Data-driven model to derive air-cooled heat sinks mass for power electronics

Thibault Abry, Jean-Christophe Crebier

Data-driven model to derive air-cooled heat sinks mass for power electronics

Thibault Abry, Research Engineer. Jean-Christophe Crebier, Senior Scientist.
Univ. Grenoble Alpes, CNRS, Grenoble INP, G2ELAB Grenoble, France
21 Avenue des Martyrs, 38000 Grenoble, thibault.abry@g2elab.grenoble-inp.fr, +33669742596

Abstract

This paper reports a new data-driven approach to model air-cooled heat sinks for active devices such as power modules or discrete package. Analysis of hundreds of references from several manufacturers allowed the establishment of new empirical relations between mass and thermal resistance of heatsinks. This work differs from previous contributions in heat sink modelling by the use of "real-world" (product driven) data to directly find relations, without the use of simulation software or analytic models. These relations can be used in electronics, and particularly in power electronics to derive the mass of a heat sink from its thermal properties. It is useful to the designer who wants to design and optimize the power densities of a power converter, or to determine the bill of material and/or its associated environmental impacts.

1 Introduction

Heat dissipation in power electronics (PE) is considered to be the most important factor influencing reliability, and it is gaining attention with emerging high-power density devices [1]. Heat sinks (HS) are commonly used to keep the device under a critical temperature to avoid premature failure. Minimizing the volume and mass of HS becomes critical to not counterpart the efforts put in maximizing power density on electric components. To choose an appropriate HS, the PE designer defines the cooling needs with the thermal resistance (R_{th}), a design parameter which dictates its capacity to transfer heat from the device to the air. It must be sufficient to keep the device below the required maximum temperature T_{max}. R_{th} depends thus on the size, the shape and the environment of the HS (ambient temperature T_{amb}, T_{max}, forced or natural convection, etc.).

Scientific literature focuses almost exclusively on estimating the R_{th} of a HS design from its parameters, often by using analytic models and simulations [2]. On one hand, analytical models based on correlation of fluid-dynamic numbers (*e.g.* Nusselt, Reynolds) are used to estimate the HS thermal resistance, from which mass and volume can be derived [3]. Implemented to simplify the effort, they can lead to over or under estimation of heat transfer [4]. On the other hand, while being more accurate, numerical simulations are time consuming and resource intensive, and can hardly be used to estimate mass from R_{th} [5].

A solution to these issues is to use machine learning, which lies in the middle. This field can bypass underlying mathematical concepts and precise simulations by creating data-based models. A successful attempt was conducted by researchers using an artificial neural network [6]. The model they created precisely predicts R_{th} from dimensions of fin-plate HS. While predictions are not resource intensive, the neural network still needs to be integrated, which isn't as straightforward as mathematical formulas. Also, their model relies on multiple dimension parameters of the HS, which might be unknown at early stage of product development.

Scientific literature focuses exclusively on estimating R_{th} from other parameters, but what could prove useful to the PE designer is the other way around: estimate physical dimensions, mass and volume of a HS given its R_{th}. As the thermal resistance needed to cool down a PE device might be estimated early in design, the size of the HS could be inferred early as well.

This paper proposes a new approach to this problem by using linear regression, which is a simple subset of machine learning, to model HS mass from R_{th}. The approach uses data extracted directly from manufacturers datasheets and catalogues, which hasn't been seen in literature. The relations found can be used to support power density estimates in converter optimization. Multiple models with different technical parameters are created, so the most appropriate one can be chosen based on the data available.

2 Model definition

2.1 General definition

Air cooling heat sinks differ in shape, fabrication process and material based on the application, which greatly affect their R_{th}, volume, and mass. The use of natural convection (NC) or forced convection (FC) with a fan also changes dramatically the thermal performance. Moreover, data availability changes from one manufacturer to another. To address this diversity and fit the need of the PE designer, multiple models are being created and presented in this paper. These models act as technology "filters" from which the most appropriate to the application can be chosen. The most common HS shape is the fin-plate, which consists in multiple fins perpendicularly "attached" to a base plate used to spread the heat among the fins. A summary of the main base parameters used to describe the fin-plate HS is presented in **Figure 1**, with dimensional parameters in black and "derived parameters" in color.

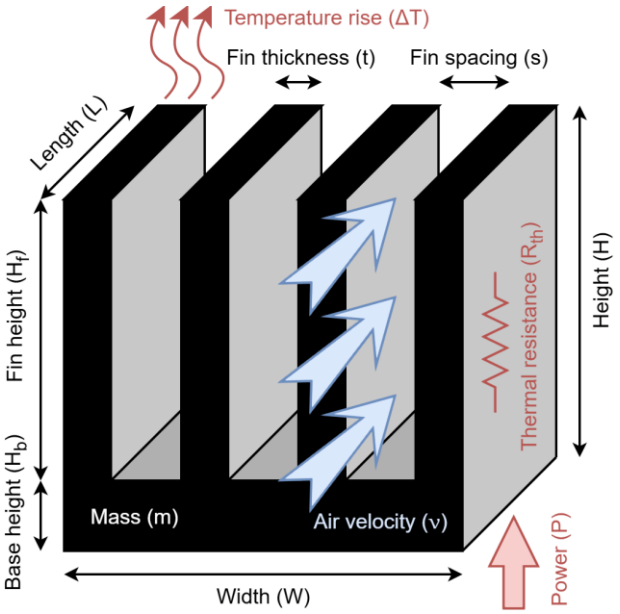

Figure 1: Fin-plate model with parameters.

