Green and Sustainable Software Engineering - a Systematic Mapping Study

Brunna C. Mourão
Federal University of Bahia
Salvador, BA, Brazil
brunna.caroline@ufba.br

Leila Karita
Federal University of Bahia
Salvador, BA, Brazil
leila.karita@ufba.br

Ivan do Carmo Machado
Federal University of Bahia
Salvador, BA, Brazil
ivan.machado@ufba.br

ABSTRACT

Understanding how the concepts of sustainability could be incorporated to the Software Engineering (SE) concepts has gained increased attention in the last years, particularly in terms of the Software Development Life Cycle (SDLC). Several studies have addressed the impact of sustainability in the SE practice, from a range of perspectives. This study presents a systematic mapping study that aggregates, summarizes and discusses the state-of-the-art approaches for sustainable SE practice. We analyzed 75 relevant primary studies addressing methods, processes, tools and metrics proposed to support the software development in a sustainable way. The included primary studies were selected using inclusion and exclusion criteria applied to studies published prior to 2017. They were analyzed based on a set of classification criteria, including contribution types, SDLC phases, evidence types, research types, application domains, publication venues, distribution between academia and industry and research methods. The results indicated a growing interest by the SE research community in the Green and Sustainable software domain. Besides, there is an observed need for more studies on techniques, tools and metrics covering construction, testing and maintenance. The results also point out a clear view of the SE community about the need for a better alignment between research and practice in this domain.

CCS CONCEPTS

• **Software and its engineering → Software organization and properties**; • **Social and professional topics** → *Sustainability*;

KEYWORDS

Systematic mapping study, green software engineering, sustainable software.

ACM Reference Format:
Brunna C. Mourão, Leila Karita, and Ivan do Carmo Machado. 2018. Green and Sustainable Software Engineering - a Systematic Mapping Study. In *17th Brazilian Symposium on Software Quality (SBQS), October 17–19, 2018, Curitiba, Brazil.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3275245.3275258

SBQS, October 17–19, 2018, Curitiba, Brazil

ACM ISBN 978-1-4503-6565-9 / 18/10 . . . $15.00
https://doi.org/10.1145/3275245.3275258

1 INTRODUCTION

Approximately 97% of the climate scientists agree that global warming trends over the last century are likely to have been the result of human activities [9]. Aligned with this concern, the Software Engineering community has increased its interest in unveiling the impacts of the Information and Communication Technology on the excessive consumption of natural resources.

Such a role has been played by the Green and Sustainable SE field, or Green SE for short. Green SE aims to create reliable and durable software that meets users' needs while reducing environmental impacts [1]. It consists of practices that enable linking Software Engineering to the principles of sustainability [7]. In this sense, software development and maintenance, when carried out in a sustainable way, might lead to produce greener software products.

Although current discussions about the topic make it possible to state that sustainability should be considered in the software development process, existing models and quality standards, such as the ISO 9126 and ISO 25010, do not consider sustainability as a quality attribute. There are important discussions reported in the SE literature about the relationship between software quality and sustainability [4, 6, 17, 19]. These authors agree with the need to think sustainability as a new quality attribute and investigate how this new attribute could be incorporated into the software development process in terms of energy and resource efficiency.

Aiming at providing future directions for investigations on Green SE, it is important to establish a common understanding about the fundamental concepts about sustainability in the SDLC. In this sense, we carried out a Systematic Mapping Study [18, 21] in order synthesize the existing Green SE approaches, to understand what have been achieved so far in terms of methods, processes, tools and metrics. This study also aims to leverage the gaps and identify opportunities for further investigations.

This research started by designing a research protocol based on the guidelines provided in [3]. We defined one main research question, which was split into eight sub-questions. The data collection included the application of inclusion and exclusion criteria. We carefully selected 4,912 candidate primary studies that were extracted from 6 digital databases and published until 2017. This process resulted in the identification of 75 relevant primary studies.

The remainder of this paper is organized as follows: Section 2 discusses related work. Section 3 describes the applied systematic mapping study methodology. Section 4 describes data collection. Section 5 reports and discusses the results of this systematic mapping study. In section 6, the threats to validity are discussed. Finally, Section 7 draws concluding remarks and points out opportunities for further research.

2 RELATED WORK

We have identified some studies developed in order to gather and assess available evidence in the Green SE field.

Penzenstadler et al. [16] reported a Systematic Literature Review (SLR) on SE practices for sustainability which provided an overview of 83 studies published prior to 2013. The study analyzed the existing literature from six perspectives, as follows: research topics, models and methods used in practice, research type, application domains, most active research groups and distribution between academics and practitioners. However, the study specifically investigates approaches about models and methods for sustainable software. Furthermore, the authors selected studies published until 2013.

Berntsen et al. [5] performed a SLR with the purpose of outlining the state-of-the-art of Sustainable SE with a focus on existing models, guidelines and practices, as well as proposals in this regard. However, this study selected 36 works until 2010 and only considered three research questions. It investigated the most cited/reported guidelines and models in Sustainable SE, the evolution of interest and the most important authors and venues reporting papers on this topic.

García-Mireles [11] performed a SLR to investigate software process improvement approaches focused on sustainability. The study analyzed 7 primary studies published prior to 2016. Four research questions were investigated: (1) publication trends in regards to sustainability enhancement based on software process improvement; (2) research approaches; (3) software processes; and (4) the main features of the identified approaches to improve sustainability in software processes. However, the study was specifically focused on sustainability from a software process improvement perspective.

Anwar and Pfahl [2] reported a SLR in the Green SE field focused on the role of software analysis. The study classified 50 studies published between 2015 and 2016 and discussed them from the following perspectives they addressed: sub-domains of Green SE, contribution types, research types, role of software analysis and the potential for future research in the area. However, the study focus specifically on software analysis and the time frame only considered two years. It is likely that several important publications were left aside from such an analysis.

Marimuthu and Chandrasekaran [13] summarizes in an SLR the body of knowledge of methods for green and sustainable software development and provides a platform for conducting future research. They analyzed 82 primary studies published between 2010 and May 2016 and defined 7 research questions: type of research, research goals, research topics, research contributing, type of validation methods, tools proposed and publication venues. However, the study did not perform any automated search through digital libraries engines, but they only applied the snowballing method within popular publications to find out the relevant studies.

