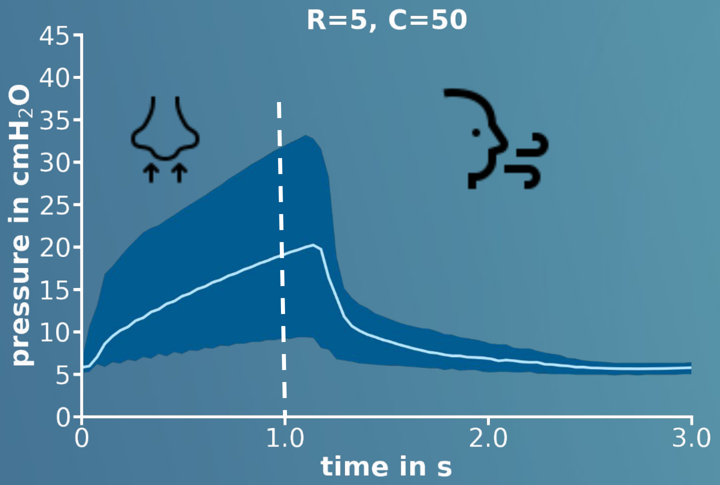
- Very different shapes of breaths



Unit of R: $cmH_2O/l/s$
Unit of C: ml/cmH_2O

Breathing profiles

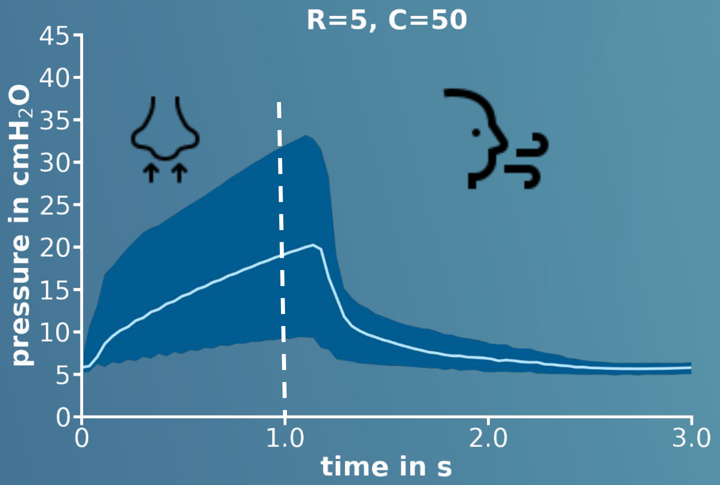
- Very different shapes of breaths

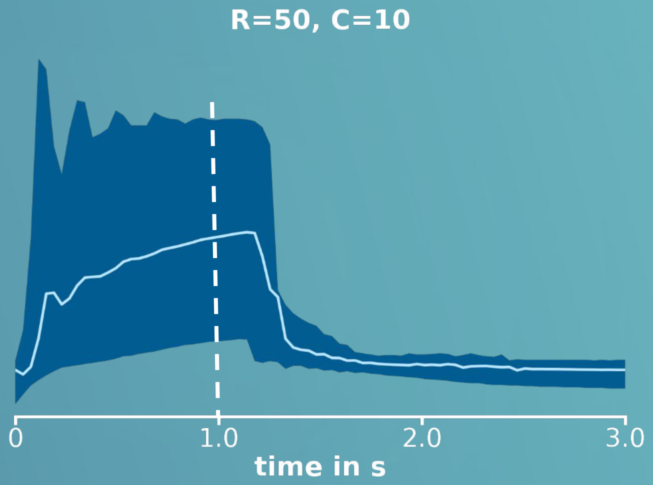


Unit of R: $cmH_2O/l/s$
Unit of C: ml/cmH_2O

Breathing profiles

- Very different shapes of breaths
- Larger deviations for unhealthy parameters





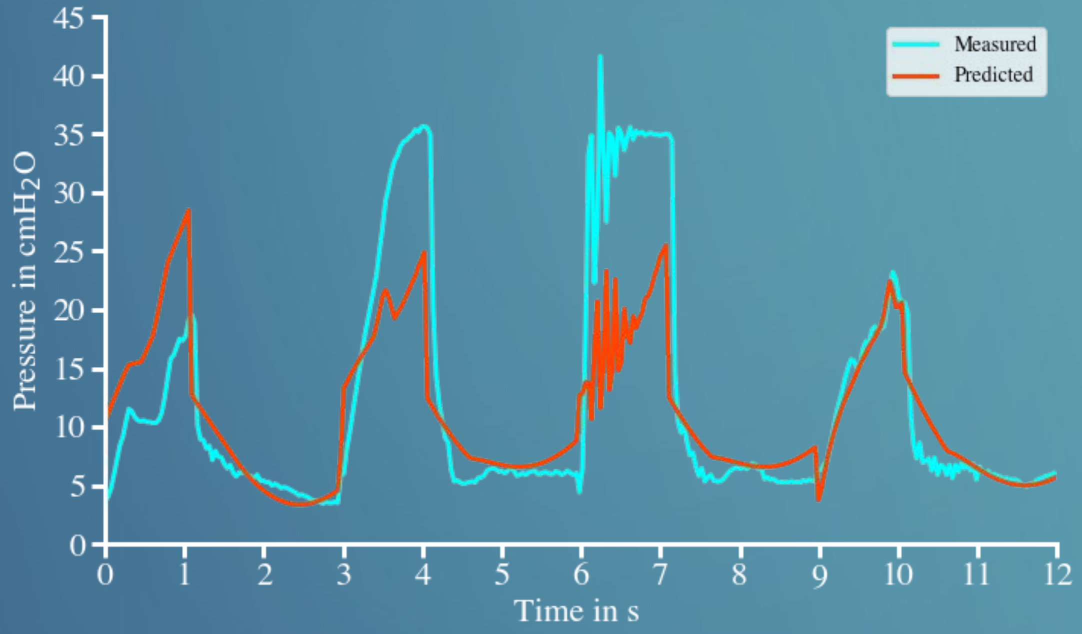
Unit of R: $cmH_2O/l/s$
Unit of C: ml/cmH_2O

02

Models

Baseline model

Polynomial function (degree 2) on input flow, resistance, compliance, output valve



MAE = 3.38 cmH_2O

More complex models needed!

