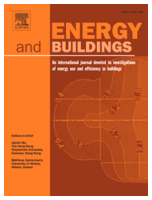
Energy & Buildings 370 (2026) 118052

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enbuild



An AHP experiment with the smart readiness indicator applied in a local reality

Emanuele Salerno

National Research Council of Italy – ISTI, Via Giuseppe Moruzzi, 1, Pisa, 56124, Italy

ARTICLE INFO

Keywords:
Sustainable buildings
Technologies for the quality of life
Smart readiness indicator
Analysis of the hierarchical process

ABSTRACT

This paper describes an experiment intended to evaluate the appropriateness of the weighting scheme used to compute the Smart Readiness Indicator (SRI), an index introduced by the European Commission and Parliament with the double aim to measure the "intelligence" of a building and to encourage the adoption of smart technologies in both new buildings and existing constructions to be retrofitted. This index is based on a hierarchical scheme including criteria related to energy savings and flexibility as well as to the comfort, convenience and wellbeing of the building occupants. In particular, the weights to assign to the criteria related to the needs of the occupants are largely subjective, and the technical support group established by the European Commission to draw up the SRI scheme did not find any scientific evidence to differentiate them, so all of them received equal weights. Exploiting Saaty's Analysis of the Hierarchical Process (AHP), we involved a local group of engineers, architects, technicians and administrators to obtain cardinal rankings for these criteria. The results are likely to depend on the local context where our inquiry has been conducted, central Italy, but also show that the opinions of our experts are diverse, even within professionally homogeneous subgroups, and do not justify the equal-weight choice made by the technical support group. Our results can also constitute one of the many expected local contributions to the revision of the present SRI scheme.

1. Introduction

By Eurostat [1], a large part of the energy consumption in Europe derives from buildings. However, an 85% of European buildings dates back to before 2000 and a 75% has poor energy performance. Thus, improving the energy performance of existing buildings is tantamount to cut significantly both energy consumption and greenhouse gasses emissions, thus enabling the goal of a fully decarbonised building stock by 2050 to be achieved, as established by the EU legislation [2,3], and contributing to the 13^{th} sustainable development goal established by the UN: Climate Action [4]. Yet, the annual energy renovation rate remains very low, at 1%. Empirical studies confirm that digital technologies contribute to reducing operational emissions in buildings, particularly by optimizing energy use [5].

The last revision of the Energy Performance of Buildings Directive (EPBD) in 2024 [3] reaffirms the smart readiness indicator (SRI) as an effective index of the capability of the building automation and control systems (BACS) to enable an intelligent management of the building operation, thus optimizing energy performance and flexibility as well as the quality of life of the occupants. This index was established by the 2018 revision of the EPBD [6] to both measure the "intelligence" of buildings and possibly guide wise technology additions to existing buildings in order to reach better energy efficiency and quality of life of their occupants. The intelligence of a building is defined as "the ability of a building or its systems to sense, interpret, communicate and actively respond in an efficient manner to changing conditions in relation to the operation of technical building systems or the external environment (including energy grids) and to demands from building occupants" [7]. Many studies focus the different aspects related to the evaluation and the implementation of the SRI; see for example [5,8–17]. The implementation of the SRI has been estimated to potentially reduce the total energy consumption by up to 198 TWh by 2050, consequently averting approximately 32 million tons of greenhouse gas emissions annually. As a measure of the capability of a building to usefully exploit information and communication technologies, the SRI would be an index of how buildings can adjust their operations to improve the overall building performance [5].

By the EC delegated and implementing regulations defining the structure and implementation of the SRI in EU member states [18,19], the evaluation of the SRI for any single building (or building unit) is assigned to a certified assessor, who issues an SRI certificate. The evaluation can also be done independently by owners or administrators, but with no right to get a certificate. The SRI computing scheme provides two assessment methods. One of them (Method A), to be used for residential or small non-residential buildings, is checklist-based and relies on a reduced catalog of smart services that can be offered by the

E-mail address: emanuele.salerno@cnr.it

https://doi.org/10.1016/j.enbuild.2026.118052
Received 24 March 2026; Received in revised form 31 July 2026; Accepted 1 August 2026
Available online 4 August 2026
0378-7788/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

building. The other method (B) needs a dedicated survey by the assessor and includes an extended catalog of smart services. This is used for large non-residential or particularly complex buildings. A further method (called Method C) has been envisaged, not yet standardized, providing specific and protracted measurements on the technological building systems (TBS).

Several studies (see [5,20] and the specific references therein) focus the inherent subjectivity of the SRI, which is our main concern here. Mainly in view of the energy aspects measured by the SRI, it has been observed that the checklist-based assessment procedure lacks precise guidelines intended to limit an excessive arbitrariness of the assessor, who could adapt the triage process preceding the evaluation of the technological levels of the different smart services, or even "manipulate" the scores to obtain favorable results [21]. Other critics complain about the lack of more quantitative indices to evaluate the actual performance of the smart services, possibly based on field measurements rather than the simple technological levels provided by the standard scheme. The latter requirement could be satisfied by the above-mentioned Method C. A similar, more general observation has also emerged during the tests conducted within the Italian project GUARDIANS, which supports partially this study: comparing the information provided by the energy performance certificate (EPC) and SRI, it is apparent that the values drawn from the EPC are all measurable and absolute, and can also be used to evaluate the cost of retrofit interventions in view of energy efficiency. This is not the case with SRI, whose information is relative, given as a percentage of the maximum achievable by the specific building to be assessed, and not always measurable. Nonetheless, studies on costs and priorities of retrofitting in view of an improvement of the SRI are present in the literature [9,16]. As far as a limitation of the assessor's arbitrariness is concerned, in our opinion, also considering the different "robust and flexible" implementation pathways listed in [7, Summary p. 9], the prescription of which services and technological domains are to be included in or excluded from the evaluation (the so-called triage process), should be left to the implementing member states depending on geographic and building specificities, that is, to the governance of the assessment system. Once this is done, the assessor must only evaluate the technological levels of all the relevant smart services; these are defined accurately in the SRI calculation sheet, so small room is left to subjectivity on their part. In any case, the resulting certification can be inconsistent in the sense that different buildings, or even similar buildings in different locations, cannot be compared in terms of SRI. The technical support group established by the European Commission to define the SRI evaluation criteria also recognized this inconvenience [7, Section 1.3.5], renouncing to comparability in favor of "transparency" (see also [22]); in other words, two buildings cannot be compared on the basis of their SRIs, but each one provides a certificate where all the choices made by the assessor and all the factors affecting the final score are open and apparent. On the other hand, the SRI is not a direct index of energy performance or occupants' quality of life, but of how the BACS can contribute to improve energy performance, energy flexibility, and needs of the occupants. A better comparability can be obtained by associating the SRI certificate with the EPC, also established by the EPBD [3, Annex V]. At present, this possibility is under study, and several models of joint SRI-EPC certificates are being proposed [7, Section 1.2.6].

Where true subjectivity can be found is in the relative importances (weights) assigned to the different criteria used to reach the final SRI value. Some of those weights are tailored to the specific case, considering factors such as the building type and climate zone, or even energy balances taken from the EPC of the particular building being assessed. Other weights are either fixed to agreed-upon values or set at equal values, maintaining that any different weighting is not supported by scientific evidence. Among the latter, the weights to assign to the criteria "Comfort" and "Convenience" are the most genuinely subjective: by Verbeke et al. [7], Comfort refers to "conscious and unconscious" perception of the physical environment, and Convenience refers to making the life easier for the occupant. This apparently varies with any individual occupant and is as such an unsuppressible source of subjectivity. On the other hand, all the criteria related to the response to the needs of occupants have been included by the EPBD, conscious of their subjectivity since its earliest versions, in order to extend the concept of building performance to aspects not related to energy efficiency. As part of the GUARDIANS work program, this paper presents the results of an experiment conducted involving various local representatives of engineers, architects, technicians and administrators to assess the appropriateness of those equal-weight factors within the geographical and social reality of the central Italian coastal region of Tuscany. Even considering the large diversity of the responses received from our stakeholders, the results obtained show that the equal-weight strategy followed by the technical support group is not tenable, and probably the most subjective weights should also be made dependent on local specificities.

Our experimental methodology followed the so-called Analysis of the Hierarchical Process, or AHP [23]. When a target decision is based on a hierarchical structure made of different levels, criteria and subcriteria, it can be difficult to decide a precise cardinal ranking among them, especially when their number is large. The mentioned solution is based on pairwise comparisons between different criteria affecting a target value, which finally bring to the desired ranking, that is, each criterion is provided with a specific weight in view of the target. A number of experts (hereafter, "judges") are requested to perform the necessary comparisons, then their judgements are combined to obtain the final rankings. This procedure has some advantages over the ones based on group decision-making, in that: 1) mathematically, the consistency of the individual judgements can be checked, the contradictory judgements can be located and possibly corrected or revised with the collaboration of the judge involved; 2) possible biases due to the opinions of particularly authoritative judges can be reduced or avoided. Of course, before starting the procedure, each judge must receive adequate information about the problem at hand and the decision to be made.

This paper is structured as follows. In Section 2, we give essential information on the SRI. Section 3 describes briefly the AHP procedure to obtain cardinal rankings from the pairwise comparisons and to assess their internal consistency. Section 4 reports our experimental procedure and its results. Finally, Section 5 concludes the paper.

