Presentation at the Central & Eastern European Cheminformatics Conference
“A Read-Across approach to model vapor-liquid equilibria of alkanethiols absorption in aqueous amine solvents”

Dr Kyrylo Klimenko

University of Chemical Technology Prague

07th September 2026

Natural gas purification challenge

- Natural gas consists of both primary components & admixtures

primary components

C_nH_{2n+2}

n = 1-4

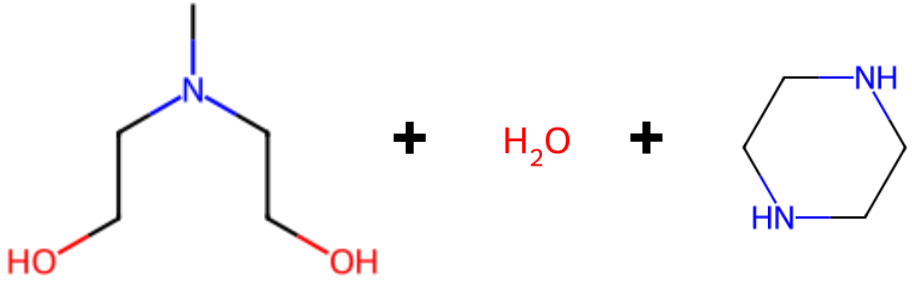
admixtures

CO_2 H_2S COS N_2 $C_nH_{2n+1}SH$ …

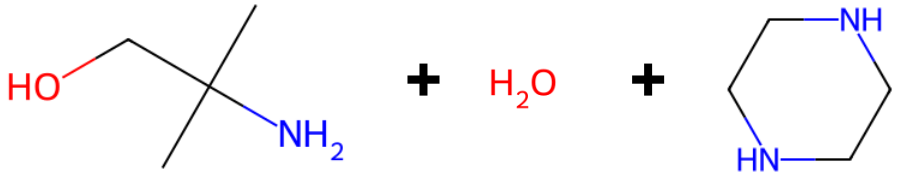
n = 1-2

- The removal of admixtures is often done by absorbance with aqueous amine(s) solvent

HySWEET™



CESAR1



- The gas purification efficiency is sensitive to change in solvent's components

<u>The selection of solution's components and their concentrations are an important business decision</u>

Optimization of solution selection

- Thorough experimental testing of eligible aqueous amine mixtures at different combinations of concentrations and temperatures is not a viable strategy due to the large potential scope of work.

- The modeling of gas absorption allows to have comparative estimation of absorbents' efficiency. The aspects of gas absorption, such as vapor-liquid equilibria of thermodynamic systems that contains >10 components is too complex to efficiently model. **Simpler systems are usually modeled**.

- The modeling of **thiol-amine-H_2O** is still extremely challenging.
  - **ML/QSPR** require more data to build and validate the model than available.
  - **Thermodynamic modeling** requires the knowledge of activity coefficients/fugacities for both 3- & 2-component systems at different temperatures. These values can be estimated from MD, however it increases modeling time and cost significantly.

<u>The modeling of thiol-amine-H_2O may need a new approach</u>

Modeling strategy for thiol-amine-H_2O

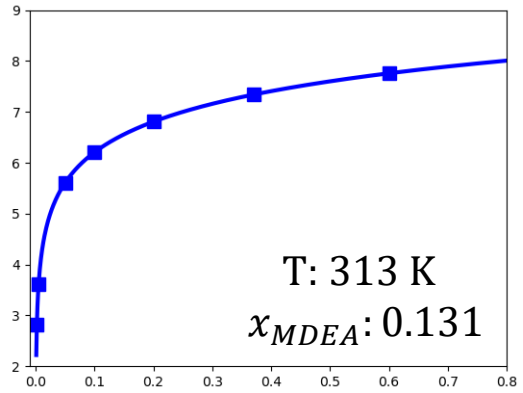
- The alternative approach is to do the modeling technique called **read-across**. In toxicology, read-across is a technique for predicting endpoint information for one substance (target substance), by using data from the same endpoint from (an)other substance. In our case, read-across would mean predicting the VLE of thiols from the VLE of H_2S, when isotherms are characterized by the same temperature and amine concentration.
- The read-across requires a shared mechanism between the target and source systems. Thiols and H_2S, have a similar absorption mechanism, which includes dissociation of species in water and interaction with the amine.

$$\mathrm{R-SH} = \mathrm{R-S^-} + H^+$$

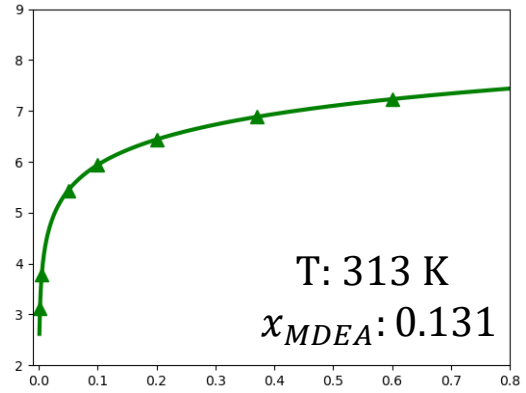
$$\mathrm{R-SH} + R'_3N = \mathrm{R-S^-} + R'_3NH^+ \quad \mathrm{R = H,\ CH_3,\ C_2H_5}$$

Fitted VLE curve

H_2S



Predicted VLE curve



CH_4S

- Since the same temperature and amine concentration become constant under this approach, the difference in VLE prediction between gasses must be reflected by a **molecular descriptor** – a number that represents chemical information of a compound.

