A Conceptual Framework for Personalization of Indoor Comfort Parameters Based on Office Workers' Preferences

Saeed Mirzaeifar[(✉)], Pedram Daee, and Vishal Singh

Aalto University, Espoo, Finland
`saeed.mirzaeifar@aalto.fi`

Abstract. Prevention of building-related illnesses and improving indoor air quality has become an emerging research area not only because of the comfort of workers in an office or the quality of the perceived air, but also because it can provide financial benefits to both employees and employers through a potential reduction in prolonged sick leaves. Therefore, building facility managers attempt to achieve the most comfortable and healthy environment conditions for the office workers. However, the parameters associated with achieved comfort vary from person to person as workers` preferences, as well physiological characteristics, are heterogeneous. In the ideal case, the indoor health parameters should be personalized based on individuals` feedback. This paper presents a computational framework for personalization of environmental parameters based on limited office workers' feedback. We propose that by using current state of the art machine learning methods it is possible to learn the preference model of individuals, by employing both the limited feedback and the relevant literature on health-related symptoms. The framework is explained and discussed in a potential example scenario. Evaluation based on real data is left as a future work.

Keywords: Personalization · Building-related symptoms
Indoor health parameters · Building information modelling · Office building
Worker preference

1 Introduction

People in developed countries spend most of their time indoors [1, 2]. Therefore, the quality of indoor environment parameters such as temperature, humidity, lighting, noise level and even chemical contaminants can affect a person's health and productivity, especially during the worktime activity [3]. A few studies have pointed that the trending emphasis on transforming building technology primarily for energy savings may trigger inaccuracy and depreciation in Indoor Air Quality (IAQ) [4]. On the other hand, IAQ is defined in the terms of the reduction in building-related health problems, and increase in the occupants' satisfaction level of comfort [5]. Based on the World Health Organization report, a group of building health-related symptoms, which are mostly well-known such as nose, eye and throat irritation; sensation of dry mucous

membrane; headache and mental fatigue and etc. can be characterized as Building Related Symptoms (BRS) [3].

In parallel with the focus on indoor air quality, nowadays, modular and distributed office facilities are becoming popular, with the aim to eliminate the problems of megalopolises such as traffic congestion, accessibility, and to increase the energy savings [6]. Furthermore, modular offices can be located in different accessible points and even as home offices. The occupants are often given the ability to control the physical environmental parameters such as Heating Ventilation and Air-Conditioning (HVAC) and lighting in the modular offices [7]. However, it may also be desirable that the HVAC and other technical systems adjust themselves based on the occupant preferences as there is the individual difference in people susceptibility in feelings and perceived thermal comfort [8]. Hence, there is a need to train a system with appropriate user feedback data complemented with knowledge in literature, to identify user preferences for providing the most comfortable environmental condition as and when needed. In this respect, the research question can be formulated as: "How to learn and personalize the indoor health parameters based on the office workers' feedback?". The first step to resolve research question is to anticipate the situations when users feel uncomfortable during work time with the aid of feedback. In this respect, a review of the health-related symptoms and associated indoor parameters provides the proposed framework with the complementary knowledge. The second step is to employ suitable mathematical models for learning, based on the literature knowledge as well as the targeted data sources. The proposed mathematical model and the initial solution are presented in Section Three, with a high-level architecture for BIM integration with the aim to automate the process of data collection and analysis. In Section Four, a potential use case example before the conclusions, is introduced.

2   Comfort and Health-Related Parameters

The previous studies [9, 10] categorized the factors threatening the comfort or even health of human into three main groups:

- BRS: nonspecific symptoms with unknown cause.
- Comfort complaints: inconveniences about the environmental situation such as thermal, noise or malodor complaints.
- Building-related illnesses or building related diseases include health problems because of pollutants and contaminants coming from outdoor air and building materials.

This categorization aids to extract and classify the contributing indoor health problems parameters in office buildings, as shown in Fig. 1.