Model comparison

	ANN	XGB
MAE	0.47	0.25
Improvement (to baseline)	**86,1%**	**92,6%**

Model comparison

	ANN	XGB	ARIMA
MAE	0.47	0.25	4.91
Improvement (to baseline)	**86,1%**	**92,6%**	**-45,2%**

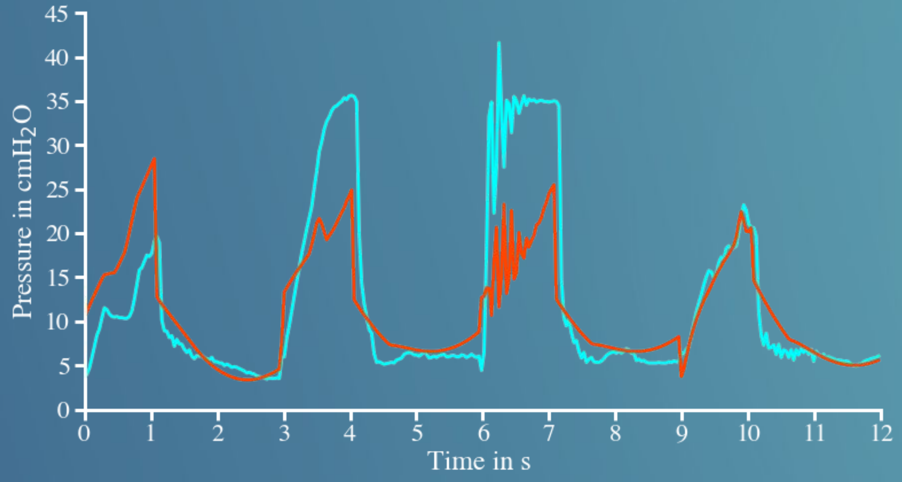
Model comparison

	ANN	XGB	ARIMA	XGB(ARF)	ANN(ARF)
MAE	0.47	0.25	4.91	0.158	0.15
Improvement (to baseline)	**86,1%**	**92,6%**	**-45,2%**	**95,3%**	**95,6%**

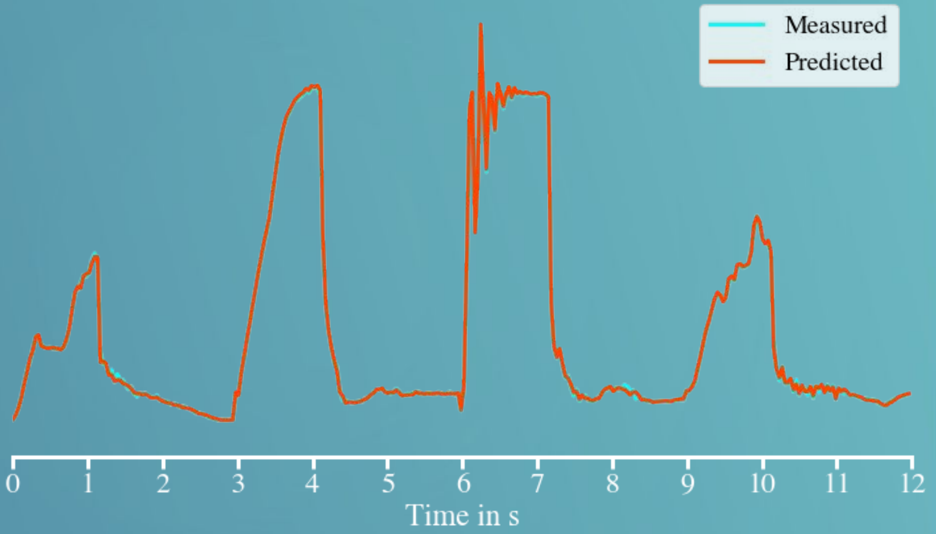
ARF Autoregressive features

Comparing the models

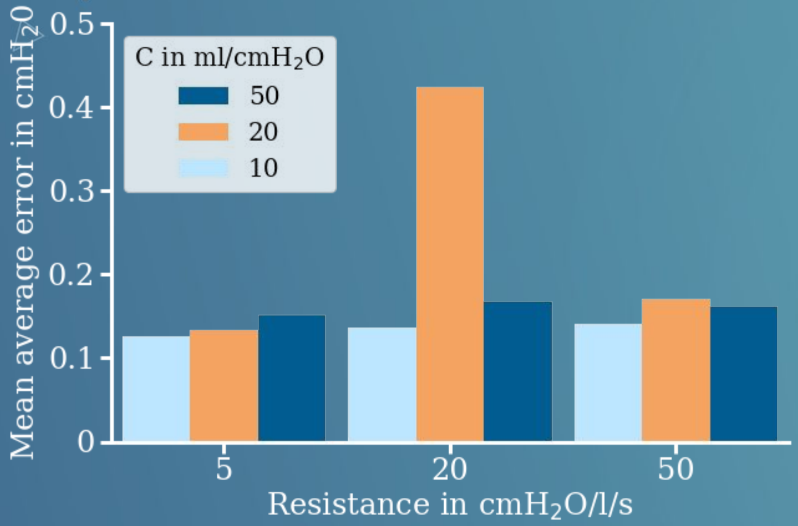
Baseline model



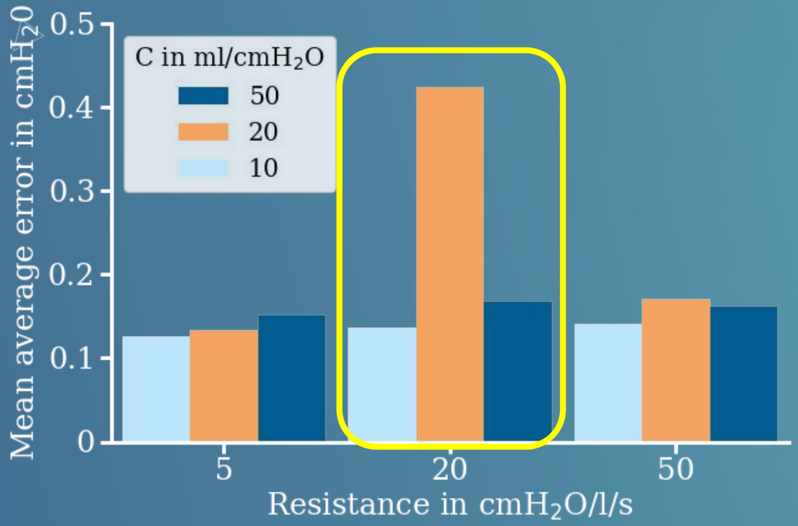
Best model: ANN(ARF)