Wolfram et al. [22], in turn, presents in an SLR the state-of-the-art of sustainability in the context of SE. The authors analyzed 168 publications published between 1980 and 2013. The study classified the primary studies from three perspectives: sub-disciplines of SE, time scope considered in the definition and the definition developed over time.

García-Mireles et al. [12] carried out a SLR aimed to provide an overview of the approaches found in the literature for dealing with interactions between software product quality and sustainability in the context of software application . In order to accomplish that, the authors classified 66 primary studies and answered two research questions about profile of papers and specific interactions between environmental sustainability goals and product quality characteristics.

In order to develop this work, we considered every mentioned study, since they bring relevant information. However, we observed that these studies were focused on particular issues such as software analysis, practices, models, methods and software process improvement, without providing a broader coverage of the field. As opposed to such studies, in our investigation we considered eight distinct perspectives, established in the set of research questions. As the research in this field is incipient, it becomes important to observe the state-of-the-art considering recent publications. The selection of primary studies counted on an automated search in the widely accepted digital libraries for Computer Science research, and also through *snowballing*. We also carried out a critical analysis and evaluation, by employing a rigorous quality assessment [10].

3 RESEARCH METHOD

The method used in this research is a Systematic Mapping Study (SMS). A SMS is a comprehensive review of primary studies in a specific topic area that seeks to identify what evidence is available on the topic [3]. This study was developed based on the systematic process proposed by Petersen et al. [18]. Each phase is further explained along this section.

In this study, the focus was to investigate the Green SE field, in particular to unveil software methods, processes, tools, and metrics. Hardware-related issues were left aside from this investigation, although some Green and Sustainable researches have been focused on such a direction.

Table 1 details the PICOC structure defined for this study. From such a structure, we could define the main research question addressed in this investigation: **“What are the existing software approaches supporting the Green SE concepts?"** To answer the research question, we derived eight following sub-questions. Table 2 lists all these questions, with associated rationales explaining their choice.

Table 1: PICOC structure.

Population	Papers published in the area of SE.
Intervention	Papers that present proposals in the area of Green SE.
Comparison	Not applied.
Outcomes	1. Emerging Domains of Green SE; 2. Lack of evidence and research opportunities in Green SE.
Context	Software (methods, models, process, SDLC, tools, methodologies and metrics).

Table 2: List of research questions.

	Research Questions	Rationale
RQ1	What contribution types of SE have been investigated?	This question is intended to investigate / determine which types of contribution among traditional SE processes, methods, tools and metrics have been investigated in light of the Green SE domain. We used the contribution types defined by Petersen et al. [18] which are: Process, method, tool, metric and model.
RQ2	If so, what SDLC phases have been applied?	This question is intended to identify where the sustainability efforts were dedicated throughout the SDLC. The steps adopted in this study are those provided in the SWEBOK [1]: requirements, design, construction, testing and maintenance.
RQ3	What evidence types have been identified?	This question seeks to identify which types of evidence are most applied in empirical studies. The levels of hierarchy of evidence were suggested by Barbara and Stuart [3], which are: No evidence, evidence obtained from demonstration or elaboration of toy examples, evidence obtained from expert opinions or observations, evidence obtained in academic studies, evidences obtained in industrial studies and evidences obtained from the industrial practice.
RQ4	What research types were conducted?	This question aims at identifying the research approaches that were presented in the primary studies, based on the classification scheme proposed by Wieringa et al. [20], which consists of evaluation, proposal of solution, validation, philosophical, opinion paper and experience paper.
RQ5	What research methods are available?	The objective of this research question is to analyze, within the research methods applied by the primary studies, which are the most used approaches to evidence the efforts of Green SE based in the study of Castellan [8], which can be defined as: case study, theory grounded, survey and meta-analysis.
RQ6	Which application domains were considered?	The objective is to investigate which application domains have been addressed by the research community, e.g., Mobile, Cloud, IoT, Embedded Systems, etc.
RQ7	Which publication venue are most commonly used?	The goal is to identify which publication venue have accepted papers on Green SE. It is believed that this data will help researchers identify the main locals of publication and find relevant studies in the field.
RQ8	What is the distribution between academia and industry?	The goal is to investigate where research efforts are being directed at Green SE, whether for academia, industry or both.

4 DATA COLLECTION

We next present the activities carried out in this SMS, which encompass the search strategy, data sources, inclusion and exclusion criteria and the selection process.

4.1 Search strategy

Initially the search for the keywords and synonyms related to the research question was carried out. The selection of the set was defined based on the guidelines proposed by Barbara and Stuart [3]. In addition, the keywords used in the Call for Papers in the International Workshop on Green and Sustainable Software (Greens) were also consulted. After performing such tasks, the final set of keywords was: **Software, Green, Sustainability, Sustainable, Energy, Eco** and **Power**.

In order to get the most relevant results from the digital libraries, we combined these keywords to form a relevant search string. The criteria used to obtain such a search string were: greater number of results retrieved from the digital libraries and articles strongly related to the topic. We carried out eleven (11) combinations until we got the *string* considered adequate for a broader search. It was then defined as follows:

((“green software engineering") OR (“software engineering" AND “eco-sustainability") OR (“sustainable software engineering"))

4.2 Data sources

We considered publications retrieved from the most widely-accepted digital libraries in the Computer Science field, as follows: IEEEXPLORE, ACM DL, SPRINGERLINK, SCIENCE DIRECT, WILEY ONLINE LIBRARY and GOOGLE SCHOLAR.

We restricted the search to studies published up to December 2017. An inferior year-limit was not defined, so as to enable the identification of the most important studies in the Green SE field. Reference papers were retrieved in the SE area whose access to authors of this SMS was possible in the research environment. The search was also performed using the *snowballing* process.

4.3 Studies selection

After the execution of the search strategy, it was necessary to establish inclusion and exclusion criteria to select the articles in the resulting set. These studies were subsequently classified and they followed to the quality assessment phase.

