Manual editing significantly improves the accuracy of motor unit recruitment threshold and discharge rate assessment from synthetic high-density surface electromyograms

Nina Murks[1], Jakob Škarabot[2], Matej Kramberger[1], Gašper Sedej[1], Tamara Valenčič[2], Christopher D Connelly[2], Haydn Thomason[2], Matjaž Divjak[1], Aleš Holobar[1],

[1] Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia
[2] School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK
E-mails: nina.murks@um.si, ales.holobar@um.si

Manual editing significantly improves the accuracy of motor unit recruitment threshold and discharge rate assessment from synthetic high-density surface electromyograms

Abstract. *This study examined the influence of manual editing on the accuracy of motor unit (MU) metrics derived from high-density surface EMG. Seven operators with varying levels of experience manually edited synthetic datasets from the Soleus and Biceps Brachii. We then calculated recruitment threshold (ReTh), derecruitment threshold (DeTh), discharge rate (DR), recruitment discharge rate (ReDR), and derecruitment discharge rate (DeDR) before and after editing. The effects of operator's experience, contraction level, initial pulse-to-noise ratio (PNR), signal-to-noise ratio (SNR), muscle type, and editing status were analyzed.*

Manual editing significantly improved accuracy only for ReTh and DR, whereas with other metrics, it only had a minor, insignificant effect. Higher initial PNR consistently reduced errors across all metrics. Contraction level strongly influenced all metrics, whereas muscle type and SNR had minimal impact—SNR was significant only for DeTh, and muscle type only for ReTh. Operator's experience did not significantly affect any metric. Overall, these results indicate that manual editing is critical for reliable estimation of ReTh and DR, but less essential for DeDR, particularly when initial PNR is high.

1 Introduction

Automatic decomposition methods for high-density electromyography (hdEMG) lack awareness of the experimental protocol or subject-specific factors [10], often making manual editing of decomposition results necessary. Manual editing is the process where operators carefully inspect all decomposed motor units (MUs) and correct errors that arise during the automatic segmentation. The algorithm may misclassify MU pulses as noise, or vice versa, due to the interference from other MUs or background noise. During manual editing, operators fix these segmentation errors by adding missing pulses or removing duplicates. They also eliminate low-quality MUs—those dominated by noise or with too few discharges to be reliable. In cases where different MUs have very similar action potential shapes, the algorithm may incorrectly merge them into a single MU. Operators then need to separate these merged MUs, ensuring their accurate and complete representation.

Although the importance of manual editing has been acknowledged, there is still a limited understanding of how this editing process affects commonly used MU metrics. Typically, MU recruitment is assessed through recruitment threshold (ReTh) and recruitment discharge rate (ReDR) [12], whereas MU derecruitment is evaluated using derecruitment threshold (DeTh) and derecruitment discharge rate (DeDR) [13].

Previous research [5], [11] has shown that effective manual editing demands substantial expertise, leading to the development of various tutorials [10], [4]. Despite these resources, editing results can vary significantly between operators, influenced by factors such as pulse-to-noise ratio (PNR) [14], operator's experience, and occasionally the muscle under study. At the same time, automatic decomposition is constantly improving, gradually decreasing the need for manual editing.

However, the implications of these differences for MU metric accuracy remain unclear. Specifically, it is not fully understood how critical expert knowledge is, and what level of error occurs if signals remain manually unedited. In this study, we examined the impact of manual editing on the accuracy of the most commonly analysed MU metrics.

2 Methods

We generated synthetic hdEMG signals using 13×21 and 10×9 electrode arrays for the Soleus (SO) and Biceps Brachii (BB), respectively, with an interelectrode distance of 5 mm. We simulated 200 MUs for the SO using the simulator from [1], and 500 MUs for the BB using the simulator from [2]. Two separate muscle realizations were created for each muscle, with MU territories randomly distributed in each realization.

MU discharge patterns followed the model described in [3]. MU sizes and recruitment thresholds were sampled from exponential distributions, resulting in many small, low-threshold MUs and progressively fewer large, high-threshold MUs. Contraction durations were set to 20 seconds, with contraction intensities at 10%, 30%, 50%, and 70% of the maximum voluntary contraction (MVC). We added the colored noise (bandwidth 20-500 Hz) to the generated signals, creating signals with three different signal-to-noise ratios (SNRs): 15 dB, 20 dB, and infinite (no noise). In total, 48 synthetic signals were produced, with 24 corresponding to each muscle.