2. Smart readiness indicator

This section only gives basic information about the structure and the evaluation method of the SRI. All the details can be found in [7]. The first three levels of the SRI hierarchy are shown in Fig. 1, with the weighting factors assigned by the technical support group to the different criteria. On top of the hierarchy, there is the SRI, expressed, as described below, as a percentage of the maximum value reachable by the building being assessed. The decision criteria at the second level are called "key functionalities", and are 1) Energy Savings and Operation, 2) Response to the Needs of Occupants, and 3) Response to the Needs of the Grid. Each functionality affects one third of the SRI value. The key functionalities are further specified by the third level of the hierarchy, whose elements are called "impact criteria". Functionality 1) subtends two criteria: i) Energy Savings, and ii) Maintenance and Fault Prediction, each one affecting a half of the value of the subtending functionality. Functionality 2) subtends four criteria: i) Comfort, ii) Convenience, iii) Information to Occupants, and iv) Health and Wellbeing, each weighting one fourth of the value of functionality 2). Finally, functionality 3) only subtends one criterion, whose score covers the entire functionality: Energy Flexibility and Storage. As apparent, the technical support group has assigned equal weight to the elements of all the groups forming this hierarchy. Taking the impact criteria into account, with respect to the SRI, we see that the two criteria Energy Savings and Maintenance and Fault Prediction weigh $1/6$, the four criteria under Response to the needs of occupants weigh $1/12$ and the only criterion under Response to the needs of the grid affects $1/3$ of the final score.

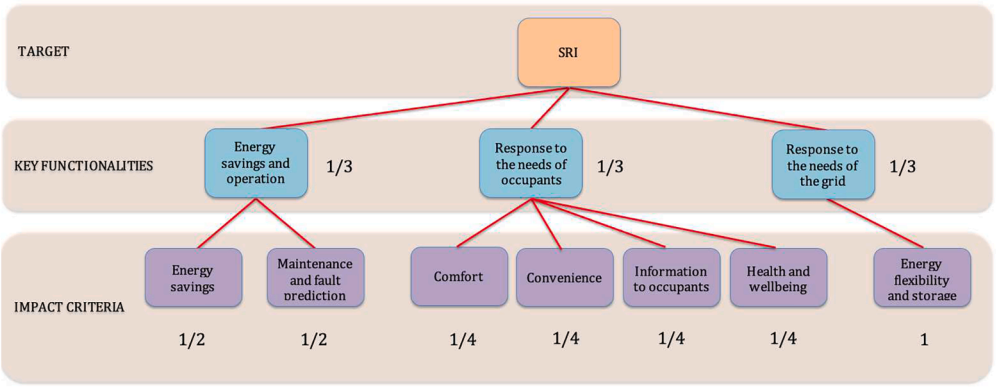


Fig. 1. The first three levels of the SRI hierarchical structure: Target, Key functionalities, and Impact criteria.

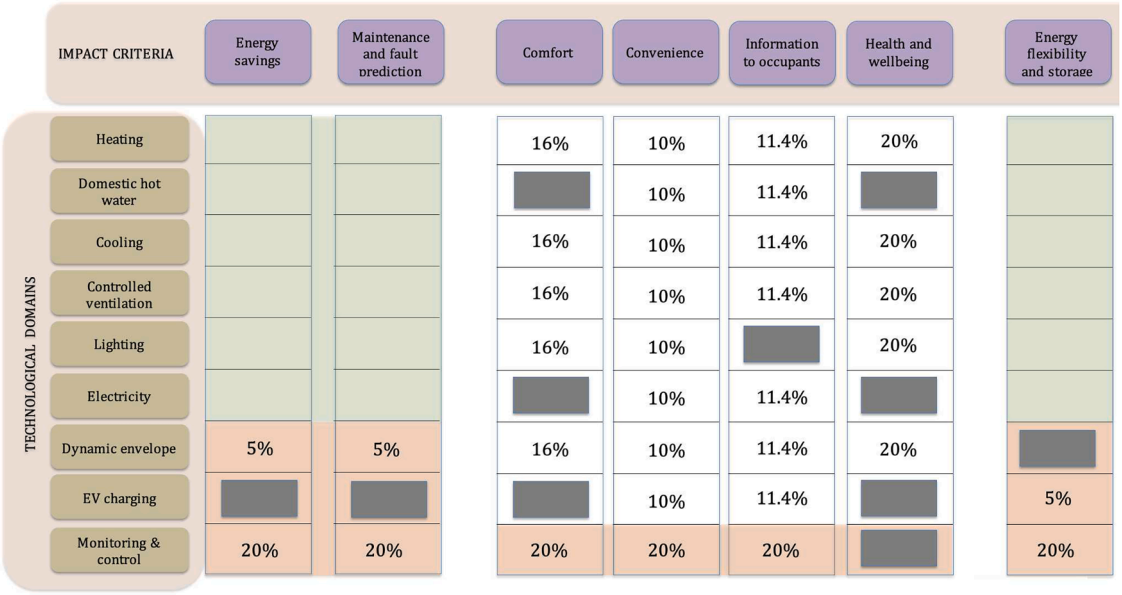


Fig. 2. The fourth level of the SRI hierarchical structure: Technological domains. Green background cells: energy-dependent weights; Orange background cells: fixed weights; White background cells: equal weights.

The BACS, including all the smart services available in the building, constitute the fourth SRI hierarchical level and are classified through nine technological domains, which affect the impact criteria in the preceding level. The technological domains established by the support group are

a) Heating
b) Domestic Hot Water
c) Cooling
d) Controlled Ventilation
e) Lighting
f) Electricity
g) Dynamic Envelope
h) EV Charging
i) Monitoring & Control

Fig. 2 shows this level of the hierarchy, with the technological domains and the associated weights placed under the corresponding impact criteria. Not all the impact criteria are affected by all technological domains: the cells marked in gray indicate that the corresponding domains do not affect the subtending impact criteria.

The colors of the cell backgrounds encode the choices made by the support group to assign an appropriate weight to each technological domain with respect to the subtending impact criterion. The green cells have been assigned weights depending on energy balances. These are based by default on the climate zone where the building is located (from the five defined by the support group, each including a subset of the EU member states), but the implementing member states can choose to either define more detailed climate zones or to let the weights be computed from energy balances characteristic of the particular building being assessed (these can be taken from the EPC). The orange cells have been assigned the fixed weights shown in the figure. Finally, as apparent from the white-background cells, the domains under each impact criterion except the fixed-weight ones have received equal weights, considering that "no scientific evidence" is available to decide otherwise [7, p. 110].

The support group has provided a catalog of 54 relevant smart services, leaving the possibility to add further ones. Each technological domain includes a subset of services from this catalog. As anticipated, the evaluation method A relies on a restricted catalog, only including 27 services. The assessor assigns to each relevant smart service one of a small number of well-defined functional levels (maximum 5, from 0 to 4). The SRI machinery then assigns a score, depending on the affected impact criterion, to each such level;[1] the scores related to the services

[1] The mapping of the functional levels into the related scores is another, well-motivated "weighting" strategy implemented by the technical support group;

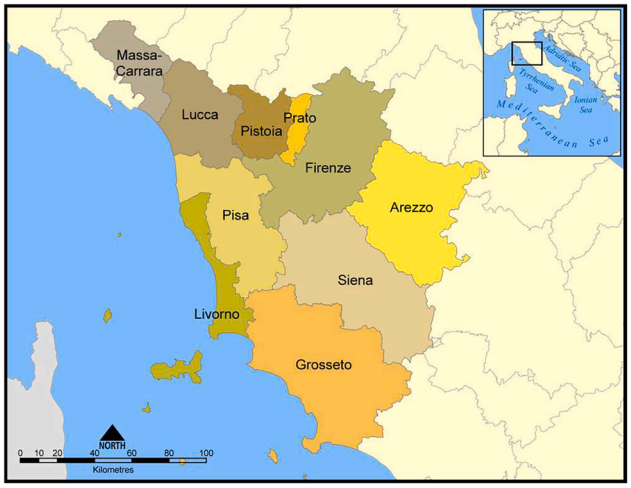


Fig. 3. The stakeholders we consulted operate in the Italian provinces of Lucca, Pisa, Livorno and Grosseto.

pertaining to each technological domain are summed up and assigned to that domain as a percentage of the maximum score obtainable by the building. Then, the scores obtained by the technological domains are weighted and summed up to form the scores of the impact criteria, and further weighted and summed up to form the key functionality scores and, finally, the SRI.

The choice of the service catalogs, the standard weighting scheme, and even the very SRI calculation method have been variously criticized in recent years (see, e.g. [5,13]). What is of interest here is the choice of equal weights shown in Fig. 1 for the key functionalities and impact criteria, and in the white cells of Fig. 2 for the technological domains. In this respect, the weaknesses of the current scheme have also been recognized by the technical support group. In synthesis, the method agreed upon by those stakeholders is based on the "predicted impact", that is, the effect of the weighted score of any service or domain on the final SRI. For non-quantifiable criteria, the study team was not able to find any objective data source that could have justified a scheme different from the equal-weight one. The resulting proposal by the study group, structured as shown in the figures above, was "perceived" as supported by most stakeholders, but no full consensus was reached.

Our purpose is to check, independently of a supposed scientific evidence, whether the equal-weight strategy is really appropriate, at least for our local reality of the Tuscan coastal area, in central Italy. To this end, we asked for the contributions of a number of stakeholders, chosen among professionals related to the building industry, that is, engineers, architects, designers, builders, technological systems installers and maintainers, surveyors, real estate brokers, building managers, and municipality technical officers. All of them operate in coastal Tuscany, in the provinces of Lucca, Pisa, Livorno and Grosseto (see map[2] in Fig. 3).

To gather and synthesize their opinions, we adopted the formal multicriteria decision method called AHP, described briefly in Section 3, which enables decision makers to evaluate the internal consistency of the opinions received, evaluate the agreement between different contributors, and merge the individual results into final rankings. Our results, along with other contributions from different groups and local realities, could contribute to form part of the "objective data source" sought by the technical support group to substantiate the weights whose effect on the result is not measurable.

here, we do not enter so deep into the SRI machinery. For details, refer to[7, Summary pp. 19ff and Sec. 1.3.6].

[2] https://wiki.openstreetmap.org/wiki/File:Provinces_of_Tuscany_map.png. Creative Commons Attribution-ShareAlike 2.0 license

Table 1
9-point judgement scale for pairwise comparisons (criterion A vs. criterion B).