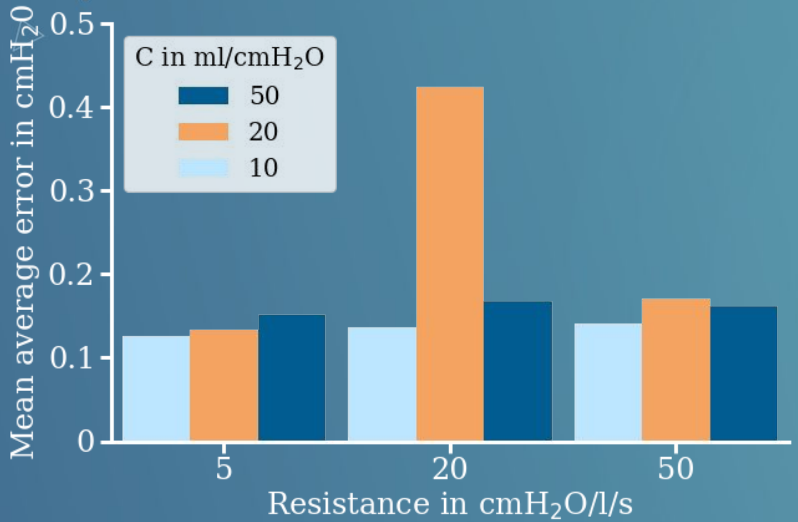
Error analysis

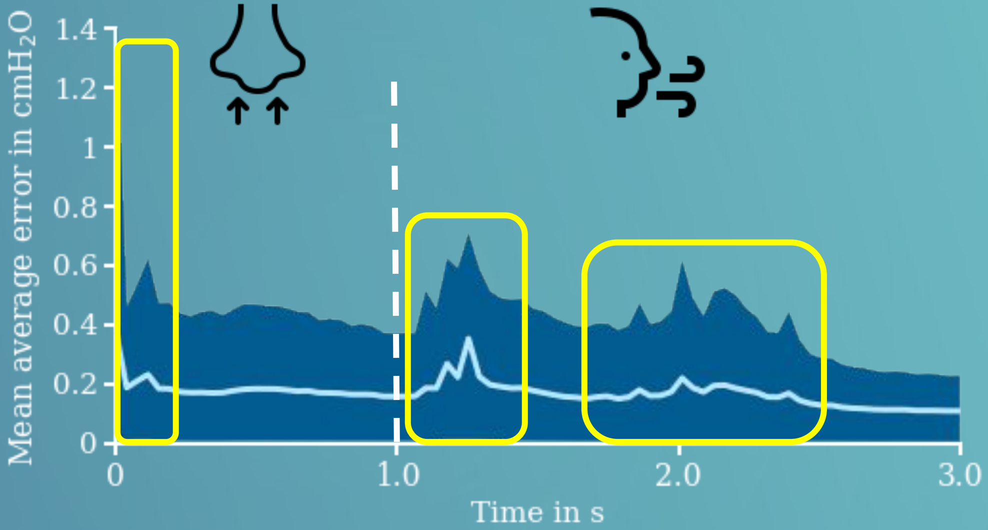


Error analysis



Error analysis





03

Next steps

Outlook

- Investigate large error for relatively unhealthy lung parameters
- Fight overfitting in combined model
- Applying LSTM (time-series model)
- Further optimise current models
- Do more feature engineering (e.g. improve function fit)

Outlook

- Investigate large error for relatively unhealthy lung parameters
- Fight overfitting in combined model
- Applying LSTM (time-series model)
- Further optimise current models
- Do more feature engineering (e.g. improve function fit)

Thank you for your attention!