All signals were decomposed automatically using the Convolution Kernel Compensation (CKC) method [6], implemented in the DEMUSE tool (version 6.0, University of Maribor, Slovenia). The decomposition was performed with the following settings: 50 runs per decomposition, band-pass filtering of the hdEMG signals using a 4th-order Butterworth filter with a 20–500 Hz passband, no spatial filtering, and automatic selection of the top 95% highest-quality hdEMG channels. The CKC decomposition identified an average of 17.6 ± 3.5 MUs for the SO muscle and 8.6 ± 3.8 MUs for the BB muscle. The average initial PNR values were 40.1 ± 12 dB for SO and 33.4 ± 8.9 dB for BB.

Seven human operators (they are all listed as coauthors) with different levels of experience edited the signals. The level of experience was determined by the number of previously edited signals; therefore, we included two beginners, who had edited fewer than 50 signals, three semi-experienced operators, who had edited 51-1000 signals, and two experts, who had edited more than 1000 signals [5]. Before the editing procedure, human operators were given a written tutorial [4] to standardise the editing process.

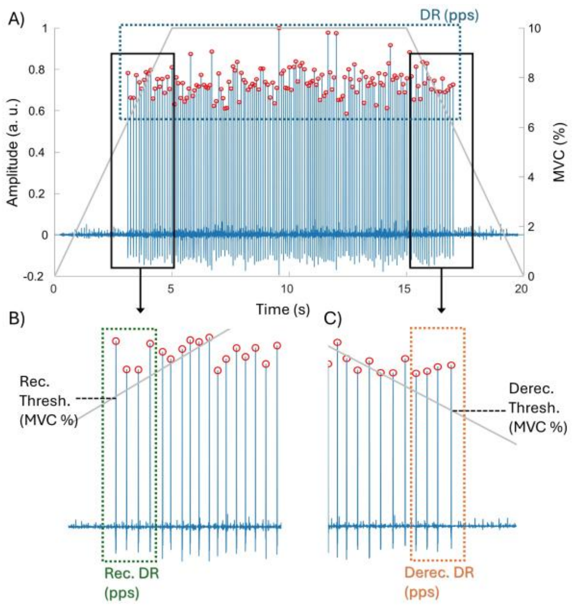

Figure 1. A) A visual representation of MU, its pulse train, discharges, and simulated excitation, along with marked discharges that were accounted for in the calculation of DR. B), C) A visualised representation of ReTh and DeTh with marked discharges that were accounted for in the calculation of reDR and deDR.

While human operators were editing the signals, log files were automatically created, saving all the used actions with a corresponding time stamp, allowing us to precisely calculate the editing time [5]. We inspected the ReTh, DeTh, discharge rate (DR), ReDR, and DeDR of signals before and after editing.

The DR was calculated by dividing the total number of discharges by the signal duration in seconds (Fig. 1A). The ReTh and DeTh were defined as the force level at the time of the first and the last discharge, respectively. The ReDR and DeDR were computed as the DR based on the first and last four MU discharges, respectively (Fig. 1B, Fig. 1C).

ReTh, DeTh, DR, ReDR, and DeDR were evaluated for all signals before and after manual editing and compared to the ground truth values obtained from the simulated reference signals. Errors were quantified as the absolute difference between measured and reference values.

Statistical analyses were conducted in R (version 4.2.2; R Foundation for Statistical Computing, Vienna, Austria) using linear mixed models (LMMs) implemented in the lme4 package [7]. The LMMs included the initial PNR value as a continuous fixed effect, muscle, operator's experience, SNR, and whether the signal was edited or not (editing status) as categorical fixed effects, and contraction level as an ordered fixed effect. Simulated subject ID and operators' ID were included as random intercepts to account for within-subject and within-operator variability. Separate models were fitted for each outcome variable: ReTh, DeTh, DR, ReDR, and DeDR.

Data distribution was evaluated using histograms, quantile-quantile plots of residuals, and plots of raw residuals against lagged residuals. Statistical significance was defined at a p-value threshold of 0.05 and assessed using the ANOVA function from the lmerTest package [8]. Pairwise post hoc comparisons of estimated marginal means were conducted using the Tukey adjustment for multiple testing, implemented via the emmeans package [9].