Is Criterion A equally or more important than Criterion B?	
Intensity of importance	Explanation
1	Equal importance
3	Weakly more important
5	Strongly more important
7	Very strongly more important
9	Extremely more important
2 - 4 - 6 - 8	Intermediate values between two adjacent judgements

3. Methods: analysis of the hierarchical process

As seen above, the weights assigned to the criteria and technological domains subtended by the functionality Response to the needs of occupants are necessarily subjective: different evaluators can give different importances to the same criteria in view of their different interests or sensitivities. To us, the lack of scientific evidence is not sufficient to justify the equal-weight decision, so we decided to apply a multicriteria decision model (MCDM) to assess the validity of this choice with the help of selected stakeholders.

Producing a cardinal ranking among different decision criteria may not be easy, especially if their number is relatively large. In several decades, specific methods have been proposed to address this class of problems. The seminal works date back to the 1970s [24]. Since then, hundreds of works have been addressing the theory and possible applications of MCDMs. The basic strength of all these methods is the capability of treating problems characterized by mixed quantitative, qualitative, objective and subjective criteria, and also marked by different conflicting interests. The different criteria to be satisfied may also be in partial conflict to each other. As the formal structuring of of a multicriteria decision problem is the first and fundamental step towards its solution, much care has been given to define sound strategies to build effective models. Brugha [25] provides guidelines to build a criteria tree that leads to a usable decision model. Saaty [23] proposes the so-called Analysis of the Hierarchical Process (AHP) to derive appropriate rankings in a hierarchical decision structure . Karthikeyan et al. [26] point out that AHP is usable and reasonable, and disentangles a complex problem into simpler steps. They also mention weaknesses, drawbacks and criticisms, also comparing AHP with some successive modifications of it. For a structure that does not provide changes in criteria or in the evaluation scale, however, the original AHP version seems to be sufficiently robust. Note that in our case we do not want to propose alterations in the SRI hierarchy as visualized in Figs. 1 and 2, so the structure of our problem is fixed. As described briefly below, AHP satisfies the requirements listed for the MCDMs, can be implemented simply, does not require consensus meetings, and provides useful means to assess the reliability of the individual evaluations received. Moreover a rich bibliography allows us to analyze thoroughly all the aspects of the method, including sensitivity [27,28] and suitability for group decisions [29,30]. These features, along with the intuitive and reasonable nature of AHP, led us to choose it for our purposes.

AHP is based on pairwise comparisons between criteria belonging to groups at the same hierarchical levels. The judges are asked to compare the relative importance of all the pairs of criteria in each such group, specifying how much one criterion is more or less important than the other, with respect to the subtending criterion in the upper hierarchical level. The judge has a scale of 17 values, the ones reported in Table 1 and their reciprocals, to specify the relative importance of the two criteria in each pair. Normally, to specify how much one criterion is more important than the other, the judge uses the odd integers from 1 to 9; the even integers in the line at the bottom of the table can be used when a refined judgement between two of the above values is deemed necessary. The table shows the importance values that can be assigned when Criterion A is equally or more important than

Criterion B. If one of these values has been assigned to this comparison, when comparing Criterion B to Criterion A, it is logical to assign the reciprocal of that value [23]. When the first criterion being compared is less important than the other, the judge can thus use the reciprocal values 1/2, 1/3, 1/4, … 1/9 to express their preference.

To rank a group of criteria, the values assigned to the pairwise comparisons can be placed in a positive reciprocal "judgement" matrix. For example, referring to the three key functionalities in Fig. 1, we would obtain

$$A = \left[\begin{array}{c|ccc} & \text{Energy} & \text{Occupants} & \text{Grid} \\ \hline \text{Energy} & 1 & a_{12} & a_{13} \\ \text{Occupants} & 1/a_{12} & 1 & a_{23} \\ \text{Grid} & 1/a_{13} & 1/a_{23} & 1 \end{array}\right], \tag{1}$$

where each a_{ij} maps the comparison between criterion i and criterion j through one of the values listed in Table 1 or their reciprocals. The diagonal elements are all unity since each of them represents the comparison of a criterion to itself. To fill up all the entries of a reciprocal matrix of order N, the judge is required to make $N(N-1)/2$ comparisons. Specifying an equal-weight ranking in terms of the judgement matrix entails having all the elements a_{ij} in A equal to 1. A positive reciprocal matrix [31,32] offers a number of properties that help us to derive a cardinal ranking from pairwise comparisons and check the internal consistency of the judgements. A property of a positive reciprocal matrix of order N is the existence of a real eigenvalue $\lambda_{\max} \geq N$. Saaty [30] proves that the associated dominant eigenvector $\mathbf{v}$ represents a cardinal ranking among the criteria being evaluated. If $\mathbf{v}$ is scaled to sum up to 1, its elements express the percent importance of the associated criteria. The judge is required to produce judgement matrices for all the groups of criteria that form the hierarchy.

When the evaluation is made by several (say, k) judges, a collective result $\mathbf{v}^{(col)}$ is obtained from the geometric mean among the individual results $\mathbf{v}^{(i)}, i = 1, \ldots k$, that is [23,30]

$$\mathbf{v}^{(col)} = \left[\prod_{i=1}^{k} \odot \mathbf{v}^{(i)}\right]^{1/k}, \tag{2}$$

where the symbol $\odot$ means element-by-element product. Even though all the $\mathbf{v}^{(i)}$s are scaled to sum up to 1, this does not need to be true for $\mathbf{v}^{(col)}$, which must be rescaled to obtain percentage scores.

Human judgements are often subjective. In our case (see Figs. 1 and 2), it is apparent that some of the choices can be derived by measured quantities, while some others depend upon personal beliefs of relative importance of one or another criterion. In any case, a quota of subjectivity, and sometimes not perfect consistency, must be tolerated. As far as consistency is concerned, the eigenvalue $\lambda_{\max}$ of A in (1) helps us to check the judgements received, through the so-called consistency index

$$ci = \frac{\lambda_{max} - N}{N - 1}, \tag{3}$$

where, as said, N is the order of the judgement matrix. The value of ci tells us how the judgements are internally inconsistent, since it represents how far the dominant eigenvalue is from the value N of perfect consistency. Any positive reciprocal matrix of order N has its ci depending on size: the larger the matrix size, the higher, in average, its ci. A small consistency index, however, can also happen by chance. This possibility is corrected through the so-called random index, $ri(N)$, which is the average consistency index for positive reciprocal matrices of order N with randomly distributed entries. This defines the consistency ratio

$$cr = \frac{ci}{ri(N)}. \tag{4}$$

Table 2 shows the values of ri computed by this author [33] for $3 \leq N \leq 15$. The consistency index is always zero for $N = 2$.

Saaty [23] suggests a maximum allowed inconsistency of $cr = 0.10$, that is, a maximum of the 10% of the average ci for a positive reciprocal matrix of order N. The analysis in [27] shows that, when the primary purpose is to obtain a general understanding of the relative importance of the different criteria, consistency ratios larger than 0.1 may not distort the results. Although subjective, a judgement can be perfectly consistent. Different levels of inconsistency can be due to several causes, including an insufficient knowledge of the precautions to take when expressing the judgements, an insufficient knowledge of the decision problem at hand, or even the judge's fatigue. As specified in Section 4, all our judges have been briefed on SRI, AHP and the rules to bear in mind when making the pairwise comparisons. The value of cr lets us check the validity of the judgement provided, but other appropriate rules [34] have also been devised to locate the most inconsistent judgements and modify them slightly without altering significantly the final rankings. When this is not possible, new, consistent judgements should be negotiated directly with the judges. Details of a procedure to locate and identify the inconsistencies are also given in [30].

AHP has long been studied and applied by researchers in diverse fields, and some important strengths and weaknesses have been recognized [35]. First of all, the possibility of treating both objective and subjective alternatives, described in plain language, is a clear advantage for the judge, provided that they have been adequately informed on the decisions to be made. Mathematically, a very easy check of judgement consistency compares the elements of the priority vector $\mathbf{v}$ to the elements of matrix A. Logically, for a perfectly consistent matrix, the ratio between two elements of its dominant eigenvector, v_i/v_j should be equal to the matrix element a_{ij}. This is the first and easiest way to check the correctness of the different judgements. The possibility to aggregate different judgements is also an advantage, as this can avoid possible biases. Decisions based on consensus could be difficult to be obtained. If all the judgement matrices provided are nearly consistent, the geometric mean proves to be theoretically one of the few possible aggregations, the one preferred in [30], on the basis of a number of required properties [36]. When the judgements obtained are too dispersed, reaching a decision could imply discussing them among all the judges to obtain consensus. Solutions to measure the dispersion and overcome this disadvantage have been proposed in [37]. In our case, we need a single AHP evaluation per group of criteria, that is, we need independent cardinal rankings for the three key functionalities and the related equal-weight groups in Fig. 1 and the four equal-weight groups shown in Fig. 2. The results of our analysis are described in the next section.

4. Results

It is not our goal to criticize the overall SRI hierarchy detailed in Figs. 1 and 2, nor to propose or modify the criteria leading to the SRI. As anticipated, we only aim to check the appropriateness of the equal weights assigned to the key functionalities, the groups of impact criteria and the groups of technological domains in our local climatic and social environment. To this purpose, we got 31 contributions from various professionals in the building industry. A few judges were GUARDIANS project partners or associates and, as such, already aware of the SRI; the others include builders, architects, engineers, TBS and BACS design, installation and maintenance technicians, building managers, real estate brokers and municipality technical officers. In our views, these professional figures cover the whole building production chain, from design, deployment and maintenance of TBS to the interests of the public administration. We chose not to involve generic end users, as it would be difficult to ascertain their understanding of the problem. However, their opinions can be reflected indirectly by those of real estate brokers and building managers, through their experience with their clients. This variety should reduce the risk of a biased collective decision. All the judges were adequately briefed on the SRI hierarchy and the basics of AHP through dedicated public seminars and face-to-face meetings. A few judges were followed step by step in filling up the AHP form; we tried to interfere as less as possibile in their task, and in no case we assisted to errors due to particular

Table 2
Random indices $ri(N)$ for $3 \le N \le 15$.