Original read-across hypothesis

- Previously an equation has been proposed for the H_2S-MDEA-H_2O VLE assessment:

$$log_{10}(P_{H2S}) = k_{H2S} \times log_2(a) + c_{H2S}$$

a is gas loading ($\frac{moles\ gas\ _{[\text{liquid phase}]}}{moles\ amine}$)

k_{gas}, c_{gas} are curve-fitting parameters

- pKa (experimental) was used as a basis of a molecular descriptor (τ_{thiol}):

$$\tau_{thiol} = \frac{(pKa\ _{[thiol]} - pKa\ _{[H2S]})}{pKa\ _{[thiol]}}$$

- The equation is modified to account for quantitative deviations in the property between the gasses. This equation reflects the tendency of thiols with higher boiling point than H_2S to have lower absorption in aqueous amine

$$log_{10}(P_{thiol}) = (k_{H2S} - \tau_{thiol}) \times log_2(a) + (c_{H2S} + \tau_{thiol})$$

- The efficiency of the RA predictions is estimated by Root-Mean-Square-Error (RMSE) and Mean Absolute Percentage Error (MAPE):

$$RMSE_t = \sqrt{\frac{1}{f}\sum_{i=1}^{f}(\widehat{y_{it}} - y_{it})^2}$$

y_{it} – *observed* P_{gas}; $\widehat{y_{it}}$ – *predicted* P_{gas};

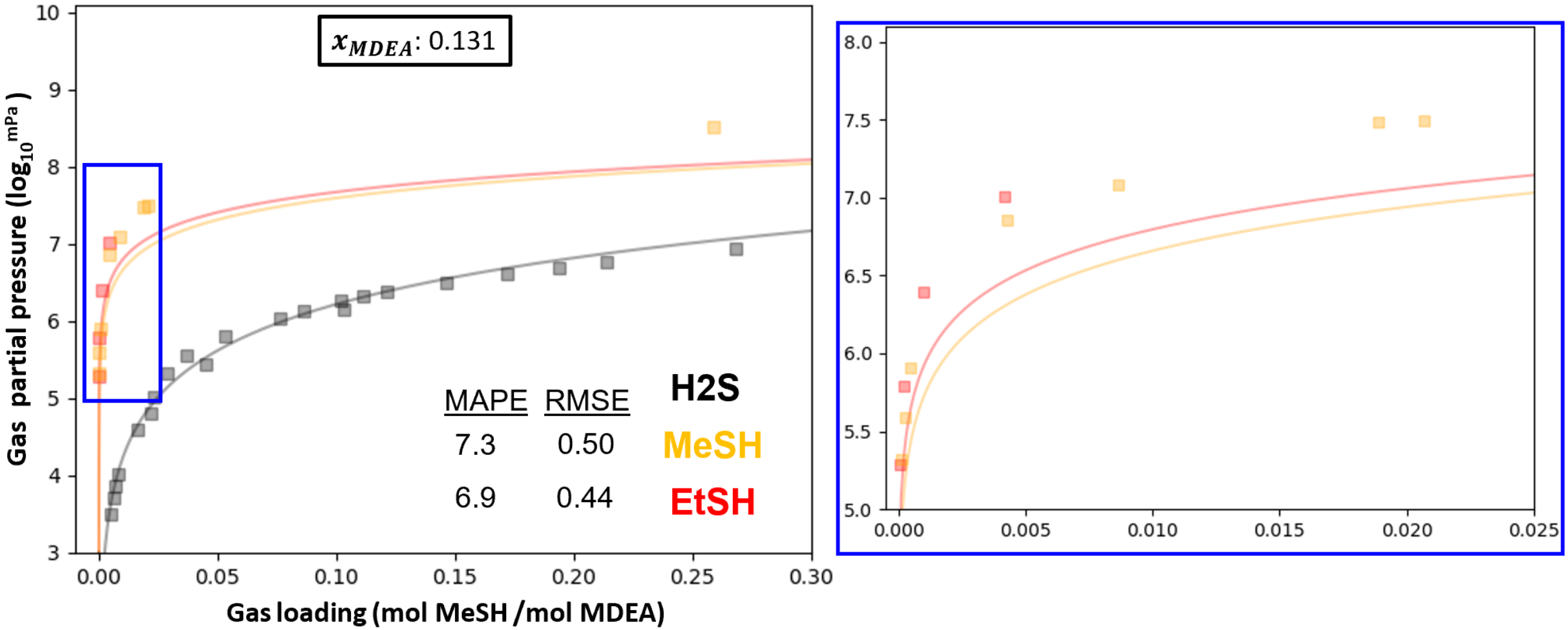
i – *data point;* t – *isotherm;*

f – *num. of points per isotherm;*

$$MAPE_t = 100\frac{1}{f}\sum_{i=1}^{f}\left|\frac{y_{it} - \widehat{y_{it}}}{y_{it}}\right|$$

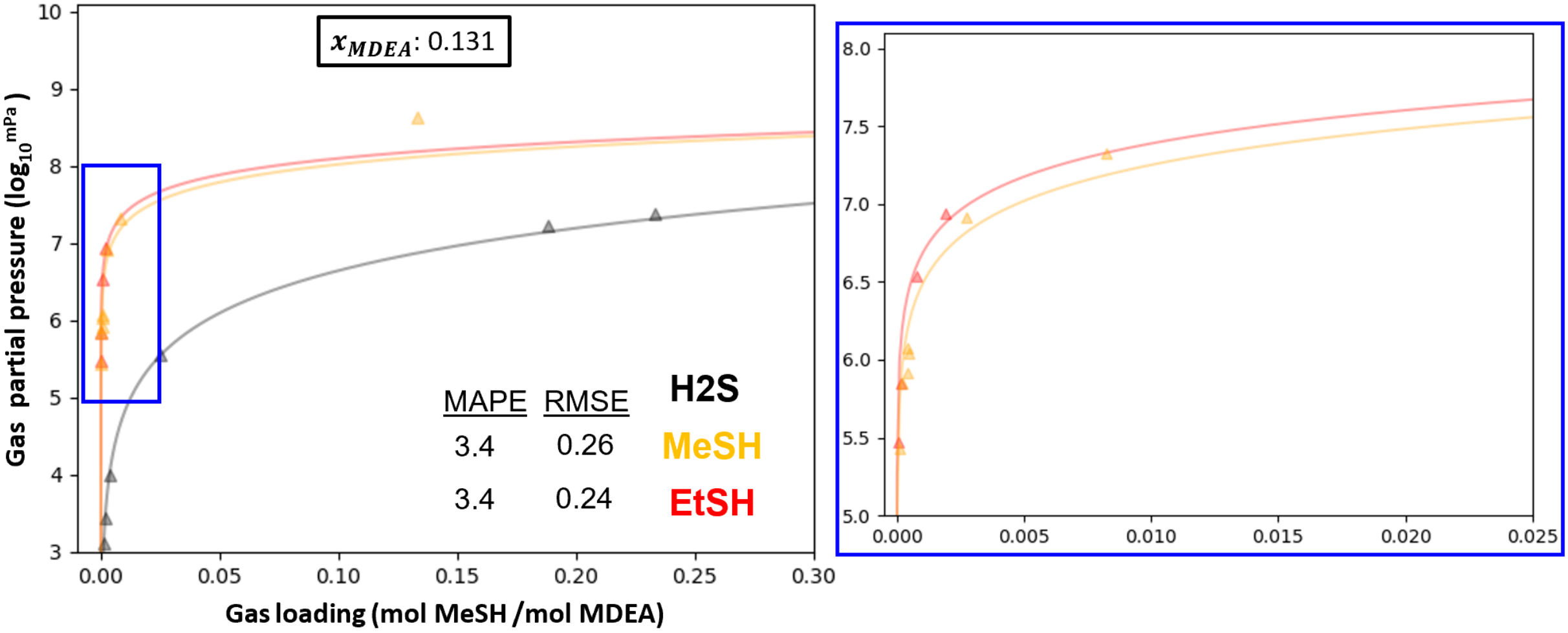
Original read-across hypothesis validation 1