The inclusion and exclusion criteria take into account the quality of the article in terms of language, type, year of publication and relation with the topic of this research. Table 3 presents the criteria defined for this SMS. It is intended to focus only on studies that address processes, methods, metrics and tools in the context of software related to Green SE.

The study selection process encompassed three phases, as Figure 1 shows. These are detailed next.

- **Phase 1.** It consisted of the application of the automatic search in the digital libraries' search engines. The particularities in the search syntax of the repositories made it impossible to execute a standard string. It was therefore necessary to prepare a specific syntax for each search engine. In the Springer Link, papers were filtered from the Computer Science area, Software Engineering sub-discipline, written in English and published until 2017. In the Science Direct, we filtered by Year of publication, “Research articles" and “Review articles". In the other databases, the default string

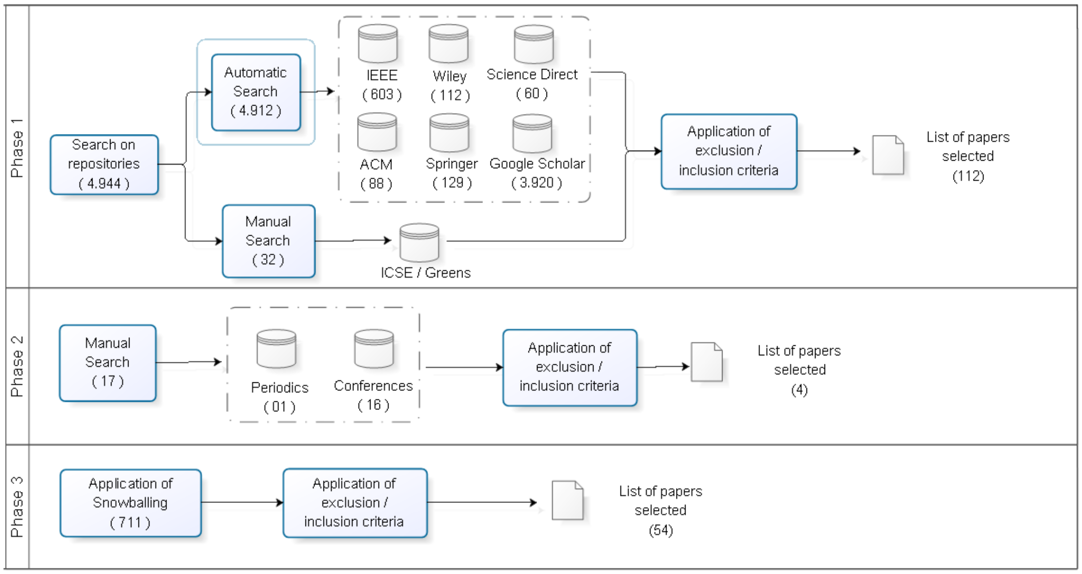

Figure 1: Selection process of primary studies.

Table 3: Exclusion and inclusion criteria.

Criteria	Class
Studies related to Green SE	I
Latest version of paper published	I
Papers published to 2017	I
Papers in English	I
Studies focused on Green, without bias in SE	E
Gray literature (theses, dissertations, reports, reports of experiences, short paper, books, magazines, expanded abstracts, opinion articles, workshops, courses, tutorials).	E
Secondary and tertiary studies	E
Duplicated papers	E

Legend: (I)nclusion and (E)xclusion criteria.

was used in the Advanced Search, and it worked seamlessly. However, we did not specify any particular filters.
The automatic search retrieved a set of 4,912 candidate primary studies. The manual search of papers was done only in the GREENS workshop, where 32 studies were retrieved. Of the 4,944 studies retrieved, the inclusion and exclusion criteria were applied. As Figure 1 shows, at the end of this phase, 112 articles remained.

- **Phase 2.** The manual search was expanded and applied on the main conferences and journals of the area in SE: ICSE, SBES and SBQS. Another 17 studies were found. After the application of the inclusion and exclusion criteria, only 4 remained.
- **Phase 3.** The *Snowballing* procedure was applied to studies retrieved from phases 1 and 2, with an additional 771 primary studies to analyze. These were scrutinized by applying the inclusion and exclusion criteria. At the end, we obtained 54 papers.

At the end of the paper selection process, the list of primary studies selected for full reading counted 170 publications.

4.3.1 Reliability of inclusion decisions. Each paper was analyzed by two researchers individually. These read the title, abstract, keywords and introduction (when necessary) and applied the inclusion and exclusion criteria. Before this procedure, it was verified the need to define a method for prevention / treatment of possible divergences.

For this purpose, a five-point Likert Scale was used. Papers marked with the “Strongly Agree" scale were accepted and those marked with “Strongly Disagree" were excluded. Articles marked with "Partially Agree", “Partially Disagree" or “Neutral" were discussed between the two researchers to decide whether they would be included or excluded from the final set. Without consensus, a third researcher acted with the analysis and the final decision making.

To ensure the reliability of agreement among researchers during the selection of primary studies, we applied Kappa statistics [14]. This procedure measures the agreement between the researchers in the evaluation of the studies. The achievement of the agreement ratio measure occurred with the application of the scores, varying from -2 to 2, in each of the Likert Scale levels. Next, the scores of the first researcher were subtracted from the scores of the second researcher and the number of zeros obtained was counted. Finally, the number of zeros was divided by the total number of studies.

The percentage agreement obtained was 70%, which corresponds to a "Strong agreement" level.

4.4 Quality evaluation

Quality criteria was employed in an attempt to ensure that the final pool of studies contained the most relevant ones. This research used the approach made by Dyba et al. [10], which address the following quality issues: reporting, credibility, rigor and relevance.

The studies were read in full by two researchers, who worked individually, applying a score to the quality issues described in Table 4. For each defined question, a score [0.0, 0.5 or 1.0] was given, according to the following criteria:

- The publication does not meet the quality criterion [0.0];
- The publication partially meets the quality criterion [0.5];
- The publication fully satisfies the quality criterion [1.0].

It was also defined that if the study met any exclusion criterion, it would be removed. Therefore, throughout the reading process, 95 articles were excluded and 75 were evaluated and targeted to the classification phase.

Table 4: Quality Criteria.

