intelligent systems TU Graz

SCIENCE ▪ PASSION ▪ TECHNOLOGY

TU Graz

Signal Processing and Speech Communications Laboratory ▪ Inffeldgasse 16C/EG, 8010 Graz, Austria ▪ fuchs@tugraz.at

Data-Scarce Condition Modeling requires Model-Based Prior Regularization

Nikolaus Mutsam Alexander Fuchs Fabio Ziegler Franz Pernkopf

Abstract

- Intermediate target values of the wear time series are not available and need to be estimated from an accumulated, end-of-process measurement.
- An iterative least-squares approach is used to estimate the hidden intermediate targets based on linear regression.
- The target estimates are used as regularization prior for DNN training.
- Experiments use real-world data of refractory wear processes in steel production.

Model-Based Prior Regularization

- Neural networks often struggle with convergence and overfitting on scarce and noisy data, but linear models cannot adequately represent complex relationships.
- We use an **Iterative Least Squares (ILS)** method to estimate intermediate targets from linear models.
- Target estimates are used to train deep neural networks (DNN).
- Estimated and measured targets are balanced by a regularization term.

- Step 1: Run ILS algorithm to generate target estimates $\mathbf{h}^T$
- Step 2: Use ILS regularization during DNN training:

$$\mathcal{L} = \alpha \frac{1}{K} \sum_{k=1}^{K} (y_c - \bar{y}_c)^2 + (1-\alpha) \frac{1}{K} \sum_{k=1}^{K} \frac{1}{N_c} \sum_{n=1}^{N_c} (h_c^T[n] - \bar{h}_c[n])^2$$

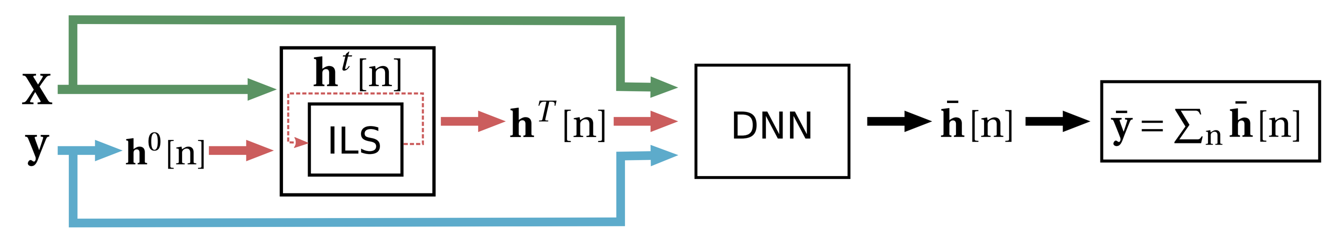


Figure 1: DNN bootstrapping process for the hidden targets $\mathbf{h}$.

