Optimization and metamodelization based on machine learning of a new neuro human thermal model

Mohamad EL KADRI[1], Fabrice DE OLIVEIRA[1], Driss SAMRI[1], and Virginie DESVIGNES[1]

1 Centre Scientifique et Technique du Bâtiment CSTB
84 Avenue Jean Jaurès
77420, Champs-sur-Marne, France
**Corresponding author: fabrice.deoliveira@cstb.fr*
Presenting author: Fabrice DE OLIVEIRA

ABSTRACT

Nowadays, due to climate change, heatwaves become stronger in terms of frequency and intensity. This phenomenon can have serious impact on the indoor environments, indoor thermal comfort and on public health. These situations of high indoor thermal conditions can expose the occupants to health risks such as hyperthermia, dehydration, and heat strokes. Then, the estimation of these risks is crucial. The currently used indices to estimate health risks such as WBGT, HSI and PHS are generally dedicated to outdoor environments and for subjects exerting heavy activities. In addition, these indices do not consider the thermophysiological responses of the human body. In 2020, a human thermoregulation model, called NHTM, was developed to mimic the thermal behavior of the human body in indoor non-uniform and transient conditions. The outputs of the NHTM are the central temperature and the water loss that can be used to assess the health risks. This model considers the interindividual differences between different populations by adjusting its parameters using genetic algorithm and choosing the values that correspond to the studied population. The present study aims to 1) optimize the NHTM model using genetic algorithm on Stolwijk and Hardy study and 2) simplify the thermophysiological calculation by developing metamodels that reduce the calculation time and the complexity of a non-uniform calculation using the NHTM model.

KEYWORDS

Climate change, thermal comfort, health risk, thermophysiology, individualization.

1 INTRODUCTION

The industrial revolution, the burning of fossil fuels, and the excessive use of resources that began since the 18th century are the main causes of the greenhouse gas emissions and the rise of global temperatures (Masson-Delmotte et al., s. d.; Valone, 2021).

The climate change is the highest global health and wellbeing threat of the 21st century. As the planet warms and extreme weather events become more frequent, the effects on human health multiply. The impacts are manifold, ranging from infectious diseases to respiratory disorders and mental health issues (Costello et al., 2009). Understanding these complex links between climate and health is crucial for better anticipating and managing the adverse consequences.

Rising average temperatures have direct implications for human health. Increasingly frequent and intense heatwaves raise the risk of cardiovascular diseases, dehydration, heat exhaustion, and other heat-related conditions (Patz et al., 2007). Furthermore, climate changes can affect air quality by increasing levels of air pollution, exacerbating respiratory problems such as asthma and lung diseases (Frumkin et al., 2008).

The assessment of the effects of rise of temperature on the wellbeing and the health of occupants is very hard due to the complexity of the human body in terms of physiological and sensory responses, and the differences of these responses between individuals. The currently standardized methods (EN 16798-1:2019, 2019) to assess the thermal comfort in indoor environments are based on the PMV (Fanger, 1970) for air-conditioned environments and the adaptative comfort for non-conditioned environments (de Dear et al., s. d.). Many other indicators are used to assess the heat stress on the human body such as UTCI (Bröde et al., 2012), and WBGT (Budd, 2008) indexes. All these indicators are based on a “mean person” and cannot consider the interindividual differences between many types of populations.

In 2020, El Kadri et al. (El Kadri, 2020) developed a thermal human model, the NHTM, which mimics the thermophysiological responses of the human body in non-uniform transient environments. This model is based on that of Wissler developed for the NASA (Wissler, 2018). The NHTM is a complementary to existing standards and indices for 2 reasons: 1) it is not limited by uniform steady environments, 2) it can simulate many types of populations according to interindividual differences. The aim of this article is to describe the NHTM, to explain how to optimize it for a target population and how to simplify it via metamodels.