N	3	4	5	6	7	8	9	10	11	12	13	14	15
$ri(N)$	0.49	0.83	1.04	1.17	1.26	1.32	1.37	1.40	1.43	1.45	1.47	1.48	1.49

Table 3
No. of cr values exceeding 0.10 and corresponding average crs for the related rankings. The two impact criteria under Energy Savings and Operation are not included since their cr is always zero.

	No. of matrices with $0.10 < cr \le 0.15$	Average cr
Key functionalities	7	0.13
Response to the needs of occupants	15	0.12
Comfort	15	0.13
Convenience	15	0.14
Information to occupants	10	0.12
Health and wellbeing	11	0.13

fatigue states. Normally, filling up the entire form took less than 45 minutes.

From each judge, we got a 3rd-order judgement matrix A_{kf} for key functionalities, a 2nd-order matrix A_{eso} for the impact criteria under Energy Saving and Operation, a 4th-order matrix A_{rno} for the impact criteria under Response to the Needs of Occupants, a 5th-order matrix A_{com} for the technological domains under the impact criterion Comfort, a 8th-order matrix A_{con} for the domains under Convenience, a 7th-order matrix A_{inf} for the criteria under Information to Occupants, and a 5th-order matrix A_{hw} for the domains under Health and Wellbeing. In some cases, we agreed small changes with the judges, following [34] and [30], with a few percent differences in the final rankings. In these cases, to adhere as much as possible to the opinions of the judges and avoid rank reversals in the final results, we accepted consistency ratios $cr \le 0.15$, rather than $cr \le 0.10$ as stated in Section 2. In other cases, even though the required cr is only exceeded in one or a few rankings, the inconsistencies were so strong that we preferred to exclude the whole judge contribution from the evaluation rather than risking to alter too much their intentions or influence their opinions. In detail, we kept modifying the values until a $cr \le 0.15$ was reached, with no rank reversal. In some cases, we could have further reduced cr, but we chose to maintain minimum alterations with respect to the original data. Limited to the values exceeding $cr = 0.10$, the results of this operation are summarized in Table 3. As apparent, less than a half of the judgements considered valid exceed the inconsistency recommended in the literature. After this selection, the number of valid contributions taken into account was 26. Ten of them are municipality technical officers (officers, hereafter), mostly architects and engineers acting in planning and monitoring public interventions, seven are engineers and technicians of a building firm with experience in masonry works supporting TBS deployment (builders), four are engineers and technicians working in design, deployment and maintenance of TBS and BACS (TBS experts), two are independent architects, one is an independent electronic engineer, one is a real estate broker and one a building manager.

In what follows, we present the results of our computation, distinguishing key functionalities, impact criteria, and technological domains. As we do not want to discuss the fixed weight of 20% assigned to the technological domain Monitoring & Control (see Fig. 2), this has not been taken into account in the evaluation, and the collective rankings for the domains under Comfort, Convenience and Information to Occupants have been scaled to sum up to 0.8 instead of 1. The corresponding histograms in Figs. 7-9, however, report values from zero to 1 for comparability. Only brief considerations are reserved to the variability of the judgements throughout the different professional subgroups. The complete raw data are available upon request.

4.1. Key functionalities

In Fig. 4, we show the histograms of the scores obtained for the three key functionalities from all the valid judgment matrices available. It is seen that the individual opinions are rather dispersed; it is also apparent, however, that most judges consider the improvement on energy efficiency obtained from BACS the most influential with respect to the building intelligence, whereas the response to the needs of the grid is generally considered the less important.

The collective ranking obtained via Eq. (2) is

$$\mathbf{v}_{kf}^{(col)} = \begin{bmatrix} \text{Energy} & 0.54 \\ \text{Occupants} & 0.31 \\ \text{Grid} & 0.15 \end{bmatrix}. \tag{5}$$

Within this general consensus, it is observed that the opinions of the broker and the manager are similar. Also the two architects show this identity of opinions. Among the officers, five judges do not agree with the priority given to energy, preferring to give more weight to occupants or grid. One builder and the independent engineer agree with this view. Only one builder assigns a higher weight to the needs of occupants than to energy savings. All four TBS experts largely agree with the result in (5). Apparently, in any case, the opinions of our judges do not justify the assumption of equal weights for the SRI key functionalities. In particular, the relatively small weight given to the functionality Response to the Needs of the Grid could reflect the observation that this alone forms one third of the final SRI in its current implementation, even though the smart grid technologies are not yet mature; this has also been identified as one of the factors that cause the low average SRIs obtained by the majority of existing buildings in Italy [13,14].

4.2. Impact criteria

The groups of impact criteria of interest here are the two under Energy Savings and Operation and the four under Response to the Needs of Occupants. These are treated separately in the next two subsections.

4.2.1. Energy savings and operation

Fig. 5 shows the two histograms related to Energy Saving and Maintenance and Fault Prediction.

In this case, the opinions are even more dispersed than before, but the final result gives a slightly larger weight to energy saving than to maintenance. The collective ranking is

$$\mathbf{v}_{eso}^{(col)} = \begin{bmatrix} \text{Energy Saving} & 0.57 \\ \text{Maintenance and Fault Prediction} & 0.43 \end{bmatrix}. \tag{6}$$

The two architects and two of the officers weight equally these two criteria; about half of the remaining judges is in favor of energy saving, and the other half favors maintenance and fault prediction. Thus, despite the dispersion in the weights assigned, the equal weighting proposed by the technical support group could be justified in this case.

4.2.2. Response to the needs of occupants

Fig. 6 shows the histograms obtained for the impact criteria under Response to the Needs of Occupants, that is, Comfort, Convenience, Information to Occupants, and Health and Wellbeing.

The collective ranking is now

$$\mathbf{v}_{rno}^{(col)} = \begin{bmatrix} \text{Comfort} & 0.32 \\ \text{Convenience} & 0.18 \\ \text{Information to Occupants} & 0.16 \\ \text{Health and Wellbeing} & 0.33 \end{bmatrix}. \tag{7}$$

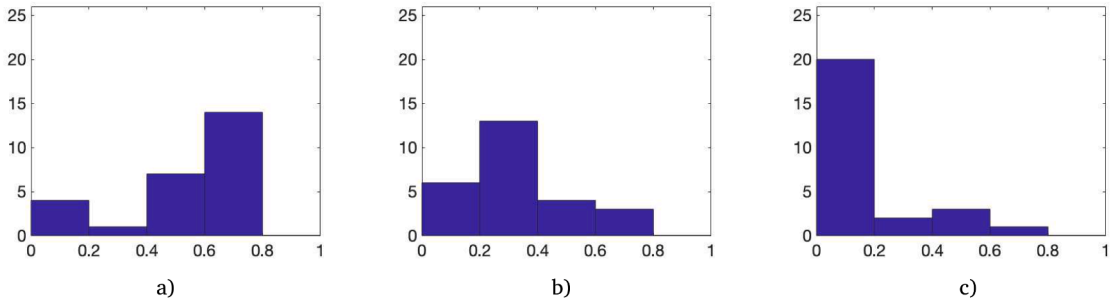


Fig. 4. Scores obtained from our data for the three key functionalities: a) Energy Savings and Operation; b) Response to the Needs of Occupants; c) Response to the Needs of the Grid.

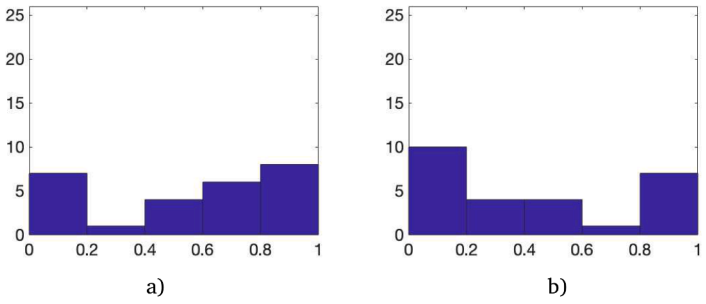


Fig. 5. Scores obtained from our data for the two impact criteria under Energy Saving and Operation: a) Energy Savings; b) Maintenance and Fault Prediction.

With respect to the equal weight choice shown in Fig. 1, comfort and wellbeing have gained importance at the expense of convenience and information to occupants. About the differences between the judges' subgroups, it is to note here that a large majority of builders and TBS experts assign a preponderant weight to health and wellbeing.

4.3. Technological domains

From Fig. 2, we see that the groups of technological domains under Response to the Needs of Occupants have received weights that are equal under each subtending impact criterion. The following subsections show the weights attributed to each of them by our judges.

4.3.1. Comfort

Fig. 7 shows the histograms for the five technological domains subtended by the impact criterion Comfort.

The collective ranking obtained is

$$\mathbf{v}_{com}^{(col)} = \begin{bmatrix} \text{Heating} & 0.25 \\ \text{Cooling} & 0.21 \\ \text{Ventilation} & 0.09 \\ \text{Lighting} & 0.16 \\ \text{Dynamic Envelope} & 0.09 \end{bmatrix}. \tag{8}$$

Whereas heating and cooling are apparently considered the most important, the domain Lighting does not lose importance with respect to the equal-weight choice, as it maintains the weight of 16% reported in Fig. 2. It is also apparent that ventilation and dynamic envelope are not considered important with respect to comfort. The broker, the manager, most of the officers and the TBS experts agree with this choice. This could depend on the fact that controlled ventilation is not so popular in Italy, even though its presence is considered important in nearly-zero energy buildings, due to their normally weatherstripped fixtures. The architects and the engineer consider ventilation nearly as important as heating and cooling, and all the builders give a considerable weight to dynamic envelope.

4.3.2. Convenience

Fig. 8 shows the eight histograms related to the impact criterion Convenience, from which it is immediate to see that Ventilation, Dynamic Envelope and Electric Vehicle Charging have not been considered very important. Almost all judges agree on this, independently of their professional subgroups.