T: 313 K



Original read-across hypothesis validation 2

T: 343 K



Modified Read-across hypothesis

- Modified Hypothesis: In **weighted average approach**, two source isotherms that have the same temperature as the target isotherm and amine concentrations that are higher and lower than the target isotherm's are selected. Then, the concentrations of the source isotherms are scaled linearly, the scaling is applied to the target isotherm's concentration to obtain **the weight**. **The weight** is then applied to the curve coefficients to produce **the weighted source isotherm to be used in RA.**

$$weight = \frac{x^{target} - x_{min}^{source}}{x_{max}^{source} - x_{min}^{source}}$$

$$\Delta k = k_{max(x_{MDEA}^{source})} - k_{min(x_{MDEA}^{source})}$$

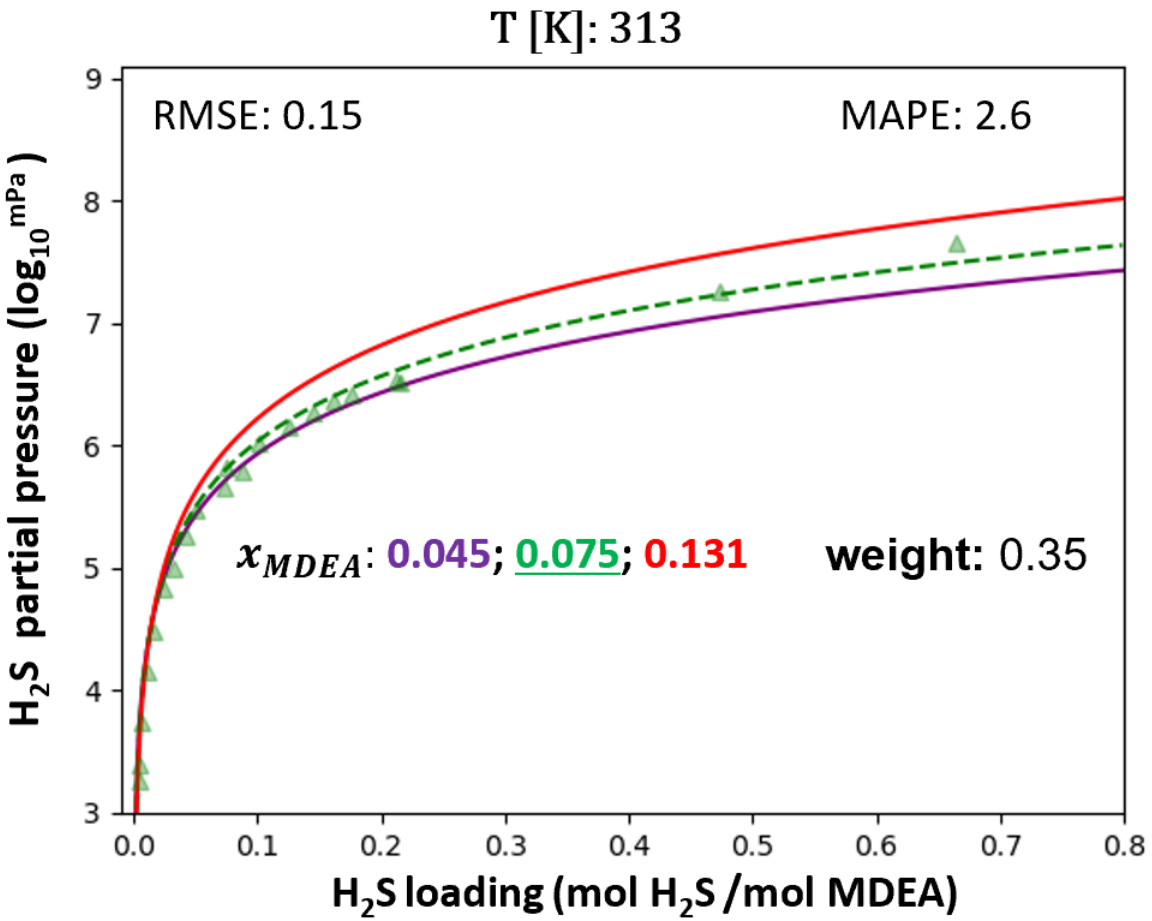
$$\Delta c = c_{max(x_{MDEA}^{source})} - c_{min(x_{MDEA}^{source})}$$

