Towards hybrid modelling of aluminium extrusion mechanical properties – a univariate representation of artificial aging

Christian Dalheim Øien[1,a *], Ole Runar Myhr[2,b], and Geir Ringen[1,c]

[1]NTNU, Department of Mechanical and Industrial Engineering, 7491 Trondheim, Norway

[2]Hydro Aluminium, Research and Technology Development, 6601 Sunndalsøra, Norway

[a]christian.d.oien@ntnu.no, [b]ole.runar.myhr@hydro.com, [c]geir.ringen@ntnu.no (*corresponding author)

Keywords: aluminium extrusion, hybrid modelling, precipitation kinetics

Abstract. In ongoing research towards hybrid modelling of the mechanical properties of age hardened aluminium extrusions one of the challenges is the numerical representation of the process of artificial aging. A theoretically formulated compound variable, or synthetic feature, based on diffusion is therefore investigated. It is further assessed to which extent this simplification can be used to univariately represent the complete temperature history in model training and inference. A comprehensive dataset of mechanical properties of Al-Mg-Si (6xxx) extrusions has been studied and machine learning models have been trained and evaluated with and without this simplification. Results indicate that this type of aging representation and dimensionality reduction is fairly accurate and may be useful for model training and process similarity assessments. In the studied dataset it led to a reduction of the number of aging variables from 9 to 1.

Introduction

As the ratio of post-consumer material (PCM) is expected to increase in aluminium extrusion, even in demanding market segments like automotive structural components, it can also be expected that the industry will benefit from enhanced modelling of the resulting material and process variations [1]. A promising strategy for achieving both precise and versatile models is hybrid modelling, where physics-based models (PBM) are combined with data-driven models (DDM). A self-inflicting challenge with utilizing PCM in industrial aluminium processing is that increasing process or material variation is incompatible with ensuring profitable production. This is greatly due to limited knowledge and accurate modelling capabilities, which again is a result of lacking experience and variation in production processes and materials. In other words, by aiding PBMs with observations we aim to increase the accuracy and extrapolation power of state-of-the-art models and combat this variation modelling paradox.

A first-off hybrid modelling implementation is in-progress work, which is based on a simple weighted average architecture. The applied PBM will be based on the work of Myhr et. al. [2, 3, 4], namely the NaMo model. The intention is that hybrid model inference will be based on an observation density assessment, from which the DDM vs. PBM weighting can be calculated. In order to ensure a robust and verifiable method for this, the input variable dimensionality should be kept at a minimum. This is the main motivation for representing the aging process univariately, whereas normally the cooling, storage and age hardening temperature history will consist of several segments of temperatures, temperature rates and durations.

In summary, the governing research question for the present research is the following: *To which extent can the Scheil integral represent artificial aging (AA) cycles in DDM of Al-Mg-Si mechanical properties?*

Theory

Through the processes extrusion, cooling, forming, and AA of Al-Mg-Si (6xxx) alloys the strength is mainly depending on the concentration of available Mg and Si in solid solution through the various stages, and commonly idealized as the precipitation sequence shown below [5, 6].:

SSSS → atomic clusters → GP-zones → β” → β’, U1, U2, B’ → β(Mg_2Si)

Here, SSSS denotes super-saturated solid solution, β”, β’, U1, U2 and B’ are different metastable phases, and β(Mg_2Si) is the equilibrium phase. The maximum strength is usually obtained for an optimum combination of particle number density and mean particle size of the hardening β” phase [5]. The maximum material strength is obtained through a T6 artificial ageing heat treatment. For prolonged ageing beyond the time corresponding to the T6 temper condition, the particles will grow, the number density will decrease and the metastable particles will transform further leading to over-aging and reduced strength [7, 8].

The basis for this study is treating AA as an isokinetic diffusion-controlled process. Diffusion is the rate controlling reaction for nucleation at lower temperatures where the nucleation barrier is negligible and also for growth and dissolution of particles. Generally, on the other hand, it must be clearly stated that the precipitate evolution in Al-Mg-Si alloys is complex and not isokinetic, as the interaction between nucleation and growth/dissolution of particles is strongly dependent on the specific temperature-time cycle.