The first step in the personalization of parameters is to identify the factors that are associated with the user feedback. There are two types of questions in the questionnaire including IAQ factors and perceived health-related symptoms. Answer to IAQ factor questions provides clear knowledge about the user preferences and do not need additional processes for knowledge elicitation and formulation. An oversimplified scenario can be if the user feels too cold or dryness of the air, the temperature and
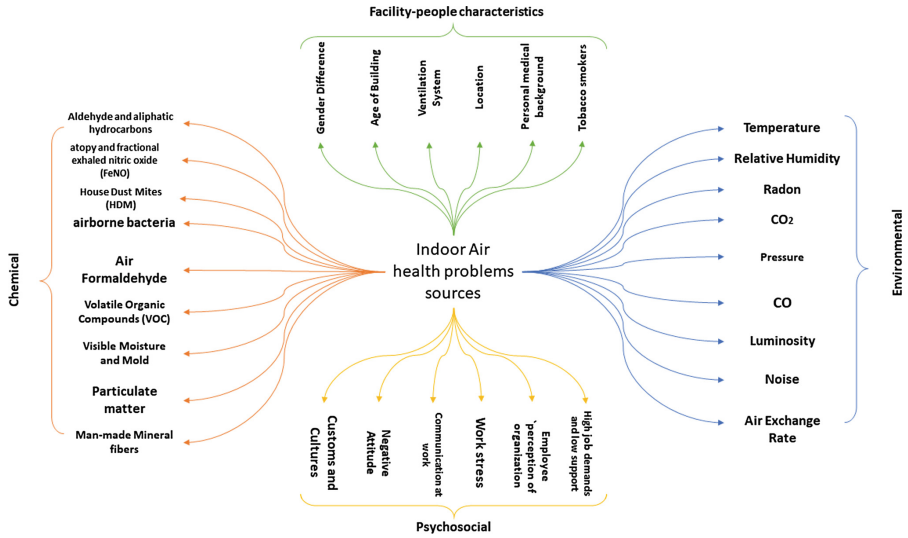
Fig. 1. Categorization of indoor health problem sources.

humidity should be regularized somehow until the uncomfortable feelings are resolved. On the contrary, the relation between building-related symptom and associated parameters needs more complicated processes to achieve the user preference knowledge. Next section explains how we aim to formulate this problem, and the initial solution is introduced. However, in this section, the association type between building-related symptoms and parameters measurable in research laboratories are discussed. These associations provide the basis to design the prediction model, building on the existing knowledge in the literature, as shown in Table 1.

Table 1. Health-related symptoms and associated (type) measurable parameters

Measurable parameter	Associated symptoms (positive+/negative−/not clear)	Reference(s)
Temperature	Nasal and skin symptoms(−) Fatigue(+) Sleepiness(+)	[11–13]
Humidity	Skin and nasal dryness(−) Eye irritation(−)	[12, 14, 15]
Lighting	Eyestrain Fatigue(−) Headache(+) Dizziness(−)	[16, 17]
CO2	Skin and eye irritation(+) Tiredness(+) Difficulty in concentration(−)	[13, 18]

(*continued*)

<div align="center">Table 1. (<i>continued</i>)</div>

Measurable parameter	Associated symptoms (positive+/negative−/not clear)	Reference(s)
TVOC	Sensory irritation(+) Asthma symptoms(+)	[18–20]
NOx	Respiratory symptoms(+) Asthma(+)	[21, 22]
Radon	Lung cancer(+)	[23]
Noise level	Headache(+) Fatigue(+) Hypertension Aggression(+)	[24, 25]
Stress level	Anxiety(+) Fatigue(+) Eye irritation(+)	[26, 27]
Posture	Connective tissues of tendons Joint capsules and ligaments	[28, 29]

3 Personalization Framework

The goal of this conceptual framework is to predict a measure of comfort preference for each individual worker based on personalized feedback and literature knowledge on known health problems and related associations. There are two key points that make this problem challenging. First, the heterogeneous nature of preference makes the prediction problem ungeneralizable to all individuals. In other words, one worker's feedback cannot directly be employed to improve the prediction for another worker. This is an important challenge as it is only possible to collect limited feedback from each worker, compared to the impersonalized case where a huge number of feedback, from the sum of all workers, is available. Second, the knowledge available in the literature, i.e. health problems and associations such as those gathered in Table 1, is sparse and unorganized, and it would be difficult and cumbersome to formulate it in the prediction problem. In the following subsections, we formally introduce and review these challenges and available solutions from the machine learning point of view. We then propose how to employ the state of the art prediction models for the preference personalization and we explicate it with an example in Sect. 4.