2.2 Technology filters

To keep the study simple, base dimensional parameters will not directly be used in the models, meaning that only relations between the derived parameters are presented.
The parameters interesting to consider in predicting are the mass m, the total thermal exchange surface area S, and the heat sink outer volume (box volume) V_b which is calculated using **Equation 1**:

$$V_b = L * W * H \quad (1)$$

R_{th} represents the thermal resistance of the HS alone, which is calculated using **Equation 2**:

$$R_{th} = \frac{\Delta T}{P} - R_{thjc} - R_{thti}, \Delta T = T_{base} - T_{amb} \quad (2)$$

With ΔT being the difference between the temperature at the base of the heatsink T_{base} and the ambient temperature T_{amb}, P the power to dissipate, R_{thjc} the junction to case thermal resistance of the semiconductor device to cool down, and R_{thti} the thermal resistance of the thermal interface material between the device and the HS. R_{th} is influenced by other factors such as surface finish (radiation), temperature spread across the base and air velocity [7].
As there is no international standard for measuring R_{th}, manufacturers usually do not give R_{thjc}, R_{thti}, and even T_{amb} of their experiment. We thus going to make some assumptions presented in 3.2.

The technology filters include the shape of the heatsink (fin-plate heatsinks and other shapes), NC and FC, but only extruded aluminium is treated in this paper, due to the lack of data for other technologies. Anodization has a non-negligible effect on R_{th} in NC, but also depends heavily on the environment. Due to the lack of reliable data, it is not considered in this paper. Finally, in FC, ducting the heatsink to limit air leaks affects R_{th}, but due to the lack of data it is not considered as well at the moment.

It is worth noting that water cooling is not treated in this paper as it relies on dramatically different HS shapes from air cooling. This will need to be addressed in a specific work.

3 Data collection

As there is no international standard for measuring and presenting R_{th} for HS, available parameters and their interpretation differ from one manufacturer to another. Most of the time, the mass is not made available by the manufacturers which complicates data collection. To solve this issue, lineic mass has been precisely calculated from extrusion profiles, when available and applicable. When available, graphs have been extracted and digitized, such as R_{th} vs extrusion length or R_{th} vs air velocity in FC. Heat sinks with a given R_{th} can then be sampled from this data and their mass calculated. Three datasets have been created from datasheets/catalogues of HS manufacturing leaders and are summarized in **Table 1**.

3.1 Data collection tools

To collect and prepare data, four main actions were done, with full automation or semi-automation:

1. Web scrapping.
2. Tabular data extraction from pdf.
3. Curve digitization from images.
4. Extrusion profile surface calculation from *.dxf*.

Action 1 is very common and was performed via a simple custom python script that is thus not shared. Actions 2, 3 and 4 were mainly performed automatically using a custom python package that is shared to the scientific community in open access [8]. This package offers the possibility to manage entries in a database and perform actions on files: digitize curves, calculate surface area, etc.

Data collected from action 1 and 2 is kept in its original condition, meaning that the action does not introduce additional uncertainty. Action 3 and 4 on the other hand do introduce additional uncertainty because of sampling.

3.1.1 Curve digitization

This action was performed using the python package *plotdigitizer* [9]. The package essentially samples points at the mean location of pixels in the graph. Visual inspection of the digitized curves shows no particular drift from the original image. The absolute uncertainty is thus calculated for each image using **Equation 3**:

$$\Delta x = \frac{W}{W_p} t_p, \ \Delta y = \frac{H}{H_p} t_p \quad (3)$$

With Δx and Δy being respectively the absolute uncertainty of the x and y component of a measurement on the graph, W and H the width and height in the respective unit on the graph, W_p and H_p the width and height in pixels and t_p the line thickness in pixels. Usually t_p scales with W_p and H_p, which themselves never go bellow $200px$. For this reason,

the uncertainty due to digitization is considered negligible for the rest of the study.

3.1.2 Extrusion profile

This action was performed using the python package *ezdxf* [10]. The package essentially reads the *.dxf* file and converts it to an image. A custom script then calculates the black and white pixel count ratio and multiplies it by the width and height of the profile to deduce the profile area. This profile area can then be multiplied by the extrusion length to get the volume, and finally multiply that by the density of the alloy of aluminium to get the mass. The perimeter is also calculated to find the surface area of the heatsink by multiplying it with the extrusion length.
This file format is parametric, meaning that sampling could in theory be as precise as wanted. The width and height of the extrusion profile is directly encoded in the file format, we thus consider it to be exact. The uncertainty of the profile area is thus considered negligible.

3.2 Datasets

Because of the lack of international standard, manufacturers do not give details about measuring conditions. Therefore, we assume the following to be true:

- All heatsinks are black anodised.
- $R_{thjc} = R_{thti} = 0$, see **Equation 2**.
- $T_{amb} = 25°C$.

The content of the 3 datasets is summarized in **Table 1**.

Table 1: Datasets summary.