Question	Criteria
1. Is the research related to any Green Software Engineering domain?	REPORT
2. Are the objectives clearly stated?	REPORT
3. Is the research context clearly expressed?	REPORT
4. Is the phase of the developmental life cycle covered by the study clearly stated?	REPORT
5. Is the application domain clearly expressed?	REPORT
6. Are the results, limitations and future work clearly described?	CREDIBILITY
7. Is the contribution clearly expressed?	CREDIBILITY
8. Is it possible to identify the publication venue of the research?	CREDIBILITY
9. Is the type of research conducted clearly expressed?	RIGOR
10. Is the research method clearly expressed? In the case of non-empirical research, are the arguments adequately presented?	RIGOR
11. Does the research make clear who contributes?	RELEVANCE

The score of each study was obtained from the sum of the points of all the evaluated criteria. The accepted studies were those with at least 63.64 % of the maximum predicted mark (11.0). 100 % of the studies evaluated reached the minimum score, an average score of 81.82 %. The studies are listed in Appendix A.

4.5 Data extraction

The data extraction forms must be designed to collect all the information needed to address the research questions and the quality criteria [15]. The following information was extracted from each study: title, authors, publication year, publication venue, SE contribution type, SDLC phases, evidence type, research type, application domain, distribution, research method and quality criteria score.

5 RESULTS

In this section, we describe the classification scheme, results and analyses of data extraction.

5.1 Classification scheme

The classification scheme was structured in eight different categories. We next describe each category:

- **Classification of contribution types.**
  The contribution types proposed by the primary studies were defined as: *process, method, tool, metric* and *model*. Throughout the readings, other contribution types have been identified and added to the selection list, they are: *framework, catalog, guide, technique* and *approach*.
- **Classification of the SDLC phases.**
  This work followed the classification suggested by the SWEBOK. Each selected study was categorized into one of the following phases: *requirements, design, construction, test* and *maintenance*.
- **Classification of evidence types.** To classify the studies according to the evidence types, the hierarchy suggested by Barbara and Stuart [3] was used. It encompasses the following levels: *No evidence, Evidence obtained from demonstration or elaboration of examples of toys, Evidence obtained from expert opinions or observations, Evidence obtained in academic studies, Evidence obtained in industrial studies* and *Evidence obtained from industrial practice*.
- **Classification of research types.** The primary studies were classified according to the approach proposed by Wieringa et al. [20]: *Evaluation, solution* and *validation*. In addition, we identified the research types: *Philosophical, Opinion papers* and *Experience papers*.
- **Classification of application domains.** The classification of the studies with respect to the field of application was made as they were identified in the readings and made available in the selection list for the categorization of the following studies. Examples of application domains are: Mobile Applications, Cloud Systems, IoT, Embedded Systems, etc.
- **Classification of publication venue.** The study followed the simpler classification for the publication venue. The studies selected were categorized among journals, conferences, symposia and international workshops, among other relevant events in the SE area. At each new identification, the event was recorded in the set of publication venue to be selected in the following classifications.
- **Classification of contribution types.** The scheme used to classify the primary studies according to their contribution was: *Academy* (the contribution of the study is intended for researchers), *Industry* (the contribution of the study is intended for industry professionals) and *Both* (the contribution of the study is intended for both researchers and industry professionals).

- **Classification of research methods.**
  Based on [8], the studies were classified into: *case study, grounded theory, survey* and *meta-analysis*. Besides such a categorization, after reading the primary studies it was possible to identify another two categories, namely *controlled experiment* and *multi-method approach*.

The non-empirical theoretical research presents itself as an approach that is independent on real-world application, it is characterized as methodology or theoretical basis of empirical research.

The classification procedure of the studies was performed after carefully reading each of the 75 primary studies. The classification was done by extracting the relevant data by using a spreadsheet.

5.2 Results and analysis

This section reports the findings of the SMS. The quantitative results and the analysis of the results are presented by each RQ.

RQ1: What contribution types of SE have been investigated?

The SMS identified 9 different contribution types, as Figure 2 shows. The contributions are distributed as follows: approaches (19%), frameworks (19%), models (16%), techniques (8%), methods (8%), tools (7%), guidelines (5%), metrics (4%) and catalog (1%). It was not possible to identify the contribution in 13% of the selected studies.

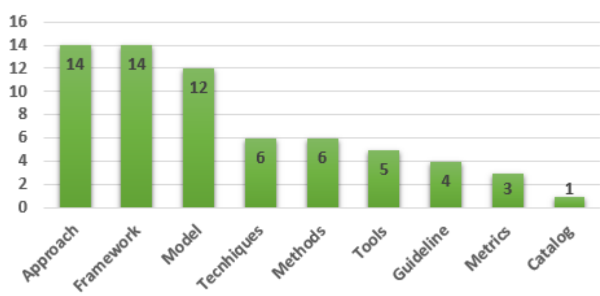

Figure 2: Number of papers by contribution type.

We can observe from the presented data that the researchers have concentrated considerably more on the exploration of approaches, frameworks and models. This may indicate that researchers expect these three contribution types would promote gains in energy efficiency and, consequently, to obtain more sustainable software products. On the other hand, the other 6 contribution types need to receive more attention by the researcher community. Such research could improve the understanding of how sustainability could affect the software development process without impact the quality of delivered products. We also noticed a lack of conceptual standards to characterize the contribution types in SE. This can be evidenced in [2, 11, 12], which use the terms "SE topics", "approaches" and "sub-domains" as synonyms for such a definition. This leads to speculation about the researchers' understanding of how to categorize their proposals.

Figure 3 shows that 13% studies propose sustainable process models. Only 2 of them use traditional SE quality standards as basis for developing it model. They are: ISO / IEC 12207 - Software Lifecycle Processes, ISO / IEC 15504 - Software Process Improvement and Capability Determination and ISO / IEC 33000 - Compliant Measurement Framework for Software Process Sustainability. These data indicate that, in relation to proposed product and process models, research is still incipient and needs to be encouraged.