As a crude approximation, let us first assume isokinetic conditions by setting a constant diffusion length

$$x = \sqrt{D(T)t} = \sqrt{D(T_r)t_r} \tag{1}$$

where a reference temperature T_r and reference time t_r has been chosen and the material’s diffusion coefficient $D(T)$ is a function of temperature. We now make use of the Arrhenius relation

$$D = D_0 \exp\left(-\frac{Q_d}{RT}\right) \tag{2}$$

where Q_d is the molar activation energy for diffusion and R is the universal gas constant, and insert this into Eq. 1. We can then express the equivalent AA time as

$$t_{AA}^* = t_r \exp\left(\frac{Q_d}{R}\left(\frac{1}{T_{AA}} - \frac{1}{T_r}\right)\right) \tag{3}$$

given an isothermal AA at temperature T_{AA}. In other words, Eq. 3 provides the equivalent diffusion time according to the reference time and temperature. For aluminium alloys it is common to set $Q_d = 130$ kJ/mol for the diffusion of Mg in Al [9]. Providing an example for intuition, if a T6 temper can be achieved by 185°C isothermally at 5 hours, then a diffusion-based approximation of corresponding aging times would be 51 minutes at 210 °C or 36 hours at 160 °C. A visualization of this is given in Fig. 1.

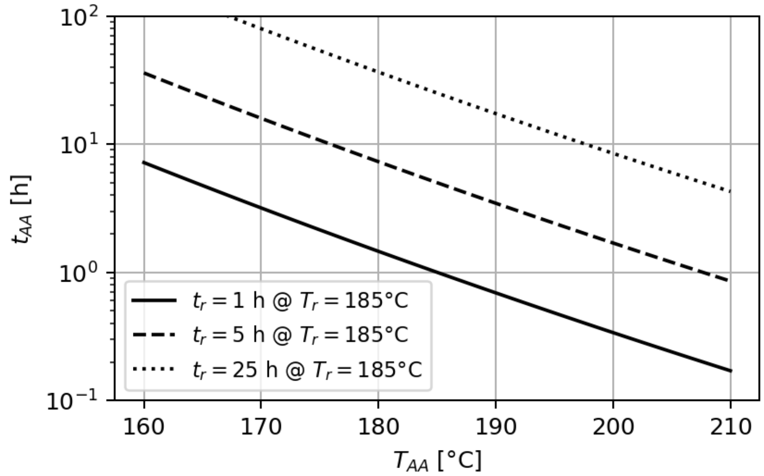

Figure 1 – Estimated equivalent aging durations t_{AA} as a function of aging temperature T_{AA} based on three different references, calculated using Eq. 3.

Moving briefly over to general transformation theory, Christian [10] showed that a reaction is additive if the rate of change of the volume fraction transformed ζ can be written as

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{h(T)}{g(\zeta)} \tag{4}$$

i.e. depending on decoupled functions of the temperature only and of the fraction transformed only. Let us now introduce the Scheil integral [11], which can be traced back to contributions to understanding cooling processes in steels. It has also been applied to Al-Mg-Si transformation kinetics specifically [12]. Given Eq. 4 and making use of the additivity property over infinitesimal time steps, it can then be showed [10] that the integral

$$I = \int_0^{t_{AA}} \frac{\mathrm{d}t}{t^*_{AA}(T)} \tag{5}$$

gives the relation between the total non-isothermal reaction time t_{AA} and the achieved fraction transformed ζ for a non-isothermal reaction, where $t^*_{AA}(T)$ is the (hypothetical) time to reach a certain reference stage ζ^* isothermally at temperature T.

In industrial AA heat treatment, a controlled continuous temperature curve is applied for practicality. Expressing such a temperature curve as a function $T(t)$ and assuming a diffusion-controlled transformation, we can then insert Eq. 3 into Eq. 5 and evaluate I numerically. This provides a dimensionless quantification of non-isothermal AA referenced as the Scheil integral. We see that $I(t_{AA} = t_r\,, T = T_r) = 1$, i.e. that it is reference normalized, and that an infinite number of temperature paths can yield $I = 1$ at any desired time t_{AA}, which would then be expected to be equivalent of a holding temperature T_r over a duration t_r isothermally. The usage of I as a synthetic data feature is explained in the following section.