3.1 Machine Learning Challenge

In many practical prediction applications, the input data with known target values, i.e., the training data, is significantly fewer than the number of attributes representing the data. In some cases, the number of data can be more than the attributes, but still, the training data may only cover some particular aspects of the search space, which would be equivalent to only having few effective training data. This can, in particular, happen in personalized systems where the feedback is limited or only few sets of configurations can be tried out. The limited effective training data poses constraints about how

accurate the predictions can be [30]. Furthermore, many powerful machine learning methods, like deep neural networks, cannot be applied in this setup since they require huge amount of data. The dominant solution for these problems is to regularize the prediction model to constrain the search space and to avoid overfitting to training data.

A parallel fruitful direction to improve these types of prediction problems is to employ alternative available sources of information, other than training data, in the prediction. In many problems, prior information about the prediction task is available through experts or relevant literature. Prior elicitation is the process of extracting the available knowledge and employing them in the prediction task [31]. This is usually done by having a data scientist or statistician interviewing the field experts and then enforcing this knowledge on the parameters of the prediction model [31]. However, the classical prior elicitation methods are expensive and would require many iterations between the experts in the field and experts in the modeling. Recent works have proposed prior elicitation methods that remove the link between the data scientist and the field expert and directly put the expert in the prediction loop. This has become possible by defining intuitive ways for the experts to input their knowledge (priors) about the problem in a Bayesian prediction model. For example, [32] ask the experts to provide information about whether an attribute is relevant in a highly regularized prediction task or not, or to provide a value as their estimate of the regression coefficient. [33] ask about pairwise similarity feedback on different attributes, [34] about the direction of relevance (positive or negative), and [35] about the probability of an attribute being relevant. All these methods also investigate applications where there are only few training data sets available. We believe these approaches can also be employed in our personalized comfort prediction problem because different associations between attributes and the target variable (see Table 1) is known in our task.

In particular, [32] proposes a Bayesian sparse linear regression as the modeling solution to handle the limited training data problem. In their model, it is possible to intuitively add external knowledge about the relevance of attributes. We believe that this would be a proper fit for our personalized preference prediction problem. To use it, first, we need to gather the necessary data including the vector of attributes (here for example sensory measurements about the work environment and personal information like age and sex or any other related information) and the corresponding target values (for example personalized feedback about symptoms). Indeed, it is impossible to bother the user for a large number of feedbacks, and therefore, the number of data is always small. This is the ideal case for that method since in nature it assumes that the number of data is even less than the number of attributes.

This model is able to handle limited feedback challenge. However, we would still need to add the literature knowledge (Table 1) into the problem at hand. To do this, we can use the modeling solution in [34] (an extension of [32]) and consider the following types of literature knowledge:

- Knowledge about the relevance of the attribute for the considered prediction task.
- Knowledge about the direction of association (positive or negative) of the attributes to the target variable.

Given these two types of knowledge for the available attributes and the limited available data from individual feedback, [34] showed that it is possible to improve the accuracy of the prediction. Section 4 demonstrates how this approach can be used in an example scenario.