ID	Natural convection	Forced air	Unsampled size	Total size
A	Yes	No	264	1986
B	Yes	No	161	1117
C	Yes	Yes	57	393

3.2.1 Dataset A

This dataset contains HS of different shapes, which have been split in two groups: *fin-plate* and *non-fin-plate* (*others*). The surface S_p and perimeter P_e of each extrusion profile has been calculated from the fabrication files, and a graph of typical thermal resistance R^*_{th0} against extrusion length L has been sampled. The mass can then be calculated using **Equation 4**:

$$m = V * \rho_{al} = L * S_p * \rho \tag{4}$$

With V the volume of aluminium, ρ_{al} the volumetric mass usually equal to $2.7 g/cm^3$.

3.2.2 Dataset B

This dataset contains HS of different shapes, which have been split in two groups: *fin-plate* and *non-fin-plate* (*others*). A graph of ΔT against P has been sampled for each HS reference.

3.2.3 Dataset C

This dataset contains only *fin-plate* heatsinks, but the manufacturers separated them in two categories: *low airflow* and *high airflow*. Each reference has a given R^*_{th0}, and a graph of R_{th} against air velocity v has been sampled each time.

4 Analysis methods

This section aims to introduce the mathematical tools and visual representations used for the present work.

4.1 Power law

The power law is a statistical distribution given by **Equation 5**:

$$y = ax^{-k} \tag{5}$$

Where $x > 0, k > 0$ and $a \in \mathbb{R}$. A lot of relationships among the parameters considered in this article are approximated using this distribution. From there, the name *variable* is used to talk about a HS *parameter*.

4.2 Multi linear regression

4.2.1 Equation

Every model of the paper has been created using a multi linear regression method. The goal of this technique is to find the coefficients of a multi variable regression line to predict the independent variable y with the dependant variables x_i. An example of such a line is given by **Equation 6**:

$$\hat{y} = b_n x_n + \cdots + b_1 x_1 + b_0 \tag{6}$$

Where $\hat{y}$ represents the predicted y values, $b_n, \dots, b_1$ the nth regression coefficient, b_0 the intercept and $x_n, \dots, x_1$ the nth prediction variable.

Usually, ordinary least squares (OLS) is used. This technique finds the regression line that minimizes the square distance along the y axis. This assumes that 100% of the error lies in the predicted variable, which is usually false in this study, because often an independent variable (such as R_{th}) has the most error. It thus provides biased coefficients, and biased error in the coefficients. Because we are more interested in prediction than the underlying model, and for its ease of implementation and interpretation, OLS is used in this study [11].

4.2.1 Uncertainty

In an effort of keeping models simple, the uncertainty of a prediction is calculated using the root mean square error (RMSE) and the 95% prediction interval (PI95), shown as a yellow overlay on graphs. A 95% confidence interval (CI95) is also shown in figures hereafter, as a blue overlay, to visually see the variability of the regression line itself.

We are going to assume normality of the residuals and constant variance (homoscedasticity), and check these assumptions visually for each model.

4.2.2 Visualization

For regressions with more than one independent variable, it is complicated to visualize the regression line fitting the data across all dimensions. A good compromise chosen in this article is to create a new variable x' made up of the sum of the opposite of each independent variable multiplied by their respective regression coefficient in log space, such as shown in **Equation 7**:

$$\log x' = \sum_1^n -b_i \log x_i = -b_1 \log x_1 - \cdots - b_n \log x_n \quad (7)$$

Plotting the independent variable against x' shows how good the regression line fits the data, allowing to plot confidence intervals.

4.3 Log transformation

In this study, all data is log transformed before analysis using the decimal logarithm $\log_{10}$. Doing so is useful for multi linear regression as it transforms power law distributions, as shown in **Equation 5**, to decreasing affine lines such as shown in **Equation 8**:

$$\log y = \log a - \mathrm{k}.\log x \quad (8)$$

A growing variance in standard space can also become constant in log space, thus helping with homoscedasticity. A multi linear regression line fitted in log space then transforms in normal space with **Equation 9**:

$$\hat{y} = x_{\mathrm{n}}^{b_n} * \ldots * x_1^{b_1} * 10^{b_0} \quad (9)$$

This transformation implies that uncertainties scale as well and become relative and unsymmetrical. **Equation 10** shows that Δy should not be too different to 0 in order to keep reasonable predictions. CI95 and PI95 also scale accordingly.

$$\log y = b_1.\log x_1 + b_0 \pm \Delta y$$

$$\Rightarrow\ y = x_1^{b_1}.10^{b_0 \pm \Delta y} = x_1^{b_1}.10^{b_0} \times^{\div} 10^{\Delta y} \quad (10)$$

In other words, we have $y \in [x_1^{b_1}.10^{b_0-\Delta y}; x_1^{b_1}.10^{b_0+\Delta y}]$.

5 Analysis and results

In this section, 4 different models are described and derived with respect to different datasets and cooling technics.

5.1 HS mass, volume, surface area from thermal resistance (Dataset A)

This section presents a relationship established between a typical thermal resistance R^*_{th0} and the mass m of fin-plate heatsinks in NC using dataset A. After sampling R^*_{th0} and the extrusion length L, three variables, whose distributions can be seen on **Figure 2**, were calculated: the mass m, the box volume V_b (**Equation 1**) and the total thermal exchange surface area $S = L * P_e$ with P_e being the perimeter of the profile. From here, the color green represents fin-plate HS, and purple represents other shapes. The four variables seem to follow a power law, with non-fin-plate heatsinks being generally smaller.

