Article
Models of Sustainable Software: A Scoping Review

Jakub Swacha

Department of IT in Management, University of Szczecin, 71-004 Szczecin, Poland; jakub.swacha@usz.edu.pl

Abstract: Information and Communication Technologies (ICTs) play a double role in the strife for sustainable development goals, as both an enabler of green solutions and a cause of excessive consumption. While the primary focus of sustainability-related research is on the hardware aspect of ICT, its software aspect also deserves attention. In order for the notion of green and sustainable software to become widespread among practitioners, models are needed, both to be used as a reference on how to develop sustainable software, and to check whether given software or its development process is sustainable. In this paper, we present the results of a scoping review of literature on sustainable software models, based on 41 works extracted from an initial set of 178 query results from four bibliographic data providers. The relevant literature is mapped using five categories (model scope, purpose, covered sustainability aspects, verification or validation, and the economic category of the country of research), allowing us to identify recent trends and research gaps, which can be addressed in future work.

Keywords: green and sustainable software; green ICT; sustainable software models; sustainable development process models; software sustainability

1. Introduction

With the rising awareness of sustainability, also the role of Information and Communication Technologies (ICTs) in achieving sustainable development goals comes to attention. There are two essential aims that are to be pursued in this regard: stopping the growth of ICT’s own footprint and applying ICT as an enabler in order to reduce the footprint of production and consumption by society [1]. The first aim corresponds to the first-order environmental effects of ICT (direct effects of the production and use of ICT), whereas the second aim corresponds to its second- and third-order environmental effects (indirect impacts through the change of production processes, products, and distribution systems, and through impacts on lifestyles and value systems) [2].

While there is an ongoing debate whether the combined impact of ICTs on energy consumption is positive or negative [3], the increased energy consumption due to the direct ICT effect has become substantial and still grows: the worst-case forecasts estimate that ICT electricity usage could contribute up to 23% of the globally released greenhouse gas emissions in 2030 [4]. This menace can be addressed with green ICT, which can be defined as “the study and practice of designing, manufacturing, using, and disposing of computers, servers, and associated subsystems—monitors, printers, storage devices, and networking and communications systems—efficiently and effectively with minimal or no impact on the environment” [5].

The focus of the definition given above is on hardware ICT components, whose environmental footprint in each stage of their lifecycle includes, respectively, e.g., materials and energy used for production, fuels in transport, energy, and consumables (e.g., replaceable batteries) during product use, and handling the waste at their end-of-life [6]. As most of the footprint causes in this case are relatively easy to identify, it is also relatively easy, in consequence, to direct effort to devising and implementing adequate, more sustainable methods and technologies. In contrast, the intangible nature of software makes it more difficult to identify and address its sustainability deficiencies. However, the role of software
in ICT is no less important than that of hardware, so the strife for green ICT cannot be limited to the latter. Hence the idea of green and sustainable software, which has been defined as software whose [7]

- “direct and indirect consumption of natural resources, which arise out of deployment and utilization, are monitored, continuously measured, evaluated and optimized already in the development process”;
- “appropriation and utilization aftermath can be continuously evaluated and optimized”;
- “development and production processes cyclically evaluate and minimize their direct and indirect consumption of natural resources and energy”.

Note that we do not differentiate the terms “green and sustainable software”, “sustainable software”, and “green software” in this paper.

An essential success factor for the development of green and sustainable software is the creation of models specifying both components and qualities of such software, as well as activities and procedures to be included in the software development process to consider it as green and/or sustainable. Such models can obviously be used both to guide the designers and developers pursuing sustainability and to evaluate whether the given software product or development process meets the sustainability criteria.