3 Results

Table 1. Mean ± standard deviation (AVG ± SD) of assessment errors for all outcome variables (ReTh, DeTh, DR, ReDR, and DeDR), excluding zero-value errors. Percentages of errors are shown in parentheses.

	Unedited (AVG ± SD)	**Edited (AVG ± SD)**
ReTh error (% MVC)	2.2 ± 3.1 (10%)	2.6 ± 3.8 (6%)
DeTh error (% MVC)	1.8 ± 2.5 (8%)	3 ± 3.7 (5%)
DR error (Hz)	0.8 ± 1.5 (70%)	0.3 ± 0.7 (17%)
ReDR error (Hz)	0.1 ± 0.4 (13%)	0.2 ± 0.5 (6%)
DeDR error (Hz)	0.1 ± 0.4 (13%)	0.2 ± 0.5 (6%)

On average, operators needed 39 ± 87 s for editing signals from BB, and 25 ± 65 s for signals from SO. The

ReTh error was influenced by the contraction level ($p = 0.003$, $F = 4.6$), with significantly higher error at 10% compared to 30% of MVC ($p = 0.0006$; Figure 2A). Muscle type also significantly ($p = 0.0002$, $F = 13.1$) affected the ReTh, with error being lower for the BB compared to the SO ($p = 0.0003$). Additionally, the initial PNR was negatively correlated with the ReTh error ($p < 0.0001$, $F = 151.2$). Editing status was significant ($p = 0.009$, $F = 6.8$), indicating that manual editing reduced errors ($p = 0.009$; Figure 3A; Table 1). In contrast, the operator's experience ($p = 0.7175$, $F = 0.5$) and SNR ($p = 0.2215$, $F = 1.5$) did not significantly affect the ReTh error.

The DeTh error was significantly affected by the contraction level ($p = 0.0354$, $F = 2.9$), with lower error observed at 10% compared to 50% of MVC ($p = 0.0011$; Figure 2B). The initial PNR was positively correlated with the DeTh error ($p < 0.0001$, $F = 52.8$), and the SNR also had a significant effect ($p < 0.0001$, $F = 10.1$), with lower errors at Inf dB compared to 15 dB ($p < 0.0001$) and 20 dB ($p = 0.0139$). In contrast, muscle type ($p = 0.72$, $F = 0.2$), operator's experience ($p = 0.5886$, $F = 0.6$), and editing status ($p = 0.9123$, $F = 0.1$; Figure 3B; Table 1) did not significantly influence the DeTh error.

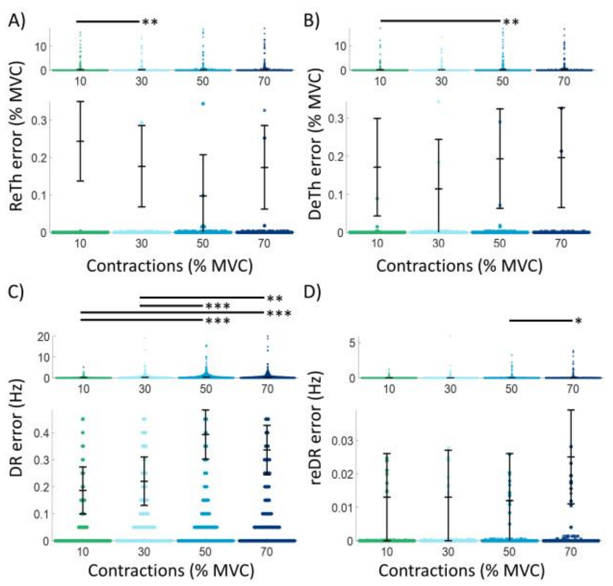

Figure 2. Comparison of A) Recruitment threshold (ReTh), B) Derecruitment Threshold (DeTh), C) Discharge rate (DR), and D) Recruitment discharge rate (ReDR) of different contraction levels as accumulated across the simulated muscle BB and SO. Black vertical whiskers denote estimated marginal means with 95% confidence intervals; Horizontal lines indicate significant differences: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