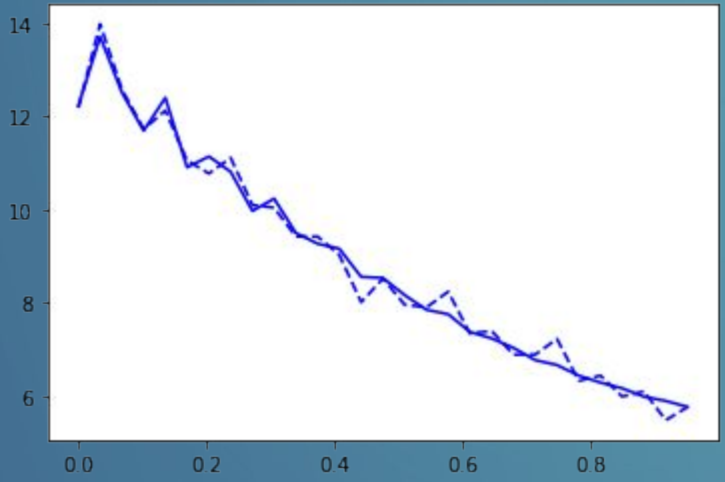
BACKUP SLIDES

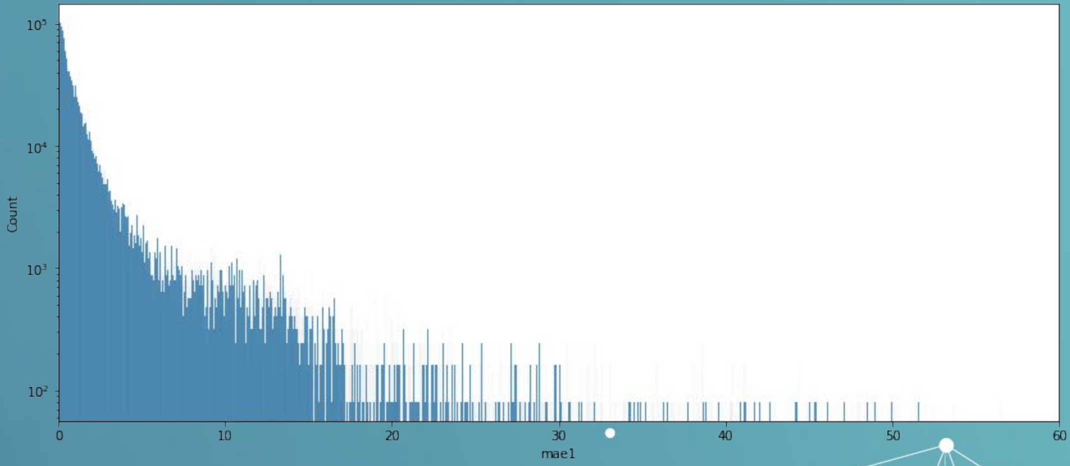
Fitting the input flow

Function of the form

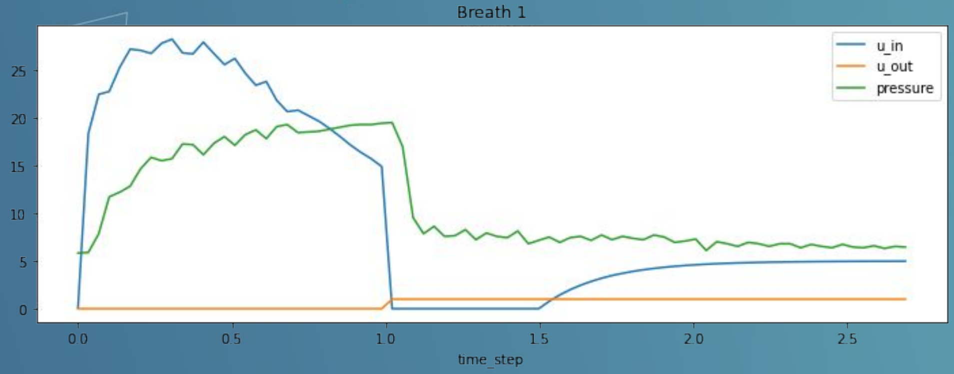
$$f(x, a, b, c, d, e, f, g, h) = a \cdot e^{b \cdot t} \cdot \sin(c \cdot t + d) + e \cdot \sin(f \cdot t + g) + h$$

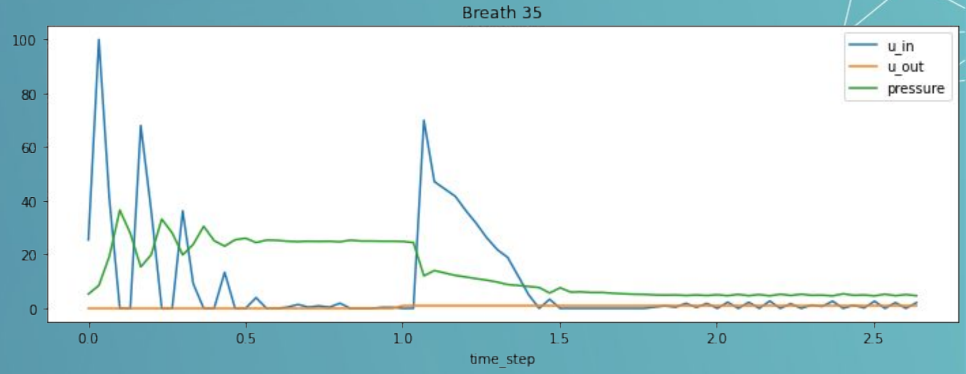
is fitted to input flow in the inspiratory phase.