In spite of the rising interest in this field, there is a lack of an up-to-date review of models proposed hitherto in the literature. We were able to identify merely four examples of similar attempts, although all of them have wider scope (i.e., are not limited in their scope to model proposals, but report all kinds of research on sustainability in software engineering) and are at least four years old, which is a considerable time in this quickly advancing field. The most recent one (from 2017) is the systematic mapping by Berntsen et al. [8] reporting 36 papers, of which 20 were describing some kind of a model (as declared by the authors, as Table 2 therein lists only 18 works). Unfortunately, the work is somewhat biased with regard to selection of sources to consider (including only those that Østfold University College has access to), and limited in the scope of reported findings, the most valuable of which is the mapping of papers to five predefined research topics (“energy efficiency, development methodology, process enhancement, organizational metrics, life cycle thinking”); the other findings include the most prolific authors and the most frequent publication venues.

A 2014 systematic literature review on sustainability studies in software engineering by Ahmad et al. [9] examined 175 papers, of which 32 were found relevant to sustainability in software engineering. The detailed results were presented for 28 papers, mapping them in three categories: research type, topic, and application domain; only 5 of them had their type categorized as “model”.

Also in 2014, Penzenstadler et al. [10] presented a systematic mapping including 83 works on both software being sustainable, agnostic of purpose, and software developed for the purpose of supporting sustainability goals. The identified works were mapped to 10 synthetic topics (obtained via unsupervised clustering based on the contents of abstracts), 15 knowledge areas (within software engineering), 7 research type facets (such as philosophical or exploratory), and 10 application domains (“software engineering and lifecycle; energy efficiency; services, mobile and cloud; business and economics; systems engineering and ICT; ULS green computing; mechanics and manufacturing; nature and agriculture; metropolitan areas and housing; software engineering education”). Moreover, the most active research groups in the area were identified.

Two years earlier, Penzenstadler et al. [11] published a systematic literature review on sustainability in software engineering, where 96 relevant works were identified and mapped to research type, topic, application domain, and category (including “sustainability in software engineering, sustainability-related application domains, sustainability concept, sustainable software solutions, and sustainable hardware solutions”). While 18 of the papers analyzed there were categorized as describing models, only 4 of them pertained to sustainable software solutions.
The aim of this study is to examine the current (i.e., at the end of 2021) state of research on models for green and sustainable software, and map key properties of the proposed models, including their scope, purpose, covered sustainability aspects, and performed verification or validation, as well as the economic category of the country where the research has been performed.

2. Materials and Methods

Considering the aim of this study, scoping review has been chosen as the main research method. Scoping review is believed to be an effective means of highlighting the relevant literature to the researcher, especially useful for rapid mapping of key concepts relevant to the area [12]. Although similar to a systematic review in that it is based on explicit and transparent search criteria, in contrast to systematic review, it aims at providing an overview of the published results rather than their formal synthesis and is neither limited to a specific study type nor to research meeting predefined quality standards [13].

Scoping review is usually performed in five subsequent steps: (1) identify research questions, (2) identify relevant studies, (3) study selection, (4) chart the data, and (5) collate, summarize, and report the results [14]. In the presented research, we followed this scheme, abstaining from the sixth optional step of asking external stakeholders to provide their insights on the findings.

Five research questions were stated:

(RQ1) What is the purpose of the models of sustainable software presented in the literature?
(RQ2) Which aspects of software sustainability are covered by the models presented in the literature?
(RQ3) Does the literature introducing the models of sustainable software provide their verification and/or validation?
(RQ4) Is the research effort on the models of sustainable software specific to developed economies only?

As our aim was a possibly broad coverage of the literature on the sustainable software models, we decided to use Google Scholar [https://scholar.google.com/, accessed date: 11 December 2021] as the first data source, considering it as the most inclusive of the available publication search engines. The main limitation of Google Scholar is its meager capabilities with regard to specifying search terms, especially limiting their scope to specific metadata fields. Currently, the keyword search in Google Scholar can either be performed on just the publication title or on its full text. For the purposes of this research, the former is too narrow (as sometimes the publication title is quite abstract, whereas its actual topic is revealed in the keywords specified by the author), the latter far too wide (returning thousands of publications that merely mention the concept sought for).