We observed that 56% of the primary studies proposed means to achieve the energy efficiency of the software in terms of consumption and energy savings. Another important finding is that 40% of studies discuss sustainability from two perspectives: (1) the need for a concrete definition to embed it in the software development process; and (2) sustainability analysis as an attribute of software quality. Finally, 19% seek to understand energy awareness, that is, how sustainability has been currently addressed in software development process in practice through the perceptions and attitudes of SE professionals.

Regarding sustainable practices throughout the SDLC, 68% studies analyze the impacts and implications of applying sustainable practices to traditional software activities. By means of these data, we noticed that, in general, the contributions are spread out in the various SDLC phases.

RQ2: If so, what SDLC phases have been applied?

The intention was to identify which were the most commonly addressed SDLC phases by the research community. The result was: requirements (24%), design (21%), coding (11%), testing (8%) and maintenance (4%). 19% of the studies performed all phases of the life-cycle and 13% did not have any evidence supporting the findings.

Figure 4 indicates that the most investigated phase is requirements, with a total of 18 studies. Next, software design, with 16 studies. These results may indicate that researchers expect sustainability-related requirements such as energy efficiency and resource efficiency to be identified and projected early in the development process prior to implementation. Results in [12, 22] indicate that our finding is similar in this respect. However, we can observe a gap in the next phases of the SDLC, especially in the maintenance one. Future research should focus on achieving sustainability in the construction, testing and maintenance phases of the software.

Other relevant result is observed by crossing the contribution types and the SDLC phases, as Figure 5 shows. The primary studies have primarily focused on understanding how the sustainability interact in design phase (50%), followed by requirements (25%), all life-cycle phases (17%) and construction (8%). Frameworks proposals, in turn, have been more targeted to the requirements phase (31%) and all life-cycle (31%), followed by design (23%), testing (15%) and maintenance (15%).

The SMS shows that the two more relevant contribution types (approaches and frameworks) seek to address issues related mainly to the requirements and design phase. This confirm the research interest in this phases and the lack of investigations in the others.

RQ3: What evidence types have been identified?

About the evidence produced by the primary studies, 88% of the studies obtained evidence from academic studies (49%) and

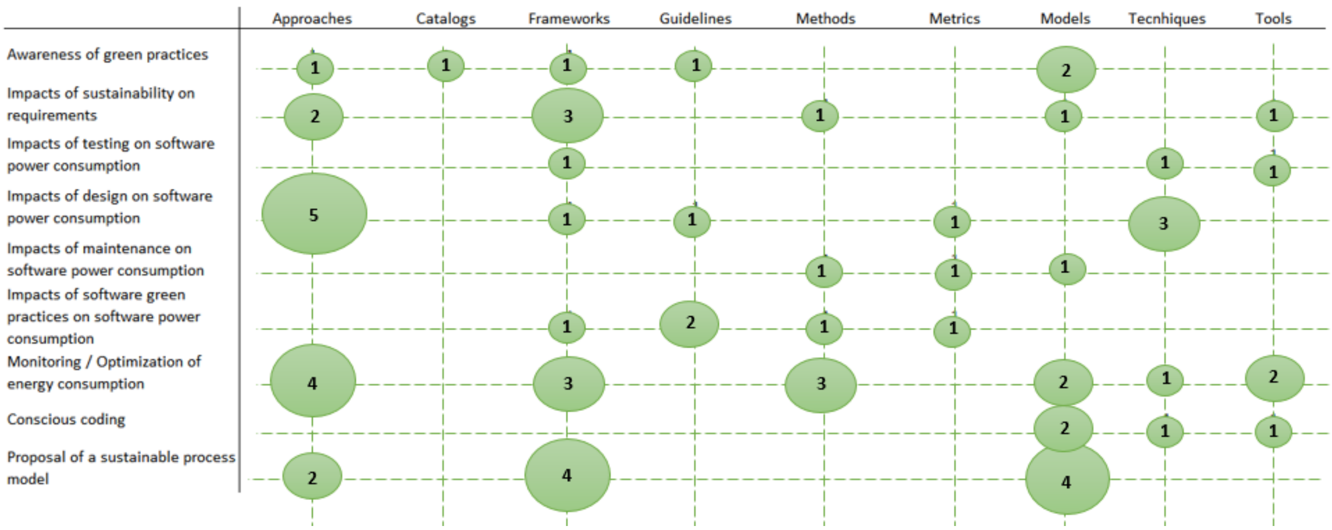

Figure 3: Number of papers by Subject X Contribution type.

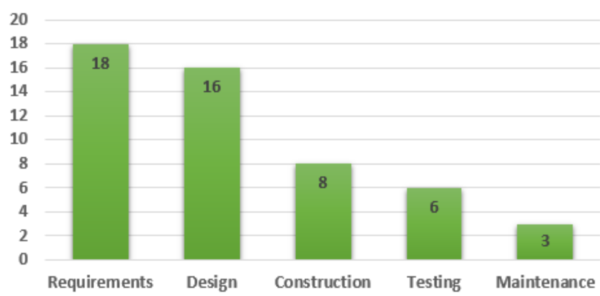

Figure 4: Quantity of papers by SDLC phases.

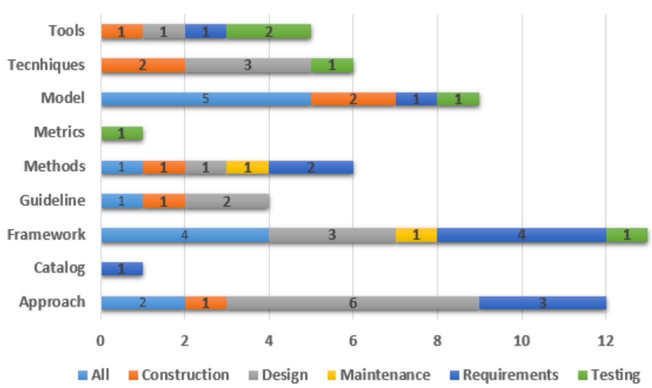

Figure 5: Contribution types x SDLC.

industry studies (39%). The others sought evidence from the opinion of experts (8%) and industry practices (4%).

Our study presents a perspective that observes the academic (simulated) and industrial (real) scenarios used by the research. It is possible to verify that most primary studies use simulated scenarios to carry out their investigations. These numbers demonstrate that the area is incipient and relies mainly on identified evidence in the theory, which needs to be evaluated in practice.