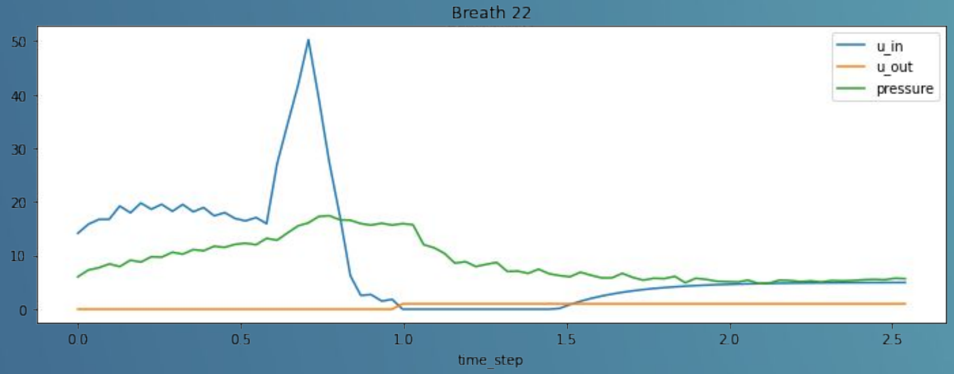


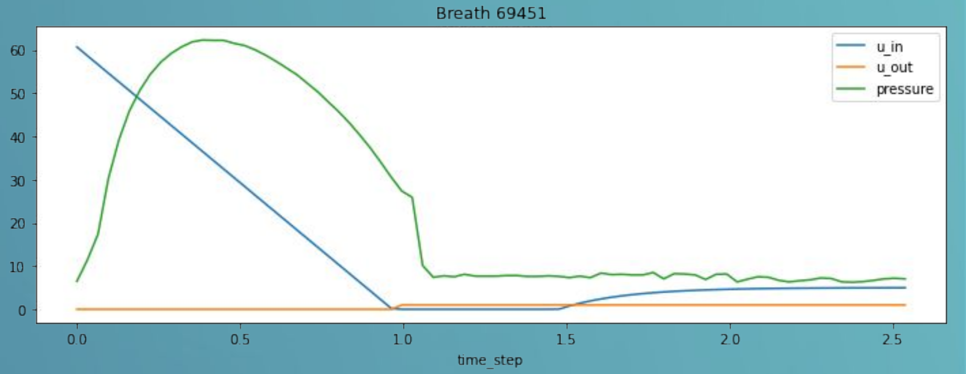


Example breaths









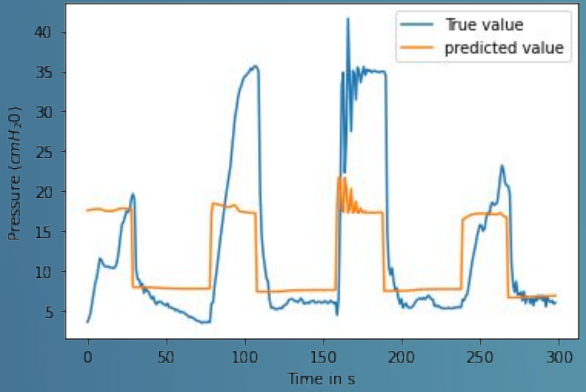
ANN Architecture

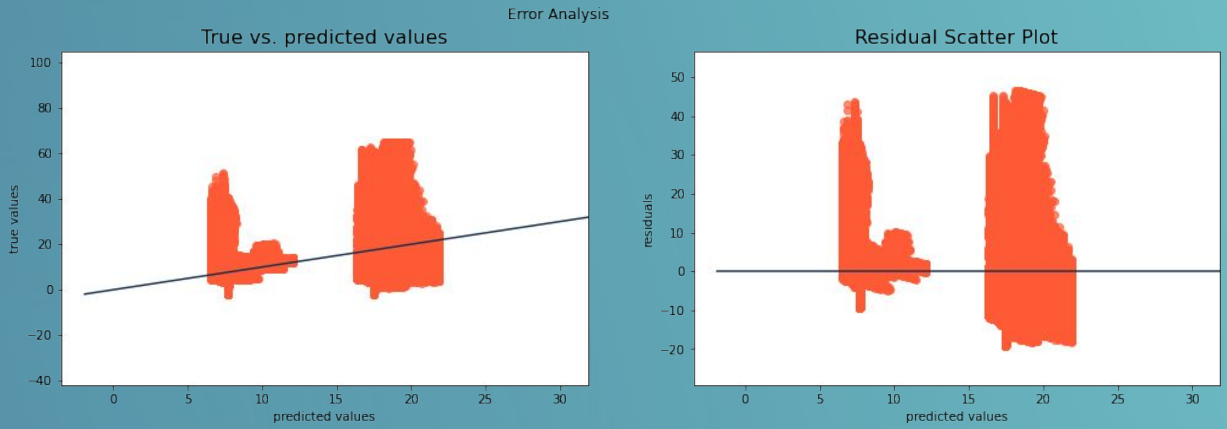
```
#Build a model
def build_and_compile_model4(norm):
    model = keras.Sequential([norm,
        layers.Dense(300, kernel_initializer = 'uniform',
activation='relu'),
        layers.Dense(300, kernel_initializer = 'uniform',
activation='relu'),
        layers.Dense(300, kernel_initializer = 'uniform',
activation='relu'),
        layers.Dense(300, kernel_initializer = 'uniform',
activation='relu'),
        layers.Dense(300, kernel_initializer = 'uniform',
activation='relu'),
        layers.Dense(1)
    ])

    model.compile(loss='mae',
                  metrics='mae',
                  optimizer=tf.keras.optimizers.Adam(0.001))
     return model
```

```
validation_split=0.2,
   verbose=1, epochs=100,
   batch_size = 500,
```

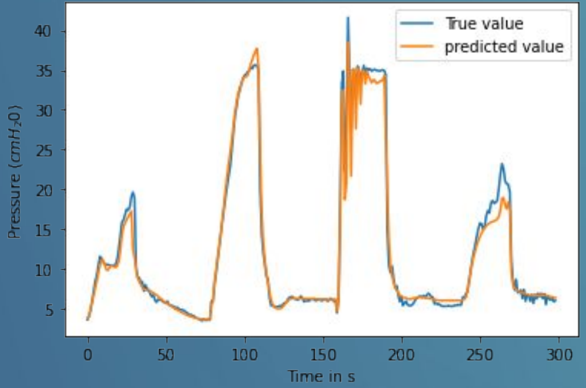
Error analysis for different models (300 timesteps)
Without AR features