Methods

The basis of the present research is a database of 2 895 extrusion experiments shared by the company Hydro Aluminium, where 5” extrusion billets have been cast with various selected Al-Mg-Si alloy compositions and extruded using a vertical press at the Department of Mechanical and Industrial Engineering at NTNU, owned and operated by SINTEF. The compositions were measured using x-ray fluorescence spectrometry. A strip profile geometry of 1.9 x 25 mm was extruded, and water quenching was applied. AA was then carried out with a recipe chosen from an array of various temperature curves, each consisting of linear segments of either constant temperature or a constant heating or cooling rate. This makes the dataset a unique combination of Al-Mg-Si alloy variations and AA heat treatments, ranging from under-aged to over-aged. Dog bone tensile test specimens were then machined from the middle of each press and tested quasi-statically at room temperature.

The resulting numerical data was compiled into a table with columns Mg, Si, Fe, Mn, and Cu content in wt. %, the AA curve in the form of 5 temperatures and 4 durations, and the measured yield and ultimate tensile stress in MPa. This gives a table of 16 columns and 2 895 rows. Additionally, as a synthetic feature the Scheil integral was calculated numerically for each row using Eq. 3 and Eq. 5, where references $t_r = 5 \cdot 3600\text{s}$, $T_r = 185 + 273\text{K}$ were chosen to represent a T6-level aging. In this sense $I = 1$ would correspond to T6, $I < 1$ to under-aging, and $I > 1$ to over-aging, respectively.

In order to assess the accuracy of representing AA univariately using the Scheil integral, machine learning was applied to train and test regression models. Models were trained and tested using Python 3.12 on a conventional laptop computer. The algorithm XGBoost [13] was chosen as it has inherent multi-output capability and fewer hyper-parameters to be tuned than typical neural network regression algorithms. It is also known to perform well in similar regression tasks [14]. The metrics

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} \tag{6}$$

and

$$\text{RMSE} = \frac{1}{n}\sqrt{\sum_i (y_i - \hat{y}_i)^2} \tag{7}$$

were chosen to quantify model performance. Both sum over the n samples of the test set where y_i and $\hat{y}_i$ are the true and predicted output of sample i and $\bar{y}_i$ is the mean of the true outputs. In essence, R^2 compares the variance of the prediction errors with the variance of the true values, ranging from negative infinity to one. RMSE, on the other hand, provides an intuitive measure of error amplitude with the same unit as the estimated variable. In order to limit implicit representation of training data points in test data a sequential data split was applied, as opposed to the more commonly applied random data split which would yield misleading results. Furthermore, in order to quantify random and data centric variation in model accuracy 5-fold cross validation was applied. Based on corresponding sets of 5-fold results 95% confidence intervals were calculated as $\text{CI} = \bar{e} \pm 2S_e$ where $\bar{e}$ is the mean and S_e is the sample standard deviation of that metric given by

$$\bar{e} = \sum_j e_j / n \tag{8}$$

and

$$S_e = \sqrt{\sum_j \left(e_j - \bar{e}_i\right)^2 / (n-1)} \tag{9}$$

Model hyper-parameters were tuned using on an exhaustive grid search which led to the values summarized in Table 1.

Table 1 – Hyper-parameter values used in XGBoost model training

Hyper-parameter	chosen value
objective	"reg:squarederror"
subsample	0.8
colsample_bytree	0.5
alpha	0.1
lambda	0.9
gamma	0.01
learning_rate	0.08
max_depth	4
n_estimators	500

To discuss one of the possible uses of I in model evaluation, the NaMo model [15] was inferred on the dataset. NaMo is a proprietary model predicting stress-strain behaviour for a broad array of 6xxx aluminium alloys. The model first predicts Particle Size Distributions for clusters and metastable particles by a precipitation model, and then linearly generates the yield strength from the contributions; 1) yield strength of pure aluminium, 2) strength contributions from elements in solid solution, 3) precipitation hardening from clusters and metastable particles, and 4) work hardening by contribution from dislocations to the yield stress. Furthermore, temperature and strain rate are integrated into the model, based on the principles of obstacle limited dislocation glide [15]. Finally, the model has been calibrated by numerous experiments to serve as a complete thermomechanical simulation model for Al-Mg-Si alloys. A Scheil integral-based assessment of the NaMo predictions with respect to under-aging and over-aging is given in the discussion section.

Results

In order to answer the previously highlighted research question, two analyses were conducted. Firstly, investigations on the validity of I and the observed relationship with peak hardness, under-aging and over-aging, and secondly; machine learning regression analyses to compare a) representing AA with all 9 available temperature and duration variables as input with b) representing AA with I as a single feature.