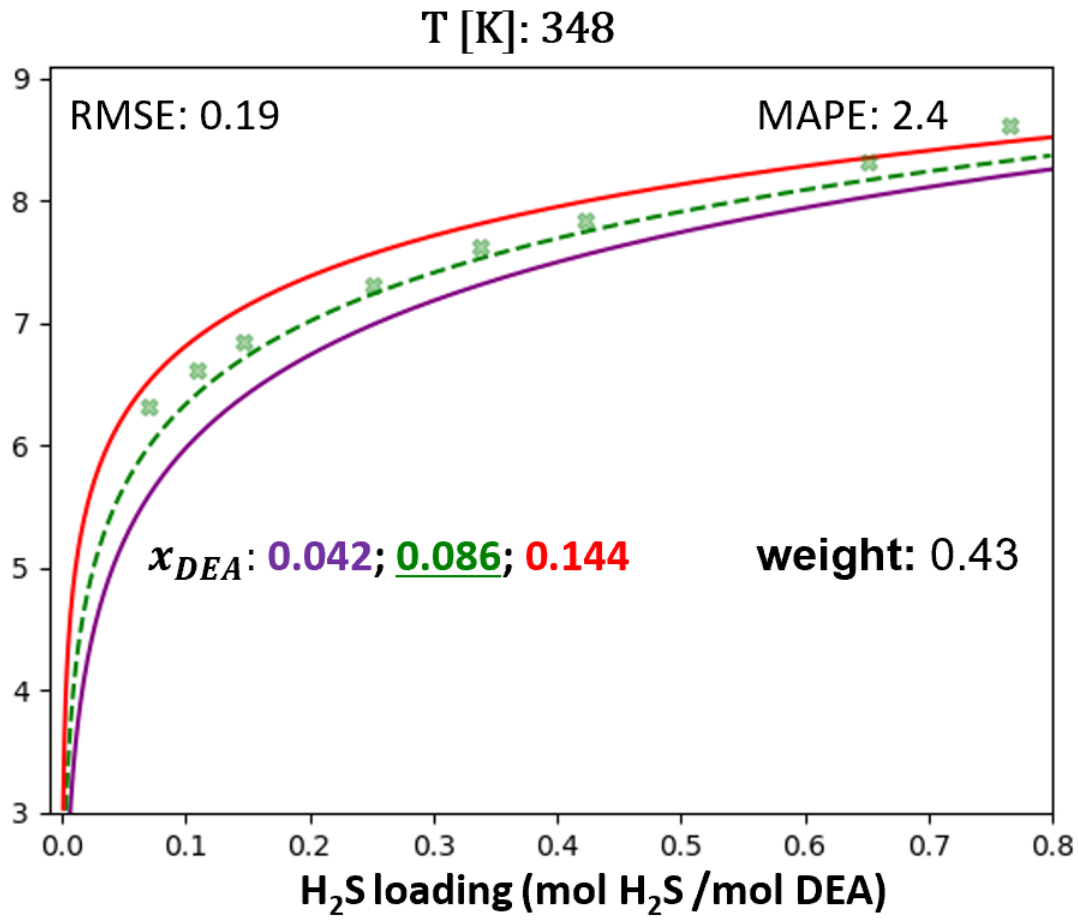
$$k_{weight} = weight*(\Delta k) + k_{min(x_{MDEA}^{source})}$$

$$c_{weight} = weight*(\Delta c) + c_{min(x_{MDEA}^{source})}$$

$$log_{10}(P_{thiol}) = (k_{weight} - \tau_{thiol}) \times log_2(a) + (c_{weight} + \tau_{thiol})$$

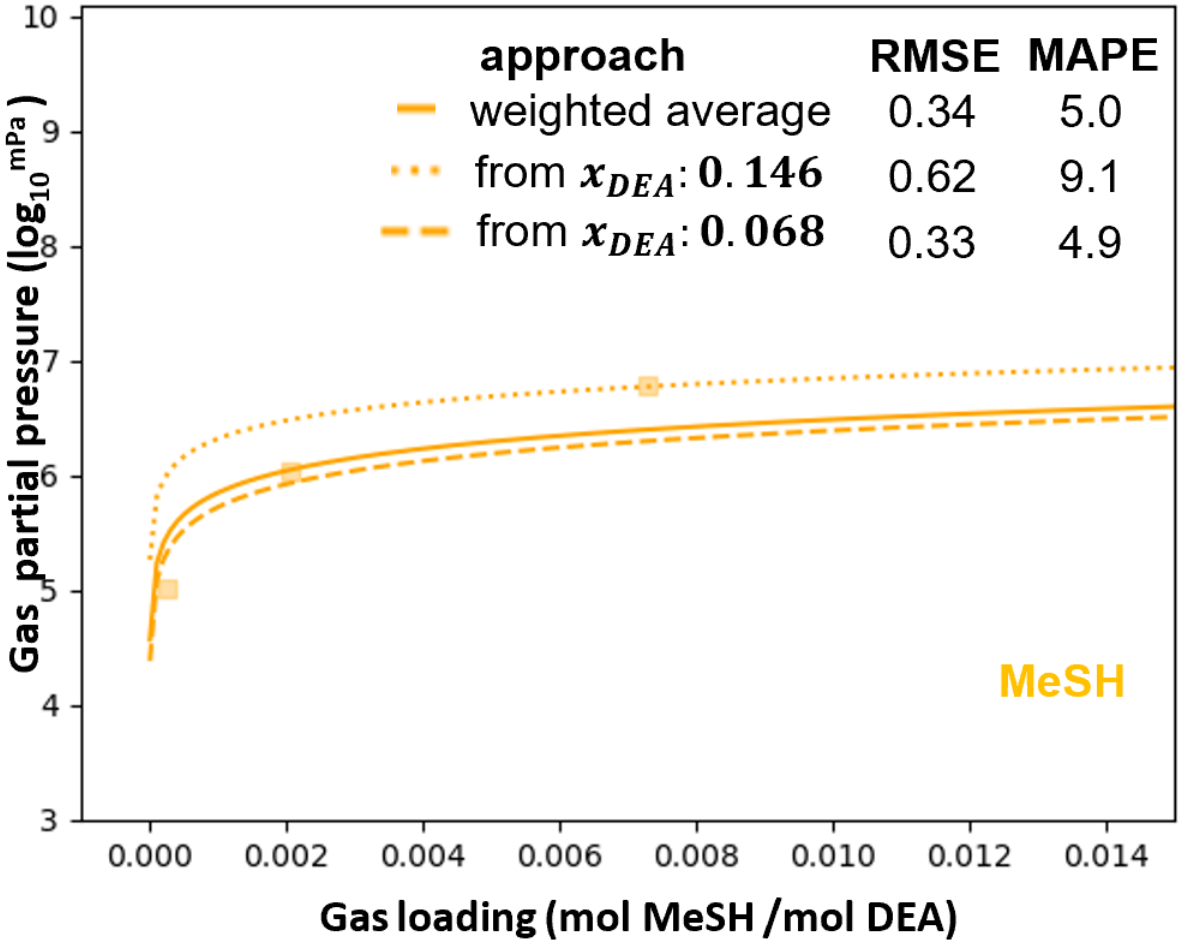
Validation of H_2S-amine-H_2O read-across