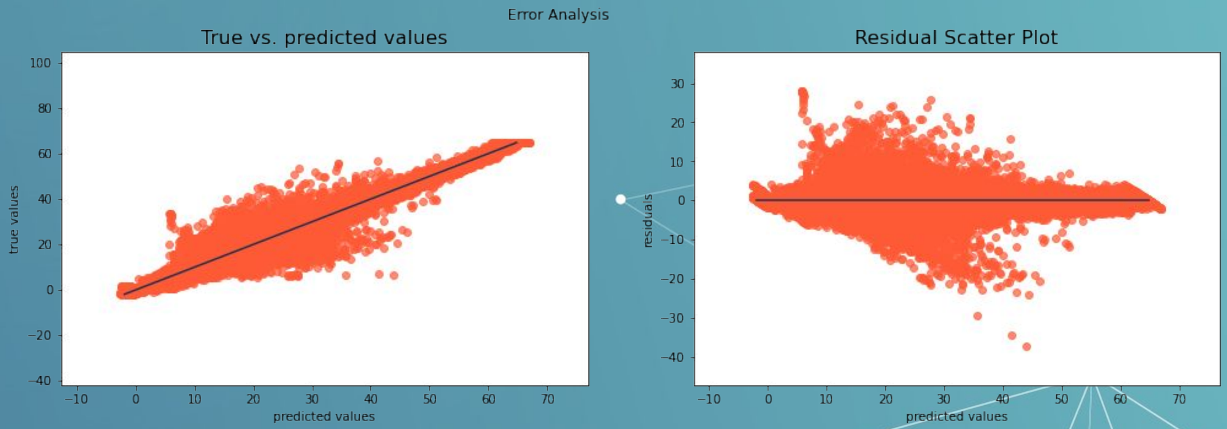
Baseline model



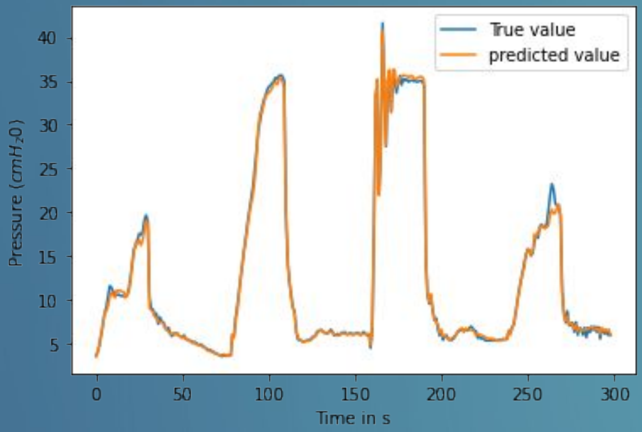


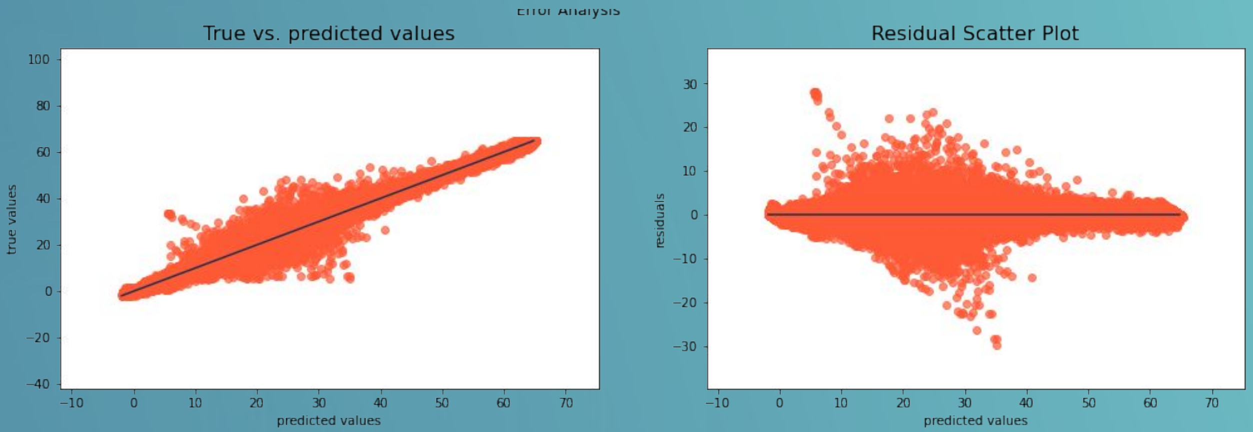
ANN



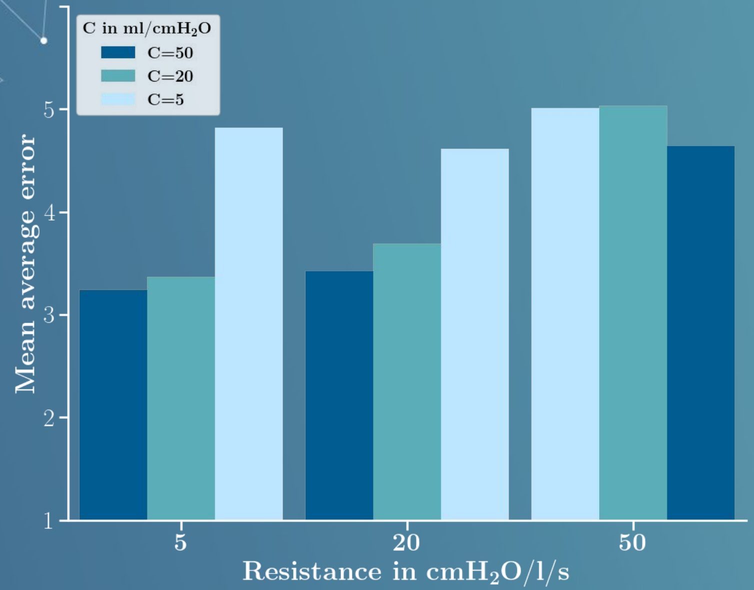


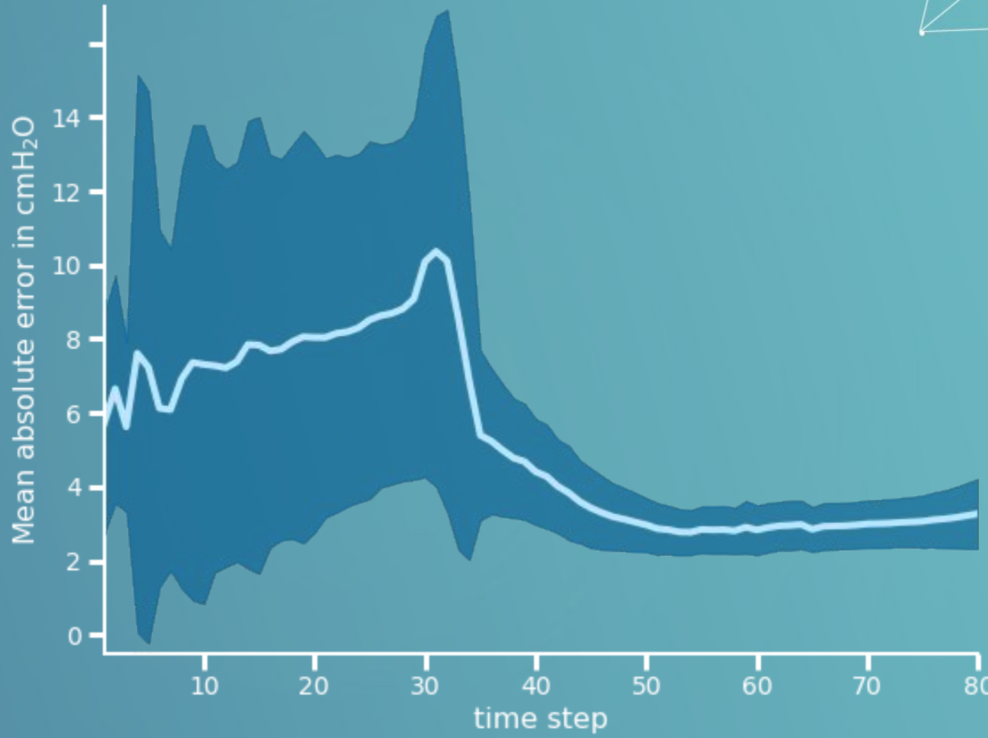
XGBoost


True value
predicted value
Pressure (cmH_2O)
Time in s


True vs. predicted values
true values
predicted values
Residual Scatter Plot
residuals
predicted values