Firstly, the observed dependency between I and achieved yield strength was studied by calculating a second synthetic feature on the dataset namely relative yield strength. To do this the dataset was split into series of experiments each consisting of one single alloying composition but with varying AA heat treatments. Within each such series the maximum achieved yield stress was determined, and then the relative yield stress was calculated as the ratio to that maximum for each experiment. In Fig. 2 all observed relative yield stresses from all studied alloy compositions have been plotted in a box plot with respect to $\log I$.

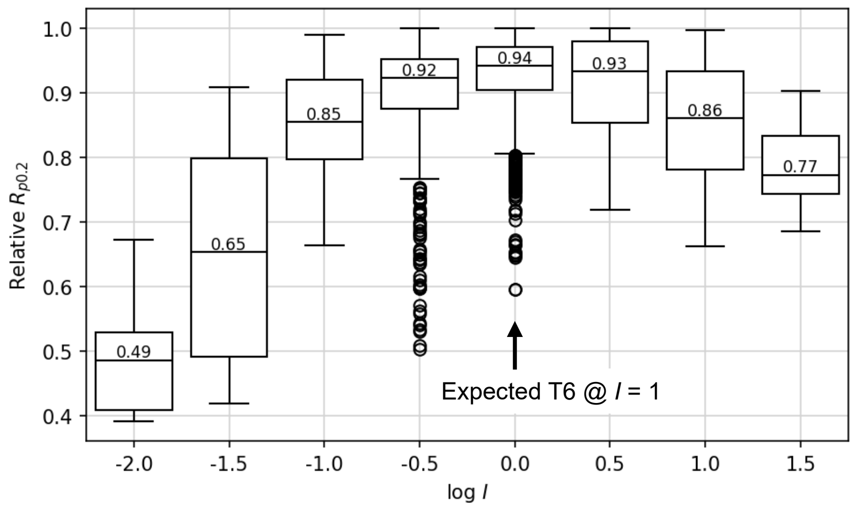

Figure 2 – Observed relative $R_{\mathrm{p0.2}}$ vs. $\log I$ across the entire dataset. This can be seen of as an empirical aging curve for the Al-Mg-Si system.

This shows that, with some consistency, AA cycles corresponding to $I \approx 1$ tend to lead to the highest yield stress despite variations in alloy composition, aging temperature and cycle type. Similarly, it can be seen that $I \leq 0.1$ corresponds to significant under-aging, and that $I \geq 10$ corresponds to significant over-aging, respectively.

Furthermore, the value of $\log I$ pertaining to each of the achieved maximum yield stresses were collected and plotted with respect to applied aging temperatures[1]. Such a box plot is shown in Fig. 3, which validates that $I \approx 1$ corresponds to T6 conditions across a wide range of AA holding temperatures.

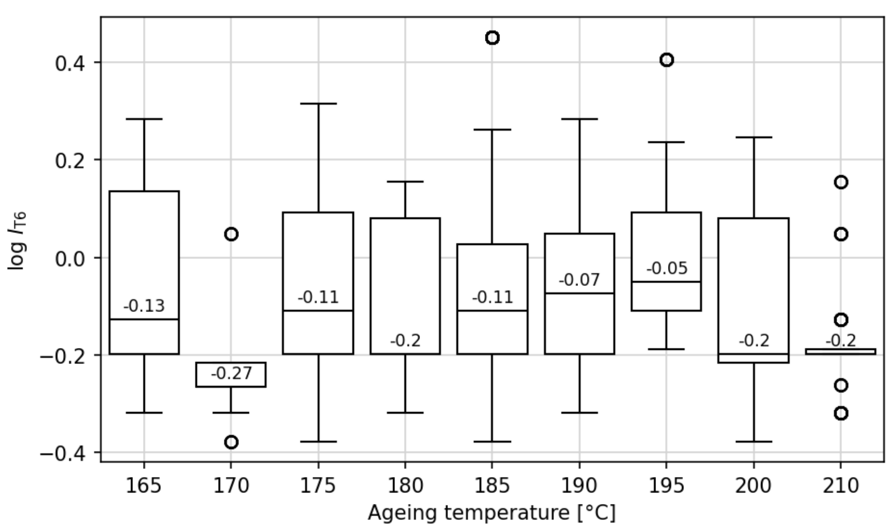

Figure 3 – Observed $\log I$ values of T6-conditions vs. AA holding temperature

Finally, the trained machine learning models according to approaches a) and b) were evaluated using the previously described metrics, and confidence intervals of each metric were calculated

[1] In fact, the maximum AA temperature in the case of a more complex heat treatment cycle.

using Eq. 8 and 9. These results are summarized in Table 2, which shows very similar performance between the two approaches. An example test set inference is shown in Fig. 4.