3.2   BIM Integration

The main aim of this study is to introduce a conceptual solution for personalization of the indoor health parameters based on the worker's preference. The building residents are considered as one of the building consituents in the context of facility management. Hence, the resident behavior can be modelled in BIM. Consequently, worker's preferences can be modeled and analyzed in building an information system for more efficient knowledge creation. In this respect, the required technical infrastructure should be provided for heterogeneous data collection and aggregation, data analysis and visualization, and knowledge creation. In addition to knowledge about the people preferences, the proposed framework is able to create the knowledge related to facility management utilizing the collected sensory data to make informed decisions about the health risks of the occupants. Through Internet of Things (IoT) communication, an infrastructure is developed where the sensor data and user feedback can be collected, and the generated knowledge can be applied for system behavior improvement [36]. BIM can play a vital role in automation and visualization of information for the built facility, and consequently, closing the information gap to provide comfort for building occupants, as shown in Fig. 2. From technical point of view, a service oriented architecture is developed, which each service oversees a module e.g. data aggregator and visualizer, in order to fulfil the technical requirements of the proposed framework. In data acquisition level, different streams of data are collected in related databases. In prototype level, the NoSQL database such as MongoDB[1] is preferred based on the volume and complexity of data. Data aggregator module will generate a timewise synchronized matrix out of data views in database. The machine learning service analyses the generated matrix which contains user feedback on collected psychological and environmental data to predict personalized model for each user. Trimble Connect[2] BIM engine is utilized to provide a knowledge base and data visualizer which can be used by the stakeholders to exploit of generated model and knowledge.

The proposed technical framework is able to aid facility managers in making decisions about building facilities based on the on knowledge created within the data analysis process e.g. the VOC data analysis can assist facility manager to identify mold growth in a building. In addition, the framework is able to provide the worker with a real-time warning about the inappropriate situation at work e.g. inconvenient posture of sitting.

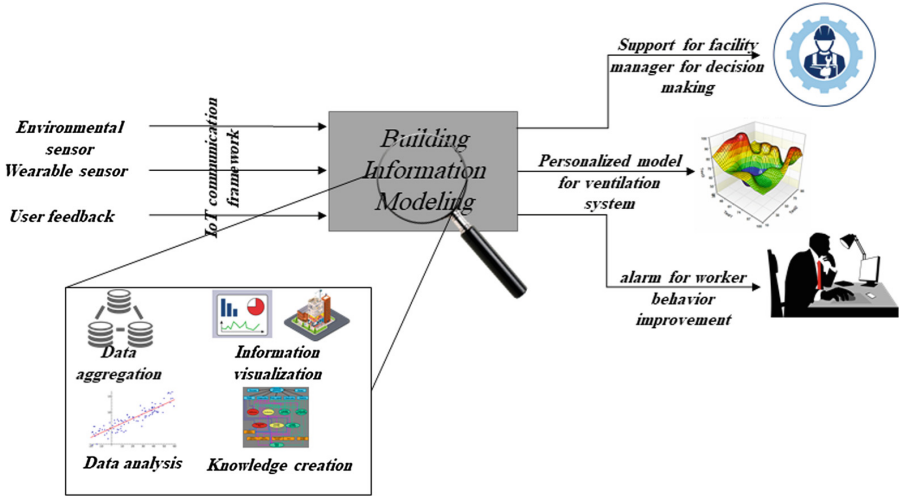[1] https://www.mongodb.com/.

[2] https://connect.trimble.com/.

Fig. 2. The high level architecture of information flow

4   Use Case Scenario

In order to better understand and validate the proposed solution, a potential example scenario is provided in this section. This example describes how the algorithm works for personalization of all measurable parameters with respect to the worker feedback. Assume that we want to have a prediction system that would estimate the level of fatigue, i.e. subjective feeling of tiredness, for each individual worker in the building. The first step is to consider a list of attributes that may contribute to fatigue. These attributes can include both the sensory measurements about the environment (such as outdoor and office temperature, humidity, CO2 level, TVOC, NOx, Radon level and other measurable parameters listed in Table 1, and time and weather attributes such as time of the day, season, weather condition, and so on), and the personal level attributes (such as age, gender, health status, allergic disease, respiratory illnesses, smoking status). In the training phase, we ask about personal feedback of individual workers on their level of fatigue at different intervals. The feedback can be provided in a discretized preference scale such as Likert-type scale (1–5). Each feedback value along with the values of the corresponding attributes create one instance of the training data. Along with the gathered training data, we include the known associations between the attributes and the target variable (fatigue) in the Bayesian model of [34]. For example, we add the directional knowledge about stress level, CO2, lighting, and temperature to the model (e.g., it is known that CO2 has a positive association with fatigue). The model uses both the training data and the knowledge to learn the best prediction solution. After the training, the model can be used to estimate the level of fatigue of that worker at any time. This knowledge can be employed to find the most pleasant attribute ranges for that worker.