This is why we decided to include two of the largest scientific bibliographic database access platforms in general: Scopus [https://www.scopus.com/, accessed date: 11 December 2021] and Web of Science [https://clarivate.com/webofsciencegroup/solutions/web-of-science/, accessed date: 11 December 2021], and the largest database of scientific publications in computer science—DBLP [https://dblp.org/search, accessed date: 11 December 2021], even though most of their content is also indexed by Google Scholar.

Having the strong intention of limiting the analysis to works dedicated to sustainable software models rather than its exemplary implementations, we decided that the search terms should include two elements: “sustainable software” and “model”. Although the results of the search for “sustainable software” should also include those mentioning “green and sustainable software”, we were aware of relevant literature using only the “green software” term, for which reason the second pair of search terms was defined (“green software” and “model”). As we considered the relative effectiveness of the respective search terms as interesting, we did not join them using an alternative operator, but instead performed two distinct queries on each of the considered data sources.

After a preliminary test of the search terms, we found that the inclusion of the full text and even the abstract in the search scope produced many false positives, thus we decided
to limit the scope to the title and keywords whenever possible (for Google Scholar and DBLP where it was not possible, the search scope was limited to only the title). As the DBLP search engine does not seem to allow for multi-word search phrases, “sustainable software” and “green software” were searched for as pairs of words. The exact search terms used to query the respective sources are listed in Table 1. All search queries were performed on 11 December 2021.

Table 1. Exact search terms used to query the respective sources.

<table>
<thead>
<tr>
<th>Source</th>
<th>Term 1</th>
<th>Term 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Scholar</td>
<td>allintitle: “sustainable software” model allintitle: “green software” model</td>
<td></td>
</tr>
<tr>
<td>DBLP</td>
<td>sustainable software model$</td>
<td>green software model$</td>
</tr>
<tr>
<td>Scopus</td>
<td>(TITLE (“sustainable software”) OR KEY (“sustainable software”)) AND (TITLE (“model”) OR KEY (“model”))</td>
<td>(TITLE (“green software”) OR KEY (“green software”)) AND (TITLE (“model”) OR KEY (“model”))</td>
</tr>
<tr>
<td>Web of Science</td>
<td>(TI = (“sustainable software”) OR AK = (“sustainable software”)) and (TI = (“model”) OR AK = (“model”))</td>
<td>(TI = (“green software”) OR AK = (“green software”)) and (TI = (“model”) OR AK = (“model”))</td>
</tr>
</tbody>
</table>

The study selection consisted in (a) combining results of both queries of each source, (b) removing duplicates (i.e., papers found by both queries), (c) combining results from the respective sources, (d) removing duplicates (i.e., papers found at more than one source), (e) identifying the location of full texts of the papers, and removing those without full text available, and (f) screening the papers’ contents to verify their relevance, and removing irrelevant papers.

The following examples of research were considered as irrelevant for the study:

- Models of software supporting sustainable processes in specific areas, e.g., in agriculture, road traffic analysis, or waste management; while such software is of primary importance for sustainable development in respective domains (see, e.g., [15]), its models provide little to no guidelines on the development of green and sustainable software in general;
- General software development, not aimed at sustainable software products or processes;
- IT industry policy making, which belongs rather to the area of sustainable government than to sustainable software;
- Business models;
- Knowledge management models;
- Network models;
- Research models;
- Models for visualization and reporting;
- Position papers that merely indicate an intention of research but do not provide its results;
- Derivative papers that describe a model published earlier without any improvements to the model, its validation, or verification (provided the original publication is included in the analyzed set).

Only texts having full text available were considered in the study. We were positively surprised that only one of the identified papers (found at Google Scholar) seemed to be no longer available at its only known location. Although there were more publications not accessible at the links provided by their respective source (or having no links provided at all), querying the main Google search engine for their title was sufficient to resolve all such cases.