Table 2 – Cross validation results of ML model trained with a) full aging representation with 9 variables and b) *I* as only aging variable

	$R^2(R_{p0.2})$	$R^2(R_m)$	$\mathrm{RMSE}(R_{p0.2})$ [MPa]	$\mathrm{RMSE}(R_m)$ [MPa]
a)	0.80 ± 0.11	0.89 ± 0.06	17.9 ± 18.2	11.6 ± 11.4
b)	0.77 ± 0.13	0.88 ± 0.11	18.2 ± 10.8	12.1 ± 10.2

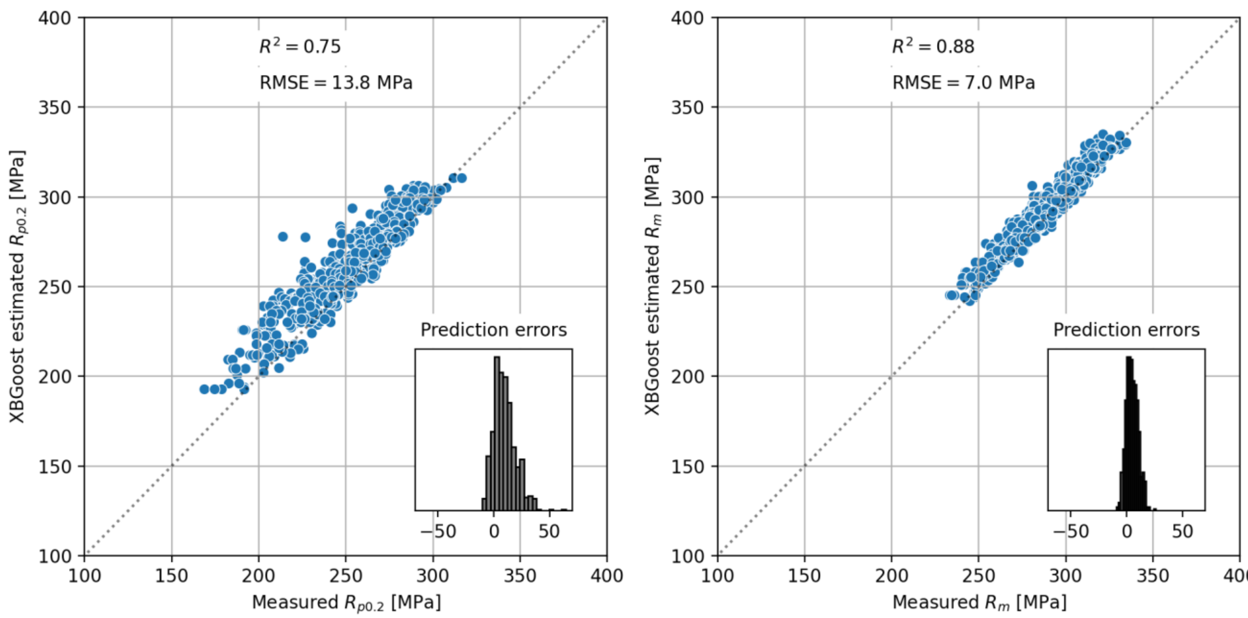

Figure 4 – Test results of ML model trained with *I* as only aging variable (from a randomly chosen cross validation fold). This scatter plot compares estimated values to measured values.

Discussion

The results show that the assumption of diffusion-controlled reaction kinetics during AA has provided a useful synthetic feature for DDM.

The presented results can be argued to have at least three implications. Firstly, the results indicate that the Scheil integral is a useful synthetic feature in numerical studies of strength properties of artificial aged Al-Mg-Si alloys to represent the aging sequence in an experiment, e.g. in the form of dimensionality reduction or data incompatibility, with a relatively low loss of information. Secondly, it can be used to categorize or compare different types of AA heat treatments, or data sets where AA is represented differently. This corresponds with the outlined motivation of the present study, were there is a need of a low-dimensional or univariate representation of aging that can be used to assess observation density in a dataset for the purpose of research on, and application of, hybrid modelling of Al-Mg-Si mechanical properties. Specifically, the intention is to create a robust method for calculating the weighting coefficient between PBM and DDM in a simple parallel hybrid model architecture. Thirdly, the Scheil integral enables an indicative assessment of PBM model accuracy with respect to under-aging/over-aging across variation in AA cycle types and temperatures. This could be a Kampmann-Wagner model or any other type of PBM.