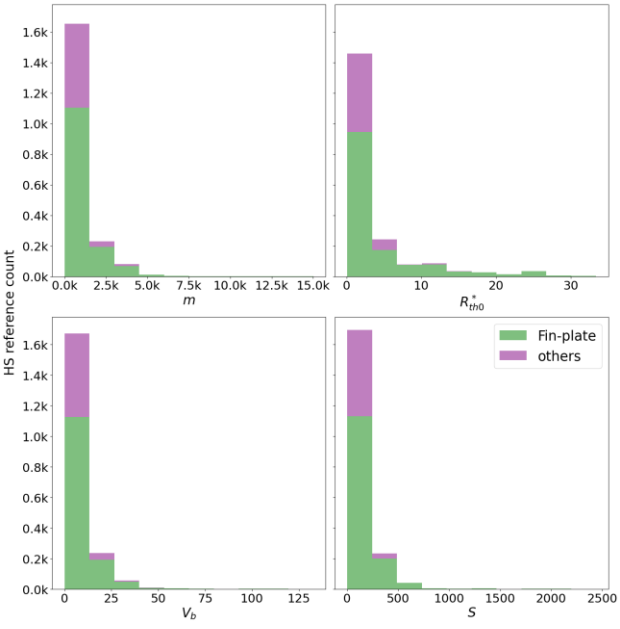

Figure 2: Distribution of $\boldsymbol{m}$, $\boldsymbol{R^*_{th0}}$, $\boldsymbol{V_b}$ and $\boldsymbol{S}$.

The correlation matrix on a log-log scale is shown on **Figure 3**. On the bottom left corner are scatter plots of all pairs of variables, on the diagonal are histograms of individual variables, and on the top right corner are correlation coefficients with p-value statistical significance. A red correlation means positive and blue means negative, while a $p < 0.005$ is considered statistically significant [11].
All the variables seem correlated to each other, and the three dependant ones are strongly negatively correlated to R^*_{th0}. This suggests that multiple univariate linear regression models can be build, and that V_b might be a good approximation of the real volume V (here represented as m, see **Equation 4**). All pairs seem homoscedastic.

The three regression lines can be visualised on **Figure 4** for the two groups at the same time. The three of them are visually a good fit ($RMSE \leq 0.17$), with residuals approximately normally distributed. Section 5.5 will summarize all regression coefficients of the models, including fin-plate-only models.

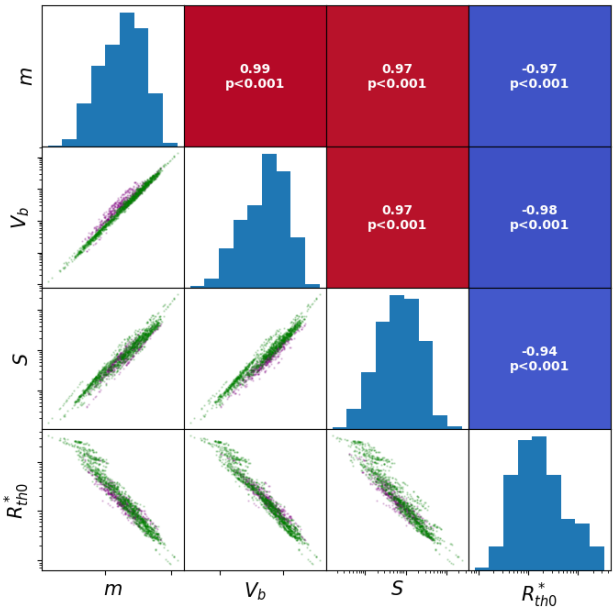

Figure 3: Correlation matrix of the variables of dataset A on log-log scale.

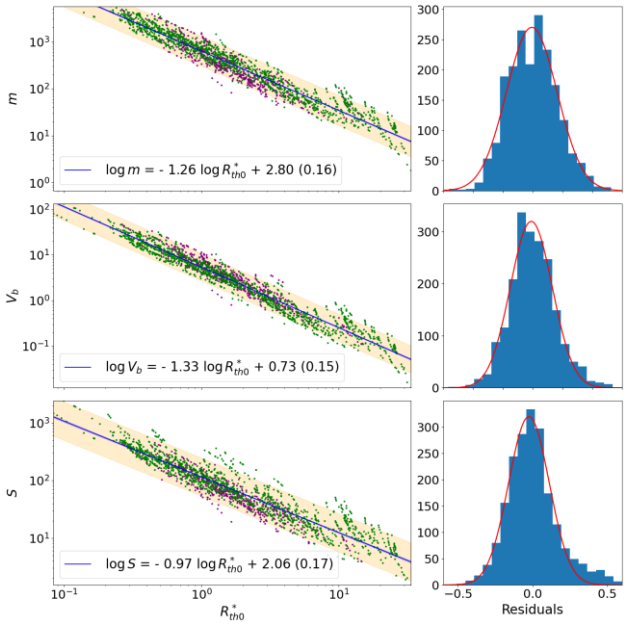

Figure 4: Regression lines of $\log m$ vs $\log R^*_{th0}$, $\log V_b$ vs $\log {R_{th0}}^*$, and $\log S$ vs $\log {R_{th0}}^*$, with residual distributions.