Contraction level significantly affected the DR error ($p < 0.0001$, $F = 18.6$), with lower errors observed at lower than higher contraction levels. Significant differences were found when comparing 10% to 50% ($p < 0.0001$) and 70% ($p < 0.0001$), as well as when comparing 30% to 50% ($p < 0.0001$) and 70% ($p = 0.001$) of MVC as visualized in Figure 2C. The initial PNR was negatively correlated with the DR error ($p < 0.0001$, $F = 331$). Editing status also had a significant effect ($p < 0.0001$, $F = 636.7$), with lower errors observed after manual editing compared to before ($p < 0.0001$; Figure 3C; Table 1). In contrast, muscle type ($p = 0.0556$, $F = 12.5$), operator's experience ($p = 0.8539$, $F = 0.2$), and SNR ($p = 0.2303$, $F = 1.5$) did not significantly influence the DR error.

The ReDR error was significantly influenced by the contraction level ($p = 0.0388$, $F = 2.8$), with higher errors observed at higher contraction levels. There was a significant difference between 70% and 50% ($p = 0.0487$; Figure 2D) of MVC. The initial PNR was negatively correlated with the ReDR error ($p < 0.0001$, $F = 39.4$). Muscle type ($p = 0.7915$, $F = 0.1$), operator's experience ($p = 0.6355$, $F = 0.5$), SNR ($p = 0.0598$, $F = 2.8$), and editing status ($p = 0.5197$, $F = 0.4$; Figure 3D; Table 1) did not significantly affect the ReDR error.

The DeDR error was positively correlated with the contraction level ($p < 0.0001$, $F = 16$) and negatively correlated with the initial PNR ($p < 0.0001$, $F = 25.9$). The error was highest at 70% compared to 10% ($p < 0.0001$), 30% ($p < 0.0001$), and 50% of MVC ($p < 0.0001$). No significant effects were found for muscle type ($p = 0.6595$, $F = 0.3$), operator's experience ($p = 0.589$, $F = 0.6$), SNR ($p = 0.1267$, $F = 2.1$), or editing status ($p = 0.3686$, $F = 0.8$; Table 1).

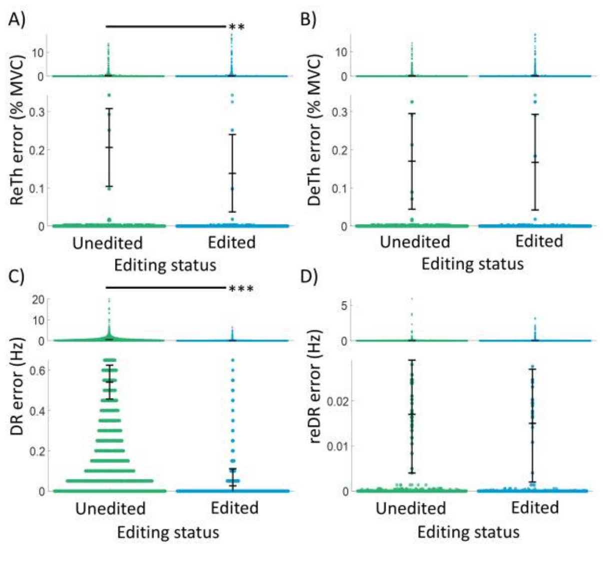

Figure 3. Comparison of A) Recruitment threshold (ReTh), B) Derecruitment Threshold (DeTh), C) Discharge rate (DR), and D) Recruitment discharge rate (ReDR) of edited and unedited signals. Black vertical whiskers denote estimated marginal means with 95% confidence intervals; Horizontal lines indicate significant differences: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

4 Discussion

In this study, we showed that manual editing of hdEMG decomposition results reduces errors in key MU metrics, particularly DR and ReTh. In contrast, the effects on DeTh, ReDR, and DeDR were less pronounced. Additionally, initial PNR consistently correlated with assessment accuracy across all evaluated parameters, which aligns with its established role as a metric of MU identification accuracy [14]. The contraction level also influenced all MU metrics, where higher errors were induced by higher contraction levels for ReTh and by lower contraction levels for DeTh, DR, ReDR, and DeDR. Higher contraction levels in most cases increase the DR, number of active MUs, and MU crosstalk, causing the manual editing to become more challenging and therefore more prone to errors for most of the MU metrics.