The results presented by [2, 11, 12, 17] show that the Green domain needs contributions from industry professionals. The participation of this professionals should be encouraged since they can contribute with their experiences and observations of the real world, allowing a better prioritization of the research objectives. Although the industry recognizes this need, there is little evidence.

RQ4: What research types were conducted?

Regarding the most important research types, most primary studies focused on proposing novel solutions (68%), followed by evaluation studies (21%) and philosophical research (3%). It is important to notice that three out of six research types proposed by [20] and used in the classification scheme of this study were not mentioned. They are: opinion, validation and experience.

Such numbers point out to the need for evaluation and experience studies, since most of the primary studies bring solution proposals. As previously mentioned, this observations indicate potentially that current solutions are still immature and that the research community is still young.

Similar findings were reported in [2, 13, 17]. This study unveils the need for assessment validation and opinion research by industry experts and practitioners to help better understand the current industry needs for Green SE. García-Mireles [11], García-Mireles et al. [12] show that the lack of empirical data becomes a constraint for the application of solutions proposals in industry.

RQ5: What research methods are available?

This question aims to identify what the most commonly used research methods by Green SE research community. According to [13], the maturity of the publication depends on the validation methods used to validate the research ideas or results.

59% of the publications carried out case studies (35%) or controlled experiments (24%). 31% of the studies were classified as follows: Non-empirical research (11%), Multi-method (7%), Survey (7%) and Meta-analysis (4%). Another 13% did not specify any research method. This result shows that the controlled experiments and case studies are the most applied research methods by the primary studies. A similar result was reported by Anwar and Pfahl [2].

When correlating the research types with the research methods, it is possible to observe that there is a predominance of solution proposals in the majority of the primary studies analyzed. 59% of papers applied both case studies and controlled experiments. This ratifies that the Green SE field have increased in attention by the research community.

From a perspective of which research types are most adopted by research methods, it can be observed from Figure 6 that both controlled experiment and case study research have been employed as empirical evaluation methods, at a same proportion, in solution proposals (75%), followed by evaluation studies (25%).

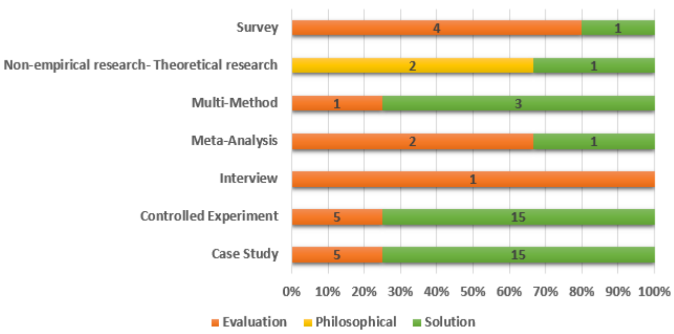

Figure 6: Research types x Research methods

RQ6: Which application domains were considered in the evaluation?

Only 24% of the primary studies addressed any application domain. Most of them (61%) presented results of applications in the mobile systems domain, followed by applications in the cloud systems domain (22%). Another 18% focused their investigations on distributed systems, legacy systems and IoT, at a same proportion.

These numbers potentially indicate that researchers have given considerable attention to mobile devices. We believe that the main reason for preference over this domain is due to the recent growth of the mobile applications market and the research for the reduction of their energy consumption.

RQ7: Which publication venues are most commonly used?

By analyzing the most popular publication venues in researches focused on Green SE, we could observe a mainstream concentration of primary studies from the ICSE main track and the Greens Workshop – held at ICSE (13.5%, each). The remainder (73%) is distributed in several other conferences and journals. In all, 41 different publication venues have been identified.

Similar results are reported in [12, 13]. In addition, other authors [13, 16] observed a small distribution of publications in journals when compared to conferences and workshops. Penzenstadler et al. [16] further states this is likely due to the fact that the research community is still forming.

In relation to the growth in the number of publications in recent years, the primary studies were published between 2003 and 2017. Most of them were published in the last five years. As Figure 7 shows, there was a small number of publications up to the year 2012, followed by a significant increase in 2013. Specifically, in this last year, the number of publications doubled and remained constant over the next two years. A considerable drop can be observed in 2016, followed by a further increase the following year.

Figure 7: Number of papers per year.

We believe this is one of the key evidences of the increasing interest of the research community in Green SE. These numbers show a clear indication of research opportunities in this field. The results also suggest the need for more publications in journals focusing on sustainable software.

RQ8: What is the distribution between academia and industry?

Most researches have industry input (56%). The academy distribution is comparatively smaller (5%). The studies also revealed that the contribution of both industry and academia is 20%. It was not possible to identify the distribution in 19% of the primary studies.

Our findings confirm that the community remains interested and focuses on solving industry-related issues. Despite this finding, when we analyze this RQ with the evidence types and the research types, together, we observe that there is still a need for the industry professionals participation to support a better targeting of researches.

6 THREATS TO VALIDITY

In this study, the following potential threats to validity were identified:

Keywords bias: The set of keywords selected in the search may not be the most representative of the domain. Therefore, it may not have returned the best set of papers aimed at Green SE. The mitigation of this bias, however, is difficult: Green SE is a relatively new discipline and there are reports that the confusion of the lack of conceptual terms and understandings hampers the accuracy of the searches for researches that permeate the domain. We attempted to mitigate this threat by researching other mapping studies in Green SE to identify which keywords were used.

Quality evaluation bias: The evaluation of the quality and the classification of the papers were influenced by the knowledge and understanding of the authors. To mitigate this threat, the inclusion, exclusion and quality criteria of the selected studies were defined. In addition, the divergences were discussed, resolved and agreed with the intervention of the third author.

7 CONCLUSIONS AND FUTURE WORK

The main motivation for this work was to investigate the state-of-the-art on Sustainability in SE in order to synthesize available evidence and identify gaps and research opportunities. We aimed to understand what been the efforts made by the SE community to incorporate sustainable practices.