The final ranking obtained is

$$\mathbf{v}_{con}^{(col)} = \begin{bmatrix} \text{Heating} & 0.17 \\ \text{Domestic hot water} & 0.15 \\ \text{Cooling} & 0.11 \\ \text{Ventilation} & 0.07 \\ \text{Lighting} & 0.12 \\ \text{Electricity} & 0.09 \\ \text{Dynamic envelope} & 0.06 \\ \text{EV charging} & 0.03 \end{bmatrix}, \tag{9}$$

which shows that, under the point of view of convenience, the only important domains are considered heating, domestic hot water, cooling and lighting. These are indeed the domains whose weights exceed the original 10% reported in Fig. 2. Only a few judges show significant disagreement with this conclusion. All the builders assign weights larger than 20% to electricity. The architects and the engineer also give a significant weight to ventilation.

4.3.3. Information to occupants

Under Information to Occupants, seven technological domains are deemed relevant. The related histograms are shown in Fig. 9, where, again, ventilation, dynamic envelope and EV charging are generally not considered important.

The final ranking is

$$\mathbf{v}_{inf}^{(col)} = \begin{bmatrix} \text{Heating} & 0.18 \\ \text{Domestic hot water} & 0.11 \\ \text{Cooling} & 0.14 \\ \text{Ventilation} & 0.11 \\ \text{Electricity} & 0.15 \\ \text{Dynamic Envelope} & 0.06 \\ \text{EV Charging} & 0.06 \end{bmatrix}. \tag{10}$$

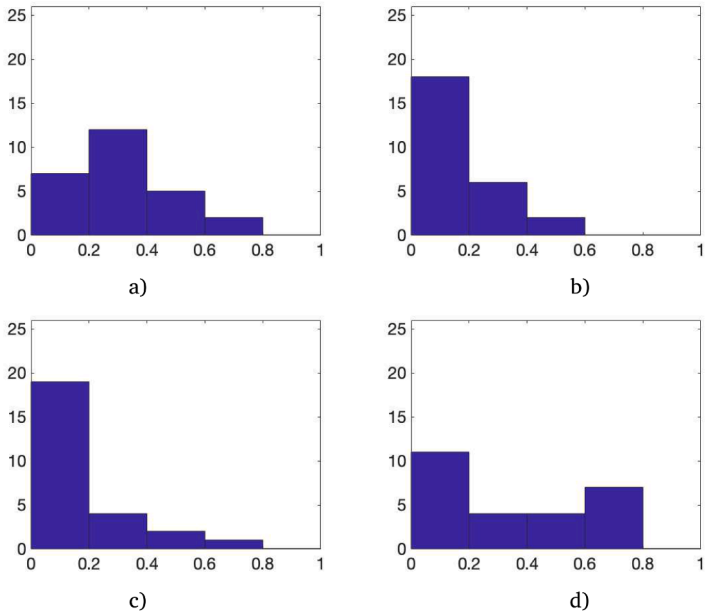


Fig. 6. Scores obtained from our data for the four impact criteria under Response to the Needs of Occupants: a) Comfort; b) Convenience; c) Information to Occupants; d) Health and Wellbeing.

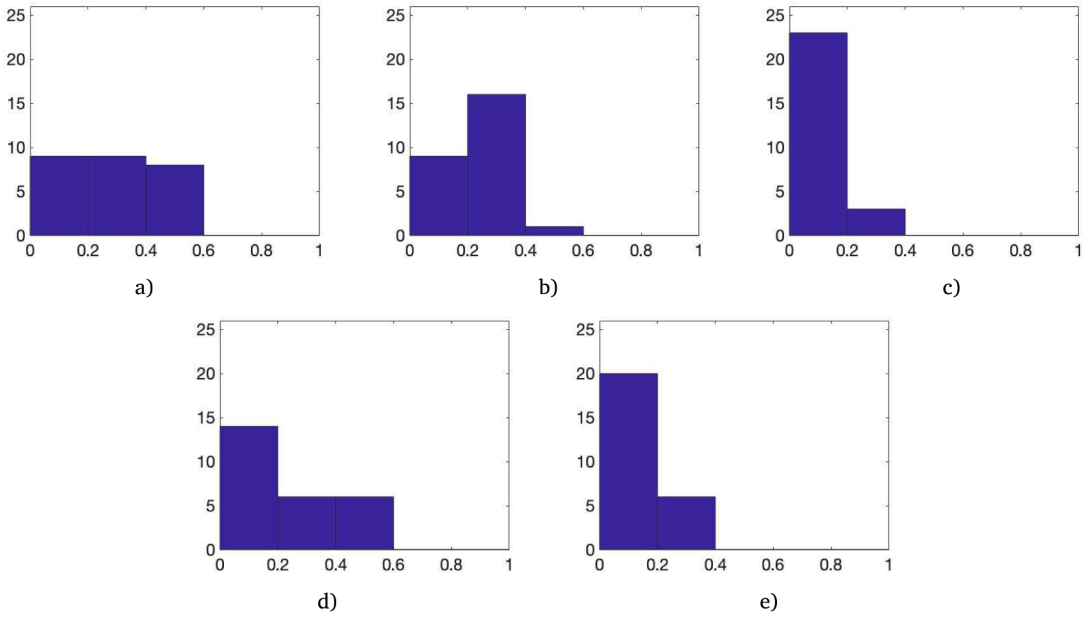


Fig. 7. Scores obtained from our data for the five technological domains under Comfort: a) Heating; b) Cooling; c) Controlled Ventilation; d) Lighting; e) Dynamic Envelope.

The only domains exceeding the original weight of 11.4% are in this case heating, cooling and electricity, for which the availability of information is considered important. Also in view of the information to occupants, all the builders assign high weights to electricity, both architects assign a 30% weight to ventilation, and one TBS expert assigns equal weights to all the domains. Note that heating, cooling and electricity are also the domains where an accurate information can improve the energy performance of the building by direct intervention of the occupants. In particular, information about the current self-production of electricity can help the user to decide the operation of most energy-consuming appliances.

4.3.4. Health and wellbeing

Fig. 10 shows the histograms obtained from the scores assigned by the judges to the five technological domains under Health and Wellbeing. The broker considers lighting important to wellbeing. The architects

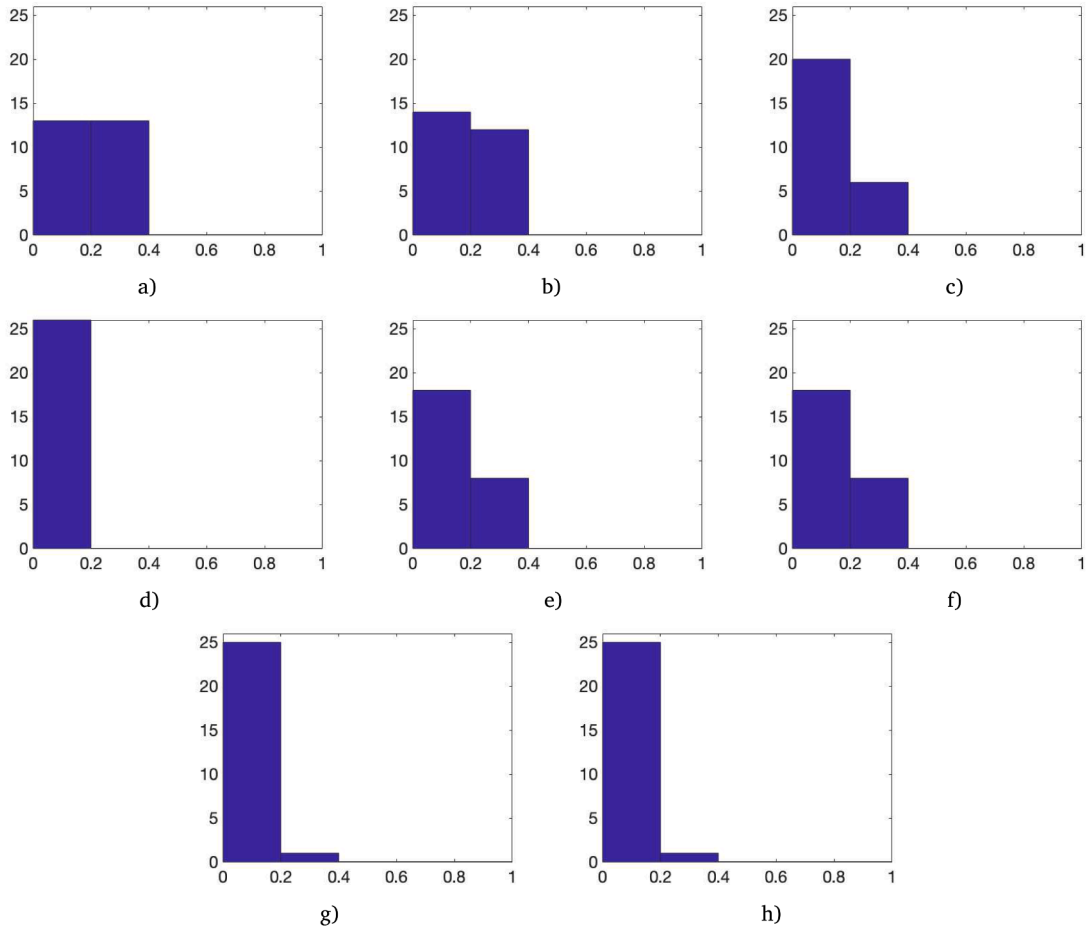


Fig. 8. Scores obtained from our data for the eight technological domains under Convenience: a) Heating; b) Domestic Hot Water; c) Cooling; d) Controlled Ventilation; e) Lighting; f) Electricity; g) Dynamic Envelope; h) EV Charging.

Table 4
Maximum perturbations in the individual rankings vs. obtained maximum perturbation in the corresponding collective rankings, from 10,000 independent simulations.