5.2 Thermal resistance, mass (Dataset C)

This section presents a relationship established between a typical NC thermal resistance R^*_{th0}, and the mass m of extruded aluminium heatsinks from dataset C. **Figure 5** shows the distribution of these two variables, which both seems to follow a power law. As it can be seen on the plot in Figure 6, there is a rupture in R^*_{th0} between the two groups *high_airflow* and *low_airflow*.

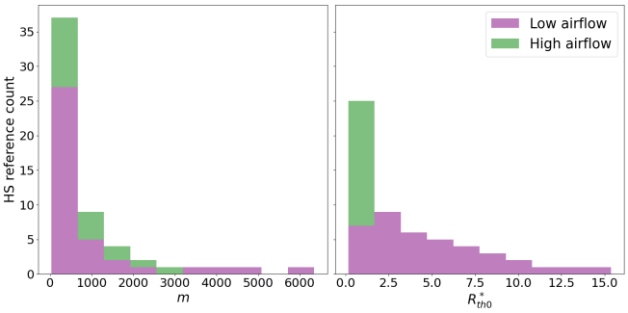

Figure 5: Distribution **of m and R^*_{th0}.**

Jumping directly to the regression analysis, the separation can be immediately seen in **Figure 6**. This can be explained by the fact that the manufacturer might have categorized a *high airflow* generally more performant heatsinks given their mass. It could also be differences in the way R^*_{th0} was measured for these two groups. Visually, the two groups seem homoscedastic, but the non-continuity makes it impossible to exploit reliably these results. This approach allows for visual inspection and validation of the data, where a neural network approach might hide flawed data.

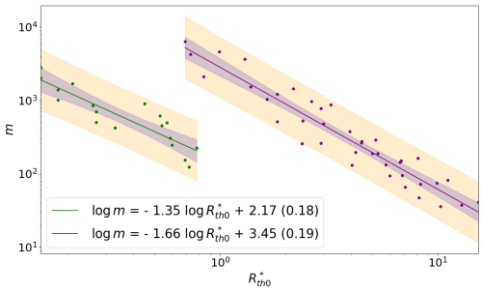

Figure 6: Regression lines of $\log m$ and $\log R^*_{th0}$.

5.3 Power, temperature, mass (Dataset B)

This section presents a relationship established between the power to dissipate P, the temperature rise above ambient ΔT and the mass of HS in NC using dataset B. A comparison with results from 5.1 is done by calculating R_{th0} from P and ΔT. The data is separated in two groups: fin-plate (green) and other (purple). After sampling P and ΔT, R_{th0} is calculated using **Equation 2** and the distribution of the four variables can be seen on **Figure 7**. m, P, and R_{th0} seem to follow a power law, while ΔT follows the uniform distribution it was sampled from.

The scatter matrix is presented on **Figure 8**. P is positively correlated to m (0.77) which is coherent with the fact that larger heatsinks are given high power to dissipate. m and ΔT are not corelated (-0.12), which is not surprising as ΔT is an index capped at $100°C$. Moreover, P is slightly correlated to ΔT (0.52) which is coherent as a high power always produces always higher temperature rise than a low power. Finally, R_{th} is highly negatively correlated to m (-0.96) which is consistent with observations from 5.1 .
Because we are focused on predicting m, it is not a problem if P is not directly correlated to ΔT, which itself is correlated to P. The rest of the analysis will be focused on the group *fin-plate* as it is dominant compared to the *other* group. Visually, all pairs seem homoscedastic.

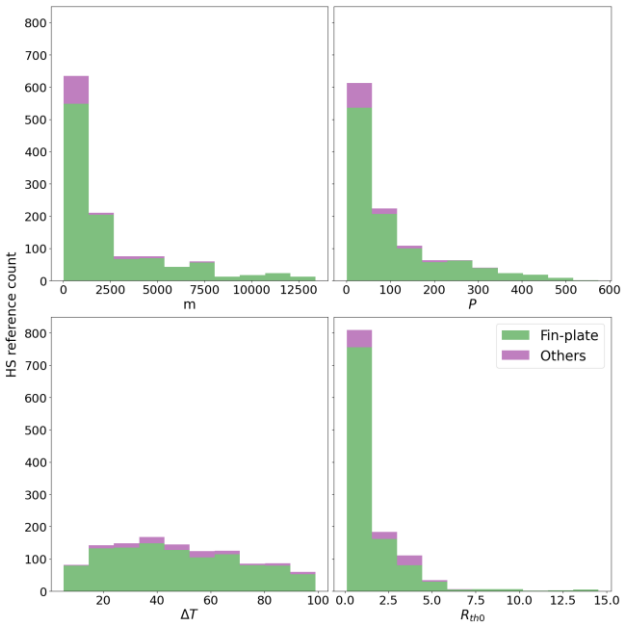

Figure 7: Distribution of m, P, ΔT, and R_{th0}.