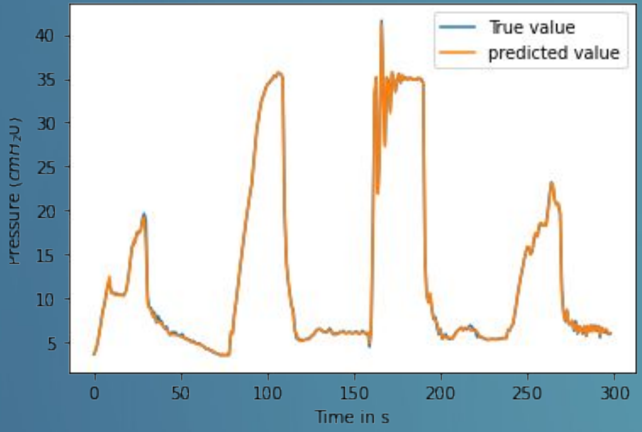
Error analysis

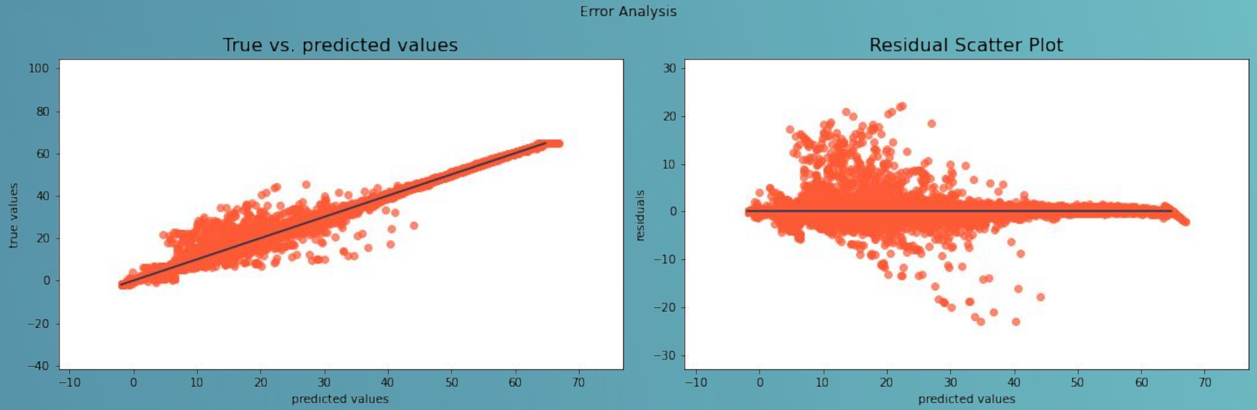




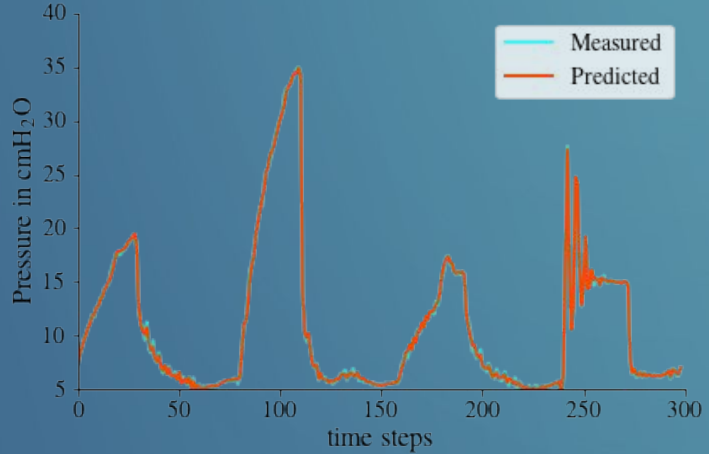
Error analysis for different models (300 timesteps)
With AR features

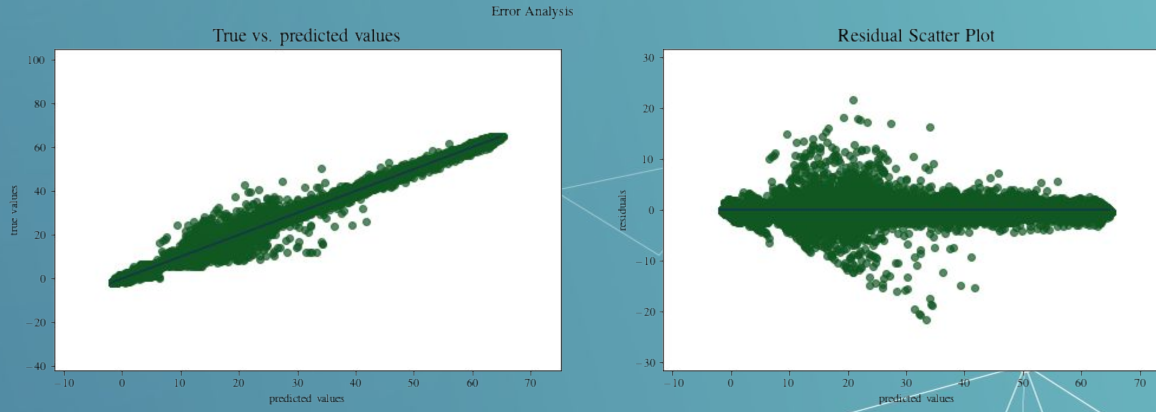
ANN


True value
predicted value
Pressure (cmH_2O)
Time in s


Error Analysis
True vs. predicted values
true values
predicted values
Residual Scatter Plot
residuals
predicted values

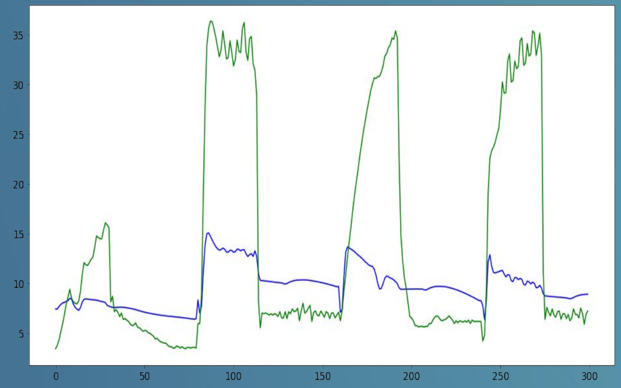
XGB


Measured
Predicted
Pressure in cmH_2O
time steps


Error Analysis
True vs. predicted values
true values
predicted values
Residual Scatter Plot
residuals
predicted values

Error analysis for different autoregressional models (300 timesteps)

ARIMA

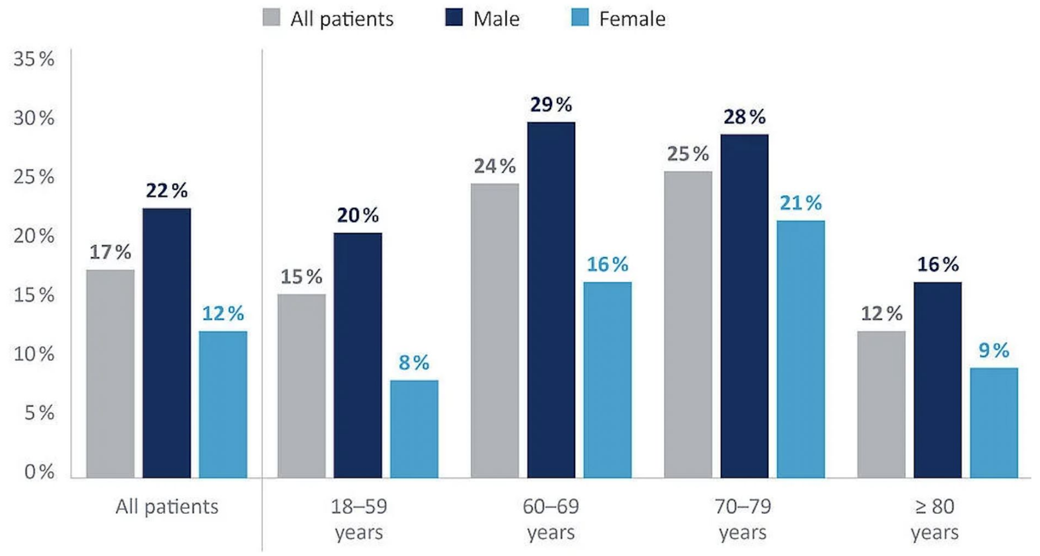


Autoregressional approach

Idea: calculate AR features on the fly and use them as input features

	LR	PF	XGB
Train MAE	1.498	0.464	0.19
Test MAE	1.495	38.51	0.38
Deviation			

Medical situation and demand



Source: Hospitalised patients with COVID-19 with completed hospital treatments admitted to hospital between 26 February and 19 April 2020 based on administrative claims data from the German Local Health Care Funds (Allgemeine Ortskrankenkassen, AOK). Research Institute of the Local Health Care Funds (WIdO).