The set of analyzed papers was eventually extended with papers categorized as models in earlier literature surveys, missing from the search results, and found to be actually relevant to the topic of sustainable software models. Three such papers were identified, all in source [9].
The complete flow of the study selection is presented in Figure 1. Interestingly, for all sources but Google Scholar, the query for green software models resulted in more results than did the query for sustainable software models.

![Figure 1. The flow of the study selection.](image)

In the context of performing systematic reviews in the future, we note that the queries yielded most results from Google Scholar (43), a little less from Scopus (41), much less from DBLP (29), and even less from Web of Science (25). If we consider, however, the relevant items only (including those found at more than one source), most were found at Scopus (27), then at Google Scholar (23), with Web of Science reporting only about half as many items as found in Scopus (15), and DBLP providing the least items (10). The relevance of search results was therefore highest at Scopus (66%), a little less at Web of Science (60%) and Google Scholar (51%), and lowest at DBLP (34%), which could be attributed to the specific search terms used in that platform (comprising words instead of phrases).

3. Results

3.1. Scope of the Identified Sustainable Software Models

In the context of RQ1, we identify two types of sustainable software models in the literature: those pertaining to software product (i.e., defining properties and/or components of sustainable software) [7,16–33], and those pertaining to software development process (i.e., defining properties and/or steps of a sustainable software development process, or of a software development process leading to the production of sustainable software) [31–55]. Three among the analyzed works address both these subjects [31–33]. In Figure 2, one can observe that models of both kinds were proposed repeatedly in the last 11 years.

![Figure 2. The number of publications in subsequent years depending on the scope of the presented model.](image)
In total, software product models seem to be more popular than those of the development process—as the former were found in 61% of the identified publications, and the latter in 46% (note the sum is over 100% due to three papers belonging to both categories). The interest in software product models achieved a high plateau in years 2014–2019, whereas the interest in software development process models had one-year spikes in 2011, 2013, and 2017. On the other hand, there were no publications on software development process models in 2018 and 2019.

3.2. Purpose of the Identified Sustainable Software Models

Answering RQ2, we identify two basic purposes of the presented models: providing a reference and guidelines for those participating in various roles in the software development life cycle in order to help them attain the qualities of sustainable software products and processes (we call them reference models) [7,16–19,22,23,25–28,30–34,36,38,46,49,53], and providing methods, measures, and tools that can be applied to verify whether given software products and processes can be considered as green or sustainable (we call them evaluation models) [17,19–21,24,26,29,30,32,33,35,37,39–45,47,48,50–52,54,55]. Six among the analyzed works address both these purposes [17,19,26,30,32,33]. Figure 3 illustrates how the number of publications dedicated to these two purposes varied in the last 11 years.

![Figure 3](https://example.com/figure3.png)

**Figure 3.** The number of publications in subsequent years depending on the purpose of the presented model.

In total, evaluation models (63%) were found to be more often described than reference models were (51%; note the sum is over 100% due to six papers belonging to both categories). The interest in the reference models visibly declined after 2017, whereas the interest in evaluation models rose in 2014 and has remained relatively constant since then.

3.3. Aspects of Sustainability Covered by the Identified Models

Software sustainability can be defined in various dimensions, such as environmental, technical, social, individual, and economic [54]. In an attempt to answer RQ3, we analyzed the identified models and found out that the papers can be categorized into two groups: those addressing environmental sustainability, in almost all cases pursuing the goal of energy efficiency ([17–19,21–23,32–44,46–52,54,55]), and those taking a holistic approach, addressing sustainability in its various aspects ([7,16,20,22–33,44,45,47,52–54]). Note that ten items ([22,23,26,29,32,33,44,47,52,54]) were included in both groups as these models consider more than one sustainability dimension but treat the topic of energy efficiency with special attention, unlike the remaining models of the second group. Figure 4 presents the number of publications from each of these categories published in respective years.
Figure 3. The number of publications in subsequent years depending on the purpose of the presented model.