2 MATERIALS AND METHODS

The NHTM (Neuro Human Thermal Model) is a new thermophysiological model based on the thermoreceptor’s signals. It consists of two systems: the passive system which accounts for phenomena of human heat transfer within the body and at its surface and the active system model which simulates the physiological responses of the human body such as the shivering, the skin blood flow and the sweating based on thermoreceptor’s signals (El Kadri, 2020). The heat transfer and the physiological responses are governed by parameters that we can change to simulate a target population. That can be done thanks to optimization which can be performed via a genetic algorithm. To do it, a database of measurements of physical and physiological variables is needed. The population should be exposed to a scenario of environmental conditions and the physical variables such as air temperature, air velocity, relative humidity, and radiant temperature (inputs of the NHTM) and the physiological variables such as skin and core temperatures (output of the NHTM) should be measured. An optimization using this algorithm is described in a previous work (El Kadri et al., 2020).
The simulation using the NHTM is costly in terms of time and resources. Hence the use of the metamodel.

Meta-models are regression or statistical functions of the calculation code, which are constructed based on 'n' simulations following a numerical experiment design. The main objective of these functions is to approximate the responses of the original code at a fraction of its computational cost. Essentially, these meta-models provide a method to predict new responses within the range of uncertain parameters with reasonable accuracy.

The approach adopted in the current research is illustrated by the figure below (Figure 1). A design of experiments, which will be presented in the subsequent paragraph, has been established. This experimental design aims to explore the space of the model parameters systematically and efficiently, enabling the construction of a reliable and accurate meta-model.

In essence, the development of these meta-models serves to simplify the thermophysiological calculations using the NHTM model, reducing the calculation time and complexity while still maintaining a high level of accuracy and reliability in predicting thermophysiological responses. This not only enhances the utility of the NHTM model in various practical applications but also facilitates its use in studying and understanding the complex thermophysiological responses of the human body under different environmental conditions.

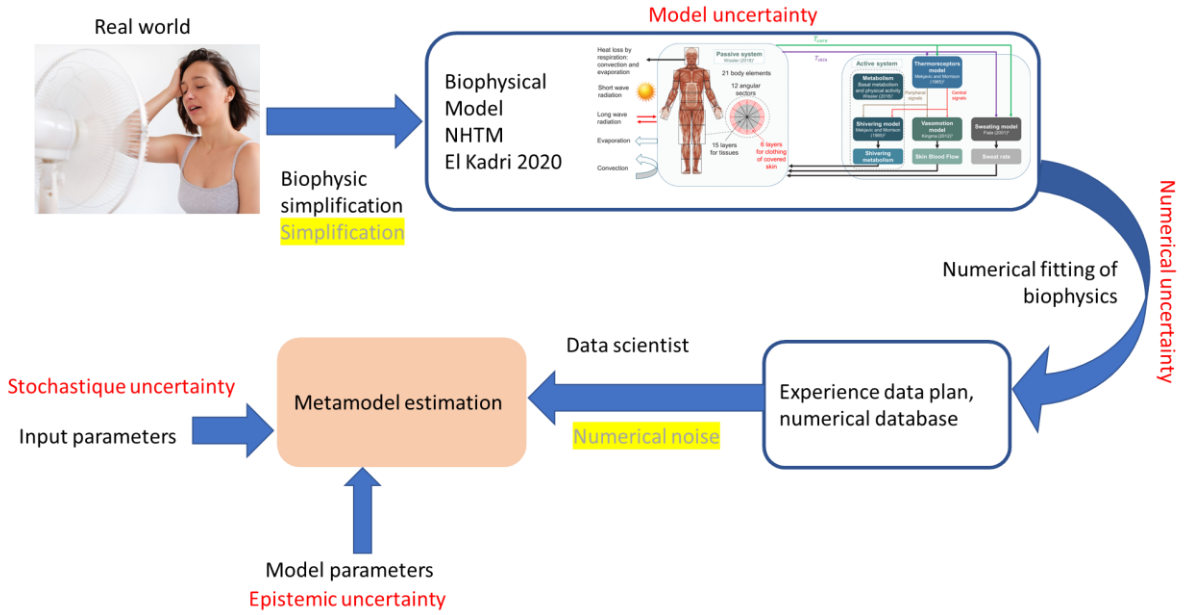

Figure 1: Uncertainty sources

2.1 Model setup and numerical experience plan

The individualization carried out within the scope of the thermal comfort research allowed the identification of parameters associated with a young and health population. The environmental variables were chosen to cover the most frequently encountered thermal conditions by the occupants but also the extreme conditions. The air and mean radiant temperatures are higher than 10 °C since such thermal conditions are rare in indoor environments. The occupant is initially considered at rest (sitting). Therefore, the metabolism does not vary. The minimums and maximums of the environmental and personal variables are given in Table 2. The total number of simulations is 16896.