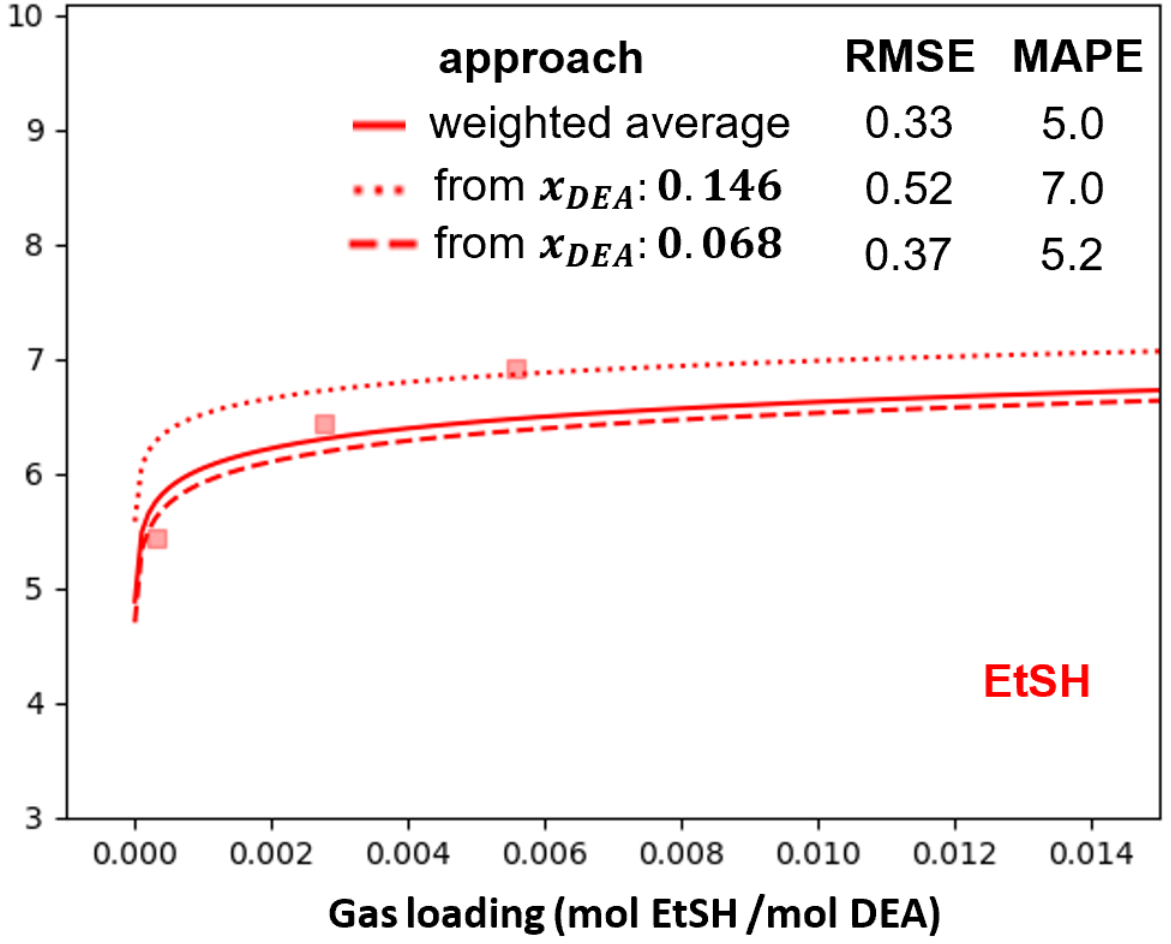




Modified read-across hypothesis validation 1