3.3. Aspects of Sustainability Covered by the Identified Models

Software sustainability can be defined in various dimensions, such as environmental, technical, social, individual, and economic [54]. In an attempt to answer RQ3, we analyzed the identified models and found out that the papers can be categorized into two groups: those addressing environmental sustainability, in almost all cases pursuing the goal of energy efficiency ([17–19,21–23,32–44,46–52,54,55]), and those taking a holistic approach, addressing sustainability in its various aspects ([7,16,20,22–33,44,45,47,52–54]). Note that ten items ([22,23,26,29,32,33,44,47,52,54]) were included in both groups as these models consider more than one sustainability dimension but treat the topic of energy efficiency with special attention, unlike the remaining models of the second group. Figure 4 presents the number of publications from each of these categories published in respective years.

Figure 4. The number of publications in subsequent years depending on the covered sustainability aspects.

In total, 51% of the sustainable software models consider more than one dimension of sustainability, the remaining 49% are focused at its environmental aspect, energy efficiency in particular, which is also treated with special attention by 48% of the models from the first group. The interest in both kinds of models was sustained throughout the entire period, with the number of models dealing with energy efficiency only attaining its peak in 2016–2017.

3.4. Verification and Validation of the Identified Sustainable Software Models

The adequateness and usefulness of models can be confirmed by their validation or verification. Answering RQ4, among the analyzed papers, there is a clean distinction between the models pertaining to software product and those pertaining to software development process (see Section 3.1): whereas the models of the first group are relatively often verified experimentally (59% of the models [34–36,38,39,41–43,46,48–50,52]), none of the latter were (including those covering both scopes); however, four of the models from the second group (21%) were validated, typically by performing a survey among experts or users acquainted with the model ([24,25,27,28]). While such a difference does not occur between reference and evaluation models, it can be observed between models limited to energy efficiency (of which 60% were verified or validated), holistic models (36% verified or validated), and models that consider more than one sustainability dimension but treat the topic of energy efficiency with special attention (10% verified or validated). Figure 5 visualizes these findings.

3.5. Location of the Research Effort on the Sustainable Software Models

In order to address RQ5, we use the classification of countries defined in United Nations’ World Economic Situation Prospects [56]. We assign a paper to a respective country group if all its authors are affiliated with an institution located in a country belonging to that group. Two of the papers were written by authors from countries belonging to different groups; 51% of the remaining ones can be attributed to developing countries, and 49% to developed ones. There is only one paper co-authored by researchers from a country in transition. As shown in Figure 6, the shares of contribution of developing and developed countries keep fluctuating, with the exemption of the last two years, when only contributions from authors from developing countries were identified.
were proven successful or failures. We have observed a small number of reference models published in recent years, which can be a base for further study, whether the reason of this situation is that earlier models have been validated, providing methods, measures, and tools for verification of such qualities.

Regarding the topic of energy efficiency with special attention (10% verified or validated), and models that consider more than one sustainability dimension but treat energy efficiency (of which 60% were verified or validated), holistic models (36% verified or validated), and multi-aspectual models (43% verified or validated), we discovered a difference between process- and product-oriented models. In the combined context of RQ1 and RQ4, we discovered a difference between process- and product-oriented models being developed, with more effort seemingly dedicated to the latter. In the combined context of RQ1 and RQ4, we discovered a difference between process- and product-oriented models being developed, with more effort seemingly dedicated to the latter. In the combined context of RQ1 and RQ4, we discovered a difference between process- and product-oriented models being developed, with more effort seemingly dedicated to the latter.

Figure 5. The number of publications providing model verification or validation in each category.

Figure 6. The number of publications in subsequent years depending on the status of the country of their authors.

4. Discussion

The identified number of works devoted to such models (41, or about 3.4 research papers per year) shows that there is a continued interest in developing models for sustainable software, but it cannot be considered a trendy research topic.