Table 1: Environmental and personal variables used in the meta-models' constructions

Variable	Minimum	Maximum	Step
Air temperature [°C]	10	45	5
Mean radiant temperature [°C]	10	45	5
Relative humidity [%]	0	100	10
Air velocity [m/s]	0	1	0.2
Clothing (Clo)	0.1	1.5	0.5
Metabolism (W/m²)	58	58	0

2.2 Regression methods

There are several types of metamodels:

- Linear regression model or generalized linear model (easy interpretability, (Nelder & Wedderburn, 1972));
- Support Vector Machine (SVM, easy interpretability for linear, complex for kernels, (Cortes & Vapnik, 1995));
- Regression trees (easy interpretability, (Breiman et al., 1984));
- Neural networks (complex interpretability, (Hinton & Salakhutdinov, 2006));
- Conditional Gaussian processes or kriging (complex interpretability, (Schwarz et al., 2009));
- Tree ensembles (complex interpretability, (Dietterich, 2000)).
- A screening of 27 regression methods was performed. The following table presents the different methods tested for the construction of metamodels

The criteria used to assess the quality of metamodels are the RMSE (Root Mean Square Error), the R^2, the MSE (Mean Square Error) and the MAE (Mean Absolute Error)

3 RESULTS OF METAMODELS FITTING

This chapter presents the results of constructing metamodels using the various methods previously discussed and the 16900 calculated cases.

3.1 Synthesis of results

The construction results of the different metamodels are shown in the Table 2. The models that provided the best fit belong to the Gaussian Process Regression (GPR) family, specifically the Matern 3/2 kernel family (Table 2).

Table 2:Results of Different Metamodels

Regression type	Regression model	RMSE	R2	MSE	MAE
Linear	Linear	1.0953	0.9	1.1988	0.83
	Interactions	0.5613	0.97	0.301	0.446
	Robust	1.103	0.89	1.2181	0.834
	Stepwise	0.5613	0.97	0.315	0.446
Tree	Fine	0.162	1	0.026	0.105
	Medium	0.245	0.99	0.06	0.174
	Coarse	0.45	0.98	0.205	0.326
SVM	Linear SVM	1.103	0.89	1.21	0.833
	Quadratic SVM	0.459	0.98	0.21	0.33
	Cubic SVM	0.299	0.99	0.0899	0.243
	Fine gaussian SVM	0.334	0.99	0.112	0.266
	Medium Gaussian SVM	0.181	1	0.327	0.148
	Coarse Gaussian SVM	0.395	0.99	0.157	0.297
Gaussian process	Rational quadratic	0.0617	1	0.0038	0.038
	Squared exponential	0.086	1	0.0074	0.056
	Matern 5/2	0.059	1	0.00346	0.0344
	Matern 3/2	0.052	1	0.0027	0.027
	Exponential	0.067	1	0.0045	0.0363
Kernel	SVM	1.19	0.88	1.415	0.779
	Least squares regression	1.08	0.9	1.164	0.77
Tree set	Boosted trees	1.47	0.81	2.16	1.37
	Bagged trees	0.27	0.99	0.071	0.194
Neural network	Narrow neural network	0.684	0.96	0.467	0.534
	Medium neural network	0.66	0.96	0.44	0.517
	Wide neural network	0.61	0.97	0.368	0.46
	Bilayered neural network	0.546	0.97	0.299	0.402
	Trilayered neural network	0.294	0.99	0.086	0.21

3.2 Results for matern 3/2 Gaussian Process

The Figure 2 shows the data points from the different simulations in blue. The orange points represent the estimations made by the GPR metamodel. These estimations are quite satisfactory from the point of mean squared error (±0.056 °C).