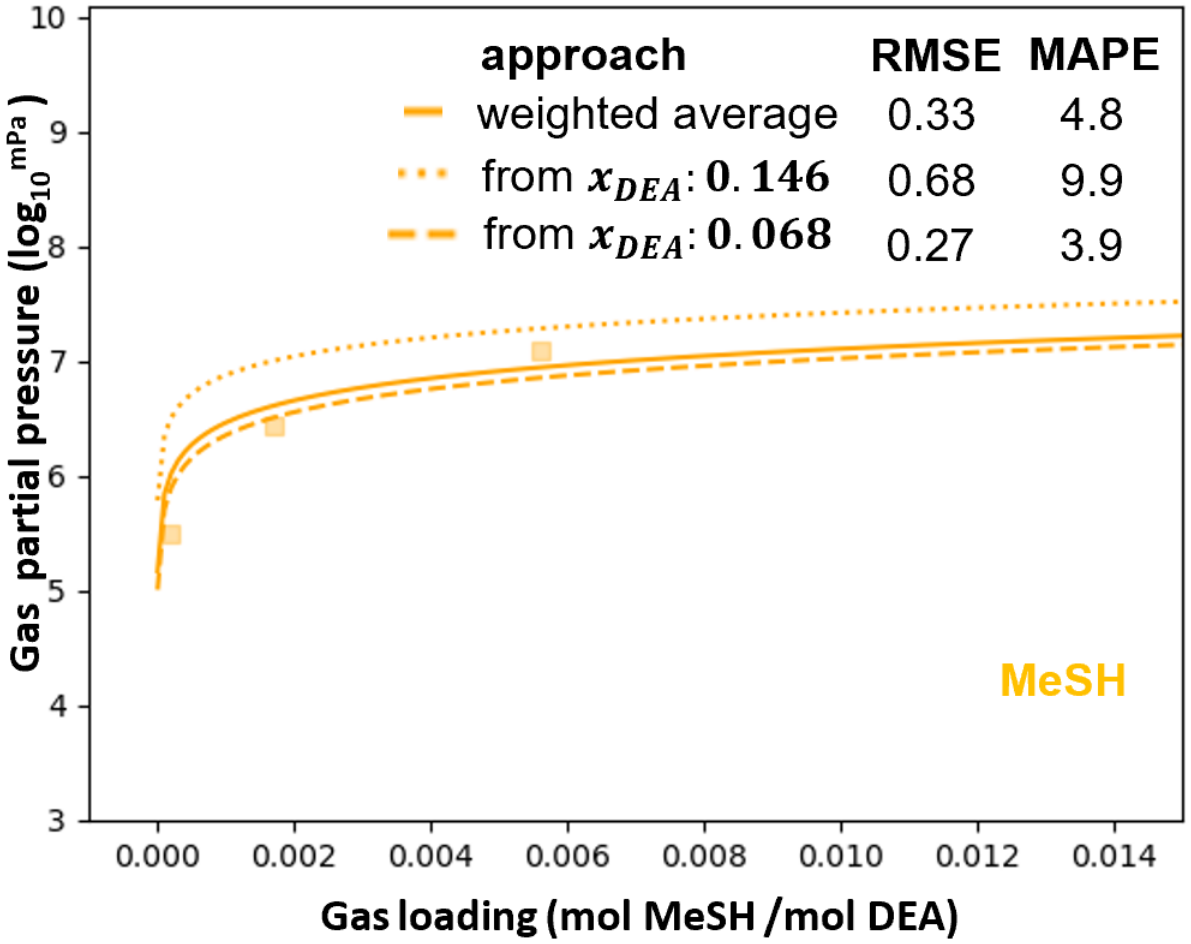
T: 313 K | x_{DEA} : 0.084

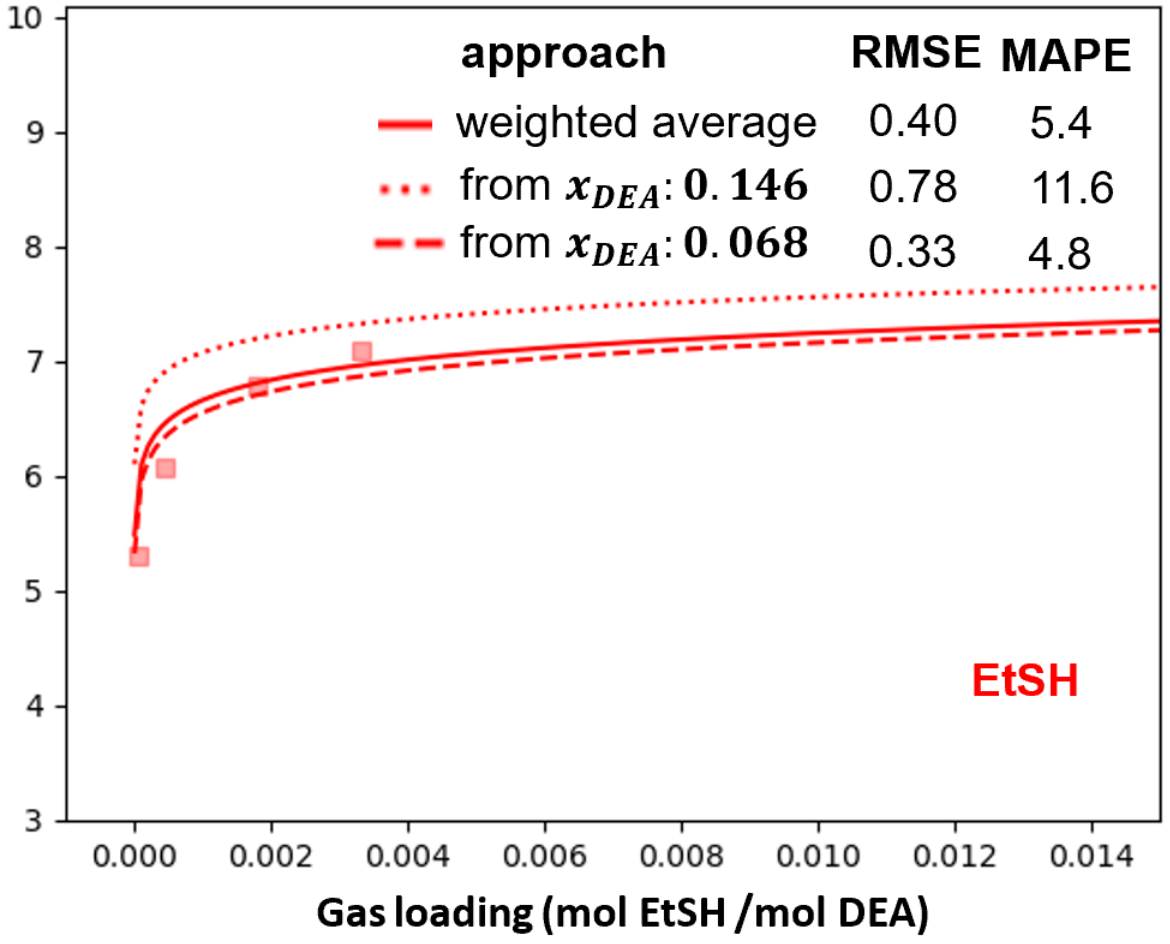