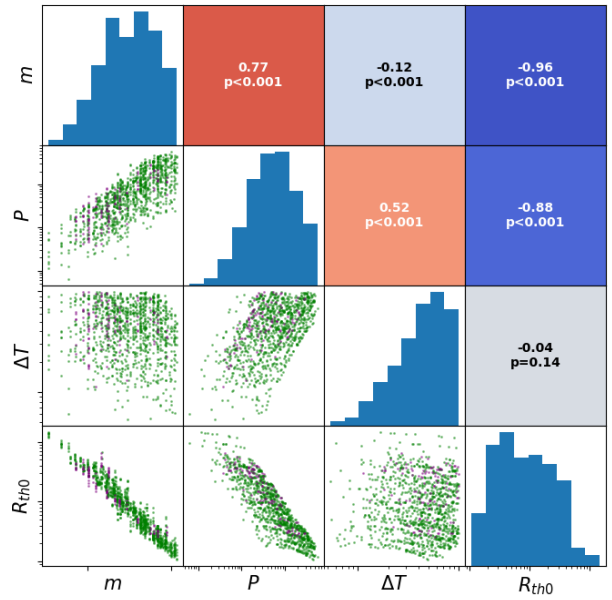

Figure 8: Correlation matrix of the variables of dataset B on a log-log scale.

The regression analysis can be visualised on **Figure 9**. The model considering P and ΔT gives better predictions than the model with only R_{th} ($RMSE = 0.13$ vs 0.17 respectively), which is coherent as a heatsink with a given mass can have multiple R_{th} in NC. The figure also shows the regression line of 5.1 for *fin-plate*. Visually, the two lines are close to each other, and coefficients are later compared in 5.5.

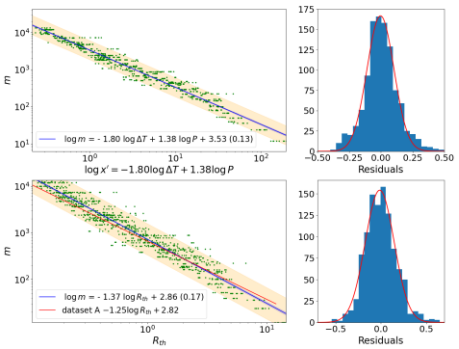

Figure 9: Regression lines of $\log m$ vs x' and $\log m$ vs $\log R_{th}$, with residual distribution.

5.4 Air velocity, thermal resistance, mass (Dataset C)

This section presents a relationship established between the thermal resistance R_{th}, the air velocity v and the mass m of fin-plate heatsinks using dataset C. After sampling v and R_{th}, the distribution of the three variables can be seen on **Figure 10**. R_{th} and m seem to follow a power law, while v was sampled using a uniform distribution, which can be seen for the two groups *low velocity* and *high velocity*.

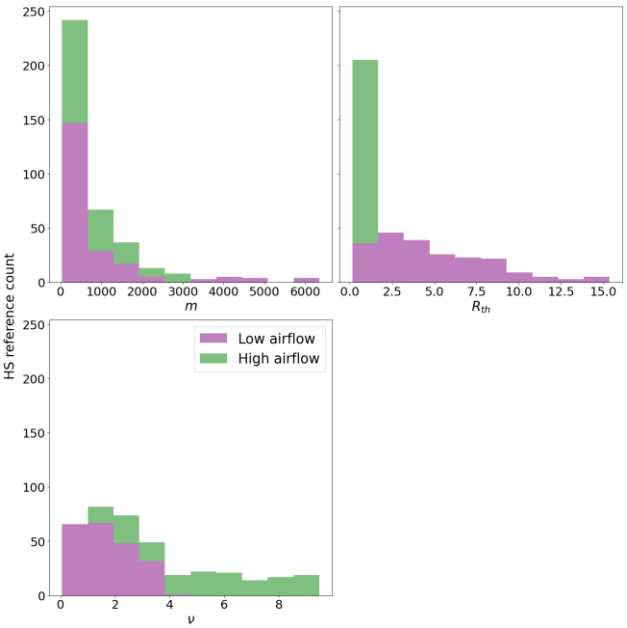

Figure 10: Distribution of m, R_{th} and v.

The correlation matrix is shown on **Figure 11**. It shows a strong negative correlation between R_{th} and v (-0.73) and a strong negative correlation between R_{th} and m (-0.79). The former is coherent with underlying thermodynamics that stipulates that R_{th} decreases with increasing air velocity, and the latter is coherent with findings from 5.1. v and m are not much correlated (0.24), which is coherent as v is sampled randomly for a given heatsink with a given mass. It is worth noting that the two groups in the scatter R_{th} vs m, with different v distributions, seem distinct from each other, but overlap on the x and y axis, which differs from 5.2 where there was a discontinuity on R_{th0}.

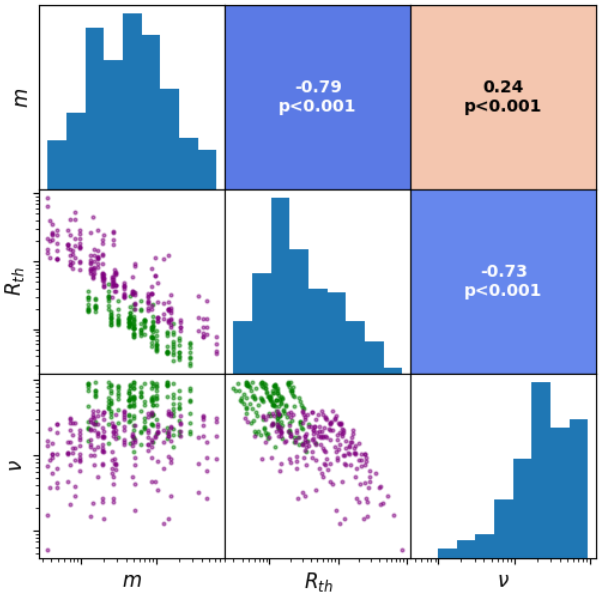

Figure 11: Correlation matrix of the variables of dataset C on a log-log scale.