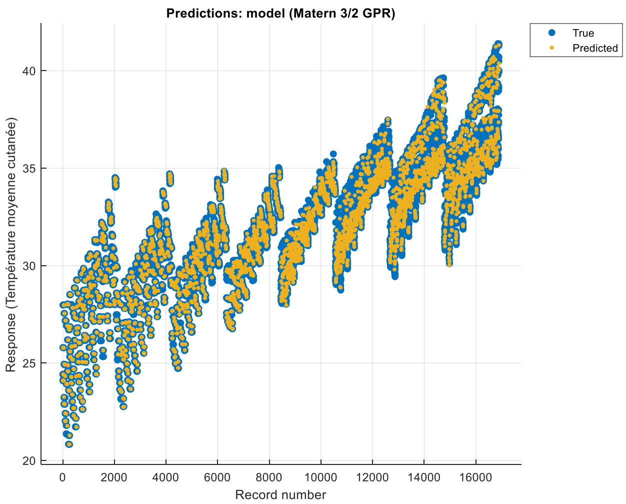

Figure 2: Simulated and Predicted Average Skin Temperature Values

The values predicted by the GPR metamodel thus show a good agreement with the simulated values (Figure 3). The main advantage of this metamodel, as previously mentioned, is its speed in estimating the variable of interest, i.e., the average skin temperature. The identified GPR model allows for an estimation of 8000 observations/s. An adaptation of the identified model is planned in Python to make the metamodel portable in a free environment.

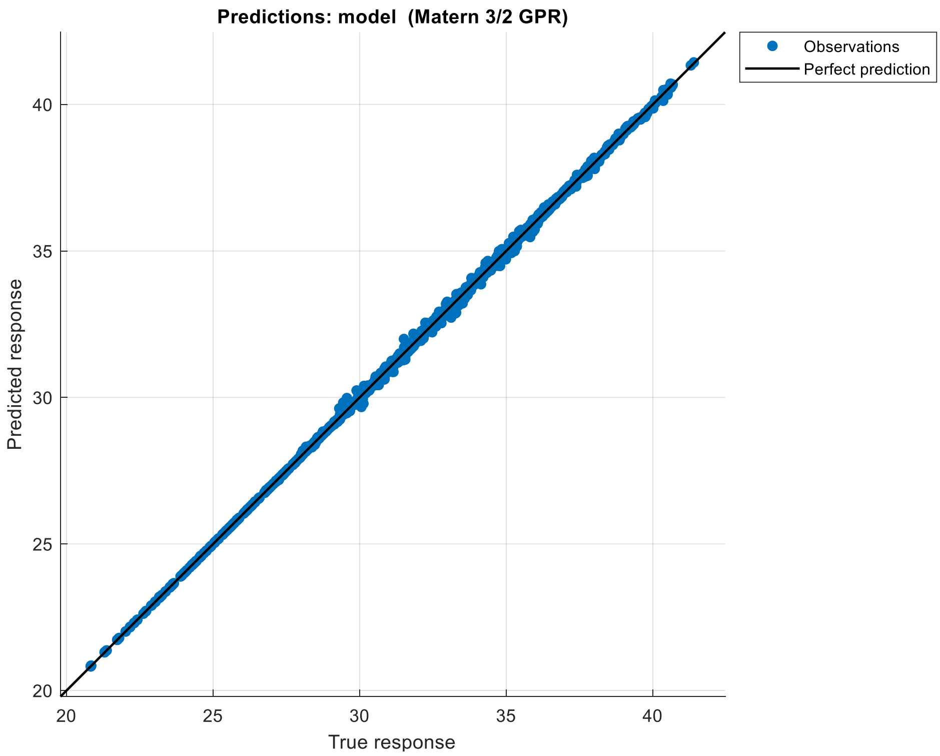

Figure 3: Function between the Simulated and Predicted Average Skin Temperature Values

The Figure 4 shows the residuals of the identified GPR model based on the simulations. 99% of the observed differences between the simulations and the estimations by the GPR model fall within the interval ±0.156°C. Moreover, as expected, the distribution of the residuals follows a normal distribution, N(0,0.052).