Figure 3 depicts this scenario where the worker provides feedback about his or her fatigue level in 3 intervals. These few data along with the association knowledge in the literature will feed the machine learning algorithm to learn the personalized fatigue prediction model.
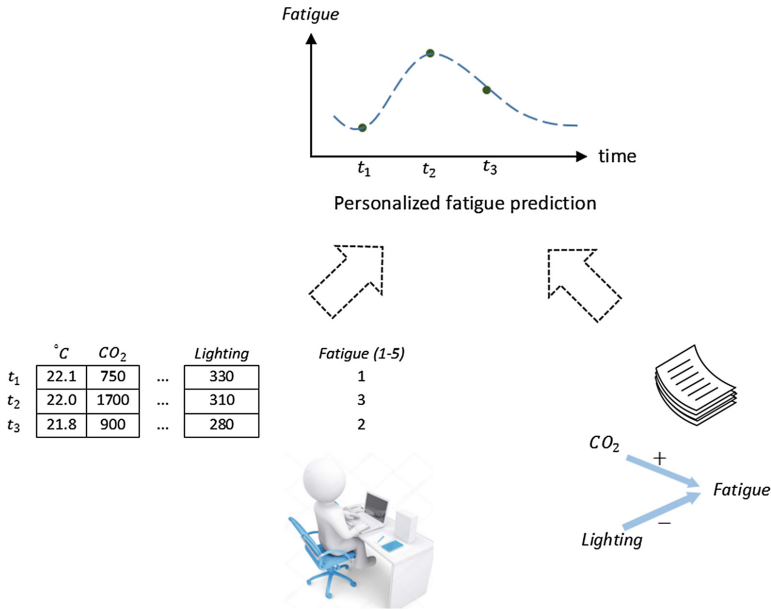

Fig. 3. Combining limited feedback from an individual worker with the available literature knowledge to achieve an accurate personalized prediction model.

5   Conclusion

Personalization in different domains is growing as people are looking for more customized solutions. Based on this preference, the expectations for the personalized indoor environment is not an exception either. In this respect, the purpose of this paper is to provide a technical framework to predict the measure of comfort preference for each individual worker based on personalized feedback and literature knowledge on known health problems and associations. The machine learning challenges are discussed in order to highlight the importance of the study. Subsequently, literature has been investigated to identify the prior knowledge about prediction models, and an example scenario is proposed to support the feasibility of the personalization model. A BIM-integrated technical solution is discussed for automating data aggregation and closing the information flow gap in built facilities.

Indoor health and well-being related parameters can affect employees' productivity during work time. So, it would be beneficial for employers to have healthier work environments to provide greater comfort for each individual worker, because different

people may fall sick in different conditions based on their physiological attributes. Personalized indoor environmental conditions can potentially enhance productivity and reduce sick leaves. The objective of the proposed framework is to overcome the aforementioned challenges and the developed system according to this framework will be evaluated with real data in further studies.