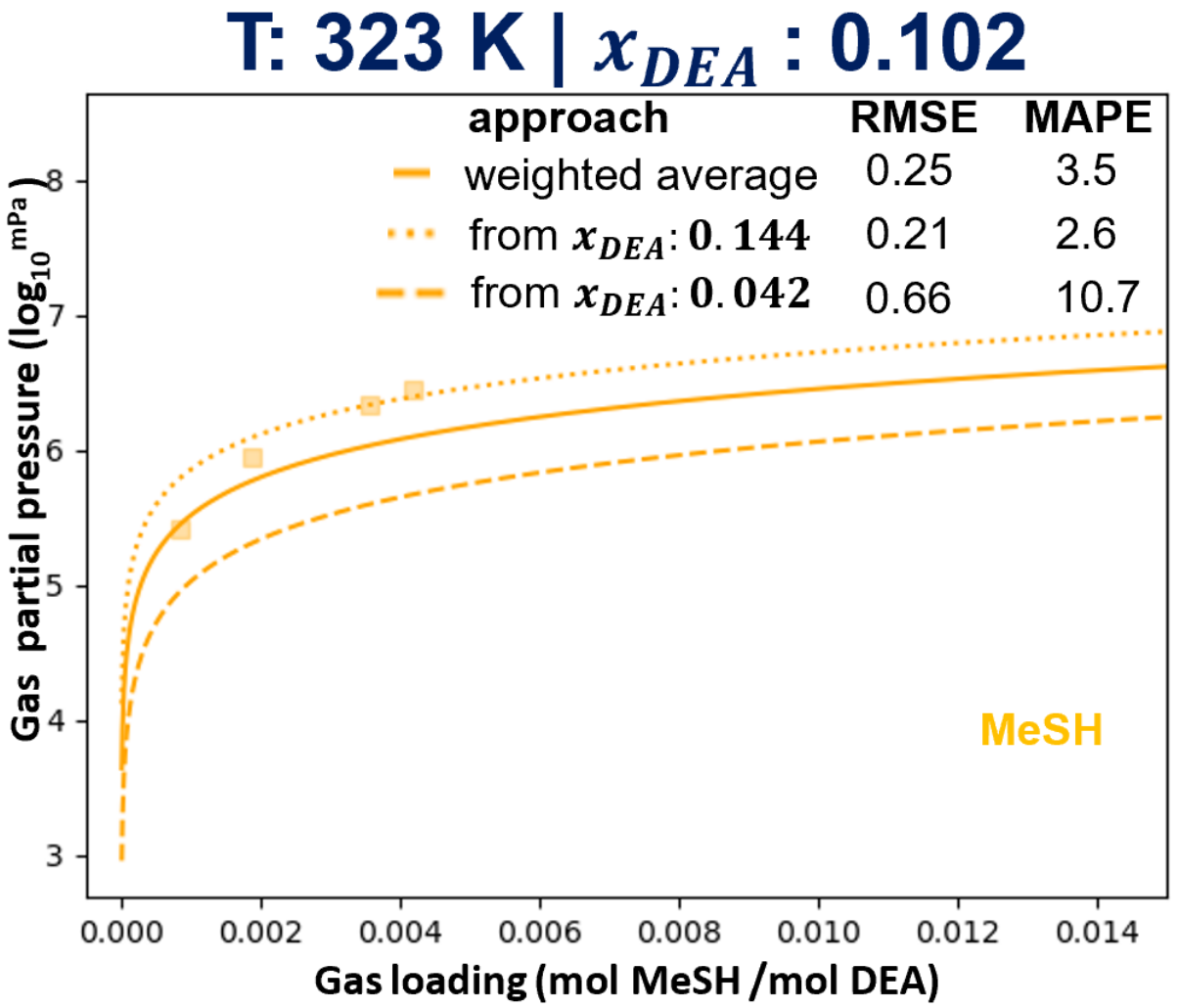
Modified read-across hypothesis validation 2