	$\mathbf{v}_{kf}$	$\mathbf{v}_{eso}$	$\mathbf{v}_{rno}$	$\mathbf{v}_{com}$	$\mathbf{v}_{con}$	$\mathbf{v}_{inf}$	$\mathbf{v}_{hw}$
Max $\lvert\Delta\mathbf{v}\rvert$	3.8%	3.3%	4.1%	5.0%	4.6%	4.5%	4.6%
Max $\lvert\nabla\mathbf{v}^{(col)}\cdot\Delta\mathbf{v}\rvert_{jj}$	$\mathbf{v}_{kf}^{(col)}$ 1.2%	$\mathbf{v}_{eso}^{(col)}$ 1.0%	$\mathbf{v}_{rno}^{(col)}$ 1.3%	$\mathbf{v}_{com}^{(col)}$ 1.5%	$\mathbf{v}_{con}^{(col)}$ 1.9%	$\mathbf{v}_{inf}^{(col)}$ 2.1%	$\mathbf{v}_{hw}^{(col)}$ 1.3%

give much importance to ventilation and dynamic envelope, in agreement with two officers. In general, by all the judges and throughout all the subgroups, ventilation is considered quite important to the occupants' wellbeing.

This is confirmed by the computed collective scores, which are

$$\mathbf{v}_{hw}^{(col)} = \begin{bmatrix} \text{Heating} & 0.31 \\ \text{Cooling} & 0.22 \\ \text{Ventilation} & 0.23 \\ \text{Lighting} & 0.14 \\ \text{Dynamic Envelope} & 0.10 \end{bmatrix}. \tag{11}$$

Indeed, the collective weight assigned to ventilation with respect to health and wellbeing is practically the same as the one assigned to cooling, and higher than the default equal weight. Lighting and dynamic architectural envelope have received a weight that is lower than the 20% assigned by the technical support group.

4.4. Stability

Once obtained the above results, it is natural to ask whether variations in the judgements would be able to reverse some of the final rankings. In other words, we would like to know whether the results obtained are robust with respect to the opinions of the judges.

Dealing with supposedly small variations in the judgements, we face this problem by computing the derivatives of the elements of vector $\mathbf{v}^{(col)}$ in (2) with respect to the scores obtained by the individual judges. With $\mathbf{v}^{(col)}$, here we mean any of the collective rankings presented in Eqs. (5)–(11), omitting the subscript. Considering the presence of the element-by-element product, the j-th component of $\mathbf{v}^{(col)}$ is

$$v_j^{(col)} = \left[\prod_{i=1}^{k} v_j^{(i)}\right]^{\frac{1}{k}}, \tag{12}$$

and its partial derivative with respect to any individual score $v_l^{(i)}$ is

$$\frac{\partial v_j^{(col)}}{\partial v_l^{(i)}} = \begin{cases} \frac{1}{k}\left[\prod_{i=1}^{k} v_l^{(i)}\right]^{\frac{1}{k}-1}\left[\prod_{i\neq j}^{1:k} v_l^{(i)}\right] & \text{if} \quad l = j \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

The $N \times k$ matrix

$$\nabla\mathbf{v}^{(col)} = \begin{bmatrix} \frac{\partial v_1^{(col)}}{\partial v_1^{(1)}} & \cdots & \frac{\partial v_1^{(col)}}{\partial v_1^{(k)}} \\ \cdots & \cdots & \cdots \\ \frac{\partial v_N^{(col)}}{\partial v_N^{(1)}} & \cdots & \frac{\partial v_N^{(col)}}{\partial v_N^{(k)}} \end{bmatrix} \tag{14}$$

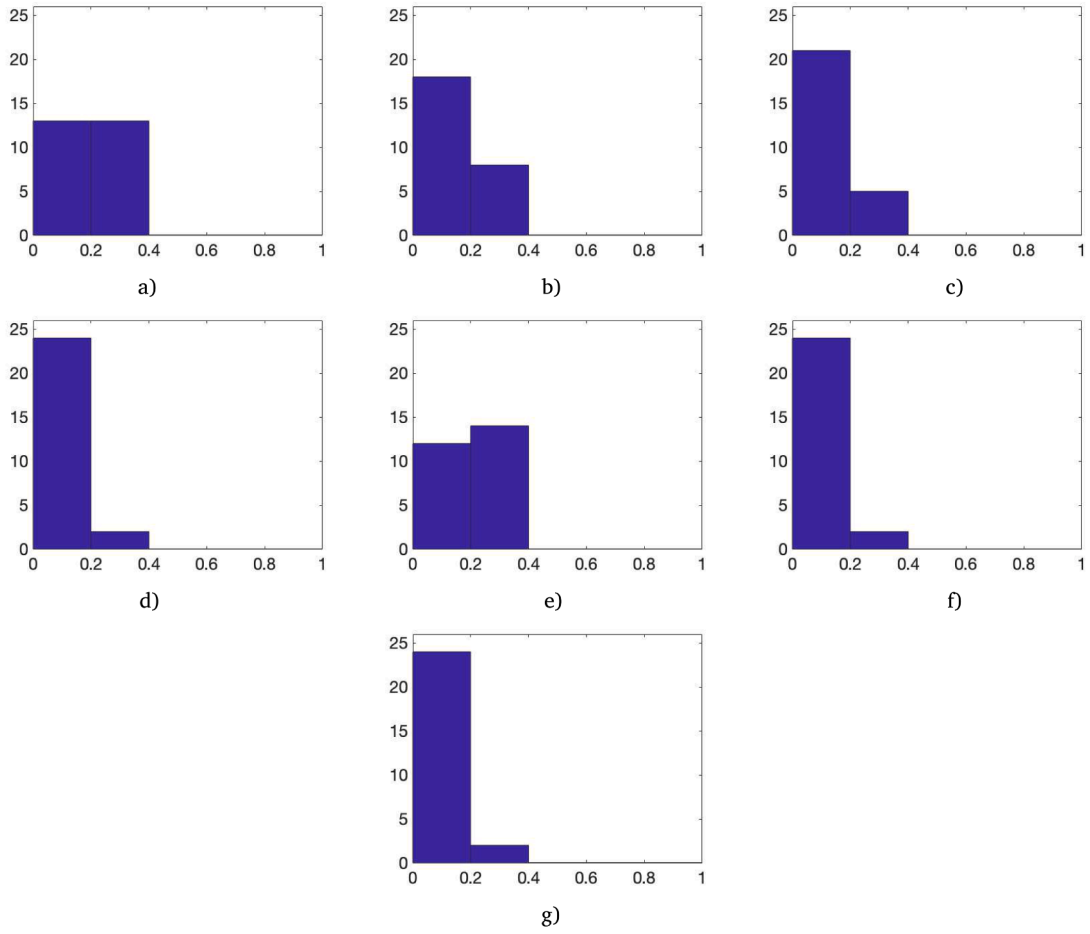


Fig. 9. Scores obtained from our data for the seven technological domains under Information to occupants: a) Heating; b) Domestic Hot Water; c) Cooling; d) Controlled Ventilation; e) Elctricity; f) Dynamic Envelope; g) EV Charging.

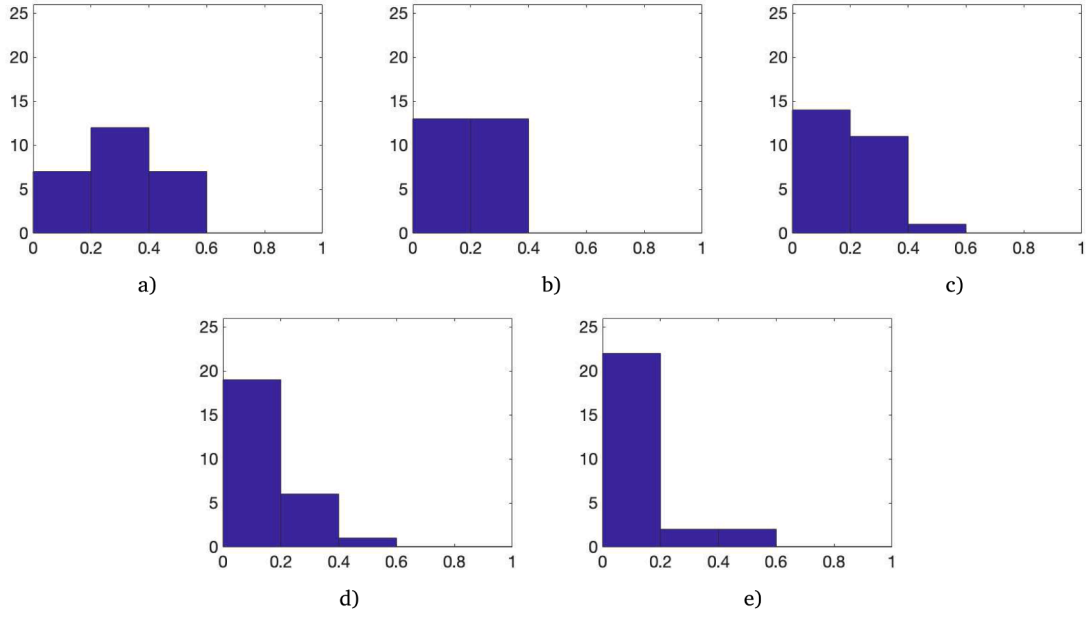


Fig. 10. Scores obtained from our data for the five technological domains under Health and Wellbeing: a) Heating; b) Cooling; c) Controlled Ventilation; d) Lighting; e) Dynamic Envelope.

Table 5
Rms errors and rank reversals in the collective rankings after removing 1 or 2 data columns.