References

1. Brasche, S., Bischof, W.: Daily time spent indoors in German homes–baseline data for the assessment of indoor exposure of German occupants. Int. J. Hyg. Environ. Health **208**(4), 247–253 (2005). https://doi.org/10.1016/j.ijheh.2005.03.003
2. Leech, J.A., Nelson, W.C., Burnett, R.T., Aaron, S., Raizenne, M.E.: It's about time: a comparison of Canadian and American time–activity patterns. J. Expo. Anal. & Environ. Epidemiol. **12**(6), 427 (2002). https://doi.org/10.1038/sj.jea.7500244
3. World Health Organization: Indoor air pollutants exposure and health effects report on a WHO meeting, Nordlingen, EURO Reports and Studies, 78 (1982)
4. Sterling, E., Sterling, T., McIntyre, D.: New health hazards in sealed buildings. Am. Inst. Archit. J. (United States) **72**(4) (1983)
5. Reijula, K., Sisäilmaongelmat, H.T.: Altistumisen arviointi ja potilaan tutkiminen [indoor air problems: Assessment of exposure and examination of the patient]. Finn. Med. J. **53**(36), 4215–4230 (1998)
6. Ghazanfari, E.: Applying product design and digital construction methodologies to conceptualize modular and distributed healthcare facilities. Master thesis, Aalto University, Finland (2016)
7. Gillespie I.: Modular Office (2002)
8. Jansz, J.: Theories and Knowledge About Sick Building Syndrome. In: Abdul-Wahab, S. (ed.) Sick Building Syndrome. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-17919-8_2
9. Hellgren, U.: Indoor air problems in Finnish hospitals-from the occupational health perspective (2012)
10. Norbäck, D.: An update on sick building syndrome. Curr. Opin. Allergy Clin. Immunol. **9** (1), 55–59 (2009). https://doi.org/10.1097/ACI.0b013e32831f8f08
11. Amin, N.D.M., Akasah, Z.A., Razzaly, W.: Architectural evaluation of thermal comfort: sick building syndrome symptoms in engineering education laboratories. Procedia - Soc. Behav. Sci. **204**, 19–28 (2015). https://doi.org/10.1016/j.sbspro.2015.08.105
12. Reinikainen, L.M., Jaakkola, J.: Significance of humidity and temperature on skin and upper airway symptoms. Indoor Air **13**(4), 344–352 (2003). https://doi.org/10.1111/j.1600-0668.2003.00155.x
13. Schiavon, S., Yang, B., Donner, Y., Chang, V., Nazaroff, W.W.: Thermal comfort, perceived air quality, and cognitive performance when personally controlled air movement is used by tropically acclimatized persons. Indoor Air **27**(3), 690–702 (2017). https://doi.org/10.1111/ina.12352
14. Shan, X., Zhou, J., Chang, V.W., Yang, E.: Comparing mixing and displacement ventilation in tutorial rooms: students' thermal comfort, sick building syndromes, and short-term performance. Build. Env. **102**, 128–137 (2016). https://doi.org/10.1016/j.buildenv.2016.03.025
15. Wolkoff, P., Nojgaard, J.K., Franck, C., Skov, P.: The modern office environment desiccates the eyes? Indoor Air **16**(4), 258–265 (2006). https://doi.org/10.1111/j.1600-0668.2006.00429.x