In summary, we noticed a growing interest by the SE Research Community on the Green context, in particular a strong focus on the field of mobile systems. Researchers are mainly interested in proposing frameworks, approaches and models. This may indicate that the research community expects these three types of contribution to be potential for producing sustainable software.

It was also possible to understand how sustainability has been addressed in the SLDC. The SMS shows that approaches and frameworks are mainly focused on the requirements and design phases. In the former, researchers support sustainability as a non-functional requirement in quality standards, which ratifies the importance of stakeholder identification, and a need to further investigate the impact of sustainability on requirements engineering. The latter mainly addresses the impacts of architecture/design on the software's power consumption.

Several approaches discussed efficiency in terms of energy consumption. This evidence may indicate the need to consider energy efficiency as an important quality attribute when designing systems architectures.

Finally, we noticed a need to carry out more validation studies and opinion surveys, mainly on the impacts of sustainability on quality models and processes in traditional software development.

There are several directions to follow towards establishing and improving Green SE practice, in particular in order to provide the SE practice with green and sustainability-aware processes, tools and methods.

ACKNOWLEDGMENTS

This research was partially funded by INES 2.0, CNPq grant 465614/2014-0 and FAPESB grants JCB0060/2016.

A APPENDIX A - QUALITY STUDIES SCORES

Table 5: Selected primary studies

P#	Primary Study	Year	QS
P1	A Catalogue of Green Architectural Tactics for the Cloud	2014	10
P2	A framework for estimating the energy consumption induced by a distributed system's architectural style	2009	10.5
P3	A generic model for sustainability with process and product-specific instances	2013	11

Continued on next column

Table 5 – Continued from previous column

P#	Primary Study	Year	QS
P4	A Green Software Development Life Cycle for Cloud Computing	2013	9
P5	A model and selected instances of green and sustainable software	2010	8
P6	A practical model for evaluating the energy efficiency of software applications	2014	8
P7	A preliminary study of the impact of software engineering on Green IT	2012	9.5
P8	A programming model for sustainable software	2015	9
P9	A software engineer's energy-optimization decision support framework	2014	11
P10	An Automated Analysis of the Branch Coverage and Energy Consumption Using Concolic Testing	2017	9
P11	An Empirical Study of Practitioners' Perspectives on Green Software Engineering	2016	10
P12	An Energy Consumption Perspective on Software Architecture	2015	9
P13	An interview study on sustainability concerns in software development projects	2017	9
P14	An investigation into energy-saving programming practices for Android smartphone app development	2014	10
P15	An ISO/IEC 33000-compliant measurement framework for software process sustainability assessment	2014	8
P16	Analyzing the harmful effect of god class refactoring on power consumption	2014	10
P17	APE: an annotation language and middleware for energy-efficient mobile application development	2014	11
P18	Awakening awareness on energy consumption in software engineering	2017	9
P19	Characterizing sustainability requirements. A new species, red herring, or just an odd fish?	2017	9
P20	Creating environmental awareness in service oriented software engineering	2011	7
P21	Derivation of Green Metrics for Software	2013	7
P22	Detecting energy bugs and hotspots in mobile apps	2014	11
P23	Developing a sustainability non-functional requirements framework	2014	8.5
P24	Empirical evaluation of two best practices for energy-efficient software development	2016	9
P25	Energy efficiency embedded service lifecycle: Towards an energy efficient cloud computing architecture	2014	10
P26	Energy-aware software: Challenges, opportunities and strategies	2013	9
P27	Energy-directed test suite optimization	2013	11
P28	Enhancing Software Engineering Processes towards Sustainable Software Product Design	2010	7
P29	Enhancing Sustainability of the Software Life Cycle via a Generic Knowledge Base	2010	8
P30	Estimating mobile application energy consumption using program analysis	2013	9.5
P31	Evaluating energy efficiency of Internet of Things software architecture based on reusable software components	2017	11
P32	Evidencing sustainability design through examples	2015	10
P33	Green computing and Software Defects in open source software: An Empirical study	2014	9
P34	Green software development model: An approach towards sustainable software development	2011	7.5
P35	Green Software Engineering with Agile Methods	2013	8
P36	Green software requirements and measurement: random decision forests-based software energy consumption profiling	2015	10.5
P37	Green software services: From requirements to business models	2013	8
P38	Green tracker: a tool for estimating the energy consumption of software	2010	10
P39	Greenadvisor: A tool for analyzing the impact of software evolution on energy consumption	2015	10

Continued on next column

Table 5 – Continued from previous column

P#	Primary Study	Year	QS
P40	GreenC5: An adaptive, energy-aware collection for green software development	2017	10
P41	Green-J Model: a novel approach to measure energy consumption of modified condition/decision coverage using concolic testing	2017	9
P42	HADAS and web services: Eco-efficiency assistant and repository use case evaluation	2017	10
P43	How do code refactorings affect energy usage?	2014	10
P44	Impacts of software and its engineering on the carbon footprint of ICT	2015	9
P45	Incorporating Sustainability Design in Requirements Engineering Process: A Preliminary Study	2016	10
P46	Infusing green: Requirements engineering for green in and through software systems	2014	9.5
P47	Initial explorations on design pattern energy usage	2012	10
P48	Integrating Aspects of Carbon Footprints and Continuous Energy Efficiency Measurements into Green and Sustainable Software Engineering	2013	9
P49	Integrating environmental sustainability in software product quality	2015	9.5
P50	"Is software green"? Application development environments and energy efficiency in open source applications"	2012	9
P51	Measuring application software energy efficiency	2012	10
P52	Measuring Software Sustainability	2003	8
P53	Measuring the Sustainability Performance of Software Projects	2010	8
P54	Requirements Prioritization Framework for Developing Green and Sustainable Software using ANP -based Decision Making	2013	8.5
P55	Requirements: the key to sustainability	2015	10
P56	Safety, security, now sustainability: the nonfunctional requirement for the 21st century	2014	10
P57	Self-optimization of the energy footprint in service-oriented architectures	2010	9
P58	Software Sustainability from a Process-Centric Perspective	2012	9
P59	Supporting Physicians by RE4S: Evaluating Requirements Engineering for Sustainability in the Medical Domain	2015	8.5
P60	Sustainability design and software: The karlskrona manifesto	2015	8
P61	Sustainability design in requirements engineering: state of practice	2016	10
P62	Sustainability guidelines for long-living software systems	2012	9
P63	Sustainable development, sustainable software, and sustainable software engineering: An integrated approach	2011	8
P64	Sustainable software engineering: process and quality models, life cycle, and social aspects	2015	8
P65	The Contexto Framework: Leveraging Energy Awareness in the Development of Context-Aware Applications	2016	11
P66	The greensoft model: A reference model for green and sustainable software and its engineering Sustainable Computing	2011	8
P67	The Impact of Improving Software Functionality on Environmental Sustainability	2013	10
P68	The impact of user choice on energy consumption	2014	10
P69	Towards power reduction through improved software design	2012	10
P70	Towards sustainable software criteria: Rescue operation and disaster management system model	2013	8
P71	Uncovering sustainability requirements: An exploratory case study in canteens	2015	10
P72	Understanding Green Software Development: A Conceptual Framework	2015	9.5
P73	Who is the Advocate?: Stakeholders for Sustainability	2013	10
P74	Sustainability requirements for connected health applications	2017	10
P75	Puzzling out Software Sustainability	2017	9