Final ranking	# rankings excluded	Mean rms difference	Standard dev rms difference	Max rms difference	# rank reversals
$\mathbf{v}_{kf}^{(col)}$	1	1.2%	0.8%	3.6%	0
	2	1.8%	1.0%	6.8%	0
$\mathbf{v}_{eso}^{(col)}$	1	1.8%	0.9%	3.1%	0
	2	2.4%	1.6%	6.4%	0
$\mathbf{v}_{rno}^{(col)}$	1	1.5%	0.6%	2.4%	9
	2	2.1%	1.0%	4.5%	125
$\mathbf{v}_{com}^{(col)}$	1	1.1%	0.4%	2.1%	0
	2	1.6%	0.7%	3.8%	39
$\mathbf{v}_{con}^{(col)}$	1	1.0%	0.3%	1.6%	1
	2	1.4%	0.5%	3.0%	46
$\mathbf{v}_{inf}^{(col)}$	1	0.9%	0.4%	1.8%	9
	2	1.3%	0.5%	3.4%	197
$\mathbf{v}_{hw}^{(col)}$	1	1.0%	0.4%	2.1%	7
	2	1.5%	0.6%	4.2%	114

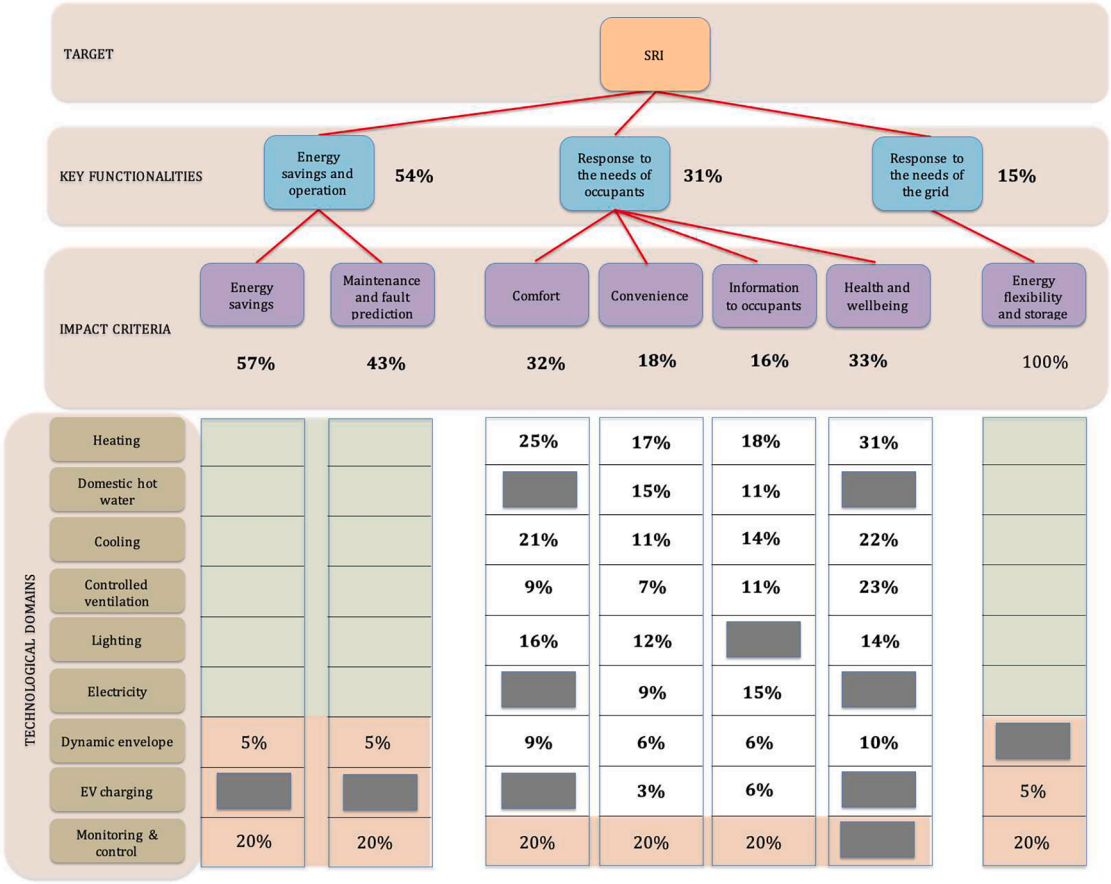


Fig. 11. The complete SRI hierarchy with the weights assigned collectively by our judges (boldfaced figures).

can be used to evaluate the variations in $\mathbf{v}^{(col)}$ produced by a difference $\Delta\mathbf{v}$ in the individual rankings:

$$\Delta\mathbf{v} = \begin{bmatrix} \delta v_1^{(1)} & \dots & \delta v_N^{(1)} \\ \dots & \dots & \dots \\ \delta v_1^{(k)} & \dots & \delta v_N^{(k)} \end{bmatrix} \tag{15}$$

The product of the j-th row of $\nabla\mathbf{v}^{(col)}$ and the j-th column of $\Delta\mathbf{v}$ estimates the variation in $v_j^{(col)}$ produced by the variations $\delta v_j^{(i)}$, with $1 \le i \le k$. To have an idea of how robust is our result, we computed the derivatives in (13) for all j and i using the values of $v_j^{(i)}$ derived by the individual judgements, then we generated 10,000 random matrices $\Delta\mathbf{v}$ for each ranking, where each column is Gaussian distributed with zero mean and standard deviation 0.01. Also, each row in $\Delta\mathbf{v}$ was subtracted of its mean value to maintain all the rankings to sum up to 1. The maximum perturbations contained in $\Delta\mathbf{v}$ and the related maximum output perturbations, for all the rankings, are reported in Table 4.

With maximum positive or negative differences of 3% to 5% in the individual scores, we obtained 1% to 2% output perturbations. This means that our results are robust against small perturbations.

Another aspect of stability is the robustness of the result against possible outliers. Since, as demonstrated by the histograms above, the indi-

vidual judgements at our disposal are quite dispersed, we characterize our results under this point of view by first excluding, one by one, all the individual rankings from each data matrix, then computing again the geometric mean in (2), the root-mean-square difference between the result and the corresponding $\mathbf{v}^{(col)}$, and comparing the two vectors to find possible rank reversals. We repeated this procedure by removing, one by one, all the 325 possible pairs of individual rankings from each data matrix. The results are shown in Table 5.

Whereas, in many cases, a significant number of rank reversals occur, the rms errors never exceed a few percent. This means that most reversals occur between final scores that are not very different from each other. In any case, it is apparent that our data contain quite a few outliers.

5. Conclusion

To offer a synoptic view of our results, Fig. 11 shows the complete SRI hierarchy, obtained by merging Figs. 1 and 2 and reporting the weights taken from our collective rankings $\mathbf{v}_{kf}^{(col)}$, $\mathbf{v}_{eso}^{(col)}$, $\mathbf{v}_{rno}^{(col)}$, $\mathbf{v}_{com}^{(col)}$, $\mathbf{v}_{con}^{(col)}$, $\mathbf{v}_{inf}^{(col)}$, and $\mathbf{v}_{hw}^{(col)}$, Eqs. (5)–(11). The weights included in the orange-background cells have been left at their fixed default values.

The first observation that can be made is that the equal-weight strategy has not generally been followed by our judges (except for a few individuals in particular rankings). The absence of scientific evidence, thus, does not seem to be a sufficient reason to choose equal weights.

Of course, the fact that our judges form a heterogeneous and unbalanced group does not allow us to say that the values in Fig. 11 can constitute a proposal for changing the SRI weighting scheme. Such a small group could well produce biased results. On the other hand, our judges are not chosen at random from the general population: all of them are experienced professionals active in the building industry at large, including in design, deployment and maintenance of TBS, and all of them have been informed about our research goals, the basics of SRI, and the mechanism of AHP. Heterogeneity could avoid biases: as far as we know, the judgements have been produced individually by all the judges, so no mutual influence occurred, and the variety of opinions detectable from the histograms in Figs. 4–10 confirms this belief. By the way, the judgements are diverse even within subgroups belonging to homogeneous categories.

Another reason why our result cannot be taken as a concrete proposal to change the present SRI weighting scheme is the local specificity of our approach, which is common to some of our references [10,13,17]: all our judges come from a circumscribed and relatively small area, so their preferences could well be influenced by the climatic and and social realities where they run their businesses. Besides the differences in energy balances, implementing the SRI certification in different regions should also take into account these differences when dealing with criteria whose effects are not measurable. The impact criteria of our main interest here, that is, the ones related to the needs of the occupants, can hardly get evaluations that are independent of the social and geographical context of the buildings being assessed. The opinion of a local judge can well be influenced by the local climate: within the North and Nort-East Europe climate zones considered in [7], it is likely that little attention is reserved to the cooling technological domain in view of comfort or health and wellbeing. Also, the local technological context may have an influence: in a region where the smart grid technologies are not available or not fully exploited, or where the monitoring and control systems are not widely adopted, the impact criterion of energy flexibility almost loses significance in the eyes of many judges. Finally, the social context, including economic aspects, might condition significantly the judgements: one of the declared goals of the introduction of the SRI is to encourage specific technological additions in building retrofitting; the opinion of a building professional is certainly influenced by the needs expressed by their clients, who could underestimate or overestimate the use of smart technologies in view of their habits or economic situations. The policies implemented by the authorities can guide the attitudes of both the public and the professionals and, eventually, of the market, thus contributing to reach the intended effects. Besides the Climate action mentioned in the Introduction, an effective implementation of the SRI would also be beneficial to the UN sustainable development goals 3, Good Health and Well-Being, 9, Industry, Innovation and Infrastructure and 11, Sustainable Cities and Communities.

The data presented here could have been assumed as a true proposal for a new weighting scheme if a consensus were reached on the basis of an open discussion among the judges, but this is out of our present scope. Furthermore, any such proposal should also be supported by similar surveys done in different local realities. Some results in this sense are presented, among others, in [10,13,17,21]. In providing the SRI calculation scheme and the related documentation, the technical support group asks all the involved stakeholders to return the results obtained by applying the scheme to actual buildings, and possibly the data used for the computation and any other document/proposal they deem is to be shared. All these data are supposed to help the support group to formulate the next revisions of the SRI scheme, thus also advising the European Commission in evolving the future policies for building intelligence.

Despite the limited scope of this study, a few remarks can be considered of general interest. First of all, subjectivity is a natural feature of human judgement, and the comparisons between the criteria taken into account here are purely subjective, although not totally arbitrary. Involving a group of experienced professionals in the evaluation ensures wise judgements to be obtained, even though from different points of view. The judgements depend on both the professional community involved and the local context where the survey is performed: probably, similar studies carried out in different contexts would have produced different results. This last remark could suggest that also the current equally weighted criteria should be replaced by something dependent on the geographical context; the choice should be left to each implementing member state.