The OLS regression analysis with ν and R_{th} as independent variables, and m as the dependant variable is shown on **Figure 12**. Unlike the analysis done in 5.2, the two groups seem to merge around the regression line in a continuous manner. The residuals visually seem normally distributed.

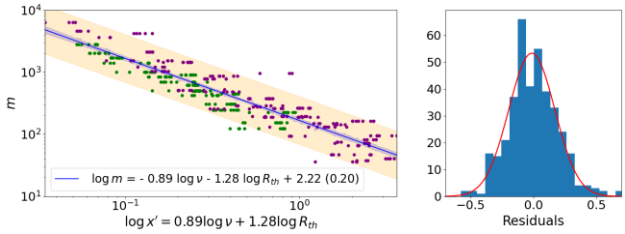

Figure 12: OLS regression line of $\log \boldsymbol{m}$ vs standardized $\boldsymbol{x'}$ with residual distribution.

5.5 Results summary

A summary of the coefficients and RMSE for each regression model can be found in **Table 2**. fp stands for *fin-plate* and all is used when all heatsink shapes are mixed together.

A few observations from these results can be made.
First, the RMSE of $V_{b,all}(R^*_{th0})$ is lower than $V_{b,fp}(R^*_{th0})$, same with $S_{all}(R^*_{th0})$ and $S_{fp}(R^*_{th0})$, which is counter-intuitive. In fact, as the all subset has more diversity than the fp subset, one should expect the RMSE to be higher as the data is more spread out. However, **Figure 4** shows the opposite, as the purple dots representing the non-fin-plate heatsinks are packed near the regression line, lowering the RMSE.
Second, the intercepts b_0 of $m_{fp}(R^*_{th0})$ (Dataset A) and $m_{fp}(R_{th0})$ (Dataset B) are close to each other while b_1 coefficients are different. This means that the regression lines are almost the same around $R_{th} = 1$ ($\log R_{th} = 0$), but then they diverge. This can be explained by differences in sample distribution, and because of the assumptions we made in 3.2: the conditions of measurements were probably different. This means that typical R^*_{th} values should not be used as a precise characteristic.
Finally, the model using both P and ΔT from Dataset B gives better predictions than the model using only R_{th0}, which can be visually seen on **Figure 9**. In fact, the upstream current of air increases with increasing temperature (thermal expansion of the gas), which changes R_{th0}, so one HS can actually have multiple thermal resistance given its power and temperature [7].

Table 2: Models summary, with coefficients and RMSE.

Dataset	Model	Coefficients	RMSE
A	$m_{all}(R^*_{th0})$	$b_0 = 2.797$ $b_{Rth0^*} = -1.259$	0.164
A	$m_{fp}(R^*_{th0})$	$b_0 = 2.819$ $b_{Rth0^*} = -1.245$	0.151
A	$V_{b,all}(R^*_{th0})$	$b_0 = 0.728$ $b_{Rth0^*} = -1.326$	0.152
A	$V_{b,fp}(R^*_{th0})$	$b_0 = 0.707$ $b_{Rth0^*} = -1.308$	0.155
A	$S_{all}(R^*_{th0})$	$b_0 = 2.062$ $b_{Rth0^*} = -0.968$	0.169
A	$S_{fp}(R^*_{th0})$	$b_0 = 2.094$ $b_{Rth0^*} = -0.955$	0.172
B	$m_{fp}(R_{th0})$	$b_0 = 2.859$ $b_{Rth0} = -1.370$	0.171
B	$m_{fp}(\Delta T, P)$	$b_0 = 3.527$ $b_{\Delta T} = -1.799$ $b_P = 1.382$	0.131
C	$m_{fp}(\nu, R_{th})$	$b_0 = 2.223$ $b_\nu = -0.889$ $b_{R_{th}} = -1.279$	0.195

6 Experimental illustration

Model $m_{fp}(P, \Delta T)$ has been tested with data taken from a datasheet of a P3 HS from SEMIKRON [12]. A plot of R_{th0} against L for multiple power P has been sampled for the test. The HS is not fin-plate, but nonetheless most of the data points fall within the PI95, as it can be seen on Figure 13.

The actual mass m might visually seem close to the predicted one $\hat{m}$, but the difference grows exponentially in normal space, as stated in **Equation 10**. Here, the mean relative difference between $\hat{m}$ and m is $\overline{\sigma_m} = -32.7\%$, which can be compared with individual samples in Table 3. While the relative difference σ_m stays stable for increasing mass, the absolute error Δm greatly increases with increasing mass, which is consistent with a power law model.