Although the analyzed time frame had not been limited, there were no relevant publications identified earlier than 2010 (the first papers to deal with the subject area were [18,39]). We also observe that the interest in the subject kept rising till 2017, when it fell to pre-2014 level, at which it stays. This may indicate that the need for such models could have been satisfied to some extent by 2017, but not all aspects were adequately addressed, hence the later publications. Nonetheless, considering the rising awareness of sustainable development, and both worldwide and regional initiatives towards its implementation, such as the European Green Deal [57], we may expect the interest in the area to increase in the forthcoming years.

Regarding RQ1, we have found that both process- and product-oriented models keep being developed, with more effort seemingly dedicated to the latter.
models, as a much larger share of the latter had been verified or validated. In fact, none of the models dealing with software development processes had actually been verified, i.e., experimentally demonstrated to provide the promised sustainability-related benefits compared to software development based on other models in controlled conditions. We consider discovering this research gap as a valuable result of this study, indicating a direction worth pursuing in future research.

With regard to RQ2, we have identified interest in both reference models, useful at attaining sustainability-relevant qualities of software products and processes, and evaluation models, providing methods, measures, and tools for verification of such qualities. We have observed a small number of reference models published in recent years, which can be a base for further study, whether the reason of this situation is that earlier models were proven successful or failures.

Regarding RQ3, we have identified energy efficiency as the single aspect of sustainability most often covered in the surveyed models. This is in line with expectations, as software is developed and runs on computers, powered with electric energy, which can be saved in a number of ways. We consider, nonetheless, a positive finding that part of research is dedicated to a more dimensional notion of sustainable software.

With regard to RQ4, we have discovered an unsatisfactory share of models published along with their verification or validation results. As mentioned earlier, it is especially evident for models of software development processes, but also models covering multiple aspects of software sustainability. Verification of already published models can therefore form an interesting vein of research, especially for industry-based researchers.

Regarding RQ5, we were surprised to find that the research effort on the models of sustainable software is almost equally shared between researchers from developed and developing countries. No less surprising is a minimal activity on behalf of authors from countries having economies in transition. This finding also calls for further research.

5. Conclusions

The key contribution of the reported research is the mapping of existing literature on green and sustainable software models using five categories (model scope, purpose, covered sustainability aspects, verification or validation, and the economic category of the country of research). This clearly addresses the gap left by prior studies on sustainable software engineering, not only by providing a more up-to-date view (the previous most recent review [8] covered only papers published till 2015, with a single included paper from 2016, which amounts to six more covered years, including the year of peak output on the topic—2017), but also by clearly focusing on sustainable software models, rather than including them as one of the results of research having a much wider scope. We have also applied much stricter inclusion criteria, considering as sustainable software models only those models that either bring sustainability-motivated improvements to software development processes and products, or allow to evaluate such processes and products in terms of sustainability criteria. Thus, we did not consider any models that were only indirectly related to sustainability or to software development.

The presented research results were obtained using a wide selection of bibliographic data sources, and by following the scoping review method—not placing any quality requirements on the considered literature—and defining only simple, general research questions that could be answered using the full set of identified relevant papers rather than its small subset. Despite these efforts, the reported research shares the limitations with other studies based on literature surveys, as there is a possibility that an inclusion of more data sources and changing the used search terms and/or even the whole procedure could yield much more relevant papers.

As several interesting research directions were indicated in the discussion of the results, there are obvious opportunities for future work, the most promising of which is the practical verification of models of sustainable software development processes.
Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data were obtained from Google Scholar, DBLP, Scopus, and Web of Science, and are available from the author with the permission of their respective providers.

Conflicts of Interest: The author declares no conflict of interest.

References

5. Murugesan, S. Harnessing green IT: Principles and practices. *IT Prof.* 2008, 10, 24–33. [CrossRef]


47. Moraga, M.; Bertoa, M. Green software measurement. In Green in Software Engineering; Springer: Cham, Switzerland, 2015; pp. 261–284. [CrossRef]