As part of the programmed activities of project GUARDIANS, we are applying the SRI evaluation to some existing buildings as a guide to design technically and economically appropriate retrofit interventions. Our future research will include protracted measurement campaigns to evaluate the actual improvements in building performance obtained after the introduction of intelligent monitoring and control technologies and additional smart services.

Funding

This work has been partially supported by European funds through the Tuscany Region R&D project GUARDIANS: Green, Utility, Accessibility, Resilience, Digital Integration, Ability, Nature, Sustainability (PR FESR 2021-2027, CUPST: 27717.29122023.043000716). Partners: BM Srl; ElettroD Srl; Thermocasa Srl; Lucense SCarl, CNR- ISTI.

Declaration of competing interest

I declare to have no competing financial interest or personal relationships that could appear to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Eurostat, Energy consumption in households (2025). https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households.
[2] European Parliament and Council, Directive (eu) 2024/1275 of the european parliament and of the council (2024). https://energy.ec.europa.eu/topics/energy-efficiency/energy-performance-buildings/energy-performance-buildings-directive_en.
[3] European Parliament And Council, Directive (EU) 2024/1275 of the european parliament and of the council, Technical Report, 2024. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024L1275.

[4] United Nations Development Programme, Sustainable Development Goals, 2021, Last checked: November 24th, 2021. https://www.undp.org/sustainable-development-goals.
[5] T. Delavar, E. Borgentorp, S. Junnila, The smart buildings revolution: a comprehensive review of the smart readiness indicator literature, Appl. Sci. 15 (2025) 1808. https://doi.org/10.3390/app15041808.
[6] European Parliament and Council, Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency, 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2018.156.01.0075.01.ENG.
[7] S. Verbeke, D. Aerts, G. Reynders, Y. Ma, P. Waide, Final report on the technical support to the development of a smart readiness indicator for buildings, Final Report MJ-03-20-335-EN-N, European Commission, Directorate C, Unit C4, 2020. https://doi.org/10.2833/41100.
[8] M. Akbari Paydar, C. Robertson, E. Burman, D. Mumovic, A comparison between the impact of dynamic envelope control strategies on the buildings' smart readiness indicator and modelled performance, Build. Serv. Eng. Res. Technol. 46 (2) (2025) 229–250. https://doi.org/10.1177/01436244241306595.
[9] V. Apostolopoulos, P. Giourka, G. Martinopoulos, K. Angelakoglou, K. Kourtzanidis, N. Nikolopoulos, Smart readiness indicator evaluation and cost estimation of smart retrofitting scenarios - A comparative case-study in european residential buildings, Sustain. Cities Soc. 82 (2022) 103921. https://doi.org/10.1016/j.scs.2022.103921.
[10] P. Autio, E. Borgentorp, L. Pulkka, S. Junnila, Smart readiness indicator: ready for business? evidence from a northern EU country, Buildings 14 (2024) 3638. https://doi.org/10.3390/buildings14113638.
[11] C. Boje, A. Guerriero, S. Kubicki, S. Thomas, Aligning IFC and SRI domains for BIM supported SRI assessement, in: 2022 IEEE 28th International Conference on Engineering, Technology and Innovation (ICE/ITMC) & 31st International Association for Management of Technology (IAMOT) Joint Conference, IEEE - IAMOT, 2022, pp. 1–8. https://doi.org/10.1109/ICE/ITMC-IAMOT55089.2022.10033263.
[12] M. Calandri, L. Tafelmaier, T. Rathje, M. Müller, T. Henzler, K. Stergiaropoulos, M. Kramp, A. Glatz, Smart building assessments: optimizing SRI calculation using a BIM-based data exchange framework, Front. Built Environ. 11 (2025) 1634162. https://doi.org/10.3389/fbuil.2025.1634162.
[13] L. Canale, M. De Monaco, B. Di Pietra, G. Puglisi, G. Ficco, I. Bertini, M. Dell'Isola, Estimating the smart readiness indicator in the italian residential building stock in different scenarios, Energies 14 (2021) 6442. https://doi.org/10.3390/en14206442.
[14] M. Lauria, M. Azzalin, F. Giglio, G.M. La Face, Responsive public policies for smart and sustainable buildings: an experimental application of the smart readiness indicator, Buildings 15 (2025) 2002. https://doi.org/10.3390/buildings15122002.
[15] L. Martinez, T. Klitou, D. Olschewski, P. Carnero Melero, P.A. Fokaides, Advancing building intelligence: developing and implementing standardized smart readiness indicator (SRI) on-site audit procedure, Energy 316 (2025) 134538. https://doi.org/10.1016/j.energy.2025.134538.
[16] A. Ntafalias, P. Papadopoulos, A.P. Ramallo-González, A.F. Skarmeta-Gómez, J. Sánchez-Valverde, M.C. Vlachou, R. Marín-Pérez, A. Quesada-Sánchez, F. Purcell, S. Wright, Smart buildings with legacy equipment: a case study on energy savings and cost reduction through an IoT platform in ireland and Greece, Results Eng. 22 (2024) 102095. https://doi.org/10.1016/j.rineng.2024.102095.
[17] A. Xenakis, A. Arsenopoulos, I. Papias, F. Serepas, S. Rizou, J. Psarras, Advancing smart building upgrades through an innovative decision support framework for smart readiness indicator implementation: a greek case study, Build. Environ. 281 (2025) 113189. https://doi.org/10.1016/j.buildenv.2025.113189.
[18] European Commission, COMMISSION DELEGATED REGULATION (EU) 2020/2155 Of 14 october 2020 supplementing directive (EU) 2010/31/EU of the european parliament and of the council by establishing an optional common european union scheme for rating the smart readiness of buildings (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020R2155&from=en.
[19] European Commission, Commission implementing regulation (EU) 2020/2156 Of 14 october 2020 detailing the technical modalities for the effective implementation of an optional common union scheme for rating the smart readiness of buildings (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020R2156.
[20] M. Halilovič, A. Kralj, J. Kurnitski, S. Domjan, A. Maček, S. Medved, Smart readiness indicator metric for building envelope structures, Energy Build. 350 (2026) 116677. https://doi.org/10.1016/j.enbuild.2025.116677.
[21] E. Janhunen, L. Pulkka, A. Säynäjoki, S. Junnila, Applicability of the smart readiness indicator for cold climate countries, Buildings 9 (4) (2019) 102. https://doi.org/10.3390/buildings9040102.
[22] M.R. Munafò, B.A. Nosek, D.V.M. Bishop, K.S. Button, C.D. Chambers, N. Percie du Sert, U. Simonsohn, E.-J. Wagenmakers, J.J. Ware, J.P.A. Ioannidis, A manifesto for reproducible science, Nat. Hum. Behav. 1 (2017) 0021. https://doi.org/10.1038/s41562-016-0021.
[23] T.L. Saaty, The Analytic Hierarchy Process, McGraw-Hill, New York, New York, 1980.
[24] A. Mardani, A. Jusoh, K. Nor, Z. Khalifah, N. Zakwan, A. Valipour, Multiple criteria decision-making techniques and their applications - a review of the literature from 2000 to 2014, Econ. Res. 28 (2015) 516–571. https://doi.org/10.1080/1331677X.2015.1075139.
[25] C.M. Brugha, Structure of multi-criteria decision-making, J. Oper. Res. Soc. 55 (2004) 1156–1168. https://doi.org/10.1057/palgrave.jors.2601777.
[26] R. Karthikeyan, K.G.S. Venkatesan, A. Chandrasekar, A comparison of strengths and weaknesses for analytical hierarchy process, J. Chem. Pharm. Sci. 9 (2016) S12–S15.
[27] B. Apostolou, J.M. Hassel, An empirical examination of the sensitivity of the analytic hierarchy process to departures from recommended consistency ratios, Math. Comput. Model. 17 (1993) 163–170. https://doi.org/10.1016/0895-7177(93)90184-Z.
[28] H. Chen, D.F. Kocaoglu, A sensitivity analysis algorithm for hierarchical decision models, Eur. J. Oper. Res. (185) (2008) 266–288. https://doi.org/10.1016/j.ejor.2006.12.029.
[29] E. Condon, B. Golden, E. Wasil, Visualizing group decisions in the analytic hierarchy process, Comput. Oper. Res. 30 (2003) 1435–1445. https://doi.org/10.1016/S0305-0548(02)00185-5.
[30] T.L. Saaty, On the measurement of intangibles. a principal eigenvector approach to relative measurement derived from paired comparisons, Not. Am. Math. Soc. 60 (2013) 192–208. https://doi.org/10.1090/noti944.
[31] J.I. Peláez, E.A. Martínez, L.G. Vargas, Consistency in positive reciprocal matrices: an improvement in measurement methods, IEEE Access 6 (2018) 25600–25609. https://doi.org/10.1109/ACCESS.2018.2829024.
[32] N. Bebiano, R. Fernandes, S. Furtado, Reciprocal matrices: properties and approximation by a transitive matrix, Comput. Appl. Math. 39 (2020) 50. https://doi.org/10.1007/s40314-020-1075-2.
[33] E. Salerno, Integration of Analysis of the Hierarchical Process and Dempster-Shafer theory for cooperative evaluation tasks, Technical Report 421098, National Research Council of Italy - ISTI, Via Moruzzi 1, 56124 Pisa, Italy, 2020.
[34] K. Kulakowski, Notes on order preservation and consistency in AHP, Eur. J. Oper. Res. 245 (2015) 333–337. https://doi.org/10.1016/j.ejor.2015.03.010.
[35] A. Ishizaka, A. Labib, Review of the main developments in the analytic hierarchy process, Expert Syst. Appl. 38 (2011) 14336–14345. https://doi.org/10.1016/j.eswa.2011.04.143.
[36] J. Aczél, F.S. Roberts, On the possible merging functions, Math. Soc. Sci. 17 (3) (1989) 205–243. https://doi.org/10.1016/0165-4896(89)90054-1.
[37] T.L. Saaty, L.G. Vargas, Dispersion of group judgments, Math. Comput. Model. 46 (2007) 918–925. https://doi.org/10.1016/j.mcm.2007.03.004.