In the example below, NaMo [15] has been inferred on each row of the dataset and the error of the estimates has been plotted with respect to *I*, viz. Fig. 5. This simple analysis indicates that the model tends to underestimate the yield strength of some of the under-aged samples and exemplifies the application of *I* in model assessment. Such as result seems to be valuable in the work with further model improvements. Without this synthetic feature, several analyses would be required since the dataset consists of multiple types of AA cycles, temperatures and durations.

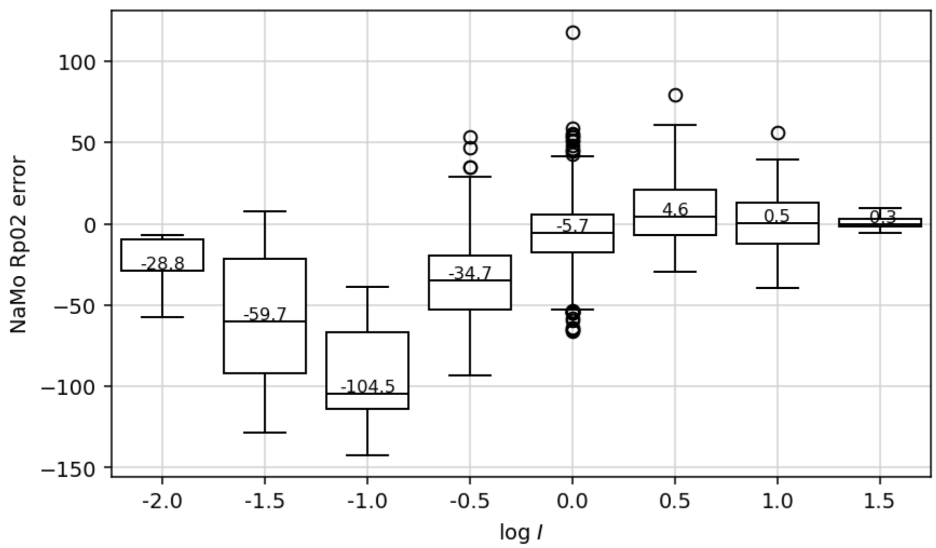

Figure 5 – Observed average $R_{p0.2}$ estimation errors vs. log *I* for the NaMo model.

It should be noted that there are several sources of variation and random error involved in the experiments that have formed the applied dataset, such as casting, homogenization and the preparation of the extrusion billets, the production conditions for extrusion such as billet heating, pressing and quenching, as well as the procedure with machining of samples and tensile testing.

Summary

The motivation for the present study has been to assess the possibility and accuracy of univariate representation of an AA heat treatment cycle of Al-Mg-Si extrusions, through approximating the precipitation kinetics as diffusion-controlled and applying the Scheil integral based on transformation additivity. This simplification has shown to lead to little information loss in the case of replacing several heat treatment variables with this one synthetic feature or variable. The results are indicated to have a three-fold impact – data characterization, DDM training, and PBM assessment. Planned further work will incorporate the studied synthetic feature in development of a weighting coefficient function for an Al-Mg-Si hybrid model based on a weighted average architecture. More specifically, it will be useful in defining a method for calculating average neighbour distances in the dataset.

Acknowledgements

This study is conducted in relation to the research project AluGreen, KSP Aluminium Green Platform, project number 328843, funded by the Norwegian Research Council.