ILS

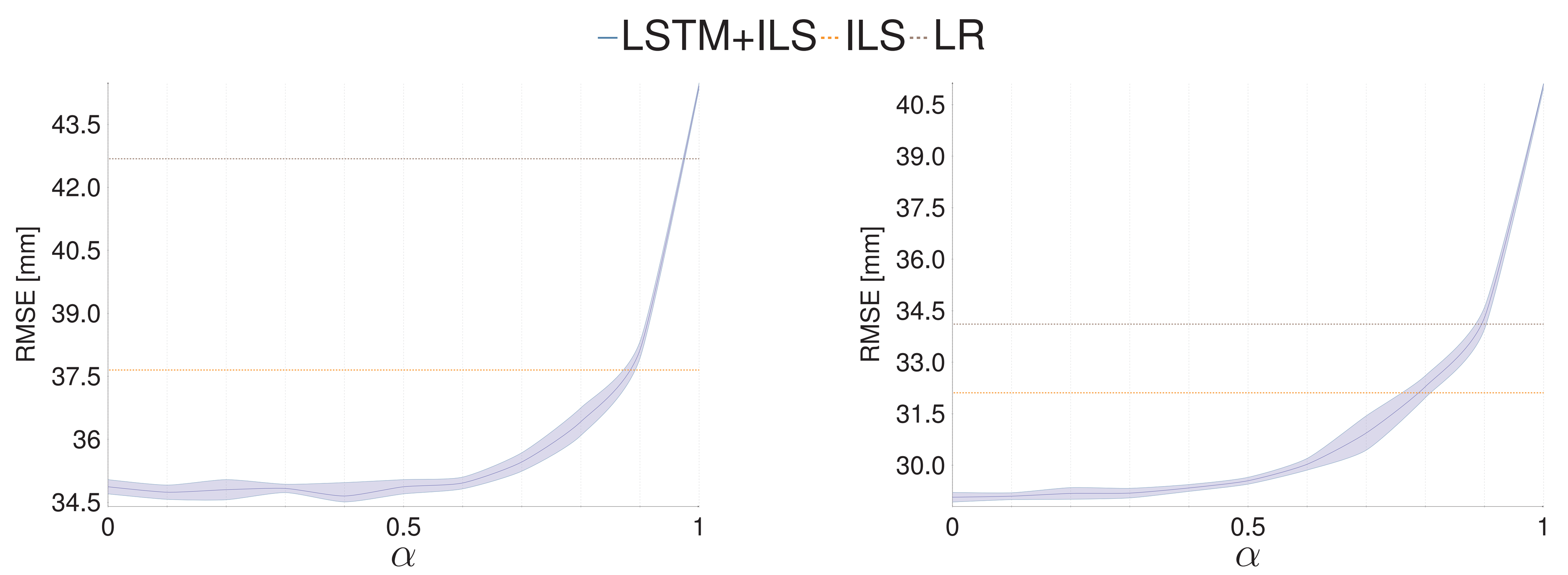
Input: $\boldsymbol{X}_c, \boldsymbol{y}_c\ \forall c$
Parameters: T (T: # iterations)
Output: $\boldsymbol{h}_c^T\ \forall c$ (estimated targets)
for $c = 1$ **to** C **do**
  Initialize: $\boldsymbol{h}_c^0 \leftarrow h_c^0[n] = \frac{y_c}{N_c}; \forall n$
end for
for $t = 1$ **to** T **do**
  for $c = 1$ **to** C **do**
    (least squares fit)
    $\boldsymbol{X}_I \leftarrow \boldsymbol{x}_i[n]\ ; \forall n, i \in \boldsymbol{C} \setminus c$
    $\boldsymbol{h}_I \leftarrow h_i^{t-1}[n]\ ; \forall n, i \in \boldsymbol{C} \setminus c$
    $\hat{\boldsymbol{\beta}}^t \leftarrow (\boldsymbol{X}_I^T \boldsymbol{X}_I)^{-1} \boldsymbol{X}_I^T \boldsymbol{h}_I$
    (predictor step)
    $\boldsymbol{X}_V \leftarrow \boldsymbol{x}_c[n]$
    $\hat{\boldsymbol{h}}_{\boldsymbol{c}} \leftarrow \boldsymbol{X}_V \hat{\boldsymbol{\beta}}^t$
    (corrector step)
    $b_c \leftarrow y_c - \sum_{n=1}^{N_c} \hat{h}_c[n]$
    $h_c^t[n] \leftarrow \frac{1}{N_c} b_c + \hat{h}_c[n]; \forall n$
    $h_c^t[n] \leftarrow \begin{cases} h_c^t[n], & \text{if } b_{min} \leq h_c^t[n] \\ b_{min}, & \text{else} \end{cases} ; \forall n$
    $\boldsymbol{h}_c^t \leftarrow h_c^t[n]\ ; \forall n$
  end for
end for
Return $\boldsymbol{h}_c^T\ \forall c$

Conclusion

- We propose a regularization prior for wear prediction of refractory material in metallurgical vessels.
- Use of a two step estimation principle: first estimate hidden targets with our proposed ILS approach, second estimate DNN parameters using hidden targets as regulization prior.
- Models trained with ILS-estimated targets outperform the same models using aggregations in the feature domain or normalizing overall measurements.
- Bootstrapping DNNs with estimates of intermediate targets is beneficial for small and noisy data sets.
- Our proposed regularized DNNs show substantial improvement over DNNs trained solely with MSE.

Results

- Variation of regularization between $\alpha = 0$ (only ILS-estimated targets) and $\alpha = 1$ (only measurements).
- Blending intermediate target estimates into the loss function clearly improves prediction accuracy.



(a) heart (b) outlet

Figure 2: Influence of α on the LSTM model performance.

- **Linear Regression (LR):** Linear regression with uniform intermediate targets
- **Iterative Least Squares (ILS):** Linear regression with iteratively estimated intermediate targets
- **Deep Neural Network (DNN) Architectures (ILS bootstrapped):** LSTM, 1D CNN, 2D CNN

Table 1: RMSE performance in [mm] and standard deviation for the DNN models.

	LSTM	1D CNN	2D CNN	ILS	LR
Metal zone	**17.33 ± 0.17**	17.56 ± 0.20	17.40 ± 0.07	19.48	22.66
Slag zone	24.54 ± 0.17	**23.25 ± 0.22**	24.66 ± 0.07	26.34	31.07
Heart	34.65 ± 0.14	35.01 ± 0.20	**34.40 ± 0.30**	37.65	42.68
Inlet	**22.24 ± 0.17**	23.21 ± 0.26	22.61 ± 0.63	24.39	29.02
Outlet	**29.07 ± 0.14**	29.84 ± 0.10	29.15 ± 0.41	32.11	34.33

Christian Doppler Forschungsgesellschaft