Rapidly-deployable ventilator needed

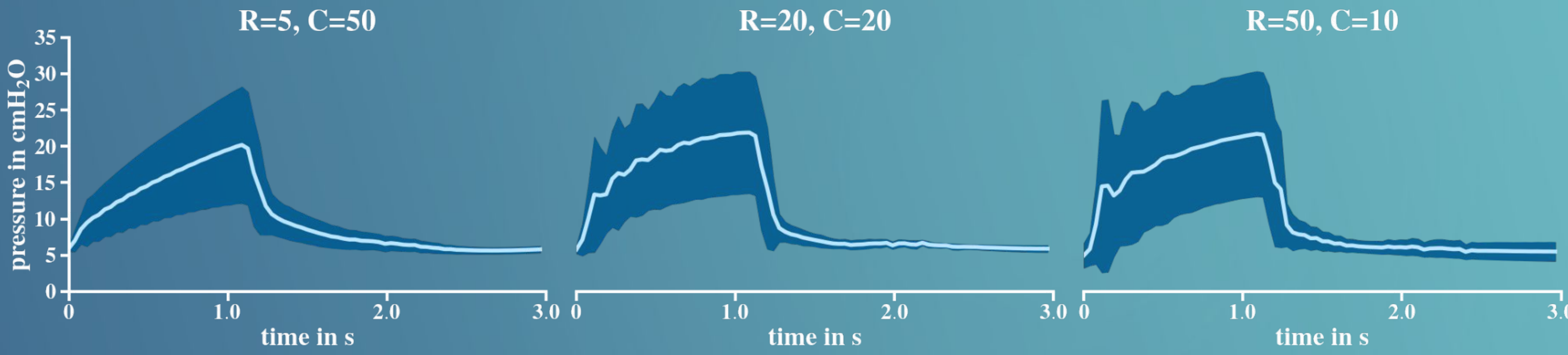
Low cost ventilator needed

Additional features(Kaggle)

```
df=df.query('u_out==0')
df['u_in_cumsum'] = (df['u_in']).groupby(df['breath_id']).cumsum()
df['minus_one']=-1.0
df['plus_one']=1.0
df['exponent']=(df['minus_one']*df['time_step'])/(df['R']*df['C'])
df['factor']=np.exp(df['exponent'])
df['vf']=(df['u_in_cumsum']*df['R'])/df['factor']
df['vt']=0
df.loc[df['time_step'] != 0,
'vt']=df['vol']/(df['C']*(df['minus_one']*df['factor']+df['plus_one']))
df['v']=df['vf']+df['vt']
```

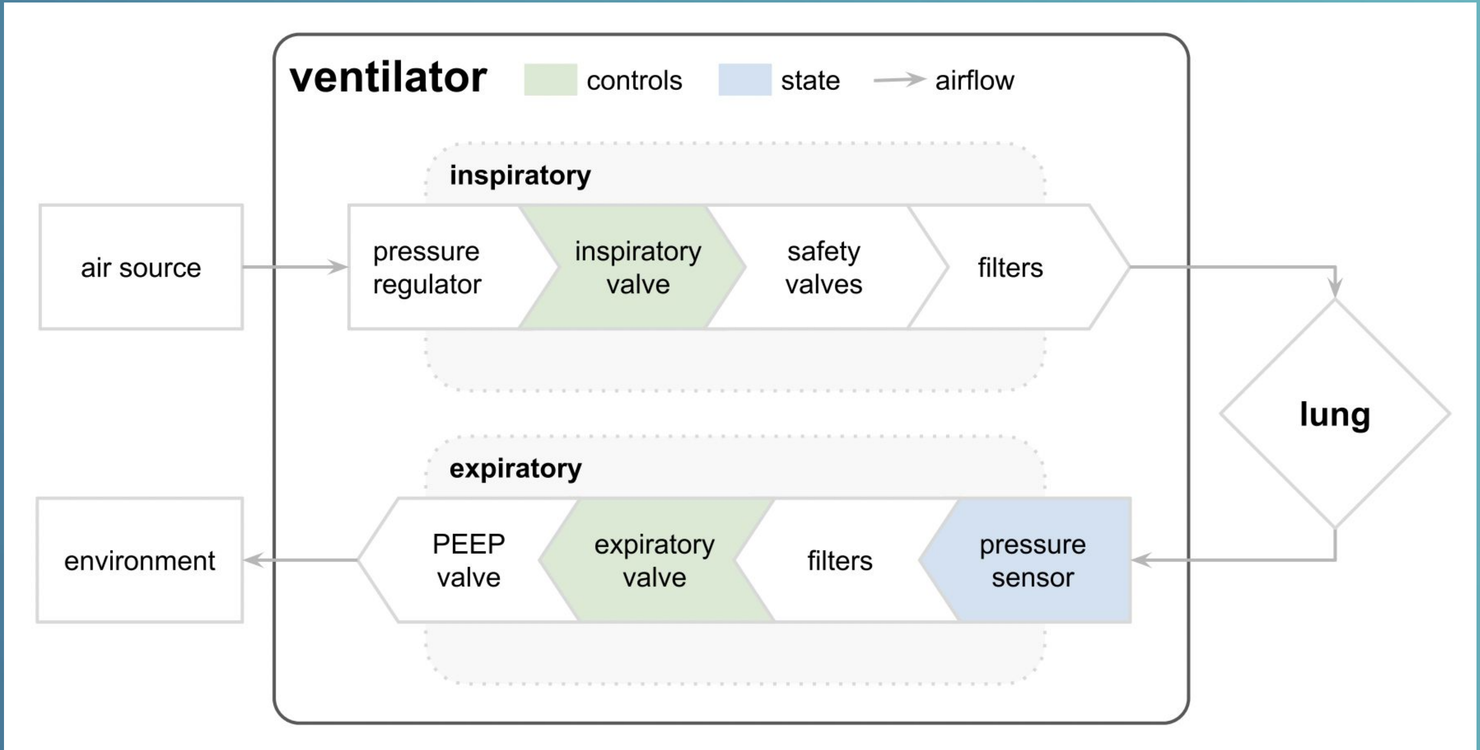
Breathing profiles

- Very different shapes of breaths
- Larger deviations for unhealthy parameters



Unit of R: $cmH_2O/l/s$
Unit of C: ml/cmH_2O

Respiratory circuit



References

PID: https://ai.googleblog.com/2022/02/machine-learning-for-mechanical.html