References

[1] D. Raabe, D. Ponge, P. J. Uggowitzer, M. Roscher, M. Paolantonio, C. Liu, H. Antrekowitsch, E. Kozeschnik, D. Seidmann, B. Gault, F. De Geuser, A. Deschamps, C. Hutchinson, C. Liu, Z. Li, P. Prangnell, J. Robson, P. Shanthraj, S. Vakili, C. Sinclair, L. Bourgeois, S. Pogatscher, *Making sustainable aluminum by recycling scrap: The science of "dirty" alloys*, Progress in Materials Science, vol. 128, 2022, https://doi.org/10.1016/j.pmatsci.2022.100947

[2] O. R. Myhr, Ø. Grong and S. J. Andersen: *Modelling of the age hardening behaviour of Al-Mg-Si alloys*. Acta Mater. Vol. 49, Jan. 2001, pp. 65-75. https://doi.org/10.1016/S1359-6454(00)00301-3

[3] O. R. Myhr, Ø. Grong, and K. O. Pedersen: *A Combined Precipitation, Yield Strength and Work Hardening Model for Al-Mg-Si Alloys*. Metall. Mater. Trans A, Vol. 41A, No. 9, 2010, pp. 2276-2289. https://doi.org/10.1007/s11661-010-0258-7

[4] O. R. Myhr, Ø. Grong and C. Schäfer: *An Extended Age-Hardening Model for Al-Mg-Si Alloys Incorporating the Room-Temperature Storage and Cold Deformation Process Stages*. Metall. Mater. Trans A, Vol. 46A, 2015, pp. 6018-6039. https://doi.org/10.1007/s11661-015-3175-y

[5] S. J. Andersen, H. W. Zandbergen, J. Jansen, C. Træholt, U. Tundal, O. Reiso,

The crystal structure of the β″ phase in Al–Mg–Si alloys, Acta Materialia, Volume 46, Issue 9, 1998, pp. 3283-3298, ISSN 1359-6454, https://doi.org/10.1016/S1359-6454(97)00493-X.

[6] A.K. Gupta, D.J. Lloyd, S.A. Court, Precipitation hardening in Al–Mg–Si alloys with and without excess Si, Materials Science and Engineering: A, Volume 316, Issues 1–2,

2001, Pages 11-17, ISSN 0921-5093, https://doi.org/10.1016/S0921-5093(01)01247-3.

[7] S. J. Andersen, C. D. Marioara, R. Vissers, A. Frøseth, H. W. Zandbergen, The structural relation between precipitates in Al–Mg–Si alloys, the Al-matrix and diamond silicon, with emphasis on the trigonal phase U1-MgAl2Si2, Materials Science and Engineering: A, Volume 444, Issues 1–2, 2007, Pages 157-169, ISSN 0921-5093.

https://doi.org/10.1016/j.msea.2006.08.084.

[8] R. Holmestad, C. D. Marioara, F. J. H. Ehlers, M. Torsæter, R. Bjørge, J. Røyset, S. J. Andersen, *Precipitation in 6xxx Aluminum Alloys*, Proceedings of the 12th International Conference on Aluminium Alloys, 2010, pp. 30-39. ISBN: 978-4-905829-11-9

[9] S. Spigarelli and C. Paoletti (2017). A Unified Physical Model for Creep and Hot Working of Al-Mg Solid Solution Alloys. Metals. 8. 9. 10.3390/met8010009. http://dx.doi.org/10.3390/met8010009

[10] J.W. Christian, The Theory of Transformations in Metals and Alloys, Ch. 12 Formal Theory of Transformation Kinetics, Pergamon, 2002, ISBN 9780080440194.

https://doi.org/10.1016/B978-008044019-4/50005-2

[11] E. Scheil, Anlaufzeit der Austenitumwandlung, Archiv für das Eisenhüttenwesen, vol. 8, issue 12, 1935, pp. 565-567, https://doi.org/10.1002/srin.193500186

[12] O. R. Myhr and Ø. Grong, Process modelling applied to 6082-T6 aluminium weldments—I. Reaction kinetics, In: Acta Metallurgica et Materialia, Volume 39, Issue 11, 1991, pp. 2693-2702.https://doi.org/10.1016/0956-7151(91)90086-G.

[13] T. Chen and C. Guestrin. *XGBoost: A Scalable Tree Boosting System*. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), 2016. https://doi.org/10.1145/2939672.2939785

[14] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz. *A Comparative Analysis of XGBoost*. In: Artificial Intelligence Review, vol. 54, 2019. http://dx.doi.org/10.1007/s10462-020-09896-5

[15] O. R. Myhr, O. S. Hopperstad, and T. Børvik, *A Combined Precipitation, Yield Stress, and Work Hardening Model for Al-Mg-Si Alloys Incorporating the Effects of Strain Rate and Temperature*, Metallurgical and Materials Transactions, 2018, Volume 49, pages 3592–3609. http://dx.doi.org/10.1007/s11661-018-4675-3