These findings align with previous reports highlighting the necessity of manual editing in decomposition workflows [5], [11], and extend current knowledge by quantifying the error reduction across a specific MU metric. While prior studies have noted variability in decomposition results between operators [5], our results show that, despite such variability, manual editing systematically improves the accuracy of MU analysis, particularly under conditions of lower PNR and higher contraction levels.

From a practical perspective, our findings suggest that manual editing remains a critical step in hdEMG decomposition, particularly when precise DR and ReTh assessments are required. However, in datasets with high initial PNR, the relative benefits of editing may be reduced, indicating a potential threshold above which manual editing may be deprioritized without substantially compromising accuracy.

This study has several limitations, including the analysis only having been performed on synthetic hdEMG signals (synthetic SO and BB muscles), where MU spike trains are usually more accurately identified by the automatic decomposition with higher initial PNR compared to the experimental signals.

In conclusion, manual editing substantially improves the accuracy of key MU parameters in hdEMG decomposition and remains an essential component of current analysis pipelines, particularly in lower-quality signals and when analyzing recruitment properties.

5 Acknowledgements

This research was funded by the European Union's Horizon Europe Research and Innovation Program [HybridNeuro project, GA No. 101079392]. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Bibliography

[1] Botter, A., & Vieira, T. M. (2017). Optimization of surface electrodes location for H-reflex recordings in soleus muscle. Journal of Electromyography and Kinesiology, 34, 14-23.

[2] Farina, D., Mesin, L., Martina, S., & Merletti, R. (2004). A surface EMG generation model with multilayer cylindrical description of the volume conductor. IEEE Transactions on Biomedical Engineering, 51(3), 415-426.

[3] Fuglevand, A. J., Winter, D. A., & Patla, A. E. (1993). Models of recruitment and rate coding organization in motor-unit pools. Journal of neurophysiology, 70(6), 2470-2488.

[4] Murks, N., Škarabot, J., & Holobar, A. (2025). Tutorial on editing the decomposition results from isometric contractions.

[5] Murks, N., Škarabot, J., Kramberger, M., Sedej, G., Valenčič, T., Connely, C. D., ... & Holobar, A. Motor unit editing time in high-density EMG decomposition decreases with the operator's experience and pulse-to-noise ratio but does not depend on simulated muscle.

[6] Holobar, A., & Zazula, D. (2007). Multichannel blind source separation using convolution kernel compensation. IEEE Transactions on Signal Processing, 55(9), 4487-4496.

[7] D. Bates, M. Mächler, B. Bolker and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," Journal of Statistical Software, 2015.

[8] A. Kuznetsova, P. B. Brockhoff and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," Journal of Statistical Software, vol. 82, pp. 1-26, 2017.

[9] R. Lenth and M. R. Lenth, "Package 'lsmeans'," *The American Statistician,* vol. 34, pp. 216-221, 2018.

[10] Del Vecchio, A., Holobar, A., Falla, D., Felici, F., Enoka, R. M., & Farina, D. (2020). Tutorial: Analysis of motor unit discharge characteristics from high-density surface EMG signals. *Journal of Electromyography and Kinesiology*, *53*, 102426.

[11] Hug, F., Avrillon, S., Del Vecchio, A., Casolo, A., Ibanez, J., Nuccio, S., ... & Farina, D. (2021). Analysis of motor unit spike trains estimated from high-density surface electromyography is highly reliable across operators. *Journal of Electromyography and Kinesiology*, *58*, 102548.

[12] Del Vecchio, A., Casolo, A., Negro, F., Scorcelletti, M., Bazzucchi, I., Enoka, R., ... & Farina, D. (2019). The increase in muscle force after 4 weeks of strength training is mediated by adaptations in motor unit recruitment and rate coding. *The Journal of physiology*, *597*(7), 1873-1887.

[13] Múgica-Esteve, M., Montero-Pardo, C., Larraga-García, B., Gutiérrez, Á., & Oliveira Barroso, F. (2024, November). Assessment of Motor Units' Recruitment/Derecruitment Discharge Rate after Phase-Dependent Peripheral Electrical Stimulation. In International Conference on NeuroRehabilitation (pp. 756-760). Cham: Springer Nature Switzerland.

[14] Holobar, A., Minetto, M. A., & Farina, D. (2014). Accurate identification of motor unit discharge patterns from high-density surface EMG and validation with a novel signal-based performance metric. *Journal of neural engineering*, *11*(1), 016008.