16. Godish, T.: Sick Buildings: Definition, Diagnosis and Mitigation. CRC Press, Boca Raton (1994)
17. Pitarma, R., Marques, G., Ferreira, B.R.: Monitoring indoor air quality for enhanced occupational health. J. Med. Syst. **41**(2), 23 (2017). https://doi.org/10.1007/s10916-016-0667-2
18. Lu, C., Lin, J., Chen, Y., Chen, Y.: Building-related symptoms among office employees associated with indoor carbon dioxide and total volatile organic compounds. Int. J. Environ. Res. Public Health **12**(6), 5833–5845 (2015). https://doi.org/10.3390/ijerph120605833
19. Dales, R., Raizenne, M.: Residential exposure to volatile organic compounds and asthma. J. Asthma **41**(3), 259–270 (2004). https://doi.org/10.1081/JAS-120026082
20. Venn, A.J., Cooper, M., Antoniak, M., Laughlin, C., Britton, J., Lewis, S.A.: Effects of volatile organic compounds, damp, and other env'l exposures in the home on wheezing illness in children. Thorax **58**(11), 955–960 (2003). https://doi.org/10.1136/thorax.58.11.955
21. Lim, F., Hashim, Z., Said, S.M., Than, L.T., Hashim, J.H., Norbäck, D.: Sick building syndrome (SBS) among office workers in a Malaysian university—Associations with atopy, fractional exhaled nitric oxide (FeNO) and the office environment. Sci. Total Environ. **536**, 353–361 (2015). https://doi.org/10.1016/j.scitotenv.2015.06.137
22. Mi, Y., Norback, D., Tao, J., Mi, Y., Ferm, M.: Current asthma and respiratory symptoms among pupils in Shanghai, China: influence of building ventilation, nitrogen dioxide, ozone, and formaldehyde in classrooms. Indoor Air **16**(6), 454–464 (2006). https://doi.org/10.1111/j.1600-0668.2006.00439.x
23. Joshi, S.M.: The sick building syndrome. Indian J. Occup. Environ. Med. **12**(2), 61 (2008). https://doi.org/10.4103/0019-5278.43262
24. Skov, P., Valbjorn, O.: The "sick" building syndrome in the office environment: the Danish town hall study. Env. Int. **13**(4–5), 339–349 (1987). https://doi.org/10.1016/0160-4120(87)90190-5
25. Teeuw, K.B., Vandenbroucke-Grauls, C.M., Verhoef, J.: Airborne gram-negative bacteria and endotoxin in sick building syndrome: a study in Dutch governmental office buildings. Arch. Intern. Med. **154**(20), 2339–2345 (1994). https://doi.org/10.1001/archinte.1994.00420200095010
26. Marmot, A.F., Eley, J., Stafford, M., Stansfeld, S.A., Warwick, E., Marmot, M.G.: Building health: an epidemiological study of "sick building syndrome" in the whitehall II study. Occup. Environ. Med. **63**(4), 283–289 (2006). https://doi.org/10.1136/oem.2005.022889
27. Nordström, K., Norbäck, D., Akselsson, R.: Influence of indoor air quality and personal factors on the sick building syndrome (SBS) in Swedish geriatric hospitals. Occup. Environ. Med. **52**(3), 170–176 (1995). https://doi.org/10.1136/oem.52.3.170
28. Grandjean, E., Hünting, W.: Ergonomics of posture-review of various problems of standing and sitting posture. Appl. Ergon. **8**(3), 135–140 (1977). https://doi.org/10.1016/0003-6870(77)90002-3
29. Leivseth, G., Drerup, B.: Spinal shrinkage during work in a sitting posture compared to work in a standing posture. Clin. Biomech. **12**(7–8), 409–418 (1997). https://doi.org/10.1016/S0268-0033(97)00046-6
30. Johnstone, I.M., Titterington D.M.: Statistical Challenges of High-Dimensional Data (2009). https://doi.org/10.1098/rsta.2009.0159
31. O'Hagan, A., et al.: Uncertain Judgements: Eliciting Experts' Probabilities. Wiley, Hoboken (2006)
32. Daee, P., Peltola, T., Soare, M., Kaski, S.: Knowledge elicitation via sequential probabilistic inference for high-dimensional prediction. Mach. Learn. **106**(9–10), 1599–1620 (2017). https://doi.org/10.1007/s10994-017-5651-7

33. Afrabandpey, H., Peltola, T., Kaski, S.: Interactive prior elicitation of feature similarities for small sample size prediction. In: Proceedings of 25th Conference on User Modeling, Adaptation and Personalization, pp. 265–269 (2017). https://doi.org/10.1145/3079628.3079698

34. Sundin, I., et al.: Improving drug sensitivity predictions in precision medicine through active expert knowledge elicitation (2017). arXiv Preprint arXiv:1705.03290

35. Daee, P., Peltola, T., Vehtari, A., Kaski, S.: User modelling for avoiding overfitting in interactive knowledge elicitation for prediction. In: Proceedings of the 23rd ACM International Conference on Intelligent User Interfaces, pp. 305–310 (2018). https://doi.org/10.1145/3172944.3172989

36. Mirzaeifar, S., Dave, B., Singh, V.: Development of systematic construction logistics using 'intelligent products' (2017). https://doi.org/10.24928/2017/0295