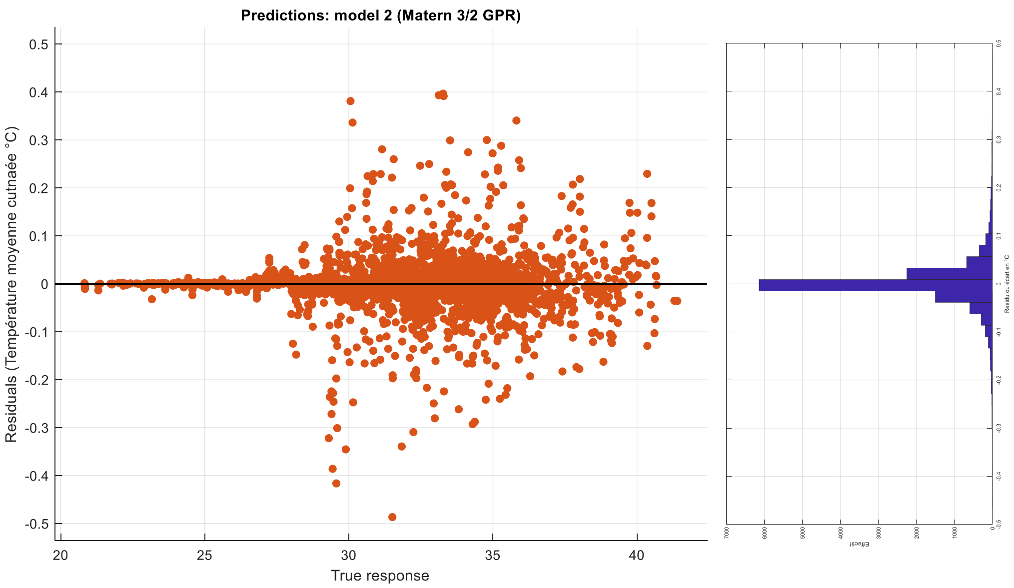

Figure 4: Residuals between the Estimates and the Simulations

4 DISCUSSION

The performance of metamodels greatly depends on the specifics of the learning task they are applied to, and there isn't a single model that is best for all tasks. However, in your case, it appears that Gaussian processes with a Matérn 3/2 kernel produced the best results.

There are several reasons why this might be the case:

1. Flexibility: The Matern kernel is very flexible and can adapt to a variety of data types. It is less smooth than the Gaussian kernel, which may enable it to better capture more complex relationships in the data.
2. Robustness to overfitting: The Matern kernel is also known for its robustness against overfitting, especially compared to other kernels such as the Radial Basis Function (RBF). This can make it more stable in situations where the number of sampling points is limited relative to the complexity of the model.
3. Effective extrapolation: Due to their smooth and continuous characteristics, Gaussian processes with Matern kernels are generally capable of effective extrapolation, i.e., making reliable predictions outside of the space covered by the sampling points. This could be a useful feature given the nature of your modeling task.

It is important to note that despite their performance in your specific case, Gaussian processes with a Matern kernel are not necessarily the best choice for all regression problems. The choice of model depends on many factors, including the nature of the data, the number of samples available, the dimensionality of the input space, and so on. A good practice in machine learning is to try several models and select the one that performs best according to an appropriate performance measure.

5 CONCLUSION AND PERSPECTIVES

In conclusion, the NHTM was built to simulate the thermophysiological behaviour of the human body in non-uniform transient thermal environments. Moreover, it can simulate many types of populations by changing its parameters. Then, the real challenge is to find the parameters that correspond to the target population. This can be done using a genetic algorithm. To do so, a database containing the measurements of physical and physiological variables of the target population exposed to a scenario of transient environmental conditions.

Metamodels are powerful tools for simplifying complex computational models and simulation studies, enabling near-instantaneous prediction of system responses, even in the face of uncertain parameters. They provide us with a systematic way to explore high-dimensional input spaces and understand the influences of different parameters on the system's behavior.

The use of metamodeling techniques allows for rapid exploration of the parameter space, sensitivity analysis, uncertainty propagation, and model calibration. They provide the means to approximate response surfaces, construct emulators, and analyze the behavior of the system across a wide range of conditions.

In our study, we have employed multiple regression methods for the construction of the metamodels and found Gaussian Process Regression, specifically with the Matern 3/2 kernel, to be most effective for our dataset. This kernel demonstrated superior performance in fitting and predicting the response of our system, thereby attesting to its practicality and reliability.

Despite their inherent complexity, metamodels serve as an indispensable tool for understanding and predicting the behavior of complex systems. By refining these models and incorporating newer, more advanced techniques, we can develop metamodels that not only improve predictive accuracy but also provide a deeper understanding of the system dynamics they represent.

Metamodels, therefore, are not merely a tool for computation; they are a cornerstone for knowledge discovery and a compass guiding us towards more insightful and meaningful conclusions in our ongoing journey to comprehend the complex systems around us.