T: 343 K | x_{DEA} : 0.084

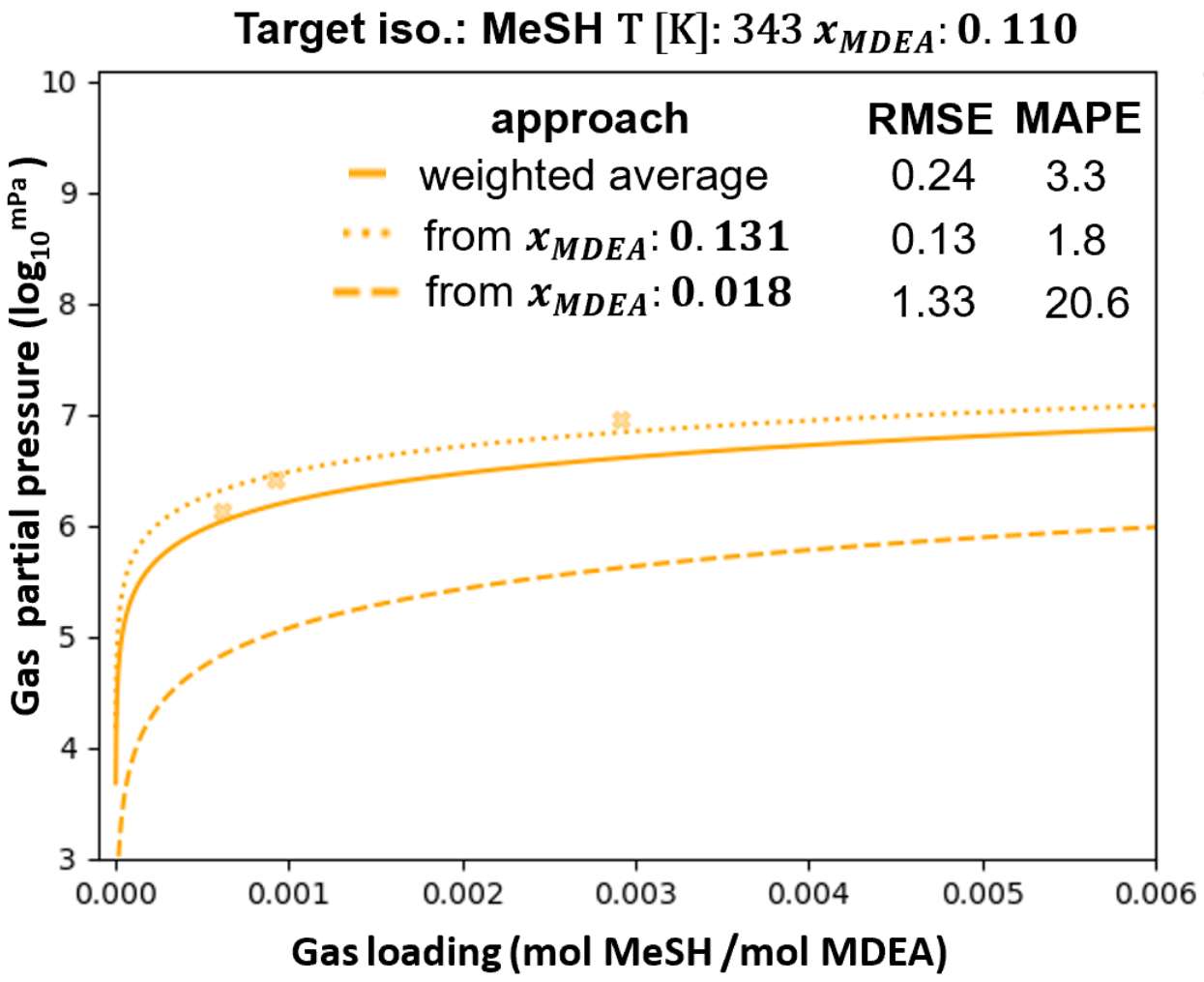


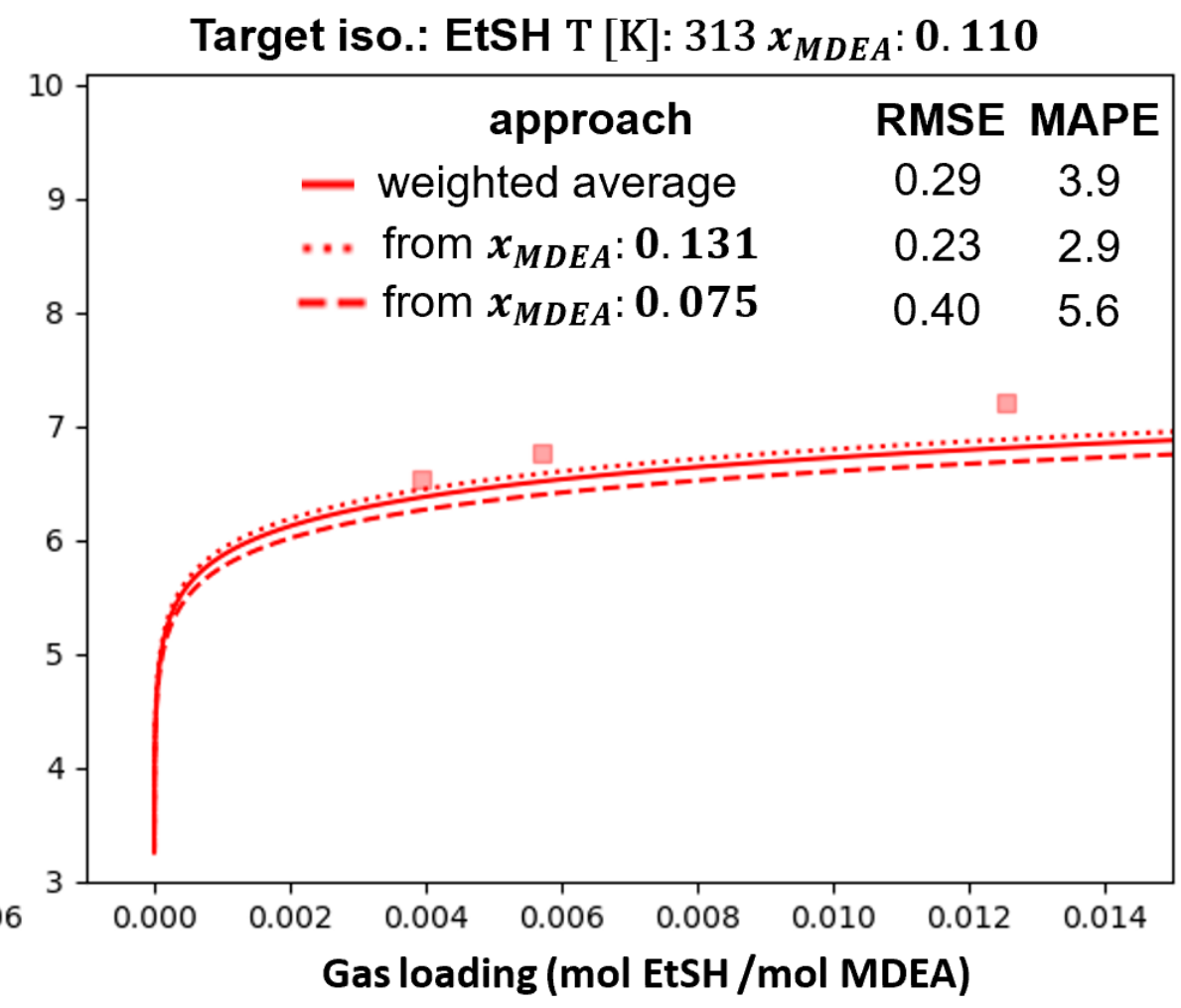


Modified read-across hypothesis validation 3



Modified read-across hypothesis validation 4





Summary

Results

- The read-across approach to model the VLE of thiol-amine-H_2O thermodynamic systems was developed and validated on the systems with aqueous MDEA and DEA at different temperatures and concentrations of the amine

Conclusions

- Applying RA methodology allows us to overcome the issue of having very little data for thiol-amine-H_2O systems that prevents successful ML modeling of VLE
- Developed equation can predict the VLE of thiol-amine-H_2O isotherms even at very low end of the gas loading range due to the continuous nature of this regression

Acknowledgements

University of Strasbourg

Prof Alexandre Varnek

Prof Dragos HORVATH

TotalEnergies

Dr Frederick DE MEYER

Dr Eric CLOAREC

Thank you for your attention!