These results are satisfying for a rough estimate of the mass given the temperature rise and power, even for non-fin-plate HS

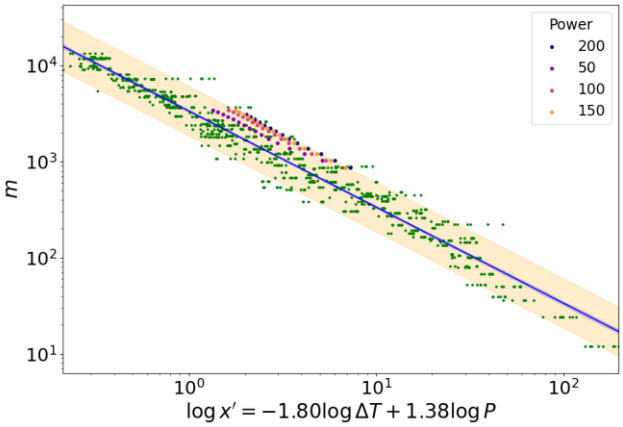

Figure 13: OLS regression line of $\log m$ vs standardized x', with test samples.

Table 3: Data samples from SEMIKRON, with prediction results.

P	R_{th0}	ΔT	m	$\hat{m}$	Δm	σ_m
50	0.75	37.5	1378	1082	-296	-21%
50	0.45	22.5	3444	2755	-689	-20%
200	0.81	161	861	509	-352	-41%
200	0.4	80	3100	1831	-1269	-41%

7 Further work and perspectives

This work is an introduction to the vast subject of data analysis applied to heat dissipation. Iterations could improve the diversity of models by including anodization and ducting effects, other heatsink technologies, other materials, better and uncertainty assessment. Observations from 5.5 show that the models presented in this paper should be used as a rough estimate, as conditions of measuring R_{th} changes from one manufacturer to another, and are likely to be different from those in an actual application.

These models could be implemented in optimization loops, not only for eco-design, but also for mass and volume power density estimates. Indeed, as mass and total volume can be estimated from R_{th} using a simple relation, integration in design is faster and computation is lighter than traditional simulations, analytic formulas or neural networks

Furthermore, parametric models such as those presented in this paper can be integrated in life cycle inventory analysis (LCIA) to support life cycle assessment (LCA). For instance, one could estimate the mass of aluminium that needs to be extruded to make a fin-plate heatsink from its thermal resistance. This is useful for exploratory LCIA, but also when the mass a particular heatsink is not given by the manufacturer and cannot be measured. Other parametric life cycle inventory (P-LCI) models can be created to further complete the LCIA, such as a relation with the surface and the thermal resistance for anodization. Uncertainty expression can seamlessly be used to generate a distribution for the inputs of the LCI, as it is commonly done in the field. The analysis could then be pushed further by evaluating the environmental impacts of the heatsink of a power conversion system at early stages of the design, mainly from the technical parameter R_{th}.

8 Conclusion

The present paper establishes new relations between thermal resistance, mass, volume and total surface area, with respect to technology parameters, by studying "real-world" heatsinks. These relations were built using data from datasheets and catalogues from manufacturers. These new relations can be integrated for example in power electronics power density optimization processes to consider heatsink mass and volume. They do not require heavy simulations or complex mathematical formulas, and can be deployed for multiple use cases such as natural or forced convection.

9 Acknowledgement

This research work is part of ARCHIMEDES project, supported by the Chips Joint Undertaking and its members, including the top-up funding by National Authorities under Grant Agreement No 101112295.

10 Literature

[1] S. M. I. Rahman, A. Moghassemi, et al., *"Emerging Trends and Challenges in Thermal Management of Power Electronic Converters: A State of the Art Review," IEEE Access*, 2024.

[2] D. Christen, M. Stojadinovic, et al., *"Energy Efficient Heat Sink Design: Natural Versus Forced Convection Cooling," IEEE Trans. Power Electron.*, 2017

[3] S. W. Churchill and H. H. S. Chu, *"Correlating equations for laminar and turbulent free convection from a vertical plate," Int. J. Heat Mass Transf.*, 1975

[4] F. H. Rafael, V. A. da Silva, et al., *"Influence of radiation view factor and geometric parameters on thermal analysis of natural convection plate-fin heat sinks: A numerical approach with experimental validation," Int. J. Therm. Sci.*, 2025

[5] O. Kosdere, Z. Sert, et al., *"Investigation of thermal performance at forced convection in plate-fin heat sink" Energy*, 2024

[6] Z. Wang, Y. Zhang, et al, *"Machine learning-based surrogate models for finned heatsink thermal modeling", Microelectronics Reliability*, 2025

[7] F. P. Incropera, D. P. DeWitt, *"Fundamentals of Heat and Mass Transfer"*. New York: *John Wiley and Sons*, 1990.

[8] T. Abry, J.-C. Crebier, *"Software and Data - Data-driven model to derive air-cooled heat sinks mass for power electronics"*. Zenodo, déc. 19, 2025. doi: 10.5281/zenodo.17987325.

[9] plotdigitizer: *Extract raw data from plots images*. Python. https://github.com/dilawar/PlotDigitizer

[10] ezdxf: *A Python package to create/manipulate DXF drawings*. Python. https://github.com/mozman/ezdxf

[11] S. Weisberg, *Applied Linear Regression*. Hoboken, NJ: Wiley, 2014.

[12] *"Datasheet P3/120 P3/180."*, SEMIKRON.