Future perspectives of this research might be oriented towards exploring non-deterministic methods such as Hamilton Monte Carlo (HMC) sampling and Markov Chain Monte Carlo (MCMC) for model fitting, speeding up the metamodel, and quantifying uncertainties. HMC is a sophisticated variant of the MCMC approach that uses gradient information to inform the sampler's proposals, thereby improving sampling efficiency, especially in high-dimensional problems. It has the potential to explore the parameter space more efficiently, thereby reducing the time for model fitting. MCMC, on the other hand, is a widely used statistical sampling technique used to estimate the probability distribution of a population from a sample. By applying MCMC sampling to the metamodeling problem, it is possible to quantify uncertainties in a more in-depth manner, thereby providing a measure of the uncertainty of the metamodel's predictions.

6 REFERENCES

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees*. Taylor & Francis.

Bröde, P., Fiala, D., Błażejczyk, K., Holmér, I., Jendritzky, G., Kampmann, B., Tinz, B., & Havenith, G. (2012). Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). *International Journal of Biometeorology*, *56*(3), 481-494. https://doi.org/10.1007/s00484-011-0454-1

Budd, G. (2008). Wet-bulb globe temperature (WBGT)-its history and limitations.pdf. *Journal of Science and Medecine in Sport*, *11*, 20-32.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273-297. https://doi.org/10.1007/BF00994018

Costello, A., Abbas, M., Allen, A., Ball, S., Bell, S., Bellamy, R., Friel, S., Groce, N., Johnson, A., Kett, M., Lee, M., Levy, C., Maslin, M., McCoy, D., McGuire, B., Montgomery, H., Napier, D., Pagel, C., Patel, J., … Patterson, C. (2009). Managing the health effects of climate change. *The Lancet*, *373*(9676), 1693-1733. https://doi.org/10.1016/S0140-6736(09)60935-1

de Dear, R., Brager, G., & Cooper, D. (s. d.). *Developing an Adaptive Model of Thermal Comfort and Preference*. 312.

Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. *Multiple Classifier Systems*, 1-15. https://doi.org/10.1007/3-540-45014-9_1

El Kadri, M. (2020). *Thermo-neurophysiological Model of the Human Body for the Study of Thermal Comfort in Non-uniform Transient Climatic Conditions (In French)*. PhD Dissertation, La Rochelle University.

El Kadri, M., Oliveira, F. D., Inard, C., & Demouge, F. (2020). Optimization of a neuro-human thermal model using a genetic algorithm. *Indoor and Built Environment*, 1420326X2097519. https://doi.org/10.1177/1420326X20975195

EN 16798-1:2019. (2019). *Performance énergétique des bâtiments—Ventilation des bâtiments* (ISSN 0335-3931).

Fanger, P. O. (1970). *Thermal comfort. Analysis and applications in environmental engineering*. DANISH TECHNICAL PRESS. Copenhagen, Denmark. http://doi.org/10.1177/146642407209200337

Frumkin, H., Hess, J., Luber, G., Malilay, J., & McGeehin, M. (2008). Climate Change : The Public Health Response. *American Journal of Public Health*, *98*(3), 435-445. https://doi.org/10.2105/AJPH.2007.119362

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*, *313*(5786), 504-507. https://doi.org/10.1126/science.1127647

Masson-Delmotte, V., Zhai, A., Pirani, S., Connors, C., & Berger Péan, S. (s. d.). *CLIMATE CHANGE 2021 THE PHYSICAL SCIENCE BASIS* (1ST ED). CAMBRIDGE UNIV PRESS UK.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), 370-384. https://doi.org/10.2307/2344614

Patz, J. A., Gibbs, H. K., Foley, J. A., Rogers, J. V., & Smith, K. R. (2007). Climate Change and Global Health : Quantifying a Growing Ethical Crisis. *EcoHealth*, *4*(4), 397-405. https://doi.org/10.1007/s10393-007-0141-1

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling : Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*(6), 632-654. https://doi.org/10.1002/tea.20311

Valone, T. F. (2021). Linear Global Temperature Correlation to Carbon Dioxide Level, Sea Level, and Innovative Solutions to a Projected 6°C Warming by 2100. *Journal of Geoscience and Environment Protection*, *9*(3), Article 3. https://doi.org/10.4236/gep.2021.93007

Wissler, E. H. (2018). *Human temperature control*. Springer Berlin Heidelberg.