Legend: P#: Primary Study ID, QS: Quality Score

REFERENCES

[1] Nadine Amsel, Zaid Ibrahim, Amir Malik, and Bill Tomlinson. 2011. Toward sustainable software engineering (nier track). In *Proceedings of the 33rd International Conference on Software Engineering*. ACM, 976–979.

[2] Hina Anwar and Dietmar Pfahl. 2017. Towards greener software engineering using software analytics: A systematic mapping. In *Software Engineering and Advanced Applications (SEAA), 2017 43rd Euromicro Conference on*. IEEE, 157–166.

[3] Kitchenham Barbara and C Stuart. 2007. Guideline for performing systematic literature reviews in software engineering (Version 2.3.). *Staffordshire, UK: University of Keele* (2007).

[4] Christoph Becker. 2014. Sustainability and longevity: Two sides of the same quality? *mental* 20 (2014), 21.

[5] Kristina Rakneberg Berntsen, Morten Rismo Olsen, Narayan Limbu, An Thien Tran, and Ricardo Colomo-Palacios. 2016. Sustainability in Software Engineering-A Systematic Mapping. In *International Conference on Software Process Improvement*. Springer, 23–32.

[6] C Calero and MF Bertoa. 2013. 25010+ S: A software quality model with sustainable characteristics. *Sustainability as an element of software quality. Green in Software Engineering Green by Software Engineering (GIBSE 2013) co-located with AOSD* (2013).

[7] Coral Calero and Mario Piattini. 2017. Puzzling out Software Sustainability. *Sustainable Computing: Informatics and Systems* 16 (2017), 117–124.

[8] Catherine M Castellan. 2010. Quantitative and qualitative research: A view for clarity. *International journal of education* 2, 2 (2010), 1.

[9] John Cook, Naomi Oreskes, Peter T Doran, William RL Anderegg, Bart Verheggen, Ed W Maibach, J Stuart Carlton, Stephan Lewandowsky, Andrew G Skuce, Sarah A Green, et al. 2016. Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters* 11, 4 (2016).

[10] Tore Dyba, Torgeir Dingsoyr, and Geir K Hanssen. 2007. Applying systematic reviews to diverse study types: An experience report. In *Empirical Software Engineering and Measurement, 2007. ESEM 2007. First International Symposium on*. IEEE, 225–234.

[11] Gabriel Alberto García-Mireles. 2016. Environmental Sustainability in Software Process Improvement: a Systematic Mapping Study. In *International Conference on Software Process Improvement*. Springer, 69–78.

[12] Gabriel Alberto García-Mireles, Mª Ángeles Moraga, Félix García, Coral Calero, and Mario Piattini. 2017. Interactions between environmental sustainability goals and software product quality: a mapping study. *Information and Software Technology* (2017).

[13] C Marimuthu and K Chandrasekaran. 2017. Software Engineering Aspects of Green and Sustainable Software: A Systematic Mapping Study. In *Proceedings of the 10th Innovations in Software Engineering Conference*. ACM, 34–44.

[14] Mary L. McHugh. 2012. Interrater Reliability: The Kappa Statistic. *Biochemia Medica* 22 (2012), 276––282. Issue 3.

[15] Paulo Anselmo da Mota Silveira Neto, Ivan do Carmo Machado, John D McGregor, Eduardo Santana De Almeida, and Silvio Romero de Lemos Meira. 2011. A systematic mapping study of software product lines testing. *Information and Software Technology* 53, 5 (2011), 407–423.

[16] Birgit Penzenstadler, Ankita Raturi, Debra Richardson, Coral Calero, Henning Femmer, and Xavier Franch. 2014. Systematic mapping study on software engineering for sustainability (SE4S). In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. ACM, 14.

[17] Birgit Penzenstadler, Ankita Raturi, Debra Richardson, and Bill Tomlinson. 2014. Safety, security, now sustainability: The nonfunctional requirement for the 21st century. *IEEE software* 31, 3 (2014), 40–47.

[18] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. 2008. Systematic Mapping Studies in Software Engineering.. In *EASE*, Vol. 8. 68–77.

[19] Colin Venters, Lydia Lau, Michael Griffiths, Violeta Holmes, Rupert Ward, Caroline Jay, Charlie Dibsdale, and Jie Xu. 2014. The blind men and the elephant: Towards an empirical evaluation framework for software sustainability. *Journal of Open Research Software* 2, 1 (2014).

[20] Roel Wieringa, Neil Maiden, Nancy Mead, and Colette Rolland. 2006. Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requirements Engineering* 11, 1 (2006), 102–107.

[21] Claes Wohlin, Per Runeson, Paulo Anselmo da Mota Silveira Neto, Emelie Engström, Ivan do Carmo Machado, and Eduardo Santana de Almeida. 2013. On the reliability of mapping studies in software engineering. *Journal of Systems and Software* 86, 10 (2013), 2594 – 2610.

[22] Nina Wolfram, Patricia Lago, and Francesco Osborne. 2017. Sustainability in Software Engineering. (2017).